

Identification, by systematic RNA sequencing, of novel candidate biomarkers and therapeutic targets in human soft tissue tumors

Anne E Sarver¹, Aaron L Sarver², Venugopal Thayanithy¹ and Subbaya Subramanian¹

Human sarcomas comprise a heterogeneous group of more than 50 subtypes broadly classified into two groups: bone and soft tissue sarcomas. Such heterogeneity and their relative rarity have made them challenging targets for classification, biomarker identification, and development of improved treatment strategies. In this study, we used RNA sequencing to analyze 35 primary human tissue samples representing 13 different sarcoma subtypes, along with benign schwannoma, and normal bone and muscle tissues. For each sarcoma subtype, we detected unique messenger RNA (mRNA) expression signatures, which we further subjected to bioinformatic functional analysis, upstream regulatory analysis, and microRNA (miRNA) targeting analysis. We found that, for each sarcoma subtype, significantly upregulated genes and their deduced upstream regulators included not only previously implicated known players but also novel candidates not previously reported to be associated with sarcoma. For example, the schwannoma samples were characterized by high expression of not only the known associated proteins *GFAP* and *GAP43* but also the novel player *GJB6*. Further, when we integrated our expression profiles with miRNA expression data from each sarcoma subtype, we were able to deduce potential key miRNA–gene regulator relationships for each. In the Ewing’s sarcoma and fibromatosis samples, two sarcomas where miR-182-5p is significantly downregulated, multiple predicted targets were significantly upregulated, including *HMCN1*, *NKX2-2*, *SCNN1G*, and *SOX2*. In conclusion, despite the small number of samples per sarcoma subtype, we were able to identify key known players; and, we discovered novel genes that may prove to be important in the molecular classification of sarcomas and in the development of novel treatments.

Laboratory Investigation (2015) 95, 1077–1088; doi:10.1038/labinvest.2015.80; published online 29 June 2015

Sarcomas, a heterogeneous group of tumors, account for about 200 000 cancers worldwide each year (about 1% of all human malignant tumors); however, they represent a disproportionately high number of all pediatric malignant tumors (15%).^{1,2} Sarcomas comprise more than 50 subtypes broadly classified into two groups: bone and soft tissue sarcomas.³ Even within a specific subtype, sarcomas are highly heterogeneous, making them diagnostically and therapeutically challenging. A few have a benign soft tumor tissue stage, such as schwannoma, which may or may not progress to malignant sarcoma. Several sarcoma subtypes are characterized by chromosomal translocations resulting in tumor-specific fusion transcripts or DNA copy number changes that are widely used as diagnostic markers.^{4,5} However, most sarcomas lack such a clearly elucidated genetic signature.

Recent advances have taken place in small RNA profiling, particularly for microRNAs (miRNAs), demonstrating that miRNAs have critical roles in the majority of canonical cellular signaling networks; dysregulation of miRNAs has now been implicated in many cancers, including breast cancer, colon cancer, gastric cancer, lung cancer, and sarcomas.^{6,7} Therefore, any comprehensive analysis of oncogenic or tumor suppressor signaling will need to include both messenger RNAs (mRNAs) and miRNAs. Understanding their roles in sarcoma will lead to new therapeutic targets and diagnostic biomarkers, opening the door to a greater understanding of the molecular mechanisms of all cancers.

In this study, we analyzed the mRNA expression profiles of 35 primary human tissue samples representing 13 different sarcoma subtypes, along with schwannoma, normal bone and

¹Division of Basic and Translational Research, Department of Surgery, University of Minnesota Medical School, University of Minnesota, Minneapolis, MN, USA and ²Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA

Correspondence: Dr S Subramanian, MS, PhD, Division of Basic and Translational Research, Department of Surgery, University of Minnesota Medical School, University of Minnesota, Moos Tower 11-212, 420 Delaware Street SE, MMC 195, Minneapolis, MN 55455, USA.

E-mail: subree@umn.edu

Received 19 December 2014; revised 17 April 2015; accepted 11 May 2015

muscle tissue. Our aim was to identify known and novel mRNAs that were significantly and uniquely upregulated in specific sarcoma subtypes. To generate the expression profiles, we used RNA-Sequencing (RNA-Seq), which allows for the generation of a superior expression data set without the concerns of probe-selection bias and other potential technical issues associated with microarray chip experiments. Then, we combined our RNA-Seq data with miRNA expression profiles taken from the Sarcoma microRNA Expression Database (S-MED), which is the most comprehensive sarcoma miRNA data set published to date; it contains miRNA expression profiles from more than 300 sarcoma primary tumor tissue samples, representing 22 different sarcoma subtypes.⁸ By integrating our RNA-Seq data with S-MED data, we aimed to identify key regulating miRNAs that could provide suitable upstream targets for more effective treatment of sarcomas.

MATERIALS AND METHODS

Tumor Samples

We collected 35 snap-frozen human patient tissue samples comprising 13 sarcoma subtypes, along with schwannoma, and normal bone and muscle tissues through the tissue procurement facility at the Masonic Cancer Center, University of Minnesota. To our knowledge, none of the normal tissue samples were patient matched. The samples were stored at -80°C until used. The University of Minnesota institutional review board approved this study. For each sarcoma subtype, we selected representative sarcoma samples per the patient's clinical diagnosis and our comprehensive miRNA expression data profiles. Then, for further RNA-Seq analysis, we selected samples representing each sarcoma subtype per their location in the dendrogram we generated (which was based on 300 sarcoma samples).⁸

mRNA RNA-Seq Expression Profiling

To perform high-resolution quantification of mRNA expression using RNA-Seq, we used at least $1\ \mu\text{g}$ of each pool of total RNA according to the *mirVana* miRNA Isolation Kit (Ambion, Austin, TX, USA). We sequenced mRNA libraries (four samples per lane, with 30–40 million paired-end reads in each sample) on a High Seq 2000 (Illumina, San Diego, CA, USA). To quality check raw-sequence FASTQ files, we used FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Then, to align those FASTQ files to the Human UCSC hg19 Assembly genome, we used the TopHat2 algorithm.⁹ Next, to quantify the aligned reads against the reference human genome hg19 transcript annotations, we used Cufflinks and Cuffdiff, according to the standard Fragments Per Kilobase of transcript per Million mapped reads (FPKM) methods.¹⁰

Data Analyses

For statistical analysis and clustering, we used the Partek Genomics Suite software package (Partek, St Louis, MO, USA). We performed a series of two-group *t*-tests to compare each sarcoma subtype with at least two unique patient

samples *versus* all remaining samples (sarcoma and normal tissue). Before comparison analysis and clustering, we filtered extremely low and invariant genes out of the data sets. Significance cutoff filters, unless otherwise specified, were set at $P < 0.001$ and an expression change of at least threefold. No transcripts in our chondrosarcoma analysis passed both significance filters, so we did not perform additional analysis on that gene set. To provide a more comprehensive reference pool, we included osteosarcoma, leiomyosarcoma, and rhabdomyosarcoma samples, but did not further analyze them, given specific concerns about heterogeneity and small sample size. Previously, our group already reported on a larger study that included gene and miRNA network analysis of osteosarcoma, MPNST, neurofibroma, and rhabdomyosarcoma.^{11–13}

For functional analysis, upstream regulator analysis, and miRNA targeting analysis, we used Ingenuity Pathway Analysis (IPA) software (Qiagen, Redwood City, CA, USA). IPA uses a right-tailed Fisher's exact test to calculate a *P*-value corresponding to the probability that a biologic function not relevant to the input data set is falsely identified as relevant. A Benjamini–Hochberg false discovery rate of 0.05 was used to correct such *P*-values. Network diagrams were generated by combining significant results derived from functional analysis, upstream regulator analysis, and miRNA targeting analysis for each sarcoma subtype.

RESULTS

Novel Overexpressed Gene Candidates

The findings of our mRNA expression analysis across the 35 tissue samples after unsupervised hierarchical clustering are highlighted in the heatmap (Figure 1a). Visual inspection easily identified some unique 'hot spots' (Figure 1b). To identify computationally all mRNAs that were significantly upregulated in each subtype (in comparison with the rest of the samples), we used a two-group *t*-test. Significantly upregulated genes and their deduced upstream regulators included both known players, and, interestingly, one or more novel candidates that have not been previously reported to be associated with sarcoma. The highly overexpressed genes in each sarcoma subtype, per their expression ratios, were as follows:

Angiosarcoma

We found that 173 genes were significantly (by at least threefold, $P < 0.001$) upregulated in our angiosarcoma samples, as compared with our non-angiosarcoma samples (Table 1 and Supplementary Table 1). The two different angiosarcoma patient samples showed distinct patterns of gene expression—but with some common elements, most notably *FLT4*. In addition to genes previously associated with cancer formation and progression, such as *CTLA4*, *CCL1*, *MADCAM1*, and *NTS*, we identified significantly upregulated genes not previously associated with angiosarcoma such as *HTN3*, *GPR182*, and *SCGB3A1* (*HIN1*). The most significantly upregulated transcript was for *D21S2088E*, a

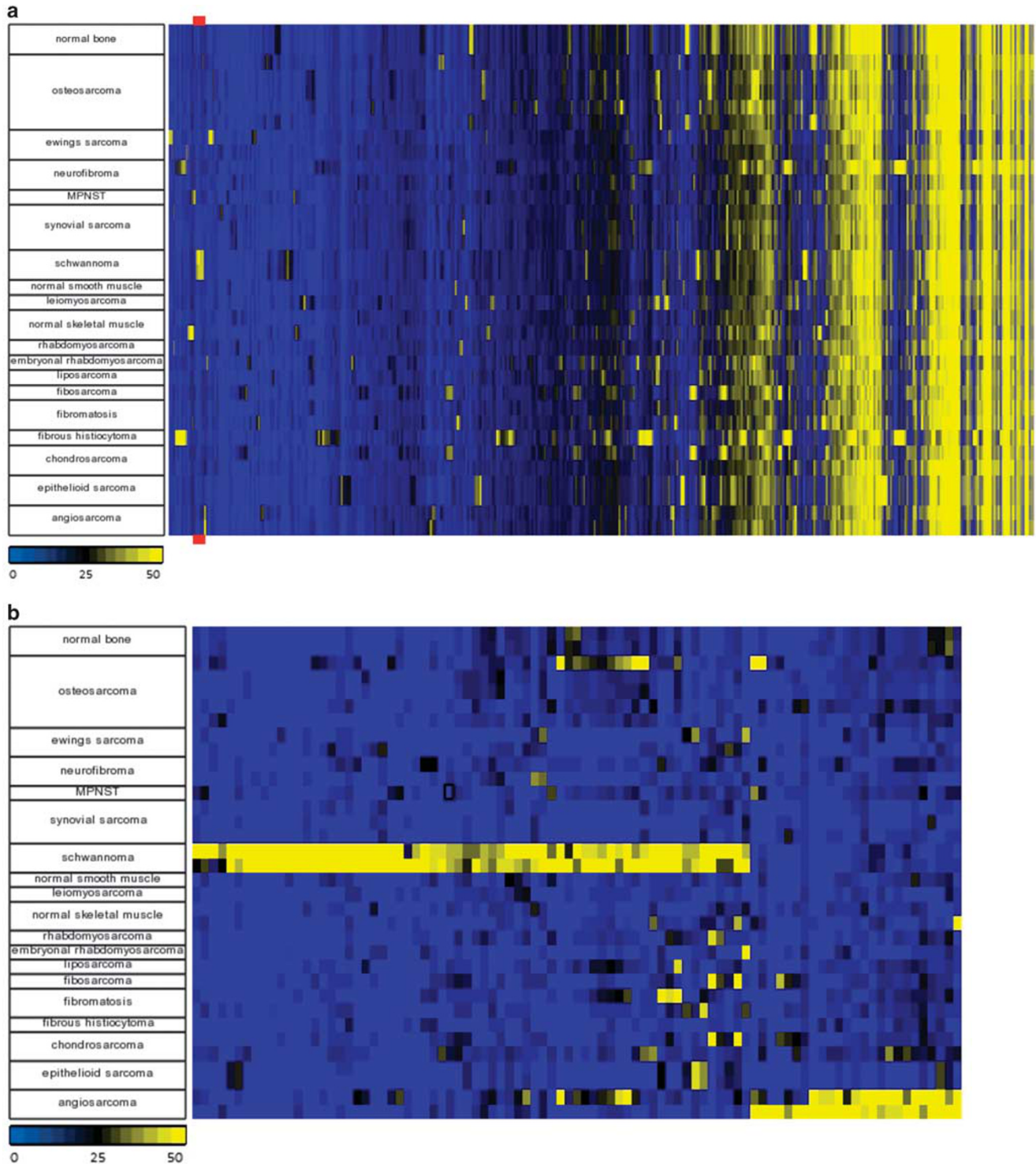


Figure 1 (a) Heatmap shows mRNA expression across the 35 tissue samples after unsupervised hierarchical clustering of Fragments Per Kilobase of transcript per Million mapped reads (FPKM) data. Red bar indicates region expanded in (b) showing representative clusters of differentially overexpressed genes in schwannoma and angiosarcoma samples.

newly annotated long noncoding RNA (lncRNA) initially described by Reymond *et al.*¹⁴ as a transcript with no obvious open reading frame specifically expressed only in the testes, uterus, fetal brain, and spinal cord.

The most significantly associated biologic functions for this gene set were typical hallmarks of cancer processes, including increases in cell migration, angiogenesis, cell invasion, and inflammatory response. The genes in this set shared multiple

Table 1 Top 10 uniquely upregulated genes in our sarcoma samples (genes in bold have not been previously associated with sarcoma; full gene lists and mean expression values are available in Supplementary Tables 1–6)

Sarcoma	Gene symbol	Ratio
Angiosarcoma	<i>D21S2088E</i>	4566
	<i>CCL1</i>	974
	<i>HTN3</i>	794
	<i>GPR182</i>	419
	<i>MADCAM1</i>	320
	<i>FLT4</i>	252
	<i>SCGB3A1</i>	238
	<i>NTS</i>	76
	<i>MMP12</i>	61
	<i>CTLA4</i>	43
Epithelioid sarcoma	<i>CST5</i>	3467
	<i>ARHGAP36</i>	1822
	<i>CDH10</i>	1642
	<i>TMEM215</i>	598
	<i>IGSF1</i>	348
	<i>ZCCHC12</i>	288
	<i>CST2</i>	211
	<i>GAL3ST3</i>	189
	<i>UNC5D</i>	181
	<i>FOXL2NB</i>	149
Ewing's sarcoma	<i>PCGEM1</i>	12063
	<i>SLC17A8</i>	4868
	<i>SCNN1G</i>	1037
	<i>C6orf15</i>	578
	<i>PGLYRP2</i>	554
	<i>CALCA</i>	546
	<i>C1QL2</i>	454
	<i>LIPI</i>	275
	<i>MAPT-AS1</i>	258
	<i>NKX2-2</i>	229
Fibromatosis	<i>SERPINB11</i>	981
	<i>LAMP5</i>	76
	<i>ADAM12</i>	52
	<i>POSTN</i>	43
	<i>ADAMTS12</i>	40
	<i>XPNPEP2</i>	35
	<i>PLA2G2A</i>	35
	<i>F2RL1</i>	30
	<i>ZNF469</i>	23
	<i>HMCN1</i>	23

Table 1 Continued

Sarcoma	Gene symbol	Ratio
Schwannoma	<i>GJB6</i>	2026
	<i>GFAP</i>	1124
	<i>GAP43</i>	340
	<i>ANGPTL7</i>	210
	<i>SOX10</i>	205
	<i>GRIA4</i>	174
	<i>SERPINA5</i>	134
	<i>ITGB8</i>	130
Synovial sarcoma	<i>S100B</i>	128
	<i>FABP7</i>	121
	<i>C5orf58</i>	122
	<i>SPATA22</i>	113
	<i>ZIC1</i>	96
	<i>FGF19</i>	39
	<i>GPR126</i>	31
	<i>ZIC4</i>	29
	<i>FAM163B</i>	29
	<i>GFRA2</i>	22
<i>DLX6-AS1</i>	20	
<i>HOXC11</i>	19	

common upstream transcriptional regulators, most significantly TNF, VEGFA, and NKX2-3.

Epithelioid sarcoma

Our analysis identified 343 genes significantly (by at least threefold, $P < 0.001$) upregulated in the epithelioid sarcoma samples versus the non-epithelioid sarcoma samples (Table 1 and Supplementary Table 2). All top 10 differentially expressed genes, including *cystatin D* (*CST5*), *ARHGAP36*, *Cadherin 10* (*CDH10*), *TMEM215*, and *IGSF1*, have not been previously associated with epithelioid sarcoma but many have been previously associated with other cancers.

Highly associated functions for this set of upregulated genes include increased cell migration and movement, increased cell proliferation/decreased cell death, and increased angiogenesis, all hallmarks of cancer. Common upstream regulators for this set of genes include NFE2L2, TGFB1, KMT2D, and PRL. *PRL* itself was significantly upregulated (by about 50-fold) in the epithelioid sarcoma samples.

Ewing's sarcoma

This comparison identified 77 genes significantly (≥ 3 -fold, P -value < 0.001) upregulated in the Ewing's sarcoma samples versus the non-Ewing's sarcoma samples (Table 1 and Supplementary Table 3). The most significantly differentially

regulated transcripts included two lncRNAs, *PCGEM1* and *MAPT-AS1*. *LIPI* and *NKX2-2* upregulation have previously been reported to be highly specific markers for Ewing's sarcoma, whereas the remaining transcripts have not been previously reported to be associated with any sarcoma subtype.

Significantly associated biologic functions include increases in cell proliferation, cell stimulation, mobilization of Ca^{2+} , and, interestingly, skin development. The most significant common upstream regulators were *KLF4*, *BDNF*, and *SOX2*. *SOX2* was upregulated (by about 20-fold) in Ewing's sarcoma compared with the non-Ewing's samples.

Fibromatosis

Our comparative analysis identified 82 genes significantly (by at least threefold, $P < 0.001$) upregulated in the fibromatosis samples (Table 1 and Supplementary Table 4). Of the top 10 upregulated genes, three of them—*ADAM12*, *POSTN*, and *PLA2G2A*—have been previously associated with a different sarcoma subtype. *SERPINB11* was the most highly differentially upregulated gene by a very large margin in this particular comparison analysis.

The most significantly enriched set of genes in this set are those involved in cellular adhesion and attachment, specifically metalloproteinases and their regulators. Highly associated biologic functions include regulation of fibroblast attachment, connective tissue disorders, tumor cell migration, and invasion. The most significant common upstream regulators were the *TGFB1*, *TGFB3*, and *IL1B* regulatory network, along with *TP53*, *HTT*, and *TNF*.

Schwannoma

In this comparison, we found a total of 357 genes significantly upregulated (by at least threefold, $P < 0.001$) in the schwannoma samples (Table 1 and Supplementary Table 5). Of the top five upregulated genes, three of them—*GFAP*, *GAP43*, and *SOX10*—have previously been experimentally confirmed to be overexpressed in schwannomas. Novel genes included *GJB6*, *ANGPTL7*, and *GRIA4*.

We found strong activating associations with many different immune and inflammatory response processes, such as leukocyte activation. Interestingly, one of the top associated disease processes was rheumatic disease, including strong activating associations with both arthropathy and arthritis. Typical cancer hallmark processes such as cell movement, recruitment, and proliferation were also highly associated with this gene set. The top common upstream regulators include *TNF*, *TGFB1*, *IFNG*, *IL1B*, *EGF*, and *IL6*. Both *IL1B* and *IL6* were significantly overexpressed (by about 20-fold) in our schwannoma samples.

Synovial sarcoma

Our analysis identified 38 genes significantly (by at least threefold, $P < 0.001$) upregulated in the synovial sarcoma samples versus the non-synovial sarcoma samples (Table 1 and Supplementary Table 6). The most significantly

differentially upregulated gene was one that is still only annotated as an open reading frame, *C5orf58*, but which is predicted by RefSeq to be protein coding. Of the 38 significantly upregulated genes, five were homeobox transcription regulators (*HOXC4*, *HOXC9*, *HOXC10*, *HOXC11*, and *HOXC12*). *HOTAIR*, the HOX transcript antisense RNA, was also significantly upregulated (by about 13-fold).

The most significantly associated biologic functions included neuronal, synapse, and epithelial cell development. The *HOXC* genes are associated with numerous skeletal system developmental processes. *FGF11*, *FGF19*, and *FZD10* are all members of the epithelial-mesenchymal transition canonical pathway. Common transcriptional regulators included *DLX1*, *DLX2*, *SOX2*, and *FGF8*. *DLX1* was significantly overexpressed (by about 10-fold) in the synovial sarcoma samples.

Upstream Transcription Factor Analysis

In addition to looking at the significant upstream regulators in each individual sarcoma data set, we also examined whether there were common upstream transcription factors that could be relevant across multiple sarcoma subtypes. In this analysis, the 10 most upregulated genes from each sarcoma were grouped and used as a superset for the upstream analysis. The top predicted transcription factors were well-known regulators of cellular growth and proliferation such as *TNF*, *IFNG*, *TGFB1*, and *IL1B* (Table 2).

Integrated miRNA Targeting Analysis

As signaling networks are regulated not only by genes but also by miRNAs, we performed an integrated miRNA targeting analysis using miRNA expression data from SMED. In general, miRNAs act by downregulating their target genes, thus one would expect that when the target genes are upregulated the targeting miRNAs are downregulated. Therefore, we looked for miRNAs predicted to target the upregulated genes that were significantly downregulated (by at least twofold, $P < 0.05$) in the sarcoma data set from S-MED (Table 3).

DISCUSSION

Sarcomas are a diverse collection of rare cancers with limited resources for research, both financially and with respect to availability of human patient tissues. In addition, many sarcoma subtypes themselves are highly heterogeneous and may have both benign and malignant stages. These obstacles provide a significant challenge in efforts to identify diagnostic biomarkers and therapeutic targets for these poorly understood tumors.

Nonetheless, our study demonstrates that, even with a small set of diverse samples, it is possible to perform a useful comparative analysis. We identified novel, uniquely expressed RNA molecules. Our findings should inform future pre-clinical studies, which may eventually lead to novel therapeutic targets and treatments. Definitive validation will

Table 2 Predicted common upstream regulators

Gene symbol	Predicted Targets
<i>TNF</i>	<i>CALCA, CCL1, CTLA4, F2RL1, GFAP, GFRA2, GPR126, MADCAM1, MMP12, NKX2-2, PLA2G2A, POSTN, SCNN1G</i>
<i>IFNG</i>	<i>CALCA, CCL1, F2RL1, FLT4, GFAP, MMP12, NKX2-2, PLA2G2A, SCGB3A1, SCNN1G, SERPINA5</i>
<i>IL1B</i>	<i>CALCA, CCL1, F2RL1, ITGB8, MMP12, NKX2-2, PLA2G2A, POSTN, S100B, SCNN1G</i>
<i>TGFB1</i>	<i>ADAM12, ADAMTS12, CALCA, CTLA4, F2RL1, GFAP, MMP12, POSTN, XPNPEP2</i>
<i>JUN</i>	<i>FABP7, GAP43, ITGB8, MMP12, NTS, PGLYRP2</i>
<i>HTT</i>	<i>FABP7, GAP43, GFAP, NTS, SCGB3A1</i>
<i>IL4</i>	<i>CCL1, F2RL1, MMP12, POSTN, SCGB3A1</i>
<i>IL6</i>	<i>CALCA, GFAP, MMP12, PLA2G2A, XPNPEP2</i>
<i>BDNF</i>	<i>GAP43, GFAP, LAMP5, SLC17A8</i>
<i>IGF1</i>	<i>GAP43, GFAP, PLA2G2A, S100B</i>
<i>LEP</i>	<i>GAP43, GFAP, MMP12, NTS</i>
<i>STAT3</i>	<i>CTLA4, GFAP, MMP12, PLA2G2A</i>

require larger association studies and functional experiments. For each of the upregulated gene sets that we focused on, we looked for evidence that they were being upregulated either due to an increase in transcription factor activity or a downregulation of a common regulatory miRNA.

In each analyzed sarcoma subtype, we were able to identify both known sarcoma players and novel candidates—beyond known translocation events—for potential therapeutic targets or diagnostic biomarkers. To better understand the transcription regulation occurring in each sarcoma subtype, we looked for common transcriptional regulators that were predicted to be activated for each sarcoma data set. Each set of upregulated genes shared a set of common upstream regulators, some of which were significantly associated with multiple sarcoma subtypes. Activation of transcription factors can occur by multiple mechanisms. We observed, in several instances, both the identification of a common transcription factor activity and significant upregulation of the transcription factor itself. Integration of upstream regulators, both protein transcription factors and predicted targeting miRNAs, may provide a window into critical oncogenic networks that could be targeted therapeutically (Figure 2).

Known players identified in our analysis included previously validated biomarkers for specific sarcomas. For example, the *NKX2-2* gene is a known target of the Ewing's-specific fusion protein EWS-FLI-1 and an established immunohistologic marker for Ewing's sarcoma cells.¹⁵ *NKX2-2* was significantly upregulated in our Ewing's sarcoma samples, along with *LIPI*, which was recently identified as a

highly specific antigen abnormally expressed in Ewing's sarcomas (and is therefore a strong candidate for immunotherapy: it could be used for targeting therapies specifically to sarcoma cells).¹⁶ *FLT4*, a marker for lymphatic vessels, was previously reported to be upregulated in sarcoma, including angiosarcoma, as we also found in our study.^{17,18} Previous studies have reported amplification of *FLT4* and *MYC* in radiation-induced secondary angiosarcomas, but not in other radiation-associated sarcomas. This study also found that *FLT4* expression was 5.3-fold higher in secondary radiation-induced angiosarcomas compared with primary angiosarcomas.^{17,18} Our angiosarcoma samples were derived from spontaneous sarcomas (not radiation induced) but still showed significantly higher *FLT4* expression compared with the other samples. It should also be noted that these tumor samples may have included additional cells such as activated tumor-invading lymphocytes, which could potentially account for some of the observed *CTLA4* expression.¹⁹

In our study, we additionally identified other previously established markers such as *GFAP*, *GAP43*, *S100B*, *FABP7*, and the *SOX10* transcription factor. In schwannoma, *GFAP* is a known marker for nerve sheath differentiation and has been established as a biomarker for certain nerve sheath tumors.^{20–22} *GAP43*, *S100B*, *FABP7*, and the *SOX10* transcription factors are all involved in neuronal development and were previously reported to be positively expressed in nerve-derived tumors such as schwannomas.^{23–27} Our reidentification of such known biomarkers adds to our confidence in our study's novel differential expression results.

Many of the top upregulated genes were previously reported to be associated with various cancers, but ours is the first report associating them with sarcomas. *CCL1*, the natural cognate ligand of *CCR8* that we identified in our angiosarcoma comparison, is known to be involved in cancer-related inflammation, malignant progression, and control of tumor cell entry into the lymph node.²⁸ Just recently, *ANGPTL7* was reported to be a hypoxia-driven proangiogenic factor upregulated in cancer cells that can induce malignant behavior in human prostate cancer cells: it could be an interesting candidate to profile across many schwannomas, to determine if its expression levels correlate with malignant progression.^{29,30} Many of the upregulated genes that we identified in our study have been functionally characterized as part of the inflammatory response, a hallmark enabling characteristic that can support multiple cancer capabilities.³¹ *MADCAM1*, another highly upregulated gene in our angiosarcoma samples, is associated with the onset and progression of multiple gastric inflammatory diseases and is also upregulated in melanoma cells.^{32,33}

Recent epigenetic studies have shown that just as expression patterns can be highly instructive, so, too, can methylation patterns. Although not examined in our study, several of the top results in our comparisons have been previously reported to be abnormally methylated in a variety

Table 3 Predicted miRNA regulators of upregulated genes (only miRNAs predicted to target three or more significantly upregulated genes are listed; for angiosarcoma, only the top 10 miRNAs are listed)

Sarcoma	miRNA	Predicted target genes
Angiosarcoma	miR-541-3p/ miR-654-5p	<i>ADAMTS7, ANO9, ARHGAP29, ATP11A, CHST15, Cxorf36, FAM163A, FGD5, FLT4, GRAP, HID1, JUP, KCTD17, LIMD2, MAF, MGST2, PLXND1, PPM1F, RAMP2, ROBO4, 100A16, SCAMP5, SCGB3A1, SH2B3, SLC7A1, TSPAN15</i>
	miR-128-3p	<i>ADAM19, APOLD1, CDH5, DLL4, DTX1, DUSP5, FAM126A, GIMAP8, MESDC1, MYCT1, NETO2, PEAK1, PLXND1, RELN, SH2D3C, SHE, SLC7A1, SOX7, TAL1</i>
	miR-1-3p	<i>ARHGAP29, CREM, EFN2, ETS1, FAM107B, FAM126A, GIMAP6, KANK3, MAF, MAP4K2, NETO2, NOL4L, NR1H3, PEAK1, PPP2R5A, TAL1</i>
	miR-485-5p	<i>ADAMTS4, ARHGAP25, ARHGEF15, CARD8, CCL21, Cxorf36, FAM163A, GRPEL2, IKBKE, KIAA1671, LIMD2, NID1, PEAR1, PPM1F, SLC41A1, TIE1</i>
	miR-370-3p	<i>ADAM19, AFAP1L1, ATP11A, CCL21, CHST15, CHST7, GRB10, GYPC, KDR, MAGI1, PEAR1, PODXL, SHE, SLC7A1</i>
	miR-381-3p	<i>ARHGAP29, B3GNT2, COL4A1, DLL4, EFN2, ETS1, FAM107B, FAM126A, FSCN1, JUP, MAGI1, RAPGEF5, SH2B3, SLC7A1</i>
	miR-193a-3p	<i>ARHGEF15, CDH5, EFN2, ETS1, FLI1, IKBKE, N4BP3, NOL4L, PPM1F, PPP1R16B, RAPGEF5, TNFAIP1, UACA</i>
	miR-377-3p	<i>ADAMTS4, APOLD1, ARHGAP29, ARRB1, COL4A1, ETS1, LIMK2, NOL4L, NOVA2, NTS, PPP1R16B, SHE, TAL1</i>
	miR-153-3p	<i>ADAM19, CD69, CD93, CHST15, CREM, DLL4, EFN2, GNG11, MAGI1, NOL4L, SCAMP5, SH2B3</i>
	miR-504-5p	<i>AKAP2, CCM2L, CDH5, F2RL3, GIMAP6, LIMD2, LYVE1, MYCT1, RPP25, SLC7A1, UACA, UBTD1</i>
	Epithelioid sarcoma	miR-344d-3p
miR-136-5p		<i>ABI3BP, ARHGEF4, B3GALNT1, CPD, CTTNBP2, GDF6, IGSF1, KDM7A, MAMDC2, MDM1, MOSPD1, MTPN, NR4A3, PLEKHA5, RIMKLB, SLC16A9, SOCS7, SYTL4, UNC5D, ZNF516</i>
miR-376a-3p		<i>ABI3BP, ARHGAP26, ARMC10, COPS7A, EFCAB14, FAM131B, HELLS, MTMR12, MYOCD, PLEKHA5, PUS7, RAPGEFL1, SALL1, SATB1, TMEM100, SLC16A1</i>
miR-196a-5p		<i>AKNA, CDKN1B, CPD, ERG, EYA4, FAM110B, FOXP2, GDNF, MAMDC2, SCUBE2, ST6GAL2, TGFB3, TUB, ZNF516</i>
miR-542-5p		<i>ARHGEF4, CA12, CMTM4, CUL4B, EPHB4, ERMP1, FBLN5, MARVELD1, PEX5, PLXNB1, SOCS2, TMEM229B, VIPR2</i>
miR-654-3p		<i>BTG1, CDH10, ESYT2, FAM115A, FOXL2NB, KCNIP4, MBOAT1, MYLIP, NT5E, SLC16A1, ST6GAL2, TUB</i>
miR-452-5p		<i>ARMC10, CDKN1B, CST2, CST5, EPS8, IMM2L, MBOAT1, RPS6KA3, SLC18A2, ST6GAL2, ZNF516</i>
Ewing's sarcoma	miR-204-5p	<i>ARX, BCL11B, CEACAM6, DKK2, IRF6, SLC17A7, ZDHHC21</i>
	miR-146a-5p	<i>KRT80, PGLYRP2, RBM11, SYNPR</i>
	miR-938	<i>CAMK1G, FAM84B, GPR64, PRKCB</i>
	miR-574-5p	<i>CAMK1G, DLG2, SOX2</i>
Fibromatosis	miR-135a-5p	<i>ADAM12, ADRA2A, ANGPTL2, CERCAM, COL5A1, EDIL3, LRRN1, MAN1A1, MAN2A1, MXRA5, PSD3, RUNX1, THBS2, VCAN, XPNPEP2</i>
	miR-141-3p	<i>ATRN, CCDC80, FAP, GOLM1, LGR4, LOXL1, RUNX1, VCAN</i>
	miR-149-5p	<i>ADAM12, ATRN, FAM198A, FBLN1, NAV1, PSD3, STAC, XPNPEP2</i>
	miR-182-5p	<i>CD55, COL5A1, FBN1, GXYLT2, HMCN1, SH3PXD2B, THBS2, WISP1</i>
	miR-130a-3p	<i>ADAM12, ATRN, CERCAM, COL6A3, MAFB, SH3D19, TIMP2</i>
	miR-17-5p	<i>ADAMTS5, CERCAM, GALNT10, LAMP5, MMP2, RUNX1, TIMP2</i>
	miR-33-5p	<i>ADAM12, ADRA2A, GALNT10, PIEZO2, POSTN, STAC, VCAN</i>
	miR-92a-3p	<i>ANGPTL2, COL5A1, FBN1, MAN2A1, SH3D19, VSNL1, ZNF469</i>

Table 3 Continued

Sarcoma	miRNA	Predicted target genes	
Sarcoma	miR-203a-3p	ADAMTSS, GALNTT10, MAFB, SPARC, VSNL1	
	miR-183-5p	ATRN, COL12A1, PSD3, SH3BD19	
	miR-200b-3p	LRP1, RUNX1, TIMP2, TMOD3	
Schwannoma	miR-563	ABLUM3, ALAD, ATP1B2, CD37, CER1, CMTM5, CNTNAP2, CORO2B, DDR1, EXTL1, GAS7, IL34, ITGB4, JUNB, NAT8L, NCDN, NR4A1, NRXN3, PLD4, RASGEF1C, S100B, SHISA9, SIMM5, SOX10, SPOCK2, ST6GAL1, SYNGR1, TMBIM1, TRAF1, TRIM58, TRIM9, WBSR17, ZCCHC24	
	miR-582-5p	ARHGAP12, CLEC7A, EPB41L3, FGF2, GPR34, SHC4, ZNRF2	
	miR-483-3p	ANGPTL7, CLDN11, LY86, QKI, SPATA13, TANC1	
	miR-1290	GPR176, GPR34, GRIA4, KLHDC8A, SORCS1, SORL1	
	miR-483-5p	CHST3, FN3K, TMCC2	
	Synovial sarcoma	miR-330-5p	BNC2, BTBD11, CRRH1, DLX1, FGF11, HOXC10, NAV3
		miR-150-5p	C5orf58, CYP4X1, HOXC9, RPRM
miR-185-5p		DLX6, FAM163B, SIX1	
miR-221-3p		CYP4X1, HOXC10, ZIC1	

of cancers. In particular, *SCGBA31* (*HIN-1*) hypermethylation has been reported in a diverse set of cancers and has been shown to activate AKT signaling.³⁴ *GRIA4*, which we found significantly upregulated in our schwannoma samples, is a target of the polycomb repressive complex 2 and is hypermethylated in follicular lymphoma.³⁵

The top mRNA in our schwannoma comparison, *GJB6* (connexin 30), whose primary pathologic association is with hearing loss, has not been previously reported in schwannoma tumors; however, some clinical reports have indicated a connection between hearing loss and schwannoma.³⁶ *GJB6* has been reported to be dysregulated in multiple other cancers but with differing directions of change that could improve its usefulness as a diagnostic marker.^{37–39} Interestingly, of the 10 most differentially upregulated genes in fibrosarcoma, three of them—*ADAM12*, *POSTN*, and *PLA2G2A*—have previously been identified in osteosarcoma, which is notoriously heterogeneous.^{40–42} The ZIC transcription factor family, two members of which were significantly upregulated in our synovial sarcoma samples, is normally involved in brain and neuronal development and is overexpressed in liposarcoma and meningioma relative to normal tissue.^{43,44}

Some of the top genes in our analysis, such as *cystatin D* (*CST5*) and *ARHGAP36* in epithelioid sarcoma, have been previously reported to be dysregulated in cancer cells.^{45,46} Yet, very little information has been published on the expression pattern or biologic function of others that we identified. *TMEM215* is a transmembrane protein and therefore likely to be involved in cell signaling. *SPATA22* was only recently identified as being required for meiosis in mammals and might have a proproliferative role in synovial sarcoma.⁴⁷ Very little has been reported on the orphan G-protein-coupled receptor GPR182, other than that it is consistently upregulated in persistently propagated tumor-specific endothelial cells, and now in angiosarcoma as well.⁴⁸ To date, no published works have described the function of *FOXL2NB*, *FAM163B*, or several of the other transcripts that still carry just the open reading frame designation, but are predicted to be protein coding. In our synovial sarcoma comparison, *SERPINB11* was the most highly differentially upregulated transcript by more than an order of magnitude, primarily because of its near-absent expression in our non-synovial sarcoma samples. Of note, *SERPINB11* is classified as a polymorphic gene/pseudogene coding for a noninhibitory serine protease inhibitor and could be functioning as a serine protease inhibitor agonist.⁴⁹ *HTN3*, a member of the antimicrobial histatin family, and *ZNF469*, traditionally associated with brittle cornea syndrome and keratoconus, have established functions or pathologic associations not immediately connected to cancer, inflammation, or cell growth.⁵⁰ These apparently non-cancer-related genes could still be important genetic markers if they are signaling a translocation event or activation of a related signaling pathway, or they might have a yet-unrecognized role in sarcoma development or progression.

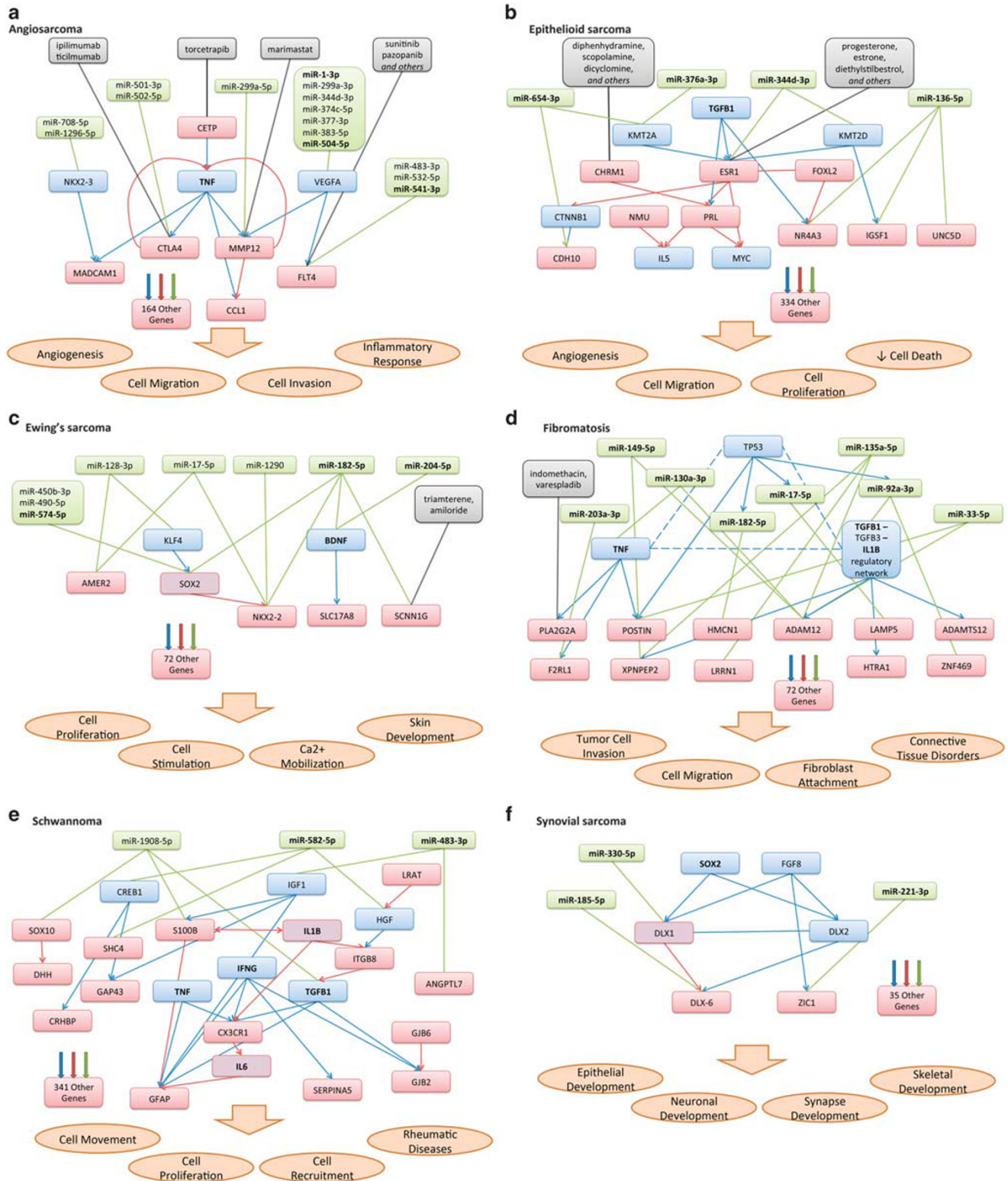


Figure 2 Interaction networks involving significantly upregulated genes in our (a) angiosarcoma, (b) epithelioid sarcoma, (c) Ewing's sarcoma, (d) fibromatosis, (e) schwannoma, and (f) synovial sarcoma datasets. Red nodes are upregulated genes. Blue nodes are significantly enriched upstream activating transcriptional regulators in each of our sarcoma datasets; bold indicates transcription factors found to be significant in multiple sarcoma datasets; bold indicates transcription factors found to be significant in multiple sarcoma datasets; red and blue striped nodes are identified upstream transcriptional regulators that were also significantly upregulated in the sarcoma dataset. Green nodes are miRNAs predicted to target key upregulated genes and/or transcriptional regulators that were significantly downregulated in sarcoma; bold indicates miRNAs predicted to target multiple upregulated genes in the sarcoma gene set. Gray nodes are compounds/agonists that effect the upregulated genes in this network. Orange nodes are key biologic and disease processes associated with the total set of upregulated genes for each sarcoma subtype.

Recently, researchers have begun to explore the biologic significance and differential expression of noncoding RNAs beyond miRNAs in cancer development and progression.⁵¹ In our study, a number of the top hits across multiple sarcoma subtypes were lncRNAs, including *D21S2088E* in angiosarcoma and *PCGEM1* in Ewing's sarcoma, both of which were the most significantly upregulated transcripts relative to all of our other samples. *PCGEM1* is normally expressed only in prostate tissue; significant upregulation of *PCGEM1* is an established biomarker for prostate cancer.⁵² This potential overlap between Ewing's sarcoma and prostate cancer renews the speculation that steroid receptors, including the androgen and progesterone receptors, may have a significant role in the development and progression of malignant skeletal tumors.⁵³ The protein FUS/Translocated in LipoSarcoma (FUS/TLS) binds to the androgen receptor and mediates androgen-dependent cell-cycle progression and prostate cancer growth.^{54,55} A loss of FUS/TLS results in decreased cellular proliferation, impaired RNA processing, and increased expression of the closely related Ewing's sarcoma protein EWS.⁵⁶ Indeed, the transactivation domains of the TET family of RNA-binding proteins (such as FUS and EWS) may be interchangeable and FUS/ERG fusions have been reported in Ewing's sarcoma patients.⁵⁷

We also observed two long antisense RNAs of unknown biologic significance, *MAPT-AS1* in Ewing's sarcoma and *DLX6-AS1* in synovial sarcoma. Expression of the *MAPT-AS1* cis-antisense RNA correlates with expression of the *MAPT* gene and the progesterone receptor (*PGR* and *TMEM26*) and has been proposed as a predictive and prognostic marker for breast cancer.⁵⁸ All of those lncRNAs, as well as other transcripts of unknown biologic function, could prove to be diagnostically useful biomarker panel members, even if their functional roles are not fully understood.

Transcripts that are significantly overexpressed in a particular sarcoma are not only useful as diagnostic biomarkers but also could indicate which chemotherapy regimen would be most efficacious for a particular patient. For example, our angiosarcoma samples showed significant overexpression of *FLT4* and *CTLA4*, both of which are established therapeutic targets in other contexts. Sorafenib, efficacious against desmoid tumors, could be therapeutically beneficial in other *FLT4*-expressing sarcomas.⁵⁹ *CTLA4*, a critical immunoregulatory gene, helps maintain tolerance to self-antigens by elevating the T-cell activation threshold.⁶⁰ An anti-*CTLA4* antibody approved to treat melanoma, ipilimumab, functions in part by releasing this immunity checkpoint inhibitor, allowing the patient's cytotoxic T lymphocytes to mount a more robust antitumor immune response—a response that could also be used in sarcoma therapy.⁶¹

Moreover, three of the other sarcoma subtypes in our study also contained a known drug target in their top 10 differentially upregulated gene lists: *PLA2G2A*, upregulated in

our fibromatosis samples, is a member of the prostaglandin family and can be downregulated by indomethacin, a nonsteroidal anti-inflammatory drug that targets COX1 and COX2, which are involved in prostaglandin synthesis.⁶² *GRIA4*, upregulated in our schwannoma samples, is one of the AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionate)-sensitive glutamate receptors; it could be regulated by AMPA agonists such as the drug tezampanel.⁶³ *HOXC11*, upregulated in our synovial sarcoma samples, is a marker of poor prognosis in patients with renal clear-cell carcinoma.⁶⁴ Treatment with the Src/Abl inhibitor dasatinib prevents recruitment of the HOXC11 transcription factor to S100 β , an established biomarker of cutaneous melanoma progression.⁶⁵ Patients with sarcomas with elevated levels of any of those genes could derive some additional benefit by adding a targeted therapeutic medication to their current treatment regimen.

Misdiagnosis of sarcoma subtype is still a significant barrier to effective treatment. Identifying and validating biomarkers for differential diagnosis will help ensure that patients receive the most effective treatments in a timely manner. Distinguishing between benign and malignant tumors will also be highly useful in determining the optimal treatment. As with all therapeutics, the potential benefits of any targeted drug regimen will need to be balanced against potential adverse events. Biomarkers that could predict sensitivity/resistance to various chemotherapy drugs would also be very valuable, particularly if they were broadly applicable to many tumor types.

We acknowledge that, given the heterogeneous nature of sarcomas (even within subtypes), the novel candidate genes we identified might not be relatively overexpressed across all samples. The number of samples we subjected to RNA-Seq was too small to identify a comprehensive biomarker signature for each sarcoma subtype, yet in our comparison analysis, we were able to identify both known sarcoma-specific markers and novel candidates. By integrating our RNA-Seq results with our previous miRNA profiling work covering those samples, we were able to identify deregulated miRNA-gene networks in each sarcoma subtype and the key predicted regulatory miRNAs. The overexpressed genes predicted to be targeted by those miRNAs represent multiple signaling pathways, supporting the notion that a single miRNA can regulate multiple pathways.

Such critical regulatory miRNAs should be further investigated to gain deeper insights into sarcoma pathobiology, which could be relevant in other cancers as well—particularly as many of the upregulated genes identified in our study have been strongly associated with higher-frequency cancers (such as breast, lung, and prostate cancer). In conclusion, our analysis is a proof-of-concept study demonstrating that high-throughput sequencing approaches and bioinformatic analysis of limited representative sarcoma samples can successfully identify novel candidates for biomarker and therapeutic targets.

Supplementary Information accompanies the paper on the Laboratory Investigation website (<http://www.laboratoryinvestigation.org>)

ACKNOWLEDGMENTS

We thank Dr Lihua Li for assistance in sarcoma tissue and RNA preparation. We also thank the Wykoff Sarcoma Foundation, the Children's Cancer Research Fund, and the American Cancer Society (RSG-13-381-01) for grant funding (to SS).

DISCLOSURE/CONFLICT OF INTEREST

The authors declare no conflict of interest.

- Taylor BS, Barretina J, Maki RG *et al*. Advances in sarcoma genomics and new therapeutic targets. *Nat Rev Cancer* 2011;11:541–557.
- Dela Cruz F, Matushansky I. MicroRNAs in chromosomal translocation-associated solid tumors: learning from sarcomas. *Discov Med* 2011;12:307–317.
- Jo VY, Fletcher CD. WHO classification of soft tissue tumours: an update based on the 2013 (4th) edition. *Pathology* 2014;46:95–104.
- Riggi N, Cironi L, Suvà ML *et al*. Sarcomas: genetics, signalling, and cellular origins. Part 1: the fellowship of TET. *J Pathol* 2007;213:4–20.
- Brisset S, Schleiermacher G, Peter M *et al*. CGH analysis of secondary genetic changes in Ewing tumors: correlation with metastatic disease in a series of 43 cases. *Cancer Genet Cytogenet* 2001;130:57–61.
- Thyanithy V, Sarver AL, Kartha RV *et al*. Perturbation of 14q32 miRNAs-cMYC gene network in osteosarcoma. *Bone* 2012;50:171–181.
- Drury R, Verghese ET, Hughes TA. The roles of microRNAs in sarcomas. *J Pathol* 2012;227:385–391.
- Sarver AL, Phalak R, Thyanithy V *et al*. S-MED: sarcoma microRNA expression database. *Lab Invest* 2010;90:753–761.
- Kim D, Peretea G, Trapnell C *et al*. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013;14:R36.
- Trapnell C, Williams BA, Pertea G *et al*. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 2010;28:511–515.
- Sarver AL, Thyanithy V, Scott MC *et al*. MicroRNAs at the human 14q32 locus have prognostic significance in osteosarcoma. *Orphanet J Rare Dis* 2013;8:7.
- Sarver AL, Li L, Subramanian S. MicroRNA miR-183 functions as an oncogene by targeting the transcription factor EGR1 and promoting tumor cell migration. *Cancer Res* 2010;70:9570–9580.
- Subramanian S, Thyanithy V, West RB *et al*. Genome-wide transcriptome analyses reveal p53 inactivation mediated loss of miR-34a expression in malignant peripheral nerve sheath tumours. *J Pathol* 2010;220:58–70.
- Reymond A, Camargo AA, Deutsch S *et al*. Nineteen additional unpredicted transcripts from human chromosome 21. *Genomics* 2002;79:824–832.
- Yoshida A, Sekine S, Tsuta K *et al*. NKX2.2 is a useful immunohistochemical marker for Ewing sarcoma. *Am J Surg Pathol* 2012;36:993–999.
- Mahlendorf DE, Staeger MS. Characterization of Ewing sarcoma associated cancer/testis antigens. *Cancer Biol Ther* 2013;14:254–261.
- Teicher BA. Searching for molecular targets in sarcoma. *Biochem Pharmacol* 2012;84:1–10.
- Guo T, Zhang L, Chang NE *et al*. Consistent MYC and FLT4 gene amplification in radiation-induced angiosarcoma but not in other radiation-associated atypical vascular lesions. *Genes Chromosomes Cancer* 2011;50:25–33.
- Krummel MF, Sullivan TJ, Allison JP. Superantigen responses and costimulation: CD28 and CTLA-4 have opposing effects on T cell expansion *in vitro* and *in vivo*. *Int Immunol* 1996;8:519–523.
- Bansal AK, Bindal R, Shetty DC *et al*. Rare occurrence of intraosseous schwannoma in a young child, its review and its pathogenesis. *J Oral Maxillofac Pathol* 2012;16:91–96.
- Voltaggio L, Murray R, Lasota J *et al*. Gastric schwannoma: a clinicopathologic study of 51 cases and critical review of the literature. *Hum Pathol* 2012;43:650–659.
- Hornick JL, Bundock EA, Fletcher CD. Hybrid schwannoma/perineurioma: clinicopathologic analysis of 42 distinctive benign nerve sheath tumors. *Am J Surg Pathol* 2009;33:1554–1561.
- Chen WS, Chen PL, Lu D, Lind AC *et al*. Growth-associated protein 43 in differentiating peripheral nerve sheath tumors from other non-neural spindle cell neoplasms. *Mod Pathol* 2014;27:184–193.
- Karamchandani JR, Nielsen TO, van de Rijn M *et al*. Sox10 and S100 in the diagnosis of soft-tissue neoplasms. *Appl Immunohistochem Mol Morphol* 2012;20:445–450.
- Hirose T, Ishizawa K, Sakaki M *et al*. Retroperitoneal schwannoma is characterized by a high incidence of cellular type and GFAP-immunoreactivity. *Pathol Int* 2012;62:456–462.
- Gershon TR, Oppenheimer O, Chin SS, Gerald WL. Temporally regulated neural crest transcription factors distinguish neuroectodermal tumors of varying malignancy and differentiation. *Neoplasia* 2005;7:575–584.
- De Rosa A, Pellegatta S, Rossi M *et al*. A radial glia gene marker, fatty acid binding protein 7 (FABP7), is involved in proliferation and invasion of glioblastoma cells. *PLoS One* 2012;7:e52113.
- Das S, Sarrou E, Podgrabinska S *et al*. Tumor cell entry into the lymph node is controlled by CCL1 chemokine expressed by lymph node lymphatic sinuses. *J Exp Med* 2013;210:1509–1528.
- Parri M, Pietrovito L, Grandi A *et al*. Angiopoietin-like 7, a novel pro-angiogenic factor over-expressed in cancer. *Angiogenesis* 2014;17:881–896.
- Sato R, Yamasaki M, Hirai K *et al*. Angiopoietin-like protein 2 induces androgen-independent and malignant behavior in human prostate cancer cells. *Oncol Rep* 2015;33:58–66.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–674.
- Sasaki M, Jordan P, Welbourne T *et al*. Troglitazone, a PPAR-gamma activator prevents endothelial cell adhesion molecule expression and lymphocyte adhesion mediated by TNF-alpha. *BMC Physiol* 2005;5:3.
- Leung E, Kanwar RK, Kanwar JR *et al*. Mucosal vascular addressin cell adhesion molecule-1 is expressed outside the endothelial lineage on fibroblasts and melanoma cells. *Immunol Cell Biol* 2003;81:320–327.
- Yu Y, Yin D, Hoque MO *et al*. AKT signaling pathway activated by HIN-1 methylation in non-small cell lung cancer. *Tumour Biol* 2012;33:307–314.
- Bennett LB, Schnabel JL, Kelchen JM *et al*. DNA hypermethylation accompanied by transcriptional repression in follicular lymphoma. *Genes Chromosomes Cancer* 2009;48:828–841.
- Del Castillo I, Moreno-Pelayo MA, Del Castillo FJ *et al*. Prevalence and evolutionary origins of the del(GJB6-D13S1830) mutation in the DFNB1 locus in hearing-impaired subjects: a multicenter study. *Am J Hum Genet* 2003;73:1452–1458.
- Wu P, Walker BA, Brewer D *et al*. A gene expression-based predictor for myeloma patients at high risk of developing bone disease on bisphosphonate treatment. *Clin Cancer Res* 2011;17:6347–6355.
- Sentani K, Oue N, Sakamoto N *et al*. Upregulation of connexin 30 in intestinal phenotype gastric cancer and its reduction during tumor progression. *Pathobiology* 2010;77:241–248.
- Ozawa H, Matsunaga T, Kamiya K *et al*. Decreased expression of connexin-30 and aberrant expression of connexin-26 in human head and neck cancer. *Anticancer Res* 2007;27:2189–2195.
- Georges S, Chesneau J, Hervouet S *et al*. A disintegrin and metalloproteinase 12 produced by tumour cells accelerates osteosarcoma tumour progression and associated osteolysis. *Eur J Cancer* 2013;49:2253–2263.
- Rao UN, Hood BL, Jones-Laughner JM *et al*. Distinct profiles of oxidative stress-related and matrix proteins in adult bone and soft tissue osteosarcoma and desmoid tumors: a proteomics study. *Hum Pathol* 2013;44:725–733.
- Mintz MB, Sowers R, Brown KM *et al*. An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. *Cancer Res* 2005;65:1748–1754.
- Brill E, Gobble R, Angeles C *et al*. ZIC1 overexpression is oncogenic in liposarcoma. *Cancer Res* 2010;70:6891–6901.
- Aruga J, Nozaki Y, Hatayama M, Odaka YS, Yokota N. Expression of ZIC family genes in meningiomas and other brain tumors. *BMC Cancer* 2010;10:79.

45. Alvarez-Díaz S, Valle N, García JM *et al*. Cystatin D is a candidate tumor suppressor gene induced by vitamin D in human colon cancer cells. *J Clin Invest* 2009;119:2343–2358.
46. Rack PG, Ni J, Payumo AY *et al*. Arhgap36-dependent activation of Gli transcription factors. *Proc Natl Acad Sci USA* 2014;111:11061–11066.
47. La Salle S, Palmer K, O'Brien M *et al*. Spata22, a novel vertebrate-specific gene, is required for meiotic progress in mouse germ cells. *Biol Reprod* 2012;86:45.
48. Xiao L, Harrell JC, Perou CM *et al*. Identification of a stable molecular signature in mammary tumor endothelial cells that persists *in vitro*. *Angiogenesis* 2014;17:511–518.
49. Askew DJ, Cataltepe S, Kumar V *et al*. SERPINB11 is a new noninhibitory intracellular serpin. Common single nucleotide polymorphisms in the scaffold impair conformational change. *J Biol Chem* 2007;282:24948–24960.
50. Abu A, Frydman M, Marek D *et al*. Deleterious mutations in the zinc-finger 469 gene cause brittle cornea syndrome. *Am J Hum Genet* 2008;82:1217–1222.
51. Brunner AL, Beck AH, Edris B *et al*. Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers. *Genome Biol* 2012;13:R75.
52. Srikantan V, Zou Z, Petrovics G *et al*. PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. *Proc Natl Acad Sci USA* 2000;97:12216–12221.
53. Walker MJ, Chaudhuri PK, Beattie CW *et al*. Steroid receptors in malignant skeletal tumors. *Cancer* 1980;45:3004–3007.
54. Haile S, Lal A, Myung JK *et al*. FUS/TLS is a co-activator of androgen receptor in prostate cancer cells. *PLoS One* 2011;6:e24197.
55. Brooke GN, Culley RL, Dart DA *et al*. FUS/TLS is a novel mediator of androgen-dependent cell-cycle progression and prostate cancer growth. *Cancer Res* 2011;71:914–924.
56. Ward CL, Boggio KJ, Johnson BN *et al*. A loss of FUS/TLS function leads to impaired cellular proliferation. *Cell Death Dis* 2014;5:e1572.
57. Shing DC, McMullan DJ, Roberts P *et al*. FUS/ERG gene fusions in Ewing's tumors. *Cancer Res*. 2003;63:4568–4576.
58. Su X, Malouf GG, Chen Y *et al*. Comprehensive analysis of long non-coding RNAs in human breast cancer clinical subtypes. *Oncotarget* 2014;5:9864–9876.
59. Gounder MM, Lefkowitz RA, Keohan ML *et al*. Activity of Sorafenib against desmoid tumor/deep fibromatosis. *Clin Cancer Res* 2011;17:4082–4090.
60. Romo-Tena J, Gómez-Martín D, Alcocer-Varela J. CTLA-4 and autoimmunity: new insights into the dual regulator of tolerance. *Autoimmun Rev* 2013;12:1171–1176.
61. Shih K, Arkenau HT, Infante JR. Clinical impact of checkpoint inhibitors as novel cancer therapies. *Drugs* 2014;74:1993–2013.
62. Hart FD, Boardman PL. Indomethacin: a new non-steroid anti-inflammatory agent. *BMJ* 1963;2:965–970.
63. Ornstein PL, Arnold MB, Augenstein NK *et al*. (3SR,4aRS,6RS,8aRS)-6-[2-(1H-tetrazol-5-yl)ethyl]decahydroisoquinoline-3-carboxylic acid: a structurally novel, systemically active, competitive AMPA receptor antagonist. *J Med Chem* 1993;36:2046–2048.
64. Liu YJ, Zhu Y, Yuan HX *et al*. Overexpression of HOXC11 homeobox gene in clear cell renal cell carcinoma induces cellular proliferation and is associated with poor prognosis. *Tumour Biol* 2014;36:2821–2829.
65. deBlacam C, Byrne C, Hughes E *et al*. HOXC11-SRC-1 regulation of S100beta in cutaneous melanoma: new targets for the kinase inhibitor dasatinib. *Br J Cancer* 2011;105:118–123.