Identification of neural stem cell gene expression signatures associated with disease progression in alveolar soft part sarcoma by integrated molecular profiling

Shamini Selvarajah[1,2], Pyne Saumyadipta[2], Eleanor Chen[1], G. Petur Nielsen[3], Glenn Dranoff[2], Edward Stack[2], Massimo Loda[1,2] and Richard Flavin[1,2]

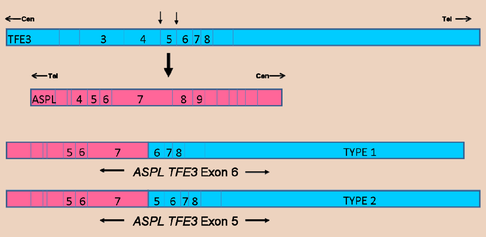
[1]Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; [2]Center for Molecular Oncologic Pathology, Dana–Farber Cancer Institute, Boston, MA 02115; [3]Massachusetts General Hospital, Boston, MA 02114

Alveolar soft part sarcoma (ASPS) is a relatively enigmatic soft tissue sarcoma with poor prognosis. Apart from the recurrent, non-reciprocal t(X;17)(p11.2;q25) translocation, there is little molecular evidence for the origin, initiation and progression of this cancer. We have applied FISH analysis, in conjunction with array comparative genomic hybridization (aCGH) and expression profiling, to examine 16 primary and metastatic ASPS samples, to derive candidate molecular pathways that may be involved in the cancer biology. FISH analysis identified the ASPL-TFE3 fusion in all cases where the N-terminal portion of TFE3 is replaced by the fused ASPL sequences, while retaining the TFE3 DNA-binding domain. High-resolution aCGH revealed a higher number of numerical aberrations in the metastatic tumors compared to the primaries, but failed to identify any consistent alterations in either of the groups. Subsequent DASL analysis revealed 1,063 genes which were differentially expressed, among which, 207 were relatively up-regulated in primaries, while 116 were up-regulated in metastases. Gene set enrichment analysis using key biological process from Gene Ontology, Molecular Signature Database, and GenesigDB identified 16 enriched genesets (p < 0.1) associated with the differentially expressed genes. Notable among these were several stem cell gene expression signatures and pathways related to differentiation. In particular, the paired box transcription factor PAX6 was up-regulated in the primary tumors, along with several genes whose mouse orthologs have previously been implicated in Pax6-DNA binding during neural stem cell differentiation. Of these, the LIM/homeobox protein Lhx6, which plays a key regulatory role in development of neural cells, is a known methylation marker in head and neck carcinomas. While aCGH did not show a metastatic genomic signature, the finding is consistent with the biology of several other translocation-associated sarcomas where transcriptional deregulation from fusion genes is implicated in the pathogenesis of the tumors, rather than extensive chromosomal instability. Of interest was the identification of stem cell gene expression signatures and pathways related to differentiation. In addition to suggesting a neural origin for ASPS, these may reveal more accessible therapeutic targets.

INTRODUCTION

Alveolar soft part sarcoma (ASPS) is a rare, high-grade, mesenchymal malignancy with a distinctive histologic and ultrastructural appearance, but still enigmatic in terms of differentiation and origin. It affects mainly adolescents and young adults in the second and third decade, with slight female predominance. In adults, ASPS occurs most commonly in the deep soft tissues of the thigh or buttock whilst, in children and infants, the head and neck regions are often involved. Despite a relatively indolent clinical course, the prognosis is poor and is often characterized by late metastases. ASPS has been the subject of considerable interest for pathologists and clinicians, owing to its unique microscopic features, uncertain line of differentiation and unpredictable clinical behaviour. The goal of this project is to gain better insights into the underlying pathogenesis of ASPS and possibly identify therapeutic molecular targets by performing gene and tissue microarrays.

Cytogenetic and Molecular Features:



ASPS is characterized by a specific unbalanced translocation: der(17)t(X:17)(p11;p25). This translocation is present as type 1 and 2 variants involving the fusion of the first seven exons of the *ASPL* (*ASPSCR1*) gene to exon 6 (type 1) or 5 (type 2) of the *TFE3* transcription factor gene. This creates a novel *ASPSCR1-TFE3* fusion protein that seems to acts as an aberrant transcription factor that induces unregulated transcription of *TFE3*-regulated genes.

Histopathology:

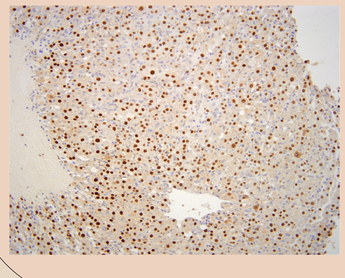
Histologically, ASPS has a distinctive appearance, usually consisting of nests of epithelioid to polygonal cells arranged in nests or bundles, sometimes with a central lack of cohesion leading to the distinctive pseudoalveolar pattern for which it is named. Mitotic activity is usually low and necrosis is infrequent. These tumors also possess a dense capillary vasculature.

Study patient cohort:

12 patients treated at DFCI/BWH and MGH from 1994-2007.
16 total samples (including 5 primary tumors &11 metastases).

Sample Code	Sample Origin	Location	Treatment before Sample Collection?	Age	Sex	Outcome	Overall Survival (months)
1M	**Met**	**Lung**	**Chemo**	**30**	**F**	**Dead**	**84**
2M	**Met**	**Lung**	**Clinical Trial: Gleevec**	**36**	**M**	**AWD**	**132**
4M	**Met**	**Lung**	**Chemo**	**27**	**M**	**Lost to F/U**	**72**
6M	**Met**	**Pancreas**	**None**	**47**	**M**	**AWD**	**156**
7P	Primary	Thigh	None	29	M	Dead	60
7M	Met	Abdominal Wall	None				
7M	Met	Colon	None				
8P	Primary	Retroorbital	None	27	M	AWD	60
8M	Met	Skull	None				
9M	**Met**	**Brain**	**Chemo**	**28**	**M**	**Dead**	**16**
10P	Primary	Thigh	None	28	F	Dead	36
10M	Met	Brain	Radiation				
11P	**Primary**	**Back**	**None**	**20**	**M**	**NED**	**24**
12M	Met	Small Intestine	Chemo	27	F	AWD	168
12M	Met	Femoral Head	Radiation				
12M	**Primary**	**Buttock**	**None**	**26**	**F**	**AWD**	**72**
14P	**Local Recurrence**	**Neck**	**None**	**47**	**F**	**AWD**	**96**

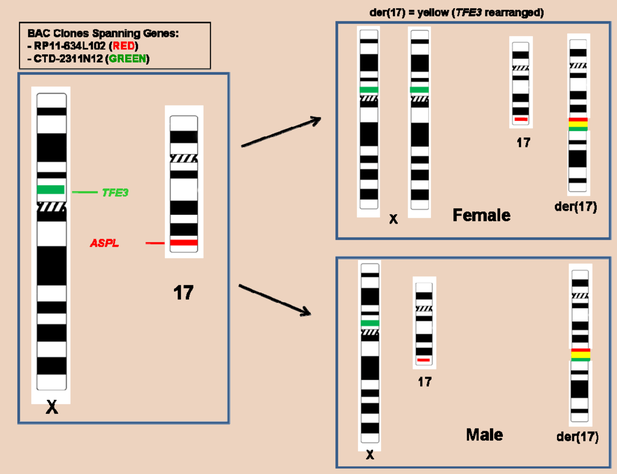
Median Age: 28 (26-47 years-old). 6 males and 6 females (M:F = 1:1). 11/12 presented with metastases at or after diagnosis of primary. Mean follow-up duration: 72 months. 5 patients died (Mean OS: 60 months)

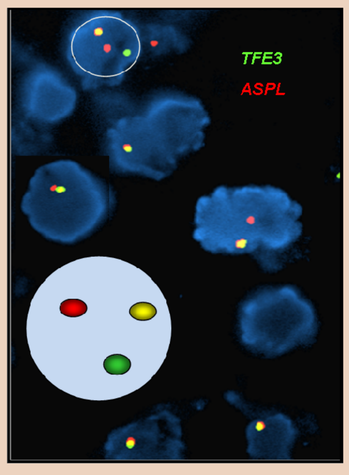


Diagnostic value of TFE3:

Nuclear expression TFE3 is seen almost exclusively in tumors harboring *TFE3* fusion, including ASPS and rare pediatric renal cell carcinoma. Although expression of TFE3 is not specific to ASPS, the presence of strong TFE3 nuclear staining along with morphologic characteristics would strong support the diagnosis of ASPS. Paraffinized human tumor tissues were sectioned at 5 microns and subsequently processed for immunohistochemical staining using a standard protocol. TFE-3 antibody (Santa Cruz, goat polyclonal) was used at 1:500 dilution and the staining was visualized using DAB.

Design of FISH Probes

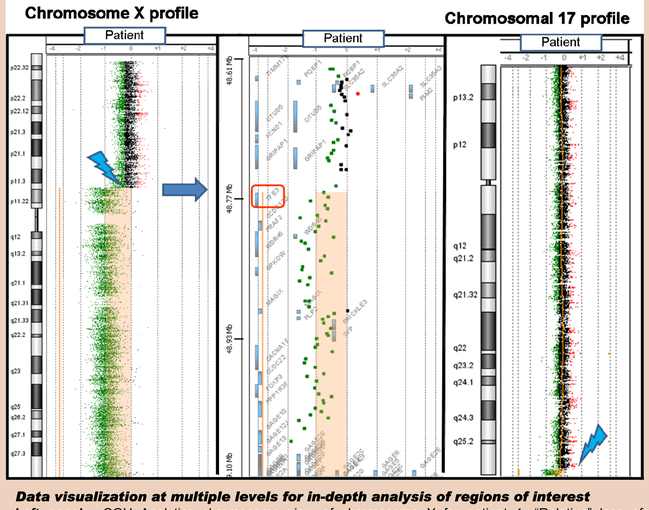




Interphase FISH with BAC-based probes harboring TFE3 and ASPL on a male patient with ASPS:

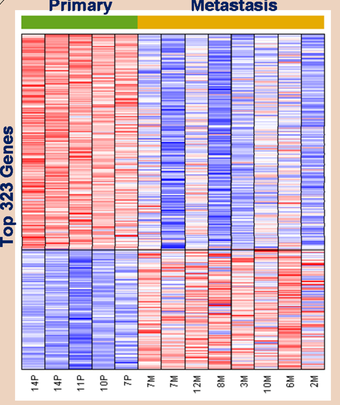
Red and green signals represent the non-rearranged *TFE3* and *ASPL* genes (SpectrumOrange; 17q25.3 and SpectrumGreen; Xp11.2), while the yellow signal derives from the fusion of *TFE3* (from extra Xp segment) and *ASPL* on the derivative chromosome 17. Hybridization on a female with ASPS will comprise of 1 yellow, 1 red and 2 green signals, reflecting the additional chromosome X.

aCGH analysis and FISH findings are consistent with presence of der(17)



Data visualization at multiple levels for in-depth analysis of regions of interest
Left panel: CGH Analytics chromosome view of chromosome X for patient 1. "Relative" loss of chromosome X is observed when male DNA is hybridized to female reference DNA (loss of sequences visualized as deviation of probes to the left of the baseline). Gain of p-arm distal to Xp11.2 can be observed following break at *TFE3* (**central panel**). **Right panel**: loss of sequences distal to *ASPS* (*ASPSCR1* gene) is observed following break at 17q25.3

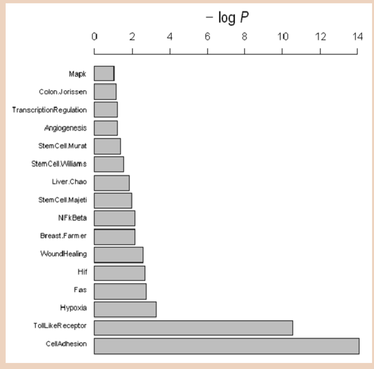
Differential Gene Expression Analysis

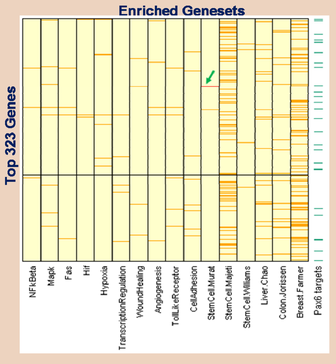


Using RNA extracted from 16 FFPE samples, we studied genome-wide expression based on the high-throughput DASL platform. 3 samples were discarded due to poor signal quality. We started with quantile-normalized data on 18,401 genes, as generated by Broad Institute's DASL platform, and performed a 2-sample t-test to detect differential expression of every gene across 5 primary and 8 metastatic sarcomas with Broad Institute's GenePattern platform. Based on t-test for differential gene expression across primary and meta-static sarcomas, 1,063 genes were found to be significant at the nominal p-value level of 0.05, of which 323 genes with t-test score greater than 3 were most differentially expressed (DE). Among these top DE genes, 207 were up-regulated in primary and 116 in metastases.

Enrichment of key biological processes, gene signatures and pathways among the top DE genes.

For geneset enrichment analysis, we obtained genesets representing key biological processes from Gene Ontology (GO), disease pathways from Molecular Signature Database (MSigDB) and gene expression signatures from the literature from GenesigDB. Genesets representing processes, signatures and pathways that are significant (p < 0.1) are listed in the y-axis. GSEA identified 16 genesets as enriched both among genes that are up-regulated in primary and metastatic sarcomas.



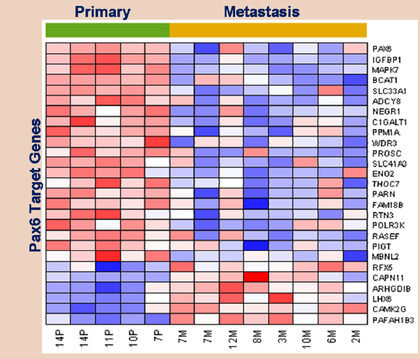


GSEA for all top DE genes

Each column represents a geneset that is either a key biological process/ pathway/ gene signature from the literature that is enriched among the top 323 DE genes. Each horizontal orange bar represents one of the top DE genes depicted in the sorted order t test score. The black horizontal line separates the genes up-regulated in primary sarcomas (above the line) from those up-regulated in metastases. The occurrence of Pax6, as a stem cell transcription factor, is marked with a green arrow and the distribution of its targets among the DE genes according to mouse ChIP studies [Sansom et al.] is shown with green marks beside the plot.

Differential expression of Pax6 targets

The heatmap depicts the top DE genes (in rows) whose mouse orthologs show Pax6 binding with DNA in ChIP studies [Sansom et al.] in the process of neural stem cell differentiation in the developing mouse cerebral cortex. The samples (in columns) are grouped as primary (green labels) and metastatic (gold labels) sarcomas. Red and blue rectangles indicate high and low gene expressions respectively. The expression of the paired box transcription factor Pax6 is shown in the topmost row.



Summary of findings:

1. aCGH/FISH studies show results consistent with the characteristic der(17)t(X;17)(p11.2;q25) translocation.
2. Gene expression profiling reveal genes differentially regulated between primary tumors and metastases. Many of the enriched genes are involved in biological processes essential for tumor progression (e.g. adhesion, hypoxia and proliferation). In addition, GSEA also reveals a unique stem cell-like signature, involving the Pax-6 regulated set of genes associated with neural stem cell biology.

Funding :

This study has been made possible through the generous support of:

iCureASPS

NUCLEABIOMARKERS