

EWSR1 fusion proteins mediate PAX7 expression in Ewing sarcoma

Gregory W Charville¹, Wei-Lien Wang², Davis R Ingram², Angshumoy Roy³, Dafydd Thomas⁴, Rajiv M Patel⁴, Jason L Hornick⁵, Matt van de Rijn¹ and Alexander J Lazar²

¹Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA; ²Departments of Pathology & Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; ³Departments of Pathology & Immunology and Pediatrics, Texas Children's Hospital, Baylor College of Medicine, Houston, TX, USA; ⁴Department of Pathology, University of Michigan School of Medicine, Ann Arbor, MI, USA and ⁵Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA

PAX7 is a paired-box transcription factor that is required for the developmental specification of adult skeletal muscle progenitors in mice. We previously demonstrated PAX7 expression as a marker of skeletal muscle differentiation in rhabdomyosarcoma. Here, using analyses of published whole-genome gene expression microarray data, we identify PAX7 as a gene with significantly increased expression in Ewing sarcoma in comparison to CIC-DUX4 round cell sarcoma. Analysis of PAX7 in a large cohort of 103 Ewing sarcoma cases by immunohistochemistry revealed expression in 99.0% of cases (102/103). PAX7 expression was noted in cases demonstrating three distinct Ewing sarcoma EWSR1 translocations involving FLI1, ERG, and NFATc2. No PAX7 expression was observed in any of 27 cases of CIC-DUX4 sarcoma by immunohistochemistry (0%; 0/27). Exploring the mechanism of PAX7 expression in Ewing sarcoma using curated RNA- and ChIP-sequencing data, we demonstrate that the EWSR1 fusion protein is required for PAX7 expression in Ewing sarcoma and identify a candidate EWSR1-FLI1-bound PAX7 enhancer that coincides with both a consensus GGAA repeat-containing binding site and a peak of regulatory H3K27 acetylation. Taken together, our findings provide mechanistic support for the utility of PAX7 immunohistochemistry in the diagnosis of Ewing sarcoma, while linking this sarcoma of uncertain histogenesis to a key transcriptional regulator of mammalian muscle progenitor cells.

Modern Pathology (2017) 30, 1312–1320; doi:10.1038/modpathol.2017.49; published online 23 June 2017

Ewing sarcoma is a malignant small, round, blue-cell neoplasm, which most often presents in children and young adults. Ewing sarcoma is often described as morphologically ambiguous, consisting of cells with finely dispersed chromatin, inconspicuous nucleoli, and scant amounts of clear to amphophilic cytoplasm, giving the appearance of a highly cellular, 'blue' tumor on hematoxylin/eosin-stained histology. Owing to this non-specific cytomorphology, the diagnosis of Ewing sarcoma was historically one of exclusion, requiring a thorough evaluation for morphologic and molecular markers of other small, round, blue-cell tumors.¹ The classification of Ewing

sarcoma was revolutionized with the identification of a unique t(11;22) translocation, which is found in approximately 85–90% of Ewing sarcoma cases, including those arising in both bone and soft tissue.^{2–4} This translocation produces an oncogenic *EWSR1-FLI1* fusion, combining the 5' transcriptional activation domain of *EWSR1* with the 3' ETS DNA-binding domain of *FLI1* to yield a potent oncogene.⁵ Several variant translocations have been implicated in the minority of cases not exhibiting the typical *EWSR1-FLI1* fusion. These variants all pair the 5' portion of *EWSR1* with the 3' DNA-binding domain of a transcription factor (Figure 1). Of these rarer variants, the majority involve other ETS family transcription factors, such as *ERG* (the most common minor variant),⁶ *ETV1*,⁷ *E1AF* (*ETV4*),⁸ or *FEV*.⁹ More recently, even rarer fusion variants involving transcription factors without an ETS domain, such as *NFATC2*, have been discovered.¹⁰

Correspondence: Dr GW Charville, MD, PhD, Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Lane 235, Stanford, CA 94305-5324, USA.
E-mail: gwc@stanford.edu

Received 4 March 2017; revised 4 May 2017; accepted 21 May 2017; published online 23 June 2017

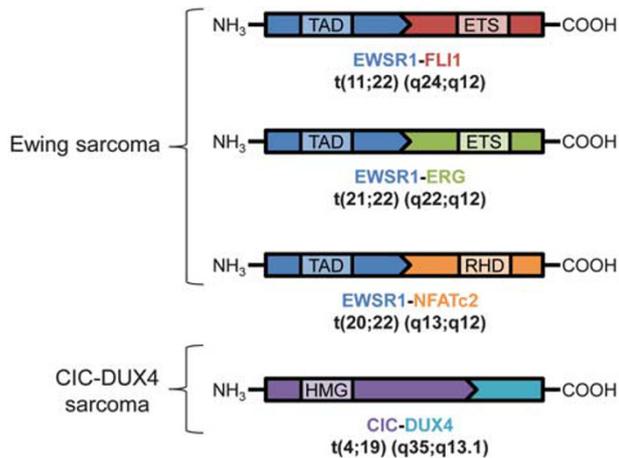


Figure 1 Schematic of translocations involved in Ewing sarcoma and CIC-DUX4 sarcoma. Representative illustrations of the EWSR1-containing Ewing sarcoma translocations in this study highlight the N-terminal transcriptional activation domain (TAD) of EWSR1. While translocations with FLI1 and ERG bear C-terminal E26 transformation-specific (ETS) DNA-binding domains, the translocation with NFATc2 carries a DNA-binding Rel homology domain (RHD) at its C terminus. The CIC-DUX4 translocation includes a high mobility group DNA-binding domain at the N terminus of CIC. Known functional domains of DUX4, including two DNA-binding homeodomains, are not present in the translocation.

Despite the identification of a pathognomonic molecular lesion in Ewing sarcoma, immunohistochemical analysis of protein expression patterns remains an important adjunctive diagnostic tool, especially in resource-limited settings, cases with scant biopsy material, or contexts in which there is clinical urgency and the relative availability of immunohistochemistry enables more rapid diagnosis.¹¹ Even before the identification of prototypical *EWSR1* translocations, it was noted that Ewing sarcoma expresses CD99 (MIC2). Although it is highly sensitive for the identification of Ewing sarcoma, CD99 is also expressed in other neoplasms, including synovial sarcoma,^{12,13} mesenchymal chondrosarcoma,¹⁴ malignant peripheral nerve sheath tumor,¹³ plasma cell myeloma,¹⁵ neuroendocrine tumors,¹⁶ and lymphoblastic lymphoma,^{17,18} among several others. In addition, the degree of CD99 membranous accentuation can vary from case-to-case, perhaps due to pre-analytic variables such as time to fixation or time of fixation. Immunohistochemistry using an antibody recognizing the ETS domain of FLI1 has also proven to be somewhat useful in the identification of Ewing sarcoma, particularly when combined with CD99 analyses.^{19–21} Much like CD99, however, FLI1 expression is not specific to Ewing sarcoma, showing expression in normal endothelium and lymphocytes,¹⁹ along with the vast majority of lymphoblastic lymphomas and vascular neoplasms.^{21–23} FLI1 is also expressed in smaller fractions of several other neoplasms, such as squamous cell carcinoma, urothelial carcinoma, Merkel cell carcinoma, lung adenocarcinoma, and melanoma.^{21,23} Akin to FLI1 immunohistochemical analysis, immunohistochemistry using an

antibody recognizing ERG is able to distinguish the subset of Ewing sarcoma with an underlying EWSR1-ERG fusion.²⁴

Recent advances in the genetic classification of undifferentiated sarcomas has led to the identification of a molecularly distinct Ewing-like tumor bearing a recurrent translocation linking the *CIC* gene with the *DUX4* or paralogous *DUX4L* gene (Figure 1).^{25–27} In addition to having an indistinct round cell morphology, the CIC-DUX4 sarcomas share many of the immunohistochemical features of Ewing sarcoma, including expression of CD99 and FLI1.²⁸ Although these classical immunohistochemical markers do not distinguish closely related round cell sarcomas, recent studies have identified several novel markers. NKX2-2 is a transcription factor expressed in over 90% of Ewing sarcomas and < 5% of CIC-DUX4 sarcomas; however, NKX2-2 expression is not entirely specific for Ewing sarcoma, with significant expression in mesenchymal chondrosarcoma and olfactory neuroblastoma, among other tumors.^{29,30} Conversely, the transcription factors WT1 and ETV4 are expressed in most CIC-DUX4 sarcomas, but not in Ewing sarcoma.^{31,32} Moreover, just as FLI1 or ERG expression is noted in ES cases bearing translocations involving these proteins, DUX4 expression has recently been shown to be a fusion-specific marker of CIC-DUX4 sarcoma.³³

In a previous study of PAX7 expression in an array of bone and soft tissue tumors, we observed that, in addition to its expression in rhabdomyosarcoma, PAX7 was also expressed in a small cohort of seven Ewing sarcoma samples.³⁴ Here, we aim to expand on this finding by exploring the diagnostic significance of PAX7 expression in a large series of Ewing sarcoma cases, along with several CIC-DUX4 sarcomas, and by studying the mechanism of PAX7 expression in Ewing sarcoma using curated high-throughput sequencing data.

Materials and methods

Cases of Ewing sarcoma, numbering 103 in all, were gathered from the surgical pathology files of The University of Texas MD Anderson Cancer Center, University of Michigan Health System, and Stanford University Medical Center. Seventy-three cases of Ewing sarcoma were evaluated as duplicate tissue cores represented as five micron-thick sections of a tissue microarray. An additional 17 Ewing sarcoma cases were evaluated as five micron-thick sections prepared from formalin-fixed paraffin-embedded whole-tissue sections. Thirteen Ewing sarcoma cases were evaluated as single tissue cores represented as five micron-thick sections of a separate tissue microarray. Twenty cases of CIC-DUX4 sarcoma were retrieved from the surgical pathology and consultation files of Brigham and Women's Hospital and were evaluated as four micron-thick sections

prepared from formalin-fixed paraffin-embedded whole-tissue sections. An additional seven CIC-DUX4 cases from University of Michigan Health System, Cleveland Clinic, and Cedars-Sinai Medical Center were evaluated as duplicate tissue cores represented as five micron-thick sections on a tissue microarray.²⁸ Classification of all CIC-DUX4 sarcomas was previously performed using fluorescent in situ hybridization (FISH) and/or reverse transcription-PCR (RT-PCR).^{28,31}

Immunohistochemical detection of PAX7 and ERG was performed as previously described.^{24,34} After antigen retrieval, histologic sections were stained with antibodies according to standard protocols for formalin-fixed paraffin-embedded material. A monoclonal antibody against PAX7 raised in mouse was purchased from the Developmental Studies Hybridoma Bank (Iowa City, IA, USA). Anti-PAX7 antibody from hybridoma supernatant was used at a dilution of 1:200 (by volume) following antigen retrieval using 0.01 M citrate buffer at pH 6.0. A monoclonal antibody against ERG (clone EP111) raised in rabbit was purchased from Epitomics (Burlingame, CA, USA) and used at a dilution of 1:2000 (by volume). Immunohistochemical analysis of NKX2-2 was performed essentially as previously described.²⁹ A monoclonal antibody against NKX2-2 raised in mouse was purchased from the Developmental Studies Hybridoma Bank (clone 74.5A5) and used at a dilution of 1:25 (by volume) following pressure cooker antigen retrieval using 0.01 M citrate buffer at pH 6.0. For PAX7, NKX2-2, and ERG, immunoreactivity was considered positive if greater than 20% of tumor cells showed staining with appropriate nuclear chromogen localization. Intensity of PAX7 staining was approximated semi-quantitatively using a scale akin to that commonly used for estimation of estrogen receptor expression in breast carcinoma: 1+ (weak but detectable above control), 2+ (distinct), and 3+ (strong).³⁵ All cases were studied according the ethical rules of each institution and processed in an anonymous coded fashion.

Publically curated microarray, chromatin immunoprecipitation-sequencing (ChIP-seq), and RNA-sequencing (RNA-seq) data were accessed via the Gene Expression Omnibus (GEO) between August 15, 2016, and November 20, 2016. Paired end RNA-seq reads from SKNMC and A673 cell lines with control GFP-targeting short hairpin-RNA (shRNA) or FLI1-targeting shRNA (GSE61950)³⁶ were aligned to the hg19 genome assembly and quantified as the number of sequenced fragments per kilobase of exon per million fragments mapped (FPKM) using the Tuxedo protocol.³⁷ All ChIP-seq data files (GSE61953) were aligned to the hg19 genome assembly and visualized using the Integrative Genomics Viewer. The y-axis scale was set at a constant value for all samples of a given ChIP-seq experiment. Microarray data sets (GSE60740) were analyzed using the Partek Genomics Suite (version

6.6) following exclusion of outlying samples by principal component analysis.

To characterize the status of the *EWSR1* locus, total RNA was extracted from 3 formalin-fixed paraffin-embedded (FFPE) PAX7-negative cases and quantified using the Qubit RNA assay (Thermo Fisher Scientific, Waltham, MA). Next generation sequencing (NGS) libraries were prepared using anchored multiplex PCR-based methodology as previously described,³⁸ using the Archer FusionPlex Sarcoma kit (ArcherDX Inc, Boulder, CO, USA). The Archer FusionPlex Sarcoma assay targets 127 exons in 26 genes for agnostic detection of gene fusions without *a priori* knowledge of the partner gene. NGS libraries were generated according to the manufacturer's protocol using 100 ng of total RNA. Amplifiable cDNA quality was assessed using the PreSeq RNA QC assay; 2 of 3 cases had amplifiable RNA based on the PreSeq assay. Final NGS libraries were quantified using the KAPA Library Quantification Kit for the Illumina platform. Illumina paired end indexed libraries were sequenced 8-plex on a MiSeq (2 × 150 bp, v2) and data analyzed on a vendor-provided virtual-machine based analysis pipeline with custom-developed output scripts. On average, sequencing with the Archer FusionPlex Sarcoma Panel generated a total of 1.07 million paired end reads with approximately 28 980 unique RNA reads per sample. Sequencing quality was assessed by % purity filtered (PF) reads, % bases >Q30, total number of reads, total number of unique RNA reads, and percent of aligned reads with high-mapping quality.

Results

To explore differential *PAX7* expression in Ewing sarcoma and CIC-DUX4 sarcoma, we first performed analyses of microarray-based genome-wide gene expression data from 12 CIC-DUX4 and 6 Ewing sarcoma samples, the latter bearing the variant *EWSR1-NFATC2* translocation (GEO data series GSE60740). Among 54675 analyzed genetic loci, we identified 7090 genes showing differential expression between Ewing sarcoma and CIC-DUX4 cases, using a fold-change of greater than two and a false discovery rate (FDR) of less than 0.01 to define significantly differential expression (Figure 2a). Ranking the differentially expressed genes according to p-value in a comparison of CIC-DUX4 sarcoma versus Ewing sarcoma (Figure 2b), we identified eight transcription factors among the 50 genes with the lowest p-values: *FOXP1*, *NR5A2*, *SOX5*, *VDR*, *WT1*, *EBF2*, *TFAP2*, and *PAX7*. Among these, *FOXP1*, *NR5A2*, *SOX5*, *VDR*, *TFAP2*, and *PAX7* showed increased expression in Ewing sarcoma, while *WT1* and *EBF2* showed increased expression in CIC-DUX4 sarcoma. *PAX7* exhibited a 73-fold increased expression in Ewing sarcoma ($P = 4.0 \times 10^{-15}$); *WT1* exhibited a 511-fold increased expression in CIC-DUX4 sarcoma

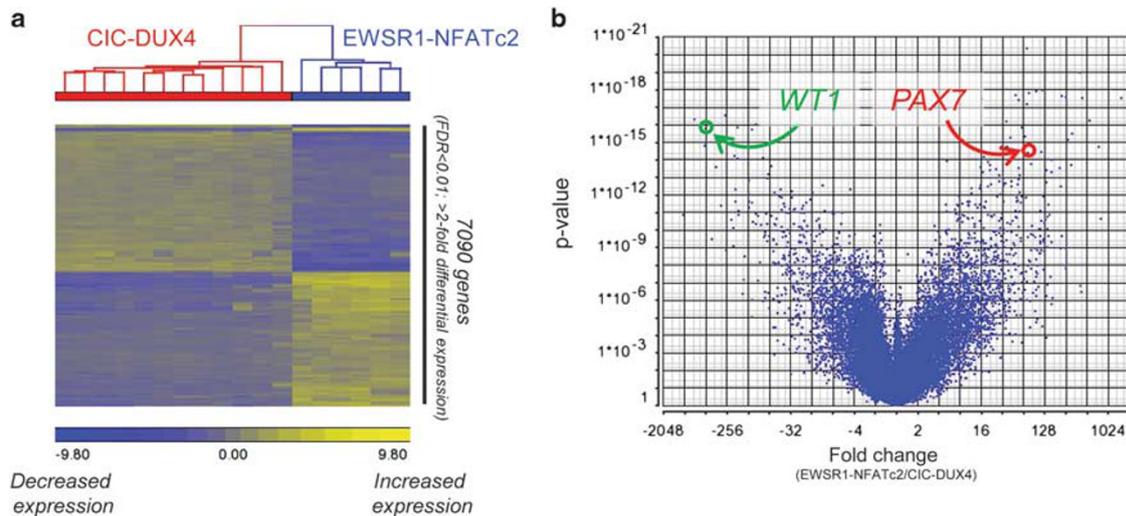


Figure 2 Whole-genome gene expression analyses identify PAX7 expression in Ewing sarcoma. (a) Heat-map of 7090 differentially expressed (>2 -fold increased or decreased, $FDR < 0.01$) genes in Ewing sarcoma versus CIC-DUX4 sarcoma. 12 CIC-DUX4 sarcomas and 6 EWSR1-NFATc2 Ewing sarcomas were clustered according to microarray-based gene expression signatures (GSE60740). (b) Volcano plot showing P -value versus fold-change in a comparison of high-throughput microarray gene expression data from 12 CIC-DUX4 sarcoma samples and 6 EWSR1-NFATc2 Ewing sarcoma samples. Each point on the plot represents one of 54675 individually analyzed loci. *WT1* and *PAX7* are highlighted as genes with significantly enriched expression in CIC-DUX4 sarcoma and Ewing sarcoma, respectively.

($P = 1.7 \times 10^{-16}$). Consistent with prior studies,^{29,30} *NKX2-2* expression was significantly increased in Ewing sarcoma (98-fold, $P = 9.7 \times 10^{-9}$), while *ETV4* expression was significantly increased in CIC-DUX4 sarcoma (12-fold, $P = 1.3 \times 10^{-8}$). Interestingly, *CD99* showed only two-fold increased expression in Ewing sarcoma, a difference that did not reach statistical significance ($P > 0.05$). Given that *PAX7* is required for the specification of adult skeletal muscle progenitors during mouse developmental myogenesis and serves as a marker of human adult skeletal muscle progenitors,^{39,40} we also examined the expression of standard markers of myogenic differentiation. *MYOD1*, *MYOG*, and *DES* each showed less than 1.25-fold differential expression in Ewing sarcoma versus CIC-DUX4 sarcoma ($P > 0.05$). Taken together, our analysis of whole-genome gene expression suggests that *PAX7* is a marker of Ewing sarcoma that distinguishes it from CIC-DUX4 sarcoma.

To further study the diagnostic utility of *PAX7* expression in Ewing sarcoma, we used immunohistochemistry to analyze *PAX7* protein in 103 Ewing sarcoma specimens. All specimens represented true biologic replicates from distinct patients and were previously diagnosed as Ewing sarcoma using a combination of morphologic, immunohistochemical, and molecular features. Of the 103 evaluated cases, 102 (99.0%) were positive for *PAX7* (Figure 3a). The lone *PAX7*-negative Ewing sarcoma case had no obvious distinguishing clinical or morphologic characteristic. Using a scale of 1–3+ to classify intensity of chromogen staining,³⁵ 98 out of 103 *PAX7*-positive cases (95.1%) demonstrated 2–3+ staining in greater than 50% of Ewing sarcoma cell nuclei, with the vast majority having anti-*PAX7* immunoreactivity in 90–100% of cells. In the remaining five

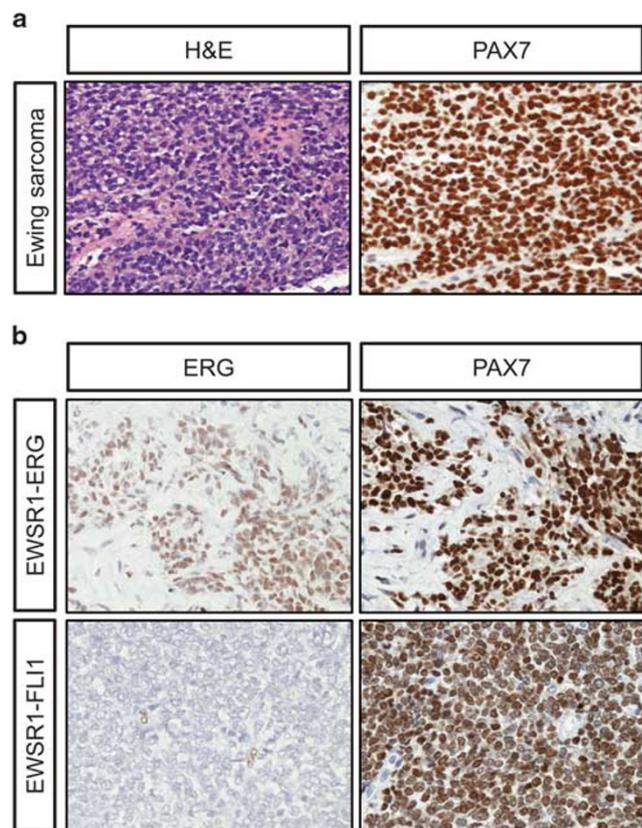


Figure 3 Strong and diffuse *PAX7* expression in Ewing sarcoma, including EWSR1-FLI1 and EWSR1-ERG variants. (a) Representative immunohistochemical detection of *PAX7* expression in Ewing sarcoma. (b) Representative immunohistochemical analysis of ERG and *PAX7* expression in Ewing sarcoma. Shown are examples of Ewing sarcoma with the EWSR1-ERG fusion (top), which shows diffuse anti-ERG immunoreactivity, and the EWSR1-FLI1 fusion (bottom), which shows no anti-ERG immunoreactivity.

PAX7-positive cases, there was, at minimum, 1–2+ staining in ~20–25% of cells. In all five cases for which a metastasis from the primary tumor was available for parallel immunohistochemical analysis, concordant PAX7 expression was identified. We, therefore, conclude that immunohistochemical detection of anti-PAX7 immunoreactivity is a sensitive marker of Ewing sarcoma.

In the initial cohort of Ewing sarcoma cases, a total of three PAX7-negative samples were identified. To further explore the molecular basis of PAX7 expression in Ewing sarcoma, and to confirm the diagnosis of Ewing sarcoma in these rare PAX7-negative cases, we performed multiplexed sequencing of PCR-enriched candidate pathogenic gene fusions, including those involving *EWSR1* (see Materials and Methods). This approach yielded informative sequencing libraries in two of three cases. Of the two sequenced cases, neither demonstrated an Ewing sarcoma-associated translocation of *EWSR1*, or any other likely pathogenic fusion. Given the lack of molecular support for the diagnosis of Ewing sarcoma, the two specimens evaluated by this sequencing approach were excluded from analyses of PAX7 expression and all subsequent analyses. Although an additional PAX7-negative specimen was not excluded given a failed sequencing analysis, it is important to note that there is no molecular confirmation of the diagnosis in this lone PAX7-negative case.

Noting that NKX2-2 is a transcription factor with robust expression in Ewing sarcoma, we were interested in studying its expression characteristics in parallel with those of PAX7. Consistent with previous reports,²⁹ NKX2-2 expression was positive (defined as expression in more than 20% of cells) in 84.9% (62/73) of cases. Four cases showed restricted expression in 10–20% of cells. Interestingly, the single PAX7-negative Ewing sarcoma case also lacked NKX2-2 expression. These findings, together with our earlier observations, indicate that there is an immunohistochemically distinct PAX7-positive, NKX2-2-negative subtype of Ewing sarcoma.

As the *EWSR1-FLI1* translocation accounts for approximately 85–90% of Ewing sarcoma, we next sought to characterize PAX7 expression in the rarer

EWSR1-ERG molecular subtype. We identified *EWSR1-ERG* variants among 73 Ewing sarcoma test cases using anti-ERG immunohistochemistry.²⁴ In all, we discovered five cases showing 2+ nuclear anti-ERG immunoreactivity in greater than 90% of cells (5/73, 6.8%), indicating an underlying *EWSR1-ERG* translocation (Figure 3b). The remaining 68 analyzed cases showed no significant ERG immunostaining within tumor cells, with vascular endothelial cells serving as a positive internal control (Figure 3b). We found that PAX7 was expressed by immunohistochemistry in all 5 ERG-positive Ewing sarcoma cases (5/5, 100%). NKX2-2 was also expressed in each of these cases. These findings indicate that PAX7 expression is a marker of both common and variant forms of Ewing sarcoma.

Our analyses of genome-wide gene expression data indicated not only that PAX7 expression is a marker of Ewing sarcoma, but also that it distinguishes Ewing sarcoma from CIC-DUX4 sarcoma. We, therefore, sought to evaluate PAX7 expression in a series of molecularly confirmed CIC-DUX4 sarcomas. Immunohistochemical analysis of PAX7 expression in 27 cases of CIC-DUX4 sarcoma revealed no significant expression in any case (0/27; Figure 4). These immunohistochemical studies confirm that PAX7 expression differentiates Ewing sarcoma from CIC-DUX4 sarcoma.

These observations, in conjunction with our previous studies,³⁴ suggest that significant PAX7 expression is a unique feature of rhabdomyosarcoma and Ewing sarcoma. In contrast to rhabdomyosarcoma, Ewing sarcoma does not show myogenic differentiation, as manifested by a lack of expression of myogenic genes, such as *DES*, in the large majority of cases.¹¹ What accounts for the observation of robust PAX7 expression in Ewing sarcoma? To address this question, we turned to a published study that used high-throughput RNA-sequencing (RNA-seq), along with chromatin immunoprecipitation-sequencing (ChIP-seq), to map *EWSR1-FLI1*-dependent chromatin remodeling and associated gene expression changes in Ewing sarcoma.³⁶ In this prior study, regulation of the *PAX7* gene was not specifically addressed and we, therefore, focused our analyses on *PAX7* using this large sequencing data set. RNA-seq studies of two Ewing sarcoma cell lines (SKNMC and A673) in this study showed expression of *PAX7* in both lines (Figure 5a). Notably, expression of *PAX7* in both Ewing sarcoma cell lines was dependent on expression of the *EWSR1-FLI1* fusion gene as evidenced by decreased *PAX7* transcript levels in the setting of shRNA-mediated *EWSR1-FLI1* knockdown (Figure 5a). The pattern of *PAX7* expression in Ewing sarcoma cells lines mimicked that of NKX2-2 (Figure 5a), which is known to be an *EWSR1-FLI1*-activated gene in Ewing sarcoma.³⁶ Despite robust *EWSR1-FLI1*-dependent *PAX7* expression in Ewing sarcoma cells, *MYOD1*, *MYOG*, and *DES* showed no significant expression by RNA-seq in Ewing sarcoma

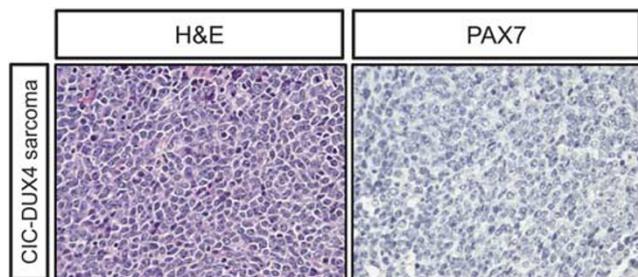


Figure 4 Absence of PAX7 expression in CIC-DUX4 sarcoma. Representative H&E and immunohistochemical analysis of PAX7 expression in a CIC-DUX4 sarcoma case.

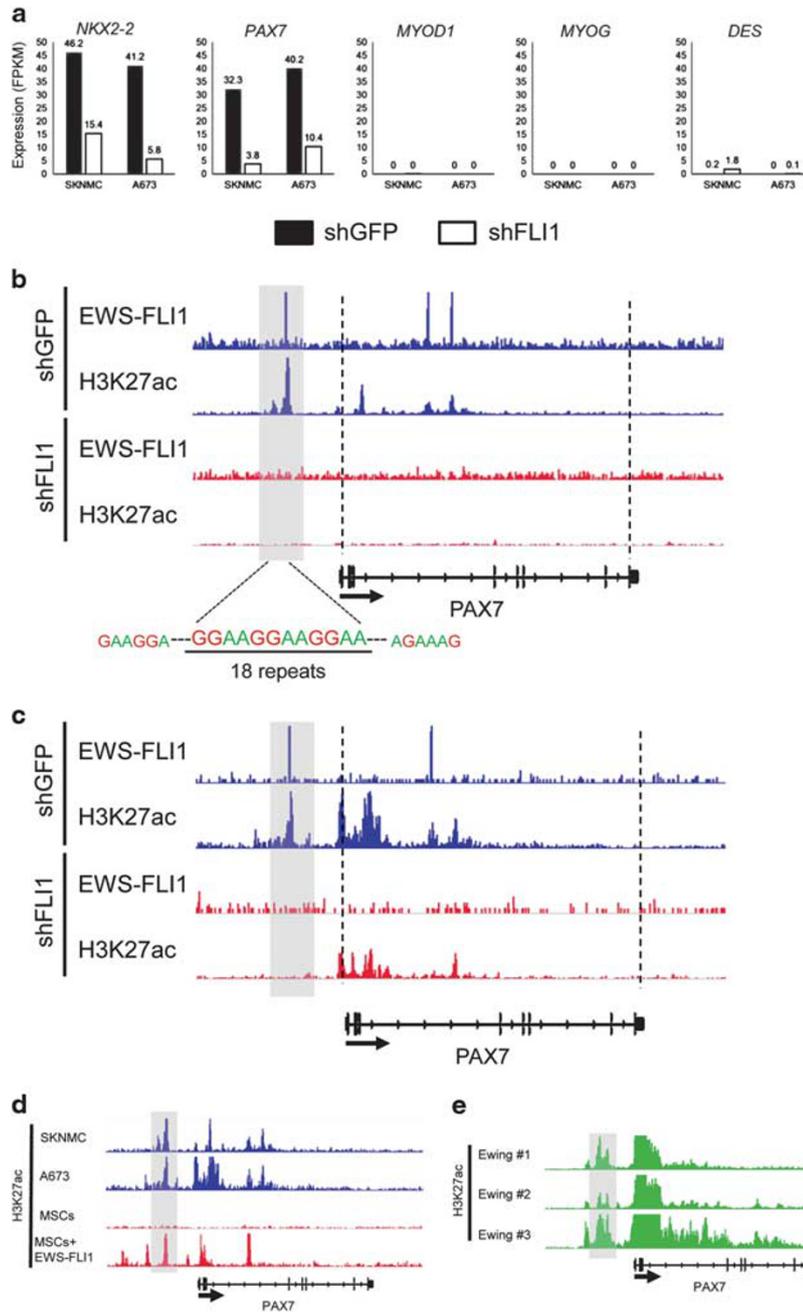


Figure 5 Identification of a *PAX7* enhancer activated by EWSR1-FLI1. (a) Quantification of expression of *NKX2-2*, *PAX7*, *MYOD1*, *MYOG*, and *DES* by RNA-sequencing in two control (shGFP) and EWSR1-FLI1-deficient (shFLI1) Ewing sarcoma cell lines (SKNMC and A673). Relative expression values are quantified as fragments per kilobase of transcript per million mapped reads (FPKM). (b) ChIP-sequencing analysis in SKNMC Ewing sarcoma cells identifies an EWSR1-FLI1-bound enhancer corresponding to a site of H3K27 acetylation proximal to the *PAX7* locus. This site corresponds to a consensus EWS-FLI1-binding site sequence with GGAA repeats. EWS-FLI1 binding and corresponding H3K27 acetylation is lost upon knockdown of EWS-FLI1 expression (shFLI1). (c) ChIP-sequencing analysis in A673 Ewing sarcoma cells identifies an EWSR1-FLI1-bound enhancer corresponding to a site of H3K27 acetylation proximal to the *PAX7* locus, analogous to that identified in SKNMC cells. Again, EWSR1-FLI1 binding and corresponding H3K27 acetylation is lost upon knockdown of EWSR1-FLI1 expression (shFLI1). (d) ChIP-sequencing demonstrates gain of H3K27 acetylation at the EWSR1-FLI1-bound *PAX7* enhancer in mesenchymal stem cells (MSCs) upon overexpression of an EWSR1-FLI1 transgene. (e) ChIP-sequencing analysis identifies a peak of H3K27 acetylation at the site of an EWSR1-FLI1-bound *PAX7* enhancer in primary Ewing sarcoma. Three biologically independent Ewing sarcoma samples are shown.

(Figure 5a), in keeping with the lack of myogenic differentiation generally seen in this tumor.¹¹ Riggi *et al.* have previously shown that EWSR1-FLI1 binds at GGAA repeat elements in the genome to induce

chromatin opening at *de novo* transcriptional enhancers.³⁶ Given that RNA-seq data revealed EWSR1-FLI1-dependent *PAX7* expression in Ewing sarcoma, we hypothesized that the fusion protein

bound an activated enhancer of *PAX7*. Analyzing EWSR1-FLI1 genome-wide ChIP-seq data, we identified a discrete EWSR1-FLI1 binding site, located approximately 20 kb upstream of the *PAX7* promoter, which was present in both SKNMC and A673 Ewing sarcoma cell lines in analyses of anti-FLI1 ChIP-seq data (Figures 5b and c). This binding site coincided with a peak of regulatory H3K27 acetylation (Figures 5b and c), providing support for its characterization as a transcriptional enhancer. Consistent with current models of EWSR1-FLI1 enhancer activation, the position of EWSR1-FLI1 binding and H3K27 acetylation also coincided with a consensus binding sequence of 18 GGAA nucleotide repeats (Figures 5b and c). In support of a model of *PAX7* enhancer activation by EWSR1-FLI1, shRNA-mediated suppression of EWSR1-FLI1 levels in both SKNMC and A673 cells caused concordant loss of H3K27 acetylation at the site of fusion protein binding (Figures 5b and c). These data indicate that EWSR1-FLI1 is necessary for H3K27 acetylation of a *PAX7* enhancer in Ewing sarcoma. As evidence that EWSR1-FLI1 is also sufficient for activation of the *PAX7* enhancer, ChIP-seq analyses of cultured mesenchymal stem cells (MSCs) showed gain of regulatory H3K27 acetylation at the aforementioned EWSR1-FLI1 binding site in MSCs upon exogenous viral expression of EWSR1-FLI1 (Figure 5d). ChIP-seq analyses of genome-wide H3K27 acetylation in three primary Ewing sarcoma samples showed that the same peak of H3K27 acetylation is present at the candidate *PAX7* enhancer in these primary tumor samples (Figure 5e). Taken together, these analyses provide evidence that a *cis* regulatory mechanism of EWSR1-FLI1 binding and enhancer activation leads to expression of *PAX7* in Ewing sarcoma.

Discussion

The results of our studies provide support for the utility of immunohistochemical analysis of *PAX7* expression as a diagnostic marker of Ewing sarcoma. The overall sensitivity of *PAX7* expression for detection of Ewing sarcoma in 103 previously diagnosed cases was 99.0% in this study. Among positive cases, the vast majority showed 2–3+ staining in greater than 50% of cells. In previous work,³⁴ we have shown that, among a diverse array of soft tissue and small round blue-cell neoplasms, expression of *PAX7* is limited to rhabdomyosarcoma and Ewing sarcoma. These previous analyses included additional *EWSR1* translocation-driven tumors—extraskelatal myxoid chondrosarcoma and desmoplastic small round cell tumor—both of which showed no *PAX7* expression.

Support for the use of *PAX7* immunohistochemistry in the diagnosis of Ewing sarcoma is provided by our analyses of high-throughput sequencing data to arrive at a mechanism that accounts for EWSR1 fusion protein-dependent *PAX7* expression.

Assessment of whole-genome EWSR1-FLI1 ChIP-seq data shows binding of the fusion protein at a position 5' to the *PAX7* promoter. This position coincides with a consensus GGAA repeat-containing EWSR1-FLI1 binding sequence and a peak of acetylated H3K27, consistent with an activated enhancer. These data support a model in which transcription of *PAX7* is activated by binding of the oncogenic EWSR1-FLI1 fusion protein at a *cis* regulatory element. To further support this model, we find that knockdown of EWSR1-FLI1 expression concordantly abrogates both H3K27 acetylation at the *PAX7* enhancer and *PAX7* expression in Ewing sarcoma cells.

In practice, the application of *PAX7* immunohistochemistry can be used to narrow a fairly broad differential diagnosis of a morphologically undifferentiated small round cell or spindle-cell tumor. If a *PAX7*-expressing neoplasm is identified, distinguishing between Ewing sarcoma and rhabdomyosarcoma may be aided by additional immunohistochemical markers. For instance, demonstration of myogenic differentiation by expression of DES, MYOD1, or MYOG would strongly favor rhabdomyosarcoma, though rare cases of Ewing sarcoma can exhibit myogenic features (see Discussion below). While expression of CD99 would typically favor Ewing sarcoma, CD99 expression has also been reported in a minor subset of embryonal rhabdomyosarcomas.⁴¹ Similarly, expression of NKX2-2 would support the diagnosis of Ewing sarcoma.²⁹ Still, identification of *EWSR1* translocation in ES or *PAX3/7-FOXO1* translocation in alveolar rhabdomyosarcoma will remain the diagnostic 'gold-standard' for cases with ambiguous immunohistochemical features. Even in the most challenging cases, we anticipate that immunohistochemical markers, such as *PAX7*, will help to curtail costs by enabling efficient triaging of cases for appropriate molecular studies.

PAX7 immunohistochemistry will also have utility in a panel of antibodies for distinguishing Ewing sarcoma and CIC-DUX4 sarcoma. In our study, 99% of Ewing sarcoma expressed *PAX7*, whereas none of 27 cases of CIC-DUX4 sarcoma showed evidence of *PAX7* expression. Although less is known about the molecular underpinnings of aberrant transcriptional pathways in CIC-DUX4 sarcoma, this distinction with regard to *PAX7* expression is presumably due to an absence of transcriptional activation of the *PAX7* locus by the CIC-DUX4 oncogenic fusion protein. Conversely, *WT1* showed significantly increased expression in CIC-DUX4 sarcoma relative to Ewing sarcoma by whole-genome gene expression analysis performed in this study, corroborating previous studies and suggesting that paired analysis of *PAX7* and *WT1* by immunohistochemistry can effectively predict the underlying genetic lesion in a large subset of morphologically ambiguous round cell sarcomas. Specifically, *WT1* immunohistochemical analyses have shown 100% specificity and at least 95% sensitivity in distinguishing CIC-DUX4 tumors

from ES;^{31,32} thus, coupling analysis of WT1 with analysis of PAX7, which is similarly sensitive and specific, would make this particular distinction even more straightforward and robust.

In the absence of PAX7, muscle-resident adult stem cells more readily adopt a non-myogenic cell fate,⁴⁰ suggesting a potential role for PAX7 in transcriptional regulation of the myogenic gene expression program. Indeed, PAX7 binds regulatory elements, induces local histone modifications, and activates transcription of downstream myogenic target genes, such as *MYF5* and *MYOD1*.^{42,43} Although ES cases only rarely exhibit significant myogenic differentiation,¹¹ several cases of so-called 'polyphenotypic' or 'hybrid' tumors have been reported.^{44,45} Given evidence that PAX7 has a role in regulating myogenic differentiation, our observation of *EWSR1* translocation-dependent PAX7 expression in Ewing sarcoma, offers an intriguing molecular explanation for these polyphenotypic tumors. At the same time, since PAX7 is expressed in the large majority of Ewing sarcoma cases, what accounts for the relatively infrequent occurrence of the myogenic phenotype in Ewing sarcoma? Mechanisms that could explain this interesting feature include expression of a repressor of myogenic regulatory factor gene transcription or direct inhibition of myogenic regulatory factor gene enhancers by the *EWSR1* fusion protein itself. Future studies of myogenic regulatory factor expression in Ewing sarcoma may provide additional insight.

Immunohistochemical studies have proven to be important tools in the diagnosis of soft tissue sarcomas defined by underlying genetic translocations. To date, these immunohistochemical assays have focused largely on the aberrant expression of one of the genes involved in the pathogenic translocation. Examples of such helpful immunohistochemical signatures include identification of WT1 in desmoplastic small round cell tumor, STAT6 in solitary fibrous tumor, and, to some extent, FLI1 in Ewing sarcoma. More recently, BCOR immunohistochemistry has been found to be a useful indicator of underlying *BCOR* gene abnormalities.⁴⁶ In contrast to these approaches, the application of immunohistochemistry presented here for the diagnosis of Ewing sarcoma uniquely exploits a characteristic transcriptional readout of the disease-causing translocation to aid in diagnosis. Presumably, unique transcriptional targets, or unique combinations of targets, exist for other translocation-driven sarcomas and might be analyzed in a similar manner. Unbiased comparisons of whole-genome gene expression data can help to identify these unique signatures, while mechanistic insights can be used both to justify the utility of a biomarker and account for features of a marker's sensitivity and specificity. We anticipate that application of these approaches will aid in discovery of a new class of biomarkers for an array of sarcomas.

Acknowledgments

The Amschwand Sarcoma Cancer Foundation is thanked for its generous support of this work (WLW, DRI, and AJL). This study was funded by Departmental Research Funds.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

- 1 Carpentieri DF, Qualman SJ, Bowen J, *et al*. Protocol for the examination of specimens from pediatric and adult patients with osseous and extraosseous Ewing sarcoma family of tumors, including peripheral primitive neuroectodermal tumor and Ewing sarcoma. *Arch Pathol Lab Med* 2005;129:866–873.
- 2 de Chadarevian JP, Vekemans M, Seemayer TA. Reciprocal translocation in small-cell sarcomas. *N Engl J Med* 1984;311:1702–1703.
- 3 Turc-Carel C, Aurias A, Mugneret F, *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;32:229–238.
- 4 Whang-Peng J, Triche TJ, Knutsen T, *et al*. Chromosome translocation in peripheral neuroepithelioma. *N Engl J Med* 1984;311:584–585.
- 5 Delattre O, Zucman J, Plougastel B, *et al*. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 1992;359:162–165.
- 6 Sorensen PH, Lessnick SL, Lopez-Terrada D, *et al*. A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat Genet* 1994;6:146–151.
- 7 Jeon IS, Davis JN, Braun BS, *et al*. A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. *Oncogene* 1995;10:1229–1234.
- 8 Urano F, Umezawa A, Yabe H, *et al*. Molecular analysis of Ewing's sarcoma: another fusion gene, EWS-E1AF, available for diagnosis. *Jpn J Cancer Res* 1998;89:703–711.
- 9 Peter M, Couturier J, Pacquement H, *et al*. A new member of the ETS family fused to EWS in Ewing tumors. *Oncogene* 1997;14:1159–1164.
- 10 Szuhai K, Ijszenga M, de Jong D, *et al*. The NFATc2 gene is involved in a novel cloned translocation in a Ewing sarcoma variant that couples its function in immunology to oncology. *Clin Cancer Res* 2009;15:2259–2268.
- 11 Folpe AL, Goldblum JR, Rubin BP, *et al*. Morphologic and immunophenotypic diversity in Ewing family tumors: a study of 66 genetically confirmed cases. *Am J Surg Pathol* 2005;29:1025–1033.
- 12 Folpe AL, Schmidt RA, Chapman D, *et al*. Poorly differentiated synovial sarcoma: immunohistochemical distinction from primitive neuroectodermal tumors and high-grade malignant peripheral nerve sheath tumors. *Am J Surg Pathol* 1998;22:673–682.
- 13 Olsen SH, Thomas DG, Lucas DR. Cluster analysis of immunohistochemical profiles in synovial sarcoma, malignant peripheral nerve sheath tumor, and Ewing sarcoma. *Mod Pathol* 2006;19:659–668.

- 14 Granter SR, Renshaw AA, Fletcher CD, *et al*. CD99 reactivity in mesenchymal chondrosarcoma. *Hum Pathol* 1996;27:1273–1276.
- 15 Shin S-J, Lee H, Jung G, *et al*. Expression of CD99 in multiple myeloma: a clinicopathologic and immunohistochemical study of 170 cases. *Korean J Pathol* 2014;48:209–216.
- 16 Pelosi G, Frassetto F, Sonzogni A, *et al*. CD99 immunoreactivity in gastrointestinal and pulmonary neuroendocrine tumours. *Virchows Arch* 2000;437:270–274.
- 17 Lucas DR, Bentley G, Dan ME, *et al*. Ewing sarcoma vs lymphoblastic lymphoma. A comparative immunohistochemical study. *Am J Clin Pathol* 2001;115:11–17.
- 18 Ozdemirli M, Fanburg-Smith JC, Hartmann DP, *et al*. Differentiating lymphoblastic lymphoma and Ewing's sarcoma: lymphocyte markers and gene rearrangement. *Mod Pathol* 2001;14:1175–1182.
- 19 Folpe AL, Hill CE, Parham DM, *et al*. Immunohistochemical detection of FLI-1 protein expression: a study of 132 round cell tumors with emphasis on CD99-positive mimics of Ewing's sarcoma/primitive neuroectodermal tumor. *Am J Surg Pathol* 2000;24:1657–1662.
- 20 Llombart-Bosch A, Navarro S. Immunohistochemical detection of EWS and FLI-1 proteins in Ewing sarcoma and primitive neuroectodermal tumors: comparative analysis with CD99 (MIC-2) expression. *Appl Immunohistochem Mol Morphol* 2001;9:255–260.
- 21 Rossi S, Orvieto E, Furlanetto A, *et al*. Utility of the immunohistochemical detection of FLI-1 expression in round cell and vascular neoplasm using a monoclonal antibody. *Mod Pathol* 2004;17:547–552.
- 22 Folpe AL, Chand EM, Goldblum JR, *et al*. Expression of FLI-1, a nuclear transcription factor, distinguishes vascular neoplasms from potential mimics. *Am J Surg Pathol* 2001;25:1061–1066.
- 23 Mhawech-Fauceglia P, Herrmann FR, Bshara W, *et al*. Friend leukaemia integration-1 expression in malignant and benign tumours: a multiple tumour tissue microarray analysis using polyclonal antibody. *J Clin Pathol* 2007;60:694–700.
- 24 Wang W-L, Patel NR, Caragea M, *et al*. Expression of ERG, an Ets family transcription factor, identifies ERG-rearranged Ewing sarcoma. *Mod Pathol* 2012;25:1378–1383.
- 25 Choi E-YK, Thomas DG, McHugh JB, *et al*. Undifferentiated small round cell sarcoma with t(4;19)(q35;q13.1) CIC-DUX4 fusion: a novel highly aggressive soft tissue tumor with distinctive histopathology. *Am J Surg Pathol* 2013;37:1379–1386.
- 26 Graham C, Chilton-MacNeill S, Zielenska M, *et al*. The CIC-DUX4 fusion transcript is present in a subgroup of pediatric primitive round cell sarcomas. *Hum Pathol* 2012;43:180–189.
- 27 Italiano A, Sung YS, Zhang L, *et al*. High prevalence of CIC fusion with double-homeobox (DUX4) transcription factors in EWSR1-negative undifferentiated small blue round cell sarcomas. *Genes Chromosomes Cancer* 2012;51:207–218.
- 28 Smith SC, Buehler D, Choi E-YK, *et al*. CIC-DUX sarcomas demonstrate frequent MYC amplification and ETS-family transcription factor expression. *Mod Pathol* 2015;28:57–68.
- 29 Hung YP, Fletcher CDM, Hornick JL. Evaluation of NKX2-2 expression in round cell sarcomas and other tumors with EWSR1 rearrangement: imperfect specificity for Ewing sarcoma. *Mod Pathol* 2016;29:370–380.
- 30 Yoshida A, Sekine S, Tsuta K, *et al*. NKX2.2 is a useful immunohistochemical marker for Ewing sarcoma. *Am J Surg Pathol* 2012;36:993–999.
- 31 Hung YP, Fletcher CD, Hornick JL. Evaluation of ETV4 and WT1 expression in CIC-rearranged sarcomas and histologic mimics. *Mod Pathol* 2016;29:1324–1334.
- 32 Specht K, Sung Y-S, Zhang L, *et al*. Distinct transcriptional signature and immunoprofile of CIC-DUX4 fusion-positive round cell tumors compared to EWSR1-rearranged Ewing sarcomas: further evidence toward distinct pathologic entities. *Genes Chromosomes Cancer* 2014;53:622–633.
- 33 Siegele B, Roberts J, Black JO, *et al*. DUX4 immunohistochemistry is a highly sensitive and specific marker for CIC-DUX4 fusion-positive round cell tumor. *Am J Surg Pathol* 2017;41:423–429.
- 34 Charville GW, Varma S, Forgó E, *et al*. PAX7 expression in rhabdomyosarcoma, related soft tissue tumors, and small round blue cell neoplasms. *Am J Surg Pathol* 2016;40:1305–1315.
- 35 McCarty KS, Szabo E, Flowers JL, *et al*. Use of a monoclonal anti-estrogen receptor antibody in the immunohistochemical evaluation of human tumors. *Cancer Res* 1986;46:4244s–4248s.
- 36 Riggi N, Knoechel B, Gillespie SM, *et al*. EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell* 2014;26:668–681.
- 37 Trapnell C, Roberts A, Goff L, *et al*. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 2012;7:562–578.
- 38 Zheng Z, Liebers M, Zhelyazkova B, *et al*. Anchored multiplex PCR for targeted next-generation sequencing. *Nat Med* 2014;20:1479–1484.
- 39 Charville GW, Cheung TH, Yoo B, *et al*. *Ex vivo* expansion and *in vivo* self-renewal of human muscle stem cells. *Stem Cell Rep* 2015;5:621–632.
- 40 Seale P, Sabourin LA, Girgis-Gabardo A, *et al*. Pax7 is required for the specification of myogenic satellite cells. *Cell* 2000;102:777–786.
- 41 Ramani P, Rampling D, Link M. Immunocytochemical study of 12E7 in small round-cell tumours of childhood: an assessment of its sensitivity and specificity. *Histopathology* 1993;23:557–561.
- 42 Relaix F, Montarras D, Zaffran S, *et al*. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J Cell Biol* 2006;172:91–102.
- 43 McKinnell IW, Ishibashi J, Le Grand F, *et al*. Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. *Nat Cell Biol* 2008;10:77–84.
- 44 Barisella M, Collini P, Orsenigo M, *et al*. Unusual myogenic and melanocytic differentiation of soft tissue pPNETs: an immunohistochemical and molecular study of 3 cases. *Am J Surg Pathol* 2010;34:1002–1006.
- 45 Sorensen PH, Shimada H, Liu XF, *et al*. Biphenotypic sarcomas with myogenic and neural differentiation express the Ewing's sarcoma EWS/FLI1 fusion gene. *Cancer Res* 1995;55:1385–1392.
- 46 Kao Y-C, Sung Y-S, Zhang L, *et al*. BCOR overexpression is a highly sensitive marker in round cell sarcomas with BCOR genetic abnormalities. *Am J Surg Pathol* 2016;40:1670–1678.