Identification of a Novel FUS/ETV4 Fusion and Comparative Analysis with Other Ewing Sarcoma Fusion Proteins

Megann A. Boone1,2, Cenny Taslim2, Jesse C. Crow2, Julia Selich-Anderson2, Mike Watson3, Peter Heppner4, James Hamill4, Andrew C. Wood4,5, Stephen L. Lessnick1,2,6, and Mark Winstanley4,5

ABSTRACT

Ewing sarcoma is a pediatric bone cancer defined by a chromosomal translocation fusing one of the FET family members to an ETS transcription factor. There have been seven reported chromosomal translocations, with the most recent reported over a decade ago. We now report a novel FET/ETS translocation involving FUS and ETV4 detected in a patient with Ewing sarcoma. Here, we characterized FUS/ETV4 by performing genomic localization and transcriptional regulatory studies on numerous FET/ETS fusions in a Ewing sarcoma cellular model. Through this comparative analysis, we demonstrate significant similarities across these fusions, and in doing so, validate FUS/ETV4 as a bona fide Ewing sarcoma translocation. This study presents the first genomic comparison of Ewing sarcoma–associated translocations and reveals that the FET/ETS fusions share highly similar, but not identical, genomic localization and transcriptional regulation patterns. These data strengthen the notion that FET/ETS fusions are key drivers of, and thus pathognomonic for, Ewing sarcoma.

Implications: Identification and initial characterization of the novel Ewing sarcoma fusion, FUS/ETV4, expands the family of Ewing sarcoma fusions and extends the diagnostic possibilities for this aggressive tumor of adolescents and young adults.

Introduction

Ewing sarcoma is an aggressive bone- and soft tissue–associated cancer diagnosed in children and young adults (1, 2). The disease is characterized by a chromosomal translocation that fuses the amino-terminal domain of a FET (FUS, EWSR1, and TAF15) protein to the carboxyl-terminal domain of an ETS transcription factor family member. Present in approximately 85% of cases, the most common chromosomal translocation, t(11;22)(q24;q22), fuses the EWSR1 and FLI1 genes to encode the EWS/FLI fusion oncoprotein (1, 2). EWS/FLI functions as an aberrant transcription factor that uses its ETS domain to bind DNA and the EWS-portion to regulate gene expression. In addition, an EWS/ERG fusion was identified in approximately 10% of cases, followed by five other fusions that are present in less than 1% of cases each: EWS/FEV, EWS/ETV1, EWS/ETV4, FUS/ERG, and FUS/FEV (3, 4). Each of these is believed to function as an aberrant transcription factor, primarily on the basis of their similar structure to EWS/FLI. Here, we report a novel eighth Ewing sarcoma fusion, FUS/ETV4, in a neonatal patient.

The ETS protein family is a group of transcription factors characterized by a highly-conserved DNA-binding domain, with structural variability outside of this region contributing to subfamily classification (5). The ETS members identified in Ewing sarcoma fusions derive from two of these subfamilies: FLI, ERG, and FEV are members of the ERG subfamily, and ETV1 and ETV4 of the PEA3 subfamily (1, 2, 5). These ETS family members bind similar high-affinity DNA sequences in vitro, but whether they have similar genomic localization in the context of the Ewing sarcoma fusions is unknown. This is of particular interest given the neomorphic capability of EWS/FLI to bind and regulate genes via GGAA-microsatellites in the human genome.

The amino-terminal intrinsically-disordered regions (IDR) of EWS and FUS have biophysical features that appear critical to the ability of FET/ETS proteins to bind DNA and regulate gene expression. These IDRs have self-association properties that are...
likely critical for FET/ETS fusion-mediated reorganization of chromatin architecture, formation of transcriptional hubs, and recruitment of necessary transcriptional co-factors, and are thus crucial for Ewing sarcomagenesis (1, 2, 6, 7).

The advent of next-generation sequencing has allowed for the ready identification of EWS-based and FUS-based genomic rearrangements (8). These technological advances have made the identification of fusion partners easier, but the vast majority of molecular studies of Ewing sarcoma fusions have focused on EWS/FLI and it is unclear how properties attributed to EWS/FLI translate to alternative FET/ETS fusions. While it makes logical sense that FET/ETS fusions have similar biologic functions, this has not been formally demonstrated. As such, the identification of alternative FET/ETS fusions has simultaneously raised new questions as to how to apply this information to clinical care. For example, a recent published survey found that only approximately 35% of clinician respondents indicated that non-EWS/FLI FUS/ETS fusions should be regarded as bona fide Ewing sarcoma or if these patients should be eligible for Ewing sarcoma clinical trials (9).

We now report the identification of a novel FUS/ETV4 fusion in a patient with Ewing sarcoma, and perform genomic localization and transcriptional studies in a Ewing sarcoma A673 knockdown/rescue model system. Here, we describe the first comparative analysis of rare, non-EWS/FLI fusions in Ewing sarcoma, and in doing so, demonstrate strong similarities between all of the fusions, indicating that they are indeed functionally similar. Simultaneously, we identify some differences between FET/ETS fusions that might represent differences in DNA-binding function and interplay between the FET and ETS portions of the fusions. These data support the conclusion that all FET/ETS translocations should be regarded as bona fide Ewing sarcoma translocations and clinically classified as such.

**Materials and Methods**

**Constructs and retroviruses**

Puromycin-resistant retroviral vectors encoding short hairpin RNAs (shRNA) targeting Luciferase (iLuc; sequence 5'-GATCCCC-TTACGCTGAGCTCCTGAATTTCGAGAGCTACTCAGC-GTAAGTTTTTGGAC-3') or the 3'-UTR of endogenous EWS/FLI mRNA (iEF; sequence 5'-GATCCCCCATAGGGGAGGCTATTA-TTTCAAAGAGAAATATACGTTCCACCATCTGGGGTGAAAAGC-3') were previously described (10, 11). FET/ETS fusions (containing amino-terminal 3xFLAG-tags) were cloned into pMSCV-Hygro (Invitrogen); sequence details provided in Supplementary Table S1.

**Cell culture methods**

HEK-293EBNA (ATCC, catalog no. CRL-10852, RRID:CVCL_6974) and A673 cells (ATCC, catalog no. CRL-1398, RRID:CVCL_0080), purchased in 2000, were cultured for 1 to 6 weeks in appropriate media (10–12). Short tandem repeat profiling and Mycoplasma testing (PCR) are performed annually on cell lines (approximately 3 months before experimentation). Retroviruses produced and used for infection as described (10–12).

**Immunodetection**

Whole-cell protein extraction, quantification, and Western blot analysis was performed as previously described (10–12). Immunoblotting was performed using anti-FLAG M2 mouse (Sigma F1804–200UG) and anti-α-Tubulin (Abcam ab7291). Membranes were imaged using the LI-COR Odyssey CLx Infrared Imaging System.

**qRT-PCR**

Total RNA was extracted from cells using the RNeasy Extraction Kit (Qiagen 74136). Reverse transcription and qPCR were performed using the iTag Universal SYBR Green 1-Step Reaction Mix (Bio-Rad 1725151) on a Bio-Rad CFX Connect Real-Time System. Primer sequences listed in Supplementary Table S2.

**Cleavage Under Targets and Tagmentation analysis**

Cleavage Under Targets and Tagmentation (CUT&Tag) was performed as described in ref. 13 on two biological replicates of knockdown/rescue A673 samples using the anti-FLAG M2 mouse antibody (1:100, Sigma F1804–200UG), and sequenced on the Illumina HiSeq4000 platform. Reads were trimmed, deduplicated using SAMTOOLS (RRID:SCR_002105), aligned to hg19 genome, spike-in normalized using DESeq2 (median ratio method, RRID:SCR_015687), tracks generated and averaged across biological replicates using Deeptools (RRID:SCR_016366), and peaks called using MACS (RRID:SCR_013291), DiffBind (RRID:SCR_012918), and DESeq2 (14–16). Peaks were called as significant with the following parameters: Irreproducible Discovery Rate of 0.01, FDR < 0.05, mean normalized counts > 80, log2(fold change) > 3 as compared with control samples (iEF + Empty Vector). Overlaps were determined using VennDiagram (RRID:SCR_002414) and GenomicRanges (RRID:SCR_000025; ref. 17).

**RNA Sequencing analysis**

RNA-seq was performed on two biological replicates of knockdown/rescue A673 cell samples. TruSeq Stranded mRNA Kit (Illumina catalog no. 20020594) was used to prepare cDNA libraries from total RNA and sequenced on Illumina HiSeq4000 to generate 150-bp paired-end reads. Reads were analyzed for quality control, trimmed, aligned to the hg19 reference genome, and analyzed for differential expression as compared with control cells (iEF + Empty Vector) using FastQC (RRID:SCR_014583), MultiQC (RRID:SCR_014982), Trim_Galore (RRID:SCR_011847), STAR (RRID:SCR_004463, version 2.5.2b), and DESeq2 (16). Venn diagrams were created for differentially expressed genes for samples as compared to control cells (iEF + Empty Vector; FDR < 0.05). Significantly regulated genes for each FET/ETS fusion were analyzed for pathway enrichment utilizing a hypergeometric test using DOSE (DOI: 10.18129/B9.bioc.DOSE), and plots generated using enrichplot (DOI: 1.18129/B9.bioc.enrichplot). Gene ontologies are from MSigDB (DOI: 10.18129/B9.bioc.msigdb).

**Statistical analysis**

PCR data is presented as mean ± SEM. Significance of data was determined using a Student t test, or as otherwise noted; P values < 0.05 were considered to be significant.

**Data availability**

The sequencing datasets generated and analyzed during the current study are available in the Gene Expression Omnibus and accessible at GSE173185. All other data generated or analyzed during this study are available from the corresponding author upon reasonable request.

**Ethics**

Nationwide Children’s Hospital Institutional Review Board determined that this project was not classified as human subjects’ research and was therefore exempt from review.
Results and Discussion

Identification of a novel FUS/ETV4 translocation

An infantile patient presented with a left posterior mediastinal mass (Fig. 1A). The mass occupied a significant portion of the left thoracic cavity and extensive intraspinal extension was observed from T3 to T8 without evidence of metastatic disease. A thoracic laminoplasty and resection of the intraspinal component was performed to manage the severely compressed spinal cord. The pathology of open-biopsy specimens revealed classic Ewing sarcoma with sheets of small round blue-staining cells with no evidence of differentiation (Fig. 1B). The tumor was CD99-positive in a diffuse membranous staining pattern (Fig. 1C), and positive for nuclear NKX2–2 expression (Fig. 1D).

EWS rearrangement was not detected, so FUS break-apart FISH was performed and identified a rearrangement. Commercial molecular genetic testing revealed a translocation between the FUS locus on chromosome 16p11.2 and the ETV4 locus on chromosome 17q21. This translocation encoded an in-frame fusion between exons 1 to 9 of FUS to exons 10–13 of ETV4. A literature search revealed the FUS/ETV4 translocation to be a novel fusion, previously unreported and undiscussed.

FUS/ETV4 has similar binding and transcriptional functions to EWS/ETV4

Neither patient-derived cell lines nor xenograft models were available to analyze biological functions of FUS/ETV4. We therefore cloned FUS/ETV4 into an expression vector to perform a functional analysis, and also cloned EWS/ETV4 as the most similar, bona fide Ewing sarcoma fusion (Fig. 2A). To allow for analysis in an isogenic background, we knocked-down endogenous EWS/FLI in A673 Ewing sarcoma cells and “rescued” expression of EWS/ETV4 or FUS/ETV4 fusion proteins through retroviral transduction (Supplementary Fig. S1A and S1B).

We first compared genome-wide localization of FUS/ETV4 and EWS/ETV4 using CUT&Tag (13). Both constructs were 3xFLAG-tagged and the use of the same anti-FLAG antibody for genomic localization allowed the data to be compared directly, without the
confounding effects of using different antibodies with different affinities and specificities. We found that FUS/ETV4-bound more than 12,000 loci and EWS/ETV4 bound more than 17,000 loci. Strikingly, more than 10,000 bound loci were shared between the fusions, and over 80% of FUS/ETV4 peaks overlapped with those of EWS/ETV4 (Fig. 2B).

We next asked whether FUS/ETV4 induced a similar transcriptional profile to EWS/ETV4. Consistent with the genomic localization studies, RNA-seq revealed that 87% of the genes regulated by FUS/ETV4 were also regulated by EWS/ETV4, although EWS/ETV4 again regulated more genes than FUS/ETV4 (Fig. 2C). Both fusions were capable of binding and regulating genes previously documented as EWS/FLI targets, including those associated with GGAmicrosatellite and high-affinity ETS-binding sites (Supplementary Fig. S1C and S1D). Taken together, these data demonstrate that the novel FUS/ETV4 fusion has transcriptional function that is similar to EWS/ETV4, and thus supports its identity as a bona fide Ewing sarcoma fusion.

**ERG- and FEV-based fusions have similar binding and transcriptional functions**

We recognized that this A673 knockdown/rescue model system could be generalized to compare other understudied Ewing sarcoma fusion proteins, particularly those ETS-family members that have both EWS and FUS versions. We therefore compared EWS/ERG with FUS/ERG and EWS/FEV with FUS/FEV (Fig. 3A; Supplementary Fig. S2A and S2B; Supplementary Fig. S3A–S3D; NB: neither EWS/FLI nor EWS/ETV1 have FUS versions identified to date). We found that almost 13,000 bound loci were shared between EWS/ERG and FUS/ERG, with more than 80% of the EWS/ERG loci also bound by FUS/ERG (Fig. 3B, left panel; NB: The higher number of FUS/ERG-bound loci likely reflects higher protein expression of FUS/ERG, see Supplementary Fig. S2B). Similarly, EWS/FEV and FUS/FEV shared almost 15,000 bound regions, accounting for approximately 70% of the regions bound by both fusions (Fig. 3B, right).

RNA-seq revealed that both EWS/ERG and FUS/ERG regulated more than 9,000 genes, approximately 80% of which were commonly regulated (suggesting that much of the “excess” FUS/ERG binding was not functionally associated with gene regulation; Fig. 3C, left panel). Similarly, EWS/FEV and FUS/FEV commonly regulated approximately 5,600 genes, representing approximately 65% of the genes regulated by EWS/FEV and approximately 90% of genes regulated by FUS/FEV (Fig. 3C, right panel).

Taken together with the ETV4-fusion data above, the finding that EWS/ETS and FUS/ETS fusions bind similar loci and regulate similar sets of genes suggest that the EWS and FUS regions of the fusions are largely interchangeable, and strengthen the notion that tumors harboring these fusions should be considered Ewing sarcomas.

**EWS- and FUS-based fusions have similar binding and transcriptional functions**

The analyses above demonstrated that fusions with the same ETS domain bind and regulate gene expression in a similar manner regardless if the fusion partner is EWS or FUS. We next sought to determine if DNA binding and gene regulation would be similar in fusions that had the same amino-terminus (either EWS or FUS), but differed in their ETS domain. We compared EWS/ETV4, EWS/ERG, and EWS/FEV alongside EWS/FLI (the most common Ewing sarcoma fusion) as a group, and FUS/ETV4, FUS/ERG, and...
FUS/FEV as a group in the A673 knockdown/rescue system (Supplementary Figs. S2A, S2B, and S3A–S3D). The EWS/ETS fusions shared almost 9,000 bound loci (Supplementary Fig. S4A), and the FUS/ETS fusions shared more than 8,700 bound loci (Supplementary Fig. S4B). RNA-seq showed similar trends, with more than 5,400 genes similarly regulated by the EWS/ETS fusions (Supplementary Fig. S4C), and approximately 2,900 genes regulated by the FUS/ETS proteins (Supplementary Fig. S4D).
Global similarities across all FET/ETS fusions support the inclusion of tumors harboring FET/ETS fusions as bona fide Ewing sarcomas

Lastly, we asked whether the similarities in DNA-binding and transcriptional regulation we observed in each “class” of fusion (grouped based on ETS domain or FET domain) would be observed across the entire group of FET/ETS fusions included herein. The data generated above was therefore analyzed in toto. Genomic localization revealed that approximately 6,600 loci were similarly bound and that approximately 2,600 genes were similarly regulated by all fusions tested (Fig. 3D). These overlaps were highly significant (P < 2.2 × 10−16). We again observed that all fusions bound and regulated both GGAA-microsatellite and high-affinity ETS site-associated genes (Supplementary Figs. S5A, SSB, S6A, and S6B). Pathway analyses revealed that FET/ETS fusions regulate pathways commonly associated with Ewing sarcoma cells, such as activation of proliferative pathways and repression of cellular adhesion, migration, and mesenchymal differentiation, and share gene-expression patterns that correlate with previously published Ewing sarcoma–specific datasets as well (Supplementary Figs. S7A, S7B, S8A, and S8B; refs. 18–20). Taken together, these data support the assertion that FET/ETS fusion proteins have similar capabilities to bind and regulate targets required for cellular processes critical for Ewing sarcomagenesis.

The most common fusion in Ewing sarcoma, EWS/FLI, has been extensively studied (1, 2). While novel functions of EWS/FLI, such as modulation of chromatin architecture and interaction with coregulators in transcriptional “hubs”, are still under investigation, it is widely accepted that fundamental properties of the fusion protein critical for Ewing sarcomagenesis include the ability to localize to specific loci in the genome, including those harboring GGAA microsatellites and/or high-affinity ETS-binding sites, and to dysregulate gene expression. Although additional Ewing sarcoma translocations have been identified, the analysis of these alternative fusions has been rudimentary at best and investigators have simply assumed similar function based on similar structure. At face value this seems reasonable, but leaves many unanswered questions, such as if there is a critical interplay between fusion type and cellular background that is required for Ewing sarcomagenesis or if there is a functional reason why some fusions are incredibly rare. Finally, the lack of important comparative analyses has allowed for confusion to arise in the clinical management of patients with likely Ewing sarcoma that harbor a rare translocation. Indeed, patients with rare translocations may not be offered entry onto clinical trials designed for patients with Ewing sarcoma and may therefore lead to subpar care (9).

In this report, we describe the identification of a novel FUS/ETV4 patient translocation. We demonstrate that this fusion shares many of the DNA-binding and gene-regulatory properties of other Ewing sarcoma–associated fusion proteins, including the well-studied EWS/FLI fusion. Through a large-scale comparison between variant Ewing fusions in an isogenic system, we find that all Ewing fusions analyzed share significant similarities in DNA binding and gene regulation. These data support the notion that the novel FUS/ETV4 fusion reported here is a bona fide Ewing sarcoma translocation, and suggest that FET/ETS translocations bind and regulate similar target genes to mediate oncogenesis. It is important to note that differences in patient tumor location and cellular state may contribute to differential activity of the FET/ETS fusions and should be investigated as fusion testing and patient sample preservation become common practice to further elucidate mechanisms of oncogenesis. At present, the data reported herein support that tumors containing FET/ETS translocations should be clinically diagnosed as Ewing sarcoma tumors and justifies the inclusion of patients with these tumors in standard and experimental Ewing sarcoma treatment protocols.

Authors’ Disclosures
M.A. Boone reports grants from NIH during the conduct of the study. J. Selich-Anderson reports grants from NIH. A.C. Wood reports a patent for PAT055863-US-PCT issued. S.L. Lessnick reports grants from NIH/NCI during the conduct of the study; personal fees from Salsarius Pharmaceuticals outside the submitted work; in addition, S.L. Lessnick has a patent for United States Patent No. US 7,939,253 B2, “Methods and compositions for the diagnosis and treatment of Ewing’s sarcoma” issued and a patent for United States Patent No. US 8,557,532, “Diagnosis and treatment of drug resistant Ewing’s sarcoma” issued; and an award U54 CA231641 during the conduct of the study. No disclosures were reported by the other authors.

Authors’ Contributions
M.A. Boone: Conceptualization, data curation, formal analysis, validation, investigation, visualization, methodology, writing—original draft, writing—review and editing. C. Taslim: Data curation, software, formal analysis, writing—review and editing. J.C. Crow: Investigation, writing—review and editing. J. Selich-Anderson: Investigation, writing—review and editing. M. Watson: Conceptualization, investigation, visualization, writing—review and editing. P. Heppner: Conceptualization, investigation, visualization, writing—review and editing. J. Hamili: Conceptualization, investigation, visualization, writing—review and editing. A.C. Wood: Conceptualization, investigation, visualization, writing—review and editing. S.L. Lessnick: Conceptualization, formal analysis, supervision, funding acquisition, methodology, writing—original draft, project administration, writing—review and editing. M. Winstanley: Conceptualization, supervision, investigation, visualization, project administration, writing—review and editing.

Acknowledgments
We sincerely thank Dr. Susan Arbuckle, Children’s Hospital at Westmead, Sydney, Australia, for assistance with NKX2-2 staining of tumor samples. We also thank Dr. Andrea K. Byrum, Dr. Emily R. Theisen, Dr. Jack Tokarsky, Ariunaa Bayanjargal, and Iftikhar A. Showpnil for thoughtful discussion concerning the hypothesis and methodology of the project. Additionally, we would like to acknowledge the generosity of the patient’s family for the use of clinical images and tumor sequencing information in this study. This work was supported by the NIH award U54 CA231461 (to S.L. Lessnick).

Received May 23, 2021; revised July 20, 2021; accepted August 26, 2021; published first August 31, 2021.
Comparison of FET/ETS Translocations in Ewing Sarcoma

Identification of a Novel FUS/ETV4 Fusion and Comparative Analysis with Other Ewing Sarcoma Fusion Proteins

Megann A. Boone, Cenny Taslim, Jesse C. Crow, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1541-7786.MCR-21-0354

Supplementary Material
Access the most recent supplemental material at:
http://mcr.aacrjournals.org/content/suppl/2021/08/28/1541-7786.MCR-21-0354.DC1

A diagrammatic summary of the major findings and biological implications:
http://mcr.aacrjournals.org/content/19/11/1795/F1.large.jpg

This article cites 20 articles, 3 of which you can access for free at:
http://mcr.aacrjournals.org/content/19/11/1795.full#ref-list-1

Sign up to receive free email-alerts related to this article or journal.

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

To request permission to re-use all or part of this article, use this link http://mcr.aacrjournals.org/content/19/11/1795.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.