The RNA-binding protein AATF coordinates rRNA maturation

Rainer W. Kaiser[1]*, Michael Ignarski[1]*, Eric L. Van Nostrand[2], Christian Frese[3], Manaswita Jain[1], Sadrija Cukoski[1], Heide Heinen[1], Melanie Schaechter[1], Konstantin Bunte[1,4], Peter Frommolt[4], Patrick Keller[5], Mark Helm[5], Katrin Bohl[1], Martin Höhne[1,8,9], Bernhard Schermer[1,8,9], Thomas Benzing[1,8,9], Katja Höpker[1], Christoph Dieterich[6,7], Gene W. Yeo[2], Roman-Ulrich Müller[1,8,9 #] and Francesca Fabretti[1]

[1] Department II of Internal Medicine and Center for Molecular Medicine Cologne, University of Cologne, 50937 Cologne, Germany.

[2] Department of Cellular and Molecular Medicine, University of California at San Diego, La Jolla, CA, USA; Institute for Genomic Medicine, University of California at San Diego, La Jolla, CA, USA.

[3] Proteomics Core Facility, Cologne Excellence Cluster on Cellular Stress Responses in Aging-associated Diseases (CECAD), University of Cologne, 50931 Cologne, Germany.

[4] Bioinformatics Core Facility, Cologne Excellence Cluster on Cellular Stress Responses in Aging-associated Diseases (CECAD), University of Cologne, 50931 Cologne, Germany.

[5] Institute of Pharmacy and Biochemistry, Johannes Gutenberg-University Mainz, Staudingerweg 5, 55128, Mainz, Germany.

[6] German Center for Cardiovascular Research (DZHK), Partner site Heidelberg/Mannheim, Im Neuenheimer Feld 669, 69120, Heidelberg, Germany

[7] Section of Bioinformatics and Systems Cardiology, Klaus Tschira Institute for Integrative Computational Cardiology and Department of Internal Medicine III, Im Neuenheimer Feld 669, 69120, Heidelberg, Germany

[8] Cologne Excellence Cluster on Cellular Stress Responses in Aging-associated Diseases (CECAD), University of Cologne, 50931 Cologne, Germany.

[9] Systems Biology of Ageing Cologne, University of Cologne, 50931 Cologne, Germany.

[#] To whom correspondence should be addressed: Roman-Ulrich Müller, Tel: 0221-478-86288, FAX: 0221-478-1423108; Email: roman-ulrich.mueller@uk-koeln.de.

*The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

ABSTRACT

Apoptosis-antagonizing transcription factor (AATF) is a predominantly nuclear protein essential for both embryonic development and tumor growth. Several studies have shown a role of AATF in the modulation of cellular signal transduction pathways such as p53-, mTOR- and HIF-signaling. However, the exact molecular functions underlying its essential nature to cell proliferation and survival have remained elusive. Interestingly, several lines of evidence point towards a pivotal role of this protein in ribosome biogenesis and the maturation of ribosomal RNA. In this study, we identify AATF in a screen for RNA-binding proteins. Importantly, CLIP-sequencing shows a predominant association with ribosomal RNA precursor molecules. Furthermore, AATF binds to mRNAs encoding for ribosome biogenesis factors as well as snoRNAs. These findings are complemented by an in-depth characterization of the protein interactome of AATF, again containing a large set of proteins known to play a role in rRNA maturation. Consequently, our multilayer analysis of the protein-RNA interactome of AATF reveals this protein to be a central hub in the coordination of ribosome biogenesis.

INTRODUCTION

AATF, also known as CHE-1, was originally identified as an RNA polymerase II interacting protein with anti-apoptotic capacities. Through its interaction with RNAPII, AATF modulates the function of a row of transcription factors including pRB, p65 and STAT3 (1-3). Taking into account its role in the prevention of apoptosis the involvement of AATF in the DNA-damage response and its impact on p53 function are explicitly interesting. The functional impact of these findings is emphasized by data showing AATF as a pro-tumorigenic factor in several tumor models (4,5). Upon DNA damage AATF gets phosphorylated by checkpoint kinases leading to its interaction with the NF-kB subunit p65 and relocalization to the p53 promoter (3,6,7). However, AATF also modulates the specificity of p53 by shifting its binding preference towards target genes leading to growth arrest over those that mediate apoptosis (5,8). Furthermore, AATF has been shown to play a role in other key pathways involved in tumorigenesis, namely mTOR- and HIF-signaling (9,10). Seeing the impact of this protein on signal transduction in apoptosis and tumor formation, its potential as a therapeutic target in cancer therapy has been discussed extensively (5,11,12). Interestingly, AATF – whilst also having been detected in cyto- and nucleoplasm – primarily appears to localize to the nucleolus with the nucleolar fraction having been associated with c-JUN-mediated apoptosis (13). With nucleoli being the sites of rRNA transcription and maturation, it is important to note that AATF was recently identified by independent RNAi screens for factors involved in ribosomal subunit production (14,15). This connection is extremely interesting due to two considerations. Firstly, cell proliferation and ribosome biogenesis rate are closely intertwined, and increased ribosome formation is linked to tumorigenesis (16,17). Secondly, blocking ribosome formation induces ribosome biogenesis stress leading to the activation of p53 mediated by inhibition of MDM2. Vice versa, induction of ribosome biogenesis inhibits p53 (16,18,19). However, the mode of action of AATF in ribosome maturation has remained unclear. A first insight came recently from a study by Bammert et al. (20) that could show AATF as part of a

nucleolar protein complex (termed ANN complex). Here, AATF, together with NOL10 and NGDN, was essential to efficient generation of the small ribosomal subunit (SSU) - in line with findings from previous screens of ribosome biogenesis factors (14). Yet, the molecular function that AATF fulfils within this complex and the question which one of the three proteins mediates binding to rRNA precursors remained elusive. Work published recently by Piñeiro et al. (21) shows AATF to be amongst 211 RNA polymerase I-dependent RNA-binding proteins and to bind to pre-rRNA molecules. In line with their findings, we identified AATF to be an RNA-binding protein. Together, their and our data reveal AATF to be the protein mediating RNA-binding in the ANN-complex. We complement this finding with the identification of AATF binding sites in the 45S rRNA precursor and an extensive characterization of its protein and RNA interaction partners placing AATF at a central position in rRNA maturation and ribosome biogenesis.

MATERIAL AND METHODS

Molecular cloning and design of small interfering RNAs (siRNA)

For the generation of GFP-tagged AATF transgenes, the AATF wild type sequence, truncated or mutated versions of the protein generated by overlap extension PCR were cloned into the AAV CAGGS GFP plasmid (Addgene #22212). siRNAs targeting the 3' UTR of human AATF were custom designed by Dharmacon (accession number AJ249940.2).

Cell culture, transfection and generation of single-copy transgenic cell lines using TALEN

Human HEK 293T, HCT116 and U2OS cells were grown in standard media at 37°C, 5% CO_2 and routinely passaged using 0.05% Trypsin. Mycoplasma contamination was checked using a commercial kit (Venor GeM, Sigma). Transfections were carried out on 60-80% confluent cells using calcium phosphate as described previously (22), or lipofection (Lipofectamine 2000 and Lipofectamine LTX) and electroporation (Amaxa Nucleofector® kit V) according to the manufacturers' instructions. For knockdown experiments using commercial siRNA pools (Dharmacon), cell lines were transfected with Lipofectamine RNAiMAX and incubated for 48 h (final concentration 20 nM).

Stably integrated transgenic cell lines were generated using TALEN technology by co-transfecting TALEN encoding plasmids specific for the human AAVS1 locus as previously described (23). 24h after transfection, cell lines were steadily selected with 2 µM Puromycin. All cell lines were genotyped by integration PCR and phenotyped by both immunoblot and immunofluorescence.

Western blot

Cells were washed once with ice-cold PBS, lysed on ice in either iCLIP lysis buffer or modified RIPA buffer for 15 minutes and then sonicated (BioRuptor Pico). Lysates were then run on home-made PAGE gels, transferred onto PVDF membranes and blocked in 5% bovine serum albumin. Primary

antibodies were incubated either for 1h at room temperature or overnight at 4°C. Secondary HRP coupled antibodies (Jackson ImmunoResearch,) were incubated for 1h at room temperature. Bands were visualized using home-made ECL, at a Fusion chemoluminescence imaging system (Peqlab) (see Suppl. Methods for a list of primary antibodies used).

Immunofluorescence/Microscopy

U2OS and HTC116 cell lines were grown on coverslips coated with 100 µg/ml Collagen I to 60-70% confluency. Cells were washed once with PBS supplemented with Ca^{2+}/Mg^{2+} (PBS+), fixed with ice-cold Methanol at -20°C for 5 minutes and rinsed three times with PBS+. Cells were then blocked in 5% donkey serum and 0.1% Triton X-100 in PBS at RT and incubated in primary antibody at 4°C overnight. Secondary antibody (Cy3-conjugated, Jackson ImmunoResearch) was added for one hour at RT, after which the stained cells were mounted on ProLong Gold mounting medium with DAPI (Invitrogen) to visualize nuclei. Images were acquired using a Zeiss epifluorescence microscope (Zeiss Axiovert 200M) as well as Zeiss ZEN software and analyzed using ImageJ/Fiji (24).

Polynucleotide Kinase Assay

HEK 293T cells transfected with pcDNA6 expressing triple FLAG tagged genes of interest were UV-C crosslinked, resuspended in lysis buffer (100 mM KCL, 5 mM $MgCl_2$, 10 mM Tris pH 7.5, 0.5% NP40, 1 mM DTT, protease inhibitors), homogenized through 23G needle on ice and sonicated (Biorupter Pico, 10 intervals: 30 s on/off). Lysates were cleared by centrifugation (20000xg), treated with 2 U/ml Turbo DNase (Invitrogen) and 40 U/ml RNase I (Ambion) for 15 min at 37 °C. Immunoprecipitation was performed with anti-FLAG M2 (Sigma) coupled to Protein G Dynabeads (ThermoFisher) for 2 h at 4 °C. Beads were washed 5 × with washing buffer (500 mM NaCl, 20 mM Tris pH 7.5, 1 mM $MgCl_2$, 0.05% NP40) and twice with PNK-buffer (50 mM Tris pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$, 0.5% NP40). Beads were resuspended in PNK buffer containing 5 mM DTT, 0.2 µCi/µl [γ-^{32}P]-ATP (Hartmann-Analytic) and 1 U/µl T4 PNK (ThermoFisher). The labeling was carried out for 20 min at 37 °C. Beads were washed 5 x with PNK buffer and boiled with Laemmli buffer. The samples were resolved on 4-12% bis-tris gels (ThermoFisher) and transferred onto Protran Nitrocellulose (Schleicher and Schuell). The blot was exposed to storage phosphor screen (Amersham) and the signal was detected with Typhoon scanner (GE Healthcare). IP efficiency was controlled by western blotting with anti-FLAG M2 (Sigma).

eCLIP-seq, Read Processing and Cluster Analysis

eCLIP of AATF in K562 and HepG2 cells using primary anti-AATF antibody (A301-032A lot 001, Bethyl) was performed as previously described (25). In addition to standard read processing and processing of reads to identify unique genomic mapping, reads mapping ribosomal RNA were quantified using a family-aware repeat element mapping pipeline that identifies reads unique to 45S pre-rRNA, 18S rRNA, or 28S rRNA respectively (Van Nostrand, E.L., *et al. in preparation*). To quantify relative enrichment between IP and input, relative information was calculated as the Kullback-Leibler

divergence: $p_i \times log_2(\frac{p_i}{q_i})$, where p_i is the fraction of total reads in IP that map to position *i*, and q_i is the fraction of total reads in input for the same position. Regarding the definition of non-rRNA interaction partners we filtered the total dataset (K562 and HePG2 cells, 2 replicates each) for significant peaks over input (log_2FC ≥3 and $-log_{10}$ p-value ≥5) and collapsed all peaks mapping to one transcript to define a list of targets. For AATF eCLIP data accessibility refer to "Availability" section below.

Coimmunoprecipitation and sample preparation for MS/MS

For each replicate seven 10 cm dishes (80% confluency) of HEK 293T Flp-In™ T-REx™ expressing FLAG/HA-tagged AATF or GFP were used. Cells were harvested in media and washed with ice-cold PBS. The harvested cells were lysed in modified RIPA buffer containing 1% NP-40 (IgPAL), 150 mM NaCl, 0.25% sodium-deoxycholate and 50 mM Tris and complete protease inhibitors (PIM; Roche). Cells were sonicated for 45s at 10% power using an Ultrasonics Sonifier (Branson) and passed through a 24G syringe 3 times. Lysates were cleared using centrifugation (16.000 g, 30 min at 4°C) and ultracentrifugation (210,000 g, 30 min at 4°C) and the supernatant was incubated with anti-FLAG beads (Miltenyi Biotec) for 2h at 4°C. The co-immunoprecipitated proteins were isolated using magnetic µMACS columns (Miltenyi) as previously described (26) and RNA was digested by adding RNase I and Benzonase to the isolated proteins. Samples were reduced (5 mM DTT, 30 min, 55°C) and alkylated (40 mM CAA, 30 min, 20°C in the dark). Proteins were digested at 37°C for 16h using trypsin and lysC (both at a 1:75 enzyme-protein ratio) using standard protocols. Formic acid was added to a final concentration of 1% to stop proteolysis. Samples were loaded onto StageTips and labeled with stable isotopes using on-column dimethyl labeling (27). Eluted peptides of the corresponding light, medium and heavy-labelled channels were mixed and dried down in a vacuum concentrator. Peptides were resuspended in 5% DMSO / 1% formic acid and stored at -20°C prior MS analysis (28,29). For details on MS processing and data analysis of the protein interactome see Suppl. Methods.

RNA isolation and RT-qPCR

Whole RNA was isolated from all cells using TRIzol® and subsequent chloroform extraction or a commercial column-based kit (ZymoResearch) according to the manufacturers' instructions. RNA for RNAseq experiments or modification analysis was DNase digested on-column. Both oligo(dT) primers and random hexamers were used for reverse transcription with either SuperScript III Reverse Transcription kit or High Capacity cDNA Reverse Transcription kit (both Thermo Scientific). Real-time quantitative PCR was performed using either SYBR Green Master Mix (Applied Biosystems) or commercial probe-based assays (Thermo Scientific, IDT) on either an Applied Biosystems 7900HT Fast Real-Time PCR System or an Applied Biosystems QuantStudio™ 12K Flex Real-Time PCR System. All experiments were performed on biological and technical triplicates. For the analysis of miRNA expression upon knockdown of AATF, TaqMan® MicroRNA Arrays A and B (Applied

Biosystems) were used according to the manufacturer's instructions. Cycle thresholds were normalized to housekeeping genes such as ACTB or GAPDH using the delta-delta-CT method as described previously (30). For all primer sequences, refer to Suppl. Methods.

RIP-qPCR

HEK 293T cells were transfected with pcDNA6 plasmids containing sequences for triple FLAG-tagged wild-type AATF, a truncation lacking both nucleolar localization sites (AATF 2ΔNoLS) and RFP. 24h after transfection the cells were washed twice with PBS and UV crosslinked (254nm, 150mJ/cm^2) on ice. Following lysis in lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, Protease Inhibitor Cocktail III and Murine RNase Inhibitor) the samples were homogenized by passing three times through a 23 G needle and sonicated (Bioruptor pico, 10x 30sec ON/ 30sec OFF) at 4°C. The samples were treated with DNase I at 37°C. FLAG antibody was coupled to protein G dynabeads. The RNA-protein complexes were immunoprecipitated over night at 4°C. The beads were washed three times with high salt buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) and five times with wash buffer (20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 0.2% Tween-20). Subsequently, the RNA was recovered by TRIzol extraction and concentrated with RNA clean and concentrator columns. An equal amount of immunoprecipitated RNA was reverse transcribed using SuperScript III as described above.

RESULTS

AATF is an RNA-binding protein

We performed an analysis of proteins associated with polyA-tailed RNA using RNA interactome capture (RIC) from 4-SU labeled and UV-A crosslinked lysates of cultured mouse inner medullary collecting duct (mIMCD3) cells (an in-depth global analysis of this dataset is under review at the *Journal of the American Society of Nephrology*). Interestingly, AATF was one of the most enriched proteins in crosslinked samples when analyzed by mass spectrometry. In order to confirm this finding and to show it is not limited to our experiment we checked for the presence of AATF in published RIC datasets. Despite the fact that these studies examined different species and cell lines – cultured human cell lines (embryonic kidney cells HEK 293, hepatocytes HuH-7, cervical cancer cells HeLa) and *S. cerevisiae* – and used different experimental strategies – 4-SU labeling/UV-A crosslinking and UV-C crosslinking – AATF was one of the most strongly enriched proteins in all of them (Fig. 1A) (31-34). To obtain further proof of its RNA-binding capacity using a different method we immunoprecipitated FLAG-tagged AATF and performed radioactive labeling with [γ-^{32}P]-ATP after partial RNase I digestion. PAGE-analysis and autoradiography shows a distinct band at the size of AATF, indeed validating AATF as an RNA binding protein (Fig. 1B).

AATF is associated with ribosomal RNA

In order to determine the identity of the RNA molecules bound to AATF we analyzed enhanced crosslinking and immunoprecipitation (eCLIP) data generated by the ENCODE consortium (Fig. 1C). On the one hand unexpectedly, with AATF having been identified in a screen for proteins associated with polyA-tailed sequences, the vast majority of transcripts bound are ribosomal (Fig. 1D). On the other hand, this finding is well in line with the subcellular localization of AATF to nucleoli, the site of rRNA generation as well as previously published data (15,20,35) and is technically explained by RNA-RNA hybridization between mRNA and rRNA molecules as set forth recently by Piñeiro et al. (21). The enrichment of rRNA species is strongly significant when compared to a similar analysis of 223 other publicly available datasets of 150 RNA binding proteins generated by the ENCODE consortium (Fig. 1D). The overrepresentation is strongest regarding the 45S rRNA precursor molecule (Fig. 1D/E).

Specific binding of AATF to cleavage sites in ribosomal precursor RNAs

In order to get a better understanding of the actual position of the binding sites in the 45S rRNA precursor, taking into account recently published data on the potential role of AATF in ribosome biogenesis (14,15), we focused on the spacer sequences outside 18S and 28S rRNA (which are essentially very abundant RNA species and may also partly be contaminants) containing the cleavage sites essential to rRNA maturation. Doing so, we discovered a number of highly specific peaks (Fig. 2A). These peaks are strongly enriched compared to the 150 ENCODE RBPs when looking at the spacer regions whilst partly less specific regarding the 18S and 28S region. When examining the peaks in the spacer regions in more detail, we noticed a close proximity to sites that are cleaved specifically during rRNA maturation (Fig. 2B) (14). More specifically this is the case for cleavage sites involved in small subunit (SSU) processing as depicted for the sites in the 5' external transcribed spacer (5' ETS) – 01, A0 and 1 – and the first sites in internal transcribed spacer 1 (ITS1) following the sequence of 18S (15,36). This finding is in line with recently published data showing that loss of AATF or other components of the so-called ANN complex results in reduced cleavage activity at these sites indicating reduced SSU processome activity (14,20). Other non SSU associated sites in ITS1, ITS2 and 3'ETS are still detected but show a much lower signal (e.g. 02, Fig. 2B). In line with a role of AATF in ribosome maturation, siRNA-mediated knockdown of AATF leads to a strong reduction of cellular rRNA (Fig. 3A). This finding was confirmed using qPCR, where 18S rRNA levels were significantly reduced after loss of AATF (Fig. 3B). In order to confirm this to be a specific effect of AATF depletion, we repeated the experiment with a siRNA targeting the 3'UTR of AATF in both wild-type cells (WT) and cells harboring a single-copy transgene of GFP-tagged AATF lacking the 3'UTR. This transgene rescued the effect of AATF knockdown on global rRNA levels (Fig. 3C). Of note, cell number (data not shown) and tubulin staining did not differ between treatments.

Nucleolar localization of AATF is required for its association with rRNA precursors

AATF has been described as being primarily nuclear with recent reports confirming an accumulation of the protein in nucleoli. Previous work indicated that the C-terminal portion of AATF is required for

this specific subcellular localization (13,20). However, the actual nucleolar localization signals (NoLS) had not been identified. Prediction of putative NoLS contained in AATF using the Nucleolar localization sequence Detector (NoD) (37) revealed the presence of two putative NoLS in the C-terminal portion of the protein (Fig. 4A, Suppl. Fig. 1A). In order to visualize the subcellular localization of AATF we generated stable cell lines harboring a single-copy transgene encoding GFP-tagged versions of either WT AATF or a mutant protein lacking the two NoLS stretches, the regions between aminoacidic position 326 and 345 and position 494 to 522 (AATF 2ΔNoLS truncation, Fig. 4A/B). The WT fusion protein clearly localizes to nucleoli as shown by the overlap DIC images as well as with the nucleolar protein Nucleolin (Fig. 4B, Suppl. Fig. 1B/C). Interestingly, an impairment of rRNA-transcription by inhibition of RNAPI using both Flavopiridol and Actinomycin D results in a marked alteration of AATF positive foci (Suppl. Fig. 1D). In contrast, the mutant protein does not show a specific nucleolar signal but is dispersed throughout the nucleo- and cytoplasm with a perinuclear enrichment (Fig. 4B). In order to confirm binding of AATF to (pre)ribosomal RNA species and examine the impact of nucleolar localization we performed RIP-qPCR after pulling down overexpressed FLAG-tagged AATF in HEK 293T cells. Here, we could show a strong enrichment of the 45S sequence whilst the enrichment of 18S rRNA was much less prominent (Fig. 4C). Loss of nucleolar localization of AATF (by deletion of the 2 NoLS) results in a markedly diminished association with both 18S and 45S rRNA (Fig. 4C).

AATF binding to other RNA species

As indicated in Fig. 1D AATF binds – even though the largest fraction is rRNA – also to other RNA-species. Here, mRNA is the most common type (69% protein coding) followed by several types of non-coding RNAs (Fig. 5A, Suppl. Table 1). Seeing the association with both ribosomal RNA and mRNAs we were wondering whether the associated coding transcripts were functionally linked to ribosome biogenesis. Interestingly, GO term and pathway analyses confirm such a link with the most overrepresented terms including “ribosome” for KEGG pathways as well as “rRNA processing”, “ribosome”, “translational initiation” and “structural constituent of ribosome” regarding gene ontologies (Fig. 5B). As shown in Fig. 5A several non-coding RNA-species co-precipitate with AATF including snoRNAs and miRNAs. This finding is especially interesting taking into account the important role snoRNAs and associated snoRNPs play in ribosome biogenesis (16,38,39). Interestingly, snoRNAs of the C/D class are overrepresented compared to H/ACA snoRNAs and scaRNAs indicating a potential specific function associated with C/D box snoRNPs (Fig. 5C/D) (40). Since snoRNAs are key players in the site-directed nucleotide modification of ribosomal RNA precursors, which play an important role in their maturation, we set forth to analyze whether loss of AATF may have a global impact on the abundance of these modifications. However, quantification of the most common variants by mass spectrometry did not reveal any significant differences between cells transfected with a siRNA targeting AATF and the respective scrambled controls (Suppl. Fig. 2C). Regarding the fact that miRNAs were also among the transcripts AATF knockdown did not show any impact on overall miRNA abundance (Suppl. Fig. 2A). However, several specific miRNAs showed a highly significant

dysregulation after loss of AATF with miR-25 and miR-509 being the most prominent examples (Suppl. Fig. 2B).

The AATF protein interactome confirms a strong link to ribosome biogenesis

Since AATF had been associated with other nucleolar proteins before (20) and recruitment of components of the ribosome biogenesis machinery to the 45S precursor may be one of the key functions of an rRNA-associated RBP, we decided to characterize the protein interactome of AATF by mass spectrometry using immunoprecipitation (AP-MS) of FLAG-tagged AATF expressed from a single-copy transgene. In order to ascertain specificity, we performed an immunoprecipitation of FLAG-tagged GFP and compared the abundance of proteins in both datasets performing a t-test. Doing so, we identify 165 proteins significantly co-immunoprecipitated with AATF ($log_2FC \geq 2$ and $-log_{10}pvalue \geq 1.3$). Those can be considered as bona-fide AATF interactors (Suppl. Table 2). Whilst some of these bona-fide interactors had been described before in the literature or were classified as physical interactors of AATF in BioGRID (V3.4.153), the majority of proteins detected had not been described to be associated with AATF before (Fig. 6A/Suppl. Table 2). Interestingly, an analysis of the most overrepresented GO terms and KEGG pathways showed primarily terms associated with the ribosome and its biogenesis (Fig. 6B). A closer look at the proteins behind these terms revealed the interactome to contain a large number of known r-proteins and rRNA processing factors (as recently identified by several screens) (14,16) as well as nucleic acid associated enzymes including helicases known to be involved with snoRNA binding or release such as DHX15 (41) (Fig. 6C). Interestingly, more than 80% of the bona fide interactors have been identified as RBPs in independent screens themselves (Fig. 6D). As the AATF interactome contains such a large number of putative or confirmed RBPs we were wondering whether many of these interactions were actually RNA-dependent, as had been shown previously for other RBPs (42). However, repeating the AP-MS experiment including an on-column RNase/benzonase treatment showed that the majority of the interactors do not depend on RNA but are rather direct protein-protein interactions (Fig. 6E, Suppl. Table 2). Using stringent thresholds ($log_2FC \geq 2$, $-log_{10}pvalue \geq 1.3$ in the t test performed between AATF IP treated with RNase versus GFP IP) 93 out of 165 interactors identified without RNase still reached our criteria in the experiment after RNase -treatment (and 147 out of 165 did so for alleviated thresholds $log_2FC \geq 1$, $-log_{10}pvalue \geq 1.0$) (Suppl. Table 2). Obviously, this does not necessarily mean that the other 72 (18) proteins are necessarily RNA-dependent interactors; when comparing the +/- RNase AATF pulldowns directly to each other only two proteins reached stringent thresholds for being overrepresented when RNA is present ($log_2FC \leq -2$, $-log_{10}pvalue \geq 1.3$ in the t test performed between AATF IP treated with RNase versus AATF IP normalized) (Fig. 6E, Suppl. Table 2). In order to obtain a better view of which proteins may at least show a partial dependency on RNA we repeated this analysis using alleviated thresholds ($log_2FC \leq -1$, $-log_{10}pvalue \geq 1.0$) which still results in only 19 putatively RNA-dependent interactors of AATF (Fig. 6E, Suppl. Table 2). Interestingly, a significant proportion of these 19 proteins (~30%) are classical ribosomal proteins, whilst this is only the case for 21 out of 146 of the other AATF interactors (14%) (data not shown). However, functional analyses (GO-terms and KEGG-

pathways) did not reveal any further obvious differences between putatively RNA-dependent and -independent protein interactors. Only two proteins were completely lost after nuclease treatment – RNA-binding protein SLIRP and protein phosphatase PPP1CB (Fig. 6E, Suppl. Table 2). Since inhibition of RNAPI altered the nuclear distribution of AATF (Suppl. Fig. 1D) we went on to ask the question whether the large number of RBPs in its interactome contained other RNAPI-dependent RBPs. Interestingly, overlapping our data with the recent global identification of RNAPI-dependent RBPs by Piñeiro et al. (21) revealed AATF to interact with both RNAPI-dependent (46) and -independent (59) RBPs (Fig. 6F).

A combinatorial approach to the RNA and protein interactome suggest AATF to be a central hub in SSU maturation

Taking into account the published data on a putative role of AATF in SSU maturation (14,15) as well as our data regarding the binding site in 45S rRNA localizing closely to SSU cleavage sites (Fig. 2A/B) we continued our analysis of the protein and RNA interactome focusing on the key complexes involved in SSU maturation (Fig. 7). Here, AATF shows a striking interaction with protein components of the three UTP complexes, the C/D and H/ACA snoRNPs and the exosome complex (Fig. 7). Furthermore, AATF does not only bind the proteins but also several mRNAs encoding for key components of these complexes - such as the enzymatic subunits FBL and DKC1 - which may add a regulatory level to the interaction of AATF with SSU processome complexes (Fig. 7). Interestingly, as determined in our AATF eCLIP experiment, U3 snoRNA itself is among the RNA interaction partners of AATF as well (Fig. 7D). Altogether, out of the snoRNAs known to be required for rRNA processing, two – U3 and U8 – are identified along with 67 other C/D snoRNAs and 19 H/ACA snoRNAs (Fig. 5C/D, 7D/E) (43).

DISCUSSION

AATF had previously been associated with ribosome biogenesis due its nucleolar localization and the impact of loss of the protein on the generation of the 40S subunit and rRNA maturation (14,15). Furthermore, AATF has been shown to be part of a protein complex (ANN complex) required for efficient generation of the SSU (20). However, the actual RBP associating this protein complex to pre-rRNA molecules had not been identified due to a lack of canonical RNA-binding domains. Very recently, work by Piñeiro et al. showed AATF to be an RNAPI-associated RBP itself (21). Their findings are in clear accordance with our results showing that AATF has been identified in a row of RNA-interactome capture screens in different species (32,33,44). A key finding of our study is the identification of the actual RNA-molecules bound by AATF by eCLIP with a clear overrepresentation of ribosomal RNAs. Furthermore, this approach allowed us to localize the binding events in the 45S precursor molecule primarily to the cleavage sites required for SSU maturation strengthening the hypothesized role of AATF in this process. Several lines of published evidence also point towards this direction. Interestingly, Bfr2 – the budding yeast orthologue of AATF – was identified in a screen that

analyzed constituents of yeast 90S particles assembled using plasmid-encoded 3'-truncated pre-18S RNAs and was shown to be specifically bound to the 5'ETS (45). Furthermore, Bammert et al. (20) had shown recently that loss of the ANN complex (containing AATF) resulted in 45S processing defects at the cleavage sites that are necessary for SSU generation. Our data provide strong indications that AATF is the actual RBP of this complex. However – just like AATF itself – both NGDN and NOL10 were detected in screens for polyA-tailed RNA-associated proteins and may possess the capacity to bind RNA themselves as well (31,33). As mentioned above, none of the ANN complex constituents – including AATF itself – possesses a canonical RNA binding domain. While Bammert et al. (20) argued that RNA binding of the ANN complex may be mediated by the WD40 domain of NOL10 which had previously been shown to mediate RNA-binding of other proteins (46), the intrinsically disordered, basic C-terminus of AATF, 20% of which is made up of lysine (K) and arginine (R) residues, appears likely to contain the RNA binding domain. This is also in line with recent studies that could emphasize the importance of intrinsically disordered regions (IDR) for RNA-protein interactions (32,47,48) and the results of recent proteome-wide screens for RNA-binding domains (48,49). Interestingly, work by He et al. showed the two NoLS, the deletion of which resulted in a loss of rRNA binding in our hands, were directly associated with RNA (49). Nevertheless, future experiments will be necessary to further dissect the exact molecular requirements of the physical interaction between AATF and RNA.

Our multilayer analysis - involving the global identification of the protein and RNA interactome - provides further evidence regarding the role of AATF in ribosome maturation. Beyond rRNA precursor molecules themselves AATF binds other RNA species important to the biogenesis of the ribosome. On the one hand, the mRNAs co-precipitating with AATF primarily encode proteins involved with RNA metabolism and ribosome maturation. Here, it is intriguing to hypothesize that AATF may exert a regulatory function regarding the post-transcriptional regulation of these proteins. On the other hand, snoRNAs are highly enriched in the AATF interactome pointing towards a role of AATF in recruiting snoRNAs to pre-rRNA molecules. The fact that AATF is associated with U3 snoRNA – taking into account that cleavage in the 5'ETS and the ITS1 of 45S pre-rRNA strongly depends on the U3 snoRNP – corroborates this hypothesis. This view is further complemented by the protein interactome containing numerous factors required for ribosome biogenesis. Since the vast majority of protein interactions – apart from SLIRP and protein phosphatase PPP1CB – does not depend on RNA but appears to be mediated by direct protein-protein binding, AATF is likely to be a central hub in the coordination of protein-RNA supercomplexes in rRNA maturation.

How is the molecular function of AATF in ribosome biogenesis linked to the previously described phenotypes regarding cellular proliferation and tumorigenesis? With rRNA availability being a central requirement for cellular survival and cell division, it is not surprising that loss of AATF in a knockout mouse model led to early embryonic lethality (35). As to human disease this protein has been implicated to play an important role in cancer biology, a hypothesis that is partly based upon its ability to inhibit apoptosis. AATF has previously been found to be amplified or overexpressed in both

hematological and solid tissue tumors, to correlate with poor prognosis and reduced survival (4-6,9,10,50) and to mediate its effects on apoptosis by the modulation of p53 abundance and function (4,6). Furthermore, increased ribosome biogenesis is not only employed by tumors to increase their proliferative potential but appears to be a risk factor for cancer onset on its own (51). However, AATF appears to ensure cellular survival not only in the setting of cancer. Since its protective role has also been in the setting of oxidative stress exposure to different cell types including renal tubular cells in a model of acute kidney injury (52-54). The known link between ribosome stress and p53 activity in the light of our new data allows for the exciting hypothesis that AATF mediates its effects on cell death employing this pathway. Based on the interaction between r-proteins and the E3 ubiquitin ligase MDM2 defects in ribosome maturation activate p53 (18,39,55-57). Loss of AATF would thus increase p53 activity linking the roles of AATF in ribosome biogenesis and p53 activity. In the light of an increasing number of studies trying to target ribosome biogenesis in this setting (58) and taking into account that AATF has been shown to sensitize cancer but not normal cells to antineoplastic drugs (5), future research shedding light on these aspects will be highly valuable.

AVAILABILITY

AATF K562 eCLIP data has been deposited at the ENCODE Data Coordination Center (*https://www.encodeproject.org/*) under accession identifier ENCSR819XBT, and HepG2 data has been deposited at the Gene Expression Omnibus (series GSE107766; samples GSM2878484 (Rep 1), GSM2878485 (Rep 2), GSM2878554 (size-matched input)). The Interactome dataset has been uploaded in ProteomeXchange via the PRIDE database (https://www.ebi.ac.uk/pride/archive/). Project accession: PXD011055; Reviewer account details: Username: reviewer43734@ebi.ac.uk; Password: psjMNW4h.

ACCESSION NUMBERS

Protein interactome PRIDE database: Project accession: PXD011055
eCLIP data GEO: series GSE107766; samples GSM2878484, GSM2878485, GSM2878554
eCLIP data ENCODE: ENCSR819XBT

ACKNOWLEDGEMENT

We thank Serena Greco-Torres, Martyna Brütting and Ruth Herzog for excellent technical assistance. Special thanks to Constantin Rill for help with TALEN transgenesis. The monoclonal antibody E7 developed by M. McCutcheon and S. Carroll was obtained from the Developmental Studies

Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa, Department of Biology, Iowa City, IA 52242.

FUNDING

This work was supported by an MD Fellowship by Boehringer Ingelheim Fonds (BIF, to R.W.K.), the Nachwuchsgruppen.NRW program of the Ministry of Science North Rhine Westfalia (MIWF, to R.-U.M.) and the German Research Foundation (MU3629/2-1 to R.-U.M., BE2212 and KFO329 to T.B., SCHE1562/6 to B.S. and HE3397/8-1, HE3397/13-2, SPP1784 to M.H.). R.-U.M. was supported by a Research Scholarship from Deutsche Nierenstiftung. HH received an MD Fellowship of the German Society of Internal Medicine (DGIM). CD acknowledges funding by the Klaus Tschira Stiftung gGmbH. This work was partially funded by the National Human Genome Research Institute ENCODE Project as a grant U54HG007005 to GWY. ELVN is a Merck Fellow of the Damon Runyon Cancer Research Foundation (DRG-2172-13) and is supported by a K99 grant from the NHGRI (HG009530). GWY was partially supported by grants from the NIH (HG007005, NS075449).

CONFLICT OF INTEREST

ELVN and GWY are co-founders and consultants for Eclipse BioInnovations Inc. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies. The authors declare no other competing financial interests.

REFERENCES

1. Bruno, T., De Angelis, R., De Nicola, F., Barbato, C., Di Padova, M., Corbi, N., Libri, V., Benassi, B., Mattei, E., Chersi, A. *et al.* (2002) Che-1 affects cell growth by interfering with the recruitment of HDAC1 by Rb. *Cancer Cell*, **2**, 387-399.
2. Fanciulli, M., Bruno, T., Di Padova, M., De Angelis, R., Iezzi, S., Iacobini, C., Floridi, A. and Passananti, C. (2000) Identification of a novel partner of RNA polymerase II subunit 11, Che-1, which interacts with and affects the growth suppression function of Rb. *FASEB J*, **14**, 904-912.
3. Ishigaki, S., Fonseca, S.G., Oslowski, C.M., Jurczyk, A., Shearstone, J.R., Zhu, L.J., Permutt, M.A., Greiner, D.L., Bortell, R. and Urano, F. (2010) AATF mediates an antiapoptotic effect of the unfolded protein response through transcriptional regulation of AKT1. *Cell Death Differ*, **17**, 774-786.
4. Welcker, D., Jain, M., Khurshid, S., Jokic, M., Hohne, M., Schmitt, A., Frommolt, P., Niessen, C.M., Spiro, J., Persigehl, T. *et al.* (2018) AATF suppresses apoptosis, promotes proliferation and is critical for Kras-driven lung cancer. *Oncogene*, **37**, 1503-1518.
5. Höpker, K., Hagmann, H., Khurshid, S., Chen, S., Hasskamp, P., Seeger-Nukpezah, T., Schilberg, K., Heukamp, L., Lamkemeyer, T., Sos, M.L. *et al.* (2012) AATF/Che-1 acts as a

phosphorylation-dependent molecular modulator to repress p53-driven apoptosis. *EMBO J*, **31**, 3961-3975.
6. Bruno, T., Desantis, A., Bossi, G., Di Agostino, S., Sorino, C., De Nicola, F., Iezzi, S., Franchitto, A., Benassi, B., Galanti, S. *et al.* (2010) Che-1 promotes tumor cell survival by sustaining mutant p53 transcription and inhibiting DNA damage response activation. *Cancer Cell*, **18**, 122-134.
7. Bruno, T., De Nicola, F., Iezzi, S., Lecis, D., D'Angelo, C., Di Padova, M., Corbi, N., Dimiziani, L., Zannini, L., Jekimovs, C. *et al.* (2006) Che-1 phosphorylation by ATM/ATR and Chk2 kinases activates p53 transcription and the G2/M checkpoint. *Cancer Cell*, **10**, 473-486.
8. Desantis, A., Bruno, T., Catena, V., De Nicola, F., Goeman, F., Iezzi, S., Sorino, C., Gentileschi, M.P., Germoni, S., Monteleone, V. *et al.* (2015) Che-1 modulates the decision between cell cycle arrest and apoptosis by its binding to p53. *Cell Death Dis*, **6**, e1764.
9. Bruno, T., Valerio, M., Casadei, L., De Nicola, F., Goeman, F., Pallocca, M., Catena, V., Iezzi, S., Sorino, C., Desantis, A. *et al.* (2017) Che-1 sustains hypoxic response of colorectal cancer cells by affecting Hif-1α stabilization. *J Exp Clin Cancer Res*, **36**, 32.
10. Desantis, A., Bruno, T., Catena, V., De Nicola, F., Goeman, F., Iezzi, S., Sorino, C., Ponzoni, M., Bossi, G., Federico, V. *et al.* (2015) Che-1-induced inhibition of mTOR pathway enables stress-induced autophagy. *EMBO J*, **34**, 1214-1230.
11. Bruno, T., Iezzi, S. and Fanciulli, M. (2016) Che-1/AATF: A Critical Cofactor for Both Wild-Type- and Mutant-p53 Proteins. *Front Oncol*, **6**, 34.
12. Iezzi, S. and Fanciulli, M. (2015) Discovering Che-1/AATF: a new attractive target for cancer therapy. *Front Genet*, **6**, 141.
13. Ferraris, S.E., Isoniemi, K., Torvaldson, E., Anckar, J., Westermarck, J. and Eriksson, J.E. (2012) Nucleolar AATF regulates c-Jun-mediated apoptosis. *Mol Biol Cell*, **23**, 4323-4332.
14. Tafforeau, L., Zorbas, C., Langhendries, J.L., Mullineux, S.T., Stamatopoulou, V., Mullier, R., Wacheul, L. and Lafontaine, D.L. (2013) The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol Cell*, **51**, 539-551.
15. Badertscher, L., Wild, T., Montellese, C., Alexander, L.T., Bammert, L., Sarazova, M., Stebler, M., Csucs, G., Mayer, T.U., Zamboni, N. *et al.* (2015) Genome-wide RNAi Screening Identifies Protein Modules Required for 40S Subunit Synthesis in Human Cells. *Cell Rep*, **13**, 2879-2891.
16. Nicolas, E., Parisot, P., Pinto-Monteiro, C., de Walque, R., De Vleeschouwer, C. and Lafontaine, D.L. (2016) Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress. *Nat Commun*, **7**, 11390.
17. Steitz, T.A. (2008) A structural understanding of the dynamic ribosome machine. *Nat Rev Mol Cell Biol*, **9**, 242-253.
18. Liu, Y., Deisenroth, C. and Zhang, Y. (2016) RP-MDM2-p53 Pathway: Linking Ribosomal Biogenesis and Tumor Surveillance. *Trends Cancer*, **2**, 191-204.
19. Boulon, S., Westman, B.J., Hutten, S., Boisvert, F.M. and Lamond, A.I. (2010) The nucleolus under stress. *Mol Cell*, **40**, 216-227.
20. Bammert, L., Jonas, S., Ungricht, R. and Kutay, U. (2016) Human AATF/Che-1 forms a nucleolar protein complex with NGDN and NOL10 required for 40S ribosomal subunit synthesis. *Nucleic Acids Res*.
21. Pineiro, D., Stoneley, M., Ramakrishna, M., Alexandrova, J., Dezi, V., Juke-Jones, R., Lilley, K.S., Cain, K. and Willis, A.E. (2018) Identification of the RNA polymerase I-RNA interactome. *Nucleic Acids Res*.
22. Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol*, **7**, 2745-2752.
23. Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B. *et al.* (2009) Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol*, **27**, 851-857.
24. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods*, **9**, 676-682.
25. Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K. *et al.* (2016) Robust

transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods*, **13**, 508-514.
26. Hubner, N.C. and Mann, M. (2011) Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). *Methods*, **53**, 453-459.
27. Boersema, P.J., Raijmakers, R., Lemeer, S., Mohammed, S. and Heck, A.J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc*, **4**, 484-494.
28. Kohli, P., Bartram, M.P., Habbig, S., Pahmeyer, C., Lamkemeyer, T., Benzing, T., Schermer, B. and Rinschen, M.M. (2014) Label-free quantitative proteomic analysis of the YAP/TAZ interactome. *American journal of physiology. Cell physiology*, **306**, C805-818.
29. Rappsilber, J., Ishihama, Y. and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Analytical chemistry*, **75**, 663-670.
30. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402-408.
31. Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M. *et al.* (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell*, **46**, 674-690.
32. Beckmann, B.M., Horos, R., Fischer, B., Castello, A., Eichelbaum, K., Alleaume, A.M., Schwarzl, T., Curk, T., Foehr, S., Huber, W. *et al.* (2015) The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat Commun*, **6**, 10127.
33. Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M. *et al.* (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, **149**, 1393-1406.
34. Liao, Y., Castello, A., Fischer, B., Leicht, S., Foehr, S., Frese, C.K., Ragan, C., Kurscheid, S., Pagler, E., Yang, H. *et al.* (2016) The Cardiomyocyte RNA-Binding Proteome: Links to Intermediary Metabolism and Heart Disease. *Cell Rep*, **16**, 1456-1469.
35. Thomas, T., Voss, A.K., Petrou, P. and Gruss, P. (2000) The murine gene, Traube, is essential for the growth of preimplantation embryos. *Dev Biol*, **227**, 324-342.
36. Phipps, K.R., Charette, J. and Baserga, S.J. (2011) The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip Rev RNA*, **2**, 1-21.
37. Scott, M.S., Troshin, P.V. and Barton, G.J. (2011) NoD: a Nucleolar localization sequence detector for eukaryotic and viral proteins. *BMC Bioinformatics*, **12**, 317.
38. Scott, M.S. and Ono, M. (2011) From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie*, **93**, 1987-1992.
39. Lindstrom, M.S., Jurada, D., Bursac, S., Orsolic, I., Bartek, J. and Volarevic, S. (2018) Nucleolus as an emerging hub in maintenance of genome stability and cancer pathogenesis. *Oncogene*.
40. Yoshihama, M., Nakao, A. and Kenmochi, N. (2013) snOPY: a small nucleolar RNA orthological gene database. *BMC Res Notes*, **6**, 426.
41. Yoshimoto, R., Kataoka, N., Okawa, K. and Ohno, M. (2009) Isolation and characterization of post-splicing lariat-intron complexes. *Nucleic Acids Res*, **37**, 891-902.
42. Brannan, K.W., Jin, W., Huelga, S.C., Banks, C.A., Gilmore, J.M., Florens, L., Washburn, M.P., Van Nostrand, E.L., Pratt, G.A., Schwinn, M.K. *et al.* (2016) SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Mol Cell*, **64**, 282-293.
43. Watkins, N.J. and Bohnsack, M.T. (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA*, **3**, 397-414.
44. Baltz, A.G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M. *et al.* (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell*, **46**, 674-690.
45. Soltanieh, S., Lapensee, M. and Dragon, F. (2014) Nucleolar proteins Bfr2 and Enp2 interact with DEAD-box RNA helicase Dbp4 in two different complexes. *Nucleic Acids Res*, **42**, 3194-3206.
46. Loedige, I., Jakob, L., Treiber, T., Ray, D., Stotz, M., Treiber, N., Hennig, J., Cook, K.B., Morris, Q., Hughes, T.R. *et al.* (2015) The Crystal Structure of the NHL Domain in Complex with RNA Reveals the Molecular Basis of Drosophila Brain-Tumor-Mediated Gene Regulation. *Cell Rep*, **13**, 1206-1220.

47. Hentze, M.W., Castello, A., Schwarzl, T. and Preiss, T. (2018) A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol*.
48. Castello, A., Fischer, B., Frese, C.K., Horos, R., Alleaume, A.M., Foehr, S., Curk, T., Krijgsveld, J. and Hentze, M.W. (2016) Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol Cell*, **63**, 696-710.
49. He, C., Sidoli, S., Warneford-Thomson, R., Tatomer, D.C., Wilusz, J.E., Garcia, B.A. and Bonasio, R. (2016) High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. *Mol Cell*, **64**, 416-430.
50. Folgiero, V., Sorino, C., Pallocca, M., De Nicola, F., Goeman, F., Bertaina, V., Strocchio, L., Romania, P., Pitisci, A., Iezzi, S. *et al.* (2018) Che-1 is targeted by c-Myc to sustain proliferation in pre-B-cell acute lymphoblastic leukemia. *EMBO Rep*, **19**.
51. Derenzini, M., Montanaro, L. and Trere, D. (2017) Ribosome biogenesis and cancer. *Acta Histochem*, **119**, 190-197.
52. Xie, J. and Guo, Q. (2006) Apoptosis antagonizing transcription factor protects renal tubule cells against oxidative damage and apoptosis induced by ischemia-reperfusion. *J Am Soc Nephrol*, **17**, 3336-3346.
53. Wang, D., Chen, T.Y. and Liu, F.J. (2018) Che-1 attenuates hypoxia/reoxygenation-induced cardiomyocyte apoptosis by upregulation of Nrf2 signaling. *Eur Rev Med Pharmacol Sci*, **22**, 1084-1093.
54. Guo, S., Chen, R., Chen, X., Xie, Z., Huo, F. and Wu, Z. (2018) Che-1 inhibits oxygen-glucose deprivation/reoxygenation-induced neuronal apoptosis associated with inhibition of the p53-mediated proapoptotic signaling pathway. *Neuroreport*.
55. Donati, G., Brighenti, E., Vici, M., Mazzini, G., Trere, D., Montanaro, L. and Derenzini, M. (2011) Selective inhibition of rRNA transcription downregulates E2F-1: a new p53-independent mechanism linking cell growth to cell proliferation. *J Cell Sci*, **124**, 3017-3028.
56. Ofir-Rosenfeld, Y., Boggs, K., Michael, D., Kastan, M.B. and Oren, M. (2008) Mdm2 regulates p53 mRNA translation through inhibitory interactions with ribosomal protein L26. *Mol Cell*, **32**, 180-189.
57. Ogawa, L.M. and Baserga, S.J. (2017) Crosstalk between the nucleolus and the DNA damage response. *Mol Biosyst*, **13**, 443-455.
58. Drygin, D., Lin, A., Bliesath, J., Ho, C.B., O'Brien, S.E., Proffitt, C., Omori, M., Haddach, M., Schwaebe, M.K., Siddiqui-Jain, A. *et al.* (2011) Targeting RNA polymerase I with an oral small molecule CX-5461 inhibits ribosomal RNA synthesis and solid tumor growth. *Cancer Res*, **71**, 1418-1430.
59. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, **4**, 44-57.

TABLE AND FIGURES LEGENDS

Figure 1 AATF is an RNA-binding protein associated with pre-rRNA

A Graphical representation of AATF enrichment comparing our dataset (striped) to published mRNA interactome datasets (black). The Log_2 fold change was calculated for crosslinked over non-crosslinked samples.

B Biochemical validation of AATF as RNA binding protein. The FLAG-tagged AATF protein was immunoprecipitated from crosslinked (+) and non-crosslinked (-) samples and the associated RNA was labeled by T4 PNK with [γ-^{32}P]-ATP. The protein-RNA complexes were separated on PAA-gels and blotted onto PVDF membranes. PNK-assay: autoradiograph of the membrane containing FLAG-AATF with the associated RNA labelled with [γ-^{32}P]-ATP. Western blot: visualization of FLAG-tagged protein by western blotting with the anti-FLAG antibody. MW: protein molecular weight marker

C Stepwise scheme of the eCLIP-protocol. 1. Intact K562 cell were UV crosslinked and lysed. 2. Representation of crosslink formed between protein and RNA after exposure to UV light. 3. The cell lysate was submitted to partial RNase I digestion. 4. The RBP-RNA complex was immunoprecipitated. 5. On the beads, the RNA was dephosphorylated and a 3´ RNA adapter was ligated. 6. The lysate was size selected via PAGE and the RNA was extracted. 7. Reverse transcription was performed with the reaction stopping at the site of crosslinking after which the RNA was removed. 8. A 3´DNA adapter was ligated and the fragments were PCR amplified. 9. The fragments were Illumina indexed and submitted to paired-end sequencing.

D Pie charts indicating the proportions of RNA species bound by AATF. Analysis of ribosomal RNA species in AATF eCLIP datasets (AATF Rep1 and Rep2) compared to size-matched input and to a combined analysis of 150 ENCODE RBP datasets (all IPs, all inputs) shows the enrichment of rRNA transcripts by AATF.

E Box plot comparison of AATF eCLIP IP data to input samples and to 150 ENCODE RBP datasets showing enrichment of rRNA species in IP (red dot) over input (black circle) is specific for AATF. The other 150 ENCODE RBP datasets show a decrease of rRNA species in the IP (white box) compared to inputs (grey box). RNA species are plotted against the reads identified per million unique fragments.

Figure 2 Identification of 45S pre-rRNA sites bound by AATF

A Relative information content on CLIPseq peak distribution along the 45 pre-rRNA and the mature 18S and 28S rRNA comparing AATF to 150 other ENCODE RBPs.

B AATF eCLIP reads map to the 45S pre-rRNA and show enrichment at cleavage sites involved in SSU maturation. Graph showing the fold enrichment of eCLIP reads of two replicates (yellow and red)

along the 13357 bp long 45S pre-rRNA. 45S rRNA cleavage sites involved in SSU maturation are indicated by arrows and dashed lines. Black arrow heads indicate regions shown below in detail for the sites: 01, A0 and 2.

Figure 3 Loss of AATF depletes total rRNA levels

A Knockdown of AATF leads to reduction of rRNA. The CDS of AATF was targeted with siRNA in mIMCD3 cells, which induced a significant depletion of endogenous AATF and was accompanied by a decrease in rRNA after 48h of incubation. Top panel: western blot with anti-AATF antibody. Middle panel: anti-β-tubulin western blot (loading control). Bottom panel: EtBr stained agarose gel. MW: protein molecular weight marker

B U2OS cells were transfected with either a siRNA pool targeting the AATF coding sequence or non-targeted control siRNA. After 48h, cells were harvested and RNA was isolated. qPCR for ACTB, AATF and 18S rRNA on RNA derived from these cells confirms the significant decrease of 18S rRNA after AATF depletion.

C Expression of AATF single-copy transgene rescues reduction of rRNA in AATF depleted U2OS cells. siRNA against the 3´UTR of AATF was transfected into wild-type U2OS cells and U2OS cells with a TALEN mediated, single-copy integration of GFP-AATF lacking the endogenous 3´UTR into the AAV locus. The 3´UTR specific knockdown of AATF in the wild type cells lead to a reduction of the 18S and 28S rRNA. The expression of the GFP-tagged transgene in the TALEN manufactured U2OS GFP::AATF cell line rescued the amount the rRNA species. Top panel: western blot with anti-AATF antibody. Middle panel: anti-β-tubulin western blot (loading control). Bottom panel: EtBr stained agarose gel. MW: protein molecular weight marker

Figure 4 Displacement from the nucleolus impairs the RNA binding capacity of AATF

A Scheme of protein domains and sequence features of wild type AATF (LZ: leucin zipper, NoLS: nucleolar localization signal) and a truncated version of the protein lacking both nucleolar localization signals.

B Localisation of WT and mutant AATF. Immunofluorescence images of transgenic U2OS cell lines expressing either stably integrated full-length GFP-AATF (upper panels) or the GFP-tagged AATF 2ΔNoLS truncation (lower panels). Deletion of both nucleolar localization sequences results in

displacement of the protein from the nucleolus. Nuclei were counterstained with DAPI. DIC: differential interference contrast.

C RIP-qPCR analysis of 18S rRNA and 45S pre-rRNA. Full-length FLAG-tagged AATF and the FLAG-tagged AATF 2ΔNoLS truncation were overexpressed in HEK 293T cells and immunoprecipitated in RNA-interaction preserving conditions. Quantification of co-precipitated rRNA revealed a significant reduction of RNA binding for both ribosomal transcripts after loss of the two NoLS sites. RQ: relative quantification (ct values for WT AATF, 2ΔNoLS AATF and RFP were normalised against the corresponding input, and consecutively against RFP). Western blot analysis (anti-FLAG antibody) shows efficient precipitation of FLAG-tagged proteins. FLAG-RFP served as negative control.

Figure 5 AATF interacts with other RNA species

A Pie chart depicting the distribution of RNA biotypes bound by AATF other than rRNA. 69 % of transcripts other than rRNA bound by AATF are protein-coding. snoRNA and scaRNA (14%) are frequently bound targets. miRNA represent 1% of the transcripts. “other” biotypes encompass lincRNA and antisense RNA.

B Bubble chart depicting the functional analysis of mRNA transcripts bound by AATF showing the terms contained in the top functional annotation cluster as identified using the DAVID Bioinformatics online tool (59) (for the 292 eCLIP targets showing significant peaks in at least two experiments, Suppl. Table 1); grey: GOBP, white: GOCC, blue: GOMF, pink: KEGG pathway. GO terms are plotted according to fold enrichment and $-\log_{10}$ of the respective p-value, with size of the bubble increasing proportionally the number of genes contained in the respective cluster.

C Pie chart showing the proportions of snoRNAs bound by AATF. Among the transcript biotype group of snoRNAs AATF preferentially binds C/D box snoRNAs, with 73% of bound snoRNAs belonging to this subtype. Box H/ACA snoRNAs comprise 20% and scaRNA 7% of transcripts bound by AATF.

D Bar chart representing the number of sno/scaRNAs bound by AATF vs the number of human annotated RNAs of this biotype. 69 of 426 known C/D box snoRNAs, 19 of 331 known H/ACA snoRNAs and 7 out of 28 known scaRNAs co-precipitated with AATF as annotated by the snoRNA Orthological Gene Database (snOPY) (40).

Figure 6 The protein interactome of AATF is strongly enriched for proteins involved in ribosome biogenesis

A Scatter plot of AATF interactome. Immunoprecipitation of FLAG-AATF (expressed from single-copy insertion in FlpIn 293T cells) and Mass Spectrometry analysis, reveals 165 protein interactors fulfilling stringent criteria when compared to a FLAG-GFP pulldown ($\log_2$FC ≥2, $-\log_{10}$ p-value ≥1.3). Known

interactors are labeled with colored dots (black dot: AATF, orange dots: physical interaction as annotated in BIOGRID for human AATF, blue dots: physical interaction as annotated for the yeast AATF orthologue red dots: interactors manually curated from the literature, white dots: previously not identified interactors).

B Bubble chart depicting the functional analysis of AATF interacting proteins showing the terms contained in the top functional annotation cluster as identified using the DAVID Bioinformatics online tool (59) (for the 165 bona fide AATF interactors); grey: GOBP, white: GOCC, blue: GOMF, pink: KEGG pathway. GO terms are plotted according to fold enrichment and $-\log_{10}$ of the respective p-value, with size of the bubble increasing proportionally the number of genes contained in the respective cluster.

C Bar chart showing the percentage of AATF interacting proteins in the protein groups of r-proteins (16), rRNA processing factors (14) and human RNA helicases (47) in grey. The numbers below indicate the total number of proteins per group, the numbers within the grey bars indicate the number of AATF interactors within this group.

D Scatter plot showing AATF protein interactors annotated as RNA binding proteins. According to the comparison with known RBPs about 80% of the proteins interacting with AATF are themselves annotated as RNA-binding proteins (red dots: known RBPs, black dot: AATF, white dots: AATF interacting proteins not annotated as RBPs)

E Scatter plot highlighting RNA dependent AATF interacting proteins. Comparing the interactome after RNase treatment revealed that only few of the protein interactions depend on RNA (black dot: AATF, red dots: RNA-dependent interactions as defined by a $\log_2$ FC ≥2 and $-\log_{10}$ p-value ≥1.3 compared to RNase treated IP, orange dots: partially RNA-dependent interactions as defined by a $\log_2$ FC ≥1 and $-\log_{10}$ p-value ≥1 compared to RNase treated IP, white dots: RNA independent interactors). See also Suppl. Figure 3A for a direct comparison and Suppl. Table 2.

F Scatter plot showing that some of the AATF interactors were shown in the recent publication by Pinero et al. to be RNA binding proteins either dependent or independent.on RNAPI inhibition Dark green: RNAPI independent interactors, light green: RNAPI dependent interactors, white dots: not identified in RNAPI RNA interactome.

Figure 7 AATF is a key hub for protein complexes involved in SSU maturation and interacts with RNAs and proteins involved

A-F In our dataset, AATF interacts with the majority of the proteins known to be members of the key protein complexes involved in the maturation of ribosomal RNA. Networks were created using Cytoscape and show constituents of the tUTP (A), the UTP-B (B), the UTP-C (C), the C/D snoRNAP and MPP10 complex (D) as well as the H/A snoRNP (E) and the exosome complex (F).

nodes: yellow circles = RNA interactors identified in AATF eCLIP, green circles = protein interactors identified in AATF AP-MS, diamonds = snoRNAs, grey diamonds / circles = members of the complex that were not detected in our interactomes.

yellow lines: AATF-RNA interaction, solid = high stringency (665 eCLIP targets containing at least one significant peak, Suppl. Table 1), dashed = low stringency (eCLIP: all targets identified, Suppl. Table 1)

green lines: AATF-protein interaction, solid = high stringency ($\log_2$FC ≥2, $\log_{10}$ p-value ≥1.3 in either AATF with or without RNAse compared to the respective GFP pulldown), dashed = low stringency (detected in our interactome with a positive FC in at least two replicates of AATF either with or without RNAse compared to the respective GFP pulldown)

Figure 1

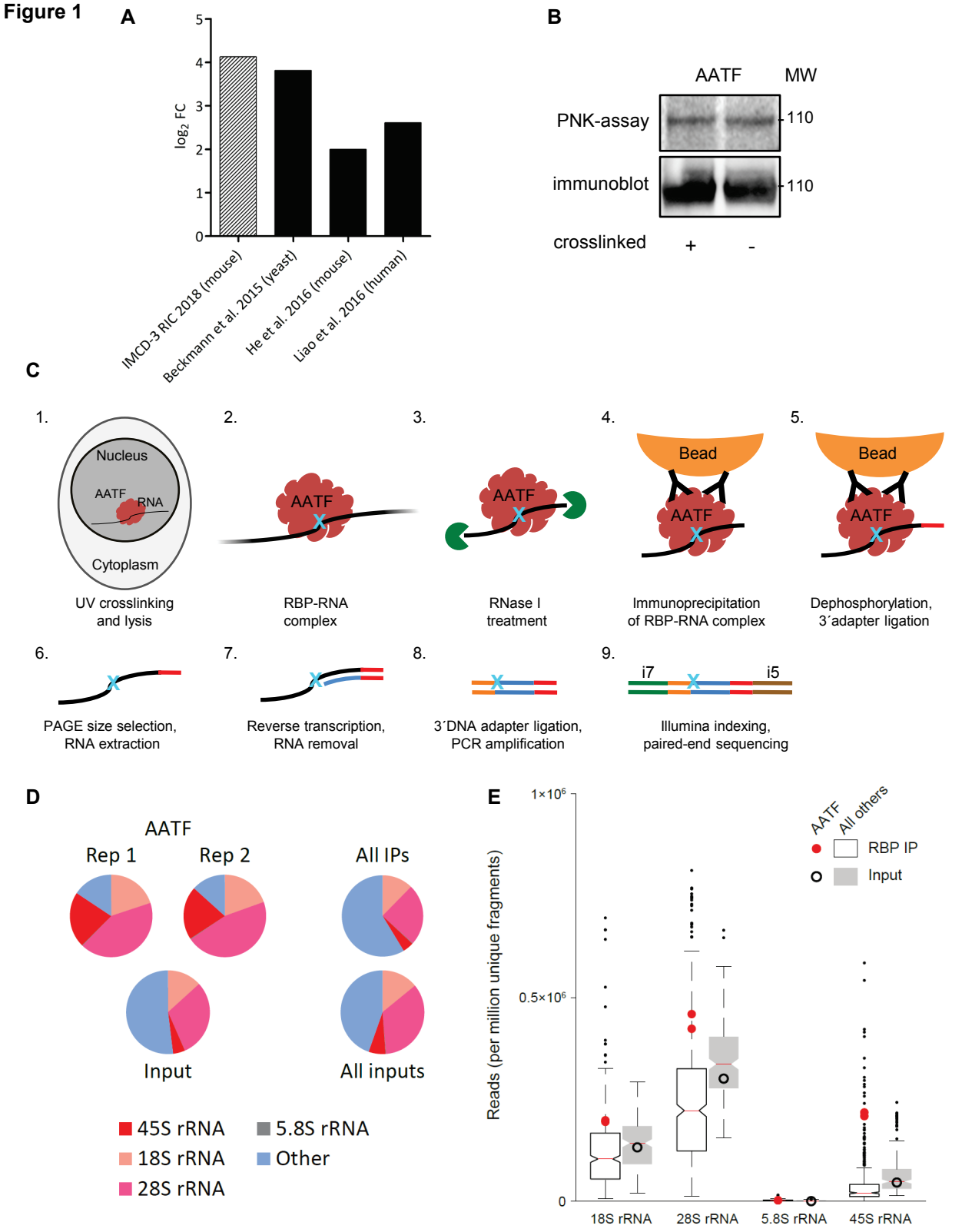

A
log2 FC
0
1
2
3
4
5
IMCD-3 RIC 2018 (mouse)
Beckmann et al. 2015 (yeast)
He et al. 2016 (mouse)
Liao et al. 2016 (human)
B
AATF
MW
PNK-assay
immunoblot
110
110
crosslinked
+
-
C
1.
Nucleus
AATF
RNA
Cytoplasm
UV crosslinking and lysis
2.
AATF
RBP-RNA complex
3.
AATF
RNase I treatment
4.
Bead
AATF
Immunoprecipitation of RBP-RNA complex
5.
Bead
AATF
Dephosphorylation, 3´adapter ligation
6.
PAGE size selection, RNA extraction
7.
Reverse transcription, RNA removal
8.
3´DNA adapter ligation, PCR amplification
9.
i7
i5
Illumina indexing, paired-end sequencing
D
AATF
Rep 1
Rep 2
All IPs
Input
All inputs
45S rRNA
18S rRNA
28S rRNA
5.8S rRNA
Other
E
Reads (per million unique fragments)
1×10⁶
0.5×10⁶
0
AATF
All others
RBP IP
Input
18S rRNA
28S rRNA
5.8S rRNA
45S rRNA precursor

Figure 2

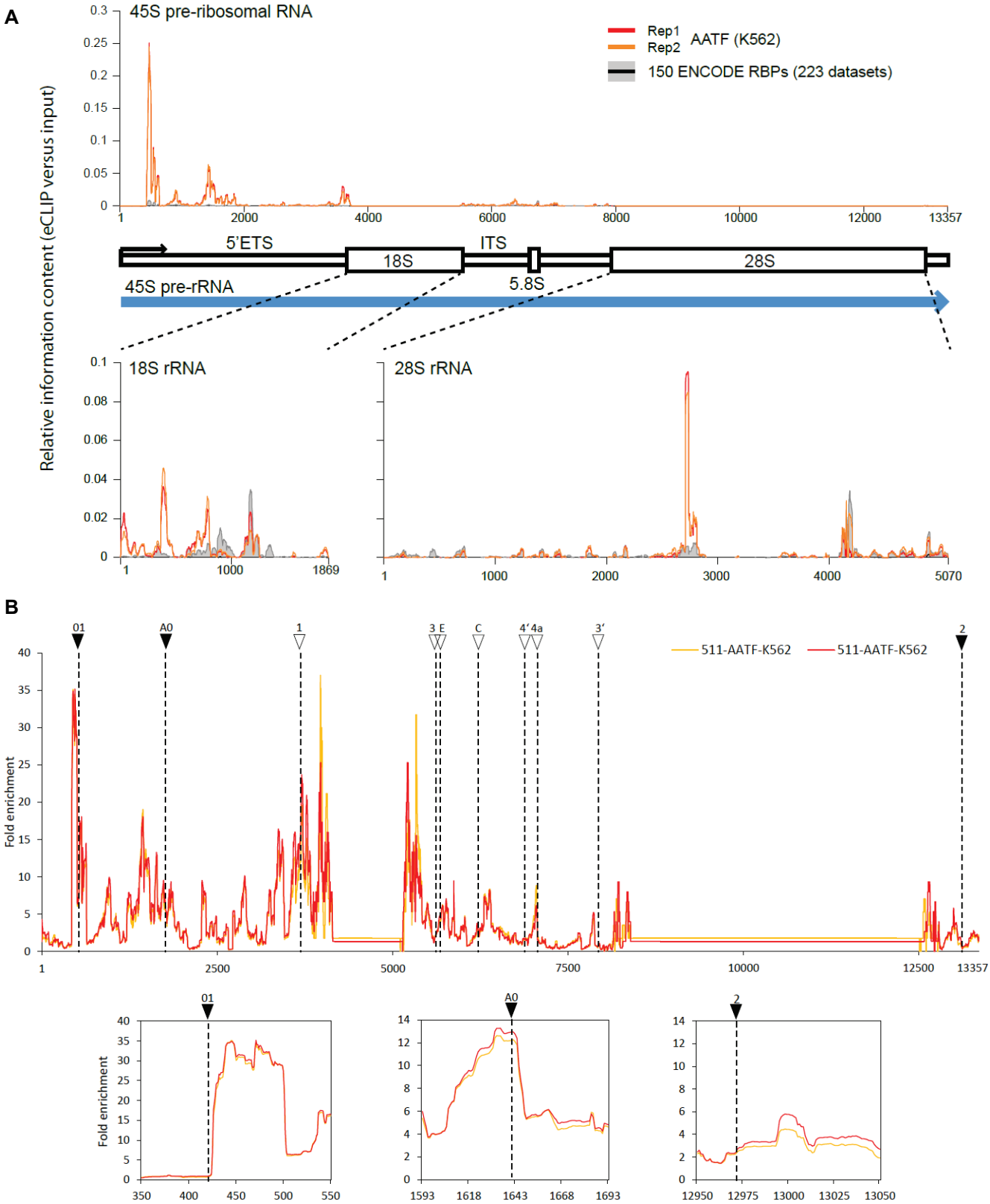
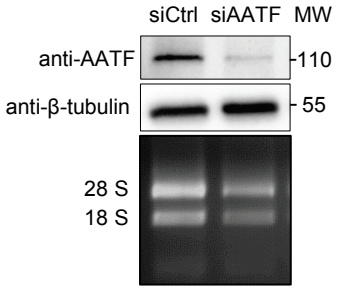
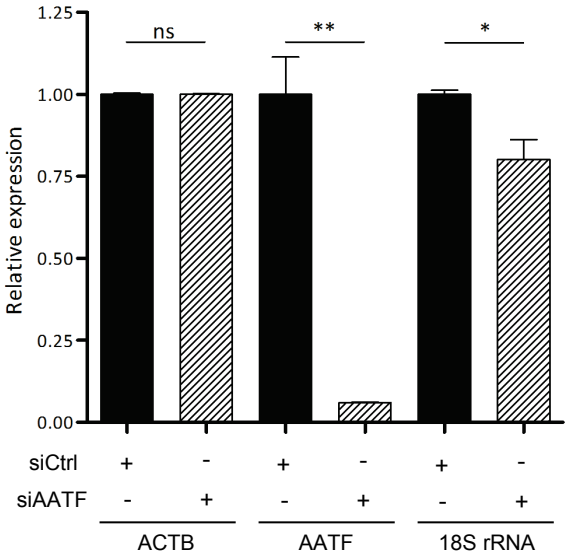

A
45S pre-ribosomal RNA
Rep1
Rep2
AATF (K562)
150 ENCODE RBPs (223 datasets)
Relative information content (eCLIP versus input)
5'ETS
18S
ITS
28S
45S pre-rRNA
5.8S
18S rRNA
28S rRNA
B
01
A0
1
3 E
C
4' 4a
3'
2
511-AATF-K562
511-AATF-K562
Fold enrichment

Figure 3

A


siCtrl
siAATF
MW
anti-AATF
110
anti-β-tubulin
55
28 S
18 S

B


1.25
1.00
0.75
0.50
0.25
0.00
Relative expression
ns
**
*
siCtrl + - + - + -
siAATF - + - + - +
ACTB
AATF
18S rRNA

C

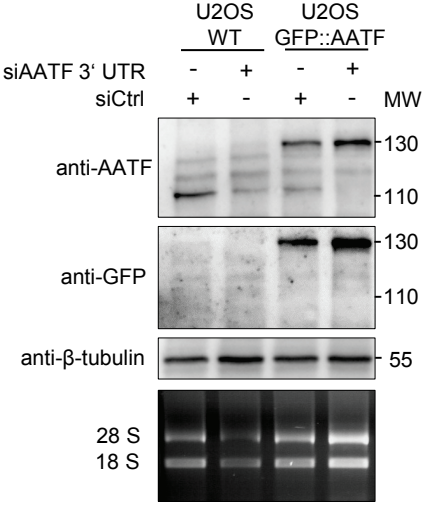
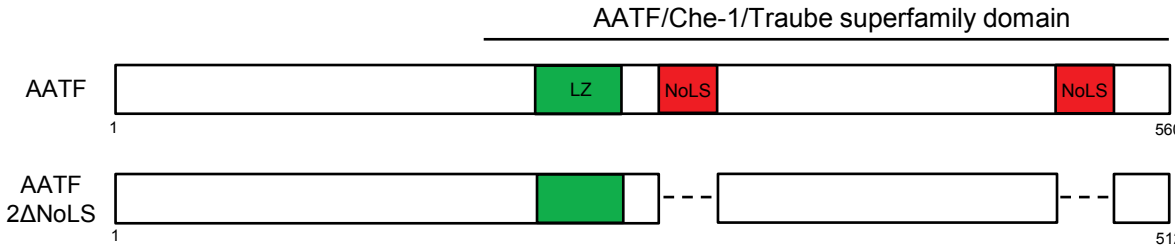
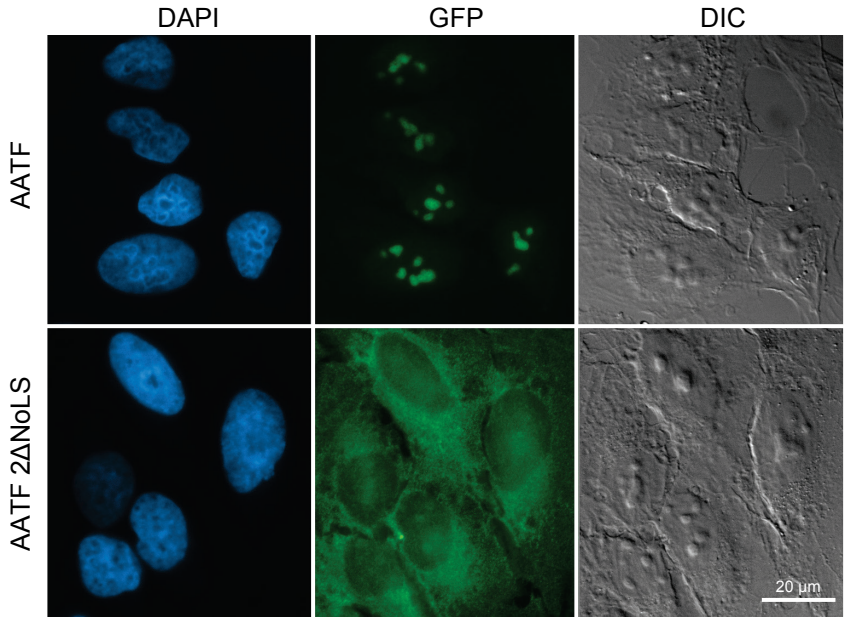

U2OS WT
U2OS GFP::AATF
siAATF 3' UTR - + - +
siCtrl + - + - MW
anti-AATF
130
110
anti-GFP
130
110
anti-β-tubulin
55
28 S
18 S

Figure 4

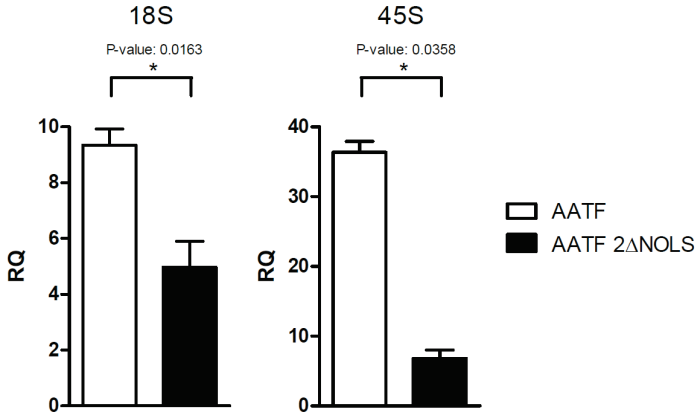
A


AATF/Che-1/Traube superfamily domain
AATF
LZ
NoLS
NoLS
1
560
AATF
2ΔNoLS
1
512

B


DAPI
GFP
DIC
AATF
AATF 2ΔNoLS
20 µm

C


18S
P-value: 0.0163
*
RQ
10
8
6
4
2
0
45S
P-value: 0.0358
*
RQ
40
30
20
10
0
AATF
AATF 2ΔNOLS

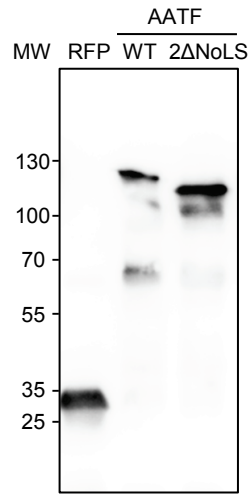
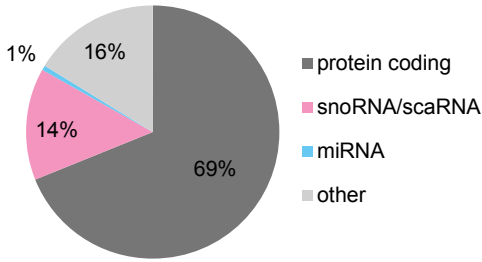
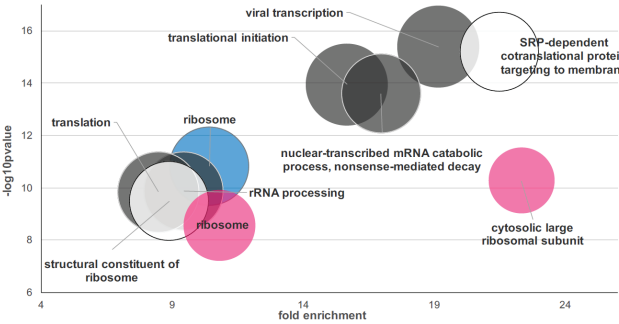

AATF
MW
RFP
WT
2ΔNoLS
130
100
70
55
35
25

Figure 5

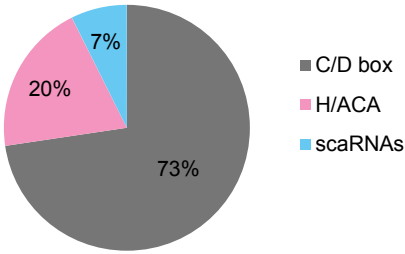
A



B



C



D

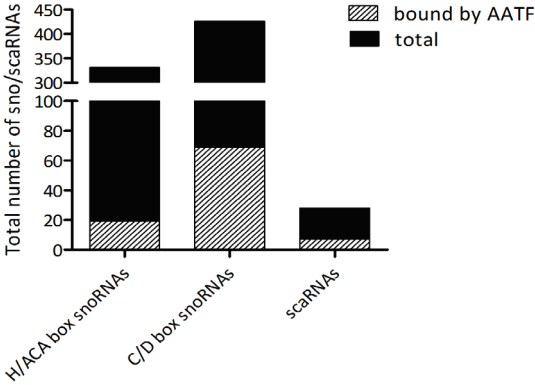


Figure 6

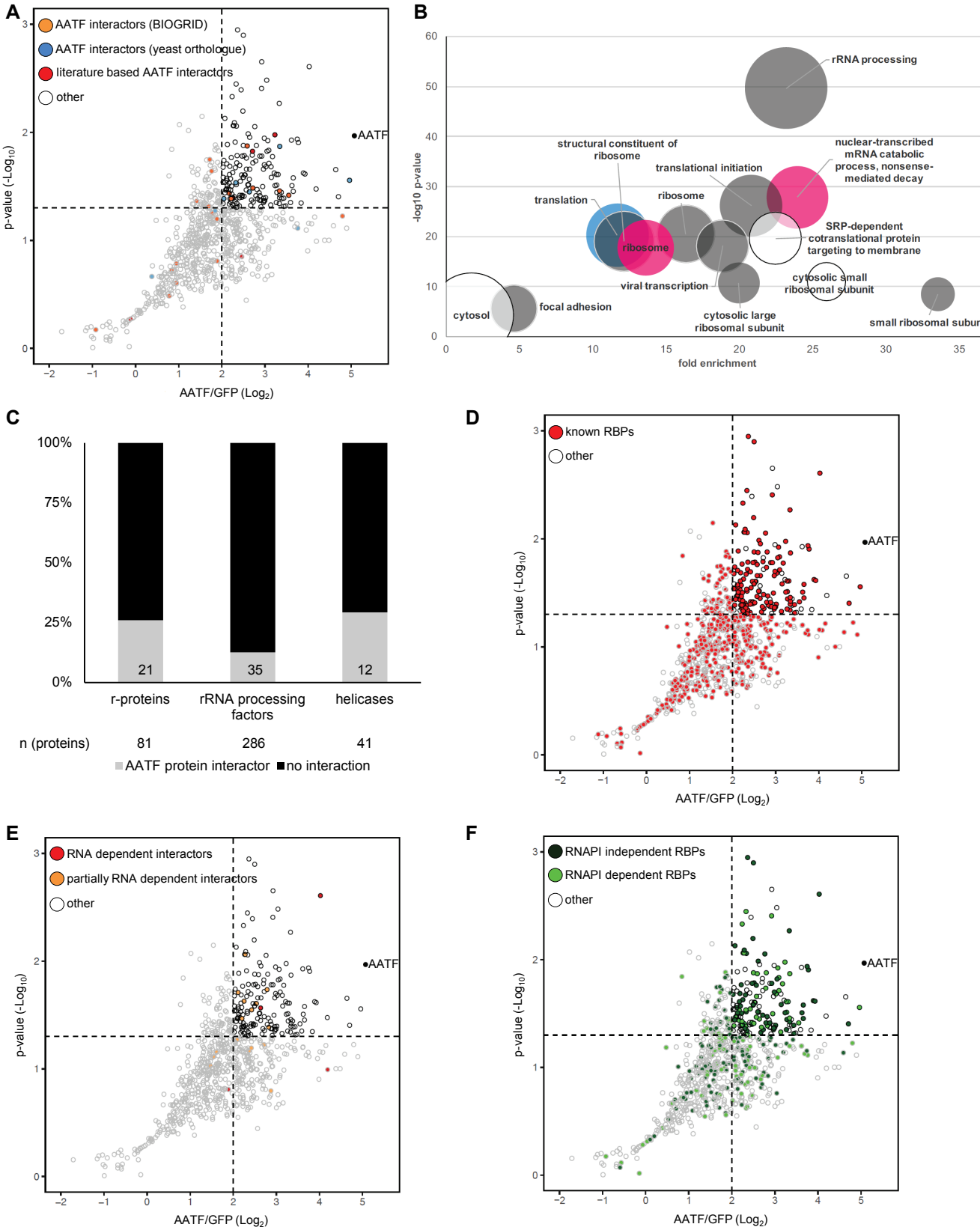
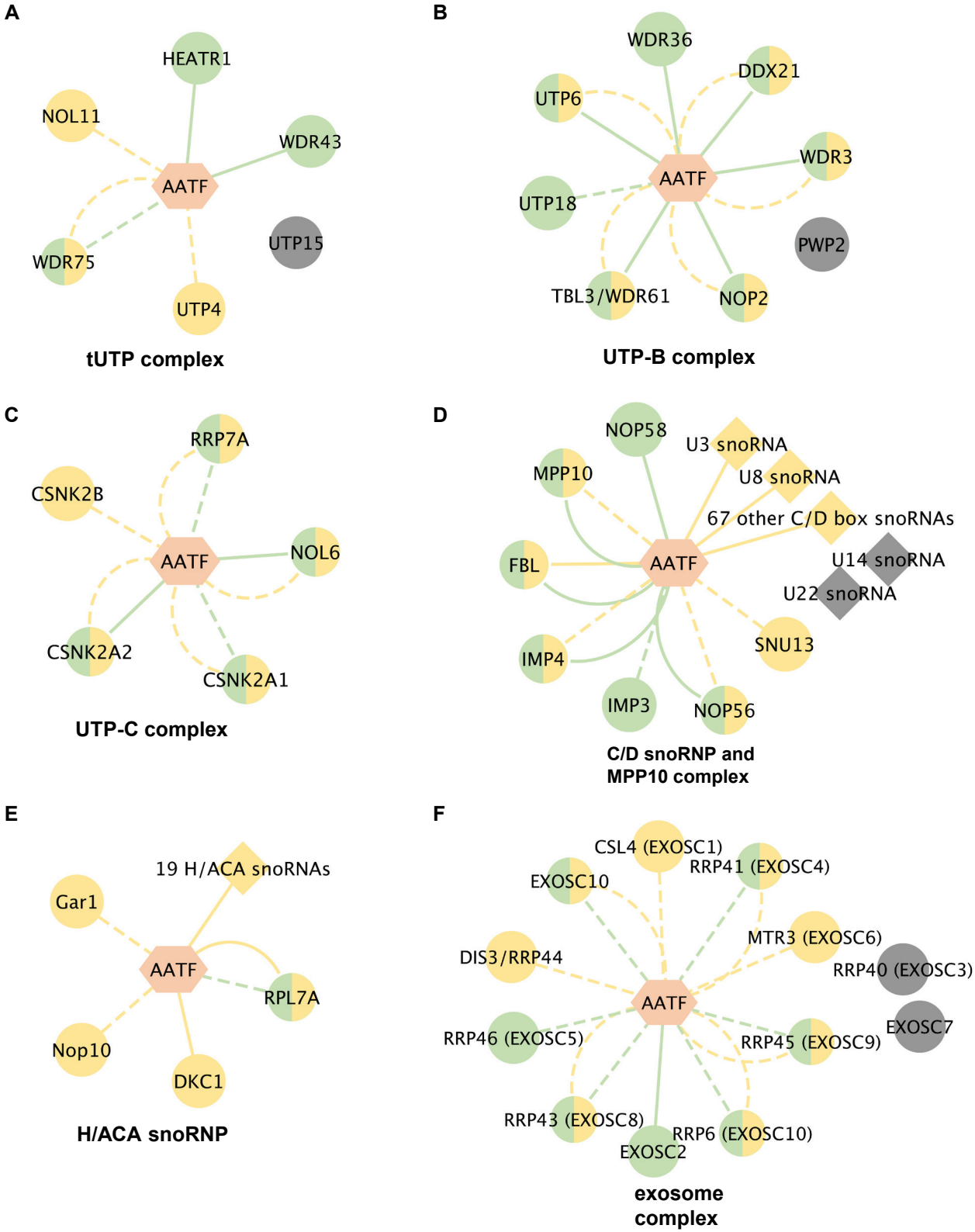

A
AATF interactors (BIOGRID)
AATF interactors (yeast orthologue)
literature based AATF interactors
other
AATF
p-value (-Log10)
AATF/GFP (Log2)
B
rRNA processing
structural constituent of ribosome
translational initiation
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay
translation
ribosome
SRP-dependent cotranslational protein targeting to membrane
viral transcription
cytosolic small ribosomal subunit
focal adhesion
cytosol
cytosolic large ribosomal subunit
small ribosomal subunit
-log10 p-value
fold enrichment
C
100%
75%
50%
25%
0%
21
35
12
r-proteins
rRNA processing factors
helicases
n (proteins)
81
286
41
AATF protein interactor
no interaction
D
known RBPs
other
AATF
p-value (-Log10)
AATF/GFP (Log2)
E
RNA dependent interactors
partially RNA dependent interactors
other
AATF
p-value (-Log10)
AATF/GFP (Log2)
F
RNAPI independent RBPs
RNAPI dependent RBPs
other
AATF
p-value (-Log10)
AATF/GFP (Log2)

Figure 7


A
HEATR1
NOL11
WDR43
AATF
WDR75
UTP15
UTP4
tUTP complex
B
WDR36
DDX21
UTP6
WDR3
AATF
UTP18
PWP2
TBL3/WDR61
NOP2
UTP-B complex
C
RRP7A
CSNK2B
AATF
NOL6
CSNK2A2
CSNK2A1
UTP-C complex
D
NOP58
U3 snoRNA
MPP10
U8 snoRNA
67 other C/D box snoRNAs
FBL
AATF
U14 snoRNA
U22 snoRNA
SNU13
IMP4
IMP3
NOP56
C/D snoRNP and
MPP10 complex
E
19 H/ACA snoRNAs
Gar1
AATF
RPL7A
Nop10
DKC1
H/ACA snoRNP
F
CSL4 (EXOSC1)
RRP41 (EXOSC4)
EXOSC10
MTR3 (EXOSC6)
DIS3/RRP44
RRP40 (EXOSC3)
AATF
EXOSC7
RRP46 (EXOSC5)
RRP45 (EXOSC9)
RRP43 (EXOSC8)
RRP6 (EXOSC10)
EXOSC2
exosome
complex