Translational Oncology
www.transonc.com

Volume 13 Number 2 February 2020 pp. 452–470 452

Hallmarks and Determinants of Oncogenic Translation Revealed by Ribosome Profiling in Models of Breast Cancer☆

Christos Vaklavas*,[1], Scott W. Blume* and William E. Grizzle†

*Department of Medicine, Division of Hematology / Oncology, University of Alabama at Birmingham, Birmingham, AL 35294, USA; †Department of Pathology, O'Neal Comprehensive Cancer Centre, University of Alabama at Birmingham, Birmingham, AL 35294, USA

Abstract

Gene expression is extensively and dynamically modulated at the level of translation. How cancer cells prioritize the translation of certain mRNAs over others from a pool of competing mRNAs remains an open question. Here, we analyze translation in cell line models of breast cancer and normal mammary tissue by ribosome profiling. We identify key recurrent themes of oncogenic translation: higher ribosome occupancy, greater variance of translational efficiencies, and preferential translation of transcriptional regulators and signaling proteins in malignant cells as compared with their nonmalignant counterpart. We survey for candidate RNA interacting proteins that could associate with the 5′untranslated regions of the transcripts preferentially translated in breast tumour cells. We identify SRSF1, a prototypic splicing factor, to have a pervasive direct and indirect impact on translation. In a representative estrogen receptor—positive and estrogen receptor—negative cell line, we find that protein synthesis relies heavily on SRSF1. SRSF1 is predominantly intranuclear. Under certain conditions, SRSF1 translocates from the nucleus to the cytoplasm where it associates with *MYC* and *CDK1* mRNAs and upregulates their internal ribosome entry site—mediated translation. Our results point to a synergy between splicing and translation and unveil how certain RNA-binding proteins modulate the translational landscape in breast cancer.

Translational Oncology (2020) 13, 452–470

Introduction

Although our understanding of transcriptional regulation and dysregulation in cancer has expanded dramatically over the recent years, comparatively less is known about the dysregulation of gene expression that occurs at the level of translation. Transcript levels have been traditionally used as a proxy of the protein abundance in a cell; however, the correlation between mRNA and protein levels is imperfect. Although a subject of intense investigation [1], large-scale genomic studies have shown that the levels of a protein in a cell can be best predicted by its translation rates [2]. Translation represents a more proximal level of control, allowing the cell to adapt swiftly to stress conditions by modulating protein synthesis from an existing pool of mRNAs, unlike the process of transcription which mediates more stable changes in cell physiology or fate [3].

Cancer cells differ from their nonmalignant counterparts not only at the level of transcription but also at the level of translation [4]. They usurp the regulatory mechanisms that govern translation to carry out translational programs that lead to the phenotypic hallmarks

Address all correspondence to: Christos Vaklavas, Department of Medicine, Division of Oncology, Huntsman Cancer Institute of the University of Utah, 2000 Circle of Hope RS2509, Salt Lake City, UT 84112, USA. E-mail: Christos.vaklavas@hci.utah.edu

☆ **Funding**: This work was supported by the Susan G. Komen Career Catalyst Award [CCR15331062 to CV]; National Cancer Institute [R01 CA108886 to S.W.B.]; and National Institutes of Health [Collaborative Human Tissue Network grant 1UM1CA183728, the Tissue Procurement Shared Facility of the O'Neal UAB Comprehensive Cancer Center P30CA13148-41, the Pulmonary Hypertension Breakthrough Initiative 1R24HL123767-01, and the Hepato/Renal Fibrocystic Diseases Core Center P30DK074038 to WEG].

[1]Present Affiliation: Department of Medicine, Division of Oncology, Huntsman Cancer Institute of the University of Utah, Salt Lake City, UT 84112, USA.

Received 28 August 2019; Revised 28 November 2019; Accepted 1 December 2019

1936-5233/19
https://doi.org/10.1016/j.tranon.2019.12.002

of malignancy [5]. Translation is a critical nexus in neoplastic transformation. The transformative impact of multiple oncogenes and signaling pathways that are activated, upregulated, or mutated in cancer converges at the level of translation [4,6,7]. Moreover, translational dysregulation endows cancer cells with the plasticity and adaptability needed to overcome a diverse array of stresses associated with a hostile microenvironment including antitumor therapies.

Leveraging the breadth and depth of coverage of massively parallel nucleic acid sequencing, we utilized the ribosome profiling strategy [8−10] to dissect the translational profiles of cell line models of breast cancer. We identify common themes of oncogenic translation across cancer cell lines that model diverse subtypes of breast cancer with distinct natural histories. We note that many more genes are differentially expressed at the level of translation than at the level of transcription and that the overlap between the two is partial. The genes and transcripts that are preferentially translated in cancer fall consistently into the same ontology categories, most notably transcriptional regulation, and signaling. We identify that the transcripts commonly transcribed in nonmalignant and malignant cell lines but preferentially translated in cancer harbor common motifs in their 5′untranslated regions, which most consistently and most significantly match the RNA-binding motifs of eIF4B and SRSF1. We uncover a novel direct regulatory function of the prototypic splicing factor SRSF1 on translation, whereby when SRSF1 translocates to the cytoplasm, it directly associates with *MYC* and *CDK1* mRNA and enhances their internal ribosome entry site−mediated translation.

Materials and Methods

Cell lines and Media

Human mammary epithelial cells (HMECs) were obtained from Lonza and cultured in the medium recommended by the manufacturer. Serum-deprived media consisted of mammary epithelial cell growth basal medium (MEBM) supplemented with amphotericin/gentamicin and hydrocortisone (as provided by the manufacturer) admixed with full serum media in a combination ratio of 9:1. Essentially the serum-deprived conditions contained 10% of the full concentration of recombinant human EGF, bovine pituitary extract (BPE), and insulin.

MCF10A cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA) and were propagated using standard techniques in DMEM/F12 media supplemented with 5% horse serum (Invitrogen, cat# 16050-122), recombinant human EGF 20 ng/mL (Peprotech, cat# AF-100-15), hydrocortisone 0.5 mg/mL (Sigma Aldrich, cat# H-0888), cholera toxin 100 ng/mL (Sigma-Aldrich, cat# C8052), insulin (ThermoFisher Scientific, cat#12585014), penicillin/streptomycin 1%. Serum-deprived conditions consisted of DMEM/F12 media plus full serum media as described previously in a combination ratio of 9:1. The concentration of cholera toxin was kept at 100 ng/mL because it regulates formation of mammary epithelial acini *in vitro* which is a feature of differentiation [11]. MCF10A cells were grown in monolayer. Because changes in translation can be immediate, we wanted to avoid any bias that may have been introduced during isolation of the MCF10A cells from Matrigel.

T47D cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA) and propagated using standard techniques in RPMI-1640 medium containing 10% fetal calf serum (FCS, HyClone, cat# SH30071.03), 10 μg/ml insulin (ThermoFisher Scientific, cat#12585014), and 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES). Charcoal--stripped serum conditions consisted of RPMI-1640 plus 10% charcoal stripped FCS and 10 mM HEPES (no insulin). FCS was charcoal stripped in accordance with the following protocol: FCS was stirred very slowly on a hot plate at 37 °C with 1% charcoal and 0.1% dextran (Sigma-Aldrich, cat#D1662, molecular weight 35,000−45,000) for 4 h and then collected into 50-mL tubes. The tubes were then spun to precipitate the charcoal, and the supernatants were collected and sterile filtered. All experiments were conducted with the same aliquot. Serum-deprived conditions (for the immunofluorescence microscopy studies) consisted of RPMI-1640 plus FCS 0.5% (no insulin).

ZR75-1 cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA) and propagated in RPMI-1640 with 20% FCS and 10 mM HEPES; charcoal-stripped serum conditions consisted of RPMI-1640 supplemented with 20% charcoal-stripped serum and 10 mM HEPES.

SUM159PT cells were obtained from Asterand Bioscience (Detroit, MI, USA) and propagated in Ham's F-12 media supplemented with 5% FCS, 10 mM HEPES, 5 μg/mL insulin (ThermoFisher Scientific, cat#12585014), 1 μg/ml hydrocortisone (Sigma-Aldrich, cat# H-0888), and penicillin/streptomycin 1%. Serum-deprived conditions consisted of Ham's F12 media supplemented with 0.5% FCS and 10 mM HEPES (no insulin and hydrocortisone).

MDA-MB-231 cells were obtained from ATCC (Manassas, VA, USA) and grown in DMEM high glucose (4.5 gr/L) pyruvate L-glutamate (Gibco 11995-065) supplemented with fetal calf serum 10%, 10 mM HEPES, and penicillin/streptomycin 1%. Serum-deprived conditions consisted of DMEM high glucose (4.5 gr/L) pyruvate L-glutamate supplemented with 1% FCS and 1% penicillin/streptomycin.

All cells were maintained in an incubator in humidified atmosphere containing 5% CO_2 at 37 °C.

Ribosome Profiling

Experiments were designed so as cells were ~80% confluent at the time of harvesting and lysis and at least 3 days after plating. At least 2 biological replicates for each cell line were performed. Media was changed to fresh media in all flasks 24 h before harvesting and to serum-deprived media or media supplemented with charcoal--stripped serum (T47D and ZR75-1 cells) 1 h before harvesting as indicated. We used cycloheximide (final concentration 100 μg/mL, incubation for 10 min before collection) to halt the translating ribosomes [12]. Ribosome profiling was conducted following the protocol from Ingolia et al. [8] with the following modifications: (1) rRNA depletion was performed upfront using the Ribo-zero Gold rRNA removal kit (Illumina, cat# MRZG126) following manufacturer's instructions and (2) the linker-ligated ribosome-protected fragments were retrieved using the RNA Clean and Concentrator Kit (Zymo Research, cat# R1015) following manufacturer's instructions. We lysed the cells without flash freezing. RNA digestion was performed with RNAse I (Ambion, cat# AM2294). We performed a series of preliminary experiments to determine the optimal RNAse I concentration for RNA digestion for each cell line. We found that RNA from HMECs and MCF10A cells is quite more sensitive to RNAse I degradation than RNA from T47D, ZR75-1, SUM159PT, and MDA-MB-231 cells. For the three-quarters of the

cell lysate from a 75-cm^2 cell culture flask intended for ribosome footprint recovery (the remaining one quarter was not digested and was used for RNA-seq), we used 4.5 μL of RNAse I for HMECs and MCF10A cells and 9 μL of RNAse I for all other cell lines.

Total RNA (after digestion with Turbo DNase, Invitrogen cat# AM2238) and digested ribosome-protected RNA was isolated using the Qiagen miRNeasy kit (Qiagen, cat# 217004) following manufacturer's protocol. Total RNA was rRNA depleted using the Ribo-zero Gold rRNA removal kit (Illumina, cat# MRZG126). Sequencing was done on an Illumina HiSeq 2500 platform.

Analyses of the Ribosome Profiling Results

The ribosome footprints were trimmed of the 3′ linker sequence (ATTGATGGTGCCTACAG) and the 5′ end nucleotide. Reads that aligned to rRNA, mtRNA, tRNA (downloaded from http://gtrnadb.ucsc.edu/), small nuclear RNA (downloaded from ensembl.org), and PhiX (mismatch cost = 2, insertion/deletion cost = 3, non-strand specific, similarity and length fraction = 0.8, maximum number of hits for a read = 30) were removed. The remaining sequences were aligned to and assembled using the reference human genome downloaded from ensemble.org (Genome Assembly used GRCh37 available at http://ftp.ensembl.org/pub/release-75/fasta/homo_sapiens/dna/) using the CLC Bio software (Qiagen Bioinformatics, Denmark) with the following parameters: mismatch cost = 2, insertion/deletion cost = 3, non−strand specific, similarity and length fraction = 0.8, maximum number of hits for a read = 10. Fragment length distribution was done using R; metagene analyses were conducted using the plastid toolkit (available at https://plastid.readthedocs.io) [13]. Gene ontology analyses were conducted with the DAVID bioinformatics resource with the default parameters [14].

The total RNA sequencing reads were analyzed as follows: reads that aligned to rRNA (using the following parameters: mismatch cost = 2, insertion/deletion cost = 3, non−strand specific, similarity and length fraction = 0.8, maximum number of hits for a read = 30) were discarded and the unmapped reads were aligned to the human genome and assembled using the GRCh37 build as a reference (parameters used: mismatch cost = 2, insertion/deletion cost = 3, non−strand specific, similarity and length fraction = 0.8, maximum number of hits for a read = 10).

Motif discovery

The 5′UTR sequences of the transcripts that were translated three times or greater more efficiently in the malignant cells than the HMECs were downloaded from ensemble.org. Transcripts for whom the 5′UTR sequence was unavailable and with 5′UTR sequences <7 nucleotides long were removed. Each sequence of this subset was then shuffled using the shuffle functionality of CLC Bio generating 10 shuffled sequences of the same length and base composition. The motifRG package in R [15] was used for discriminative motif discovery using the 5′UTR sequence subsets as foreground and the 10x shuffled sequences as background. The candidate RNA-binding proteins for the identified motifs were retrieved from http://cisbp-rna.ccbr.utoronto.ca using the Motif Scan tool [16]. Motif matching for the top-scoring motifs (eIF4B: SUYGGAM and SRSF1: AGGASM) to confirm that the identified motifs match to 5′UTR sequences of the 3x differentially translated transcripts plus identify the individual transcripts that contain these motifs in their respective 5′UTR was done using FIMO [17] (available at http://meme-suite.org/).

CRISPR-Cas9 Mediated Deletion of SRSF1

The crRNAs (Invitrogen TrueGuide™ crRNA; ID: CRISPR248724_CR; ID: CRISPR963384_CR, cat# A35509; control: TrueGuide™ crRNA Negative Control, non-targeting 1, cat# A35519) and tracrRNA (Invitrogen TrueGuide™, cat# A35507) were reconstituted and annealed following the manufacturer's instructions. Transfections were performed in a 24-well plate using the Lipofectamine™ CRISPRMAX™ (Invitrogen, cat# CMAX00003) and the TrueCut™ Cas9 Protein v2 (Invitrogen cat# A36497). Our best results with SUM159PT cells were obtained with Cas9 protein and gRNA of 7.5 and 10 pmol, respectively, and 1.5 μL of Lipofectamine CRISPRMAX per well.

Gene editing efficiency in a pooled cell population was evaluated using the GeneArt™ Genomic Cleavage Detection Kit (Invitrogen, cat# A24372) following manufacturer's instructions and using the following primers to amplify the SRSF1 locus encompassing the predicted cleavage sites (Supplemental Figure 6A):

Forward: GTGCGCCGAGCTGATAAAGG; Reverse: ACTCAGCTCCTTACTCGACTCCTGC.

We observed higher gene editing efficiency with repeated transfections at optimal conditions. To obtain the desired knockout clonal cell line, we proceeded with limited dilution cloning at ~0.8 cells/well of a 96-well plate. Confirmation of SRSF1 deletion of the selected clones was done with immunoblotting.

Transfections

We used lipofectamine RNAiMAX (ThermoFisher Scientific, cat# 13778030) for our siRNA transfections following manufacturer's instructions. siRNA-lipofectamine complexes were incubated at room temperature for 40 min before being added to the cells. SRSF1 siRNAs used were siRNA1: s12725 and siRNA2: s12727 (both SilencerSelect ThermoFisher Scientific, cat# 4392420); control siRNA was SilencerSelect Negative Control No. 1 (ThermoFisher Scientific, cat# 4390843). SilencerSelect technology uses an improved prediction algorithm and chemical modifications to minimize off-target effects. SRSF1 has multiple transcript variants; siRNA 2 was predicted to target all variants while some isoforms were spared by siRNA 1. This differential isoform targeting by our siRNAs may in part account for the milder effect of siRNA 1. For transfections in a 24-well plate (cell viability assays, luciferase assays), we used 5 pmol of siRNA and 1.5 μL of lipofectamine RNAiMAX per well. For transfections in a 6-well plate (immunoblotting and puromycin incorporation assay), we used 40 pmol of siRNA and 12 μL of lipofectamine RNAiMAX per well. The final concentration of siRNA and lipofectamine was kept the same in all conditions (1 μM and 3 μl/mL media, respectively).

We used lipofectamine 2000 (ThermoFisher Scientific cat# 11668030) for stable and transient transfections using the same concentration of control and vector DNA. Clonal cell populations stably transfected with the respective linearized plasmids were cotransfected with a linear puromycin resistance cassette (Takara, cat# 631626) and selected via limited dilution cloning.

Immunoblotting

Whole-cell lysates were rapidly prepared by adding lysis buffer (4% SDS and 720 mM 2-mercaptoethanol, preheated to 100 °C) directly to cell monolayer; the lysate was recovered and heated for 5 more minutes after the addition of glycerol to 10%. For instances in which cell adhesion was compromised (e.g., downregulation of SRSF1 in

SUM159PT cells and treatment with nocodazole in T47D cells), care was taken (using centrifugation) to ensure recovery of loosely adherent or floating cells. Equivalent aliquots (by protein content, unless otherwise indicated) were separated on 10% SDS/PAGE gels, transferred to 0.2-μm nitrocellulose (BioRad, cat# 1620146) or PVDF (BioRad, cat#1620174) membranes, and subjected to standard immunoblotting procedures followed by chemiluminescence (GE Healthcare, cat# RPN2109) image capture. Image analysis and band densitometry was performed using the Genetools Analysis software (Syngene, Frederick, MD).

Cell Viability Assay

Cell viability was determined using the CellTiter-Glo Luminescent Cell Viability Assay (Promega, cat# G7570) following manufacturer's instructions. Negative controls (media with buffer alone) were subtracted from readings. Luminescence measurements were obtained using a Glomax 20/20 luminometer (Promega, Madison, WI, USA).

Low-density (clonogenic) Survival Assay

Cells were plated on an ultralow attachment plate (Corning Costar, cat# 3471) at a concentration of 1000 cells/mL in serum deprived media and incubated for 10 days. At the end of the incubation time, MTT was added in each well at a final concentration of 250 μg/mL. The plate was incubated for 3 h, and the contents of the wells were aspirated, centrifuged, and the supernatants aspirated. The precipitated formazan crystals were resuspended in equal volumes of dimethyl sulfoxide (DMSO), transferred in dedicated cuvettes, and the absorbance was measured at 570 nm in a spectrophotometer.

Indirect Immunofluorescence Staining and Confocal Imaging

The cells were seeded in 8-well chamber slides (Nunc; Nalge Nunc International, Penfield, NY, USA) and allowed for at least 24 h to recover and resume proliferation before changing the media to serum-deprived media (SUM159PT and T47D) or media with charcoal-stripped serum (T47D). Cells were fixed with 4% paraformaldehyde for 20 min at room temperature followed by permeabilization with 0.2% Triton X-100 for 10 min, washing in PBS with 75 mM glycine at least 3 times, and blocking in 5% normal goat serum. Then, the primary antibody was added and incubated for 1 h at room temperature. After two washes in PBS and a 10-min reblocking step, secondary antibodies (1:200) were added and incubated for 50 min in the dark. After two washes with PBS, nuclei were stained with 4′6-Diamidino-2-phenylindole dihydrochloride (DAPI) (0.2 μg/ml, Sigma–Aldrich) and mounted using ProLong Gold (Life Technologies; Thermo Fischer Scientific).

Images were captured using a Nikon A1R confocal instrument (Nikon Corporation, Tokyo, Japan). Fields were randomly selected for imaging on the basis of the DAPI staining pattern alone. Paired images of control and experimental wells were acquired sequentially, and all settings including laser power, PMT voltage, and pinhole were held constant between samples.

Plasmids

The bicistronic reporter vector in which the two cistrons (Renilla and Firefly luciferase) are separated by the *MYC* 5′UTR (pRF) was a kind gift from Dr. Willis and Dr. Thompson [18,19]. The empty control vector was provided by Dr. Thompson. The T7-SRSF1 and T7-SRSF1-NRS1 plasmids were a kind gift from Dr. Krainer [20–22]. The control pCGT vector was a gift from Winship Herr (Addgene plasmid # 99030; http://n2t.net/addgene:99030). The pRF_*CDK1*_5UTR bicistronic vector was generated by cloning the *ENST00000448257* 5′UTR (IDT DNA, gBlocks. Sequence used: *CTGGAATTCGCTGGCTCTTGGAAATTGAGCGGAGAGCGACGCGGTTGTTGTAGCTGCCGCTGCGGCCGCCGCGGAATAATAAGCCGGGTACAGTGGCTGGGGTCAGGGTCGTGTCTAGGGGACGGCCGAGGGCCTCGGAGGGCGAGTATTGAGGAACGGGGTCCTCTAAGAAGGCCGGACTGGAGGATCTACCATACCCATTGACTAACCATGGAAG*) into the pRF empty vector at the NcoI and EcoR1 restriction sites immediately upstream of the Firefly cistron. All plasmids were confirmed by sequencing.

Puromycin Incorporation Assay

Cells were transfected with the indicated siRNAs and grown in 6-well plates in full-serum or serum-deprived conditions for 24 h (SUM159PT cells) or full serum media with nocodazole (final concentration 0.4 μg/mL) for 16 h (T47D cells). Media was changed to puromycin-containing media to a final puromycin concentration of 10 μg/mL, and cells were incubated for 10 min. Puromycin functions as an amino acyl tRNA analog; it is incorporated into the C-terminal end of a nascent peptide during ribosomal elongation and induces premature termination of protein synthesis. The plates were then placed on ice, the media was aspirated, 250 μL of whole-cell cold lysis buffer (150 mM NaCl, 1% Triton X-100, 50 mM Tris pH 8) was added, and the adherent cells were scraped with prechilled scrapers and transferred to prechilled cytocentrifuge tubes. The cell suspension was then maintained in constant agitation at 4 °C for 30 min and then centrifuged at 12,000 g for 15 min also at 4 °C. The supernatant was aspirated into a new tube and protein concentration was quantified using the *DC* protein assay (BioRad, Hercules, CA) per manufacturer's instructions, and a Take3 plate with Synergy H1 reader (BioTek, Winooski, VT). Equal amounts of total protein from each sample were then separated by SDS-PAGE 10% and transferred onto 0.2-μm nitrocellulose membranes (BioRad, cat# 1620146). The membranes were stained first with ponceau S (Sigma–Aldrich; 0.1% in 5% glacial acetic acid) for 1 h, imaged, and then destained in water. To capture new protein synthesis, membranes were incubated with an anti-puromycin mouse monoclonal antibody (MABE343 clone 12D10, EMD Millipore). All images were captured using a G:Box Chemi XX6 (Syngene, Frederick, MD).

Luciferase assay

Stably transfected IRES reporter cells were treated in parallel in a 24-well plate as indicated. We stably transfected the same cell lines with the empty bicistronic vector for control. In the control cell lines, the firefly/Renilla luciferace activity ratio was several logs lower than that of the cell lines transfected with the bicistronic reporter vector. Lysates were prepared (1X passive lysis buffer), and firefly and Renilla luciferase activities were measured using the dual luciferase system as recommended by the manufacturer (Promega, cat# E1910). Luminescence measurements were obtained with a Glomax 20/20 luminometer (Promega, Madison, WI, USA).

Flow Cytometry and Cell Cycle Analysis

T47D cells growing in a 6-well plate and transfected with the respective siRNAs as indicated were trypsinized, washed in PBS, and fixed for 30 min at 4C with 70% ice cold ethanol. The cells were then washed again with PBS and stained with propidium iodine (200 μL of 50 μg/μL stock solution) for 20 min. Unstained cells constituted a

negative control. Flow cytometry was carried out in a LSRII Flow Cytometer (Becton Dickinson) using BD FACSDiva, version 8.0, (BD Biosciences), and results were analyzed using the ModFit LT software (Verity Software House, Topsham, Maine). Doublets or cell aggregates, which can compromise the results of cell cycle analysis, were discriminated on the basis of size. In our studies, doublets or aggregates constituted <1% (0.21−0.94%) of events.

Ultraviolet Cross-linking and Immunoprecipitation

Equal number of cells were seeded in 10-cm polystyrene cell culture dishes in full-serum media. Media was changed to full-serum media (control cultures) or serum-deprived media (SUM159PT) or nocodazole-containing (final concentration 0.4 μg/mL) media (T47D) for 24 and 16 h, respectively. At the end of the treatment, media was aspirated and cells were gently washed with ice-cold PBS. The plates were then irradiated at 1500 mJ/cm^2 once in ice-cold PBS in a HybriLinker HL-2000 (Analytik Jena AG, Jena, Germany). Cells were then scraped from the plates with a rubber policeman and transferred to prechilled tubes. The tubes were centrifuged at 800 g at 4 °C, and the supernatant was aspirated. The pellet was resuspended in ice-cold PBS, transferred to a new tube, and centrifuged at 2400 g at 4 °C for 1 min. The supernatant was aspirated, and the pellet was snap frozen in dry ice. The pellet was resuspended in lysis buffer (NaCl 100 mM, $MgCl_2$ 10 mM, Tris pH 7.6 30 mM, Triton X-100 0.5%, DTT 1 mM), homogenized with a 21-g needle and syringe 15 times, and incubated with 2 μL of Turbo DNase (Invitrogen, cat# AM2238) for 10 min at 37 °C and 800 RPM in a shaking incubator. Lysates were then spun at 13,000 g for 10 min at 4 °C. The supernatant was recovered and precleared by incubation with prewashed and resuspended in lysis buffer Dynabeads protein G (Invitrogen, cat# 10003D) for 1 h at 4 °C on continuous rotation. The precleared sample was then incubated with the anti-SRSF1 mouse monoclonal antibody (5 μg per sample, Invitrogen, Catalog #32−4500) and Dynabeads protein G (30 μL per sample) overnight at 4 °C on continuous rotation. Nonimmunoprecipitated supernatant was retained as control. The magnetic bead−immune complex was then washed 4 times with lysis buffer, reincubated with Turbo DNase for 10 min at room temperature, rewashed two more times, and resuspended in 45 μL of lysis buffer. Incubation with 2.5 μL of proteinase K (20 μg/μL) at 37 °C for 40 min ensued, followed by centrifugation at 12,000 g for 1 min. Total RNA from the supernatant was isolated and purified using the Qiagen miRNeasy kit following manufacturer's instructions. Equal volumes of purified RNA eluate were reverse transcribed and PCR amplified in parallel for the same number of cycles using the following primers:

Forward *MYC* 5′UTR primer: TAAGGTACCTAATTCCAGC GAGAGGCA.
Reverse *MYC* 5′UTR primer: GCGAGATCTGTCGCGGGAG GCTGCTGGTT.
Forward *CDK1* 5′UTR primer: TAAGAATTCGCTGGCTCTT GGAAATTGAG.
Reverse *CDK1* 5′UTR primer: GTGCCATGGTTAGTCAAT GGGTATGGTA.

The PCR products were then separated on 1% agarose gel and stained with SYBR gold (Invitrogen, cat#S11494), and images were captured using a G:Box Chemi XX6 (Syngene, Frederick, MD).

Antibodies

Antibody	Source	Identifier
Mouse monoclonal anti-SRSF1 (clone 96)	ThermoFisher Scientific	32−4500
Mouse monoclonal anti-tubulin (clone B-5-1-2)	Sigma Aldrich	T6074
Mouse monoclonal anti-c-myc (clone 9E10)	Santa Cruz Biotechnology	sc-40
Mouse monoclonal anti-puromycin (12D10)	EMD Millipore	MABE343
Mouse monoclonal anti-CDK1 (clone 17)	Santa Cruz Biotechnology	sc-54
Mouse monoclonal anti-CDKN1A/p21 (clone F-5)	Santa Cruz Biotechnology	sc-6246
Rabbit polyclonal anti-calreticulin	ThermoFisher Scientific	PA3-900
Mouse monoclonal anti-RACK1 (clone 20)	BD Biosciences	610177
Rabbit polyclonal anti-Climp-63	Abcam	ab152154
Goat polyclonal HPR Goat Anti-Rabbit IgG	BD Pharmingen	554021
Goat polyclonal HPR Goat Anti-Mouse Ig	BD Pharmingen	554002

Statistical Considerations

Genomic data were analyzed using statistical packages in R. Statistical tests and numbers of replicates (*n*) are indicated in the figures and figure legends. They are also summarized as follows:

1) Correlations in gene expression were calculated using the Spearman correlation.
2) Translational efficiencies between HMECs vs. nontumorigenic (MCF10A) and breast tumor cell lines were compared using the two-sample Wilcoxon test.
3) The variances of the translational efficiencies were compared using the Fligner−Killeen test (for multiple comparisons) and Fisher's F test (pairwise comparisons).
4) Differential gene expression and principle component analysis were conducted using the DESeq2 statistical package in R [23]. The regularized-logarithm transformation was used, and *P* values for the differentially expressed genes were adjusted by the Bergamini-Hochberg method for multiple comparisons.

Non-genomic data analyses were performed using GraphPad Prism, version 6. Data are presented as means ± standard deviation. For comparisons between two groups, we used two-tailed paired Student t-tests. No statistical method was used to predetermine the sample size. The investigators were not blinded to allocation during experiments and outcome assessment.

Results

Correlations Between mRNA Abundance, Ribosome Protected Fragments, and Ribosome Occupancy

We profiled six cell lines representing tumor subtypes (estrogen receptor [ER]−positive: T47D and ZR75-1; and ER-negative: SUM159PT and MDA-MB-231) arising from the same tissue, i.e., mammary and known to be fundamentally distinct in terms of natural history, therapeutic strategies, and clinical outcomes. Translational profiles of a nonmalignant (human mammary epithelial cells [HMECs]) and a nontumorigenic (MCF10A cells) counterparts were also analyzed.

We performed at least two biological replicates and achieved good reproducibility for both the RNA and ribosome-protected fragment (RPF) sequencing reads (Supplemental Figure 1A). The fragment length distribution of the RPF reads follows a very consistent pattern across all cell lines and conditions [24] (Supplemental Figure 1B).

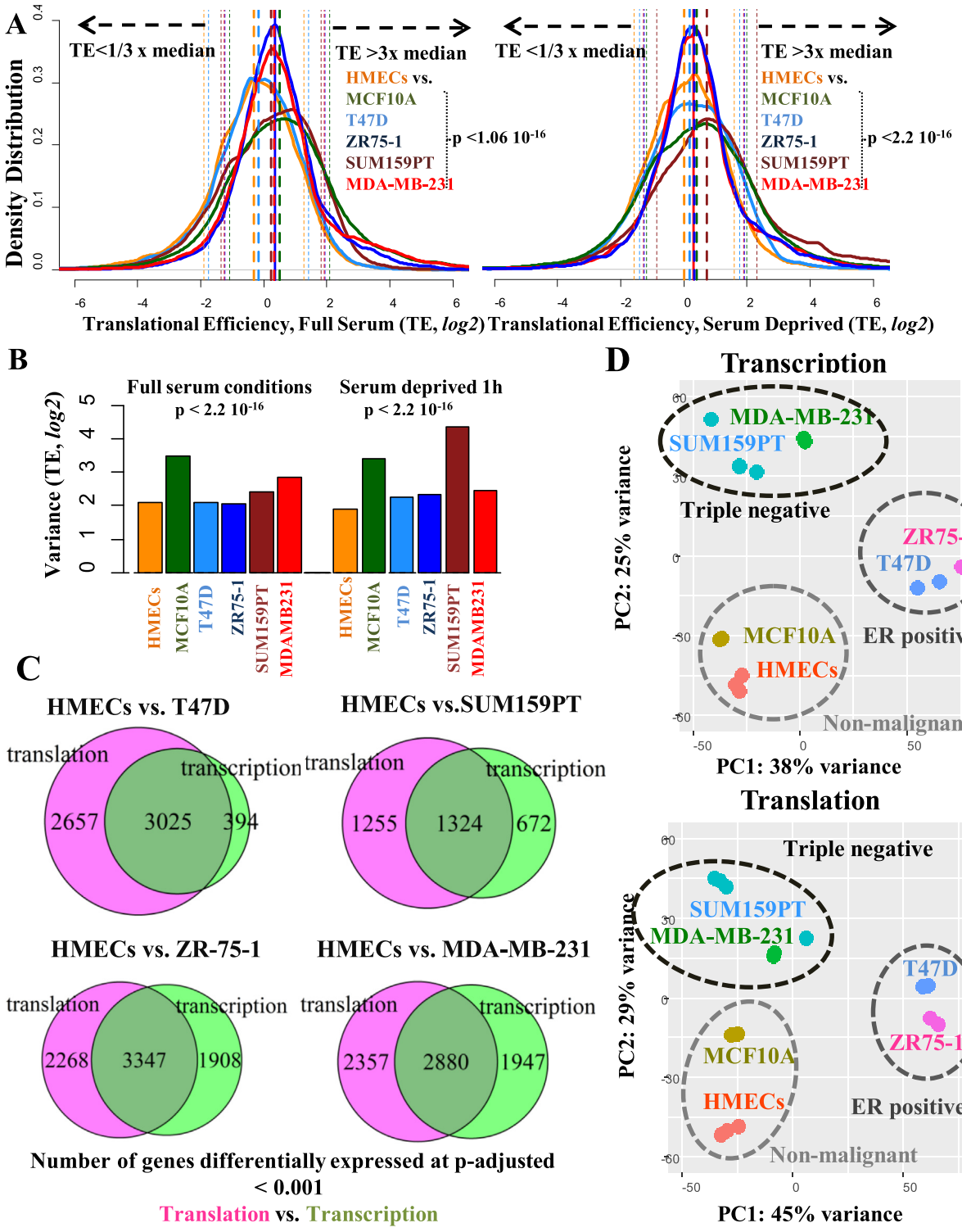

Figure 1. Correlations between mRNA abundance, ribosome-protected fragments, and ribosome occupancy. (A) Density plots color-coded for each cell line of the log2 transformed translational efficiencies (TEs) of protein-encoding genes under normal growth (left panel) and serum-deprived (charcoal stripped serum for the ER + cell lines; right panel) conditions for 1 h conditions. Thick dashed lines mark the median TE for each cell line; thin dashed lines mark the 0.333x and 3x median TE points along the x axis for each cell line. The median TE of the malignant cells (including MCF10A) was higher than that of the nonmalignant HMECs (Wilcoxon Rank Sum test). Although the TE for most genes falls in proximity to the median TE for each cell line, up to one-third of the genes were found to have TE that varied from the median by a factor greater than 3. (B) Variance of the log2 transformed translational efficiencies in malignant and nonmalignant cell lines in normal growth (full-serum, left panel) and stress (serum-deprived; charcoal stripped serum for the ER + cell lines, left panel) conditions. The dispersion of translational efficiencies is greater in malignant cells especially of the ER-negative cell lines and under conditions of stress than in HMECs (Fligner—Killeen test $P < 2.2\ 10^{-16}$). In pairwise comparisons of the variances of the log2 transformed translational efficiencies between HMECs vs. MCF10A, T47D, ZR75-1, SUM159PT, and MDA-MB-231 under normal growth and stress conditions (Fisher's F test), the variances

Metagene analyses [13] produced a peak at the translation initiation and termination sites of the annotated coding sequences consistent with a profile of active translation [12,25] (Supplemental Figure 2). Moreover, we consistently appreciated a second peak on the 5th codon downstream of the translation initiation site; this recognized postinitiation ribosomal pausing has been proposed as an inherent signature of the translation machinery to ensure productive translation [26].

Our results show that the median ribosome occupancy of the mRNAs of protein-encoding genes is consistently higher in malignant cells than in HMECs (Figure 1A and Supplemental Figure 3A). Although ribosome profiling is not the optimal technique to capture global changes in translation, i.e., a global upregulation or downregulation in protein synthesis [27], our observations suggest that in breast cancer besides the upregulation of virtually all components of the translational machinery [5] including the ribosomes [28], higher translational efficiency of the existing pool of mRNAs can also account for the global upregulation of protein synthesis.

We find a good correlation between RPFs and their respective mRNA levels (Supplemental Figure 3B, Spearman correlations 0.88−0.942). Accordingly, the translational efficiencies of most protein-encoding genes were in proximity to the median translational efficiency of each respective cell line (Figure 1A). This finding is consistent with reported results from model genes [29] and suggests that transcription and translation are for the most part coupled processes. However, the variance of translational efficiencies is greater in malignant cells, especially in ER-negative cell lines and under conditions of stress as compared with HMECs (Figure 1B). In addition, with the exception of the nonmalignant HMECs and the triple-negative MDA-MB-231 cells, the variance of translational efficiencies was greater under conditions of stress than under normal growth conditions. The greater dispersion of translational efficiencies after short (1 h) exposure to stress conditions is consistent with the fact that changes in translation allow the cell to adapt swiftly to stress conditions by modulating protein synthesis from an existing transcriptional template [3,30]. The physiologically relevant stress condition tested was relative growth factor deprivation, achieved by restricting serum concentration (HMECs, MCF10A, SUM159PT, and MDA-MB-231) or restricting steroidal hormone concentration (ER-positive cell line models: T47D and ZR75-1).

Discordant differential gene expression at the level of transcription and translation

To identify genes that are differentially expressed at the transcriptional (mRNA) and translational (RPF) level, we performed differential gene expression analyses using the DESeq2 package in R [23] comparing HMECs with the respective malignant counterparts.

The results show that the number of genes that are differentially expressed at the translational level is consistently higher than at the transcriptional level (Figure 1C). There is only a partial concordance in the genes that are differentially expressed at the transcriptional and translational level. For a level of statistical significance of P (adjusted for multiple comparisons by the Bergamini-Hochberg method) of <0.001, up to 48.7% of genes differentially expressed at the level of translation, are not differentially expressed at the level of transcription. In accord with this result, principle component analysis on the basis of unique reads of both protein-encoding (Figure 1D) as well as all genes (Supplemental Figure 3C) shows that the variances across principle component 1 and 2 are greater at the translational level than at the transcriptional level. The coordinate clustering of triple-negative, ER-positive, and nonmalignant cell lines suggests that despite the greater variation in gene expression conferred by translation, each cell line falls into a cluster defined by the tumor subtype it represents.

The Expression of Many Genes and Transcripts Involved in or Regulating Cancer-relevant Processes is Upregulated at the Level of Translation

To dissect further the partial discordance in differential gene expression at the transcriptional vs. translational level, we performed ontology analyses of the genes and transcripts that are transcribed in both malignant and nonmalignant cells but are preferentially translated in malignant cells by a threefold or greater efficiency. We reasoned that the preferential translation of certain groups of genes and transcripts may underpin the greater differences in gene expression at the level of translation.

Using the translational efficiencies of the assembled transcripts, our analyses show that the transcripts commonly transcribed in HMECs and malignant cells but preferentially translated in malignant cells fall into consistently similar ontology categories across all malignant cell lines evaluated: transcription, signal transduction, regulation of apoptosis, cell division, cell adhesion, DNA repair, and translation itself. These ontologies are highly pertinent for cancer-related processes (Figure 2).

We obtained similar results using the translational efficiencies at the level of genes (Supplemental Figure 4). The ontologies of genes preferentially translated in cancer were remarkably similar across all malignant cell lines irrespective of their ER status: transcription and transcriptional regulation, signaling, cell adhesion, and immune or inflammatory response. The clustering of these genes and transcripts in the upper left quartile of the ribosome footprint vs. mRNA

were statistically significantly different ($P < 2.2\ 10^{-16}$) except between HMECs vs. T47D (P, 0.7228) and HMECs vs. ZR75-1 (P, 0.059) under normal growth conditions. In pairwise comparisons of the variances of the log2 transformed translational efficiencies of the same cell line under normal growth and stress conditions (Fisher's F test), the variances were statistically significantly different ($P < 1.98\ 10^{-7}$) except in MCF10A cells (P, 0.214). (C) Venn diagrams of the differentially expressed protein-encoding genes at the level of transcription (green, mRNA reads) vs. translation (magenta, RPF reads) for a level of significance, P-adjusted (adjusted for multiple comparisons by the Bergamini-Hochberg method) < 0.001. Each comparator arm has the same number of biological replicates: HMECs vs. T47D, $n = 2$; HMECs vs ZR75-1, $n = 2$; HMECs vs. SUM159PT, $n = 3$; HMECs vs MDA-MB-231, $n = 2$. Nearly one-half of genes that are differentially expressed at the level of translation are not differentially expressed at the level of transcription. (D) Principle component analysis of mRNA (upper panel) and RPF (lower panel) unique reads of protein encoding genes (HMECs $n = 3$; MCF10A $n = 2$; T47D $n = 2$; ZR75-1 $n = 2$; SUM159PT $n = 4$; MDA-MB-231 $n = 2$). The variances across principle component 1 and 2 are greater at the translational than at the transcriptional level. Despite the greater variation in gene expression conferred by translation, each cell line falls into a cluster defined by the cancer subtype it represents. HMECs, Human mammary epithelial cells.

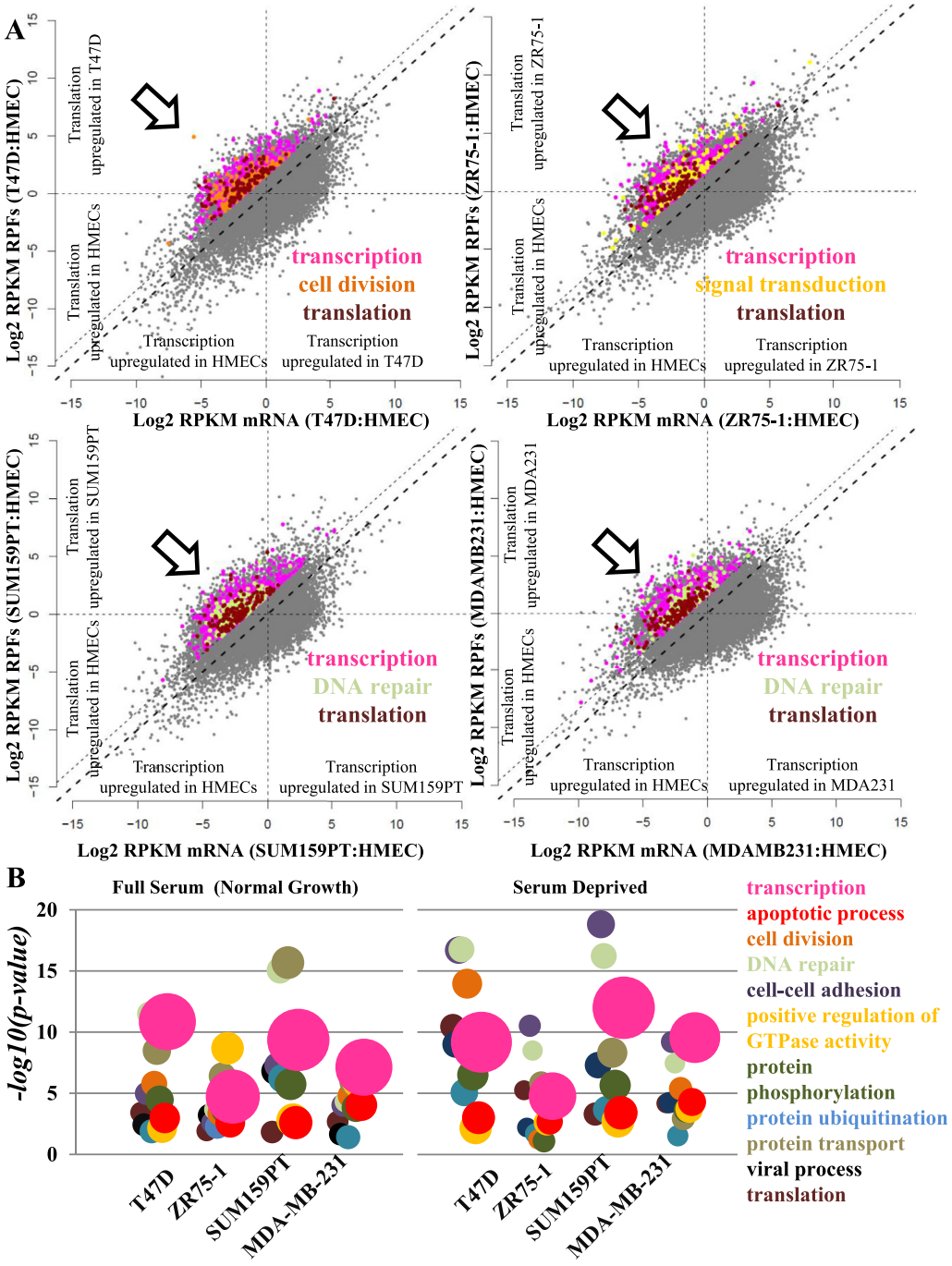

Figure 2. The expression of many transcripts involved in or regulating cancer-relevant processes is upregulated at the level of translation. (A) Dot plots of transcript RPF (y-axis, expression of translation) and mRNA RPKM (reads per kilobase of exon model per million mapped reads) ratios (x-axis, expression of transcription) between ER(+) (T47D and ZR75-1) and triple-negative (SUM159PT and MDA-MB-231) breast cancer cells vs. nonmalignant counterpart (HMECs). Transcripts upregulated at the transcriptional level in malignant cells are in the right quartiles; transcripts upregulated at the translational level in malignant cells are in the upper quartiles; transcripts coordinately upregulated or downregulated at the transcriptional and translational level in malignant cells lie along the mid-diagonal line. The thick dashed line marks the mid-diagonal line; the thin dashed line marks the cutoff of threefold or greater translational efficiency in malignant cells vs. HMECs. Note the clustering of translationally upregulated transcripts involved in transcription, cell division, signal transduction, DNA repair, and translation in the upper left quartile (empty arrow, dots color-coded on the basis of ontology). For many of these transcripts, expression is upregulated at the translational level without always a coordinate upregulation at the transcriptional level. (B) Bubble plot of the ontology of transcripts whose translational efficiency is at least 3 times greater in malignant cells than in HMECs under normal growth and serum-deprived (charcoal-stripped serum for T47D and ZR75-1 cells) conditions (left and right panel, respectively). Bubbles are color-coded by ontology category. The size of each bubble is proportionate to the number of transcripts falling into this category. Transcript ontologies are plotted against the significance of gene-term enrichment (-log10(*P*-value), y-axis). Translationally upregulated transcripts fall into remarkably similar ontology categories across all malignant cell lines tested. HMECs, Human mammary epithelial cells.

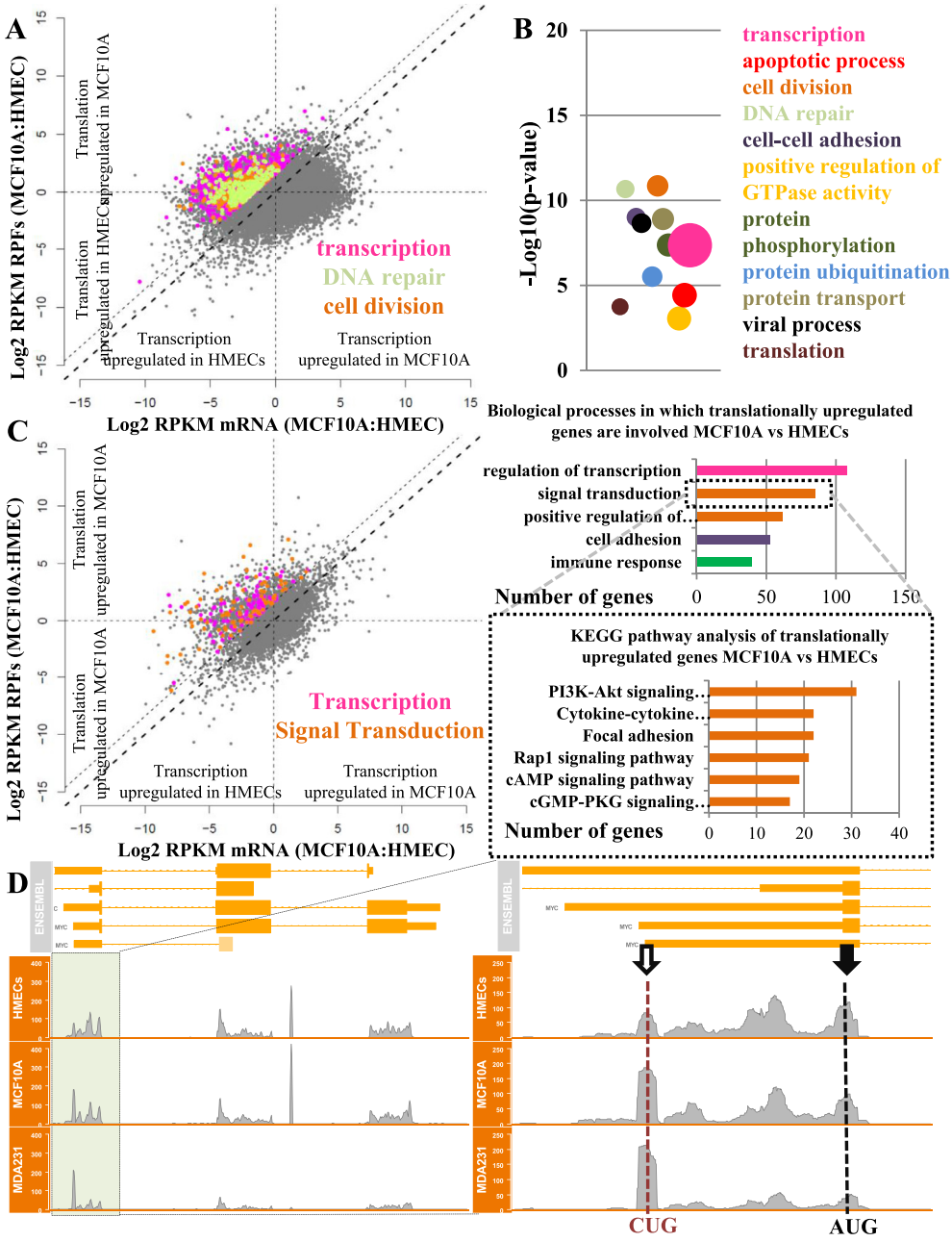

Figure 3. The translational landscape of MCF10A cells aligns with the profile of its malignant counterparts. (A) Dot plots of transcript RPF (y-axis, expression of translation) and mRNA RPKM ratios (x-axis, expression of transcription) between the nontransformed MCF10A cell line and HMECs. Thick dashed line marks the mid-diagonal line; thin dashed line marks the cutoff of more than threefold translational efficiency in MCF10A cells vs. HMECs. Translationally upregulated transcripts involved in transcription, DNA repair, and cell division are highlighted. For many of these transcripts, expression is upregulated at the translational level without a coordinate upregulation at the transcriptional level. (B) Bubble plot of the ontology of transcripts whose translational efficiency is $\geq$ 3x greater in MCF10A cells than in HMECs. Bubbles are color-coded by ontology category. The size of each bubble is proportionate to the number of transcripts falling into this category. Transcript ontologies are plotted against the significance of gene-term enrichment (-log10(*P*-value), y-axis). The ontologies of the translationally upregulated transcripts are very similar between MCF10A and malignant cells. (C) Dot plots of gene RPF (y-axis, expression of translation) and mRNA RPKM ratios (x-axis, expression of transcription) between the nontransformed MCF10A cell line and HMECs (left panel). Translationally upregulated genes involved in transcription and signal transduction are highlighted. For many of these genes, expression is upregulated at the level of translation without a coordinate upregulation at the level of transcription. The ontologies of genes commonly transcribed in MCF10A cells and HMECs but preferentially translated in MCF10A cells by $\geq$ 3x greater efficiency are shown in the right panel. KEGG pathway analysis of the subgroup of genes involved in signaling shows that these genes are mostly involved in the PI3K-AKT pathway. (D) Genomic tracks for *MYC* (upper panel) of the RPFs in HMECs, MCF10A, and MDA-MB-231 (MDA231) cells in descending order (lower panel). The alignments along the entire gene are shown in the left panel, while a focused view of the 5′UTR and first exon is shown in the right panel. A peak of RPFs surrounding an upstream noncanonical CUG translation initiation site dominates over the canonical AUG peak in the MCF10A and MDA-MB-231 cells in contradiction to HMECs. Graph generated using the same number of representative rRNA/mtRNA/tRNA/noncoding RNA/PhiX depleted RPF reads. HMECs, Human mammary epithelial cells; RPF, ribosome-protected fragment.

abundance plots in Figure 2 (and Supplemental Figure 4) denotes that the translational upregulation of these transcripts and genes is not always associated with a coordinate change of their respective mRNA levels. Collectively, these results are consistent with the literature that suggests that cancer cells preferentially translate genes and transcripts that support the phenotypic hallmarks of malignancy [31].

The Translational Landscape of MCF10A Cells Aligns With the Profile of its Malignant Counterparts

We included MCF10A cells in our studies as a second nontumorigenic cell line alongside HMECs. MCF10A cells have been used extensively to study mammary morphogenesis [32,33]. They have been traditionally thought to represent a nontransformed cell line model of mammary tissue; however, their reliability to do so has been questioned [34]. These ER-negative cells were originally derived from human fibrocystic mammary tissue and spontaneously immortalized without viral or chemical intervention [35]. MCF10A cells do not express estrogen receptor [34,35], and their known molecular characteristics include the depletion of the chromosomal locus containing the p16 and p14ARF genes (both of which are critical in regulating senescence), and amplification of *MYC* [34,35]. HMECs, on the other hand, are isolated from adult female breast tissue with no apparent disease in the context of reduction mammoplasties. HMECs become progressively senescent with repeated passages [36, 37].

Our results show that at the translational level, MCF10A cells resemble breast tumor cells in many respects and more specifically ER(−) cells : (1) their median ribosome occupancy is greater than that in HMECs and approximates that of malignant counterparts (Figure 1A and Supplemental Figure 3A); (2) the variance of the log2 transformed translational efficiencies is similar to the variance of the triple-negative breast cancer counterparts (Figure 1B); (3) MCF10A cells upregulate the translation of transcripts and genes involved in the same cancer-relevant processes as their malignant counterparts (Figure 3A−C). In fact, pathway analysis of the translationally upregulated genes involved in signaling shows that these genes are mostly involved in the PI3K-AKT pathway; (4) while by principle component analysis on the basis of RNA reads, MCF10A cells cluster in proximity to the HMECs, on the basis of RPFs, MCF10A cells approach the triple-negative MDA-MB-231 cells (Figure 1D and Supplemental Figure 3C).

Moreover, studies have shown that during tumor initiation, the translational apparatus is redirected toward noncanonical upstream initiation sites, enhancing the translational efficiency of oncogenic mRNAs [38]. This appears to be the case with the *MYC* mRNA: peaks of RPFs that could represent translation initiation sites [12,25] are differentially distributed along the *MYC* 5′UTR among HMECs, MCF10A, and MDA-MB-231 cells (Figure 3D). A peak surrounding a noncanonical CUG translation initiation site dominates over the canonical AUG peak in the MCF10A and MDA-MB-231 cell lines in contradiction to HMECs. Although the mechanism and the implications of this shift of translation initiation on oncogenesis have been established, the impact on the proteome remains elusive [38].

Collectively, these results suggest that MCF10A cells have acquired features of oncogenic translation while their transcriptome still resembles the transcriptome of HMECs. Altogether, ribosome profiling data suggest that gene expression may be significantly skewed at the level of translation and discordantly to transcription.

Transcripts whose Translation is Preferentially Upregulated in Cancer Harbor Recurrent Motifs in their 5′UTRs

The majority of translational regulation occurs at the level of initiation [3]. To gain a mechanistic insight into the preferential diversion of the translational apparatus in breast cancer, we sought to identify whether the transcripts commonly expressed in HMECs and malignant cells but preferentially translated in malignant cells by a threefold or greater efficiency harbor certain motifs in their 5′UTRs. We then sought to identify candidate RNA-binding proteins that would potentially associate with those motifs motivated by the hypothesis that such RNA-binding proteins may contribute to the aberrant translational upregulation of their client transcripts. We also performed the reverse analysis, whereby we sought to identify which transcripts among the translationally upregulated ones harbor the highest ranked matching motifs.

Our analyses show remarkable similarities in the overrepresented motifs across all malignant cell lines including the nontransformed MCF10A cells (Supplemental Figure 5A and Supplemental Table 1). In surveying candidate RNA-binding proteins that could associate with the retrieved motifs, two proteins ranked the highest and consistently across all malignant cell lines, including MCF10A cells: eIF4B and SRSF1 (Figure 4A). This consistency suggests that an at least partially shared set of RNA-binding proteins may be responsible for the translational dysregulation in breast cancer. The limited overlap between translationally upregulated transcripts harboring eIF4B- and SRSF1-binding motifs (Figure 4B and Supplemental Figure 5B) suggests that eIF4B and SRSF1 may upregulate the translation of different groups of transcripts. The translation of some transcripts however may be dysregulated by either or both SRSF1 and eIF4B. This is consistent with the notion that translational dysregulation in cancer may result from multiple mechanisms operating in parallel rather than a single one. Multiple translationally upregulated transcripts encoding for proteins with unequivocal relevance to Breast Cancer Biology harbor eIF4B- and SRSF1-binding motifs (Figure 4C). There was also an overlap between the translationally upregulated transcripts harboring eIF4B- and SRSF1-binding motifs in T47D and SUM159PT cells despite the substantially different breast tumors that they represent (Figure 4C). We should acknowledge that for many retrieved motifs, no known cognate RNA-binding protein was identified and the presence of a motif in a transcript does not imply association.

The role of eIF4B in modulating the translational landscape has been explored [39]. eIF4B stimulates the RNA helicase activity of eIF4A in unwinding secondary structures in the 5′UTR. In doing so, eIF4B enables cancer cells to selectively translate transcripts with structured 5′UTRs; these transcripts encode for proteins involved in cell proliferation and survival [40]. Indeed, the 5′UTRs of the translationally upregulated mRNAs with eIF4B-binding motifs have consistently lower minimum free energies than the minimum free energies of the 5′UTRs of all translationally upregulated transcripts (Figure 4D and Supplemental Figure 5C). In accordance with the literature [39], the translationally upregulated mRNAs with eIF4B--binding motifs in their 5′UTRs encode for antiapoptotic proteins and positive regulators of cell division (Figure 4C). The upregulation of translation of many of these transcripts by eIF4B has been previously experimentally shown [39]. In addition, our results expand the repertoire of ontologies of these transcripts (Figure 4C).

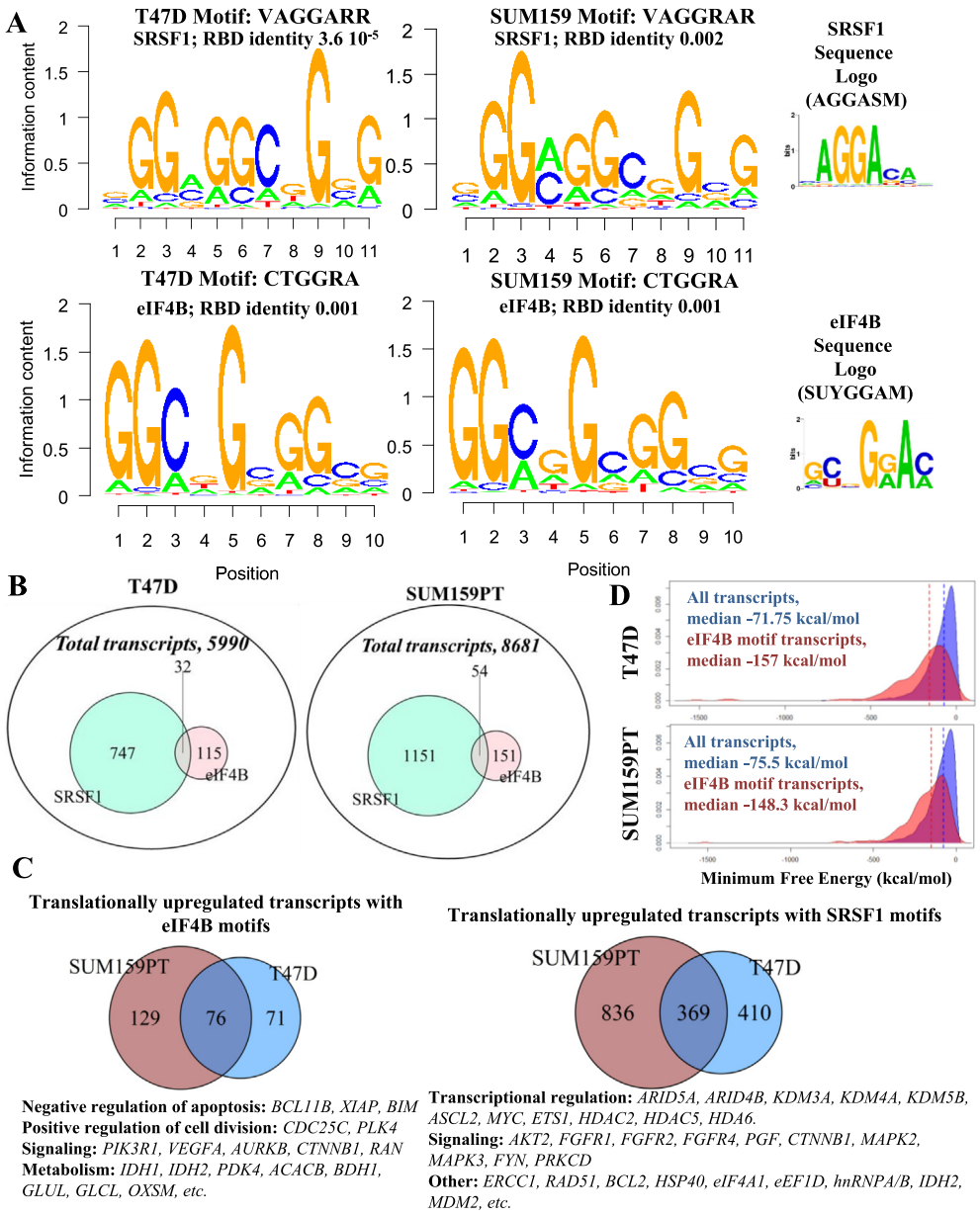

Figure 4. The 5′UTRs of transcripts that are preferentially translated in tumour cells harbor common motifs which have high RBD identity scores to SRSF1 and eIF4B binding motifs. (A) Discriminative motif analysis of the 5′UTRs of transcripts that are preferentially translated in T47D and SUM159PT cells. Note the very similar VAGGARR and VAGGRAR motifs and identical CTGGRA motifs retrieved from this analysis in T47D and SUM159PT cells with high RNA-binding domain (RBD) identity scores to SRSF1 and eIF4B-binding motifs, respectively. The sequence contexts in the 5′UTRs where these motifs are found are shown below the RBD identity scores as well as the sequence logos (right) of the SRSF1 and eIF4B RNA-binding motifs (AGGASM and SUYGGAM, respectively). Note the GG-rich context in which these motifs are found. (B) Venn diagrams of the total number of translationally upregulated transcripts vs. translationally upregulated transcripts harboring SRSF1- and eIF4B-binding motifs. Note the significantly larger number of translationally upregulated transcripts harboring SRSF1- vs. eIF4B-binding motifs and the small overlap between the two subgroups. (C) Venn diagrams of the transcripts harboring eIF4B and SRSF1 motifs in their 5′UTRs and whose translation is at least 3 times greater in the T47D and SUM159PT cells than in HMECs. Multiple translationally upregulated transcripts encoding for proteins with unequivocal relevance to Breast Cancer Biology harbor eIF4B- and SRSF1-binding motifs. The upregulation of translation of transcripts encoding for proteins that promote proliferation (*CDC25C*) and inhibit apoptosis (*XIAP*) by eIF4B has been previously shown (Ref 39). Note the overlap between the translationally upregulated transcripts harboring eIF4B and SRSF1 binding motifs in T47D and SUM159PT cells despite the substantially different breast tumors that they represent. (D) Density plots of the minimum free energy of the 5′UTRs of all transcripts (blue) vs. transcripts that harbor an eIF4B-binding motif (red). Medians shown in color-coded dotted lines. The translationally upregulated mRNAs with eIF4B-binding motifs in their 5′UTRs have consistently lower minimum free energies than the minimum free energies of the 5′UTRs of all translationally upregulated transcripts. This result is consistent with the notion that eIF4B enables tumour cells to translate transcripts with structured 5′UTRs.

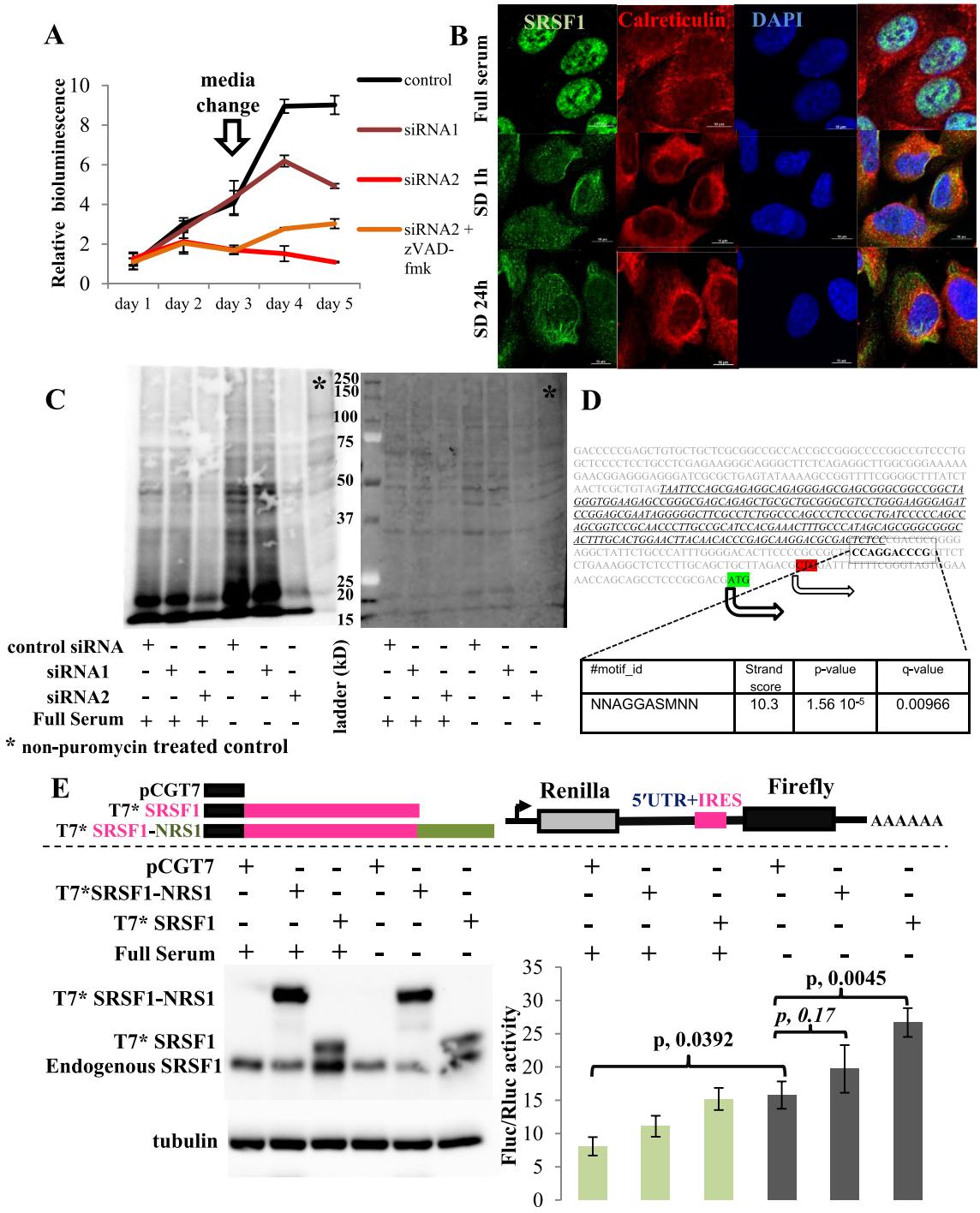

Figure 5. The role of SRSF1 on translation in a representative ER(−) breast cancer cell line (SUM159PT). (A) Cell viability assay of SUM159PT cells. Downregulation of SRSF1 with siRNAs is associated with decreased viability of SUM159PT cells and apoptosis. Cell death is prevented with a pan-caspase inhibitor, zVAD-fmk (data presented as mean ± standard deviation, $n = 3$). (B) Immunofluorescent confocal microscopy of SRSF1 (green) and calreticulin (red, a component of the endoplasmic reticulum). Upon exposure to conditions of relative growth factor deprivation (serum-deprived conditions 1 h and 24 h; SD 1 h and SD 24 h, respectively), SRSF1 promptly translocates from the nucleus to the cytoplasm. Scale bars: 10 μm. (C) Puromycin incorporation assay (left panel) to capture global new protein synthesis. Ponceau S staining (right panel) serves as the control to confirm equal protein loading. Downregulation of SRSF1 is associated with downregulation of global new protein synthesis under both full-serum and serum-deprived (stress) conditions. (D) The *MYC* 5′UTR harbors an SRSF1-binding site (AGGASM) downstream and in proximity to its IRES (in italics and underlined). The 2 alternative translation initiation sites are illustrated. (E) Immunoblot for SRSF1 (left lower panel) and luciferase assay (right lower panel) of SUM159PT *MYC* IRES reporter cells. SUM159PT cells stably transfected with the *MYC* IRES bicistronic reporter vector (right upper panel) were transiently transfected with empty pCGT7 vector or the same vector containing N-terminus T7 tagged SRSF1 or T7 tagged SRSF1 fused with the NRS1 sequence of SRSF2 at the C-terminus (left upper panel). IRES activity is defined as the ratio of Firefly:Renilla luciferase (data presented as mean ± standard deviation, $n = 2$; paired Student's t-test).

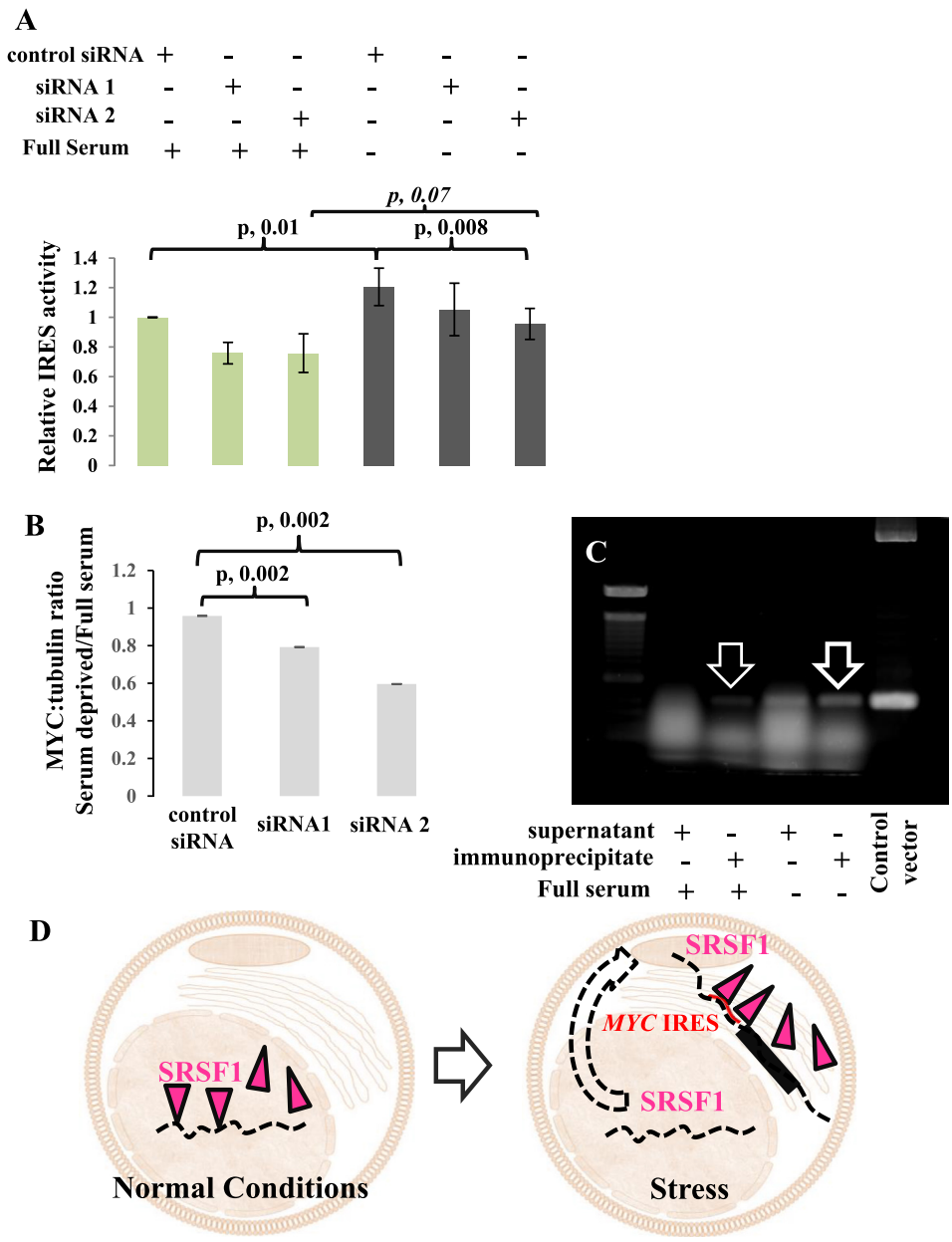

Figure 6. Cytoplasmic SRSF1 operates as an ITAF for *MYC* IRES. (A) Luciferase assay of SUM159PT *MYC* IRES reporter cells. SRSF1 expression was downregulated with siRNAs and cells were grown under full-serum or serum-deprived (for 24 h) conditions as indicated. siRNA mediated downregulation of SRSF1 is associated with a decrease in the *MYC* IRES activity and leads to smaller increments of the *MYC* IRES activity under conditions of stress. IRES activity (ratio of Firefly:Renilla luciferase) is normalized to control siRNA full-serum condition (data presented as mean ± standard deviation, $n = 5$; paired Student's t-test). (B) MYC band intensity in an immunoblot (Supplemental Figure 9) was divided by the respective tubulin band intensity to provide the relative MYC protein levels. The histogram illustrates the ratio of the relative MYC protein levels between serum-deprived (stress) and full-serum conditions (mean ± standard deviation, $n = 2$; paired Student's t-test). SUM159PT cells in which SRSF1 is downregulated cannot maintain MYC protein expression under conditions of stress. (C) UV CLIP assay of SUM159PT *MYC* IRES reporter cells for SRSF1 under full-serum and serum-deprived (for 24 h) conditions. SRSF1 directly associates with the *MYC* 5′UTR and this association is more pronounced under serum deprived conditions. Nonimmunoprecipitated supernatant and the vector used for the stable transfection serve as control. (D) Model for the translational activity of SRSF1 in SUM159PT cells: under normal growth conditions, SRSF1 is predominantly intranuclear and operates as a splicing factor leaving its splicing "fingerprint" on the transcriptome. Under conditions of stress, SRSF1 translocates to the cytoplasm where it operates as an IRES *trans*-acting factor and upregulates the IRES mediated translation of *MYC* mRNA (and probably of other transcripts).

SRSF1, on the other hand, is a splicing factor and a known oncogene in breast cancer through its activity in alternative splicing [20,41]. How SRSF1 contributes to the dysregulation of the translational landscape is less well understood. We reasoned that the impact of SRSF1 on translation may be direct through mRNA associations [42] or indirect by splicing in or out translational enhancer or repressor elements, respectively. In the latter case, the retrieval of its binding motif in our analyses may represent SRSF1's splicing signature. Motivated by this gap in knowledge and by the observation that approximately one-eighth (potentially more given

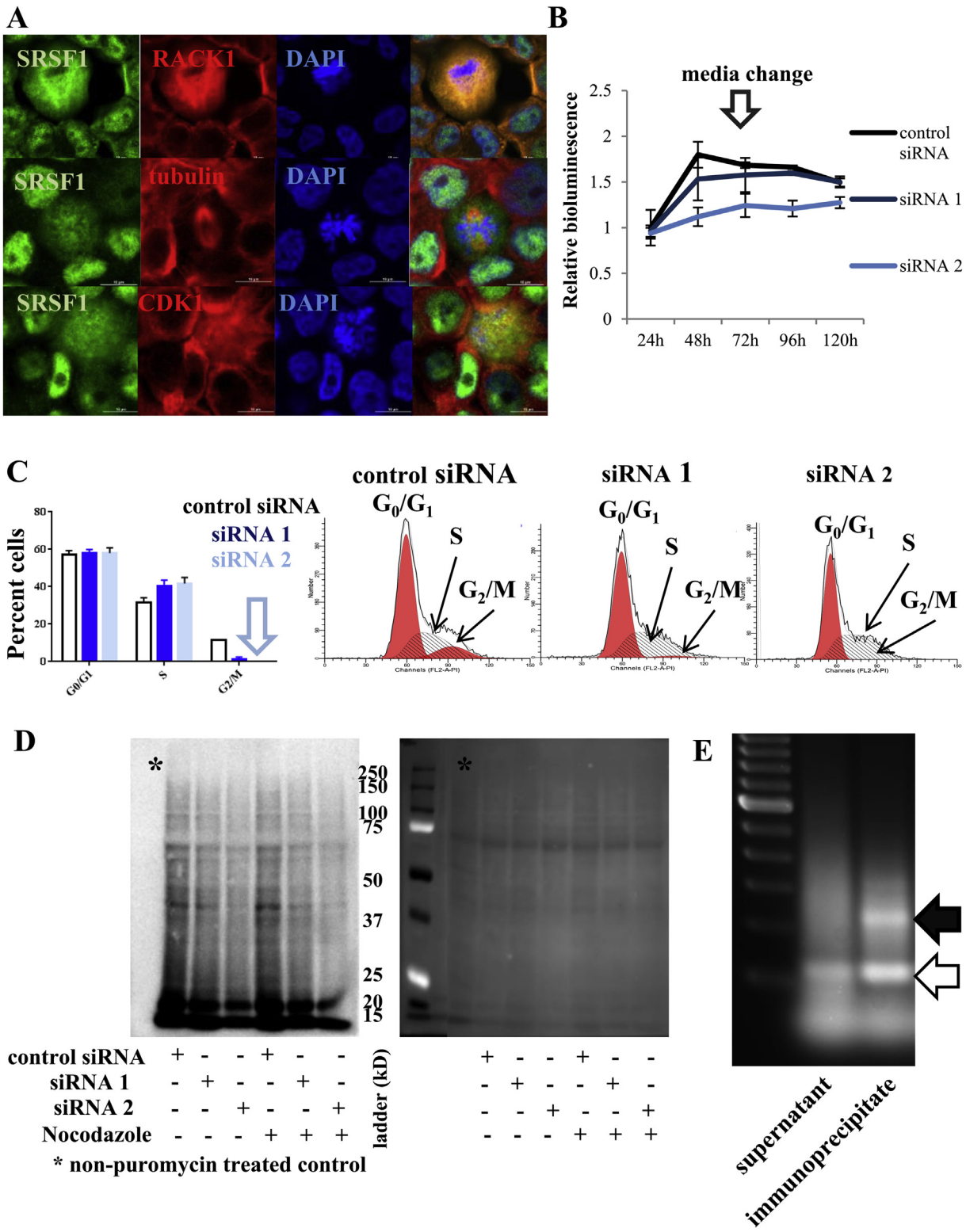

Figure 7. The role of SRSF1 on translation in a representative ER(+) breast cancer cell line (T47D). (A) Immunofluorescence confocal microscopy of T47D cells. Images focused on dividing cells surrounded by nondividing cells. In T47D cells, SRSF1 is predominantly intranuclear. In dividing cells, however, SRSF1 localizes diffusely in the cytoplasm. SRSF1 colocalizes with RACK1 (a component of the IRES translation initiation complex, upper panel). Note the mitotic spindle (middle panel) and the diffusely localized CDK1 and SRSF1 in the dividing cells (lower panel). Scale bars: 10 μm. (B) Cell viability assay of T47D cells. siRNA-mediated downregulation of SRSF1 is associated with growth arrest in T47D cells (data presented as mean ± standard deviation, $n = 3$). (C) Cell cycle analysis by flow cytometry of T47D cells. siRNA mediated downregulation of SRSF1 in T47D cells is associated with an S-phase arrest (furthest left, data presented as mean ± standard deviation, $n = 2$). Note the nearly 0% cells in G2/M phase with siRNA2. A representative flow cytometry histogram is shown with each siRNA on the right. (D) Puromycin incorporation assay (left panel). Ponceau S staining (right panel) serves as control to confirm equal protein loading. Downregulation of SRSF1 is associated with a global

the intrinsic flexibility of SRSF1 in RNA motif recognition [43]) of the translationally upregulated transcripts harbored SRSF1-binding motifs (Figure 4B and Supplemental Figure 5B), we sought to dissect the role of this RNA-binding protein in a representative ER-negative (SUM159PT) and ER-positive (T47D) breast cancer cell line.

The Role of SRSF1 on Translation in a Representative ER(−) Breast Cancer Cell Line

In dissecting the role of SRSF1 in SUM159PT cells, we first attempted to delete the gene using a CRISPR-Cas9 system (Supplemental Figure 6A). Our efforts were unsuccessful; the surviving clones uniformly expressed SRSF1 at levels comparable with controls (Supplemental Figure 6B). This negative result suggested that SRSF1 is a critical gene for triple-negative breast cancer cells. Indeed, downregulation of SRSF1 with siRNAs was associated with decreased viability and apoptosis as cell death was prevented with a pan-caspase inhibitor, zVAD-fmk (Figure 5A and Supplemental Figure 7). We made the same observation with MDA-MB-231 cells as well (Supplemental Figure 8A). The impact of SRSF1 downregulation on the viability of SUM159PT cells was even more pronounced under conditions of stress (Supplemental Figure 6C).

SRSF1 shuttles between the nucleus and the cytoplasm where it is known to operate on splicing and translation, respectively [44]. To identify the subcellular localization of SRSF1 and consequently the predominant mechanism by which it modulates translation, we performed confocal immunofluorescence microscopy. Our studies show that under normal growth conditions SRSF1 predominantly localizes to the nucleus; under conditions of stress though, it promptly translocates to the cytoplasm (Figure 5B). Irrespective of its subcellular localization and growth conditions, SRSF1 has a strong impact on global new protein synthesis in our triple-negative breast cancer cells (Figure 5C, Supplemental Figure 8B).

To dissect the direct effect of SRFS1 on translation, we selected a representative mRNA with an SRSF1-binding motif in its 5′UTR. The *MYC* 5′UTR harbors an SRSF1 binding motif which lies in proximity and downstream of the *MYC* internal ribosome entry site (IRES, Figure 5D). We hypothesized that, through a direct association, SRSF1 may modulate the *MYC* IRES activity, essentially operating as an IRES *trans*-acting factor (ITAF). To investigate this hypothesis, we used SUM159PT cells stably transfected with a bicistronic reporter vector where the two cistrons are separated by the *MYC* 5′UTR [18,19] (Figure 5E). Our studies show that under normal growth conditions, overexpression of SRSF1 is associated with a mild increase in the *MYC* IRES activity. Consistent with the literature [45], under conditions of stress, *MYC* IRES activity is upregulated. This upregulation was significantly more pronounced with SRSF1 overexpression. To dissect the role of the nucleocytoplasmic translocation of SRSF1 in this upregulation, we also used a C-terminal fusion of SRSF1 with the nuclear retention signal (NRS1) of SRSF2 [20–22]. The NRS1 sequence from SRSF2 (a serine- and arginine-rich protein within the same family as SRSF1 which does not translocate to the cytoplasm) blunts the nucleocytoplasmic translocation of SRSF1 [20]. Our results show that transfection of our reporter cells with the SRSF1-NRS1 mutant was associated with smaller increments in *MYC* IRES activity under both normal growth but especially stress conditions (Figure 5E). The reverse was also true, whereby siRNA-mediated downregulation of SRSF1 was associated with a decrease in the *MYC* IRES activity and blunted increments in the *MYC* IRES activity under conditions of stress (Figure 6A). These smaller increments under conditions of stress were also associated with an inability to maintain MYC protein expression levels under conditions of stress (Figure 6B, Supplemental Figure 9). Finally, our UV CLIP assay shows that SRSF1 does associate with the *MYC* 5′UTR, and this association is more pronounced under conditions of stress (Figure 6C).

Collectively, our results show that SRSF1 is uniformly critical for protein synthesis in SUM159PT cells and the mechanism by which it is involved differs depending on the growth conditions (Figure 6D). Under normal growth conditions, SRSF1 is predominantly nuclear and its effect on translation is indirect, presumably mediated by splicing in or out elements that promote or repress translation, respectively. Under conditions of stress, SRSF1 translocates to the cytoplasm where it directly prioritizes the translation of *MYC* mRNA and potentially other mRNAs by promoting their IRES-mediated translation, essentially operating as an "at-risk" ITAF.

The Role of SRSF1 on Translation in a Representative ER(+) Breast Cancer Cell Line

In T47D cells, SRSF1 is predominantly nuclear and, unlike the SUM159PT cells, growth with media with charcoal-stripped serum or under serum-deprived conditions, was not associated with nucleocytoplasmic translocation of SRSF1 (Supplemental Figure 10A and B). We did note however that in cells undergoing mitosis SRSF1 was diffusely localized in the cytoplasm (Figure 7A). We made the same observation in ZR75-1 cells as well (Supplemental Figure 10C). siRNA-mediated downregulation of SRSF1 was associated with inhibition of cell proliferation, and this inhibition was due to a pronounced S-phase arrest (Figure 7B and C). These observations suggested that SRSF1 plays an important role in cellular division and the cytoplasmic diffusion accompanying the nuclear envelope disassembly may be an integral component of its function during mitosis. Indeed, many mRNAs involved in cell division have been shown to be direct translational targets of SRSF1 [46]. Similar to the case of SUM159PT cells, the siRNA-mediated downregulation of SRSF1 is associated with pronounced global downregulation of protein synthesis and more so when the cells were synchronized to G2/M phase with nocodazole (Figure 7D).

Progression through the stages of cell cycle is regulated by specific cyclin-cyclin–dependent kinase (CDK) complexes [47]. Cyclin A–CDK1 and cyclin B–CDK1 complexes facilitate the onset and progression of mitosis, respectively [48]. CDKN1A/p21, on the other hand, prevents or limits the activity of cyclin-CDKs and induces growth arrest [49]. While CDK1 is expressed in all cell lines (more specifically overexpressed in T47D), CDKN1A/p21 was minimally expressed, or not expressed at all, in the breast cancer cell lines (except ZR75-1) in contradiction to HMECs and MCF10A cells (Supplemental Figure 11). The inability of T47D cells to progress to the G2/

downregulation in new protein synthesis especially when cells are synchronized to G2/M phase. (E) UV CLIP assay of the T47D *CDK1* IRES reporter cells. The assay identifies the endogenous (empty arrow) and the stably transfected (black arrow) *CDK1* 5′UTR in the SRSF1 immunoprecipitate. RNA extracted from the non-immunoprecipitated supernatant serves as control.

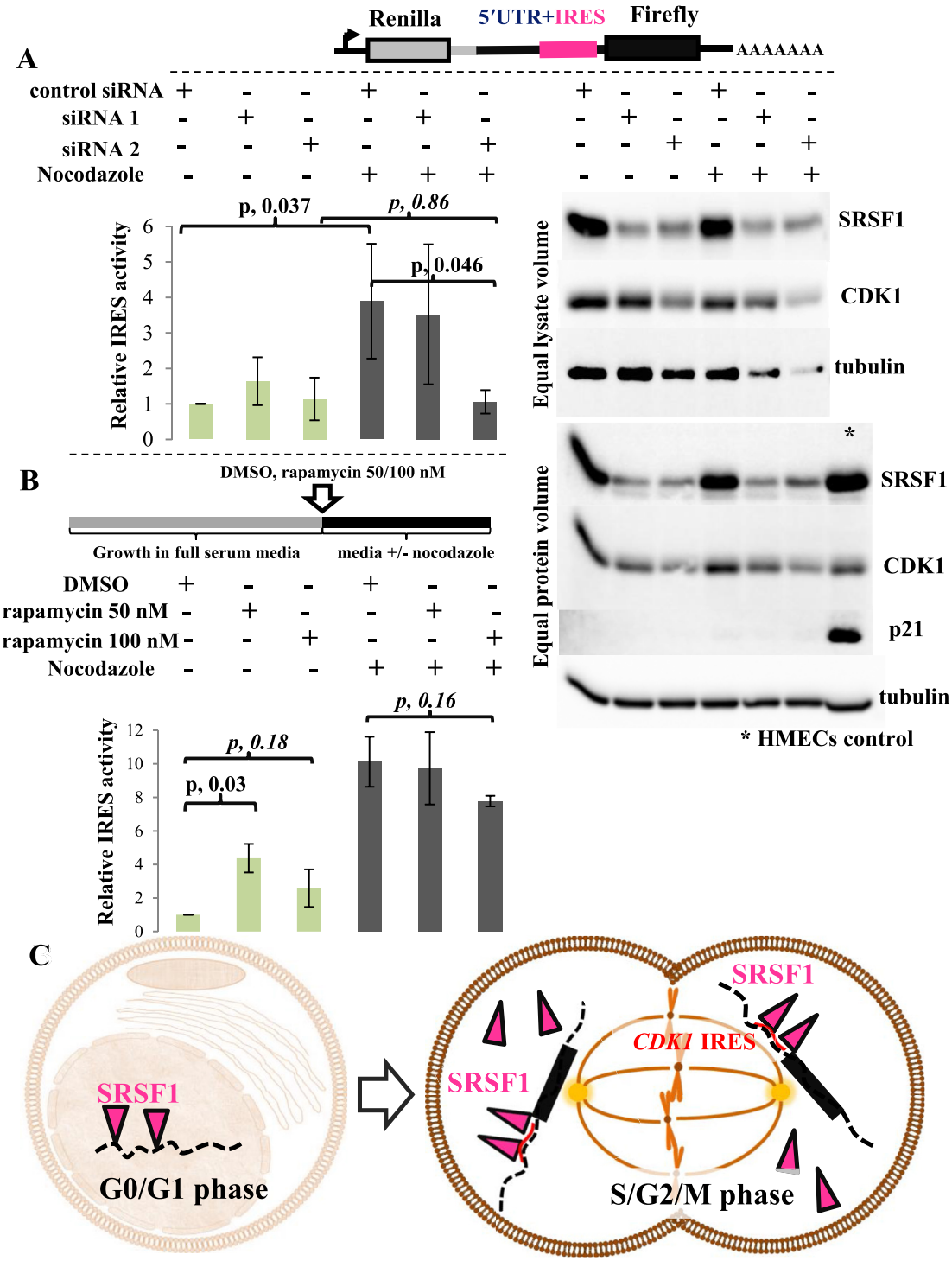

Figure 8. During cell division, cytoplasmic SRSF1 operates as an ITAF for the *CDK1* IRES. (A) Luciferase assay (left panel) and immunoblot for SRSF1, CDK1, and p21/CDKN1A (right panel) of T47D *CDK1* IRES reporter cells. T47D cells were stably transfected with the *CDK1* IRES bicistronic reporter vector (upper panel). Left panel. SRSF1 expression was downregulated with siRNAs and cells were asynchronous or synchronized with nocodozale (16 h) as indicated. siRNA mediated downregulation of SRSF1 was not associated with an apparent effect on the *CDK1* IRES activity in an asynchronous T47D cell population. G2/M synchronization with nozodazole captures the cell cycle dependent upregulation of *CDK1* IRES mediated translation. This upregulation is blocked by siRNA mediated downregulation of the SRSF1. IRES activity (ratio of Firefly:Renilla luciferase) is normalized to control siRNA-no nocodazole (asynchronous) condition. Data presented as mean ± standard deviation, $n = 4$; paired Student's t-test. Right panel. CDK1 protein levels change coordinately with SRSF1 especially when cells are synchronized to the G2/M phase. Two replicates of the same experiment are shown. Note the downregulation of tubulin with SRSF1 downregulation in the synchronized cells consistent with the fact that tubulin constitutes the main component of the mitotic spindle. The downregulation of SRSF1 does not impact p21 levels in T47D cells. (B) Luciferase assay of T47D *CDK1* IRES reporter cells, asynchronous or synchronized to the G2/M phase with nocodazole and concurrently treated with rapamycin (50 or 100 nM) or DMSO control (upper panel). Lower panel. G2/M

M phase upon SRSF1 downregulation led us to hypothesize that SRSF1 directs the translation of *CDK1* mRNA during mitosis by operating as an ITAF akin to the case of *MYC* mRNA in SUM159PT cells under conditions of stress. During mitosis, IRES-mediated translation predominates [45,50,51]; *CDK1* mRNA contains an IRES [52] and is a translational target of SRSF1 along with other mRNAs involved in cell division [46]. Moreover, our studies show that SRSF1 colocalizes with RACK1, a component of the IRES translation initiation complex (Figure 7A) [53].

To investigate this hypothesis, we used T47D cells stably transfected with a bicistronic reporter vector where the Renilla and firefly cistrons are separated by the *CDK1* 5′UTR. We confirmed by UV CLIP a direct association between SRSF1 and *CDK1* 5′UTR (Figure 7E). Our bioluminescence studies show that, indeed, during mitosis, the IRES-mediated translation of *CDK1* is upregulated and this upregulation is blunted when SRSF1 is downregulated by siRNA (Figure 8A). This inability to upregulate the IRES-mediated translation of *CDK1* mRNA is in turn associated with lower levels of CDK1, especially during cell division (Figure 8A). Cell cycle stalling upon SRSF1 downregulation cannot be attributed to CDKN1A/p21 upregulation (Figure 8A). Furthermore, rapamycin, an mTOR inhibitor and inhibitor of cap-dependent translation, does not blunt the increments of IRES activity during cell division. This result confirms that indeed, during cell division, the translation of the *CDK1* mRNA is IRES mediated (Figure 8B). Of note, consistent with the literature [54,55], mTOR inhibition may lead to a compensatory increase of IRES-mediated translation of *CDK1* mRNA.

Collectively, these results suggest that in T47D cells, SRSF1 has pervasive impact on translation and the mechanism by which SRSF1 modulates translation is cell cycle dependent. During G0 and G1 phase, SRSF1 may modulate translation indirectly through splicing. During mitosis though, the disassembly of the nuclear envelope allows SRSF1 to diffuse in the cytoplasm and orchestrate the mitotic translational landscape by operating as an ITAF (Figure 8C).

Discussion

The fundamental, yet elusive, biologic question that this work intends to address is how, from a pool of competing mRNAs, breast tumor cells prioritize the translation of certain transcripts over others. Gene expression can be significantly skewed at the level of translation accounting at least in part for the imperfect correlation between transcript and protein levels in a cell [56]. In cancer, the translational apparatus is diverted toward preferential translation of genes and transcripts that support the phenotypic hallmarks of malignancy [6,31]. The clinical-translational (bench to bedside) implications of this work is that it provides clues as to why gene expression is dysregulated in cancer, especially when there is a discrepancy between mRNA and protein levels in a tumor or when no "actionable" mutations are identified. Cancer cells may prioritize the translation of mRNAs encoding for components of a signaling pathway, e.g., PI3K-AKT pathway, without the tumor itself harboring an actionable mutation at the DNA level. In addition, distinctive changes in translation may define a premalignant state and precede the development of a frank malignant phenotype. This knowledge will help us understand the natural history and heterogeneity of premalignant lesions such as ductal carcinoma *in situ* and rationalize our therapeutic approaches.

Ribosome profiling and discriminative motif analysis have enabled us to reach a small list of variably common candidate RNA-binding proteins that could constitute key determinants of the translational output in breast cancer. A distinctive feature of our approach is that instead of focusing on a particular mechanism, we focus on a certain protein (SRSF1) that is commonly expressed in ER(+) and triple-negative breast tumor cells and has pleiotropic functions. Our results suggest that cancer cells rely exquisitely on SRSF1 to maintain protein synthesis and point to a vulnerability at the posttranscriptional level that can be therapeutically exploited. We uncover a novel mechanism by which SRSF1 can directly modulate the translational output, i.e., operating as a condition-specific ITAF, thereby expanding the repertoire of SRSF1's functions. SRSF1 undergoes complex posttranslational modifications, which in turn determine its subcellular localization and functions [44]. More specifically, the subcellular localization of SRSF1 seems to be determined by its phosphorylation status [44,57,58] and methylation status of 3 Arg residues [59]. Differential expression and activity of the enzymes that mediate these posttranslational modifications may underpin the differential cellular contexts in which SRSF1 operates as an ITAF. Our results lend further support to the therapeutic potential of strategies that interfere with SRSF1's RNA interaction-mediated functions, such as with decoy oligonucleotides [60]. In addition, targeting IRES-mediated translation, the translational program that SRSF1 directly and potentially indirectly modulates, is another promising approach [55,61].

SRSF1 was originally characterized as a splicing factor, and its aberrant splicing activity in cancer is a subject of ongoing investigations. In our studies, downregulation of SRSF1 under normal growth conditions (SUM159PT) or in an asynchronous cell population (T47D) when it is predominantly intranuclear brought new protein synthesis to a grinding halt. This result points to a biologically meaningful synergy between aberrant splicing and oncogenic translation. While alternative splicing results in protein isoforms with different, even opposing functions [20,62], the implications of aberrant splicing on translation have not been previously explored. The immediate question that arises pertains to the 5′UTR elements that are spliced in or out by SRSF1 and may operate differentially as translational enhancers or repressors, respectively, in malignant cells vs. nonmalignant counterparts. Posttranscriptional processes that are differentially operational between malignant and nonmalignant cells may recognize these 5′UTR elements and accordingly direct the translational apparatus. Indeed, hnRNPA1 and hnRNPA2, PABPN1, and FXR1 also

synchronization with nozodazole captures the cell cycle dependent upregulation of *CDK1* IRES mediated translation. Treatment with rapamycin (an mTOR inhibitor, i.e. inhibitor of cap-dependent translation), does not block this upregulation. Note also the upregulation of *CDK1* IRES activity induced by rapamycin in the asynchronous population (data presented as mean ± standard deviation, $n = 2$; paired Student's t-test). (C) Model for the translational activity of SRSF1 in T47D cells: during the G0/G1 phase, SRSF1 is predominantly intranuclear and operates as a splicing factor. During the S, G2, and M phase, the disassembly of the nuclear membrane releases SRSF1 to the cytoplasm where it operates as an IRES *trans*-acting factor and upregulates the IRES mediated translation of *CDK1* (and probably of other transcripts).

matched to the retrieved 5′UTR SRSF1 matching motifs, VAGGARR and VAGGRAR, and are involved in IRES-mediated translation [63]; RNA processing, alternative cleavage and polyadenylation [64]; and DAP5-mediated translation initiation [65], respectively. By splicing in the recognition motifs for these proteins, SRSF1-mediated splicing may set the stage for these aberrant processes and translational programs to operate. In addition, the SRSF1 matching motifs were found in GG-rich sequence contexts which may potentially represent G-quadruplex structures. In agreement with our findings, an alternatively spliced-in IRES element containing a G-quadruplex as a central motif has been shown at a mechanistic level to support IRES-mediated translation under conditions of stress [66]. Further studies are needed to establish the co-operativity between alternative splicing and aberrant translation.

Conclusions

Our studies provide a genome-wide view of the dysregulation in gene expression that occurs at the translational level in breast cancer, reveal parsimonious themes of oncogenic translation, identify RNA-binding proteins that cast a large shadow on the translational landscape, point to an association between aberrant splicing and oncogenic translation, and uncover a novel direct mechanism by which the splicing factor SRSF1 can modulate translation.

Data statement

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE126736 (BioProject: PRJNA523167, SRA: SRP186252).

Analyzed data and the R code used to generate the main and supplemental figures have been deposited at the Mendeley Data Repository: https://data.mendeley.com/datasets/zcsn75knky/draft?a=7308d74e-0d59-47d5-a797-1a8aceabd9f2.

DOI (main figures) https://doi.org/10.17632/zcsn75knky.1#folder-353df18e-7b6e-4175-8463-1ead47a4581f.

DOI (supplemental figures) https://doi.org/10.17632/zcsn75knky.1#folder-40ce3ba2-5708-4d32-bc5d-439104e16ac8.

Link to a UCSC genome browser session with uploaded sequence tracks: https://genome.ucsc.edu/s/cvaklavas/hg19_Vaklavas.

Cell cycle flow cytometry data have been deposited at FlowRepository (Repository ID, FR-FCM-Z2Z6) and can be accessed at http://flowrepository.org/id/RvFrdZn2XvxxW0b0pEMDX5DBQDrW5e4L9bNQZwbdMMLaMcnlQ2I4maIhPfLm1i9Y.

Acknowledgments

The authors thank the High Resolution Imaging Facility, the UAB Comprehensive Flow Cytometry Core, and the Heflin Center for Genomic Sciences Core Facilities at UAB for the studies involving confocal microscopy, flow cytometry, and sequencing, respectively.

The authors are deeply thankful to Dr. Anne Willis, Dr. Sunnie Thompson, and Dr. Adrian Krainer for sharing critical reagents (plasmids) for this study as well as technical advice. The authors thank the Breast Cancer Research Foundation of Alabama for covering the publication costs.

Competing Interests

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tranon.2019.12.002.

References

[1] Liu Y, Beyer A and Aebersold R (2016). On the dependency of cellular protein levels on mRNA abundance. *Cell* **165**(3), 535–550.
[2] Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J and Wolf J, et al (2011). Global quantification of mammalian gene expression control. *Nature* **473**(7347), 337–342.
[3] Sonenberg N and Hinnebusch AG (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136**(4), 731–745.
[4] Spilka R, Ernst C, Mehta AK and Haybaeck J (2013). Eukaryotic translation initiation factors in cancer development and progression. *Cancer Lett* **340**(1), 9–21.
[5] Vaklavas C, Blume SW and Grizzle WE (2017). Translational dysregulation in cancer: molecular insights and potential clinical Applications in biomarker development. *Front Oncol* **7**, 158.
[6] Truitt ML and Ruggero D (2016). New frontiers in translational control of the cancer genome. *Nat Rev Cancer* **16**(5), 288–304.
[7] Barna M, Pusic A, Zollo O, Costa M, Kondrashov N and Rego E, et al (2008). Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. *Nature* **456**(7224), 971–975.
[8] Ingolia NT, Brar GA, Rouskin S, McGeachy AM and Weissman JS (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* **7**(8), 1534–1550.
[9] Ingolia NT, Ghaemmaghami S, Newman JR and Weissman JS (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**(5924), 218–223.
[10] Ingolia NT, Lareau LF and Weissman JS (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**(4), 789–802.
[11] Nedvetsky PI, Kwon SH, Debnath J and Mostov KE (2012). Cyclic AMP regulates formation of mammary epithelial acini in vitro. *Mol Biol Cell* **23**(15), 2973–2981.
[12] Fields AP, Rodriguez EH, Jovanovic M, Stern-Ginossar N, Haas BJ and Mertins P, et al (2015). A Regression-Based analysis of ribosome-profiling data reveals a Conserved complexity to mammalian translation. *Molecular cell* **60**(5), 816–827.
[13] Dunn JG and Weissman JS (2016). Plastid: nucleotide-resolution analysis of next-generation sequencing and genomics data. *BMC Genomics* **17**(1), 958.
[14] Huang da W, Sherman BT and Lempicki RA (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**(1), 44–57.
[15] Yao Z, Macquarrie KL, Fong AP, Tapscott SJ, Ruzzo WL and Gentleman RC (2014). Discriminative motif analysis of high-throughput dataset. *Bioinformatics* **30**(6), 775–783.
[16] Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS and Li X, et al (2013). A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**(7457), 172–177.
[17] Grant CE, Bailey TL and Noble WS (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**(7), 1017–1018.
[18] Hertz MI, Landry DM, Willis AE, Luo G and Thompson SR (2013). Ribosomal protein S25 dependency reveals a common mechanism for diverse internal ribosome entry sites and ribosome shunting. *Mol Cell Biol* **33**(5), 1016–1026.
[19] Stoneley M, Paulin FE, Le Quesne JP, Chappell SA and Willis AE (1998). C-Myc 5' untranslated region contains an internal ribosome entry segment. *Oncogene* **16**(3), 423–428.
[20] Anczukow O, Rosenberg AZ, Akerman M, Das S, Zhan L and Karni R, et al (2012). The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. *Nat Struct Mol Biol* **19**(2), 220–228.
[21] Caceres JF, Misteli T, Screaton GR, Spector DL and Krainer AR (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J Cell Biol* **138**(2), 225–238.

[22] Cazalla D, Zhu J, Manche L, Huber E, Krainer AR and Caceres JF (2002). Nuclear export and retention signals in the RS domain of SR proteins. *Mol Cell Biol* **22**(19), 6871–6882.
[23] Love MI, Huber W and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**(12), 550.
[24] Ingolia NT, Brar GA, Stern-Ginossar N, Harris MS, Talhouarne GJ and Jackson SE, et al (2014). Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep* **8**(5), 1365–1379.
[25] Johnstone TG, Bazzini AA and Giraldez AJ (2016). Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J* **35**(7), 706–723.
[26] Han Y, Gao X, Liu B, Wan J, Zhang X and Qian SB (2014). Ribosome profiling reveals sequence-independent post-initiation pausing as a signature of translation. *Cell Res* **24**(7), 842–851.
[27] Ingolia NT (2016). Ribosome footprint profiling of translation throughout the genome. *Cell* **165**(1), 22–33.
[28] Belin S, Beghin A, Solano-Gonzalez E, Bezin L, Brunet-Manquat S and Textoris J, et al (2009). Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. *PLoS One* **4**(9):e7147.
[29] Slobodin B, Han R, Calderone V, Vrielink J, Loayza-Puch F and Elkon R, et al (2017). Transcription impacts the efficiency of mRNA translation via Co-transcriptional N6-adenosine methylation. *Cell* **169**(2), 326–337 e12.
[30] Rajasekhar VK, Viale A, Socci ND, Wiedmann M, Hu X and Holland EC (2003). Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes. *Mol Cell* **12**(4), 889–901.
[31] Robichaud N, Sonenberg N, Ruggero D and Schneider RJ (2019). Translational control in cancer. *Cold Spring Harb Perspect Biol* **11**(7).
[32] Debnath J, Mills KR, Collins NL, Reginato MJ, Muthuswamy SK and Brugge JS (2002). The role of apoptosis in creating and maintaining luminal space within normal and oncogene-expressing mammary acini. *Cell* **111**(1), 29–40.
[33] Debnath J and Brugge JS (2005). Modelling glandular epithelial cancers in three-dimensional cultures. *Nat Rev Cancer* **5**(9), 675–688.
[34] Qu Y, Han B, Yu Y, Yao W, Bose S and Karlan BY, et al (2015). Evaluation of MCF10A as a reliable model for normal human mammary epithelial cells. *PLoS One* **10**(7):e0131285.
[35] Soule HD, Maloney TM, Wolman SR, Peterson Jr WD, Brenz R and McGrath CM, et al (1990). Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Res* **50**(18), 6075–6086.
[36] Dimri GP, Lee X, Basile G, Acosta M, Scott G and Roskelley C, et al (1995). A biomarker that identifies senescent human cells in culture and in aging skin in vivo. *Proc Natl Acad Sci U S A* **92**(20), 9363–9367.
[37] Prat A, Karginova O, Parker JS, Fan C, He X and Bixby L, et al (2013). Characterization of cell lines derived from breast cancers and normal mammary tissues for the study of the intrinsic molecular subtypes. *Breast Cancer Res Treat* **142**(2), 237–255.
[38] Sendoel A, Dunn JG, Rodriguez EH, Naik S, Gomez NC and Hurwitz B, et al (2017). Translation from unconventional 5' start sites drives tumour initiation. *Nature* **541**(7638), 494–499.
[39] Shahbazian D, Parsyan A, Petroulakis E, Topisirovic I, Martineau Y and Gibbs BF, et al (2010). Control of cell survival and proliferation by mammalian eukaryotic initiation factor 4B. *Mol Cell Biol* **30**(6), 1478–1485.
[40] Shahbazian D, Parsyan A, Petroulakis E, Hershey J and Sonenberg N (2010). eIF4B controls survival and proliferation and is regulated by proto-oncogenic signaling pathways. *Cell Cycle* **9**(20), 4106–4109.
[41] Anczukow O, Akerman M, Clery A, Wu J, Shen C and Shirole NH, et al (2015). SRSF1-Regulated alternative splicing in breast cancer. *Mol Cell* **60**(1), 105–117.
[42] Michlewski G, Sanford JR and Caceres JF (2008). The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1. *Mol Cell* **30**(2), 179–189.
[43] Feng H, Bao S, Rahman MA, Weyn-Vanhentenryck SM, Khan A and Wong J, et al (2019). Modeling RNA-binding protein specificity in vivo by Precisely Registering protein-RNA Crosslink sites. *Mol Cell* **74**(6), 1189–11204 e6.
[44] Das S and Krainer AR (2014). Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Mol Cancer Res* **12**(9), 1195–1204.
[45] Spriggs KA, Stoneley M, Bushell M and Willis AE (2008). Re-programming of translation following cell stress allows IRES-mediated translation to predominate. *Biol Cell* **100**(1), 27–38.
[46] Maslon MM, Heras SR, Bellora N, Eyras E and Caceres JF (2014). The translational landscape of the splicing factor SRSF1 and its role in mitosis. *Elife 2014:e02028.*
[47] Malumbres M and Barbacid M (2009). Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* **9**(3), 153–166.
[48] Tripathy D, Bardia A and Sellers WR (2017). Ribociclib (LEE011): mechanism of action and clinical impact of this selective cyclin-dependent kinase 4/6 inhibitor in various Solid tumors. *Clin Cancer Res* **23**(13), 3251–3262.
[49] Xiong Y, Hannon GJ, Zhang H, Casso D, Kobayashi R and Beach D (1993). p21 is a universal inhibitor of cyclin kinases. *Nature* **366**(6456), 701–704.
[50] Pyronnet S, Pradayrol L and Sonenberg N (2000). A cell cycle-dependent internal ribosome entry site. *Mol Cell* **5**(4), 607–616.
[51] Qin X and Sarnow P (2004). Preferential translation of internal ribosome entry site-containing mRNAs during the mitotic cycle in mammalian cells. *J Biol Chem* **279**(14), 13721–13728.
[52] Marash L, Liberman N, Henis-Korenblit S, Sivan G, Reem E and Elroy-Stein O, et al (2008). DAP5 promotes cap-independent translation of Bcl-2 and CDK1 to facilitate cell survival during mitosis. *Mol Cell* **30**(4), 447–459.
[53] Yu Y, Ji H, Doudna JA and Leary JA (2005). Mass spectrometric analysis of the human 40S ribosomal subunit: native and HCV IRES-bound complexes. *Protein Sci* **14**(6), 1438–1446.
[54] Shi Y, Sharma A, Wu H, Lichtenstein A and Gera J (2005). Cyclin D1 and c-myc internal ribosome entry site (IRES)-dependent translation is regulated by AKT activity and enhanced by rapamycin through a p38 MAPK- and ERK-dependent pathway. *J Biol Chem* **280**(12), 10964–10973.
[55] Vaklavas C, Meng Z, Choi H, Grizzle WE, Zinn KR and Blume SW (2015). Small molecule inhibitors of IRES-mediated translation. *Cancer Biol Ther* **16**(10), 1471–1485.
[56] Vogel C and Marcotte EM (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* **13**(4), 227–232.
[57] Colwill K, Pawson T, Andrews B, Prasad J, Manley JL and Bell JC, et al (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J* **15**(2), 265–275.
[58] Gui JF, Lane WS and Fu XD (1994). A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature* **369**(6482), 678–682.
[59] Sinha R, Allemand E, Zhang Z, Karni R, Myers MP and Krainer AR (2010). Arginine methylation controls the subcellular localization and functions of the oncoprotein splicing factor SF2/ASF. *Mol Cell Biol* **30**(11), 2762–2774.
[60] Denichenko P, Mogilevsky M, Clery A, Welte T, Biran J and Shimshon O, et al (2019). Specific inhibition of splicing factor activity by decoy RNA oligonucleotides. *Nat Commun* **10**(1), 1590.
[61] Vaklavas C, Grizzle WE, Choi H, Meng Z, Zinn KR and Shrestha K, et al (2016). IRES inhibition induces terminal differentiation and synchronized death in triple-negative breast cancer and glioblastoma cells. *Tumour Biol* **37**(10), 13247–13264.
[62] Zhang J and Manley JL (2013). Misregulation of pre-mRNA alternative splicing in cancer. *Cancer Discov* **3**(11), 1228–1237.
[63] Jo OD, Martin J, Bernath A, Masri J, Lichtenstein A and Gera J (2008). Heterogeneous nuclear ribonucleoprotein A1 regulates cyclin D1 and c-myc internal ribosome entry site function through Akt signaling. *J Biol Chem* **283**(34), 23274–23287.
[64] Jenal M, Elkon R, Loayza-Puch F, van Haaften G, Kuhn U and Menzies FM, et al (2012). The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **149**(3), 538–553.
[65] de la Parra C, Ernlund A, Alard A, Ruggles K, Ueberheide B and Schneider RJ (2018). A widespread alternate form of cap-dependent mRNA translation initiation. *Nat Commun* **9**(1), 3068.
[66] Al-Zeer MA, Dutkiewicz M, von Hacht A, Kreuzmann D, Rohrs V and Kurreck J (2019). Alternatively spliced variants of the 5'-UTR of the ARPC2 mRNA regulate translation by an internal ribosome entry site (IRES) harboring a guanine-quadruplex motif. *RNA Biol* **16**(11), 1622–1632.