FUS controls the processing of snoRNAs into smaller RNA fragments that can regulate gene expression

Patrycja Plewka[1], Michal Szczesniak[2], Agata Stepien[1], Marek Zywicki[3], Andrzej Pacak[1], Martino Colombo[4], Izabela Makalowska[2], Marc-David Ruepp[4,5], Katarzyna Dorota Raczynska[1]*

[1]Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland; [2]Department of Integrative Genomics, Institute of Anthropology, Adam Mickiewicz University in Poznan, Poznan, Poland; [3]Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland; [4]Department of Chemistry and Biochemistry, University of Bern, Bern, Switzerland; [5]UK Dementia Research Institute Centre at King's College London, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

*To whom the correspondence should be addressed: doracz@amu.edu.pl

ABSTRACT

FUS is a multifunctional protein involved in many pathways of RNA metabolism in human cells, including transcription, splicing, miRNA processing and replication-dependent histone gene expression. In this paper, we show for the first time that in human cells FUS can mediate the biogenesis of sdRNA, snoRNA-derived RNAs, that can be further involved in the regulation of gene expression. Using RNA immunoprecipitation followed by high throughput sequencing we identified snoRNAs in FUS-immunoprecipitated fraction. The interaction of FUS with a snoRNA fragment was further confirmed by EMSA and double filter binding assay. We observed that FUS negatively influences the level of selected mature snoRNAs in cells. Scanning of available human small RNAs databases revealed the existence of sdRNAs with the length of 19-33 nucleotides, that can be derived from FUS-dependent snoRNAs. Further *in silico* approach enabled us to predict putative targets for these sdRNAs. Our preliminary results indicate that sdRNAs may bind to the untranslated region of target mRNAs and influence their posttranscriptional stability or translation. Moreover, we identified a sdRNA that can interact with noncoding transcript and destabilize it, which in turn, might stabilize the level of mRNA transcribed from the same genomic region.

INTRODUCTION

Small nucleolar ribonucleoproteins (snoRNPs) are nucleolus-localized complexes consisting of small nucleolar RNAs (snoRNAs) associated with highly conserved core proteins. They are grouped in two major classes referred to as "box C/D" and "box H/ACA" snoRNPs, depending on the presence of "C/D" or "H/ACA" sequence motif in the snoRNAs. Most snoRNPs function in precursor rRNA (pre-rRNA) processing by introducing sequence-specific modifications guided by snoRNAs. C/D snoRNPs are responsible for site-specific 2'-*O*-ribose methylation, while the H/ACA snoRNPs catalyze isomerization of uridine to pseudouridine. Moreover, a set of snoRNPs is Cajal body-specific (scaRNPs) and serve to modify the snRNA component of U snRNPs. Two enzymes are responsible for the catalytic activity of snoRNPs: 2'-*O*-methyltransferase fibrillarin and pseudouridine synthase dyskerin (DKC1) [1-3]. Interestingly, a subset of snoRNAs that pair with rRNAs do not direct their modification but rather act as molecular chaperones required for correct folding of the pre-rRNAs (reviewed in [4]). Furthermore, in higher eukaryotes, many so called "orphan" snoRNAs were discovered. They contain either the box C/D or box H/ACA motif but do not contain an rRNA antisense region, suggesting that they may play different roles in the cell [1].

The snoRNP protein core includes four proteins. C/D box snoRNPs consist of: fibrillarin, NOP56, NOP5/NOP58 and a 15.5 kDa protein; H/ACA box snoRNPs consist of: dyskerin, GAR1, NHP2 and NOP10 [3, 5]. The assembly of snoRNPs occurs entirely in the nucleus and is a complex multiple step process

that requires numerous transient 'assembly factors' not found in the mature, catalytically active snoRNP complexes. These assembly factors ensure efficiency, specificity and quality control of the snoRNP production. The assembly occurs in three major stages: i) formation of protein-only complex, that contains particular core proteins and assembly factors, ii) incorporation of nascent snoRNAs, and iii) release of the assembly factors and activation of snoRNP catalytic activity [6, 7]. Formation of the complete core snoRNP complex is required for its nucleolar localization [8].

Most vertebrate snoRNAs are encoded within introns of pre-mRNAs, and the assembly of box C/D snoRNPs is generally dependent on splicing [9, 10]. In contrast, the assembly of box H/ACA snoRNPs is believed to occur on nascent pre-mRNAs, in a splicing independent manner, and is connected to RNA polymerase II (RNAP2) transcription. This connection is mediated by Nuclear Assembly Factor 1 (NAF1), which interacts with the C-terminal domain of the large subunit of RNAP2, and which may recruit H/ACA core proteins to the newly synthesized snoRNAs [11]. Nevertheless, in the case of both classes of snoRNPs, splicing of the host pre-mRNA is essential for providing linear precursor snoRNA, that is generated from debranched intron lariat and further processed by exonucleases: XRN1/2 families from the 5'-end, and RNA exosome with NEXT complex from the 3'-end (reviewed in [7]).

Recent studies based on human cells have identified numerous circular intronic long noncoding RNAs as well as long noncoding RNAs produced from introns with two imbedded snoRNAs genes [12, 13]. Similarly, stable intronic sequence RNA (sisRNAs) have been described in *Xenopus leavis* [14]. These sisRNAs can be linear or form debranching-resistant lariat structures and can be accumulated both in the nucleus and in the cytoplasm. Interestingly, the cytoplasmic lariat sisRNAs originate from small introns (<200 nt) and can encode snoRNA genes [14]. The function of intronic RNAs is largely unknown, they may play a role in local gene transcription, sequester proteins or even act as depositories of snoRNAs in cells.

A growing amount of evidence indicates that many snoRNAs are processed into shorter functional forms with the length of 19-40 nucleotides (reviewed in [4, 15]). snoRNA-derived RNAs (sdRNAs) have been identified in animals (human [16-22]), rodents [18, 23], Drosophila, chicken [18]) as well as in plants, fission yeast (*Saccharomyces cerevisiae*) [18], protozoa (*Giardia lamblia*) [24] and the Epstein-Barr virus [25]. The mechanism of sdRNAs generation is still not completely understood. It was shown that many H/ACA-box sdRNAs are produced in a Dicer-dependent manner and associate with Argonaute (AGO) proteins [16-19, 26, 27]. On the contrary, C/D box-derived small RNAs are not efficiently incorporated into the AGO2 protein. Most of them originate from the termini of mature snoRNAs, and hence carry "C" and "D" box motifs [28, 29].

Functionally, many sdRNAs display microRNA-like properties in posttranscriptional gene silencing activity [16-19]. However, in recent years numerous non-canonical functions have been ascribed to sdRNAs, such as mediating in mRNA editing and splicing [15, 27, 30, 31]. Interestingly, a Dicer-dependent sdRNA was identified in the cytoplasmic fraction of *Giardia lamblia* cells, and its function in translational regulation was experimentally validated [24]. The presence of sdRNA (originating from both box C/D and H/ACA snoRNAs) was also demonstrated in *S. cerevisiae*. Interestingly, their processing events were most prominent under non-optimal conditions for yeast growth, which include UV irradiation, oxygen deprivation, high/low pH or medium with no amino acids or sugars. Such response to environmental conditions may therefore indicate that the sdRNA processing may play a crucial role in the regulation of stress-dependent metabolism [32]. Significantly, these sdRNAs have been co-purified with the yeast ribosomes implicating a novel, yet undiscovered regulatory role of sdRNAs in ribosome biosynthesis [32].

In this paper we show for the first time that Fused in Sarcoma/Translocated in Liposarcoma (FUS/TLS; named FUS hereafter) is involved in the processing of small RNAs derived from mature snoRNAs in human cells. FUS belongs to the FET family of proteins which includes three highly conserved, abundant and ubiquitously expressed RNA-binding proteins: FUS, EWS and TAF15 [33]. FUS is predominantly present in the nuclear matrix, although it is also found in cytoplasmic fraction and is supposed to participate in nucleo-cytoplasmic shuttling and transport of mRNA [34]. The protein also binds to DNA (ssDNA as well as dsDNA) facilitating DNA annealing and D-loop formation thus mediating

in genomic maintenance, DNA recombination and the DNA repair pathway [35-37]. Moreover, FUS regulates several key steps of RNA metabolism, including: transcription, splicing and alternative splicing (reviewed in [38, 39]). Additionally, FUS is involved in replication dependent histone gene expression [40] and may play a role in miRNA processing as a component of the large Drosha microprocessor complex [41]. Notably, several FUS/TLS mutations have been found in familial forms of Amyotrophic Lateral Sclerosis (ALS) and Fronto Temporal Lobar Degeneration (FTLD), implicating a role of this protein in neurodegenerative diseases [33, 39, 42-45].

Here we show that FUS can bind snoRNAs in human cells. Using RNA immunoprecipitation with anti-FUS antibodies followed by high throughput sequencing (RIP-seq) we identified all three classes of snoRNAs in immunoprecipitated fraction. The interaction of FUS with snoRNA fragment was further confirmed by EMSA and double filter binding assay. Surprisingly, we observed that FUS negatively influences the level of mature snoRNAs in cells although the splicing efficiency of snoRNA-hosting introns and the level of snoRNA precursors were not altered. Scanning of available human small RNAs databases revealed existence of sdRNAs with the length of 19-33 nucleotides, that can be derived from selected snoRNAs. Therefore we suggest that FUS does not affect snoRNAs biogenesis but rather competes with snoRNP proteins to regulate the synthesis of FUS-dependent sdRNAs. Further *in silico* approach enabled to predict putative targets for our sdRNAs. Our preliminary results show that FUS-dependent sdRNAs might regulate gene expression at the level of transcripts stability and translation.

MATERIALS AND METHODS

Cell culture

HeLa and HEK 293T cells were grown in Dulbecco's modified Eagle medium with L-glutamine and 4.5 g/L glucose (DMEM; Lonza) supplemented with 10% fetal calf serum (Gibco) and antibiotics (100 U/ml penicillin, 100 µg/ml streptomycin, 0.25 µg/ml amphotericin B (Sigma)) at 37°C in a moist atmosphere containing 5% CO2. HEK293T cells with FUS overexpression (FUS OE), HEK293T cells with EBFP overexpression (EBFP OE) and HeLa cells with inducible knockdown of FUS (FUS KD) were prepared as described previously [40] (Supplementary Fig. S1). HeLa cells with FUS knockout (FUS KO) were kindly provided from Dr. Marc-David Ruepp laboratory [46] (Supplementary Fig. S1). Plasmid DNA transfections were performed with VIROMER® RED (Lipocalyx) and Fugene® HD Transfection Reagent (Roche), according to manufacturer's instruction.

Antibodies

The following primary antibodies were used in this work: anti-FUS (homemade), anti-actin (MP Biomedicals), anti-Symplekin (Bethyl Laboratories), anti-FLAG (Sigma Aldrich), anti-BRE (Santa Cruz Biotechnology), anti KCNQ1 (Santa Cruz Biotechnology). The following secondary antibodies were used: goat anti-rabbit IgG-HRP, goat anti-mouse IgG-HRP (Santa Cruz Biotechnology).

RNA isolation, cDNA preparation, PCR and qPCR

RNA isolation, PCR and qPCR were performed as described in [40]. The statistical significance of qPCR results was determined by Student's T test. cDNA was synthetized using random hexamers or in coupled polyadenylation reverse transcription reaction [40]. To detect lariat structure cDNA was amplified by PCR using outward-facing primers and detected product with expected size was further cloned and sequenced, as described in [14]. Primers used for PCR and qPCR are available on request.

Northern blot

Northern blot analysis was performed as previously described [47]. Shortly, total RNA (40µg) was separated in 7M urea 15% denaturing polyacrylamide gel. The electrophoresis was run in 20 mM MOPS–NaOH buffer (pH 7.0) and RNA Decade™ Marker (Thermo Fisher Scientific) was loaded as a size marker. After electrophoresis RNA was transferred onto Amersham Hybond-NX nylon membrane with the use of a Trans-Blot SD Semi-Dry Transfer Cell (BioRad) by applying 20V for 1 h and then fixed using EDC-mediated (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide) chemical cross-linking method (2h in 55°C). For detection of snoRNAs and sdRNAs, a 1h pre-hybridization and a 16h hybridization at 37°C

were performed in hybridization buffer (3.5% SDS, 0.375 M sodium phosphate dibasic, 0.125 M sodium phosphate monobasic) with DNA probes (Sigma) labelled at 5'end with [γ-^{32}P]-ATP. The membranes were then washed twice in 2x SSC, 0.1% SDS buffer to remove any unbound probes and blots were exposed for three days to phosphorimaging screen (Fujifilm) and scanned with Fujifilm FLA5100 reader (Fujifilm). The hybridization signals were quantified with the use of Multi Gauge V2.2 software.

Electrophoretical mobility Shift Assay and Double Filter Binding Assay

RNA fragment encompassing C/D motifs of SNORD45c and control sequence (control RNA) encompassing GGUG motif recognized by FUS (as described by [48]) were labeled by *in vitro* transcription in the presence of [α-^{32}P]-UTP. Recombinant FLAG-FUS and FLAG-EBFP (as negative control) proteins were overexpressed and purified from HEK293T cells. For EMSA and Double Filter Binding Assay 150 ng of protein were incubated with labeled RNA fragment (10 000 cpm) for 30 min at RT in buffer containing 20 mM Tris-HCl pH 7.4, 50 mM KCl and 10% glycerol. To test the binding specificity, unlabeled RNA fragment or tRNA (30 ng and 60 ng) was added to the sample 5 min before labeled RNA was added. All RNAs were denatured for 1 min in 95°C and cooled to RT before adding to the reaction. Samples were loaded on 5% polyacrylamide gel and electrophoresis was run in 0.5 TBE for about 2.5 hours in 4°C at 200 V, 15-20 mA. For Double Filter Binding Assay, after incubation, samples were passed through Nitrocelulose/Amersham Hybond membranes using vacuum. The gel and membranes were dried and exposed to phosphorimaging screen (Fujifilm).

Luciferase Reporter Assay

The 3'UTR region of Brain and Reproductive organ-Expressed protein (BRE) mRNA was amplified by PCR with primers introducing *NheI* and *SalI* restriction sites and inserted into pmirGLO vector after linearization by *NheI* and *SalI*, leading to the formation of LUC-3'UTR BRE luciferase reporter plasmid. PCR primer sequences are available on request. Control cells FUS KO cells and FUS OE cells (around $2x10^4$ cells/well) were co-transfected with 0.6 µg of LUC-3'UTR BRE luciferase reporter plasmid or pmirGLO as control plasmid (0.4 µg) using VIROMER® RED reagent (Lipocalyx), according to manufacturer's instructions. Luciferase activity was determined after 48h using the Dual Glo® Luciferase Assay (Promega) and Infinite M200 PRO TECAN Luminometer. Relative luciferase activities were calculated as the ratios between *Renilla* and firefly Luciferase activities, and further normalized to control treatments.

Protein isolation, RNA immunoprecipitation, Western blot

Total protein extracts were prepared by lysis using hypotonic lysis buffer or RIPA buffer without SDS as described in [49]. Cytoplasmic and nuclear fractions were prepared as described previously [40]. For RNA immunoprecipitation total and nuclear protein extracts from cells overexpressing FLAG-tagged FUS and EBFP proteins were gently rotated overnight at 4°C with anti-FLAG antibody-coupled magnetic beads (Sigma-Aldrich), then washed five times with PBS-T. After washing, co-precipitated RNA was incubated with 1 U RNase T1 (Ambion) for 20 min at 25°C with 500 rpm shaking. Next, samples were washed three times with PBS-T and then treated with 50 U of alkaline phosphatase (Thermo Fisher Scientific) for 20 min at 25°C with 500 rpm shaking to remove phosphate groups at the 3' ends of RNA followed by three more PBS-T washes. RNA was then treated with 20 U of kinase (OptiKinase Affimetrix) to add phosphate groups at its 5' ends. PBS-T washing was repeated thrice and RNA was eluted from the beads with TRIZOL and used for library preparation. For Western blot analysis samples were separated by SDS-polyacrylamide gel electrophoresis (PAGE), transferred to polyvinylidene difluoride (PVDF) membrane (Millipore) and incubated first with specific primary antibodies and then with corresponding species-specific horseradish peroxidase (HRP)-coupled secondary antibodies followed by detection using the enhanced chemiluminescence method (ECL, GE Healthcare).

Library preparation for RNA-seq

Libraries were prepared using TruSeq Small RNA Library Prep Kit (Illumina). In brief immunoprecipitated RNAs derived from particular samples were ligated to 3' and 5' RNA adapters. In

further steps reverse transcription and PCR reactions were performed. PCR products were indexed by utilization of specific RNA PCR Index Primers (RPI, Illumina). PCR products were separated using electrophoresis in 6% polyacrylamide gel containing 1% glycerol. After 10 min of staining in SYBR™ Gold Nucleic Acid Gel Stain (Invitrogen, Thermo Fisher Scientific)/0.5xTBE buffer, DNA fragments: 140 bp - 300 bp in length were cut and eluted overnight at 28°C with 400 rpm shaking in 400µl elution buffer (50 mM Mg-acetate, 0.5 M ammonium acetate, 1 mM EDTA, 0.1% SDS). After chloroform/phenol purification libraries were precipitated using 3 volumes of 100% EtOH in the presence of 1.5 µl GlycoBlue™ coprecipitant (15 mg/mL) (Ambion, Thermo Fisher Scientific). Purified libraries were quantified using Infinite® 200 PRO microplate plate reader with NanoQuant Plate™ (Tecan) and Quant-iT™ PicoGreen™ dsDNA Assay Kit dye (Invitrogen, Thermo Fisher Scientific). Libraries were then pooled in equimolar ratio before sequencing. Sequencing was performed using TruSeq SR Cluster kit v3-cBot-HS (Illumina), TruSeq SBS kit v3-HS (50 cycles) (Illumina) and HiScan™SQ System platform (Illumina).

Bioinformatic analysis

Raw sequencing reads were trimmed for Illumina 3' adaptor sequence using Cutadapt (Available online at: http://journal.embnet.org/index.php/embnetjournal/article/view/200). Quality of the cleaned libraries was verified using FastQC [Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc]. Next, Bowtie [50] (adjusted to return only hits within best strata) was used to align cleaned reads to the reference human genome assembly GRCh37 obtained from Ensembl database. SeqMonk software was used to perform the quantification. Briefly, overlapping reads were joined into contigs and RPM (reads per million of reads) for every contig was calculated. Next, contigs were annotated with overlapping genomic features using BEDTools suite and in-house scripts, allowing for multiple annotation of single contig. Finally, contig enrichment in immunoprecipitated sample was calculated as a $\log_2$ fold change over control sample. Similarly, nuclear enrichment of the contigs was calculated as a $\log_2$ fold change over total RNA sample.

Small RNA sequencing datasets were downloaded from Sequence Read Archive [51] (run IDs: SRR1586016 (HEK293, GSM1513689, 185.7 M bases); SRR3931966 (human neuroblastoma cells treated with Seyoflurane, 661.6 M bases) and SRR2728234 (blood (PTSG with comorbid depression, GSM1912509, 683 M bases). The reads were subjected to adapter clipping with FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) and only > 18 bases long, adapter-containing small RNAs were kept. Then, the reads were mapped to a set of human snoRNA sequences from ENSEMBL [52] with Bowtie. The mappings were then subjected to manual inspection to keep only putative snoRNA-derived miRNAs (length <= 24 nt, raw count > 10, most abundant per sequence arm) and most abundant sequences longer than 24 bases, with raw count > 50, representing most abundant mapped reads per sequence arm (Supplementary Table S2). To predict and analyse targets for short and long (>24 nt) sRNAs, Miranda [53] was used. This was done in two ways: i) using the whole transcriptome and ii) using only 3' UTRs parts of the human transcripts. The human transcriptome was downloaded from ENSEMBL 86 [52]. Then, top 100 genes with highest Miranda scores (i.e. those predicted to form the highest number and/or most stable interactions with sRNAs) were selected. The gene IDs obtained in this way were then used as an input in over-representation analysis for pathways and gene ontologies using ConsensusPathDB (http://consensuspathdb.org/) [54]. The predicted sRNA target regions were then compared to HEK293-specific data from PAR-CLIP and HITS-CLIP experiments against AGO1, AGO2, AGO3 and AGO4 from StarBase 2.0 database [55], after converting the genome coordinate system from hg19 to hg38, using LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver). This was done with Intersect tool from BEDTools suite [56], requiring that at least 50% of the targeted region was falling into the CLIP-Seq cluster. Only Miranda sRNA target predictions that overlap CLIP-Seq peaks were kept for further consideration.

Gene expression analysis after high throughput sequencing of RNA isolated from SH-SY5Y FUS KO cells was performed as described in [57].

RESULTS

FUS directly interacts with snoRNAs in human cells

Total and nuclear protein extracts from HEK293T cells overexpressing FLAG-FUS and FLAG-EBFP (as a negative control) were subjected to immunoprecipitation with anti-FLAG antibodies followed by isolation of RNA from immunoprecipitated fractions. Co-precipitated RNAs were analysed by high throughput sequencing. The analysis of small noncoding RNAs bound by FUS revealed a large fraction of small nucleolar RNAs. They belong to three classes of snoRNAs: C/D box snoRNAs (SNORD), H/ACA box-snoRNAs (SNORA) and scaRNAs. FUS-bound snoRNAs were enriched in fraction immunoprecipitated from nuclear extract and their distribution resembled the distribution of endogenous snoRNAs in the cell, that is around 67% SNORD, 27% SNORA and 6% scaRNA (Fig. 1, Supplementary Table S1). This suggests that FUS binds to snoRNAs in a class independent manner.

To confirm that FUS can directly interact with snoRNAs we performed double filter binding assay (Fig. 2A) and EMSA assay (Fig. 2B). For this purpose isotope labeled RNA fragment encompassing C/D motifs of SNORD45c and control sequence encompassing GGUG motif recognized by FUS (positive control, as described by [48]) were incubated with recombinant FLAG-FUS and FLAG-EBFP (as negative control). The binding specificity was tested by adding specific (unlabeled RNA) and nonspecific (tRNAs) competitors. As shown in Fig. 2 we confirmed that FUS can directly interact with snoRNAs.

FUS negatively regulates the level of mature snoRNAs

As FUS interacts directly with snoRNAs we addressed the question whether FUS can affect the level of mature snoRNAs in the cell. For that purpose we isolated total RNA from control cells and cells with FUS overexpression (FUS OE), inducible FUS knockdown (FUS KD) and FUS knockout (FUS KO) (Supplementary Fig. S1). cDNA synthesized in coupled polyadenylation reverse transcription reaction was used as a template in qPCR with universal reverse primer and snoRNA-specific forward primer. As a result, we observed that the level of selected mature snoRNAs changed inversely to the amount of FUS in the cell, suggesting that FUS acts as a negative regulator of mature snoRNAs (Fig. 3).

In vertebrates, great majority of box C/D and H/ACA snoRNAs are encoded within intronic sequences of pre-mRNAs and their maturation depends on splicing [9, 10]. Therefore, we first assumed that FUS affects snoRNA biogenesis by changing the splicing profile of snoRNA-encoding introns. To test this hypothesis we analysed the ratio between spliced and unspliced isoforms of snoRNA-hosting introns. We could not observe any changes in their splicing efficiency, either in cells with FUS overexpression or in cells with FUS depletion (Fig. 4A). Furthermore, as mature box C/D snoRNAs are processed from the removed and debranched introns by exonucleolytic activities we also tested the level of snoRNA precursors in cells with FUS overexpression and cells with FUS depletion (Fig. 4B). Again, we could not observe significant changes, suggesting that FUS influences snoRNA levels at later steps of snoRNA biogenesis.

FUS might induce processing of mature snoRNAs into smaller RNA fragments, sdRNAs

There have been many reports published within last few years concerning sdRNAs and their origin from snoRNA (reviewed in [4, 15]. Biogenesis of sdRNAs should lead to downregulation of mature snoRNAs in the cell. Therefore we addressed the question whether FUS can be involved in generation of small RNA fragments from selected mature snoRNAs. To answer this question we searched through available data from high throughput sequencing of small RNAs isolated from three different human cell lines (HEK293T cells (SRR1586016), neuroblastoma cells (SRR3931966) and whole blood (SRR2728234)). We looked for sdRNAs that could be derived from the following snoRNAs: SNORD104, SNORD45c, SNORD14c, SNORD68 (HBII-202), SNARNA2 (HBII-382), SNARNA4 (ACA26), as the level of these snoRNAs was highly affected by FUS (Fig. 3A-C). We observed that small RNAs could be mapped to all 6 snoRNAs in all three libraries, except for SNARNA4 (ACA26) in whole blood library. As shown in Fig. 5A, small RNAs map typically to the sequences in two clusters: one to the 3' side of the snoRNA with reads having quite homogeneous 3' end; the other close to the 5' end with homogeneous 5' end; with the exception of SNORD68 (HBII-202) [28]. Regarding the length, our *in silico* approach enabled us to identify putative short sdRNAs (with length <25 nt) embedded in SNORD45c, HBII-382 (SNARNA2), SNORD14c, SNORD104 (Supplementary Table S2). Besides short, miRNA-like reads, higher number of

longer reads were also mapped to the snoRNAs (with length ≥25 nt), embedded in SNARNA2(HBII-382), SNORD68 (HBII-202), SNORD104 (Fig. 5A) (Supplementary Table S2).

Next, by Northern blot hybridization we confirmed that sdRNA104 generated from the 5' end of SNORD104 indeed exists in the cell and its level depends on the FUS expression (Fig. 5B). We could not detect any other sdRNA by Northern blot hybridization, probably because of very low level of those molecules in cells.

sdRNAs can regulate stability of noncoding transcripts and translation of mRNAs

The function of many sdRNAs has not been elucidated yet. Some of them were shown to act like miRNAs as they suppress target gene expression through the inhibition of translation or acceleration of mRNA degradation after complementary Watson-Crick base pairing with target transcripts [58]. Others bind to introns/exons and can influence splicing profile of targeted pre-mRNAs [27, 30, 31]. To bring closer the function of FUS-dependent sdRNAs we first searched for their putative targets using Miranda software and either the whole transcriptome or only 3' UTRs as an input. The criteria of sdRNA:RNA hybrid formation resemble ones used for miRNAs hybridization to their targets. To support *in silico* target predictions, available datasets from CLIP-Seq experiment against AGO proteins were used. Next, we tested the level of expression of putative targets in cells with FUS knockout using high throughput sequencing results of RNA isolated from FUS KO cells. According to our hypothesis in these cells the level of FUS-dependent sdRNAs should be downregulated and accompanied by altered levels of their putative targets.

Our results revealed that FUS-dependent sdRNA68, originating from SNORD68 (HBII-202), might target KCNQ10T1-001 antisense transcripts encoded by the opposite strand of KCNQ1 protein coding gene (Fig. 6A). The sdRNA:target duplex is formed within the unique region that distinguishes noncoding RNA from mRNA transcribed from the same genomic region. KCNQ10T1 antisense transcripts (ENSG00000269821) encompass intron10, exon11 and intron11 of KCNQ1 (ENSG00000269821) mRNA and sdRNA-targeted region includes intron10 and intron11. The question arose whether FUS-dependent sdRNA can regulate the level of noncoding transcripts and affect protein coding RNA. RT-qPCR analysis was done using set of primers designed to amplify specifically noncoding or coding transcripts in cells with FUS depletion (in which we expect decreased level of sdRNAs) and cells with FUS overexpression (in which we expect increased level of sdRNAs). As shown in Fig. 6A in FUS KO cells the level of KCNQ1 mRNA was downregulated, which corresponds with the elevated level of noncoding, antisense KCNQ10T1 transcript. The opposite effect was observed in cells with FUS overexpression. Such a result suggested that sdRNA68 targets noncoding RNA for degradation that in turn promotes the synthesis of protein coding mRNA. Indeed, decreased level of KCNQ1 protein in FUS KO cells and slightly increased KCNQ1 protein level in FUS OE cells were confirmed by Western blot (Fig. 6B). Therefore, one of the functions of FUS-dependent sdRNA can be to regulate the ratio of noncoding/protein coding transcripts that can be crucial for protein synthesis.

Moreover, our *in silico* approach revealed that selected FUS-dependent sdRNAs can hybridize to 3'UTRs of mRNAs. This in turn may have an impact on translation efficiency. To evaluate the potential influence of sdRNAs on expression of targeted transcript we decided to test the effect of interaction between sd104 and *BRE* mRNA (ENSG00000158019). We cloned 3'UTR region of *BRE* transcript downstream of Luciferase (LUC) coding sequence. Then, the effect on LUC synthesis was tested in FUS KO cells, where we expect lower level of sdRNA. Interestingly, as shown in Fig. 6C, 48 hours after plasmid transfection we observed decreased level of LUC protein. The results were confirmed by Western blot using anti-BRE antibodies (Fig. 6D). Interestingly, this was contrary to the elevated level of target transcript detected in RNA-seq results from FUS KO cells, where we observed 1.27 fold enrichment of *BRE* mRNA. We concluded that FUS-dependent sdRNAs might stimulate translation or stabilize the binding of *BRE* transcripts to ribosomes. Therefore, the protein synthesis is downregulated when the sdRNA production is disrupted. However, we observed that expected upregulation of sdRNAs in FUS OE cells was not coupled with increased mRNA translation as we did not detect any significant changes in the protein level (Fig. 6C, D).

We also tested whether snoRNA can be trapped by circular lariat intron that are derived from spliced introns and escape debranching. Indeed, when we amplified cDNA by PCR using outward-facing primers we detected a product with expected size of lariat intron encoding SNORD14c (lariat 14c). The lariat 14c product was further cloned and its origin was confirmed by sequencing. Similarly, as described in [14], we figured out that although its sequence mapped to the region between the 5' and the 3' end of the intron, one end mapped precisely to the 5' splice site whereas the other end mapped to a region 14 nucleotides upstream, at the presumed branchpoint (Fig. 6E lower panel). This suggests that lariat 14c was derived by splicing, avoided debranching and was stabilized in the form of lariat without tail. Interestingly, lariat 14c level was downregulated in FUS KO cells and upregulated in FUS OE cells, thus its formation could be induced by FUS (Fig. 6E upper panel). In regards to the function of lariat formation, we assume that it might block mature snoRNA production or protect snoRNA from further processing into sdRNA (see Discussion).

DISSCUSSION

As it was suggested by Bratkovic and Rogelj, sdRNAs are not simply degradation products of snoRNAs for at least three reasons: i) common processing patterns have been observed, ii) the extent of processing of different snoRNAs may differ in the same cell line, and iii) posttranscriptional gene regulation activity was confirmed for some sdRNAs [4]. In this paper we show for the first time that in human cells FUS can mediate the biogenesis of sdRNA, that can be further involved in the regulation of gene expression at different level.

The mechanism of the sdRNAs production is still not fully elucidated. According to our preliminary results trimming of selected snoRNAs into sdRNAs is mediated by FUS and this process involves mature snoRNAs. Mature snoRNAs are processed from the removed and debranched introns by exonucleolytic activities, and we observed no effect of FUS on either the splicing profile of their host genes, or the level of snoRNA precursors (Fig. 2B, C). This suggests that FUS has influence on the snoRNAs level at later steps of snoRNPs biogenesis.

The assembly of box C/D and H/ACA snoRNPs is a multistep mechanism, that requires both core proteins and assembly factors [6, 7]. During the biogenesis of box C/D snoRNPs, a protein-only complex containing SNU13 and NOP58, in association with assembly factors: AAAC ATPase RUVBL1/2, NUFIP, ZNHIT3 and ZNHIT6, is pre-formed with the help of HSP90/R2TP complex, and loaded on snoRNA in a co-splicing dependent manner. Pre-snoRNP particles are then transported to Cajal Bodies (CBs) where final processing occurs, and then catalytically active snoRNPs are transported to the nucleoli. However, whether assembly factors leave pre-snoRNP before or after arriving in CBs, is not clear. During the assembly of box H/ACA snoRNPs, two conserved proteins, NAF1 and SHQ1, are required for the stability of box H/ACA snoRNAs, without being part of the mature particles. At a final step of biogenesis, NAF1 is replaced by GAR1, that leads to production of mature and functional H/ACA snoRNPs [6, 7]. In mammalian cells, the SMN complex has been suggested to play a role in this exchange [59]. We assume that FUS can interact with any of these proteins to impede the complete snoRNP assembly. Alternatively, by such interaction, FUS can recruit or facilitate the action of other proteins involved in exo- or endonucleolytic cleavage of snoRNAs into smaller fragments. As a third possibility, FUS can induce sdRNAs production at the stage of snoRNP disassembly. Nevertheless, in both cases, whether FUS impedes the snoRNP assembly or induces snoRNP disassembly, the sdRNAs production leads to decreased level of mature snoRNAs/snoRNPs in the cells.

In an alternative hypothesis, snoRNAs can be trapped by circular lariat introns that are derived from spliced introns, which escape debranching and are protected from further processing. Formation of such lariat introns containing snoRNAs could be mediated by FUS. Also in this case, the interaction between FUS and snoRNP proteins might be necessary. It cannot be excluded that this step is required to generate sdRNA. Alternatively, FUS can induce generation of lariat structures to trap snoRNAs and block their maturation or processing. This could act as another mechanism for regulation of snoRNA/sdRNA levels, similar to the pathways described in *Arabidopsis thaliana,* where intron lariat RNAs inhibits miRNA biogenesis [60].

In our RIP-seq experiment we also identified small fraction of tRFs (tRNA-derived fragments), so the possibility that FUS also induces the synthesis of other small regulatory RNAs cannot be excluded. The biogenesis of most of the tRFs is also not fully described. It is suggested that the role of Dicer, DGCR8 or Drosha is an exception, rather than the rule, however, enzymes involved in the liberation of tRFs from the 5' and 3' end of mature tRNAs are still not known [61]. Thus, the tRFs and sdRNAs generation might be similar and they both might involve FUS protein. A deeper understanding of the molecular mechanism of the sdRNAs biogenesis might therefore shed a light on the processing of tRFs, and other small RNAs mediated by FUS as well.

Interestingly, among FUS-dependent-sdRNA targets that we identified, there are transcripts highly expressed in brain (The Human Protein Atlas; http://www.proteinatlas.org). For example, BRE (Brain and Reproductive organ-Expressed protein) functions in DNA damage repair as well as in anti-apoptosis as a component of the BRCA1-A complex. BRE is supposed to play a role in homeostasis or cellular differentiation in cells of neural, epithelial and germline origins. Recent research shed new light on the importance of snoRNA fragments in cancer tissues. Bangma and Jenster group observed that sdRNAs display strong differential expression, and are massively upregulated in prostate cancer [62]. It is also worth to underline that BRE and KCNQ10T1 are disease-related genes. Marked overexpression of BRE was detected in many tumors, suggesting that it promotes cells proliferation and local tumor growth [63, 64]. Epigenetic status of KCNQ10T1 was shown to be correlated with corectal carcinogenesis [65]. In summary, a link between the sdRNAs and human diseases exists. On the other hand, FUS is also implicated in neurodegeneration, tumors and cellular stress response through errors in multiple steps of RNA processing. Therefore, our preliminary results showing connection between FUS-dependent sdRNAs and disease-related genes are promising and might bring closer the role of FUS and FUS-dependent sdRNA in oncogenic networks of selected tumors.

ACKNOWLEDGEMENT

We would like to acknowledge Susheel Bhat for his input during manuscript preparation. The research and related results of M.-D.R. were made possible by the support of the NOMIS Foundation.

FUNDING

This work was supported by: KNOW RNA Research Centre in Poznan under Grant 01/KNOW2/2014 (to PP, MS, AS, MZ, AP, IM, KDR).

All authors have seen and approved the manuscript, and it hasn't been accepted or published elsewhere.

REFERENCES

1. Filipowicz, W. and V. Pogacic, *Biogenesis of small nucleolar ribonucleoproteins.* Curr Opin Cell Biol, 2002. **14**(3): p. 319-27.
2. Kiss, T., et al., *Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs.* Cold Spring Harb Symp Quant Biol, 2006. **71**: p. 407-17.
3. Reichow, S.L., et al., *The structure and function of small nucleolar ribonucleoproteins.* Nucleic Acids Res, 2007. **35**(5): p. 1452-64.
4. Bratkovic, T. and B. Rogelj, *The many faces of small nucleolar RNAs.* Biochim Biophys Acta, 2014. **1839**(6): p. 438-43.
5. Watkins, N.J., A. Dickmanns, and R. Luhrmann, *Conserved Stem II of the Box C/D Motif Is Essential for Nucleolar Localization and Is Required, Along with the 15.5K Protein, for the Hierarchical Assembly of the Box C/D snoRNP.* Molecular and Cellular Biology, 2002. **22**(23): p. 8342-8352.
6. Bizarro, J., et al., *Proteomic and 3D structure analyses highlight the C/D box snoRNP assembly mechanism and its control.* J Cell Biol, 2014. **207**(4): p. 463-80.
7. Massenet, S., E. Bertrand, and C. Verheggen, *Assembly and trafficking of box C/D and H/ACA snoRNPs.* RNA Biol, 2017. **14**(6): p. 680-692.

8. Verheggen, C., et al., *Box C/D small nucleolar RNA trafficking involves small nucleolar RNP proteins, nucleolar factors and a novel nuclear domain.* EMBO J, 2001. **20**(19): p. 5480-90.
9. Tycowski, K.T., M.D. Shu, and J.A. Steitz, *A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3.* Genes Dev, 1993. **7**(7A): p. 1176-90.
10. Hirose, T., M.-D. Shu, and J.A. Steitz, *Splicing-Dependent and -Independent Modes of Assembly for Intron-Encoded Box C/D snoRNPs in Mammalian Cells.* Molecular Cell, 2003. **12**(1): p. 113-123.
11. Richard, P., et al., *Cotranscriptional recognition of human intronic box H/ACA snoRNAs occurs in a splicing-independent manner.* Mol Cell Biol, 2006. **26**(7): p. 2540-9.
12. Zhang, Y., et al., *Circular intronic long noncoding RNAs.* Mol Cell, 2013. **51**(6): p. 792-806.
13. Yin, Q.F., et al., *Long noncoding RNAs with snoRNA ends.* Mol Cell, 2012. **48**(2): p. 219-30.
14. Talhouarne, G.J. and J.G. Gall, *Lariat intronic RNAs in the cytoplasm of Xenopus tropicalis oocytes.* RNA, 2014. **20**(9): p. 1476-87.
15. Mleczko, A.M. and K. Bakowska-Zywicka, *When small RNAs become smaller: emerging functions of snoRNAs and their derivatives.* Acta Biochim Pol, 2016. **63**(4): p. 601-607.
16. Kawaji, H., et al., *Hidden layers of human small RNAs.* BMC Genomics, 2008. **9**: p. 157.
17. Ender, C., et al., *A human snoRNA with microRNA-like functions.* Mol Cell, 2008. **32**(4): p. 519-28.
18. Taft, R.J., et al., *Small RNAs derived from snoRNAs.* RNA, 2009. **15**(7): p. 1233-40.
19. Brameier, M., et al., *Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs.* Nucleic Acids Res, 2011. **39**(2): p. 675-86.
20. Langenberger, D., et al., *Identification and classification of small RNAs in transcriptome sequence data.* Pac Symp Biocomput, 2010: p. 80-7.
21. Ono, M., et al., *Identification of human miRNA precursors that resemble box C/D snoRNAs.* Nucleic Acids Res, 2011. **39**(9): p. 3879-91.
22. Scott, M.S., et al., *Human miRNA precursors with box H/ACA snoRNA features.* PLoS Comput Biol, 2009. **5**(9): p. e1000507.
23. Politz, J.C., E.M. Hogan, and T. Pederson, *MicroRNAs with a nucleolar location.* RNA, 2009. **15**(9): p. 1705-15.
24. Saraiya, A.A. and C.C. Wang, *snoRNA, a novel precursor of microRNA in Giardia lamblia.* PLoS Pathog, 2008. **4**(11): p. e1000224.
25. Hutzinger, R., et al., *Expression and processing of a small nucleolar RNA from the Epstein-Barr virus genome.* PLoS Pathog, 2009. **5**(8): p. e1000547.
26. Langenberger, D., et al., *Dicer-processed small RNAs: rules and exceptions.* J Exp Zool B Mol Dev Evol, 2013. **320**(1): p. 35-46.
27. Kishore, S., et al., *The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing.* Hum Mol Genet, 2010. **19**(7): p. 1153-64.
28. Li, Z., et al., *Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs.* Nucleic Acids Res, 2012. **40**(14): p. 6787-99.
29. Kishore, S., et al., *Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing.* Genome Biol, 2013. **14**(5): p. R45.
30. Falaleeva, M. and S. Stamm, *Processing of snoRNAs as a new source of regulatory non-coding RNAs: snoRNA fragments form a new class of functional RNAs.* Bioessays, 2013. **35**(1): p. 46-54.
31. Scott, M.S., et al., *Human box C/D snoRNA processing conservation across multiple cell types.* Nucleic Acids Res, 2012. **40**(8): p. 3676-88.
32. Zywicki, M., K. Bakowska-Zywicka, and N. Polacek, *Revealing stable processing products from ribosome-associated small RNAs by deep-sequencing data analysis.* Nucleic Acids Res, 2012. **40**(9): p. 4013-24.
33. Tan, A.Y. and J.L. Manley, *The TET family of proteins: functions and roles in disease.* J Mol Cell Biol, 2009. **1**(2): p. 82-92.

34. Zinszner, H., et al., *TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling.* J Cell Sci, 1997. **110 (Pt 15)**: p. 1741-50.
35. Perrotti, D., et al., *TLS/FUS, a pro-oncogene involved in multiple chromosomal translocations, is a novel regulator of BCR/ABL-mediated leukemogenesis.* EMBO J, 1998. **17**(15): p. 4442-55.
36. Baechtold, H., et al., *Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein TLS/FUS and is able to promote D-loop formation.* J Biol Chem, 1999. **274**(48): p. 34337-42.
37. Gardiner, M., et al., *Identification and characterization of FUS/TLS as a new target of ATM.* Biochem J, 2008. **415**(2): p. 297-307.
38. Yang, S., et al., *Fused in sarcoma/translocated in liposarcoma: a multifunctional DNA/RNA binding protein.* Int J Biochem Cell Biol, 2010. **42**(9): p. 1408-11.
39. Sama, R.R., C.L. Ward, and D.A. Bosco, *Functions of FUS/TLS from DNA repair to stress response: implications for ALS.* ASN Neuro, 2014. **6**(4).
40. Raczynska, K.D., et al., *FUS/TLS contributes to replication-dependent histone gene expression by interaction with U7 snRNPs and histone-specific transcription factors.* Nucleic Acids Res, 2015. **43**(20): p. 9711-28.
41. Gregory, R.I., et al., *The Microprocessor complex mediates the genesis of microRNAs.* Nature, 2004. **432**(7014): p. 235-240.
42. Kwiatkowski, T.J., Jr., et al., *Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis.* Science, 2009. **323**(5918): p. 1205-8.
43. Vance, C., et al., *Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6.* Science, 2009. **323**(5918): p. 1208-1211.
44. Dormann, D. and C. Haass, *Fused in sarcoma (FUS): an oncogene goes awry in neurodegeneration.* Mol Cell Neurosci, 2013. **56**: p. 475-86.
45. Lagier-Tourenne, C., M. Polymenidou, and D.W. Cleveland, *TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration.* Hum Mol Genet, 2010. **19**(R1): p. R46-64.
46. Reber, S., et al., *CRISPR-Trap: a clean approach for the generation of gene knockouts and gene replacements in human cells.* Mol Biol Cell, 2018. **29**(2): p. 75-83.
47. Pall, G.S. and A.J. Hamilton, *Improved northern blot method for enhanced detection of small RNA.* Nat Protoc, 2008. **3**(6): p. 1077-84.
48. Lerga, A., et al., *Identification of an RNA binding specificity for the potential splicing factor TLS.* J Biol Chem, 2001. **276**(9): p. 6807-16.
49. Romeo, V. and D. Schumperli, *Cycling in the nucleus: regulation of RNA 3' processing and nuclear organization of replication-dependent histone genes.* Curr Opin Cell Biol, 2016. **40**: p. 23-31.
50. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.
51. Coordinators, N.R., *Database resources of the National Center for Biotechnology Information.* Nucleic Acids Res, 2016. **44**(D1): p. D7-19.
52. Aken, B.L., et al., *The Ensembl gene annotation system.* Database (Oxford), 2016. **2016**.
53. Betel, D., et al., *Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites.* Genome Biol, 2010. **11**(8): p. R90.
54. Kamburov, A., et al., *The ConsensusPathDB interaction database: 2013 update.* Nucleic Acids Res, 2013. **41**(Database issue): p. D793-800.
55. Li, J.H., et al., *starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data.* Nucleic Acids Res, 2014. **42**(Database issue): p. D92-7.
56. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-2.
57. Reber, S., et al., *Minor intron splicing is regulated by FUS and affected by ALS-associated FUS mutants.* EMBO J, 2016. **35**(14): p. 1504-21.
58. Pasquinelli, A.E., *MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship.* Nat Rev Genet, 2012. **13**(4): p. 271-82.

59. Pellizzoni, L., et al., *The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1.* Curr Biol, 2001. **11**(14): p. 1079-88.
60. Li, Z., et al., *Intron Lariat RNA Inhibits MicroRNA Biogenesis by Sequestering the Dicing Complex in Arabidopsis.* PLoS Genet, 2016. **12**(11): p. e1006422.
61. Kumar, P., C. Kuscu, and A. Dutta, *Biogenesis and Function of Transfer RNA-Related Fragments (tRFs).* Trends Biochem Sci, 2016. **41**(8): p. 679-689.
62. Martens-Uzunova, E.S., et al., *C/D-box snoRNA-derived RNA production is associated with malignant transformation and metastatic progression in prostate cancer.* Oncotarget, 2015. **6**(19): p. 17430-44.
63. Chan, B.C., et al., *BRE enhances in vivo growth of tumor cells.* Biochem Biophys Res Commun, 2005. **326**(2): p. 268-73.
64. Chui, Y.L., et al., *BRE over-expression promotes growth of hepatocellular carcinoma.* Biochem Biophys Res Commun, 2010. **391**(3): p. 1522-5.
65. Nakano, S., et al., *Expression profile of LIT1/KCNQ1OT1 and epigenetic status at the KvDMR1 in colorectal cancers.* Cancer Sci, 2006. **97**(11): p. 1147-54.

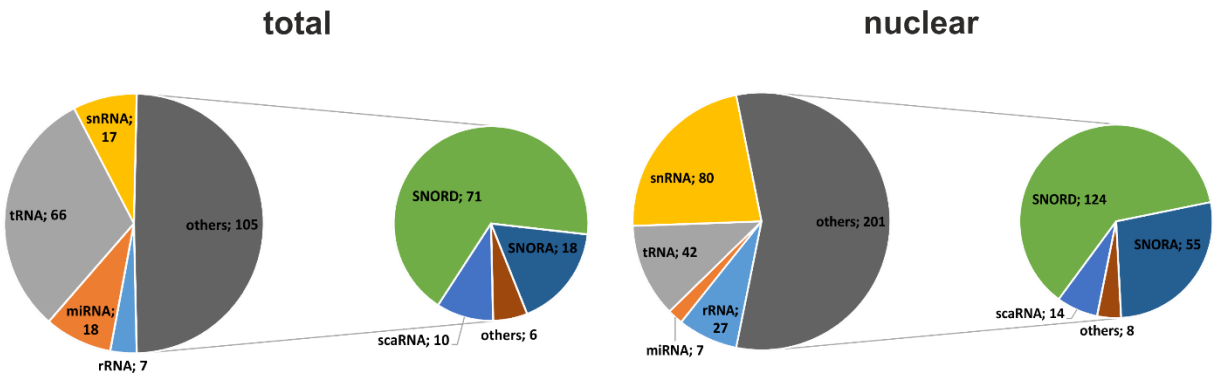


Figure 1. Distribution of small noncoding RNAs bound to FUS identified in fractions immunoprecipitated from total and nuclear extracts. The numbers represent snRNAs that were ≥4 fold enriched in comparison to control sample.

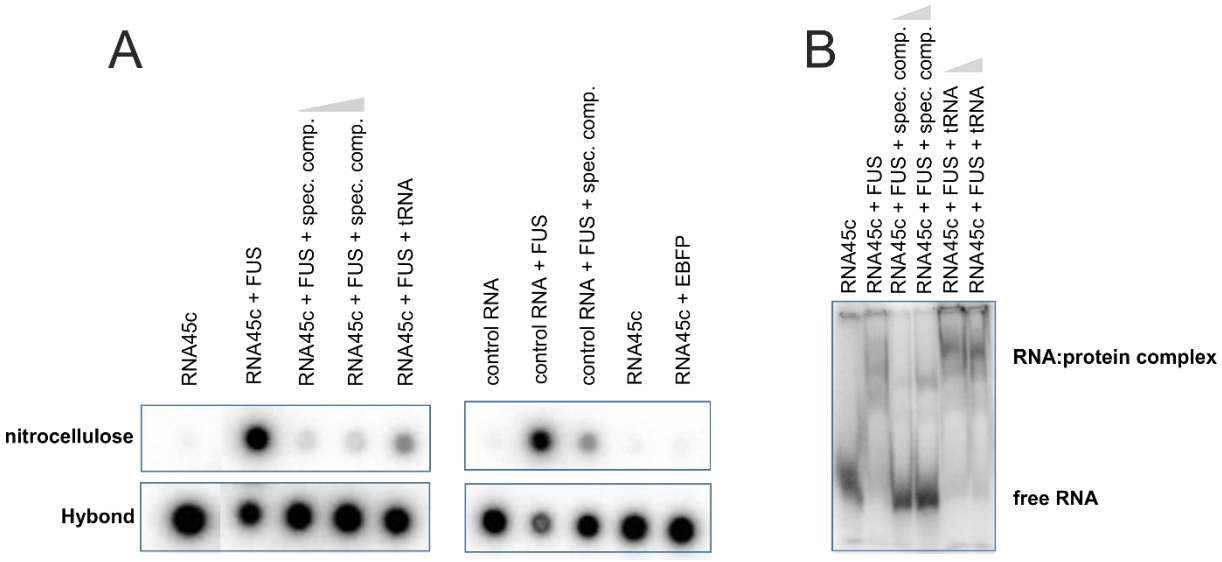
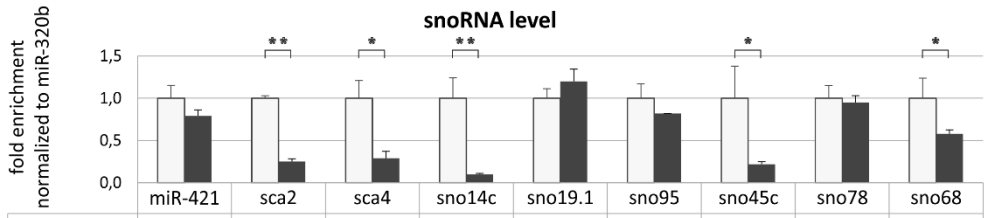
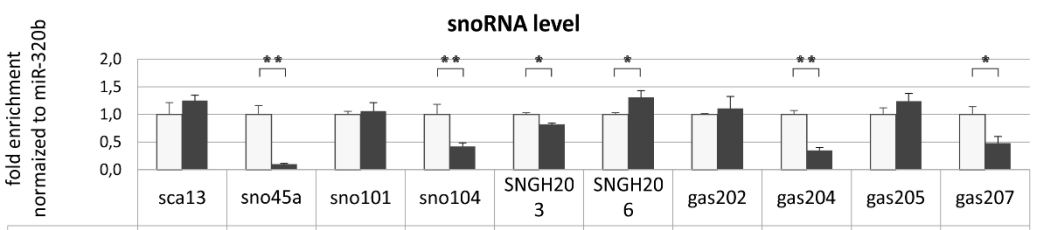


Figure 2. Double filter binding assay (A) and EMSA assay (B) confirmed that FUS can directly interact with snoRNAs. RNA fragment encompassing C/D motifs of SNORD45c (RNA45c) and positive control sequence (control RNA) were subjected to interaction with purified recombinant FLAG-FUS and FLAG-EBFP (as negative control). The binding specificity was tested by adding specific (unlabeled RNA fragment encompassing C/D motifs of SNORD45c) and nonspecific (tRNAs) competitors.

A

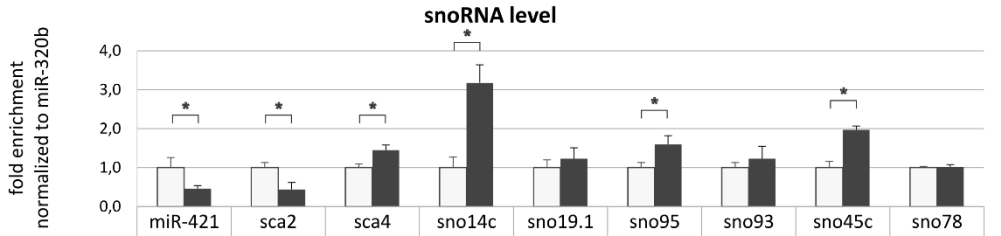


	miR-421	sca2	sca4	sno14c	sno19.1	sno95	sno45c	sno78	sno68
control cells	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
FUS OE	0,79	0,25	0,29	0,10	1,20	0,82	0,22	0,95	0,58

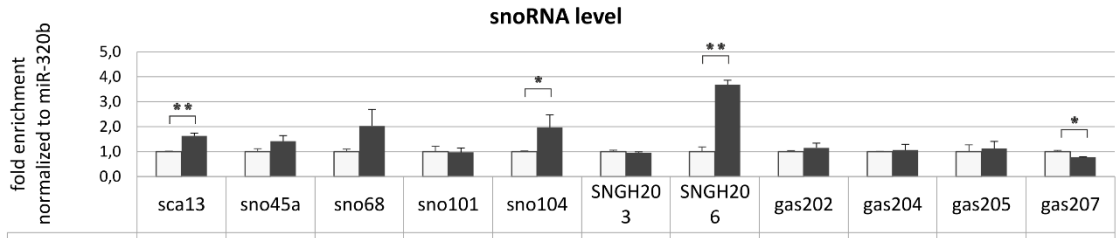


	sca13	sno45a	sno101	sno104	SNGH20 3	SNGH20 6	gas202	gas204	gas205	gas207
control cells	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
FUS OE	1,25	0,10	1,06	0,42	0,82	1,31	1,11	0,35	1,24	0,48

B

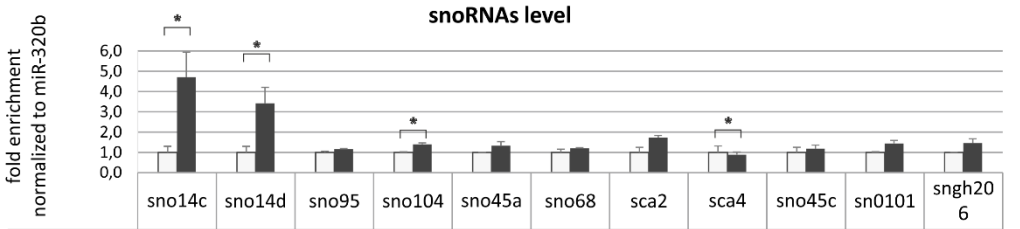


	miR-421	sca2	sca4	sno14c	sno19.1	sno95	sno93	sno45c	sno78
control cells	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
FUS KD	0,46	0,44	1,45	3,18	1,23	1,60	1,23	1,97	1,02



	sca13	sno45a	sno68	sno101	sno104	SNGH20 3	SNGH20 6	gas202	gas204	gas205	gas207
control cells	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
FUS KD	1,63	1,42	2,03	0,97	1,97	0,96	3,68	1,15	1,06	1,12	0,77

C



	sno14c	sno14d	sno95	sno104	sno45a	sno68	sca2	sca4	sno45c	sn0101	sngh20 6
control cells	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
FUS KO	4,70	3,42	1,16	1,40	1,33	1,20	1,74	0,89	1,18	1,44	1,46

Figure 3.The effect of FUS overexpression (FUS OE) (A), and FUS depletion (FUS KD) (B), (FUS KO) (C), on the level of selected mature snoRNAs was tested by RT-qPCR. Error bars indicate standard deviations (SD) of three biological replicates. *P*-values were calculated using Student's T-test, and statistical significance is represented as follows: *$p \leq 0.05$; **$p \leq 0.01$.

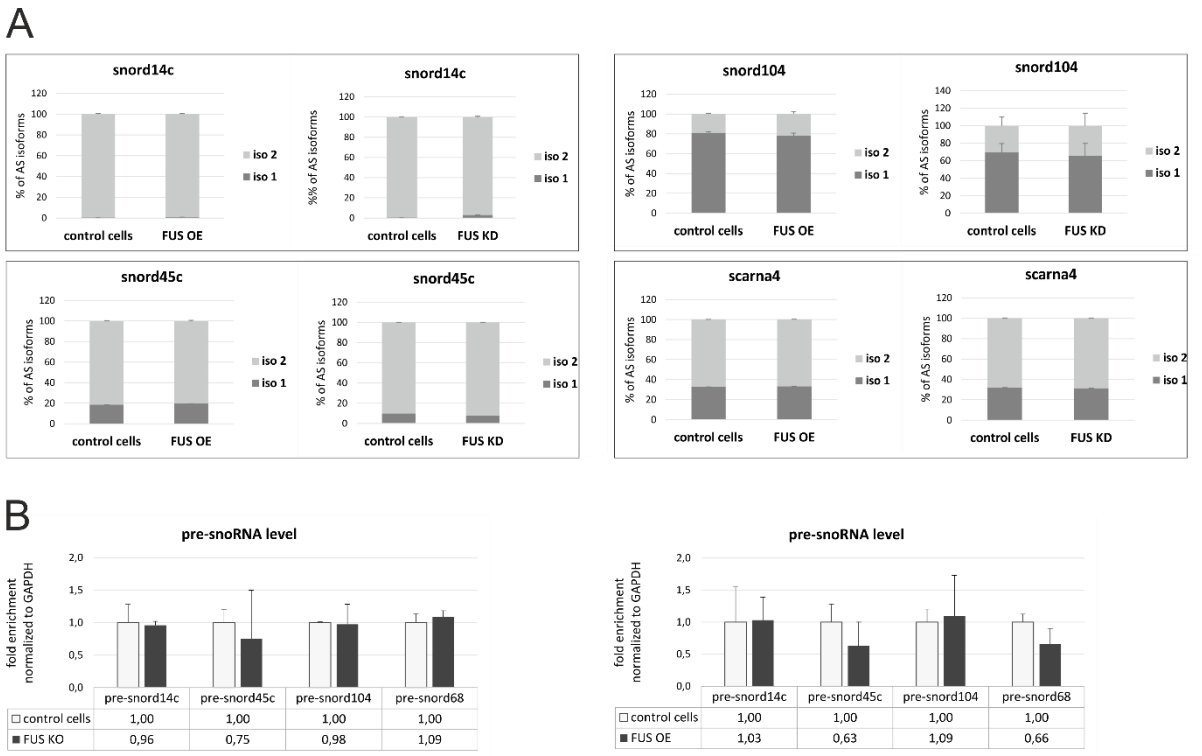


Figure 4. A) The effect of FUS overexpression (FUS OE) and FUS depletion (FUS KO) on splicing of introns encoding snoRNAs genes was analysed by RT-qPCR. Primers were designed to amplify alternative isoforms (iso1 (spliced isoform) and iso2 (unspliced isoform)). B) The effect of FUS overexpression (FUS OE) and FUS depletion (FUS KO) on the level of snoRNAs precursors was analysed by RT-qPCR.

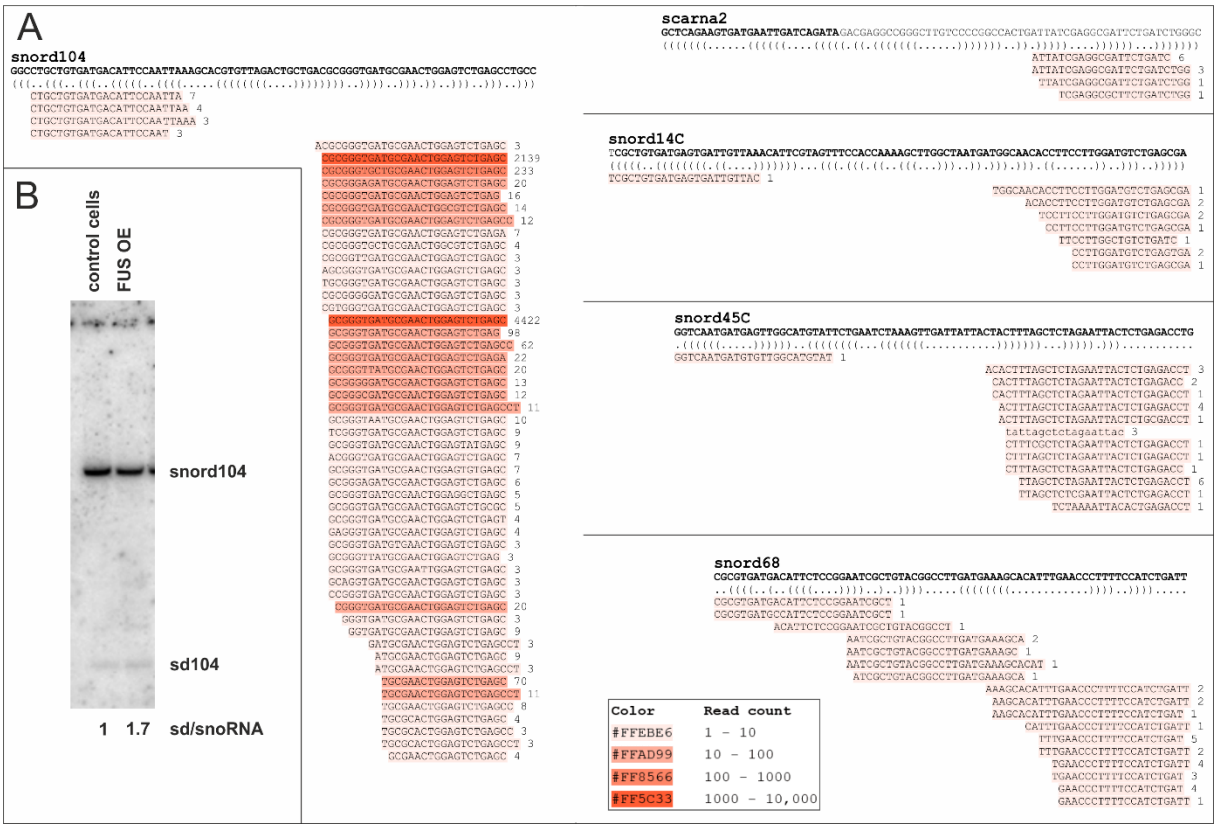


Figure 5. A) Scanning of human small RNA databases (derived from cultured HEK293 cells, human neuroblastoma cells and whole blood) revealed existence of sdRNAs with the length of 19-33 nt, which can be derived from selected snoRNAs. Numbers on the right indicates numbers of reads. B) Northern blot hybridization confirmed the existence of sdRNA104 in cells. The ratio of sd104/sno104 was calculated by dividing hybridization signal intensity of sdRNA by snoRNA. Secondary structures were predicted with RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi): minimum free energy prediction, optimal secondary structure.

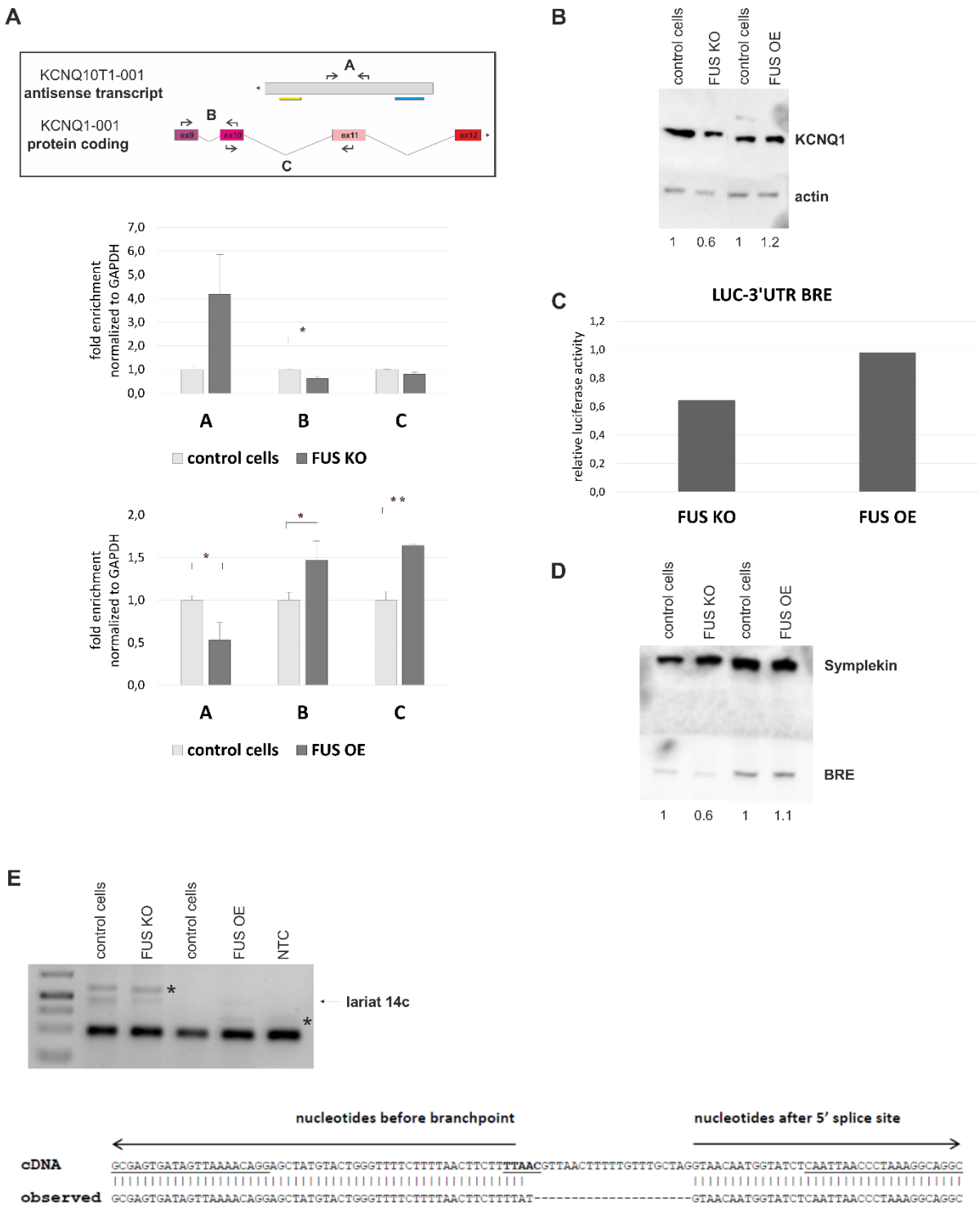


*Figure 6. A) FUS-dependent sdRNAs (indicated by yellow and blue line) can target KCNQ10T1 antisense transcripts. The sdRNA:target duplex is formed within a unique region that distinguishes noncoding RNA from mRNA transcribed from the same genomic region. RT-qPCR analysis was done using a set of primers designed as shown in the scheme. Error bars represent the SD of three biological replicates. P-values were calculated using Student's T-test, and the statistical significance is represented as follows: *P ≤ 0.05; **P ≤ 0.01. B) Western blot followed by immunodetection with anti-KCNQ1 and anti-actin (loading control) antibodies was performed using protein extract isolated from control cells and cells with FUS overexpression (FUS OE) or FUS knockout (FUS KO). Relative amount of FUS protein is listed below. C) Luciferase assays confirmed positive effect of sdRNAs on mRNA translation. 3'UTR of BRE*

mRNA targeted by FUS-dependent 104sdRNA was cloned downstream of Luciferase coding sequence and its effect on translation was tested in FUS KO and FUS OE cells. D) Western blot followed by immunodetection with anti-BRE and anti-Symplekin (loading control) antibodies were performed using protein extract isolated from control cells and FUS OE and FUS KO cells. Relative amount of FUS protein is listed below. E) (upper panel) RT-PCR using outward-facing primers was performed to amplify lariat 14c in FUS KO, FUS OE, control cells and non-template control (NTC). (Lower panel) Sequencing of lariat 14 product revealed 14-nucleotides gap between branchpoint and 5'splice site. Sequence of predicted branchpoint is in bold and underlined, sequence of primers is also underlined.