Identification of a PAX-FKHR Gene Expression Signature that Defines Molecular Classes and Determines the Prognosis of Alveolar Rhabdomyosarcomas

Elai Davicioni,[1,3] Friedrich Graf Finckenstein,[5] Violette Shahbazian,[5] Jonathan D. Buckley,[2,3,4] Timothy J. Triche,[1,3,5] and Michael J. Anderson[1,3,5]

Departments of [1]Pathology and [2]Preventive Medicine, [3]Keck School of Medicine, and [4]Norris Comprehensive Cancer Center, University of Southern California; [5]Department of Pathology and Laboratory Medicine, Saban Research Institute, Children's Hospital Los Angeles, Los Angeles, California

Abstract

Alveolar rhabdomyosarcomas (ARMS) are aggressive soft-tissue sarcomas affecting children and young adults. Most ARMS tumors express the *PAX3-FKHR* or *PAX7-FKHR* (*PAX-FKHR*) fusion genes resulting from the t(2;13) or t(1;13) chromosomal translocations, respectively. However, up to 25% of ARMS tumors are fusion negative, making it unclear whether ARMS represent a single disease or multiple clinical and biological entities with a common phenotype. To test to what extent PAX-FKHR determine class and behavior of ARMS, we used oligonucleotide microarray expression profiling on 139 primary rhabdomyosarcoma tumors and an *in vitro* model. We found that ARMS tumors expressing either *PAX-FKHR* gene share a common expression profile distinct from fusion-negative ARMS and from the other rhabdomyosarcoma variants. We also observed that PAX-FKHR expression above a minimum level is necessary for the detection of this expression profile. Using an ectopic PAX3-FKHR and PAX7-FKHR expression model, we identified an expression signature regulated by PAX-FKHR that is specific to *PAX-FKHR*-positive ARMS tumors. Data mining for functional annotations of signature genes suggested a role for PAX-FKHR in regulating ARMS proliferation and differentiation. Cox regression modeling identified a subset of genes within the PAX-FKHR expression signature that segregated ARMS patients into three risk groups with 5-year overall survival estimates of 7%, 48%, and 93%. These prognostic classes were independent of conventional clinical risk factors. Our results show that PAX-FKHR dictate a specific expression signature that helps define the molecular phenotype of PAX-FKHR-positive ARMS tumors and, because it is linked with disease outcome in ARMS patients, determine tumor behavior. (Cancer Res 2006; 66(14): 6936-46)

Introduction

Rhabdomyosarcomas are the most common soft-tissue sarcomas of childhood, consisting of a highly heterogeneous family of tumors that show varying degrees of skeletal muscle differentiation (1). Although the etiology and cell-of-origin of rhabdomyosarcoma remain speculative, as rhabdomyosarcoma tumors are found at diverse anatomic sites not obviously associated with skeletal muscle, rhabdomyosarcoma is thought to arise from primitive mesenchymal progenitors that have undergone a limited program of myogenic differentiation (2). Similar to other pediatric tumors, such as acute myeloid leukemia, this diverse group of tumors has several histologic and genetic subtypes that are clinically associated with diverse patient outcomes (3). Embryonal rhabdomyosarcoma (ERMS) and the morphologic spindle/botryoid variants are associated with intermediate and superior patient prognosis, respectively. In contrast, alveolar rhabdomyosarcoma (ARMS) is an aggressive variant with a high frequency of metastasis at the time of initial diagnosis; therefore, patients with these tumors have a worse prognosis (3–5).

Cytogenetic analysis of ARMS tumors has identified two distinct chromosomal translocations that are used for the differential diagnosis from ERMS as well as the other childhood solid tumors. The t(2;13) translocation, which generates a chimeric gene fusing *PAX3* on chromosome 2 to *FKHR* on chromosome 13, is found in ~55% of the tumors. A variant t(1;13) translocation leads to the fusion of *PAX7* on chromosome 1 to *FKHR* and is found in 22% of these tumors (6). Despite exhibiting the classic alveolar histology, however, ~25% of ARMS tumors lack either translocation. In fact, it seems that these fusion-negative ARMS tumors are more similar to ERMS tumors with respect to other covariates, such as age at diagnosis and patient survival (7). Furthermore, there are indications that *PAX3-FKHR* and *PAX7-FKHR* ARMS tumors represent high- and low-risk subgroups, respectively, among patients presenting with metastatic disease (6, 8). Lastly, there are a few rhabdomyosarcoma tumors displaying mixed alveolar/embryonal histology. The *PAX3-FKHR* or *PAX7-FKHR* (*PAX-FKHR*) chromosomal translocations are observed in only some of these mixed histology tumors (9). Presently, tumors with any evidence of alveolar histology are considered to be one pathologic entity, although it is clear that they are genetically and clinically heterogeneous.

The PAX-FKHR fusion proteins contain the NH_2-terminal region, including the intact paired domain and homeodomain DNA-binding elements of PAX3 or PAX7, and the potent COOH-terminal transactivation domain of FKHR (10). The forkhead DNA-binding domain of FKHR is truncated in the fusion protein and does not seem to influence target gene recognition (11). Transcriptional activity attributed to PAX-FKHR fusion proteins is therefore likely due to deregulation of both wild-type PAX and FKHR function (12, 13). Several studies have made efforts toward identifying PAX-FKHR target genes using heterologous cell lines, including NIH3T3 fibroblasts (14), RD ERMS (15), and SAOS-2 osteosarcoma (16).

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

E. Davicioni and F. Graf Finckenstein contributed equally to this work.

Requests for reprints: Michael J. Anderson, Department of Pathology and Laboratory Medicine, Saban Research Institute, Children's Hospital Los Angeles, 4650 Sunset Boulevard, MS 103, Los Angeles, CA 90027. Phone: 323-669-5624; Fax: 323-906-8081; E-mail: manderson@chla.usc.edu.

doi:10.1158/0008-5472.CAN-05-4578

These studies, and others, have shown numerous consequences of PAX-FKHR expression, including activation of a myogenic transcription program (14), mesenchymal-to-epithelial transition (16), and transformation and growth suppression (17). However, few targets have been identified and shown to have relevance to ARMS biology or clinical value in predicting patient outcome.

In the present study, we focused on the role of the *PAX-FKHR* fusion genes in the diagnosis, tumor behavior, and prognosis of ARMS. First, we addressed the question of the proper diagnostic classification of ARMS tumors. We found that this subgroup of tumors should be grouped by their genotype as determined by the mRNA expression of the *PAX-FKHR* genes rather than by their morphologic phenotype. Second, to identify the molecular targets of the fusion transcription factors, we characterized a PAX-FKHR expression signature detectable both in tumors and in an *in vitro* model of ERMS cells expressing PAX-FKHR. To describe the biological role of PAX-FKHR, we mined its expression signature for overrepresented functional annotations and biological networks, finding indications for a role in regulation of proliferation and differentiation through suppression of apoptosis and muscle cell differentiation. Finally, we showed that the expression patterns of a subset of the PAX-FKHR signature genes correlate with patient outcome. Thus, we not only identified novel prognostic markers that are independent of conventional criteria but also showed the importance of PAX-FKHR transcriptional control for the behavior of ARMS.

Materials and Methods

Primary tumors and cell lines. Frozen tumor samples from patients enrolled in the Intergroup Rhabdomyosarcoma Study Group (IRS-IV and IRS-V) Children's Oncology Group clinical trials were obtained from the Pediatric Cooperative Human Tissue Network (CHTN) tumor bank (Columbus, OH). Additional tumor samples were obtained from the Children's Hospital Los Angeles (CHLA) institutional tumor bank (Supplementary Table S1). Histopathologic diagnoses were based on the International Classification of Rhabdomyosarcoma criteria (18). All tumor samples contained >80% tumor cells. Rhabdomyosarcoma cell lines used for microarray analysis included four embryonal cell lines (TTC-442, TTC-516, Birch, and RD), one alveolar fusion-negative cell line (RH18), and four alveolar *PAX3-FKHR* fusion-positive cell lines (HR, JR, RH28, and RH30). All cell lines were cultured in RPMI 1640 supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA).

Retroviral transduction and fluorescence-activated cell sorting. Hemagglutinin (HA) epitope-tagged *PAX3-FKHR* and *PAX7-FKHR* (19) were subcloned into the retroviral expression vector murine stem cell virus (MSCV)-internal ribosome entry site (IRES)-green fluorescent protein (GFP; refs. 20, 21). Supernatants containing virus were made in 293T cells by cotransfecting expression vector with plasmids pHIT60 (22) or pCgp (23) encoding gag-pol and pHIT456 (24) encoding amphotropic env. RD cells were transduced with retrovirus, subjected to fluorescence-activated cell sorting (FACS) 48 to 72 hours later for GFP expression, and expanded for RNA and protein isolation.

RNA isolation, expression arrays, and quantitative reverse transcription-PCR. RNA was extracted from frozen tumor tissue using RNA STAT-60 (Tel-Test, Inc., Friendswood, TX) and purified using the RNeasy Protect kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. Total mRNA from the transduced polyclones was harvested using TRIzol (Invitrogen). Biotin-labeled cRNA was prepared from total RNA and hybridized to Affymetrix GeneChip Human U133A Expression Arrays according to the manufacturer's protocol (Affymetrix, Santa Clara, CA). To determine PAX-FKHR mRNA expression in primary tumors, aliquots of the microarray *in vitro* transcription reactions were subjected to quantitative PCR using the QuantiTect SYBR Green PCR kit (Qiagen) on a Smart Cycler (Cepheid, Sunnyvale, CA; see Supplementary Information and Supplementary Fig. S1; ref. 25).

Immunoblotting. Protein was isolated as described previously (26). Equal amounts of protein (40 μg) were fractionated by 7.5% to 12% SDS-PAGE, transferred to Immobilon-P membrane (Millipore, Billerica, MA), and reacted with antibodies against HA (HA.11; Covance, Princeton, NJ) and β-actin (AC-15; Sigma, St. Louis, MO). Immunodetection was done using horseradish peroxidase–conjugated secondary antibodies (Bio-Rad, Hercules, CA) and the enhanced chemiluminescence detection system (Amersham, Piscataway, NJ).

Transient reporter assays. Transfections were done using Lipofectamine and Plus reagent according to the manufacturer's protocol (Invitrogen). For the PRS9 reporter assays, transfections were done using 1 μg pcDNA3-lacZ and 1 μg pTK-PRS9 (19, 27). Cells were harvested 72 hours after transfection, and chloramphenicol acetyltransferase (CAT) assays done as described previously (19). CAT activity was normalized to β-galactosidase activity; experiments were done within the linear range of the assays and done thrice in duplicate.

Gene expression and survival analysis. All data management and analysis was conducted using the Genetrix suite of tools for microarray analysis (EPIcenter Software, Pasadena, CA).[6] The complete tumor and cell line microarray data set can be found on the National Cancer Institute Cancer Array Database.[7] Probe set modeling and data preprocessing were derived using the Robust Multiarray algorithm implemented within the ProbeProfiler module (28). ANOVA was used for gene selection in the initial analysis of primary tumors. A semisupervised routine called meta-clustering was implemented that determined cluster centroid identification under leave-n-out sample cross-validation with a fuzzy k-means algorithm. Gene selection for the *in vitro* PAX-FKHR model system was based on the fold change in the average expression (at least 1.5-fold change) and the associated t test, with a P cutoff of <0.001. Differentially expressed genes identified by analysis of *in vitro* expression profiles were validated for differential expression in primary tumors with reiterative cross-validation again using the meta-clustering algorithm, except that a t test was used for gene selection between fusion-positive and fusion-negative primary rhabdomyosarcoma tumors. Comparison of survival times was carried out using Kaplan-Meier survival plots and log-rank tests of significance. Prognostic metagenes were created by weighting selected genes by their Cox regression test statistics. Multivariate analysis of metagene predictor scores was done using Cox regression proportional hazards modeling to test for independence from previously characterized clinical risk factors (see Supplementary Information for complete analysis details).

Functional annotation of gene clusters. Functional annotation was done using the Expression Analysis Systematic Explorer (EASE)[8] software package for overrepresentation analysis of functional gene categories and for multidatabase annotation of the differentially expressed genes (29). Pathway analysis was done with the online Ingenuity Pathways analysis (IPA) tool (version 3.0; ref. 30).[9] Literature data mining was done with the online PubMatrix tool (31).[10]

Results

PAX-FKHR fusion genes dictate the expression profile of ARMS. The histologic variants of rhabdomyosarcoma are associated with distinct genetic abnormalities and diverse patient outcomes. We therefore hypothesized that each of these rhabdomyosarcoma variants would exhibit specific gene expression patterns and that the molecular expression profiles of these tumors underlie their histopathologic phenotype. To test this, we did microarray analysis on a large data set of primary pretreatment rhabdomyosarcoma

[6] http://www.epicentersoftware.com.
[7] http://caarraydb.nci.nih.gov.
[8] http://david.niaid.nih.gov/david/ease.htm.
[9] http://www.ingenuity.com.
[10] http://pubmatrix.grc.nia.nih.gov.

tumors (Supplementary Table S1). Using a semisupervised class discovery algorithm, called meta-clustering analysis, we generated gene expression profiles for these 139 rhabdomyosarcoma tumors. The *k*-means algorithm was employed to identify cluster centroids from samples with similar expression profiles and was initialized with *k* set to 4 using ANOVA for gene selection among the four rhabdomyosarcoma subtypes (alveolar, mixed alveolar/embryonal, embryonal, and spindle/botryoid). This analysis resulted in expression profiles derived from all 22,215 queried probe sets on the Affymetrix U133A GeneChip. After 1,000 rounds of meta-clustering, 534 genes represented by 650 probe sets were found to participate in centroid clustering at a false discovery rate of 0.1% (Supplementary Table S2). In contrast to our initial hypothesis, the cross-validated data set visualized with a multidimensional scaling plot showed only two, and not four, main tumor clusters (Fig. 1*A*). Most of the tumors with alveolar histology, including three of the tumors with mixed alveolar/embryonal histology, clustered together into a dense group away from the origin. Tumors with embryonal or spindle/botryoid histology clustered into a more heterogeneous spread. However, 15 tumors with alveolar or mixed alveolar/embryonal histology, representing 21% of all the tumors with alveolar histology, clustered with the non-ARMS variants.

Previous studies have shown that the *PAX-FKHR* genes are expressed in most rhabdomyosarcoma tumors with alveolar histology, although ~25% do not express either gene (6, 7). To determine whether PAX-FKHR expression differences explain why the 15 samples did not cluster among the other ARMS tumors, we measured PAX-FKHR mRNA expression by quantitative reverse transcription-PCR (QRT-PCR) in 59 of the 70 tumors with alveolar or mixed alveolar/embryonal histology (Supplementary Fig. S1B). Figure 1*B* shows the QRT-PCR data superimposed on the unsupervised multidimensional scaling plot. PAX-FKHR-expressing alveolar tumors (orange to pink dots) clustered together, whereas alveolar tumors in which PAX-FKHR expression was not detected (black dots) clustered instead with the other fusion-negative rhabdomyosarcoma tumors. Two tumors (blue dots) with very low but detectable PAX-FKHR expression, at levels below 0.1% of the median PAX-FKHR expression level for all tumors assayed, also clustered with other fusion-negative tumors. Five additional alveolar tumors, for which material was not available for QRT-PCR but were determined previously by conventional RT-PCR analysis to be fusion negative, similarly clustered with the other fusion-negative rhabdomyosarcoma tumors (determined at diagnosis to be fusion negative; see ref. 6). These data suggest that despite their common histology *PAX-FKHR* fusion-positive and fusion-negative ARMS tumors have distinct expression profiles that are dependent on the expression of the *PAX-FKHR* genes. Furthermore, it seems that a minimum expression level of PAX-FKHR is required in order for a tumor with alveolar histology to cluster among the PAX-FKHR-expressing ARMS tumors.

An *in vitro* model identifies an expression profile regulated by PAX-FKHR. If PAX-FKHR is in fact a major determinant of the expression profiles of fusion-positive ARMS tumors, we reasoned that ectopic expression of PAX-FKHR in an ERMS cell line would significantly alter its expression profile such that this non-ARMS cell line would now cluster with PAX-FKHR-expressing ARMS cell lines. To test this, we stably expressed PAX3-FKHR and PAX7-FKHR in the embryonal RD cell line using retroviral transduction. Retroviral constructs contained HA-tagged *PAX3-FKHR* or *PAX7-FKHR* cDNAs cloned upstream of IRES-GFP sequences, facilitating the rapid screening of transduced cells by FACS (Fig. 2*A*). Four independent polyclonal populations infected with PAX3-FKHR, PAX7-FKHR, or vector control virus were isolated. Expression of PAX-FKHR was detected by Western blot in all of the PAX-FKHR polyclones (Fig. 2*B* and *C*). Transactivation of the pTK-PRS9 reporter construct (27), which contains six tandem copies of a consensus PAX and PAX-FKHR binding sequence, in representative PAX3-FKHR and PAX7-FKHR transduced polyclones confirmed the functionality of the ectopic proteins (Fig. 2*D*). CAT activity in the PAX3-FKHR- and PAX7-FKHR-expressing cells was 4-fold higher than in vector transduced cells, which showed basal CAT activity due to the presence of endogenous PAX3 and PAX7 (32).

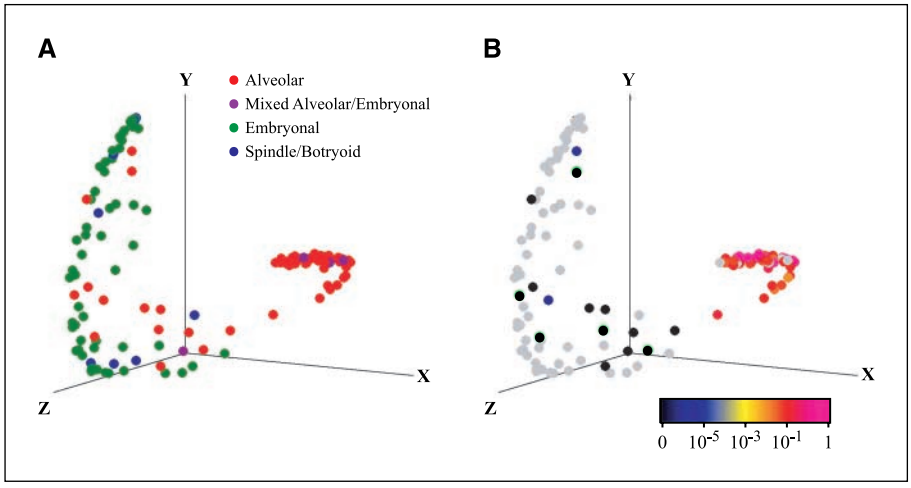

Figure 1. PAX-FKHR expression in ARMS is associated with a unique gene expression profile that is independent of tumor histology. *A,* multidimensional scaling analysis of 139 primary rhabdomyosarcoma tumors based on semisupervised analysis reveals tight clustering of most alveolar tumors, in contrast to the heterogeneous distribution of embryonal and spindle/botryoid tumors. Included in the main alveolar cluster are three mixed histology alveolar/embryonal tumors. The legend indicates the histologic diagnoses. *B,* replot of (*A*) based on normalized QRT-PCR expression levels of PAX3-FKHR and PAX7-FKHR for 59 tumors with alveolar or mixed alveolar/embryonal histology. Relative PAX-FKHR mRNA levels are depicted by colored dots as indicated in the color scale. Note that the two tumors with low but detectable PAX-FKHR expression (*blue dots*) had median PAX-FKHR expression levels 1,000-fold less than alveolar tumors expressing high levels of PAX-FKHR (*orange to pink dots*). Rhabdomyosarcoma samples for which no QRT-PCR data was available are also indicated (*gray dots*). Note that all non-ARMS tumors were shown at diagnosis to be fusion negative by conventional RT-PCR (data not shown).

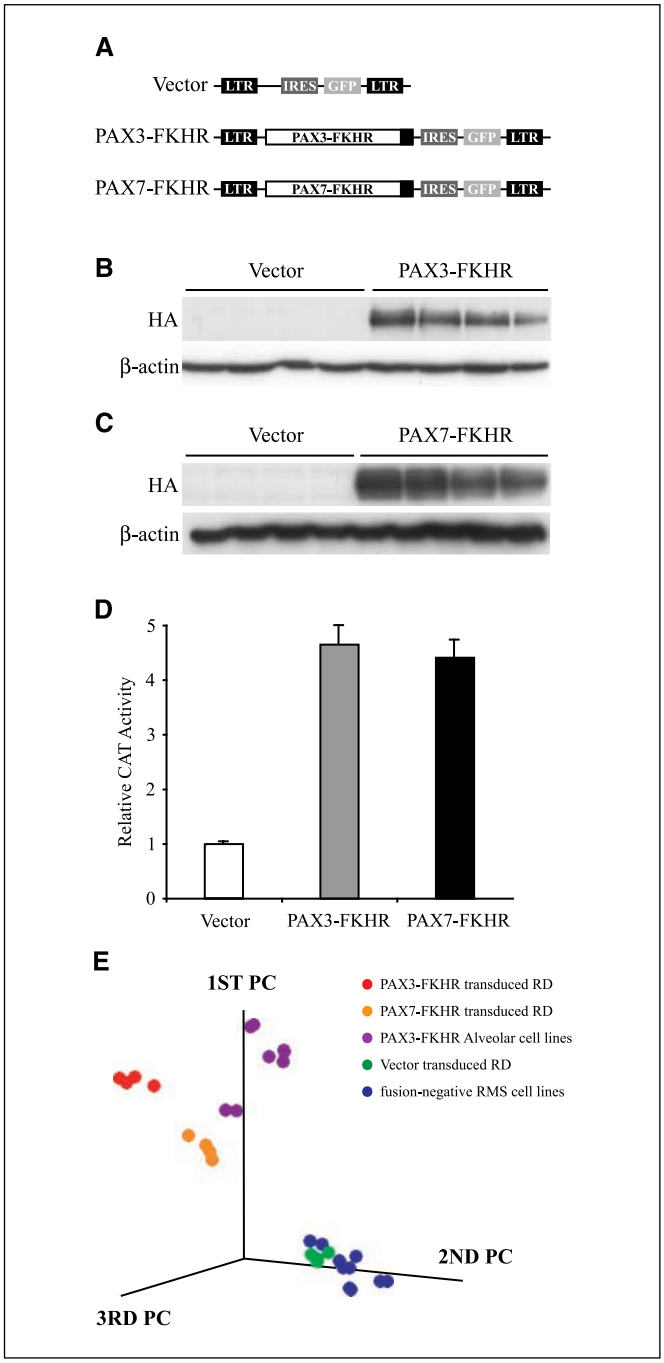

Figure 2. Identification of a PAX-FKHR expression profile using an *in vitro* model. *A,* diagram of the retroviral constructs used to express PAX-FKHR. Three tandem HA epitope tags are represented by the *black boxes* at the 3′ end of the *PAX-FKHR* cDNAs. *B* and *C,* Western blot analysis of FACS polyclones transduced with vector, PAX3-FKHR, or PAX7-FKHR retrovirus. PAX-FKHR expression was determined using an anti-HA antibody. β-Actin was used as a loading control. *D,* transactivation of the pTK-PRS9 reporter construct in representative vector, *PAX3-FKHR*, and *PAX7-FKHR* polyclones confirmed the functionality of the ectopic proteins. The activity in the vector polyclone was arbitrarily set to 1, with all other activities related to this value. *Columns,* mean of multiple experiments; *bars,* SD. *E,* principal components (*PC*) analysis of rhabdomyosarcoma cell lines and transduced RD polyclones using 85 genes differentially expressed between both *PAX-FKHR* and vector transduced RD polyclones and between fusion-positive and fusion-negative rhabdomyosarcoma cell lines. *PAX3-FKHR* and *PAX7-FKHR* transduced samples cocluster with cell lines derived from fusion-positive alveolar tumors, in contrast to vector transduced samples that cluster with the fusion-negative rhabdomyosarcoma cell lines.

We next analyzed the microarray expression profiles of these PAX3-FKHR, PAX7-FKHR, and vector transduced RD populations. A simple *t* test identified 334 genes (represented by 389 probe sets) differentially expressed between the PAX-FKHR populations and the vector populations (Supplementary Table S3). Approximately 80% of the 334 genes identified in this screen were defined as "up-regulated," showing increased expression in the PAX-FKHR-expressing polyclones relative to vector polyclones, and 68 genes were defined as "down-regulated." Using this *in vitro*–derived PAX-FKHR expression profile, we next determined how the ectopic expression of PAX-FKHR relates to expression patterns seen in rhabdomyosarcoma-derived cell lines. To do this, we analyzed a panel of established rhabdomyosarcoma cell lines derived from primary tumors, including four PAX3-FKHR-expressing alveolar and five fusion-negative cell lines. Within the 334 differentially expressed genes that comprised the *in vitro* PAX-FKHR expression profile, we identified a subset of 85 genes (represented by 106 probe sets) that were differentially expressed between the PAX-FKHR-expressing and fusion-negative rhabdomyosarcoma cell lines (Supplementary Table S4). Principle component analysis based on the expression of these 85 genes separated the cell lines into two clusters (Fig. 2*E*). PAX3-FKHR and PAX7-FKHR transduced RD clustered with the alveolar fusion-positive cell lines. As expected, vector transduced populations clustered with the other fusion-negative rhabdomyosarcoma cell lines. This subset of 85 genes represents ~12% of the total variation and therefore reflects only some of the inherent differences between PAX-FKHR-expressing and fusion-negative rhabdomyosarcoma cell lines (data not shown). However, this analysis highlights the fact that the expression profile of an embryonal cell line can be shifted toward that of a fusion-positive alveolar cell line by ectopic expression of PAX-FKHR, resulting from the activation of a PAX-FKHR transcriptional program. Together, our analyses of primary tumors and model cell lines suggest the existence of only two molecular classes of rhabdomyosarcoma. Reflecting these findings, hereafter, we refer to PAX-FKHR-expressing tumors as molecular ARMS (mARMS) and all fusion-negative rhabdomyosarcoma tumors, independent of tumor histology, as molecular ERMS (mERMS).

Identification of a PAX-FKHR expression signature in primary ARMS tumors derived from the *in vitro* expression profile. Based on the above, we sought to determine whether the *in vitro* PAX-FKHR expression profile is present in primary ARMS tumors as well. By screening this profile for differential expression between the newly defined mARMS and mERMS tumor classes, we sought to identify genes relevant to the pathobiology of fusion-positive ARMS tumors. As before, we applied the meta-clustering algorithm to identify genes that were differentially expressed: genes were selected using a *t* test comparison of the mARMS (n = 55) and mERMS (n = 84; including fusion-negative ARMS) cases in a series of 1,000 leave-n-out analyses. The P threshold was set to a value that provided an estimated false discovery rate of 0.1%. Genes selected in at least 50% of these randomized iterations were chosen resulting in a total of 109 genes (~33% of the *in vitro* expression profile). Of these 109 genes, 28 were expressed incongruently between the transduced polyclones and the primary tumors (i.e., up-regulated by PAX-FKHR expression in RD cells but expressed at decreased levels in *PAX-FKHR* primary tumors relative to fusion-negative rhabdomyosarcoma tumors or vice versa). These genes were removed from our analysis leaving a PAX-FKHR expression signature consisting of 61 up-regulated and 20 down-regulated genes (Table 1) represented by 102 probe sets. As a measure of the

Table 1. PAX-FKHR expression signature

Affymetrix ID	Symbol	Gene name	PAX-FKHR regulation*	Ref. †
209460_at	*ABAT*	*4-Aminobutyrate aminotransferase*	Up	
214895_s_at	***ADAM10***	***A disintegrin and metalloproteinase domain 10***	Up	
208212_s_at	*ALK*	*Anaplastic lymphoma kinase (Ki-1)*	Up	
202920_at	***ANK2***	***Ankyrin 2, neuronal***	Up	
202207_at	*ARL7*	*ADP-ribosylation factor-like 7*	Up	
207076_s_at	*ASS*	*Argininosuccinate synthetase*	Up	(14)
205444_at	*ATP2A1*	*ATPase, Ca^{2+} transporting, cardiac muscle, fast twitch 1*	Down	
205431_s_at	***BMP5***	***Bone morphogenetic protein 5***	Up	(57)
216598_s_at	*CCL2*	*Chemokine (C-C motif) ligand 2*	Down	
201005_at	***CD9***	***CD9 antigen (p24)***	Up	
212977_at	***CMKOR1***	***Chemokine orphan receptor 1***	Up	
209082_s_at	*COL18A1*	*Collagen, type XVIII, α 1*	Up	
204850_s_at	*DCX*	*Doublecortex; lissencephaly, X-linked (doublecortin)*	Up	
201581_at	***DJ971N18.2***	***Hypothetical protein DJ971N18.2***	Up	
222154_s_at	*DNAPTP6*	*DNA polymerase-transactivated protein 6*	Up	
204014_at	*DUSP4*	*Dual-specificity phosphatase 4*	Down	
211237_s_at	*FGFR4*	*Fibroblast growth factor receptor 4*	Up	
219147_s_at	*FLJ20559*	*Chromosome 9 open reading frame 95*	Up	
203689_s_at	*FMR1*	*Fragile X mental retardation 1*	Up	
203725_at	*GADD45A*	*Growth arrest and DNA damage inducible, α*	Up	
205848_at	*GAS2*	*Growth arrest specific 2*	Down	
207145_at	***GDF8***	***Growth differentiation factor 8***	Down	
209168_at	***GPM6B***	***Glycoprotein M6B***	Up	
202455_at	*HDAC5*	*Histone deacetylase 5*	Up	
205163_at	***HUMMLC2B***	***Myosin light chain 2***	Down	(14)
210095_s_at	*IGFBP3*	*Insulin-like growth factor–binding protein 3*	Down	
203233_at	***IL4R***	***Interleukin-4 receptor***	Up	
202794_at	*INPP1*	*Inositol polyphosphate-1-phosphatase*	Up	
209185_s_at	*IRS2*	*Insulin receptor substrate 2*	Up	
205902_at	*KCNN3*	*Potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3*	Up	
205968_at	*KCNS3*	*Potassium voltage-gated channel, delayed rectifier, subfamily S, member 3*	Up	
205888_s_at	*KIAA0555*	*Jak and microtubule-interacting protein 2*	Up	
205151_s_at	*KIAA0644*	*KIAA0644 gene product*	Down	
212956_at	*KIAA0882*	*KIAA0882 protein*	Up	
210102_at	*LOH11CR2A*	*Loss of heterozygosity, 11, chromosomal region 2, gene A*	Up	
214110_s_at	***LSP1***	***Similar to lymphocyte-specific protein 1***	Up	
208786_s_at	*MAP1LC3B*	*Microtubule-associated protein 1 light chain 3β*	Up	
213256_at	***MARCH3***	***Membrane-associated RING-CH protein III***	Up	
211042_x_at	***MCAM***	***Melanoma cell adhesion molecule***	Up	
210794_s_at	***MEG3***	***Maternally expressed 3***	Up	
203510_at	***MET***	***Met proto-oncogene (hepatocyte growth factor receptor)***	Up	(58, 59)
205330_at	*MN1*	*Meningioma (disrupted in balanced translocation) 1*	Up	
219038_at	*MORC4*	*MORC family CW-type zinc finger 4*	Up	
209708_at	***MOXD1***	***Monooxygenase, DBH-like 1***	Down	
209757_s_at	***MYCN***	***V-myc myelocytomatosis viral-related oncogene, neuroblastoma***	Up	
206657_s_at	***MYOD1***	***Myogenic factor 3***	Up	(60) ‡
203962_s_at	*NEBL*	*Nebulette*	Up	
205113_at	*NEF3*	*Neurofilament 3 (150-kDa medium)*	Down	
206089_at	***NELL1***	***NEL-like 1 (chicken)***	Up	
204105_s_at	*NRCAM*	*Neuronal cell adhesion molecule*	Up	
218162_at	*OLFM3*	*Olfactomedin 3*	Down	
221969_at	*PAX5*	*Paired box gene 5 (B-cell lineage-specific activator protein)*	Up	
219148_at	*PBK*	*PDZ-binding kinase*	Up	
217996_at	*PHLDA1*	*Pleckstrin homology-like domain, family A, member 1*	Down	
204612_at	*PKIA*	*Protein kinase (cyclic AMP–dependent, catalytic) inhibitor α*	Down	
201939_at	*PLK2*	*Polo-like kinase 2 (Drosophila)*	Down	
210830_s_at	*PON2*	*Paraoxonase 2*	Up	

(*Continued on the following page*)

Table 1. PAX-FKHR expression signature (Cont'd)

Affymetrix ID	Symbol	Gene name	PAX-FKHR regulation*	Ref. †
211341_at	*POU4F1*	*POU domain, class 4, transcription factor 1*	Up	
212680_x_at	***PPP1R14B***	***Protein phosphatase 1, regulatory (inhibitor) subunit 14B***	Up	
203680_at	*PRKAR2B*	*Protein kinase, cyclic AMP–dependent, regulatory, type II, β*	Up	(16)
213093_at	*PRKCA*	*Protein kinase C, α*	Up	
211373_s_at	*PSEN2*	*Presenilin 2 (Alzheimer's disease 4)*	Up	
202388_at	*RGS2*	*Regulator of G-protein signaling 2, 24 kDa*	Down	
206850_at	***RRP22***	***RAS-related on chromosome 22***	Up	
201739_at	*SGK*	*Serum/glucocorticoid-regulated kinase*	Up	(16)
203625_x_at	***SKP2***	***S-phase kinase-associated protein 2 (p45)***	Up	(61)
221489_s_at	***SPRY4***	***Sprouty homologue 4 (Drosophila)***	Down	
212353_at	***SULF1***	***Sulfatase 1***	Up	
212382_at	***TCF4***	***Transcription factor 4***	Down	
216511_s_at	*TCF7L2*	*Transcription factor 7-like 2 (T-cell-specific, HMG box)*	Up	
217853_at	*TENS1*	*Tensin-like SH2 domain-containing 1*	Up	
202039_at	*TIAF1*	*Transforming growth factor B1–induced antiapoptotic factor 1*	Up	
209656_s_at	***TM4SF10***	***Transmembrane 4 superfamily member 10***	Up	
205123_s_at	*TMEFF1*	*Transmembrane protein with epidermal growth factor–like and two follistatin-like domains*	Up	
202643_s_at	*TNFAIP3*	*Tumor necrosis factor-α-induced protein 3*	Up	
205388_at	***TNNC2***	***Troponin C2, fast***	Down	(14)
202369_s_at	***TRAM2***	***Translocation-associated membrane protein 2***	Up	
202478_at	*TRIB2*	*Tribbles homologue 2*	Down	
221861_at	*Unknown*	*Homo sapiens mRNA; cDNA DKFZp762M127*	Up	
201760_s_at	*WSB2*	*WD repeat and SOCS box-containing protein 2*	Up	
201368_at	*ZFP36L2*	*Zinc finger protein 36, C3H type-like 2*	Down	

NOTE: Genes in bold were used to create the prognostic 33 probe set metagene (Fig. 4).
*Gene expression up-regulated or down-regulated by ectopic PAX-FKHR expression.
†References refer to articles that have characterized the gene as a wild-type PAX or PAX-FKHR target.
‡In preparation.

significance of these PAX-FKHR signature genes to the expression profiles of mARMS tumors, we compared this 81-gene PAX-FKHR expression signature to the 81 top-ranked differentially expressed genes between mARMS and mERMS tumors (ranked by *t* test statistic) and found 16 (20%) genes in common between the two gene lists (Supplementary Table S2). Notably absent from the PAX-FKHR expression signature are some of the top-ranked discriminating genes, such as *TFAP2β*, *CDH3*, and *CNR1*. The PAX-FKHR expression signature, depicted in a two-way hierarchical clustering dendrogram and corresponding expression matrix, shows that all of the mARMS tumors cluster on a separate branch of the dendrogram from the fusion-negative mERMS tumors (Fig. 3*A* and *B*, respectively). As expected, fusion-negative ARMS tumors cluster with the ERMS tumors.

Gene Ontology and pathway analysis of the PAX-FKHR expression signature. To investigate the functional consequences of the PAX-FKHR expression signature, in terms of "biological themes," we did overrepresentation analysis using the EASE program (29). This type of analysis was aimed at identifying functional categories from the Gene Ontology (GO) and public gene annotation databases that are present in the PAX-FKHR expression signature more frequently than expected by chance alone ($P < 0.05$). GO categories overrepresented in both up-regulated and down-regulated components of the expression signature included "development," "morphogenesis," and "organogenesis" (Fig. 3*C*; Supplementary Table S5). However, the up-regulated component differed from the down-regulated component with overrepresentation of GO terms "cell adhesion," "cell proliferation," "programmed cell death," "cell differentiation," "apoptosis," and "neurogenesis." In contrast, the down-regulated component showed overrepresentation of GO categories, such as "muscle development" and "regulation of muscle contraction."

Although the EASE analysis identified functional categories that were regulated by PAX-FKHR, it was not informative with respect to how the individual genes interact within signaling pathways. To identify gene networks regulated by PAX-FKHR, we used the IPA tool. This proprietary database devises networks based on gene associations derived from peer-reviewed publications (30). We selected our PAX-FKHR expression signature to represent Focus Genes, which serve as nodes to derive biological networks that contain 35 genes. Of the 81 genes that comprise the PAX-FKHR expression signature, 49 mapped to four genetic networks that were generated according to statistical criteria using the IPA tool (Supplementary Tables S6-S11; Supplementary Fig. S2). Ten to 13 expression signature genes were used to build each of the four networks, representing approximately a third of the genes within a given network. Two of the networks confirmed our observations from the EASE analysis. Network 1 included 13 PAX-FKHR expression signature genes in a 35-gene network that featured regulators of apoptosis (Supplementary Fig. S2A). Functional annotation of this network revealed overrepresentation of GO categories, such as "cell proliferation" and "apoptosis"

(Supplementary Table S8). Network 4 featured MyoD as a central node and included 10 PAX-FKHR expression signature genes (Supplementary Fig. S2D). The GO categories of "muscle development" and "regulation of cell cycle" were overrepresented in this network (Supplementary Table S11). In summary, the EASE and IPA network data mining analyses indicate that the functional consequences of PAX-FKHR expression include repression of myogenic differentiation and maintenance of a proliferative state through regulation of apoptosis.

A subset of the PAX-FKHR expression signature predicts patient outcome. Establishing a correlation between the PAX-FKHR signature genes and patient survival would not only indicate their value for prognostic classification of ARMS but also show their relevance for tumor behavior. We observed by QRT-PCR analysis that although PAX-FKHR was expressed in all mARMS tumors their expression levels varied over an 80-fold range (Supplementary Fig. S1B). Tumors were also heterogeneous with respect to patient survival (range, 0.9-11.47 years). However, PAX-FKHR expression itself failed to correlate significantly to patient survival in a univariate model (data not shown). Therefore, we hypothesized that downstream target genes (i.e., the PAX-FKHR expression signature) would be better suited to analyze patient survival in multivariate models. To test this, we used Cox regression modeling to identify genes from our PAX-FKHR expression signature whose expression was correlated with patient outcome in 50 mARMS tumor patients (those with available survival data). A Cox regression proportional hazards model was applied to the PAX-FKHR expression signature using leave-*n*-out sample cross-validation. For each iteration, a randomly generated "training" subset of 25 patients was used to identify the best single-gene predictor of outcome, which was subsequently evaluated on the remaining "test" subset. This process was repeated to generate 2,500 cross-validated univariate models. A total of 57 probe sets (56% of the 102 PAX-FKHR expression signature probe sets) were used in at least one of the cross-validated models. The best single-gene predictor, TNNC2 (a muscle-specific troponin isoform), was used in about a third of all models. Five-year overall survival estimates from Kaplan-Meier analysis of the groups generated from splitting the samples about the median TNNC2 expression level were 30% and 70% (log-rank test, $P < 0.02$; Fig. 4*A* and *B*). When patients with evidence of metastatic disease at presentation were omitted from the analysis, TNNC2 expression levels were still informative with respect to outcome (Fig. 4*C*). However, we found that combining PAX-FKHR signature genes into multigene data vectors or "metagenes" greatly improved the statistical power of the resulting multivariate model compared with any single-gene

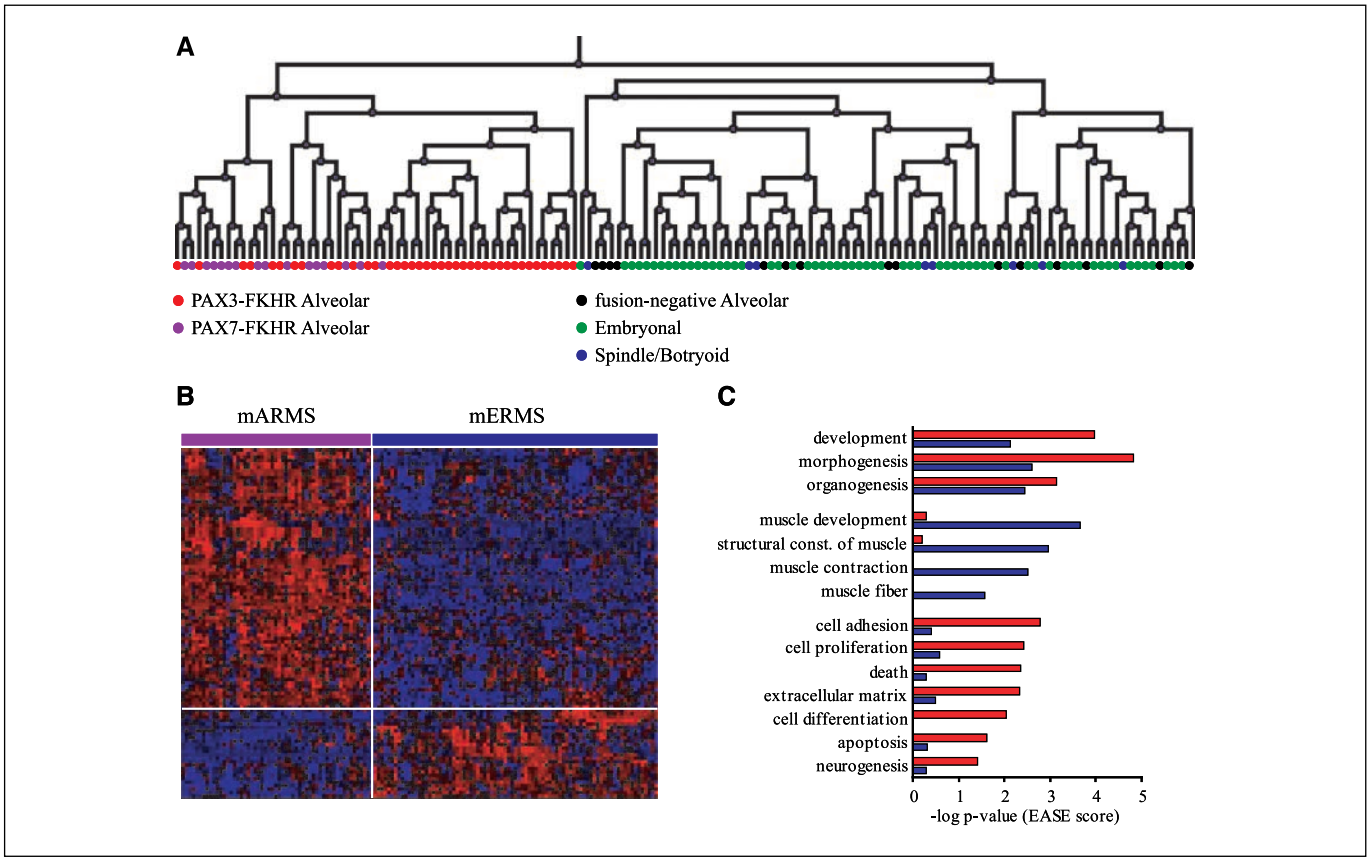

Figure 3. Microarray analysis and functional annotation of a PAX-FKHR expression signature active in primary rhabdomyosarcoma tumors. *A* and *B,* dendrogram (*A*) and expression matrix (*B*) derived from two-way hierarchical clustering analysis depicts sample clustering of 139 primary rhabdomyosarcoma tumors by 81 genes determined to be differentially expressed both between *PAX-FKHR* and vector transduced RD polyclones and between PAX-FKHR ARMS tumors and fusion-negative rhabdomyosarcoma tumors. The legend indicates the fusion gene status and histology for each tumor. This subset of genes accurately segregates all PAX-FKHR ARMS on the *left* of the dendrogram (also *purple bar* in *B*) and all other fusion-negative ARMS and ERMS on the *right* (also *blue bar* in *B*). *B,* the expression of each gene in each sample was normalized in the pseudocolored expression matrix based on the number of SDs above (*red*) and below (*blue*) the median expression value (*black*) across all samples. *C,* overrepresentation analysis of GO annotations shows functional categories enriched in the up-regulated and down-regulated components of the PAX-FKHR expression signature (*red* and *blue columns*, respectively) plotted by the negative log of the EASE score *P*.

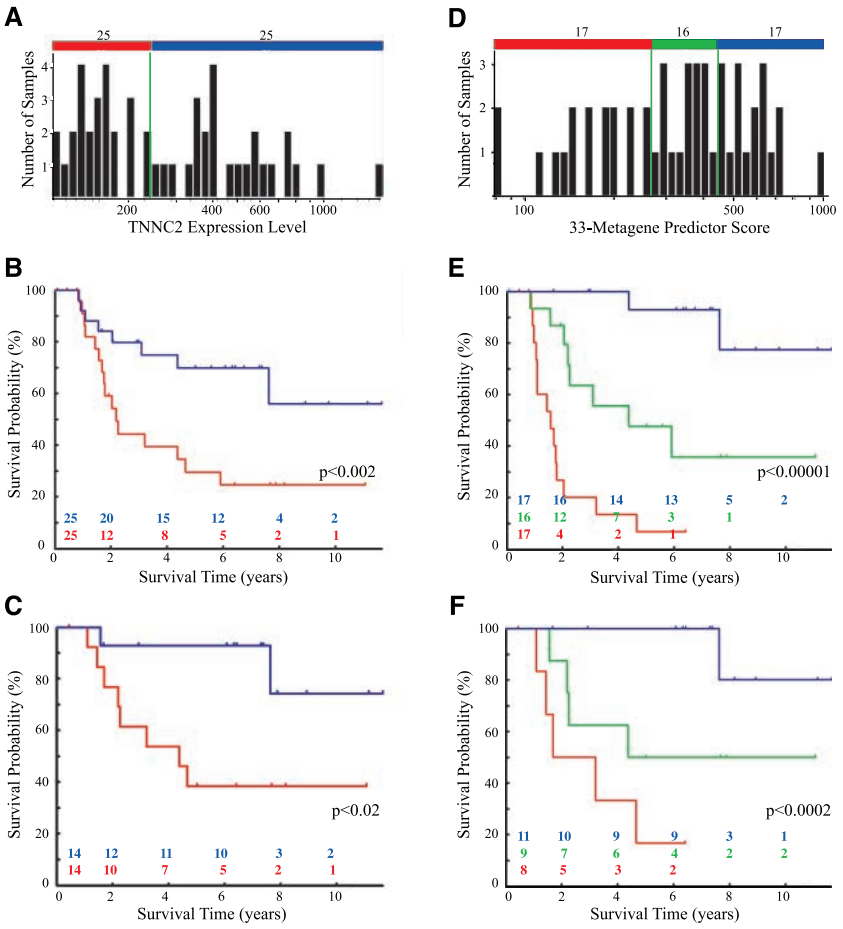

Figure 4. PAX-FKHR expression signature determines clinical outcome. *A* and *D,* binned distribution of the best single-gene predictor (*A*) or 33-metagene predictor scores (*D*) for 50 patients with mARMS tumors. *Vertical green lines,* median *TNNC2* expression level (*A*) or tertile cut points for the 33-metagene scores (*D*). *B, C, E*, and *F,* Kaplan-Meier survival analysis for all mARMS patients (*B* and *E*) and nonmetastatic mARMS patients only (*C* and *F*) grouped by TNNC2 expression level median cut point (*B* and *C*) or 33-metagene predictor tertiles (*E* and *F*). *Numbers below the curves,* number of patients at risk in each group about the median (*B* and *C*; *red curve,* below median; *blue curve,* above median) or tertile (*E* and *F*; *red curve,* first; *green curve,* second; *blue curve,* third). *P*s are from statistical analysis of Kaplan-Meier survival curves by log-rank test.

univariate model. Cox regression χ^2 test statistics were determined for each multivariate model and showed that a 33-metagene model (representing 25 genes; Table 1) had the highest χ^2 test statistic. Metagenes generated after sample permutations show that these results were not the result of chance alone (Supplementary Fig. S3). Further addition of PAX-FKHR expression signature genes to the model decreased the performance of the metagene due to the increased noise associated with genes that were not significantly correlated to outcome (see Supplementary Information).

To evaluate the predictive power of the 33-metagene model, mARMS patients were split into three groups by their metagene scores (Fig. 4*D*). Kaplan-Meier analysis revealed that patients in the first tertile (n = 17, red curve) had a 5-year overall survival estimate of 7% (Fig. 4*E*). In contrast, patients in the second (n = 16; green curve) and third (n = 17; blue curve) tertiles had a 5-year overall survival estimate of 48% and 93%, respectively (log-rank test, $P < 0.00001$). The highest risk group was comprised entirely of patients with *PAX3-FKHR* tumors (Fisher exact, $P < 0.001$) and 6 of 14 patients with metastatic disease at presentation, although this was not statistically significant (Fisher exact, $P < 0.48$). The lowest risk group contained nearly the same number of patients with *PAX3-FKHR* (n = 8) and *PAX7-FKHR* (n = 9) tumors, but overall this group accounted for 56% of all *PAX7-FKHR* tumors (Fisher exact, $P < 0.02$). Similarly, event-free survival (EFS) at 5 years showed significant differences among the first, second, and third tertiles with EFS of 9%, 37%, and 86%, respectively (log-rank test, $P < 0.0001$; data not shown). Again, the metagene was still predictive after omitting patients with evidence of metastatic disease at diagnosis from the analysis (Fig. 4*F*). Patients with nonmetastatic disease in the first, second, and third tertiles had an overall survival estimate at 5 years of 17%, 50%, and 100% (log-rank test, $P < 0.001$).

Adjusting for previously characterized prognostic factors, such as *PAX-FKHR* translocation variant, tumor size, patient age, anatomic site, and staging in a multivariate analysis, did not affect the performance of the metagene significantly (Table 2). Therefore, the prognostic metagene identifies differences in patient outcome independent of other known prognostic variables. These results add to our understanding of ARMS tumor biology, because they link malignant tumor behavior (i.e., growth characteristics of ARMS

Table 2. Metagene survival predictor is independent of clinical risk factors

Covariate	Variable	5-y Overall survival estimates (%)	*P**	
			Univariate	Adjusted 33-metagene
33-Metagene score	1st, 2nd, 3rd tertiles	7, 48, 93	<0.0001	—
TNNC2 †	Low vs high expression	30 vs 70	0.02	0.0002
Translocation	PAX3-FKHR vs PAX7-FKHR	34 vs 92	0.0007	<0.0001
Metastasis	M_0 vs M_1	63 vs 12	0.006	<0.0001
IRS risk group ‡	Intermediate vs high-risk group	65 vs 14	0.004	<0.0001
Stage	I, II, III, IV	0, 100, 52, 14	0.001	0.0002
Local invasiveness	T_1 vs T_2	72 vs 30	0.02	0.0005
Nodal involvement	N_0 vs N_1	69 vs 29	0.03	<0.0001
Tumor size (cm)	≤5 vs >5	67 vs 38	0.04	0.0002
Anatomic site§	Favorable vs unfavorable	33 vs 53	0.34	<0.0001
Age (y)	≤10 vs >10	57 vs 40	0.72	0.8

*Univariate: Cox regression *P*s for individual clinical risk factors. Adjusted *P*: multivariate Cox regression *P*s for the 33-metagene adjusted for each of the clinical risk factors or covariates.

†Best single-gene predictor of outcome as determined by cross-validated Cox proportional hazards modeling. Low versus high expression groups were split by the median expression level.

‡IRS risk groups as determined by IRS-V study guide published by the Children's Oncology Group.

§Favorable = orbit/eye lid, head and neck (excluding parameningeal), genitourinary (not bladder/prostate). Unfavorable = bladder, prostate, extremity, parameningeal, other (trunk, retroperitoneal, etc.).

cells) to a group of genes suggested to be regulated by PAX-FKHR in our expression array analyses. Therefore, the prognostic value of the PAX-FKHR signature ultimately proves its functional relevance for ARMS tumors.

Discussion

In this study, we used oligonucleotide microarray analysis to identify classes of ARMS tumors based on their expression profiles and to discover transcriptional targets of the *PAX-FKHR* fusion genes. We believe our results will be useful in elucidating the genetic components of this disease. Our studies show that fusion-positive ARMS are defined by an expression profile that distinguishes them from all fusion-negative rhabdomyosarcoma rather than by their morphologic appearance. A significant portion of this expression profile is dictated by PAX-FKHR as established by our *in vitro* model. We also show that a PAX-FKHR expression signature can be used to predict tumor behavior as assessed by clinical outcome, confirming the previously proposed correlation between PAX-FKHR expression and increased aggressiveness of ARMS (6, 7, 33, 34).

Using microarray analysis, Wachtel et al. reported that rhabdomyosarcoma tumors can be divided into three molecular classes: embryonal, alveolar fusion-negative, and alveolar fusion-positive (35). Their study used a small sample set (n = 29), and no cross-validation was done. In contrast, our findings are based on a larger data set (n = 139), facilitating a high degree of reproducibility owing to the utilization of sample cross-validation techniques. The data from our microarray clustering and QRT-PCR analyses show that histologically isomorphic ARMS tumors form two molecular classes of tumors, those that express PAX-FKHR and those that do not. Furthermore, in contrast to Wachtel et al., we find that fusion-negative alveolar tumors share a common expression profile with the other fusion-negative rhabdomyosarcoma variants. These findings are further supported by whole genome loss-of-heterozygosity and immunohistochemical analyses.[11] Collectively, our data support the reevaluation of the current histology-based classification scheme to include findings from genome-wide expression analyses of rhabdomyosarcoma tumors.

Although previous genome-wide studies aimed at identifying PAX-FKHR target genes have been reported (14–16), the relevance of these genes to the expression profiles of primary ARMS tumors was not confirmed. In our study, the differentially expressed PAX-FKHR target genes identified *in vitro* were screened against those genes differentially expressed between fusion-positive ARMS and fusion-negative rhabdomyosarcoma tumors. Our approach is analogous to the one taken by Hu-Lieskovan et al., wherein they identified EWS-FLI1 target genes using an *in vitro* model and primary Ewing's tumor samples (36). Taking this approach, we identified a PAX-FKHR expression signature with a significant overlap with the top-ranked discriminators of molecular class (i.e., between mARMS and mERMS tumors). This overlap suggests that a significant portion of the variation in gene expression found between the two molecular classes of rhabdomyosarcoma can be directly attributed to a PAX-FKHR transcriptional response. However, many of the genes differentially expressed between fusion-positive and fusion-negative rhabdomyosarcoma tumors, but absent from the PAX-FKHR signature, can conceivably be attributed to other sources of genetic variation at the transcriptional level, such as the genetic backgrounds of the yet unidentified rhabdomyosarcoma progenitor cells (32, 37, 38).

Overrepresentation analysis provided us with a statistical means to comprehend the biological themes associated with our PAX-FKHR expression signature. Genes down-regulated by PAX-FKHR were overrepresented by "muscle development" and "muscle contraction." This is in contrast to Khan et al., who found induction of

[11] E. Davicioni et al., in preparation.

a myogenic transcription program in NIH3T3 transduced with PAX3-FKHR (14). This myogenic transcription program included genes, such as *ACTC, MYL1, MYOG, SNAI2*, and *TNNC2*. We found all of these genes to be down-regulated in our model system. One gene, *PRRX1*, found by Khan et al. to be repressed was actually up-regulated in our model system. This suggests that in the myogenic RD background, in contrast to the nonmyogenic NIH3T3 background, PAX-FKHR expression represses myogenic differentiation as has been shown previously for PAX3 and PAX3-FKHR in C2C12 myoblasts (39).

Despite the down-regulation of a large number of myogenesis-related genes, MyoD was expressed 2-fold higher in PAX-FKHR transduced RD compared with vector control. This confirms our recent finding that *MyoD* is a PAX-FKHR target gene.[12] This apparent discrepancy can be explained by findings from a recent report that showed that MyoD regulates distinct subsets of genes in proliferating versus differentiating myoblasts (40). The authors also used EASE analysis and found that in proliferating myoblasts the GO categories "synaptic transmission" and "transmission of nerve impulse" were overrepresented within MyoD target genes. In contrast, in terminally differentiating myotubes, the GO categories "muscle development," "muscle contraction," and "regulation of striated muscle contraction" were overrepresented within MyoD target genes. These results are remarkably similar to our findings. We speculate that following MyoD induction by PAX-FKHR only a subset of MyoD target genes is activated, similar to those found in proliferating myoblasts where "neural phenotype" GO categories were overrepresented in the MyoD expression signature. PAX-FKHR, possibly through activation of MyoD, may promote a limited degree of myogenic determination and at the same time actively repress myogenic differentiation (as assessed by the expression of markers of muscle terminal differentiation, such as *MYL2* and *TNNC2*; refs. 41, 42). *MyoD* was in fact a central node in one of the IPA networks that included the cell cycle G_1-S regulator *RB1* as well as *MYOG, HDAC5, TWIST1, ID2, GDF8*, and *IGF1*. These genes are all involved in regulating the transition between proliferating and differentiating myoblasts (2, 40, 43). HDAC5, for example, is a crucial regulator of myogenic differentiation, acting as a transcriptional corepressor for MEF2C at MyoD target gene promoters (44). Other previously described MyoD downstream target genes, such as *ASS, PKIA*, and *TNNC2* (41), were also part of this network and of the PAX-FKHR expression signature.

Overrepresentation analysis identified the GO categories "programmed cell death," "apoptosis," and "cell proliferation." Confirming this, pathway analysis yielded a network that integrated genes involved in suppressing apoptosis [e.g., up-regulation of *MYCN* (45) and down-regulation of *IGFBP3* (46)] and promoting cell survival [e.g., down-regulation of *DUSP4* (47) and *SPRY4* (48) and up-regulation of *FGFR4* (49)]. In addition, some of these genes, such as *FGFR4* and *PRKCA*, seem to play dual roles in myogenic cells, serving to both repress differentiation and promote cell survival (50, 51). Repression of myogenic differentiation and activation of cell survival pathways, therefore, seem to be hallmarks of the PAX-FKHR-mediated transcriptional program.

Previous work with a similar *in vitro* model system showed that ectopic PAX3-FKHR expression in ERMS cell lines (including RD) leads to increased tumorigenicity and aggressiveness in mouse tumor models (33). We could not, however, correlate PAX-FKHR mRNA expression to mARMS patient survival. Instead, we found that a third of the PAX-FKHR expression signature (i.e., the 33-metagene) could identify outcome differences in mARMS patients. This shows that a PAX-FKHR transcriptional program plays an important role in determining not only the molecular phenotype but also the clinical behavior of ARMS tumors. In addition, these results lend further support to the validity of our approach to identify PAX-FKHR target genes that are biologically relevant to primary ARMS tumors.

The clinical relevance of making accurate diagnoses and predicting outcome in rhabdomyosarcoma is not trivial, as some of the progress made in rhabdomyosarcoma treatment over the last 30 years is overshadowed by misdiagnosis that often leads to suboptimal treatment of patients (52, 53). The diversity of outcomes for rhabdomyosarcoma patients has led to an intense search for prognostic indicators useful for therapeutic stratification. Currently, the most useful prognostic variables are the extent of disease at presentation, age at diagnosis, anatomic site, and histology (54). Studies involving other malignancies have successfully used expression profiling, including techniques, such as Cox regression modeling, for patient risk stratification based on the expression of a few genes (55, 56). The 33-metagene developed in this study clearly discriminates ARMS patients into risk groups, the lowest-risk group with a 5-year overall survival estimate of 93% and the highest-risk group with a 5-year overall survival estimate of only 7%. In addition, multivariate analysis showed that the 33-metagene is predictive of patient outcome independent of known clinical risk factors. When we excluded metastatic disease patients, we still found significant differences in outcome among the nonmetastatic patients, suggesting that the PAX-FKHR transcriptional program is a major determinant of outcome independent of the clinical manifestation of metastases. Our analysis also indicates that the expression pattern of just a single gene, such as the skeletal muscle–specific troponin isoform, *TNNC2*, can provide prognostic information, which could be amenable to a routine clinical assay (i.e., QRT-PCR). In conclusion, it seems that the PAX-FKHR expression signature genes can be used for patient risk assessment, yet its clinical usefulness has ultimately to be proven by prospective evaluation. Finally, new targeted therapies are needed to treat those tumors that do not respond to traditional modalities, especially within the subgroup of highly metastatic ARMS tumors. The identification of a defined group of genes regulated in a tumor type–specific manner, as described by our study, may provide promising candidates for such gene-specific therapies.

Acknowledgments

Received 12/27/2005; revised 3/8/2006; accepted 5/9/2006.

Grant support: NIH Director's Challenge U01 (T.J. Triche and J.D. Buckley), Saban Research Institute Faculty Research Career Development Award, and Children's Oncology Group Young Investigator Grant (M.J. Anderson).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Suzanne Baker and Derek Persons for the MSCV-IRES-GFP construct; Paula Cannon for the viral packaging constructs; Martyn Goulding for the PRS9 construct; Jerry Barnhart for FACS; and Dr. Deborah Schofield, Betty Schaub, Sitara Waidyaratne, Xuan Chen (University of Southern California/CHLA Microarray Core), Jackie Smith, Julie Bridge, Thomas Barr (CHTN Biopathology Center), Xian Fang Liu (CHLA Pathology), and Ingenuity Systems for access to the β test of the Pathway Analysis Tool version 3.0. F. Graf Finckenstein thanks the Foerdergemeinschaft Kinder-Krebs-Zentrum Hamburg e.V. (Germany) for previous support.

[12] F. Graf Finckenstein et al., in preparation.

References

1. Pappo AS, Shapiro DN, Crist WM, Maurer HM. Biology and therapy of pediatric rhabdomyosarcoma. J Clin Oncol 1995;13:2123–39.
2. Merlino G, Helman LJ. Rhabdomyosarcoma—working out the pathways. Oncogene 1999;18:5340–8.
3. Qualman SJ, Morotti RA. Risk assignment in pediatric soft-tissue sarcomas: an evolving molecular classification. Curr Oncol Rep 2002;4:123–30.
4. Breneman JC, Lyden E, Pappo AS, et al. Prognostic factors and clinical outcomes in children and adolescents with metastatic rhabdomyosarcoma—a report from the Intergroup Rhabdomyosarcoma Study IV. J Clin Oncol 2003;21:78–84.
5. Tobar A, Avigad S, Zoldan M, Mor C, Goshen Y, Zaizov R. Clinical relevance of molecular diagnosis in childhood rhabdomyosarcoma. Diagn Mol Pathol 2000;9:9–13.
6. Sorensen PH, Lynch JC, Qualman SJ, et al. PAX3-FKHR and PAX7-FKHR gene fusions are prognostic indicators in alveolar rhabdomyosarcoma: a report from the Children's Oncology Group. J Clin Oncol 2002;20:2672–9.
7. Barr FG, Qualman SJ, Macris MH, et al. Genetic heterogeneity in the alveolar rhabdomyosarcoma subset without typical gene fusions. Cancer Res 2002;62:4704–10.
8. Anderson J, Gordon T, McManus A, et al. Detection of the PAX3-FKHR fusion gene in paediatric rhabdomyosarcoma: a reproducible predictor of outcome? Br J Cancer 2001;85:831–5.
9. Kilpatrick SE, Teot LA, Geisinger KR, et al. Relationship of DNA ploidy to histology and prognosis in rhabdomyosarcoma. Comparison of flow cytometry and image analysis. Cancer 1994;74:3227–33.
10. Galili N, Davis RJ, Fredericks WJ, et al. Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma. Nat Genet 1993;5:230–5.
11. Sublett JE, Jeon IS, Shapiro DN. The alveolar rhabdomyosarcoma PAX3/FKHR fusion protein is a transcriptional activator. Oncogene 1995;11:545–52.
12. del Peso L, Gonzalez VM, Hernandez R, Barr FG, Nunez G. Regulation of the forkhead transcription factor FKHR, but not the PAX3-FKHR fusion protein, by the serine/threonine kinase Akt. Oncogene 1999;18:7328–33.
13. Bennicelli JL, Fredericks WJ, Wilson RB, Rauscher FJ III, Barr FG. Wild type PAX3 protein and the PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma contain potent, structurally distinct transcriptional activation domains. Oncogene 1995;11:119–30.
14. Khan J, Bittner ML, Saal LH, et al. cDNA microarrays detect activation of a myogenic transcription program by the PAX3-FKHR fusion oncogene. Proc Natl Acad Sci U S A 1999;96:13264–9.
15. Barber TD, Barber MC, Tomescu O, Barr FG, Ruben S, Friedman TB. Identification of target genes regulated by PAX3 and PAX3-FKHR in embryogenesis and alveolar rhabdomyosarcoma. Genomics 2002;79:278–84.
16. Begum S, Emani N, Cheung A, Wilkins O, Der S, Hamel PA. Cell-type-specific regulation of distinct sets of gene targets by Pax3 and Pax3/FKHR. Oncogene 2005; 24:1860–72.
17. Xia SJ, Barr FG. Analysis of the transforming and growth suppressive activities of the PAX3-FKHR oncoprotein. Oncogene 2004;23:6864–71.
18. Newton WA, Jr., Gehan EA, Webber BL, et al. Classification of rhabdomyosarcomas and related sarcomas. Pathologic aspects and proposal for a new classification-an Intergroup Rhabdomyosarcoma Study. Cancer 1995;76:1073–85.
19. Anderson MJ, Shelton GD, Cavenee WK, Arden KC. Embryonic expression of the tumor-associated PAX3-FKHR fusion protein interferes with the developmental functions of Pax3. Proc Natl Acad Sci U S A 2001;98: 1589–94.
20. Persons DA, Allay JA, Allay ER, et al. Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. Blood 1999;93:488–99.
21. Torchia EC, Jaishankar S, Baker SJ. Ewing tumor fusion proteins block the differentiation of pluripotent marrow stromal cells. Cancer Res 2003;63:3464–8.
22. Soneoka Y, Cannon PM, Ramsdale EE, et al. A transient three-plasmid expression system for the production of high titer retroviral vectors. Nucleic Acids Res 1995;23:628–33.
23. Han JY, Cannon PM, Lai KM, Zhao Y, Eiden MV, Anderson WF. Identification of envelope protein residues required for the expanded host range of 10A1 murine leukemia virus. J Virol 1997;71:8103–8.
24. Page KA, Landau NR, Littman DR. Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity. J Virol 1990;64:5270–6.
25. Peter M, Gilbert E, Delattre O. A multiplex real-time PCR assay for the detection of gene fusions observed in solid tumors. Lab Invest 2001;81:905–12.
26. Plattner R, Anderson MJ, Sato KY, Fasching CL, Der CJ, Stanbridge EJ. Loss of oncogenic ras expression does not correlate with loss of tumorigenicity in human cells. Proc Natl Acad Sci U S A 1996;93:6665–70.
27. Chalepakis G, Fritsch R, Fickenscher H, Deutsch U, Goulding M, Gruss P. The molecular basis of the undulated/Pax-1 mutation. Cell 1991;66:873–84.
28. James AC, Veitch JG, Zareh AR, Triche T. Sensitivity and specificity of five abundance estimators for high-density oligonucleotide microarrays. Bioinformatics 2004;20:1060–5.
29. Hosack DA, Dennis G, Jr., Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. Genome Biol 2003;4:R70.
30. Zeng F, Schultz RM. RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. Dev Biol 2005;283:40–57.
31. Becker KG, Hosack DA, Dennis G, Jr., et al. PubMatrix: a tool for multiplex literature mining. BMC Bioinformatics 2003;4:61.
32. Tiffin N, Williams RD, Shipley J, Pritchard-Jones K. PAX7 expression in embryonal rhabdomyosarcoma suggests an origin in muscle satellite cells. Br J Cancer 2003;89:327–32.
33. Anderson J, Ramsay A, Gould S, Pritchard-Jones K. PAX3-FKHR induces morphological change and enhances cellular proliferation and invasion in rhabdomyosarcoma. Am J Pathol 2001;159:1089–96.
34. Barr FG. Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma. Oncogene 2001;20:5736–46.
35. 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;64:5539–45.
36. Hu-Lieskovan S, Zhang J, Wu L, Shimada H, Schofield DE, Triche TJ. EWS-FLI1 fusion protein up-regulates critical genes in neural crest development and is responsible for the observed phenotype of Ewing's family of tumors. Cancer Res 2005;65:4633–44.
37. Keller C, Hansen MS, Coffin CM, Capecchi MR. Pax3:Fkhr interferes with embryonic Pax3 and Pax7 function: implications for alveolar rhabdomyosarcoma cell of origin. Genes Dev 2004;18:2608–13.
38. Linardic CM, Downie DL, Qualman S, Bentley RC, Counter CM. Genetic modeling of human rhabdomyosarcoma. Cancer Res 2005;65:4490–5.
39. Epstein JA, Lam P, Jepeal L, Maas RL, Shapiro DN. Pax3 inhibits myogenic differentiation of cultured myoblast cells. J Biol Chem 1995;270:11719–22.
40. Blais A, Tsikitis M, Acosta-Alvear D, Sharan R, Kluger Y, Dynlacht BD. An initial blueprint for myogenic differentiation. Genes Dev 2005;19:553–69.
41. Bergstrom DA, Penn BH, Strand A, Perry RL, Rudnicki MA, Tapscott SJ. Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. Mol Cell 2002;9:587–600.
42. Tonin PN, Scrable H, Shimada H, Cavenee WK. Muscle-specific gene expression in rhabdomyosarcomas and stages of human fetal skeletal muscle development. Cancer Res 1991;51:5100–6.
43. Berkes CA, Tapscott SJ. MyoD and the transcriptional control of myogenesis. Semin Cell Dev Biol 2005; 16:585–95.
44. McKinsey TA, Zhang CL, Olson EN. Control of muscle development by dueling HATs and HDACs. Curr Opin Genet Dev 2001;11:497–504.
45. Slack A, Chen Z, Tonelli R, et al. The p53 regulatory gene MDM2 is a direct transcriptional target of MYCN in neuroblastoma. Proc Natl Acad Sci U S A 2005;102: 731–6.
46. Butt AJ, Williams AC. IGFBP-3 and apoptosis—a license to kill? Apoptosis 2001;6:199–205.
47. Chen P, Hutter D, Yang X, Gorospe M, Davis RJ, Liu Y. Discordance between the binding affinity of mitogen-activated protein kinase subfamily members for MAP kinase phosphatase-2 and their ability to activate the phosphatase catalytically. J Biol Chem 2001;276:29440–9.
48. Sasaki A, Taketomi T, Kato R, et al. Mammalian Sprouty4 suppresses Ras-independent ERK activation by binding to Raf1. Nat Cell Biol 2003;5:427–32.
49. Hart KC, Robertson SC, Kanemitsu MY, Meyer AN, Tynan JA, Donoghue DJ. Transformation and Stat activation by derivatives of FGFR1, FGFR3, and FGFR4. Oncogene 2000;19:3309–20.
50. Li L, Zhou J, James G, Heller-Harrison R, Czech MP, Olson EN. FGF inactivates myogenic helix-loop-helix proteins through phosphorylation of a conserved protein kinase C site in their DNA-binding domains. Cell 1992;71:1181–94.
51. Shaoul E, Reich-Slotky R, Berman B, Ron D. Fibroblast growth factor receptors display both common and distinct signaling pathways. Oncogene 1995;10: 1553–61.
52. Asmar L, Gehan EA, Newton WA, et al. Agreement among and within groups of pathologists in the classification of rhabdomyosarcoma and related childhood sarcomas. Report of an international study of four pathology classifications. Cancer 1994;74:2579–88.
53. Ruymann FB, Grovas AC. Progress in the diagnosis and treatment of rhabdomyosarcoma and related soft tissue sarcomas. Cancer Invest 2000;18:223–41.
54. Crist WM, Garnsey L, Beltangady MS, et al. Prognosis in children with rhabdomyosarcoma: a report of the Intergroup Rhabdomyosarcoma Studies I and II. Intergroup Rhabdomyosarcoma Committee. J Clin Oncol 1990;8:443–52.
55. Dave SS, Wright G, Tan B, et al. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. N Engl J Med 2004;351: 2159–69.
56. Pittman J, Huang E, Dressman H, et al. Integrated modeling of clinical and gene expression information for personalized prediction of disease outcomes. Proc Natl Acad Sci U S A 2004;101:8431–6.
57. Mayanil CS, George D, Freilich L, et al. Microarray analysis detects novel Pax3 downstream target genes. J Biol Chem 2001;276:49299–309.
58. Ginsberg JP, Davis RJ, Bennicelli JL, Nauta LE, Barr FG. Up-regulation of MET but not neural cell adhesion molecule expression by the PAX3-FKHR fusion protein in alveolar rhabdomyosarcoma. Cancer Res 1998;58: 3542–6.
59. Epstein JA, Shapiro DN, Cheng J, Lam PY, Maas RL. Pax3 modulates expression of the c-Met receptor during limb muscle development. Proc Natl Acad Sci U S A 1996;93:4213–8.
60. Maroto M, Reshef R, Munsterberg AE, Koester S, Goulding M, Lassar AB. Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. Cell 1997;89:139–48.
61. Zhang L, Wang C. PAX3-FKHR transformation increases 26S proteasome-dependent degradation of p27Kip1, a potential role for elevated Skp2 expression. J Biol Chem 2003;278:27–36.