Clinical Application of Prognostic Gene Expression Signature in Fusion Gene–Negative Rhabdomyosarcoma: A Report from the Children’s Oncology Group


Abstract

Purpose: Pediatric rhabdomyosarcoma (RMS) has two common histologic subtypes: embryonal (ERMS) and alveolar (ARMS). PAX–FOXO1 fusion gene status is a more reliable prognostic marker than alveolar histology, whereas fusion gene–negative (FN) ARMS patients are clinically similar to ERMS patients. A previous expression signature (MG5) previously identified two diverse risk groups within the fusion gene–negative RMS (FN-RMS) patients, but this has not been independently validated. The goal of this study was to test whether expression of the MG5 metagene, measured using a technical platform that can be applied to routine pathology material, would correlate with outcome in a new cohort of patients with FN-RMS.

Experimental Design: Cases were taken from the Children’s Oncology Group (COG) D9803 study of children with intermediate-risk RMS, and gene expression profiling for the MG5 genes was performed using the nCounter assay. The MG5 score was correlated with clinical and pathologic characteristics as well as overall and event-free survival.

Results: MG5 standardized score showed no significant association with any of the available clinicopathologic variables. The MG5 signature score showed a significant correlation with overall survival (N = 57; HR, 7.3; 95% CI, 1.9–27.0; P = 0.003) and failure-free survival (N = 57; HR, 6.1; 95% CI, 1.9–19.7; P = 0.002).

Conclusions: This represents the first, validated molecular prognostic signature for children with FN-RMS who otherwise have intermediate-risk disease. The capacity to measure the expression of a small number of genes in routine pathology material and apply a simple mathematical formula to calculate the MG5 metagene score provides a clear path toward better risk stratification in future prospective clinical trials.

Introduction

Rhabdomyosarcoma (RMS) is the most common soft tissue sarcoma in children and adolescents. Traditional risk stratification of RMS patients is based on a complex algorithm of clinical "group" and "stage," which is determined by histology (embryonal vs. alveolar), primary tumor site (favorable vs. unfavorable), size, presence or absence of nodal or distant metastases, and degree of primary tumor resection at diagnosis. On the basis of the algorithm, the Children’s Oncology Group (COG) classifies patients into low-, intermediate-, and high-risk groups that are used for treatment allocation (1). Risk stratification varies between the North American COG and European national sarcoma groups, complicating the comparison of clinical trials results. In general, the prognosis for children with embryonal RMS (ERM) is more favorable than those with alveolar RMS (ARM) for similar stage/clinical group of patients (1, 2).

In the era of molecular profiling, several sarcomas have been reclassified for treatment purposes based on molecular prognostic markers. In RMS, the hallmark molecular signature was the identification of PAX3–FOXO1 gene fusion in 1993 (3). A less frequent PAX7–FOXO1 gene fusion was identified in 1994 (4). Although these gene fusions are essentially unique to ARMS, they are only found in about 70% to 80% of histologically defined ARMS (5, 6). Several studies have suggested that PAX–FOXO1 fusion gene status portends a worse prognosis than fusion gene–negative (FN) ARMS (7, 8). The PAX3–FOXO1 fusion may signify an inferior outcome than the PAX7–FOXO1 fusion, although these data remain conflicting (9–11). Furthermore, patients with FN-ARMS have clinical outcomes similar to ERMS patients,
Translational Relevance
To date, there are no known prognostic markers identified in patients with fusion gene–negative rhabdomyosarcoma (FN-RMS). This study validates the five-gene (MG5) signature as a prognostic marker in patients with fusion gene–negative intermediate-risk RMS, clearly stratifying this otherwise clinically homogenous population of patients into two risk groups based on outcome. In addition, this analysis was performed using nCounter assay on paraffin-embedded tissues, and the results were concordant with previously published results using frozen tissues in a different patient cohort. Therefore, this work holds tremendous translational relevance, as the MG5 signature can be reliably assessed in readily available paraffin-embedded tissues of FN-RMS patients in prospective clinical trials to stratify them into prognostic risk groups as well as to potentially tailor future therapy based on these risk groups.

Materials and Methods

Patients and samples
This study involved 68 primary FN-RMS samples from patients treated on the COG study, D9803, from 1999 to 2005. The design and results from COG D9803 have been reported previously (14). All study subjects had intermediate-risk disease, based on accepted clinicopathology criteria to include the following: patients with nonmetastatic ARMS, patients with stage II or III, Clinical Group III ERMS, and patients <10 years of age with metastatic ERMS. On D9803, patients were randomized to receive vincristine, dactinomycin, and cyclophosphamide (VAC) or VAC alternating with vincristine, topotecan and cyclophosphamide (VTC). As part of D9803, paraffin-embedded primary biopsy/resection samples for patients were banked at the Cooperative Human Tissue Network Bank (Columbus, OH); these blocks were used for this study.

Fusion gene status in all cases was confirmed by either reverse-transcription polymerase chain reaction (RT-PCR) or fluorescent in-situ hybridization (FISH), as previously described. Cases with FN-RMS and adequate tissue, as well as clinicopathologic annotation, were included in the analysis.

RNA extraction from FFPE samples
FFPE tissue scraps were used for RNA extraction. RNA was extracted using the Qiagen AllPrep DNA/RNA FFPE Kit (Qiagen). Briefly, scraps were deparaffinized using xylene. Samples were then incubated with an optimized lysis buffer to release the RNA and precipitate the DNA. RNA containing supernatant was separated with centrifugation and then purified using the RNeasy MiniElute spin column. Purified RNA was treated with DNase to remove any contaminating DNA, then washed and eluted. Up to 400 ng of RNA was used with the nCounter assay (NanoString Technologies), and the Digital Analyzer was run at 1,150 fields of view for all samples.

Gene expression profiling using nCounter assay
The nCounter assay was used to measure the expression of 83 RNA transcripts, the expression of which was previously identified by us and others as having prognostic value for children with RMS (12, 13; Triche, unpublished data); additional transcripts included four housekeeping genes, and eight negative and six positive controls, in 84 samples representative of 68 primary FN RMS tumors. Here, we limit the analysis to the five genes that are represented in MG5 signature (EPHA2 (Ephrin Receptor A2), EED (Embryonic Ectoderm Development), NSMF (NMDA receptor synaptosomal signaling and neuronal migration factor), CBS (Cystathionine-β-synthase), and EPB41L1B (Erythrocyte Membrane Protein Band 4.1-like 4B; Supplementary Table S1).

Taking advantage of the microarray expressions available from two cohorts of primary RMS samples (COG/IRSG (N = 134; ref. 11) and ITCC/CIT (N = 101; ref. 8), we explored the variability of the genes included in the nCounter platform in order to identify potential invariant endogenous controls that could be suitable for data normalization. We selected 10 genes that showed standard deviation (SD) < 0.6 as well as consistent level of expression in both datasets (Supplementary Fig. S1), including DLG2, GRIK5, PGLI, ZNF671, CDYL, HPRT1, COMMD3, DYNLT1, RPL19, and ACTB. The samples went through a first quality assessment by evaluating distribution and level of expression of endogenous genes as well as positive and negative controls (Supplementary Fig. S2A–S2C). Normalization was performed by summarizing (by geometric mean) the positive and the invariant endogenous controls and adjusting samples by a relative factor. Background correction was performed using mean level of the negative controls plus 2 SDs. A second quality assessment was performed, including plotting of the log2 of the average raw counts versus the interquartile range (IQR) values (in log scale; Supplementary Fig. S2D–S2F). Samples showing low intensity associated with high variability (IQR > 7.6) were removed from the analysis. When sample was performed with technical replicates, only the one with the highest average endogenous gene expression was retained.
Nanostring analysis was performed using "NanoStringNorm" R package version 1.1.17. The gene expression profile data have been deposited at the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-3580.

Data analysis and correlation with clinicopathologic characteristics
MG5 score was computed from the log transformed data using the model previously described (13). In summary, the risk score is the weighted sum of the log expression intensity of the genes, using the weights determined in the original study. The score was standardized by subtracting the score median and dividing by the score IQR. Correlation of MG5 score with clinicopathologic variables was tested by Wilcoxon or Kruskal–Wallis rank sum tests. Correlation of the variables with OS and failure-free survival (FFS) at univariate or multivariate levels was tested using Cox proportional hazard regression models. Proportional hazard assumption was tested using the scaled Schoenfeld residual method. All analyses were performed using the R system for statistical computing, version 3.1.0 (R Core Team, 2013).

Results
Patient and sample characteristics
Primary tumor specimens were available for 68 FN-RMS treated on D9803 study, of which 68 specimens were FN RMS. Raw gene expression profile data were obtained from 84 samples, which included 16 technical replicates (derived from 11 patients). Twelve samples from 11 patients were discarded because of low quality. The nCounter assay was showed to be highly reproducible between replicated samples of 0.96 (range, 0.77–0.99). Upon removal of duplicated samples, the final dataset included 57 unique FN specimens (6 ARMS, 3 mixed, and 48 ERMS). Table 1 summarizes the clinicopathologic features of these 57 patients.

<table>
<thead>
<tr>
<th>Variable</th>
<th>OS</th>
<th>FFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histology (ERMS vs. ARMS_Neg)</td>
<td>1.29 (0.16–10.08)</td>
<td>8.11E–001</td>
</tr>
<tr>
<td>Histology (MIXED_Neg vs. ARMS_Neg)</td>
<td>1.74 (0.11-28.20)</td>
<td>6.96E–001</td>
</tr>
<tr>
<td>Gender (male vs. female)</td>
<td>0.55 (0.17-1.75)</td>
<td>3.04E–001</td>
</tr>
<tr>
<td>Age.cat (unfavorable vs. favorable)</td>
<td>1.33 (0.36–4.92)</td>
<td>6.70E–001</td>
</tr>
<tr>
<td>Tumor size (&lt;5 cm vs. &lt;5 cm)</td>
<td>0.59 (0.15–2.50)</td>
<td>4.49E–001</td>
</tr>
<tr>
<td>Tumor location (unfavorable vs. favorable)</td>
<td>0.47 (0.13–1.65)</td>
<td>2.33E–001</td>
</tr>
<tr>
<td>Group</td>
<td>0.62 (0.32–1.22)</td>
<td>1.66E–001</td>
</tr>
<tr>
<td>IRS stage</td>
<td>0.63 (0.37–1.09)</td>
<td>1.00E–001</td>
</tr>
<tr>
<td>Metastasis (M1 vs. M0)</td>
<td>0.28 (0.04–2.24)</td>
<td>2.33E–001</td>
</tr>
<tr>
<td>N stage (N1 vs. N0)</td>
<td>0.44 (0.06–3.47)</td>
<td>4.35E–001</td>
</tr>
</tbody>
</table>

| MGS score | 2.87 (1.52-5.27) |

Table 1. Summary of clinicopathologic features of the 57 FN-RMS patients included in the study.

NOTE: Boldface values are statistically significant.
None of the available clinicopathologic features showed a significant correlation with OS and FFS in univariate survival regression models in the cohort of 57 patients (Table 2).

**Correlation of MG5 risk score with clinicopathologic features and outcome**

MG5 standardized score showed no significant association with any of the clinicopathologic variables available, including gender, age, stage, risk group, tumor histology, size, and location (Supplementary Fig. S3). We confirmed that the expression of three out of the five genes included in the MG5 signature were significantly associated with a poorer outcome (Table 3). Moreover, we observed a significant association of the MG5 score, as a continuous variable, with both OS and FFS (Table 2). Proportional hazard assumption was violated for neither OS nor FFS. The AUC of the ROC was 0.68 and 0.71 at 3 years and 0.69 and 0.74 at 5 years for OS and FFS, respectively. To investigate the dependence of survival on MG5 score level, we plotted the expected OS and FFS probabilities at 3 and 5 years versus the MG5 score (Fig. 1A and B). As observed in the previous study, a low MG5 score was associated with a low risk for relapse or death, and the risk increased greatly in the top tertile. We split the low MG5 score was associated with a low risk for relapse or death, and the risk increased greatly in the top tertile. We split the low MG5 score group into two lower score groups (Low_MG5) to compare with the top tertile (High_MG5). Figure 2 shows a remarkable difference in OS and FFS between the low and high MG5 score groups, with HR values in the same range as those found in the previous study.

Multivariable analysis confirmed that MG5 score is an independent prognostic marker when combined with tumor histology (ARMS, ERMS) and metastatic status (M0, M1; Table 4).

**Discussion**

Although PAX-FOXO1 fusion status has recently become a well-accepted prognostic marker in ARMS, there are no similarly validated molecular prognostic features for FN-RMS. The results of our study confirm that MG5 is an independent prognostic marker in FN RMS and that a high MG5 score portends a poor prognosis in this group of patients.

Previously, the MG5 signature was shown to have prognostic implications in FN patients and seemingly divided the COG patients with high-risk disease into two separate risk groups (13). However, this study used the Affymetrix GeneChip Human U133 microarray platform to measure gene expression, which ideally requires availability of frozen tissues. To date, the need for frozen specimens generally limited the capacity for a molecular prognostic factor to be applied to very large numbers of specimens or incorporated into clinical practice. On the other hand, the nCounter assay allows for RNA extracted from FFPE tissues to be analyzed for gene expression. This approach shows excellent correlation with TaqMan based RT-PCR approach (19), and RNA expression quantified by nCounter has been used to discriminate FN from FP-RMS (20). The nCounter assay has also been used and validated in other tumor types such as breast cancer and neuroblastoma (21, 22).

In this study, we aimed to validate the MG5 signature using the FFPE-based nCounter approach in a cohort of 68 COG intermediate-risk patients representing a cohort distinct from those used to generate MG5. Our results showed that high-quality RNA was available for gene expression analysis in the majority of FFPE specimens (84%) of the patients and that the gene expression pattern was highly reproducible across technical replicates. More importantly, the gene signature divided a uniformly treated population of COG intermediate-risk FN-RMS patients into two distinct groups for both OS and FFS, thus validating that MG5 is indeed prognostic in FN RMS. Of note, our results did not show any correlation of survival with the available clinicopathologic features but this was expected as all patients included in this study were homogenously classified as intermediate-risk based on the clinical and histologic features.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Probability of OS (A) and FFS (B) survival at 3 (black line) and 5 years (gray line), depending on the MG5 score level. Dotted lines represent 95% CI. The box-and-whisker plot and density plot at the bottom of the graphs represent the distribution of the score levels. Each of the vertical segments at the bottom margin represents a patient. The vertical dotted line represents the MG5 top tertile.

<table>
<thead>
<tr>
<th>Gene</th>
<th>HR (95% CI) OS</th>
<th>P</th>
<th>HR (95% CI) FFS</th>
<th>P</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPHA2</td>
<td>2.65 (1.16–6.03)</td>
<td>2.06E-002</td>
<td>2.76 (1.31–5.81)</td>
<td>7.45E-003</td>
<td>57</td>
</tr>
<tr>
<td>EED</td>
<td>1.00 (0.29–3.47)</td>
<td>9.99E-001</td>
<td>1.08 (0.34–3.40)</td>
<td>8.98E-001</td>
<td>57</td>
</tr>
<tr>
<td>NSMF</td>
<td>2.08 (1.16–3.74)</td>
<td>1.46E-002</td>
<td>2.59 (1.36–4.91)</td>
<td>3.67E-003</td>
<td>57</td>
</tr>
<tr>
<td>CBS</td>
<td>1.10 (0.73–1.66)</td>
<td>6.44E-001</td>
<td>1.09 (0.74–1.59)</td>
<td>6.77E-001</td>
<td>57</td>
</tr>
<tr>
<td>EPB41L4B</td>
<td>1.34 (1.02–1.77)</td>
<td>3.72E-002</td>
<td>1.29 (1.00–1.66)</td>
<td>4.89E-002</td>
<td>57</td>
</tr>
</tbody>
</table>

**Table 3.** Association between expression of the genes included in the MG5 signature with OS and FFS in fusion-negative RMS patients—univariate Cox proportional hazard regression model.

NOTE: Boldface values are statistically significant.
The five analytes identified in the MG5 signature include EPHA2 (chromosome 1), EED (chromosome 11), NSMF (chromosome 9), CBS (chromosome 21), and EPB41L4B (chromosome 9). Although the role of these genes in RMS remains to be specifically defined, majority of them have been implicated in cancer. EPHA2 receptor overexpression in gastric cancer stromal cells is a prognostic factor for relapse (23). It also promotes tumor cell proliferation and motility in non–small cell lung cancer (24); and increases infiltrative capacity of glioma stem cells in glioblastoma multiforme models (25). EPHA2 is also the most abundant surface receptor expressed in osteosarcoma, and patients with EPHA2-positive disease tend to have inferior OS (26). In vitro models have shown EPHA2 can foster angiogenesis in Ewing sarcoma (27). In RMS, it is a downstream target of the PAX3–FOXO1 fusion protein as PAX3–FOXO1 fusion protein has been shown to directly bind the EPHA2 promoter (28). Therefore, overexpression of EPHA2 in FN RMS might be a potential parallel to fusion positive RMS in the context of poor patient outcome.

EED is a member of the Polycomb repressor complex 2 (PRC2), which is important in epigenetic regulation of cancer. It is upregulated in breast cancer lymph node metastasis and correlates with tumor proliferation (29). PRC2 is also highly expressed in mesothelioma and knockdown of EED leads to decreased proliferation, migration, and tumorigenicity of mesothelioma cells (30). PRC2 overexpression also occurs in prostate cancer and adenoid cystic carcinoma of the salivary gland and portends a poor prognosis (31, 32). Recently, RMS was shown to highly express JARID2, a gene that encodes a protein that recruits histone-methylating complexes to their target genes. Furthermore, JARID2 is a direct transcriptional target of PAX3–FOXO1 fusion protein that through PRC2 represses expression of genes involved in myogenic differentiation and maintains the undifferentiated phenotype seen in RMS. Interaction of PRC2 with myogenic gene promoters is dependent on EED (33). The polycomb group protein, Enhancer of Zeste Homolog 2 (EZH2), also supports survival and proliferation of ARMS cells by repressing F-box protein 32 (FBXO32), a gene associated with muscle homeostasis (34, 35). CBS encodes cystathionine β-synthase, an enzyme in the folate metabolic pathway, which is closely linked to DNA methylation. Dysregulation of this pathway contributes to cancer pathogenesis, and induction of CBS promotes cellular proliferation in colorectal cancer (36). It also promotes ovarian tumor growth and drug resistance (37). How EPB41L4B might relate to cancer is unclear, but Ehm2, a member of the NF2/ERM/4.1 superfamily, has been implicated in breast and prostate cancer invasiveness (38, 39). Thus, beyond biomarker use, these genes might also contribute to RMS pathogenesis.

Our study confirms the ability of MG5 signature to identify different risk groups within FN-RMS from a different patient cohort composed of those with intermediate-risk disease and signifies the ability of this signature to be able to identify poor-risk patients in this FN cohort. Nonetheless, it will be important to validate the potential clinical usefulness of this prognostic biomarker further in the context of a larger prospective clinical trial.

### Table 4. Multivariate Cox proportional hazard regression model including MG5 score (as continuum), tumor histology (ARMS, ERMS), and metastatic status (M0, M1)

<table>
<thead>
<tr>
<th>Variables</th>
<th>OS</th>
<th>P</th>
<th>N</th>
<th>FFS</th>
<th>P</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG5 score (M1 vs. M0)</td>
<td>2.7 (1.39–5.18)</td>
<td>3.40E-003</td>
<td>53</td>
<td>2.95 (1.53–5.68)</td>
<td>1.27E-003</td>
<td>53</td>
</tr>
<tr>
<td>Metastasis (M1 vs. M0)</td>
<td>0.4 (0.05–2.92)</td>
<td>3.42E-001</td>
<td>53</td>
<td>0.72 (0.16–3.34)</td>
<td>6.72E-001</td>
<td>53</td>
</tr>
<tr>
<td>Histology (ERMS vs. ARMS_Neg)*</td>
<td>1.6 (0.17–14.45)</td>
<td>6.94E-001</td>
<td>53</td>
<td>1.46 (0.17–12.33)</td>
<td>7.26E-001</td>
<td>53</td>
</tr>
</tbody>
</table>

**Note:** Boldface values are statistically significant.

*In this analysis, the 3 patients with Mixed histology were pooled with the alveolar histology.

---

**Figure 2.** The Kaplan–Meier curves for OS (A) and FFS (B) splitting MG5 signature score in two groups. MG5 was categorized using tertiles: High_MG5 were patients whose MG5 score expression was within the top tertile; all the other were defined as Low_MG5.
Ultimately, our goal will be to apply the “MG5 score” to an individual patient with FN-RMS. How robustly this discriminates individuals destined to fare well or poorly with a given therapy also remains to be determined. It is also important to emphasize that the survival advantage for that subset of children with a favorable MG5 score is based on the therapy that is currently applied to patients with intermediate risk disease. Whether therapy intensity can be decreased while maintaining the excellent survival—or whether intensifying therapy can improve survival in that subset with an unfavorable MG5 score—must also be prospectively tested.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Development of methodology: T.J. Triche, D.S. Hawkins
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J.R. Anderson, T.J. Triche, J. Gastier-Foster, M. Wing, S.X. Skapek

References


Clinical Application of Prognostic Gene Expression Signature in Fusion Gene–Negative Rhabdomyosarcoma: A Report from the Children’s Oncology Group

Pooja Hingorani, Edoardo Missiaglia, Janet Shipley, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/21/20/4733

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2015/10/13/21.20.4733.DC1

Cited articles
This article cites 37 articles, 13 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/21/20/4733.full.html#ref-list-1

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

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

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.