Recurrent \textit{PRDM10} Gene Fusions in Undifferentiated Pleomorphic Sarcoma

Jakob Hofvander$^1$, Johnbosco Tayebwa$^1$, Jenny Nilsson$^1$, Linda Magnusson$^1$, Otte Brosjö$^2$, Olle Larsson$^3$, Fredrik Vult von Steyern$^4$, Nils Mandahl$^1$, Christopher D. M. Fletcher$^5$, Fredrik Mertens$^1$

1 Department of Clinical Genetics, University and Regional Laboratories, Lund University, SE-221 85 Lund, Sweden;

2 Department of Orthopedics, Karolinska University Hospital, SE-171 76 Solna, Sweden;

3 Department of Pathology, Karolinska University Hospital, SE-171 76 Solna, Sweden;

4 Department of Orthopedics, Skåne University Hospital, Lund University, SE-221 85 Lund, Sweden;

5 Department of Pathology, Brigham and Women’s Hospital, Boston, MA, USA.

\textbf{Keywords:} Sarcoma, Undifferentiated, Fusion, \textit{PRDM10}, \textit{MED12}, \textit{CITED2}

\textbf{Running head:} \textit{PRDM10} fusions in UPS

\textbf{Financial support:} The study was supported by grants to F. Mertens from the Swedish Cancer Society, the National Research Council of Sweden, the Gunnar Nilsson Cancer Foundation, the IngaBritt and Arne Lundberg Foundation, and the Medical Faculty of Lund University.

\textbf{Disclosure/conflict of interest:} The authors declare no conflict of interest.

*Corresponding author: Jakob Hofvander, Department of Clinical Genetics, Lund University Hospital, SE-221 85 Lund, Sweden. E-mail: jakob.hofvander@med.lu.se
Translational relevance

Undifferentiated pleomorphic sarcoma (UPS) is one of the most common subtypes of soft tissue sarcomas. The clinical behaviour is unpredictable, and metastases occur in about 1/3 of the patients. Treatment is based on surgery with wide margins. Depending on surgical margins, location and tumor-associated risk factors, adjuvant treatment including radiotherapy and chemotherapy is considered. Biomarkers that could distinguish UPS from other types of sarcoma, as well as improve treatment stratification are needed. Previous genetic analyses have failed to reveal any consistent or tumor-specific aberrations. We here describe the finding of novel, and so far tumor-specific, gene fusions – MED12/PRDM10 and CITED2/PRDM10 – in a subset of UPS. None of the patients with these gene fusions has developed any metastases and all tumors were diagnosed as low-grade malignant at morphological re-review, suggesting that fusion-positive tumors may represent a less aggressive subset of UPS.

(141 words)
Abstract

Purpose: Undifferentiated pleomorphic sarcoma (UPS) is defined as a sarcoma with cellular pleomorphism and no identifiable line of differentiation. It is typically a high-grade lesion with a metastatic rate of about 1/3. No tumor-specific rearrangement has been identified and genetic markers that could be used for treatment stratification are lacking. We performed transcriptome sequencing (RNA-Seq) to search for novel gene fusions.

Experimental design: RNA-Seq, fluorescence in situ hybridization, and/or various PCR methodologies were used to search for gene fusions and rearrangements of the PRDM10 gene in 84 soft tissue sarcomas.

Results: Using RNA-Seq, two cases of UPS were found to display novel gene fusions, both involving the transcription factor PRDM10 as the 3’-partner and either MED12 or CITED2 as the 5’ partner gene. Further screening of 82 soft tissue sarcomas for rearrangements of the PRDM10 locus, revealed one more UPS with a MED12/PRDM10 fusion. None of these genes has been implicated in neoplasia-associated gene fusions before.

Conclusions: Our results suggest that PRDM10-fusions are present in around 5% of UPS. Although the fusion-positive cases in our series showed the same nuclear pleomorphism and lack of differentiation as other UPS, it is noteworthy that all three were morphologically low-grade and that none of the patients developed metastases. Thus, PRDM10 fusion-positive sarcomas may constitute a clinically important subset of UPS.

(216 words)
Introduction

Undifferentiated sarcoma is defined as a sarcoma with no identifiable line of differentiation, excluding dedifferentiated types of specific sarcomas (1). Undifferentiated sarcomas, accounting for approximately 20% of all soft tissue sarcomas, may be further subdivided according to cellular shape (round cell, spindle cell, epithelioid, or pleomorphic). The pleomorphic variant (UPS) is particularly common among adults, and most frequently arises in the lower extremities (2). It is typically a high-grade lesion with a local recurrence rate ranging between 19-31%, a metastatic rate of 31-35% and a five-year survival of 65-70% (3). UPS have a highly variable morphology, all sharing a marked pleomorphism often admixed with spindle cells and bizarre multinucleated giant cells. Treatment is based on the same strategy as for most other soft tissue sarcomas, i.e., surgery with wide margins. Depending on surgical margins, location, and tumor-associated risk factors, adjuvant treatment including radiotherapy and chemotherapy is considered.

The genetic aspects of UPS are still poorly defined, partly due to shifting diagnostic criteria; although many sarcomas now diagnosed as UPS were previously classified as malignant fibrous histiocytoma (MFH), a substantial subset of MFH tumors were shown to constitute poorly differentiated forms of other sarcomas, such as leiomyosarcoma or liposarcoma (1). The karyotypes and copy number profiles for UPS tend to be highly complex, with extensive intercellular variation, and a complete description of all chromosomal aberrations is rare (4-7). However, the level of cytogenetic complexity varies considerably, with a subset showing only a few structural and/or numerical aberrations. Still, no specific recurrent aberration has so far been identified, and there are no good genetic markers that could be used for treatment stratification.
In an attempt to identify clinically and biologically relevant subgroups of UPS, we performed transcriptome sequencing (RNA-Seq) and we here report the finding of two novel gene fusions in UPS, both involving the transcription factor *PRDM10* as the 3’-partner and either *MED12* or *CITED2* as the 5’ partner gene.

**Materials and methods**

**Patients and tumors**

RNA-Seq of two UPS (Cases 1 and 2), selected on the basis of their simple karyotypes, were found to harbour gene fusions involving the *PRDM10* gene. In order to evaluate the frequency and distribution of *PRDM10* fusions in UPS and other soft tissue sarcomas, a cohort of 82 additional soft tissue sarcomas (26 UPS, 22 myxofibrosarcomas, 10 leiomyosarcomas, 5 low-grade fibromyxoid sarcomas, 5 myofibroblastic sarcomas, 3 myxoid liposarcomas, 2 malignant peripheral nerve sheath tumors, 1 solitary fibrous tumor, 4 spindle cell sarcomas, 1 fibroblastic sarcoma, and 3 unclassifiable sarcomas) was analyzed. The tumors in this extended cohort were partly selected on the basis of their karyotypes. Thus, tumors with structural rearrangements of chromosome arms Xq, 6q, and 11q, i.e., the locations of the *MED12*, *CITED2*, and *PRDM10* genes, respectively, at G-banding analysis were retrieved from the archives of the Dept of Clinical Genetics in Lund; Xq, 6q, and/or 11q rearrangements were present in 12, 15, and 29 cases, respectively. We also specifically retrieved 16 tumors that had been diagnosed as low-grade malignant UPS, myxofibrosarcoma, or leiomyosarcoma by querying the Scandinavian Sarcoma Group registry. All tumors were diagnosed according to established criteria (1, 8). Clinical, morphological, and cytogenetic data are presented in Supplementary Table S1. All samples were obtained after written consent and all studies were approved by the institutional ethical committees.
Cytogenetic and FISH analyses

Cell culturing, harvesting and G-banding were performed as described and the karyotypes were written following the recommendations of the International System for Human Cytogenetic Nomenclature (9, 10).

FISH was performed on interphase nuclei from Cases 2, 27, 29, 52, and 75 using bacterial artificial chromosome (BAC) clones flanking the PRDM10 locus obtained from the BAC PAC resources, Oakland, USA. 5´-probes were RP11-664J16, RP11-237N19, and RP11-61J24 and 3´-probes were RP11-1104M18, RP11-121M22, and RP11-110K10. Clone preparation, hybridization, and analysis were performed as described (11). No material for FISH was available from Case 1.

RNA-Seq

RNA-Seq and bioinformatic analysis to identify candidate fusion transcripts were performed on Cases 1, 2, 35, 36, 44, and 49. mRNA libraries were prepared for sequencing using the Truseq RNA sample preparation kit v 2 (Illumina, San Diego, USA) as previously described (12). Briefly, poly-A-tailed RNA was enriched from total RNA using magnetic oligo-dT beads. RNA was fragmented to a median size of 200 nucleotides and cDNA was synthesized from these fragments using Superscript II reverse transcriptase (Invitrogen, Foster City, USA). Double stranded cDNA was produced using DNA polymerase I and RNase H. Oligonucleotide adaptors were ligated to the double stranded cDNA and the adaptor-bound fragments were enriched using a 15 cycle PCR. Paired-end 101 bp reads were generated from the mRNA libraries using the HiScanSQ System (Illumina).

To identify candidate fusion transcripts from the sequence data, analyses were performed on fastq files using Chimerascan (13) version 0.4.5, SOAPfuse (14) version 1.26
and TopHat (15) version 2.0.7. The GRCh37/hg19 build was used as the human reference genome.

Quantitative real-time PCR (qPCR)

To evaluate differences in the expression levels of the 5’- and 3’-parts of PRDM10, indicative of a chromosomal breakage within the gene, TaqMan gene expression assays were performed with: Hs00360640 (PRDM10 5’) covering exons 5-6 and Hs000999748 (PRDM10 3’) covering exons 20-21. The TBP gene was used as endogenous control. qPCR was performed according to the manufacturer’s instructions and all reactions were run in triplicate (Applied Biosystems, Foster City, USA). Calculations were done using the comparative C_{T} method (i.e. ΔΔC_{T} method; 16) using the SDS software 1.3.1 (Applied Biosystems).

RT-PCR

Total RNA was extracted from frozen tumor samples using the RNeasy Lipid tissue kit (Qiagen, Valencia, CA, USA). Reverse transcription and PCR amplifications were performed as described (11, 17). Primers specific for MED12, CITED2 and PRDM10 were designed to detect possible fusion transcripts (Supplementary Table S2). Transcripts were amplified using an initial denaturation for 2 min at 94°C, followed by 30 cycles of 30 sec at 94°C, 30 sec at 58°C, and 3 min at 72°C, and a final extension for 3 min at 72°C. Amplified fragments were purified from agarose gels and directly sequenced using the Big Dye v1.1 cycle sequencing kit (Applied Biosystems) on an ABI-3130 genetic analyzer (Applied Biosystems). The BLASTN software (http://www.ncbi.nlm.nih.gov/blast) was used for the analysis of MED12, CITED2 and PRDM10 sequence data.
Results

Genetic findings in the two index cases

RNA-Seq resulted in 13,955,975 reads in Case 1 and 12,758,033 reads in Case 2. In Case 1, Chimerascan identified a MED12/PRDM10 fusion, supported by 3 unique flanking reads, and in Case 2 SOAPfuse identified a CITED2/PRDM10 fusion supported by 2 spanning reads and 6 junction reads. In both cases, the genes implicated in the fusions map to breakpoints identified at G-banding analysis: MED12 maps to Xq13, PRDM10 to 11q24, and CITED2 to 6q24. Thus, both fusions were in agreement with the karyotypes, i.e., a t(X;1;11)(q13;p36;q23) in Case 1 and a t(6;11)(q24;q24) in Case 2 (Fig. 1, Supplementary Table S1). Additional detected potential fusion transcripts were considered read-through transcripts or other artefacts.

RT-PCR and subsequent sequencing of amplified products from Cases 1 and 2 identified in-frame MED12/PRDM10 and CITED2/PRDM10 fusions, respectively (Fig. 1). No reciprocal transcript, i.e., PRDM10/MED12 or PRDM10/CITED2, could be detected (data not shown). FISH with PRDM10-specific probes in Case 2 verified the break in PRDM10 also at the genomic level (Fig. 2).

The breakpoints in the two 5’-genes (MED12 and CITED2) were located towards the ends of their coding parts. In MED12, the breakpoint was located in the intron between exons 43 and 44. MED12 thus only loses two of its 45 exons in the fusion event. CITED2 has two exons and the breakpoint was located within exon 2, at nucleotide (nt) position 1047 (NM_006079.4) which is only 9 nt from the stop codon. The shared 3’-partner, PRDM10, has 22 exons. In Case 1, the fusion breakpoint was located between exons 12 and 13 and in Case 2 between exons 13 and 14.
qPCR showed higher expression of the 3’-part of PRDM10 in both cases. The ratios between the expression levels of the 3’- and 5’-probes were 1.82 and 4.15 in Cases 1 and 2, respectively.

**Genetic findings in an extended cohort of soft tissue sarcomas**

Because of the possibility of multiple 5’-partners to PRDM10 and the finding of differential expression of the 5’- and 3’-parts of PRDM10 in the two fusion-positive index tumors (Cases 1 and 2), 78 additional soft tissue sarcomas were analyzed by qPCR. Neither the 3’ nor the 5’ expression levels were consistently higher among fusion-positive tumors than among fusion-negative tumors (Supplementary Fig. S1). None of the tumors showed a 3’/5’ ratio above 1.2, while six had ratios below 0.7. All these six cases were analyzed by RT-PCR for MED12/PRDM10 and CITED2/PRDM10 fusion transcripts, using multiple primer pairs (Supplementary Table S2), revealing a MED12/PRDM10 fusion in one (Case 27, an UPS). Sequencing confirmed a fusion between the last nt of MED12 exon 43 and the first nt of PRDM10 exon 14 (Fig. 1). Three of the five RT-PCR negative cases could be analyzed also by interphase FISH using a break-apart probe for PRDM10; all were negative. Finally, four myxofibrosarcomas were subjected to RNA-Seq but did not display any fusion transcript involving PRDM10. Thus, only one additional PRDM10 gene fusion was detected among the 82 soft tissue sarcomas, including 26 UPS, in the extended cohort (Supplementary Table S1).

**Morphology of PRDM10-positive tumors**

All three cases showed features of an UPS with neither morphologic nor immunophenotypic evidence of any specific line of differentiation (Fig. 3). Each consisted of eosinophilic spindled, ovoid or multinucleate cells with bizarre, irregular, vesicular nuclei. Each had a variably prominent collagenous stroma containing multifocally scattered lymphocytes. In contrast to most pleomorphic sarcomas, in each case, mitoses numbered less than 1 per 10 high power fields and there was no necrosis. These unusual tumors were graded...
subjectively as low-grade based on the experience of one of the authors (CDMF). Aside from this finding, there were no features that distinguished these tumors from other UPS in general. One case each had focally myxoid matrix (Case 2), prominent pseudovascular clefts (Case 1) and numerous multinucleate giant cells (Case 27), respectively. Tumor cells expressed only CD34, which is not lineage-specific.

**Discussion**

Although UPS is one of the most common sarcoma subtypes, its genetic features remain poorly explored. Marked differences in clinical outcome that cannot be explained merely by differences in tumor size or location, combined with a lack of targeted treatments, provide compelling arguments for more comprehensive attempts to delineate genetic subgroups of UPS. We have in this study been able to identify a small but significant subset of UPS showing gene fusions in which either \textit{MED12} or \textit{CITED2} is fused with \textit{PRDM10}. None of these gene fusions has been described in any other neoplasm, suggesting that they are specific for UPS.

The cohort studied here included a total of 84 soft tissue sarcomas, 28 of which were diagnosed as UPS. The three sarcomas that were positive for \textit{PRDM10} fusions were all diagnosed as UPS, but we find it unlikely that \textit{PRDM10} fusions are present in as much as 10\% (3/28) of UPS. Only one of the 26 UPS cases in the extended cohort was positive, and that case had been selected because it had been classified as a low-grade malignant tumor; low-grade malignant lesions constitute a minority of all UPS (1). Also a comparison between the features of the present cases and previous cytogenetic data on UPS indicate that \textit{PRDM10} fusions are rare events. In the present study, two of three fusion-positive cases had simple karyotypes with a balanced translocation, either as the sole change or together with a few
numerical aberrations; the cytogenetic analysis failed in the third case. Abnormal karyotypes have been described in 85 cases of UPS, the majority (57/85) showing highly complex karyotypes with 50-100 chromosomes and multiple structural and numerical changes (4); only 13 of the cases had a near-diploid karyotype with less than five structural rearrangements and without any sign of gene amplification (ring chromosomes or double minutes). Thus, we estimate the frequency of PRDM10 fusion-positive tumors to be around 5% of all UPS.

Even if the PRDM10 fusion-positive cases constitute a minority of all UPS, it may be clinically important to identify them. The high metastasis rate of UPS, approximately one-third, calls for aggressive treatment. Possibly, PRDM10 fusion-positive tumors have a lower propensity for metastasizing; none of our three patients has developed metastases and they were all in complete remission after 41 months to 21 years of follow-up. Furthermore, and in agreement with the favourable outcome, all three were classified as low-grade malignant tumors when re-reviewed; it should be emphasized, though, that two had initially been diagnosed as high-grade lesions. However, there were no distinct morphologic features among the PRDM10 fusion-positive cases setting them apart from other UPS. Thus, fusions involving PRDM10 could possibly function as a marker to identify a patient subset with favourable clinical outcome. Needless to say, however, the behaviour of PRDM10 fusion-positive tumors needs to be evaluated in a much larger series of cases, before it can be decided whether they should be treated in other ways than other UPS.

PRDM10 is a poorly studied member of the PRDM (PRDI-BF1 and RIZ homology domain containing) family of proteins. It lacks enzymatic activity and is believed to function as a transcriptional co-factor by recruiting histone-modifying enzymes to target promoters and is suggested to have an important role during development of the central nervous system (18). The protein is characterized by multiple zinc-finger domains and an N-terminal PR domain (19). Several other members of the PRDM family are associated with
cancer and gene fusions involving PRDM16 have been reported in cases of acute myeloid leukemia and myelodysplastic syndrome. PRDM16 can have several partner genes and all reported fusions lead to overexpression of parts of the gene, usually not containing the PR domain, or the complete gene by promoter swapping (20).

MED12 is part of a large multiprotein complex known as the mediator complex which functions as a protein bridge between transcription factors and RNA polymerase II to initiate transcription (21). This complex also affects later stages of the transcription process including elongation and termination. MED12, MED13, Cyclin C and cyclin-dependent kinase 8 together form a dissociable part of the mediator complex known as the CDK8 module (22). The CDK8 module functions as a negative regulator of transcription by competing for the same binding site as RNA polymerase II on the core mediator complex. However, there are also reports implicating CDK8 as a transcriptional activator (23). This multifunctional module plays major roles in proliferation and differentiation and participates in various molecular pathways including the p53 and Wnt/β pathways (24). MED12 regulates the kinase-activity of the Cdk8 module and mutations in MED12 are associated with several diseases, including neoplasia. Mutations, especially in exon 2, are found at high frequencies in uterine leiomyoma and fibroadenoma of the breast (25, 26), as well as in malignancies, such as colorectal cancer, leiomyosarcoma, and prostate cancer (21, 27).

CITED2 is a non-DNA binding transcriptional co-activator that affects the activity of multiple genes by recruiting CBP/p300 to chromatin via the DNA-binding transcription factor AP2. CITED2 also competitively inhibits the transcription of hypoxia-activated genes by blocking the interactions between HIF-A1 and CBP/p300 (28). It is a multifunctional protein best known for its importance during development but also in cancer. It has been reported to be overexpressed in breast cancer where it modulates the transcriptional activity of the estrogen receptor (29).
It is difficult to make predictions on the functional outcome of fusion genes without further analysis at the protein level. However, it is reasonable to assume that both MED12/PRDM10 and CITED2/PRDM10 act as driver mutations; all previously identified recurrent gene fusions occurring in sarcomas with simple karyotypes, i.e., with few or no additional aberrations other than the translocations underlying the fusions, have been shown to be strong driver mutations (30). It is also worth noting that all genes involved in the PRDM10 fusions play important roles in gene regulation. The breakpoint in PRDM10 reveals that the PR domain is lost but nine of the ten zinc-finger domains are included in the fusion. The breakpoints in MED12 and CITED2 are located close to the 3’ end of the genes which might indicate that the functions of these proteins are still intact despite the fusion events. Recruiting functional transcription regulators to a new set of target genes by fusing them to the zinc-finger domains of PRDM10 could potentially be a mode of action to promote tumor development in these cases.

Acknowledgments

We gratefully acknowledge the help from Elisabeth Johansson at the Scandinavian Sarcoma Group central registry.
References


18. Hohenauer T, Moore AW. The Prdm family: expanding roles in stem cells and
19. Fog CK, Galli GG, Lund AH. PRDM proteins: important players in differentiation and

   estrogen receptor transcriptional activity in breast cancer cells. Biochem Biophys Res

30. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions
Legends to figures

**Figure 1.** A: Illustration of the *MED12, CITED2* and *PRDM10* genes with vertical arrow heads indicating the breakpoint locations, horizontal arrows indicating the locations of PCR primers, and braces indicating the locations of probes for quantitative real-time PCR. The coding parts of the genes are indicated in dark green color. B-D: Partial chromatograms of amplified fragments corresponding to in-frame *MED12/PRDM10* and *CITED2/PRDM10* fusion transcripts.

**Figure 2.** A: Representative karyotype of Case 1 showing a translocation t(X;1;11)(q13;p36;q23) correlating with the genomic location of the *MED12* and *PRDM10* genes. B-C: Interphase FISH analysis revealing split signals with BAC probes covering 5´and 3´regions neighboring the *PRDM10* gene.

**Figure 3.** Morphology of *PRDM10* fusion-positive undifferentiated pleomorphic sarcomas. A: Case 2 showing non-distinctive pleomorphic spindle and ovoid cells. B: High power of Case 1 highlights bizarre nuclear morphology. C-E: Individual cases showing focally myxoid stroma (Case 2), pseudovascular clefts (Case 1) and prominent tumor giant cells (Case 27).
Figure 1.
Figure 2.
Figure 3.
Recurrent PRDM10 Gene Fusions in Undifferentiated Pleomorphic Sarcoma


Clin Cancer Res Published OnlineFirst December 16, 2014.