Cell Reports
Report

NSun2-Mediated Cytosine-5 Methylation of Vault Noncoding RNA Determines Its Processing into Regulatory Small RNAs

Shobbir Hussain,[1] Abdulrahim A. Sajini,[1] Sandra Blanco,[1] Sabine Dietmann,[1] Patrick Lombard,[1] Yoichiro Sugimoto,[2] Maike Paramor,[1] Joseph G. Gleeson,[3] Duncan T. Odom,[4] Jernej Ule,[2,5,*] and Michaela Frye[1,*]
[1]Wellcome Trust – Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QR, UK
[2]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK
[3]Howard Hughes Medical Institute, University of California, San Diego School of Medicine, La Jolla, CA 92093, USA
[4]University of Cambridge, CR-UK, Cambridge Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK
[5]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK
*Correspondence: j.ule@ucl.ac.uk (J.U.), mf364@cam.ac.uk (M.F.)
http://dx.doi.org/10.1016/j.celrep.2013.06.029

SUMMARY

Autosomal-recessive loss of the *NSUN2* gene has been identified as a causative link to intellectual disability disorders in humans. NSun2 is an RNA methyltransferase modifying cytosine-5 in transfer RNAs (tRNAs), yet the identification of cytosine methylation in other RNA species has been hampered by the lack of sensitive and reliable molecular techniques. Here, we describe miCLIP as an additional approach for identifying RNA methylation sites in transcriptomes. miCLIP is a customized version of the individual-nucleotide-resolution crosslinking and immunoprecipitation (iCLIP) method. We confirm site-specific methylation in tRNAs and additional messenger and noncoding RNAs (ncRNAs). Among these, vault ncRNAs contained six NSun2-methylated cytosines, three of which were confirmed by RNA bisulfite sequencing. Using patient cells lacking the NSun2 protein, we further show that loss of cytosine-5 methylation in vault RNAs causes aberrant processing into Argonaute-associated small RNA fragments that can function as microRNAs. Thus, impaired processing of vault ncRNA may contribute to the etiology of NSun2-deficiency human disorders.

INTRODUCTION

Cytosine-5 methylation (m^5C) is a common epigenetic modification found in DNA with important regulatory roles in transcription (Suzuki and Bird, 2008). The cellular and molecular functions of m^5C-modified nucleobases in RNA, however, remain largely unknown. Dnmt2 and NSun2 are currently the only known m^5C RNA methyltransferases in higher eukaryotes, and transfer RNA (tRNA) is the confirmed target substrate for both enzymes (Brzezicha et al., 2006; Goll et al., 2006). The regulatory functions of m^5C modifications in tRNA are not fully understood but have been reported to regulate tRNA stability and cleavage (Schaefer et al., 2010; Tuorto et al., 2012). Deletion of Dnmt2 or NSun2 in yeast, flies, and mice impairs cellular differentiation pathways in skin, testes, and brain (Blanco et al., 2011; Hussain et al., 2013; Rai et al., 2007; Tuorto et al., 2012). In humans, mutations in the *NSUN2* gene can cause disorders that are associated with intellectual disability (Abbasi-Moheb et al., 2012; Khan et al., 2012; Martinez et al., 2012).

Although NSun2-dependent deposition of m^5C into tRNAs has been widely confirmed, global identification of m^5C in RNA has been hampered by the lack of suitable molecular techniques. Recent high-throughput RNA methylation profiling by bisulfite sequencing and the chemical modification of cytosine-5 by 5-azacytidine increased the repertoire of RNAs carrying m^5C modifications (Khoddami and Cairns, 2013; Squires et al., 2012). In this report, we combine various transcriptome-wide methodologies to identify NSun2-specific RNA methylation sites independent of any chemical modification of RNA. CLIP (crosslinking immunoprecipitation) is a stringent technique devised to identify RNA-protein interactions and uses UV crosslinking to induce a covalent bond between protein and RNA (Ule et al., 2003). Combined with next-generation sequencing, the iCLIP protocol enables genome-wide analysis of crosslink sites at nucleotide resolution (iCLIP) (König et al., 2010).

We modified the iCLIP protocol to identify additional RNA methylation targets of NSun2 and termed it miCLIP (methylation iCLIP). In addition to the established tRNA target substrates of NSun2, miCLIP identified coding RNAs and noncoding RNAs (ncRNAs). We establish vault ncRNAs as NSun2-specific methylated targets and confirm the deposition of m^5C by RNA bisulfite sequencing. Finally, we provide evidence that m^5C controls the processing of vault ncRNAs into small regulatory RNAs with microRNA functions.

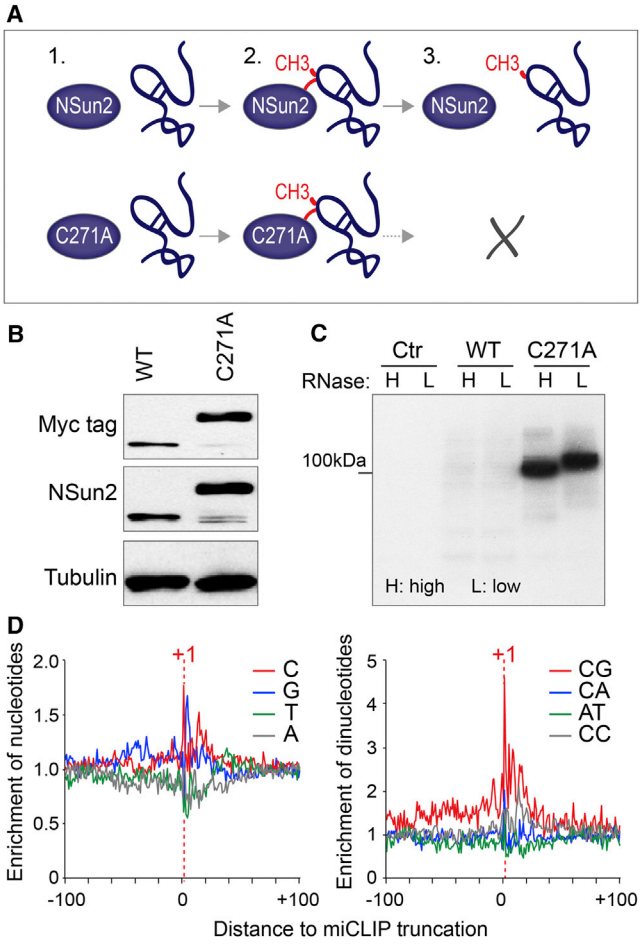


Figure 1. miCLIP Identifies Cytosine-5-Methylated Nucleosides

(A) Schematics of NSun2-mediated cytosine-5 methylation and how the C271A mutation causes irreversible covalent crosslinks between the protein and substrate.

(B) Western blot detecting wild-type (WT) and mutant (C271A) NSun2 proteins using an antibody for the Myc tag (top) or NSun2 (middle). Tubulin (bottom) serves as a loading control.

(C) Detection of radiolabeled immunoprecipitated protein-RNA complexes (^{32}P-ATP) after transfection of an empty vector control (Ctr), wild-type NSun2 (WT), or mutant NSun2 (C271A) using a Myc antibody. Lysates were incubated with high (H) or low (L) concentration of RNase.

(D) Enrichment of nucleotides (left) and dinucleotides (right) in the region up to 100 nt around all crosslink sites. Only the top 4 dinucleotides at position +1 are shown (see also Table S1).

See also Figure S1.

RESULTS

miCLIP: A Technique to Identify m^{5}C in the Transcriptome at Nucleotide Resolution

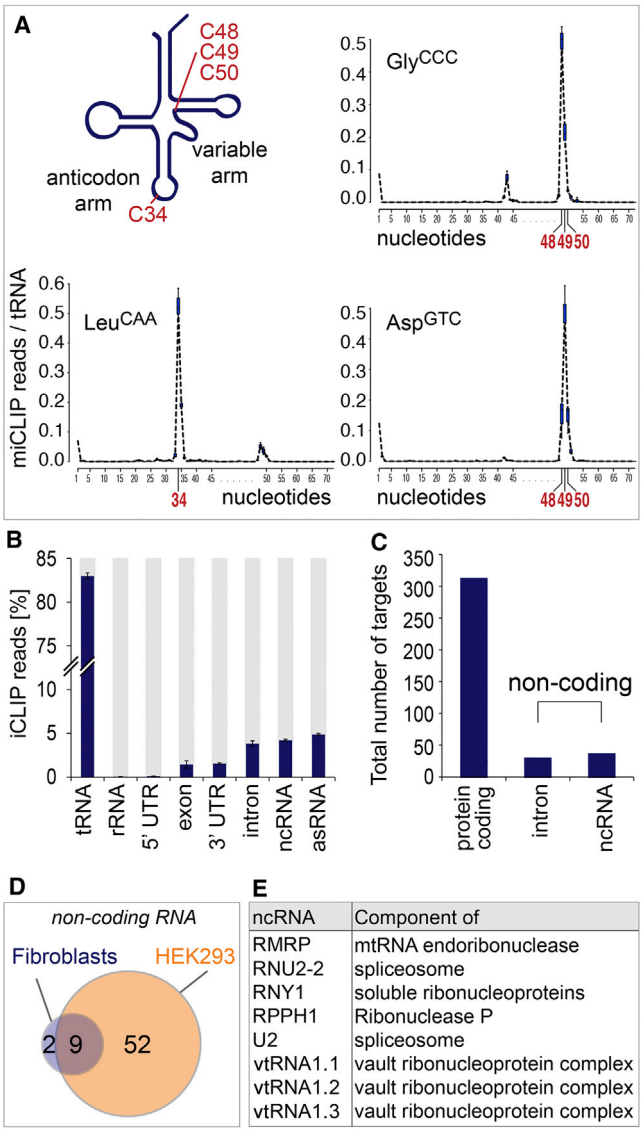
Cytosine methylation at carbon 5 (m^{5}C) is initiated by the formation of a covalent bond between cysteine 321 of NSun2 and the cytosine pyrimidine ring (Figure 1A) (Liu and Santi, 2000). The release of the methylated RNA depends on a second conserved cysteine at position 271 (C271) (Figure 1A) (King and Redman, 2002; Redman, 2006). Mutation of C271 (C271A) stabilizes the covalently linked protein-RNA catalytic intermediate, which can be detected as higher-molecular-weight complexes by western blot (Figures 1A and 1B) (Hussain et al., 2009).

Because the formation of the protein-RNA covalent bond allowed direct immunoprecipitation of the Myc-tagged C271A NSun2 without UV crosslinking, we named our method miCLIP (methylation iCLIP). The protein-RNA complex was detected by radiolabeling, and a shift in molecular weight in response to a high concentration of RNase I confirmed the presence of the NSun2-RNA complex (Figure 1C). We extracted the RNA from the purified complex and amplified the libraries for 25 or 35 PCR cycles, followed by high-throughput sequencing (Figures S1A and S1B) (König et al., 2010; Sugimoto et al., 2012). We used at least three independent replicates per cell line for all analyses. Analyses of the complementary DNA libraries showed strong cytosine enrichment at position +1 (Figure 1D, left panel), which corresponds to the first nucleotide of all sequence reads (Sugimoto et al., 2012). Thus, reverse transcription terminates precisely at the polypeptide-nucleotide (C271A-cytosine-5) crosslink site. We further detected enrichment of CG dinucleotide at position +1, indicating that deposition of m^{5}C occurs preferably at this dinucleotide (Figure 1D, right; Table S1).

miCLIP-Identified NSun2 Targets Are tRNAs, mRNAs, and ncRNAs

The vast majority of miCLIP reads (>80%) mapped to tRNAs (Figures 2A and 2B). RNA bisulfite conversion identified tRNA Asp^{GTC}, Val^{AAC}, Gly^{GCC}, and Leu^{CAA} as methylation substrates of NSun2 in mouse (Tuorto et al., 2012), and miCLIP precisely mapped the expected m^{5}C sites in these tRNAs (Figure 2A; Figure S2A). When all tRNA reads were mapped, miCLIP identified a total of 41 isoacceptors (Figure S2B). These results are in good agreement with the recently developed 5-azacytidine-mediated RNA immunoprecipitation method (Aza-IP), where the majority of tRNAs were found to be methylated by NSun2 (Khoddami and Cairns, 2013). miCLIP consistently detected NSun2-targeted sites within the variable arm at cytosines 48, 49, and 50 (Figure S2B). However, it does not detect any additional NSun2 target sites outside the variable arm (Khoddami and Cairns, 2013).

The specificity of the interaction between the C271A mutant protein and its target RNA was further confirmed by the very low number of reads mapped to ribosomal RNAs (rRNAs) (Figure 2B). The total number of reads mapping to other ncRNAs and messenger RNAs (mRNA) was consistently less than 20% (Figure 2C; Table S2). It has been recently suggested that NSun2-mediated methylation of mRNAs may increase their half-life (Zhang et al., 2012); yet gene expression assays in various tissues including testis and liver failed to uncover any major changes in mRNA abundance when NSun2 was deleted (Hussain et al., 2013; Tuorto et al., 2012). The only mRNA identified by miCLIP that was differentially expressed when NSun2 was inhibited by RNAi in HEK293 cells was NSun2 itself (Figure S3A; Table S3). We also sequenced cDNA from total RNA isolated from human skin fibroblasts carrying a heterozygous or homozygous loss-of-function mutation in the *NSUN2* gene (Figure S3B; Table S4) (Martinez et al., 2012). The vast majority of the 312 miCLIP-identified mRNAs (>90%) remained



E

ncRNA	Component of
RMRP	mtRNA endoribonuclease
RNU2-2	spliceosome
RNY1	soluble ribonucleoproteins
RPPH1	Ribonuclease P
U2	spliceosome
vtRNA1.1	vault ribonucleoprotein complex
vtRNA1.2	vault ribonucleoprotein complex
vtRNA1.3	vault ribonucleoprotein complex

Figure 2. Detection of Coding and Noncoding RNA by miCLIP

(A) Schematics of tRNA carrying m^5C at positions 48, 49, and 50 in the variable arm or at position 34 in the anticodon arm (top left). Frequency of miCLIP reads per tRNA identifies all cytosine-5 methylated sites in tRNA Gly^{CCC}, Leu^{CAA}, and Asp^{GTC} (top right and bottom). See also Figure S2.

(B) Percentage of miCLIP reads in noncoding (tRNA, rRNA, 5′ and 3′ UTR, intron, ncRNA, and asRNA) and protein-coding RNAs. Shown are common miCLIP targets of three replicates after 25 cycles of amplification. Error estimates represent SD of the mean. See also Figure S2.

(C) Total number of protein-coding, intron, and other noncoding RNAs (ncRNAs) in three replicates after 35 cycles of amplification. See also Figure S3.

(D) Venn diagram of common miCLIP-identified ncRNAs in HEK293 (orange) and human fibroblasts (blue). See also Figure S3.

(E) Description of eight ncRNAs with differential abundance in human fibroblast carrying a heterozygous (+/−) or homozygous (−/−) loss-of-function mutation for *NSUN2*.

unchanged in the absence of NSun2 (Figure S3B). In conclusion, we do not find evidence for a major role of NSun2 or NSun2-mediated m^5C in mRNA stability.

NSUN2 Methylates Several ncRNAs

In addition to the tRNAs, miCLIP identified several other ncRNAs. Most of these ncRNAs were identified in both cell types and under both PCR amplification conditions (Figure 2D; Figure S3C). Although the number of ncRNAs found in $NSUN2^{+/-}$ human fibroblasts was low, they largely overlapped with those identified in HEK293 cells (Figure 2D; Table S2). miCLIP identified a maximum of 61 common ncRNAs, out of which only 43 were annotated (Table S5). We sequenced cDNA libraries prepared from small RNA isolated from $NSUN2^{+/-}$ and $NSUN2^{-/-}$ human fibroblasts and found eight of the ncRNAs differentially regulated (Figure 2E; Table S6; Figures S3D and S3F). For instance, miCLIP identified NSun2-target cytosine 174 in RPPH1 and cytosine 92 in 5S rRNA (Figures 3A and 3B; Figure S4A), which agrees with previous RNA bisulfite sequencing and Aza-IP in HeLa cells, respectively (Khoddami and Cairns, 2013; Squires et al., 2012). Additionally, we identified NSun2 target sites in 7SK and the three vault RNAs (vtRNAs) (vtRNA1.1, vtRNA1.2, and vtRNA1.3) (Figures 2E, 3A, and 3C; Figure S4B). Of these, only cytosine 69 in vtRNA1.1 has previously been reported to be methylated by NSun2 using Aza-IP and RNA bisulfite sequencing (Khoddami and Cairns, 2013; Squires et al., 2012).

Vault ncRNAs were first described as components of macromolecular ribonucleoprotein complexes termed vaults, which are conserved organelles found in most species (Kedersha et al., 1990; Kedersha and Rome, 1986). The function of vaults is unknown, yet they have been associated with resistance to chemotherapy in cancer (Mossink et al., 2003). Although vtRNA1.1 is expressed at lower levels than vtRNA1.2 and vtRNA1.3 in HEK293 cells (Stadler et al., 2009), it was the most prominent target identified by miCLIP (Figure 3A). vtRNA1.1 showed a methylated site at C69 and vtRNA1.2 and vtRNA1.3 at C59 (Figure 3C). vtRNA1.2 and vtRNA1.3 shared methylation sites at C27 and C59, and only vtRNA1.3 was additionally methylated at C15 (Figure 3C).

To confirm the NSun2-mediated methylation sites independently, we performed RNA bisulfite sequencing in $NSUN2^{+/-}$ and $NSUN2^{-/-}$ human fibroblasts. In line with the miCLIP cDNA counts, vtRNA1.1 showed the highest levels of m^5C at position 69 (Figures 3A and 3D, right panel). We further confirmed m^5C at C27 and C59 in vtRNA1.3, whereas RNA bisulfite sequencing failed to detect m^5C in vtRNA1.2 (Figure 3D, left and middle panels; Figures S4C and S4D). All vtRNAs isolated from $NSUN2^{-/-}$ fibroblasts lacked m^5C at the corresponding positions (Figure 3D, lower panels). vtRNAs share high sequence homology and the potentially methylated cytosines C15, C27, and C59/69 are present in all three ncRNAs (Figure S5A). Using LocARNA, a tool that simultaneously folds and aligns input RNA sequences (Smith et al., 2010), we predict that all three cytosines are located in vtRNA stem structures but are either unpaired or next to an unpaired nucleoside (Figure 3C; Figures S5A and S5B).

We noted that the consensus sequence for miCLIP target sites in vtRNAs was TCG (Figure 3C). Although this consensus sequence can vary in tRNAs and other RNAs (data not shown), it raises the possibility that NSun2-mediated methylation may have different molecular functions defined by the context of

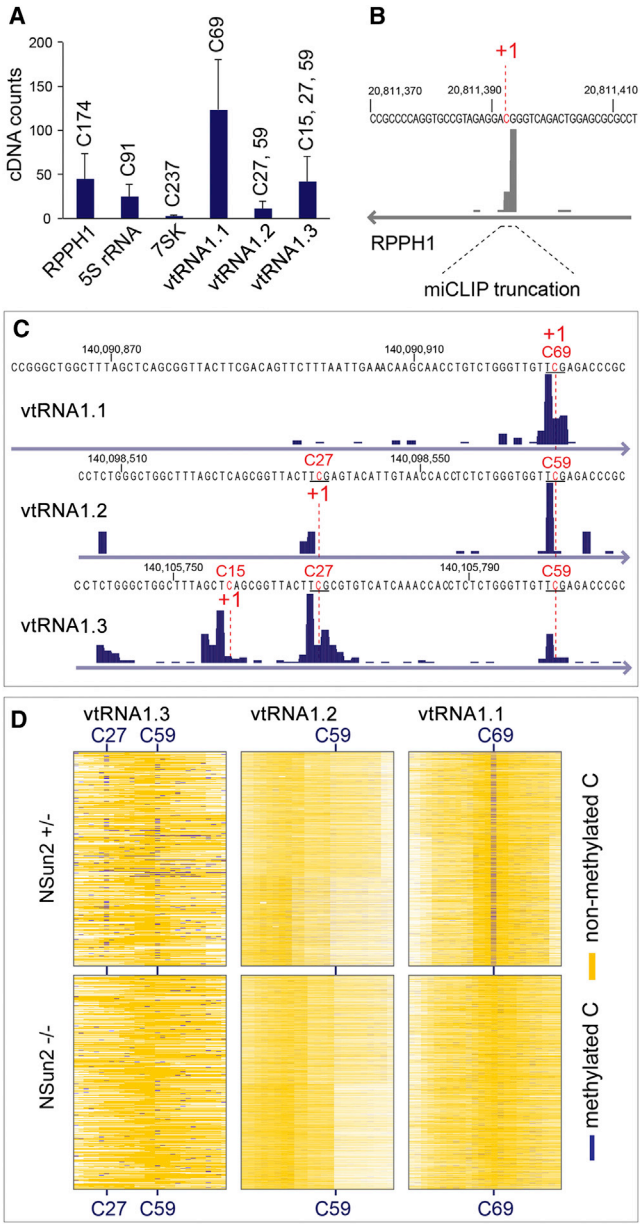


Figure 3. Identification of m^{5}C in Noncoding RNAs

(A) Total number of cDNAs and position of the m^{5}C modification mapping to RPPH1, 5S rRNA, 7SK, and vtRNA (vtRNA1.1, vtRNA1.2, vtRNA1.3) in three independent miCLIP experiments after 25 cycles of amplification. Error estimates represent SD of the mean.

(B and C) Detection of miCLIP sites mapped as a custom track on the UCSC genome browser in RPPH1 (B) and vtRNA1.1, vtRNA1.2, and vtRNA1.3 (C). +1 indicates the crosslinked cytosine. Underlined is the potential consensus site T(m^{5}C)G.

(D) RNA bisulfite sequencing showing the total number of reads with methylated (blue) and nonmethylated (yellow) cytosines in vtRNA1.1, vtRNA1.2, and vtRNA1.3 in *NSUN2*$^{+/-}$ and *NSUN2*$^{-/-}$ human fibroblasts.

See also Figure S4.

the surrounding sequence. T(m^{5}C)G might act as a vtRNA-specific recognition site for processing factors. Human vtRNAs can be processed into small RNAs (svRNAs) by a mechanism different from the canonical microRNA (miRNA) pathway, but at least some of these svRNAs can regulate gene expression similar to miRNAs (Persson et al., 2009). How the processing of vtRNAs into svRNAs is controlled was unknown, and we speculated that the deposition of m^{5}C into vtRNA might determine its processing into svRNAs.

Fragmentation of vtRNA into svRNAs Is Altered in the Absence of m^{5}C

To investigate whether methylation of the vtRNAs might affect their processing into svRNAs, we performed RNA sequencing of purified 15–35 nt fragments isolated from *NSUN2*$^{+/-}$ and *NSUN2*$^{-/-}$ human fibroblast. The sequence reads were aligned to vtRNAs, allowing no mismatch, and we only considered fragments with a total sum of 40 reads across four replicates in at least one of the conditions. We identified six vault ncRNA fragments, five of which mapped to vtRNA1.1 and one to vtRNA1.2 (Figure S6A).

Four out the five fragments derived from vtRNA1.1 exhibited statistically significant (false discovery rate [FDR] < 0.05) differential abundances in *NSUN2*$^{-/-}$ and *NSUN2*$^{+/-}$ fibroblasts (Figures 4A and 4B). Only one of these fragments (svRNA4) was strongly reduced in the absence of NSun2, and its 5′ start site precisely coincided with the NSun2-mediated methylation site at position C69 (Figures 4A and 4B). We confirmed differential abundances of svRNA1 and svRNA4 in *NSUN2*$^{-/-}$ cells by quantitative PCR (qPCR) (Figure 4C). To test whether the presence of svRNA4 depended on NSun2, we re-expressed full-length NSun2 (pB-NSun2) in *NSUN2*$^{-/-}$ fibroblasts via retroviral infection (Figure 4D; Figure S6B). The level of svRNA4 increased by more than 10-fold when NSun2 was re-expressed compared to control cells infected by an empty vector (pB-empty) (Figure 4E).

We speculated that m^{5}C might affect processing of vtRNA1.1. Deposition of m^{5}C at position 69 may induce the generation of svRNA4 by recruiting site-specific endonucleases. To test our hypothesis, we incubated synthetic vtRNA1.1 RNAs carrying or lacking m^{5}C at position 69 with cell lysate from *NSUN2*$^{-/-}$ fibroblasts, to avoid methylation of the synthetic RNA, and measured the abundance of svRNA4 by qPCR (Figure 4F). The levels of svRNA4 were highest when vtRNA1.1 was methylated in two independent experiments (Figure 4F; Figure S6C).

svRNA4 shares a high sequence homology to vtRNA fragments bound to Argonaute (GACCCGCGGGCGCUCUCCAGU CCUUU) (Burroughs et al., 2011). We isolated small RNAs that copurified with Argonaute proteins in *NSUN2*$^{+/-}$ fibroblasts and confirmed that svRNA4 bound to both Argonaute 2 and Argonaute 3 (Figure 4G). Given that vtRNA fragments can repress gene expression similarly to miRNAs (Persson et al., 2009), we reasoned that svRNA4 might have miRNA-like functions. We computationally predicted potential svRNA4 target mRNAs (Kertesz et al., 2007) and selected the top 100 potential target mRNAs for further analysis (Table S7). Because the loss of NSun2 function causes neurodevelopmental disorders in humans, we focused on mRNAs related to human diseases (Abbasi-Moheb et al., 2012; Khan et al., 2012; Martinez et al., 2012). Potential target mRNAs included CACNG7 and CACNG8, both of which encode voltage-gated calcium channels (Burgess

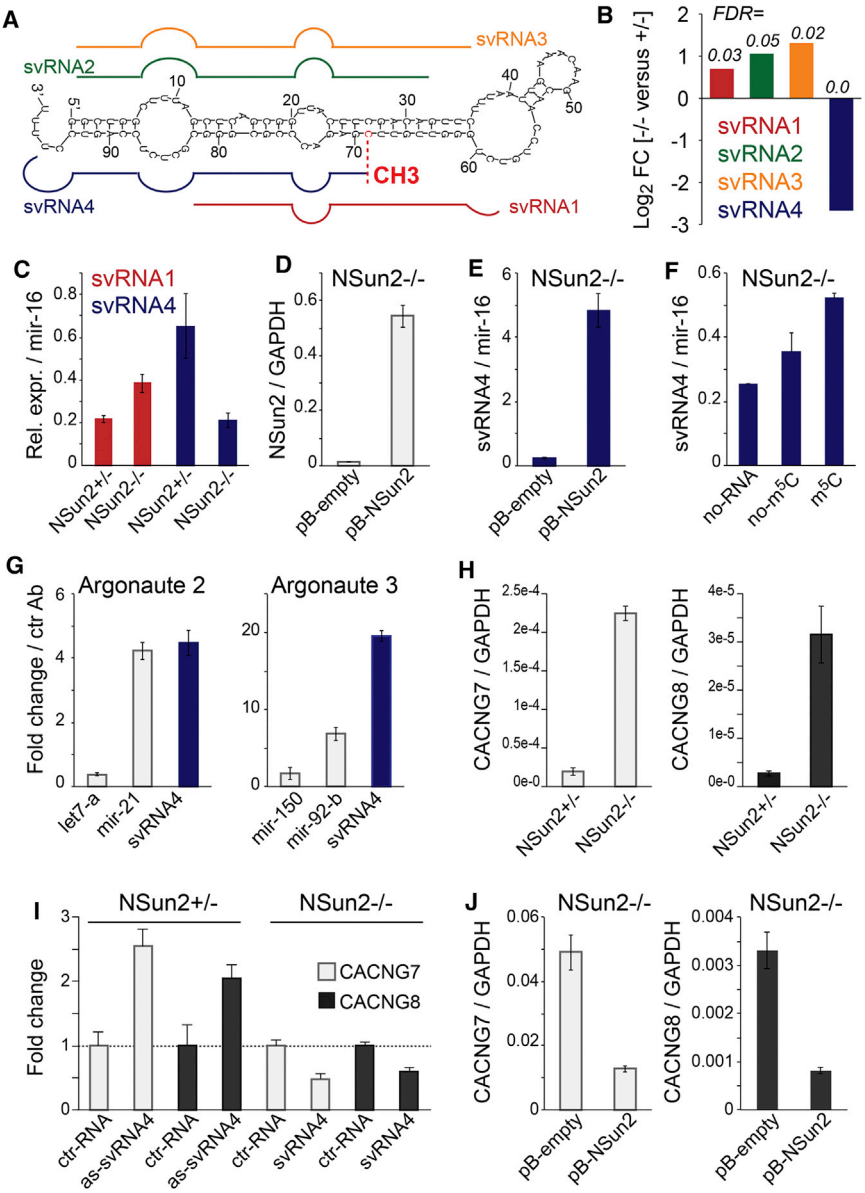


Figure 4. Differential Processing of vtRNA1.1 into svRNAs in the Absence of m^{5}C

(A) Schematics of secondary structure of vtRNA1.1 and small vault RNA (svRNA) found to be differentially abundant in *NSUN2*$^{+/-}$ and *NSUN2*$^{-/-}$ fibroblasts. CH3, cytosine-5 methylated site at position 69.

(B) Fold-change (log_2) and false discovery rate (FDR) values for reads of svRNA1-4 in *NSUN2*$^{+/-}$ versus *NSUN2*$^{-/-}$ human fibroblasts.

(C) Detection of svRNA1 and svRNA4 in *NSUN2*$^{+/-}$ and *NSUN2*$^{-/-}$ cells using qPCR.

(D and E) RNA levels of NSun2 (D) and svRNA4 (E) in NSun2 null (−/−) fibroblasts rescued by viral infection of NSun2 (pB-NSun2) compared to the empty vector control (pB-empty).

(F) Abundance of svRNA4 in *NSUN2*$^{-/-}$ cell lysates (no-RNA) or incubated with synthetic vtRNA1.1 carrying (m^{5}C) or lacking (no-m^{5}C) at position 69.

(G) Detection of svRNA4 in small RNA pool copurified with Argonaute 2 (left) and Argonaute 3 (right). Let7-a and mir-150 are negative and mir-21 and mir-92-b are positive controls for Argonaute 2- and 3-bound microRNAs, respectively.

(H) qPCR showing expression of CACNG7 and CACNG8 RNA relative to GAPDH in *NSUN2*$^{-/-}$ and *NSUN2*$^{+/-}$ fibroblasts.

(I) Fold-change expression of CACNG7 and CACNG8 RNA in *NSUN2*$^{-/-}$ and *NSUN2*$^{+/-}$ fibroblasts transduced with svRNA antagomirs (as-svRNA4) or svRNA4 microRNA mimics (svRNA) versus respective control RNAs (ctr-RNA).

(J) RNA levels of CACNG7 and CACNG8 relative to GAPDH in NSun2 null (−/−) fibroblasts rescued by viral infection of NSun2 (pB-NSun2) compared to the empty vector control (pB-empty). Error estimates represent SEM (C–J).

See also Figures S5 and S6.

et al., 2001; Waithe et al., 2011). Both CACNG7 and CACNG8 mRNA levels increased in *NSUN2*$^{-/-}$ fibroblasts, which was accompanied by increased CACNG8 protein levels (Figure 4H; Figure S6D).

To test whether svRNA4 exhibited miRNA-like functions, we transfected svRNA4 antagomirs (as-svRNA4) and miRNA mimics (svRNA4) into *NSUN2*$^{+/-}$ and *NSUN2*$^{-/-}$ fibroblasts (Figure 4I). Inhibition of svRNA4 by its antagomir increased the levels of CACNG7 and CACNG8 mRNAs in *NSUN2*$^{+/-}$ fibroblasts, whereas transfection of svRNA4 mimics decreased mRNA levels in *NSUN2*$^{-/-}$ fibroblasts (Figure 4I). Finally, we confirmed that rescue of *NSUN2*$^{-/-}$ fibroblasts by overexpressing full-length NSun2 not only increased the levels of svRNA4 but also reduced the levels of CACNG7 and CACNG8 mRNAs and CACNG8 protein (Figures 4D, 4E, and 4J; Figures S6B, S6C, and S6E). Together, these results demonstrate that the loss of NSun2-mediated methylation of vtRNAs alters their processing into svRNAs and indicates that this affects the levels of svRNA-regulated mRNAs.

DISCUSSION

Although the occurrence of methylated cytosine-5 was simultaneously discovered in DNA and RNA, the functional analysis of m^{5}C in RNA was hampered by the lack of suitable technical tools. Bisulfite sequencing to chemically identify m^{5}C in nucleosides was only recently adapted to RNA and confirmed site-specific methylation activity of NSun2 in tRNAs (Blanco et al., 2011; Martinez et al., 2012; Schaefer et al., 2009; Tuorto et al., 2012). System-wide RNA bisulfite sequencing confirmed NSun2-directed methylation of tRNAs and additionally identified CINP and NAPRT1 mRNAs as well as the RNA subunit RPPH1 of

RNaseP as potential NSun2-methylated targets (Squires et al., 2012). The recent development of the Aza-IP method, a technique that exploits covalent bond formation between RNA methylases and the cytidine analogue 5-azacytidine, further identified SCARNA2 and vtRNA1.1 as bona fide target RNAs of NSun2 (Khoddami and Cairns, 2013). However, both RNA bisulfite conversion and Aza-IP rely on the chemical modification of the target RNAs, which may compromise RNA stability and integrity. Thus, the development of complementary methods to conclusively determine global NSun2-specific methylated transcriptomes was essential.

We have successfully used a mutated version of the NSun2 protein to induce the formation of an irreversible covalent cross-link between NSun2 and the methylated cytosine in its target RNA. Whereas miCLIP requires expression of the mutant NSun2 protein, it is independent of any chemical modification of nucleosides. We successfully mapped NSun2-specific binding to m^5C in coding RNAs and ncRNAs. Our analysis identified around 300 common mRNAs among three replicates, indicating that NSun2-dependent methylation of coding RNAs is relatively rare. None of the miCLIP-identified mRNAs were differentially regulated in NSun2-depleted cells.

We confirmed vtRNAs as methylation substrates for NSun2 by RNA bisulfite sequencing using human skin fibroblast carrying a homozygous loss-of-function mutation in the *NSun2* gene (Martinez et al., 2012). vtRNAs are ncRNAs found as part of the vault ribonucleoprotein complex of unknown function (Kedersha and Rome, 1986). Interestingly, vtRNAs are significantly upregulated during neural differentiation (Skreka et al., 2012), and NSun2 deficiency in humans causes neurodevelopmental symptoms (Abbasi-Moheb et al., 2012; Khan et al., 2012; Martinez et al., 2012). However, the role of site-specific loss of m^5C in vtRNAs in the NSun2-mediated functions remained unclear.

vtRNAs can be processed into regulatory small RNAs in a Dicer-dependent mechanism (Friedländer et al., 2012; Langenberger et al., 2013; Persson et al., 2009). Our study suggests that vtRNA methylation may add another layer of regulation to this process. We identified four differential abundant vtRNA-derived svRNAs in patient fibroblasts lacking the NSun2 protein, yet only one of those (svRNA4) decreased in the absence of NSun2. Our study suggests a mechanism whereby vtRNA methylation may act as a molecular switch to produce specific svRNAs, which in turn influence distinct sets of mRNAs. We identified CACNG7 and CACNG8 as potential mRNAs regulated via the vtRNA-derived small RNA svRNA4. In conclusion, our data indicate that NSun2-mediated m^5C of vtRNAs regulates their processing into specific small RNAs and alterations in this pathway may contribute to symptoms found in humans with NSun2 deficiency.

EXPERIMENTAL PROCEDURES

miCLIP

Ethics approval was provided by the UCSD institutional review board. The Myc-tagged NSun2 C271A mutated construct (Hussain et al., 2009) or an empty vector control was transfected into COS7 or HEK293 cells using Lipofectamine 2000 (Invitrogen) and cells were harvested 24 hr later. NSun2 was immunoprecipitated with monoclonal Myc antibody (9E10), and iCLIP was subsequently performed as described before (König et al., 2010).

Bisulfite Sequencing

Total RNA was extracted from human fibroblasts with TRIzol (Invitrogen), DNase (Ambion), and Ribo-zero treated (Epicenter, Illumina). The remaining RNA fraction was bisulfite-converted as previously described (Blanco et al., 2011). Bisulfite-converted libraries were generated according to the TruSeq Small RNA Preparation Kit (Illumina), after 3′ and 5′ ends were repaired with T4 PNK (New England Biolabs). RNA was then reverse transcribed followed by 18-cycle PCR amplification.

Libraries for Deep Sequencing

For mRNA sequencing experiments, total RNA was obtained using TRIzol reagent (Invitrogen) and OligodT magnetic dynabeads were then used to isolate the mRNA fraction. For small RNA sequencing, total RNA from human fibroblasts was purified with the MirVana kit (Ambion) and then size-selected following separation on a 15% Novex TBE-Urea gel (Invitrogen). Libraries for sequencing were then prepared using the Illumina TruSeq Prep Kits and sequenced on the Illumina GAII platform.

Real-Time qPCR

qPCR was performed using TaqMan assay sets purchased from Applied Biosystems or using the QuantifastSYBR green system (QIAGEN).

A detailed description of all experimental procedures is available in the Extended Experimental Procedures.

ACCESSION NUMBERS

Sequencing data for cells derived from human patients have been deposited in the NIH Database of Genotypes and Phenotypes (dbGaP). The GEO accession number for the HEK293 sequencing data reported in this paper is GSE44386 (Table S8).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and eight tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.06.029.

ACKNOWLEDGMENTS

We are most grateful to everyone who provided us with reagents. We thank the CI Genomics and Bioinformatics Core Facilities. We gratefully acknowledge the support of the Cambridge Stem Cell Initiative and Stephen Evans-Freke. This work was funded by Cancer Research UK, the Medical Research Council, and the European Research Council. Y.S. is supported by the Nakajima Foundation.

Received: February 19, 2013
Revised: May 20, 2013
Accepted: June 21, 2013
Published: July 18, 2013

REFERENCES

Abbasi-Moheb, L., Mertel, S., Gonsior, M., Nouri-Vahid, L., Kahrizi, K., Cirak, S., Wieczorek, D., Motazacker, M.M., Esmaeeli-Nieh, S., Cremer, K., et al. (2012). Mutations in NSUN2 cause autosomal-recessive intellectual disability. Am. J. Hum. Genet. *90*, 847–855.

Blanco, S., Kurowski, A., Nichols, J., Watt, F.M., Benitah, S.A., and Frye, M. (2011). The RNA-methyltransferase Misu (NSun2) poises epidermal stem cells to differentiate. PLoS Genet. *7*, e1002403.

Brzezicha, B., Schmidt, M., Makalowska, I., Jarmolowski, A., Pienkowska, J., and Szweykowska-Kulinska, Z. (2006). Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). Nucleic Acids Res. *34*, 6034–6043.

Burgess, D.L., Gefrides, L.A., Foreman, P.J., and Noebels, J.L. (2001). A cluster of three novel Ca2+ channel gamma subunit genes on chromosome 19q13.4: evolution and expression profile of the gamma subunit gene family. Genomics *71*, 339–350.

Burroughs, A.M., Ando, Y., de Hoon, M.J., Tomaru, Y., Suzuki, H., Hayashizaki, Y., and Daub, C.O. (2011). Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. RNA Biol. *8*, 158–177.

Friedländer, M.R., Mackowiak, S.D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. *40*, 37–52.

Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.L., Zhang, X., Golic, K.G., Jacobsen, S.E., and Bestor, T.H. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science *311*, 395–398.

Hussain, S., Benavente, S.B., Nascimento, E., Dragoni, I., Kurowski, A., Gillich, A., Humphreys, P., and Frye, M. (2009). The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability. J. Cell Biol. *186*, 27–40.

Hussain, S., Tuorto, F., Menon, S., Blanco, S., Cox, C., Flores, J.V., Watt, S., Kudo, N.R., Lyko, F., and Frye, M. (2013). The mouse cytosine-5 RNA methyltransferase NSun2 is a component of the chromatoid body and required for testis differentiation. Mol. Cell. Biol. *33*, 1561–1570.

Kedersha, N.L., and Rome, L.H. (1986). Isolation and characterization of a novel ribonucleoprotein particle: large structures contain a single species of small RNA. J. Cell Biol. *103*, 699–709.

Kedersha, N.L., Miquel, M.C., Bittner, D., and Rome, L.H. (1990). Vaults. II. Ribonucleoprotein structures are highly conserved among higher and lower eukaryotes. J. Cell Biol. *110*, 895–901.

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007). The role of site accessibility in microRNA target recognition. Nat. Genet. *39*, 1278–1284.

Khan, M.A., Rafiq, M.A., Noor, A., Hussain, S., Flores, J.V., Rupp, V., Vincent, A.K., Malli, R., Ali, G., Khan, F.S., et al. (2012). Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. Am. J. Hum. Genet. *90*, 856–863.

Khoddami, V., and Cairns, B.R. (2013). Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat. Biotechnol. *31*, 458–464.

King, M.Y., and Redman, K.L. (2002). RNA methyltransferases utilize two cysteine residues in the formation of 5-methylcytosine. Biochemistry *41*, 11218–11225.

König, 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.

Langenberger, D., Çakir, M.V., Hoffmann, S., and Stadler, P.F. (2013). Dicer-processed small RNAs: rules and exceptions. J. Exp. Zoolog. B Mol. Dev. Evol. *320*, 35–46.

Liu, Y., and Santi, D.V. (2000). m5C RNA and m5C DNA methyl transferases use different cysteine residues as catalysts. Proc. Natl. Acad. Sci. USA *97*, 8263–8265.

Martinez, F.J., Lee, J.H., Lee, J.E., Blanco, S., Nickerson, E., Gabriel, S., Frye, M., Al-Gazali, L., and Gleeson, J.G. (2012). Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. J. Med. Genet. *49*, 380–385.

Mossink, M.H., van Zon, A., Scheper, R.J., Sonneveld, P., and Wiemer, E.A. (2003). Vaults: a ribonucleoprotein particle involved in drug resistance? Oncogene *22*, 7458–7467.

Persson, H., Kvist, A., Vallon-Christersson, J., Medstrand, P., Borg, A., and Rovira, C. (2009). The non-coding RNA of the multidrug resistance-linked vault particle encodes multiple regulatory small RNAs. Nat. Cell Biol. *11*, 1268–1271.

Rai, K., Chidester, S., Zavala, C.V., Manos, E.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2007). Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish. Genes Dev. *21*, 261–266.

Redman, K.L. (2006). Assembly of protein-RNA complexes using natural RNA and mutant forms of an RNA cytosine methyltransferase. Biomacromolecules *7*, 3321–3326.

Schaefer, M., Pollex, T., Hanna, K., and Lyko, F. (2009). RNA cytosine methylation analysis by bisulfite sequencing. Nucleic Acids Res. *37*, e12.

Schaefer, M., Pollex, T., Hanna, K., Tuorto, F., Meusburger, M., Helm, M., and Lyko, F. (2010). RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. Genes Dev. *24*, 1590–1595.

Skreka, K., Schafferer, S., Nat, I.R., Zywicki, M., Salti, A., Apostolova, G., Griehl, M., Rederstorff, M., Dechant, G., and Hüttenhofer, A. (2012). Identification of differentially expressed non-coding RNAs in embryonic stem cell neural differentiation. Nucleic Acids Res. *40*, 6001–6015.

Smith, C., Heyne, S., Richter, A.S., Will, S., and Backofen, R. (2010). Freiburg RNA Tools: a web server integrating INTARNA, EXPARNA and LOCARNA. Nucleic Acids Res. *38*(Web Server issue), W373-7.

Squires, J.E., Patel, H.R., Nousch, M., Sibbritt, T., Humphreys, D.T., Parker, B.J., Suter, C.M., and Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res *40*, 5023–5033.

Stadler, P.F., Chen, J.J., Hackermüller, J., Hoffmann, S., Horn, F., Khaitovich, P., Kretzschmar, A.K., Mosig, A., Prohaska, S.J., Qi, X., et al. (2009). Evolution of vault RNAs. Mol. Biol. Evol. *26*, 1975–1991.

Sugimoto, Y., König, J., Hussain, S., Zupan, B., Curk, T., Frye, M., and Ule, J. (2012). Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. Genome Biol. *13*, R67.

Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. *9*, 465–476.

Tuorto, F., Liebers, R., Musch, T., Schaefer, M., Hofmann, S., Kellner, S., Frye, M., Helm, M., Stoecklin, G., and Lyko, F. (2012). RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat. Struct. Mol. Biol. *19*, 900–905.

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.

Waithe, D., Ferron, L., and Dolphin, A.C. (2011). Stargazin-related protein γ_7 is associated with signalling endosomes in superior cervical ganglion neurons and modulates neurite outgrowth. J. Cell Sci. *124*, 2049–2057.

Zhang, X., Liu, Z., Yi, J., Tang, H., Xing, J., Yu, M., Tong, T., Shang, Y., Gorospe, M., and Wang, W. (2012). The tRNA methyltransferase NSun2 stabilizes p16INK^4 mRNA by methylating the 3′-untranslated region of p16. Nat. Commun. *3*, 712.

Cell PRESS
Open ACCESS

EXTENDED EXPERIMENTAL PROCEDURES

Cell Culture

HEK293 cells were grown in Dulbecco's Minimal Essential Medium (DMEM) (Life Technologies) supplemented with 10% fetal bovine serum. Human fibroblasts were grown in Minimal Essential Medium (MEM) (Life Technologies) supplemented with 20% fetal bovine serum. All cells were cultured in humidified atmospheres with 5% CO_2.

miCLIP

The Myc-tagged NSun2 C271A mutated construct (Hussain et al., 2009) or an empty vector control were transfected into COS7 or HEK293 cells using Lipofectamine 2000 (Invitrogen) and cells were harvested 24 hr later. miCLIP was subsequently performed as described previously (König et al., 2010). Briefly, cells were lysed in lysis buffer consisting of 50 mM Tris-HCL pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate. Lysates were then treated with high concentration of DNase and low concentration of RNaseI to partially fragment RNAs. Lysates were then cleared by centrifugation at 13,000 rpm for 15 min at 4°C and then incubated with Protein G Dynabeads (Invitrogen) in the presence of an anti-Myc antibody (9E10, Sigma). Following stringent washing, 3′ end dephosphorylation was performed with T4 polynucleotide kinase (New England Biolabs) before addition of a pre-adenylated linker using RNA ligase (New England Biolabs). 5′ end labeling was then performed using T4 PNK and ^{32}P-ATP before protein-RNA complexes were eluted and run on denaturing gels. Nitrocellulose transfer was then performed and the radioactive signal was used to dissect nitrocellulose pieces that contained NSun2-partially digested RNA complexes. RNA was recovered by incubating the nitrocellulose pieces in a buffer containing Proteinase K and 3.5 M urea. Next, reverse transcription was performed using oligonucleotides for reverse transcription containing two inversely oriented adaptor regions separated by a BamHI restriction site as well as a barcode region at their 5′ end containing a 5-nucleotide random barcode to enable tracing of individual cDNAs. cDNAs were size-purified on TBE-Urea gels before being circularized by CircLigase II (Epicenter). Circularized cDNAs were then annealed to an oligonucleotide complementary to the BamHI site and then BamHI digested. Linearized cDNAs were then PCR-amplified using primers complementary to the adaptor regions using 25 or 35 cycles of PCR. Libraries were then subjected to high-throughput sequencing using the Illumina GA2 platform.

Libraries for Deep Sequencing

For mRNA sequencing experiments, siRNA-mediated knockdown of NSun2 in HEK293 cells was performed by transfecting NSun2 siRNA (5′-CACGTGTTCACTAAACCCTAT-3′) or a control scrambled sequence siRNA using RNAiMAX transfection reagent (Invitrogen). Human fibroblasts, and siRNA-transfected HEK293 cells were harvested 48 hr post-transfection, and total RNA isolated using TRIzol reagent (Invitrogen). OligodT magnetic dynabeads were then used to isolate the mRNA fraction and libraries for sequencing were then prepared using the Illumina TruSeq mRNA Prep Kit.

For small RNA sequencing, total RNA from human fibroblasts was purified with the MirVana kit (Ambion) in order to isolate the < 200bp fraction of RNA. RNAs were then separated on a 15% Novex TBE-Urea gel. Gel pieces containing RNAs of size 15-35 nucleotides were then excised and RNAs were eluted by incubating crushed gel pieces in water. The sample was then ethanol precipitated to recover RNA and these samples were then used as the input for the Illumina small RNA Prep Kit. All next generation sequencing was performed on the Illumina GA2 platform. All experiments were performed in quadruplicates.

Real-Time qPCR

RNA was isolated from human fibroblasts using TRIzol and cDNA synthesis performed using the SuperscriptIII reverse transcriptase kit (Invitrogen). Unless otherwise stated, qPCR was then performed using TaqMan assay sets purchased from Applied Biosystems as per manufacturers recommendations. TaqMan probe sets used were CACNG7: Hs00259061_m1, CACNG8: Hs01100182_m1, NSUN2: Hs00214829_m1, and GAPDH (4326317E) as an internal control.

Purification of MicroRNA and Detection by qPCR

Total RNA from *NSUN2*$^{+/-}$ and *NSUN2*$^{-/-}$ dermal fibroblasts was isolated using TRIzol (Invitrogen). RNAs of < 200 base pairs were enriched using the mirVana kit (Ambion). The RNAs were then separated using 15% acrylamide-urea gels (Invitrogen) and gel slices containing10 to 40 base pairs were excised and incubated with RNA free water for 2 hr at 4°C to elute RNA. RNA was then purified using spin-X centrifuge tube filters (0.22 μM) (Costar) and ethanol precipitated over night at −20°C. Small RNA Reverse transcription and qPCR was performed as previously described (Shi and Chiang, 2005). Briefly, polyA tails were added to small RNAs using poly(A) polymerase (Ambion). The poly(T) primer 5′-GCG AGC ACA GAA TTA ATA CGA CTC ACT ATA GG(T)12VN-3′ was then used for cDNA synthesis. QuantifastSYBR® green master mix (QIAGEN) was used for QPCR reactions. The forward primers were: miR-16 (5′-TAG CAG CAC GTA AAT ATT GGC G-'3), miR-150 (5′ TCT CCC AAC CCT TGT ACC AGT G 3′), miR-92b (5′ AGG GAC GGG ACG CGG TGC AGT G 3′), svRNA1 (5′-TGT CTG GGT TGT TCG AGA CCC GCG GGC-3′), svRNA4 (5′-CGA GAC CCG CGG GCG CTC TCC AGT CCT TTT-'3). The reverse primer for all reactions was: 5′-GCG AGC ACA GAA TTA ATA CGA C-3′.

In Vitro Cleavage Assay of VTRNA1.1

Nsun2 −/− dermal fibroblasts were lysed in a buffer consisting of 1% NP-40, 25 mM Tris-HCl at pH 7.4, 30 mM NaCl and 10 mM $MgCl_2$ for 30 min on ice and lysates were then pooled and divided equally between samples. 100 pmol of synthetic VTRNA1.1 RNA that either was unmodified or contained the m^5C modification at position 69 (Thermo Scientific) was then added to samples, whereas no oligo was added to control samples. Incubation was carried out at room temperature or 37°C for 1 hr and phenol-chloroform extraction was subsequently performed to recover RNA. Purification of microRNAs and their quantification was then performed as described earlier. To account for possible subtle variations in microRNA recovery efficiencies between samples, svRNA4 abundance was always measured relative to miR-16 levels.

Argonaute Coimmunoprecipitations

Human dermal fibroblasts were lysed in a buffer containing 1% NP-40, 25 mM Tris-HCl at pH 8.0, 150 mM NaCl, 2 mM $MgCl_2$ and 1 mM DTT and incubated on ice for 30 min. Rat monoclonal Ago2 (Sigma, 11A9) or mouse monoclonal Ago3 (Merck Millipore, 03-250) or control antibodies were pre-incubated with Protein G magnetic dynabeads in lysis buffer for 30 min. Following washing of beads, cleared lysates were added to the beads and immunoprecipitation was carried out for 2 hr at 4°C with gentle mixing. Beads were then washed extensively in lysis buffer and precipitated RNAs were recovered from the beads using TRIzol. Further purification of microRNAs and subsequent qPCR quantifications were performed as described earlier.

Small RNA Transfections

NSun2 -/+ and −/− dermal fibroblasts were plated onto six well plates. When they had reached 40% confluence, cells were transfected with a control oligo or svRNA4 antagomirs (2′-O-methyl-anti-svRNA4: 5′ AAA AGG ACU GGA GAG CGC CCG CGG GUC UCG); or control microRNA miRIDIAN mimic or svRNA4 microRNA miRIDIAN mimic (5′ CGA GAC CCG CGG GCG CUC UCC AGU CCU UUU) (designed and synthesized by Thermo Scientific) using RNAimax transfection reagent (Invitrogen). After 48 hr, cells were washed in PBS and RNA was isolated using TRIzol. CACNG7 and CACNG8 qPCR quantification was then performed as described earlier.

Rescue of NSun2 Null Cells

To rescue expression of NSun2 in *NSUN2*$^{-/-}$ dermal fibroblasts, full length NSun2 (pB-NSun2) or the empty vector (pB-empty) were infected via retrovirus as described previously (Hussain et al., 2009).

Western Blotting

Cells were harvested and lysed cells in lysis buffer consisting of 50 mM Tris-HCL pH 7.4, 250 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate before they were cleared by centrifugation at 13,000 rpm for 15 min at 4°C. Lysates were then mixed with SDS protein sample buffer and electrophoresed on a 10% SDS polyacrylamide gel. Gels were blotted onto nitrocellulose membranes, which were incubated in TBST-blocking solution (Tris-buffered saline, pH 8.8, with 5% marvel milk powder). Blots were incubated with primary antibodies in blocking solution followed by incubation with the appropriate HRP-conjugated secondary antibodies. The chemiluminescent signal was detected using an ECL chemiluminescent kit (GE Healthcare) according to the manufacturer's instructions. Antibodies used were Actin from Santa Cruz (sc-8432), CACNG8 from Santa Cruz (sc-168396), Myc 9E10 from Sigma (C6594).

Bioinformatics Analyses

For miCLIP data, potential PCR duplicates were removed, and reads with identical barcodes were identified to determine exact cDNA counts. Adapters and barcodes were removed, and reads were aligned to the human reference genome (GRCh37/hg19) using Bowtie (Langmead et al., 2009). Only reads uniquely aligning to the genome with one mismatch were considered for further analysis. The reads that aligned to tRNA genes annotated in the Genomic tRNA database (GtRNAdb, October 2012; http://lowelab.ucsc.edu/GtRNAdb/) were extracted, and the miCLIP read stop positions in the tRNA genes were determined. Read counts were normalized based on the total number of reads in the replicate samples per thousand reads. Alignments of the tRNA genes according to their secondary structures were generated manually, and cDNA counts per miCLIP stop position in the genes were mapped to the corresponding position in the alignment using custom scripts. Non-tRNA data were mapped as described previously (Sugimoto et al., 2012).

All protein-coding and further noncoding gene set analyses were conducted based on the Ensembl annotation (release 68). For mRNaseq, reads were mapped to the human reference genome (Ensembl release 68, GRCh37) using Tophat (Trapnell et al., 2009). The parameters used were a mate-inner-dist of 70, mate-std-dev of 100 and splice-mismatches of 2. Raw counts per gene were determined by using the Python HTSeq package (http://www-huber.embl.de/users/anders/HTSeq). Normalization and differential expression analysis was performed using the Bioconductor/R package DESeq (Anders and Huber, 2010). DESeq estimates the fold change for each gene, p values and the Benjamini-Hochberg FDR are calculated to statistically test the measured differential expression. The cut-off used for a statistically significantly differentially expressed gene was an adjusted p-value of less than 0.1.

For mRNaseq, Fastq files were mapped to the Ensembl release 68 human reference genome (GRCh37 assembly) using Tophat (Trapnell et al., 2009). The parameters used were a mate-inner-dist of 70, mate-std-dev of 100 and splice-mismatches of 2. In order

to measure gene expression from mapped data, the BAM files from Tophat mapping were sorted using SAMtools. Raw counts per gene were estimated by the Python script HTSeq count and a list of genomic features downloaded from Ensembl (http://www-huber.embl.de/users/anders/HTSeq/). The counts were combined for each sample and differential expression analysis was performed using the R package DESeq (Anders and Huber, 2010). DEseq generates a fold change for each gene, p values and the Benjamini-Hochberg FDR are calculated to statistically test the measured DE. The cut-off used for a statistically significantly DE gene was an adjusted p-value of less than 0.1.

For small RNA sequencing data, adapters were removed, and the 20-35 nucleotide fragments were aligned to the human reference genome (UCSC GRCh37/hg19) using Bowtie (Langmead et al., 2009), allowing no mismatch and discarding all non-unique alignments. For all small RNA fragments aligning at ncRNA gene positions, individual counts were normalized and differential abundance of the fragments was evaluated with the Bioconductor/R package DESeq (Anders and Huber, 2010).

For bisulfite sequencing data, reads were mapped to the Ensembl reference genome using Bismark (Krueger and Andrews, 2011). The reference has been processed using the script bismark_genome_preparation and the mapper transforms reads into a C-to-T and G-to-A version. One non-bisulfite mismatch per read was tolerated. The average percentage of reads aligned to the reference was ~70%. Aligned reads were annotated using the HTseq package (http://www-huber.embl.de/users/anders/HTSeq/) based on Ensembl gene annotations. Using custom scripts, alignments of bisulfite-treated reads for each gene were extracted. These alignments were then converted to a binary matrix with a −1/+1 determining if a methylcytosine or cytosine was at a position in a protein-coding or ncRNA gene. Heatmaps were generated using the program matrix2png from this matrix (http://www.chibi.ubc.ca/matrix2png/).

To ensure patient confidentiality, all sequencing data for human fibroblasts will be deposited on dbGaP or are available upon request.

SUPPLEMENTAL REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572.

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.

Shi, R., and Chiang, V.L. (2005). Facile means for quantifying microRNA expression by real-time PCR. Biotechniques *39*, 519–525.

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

Cell PRESS
Open ACCESS

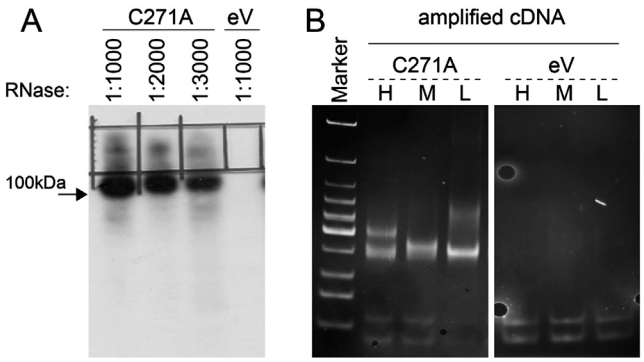


Figure S1. NSun2-RNA Complex Detection and cDNA Amplification, Related to Figure 1

(A) Radiolabeled NSun2-RNA complexes run on denaturing gels and blotted onto nitrocellulose membranes. Marking indicates NSun2-RNA complexes of higher molecular weight dissected for library preparation. A range of RNaseI conditions were tested and a dilution of 1:1000 gave the best results. HEK293 cells transfected with empty vector (eV) used as a negative control to exclude contamination in the subsequent PCR step during library preparation.

(B) TBE polyacrylamide gel showing PCR miCLIP cDNA fragments after 25 cycles of amplification. DNA was stained using ethidium bromide. High- (H), Medium- (M), and Low- (L) indicates the size of excised TBE-Urea gels fragments before the amplification step.

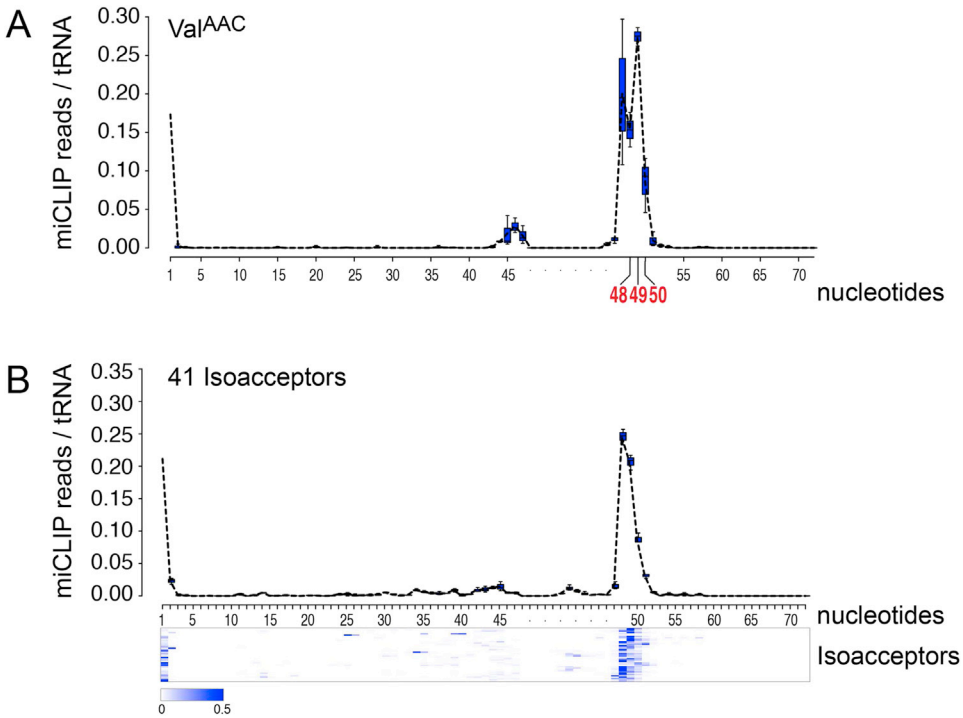


Figure S2. Identification of m^{5}C in tRNAs Using miCLIP, Related to Figures 2A and 2B

(A) Frequency of miCLIP reads in tRNA Val^{AAC}.

(B) Frequency of miCLIP reads and heat map identify position 48, 49 and 50 as main NSun2-targeted site in tRNA isoacceptors.

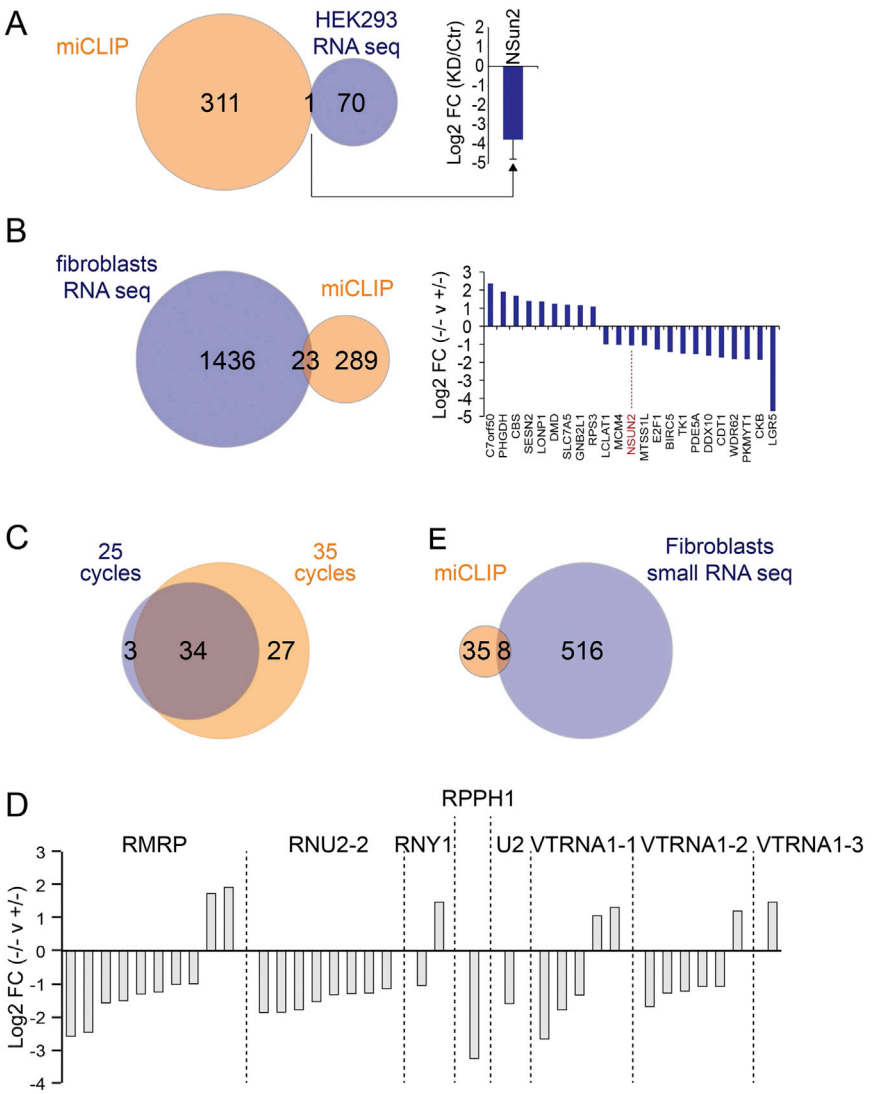


Figure S3. miCLIP Identified RNAs, Related to Figures 2C and 2D

(A) Venn diagram (left hand panel) of miCLIP identified mRNAs and differentially expressed mRNAs in HEK293 cells overexpressing an NSun2 RNAi versus a scrambled vector control identifies NSun2 as the only common target (right hand panel). KD: knockdown; Ctr: scrambled vector control).

(B) Venn diagram (left hand panel) showing miCLIP mRNA targets and differentially expressed mRNAs in *NSUN2* (+/−) versus (−/−) human fibroblasts reveal a small over-lap of 23 (7%) that can be both up- and downregulated in *NSUN2* (+/−) versus (−/−) human fibroblasts (right hand panel).

(C) Venn diagram miCLIP identified non-coding RNA in HEK293 cells after 25 or 35 cycles of amplification.

(D) Venn diagram miCLIP non-coding RNA targets with differentially abundant small RNAs in *NSUN2* (+/−) versus (−/−) human fibroblasts.

(E) Log_2-fold change (FC) of non-coding RNAs obtained by small RNA sequencing in NSun2 null (−/−) versus NSun2-expressing (+/−) fibroblasts identified in Figure 2E.

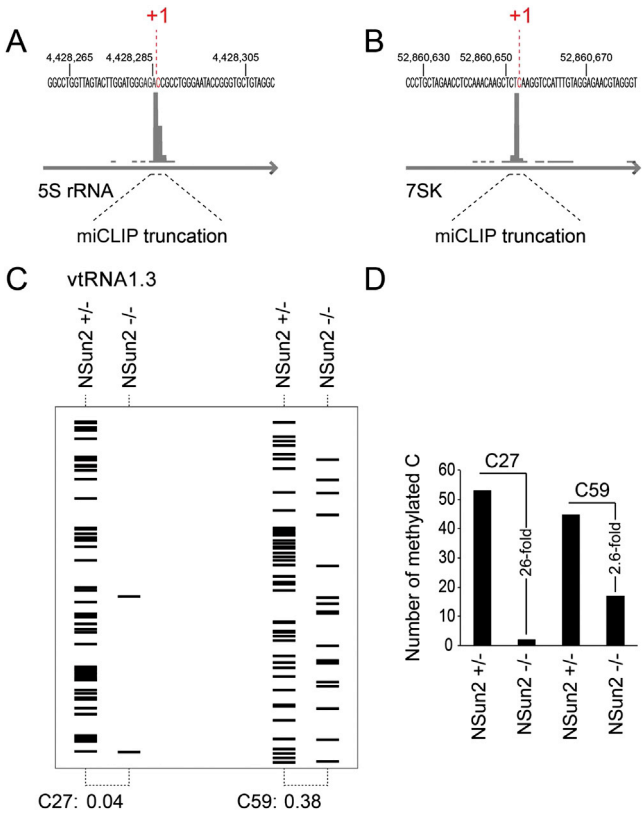


Figure S4. Identification of m^{5}C Using miCLIP and Bisulfite Conversion, Related to Figure 3

(A and B) Detection of miCLIP sites using UCSC genome browser in 5S rRNA (A) and 7SK (B).

(C) Schematic representation of the number of m^{5}C modifications in bisulfite converted VTRNA1.3 at positions 27 and 59 in *NSUN2*$^{+/-}$ and *NSUN2*$^{-/-}$ cells shown in Figure 3D.

(D) Quantification of (C).

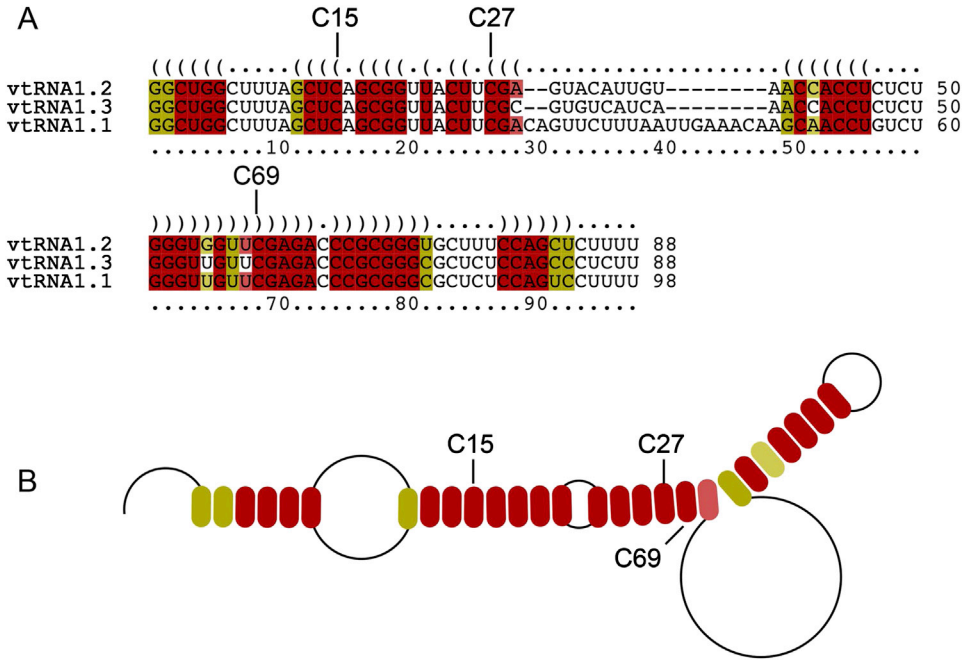


Figure S5. Structural Position of Methylated Cytosines in VtRNAs, Related to Figure 4

Sequence (A) and structural (B) alignment of VTRNA1.1, VTRNA1.2 and VTRNA1.3 using LocARNA (http://rna.informatik.uni-freiburg.de). All miCLIP identified sites are located within the RNA stem structure (C15, C27,C59/69).

A

Vault ncRNA	*Sequence*	*svRNA*
vtRNA1.1	TGTCTGGGTTGTTCGAGACCCGCGGGC	svRNA1
vtRNA1.1	GGGCTGGCTTTAGCTCAGCGGTTACTTCGACAG	svRNA2
vtRNA1.1	GGGCTGGCTTTAGCTCAGCGGTTACTTCGACAGTTC	svRNA3
vtRNA1.1	CGAGACCCGCGGGCGCTCTCCAGTCCTTTT	svRNA4
vtRNA1.1	TTGTTCGAGACCCGCGGGCGCTCTCCAGTCCTT	--
vtRNA1.2	CGCGGGTGCTTTCCAGCTCTTT	--

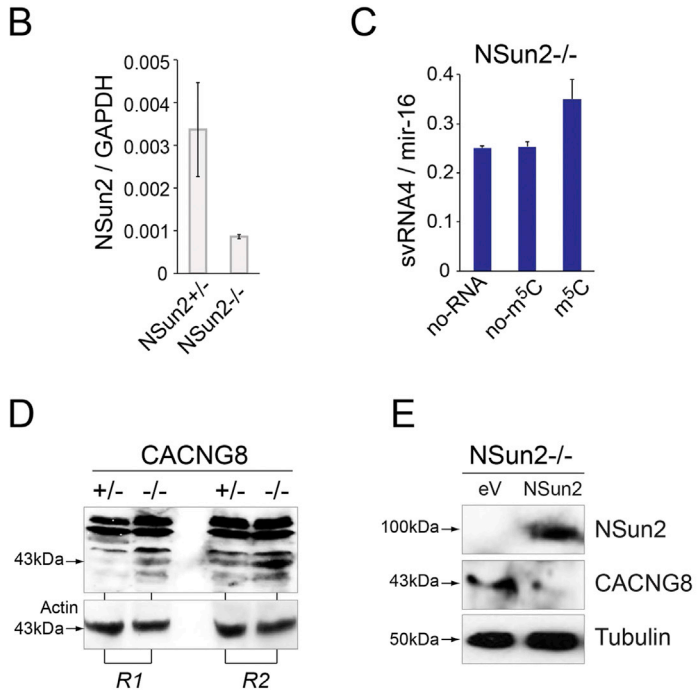


Figure S6. Processing of VTRNA1.1 into svRNAs, Related to Figure 4

(A) Complete list of svRNAs derived from VtRNAs found in human fibroblasts.

(B) QPCR showing expression of NSun2 RNA relative to GAPDH in *NSUN2* (−/−) and (+/−) fibroblasts.

(C) Abundance of svRNA4 in *NSUN2*$^{-/-}$ cell lysates (no-RNA) or incubated with synthetic VTRNA1.1 carrying (m^5C) or lacking (no-m^5C) at position 69 at 37°C.

(D) Western Blot detecting CACNG8 protein (43kDa) (upper panel) in *NSUN2* (−/−) and (+/−) fibroblasts in two independent replicates (R1, R2). Bands at higher molecular weight are unspecific. Actin served as a loading control (lower panel).

(E) Western Blot detecting NSun2 (100kDa) (upper panel) and CACNG8 protein (43kDa) (middle panel) in *NSUN2* −/− cells infected with an empty vector control (eV) or full length NSun2 (NSun2). Tubulin (50kDa) (lower panel) served as loading control.