



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

Marc Ladanyi^{*1,2}, Man Yee Lui¹, Cristina R Antonescu¹, Amber Krause-Boehm³, Alfons Meindl⁴, Pedram Argani⁵, John H Healey⁶, Takafumi Ueda⁷, Hideki Yoshikawa⁷, Aurelia Meloni-Ehrig⁸, Poul HB Sorensen⁹, Fredrik Mertens¹⁰, Nils Mandahl¹⁰, Herman van den Berghe¹¹, Raf Sciot¹², Paola Dal Cin¹² and Julia Bridge^{3,13,14}

¹Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York, USA; ²Department of Human Genetics, Memorial Sloan-Kettering Cancer Center, New York, New York, USA; ³Department of Pathology, CHMG/MMI, University of Nebraska Medical Center, Omaha, Nebraska, USA; ⁴Department of Medical Genetics, Ludwig-Maximilians-Universität, München, Germany; ⁵Department of Pathology, The Johns Hopkins Hospital, Baltimore, Maryland, USA; ⁶Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, New York, USA; ⁷Department of Orthopedic Surgery, Osaka Medical Center for Cancer and Cardiovascular Diseases, Japan; ⁸Department of Human Genetics, Huntsman Cancer Institute, University of Utah Medical Center, Salt Lake City, Utah, USA; ⁹Department of Pathology, British Columbia's Children's Hospital, Vancouver, Canada; ¹⁰Department of Clinical Genetics, University Hospital, Lund, Sweden; ¹¹Center for Human Genetics, University of Leuven, Belgium; ¹²Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts, USA; ¹³Department of Pediatrics, CHMG/MMI, University of Nebraska Medical Center, Omaha, Nebraska, USA; ¹⁴Department of Orthopedic Surgery, CHMG/MMI, University of Nebraska Medical Center, Omaha, Nebraska, USA

Alveolar soft part sarcoma (ASPS) is an unusual tumor with highly characteristic histopathology and ultrastructure, controversial histogenesis, and enigmatic clinical behavior. Recent cytogenetic studies have identified a recurrent der(17) due to a non-reciprocal t(X;17)(p11.2;q25) in this sarcoma. To define the interval containing the Xp11.2 break, we first performed FISH on ASPS cases using YAC probes for *OATL1* (Xp11.23) and *OATL2* (Xp11.21), and cosmid probes from the intervening genomic region. This localized the breakpoint to a 160 kb interval. The prime candidate within this previously fully sequenced region was *TFE3*, a transcription factor gene known to be fused to translocation partners on 1 and X in some papillary renal cell carcinomas. Southern blotting using a *TFE3* genomic probe identified non-germline bands in several ASPS cases, consistent with rearrangement and possible fusion of *TFE3* with a gene on 17q25. Amplification of the 5' portion of cDNAs containing the 3' portion of *TFE3* in two different ASPS cases identified a novel sequence, designated *ASPL*, fused in-frame to *TFE3* exon 4 (type 1 fusion) or exon 3 (type 2 fusion). Reverse transcriptase PCR using a forward primer from *ASPL* and a *TFE3* exon 4 reverse primer detected an *ASPL-TFE3* fusion transcript in all ASPS cases (12/12: 9 type 1, 3 type 2), establishing the utility of this assay in the diagnosis of ASPS. Using appropriate primers, the reciprocal fusion transcript, *TFE3-ASPL*, was detected in only one of 12 cases, consistent with the non-reciprocal nature of the translocation in most cases, and supporting *ASPL-TFE3*

as its oncogenically significant fusion product. *ASPL* maps to chromosome 17, is ubiquitously expressed, and matches numerous ESTs (Unigene cluster Hs.84128) but no named genes. The *ASPL* cDNA open reading frame encodes a predicted protein of 476 amino acids that contains within its carboxy-terminal portion of a UBXL-like domain that shows significant similarity to predicted proteins of unknown function in several model organisms. The *ASPL-TFE3* fusion replaces the N-terminal portion of *TFE3* by the fused *ASPL* sequences, while retaining the *TFE3* DNA-binding domain, implicating transcriptional deregulation in the pathogenesis of this tumor, consistent with the biology of several other translocation-associated sarcomas. *Oncogene* (2001) 20, 48–57.

Keywords: chromosomal translocation; gene fusion; molecular diagnosis; UBXL domain; soft tissue sarcoma

Introduction

Alveolar soft part sarcoma (ASPS) is an unusual tumor with highly characteristic histopathology (Figure 1) and ultrastructure, controversial histogenesis, and enigmatic clinical behavior (Lieberman *et al.*, 1989; Ordonez, 1999). Most cases of ASPS occur in the second and third decade of life, with a slight female predilection (Lieberman *et al.*, 1989; Ordonez, 1999). It usually involves the muscle and deep soft tissues of the extremities, but has also been reported in tissues where skeletal muscle is absent. Its cell of origin and biology have remained unclear. Ultrastructural analysis shows secretory-like granules and, in about half of cases, characteristic rectangular or rhomboid crystalline cytoplasmic deposits of unknown composition (Ordonez, 1999). The treatment is primarily surgical (Lieberman *et al.*, 1989). Distant metastases are common but often

*Correspondence: M Ladanyi, Department of Pathology, Memorial Sloan-Kettering Cancer Center, 1275 York Ave., New York, NY 10021, USA

Received 25 August 2000; revised 25 October 2000; accepted 1 November 2000

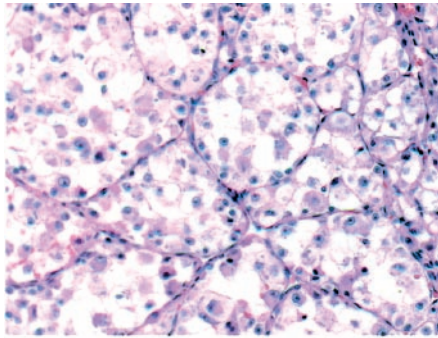


Figure 1 (a) Typical histology of ASPS showing well-defined nests of cells with abundant pink cytoplasm. The loss of central cohesion produces a pseudoalveolar appearance. (Case ASPS-1, hematoxylin and eosin stained section). (b) Partial karyotype of tumor ASPS-6 which showed the abnormal chromosomal complement: 46,XX,der(17)t(X;17)(p11.2;q25)

indolent. ASPS appears refractory to all chemotherapy agents tested so far. Although prolonged survival is possible even in patients with metastases, the long term disease-specific mortality is high (Lieberman *et al.*, 1989).

Cytogenetic studies of ASPS prior to 1998 (reviewed in Joyama *et al.*, 1999) described a recurrent add(17)(q25) representing unidentified additional chromosomal material at band 17q25. In 1998, Heimann *et al.* (1998) defined this alteration more clearly as a der(17)t(X;17)(p11;q25), and confirmed its non-reciprocal nature by fluorescent *in situ* hybridization (FISH) analysis showing loss of 17q25-qter sequences. Re-evaluation of prior published cases established this as a specific recurrent unbalanced translocation in ASPS (Heimann *et al.*, 1998; Joyama *et al.*, 1999). The localization of the chromosome X breakpoint was subsequently refined to Xp11.2 (Joyama *et al.*, 1999). We show here that this translocation fuses the *TFE3* transcription factor gene at Xp11 to a novel gene at 17q25, designated *ASPL*, in all cases of ASPS tested, implicating transcriptional deregulation in the pathogenesis of this tumor. *ASPL* is widely expressed in all adult tissues and encodes a predicted protein of unknown function containing a conserved domain possibly related to the ubiquitylation pathway.

Results

Cytogenetic data

Basic clinical data are presented in Table 1 for the 12 ASPS cases with material available for the molecular

Table 1 ASPS cases with molecular data

ASPS Case#	Origin	Age, sex	Primary site	ASPL- <i>TFE3</i> fusion type
1	USA	19 F	Thigh	1
2	USA	36 F	Axilla	2
3	USA	31 F	Thigh	1
4	USA	38 M	Arm	2
5	USA	29 F	Thigh	1
6	USA	14 F	Thigh	1
7	USA	40 F	Thigh	1
8	Sweden	40 M	Hand	2
9	Japan	7 M	Thigh	1
10	Japan	28 F	Thigh	1
11	Canada	9 M	Lower leg	1
12	USA	20 F	Thigh	1

studies described below. Cases ASPS-6, ASPS-8, and ASPS-9 from Table 1 had informative cytogenetics. ASPS-6 showed 46,XX,der(17)t(X;17)(p11;q25) [6] (Figure 1). The karyotype of ASPS-8 was 45,Y,del(X)(p11),dic(1;14)(p11;p11),+5,der(7)t(1;7)(q11;q36),-13,der(17)t(X;17)(p11;q25),+20,-22 [6] (not shown). In the latter case, we could not exclude cytogenetically the presence of the very distal part of 17q on the der(X) because there was no material available for FISH analysis, but the molecular studies reported below support the unbalanced nature of the t(X;17) in this case. In contrast to most other ASPS, there was no net gain of Xp material in ASPS-8. The karyotype in case ASPS-9, 46,XY,der(17)t(X;17)(p11.2;q25) [23], has been reported previously (Joyama *et al.*, 1999).

New cytogenetic data were available in one additional previously published case of a thigh tumor in a 21 year old woman (Scioto *et al.*, 1993), in which the 17q abnormality was originally described as add(17)(q25). The revised karyotype in this case is: 45,XX,del(1)(p11),der(9)t(9;15)(p11;q11), der(17)t(X;17)(p11;q25),-22. This case had insufficient material for the molecular analysis described below and is not included in Table 1. Nonetheless, all four cases whose karyotypes are described here contained an unbalanced der(17)t(X;17)(p11;q25), confirming the consistency of this finding in ASPS samples with informative cytogenetics.

FISH analysis of Xp11 breakpoint

To define the interval containing the Xp11.2 rearrangement, we first performed FISH using probes from the *OATL1* (Xp11.23) and *OATL2* (Xp11.21) regions. A single unstained metaphase cell preparation from case ASPS-6 was available for tricolor FISH studies. Tricolor FISH performed on this case revealed the presence of the *OATL1* YAC2 probe on the der(17) chromosome as well as on the two cytogenetically normal X chromosomes in 10 metaphase cells analysed (Figure 2). This was confirmed by bicolor FISH experiments performed on cytologic touch preparations of cases ASPS-1 and ASPS-2 showing three signals for the *OATL1* probe and two signals for an alloid sequence probe specific for the

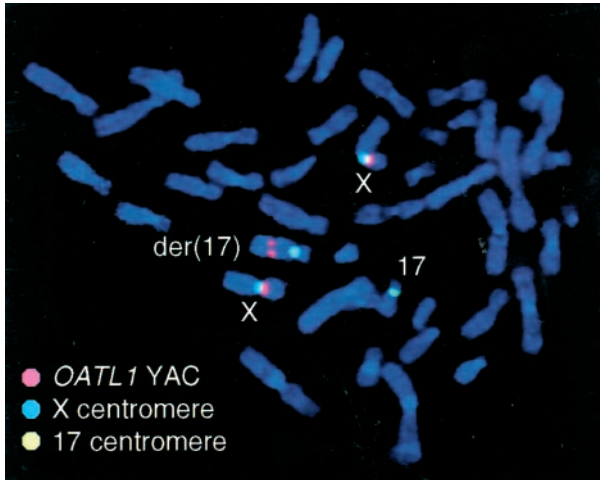


Figure 2 Tricolor metaphase FISH analysis of ASPS-6 shows an *OATL1* probe signal (red) on the derivative chromosome 17 (green) in addition to a signal on each of the normal X chromosome X homologs (aqua)

X centromere in the vast majority of cells analysed (Figure 3). Similar interphase FISH studies using the *OATL2* probe showed only two signals, both associated with the X centromere (Figure 3) in ASPS-1 and ASPS-2. These results were consistent with an unbalanced translocation involving a gene between *OATL1* and *OATL2*, a region of approximately 4 Mb. To further narrow the interval we performed FISH on ASPS-1 and ASPS-6 using four cosmids (A02167, U164B6, B1125, I1135) from a previously established 1.1 Mb cosmid contig immediately centromeric to the region covered by *OATL1* YAC2 (Schindelbauer *et al.*, 1996). Bicolor interphase FISH experiments revealed three signals for probes A02167 and U164B6, and two signals for the centromere X-specific probe in the vast majority of cells analysed suggesting that, similar to *OATL1*, these loci were present on the der(17) in addition to the two normal X chromosomes (Figure 3). Only two signals for B1125 could be detected, however, indicating that this locus was proximal to the breakpoint on the X chromosome (Figure 3). This allowed us to localize the breakpoint to a 160 kb interval between cosmids U164B6 and B1125. The sequencing of this region is essentially complete (GenBank #AF207550). A previous detailed sequence analysis identified five genes within this region, including UDP-galactose transporter, *PIM-2* protooncogene homolog *PIM-2H*, shal-type potassium channel, JM12 protein, and *TFE3* (project p23-3 at <http://genome.imb-jena.de/>). Among these, *TFE3*, a transcription factor gene known to be fused to translocation partners on 1 and X in a subset of papillary renal adenocarcinomas (Clark *et al.*, 1997; Sidhar *et al.*, 1996; Weterman *et al.*, 1996), was considered as the prime candidate. The orientation of *TFE3* is with its 3' end telomeric. Further interphase FISH analysis with cosmid I1135, which is located

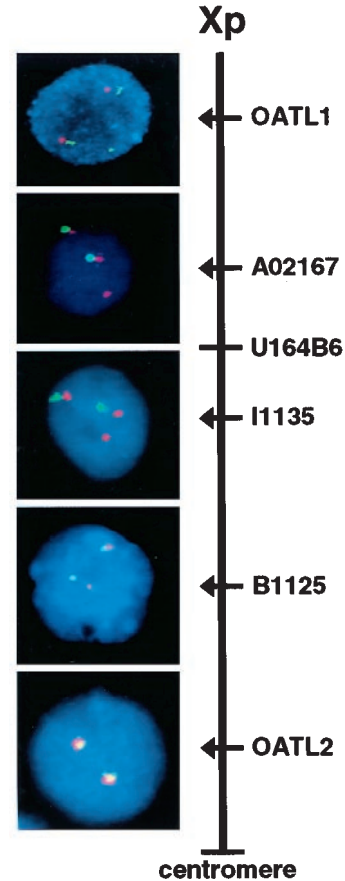


Figure 3 FISH studies of Xp11 narrowing the breakpoint region. The relative order of the Xp11.2 probes used is shown in the schematic diagram (right). Bicolor FISH analyses with YAC probe *OATL1* (in green with CEP X in red; case ASPS-1) and with cosmid probes A02167 or I1135 (in red with CEP X or DXZ1 in green; case ASPS-6) show three signals for the *OATL1*, A02167 and I1135 probes but only two signals for the X centromeric probe indicating extra copies of the *OATL1*, A02167 and I1135 loci (likely on the der(17)). Equivalent results were obtained with cosmid probe U164B6 (not shown). In contrast, bicolor FISH studies using probes B1125 and *OATL2* (in green with CEP X or DXZ1 in red; case ASPS-6) show the normal copy number signals for B1125 and *OATL2* adjacent to the X centromeric signals. These results narrow the breakpoint region to the I1135-B1125 interval containing the *TFE3* gene (see text)

within this 160 kb interval and known by sequence analysis to contain the 3' end of *TFE3* (exons 4–8, see GenBank #AF207550), also showed three signals, consistent with its translocation to chromosome 17 (Figure 3).

Southern blot analysis of *TFE3*

Because *TFE3* rearrangements in papillary renal adenocarcinomas most often occur within intron 3 (and less commonly within intron 1) (Clark *et al.*, 1997; Sidhar *et al.*, 1996), we used the available genomic sequence of *TFE3* to design a 728 bp repeat-free probe from intron 3 of *TFE3*. This *TFE* intron 3 probe was synthesized by PCR on normal DNA using primers ATCAAGCAGATTCCCTGACAC and TAGGCT-

GACATAAAGCAGTGC. Southern blotting using this probe identified rearranged bands in multiple enzyme digests in ASPS-1 (Figure 4) and several other ASPS cases (results not shown). Mapping of the breakpoint in case ASPS-1 narrowed the breakpoint to a 4.4 kb *PvuII* restriction fragment corresponding essentially to *TFE3* intron 3 (Figure 5). The genomic breaks in other ASPS cases were not systematically mapped.

Isolation of a novel fusion transcript containing the 3' end of TFE3

Based on the mapping of the breakpoint to *TFE3* intron 3 and on the observation that the functional products of *TFE3* rearrangements in papillary renal adenocarcinomas fuse the 3' portion of *TFE3* (including at least *TFE3* exon 4 and more 3' regions) to 5' portions of the different translocation partners (Clark et al., 1997; Sidhar et al., 1996), we performed a 5' rapid amplification of cDNA ends (RACE) procedure from the 3' portion of *TFE3* on total RNA from cases ASPS-1 and ASPS-8. Reverse transcription was primed using a *TFE3* exon 5 or exon 7 reverse primer, followed by heminested PCR with nested *TFE3* exon 4 reverse primers (see Materials and methods). Among multiple different RACE products obtained from both cases, some corresponded to normal *TFE3*, but in both

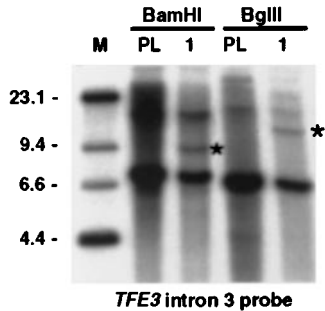


Figure 4 *TFE3* genomic rearrangement in case ASPS-1. Southern blot analysis with a *TFE3* intron 3 probe shows non-germline bands in both enzyme digests of ASPS-1 genomic DNA, compared to the placental control lanes (PL)

cases transcripts containing the same novel sequence fused in-frame to *TFE3* exon 3 (in ASPS-8) or exon 4 (in ASPS-1) were obtained (Figure 6). This novel sequence, which we designated *ASPL* (alveolar soft part sarcoma locus), did not match any named genes in GenBank. The full symbol approved by the HUGO Gene Nomenclature Committee is *ASPSCR1*.

RT-PCR detection of ASPL-TFE3 fusion transcripts

To confirm the consistent presence of an *ASPL-TFE3* fusion transcript in these tumors, we performed reverse transcriptase PCR (RT-PCR) using a forward primer from the *ASPL* sequence (AAAGAAGTC-CAAGTCGGGCCA) and *TFE3* exon 4 reverse primer (CGTTTGATGTTGGGCAGCTCA). A sequence-ver-

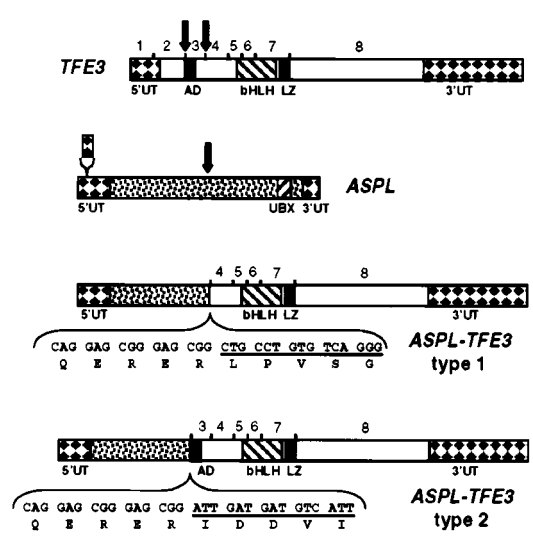


Figure 6 Schematic diagram of normal *TFE3* and *ASPL-TFE3* fusion products (type 1 and type 2) identified by 5' RACE with sequences of fusion junctions. Relative dimensions are approximate. Arrows indicate fusion points in *TFE3* and *ASPL* cDNAs. A minor form of the *ASPL* transcript contains a 47 nt alternative splice within the 5' untranslated region. UT: untranslated, AD: activation domain, bHLH: basic helix-loop-helix DNA-binding domain, LZ: leucine zipper dimerization domain, UBX: UBX domain

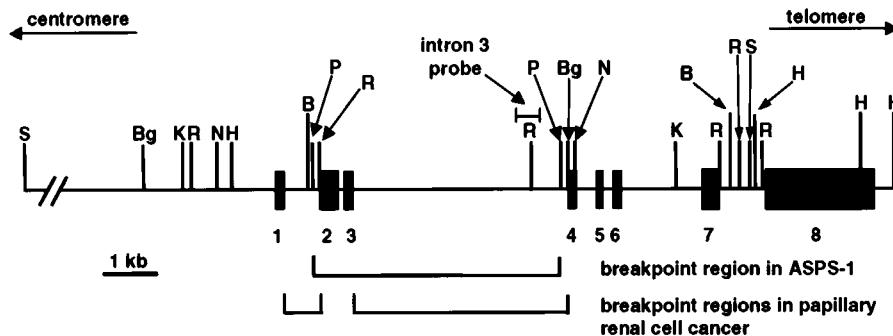


Figure 5 *TFE3* genomic map. Restriction sites are derived from published data (Macchi et al., 1995; Sidhar et al., 1996) and from the complete genomic sequence (GenBank #AF207550). For some enzymes only the sites bracketing the region probed are shown. The smallest rearranged restriction fragment detected using the intron 3 probe in case ASPS-1 was the 4.4 kb *PvuII* fragment. The locations of genomic breaks previously described in papillary renal cancers (see text) are also shown

ified *ASPL-TFE3* fusion transcript was detected in all ASPS cases (12/12) (Table 1), including the two cases used for 5' RACE, but was absent in control RNA (K562 leukemia cell line) (Figure 7), establishing the potential clinical utility of this RT-PCR assay in confirming the diagnosis of ASPS. Two mutually exclusive types of *ASPL-TFE3* fusion transcripts were observed, corresponding to the two types of RACE products described above: in three cases, the fusion transcript included *TFE3* exon 3 ('type 2') while in the remaining nine cases it did not ('type 1').

RT-PCR for the reciprocal fusion product, *TFE3-ASPL*, using *TFE3* forward and *ASPL* reverse primers appropriate to the type of *ASPL-TFE3* rearrangement in each case, was positive in ASPS-7 and negative in the other 11 cases, including ASPS-8 (results not shown), consistent with the non-reciprocal nature of this translocation observed cytogenetically in most cases to date.

We also evaluated the expression of the remaining unrearranged alleles of *TFE3* and *ASPL* by RT-PCR using appropriate *TFE3* or *ASPL* primer pairs flanking the *ASPL-TFE3* fusion point (results not shown). Expression of normal *TFE3* was detected in five of seven cases tested. The positive tumors were from four females and one male. Both negative tumors were in males, including case ASPS-8, where cytogenetic data were available and were consistent with this finding (see above). Normal *ASPL* transcripts were detected in five of five cases tested. Although these tumor samples typically contain >90% tumor cells, the possibility that the normal transcripts detected in this PCR-based analysis were derived from non-neoplastic tumor stromal cells cannot be completely excluded.

Isolation of the full length *ASPL* cDNA

Separate 5' and 3' portions of the partial *ASPL* sequence obtained from the *TFE3* 5' RACE (see above) matched two non-overlapping human expressed sequence tags (ESTs) (GenBank #AA489553 and #AA056378), respectively from IMAGE clones 843309 (HeLa) and 509449 (normal colon). These two IMAGE clones belong to the anonymous Unigene cluster Hs.84128, composed of ESTs from a wide variety of source tissues. We first extended the *ASPL* sequence by a 5' and 3' walk in the ESTs of this human Unigene cluster. To complete the 5' sequence, we

performed 5' RACE using nested reverse primers, on double-strand adaptor-ligated HeLa cDNA (Marathon cDNA, Clontech). The sequence of the 3' end, available in multiple ESTs, was also confirmed by 3' RACE on the tailed HeLa cDNA. The assembled complete sequence of the normal *ASPL* cDNA was confirmed by resequencing.

The 1872 nucleotide (nt) *ASPL* cDNA (Figure 8) lacks significant homology to previously characterized human genes. A search of the *ASPL* cDNA sequence against the draft human genome sequence showed a near perfect head-to-head match of its 5' end with the 5' end of BAC clone RP11-489C9 recently assigned to 17q25, approximately 1.7 Mb from the telomere (GenBank #AC069004). The only sequence discrepancy was a silent difference, CCA vs CCG, at codon 45, presumably representing a polymorphism. The match with the portion of the *ASPL* genomic sequence allowed us to define the first four exon boundaries within the *ASPL* cDNA (Figure 8).

We also identified major and minor splice forms of the *ASPL* transcript which differ respectively by the absence or presence of a 47 nt segment from exon 2, which encodes part of the 5' untranslated portion of *ASPL* (Figures 6 and 8). This 47 nt segment seem to represent a 'cryptic intron' (Bruce *et al.*, 1999; Haseloff *et al.*, 1997), yet it lacks canonical exon splice junctions. That both splice forms are present in cDNA rather than deriving from contaminating genomic DNA was confirmed by RT-PCR assays spanning exon boundaries. Specifically, the expression of the minor splice form containing the additional 47 bp portion of exon 2 was detected by PCR using a reverse primer from this region (GCTTTTCCAAGTGA-GAGCCAG) with an exon 1 forward primer (GGGTCACGTGAGCGGAAAATG) in cDNA from HeLa cells and two ASPS cases (ASPS-7 and ASPS-8) (results not shown). However, PCR primers bracketing this alternatively spliced 47 bp fragment, GCTTCAGGTTCTGGAGGACAC (straddling the boundary of exons 1 and 2) and ATCCCTCGTGTA-CACGCAGACT (from exon 5), produced only the 360 bp fragment expected from transcripts lacking the 47 bp portion, suggesting that this is the major splice form of *ASPL*, at least in the cDNAs tested (HeLa, K562, ASPS-2, ASPS-7).

Sequence analysis of *ASPL* predicted protein

The *ASPL* cDNA open reading frame encodes a predicted protein of 476 amino acids that shows significant identity (29–33%) and similarity (46–51%) within its carboxy-terminal portion (approximate residues 263–385) to predicted proteins of unknown function in *Saccharomyces cerevisiae* (YMV7 product; SwissProt #P54730), *Arabidopsis thaliana* (GenBank #AAD12706), *Drosophila melanogaster* (CG5469 product; GenBank #AAF47656), and *Caenorhabditis elegans* (B0024.10 and H06H21.6 products; GenBank #T18645 and #AAC69496) (Figure 9). By ProfileScan analysis, this carboxy-terminal portion of *ASPL*

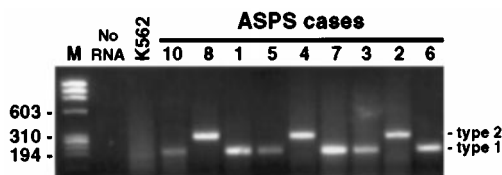


Figure 7 Detection of *ASPL-TFE3* in all ASPS cases by RT-PCR. Agarose gel showing specific *ASPL-TFE3* products in 9/9 ASPS cases, but not in negative controls (K562 and reaction lacking RNA). Type 2 products differ by the inclusion of *TFE3* exon 3 in the chimeric transcript (see Figure 6 and text)

```

GGCGGGTCACGTGAGCGGAAAATGGCGGCCCGCCAGCGGGCGGAGGCTCCGCGGTGTCG 60
GTGCTGGCCCCGAACCGCGCGGCCACACGGTGAAGGTGACGCCGAGCACCGTGTGCTT 120
▼
CAGGTTCTGGAGGACACGTCGCCGGCGGCGAGACTTCAACCCCTGTGAatgatgatgaag 180
tgagtttgcctccagctcagcagcaggggtctgagtATATCTGTGCCCTGCCCTGAACA 240
TACTCGCTCTCACTTGGAAAAGCCCCAGTTTCAGAGGAGCGTGCCTCGACCTTCTCTC 300
CAGTGGAGATTGCCAACCTGCCAACAAATGCCAAGCTGGAGTGGTGCCCGCTCCCGG 360
M V P A S R 6
AGCCGTGAGGGGCTGAGAACATG▼TTCGCATCGCTTTGCAGCTGGACGATGGCTCGAGG 420
S R E G P E N M V R I A L Q L D D G S R 26
TTGCGAGACTCTTCTGTTCAGGCGAGCCCTCTGGGACTCTCTCAGCCATTTTCCACAG 480
L Q D S F C S G Q T L W E L L S H F P Q 46
▼
ATCAGGGAGTGCCTGCBCACCCCGGGGCCACCCAGCTCTGCGTGTACACGAGGGAT 540
I R E C L Q H P G G A T P V C V Y T R D 66
GAGTGCAGGGTGAAGCTGCCCTGCGGGGACGACGCTGCAGTCTGCGTGGCCCTGACCGG 600
E V T G E A A L R G T T L Q S L G L T G 86
GGCAGCGCCACCATCAGGTTTTCATGAAGTGTACGACCCCGTGGGCAAGACCCAGGA 660
G S A T I R F V M K C Y D P V G K T P G 106
AGCCTGGGCTCGTACGCTCGGCTGGCCAGGACCGCCAGCGCTCCACTTCCCTGGAA 720
S L G S S A S A G Q A A A S A P L P L E 126
TCTGGGAGCTCAGCGCGCGGACTGTAGCCGTCGGGAGGACCGGACACCTCAGGGCCC 780
S G E L S R G D L S R P E D A D T S G P 146
TGCCTGCAGCAGACTCAGGAGAGCAGAGACAAGGGCACCCGAGCTGCCCOCTTTGTT 840
C C E H T Q E K Q S T R A P A A A P F V 166
CCTTCTCGGGTGGGGACAGAGACTGGGGGCCCTCTGGGCCACGAGGCTCTGACA 900
P F S G G G Q R L G G P P G P T R P L T 186
TCATCTCAGCTAAGTTGCCAAGTCCCTCTCCAGCCCTGGAGGCCCTCCAGCCAAAG 960
S S S A K L P K S L S S P G G P S K P K 206
AAGTCCAGTCCGGCCAGGATCCCGCAGCAGGAGCAGGAGCAGGAGCGGGAGCGGGATCC 1020
K S K S G Q D P Q Q E Q E Q E R F R D P 226
▼
CAGCAGGAGCAGGAGCGGGAGCGCCGCTGGACCGGACCCCTGGACCGGGAGCGGGTG 1080
Q Q E E R E R P V D R E P V 246
CTGTGCCACCCGACCTGGAGGAGCGGCTGCAGCGCTGGCCAGCGGAGTGCCTGATGAG 1140
V C H P D L E E R L Q A W P A E L P D E 266
TTCTTTAGCTGACGCTGGACGCTGAGAAAGCAGCTTGGCCAGCTCAAGAGTGAAGCG 1200
F F E L T V D D V R R R L A Q L K S E R 286
AAGCCCTGGAAGAAGCCCCTTGGTACCAAGGCCCTTACAGGAGGCGCAGATAAAGGAG 1260
K R L E E A P L V T K A F R E A Q I K E 306
AAGTGGAGCGCTACCCAAAGTGGCTCTGAGGGTCTGTTCGCCAGCCGCTACGTCCTA 1320
K L E R Y P K V A L R V L F P D R Y V L 326
CAGGGCTCTTCCGCCCCAGGACAGTGGGGACTTGCAGACTTCGTGAGGAGCCAC 1380
Q G F F R P S E T L R D F V R S H 346
CTGGGGAACCCGAGCTGCATTTTACCTTTCATCACCCTCCAAAAACAGTCCCTGGAC 1440
L G N P E L S F Y L F I T P P K T V L D 366
GACCACAGCAGACCCCTCTTTCAGGCAACCTCTTCCCGCCGCTCTGCTGCACTTGGGA 1500
D H T Q T L F Q A N L F P A A L V H L G 386
GCCGAGGAGCGCGGAGTGTCTACCTGGAGCCCTGGCCGCTGGAGATGCCATCCCCCA 1560
A E E P A G V L E P L E H A I S P 406
TCTCGCCGCGAGTGTGTTGGCCAGGTACATGTCCAGGGCCCGGGTCCCOCTCCCCCA 1620
S A A D V L V A R Y M S R A A G S P S P 426
TTGCCAGCCCTGACCTGCACCTAAGTCTGAGCCAGCTGCTGAGGAGGGGGCGCTGGTC 1680
L P A P D P A P K S E P A A E G A L V 446
CCCCTGAGCCCATCCAGGGAGCGCCAGCCCGTGAAGAGGAGCCTGGCAAGGTGCC 1740
P P E P I P G T A Q P V K R S L G K V P 466
AAGTGGCTGAAGCTGCCCGCCAGCAAGAGTGAAGAGTGCACGCTGAGGTGCCACTCC 1800
K W L K L P A S K R * 477
GCCACCACAGGACCCCTCTCTCCAGCAGGATAAAGACTTGTGCATCCCTCAAAAA 1860
AAAAAAAAAAAA 1872

```

Figure 8 *ASPL* cDNA sequence and predicted protein. The predicted boundaries of the first four exons (arrowheads) are based on matches to the partial *ASPL* genomic sequence contained at the 5' end of BAC clone RP11-498C9 (see text). The 47 nt alternatively spliced portion of exon 2 is shown in lowercase (see text). There are two potential in-frame initiation codons (underlined), but the first initiation codon is phylogenetically more conserved (analysis not shown). The second one presents a more favorable Kozak consensus context for

(residues 322–398) contained a significant match for a UBX domain (Pfam accession #PF00789, InterPro accession #IPR001012), a poorly characterized domain seen mainly in the carboxy-terminal half of a variety of proteins. An even stronger match was obtained with the UX domain profile (PS50033) which represents a subset of UBX domains.

No other complex motifs could be identified in the *ASPL* predicted protein. Notably, the portion of *ASPL* included in the fusion protein contains no regions reminiscent of transcriptional activation domains, with the possible exception of a partly repetitive acidic and glutamine-rich sequence, QDPQQEQEQERERDPQ-QEQERER, located immediately proximal to the *ASPL*-*TFE3* fusion point. *ASPL* shows no sequence homology to other *TFE3* fusion partners, namely *PRCC*, *PSR*, or *NonO* (analysis not shown).

Chromosomal assignment and expression analysis of *ASPL*

We identified a reverse primer from *ASPL* (CTCCTGCTCCTGCTGGGGATC) which appeared to be located on the same exon as the forward primer AAAGAAGTCCAAGTCCGGGCCA used for the *ASPL*-*TFE3* RT-PCR. This was evidenced by the amplification of a product of the expected size (80 bp) from normal genomic DNA. Using this primer pair, PCR analysis of monochromosomal somatic cell hybrid DNAs (Coriell Institute, Camden, NJ, USA) confirmed that *ASPL* maps to chromosome 17 (results not shown).

To assess the normal expression pattern of native *ASPL*, we hybridized Northern blots of various adult and embryonal human tissues (Clontech, Palo Alto, CA, USA) with an *ASPL* partial cDNA probe. The predominant normal *ASPL* transcript is about 1.9 kb in size and is detected in all tissues tested (Figure 10). When normalized to β -actin transcript level (not shown), the highest *ASPL* expression was seen in heart, skeletal muscle, pancreas (Figure 10), and testis (not shown). Slightly lower expression was seen in the remaining tissues examined [kidney, liver, lung, placenta, brain (Figure 10), leukocytes, colon, small bowel, ovary, prostate, thymus, spleen (not shown)]. Expression of *ASPL* in fetal tissues, including kidney, liver, lung, brain, and heart, appeared notably lower than in adult tissues (not shown). In a Northern blot panel of human cancer cell lines (Clontech), *ASPL* was expressed in all cell lines, with somewhat higher levels in K562, A549, G361, and SW480, and lower levels in Raji, MOLTA, HeLa, and HL60 (Figure 10). The ubiquitous expression of *ASPL* by Northern blot analysis is consistent with the wide variety of source

translation initiation (Kozak, 1991). A silent mismatch with the genomic sequence from BAC RP11-498C9 at codon 45 is shown in italic. The fusion point with *TFE3* is shown by the arrow. The portion of the *ASPL* putative protein showing a significant match to the UBX domain profile is in boldface. The polyadenylation site is underlined. GenBank accession numbers for *ASPL* are AF324219 and AF324220

```

At2g01650 PKKIDRENGCFLVQIRVFFSVSENVASRIEVPDSFYSLSADEIKREADLRKKAIAESQLLIPRSYKEKQAKAARKRYKRSMIRVQFPDGVVLQGVFAPWEPTFA-----LYE
CG5469    PLELDR-N--IKVLL-----PSQACRVALPDEFYRLSPREEIKKEQQLRSEATAQSQMLRTKAMREREQRNLRMYRYALVRVKFPNGLFIQGTENVYKISD-----VFE
YMW7     ANNLPKKN--KAISE-----CLRVRKFPDRSHIQIA--FKPNEDMRTVYNVVSQFLIDENMP-----FTLNQSHPFKPLAKDDKLLDDLEFGSKTML-----LFE
ASPL     KEKLERYP-----K-----VALRVLFPPDR--YVLQG--FFRPSETVGDRLDFVRSHLGNPELSFYLFITPPKTVLDDHTQTFLQANLFPALVHLGAEFPAGVYLEPGLLE
          :: :                               * : . * :           : : . :           *           :           : *
UBX     |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
UX     |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
    
```

Figure 9 CLUSTER W (1.8) multiple sequence alignment of ASPL UBX domain with related proteins in model organisms. Predicted proteins from *Arabidopsis thaliana* (At2g01650 product; GenBank #AAD12706), *Drosophila melanogaster* (CG5469 product; GenBank #AAF57656), and *Saccharomyces cerevisiae* (YMW7 product; SwissProt #P54730), are aligned with a portion of ASPL. This portion of ASPL shows significant matches to the profiles of the UBX domain (PF00789) and the UX domain (PS50033) using ProfileScan (Nscores of 9.989 and 12.443, respectively), in the portions indicated for each. Key: * = single, fully conserved residue; := conservation of strong groups, . = conservation of weak groups

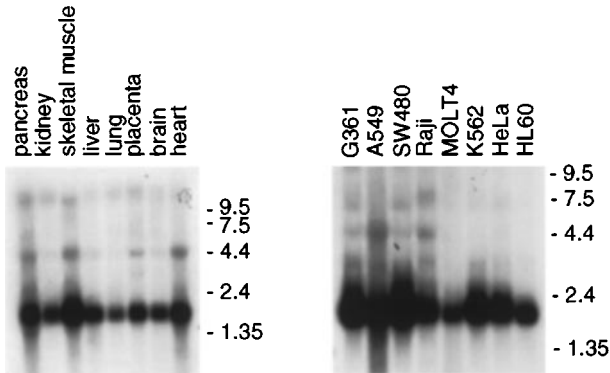


Figure 10 Northern blot analysis of expression of *ASPL* in normal adult human tissues and human cancer cell lines. Even loading of lanes was confirmed by rehybridization with a β -actin probe (not shown). HL60: promyelocytic leukemia, HeLa: cervical carcinoma, K562: chronic myelogenous leukemia, MOLT4: lymphoblastic leukemia, Raji: Burkitt's lymphoma, SW480: colorectal adenocarcinoma, A549: lung carcinoma, G361: melanoma

tissues of *ASPL* ETSs within Unigene cluster Hs.84128.

Discussion

Since the recognition of ASPS as a distinct clinicopathologic entity in 1952 (Christopherson *et al.*, 1952), its cell of origin and biology have remained unclear. Its typical presentation is as an intra- or perimuscular thigh mass in a young woman, an observation confirmed in the present group of patients (Table 1), but it has also been reported in tissues where skeletal muscle is absent, such as stomach, retroperitoneum, breast, and female genital tract, among others. Its proposed origin from skeletal muscle precursor cells has been the subject of several studies and considerable controversy (Gomez *et al.*, 1999; Rosai *et al.*, 1991; Tallini *et al.*, 1994; Wang *et al.*, 1996). In some sarcomas (e.g. alveolar rhabdomyosarcoma and myxoid liposarcoma), specific translocations subvert transcription factors involved the normal differentiation of their respective cell types (e.g. PAX3 and CHOP). In contrast, in ASPS neither translocation partner shows a tissue-specific expression pattern, and therefore the *ASPL-TFE3* fusion so far does not provide any additional clues to the cell of origin or differentiation program of this enigmatic tumor. The consistent

detection of the *ASPL-TFE3* fusion in ASPS identifies it as a marker of potential clinical utility to add to the growing list of translocation-based molecular diagnostic markers in sarcomas (Ladanyi and Bridge, 2000).

TFE3, along with *TFEB*, *TFEC*, and *MITF*, forms the microphthalmia-TFE (MiT) subfamily of basic helix-loop-helix leucine zipper (bHLH-LZ) transcription factors (Hemesath *et al.*, 1994; Rehli *et al.*, 1999). *TFE3* was initially identified through its binding to the immunoglobulin heavy chain gene enhancer μ E3 motif (Beckmann *et al.*, 1990), where it forms part of a complex that includes tissue-restricted ETS domain transcription factors (Tian *et al.*, 1999). The ubiquitous expression of *TFE3* suggests a wider physiologic role (Beckmann *et al.*, 1990). Indeed, *TFE3* has also recently been implicated in cell type-specific expression of target genes in the melanocytic and peripheral neuroendocrine lineages (Verastegui *et al.*, 2000; Wong *et al.*, 1995). *TFE3* also appears to cooperate with Smad proteins in at least one TGF-beta-activated signal transduction pathway, the one leading to activation of the plasminogen activator inhibitor-1 gene (Hua *et al.*, 1998, 1999). Thus, *TFE3*, although broadly expressed, may achieve functional specificity through interactions with more tightly regulated proteins.

As stated above, *TFE3* is rearranged within its first or third intron by specific reciprocal translocations involving Xp11.2 in a subset of papillary renal adenocarcinomas (Sidhar *et al.*, 1996; Weterman *et al.*, 1996). In most cases, the translocation partner is a novel gene designated *PRCC* at 1q21.2 and the *PRCC-TFE3* fusion transcript includes all but exon 1 of *TFE3*, thereby containing the entire coding region of *TFE3* (Sidhar *et al.*, 1996). In some papillary renal adenocarcinomas, *TFE3* fuses instead with the splicing factor genes, *PSF* at 1p34 or *NonO* at Xq12 (Clark *et al.*, 1997). *ASPL* shows no sequence similarities to *PRCC*, *PSF*, or *NonO*. The *PSF-TFE3* and *NonO-TFE3* fusions exclude the first three exons of *TFE3*, thereby replacing its activation domain by the fused sequences, while retaining the *TFE3* bHLH-LZ DNA-binding domain, encoded by exons 5 to 7. Thus, they are structurally similar to the *ASPL-TFE3* type 1 fusion, although *ASPL* shows no sequence homology to *PRCC*, *PSF*, or *NonO*. In the predicted *ASPL-TFE3* fusion protein, the amino terminal 234 amino acids are fused with the carboxy terminal 280 or 315 amino acids of *TFE3* in the type 1 and type 2 fusions,

respectively. We mapped the *TFE3* genomic break in one case with a type 1 fusion and it was localized to intron 3, as would be expected. We did not map the genomic break in any case with the type 2 fusion. These could affect intron 2 which is very small (81 bp) or may occur further upstream and produce a fusion transcript involving *TFE3* exon 3 by splicing out of more upstream exons.

TFE3 contains a nuclear localization signal located between the basic domain and the first helix (Weterman *et al.*, 2000), a region retained within the predicted *ASPL-TFE3* fusion protein. In transactivation assays using a *TFE3*-specific reporter system, PRCC-*TFE3* is a stronger activator than *TFE3* (Weterman *et al.*, 2000). These considerations, along with the apparently constitutive activity of the *ASPL* promoter, suggest that in ASPS, *ASPL-TFE3* may similarly exert its pathogenic effect through transcriptional deregulation, consistent with the biology of several other translocation-associated sarcomas. Whether this deregulation takes the form of a gain-of-function or a dominant negative remains to be determined. By sequence analysis, *ASPL* lacks any obvious activation domain and *ASPL-TFE3* type 1 fusions also lack the *TFE3* activation domain, which has been mapped to exon 3. Thus, the *ASPL-TFE3* type 1 fusion product could function as a dominant negative, akin to *TFE3* splice forms lacking exon 3 (Roman *et al.*, 1991), through homo-di- or tetra-merization mediated by its bHLH-LZ domain (Artandi *et al.*, 1994). MiT family members are known to heterodimerize (Hemesath *et al.*, 1994), and this could extend the transcriptional deregulation effected by *ASPL-TFE3* beyond the targets of normal *TFE3*.

ASPL, along with similar predicted proteins in several model organisms, may define a new UBX domain subfamily related to the yeast YMV7 product (Figure 9). The UBX domain has not yet been defined in terms of function or tertiary structure. Its identification in some proteins [human predicted proteins Y33K (Swiss Prot #Q04323) and HUMORF (GenBank #M68864)] also containing the ubiquitin-regulatory UBA domain (Hofmann and Bucher, 1996) suggests that it may be somehow related to the ubiquitylation pathway, a role that remains to be proven experimentally. Its co-occurrence with the UBA domain in selected proteins led to its designation as a UBA-associated domain, but most proteins containing UBX domains actually lack a UBA domain. Named proteins containing UBX domains include human FAF1 and REP8, and yeast shp1.

Most translocations producing fusion genes commonly appear cytogenetically reciprocal, with no net gain or loss of genetic material, even though only one of two resulting chimeric genes is pathogenetically significant. Indeed, few if any such translocations are consistently unbalanced. An unusual feature of the t(X;17) in ASPS is its consistently non-reciprocal nature, observed cytogenetically in all cases to date (Heimann *et al.*, 1998; Joyama *et al.*, 1999), and supported by the absence of a reciprocal *TFE3-ASPL*

fusion transcript in 11/12 cases tested in the present series. Incidentally, the presence of a normal X in some males with the t(X;17) (e.g. case ASPS-9 described above and in Joyama *et al.* (1999)) suggests that the translocation occurs in the G2 phase of the cell cycle, at least in these cases. Interestingly, in contrast to the unbalanced t(X;17)(p11.2;q11) of soft tissue ASPS, a balanced t(X;17)(p11.2;q11) has been reported in rare tumors described as pediatric renal adenocarcinomas (Heimann *et al.*, 1998; Hernandez-Marti *et al.*, 1995; Tomlinson *et al.*, 1991). We now find that all such cases tested display an *ASPL-TFE3* fusion transcript and may be morphologically and biologically related to ASPS (P Argani, M Ladanyi, unpublished observations).

Because of the unbalanced nature of the t(X;17), almost all ASPS harbor allelic loss of 17q25 sequences telomeric to *ASPL* and most show a gain of Xp sequences telomeric to *TFE3* (representing most of the short arm). Indeed, gain of Xp11-pter has been detected by comparative genomic hybridization in a case of ASPS (Kiuru-Kehlefeldt *et al.*, 1998). The non-reciprocal structure of the *ASPL-TFE3* rearrangement suggests a special biology in ASPS combining fusion protein formation with pathogenetic losses and/or gains of other as yet uncharacterized genes, akin to *ETV6-AML1* formation and *ETV6* loss in pediatric acute lymphoblastic leukemias (Golub *et al.*, 1997). The possible additional role of these recurrent translocation-associated genetic imbalances in the biology of this sarcoma remains to be elucidated.

Materials and methods

Tumor samples

Frozen tissue, available in 12 cases histologically diagnosed as ASPS (Table 1), was used for nucleic acid extraction in all cases and touch preparations (for FISH) in some cases. The basic patient data in the 12 cases which had sufficient material for studies of the *ASPL-TFE3* fusion are provided in Table 1. Positivity for *ASPL-TFE3* was not a criterion for inclusion in the study group.

Cytogenetic analysis

Cytogenetic analysis was performed on sterile representative samples of cases ASPS-6 and ASPS-7 using standard culture and harvesting procedures. Briefly, the tissue was disaggregated mechanically and enzymatically, then cultured in RPMI 1640 media supplemented with 20% fetal bovine serum and 1% penicillin/streptomycin-L-glutamine (Irvine Scientific, Santa Ana, CA, USA) for 3–5 days. Two to four hours prior to harvest, cells were exposed to colcemid (0.02 µg/ml). Following hypotonic treatment (0.074 M KCl for 30 min for flasks and 0.8% Na citrate for 25 min for coverslips), the preparations were fixed three times with methanol:glacial acetic acid (3:1). Metaphase cells were banded with Giemsa trypsin. The karyotypes were expressed according to the International System for Human Cytogenetic Nomenclature 1995.

Fluorescence in situ hybridization

Six probes localized to subregions of Xp11.2 (*OATL1* YAC2, cosmid A02167, cosmid U164B6, cosmid I1135, cosmid B1125, and *OATL2*) were used. The *OATL1* and *OATL2* probes were kindly provided by A Geurts van Kessel. The relative locations of these probes are illustrated in Figure 3. One μg of each DNA probe was directly labeled with SpectrumOrange or SpectrumGreen using a nick translation kit (Vysis, Downer's Grove, IL, USA) according to the manufacturer's instructions. The labeled DNA was purified through a Sephadex G50 column and precipitated in the presence of sonicated herring sperm DNA (15 μg). The labeled DNA was then co-precipitated for annealing purposes with 2 μg Cot-1 DNA (sonicated total human DNA).

Bicolor FISH studies were performed with probe mixtures of the directly labeled YAC or cosmid clone probes listed in Figure 3 and directly or indirectly labeled centromere X-specific alphoid sequence (CEP X (Vysis) or DXZ1 (Ventana Medical Systems, Tucson, AZ, USA)) on cytologic touch preparations of two of the ASPS specimens (ASPS-1 and ASPS-6). Tricolor FISH studies utilizing a SpectrumOrange labeled CEP X probe, a SpectrumAqua labeled CEP 17 probe (CEP X and CEP 17 probes, Vysis Inc., Downer's Grove, IL, USA) and a SpectrumGreen labeled *OATL1* YAC2 probe. The metaphase or interphase slides to be hybridized were pretreated with RNase A (100 $\mu\text{g ml}^{-1}$ in $2\times\text{SSC}$ at 37°C for 1 h) and eventually, porcine pepsin (0.01% (w/v) in 0.01 M hydrochloric acid, 5 min at 37°C), followed by formaldehyde post-fixation (1% (v/v) in MgCl_2 phosphate-buffered saline (PBS) for 10 min at room temperature). Subsequently, the slides were denatured in 70% formamide, $2\times\text{SSC}$, pH 7.0, at 73°C for 2–3 min and placed immediately into cold 70% EtOH, followed by a cold ethanol series of 80, 90 and 100%. The slides were then placed on a slidewarmer for 10 min at 45°C prior to hybridization. The probes tested were denatured at 75°C for 5 min, chilled on ice and incubated at 37°C for 10 min. The cells and probes were sealed under an 18×18 mm coverslip and incubated for 16–18 h at 37°C in a humidity chamber.

Images were prepared utilizing the Applied Image Analysis System (Applied Imaging, Pittsburgh, PA, USA). The number of hybridization signals for each probe was assessed in a minimum of 200 interphase nuclei with strong and well-delineated signals by two different individuals. As negative controls, normal peripheral-blood lymphocytes were simultaneously hybridized with these probes.

Southern and Northern blot analysis

For Southern blotting, genomic DNA was isolated by standard organic methods. Aliquots (10 μg) of DNA were digested with specific restriction enzymes, separated on 0.8% agarose gels and transferred onto nylon membranes, and subjected to ultraviolet cross-linking for 5 min. Southern blots and commercially obtained Northern blots (Clontech, Palo Alto, CA, USA) were prehybridized for at least 1 h and subsequently hybridized in Hybrisol (Intergen, Purchase, NY, USA) with probes radiolabeled to a high specific activity by the random primer technique. The filters were washed at high stringency and imaged using X-ray film or a phosphorimager (BioRad, Hercules, CA, USA).

Isolation of *ASPL-TFE3* and *ASPL* cDNAs

We used the 5' rapid amplification of cDNA ends (RACE) system from GIBCO BRL (Rockville, MD, USA). Briefly, after annealing of a *TFE3*-specific reverse primer from exon 5 (CAAAAGGGCCTTTGCCTCGGTC) or exon 7 (CAGGATGGTGCCCTTGTT), the RNA was copied into cDNA with Superscript II reverse transcriptase. The RNA was then degraded with RNase and the cDNA was purified by GlassMAX Spin Cartridge (GIBCO BRL). The purified cDNA was then tailed with dCTP using TdT. The dC-tailed cDNA was then subjected to PCR amplification using the manufacturer's Abridged Anchor Primer as forward primer and a nested *TFE3*-specific reverse primer from exon 4 (CGTTTGATGTTGGGCAGCTCA). For the first PCR step, either AmpliTaq (Perkin Elmer, Norwalk, CT, USA) or HotStarTaq (Qiagen, Valencia, CA, USA) were used. Five μl of 100-fold dilution of the first PCR products was then reamplified using AmpliTaq in a second PCR with heminested primers, consisting of the manufacturer's Abridged Universal Anchor Primer, and another *TFE3* exon 4 reverse primer (GCAGGAGTTGCTGACAGTGAT). Discrete product bands observed on the agarose gel were excised and subjected to direct sequencing on an automated sequencer using the latter primer or a further nested primer from *TFE3* exon 4, (CCTTGACTACTGTACACATC).

To isolate the 5' end and confirm the 3' end of the normal *ASPL* cDNA, we used 5' and 3' tailed HeLa cDNA (Marathon cDNA, Clontech). Appropriate reverse (CAAA-GAATCAATCATGGCAG) and forward (AAAGAAGTCCAAGTCGGGCCA) primers from *ASPL* were used respectively for the 5' and 3' RACE, in combination with the supplied API anchor primer for the tailed ends.

RT-PCR for *ASPL-TFE3*

First, reverse transcription was performed for 30 min at 42°C on 1 μg of RNA using $10\times$ buffer II, 25 mM MgCl_2 , 50 mM random hexamers, 10 mM dNTP, 40 U/ μl RNase inhibitor, 200 U/ μl reverse transcriptase and DEPC-treated H_2O for a final volume of 20 μl . A negative control (10 μl of DEPC-water) was included at this stage. The reverse transcriptase was inactivated at 99°C for 5 min. Eighty μl of master mix (water, $10\times$ buffer II, MgCl_2 for a final concentration of 1.5 mM, 15 pMol of each primer (see Results for primer sequences), 5 U/ μl of AmpliTaq DNA Polymerase (Perkin Elmer) were added to the tube). PCR consisted of 35 cycles of 95°C for 1 min, 60°C for 1 min, 72°C for 1 min, final extension of 72°C for 10 min. The PCR products were electrophoresed in 2% NuSieve agarose gel (FMC Bioproducts, Rockland, ME, USA) and visualized by ethidium bromide staining.

To detect the presence of an *ASPL-TFE3* fusion transcript, we performed RT-PCR using a forward primer from *ASPL* (AAAGAAGTCCAAGTCGGGCCA) and a *TFE3* exon 4 reverse primer (CGTTTGATGTTGGGCAGCTCA), as previously described. RT-PCR for the reciprocal fusion product, *TFE3-ASPL*, was performed using an *ASPL* reverse primer (CACCGTCAGCTCAAAGAACTC) and a *TFE3* forward primer appropriate to the type of *ASPL-TFE3* rearrangement in each case: for cases with a type 1 *ASPL-TFE3* fusion, a *TFE3* exon 3 forward primer (TTGATGATGTCATTGATGAGATC), and in cases with a type 2 *ASPL-TFE3* fusion, a more upstream *TFE3* primer (GCTCAAAAGCCAACCCTTAC).

Acknowledgments

We would like to thank William Gerald, Murray Brennan, Cheryl Coffin, James Neff, Richard Terek, and Henrik Bauer for specimen procurement; Aimée Hamelin, Gisele Colleoni, Marilu Nelson, Patty Cattano, Rebecca Shephard, Jian Liu, and Zhong Chen for technical assistance; Heide Hellebrand for assistance with Xp11 sequence analysis; Debbie MacDougall and Kimberly Christian for assistance with artwork; Kin Kong and Allyne Manzo for photographic work, and Philip H Lieberman for longstanding support. We thank Ad Geurts van Kessel for kindly providing the *OATL1* and *OATL2* YAC probes. We would also like to thank the following individuals for providing materials (Xp11

probes or ASPS tumor samples) used in earlier stages of this work which did not contribute to the present results: Michael Geraghty, Juan Rosai, and Carol Kruse. The genomic sequence analysis of this region of Xp11.2 deposited in GenBank #AF207550 was performed at the Institute of Molecular Biotechnology, Jena, Germany. ASPS tumor procurement at MSKCC was performed under IRB protocol #90-024, supported by NIH PO1 CA47179. JA Bridge is supported in part by the John A Wiebe Jr Children's Health Care Fund and the Nebraska State Department of Health, LB595. A Meindl is supported by the German Ministry of Research and Education. M Ladanyi is supported in part by NIH PO1 CA47179.

References

- Artandi SE, Cooper C, Shrivastava A and Calame K. (1994). *Mol. Cell Biol.*, **14**, 7704–7716.
- Beckmann H, Su LK and Kadesch T. (1990). *Genes Dev.*, **4**, 167–179.
- Bruce SR, Kaetzel CS and Peterson ML. (1999). *Nucleic Acids Res.*, **27**, 3446–3454.
- Christopherson WM, Foote FW and Stewart FW. (1952). *Cancer*, **5**, 100–111.
- Clark J, Lu YJ, Sidhar SK, Parker C, Gill S, Smedley D, Hamoudi R, Linehan WM, Shipley J and Cooper CS. (1997). *Oncogene*, **15**, 2233–2239.
- Golub TR, Barker GF, Stegmaier K and Gilliland DG. (1997). *Curr. Top. Microbiol. Immunol.*, **220**, 67–79.
- Gomez JA, Amin MB, Ro JY, Linden MD, Lee MW and Zarob RJ. (1999). *Arch. Pathol. Lab. Med.*, **123**, 503–507.
- Haseloff J, Siemering KR, Prasher DC and Hodge S. (1997). *Proc. Natl. Acad. Sci. USA*, **94**, 2122–2127.
- Heimann P, Devalck C, Dubusscher C, Sariban E and Vamos E. (1998). *Genes Chromosom. Cancer*, **23**, 194–197.
- Hemesath TJ, Steingrimsson E, McGill G, Hansen MJ, Vaught J, Hodgkinson CA, Arnheiter H, Copeland NG, Jenkins NA and Fisher DE. (1994). *Genes Dev.*, **8**, 2770–2780.
- Hernandez-Martí MJ, Orellana-Alonso C, Badia-Garrabou L, Verdeguer MA and Paradis-Alos A. (1995). *Cancer Genet. Cytogenet.*, **83**, 82–83.
- Hofmann K and Bucher P. (1996). *Trends Biochem. Sci.*, **21**, 172–173.
- Hua X, Liu X, Ansari DO and Lodish HF. (1998). *Genes Dev.*, **12**, 3084–3095.
- Hua X, Miller ZA, Wu G, Shi Y and Lodish HF. (1999). *Proc. Natl. Acad. Sci. USA*, **96**, 13130–13135.
- Joyama S, Ueda T, Shimizu K, Kudawara I, Mano M, Funai H, Takemura K and Yoshikawa H. (1999). *Cancer*, **86**, 1246–1250.
- Kiuru-Kuhlefelt S, el-Rifai W, Sarlomo-Rikala M, Knuutila S and Miettinen M. (1998). *Mod. Pathol.*, **11**, 227–231.
- Kozak M. (1991). *J. Biol. Chem.*, **266**, 19867–19870.
- Ladanyi M and Bridge JA. (2000). *Hum. Pathol.*, **31**, 532–538.
- Lieberman PH, Brennan MF, Kimmel M, Erlandson RA, Garin-Chesa P and Flehinger BY. (1989). *Cancer*, **63**, 1–13.
- Macchi P, Notarangelo L, Giliani S, Strina D, Repetto M, Sacco MG, Vezzoni P and Villa A. (1995). *Genomics*, **28**, 491–494.
- Ordóñez NG. (1999). *Adv. Anat. Pathol.*, **6**, 125–139.
- Rehli M, Den Elzen N, Cassidy AI, Ostrowski MC and Hume DA. (1999). *Genomics*, **56**, 111–120.
- Roman C, Cohn L and Calame K. (1991). *Science*, **254**, 94–97.
- Rosai J, Dias P, Parham DM, Shapiro DN and Houghton P. (1991). *Am. J. Surg. Pathol.*, **15**, 974–981.
- Schindelbauer D, Hellebrand H, Grimm L, Bader I, Meitinger T, Wehnert M, Ross M and Meindl A. (1996). *Genome Res.*, **6**, 1056–1069.
- Sciot R, Dal Cin P, De Vos R, Van Damme B, de Wever I, Van Den Berghe H and Desmet VJ. (1993). *Histopathology*, **23**, 439–444.
- Sidhar SK, Clark J, Gill S, Hamoudi R, Crew AJ, Gwilliam R, Ross M, Linehan WM, Birdsall S, Shipley J and Cooper CS. (1996). *Hum. Mol. Genet.*, **5**, 1333–1338.
- Tallini G, Parham DM, Dias P, Cordon-Cardo C, Houghton PJ and Rosai J. (1994). *Am. J. Pathol.*, **144**, 693–701.
- Tian G, Erman B, Ishii H, Gangopadhyay SS and Sen R. (1999). *Mol. Cell Biol.*, **19**, 2946–2957.
- Tomlinson GE, Nisen PD, Timmons CF and Schneider NR. (1991). *Cancer Genet. Cytogenet.*, **57**, 11–17.
- Verastegui C, Bertolotto C, Bille K, Abbe P, Ortonne JP and Ballotti R. (2000). *Mol. Endocrinol.*, **14**, 449–456.
- Wang NP, Bacchi CE, Jiang JJ, McNutt MA and Gown AM. (1996). *Mod. Pathol.*, **9**, 496–506.
- Weternan MA, Wilbrink M and Geurts van Kessel A. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 15294–15298.
- Weternan MAJ, van Groningen JJM, Jansen A and Geurts van Kessel A. (2000). *Oncogene*, **19**, 69–74.
- Wong SC, Moffat MA, Coker GT, Merlie JP and O'Malley KL. (1995). *J. Neurochem.*, **65**, 23–31.