Alveolar soft-part sarcoma (ASPS) resembles a mesenchymal stromal progenitor: Evidence from meta-analysis of transcriptomic data

Luke H Stockwin

1 Unaffiliated, Frederick, Maryland

Corresponding Author: Luke H Stockwin
Email address: luke.stockwin@gmail.com

Alveolar soft-part sarcoma (ASPS) is an extremely rare malignancy characterized by the unbalanced translocation der(17)t(X;17)(p11;q25). This translocation generates a fusion protein, ASPL-TFE3, that drives pathogenesis through aberrant transcriptional activity. Although considerable progress has been made in identifying ASPS therapeutic vulnerabilities (e.g., MET inhibitors), basic research efforts are hampered by the lack of appropriate in vitro reagents with which to study the disease. In this report, previously unmined microarray data for the ASPS cell line, ASPS-1, was analyzed relative to the NCI sarcoma cell line panel. These data were combined with meta-analysis of pre-existing ASPS patient microarray and RNA-seq data to derive a platform-independent ASPS transcriptome. Results demonstrated that ASPS-1, in the context of the NCI sarcoma cell panel, had some similarities to normal mesenchymal cells and connective tissue sarcomas. The cell line was characterized by high relative expression of transcripts such as CRYAB, MT1G, GCSAML, and SV2B. Notably, ASPS-1 lacked mRNA expression of myogenesis-related factors MYF5, MYF6, MYOD1, MYOG, PAX3, and PAX7. Furthermore, ASPS-1 had a predicted mRNA surfaceome resembling an undifferentiated mesenchymal stromal cell through expression of GPNMB, CD9 (TSPAN29), CD26 (DPP4), CD49C (ITGA3), CD54 (ICAM1), CD63 (TSPAN30), CD68 (SCARD1), CD130 (IL6ST), CD146 (MCAM), CD147 (BSG), CD151 (SFA-1), CD166 (ALCAM), CD222 (IGF2R), CD230 (PRP), CD236 (GPC), CD243 (ABCB1), and CD325 (CDHN). Subsequent re-analysis of ASPS patient data generated a consensus expression profile with considerable overlap between studies. In common with ASPS-1, elevated expression was noted for CTSK, DPP4, GPNMB, INHBE, LOXL4, PSG9, SLC20A1, STS, SULT1C2, SV2B, and UPP1. Transcripts over-expressed only in ASPS patient samples included ABCB5, CYP17A1, HIF1A, MDK, P4HB, PRL, and PSAP. These observations are consistent with that expected for a mesenchymal progenitor cell with adipogenic, osteogenic, or chondrogenic potential. In summary, the consensus data generated in this study highlight the unique and highly conserved nature of the ASPS transcriptome.
Although the ability of the ASPL-TFE3 fusion to perturb mRNA expression must be acknowledged, the prevailing ASPS transcriptome resembles that of a mesenchymal stromal progenitor.
Alveolar soft-part sarcoma (ASPS) resembles a mesenchymal stromal progenitor: evidence from meta-analysis of transcriptomic data

Luke H Stockwin

1 Unaffiliated

Corresponding Author:
Luke Stockwin
1021 N. Market St., Frederick, Maryland, 21701, United States.
Email address: luke.stockwin@gmail.com

Abstract

Alveolar soft-part sarcoma (ASPS) is an extremely rare malignancy characterized by the unbalanced translocation der(17)t(X;17)(p11;q25). This translocation generates a fusion protein, ASPL-TFE3, that drives pathogenesis through aberrant transcriptional activity. Although considerable progress has been made in identifying ASPS therapeutic vulnerabilities (e.g., MET inhibitors), basic research efforts are hampered by the lack of appropriate *in vitro* reagents with which to study the disease. In this report, previously unmined exon array data for the ASPS cell line, ASPS-1, was analyzed relative to the NCI sarcoma cell line panel. These data were combined with meta-analysis of pre-existing ASPS patient microarray and RNA-seq data to derive a platform-independent ASPS transcriptome. Results demonstrated that ASPS-1, in the context of the NCI sarcoma cell panel, had some similarities to normal mesenchymal cells and connective tissue sarcomas. The cell line was characterized by high relative expression of transcripts such as \(\text{CRYAB}\), \(\text{MT1G}\), \(\text{GCSAML}\), and \(\text{SV2B}\). Notably, ASPS-1 lacked mRNA expression of myogenesis-related factors \(\text{MYF5}\), \(\text{MYF6}\), \(\text{MYOD1}\), \(\text{MYOG}\), \(\text{PAX3}\), and \(\text{PAX7}\). Furthermore, ASPS-1 had a predicted mRNA surfaceome resembling an undifferentiated mesenchymal stromal cell through expression of \(\text{GPNMB}\), \(\text{CD9}\) (\(\text{TSPAN29}\)), \(\text{CD26}\) (\(\text{DPP4}\)), \(\text{CD49C}\) (\(\text{ITGA3}\)), \(\text{CD54}\) (\(\text{ICAM1}\)), \(\text{CD63}\) (\(\text{TSPAN30}\)), \(\text{CD68}\) (\(\text{SCARD1}\)), \(\text{CD130}\) (\(\text{IL6ST}\)), \(\text{CD146}\) (\(\text{MCAM}\)), \(\text{CD147}\) (\(\text{BSG}\)), \(\text{CD151}\) (\(\text{SFA-1}\)), \(\text{CD166}\) (\(\text{ALCAM}\)), \(\text{CD222}\) (\(\text{IGF2R}\)), \(\text{CD146}\) (\(\text{PRP}\)), \(\text{CD236}\) (\(\text{GPC}\)), \(\text{CD243}\) (\(\text{ABCB1}\)), and \(\text{CD325}\) (\(\text{CDHN}\)). Subsequent re-analysis of ASPS patient data generated a consensus expression profile with considerable overlap between studies. In common with ASPS-1, elevated expression was noted for \(\text{CTSK}\), \(\text{DPP4}\), \(\text{GPNMB}\), \(\text{INHBE}\), \(\text{LOXL4}\), \(\text{PSG9}\), \(\text{SLC20A1}\), \(\text{STS}\), \(\text{SULT1C2}\), \(\text{SV2B}\), and \(\text{UPP1}\). Transcripts over-expressed only in ASPS patient samples included \(\text{ABCB5}\), \(\text{CYP17A1}\), \(\text{HIF1A}\), \(\text{MDK}\), \(\text{P4HB}\), \(\text{PRL}\), and \(\text{PSAP}\). These observations are consistent with that expected for a mesenchymal progenitor cell with adipogenic, osteogenic, or chondrogenic potential. In summary, the consensus data generated in this study highlight the unique and highly conserved nature of the ASPS transcriptome. Although the ability of the ASPL-TFE3 fusion to perturb mRNA expression must be acknowledged, the prevailing ASPS transcriptome resembles that of a mesenchymal stromal progenitor.
Introduction

Alveolar Soft-Part Sarcoma (ASPS) is an extremely rare soft tissue sarcoma of adolescents and young adults (Christopherson et al. 1952; Paoluzzi & Maki 2019). ASPS usually manifests as a soft, painless, slow-growing mass and although disease follows an indolent course, it has the potential to metastasize to several sites (Portera et al. 2001). ASPS is characterized by an unbalanced translocation (X;17)(p11;q25) that fuses the ASPSCR1 and TFE3 genes, generating a fusion protein that drives pathogenesis (Ladanyi et al. 2001). Evidence suggests that the fusion protein accumulates in the nucleus and directs transcriptional activity (Argani et al. 2003; Betschinger et al. 2013; Hirobe et al. 2013). For example, ASPL-TFE3 binds and activates MET transcription, resulting in an overall enhancement in kinase activity in the presence of hepatocyte growth factor (Tsuda et al. 2007). As a consequence, some clinical benefit is being achieved with kinase inhibitors targeting MET (Paoluzzi & Maki 2019; Schoffski et al. 2018).

Despite this progress, the origin of disease is still the subject of intense speculation (Folpe & Deyrup 2006). A longstanding hypothesis posits that ASPS has a myogenous origin (Fisher & Reidbord 1971; Folpe & Deyrup 2006; Mukai et al. 1983). However, ASPS tumors do not appear to express markers of muscle cell differentiation such as the myogenic nuclear regulatory proteins MyoD1 and myogenin (Gomez et al. 1999; Hoshino et al. 2009; Tallini et al. 1994; Wang et al. 1996). Several transcriptomic studies have also been published that speculate on the origin of disease (Goodwin et al. 2014; Selvarajah et al. 2014; Stockwin et al. 2009; Tanaka et al. 2017). In 2009 we undertook one of the first microarray studies of ASPS and identified expression of several muscle-restricted transcripts (ITGB1BP3/MIBP, MYF5, MYF6, and TRIM63). However, these data were generated using universal RNA as a reference, which may have biased towards skeletal muscle expressed transcripts (Stockwin et al. 2009). Selvarajah et al., showed that the transcription factor PAX6 was upregulated in primary ASPS, suggesting a “tentative neural line of differentiation for ASPS” (Selvarajah et al. 2014). Goodwin et al., generated microarray data from a mouse model of ASPS and also human patient samples. These authors speculated that “some mesenchymal progenitor, possibly pericyte/endothelial in character, provides one potential cell of origin” (Goodwin et al. 2014). Similarly, Tanaka et al. were able to model ASPS through ectopic expression of ASPL-TFE3 in murine embryonic, but not adult, mesenchymal cells (Tanaka et al. 2017). These observations underscore the current lack of clarity with respect to ASPS ontogeny and lend support to the suggestion that ASPS cells represent a “scrambled” phenotype where the ASPL-TFE3 fusion impairs differentiation (Folpe & Deyrup 2006; Naka et al. 2013).

In 2011, a multi-year study culminated in the development of an ASPS cell line designated ASPS-1 (Kenney et al. 2011). This reagent provided the first opportunity to study ASPS gene expression without interference from contaminating cell types. Microarray data was subsequently generated for ASPS-1 as part of the NCI sarcoma cell line panel (Teicher et al. 2015). In the current study, ASPS-1 data was mined relative to the entire NCI sarcoma cell line panel. These efforts were combined with re-analysis of microarray and RNA-seq studies focusing on ASPS patient samples (Goodwin et al. 2014; Kummar et al. 2013; Stockwin et al. 2009). In this regard, we aimed to unify current publicly available transcriptomic data into a consensus profile that can be used as a basis for exploring disease ontogeny and therapeutic...
vulnerabilities. Results obtained in this study show that, at the mRNA level, ASPS resembles a mesenchymal stromal progenitor.

*Throughout this report the abbreviation MSC is used interchangeably for “mesenchymal stromal cells” and “mesenchymal stem cell”, although the former is preferred.

Materials and Methods

This study utilized six datasets downloaded from the gene expression omnibus (https://www.ncbi.nlm.nih.gov/geo/). GSE68591 comprises exon expression data (Affymetrix Human Exon 1.0 v2 ST platform) for the NCI sarcoma cell line panel (includes data from ASPS-1, 67 sarcoma lines, and five normal tissues) (Teicher et al. 2015). GSE13433 comprises mRNA expression data for seven ASPS patient tumors analyzed using the Affymetrix U133 plus 2.0 platform (Stockwin et al. 2009). For the analyses of data from GSE13433, additional U133 plus 2.0 control arrays were obtained from GSE17070 (normal skeletal muscle) and GSE118370 (normal lung). GSE32569 comprises a set of U133 plus 2.0 microarrays generated from six patients pre- and post- treatment with Cediranib (Kummar et al. 2013). Lastly, GSE54729 comprises RNA-seq data (HiSeq 2000) from five ASPS patients and three skeletal muscle controls. The overall study design is illustrated in Figure 1. For experiments involving Affymetrix human Exon 1.0 ST and U133 plus 2.0 arrays differentially expressed genes were identified using the Transcriptome Analysis Console (TAC 4.0, ThermoFisher Scientific) using standard algorithm and comparison settings (RMA normalization, P<0.05, FDR <0.05, fold change +/-2). The TAC was also used to generate hierarchical clusters using the automated workflow. For RNA-seq data, differential expression values relative to skeletal muscle were determined using the GSE54729_10408R.txt spreadsheet that accompanies the submission. In detail, normalized FKPM values were averaged for the 5 human ASPS samples and the 3 normal human skeletal muscles samples; fold changes were then calculated from these values. In terms of utilities; the GTEX portal (https://www.gtexportal.org/) multi gene query option was used to inform tissue of origin from the top 50 differentially expressed ASPS-1 transcripts. Similarly, the GTEX Top 50 expressed genes search function (https://gtexportal.org/home/topExpressedGenePage) was used to identify genes expressed selectively in skeletal muscle. The Protein Atlas (https://www.proteinatlas.org/) was used to investigate both mRNA and protein expression in normal and cancerous samples for specific transcripts. The in silico surfaceome (http://wlab.ethz.ch/surfaceome/) (Bausch-Fluck et al. 2018) was used to predict the hierarchy of cell surface protein expression for ASPS-1. In detail, a file containing the published human surfaceome (table_S3_surfaceome.xlsx) was downloaded and merged, using MS excel, with the list of differentially expressed ASPS-1 transcripts. Transcripts appearing in both datasets were then extracted and sorted according to ASPS-1 expression. The VENN diagram utility InteractiVenn (http://www.interactivenn.net/) (Heberle et al. 2015) was used to determine the extent of overlap between the 4 different experimental approaches.

Results and Discussion

In the first analysis, ASPS-1 exon-level data from the NCI sarcoma cell line panel was analyzed relative to all samples (cancer and normal). Results shown in Table 1 list the top fifty transcripts...
over-expressed in ASPS-1 relative to the average (raw data in Supplemental Table 1). Results
demonstrated that crystallin alpha beta (CRYAB) mRNA showed the highest expression in ASPS-
1 relative to the average (227 fold). Metallothionein 1G (MT1G) was the next most elevated (202
fold). Following this were several lower-abundance transcripts coding for C7orf69, Synaptic
Vesicle Glycoprotein 2B (SV2B) and Germinal Center-Associated Signaling and Motility-Like
Protein (GCSAML). In common with the initial published report of ASPS-1, results also
confirmed expression of GPNMB (Kenney et al. 2011). Likewise, ASPS-1 had some of the
highest levels of MET, and VEGFR2 in the panel, both of which are previously noted
characteristics of disease (Jun et al. 2010; Stockwin et al. 2009; Tsuda et al. 2007). This analysis
served to confirm the ASPS origin of ASPS-1 as outlined in Kenney et al. (Kenney et al. 2011).

Following this, ASPS-1 was subjected to hierarchical clustering relative to the entire NCI
sarcoma cell line panel (Figure 2A). Results showed that ASPS-1 was an outlier that did not
closely align with any lines/tissues in the panel. Those specimens with the nearest similarity are
shown in Figure 2B; this includes normal cell/tissues such as knee articular chondrocytes, dermal
fibroblasts, skeletal muscle, mesenchymal stem cells and osteoblasts. Cancer line relatives
included the spindle cell sarcoma Hs 132.T, the fibrosarcomas Hs 93.T and Hs 414.T along with
the chondrosarcoma Hs 819. In order to define which changes were guiding this clustering,
several breakout analyses were performed. Over-expressed transcripts found in the nearest
cluster group and ASPS-1 included GPNMB, CRYAB, FABP3, and CTSK, which are markers of
mesenchymal cells (Debnath et al. 2018; Kulterer et al. 2007; Wang et al. 2014; Yu et al. 2016).
Transcripts expressed only in ASPS-1 relative to the nearest cluster included SEPT3, C7orf69,
MT1G, and ACP5, many of which are implicated in the differentiation of mesenchymal stromal
cells (Dohi et al. 2005; Hayman et al. 2000; Moller et al. 2018; Tan et al. 2015). Lastly, the
nearest cluster group could be distinguished from ASPS-1 through higher expression of the
canonical fibroblast markers GREM1, LOX, THY1, and POSTN (Hortells et al. 2019; Jiang &
Rinkevich 2018; Karagiannis et al. 2012). These data point towards a mesenchymal stromal
origin that had not undergone significant fibroblast lineage specialization.

The Genotype-Tissue Expression (GTEx) portal (https://www.gtexportal.org/) was then used in
order to determine whether the ASPS-1 expression data suggested a tissue of origin. Results
obtained using the top fifty over-expressed transcripts suggested that cardiac/skeletal muscle was
a likely origin through expression of transcripts such as CRYAB, FABP3, SLN, and CKMT2
(Figure 2C). The GTEx database also lists the top 50 transcripts expressed by normal skeletal
muscle, where 28 (e.g. creatine kinase M-Type CKM and myoglobin MB) have a considerable
degree of muscle-specificity. Results showed that only two transcripts from this set, EEF1A2 and
SLN are over-expressed in ASPS-1. As a consequence, we then undertook an expansive study of
whether ASPS-1 had any hallmarks of myogenic differentiation. Figure 3 plots raw expression
data for all cell lines in the sarcoma panel, where lines are grouped according to histology.
Results showed that ASPS-1, unlike some rhabdosarcoma lines, did not express myogenic
regulatory factor mRNAs including MYF5, MYF6, MYOD1, MYOG, and similarly did not
express PAX3 and PAX7. Taken together these observations suggest that although ASPS-1 has
hallmarks of a muscle resident cell, it had not undergone myogenic differentiation.

These data also provide a unique opportunity to assess the potential cell surface phenotype of
ASPS-1. Bausch-Fluck et al. identified 2886 proteins that are known, or are predicted by
machine learning, to be expressed on the cell surface (Bausch-Fluck et al. 2018). Here, ASPS-1 raw microarray data was filtered for these targets and the list sorted in terms of expression. The resultant ASPS-1 ‘surfaceome’ is shown in Supplemental data 2. As could be anticipated, GPNMB was the highest expressed mRNA; followed by novel surface makers such as the glutamate transporter SLC38A1 and the amyloid beta (A4) precursor protein APP. In terms of CD antigens, the following mRNAs were highly expressed in ASPS-1; CD9 (TSPAN29), CD26 (DPP4), CD49C (ITGA3), CD54 (ICAM1), CD63 (TSPAN30), CD68 (SCARD1), CD130 (IL6ST), CD146 (MCAM), CD147 (BSG), CD151 (SFA-1), CD166 (ALCAM), CD222 (IGF2R), CD230 (PRP), CD236 (GPC), CD243 (ABCB1), and CD325 (CDHN).

Many of these observations are also compatible with a mesenchymal stromal cell. For example, CD9 (TSPAN29) and CD243 (ABCB1), although widely expressed, are found to varying degrees on MSC (Islam et al. 2005; Kim et al. 2007). CD49C (ITGA3) and CD151 (SFA-1) are both markers of chondrogenic differentiation in MSC (Grogan et al. 2007; Lee et al. 2009a). Expression of CD54 (ICAM1) can be induced in MSC (Ren et al. 2010), CD63 (TSPAN30) is expressed by bone marrow MSC (McBride et al. 2017), CD68 (SCARD1) expression has been shown on MSC from human umbilical cord (La Rocca et al. 2009) and CDH2 is a regulator of mesenchymal stem cell fate (Alimperti & Andreadis 2015). Exosomes expressing Basigin, BSG (CD147), have been shown to promote angiogenesis in MSC (Vrijsen et al. 2016). CD147 is also a major constituent of the pre-crystalline granules present in ASPS (Ladanyi et al. 2002).

In summary, the results presented here demonstrate that the ASPS-1 transcriptome is unique amongst the NCI sarcoma panel, where the closest relatives are normal mesenchymal cells and connective tissue sarcomas. Although ASPS-1 has an expression signature with some similarity to skeletal/cardiac muscle tissue, markers of myogenesis were not detected in this cell line. Furthermore, the ASPS-1 surfaceome does not immediately speak to a tissue derivation but suggests an undifferentiated mesenchymal state.

The next phase of the project involved re-analyzing microarray and RNA-seq data from ASPS tumor resections. GSE13433 comprises microarray data (Affymetrix U133 plus 2.0) for seven patients with primary or metastatic ASPS (Stockwin et al. 2009). In the original study, universal RNA (representing a collection of adult human tissues) was used as a reference. However, patient samples 1, 3, 5 and 6 were obtained from skeletal muscle biopsies whereas samples 4 and 7 were isolated from lung. As a consequence, microarray data from normal skeletal muscle and lung represent more appropriate controls. Therefore, skeletal muscle arrays were obtained from GSE17070 and normal lung samples from GSE118370. Patient 2 data, derived from the mandible, was excluded from the analysis for lack of an appropriate control. Two lists of differentially expressed transcripts were then generated for patients 1,3,5,6 vs skeletal muscle and 4,7 vs. normal lung. Results from these two experiments were largely concordant with the profiles obtained in the original study (Stockwin et al. 2009). However, in Stockwin et al.,
muscle-differentiation associated transcripts ITGB1BP3/MIBP, MYF5 and MYF6 were identified as overexpressed. In our analysis, only MYF6 was identified, and only in the experiment involving patients 4 and 7 vs normal lung; supporting our inference that the published study over-emphasized myogenic differentiation in patient ASPS.

GSE32569 is a similar ASPS dataset where U133 plus 2.0 microarrays were generated from patients treated with Cediranib (Kummar et al. 2013). We undertook to use this data to generate a list of differentially expressed transcripts from pre-treatment arrays relative to GSE17070 skeletal muscle samples. Results again showed a similar profile to that obtained from the analysis of GSE13433. The final experiment was performed using RNA-seq data from GSE54729. In this published study, data was generated from five human ASPS tumor samples in order to compare the transcriptome with five mouse tumors generated through ectopic expression of ASPSCR1-TFE3 (Goodwin et al. 2014). Here, FKPM values for the five human tumors and three skeletal muscle controls were used to generate a list of differentially expressed transcripts. The top fifty upregulated transcripts generated from each of the four experiments using GSE13433, GSE32569 and GSE54729 are shown in Table 2. Meta-analysis of data from these four in vivo studies had considerable overlap, emphasizing the consistent upregulation of mRNAs such as GPNMB, ABCB5, PSG9, CYP17A1, PRL, SULT1C2, and SV2B. As with ASPS-1, over-expression of myogenic regulatory factor mRNA was not consistently seen in any of the experiments involving patient samples (results not shown). Lastly, lists of differentially expressed genes from both ASPS-1 and patient sample experiments were combined (at a five-fold cut-off*), and a VENN diagram generated in order to determine the extent of overlap (Figure 4). Results demonstrated that twenty-five transcripts were elevated in all of the meta-analyses, whereas seventy-three were expressed at the intersection between all in vivo analyses.

An exploration of the twenty-five conserved transcripts in the context of stem cell biology provides further insights into MSC lineage potential. For example; angiopoietin Like 2 (ANGPTL2) is a regulator of stem cell adipogenesis, chondrogenesis and osteogenesis (Takano et al. 2017; Tanoue et al. 2018). Expression of Cathepsin K (CTSK) is compelling given that in mice CTSK-mGFP cells label the periosteal mesenchyme and have been used to identify periosteal stem cells (Debnath et al. 2018). As previously noted, Dipeptidyl Peptidase 4 (DPP4), also known as CD26, marks mesenchymal preadipocyte progenitors (Merrick et al. 2019). Glycoprotein Nmb (GPNMB) should be recognized as the prototypic cell surface marker for ASPS. As stated, GPNMB is recognized as a marker of mesenchymal cells (Kuci et al. 2019). Interestingly, within protein atlas, the cell line designated ‘ASC diff’, a differentiated adipose-derived mesenchymal stem cell line has the highest expression of GPNMB and also expresses TRIM63, CRYAB, FABP3, and CTSK. These observations would appear to favor the concept that ASPS resembles an MSC capable of adipogenic, chondrogenic or osteogenic differentiation.

Several inferences can also be made for the seventy-three conserved transcripts identified in all ASPS patient experiments. The multi-drug resistance transporter ABCB5, in addition to being expressed by melanoma, also defines a subset of MSC in the cornea and skin (Frank et al. 2003; Ksander et al. 2014; Vander Beken et al. 2019). The hormone prolactin (PRL) has been shown to stimulate proliferation of MSC and also to direct chondrogenic and osteogenic differentiation (Ogueta et al. 2002; Seriwatanachai et al. 2012; Surarit et al. 2016). Increased expression of the growth factor midkine (MDK) has been noted in previous ASPS gene expression studies and is
an MSC survival factor (Stockwin et al. 2009; Zhao et al. 2014). Upregulation of hypoxia-related transcripts such as HIF1A suggests that this pathway is active in ASPS and, although ubiquitous, HIF1A plays an important role in the control of multipotency for MSC (Palomaki et al. 2013). It was similarly interesting that the ASPS patient experiments showed increased expression of THY1 (CD90) relative to control samples. This target is regarded as a classical marker of MSC and has recently been shown to promote osteogenic differentiation over an adipogenic fate (Saalbach & Anderegg 2019). In summary, re-analysis of microarray and RNA-seq data for ASPS patient samples yielded transcriptomes with considerable overlap between studies irrespective of platform technology; and the final consensus ASPS transcriptome resembles an undifferentiated mesenchymal stromal cell.

A supplemental dataset (file 3) is also provided that comprises low-stringency (2-fold) expression data provided so that any gene of interest can also be analyzed for expression over the entire set of experiments.

Conclusions

Alveolar-soft part sarcoma is an example of a malignancy that has, despite several immunohistochemical and genomics studies, evaded classification (Fisher & Reidbord 1971; Folpe & Deyrup 2006; Gomez et al. 1999; Goodwin et al. 2014; Hoshino et al. 2009; Mukai et al. 1983; Selvarajah et al. 2014; Stockwin et al. 2009; Tallini et al. 1994; Tanaka et al. 2017; Wang et al. 1996). This study was prompted by the public release of exon expression data for the cell line ASPS-1, which offers a unique opportunity to study ASPS in isolation (Kenney et al. 2011; Teicher et al. 2015). We were similarly interested in revisiting the genomic studies of ourselves and others to generate a consensus expression profile independent of platform technology.

The central finding of the current study was that the ASPS transcriptome is indicative of an undifferentiated mesenchymal stromal cell (MSC). Specifically, The ASPS-1 cell line exhibited a mesenchymal expression signature, where expression data clustered with normal and malignant mesenchymal cells within the NCI sarcoma cell line panel. The ASPS-1 surfaceome was similarly suggestive of an undifferentiated mesenchymal cell. Generation of an ASPS consensus transcriptome from previously reported patient studies highlighted the importance of targets such as GPNMB, ABCB5, CSTK, DPP4, BSG, ALCAM, PRL, and CDHN; all of which were consistent with an undifferentiated MSC. Conversely, the ASPS transcriptome lacked expression of myogenesis-related genes and did not feature transcripts indicative of neural, pericyte or endothelial differentiation.

MSC are found in most tissues, these cells are capable of multipotent differentiation into bone, muscle, cartilage, adipocytes, marrow stromal cells, tenocytes, fibroblasts, endothelial and neural cells (Caplan 2007; Pittenger et al. 2019). Tissues maintain a pool of MSC, with varying degrees of specialization, ready to dynamically replenish differentiated cells in response to signals associated with growth, homeostasis or damage (Rubenstein et al. 2020). Prior to this study, ASPL-TFE3 had already been shown to immortalize embryonic mesenchymal cells (Tanaka et al. 2017). The suggestion that ASPS resembles a mesenchymal stromal progenitor provides a plausible explanation for the failure of previous studies to pinpoint cellular origin, given that the
cell retains an undifferentiated state. Evidence from this study favors an MSC capable of adipogenic, osteogenic or chondrogenic differentiation, but not necessarily at the exclusion of other lineages.

If an MSC origin for ASPS could ultimately be confirmed, there would be important consequences for therapeutic development. Foremost is the suggestion that ASPS growth may be inhibited by factors that promote MSC differentiation. For example, several high-throughput studies have identified clinically relevant small molecules capable of promoting or inhibiting differentiation of MSC (Brey et al. 2011; Huang et al. 2008). Re-screening these compounds for their ability to inhibit the growth of ASPS-1 may yield clinically tractable candidates for ASPS treatment. From the authors perspective, the effect of HDAC inhibitors, steroids and retinoids on ASPS-1 growth are of particular interest (Lee et al. 2009b; Salloum et al. 2013).

The findings of this study have a central caveat; all speculation regarding cellular origin must be moderated until the inference of ASPL-TFE3 is removed. Given the ability of this fusion protein to re-direct transcription, the observed transcriptomes may mask the true cellular origin. For example, GPNMB, CRYAB, CYP17A1, SULT1C2, UPP1 and SV2B have been shown to be up-regulated following expression of ASPL-TFE3 in inducible 293 cells (Kobos et al. 2013). Therefore, the current study only suggests that ASPS resembles an MSC and no firm conclusion can be made regarding origin. A straightforward approach to address this central question involves generating an ASPL-TFE3 knockout in ASPS-1 perhaps with re-introduction of wild-type TFE3 to maintain viability. The resultant line could then be characterized by RNA-seq and FACS phenotyping. These experiments could be accompanied by the addition of defined media to determine whether differentiation can be directed toward specific MSC lineages. In the interim, the data presented here provides a unified picture of ASPS mRNA expression, where considerable similarity with mesenchymal stromal progenitors is evident.

Acknowledgments

I would like to thank Dr. Francesco Tomassoni, Dr. C. Andrew Stewart and Marie Stockwin, PAC for scientific and editorial input.

References

Alimperti S, and Andreadis ST. 2015. CDH2 and CDH11 act as regulators of stem cell fate decisions. Stem Cell Res 14:270-282. 10.1016/j.scr.2015.02.002
Argani P, Lal P, Hutchinson B, Lui MY, Reuter VE, and Ladanyi M. 2003. Aberrant nuclear immunoreactivity for TFE3 in neoplasms with TFE3 gene fusions: a sensitive and specific immunohistochemical assay. Am J Surg Pathol 27:750-761. 10.1097/00000478-200306000-00005
Bausch-Fluck D, Goldmann U, Muller S, van Oostrum M, Muller M, Schubert OT, and Wollscheid B. 2018. The in silico human surfaceome. Proc Natl Acad Sci U S A 115:E10988-E10997. 10.1073/pnas.1808790115
Betschinger J, Nichols J, Dietmann S, Corrin PD, Paddison PJ, and Smith A. 2013. Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. Cell 153:335-347. 10.1016/j.cell.2013.03.012
Brey DM, Motlekar NA, Diamond SL, Mauck RL, Garino JP, and Burdick JA. 2011. High-throughput screening of a small molecule library for promoters and inhibitors of mesenchymal stem cell osteogenic differentiation. Biotechnol Bioeng 108:163-174. 10.1002/bit.22925

Bruder SP, Ricalton NS, Boynton RE, Connolly TJ, Jaiswal N, Zaia J, and Barry FP. 1998. Mesenchymal stem cell surface antigen SB-10 corresponds to activated leukocyte cell adhesion molecule and is involved in osteogenic differentiation. J Bone Miner Res 13:655-663. 10.1359/jbmr.1998.13.4.655

Caplan AI. 2007. Adult mesenchymal stem cells for tissue engineering versus regenerative medicine. J Cell Physiol 213:341-347. 10.1002/jcp.21200

Christopherson WM, Foote FW, Jr., and Stewart FW. 1952. Alveolar soft-part sarcomas; structurally characteristic tumors of uncertain histogenesis. Cancer 5:100-111. 10.1002/1097-0142(195201)5:1<100::aid-cncr2820050112>3.0.co;2-k

Covas DT, Panepucci RA, Fontes AM, Silva WA, Jr., Orellana MD, Freitas MC, Neder L, Santos AR, Peres LC, Jamur MC, and Zago MA. 2008. Multipotent mesenchymal stromal cells obtained from diverse human tissues share functional properties and gene-expression profile with CD146+ perivascular cells and fibroblasts. Exp Hematol 36:642-654. 10.1016/j.exphem.2007.12.015

Debnath S, Yallowitz AR, McCormick J, Lalani S, Zhang T, Xu R, Li N, Liu Y, Yang YS, Eiseman M, Shim JH, Hameed M, Healey JH, Bostrom MP, Landau DA, and Greenblatt MB. 2018. Discovery of a periosteal stem cell mediating intramembranous bone formation. Nature 562:133-139. 10.1038/s41586-018-0554-8

Dohi Y, Shimaoka H, Ikeuchi M, Ohgushi H, Yonemasu K, and Minami T. 2005. Role of metallothionein isoforms in bone formation processes in rat marrow mesenchymal stem cells in culture. Biol Trace Elem Res 104:57-70. 10.1385/BTER:104:1:057

Espagnolle N, Guilloton F, Deschaseaux F, Gadelorge M, Sensebe L, and Bourin P. 2014. CD146 expression on mesenchymal stem cells is associated with their vascular smooth muscle commitment. J Cell Mol Med 18:104-114. 10.1111/jcm.12168

Fisher ER, and Reidbord H. 1971. Electron microscopic evidence suggesting the myogenous derivation of the so-called alveolar soft part sarcoma. Cancer 27:150-159. 10.1002/1097-0142(197101)27:1<0150::aid-cncr2820270123>3.0.co;2-0

Folpe AL, and Deyrup AT. 2006. Alveolar soft-part sarcoma: a review and update. J Clin Pathol 59:1127-1132. 10.1136/jcp.2005.031120

Frank NY, Pendse SS, Lapchak PH, Margaryan A, Shlain D, Doeing C, Sayegh MH, and Frank MH. 2003. Regulation of progenitor cell fusion by ABCB5 P-glycoprotein, a novel human ATP-binding cassette transporter. J Biol Chem 278:47156-47165. 10.1074/jbc.M308700200

Gomez JA, Amin MB, Ro JY, Linden MD, Lee MW, and Zarbo RJ. 1999. Immunohistochemical profile of myogenin and MyoD1 does not support skeletal muscle lineage in alveolar soft part sarcoma. Arch Pathol Lab Med 123:503-507. 10.1043/0003-9985(1999)123<0503:IPOMAM>2.0.CO;2

Goodwin ML, Jin H, Straessler K, Smith-Fry K, Zhu JF, Monument MJ, Grossmann A, Randall RL, Capecchi MR, and Jones KB. 2014. Modeling alveolar soft part sarcomagenesis in the mouse: a role for lactate in the tumor microenvironment. Cancer Cell 26:851-862. 10.1016/j.ccell.2014.10.003
Grogan SP, Barbero A, Diaz-Romero J, Cleton-Jansen AM, Soeder S, Whiteside R, Hogendoorn PC, Farhadi J, Aigner T, Martin I, and Mainil-Varlet P. 2007. Identification of markers to characterize and sort human articular chondrocytes with enhanced in vitro chondrogenic capacity. *Arthritis Rheum* 56:586-595. 10.1002/art.22408

Hayman AR, Bune AJ, Bradley JR, Rashbass J, and Cox TM. 2000. Osteoclastic tartrate-resistant acid phosphatase (Acp 5): its localization to dendritic cells and diverse murine tissues. *J Histochem Cytochem* 48:219-228. 10.1177/002215540048002007

Heberle H, Meirelles GV, da Silva FR, Telles GP, and Minghim R. 2015. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 16:169. 10.1186/s12859-015-0611-3

Hirobe M, Masumori N, Tanaka T, Kitamura H, and Tsukamoto T. 2013. Establishment of an ASPL-TFE3 renal cell carcinoma cell line (S-TFE). *Cancer Biol Ther* 14:502-510. 10.4161/cbt.24344

Hortells L, Johansen AKZ, and Yutzey KE. 2019. Cardiac Fibroblasts and the Extracellular Matrix in Regenerative and Nonregenerative Hearts. *J Cardiovasc Dev Dis* 6. 10.3390/jcdd6030029

Hoshino M, Ogose A, Kawashima H, Izumi T, Hotta T, Hatano H, Morita T, Otsuka H, Umezu H, Yanoma S, Tsukuda M, and Endo N. 2009. Molecular analyses of cell origin and detection of circulating tumor cells in the peripheral blood in alveolar soft part sarcoma. *Cancer Genet Cytogenet* 190:75-80. 10.1016/j.cancergenyto.2008.11.014

Hu X, Garcia M, Weng L, Jung X, Murakami JL, Kumar B, Warden CD, Todorov I, and Chen CC. 2016. Identification of a common mesenchymal stromal progenitor for the adult haematopoietic niche. *Nat Commun* 7:13095. 10.1038/ncomms13095

Huang AH, Motlekar NA, Stein A, Diamond SL, Shore EM, and Mauck RL. 2008. High-throughput screening for modulators of mesenchymal stem cell chondrogenesis. *Ann Biomed Eng* 36:1909-1921. 10.1007/s10439-008-9562-4

Islam MO, Kanemura Y, Tajria J, Mori H, Kobayashi S, Shofuda T, Miyake J, Hara M, Yamasaki M, and Okano H. 2005. Characterization of ABC transporter ABCB1 expressed in human neural stem/progenitor cells. *FEBS Lett* 579:3473-3480. 10.1016/j.febslet.2005.05.019

Jiang D, and Rinkevich Y. 2018. Defining Skin Fibroblastic Cell Types Beyond CD90. *Front Cell Dev Biol* 6:133. 10.3389/fcell.2018.00133

Jin HJ, Kwon JH, Kim M, Bae YK, Choi SJ, Oh W, Yang YS, and Jeon HB. 2016. Downregulation of Melanoma Cell Adhesion Molecule (MCAM/CD146) Accelerates Cellular Senescence in Human Umbilical Cord Blood-Derived Mesenchymal Stem Cells. *Stem Cells Transl Med* 5:427-439. 10.5966/sctm.2015-0109

Jun HJ, Lee J, Lim DH, Park JO, Ahn G, Seo SW, Sung KS, Lim DH, Yoo KH, and Choi YL. 2010. Expression of MET in alveolar soft part sarcoma. *Med Oncol* 27:459-465. 10.1007/s12032-009-9234-8

Karagiannis GS, Poutahidis T, Erdman SE, Kirsch R, Riddell RH, and Diamandis EP. 2012. Cancer-associated fibroblasts drive the progression of metastasis through both paracrine and mechanical pressure on cancer tissue. *Mol Cancer Res* 10:1403-1418. 10.1158/1541-7786.MCR-12-0307

Kenney S, Vistica DT, Stockwin LH, Burkett S, Hollingshead MG, Borgel SD, Butcher DO, Schrump DS, and Shoemaker RH. 2011. ASPS-1, a novel cell line manifesting key features of alveolar soft part sarcoma. *J Pediatr Hematol Oncol* 33:360-368. 10.1097/MPH.0b013e3182002f9f
Kim YJ, Yu JM, Joo HJ, Kim HK, Cho HH, Bae YC, and Jung JS. 2007. Role of CD9 in proliferation and proangiogenic action of human adipose-derived mesenchymal stem cells. *Pflugers Arch* 455:283-296. 10.1007/s00424-007-0285-4

Kobos R, Nagai M, Tsuda M, Merl MY, Saito T, Lae M, Mo Q, Olshen A, Lianoglou S, Leslie C, Ostrovnya I, Antczak C, Djaballah H, and Ladanyi M. 2013. Combining integrated genomics and functional genomics to dissect the biology of a cancer-associated, aberrant transcription factor, the ASPSCR1-TFE3 fusion oncoprotein. *J Pathol* 229:743-754. 10.1002/path.4158

Ksander BR, Kolovou PE, Wilson BJ, Saab KR, Guo Q, Ma J, McGuire SP, Gregory MS, Vincent WJ, Perez VL, Cruz-Guilloty F, Kao WW, Call MK, Tucker BA, Zhan Q, Murphy GF, Lathrop KL, Alt C, Mortensen LJ, Lin CP, Zieske JD, Frank MH, and Frank NY. 2014. ABCB5 is a limbal stem cell gene required for corneal development and repair. *Nature* 511:353-357. 10.1038/nature13426

Kuci S, Kuci Z, Schafer R, Spohn G, Winter S, Schwab M, Salzmann-Manrique E, Klingebiel T, and Bader P. 2019. Molecular signature of human bone marrow-derived mesenchymal stromal cell subsets. *Sci Rep* 9:1774. 10.1038/s41598-019-38517-7

Kulterer B, Friedl G, Jandrositz A, Sanchez-Cabo F, Prokesch A, Paar C, Scheideler M, Windhager R, Preisegger KH, and Trajanoski Z. 2007. Gene expression profiling of human mesenchymal stem cells derived from bone marrow during expansion and osteoblast differentiation. *BMC Genomics* 8:70. 10.1186/1471-2164-8-70

Kummar S, Allen D, Monks A, Polley EC, Hose CD, Ivy SP, Turkbey IB, Lawrence S, Kinders RJ, Choyke P, Simon R, Steinberg SM, Doroshow JH, and Helman L. 2013. Cediranib for metastatic alveolar soft part sarcoma. *J Clin Oncol* 31:2296-2302. 10.1200/JCO.2012.47.4288

La Rocca G, Anzalone R, and Farina F. 2009. The expression of CD68 in human umbilical cord mesenchymal stem cells: new evidences of presence in non-myeloid cell types. *Scand J Immunol* 70:161-162. 10.1111/j.1365-3083.2009.02283.x

Ladanyi M, Antonescu CR, Drobnjak M, Baren A, Lui MY, Golde DW, and Cordon-Cardo C. 2002. The precrystalline cytoplasmic granules of alveolar soft part sarcoma contain monocarboxylate transporter 1 and CD147. *Am J Pathol* 160:1215-1221. 10.1016/S0002-9440(10)62548-5

Ladanyi M, Lui MY, Antonescu CR, Krause-Boehm A, Meindl A, Argani P, Healey JH, Ueda T, Yoshikawa H, Meloni-Ehrig A, Sorensen PH, Mertens F, Mandahl N, van den Bergh H, Sciot R, Dal Cin P, and Bridge J. 2001. The der(17)(t(X;17)(p11;q25) of human alveolar soft part sarcoma fuses the TFE3 transcription factor gene to ASPL, a novel gene at 17q25. *Oncogene* 20:48-57. 10.1038/sj.onc.1204074

Lee HJ, Choi BH, Min BH, and Park SR. 2009a. Changes in surface markers of human mesenchymal stem cells during the chondrogenic differentiation and dedifferentiation processes in vitro. *Arthritis Rheum* 60:2325-2332. 10.1002/art.24786

Lee S, Park JR, Seo MS, Roh KH, Park SB, Hwang JW, Sun B, Seo K, Lee YS, Kang SK, Jung JW, and Kang KS. 2009b. Histone deacetylase inhibitors decrease proliferation potential and multilineage differentiation capability of human mesenchymal stem cells. *Cell Prolif* 42:711-720. 10.1111/j.1365-2184.2009.00633.x

McBride JD, Rodriguez-Menocal L, Guzman W, Candanedo A, Garcia-Contreras M, and Badiavas EV. 2017. Bone Marrow Mesenchymal Stem Cell-Derived CD63(+) Exosomes...
Transport Wnt3a Exteriorly and Enhance Dermal Fibroblast Proliferation, Migration, and Angiogenesis In Vitro. Stem Cells Dev 26:1384-1398. 10.1089/scd.2017.0087

Merrick D, Sakers A, Irgebay Z, Okada C, Calvert C, Morley MP, Percec I, and Seale P. 2019. Identification of a mesenchymal progenitor cell hierarchy in adipose tissue. Science 364. 10.1126/science.aav2501

Moller AMJ, Fuchtbauer EM, Bruel A, Andersen TL, Borggaard XG, Pavlos NJ, Thomsen JS, Pedersen FS, Delaisse JM, and Soe K. 2018. Septins are critical regulators of osteoclastic bone resorption. Sci Rep 8:13016. 10.1038/s41598-018-31159-1

Mukai M, Iri H, Nakajima T, Hirose S, Torikata C, Kageyama K, Ueno N, and Murakami K. 1983. Alveolar soft-part sarcoma. A review on its histogenesis and further studies based on electron microscopy, immunohistochemistry, and biochemistry. Am J Surg Pathol 7:679-689.

Naka A, Iida KT, Nakagawa Y, Iwasaki H, Takeuchi Y, Satoh A, Matsuzaka T, Ishii KA, Kobayashi K, Yatoh S, Shimada M, Yahagi N, Suzuki H, Sone H, Yamada N, and Shimano H. 2013. TFE3 inhibits myoblast differentiation in C2C12 cells via down-regulating gene expression of myogenin. Biochem Biophys Res Commun 430:664-669. 10.1016/j.bbrc.2012.11.094

Ogueta S, Munoz J, Obregon E, Delgado-Baeza E, and Garcia-Ruiz JP. 2002. Prolactin is a component of the human synovial liquid and modulates the growth and chondrogenic differentiation of bone marrow-derived mesenchymal stem cells. Mol Cell Endocrinol 190:51-63. 10.1016/s0303-7207(02)00013-8

Palomaki S, Pietila M, Laitinen S, Pesala J, Sormunen R, Lehenkari P, and Koivunen P. 2013. HIF-1alpha is upregulated in human mesenchymal stem cells. Stem Cells 31:1902-1909. 10.1002/stem.1435

Paoluzzi L, and Maki RG. 2019. Diagnosis, Prognosis, and Treatment of Alveolar Soft-Part Sarcoma: A Review. JAMA Oncol 5:254-260. 10.1001/jamaoncol.2018.4490

Pittenger MF, Discher DE, Peault BM, Phinney DG, Hare JM, and Caplan AI. 2019. Mesenchymal stem cell perspective: cell biology to clinical progress. NPJ Regen Med 4:22. 10.1038/s41536-019-0083-6

Portera CA, Jr., Ho V, Patel SR, Hunt KK, Feig BW, Respondek PM, Yasko AW, Benjamin RS, Pollock RE, and Pisters PW. 2001. Alveolar soft part sarcoma: clinical course and patterns of metastasis in 70 patients treated at a single institution. Cancer 91:585-591. 10.1002/1097-0142(20010201)91:3<585::aid-cncr1038>3.0.co;2-0

Ren G, Zhao X, Zhang L, Zhang J, L'Huillier A, Ling W, Roberts AI, Le AD, Shi S, Shao C, and Shi Y. 2010. Inflammatory cytokine-induced intercellular adhesion molecule-1 and vascular cell adhesion molecule-1 in mesenchymal stem cells are critical for immunosuppression. J Immunol 184:2321-2328. 10.4049/jimmunol.0902023

Rubenstein AB, Smith GR, Raue U, Begue G, Minchev K, Ruf-Zamojski F, Nair VD, Wang X, Zhou L, Zaslavsky E, Trappe TA, Trappe S, and Sealfon SC. 2020. Single-cell transcriptional profiles in human skeletal muscle. Sci Rep 10:229. 10.1038/s41598-019-57110-6

Saalbach A, and Anderegg U. 2019. Thy-1: more than a marker for mesenchymal stromal cells. FASEB J 33:6689-6696. 10.1096/fj.201802224R

Salloum RH, Rubin JP, and Marra KG. 2013. The role of steroids in mesenchymal stem cell differentiation: molecular and clinical perspectives. Horm Mol Biol Clin Investig 14:3-14. 10.1515/hmbci-2013-0016
Schoffski P, Wozniak A, Kasper B, Aamdal S, Leahy MG, Rutkowski P, Bauer S, Gelderblom H, Italiano A, Lindner LH, Hennig I, Strauss S, Zakotnik B, Anthony A, Albiges L, Blay JY, Reichardt P, Sufliarsky J, van der Graaf WTA, Debiec-Rychter M, Sciot R, Van Cann T, Marreaud S, Raveloarivahy T, Collette S, and Stacchiotti S. 2018. Activity and safety of crizotinib in patients with alveolar soft part sarcoma with rearrangement of TFE3: European Organization for Research and Treatment of Cancer (EORTC) phase II trial 90101 ‘CREATE’. *Ann Oncol* 29:758-765. 10.1093/annonc/mdx774

Selvarajah S, Pyne S, Chen E, Sompallae R, Ligon AH, Nielsen GP, Dranoff G, Stack E, Loda M, and Flavin R. 2014. High-resolution array CGH and gene expression profiling of alveolar soft part sarcoma. *Clin Cancer Res* 20:1521-1530. 10.1158/1078-0432.CCR-13-2090

Seriwatanachai D, Krishnamra N, and Charoenphandhu N. 2012. Chondroregulatory action of prolactin on proliferation and differentiation of mouse chondrogenic ATDC5 cells in 3-dimensional micromass cultures. *Biochem Biophys Res Commun* 420:108-113. 10.1016/j.bbrc.2012.02.123

Stockwin LH, Vistica DT, Kenney S, Schrupp DS, Butcher DO, Raffeld M, and Shoemaker RH. 2009. Gene expression profiling of alveolar soft-part sarcoma (ASPS). *BMC Cancer* 9:22. 10.1186/1471-2407-9-22

Stopp S, Bornhauser M, Ugarte F, Wobus M, Kuhn M, Brenner S, and Thieme S. 2013. Expression of the melanoma cell adhesion molecule in human mesenchymal stromal cells regulates proliferation, differentiation, and maintenance of hematopoietic stem and progenitor cells. *Haematologica* 98:505-513. 10.3324/haematol.2012.065201

Surarit R, Krishnamra N, and Seriwatanachai D. 2016. Prolactin receptor and osteogenic induction of prolactin in human periodontal ligament fibroblasts. *Cell Biol Int* 40:419-427. 10.1002/cbin.10580

Tan SL, Ahmad TS, Ng WM, Azlina AA, Azhar MM, Selvaratnam L, and Kamarul T. 2015. Identification of Pathways Mediating Growth Differentiation Factor5-Induced Tenogenic Differentiation in Human Bone Marrow Stromal Cells. *PLoS One* 10:e0140869. 10.1371/journal.pone.0140869

Tanaka M, Homme M, Yamazaki Y, Shimizu R, Takazawa Y, and Nakamura T. 2017. Modeling Alveolar Soft Part Sarcoma Unveils Novel Mechanisms of Metastasis. *Cancer Res* 77:897-907. 10.1158/0008-5472.CAN-16-2486

Tanoue H, Morinaga J, Yoshizawa T, Yugami M, Itoh H, Nakamura T, Uehara Y, Masuda T, Odagiri H, Sugizaki T, Kadomatsu T, Miyata K, Endo M, Terada K, Ochi H, Takeda S, Yamagata K, Fukuda T, Mizuta H, and Oike Y. 2018. Angiopoietin-like protein 2 promotes chondrogenic differentiation during bone growth as a cartilage matrix factor. *Osteoarthritis Cartilage* 26:108-117. 10.1016/j.joca.2017.10.011

Teicher BA, Polley E, Kunkel M, Evans D, Silvers T, Delosh R, Laudeman J, Ogle C, Reinhart R, Selby M, Connelly J, Harris E, Monks A, and Morris J. 2015. Sarcoma Cell Line Screen of Oncology Drugs and Investigational Agents Identifies Patterns Associated with Gene
and microRNA Expression. *Mol Cancer Ther* 14:2452-2462. 10.1158/1535-7163.MCT-15-0074

Tsuda M, Davis IJ, Argani P, Shukla N, McGill GG, Nagai M, Saito T, Lae M, Fisher DE, and Ladanyi M. 2007. TFE3 fusions activate MET signaling by transcriptional up-regulation, defining another class of tumors as candidates for therapeutic MET inhibition. *Cancer Res* 67:919-929. 10.1158/0008-5472.CAN-06-2855

Vander Beken S, de Vries JC, Meier-Schiesser B, Meyer P, Jiang D, Sindrilaru A, Ferreira FF, Hainzl A, Schatz S, Muschhammer J, Scheurmann NJ, Kampilafkos P, Seitz AM, Durselen L, Ignatius A, Kluth MA, Ganss C, Wlaschek M, Singh K, Maity P, Frank NY, Frank MH, and Scharffetter-Kochanek K. 2019. Newly Defined ATP-Binding Cassette Subfamily B Member 5 Positive Dermal Mesenchymal Stem Cells Promote Healing of Chronic Iron-Overload Wounds via Secretion of Interleukin-1 Receptor Antagonist. *Stem Cells* 37:1057-1074. 10.1002/stem.3022

Vrijsen KR, Maring JA, Chamuleau SA, Verhage V, Mol EA, Deddens JC, Metz CH, Lodder K, van Eeuwijk EC, van Dommelen SM, Smits AM, Goumans MJ, and Sluijter JP. 2016. Exosomes from Cardiomyocyte Progenitor Cells and Mesenchymal Stem Cells Stimulate Angiogenesis Via EMMPRIN. *Adv Healthc Mater* 5:2555-2565. 10.1002/adhm.201600308

Wang NP, Bacchi CE, Jiang JJ, McNutt MA, and Gown AM. 1996. Does alveolar soft-part sarcoma exhibit skeletal muscle differentiation? An immunocytochemical and biochemical study of myogenic regulatory protein expression. *Mod Pathol* 9:496-506.

Wang S, Zhou Y, Andreyev O, Hoyt RF, Jr., Singh A, Hunt T, and Horvath KA. 2014. Overexpression of FABP3 inhibits human bone marrow derived mesenchymal stem cell proliferation but enhances their survival in hypoxia. *Exp Cell Res* 323:56-65. 10.1016/j.yexcr.2014.02.015

Yu B, Sondag GR, Malcuit C, Kim MH, and Safadi FF. 2016. Macrophage-Associated Osteoactivin/GPNMB Mediates Mesenchymal Stem Cell Survival, Proliferation, and Migration Via a CD44-Dependent Mechanism. *J Cell Biochem* 117:1511-1521. 10.1002/jcb.25394

Zhao SL, Zhang YJ, Li MH, Zhang XL, and Chen SL. 2014. Mesenchymal stem cells with overexpression of midkine enhance cell survival and attenuate cardiac dysfunction in a rat model of myocardial infarction. *Stem Cell Res Ther* 5:37. 10.1186/scrt425
Study Design

This study focuses on the analysis of four publicly available GEO gene-expression datasets. GSE68591 comprises exon level expression data for the NCI sarcoma cell line panel, which includes data for the ASPS cell line ASPS-1. The remaining three studies comprise transcriptomic data for ASPS patient samples. GSE13433 comprises Affymetrix U133 plus 2.0 microarray data from our initial gene expression study of seven ASPS patients. GSE32569 uses the same array platform to study ASPS patient sample responses to Cedirinib. Lastly, GSE54729 comprises Illumina Hiseq 2000 RNA-seq data for ASPS patient samples generated as part of an ASPS mouse modeling study. These data were re-analyzed using appropriate controls in order to generate a consensus transcriptome and gain insights into ASPS pathobiology.
ASPS transcriptomic studies identified through GEO search (n=4)

- **GSE68591**
  - "Exon expression for NCI Sarcoma cell line panel"
  - **ASPS-1 Cell Line**
  - Affymetrix Human Exon 1.0 ST Array

- **GSE13433**
  - "Gene Expression Profiling of Alveolar Soft-Part Sarcoma (ASPS)"
  - **ASPS patient samples**
  - Affymetrix Human U133 Plus 2.0

- **GSE32569**
  - "Gene expression from pre- and post-Cediranib treated patients with metastatic Alveolar Soft Part Sarcoma (ASPS)"
  - **ASPS patient samples**
  - Affymetrix Human U133 Plus 2.0

- **GSE54729**
  - "Transcriptome profiling of a mouse model of alveolar soft part sarcoma"
  - **ASPS patient samples**
  - Illumina HiSeq 2000 (Homo sapiens)

  - **FKPM Differential analysis**
  - Expression relative to skeletal muscle control

Expression of myogenesis-related transcripts?

- **ASPS-1 mRNA Surfaceome**

Platform-Independent Consensus ASPS Transcriptome
Figure 2

Analysis of ASPS-1 transcriptomic data.

(A) Hierarchical clustering of the NCI sarcoma cell line panel exon array data (B) tissue or tumor derivation for cluster nearest ASPS-1. (C) GTEx tissue of origin analysis for top fifty ASPS-1 transcripts.
Figure 3

Expression of myogenesis-related transcripts in ASPS-1 relative to the other sarcoma cell lines

Cell lines in the NCI sarcoma panel were segregated according to disease type and the average of exon expression data plotted for transcripts encoding myogenesis-related transcription factors and muscle structural proteins. ASPS-1 data is shown first; then RMS = Rhabdomyosarcoma, RT = Rhabdoid tumor, LMS = Leiomyosarcoma, CS = Chondrosarcoma, FS = Fibrosarcoma, EWS = Ewings Sarcoma, OS = Osteosarcoma, SS = Synovial Sarcoma, US = Uterine Sarcoma, SNOS = Sarcoma not otherwise specified, GCS = Giant cell sarcoma, MNS = Malignant peripheral nerve sheath, LPS = Liposarcoma, SPS = Spindle cell sarcoma, EP = Epithelioid, NC = Normal cells, SkMc = skeletal muscle cells. For transcripts; MYF5 = Myogenic Factor 5, MYF6 = Myogenic Factor 6, MYOD1 = Myogenic Differentiation 1, MYOG = Myogenin, PAX3 = Paired Box 3, PAX7 = Paired Box 7. Structural proteins; DES = desmin, NEB = nebulin, TNNT1 = Troponin T1 - Slow Skeletal Type, TRIM63 = Tripartite Motif Containing 63, TTN = titin and MSTN = Myostatin. Transcripts evaluated but not shown included; EYA1, LBX1, MEF2B, MEOX2, MITF, MSX1, PITX1, SIM2, SIX1, SIX4, TFE3 and TFEB.
Figure 4

VENN diagram showing overlap between analyses.

Lists of over-expressed transcripts (five-fold cut off) were used to determine extent of overlap between the five datasets. The number of differentially expressed transcripts at five-fold is underlined. Callouts show the 25 transcripts over-expressed in all experiments and the 73 found in all ASPS patient analyses.
Table 1 (on next page)

The top fifty over-expressed genes in ASPS-1 relative to the average of the NCI sarcoma cell line panel.
| Gene Symbol          | Affy ID | Description                                                                 | FCASPS-1 vs. Av |
|---------------------|---------|------------------------------------------------------------------------------|----------------|
| CRYAB; FDXACB1       | 3391149 | crystallin alpha B; ferredoxin-fold anticodon binding domain containing 1    | 227.1          |
| MT1G                | 3692999 | metallothionein 1G                                                             | 202.9          |
| C7orf69             | 3000905 | chromosome 7 open reading frame 69                                             | 202.0          |
| GCSAML              | 2390102 | germinal center-associated, signaling and motility-like                        | 177.6          |
| SV2B                | 3608638 | synaptic vesicle glycoprotein 2B                                              | 132.6          |
| ADGRL4              | 2419432 | adhesion G protein-coupled receptor 4                                         | 129.9          |
| PLASG7              | 2955827 | phospholipase A2, group VII                                                   | 93.5           |
| SULT1C2             | 2498911 | sulfotransferase family 1C member 2                                           | 91.3           |
| SLN                 | 3388954 | sarcolipin                                                                    | 88.9           |
| CFAEP61             | 3878972 | cilia and flagella associated protein 61                                      | 85.2           |
| PPEF1               | 3970714 | protein phosphatase, EF-hand calcium binding domain 1                         | 82.4           |
| ACP5                | 3851072 | acid phosphatase 5, tartrate resistant                                        | 79.8           |
| CD36                | 3010503 | CD36 molecule (thrombospondin receptor)                                       | 79.1           |
| MT1G                | 3360239 | metallothionein 1G                                                             | 78.9           |
| ASB11               | 4000485 | ankyrin repeat and SOCS box containing 11, E3 ubiquitin protein ligase        | 78.1           |
| BMP5                | 2958172 | bone morphogenetic protein 5                                                  | 74.9           |
| PRUNE2              | 3210616 | prune homolog 2 (Drosophila)                                                  | 72.3           |
| SUCNR1              | 2648098 | succinate receptor 1                                                           | 66.5           |
| PSF9                | 3864286 | pregnancy specific beta-1-glycoprotein 9                                      | 66.0           |
| CKMT2               | 2818035 | creatine kinase, mitochondrial 2 (sarcomeric)                                 | 61.7           |
| DPP4                | 2584018 | dipeptidyl-peptidase 4                                                        | 61.2           |
| ABC81               | 3060283 | ATP binding cassette subfamily B member 1                                     | 59.7           |
| SCIN                | 2990404 | scinderin                                                                     | 58.7           |
| FABP3               | 2404148 | fatty acid binding protein 3, muscle and heart                                | 55.8           |
| PRUNE2              | 3210497 | prune homolog 2 (Drosophila)                                                  | 55.7           |
| SLC27A2             | 3593575 | solute carrier family 27 (fatty acid transporter), member 2                   | 54.7           |
| SLC24C1             | 2869096 | solute carrier organic anion transporter family, member 4C1                   | 51.9           |
| PSG11,5,4,2         | 3863929 | pregnancy specific beta-1-glycoprotein 11,5,4,2                               | 50.4           |
| PRLR                | 2853102 | prolactin receptor                                                            | 47.9           |
| NPY6R               | 2830450 | neuropeptide Y receptor Y6 (pseudogene)                                       | 44.6           |
| ANXA3               | 2732844 | annexin A3                                                                    | 43.8           |
| TRPC7               | 2876793 | transient receptor potential cation channel, subfamily C, member 7            | 42.9           |
| CDS5                | 2439138 | CDS molecule-like                                                             | 41.0           |
| AKR1C2              | 3274758 | aldo-keto reductase family 1, member C2                                        | 40.9           |
| GPMB                | 2992814 | glycoprotein (transmembrane) nmb                                              | 40.3           |
| IL1RA2              | 4018729 | interleukin 13 receptor, alpha 2                                              | 39.1           |
| LRRIC3              | 2425173 | leucine rich repeat containing 39                                             | 38.8           |
| CST1                | 3901361 | cystatin SN                                                                   | 37.4           |
| CDH7                | 3792273 | cadherin 7, type 2                                                            | 36.8           |
| DOK5                | 3889833 | docking protein 5                                                              | 35.1           |
| SEPT3; WBP2NL        | 3947227 | septin 3; WBP2 N-terminal like                                                 | 34.8           |
| GCNT3               | 3596147 | glucosaminyl (N-acetyl) transferase 3, mucin type                             | 34.8           |
| ENP5                | 2955673 | ectonucleotide pyrophosphatase/phosphodiesterase 5 (putative)                 | 34.7           |
| DOK3                | 2888879 | docking protein 3                                                              | 34.2           |
| LCP1                | 3512874 | lymphocyte cytotoxic protein 1 (L-plastin)                                     | 34.0           |
| CDA                 | 2324084 | cytidine deaminase                                                            | 33.4           |
| KHL4                | 3988324 | kelch-like family member 4                                                    | 33.3           |
| CTSK                | 2434609 | cathepsin K                                                                   | 31.5           |
| LIPC                | 3593691 | lipase, hepatic                                                                | 31.0           |
| RPSA                | 3827218 | ribosomal protein SA                                                          | 30.0           |
Table 2 (on next page)

The top fifty upregulated transcripts generated from each of the four experiments utilizing ASPS patient data from GSE13433, GSE32569 and GSE54729
| Ensembl ID | FC | GENE ID | AFFY ID | FC | GENE ID | AFFY ID | FC | GENE ID | AFFY ID | FC | GENE ID |
|------------|----|---------|---------|----|---------|---------|----|---------|---------|----|---------|
| 210397_at  | 71.0 | AHNAK2  | 205852_at | 61.8 | PCSK1  | 207733_s_at | 60.8 | PSG9  | 205052_at | 420.6 | CYP17A1 |
| 205852_at  | 61.8 | PCSK1  | 205052_at | 420.6 | CYP17A1 | 205776_s_at | 365.1 | PCYT1  | 205511_at | 85.1 | SV2B  |
| 205511_at  | 85.1 | SV2B  | 205545_at | 572.3 | PRL  | 204830_at | 106.0 | ABCC5  | 204830_at | 194.0 | ETP1  |
| 204830_at  | 106.0 | ABCC5  | 204830_at | 194.0 | ETP1  | 207376_s_at | 365.1 | PCYT1  | 205511_at | 85.1 | SV2B  |
| 205511_at  | 85.1 | SV2B  | 205545_at | 572.3 | PRL  | 204830_at | 106.0 | ABCC5  | 204830_at | 194.0 | ETP1  |