Distinct methylation profiles characterize fusion-positive and fusion-negative rhabdomyosarcoma

Wenyue Sun[1], Bishwanath Chatterjee[1], Yonghong Wang[2], Holly S Stevenson[2], Daniel C Edelman[2], Paul S Meltzer[2] and Frederic G Barr[1]

[1]Cancer Molecular Pathology Section, Laboratory of Pathology, Bethesda, MD, USA and [2]Clinical Molecular Profiling Core, Genetics Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA

Rhabdomyosarcoma comprises two major subtypes, fusion positive (*PAX3–FOXO1* or *PAX7–FOXO1*) and fusion negative. To investigate the significance of DNA methylation in these subtypes, we analyzed methylation profiles of 37 rhabdomyosarcoma tumors and 10 rhabdomyosarcoma cell lines, as well as 8 normal tissues. Unsupervised clustering of DNA methylation clearly distinguished the fusion-positive and fusion-negative subsets. The fusion-positive tumors showed substantially lower overall levels of methylation compared with fusion-negative tumors. Comparison with the methylation pattern of normal skeletal muscle and bone marrow indicates that fusion-negative rhabdomyosarcoma is more similar to these normal tissues compared with fusion-positive rhabdomyosarcoma, and suggests that many of the methylation differences between these subtypes arise from 'aberrant' hyper- and hypomethylation events in fusion-positive rhabdomyosarcoma. Integrative methylation and gene expression analysis revealed that methylation differences between fusion-positive and fusion-negative tumors could either be positively or negatively associated with mRNA expression. There was no significant difference in the distribution of PAX3–FOXO1-binding sites between genes with and without differential methylation. However, the finding that PAX3–FOXO1-binding sites were enriched among genes that were both differentially methylated and differentially expressed suggests that the fusion protein interacts with DNA methylation to regulate target gene expression. An 11-gene DNA methylation signature, classifying the rhabdomyosarcoma tumors into fusion-positive and fusion-negative subsets, was established and validated by pyrosequencing assays. Notably, *EMILIN1* (part of the 11-gene signature) showed higher methylation and lower mRNA expression in fusion-positive compared with fusion-negative tumors, and demonstrated demethylation and re-expression in multiple fusion-positive cell lines after treatment with 5-aza-2′-deoxycytidine. In conclusion, our study demonstrates that fusion-positive and fusion-negative rhabdomyosarcoma tumors possess characteristic methylation profiles that contribute to the expression differences between these fusion subtypes. These findings indicate an important relationship between fusion status and epigenetic changes in rhabdomyosarcoma, present a novel approach for ascertaining fusion status, and may identify new therapeutic targets in rhabdomyosarcoma.
Modern Pathology (2015) **28**, 1214–1224; doi:10.1038/modpathol.2015.82; published online 31 July 2015

The pediatric soft tissue cancer rhabdomyosarcoma has been traditionally classified by histology into two major subtypes, alveolar (~30%) and embryonal (~70%). Most alveolar rhabdomyosarcoma cases (~80%) have chromosomal translocations that join the DNA-binding domain of *PAX3* or *PAX7* to the

transactivation domain of *FOXO1*. The *PAX3–FOXO1* and *PAX7–FOXO1* fusions (henceforth referred to as PAX3/7–FOXO1) encode potent transcription factors that contribute to tumorigenesis by altering growth and apoptotic pathways, modulating myogenic differentiation, and stimulating metastatic pathways.[1,2] Several studies have shown that the *PAX3/7–FOXO1* fusion is a predictor of unfavorable outcome in children with rhabdomyosarcoma, and that fusion status is superior to the histologic subtype in predicting rhabdomyosarcoma outcome.[3–5] In addition, subcategorization by fusion status more

Methylation differences in rhabdomyosarcoma
W Sun *et al*
1215

accurately captures the genomic landscape, expression pattern and biology of rhabdomyosarcoma.[4,6,7]

Epigenetic alterations such as DNA methylation can also influence tumor development by altering the expression of tumor suppressor genes and oncogenes. DNA methylation occurs in CpG islands, CpG shores and distal regulatory regions. Previous methylation studies identified genes epigenetically silenced in rhabdomyosarcoma tumors, such as *RASSF1* (ref 8) and *HIC1*.[9] In addition, other studies suggested that there may be distinct DNA methylation patterns in alveolar and embryonal rhabdomyosarcoma.[10,11]

In this study, we undertook the first systematic, genome-wide comparison of DNA methylation profiles between fusion-positive and fusion-negative rhabdomyosarcoma. We delineated the association between *PAX3/7–FOXO1* fusion status and DNA methylation through a combination of DNA methylation and gene expression assays, including array-based profiling. In addition, we report an 11-probe DNA methylation signature that is sufficient to classify fusion-positive and fusion-negative rhabdomyosarcoma tumors. These methylation findings were validated by pyrosequencing and extended by 5-aza-2′-deoxycytidine inhibition studies.

Materials and methods

Human Tissue Samples

Frozen tissue samples were received from the Children's Oncology Group Biopathology Center, as part of a protocol approved by the University of Pennsylvania Institutional Review Board. The far majority of the specimens were collected at the time of diagnosis before any therapy was started. Samples were reviewed to identify cases with > 70% tumor cells. Fusion status was determined by reverse transcriptase-PCR or fluorescence *in situ* hybridization for all samples without unambiguous embryonal rhabdomyosarcoma histology.[12,13] The normal skeletal muscle and bone marrow samples were obtained post mortem from eight anonymized individuals with no history or signs of cancer involvement (Supplementary Table 1).

Cell Lines

Six fusion-positive cell lines (JR, MP4, RH5, RH28, RH30, and RH41) and five fusion-negative cell lines (RD, RH6, RH18, Birch, and SMS-CTR) were used in this study. The source of these cell lines was as follows: RD, Dr L Helman; MP4, Dr T Cripe; RH5, Dr J Khan; RH6, Dr P Houghton; RH18 and Birch, Dr M Tsokos; RH28, Dr B Emanuel; RH30, American Type Culture Collection; RH41, JR, and SMS-CTR, Dr C Linardic. Short tandem repeat analysis revealed that Birch and RH6 were clonally related and all other lines were clonally independent. Genotyping results

were also compared with data from the Children's Oncology Group Cell Culture and Xenograft Repository (http://www.cogcell.org) to verify identity of the cell lines. Cells were cultured using DMEM or RPMI-1640 media (Life Technologies) supplemented with 10% or 20% FBS and antibiotic antimycotic at 37 °C in 5% CO_2.

DNA Methylation Profiling

Genomic DNA was extracted using the QIAamp DNA Mini Kit (Qiagen) or Isolate II Genomic DNA Kit (Bioline). Bisulfite conversion of genomic DNA was done with the EZ-96 DNA Methylation Kit (Zymo Research) according to the manufacturer's protocol with modifications for the Illumina Infinium Methylation Assay.

Bisulfite-converted genomic DNA was analyzed using Illumina's Infinium HM27 or HM450 methylation platform. Data were extracted using the lumi or methylumi package available through Bioconductor.[14] The β-value was chosen as the measure of methylation, ranging from 0 (no methylation of any allele) to 1.0 (complete methylation of both alleles). Probes that correspond to sequences on the X and Y chromosomes, overlap a single nucleotide polymorphism or repetitive element, or have a detection P-value > 0.05 in any sample were excluded. We obtained transcription start site and exon coordinates for RefSeq genes, and CpG island annotations from the UCSC hg19 reference genome. Promoters were defined as a region encompassing 2 kb upstream and 500 bp downstream of the RefSeq transcription start site. CpG shores were defined as 2 kb regions directly upstream and downstream of CpG islands.

Empirical Bayes comparisons were performed using the limma package in R/Bioconductor to determine the difference in methylation between categorical groups.[15] A probe was considered differentially methylated if the Benjamini and Hochberg-adjusted P-value was > 0.05 and $|\Delta\beta| \geq 0.2$.[16] To develop a DNA methylation signature, Lasso regularization was applied to identify CpGs for which methylation levels predicted classification of fusion-positive and fusion-negative rhabdomyosarcoma.[17]

Expression Analysis

Total RNA was isolated with RNA STAT-60 (Tel-Test). RNAs isolated from tumor samples were run on Affymetrix GeneChip Human Genome U133 Plus 2.0 Arrays. The expression microarray data were normalized by the Robust Multiarray Average method.[18] The Significance Analysis of Microarrays technique was applied to identify differentially expressed genes.[19] For quantitative reverse transcriptase-PCR, reverse transcription was performed with the SuperScript III First-Strand Synthesis System (Life Technologies). *EMILIN1* RNA expression was quantified with a Taqman assay

(Life Technologies); and normalized for 18S ribosomal RNA expression, using the $2^{-\Delta\Delta Ct}$ method.[20]

Additional Data Analysis

Hierarchical clustering was performed using Pearson's correlation distance and Ward's agglomeration method. The principal component analysis was performed using the prcomp function in the statistical package of R. Gene Set Analysis was performed with Molecular Signatures Database version 2.5.1. We determined the statistical significance of the overlap with each functional category using the Fisher's test. A Benjamini and Hochberg-adjusted *P*-value < 0.05 was considered significant. Correlation of methylation and expression was computed using the Pearson's method, and significance (*P* < 0.05) was assessed using permutation testing.

Pyrosequencing

Pyrosequencing was carried out on the PyroMark MD Pryosequencing System (Qiagen).[21] Methylation was quantified using Pyro-Q-CpG Software that calculates the ratio of converted to unconverted cytosines at each CpG.

Treatment of Cells with 5-Aza-2′-Deoxycytidine

All cell lines were seeded (5×10^4) in their respective culture medium in 96-well plates and maintained for 24 h before treating with various concentrations of 5-aza-2′-deoxycytidine (dissolved in 50% DMSO) for 5 days.[22] Medium containing 5-aza-2′-deoxycytidine was replaced every 24 h, and growth rate was measured by CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega).

Results

Comparison of Methylation Profiles between Fusion-Positive and Fusion-Negative Rhabdomyosarcoma

To explore the DNA methylome of fusion-positive and fusion-negative rhabdomyosarcoma, we used the HM 27 K methylation platform to examine DNA methylation profiles in a discovery cohort including 20 fusion-positive and 17 fusion-negative rhabdomyosarcoma samples (Supplementary Table 1). We first selected the top 10% of probes whose DNA methylation level (*β*-value) varied the most across the 37 rhabdomyosarcoma samples, and conducted an unsupervised hierarchical clustering. This analysis separated the rhabdomyosarcoma tumors into two groups consisting exclusively of either fusion-positive or fusion-negative rhabdomyosarcoma cases (Figure 1a, *P* < 0.001). Within the fusion-positive tumors, there was no significant difference in the distribution of *PAX3–FOXO1*-positive *vs PAX7–FOXO1*-positive cases (*P* = 0.11). A principal

component analysis revealed a similar finding as the hierarchical clustering analysis (Figure 1b).

To validate our findings in an independent set of rhabdomyosarcoma samples, we used the HM 450 K methylation platform to analyze genome-wide methylation in five fusion-positive (MP4, RH5, RH28, RH30, and RH41) and five fusion-negative (RD, RH6, RH18, Birch, and SMS-CTR) rhabdomyosarcoma cell lines. In an unsupervised analysis of this cell line panel, the fusion-positive and fusion-negative cell lines clustered separately based on their methylation differences (Figure 1c).

We next compared the relative methylation (*β*) density distribution of these CpG sites (10% with highest variability) between fusion-positive and fusion-negative rhabdomyosarcoma. In comparison with fusion-negative tumors, fusion-positive rhabdomyosarcoma had an increased frequency of CpG sites with a lower level of methylation (*β* < 0.3) and a decreased frequency of CpG sites with a higher level of methylation (*β* > 0.7; Figure 1d). The overall methylation levels in fusion-positive and fusion-negative rhabdomyosarcoma also differed, with fusion-positive cases showing substantially lower levels of methylation (*P* < 0.01; Figure 1e).

Identification of Differentially Methylated Genes in Fusion-Positive *vs* Fusion-Negative Rhabdomyosarcoma

We next performed a supervised analysis of our methylation data set to identify differentially methylated probes. Using our criteria of a Benjamini and Hochberg-adjusted *P*-value < 0.05 and $|\Delta\beta| \geq 0.20$, we identified 348 probes (268 genes) in our discovery cohort that were significantly hypermethylated in fusion-positive compared with fusion-negative rhabdomyosarcoma, and 1055 probes (770 genes) that were significantly hypomethylated in fusion-positive tumors (Figure 2a; Supplementary Tables 2 and 3). Gene set analysis of the differentially methylated genes in fusion-positive *vs* fusion-negative rhabdomyosarcoma demonstrated the enrichment of GO BP terms such as development, immune response, and signal transduction (Supplementary Table 4).

A recent genome-wide analysis of fusion-positive rhabdomyosarcoma presented a high-resolution map of PAX3–FOXO1-binding sites in 1072 genes.[23] Comparison of PAX3–FOXO1-binding sites with methylation status revealed that PAX3–FOXO1-binding sites are similarly distributed between genes with (70/1038) and without (650/12 020) differential methylation (*P* = 0.076). This finding provides evidence that PAX3–FOXO1 does not directly contribute to the differential methylation pattern.

Comparison with Methylation Pattern in Normal Tissues

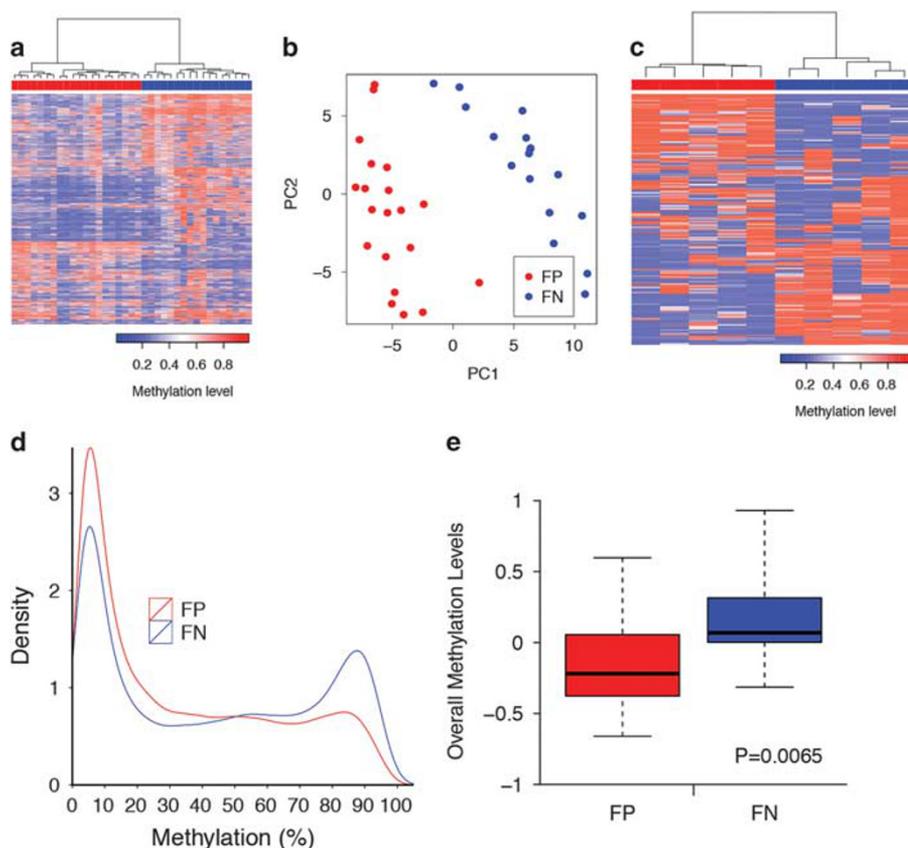In our comparative analysis of fusion-positive and fusion-negative rhabdomyosarcoma, it is not

Figure 1 Unsupervised analyses of methylation profiles distinguish fusion-positive from fusion-negative rhabdomyosarcoma. (**a**) Unsupervised hierarchical clustering of methylation profiles from 37 rhabdomyosarcoma tumors. Pearson's correlation and ward linkage were applied to the top 10% of probes with greatest variation (based on standard deviation) in the clustering analysis. Each column represents an individual sample with the 20 fusion-positive rhabdomyosarcoma cases shown as red rectangles and the 17 fusion-negative rhabdomyosarcoma cases shown as blue rectangles above the heat map. The blue and red lines in the heat map indicate low and high methylation levels according to the scale below the heat map. (**b**) Principal component analysis of rhabdomyosarcoma tumor methylation data. 20 fusion-positive (red dots) and 17 fusion-negative (blue dots) rhabdomyosarcoma cases were clustered with the 10% of probes with greatest sample-to-sample variation (based on standard deviation). (**c**) Unsupervised hierarchical clustering of DNA methylation profiles from rhabdomyosarcoma cell lines. Clustering of the cases was based on the top 1000 probes with the greatest sample-to-sample variability in methylation across the 10 rhabdomyosarcoma cell lines. In the dendrogram, each cell line is shown in one column, with the fusion-positive rhabdomyosarcoma lines indicated by red rectangles and fusion-negative rhabdomyosarcoma lines indicated by blue rectangles above the heat map. (**d**) Distribution of DNA methylation (*β*) levels. The most variable (top 10%) CpG sites were examined in 20 fusion-positive (red) and 17 fusion-negative (blue) rhabdomyosarcoma tumors. The *y* axis corresponds to frequency of a methylation level (expressed as *β*) such that the total probability (area under curve) is equal to one. Density plots for *β*-values were computed with the density function (kernel density estimation with a Gaussian smoothing function). (**e**) Comparison of overall methylation levels. The methylation levels of the top 10% of probes with greatest variation (based on standard deviation) are depicted by standardized average methylation levels across the genome in the 20 fusion-positive (red) and 17 fusion-negative (blue) samples. FN, fusion-negative rhabdomyosarcoma; FP, fusion-positive rhabdomyosarcoma.

possible to determine whether any given difference represents an aberrant event in fusion-positive or fusion-negative rhabdomyosarcoma, or possibly in both subtypes. To investigate which methylation states may be representative of a 'normal' state, we compared methylation in rhabdomyosarcoma with normal skeletal muscle and normal bone marrow, which are both mesodermal derivatives.[24] We selected perinatal muscle because of the higher fraction of undifferentiated mononuclear muscle cells.[25] In an unsupervised analysis based on differentially methylated CpG sites in fusion-positive *vs* fusion-negative rhabdomyosarcoma (Figure 2b), the four samples of each normal tissue

tightly clustered indicating a high level of concordance in methylation patterns. All fusion-positive samples were on one main branch, while the normal muscle and marrow samples were grouped together with fusion-negative samples on the other main branch. Fusion-negative samples EN-4 and EN-5 clustered more closely to normal muscle than to the other fusion-negative rhabdomyosarcoma cases; the categorization of EN-4 as a well-differentiated rhabdomyosarcoma confirms the close association of this case to normal muscle.

The overall clustering indicates that fusion-negative rhabdomyosarcoma is more similar to normal muscle and/or bone marrow than fusion-positive
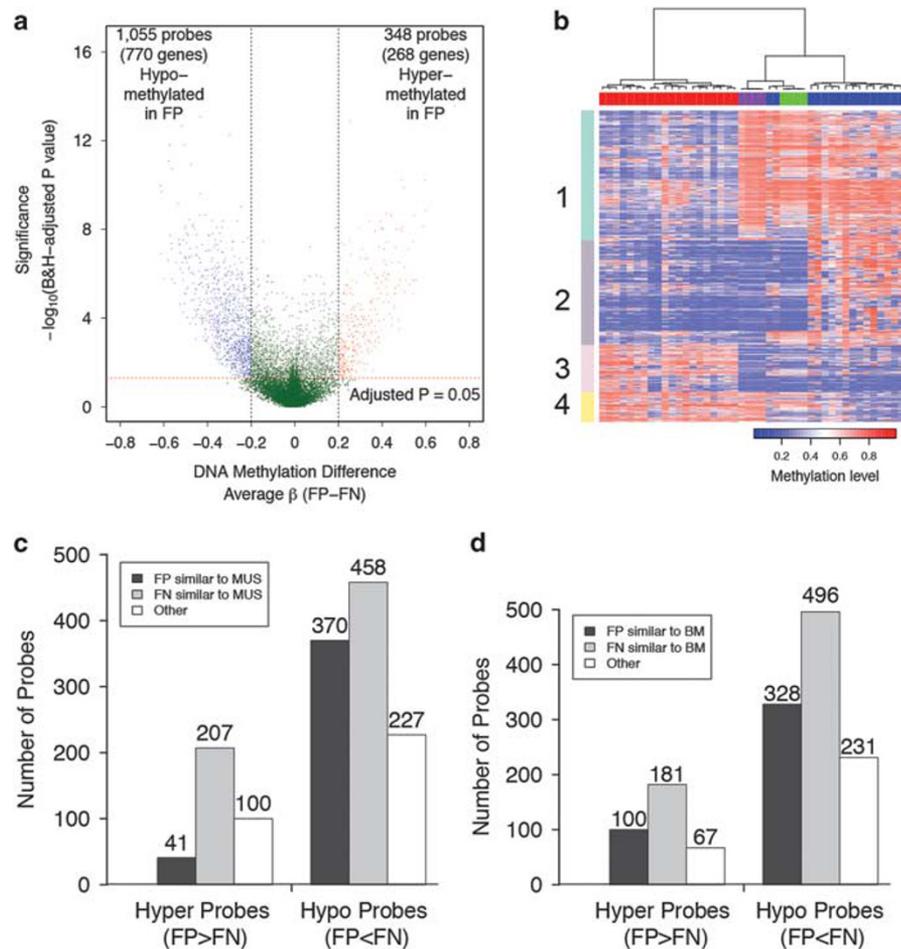
Figure 2 Identification of differentially methylated probes in fusion-positive and fusion-negative rhabdomyosarcoma. (**a**) Volcano plot displaying significant, differentially methylated CpG sites. The β-value difference in DNA methylation between 20 fusion-positive and 17 fusion-negative rhabdomyosarcoma cases is plotted on the *x* axis, and the *P*-value for FDR-corrected Limma test of differences between fusion-positive and fusion-negative rhabdomyosarcoma ($-1 \times \log_{10}$ scale) is plotted on the *y* axis. Probes that are significantly different between fusion-positive and fusion-negative rhabdomyosarcoma are colored in red (hypermethylated in fusion-positive rhabdomyosarcoma) or blue (hypomethylated in fusion-positive rhabdomyosarcoma). (**b**) Heat map comparing DNA methylation profiles of normal skeletal muscle and bone marrow with fusion-positive and fusion-negative rhabdomyosarcoma. The clustering analysis was performed using the differentially methylated genes/probes identified in the comparison of fusion-positive *vs* fusion-negative rhabdomyosarcoma. Pearson correlation distance and Ward linkage were applied. Each individual sample is shown in one column with the 20 fusion-positive rhabdomyosarcoma cases, 17 fusion-negative rhabdomyosarcoma cases, 4 normal muscle cases, and 4 normal bone marrow cases labeled as red, blue, green, and purple rectangles, respectively, above the heat map. The four numbered rectangles on the left side highlight groups of differentially methylated genes, as described in the text. (**c** and **d**) Comparison of methylation of differentially methylated probes (in fusion-positive *vs* fusion-negative rhabdomyosarcoma) to methylation in normal muscle (**c**) or bone marrow (**d**). B&H, Benjamini and Hochberg; BM, normal bone marrow; FN, fusion-negative rhabdomyosarcoma; FP, fusion-positive rhabdomyosarcoma; Hyper, hypermethylated; Hypo, hypomethylated; MUS, normal skeletal muscle.

rhabdomyosarcoma. Though the methylation status of numerous differentially methylated sites is similar in fusion-negative tumors and the normal tissues (Figure 2b, groups 1 and 3), there are also differentially methylated sites for which the methylation status in fusion-positive tumors closely resembles either or both normal tissues (Figure 2b, groups 2 and 4). Of the probes that were hypermethylated in fusion-positive rhabdomyosarcoma, 41 (12%) showed similar methylation levels in fusion-positive rhabdomyosarcoma and normal muscle whereas 207 (59%) of the probes exhibited similar methylation levels in fusion-negative

rhabdomyosarcoma and muscle. In addition, of the probes that were hypomethylated in the fusion-positive cases, 370 (35%) and 458 (43%) probes showed similarities between fusion-positive rhabdomyosarcoma and muscle, and between fusion-negative rhabdomyosarcoma and muscle, respectively (Figure 2c). A comparison of methylation in the rhabdomyosarcoma cases and bone marrow samples reveals comparable results (Figure 2d). These methylation studies thus reveal similarities of both rhabdomyosarcoma subtypes to the normal tissues for different sets of genes. We hypothesize that groups 1 and 3 represent 'aberrant' hypo- and

hypermethylation events occurring in fusion-positive tumors, and that groups 2 and 4 represent 'aberrant' hyper- and hypomethylation events occurring in fusion-negative tumors.

Correlation of Methylation and Gene Expression in Fusion-Positive and Fusion-Negative Rhabdomyosarcoma

We next explored associations between DNA methylation and gene expression to understand the functional consequences of DNA methylation alterations in rhabdomyosarcoma. For this analysis, we measured genome-wide gene expression in 33 fusion-positive and 25 fusion-negative rhabdomyosarcoma cases, using HG-U133 Plus 2.0 arrays. Our unsupervised analysis confirmed that fusion-positive and fusion-negative rhabdomyosarcoma cases clustered separately according to the presence or absence of the PAX3/7–FOXO1 fusion[2,7] (Supplementary Figure 1).

We next used the significance analysis of microarrays technique to identify genes that are differentially expressed between fusion-positive and fusion-negative rhabdomyosarcoma. This analysis (with criteria of $\Delta = 1.2$, absolute fold change > 2, and false discovery rate < 0.01) identified 548 overexpressed and 454 underexpressed genes in fusion-positive rhabdomyosarcoma. In total, we found 83 genes that were both differentially expressed and differentially methylated between fusion-positive and fusion-negative rhabdomyosarcoma. These 83 genes represent 12.8% of differentially expressed genes, and 8.6% of differentially methylated genes, and include 13 genes hypermethylated and underexpressed in fusion-positive compared with fusion-negative rhabdomyosarcoma, 39 hypomethylated overexpressed genes, 17 hypermethylated overexpressed genes, and 14 hypomethylated underexpressed genes (Supplementary Table 5).

We investigated the correlation between DNA methylation and gene expression among nine fusion-positive and seven fusion-negative rhabdomyosarcoma tumors that had both DNA methylation and gene expression data available. We identified an overall inverse correlation between promoter methylation and RNA expression across all 16 rhabdomyosarcoma samples and all 12 247 evaluable genes (Supplementary Figure 2). Further, we analyzed the correlation between promoter methylation and gene expression for the 966 differentially methylated genes available on the expression array. Pearson's correlation analysis revealed 258 genes with significant positive and/or negative methylation expression correlation based on permutation tests ($P < 0.05$) and 708 genes without significant correlation.

For each differentially methylated gene, we performed association analysis between the presence of PAX3–FOXO1-binding sites and the presence of a correlation between methylation with expression

(Supplementary Table 6). Genes with PAX3–FOXO1-binding sites were significantly enriched in the differentially methylated genes that showed a correlation of methylation and expression (25/258) *vs* those differentially methylated genes that did not show such a correlation (41/708; $P = 0.043$).

Development of a Methylation Signature Distinguishing Fusion-Positive From Fusion-Negative Rhabdomyosarcoma

We hypothesized that a small panel of DNA methylation biomarkers could be identified to serve as a signature that discriminates fusion-positive from fusion-negative rhabdomyosarcoma. We applied the Lasso method (binomial regression with L1 penalty) to the methylation data from our discovery cohort using the glmnet package developed by Tibshirani *et al*.[17] The Lasso method resulted in the identification of a panel of 11 CpG sites (corresponding to 11 genes), 7 of which are hypomethylated and 4 of which are hypermethylated in fusion-positive rhabdomyosarcoma (Figure 3a, Supplementary Table 7). A heat map depicts the clear separation of fusion-positive and fusion-negative tumors using the 11-gene signature (Figure 3b). Moreover, unsupervised clustering analysis of methylation levels of these 11 genes clearly distinguished the 5 fusion-positive from the 5 fusion-negative rhabdomyosarcoma cell lines (Figure 3c).

We also performed clustering analysis with the expression levels of these 11 genes from our expression array data set of 58 rhabdomyosarcoma cases. Our results showed that the rhabdomyosarcoma cases could be confidently separated into two main groups associated with PAX3/7–FOXO1 fusion status (Fisher test, $P < 0.05$, Figure 3d). Eight of these 11 genes (*ARHGEF3*, *DIRAS3*, *EMILIN1*, *GATA4*, *MYOG*, *NHLH1*, *NOS1*, and *PIPTNM3*) were significantly differentially expressed in the 33 fusion-positive compared with the 25 fusion-negative tumors (Supplementary Figure 3).

Validation of the Methylation Signature by Pyrosequencing

To validate this 11-gene methylation signature, we performed pyrosequencing analysis for 6 of the 11 genes (*DIRAS3*, *EMILIN1*, *LCP1*, *MYOG*, *NHLH1*, and *TGFB1*) in six fusion-positive and five fusion-negative rhabdomyosarcoma samples that were used in our methylation array analysis. This pyrosequencing analysis assayed the CpG sites found on the methylation array, as well as nearby CpG sites for these six genes (Supplementary Figure 4). We observed a high concordance in the methylation levels at the six CpG sites assayed by both methylation arrays and pyrosequencing ($R^2 = 0.82$; Figure 4a). In all six genes, differential methylation at other CpG sites was in agreement with that found
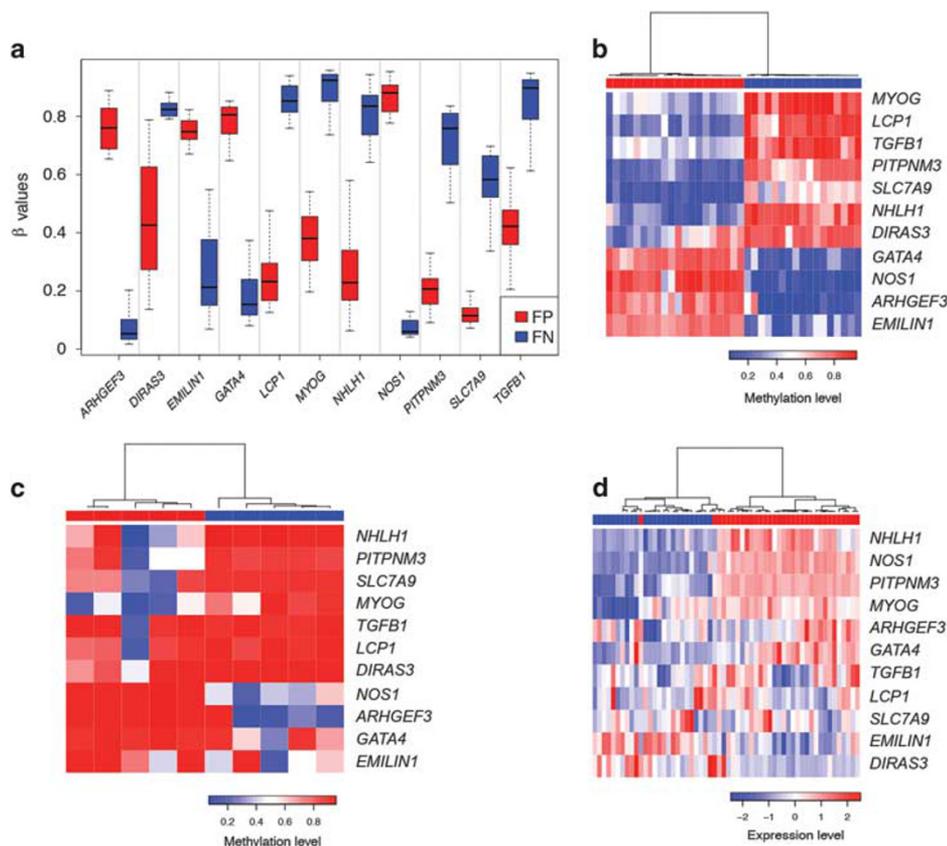
Methylation differences in rhabdomyosarcoma
1220
W Sun *et al*

Figure 3 Development of a DNA methylation signature distinguishing fusion-positive from fusion-negative rhabdomyosarcoma. (**a**) DNA methylation levels of the 11 CpG sites in this DNA methylation signature. (**b** and **c**) Unsupervised clustering of these 11 CpG sites in 37 rhabdomyosarcoma tumors (**b**) and 10 rhabdomyosarcoma cell lines (**c**). The fusion-positive and fusion-negative rhabdomyosarcoma samples are shown as red and blue rectangles, respectively, above the heat map. (**d**) Unsupervised clustering of gene expression of these 11 genes in 58 rhabdomyosarcoma tumor samples. The 33 fusion-positive rhabdomyosarcoma samples and 25 fusion-negative rhabdomyosarcoma samples are labeled as red and blue rectangles above the heat map. The clustering algorithm was based on Pearson's correlation distance metric and Ward linkage on scaled probe-expression data. FN, fusion-negative rhabdomyosarcoma; FP, fusion-positive rhabdomyosarcoma.

by methylation array analysis (Figure 4b), and thus the methylation status of the selected CpG site on the array is predictive of the methylation status of nearby CpGs.

Promoter Demethylation and Re-expression of *EMILIN1* after 5-Aza-2′-Deoxycytidine Treatment

To further examine whether regulation of methylation has a role in fusion-positive and fusion-negative rhabdomyosarcoma development, we treated six fusion-positive and five fusion-negative rhabdomyosarcoma cell lines with the DNA methyltransferase inhibitor 5-aza-2′-deoxycytidine at varying dosages (0, 0.1, 1, and 10 μM) for 5 days.[26] These fusion-positive and fusion-negative cell lines exhibited significant but variable reduction in proliferation after 5-aza-2′-deoxycytidine treatment (Figure 5a).

To further study the relationship between DNA methylation and gene expression in these cell lines, we focused on the *EMILIN1* gene. We determined *EMILIN1* promoter methylation levels by

pyrosequencing the six fusion-positive and five fusion-negative cell lines treated with either 1 μM 5-aza-2′-deoxycytidine or control DMSO for 5 days. In control-treated cells, we found very-high ($>80\%$) DNA methylation levels in five of six fusion-positive lines (JR, MP4, RH5, RH28, and RH41) and one of five fusion-negative lines (RD; Figure 5b). Comparing 5-aza-2′-deoxycytidine and control treatments, we observed a significant decrease in *EMILIN1* promoter methylation levels ($P < 0.05$, methylation change $>30\%$) in the same five fusion-positive and one fusion-negative cell lines.

To determine the relationship of *EMILIN1* promoter methylation to mRNA expression, we quantified *EMILIN1* expression in 5-aza-2′-deoxycytidine and control-treated cells by quantitative reverse transcriptase-PCR (Figure 5c). In control-treated cells, very-low *EMILIN1* expression levels were observed in the same rhabdomyosarcoma lines noted above (five fusion positive and one fusion negative) that had high DNA methylation levels. Furthermore, significantly increased mRNA expression ($\geq$2-fold upregulation and $P < 0.05$) in 5-aza-2′-
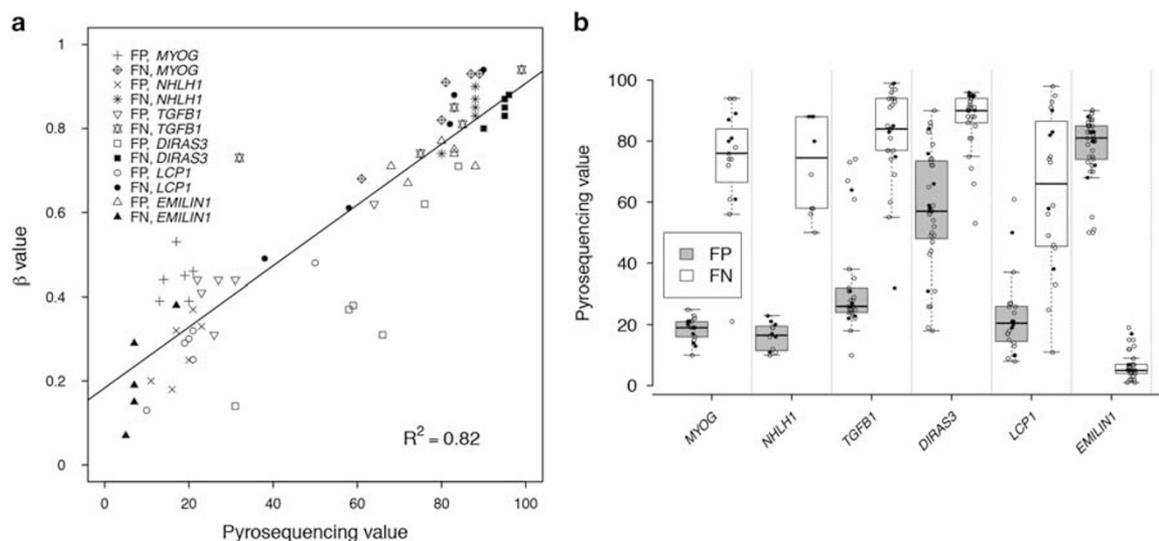
Figure 4 Pyrosequencing studies of six genes from the 11-gene signature. (**a**) Correlation between Infinium methylation score and pyrosequencing values. The methylation status of the specific CpG sites queried on the methylation arrays for six genes (*MYOG*, *NHLH1*, *TGFB1*, *DIRAS3*, *LCP1*, and *EMILIN1*) was analyzed by pyrosequencing in six fusion-positive and five fusion-negative samples, and compared with the β-value from the methylation array. (**b**) Distribution of methylation levels for multiple CpG sites in the six genes in fusion-positive and fusion-negative tumors. For each gene, methylation at the corresponding CpG site (filled circle) in the DNA methylation signature as well as its surrounding CpG sites (open circles) was measured by pyrosequencing in six fusion-positive and five fusion-negative samples. FN, fusion-negative rhabdomyosarcoma; FP, fusion-positive rhabdomyosarcoma.

deoxycytidine-treated cells compared with control-treated cells was found in the same five fusion-positive and one fusion-negative cell lines with high basal methylation and low basal expression levels. These findings indicate that *EMILIN1* promoter methylation and expression are inversely correlated, and that loss of methylation results in increased *EMILIN1* mRNA expression.

Discussion

In this study we performed methylation and expression profiling of fusion-positive and fusion-negative rhabdomyosarcoma tumors to determine whether epigenetic differences exist between these rhabdomyosarcoma subtypes and to investigate how these epigenetic differences contribute to biological differences between these two subtypes. Furthermore, we explored the application of these methylation findings to molecular classification of rhabdomyosarcoma tumors. Ultimately, we hope that these findings will aid in the identification of new targets for rhabdomyosarcoma therapy.

We and others have shown that mRNA expression profiling clearly discriminates fusion-positive and fusion-negative rhabdomyosarcoma, and the differences in overall gene expression are associated with distinct clinicopathologic phenotypes.[2,4,7,27] Some of these expression differences are attributable to the PAX3/PAX7–FOXO1 fusion protein acting as a transcription factor to modulate expression of downstream targets. Other gene expression differences may be related to additional genetic differences, such as point mutations and gene amplification, between fusion-positive and fusion-negative rhabdomyosarcoma.[6] In this study, we provide evidence showing that epigenetic alterations, such as DNA methylation, serve as an additional factor contributing to expression differences between fusion-positive and fusion-negative rhabdomyosarcoma.

The mechanism leading to the DNA methylation differences in fusion-positive and fusion-negative rhabdomyosarcoma is not well understood. We propose that there are 'aberrant' hypomethylation and hypermethylation events in both fusion-positive and fusion-negative tumors. There are a number of possible explanations for these aberrant events. One possibility is that the PAX3/7–FOXO1 fusion protein directly modulates the methylation pattern. However, our finding that differentially methylated genes are not preferentially associated with PAX3/7–FOXO1-binding sites does not support a major role for the fusion proteins in generating methylation differences. However, for at least a subset of genes, the fusion protein interacts with promoter methylation to provide an additional mechanism for regulating gene expression.

A second possibility is that the differential methylation events represent alterations that were selected to collaborate during rhabdomyosarcoma tumorigenesis with the *PAX3/7–FOXO1* fusion or with RAS pathway mutations in fusion-positive and fusion-negative rhabdomyosarcoma, respectively. In particular, *PAX3–FOXO1*-positive rhabdomyosarcoma, which lacks recurrent somatic point mutations
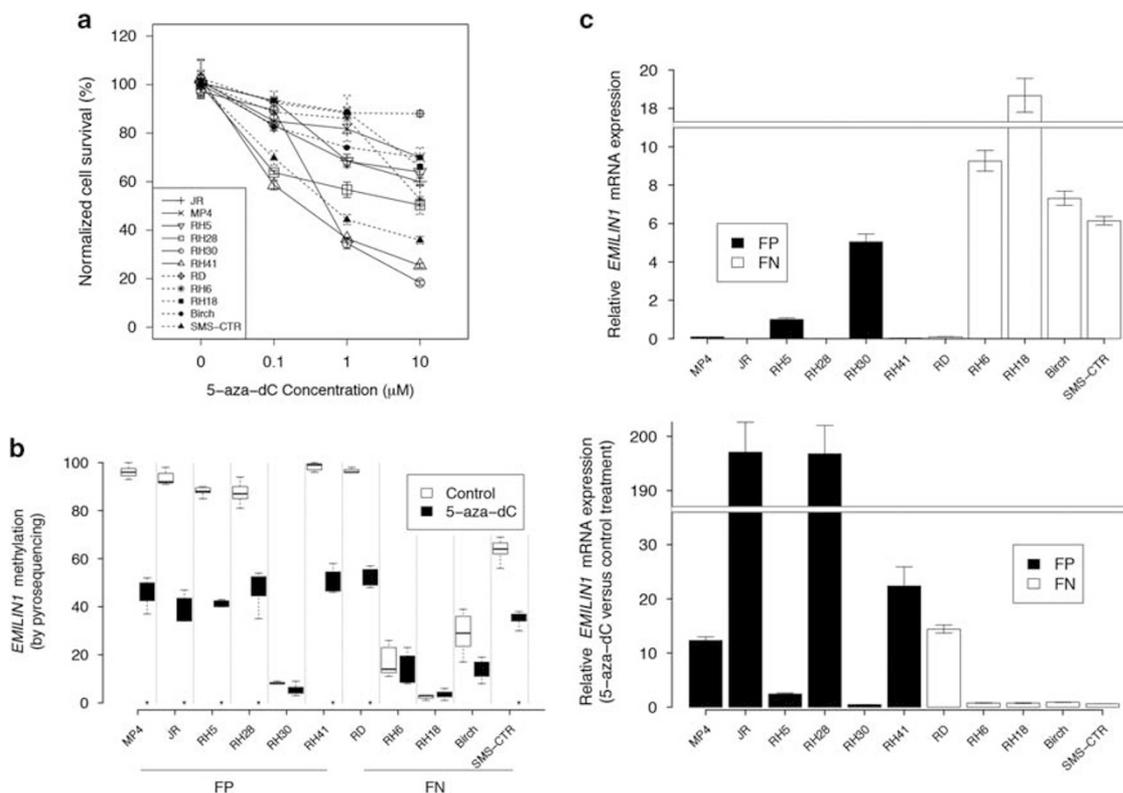
Figure 5 5-aza-2′-deoxycytidine treatment induced demethylation and increased expression of *EMILIN1* in fusion-positive rhabdomyosarcoma cells. (**a**) Cellular proliferation of six fusion-positive and five fusion-negative rhabdomyosarcoma cell lines in response to 5 -day treatment with varying concentrations of 5-aza-2′-deoxycytidine. At the indicated time points, triplicate determinations were quantified for each cell line using CellTiter 96 aqueous one solution cell proliferation assay according to the manufacturer's instructions. (**b**) *EMILIN1* promoter methylation at multiple CpG sites was measured by pyrosequencing in cells treated with 1 μM 5-aza-2′-deoxycytidine or vehicle alone for 5 days. Asterisks denote statistical difference at a *P*-value of <0.05. (**c**) *EMILIN1* mRNA expression levels in fusion-positive and fusion-negative rhabdomyosarcoma cell lines. Quantitative reverse transcriptase-PCR was utilized to measure mRNA expression of test and control genes in triplicate. Upper panel, baseline *EMILIN1* expression normalized to the level in RH5. Lower panel, changes in *EMILIN1* mRNA expression levels in rhabdomyosarcoma cell lines treated with 5-aza-2′-deoxycytidine as compared with the corresponding control DMSO-treated cells. FN, fusion-negative rhabdomyosarcoma; FP, fusion-positive rhabdomyosarcoma; 5-aza-dC, 5-aza-2′-deoxycytidine.

in nearly all cases and amplification events in $>50\%$ of cases,[6] may acquire and accumulate epigenetic changes that serve to collaborate with the fusion protein during tumorigenesis. The smaller number of recurrent secondary genetic events in fusion-positive compared with fusion-negative rhabdomyosarcoma tumors may explain the larger number of methylation differences relative to the normal tissues in fusion-positive tumors.

A final possibility for the molecular basis of the differential methylation events in fusion-positive and fusion-negative tumors is an intrinsic difference between the progenitor cells that give rise to these rhabdomyosarcoma tumors. Some of the 'aberrant' methylation events in rhabdomyosarcoma tumors may represent a methylation difference, characteristic of this specific normal cell type. The fundamental question of which cells in the body give rise to rhabdomyosarcoma is still debated. Mouse studies propose the Myf6-expressing fetal myoblast as the cell of origin for fusion-positive tumors.[28] In contrast, similar studies suggest that fusion-negative embryonal rhabdomyosarcoma can form in several myogenic lineages, both prenatal and postnatal, including muscle stem cells or downstream myogenic precursor lineages.[29] It is also possible that a subset of fusion-negative embryonal rhabdomyosarcoma cases may arise from non-myogenic cells, such as the adipogenic lineage.[30]

Our data support the premise that epigenetic regulation is an important mechanism in tumorigenesis for modulating mRNA expression, thereby contributing to the expression differences between fusion-positive and fusion-negative rhabdomyosarcoma. Of interest, we noted a sizable proportion of genes showing a direct correlation between DNA methylation and mRNA expression, which does not conform to the classical paradigm of an inverse correlation. The direct correlation between DNA methylation and gene expression may be due to our definition of promoter regions (2 kb upstream and 500 bp downstream of RefSeq transcription start

site), which includes portions of gene bodies and insulator regions, or might be due to complex *cis*- and *trans* regulation leading to activation of hyper-methylated genes and repression of hypomethylated genes.[31,32] Based on a recent report showing that DNA methylation in the gene body increases mRNA expression,[33] gene body DNA methylation is proposed to increase transcriptional activity by blocking initiation of intragenic promoters or by affecting activities of repetitive DNA elements within the transcriptional units.[34]

The development of an 11-gene methylation signature presents another approach to discriminate fusion-positive and fusion-negative rhabdomyosarcoma. Currently, reverse transcriptase-PCR and/or fluorescence *in situ* hybridization are used to detect the *PAX3/7–FOXO1* fusion genes as a part of the diagnosis and classification of rhabdomyosarcoma. However, fluorescence *in situ* hybridization is limited to samples that maintain nuclear integrity, and reverse transcriptase-PCR is limited to samples with intact RNA. Though immunohistochemical analysis can also be used as a surrogate marker of fusion positivity, this methodology requires intact tissue sections.[35] DNA is an ideal substrate for molecular diagnosis because it readily survives various adverse conditions found in clinical specimens. We propose that our methylation signature is a novel approach for ascertaining fusion status that may complement the methods mentioned above. It should also be noted that the 11-gene methylation signature was developed in rhabdomyosarcoma without any concurrent analysis of other malignant small round cell tumors, and thus the overall utility of the panel in differential diagnosis remains to be established.

Our findings highlight *EMILIN1* as an example of a gene in rhabdomyosarcoma whose expression is regulated by DNA methylation. *EMILIN1* encodes an extracellular matrix glycoprotein that has been proposed to suppress growth.[36,37] Using a chemically induced two-stage model of skin carcinogenesis, homozygous *EMILIN1* knockout mutations resulted in accelerated tumor formation and a higher tumor incidence.[38] Our data support that promoter DNA hypermethylation downregulates *EMILIN1* expression and thereby may reduce its growth suppressive function during fusion-positive rhabdomyosarcoma development.

In conclusion, this study provides the first systematic, genome-wide comparison of DNA methylation between fusion-positive and fusion-negative rhabdomyosarcoma. Our findings provide valuable insights into the epigenetic control of gene expression in fusion-positive and fusion-negative rhabdomyosarcoma tumors, present an additional approach for determining fusion status in rhabdomyosarcoma, and may identify new therapeutic targets in fusion-positive and fusion-negative rhabdomyosarcoma.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

1 Parham DM, Barr FG. Classification of rhabdomyosarcoma and its molecular basis. Adv Anat Pathol 2013;20: 387–397.

2 Davicioni E, Anderson MJ, Finckenstein FG et al. Molecular classification of rhabdomyosarcoma–genotypic and phenotypic determinants of diagnosis: a report from the Children's Oncology Group. Am J Pathol 2009;174:550–564.

3 Missiaglia E, Williamson D, Chisholm J et al. PAX3/FOXO1 fusion gene status is the key prognostic molecular marker in rhabdomyosarcoma and significantly improves current risk stratification. J Clin Oncol 2012;30:1670–1677.

4 Williamson D, Missiaglia E, de Reynies A et al. Fusion gene-negative alveolar rhabdomyosarcoma is clinically and molecularly indistinguishable from embryonal rhabdomyosarcoma. J Clin Oncol 2010;28:2151–2158.

5 Skapek SX, Anderson J, Barr FG et al. PAX-FOXO1 fusion status drives unfavorable outcome for children with rhabdomyosarcoma: a children's oncology group report. Pediatr Blood Cancer 2013;60:1411–1417.

6 Shern JF, Chen L, Chmielecki J et al. Comprehensive genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a common genetic axis in fusion-positive and fusion-negative tumors. Cancer Discov 2014;4:216–231.

7 Wachtel M, Dettling M, Koscielniak E et al. Gene expression signatures identify rhabdomyosarcoma subtypes and detect a novel t(2;2)(q35;p23) translocation fusing PAX3 to NCOA1. Cancer Res 2004;64: 5539–5545.

8 Harada K, Toyooka S, Maitra A et al. Aberrant promoter methylation and silencing of the RASSF1A gene in pediatric tumors and cell lines. Oncogene 2002;21: 4345–4349.

9 Rathi A, Virmani AK, Harada K et al. Aberrant methylation of the HIC1 promoter is a frequent event in specific pediatric neoplasms. Clin Cancer Res 2003;9:3674–3678.

10 Mahoney SE, Yao Z, Keyes CC et al. methylation studies suggest distinct DNA methylation patterns in pediatric embryonal and alveolar rhabdomyosarcomas. Epigenetics 2012;7:400–408.

11 Chen X, Stewart E, Shelat AA et al. Targeting oxidative stress in embryonal rhabdomyosarcoma. Cancer Cell 2013;24:710–724.

12 Barr FG, Smith LM, Lynch JC et al. Examination of gene fusion status in archival samples of alveolar rhabdomyosarcoma entered on the Intergroup Rhabdomyosarcoma

Methylation differences in rhabdomyosarcoma
1224
W Sun et al

Study-III trial: a report from the Children's Oncology Group. J Mol Diagn 2006;8:202–208.

13 Dumont SN, Lazar AJ, Bridge JA *et al.* PAX3/7-FOXO1 fusion status in older rhabdomyosarcoma patient population by fluorescent in situ hybridization. J Cancer Res Clin Onco 2012;138:213–220.

14 Triche TJ Jr., Weisenberger DJ, Van Den Berg D *et al.* Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res 2013;41:e90.

15 Du P, Zhang X, Huang CC *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics 2010;11:587.

16 Ellis RJ, Wang Y, Stevenson HS *et al.* Genome-wide methylation patterns in papillary thyroid cancer are distinct based on histological subtype and tumor genotype. J Clin Endocrinol Metab 2014;99:E329–E337.

17 Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1–22.

18 Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics 2010;26:2363–2367.

19 Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 2001;98:5116–5121.

20 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402–408.

21 Alisch RS, Barwick BG, Chopra P *et al.* Age-associated DNA methylation in pediatric populations. Genome Res 2012;22:623–632.

22 Hoque MO, Kim MS, Ostrow KL *et al.* Genome-wide promoter analysis uncovers portions of the cancer methylome. Cancer Res 2008;68:2661–2670.

23 Cao L, Yu Y, Bilke S *et al.* Genome-wide identification of PAX3-FKHR binding sites in rhabdomyosarcoma reveals candidate target genes important for development and cancer. Cancer Res 2010;70:6497–6508.

24 Hettmer S, Wagers AJ. Muscling in: Uncovering the origins of rhabdomyosarcoma. Nat Med 2010;16:171–173.

25 Carter RL, Jameson CF, Philp ER *et al.* Comparative phenotypes in rhabdomyosarcomas and developing skeletal muscle. Histopathology 1990;17:301–309.

26 Kaminskas E, Farrell AT, Wang YC *et al.* FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. Oncologist 2005;10:176–182.

27 Lae M, Ahn EH, Mercado GE *et al.* Global gene expression profiling of PAX-FKHR fusion-positive alveolar and PAX-FKHR fusion-negative embryonal rhabdomyosarcomas. J Pathol 2007;212:143–151.

28 Abraham J, Nunez-Alvarez Y, Hettmer S *et al.* Lineage of origin in rhabdomyosarcoma informs pharmacological response. Genes Dev 2014;28:1578–1591.

29 Rubin BP, Nishijo K, Chen HI *et al.* Evidence for an unanticipated relationship between undifferentiated pleomorphic sarcoma and embryonal rhabdomyosarcoma. Cancer Cell 2011;19:177–191.

30 Hatley ME, Tang W, Garcia MR *et al.* A mouse model of rhabdomyosarcoma originating from the adipocyte lineage. Cancer Cell 2012;22:536–546.

31 Aryee MJ, Liu W, Engelmann JC *et al.* DNA methylation alterations exhibit intraindividual stability and inter-individual heterogeneity in prostate cancer metastases. Sci Transl Med 2013;5169ra10.

32 Bert SA, Robinson MD, Strbenac D *et al.* Regional activation of the cancer genome by long-range epigenetic remodeling. Cancer Cell 2013;23:9–22.

33 Yang X, Han H, De Carvalho DD *et al.* Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell 2014;26:577–590.

34 Maunakea AK, Nagarajan RP, Bilenky M *et al.* Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 2010;466:253–257.

35 Rudzinski ER, Anderson JR, Lyden ER *et al.* Myogenin, AP2β, NOS-1, and HMGA2 are surrogate markers of fusion status in rhabdomyosarcoma: a report from the soft tissue sarcoma committee of the children's oncology group. Am J Surg Pathol 2014;38:654–659.

36 Verdone G, Doliana R, Corazza A *et al.* The solution structure of EMILIN1 globular C1q domain reveals a disordered insertion necessary for interaction with the alpha4beta1 integrin. J Biol Chem 2008;283:18947–18956.

37 Danussi C, Del Bel Belluz L, Pivetta E *et al.* EMILIN1/alpha9beta1 integrin interaction is crucial in lymphatic valve formation and maintenance. Mol Cell Biol 2013;33:4381–4394.

38 Danussi C, Petrucco A, Wassermann B *et al.* An EMILIN1-negative microenvironment promotes tumor cell proliferation and lymph node invasion. Cancer Prev Res 2012;5:1131–1143.