Characterization of FET and ETS domain contributions to fusion oncoprotein activity in Ewing sarcoma

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Megann Alexandra Boone, B.S.

Biomedical Sciences Graduate Program

The Ohio State University

2021

Dissertation Committee

Stephen Lessnick, MD/PhD, Advisor

Timothy Cripe, MD/PhD

Lawrence Kirschner, MD/PhD

Mark Parthun, PhD

Abstract

Ewing sarcoma is an aggressive bone and soft tissue-associated cancer affecting pediatric, adolescent, and young adult patients. Despite general improvement in pediatric cancer outcomes due to novel therapeutic options, Ewing sarcoma treatment, which consists of high-dose chemotherapy, radiation, and/or local surgical control, has remained largely unchanged for several decades and patients with metastatic disease continue to see poor outcomes. Although pediatric cancers often have far fewer mutational events than adult cancers, Ewing sarcoma is particularly interesting as the disease is often characterized by a sole chromosomal translocation event: These chromosomal translocations fuse one of the FET protein family members, a group of putative RNA-binding proteins, to a member of the ETS transcription factor family. As these FET/ETS fusion proteins have been determined to function as oncogenic transcription factors responsible for driving Ewing sarcomagenesis, it is critical that the biological mechanisms these fusions utilize to facilitate this process are elucidated.

Despite discovery of several FET/ETS translocations, the majority of studies in the field focus on EWS/FLI, as it is the most common fusion observed in patients. Although these studies have provided a breadth of knowledge surrounding oncogenic function of the protein, there is a great deal of uncertainty how alternative FET/ETS fusions should be diagnosed and treated in the clinic. Herein, we characterize a novel FET/ETS fusion and perform the first comparative analysis on EWS/FLI and alternative, rarer FET/ETS fusion proteins. Our results reveal general similarities in DNA-binding and transcriptional regulation properties between the broad FET/ETS fusion group and provide the first tangible body of evidence to support that these fusions should indeed be classified as *bona fide* Ewing sarcoma tumors.

Furthermore, we sought to characterize contributions of the FLI protein to overall EWS/FLI function. Previously published data surrounding EWS/FLI function supports that the EWS domain is responsible for protein-protein interactions and transcriptional regulatory properties observed for the fusion protein. The FLI domain reportedly confers a sole function of DNA-binding to EWS/FLI activity. The vast majority of studies surrounding FLI contributions to the fusion protein were completed in alternative, non-Ewing sarcoma model systems. As new data surrounding other ETS factors have implicated regions surrounding the DNA-binding domain of the protein in modulation of overall activity, we sought to characterize regions surrounding the FLI DNA-binding domain in an appropriate Ewing sarcoma cellular model. Herein, we identify a novel role of FLI in EWS.FLI-mediated transcriptional regulation to the structural regions flanking

the FLI DNA-binding domain. These structural features are critical for overall EWS/FLI protein activity and downstream oncogenesis, revealing a possible novel therapeutic vulnerability that may be utilized in drug development of targeted inhibitors in the future.

Although it is widely accepted that FET/ETS fusion proteins act as oncogenic drivers of disease in Ewing sarcomagenesis, specific contributions of the FET and ETS domains to protein activity have yet to be fully elucidated. This body of work successfully determined that FET/ETS fusion proteins function similarly in Ewing sarcoma cells, and also identified a novel role of the ETS DNA-binding domain in transcriptional regulatory function of the fusion protein(s) that is required for oncogenesis. Together, these findings inform the clinical diagnosis process for Ewing sarcoma patients who present with a rare, alternative FET/ETS fusion and simultaneously identify a potential motif for FET/ETS protein inhibition that may be useful in the treatment of this aggressive disease, thus impacting patient outcomes.

Dedication

This work is dedicated to both of my grandmothers, who taught me to work hard, to stand up for what you believe in, and that laughter is the best medicine.

Acknowledgements

I would like to acknowledge my advisor, Dr. Stephen Lessnick, for encouraging me, providing me room to grow, teaching me about the scientific process, and for teaching me about life and who I want to be. I would also like to acknowledge my lab mates, who have offered never-ending advice, made science more fun, and become lifelong friends. I will surely look back on these years quite fondly. It was a pleasure to work in the Center for Childhood Cancer and Blood Diseases at The Abigail Wexner Research Institute at Nationwide Children's Hospital. The open, welcoming environment and wonderful colleagues I have found here have made graduate school that much better.

I would also like to acknowledge the teachers and professors who have dedicated their lives to their students, such as Mr. Charles Toumazous, Mrs., Shirley Shuman, Dr. Richard Weeks, and Dr. Charles Anthony. Without their belief and encouragement, I would not be where I am or who I am today. Lastly, I would like to thank my family and partner for their never-ending support. I am forever indebted to the sacrifices they have made to allow my sister and me to succeed in life. I can only hope that I will grow to be a strong, kind person deserving of all that I have been given, and will continually strive to pay it forward.

Vita

2012 to 2016 ..... B.S. Biochemistry,
West Virginia Wesleyan College

2016 to present ..... Ph.D. Candidate,
Biomedical Sciences Graduate Program
The Ohio State University

Publications

Boone MA, Taslim C, Crow JC, Selich-Anderson J, Watson M, Heppner P, Hamill J, Wood AC, Lessnick SL, Winstanley M. Identification of a novel FUS/ETV4 fusion and comparative analysis with other Ewing sarcoma fusion proteins. *Molecular Cancer Research, In Review.* 2021.

Boone MA, Taslim C, Crow JC, Selich-Anderson J, Byrum AK, Showpnil IA, Sunkel BD, Wang M, Stanton BZ, Theisen ER, Lessnick SL. The FLI portion of EWS/FLI contributes a transcriptional regulatory function that is distinct and separable from its DNA-binding function in Ewing sarcoma. *Oncogene, In Press*. 2021.

Murphy RA, Stafford RM, Petrasovits BA, **Boone MA**, Valentovic MA. Establishment of HK-2 cells as a relevant model to study tenofovir-induced cytotoxicity. *International Journal of Molecular Sciences*. 2017;18(3).

Fields of Study

Major Field: Biomedical Sciences

Minor Field: Clinical and Preclinical Research

Table of Contents

List of Tables

List of Figures

Chapter 1:

Introduction

1.1 Introduction to Ewing sarcoma

Ewing sarcoma is an aggressive bone and soft tissue-associated cancer typically seen in pediatric, adolescent and young adult patients (1-5). The disease is characterized by the presence of a chromosomal translocation fusing one of the FET protein family members to a member of the ETS transcription factor family (2, 4-9). Despite improvement in survival for many cancer types due to the advent of targeted therapeutic options, Ewing sarcoma patients have seen little improvement in outcomes in the past several decades (3, 10-15). Poor patient outcomes demand a better understanding of the basic biology underlying the disease to facilitate development of novel treatment options to improve both patient survival and quality-of-life.

1.1.1 History and presentation of Ewing sarcoma

Dr. James Ewing, a renowned physician and cancer pioneer, first described a case of "diffuse endothelioma of bone" in a young girl in 1921, which would later become known as Ewing sarcoma after his namesake (16, 17). Histological analysis of the tumor revealed a disease distinctly separate from the more widely known osteosarcoma (1, 3, 10, 15, 16). Ewing sarcoma is considered to be a "small, round blue cell tumor" and most often occurs in the long bones or flat bones of patients, such as the femur or pelvis (1, 3, 10, 15, 16). Peak incidence of the disease is seen between 10 and 26-years of age, with a slightly higher prevalence in male versus female patients (1, 3, 10, 15, 16). Current treatment regimens

consist of high-dose chemotherapy, intensive radiation, and local surgical control (1, 3, 10, 15, 16). Although patients who present with localized disease have a ~70% chance of five-year survival, almost 30% of patients present with metastatic disease. Despite decades of research, five-year survival rates remain at ~20% (1, 3, 10, 15, 16). Chances of relapse after initial treatment are high as well and these patients also face poor survival outcomes (1, 3, 10, 15, 16). Due to the highly aggressive treatment regimen used, concern of long-term toxicity issues is high for patients, including cardiotoxicity, secondary tumors, infertility, and overall decreased quality of life (1, 3, 10, 12, 15, 16). Although the use of adjuvant chemotherapy, which began a half century ago, resulted in marked outcome improvements, our inability to shift overall survival curves over the past few decades highlights the importance of better understanding Ewing sarcoma disease biology to develop more effective, less toxic novel treatment options.

1.1.2 Ewing sarcoma-associated chromosomal translocations

In the early 1980s, molecular analysis of Ewing sarcoma patient samples revealed a chromosomal translocation, t(11;22)(q24;q12), in a significant number of patient cases (4-7). It took another decade and the advent of polymerase chain reaction (PCR) technology to reveal the genes associated with this translocation: *EWSR1 and FLI1* (2). While FLI1 was determined to belong to the large, highly conserved ETS (E-26 transformation specific) transcription factor family, EWSR1 was a novel gene that would eventually be placed in the FET protein family (FUS, EWSR1, TAF15) (2, 9, 18-26). It is now known that the *EWSR1/FLI1*, or EWS/FLI, translocation occurs in 85-90% of Ewing sarcoma

patients (2, 3, 5, 7, 15). After initial characterization of this translocation, another six additional translocations fusing one of the FET protein family members to an ETS transcription factor were discovered in Ewing sarcoma patient tumors, including EWS/ERG, EWS/FEV, EWS/ETV1, and EWS/ETV4, as well as FUS/ERG and FUS/FEV (Figure 1.1) (26, 27). Along with other Ewing sarcoma-specific clinical features, presence of one of the EWS/ETS translocations in a tumor is typically regarded as sufficient justification for a Ewing sarcoma diagnosis (28, 29). Because of the rarity of FUS/ETS translocations, more uncertainty exists in the field regarding how these alternative translocations should be diagnosed, creating a clinical dilemma for how patients should be treated and stratified in clinical trials (28). As sequencing technologies become more widely available and these rare variants will likely be detected more often, information surrounding how these alternative fusions should be classified is desperately needed.

1.1.2.1 Mechanisms driving chromosomal translocation

Although both solid and liquid tumors contain gene fusions that are the product of chromosomal rearrangements, the mechanisms driving genomic events that result in these fusion proteins have yet to be fully understood. There are several possible mechanisms by which these rearrangements can arise, including chromothripsis and chromoplexy (30-33). Chromothripsis is a phenomenon characterized by chromosomal "shattering," in which hundreds of breakpoints arise in one or two neighboring chromosomes that are then rejoined in an attempt to repair this damage (31, 32). Chromothriptic events are likely the result of replication stress or mitotic errors, such as formation of micronuclei or premature

chromatin compaction (31, 32). Due to the severe, detrimental consequences of chromothripsis on the cell replication and survival, it is unsurprising that chromothripsis has been detected in only a small number of cancer types. In contrast to chromothripsis, chromoplexy has a smaller number of rearrangement events, typically in the range of tens instead of hundreds (30, 31). Chromoplexy also typically involves several chromosomes (30, 31). Recent studies have revealed that the TMRPSS2-ERG fusion found in >50% of prostate cancer cells is likely the result of chromoplectic events (32-34). Evidence of chromoplexy has also been detected in non-small cell lung cancers, head and neck cancer, and melanoma (32-34). Mechanistic studies have revealed that chromoplexy is likely due to DNA damage that may arise during transcription factor binding and transcriptional activation (32-34). A study performed on Ewing sarcoma patient tumors found that chromoplectic events were detected in a portion of samples and that chromoplexy loops could almost always be traced to an EWS/ETS origin rearrangement (30). Direct mechanisms driving these chromoplectic events have yet to be elucidated. Despite implication of chromoplexy driving Ewing sarcoma rearrangements in a fraction of patient tumors, evidence of chromoplectic events were not observed in all samples and further studies characterizing alternative mechanisms of translocation must be performed (30).

1.1.3 Ewing sarcoma cell-of-origin is unknown

Despite classification as a sarcoma, the cell-of-origin of Ewing sarcoma is unknown and remains an area of great contention. Several models have been proposed over the years, including neuroectodermal progenitor and mesenchymal stem cells (35, 36). Small,

neurosecretory granules have been identified in the nucleus of several Ewing sarcoma patient tumor samples, similar to those observed in neuroectodermal cells (36, 37). Upon EWS/FLI knock-down, Ewing sarcoma cells also mimic the histological appearance of other neuroectodermal cells. Despite this, neuroectodermal cells do not tolerate introduction of EWS/FLI, undergoing senescence and eventual cell death, suggesting that this is not the cell of origin (35, 36). To data, mesenchymal stem cells (MSCs) are believed to be the most likely cell of origin (35). MSCs are the only cell type to tolerate the introduction of the EWS/FLI translocation and continue to proliferate (38). Upon EWS/FLI knock-down, the transcriptional profile of Ewing sarcoma cells looks most like that of MSCs (3). Additionally, upon EWS/FLI knock-down in Ewing sarcoma cells, the cells are able to differentiate into adipogenic, neuronal, and osteogenic lineages; the same cell lineages also observed in MSCs (3, 21). Despite a handful of studies published on the MSC model, tolerance of EWS/FLI is not always observed in MSCs and the transcriptional profile of Ewing sarcoma cells with EWS/FLI depletion does not perfectly match that of MSCs (3, 35). The obvious caveat of an unknown cell-of-origin is that there is no cellular model system in which these fusion proteins can be studied in culture or in animal models. As such, laboratory studies focus on patient-derived Ewing sarcoma cell lines as the preferential model system.

1.1.4 Evolution of molecular diagnosis of Ewing sarcoma

In addition to the presence of a FET/ETS fusion, Ewing sarcoma cells typically demonstrate high expression of CD99, a transmembrane protein that is typically associated

with cell surface interactions in the vasculature (3, 39-41). Although the genomes of many pediatric tumors have been shown to be "mutationally-quiet", a fraction of Ewing sarcoma patients do possess additional mutations: The most common secondary mutation, found in 5-20% of patients, occurs in *TP53*, a ubiquitously expressed tumor suppressor gene (3, 42-44). The next most common mutation found in Ewing sarcoma tumors occurs in *STAG2*, an integral protein involved in the cohesin complex, and is found in 10-12% of patients (3, 43-47). Although the mechanisms that make these mutations preferential in Ewing sarcoma cells are actively being investigated, there is evidence to support that patients with *TP53* and *STAG2* mutations have more aggressive disease and poorer outcomes overall (3, 43-47). Because of the rare occurrence of additional mutations, Ewing sarcoma represents an interesting disease paradigm in which one oncoprotein is solely responsible for driving disease.

1.1.5 Targeted therapy for treatment of Ewing sarcoma

Despite advancements in targeted therapeutics, multi-agent chemotherapy and radiation treatment regimens, combined with local surgical control, remain the standard of care for Ewing sarcoma patients (3, 10, 14, 15). Several targeted drugs have been developed for Ewing sarcoma over the years, with mixed and often disappointing clinical results (13, 44, 48-52). One of these inhibitors, YK-4-279, reportedly binds EWS/FLI directly to disrupt a crucial protein-protein interaction required for EWS/FLI function (48, 49). Despite potential promise, an unfeasible dosage regimen (14-day continuous infusion) has prevented it from moving forward quickly in the clinic. Additional inhibitors targeting

EWS/FLI protein interactors, such as LSD1 (lysine-specific demethylase 1) or HDACs (histone deacetylases), are still in development, though inhibition of these targets have had lackluster efficacy in the treatment of other cancers (13, 52). This highlights the need for basic science studies to better understand the biological mechanisms in which FET/ETS fusions drive Ewing sarcomagenesis.

1.2 FET protein family

The FET protein family is a group of ubiquitously expressed putative RNA-binding proteins and is composed of three proteins in humans: FUS, EWS, and TAF15. The protein family was initially discovered in 1992 with the identification of *EWSR1* as one of the genes involved in Ewing sarcoma-associated chromosomal translocations (2, 21). Soon thereafter, FUS (Fused in Sarcoma or Translocated in Liposarcoma (TLS)) was identified in translocations found in human myxoid and round cell liposarcomas, as well as leukemia (21, 53). TAF15 (TATA-binding protein Associated Factor 15), also known as $hTAF_{II}68$, was subsequently discovered as a closely homologous protein that was localized to the nucleus and associated with ribonucleoproteins (24). Further analyses of FET protein family structure and normal protein function will allow a better understanding of their oncogenic involvement in Ewing sarcomagenesis.

1.2.1 FET protein structure and function

Proteins are typically classified into a family based on similar function and/or sequential and structural motifs. The FET protein family is characterized by several structural motifs. The amino-terminal domains of this protein family consist of degenerate hexapeptide repeats (DHRs) and a strong transactivation domain (AD) (3, 21, 24, 53, 54). These DHRs, composed of a serine-tyrosine-glycine-glutamine-glutamine-serine (SYGQQS) sequence, form low-complexity, intrinsically-disordered domains that do not readily form a steady structure (55, 56). As such, structural studies have been largely unsuccessful for this region of the protein family. The carboxyl-terminal domains of the FET proteins are more structured and contain several arginine-glycine-glycine (RGG) boxes, as well as a RNA-recognition motif (RRM) and a zinc-finger domain (ZnF; Figure 1.2) (24, 53, 54, 57).

The FET proteins are primarily localized to the nucleus and functional studies have determined they likely play key roles in several homeostatic processes, including RNA-binding, transport, and mRNA maturation (3, 21, 53). Additionally, all three FET members have been determined to be involved in RNA transcription: The DHRs in the amino-terminal domain resemble those in the carboxyl-terminal domain of RNA polymerase II and all three protein members have been shown to interact with both transcription factor IID (TFIID) and RNA polymerase II (21, 53). Localization studies have also shown that the FET proteins can localize to the promoter regions of genes that are activated downstream (21). Additionally, recent studies have revealed that the IDRs in the amino-terminal domain of FET proteins confer a unique ability for the proteins to undergo liquid-liquid phase separation, a process in which proteins form a highly-concentrated condensate

similar to a liquid droplet (55, 56, 58). It is believed that these phase separated droplets are a normal physiological event that allow multiple proteins to interact in high concentrations to facilitate cellular processes, such as transcriptional regulation and RNA processing (55, 56, 58).

1.2.2 Involvement of FET proteins in disease

Despite clear roles in normal cellular processes, the major focus on the FET protein family lies in their involvement in disease biology. These proteins have been implicated in neurodegenerative diseases, such as amyotrophic lateral sclerosis (ALS), and numerous cancer types (21, 54). For instance, mutations in the FUS protein have been discovered in several cases of familial ALS (54). These mutations lead to changes in the localization of FUS from the nucleus to the cytoplasm, alterations in its capacity to process RNA, and a change in the ability of FUS to undergo the LLPS process (21, 54). These altered properties likely contribute to ALS disease etiology (54-56).

Translocations involving one of the FET protein members have been discovered in a number of cancer types, including acute myeloid and acute lymphoblastic leukemia, chondrosarcoma, liposarcoma, clear cell sarcoma, bone sarcoma, and Ewing sarcoma (21). Although all three members have been identified in cancer-associated translocations, the vast majority of these translocations involve EWS (21). Curiously, these translocations often fuse the amino-terminal activation domain (AD) of a FET protein to the carboxyl-terminal domain of a transcription factor, thus combining a strong AD with a DNA-binding

domain that results in the formation of an oncogenic transcription factor (21). The majority of functional studies on FET-involved translocations have focused on EWS/FLI in the context of Ewing sarcoma and will be discussed below.

1.3 ETS transcription factor family

The ETS protein family is a large group of transcription factors that are highly conserved across species and have been implicated in numerous critical developmental processes (20, 23). ETS1, the founding member of the family, was first discovered as an integration site for the avian retrovirus E26 that eventually results in the development of leukemia in chickens (59). Subsequently, many members of the ETS family have also been implicated in human cancers as well (20, 23).

1.3.1 ETS protein structure and subfamily classification

In humans, the ETS protein family consists of 28 members that all share a highly-conserved winged helix-turn-helix ETS DNA-binding domain (DBD) (19, 60). The ETS DBD is generally regarded as an ~85 amino acid sequence. The third alpha-helix of the ETS DBD makes direct contact with DNA in the major groove of the double-helical structure (19, 23, 60, 61) . Common structural features of the group include the ETS DBD in the carboxyl-terminus and a helix-loop-helix pointed (PNT) domain in the amino-terminus (34, 61). The PNT domain is a member of the sterile alpha motif (SAM) family that facilitates a variety of protein-protein interactions, including self-association (Figure 1.3) (34, 61). Proteins are

further classified into subfamilies based on sequence homology both inside and outside of the ETS DBD (19, 20, 22, 23). Although the DBDs share high amounts of sequence homology, protein structure outside of this region can vary significantly. For example, the DBD of ERG and FLI are highly similar with 98% sequence homology. Outside of the DBD, ERG and FLI share ~70% sequence homology and are grouped together into the ERG subfamily. Interestingly, the ETS factors that are involved in Ewing sarcoma translocations are found in only two subfamilies: the ERG subfamily, which contains ERG, FLI, and FEV, and the PEA3 subfamily, which contains both ETV1 and ETV4 (22, 23).

The function of the ETS family is closely tied with the DBD found in the carboxyl-terminus: The high-affinity (HA) consensus binding sequence for the ETS DBD contains a GGA(A/T) core and is flanked by several base pairs that vary depending on family member (18, 19, 60, 62, 63). Although binding to similar HA sites could potentially lead to complete functional redundancy, it has been determined that this is not the case: Flanking sequences surrounding the DBD vary for each ETS member, governing each protein's DNA-binding capability and protein-protein interactions with other transcriptional and epigenetic regulators (64-66). For example, several members of the ETS family have demonstrated strong auto-inhibitory properties that govern DNA-binding affinity (67-71). Mechanisms of autoinhibition often involve structural features immediately flanking the ETS DBD directly interacting with the DBD to modulate DNA binding (67-71). Although strong autoinhibitory activity has been observed for other ETS factors, such as members of the PEA3 subfamily, ERG has moderate autoinhibitory activity and FLI has weak

autoinhibitory activity governing DNA-binding in vitro (66, 70, 72-74). These different structural factors that affect function contribute to the wide-range of biological functions observed for this family, as discussed below.

1.3.2 Biological function of Ewing sarcoma-associated ETS factors

The ETS family of transcription factors are expressed across many cell types and are involved in a number of biological processes. Indeed, analysis of numerous tissue types has shown that several ETS factors are expressed in any one-cell type (23). The protein family has been implicated in many important processes, such as cellular proliferation, differentiation, apoptosis, tissue remodeling, and angiogenesis (23, 34). Specifically, FLI has been shown to be a critical regulator of hematopoietic development, including erythrocytes, T-cells, B-cells, and megakaryocytes. FLI has also been implicated in both vasculogenesis and angiogenesis (23, 63). Similarly, ERG is also a key regulator in hematopoiesis and has been implicated in vasculogenesis and vascular inflammation (61). Both FLI and ERG are crucial for embryonic development and deletion of either gene in mice results in embryonic lethality (23, 63).

1.3.3 ETS factors involved in cancer

As the ETS protein family is implicated in a variety of crucial developmental and maintenance roles in numerous tissue types, it is unsurprising that aberrant regulation of these proteins can contribute to disease. ETS factors have been implicated in solid tumors, such as breast, lung, brain, and ovarian cancer, as well as various hematological

malignancies (23). Specifically, mechanisms that result in ETS factor involvement in tumorigenesis include overexpression, amplification, loss- or gain-of-function mutations, as well as chromosomal translocations (23). Most famously, prostate cancer and Ewing sarcoma tumors have chromosomal translocations involving one of a few ETS family members (23, 61, 75). Fusions involving the androgen-receptor *TMPRSS2* and *ERG*, both located on 22q12, have been identified in over 50% of prostate cancers (33, 34, 61, 75). Alternative *TMPRSS2* fusions with *ETV1*, *ETV4*, *ETV5*, and *FEV* also exist in an additional group of prostate cancer patients (33, 34, 61, 75). As mentioned above, Ewing sarcoma is characterized by a chromosomal translocation involving one of the FET protein family members and an ETS transcription factor (2, 8, 18). EWS/FLI is the most common chromosomal translocation found in 85-90% of patients, EWS/ERG is the second most common found in 10-15% of patients, and the remainder of these fusions are found in approximately 5% of patients (3, 15). The translocation of these FET and ETS proteins results in an interesting combination of protein properties that can be attributed to either native protein, as well as novel functions unique to the fusion protein.

1.4 Ewing sarcoma-associated FET/ETS translocations

The presence one of the FET/ETS translocations is the defining factor of Ewing sarcoma tumors and it has been determined that the encoded fusion protein functions as an oncogenic transcription factor that drives disease (1-3, 15, 45, 65, 76, 77). As the fusion of EWS to FLI is the most common translocation in Ewing sarcoma tumors, almost the

entirety of functional studies annotating mechanisms of oncogenesis have focused on EWS/FLI.

1.4.1 Functions of EWS/FLI

EWS/FLI is widely-accepted to be the driver protein in Ewing sarcomagenesis (1, 3, 5, 15, 26, 77). EWS/FLI acts as an oncogenic transcription factor, creating a molecular landscape tolerant of aberrant regulation that leads to tumorigenesis. Functional studies of the EWS/FLI fusion protein have identified several properties as crucial for this process, including DNA binding, transcriptional regulation, pioneer factor activity, and modulation of chromatin architecture. Each of these functions will be discussed in detail below.

1.4.1.1 DNA-binding

As the EWS/FLI fusion protein retains the intact DNA-binding domain of FLI, the ability to bind DNA is also retained (2, 8, 18, 19, 62, 78, 79). Binding studies have revealed several thousand genomic binding sites for EWS/FLI, including promoter and enhancer sequences located both proximally and distally to associated genes (3, 21, 77, 80, 81). A discussion of preferential DNA binding sequences and mechanisms governing DNA-binding enables a better understanding of EWS/FLI function.

The native FLI HA site, ACCGGAAGTG, is one of the prominent sequences EWS/FLI binds to in the genome (21, 80). EWS/FLI, like native FLI, can bind sequences similar to the HA site that contain alternative base pairs flanking the “GGA” core; these are referred

to as HA site-like (80). In addition to the HA site, the fusion of EWS to FLI confers the novel ability to bind highly repetitive GGAA sequences, termed microsatellites, to the fusion protein (77, 81-87). These GGAA-microsatellites are typically located in non-coding regions of the genome that were often referred to as "junk DNA" in the past (21, 82). In fact, the ability of EWS/FLI to bind and regulate at these regions has been determined to be a crucial factor necessary to facilitate oncogenesis (77, 81-87). GGAA-microsatellites can vary in length based on tissue and cell type, as well as location in the genome. Sequencing studies have provided evidence to further support the criticality of these microsatellites in Ewing sarcomagenesis and reveal biological differences that contribute to observed disparities in racial and ethnic backgrounds of those affected by the disease: Ewing sarcoma is almost ten-times more likely to develop in people of European-descent compared to Asian or African ethnicity (85). When patient and control samples were analyzed for GGAA-microsatellite length, it was determined that people of European-descent had significantly shorter microsatellites in the *NR0B1* promoter and that the presence of specific "sweet spot"-length microsatellites contributes to EWS/FLI oncogenic activity and disease formation (85, 86).

1.4.1.2 Gene regulation

The ability of EWS/FLI to drive oncogenesis is closely tied with its ability to dysregulate the expression of several thousands of genes. Transcriptomic studies of Ewing sarcoma cells have shown that EWS/FLI is capable of activating and repressing genes through both

direct and indirect mechanisms (3, 15, 21, 77). An analysis of regulated pathways and mechanisms of regulation reveals much about the oncogenic function of EWS/FLI.

Although native FLI is a transcription factor capable of gene regulation, it is the translocation of the amino-terminus of EWS that allows EWS/FLI to act as a potent transcriptional activator: Indeed, when luciferase assays were used to measure the transcriptional activation capacity of different EWS/FLI mutants, it was determined that FLI alone was a weak transcriptional activator and that the majority of EWS was required for strong transcriptional activity (78, 79). Recent studies have also demonstrated that the intrinsically-disordered SYGQ-regions and tyrosine residues of EWS are required for transcriptional activity of EWS/FLI in Ewing sarcoma cells as well (38, 80). EWS/FLI-activated genes are typically associated with proliferation and cell-survival pathways (77).

Though EWS/FLI represses slightly more genes than it activates, mechanisms of repression are often more complex and not fully understood. It is believed that EWS/FLI-associated gene repression is achieved through both direct and indirect or downstream mechanisms: EWS/FLI is capable of binding to promoter and/or enhancer elements required for gene activation to block transcriptional machinery from transcribing these genes to directly repress these targets (77). EWS/FLI may indirectly repress genes by interacting with and redirecting activating proteins to prevent their typical transcriptional activation functions. While EWS/FLI-activated genes are more likely to have an ETS HA site or GGAA-microsatellite in their promoter and/or enhancer, repressed genes are typically devoid of

either sequence (21, 77). This further strengths the theory that EWS/FLI-mediated repression largely occurs through indirect mechanisms. Additionally, EWS/FLI actively recruits epigenetic modifiers, such as demethylases or deacetylases, to modify DNA structure to inhibit gene transcription (52, 88-91). EWS/FLI also upregulates repressive proteins, such as DAX (transcribed by the *NR0B1* gene), that will alter gene expression of downstream targets (92). As Ewing sarcoma cells are believed to derive from a stem or progenitor cell origin and many cancer cells work to actively retain this "stemness", it is unsurprising that repressed genes are typically associated with developmental and differentiation processes (3, 21, 77).

Another important property that contributes to EWS/FLI's ability to drive transcriptional regulation is its apparent ability to undergo phase separation. EWS/FLI has been shown to interact with numerous protein partners to activate or repress transcription. Protein partners include wild-type EWS, RING1B, BARD1, NKX2.2, LSD1 and the BAF complex (38, 52, 65, 88, 92-97). Recently, it was demonstrated that EWS/FLI likely interacts with these proteins in transcriptional "hubs": EWS/FLI is known to self-associate through the EWS-domain to form multimers. In high concentrations, these EWS/FLI molecules reportedly undergo phase separation and can interact with crucial protein partners here to mediate transcription (Figure 1.4) (57). Further strengthening this theory, disruption of the SYGQ-residues in the EWS domain prevents phase separation and a loss of transcriptional control is observed (38).

1.4.1.3 Pioneer factor activity

Chromatin compaction is a mechanism used to govern transcriptional regulation across all cell types (98-102). DNA is tightly wound around histone proteins to form a nucleosome, which can then be modified through epigenetic changes to facilitate further DNA compaction until the chromatin exists in a form that is inaccessible for binding and interaction with most proteins (98-101). As a large majority of chromatin exists in this tightly packed "heterochromatin" form, it is crucial that certain proteins are able to reverse this compaction process to access these "closed" regions of chromatin to allow for transcriptional activation necessary for cell differentiation, replication, and normal homeostatic processes (98-102). Such proteins are referred to as "pioneer factors" (100-103). Although previous studies on pioneer factors determined that these proteins are able to interact with and bind to nucleosomal-wrapped DNA, only a small portion of the encoded proteins in a cell are able to do so (103, 104). Certain structural features of a transcription factor indicate that they are capable of binding to nucleosomal-wrapped DNA. For example, the presence of a short alpha-helix in the DNA-binding domain of a protein that comes into contact with the major groove of DNA is indicative that this protein may be able to bind closed chromatin (103, 104). To note, this type of DNA-binding domain is present in the ETS transcription factor family. Studies performed on ETS factors PU.1, ELF1, and ELF2 revealed that these ETS factors can indeed bind DNA in the presence of nucleosomes, thus, indicating that the larger ETS family of proteins may act as pioneer factors as well (103, 104). Although pioneer factor studies have yet to be performed on wild-type FLI protein, EWS/FLI was recently shown to demonstrate pioneer factor

activity in a mesenchymal stem cell model (105). As GGAA-microsatellites are often found in heterochromatic regions, the reported pioneer factor activity of EWS/FLI is likely crucial in its ability to regulate transcription of these targets. Although the preliminary results studied in a MSC model are useful, these results have yet to be replicated in a Ewing sarcoma cellular model.

1.4.1.4 Modulation of chromatin architecture

DNA-bound nucleosomes are condensed into chromatin, which can be further compacted into topologically associated domains (TADs) (98, 99, 104, 106). TADs are semi-insulated regions where interactions between DNA, transcription factors, and additional modifiers can occur to regulate transcription (99, 107). For example, DNA containing a regulatory element for a downstream target, such as an enhancer, can be looped to its target inside of the TAD. TADs can be further compacted into compartments: Compartments are generally regarded to contain "open" or active chromatin, where gene transcription can occur, or "closed" chromatin, where genes are typically repressed (99, 107). Numerous processes regulate chromatin compaction and very strongly impact transcriptional regulation. Certain proteins and complexes, such as CTCF, YY1, and cohesin, have been demonstrated to play crucial roles in maintenance of these compartments (47, 106). For example, a previous study revealed that removal of one CTCF-binding site can be sufficient to disrupt a TAD boundary (99). Additional mechanisms governing chromatin state and transcriptional regulation are those of epigenetic modifications: Chromatin can be acetylated, methylated,

phosphorylated, etc. as a means to control the "open" or "closed" state to further regulate gene transcription (108).

In addition to its apparent role of a pioneer factor, interaction between EWS/FLI and various proteins governing chromatin architecture have been detected (46, 47, 91, 95). For example, EWS/FLI actively recruits the BAF complex, an ATP-dependent chromatin remodeler, that regulates genomic architecture and DNA accessibility (38). If this recruitment and interaction is disrupted through mutation of the EWS-domain of EWS/FLI, BAF is no longer recruited to GGAA-microsatellites and expression of many typical EWS/FLI-regulated genes are lost (38). EWS/FLI also interacts with LSD1, a lysine-specific demethylase protein, that is normally found in the NuRD (Nucleosome Remodeling and Deactylatse) repressive complex (52, 88). Despite its typical role in transcriptional repression through histone modification, LSD1 has been linked to the activation of gene targets in Ewing sarcoma cells (52, 80, 90). Similarly, EWS/FLI also recruits and interacts with RING1B, a member normally found in the polycomb repressive complex, to participate in gene activation (95). These findings indicate that EWS/FLI can both recruit and redirect protein function to alter chromatin architecture and drive its transcriptional regulatory program.

1.4.2 Alternative FET/ETS Ewing sarcoma translocations

Despite extensive studies on EWS/FLI, very few studies on alternative FET/ETS fusion proteins have been completed. It is known that ETS factors have highly conserved DNA-

binding domains and that this may imply similar function in the case of these translocations, but without studies investigating these ideas, it is unclear if this is a correct or incorrect assumption. Additionally, EWS and FUS are both members of the FET protein family. Involvement of FUS in the etiology of neurodegenerative diseases has revealed properties of this protein, such as its ability to form aggregates in the cell that disrupt normal molecular processes. Despite the assumption that EWS would have similar properties, novel studies indicate that the two proteins may function in similar, but not identical capacities. Indicating as much, there is uncertainty in the clinical diagnosis and treatment decisions made for those who present with a tumor harboring a FUS/ETS fusion (28). As a result of the lack of studies exploring alternative FET/ETS fusion protein function, it is unclear if information on the function of EWS/FLI should be generalized to these rare, alternative fusions. Because these FET/ETS translocations involve numerous members of two specific protein families, it is likely that novel additional FET/ETS translocations will also be detected in Ewing sarcoma tumors in the future, further highlighting the importance of comparative studies.

1.5 Rationale and Approach

1.5.1 Rationale

For both solid and liquid tumors, targeted inhibitors have greatly improved the specificity of cancer treatment options, while also decreasing off-target side effects. For example, the direct inhibition of alternative gene fusions, such as BCR-ABL in chronic myeloid

leukemia, have revolutionized cancer care and greatly improved disease outcomes for these patients. Targeting crucial protein interactors or cell markers has also proven to be a successful tactic in inhibition of cancer cell proliferation for various cancers. Despite these therapeutic successes, such triumphs have yet to be declared for patients with Ewing sarcoma tumors. In general, transcription factors have been difficult to target, as the convex structure of DNA-binding domains are not easily bound by small molecules (3, 15). Additionally, the intrinsic disorder of the amino-terminal domain of FET proteins completely lack any definitive structure, making it incredibly difficult to target this domain with a small molecule (3, 15). As such, Ewing sarcoma patients are faced with a chemotherapy regimen similar to original treatment regimens introduced in the 1960s. These therapeutic cassettes have variable responses and potential long-term toxicity concerns as well, highlighting the desperate need for treatment advances from those suffering with this aggressive tumor type. A thorough biological understanding of EWS/FLI- and FET/ETS-driven oncogenic processes is required to better understand the disease and pinpoint feasible drug targets for future therapeutic development.

Rationale: *Ewing sarcoma is a disease defined by the presence of a FET/ETS chromosomal translocation. Despite extensive studies on EWS/FLI, the most common translocation found in patients, it is unclear if oncogenic properties of the protein translate to other FET/ETS translocations. If this information is indeed generalizable, determination of protein function may identify targetable vulnerabilities of the fusion protein(s) to create more efficacious targeted therapies.*

1.5.2 Hypothesis and Approach

The primary objective of this work is to annotate the function of FET and ETS domain contributions to fusion oncoprotein activity in a Ewing sarcoma cell line. We must first identify whether FET/ETS proteins function in a similar manner or if they differ drastically in molecular function. This information will allow us to know if structural and mechanistic studies that focus on EWS/FLI are potentially generalizable to the larger FET/ETS translocation group. Additionally, we hope to identify structural features of the EWS/FLI protein that are critical for oncogenic function. This work is crucial for our ability to efficiently translate basic science studies into clinically relevant information that can be used to improve Ewing sarcoma survival outcomes.

Chapter 2*: Identification of a novel FUS/ETV4 fusion and comparative analysis with other Ewing sarcoma fusion proteins*

Hypothesis 1: Due to the high sequence homology and similar reported functions between members of the FET and ETS protein families, we hypothesize that the broad FET/ETS translocation group will exhibit similar biological function(s) in Ewing sarcoma cells.

> **Approach:** Compare the global DNA-binding and transcriptional regulatory patterns of reported Ewing sarcoma-associated FET/ETS fusion proteins using CUT&Tag genomic localization and RNA-sequencing technologies.

Hypothesis 2: We predict that similarities in the function of FET/ETS fusion proteins will extend to novel translocations involving FET and ETS family members in Ewing sarcoma patient tumors, and that these similarities will justify the diagnostic inclusion of novel non-EWS/FLI FET/ETS fusions as Ewing sarcoma tumors.

Approach: Compare the genomic localization and transcriptional regulation profiles of a novel FET/ETS fusion, FUS/ETV4, with that of previously identified FET/ETS fusions.

Chapter 3*: The FLI portion of EWS/FLI contributes a transcriptional regulatory function that is distinct and separable from its DNA-binding function in Ewing sarcoma*

Hypothesis: As implicated in the function of other ETS factors, we hypothesize that protein domains immediately flanking the DNA-binding domain of FLI modulate the ability of EWS/FLI to function as an oncogenic transcription factor.

Approach 1: Identify crucial regions of FLI required for EWS/FLI transcriptional activation function utilizing luciferase reporter assay screens through creation of mutant EWS/FLI protein constructs containing varying regions of the FLI domain.

Approach 2: Determine function(s) the FLI domain contributes to EWS/FLI through studies of EWS/FLI mutant constructs utilizing in vitro and in vivo study methods, including protein:DNA binding studies, genomic localization studies,

RNA-sequencing, ATAC-sequencing, and transformation assays in a relevant Ewing sarcoma cellular model.

These hypotheses and approaches are discussed in detail in their respective chapters. The results from Chapter 2 have been submitted to *Molecular Cancer Research* for publication. The results from Chapter 3 have been peer-reviewed and accepted for publication in *Oncogene*.

1.6 References

1. Cripe TP. Ewing sarcoma: an eponym window to history. Sarcoma. 2011;2011:457532.

2. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature. 1992;359(6391):162-5.

3. Grunewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H, et al. Ewing sarcoma. Nat Rev Dis Primers. 2018;4(1):5.

4. Aurias A, Rimbaut C, Buffe D, Zucker JM, Mazabraud A. Translocation involving chromosome 22 in Ewing's sarcoma. A cytogenetic study of four fresh tumors. Cancer Genet Cytogenet. 1984;12(1):21-5.

5. Turc-Carel C, Aurias A, Mugneret F, Lizard S, Sidaner I, Volk C, et al. Chromosomes in Ewing's sarcoma. I. An evaluation of 85 cases of remarkable consistency of t(11;22)(q24;q12). Cancer Genet Cytogenet. 1988;32(2):229-38.

6. Aurias A, Rimbaut C, Buffe D, Dubousset J, Mazabraud A. [Translocation of chromosome 22 in Ewing's sarcoma]. C R Seances Acad Sci III. 1983;296(23):1105-7.

7. Turc-Carel C, Philip I, Berger MP, Philip T, Lenoir GM. Chromosome study of Ewing's sarcoma (ES) cell lines. Consistency of a reciprocal translocation t(11;22)(q24;q12). Cancer Genet Cytogenet. 1984;12(1):1-19.

8. May WA, Gishizky ML, Lessnick SL, Lunsford LB, Lewis BC, Delattre O, et al. Ewing sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation. Proc Natl Acad Sci U S A. 1993;90(12):5752-6.

9. Zucman J, Melot T, Desmaze C, Ghysdael J, Plougastel B, Peter M, et al. Combinatorial generation of variable fusion proteins in the Ewing family of tumours. EMBO J. 1993;12(12):4481-7.

10. Pishas KI, Lessnick SL. Recent advances in targeted therapy for Ewing sarcoma. F1000Res. 2016;5.

11. Jain S, Kapoor G. Chemotherapy in Ewing's sarcoma. Indian J Orthop. 2010;44(4):369-77.

12. Longhi A, Ferrari S, Tamburini A, Luksch R, Fagioli F, Bacci G, et al. Late effects of chemotherapy and radiotherapy in osteosarcoma and Ewing sarcoma patients: the Italian Sarcoma Group Experience (1983-2006). Cancer. 2012;118(20):5050-9.

13. Lawlor ER, Sorensen PH. Twenty Years on: What Do We Really Know about Ewing Sarcoma and What Is the Path Forward? Crit Rev Oncog. 2015;20(3-4):155-71.

14. Dyson KA, Stover BD, Grippin A, Mendez-Gomez HR, Lagmay J, Mitchell DA, et al. Emerging trends in immunotherapy for pediatric sarcomas. J Hematol Oncol. 2019;12(1):78.

15. Riggi N, Suva ML, Stamenkovic I. Ewing's Sarcoma. N Engl J Med. 2021;384(2):154-64.

16. Ewing J. Classics in oncology. Diffuse endothelioma of bone. James Ewing. Proceedings of the New York Pathological Society, 1921. CA Cancer J Clin. 1972;22(2):95-8.

17. Ewing J. The Classic: Diffuse endothelioma of bone. Proceedings of the New York Pathological Society. 1921;12:17. Clin Orthop Relat Res. 2006;450:25-7.

18. Bailly RA, Bosselut R, Zucman J, Cormier F, Delattre O, Roussel M, et al. DNA-binding and transcriptional activation properties of the EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing sarcoma. Mol Cell Biol. 1994;14(5):3230-41.

19. Liang H, Olejniczak ET, Mao X, Nettesheim DG, Yu L, Thompson CB, et al. The secondary structure of the ets domain of human Fli-1 resembles that of the helix-turn-helix DNA-binding motif of the Escherichia coli catabolite gene activator protein. Proc Natl Acad Sci U S A. 1994;91(24):11655-9.

20. Laudet V, Hanni C, Stehelin D, Duterque-Coquillaud M. Molecular phylogeny of the ETS gene family. Oncogene. 1999;18(6):1351-9.

21. Kovar H. Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. Sarcoma. 2011;2011:837474.

22. Li Y, Luo H, Liu T, Zacksenhaus E, Ben-David Y. The ets transcription factor Fli-1 in development, cancer and disease. Oncogene. 2015;34(16):2022-31.

23. Sizemore GM, Pitarresi JR, Balakrishnan S, Ostrowski MC. The ETS family of oncogenic transcription factors in solid tumours. Nat Rev Cancer. 2017;17(6):337-51.

24. Bertolotti A, Lutz Y, Heard DJ, Chambon P, Tora L. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. EMBO J. 1996;15(18):5022-31.

25. Hackl W, Luhrmann R. Molecular cloning and subcellular localisation of the snRNP-associated protein 69KD, a structural homologue of the proto-oncoproteins TLS and EWS with RNA and DNA-binding properties. J Mol Biol. 1996;264(5):843-51.

26. Sankar S, Lessnick SL. Promiscuous partnerships in Ewing's sarcoma. Cancer Genet. 2011;204(7):351-65.

27. Ng TL, O'Sullivan MJ, Pallen CJ, Hayes M, Clarkson PW, Winstanley M, et al. Ewing sarcoma with novel translocation t(2;16) producing an in-frame fusion of FUS and FEV. J Mol Diagn. 2007;9(4):459-63.

28. Kinnaman MD, Zhu C, Weiser DA, Mohiuddin S, Hingorani P, Roth M, et al. Survey of Paediatric Oncologists and Pathologists regarding Their Views and Experiences with Variant Translocations in Ewing and Ewing-Like Sarcoma: A Report of the Children's Oncology Group. Sarcoma. 2020;2020:3498549.

29. Machado I, Noguera R, Pellin A, Lopez-Guerrero JA, Piqueras M, Navarro S, et al. Molecular diagnosis of Ewing sarcoma family of tumors: a comparative analysis of 560 cases with FISH and RT-PCR. Diagn Mol Pathol. 2009;18(4):189-99.

30. Anderson ND, de Borja R, Young MD, Fuligni F, Rosic A, Roberts ND, et al. Rearrangement bursts generate canonical gene fusions in bone and soft tissue tumors. Science. 2018;361(6405).

31. Ly P, Brunner SF, Shoshani O, Kim DH, Lan W, Pyntikova T, et al. Chromosome segregation errors generate a diverse spectrum of simple and complex genomic rearrangements. Nat Genet. 2019;51(4):705-15.

32. Shen MM. Chromoplexy: a new category of complex rearrangements in the cancer genome. Cancer Cell. 2013;23(5):567-9.

33. Ballas LK, Hu BR, Quinn DI. Chromoplexy and hypoxic microenvironment drives prostate cancer. Lancet Oncol. 2014;15(13):1419-21.

34. Gasi Tandefelt D, Boormans J, Hermans K, Trapman J. ETS fusion genes in prostate cancer. Endocr Relat Cancer. 2014;21(3):R143-52.

35. Lin PP, Wang Y, Lozano G. Mesenchymal Stem Cells and the Origin of Ewing's Sarcoma. Sarcoma. 2011;2011.

36. Tu J, Huo Z, Gingold J, Zhao R, Shen J, Lee DF. The Histogenesis of Ewing Sarcoma. Cancer Rep Rev. 2017;1(2).

37. Suh CH, Ordonez NG, Hicks J, Mackay B. Ultrastructure of the Ewing's sarcoma family of tumors. Ultrastruct Pathol. 2002;26(2):67-76.

38. Boulay G, Sandoval GJ, Riggi N, Iyer S, Buisson R, Naigles B, et al. Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. Cell. 2017;171(1):163-78 e19.

39. Rocchi A, Manara MC, Sciandra M, Zambelli D, Nardi F, Nicoletti G, et al. CD99 inhibits neural differentiation of human Ewing sarcoma cells and thereby contributes to oncogenesis. J Clin Invest. 2010;120(3):668-80.

40. Louati S, Senhaji N, Chbani L, Bennis S. EWSR1 Rearrangement and CD99 Expression as Diagnostic Biomarkers for Ewing/PNET Sarcomas in a Moroccan Population. Disease Markers. 2018;2018:7971019.

41. Baldauf MC, Orth MF, Dallmayer M, Marchetto A, Gerke JS, Rubio RA, et al. Robust diagnosis of Ewing sarcoma by immunohistochemical detection of super-enhancer-driven EWSR1-ETS targets. Oncotarget. 2018;9(2):1587-601.

42. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.

43. Brohl AS, Solomon DA, Chang W, Wang J, Song Y, Sindiri S, et al. The genomic landscape of the Ewing Sarcoma family of tumors reveals recurrent STAG2 mutation. PLoS Genet. 2014;10(7):e1004475.

44. Pishas KI, Drenberg CD, Taslim C, Theisen ER, Johnson KM, Saund RS, et al. Therapeutic Targeting of KDM1A/LSD1 in Ewing Sarcoma with SP-2509 Engages the Endoplasmic Reticulum Stress Response. Mol Cancer Ther. 2018;17(9):1902-16.

45. Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, Calicchio ML, et al. The genomic landscape of pediatric Ewing sarcoma. Cancer Discov. 2014;4(11):1326-41.

46. Tirode F, Surdez D, Ma X, Parker M, Le Deley MC, Bahrami A, et al. Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. Cancer Discov. 2014;4(11):1342-53.

47. Surdez D, Zaidi S, Grossetete S, Laud-Duval K, Ferre AS, Mous L, et al. STAG2 mutations alter CTCF-anchored loop extrusion, reduce cis-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma. Cancer Cell. 2021.

48. Barber-Rotenberg JS, Selvanathan SP, Kong Y, Erkizan HV, Snyder TM, Hong SP, et al. Single enantiomer of YK-4-279 demonstrates specificity in targeting the oncogene EWS-FLI1. Oncotarget. 2012;3(2):172-82.

49. Minas TZ, Han J, Javaheri T, Hong SH, Schlederer M, Saygideger-Kont Y, et al. YK-4-279 effectively antagonizes EWS-FLI1 induced leukemia in a transgenic mouse model. Oncotarget. 2015;6(35):37678-94.

50. Loganathan SN, Tang N, Fleming JT, Ma Y, Guo Y, Borinstein SC, et al. BET bromodomain inhibitors suppress EWS-FLI1-dependent transcription and the IGF1 autocrine mechanism in Ewing sarcoma. Oncotarget. 2016;7(28):43504-17.

51. Gollavilli PN, Pawar A, Wilder-Romans K, Natesan R, Engelke CG, Dommeti VL, et al. EWS/ETS-Driven Ewing Sarcoma Requires BET Bromodomain Proteins. Cancer Res. 2018;78(16):4760-73.

52. Sankar S, Theisen ER, Bearss J, Mulvihill T, Hoffman LM, Sorna V, et al. Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth. Clin Cancer Res. 2014;20(17):4584-97.

53. Iko Y, Kodama TS, Kasai N, Oyama T, Morita EH, Muto T, et al. Domain architectures and characterization of an RNA-binding protein, TLS. J Biol Chem. 2004;279(43):44834-40.

54. Tan AY, Manley JL. TLS/FUS: a protein in cancer and ALS. Cell Cycle. 2012;11(18):3349-50.

55. Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovsky A, Poser I, et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science. 2018;360(6391):918-21.

56. Lin Y, Protter DS, Rosen MK, Parker R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol Cell. 2015;60(2):208-19.

57. Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science. 2018;361(6400).

58. Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell. 2018;175(7):1842-55 e16.

59. Lautenberger JA, Papas TS. Inversion of a chicken ets-1 proto-oncogene segment in avian leukemia virus E26. J Virol. 1993;67(1):610-2.

60. Szymczyna BR, Arrowsmith CH. DNA binding specificity studies of four ETS proteins support an indirect read-out mechanism of protein-DNA recognition. J Biol Chem. 2000;275(37):28363-70.

61. Adamo P, Ladomery MR. The oncogene ERG: a key factor in prostate cancer. Oncogene. 2016;35(4):403-14.

62. Mao X, Miesfeldt S, Yang H, Leiden JM, Thompson CB. The FLI-1 and chimeric EWS-FLI-1 oncoproteins display similar DNA binding specificities. J Biol Chem. 1994;269(27):18216-22.

63. Truong AH, Ben-David Y. The role of Fli-1 in normal cell function and malignant transformation. Oncogene. 2000;19(55):6482-9.

64. Ibarra IL, Hollmann NM, Klaus B, Augsten S, Velten B, Hennig J, et al. Mechanistic insights into transcription factor cooperativity and its impact on protein-phenotype interactions. Nat Commun. 2020;11(1):124.

65. Erkizan HV, Uversky VN, Toretsky JA. Oncogenic partnerships: EWS-FLI1 protein interactions initiate key pathways of Ewing's sarcoma. Clin Cancer Res. 2010;16(16):4077-83.

66. Hollenhorst PC, McIntosh LP, Graves BJ. Genomic and biochemical insights into the specificity of ETS transcription factors. Annu Rev Biochem. 2011;80:437-71.

67. Gu TL, Goetz TL, Graves BJ, Speck NA. Auto-inhibition and partner proteins, core-binding factor beta (CBFbeta) and Ets-1, modulate DNA binding by CBFalpha2 (AML1). Mol Cell Biol. 2000;20(1):91-103.

68. Goetz TL, Gu TL, Speck NA, Graves BJ. Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor alpha2. Mol Cell Biol. 2000;20(1):81-90.

69. De S, Okon M, Graves BJ, McIntosh LP. Autoinhibition of ETV6 DNA Binding Is Established by the Stability of Its Inhibitory Helix. J Mol Biol. 2016;428(8):1515-30.

70. Currie SL, Lau DKW, Doane JJ, Whitby FG, Okon M, McIntosh LP, et al. Structured and disordered regions cooperatively mediate DNA-binding autoinhibition of ETS factors ETV1, ETV4 and ETV5. Nucleic Acids Res. 2017;45(5):2223-41.

71. Currie SL, Warner SL, Vankayalapati H, Liu X, Sharma S, Bearss DJ, et al. Development of High-Throughput Screening Assays for Inhibitors of ETS Transcription Factors. SLAS Discov. 2019;24(1):77-85.

72. De S, Chan AC, Coyne HJ, 3rd, Bhachech N, Hermsdorf U, Okon M, et al. Steric mechanism of auto-inhibitory regulation of specific and non-specific DNA binding by the ETS transcriptional repressor ETV6. J Mol Biol. 2014;426(7):1390-406.

73. Regan MC, Horanyi PS, Pryor EE, Jr., Sarver JL, Cafiso DS, Bushweller JH. Structural and dynamic studies of the transcription factor ERG reveal DNA binding is allosterically autoinhibited. Proc Natl Acad Sci U S A. 2013;110(33):13374-9.

74. Sharma R, Gangwar SP, Saxena AK. Comparative structure analysis of the ETSi domain of ERG3 and its complex with the E74 promoter DNA sequence. Acta Crystallogr F Struct Biol Commun. 2018;74(Pt 10):656-63.

75. Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia. 2008;10(2):177-88.

76. Hajdu SI. The enigma of Ewing's sarcoma. Ann Clin Lab Sci. 2006;36(1):108-10.

77. Kovar H. Downstream EWS/FLI1 - upstream Ewing's sarcoma. Genome Med. 2010;2(1):8.

78. May WA, Lessnick SL, Braun BS, Klemsz M, Lewis BC, Lunsford LB, et al. The Ewing's sarcoma EWS/FLI-1 fusion gene encodes a more potent transcriptional activator and is a more powerful transforming gene than FLI-1. Mol Cell Biol. 1993;13(12):7393-8.

79. Lessnick SL, Braun BS, Denny CT, May WA. Multiple domains mediate transformation by the Ewing's sarcoma EWS/FLI-1 fusion gene. Oncogene. 1995;10(3):423-31.

80. Theisen ER, Miller KR, Showpnil IA, Taslim C, Pishas KI, Lessnick SL. Transcriptomic analysis functionally maps the intrinsically disordered domain of EWS/FLI and reveals novel transcriptional dependencies for oncogenesis. Genes Cancer. 2019;10(1-2):21-38.

81. Machiela MJ, Grunewald TGP, Surdez D, Reynaud S, Mirabeau O, Karlins E, et al. Genome-wide association study identifies multiple new loci associated with Ewing sarcoma susceptibility. Nat Commun. 2018;9(1):3184.

82. Gangwal K, Lessnick SL. Microsatellites are EWS/FLI response elements: genomic "junk" is EWS/FLI's treasure. Cell Cycle. 2008;7(20):3127-32.

83. Gangwal K, Sankar S, Hollenhorst PC, Kinsey M, Haroldsen SC, Shah AA, et al. Microsatellites as EWS/FLI response elements in Ewing's sarcoma. Proc Natl Acad Sci U S A. 2008;105(29):10149-54.

84. Gangwal K, Close D, Enriquez CA, Hill CP, Lessnick SL. Emergent Properties of EWS/FLI Regulation via GGAA Microsatellites in Ewing's Sarcoma. Genes Cancer. 2010;1(2):177-87.

85. Beck R, Monument MJ, Watkins WS, Smith R, Boucher KM, Schiffman JD, et al. EWS/FLI-responsive GGAA microsatellites exhibit polymorphic differences between European and African populations. Cancer Genet. 2012;205(6):304-12.

86. Johnson KM, Mahler NR, Saund RS, Theisen ER, Taslim C, Callender NW, et al. Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. Proc Natl Acad Sci U S A. 2017;114(37):9870-5.

87. Johnson KM, Taslim C, Saund RS, Lessnick SL. Identification of two types of GGAA-microsatellites and their roles in EWS/FLI binding and gene regulation in Ewing sarcoma. PLoS One. 2017;12(11):e0186275.

88. Sankar S, Bell R, Stephens B, Zhuo R, Sharma S, Bearss DJ, et al. Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. Oncogene. 2013;32(42):5089-100.

89. Theisen ER, Pishas KI, Saund RS, Lessnick SL. Therapeutic opportunities in Ewing sarcoma: EWS-FLI inhibition via LSD1 targeting. Oncotarget. 2016;7(14):17616-30.

90. Theisen ER, Selich-Anderson J, Miller KR, Tanner JM, Taslim C, Pishas KI, et al. Chromatin profiling reveals relocalization of lysine-specific demethylase 1 by an oncogenic fusion protein. Epigenetics. 2021;16(4):405-24.

91. Tomazou EM, Sheffield NC, Schmidl C, Schuster M, Schonegger A, Datlinger P, et al. Epigenome mapping reveals distinct modes of gene regulation and widespread enhancer reprogramming by the oncogenic fusion protein EWS-FLI1. Cell Rep. 2015;10(7):1082-95.

92. Mendiola M, Carrillo J, Garcia E, Lalli E, Hernandez T, de Alava E, et al. The orphan nuclear receptor DAX1 is up-regulated by the EWS/FLI1 oncoprotein and is highly expressed in Ewing tumors. Int J Cancer. 2006;118(6):1381-9.

93. Gorthi A, Bishop AJR. Ewing sarcoma fusion oncogene: At the crossroads of transcription and DNA damage response. Mol Cell Oncol. 2018;5(4):e1465014.

94. Gorthi A, Romero JC, Loranc E, Cao L, Lawrence LA, Goodale E, et al. EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. Nature. 2018;555(7696):387-91.

95. Sanchez-Molina S, Figuerola-Bou E, Blanco E, Sanchez-Jimenez M, Taboas P, Gomez S, et al. RING1B recruits EWSR1-FLI1 and cooperates in the remodeling of chromatin necessary for Ewing sarcoma tumorigenesis. Sci Adv. 2020;6(43).

96. Kinsey M, Smith R, Lessnick SL. NR0B1 is required for the oncogenic phenotype mediated by EWS/FLI in Ewing's sarcoma. Mol Cancer Res. 2006;4(11):851-9.

97. Spahn L, Siligan C, Bachmaier R, Schmid JA, Aryee DN, Kovar H. Homotypic and heterotypic interactions of EWS, FLI1 and their oncogenic fusion protein. Oncogene. 2003;22(44):6819-29.

98. Becker JS, McCarthy RL, Sidoli S, Donahue G, Kaeding KE, He Z, et al. Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. Mol Cell. 2017;68(6):1023-37 e15.

99. Szabo Q, Bantignies F, Cavalli G. Principles of genome folding into topologically associating domains. Sci Adv. 2019;5(4):eaaw1668.

100. Zaret KS. Pioneer Transcription Factors Initiating Gene Network Changes. Annu Rev Genet. 2020;54:367-85.

101. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol Cell. 2002;9(2):279-89.

102. Mayran A, Sochodolsky K, Khetchoumian K, Harris J, Gauthier Y, Bemmo A, et al. Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening. Nat Commun. 2019;10(1):3807.

103. Fernandez Garcia M, Moore CD, Schulz KN, Alberto O, Donague G, Harrison MM, et al. Structural Features of Transcription Factors Associating with Nucleosome Binding. Mol Cell. 2019;75(5):921-32 e6.

104. Zhu F, Farnung L, Kaasinen E, Sahu B, Yin Y, Wei B, et al. The interaction landscape between transcription factors and the nucleosome. Nature. 2018;562(7725):76-81.

105. Riggi N, Knoechel B, Gillespie SM, Rheinbay E, Boulay G, Suva ML, et al. EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. Cancer Cell. 2014;26(5):668-81.

106. Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, et al. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell. 2017;171(7):1573-88 e28.

107. Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, et al. Two independent modes of chromatin organization revealed by cohesin removal. Nature. 2017;551(7678):51-6.

108. Zhao Z, Shilatifard A. Epigenetic modifications of histones in cancer. Genome Biol. 2019;20(1):245.

1.7 Figures

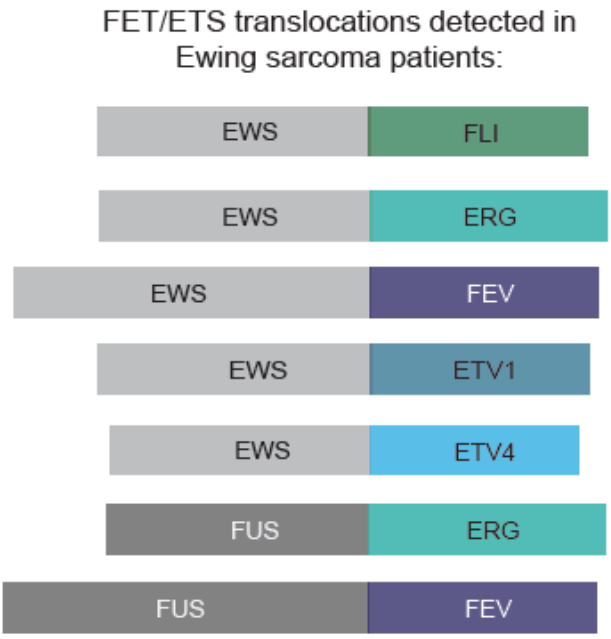

Figure 1.1: EWS/ETS and FUS/ETS translocations detected in Ewing sarcoma patient tumors. EWS and FUS translocations have been detected with ETS members found in the ERG (FLI, ERG, FEV) and PEA3 (ETV1 and ETV4) subfamilies. EWS/FLI is the most common translocation found in 85-90% of patients, EWS/ERG is found in ~10% of patients, and the additional 5 fusions are found in >5% of patients. Several translocation breakpoints have been detected for each fusion type and the schematics here represent the to-scale length of protein included in real patient tumor translocations.

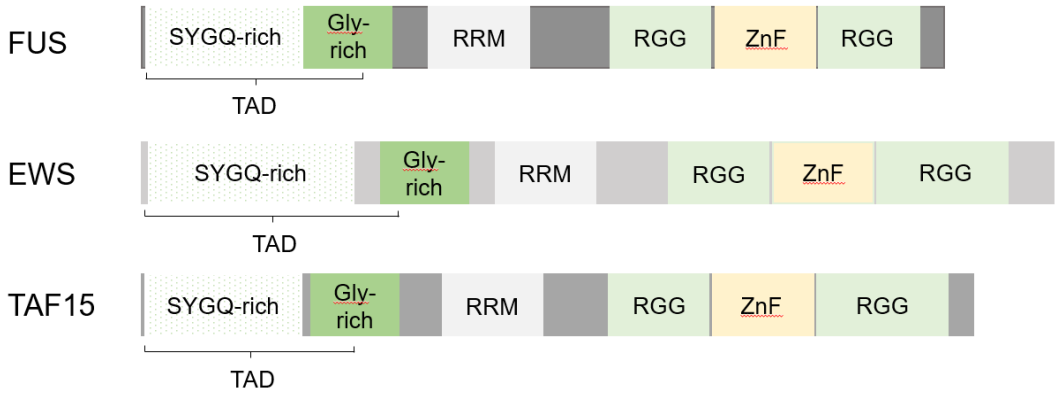

Figure 1.2. Structural features of the FET family of proteins. The FET protein family is characterized by the presence of several different domains, including SYGQ-rich and Gly(cine)-rich regions in the amino-terminal domains that form strong transcriptional activation domains. FET protein also contain RNA-recognition motifs (RRM), several arginine-glycine-glycine (RGG) boxes, and zinc finger domains (ZnF) in the carboxyl-terminal domains.

Figure 1.3. Structural characteristics of the ETS transcription factor family. The ETS family of proteins are characterized by the presence of a pointed (PNT) domain and an ETS DNA-binding domain (DBD) in the amino-terminal and carboxyl-terminal domains, respectively. Although present in <50% of the ETS proteins, the PNT domain is crucial for protein-protein interactions and the ETS DBD enables the ETS family to bind to DNA with a GGA(A/T) core.

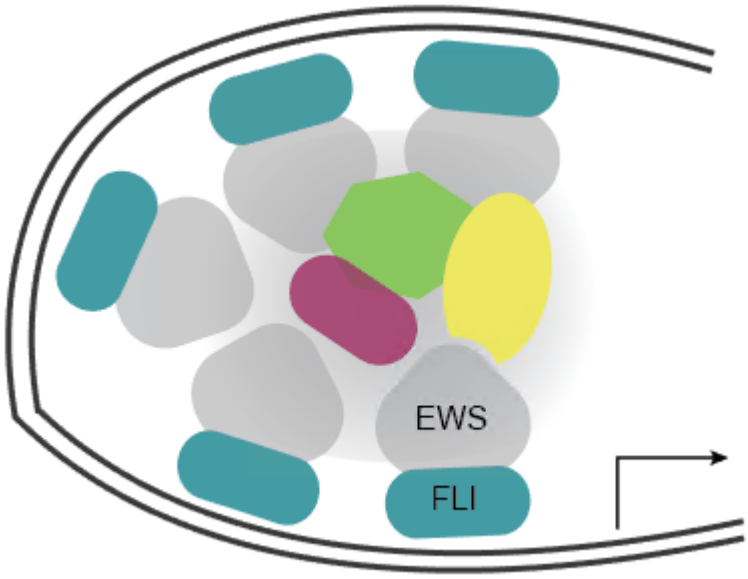

Figure 1.4. EWS/FLI interactions occur in transcriptional hubs. When multiple EWS/FLI molecules bind to preferential DNA sequences in the genome, such as those found in promoter or enhancer regions upstream of the transcription start site of a downstream gene target, the fusion protein interacts with itself to multimerize. EWS/FLI is believed to form transcriptional "hubs", where it interacts with other proteins and molecules to facilitate transcriptional regulation.

Chapter 2:

Identification of a novel FUS/ETV4 fusion and comparative analysis with other Ewing sarcoma fusion proteins

In review process at *Molecular Cancer Research* as:

Boone MA, Taslim C, Crow JC, Selich-Anderson J, Watson M, Heppner P, Hamill J, Wood AC, Lessnick SL, Winstanley M. Identification of a novel FUS/ETV4 fusion and comparative analysis with other Ewing sarcoma fusion proteins. 2021.

2.1 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 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 Ewing sarcoma patient. 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.

2.2 Implications

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

2.3 Introduction

Ewing sarcoma is an aggressive bone- and soft tissue-associated cancer primarily diagnosed in children and young adults (1, 2). The disease is characterized by the presence of a chromosomal translocation that encodes fusions between the amino-terminal domain of a FET (FUS, EWSR1, and TAF15) protein to the carboxyl-terminal domain of an ETS (E26 Transformation-Specific) transcription factor family member. The most common chromosomal translocation, present in ~85% of cases, is the t(11;22)(q24:q12), that fuses the *EWSR1* gene to *FLI1* 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. Following the identification of EWS/FLI nearly three decades ago, an EWS/ERG fusion was found in ~10% of cases, followed by five other fusions that are present in <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 domain structure to EWS/FLI. Here, we report a novel eighth Ewing sarcoma fusion, FUS/ETV4, identified in a neonatal patient.

The ETS protein family is a large 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 five 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). It is believed that these ETS family members bind similar high-affinity target 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 (IDRs) 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 mediate phase separation and/or "hub" formation (6, 7). These unique properties are likely critical for FET/ETS fusion oncoprotein-mediated reorganization of chromatin architecture, formation of transcriptional hubs, and recruitment of necessary transcriptional co-factors, such as the BAF complex, and are thus likely crucial for Ewing sarcomagenesis (1, 2).

The vast majority of molecular studies of Ewing sarcoma fusions have focused on EWS/FLI, but there are few, if any, detailed analyses of other fusion proteins. While it makes logical sense that FET/ETS fusions will have similar biologic functions, this has not been formally demonstrated. This focus on EWS/FLI and relative lack-of-focus on the other Ewing fusions has significant impact on the clinical management of patients (8). The advent of next-generation sequencing has allowed for the ready identification of EWS-based and FUS-based fusion transcripts or genomic-rearrangements (9). These

technological advances have made the identification of fusion partners easier, and simultaneously raised new questions as to how to apply this information to clinical care. As an example, a recent survey by the Children's Oncology Group found that only ~35% of clinician respondents indicated that non-EWS/FLI FUS/ETS fusions should be classified as Ewing sarcoma (8). Importantly, a significant portion of respondents indicated they were unsure whether alternative FET/ETS fusions should even be used to diagnose Ewing sarcoma or to allow patients to be included in Ewing sarcoma clinical trials (8).

We now report the initial identification of a novel FUS/ETV4 fusion in a patient with Ewing sarcoma, and perform genomic localization and transcriptional studies in an Ewing sarcoma A673 knock-down/rescue model system. We used this same system to perform the first comparative analysis of other rare, non-EWS/FLI, fusions in Ewing sarcoma, and in doing so, we demonstrate strong similarities between all of the fusions, demonstrating that they are indeed functionally similar. At the same time, we find 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.

2.4 Materials and Methods

Constructs and retroviruses

Puromycin-resistant retroviral vectors encoding shRNAs targeting Luciferase (iLuc; sequence: 5'-GATCCCCCTTACGCTGAGTACTTCGATTCAAGAGATCGAAGTACTCAGCGTAAGTTTTTGGAAC-3") or the 3'-UTR of endogenous EWS/FLI mRNA (iEF; sequence: 5'-GATCCCCATAGAGGTGGGAAGCTTATTTCAAGAGAATAAGCTTCCCACCTCTATTTTTTGGAAC-3') were previously described (10, 11). FET/ETS fusions (containing amino-terminal 3xFLAG-tags) were cloned into pMSCV-Hygro (Invitrogen); sequence details provided in Table 2.1.

Cell culture methods

HEK-293EBNA (Invitrogen) and A673 cells (ATCC), cultured for 1-6 weeks, in appropriate media and retroviruses produced and used for infection as described (10-12). STR profiling and mycoplasma testing are performed annually on all cell lines.

Immunodetection

Whole-cell protein extraction, protein 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 LiCor 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 iTaq Universal SYBR Green 1-Step Reaction Mix (BioRad 1725151) on a Bio-Rad CFX Connect Real-Time System. Primer sequences listed in Table 2.2.

CUT&Tag and Analysis

CUT&Tag (Cleavage Under Targets and Tagmentation) was performed as described by (13) on two biological replicates of knock-down/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, de-duplicated using SAMTOOLS (RRID:SCR_002105), aligned to hg19 reference genome, spike-in normalized using DESeq2 (median ratio method, RRID:SCR_015687), tracks were generated and averaged across biological replicates using Deeptools (RRID:SCR_016366), and peaks were 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 (False Discovery Rate) < 0.05, log2(fold-change) > 3 over control samples (iEF+Empty Vector), mean normalized counts > 80. Overlaps were determined using VennDiagram (RRID:SCR_002414) and GenomicRanges (RRID:SCR_000025) (17).

RNA-sequencing and Analysis

RNA-sequencing was performed on two biological replicates of knock-down/rescue A673 cell samples. TruSeq Stranded mRNA Kit (Illumina Cat. 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 human genome and analyzed for differential expression 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 compared to control cells (iEF+Empty Vector) (FDR < 0.05).

Statistical Analysis

PCR data is presented as mean ± SEM. Significance of soft agar assays was determined using a Student's t-test, or as otherwise noted; p-values < 0.05 were considered to be significant.

2.5 Results and Discussion

Identification of a novel FUS/ETV4 translocation

An infantile patient presented with a left posterior mediastinal mass (Figure 2.1A). 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 (Figure 2.1B). The tumor was CD99-positive in a diffuse membranous staining pattern (Figure 2.1C), and positive for nuclear NKX2-2 expression (Figure 2.1D). *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 1-9 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

There were no cell lines nor patient-derived xenograft models available from the patient in which to analyze the transcriptional functions of FUS/ETV4. We therefore cloned FUS/ETV4 into a retroviral expression vector, and also cloned EWS/ETV4 as the most similar *bona fide* rare Ewing sarcoma fusion (Figure 2.2A). To allow for analysis in an isogenic background, we knocked-down endogenous EWS/FLI in A673 Ewing sarcoma cells and expressed either EWS/ETV4 or FUS/ETV4 fusion proteins through retroviral transduction (Figure 2.3A-B).

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 >12,000 loci and EWS/ETV4 bound >17,000 loci. Strikingly, >10,000 bound loci were shared between the two proteins, and over 80% of FUS/ETV4 peaks overlapped with those of EWS/ETV4 (Figure 2.2B).

We next asked whether FUS/ETV4 induced a similar transcriptional profile to EWS/ETV4. Consistent with the genomic localization studies, RNA-sequencing 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 (Figure 2.2C). Both fusions were capable of binding and regulating genes previously documented as EWS/FLI targets, including those associated with both high-affinity and GGAA-microsatellite binding sites (Figure 2.3C-D). Taken together, these data demonstrate that the novel FUS/ETV4 fusion has transcriptional function that are 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 the A673 knock-down/rescue 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 to FUS/ERG, and EWS/FEV to FUS/FEV (Figure 2.4A and Figure 2.5A-B; 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 >80% of the EWS/ERG loci also bound by FUS/ERG (Figure 2.4B, left panel; NB: The higher number of FUS/ERG-bound loci likely reflects higher protein expression of FUS/ERG, see Figure 2.5B). Similarly, EWS/FEV and FUS/FEV shared almost 15,000 bound regions, accounting for ~70% of the regions bound by both fusion proteins (Figure 2.4B, right panel).

RNA-sequencing revealed an ~80% overlap between genes regulated by EWS/ERG and FUS/ERG and that each regulated >9,000 genes (suggesting that much of the "excess" FUS/ERG binding was not-functionally associated with gene regulation; Figure 2.4C, left panel). Similarly, EWS/FEV and FUS/FEV regulated ~5,600 genes in common, representing ~65% of the genes regulated by EWS/FEV and ~90% of genes regulated by FUS/FEV (Figure 2.4C, 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 all 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 of whether 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 knock-down/rescue system (Figure 2.5A-B). The EWS/ETS fusions shared almost 9,000 bound loci (Figure 2.6A), and the FUS/ETS fusions shared >8,700 bound loci (Figure 2.6B). RNA-sequencing showed similar trends, with >5,400 genes similarly regulated by each of the EWS/ETS fusions (Figure 2.6C), and ~2,900 genes regulated by the FUS/ETS proteins (Figure 2.6D).

Global similarities across all FET/ETS fusions support the inclusion of all 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 on amino-terminal 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 over 6,600 loci were similarly bound and that ~2,600 genes were similarly regulated by all fusion proteins tested (Figure 2.4D). These overlaps were highly significant ($p<2.2x10^{-16}$). We again observed that all fusions bound and regulated both GGAA-microsatellite associated genes, and genes associated with high-affinity ETS binding sites (Figures 2.7A-B and 2.8A-B). Taken together, these data support the assertion that all FET/ETS fusion proteins have similar capabilities to bind DNA and regulate gene expression.

The most common fusion in Ewing sarcoma, EWS/FLI, has been extensively studied (1, 2, 18). This work has led to development of novel concepts for EWS/FLI protein function, including the function of the EWS-portion of the fusion as a transcriptional regulatory domain, likely through the assembly of transcriptional hubs via self-association properties, the ability of the EWS-portion to recruit transcriptional co-regulators, such as BAF and LSD1, and the ability of the fusion to alter chromatin architecture (1, 2, 6, 19). Undergirding these properties lies the ability of the fusion 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 ultimately resulting in the formation of Ewing sarcoma. Although additional Ewing sarcoma translocations have been identified, the analysis of these 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 functional differences in the fusions exist that might result in some being more rarely associated with Ewing sarcoma, or whether there is a critical interplay between fusion type and cellular background that is required for Ewing sarcoma development. 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 one of the rare translocations. 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 (8).

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 of the 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. Accordingly, these data 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, as well as clinical trials.

2.6 Declarations

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.

Acknowledgements

We 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 Iftekhar Showpnil for thoughtful discussion concerning the hypothesis and methodology of the project, as well as copyediting this manuscript. Additionally, we would like to acknowledge the generosity of the patient's family for the use of the tumor sequencing information in this study.

Authors Contributions

MAB, SLL, and MW(instanley) are responsible for conceptualization of the project. For patient care, MW(instanley) was acting oncologist, MW(atson) was acting pathologist, PH was acting neurosurgeon, JH was surgeon responsible for local tumor control, and AW provided patient sequencing guidance. Methodology for laboratory studies was formulated by MAB and SLL. Investigation was performed by MAB, JCC, and JSA. Data analysis was performed by MAB and CT. Manuscript preparation was completed by MAB and

reviewing and editing was performed by all authors. Funding acquisition and supervision completed by SLL.

Availability of data

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.

2.7 References

1. Riggi N, Suva ML, Stamenkovic I. Ewing's Sarcoma. N Engl J Med. 2021;384(2):154-64.

2. Grunewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H, et al. Ewing sarcoma. Nat Rev Dis Primers. 2018;4(1):5.

3. Sankar S, Lessnick SL. Promiscuous partnerships in Ewing's sarcoma. Cancer Genet. 2011;204(7):351-65.

4. Ng TL, O'Sullivan MJ, Pallen CJ, Hayes M, Clarkson PW, Winstanley M, et al. Ewing sarcoma with novel translocation t(2;16) producing an in-frame fusion of FUS and FEV. J Mol Diagn. 2007;9(4):459-63.

5. Sizemore GM, Pitarresi JR, Balakrishnan S, Ostrowski MC. The ETS family of oncogenic transcription factors in solid tumours. Nat Rev Cancer. 2017;17(6):337-51.

6. Boulay G, Sandoval GJ, Riggi N, Iyer S, Buisson R, Naigles B, et al. Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. Cell. 2017;171(1):163-78 e19.

7. Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science. 2018;361(6400).

8. Kinnaman MD, Zhu C, Weiser DA, Mohiuddin S, Hingorani P, Roth M, et al. Survey of Paediatric Oncologists and Pathologists regarding Their Views and Experiences with Variant Translocations in Ewing and Ewing-Like Sarcoma: A Report of the Children's Oncology Group. Sarcoma. 2020;2020:3498549.

9. Machado I, Noguera R, Pellin A, Lopez-Guerrero JA, Piqueras M, Navarro S, et al. Molecular diagnosis of Ewing sarcoma family of tumors: a comparative analysis of 560 cases with FISH and RT-PCR. Diagn Mol Pathol. 2009;18(4):189-99.

10. Johnson KM, Mahler NR, Saund RS, Theisen ER, Taslim C, Callender NW, et al. Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. Proc Natl Acad Sci U S A. 2017;114(37):9870-5.

11. Theisen ER, Miller KR, Showpnil IA, Taslim C, Pishas KI, Lessnick SL. Transcriptomic analysis functionally maps the intrinsically disordered domain of EWS/FLI and reveals novel transcriptional dependencies for oncogenesis. Genes Cancer. 2019;10(1-2):21-38.

12. Pishas KI, Drenberg CD, Taslim C, Theisen ER, Johnson KM, Saund RS, et al. Therapeutic Targeting of KDM1A/LSD1 in Ewing Sarcoma with SP-2509 Engages the Endoplasmic Reticulum Stress Response. Mol Cancer Ther. 2018;17(9):1902-16.

13. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019;10(1):1930.

14. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.

15. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.

16. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

17. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013;9(8):e1003118.

18. Lawlor ER, Sorensen PH. Twenty Years on: What Do We Really Know about Ewing Sarcoma and What Is the Path Forward? Crit Rev Oncog. 2015;20(3-4):155-71.

19. Sankar S, Theisen ER, Bearss J, Mulvihill T, Hoffman LM, Sorna V, et al. Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth. Clin Cancer Res. 2014;20(17):4584-97.

2.8 Figures

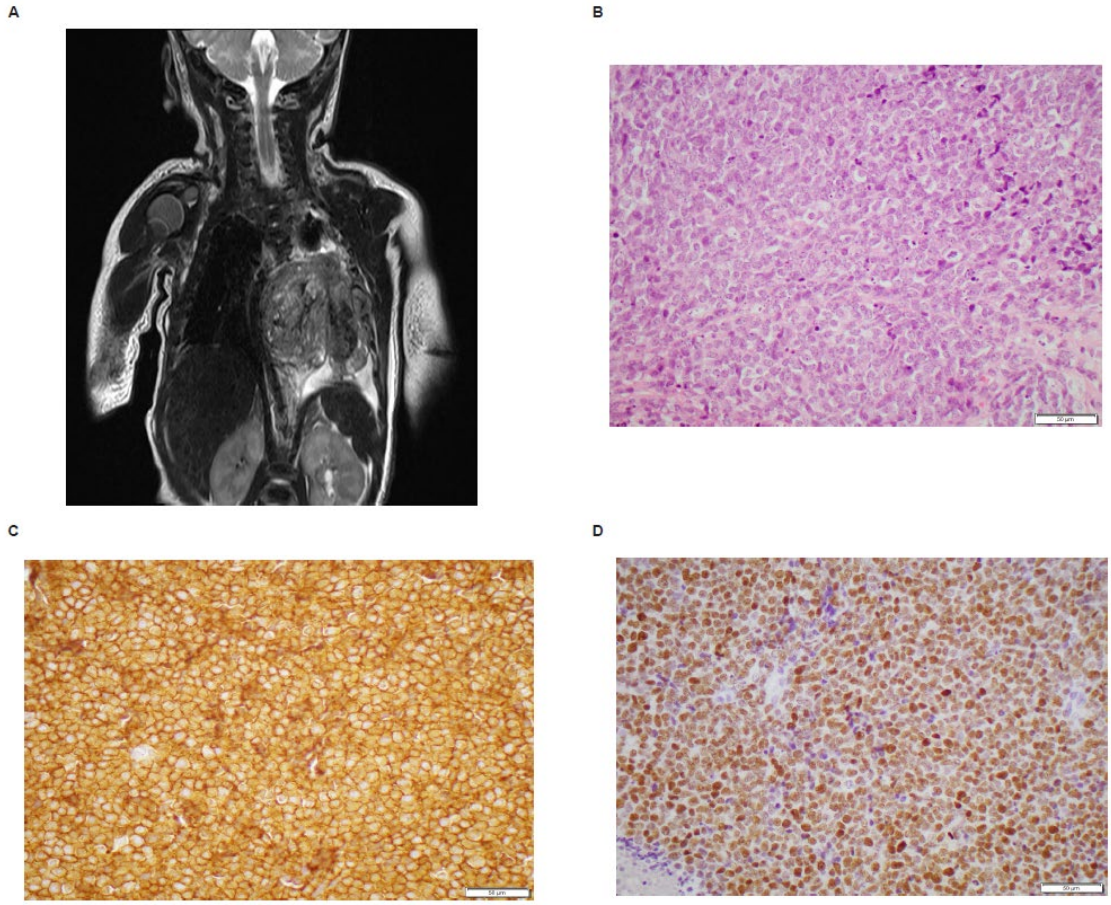

Figure 2.1. Neonatal patient presenting with Ewing sarcoma tumor

(A) Coronal magnetic resonance imaging (MRI) scan revealed a left posterior mediastinal mass. (B) Hematoxylin and eosin staining of patient tumor biopsy revealed sheets of undifferentiated, mitotically active small, round blue cells with dispersed chromatin and minimal amphophilic cytoplasm (50 μM scale bar depicted on image). (C) CD99 immunochemistry reveals diffuse membranous expression (50 μM scale bar depicted on

image). (D) NKX2-2 immunohistochemistry shows diffuse strong nuclear immunoreactivity (50 μM scale bar depicted on image).

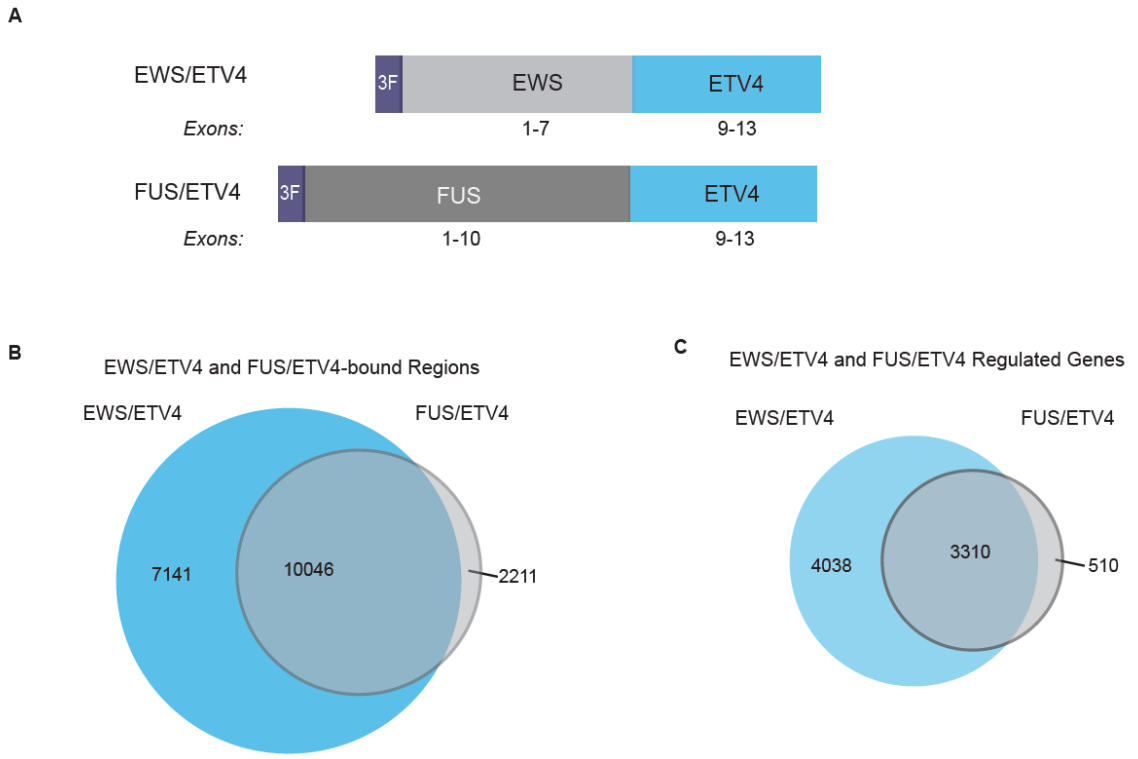

Figure 2.2. EWS/ETV4 and FUS/ETV4 DNA-binding and transcriptional profile overlap reveals similar biological functions

Protein schematic of 3xFLAG-tagged (3F) EWS/ETV4 and FUS/ETV4 constructs. EWS is represented in light grey, FUS in dark grey, and ETV4 in light blue. Exons included in each fusion are noted. (B) Venn diagram overlap analysis performed on CUT&Tag-detected genomic localization data for EWS/ETV4 and FUS/ETV4 expressed in A673 knock-down/rescue cells, as compared to control cells (Control: iEF + Empty Vector; EWS/ETV4: iEF + EWS/ETV4; FUS/ETV4: iEF + FUS/ETV4) (N=2). The number of peaks uniquely bound by each construct or those that are similarly bound are indicated in the figure. Significance of overlap: $p < 2.2 \times 10^{-16}$. (C) Venn diagram analysis of RNA-

sequencing results depicting significantly regulated genes for EWS/ETV4 and FUS/ETV4-expressing A673 knock-down rescue cells, as compared to iEF + Empty Vector control cells (N=2). Number of regulated genes for each construct is indicated in the figure. Significance of overlap: $p < 2.2 \times 10^{-16}$.

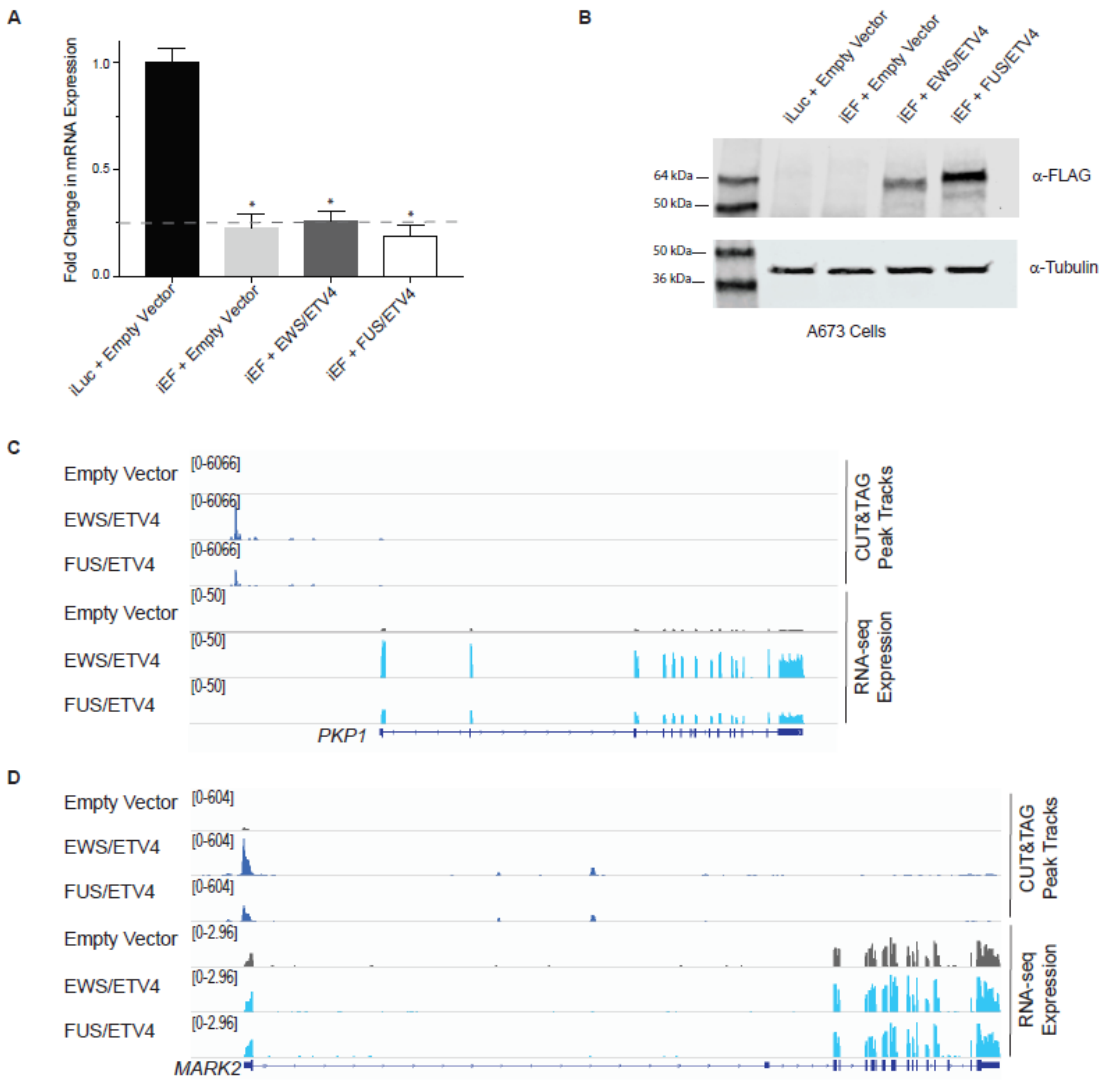

Figure 2.3. EWS/ETV4 and FUS/ETV4 fusion proteins studied in A673 knock-down/rescue model system

(A) Representative qRT-PCR results of endogenous EWS/FLI mRNA levels in A673 cells harboring the indicated constructs (iLuc: control shRNA; iEF: shRNA targets the 3'UTR of endogenous EWS/FLI). EWS/FLI mRNA values were normalized to RPL30 mRNA control values. Asterisks indicate samples are statistically different as compared to control

iLuc + Empty Vector cells (p-value < 0.05, N=1). (B) Western blot of 3xFLAG-tagged EWS/ETV4 and FUS/ETV4 protein expression in A673 cells. Membranes were probed with α-FLAG or α-tubulin (loading control) antibodies. (C-D) CUT&TAG and RNA-sequencing peak tracks visualized for Empty Vector cells (iEF + Empty Vector), EWS/ETV4-containing cells, and FUS/ETV4-containing cells (N=2 for each sample). Example genes include those associated with both microsatellite (*PKP1*) and high-affinity (HA) site (*MARK2*)-regulated genes. Peak track scales are depicted on the left.

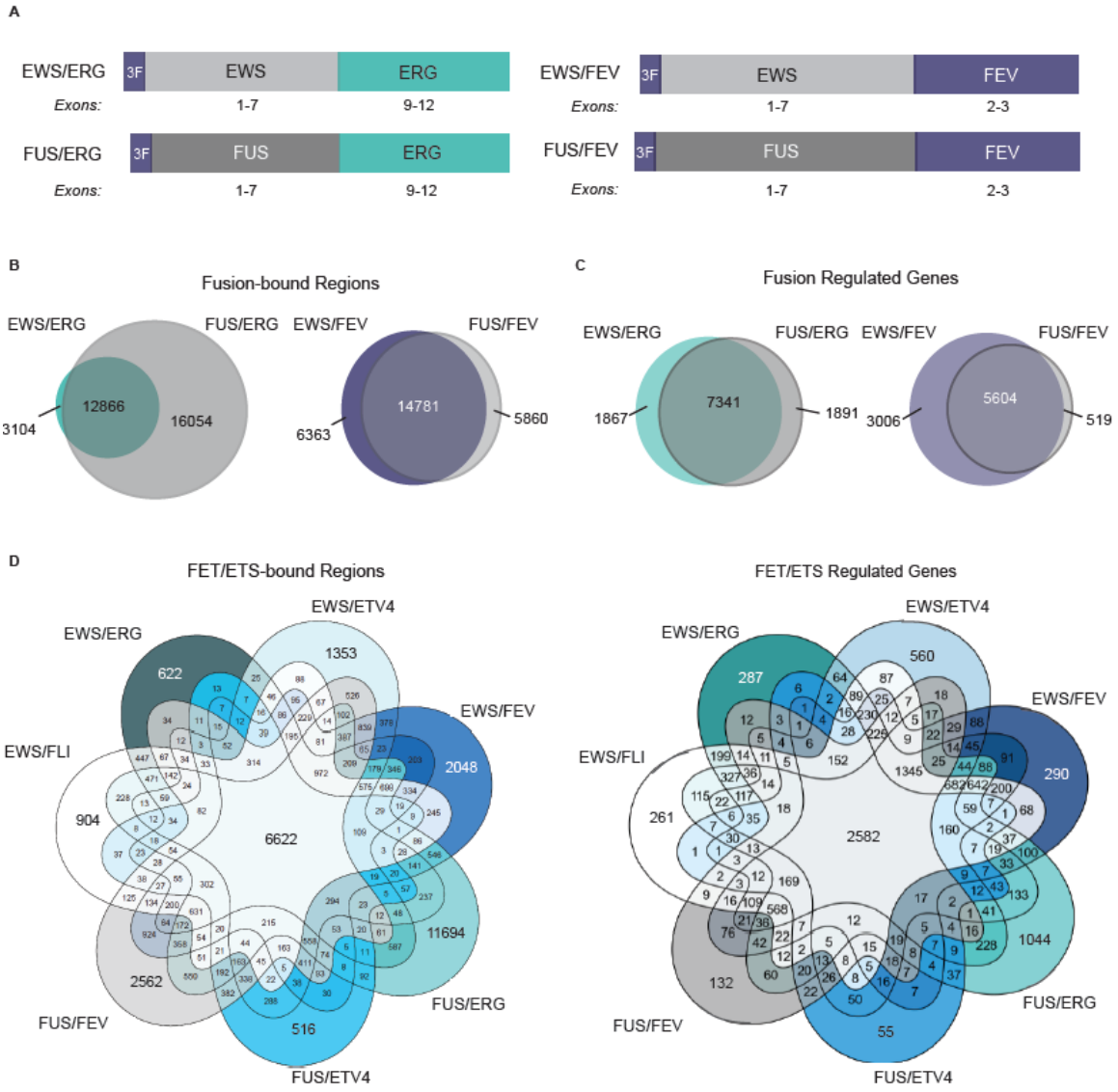

Figure 2.4. Comparison of FET/ERG and FET/FEV fusions shows similar genomic localization and transcriptional regulatory profiles

(A) Protein schematic of 3xFLAG-tagged (3F) cDNA constructs, including EWS/ERG, FUS/ERG, EWS/FEV, and FUS/FEV. EWS is depicted in light grey, FUS in dark grey, ERG in teal, and FEV in indigo. Exons included in each fusion are noted. (B) Venn diagram

overlap analysis of CUT&Tag genomic localization data for the corresponding fusion protein listed after expression in A673 knock-down/rescue cells (iEF + Construct), as compared to control cells (iEF + Empty Vector) (N=2). Number of bound regions for each construct depicted in figure. Significance of overlap: $p < 2.2 \times 10^{-16}$. (C) Venn diagram overlap analysis of RNA-sequencing expression data for genes called as significantly regulated by the corresponding construct listed in A673 knock-down/rescue cells, as compared to control cells (iEF + Empty Vector) (N=2). Number of significantly regulated genes by each fusion listed in figure. Significance of overlap: $p < 2.2 \times 10^{-16}$. (D) Venn diagram overlap analysis of CUT&Tag genomic localization binding data of FET/ETS translocations in A673 knock-down/rescue cells (N=2, left panel). All bound regions are called as significant for the corresponding translocation as compared to control cells (iEF + Empty Vector). Significance of overlap: $p < 2.2 \times 10^{-16}$. Venn diagram analysis of significantly regulated genes by corresponding FET/ETS translocations, as compared to control cells (iEF + Empty Vector) determined using RNA-sequencing (N=2, right panel). Significance of overlap: $p < 2.2 \times 10^{-16}$.

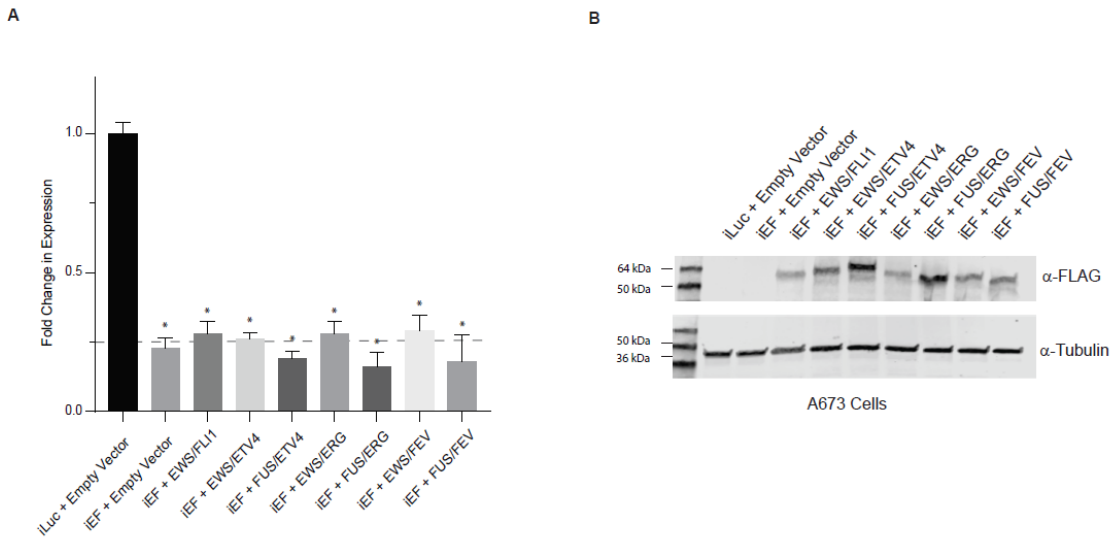

Figure 2.5. Successful expression of FET/ETS fusion proteins in A673 knock-down/rescue model system

(A) Representative qRT-PCR results determining endogenous EWS/FLI mRNA knock-down in A673 cells. iLuc + Empty Vector cells contain endogenous EWS/FLI mRNA, whereas iEF + Construct samples contain shRNA targeting the 3'UTR of endogenous EWS/FLI mRNA. All samples were normalized to RPL30 mRNA control samples. Statistical significance as compared to iLuc + Empty Vector is indicated by asterisks (p-value < 0.05, $N = 1$). (B) Western blot analysis demonstrating protein expression of FET/ETS fusion proteins in A673 knock-down/rescue cells. Membranes were probed for protein expression (α-FLAG) and a loading control (α-tubulin).

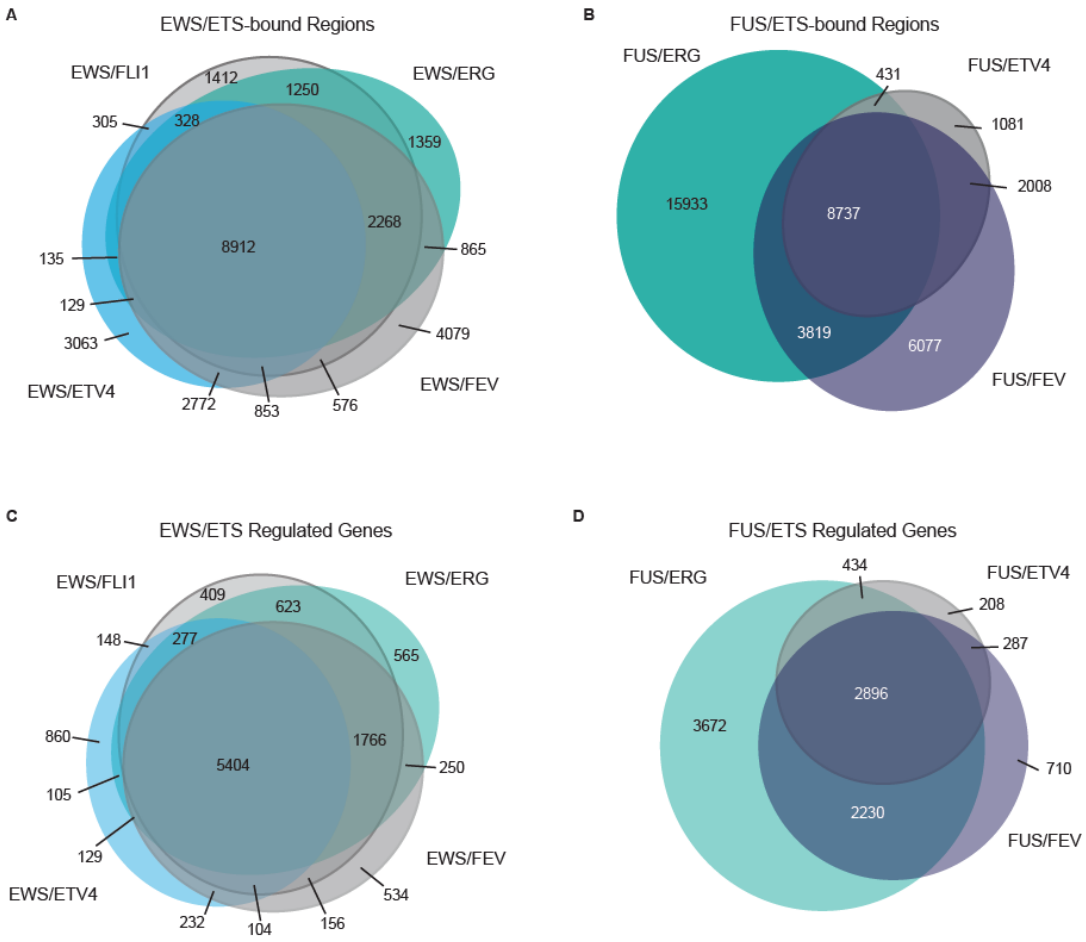

Figure 2.6. Overlap of EWS/ETS and FUS/ETS fusions reveals similar DNA-binding and transcriptional profiles

(A-B) DNA-bound regions called as significant over background for the (A) EWS/ETS fusions and (B) FUS/ETS fusions were overlapped (N=2 for each sample). Number of individually bound and shared bound regions are indicated in each circle. Significance of overlap: $p < 2.2 \times 10^{-16}$. (C-D) Venn diagram analysis depicts significantly regulated genes for (C) EWS/ETS and (D) FUS/ETS fusion proteins. The number of genes regulated by

each protein is indicated in the figure (N=2 for each sample). Significance of overlap: $p < 2.2 \times 10^{-16}$.

Figure 2.7 FET/ETS fusion proteins bind at known EWS/FLI response elements in Ewing sarcoma cells

(A-B) Representative peak tracks from CUT&Tag DNA-binding analysis are shown for Empty Vector (iEF + Empty Vector) A673 control cells, as well as A673 knock-down/rescue cells containing each of the FET/ETS fusions listed (N=2 for each sample). Examples of (A) GGAA-microsatellite (*VRK1*) and (B) high-affinity site (*BIRC2*) bound peaks typically associated with EWS/FLI function in Ewing sarcoma cells depicted here. Peak track scales displayed on the right.

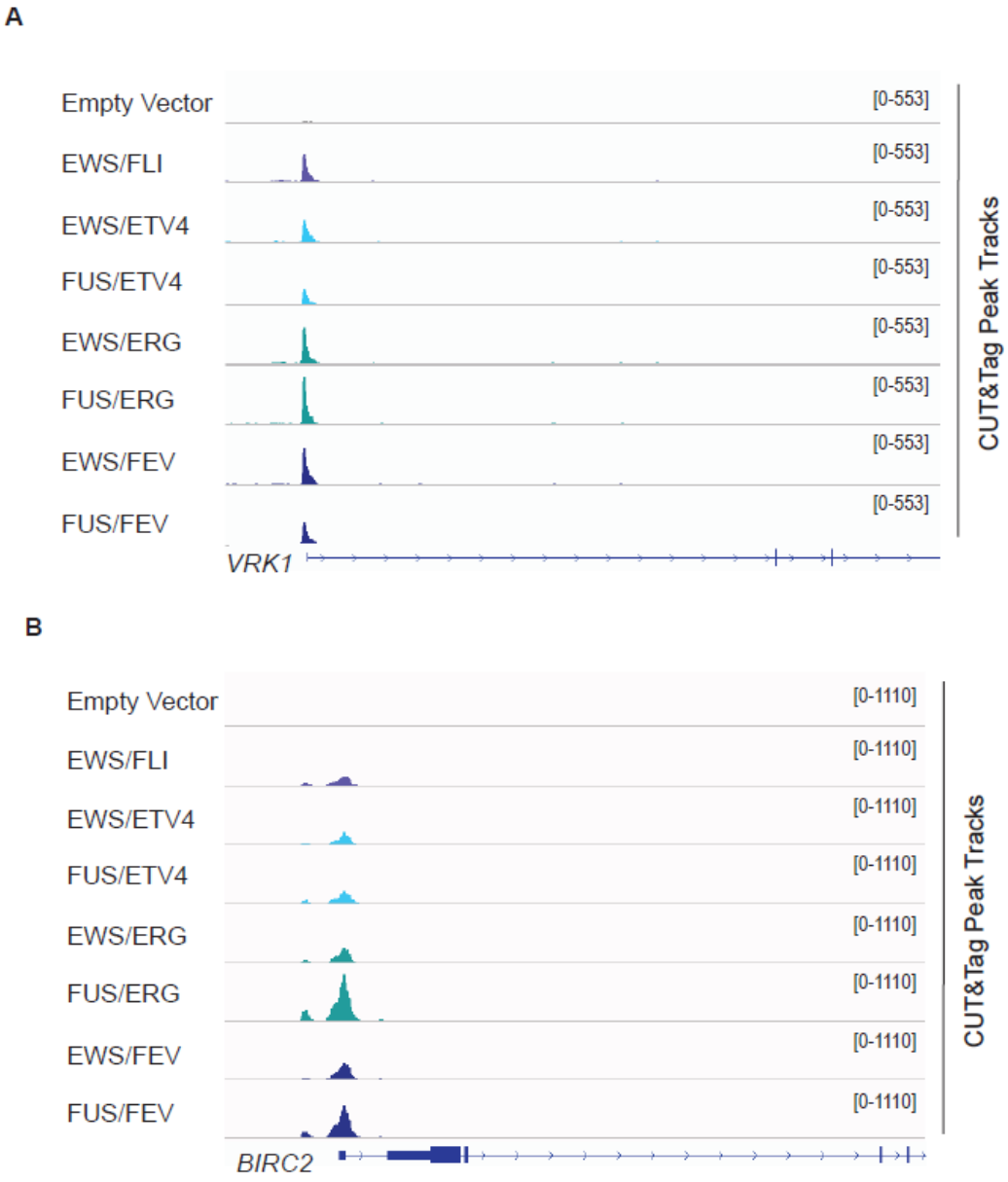

Figure 2.8. RNA-sequencing analysis reveals FET/ETS fusions regulate genes typically associated with Ewing sarcoma cells

(A-B) Representative tracks of RNA-sequencing expression data from IGV are shown for Empty Vector (iEF + Empty Vector) and rescue samples (iEF + FET/ETS translocation)

(N=2 for each sample). Examples of expression data are associated with EWS/FLI regulation in Ewing sarcoma cells via (A) GGAA-microsatellite (*CAV2*) and (B) high-affinity site (*LIPH*). Peak tracks scales are depicted on the right.

2.9 Tables

Table 2.1. Sequence of FET/ETS cDNA constructs.

Translocation	FET Exons	FET Amino Acids	ETS Exons	ETS Amino Acids
EWS/FLI1 EWSR1:NP_001156757.1 FLI1: NP_002008.2	1-7	1-265	7-9	242-452
EWS/ETV4 EWSR1:NP_001156757.1 ETV4: NP_001073143	1-7	1-265	9-13	271-484
FUS/ETV4 FUS: NP_004951.1 ETV4: NP_001073143	1-10	1-355	9-13	271-484
EWS/ERG EWSR1:NP_001156757.1 ERG: NP_891548.1	1-7	1-265	9-12	250-479
FUS/ERG FUS: NP_004951.1 ERG: NP_891548.1	1-7	1-255	9-12	250-479
EWS/FEV EWSR1:NP_001156757.1 FEV: NP_059991.1	1-10	1-347	2-3	18-238
FUS/FEV FUS: NP_004951.1 FEV: NP_059991.1	1-10	1-355	2-3	18-238

Table 2.1. Sequence of FET/ETS protein-encoding cDNA constructs used for manuscript, including corresponding exon and amino acid information. All translocations studied here directly correlate to translocations identified and reported in the literature found in Ewing sarcoma patient tumors.

Table 2.2. Sequences for primers used in qRT-PCR experiments.

Gene	Forward Primer	Reverse Primer
EWS/FLI	5'-CAGTCACTGCACCTCCATCC	5'-TTCATGTTATTGCCCCAAGC
RPL30	5'-GGGGTACAAGCAGACTCTGAAG	5'-ATGGACACCAGTTTTAGCCAAC

Table 2.2. Sequences of primers used for qRT-PCR experiments to determine knock-down of endogenous EWS/FLI mRNA. RPL30 is used as a control to normalize data for all samples.

Chapter 3:

The FLI portion of EWS/FLI contributes a transcriptional regulatory function that is distinct and separable from its DNA-binding function in Ewing sarcoma

Accepted for publication at *Oncogene* as:

Boone MA, Taslim C, Crow JC, Selich-Anderson J, Byrum AK, Showpnil IA, Sunkel BD, Wang M, Stanton BZ, Theisen ER, Lessnick SL. The FLI portion of EWS/FLI contributes a transcriptional regulatory function that is distinct and separable from its DNA-binding function in Ewing sarcoma. *Oncogene*. 2021.

3.1 Abstract

Ewing sarcoma is an aggressive bone cancer of children and young adults defined by the presence of a chromosomal translocation: t(11;22)(q24;q12). The encoded protein, EWS/FLI, fuses the amino-terminal domain of EWS to the carboxyl-terminus of FLI. The EWS portion is an intrinsically-disordered transcriptional regulatory domain, which the FLI portion contains an ETS DNA-binding domain and two flanking regions of unknown function. Early studies using non-Ewing sarcoma models provided conflicting information on the roles of each domain of FLI in EWS/FLI oncogenic function. We therefore sought to define the specific contributions of each FLI domain to EWS/FLI activity in a well-validated Ewing sarcoma model and, in doing so, to better understand Ewing sarcoma development mediated by the fusion protein.

We analyzed a series of engineered EWS/FLI mutants with alterations in the FLI portion using a variety of assays. Fluorescence anisotropy, CUT&RUN, and ATAC-sequencing experiments revealed that the isolated ETS domain is sufficient to maintain the normal DNA-binding and chromatin accessibility function of EWS/FLI. In contrast, RNA-sequencing and soft agar colony formation assays revealed that the ETS domain alone was insufficient for transcriptional regulatory and oncogenic transformation functions of the fusion protein. We found that an additional alpha-helix immediately downstream of the ETS domain is required for full transcriptional regulation and EWS/FLI-mediated

oncogenesis. These data demonstrate a previously unknown role for FLI in transcriptional regulation that is distinct from its DNA-binding activity. This activity is critical for the cancer-causing function of EWS/FLI and may lead to novel therapeutic approaches.

3.2 Introduction

Ewing sarcoma is a bone-tumor of children and young adults (2). These tumors contain chromosomal translocations that encode fusions between members of the FET and ETS protein families (1, 3). In ~85% of patients, this translocation occurs at t(11;22)(q24;q12), fusing *EWSR1* to *FLI1* and effectively encoding the EWS/FLI protein (1, 3-6). Numerous studies have demonstrated that EWS/FLI has oncogenic function and serves as the driver of Ewing sarcoma (1, 4, 7). Indeed, EWS/FLI is often the only genetic abnormality in these otherwise "genomically-quiet" tumors (8). Thus, determining the mechanisms underlying the oncogenic function of EWS/FLI is critical to understanding Ewing sarcoma tumorigenesis, identifying new therapeutic approaches, and may also shed light on the oncogenic mechanisms of other "ETS-associated" tumors.

EWS/FLI functions as an aberrant transcription factor that dysregulates several thousand genes (9, 10). EWS contributes strong transcriptional activating and repressing functions to the fusion (11-13). The mechanisms by which the EWS-portion mediates these functions are only beginning to be understood, but include the recruitment of epigenetic co-regulators

and RNA-polymerase II, perhaps via the formation of transcriptional "hubs", phase-separated droplets, or even polymerized fibrils (9, 14-17).

FLI is a member of the ETS transcription factor family (18-20). The ETS family is defined by the presence of highly-conserved winged helix-turn-helix DNA-binding domains (DBD) (18). The preferred high-affinity (HA) binding sequence for FLI is "ACCGGAAGTG", while other family members bind similar sequences containing a "GGA(A/T)" core surrounded by additional base pairs (18, 21). In addition to binding classic ETS HA sites, EWS/FLI gains the neomorphic ability to bind microsatellite sequences consisting of multiple "GGAA" repeats (22-24). Thousands of GGAA-microsatellite sequences are scattered throughout the human genome, many of which serve as EWS/FLI-response elements associated with genes critical for Ewing sarcomagenesis (22-24). Along with the ETS DNA-binding domain, the FLI portion of the fusion contains additional amino-terminal and carboxyl-terminal regions of uncertain function.

The cell of origin of Ewing sarcoma is unknown (25). Early studies of the FLI portion of EWS/FLI used heterologous cell types, such as NIH3T3 murine fibroblasts, with conflicting results (25). For example, May et al. found that expression of EWS/FLI induced oncogenic transformation of NIH3T3 cells in a manner dependent on the FLI DNA-binding domain (7). In contrast, Welford et al. showed the DNA-binding domain of FLI was not required for EWS/FLI-mediated oncogenic transformation (26). Subsequent studies in patient-derived Ewing sarcoma cells showed that a DNA-binding defective mutant of

EWS/FLI was unable to mediate oncogenic transformation, demonstrating that DNA-binding is absolutely required for EWS/FLI-mediated transformation in a more relevant Ewing cellular model (13). The carboxyl-terminal region of FLI (outside of the DNA-binding domain) was also evaluated in the NIH3T3 model and determined to be important for transcriptional control and oncogenic transformation mediated by EWS/FLI, though this has not been tested in a Ewing sarcoma model (27). Furthermore, gene expression patterns mediated by EWS/FLI in the NIH3T3 model were drastically different from those in Ewing sarcoma cellular models, suggesting that EWS/FLI may utilize alternative mechanisms to drive oncogenesis in different systems and that model system selection is important (25). To date, a systematic evaluation of the FLI portion of EWS/FLI in Ewing sarcoma cells has not been reported and so the roles of various regions of FLI in EWS/FLI-mediated oncogenic transformation remain unknown.

To address this, we analyzed the FLI portion of EWS/FLI in Ewing sarcoma cells using molecular and genomic techniques in our well-validated "knock-down/rescue" system. This model allowed us to identify a previously uncharacterized region just outside of the FLI DNA-binding domain as essential for EWS/FLI function. Mechanistic studies demonstrate a unique contribution of this region in mediating gene expression and subsequent oncogenic transformation that is independent of DNA-binding or the modulation of open chromatin states.

3.3 Materials and Methods

Constructs and retroviruses

Puromycin-resistant retroviral vectors encoding shRNAs targeting Luciferase (iLuc; sequence: 5'-GATCCCCCTTACGCTGAGTACTTCGATTCAAGAGATCGAAGTACTCAGCGTA AGTTTTTGGAAC-3") or the 3'-UTR of endogenous EWS/FLI mRNA (iEF; sequence: 5'-GATCCCCATAGAGGTGGGAAGCTTATTTCAAGAGAATAAGCTTCCCACCTCT ATTTTTTGGAAC-3') were previously described (24, 28). Full-length EWS/FLI and mutants (all containing amino-terminal 3xFLAG-tags) were cloned into pMSCV-Hygro (Invitrogen) with sequence details provided in Table 3.1. Luciferase reporter constructs (in pGL3 vectors; Promega Corporation) were previously described (24). Recombinant proteins (with a carboxyl-terminal 6xHistidine tag) were expressed using pET28a plasmids (EMD Chemicals).

Cell culture methods

HEK-293EBNA (Invitrogen) and A673 cells (ATCC) were grown, retroviruses produced and used for infection, and soft agar assays were performed as described (24, 28, 29). STR profiling and mycoplasma testing are performed annually on all cell lines. Dual luciferase reporter assays were performed in HEK-293EBNA cells as previously described (24). 3.75-5.0 microgram of cDNA constructs were transfected into HEK-293EBNA cells and collected 48 hours later for RNA-sequencing analysis.

Immunodetection

Whole-cell or nuclear protein extraction, protein quantification, and Western blot analysis was performed as previously described (24, 28, 29). Immunoblotting was performed using anti-FLAG M2 mouse (Sigma F1804-200UG), anti-α-Tubulin (Abcam ab7291), and anti-Lamin B1 (Abcam ab133741). Membranes were imaged using the LiCor 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 iTaq Universal SYBR Green 1-Step Reaction Mix (BioRad 1725151) on a Bio-Rad CFX Connect Real-Time System. Primer sequences are found in Table 3.2.

Recombinant protein purification

Recombinant 6xHistidine-tagged proteins were prepared from *E.coli* BL21(DE3) cells transformed with pET28a plasmids. Cells were resuspended (25mM Tris-HCl, pH 7.9, 1M NaCl, 0.1mM EDTA, 1mM PMSF, 5mM imidazole, proteinase inhibitors (Roche 4693159001)) and lysed via sonication. The lysate was centrifuged at 10 000xg for 30 minutes and the supernatant incubated with Ni-NTA resin (Qiagen) for 1-hour at 4ºC. Resin-bound protein was washed over a column with 90mL of lysis buffer and eluted using lysis buffer containing 500mM imidazole. Eluted protein was dialyzed overnight (300mM

KCl, 25mM Bis-Tris, 0.05% sodium azide, 5mM β-ME), treated with nuclease (Pierce 88700), and purified by ion-exchange chromatography as previously described (30). IEC fractions were combined, dialyzed into storage buffer (10% glycerol, 65mM KCI, 25mM Tris-HCl-pH 7.9, 6mM MgCl2, 0.5mM EDTA, 0.2mM PMSF, 1mM DTT), and concentrated using Amicon Ultra centrifugal filter units. A260/A280 ratio for purified proteins were determined to be between 0.55-0.58.

Fluorescence anisotropy

Fluorescence anisotropy was performed as previously described (24). Recombinant protein sequences and fluorescein-labelled DNA duplex sequences (ordered from IDT) are found in Table .3.1 and 3.3, respectively.

CUT&RUN and Analysis

Two biological replicates for each knock-down/rescue sample were analyzed by CUT&RUN using the anti-FLAG M2 mouse antibody (Sigma F1804-200UG) as described and sequenced with the Illumina HiSeq4000 (28). Raw reads were trimmed, de-duplicated, aligned to hg19 reference genomes, and peaks were called using macs2 and DiffBind (Bioconductor) using "iEF + Empty Vector" samples as controls (31). Bigwig files combining two replicates with normalization option "RPGC" were created using Deeptools (32). Overlapping peak analysis was completed using R packages ChIPpeakAnno and Genomic Ranges (33, 34).

RNA-sequencing and Analysis

RNA-sequencing was performed on three biological replicates for knock-down/rescue A673 samples in three separate experiments (Figure 3.6, Figure 3.8, and Figure 3.12, respectively). TruSeq Stranded mRNA Kit (Illumina Cat. 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 human genome and analyzed for differential analysis (using FASTQC, Multiqc, Trim_galore, STAR version 2.5.2b, DESeq2) (35). GSEA (Version 4.0.3) analysis was performed: significantly activated and repressed genes were defined using an FDR < 0.05 cut-off for EF DBD to create gene sets. EF DBD+ or EF genes were used as the rank-ordered gene list to compare with these gene sets (36). RNA-expression scatterplot analysis was performed as previously described (28).

ATAC-sequencing and Analysis

ATAC-sequencing was performed on two separate biological replicates for knock-down/rescue A673 cells as previously described and sequenced with Illumina HiSeq4000 (37, 38). The ENCODE pipeline was used for trimming, alignment to hg19 reference genome, and peak calling on individual replicates (ENCODE Project). RegioneR was used to perform permutation test and test significance of overlapping ATAC peaks in different samples (34). EnrichedHeatmap, ggplot2, ChIPpeakAnno, and GenomicRanges were used to calculate overlapping regions and create heatmaps (33, 34, 39, 40). Differential ATAC

peak analysis was completed using DiffBind (Bioconductor) and DESeq2 with an FDR<0.05 (35).

Code Availability

All code used to analyze sequencing data as described above are from publically available resources.

Statistical Analysis

Luciferase assay, soft agar assay, and PCR data are presented as mean ± SEM. Fluorescence anisotropy data are presented as mean ± SEM. Significance of experimental results was determined using a two-sided Student's t-test for comparison between groups. P-values less than 0.05 were considered to be significant.

3.4 Results

Amino- and carboxyl-terminal regions of FLI are dispensable for EWS/FLI-mediated transcriptional activation in luciferase reporter assays

We first sought to determine the role of the amino- and carboxyl-regions of FLI in EWS/FLI-mediated transcriptional activation using a luciferase reporter assay containing a 20xGGAA-repeat microsatellite (24). We used a "type IV-breakpoint" EWS/FLI fusion containing regions encoded by exons 1-7 of *EWSR1* fused to exons 7-9 of *FLI1* as the full-length protein with a 3xFLAG-tag (24, 28). We also created 3xFLAG-tagged "EF ΔN-FLI"

and “EF ΔC-FLI” mutants harboring deletions amino-terminal or carboxyl-terminal to the FLI DNA-binding domain, respectively (Figure 3.1A) (27). Expression plasmids encoding these proteins were co-transfected with the 20xGGAA-microsatellite luciferase reporter into HEK-293EBNA cells (Figure 3.1B). We found that all three versions of EWS/FLI were capable of activating luciferase reporter gene expression to similar levels (Figure 3.1C). These data demonstrate that neither the amino-terminal nor the carboxyl-terminal region of FLI is required for transcriptional activation mediated by EWS/FLI *in vitro*.

Flanking regions of the FLI DNA-binding domain of FLI are required for oncogenic function of EWS/FLI in a Ewing sarcoma cellular model

We next hypothesized that the only region of FLI critical for EWS/FLI activity is the ETS DNA-binding domain itself. The DNA-binding domain of FLI is not well-defined in the published literature. The ETS domain is often referred to as an 85-amino acid sequence (18, 19, 21). However, other structural and functional studies of FLI used a larger region of FLI as the ETS domain that included short amino- and carboxyl-extensions to the 85-amino acid “core” (7, 41). To test both “ETS domains”, we created two new mutant forms of EWS/FLI: “EF DBD” that fuses EWS directly to the 85-amino acid ETS domain and “EF DBD+” that fused EWS to a 102-amino acid ETS domain (containing 7- and 10-amino acid extensions on the amino-terminal and carboxyl-terminal sides of DBD, respectively) that has been used in prior studies (Figure 3.2A) (22).

Constructs were transfected into HEK-293EBNA cells and luciferase reporter assays using the 20xGGAA-microsatellite revealed that both EF DBD and EF DBD+ induced robust transcriptional activation and were even more active than full-length EWS/FLI (EF) itself (Figure 3.3A; Figure 3.2B).

To determine if the luciferase reporter results would translate to a more relevant Ewing sarcoma cellular model, we used our "knock-down/rescue" system to replace endogenous EWS/FLI with exogenous constructs in patient-derived A673 Ewing sarcoma cells (42). Retrovirally-expressed shRNAs targeting firefly luciferase (iLuc) or the 3'-UTR of endogenous EWS/FLI (iEF) were used to knock-down endogenous EWS/FLI (Figure 3.2C). EWS/FLI was subsequently rescued through retroviral expression of cDNA constructs (Empty Vector, EF, EF DBD, or EF DBD+) (Figure 3.2D). These cells were seeded into soft agar to measure anchorage-independent colony formation as a measure of oncogenic transformation (Figure 3.2E-F). Positive control cells (iLuc + Empty Vector) showed high rates of colony formation, while cells lacking EWS/FLI (iEF + Empty Vector) showed a near total loss of transformation capacity that was rescued by re-expression of full-length EWS/FLI (iEF + EF; Figure 3.2E-F). Interestingly, expression of EF DBD+ (iEF + EF DBD+) rescued colony formation to the same level as full-length EF, but the smaller EF DBD construct (iEF + EF DBD) failed to rescue colony formation (Figure 3.2E-F). These data define a significant functional difference between EF DBD and EF DBD+ in the A673 Ewing sarcoma model that is not correlated to their transcriptional activity in the luciferase reporter assay.

DNA-binding and genomic localization of EWS/FLI are nearly identical in FLI domain mutants

The inability of EF DBD to rescue A673 cell colony growth suggested a loss of a critical function as compared to EF DBD+, with the only difference between the two constructs being the 17-amino acids flanking the 85-amino acid DNA-binding domain core. We therefore reasoned that these flanking amino acids may contribute to EWS/FLI DNA-binding affinity. To test this, we performed fluorescence anisotropy studies to compare the ability of FLI DBD and FLI DBD+ recombinant protein to bind fluorescein-labeled DNA (Figure 3.4A, Figure 3.5A-B). We tested an ETS high-affinity (HA) site, a 2xGGAA-repeat microsatellite, and a 20xGGAA-repeat microsatellite (Figure 3.4B-D). We found that both FLI DBD and FLI DBD+ bound each DNA element with similar dissociation constants (K_D; Figure 3.4B-D).

Although *in vitro* DNA-binding was similar between FLI DBD and FLI DBD+ recombinant proteins, we next considered if differences in DNA-binding would be revealed in the context of a chromatinized human genome. To assess this, we performed CUT&RUN (Cleavage Under Targets & Release Under Nuclease) to determine the genomic localization of 3xFLAG-tagged EF, EF DBD, and EF DBD+ proteins in A673 cells using our knock-down/rescue system (28, 43). An anti-FLAG antibody was used to ensure we evaluated the localization of exogenous constructs and not any low-level residual EWS/FLI remaining after knock-down. We found that CUT&RUN identified a similar number of

binding peaks between EF (14 040), EF DBD+ (14 970), and EF DBD (14 394). Comparison of the binding locations for each construct demonstrated that 90% of EF DBD peaks overlap with those of EF and EF DBD+ (Figure 3.4E). Further exploration of EWS/FLI-bound high-affinity sites and microsatellites did not identify any significant differences between EF DBD and EF or EF DBD+ (Figure 3.4F-H). Taken together, these data indicate that there are no large-scale changes in DNA-binding capabilities that might explain the inability of EF DBD to rescue oncogenic transformation in Ewing sarcoma cells.

EF DBD exhibits a hypomorphic gene regulatory capability in Ewing sarcoma cells

The above studies demonstrated that genome-wide localization is nearly-identical between the EWS/FLI constructs. Although luciferase assays showed strong transcriptional activation by EF DBD, we considered whether the transcriptional regulatory function of EF DBD might be disrupted in a more relevant Ewing sarcoma model. To test this hypothesis, we performed RNA-sequencing on knock-down/rescue A673 cells expressing EF, EF DBD, or EF DBD+.

EF regulated 4 124 genes and EF DBD+ regulated 3 374 genes (at adjusted p-values < 0.05). Importantly, 90% of the genes regulated by EF DBD+ were also regulated by EF. In contrast, EF DBD demonstrated a loss in transcriptional regulation of both activated and repressed genes, regulating only 964 genes in total (Figure 3.6A-B).

We next performed a more detailed evaluation of the RNA-sequencing data using Gene Set Enrichment Analysis (GSEA). We asked where the activated and repressed gene sets of EF DBD fall in comparison to the rank-ordered gene expression list of EF DBD+. We found very strong correlations of both the activated and repressed gene sets (|NES| of 3.5 and 2.65, respectively; Figure 3.6C-D). Even stronger correlations were observed when EF DBD-regulated gene sets were compared with EF activated and repressed genes (|NES| of 7.09 and 5.65; Figure 3.7A-B).

The GSEA results revealed a near-complete "stacking" of the EF DBD-regulated genes at the furthest edges of the EF DBD+ (or EF) rank-ordered lists. This suggests that EF DBD significantly rescues a portion of the EWS/FLI-regulated genes, while other genes are still regulated, but to a not statistically-significant lower level. We therefore hypothesized that EF DBD functions as an attenuated, hypomorphic version of EWS/FLI. To test this hypothesis, we performed a scatterplot analysis to compare the ability of these constructs to rescue previously-reported EWS/FLI-regulated genes (44). Transcriptional regulation by EF DBD+ was highly correlated with regulation by EF for both activated (slope=0.88, R=0.93) and repressed genes (slope=0.94, R=0.97; Figure 3.6E and Figure 3.7C). In contrast, EF DBD demonstrated much weaker correlations (slope=0.32 with R=0.54 for activated genes; slope=0.54 with R=0.78 for repressed genes; Figure 3.6E and Figure 3.7C). These data suggest that EF DBD is regulating a similar set of genes, albeit more weakly than EF or EF DBD+.

To determine if the diminished activity of EF DBD was specific to the A673 knock-down/rescue model, we next sought to study transcriptional regulation of EF DBD and EF DBD+ in an alternative cell line. EF DBD and EF DBD+ constructs were transfected into the previously published HEK-293EBNA model system and RNA-sequencing analysis was performed (Figure 3.8A) (45). Venn diagram analysis of significantly regulated genes for EF DBD and EF DBD+ demonstrated that a loss of activity was again observed with EF DBD, but a majority of genes regulated by EF DBD overlapped with those regulated by EF DBD+ (Figure 3.8B-C). GSEA analysis revealed a highly significant correlation of EF DBD-activated and repressed genes when compared to EF DBD+-regulated genes (NES= 1.78 and -2.90, respectively; Figure 3.8D-E).

Taken together, these data indicate that EF DBD is significantly attenuated in its ability to regulate expression in multiple cell types. Thus, EF DBD is best considered a transcriptional regulatory hypomorph, even though its DNA-binding function is intact. The loss of oncogenic potential of EF DBD appears to be due to an underlying defect in transcriptional regulatory capability. This is an unanticipated result, as the transcriptional regulation function of EWS/FLI was believed to be mediated solely by the EWS-portion of the fusion with the FLI-portion contributing only DNA-binding function.

Capacity of EWS/FLI to mediate chromatin state is unaltered by deletions surrounding the FLI DNA-binding domain

It was recently reported that EWS/FLI functions as a pioneer transcription factor to open regions of chromatin that were previously closed (9, 15). As chromatin accessibility is a general necessity for transcriptional regulation, we next evaluated the role of EWS/FLI and its mutants on creation (or maintenance) of open chromatin states by performing ATAC-sequencing in our knock-down/rescue system. To focus on the role of the EWS/FLI mutants on chromatin accessibility, we overlapped EWS/FLI-bound DNA regions (identified in our CUT&RUN analysis) with the ATAC-sequencing data. We found that ~95% of the nearly 13 000 EWS/FLI-bound sites had detectable ATAC signal (Figure 3.9A), indicating that most EWS/FLI binding peaks are associated with open chromatin states.

To determine if EF DBD is defective in opening chromatin, we compared the ATAC signal at regions bound by EF DBD and those bound by EF DBD+. We found that almost 95% of ATAC peaks were shared between the two (Figure 3,9B), suggesting that there were not significant differences in EWS/FLI-associated accessible chromatin in EF DBD-containing cells.

To determine if more subtle differences in open chromatin might be associated with the capability of each mutant to regulate gene expression, we performed a heatmap analysis (Figure 3.10A-B). At EWS/FLI-bound loci near genes regulated by EF DBD+, we found that ATAC signal was similar between cells, regardless if EF DBD regulated the same genes or not. We also noted that the ATAC signal was similar at these sites in EWS/FLI

knock-down cells (EF KD), indicating that the loss of EWS/FLI is not associated with a closing of the open chromatin state, at least in this system (Figure 3.9C-D). These data indicate that the dysfunction of EF DBD in mediating gene regulation is not a consequence of altered pioneer-type function to induce or maintain an open chromatin state at regulated genes.

A fourth alpha-helix of the FLI DNA-binding domain is essential for EWS/FLI-mediated oncogenic transformation

Finally, we sought to determine which flanking region of EF DBD+ is critical for its oncogenic transformation function. We first engineered FLI DBD+ ΔN and ΔC recombinant proteins harboring deletions of either the amino-terminal 7-amino acids or the carboxyl-terminal 10-amino acids surrounding the core 85-amino acid FLI DNA-binding domain (Figure 3.11A-B). Fluorescence anisotropy performed on HA site, 2xGGAA-repeat microsatellite, and 20xGGAA-repeat microsatellite DNA revealed generally similar DNA-binding with slight differences for each construct on each target DNA (Figure 3.11C-F).

To study the role of the flanking regions of the FLI DNA-binding domain on EWS/FLI activity in the A673 knock-down/rescue system, we created EF DBD+ constructs harboring the same deletions as described above (EF DBD+ ΔN or EF DBD+ ΔC; Figure 3.12A-B). RNA-sequencing revealed that while the EF DBD+ ΔN protein retained transcriptional regulation activity similar to EF and EF DBD+, the EF DBD and EF DBD+ ΔC proteins

showed a similar loss of regulatory capacity (Figure 3.12C). This loss in transcriptional regulation correlated with oncogenic transformation capacity. Soft agar assays demonstrated that EF DBD+ ΔN was fully-functional, while EF DBD+ ΔC lost the ability to transform A673 cells (Figure 3.12D). These results define the 10-amino acids downstream of the FLI DNA-binding domain as essential for EWS/FLI-mediated transcriptional regulation and oncogenic transformation.

Analysis of a previously published FLI protein crystal structure revealed that this 10-amino acid sequence forms an additional fourth alpha-helix immediately downstream of the DNA-binding domain of FLI (41). To determine if this structure is necessary for EWS/FLI-driven oncogenic transformation, we created several amino-acid mutations to disrupt the α_4-helix of the EF DBD+ ΔN construct, which contains the smallest amount of FLI determined to retain full protein activity (EF DBD+ ΔN α-helix Mutant and EF DBD+ ΔN α–helix Pro Mutant; Figure 3.13A). Like EF DBD+ ΔC, these constructs failed to induce colony formation in A673 cells (Figure 3.13B-E). This suggests that the alpha-helix immediately downstream of the FLI DNA-binding domain is indeed contributing a critical function required for transcriptional regulation and oncogenic transformation properties of EWS/FLI.

A recent study demonstrated that recombinant FLI dimerizes via interactions between the α_4-helix of one FLI molecule with the a_1-helix of another FLI molecule (41). We found that introduction of an F362A mutation, shown to disrupt FLI homodimerization, to our

EF DBD+ construct had no effect on oncogenic transformation in A673 cells (Figure 3.14A-D). This suggests that homodimerization is not required for the oncogenic potential of EWS/FLI and this alpha-helical region must act in other capacities.

3.5 Discussion

Although several studies have suggested that the regions outside of the DNA-binding domain of FLI may be important for overall EWS/FLI function, the FLI-portion of the fusion has largely been viewed as simply contributing DNA-binding function. In the current study, we took a systematic approach to understand the contributions of FLI to EWS/FLI activity in a Ewing sarcoma cellular background. This allowed us to define a previously unappreciated role for the fourth alpha-helix of the extended FLI DNA-binding domain in transcriptional regulation. This alpha-helix does not appear to be important for the DNA-binding, genomic localization, or chromatin accessibility functions of EWS/FLI. Instead, loss of this helix results in a significant loss of gene-regulatory function that culminates in a complete loss of oncogenic transformation mediated by EWS/FLI.

The mechanism(s) by which the fourth alpha-helix participates in gene regulation will require additional studies. One possibility is this fourth alpha-helix is involved in protein-protein interactions with adjacent transcription factors. Several transcription factors interact with the FLI portion of EWS/FLI, including SRF and AP-1 members that form ternary complexes with EWS/FLI on DNA (46, 47). Published interaction sites for these

factors do not map to this critical alpha-helical region and so do not readily explain the differences in activity observed between EF DBD and EF DBD+ proteins. EWS/FLI may interact with other transcription factors via this region; however, we do not favor a loss of such EWS/FLI-transcription factor interactions as the most likely cause of the massive loss of transcriptional function by EF DBD. We reason that if there were losses of EWS/FLI interactions with specific transcription factors, we may have expected a more limited loss of gene expression (rather than the ~70% loss observed for EF DBD). Furthermore, the formation of ternary complexes between pairs of transcription factors with DNA tend to stabilize DNA binding, so we might also have anticipated a significant change in genomic localization of EF DBD, which was not observed. We currently favor a model whereby the fourth alpha-helix interacts with epigenetic regulators and/or components of the core transcriptional machinery that are required for global gene regulation, rather than regulation limited to specific loci.

Work in NIH3T3 murine fibroblasts suggested a role for the carboxyl-terminal region of FLI in mediating transcriptional down-regulation by EWS/FLI (27). Our work here rules out a significant role for this region in EWS/FLI-mediated oncogenesis. Additionally, luciferase reporter assays have long been used as functional screens, but our results demonstrate that activation on a luciferase reporter does not necessarily reflect function in a Ewing sarcoma cellular background. Indeed, we also note that we did not see direct evidence of the pioneer-type function of EWS/FLI in the Ewing sarcoma model, which had been previously observed in a mesenchymal stem cell model (9). In our system, EWS/FLI-

occupied sites remained open and accessible following knock-down of EWS/FLI. It may be that the 80-90% knock-down we achieved was insufficient to allow for chromatin closing of those loci or perhaps insufficient time was provided to allow for chromatin closing. Nevertheless, changes in chromatin accessibility were not associated with the transcriptional dysfunction exhibited by EF DBD. These findings highlight the importance of analyzing EWS/FLI activity in a relevant Ewing sarcoma cellular context.

A detailed comparison of ETS protein structures revealed that many harbor this additional fourth alpha-helix downstream of their DNA-binding domains. As such, the work presented here may have relevance beyond an EWS/FLI context. For example, Ewing sarcoma translocations involve one of five closely-homologous ETS family members (FLI, ERG, FEV, ETV1, and ETV4) (11). Additionally, *TMPRSS2-ERG* fusions exist in approximately 50% of prostate cancer cases, with *TMPRSS2-FEV*, *-ETV1*, *-ETV4*, and *-ETV5* fusions found in other patients (48). In fact, ETS family members have been implicated in numerous solid and liquid tumors via over-expression, amplification, mutations, and translocations (20). As the functional motif we identified as crucial for EWS/FLI activity is conserved in numerous ETS factors, the data presented in this report may have wide-ranging implications for oncogenesis in multiple tumor types.

In summary, we have taken a systematic structure-function approach to identify a previously unappreciated region in the extended FLI DNA-binding domain that is required for transcriptional regulation and oncogenic transformation mediated by EWS/FLI. This

transcriptional function is distinct from the DNA-binding and genomic localization functions typically associated with the ETS domain. This work has implications not only for the development of Ewing sarcoma, but may also be useful in understanding the development of other ETS-associated tumors and, perhaps, even normal ETS transcriptional function. A better understanding of this newly-defined region may lead to novel approaches for therapeutically-targeting EWS/FLI, as well as other ETS factors.

3.6 Abbreviations

ATAC-sequencing: Assay for Transposase-Accessible Chromatin using sequencing; CUT&RUN: Cleave Under Targets & Release Under Nuclease; cDNA: complementary DNA; DBD: DNA-binding domain; EF: experimental EWS/FLI cDNA constructs; ERG: ETS-related gene; ETS: E26 transformation specific; ETV1: ETS variant transcription factor 1; ETV4: ETS variant transcription factor 4; ETV5: ETS variant transcription factor 5; *EWSR1* (EWS): Ewing sarcoma breakpoint region 1; FET: FUS/TLS, EWSR1, TAF15; FEV: Fifth Ewing variant protein; *FLI1* (FLI): Friend leukemia integration 1; GSEA: Gene Set Enrichment Analysis; HA: high-affinity; HEK-293EBNA: Human embryonic kidney-293 cell line expressing Epstein Barr nuclear antigen; IEC: ion exchange chromatography; IGV: Integrated Genome Viewer; KD: knock-down; K_D= dissociation constant; log2(FC): log2(Fold Change); μSat: microsatellite; NES: Normalized Enrichment Score; qRT-PCR: quantitative Reverse Transcriptase-Polymerase Chain Reaction; SEM: standard error of the mean; STR: short tandem repeats; TMPRSS2: Transmembrane protease, serine 2

3.7 Declarations

Acknowledgements

We thank Dr. Kirsten N. Johnson, Dr. Kathleen I. Pishas, Dr. Jack Tokarsky, and Ariunaa Bayanjargal for thoughtful discussion concerning the hypothesis and methodology of this manuscript.

Competing interests

Research reported in this publication was supported by the National Institute of Health award T32 GM068412 to MAB, and U54 CA231641 to SLL. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. SLL declares a competing interest as a member of the advisory board for Salarius Pharmaceuticals. SLL is also a listed inventor on United States Patent No. US 7,939,253 B2, "Methods and compositions for the diagnosis and treatment of Ewing's sarcoma," and United States Patent No. US 8,557,532, "Diagnosis and treatment of drug-resistant Ewing's sarcoma." This does not alter our adherence to *Oncogene* policies on sharing data and materials.

Authors Contributions

MAB and SLL are responsible for conceptualization of the project. Investigation was performed by MAB, JCC, JSA, AKB, BDS, and BZS. Methodology was formulated by

MAB, JSA, AKB, ERT, IS, and BZS. Data analysis was performed by MAB, CT, ERT, IS, and MW. Manuscript preparation was completed by MAB and reviewing and editing was performed by all authors. Funding acquisition was completed by MAB and SLL. Supervision was provided by SLL.

Availability of data

The sequencing datasets generated and analyzed during the current study are available in the Gene Expression Omnibus and accessible at GSE160898. All other data generated or analyzed during this study are available from the corresponding author on reasonable request.

3.8 References

1. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature. 1992;359(6391):162-5.

2. Longhi A, Ferrari S, Tamburini A, Luksch R, Fagioli F, Bacci G, et al. Late effects of chemotherapy and radiotherapy in osteosarcoma and Ewing sarcoma patients: the Italian Sarcoma Group Experience (1983-2006). Cancer. 2012;118(20):5050-9.

3. Zucman J, Melot T, Desmaze C, Ghysdael J, Plougastel B, Peter M, et al. Combinatorial generation of variable fusion proteins in the Ewing family of tumours. EMBO J. 1993;12(12):4481-7.

4. Bailly RA, Bosselut R, Zucman J, Cormier F, Delattre O, Roussel M, et al. DNA-binding and transcriptional activation properties of the EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing sarcoma. Mol Cell Biol. 1994;14(5):3230-41.

5. Aurias A, Rimbaut C, Buffe D, Dubousset J, Mazabraud A. [Translocation of chromosome 22 in Ewing's sarcoma]. C R Seances Acad Sci III. 1983;296(23):1105-7.

6. Turc-Carel C, Aurias A, Mugneret F, Lizard S, Sidaner I, Volk C, et al. Chromosomes in Ewing's sarcoma. I. An evaluation of 85 cases of remarkable consistency of t(11;22)(q24;q12). Cancer Genet Cytogenet. 1988;32(2):229-38.

7. May WA, Gishizky ML, Lessnick SL, Lunsford LB, Lewis BC, Delattre O, et al. Ewing sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation. Proc Natl Acad Sci U S A. 1993;90(12):5752-6.

8. Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, Calicchio ML, et al. The genomic landscape of pediatric Ewing sarcoma. Cancer Discov. 2014;4(11):1326-41.

9. Riggi N, Knoechel B, Gillespie SM, Rheinbay E, Boulay G, Suva ML, et al. EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. Cancer Cell. 2014;26(5):668-81.

10. Braun BS, Frieden R, Lessnick SL, May WA, Denny CT. Identification of target genes for the Ewing's sarcoma EWS/FLI fusion protein by representational difference analysis. Mol Cell Biol. 1995;15(8):4623-30.

11. Sankar S, Lessnick SL. Promiscuous partnerships in Ewing's sarcoma. Cancer Genet. 2011;204(7):351-65.

12. Lessnick SL, Braun BS, Denny CT, May WA. Multiple domains mediate transformation by the Ewing's sarcoma EWS/FLI-1 fusion gene. Oncogene. 1995;10(3):423-31.

13. Sankar S, Bell R, Stephens B, Zhuo R, Sharma S, Bearss DJ, et al. Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. Oncogene. 2013;32(42):5089-100.

14. Erkizan HV, Uversky VN, Toretsky JA. Oncogenic partnerships: EWS-FLI1 protein interactions initiate key pathways of Ewing's sarcoma. Clin Cancer Res. 2010;16(16):4077-83.

15. Boulay G, Sandoval GJ, Riggi N, Iyer S, Buisson R, Naigles B, et al. Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. Cell. 2017;171(1):163-78 e19.

16. Spahn L, Siligan C, Bachmaier R, Schmid JA, Aryee DN, Kovar H. Homotypic and heterotypic interactions of EWS, FLI1 and their oncogenic fusion protein. Oncogene. 2003;22(44):6819-29.

17. Kovar H. Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. Sarcoma. 2011;2011:837474.

18. Liang H, Olejniczak ET, Mao X, Nettesheim DG, Yu L, Thompson CB, et al. The secondary structure of the ets domain of human Fli-1 resembles that of the helix-turn-helix DNA-binding motif of the Escherichia coli catabolite gene activator protein. Proc Natl Acad Sci U S A. 1994;91(24):11655-9.

19. Laudet V, Hanni C, Stehelin D, Duterque-Coquillaud M. Molecular phylogeny of the ETS gene family. Oncogene. 1999;18(6):1351-9.

20. Sizemore GM, Pitarresi JR, Balakrishnan S, Ostrowski MC. The ETS family of oncogenic transcription factors in solid tumours. Nat Rev Cancer. 2017;17(6):337-51.

21. Mao X, Miesfeldt S, Yang H, Leiden JM, Thompson CB. The FLI-1 and chimeric EWS-FLI-1 oncoproteins display similar DNA binding specificities. J Biol Chem. 1994;269(27):18216-22.

22. Gangwal K, Close D, Enriquez CA, Hill CP, Lessnick SL. Emergent Properties of EWS/FLI Regulation via GGAA Microsatellites in Ewing's Sarcoma. Genes Cancer. 2010;1(2):177-87.

23. Gangwal K, Lessnick SL. Microsatellites are EWS/FLI response elements: genomic "junk" is EWS/FLI's treasure. Cell Cycle. 2008;7(20):3127-32.

24. Johnson KM, Mahler NR, Saund RS, Theisen ER, Taslim C, Callender NW, et al. Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. Proc Natl Acad Sci U S A. 2017;114(37):9870-5.

25. Braunreiter CL, Hancock JD, Coffin CM, Boucher KM, Lessnick SL. Expression of EWS-ETS fusions in NIH3T3 cells reveals significant differences to Ewing's sarcoma. Cell Cycle. 2006;5(23):2753-9.

26. Welford SM, Hebert SP, Deneen B, Arvand A, Denny CT. DNA binding domain-independent pathways are involved in EWS/FLI1-mediated oncogenesis. J Biol Chem. 2001;276(45):41977-84.

27. Arvand A, Welford SM, Teitell MA, Denny CT. The COOH-terminal domain of FLI-1 is necessary for full tumorigenesis and transcriptional modulation by EWS/FLI-1. Cancer Res. 2001;61(13):5311-7.

28. Theisen ER, Miller KR, Showpnil IA, Taslim C, Pishas KI, Lessnick SL. Transcriptomic analysis functionally maps the intrinsically disordered domain of EWS/FLI and reveals novel transcriptional dependencies for oncogenesis. Genes Cancer. 2019;10(1-2):21-38.

29. Pishas KI, Drenberg CD, Taslim C, Theisen ER, Johnson KM, Saund RS, et al. Therapeutic Targeting of KDM1A/LSD1 in Ewing Sarcoma with SP-2509 Engages the Endoplasmic Reticulum Stress Response. Mol Cancer Ther. 2018;17(9):1902-16.

30. Regan MC, Horanyi PS, Pryor EE, Jr., Sarver JL, Cafiso DS, Bushweller JH. Structural and dynamic studies of the transcription factor ERG reveal DNA binding is allosterically autoinhibited. Proc Natl Acad Sci U S A. 2013;110(33):13374-9.

31. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.

32. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44(W1):W160-5.

33. Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics. 2010;11:237.

34. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013;9(8):e1003118.

35. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

36. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.

37. Hodges HC, Stanton BZ, Cermakova K, Chang CY, Miller EL, Kirkland JG, et al. Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. Nat Struct Mol Biol. 2018;25(1):61-72.

38. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10(12):1213-8.

39. Gu Z, Eils R, Schlesner M, Ishaque N. EnrichedHeatmap: an R/Bioconductor package for comprehensive visualization of genomic signal associations. BMC Genomics. 2018;19(1):234.

40. Maag JLV. gganatogram: An R package for modular visualisation of anatograms and tissues based on ggplot2. F1000Res. 2018;7:1576.

41. Hou C, Tsodikov OV. Structural Basis for Dimerization and DNA Binding of Transcription Factor FLI1. Biochemistry. 2015;54(50):7365-74.

42. Martinez-Ramirez A, Rodriguez-Perales S, Melendez B, Martinez-Delgado B, Urioste M, Cigudosa JC, et al. Characterization of the A673 cell line (Ewing tumor) by molecular cytogenetic techniques. Cancer Genet Cytogenet. 2003;141(2):138-42.

43. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife. 2017;6.

44. Sankar S, Theisen ER, Bearss J, Mulvihill T, Hoffman LM, Sorna V, et al. Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth. Clin Cancer Res. 2014;20(17):4584-97.

45. Mendiola M, Carrillo J, Garcia E, Lalli E, Hernandez T, de Alava E, et al. The orphan nuclear receptor DAX1 is up-regulated by the EWS/FLI1 oncoprotein and is highly expressed in Ewing tumors. Int J Cancer. 2006;118(6):1381-9.

46. Watson DK, Robinson L, Hodge DR, Kola I, Papas TS, Seth A. FLI1 and EWS-FLI1 function as ternary complex factors and ELK1 and SAP1a function as ternary and quaternary complex factors on the Egr1 promoter serum response elements. Oncogene. 1997;14(2):213-21.

47. Kim S, Denny CT, Wisdom R. Cooperative DNA binding with AP-1 proteins is required for transformation by EWS-Ets fusion proteins. Mol Cell Biol. 2006;26(7):2467-78.

48. Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia. 2008;10(2):177-88.

3.9 Figures

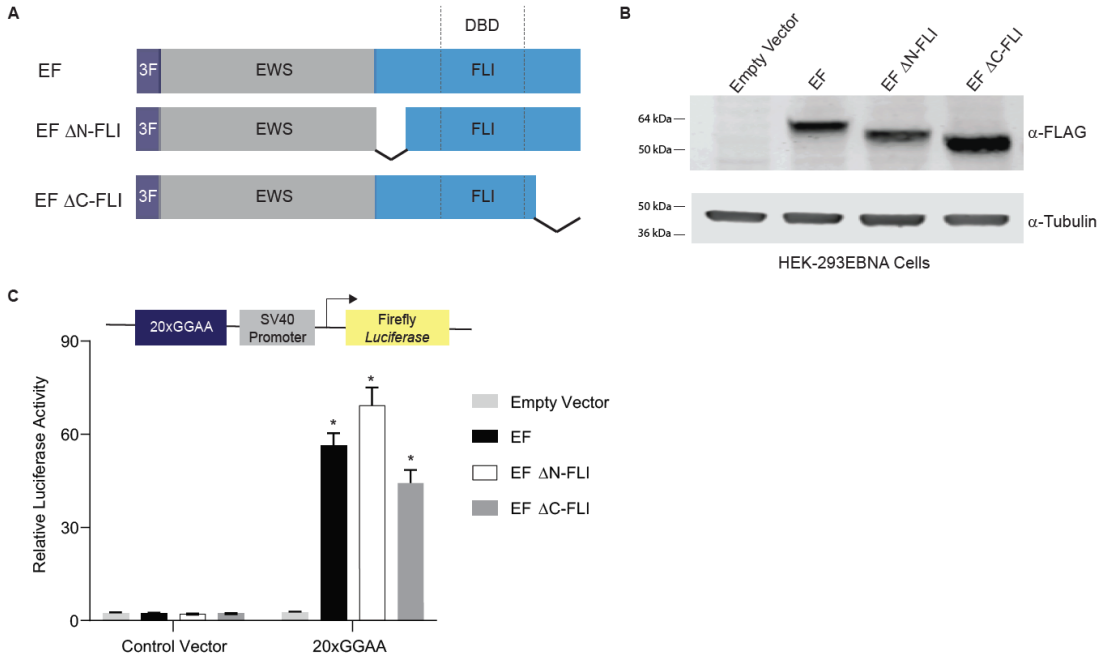

Figure 3.1. Amino- and carboxyl-terminal regions of FLI are not required for EWS/FLI-mediated transcriptional activation.

(A) Protein schematic of 3xFLAG-tagged (3F) EWS/FLI (EF) cDNA constructs. EWS is represented in grey, FLI is represented in blue, and dashed lines in the FLI portion represent the 85-amino acid ETS DNA-binding domain (DBD) of FLI. In each construct, EWS is fused directly to the FLI portion, but connecting lines are shown here to represent regions of FLI that are eliminated in each construct. EF represents a full-length "type IV" EWS/FLI translocation. EF ΔN-FLI and EF ΔC-FLI indicate constructs where EWS was fused to a version of FLI with a deletion in the amino- or carboxyl-terminal region, respectively. (B)

Western blot of 3xFLAG-tagged EWS/FLI protein expression in HEK-293EBNA cells. Membranes were probed with either α-FLAG or α−tubulin antibodies. (C) Dual luciferase reporter assay results for the indicated cDNA constructs co-transfected into HEK-293EBNA cells with a Control Vector harboring no GGAA-repeats, or a vector containing 20xGGAA-repeats (represented above the graph). Data are presented as mean ± SEM (N= 6 biological replicates with 3 technical replicates each). Asterisks indicate that the activity of EF, EF ΔN-FLI, and EF ΔC-FLI are each statistically significant when compared to Empty Vector on a 20xGGAA μSat (p-value < 0.05). The activity of EF ΔN-FLI and EF ΔC-FLI are not statistically different from EF on the 20xGGAA μSat (p-value = 0.8).

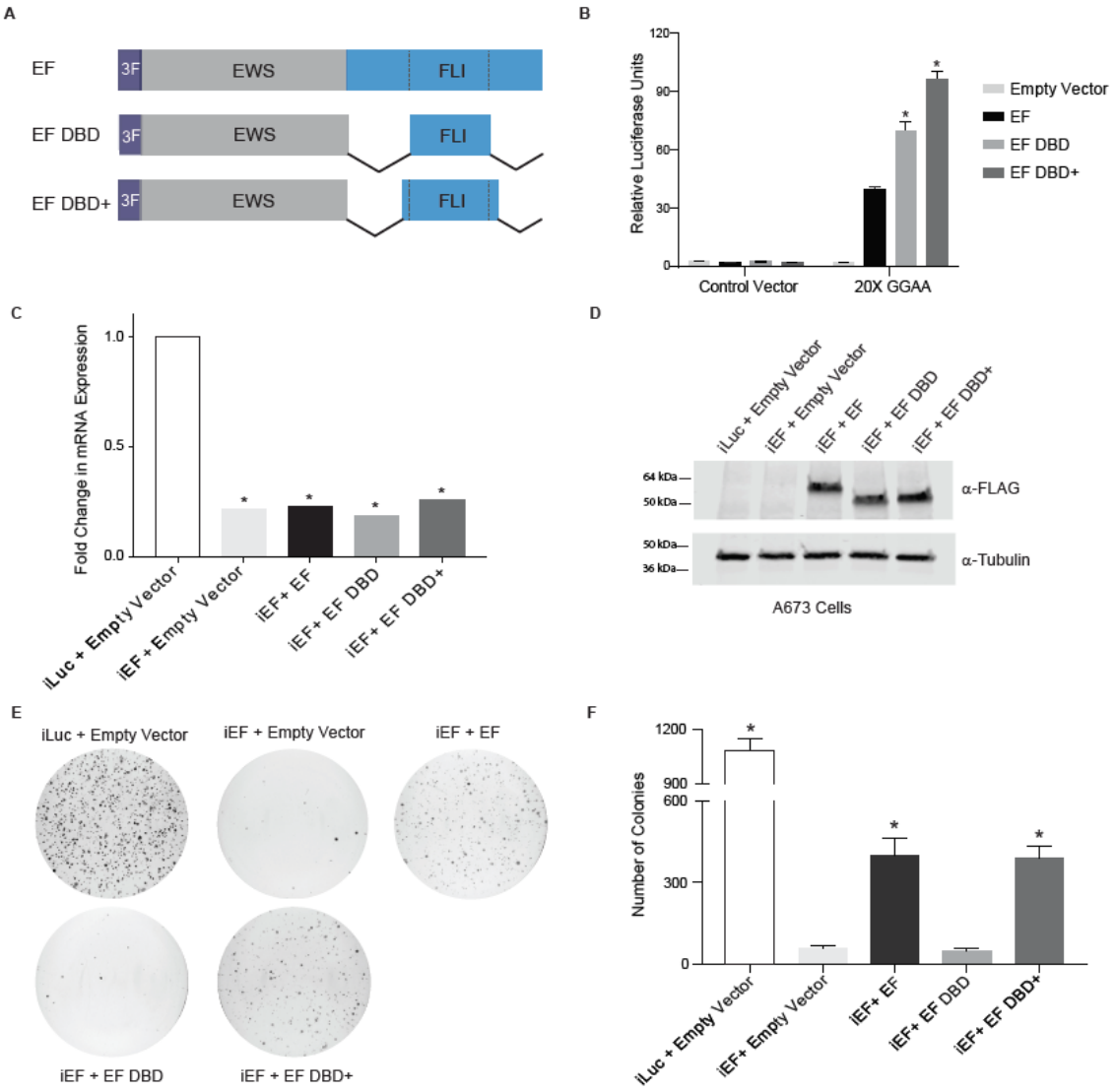

Figure 3.2. Oncogenic transformation capacity of EWS/FLI affected by short regions surrounding the FLI DBD.

(A) Protein schematic of 3xFLAG-tagged (3F) EWS/FLI cDNA constructs with deleted FLI domain regions. EF represents a full-length type IV EWS/FLI, EF DBD represents EWS fused directly to the 85-amino acid DNA-binding domain of FLI, and EF DBD+ represents EWS fused to a 102-amino acid region of FLI that contains the 85 amino-acid

DNA-binding domain with 7 additional amino-acids on the amino-terminal side and 10 additional amino-acids on the carboxyl-terminal side (B) Dual luciferase reporter assay results for the indicated constructs tested on control and 20xGGAA μSat-containing plasmids (as described in Figure 3.1). Data are presented as mean ± SEM (N= 6 biological replicates with 3 technical replicates each). Asterisks indicate that the activity of EF DBD and EF DBD+ are each statistically higher than EF (p-value < 0.001). (C) Representative qRT-PCR results of endogenous EWS/FLI in A673 cells harboring the indicated constructs (iLuc is a control shRNA targeting luciferase and iEF is a shRNA targeting the 3'UTR of endogenous EWS/FLI; N= 1 biological replicate with 3 technical replicates for each sample). EWS/FLI mRNA values are normalized to RPL30 mRNA control values. Asterisks indicate samples are statistically different as compared to control (iLuc + Empty Vector) cells (p-value < 0.001). (D) Western blot analysis of exogenous EWS/FLI protein expression in the A673 knock-down/rescue cells. Protein constructs were detected using α-FLAG antibody and α-Tubulin was used as a loading control. (E) Representative soft agar assay results of A673 Ewing sarcoma cells containing the indicated constructs. (F) Soft agar assay colony formation quantification. Data presented as mean ± SEM (N= 9 biological replicates with 2 technical replicates each). Asterisks indicate p-value <0.001 as compared to iEF + Empty Vector cells.

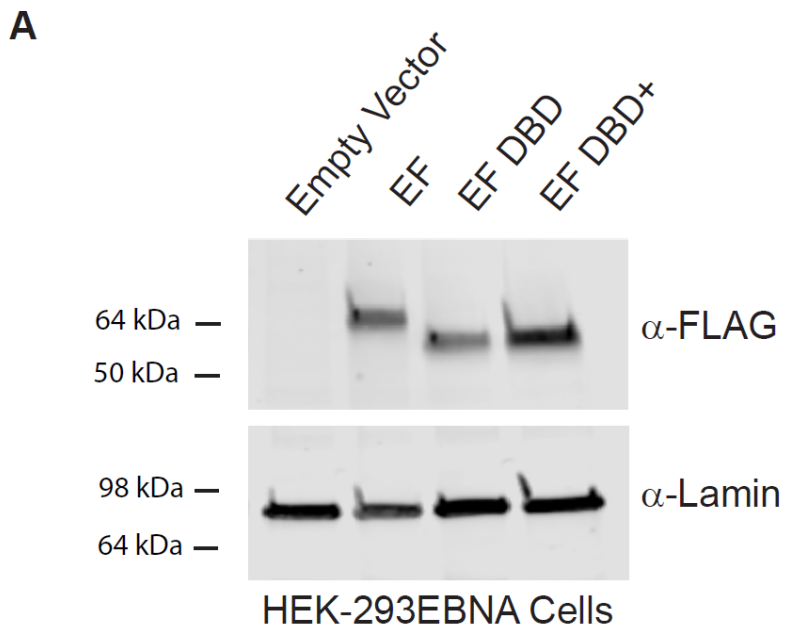

Figure 3.3. EWS/FLI mutant construct expression in HEK-293EBNA cells.

(A) 3xFLAG-tagged full-length EWS/FLI (EF), EF DBD, or EF DBD+ constructs were expressed in HEK-293EBNA cells. Western blot analysis was used to determine expression of these proteins utilizing an α-FLAG antibody. α-Lamin antibody was used as a loading control.

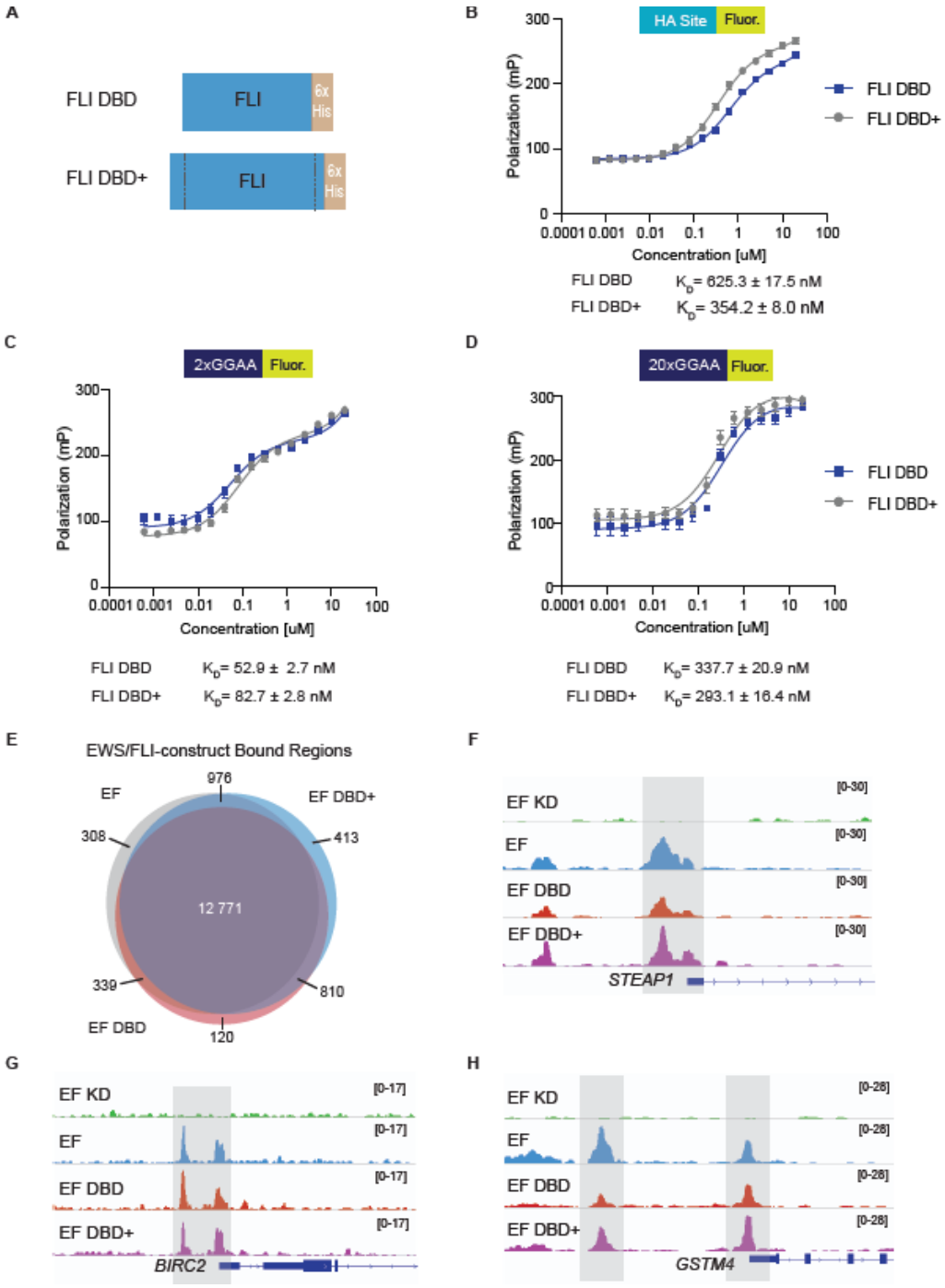

Figure 3.4. DNA-binding and genomic localization properties of EWS/FLI unaltered by deletions flanking the FLI DNA-binding domain.

(A) Protein schematic of FLI DBD or FLI DBD+ recombinant protein (with C-terminal 6xHistidine-tag [6xHis]). (B-D) Fluorescence anisotropy assay results for FLI DBD and FLI DBD+ recombinant proteins (0-20 μM) on 5 nM fluorescein-labeled DNA sequences: (B) ETS high-affinity (HA) site DNA, (C) 2x-repeat GGAA μSat DNA, and (D) 20x-repeat GGAA μSat DNA (N= 2 biological replicates, 3 technical replicates each). Dissociation constants (K_D) for FLI DBD and FLI DBD+ are noted for each DNA response element. (E) Venn diagram comparing peaks called in CUT&RUN for EWS/FLI construct localization in knock-down/rescue cells (EF = iEF + EF; EF DBD = iEF + EF DBD; EF DBD+ = iEF + EF DBD+) when compared to cells that did not contain a rescue construct (iEF + Empty Vector) (adjusted p-value (FDR) < 0.05; N=2 biological replicates each). The number of peaks overlapping between constructs are indicated on the Venn diagram. (F-H) Representative CUT&RUN peak tracks from IGV are shown for iEF + Empty Vector (EF KD), EF, EF DBD, and EF DBD+ samples. Examples of peaks from EWS/FLI-associated HA-site regulated genes ([F] *STEAP1* and [G] *BIRC2*) and GGAA-μSat-regulated genes ([H] *GSTM4*) are highlighted. Peak track scales are shown on the right.

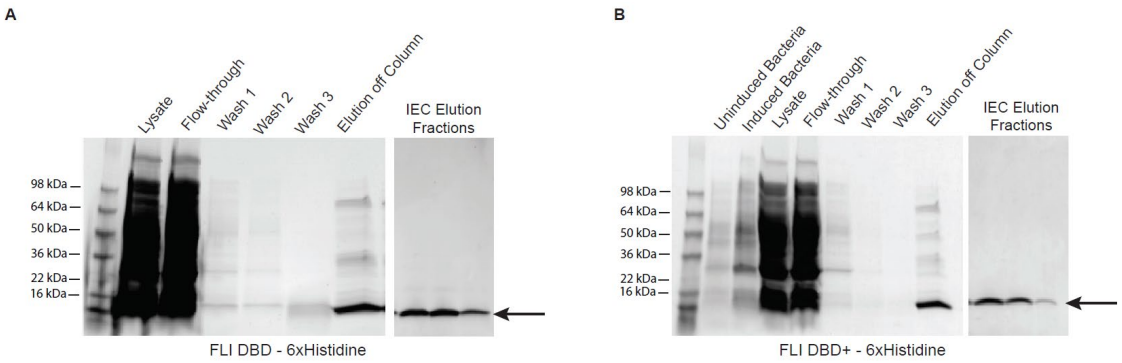

Figure 3.5. Recombinant FLI DBD and FLI DBD+ protein purification.

(A-B) Samples were taken at several stages of recombinant protein purification for (A) FLI DBD - 6xHisitidine and (B) FLI DBD+ - 6xHistidine, including: uninduced bacteria, induced bacteria, after bacterial lysis, after wash 1, wash 2, and wash 3, eluted fraction from the column, and eluted fractions after performing ion-exchange chromatography (IEC). Samples were run on SDS-PAGE gels and visualized using Coomassie blue staining.

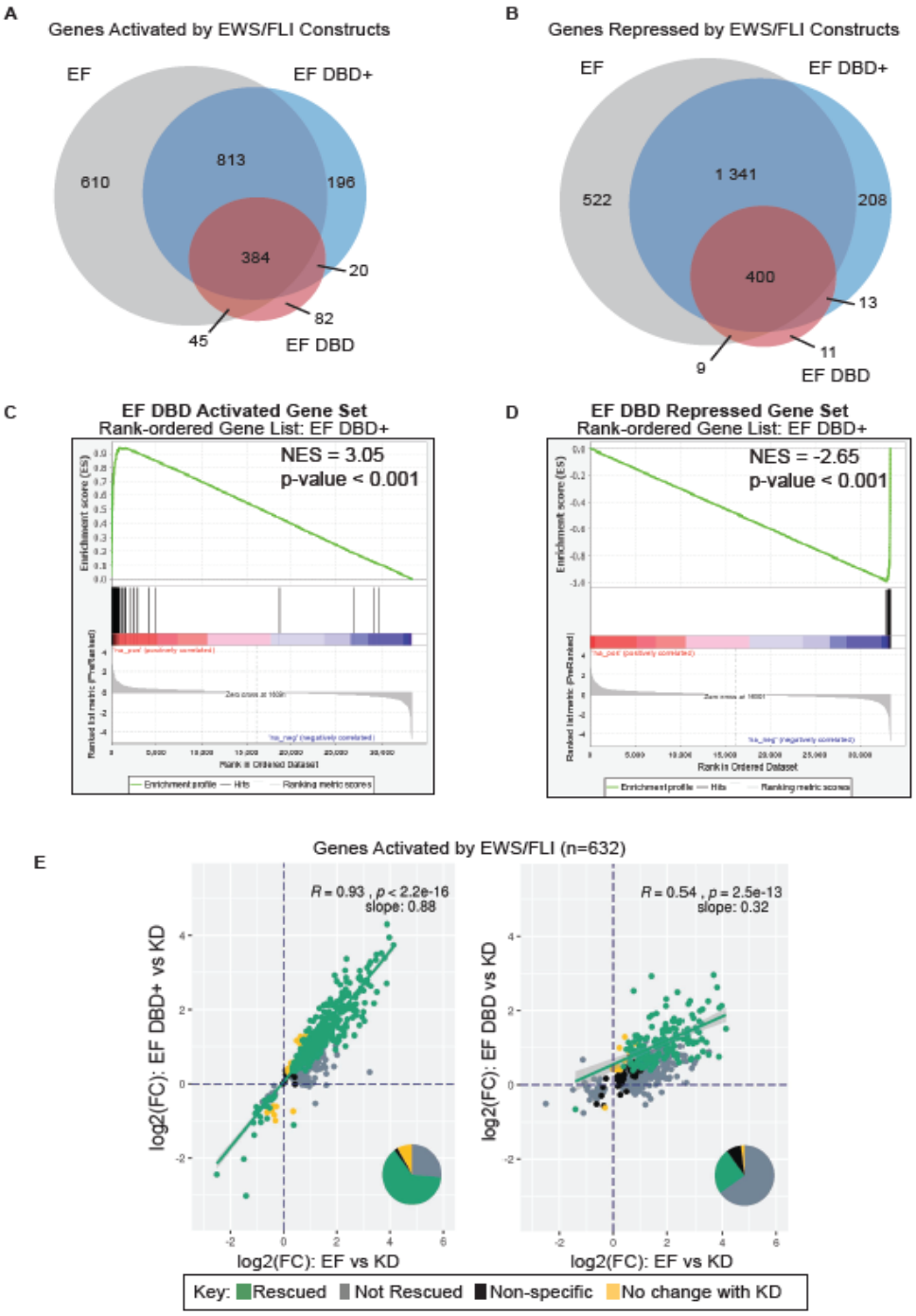

Figure 3.6. EWS/FLI-driven transcriptional regulation diminished by FLI DBD flanking deletions in Ewing sarcoma cells.

(A-B) Venn diagram analysis of RNA-sequencing data comparing genes significantly (A) activated or (B) repressed in A673 cells rescued with the indicated constructs (full-length EWS/FLI [EF], EF DBD, and EF DBD+) when compared to A673 cells with no exogenous EWS/FLI construct (iEF + Empty Vector) (adjusted p-value (FDR) < 0.05; N=3 biological replicates each). (C-D) GSEA analysis comparing all genes regulated by EF DBD+ as the rank-ordered gene list to a gene set of (C) genes activated by EF DBD (log2(FC) > 1.5, FDR < 0.05) or (D) genes repressed by EF DBD (log2(FC) < -1.5, FDR < 0.05) as the gene set. (E) Genes significantly activated by endogenous EWS/FLI were defined using a previous RNA-sequencing dataset (44). Genes activated by EF, EF DBD, and EF DBD+ in A673 knock-down/rescue cells were compared to this list of EWS/FLI-activated genes. Scatterplots comparing genes activated by EF (on the x-axis) to EF DBD+ (left) or EF DBD (right) (on the y-axis) were plotted to determine the ability of these constructs to rescue expression these genes. Significance was defined by a log2(FC) > 0 and an adjusted p-value < 0.05. Pearson correlation coefficient and associated p-values with slope are noted on the plots. Pie charts represent the proportion of genes found in each of the described groups.

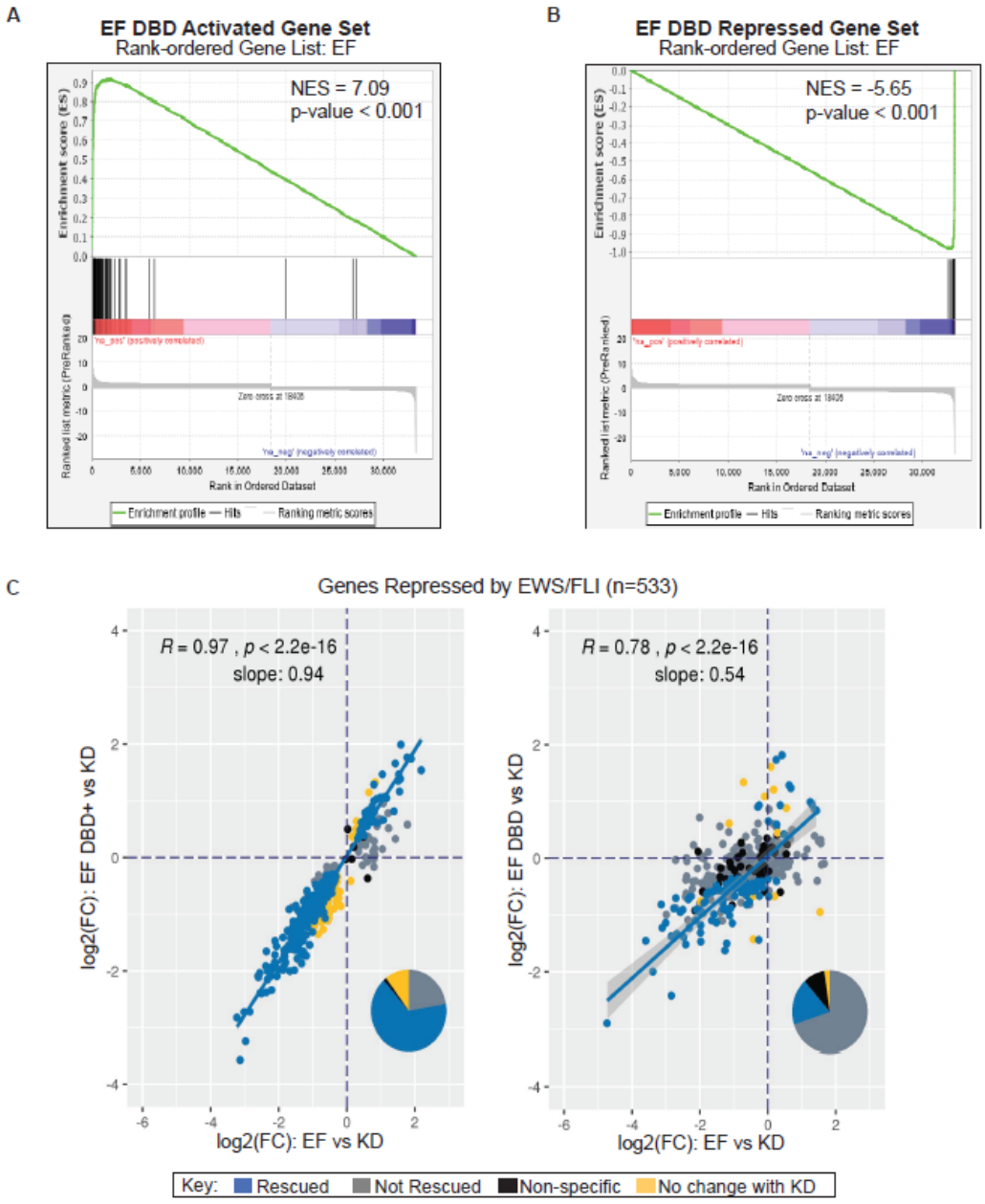

Figure 3.7. Deletions surrounding the FLI DBD of EWS/FLI result in weaker transcriptional regulation.

(A-B) GSEA analysis comparing genes regulated by EF as the rank-ordered gene list to a gene set of (A) activated genes by EF DBD (log2(FC) > 1.5, FDR < 0.05) or (B) repressed

genes by EF DBD (log2(FC) < -1.5, FDR < 0.05). (C) Genes significantly repressed by endogenous EWS/FLI were defined using a previous RNA-sequencing dataset (44). Genes repressed by EF, EF DBD, and EF DBD+ in A673 knock-down/rescue cells were compared to this list of EWS/FLI-repressed genes. Scatterplots comparing genes repressed by exogenous EF (on the x-axis) to EF DBD+ (left) or EF DBD (right) (on the y-axis) were plotted to determine the ability of these constructs to rescue repression of these genes. Significance was defined by a log2(FC) < 0 and an adjusted p-value < 0.05. Pearson correlation coefficient and associated p-values with slope are noted on the plots. Pie charts represent the proportion of genes found in each of the described groups.

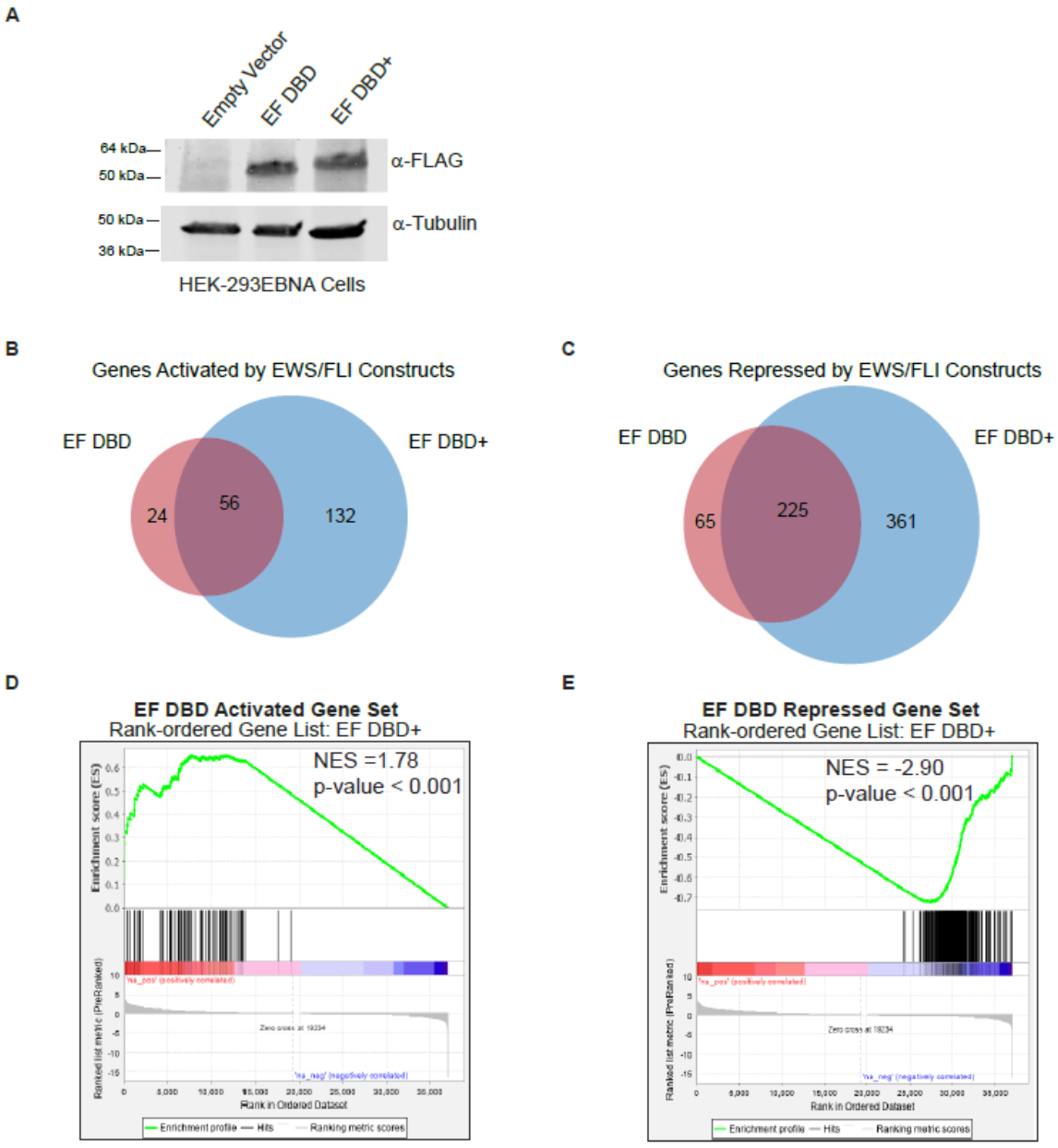

Figure 3.8. EF DBD exhibits hypomorphic ability to drive transcription in alternative HEK-293EBNA cell model.

(A) Western blot analysis of EF DBD and EF DBD+ cDNA-containing constructs transfected into HEK-293EBNA cells before collection for RNA-sequencing analysis. α-

FLAG and α−tubulin were used to probe blots to ensure equivalent protein expression for each protein. (B-C) Venn diagram overlap analysis was performed to determine the number of significantly regulated activated and repressed genes for both EF DBD and EF DBD+ in the HEK-293EBNA cells (N=3 biological replicates each). Significantly regulated genes were determined using a FDR cut-off of 0.05. (D-E) GSEA analysis comparing genes regulated by EF DBD+ as the rank-ordered gene list to a gene set of (D) activated genes by EF DBD (log(2)FC > 0, FDR < 0.05) or (E) repressed genes by EF DBD (log2(FC) < 0, FDR < 0.05).

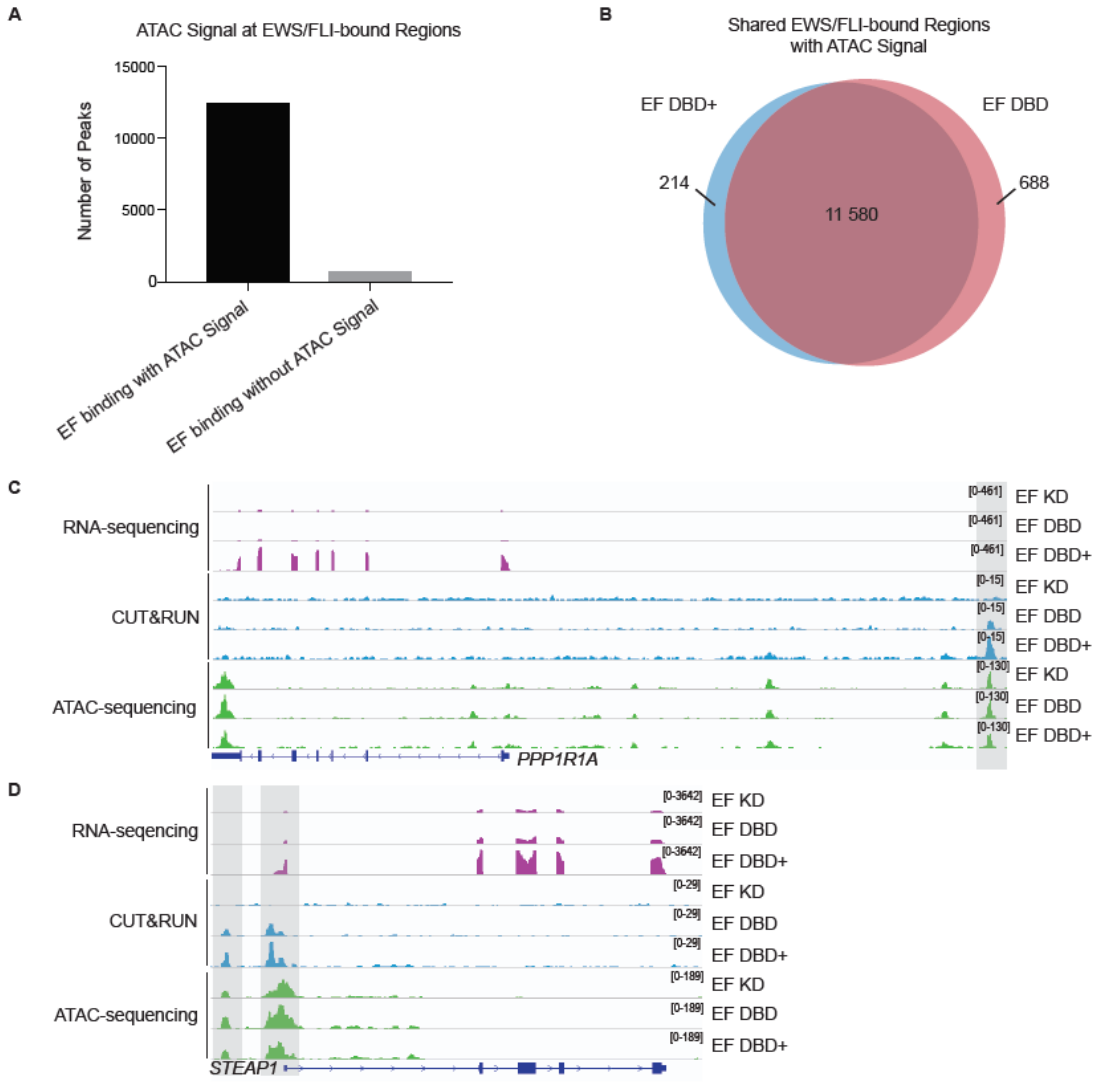

Figure 3.9. Chromatin-opening ability of EWS/FLI is unaltered by deletions flanking the FLI DNA-binding domain.

(A) All EWS/FLI-bound loci in A673 cells (determined by CUT&RUN of knock-down/rescue cells expressing full-length EWS/FLI [EF]) were compared to loci harboring ATAC signal peaks and shown in graphical format (ATAC performed on N=2 biological replicates each). There were 12 482 EF-bound peaks with ATAC signal and 696 EF-bound

peaks without ATAC signal. (B) Venn diagram analysis of regions bound by EF DBD+ and/or EF DBD that also had overlapping ATAC signals. (C-D) Representative tracks of RNA-sequencing, CUT&RUN genomic localization, and ATAC-sequencing signals for the indicated knock-down/rescue A673 cells (EF KD = iEF + EF; EF DBD = iEF + EF DBD; EF DBD+ = iEF + EF DBD+). Scales to view tracks were kept consistent across sequencing type in each panel and are represented on the right. Representative genes *PPP1R1A* (C) and *STEAP1* (D) are regulated by EF DBD+ but not EF DBD (adjusted p-value <0.05) and overlapping CUT&RUN and ATAC-sequencing peaks are highlighted.

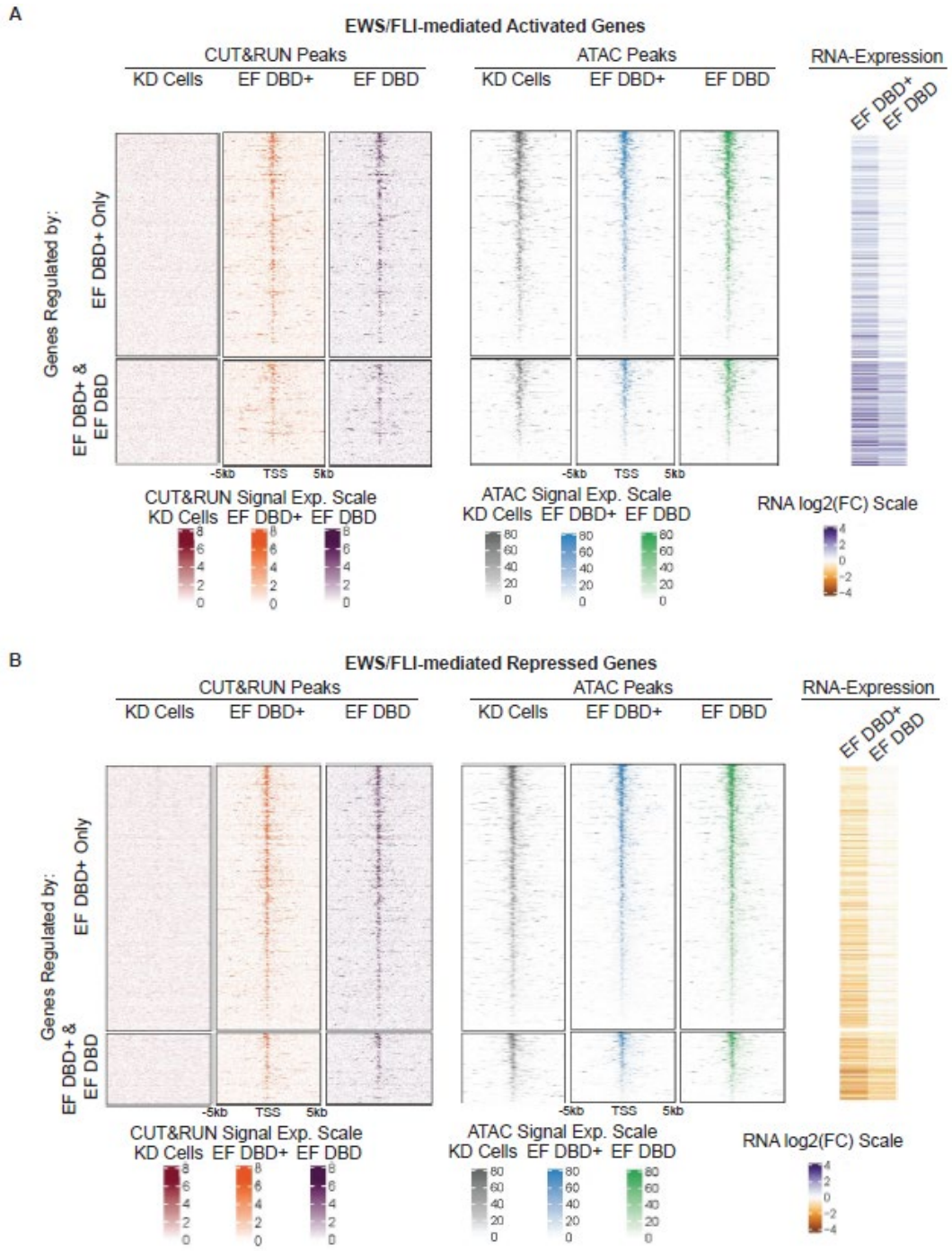

Figure 3.10. Heatmap analysis of ATAC-signal at EWS/FLI-mutant bound activated and repressed genes.

(A-B) Heatmaps depicting CUT&RUN and ATAC-sequencing signals, centered on the nearest transcriptional start site (TSS) of genes regulated by EF DBD+ only or both EF DBD+ and EF DBD. EWS/FLI-mediated activated genes are visualized in (A) and repressed genes in (B). Knock-down cells (KD; iEF + Empty Vector), EF DBD+ (iEF + EF DBD+), and EF DBD (iEF + EF DBD) were compared (scales for peak height are depicted below heatmaps). The log2(FC) of RNA expression for EF DBD+ and EF DBD (compared to KD) is pictured on the right with log2(FC) scale depicted below.

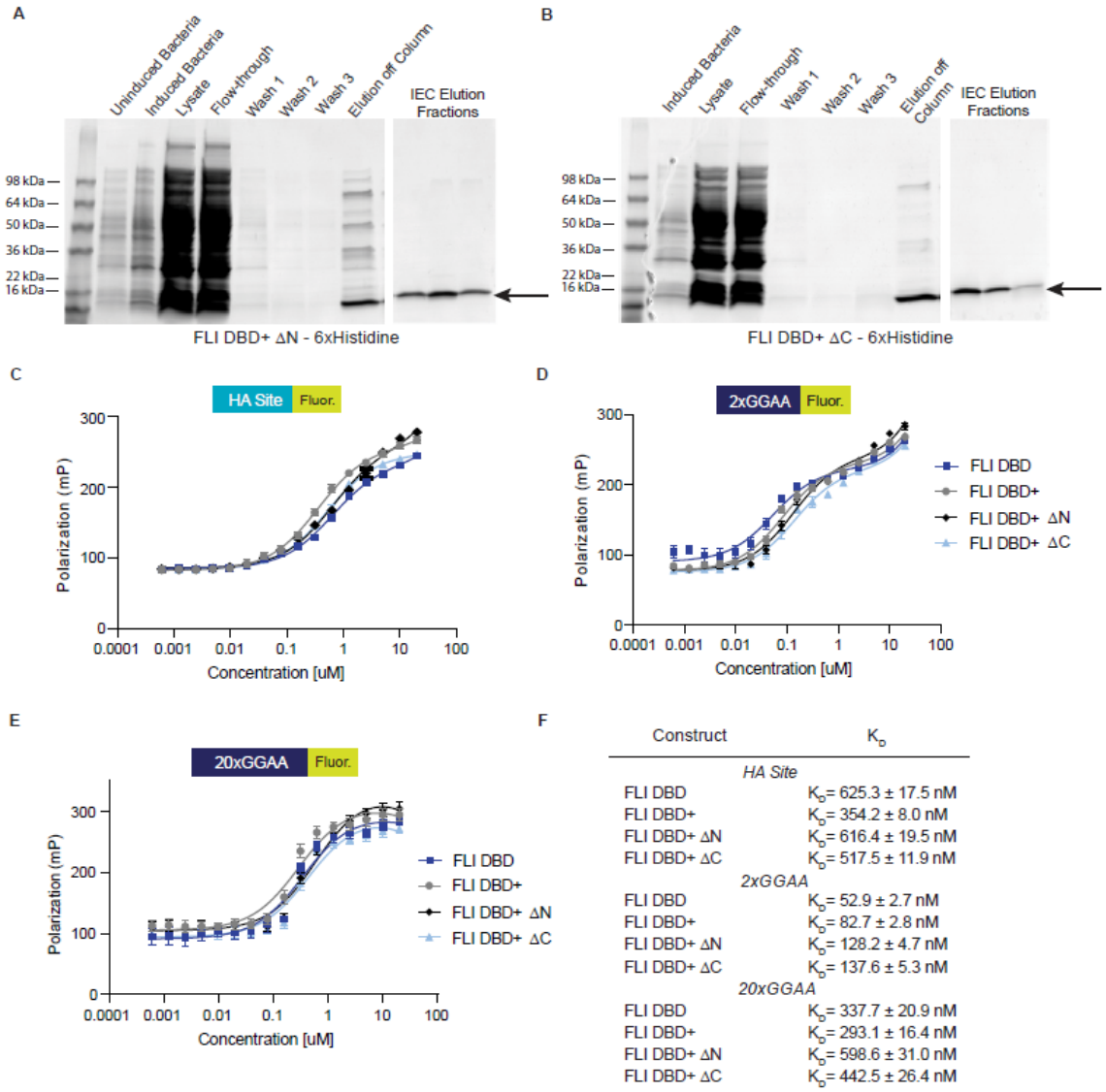

Figure 3.11. Flanking regions outside of the FLI DNA-binding domain are not crucial for the ability of FLI to bind DNA.

(A-B) Coomassie blue staining was performed on SDS-PAGE gels to analyze FLI DBD+ ΔN (B) and FLI DBD+ ΔC (C) recombinant protein purification. Samples were taken during various stages of purification, including eluted fractions after ion exchange chromatography (IEC). (C-E) Fluorescence anisotropy was performed using 0-20 μM

recombinant protein and 5 nM of fluorescein-labelled DNA duplex on the following DNA sequences: (C) ETS HA site, (D) 2xGGAA-μSat, and (E) 20xGGAA-μSat. Each measurement was completed with 2 biological replicates, with 3 technical replicates in each. (F) Binding dissociation constants (K_D) of recombinant FLI proteins were determined from fluorescence anisotropy data and quantification recorded here.

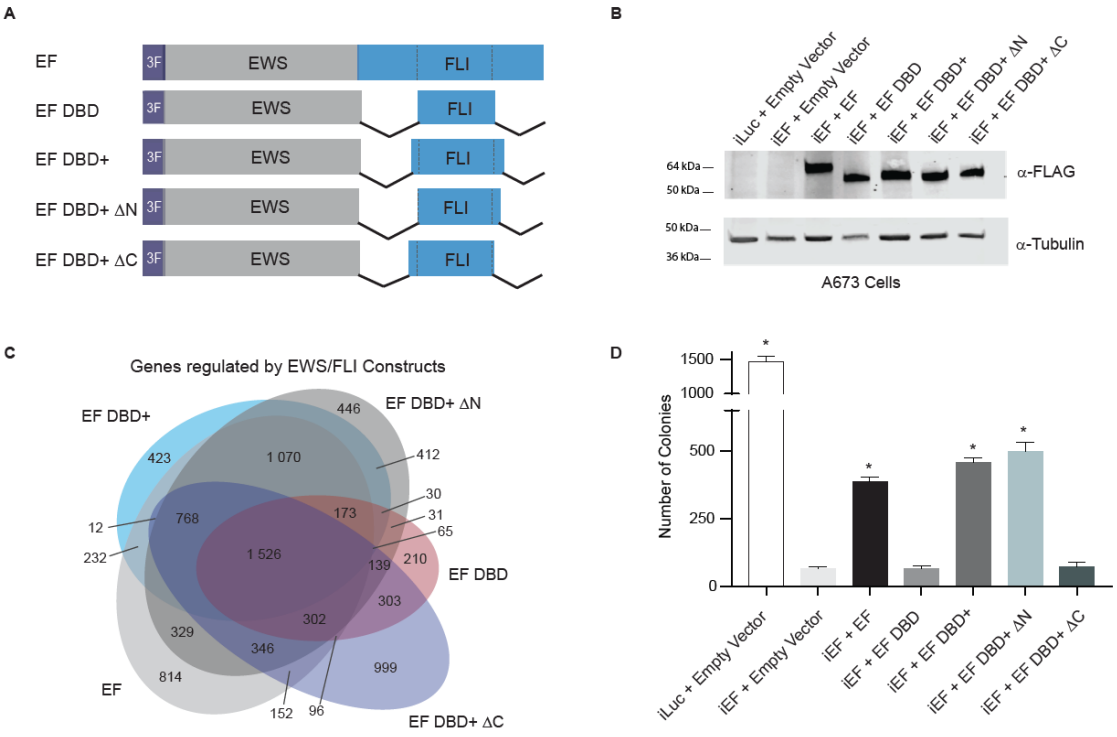

Figure 3.12. The carboxyl-terminal amino acids flanking the FLI DNA-binding domain are essential for EWS/FLI-mediated oncogenic transformation.

(A) Protein schematics of 3xFLAG-tagged (3F) EWS/FLI constructs: EF, EF DBD and EF DBD+ are described in Figure 3.2. EF DBD+ ΔN represents an EWS/FLI mutant where EWS is fused to the DBD+ version of FLI missing the 7-amino-terminal amino acids to the DNA-binding domain; EF DBD+ ΔC represents an EWS/FLI mutant where EWS is fused to the DBD+ version of FLI missing the 10 carboxyl-terminal amino acids to the DNA-binding domain. (B) Western blot analysis of constructs expressed in A673 cells using our knock-down/rescue system. (C) Venn diagram overlap analysis of RNA-sequencing results (N=3 biological replicates each). Overlap depicts genes called as

significantly regulated by the listed construct compared to control cells (iEF + Empty Vector Cells). Genes were called as significantly regulated using an FDR cut-off of 0.05 (p-value of overlap < 2.2e-16). (D) Soft agar assay colony formation quantification of A673 knock-down/rescue cells containing the listed knock-down and rescue constructs. Data represented by mean ± SEM (N= 3 biological replicates with 2 technical replicates each). Asterisks indicate p-value < 0.05 as compared to negative control iEF + Empty Vector sample with no EWS/FLI expression.

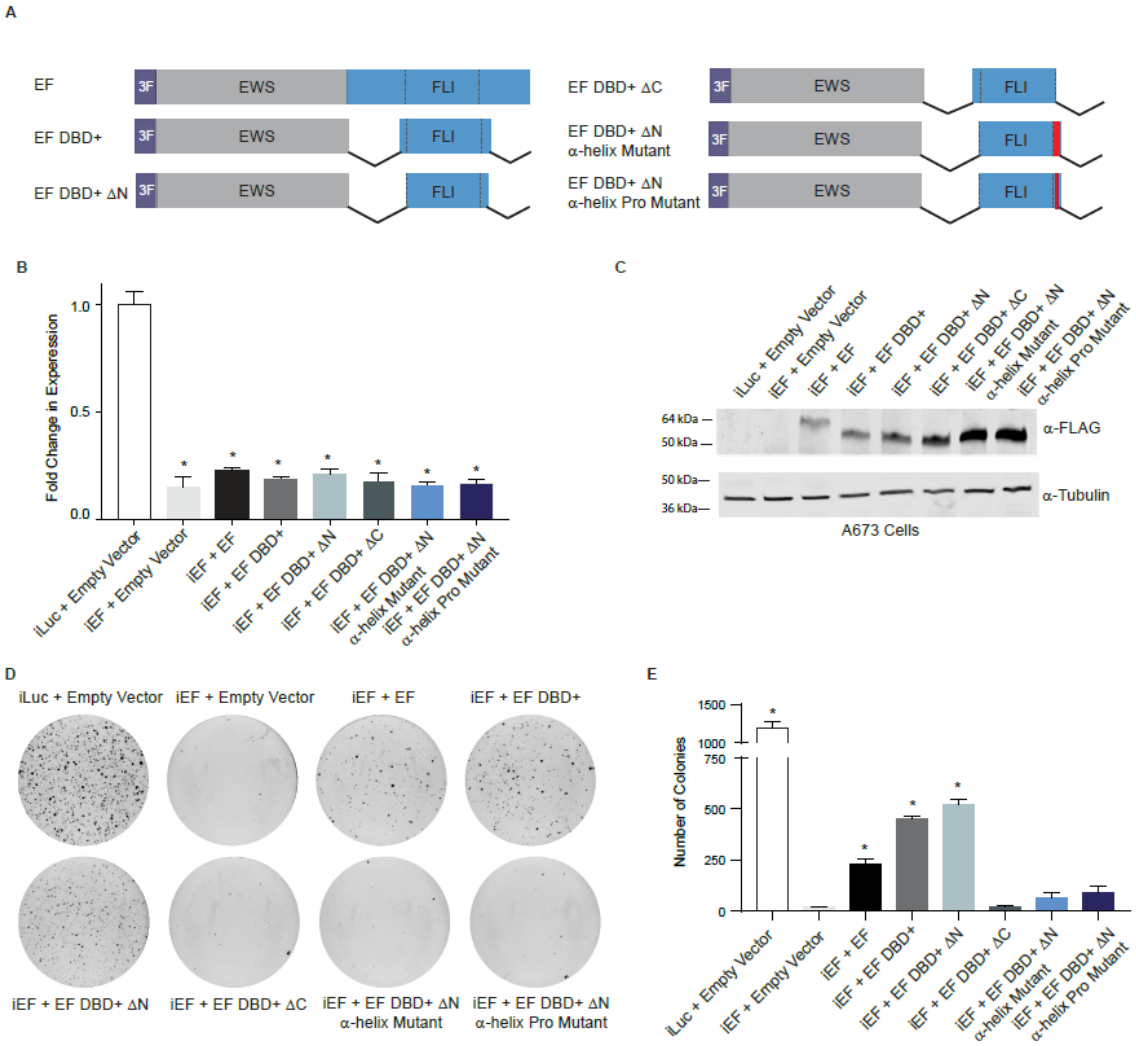

Figure 3.13. Disruption of the fourth alpha-helix of the FLI DNA-binding domain results in loss of oncogenic transformation potential of EWS/FLI.

(A) Protein schematics of 3xFLAG-tagged EWS/FLI cDNA constructs, including versions of EF DBD+ ΔN that have mutations in the fourth alpha-helix of the DNA-binding domain of FLI. EF DBD+ ΔN α-helix Mutant includes four amino acid substitutions and EF DBD+ ΔN α-helix Pro(line) Mutant includes five proline-specific amino acid substitutions, both

of which are predicted to disrupt the alpha-helical structure of this region. (B) Representative qRT-PCR results depicting knock-down of endogenous EWS/FLI mRNA in the A673 knock-down/rescue system; completed using EWS/FLI-specific primers and normalized to RPL30 mRNA values for each sample. Data are presented as mean ± SEM (N= 1 biological replicate with 3 technical replicates). Asterisks indicate significant knock-down of EWS/FLI as compared to iEF + Empty Vector control cells (p-value < 0.05). (C) Western blot analysis of A673 knock-down/rescue samples: α-FLAG used to probe for expression of the cDNA constructs and α-tubulin used as loading control. (D) Representative soft agar assay results showing oncogenic transformation potential of EWS/FLI cDNA constructs in A673 knock-down/rescue cells. (E) Quantification of colonies in soft agar assays represented by mean ± SEM (N= 3 biological replicates with 2 technical replicates each). Asterisks indicate significant transformation over control cells (iEF + Empty Vector) (p-value < 0.05).

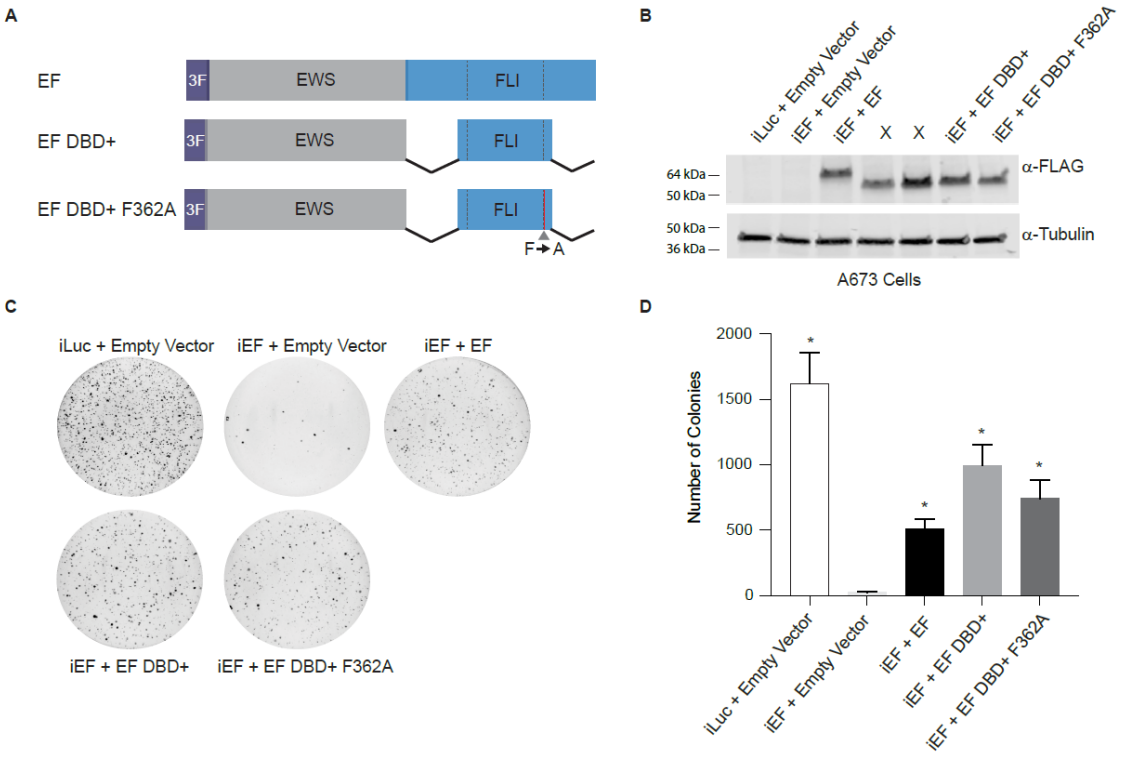

Figure 3.14. Homodimerization motif is dispensable for EWS/FLI-mediated oncogenic transformation.

(A) Protein schematics of 3xFLAG-tagged (3F) EWS/FLI constructs: EF represents full-length EWS/FLI; EF DBD+ contains EWS fused to the DBD+ version of FLI; EF DBD+ F362A represents EWS fused to the DBD+ version of FLI with the phenylalanine (F) residue at residue 362 mutated to alanine (A). (B) Constructs were expressed in A673 cells using our knock-down/rescue system and Western blot analysis was used to determine efficient expression of these proteins using α-FLAG antibody for detection of EWS/FLI constructs and α-Tubulin as a loading control. (Samples labeled as X are not relevant to this current set of experiments.) (C) Representative soft agar assay results are depicted,

including controls and experimental samples. (D) Soft agar assay colony formation quantification. Data is represented by mean ± SEM (N= 3 biological replicates with 2 technical replicates each). Asterisks indicate p-value < 0.005 as compared to samples with EWS/FLI knock-down (iEF + Empty Vector).

3.10 Tables

Table 3.1. Amino acids references for EWS/FLI or FLI recombinant protein-encoding constructs.

	Amino acids residues according to protein:		
Construct Name	**EWSR1** (NP_001156757.1)	**FLI1** (NP_002008.2)	**"Type IV" EWS/FLI** (7)
Full-length "Type IV" EF	1-265	242-452	266-476
EF ΔN-FLI	1-265	275-452	299-476
EF ΔC-FLI	1-265	242-373	266-397
EF DBD	1-265	277-361	302-386
EF DBD+	1-265	270-371	295-396
FLI DBD	-	277-361	302-386
FLI DBD+	-	270-371	295-396
FLI DBD+ ΔN	-	277-371	302-396
FLI DBD+ ΔC	-	270-361	295-386
EF DBD+ ΔN	1-265	277-371	302-396
EF DBD+ ΔC	1-265	270-361	295-386
EF DBD+ F362A	1-265	270-371, F362A	295-396, F362A
EF DBD+ ΔN α-helix Mutant	1-265	277-371, H363D, I365S, Q367D, L369S	302-396, H388D, I390S, Q392D, L394S
EF DBD+ ΔN α–helix Pro Mutant	1-265	277-371, G364P, I365P, G366P, R367P, L368P	302-396, G389P, I390P, G391P, R392P, L393P

Table 3.1. Amino acids references for EWS/FLI of FLI recombinant protein-encoding constructs.

Amino acid residues composition of each EWS/FLI cDNA or FLI recombinant protein-encoding construct are denoted here, according to the residue number in native EWSR1 (NP_001156757.1), native FLI1 (NP_002008.2), or "type IV" EWS/FLI translocation (identified by May et al., 1993).

Table 3.2. Sequences for primers used in qRT-PCR experiments

Gene	Forward Primer	Reverse Primer
EWS/FLI	5'-CAGTCACTGCACCTCCATCC	5'-TTCATGTTATTGCCCCAAGC
RPL30	5'-GGGGTACAAGCAGACTCTGAAG	5'-ATGGACACCAGTTTTAGCCAAC

Table 3.2. Sequences for primers used in qRT-PCR experiments.

Forward and reverse primer sequences used for qRT-PCR experiments to determine levels of endogenous EWS/FLI or RPL30 mRNA.

Table 3.3. Fluorescein-labeled DNA-duplex oligonucleotides used for fluorescence anisotropy experiment.

DNA Duplex	Sequence (forward strand of duplex)
High Affinity (HA) Site	/56FAM/TTTACCGGAAGTGTT
2X GGAA	/56FAM/TTTGGAAGGAATTT
20X GGAA	/56FAM/ TTTGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG GAATTT

Table 3.3. Fluorescein-labeled DNA-duplex oligonucleotides used for fluorescence anisotropy experiments

Forward sequences of fluorescein-labeled DNA-duplex oligonucleotides used for fluorescence anisotropy experiments.

Chapter 4:

Conclusions and Future Directions

4.1 Conclusions

Ewing sarcoma is an aggressive pediatric bone- and soft tissue-associated cancer that is characterized by a chromosomal translocation involving a FET and ETS protein member, which encodes an oncogenic transcription factor responsible for driving disease (1-4). In this body of work, we sought to determine novel properties and functions that can be attributed to either the FET or ETS domain found in Ewing sarcoma-associated chromosomal translocations. Accordingly, chapter 2 describes the characterization of a novel FUS/ETV4 translocation detected in a Ewing sarcoma patient, as well as the first study to compare genomic localization and transcriptional regulatory functions of the broad FET/ETS translocation group. Chapter 3 describes results of a structure-function mapping study employed to determine functions that the FLI domain contributes to EWS/FLI function. Results discussed herein outline the novel findings that FET/ETS translocations possess similar biological functions in Ewing sarcoma cells and that an alpha-helix immediately downstream of the ETS DNA-binding domain contributes to transcriptional regulatory properties observed for the EWS/FLI fusion protein. Chapter 4 will discuss these findings in the context of the Ewing sarcoma field, implications of each, and future directions.

4.1.1 Clinical uncertainty surrounding diagnosis and treatment of FET/ETS fusions

Despite report of various alternative FET/ETS fusions in Ewing sarcoma patients, the vast majority of studies focus on elucidating functional properties of the EWS/FLI fusion. Although the EWS/FLI fusion is detected in 85-90% of Ewing sarcoma cases, this statistic may not accurately represent the number of patients whose tumors harbor EWS/FLI or alternative FET/ETS fusions: The EWS and FLI proteins were first shown to be involved in the chromosomal translocation observed in a majority of patient tumors in 1992, following the advent of PCR technologies (1). Today, patient tumor samples obtained from biopsy are often quite small and fixed in formaldehyde for analysis and preservation. Due to the inadequate tumor sample size and poor RNA quality, the utility of performing PCR on these samples has been limited (5). As such, the gold standard for translocation detection in suspected Ewing sarcoma tumors is break-apart fluorescence in-situ hybridization (FISH) (2, 5-9). In many cases, FISH probes specifically target the *EWSR1* gene to determine if a translocation has occurred there and if so, it is typically assumed this is indicative of the presence of a EWS/FLI translocation (2, 6-9). Probing only for *EWSR1* likely misses alternative ETS member fusions, though. A survey published by the Children's Oncology Group found that a majority of clinicians believe that sequencing of the tumor was only necessary if a translocation was present that did not involve *EWSR1* (10). Even is sequencing is desired, a lack of availability of sequencing services and/or lack of insurance coverage for additional diagnostic tests may prevent it from being performed. This will likely also contribute to non-EWS/FLI fusions going undetected.

Because of inconsistencies in Ewing sarcoma tumor sample analysis, it is likely that the proportion of alternative FET/ETS fusions may make up a larger percentage of patient-detected translocations than what is currently reported. Furthermore, it is also likely that other yet-to-be-discovered FET/ETS fusions exist in Ewing sarcoma patients. Herein, we have described a novel FUS/ETV4 fusion in a neonatal Ewing sarcoma patient. As sequencing services become more widely available and will some become recommended for diagnosis, additional FET/ETS fusions will surely be detected as well. As much uncertainty revolves around how these alternative EWS/ETS and FUS/ETS fusions should be diagnosed and treated, guidance on how alternative FET/ETS fusions should be regarded is of critical importance.

4.1.2 FET/ETS fusions share similar biological functions in Ewing sarcoma cells

To address functional questions surrounding alternative FET/ETS fusion protein activity, we sought to perform a comparative analysis of crucial protein properties typically assigned to EWS/FLI, including DNA-binding and transcriptional regulation. These studies have revealed that the broad FET/ETS fusion group bind to similar DNA sequences, including the ETS high affinity site and GGAA-microsatellites that are critical for EWS/FLI-driven oncogenesis. We also observed that transcriptional regulation of a large number of gene targets are also shared between FET/ETS fusions, suggesting that the general FET/ETS fusion class is capable of driving Ewing sarcomagenesis through similar mechanisms. Although we hypothesized that FET/ETS fusions would share some degree

of functional overlap, the high degree of functional similarity observed in both DNA-binding and transcriptomic studies was unexpected. Indeed, it was quite surprising to find that FET and ETS fusion partners were largely interchangeable in this context. This finding implies that on a gene regulatory level FET/ETS fusions should be considered *bona fide* Ewing sarcoma translocations and treated as such.

Although this study investigated FET/ETS fusions in a solitary Ewing sarcoma cell line, these results indicate that similar DNA-binding and transcriptional regulatory properties may be shared across cell type and tumor location. Therefore, we propose a model in which FET/ETS chromosomal translocations expressed in the Ewing sarcoma cell-of-origin will bind similar DNA sequences and regulate similar gene targets to ultimately drive Ewing sarcomagenesis (Figure 4.1).

4.1.3 The ETS DNA-binding domain of FLI contributes a novel transcriptional regulatory function to EWS/FLI

ETS transcription factors have been implicated in the disease etiology of numerous solid and liquid cancers types. As previous studies have found the flanking regions surrounding the DNA-binding domain of several ETS factors are involved in both protein-protein interactions and modulation of DNA-binding activity, we sought to investigate roles of the DNA-binding domain-flanking sequences of FLI in EWS/FLI activity. By utilizing a structure-function mapping approach and various in vitro and in vivo techniques, we found

that a fourth alpha-helix immediately downstream of the ETS DBD of FLI is crucial for EWS/FLI-driven transcriptional regulation and oncogenic transformation. Structural analysis of FLI and other ETS factors reveals that this alpha-helix is conserved largely throughout the entire family. A high degree of conservation here likely indicates that it is crucial for the function of this group of transcription factors.

Supporting this theory, several published studies have reported protein interactions that occur with the DNA-binding domain or regions immediately flanking it that are critical for ETS-regulated transcription (11-15). Additionally, it was recently reported that the same region immediately downstream of the DNA-binding domain of ETS1 binds and alters DNA conformation to allow binding of FOXO1, a key transcription factor often translocated with PAX3 in alveolar rhabdomyosarcoma (11). Without modulation of DNA conformation by ETS1, FOXO1 has a low binding affinity to these DNA sequences and consequently, the transcriptional profile of each protein is significantly altered (11). Unfortunately, analysis of these reported ETS-family protein interactors reveals that either these interactions do not occur with EWS/FLI or the interactions occur outside of the FLI DNA-binding domain. Although these studies cannot directly explain the loss of transcriptional regulation observed for EWS/FLI in our study, the data does provide evidence to further support the hypothesis that a critical protein interaction is occurring here. As such, we propose a model in which protein-protein interactions are occurring with this DBD-flanking alpha-helix of FLI, and that these interactions are crucial for EWS/FLI function and Ewing sarcoma tumorigenesis (Figure 4.2).

4.1.4 Structural similarity between FLI and other ETS factors indicate findings may be generalizable

As a structural analysis revealed that the fourth alpha-helix identified herein is conserved across all ETS factors involved in Ewing sarcoma and prostate cancer translocations, this indicates that protein interactors or ETS-DNA interactions may also be shared across these fusions. Deletion or mutation of this alpha-helix results in the loss of oncogenic transformation potential of EWS/FLI. As such, this finding supports that small molecule inhibitors targeted to this region may be useful in the treatment of all FET/ETS-driven Ewing sarcoma tumors, as well as prostate cancer tumors.

4.2 Future Directions

The findings reported herein have clear implications to warrant future studies that will further contribute to knowledge surrounding Ewing sarcoma etiology, with the ideal outcome that this knowledge will lead to the development of targeted therapeutics to better patients outcomes for those diagnosed with this aggressive disease.

4.2.1 Analysis of FET/ETS fusion protein function in patient tumor samples

The FET/ETS fusion proteins studied here exhibited similar function in our A673 Ewing sarcoma cellular model. Though this is a patient-derived Ewing sarcoma model, it is also important to account for tumoral heterogeneity, as each patient's tumor microenvironment and molecular landscape is unique. As such, the DNA-bound regions and transcriptomic profile of these respective fusions should be also investigated in their "native" environments through analysis of patient tumor samples. Although it is likely that similar DNA-binding function will be retained in these native tumor samples, it is important to note that studies of gene expression profiling in several Ewing sarcoma cell lines have reported differences in the observed transcriptomic profiles of each (16). As we now know that the FET/ETS fusions are capable of regulating similar genes in the same cellular context, it possible that FET/ETS fusions may regulate only a "core" set of essential gene targets across tumor samples and that the observed diversity across cell types may be due to differential regulation of non-essential targets. If this is the case, cross-comparison of transcriptomic data of numerous patient samples may reveal what these essential targets are. Thus, these studies can potentially identify FET/ETS-gene or protein dependencies that may prove to be therapeutic vulnerabilities for further study.

4.2.2 Investigation of Ewing sarcoma translocation incidence across patient populations

Because of the overall low incidence of pediatric cancer, the patient populations in which to study these diseases is quite small for some cancer types, including Ewing sarcoma. It is estimated that only 400-600 patients are diagnosed with Ewing sarcoma annually in the

United States, thereby significantly inhibiting our ability to perform large-scale genomic studies (2, 3, 9, 17, 18). As those who are affected by the disease do face poor outcomes, it is crucial that patient data and samples be used efficiently to study the disease. This is of particular importance as racial and ethnic differences have been observed in Ewing sarcoma incidence rates (2, 3, 19). These differences can partially be attributed to the presence of GGAA-microsatellites of a particular length, but alternative trends that affect disease frequency have yet to be elucidated (19). This further highlights the importance of studying the FET/ETS fusions in a native tumor background to detect any potential biological differences that may exist between patient populations that contribute to Ewing sarcomagenesis.

Interestingly, our collaborators from New Zealand (NZ), who detected the novel FUS/ETV4 patient translocation, report that the indigenous Māori group of NZ have higher incidence rates of Ewing sarcoma than other local populations (unpublished data). Additionally, they report that FUS/ETS fusions are detected more often in Māori people (unpublished data). The higher incidence rates observed may be casual relationships, or there could be a significant link between translocation status and Ewing sarcomagenesis in these patients. As patient tumor sequencing and biobanking of samples are now becoming common practice, there is a clear need for the investigation of FET/ETS translocation occurrence across patient populations. These studies may reveal that certain groups are more prone to the chromoplectic events that drive chromosomal translocation occurrence in Ewing sarcoma, whether that be due to genetic predisposition or environmental stressors.

It is also possible that particular cell types are more tolerant of FET/ETS fusion proteins after translocation events occur. Similar to those performed here, DNA-binding and RNA-sequencing studies would be immediately useful to profile similarities and differences of patient tumors that could explain the observed trends in disease formation. Detection, collaboration and data sharing of Ewing sarcoma patient samples containing one of these rare, alternative FET/ETS fusions is crucial to establish a better understanding of disease biology across patient populations.

4.2.3 Establishing a definitive role for the fourth alpha-helix of the ETS DNA-binding domain

As deletion of an alpha-helix flanking the FLI DBD results in the loss of EWS/FLI-driven oncogenic transformation, this structural feature contributes an absolutely essential function to the fusion protein. Future studies to investigate interaction partners at the fourth alpha-helix of the DBD of FLI include both in vitro and in vivo methods: In vitro, recombinant FLI DBD and FLI DBD+ proteins can be purified and incubated with a nuclear lysate isolated from Ewing sarcoma cells. Immunoprecipitation followed by Western blot or mass spectrometry may reveal differences in protein interactions between the two FLI proteins, identifying the missing crucial protein for EWS/FLI function. Additionally, recombinant protein could be used to determine if this alpha-helix is responsible for altering DNA confirmation through crystal structure or nuclear magnetic resonance studies, as is reported for ETS1. In vivo, mutant EWS/FLI protein, either

containing or lacking the fourth alpha-helix of FLI, could be studied using fluorescence resonance electron transfer (FRET) or proximity-ligation assays to confirm a protein interaction occurs at this region if a protein of interest is identified. If a protein of interest is suspected to bind near EWS/FLI to interact with this alpha-helix, DNA-binding studies may also be used to confirm that the presence of the EWS/FLI mutant protein alters binding of this interaction partner. Determination of the crucial role this alpha-helix acts in for EWS/FLI function is crucial, as identification of these potential interaction partners may lead to a targetable interaction through development of a small molecule inhibitor.

4.3 Closing Remarks

Ewing sarcoma is an aggressive pediatric bone cancer defined by the presence of a chromosomal translocation fusing a FET family member to a member of the ETS transcription factor family. Although Ewing sarcoma patients have a characteristically low mutational burden, a look into disease biology mediated by these fusion proteins reveals a highly complex, multi-faceted set of properties attributed to those oncogenic transcription factors that we have only just begun to understand. The aggressiveness of current Ewing sarcoma treatment regimens, coupled with poor outcomes, highlights the importance of understanding the basic biology of the disease in hopes of ultimately creating better treatments for patients. The novel findings reported herein contribute to our basic

understanding of Ewing sarcomagenesis and through collective efforts, those in the Ewing sarcoma research field will ensure better outcomes become a reality for all.

4.4 References

1. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature. 1992;359(6391):162-5.

2. Grunewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H, et al. Ewing sarcoma. Nat Rev Dis Primers. 2018;4(1):5.

3. Riggi N, Suva ML, Stamenkovic I. Ewing's Sarcoma. N Engl J Med. 2021;384(2):154-64.

4. Sankar S, Lessnick SL. Promiscuous partnerships in Ewing's sarcoma. Cancer Genet. 2011;204(7):351-65.

5. Bridge RS, Rajaram V, Dehner LP, Pfeifer JD, Perry A. Molecular diagnosis of Ewing sarcoma/primitive neuroectodermal tumor in routinely processed tissue: a comparison of two FISH strategies and RT-PCR in malignant round cell tumors. Mod Pathol. 2006;19(1):1-8.

6. Aurias A, Rimbaut C, Buffe D, Zucker JM, Mazabraud A. Translocation involving chromosome 22 in Ewing's sarcoma. A cytogenetic study of four fresh tumors. Cancer Genet Cytogenet. 1984;12(1):21-5.

7. Turc-Carel C, Philip I, Berger MP, Philip T, Lenoir GM. Chromosome study of Ewing's sarcoma (ES) cell lines. Consistency of a reciprocal translocation t(11;22)(q24;q12). Cancer Genet Cytogenet. 1984;12(1):1-19.

8. Turc-Carel C, Aurias A, Mugneret F, Lizard S, Sidaner I, Volk C, et al. Chromosomes in Ewing's sarcoma. I. An evaluation of 85 cases of remarkable consistency of t(11;22)(q24;q12). Cancer Genet Cytogenet. 1988;32(2):229-38.

9. Tu J, Huo Z, Gingold J, Zhao R, Shen J, Lee DF. The Histogenesis of Ewing Sarcoma. Cancer Rep Rev. 2017;1(2).

10. Kinnaman MD, Zhu C, Weiser DA, Mohiuddin S, Hingorani P, Roth M, et al. Survey of Paediatric Oncologists and Pathologists regarding Their Views and Experiences with Variant Translocations in Ewing and Ewing-Like Sarcoma: A Report of the Children's Oncology Group. Sarcoma. 2020;2020:3498549.

11. Ibarra IL, Hollmann NM, Klaus B, Augsten S, Velten B, Hennig J, et al. Mechanistic insights into transcription factor cooperativity and its impact on protein-phenotype interactions. Nat Commun. 2020;11(1):124.

12. Li R, Pei H, Watson DK. Regulation of Ets function by protein - protein interactions. Oncogene. 2000;19(55):6514-23.

13. Nagulapalli S, Pongubala JM, Atchison ML. Multiple proteins physically interact with PU.1. Transcriptional synergy with NF-IL6 beta (C/EBP delta, CRP3). J Immunol. 1995;155(9):4330-8.

14. Garvie CW, Hagman J, Wolberger C. Structural studies of Ets-1/Pax5 complex formation on DNA. Mol Cell. 2001;8(6):1267-76.

15. McNagny KM, Sieweke MH, Doderlein G, Graf T, Nerlov C. Regulation of eosinophil-specific gene expression by a C/EBP-Ets complex and GATA-1. EMBO J. 1998;17(13):3669-80.

16. Aryee DN, Sommergruber W, Muehlbacher K, Dockhorn-Dworniczak B, Zoubek A, Kovar H. Variability in gene expression patterns of Ewing tumor cell lines differing in EWS-FLI1 fusion type. Lab Invest. 2000;80(12):1833-44.

17. Ewing J. Classics in oncology. Diffuse endothelioma of bone. James Ewing. Proceedings of the New York Pathological Society, 1921. CA Cancer J Clin. 1972;22(2):95-8.

18. Cripe TP. Ewing sarcoma: an eponym window to history. Sarcoma. 2011;2011:457532.

19. Beck R, Monument MJ, Watkins WS, Smith R, Boucher KM, Schiffman JD, et al. EWS/FLI-responsive GGAA microsatellites exhibit polymorphic differences between European and African populations. Cancer Genet. 2012;205(6):304-12.

4.5 Figures

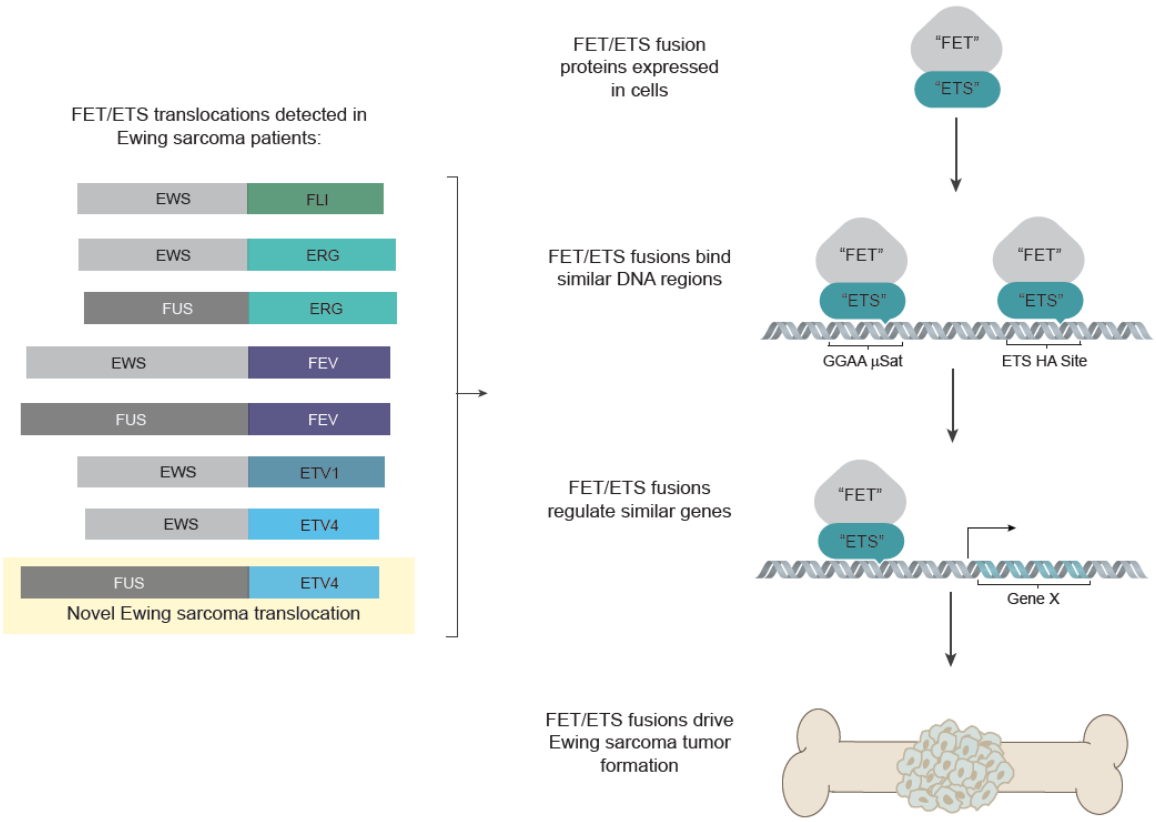

Figure 4.1. FET/ETS fusion proteins share similar DNA-binding and transcriptional regulatory properties. A comparative analysis reveals that FET/ETS translocations exhibit similar DNA-binding and transcriptional regulation profiles in a Ewing sarcoma cell line. As these fusions share highly similar functions, this can likely be extended to novel FET/ETS translocations, including the FUS/ETV4 translocation reported here.

Based on these data, FET/ETS translocations should be considered to be *bona fide* Ewing sarcoma translocations.

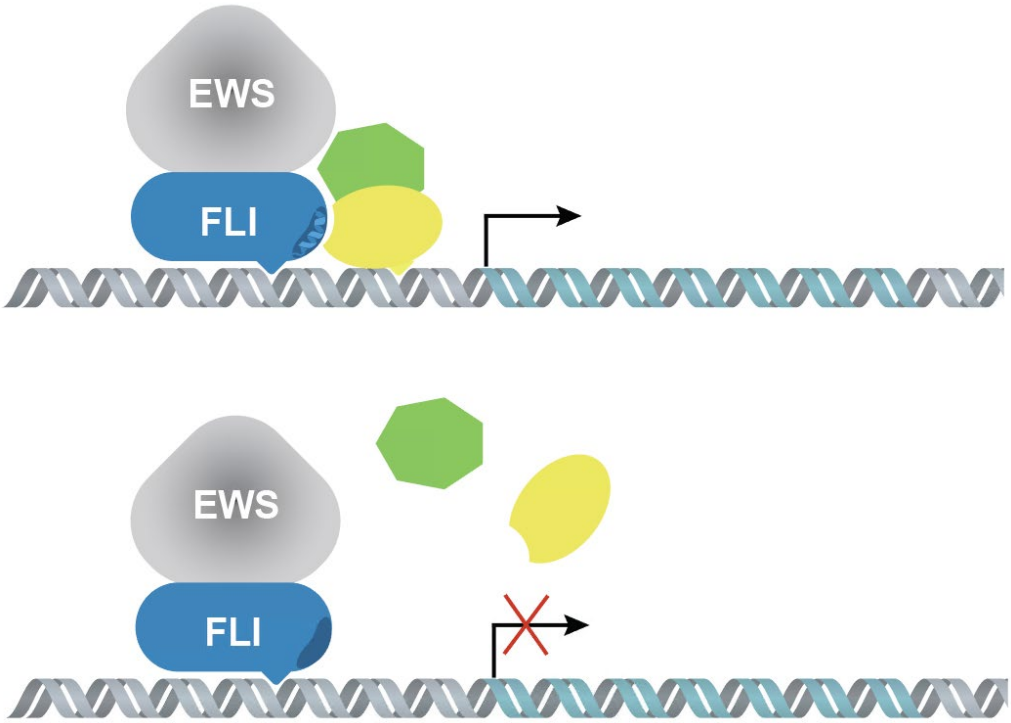

Figure 4.2. Alpha-helix flanking the FLI DNA-binding domain is crucial for transcriptional regulation by the EWS/FLI fusion protein required for oncogenic transformation.

Analysis of FLI domain contributions to EWS/FLI activity reveals that a fourth alpha-helix flanking the DNA-binding domain is required for transcriptional regulation properties of the fusion protein. Data reported herein support that this alpha-helix likely mediates interaction with other transcriptional regulators or epigenetic modifiers that are required for the EWS/FLI-driven transcription program and oncogenic transformation. Comparison of FLI and other ETS factors reveal that this alpha-helix is conserved across many members of the family, further indicating that this structural feature is essential for function of the ETS transcription factor family.

Compiled References List

Compiled References List

1. Cripe TP. Ewing sarcoma: an eponym window to history. Sarcoma. 2011;2011:457532.

2. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature. 1992;359(6391):162-5.

3. Grunewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H, et al. Ewing sarcoma. Nat Rev Dis Primers. 2018;4(1):5.

4. Aurias A, Rimbaut C, Buffe D, Zucker JM, Mazabraud A. Translocation involving chromosome 22 in Ewing's sarcoma. A cytogenetic study of four fresh tumors. Cancer Genet Cytogenet. 1984;12(1):21-5.

5. Turc-Carel C, Aurias A, Mugneret F, Lizard S, Sidaner I, Volk C, et al. Chromosomes in Ewing's sarcoma. I. An evaluation of 85 cases of remarkable consistency of t(11;22)(q24;q12). Cancer Genet Cytogenet. 1988;32(2):229-38.

6. Aurias A, Rimbaut C, Buffe D, Dubousset J, Mazabraud A. [Translocation of chromosome 22 in Ewing's sarcoma]. C R Seances Acad Sci III. 1983;296(23):1105-7.

7. Turc-Carel C, Philip I, Berger MP, Philip T, Lenoir GM. Chromosome study of Ewing's sarcoma (ES) cell lines. Consistency of a reciprocal translocation t(11;22)(q24;q12). Cancer Genet Cytogenet. 1984;12(1):1-19.

8. May WA, Gishizky ML, Lessnick SL, Lunsford LB, Lewis BC, Delattre O, et al. Ewing sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation. Proc Natl Acad Sci U S A. 1993;90(12):5752-6.

9. Zucman J, Melot T, Desmaze C, Ghysdael J, Plougastel B, Peter M, et al. Combinatorial generation of variable fusion proteins in the Ewing family of tumours. EMBO J. 1993;12(12):4481-7.

10. Pishas KI, Lessnick SL. Recent advances in targeted therapy for Ewing sarcoma. F1000Res. 2016;5.

11. Jain S, Kapoor G. Chemotherapy in Ewing's sarcoma. Indian J Orthop. 2010;44(4):369-77.

12. Longhi A, Ferrari S, Tamburini A, Luksch R, Fagioli F, Bacci G, et al. Late effects of chemotherapy and radiotherapy in osteosarcoma and Ewing sarcoma patients: the Italian Sarcoma Group Experience (1983-2006). Cancer. 2012;118(20):5050-9.

13. Lawlor ER, Sorensen PH. Twenty Years on: What Do We Really Know about Ewing Sarcoma and What Is the Path Forward? Crit Rev Oncog. 2015;20(3-4):155-71.

14. Dyson KA, Stover BD, Grippin A, Mendez-Gomez HR, Lagmay J, Mitchell DA, et al. Emerging trends in immunotherapy for pediatric sarcomas. J Hematol Oncol. 2019;12(1):78.

15. Riggi N, Suva ML, Stamenkovic I. Ewing's Sarcoma. N Engl J Med. 2021;384(2):154-64.

16. Ewing J. Classics in oncology. Diffuse endothelioma of bone. James Ewing. Proceedings of the New York Pathological Society, 1921. CA Cancer J Clin. 1972;22(2):95-8.

17. Ewing J. The Classic: Diffuse endothelioma of bone. Proceedings of the New York Pathological Society. 1921;12:17. Clin Orthop Relat Res. 2006;450:25-7.

18. Bailly RA, Bosselut R, Zucman J, Cormier F, Delattre O, Roussel M, et al. DNA-binding and transcriptional activation properties of the EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing sarcoma. Mol Cell Biol. 1994;14(5):3230-41.

19. Liang H, Olejniczak ET, Mao X, Nettesheim DG, Yu L, Thompson CB, et al. The secondary structure of the ets domain of human Fli-1 resembles that of the helix-turn-helix DNA-binding motif of the Escherichia coli catabolite gene activator protein. Proc Natl Acad Sci U S A. 1994;91(24):11655-9.

20. Laudet V, Hanni C, Stehelin D, Duterque-Coquillaud M. Molecular phylogeny of the ETS gene family. Oncogene. 1999;18(6):1351-9.

21. Kovar H. Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. Sarcoma. 2011;2011:837474.

22. Li Y, Luo H, Liu T, Zacksenhaus E, Ben-David Y. The ets transcription factor Fli-1 in development, cancer and disease. Oncogene. 2015;34(16):2022-31.

23. Sizemore GM, Pitarresi JR, Balakrishnan S, Ostrowski MC. The ETS family of oncogenic transcription factors in solid tumours. Nat Rev Cancer. 2017;17(6):337-51.

24. Bertolotti A, Lutz Y, Heard DJ, Chambon P, Tora L. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. EMBO J. 1996;15(18):5022-31.

25. Hackl W, Luhrmann R. Molecular cloning and subcellular localisation of the snRNP-associated protein 69KD, a structural homologue of the proto-oncoproteins TLS and EWS with RNA and DNA-binding properties. J Mol Biol. 1996;264(5):843-51.

26. Sankar S, Lessnick SL. Promiscuous partnerships in Ewing's sarcoma. Cancer Genet. 2011;204(7):351-65.

27. Ng TL, O'Sullivan MJ, Pallen CJ, Hayes M, Clarkson PW, Winstanley M, et al. Ewing sarcoma with novel translocation t(2;16) producing an in-frame fusion of FUS and FEV. J Mol Diagn. 2007;9(4):459-63.

28. Kinnaman MD, Zhu C, Weiser DA, Mohiuddin S, Hingorani P, Roth M, et al. Survey of Paediatric Oncologists and Pathologists regarding Their Views and Experiences with Variant Translocations in Ewing and Ewing-Like Sarcoma: A Report of the Children's Oncology Group. Sarcoma. 2020;2020:3498549.

29. Machado I, Noguera R, Pellin A, Lopez-Guerrero JA, Piqueras M, Navarro S, et al. Molecular diagnosis of Ewing sarcoma family of tumors: a comparative analysis of 560 cases with FISH and RT-PCR. Diagn Mol Pathol. 2009;18(4):189-99.

30. Anderson ND, de Borja R, Young MD, Fuligni F, Rosic A, Roberts ND, et al. Rearrangement bursts generate canonical gene fusions in bone and soft tissue tumors. Science. 2018;361(6405).

31. Ly P, Brunner SF, Shoshani O, Kim DH, Lan W, Pyntikova T, et al. Chromosome segregation errors generate a diverse spectrum of simple and complex genomic rearrangements. Nat Genet. 2019;51(4):705-15.

32. Shen MM. Chromoplexy: a new category of complex rearrangements in the cancer genome. Cancer Cell. 2013;23(5):567-9.

33. Ballas LK, Hu BR, Quinn DI. Chromoplexy and hypoxic microenvironment drives prostate cancer. Lancet Oncol. 2014;15(13):1419-21.

34. Gasi Tandefelt D, Boormans J, Hermans K, Trapman J. ETS fusion genes in prostate cancer. Endocr Relat Cancer. 2014;21(3):R143-52.

35. Lin PP, Wang Y, Lozano G. Mesenchymal Stem Cells and the Origin of Ewing's Sarcoma. Sarcoma. 2011;2011.

36. Tu J, Huo Z, Gingold J, Zhao R, Shen J, Lee DF. The Histogenesis of Ewing Sarcoma. Cancer Rep Rev. 2017;1(2).

37. Suh CH, Ordonez NG, Hicks J, Mackay B. Ultrastructure of the Ewing's sarcoma family of tumors. Ultrastruct Pathol. 2002;26(2):67-76.

38. Boulay G, Sandoval GJ, Riggi N, Iyer S, Buisson R, Naigles B, et al. Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. Cell. 2017;171(1):163-78 e19.

39. Rocchi A, Manara MC, Sciandra M, Zambelli D, Nardi F, Nicoletti G, et al. CD99 inhibits neural differentiation of human Ewing sarcoma cells and thereby contributes to oncogenesis. J Clin Invest. 2010;120(3):668-80.

40. Louati S, Senhaji N, Chbani L, Bennis S. EWSR1 Rearrangement and CD99 Expression as Diagnostic Biomarkers for Ewing/PNET Sarcomas in a Moroccan Population. Disease Markers. 2018;2018:7971019.

41. Baldauf MC, Orth MF, Dallmayer M, Marchetto A, Gerke JS, Rubio RA, et al. Robust diagnosis of Ewing sarcoma by immunohistochemical detection of super-enhancer-driven EWSR1-ETS targets. Oncotarget. 2018;9(2):1587-601.

42. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.

43. Brohl AS, Solomon DA, Chang W, Wang J, Song Y, Sindiri S, et al. The genomic landscape of the Ewing Sarcoma family of tumors reveals recurrent STAG2 mutation. PLoS Genet. 2014;10(7):e1004475.

44. Pishas KI, Drenberg CD, Taslim C, Theisen ER, Johnson KM, Saund RS, et al. Therapeutic Targeting of KDM1A/LSD1 in Ewing Sarcoma with SP-2509 Engages the Endoplasmic Reticulum Stress Response. Mol Cancer Ther. 2018;17(9):1902-16.

45. Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, Calicchio ML, et al. The genomic landscape of pediatric Ewing sarcoma. Cancer Discov. 2014;4(11):1326-41.

46. Tirode F, Surdez D, Ma X, Parker M, Le Deley MC, Bahrami A, et al. Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. Cancer Discov. 2014;4(11):1342-53.

47. Surdez D, Zaidi S, Grossetete S, Laud-Duval K, Ferre AS, Mous L, et al. STAG2 mutations alter CTCF-anchored loop extrusion, reduce cis-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma. Cancer Cell. 2021.

48. Barber-Rotenberg JS, Selvanathan SP, Kong Y, Erkizan HV, Snyder TM, Hong SP, et al. Single enantiomer of YK-4-279 demonstrates specificity in targeting the oncogene EWS-FLI1. Oncotarget. 2012;3(2):172-82.

49. Minas TZ, Han J, Javaheri T, Hong SH, Schlederer M, Saygideger-Kont Y, et al. YK-4-279 effectively antagonizes EWS-FLI1 induced leukemia in a transgenic mouse model. Oncotarget. 2015;6(35):37678-94.

50. Loganathan SN, Tang N, Fleming JT, Ma Y, Guo Y, Borinstein SC, et al. BET bromodomain inhibitors suppress EWS-FLI1-dependent transcription and the IGF1 autocrine mechanism in Ewing sarcoma. Oncotarget. 2016;7(28):43504-17.

51. Gollavilli PN, Pawar A, Wilder-Romans K, Natesan R, Engelke CG, Dommeti VL, et al. EWS/ETS-Driven Ewing Sarcoma Requires BET Bromodomain Proteins. Cancer Res. 2018;78(16):4760-73.

52. Sankar S, Theisen ER, Bearss J, Mulvihill T, Hoffman LM, Sorna V, et al. Reversible LSD1 inhibition interferes with global EWS/ETS transcriptional activity and impedes Ewing sarcoma tumor growth. Clin Cancer Res. 2014;20(17):4584-97.

53. Iko Y, Kodama TS, Kasai N, Oyama T, Morita EH, Muto T, et al. Domain architectures and characterization of an RNA-binding protein, TLS. J Biol Chem. 2004;279(43):44834-40.

54. Tan AY, Manley JL. TLS/FUS: a protein in cancer and ALS. Cell Cycle. 2012;11(18):3349-50.

55. Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovsky A, Poser I, et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science. 2018;360(6391):918-21.

56. Lin Y, Protter DS, Rosen MK, Parker R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol Cell. 2015;60(2):208-19.

57. Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science. 2018;361(6400).

58. Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell. 2018;175(7):1842-55 e16.

59. Lautenberger JA, Papas TS. Inversion of a chicken ets-1 proto-oncogene segment in avian leukemia virus E26. J Virol. 1993;67(1):610-2.

60. Szymczyna BR, Arrowsmith CH. DNA binding specificity studies of four ETS proteins support an indirect read-out mechanism of protein-DNA recognition. J Biol Chem. 2000;275(37):28363-70.

61. Adamo P, Ladomery MR. The oncogene ERG: a key factor in prostate cancer. Oncogene. 2016;35(4):403-14.

62. Mao X, Miesfeldt S, Yang H, Leiden JM, Thompson CB. The FLI-1 and chimeric EWS-FLI-1 oncoproteins display similar DNA binding specificities. J Biol Chem. 1994;269(27):18216-22.

63. Truong AH, Ben-David Y. The role of Fli-1 in normal cell function and malignant transformation. Oncogene. 2000;19(55):6482-9.

64. Ibarra IL, Hollmann NM, Klaus B, Augsten S, Velten B, Hennig J, et al. Mechanistic insights into transcription factor cooperativity and its impact on protein-phenotype interactions. Nat Commun. 2020;11(1):124.

65. Erkizan HV, Uversky VN, Toretsky JA. Oncogenic partnerships: EWS-FLI1 protein interactions initiate key pathways of Ewing's sarcoma. Clin Cancer Res. 2010;16(16):4077-83.

66. Hollenhorst PC, McIntosh LP, Graves BJ. Genomic and biochemical insights into the specificity of ETS transcription factors. Annu Rev Biochem. 2011;80:437-71.

67. Gu TL, Goetz TL, Graves BJ, Speck NA. Auto-inhibition and partner proteins, core-binding factor beta (CBFbeta) and Ets-1, modulate DNA binding by CBFalpha2 (AML1). Mol Cell Biol. 2000;20(1):91-103.

68. Goetz TL, Gu TL, Speck NA, Graves BJ. Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor alpha2. Mol Cell Biol. 2000;20(1):81-90.

69. De S, Okon M, Graves BJ, McIntosh LP. Autoinhibition of ETV6 DNA Binding Is Established by the Stability of Its Inhibitory Helix. J Mol Biol. 2016;428(8):1515-30.

70. Currie SL, Lau DKW, Doane JJ, Whitby FG, Okon M, McIntosh LP, et al. Structured and disordered regions cooperatively mediate DNA-binding autoinhibition of ETS factors ETV1, ETV4 and ETV5. Nucleic Acids Res. 2017;45(5):2223-41.

71. Currie SL, Warner SL, Vankayalapati H, Liu X, Sharma S, Bearss DJ, et al. Development of High-Throughput Screening Assays for Inhibitors of ETS Transcription Factors. SLAS Discov. 2019;24(1):77-85.

72. De S, Chan AC, Coyne HJ, 3rd, Bhachech N, Hermsdorf U, Okon M, et al. Steric mechanism of auto-inhibitory regulation of specific and non-specific DNA binding by the ETS transcriptional repressor ETV6. J Mol Biol. 2014;426(7):1390-406.

73. Regan MC, Horanyi PS, Pryor EE, Jr., Sarver JL, Cafiso DS, Bushweller JH. Structural and dynamic studies of the transcription factor ERG reveal DNA binding is allosterically autoinhibited. Proc Natl Acad Sci U S A. 2013;110(33):13374-9.

74. Sharma R, Gangwar SP, Saxena AK. Comparative structure analysis of the ETSi domain of ERG3 and its complex with the E74 promoter DNA sequence. Acta Crystallogr F Struct Biol Commun. 2018;74(Pt 10):656-63.

75. Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia. 2008;10(2):177-88.

76. Hajdu SI. The enigma of Ewing's sarcoma. Ann Clin Lab Sci. 2006;36(1):108-10.

77. Kovar H. Downstream EWS/FLI1 - upstream Ewing's sarcoma. Genome Med. 2010;2(1):8.

78. May WA, Lessnick SL, Braun BS, Klemsz M, Lewis BC, Lunsford LB, et al. The Ewing's sarcoma EWS/FLI-1 fusion gene encodes a more potent transcriptional activator and is a more powerful transforming gene than FLI-1. Mol Cell Biol. 1993;13(12):7393-8.

79. Lessnick SL, Braun BS, Denny CT, May WA. Multiple domains mediate transformation by the Ewing's sarcoma EWS/FLI-1 fusion gene. Oncogene. 1995;10(3):423-31.

80. Theisen ER, Miller KR, Showpnil IA, Taslim C, Pishas KI, Lessnick SL. Transcriptomic analysis functionally maps the intrinsically disordered domain of EWS/FLI and reveals novel transcriptional dependencies for oncogenesis. Genes Cancer. 2019;10(1-2):21-38.

81. Machiela MJ, Grunewald TGP, Surdez D, Reynaud S, Mirabeau O, Karlins E, et al. Genome-wide association study identifies multiple new loci associated with Ewing sarcoma susceptibility. Nat Commun. 2018;9(1):3184.

82. Gangwal K, Lessnick SL. Microsatellites are EWS/FLI response elements: genomic "junk" is EWS/FLI's treasure. Cell Cycle. 2008;7(20):3127-32.

83. Gangwal K, Sankar S, Hollenhorst PC, Kinsey M, Haroldsen SC, Shah AA, et al. Microsatellites as EWS/FLI response elements in Ewing's sarcoma. Proc Natl Acad Sci U S A. 2008;105(29):10149-54.

84. Gangwal K, Close D, Enriquez CA, Hill CP, Lessnick SL. Emergent Properties of EWS/FLI Regulation via GGAA Microsatellites in Ewing's Sarcoma. Genes Cancer. 2010;1(2):177-87.

85. Beck R, Monument MJ, Watkins WS, Smith R, Boucher KM, Schiffman JD, et al. EWS/FLI-responsive GGAA microsatellites exhibit polymorphic differences between European and African populations. Cancer Genet. 2012;205(6):304-12.

86. Johnson KM, Mahler NR, Saund RS, Theisen ER, Taslim C, Callender NW, et al. Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. Proc Natl Acad Sci U S A. 2017;114(37):9870-5.

87. Johnson KM, Taslim C, Saund RS, Lessnick SL. Identification of two types of GGAA-microsatellites and their roles in EWS/FLI binding and gene regulation in Ewing sarcoma. PLoS One. 2017;12(11):e0186275.

88. Sankar S, Bell R, Stephens B, Zhuo R, Sharma S, Bearss DJ, et al. Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. Oncogene. 2013;32(42):5089-100.

89. Theisen ER, Pishas KI, Saund RS, Lessnick SL. Therapeutic opportunities in Ewing sarcoma: EWS-FLI inhibition via LSD1 targeting. Oncotarget. 2016;7(14):17616-30.

90. Theisen ER, Selich-Anderson J, Miller KR, Tanner JM, Taslim C, Pishas KI, et al. Chromatin profiling reveals relocalization of lysine-specific demethylase 1 by an oncogenic fusion protein. Epigenetics. 2021;16(4):405-24.

91. Tomazou EM, Sheffield NC, Schmidl C, Schuster M, Schonegger A, Datlinger P, et al. Epigenome mapping reveals distinct modes of gene regulation and widespread enhancer reprogramming by the oncogenic fusion protein EWS-FLI1. Cell Rep. 2015;10(7):1082-95.

92. Mendiola M, Carrillo J, Garcia E, Lalli E, Hernandez T, de Alava E, et al. The orphan nuclear receptor DAX1 is up-regulated by the EWS/FLI1 oncoprotein and is highly expressed in Ewing tumors. Int J Cancer. 2006;118(6):1381-9.

93. Gorthi A, Bishop AJR. Ewing sarcoma fusion oncogene: At the crossroads of transcription and DNA damage response. Mol Cell Oncol. 2018;5(4):e1465014.

94. Gorthi A, Romero JC, Loranc E, Cao L, Lawrence LA, Goodale E, et al. EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. Nature. 2018;555(7696):387-91.

95. Sanchez-Molina S, Figuerola-Bou E, Blanco E, Sanchez-Jimenez M, Taboas P, Gomez S, et al. RING1B recruits EWSR1-FLI1 and cooperates in the remodeling of chromatin necessary for Ewing sarcoma tumorigenesis. Sci Adv. 2020;6(43).

96. Kinsey M, Smith R, Lessnick SL. NR0B1 is required for the oncogenic phenotype mediated by EWS/FLI in Ewing's sarcoma. Mol Cancer Res. 2006;4(11):851-9.

97. Spahn L, Siligan C, Bachmaier R, Schmid JA, Aryee DN, Kovar H. Homotypic and heterotypic interactions of EWS, FLI1 and their oncogenic fusion protein. Oncogene. 2003;22(44):6819-29.

98. Becker JS, McCarthy RL, Sidoli S, Donahue G, Kaeding KE, He Z, et al. Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. Mol Cell. 2017;68(6):1023-37 e15.

99. Szabo Q, Bantignies F, Cavalli G. Principles of genome folding into topologically associating domains. Sci Adv. 2019;5(4):eaaw1668.

100. Zaret KS. Pioneer Transcription Factors Initiating Gene Network Changes. Annu Rev Genet. 2020;54:367-85.

101. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol Cell. 2002;9(2):279-89.

102. Mayran A, Sochodolsky K, Khetchoumian K, Harris J, Gauthier Y, Bemmo A, et al. Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening. Nat Commun. 2019;10(1):3807.

103. Fernandez Garcia M, Moore CD, Schulz KN, Alberto O, Donague G, Harrison MM, et al. Structural Features of Transcription Factors Associating with Nucleosome Binding. Mol Cell. 2019;75(5):921-32 e6.

104. Zhu F, Farnung L, Kaasinen E, Sahu B, Yin Y, Wei B, et al. The interaction landscape between transcription factors and the nucleosome. Nature. 2018;562(7725):76-81.

105. Riggi N, Knoechel B, Gillespie SM, Rheinbay E, Boulay G, Suva ML, et al. EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. Cancer Cell. 2014;26(5):668-81.

106. Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, et al. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell. 2017;171(7):1573-88 e28.

107. Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, et al. Two independent modes of chromatin organization revealed by cohesin removal. Nature. 2017;551(7678):51-6.

108. Zhao Z, Shilatifard A. Epigenetic modifications of histones in cancer. Genome Biol. 2019;20(1):245.

109. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019;10(1):1930.

110. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.

111. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.

112. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

113. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013;9(8):e1003118.

114. Braun BS, Frieden R, Lessnick SL, May WA, Denny CT. Identification of target genes for the Ewing's sarcoma EWS/FLI fusion protein by representational difference analysis. Mol Cell Biol. 1995;15(8):4623-30.

115. Braunreiter CL, Hancock JD, Coffin CM, Boucher KM, Lessnick SL. Expression of EWS-ETS fusions in NIH3T3 cells reveals significant differences to Ewing's sarcoma. Cell Cycle. 2006;5(23):2753-9.

116. Welford SM, Hebert SP, Deneen B, Arvand A, Denny CT. DNA binding domain-independent pathways are involved in EWS/FLI1-mediated oncogenesis. J Biol Chem. 2001;276(45):41977-84.

117. Arvand A, Welford SM, Teitell MA, Denny CT. The COOH-terminal domain of FLI-1 is necessary for full tumorigenesis and transcriptional modulation by EWS/FLI-1. Cancer Res. 2001;61(13):5311-7.

118. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44(W1):W160-5.

119. Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics. 2010;11:237.

120. Hodges HC, Stanton BZ, Cermakova K, Chang CY, Miller EL, Kirkland JG, et al. Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. Nat Struct Mol Biol. 2018;25(1):61-72.

121. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10(12):1213-8.

122. Gu Z, Eils R, Schlesner M, Ishaque N. EnrichedHeatmap: an R/Bioconductor package for comprehensive visualization of genomic signal associations. BMC Genomics. 2018;19(1):234.

123. Maag JLV. gganatogram: An R package for modular visualisation of anatograms and tissues based on ggplot2. F1000Res. 2018;7:1576.

124. Hou C, Tsodikov OV. Structural Basis for Dimerization and DNA Binding of Transcription Factor FLI1. Biochemistry. 2015;54(50):7365-74.

125. Martinez-Ramirez A, Rodriguez-Perales S, Melendez B, Martinez-Delgado B, Urioste M, Cigudosa JC, et al. Characterization of the A673 cell line (Ewing tumor) by molecular cytogenetic techniques. Cancer Genet Cytogenet. 2003;141(2):138-42.

126. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife. 2017;6.

127. Watson DK, Robinson L, Hodge DR, Kola I, Papas TS, Seth A. FLI1 and EWS-FLI1 function as ternary complex factors and ELK1 and SAP1a function as ternary and quaternary complex factors on the Egr1 promoter serum response elements. Oncogene. 1997;14(2):213-21.

128. Kim S, Denny CT, Wisdom R. Cooperative DNA binding with AP-1 proteins is required for transformation by EWS-Ets fusion proteins. Mol Cell Biol. 2006;26(7):2467-78.

129. Li R, Pei H, Watson DK. Regulation of Ets function by protein - protein interactions. Oncogene. 2000;19(55):6514-23.

130. Nagulapalli S, Pongubala JM, Atchison ML. Multiple proteins physically interact with PU.1. Transcriptional synergy with NF-IL6 beta (C/EBP delta, CRP3). J Immunol. 1995;155(9):4330-8.

131. Garvie CW, Hagman J, Wolberger C. Structural studies of Ets-1/Pax5 complex formation on DNA. Mol Cell. 2001;8(6):1267-76.

132. McNagny KM, Sieweke MH, Doderlein G, Graf T, Nerlov C. Regulation of eosinophil-specific gene expression by a C/EBP-Ets complex and GATA-1. EMBO J. 1998;17(13):3669-80.

133. Aryee DN, Sommergruber W, Muehlbacher K, Dockhorn-Dworniczak B, Zoubek A, Kovar H. Variability in gene expression patterns of Ewing tumor cell lines differing in EWS-FLI1 fusion type. Lab Invest. 2000;80(12):1833-44.