EWSR1 and FUS fusion genes in tumorigenesis

Möller, Emely

2010

EWSR1 AND *FUS* FUSION GENES IN TUMORIGENESIS

Akademisk avhandling av

EMELY MÖLLER

som med vederbörligt tillstånd av Medicinska fakulteten vid Lunds universitet för avläggande av doktorsexamen i medicinsk vetenskap kommer att offentligen försvaras i föreläsningssal F1, centralblocket, Universitetssjukhuset i Lund, torsdagen den 27 maj 2010 09.30

FAKULTETSOPPONENT

Enrique de Álava

Laboratory of Molecular Pathology, Centro de Investigación del Cáncer-IBMCC, Universidad de Salamanca-CSIC, Salamanca, Spain

EWSR1 AND *FUS* FUSION GENES IN TUMORIGENESIS

EMELY MÖLLER

DEPARTMENT OF CLINICAL GENETICS
FACULTY OF MEDICINE
LUND UNIVERSITY
SWEDEN

2010

Printed in Sweden by Media Tryck AB, Lund 2010

ISSN 1652-8220

ISBN 978-91-86443-77-1

Lund University, Faculty of Medicine Doctoral Dissertation Series 2010:62

Cover: The firefly (*Photinus pyralis*). The image is reproduced with permission from the copyright owner.

Source: http://www.flickr.com/photos/16849297@N00/181210602

CONTENTS

ORIGINAL ARTICLES

I. Möller E, Mandahl N, Iliszko M, Mertens F and Panagopoulos I (2009). Bidirectionality and transcriptional activity of the *EWSR1* promoter region. *Oncology Reports* **21:**641-648

II. Möller E, Stenman G, Mandahl N, Hamberg H, Mölne L, van den Oord JJ, Brosjö O, Mertens F and Panagopoulos I (2008). *POU5F1*, encoding a key regulator of stem cell pluripotency, is fused to *EWSR1* in hidradenoma of the skin and mucoepidermoid carcinoma of the salivary glands. *The Journal of Pathology* **215:**78-86

III. Panagopoulos I, Möller E, Dahlén A, Isaksson M, Mandahl N, Vlamis-Gardikas A and Mertens F (2007). Characterization of the native CREB3L2 transcription factor and the FUS/CREB3L2 chimera. *Genes, Chromosomes & Cancer* **46:**181-191

IV. Möller E, Hornick JL, Veerla S, Domanski HA and Mertens F. Low-grade fibromyxoid sarcoma has a specific expression profile with upregulation of *CD24* and *FOXL1* (*Manuscript*)

Articles are referred to in the text by their Roman numerals. Articles I-III are reproduced with kind permission from Spandidos Publications and Wiley InterScience.

ABBREVIATIONS

BSTS	Bone and soft tissue sarcoma
CTD	Carboxy-terminal domain
DBD	DNA-binding domain
DLR	Dual-luciferase assay
ESFT	Ewing sarcoma family tumor
FISH	Fluorescence *in situ* hybridization
FLuc	Firefly (*Photinus pyralis*) luciferase
IHC	Immunohistochemistry
LGFMS	Low-grade fibromyxoid sarcoma
MEC	Mucoepidermoid carcinoma
NTD	Amino-terminal domain
PCR	Polymerase chain reaction
RBD	RNA-binding domain
RLuc	*Renilla reniformis* luciferase
RT-PCR	Reverse transcriptase-PCR
TAD	Transcriptional activation domain
TF	Transcription factor
TFBS	Transcription factor binding site
wt	Wild-type

PREFACE

Translocations involving *EWSR1* and *FUS* typically characterize types of malignant bone and soft tissue tumors (sarcomas) that display few other aberrations. The translocation event and associated gene fusion are believed to be crucial for tumorigenesis and thus, these sarcomas provide excellent backgrounds for studying fusion-gene mediated tumorigenic mechanisms. The first sarcoma-associated fusion gene, *EWSR1-FLI1*, was identified almost 20 years ago and since then, several putative tumorigenic functions of EWSR1 and FUS chimeras have been unravelled. However, the fact that the cell-of-origin of most sarcomas is unknown poses a major problem for studying sacomagenesis. This is apparent because the tumorigenic properties of fusion proteins have been found to be dependent on specific cellular environments. Moreover, the genetic backgrounds of many sarcomas, as well as the tumorigenic roles of their associated chimeras, are still to a large extent unknown.

The general aim of this thesis has been to understand more about the role of EWSR1 and FUS chimeras in tumorigenesis. The thesis is divided in three sections; the first provides an introduction to the wild-type and tumorigenic functions of the EWSR1 and FUS proteins, the second contains a summary, including aims, materials and methods and conclusions, of the present work and the third includes the articles on which the thesis is based.

Lund, April 2010

INTRODUCTION

The acquired genetic events that underlie the tumorigenic process can be summarized to involve the activation of oncogenes or inactivation of tumor suppressor genes (Vogelstein and Kinzler, 2004). Oncogenic activation sometimes occurs through balanced chromosomal translocations of which numerous have been reported since the identification of the Philadelphia chromosome in chronic myeloid leukemia (Rowley, 1973). Many of these aberrations are of considerable diagnostic and prognostic importance, particularly for hematological malignancies and childhood sarcomas, and are believed to constitute early and important events in tumorigenesis (Mitelman *et al*, 2007).

Balanced translocations result either in the deregulation of a gene at one of the breakpoints, or the fusion of two genes, one from each breakpoint. *EWSR1* and *FUS* (alias *TLS*) were first identified as the targets of translocations t(11;22)(q24;q12) in the Ewing sarcoma family tumor (ESFT) and t(12;16)(q13;p11) in myxoid liposarcoma, respectively (Delattre *et al*, 1992, Crozat *et al*, 1993, Rabbitts *et al*, 1993). These translocations result in the formation of fusion oncogenes where the 5´-part of *EWSR1* or *FUS* is joined in-frame to the 3´-partner gene, in this case *FLI1* or *DDIT3* (alias *CHOP*) (Figure 1A). Subsequently, various analogous chimeric genes involving *EWSR1, FUS* or the related gene *TAF15* (alias *hTAF$_{II}$68* or *RBP56*), have been identified in bone and soft tissue tumors and leukemias.

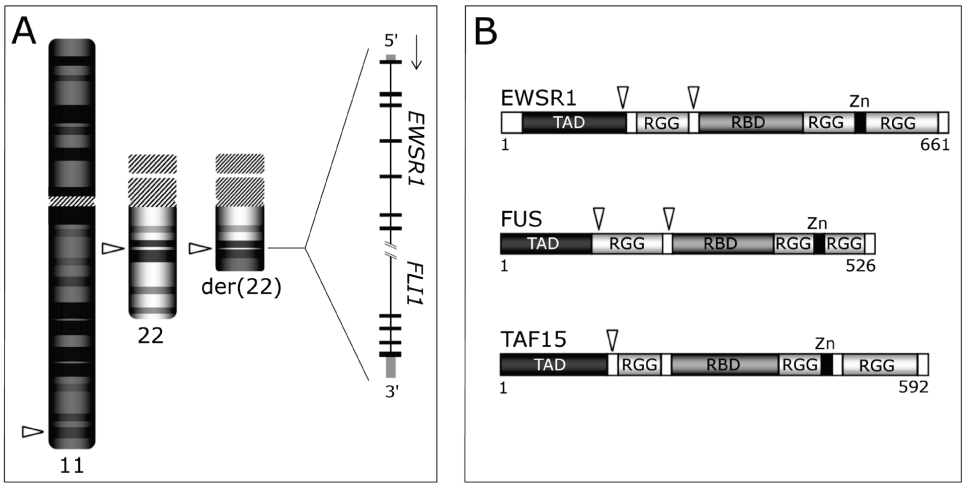

Figure 1. A) Translocation t(11;22)(q24;q12) in ESFT. The breakpoints on chromosomes 11 and 22 are marked with arrowheads. The fusion gene *EWSR1-FLI1* is expressed from the *EWSR1* promoter (arrow) on the derivative chromosome 22 (der(22)). B) Structure of the wild-type TET proteins. Arrowheads show the approximate locations of the most common breakpoints. In the chimeric proteins, the transcriptional activation domain (TAD) is retained while the RNA binding domain (RBD), zinc finger motif (Zn) and Arg-Gly-Gly (RGG) repeat rich regions are replaced with the DNA binding domains of the fusion partner. The number of amino acid residues is denoted below.

The TET proteins

The structurally and functionally related proteins EWSR1, FUS and TAF15 constitute the TET-family. Their structural features include a central RNA binding domain (RBD), Arg-Gly-Gly (RGG) repeat rich regions and a, possibly DNA-binding, Cys_2-Cys_2-type zinc finger motif in the carboxy-terminal domain (CTD). EWSR1, FUS and TAF15 also have in common a Ser-Tyr-Gly-Gln (SYGQ) rich amino-terminal domain (NTD) which functions as a transcriptional activation domain (TAD) (Delattre *et al*, 1992, Crozat *et al*, 1993, May *et al*, 1993, Prasad *et al*, 1994, Sánchez-García and Rabbitts, 1994, Bertolotti *et al*, 1996, Bertolotti *et al*, 1999). Typically, the TET

fusion genes encode chimeric transcription factors (TFs) in which the TAD is retained, while the RBD and RGG portions are replaced with the DNA-binding domain (DBD) of the fusion partner TF (Fig. 1B). Removal of the EWSR1 RBD and RGG boxes, which can repress transactivation, renders the TAD more potent in the chimeric protein (Bailly *et al*, 1994, Li and Lee, 2000).

The TET proteins are involved in transcription through their interaction with co-factors or DNA-binding proteins. Their direct interactions with distinct subpopulations of TFIID and subunits of RNA polymerase II link them to transcription initiation (Bertolotti *et al*, 1996, Bertolotti *et al*, 1998). EWSR1 functions as a transcriptional activator that associates with the co-activators p300 or CREB-binding protein (CBP) and is involved in POU5F1 (alias OCT-3/4)-, BRNA3A- and HNF4-mediated transcription (Rossow and Janknecht, 2001, Thomas and Latchman, 2002, Araya *et al*, 2003, Lee *et al*, 2005). FUS is involved in NFκB-mediated transcription and associates with the DBD of hormone receptors (Powers *et al*, 1998, Uranishi *et al*, 2001). TET proteins have also been linked to RNA splicing through their interaction with splicing factors, their presence in the spliceosome and involvement in splice site selection, possibly connecting the processes of transcription and splicing (Tan and Manley, 2009).

Mice with homozygous loss of *EWSR1* and *FUS* displayed postnatal mortality, male and/or female sterility and defective B-cell development and/or activation (Hicks *et al*, 2000, Kuroda *et al*, 2000, Li *et al*, 2007). The phenotypes were similar but gene-specific, and loss of *EWSR1* did not cause compensatory expression of *FUS* (Li *et al*, 2007). The mice also showed defective meiotic and homologous recombination processes and increased sensitivity to ionizing radiation, linking these proteins to DNA pairing and repair mechanisms. Moreover, FUS has been suggested to bind ssDNA and

mediate displacement-loop (D-loop) formation during DNA recombination and EWSR1 and FUS were found to be phosphorylated by DNA damage response proteins upon induction of DNA double strand breaks (Baechtold *et al*, 1999, Gardiner *et al*, 2008, Klevernic *et al*, 2009). Studies of the zebrafish (*Danio rerio*) EWSR1 orthologs suggest a role for EWSR1 also in maintaining mitotic integrity, as depletion of the proteins caused defective mitotic spindles and apoptosis (Azuma *et al*, 2007).

FUS has been found to be mutated in the neurodegenerative disease amyotrophic lateral sclerosis (Kwiatkowski *et al*, 2009, Vance *et al*, 2009). Through missense mutations, Arg residues in the CTD are replaced with other amino acids, often Cys, which leads to the mislocalization of FUS to the cytoplasm in the affected motor neurons. FUS has been found to bind RNA, presumably mRNA or pre-mRNA, both in the nucleus and cytoplasm and to be part of a kinesin-associated protein complex which can transport mRNA along neuronal microtubuli (Zinszner *et al*, 1997, Kanai *et al*, 2004). The mislocalization may impact mRNA transport and ultimately protein synthesis in the affected cells.

TET proteins can be post-transcriptionally regulated through Arg residue methylation, Ser/Thr phosphorylation or glycosylation, likely to affect their function, stability and localization (Tan and Manley 2009). However, very little is known about their transcriptional regulation, although it also directs the expression of the TET fusion genes. In Article I, we therefore characterized the activity of the *EWSR1* promoter region in relation to conservation patterns and localization of TF binding sites (TFBSs).

Rearrangement of TET genes in sarcomas

EWSR1, *FUS* and *TAF15* are involved in translocations that typically characterize different types of bone and soft tissue sarcomas (BSTSs) (Fig. 2). BSTSs are rare connective tissue malignancies that are heterogeneous in respect to their appearance, epidemiology, genetic alterations and progression (Fletcher *et al*, 2002, Osuna and de Álava, 2009). With few exceptions, the cell of origin and etiology of the tumor are unknown. As the diagnostic entities overlap morphologically, the distinction between them, or even between BSTSs and benign or non-neoplastic lesions, may be difficult. The identification of recurrent, tumor-type specific translocations in BSTSs with relatively simple karyotypes has aided the diagnosis and choice of treatment of these tumors and contributed to the unraveling of their pathogenesis (Osuna and de Álava, 2009). The classification of BSTSs is currently based on tumor morphology and molecular genetics, using a combination of techniques such as immunohistochemistry (IHC), cytogenetics and fluorescence *in situ* hybridization (FISH), and polymerase chain reaction (PCR)-based methods (Fletcher *et al*, 2002). Also, genome-wide analyses such as global gene expression arrays are increasingly used as classification tools and to identify markers of prognostic importance (Tschoep *et al*, 2007).

Of the TET members, *EWSR1* is most often rearranged as it has almost 20 identified partners, most of which are associated with BSTSs (Fig. 2) (Mitelman *et al*, 2010). Some of the TET fusion genes are remarkably specific for, and common in, certain tumor types, e. g. *EWSR1-FLI1* and *FUS-CREB3L2* in ESFT and low-grade fibromyxoid sarcoma (LGFMS), respectively (Turc-Carel *et al*, 1988, Panagopoulos *et al*, 2004). Other tumor types are characterized by TET fusion genes with alternating TET 5´-partners, e. g. *EWSR1-DDIT3* and *FUS-DDIT3* in myxoid liposarcoma, or *EWSR1-*

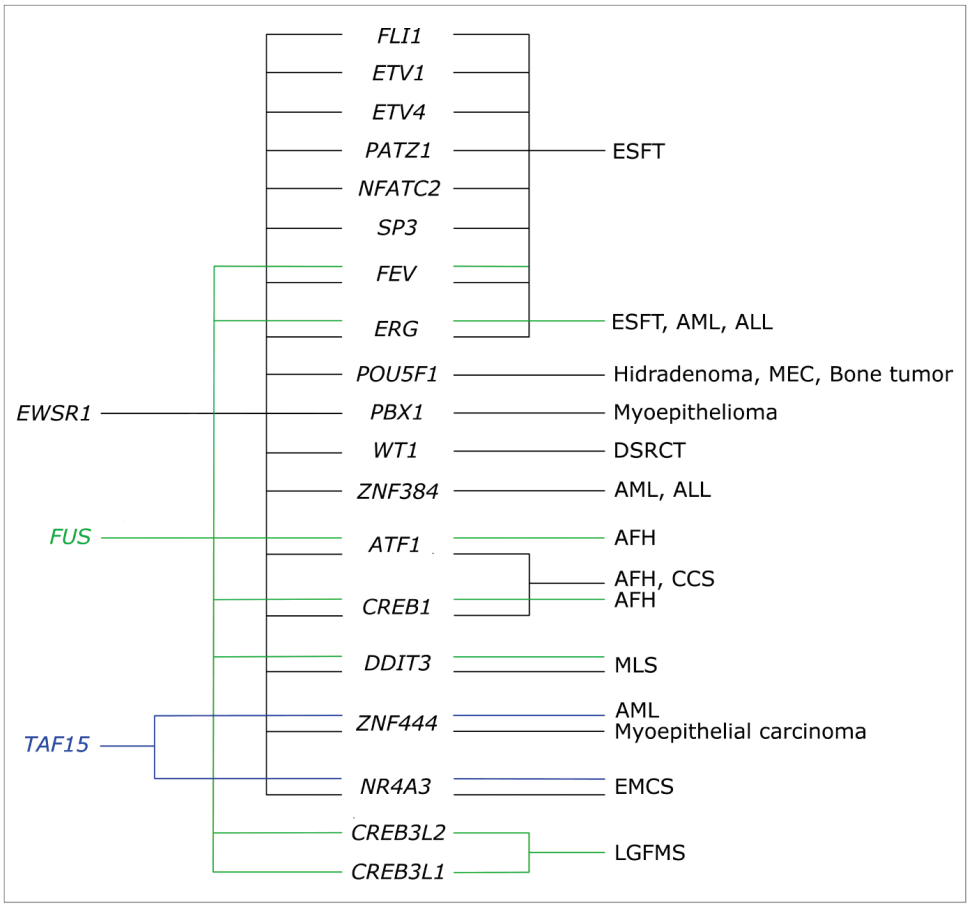

Figure 2. Summary of the known TET fusion genes (Mitelman *et al*, 2010). Horizontal lines connect the TET 5´-partner gene (*EWSR1*, *FUS*, *TAF15*) with their 3´-partners (in the middle), and the tumor type(s) which they characterize (to the right). AML = Acute myeloid leukemia, ALL = Acute lymphoblastic leukemia, DSRCT = Desmoplastic small round cell tumor, AFH = Angiomatoid fibrous histiocytoma, CCS = Clear cell sarcoma, EMCS = Extraskeletal myxoid chondrosarcoma. The TET fusions are most common in BSTSs and *EWSR1* is, among the TET genes, most often found in fusions.

NR4A3 and *TAF15-NR4A3* in extraskeletal myxoid chondrosarcoma (EMCS), reflecting the analogous functions of the TET proteins (Fig. 2). The observed tumor-type specificity suggests that the fusion protein is dependent

on a certain cell type to initiate tumorigenesis. However, the same fusion gene may be present in vastly different tumor types; the *EWSR1-POU5F1* fusion was found in an undifferentiated bone tumor (Yamaguchi *et al*, 2005) as well as in skin hidradenomas and salivary gland carcinomas (Article II). *FUS-ERG* was first identified in acute myeloid leukemia and subsequently in ESFT (Ichikawa *et al*, 1994, Panagopoulos *et al*, 1994, Shing *et al*, 2003), and it has been suggested that this chimera exerts different effects in fibroblasts and hematopoietic cells (Zou *et al*, 2005). It is possible that the chimeric protein is capable of initiating different tumorigenic programs depending on the cell type, giving rise to the specific tumor phenotype.

Tumorigenic mechanisms of TET chimeras

In translocation-positive cells, one wild-type (wt) allele of each fusion partner is lost and theoretically, two fusion genes (e. g. *EWSR1-FLI1* and *FLI1-EWSR1*) are created, both of which could contribute to tumorigenesis. Mice heterozygous for the *EWSR1* or *FUS* knock-out allele were indistinguishable from their litter mates (Hicks *et al*, 2000, Kuroda *et al*, 2000, Li *et al*, 2007), suggesting that the wt functions of EWSR1 and FUS are retained in the tumor cells. However, it was recently described that EWSR1-FLI1 can directly interfere with wt EWSR1 function and thereby cause mitotic defects, which may contribute to tumorigenesis through increased genomic instability (Embree *et al*, 2009). The reciprocal fusion gene (e. g. *FLI1-EWSR1*) is considered to be dispensable for tumor growth since it is typically not expressed or does not result in the production of a full length chimeric protein (Delattre *et al*, 1992, Panagopoulos *et al*, 2002, Panagopoulos *et al*, 2004). Hence, gained expression of the TET chimeric protein is believed to be crucial for tumorigenesis.

Ectopically expressed EWSR1 and FUS chimeras are able to transform cells and/or cause tumor formation in mice, and tumorigenicity is reduced when the fusion gene expression is abolished with siRNA (Herrero-Martín *et al*, 2009). Also, the tumor-type specificity of the fusion genes is reflected in the tumorigenic properties of the fusion proteins. The EWSR1 and FUS NTDs can substitute for each other in tumorigenic chimeric proteins, while the tumor phenotype depends on which partner, DDIT3 or FLI1, domain the chimera contains (Zinszner *et al*, 1994). Transgenic mice expressing FUS-DDIT3 in all tissues only develop adipocytic tumors resembling human liposarcomas, while mice expressing the DDIT3 or FUS domain alone are normal (Pérez-Losada *et al*, 2000a, Pérez-Losada *et al*, 2000b, Pérez-Mancera *et al*, 2002). Human fibrosarcoma cells expressing FUS-DDIT3 or full length DDIT3 induce liposarcomas in mice, whereas cells expressing the FUS NTD induce poorly differentiated sarcomas (Engström *et al*, 2006).

EWSR1-FLI1 and *EWSR1-ERG* induce a common ESFT-like (based on morphology, IHC and gene expression) transformation of human mesenchymal progenitor cells (MPCs) (Miyagawa *et al*, 2008, Riggi *et al*, 2008). However, these cells are not tumorigenic in mice, while *EWSR1-FLI1* transformed murine MPCs are (Riggi *et al*, 2005). Moreover, *FUS-DDIT3*, but not *EWSR1-ATF1*, expression cause murine MPCs to be tumorigenic in mice (Riggi *et al*, 2006). These findings suggest that MPCs may not be the appropriate recipient cell type for some fusion genes, or that the fusion gene is not sufficient for tumorigenesis. Secondary events are observed in BSTSs with TET fusion genes. The majority of ESFT cases carries additional recurrent genomic imbalances, including gains of the *EWSR1* and *FLI1* loci (Savola *et al*, 2009). However, few tumor types have been subjected to detailed genomic characterization. We therefore analyzed LGFMS tumors with SNP array

(Article IV) to investigate the presence of consistent genomic imbalances in this tumor type.

The TET chimeric proteins are believed to function as aberrant TFs that directly or indirectly cause anomalous expression of target genes. Through binding DNA directly, chimeras have been found to target several genes that are important for cell differentiation, survival or growth (Riggi *et al*, 2007, Kovar, 2010). Chimeras are generally thought of as transcriptional activators that may induce target gene expression by means of the potent TET TAD. EWSR1-FLI1, for example, is a stronger activator than the wt proteins through ETS sites, i. e. sequence motifs that are recognized by FLI1 and other members of the ETS family, which are also enriched in EWSR1-FLI1 immunoprecipitates (Bailly *et al*, 1994, Guillon *et al*, 2009). However, it has been suggested that EWSR1-FLI1 and EWSR1-ATF1 may also act as transcriptional repressors of some genes (Brown *et al*, 1995, Jishage *et al*, 2003, Kauer *et al*, 2009). Since chimeras can bind sequence motifs that are normally recognized by the wt partner TF, the target genes of the chimera can be expected to overlap with those of the fusion partner. Even so, the chimera may expand the target gene spectrum through the inappropriate expression of the fusion partner DBD, which can be achieved in different ways. For instance, since FLI1 expression is very restricted, the FLI1 domains become abnormally expressed through the fusion which is expressed from the ubiquitously active *EWSR1* promoter (Mhawech-Fauceglia *et al*, 2007, Andersson *et al*, 2008). FUS-DDIT3 is localized to the nucleus, whereas DDIT3 is expressed at low levels and localized to the cytosol under normal conditions. DDIT3 localizes to the nucleus only upon endoplasmatic reticulum (ER) stress caused by the accumulation of unfolded proteins in the ER (Oyadomari *et al*, 2004). Hence, FUS-DDIT3 expression brings the

DDIT3 domains to an abnormal nuclear localization. ATF1 activity is regulated by cyclic AMP (cAMP) through a phosphorylation site in the NTD which is not retained in EWSR1-ATF1, and the chimera is therefore rendered constitutively active (Fujimura *et al*, 1996).

The TET chimeras are also involved in transcription through protein interactions, just like the wt proteins. The association of the EWSR1 or FUS NTD with RNA polymerase II is maintained in the chimeras, while the interaction with splicing factors, which normally occurs through the CTD, is lost, inhibited or altered (Yang *et al*, 2000a, Yang *et al*, 2000b, Chansky *et al*, 2001). Also, EWSR1-FLI1 has been suggested to affect *CCND1* transcription in ESFT through lowering the RNA polymerase II elongation rate, resulting in the utilization of an upstream polyadenylation site and subsequent production of a more growth promoting CCND1 isoform (Sanchez *et al*, 2008). EWS-ETS chimeras have been found to bind some promoters in a ternary complex with AP-1 (FOS/JUN) (Kim *et al*, 2006), and EWSR1-FLI1 and EWSR1-ATF1 interact with CBP, potentially affecting the downstream apoptotic pathways (Fujimura *et al*, 2001, Ramakrishnan *et al*, 2004). Moreover, binding sites for other TFs have been found to coincide with ETS sites in EWSR1-FLI1 regulated genes, indicating that EWSR1-FLI1 may need to cooperate with these factors in target gene regulation (Kauer *et al*, 2009). FUS-DDIT3 interacts with NFKBIZ which affects NFκB-mediated transcription, interferes with a natural dimerization partner of DDIT3 and dysregulates key genes to block adipocytic differentiation (Peéez-Mancera *et al*, 2008, Göransson *et al*, 2009).

Nontheless, the knowledge about the tumorigenic mechanisms of many chimeras, e. g. FUS-CREB3L2, is still very limited. In Articles III and IV, we therefore studied the functions and DNA-binding preferences of the

wt CREB3L2 TF and the FUS-CREB3L2 chimera, and identified FUS-CREB3L2-dependent global gene expression patterns in LGFMS tumors.

THE PRESENT STUDY

AIMS

- To characterize the 5´-flanking regulatory region of *EWSR1* and delineate regions that are important for transcriptional activity in ESFT cell lines (Article I)
- To investigate the involvement of *EWSR1* rearrangements and *EWSR1-POU5F1* fusion in skin hidradenomas and salivary gland mucoepidermoid carcinomas (MECs) (Article II)
- To characterize the properties of the FUS-CREB3L2 chimera and the wt CREB3L2 transcription factor (Article III)
- To identify specific expression patterns that are associated with the presence of translocation t(7;16) and FUS-CREB3L2 chimera in LGFMS (Article IV)

MATERIALS AND METHODS

Patients and tumor samples

The majority of the tumor samples was obtained from patients treated at the orthopedic tumor centers at the Lund University Hospital or at the Karolinska Hospital in Stockholm. The samples were archived at the Department of Clinical Genetics, Lund. Paraffin-embedded tumor material was histopathologically reviewed and tumor cells were short-term cultured and subjected to cytogenetic characterization with chromosome banding analysis and FISH. When applicable, RNA and/or DNA was extracted from frozen specimens using standard molecular techniques for reverse transcriptase PCR (RT-PCR), real-time PCR, gene expression array or single nucleotide polymorphism (SNP) array.

In addition to the samples from Swedish patients, ESFT tumor biopsies were obtained from patients treated at the Medical University of Gdansk, Poland (Article I). The majority of ESFT cases carries the translocation t(11;22)(q24;q12) and expresses the *EWSR1-FLI1* fusion transcript. RT-PCR was used to confirm *EWSR1-FLI1* expression in the present cases. For Article II, we had access to RNA or cDNA samples from additional hidradenomas and MECs for RT-PCR analysis and interphase FISH. These samples were obtained from the Department of Pathology at the Sahlgrenska University Hospital in Gothenburg or from the Department of Pathology at the University Hospital in Leuven, Belgium. For Article IV, material from 24 LGFMS cases was available from Sweden and seven other tumor centers in Europe and the U. S. A. (listed in Mertens *et al*, 2005), and used for gene expression array (19 cases), IHC analyses (nine cases) and/or SNP array analysis (nine cases). All cases but one harbored the LGFMS-specific translocation t(7;16)(q33;p11) and/or expressed the *FUS-CREB3L2*

fusion transcript. The remaining case harbored the rare variant translocation t(11;16)(p11;p11) and expressed *FUS-CREB3L1*. A few cases lacked the t(7;16) at cytogenetic analysis and instead had one or more supernumerary ring chromosomes. Myxofibrosarcoma (MFS), desmoid fibromatosis (DFM), solitary fibrous tumor (SFT) and EMCS tumors, which share histological features with LGFMS (Billings *et al*, 2005, Guillou *et al*, 2007, Weiss and Goldblum, 2008, Meng *et al*, 2009), were included as comparison to the LGFMS group in the gene and protein expression analyses. All but one of these samples were obtained from patients treated at the orthopedic tumor centers in Lund or Stockholm. One EMCS sample was obtained from the Department of Cancer Genetics at the Norwegian Radium Hospital in Oslo. All EMCS were characterized by the translocations t(9;22)(q22;q12) or t(9;17)(q22;q11) and showed expression of the *EWSR1-NR4A3* or *TAF15-NR4A3* fusion transcripts. The MFS cases were of the low-grade variant with 37-58 chromosomes (Mentzel *et al*, 1996). For Article IV, we also constructed a tissue microarray (TMA) with material from LGFMS, MFS, DFM, SFT and EMCS tumors to facilitate the IHC analyses. For the TMA construction, tissue columns were cut from histologically representative areas of the paraffin-embedded material and inserted into recipient TMA paraffin blocks with 50 cases on each block.

All samples were obtained after informed consent and the studies were approved by the local ethics committees.

Cell lines

Established cell lines are preferred for *in vitro* studies since they are usually easily cultured and transfected, and generally generate reproducible results. For Article I, several ESFT cell lines were available. These cell lines

may carry additional and different aberrations compared to ESFT tumors, but at least have the expression of *EWSR1-FLI1* in common which is specific and important for ESFT tumorigenesis. However, for many BSTSs, established cell lines are not readily available and human embryonic kidney (HEK293) or murine embryonic fibroblast (NIH3T3) cells that grow rapidly and are easily manipulated were chosen in these instances.

For Article I, the established ESFT cell lines A-673, RD-ES, SK-ES-1, SK-NEP-1 and SK-PN-DW were purchased from the American Type Culture Collection (ATCC) and MHH-ES-1, RH-1, SK-N-MC and TC-71 from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ). The cell line TTC-466 was donated by Dr P. H. Sorensen at the Department of Pathology, Children's and Women's Hospital of British Columbia, Vancouver, Canada. The HEK293 and NIH3T3 cell lines, purchased from the Interlab Cell Line Collection (ICLC), were used for Articles I-IV. All cells were cultured according to the manufacturer's recommendations. When required, RNA and/or protein was extracted from cultured cells using standard molecular techniques and used for RT-PCR, real-time PCR, gene expression array or western blot, respectively. The cells were transfected with various reporter gene constructs and/or expression plasmids to investigate promoter activity (Article I), transcriptional activation potential and activation of transcription through specific regulatory sequences (Articles II-IV) and subcellular localization of proteins (Article III), as discussed below. HEK293 cells were also transfected with FUS-CREB3L2-pCMV-Tag2B (Stratagene, La Jolla, CA, USA) constructs (Article IV) which allows for G418 (geniticin) selection of stable transfectants. In this construct, the *FUS-CREB3L2* cDNA is cloned adjacent to the FLAG epitope, downstream of the CMV promoter (cytomegalovirus immediate early promoter). Nuclear expression of NTD

FLAG-tagged FUS-CREB3L2 was detected with western blot and immunocytochemistry using an anti-FLAG M2 antibody (Stratagene). Cells were transfected using Effectene, Polyfect (Qiagen, Hilden, Germany) or FuGENE (Roche Applied Science, Indianapolis, IN, USA) according to the manufacturers' recommendations.

In silico analyses of regulatory sequences

The basal (or core) promoter regions, where the general transcriptional machinery assembles, are perhaps the most straight-forward to study among the regulatory sequences due to their location directly upstream of transcription start sites (TSS). The basal transcription is modulated by TFs that bind sites in the proximal promoter region, which extends a few hundred bp upstream of the core promoter. Eukaryotic promoters are diverse and it is now understood that only a minority has the characteristic TATA(T/A)/(T/A) and/or CCAAT sequences in the basal and proximal promoters, respectively. Instead, the presence of CpG islands was found to correlate the best with promoter activity (Cooper *et al*, 2006). The distal promoter region contains TFBSs that further tune the expression of the downstream gene, perhaps in a cell-type specific or temporal manner (Farnham, 2009). Distal regions may extend downstream of the TSS to involve intronic regions and the regulatory elements of a gene can be located far away, even on a different chromosome (Miele and Dekker, 2008). Moreover, alternating and bi-directional promoters can regulate the same gene and are abundant in the genome (Trinklein *et al*, 2004, Kimura *et al*, 2006). These aspects make the definition of promoter outer boundaries arbitrary and the selection of putatively biologically important TFBSs challenging. Utilizing a combination of different programs that predict the localization of TSSs,

promoter sequences, TFBSs, CpG islands and conserved regions may aid this decision. Naturally, coupling the *in silico* analysis with *in vitro* experiments is necessary for evaluating biological relevance.

Most of the TFBS prediction programs, e. g. MatInspector (Quandt *et al*, 1995), utilize the TRANSFAC or JASPAR databases in which TFBSs are described as positional weight matrices (PWMs). A PWM is generated through aligning a set of TFBS sequences, identified from *in vitro* studies, and calculating the frequency of each base at each sequence position to create a nucleotide distribution matrix. Each nucleotide is then given a value (weight) depending on the conservation of the individual nucleotide position in the matrix; the maximum value 100 for total conservation and the minimum value zero for equal distribution of the four bases. The consecutive nucleotides with the highest weights constitute the TFBS core sequence (Quandt *et al*, 1995). MatInspector and Match (Kel *et al*, 2003) utilize PWMs to search an unknown sequence for matches to the consensus matrix descriptions and core sequences. In the simplest case, TFBSs can be identified by sequence comparison to the consensus sequence utilizing the program Patch (Wingender, 2008).

For the study of the *EWSR1* promoter region (Article I), the sequence which extends from 1600 bp upstream (-1600) to 400 bp downstream (+400) of the TSS that covers the most conserved regions, was characterized. The TFBS prediction programs MatInspector, Match and Patch, the CpG rich region identification tool CpGPlot (Larsen *et al*, 1992), the comparative genomics tool GenomeVISTA (Couronne *et al*, 2003) and multiple alignment tool MultiAlin (Corpet, 1988) were used for the *in silico* analyses of the sequence. MultiAlin was also used for the alignment of the amino acid sequences of the CREB3 family basic and leucine zipper regions

(Article III). In Article IV, the pattern of TFBSs in the promoters of the genes that were specifically expressed in LGFMS tumors and FUS-CREB3L2 expressing cells was analyzed with the program SMART (systematic motif analysis retrieval tool) (Veerla *et al*, 2010). RefSeq mRNA and UCSC promoter sequence information and TRANSFAC v 2009.1 PWMs were utilized to identify TFBSs in the -1500 to +500 sequence of each gene's TSS, as well as in the conserved (between human and mouse) regions within the -5000 +1000 sequence. TFBSs with a significant enrichment (P_E) and presence (P_P) in the gene list were identified. The P_E value is the probability of finding the observed number of instances of a given TFBS, and P_P the probability of finding the observed number of promoters with at least one instance of a given TFBS, in randomly generated gene lists. SMART was also used to cluster genes that display a similar TFBS pattern in their promoters (promoter clustering). Moreover, the regulatory sequences of *CD24* and *FOXL1* were investigated *in silico* using MatInspector, Patch, CpGPlot, GenomeVISTA and Promoterscan (Prestridge, 1995) (Article IV). Promoterscan identifies regions which contain TFBSs that are more frequent in eukaryotic RNA polymerase II promoters than in non-promoter sequences (Prestridge, 1995).

Studying transcriptional activity *in vitro*

The utilization of reporter genes has greatly facilitated the investigation of gene expression and protein interactions in living cells and organisms. The green fluorescent protein (GFP) and firefly luciferase (FLuc), derived from the jellyfish *Aequorea victoria* and firefly *Photinus pyralis*, respectively, have been widely used for these purposes. Enhanced GFP (EGFP) exhibits green fluorescence when exposed to blue light and was utilized to monitor the subcellular localization of the full length and truncated

CREB3L2 and FUS-CREB3L2 proteins in Article III. Throughout the study (Articles I-IV), FLuc was utilized to measure transcriptional activity using the dual luciferase reporter (DLR) assay system (Promega, Madison, WI, USA), where the Fluc enzymatic activity is measured as light emission upon addition of the FLuc substrate to the sample (cell lysate). The subsequently added *Renilla reniformis* luciferase (RLuc) substrate simultaneously quenches the FLuc reaction and initiates the RLuc reaction, and light of a different wavelength is emitted. Typically, the expression of *FLuc*, and thus also the FLuc activity, is influenced by the experimental conditions whereas the *RLuc* expression (and activity) serves as the internal control. The FLuc luminescence is normalized to that of RLuc, expressing the results as ratios; FLuc:RLuc.

Promoter activity

The two kbp *EWSR1* upstream sequence was characterized in vector pFLhRL where the luciferase genes are transcribed in opposite directions with *RLuc* under the control of the SV40 promoter (simian virus 40 early enhancer/promoter) (Fig. 3A, top) (Möller *et al*, 2007). Thereby, the use of two individual *Luc* reporter gene plasmids, which may influence the transcription of each other (Huszár *et al*, 2001), is avoided. Nine deletion constructs were generated that contained specific but overlapping portions of the *EWSR1* promoter region, cloned upstream of *FLuc*. In this way, functional regions of the sequence can be identified that are important for promoter function or are likely to recruit negatively or positively acting factors. In addition, different portions of the sequence surrounding the TSSs of *EWSR1* and the nearby gene *RHBDD3* were cloned in the opposite direction in order to estimate their bidirectional activity.

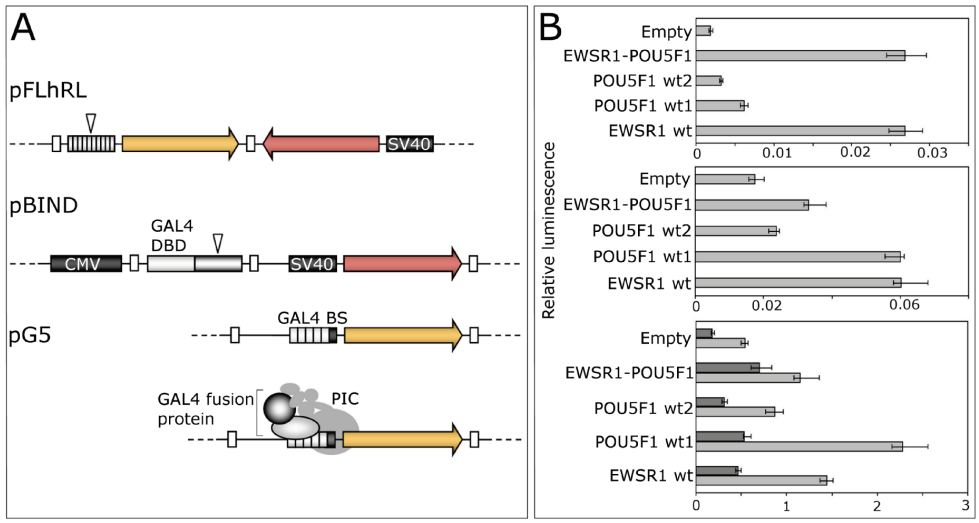

Figure 3. A) In pFLhRL, a promoter region can be cloned upstream (arrowhead) of *FLuc* (yellow), while *RLuc* (red) is constitutively expressed from the SV40 promoter in the same vector. In pBIND, the cDNA encoding a presumed transcriptional activator (arrowhead) is cloned in-frame to the GAL4 DBD cDNA, downstream of the CMV promoter. The GAL4 fusion protein binds the GAL4 binding sites (BS) upstream of the basal promoter (small grey box), of the pG5 vector and may stimulate the assembly of the preinitiation complex (PIC) to enhance the expression of *FLuc*. Small white boxes = SV40 polyadenylation signals. B) In Article II, EWSR1-POU5F1 was found to be a stronger transcriptional activator than the two wt POU5F1 isoforms (top) using the pBIND system. Also, EWSR1-POU5F1 activated transcription through the POU5F1 binding sites of the *NANOG* promoter (middle) and *POU5F1* distal enhancer (bottom), compared to the regulatory sequence itself ("empty"). However, POU5F1 wt1 was the strongest activator through these sites. The *POU5F1* enhancer was cloned upstream of the SV40 promoter, grey bars represent the activity of the SV40 promoter alone. X axes show the relative luminescence (FLuc:RLuc ratios).

Transcriptional activation potential

To investigate if EWSR1-POU5F1 and wt POU5F1 (Article II), wt CREB3L2 and FUS-CREB3L2 (Article III), can function as transcriptional activators, their respective cDNAs were cloned in-frame to the GAL4 DBD

encoding sequence of the pBIND vector (Promega) (Fig. 3A, middle). The GAL4-fusion transcript is constitutively expressed from the CMV promoter. The pBIND constructs were co-transfected with the pG5*luc* vector (Promega) which contains GAL4 DBD binding sites upstream of a minimal basal promoter that drives the expression of *FLuc* (Fig. 3A, middle). A transcriptional activator will stimulate the assembly and stabilization of the preinitiation complex (PIC) and basal transcription machinery and thereby increase the expression of *FLuc*, as compared to the GAL4 DBD alone (Fig. 3A, bottom, and 3B).

Activation of transcription through regulatory sequences

The DNA-binding preferences of the EWSR1-POU5F1, wt POU5F1, FUS-CREB3L2 and wt CREB3L2 proteins were investigated in Articles II and III-IV, respectively. Upon removal of the GAL4 DBD, the pBIND constructs mentioned above (constitutively expressing each of the proteins) were co-transfected with plasmids containing specific TFBSs, enhancer elements or promoter regions of putative target genes cloned upstream of *Fluc*. In Article IV, the regulatory sequences of *CD24* or *FOXL1* were cloned in pFLhRL and co-transfected with pCR3.1-based FUS-CREB3L2 expression plasmids (without reporter gene). Regardless of the vector system used, the regulatory sequence may activate the *FLuc* expression on its own. However, if the transcriptional activator acts directly or indirectly through the cloned site to stimulate transcription, the expression of *FLuc* markedly increases (Fig. 3B, middle and bottom).

Determining gene expression patterns and levels

Tumor RNA was reverse transcribed into cDNA and used as template in RT-PCR for the detection of fusion transcripts (Articles I and II) and in quantitative real-time PCR for the estimation of gene expression (Articles I and IV). Northern blot, in which the RNA is size separated through gel electrophoresis and hybridized to a specific probe, was used in Article III to analyze *CREB3L2* expression in human tissues. This way, the complete transcript length can be determined. For the global detection of gene expression in Article IV, tumor cDNA was labeled and hybridized onto the Human GeneChip® Gene 1.0 ST array which detects whole transcripts with 764,885 distinct probe sets (Affymetrix, Santa Clara, CA, USA).

Gene expression levels of a sample are always formulated relative the expression levels of another sample. Ideally in gene expression analysis, a tumor sample is compared to the corresponding normal tissue from the same patient (paired samples). But even then, between samples variation may remain due to different routines in handling biopsies or extracting RNA, interfering with the downstream data analyses. The etiology, as well as the corresponding normal tissue, is unknown for many BSTSs and thus, proper paired samples are rarely available. In the real-time PCR analyses, the expression level of the gene-of-interest was compared to that of a housekeeping gene, giving a relative estimate of the gene expression levels (Articles I and IV). As the LGFMS cell of origin is uncertain, the LGFMS group was compared to tumor types with a similar morphology in the gene expression array analysis (Article IV). Previous studies have emphasized that the distinction of LGFMS from MFS and DFM may be difficult (Folpe *et al*, 2000, Billings *et al*, 2005, Guillou *et al*, 2007), thereby motivating the current comparison. Moreover, the program and approach chosen for data analysis,

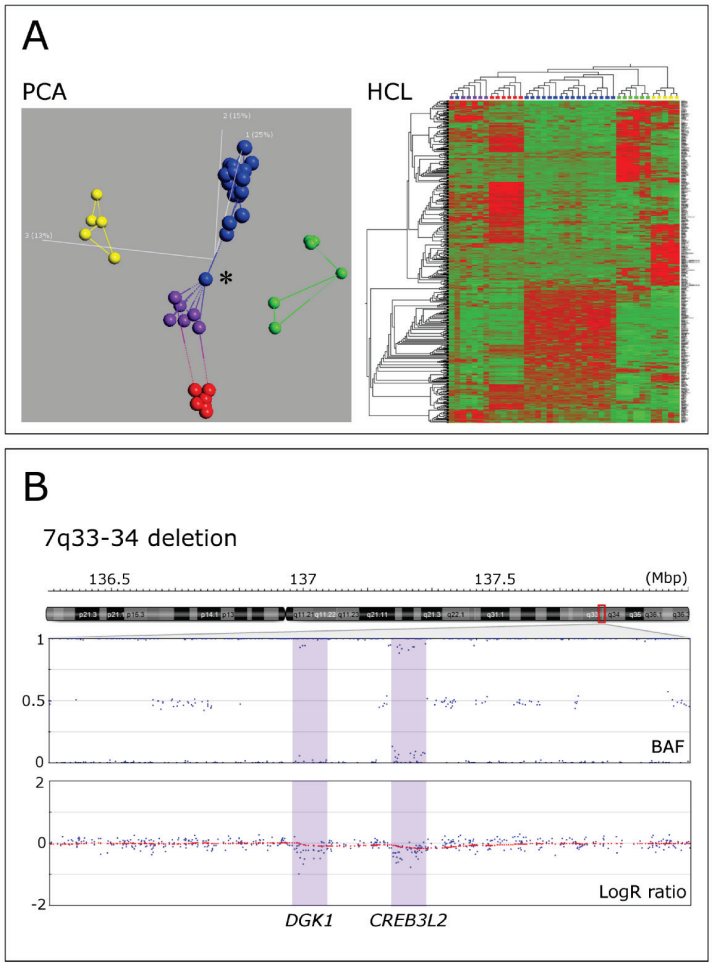

Figure 4. A) Analysis of gene expression data with PCA (3D) and HCL (2D). The tumor groups; LGFMS (blue), MFS (purple), DFM (red), EMCS (green) and SFT (yellow) were clearly different from each other and an outlier LGFMS sample (asterisk) was identified (Article IV). B) Identification of recurrent deletions (shaded areas) at and nearby the *CREB3L2* breakpoints in LGFMS tumors, with SNP array (Article IV). The deletions are visible by deviating B allele frequency (BAF) and LogR ratio. The mean Log R ratio is marked with a red line.

which involves finding similarity patterns in the data and performing statistical tests, can influence the interpretation of the data. This potential problem may be circumvented by unsupervised examination of the data, i. e. allowing the data to organize themselves without presumptions, using different programs and methods, e. g. hierarchical clustering analysis (HCL) and principal component analysis (PCA). PCA with the program Qlucore Omics Explorer v

2.0 (Qlucore AB, Lund, Sweden) allows for interactive imaging of multidimensional data (e. g. gene expression data) in a three dimensional (3D) space, which facilitates the visualization of similarity patterns in the data set (Fig. 4A).

Detection of genomic imbalances

Genomic imbalances such as amplifications and deletions are present in many cancers and typically point out regions that contain oncogenes or tumor suppressor genes, respectively. As there are more than one million known SNPs distributed throughout the genome, they provide a high-resolution platform for the estimation of copy number. Since SNPs provide allele-specific information, copy number neutral loss-of-heterozygosity (LOH) events, e. g. uniparental disomy, or mono-allelic amplifications can also be detected.

In Article IV, we used SNP array analysis to identify genomic imbalances in LGFMS tumor samples (Fig. 4B). Tumor DNA was amplified, fragmented, hybridized and labeled on the Human Omni 1-Quad v 1.0 array (Illumina, San Diego, CA, USA) which assays more than 1 million selected SNPs. Access to a patient-matched normal DNA sample is helpful for the identification of normal copy number variations. However, such samples are not always available. In the normalization procedure, the tumor sample data are compared to those of 120 normal individuals, thereby aiding the identification of normal genomic variation.

RESULTS AND DISCUSSION

Article I

Bidirectionality and transcriptional activity of the *EWSR1* promoter region.

The TET proteins are ubiquitously expressed in human tissues, displaying both specific and overlapping expression patterns (Andersson *et al*, 2008). Investigation of the first 600 bp of the *EWSR1* 5´-regulatory region revealed that it lacked TATA and CAAT-boxes, was GC-rich and had several putative TSSs (Plougastel *et al*, 1993). However, the close head-to-head arrangement of *EWSR1* and *RHBDD3* was not recognized in that study. When analyzing two kbp of the *EWSR1* regulatory sequence in the present study, a putative CpG-island was identified that covered the intergenic region and extended beyond the first exons of both genes. This, in addition to the short distance (100-400 bp depending on the definition of the *EWSR1* TSS) between the *EWSR1* and *RHBDD3* TSSs, are recognized features of bi-directional promoters (Adachi and Lieber, 2002, Trinklein *et al*, 2004). Peaks in conservation were observed for the intergenic region and sequence directly upstream of the *EWSR1* TSS. These regions are included in the -400 to +400 portion which was found to be important for promoter activity, also when cloned in the reverse orientation. Deleting part of the upstream sequence (-1600 to -440) had little effect on promoter activity. By real-time PCR, *RHBDD3* was expressed in ESFT biopsies and cell lines and ubiquitously expressed across tissues, just like *EWSR1*.

When cloned in the reverse orientation, fragments including the 400 bp sequence downstream of the *EWSR1* TSS had the highest activity. In contrast, deletion of the +65 to +400 portion was necessary to achieve high activity with fragments cloned in the forward orientation. The discrepancy between these observations could be explained by the involvement of

orientation-dependent TFBSs in the regulation of one of the genes. Conserved SP binding sites and an E-box, which is recognized by MYC, MAX and USF factors, are present in this region and may be involved in the regulation of this sequence. The narrow region directly downstream of the *EWSR1* TSS was most important for promoter activity and is likely to be required for basal transcription. A binding site for the ubiquitously expressed factor YY1 can be found in this region, perhaps suggesting its involvement in transcription initiation. However, in order to identify active sites that are crucial for promoter activity, TFBS mutation analyses would be required.

The presence of SP sites and E-box sequences, and lack of a TATA box, within a CpG island that covers the TSS, is more common in house-keeping gene promoters than in those with tissue-specific activity (Rozenberg *et al*, 2008, Zhu *et al*, 2008). Taken together, our results suggest that *EWSR1* and *RHBDD3* are regulated in a house-keeping manner through a common bi-directional promoter. Thus, it is likely that also the expression of *EWSR1* fusion genes is regulated in the same way. Decreased *RHBDD3* expression has been associated with cancer. In pituitary tumors, CpG-island methylation seems to be the causative mechanism behind the decreased expression, whereas in colorectal tumors the mechanism is unknown (Bahar *et al*, 2004, Bahar *et al*, 2007). As the *EWSR1* promoter is predicted to regulate two genes with divergent involvement in several different tumor types, there is a need for understanding the regulatory networks acting through this sequence.

Article II

POU5F1, encoding a key regulator of stem cell pluripotency, is fused to *EWSR1* in hidradenoma of the skin and mucoepidermoid carcinoma of the salivary glands.

The index case was diagnosed as a hidradenoma based on the histopathological examination. The cytogenetic and molecular genetic analyses revealed the presence of a translocation t(6;22)(p21;q12) and expression of an *EWSR1-POU5F1* fusion transcript as the sole aberration. We could also show that the tumor cells displayed strong nuclear, and weak cytoplasmic, expression of POU5F1 by IHC with an antibody against the POU5F1 CTD. Thereby, this was the first report of an *EWSR1* fusion gene in epithelial tumors.

Hidradenoma, which primarily affects adults, is a benign, epidermal sweat gland neoplasm which contains different cell types of varying proportions (McNiff *et al*, 2006). The variant in which glycogen-containing clear cells predominate (clear cell hidradenoma) has been found to harbor the translocation t(11;19)(q21;p13) and *CRTC1-MAML2* fusion, while fusion-negative tumors were mainly composed of poroid and cuticular cells (Winnes *et al*, 2007). When analyzing additional cases of *CRTC1-MAML2*-negative hidradenomas with no or few (≤ 10-15%) clear cells, we identified the *EWSR1-POU5F1* fusion in two cases and *EWSR1* split signals in two cases. MEC is another glandular tumor which primarily affects adults, contains different cell types at varying proportions and harbors the translocation t(11;19)(q21;p13) and the *CTRC1-MAML2* fusion. Previous studies suggested that *CTRC1-MAML2*-positive MEC tumors occur in younger individuals and are highly differentiated, low-grade and smaller (Behboudi *et al*, 2006, Okabe *et al*, 2006). Another study identified the *CTRC1-MAML2* fusion in tumors of

all grades (Tirado *et al*, 2007) and recently, *CDKN2A* (*p16*) deletions were associated with unfavorable outcomes of *CTRC1-MAML2*-positive MEC tumors, suggesting that the fusion is an early event and that additional aberrations are needed to develop metastatic disease (Anzick *et al*, 2010). Due to the genetic and biological similarities between hidradenoma and MEC, we included poorly differentiated *CTRC1-MAML2*-negative MECs in the present study and found an *EWSR1* split signal and *EWSR1-POU5F1* fusion in one case. Our findings provide further evidence of a genetic link between hidradenoma and MEC, indicating that common molecular pathways may be involved in the development of these tumors.

In the fusion transcripts detected in the present study, a part of *EWSR1* exon 6 was fused in-frame to a part of *POU5F1* exon 2. Previously, an *EWSR1-POU5F1* fusion was identified in one case of undifferentiated bone tumor where exon 6 of *EWSR1* was fused to exon 1 of *POU5F1* (Yamaguchi *et al*, 2005). Both fusion transcript variants are predicted to encode the NTD EWSR1 and the DBD of POU5F1, analogous to other EWSR1 chimeric proteins. As expected, we could show that the EWSR1-POU5F1 chimera is a stronger transcriptional activator than wt POU5F1 using the GAL4-based vector system (Fig. 3B). POU5F1 is one of the key regulators of embryonic stem cell pluripotency and is known to bind the pou5f1/sox2 composite element in the *Pou5f1* distal enhancer and in the promoter of the target gene *Nanog* (Chew *et al*, 2005, Rodda *et al*, 2005). The EWSR1-POU5F1 chimera activated transcription more efficiently than the POU5F1 isoform 2 (wt2) through these sites, yet weaker than the longer wt POU5F1 isoform 1 (wt1) and wt EWSR1 (Fig. 3B). It cannot be excluded that the chimera might prefer other sequence motifs, or primarily act through indirect mechanisms. As EWSR1 lacks an obvious DBD, it seems likely that

the activation through these elements by EWSR1 occurs through protein interactions. Interestingly, the NTD of EWSR1 has been found to associate with the POU5F1 central POU domain, which is retained in the chimera, and enhance POU5F1-mediated transcription (Lee *et al*, 2005).

The chimera identified in the present study (called EWS-OCT-4B) was recently found to be localized to the nucleus, to be a weaker transcriptional activator than the first identified EWSR1-POU5F1 chimera, to directly bind the POU5F1 octamer motif and to be tumorigenic in mice (Kim *et al*, 2010). Also, upregulation of the POU5F1 target genes *FGF-4* and *NANOG* was observed in cells with stable expression of EWS-OCT-4B. However, the DNA-binding and regulatory actions of the chimera were not compared to those of the full length wt proteins, and it is not established whether the target gene activation is direct or indirect. The results from this study are therefore consistent with our results in that the chimera is capable of activating transcription through the octamer motif, although it may be a relatively weak transcriptional activator.

POU5F1 is expressed during embryogenesis and maintained only in the germ cell lineage of the adult. It has been proposed that POU5F1 is expressed in adult differentiated tissues and tumors, however not all reports securely distinguish *POU5F1* from its transcribed pseudogenes, POU5F1 wt 1 from the putative pseudogene-encoded protein POU5F1P1, or between the POU5F1 isoforms (de Jong and Loijenga, 2006, Kotoula *et al*, 2008, Panagopoulos *et al*, 2008). The translocation t(6;22) may constitute a mechanism of POU5F1 reactivation, whereby the POU5F1 DBD becomes inappropriately expressed, causing the anomalous regulation of target genes and contributing to de-differentiation.

Article III

Characterization of the native CREB3L2 transcription factor and the FUS/CREB3L2 chimera.

CREB3L2 (alias *BBF2H7*) was first identified as part of the *FUS-CREB3L2* fusion gene, the result of translocation t(7;16)(q32-34;p11), in LGFMS (Storlazzi *et al*, 2003). The t(7;16) or variant t(11;16)(p11;p11) and the corresponding chimeras *FUS-CREB3L2* or *-CREB3L1* (alias *OASIS*) were subsequently found to be specific for LGFMS (Panagopoulos *et al*, 2004). The knowledge about wt CREB3L2 is limited, but CREB3L1 is structurally and functionally related to CREB3L3 (CREB-H), CREB3L4 (CREB4), CREB3 (Luman) and ATF6. These proteins are ER resident through their transmembrane (TM) domains and are activated by regulated intramembrane proteolysis (RIP) in response to the accumulation of misfolded proteins in the ER (ER-stress) (Bailey and O´Hare, 2007). Upon ER-stress, the cleaved fragment which contains the basic and leucine zipper (bZIP) domains, is translocated to the nucleus where it can activate transcription of target genes, e. g. *HSPA5* (*GRP78*). It has been suggested that the target gene activation acts to rescue the cell from ER-stress induced apoptosis (Kondo *et al*, 2005, Kondo *et al*, 2007).

In the present study, we showed that also CREB3L2 belongs to this protein family. By *in silico* analysis, the predicted bZIP and TM amino acid sequences of CREB3L2 are highly similar to those of CREB3L1/L3/L4, CREB3 and ATF6, with the highest similarity to CREB3L1. Like *CREB3L1*, *CREB3L2* expression was detected in various normal tissues. Moreover, the *in vitro* subcellular localization experiments showed that CREB3L2 was localized to the ER, while the truncated protein CREB3L2ΔTM, which lacks the TM and luminal domains, was localized to the nucleus. CREB3L2

corresponds to the cleaved fragment which would be produced by an activated RIP pathway. Analogous to CREB3L2, we found that FUS-CREB3L2 was cytosolic and FUS-CREB3L2ΔTM nuclear. This suggests that CREB3L2 and FUS-CREB3L2 may undergo RIP, just like CREB3L1.

By utilizing the GAL4 system, we found that CREB3L2 can function as a transcriptional activator and the TAD was mapped to the first 120 aa. This part is replaced with the more potent TAD of the FUS NTD in FUS-CREB3L2, and the chimera was found to be a stronger transcriptional activator than CREB3L2. The full length andΔTM CREB3L2 and FUS - CREB3L2, respectively, proteins were equally efficient as activators. We also investigated the transcriptional activity of CREB3L2 and FUS-CREB3L2 through the box-B element, cAMP responsive element (CRE), ATF6 binding sites and the *HSPA5* (*GRP78*) promoter, which have been found to be activated by CREB3L1 and CREB3L1ΔTM (Omori *et al*, 2002, Kondo *et al*, 2005). CREB3L2 activated transcription through these sites and CREB3L2ΔTM was a stronger activator than the full length protein, analogous to CREB3L1ΔTM. FUS-CREB3L2 was the strongest activator through the cloned sites but, surprisingly, FUS-CREB3L2ΔTM was the weakest. Another study identified CREB3L2 to be an ER-stress induced TF and that CREB3L2ΔTM directly binds CRE sites in the *HSPA5* promoter, which is in agreement with our results (Kondo *et al*, 2007).

The structural and functional similarities between CREB3L2 and CREB3LI suggest that the FUS-CREB3L2/L1 chimeras might have analogous functions in LGFMS. The oncogenic role of these chimeras is unknown but can be hypothesized to entail anomalous regulation of genes normally controlled by CREB3L2/L1. Although FUS-CREB3L2ΔTM was not the strongest activator through the selected TFBSs, it was located in the

nucleus and was capable of activating transcription through CREB3L2/L1 response elements. As it has been suggested that the CREB3L2 protein is not expressed under normal conditions (Kondo *et al*, 2005, Kondo *et al*, 2007), the inappropriate expression of the CREB3L2 bZIP DBD through the chimera may be sufficient for anomalous target gene regulation, despite a relatively weak transactivation capability.

Article IV

Low-grade fibromyxoid sarcoma has a specific expression profile with upregulation of *CD24* and *FOXL1*

LGFMS is a rare, indolent, soft tissue tumor which typically arises in the deep, intramuscular tissue of young adults. The tumor cells are bland, spindle-shaped, of fibroblastic or myofibroblastic differentiation and surrounded by a myxoid or collagenized matrix (Folpe *et al*, 2002). However, LGFMS may also display areas of hypercellularity, nuclear pleomorphism and capillary infiltration, and has a potential for local recurrence and metastasis. Consequently, LGFMS may be confused with more benign tumors such as DFM, or more aggressive tumors such as MFS (Mentzel *et al*, 1996, Billings *et al*, 2005, Weiss and Goldblum, 2008). In the present study, we compared the gene expression profile of LGFMS with that of MFS, DFM, SFT and EMCS, and could show that LGFMS is clearly distinct from the other entities (Fig. 4A). LGFMS was most similar to MFS and secondly to DFM on the gene expression level.

The detection of t(7;16) or t(11;16) and the *FUS-CREB3L2/L1* fusion transcripts, are the most solid diagnostic criteria for LGFMS (Panagopoulos *et al*, 2004). A subset of LGFMS cases expresses *FUS-CREB3L2* but lacks the t(7;16) and instead harbors one or more

supernumerary ring chromosomes, and it has been shown that the fusion gene may reside in the ring (Bartuma *et al,* 2010). By SNP array analysis of LGFMS tumors, we found no other recurrent genomic aberrations other than translocation t(7;16)-associated deletions and gain of 7q in cases with ring chromosomes, emphasizing the importance of the fusion gene event in LGFMS (Fig. 4B). The detected deletions remove most of the *FUS* and *CREB3L2* portions not involved in the fusion gene, thereby explaining the absence of reciprocal fusion gene expression (Panagopoulos *et al*, 2004). Gain of 7q in two cases, and of a small 16p11 segment in one case, with supernumerary ring chromosomes is in agreement with previous analyses of ring chromosomes in LGFMS (Mezzelani *et al*, 2000, Bartuma *et al,* 2010). Gain of 7q has been identified as a recurrent secondary event also in translocation t(12;16)-positive cases of myxoid liposarcoma (Mandahl *et al*, 1994).

The FUS-CREB3L2 chimera may cause the anomalous expression of CREB3L2 target genes. However, little is known about these target genes and the known CREB3L2/L1 target *HSPA5* was not upregulated in the LGFMS tumors or in cells with stable expression of FUS-CREB3L2 (FC-HEK) of the present study. Since CREB3L2 and CREB3L2ΔTM were found to preferentially bind CRE sites (Kondo *et al*, 2007), we hypothesized that CREB3L2 and FUS-CREB3L2 regulate genes that contain CRE sites in their promoters. Indeed, many of the LGFMS upregulated genes contained CRE half-sites in their promoters, although these sites were not significantly overrepresented. Instead, we found significant enrichment (P_E) and presence (P_P) of binding sites for FOX factors. Of these, *FOXL1* and *FOXF1* are located close to each other on 16q24 and were specifically upregulated in LGFMS (Fig. 5A), indicating that FOX factors are important for LGFMS.

FOXL1 was the top upregulated gene in LGFMS but was not specifically upregulated in FC-HEK, suggesting that the regulation of this gene is not directly caused by FUS-CREB3L2. In support of this, FUS-CREB3L2ΔTM only had a small effect on the *FOXL1* promoter (which contains one CRE half-site) in luciferase assays (Fig. 5B, right). *Foxl1* and *Foxf1* are regulated by Gli1 and Gli2 in the Hedgehog (Hh) pathway in the mouse (Madison *et al*, 2009). Hh ligands and targets, e. g. *FOXL1*, *FOXF1*, *PTCH1*, *SHH*, *BCL2* and to some extent *GLI1* and *GLI2*, were expressed at higher levels in LGFMS, although not displaying LGFMS-specific expression, suggesting a role for this pathway in LGFMS.

The subset of genes that was specifically upregulated both in LGFMS and FC-HEK may contain putative FUS-CREB3L2 target genes. *CD24* was among those genes and we could show that FUS-CREB3L2ΔTM and CREB3L2ΔTM activated transcription from a *CD24* regulatory sequence which contains two CRE half-sites (Fig. 5B, left). CD24 is a glycosylated cell surface mucin which is expressed in various carcinomas where cytoplasmic CD24 expression, compared to membraneous, has been associated with poor prognosis and metastasis (Kristiansen *et al*, 2004, Lee *et al,* 2009). Additional experiments such as TFBS mutation analyses and/or ChIP-chip or gel-shift assay investigation would be needed to characterize the FUS-CREB3L2 interaction with *CD24*.

The expression of selected LGFMS-specific genes on the protein level (CASR, CD24, CREB3L2, FOXL1, MUC4, PRKCB1, PTN, SCUBE1 and TMPRSS2) is currently being examined. Proteins that show LGFMS-specific expression could potentially be used as LGFMS markers, but this would, however, need to be evaluated in larger series of LGFMS and differential diagnostic entities.

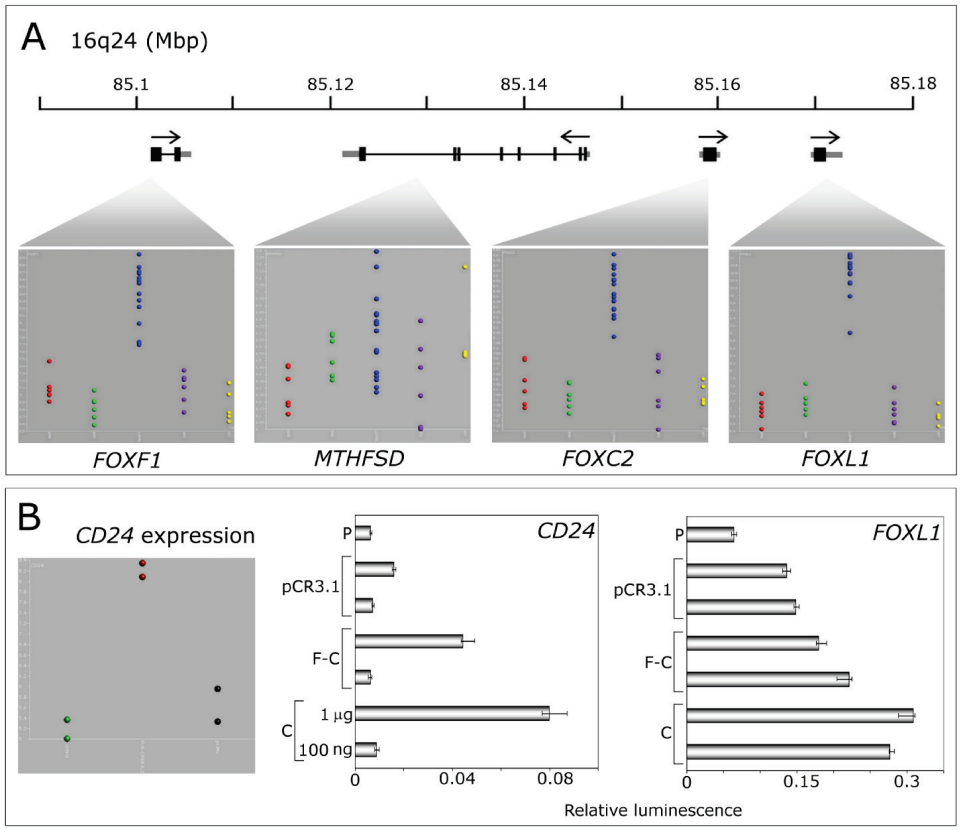

Figure 5. A) *FOXL1* was the top upregulated gene in LGFMS (blue), compared to the DFM (red), EMCS (green), MFS (purple) and SFT (yellow) groups. *FOXL1*, *FOXF1* and *FOXC2* are located at 16q24 and all upregulated in LGFMS. Binding sites for FOXL1 and FOXF1 were significantly enriched in the promoters of LGFMS upregulated genes. *MTHFSD* is transcribed in the opposite direction and was not specifically upregulated in LGFMS. B) *CD24* was specifically upregulated in FUS-CREB3L2 expressing cells (dark red), compared to cells transfected with empty vector (dark grey) and control cells (green). The regulatory sequences of *CD24* (middle) and *FOXL1* (right), which contain CRE half-sites, were cloned in pFLhRL and transfected together with 100 ng or 1 µg of FUS-CREB3L2ΔTM (F-C)- or CREB3L2ΔTM (C)-expression plasmids, or empty vector (pCR3.1). FUS-CREB3L2ΔTM activated transcription much more efficiently through the *CD24* sequence, compared to that of *FOXL1*, but weaker than CREB3L2ΔTM. P = the activity of the sequence alone.

CONCLUSIONS

Article I

- The short, conserved intergenic sequence between the *RHBDD3* and *EWSR1* TSSs is important for promoter activity
- Regions of most importance for promoter activity, in both orientations, are localized to the area surrounding the TSSs
- The promoter activity of some regions are likely to be orientation or cell type-dependent
- *EWSR1* and *RHBDD3* are expressed across normal tissues and in ESFT tumors and cell lines
- The *EWSR1* promoter region shares features with bi-directional and house-keeping gene promoters, and is likely to regulate also *RHBDD3*

Article II

- Recurrent *EWSR1* rearrangments may be found also in epithelial tumor types
- The *EWSR1-POU5F1* fusion provides a further genetic link between hidradenoma and MEC
- The EWSR1-POU5F1 chimera iss a stronger transcriptional activator than wt POU5F1 and can activate transcription from POU5F1 binding sites, although not as efficiently as the wt proteins
- Translocation may be a mechanism of POU5F1 reactivation

Article III

- CREB3L2 is similar to CREB3L1 and related proteins, and expressed across normal tissues
- CREB3L2 and FUS-CREB3L2 seem to undergo RIP, just like CREB3L1
- The TAD of CREB3L2 is weaker than that of FUS and accordingly, the FUS-CREB3L2 chimera is a stronger transcriptional activator than CREB3L2
- FUS-CREB3L2 is the most efficient activator through the selected TFBSs and the *HSPA5* promoter, whereas FUS-CREB3L2ΔTM is the weakest

Article IV

- The gene expression profile of LGFMS is distinct from histologically similar tumors
- The only recurrent genomic aberrations in LGFMS are associated with the formation of the fusion gene or supernumerary ring chromosomes
- *FOXL1* is specifically upregulated in LGFMS and may be involved in the regulation of a subset of the LGFMS upregulated genes that contain overrepresented FOX binding sites
- *CD24* is specifically upregulated in LGFMS and FUS-CREB3L2 expressing cells
- FUS-CREB3L2 activated transcription of *CD24*, suggesting that the expression of this gene is FUS-CREB3L2 dependent, whereas the expression of *FOXL1* is not

SUMMARY IN SWEDISH

(SVENSK SAMMANFATTNING)

Cancer är en genetisk sjukdom som orsakas av förvärvade förändringar i arvsmassan. Translokationer, vilka innebär att delar av två kromosomer bryts av och byter plats, är exempel på sådana förändringar. Konsekvensen av denna händelse kan vara att delar av gener vid kromosombrottpunkterna sammanförs till att bilda en ny gen, en s. k. fusionsgen. Fusionsgenen producerar ett fusionsprotein som tros vara mycket viktigt för tumöruppkomst. Avhandlingen har fokuserat på fusionsgener som involverar generna *EWSR1* och *FUS*, vilka karakteriserar olika typer av ben- och mjukdelssarkom, maligna tumörer som drabbar kroppens stödjevävnad. Eftersom det kan vara svårt att skilja sarkom från benigna tumörtyper, samt mellan olika typer av sarkom, har de tumörtypspecifika fusionsgenerna varit viktiga för att ställa diagnos och bestämma behandlingsstrategi.

I Studie I karakteriserades *EWSR1* genens promotor eftersom den styr uttrycket av de fusionsgener som involverar *EWSR1*. Vi identifierade små regioner som var viktiga för promotoraktivitet och fann att *EWSR1* promotorn troligtvis också reglerar *RHBDD3*, en annan cancerrelaterad gen. Våra resultat utgör ett viktigt fundament för att förstå hur dessa gener regleras. I Studie II identifierades fusionsgenen *EWSR1-POU5F1* i två olika körteltumörtyper, det första fyndet av en *EWSR1* fusionsgen i epiteliala tumörtyper. Utöver de histopatologiska likheter som dessa tumörtyper uppvisar, indikerar de genetiska fynden att de aktuella tumörtyperna har liknande ursprung. Vi visade också att fusionsproteinet EWSR1-POU5F1 var

en effektiv trankriptionsaktivator, i likhet med andra EWSR1 fusionsprotein, som kunde aktivera transkription genom DNA sekvenser som normalt påverkas av POU5F1. Det är möjligt att EWSR1-POU5F1 bidrar till tumöruppkomst genom att aktivera gener som normalt regleras av POU5F1.

I Studierna III och IV undersöktes fusionsgenen *FUS-CREB3L2* som är specifik för låggradiga fibromyxoidsarkom (LGFMS). LGFMS drabbar yngre vuxna, har malign potential och kan vara svår att skilja från både mer benigna och mer maligna tumörtyper. I Studie III fann vi att fusionsproteinet FUS-CREB3L2 var en effektivare transkriptionsaktivator än vildtypsproteinet CREB3L2, och att det kunde aktivera transkription genom DNA sekvenser som normalt känns igen av CREB3L2. I Studie IV analyserades LGFMS tumörmaterial med genexpressionsarray för att identifiera gener som är specifikt upp- eller nedreglerade i LGFMS. Bl a *FOXL1* och *CD24* var signifikant uppreglerade. Celler med stabilt uttryck av FUS-CREB3L2 genererades och vi fann att *CD24* var uppreglerad även i dessa celler. FUS-CREB3L2 kunde också påvisas aktivera transkription genom DNA sekvenser som tros reglera *CD24* uttryck. Det är möjligt att *CD24* uttrycket i LGFMS är beroende av FUS-CREB3L2 och att FUS-CREB3L2 generellt bidrar till tumöruppkomst genom att direkt påverka uttrycket av vissa gener. Våra resultat är viktiga för vidare studier av FUS-CREB3L2 fusionsproteinets roll i tumörutveckling.

ACKNOWLEDGEMENTS

I am sincerely grateful to everyone who has contributed to the completion of this thesis work;

Fredrik Mertens, who was actually my first contact with the Department of Clinical Genetics, and more importantly, who has been my supervisor. You are impressively supportive, open, trusting, knowledgeable and positive (yes, actually!), and I am happy to have worked with you.

Ioannis Panagopoulos for always taking the time for whatever is needed, being supportive and teaching me a lot of the things I have learned as a PhD student.

My past and present co-supervisors **Nils Mandahl** and **Thoas Fioretos** for generously sharing your experiences, reading manuscripts and having time for methodological discussions.

Co-authors and **collaborators** for your invaluable experiences, inputs and efforts.

Felix Mitelman, the former head of the Department of Clinical Genetics, for establishing a truly welcoming, supportive, non-competitive and inspiring environment.

The technical staff, **Andrea Horvat**, **Carin Lassen**, **Jenny Nilsson**, **Linda Magnusson, Margareth Isaksson** and **Marianne Rissler** for their skills which they gladly and patiently share with anyone who may need it, and for being great company! A special thanks to **Margareth** who has helped out a lot with the practical lab work.

The senior staff **Anna Andersson**, **Anna Collin**, **Anette Welin**, **Bertil Johansson**, **Catarina Lundin**, **Charlotte Jin**, **David Gisselsson Nord**, **Marcus Heidenblad**, **Maria Soller**, **Nina Larsson**, **Petra Johnels**, **Samuel Gebre-Medhin**, **Tord Jonson**, **Ulf Kristoffersson** and **Yuesheng Jin**. Thank you for the nice activities ouside the lab. A special thanks to **Anette** and **Eva Lindstedt** for helping out with everything without hesitation.

Past and present PhD and post doc collegues; **Björn Nilsson**, **David Lindgren**, **Gisela Lundberg**, **Hammurabi Bartuma**, **Helena Ågerstam**, **Henrik Liljebjörn**, **Jenny Karlsson**, **Josef Davidsson**, **Kajsa Paulsson**, **Karolin Hansén Nord**, **Kristina Karrman**, **Linda Holmquist Mengelbier**, **Marcus Järås**, **Nils Hansen**, **Srinivas Veerla** and **Ylva Stéwenius**. Thank you for all the "fika" moments, town lunches and out-of-the lab get togethers!

Students **Ildikó Frigyesi** and **Lina Andersson** for working very hard on their master thesis projects with me.

The staff at the clinic for your nice company, hard work and contributions to creating a happy work environment!

Mikael, for being amazingly supportive, encouraging, generous, kind, loving and such a good friend.

My family and **friends** for endless support and joy.

The research presented in this thesis was supported by the Swedish Children's Cancer Foundation, the Swedish Cancer Society, the Swedish Research Council, the Gunnar Nilsson´s Cancer Foundation, the IngaBritt and Arne Lundberg Research Foundation and the Royal Physiographic Society of Lund.

REFERENCES

Adachi N, Lieber MR. Bidirectional gene organization: a common architectural feature of the human genome. *Cell.* 2002;**109**:807-9.

Andersson MK, Ståhlberg A, Arvidsson Y, Olofsson A, Semb H, Stenman G, Nilsson O, Åman P. The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. *BMC Cell Biol.* 2008;**9**:37.

Anzick SL, Chen WD, Park Y, Meltzer P, Bell D, El-Naggar AK, Kaye FJ. Unfavorable prognosis of *CRTC1-MAML2* positive mucoepidermoid tumors with *CDKN2A* deletions. *Genes Chromosomes Cancer.* 2010;**49**:59-69.

Araya N, Hirota K, Shimamoto Y, Miyagishi M, Yoshida E, Ishida J, Kaneko S, Kaneko M, Nakajima T, Fukamizu A. Cooperative interaction of EWS with CREB-binding protein selectively activates hepatocyte nuclear factor 4-mediated transcription. *J Biol Chem.* 2003;**278**:5427-32.

Azuma M, Embree LJ, Sabaawy H, Hickstein DD. Ewing sarcoma protein Ewsr1 maintains mitotic integrity and proneural cell survival in the zebrafish embryo. *PLoS ONE.* 2007;**2**:e979.

Baechtold H, Kuroda M, Sok J, Ron D, Lopez BS, Akhmedov AT. Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein TLS/FUS and is able to promote D-loop formation. *J Biol Chem.* 1999;**274**:34337-42.

Bahar A, Simpson DJ, Cutty SJ, Bicknell JE, Hoban PR, Holley S, Mourtada-Maarabouni M, Williams GT, Clayton RN, Farrell WE. Isolation and characterization of a novel pituitary tumor apoptosis gene. *Mol Endocrinol.* 2004;**18**:1827-39.

Bahar A, Whitby P, Holley S, Hoban PR, Elder JB, Deakin M, Hall C, Clayton RN, Williams GT, Farrell WE. Primary colorectal tumors fail to express the proapoptotic mediator PTAG and its reexpression augments drug-induced apoptosis. *Genes Chromosomes Cancer.* 2007;**46**:202-12.

Bailey D, O´Hare P. Transmembrane bZIP transcription factors in ER stress signaling and the unfolded protein response. *Antioxid Redox Signal.* 2007;**9**:2305-21.

Bailly RA, Bosselut R, Zucman J, Cormier F, Delattre O, Roussel M, Thomas G, Ghysdael J. DNA-binding and transcriptional activation properties of the EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing sarcoma. *Mol Cell Biol.* 1994;**14**:3230-41.

Bartuma H, Möller E, Collin A, Domanski HA, Vult von Steyern F, Mandahl N, Mertens F. Fusion of the *FUS* and *CREB3L2* genes in a supernumerary ring chromosome in low-grade fibromyxoid sarcoma. *Cancer Genet Cytogenet.* 2010, *in press.*

Behboudi A, Enlund F, Winnes M, Andrén Y, Nordkvist A, Leivo I, Flaberg E, Szekely L, Mäkitie A, Grenman R, Mark J, Stenman G. Molecular classification of mucoepidermoid carcinomas-prognostic significance of the *MECT1-MAML2* fusion oncogene. *Genes Chromosomes Cancer.* 2006;**45**:470-81.

Bertolotti A, Lutz Y, Heard DJ, Chambon P, Tora L. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *Embo J.* 1996;**15**:5022-31.

Bertolotti A, Melot T, Acker J, Vigneron M, Delattre O, Tora L. EWS, but not EWS-FLI-1, is associated with both TFIID and RNA polymerase II: interactions between two members of the TET family, EWS and hTAFII68, and subunits of TFIID and RNA polymerase II complexes. *Mol Cell Biol.* 1998;**18**:1489-97.

Bertolotti A, Bell B, Tora L. The N-terminal domain of human $TAF_{II}68$ displays transactivation and oncogenic properties. *Oncogene.* 1999;**18**:8000-10.

Billings SD, Giblen G, Fanburg-Smith JC. Superficial low-grade fibromyxoid sarcoma (Evans tumor). A clinicopathologic analysis of 19 cases with a unique observation in the pediatric population. *Am J Surg Pathol.* 2005;**29**:204-10.

Brown AD, Lopez-Terrada D, Denny C, Lee KA. Promoters containing ATF-binding sites are de-regulated in cells that express the EWS/ATF1 oncogene. *Oncogene.* 1995;**10**:1749-56.

Chansky HA, Hu M, Hickstein DD, Yang L. Oncogenic TLS/ERG and EWS/Fli-1 fusion proteins inhibit RNA splicing mediated by YB-1 protein. *Cancer Res.* 2001;**61**:3586-90.

Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, Ng HH. Reciprocal transcriptional regulation of *Pou5f1* and *Sox2* via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol.* 2005;**25**:6031-46.

Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res.* 2006;**16**:1-10.

Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 1988;**16**:10881-90.

Couronne O, Poliakov A, Bray N, Ishkhanov T, Ryaboy D, Rubin E, Pachter L, Dubchak I. Strategies and tools for whole-genome alignments. *Genome Res.* 2003;**13**:73-80.

Crozat A, Åman P, Mandahl N, Ron D. Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. *Nature.* 1993;**363**:640-4.

De Jong J, Looijenga LHJ. Stem cell marker OCT3/4 in tumor biology and germ cell tumor diagnostics: History and future. *Crit Rev Oncog.* 2006;**12**:171-203.

Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G, Aurias A, Thomas G. Gene fusion with an *ETS* DNA-binding domain caused by chromosome translocation in human tumours. *Nature.* 1992;**359**:162-5.

Embree LJ, Azuma M, Hickstein DD. Ewing sarcoma fusion protein EWSR1/FLI1 interacts with EWSR1 leading to mitotic defects in zebrafish embryos and human cell lines. *Cancer Res.* 2009;**69**:4363-71.

Engström K, Willén H, Kåbjörn-Gustafsson C, Andersson C, Olsson M, Göransson M, Järnum S, Olofsson A, Warnhammar E, Åman P. The myxoid/round cell liposarcoma fusion oncogene *FUS-DDIT3* and the normal

DDIT3 induce a liposarcoma phenotype in transfected human fibrosarcoma cells. *Am J Pathol.* 2006;**168**:1642-53.

Farnham PJ. Insights from genomic profiling of transcription factors. *Nat Rev Genet.* 2009;**10**:605-16.

Fletcher CDM, Unni KK, Mertens F (eds). *World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Soft Tissue and Bone.* IARC Press: Lyon, 2002.

Folpe AL, Lane KL, Paull G, Weiss SW. Low-grade fibromyxoid sarcoma and hyalinizing spindle cell tumor with giant rosettes. A clinicopathologic study of 73 cases supporting their identity and assessing the impact of high-grade areas. *Am J Surg Pathol.* 2000;**24**:1353-60.

Folpe, van den Berg E, Molenaar WM. Low grade fibromyxoid sarcoma. In: *World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Soft Tissue and Bone.* Fletcher CDM, Unni KK, Mertens F (eds). IARC Press: Lyon, 2002;104-5.

Fujimura Y, Ohno T, Siddique H, Lee L, Rao VN, Reddy ES. The *EWS-ATF-1* gene involved in malignant melanoma of soft parts with t(12;22) chromosome translocation, encodes a constitutive transcriptional activator. *Oncogene.* 1996;**12**:159-67.

Fujimura Y, Siddique H, Lee L, Rao VN, Reddy ES. EWS-ATF-1 chimeric protein in soft tissue clear cell sarcoma associates with CREB-binding protein and interferes with p53-mediated *trans*-activation function. *Oncogene.* 2001;**20**:6653-9.

Gardiner M, Toth R, Vandermoere F, Morrice NA, Rouse J. Identification and characterization of FUS/TLS as a new target of ATM. *Biochem J.* 2008;**415**:297-307.

Guillou L, Benhattar J, Gengler C, Gallagher G, Ranchère-Vince D, Collin F, Terrier P, Terrier-Lacombe MJ, Leroux A, Marquès B, de Saint Aubain Somerhausen N, Keslair F, Pedeutour F, Coindre JM. Translocation-positive low-grade fibromyxoid sarcoma: Clinicopathologic and molecular analysis of a series expanding the morphologic spectrum and suggesting potential

relationship to sclerosing epithelioid fibrosarcoma. A study from the French Sarcoma Group. *Am J Surg Pathol*. 2007;**31**:1387-402.

Guillon N, Tirode F, Boeva V, Zynovyev A, Barillot E, Delattre O. The oncogenic EWS-FLI1 protein binds *in vivo* GGAA microsatellite sequences with potential transcriptional activation function. *PLoS One*. 2009;**4**:e4932.

Göransson M, Andersson MK, Forni C, Ståhlberg A, Andersson C, Olofsson A, Mantovani R, Åman P. The myxoid liposarcoma FUS-DDIT3 fusion oncoprotein deregulates NF-κB target genes by interaction with NFKBIZ. *Oncogene*. 2009;**28**:270-8.

Herrero-Martín D, Osuna D, Ordóñez JL, Sevillano V, Martins AS, Mackintosh C, Campos M, Madoz-Gúrpide J, Otero-Motta AP, Caballero G, Amaral AT, Wai DH, Braun Y, Eisenacher M, Schaefer KL, Poremba C, de Álava E. Stable interference of EWS-FLI1 in an Ewing sarcoma cell line impairs IGF-1/IGF-1R signalling and reveals TOPK as a new target. *Br J Cancer*. 2009;**101**:80-90.

Hicks GG, Singh N, Nashabi A, Mai S, Bozek G, Klewes L, Arapovic D, White EK, Koury MJ, Oltz EM, Van Kaer L, Ruley HE. *Fus* deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nat Genet*. 2000;**24**:175-9.

Huszár T, Mucsi I, Terebessy T, Masszi A, Adamkó S, Jeney C, Rosivall L. The use of a second reporter plasmid as an internal standard to normalize luciferase activity in transient transfection experiments may lead to a systematic error. *J Biotechnol*. 2001;**88**:251-8.

Ichikawa H, Shimizu K, Hayashi Y, Ohki M. An RNA-binding protein gene, *TLS/FUS*, is fused to *ERG* in human myeloid leukemia with t(16;21) chromosomal translocation. *Cancer Res*. 1994;**54**:2865-8.

Jishage M, Fujino T, Yamazaki Y, Kuroda H, Nakamura T. Identification of target genes for EWS/ATF-1 chimeric transcription factor. *Oncogene*. 2003;**22**:41-9.

Kanai Y, Dohmae N, Hirokawa N. Kinesin transports RNA: Isolation and characterization of an RNA-transporting granule. *Neuron*. 2004;**43**:513-25.

Kauer M, Ban J, Kofler R, Walker B, Davis S, Meltzer P, Kovar H. A molecular function map of Ewing´s sarcoma. *PLoS One*. 2009;**4**:e5415.

Kel AE, Göβling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, Wingender E. MATCH^{TM}: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res*. 2003;**31**:3576-9.

Kim S, Denny CT, Wisdom R. Cooperative DNA binding with AP-1 proteins is required for transformation by EWS-Ets fusion proteins. *Mol Cell Biol*. 2006;**26**:2467-78.

Kim S, Lim B, Kim J. *EWS-Oct-4B*, an alternative *EWS-Oct-4* fusion gene, is a potent oncogene linked to human epithelial tumours. *Br J Cancer*. 2010;**102**:436-46.

Kimura K, Wakamatsu A, Suzuki Y, Ota T, Nishikawa T, Yamashita R, Yamamoto JI, Sekine M, Tsuritani K, Wakaguri H, Ishii S, Sugiyama T, Saito K, Isono Y, Irie R, Kushida N, Yoneyama T, Otsuka R, Kanda K, Yokoi T, Kondo H, Wagatsuma M, Murakawa K, Ishida S, Ishibashi T, Takahashi-Fujii A, Tanase T, Nagai K, Kikuchi H, Nakai K, Isogai T, Sugano S. Diversification of transcriptional modulation: Large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res*. 2006;**16**:55-65.

Klevernic IV, Morton S, Davis RJ, Cohen P. Phosphorylation of Ewing´s sarcoma protein (EWS) and EWS-Fli1 in response to DNA damage. *Biochem J*. 2009;**419**:625-34.

Kondo S, Murakami T, Tatsumi K, Ogata M, Kanemoto S, Otori K, Iseki K, Wanaka A, Imaizumi K. OASIS, a CREB/ATF-family member, modulates UPR signalling in astrocytes. *Nat Cell Biol*. 2005;**7**:186-94.

Kondo S, Saito A, Hino SI, Murakami T, Ogata M, Kanemoto S, Nara S, Yamashita A, Yoshinaga K, Hara H, Imaizumi K. BBF2H7, a novel transmembrane bZIP transcription factor, is a new type of endoplasmic reticulum stress transducer. *Mol Cell Biol*. 2007;**27**:1716-29.

Kotoula V, Papamichos SI, Lambropoulos AF. Revisiting OCT4 expression in peripheral blood mononuclear cells. *Stem Cells*. 2008;**26**:290-1.

Kovar H. Downstream EWS/FLI1 - upstream Ewing´s sarcoma. *Genome Med.* 2010;**2**:8.

Kristiansen G, Sammar M, Altevogt P. Tumour biological aspects of CD24, a mucin-like adhesion molecule. *J Mol Histol.* 2004;**35**:255-62.

Kuroda M, Sok J, Webb L, Baechtold H, Urano F, Yin Y, Chung P, de Rooij DG, Akhmedov A, Ashley T, Ron D. Male sterility and enhanced radiation sensitivity in *TLS*$^{-/-}$ mice. *Embo J.* 2000;**19**:453-62.

Kwiatkowski TJ Jr, Bosco DA, LeClerc AL, Tamrazian E, Vanderburg CR, Russ C, Davis A, Gilchrist J, Kasarskis EJ, Munsat T, Valdmanis P, Rouleau GA, Hosler BA, Cortelli P, de Jong PJ, Yoshinaga Y, Haines JL, Pericak-Vance MA, Yan J, Ticozzi N, Siddique T, McKenna-Yasek D, Sapp PC, Horvitz HR, Landers JE, Brown RH Jr. Mutations in the *FUS/TLS* gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science.* 2009;**323**:1205-8.

Larsen F, Gundersen G, Lopez R, Prydz H. CpG islands as gene markers in the human genome. *Genomics.* 1992;**13**:1095-107.

Lee J, Rhee BK, Bae GY, Han YM, Kim J. Stimulation of Oct-4 activity by Ewing´s sarcoma protein. *Stem Cells.* 2005;**23**:738-51.

Lee JH, Kim SH, Lee ES, Kim YS. CD24 overexpression in cancer development and progression: a meta-analysis. *Oncol Rep.* 2009;**22**:1149-56.

Li H, Watford W, Li C, Parmelee A, Bryant MA, Deng C, O´Shea J, Lee SB. Ewing sarcoma gene *EWS* is essential for meiosis and B lymphocyte development. *J Clin Invest.* 2007;**117**:1314-23.

Li KKC, Lee KAW. Transcriptional activation by the Ewing´s sarcoma (EWS) oncogene can be cis-repressed by the EWS RNA-binding domain. *J Biol Chem.* 2000;**275**:23053-8.

Madison BB, McKenna LB, Dolson D, Epstein DJ, Kaestner KH. FoxF1 and FoxL1 link hedgehog signaling and the control of epithelial proliferation in the developing stomach and intestine. *J Biol Chem.* 2009;**284**:5936-44.

Mandahl N, Mertens F, Åman P, Rydholm A, Brosjö O, Willén H, Mitelman F. Nonrandom secondary chromosome aberrations in liposarcomas with t(12;16). *Int J Oncol.* 1994;**4**:307-10.

May WA, Lessnick SL, Braun BS, Klemsz M, Lewis BC, Lunsford LB, Hromas R, Denny CT. The Ewing´s sarcoma *EWS/FLI-1* fusion gene encodes a more potent transcriptional activator and is a more powerful transforming gene than *FLI-1*. *Mol Cell Biol.* 1993;**13**:7393-8.

McNiff J, McCalmont T, Requena L, Sangueza O, Vassallo C, Rosso R, Borroni G, Glusac E, Pichardo R. Benign Tumours with apocrine and eccrine differentiation. In: *World Health Organization Classification of Tumours: Pathology and Genetics of Skin Tumours*. LeBoit P, Burg G, Weedon D, Sarasin A (eds). IARC Press: Lyon, 2006;139-51.

Meng GZ, Zhang HY, Bu H, Geng JG. Low-grade fibromyxoid sarcoma versus fibromatosis: A comparative study of clinicopathological and immunohistochemical features. *Diagn Cytopathol.* 2009;**37**:96-102.

Mentzel T, Calonje E, Wadden C, Camplejohn RS, Beham A, Smith MA, Fletcher CDM. Myxofibrosarcoma. Clinicopathologic analysis of 75 cases with emphasis on the low-grade variant. *Am J Surg Pathol.* 1996;**20**:391-405.

Mertens F, Fletcher CDM, Antonescu CR, Coindre JM, Colecchia M, Domanski HA, Downs-Kelly E, Fisher C, Goldblum JR, Guillou L, Reid R, Rosai J, Sciot R, Mandahl N, Panagopoulos I. Clinicopathologic and molecular genetic characterization of low-grade fibromyxoid sarcoma, and cloning of a novel *FUS/CREB3L1* fusion gene. *Lab Invest.* 2005;**85**:408-15.

Mezzelani A, Sozzi G, Nessling M, Riva C, Della Torre G, Testi MA, Azzarelli A, Pierotti MA, Lichter P, Pilotti S. Low-grade fibromyxoid sarcoma: a further low-grade soft tissue malignancy characterized by a ring chromosome. *Cancer Genet Cytogenet.* 2000;**122**:144-8.

Mhawech-Fauceglia P, Herrmann FR, Bshara W, Odunsi K, Terracciano L, Sauter G, Cheney RT, Groth J, Penetrante R. Friend leukaemia integration-1 expression in malignant and benign tumours: A multiple tumour tissue microarray analysis using polyclonal antibody. *J Clin Pathol.* 2007;**60**:694-700.

Miele A, Dekker J. Long-range chromosomal interactions and gene regulation. *Mol Biosyst*. 2008;**4**:1046-57.

Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer*. 2007;**7**:233-45.

Mitelman F, Johansson B, Mertens F (eds). Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer, *http://cgap.nci.nhi.gov/Chromosomes/Mitelman*. 2010.

Miyagawa Y, Okita H, Nakaijima H, Horiuchi Y, Sato B, Taguchi T, Toyoda M, Katagiri YU, Fujimoto J, Hata J, Umezawa A, Kiyokawa N. Inducible expression of chimeric EWS/ETS proteins confers Ewing´s family tumor-like phenotypes to human mesenchymal progenitor cells. *Mol Cell Biol*. 2008;**28**:2125-37.

Möller E, Isaksson M, Mandahl N, Mertens F, Panagopoulos I. Comparison of the proximal promoter regions of the *PAX3* and *PAX7* genes. *Cancer Genet Cytogenet*. 2007;**178**:114-9.

Okabe M, Miyabe S, Nagatsuka H, Terada A, Hanai N, Yokoi M, Shimozato K, Eimoto T, Nakamura S, Nagai N, Hasegawa Y, Inagaki H. *MECT1-MAML2* fusion transcript defines a favorable subset of mucoepidermoid carcinoma. *Clin Cancer Res*. 2006;**12**:3902-7.

Omori Y, Imai JI, Suzuki Y, Watanabe S, Tanigami A, Sugano S. OASIS is a transcriptional activator of CREB/ATF family with a transmembrane domain. *Biochem Biophys Res Commun*. 2002;**293**:470-7.

Osuna D, de Álava E. Molecular pathology of sarcomas. *Rev Recent Clin Trials*. 2009;**4**:12-26.

Oyadomari S, Mori M. Roles of CHOP/GADD153 in endoplasmic reticulum stress. *Cell Death Differ*. 2004;**11**:381-9.

Panagopoulos I, Åman P, Fioretos T, Höglund M, Johansson B, Mandahl N, Heim S, Behrendtz M, Mitelman F. Fusion of the *FUS* gene with *ERG* in acute myeloid leukemia with t(16;21)(p11;q22). *Genes Chromosomes Cancer*. 1994;**11**:256-62.

Panagopoulos I, Mertens F, Dêbiec-Rychter M, Isaksson M, Limon J, Kardas I, Domanski HA, Sciot R, Perek D, Crnalic S, Larsson O, Mandahl N. Molecular genetic characterization of the EWS/ATF1 fusion gene in clear cell sarcoma of tendons and aponeuroses. *Int J Cancer*. 2002;**99**:560-7.

Panagopoulos I, Storlazzi CT, Fletcher CDM, Fletcher JA, Nascimento A, Domanski HA, Wejde J, Brosjö O, Rydholm A, Isaksson M, Mandahl N, Mertens F. The chimeric *FUS/CREB3L2* gene is specific for low-grade fibromyxoid sarcoma. *Genes Chromosomes Cancer*. 2004;**40**:218-28.

Panagopoulos I, Möller E, Collin A, Mertens F. The *POU5F1P1* pseudogene encodes a putative protein similar to POU5F1 isoform 1. *Oncol Rep*. 2008;**20**:1029-33.

Pérez-Losada J, Pintado B, Gutiérrez-Adán A, Flores T, Bañares-González B, del Campo JC, Martín-Martín JF, Battaner E, Sánchez-García I. The chimeric FUS/TLS-CHOP fusion protein specifically induces liposarcomas in transgenic mice. *Oncogene*. 2000a;**19**:2413-22.

Pérez-Losada J, Sánchez-Martin M, Rodríguez-García MA, Pérez-Mancera PA, Pintado B, Flores T, Battaner E, Sánchez-García I. Liposarcoma initiated by FUS/TLS-CHOP: the FUS/TLS domain plays a critical role in the pathogenesis of liposarcoma. *Oncogene*. 2000b;**19**:6015-22.

Pérez-Mancera PA, Pérez-Losada J, Sánchez-Martín M, Rodriguez-García MA, Flores T, Battaner E, Gutiérrez-Adán A, Pintado B, Sánchez-García I. Expression of the FUS domain restores liposarcoma development in CHOP transgenic mice. *Oncogene*. 2002;**21**:1679-84.

Pérez-Mancera PA, Bermejo-Rodríguez C, Sánchez-Martín M, Abollo-Jiménez F, Pintado B, Sánchez-García I. *FUS-DDIT3* prevents the development of adipocytic precursors in liposarcoma by repressing PPARgamma and C/EBPalpha and activating eIF4E. *PLoS One*. 2008;**3**:e2569.

Plougastel B, Zucman J, Peter M, Thomas G, Delattre O. Genomic structure of the *EWS* gene and its relationship to EWSR1, a site of tumor-associated chromosome translocation. *Genomics*. 1993;**18**:609-15.

Powers CA, Mathur M, Raaka BM, Ron D, Samuels HH. TLS (translocated-in-liposarcoma) is a high-affinity interactor for steroid, thyroid hormone, and retinoid receptors. *Mol Endocrinol.* 1998;**12**:4-18.

Prasad DD, Ouchida M, Lee L, Rao VN, Reddy ES. TLS/FUS fusion domain of TLS/FUS-erg chimeric protein resulting from the t(16;21) chromosomal translocation in human myeloid leukemia functions as a transcriptional activation domain. *Oncogene.* 1994;**9**:3717-29.

Prestridge DS. Predicting Pol II promoter sequences using transcription factor binding sites. *J Mol Biol.* 1995;**249**:923-32.

Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* 1995;**23**:4878-84.

Rabbitts TH, Forster A, Larson R, Nathan P. Fusion of the dominant negative transcription regulator *CHOP* with a novel gene *FUS* by translocation t(12;16) in malignant liposarcoma. *Nat Genet.* 1993;**4**:175-80.

Ramakrishnan R, Fujimura Y, Zou JP, Liu F, Lee L, Rao VN, Reddy ES. Role of protein-protein interactions in the antiapoptotic function of EWS-Fli-1. *Oncogene.* 2004;**23**:7087-94.

Riggi N, Cironi L, Provero P, Suvà ML, Kaloulis K, Garcia-Echeverria C, Hoffmann F, Trumpp A, Stamenkovic I. Development of Ewing´s sarcoma from primary bone marrow-derived mesenchymal progenitor cells. *Cancer Res.* 2005;**65**:11459-68.

Riggi N, Cironi L, Provero P, Suvà ML, Stehle JC, Baumer K, Guillou L, Stamenkovic I. Expression of the FUS-CHOP fusion protein in primary mesenchymal progenitor cells gives rise to a model of myxoid liposarcoma. *Cancer Res.* 2006;**66**:7016-23.

Riggi N, Cironi L, Suvà ML, Stamenkovic I. Sarcomas: genetics, signalling, and cellular origins. Part 1: The fellowship of TET. *J Pathol.* 2007;**213**:4-20.

Riggi N, Suvà ML, Suvà D, Cironi L, Provero P, Tercier S, Joseph JM, Stehle JC, Baumer K, Kindler V, Stamenkovic I. EWS-FLI-1 expression triggers a

Ewing's sarcoma initiation program in primary human mesenchymal stem cells. *Cancer Res.* 2008;**68**:2176-85.

Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P. Transcriptional regulation of *Nanog* by OCT4 and SOX2. *J Biol Chem.* 2005;**280**:24731-7.

Rossow KL, Janknecht R. The Ewing´s sarcoma gene product functions as a transcriptional activator. *Cancer Res.* 2001;**61**:2690-5.

Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature.* 1973;**243**:290-3.

Rozenberg JM, Shlyakhtenko A, Glass K, Rishi V, Myakishev MV, FitzGerald PC, Vinson C. All and only CpG containing sequences are enriched in promoters abundantly bound by RNA polymerase II in multiple tissues. *BMC Genomics.* 2008;**9**:67.

Sanchez G, Bittencourt D, Laud K, Barbier J, Delattre O, Auboeuf D, Dutertre M. Alteration of cyclin D1 transcript elongation by a mutated transcription factor up-regulates the oncogenic D1b splice isoform in cancer. *Proc Natl Acad Sci U S A.* 2008;**105**:6004-9.

Sánchez-García I, Rabbitts TH. Transcriptional activation by TAL1 and FUS-CHOP proteins expressed in acute malignancies as a result of chromosomal abnormalities. *Proc Natl Acad Sci U S A.* 1994;**91**:7869-73.

Savola S, Klami A, Tripathi A, Niini T, Serra M, Picci P, Kaski S, Zambelli D, Scotlandi K, Knuutila S. Combined use of expression and CGH arrays pinpoints novel candidate genes in Ewing sarcoma family of tumors. *BMC Cancer.* 2009;**9**:17.

Shing DC, McMullan DJ, Roberts P, Smith K, Chin SF, Nicholson J, Tillman RM, Ramani P, Cullinane C, Coleman N. *FUS/ERG* gene fusions in Ewing´s tumors. *Cancer Res.* 2003;**63**:4568-76.

Storlazzi CT, Mertens F, Nascimento A, Isaksson M, Wejde J, Brosjö O, Mandahl N, Panagopoulos I. Fusion of the *FUS* and *BBF2H7* genes in low grade fibromyxoid sarcoma. *Hum Mol Genet.* 2003;**12**:2349-58.

Tan AY, Manley JL. The TET family of proteins: Functions and roles in disease. *J Mol Cell Biol.* 2009;**1**:82-92.

Thomas GR, Latchman DS. The pro-oncoprotein EWS (Ewing´s Sarcoma protein) interacts with the Brn-3a POU transcription factor and inhibits its ability to activate transcription. *Cancer Biol Ther.* 2002;**1**:428-32.

Tirado Y, Williams MD, Hanna EY, Kaye FJ, Batsakis JG, El-Naggar AK. *CRTC1/MAML2* fusion transcript in high grade mucoepidermoid carcinomas of salivary and thyroid glands and Warthin´s tumors: implications for histogenesis and biologic behavior. *Genes Chromosomes Cancer.* 2007;**46**:708-15.

Trinklein ND, Aldred SF, Hartman SJ, Schroeder DI, Otillar RP, Myers RM. An abundance of bidirectional promoters in the human genome. *Genome Res.* 2004;**14**:62-6.

Tschoep K, Kohlmann A, Schlemmer M, Haferlach T, Issels RD. Gene expression profiling in sarcomas. *Crit Rev Oncol Hematol.* 2007;**63**:111-24.

Turc-Carel C, Aurias A, Mugneret F, Lizard S, Sidaner I, Volk C, Thiery JP, Olschwang S, Philip I, Berger MP, Philip T, Lenoir GM, Mazabraud A. Chromosomes in Ewing´s sarcoma. I. An evaluation of 85 cases and remarkable consistency of t(11;22)(q24;q12). *Cancer Genet Cytogenet.* 1988;**32**:229-38.

Uranishi H, Tetsuka T, Yamashita M, Asamitsu K, Shimizu M, Itoh M, Okamoto T. Involvement of the pro-oncoprotein TLS (translocated in liposarcoma) in nuclear factor-*κ*B p65-mediated transcription as a coactivator. *J Biol Chem.* 2001;**276**:13395-401.

Vance C, Rogelj B, Hortobágyi T, De Vos KJ, Nishimura AL, Sreedharan J, Hu X, Smith B, Ruddy D, Wright P, Ganesalingam J, Williams KL, Tripathi V, Al-Saraj S, Al-Chalabi A, Leigh PN, Blair IP, Nicholson G, de Belleroche J, Gallo JM, Miller CC, Shaw CE. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science.* 2009;**323**:1208-11.

Veerla S, Rignér M, Höglund M. Genome-wide transcription factor binding site/promoter databases for the analysis of gene sets and co-occurrence of transcription factor binding motifs. *BMC Genomics*. 2010;**11**:145.

Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med*. 2004;**10**:789-99.

Weiss S, Goldblum J (eds). Fibrosarcoma, fibromyxoid type. In: *Enzinger and Weiss´s Soft tissue Tumors*. 5th edition. Mosby Elsevier Inc.: Philadelphia, 2008;316-25.

Wingender E. The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Brief Bioinform*. 2008;**9**:326-32.

Winnes M, Mölne L, Suurküla M, Andrén Y, Persson F, Enlund F, Stenman G. Frequent fusion of the *CRTC1* and *MAML2* genes in clear cell variants of cutaneous hidradenomas. *Genes Chromosomes Cancer*. 2007;**46**:559-63.

Yamaguchi S, Yamazaki Y, Ishikawa Y, Kawaguchi N, Mukai H, Nakamura T. *EWSR1* is fused to *POU5F1* in a bone tumor with translocation t(6;22)(p21;q12). *Genes Chromosomes Cancer*. 2005;**43**:217-22.

Yang L, Chansky HA, Hickstein DD. EWS.Fli-1 fusion protein interacts with hyperphosphorylated RNA polymerase II and interferes with serine-arginine protein-mediated RNA splicing. *J Biol Chem*. 2000a;**275**:37612-8.

Yang L, Embree LJ, Hickstein DD. TLS-ERG leukemia fusion protein inhibits RNA splicing mediated by serine-arginine proteins. *Mol Cell Biol*. 2000b;**20**:3345-54.

Zhu J, He F, Hu S, Yu J. On the nature of human housekeeping genes. *Trends Genet*. 2008;**24**:481-4.

Zinszner H, Albalat R, Ron D. A novel effector domain from the RNA-binding protein TLS or EWS is required for oncogenic transformation by CHOP. *Genes Dev*. 1994;**8**:2513-26.

Zinszner H, Sok J, Immanuel D, Yin Y, Ron D. TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *J Cell Sci*. 1997;**110**:1741-50.

Zou J, Ichikawa H, Blackburn ML, Hu HM, Zielinska-Kwiatkowska A, Mei Q, Roth GJ, Chansky HA, Yang L. The oncogenic TLS-ERG fusion protein exerts different effects in hematopoietic cells and fibroblasts. *Mol Cell Biol.* 2005;**25**:6235-46.