Published in final edited form as:

Genes Chromosomes Cancer. 2011 May ; 50(5): 338–347. doi:10.1002/gcc.20858.

mRNA and Protein Levels of FUS, EWSR1 and TAF15 are Upregulated in Liposarcoma

Jessica I. Spitzer^{1,2}, Stacy Ugras³, Simon Runge¹, Penelope Decarolis³, Christina Antonescu⁴, Tom Tuschl^{1,*}, and Samuel Singer^{3,*}

¹ Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, New York, NY 10065

² Department of Pediatric Oncology, Hematology and Clinical Immunology, Center for Child and Adolescent Health, Heinrich Heine University, Duesseldorf, Germany

³ Sarcoma Biology Laboratory, Sarcoma Disease Management Program, Memorial Sloan-Kettering Cancer Center, New York, NY

⁴ Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, NY

Abstract

Translocations or mutations of *FUS*, *EWSR1* and *TAF15* (*FET*) result in distinct genetic diseases. N-terminal translocations of any FET protein to a series of transcription factors yields chimeric proteins that contribute to sarcomagenesis, whereas mutations in the conserved C-terminal domain of wild-type FUS were recently shown to cause familial amyotrophic lateral sclerosis. We thus investigated whether the loss of one *FUS* allele by translocation in liposarcoma may be followed by mutations in either the remaining *FUS* allele or the paralogous *EWSR1*. Furthermore, we investigated the strength of the FET promoters and their contributions to sarcomagenesis given the proteins' frequent involvement in oncogenic translocations. We sequenced the respective genomic regions of both *FUS* and *EWSR1* in 96 liposarcoma samples. Additionally, we determined FET transcript and protein levels in several liposarcoma cell lines. We did not observe sequence variations in either *FUS* or *EWSR1*. However, protein copy numbers reached an impressive 0.9 and 5.5 Mio of FUS and EWSR1 per tumor cell, respectively. Compared with adipose-derived stem cells, FUS and EWSR1 protein expression levels were elevated on average 28.6-fold and 7.3-fold, respectively. *TAF15* mRNA levels were elevated on average 3.9-fold, though with a larger variation between samples. Interestingly, elevated *TAF15* mRNA levels did not translate to strongly elevated protein levels, consistent with its infrequent occurrence as translocation partner in tumors. These results suggest that the powerful promoters of FET genes are predominantly responsible for the oncogenic effect of transcription factor translocations in sarcomas.

INTRODUCTION

FUS, EWSR1 and TAF15 constitute the FET family of RNA-binding proteins (RBPs). All three proteins predominantly localize to the nucleus and are ubiquitously expressed at high levels in human fetal and adult tissues (Wu et al., 2009). Their C-terminus is rich in serine, tyrosine, glutamine and glycine, and followed by several RGG motifs, a Cys₂ - Cys₂ zinc finger and an RNA-binding domain (Fig. 1). The direct RNA targets of these proteins

*Correspondence to: Tom Tuschl, Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, New York, NY 10065, ttuschl@rockefeller.edu; Samuel Singer, Sarcoma Biology Laboratory, Sarcoma Disease Management Program, Memorial Sloan-Kettering Cancer Center, New York, NY, singers@mskcc.org.

remain to be defined, but functionally, FET proteins have been implicated in regulating transcription and splicing (Tan and Manley, 2009).

Recently, mutations in the very C-terminus of FUS have been described in familial amyotrophic lateral sclerosis (FALS) in about 5 – 10 % of all cases (Kwiatkowski et al., 2009; Vance et al., 2009). These mutations coincide with a mostly cytoplasmic redistribution of FUS, which may be interpreted as a gain-of-function with respect to increased regulation of cytoplasmic targets or a loss-of-function with respect to regulation of nuclear targets.

Prior to the discovery that FUS mutations cause FALS, recurrent translocations of both FUS and EWSR1 were recognized in sarcomas (Delattre et al., 1992; Crozat et al., 1993; Rabbitts et al., 1993). The fusions join the N-terminal domain of FET proteins to various DNA-binding proteins, thereby generating transforming transcription factors expressed under the control of the strong FET gene family promoters. Liposarcoma is the single most common soft tissue sarcoma and accounts for at least 20% of all sarcomas in adults (Singer et al., 2007). Myxoid-round cell liposarcoma (MRCLS) is the second most common biological group of liposarcoma and is characterized by the presence of a reciprocal chromosomal translocation, t(12;16)(q13;p11), involving *FUS* and *DDIT3/CHOP*, found in more than 90% of MRCLS cases. *DDIT3* is a DNA-binding protein and its entire coding sequence tends to be merged with the N-terminal FUS coding region, which is believed to act as a transactivation domain causing *DDIT3*-target-specific transcriptional activation. Another, although much rarer, translocation is found in about 5 % of MRCLS (t(12;22)(q13;q12)), in which *DDIT3* is fused to *EWSR1* (Mertens et al., 2009). Furthermore, *FUS* and *EWSR1* are also frequent translocation partners other sarcoma types, including Ewing's sarcoma/peripheral primitive neuroectodermal tumor (Delattre et al., 1992), in which *EWSR1* is translocated in virtually all the cases, desmoplastic small-round cell tumor (Ladanyi and Gerald, 1994), and low grade fibromyxoid sarcoma (Mertens et al., 2005). Finally, TAF15 has also been shown to form fusion proteins in different types of cancer, although much less frequently compared with FUS and EWSR1 (Martini et al., 2002).

Interestingly, there are clear “tumor-preferences” for both *EWSR1* and *FUS* translocations. As already mentioned, the characteristic *FUS-DDIT3* translocation is found in 90% of all MRCLS whereas its *EWSR1* counterpart in only 5%. All of the MRCLS samples that we analyzed (in total 35 tissue samples and 7 cell lines) harbored the *FUS-DDIT3* translocation. Looking at Ewing's sarcoma, however, this picture is reversed. *EWSR1-FLI1* accounts for over 90% of all cases, *EWSR1-ERG* for about 5%, whereas *FUS-ERG* fusions are exceedingly rare (Mertens et al., 2009).

To date, due to a lack of powerful animal models of sarcomagenesis, it remains uncertain whether the gene fusions described above are sufficient for sarcomagenesis (Helman and Meltzer, 2003). Drawing parallels to leukemia, chromosome translocations are necessary, but not sufficient to directly cause leukemia (Greaves and Wiemels, 2003). Experiments in transgenic mice indicate that progression of chronic myelogenous leukemia (which frequently harbors the BCR-ABL1 fusion protein) to an acute phase or blast-cell crisis with differentiation arrest is associated with additional and diverse genetic changes.

The discovery of FALS-causing mutations in *FUS* (Kwiatkowski et al., 2009; Vance et al., 2009), but also the reduction in FET protein levels due to the loss of one allele in the translocation, prompted us to investigate whether mutations in the FET proteins or unexpected changes in the FET protein or transcript abundance play an additional role in sarcomagenesis.

Oligonucleotides

FLAG/HA *Bam*HI for,

5'GATCCAGCGTAAATCGGGCACGTCATAAGGGTACTTGTCATCGTCGTCCTTGTA
GTC ACCGGTG; FUS_pET23a *SalI* for,

5'ACGCGTCGACCCATGGCCTCAAACGATTATACC; FUS_pET23a_NotI_rev,

5'ATAGTTTACGCGCCGCATACGGCCTCTCCCTGCGATC; EWSR1_pET23a_ *EcoRI* for, 5' GCAATCCGAATTCATGGCGTCCACGGATTACAG;

EWSR1_pET23a_EcoRI_rev, 5'GCAATCCGAATTCGTGTAGGGCCGATCTCTGCGC;

NcoI for, 5'CAATCCCATGGACTACAAGGACGACGATGAC; *EcoRV* rev,

5'GCAATCGATATCTCAGTGGTGGTGGTGGTGGTG; FUS_PCR_for,

5'ACGCGTCGACATGGCCTCAAACGATTATACCCAAC; FUS PCR rev,

5'ATAGTTTACGCGCCGCTTAATACGGCCTCTCCCTGC; EWSR1 PCR for,

5'ACGCGTTCGACATGGCGTCCACGGATTACAGTAC: EWSR1 PCR rev.

5'ATAGTTT TAGCGGCCGCTAGTAGGGCCGATCTCTGC; TAF15 PCR for.

5'ACGCGTCGACATGTCGGATTCTGGAAG: TAF15 PCR rev.

5'ATAGTTTtagcggccgctcagtatgggtcgggttgc: FUS 1 for.

5'GTAAACGACGGCCAGTAGGTCTTGCCTATTCCCCAT: FUS 1 rev.

5' CAGGAAACAGCTATGACTCCACCTAGCCCTCAAAATG: FUS 2 for.

5'GTAAAAACGACGGGCCAGTCATTTTGAGGGGCTAGGTGGA: FUS 2 rev.

5' CAGGAAAACAGCTATGACCTCACCATTAAAAAGGGCCAA: EWSR1 1 for.

5'GTAAAAACGACGGCCAGTCATTCACTGGACGCTTCAGA: EWSR1_1 rev.

5'GTAATAACGACGGCCAGTCAATTCACCTGGACGCTTCAGG, EWSR1_1 for
5'CAGGAAAACAGCTATGACCCACTTCACTCAATCTCGGCA: EWSB1_2 for

5'GTAAAAACGACGGGCCAGTGCAGTTGCCCTCTGCTTAAC: EWSR1_2 rev

5'GTAATAACGACGGCCAGTGCAGTTGCCCTCTGCTTAAC; EWSR1-2-10v,
5'CAGGAAAACAGCTATGACCGCTCCTCCCCAACCACTATCA; EWSR1-3 for

5'GTAAAAACGACGGCCAGTTGATAGTGGTTGGGAGGAGC: EWSB1 3' rev

5'GTAAAACGACGGCCAGTTCATAGTGGTTCGGGAGGAGC; EWSR1_5-1c, 5'CAGGAAAACAGCTATGACCCTGTTTCCATGTTCACAACCG; EWSR1_4 for

5'GTAAAAACGACGGGCCAGTGCCGCCAGGCACAGTAA: EWSR1 4 rev

5' CAGGAAAACAGCTATGACCTCCGCACACTACCATTTTAACTT: EWSB1 5' for

5'GTAAAAACGACGGCCAGTATTGTGGAGAACCAAGAGGG: EWSB1 5' rev

5'GTAAAACGACGGCCAGTATTTGTGGAGAACCAAGAGGG, EWSR1_5_rev,
5'CAGGAAAACAGCTATGACCCTGTTTCCATGTTCAGAACCG: EWSR1_6 for

5'GTAAAAACGACGGCCAGTTTCCTTCTTTTAAAAAATGGTTGTTT: EWSB1.6 rev

5'GTAAAACGACGGCCAGTTCCTTCCTTTTAAAAATGTTTGTGTT, EWS
5'CAGGAAACAGCTATGACCCCATGAACATTGTTACAGTTAAGAGG:

5' CAGGAAACAGGCTATGACCCCATGAACATTGTTACAGTTAAGAGG,
FWSB1_7 for 5'GTAAAACGACGGCCAGTTCCTAAATGTTTATAAAGAGTCATCC:

EWSR1_7_fol, 5'GTAAAACGACGGCCAGTTCCTCAAATGTTTATATAAAAGATCAT
EWSR1_7_rev, 5'CAGGAAACAGCTATGACCCCATGGAAAGGCAAAGTTATTTC:

The pET23(a) vector (Novagen, Madison, WI, #69745) was modified to contain an N-terminal FLAG/HA-tag in addition to its C-terminal His6-tag yielding pET23(a)_mod. The pET23(a) plasmid was digested with *Bam*HI followed by ligation of pre-annealed oligodeoxynucleotides FLAG/HA_*Bam*HI_for and FLAG/HA_*Bam*HI_rev. PCR amplification using FUS_pET23a_*Sall*_for and FUS_pET23a_*Not*I_rev from cDNA derived from total HEK293 RNA yielded the coding sequence (CDS) of *FUS*. The PCR product was *Sall*/*Not*I digested and ligated into the *Sall*/*Not*I digested pET23(a)_mod vector. Another PCR was performed (primers *Nco*I_for/*Eco*RV_rev) and the FLAG/HA_FUS_His6 cDNA was amplified. Regular pENTR4 vector (Invitrogen) was *Nco*I/*Eco*RV digested. A fill-in reaction with T4 DNA polymerase was performed to create blunt-ended products of both the FLAG/HA_FUS_His6 cDNA and the pENTR4 vector. Then, the FLAG/HA_FUS_His6

cDNA was ligated into the pENTR4 vector. The pENTR4_FLAG/HA_FUS_His6 construct was recombined into pDEST8 destination vector using GATEWAY LR recombinase (Invitrogen, Carlsbad, CA). Similar cloning steps were performed to obtain the pDEST8_FLAG/HA_EWSR1R1_His6 plasmid except that EcoRI was used instead of SalI/NotI. Plasmids used in this study are available from www.addgene.com.

Plasmids for the Creation of Stable Cell Lines

Plasmids pENTR4 FUS, EWSR1 and TAF15 were generated by PCR amplification of the respective coding sequences (CDS) followed by restriction digest with *SalI* and *NotI* and ligation into pENTR4 (Invitrogen) (primers: FUS_PCR_for, FUS_PCR_rev, EWSR1_PCR_for, EWSR1_PCR_rev, TAF15_PCR_for, TAF15_PCR_rev). pENTR4 FUS, EWSR1 and TAF15 were recombined into pFRT_FLAGHA-DEST modified destination vectors (Invitrogen) using Gateway LR Clonase II enzyme mix according to the manufacturer's instructions (Invitrogen) to allow for constitutive expression of stably transfected FLAGHA-tagged protein in Flp-In T-REx HEK293 cells (Invitrogen) from the CMV promoter.

Cell Culture

Liposarcoma, malignant fibrous histiocytoma and adipose-derived stem cell lines were established from tissue samples. MPNST724, ST88-14 and GIST882 cell lines were a kind gift from Jonathan Fletcher (Brigham and Women's Hospital) as was the HSSYII cell line from Dr. Ladanyi / MSKCC. All of these cell lines were grown in a 1:1 ratio of F-12 and DMEM (Invitrogen), supplemented with 10% fetal bovine serum, 100 U/ml of penicillin, 100 µg/ml streptomycin and maintained at 37°C in 5% CO₂.

T-REx HEK293 Flp-In cells (Invitrogen) were grown in D-MEM high glucose (1x) with 10% fetal bovine serum, 100 U/ml of penicillin, 100 µg/ml streptomycin, 100 µg/ml zeocin and 15 µg/ml blasticidin. Cell lines stably expressing FLAGHA-tagged proteins were generated by co-transfection of pFRT_FLAGHA_gene_of_interest constructs with pOG44 (Invitrogen). Cells constitutively expressing either of three FLAGHA-tagged proteins were cultivated in D-MEM high glucose (1x) with 10% fetal bovine serum, 100 U/ml of penicillin, 100 µg/ml streptomycin and 100 µg/ml hygromycin.

Sf9 cells were grown in Grace's Insect Medium (Invitrogen), supplemented with 10% fetal bovine serum, 1 % Pluronic F-68 (Invitrogen), 100 U/ml of penicillin, 100 µg/ml streptomycin and maintained in room air at 26°C in spinner culture (80 rpm).

Recombinant Protein Expression and Purification in *Spodoptera Frugiperda* (Sf9) Cells

pDEST8_FLAG/HA_FUS_His6 was transformed into MAX Efficiency DH10Bac competent *E. coli* (Invitrogen). Bacmid DNA was isolated using PureLink HQ Mini Plasmid Purification (Invitrogen) and transfected into Sf9 cells using Cellfectin II Reagent (both from Invitrogen). Three rounds of viral amplifications yielded 250 ml of cell supernatant containing 2×10^8 plaque forming units/ml virus. 25 ml of this solution were used for infection of 1 liter of Sf9 culture maintained at a density of 1×10^6 cells/ml. Four days after infection, Sf9 cells were washed by centrifugation (500 x g) in 1x PBS, and pellets were suspended in 5x pellet volume of ice-cold lysis buffer (50 mM Tris-HCl, pH 8.0, 5 mM MgCl₂, 1 M KCl, 10% glycerol, 5 mM imidazol, 0.1% triton-X-100, 1 mM beta-mercaptoethanol and Complete EDTA-free protease inhibitor cocktail (Roche AG, Basel, Switzerland)). The suspension was incubated on ice for 10 min and then suspended by 20 strokes with a Dounce homogenizer. Insoluble material was removed by centrifugation for 20 min at 20,000 x g and the supernatant was further cleared by passing through a 5µm Supor membrane syringe filter (Pall Acrodisc, Ann Arbor, MI). Recombinant EWSR1 was

purified analogously. Sf9-expressed FET proteins were purified using the AektaExplorer (GE Healthcare, Piscataway, NJ). 10 ml of Co²⁺ TALON beads (GE Healthcare) were washed 3x with de-ionized water and packed into a XK 16 column (GE Healthcare). The column was equilibrated with 4 column volumes (CV) lysis buffer supplemented with 5 mM imidazole and 1 mM DTT. After loading the cell lysate, the column was washed with 2 CV lysis buffer supplemented with 5 mM imidazole and 2 CV supplemented with 13 mM imidazole. Proteins were eluted from the column in 4 CV elution buffer (50 mM Tris-HCl, pH 8.0, 5 mM MgCl₂, 1 M KCl, 10% glycerol, 400 mM imidazol, 0.1% triton-X-100, 1 mM beta-mercaptoethanol) running a gradient with a final concentration of 400 mM imidazole. Fractions containing the protein of interest were pooled and dialyzed overnight in 2 x 1 l of dialysis buffer (20 mM Tris-HCl, pH 8.0, 300 mM KCl, 5 mM MgCl₂, 0.1% v/v Triton X-100, 50% v/v glycerol, 1 mM DTT) using dialysis bags (Spectrum Laboratories, Rancho Dominguez, CA, Spectra/Por, 08-667E) with a membrane molecular cutoff of 12 – 14 kDa. Protein concentrations were estimated by comparing Coomassie stain intensity against a BSA standard on a 10% SDS-PAGE gel.

Patient Sample Acquisition, RNA Isolation, Reverse Transcription and Real-Time Quantitative PCR (qRT-PCR)

Informed consent was obtained and the project was approved by the local review board (IRB protocol 02-060). Tissue samples were macro-dissected from cryomolds following histological review by Dr. Antonescu to assure subtype uniformity and removal of necrotic/normal tissue. Tumor tissue was lysed using QIAzol lysis reagent, as previously described (Singer et al., 2007). RNA isolation from frozen tissue was performed using RNeasy Lipid Tissue Mini Kit (Qiagen, Vamencia, CA) and from cell lines using the RNeasy Kit (Qiagen).

Reverse transcription was performed with 1.5 µg total RNA in a 100 µl reaction using random hexamer priming and TaqMan reverse transcription reagents (Applied Biosystems (AB), Foster City, CA) on a Thermo Hybaid thermocycler. All qRT-PCR assays were done on the ABI Prism 7900HT Sequence Detection System and analyzed using SDS version 2.1 software (AB). TaqMan Gene Expression Assays (AB), which include gene-specific probe and primer sets, were used according to the manufacturer's protocol to detect *EWSR1* (Hs01580532_g1), *FUS* (Hs01100216_g1), *TAF15* (Hs00896635_m1) and 18s rRNA (Hs99999901_s1). The relative expression levels of the above-mentioned genes were determined with the $\Delta\Delta C_t$ method using 18s rRNA as an endogenous control.

Preparation of Whole Cell Extracts and Western Blotting

Cells were washed with ice-cold PBS and lysed in 3x pellet volumes 10 mM HEPES-KOH, pH 7.5, 150 mM KCl, 2 mM EDTA, 1 mM NaF, 0.5% NP-40, 0.5 mM DTT and complete EDTA-free protease inhibitor cocktail (Roche). The lysate was incubated for 10 min on ice, centrifuged for 10 min at 20,000 x g (4°C) and total protein concentration was measured by Bradford protein assay.

Whole cell lysates and recombinant purified proteins were analyzed on a 10% SDS-PAGE gel. For a quantitative approach, cells were counted before cell lysis and the exact volume after cell lysis was also recorded. Purified recombinant protein (FUS/EWSR1) was also loaded on the same SDS-PAGE gels in decreasing concentrations as standards. Protein samples were then transferred to nitrocellulose membrane (BioRAD, Hercules, CA; Trans-Blot; 1.5 mAmp per cm² membrane for 1.5 h) and probed with the indicated antibodies. The incubation with both primary and secondary antibody was kept short to avoid binding saturation. Signals were developed using the ECL kit (GE Healthcare) under standard conditions. The luminescence signal was recorded with a Fujifilm Image Reader LAS-3000 and quantified using ImageJ.

After measuring the band intensities, the point of equivalence was determined at which the intensity of the standard equaled that of the expression of the protein of interest in the whole cell lysate. Next, the mass of the protein of interest in the whole cell lysate was determined (mass in g of protein = average concentration in g per μ l lysate x total μ l lysate). With that, the number of moles of the protein of interest (number of moles = mass in g of protein in the whole lysate / molecular weight in kilo Dalton) and then the molecules in the whole lysate (number of molecules = number of moles / 6.02×10^{23}) were calculated. By dividing the resulting number by the cell number in the original cell pellet the copy number of EWSR1 / FUS protein molecules per cell was obtained.

Antibodies

Monoclonal anti-EWSR1 (Abcam, AB54708), polyclonal anti-FUS (Abcam, Cambridge, MA, AB23439 (Deng et al., 2010)), polyclonal anti-TAF15 (Sigma, St. Louis, MO, AV30112) and monoclonal beta-tubulin (Sigma, T4026 (Wolff et al., 1988)) were used as primary antibodies at a 1:1000 dilution. HRP-conjugated anti-rabbit Ig and anti-mouse Ig were used as secondary antibodies for Western blot analysis.

Genomic DNA Isolation, Sequencing and Mutation Detection

Genomic DNA from patient samples and cell lines was extracted using the Qiagen DNeasy Blood & Tissue Kit. PCR reactions were carried out in 384 well plates, in a Duncan DT-24 water bath thermal cycler, with 10 ng of whole genome amplified DNA as template, using a touchdown PCR protocol with HotStart Kapa Fast Taq (Kapa Biosystems, Woburn, MA): 1 cycle of 95°C for 5 min; 3 cycles of 95°C for 30 sec, 64°C for 15 sec, 72°C for 30 sec; 3 cycles of 95°C for 30 sec, 62°C for 15 sec, 72°C for 30 sec; 3 cycles of 95°C for 30 sec, 60°C for 15 sec, 72°C for 30 sec; 37 cycles of 95°C for 30 sec, 58°C for 15 sec, 72°C for 30 sec; 1 cycle of 70°C for 5 min. Templates were purified using AMPure (Agencourt Biosciences, Beverly, MA). The purified PCR reactions were split into two, and sequenced bidirectionally using Big Dye Terminator Kit v.3.1 (Applied Biosystems, Foster City, CA), at Agencourt Biosciences. Dye terminators were removed using the CleanSEQ kit (Agencourt Biosciences), and sequence reactions were run on ABI PRISM 3730xl sequencing apparatus (Applied Biosystems).

Mutations were detected using an automated pipeline at the MSKCC Bioinformatics Core. Bi-directional reads and mapping tables (to link read names to sample identifiers, gene names, read direction, and amplicon) were subjected to a QC filter which excludes reads that have an average phred score of < 10 for bases 100–200. Passing reads were assembled against the reference sequences for each gene, containing all coding and UTR exons including 5Kb upstream and downstream of the gene, using command line Consed 16.0 (PMID: 9521923). Assemblies were passed on to Polyphred 6.02b (PMID: 9207020) which generated a list of putative candidate mutations, and to Polyscan 3.0 (PMID: 17416743) which generated a second list of putative mutations. To reduce the number of false-positives, only point mutations which were supported by at least one bi-directional read pair and at least one sample mutation called by Polyphred were considered and manually inspected.

RESULTS

Absence of Sequence Variation in the Last Three Exons of FUS in 96 Human Liposarcoma Samples

To search for mutations similar to those recently published in familial amyotrophic lateral sclerosis (FALS), we sequenced the last three FUS exons 13 to 15 (Fig. 1) from DNA isolated from 43 dedifferentiated (DDLs) and 35 myxoid-round cell (MRCLS) liposarcoma tissue samples as well as 4 well differentiated (WDLs), 7 DDLs and 5 MRCLS cell lines

(Table 1). One of the DDLS and one of the MRCLS cell lines had been established from patient samples that were simultaneously sequenced. All MRCLS patient samples and cell lines harbor the *FUS-DDIT3* translocation described above.

We detected no mutations and also found no single nucleotide polymorphisms (SNPs) in the *FUS* segments of the 96 samples that we sequenced, despite several reported SNPs in the selected genomic region (NCBI database of single nucleotide polymorphisms (Sherry et al., 2001): rs68121698, rs72550884, rs72550885, rs68121698, rs13331793). However, those SNPs occurrences are biased towards African populations whereas our patient population was mostly Caucasian. Another gene, *ZIC1*, was fully sequenced simultaneously in the same samples and revealed polymorphic variation, thereby validating the experimental approach (Brill et al., 2010).

We also sequenced the paralogous region of the *EWSR1* gene, but although we confirmed a previously reported SNP (rs2518683; average heterozygous frequency 19%) in three patient tissues (3.1%), we did not find any new mutations. The difference in reported and observed polymorphic frequency is most likely explained by ethnical differences in patient population and our small sample size. Additionally, we found two previously not reported, intronic sequence variations (hg19: chr22:28,025,421 and chr22:28,025,871) in two independent DDLS samples, the first was a homozygous base change from C to T, the second a heterozygous base change from C to T.

mRNA Expression Analysis of FET Family Members

Next, we examined how the loss of one FET family member allele affected FET family mRNA and protein levels in various liposarcoma cell lines compared to reference samples.

First, we determined relative mRNA levels by real-time quantitative PCR (qRT-PCR) (Fig. 2A). In total, we tested 20 sarcoma cell lines, 14 of which were derived from liposarcoma, 4 from malignant fibrous histiocytoma, a synovial sarcoma cell line and a malignant peripheral nerve sheath tumor cell line. We also profiled a gastrointestinal stromal tumor cell line and a Schwannoma cell line. For reference, we studied 3 HEK293 Flp-In T-REx cell lines stably overexpressing one FET member each, and two adipose-derived stem cells (ASC) from different individuals, each at three stages of differentiation (ASC1 and ASC2 at 0, 4 and 10 days of differentiation) (Table 1).

Comparing the expression levels across the ASCs, *EWSR1* was the most highly expressed mRNA ($\Delta\Delta\text{Ct}$ value 25.5) compared to *FUS* and *TAF15* (both: $\Delta\Delta\text{Ct}$ value 26.8).

Since FET mRNA levels varied only slightly at the different differentiation stages, we decided to set the $\Delta\Delta\text{Ct}$ value for ASC1day 0 to 1 and adjusted all other $\Delta\Delta\text{Ct}$ values accordingly. Compared to their levels in undifferentiated ASCs, FET mRNA levels were elevated 1.3-fold, 2.4-fold and 3.9-fold in *EWSR1*, *FUS* and *TAF15*, respectively across all tested tumor cell lines (Fig. 2A).

Looking specifically at liposarcoma cell lines with *FUS-DDIT3* translocations (MRCLS) and those characterized by 12q amplification without *FUS-DDIT3* translocations (DDLS, WDLS), there was a 5.3% decrease in mRNA levels for *FUS*, and an 8.3% increase of *TAF15* compared with ASCs. Disregarding the outlying value for the ML3184 cell line which most likely represents a technical failure, results in an 11.2% increase in *EWSR1* mRNA levels.

The fact that the mRNA level for *FUS* in the MRCLS cell line harboring a *FUS-DDIT3* translocation was only 9.3% reduced instead of the expected 50% down-regulation suggests

that posttranscriptional regulation compensates for the loss of one allele upon translocation to a certain degree.

Protein expression analysis of FET proteins

We selected a cell line derived from DDLS characterized by 12q amplification and one from MRCLS that contained a *FUS-DDIT3* translocation. Protein copy numbers per cell were determined by quantitative Western blotting using recombinantly expressed and purified proteins as reference standards (Fig. 2B). TAF15 was not examined, as its involvement in translocation is rare compared with *FUS* or *EWSR1*.

Antibodies were specific to the C-termini of *FUS* and *EWSR1* proteins and are unable to cross-react with *FUS* or *EWSR1* fusion proteins. *EWSR1* copy numbers were 5.4 and 5.6 Mio in the DDLS the MRCLS cell line, respectively, and *FUS* copy numbers were 0.87 and 0.47 Mio respectively.

To get a better sense of the variability in FET protein expression, we expanded the analysis to include one more DDLS and MRCLS cell line as well as undifferentiated ASCs. The differences in FET protein levels were calculated after normalizing the Western blotting signals to those of beta-tubulin SDS-PAGE gel loading controls. Both MRCLS cell lines carry the described *FUS-DDIT3* translocation.

FET proteins, when compared with their abundance in the ASCs, were elevated 4.5- to 10.5-fold for *EWSR1*, 16- to 41-fold for *FUS*, and 2.1- to 2.9-fold for TAF15 (Fig. 2C and D). These increases in protein level were also mirrored by increased mRNA levels as determined above (Fig. 2A).

The increase in protein level (values for mRNA expression level increases are given in brackets) specifically in DDLS cell lines for *EWSR1* were 7.4 (1.4), for *FUS* 39.3 (2.2) and for TAF15 2.5 (3.8). The results for the MRCLS cell lines which both harbor a *FUS* translocation were as follows: *EWSR1* 5.7 (1.5), *FUS* 17.2 (2.7) and TAF15 2.5 (4.0).

DISCUSSION

Translocations involving the promoter and N-terminal domains of the FET family of RNA-binding proteins have long been recognized as important players in human sarcomagenesis, and are now widely used by pathologists to improve diagnostic accuracy and tumor classification. While most molecular studies of sarcomagenesis focused on the C-terminal translocation partners, typically DNA-binding proteins, less effort has been dedicated to the analysis of FET proteins and their function in sarcomagenesis. Nevertheless, while the chimeric translocation proteins are considered to act as oncogenic transcription factors activating downstream targets important in sarcomagenesis, FET proteins may also function as tumor suppressors, and the concomitant loss of one of its alleles may be followed by mutations in its remaining allele. The latter hypothesis has never been tested, but when C-terminal domain mutations in *FUS* were shown to cause familial amyotrophic sclerosis (FALS) (Kwiatkowski et al., 2009; Vance et al., 2009), we were interested in assessing the mutational status of *FUS* and its paralog *EWSR1* in this region. Mutations in the C-terminal domain of *FUS* alter its subcellular localization from nuclear to cytoplasmic. Mutations in *FUS* explain about 5 – 10 % of cases of FALS (Kwiatkowski et al., 2009; Vance et al., 2009).

We were unable to detect *FUS* mutations or SNPs in any of the 96 clinical samples that we sequenced. 40 out of these 96 samples have a *FUS-DDIT3* translocation. Given the frequency at which FALS is caused by mutations in *FUS*, we might have expected at least a

few patients to carry a mutation, especially if the mutations functioned as tumor suppressors, yet we detected none. It remains a possibility that mutations outside of the sequenced region might be found, but to date, our sequenced region harbored the vast majority of known mutations linked to FALS. It is probably safe to assume that mutations in *FUS* as found in FALS are not oncogenic and control pathways unrelated to tumorigenesis but rather contribute to the death of motor neurons. With two exceptions, we could also not detect sequence variations in *EWSR1*. Since both of these were intronic and not close to splice junctions, their pathogenetic relevance remains unclear.

Another interesting aspect of FET family genes is their involvement in translocations that contribute to sarcomagenesis and leukemogenesis. This suggested strong promoters and expression of FET proteins in tumors or their progenitor cells in order for the fusion proteins to drive oncogenic events. Indeed, FET family member mRNA and protein levels were elevated in liposarcoma cell lines compared with adipose-derived stem cells. Protein copy numbers per cell were extremely high, reaching 5,400,000 for *EWSR* and 870,000 for *FUS* in a dedifferentiated liposarcoma cell line with 12q amplification that does not carry a *FUS* translocation. The copy numbers of the myxoid-round liposarcoma cell line, which harbors the characteristic *FUS-DDIT3* translocation are 5,555,753 and 472,907, respectively. In comparison, published protein copy numbers include e.g. PABP with 800,000 molecules per HeLa cell (Gorlach et al., 1994) and MYC with 33,000 copies per cell in primary human lung fibroblasts (Rudolph et al., 1999).

Interestingly, our results in liposarcoma show that *EWSR1* has more than 6 times as many copies per cell as *FUS* indicating that *EWSR1* is under the control of an even more powerful promoter. Overall, *EWSR1* translocations in sarcomas - with the exception of liposarcoma - are much more common than *FUS* translocation (Mertens et al., 2009). Our results indicate that this might be due to the fact that *EWSR1* fusions with their very strong promoter have even more oncogenic potential than those of *FUS*.

Compared with *FUS* and *EWSR1*, *TAF15* mRNA levels were on average stronger elevated across samples which interestingly did not translate into equally highly elevated protein levels. One possible reason might be autoregulation of FET proteins for which we found evidence in our other studies (unpublished results). At the same time, this explains why *TAF15* is much less frequently used as a translocation partner in tumors since translocated protein levels do not reach similar levels compared to the much more frequently involved *FUS* and *EWSR1*.

In the existing liposarcoma mouse model, the *FUS-DDIT3* translocation was expressed under the control of the EF1 promoter (Perez-Losada et al., 2000). Since this achieved only a modest overexpression it is likely that the resulting *FUS-DDIT3* levels are lower than in actual human tumors. This might explain why, apart from this one disease model, no other murine liposarcoma models have been established so far.

Overall, this raises the question whether the N-terminal portions of the *FUS* and *EWSR1* translocations serve a biological purpose other than expressing the C-terminal translocation partner under the extremely strong FET protein promoter.

Acknowledgments

We are grateful to I. Dolgalev (MSKCC Genomic Core Facility) for genomic DNA sequencing.

Supported by: Deutsche Forschungsgemeinschaft (J.S.), NIH grants MH080442 and 1CA145442 in addition to several Starr Cancer Consortium grants (T.T.). T.T. is an HHMI investigator as well as cofounder and scientific advisor to Alnylam Pharmaceuticals and an advisor to Regulix Therapeutics. S.S. is PI of the soft tissue sarcoma

PO1 and work in his laboratory was supported by the National Cancer Institute at the National Institutes of Health through a Soft Tissue Sarcoma Program Project grant: P01 CA047179 and Starr Cancer Consortium Grant I3-A151.

References

- Brill E, Gobble R, Angeles C, Lagos-Quintana M, Crago A, Laxa B, Decarolis P, Zhang L, Antonescu C, Socci ND, Taylor BS, Sander C, Koff A, Singer S. ZIC1 overexpression is oncogenic in liposarcoma. *Cancer Res* 2010;70:6891–6901. [PubMed: 20713527]
- Crozat A, Aman P, Mandahl N, Ron D. Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. *Nature* 1993;363:640–644. [PubMed: 8510758]
- Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G, Aurias A, Thomas G. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 1992;359:162–165. [PubMed: 1522903]
- Deng HX, Zhai H, Bigio EH, Yan J, Fecto F, Ajroud K, Mishra M, Ajroud-Driss S, Heller S, Sufit R, Siddique N, Mugnaini E, Siddique T. FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis. *Ann Neurol* 2010;67:739–748. [PubMed: 20517935]
- Gorlach M, Burd CG, Dreyfuss G. The mRNA poly(A)-binding protein: localization, abundance, and RNA-binding specificity. *Exp Cell Res* 1994;211:400–407. [PubMed: 7908267]
- Greaves MF, Wiemels J. Origins of chromosome translocations in childhood leukaemia. *Nat Rev Cancer* 2003;3:639–649. [PubMed: 12951583]
- Helman LJ, Meltzer P. Mechanisms of sarcoma development. *Nat Rev Cancer* 2003;3:685–694. [PubMed: 12951587]
- Kwiatkowski TJ Jr, Bosco DA, Leclerc AL, Tamrazian E, Vanderburg CR, Russ C, Davis A, Gilchrist J, Kasarskis EJ, Munsat T, Valdmanis P, Rouleau GA, Hosler BA, Cortelli P, de Jong PJ, Yoshinaga Y, Haines JL, Pericak-Vance MA, Yan J, Ticozzi N, Siddique T, McKenna-Yasek D, Sapp PC, Horvitz HR, Landers JE, Brown RH Jr. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 2009;323:1205–1208. [PubMed: 19251627]
- Ladanyi M, Gerald W. Fusion of the EWS and WT1 genes in the desmoplastic small round cell tumor. *Cancer Res* 1994;54:2837–2840. [PubMed: 8187063]
- Martini A, La Starza R, Janssen H, Bilhou-Nabera C, Corveleyn A, Somers R, Avenirin A, Foa R, Hagemijer A, Mecucci C, Marynen P. Recurrent rearrangement of the Ewing's sarcoma gene, EWSR1, or its homologue, TAF15, with the transcription factor CIZ/NMP4 in acute leukemia. *Cancer Res* 2002;62:5408–5412. [PubMed: 12359745]
- Mertens F, Antonescu CR, Hohenberger P, Ladanyi M, Modena P, D'Incalci M, Casali PG, Aglietta M, Alvegard T. Translocation-related sarcomas. *Semin Oncol* 2009;36:312–323. [PubMed: 19664492]
- Mertens F, Fletcher CD, Antonescu CR, Coindre JM, Colecchia M, Domanski HA, Downs-Kelly E, Fisher C, Goldblum JR, Guillou L, Reid R, Rosai J, Sciort R, Mandahl N, Panagopoulos I. Clinicopathologic and molecular genetic characterization of low-grade fibromyxoid sarcoma, and cloning of a novel FUS/CREB3L1 fusion gene. *Lab Invest* 2005;85:408–415. [PubMed: 15640831]
- Perez-Losada J, Pintado B, Gutierrez-Adan A, Flores T, Banares-Gonzalez B, del Campo JC, Martin-Martin JF, Battaner E, Sanchez-Garcia I. The chimeric FUS/TLS-CHOP fusion protein specifically induces liposarcomas in transgenic mice. *Oncogene* 2000;19:2413–2422. [PubMed: 10828883]
- Rabbitts TH, Forster A, Larson R, Nathan P. Fusion of the dominant negative transcription regulator CHOP with a novel gene FUS by translocation t(12;16) in malignant liposarcoma. *Nat Genet* 1993;4:175–180. [PubMed: 7503811]
- Rudolph C, Adam G, Simm A. Determination of copy number of c-Myc protein per cell by quantitative Western blotting. *Anal Biochem* 1999;269:66–71. [PubMed: 10094776]
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29:308–311. [PubMed: 11125122]
- Singer S, Socci ND, Ambrosini G, Sambol E, Decarolis P, Wu Y, O'Connor R, Maki R, Viale A, Sander C, Schwartz GK, Antonescu CR. Gene expression profiling of liposarcoma identifies

- distinct biological types/subtypes and potential therapeutic targets in well-differentiated and dedifferentiated liposarcoma. *Cancer Res* 2007;67:6626–6636. [PubMed: 17638873]
- Tan AY, Manley JL. The TET family of proteins: functions and roles in disease. *J Mol Cell Biol* 2009;1:82–92. [PubMed: 19783543]
- Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, Hu X, Smith B, Ruddy D, Wright P, Ganesalingam J, Williams KL, Tripathi V, Al-Saraj S, Al-Chalabi A, Leigh PN, Blair IP, Nicholson G, de Belleruche J, Gallo JM, Miller CC, Shaw CE. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 2009;323:1208–1211. [PubMed: 19251628]
- Wolff A, de Nechaud B, Chillet D, Gros F. The aminoterminal domain of alpha-tubulin probed by monoclonal antibodies. Recognition of a rarely exposed epitope by the monoclonal antibody 111 B52 C2. *Biol Cell* 1988;63:319–326. [PubMed: 2465048]
- Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, Hodge CL, Haase J, Janes J, Huss JW 3rd, Su AI. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol* 2009;10:R130. [PubMed: 19919682]

Figure 1.

Domain organization of the FET family members. The black vertical arrows indicate common breakpoints in the above-mentioned cancer types. The main region of mutations causing FALS is indicated by a pink arrow. The outmost sequencing primers (Results) used for the *EWSR1* and *FUS* sequencing are indicated by red horizontal arrows. For *FUS*, two primer pairs were used, in the case of *EWSR1* (due to difficult template) a total of seven pairs were employed. All protein domains implicated in RNA-binding are highlighted in yellow. The numbers refer to exons.

Figure 2.

mRNA and protein levels of EWSR1, FUS and TAF15. (A) FET mRNA expression levels across all samples. FET mRNA levels were increased compared with ASCs at different days of differentiation. Cell lines from left to right: 2 ASCs isolated from two individuals (each at three stages of differentiation), 5 WD, 4 DD, 5 MRC (which all carry a *FUS* translocation); 4 malignant fibrous histiocytoma cell lines (MFH); a malignant peripheral nerve sheath tumor cell line (MPNST724); a malignant Schwannoma cell line (ST88–14); a gastrointestinal stromal tumor cell line (GIST882); a synovial sarcoma cell line (HSSYII); 3 HEK293 cell lines each overexpressing one of the three FET proteins. The $\Delta\Delta C_t$ values for ASC1day 0 were set to 1 and all other values adjusted accordingly. The corresponding qRT-PCR cycle numbers were as follows: 25.14 (*EWSR1*), 26.32 (*FUS*) and 27.05 (*TAF15*). The eight liposarcoma cell lines which had been among those previously sequenced are in bold. (B) Determination of EWSR1 and FUS protein copy numbers. The amounts of recombinant protein (in ng) and volumes (in μ l) of whole cell lysate of the individual cell lines loaded on gels are indicated. The recombinant protein has a slightly larger molecular weight than the wildtype protein due to its two affinity tags (FLAG/HA, His6). Images were cropped but no other bands were visible on the full-length images. (C) and (D) FET protein expression levels of six selected cell lines. Western blot of ASCs at differentiation day 0 and five liposarcoma cell lines. Whole cell lysates were normalized (1.63 μ g/ μ l) and 23.4 μ g total protein each was analyzed by SDS-PAGE. (C) shows the Western blot result and (D) the densitometrical determination of the band intensities normalized to the beta-tubulin band intensities. The intensities across all visible bands were measured since these most likely represent different isoforms of the two proteins.

Table 1

Overview of tissue samples and cell lines used in the study.

Name	Abbreviation	Names of cell lines	#	Analysis Performed
<i>Tissue Samples and Cell Lines Used in Sequencing</i>				
well differentiated liposarcoma	WDLS	RWD8000, RWD3051, WD0082, RWD5700	4	sequencing
myxoid-round cell liposarcoma	MRCLS	n/a (tissue samples)	35	sequencing
myxoid-round cell liposarcoma	MRCLS	RC5397, ML2308, ML4415, ML6468, MRC2786	5	sequencing
dedifferentiated liposarcoma	DDLS	n/a (tissue samples)	43	sequencing
dedifferentiated liposarcoma	DDLS	DD5577p6, DDLS8817, LPS141, RDD8107-2, RDD6960-2, DD5590-EP, DD8200	7	sequencing
<i>Cell Lines Used in Other Applications</i>				
adipose-derived stem cells (each at three stages of differentiation)	ASC	ASC1, ASC2	2	qRT-PCR ^a , WB ^b (ASC1) ^c
well differentiated liposarcoma	WDLS	RWD3051, WD0082, WD0557, WD6736, WDLs9509	5	qRT-PCR, WB (WD0082)
myxoid-round cell liposarcoma	MRCLS	ML2308, ML3184, ML4415, ML8140, RC5397	5	qRT-PCR, qWB ^d (ML2308), WB (ML2308, RC5397)
dedifferentiated liposarcoma	DDLS	DD6960, DDLS8817, LPS141, RDD8107	4	qRT-PCR, qWB (DDLS8817), WB (DDLS8817, LPS141)
malignant fibrous histiocytoma	n/a	MFH0021, MFH2068, MFH3438, RMFH8500	4	qRT-PCR
synovial sarcoma	n/a	HSSYII	1	qRT-PCR
malignant peripheral nerve sheath tumor	n/a	MPNST724	1	qRT-PCR
malignant Schwannoma	n/a	ST88-14	1	qRT-PCR
gastrointestinal stromal tumor	n/a	GIST882	1	qRT-PCR
human embryonic kidney cell line stably overexpressing FLAGHA-tagged EWSR1	n/a	HEK293-EWSR1	1	qRT-PCR, WB
human embryonic kidney cell line stably overexpressing FLAGHA-tagged FUS	n/a	HEK293-FUS	1	qRT-PCR, WB
human embryonic kidney cell line stably overexpressing FLAGHA-tagged TAF15	n/a	HEK293-TAF15	1	qRT-PCR, WB

^a real-time quantitative PCR^b Western blotting^c If certain analyses were only performed on a subset of cell lines, their names are given in brackets.^d quantitative Western blotting