Novel Markers of Subclinical Disease for Ewing Family Tumors from Gene Expression Profiling

Irene Y. Cheung, Yi Feng, Karen Danis, Neerav Shukla, Paul Meyers, Marc Ladanyi, and Nai-Kong V. Cheung

Abstract

**Purpose:** Targeting subclinical disease in the bone marrow is particularly relevant in metastatic Ewing family tumors (EFT) where cure is difficult. Genome-wide expression arrays can uncover novel genes differentially expressed in tumors over normal marrow/blood, which may have potentials as markers of subclinical disease.

**Experimental Design:** Gene expression array data were obtained on 28 EFT tumors using the Affymetrix U133 gene chip and compared with 10 normal blood samples. Ten genes with high tumor to blood ratios were identified. Quantitative reverse transcription-PCR was done to study (a) the dynamic range of detection of rare tumor cells, (b) the gene expression in normal blood/marrow samples, (c) the gene expression among EFT tumors, and (d) the detection and prognostic impact of marker positivity in histology-negative diagnostic marrows of EFT patients.

**Results:** Five of 10 genes (i.e., six-transmembrane epithelial antigen of the prostate 1 [STEAP1], cyclin D1 [CCND1], NNX2-2 transcription factor [NKKX2-2], plakophilin 1 [PKP1], and transmembrane protein 47 [TMEM47]) were chosen for further analyses based on their steep linear dynamic range in detecting tumor cells seeded in normal mononuclear cells and on their homogeneous expression among EFT tumors. Prognostic effect was evaluated in 35 histology-negative diagnostic marrows. Marker negativity of STEAP1, CCND1, or NNX2-2, as well as three markers in combination, was strongly correlated with patient survival as well as survival without new metastases.

**Conclusions:** This gene expression array-based approach identified novel markers that may be informative at diagnosis for risk group assessment. Their clinical utility needs to be tested in large patient cohorts.

Before the era of effective chemotherapy, 5-year overall survival for Ewing family tumors (EFT) was ≤20%, largely because of uncontrolled subclinical metastatic disease (1). Despite the advent of modern chemoradiotherapy, the 5-year progression-free survival was only 30% among patients with overt metastases at the time of diagnosis, although many of them had achieved near complete clinical remission following induction therapy. In fact, even for patients presenting with localized disease at diagnosis, their overall survival decreased from 70% at 5 years to ~50% at 10 years (1). The failure to cure is likely due to occult cancer cells that metastasize by hematogenous spread (2). Accurate and sensitive markers of EFT in the marrow or blood may help in the early detection of distant metastasis. More importantly, testing marrow or blood can facilitate the application of novel treatment approaches directed at minimal residual disease in the adjuvant setting, when there is no radiographic or histologic evidence of disease.

All EFT tumors are characterized by similar chromosomal translocations that result in the fusion of the EWS gene on chromosome 22q12 with different ETS-related genes. More than 85% of analyzed cases of EFT are characterized by a t(11;22)(q24;q12) translocation. The most common fusion is made up of exon 7 of the EWS gene and exon 6 of the FLI1 gene (type 1 fusion), which was associated with favorable prognosis (3). Less common translocations (≤10%) in EFT involve other members of the ETS family of transcription factors. These include ERG (located on chromosome 21), ETVI (located on chromosome 7), EIAF (located on chromosome 17), and FEV (located on chromosome 2), resulting in t(21;22), t(7;22), t(17;22), and t(2;22) translocations, respectively (1). Among the chimeric transcripts resulting from such fusion, EWS-FLI1 (type 1 or 2) and EWS-ERG fusion transcripts are most widely used as molecular markers. Nested reverse transcription-PCR (RT-PCR) of fusion sequences (4) has enabled tumor detection at frequency of 10^-6 and has provided highly specific and sensitive tools for detecting circulating EFT in the bone marrow and blood. In fact, micrometastatic disease...
Table 1. Median expression levels and ratios to peripheral blood of candidate markers of EFTs

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene symbol</th>
<th>Tumor</th>
<th>PB</th>
<th>Tumor:PB</th>
<th>ABI assay ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclin D1</td>
<td>CCND1</td>
<td>9,720</td>
<td>48</td>
<td>201</td>
<td>HS00277039_m1</td>
</tr>
<tr>
<td>Plakophilin 1</td>
<td>PKP1</td>
<td>3,311</td>
<td>20</td>
<td>164</td>
<td>HS0040873_m1</td>
</tr>
<tr>
<td>NK2 transcription factor</td>
<td>NKX2-2</td>
<td>6,077</td>
<td>46</td>
<td>132</td>
<td>HS00159616_m1</td>
</tr>
<tr>
<td>Six-transmembrane epithelial antigen of the prostate 1</td>
<td>STEAP1</td>
<td>7,805</td>
<td>123</td>
<td>64</td>
<td>HS00185180_m1</td>
</tr>
<tr>
<td>Transmembrane protein 47</td>
<td>TMEM47</td>
<td>3,594</td>
<td>57</td>
<td>63</td>
<td>HS00230190_m1</td>
</tr>
<tr>
<td>Chondroitin sulfate proteoglycan 5</td>
<td>CSPG5</td>
<td>3,035</td>
<td>81</td>
<td>37</td>
<td>HS00198108_m1</td>
</tr>
<tr>
<td>Fibronectin leucine-rich transmembrane protein 2</td>
<td>FLRT2</td>
<td>3,803</td>
<td>104</td>
<td>37</td>
<td>HS00544171_s1</td>
</tr>
<tr>
<td>Microtubule-associated protein (\tau)</td>
<td>MAPT</td>
<td>5,050</td>
<td>184</td>
<td>27</td>
<td>HS00213484_m1</td>
</tr>
<tr>
<td>Wingless-related MMTV integration site 5A</td>
<td>WNT5A</td>
<td>2,064</td>
<td>74</td>
<td>28</td>
<td>HS00180103_m1</td>
</tr>
<tr>
<td>FAT tumor suppressor homologue 4</td>
<td>FAT4</td>
<td>2,026</td>
<td>93</td>
<td>22</td>
<td>HS00707110_s1</td>
</tr>
</tbody>
</table>

Abbreviations: PB, peripheral blood; ABI, Applied Biosystems.

Materials and Methods

Identification of potential markers of subclinical EFT by genome-wide gene expression array analyses. Affymetrix Human U133 oligonucleotide array data on 28 fresh-frozen EFT and 10 EFT cell lines (CADO-ES-1, SK-PN-DW, 6647, A673, MHH-ES-1, RD-ES, SK-ES, SK-ES-S1, SK-N-MC, and TC32) as previously described (11) were compared with publicly available data using the Affymetrix U133 chip from 10 normal blood and 10 normal bone marrow samples (12, 13). An EFT cell line served as a positive control in each experiment with the coefficient of variance within 5%. This study was done blinded to the patient outcome.

Cell lines and buffy coat for sensitivity studies. SK-EKT, an EFT cell line established at Memorial Sloan-Kettering Cancer Center, and A673 from the American Type Culture Collection were cultured in RPMI 1640 with 10% calf serum and penicillin/streptomycin at 37°C. Both cell lines have t(11;22)(q24;q12) translocation: SK-EKT at exon 10 of the EWS gene and exon 5 of the FLI1 gene and A673 at exon 7 of the EWS gene and exon 6 of the FLI1 gene (type 1 fusion). Buffy coat was obtained from New York Blood Center (New York, NY).

Characteristics of EFT patient tumors and marrow samples for quantitative RT-PCR studies. Twenty EFT tumor samples obtained at diagnosis and relapse at Memorial Sloan-Kettering Cancer Center were snap frozen in liquid nitrogen. Archived diagnostic bone marrow samples (n = 35) from EFT patients treated at Memorial Sloan-Kettering Cancer Center from 2001 to 2005 with no evidence of EFT by histologic examination were included for this study in accordance with the guidelines of the institutional review board. Marrow samples obtained at relapse were excluded from this study. The sample size was determined by tissue availability and not by power calculations.

Among this retrospective cohort, there were 25 males and 10 females. Their median age at diagnosis was 16 years; 21 patients were older than 14 years at diagnosis. Primary tumors of 23 patients were localized in the axial skeleton, whereas extremities were the primary sites for the rest of the patients. Seven of the 35 patients had primaries in the pelvis. Fifteen patients had primary tumor measuring ≥8 cm, 13 patients had <8 cm, and no measurement information on 7 patients. Eighteen of 35 patients had elevated serum lactate dehydrogenase (LDH; >200 units/mL). Three patients were diagnosed with lung metastasis and two with metastasis to the lung and bony sites. Treatment received was P6 protocol (n = 17; ref. 12) and modified P6 protocol with an augmented dose of ifosfamide (n = 18). The median time of follow-up among survivors was 36.1 months. March 2007 was the end of the follow-up period.

Molecular analysis. Mononuclear cells were isolated from marrow samples, total RNA was isolated, and quality was assessed as previously described (13–15). An aliquot was used for EWS-FLI1 studies and the rest was cryopreserved at -80°C. cDNA was synthesized from 1 μg of total RNA. cDNA (1 μL) was used for real-time quantitative PCR using Applied Biosystems Sequence Detection System 7300. Housekeeping gene \(\beta_2\) microglobulin was used for normalization, and each sample was quantified using the comparative \(C_T\) method (Applied Biosystems) as a relative fold difference to the positive control cell line SK-EKT or A673. All gene expression assay reagents were purchased from Applied Biosystems; their assay ID is shown in Table 1. An EFT cell line served as a positive control in each experiment with the coefficient of variance of the \(C_T\) within 5%. This study was done blinded to the patient outcome.

For the detection of fusion transcript EWS-FLI1 type 1 fusion by quantitative PCR, the forward primer was ACTATGTACCCACCC-CAAACATTG (EWS exon 7), the antisense primer was TTACTGACGTGGTGGCCCTCT (FLI1 exon 7), and the probe was FAM-AAGCCCGAGTCAATAGGCTCATMRA (EWS exon 7). Statistical analysis. Bone marrow was classified as marker positive if the gene transcript level was greater than the upper limit of normal as defined as mean + 2 SD of 26 normal marrow and blood samples. Two
clinical endpoints were used, overall survival and survival without new metastases, both measured from diagnosis (i.e., the date of marrow sampling). The following prognostic factors were tested in univariate analyses: serum LDH level, age at diagnosis, primary tumor size, primary site, metastatic disease (16), and novel marker positivity. The effect of prognostic factors on time to death or to develop new metastases was calculated using Kaplan-Meier method and compared by the log-rank test. Only those factors that were statistically significant ($P < 0.05$) in the univariate model were included in the multivariate Cox regression model. We regarded the multivariate analysis with small sample size as preliminary because 12 relapses and 11 deaths represented relatively few events.

Results

Candidate markers of subclinical EFT from gene expression array analyses. For each probe in the U133 chip, the gene expression levels of 28 EFT tumors and 10 EFT cell lines were compared with their levels in 10 normal peripheral blood samples. Only genes with highly significant tumor to blood ratio ($>20; P < 2 \times 10^{-6}$ using Bonferroni correction for multiple comparisons) were chosen. After excluding genes of ubiquitous nature, such as collagen and pseudogenes, as well as genes with known expression in marrow, 10 genes with median expression level of $>2,000$ units were filtered from $\sim 22,000$ gene probes. Their expression levels were compared with blood and ranked in descending tumor to blood ratio (Table 1).

Sensitivity of novel EFT markers. Cells from EFT cell lines SK-ERT or A673 (defined as 100,000 transcript units) were seeded into 10$^7$ normal peripheral blood mononuclear cells, ranging from 1 to 1,000 tumor cells. EWS-FLI type 1 fusion transcript expressed by A673 and not SK-ERT was included as the “gold standard” for comparison. The performance of five genes (i.e., CSPG5, FLRT2, MAPT, WNT5A, and FAT4) was suboptimal because of their relatively high expression among normal mononuclear cells in the seeding experiments (Fig. 1A). When NKX2-2 transcription factor (NKX2-2) and EWS-FLI type 1 fusion were tested, their sensitivities were 1 in 10$^5$ (Fig. 1B). The detection limit of plakophilin 1 (PKP1) and transmembrane protein 47 (TMEM47) was 1 in 10$^6$ (Fig. 1C), whereas the sensitivity of six-transmembrane epithelial antigen of the prostate 1 (STEAP1) and cyclin D1 (CCND1) was 1 tumor cell in 10$^7$ normal hematopoietic cells (Fig. 1D).

Marker specificity and expression in EFT tumors. Based on the sensitivity studies, we chose the best five markers for further

| Table 2. Univariate analyses on overall survival and survival without new metastases |
|-----------------|-----------------|-----------------|
| Marker          | Overall survival ($P$) | Survival without new metastases ($P$) |
| STEAP1          | 0.0012           | 0.0009          |
| CCND1           | 0.0077           | 0.0014          |
| NKX2-2          | 0.0017           | 0.0008          |
| Marker panel*   | 0.0006           | 0.0001          |
| PKP1            | 0.13             | 0.26            |
| TMEM47          | 0.11             | 0.11            |

NOTE: Negative panel: negative in all three markers; positive panel: positive in any one of these three markers.

*Marker panel includes STEAP1, CCND1, and NKX2-2.
analyses (i.e., STEAP1, CCND1, PKP1, TMEM47, and NKX2-2). The upper limit of normal, defined as mean + 2 SD of 26 normal marrow and blood samples, was 11, 7, 13, 14, and 0 units, respectively. A marker was classified as positive if the transcript level was greater than the upper limit of normal. Twenty EFT tumors were tested for the expression of these five potential markers as well EWS-FLI (type 1). Tumors (100%) were positive for each of the five markers and 80% had detectable type 1 fusion transcript.

Detection of gene expression of novel markers in diagnostic marrows. Diagnostic bone marrow samples from 35 EFT patients were evaluated using these five markers plus EWS-FLI type 1 fusion transcript for comparison. The patient characteristics were detailed in Materials and Methods. Median time of follow-up was 36.1 months, and the median time to develop new metastases was 13.4 months. All bone marrows were negative for histology and undetectable EWS-FLI type 1 fusion transcript by quantitative PCR. In contrast, all five novel markers had detectable expression in a subset of samples: CCND1 had the highest number of positives samples (n = 12), STEAP1 and TMEM47 had six positives each, PKP1 had four, and NKX2-2 had two.

patients were evaluated using these five markers plus EWS-FLI type 1 fusion transcript for comparison. The patient characteristics were detailed in Materials and Methods. Median time of follow-up was 36.1 months, and the median time to develop new metastases was 13.4 months. All bone marrows were negative for histology and undetectable EWS-FLI type 1 fusion transcript by quantitative PCR. In contrast, all five novel markers had detectable expression in a subset of samples: CCND1 had the highest number of positives samples (n = 12), STEAP1 and TMEM47 had six positives each, PKP1 had four,
and NKX2-2 had two. The marrows of 16 patients were positive in one or more of these markers.

**Prognostic impact of marker status on patient survival.**

Demographics of the 35 patients were detailed in Materials and Methods. Using univariate analysis to test these five markers, positivity of STEAP1, CCND1, or NKX2-2, individually and in combination (marker panel), was highly prognostic with respect to patient survival ($P = 0.0006$), as well as surviving without new metastases ($P = 0.0001$; Table 2). As expected, patients who presented with clinically detectable metastasis at diagnosis were statistically more likely to develop new metastases ($P = 0.02$; Fig. 2A) and to eventually succumb to the disease ($P = 0.04$; Fig. 3A). In this cohort, the marrow marker status at diagnosis was highly prognostic (Figs. 2B and 3B). Among the 20 patients whose diagnostic marrows were negative for all three markers, only 2 patients had developed new metastases and eventually died. This suggests that marrow marker status may be informative at diagnosis for risk group assessment.

Univariate analyses on overall survival and survival without developing new metastasis were also tested on standard prognostic variables, including elevated serum LDH at diagnosis (>200 units/mL), age at diagnosis (>14 years of age), primary tumor size (>8 cm), and primary tumor site (axial versus extremity, pelvis versus nonpelvis). All variables, except serum LDH at diagnosis, were statistically insignificant. As shown in Table 3, among the three variables (marker panel, clinical metastasis, and serum LDH) included in the multivariate Cox regression analysis, marker panel status (positive versus negative) was the only statistically significant variable associated with overall survival and survival without new metastasis.

### Discussion

A widely used strategy to identify tumor markers has been to survey the world’s published literature for proteins or genes expressed in tumors. This approach, however, was shown to be highly inadequate because specificity and sensitivity of detection were not part of their discovery criteria. Such deficiencies particularly hold true for orphan diseases such as solid tumors in children where expressed proteins are generally not well known. In the past decade, molecular methods for the detection of cancer have greatly energized the cancer community because of their promise of exquisite specificity and sensitivity (17). Genomics and proteomics analysis of normal and diseased tissues has yielded many genes and their products for cancer profiling as well as the development of targeted therapeutics (18–20). Nevertheless, previous efforts to identify markers of subclinical disease have often failed to consider the tissue compartments (i.e., bone marrow and peripheral blood) in which the genomic information will be most useful.

The marker discovery outlined in this report took into account the gene expression of EFT tumors to normal blood ratios as well as expression level. Approximately 22,000 gene probes were filtered to a list of 10 genes as potential markers. Among the 10 genes identified by this genome-wide gene expression array-based approach, the sensitivity performance of five genes (i.e., CSPG5, FLRT2, MAPT, WNT5A, and FAT4) was suboptimal because of their relatively high expression in normal hematopoietic cells. Their tumor to blood ratios also ranked lowest in the gene array analysis (Table 1). Among the top four novel genes (i.e., STEAP1, CCND1, PKP1, and TMEM447), their sensitivity was 10- to 100-fold superior to the EWS-FLI1 type 1 fusion transcript when tested by quantitative RT-PCR, whereas NKX2-2 had comparable sensitivity. Being the most common hybrid transcript, EWS-FLI1 type 1 fusion was a logical choice selected for comparison with these novel markers.

Previous studies suggested that fusion transcript positivity was associated with a greater risk for recurrence as well as for distant metastasis (7, 8). However, not all EFT tumors have this specific chromosomal translocation; thus, the marrows of not all patients could have detectable fusion transcript. In contrast, all five novel markers were uniformly expressed in EFT tumor samples tested.

A cohort of diagnostic marrows that were histologically negative was used to determine the clinical relevance of these novel markers in the subclinical setting. In contrast to the absence of fusion transcript in all the marrows, our findings suggest that the detection of STEAP1, CCND1, and/or NKX2-2 at diagnosis was informative and was predictive of whether patients had a higher likelihood to eventually develop new metastases and to die of EFT. These three markers remained prognostic even after Bonferroni adjustment for multiple comparisons as well as multivariate analysis. These results support our hypothesis that differential tumor gene expression analyses can identify sensitive and specific tumor markers of subclinical metastatic disease.

STEAP1 is a 339-amino acid protein with six potential membrane-spanning regions flanked by hydrophilic NH2-terminal and COOH-terminal domains, suggesting that it folds in a serpentine manner into three extracellular and two intracellular loops (21). STEAP1 is expressed predominantly in human prostate tissue and is up-regulated in multiple cancer cell lines, including prostate, bladder, colon, ovarian, melanoma, rhabdomyosarcoma, and EFT (21, 22). It is highly restricted to secretory epithelium of prostate and bladder tissue in normal tissues. Because of its tissue specificity, STEAP may have

### Table 3. Multivariate analysis of overall survival and survival with new metastases

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Overall survival</th>
<th>Survival without new metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard ratio (95% CI)</td>
<td>$P$</td>
</tr>
<tr>
<td>Marker panel</td>
<td>8.19 (1.44-46.43)</td>
<td>0.018</td>
</tr>
<tr>
<td>Serum LDH</td>
<td>1.01 (0.22-4.65)</td>
<td>0.99</td>
</tr>
<tr>
<td>Metastasis at diagnosis</td>
<td>1.78 (0.44-7.17)</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Abbreviation: 95% CI, 95% confidence interval.
potential as a tumor target for CTLs (22–24). CCND1 plays a pivotal role in controlling cyclin-dependent kinases during cell cycle progression (25), and it has overexpression and adverse prognostic effect in human cancers, including neuroblastoma (26), rhabdomyosarcoma, and EFT (27). It was recently reported to have clinical utility as a marker of minimal residual disease in the bone marrows of patients with stage 4 neuroblastoma (28). Both STEAP1 and CCND1 were also identified as EFT-specific genes in DNA microarray analyses (29). As to homeodomain-containing NK2 transcription factor NKX2-2, it was recently identified as a critical target gene of the EWS-FLI fusion protein and is necessary for oncogenic transformation in this tumor (30). The remaining two novel markers, PKP1, a plakophilin desmosomal protein found in squamous cell carcinomas of the oropharynx (31), and TMEM47, which is distantly related to peripheral myelin protein 22 (32), were not prognostic in our study, although they may possibly have clinical significance in other larger set of marrow samples. These last two genes have not been previously reported to be associated with EFT.

In many human malignancies, the presence of subclinical disease portends poor outcome and is often a major hurdle to cure. Our findings show the feasibility of using gene expression array approach to discover novel markers and rank their potentials by sensitivity and specificity assays. More importantly, by using well-annotated clinical samples (marrows that were histology negative from diagnostic patients treated at a single institution using the same protocol), their potential clinical relevance as markers of subclinical disease can be quickly explored. Our findings suggest that three of the five novel markers can be potential markers of subclinical disease. This is particularly relevant for localized EFT because fusion transcripts are below detection in the bone marrow samples at diagnosis. The availability of sensitive generic EFT markers without prior knowledge of the fusion type should facilitate the detection of subclinical metastatic EFT and complement clinical staging for risk group assessment and patient management. Although encouraging, our conclusions based on a relatively small number of patients will require further testing in large patient cohorts.

References
Novel Markers of Subclinical Disease for Ewing Family Tumors from Gene Expression Profiling

Irene Y. Cheung, Yi Feng, Karen Danis, et al.


Updated version  Access the most recent version of this article at: http://clincancerres.aacrjournals.org/content/13/23/6978

Cited articles  This article cites 32 articles, 14 of which you can access for free at: http://clincancerres.aacrjournals.org/content/13/23/6978.full#ref-list-1

Citing articles  This article has been cited by 12 HighWire-hosted articles. Access the articles at: http://clincancerres.aacrjournals.org/content/13/23/6978.full#related-urls

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.