Loss of 5-methylcytosine alters the biogenesis of Vault-derived small RNAs to coordinate epidermal differentiation

Abdulrahim A. Sajini[1,2,3], Nila Roy Choudhury[4], Rebecca E. Wagner[1], Susanne Bornelöv[5], Tommaso Selmi[1], Christos Spanos[6], Sabine Dietmann[5], Juri Rappsilber[6,7], Gracjan Michlewski[4,6,8*], and Michaela Frye[1,9*]

[1] Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, United Kingdom; [2] Department of Biomedical Engineering, Khalifa University of Science and Technology, P.O. Box 127788, Abu Dhabi, United Arab Emirates; [3] Department of Medical Laboratory Technology, University of Tabuk, Tabuk, P.O. Box 71491, Saudi Arabia. [4] Division of Infection and Pathway Medicine, University of Edinburgh, The Chancellor's Building, 49 Little France Crescent, Edinburgh, EH16 4SB, United Kingdom; [5] Wellcome MRC Cambridge Stem Cell Institute, Tennis Court Road, Cambridge, CB2 1QR, United Kingdom; [6] Wellcome Centre for Cell Biology, University of Edinburgh, Michael Swann Building, Max Born Crescent, Edinburgh, EH9 3BF, United Kingdom; [7] Department of Biotechnology, Technische Universität Berlin, Gustav-Meyer-Allee 25, 13355 Berlin, Germany; [8] ZJU-UoE Institute, Zhejiang University, 718 East Haizhou Road, Haining, Zhejiang 314400, P.R. China; [9] German Cancer Research Centre (Deutsches Krebsforschungszentrum, DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany.

* corresponding authors: Michaela Frye (m.frye@dkfz.de); Gracjan Michlewski (Gracjan.Michlewski@ed.ac.uk)

Sequencing data are up-loaded onto dbGaP (phs000645.v4.p1) and GEO (GSE122600, GSE125046)

Abstract

The presence and absence of RNA modifications regulates RNA metabolism by modulating the binding of writer, reader, and eraser proteins. For 5-methylcytosine (m^5C) however, it is largely unknown how it recruits or repels RNA-binding proteins. Here, we decipher the consequences of m^5C deposition into the abundant non-coding vault RNA VTRNA1.1. Methylation of cytosine 69 in VTRNA1.1 occurs frequently in human cells, is exclusively mediated by NSUN2, and determines the processing of VTRNA1.1 into small-vault RNAs (svRNAs). We identify the serine/arginine rich splicing factor 2 (SRSF2) as a novel VTRNA1.1-binding protein that counteracts VTRNA1.1 processing by binding the non-methylated form with higher affinity. Both NSUN2 and SRSF2 orchestrate the production of distinct svRNAs. Finally, we reveal a functional role of svRNAs in regulating the epidermal differentiation programme. Thus, our data reveal a direct role for m^5C in the processing of VTRNA1.1 that involves SRSF2 and is crucial for efficient cellular differentiation.

Introduction

The post-transcriptional deposition of chemical modifications into RNA emerged as a crucial regulator of gene expression programs [1]. 5-methylcytosine (m^5C) occurs in various RNA molecules and is mediated by at least eight, highly conserved enzymes (NSUN1-7, and DNMT2) in mammals [2]. One of the best characterized m^5C methyltransferase is NSUN2, which targets the majority of cytoplasmic transfer RNAs (tRNAs) and a smaller number of coding and non-coding RNAs including VTRNAs [3-7].

In tRNAs, NSUN2-mediated formation of m^5C protects from endonucleolytic cleavage [7,8]. Loss of tRNA methylation enhances the affinity to the endonuclease angiogenin, which then cleaves the tRNAs causing a global reduction in protein synthesis [7,8]. The cellular consequences of NSUN2-deletion are reduced cellular migration and delayed activation of stem cell differentiation [8-10]. NSUN2-mediated RNA methylation is required for normal development [7], and loss-of-function mutations in human *NSUN2* gene is associated with neuro-developmental disorders [11-14].

The functional role of m^5C in VTRNAs is less clear. VTRNAs are integral components of large ribonucleoprotein vault particles found in the cytoplasm of most eukaryotic cells [15,16]. However, only about 5% of cytoplasmic VTRNA is directly associated to vault particles and similarly small amounts of VTRNAs are reported to reside in the nucleus [17,18]. In humans, four VTRNAs are expressed VTRNA1.1, VTRNA1.2, VTRNA1.3, and VTRNA2.1 [16], two of which (VTRNA1.1 and VTRNA1.3) are methylated by NSUN2 [3]. VTRNAs have been implicated in the cellular immune response, cell survival and oncogenic multi-drug resistance, indicating a functional role in several fundamental biological processes [17,19-23].

VTRNAs are also processed into smaller regulatory RNAs (svRNA) by a pathway different from microRNA (miRNA) biogenesis [21]. VTRNA-derived svRNAs are highly abundant in exosomes, and at least some of them regulate gene expression similarly to miRNAs [3,21,24,25]. We previously revealed that the processing of full-length VTRNA1.1 into svRNAs depended on the methylation of cytosine 69 (C69) [3], yet the underlying molecular mechanisms and the functional role of the svRNAs remained unknown.

Here, we performed mass spectrometry-based quantitative proteomics to identify all proteins whose binding affinity is directly determined by the presence or absence of m^5C69 in VTRNA1.1. We identify SRSF2 as a novel VTRNA-binding protein that is repelled by m^5C69. By binding the un-methylated form with higher affinity, SRSF2 protects VTRNA1.1 from processing. We confirm that both NSUN2 and SRSF2 coordinate the processing of VTRNA1.1 into specific svRNAs. Functionally, we show that the presence of one specific VTRNA-derived small non-coding RNA (svRNA4) is sufficient to alter the transcriptional program needed to induce epidermal differentiation. Together, we demonstrate that the deposition of m^5C orchestrates VTRNA1.1 processing and thereby determines its downstream biological function.

Results

Methylation of VTRNA1.1 requires NSUN2

NSUN2 methylates the vast majority of tRNAs and a small number of coding and non-coding RNAs [1]. To determine which of these methylated sites solely depended on NSUN2, we rescued human dermal fibroblasts lacking a functional NSUN2 protein (*NSUN2*-/-) by re-expressing NSUN2 or an enzymatic dead version of the enzyme (K190M) [11,26]. We confirmed cytosine-methylation using RNA bisulfite (BS) sequencing [7]. As expected, the methylation levels between *NSUN2*-/- cells infected with the empty vector control and the enzymatic dead NSUN2 (K190M) highly correlated (Fig. 1a). In contrast, methylation of more than 100 sites significantly increased when NSUN2 was re-expressed (Fig. 1b; Supplementary Data 1). We confirmed that the NSUN2 and K190M proteins were equally expressed in the rescued *NSUN2*-/- cells (Supplementary Figure 1a).

In addition to tRNAs, we confirmed NSUN2-specific methylation of cytosine (C) 69 in VTRNA1.1 (Fig. 1c) [3]. Furthermore, we identified a small number of novel high confidence sites in both coding and non-coding RNAs (Supplementary Figure 1b-e). NSUN2-dependent methylation sites in VTRNAs, RPPH1, and HECTD1 for instance, were consistently methylated in human cells, including HEK293 and human embryonic stem cells (H9) (Fig. 1d, e). Notably, the methylation levels of C69 in VTRNA1.1 varied and were usually lower than 50%, even when NSUN2 was over-expressed (Fig. 1c-e; left hand panels). We concluded that methylation of VTRNA1.1 at C69 occurred at dynamic levels but was widely present in human cells.

VTRNA1.1 methylation determines the biogenesis of svRNA4

The functional relevance of C69 methylation in VTRNA1.1 was unknown. However, our previous study demonstrated that the presence and absence of m^5C correlated with the

differential processing of VTRNA1.1 into small non-coding RNA fragments (svRNA1-4) (Supplementary Figure 1f) [3]. Of these svRNAs, only the length of svRNA4 coincided with C69 (Fig. 1f; Supplementary Figure 1f). SvRNA4 was more abundant in human dermal fibroblasts expressing NSUN2 (Fig. 1g), indicating that the formation of svRNA4 was enhanced when VTRNA1.1 carried m^5C69 [11]. We asked whether the presence of m^5C69 increased the formation of svRNA4 and quantified svRNA4 in *NSUN2*-/- cells re-expressing the wild-type (wt) or enzymatic dead versions of NSUN2 (C321A; C271A) [8,26]. The processing of VTRNA1.1. into svRNA4 depended on the methylation activity of NSUN2 because only the wild-type construct of NSUN2 increased svRNA4 production (Fig. 1g). All over-expressed constructs were equally up-regulated in the *NSUN2*-/- cells (Fig. 1h) [8]. Thus, the presence of a methylation group at C69 enhanced the processing of VTRN1.1 into svRNA4.

Proteins binding to un-methylated and methylated VTRNA1.1

To dissect how VTRNA1.1 processing was regulated, we sought to identify all RNA-binding proteins showing a higher affinity to methylated or un-methylated VTRNA1.1. We performed quantitative RP-SMS (RNA pull-down SILAC (stable isotope labeling with amino acids in cell culture) mass spectrometry) in two independent experiments (Supplementary Figure 2a; Supplementary Data 2 and 3) [27]. We found a high correlation of identified proteins between the technical replicates (Supplementary Figure 2b) and identified a total of 144 proteins commonly bound to VTRNA1.1 in two independent experiments (Fig. 2a; Supplementary Figure 2c). Gene Ontology analyses confirmed that we significantly enriched for proteins binding to single and double stranded RNAs (Fig. 2b; Supplementary Data 4).

As expected, the majority of RNA-binding proteins were not differentially bound to methylated or un-methylated VTRNA1.1, this included proteins related to m^5C deposition (NSUN2) and '*reading*' (ALYREF) (Fig. 2a) [4]. In contrast, pseudouridine synthase 7 (PUS7) bound

methylated VTRNA1.1 with higher affinity, whereas the serine / arginine rich (SR) splicing factor 2 (SRSF2) was consistently repelled by m^5C in VTRNA1.1 (Fig. 2a). To confirm that we pulled down VTRNA1.1-specific RNA-binding proteins, we performed RNA pull-down assays followed by Western blot, and found that SRSF2 bound non-methylated (C69) with higher affinity than methylated (m^5C69) VTRNA1.1 (Fig. 2c). In contrast, PUS7 bound methylated VTRNA1.1 with higher affinity (Supplementary Figure 2d; C69, m^5C69). The heterogeneous nuclear RNP (hnRNP) A1 protein served as a loading control (Fig. 2c) [27].

m^5C directly influences the affinity of SRSF2 to VTRNA1.1

SRSF2 is best-known for its role in splicing where it mediates exon inclusion and exclusion equally well [28]. SRSF2 was a promising candidate to regulate VTRNA1.1 processing because SR protein binding is not limited to pre-mRNA, they can also associate with non-coding RNAs, such as 7SK and MALAT1 [29-32]. SRSF2 binds to pre-mRNA via its RNA recognition motif domain [33-35], and VTRNA1.1 contained two putative SRSF2 RNA binding motifs (RRM1 and RRM2) (Fig. 2d; Supplementary Figure 2e) [28,36]. RRM1 overlapped with the methylated cytosine 69 (Fig. 2d). To validate the functional importance of the SRSF2 binding motifs, we point mutated C69 (C69A) and C88 (C88U) in VTRNA1.1 and performed RNA pull-down experiments followed by Western blot for SRSF2. Both mutations decreased the binding affinity to SRSF2 (Fig. 2d; Supplementary Figure 2f). Interestingly, binding of PUS7 to the mutated VTRNA1.1-constructs was also decreased (Supplementary Figure 2d).

To further confirm a direct RNA–protein interaction between VTRNA1.1 and SRSF2, we performed electrophoretic mobility shift assays (EMSA) using methylated and un-methylated VTRNA1.1 (Fig. 2e). Purified, recombinant GST-SRSF2 RNA recognition motif (GST-SRSF2 RRM) showed high binding affinity to un-methylated VTRNA1.1 (C69) that correlated with increased concentration of the recombinant protein. Importantly, the binding affinity of GST-

SRSF2 RRM was significantly lower when VTRNA1.1 was methylated at C69 (m^5C69) (Fig. 2e, f).

Finally, we asked whether SRSF2-binding affinity to VTRNA1.1 also decreased upon NSUN2-mediated methylation *in vivo*. First, we confirmed comparable protein expression levels of SRSF2 in NSUN2-expressing (+/+; +/-) or –lacking (-/-) cells by immunoprecipitation or Western blotting (Fig. 3a, b; Supplementary Figure 3a). Next, we co-immunoprecipitated SRSF2 and measured bound VTRNA1.1 by qRT-PCR (Fig. 3c). The amount of VTRNA1.1 bound to SRSF2 was highest in the absence of NSUN2 (Fig. 3c), confirming that SRSF2 preferentially bound un-methylated VTRNA1.1.

In conclusion, we identified SRSF2 as a novel VTRNA1.1-binding protein, whose affinity was reduced by post-transcriptional methylation of VTRNA1.1 by NSUN2.

VTRNA1.1 processing is altered in the absence of SRSF2

To test how SRSF2 modulated the processing of VTRNA1.1, we depleted SRSF2 in *NSUN2*-lacking cells for two reasons. First, the affinity of SRSF2 to VTRNA1.1 was highest in *NSUN2*-/- cells (Fig. 3c). Second svRNA4 production was lowest in *NSUN2*-/- cells (Fig. 1g). While SRSF1, a close protein family member of SRSF2, was efficiently down-regulated using siRNA and shRNAs, SRSF2 was consistently repressed by only 50% (Supplementary Figure 3b, c). However, both proteins were similarly down-regulated on protein levels (Fig. 3d). The processing of VTRNA1.1 into svRNA4 was significantly increased when SRSF2 was knocked-down (Fig. 3e). Together, these results indicated that (*i*) SRSF2 protected un-methylated VTRNA1.1 from processing and (*ii*) loss of SRSF2 was sufficient to rescue svRNA4 production in the absence of NSUN2-driven methylation. This data indicated that SRSF2-binding to VTRNA1.1 was likely to occur up-stream of VTRNA1.1 methylation. The

196 processing into svRNA1 was slightly reduced by down-regulation of SRSF1 and 2, yet this
197 change was not significant (Fig. 3e). In summary, our data demonstrated that SRSF2-binding
198 to un-methylated VTRNA1.1 down-regulated the level of processing into svRNA4.
199
200 ***VTRNA1.1, svRNA4 and SRSF2 levels are dynamically regulated***
201 We next asked whether SRSF2-regulated VTRNA1.1 processing was physiologically relevant.
202 Because removal of SRSF2 can cause cell death [37], we turned our studies to primary human
203 keratinocytes (HK), which are highly resistant to apoptosis [38]. Primary HK can be cultured as
204 undifferentiated, lineage committed progenitor cells and induced to terminally differentiate by
205 exposure to high calcium concentration in the culture medium [39]. Using the calcium switch
206 assay, we differentiated epidermal cells for two and six days (Fig. 4a).
207
208 First, we confirmed that RNA expression levels of the epidermal differentiation markers
209 keratin 10 (*Krt10*), transglutaminase I (*Tgm*), involucrin (*Inv*), and the epidermal differentiation
210 regulator Ovo Like Transcriptional Repressor 1 (*Ovol1*) were all up-regulated at both time
211 points (Fig. 4b) [40,41]. Notably, also the abundance of full length VTRNA1.1 significantly
212 increased upon differentiation (Fig. 4b). In contrast, the RNA levels of *Srsf2*, *Nsun2* and
213 svRNA4 were all repressed upon differentiation (Fig. 4c). Together, these data indicated that
214 both methylation of RNAs by NSUN2 and VTRNA1.1 processing was highest in
215 undifferentiated epidermal progenitor cells.
216
217 ***m^{5}C levels are dynamic and depend on the type of RNA***
218 To directly determine the methylation differences in undifferentiated and differentiated
219 epidermal cells, we performed RNA bisulfite (BS) sequencing [7,42]. We only quantified sites
220 with a coverage of more than 100 reads and a minimum of 20% methylation in at least one of
221 the two conditions (Supplementary Data 5). The average methylation level of the around 350

identified sites was significantly higher in the undifferentiated cells (Fig. 4d). We confirmed comparable coverage of these sites in the two conditions (Fig. 4e). Next, we compared the methylation levels of NSUN2-dependent and high confidence sites, which we defined as all rescued sites found in the *NSUN2*-/- human dermal fibroblasts (n=76) (Supplementary Data 6). The vast majority of these sites were located in tRNAs and a small number occurred in other non-coding RNAs (Fig. 4f). In contrast to all sites detected in the epidermal cells, most of these NSUN2-methylated RNAs, including VTRNA1.1, showed higher methylation levels in the differentiated condition (Fig. 4f-h). The higher level of methylation was consistent in all replicates (Fig. 4g, h; right hand panels).

One explanation why specifically these NSUN2-dependent methylation sites were more abundant in differentiated cells was that they all occurred in abundant and stable non-coding RNAs. Due to their long half-life, the methylated forms might accumulate during differentiation, provided the processing machinery '*reading'* the presence or absence of the methyl mark was less active in differentiating cells. This hypothesis was supported by our finding that full length VTRNA1.1 was more abundant in differentiated epidermal cells, while svRNA4 decreased (Fig. 4b, c). In addition, VTRNA1.1 was consistently better covered in differentiated cells by BS sequencing (Supplementary Figure 4a). Similarly, the coverage of tRNA Leu^{CAA} was reduced in undifferentiated cells and it also showed higher methylation levels at NSUN2-specific sites in the differentiated cell state (Fig. 4h; Supplementary Figure 4b). Moreover, higher coverage also negatively correlated with expression of NSUN2 in the human dermal fibroblasts (Supplementary Figure 4a, b). Thus, our data indicated that the overall methylation levels of stable non-coding RNAs depended on the presence of both the methylating enzyme and the RNA processing machinery.

svRNA4 maintains undifferentiated transcriptional programme

To provide direct evidence that both the presence of NSUN2 and VTRNA1.1 processing was required in the undifferentiated progenitor cells, we transfected a svRNA4 mimic into primary HK and differentiated them using the calcium switch assay (Fig. 4i). svRNA4-transduced keratinocytes failed to undergo the morphological changes that are normally associated with a stratified squamous epithelium (Fig. 4i) [43]. To confirm a reduced capacity to differentiate in the presence of svRNA4, we measured RNA and protein expression levels of terminal differentiation markers. Western blot for the differentiation marker KRT10 and OVOL1 revealed repression of both markers in the presence of svRNA4 (Fig. 4j). QRT-PCR further confirmed down-regulation of RNA expression levels of the terminal differentiation markers *Krt10*, *Inv*, and *Tgm* as well as *Nsun2* and *Myc*, both known to promote lineage commitment in skin (Fig. 4k) [9,44-46].

Importantly, inhibition of endogenous svRNA4 using an anti-svRNA construct was sufficient to reverse the abundance of the terminal differentiation markers of *Tgm*, *Inv*, and *Ovol1* even in low calcium growth conditions (Fig. 4l). The low abundance of full length VTRNA1.1 in both experiments confirmed that both constructs efficiently repressed their target RNA sequences (Fig. 4k, l). Together, our data indicated that the presence of svRNA4 was sufficient to maintain the transcriptional programme of a committed, yet undifferentiated progenitor state.

SRSF2 maintains cell cycling of progenitor cells

Since the presence of svRNA4 modulated human epidermal differentiation and SRSF2 influenced svRNA4 processing, we next determined the functional role of SRSF2 during epidermal cell differentiation. We transfected the epidermal cells with SRSF2 siRNAs, and then induced them to differentiate by increasing the calcium concentration in the growth medium (Fig. 5a). As a control, we also transfected a SRSF1 siRNA, and confirmed that only

SRSF2 expression levels were reduced after knock-down with the SRSF2-specific siRNA on both RNA and protein levels (Fig. 5b; Supplementary Figure 5a). As expected, the human keratinocytes underwent efficient differentiation and only *Ovol1* RNA levels were significantly increased in the absence of SRSF2 (Fig. 5c). These data indicated that the differentiation programme was largely unaffected by depletion of SRSF2.

Terminal differentiation of epidermal cells is defined by both the up-regulation of differentiation markers and exit from the cell cycle [47]. We noted that the SRSF2-depleted HK showed differences in the distribution of cell cycle phases when compared to the differentiated control keratinocytes (Supplementary Figure 5b). Therefore, we next asked whether SRSF2 influenced cell division and removed *Srsf2* in undifferentiated HK for 48 and 72 hours (Supplementary Figure 5c). Only after 72 hours of knock-down, we measured a significant up-regulation of the terminal differentiation markers *Krt10* and *Tgm* (Fig. 5d). In contrast, we found a consistent down-regulation of major cell cycle regulators as early as 48 hours after transfection (Fig. 5e) [48,49]. Cell cycle analyses revealed that repression of SRSF2 led to a significant increase of the Sub-G1 phase of the cell cycle, indicating enhanced cell death of the transfected HK (Fig. 5f). Indeed, keratinocytes failed to survive longer than 72 hours after SRSF2 depletion (Fig, 5g). Thus, SRSF2 is required for cell division and survival of primary human keratinocytes.

NSUN2 and SRSF2 act in concert to process VTRNA1.1

Finally, we asked whether the production of svRNA4 required both NSUN2 and SRSF2 in the primary human keratinocytes. Since depletion of SRSF2 caused substantial cell death after 72 hours, we removed NSUN2 and SRSF2 for only 24 hours to obtain the maximum number of viable cells (Fig. 5g). As described for the human dermal fibroblasts (hDF) (Fig. 1g), knock-down of NSUN2 in keratinocytes reduced the processing of VTRNA1.1 into svRNA4.

However, the svRNA4 level was restored to normal, when we simultaneously repressed NSUN2 and SRSF2 for 24 hours (Fig. 5h). The data was in line with our observation in human dermal fibroblasts, where removal of SRSF2 rescued the levels of svRNA4 in *NSUN2*-/- patient-derived cells (Fig. 3e). and confirmed that binding of SRSF2 to un-methylated VTRNA1.1 reduced its processing into svRNA4. We concluded that VTRNA1.1 processing required both NSUN2 and SRSF2 to maintain an undifferentiated cell state.

In summary, our study revealed that SRSF2 bound un-methylated VTRNA1.1 with higher affinity and thereby protected it from both methylation and cleavage. High levels of svRNA4 in the presence of both proteins allowed cell cycle progression but reduced expression of the differentiation-promoting transcription factor OVOL1 and other terminal differentiation markers (Fig. 5i). Thus, the presence and expression levels of both NSUN2 and SRSF2 coordinated how much methylated VTRNA1.1 was processed into svRNA4 (Fig. 6a-c).

Discussion

The proper formation of 5-methylcytosine in tRNA is required for normal development [7]. Aberrant deposition of m^5C into tRNAs causes neuro-developmental deficits by impairing the translation machinery [2,8]. While the functional role of m^5C in tRNA is now increasingly understood, the importance of m^5C in other non-coding RNAs remains unclear. NSUN2 is one of the best-characterized cytosine-5 RNA methylases and methylates the vast majority of tRNAs. Here, we confirm that NSUN2 also methylates the vault RNA VTRNA1.1 in a wide range of human cells. NSUN2 is the sole enzyme to methylate cytosine 69 in VTRNA1.1 and thereby regulates its processing into multiple regulatory small RNAs (svRNAs).

RNA modifications control the fate and function of RNA molecules by recruiting or repelling specific RNA binding proteins [50]. To determine the functional relevance of m^5C69 on VTRNA1.1 metabolism, we identified all RNA-binding proteins that bind methylated and un-methylated VTRNA1.1 with different affinity. Our quantitative mass spectrometry-based approach identified SRSF2 as a novel VTRNA1.1 binding protein that was repelled by the presence of m^5C69. During gene transcription, SRSF2 contributes to constitutive and alternative splicing by binding to exonic splicing enhancer sequences (ESE), predominantly within intron-containing pre-mRNA [33,34]. In addition, SRSF2 binds mRNAs via non-ESE sites, for instance the HIV-1 tat mRNA[51]. We identified two SRSF2 binding sites in VTRNA1.1, one of which overlapped with the methylated C69. SRSF2 has been described to bind other non-coding RNAs such as 7SK and MALAT1[30-32]. The binding of MALAT1 or 7SK is thought to influence the recruitment of SRSF2 to distinct active transcriptional regions [30,32]. Nuclear un-methylated VTRNA1.1 might have a similar function.

In addition to their roles in splicing, SR proteins are thought to have a broader role in RNA metabolism. SRSF1 for instance facilitates the processing of certain microRNAs independent of its function in splicing [52]. Furthermore, CLIP assays revealed little direct correlation between SRSF2 protein binding and induced splicing changes [31]. Here, we propose that SRSF2 contributes to VTRNA1.1 processing into svRNA4 binding to un-methylated VTRNA1.1 and thereby protecting it from methylation by NSUN2 (Fig. 6a). For instance, epidermal progenitors express high levels of both NSUN2 and SRSF2 and consequently, methylate a substantial fraction of VTRNA1.1. Methylated VTRNA1.1 binds SRSF2 with lower affinity and the levels of svRNA4 increase. Deletion of NSUN2 leads to loss of m^5C at C69 and enhances the binding of VTRNA1.1 to SRSF2 (Fig. 6b). Binding to SRSF2 reduced the level of VTRNA1.1 entering the RNA processing machinery leading to reduced production of svRNA4. Finally, we tested a scenario where both proteins are repressed (Fig. 6c). This

enhanced the fraction of free VTRNA1.1 that entered the RNA processing machinery leading to increased levels of svRNA4. Together, VTRNA1.1 methylation at C69 facilitates the production of svRNA4 by reducing SRSF2 affinity to VTRNA1.1.

Deciphering the precise function of SRSF2 in modulating the processing of VTRNA1.1 was hampered by the fact that SRSF2 is an essential protein upstream of both RNA methylation and processing. Homozygous germ line deletion of SRSF2 is embryonic lethal [53], and conditional knock-out mice display tissue-specific phenotypes [53-56]. Similar to repression of SRSF2 in epidermal progenitors, loss of SRSF2 in mouse embryonic fibroblasts induced G2/M cell cycle arrest and genomic instability [54]. In addition, SRSF2 is required to maintain human embryonic stem cell pluripotency [57]. Mutations in the human *SRSF2* gene altering its RNA-binding affinity impairs hematopoietic differentiation *in vivo* and is frequently (40% incidence) found in patients with myelodysplastic syndromes and certain leukemias [28,58,59].

The function of SRSF2 in splicing might be modulated by VTRNA1.1, yet VTRNA1.1-methylation and processing are probably not linked to splicing. Repression of SRSF2 in primary human keratinocytes caused cell cycle arrest and cell death, and this effect was independent of VTRNA1.1-methylation and processing. However, simultaneous deletion of SRSF2 and NSUN2 increased svRNA4 levels confirming that SRSF2 is involved in regulating VTRNA1.1 processing in primary human keratinocytes. Thus, NSUN2 and SRSF2 together determined how much VTRNA1.1 was processed into svRNAs. Finally, we confirmed a functional relevance of svRNA4 in regulating epidermal differentiation. One explanation for how svRNA4 modulated epidermal differentiation is through post-transcriptional silencing of mRNA targets. Our previous bioinformatic analysis identified OVOL1 as a high confidence target of svRNA4 [3]. While we indeed found an inverse relationship between the abundances of OVOL1 and svRNA4, more work is needed to confirm that svRNA4 indeed acts similarly to

377 miRNAs. However, we identified a functional relevance for svRNA4 in regulating the terminal
378 differentiation program. While enhanced levels of svRNA4 repressed the terminal
379 differentiation program, sequestering svRNA4 was sufficient to trigger differentiation.
380
381
382

Methods

Cell culture, transfections and infections

HEK293 and Hela cells (ATCC) were grown in Dulbecco's Minimal Essential Medium (DMEM) (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) (Sigma), Penicillin, and Streptomycin. Human dermal fibroblasts [11] were grown in Minimal Essential Medium (MEM) (Thermo Fisher Scientific) supplemented with 20% FBS (Sigma), 1% Penicillin, and Streptomycin. The human embryonic stem cell line Hues9 (H9) was obtained from the WiCell. H9 cells were maintained in Essential 8 media (Thermo Fisher Scientific) on human embryonic stem cell-qualified matrigel (Corning) coated plates at 37°C, 5% CO2. Primary human keratinocytes isolated from neonatal foreskins (Cellworks distributed, ZHC-1116) were cultured on collagen (BD Biosciences) coated plates in KGM-gold (LONZA) medium. For cell propagation calcium concentrations were adjusted to 0.06 mM using 1.2 M $CaCl_2$ stock solution. Medium was changed every other day and cells were passaged when reaching 60-70% confluency. For epidermal differentiation, KGM-gold calcium concentrations were adjusted to 1.2 mM using 1.2 M $CaCl_2$ stock solution. Primary human epidermal cells were only used until passage 6-7 after thawing. All cells were cultured in humidified atmospheres with 5% CO2.

For the transfection of small interfering RNA (siRNA), human dermal fibroblasts and primary human keratinocytes (HK) cells were cultured until reaching 50-60 % confluency. Cells were then transfected with a control siRNAs or svRNA4 antagomirs (2′-O-methyl-anti-svRNA4: 5′ AAA AGG ACU GGA GAG CGC CCG CGG GUC UCG); control microRNA miRIDIAN mimic or svRNA4 microRNA miRIDIAN mimic (5′ CGA GAC CCG CGG GCG CUC UCC AGU CCU UUU) (Dharmacon-GE); SRSF1 siRNAs (QIAGEN) or SFSF2 siRNAs (QIAGEN, and Dharmacon-GE) using RNAimax transfection kit (Thermo Fisher Scientific). 24, 48, or 72 hours post-transfection, cells were washed in PBS and RNA or protein was isolated.

For the infection of the short hairpin (sh) RNA, the most effect siRNAs were clones into PLKO.1 puro plasmids (Addgene). To produce the lentivirus, HEK293T cells were grown in DMEM supplemented with 10% FBS on 60 cm cell dishes until 40-50% confluent and then transfected with 15 μg lentiviral PLKO.1 puro vectors containing shRNA for SRSF1 (shRNA1.1, shRNA1.6) or SRSF2 (shRNA2.4, shRNA2.5) or the empty PLKO.1 puro vector and 7.5 μg of both packaging vectors (Pol and Gag) using the calcium-phosphate transfection kit (Thermo Fisher Scientific). The next day the medium was carefully changed to avoid detaching HeK293T cells and kept in culture for another 24 hours. Subsequently, supernatant containing viral particles were collected and filtered (0.22 μm) before used to infect human dermal fibroblast. The next day, the medium was changed, and cells were kept in culture for 24 hours before selecting with 1 μg/ml puromycin for 4 days with medium changed after day 2. After selection, the fibroblasts were propagated without purpmycin for further experiments.

To rescue expression of NSUN2 protein or its enzymatic dead version, full length human NSUN2 (pB-NSUN2), inactive mutants C271A (pB-NSUN2-C271A), K190M (pB-NSUN2-K190M) or C321A (pB-NSUN2-C321A), and the empty vector (pB-empty) were infected retrovirally [26].

RNA bisulfite sequencing

All RNA bisulfite conversion experiments were performed in five independent replicates. Total RNA of about 4 μg was extracted using Trizol (Thermo Fisher Scientific) and then DNase (Ambion) and Ribo-Zero (Illumina) treated according to the manufacturers' instruction. The remaining RNA fraction was bisulfite-converted. Briefly, ribosomal depleted RNA was mixed with 70 μl of 40% sodium bisulfite pH 5.0 and DNA protection buffer (EpiTect Bisulfite Kit, Qiagen). The reaction mixture was incubated for three cycles of 5 minutes at 70°C followed

by 1 hour at 60°C and then desalted with Micro Bio-spin 6 chromatography columns (Bio-Rad). RNA was desulphonated by adding an equal volume of 1 M Tris (pH 9.0) to the reaction mixture and incubated for 1 hour at 37°C, followed by ethanol precipitation. Bisulfite converted RNA was then treated with T4 PNK (New England Biolabs) to repair both 5' and 3' ends for library preparation. Repaired RNA quality and concentration was measured on a Bioanalyzer 2100 RNA nano-chip (Agilent). About 100 ng of RNA was used to generate the libraries using the TruSeq Small RNA preparation kit (Illumina). RNA adapters were then ligated, reverse-transcribed and amplified by 18 cycles of PCR before sequencing on a HiSeq4000 (Illumina).

Small qRT-PCR and qRT-PCR

To measure the abundance of svRNA1 and svRNA4 we performed small quantitative PCR (qRT-PCR). Total RNA from human epidermal cells, *NSUN2+/-*, *NSUN2-/-* dermal fibroblasts was isolated using TRIzol according to the manufacturer's instruction (Thermo Fisher Scientific). RNAs of smaller than 200 base pairs were enriched using mirVana kit (Ambion) according manufacturer's instructions. Enriched RNAs were then separated using 15% acrylamide-urea gels (Thermo Fisher Scientific) and gel slices spanning 10-40 bps were excised and crushed in RNase free water and incubated overnight at 4°C. Eluted 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 qRT-PCR was performed as follows [60]. 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 qRT-PCR reactions. cDNA was incubated at 90°C for 10 minutes followed by 40 cycles of 30 seconds denaturation at 90°C, and one minute of extension at 60°C. The forward primers were: miR-16 (5′-TAG CAG CAC GTA AAT ATT GGC G-'3), svRNA1 (5′-TGT CTG GGT TGT

TCG AGA CCC GCG GGC-3′), and 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′.

For conventional quantitative PCR (qRT-PCR) total RNA was isolated from human fibroblasts, and primary human keratinocytes cells using TRIzol (Thermo Fisher Scientific). cDNA synthesis was performed using the SuperscriptIII reverse transcriptase kit (Thermo Fisher Scientific) according to manufacturer’s instructions. qRT-PCRwas performed using TaqMan assay sets purchased from Thermo Fisher Scientific and used as per manufacturer’s recommendations or pre-designed primers (Sigma-Aldrich) and Sybr master mix 2x (Life Technologies) were used. The following probes were used to amplify selected genes: *Gapdh* (4326317E), *Nsun2* (Hs00214829_m1), *Srsf2* (Hs00958207_cn), *Srsf1* (Hs00199471_m1), *Ovol1* (Hs00190060_m1), *vt-RNA1.1* (Hs03676993_s1), *Inv* (Hs00846307_s1), *Tgm1* (Hs00165929_m1 and Hs01070316_m1), *Krt10* (Hs01051614_g1 and Hs00166289_m1), Myc (Hs00153408_m1), *Itgα6* (Hs01041013_m1), *Cdc45* (Hs00907337_m1), *Mcm10* (Hs00960349_m1), *Cdc25* (Hs00156411_m1), and *Cdk1* (Hs00938777_m1). The following pre-designed primers were used: *Gapdh* (forward: ATC TTC CAG GAG CGA GAT CC, reverse: ACC ACT GAC ACG TTG GCA GT), *Srsf2* (forward: CCT AAT TTG TGG CCT CCT GA, reverse: TCA ATC TCT TGA CAG CT TAG GC).

SILAC

HeLa cells were grown in SILAC DMEM (Thermo Fisher Scientific) supplemented with either “light” L-lysine-2HCl and L-arginine-HCl or “heavy” 13C6-L-lysine-2HCl and 13C615N4-L-arginine-HCl. For full incorporation of “light” or “heavy” L-Lys/L-Arg, for at least six passages. Subsequently, 3×10^6 cells were re-suspended in 1 ml of buffer-D (100 mM Tris-HCl at pH 8.0; 100 mM KCl; 0.2 mM EDTA; 0.5 mM DTT; 0.2 mM PMSF; 20% (w/v)

glycerol), scraped and sonicated (Diagenode). The suspension was centrifuged for 5 minutes at 10.000 g, and the supernatant was used for VTRNA pull-down assays followed by mass spectrometry detection [61]

RNA pull-down and SILAC mass-spectrometry (RP-SMS)

We performed agarose mediated VT-RNA1.1 pull-downs [62],[27]. Briefly, 10 µg of methylated, or un-methylated VT-RNA1.1 (Dharmacon-GE) was treated with sodium m-periodate (Sigma) for 1 hour at room temperature while rotating. VT-RNAs were then precipitated by 3M of sodium acetate (Ambion) and ethanol. Recovered VT-RNAs were then covalently coupled with adipic acid dihydrazide agarose beads (Sigma) overnight at 4°C. VT-RNA-bead complex was incubated with 40% v/v of total SILAC "heavy" HeLa lysates supplemented with 1.5 mM $MgCl_2$, 25 mM creatine-phosphate (Millipore), 200 units per ml RNAseOUT (ThermoFisher Scientific) and 5 mM ATP (Sigma) for 30 minutes at 37°C. The beads only control was incubated with SILAC "light" HeLa lysates. The reactions were washed four times with Buffer-G (20 mM Tris-HCl at PH7.5; 137 mM NaCl; 1mM EDTA; 1% Triton x100; 10% glycerol; 1.5 mM $MgCl_2$) or Buffer-A (20 mM Tris-HCl at pH: 7.5; 50 mM NaCl; 1mM EDTA; 1% NP-40; 10% glycerol; 1.5 mM $MgCl_2$), for less stringent RNA-protein interactions. After the final wash, RNA-bead complex was mixed with beads only control and incubated in 60-µl H_2O containing 4x LDS protein sample buffer (Novex), 10X reducing reagent buffer (Novex), then incubated at 70°C for 10 minutes. The end solution was centrifuged at maximum speed for 1 minute, and the supernatant was collected and analysed either by quantitative mass-spectrometer or SDS-PAGE western blot for proteins detection [27].

RNA immunoprecipitation (RIP):

Pelleted human dermal fibroblasts were lysed in RIP buffer (1% NP-40; 25 mM Tris-HCl at pH 8.0; 150 mM NaCl; 2 mM MgCl2; 1 mM DTT) and incubated on ice for 30 minutes. SRSF2

antibody (5µ ab11826; Abcam), or control pre-bled serum were pre-incubated with Protein G magnetic dynabeads (Thermo Fisher Scientific) in lysis buffer for 30 minutes. Following washing of beads, cleared lysates were added to the beads and immunoprecipitation was carried out for 2 hours at 4°C with gentle mixing. Beads were then washed extensively in lysis buffer and precipitated RNAs were recovered from the beads using TRIzol (Thermo Fisher Scientific). Immunoprecipitation reactions were performed in the presence of 200 units per ml of RNAse inhibitor.

Co-immunoprecipitations and Western blot

Confluent 150 cm dishes of cells were washed twice with ice-cold PBS before scraped using 300 µl of ice-cold CHAPS (FIVE photon Biochemicals) lysis buffer supplemented with protease inhibitors (Roche) and phosphatase inhibitors (Roche). Lysates were transferred to pre-chilled Eppendorf tubes and incubated on ice for 10 minutes. Tubes were strongly taped several times during the incubation period to facilitate cell membrane lysis. Lysates were centrifuged for 15 minutes at maximum speed under cool conditions. The supernatant was stored until used for co-IP analysis. About 3 µg of primary antibodies or control pre-bled serum were washed 3 times with PBS before incubated with CHAPS prepared HeLa lysates at 4°C for 1 hour. During incubation period 50 µl of Agarose Anti-Rabbit IgG IP beads (ROCKLAND) was washed extensively with pre-chilled lysis buffer (50mM Tris-HCl pH: 8; 150 mM NaCl; 1% NP-40; protease inhibitor; phosphatase inhibitor). After the incubation, Agarose Anti-Rabbit IgG IP beads were added to the mixture and left rotating overnight at 4°C. Next morning the samples were centrifuged at 1000xg for 1 minute and the supernatant was removed. The pelleted bead-antibody complex was washed 4 time with lysis buffer. The bead-antibody complex was then mixed with 4X LDS sample buffer (Novex), and 10X reducing reagent buffer (Novex) and incubated at 70°C for 10 minutes for western blotting detection.

For western blotting, cells were harvested and washed twice with PBS before being lysed with lysis buffer (50 mM Tris-HCL at pH 7.4; 250 mM NaCl; 1% NP-40; 0.1% SDS; 0.5% sodium deoxycholate) and were cleared by centrifugation at full speed for 10 minutes at 4°C. Supernatant was mixed with 4x LDS protein sample buffer (Novex), 10X reducing reagent buffer (Novex) and incubated at 70°C for 10 minutes. Samples were run on running buffer (Novex) containing 40x NuPAGE Antioxidant (Novex) then electrophoresed on a 4-12% Bis-Tris precast polyacrylamide gels (Novex). Proteins were transferred onto nitrocellulose membranes (GE Healthcare) using transfer buffer (191mM Glycine; 25mM Tris-hydrochloride; 10% Methanol; 0.1% SDS). Nitrocellulose membranes were blocked in 10% Western Blocking Reagent (Roche). Blots were then incubated with primary antibodies in blocking solution overnight at 4°C then followed by incubation with the appropriate HRP-conjugated secondary antibodies. The chemiluminescent signal was detected using an ECL chemiluminescent kit (GE Healthcare) according to instructions. The following antibodies were used: NSUN2 (1:1000; Met-A) [63], hnRNP-A1 (1:2000; D21H11; Cell Signalling Technology), SRSF2 (1:1000; ab11826; Abcam), SRSF1 (1:2000; 32-4500; Thermo Fisher Scientific), OVOL1 (1:1000; ab65023; Abcam), KRT10 (1:2000; PRB-159P; Covance), Tubulin (1:5000; clone DM1A; Sigma), PUS7 (1:2000; ab118039, Abcam).

EMSA

EMSA was performed with end-labelled vtRNA1.1 and indicated amounts of recombinant GST-SRSF2 RRM. Probes (50×10^3 counts per minute, ~0.1 pmol) were incubated in 16 µl reactions with recombinant protein in elution buffer (25mM Tris-HCl, 5 mM reduced glutathione, pH 8.0) supplemented with 3 mM $MgCl_2$, 0.5mM ATP, 37.5 mM creatine phosphate and 10 ng tRNA for one hour on ice. Reactions were mixed with native loading buffer and analysed on a 6% (w/v) non-denaturing polyacrylamide gel run in 0.5xTBE at 8W for 1 hour and 10 minutes. The signal was registered using radiographic X-ray film.

Cell cycle analyses

Flow cytometry was used to probe cell cycle stage. Flow cytometry analysis was performed with the LSRFortessa Flow Cytometer (BD Biosciences). Cells were washed in PBS and collected with Trypsin-EDTA (1:1 in PBS). Cells were then fixed by resuspending in ice-cold 70% Ethanol. Samples were kept at 4°C until processing. Before processing cells were centrifuged at 12,000 g for 5 minutes and resuspended in 3 mL PBS with DAPI (1:3000). Fluorescence of each samples was measured on the flow cytometer at 450/50 405nm. Data was analysed using FCS express6 (DeNovo Software). All samples were gated using forward versus side scatter to eliminate debris.

Bioinformatic analyses

To assess the transcriptome-wide methylation levels and to create the methylation level heatmaps, Trim Galore! (v0.4.0) with parameters "--stringency 3 -e 0.2 -a TGGAATTCTCGGGTGCCAAGGA" was used to remove sequencing adapters and reads shorter than 20 nucleotides. Alignment to the hg38 reference genome was done with Bismark (v0.14.4) with parameters "-n 2 -l 50 --un --ambiguous --bowtie1 --chunkmbs 2048", to allow for up to two mismatches and to save unaligned and ambiguously mapping reads separately. Seqtk with parameters "-e 3" was used to remove the last three bases (a potential 'CCA' tail) from the unaligned and ambiguous reads followed by a second alignment attempt using Bismark. Finally, ngsutils (v0.5.9) in the "junction" mode was used to extract splice junctions from known genes (Gencode v28) and unaligned reads from the second attempt were aligned to the junctions using Bismark. The aligned reads were converted back to genomic coordinates using bamutils in "convertregion" mode. 'N' in the cigar string was replaced with 'D' for compatibility with bismark_methylation_extractor.

To identify novel high-confidence methylation sites we selected all sites with read coverage above 10 in all replicates. Welch's t-test was used to compare the methylation levels of either the empty vector (e. Vector) or the enzymatic dead version (K190M) to the Nsun2 rescue (NSUN2); five replicates per group. False discovery rate was used to correct for multiple testing. Methylation sites with padj<0.05 were considered to be differently methylated.

Samtools merge was used to combine aligned reads from all three steps. Reads with >1/3 methylated cytosines were discarded as they are likely artifacts from highly structured regions. The bismark_methylation_extractor with the "--bedGraph --counts --CX_context" options was used to extract methylated cytosines.

Sequencing data from patient cells are deposited on dbGaP under the accession number phs000645.v4.p1. The sequencing data are available on GEO with the accession numbers GSE122600 and GSE125046.

Gene set enrichment analyses were done using GOrilla [http://cbl-gorilla.cs.technion.ac.il/] [64]. As running mode, we used '*two ranked lists of genes*' using all identified binding proteins as background list.

Code availability

The scripts used for the alignment and processing of the bisulfite sequencing data are available at [https://github.com/susbo/trans-bsseq].

Data availability

Patient-related sequencing data are available on dbGaP: phs000645.v4.p1 [https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000645.v4.p1].

All other sequencing data are available on GEO: GSE122600 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122600] and GSE125046 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125046].

The source data underlying Figures 1c, g, h; 2c-f; 3a-e; 4b, c, j-l; 5b-f, h and Supplementary Figures 1a; 2d, f; 3a-c; 5a and c are provided as a Source Data file.

Acknowledgments

We thank everybody who provided us with reagents, in particular we thank James Stevenin for sending us recombinant SRSF2. We gratefully acknowledge the support of all the WT-MRC Stem Cell Institute core facility managers. This work was funded by Cancer Research UK (CR-UK) and the European Research Council (ERC). Parts of this research in Michaela Frye's laboratory was supported by core funding from Wellcome and MRC to the Wellcome-MRC Cambridge Stem Cell Institute. Juri Rappsilber's laboratory was supported by Wellcome Trust Senior Research Fellowship (084229). Gracjan Michlewski's laboratory was supported by the MRC Career Development Award (G10000564), Wellcome Trust Seed Award (210144/Z/18/Z) and Wellcome Trust Centre for Cell Biology Core Grants (077707 and 092076). Abdulrahim Sajini was supported by a scholarship from the University of Tabuk and Khalifa University of Science and Technology Faculty start-up award number FSU-2018-01. Rebecca Wagner was supported by the Wellcome Trust PhD Programme in Stem Cell Biology & Medicine.

Author contributions

MF and GM designed and analysed data and wrote manuscript. AAS, NRC, REW, TS performed experiments and analysed data. SB and SD performed computational analyses. CS and JR performed and supervised analyses.

Competing interests

MF consults for STORM Therapeutics. The remaining authors declare no competing interests.

References:

1 Frye, M., Harada, B. T., Behm, M. & He, C. RNA modifications modulate gene expression during development. *Science* **361**, 1346-1349, doi:10.1126/science.aau1646 (2018).

2 Frye, M. & Blanco, S. Post-transcriptional modifications in development and stem cells. *Development* **143**, 3871-3881, doi:10.1242/dev.136556 (2016).

3 Hussain, S. *et al.* NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell reports* **4**, 255-261, doi:10.1016/j.celrep.2013.06.029 (2013).

4 Yang, X. *et al.* 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. *Cell Res* **27**, 606-625, doi:10.1038/cr.2017.55 (2017).

5 Khoddami, V. & Cairns, B. R. Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nat Biotechnol* **31**, 458-464, doi:10.1038/nbt.2566 (2013).

6 Squires, J. E. *et al.* Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res* **40**, 5023-5033, doi:10.1093/nar/gks144 (2012).

7 Blanco, S. *et al.* Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. *The EMBO journal* **33**, 2020-2039, doi:10.15252/embj.201489282 (2014).

8 Blanco, S. *et al.* Stem cell function and stress response are controlled by protein synthesis. *Nature* **534**, 335-340, doi:10.1038/nature18282 (2016).

9 Blanco, S. *et al.* The RNA-methyltransferase Misu (NSun2) poises epidermal stem cells to differentiate. *PLoS Genet* **7**, e1002403, doi:10.1371/journal.pgen.1002403
PGENETICS-D-11-00619 [pii] (2011).

10 Flores, J. V. *et al.* Cytosine-5 RNA Methylation Regulates Neural Stem Cell Differentiation and Motility. *Stem Cell Reports* **8**, 112-124, doi:10.1016/j.stemcr.2016.11.014 (2017).

11 Martinez, F. J. *et al.* Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. *J Med Genet* **49**, 380-385, doi:10.1136/jmedgenet-2011-100686 (2012).

12 Khan, M. A. *et al.* Mutation in NSUN2, which Encodes an RNA Methyltransferase, Causes Autosomal-Recessive Intellectual Disability. *Am J Hum Genet* **90**, 856-863, doi:S0002-9297(12)00203-0 [pii]
10.1016/j.ajhg.2012.03.023 (2012).

13 Fahiminiya, S. *et al.* Whole exome sequencing unravels disease-causing genes in consanguineous families in Qatar. *Clinical genetics*, doi:10.1111/cge.12280 (2013).

14 Abbasi-Moheb, L. *et al.* Mutations in NSUN2 cause autosomal-recessive intellectual disability. *Am J Hum Genet* **90**, 847-855, doi:S0002-9297(12)00199-1 [pii]
10.1016/j.ajhg.2012.03.021 (2012).

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

16 Stadler, P. F. *et al.* Evolution of vault RNAs. *Mol Biol Evol* **26**, 1975-1991, doi:msp112 [pii]
10.1093/molbev/msp112 (2009).

17 Kickhoefer, V. A. *et al.* Vaults are up-regulated in multidrug-resistant cancer cell lines. *J Biol Chem* **273**, 8971-8974 (1998).

18 Abbondanza, C. *et al.* Interaction of vault particles with estrogen receptor in the MCF-7 breast cancer cell. *J Cell Biol* **141**, 1301-1310 (1998).

19 Amort, M. *et al.* Expression of the vault RNA protects cells from undergoing apoptosis. *Nat Commun* **6**, 7030, doi:10.1038/ncomms8030 (2015).

20 Scheffer, G. L. *et al.* The drug resistance-related protein LRP is the human major vault protein. *Nat Med* **1**, 578-582 (1995).

21 Persson, H. *et al.* The non-coding RNA of the multidrug resistance-linked vault particle encodes multiple regulatory small RNAs. *Nat Cell Biol* **11**, 1268-1271, doi:ncb1972 [pii]
10.1038/ncb1972 (2009).

22 Ma, Y., Ouyang, J., Wei, J., Maarouf, M. & Chen, J. L. Involvement of Host Non-Coding RNAs in the Pathogenesis of the Influenza Virus. *Int J Mol Sci* **18**, doi:10.3390/ijms18010039 (2016).

23 Li, F. *et al.* Robust expression of vault RNAs induced by influenza A virus plays a critical role in suppression of PKR-mediated innate immunity. *Nucleic Acids Res* **43**, 10321-10337, doi:10.1093/nar/gkv1078 (2015).

24 van Balkom, B. W., Eisele, A. S., Pegtel, D. M., Bervoets, S. & Verhaar, M. C. Quantitative and qualitative analysis of small RNAs in human endothelial cells and exosomes provides insights into localized RNA processing, degradation and sorting. *J Extracell Vesicles* **4**, 26760, doi:10.3402/jev.v4.26760 (2015).

25 Nolte-'t Hoen, E. N. *et al.* Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res* **40**, 9272-9285, doi:10.1093/nar/gks658 (2012).

26 Hussain, S. *et al.* The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability. *J Cell Biol* **186**, 27-40 (2009).

27 Choudhury, N. R. & Michlewski, G. Quantitative identification of proteins that influence miRNA biogenesis by RNA pull-down-SILAC mass spectrometry (RP-SMS). *Methods*, doi:10.1016/j.ymeth.2018.06.006 (2018).

28 Zhang, J. *et al.* Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. *Proc Natl Acad Sci U S A* **112**, E4726-4734, doi:10.1073/pnas.1514105112 (2015).

29 Anko, M. L. Regulation of gene expression programmes by serine-arginine rich splicing factors. *Semin Cell Dev Biol* **32**, 11-21, doi:10.1016/j.semcdb.2014.03.011 (2014).
30 Ji, X. *et al.* SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* **153**, 855-868, doi:10.1016/j.cell.2013.04.028 (2013).
31 Pandit, S. *et al.* Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol Cell* **50**, 223-235, doi:10.1016/j.molcel.2013.03.001 (2013).
32 Tripathi, V. *et al.* The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* **39**, 925-938, doi:10.1016/j.molcel.2010.08.011 (2010).
33 Graveley, B. R. & Maniatis, T. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Mol Cell* **1**, 765-771 (1998).
34 Liu, H. X., Chew, S. L., Cartegni, L., Zhang, M. Q. & Krainer, A. R. Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. *Mol Cell Biol* **20**, 1063-1071 (2000).
35 Long, J. C. & Caceres, J. F. The SR protein family of splicing factors: master regulators of gene expression. *Biochem J* **417**, 15-27, doi:10.1042/BJ20081501 (2009).
36 Daubner, G. M., Clery, A., Jayne, S., Stevenin, J. & Allain, F. H. A syn-anti conformational difference allows SRSF2 to recognize guanines and cytosines equally well. *The EMBO journal* **31**, 162-174, doi:10.1038/emboj.2011.367 (2012).
37 Cheng, Y. *et al.* Liver-Specific Deletion of SRSF2 Caused Acute Liver Failure and Early Death in Mice. *Mol Cell Biol* **36**, 1628-1638, doi:10.1128/MCB.01071-15 (2016).
38 Russell, D., Ross, H. & Lane, E. B. ERK involvement in resistance to apoptosis in keratinocytes with mutant keratin. *J Invest Dermatol* **130**, 671-681, doi:10.1038/jid.2009.327 (2010).
39 Sharpe, G. R., Gillespie, J. I. & Greenwell, J. R. An increase in intracellular free calcium is an early event during differentiation of cultured human keratinocytes. *FEBS Lett* **254**, 25-28 (1989).
40 Watt, F. M. Terminal differentiation of epidermal keratinocytes. *Curr Opin Cell Biol* **1**, 1107-1115 (1989).
41 Nair, M. *et al.* Ovol1 regulates the growth arrest of embryonic epidermal progenitor cells and represses c-myc transcription. *J Cell Biol* **173**, 253-264, doi:jcb.200508196 [pii]
10.1083/jcb.200508196 (2006).
42 Schaefer, M., Hagemann, S., Hanna, K. & Lyko, F. Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines. *Cancer Res* **69**, 8127-8132, doi:10.1158/0008-5472.CAN-09-0458 (2009).
43 Bikle, D. D., Xie, Z. & Tu, C. L. Calcium regulation of keratinocyte differentiation. *Expert Rev Endocrinol Metab* **7**, 461-472, doi:10.1586/eem.12.34 (2012).
44 Frye, M., Gardner, C., Li, E. R., Arnold, I. & Watt, F. M. Evidence that Myc activation depletes the epidermal stem cell compartment by modulating adhesive interactions with the local microenvironment. *Development* **130**, 2793-2808 (2003).
45 Arnold, I. & Watt, F. M. c-Myc activation in transgenic mouse epidermis results in mobilization of stem cells and differentiation of their progeny. *Curr Biol* **11**, 558-568 (2001).
46 Gandarillas, A. & Watt, F. M. c-Myc promotes differentiation of human epidermal stem cells. *Genes Dev* **11**, 2869-2882 (1997).
47 Kolly, C., Suter, M. M. & Muller, E. J. Proliferation, cell cycle exit, and onset of terminal differentiation in cultured keratinocytes: pre-programmed pathways in control of C-Myc and Notch1 prevail over extracellular calcium signals. *J Invest Dermatol* **124**, 1014-1025, doi:10.1111/j.0022-202X.2005.23655.x (2005).
48 Enserink, J. M. & Kolodner, R. D. An overview of Cdk1-controlled targets and processes. *Cell Div* **5**, 11, doi:10.1186/1747-1028-5-11 (2010).
49 Homesley, L. *et al.* Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins. *Genes Dev* **14**, 913-926 (2000).
50 Roundtree, I. A. & He, C. RNA epigenetics--chemical messages for posttranscriptional gene regulation. *Curr Opin Chem Biol* **30**, 46-51, doi:10.1016/j.cbpa.2015.10.024 (2016).
51 Erkelenz, S. *et al.* Balanced splicing at the Tat-specific HIV-1 3'ss A3 is critical for HIV-1 replication. *Retrovirology* **12**, 29, doi:10.1186/s12977-015-0154-8 (2015).
52 Wu, H. *et al.* A splicing-independent function of SF2/ASF in microRNA processing. *Mol Cell* **38**, 67-77, doi:10.1016/j.molcel.2010.02.021 (2010).
53 Ding, J. H. *et al.* Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart. *The EMBO journal* **23**, 885-896, doi:10.1038/sj.emboj.7600054 (2004).
54 Xiao, R. *et al.* Splicing regulator SC35 is essential for genomic stability and cell proliferation during mammalian organogenesis. *Mol Cell Biol* **27**, 5393-5402, doi:10.1128/MCB.00288-07 (2007).
55 Wang, H. Y., Xu, X., Ding, J. H., Bermingham, J. R., Jr. & Fu, X. D. SC35 plays a role in T cell development and alternative splicing of CD45. *Mol Cell* **7**, 331-342 (2001).
56 Komeno, Y. *et al.* SRSF2 Is Essential for Hematopoiesis, and Its Myelodysplastic Syndrome-Related Mutations Dysregulate Alternative Pre-mRNA Splicing. *Mol Cell Biol* **35**, 3071-3082, doi:10.1128/MCB.00202-15 (2015).
57 Lu, Y. *et al.* Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. *Cell Stem Cell* **15**, 92-101, doi:10.1016/j.stem.2014.04.002 (2014).
58 Kim, E. *et al.* SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer Cell* **27**, 617-630, doi:10.1016/j.ccell.2015.04.006 (2015).

59 Ball, M., List, A. F. & Padron, E. When clinical heterogeneity exceeds genetic heterogeneity: thinking outside the genomic box in chronic myelomonocytic leukemia. *Blood*, doi:10.1182/blood-2016-07-692988 (2016).
60 Shi, R. & Chiang, V. L. Facile means for quantifying microRNA expression by real-time PCR. *BioTechniques* **39**, 519-525 (2005).
61 Choudhury, N. R. *et al.* Trim25 Is an RNA-Specific Activator of Lin28a/TuT4-Mediated Uridylation. *Cell reports* **9**, 1265-1272, doi:10.1016/j.celrep.2014.10.017 (2014).
62 Michlewski, G. & Caceres, J. F. RNase-assisted RNA chromatography. *RNA* **16**, 1673-1678, doi:10.1261/rna.2136010 (2010).
63 Frye, M. & Watt, F. M. The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Curr Biol* **16**, 971-981 (2006).
64 Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48, doi:10.1186/1471-2105-10-48 (2009).

803 ***Figure legends***

804 ***Figure 1***. **Methylation of VTRNA1.1 by NSUN2 determines the biogenesis of svRNA4.** (**a-**
805 **c**) Correlation of site-specific methylation (m^5C) levels (a, b) and methylation level at all
806 covered cytosines in VTRNA1.1. (c) in *NSUN2*-/- cells infected with the empty (e.) vector
807 (ctr), the enzymatic dead construct K190M or the wild-type NSUN2 construct. NSUN2-
808 specific sites are highlighted in red. (**d, e**) Heatmaps (upper panels) and methylation level
809 (bottom panels) of VTRNA1.1, VTRN1.3, RPPH1, and HECTD1 in infected *NSUN2*-/- cells
810 (d), human embryonic fibroblasts (H9) and HEK293 cells (e). Shown are 5 independent
811 bisulfite conversion experiments. (**f**) Schematic illustration of NSUN2-dependent methylation
812 (CH3) of VTRNA1.1 and the small regulatory non-coding fragments svRNA4. (**g**) Abundance
813 of svRNA4 in the presence (+/-) and absence (-/-) of NSUN2. Methylation dead NSUN2-
814 mutant constructs (C271A; C321A) fail to rescue svRNA4 levels in *NSUN2*-/- cells. Error bars
815 indicate s.d. (n = 3 qRT-PCR reactions). *** $p<0.001$ unpaired student's t-test. (**h**) Log_2 fold-
816 change of *NSUN2* in the indicated cells compared to *NSUN2*+/+ control cells measured by
817 ribosome profiling. Source data are provided as a Source Data file.
818
819 ***Figure 2***. **SRSF2 preferentially binds un-methylated human VTRNA1.1.** (**a**) Of the 144
820 common proteins binding to VTRNA1.1 in two different RP-SMS experiments, a small number
821 bound methylated (red) or unmethylated (blue) VTRNA1.1 with higher affinity. (**b**) Gene
822 Ontology (GO) analyses of the 144 commonly bound proteins. (**c**) Western blot and Coomassie
823 stain for SRSF2 in HeLa cell lysates pulled-down with agarose beads coupled to methylated
824 (m^5C69) or un-methylated (C69) Vault-RNA1.1 (upper panel). hnRNP A1 serves as a loading
825 and RNA-binding control (lower panel). Numbers indicate band intensity versus loading
826 control. (**d**) Location of the putative SRSF2 RNA-binding motifs (RRM1 and RRM2) in
827 VTRNA1.1 (upper panel) and RNA pulldowns using wildtype or mutated (C69A; C88U) VT-
828 RNA1.1-constructs to confirm both putative SRSF2 binding sites are necessary for SRSF2

binding. Shown is mean and range (n = 2 independent experiments). Quantification in (**c, d**) was done using ImageJ. (**e**) EMSA assay using methylated (m^5C69) and unmethylated (C69) VTRNA1.1 to measure binding of recombinant SRSF2. (**f**) Quantification of (e). Error bars indicate s.d. (n = 3 experiments). ** $p<0.01$, * $p<0.05$ students t-test. Source data are provided as a Source Data file.

Figure 3. **Methylation-guided VT-RNA1.1 processing is altered in the absence of SRSF2**. (**a, b**) Western blot for endogenous SRSF2 (**a**) and NSUN2 (**b**) in NSUN2-expressing (+/+, +/-) and -lacking (-/-) human fibroblasts. Tubulin served as a loading control. (**c**) qRT-PCR measuring SRSF2-bound VTRNA1.1 normalized to the control (ctr; Rabbit serum conjugated with Dynabeads) after recovering the pulled down RNA in the SRSF2 immunoprecipitation. (**d**) Western blot for SRSF1 and SRSF2 in *NSUN2*-/- human fibroblasts infected with shRNAs (1.1, 1.6, 2.4, 2.5). Tubulin served as a loading control. (**e**) Fold-change (FC) of svRNA1 and 4 abundances after knock-down of SRSF1 and 2, relative to *NSUN2*-/- cells infected with the empty vector (e.V.). Shown are the pooled values using the two shRNA constructs shown in (d). Error bars indicate s.d. (n = 3-6 qRT-PCRs). **$p<0.01$ students t-test. Source data are provided as a Source Data file.

Figure 4. **VTRNA1.1 methylation and processing are altered during cell differentiation**. (**a**) Treatment regime of keratinocytes using calcium switch assay. (**b, c**) qRT-PCR to measure RNA levels of up-regulated (b) and down-regulated (c) markers at 2and 6 days after calcium treatment compared to the 0 day control. Error bars indicate s.d. (n=3 qRT-PCRs) ****$p<0.0001$, ***$p<0.001$, **$p<0.01$ multiple t-tests. (**d**) Methylation levels (n= 5 BS conversion reactions) at cytosine in RNA isolated from undifferentiated (undiff) and differentiated (diff) primary HK shown as box plots showing all points with minimum to maximum values. ****$p<0.0001$ Mann Whitney test. (**e**) Log_2 coverage (n=5 BS conversion

reactions) of sites in RNA isolated from undifferentiated (undiff) and differentiated (diff) primary HK. (**f**) Correlation between methylation levels at cytosines in undifferentiated and differentiated primary HK. Elevated methylation levels at tRNAs (examples in blue) and VTRNA1.1 (red). (**g, h**) Methylation levels in VTRNA1.1 (g) and tRNA Leu CAA (h) in the indicated cells (left hand panels) and heat maps (right hand panels) showing methylation levels in the individual replicates. (**i**) Light microscope image comparing the morphology of primary HK transfected with a control siRNA (Ctr) or svRNA4 after 4 days of differentiation in high $CaCl_2$. Scale bar: 50 μm. (**j**) Western blot detecting KRT10 and OVOL1 in HK transfected with Ctr siRNA or svRNA4 four days after calcium-induction. Tubulin served as loading control. (**k, l**) Treatments regimes and transfection (upper panels) of svRNA4 (K) or anti-svRNA4 (L) and qRT-PCR (lower panels) to measure RNA levels of the indicated markers 4 days after calcium treatment. Error bars indicate s.d. (n=3 qRT-PCRs). ****p<0.0001, ***p<0.001, **p<0.01, *p<0.05 student's t-test. Source data are provided as a Source Data file.

Figure 5. **SRSF2 is required for cell cycle and survival of undifferentiated cells**. (**a**) Treatment regime and transfection of differentiating primary human keratinocytes. (**b**) Western blot detecting SRSF2 after treatment with *Srsf2* siRNA. Tubulin (TUBB) serves as loading control. (**c**) Quantification of RNA expression levels of *Ovol1, Vtrna1.1, Tgm*, *Krt10,* and *Ivl* after 6 days of calcium-induced differentiation versus untreated control (0 days). FC: Fold-change. Error bars indicate s.d. (n = 3 qRT-PCRs). **p<0.01 two-way ANOVA. (**d, e**) Log_2 RNA fold-change (FC) of differentiation markers (d) and cell cycle regulators (e) after 48 and 72 hours of *Srsf2* knock-down. Data shown as box plot with mean showing all data from minimum to maximum (n = 4). ****p<0.0001, ***p<0.001, **p<0.01, *p<0.05 Two-way ANOVA. (**f**) Cell cycle distribution of primary HK transfected with the *Srsf2* siRNA for 72 hours. Error bars Data shown as box plot showing all data from minimum to maximum (n = 10

Flow sorts). ****p<0.0001 Two-way ANOVA. (**g**) Light microscope images of HK transfected with a control (ctr) siRNA (upper panels) and a *Srsf2* siRNA (lower panels) after 48 and 72 hours in low calcium medium. Scale bar: 50 μm. (**h**) Small RNA qRT-PCR measuring the abundance of svRNA4 in primary HK transfected with the indicated siRNA constructs. Data shown as box plot with mean showing all data from minimum to maximum. Error bars represent s.d. (n = 4 qRT-PCRs). **p<0.01 unpaired student's t-test. (**i**) Illustration how levels of NSUN2, svRNA4, SRSF2, and OVOL1 change upon terminal differentiation in keratinocytes. Source data are provided as a Source Data file.

Figure 6. **Summary of VTRNA1.1 processing into svRNA4.** (**a**) Expression of both NSUN2 and SRSF2 (e.g. in progenitor cells) result in high levels of VTRNA1.1 methylation (CH3) and high levels of svRNA4 and down-regulation of OVOL1. (**b**) No NSUN2 in the presence of SRSF2 suppresses formation of svRNA4. (**c**) Lack of expression of both NSUN2 and SRSF2 release VTRNA1.1 from SRSF2 binding and increases the levels of svRNA4.

Figure 1

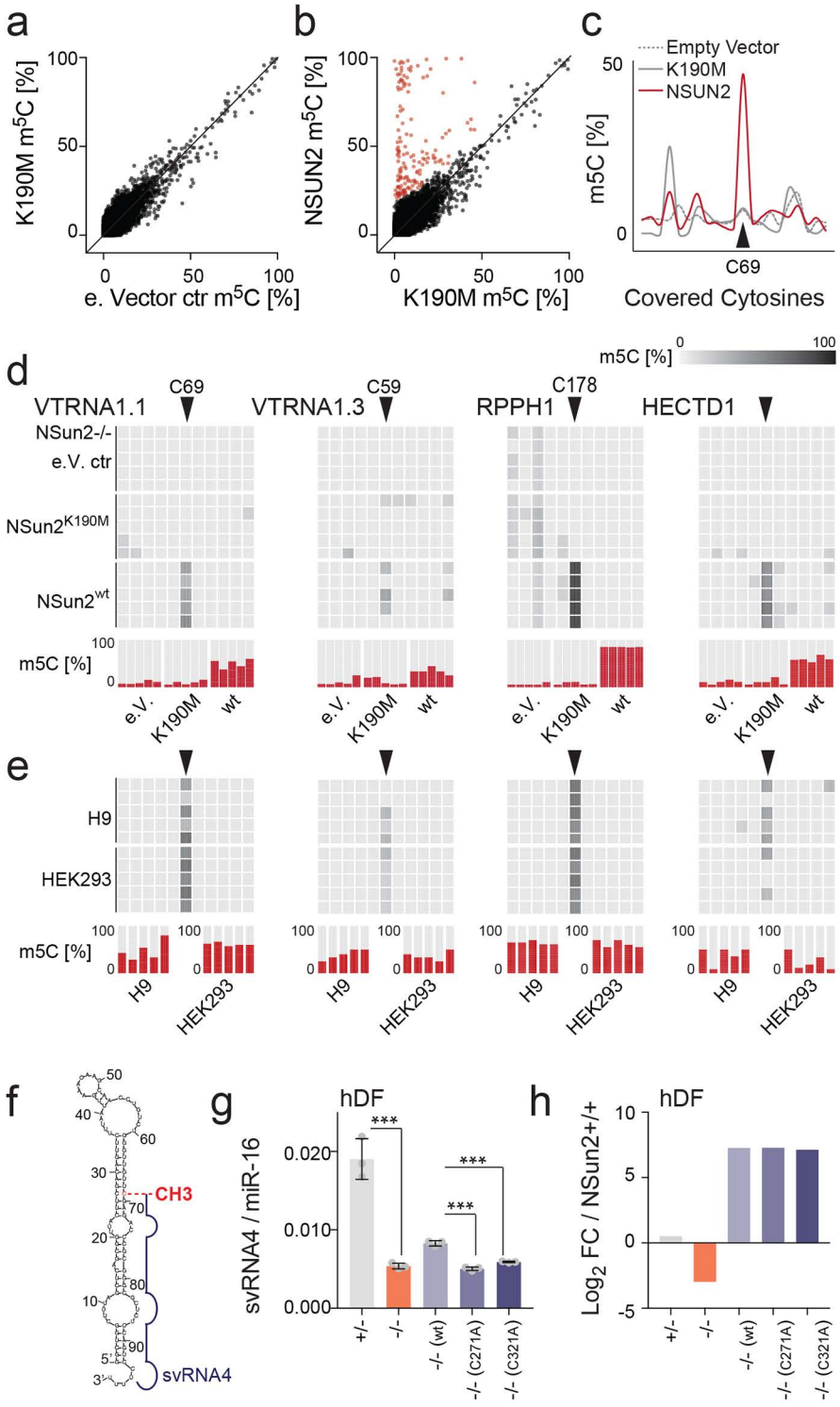

a
b
c
K190M m5C [%]
e. Vector ctr m5C [%]
NSUN2 m5C [%]
K190M m5C [%]
Empty Vector
K190M
NSUN2
m5C [%]
C69
Covered Cytosines
d
m5C [%]
VTRNA1.1
C69
VTRNA1.3
C59
RPPH1
C178
HECTD1
NSun2-/-
e.V. ctr
NSun2K190M
NSun2wt
m5C [%]
e.V.
K190M
wt
e
H9
HEK293
m5C [%]
f
CH3
svRNA4
g
hDF
svRNA4 / miR-16
+/-
-/-
-/- (wt)
-/- (C271A)
-/- (C321A)
h
hDF
Log2 FC / NSun2+/+

Figure 2

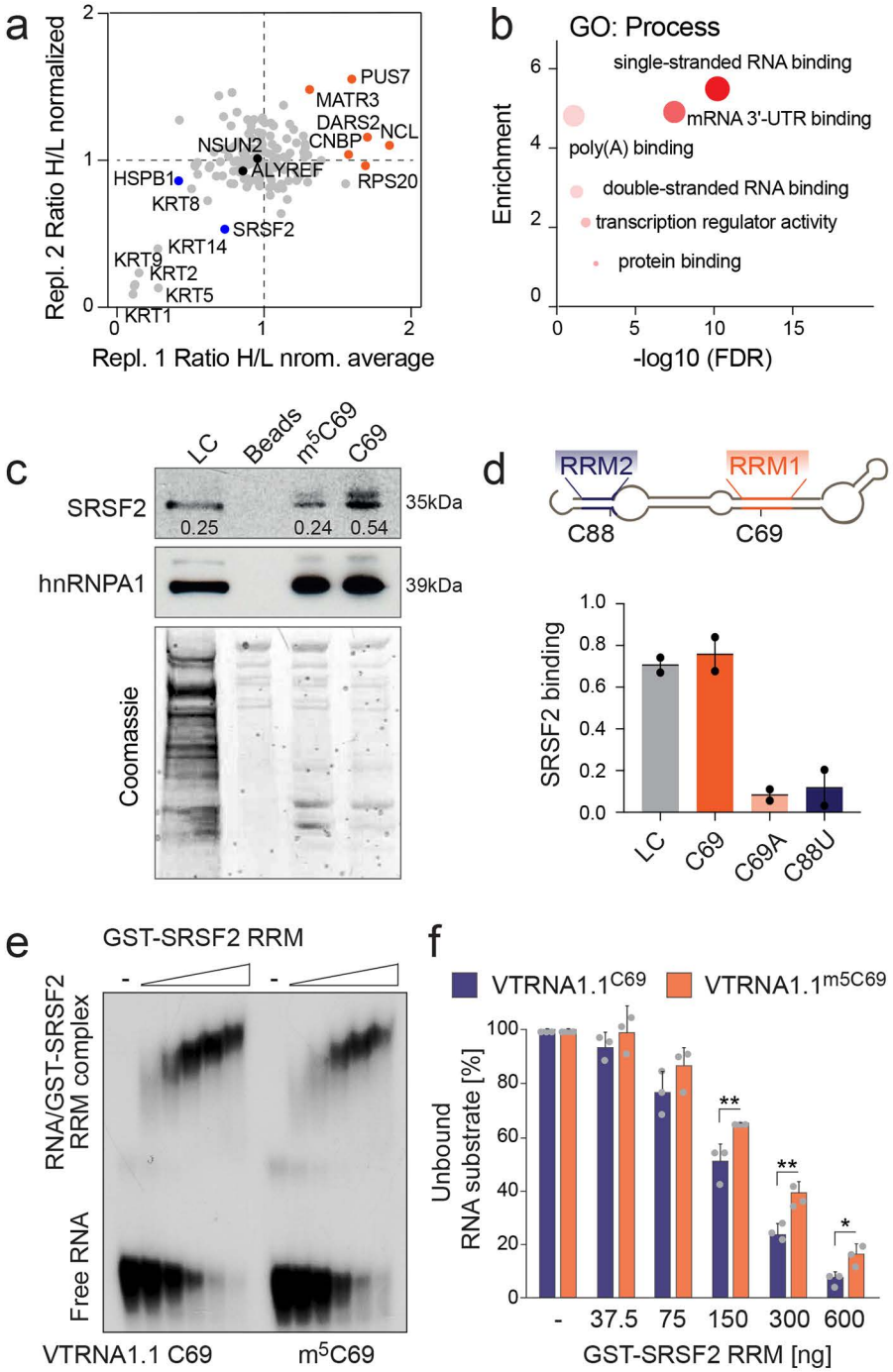

a
Repl. 2 Ratio H/L normalized
Repl. 1 Ratio H/L nrom. average
PUS7
MATR3
DARS2
NCL
CNBP
NSUN2
ALYREF
RPS20
HSPB1
KRT8
SRSF2
KRT14
KRT9
KRT2
KRT5
KRT1
b
GO: Process
single-stranded RNA binding
mRNA 3'-UTR binding
poly(A) binding
double-stranded RNA binding
transcription regulator activity
protein binding
Enrichment
-log10 (FDR)
c
LC
Beads
m5C69
C69
SRSF2
35kDa
0.25
0.24
0.54
hnRNPA1
39kDa
Coomassie
d
RRM2
RRM1
C88
C69
SRSF2 binding
LC
C69
C69A
C88U
e
GST-SRSF2 RRM
RNA/GST-SRSF2 RRM complex
Free RNA
VTRNA1.1 C69
m5C69
f
VTRNA1.1C69
VTRNA1.1m5C69
Unbound RNA substrate [%]
GST-SRSF2 RRM [ng]

Figure 3

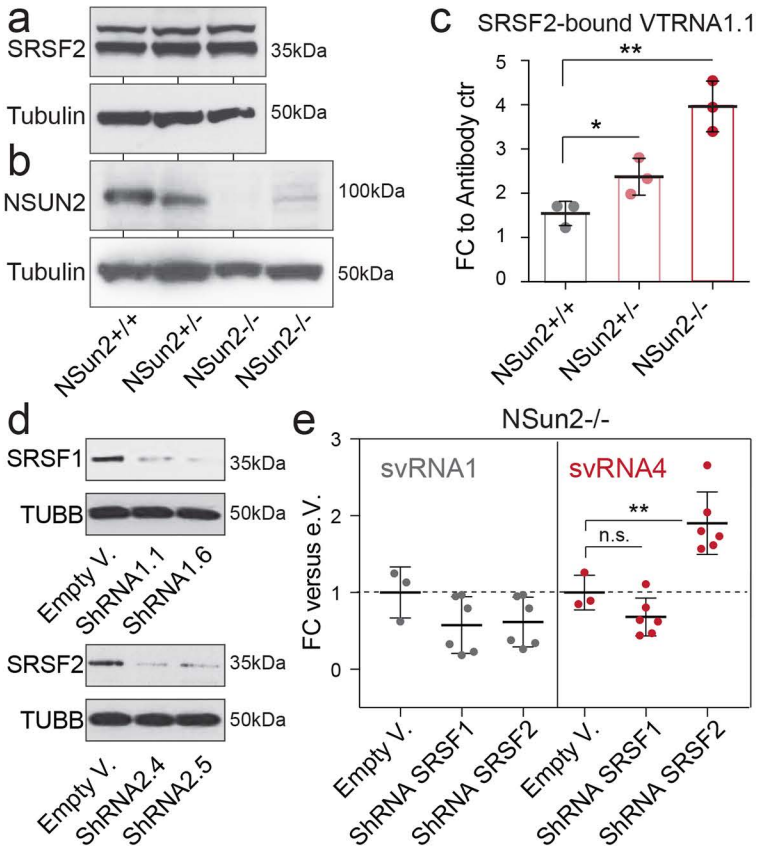

a
SRSF2
35kDa
Tubulin
50kDa
b
NSUN2
100kDa
Tubulin
50kDa
NSun2+/+
NSun2+/-
NSun2-/-
NSun2-/-
c
SRSF2-bound VTRNA1.1
FC to Antibody ctr
**
*
NSun2+/+
NSun2+/-
NSun2-/-
d
SRSF1
35kDa
TUBB
50kDa
Empty V.
ShRNA1.1
ShRNA1.6
SRSF2
35kDa
TUBB
50kDa
Empty V.
ShRNA2.4
ShRNA2.5
e
NSun2-/-
svRNA1
svRNA4
FC versus e.V.
**
n.s.
Empty V.
ShRNA SRSF1
ShRNA SRSF2
Empty V.
ShRNA SRSF1
ShRNA SRSF2

Figure 4

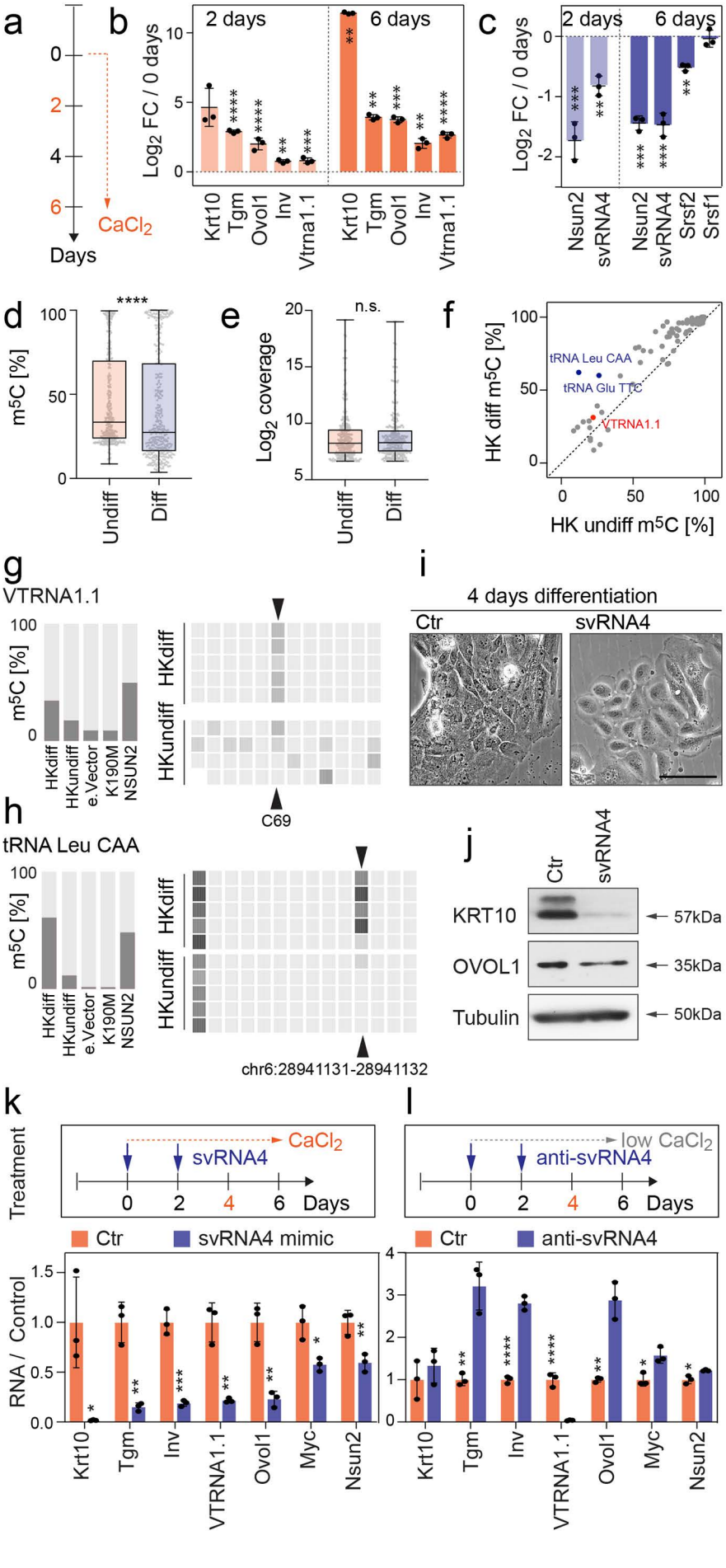

a
b
c
2 days
6 days
Log2 FC / 0 days
CaCl2
Days
Krt10
Tgm
Ovol1
Inv
Vtrna1.1
Nsun2
svRNA4
Srsf2
Srsf1
d
m5C [%]
Undiff
Diff
e
Log2 coverage
n.s.
f
HK diff m5C [%]
HK undiff m5C [%]
tRNA Leu CAA
tRNA Glu TTC
VTRNA1.1
g
VTRNA1.1
HKdiff
HKundiff
e.Vector
K190M
NSUN2
C69
h
tRNA Leu CAA
chr6:28941131-28941132
i
4 days differentiation
Ctr
svRNA4
j
KRT10
OVOL1
Tubulin
57kDa
35kDa
50kDa
k
l
Treatment
svRNA4
anti-svRNA4
low CaCl2
svRNA4 mimic
anti-svRNA4
RNA / Control
Krt10
Tgm
Inv
VTRNA1.1
Ovol1
Myc
Nsun2

Figure 5

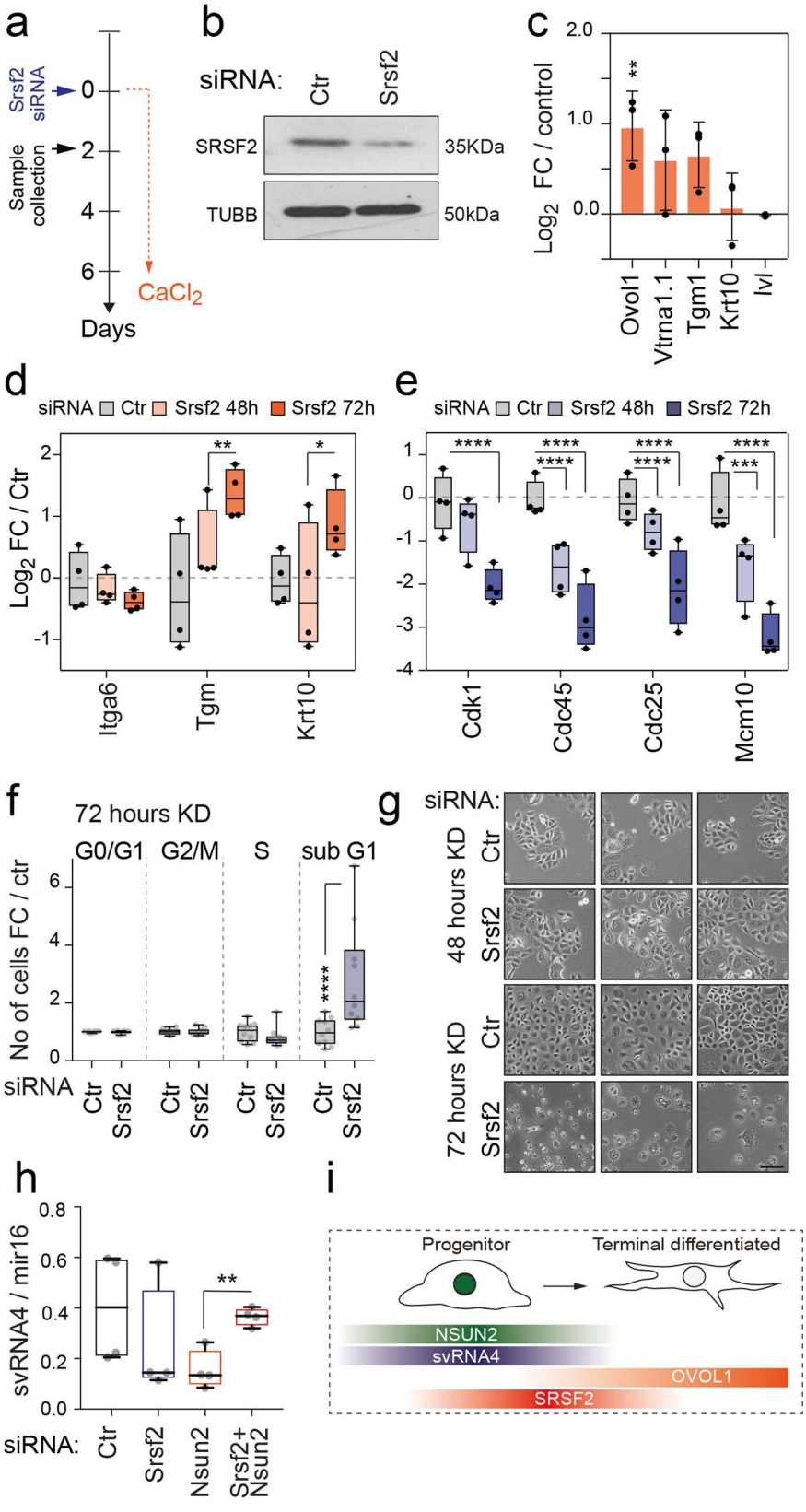
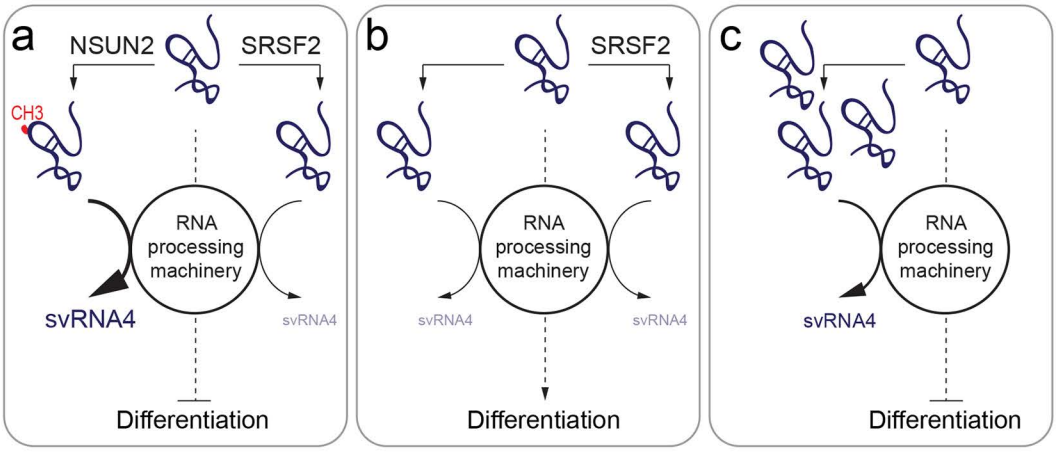

a
Srsf2 siRNA
Sample collection
0
2
4
6
Days
CaCl2
b
siRNA:
Ctr
Srsf2
SRSF2
35KDa
TUBB
50kDa
c
Log2 FC / control
2.0
1.0
0.0
**
Ovol1
Vtrna1.1
Tgm1
Krt10
Ivl
d
siRNA Ctr Srsf2 48h Srsf2 72h
Log2 FC / Ctr
**
*
Itga6
Tgm
Krt10
e
siRNA Ctr Srsf2 48h Srsf2 72h
****
***
Cdk1
Cdc45
Cdc25
Mcm10
f
72 hours KD
G0/G1
G2/M
S
sub G1
No of cells FC / ctr
****
siRNA
Ctr
Srsf2
g
siRNA:
Ctr
Srsf2
48 hours KD
72 hours KD
h
svRNA4 / mir16
**
siRNA:
Ctr
Srsf2
Nsun2
Srsf2+ Nsun2
i
Progenitor
Terminal differentiated
NSUN2
svRNA4
OVOL1
SRSF2

Figure 6

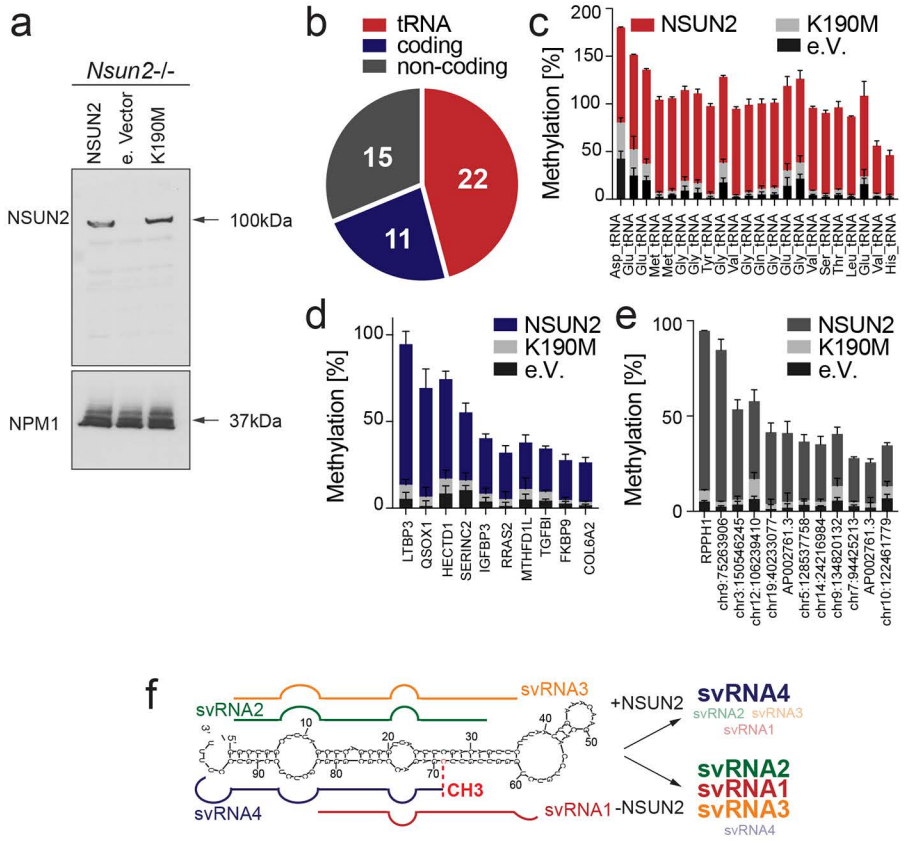

a
NSUN2
SRSF2
CH3
RNA processing machinery
svRNA4
svRNA4
Differentiation
b
SRSF2
RNA processing machinery
svRNA4
svRNA4
Differentiation
c
RNA processing machinery
svRNA4
Differentiation

Supplementary Information

Loss of 5-methylcytosine alters the biogenesis of Vault-derived small RNAs to coordinate epidermal differentiation

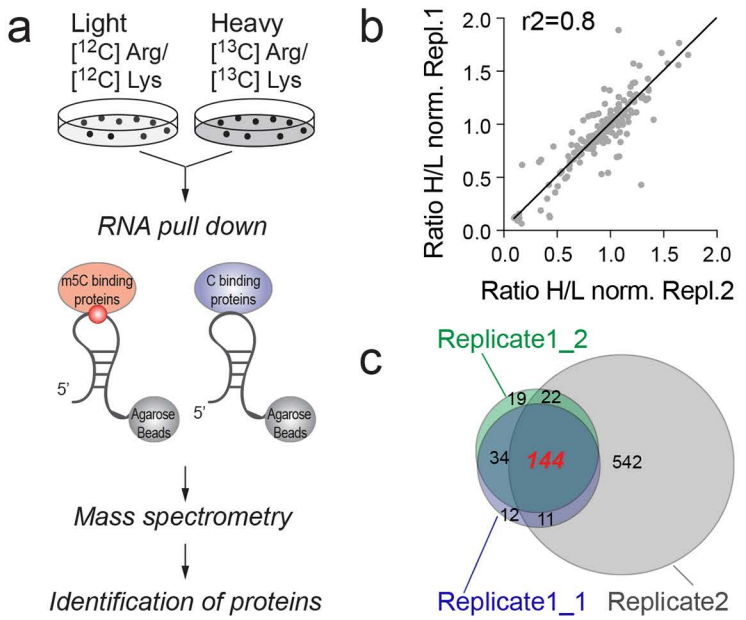
Sajini et al.

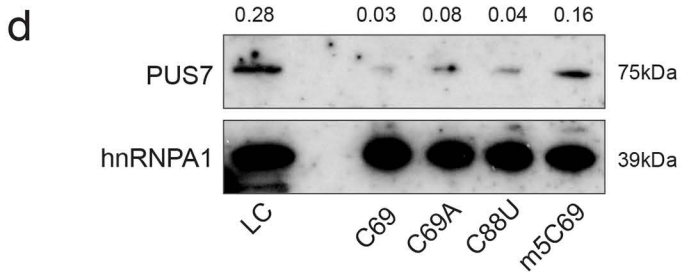
Supplementary Figure 1

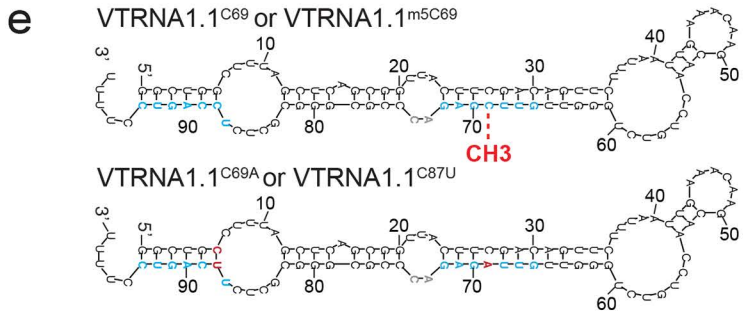


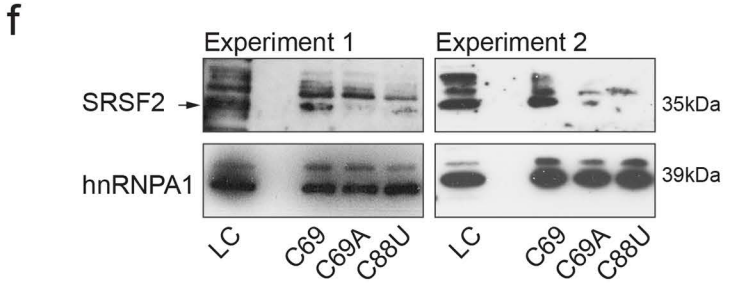
Supplementary Figure 1. NSUN2-dependent RNA methylation. (**a**) Western blot detecting NSUN2 in *NSUN2*-/- cells infected with the empty (e.) vector (ctr), the enzymatic dead construct K190M, and the wild-type NSUN2 construct. (**b**) Number of NSUN2-dependent methylation sites in tRNA (red), coding (blue) and non-coding (grey). (**c-e**) Methylation of tRNA (c), coding RNA (d) and non-coding RNA (e) in *NSUN2*-/- cells infected with the empty (e.) vector (ctr), the enzymatic dead construct K190M, or the wild-type NSUN2 construct. (**f**) Schematic illustration of NSUN2-dependent methylation (CH3) of VTRNA1.1 and the small regulatory non-coding fragments svRNA1-4. Error bars indicate s.d. (n = 5 bisulfite conversion assays) (c-e). Source data are provided as a Source Data file.

Supplementary Figure 2



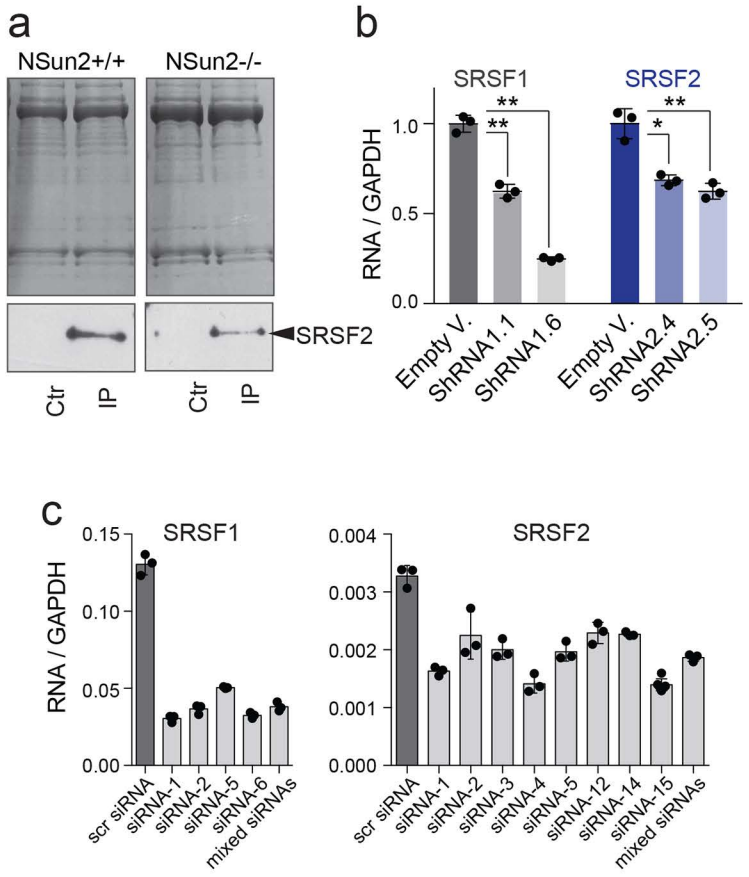






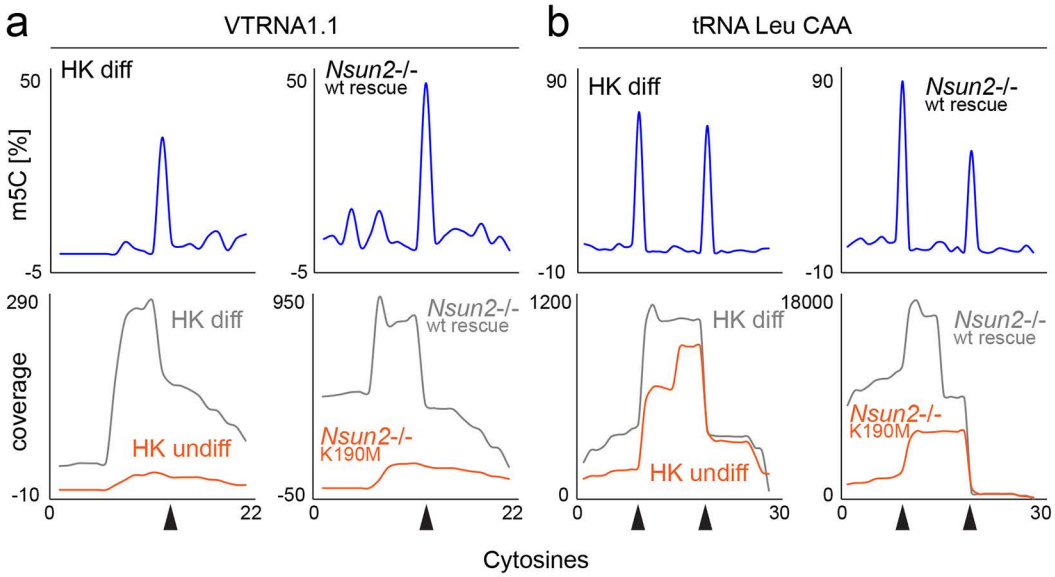
Supplementary Figure 2. SRSF2 and PUS7 bind methylated VTRNA1.1. differentially. (a) Schematic illustration of the SILAC-based quantitative mass-spectrometry assay. (**b**) Correlation of differential binding of proteins to VTRNA1.1 in two technical replicates. (**c**) Venn diagram showing a total 144 commonly identified proteins in two independent replicates. (**d**) RNA pull-down of methylated (m^5C69) and unmethylated (C69) VTRNA1.1 or VTRNA1.1 carrying the indicated point mutations followed by Western blot for PUS7. Numbers indicate band intensity (ImageJ). (**e**) Human VTRNA1.1-constructs with highlighted (blue) putative SRSF2-binding sites (upper panel) and the constructs carrying the point mutations to interrupt the binding sites (lower panel). (**f**) RNA pull-downs using wildtype or mutated (C69A; C88U) VT-RNA1.1-constructs confirm both putative SRSF2 binding sites are necessary for SRSF2 binding. hnRNPA1 served as a loading control in (d, f). Source data are provided as a Source Data file.

Supplementary Figure 3



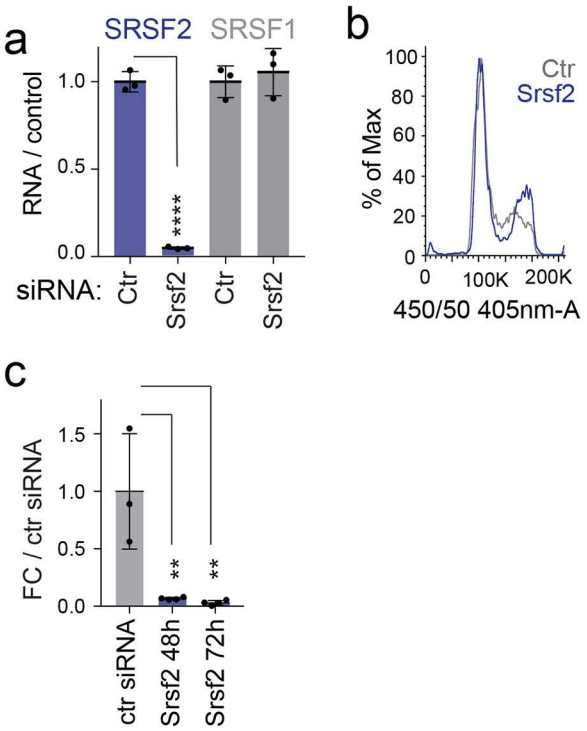
Supplementary Figure 3. Biogenesis of svRNA4 is altered in the absence of SRSF2. (**a**) Coomassie stained gel (upper panels) and immunoprecipitation (IP) detecting SFSR2 in *NSUN2*+/+ and *NSUN2*-/- lysates (lower panels). Ctr: Rabbit serum conjugated with Dynabeads. (**b**) SRSF1 and 2 RNA levels in *NSUN2*-/- human fibroblasts infected with shRNA empty vector (Empty V.), SRSF1 specific shRNAs (1.1 and 1.6) and SRSF2 specific shRNAs (2.4 and 2.5). Values were normalized to GAPDH. **p<0.01, *p<0.05 ANOVA. (**c**) Knock-down of SRSF1 (left hand panel) and SRSF2 (right hand panel) using different siRNA constructs. Shown are mean values of RNA levels versus GAPDH. Error bars indicate s.d. (n=3 qRT-PCRs). Source data are provided as a Source Data file.

Supplementary Figure 4



Supplementary Figure 4. Coverage of NSUN2-dependent methylation. (**a, b**) Methylation level (n=5 bisulfite conversion experiments) (upper panels) and coverage (reads) of the sites (lower panels) in VTRNA1.1 (a) and tRNA Leu CAA (b) in undifferentiated (undiff) and differentiated (diff) primary human keratinocytes (HK) (left hand panels) and *NSUN2*-/- human dermal fibroblasts rescued with the wild-type (wt) or enzymatic dead (K190M) construct of NSUN2.

Supplementary Figure 5



Supplementary Figure 5. SRSF2 is required for cell divisions. (**a**) RNA expression levels of *Srsf1* and *Srsf2* after transfection of a control siRNA (ctr) or a siRNA for *Srsf1* and *Srsf2*. Error bars represent s.d. (n =3 qRT-PCRs). ****p<0.0001 unpaired t-test. (**b**) Cell cycle profile of primary HK after 24 hours of *Srsf2* knock-down compared to control siRNA (Ctr) treated cells in high calcium medium. (**c**) RNA expression levels of *Srsf2* after transfection of a control siRNA (ctr) or a siRNA for *Srsf2*. Error bars represent s.d. (n = 3-4 qRT-PCRs). **p<0.01 One-way ANOVA. Source data are provided as a Source Data file.