Published online 8 August 2019 *Nucleic Acids Research, 2019, Vol. 47, No. 18* **9619–9636**
doi: 10.1093/nar/gkz699

EWS–FLI1 modulated alternative splicing of ARID1A reveals novel oncogenic function through the BAF complex

Saravana P. Selvanathan[1,*], Garrett T. Graham[1], Alexander R. Grego[1], Tabari M. Baker[2], J. Robert Hogg[3], Mark Simpson[4], Mona Batish[4,5], Brian Crompton[6,7], Kimberly Stegmaier[6,7], Eleni M. Tomazou[8], Heinrich Kovar[8,9], Aykut Üren[1] and Jeffrey A. Toretsky[1,*]

[1]Departments of Oncology and Pediatrics, Georgetown University, Washington, DC 20057, USA, [2]Caris Life Sciences, Phoenix, AZ, USA, [3]Biochemistry and Biophysics Center, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA, [4]Department of Microbiology, Biochemistry and Molecular Genetics, Rutgers University, Newark, NJ 07103, USA, [5]Department of Medical and Molecular Sciences, University of Delaware, Newark, DE 19716, USA, [6]Department of Pediatric Oncology, Dana-Farber Cancer Institute and Boston Children's Hospital, Boston, MA, USA, [7]Broad Institute, Cambridge, MA, USA, [8]Children's Cancer Research Institute, St. Anna Kinderkrebsforschung, Vienna, Austria and [9]Department of Pediatrics, Medical University of Vienna, Vienna, Austria

Received December 11, 2018; Revised July 23, 2019; Editorial Decision July 29, 2019; Accepted August 01, 2019

ABSTRACT

Connections between epigenetic reprogramming and transcription or splicing create novel mechanistic networks that can be targeted with tailored therapies. Multiple subunits of the chromatin remodeling BAF complex, including ARID1A, play a role in oncogenesis, either as tumor suppressors or oncogenes. Recent work demonstrated that EWS–FLI1, the oncogenic driver of Ewing sarcoma (ES), plays a role in chromatin regulation through interactions with the BAF complex. However, the specific BAF subunits that interact with EWS–FLI1 and the precise role of the BAF complex in ES oncogenesis remain unknown. In addition to regulating transcription, EWS–FLI1 also alters the splicing of many mRNA isoforms, but the role of splicing modulation in ES oncogenesis is not well understood. We have identified a direct connection between the EWS–FLI1 protein and ARID1A isoform protein variant ARID1A-L. We demonstrate here that ARID1A-L is critical for ES maintenance and supports oncogenic transformation. We further report a novel feed-forward cycle in which EWS–FLI1 leads to preferential splicing of *ARID1A-L,* promoting ES growth, and ARID1A-L reciprocally promotes EWS–FLI1 protein stability. Dissecting this interaction may lead to improved cancer-specific drug targeting.

INTRODUCTION

The AT-rich interactive domain-containing protein 1A (*ARID1A*, *BAF250a*, *SMARCF1*) gene encodes a central component of the BAF complex (**B**RG1/BRM **A**ssociated **F**actor). The BAF complex is a set of evolutionarily conserved ATP-dependent chromatin remodeling proteins organized around core scaffold proteins (1). The BAF complex plays a critical role in many cellular processes including proliferation, differentiation, motility and DNA repair (2). The BAF complex has both oncogenic and tumor suppressive activities, depending on the presence of subunit mutations and the biological context (2). This apparent paradoxical activity may be better understood with emerging knowledge that BAF complexes collaborate with both normal and aberrant transcription factors such as AP-1, PU.1 and EWS–FLI1 (3–6).

The first evidence of the involvement of BAF complexes in carcinogenesis was provided by the identification of biallelic, truncating mutations of the *BAF47* (*INI1, SNF5*) gene in malignant rhabdoid tumors, a highly aggressive childhood cancer (7). Kadoch *et al.* later demonstrated that the SS18-SSX fusion protein functions primarily by ejecting BAF47 from BAF complexes resulting in oncogenic transformation (8). In recent studies, many cancers have been

*To whom correspondence should be addressed. Tel: +1 202 687 8655; Fax: +1 202 687 1434; Email: jat42@georgetown.edu
Correspondence may also be addressed to Saravana P. Selvanathan. Tel: +1 202 687 2518; Fax: +1 202 687 1434; Email: sps63@georgetown.edu

identified to have BAF complex mutations or loss of BAF complex proteins (9). Another recent study found that loss of ARID1A led to development of colon carcinoma in mice (10). Taken together, these results indicate that the BAF complexes have an impact on a broad range of cellular processes vital to oncogenesis through their role in chromatin remodeling.

Ewing sarcoma (ES) is a highly aggressive pediatric cancer of the bone and soft tissue characterized by the t(11;22)(q24;q12) *EWSR1-FLI1* chromosomal fusion oncogene that produces the EWS–FLI1 fusion protein. EWS–FLI1 is a well-described transcriptional regulator and modulator of RNA splicing (11–14). Recent transcriptome analyses showed that EWS–FLI1 alters the production of many mRNA isoforms; however, the role of RNA splicing and the associated changes in isoform ratios in ES oncogenesis remain unknown (11). For example, alternative splicing of BCL family mRNA, long known to modulate cell survival, alters the chemosensitivity of ES (15). A small molecule inhibitor of EWS–FLI1 protein interactions, YK-4-279, directly binds to EWS–FLI1; YK-4-279 reverses EWS–FLI1-induced alternative splicing and has become useful for dissecting functional activities (11,16). Previous studies showed that YK-4-279 blocks RNA Helicase A (RHA) and p68 (DDX5) from binding to EWS–FLI1 (11), however other protein interactions have not been extensively studied. EWS–FLI1 remains an ideal therapeutic target for ES and with both direct (TK216; NCT02657005) and indirect (Trabectedin; NCT00070109, SP-2577; NCT03600649, and mithramycin; NCT01610570) inhibitors in various stages of clinical development.

EWS–FLI1 is an oncoprotein that both serves as a pioneer transcription factor and modulator of mRNA splicing (17,18); the link between these biologic processes has been described previously in other contexts (19,20). We have previously reported many novel protein interactions with EWS–FLI1, including BAF complex proteins (11). Further, a recently discovered connection between EWS–FLI1 and the BAF complex suggests chromatin regulation is a function of this oncogene (3). While interactions between EWS–FLI1 and multiple members of the BAF complex were reported, specific interactions and downstream effect on oncogenic function have not been described (3).

Here, we report a novel feed-forward mechanism describing how EWS–FLI1–BAF complex interactions contribute to ES oncogenesis and growth. This mechanism involves both the role of EWS–FLI1 in alternative splicing and interactions with the BAF complex that lead to ES maintenance. Our data indicates that EWS–FLI1 modulates the splicing of the BAF complex protein ARID1A to produce the ARID1A-L isoform in ES cells. We also show that EWS–FLI1 directly binds to the region of ARID1A-L protein that is maintained in the oncogenic isoform. We find that BRG1 ATPase activity is maintained in this complex, indicating that the EWS–FLI1-induced ARID1A isoform remains functional in the BAF complex. We identify ARID1A-L as an EWS–FLI1 regulated isoform required for ES growth that integrates EWS–FLI1 into a functional BAF complex. We also show a novel dependence of EWS–FLI1 on ARID1A-L and vice versa for protein stability in ES and human mesenchymal stem cells (hMSC). Our results provide new insights into the mechanism by which EWS–FLI1 induces alternative splicing of ARID1A to promote BAF complex activity in ES oncogenesis and growth.

MATERIALS AND METHODS

Cell lines and reagents

ES cell lines TC32, TC71, A4573 and STA-ET-7.2, were grown in RPMI with 10% FBS and 1% HEPES. SKES and RDES cells were grown in McCoy's 5A medium with 15% FBS. hMSC were obtained from Lonza and maintained in StemPro MSC SFM supplement CTS and GlutaMax for complete media (Gibco by Life technologies). hMSC cells were grown in the flasks coated with CELLstart (Gibco by Life technologies). hMSC cells were subcultured when cell confluency reaches 60–80%, cells were in mid-logarithmic phase of growth and cell viability is at least 90%. All cell lines were grown at 37°C in 5% CO_2 and ES cells were passaged every 2–4 days. Cell line integrity was confirmed by fingerprinting. Cell lines were tested mycoplasma negative *in domo*.

RNA-Seq isoform proportion analysis

Human RNA-Seq data was generated from primary patient tumor samples as previously described including tumor tissue from patients with ES which were collected with informed consent under a Hospital Sant Joan de Déu (Barcelona, Spain), Institutional Review Board (IRB)–approved protocol, dbGaP Study Accession: phs000804.v1.p1 (21). ES cell line TC32 RNA-seq 100 million paired-end reads were aligned to hg19 using TopHat2/Bowtie2 (22,23) with default parameters for genome-guided transcriptome reconstruction. Data was processed using Partek Genomic Suite software (PGS v6.6, Partek Inc., St. Louis, MO, USA) with the isoform expression quantification tools to calculate transcript-level expression of mRNA isoforms and determine alternative splicing ratios. PGS estimates the most likely relative expression levels of each isoform using an expectation/maximization algorithm similar to the one given in Xing *et al.* (24). Partek's algorithm differs from Xing *et al.*, in that it quantifies isoform expression levels across the whole genome at the same time rather than each gene separately, and normalizes by transcript length to account for the transcript fragmentation step in RNA-seq.

RNA-seq validation by qRT-PCR

RNA-seq findings (both expression and alternative splicing) were validated using qRT-PCR of genes that were the foci in this study. Validation of alternative splicing was performed using two different primer-pair combinations within the same transcript (RefSeq – release 61). One pair of primers targeted an area that can be used to differentiate between isoforms, and a second pair targeted an area common to all transcripts of the gene. The latter was used to normalize endogenous expression. Transcript expression validation used 18S as the across-sample reference gene relative to the primer set for the transcript of interest. Sequences of the primers used are provided in Supplementary Table S1.

Gene silencing with shRNA/exogenous expression

The expression of ARID1A was suppressed by using specific short hairpin (shRNA-TRCN0000059090) lentiviral plasmid which was obtained from the RNA Interference (RNAi) Consortium. The second shRNA targeting ARID1A 3′UTR region was obtained from OriGene Technologies Inc (Rockville, MD, USA). A non-silencing scramble shRNA construct was purchased from GE Healthcare Dharmacon, Inc (Lafayette, CO, USA). EWS–FLI1 shRNA was a generous gift from Christopher T. Denny (University of California, Los Angeles, USA). Lentiviral vector delivery of shRNA encoding ARID1A and EWS–FLI1 in different ES cell lines was performed as previously established (11). The ARID1A-L exogenous expression vector used in this study was a generous gift from Weidong Wang (NIA, NIH, Baltimore, MD 21224, USA) and few mutations were corrected by site-directed mutagenesis. We generated an ARID1A-S expression vector from ARID1A-L by deleting 650 nucleotides on the 5′ end of exon 18 by site-directed mutagenesis.

Oncomine data sets

The expression level of *ARID1A* in ES patient tumor samples were acquired from Oncomine Compendium of Expression Array data (https://www.oncomine.org/resource/login.html). The following seven data sets were used from published oncomine database—Khan (25), Ohali (26), Henderson (27), Baird (28), Ferreira (29), Schaefer (30) and Postel-Vinay (31).

Single molecule fluorescent *in situ* hybridization (FISH)

Probe sets of 40 small (20nt long) probes complementary to (i) unique 651 bp of exon 18 or (ii) to exon 20 were designed with 3′ amino modifications. The probes for each set were pooled and labeled en mass either with Alexa 594 (exon 18L) or cy5 (exon20). The labeled probes were purified using HPLC (32). TC32 cells were grown on glass coverslips and treated with no drug, or 30 μM of active drug or 30 μM of inactive drug for 16 h. The cells were fixed using 4% paraformaldehyde for 10 min, permeabilized with 70% ethanol at 4°C and hybridized overnight at 37°C with labeled probes. The coverslips were washed to remove unbound probes, stained with DAPI and mounted using deoxygenated media (32,33). The coverslips were imaged using an 100× oil objective in a Nikon TiE inverted fluorescence microscope equipped with cooled CCD camera. The *z*-stacks (3 μm with 0.2 μm apart) that were acquired using Metamorph software and the images were analyzed using custom written programs in MATLAB (Mathworks Inc.) (33,34).

Cell proliferation assay

Cell proliferation was monitored in the *xCELLigence* Real-Time Cell Analysis (RTCA; ACEA Biosciences Inc.) system using E-Plate 16 (ACEA Biosciences Inc.). A total of 100 μl of culture media was added into the plates and baseline measurements were taken. Cells were then seeded into the wells (10 000 cells/well in 100 μl) and allowed to grow up to 72 h. The electronic impedance of sensor electrodes is measured to allow monitoring and detection of physiological changes of the cells on the electrodes. Impedance was measured every 10 min during 72 h and is represented as cell index by the RTCA-integrated software of the *xCELLigence* system. Data were collected from three independent experiments. For two-way ANOVA calculation, readings up to 54 h were used.

Anchorage independent growth assay

The ability of transfected cell lines to divide in the absence of anchorage was determined by scoring the number of colonies formed in 0.4% agarose supplemented with 10% fetal bovine serum (with a 0.6% agarose underlay). Colonies larger than 0.12 mm were counted after 2–3 weeks in culture.

Caspase assay

Cells were lysed with lysis buffer containing 50 mM Tris (pH 7.5), 150 mM NaCl and 0.5% Triton X100. We added 10 μg of protein with caspase assay buffer (20 mM HEPES, 2 mM DTT and 10% glycerol) containing Ac-DEVD-AMC (MedChemExpress, USA) and incubated for 2–6 h. Reactions were excited at 360 nm, and emission at 460 nm was measured in a fluorescence plate reader.

EWS–FLI1 interacting partners identified through MS

We developed an unbiased screen for potential partners that interact with EWS–FLI1 by adding TC32 nuclear lysate to immobilized EWS–FLI1, followed by elution and PAGE as previously described (11).

Nuclear co-immunoprecipitation

Co-immunoprecipitation (CoIP) was performed to validate BAF complex protein–protein interaction. Active motif universal magnetic Co-IP kit (Active Motif) was used to make nuclear extract from TC32 cells according to the manufacturer's instructions. The protein G magnetic beads were used for Co-IP and the IP was performed on 300 μg samples using 2 μg of FLI1 polyclonal antibody and rabbit IgG (as a negative control). Western blots were performed individually using the following antibodies: FLI1 (Santa Cruz Biotechnology, C-19 SC365), ARID1A (Santa Cruz Biotechnology, sc-32761), ARID1B (Abcam, ab54761), SMARCA4 (BRG1; Santa Cruz Biotechnology, sc-17796), SMARCC2 (BAF170; Bethyl Laboratories, A301-039A), SMARCC1 (BAF155; Santa Cruz Biotechnology, sc-9746), SMARCB1 (SNF5; Bethyl Laboratories, A301-087A), and ACTL6A (BAF53A; Abcam ab131272). Detection was carried out using Millipore Immobilon Western Chemiluminescent HRP Substrate per the manufacturer's instructions (Millipore Corp.) using a Fujifilm LAS-3000 imaging system.

ARID1A fragments synthesis

ARID1A long and short isoform fragments were synthesized in cell-free protein expression system using TNT T7

quick for PCR DNA *in vitro* transcription/translation kit (Promega, Madison, WI, USA). The reactions were carried out as described in the manufacturer's protocol. In brief, ARID1A PCR fragments containing a T7 promoter were added to the TNT T7 PCR quick master mix and incubated for 60–90 min at 30°C. Using the Transcend systems protocol, biotinylated lysine residues are incorporated into nascent proteins during translation. The translated proteins were analyzed by SDS-PAGE and detected by western blot using an anti-ARID1A antibody (Santa Cruz Biotechnology, USA, sc-373784).

ELISA

ELISA was performed on Maxi-sorb 96-well plate (Nunc). Recombinant full length EWS–FLI1 were coated on the bottom surface of the ELISA plate overnight and followed by blocking with 4% BSA. *In vitro* transcribed/translated ARID1A fragments were used as the ligand to detect the specific protein interaction using anti-ARID1A HRP-conjugated monoclonal antibody (Santa Cruz Biotechnology, USA). For the reciprocal ELISA, *in vitro* transcribed/translated ARID1A fragments were coated on the bottom surface followed by blocking and recombinant full length EWS–FLI1 as ligand. Protein binding was detected using 3,3′,5,5′-tetramethylbenzidine (TMB) peroxidase substrate kit (Bio-Rad, CA, USA) per the manufacturer's instructions. The color development was monitored by absorbance reading at 450 nm using a microplate reader. Negative controls for each ELISA included the absence of EWS–FLI1, absence of second recombinant protein, and blocking alone.

ATPase activity assay

EWS–FLI1-associated complexes were isolated using nuclear co-IP as described previously and followed by measuring the ATPase activity using ADP-Glo Max assay from Promega, according to the manufacturer's instructions. EWS–FLI1-associated complexes were eluted and pH was adjusted by adding 12.5 μl of 1× reaction buffer containing 50 mM Tris–HCl (pH 7.5), 20 mM $MgCl_2$, 2 mM DTT and 2 mM ATP for the ATPase assay. The assay was performed in two steps. (i) After the ATPase reaction, ADP-Glo reagent was added to terminate the enzymatic reaction and to deplete the remaining ATP and (ii) the ADP-Glo Max detection reagent was added to convert ADP to ATP and newly synthesized ATP was measured using a luciferase/luciferin reaction. The fluorescence generated correlates with the amount of ADP generated in the ATPase assay, which indicates the ATPase activity.

Mutation analysis

The first study cohort consisted of 37 ES cases consecutively submitted to Caris Life Sciences (Phoenix, AZ, USA) for molecular profiling using multiple technologies, including next generation sequencing (NGS), immunohistochemistry (IHC), and *in situ* hybridization (ISH). All patient data was provided under IRB exemption 4, such that only final analysis of anonymous data was included in this manuscript. Next generation sequencing on a targeted panel of 592 genes (targeted exome sequencing) was performed on formalin-fixed, paraffin embedded tumor tissue. In brief, genomic DNA was isolated from microdissected tumor tissue. Library preparation and target enrichment for all exons of the 592 gene panel was performed using a customized SureSelect XT target enrichment kit (Agilent Technologies; Santa Clara, CA, USA). Sequencing of the prepared library was performed using the Illumina NextSeq platform to a mean read depth of $>$ 500X. Variants with $>$5% frequency with 99% confidence in a binomial model were reported. Single nucleotide polymorphisms (SNPs) categorized as common SNPs in dbSNP database (https://www.ncbi.nlm.nih.gov/projects/SNP/) were filtered and excluded from analyses. Demographic data for patients and mutations found in BAF complex genes (ARID1A, ARID2, BCL11A, BCL11B, BCL7A, PBRM1, SMARCA4, SMARCB1, SMARCE1, SS18, SS18L1) were analyzed and reported in Supplementary Table S2. The second study cohort consisted of 24 ES cases as previously described (21), Demographic data for patients and mutations found in BAF complex genes were analyzed and reported in Supplementary Table S2.

Statistical analysis

We performed statistical analyses with GraphPad Prism. All statistical analysis was performed with Student's *t*-test/two-way ANOVA; data is expressed as means and SD; asterisks denote $P < 0.05$.

RESULTS

ARID1A is alternatively spliced in ES by EWS–FLI1 to maintain *ARID1A-L*

We analyzed novel protein interactions with EWS–FLI1 by affinity mass spectrometry and discovered eight BAF complex components among EWS–FLI1 interacting proteins (Table 1). Given the substantial number of BAF complex proteins interacting with EWS–FLI1, we focused further on involvement of the BAF complex in ES oncogenesis. There were no recurrent BAF complex mutations in 61 patient samples analyzed (Supplementary Table S2). We previously investigated exon array data and found that *ARID1A* showed a marked change in relative isoform ratios when EWS–FLI1 expression was reduced (11). *ARID1A* expression was confirmed in seven published Oncomine datasets generated from ES tumors (Supplementary Figure S1A-D). *ARID1A* mRNA was more highly expressed in ES relative to many other cancer types, thus our experiments focused on contributions of *ARID1A* to ES oncogenesis (Supplementary Figure S1E and F).

RefSeq annotation indicates that there are two major *ARID1A* isoforms generated by alternative 3′ splice site usage, long (*ARID1A-L*) and short (*ARID1A-S*), where *ARID1A-S* lacks 651 nucleotides on the 5′ portion of exon 18 (Figure 1A). We show that TC32 *ARID1A* mRNA undergoes a significant isoform switch when EWS–FLI1 expression is reduced by shRNA in this ES cell line (Figure 1B). A panel of 21 ES tumor RNAseq samples confirmed

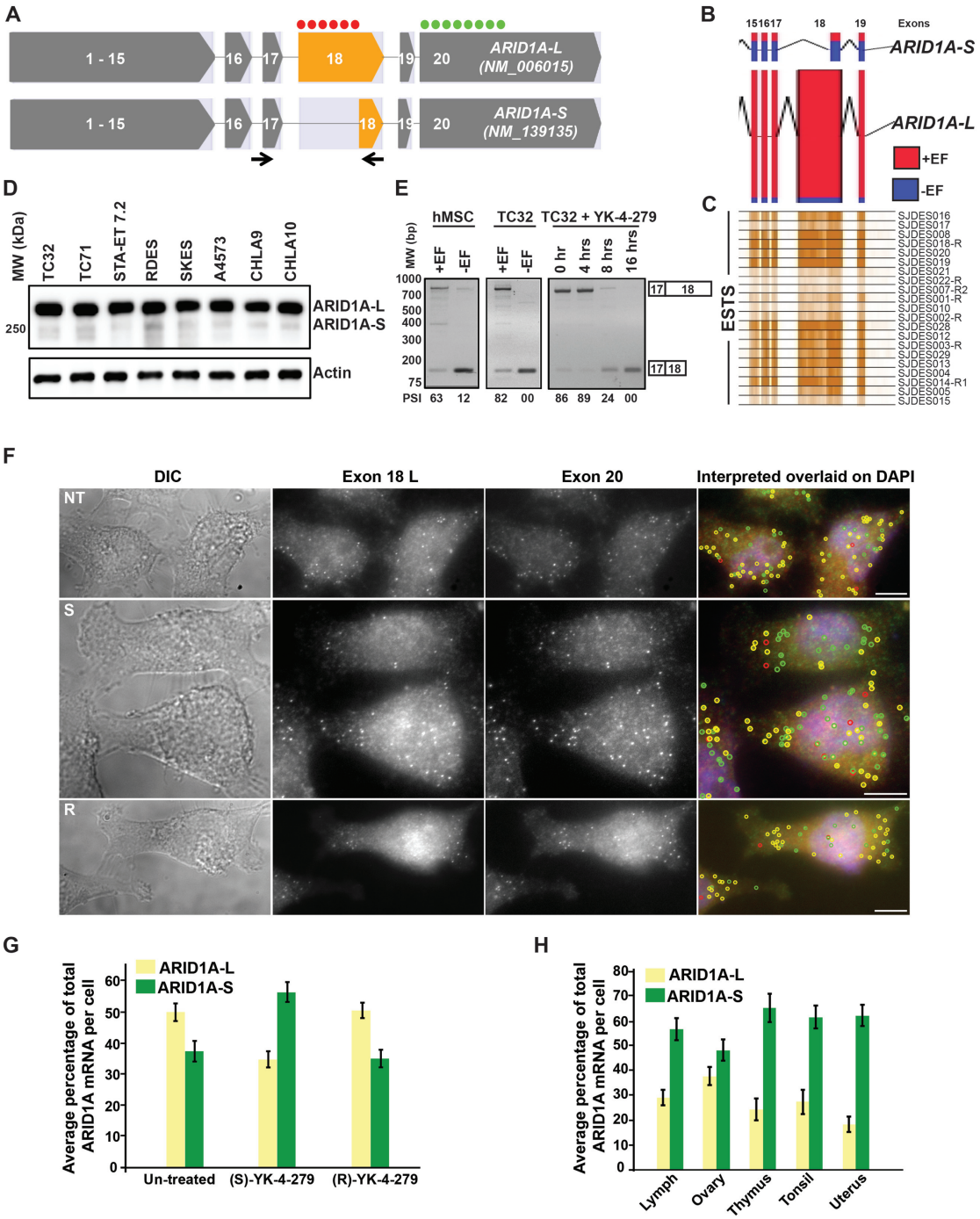

Figure 1. EWS–FLI1 alters the splicing of ARID1A. (**A**) *ARID1A* is annotated in two major isoforms long (*ARID1A-L*) and short (*ARID1A-S*); *ARID1A-S* lacks 651 nucleotides in the N-terminal portion of exon 18. Primer locations are indicated with arrows and colored dots indicate the FISH probe locations. (**B**) Analysis of TC32 wild-type expressing EWS–FLI1 (+EF, red bar) and EWS–FLI1 shRNA reduced cells (-EF, blue bar) shown from exon 15 to 19 of ARID1A. Isoform proportion analysis using Partek Genomics Suite shows the height of the bar is directly proportional to isoform expression level. (**C**) Heat maps of 21 ES patient tumor *ARID1A* sequences from exon 15 to exon 19 are shown. The intensity of the brown bar indicates expression of exon. (**D**) ARID1A protein isoform expression in 8 ES cell lines. Actin was used as a loading control. (**E**) human mesenchymal stem cells (hMSC) with exogenous EWS–FLI1 (+EF) or wild-type (–EF) and TC32 cells (+EF) or shRNA EWS–FLI1 (–EF) show semi-quantitative RT-PCR products using specific primers from (A, arrows) and calculation the percent of spliced-in (PSI). TC32 treated with 3 μM YK-4-279 over time was similarly analyzed. (**F**) Representative images for single molecule fluorescent *in situ* RNA hybridization with *ARID1A* probes in TC32 cells with (A) for no treatment (NT), (*S*)-YK-4-279 or (*R*)-YK-4-279. The first panel is differential interference contrast (DIC), the second and third panels are a merge of z stacks from raw images from the independent probes 18L and 20, respectively; the fourth panel is an overlay of DAPI staining with merge of both channel images. The signal from the exon 18L probe is labeled red and exon 20 is labeled green. *ARID1A-L* RNA signal is identified as yellow (hybridized to both red and green) and *ARID1A-S* is identified as green. Scale bar is 5μm. (**G**) Quantification obtained after MATLAB analysis of images of at least 100 cells for each condition. (**H**) Similar data for 5 different normal tissues (images in Supplementary Figure. S3). The error bars (G and H) represent 95% confidence intervals from 100 cells.

Table 1. LC–MS/MS identified potential protein participants in EWS–FLI1–BAF complexes[a]

Proteins	Accession number	Protein description	Percentage sequence covered	Peptides matched
ACL6A	O96019	Actin-like protein 6A	39.2	9
ARI1A	O14497	AT-rich interactive domain-containing protein 1A	21.2	12
ARI1B	Q8NFD5	AT-rich interactive domain-containing protein 1B	6.5	2
SMCA4	P51532	Transcription activator BRG1	24.5	11
SMCA5	O60264	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5	11.3	7
SMRC1	Q92922	SWI/SNF complex subunit SMARCC1	31.4	11
SMRC2	Q8TAQ2	SWI/SNF complex subunit SMARCC2	33.3	12
SNF5	Q12824	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily B member 1	35.8	12

[a]Recombinant EWS–FLI1 was immobilized on an affinity column followed by the addition of ES nuclear lysate, elution, PAGE separation and identification of bound proteins using liquid chromatography–mass spectrometry (MS). For details see, Selvanathan, PNAS 2015.

high relative expression of *ARID1A-L* (Figure 1C). Expression of ARID1A-L protein was also confirmed in eight ES cell lines (Figure 1D) with validation of isoform sizes detected by the antibody using HEK293 which have little endogenous protein (Supplementary Figure S1G).

Mesenchymal stem cells (MSC) are the putative cell of origin of ES (35,36) and expression of EWS–FLI1 in human MSC (hMSC) led to a 2.5-fold increase in total *ARID1A* expression (Supplementary Figure S1F). Analysis of mRNA by RT-PCR (using isoform-specific primers) and quantification of isoforms by gel densitometry confirmed similar *ARID1A* isoform levels between hMSC expressing EWS–FLI1 and wild-type ES TC32 cells (percent spliced-in (PSI) 63% vs. 82% respectively, Figure 1E). Wild-type hMSC (without EWS–FLI1) had similar low levels of *ARID1A-L* expression compared to TC32 cells with shRNA reduction of EWS–FLI1 (11) with a PSI of 12% and 0%, respectively (Figure 1E). The *ARID1A* isoform ratios with reduction of EWS–FLI1 were confirmed in four of five additional ES cell lines (Supplementary Figure S2A and B).

A series of experiments validated that EWS–FLI1 was responsible for the isoform pattern change. Reduction of *EWSR1* mRNA expression by an shRNA targeting a region that is not part of the fusion message (Supplementary Figure S2C) did not elicit any change in *ARID1A* isoform expression ratio (Supplementary Figure S2D). We confirmed that EWS–FLI1 influences *ARID1A* isoform expression through protein partner interactions by using the small molecule inhibitor YK-4-279 to inhibit EWS–FLI1 interactions (16,37). Treatment of ES cells with YK-4-279 showed a time-dependent switch in the *ARID1A* from long to short isoform (Figure 1E). As a control for non-specific isoform switching we included doxorubicin, an agent frequently used to treat ES that does have a transcriptional signature change similar to reduction of EWS–FLI1 (38), but does not affect EWS–FLI1-directed splicing (11). We found that doxorubicin treatment has no effect on *ARID1A* isoform expression ratio (Supplementary Figure S2E). Treatment of non-ES cells with YK-4-279 also does not influence the *ARID1A* isoform expression ratio (Supplementary Figure S2F and G).

Since RT-PCR is a semi-quantitative technique, we confirmed our data using single molecule resolution *in-situ* RNA hybridization probes targeting the isoform-specific region of exon 18 as well as exon 20 to directly evaluate this isoform ratio shift in cells (Figure 1A). We found that the observed isoform ratio shift from *ARID1A-L* to *ARID1A-S* was present with treatment using the active enantiomer ((S)-YK-4-279) but not the inactive enantiomer ((R)-YK-4-279) (Figure 1F and G). All non-tumor tissues tested had a higher ratio of *ARID1A-S* relative to *ARID1A-L* (Figure 1H and Supplementary Figure S3). Taken together, our data show a preferential expression of *ARID1A-S* in hMSC and a panel of normal tissues, but *ARID1A-L* in ES cells regulated by EWS–FLI1, indicating that the ARID1A-L isoform may play a role in EWS–FLI1-driven tumorgenesis.

We also considered that mRNA turnover via mechanisms such as nonsense-mediated mRNA decay (NMD) may contribute to the regulation of *ARID1A* isoforms by EWS–FLI1. Although the alternative splice isoforms do not encode predicted premature termination codons, NMD susceptibility may arise from isoform-specific usage of upstream open reading frames or ribosomal frameshifting. Cycloheximide (CHX) (39) treatment blocks NMD (40), such that treatment of cells increases the abundance of transcripts that are normally degraded. However, when ES cells were treated with CHX, we measured only a small rise in ARID1A mRNA levels while positive controls indicated NMD was effectively blocked (Supplementary Figure S2H). Reduction of EWS–FLI1 demonstrated the expected changes in transcript levels (Supplementary Figure S2I). Given the small increase in ARID1A, we evaluated *ARID1A* relative isoform ratio after CHX treatment and found no significant isoform ratio differences (Supplementary Figure S2J).

Isoform specific function of ARID1A on ES cell growth

To determine the function of individual ARID1A isoforms, TC32 cells were depleted of both ARID1A isoforms using shRNA followed by exogenous expression of full-length *ARID1A-L* or *ARID1A-S* (Figure 2A and B). Using continuous growth monitoring in the xCELLigence Sys-

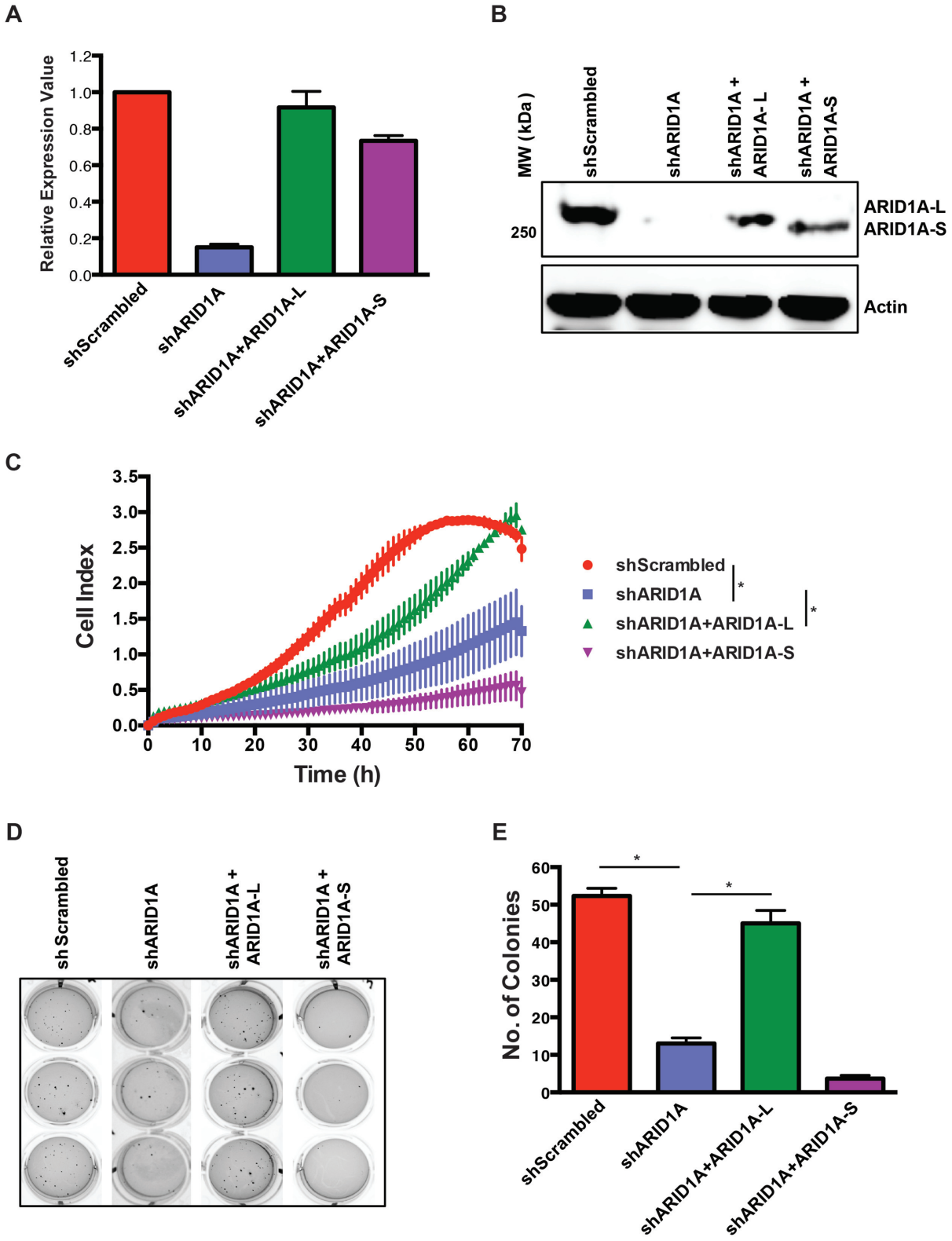

Figure 2. ARID1A-L is necessary for ES oncogenesis. Total *ARID1A* reduced (sh*ARID1A*) in TC32 cells followed by exogenously expressed *ARID1A-L* or *ARID1A-S*. (**A**) *ARID1A* expression using primers to evaluate mRNA expression across all isoforms. (**B**) ARID1A protein levels after either shScrambled control or shARID1A followed by individual isoform exogenous re-expression (ARID1A-L or ARID1A-S). (**C**) Cell proliferation assay using *xCELLigence* system using electric impedance as a measure of ES cancer cell proliferation in various conditions as in (A, B), and data were collected from three independent experiments. *Indicates significantly different comparisons ($P < 0.05$, Two-way ANOVA). (**D**) Anchorage-independent growth assays showing colony formations for same conditions as in (A, B). (**E**) Enumeration of anchorage-independent colonies (three independent experiments, *$P < 0.05$).

tem, we found that shRNA reduction of both *ARID1A* isoforms significantly reduced cell proliferation (Figure 2C). shARID1A followed by expression of *ARID1A-L* (shARID1A + *ARID1A-L*) had a longer lag time but eventually matched the growth of shScramble control cells, whereas shARID1A followed by expression of *ARID1A-S* (shARID1A + *ARID1A-S*) repressed growth (Figure 2C). We tested a second ES cell line (A4573) for ARID1A cell growth effects following shRNA reduction and re-expression of individual isoforms (Supplementary Figure S4A and B). Similar to TC32, the data show that ARID1A-L but not ARID1A-S supports ES growth in A4573 (Supplementary Figure S4C). Anchorage independent growth, a frequent measure of transformation, was restored to levels similar to wild-type cells after re-expression of ARID1A-L following ARID1A knockdown while expression of ARID1A-S alone or following knockdown of ARID1A inhibited colony formation of TC32 and A4573 ES cell lines in soft agar (Figure 2D and E, and Supplementary Figure S4D and E). As a further specificity control, an additional shRNA targeting the ARID1A 3′-UTR showed similar isoform-specific results for both cell lines (Supplementary Figure S5).

To validate the role of ARID1A-L as a cooperative partner for oncogenic transformation, human MSC were transfected with ARID1A-L or doxycycline-inducible EWS–FLI1 expression vectors individually or together (Figure 3A–C). It is noteworthy that EWS–FLI1 protein levels were higher in hMSC that had constitutive exogenous ARID1A-L (Figure 3C). Induction of EWS–FLI1 expression in hMSC with exogenous ARID1A-L expression resulted in 1.6-fold more growth than in cells with either gene expressed alone in an adherent cell assay (Figure 3D). Using these same cells, with permutations for necessary controls, we show a 3.5-fold increase in colony number as well as a significant increase in colony size when ARID1A-L was present to stabilize the EWS–FLI1 expression (Figure 3E and F). Based on our data, it is evident that ARID1A-L is necessary for anchorage independent ES cell proliferation, implying that the two proteins (ARID1A-L and EWS–FLI1) work in concert driving a key property of oncogenesis.

ARID1A-L and EWS–FLI1 demonstrate co-dependence for maintenance of protein levels

Given the apparent requirement for ARID1A-L for ES cell growth and transformation, we evaluated ARID1A protein levels after reduction of EWS–FLI1 expression. We analyzed ARID1A protein levels to determine the degree of correlation with *ARID1A-L* and *ARID1A-S* mRNA expression. As expected, expression of ARID1A-L protein was significantly decreased upon EWS–FLI1 knockdown, which is consistent with reduction of its mRNA levels (Figure 1E); however, unexpectedly, there was no concomitant increase in ARID1A-S protein despite a proportional increase of its mRNA isoform (Figure 4A). We further observed that exogenous expression of *ARID1A-S* mRNA in two ES cell lines only led to minimal expression of ARID1A-S protein; however, surprisingly ARID1A-L protein was almost absent (Supplemental Figure S6A and B).

To investigate this perplexing finding, we looked at the cellular effects of exogenous *ARID1A-S* expression. Two ES cell lines demonstrated activation of caspase-3 leading to cellular apoptosis by 18 hours after transfection with the *ARID1A-S* expression vector (Supplementary Figure S6C and D). This rapid apoptosis was confirmed with cells expressing ARID1A-S in the absence of ARID1A-L having no growth (Supplementary Figure S6E and F). These data suggest that ARID1A-S is not tolerated by ES cells when ARID1A-L is absent with induction of apoptosis leading to cell death.

We further evaluated the effects of EWS–FLI1 suppression on the ARID1A-L protein isoform. We created TC32 and A4573 cell line models that stably expressed *ARID1A*-L without introns from an exogenous EWS–FLI1-independent promoter, and modulated EWS–FLI1 expression by shRNA (Figure 4B). We then reduced EWS–FLI1 mRNA expression in the *ARID1A-L* expressing cells as well as the parental cell line. On day 6, *EWS–FLI1* mRNA levels showed significant reduction following shEWS–FLI1 (Supplementary Figure S6G, blue). As expected, we found little to no ARID1A-L mRNA following EWS–FLI1 reduction in control parental cells (Figure 4C, blue). However, in shEWS–FLI1 reduced cells that had exogenous *ARID1A-L* expression, 50% of *ARID1A-L* RNA was retained, consistent with the loss of only endogenous ARID1A mRNA expression (Figure 4C, brown). However, we observed a near complete loss of ARID1A-L protein with EWS–FLI1 reduction not only in the parental but also the ARID1A-L overexpression cell lines (Figure 4D and E). The presence of modest amounts of ARID1A-L protein, only when exogenously expressed, sufficed to partially rescue ES cell growth following the reduction of EWS–FLI1 in both ES cell lines, TC32 and A4573 (Figure 4F and G; brown), while shEWS–FLI1 alone markedly impaired growth (Figure 4F and G; blue). This indicates that in the absence of EWS–FLI1, ARID1A-L can support modest ES cell growth.

We then evaluated whether reduction of EWS–FLI1 expression reduces ARID1A protein levels through protein stability modulation treating shEWS–FLI1 ES cells with the proteasome inhibitor MG-132. EWS–FLI1 repression by shEWS–FLI1 led to reduction of ARID1A-L protein; however, MG-132 increased ARID1A-L protein levels to 46% and 27% compared with shScrambled control in both A4573 and TC32 ES cell lines, respectively (Figure 5A and B). Given that EWS–FLI1 contributed to the continued protein stability of ARID1A-L, we wanted to determine whether the reciprocal effect was also true: if ARID1A-L contributed to the continued protein stability of EWS–FLI1. When *ARID1A* expression was reduced by shRNA, we found a concomitant decrease in EWS–FLI1 protein levels (Figure 5C and D) but not in mRNA levels (Figure 5E and F). In a follow-up experiment, when ARID1A expression was reduced, EWS–FLI1 levels increased in both A4573 and TC32 cells, by 17% and 23% respectively, with MG-132 treatment (Figure 5G and H). Taken together, these data support a reciprocal protein stability interaction between EWS–FLI1 and ARID1A-L. This suggests a potential feedback mechanism that functions to maintain expression of ARID1A-L in cells expressing EWS–FLI1.

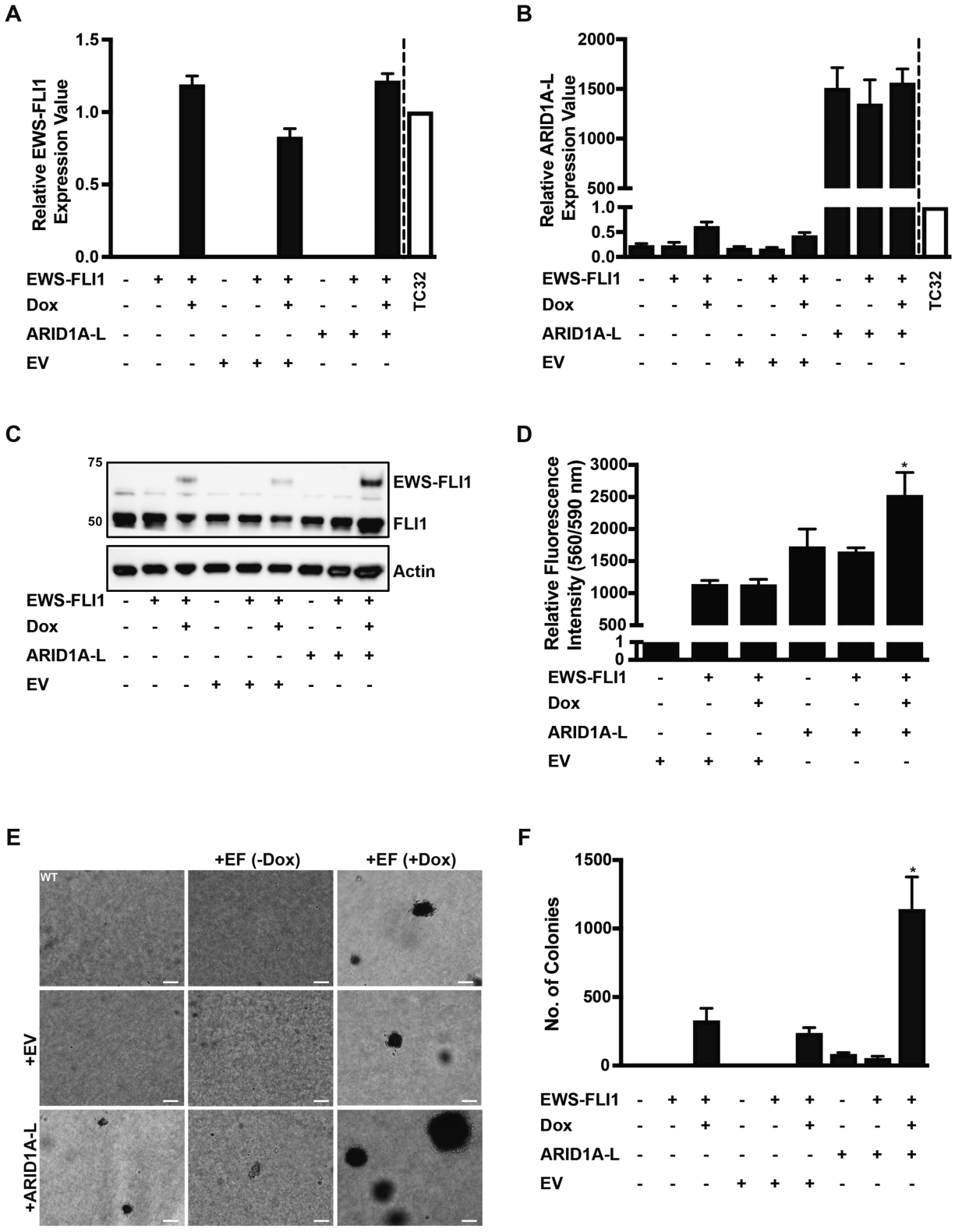

Figure 3. ARID1A-L co-operates with EWS–FLI1 expression to support oncogenesis in hMSC. hMSC were transfected with individual ARID1A-L or dox-inducible EWS–FLI1 expression vectors in addition to the combination of the two. (**A**) qRT-PCR showing EWS–FLI1 expression. (**B**) qRT-PCR showing ARID1A-L expression. (**C**) Immunoblot shows EWS–FLI1 protein expression. Actin is a loading control. (**D**) Growth of transfected hMSC (cell proliferation) after four days followed by Cell-titer Blue assay and data were collected from three independent experiments. *Indicates significantly different comparisons with rest of the groups ($P < 0.05$, two-way ANOVA). (**E**) Anchorage-independent growth assays showing colony formations for same conditions as in (A). Scale bar is 100 μm. (**F**) Quantification of anchorage-independent colonies (three independent experiments, *$P < 0.05$).

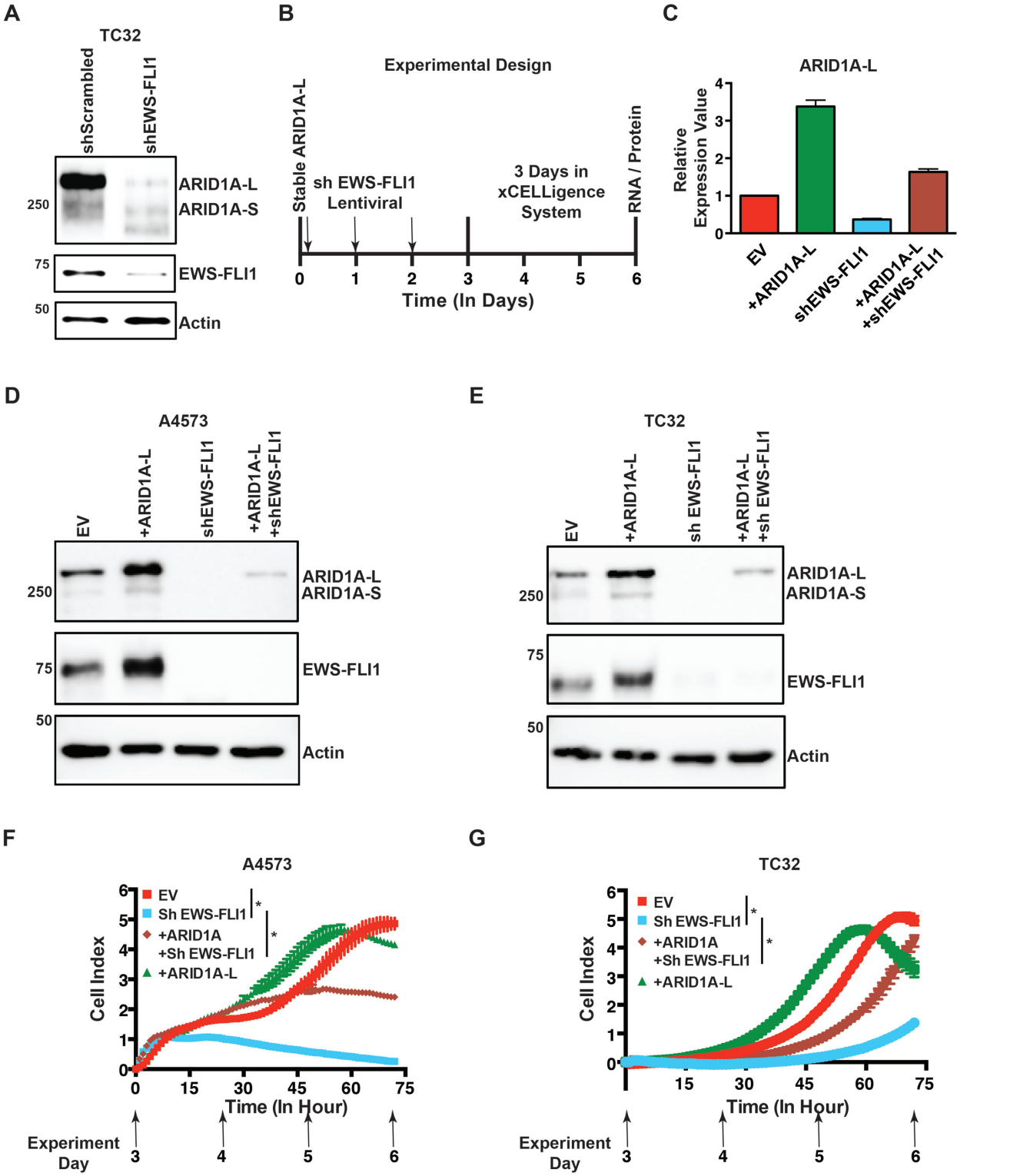

Figure 4. ARID1A-L protein can rescue ES cell growth following reduction of EWS–FLI1. (**A**) Immunoblot for ARID1A in shScrambled and EWS–FLI1 reduced (shEWS–FLI1) ES TC32 cells. (**B**) Experimental design for ARID1A-L expression followed by shEWS–FLI1. On day 0, all cells were stably expressing ARID1A-L or were selected with an empty vector (EV). Starting on day 1 and for three consecutive days, cells were infected with shEWS–FLI1 lentivirus (arrows). The growth assay using xCelligence began on day 3, while other cells remained in standard culture. On day 6, mRNA and proteins were extracted from the remaining cells for analysis. (**C**) qRT-PCR analysis of cell populations derived from Experiment (B) with specific primers in ARID1A exons 17 and the 5′ region of exon 18 to measure only the long isoform expression. (**D**, **E**) Immunoblots with ARID1A, EWS–FLI1, and actin antibodies for protein extracted from Experiment (B) on day 6 in A4573 and TC32. (**F**, **G**) xCelligence assay from cells prepared in Experiment (B) starting from day 3 in A4573 and TC32. Red is empty vector (EV). Green is stable expression of ARID1A-L. Blue is EV + shEWS–FLI1. Brown is ARID1A-L followed by shEWS–FLI1. *Indicates significantly different comparisons ($P < 0.05$, two-way ANOVA)

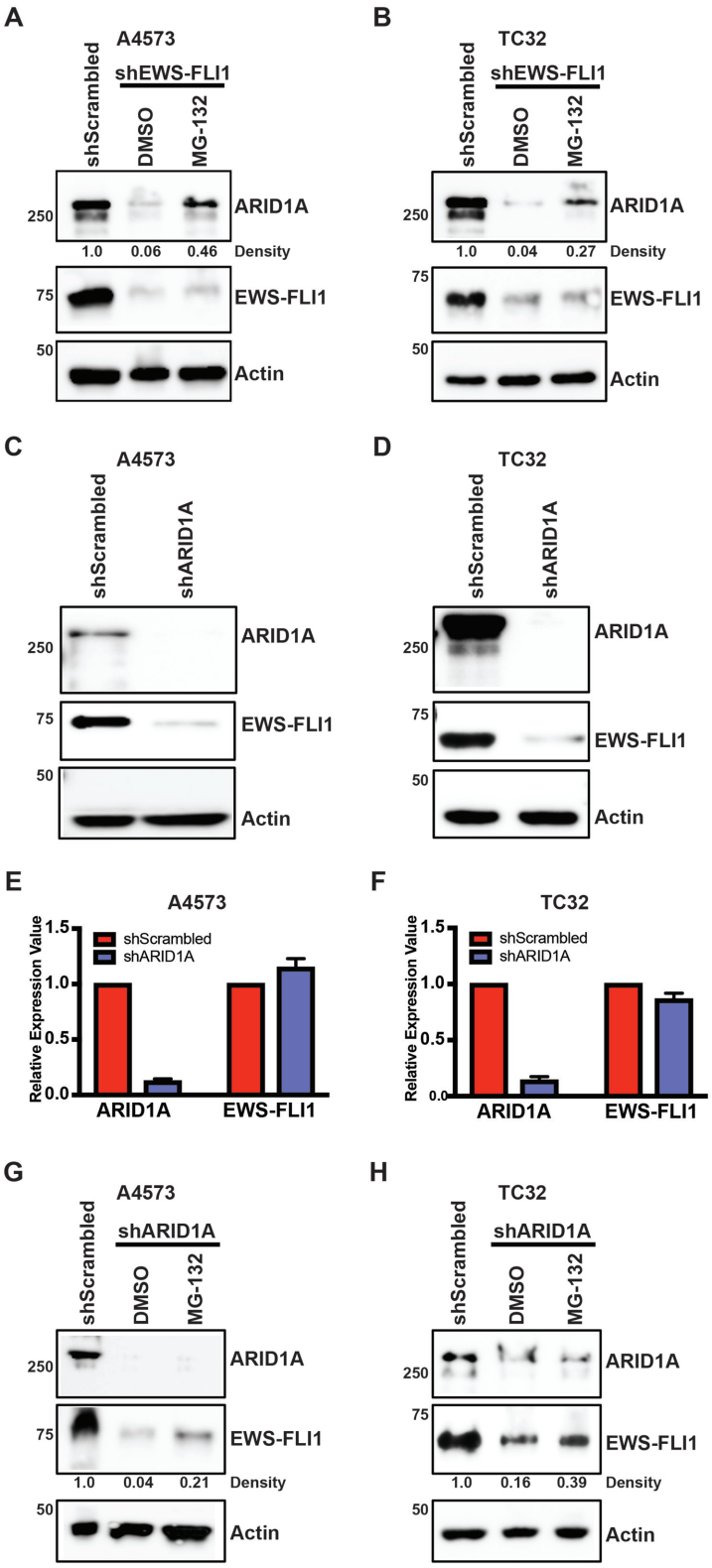

Figure 5. ARID1A-L is required for EWS–FLI1 protein stability. (**A, B**) shEWS–FLI1 reduced protein in DMSO control or 20 μM MG-132 for 8 h in A4573 and TC32 followed by evaluation of protein levels. Numbers indicate relative densitometry for ARID1A bands comparing shEF to shScrambled. (**C**, **D**) shARID1A or shScrambled followed by protein extraction and immunoblot for ARID1A, EWS–FLI1, and actin in A4573 and TC32. (**E**, **F**) qRT-PCR analysis of *ARID1A* and *EWS–FLI1* expression in A4573 and TC32. (**G**, **H**) shARID1A reduced protein in DMSO control or 20 μM MG-132 for 8 h in A4573 and TC32 followed by evaluation of protein levels. Numbers indicate relative densitometry for EWS–FLI1 bands comparing shARID1A to shScrambled. All protein blots show actin controls.

Because we established that ARID1A-L is a critical isoform for ES and some presence appears necessary for EWS–FLI1, we sought to determine the nature of EWS–FLI1 interaction with the BAF complex.

EWS–FLI1 binds directly to the BAF complex through ARID1A-L

We evaluated BAF complex protein interactions with EWS–FLI1 identified by our previous screen (Table 1). Six of eight BAF complex proteins were confirmed as part of an EWS–FLI1 complex by co-immunoprecipitation (CoIP) in TC32 cells (Figure 6A). These results demonstrate that EWS–FLI1 – BAF complexes include ARID1A (only the long isoform, ARID1A-L), BRG1 (SMARCA4), BAF170 (SMARCC2), BAF155 (SMARCC1), BAF47 (SMARCB1) and BAF53, but not ARID1B. ARID1A and ARID1B are 60% homologous and are mutually exclusive within the BAF complex (41). Reciprocal CoIP with ARID1A was not possible due to inadequate antibodies; however, reciprocal BRG1 CoIP was successful and confirmed BAF complex association with EWS–FLI1 (Supplementary Figure S7A). EWSR1 protein, using a C-terminal antibody, also pulled down small amounts of the BAF proteins (Supplementary Figure S7B). However, the same EWSR1 antibody did not co-IP BAF complex proteins in Jurkat cells, a leukemia cell line that lacks EWS–FLI1 (Supplementary Figure S7C) indicating the EWS–FLI1 fusion, not wild-type EWSR1, is required for substantial BAF complex interaction.

YK-4-279 treatment (3 μM) of ES cells for 16 hours followed by EWS–FLI1 co-IP revealed that YK-4-279 prevented precipitation of ARID1A-L, as well as other BAF complex proteins, without reduction of EWS–FLI1 protein expression levels in two independent experiments (Supplementary Figure S7D). Further, ES cells treated with YK-4-279 showed that BRG1 co-IP of EWS–FLI1 was reduced; however, the BAF complex itself remained intact in ES and Jurkat leukemia cells (Supplementary Figure S7E and F, respectively). These combined co-IP data show that YK-4-279 prevents EWS–FLI1 interactions with the BAF complex similar to previously reported disruption of protein interactions (11).

We next sought to determine whether ARID1A is the BAF subunit that directly binds to EWS–FLI1. We prepared recombinant protein fragments of both *ARID1A-L* and *ARID1A-S*, designated ARID1A-#FL and ARID1A-#FS, to determine whether EWS–FLI1 directly binds to either isoform (Figure 6B, top). ES nuclear lysate containing EWS–FLI1 was incubated with ARID1A recombinant protein fragments. We found that ARID1A-#FL but not ARID1A-#FS binds directly to EWS–FLI1 (Figure 6B, bottom). To confirm that the interaction is directly through ARID1A-L, an ELISA was performed in which we found that recombinant EWS–FLI1 only bound ARID1A-#FL but not ARID1A-#FS (Figure 6C). Reciprocal ELISA confirmed binding between EWS–FLI1 and ARID1A-L (Supplementary Figure S8A). Full-length EWSR1 did not bind to either of the ARID1A fragments (Supplementary Figure S8B). Direct ELISA binding of ARID1A-#FL was significantly reduced by YK-4-279 in a dose-dependent fashion (Supplementary Figure S8C). Together, this data suggests that EWS–FLI1 interacts with the BAF complex through ARID1A-L. However, our data does not exclude the possibility that multiple BAF complex proteins may also interact directly with EWS–FLI1.

BRG1 ATPase activity is linked to EWS–FLI1 through ARID1A

ARID1A is thought to determine BAF targets and provide the scaffolding through which other proteins interact, including the catalytic component BRG1 (SMARCA4). BRG1 is an ATPase and either BRG1 or BRM are necessary for BAF complex function (42). We sought to determine whether BRG1 ATPase activity is retained in BAF complexes interacting with EWS–FLI1. We demonstrated with co-IP that BRG1 was in a complex with EWS–FLI1 in ES cells (Supplementary Figure S7A). To determine if ATPase activity is retained in the EWS–FLI1 complex, we assayed ATPase activity of EWS–FLI1 precipitated complexes. We also assayed ATPase activity after shARID1A. Complexes that co-precipitated with EWS–FLI1 retained ATPase activity while suppression of ARID1A with shRNA led to a 50% reduction in EWS–FLI1-associated ATPase activity (Figure 6D). Analysis of the precipitated proteins indicated that BRG1 interactions with EWS–FLI1 were lost when ARID1A expression was reduced by shARID1A (Figure 6E). Total ATPase activity was similar in nuclear lysates with or without shARID1A (Figure 6D). Residual ATPase activity associated with EWS–FLI1 following ARID1A reduction could be attributed in part to RHA (43), which associated with EWS–FLI1 independent of shARID1A (Figure 6E). Our data suggests that EWS–FLI1 interacts with functional BAF complexes via ARID1A and these complexes include active BRG1.

ARID1A-L compensates for EWS–FLI1 loss at specific target genes

EWS–FLI1 modulates mRNA expression of many target genes, producing a specific transcriptional signature pattern that has been associated with EWS–FLI1 transformation. We evaluated the effect of ARID1A reduction on known EWS–FLI1 transcriptional targets using shRNA of ARID1A followed by qRT-PCR. Several genes whose expression is increased by EWS–FLI1 (*VEGFA, EZH2, UPP1, TERT, NROB1, GLI1, PTPL1* and *ID2*) were decreased with shARID1A in both TC32 and A4573, while expression of *TGFβR2* and *LOX*, normally suppressed by EWS–FLI1, were increased (Figure 7A, Supplementary Figure S9A, respectively). We noticed a discordance with *IGFBP3* (normally suppressed by EWS–FLI1) in that neither cell line activated expression in the absence of ARID1A. Since exogenous expression of ARID1A-L partially rescued shEWS–FLI1-induced proliferation reduction in ES cells (Figure 4F and G), we evaluated expression of EWS–FLI1 regulated gene targets after exogenous expression of ARID1A-L. In each graph, shEWS–FLI1 shows the increased expression of a normally repressed target (*TGFβR2*) with decreased expression of normally enhanced targets (*VEGFA*, *EZH2*, *UPP1*, *TERT*, *NROB1*, *GLI1*, *PTPL1* and *ID2*) (Figure 7B–K, Supplementary Figure S9B–K). Ectopically expressing ARID1A-L reversed

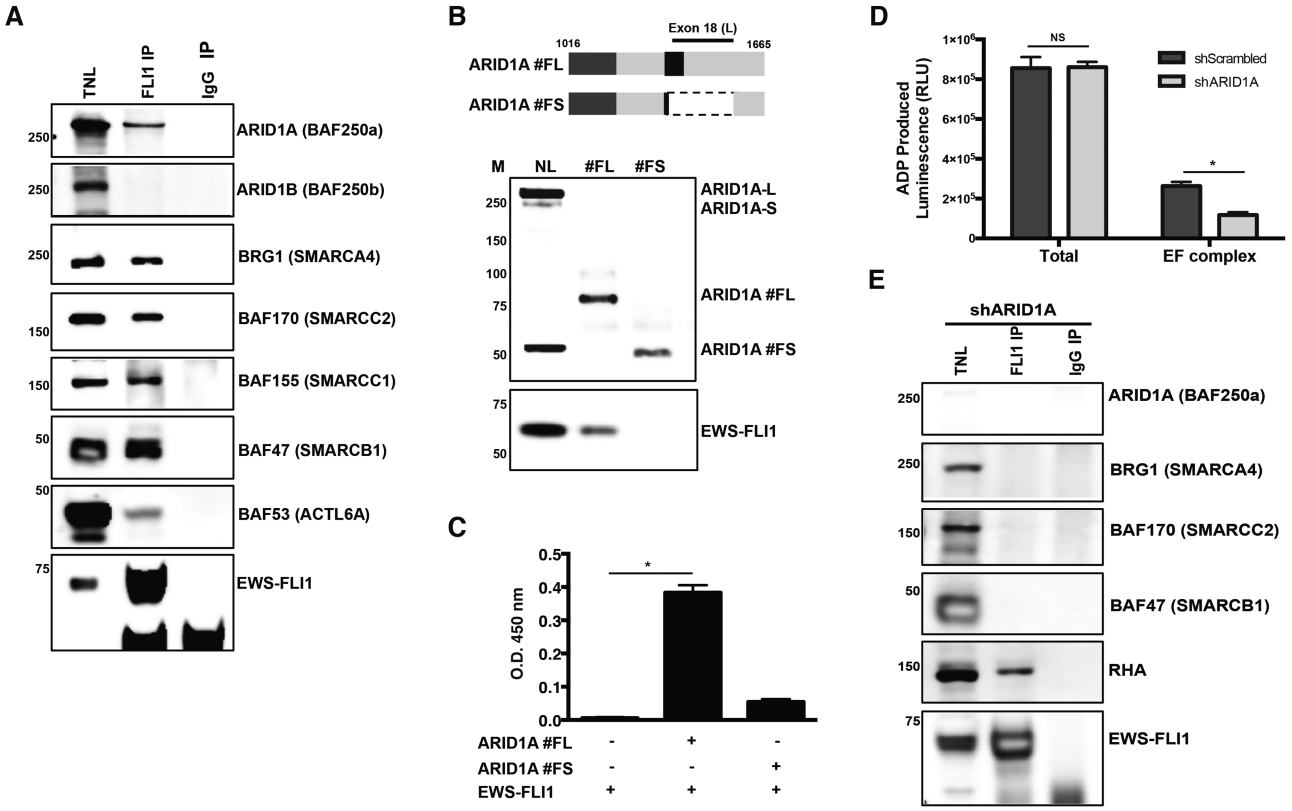

Figure 6. EWS–FLI1 directly interacts with BAF complexes, including association with ATPase activity of BRG1, through ARID1A-L. (**A**) FLI1 antibody co-immunoprecipitation (co-IP) of EWS–FLI1 and antibodies used for the western blots to evaluated BAF proteins are indicated on the right. 10% of total nuclear lysate (TNL) was used as input (lane 1), FLI1 antibody precipitates show the presence of proteins as a part of the complex (lane 2). For the negative control, rabbit polyclonal IgG antibody (Ab) was used to show specificity (lane 3). (**B**) (top) Construct fragments of ARID1A indicated by amino acid number are synthesized by *in vitro* transcribed/translated (IVTT) system. The dashed line in #FS shows the missing portion of exon 18. (bottom) In vitro transcribed/ translated ARID1A fragments containing biotin-lysine are used as bait to bind Streptavidin magnetic beads. EWS–FLI1 from ES cell lysate is combined with IVTT ARID1A protein fragments and resolved by PAGE followed by blotting with ARID1A and FLI1 antibody. (**C**) ELISA using recombinant EWS–FLI1 and ARID1A fragments from B. Error bars show mean ± SD from three individual experiments (*$P < 0.05$). (**D**) The ATPase activity of proteins immunoprecipitated with EWS–FLI1 was measured using ADP-Glo Max assay in TC32 shScrambled (dark grey) and shRNA reduced ARID1A (light grey). *Indicates significantly different comparisons ($P < 0.05$), NS indicates non-significant. (**E**) TC32 cells with shRNA reduction of ARID1A followed by co-IP using FLI1 antibody. As a positive control, we show RHA remains in the EWS–FLI1 complex.

the effect of EWS–FLI1 knockdown on EWS–FLI1 regulated genes in TC32 and A4573 cell lines (Figure 7B–K, Supplementary Figure S9B-K). Consistent with the absent effect on *IGFBP3* by shARID1A, no rescue of expression is seen in TC32 (Figure 7C) and very minimal rescue in A4573 (Supplementary Figure S9C). These results implicate ARID1A-L, and its role in the BAF complex, as a potential factor in EWS–FLI1 access to DNA regulatory sequences.

DISCUSSION

BAF complex activity is thought to be critical to cancer cell epigenetic reprogramming, yet the specific contributions of this complex are not yet fully understood. In the present study, we show that the oncogenic fusion protein EWS–FLI1 associates with the BAF complex through direct interaction with the long isoform of ARID1A, which is created by EWS–FLI1-mediated alternative ARID1A splicing. We also find that the observed isoform expression is consistent in ES patient tumor samples. ARID1A-L contains the full ARID1A exon 18 that codes for the region in which EWS–FLI1 interacts with ARID1A. Our data indicates that in addition to altered isoform ratios, there is mutual protein stabilization between EWS–FLI1 and ARID1A-L. We further show that expression of ARID1A-L increases proliferation and supports anchorage-independent growth while ARID1A-S decreases cell viability. A subset of EWS–FLI1-regulated gene targets are co-regulated by ARID1A-L, and expression can be re-established when ARID1A-L is expressed in cells depleted of endogenous EWS–FLI1. This suggests chromatin regulation as an emerging function of EWS–FLI1. However, we found one EWS–FLI1 regulated gene, *IGFBP3*, which did not follow the pattern suggesting an ARID1A independent mechanism. These data strongly suggests a role for EWS–FLI1 in modulating alternative splicing of ARID1A to alter the BAF complex and thereby maintaining ES growth and oncogenesis. This discussion will comment on the novel direct interactions of EWS–FLI1 with the BAF complex and the role of isoform specificity in ES oncogenesis.

Understanding the interactive proteome with EWS–FLI1 has evolved with new technologies and capacity for

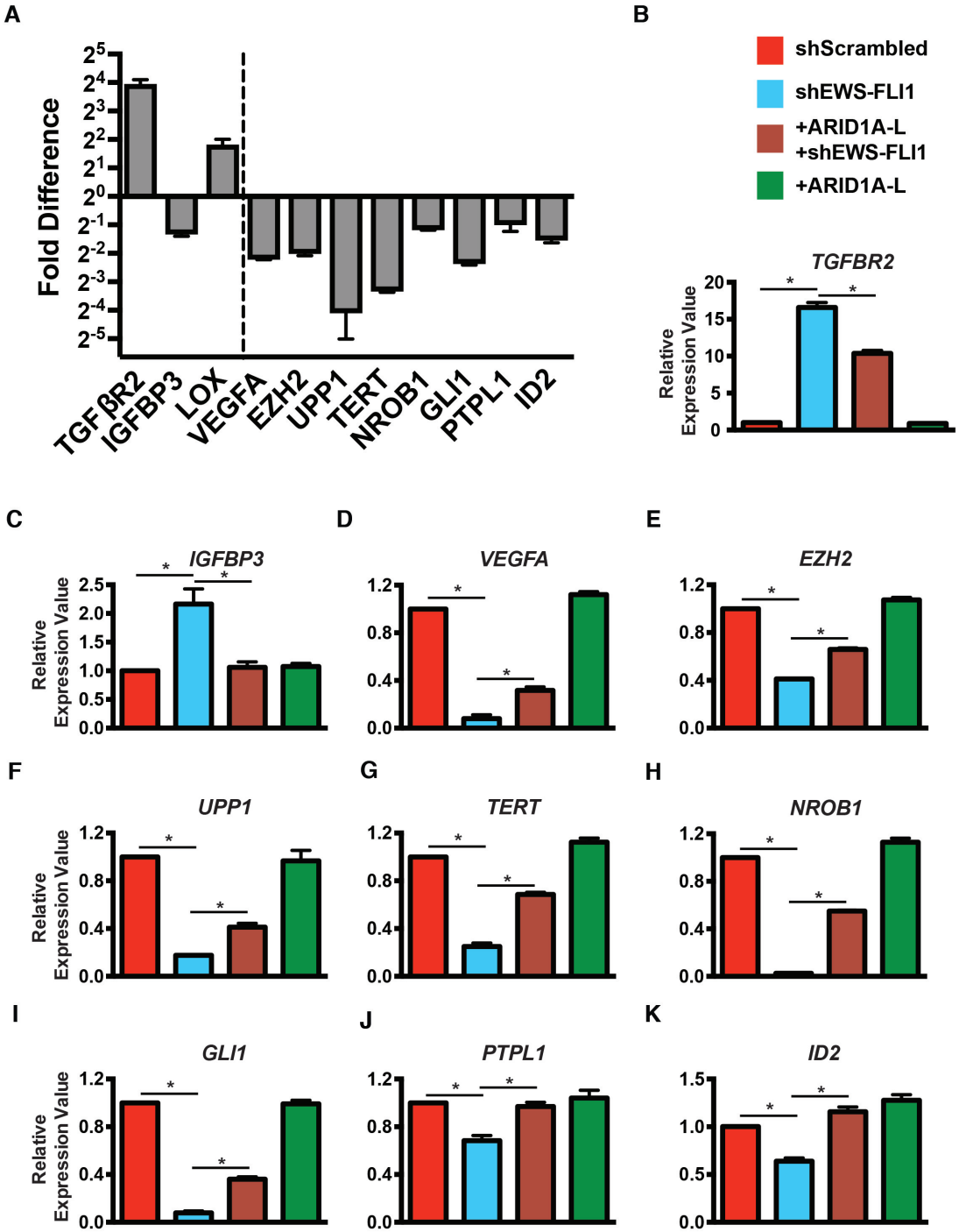

Figure 7. ARID1A-L compensates for EWS–FLI1 loss in regulating canonical target genes. (**A**) qRT-PCR showing the relative expression of EWS–FLI1 canonical target genes in shRNA reduced ARID1A in TC32 cells. (**B–K**) qRT-PCR relative expression of EWS–FLI1 canonical target genes that are normally suppressed in wild-type ES: *TGFBR2, IGFBP3*; or activated in wild-type ES: *VEGFA, EZH2, UPP1, TERT, NROB1, GLI1, PTPL1* and *ID2*. Red is empty vector (EV). Green is stable exogenous expression of ARID1A-L, Blue is EV + shEWS–FLI1, Brown is stable ARID1A-L followed by shEWS–FLI1. *Indicates significantly different comparisons ($P < 0.05$, *t*-test).

validation. The earliest studies used yeast two-hybrid systems (44) while advances in mass spectrometry (MS), especially following HPLC resolution of proteins, demonstrated that significant numbers of proteins interact with EWS–FLI1 (11). Our prior investigations show that splice variants, just as with transcripts, demonstrate some variance across cell lines that do not correlate with EWS–FLI1 fusion type (11). This current study focuses on proteins of the BAF complex identified in our earlier work, but here we validate these interactions through multiple co-immunoprecipitation experiments. We focused on ARID1A given our interest in its isoform specificity.

This interaction with ARID1A appears to be through unique properties of EWS–FLI1 as EWSR1 fails to co-IP the BAF complex in translocation-negative Jurkat cells. Proteins that interact with the BAF complex likely vary by cell type, including EWSR1 (45). While our experiments have focused on ARID1A due to the observed mRNA isoform expression changes mediated by EWS–FLI1 expression, it is also possible that other BAF-complex proteins may directly interact with EWS–FLI1. We identified a region including exon 18 that is spliced into *ARID1A* in the presence of EWS–FLI1 in ES. This region of exon 18 was reciprocally found to bind to EWS–FLI1 in two assays. Unfortunately, there is no clear function of the ARID1A-L specific protein sequence from exon 18, however, some reports suggested that this is a binding region for HIC1 (46). ARID1A-S lacks this region.

Very few cancers or normal differentiated tissues express high levels of ARID1A-S, suggesting that this isoform may be toxic through non-EWS–FLI1-mediated mechanisms as well. Since we show that hMSC are more tolerant of ARID1A-S expression, we hypothesize that less differentiated cells utilize the BAF complex differently. One possibility for future investigation is the role of ARID1A-L/ARID1A-S in reprogramming BAF complexes based on the state of differentiation of the cell. Our data suggests over-expression of ARID1A-S may be a form of dominant negative by altering the stoichiometry of non-EWS–FLI1 binding proteins from the BAF complex. In ES, expression of ARID1A-S was acutely toxic, activating caspase-3 shortly after expression. The apoptotic activation of proteases is a possible explanation for the lack of measurable proteins following ARID1A-S expression. A deeper understanding of ARID1A-S is required and may reveal additional insights into BAF complex formation and function as it pertains to cellular differentiation and oncogenesis.

In ES cells, wild-type EWSR1 does form a complex with BAF proteins; however this may be indirect through heterotypic EWSR1/EWS–FLI1 interactions, as suggested by prior reports and the lack of interactions in fusion-negative cells (47). Interactions between EWS–FLI1 and chromatin writers and erasers have been hypothesized to occur through intrinsically disordered regions (48). Of note, low-complexity peptide sequences from EWSR1 are necessary for EWS–FLI1 interactions with BAF complexes (3). These unique properties could be due to liquid-liquid phase separation of EWSR1 as clear canonical binding domains are not identifiable in either EWS–FLI1 nor ARID1A in the region of interaction (48). Further dissection of these emerging protein interaction properties will be required to better understand the functional interactions.

Expression of mRNA isoform variants are an acknowledged mode of cellular regulation, and in some cases lead to oncogenesis (49). Our prior work demonstrated a genome-wide discovery of EWS–FLI1-regulated alternative splicing; however, the importance of alternative isoforms generated by EWS–FLI1 was unanswered (11). In our prior work, we showed that EWS–FLI1 directly interacted with multiple splicing factors (11), but in this work we chose to study the functional effects of one of these isoforms. We speculate that a U1-binding site in *ARID1A* exon 18 is bound by a splice site regulator that is modulated by the presence of EWS–FLI1. Additional investigations will identify specific regulators of splice site utilization.

Isoform-specificity in proteins has been previously recognized in oncogenesis; however, partnering of an oncogene with an isoform-specific protein is less common (50). To our knowledge, this is an early description of an isoform splice variant that is both maintained by an oncoprotein and required for the stability of that oncoprotein, as we have observed with EWS–FLI1 and ARID1A-L. Our model describes a requirement for ARID1A-L, which is preferentially expressed by EWS–FLI1, through interactions with the spliceosome (Figure 8). In the first panel, we show that in the presence of EWS–FLI1, ARID1A-L is expressed and binds directly to EWS–FLI1. In the middle panel, reduction of EWS–FLI1 or treatment with the EWS–FLI1 inhibitor YK-4-279 in ES leads to an isoform shift from *ARID1A-L* to *ARID1A-S*. In either case, ES cells die and we show one possibility is that ARID1A-S causes apoptosis ES. In the third panel, the reduction of ARID1A, both isoforms, leads to EWS–FLI1 degradation and cell death; the loss of ARID1A-L to stabilize EWS–FLI1 is one explanation that is demonstrated. This mutual stabilization is supported through experiments with MG-132 showing persistent protein levels when the proteasome is inhibited. In fact, through a number of experiments, including the use of protease inhibitors, EWS–FLI1 appears to be stabilized by binding to ARID1A-L. There is data to support the reciprocal condition in which EWS–FLI1 may stabilize ARID1A in ES as well. In ES cells, ARIDA1A-S leads to apoptosis, wherein the proteases of apoptosis may lead to EWS–FLI1 degradation. As suggested above, future work with ARID1A-S, while challenging due to its toxicity, will be important to resolve these questions.

While this model applies to ES, we found that in hMSC, expression of ARID1A-L was required to stabilize the EWS–FLI1 expression to optimize anchorage-independent growth. The presence of high levels of ARID1A may be an artifact of our expression model, however if ARID1A-L is required for binding to EWS–FLI1, the isoform switch may not happen soon enough in the model system. This phenomenon requires further study.

Our data supports YK-4-279 dissociating EWS–FLI1 from ARID1A-L as a novel inhibition of a protein-protein interaction by this small molecule. However, when ES cells were treated with YK-4-279, followed by IP of BRG1, ARID1A was present in the BAF complex despite the dissociation of EWS–FLI1 (Supplementary Figure S7E). Since the BAF complex is stable in 5M urea (51) it is possible

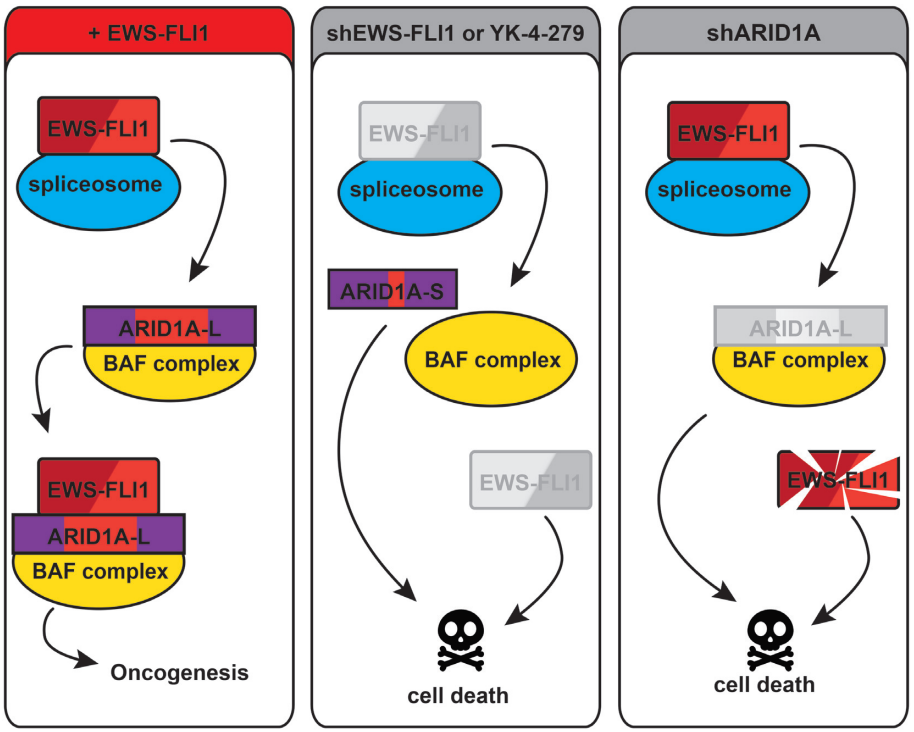

Figure 8. Feed-forward oncogenic cycle of EWS–FLI1 and ARID1A-L. In the first panel, we show that in the presence of EWS–FLI1, ARID1A-L is expressed and binds directly to EWS–FLI1 to support ES oncogenesis. In the middle panel, reduction of EWS–FLI1 or treatment with the EWS–FLI1 inhibitor YK-4-279 in ES leads to an isoform shift from *ARID1A-L* to *ARID1A-S*. In either case, ES cells die and we show one possibility is that ARID1A-S cause apoptosis ES. In the third panel, the reduction of ARID1A, both isoforms, leads to EWS–FLI1 degradation and cell death; the loss of ARID1A-L to stabilize EWS–FLI1 is one explanation that is demonstrated.

that the cohesion of the BAF complex, and temporal protein stability explain this finding. However, the impact of the protein interactions leading to BAF complex stability requires further resolution and could also be due to other post-translational modifications or protein interactions not yet defined.

Our results provide a novel connection between a tumorigenic fusion protein and protein isoform variants created through alternative splicing. The ARID1A-L isoform described here provides a point of interaction between EWS–FLI1 and the BAF complex. This requirement of BAF complex interactions for EWS–FLI1 to function as an oncogene is a critical piece of evidence in understanding the interactome of this fusion protein. We further provide evidence implicating a critical protein interaction partner within the BAF complex, ARID1A-L. The alternative splicing mechanism through which this specific interaction partner is produced requires further investigation. The diversity of EWS–FLI1 functions continue to grow, further supporting EWS–FLI1 and its protein interactions as rational therapeutic targets. An understanding of EWS–FLI1 protein interaction network at the molecular and biophysical level are key to advancing both knowledge of tumorigenesis and identifying ES tumor vulnerabilities.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Alexandra Logerfo for manuscript assistance and Cigall Kadoch for helpful discussions on assay design.
Author contributions: S.P.S. designed and conducted the experiments, performed data analysis and wrote the manuscript. A.R.G. maintained the cell lines. T.M.B. contributed human mutation data and analysis. J.R.H. provided the NMD substrates. M.S. and M.B. performed FISH experiments and data analysis. G.T.G., E.M.T., B.C. and K.S. contributed human RNA-seq data and analysis. A.Ü. and H.K. provided scientific input and edited the manuscript. J.A.T. designed the study and edited the manuscript. All authors reviewed and contributed to editing the manuscript.

FUNDING

St. Baldrick's Foundation that provided key pilot funding; Children's Cancer Foundation (Baltimore, MD), Hyundai Hope on Wheels; Alan B. Slifka Foundation, Go4theGoal; Burroughs Wellcome Clinical Scientist Award in Translational Research (to J.A.T.); NIH [RC4CA156509 to J.A.T., R01CA133662 to J.A.T., R01CA138212 to J.A.T.; DP5 OD012160 to M.B., R01CA204915 to K.S.]; BBSR, GESR, TCSR, and PMSR through LCCC CCSG Grant [P30 CA051008-16 to L.W., P.I.]; E.M.T. was supported by fellowships of the Austrian Science Fund [FWF, Lise Meitner Fellowship M1448-B13; Elise Richter Fellowship V506-

B28]. Funding for open access charge: Children's Cancer Foundation (Baltimore MD).

Conflict of interest statement. USPTO awarded for YK-4-279 to Georgetown University, inventors include A.Ü. and J.T. Georgetown University executed license with Onctern Therapeutics, Inc., in which J.T. and A.Ü. are founding share-holders and J.T. is a scientific consultant. K.S. participates in the DFCI/Novartis Drug Discovery Program which includes grant support and previously included consulting and has consulted for Rigel Pharmaceuticals on a topic unrelated to this manuscript.

REFERENCES

1. Narayanan,R., Pirouz,M., Kerimoglu,C., Pham,L., Wagener,R.J., Kiszka,K.A., Rosenbusch,J., Seong,R.H., Kessel,M., Fischer,A. *et al.* (2015) Loss of BAF (mSWI/SNF) complexes causes global transcriptional and chromatin state changes in forebrain development. *Cell Rep.*, **13**, 1842–1854.
2. Hodges,C., Kirkland,J.G. and Crabtree,G.R. (2016) The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harbor Perspect. Med.*, **6**, a026930.
3. Boulay,G., Sandoval,G.J., Riggi,N., Iyer,S., Buisson,R., Naigles,B., Awad,M.E., Rengarajan,S., Volorio,A., McBride,M.J. *et al.* (2017) Cancer-Specific retargeting of BAF complexes by a Prion-like domain. *Cell*, **171**, 163–178.
4. Kim,S., Denny,C.T. and Wisdom,R. (2006) Cooperative DNA binding with AP-1 proteins is required for transformation by EWS-Ets fusion proteins. *Mol. Cell. Biol.*, **26**, 2467–2478.
5. Vierbuchen,T., Ling,E., Cowley,C.J., Couch,C.H., Wang,X., Harmin,D.A., Roberts,C.W.M. and Greenberg,M.E. (2017) AP-1 transcription factors and the BAF complex mediate signal-dependent enhancer selection. *Mol. Cell*, **68**, 1067–1082.
6. McAndrew,M.J., Gjidoda,A., Tagore,M., Miksanek,T. and Floer,M. (2016) Chromatin remodeler recruitment during macrophage differentiation facilitates transcription factor binding to enhancers in mature cells. *J. Biol. Chem.*, **291**, 18058–18071.
7. Roberts,C.W., Galusha,S.A., McMenamin,M.E., Fletcher,C.D. and Orkin,S.H. (2000) Haploinsufficiency of Snf5 (integrase interactor 1) predisposes to malignant rhabdoid tumors in mice. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 13796–13800.
8. Kadoch,C. and Crabtree,G.R. (2013) Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. *Cell*, **153**, 71–85.
9. Arnaud,O., Le Loarer,F. and Tirode,F. (2018) BAFfling pathologies: Alterations of BAF complexes in cancer. *Cancer Lett.*, **419**, 266–279.
10. Mathur,R., Alver,B.H., San Roman,A.K., Wilson,B.G., Wang,X., Agoston,A.T., Park,P.J., Shivdasani,R.A. and Roberts,C.W. (2017) ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat. Genet.*, **49**, 296–302.
11. Selvanathan,S.P., Graham,G.T., Erkizan,H.V., Dirksen,U., Natarajan,T.G., Dakic,A., Yu,S., Liu,X., Paulsen,M.T., Ljungman,M.E. *et al.* (2015) Oncogenic fusion protein EWS–FLI1 is a network hub that regulates alternative splicing. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E1307–E1316.
12. Chansky,H.A., Hu,M., Hickstein,D.D. and Yang,L. (2001) Oncogenic TLS/ERG and EWS/Fli-1 fusion proteins inhibit RNA splicing mediated by YB-1 protein. *Cancer Res.*, **61**, 3586–3590.
13. Knoop,L.L. and Baker,S.J. (2000) The splicing factor U1C represses EWS/FLI-mediated transactivation. *J. Biol. Chem.*, **275**, 24865–24871.
14. Kovar,H. (2014) Blocking the road, stopping the engine or killing the driver? Advances in targeting EWS/FLI-1 fusion in Ewing sarcoma as novel therapy. *Expert Opin. Ther. Targets*, **18**, 1315–1328.
15. Zollner,S.K., Selvanathan,S.P., Graham,G.T., Commins,R.M.T., Hong,S.H., Moseley,E., Parks,S., Haladyna,J.N., Erkizan,H.V., Dirksen,U. *et al.* (2017) Inhibition of the oncogenic fusion protein EWS–FLI1 causes G2-M cell cycle arrest and enhanced vincristine sensitivity in Ewing's sarcoma. *Science Signaling*, **10**, eaam8429.
16. Erkizan,H.V., Kong,Y., Merchant,M., Schlottmann,S., Barber-Rotenberg,J.S., Yuan,L., Abaan,O.D., Chou,T.H., Dakshanamurthy,S., Brown,M.L. *et al.* (2009) A small molecule blocking oncogenic protein EWS–FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma. *Nat. Med.*, **15**, 750–756.
17. Tomazou,E.M., Sheffield,N.C., Schmidl,C., Schuster,M., Schonegger,A., Datlinger,P., Kubicek,S., Bock,C. and Kovar,H. (2015) Epigenome mapping reveals distinct modes of gene regulation and widespread enhancer reprogramming by the oncogenic fusion protein EWS–FLI1. *Cell Rep.*, **10**, 1082–1095.
18. Riggi,N., Knoechel,B., Gillespie,S.M., Rheinbay,E., Boulay,G., Suva,M.L., Rossetti,N.E., Boonseng,W.E., Oksuz,O., Cook,E.B. *et al.* (2014) EWS–FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell*, **26**, 668–681.
19. Brown,S.J., Stoilov,P. and Xing,Y. (2012) Chromatin and epigenetic regulation of pre-mRNA processing. *Hum. Mol. Genet.*, **21**, R90–R96.
20. Luco,R.F., Allo,M., Schor,I.E., Kornblihtt,A.R. and Misteli,T. (2011) Epigenetics in alternative pre-mRNA splicing. *Cell*, **144**, 16–26.
21. Crompton,B.D., Stewart,C., Taylor-Weiner,A., Alexe,G., Kurek,K.C., Calicchio,M.L., Kiezun,A., Carter,S.L., Shukla,S.A., Mehta,S.S. *et al.* (2014) The genomic landscape of pediatric ewing sarcoma. *Cancer Discov.*, **4**, 1326–1341.
22. Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.
23. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
24. Xing,Y., Yu,T., Wu,Y.N., Roy,M., Kim,J. and Lee,C. (2006) An expectation-maximization algorithm for probabilistic reconstructions of full-length isoforms from splice graphs. *Nucleic Acids Res.*, **34**, 3150–3160.
25. Khan,J., Wei,J.S., Ringner,M., Saal,L.H., Ladanyi,M., Westermann,F., Berthold,F., Schwab,M., Antonescu,C.R., Peterson,C. *et al.* (2001) Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat. Med.*, **7**, 673–679.
26. Ohali,A., Avigad,S., Zaizov,R., Ophir,R., Horn-Saban,S., Cohen,I.J., Meller,I., Kollender,Y., Issakov,J. and Yaniv,I. (2004) Prediction of high risk Ewing's sarcoma by gene expression profiling. *Oncogene*, **23**, 8997–9006.
27. Henderson,S.R., Guiliano,D., Presneau,N., McLean,S., Frow,R., Vujovic,S., Anderson,J., Sebire,N., Whelan,J., Athanasou,N. *et al.* (2005) A molecular map of mesenchymal tumors. *Genome Biol.*, **6**, R76.
28. Baird,K., Davis,S., Antonescu,C.R., Harper,U.L., Walker,R.L., Chen,Y., Glatfelter,A.A., Duray,P.H. and Meltzer,P.S. (2005) Gene expression profiling of human sarcomas: insights into sarcoma biology. *Cancer Res.*, **65**, 9226–9235.
29. Ferreira,B.I., Alonso,J., Carrillo,J., Acquadro,F., Largo,C., Suela,J., Teixeira,M.R., Cerveira,N., Molares,A., Gomez-Lopez,G. *et al.* (2008) Array CGH and gene-expression profiling reveals distinct genomic instability patterns associated with DNA repair and cell-cycle checkpoint pathways in Ewing's sarcoma. *Oncogene*, **27**, 2084–2090.
30. Schaefer,K.L., Eisenacher,M., Braun,Y., Brachwitz,K., Wai,D.H., Dirksen,U., Lanvers-Kaminsky,C., Juergens,H., Herrero,D., Stegmaier,S. *et al.* (2008) Microarray analysis of Ewing's sarcoma family of tumours reveals characteristic gene expression signatures associated with metastasis and resistance to chemotherapy. *Eur. J. Cancer*, **44**, 699–709.
31. Postel-Vinay,S., Veron,A.S., Tirode,F., Pierron,G., Reynaud,S., Kovar,H., Oberlin,O., Lapouble,E., Ballet,S., Lucchesi,C. *et al.* (2012) Common variants near TARDBP and EGR2 are associated with susceptibility to Ewing sarcoma. *Nat. Genet.*, **44**, 323–327.
32. Batish,M., Raj,A. and Tyagi,S. (2011) Single molecule imaging of RNA in situ. *Methods Mol. Biol.*, **714**, 3–13.
33. Markey,F.B., Ruezinsky,W., Tyagi,S. and Batish,M. (2014) Fusion FISH imaging: single-molecule detection of gene fusion transcripts in situ. *PLoS One*, **9**, e93488.

34. Batish,M., van den Bogaard,P., Kramer,F.R. and Tyagi,S. (2012) Neuronal mRNAs travel singly into dendrites. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 4645–4650.
35. Riggi,N., Cironi,L., Provero,P., Suva,M.L., Kaloulis,K., Garcia-Echeverria,C., Hoffmann,F., Trumpp,A. and Stamenkovic,I. (2005) Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. *Cancer Res.*, **65**, 11459–11468.
36. Tirode,F., Laud-Duval,K., Prieur,A., Delorme,B., Charbord,P. and Delattre,O. (2007) Mesenchymal stem cell features of Ewing tumors. *Cancer Cell*, **11**, 421–429.
37. Barber-Rotenberg,J.S., Selvanathan,S.P., Kong,Y., Erkizan,H.V., Snyder,T.M., Hong,S.P., Kobs,C.L., South,N.L., Summer,S., Monroe,P.J. *et al.* (2012) Single enantiomer of YK-4-279 demonstrates specificity in targeting the oncogene EWS–FLI1. *Oncotarget*, **3**, 172–182.
38. Stegmaier,K., Wong,J.S., Ross,K.N., Chow,K.T., Peck,D., Wright,R.D., Lessnick,S.L., Kung,A.L. and Golub,T.R. (2007) Signature-based small molecule screening identifies cytosine arabinoside as an EWS/FLI modulator in Ewing sarcoma. *PLoS Med.*, **4**, e122.
39. DelBove,J., Kuwahara,Y., Mora-Blanco,E.L., Godfrey,V., Funkhouser,W.K., Fletcher,C.D., Van Dyke,T., Roberts,C.W. and Weissman,B.E. (2009) Inactivation of SNF5 cooperates with p53 loss to accelerate tumor formation in Snf5(+/−);p53(+/−) mice. *Mol. Carcinog.*, **48**, 1139–1148.
40. Hogg,J.R. and Goff,S.P. (2010) Upf1 senses 3′UTR length to potentiate mRNA decay. *Cell*, **143**, 379–389.
41. Helming,K.C., Wang,X., Wilson,B.G., Vazquez,F., Haswell,J.R., Manchester,H.E., Kim,Y., Kryukov,G.V., Ghandi,M., Aguirre,A.J. *et al.* (2014) ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat. Med.*, **20**, 251–254.
42. Stanton,B.Z., Hodges,C., Calarco,J.P., Braun,S.M., Ku,W.L., Kadoch,C., Zhao,K. and Crabtree,G.R. (2017) Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. *Nat. Genet.*, **49**, 282–288.
43. Lee,C.G. and Hurwitz,J. (1993) Human RNA helicase A is homologous to the maleless protein of Drosophila. *J. Biol. Chem.*, **268**, 16822–16830.
44. Petermann,R., Mossier,B.M., Aryee,D.N., Khazak,V., Golemis,E.A. and Kovar,H. (1998) Oncogenic EWS-Fli1 interacts with hsRPB7, a subunit of human RNA polymerase II. *Oncogene*, **17**, 603–610.
45. Linden,M., Thomsen,C., Grundevik,P., Jonasson,E., Andersson,D., Runnberg,R., Dolatabadi,S., Vannas,C., Luna Santamariotaa,M., Fagman,H. *et al.* (2019) FET family fusion oncoproteins target the SWI/SNF chromatin remodeling complex. *EMBO Rep.*, **20**, e45766.
46. Van Rechem,C., Boulay,G. and Leprince,D. (2009) HIC1 interacts with a specific subunit of SWI/SNF complexes, ARID1A/BAF250A. *Biochem. Biophys. Res. Commun.*, **385**, 586–590.
47. Spahn,L., Siligan,C., Bachmaier,R., Schmid,J.A., Aryee,D.N. and Kovar,H. (2003) Homotypic and heterotypic interactions of EWS, FLI1 and their oncogenic fusion protein. *Oncogene*, **22**, 6819–6829.
48. Toretsky,J.A. and Wright,P.E. (2014) Assemblages: Functional units formed by cellular phase separation. *J. Cell Biol.*, **206**, 579–588.
49. David,C.J. and Manley,J.L. (2010) Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev.*, **24**, 2343–2364.
50. Srinivasan,R.S., de Erkenez,A.C. and Hemenway,C.S. (2003) The mixed lineage leukemia fusion partner AF9 binds specific isoforms of the BCL-6 corepressor. *Oncogene*, **22**, 3395–3406.
51. Zhao,K., Wang,W., Rando,O.J., Xue,Y., Swiderek,K., Kuo,A. and Crabtree,G.R. (1998) Rapid and phosphoinositol-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. *Cell*, **95**, 625–636.