Original Article Targeting the oncogenic TBX3:nucleolin complex to treat multiple sarcoma subtypes

Tarryn Willmer^{1,2,3*}, Victoria Damerell^{1*}, Shannon Smyly¹, Danica Sims¹, Michelle Du Toit¹, Stephanie Ncube¹, Musalula Sinkala⁴, Dhirendra Govender^{5,6}, Edward Sturrock⁷, Jonathan M Blackburn⁷, Sharon Prince¹

¹Division of Cell Biology, Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Cape Town 7925, South Africa; ²Biomedical Research and Innovation Platform, South African Medical Research Council, Tygerberg 7505, South Africa; ³Division of Medical Physiology, Faculty of Health Sciences, Stellenbosch University, Tygerberg 7505, South Africa; ⁴Division of Computational Biology, Department of Integrative Biomedical Sciences, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town 7925, South Africa; ⁵Anatomical Pathology, Pathcare, Cape Town 7925, South Africa; ⁶Division of Anatomical Pathology, Faculty of Health Sciences, University of Cape Town, NHLS-Groote Schuur Hospital, Cape Town 7925, South Africa; ⁷Division of Chemical and Systems Biology, Department of Integrative Biomedical Sciences, Faculty of Health Sciences, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town 7925, South Africa; ⁸Division of Chemical and Systems Biology, Department of Integrative Biomedical Sciences, Faculty of Health Sciences, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town 7925, South Africa. *Equal contributors.

Received July 9, 2021; Accepted October 8, 2021; Epub November 15, 2021; Published November 30, 2021

Abstract: Sarcomas are diverse cancers of mesenchymal origin, with compromised clinical management caused by insufficient diagnostic biomarkers and limited treatment options. The transcription factor TBX3 is upregulated in a diverse range of sarcoma subtypes, where it plays a direct oncogenic role, and it may thus represent a novel therapeutic target. To identify versatile ways to target TBX3, we performed affinity purification coupled by mass spectrometry to identify putative TBX3 protein cofactors that regulate its oncogenic activity in sarcomas. Here we identify and validate the multifunctional phosphoprotein nucleolin as a TBX3 cofactor. We show that nucleolin is co-expressed with TBX3 in several sarcoma subtypes and their expression levels positively correlate in sarcoma patients which are associated with poor prognosis. Furthermore, we demonstrate that nucleolin and TBX3 interact in chondrosarcoma, liposarcoma and rhabdomyosarcoma cells where they act together to enhance proliferation and migration and regulate a common set of tumor suppressor genes. Importantly, the nucleolin targeting aptamer, AS1411, exhibits selective anti-cancer activity in these cells and mislocalizes TBX3 and nucleolin to the cytoplasm which correlates with the re-expression of the TBX3/nucleolin target tumor suppressors *CDKN1A* (p21^{CIP1}) and *CDKN2A* (p14^{ARF}). Our findings provide the first evidence that TBX3 requires nucleolin to promote features of sarcomagenesis and that disruption of the oncogenic TBX3-nucleolin interaction by AS1411 may be a novel approach for treating sarcomas.

Keywords: T-box transcription factor 3 (TBX3), nucleolin, oncogene, sarcomas, nucleolin aptamer, AS1411

Introduction

Sarcomas are mesenchymal malignancies which account for approximately 1% of adult and 20% of pediatric solid tumors [1, 2]. A major challenge to the clinical management of sarcomas is the existence of more than 100 different histologic subtypes which differ in molecular characteristics, pathology, clinical features, and responses to treatment. Furthermore, sarcomas are highly aggressive and insensitive to current treatments which include surgery in combination with chemotherapy and/or radiation [3, 4]. Importantly, sarcomas metastasize in one third of all patients, and in these instances, surgery is no longer effective and the response to chemotherapeutics is limited [5]. Patients with metastatic sarcoma therefore have a poor prognosis and an overall five-year survival rate of less than 10%. Targeted therapies have gained much traction to address the burden of sarcoma but because of a limited understanding of the molecular basis underpinning sarcomagenesis there have not been any major clinical breakthroughs in targeted sarcoma treatments [4].

The T-box transcription factor TBX3 contributes critically to embryogenesis by regulating programs of gene expression. Its deregulated expression is associated with human disorders including ulnar mammary syndrome, obesity and rheumatoid arthritis, as well as the development of numerous sarcomas and carcinomas [6, 7]. Indeed, TBX3 is overexpressed in fibrosarcoma, rhabdomyosarcoma, liposarcoma, chondrosarcoma and cancers of the breast, pancreas, liver, lung, stomach, head and neck, ovaries, bladder and skin [6]. When overexpressed TBX3 can bypass senescence, promote proliferation and immortalization, and prevent apoptosis by interfering with multiple points of the p19^{ARF}/p14^{ARF}-Mdm2-p53-p21 pathway. TBX3 exerts these roles by repressing the p21 and p19^{ARF}/p14^{ARF} promoters either directly [8], or by recruiting histone deacetylases [9]. TBX3 also mediates tumor formation, invasion, and migration in a range of carcinoma and sarcoma subtypes [10-16]. While TBX3 has been identified as a novel therapeutic target, direct targeting of transcription factors is notoriously challenging. A more viable approach to inhibiting TBX3 oncogenic activity may involve targeting its interacting partners.

Nucleolin is a multifunctional phosphoprotein which is predominantly located in the nucleolus but is also present in other regions of the nucleus and in the cytoplasm and, in rapidly dividing cells, it is also found in the cell membrane [17, 18]. It is best known for its roles in ribosome biogenesis including ribosomal RNA transcription, organization of nucleolar chromatin, pre-ribosome packaging and ribosome assembly. In addition, nucleolin has been implicated in other processes such as DNA metabolism, replication and repair, the maintenance of telomeres as well as epigenetic and post-transcriptional processes including chromatin remodeling, histone chaperoning and mRNA stability [17]. Nucleolin is aberrantly expressed at the cell surface of a range of cancer cell types where it complexes with oncoproteins to contribute to several aspects of the cancer process such as cell survival, proliferation and invasion [19-23]. Its expression on the surface of cancer cells makes nucleolin a promising anti-cancer target because it can serve as a receptor which allows for tumor-specific uptake of targeted chemotherapeutic drugs without affecting normal cells. There are many

similarities between the oncogenic functions and regulation of nucleolin and TBX3. For example, nucleolin blocks apoptosis and prevents cell cycle arrest by inhibiting p53 and p21 expression and co-operates with oncogenic Ras to transform rat fibroblasts [21, 24-27]. Like TBX3, nucleolin levels also peak in the S-phase, and it is transcriptionally activated by c-Myc and post-translationally stabilized by cyclin A/CDK2a [28-30]. Importantly, the 26-base guanine-rich DNA oligonucleotide aptamer to nucleolin, AS1411, has selective anti-proliferative activity against a wide range of cancer cells [31]. Furthermore, AS1411 has been evaluated in phase II clinical trials for acute myeloid leukemia, with positive preliminary results (clinicaltrials.gov identifiers, NCT01034410 and NCT00512083). There is, however, a paucity of information regarding the molecular mechanisms by which AS1411 exhibits anti-cancer activity.

In this study, we used affinity purification mass spectrometry to identify TBX3 interacting proteins, and we validated and characterized the interaction between TBX3 and nucleolin. We measured the expression of nucleolin and TBX3 in a range of sarcoma subtypes and identified increased proliferation and migration as a co-operative function of TBX3 and nucleolin in multiple sarcoma subtypes. Furthermore, we show that the nucleolin targeting aptamer AS1411 inhibits the pro-proliferative and pro-migratory effects of TBX3 and nucleolin in chondrosarcoma, rhabdomyosarcoma and liposarcoma cell lines. Taken together, our findings support the exploration of therapeutic strategies targeting TBX3 via its interacting partners and highlight AS1411 as a promising candidate.

Materials and methods

Plasmids and DNA constructs

The following constructs were gifts: human pCMV-TBX3-HA (Dr Christine Campbell, Cleveland Clinic Foundation, USA), FLAG-tagged pCMV-Tbx3 (Prof Colin Goding, Ludwig Institute, University of Oxford, UK) and pEGFPnucleolin vector (Prof Michael Kastan, Addgene plasmid # 28176; http://n2t.net/addgene:28176; RRID:Addgene_28176) [22]. The human TBX3 N-terminal construct (N-term) lacks the activation and dominant repression domains and the human TBX3 DNA-binding domain mutant (DBM) construct contains an arginine to glycine substitution at position 133 (R133G) which disrupts the DNA-binding domain. Both N-term and DBM constructs were generated in the laboratory as described in [32].

Cell culture and transfections

FGO and DMB human skin fibroblasts (provided by Associate Professor Denver Hendricks, University of Cape Town, South Africa) [33] and HT1080 human fibrosarcoma (ATCC®CCL-120), SW1353 human chondrosarcoma (ATCC®HTB-94), SW872 human liposarcoma (ATCC®HTB-92), SW982 human synovial sarcoma (AT-CC®HTB-93) and RD human embryonal rhabdomyosarcoma (ATCC®CCL-136) cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma Aldrich, USA). All culture media were supplemented with 10% heat-inactivated foetal bovine serum, 100 U/ml penicillin and 100 µg/ml streptomycin. Cells were maintained at 37°C in an atmosphere of 5% CO and 65% humidity. Cells (60-70% confluent) were transfected using a 3:1 transfection reagent (FuGENE HD, Roche, Germany or X-tremeGENE HP, Roche, Germany): DNA ratio with according to the manufacturer's instructions. For transient nucleolin overexpression, cells were transfected with 250 ng of the pEGFP-nucleolin or pEGFP-Empty vector. Transient silencing of TBX3 or nucleolin was achieved by siTBX3 (SI00083503, Qiagen, USA) or siNuc (Dharmacon, Lafayette, USA and SI02654925, Qiagen, USA) respectively, and a nonspecific control siRNA (siCtrl, 1027310; Qiagen, USA) was included in these experiments, using HiPerFect (Qiagen, USA) according to manufacturer's instructions.

Generation of stable cell lines

Stable TBX3 knockdown (shTBX3) SW1353 cells were generated by transfection with a pSuper.neo/GFP expression vector encoding a sequence targeting TBX3 or a nonspecific control sequence [34]. Stable transfectants were selected using 800 µg/ml G418 (Promega, USA). Cell lines stably overexpressing TBX3 were generated by transfecting SW1353, SW-872 and RD cells with a FLAG-tagged pCMV-Tbx3 or FLAG-tagged pCMV-empty vector using FuGENE HD (Roche, Germany) according to the manufacturer's instructions. G418-selection of stable transfectants was performed using 800 μ g/ml G418 antibiotic (Promega, USA). Effective knockdown and overexpression of TBX3 was assessed by western blotting.

Treatment with the oligonucleotide AS1411 aptamer

The oligodeoxynucleotides AS1411, 5'-d(GGTG GTGGTGGTGGTGGTGGTGGTGGTGGTGG)-3', and an inactive control oligonucleotide, 5'-d(CCTCCTCCTCC TTCTCCTCCTCC)-3' were purchased in the desalted form from Integrated DNA Technologies. A stock solution (1,000 μ mol/L) was prepared in water and added directly to the cell culture medium to obtain the desired final concentration. Unless otherwise stated, cells were treated with a final oligonucleotide concentration of 10 μ mol/L for 48 h, followed by preparation of cell extracts.

Immunofluorescence

Immunofluorescence was performed as described previously [35]. Briefly, cells were incubated with rabbit polyclonal anti-TBX3 (1:25 dilution; 42-4800; Zymed, USA) or mouse polyclonal anti-nucleolin (1:50 dilution, sc-8031, Santa Cruz Biotechnology, USA) antibodies at 4°C overnight. Subsequently, cells were incubated with either fluor-conjugated secondary Cy3 donkey anti-rabbit IgG (Jackson Immuno-Research Laboratories, Inc., USA) or Alexa 488 goat anti-mouse secondary (1:1000; Molecular probes, USA) antibodies for 2 hours at room temperature. To visualise the nuclei, cells were incubated with 1 µg/ml DAPI (Sigma Aldrich, USA) or 1 mg/ml Hoechst 33258 (Thermofisher Scientific, USA). Fluorescent cells were viewed using Axiovert confocal microscope (Zeiss, USA).

Western blot analysis

Protein extracts were prepared as previously described [36]. The primary antibodies used were rabbit polyclonal anti-TBX3 (1:250; 42-4800, Zymed, USA), mouse monoclonal anti-Flag M2 (1:1000, F1804, Sigma Aldrich, USA), rabbit polyclonal anti-p38 (1:5000, Sigma Aldrich, USA), anti-HA (1:1000, Sigma Aldrich, USA), mouse monoclonal anti-nucleolin (1:1000, sc-8031, Santa Cruz Biotechnology, USA), mouse monoclonal anti-p84 (1:1000, Genetex, USA) and mouse mono-

clonal anti-GAPDH (1:1000, Genetex, USA). Secondary antibodies were used at 1:5000 dilution and included horseradish peroxidase (HRP)-conjugated goat anti-rabbit (Biorad, USA), goat anti-mouse (Biorad, USA) or donkey anti-goat (Santa Cruz Biotechnology, USA).

Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was extracted using the RNeasy Plus Mini kit (Qiagen, USA). The RNA quality and concentration were determined using a NanoDrop® ND-1000 spectrophotometer (Agilent Technologies, Boeblingen, Germany). RNA was reverse transcribed using the InProm-II™ reverse transcription system (Promega A3800, USA), following the manufacturer's instructions. The reactions were carried out using 2 µl of cDNA, 2× SYBR green master mix (Applied Biosystems, Carlsbad, CA, USA) and primers to amplify the CDKN2A (p14ARF) (Forward-5'-ACC-AGCCGCTTCCTAGAAGAC-3', Reverse-5'-CACG-GGTCGGGTGAGAGT-3'), CDKN1A (p21^{CIP1}) (Forward-5'-GGGGCGGTTGTATATCAGGG-3'. Reverse-5'-TCTCACCTCCTCTGAGTGCC-3') and GUSB (QT00046046; Qiagen, USA) genes. gRT-PCR was conducted using the StepOnePlus[™] Real-Time PCR system (Applied Biosystems). PCR cycle parameters were denaturation for 15 min at 95°C and combined annealing and extension for 35 cycles at 60°C for 1 min. Samples were prepared and guantified in triplicate and non-template controls were included in each run to detect nonspecific amplification or the presence of contamination. The results were analysed using the 2-AACt method and relative CDKN1A and CDKN2A expression levels were normalised to mRNA levels of glucuronidasebeta (GUSB).

Proliferation assay

Short-term growth was measured as described previously [36]. Briefly, cells were seeded in triplicate in 12-well or 24-well plates at densities of 2×10^4 cells/mL, 1×10^4 cells/mL or 2×10^5 cells/mL for SW1353, RD and SW872 cells, respectively. Cells were collected by trypsinization at 48-72-hour intervals and counted on a haemocytometer as described previously [36]. As an alternative assay for proliferation, cell viability was quantified using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) Cell Proliferation Kit (Roche, Germany) according to the manufacturer's instructions. Briefly, cells were seeded in quadruplicate in a 96-well plate (1000 cells/well) and cell viability determined over 96 h. Absorbance (595 nm) was determined and the absorbance of the medium only control was subtracted from the samples. At least three independent experiments were performed in quadruplicate, from which the half maximal inhibitory concentration (IC_{50}) was calculated using GraphPad Prism version 6.0 (GraphPad Software, USA).

In vitro cell migration assays

Cell migration was assessed using a 2-dimensional in vitro scratch motility assay, which has been described in detail previously [34]. Images of the wound were taken over time using a phase contrast microscope and migration areas were calculated using ImageJ software (National Institutes of Health, Bethesda, MD).

Subcellular fractionation

Nuclear and cytoplasmic extracts were prepared from sub-confluent SW1353 chondrosarcoma cells in 10 cm cell culture dishes. Briefly, cells were washed twice in cold 1× PBS, harvested in cold 1× PBS by scraping and pelleted at 5000 g for 3 min at 4°C. The cell pellet was resuspended in 5 volumes of solution 1 (20 mM PIPES pH 6.8, 1 mM EGTA pH 6.8, 1 mM MgCl_a) plus protease inhibitors and the cells allowed to swell on ice for 5 min, then 0.5% Triton X-100 was added to lyse the cell membranes and left to incubate for further 5 min. The cells were centrifuged at 900 g and the cytosolic fraction (supernatant) was transferred to fresh tubes and stored at -80°C. The pellet was washed in 5 volumes of solution 1 with protease inhibitors and centrifuged at 900 g. The supernatant was discarded and the cell nuclei (pellet) lysed in 5 volumes of solution 2 (100 mM KCl, 300 mM sucrose, 10 mM PIPES pH 6.8, 1 mM EGTA pH 6.8, 3 mM MgCl₂) with protease inhibitors and RNASE-free DNASE (100 µg/ml) at 30°C for 45 mins followed by centrifugation at 1500 g for 5 min at 4°C. The nuclear extract (supernatant) was transferred to fresh tubes and stored at -80°C. Protein concentration was assayed by BCA assay (Pierce, USA) with bovine serum albumin as the standard.

Immunoprecipitation assay

FLAG-Tbx3 or FLAG-Empty cells were plated in 15 cm tissue culture dishes at 90% confluence. Cells were lysed in 150 mM lysis buffer [50 mM Tris HCI (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1% TRITON X-100, 10 mM NaF, 0.2 mM PMSF, 0.7 µg/ml pepstatin A, 2 µg/ml aprotinin, 2 mM NaF, 0.01 mM sodium orthovanadate, 2 complete proteinase inhibitor tablets (Roche, Germany)] and cell extracts were centrifuged at 12000 rpm for 20 min at 4°C to remove cellular debris, The supernatant was incubated with anti-Flag antibody conjugated to agarose beads (anti-Flag M2 affinity gel, A2220, Sigma Aldrich, USA) with constant rotation overnight at 4°C. Immune complexes were resuspended in 20 µl 2× SDS loading buffer and boiled for 10 min at 100°C. Protein samples were either analysed by SDS-PAGE and western blotting using antibodies to TBX3 (1:500; 42-4800, Zymed, USA), nucleolin (1:1000, sc-8031, Santa Cruz Biotechnology, USA), or subjected to filter aided sample preparation for mass spectrometry analysis.

Sample preparation for mass spectrometry analysis

All buffer exchanges were conducted by centrifugation at 14000 g for 15 min at RT. Approximately 200 µg of affinity purified protein was denatured by boiling for 10 min at 95°C in 2× SDS loading buffer without bromophenol blue. The BCA assay (Pierce, Thermo Scientific, USA) was used to estimate protein extract concentrations. To reduce cysteine bonds protein extracts were treated with 100 mM DTT for 1 h at RT. Reduced proteomic material was loaded into an Ultracel centrifugal unit with a molecular weight cut off of 30000 Da (Amicon Ultra, Merck). Buffer-exchange was performed three times with 200 µL of UT buffer (8 M urea, 0.1 M Tris, pH 8.5). Alkylation of reduced cysteine residues was performed by adding 200 µL of UT buffer containing 0.05 M iodacetamide (Sigma Aldrich, USA) onto the filter and incubating it in the dark for 20 min. Two buffer exchanges using 200 µL of UT were performed to remove the alkylating agent, followed by three buffer exchanges with 100 µL of ABC buffer (50 mM ammonium bicarbonate buffer, pH 8). Proteins were proteolyzed by the addition of 40 µL of ABC buffer containing sequencing-grade modified trypsin (Promega, USA) at a ratio of 1:50 to the amount of protein retained on the filter. Proteolysis was carried out at 37°C for 18 h in a wet chamber. Peptides were eluted from the filters using 150 µL of ABC buffer. The resulting peptides were desalted using a homemade stage tip containing Empore Octadecyl C18 solid-phase extraction disk (Supelco) [37]. All activation, equilibration, peptide wash and elution steps were performed by centrifugation at 5000 g for 5 min. Activation and equilibration of the C18 disk was performed using three rinses with 80% acetonitrile (ACN), followed by three rinses with 2% ACN. 10 ug of peptide-rich solution was loaded onto the disk and centrifuged. Desalting was conducted using three washes of 2% ACN containing 0.1% formic acid (Sigma Aldrich, USA). The desalted peptides were then eluted into glass capillary tubes using three rounds of 100 µL of 60% ACN and 0.1% formic acid. Peptides were dried under vacuum after which they were resuspended in 2% ACN, 0.1% formic acid at a concentration of 200 ng/µL.

Liquid chromatography-mass spectrometry analysis (LC-MS)

LC-MS was performed as described previously [38] and mass spectrometry was performed using LC-MS with a Dionex Ultimate3500 nano-LC coupled to a Q Exactive hybrid quadrupole-Orbitrap.

Mass spectrometry data analysis

All MS RAW files were processed using MaxQuant (version 1.5.4.1) [39] against the Uniprot human database (Proteome ID: UP00-0005640) using default settings and with match-between-runs functionality enabled. Carbamidomethylation of cysteine residues was specified as a fixed modification; variable modifications considered were oxidation of methionine and acetylation of protein Nterminus: trypsin was selected as digestion enzyme, with two missed cleavages allowed. Reverse hits to a target-decoy database and common contaminants were removed from the data sets and only protein identifications with a q-value < 0.01 were considered for further analysis. Protein groups identified in at least two replicate FLAG-Tbx3 affinity purification datasets and absent in all empty-vector control affinity purification datasets were retained for further analysis.

Immunohistochemistry (IHC)

Paraffin-embedded human sarcoma tissue sections (N = 20) were obtained from the Division of Anatomical Pathology, University of Cape Town after approval by the University of Cape Town Human Research Ethics Committee. Immunohistochemical staining was performed as previously described [16]. Sections were incubated with mouse polyclonal antinucleolin (1:25; sc-8031; Santa Cruz Biotechnology, USA) primary antibody prepared with an antibody diluent (K8006, Dako, Denmark), added to the sections and incubated for 1 h. IHC assessment was performed by a senior pathologist. Staining was scored according to four categories: 0 for "negative", 1+ for "mild staining", 2+ for "moderate staining" and 3+ for "strong staining". H-score was calculated based on the percentage of positive cells and different staining intensities using the formula H-score = [1× (% cells 1+) + 2× (% cells 2+) + 3× (% cells 3+)]. An average H-score of <150 was considered "low expression" while an H-score of \geq 150 was considered "high expression".

The cancer genome atlas (TCGA) data analysis

TCGA sarcoma cohort gene expression data (TCGA-SARC) (n = 261) was accessed via the TCGA [40] website; https://portal.gdc.cancer. gov/repository. We used Pearson's linear correlation coefficient score to evaluate the correlation between the mRNA transcripts of TBX3 and NCL in sarcoma patients. Kaplan-Meier method [41] was used to compare the overall survival duration between groups of patients with tumors that express high TBX3 and high nucleolin versus tumors that express low TBX3 and low nucleolin. Gene set enrichment analysis (GSEA) [42] was employed to determine enriched KEGG pathways between tumors that expressed high TBX3 and high nucleolin versus tumors that express low TBX3 and low nucleolin. To identify differentially expressed genes between (1) sarcomas expressing high TBX3 compared to low TBX3 and (2) sarcomas expressing high nucleolin compared to low nucleolin, we used the moderated student t-test based on the negative binomial model [43, 44].

Gene dependency analysis

To evaluate the correlation between CRISPRderived gene dependency score and the mRNA transcription abundance of NCL, we obtained data from the Achilles project [45] on the fitness of 24 sarcoma cell lines following CRI-SPR knockouts of NCL and the cell line's NCL mRNA transcription levels from the Cancer Cell Line Encyclopedia [46]. Pearson's correlation between the mRNA levels and the CRISPR dependency score was measured.

Statistical analysis

Statistical significance was determined using Two-way ANOVA or student's t-test (Excel, Microsoft, Redmond, WA, USA). Significance was accepted at *P<0.05, **P<0.01, ***P< 0.001. Unless otherwise stated, all data were obtained from at least three independent experimental repeats with error bars representing standard error of the mean (SEM). Graphs were generated using GraphPad Prism software 6.0 (GraphPad Prism software, USA).

Results

Identification of TBX3 protein partners

To identify TBX3 interacting proteins that cooperate with it in sarcomas we used an affinity pulldown assay followed by mass spectrometry (AP-MS) [47]. A fusion protein comprised of three tandem FLAG peptides fused to the TBX3 N-terminus (FLAG-Tbx3) [10] was used as bait to pull down TBX3 binding partners in chondrosarcoma cell extracts and 37 proteins were identified that met the criteria indicated in the Materials and Methods (**Table 1**).

Gene Ontology (GO) analysis of the protein partners identified showed a significant enrichment in gene expression, nucleic acid metabolism, RNA biosynthesis, processing and splicing, complex assembly, protein translation, intracellular transport, viral transcription and senescence (**Figure 1A**). The observation that 80% of identified proteins are involved in gene expression is not surprising considering the role of TBX3 as a transcription factor. A number of interacting proteins implicated in RNA splicing were also identified, which is consistent with the findings of Kumar et al. [48]. Importantly, a protein interaction network map

Gene	Protein name	Number of peptides identified	P-value
ADSSL1	adenylosuccinate synthase like 1	3	3.24E-11
CALR	calreticulin	6	3.69E-03
CBX1	chromobox homolog 1	7	1.69E-03
CBX3	chromobox homolog 3	4	1.69E-04
COPR5	cooperator of PRMT5	5	7.93E-13
DDX5	DEAD (Asp-Glu-Ala-Asp) box helicase 5	9	1.21E-12
EN03	enolase 3 (beta, muscle)	3	5.79E-10
ERH	enhancer of rudimentary homolog (Drosophila)	4	9.70E-15
H2AFX	H2A histone family, member X	10	1.27E-09
HIST1H2BA	histone cluster 1, H2ba	18	1.73E-08
HIST1H4A	histone cluster 1, H4a	5	1.36E-58
HNRNPH1	heterogeneous nuclear ribonucleoprotein H1 (H)	12	3.29E-05
HNRNPK	heterogeneous nuclear ribonucleoprotein K	6	3.17E-11
HNRNPU	heterogeneous nuclear ribonucleoprotein U	17	1.36E-03
HSPA8	Heat shock cognate 71 kDa protein	17	2.54E-29
KAT6A	K(lysine) acetyltransferase 6A	5	7.14E-03
LIMA1	LIM domain and actin binding 1	7	7.16E-03
LUC7L2	LUC7-like 2 (S. cerevisiae)	10	6.02E-11
NCL	nucleolin	5	1.57E-15
NPM1	nucleophosmin	5	3.91E-07
POU3F1	POU class 3 homeobox 1	11	2.38E-36
PPM1A	protein phosphatase, Mg2+/Mn2+ dependent, 1A	11	8.51E-06
PPM1B	protein phosphatase, Mg2+/Mn2+ dependent, 1B	11	1.69E-04
PRDX1	peroxiredoxin 1	3	9.59E-03
PRMT5	protein arginine methyltransferase 5	4	2.68E-09
PRPF19	pre-mRNA processing factor 19	10	6.00E-05
PUF60	poly-U binding splicing factor 60 KDa	9	2.17E-03
RPL18	ribosomal protein L18	6	3.63E-08
RPL23	ribosomal protein L23	9	3.42E-84
RPS15	ribosomal protein S15	7	1.19E-11
RPS3	ribosomal protein S3	9	4.89E-12
RPS5	ribosomal protein S5	4	2.08E-05
RPS8	ribosomal protein S8	4	2.67E-04
SCAPER	S-phase cyclin A-associated protein in the ER	5	9.70E-04
SNAPC1	small nuclear RNA activating complex, polypeptide 1	3	9.59E-03
TOP1	topoisomerase (DNA) I	4	4.23E-03
U2AF2	U2 small nuclear RNA auxiliary factor 2	6	2.20E-03

Table 1. TBX3-interacting proteins

generated using the STRING (Search Tool for the Retrieval of Interacting Genes/proteins) database revealed high confidence interactions between various TBX3 interacting proteins within RNA splicing and chromatin remodelling networks (**Figure 1B**). This suggests that TBX3 may function as a component of multiprotein complexes. Furthermore, the association of TBX3 with factors involved in senescence and viral transcription coincide with reports implicating TBX3 in these processes. TBX3 and nucleolin levels correlate in sarcoma patients and correlate with poor overall survival

Among the putative TBX3 interacting partners identified by detection of multiple peptides, nucleolin (**Figure 1C**) was of interest because of its overlapping roles with TBX3 in cancer cell survival, proliferation, and migration. We therefore firstly explored the clinical significance of TBX3 and nucleolin expression in sar-

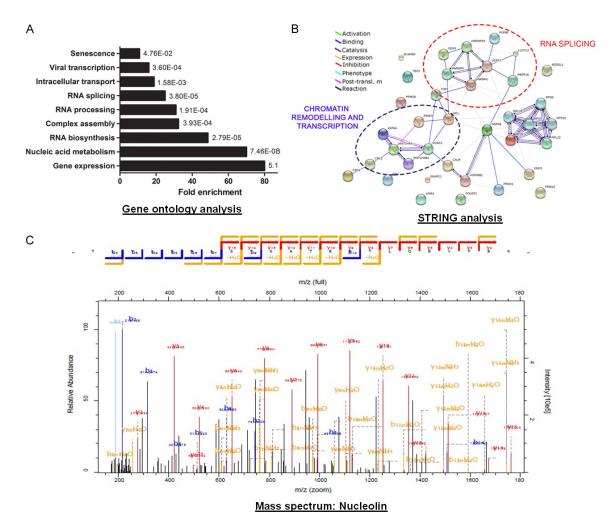
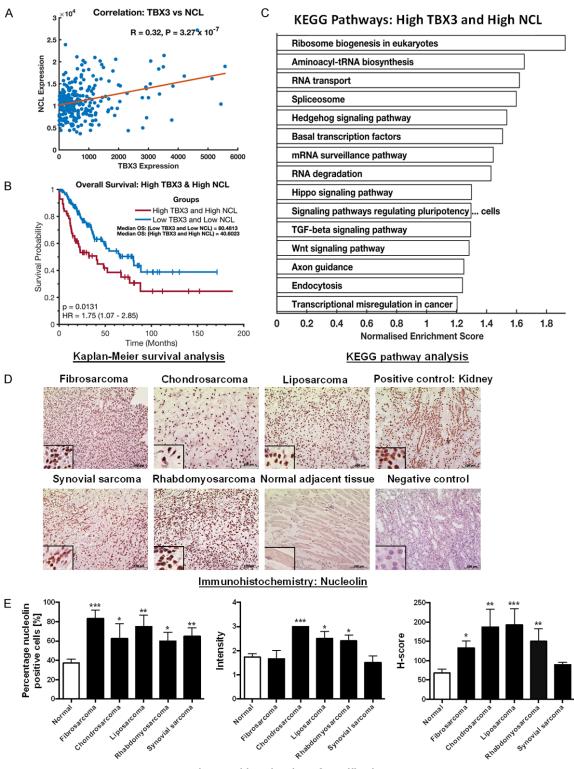


Figure 1. Identification of TBX3 protein partners. A. Functional classification of the TBX3 interactome. TBX3 interactors were categorised according to their biological function using GO analysis and plotted as fold enrichment compared to the total interactome. B. A network integrating physical interactions of identified TBX3 protein partners was generated using STRING. Nodes represent proteins; edges represent interactions. Clusters of interacting partners involved in RNA splicing and chromatin remodelling/transcription are circled in red and purple respectively. C. Representative fragmentation spectrum of one of the nucleolin peptides identified by mass spectrometry with the amino acid sequence TLVLSNLSYSATEETLQEVFEK.

coma tissues by analyzing publicly available RNA sequencing data from The Cancer Genome Atlas sarcoma cohort (TCGA-SARC). Our results showed a significant positive association between TBX3 and nucleolin expression in sarcoma patients (Pearson's rank correlation; r = 0.32, P = 3.27E-7, **Figure 2A**). Kaplan-Meier survival analysis revealed that high TBX3 and high nucleolin expression (P = 0.013) associated with poor overall patient survival (**Figure 2B**). To further explore the biological significance of high TBX3 and high nucleolin expression in sarcoma patients, we performed enrichment analysis with Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways with upregulated genes (**Figure 2C**). Several cancer related signaling pathways were enriched when both TBX3 and nucleolin are highly expressed, including the Hedgehog, Hippo, TGF-beta, and WNT signaling pathways.

Nucleolin is overexpressed in patient-derived sarcoma tissue

We have previously reported that TBX3 is expressed in patient-derived fibrosarcoma, liposarcoma, rhabdomyosarcoma, chondrosarcoma and synovial sarcoma tissues [10]. Here we show using immunohistochemistry that nucleolin is expressed in the same sarcoma



Immunohistochemistry Quantification

Figure 2. TBX3 and nucleolin (NCL) expression correlates in sarcoma patients and is associated with poor prognosis. A. Positive gene correlation between TBX3 and NCL mRNA levels across 261 TCGA sarcoma samples. Correlation coefficient R and *P*-values were calculated using Pearson correlation test. B. Kaplan-Meier plot displaying overall survival of TCGA sarcoma patients with high TBX3 and high NCL expression (red line) compared to low expression (blue line). *P*-values were calculated using the log-rank test statistic. C. Enriched KEGG pathways in TCGA sarcoma patients with high TBX3 and high NCL expression. D. Immunohistochemistry of patient-derived fibrosarcoma (N =

TBX3 and nucleolin co-operate to promote sarcoma proliferation and migration

3), chondrosarcoma (N = 4), liposarcoma (N = 4), rhabdomyosarcoma (N = 5) and synovial sarcoma (N = 4) tissue sections and adjacent normal tissues (where present) using an antibody specific to nucleolin. Human kidney tissue was included as a positive control for nucleolin. A technical negative control was included and incubated with bovine serum albumin instead of primary antibody. Haematoxylin was used as a nuclear counterstain. Representative images are shown (scale bars, 100 μ m; insets are magnified images from selected areas). E. Left panel: Column bar graph representing percentage (%) of nucleolin positive cells across tested sarcoma tissue sections. Middle panel: Scatterplot of staining intensity analysis across tested sarcoma tissue samples. Cells were quantified and graded based on staining intensity on a scale from 0 to 3 (0 = negative; 1 = mild staining; 2 = moderate staining; 3 = strong staining). Right panel: Scatterplot of H-scores across tested sarcoma tissue samples. H-score was calculated based on the percentage of positive cells and staining intensity using the following formula H-score = [1× (% cells 1+) + 2× (% cells 2+) + 3× (% cells 3+)]. The H-score ranges from 0-300. An average H-score of between 0-149 was considered "low expression" while an H-score of 150-300 was considered "high expression". Results are expressed as mean \pm SEM (bars).

tissues (representative images are shown in Figure 2D). The percentage of nucleolin positive cells was higher in all the sarcoma subtype tissues compared to normal adjacent tissues (Figure 2E left panel) and the chondrosarcoma, liposarcoma and rhabdomyosarcoma tissue sections showed a higher staining intensity (2+/3+) compared to normal adjacent tissue (1+/2+) (Figure 2E middle panel). The calculated H-scores revealed that, compared to normal adjacent tissue, nucleolin expression was significantly higher in fibrosarcoma, chondrosarcoma, liposarcoma and rhabdomyosarcoma tissue and it was on average >150 in chondrosarcoma, liposarcoma and rhabdomyosarcoma tissues revealing that these sarcoma tissues have high nucleolin expression (Figure 2E right panel).

TBX3 interacts with nucleolin via its DNAbinding domain in multiple sarcoma subtypes

To further investigate the relationship between TBX3 and nucleolin we screened sarcoma cell lines that we have previously reported to express TBX3 [10] for nucleolin expression. Like TBX3, nucleolin was expressed in cell lines derived from fibrosarcoma, chondrosarcoma, synovial sarcoma, liposarcoma and rhabdomyosarcoma (**Figure 3A**). Co-immunoprecipitation assays (**Figure 3B**) and immunofluorescence (**Figure 3C**) showed that TBX3 and nucleolin interacted and co-localized in the nucleus, and in particular with structures assumed to be nucleoli, respectively in SW1353 chondrosarcoma, RD rhabdomyosarcoma and SW872 liposarcoma cells.

To determine the domain(s) of TBX3 that are responsible for its interaction with nucleolin, we performed a series of co-immunoprecipitation experiments in which SW1353 chondrosarcoma cells were transiently transfected with full length TBX3 (TBX3 FL), full length TBX3 harbouring a disrupted DNA binding domain (TBX3 DBM) or the TBX3 N-terminus lacking the activation and dominant R1 repression domains (referred to as TBX3 N-term) (**Figure 3D**). All TBX3 constructs were expressed (**Figure 3E** left panel) and endogenous nucleolin co-immunoprecipitated with TBX3 FL and TBX3 N-term but not with the TBX3 DBM (**Figure 3E** right panel). These results suggest that TBX3 directly interacts with nucleolin through the DNA binding domain.

Nucleolin promotes proliferation and migration of chondrosarcoma, rhabdomyosarcoma and liposarcoma cells

TBX3 promotes proliferation and migration of chondrosarcoma [10, 32], rhabdomyosarcoma [16] and liposarcoma cells (<u>Supplementary</u> <u>Figure 1</u>) and to explore whether TBX3 requires nucleolin for these functions, we first investigated whether nucleolin phenocopies TBX3 in these sarcoma subtypes. We thus depleted nucleolin in these cells by siRNA (**Figure 4A**) and show that knockdown of nucleolin (siNuc) significantly inhibited their proliferation (**Figure 4B**) and migration (**Figure 4C**). These results demonstrate that depletion of nucleolin phenocopies TBX3 knockdown in sarcomas.

TBX3 and nucleolin co-operate to promote proliferation and migration of chondrosarcoma cells

We next investigated whether the pro-proliferative and pro-migratory roles of TBX3 in sarcomas are nucleolin-dependent. To this end, we determined the impact of depleting nucleolin by siRNA on FLAG-Empty and FLAG-Tbx3 SW-1353 cell proliferation and migration. Western

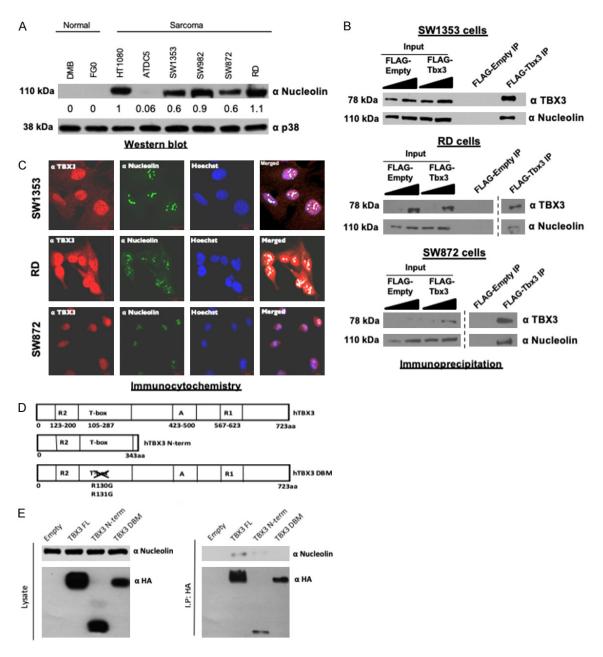


Figure 3. Nucleolin is overexpressed in sarcoma cell lines and directly interacts with TBX3. A. Western blot analysis of nucleolin expression in DMB and FGO normal human skin fibroblast, HT1080 fibrosarcoma, ATDC5 and SW1353 chondrosarcoma, SW982 synovial sarcoma, SW872 liposarcoma and RD rhabdomyosarcoma cell lines. p38 was used as a loading control. B. Immunoprecipitation with an anti-FLAG antibody and SW1353, RD and SW872 protein extracts. Purified immune complexes were analysed by western blotting with antibodies to nucleolin and TBX3. C. Nucleolin co-localizes with TBX3. SW1353, RD and SW872 cells were processed for immunofluorescence using antibodies specific to nucleolin (Alexa 488; green channel) and TBX3 (Cy3; red channel). Confocal microscopic images from 20 fields of view were captured and representative images taken at 400× magnification are shown. Nuclei were visualized by co-staining cells with Hoechst. Merged images from Zeiss software co-localization analysis with white pixels indicating TBX3 and nucleolin co-localization. D. The TBX3 DNA-binding domain is required for its association with nucleolin. Schematic diagram of the human full length (FL) TBX3 protein (hTBX3), TBX3 N terminal (N-term) and TBX3 DNA-binding domain mutant (DBM) expression constructs. "A" and "R" denotes activation and repression domains, respectively. E. Left panel: western blots showing expression of transfected HA-tagged TBX3 FL, TBX3 N-term and TBX3 DBM proteins. Right panel: protein extracts from HA-tagged empty, TBX3 FL, TBX3 N-term, and TBX3 DBM cells immunoprecipitated using an anti-HA antibody. Purified immune complexes were analysed by western blotting with the indicated antibodies.

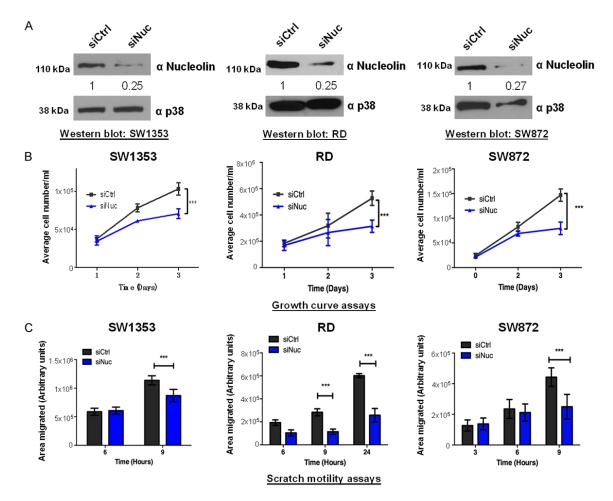


Figure 4. Nucleolin impacts sarcoma cell proliferation and migration. (A) Western blots confirming successful nucleolin knockdown in SW1353, RD and SW872 cells after transfection with siNuc for 48 h. Total p38 served as a loading control. (B) Growth curves and (C) 2D-Scratch motility assays show that knockdown of nucleolin inhibits SW1353, RD and SW872 cell proliferation and migration respectively. Two-way ANOVA was used to compare between groups, and data are the mean ± SEM of three independent experiments, ***P<0.001.

blotting was used to confirm the overexpression of TBX3 and the knockdown of nucleolin (Figure 5A). Consistent with our previous observations, overexpressing TBX3 enhanced the proliferative (Figure 5B) and migratory (Figure 5C) ability of the SW1353 cells (FLAG-Tbx3 siCtrl) and knocking down nucleolin (FLAG-Empty siNuc) inhibited their proliferation (Figure 5B) and migration (Figure 5C). Importantly, depleting nucleolin in TBX3 overexpressing cells (FLAG-Tbx3 siNuc) abrogated the positive effect of TBX3 on proliferation and migration (Figure 5B, 5C). This suggests that nucleolin and TBX3 function together to promote cell proliferation and migration. To confirm these findings, the impact of ectopically overexpressing nucleolin (GFP-Nucleolin) on SW1353 shTBX3 cell proliferation and migration was assessed. Indeed, the results show that the TBX3 knockdown phenotype could not be rescued by overexpressing nucleolin (**Figure 5E**, **5F**). **Figure 5D** confirms the overexpression of GFP-Nucleolin and knockdown of TBX3 (sh-TBX3). Taken together, these results demonstrate that both TBX3 and nucleolin positively impact cell proliferation and migration and require each other to do so.

The nucleolin targeting aptamer, AS1411, reduces sarcoma cell viability and migration

Figure 6A shows that a large number of cancer cell lines (green) including several sarcoma cell lines (blue) are highly dependent on nucleolin for their survival suggesting that nucleolin is a promising drug target for sarcoma treatment.

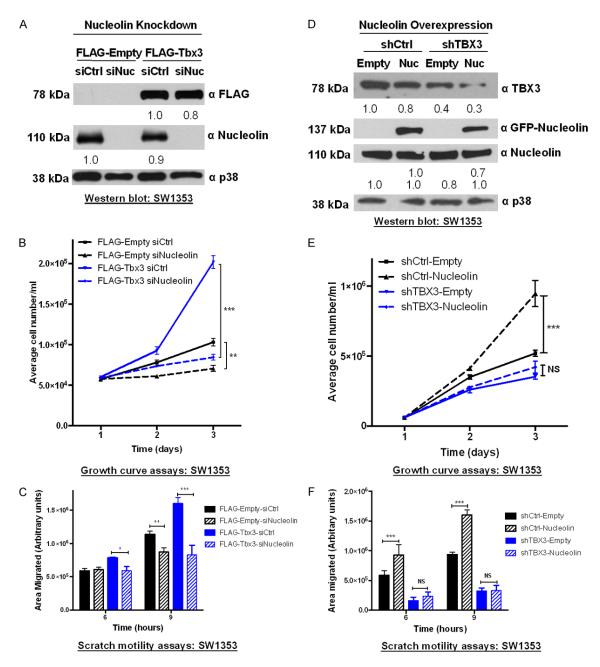


Figure 5. Nucleolin and TBX3 co-operate to promote chondrosarcoma cell proliferation and migration. (A-C) Knockdown of nucleolin inhibits TBX3-induced cell (B) proliferation and (C) migration. SW1353 FLAG-Empty and FLAG-Tbx3 cells were transiently transfected with 50 nM siNucleolin (siNuc) or siControl (siCtrl). (A) Western blotting confirmed overexpression of FLAG-Tbx3 and knockdown of nucleolin. p38 was used as a loading control. (D-F) Ectopic overexpression of nucleolin cannot rescue the effect of TBX3 knockdown on SW1353 cell (E) proliferation and (F) migration. SW1353 shTBX3 and shCtrl cells were transiently transfected with either pEGFP-Empty (Empty) or pEGFP-Nucleolin (Nuc). (D) Efficacy of TBX3 knockdown and nucleolin overexpression was confirmed by western blotting. Each result represents two independent experiments performed in triplicate. Two-way ANOVA was used to compare between groups, data are the mean ± SEM of two independent experiments performed in triplicate, *P<0.05; **P<0.01; ***P<0.001.

Here, we show that the nucleolin-targeting oligonucleotide aptamer AS1411 significantly inhibits the viability of SW1353, RD and SW872 sarcoma cell lines in a dose-dependent manner without affecting the viability of the nonmalignant FGO and DMB human fibroblasts

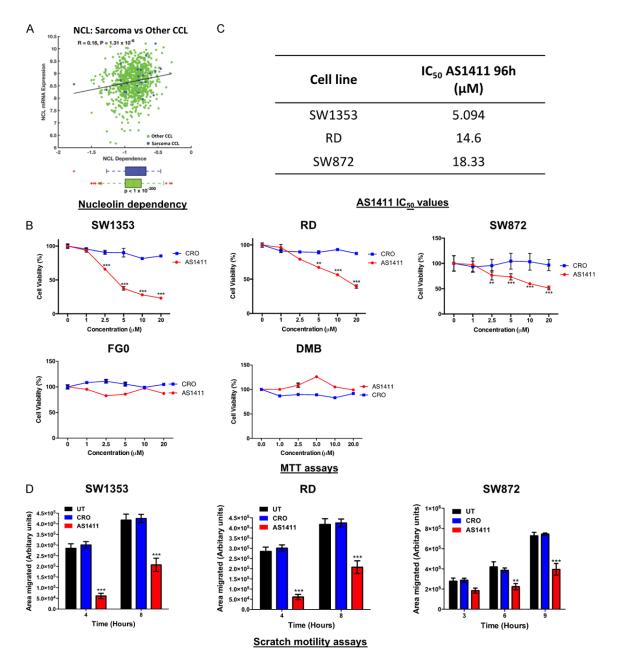


Figure 6. The nucleolin targeting aptamer, AS1411, inhibits SW1343, RD and SW872 cell viability and migration. A. Pearson's correlation between nucleolin (NCL) mRNA expression levels and CRISPR-derived gene dependency score of NCL across 850 cancer cell lines (CCL) (green) and 24 sarcoma cell lines (blue). B. Cells were treated with 10 μ M AS1411 or CR0 for 96 h after which an MTT assay was performed. Two-way ANOVA was used to compare groups, data are the mean ± SEM of three independent experiments performed in quadruplicate, *P<0.05; **P<0.01; ***P<0.001. C. IC₅₀ values for AS1411-treated SW1353, SW872 and RD cells. D. Cells were treated with 10 μ M AS1411 or CR0 for 96 h after which 2D-Scratch motility assays were performed. Bar graphs show the distance migrated between 3-9 h. Results are representative of three independent experiments. Student's t-test was used to compare between groups, **P<0.01; ***P<0.001; error bars represent mean ± SEM.

(Figure 6B). These results suggest that AS14-11 is highly selective for the sarcoma cell lines tested and its IC_{50} values for SW1353, RD and SW872 cells are presented in Figure 6C. Moreover, at a clinically relevant concentration of

AS1411 (10 μ M) [49], sarcoma cell migration was significantly inhibited compared to CRO- or untreated cells (**Figure 6D**). These results suggest that AS1411 may be an effective strategy to treat a diverse range of sarcomas.

AS1411 promotes cytoplasmic accumulation of TBX3 and nucleolin

AS1411 disrupts nucleolin protein-interactions, and alters the localization of nucleolinbinding partners [50, 51]. We therefore examined the subcellular localization of nucleolin and TBX3 after treatment of SW1353 cells with AS1411 or CRO for 48 h. Confocal microscopy revealed that, while TBX3 predominantly localized to the nucleus in untreated and CROtreated cells, AS1411 treatment led to an accumulation of TBX3 in the cytoplasm (Figure 7A). These results were confirmed by western blotting which showed that whereas levels of nuclear TBX3 and nucleolin decreased, cytoplasmic levels of both proteins increased (Figure 7B). Importantly, in the CRO treated cells there was a negligible change in protein distribution.

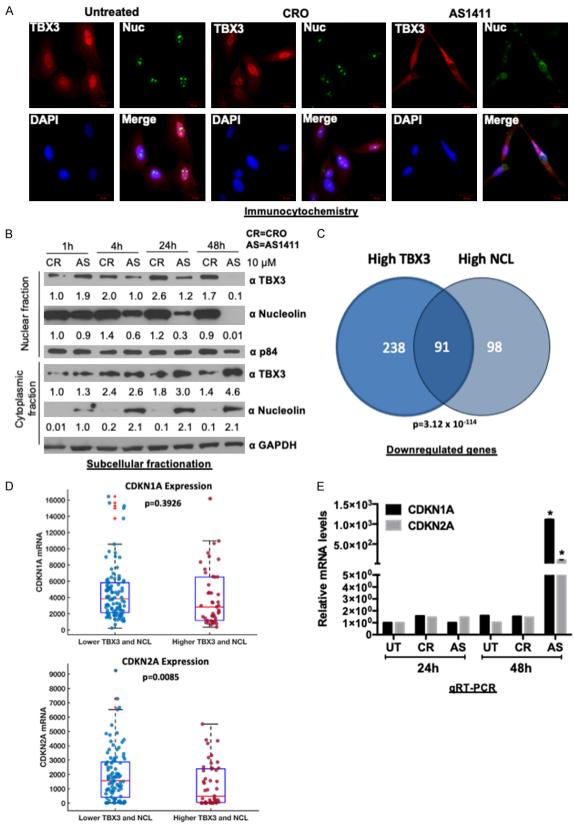
AS1411 causes upregulation of TBX3 tumor suppressor target genes

Previous studies have demonstrated that TBX3 and nucleolin (NCL) promote proliferation by repressing key tumor suppressor genes, including CDKN1A (p21) and CDKN2A (p14/ 19^{ARF}) [8, 10, 19, 32]. To determine whether these genes are potentially co-regulated by TBX3 and NCL, we analysed the TCGA-SARC database for differentially expressed genes in sarcoma patients expressing high vs low TBX3 or high vs low NCL. We found 329 and 189 genes downregulated in high-TBX3 and high-NCL sarcomas respectively (Figure 7C). When the two datasets were compared, 91 of the downregulated genes were common to both (Figure 7C), indicating that TBX3 and NCL may regulate a common set of target genes. Out of the 91 overlapping genes, almost half (39 genes) were identified to be known or predicted tumor suppressor genes (Supplementary Table 1). Furthermore, in sarcomas that express high levels of TBX3 and high levels of NCL their known target genes CDKN1A and CDKN2A were downregulated (Figure 7D). Based on these findings and our observations that TBX3 and NCL levels decrease in the nuclei of AS1411-treated SW1353 cells, we speculated that CDKN1A and CDKN2A would be upregulated in AS1411-treated cells. Indeed, gRT-PCR analyses show that compared to untreated and CRO-treated SW1353 cells, AS1411-treated SW1353 cells exhibited a significantly higher expression of CDKN1A and CDKN2A after 48 h (Figure 7E). These results suggest that TBX3 and NCL cooperatively repress *CDKN1A* and *CDKN2A* and that treatment with AS1411 relieves this repression resulting in inhibition of SW1353 cell proliferation.

Discussion

Sarcomas are highly aggressive cancers that mostly affect children and are notoriously recalcitrant to current standard treatment regimes. This has driven a mechanism-based search for novel molecular therapeutic targets. Recently, the transcription factor TBX3 was shown to promote several features of tumorigenesis in chondrosarcoma, rhabdomyosarcoma and liposarcoma cells including cell proliferation, migration, invasion and tumor formation [10, 16, 32]. These findings highlight the potential of targeting TBX3 as a therapeutic approach to treat sarcomas. However, direct targeting of transcription factors is challenging. This study shows that nucleolin is a TBX3 cofactor and that they co-operate to promote sarcoma cell proliferation and migration and that the nucleolin-targeting aptamer, AS1411, may be a promising therapeutic to treat TBX3driven sarcomas.

In the present study, we identify 37 candidate TBX3 interacting partners that potentially cooperate with it to promote proliferation and migration of sarcomas. Of these, nucleolin was of interest because, in cancer, there is a significant overlap in the mechanisms that regulate TBX3 and nucleolin and their functions. For example, nucleolin and TBX3 are upregulated by cyclin A-CDK2, c-Myc, and AKT, and their levels peak during S-phase [16, 19, 25, 29-31, 35, 52-54]. Furthermore, like TBX3, nucleolin inhibits apoptosis and prevents cell cycle arrest by suppressing the expression of p53 and p21 and when phosphorylated by AKT, it promotes migration and EMT [21-23]. In addition, knockdown of nucleolin by siRNA, an antibody, or the peptide HB-19, dramatically blocked migration of gastric cancer, multiple glioma and endothelial cells [55, 56]. In the present study, we also show that depleting nucleolin by siRNA inhibits proliferation and migration of chondrosarcoma, rhabdomyosarcoma and liposarcoma cells. Importantly, we show that nucleolin and TBX3 require each other to induce sarcoma cell proliferation and migration. Indeed, knocking down nucleolin in TBX3 overexpressing cells abrogated TBX3induced cell proliferation and migration, while overexpressing nucleolin in shTBX3 cells could not rescue TBX3's impact on these processes.



TBX3 and nucleolin co-operate to promote sarcoma proliferation and migration

Figure 7. AS1411 mislocalizes TBX3 and nucleolin to the cytoplasm and causes upregulation of TBX3 and nucleolin tumor suppressor target genes. A. Confocal microscopy images of SW1353 cells treated with 10 μ M AS1411 or CRO for 48 h and processed for immunofluorescence with antibodies to TBX3 (Cy3; red channel) and nucleolin (Alexa 488; green channel). Confocal microscopic images from 20 fields of view were captured and representative images taken at 400× magnification are shown (scale bars, 20 μ m). Nuclei were visualized by co-staining cells with DAPI. B. SW1353 cells were treated with 10 μ M AS1411 or CRO for 48 h. Nuclear and cytoplasmic proteins were separated and analyzed by western blotting using primary antibodies to TBX3, nucleolin, GAPDH (cytoplasmic fraction marker) and p84 (nuclear fraction marker). C. Venn diagram analysis across two groups of genes showing downregulated genes in TCGA sarcoma patients expressing high TBX3 (n = 329) versus high nucleolin (NCL) (n = 189). 91 overlapping genes (hypergeometric *P*-value = 3.12×10^{-114}). D. Analysis of *CDKN1A* and *CDKN2A* expression in TCGA sarcoma patient samples expressing low TBX3 and NCL versus high TBX3 and NCL. E. qRT-PCR analysis of TBX3 tumor suppressor target genes (*CDKN1A* (p21^{CIP1}) and *CDKN2A* (p14^{ARF})) after 24 h and 48 h treatment with 10 μ M AS1411. Student's t-test was used to compare between groups, *P<0.05; error bars represent mean \pm SEM.

Nucleolin is an ubiquitously expressed protein which has been implicated in many cellular functions [17]. Interestingly, nucleolin is overexpressed in a growing list of cancers where it is exclusively localised at the cell surface and complexes with several oncoproteins to promote cancer cell proliferation, survival, angiogenesis, and metastasis. Due to its expression on the surface of cancer cells nucleolin has been identified as a promising anti-cancer drug target because it can serve as a receptor which allows for tumor-specific uptake of targeted chemotherapeutic drugs without affecting normal cells. Indeed, numerous strategies aimed at controlling nucleolin abundance and relocation are being evaluated in pre-clinical and clinical trials. These include polyclonal antibodies and peptides such as HB-19, which inhibit cell proliferation and migration and induce cell death of leukaemia and lymphoma cells [57, 58]. More recently, oligonucleotide aptamers have also been a focus of translational research and aptamers targeting nucleolin exhibit anti-cancer activities in a growing list of neoplasms. The most successful is AS1411 which has a high binding affinity for cell-surface nucleolin and is internalized into cancer cells by macropinocytosis [59]. AS1411 exhibits anticancer activity in almost every cancer cell type tested without affecting normal cells and in xenograft mouse models bearing tumors of diverse histological origin, including lung, prostate and renal cancer, glioma and leukemia [31, 60]. Importantly, findings from phase I clinical trials revealed good overall tolerability in patients with advanced solid tumors and acute myeloid leukemia [61]. This is further supported by phase II studies, which demonstrated promising anti-cancer activity of AS1411 in a patient with metastatic renal cell carcinoma (NCT00740441) as well as in combination with cytarabine for the treatment of patients with acute myeloid leukemia (NCT00512083), where an improved overall survival and acceptable safety profile was observed [61]. Taken together, these studies strengthen the candidacy of AS1411 as an anti-cancer agent and additional studies on the effect of AS1411 in sarcomas are warranted.

There is evidence to suggest that AS1411 disrupts nucleolin protein-protein interactions and consequently their oncogenic functions, although the exact downstream mechanism(s) involved is not fully understood [26, 50]. Here we provide evidence that AS1411 reduced the levels of nuclear TBX3 and nucleolin and increased cytoplasmic TBX3 and nucleolin. This correlated with an upregulation in expression of the TBX3 and nucleolin target genes CDKN1A (p21^{CIP1}) and CDKN2A (p14^{ARF}) which is likely due to reduced occupancy of TBX3/ nucleolin at their promoter regions. Nucleolin has previously been shown to act as a molecular chaperone, shuttling various proteins, including transcription factors, between the cytoplasm and the nucleus and TBX3 may therefore be one of these transcription factors [17, 62]. Our findings agree with another study that reported that AS1411 induced a nuclearto-cytoplasmic shift of a protein arginine methyltransferase 5 (PRMT5) repressor complex which led to increased levels of the PRMT5 target genes, suppressor of tumorigenicity 7 (ST7) and cyclin E2 [50]. Furthermore, using bioinformatic tools we show that high TBX3 and high nucleolin share 91 overlapping genes of which 39 are either well known or predicted tumor suppressor genes. This suggests that TBX3 and nucleolin co-operate to regulate a common set of target genes including CDKN1A (p21^{CIP1}) and CDKN2A (p14^{ARF}). Future work will need to determine whether nucleolin and TBX3 co-operate to directly repress the common set of tumor suppressor genes identified in this study.

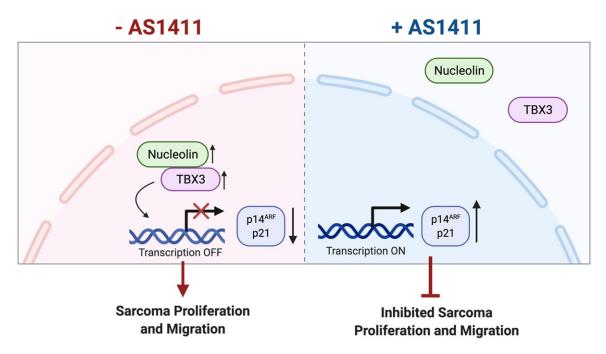


Figure 8. Oncogenic activity of TBX3 and nucleolin in sarcomas and proposed model of action of the nucleolin targeting aptamer AS1411. The oncogenic TBX3 co-operates with nucleolin to promote proliferation and migration of chondrosarcoma, rhabdomyosarcoma and liposarcoma cells. Treatment with the nucleolin targeting aptamer AS1411 leads to mislocalization of TBX3 and nucleolin to the cytoplasm, de-repression of TBX3/nucleolin target tumor suppressor genes as well as inhibited sarcoma proliferation and migration. Hence, AS1411 may represent a viable option for treating a diverse range of sarcomas.

Both TBX3 and nucleolin can bind RNA and are constituents of the spliceosome and it is therefore plausible that TBX3 and nucleolin also co-operate to post-transcriptionally regulate gene expression [17, 48]. Indeed, nucleolin promotes proliferation by repressing p53 translation [22, 27] and during development TBX3 regulates the alternative splicing of Disks large homolog (3Dlg3) and Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (Nfkb1) by directly binding T-elements in their mRNA transcripts [48]. Furthermore, nucleolin can regulate multiple genes involved in tumorigenesis, including gastrin, AKT1, CCN1 (cyclin I) and Sp1, by binding their mRNA and either enhancing their translation or protein stability [19]. The impact of TBX3 and/or nucleolin on the regulation of these mRNAs could be relevant to their oncogenic roles. It is also worth noting that in the nucleoli, nucleolin has an established and pivotal role in ribosome biogenesis and cancer cells have increased ribosome biogenesis which leads to accelerated translational capacity and thus enables high proliferation rates [17, 63]. It has therefore been suggested that targeting the nucleoli machinery may be a promising anti-cancer therapy, and 20 out of 36 chemotherapeutic drugs currently in clinical use inhibit ribosome biogenesis [64]. Considering the accumulating reports linking ribosome biogenesis and cancer it will be worth exploring if TBX3 co-operates with nucleolin to regulate ribosome biogenesis. This is a particularly intriguing possibility given our observation that the majority of co-localization between TBX3 and nucleolin occurs within the nucleoli.

Conclusion

Taken together, findings from this study provide valuable insights into sarcoma biology and add to our current knowledge of the mechanisms that regulate the molecular role of TBX3 in sarcoma development and progression. Indeed, we show that together. TBX3 and nucleolin are key to promoting cell proliferation and migration in chondrosarcoma, liposarcoma and rhabdomyosarcoma cells and disrupting their interaction with the nucleolin targeting aptamer AS1411 reverses these oncogenic processes. Furthermore, our observations suggest that AS1411-induced perturbations of the TBX3:nucleolin complex and de-repression of tumor suppressor genes may contribute to the biological effects of AS1411 (Figure 8). Our findings provide a strong motivation for investigating the use of AS1411 to treat multiple sarcomas driven by the TBX3-nucleolin interaction.

Acknowledgements

We would like to thank the Confocal and Light Microscope Imaging Facility at the University of Cape Town (UCT), South Africa for providing microscopy services and Mrs Subash Govender for help with the immunohistochemistry protocol. The graphical abstract (Figure 8) was created using BioRender (https://biorender.com/) as part of the Academic License. This work was supported by grants from the South Africa Medical Research Council (SA MRC), the South Africa National Research Foundation (NRF ZA), Cancer Association of South Africa (CANSA) and the University of Cape Town. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. JMB thanks the NRF for a South African Research Chair grant.

Disclosure of conflict of interest

None.

Address correspondence to: Sharon Prince, Division of Cell Biology, Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Cape Town 7925, South Africa. Tel: +27-21-406-6240; E-mail: sharon.prince@uct.ac.za

References

- [1] Burningham Z, Hashibe M, Spector L and Schiffman JD. The epidemiology of sarcoma. Clin Sarcoma Res 2012; 2: 14.
- [2] Bray F, Colombet M and Mery L. Cancer incidence in five continents. Lyon: International Agency for Research on Cancer; 2017.
- [3] Fletcher CDM, Bridge JA, Hogendoorn P and Mertens F. WHO classification of tumours of soft tissue and bone. Lyon: IARC Press; 2013.
- [4] Damerell V, Pepper MS and Prince S. Molecular mechanisms underpinning sarcomas and implications for current and future therapy. Signal Transduct Target Ther 2021; 6: 246.
- [5] Bleloch JS, Ballim RD, Kimani S, Parkes J, Panieri E, Willmer T and Prince S. Managing sarcoma: where have we come from and where are we going? Ther Adv Med Oncol 2017; 9: 637-659.
- [6] Khan SF, Damerell V, Omar R, Du Toit M, Khan M, Maranyane HM, Mlaza M, Bleloch J, Bellis C, Sahm BDB, Peres J, ArulJothi KN and Prince S. The roles and regulation of TBX3 in development and disease. Gene 2020; 726: 144223.

- [7] Willmer T, Cooper A, Peres J, Omar R and Prince S. The T-box transcription factor 3 in development and cancer. Biosci Trends 2017; 11: 254-266.
- [8] Brummelkamp TR, Kortlever RM, Lingbeek M, Trettel F, MacDonald ME, van Lohuizen M and Bernards R. TBX-3, the gene mutated in ulnarmammary syndrome, is a negative regulator of p19ARF and inhibits senescence. J Biol Chem 2002; 277: 6567-6572.
- [9] Yarosh W, Barrientos T, Esmailpour T, Lin L, Carpenter PM, Osann K, Anton-Culver H and Huang T. TBX3 is overexpressed in breast cancer and represses p14 ARF by interacting with histone deacetylases. Cancer Res 2008; 68: 693-699.
- [10] Willmer T, Cooper A, Sims D, Govender D and Prince S. The T-box transcription factor 3 is a promising biomarker and a key regulator of the oncogenic phenotype of a diverse range of sarcoma subtypes. Oncogenesis 2016; 5: e199.
- [11] Dong L, Dong Q, Chen Y, Li Y, Zhang B, Zhou F, Lyu X, Chen GG, Lai P, Kung HF and He ML. Novel HDAC5-interacting motifs of Tbx3 are essential for the suppression of E-cadherin expression and for the promotion of metastasis in hepatocellular carcinoma. Signal Transduct Target Ther 2018; 3: 22.
- [12] Dong L, Lyu X, Faleti OD and He MI. The special stemness functions of Tbx3 in stem cells and cancer development. Semin Cancer Biol 2019; 57: 105-110.
- [13] Feng X, Yao W, Zhang Z, Yuan F, Liang L, Zhou J, Liu S and Song J. T-box transcription factor Tbx3 contributes to human hepatocellular carcinoma cell migration and invasion by repressing E-cadherin expression. Oncol Res 2018; 26: 959-966.
- [14] Wang Y. Expression level of TBX3 gene in renal carcinoma and its clinical significance. Oncol Lett 2018; 15: 4325-4240.
- [15] Krstic M, Kolendowski B, Cecchini MJ, Postenka CO, Hassan HM, Andrews J, MacMillan CD, Williams KC, Leong HS, Brackstone M, Torchia J, Chambers AF and Tuck AB. TBX3 promotes progression of pre-invasive breast cancer cells by inducing EMT and directly up-regulating SLUG. J Pathol 2019; 248: 191-203.
- [16] Sims D, Maranyane HM, Damerell V, Govender D, Isaacs AW, Peres J and Prince S. The c-Myc/ AKT1/TBX3 axis is important to target in the treatment of embryonal rhabdomyosarcoma. Cancers 2020; 12: 501.
- [17] Tuteja R and Tuteja N. Nucleolin: a multifunctional major nucleolar phosphoprotein. Crit Rev Biochem Mol Biol 1998; 33: 407-436.
- [18] Ginisty H, Sicard H, Roger B and Bouvet P. Structure and functions of nucleolin. J Cell Sci 1999; 112: 761-772.
- [19] Abdelmohsen K and Gorospe M. RNA-binding protein nucleolin in disease. RNA Biol 2012; 9: 799-808.

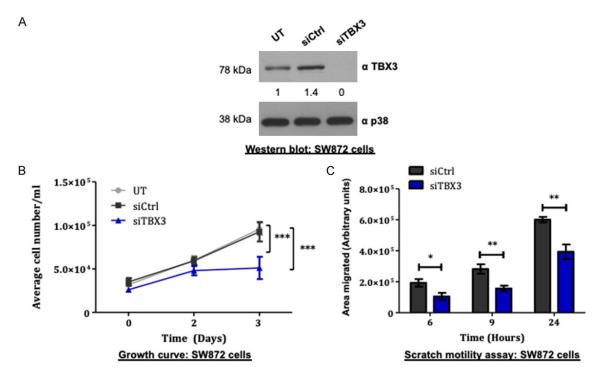
- [20] Berger CM, Gaume X and Bouvet P. The roles of nucleolin subcellular localization in cancer. Biochimie 2015; 113: 78-85.
- [21] Bhatt P, D'Avout C, Kane NS, Borowiec JA and Saxena A. Specific domains of nucleolin interact with Hdm2 and antagonize Hdm2mediated p53 ubiquitination. FEBS J 2012; 279: 370-383.
- [22] Takagi M, Absalon MJ, McLure KG and Kastan MB. Regulation of p53 translation and induction after DNA damage by ribosomal protein L26 and nucleolin. Cell 2005; 123: 49-63.
- [23] Wu DM, Zhang P, Liu RY, Sang YX, Zhou C, Xu GC, Yang JL, Tong AP and Wang CT. Phosphorylation and changes in the distribution of nucleolin promote tumor metastasis via the PI3K/Akt pathway in colorectal carcinoma. FEBS Lett 2014; 588: 1921-1929.
- [24] Li Q, Zhu Y, Hou L, Wang J, Hu G, Fang X, Hu Y, Tao T, Wei X, Tang H, Huang B and Hu W. C23 promotes tumorigenesis via suppressing p53 activity. Oncotarget 2016; 7: 58274-58285.
- [25] Grinstein E, Shan Y, Karawajew L, Snijders PJ, Meijer CJ, Royer HD and Wernet P. Cell cyclecontrolled interaction of nucleolin with the retinoblastoma protein and cancerous cell transformation. J Biol Chem 2006; 281: 22223-22235.
- [26] Schokoroy S, Juster D, Kloog Y and Pinkas-Kramarski R. Disrupting the oncogenic synergism between nucleolin and Ras results in cell growth inhibition and cell death. PLoS One 2013; 8: e75269.
- [27] Yang A, Shi G, Zhou C, Lu R, Li H, Sun L and Jin Y. Nucleolin maintains embryonic stem cell self-renewal by suppression of p53 proteindependent pathway. J Biol Chem 2011; 286: 43370-43382.
- [28] González V, Guo K, Hurley L and Sun D. Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. J Biol Chem 2009; 284: 23622-23635.
- [29] Sarcevic B, Lilischkis R and Sutherland RL. Differential phosphorylation of T-47D human breast cancer cell substrates by D1-, D3-, E-, and A-type cyclin-CDK complexes. J Biol Chem 1997; 272: 33327-33337.
- [30] Sirri V, Roussel P, Gendron MC and Hernandez-Verdun D. Amount of the two major Ag-NOR proteins, nucleolin, and protein B23 is cell-cycle dependent. Cytometry 1997; 28: 147-156.
- [31] Bates PJ, Laber DA, Miller DM, Thomas SD and Trent JO. Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. Exp Mol Pathol 2009; 86: 151-164.
- [32] Willmer T, Hare S, Peres J and Prince S. The T-box transcription factor TBX3 drives proliferation by direct repression of the p21(WAF1) cyclin-dependent kinase inhibitor. Cell Div 2016; 11: 6.

- [33] van der Westhuyzen DR, Coetzee GA, Demasius IP, Harley EH, Gevers W, Baker SG and Seftel HC. Low density lipoprotein receptor mutations in South African homozygous familial hypercholesterolemic patients. Arteriosclerosis 1984; 4: 238-247.
- [34] Peres J, Davis E, Mowla S, Bennett DC, Li JA, Wansleben S and Prince S. The highly homologous T-box transcription factors, TBX2 and TBX3, have distinct roles in the oncogenic process. Genes Cancer 2010; 1: 272-282.
- [35] Willmer T, Peres J, Mowla S, Abrahams A and Prince S. The T-box factor TBX3 is important in S-phase and is regulated by c-Myc and cyclin A-CDK2. Cell Cycle 2015; 14: 3173-3183.
- [36] Prince S, Wiggins T, Hulley PA and Kidson SH. Stimulation of melanogenesis by tetradecanoylphorbol 13-acetate (TPA) in mouse melanocytes and neural crest cells. Pigment Cell Res 2003; 16: 26-34.
- [37] Rappsilber J, Ishihama Y and Mann M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem 2003; 75: 663-670.
- [38] Nel AJ, Garnett S, Blackburn JM and Soares NC. Comparative reevaluation of FASP and enhanced FASP methods by LC-MS/MS. J Proteome Res 2015; 14: 1637-1642.
- [39] Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV and Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res 2011; 10: 1794-1805.
- [40] Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C and Stuart JM. The cancer genome atlas pan-cancer analysis project. Nat Genet 2013; 45: 1113-1120.
- [41] Goel MK, Khanna P and Kishore J. Understanding survival analysis: Kaplan-Meier estimate. Int J Ayurveda Res 2010; 1: 274-278.
- [42] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genomewide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.
- [43] Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008; 5: 621-628.
- [44] Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 2013; 4: 392.
- [45] Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, East A, Ali LD, Lizotte PH, Wong TC, Jiang G, Hsiao J, Mermel CH, Getz

G, Barretina J, Gopal S, Tamayo P, Gould J, Tsherniak A, Stransky N, Luo B, Ren Y, Drapkin R, Bhatia SN, Mesirov JP, Garraway LA, Meyerson M, Lander ES, Root DE and Hahn WC. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineagespecific dependencies in ovarian cancer. Proc Natl Acad Sci U S A 2011; 108: 12372-12377.

- [46] Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, Barretina J, Gelfand ET, Bielski CM, Li H, Hu K, Andreev-Drakhlin AY, Kim J, Hess JM, Haas BJ, Aguet F, Weir BA, Rothberg MV, Paolella BR, Lawrence MS. Akbani R. Lu Y. Tiv HL. Gokhale PC. de Weck A, Mansour AA, Oh C, Shih J, Hadi K, Rosen Y, Bistline J, Venkatesan K, Reddy A, Sonkin D, Liu M, Lehar J, Korn JM, Porter DA, Jones MD, Golji J, Caponigro G, Taylor JE, Dunning CM, Creech AL, Warren AC, McFarland JM, Zamanighomi M, Kauffmann A, Stransky N, Imielinski M, Maruvka YE, Cherniack AD, Tsherniak A, Vazquez F, Jaffe JD, Lane AA, Weinstock DM, Johannessen CM, Morrissey MP, Stegmeier F, Schlegel R, Hahn WC, Getz G, Mills GB, Boehm JS, Golub TR, Garraway LA and Sellers WR. Next-generation characterization of the cancer cell line encyclopedia. Nature 2019; 569: 503-508.
- [47] Nakedi KC, Nel AJ, Garnett S, Blackburn JM and Soares NC. Comparative Ser/Thr/Tyr phosphoproteomics between two mycobacterial species: the fast growing mycobacterium smegmatis and the slow growing mycobacterium bovis BCG. Front Microbiol 2015; 6: 237.
- [48] Kumar P, Franklin S, Emechebe U, Hu H, Moore B, Lehman C, Yandell M and Moon AM. TBX3 regulates splicing in vivo : a novel molecular mechanism for ulnar-mammary syndrome. PLoS Genet 2014; 10: e1004247.
- [49] Rosenberg JE, Bambury RM, Van Allen EM, Drabkin HA, Lara PN Jr, Harzstark AL, Wagle N, Figlin RA, Smith GW, Garraway LA, Choueiri T, Erlandsson F and Laber DA. A phase II trial of AS1411 (a novel nucleolin-targeted DNA aptamer) in metastatic renal cell carcinoma. Invest New Drugs 2014; 32: 178-187.
- [50] Teng Y, Girvan AC, Casson LK, Pierce WM Jr, Qian M, Thomas SD and Bates PJ. AS1411 alters the localization of a complex containing protein arginine methyltransferase 5 and nucleolin. Cancer Res 2007; 67: 10491-10500.
- [51] Wolfson E, Solomon S, Schmukler E, Goldshmit Y and Pinkas-Kramarski R. Nucleolin and ErbB2 inhibition reduces tumorigenicity of ErbB2-positive breast cancer. Cell Death Dis 2018; 9: 47.
- [52] Fu Z and Fenselau C. Proteomic evidence for roles for nucleolin and poly[ADP-ribosyl] transferase in drug resistance. J Proteome Res 2005; 4: 1583-1591.

- [53] Cheng M, Wang D and Roussel MF. Expression of c-Myc in response to colony-stimulating factor-1 requires mitogen-activated protein kinase kinase-1. J Biol Chem 1999; 274: 6553-6558.
- [54] Hovanessian AG, Soundaramourty C, El Khoury D, Nondier I, Svab J and Krust B. Surface expressed nucleolin is constantly induced in tumor cells to mediate calcium-dependent ligand internalization. PLoS One 2010; 5: e15787.
- [55] Fujiki H, Watanabe T and Suganuma M. Cellsurface nucleolin acts as a central mediator for carcinogenic, anti-carcinogenic, and disease-related ligands. J Cancer Res Clin Oncol 2014; 140: 689-699.
- [56] Koutsioumpa M, Polytarchou C, Courty J, Zhang Y, Kieffer N, Mikelis C, Skandalis SS, Hellman U, Iliopoulos D and Papadimitriou E. Interplay between $\alpha\nu\beta3$ integrin and nucleolin regulates human endothelial and glioma cell migration. J Biol Chem 2013; 288: 343-354.
- [57] Huang Y, Shi H, Zhou H, Song X, Yuan S and Luo Y. The angiogenic function of nucleolin is mediated by vascular endothelial growth factor and nonmuscle myosin. Blood 2006; 107: 3564-3571.
- [58] Krust B, El Khoury D, Nondier I, Soundaramourty C and Hovanessian AG. Targeting surface nucleolin with multivalent HB-19 and related Nucant pseudopeptides results in distinct inhibitory mechanisms depending on the malignant tumor cell type. BMC Cancer 2011; 11: 333.
- [59] Reyes-Reyes EM, Teng Y and Bates PJ. A new paradigm for aptamer therapeutic AS1411 action: uptake by macropinocytosis and its stimulation by a nucleolin-dependent mechanism. Cancer Res 2010; 70: 8617-8629.
- [60] Yazdian-Robati R, Bayat P, Oroojalian F, Zargari M, Ramezani M, Taghdisi SM and Abnous K. Therapeutic applications of AS1411 aptamer, an update review. Int J Biol Macromol 2020; 155: 1420-1431.
- [61] Romano S, Fonseca N, Simões S, Gonçalves J and Moreira JN. Nucleolin-based targeting strategies for cancer therapy: from targeted drug delivery to cytotoxic ligands. Drug Discov Today 2019; 24: 1985-2001.
- [62] Jerke U, Tkachuk S, Kiyan J, Stepanova V, Kusch A, Hinz M, Dietz R, Haller H, Fuhrman B and Dumler I. Stat1 nuclear translocation by nucleolin upon monocyte differentiation. PLoS One 2009; 4: e8302.
- [63] van Sluis M and McStay B. Ribosome biogenesis: achilles heel of cancer? Genes Cancer 2014; 5: 152-153.
- [64] Burger K, Mühl B, Harasim T, Rohrmoser M, Malamoussi A, Orban M, Kellner M, Gruber-Eber A, Kremmer E, Hölzel M and Eick D. Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. J Biol Chem 2010; 285: 12416-12425.



Supplementary Figure 1. TBX3 promotes proliferation and migration of liposarcoma cells. A. Western blots confirming successful TBX3 knockdown in SW872 cells after treatment with siTBX3 for 72 h, with p38 as a loading control. B. Knockdown of TBX3 inhibits cell proliferation of SW872 cells. Growth curve assays were performed over a 3-day period and cells harvested by trypsinisation and counted on a haemocytometer. C. Knockdown of TBX3 inhibits cell migration of SW872 cells. 2D-Scratch motility assays were performed and migration was measured at 6-24 h intervals. Two-way ANOVA was used to compare between groups, and data are the mean ± SEM of three independent experiments, ***P<0.001.

Gene Symbol	Gene Name	Cancer Type	Reference
CTSG	cathepsin G	Acute myeloid leukemia	[1]
MYL1	myosin light chain 1	Rhabdomyosarcoma	[2]
PGC	progastricsin	Prostate, ovarian and breast cancer	[3-6]
GKN1	gastrokine 1	Gastric cancer	[7]
LIPF	lipase F, gastric type	Gastric cancer	[8]
TFF2	trefoil factor 2	Colon cancer	[9]
TNNC2	troponin C2	Rhabdomyosarcoma, colon cancer, head and neck squamous cell carcinoma	[10-12]
PIGR	polymeric immunoglobulin receptor	Pancreatic and periampullary cancer	[13]
TFF1	trefoil factor 1	Gastric cancer, pancreatic cancer	[14-17]
MMP12	matrix metallopeptidase 12	Colon, colorectal, gastric, breast and lung cancer, melanoma	[18]
APOA4	apolipoprotein A4	Downregulated in hepatocellular carcinoma and ovarian cancer	[19, 20]
CSN1S1	casein alpha s1	Breast cancer	[21, 22]
CTSE	cathepsin E	Prostate cancer, melanoma	[23]
NCR2	natural cytotoxicity triggering receptor 2	Breast, cervical, colorectal and prostate cancer, melanoma	[24, 25]
CLDN18	claudin 18	Lung adenocarcinoma, gastric cancer	[26, 27]
GKN2	gastrokine 2	Gastric cancer	[28]
AGR2	anterior gradient 2	Lung and pancreatic adenocarcinoma, breast cancer	[29]
SLC26A3	solute carrier family 26 member 3	Colon cancer	[30]
PSCA	prostate stem cell antigen	Esophageal squamous cell carcinoma, prostate cancer, gallbladder cancer	[31-33]
TRPM1	transient receptor potential cation channel subfamily M member 1	Melanoma	[34]

Supplementary Table 1. Overlapping known and predicted tumor suppressor ger

TBX3 and nucleolin co-operate to promote sarcoma proliferation and migration

RPL3L	ribosomal protein L3 like	Endometrial cancer	[35]
REG3A	regenerating family member 3 alpha	Gastric and breast cancer	[36, 37]
ITLN1	intelectin 1	Gastric cancer, neuroblastoma, ovarian cancer	[38-40]
MSMB	microseminoprotein beta	Prostate and ovarian cancer	[41, 42]
TRIM40	tripartite motif containing 40	Gastric and colorectal cancer	[43, 44]
GHRL	ghrelin and obestatin prepropeptide	Gastric and breast cancer	[45, 46]
CLEC1B	C-type lectin domain family 1 member B	Hepatocellular carcinoma	[47]
SERPINB11	serpin family B member 11	Oral squamous cell carcinomas	[48]
SERPINA4	serpin family A member 4	Colorectal, lung, and breast cancer, hepatocellular carcinoma	[49-53]
HAND1	heart and neural crest derivatives expressed 1	Colorectal cancer	[54]
ERN2	endoplasmic reticulum to nucleus signaling 2	Ovarian cancer	[55]
MUC17	mucin 17	Gastric and colorectal cancer	[56, 57]
FAM3D	family with sequence similarity 3 member D	Colon cancer	[58]
CLRN1	clarin 1	Pituitary prolactinoma	[59]
PPP1R3A	protein phosphatase 1 regulatory subunit 3A	Hematological malignancies, non-small cell lung carcinoma	[60, 61]
ANXA10	annexin A10	Prostate, bladder, lung, cervical and cancer, hepatocellular carcinoma	[62-66]
PRLHR	prolactin releasing hormone receptor	Breast cancer	[67]
A4GNT	alpha-1,4-N-acetylglucosaminyltransferase	Gastric cancer	[68]
NR0B2	nuclear receptor subfamily 0 group B member 2	Clear Cell Renal Cell Carcinoma, hepatocellular carcinoma	[69, 70]

Supplementary references

- [1] Jin W, Wu K, Li YZ, Yang WT, Zou B, Zhang F and Wang KK. AML1-ETO targets and suppresses cathepsin G, a serine protease, which is able to degrade AML1-ETO in t(8;21) acute myeloid leukemia. Oncogene 2013; 32: 1978-87.
- [2] Walters ZS, Villarejo-Balcells B, Olmos D, Buist TW, Missiaglia E, Allen R, Al-Lazikani B, Garrett MD, Blagg J and Shipley J. JARID2 is a direct target of the PAX3-FOXO1 fusion protein and inhibits myogenic differentiation of rhabdomyosarcoma cells. Oncogene 2014; 33: 1148-57.
- [3] Shen S, Jiang J and Yuang Y. Pepsinogen C expression, regulation and its relationship with cancer. Cancer Cell Int 2017; 17: 57.
- [4] Serra C, Vizoso F, Lamelas ML, Rodríguez JC, González LO, Merino AM, Baltasar A, Pérez-Vázquez MT and Medrano J. Comparative study of two androgen-induced markers (apolipoprotein D and pepsinogen C) in female and male breast carcinoma. Int J Surg Investig 2000; 2: 183-92.
- [5] Serra Díaz C, Vizoso F, Rodríguez JC, Merino AM, González LO, Baltasar A, Pérez-Vázquez MT and Medrano J. Expression of pepsinogen C in gynecomastias and male breast carcinomas. World J Surg 1999; 23: 439-45.
- [6] Rojo JV, Merino AM, González LO and Vizoso F. Expression and clinical significance of pepsinogen C in epithelial ovarian carcinomas. Eur J Obstet Gynecol Reprod Biol 2002; 104: 58-63.
- [7] Yoon JH, Kang YH, Choi YJ, Park IS, Nam SW, Lee JY, Lee YS and Park WS. Gastrokine 1 functions as a tumor suppressor by inhibition of epithelial-mesenchymal transition in gastric cancers. J Cancer Res Clin Oncol 2011; 137: 1697-704.
- [8] Kong Y, Zheng Y, Jia Y, Li P and Wang Y. Decreased LIPF expression is correlated with DGKA and predicts poor outcome of gastric cancer. Oncol Rep 2016; 36: 1852-60.
- [9] Dubeykovskaya Z, Dubeykovskiy A, Worthley DL, Westphalen CB, Asfaha S, Kitur K and Wang TC. Mo2079 Tff2 is a novel tumor suppressor that inhibits expansion of Gr1+CD11b+ myeloid-derived suppressor cells and blocks colon carcinogenesis. Gastroenterology 2012; 143: e25.
- [10] Jin Y and Yang Y. Bioinformatics-based discovery of PYGM and TNNC2 as potential biomarkers of head and neck squamous cell carcinoma. Biosci Rep 2019; 39: BSR20191612.
- [11] Liu Z, Zhang X, Lei H, Lam N, Carter S, Yockey O, Xu M, Mendoza A, Hernandez ER, Wei JS, Khan J, Yohe ME, Shern JF and Thiele CJ. CASZ1 induces skeletal muscle and rhabdomyosarcoma differentiation through a feed-forward loop with MYOD and MYOG. Nat Commun 2020; 11: 911.
- [12] Sun Y, Mironova V, Chen Y, Lundh EPF, Zhang Q, Cai Y, Vasiliou V, Zhang Y, Garcia-Milian R, Khan SA and Johnson CH. Molecular pathway analysis indicates a distinct metabolic phenotype in women with right-sided colon cancer. Transl Oncol 2020; 13: 42-56.
- [13] Fristedt R, Elebro J, Gaber A, Jonsson L, Heby M, Yudina Y, Nodin B, Uhlén M, Eberhard J and Jirström K. Reduced expression of the polymeric immunoglobulin receptor in pancreatic and periampullary adenocarcinoma signifies tumour progression and poor prognosis. PLoS One 2014; 9: e112728.

TBX3 and nucleolin co-operate to promote sarcoma proliferation and migration

- [14] Yamaguchi J, Yokoyama Y, Kokuryo T, Ebata T, Enomoto A and Nagino M. Trefoil factor 1 inhibits epithelialmesenchymal transition of pancreatic intraepithelial neoplasm. J Clin Invest 2018; 128: 3619-29.
- [15] Lefebvre O, Chenard MP, Masson R, Linares J, Dierich A, LeMeur M, Wendling C, Tomasetto C, Chambon P and Rio MC. Gastric mucosa abnormalities and tumorigenesis in mice lacking the pS2 trefoil protein. Science 1996; 274: 259-62.
- [16] Soutto M, Belkhiri A, Piazuelo MB, Schneider BG, Peng D, Jiang A, Washington MK, Kokoye Y, Crowe SE, Zaika A, Correa P, Peek RM Jr and El-Rifai W. Loss of TFF1 is associated with activation of NF-κB-mediated inflammation and gastric neoplasia in mice and humans. J Clin Invest 2011; 121: 1753-67.
- [17] Heuer J, Heuer F, Stürmer R, Harder S, Schlüter H, Braga Emidio N, Muttenthaler M, Jechorek D, Meyer F and Hoffmann W. The tumor suppressor TFF1 occurs in different forms and interacts with multiple partners in the human gastric mucus barrier: indications for diverse protective functions. Int J Mol Sci 2020; 21: 2508.
- [18] Decock J, Thirkettle S, Wagstaff L and Edwards DR. Matrix metalloproteinases: protective roles in cancer. J Cell Mol Med 2011; 15: 1254-65.
- [19] Bharali D, Banerjee B, Bharadwaj M, Husain S and Kar P. Expression analysis of apolipoproteins AI & AIV in hepatocellular carcinoma: a protein-based hepatocellular carcinoma-associated study. Indian J Med Res 2018; 147: 361.
- [20] Dieplinger H, Ankerst DP, Burges A, Lenhard M, Lingenhel A, Fineder L, Buchner H and Stieber P. Afamin and apolipoprotein A-IV: novel protein markers for ovarian cancer. Cancer Epidemiol Biomarkers Prev 2009; 18: 1127-33.
- [21] Mou MA, Keya NA, Islam M, Hossain MJ, Al Habib MS, Alam R, Rana S, Samad A and Ahammad F. Validation of CSN1S1 transcriptional expression, promoter methylation, and prognostic power in breast cancer using independent datasets. Biochem Biophys Rep 2020; 24: 100867.
- [22] Bonuccelli G, Castello-Cros R, Capozza F, Martinez-Outschoorn UE, Lin Z, Tsirigos A, Xuanmao J, Whitaker-Menezes D, Howell A, Lisanti MP and Sotgia F. The milk protein α-casein functions as a tumor suppressor via activation of STAT1 signaling, effectively preventing breast cancer tumor growth and metastasis. Cell Cycle 2012; 11: 3972-82.
- [23] Kawakubo T, Okamoto K, Iwata J, Shin M, Okamoto Y, Yasukochi A, Nakayama KI, Kadowaki T, Tsukuba T and Yamamoto K. Cathepsin E prevents tumor growth and metastasis by catalyzing the proteolytic release of soluble TRAIL from tumor cell surface. Cancer Res 2007; 67: 10869-78.
- [24] Eisenberg V, Shamalov K, Meir S, Hoogi S, Sarkar R, Pinker S, Markel G, Porgador A and Cohen CJ. Targeting multiple tumors using T-cells engineered to express a natural cytotoxicity receptor 2-based chimeric receptor. Front Immunol 2017; 8: 1212.
- [25] Barrow AD, Edeling MA, Trifonov V, Luo J, Goyal P, Bohl B, Bando JK, Kim AH, Walker J, Andahazy M, Bugatti M, Melocchi L, Vermi W, Fremont DH, Cox S, Cella M, Schmedt C and Colonna M. Natural killer cells control tumor growth by sensing a growth factor. Cell 2018; 172: 534-548, e19.
- [26] Luo J, Chimge NO, Zhou B, Flodby P, Castaldi A, Firth AL, Liu Y, Wang H, Yang C, Marconett CN, Crandall ED, Offringa IA, Frenkel B and Borok Z. CLDN18.1 attenuates malignancy and related signaling pathways of lung adenocarcinoma in vivo and in vitro. Int J Cancer 2018; 143: 3169-80.
- [27] Oshima T, Shan J, Okugawa T, Chen X, Hori K, Tomita T, Fukui H, Watari J and Miwa H. Down-regulation of Claudin-18 is associated with the proliferative and invasive potential of gastric cancer at the invasive front. PLoS One 2013; 8: e74757.
- [28] Ouyang J, Pan X, Lin H, Hu Z, Xiao P and Hu H. GKN2 increases apoptosis, reduces the proliferation and invasion ability of gastric cancer cells through down-regulating the JAK/STAT signaling pathway. Am J Transl Res 2017; 9: 803-11.
- [29] Sommerova L, Ondrouskova E, Vojtesek B and Hrstka R. Suppression of AGR2 in a TGF-β-induced smad regulatory pathway mediates epithelial-mesenchymal transition. BMC Cancer 2017; 17: 546.
- [30] Bhutia YD, Babu E, Ramachandran S, Yang S, Thangaraju M and Ganapathy V. SLC transporters as a novel class of tumour suppressors: identity, function and molecular mechanisms. Biochem J 2016; 473: 1113-24.
- [31] Zhang LY, Wu JL, Qiu HB, Dong SS, Zhu YH, Lee VH, Qin YR, Li Y, Chen J, Liu HB, Bi J, Ma S, Guan XY and Fu L. PSCA acts as a tumor suppressor by facilitating the nuclear translocation of RB1CC1 in esophageal squamous cell carcinoma. Carcinogenesis 2016; 37: 320-32.
- [32] Heinrich MC, Göbel C, Kluth M, Bernreuther C, Sauer C, Schroeder C, Möller-Koop C, Hube-Magg C, Lebok P, Burandt E, Sauter G, Simon R, Huland H, Graefen M, Heinzer H, Schlomm T and Heumann A. PSCA expression is associated with favorable tumor features and reduced PSA recurrence in operated prostate cancer. BMC Cancer 2018; 18: 612.
- [33] Ono H, Hiraoka N, Lee YS, Woo SM, Lee WJ, Choi IJ, Saito A, Yanagihara K, Kanai Y, Ohnami S, Chiwaki F, Sasaki H, Sakamoto H, Yoshida T and Saeki N. Prostate stem cell antigen, a presumable organ-dependent tumor suppressor gene, is down-regulated in gallbladder carcinogenesis. Genes Chromosomes Cancer 2012; 51: 30-41.

- [34] Guo H, Carlson JA and Slominski A. Role of TRPM in melanocytes and melanoma. Exp Dermatol 2012; 21: 650-4.
- [35] Takai N, Kawamata N, Walsh CS, Gery S, Desmond JC, Whittaker S, Said JW, Popoviciu LM, Jones PA, Miyakawa I and Koeffler HP. Discovery of epigenetically masked tumor suppressor genes in endometrial cancer. Mol Cancer Res 2005; 3: 261-9.
- [36] Qiu YS, Liao GJ and Jiang NN. REG3A overexpression suppresses gastric cancer cell invasion, proliferation and promotes apoptosis through PI3K/Akt signaling pathway. Int J Mol Med 2018; 41: 3167-3174.
- [37] Zhang P, Zhang J, Yu M, Zhang X and Li H. Over-expression of REG3A gene can inhibit the proliferation, invasion and migration of human breast cancer. Int J Clin Exp Pathol 2016; 9: 12104-13.
- [38] Li D, Zhao X, Xiao Y, Mei H, Pu J, Xiang X, Jiao W, Song H, Qu H, Huang K, Zheng L and Tong Q. Intelectin 1 suppresses tumor progression and is associated with improved survival in gastric cancer. Oncotarget 2015; 6: 16168-82.
- [39] Li D, Mei H, Pu J, Xiang X, Zhao X, Qu H, Huang K, Zheng L and Tong Q. Intelectin 1 suppresses the growth, invasion and metastasis of neuroblastoma cells through up-regulation of N-myc downstream regulated gene 2. Mol Cancer 2015; 14: 47.
- [40] Au-Yeung CL, Yeung TL, Achreja A, Zhao H, Yip KP, Kwan SY, Onstad M, Sheng J, Zhu Y, Baluya DL, Co NN, Rynne-Vidal A, Schmandt R, Anderson ML, Lu KH, Wong STC, Nagrath D and Mok SC. ITLN1 modulates invasive potential and metabolic reprogramming of ovarian cancer cells in omental microenvironment. Nat Commun 2020; 11: 3546.
- [41] Beke L, Nuytten M, Van Eynde A, Beullens M and Bollen M. The gene encoding the prostatic tumor suppressor PSP94 is a target for repression by the Polycomb group protein EZH2. Oncogene 2007; 26: 4590-5.
- [42] Henriksen R, Lundwall Å, Udby L and Fernlund P. The expression of β-microseminoprotein but not CRISP3 is reduced in ovarian cancer and correlates to survival. Anticancer Res 2012; 32: 3993-9.
- [43] Noguchi K, Okumura F, Takahashi N, Kataoka A, Kamiyama T, Todo S and Hatakeyama S. TRIM40 promotes neddylation of IKK and is downregulated in gastrointestinal cancers. Carcinogenesis 2011; 32: 995-1004.
- [44] Eberhardt W, Haeussler K, Nasrullah U and Pfeilschifter J. Multifaceted roles of TRIM proteins in colorectal carcinoma. Int J Mol Sci 2020; 21: 7532.
- [45] Hu XL, Zhu YJ, Hu CH, You L, Wu J, He XY, Huang WJ and Wu ZH. Ghrelin affects gastric cancer progression by activating AMPK signaling pathway. Biochem Genet 2021; 59: 652-67.
- [46] Au CC, Furness JB, Britt K, Oshchepkova S, Ladumor H, Soo KY, Callaghan B, Gerard C, Inghirami G, Mittal V, Wang Y, Huang XY, Spector JA, Andreopoulou E, Zumbo P, Betel D, Dow L and Brown KA. Three-dimensional growth of breast cancer cells potentiates the anti-tumor effects of unacylated ghrelin and AZP-531. Elife 2020; 9: e56913.
- [47] Hu K, Wang ZM, Li JN, Zhang S, Xiao ZF and Tao YM. CLEC1B expression and PD-L1 expression predict clinical outcome in hepatocellular carcinoma with tumor hemorrhage. Transl Oncol 2018; 11: 552-8.
- [48] Shiiba M, Nomura H, Shinozuka K, Saito K, Kouzu Y, Kasamatsu A, Sakamoto Y, Murano A, Ono K, Ogawara K, Uzawa K and Tanzawa H. Down-regulated expression of SERPIN genes located on chromosome 18q21 in oral squamous cell carcinomas. Oncol Rep 2010; 24: 241-9.
- [49] Sun HM, Mi YS, Yu FD, Han Y, Liu XS, Lu S, Zhang Y, Zhao SL, Ye L, Liu TT, Yang DH, Sun XF, Qin XB, Zhou ZG, Tang HM and Peng ZH. SERPINA4 is a novel independent prognostic indicator and a potential therapeutic target for colorectal cancer. Am J Cancer Res 2016; 6: 1636-49.
- [50] Wang NQ, Zou J and Diao Y. Plasmid-mediated expression of kallistatin and its biological activity in lung cancer related cells. Yao Xue Xue Bao 2013; 48: 359-65.
- [51] Shiau AL, Teo ML, Chen SY, Wang CR, Hsieh JL, Chang MY, Chang CJ, Chao J, Chao L, Wu CL and Lee CH. Inhibition of experimental lung metastasis by systemic lentiviral delivery of kallistatin. BMC Cancer 2010; 10: 245.
- [52] Li P, Guo Y, Bledsoe G, Yang Z, Chao L and Chao J. Kallistatin induces breast cancer cell apoptosis and autophagy by modulating Wnt signaling and microRNA synthesis. Exp Cell Res 2016; 340: 305-14.
- [53] Tse LY, Sun X, Jiang H, Dong X, Fung PW, Farzaneh F and Xu R. Adeno-associated virus-mediated expression of kallistatin suppresses local and remote hepatocellular carcinomas. J Gene Med 2008; 10: 508-17.
- [54] Tan J, Yang X, Jiang X, Zhou J, Li Z, Lee PL, Li B, Robson P and Yu Q. Integrative epigenome analysis identifies a Polycomb-targeted differentiation program as a tumor-suppressor event epigenetically inactivated in colorectal cancer. Cell Death Dis 2014; 5: e1324.
- [55] Ow GS, Ivshina AV, Fuentes G and Kuznetsov VA. Identification of two poorly prognosed ovarian carcinoma subtypes associated with CHEK2 germ-line mutation and non-CHEK2 somatic mutation gene signatures. Cell Cycle 2014; 13: 2262-80.
- [56] Yang B, Wu A, Hu Y, Tao C, Wang JM, Lu Y and Xing R. Mucin 17 inhibits the progression of human gastric cancer by limiting inflammatory responses through a MYH9-p53-RhoA regulatory feedback loop. J Exp Clin Cancer Res 2019; 38: 283.

- [57] Katkoori V, Shanmugam C, Lucas BL, Batra SK, Grizzle WE, Siegal G and Manne U. Abstract #4362: MUC17 is a potential tumor suppressor gene in colorectal adenocarcinoma. Cancer Res 2009; 69.
- [58] Liang W, Peng X, Li Q, Wang P, Lv P, Song Q, She S, Huang S, Chen K, Gong W, Yuan W, Thovarai V, Yoshimura T, O'huigin C, Trinchieri G, Huang J, Lin S, Yao X, Bian X, Kong W, Xi J, Wang JM and Wang Y. FAM3D is essential for colon homeostasis and host defense against inflammation associated carcinogenesis. Nat Commun 2020; 11: 5912.
- [59] Wang C, Tan C, Wen Y, Zhang D, Li G, Chang L, Su J and Wang X. FOXP1-induced IncRNA CLRN1-AS1 acts as a tumor suppressor in pituitary prolactinoma by repressing the autophagy via inactivating Wnt/β-catenin signaling pathway. Cell Death Dis 2019; 10: 499.
- [60] Inoue K, Kohno T, Takakura S, Morishita K, Takita J, Hayashi Y, Mizoguchi H and Yokota J. Alterations of the PPP1R3 gene in hematological malignancies. Int J Oncol 2000; 17: 717-21.
- [61] Kohno T, Takakura S, Yamada T, Okamoto A, Tanaka T and Yokota J. Alterations of the PPP1R3 gene in human cancer. Cancer Res 1999; 59: 4170-4.
- [62] Hung MS, Chen YC, Lin PY, Li YC, Hsu CC, Lung JH, You L, Xu Z, Mao JH, Jablons DM and Yang CT. Cul4A modulates invasion and metastasis of lung cancer through regulation of ANXA10. Cancers (Basel) 2019; 11: 618.
- [63] Miyazawa Y, Sekine Y, Kato H, Furuya Y, Koike H and Suzuki K. Simvastatin up-regulates annexin A10 that can inhibit the proliferation, migration, and invasion in androgen-independent human prostate cancer cells. Prostate 2017; 77: 337-49.
- [64] Munksgaard PP, Mansilla F, Brems Eskildsen AS, Fristrup N, Birkenkamp-Demtröder K, Ulhøi BP, Borre M, Agerbæk M, Hermann GG, Orntoft TF and Dyrskjøt L. Low ANXA10 expression is associated with disease aggressiveness in bladder cancer. Br J Cancer 2011; 105: 1379-87.
- [65] Quiskamp N, Poeter M, Raabe CA, Hohenester UM, König S, Gerke V and Rescher U. The tumor suppressor annexin A10 is a novel component of nuclear paraspeckles. Cell Mol Life Sci 2014; 71: 311-29.
- [66] Peng X, Huang C, Sun Z, Zhang G and Zhu Y. ANXA10 plays a suppressor role by regulating RhoA-ROCK signaling pathway in hepatocellular cancer. Int J Clin Exp Med 2019; 12: 12068-76.
- [67] Nouhi Z, Chughtai N, Hartley S, Cocolakis E, Lebrun JJ and Ali S. Defining the role of prolactin as an invasion suppressor hormone in breast cancer cells. Cancer Res 2006; 66: 1824-32.
- [68] Karasawa F, Shiota A, Goso Y, Kobayashi M, Sato Y, Masumoto J, Fujiwara M, Yokosawa S, Muraki T, Miyagawa S, Ueda M, Fukuda MN, Fukuda M, Ishihara K and Nakayama J. Essential role of gastric gland mucin in preventing gastric cancer in mice. J Clin Invest 2012; 122: 923-34.
- [69] Zhang Y, Xu P, Park K, Choi Y, Moore DD and Wang L. Orphan receptor small heterodimer partner suppresses tumorigenesis by modulating cyclin D1 expression and cellular proliferation. Hepatology 2008; 48: 289-98.
- [70] Kudryavtseva AV, Nyushko KM, Zaretsky AR, Shagin DA, Sadritdinova AF, Fedorova MS, Savvateeva MV, Guvatova ZG, Pudova EA, Alekseev BY, Dmitriev AA and Snezhkina AV. Suppression of NR0B2 gene in clear cell renal cell carcinoma is associated with hypermethylation of its promoter. Mol Biol (Mosk) 2018; 52: 482-8.