DEPARTMENT'S RATIFICATION

This is to ascertain that the Department has no objection to the candidate's options regarding access to the Library thesis copy. If so, please sign below and return the completed form to the **Monash Graduate School Research Services, Building 3D, Clayton Campus.**

Supervisor's signature: .....[redacted]..... Date: 15/12/00

(Please print name).....MELANIE PRITCHARD

H24/3046

MONASH UNIVERSITY
THESIS ACCEPTED IN SATISFACTION OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
ON..... 5 October 2001 .....
.....[redacted]....
for Sec. Research Graduate School Committee

Characterization and Functional Study of a Novel Epithelial-Specific ETS Transcription Factor - ELF5

A thesis submitted for the degree of
DOCTOR OF PHILOSOPHY (MED)

by

Jiong Zhou
B.Sc. (HONS)

Centre for Functional Genomics and Human Disease
Monash Institute of Reproduction and Development
Monash University
Australia

June 2001

Table of Contents

List of Figures

Figure 7.10 Mammary gland development is defective in pregnant and lactating *Elf5*$^{+/-}$ females (second pregnancy) ff116

List of Tables

Abstract

ETS (E26 transformation specific or E-twenty-six specific) factors have important developmental roles and many ETS factors have also been implicated in the control of cellular proliferation and tumorigenesis. Most ETS factors identified to date appear to function predominantly in hematopoietic lineages. However, a subclass of ETS factors that are epithelial-specific has emerged suggesting novel epithelial functions for the ETS transcription factor family.

This thesis describes the identification of a novel *ETS* family member, *ELF5,* whose expression is highly restricted to tissues rich in epithelial cells. *In vitro* studies have demonstrated that Elf5 can function through its ETS domain as a transcriptional activator. The human *ELF5* gene is localized to chromosome 11p13-15, a region that frequently undergoes loss of heterozygosity (LOH) in many types of cancer, including those of breast, kidney, lung and prostate. *ELF5* expression appears to be lost in many cancer cell lines, a subset of which were found to have lost an allele or have rearrangement of the *ELF5* gene.

In order to elucidate the biological functions of Elf5, we have generated an *Elf5*$^{-/-}$ mouse model. Knockout experiments have shown that Elf5 has crucial cellular proliferation and differentiation functions during early mouse embryogenesis. This function may be related to the implantation of the mouse embryo, where the loss of function leads to embryonic lethality around E3.5-7.5. Gene targeting experiments also demonstrated an important role for Elf5 in the proliferation and differentiation of mouse mammary alveolar epithelial cells during pregnancy and lactation. Mammary gland development during these stages is extremely sensitive to the level of Elf5 expression. The loss of one functional allele leads to complete developmental arrest of the mammary gland. In addition, this study has identified possible functional relationships of ELF5 with a hormone signaling pathway (PRL/PRLR), a cell cycle regulator (CYCLIN D1), and a breast cancer suppressor protein (BRCA1).

The work presented in this thesis may contribute towards the understanding of the functions of ETS transcription factors in mammalian development, and possibly in human disease.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University or other Institution and, to the best of my knowledge, contains no material previously published or written by another person, except where due reference is made in the text. Where the work in this thesis is part of joint research, the relative contribution of the respective person is acknowledged in the text.

Jiong Zhou

Acknowledgements

First of all, I would like to express my sincerest gratitude to my first supervisor, Professor Ismail Kola, for his confidence in my abilities over the years. His encouragement and support has been the major driving force in accomplishing the results described in this thesis. His involvement and contributions in this project has been tremendous and greatly appreciated.

I must thank my supervisors Associate Professor Paul Hertzog and Dr. Melanie Pritchard for their great support, effort and contributions to this project. The accomplishment of the results described in this thesis will not be possible without their continual support. Dr. Martin Tymms and Dr. Ross Thomas must also be thanked for their roles as co-supervisors of this project, and their direct involvement in producing some of the results. To Dr. Trevor Wilson, Ms. Silva Zarvasek, Dr. Dakang Xu, Dr. Annie Ng, Ms. Sue Tsao, Ms. Irene Tellis, Ms. Carol Wang, Ms. Erika Lapinskas, Dr. Sika Ristevski, and Dr. Ernst Wolvetang, I would like to thank them for their involvement and contributions to this project. I thank Dr. Daniel Lazner and Dr. Mireille Lahoud who have unselfishly provided me with valuable reagents. I wish to thank all the members of the Centre for Functional Genomics and Human Disease for creating a stimulating scientific environment. In addition, I would like to thank my language teachers Ms. Antonietta Giudice, Ms. Elaine Sanij, and Ms. Laureane Mittaz for teaching me Italian, Arabic, and French, and for their continued friendship and support.

I must say a big-big **THANK YOU!!!** to Dr. Melanie Pritchard and Dr. Josephine Tkalcevic for their great patience, constructive input and excellent discussions in proof-reading of this thesis.

Special thanks to mum and dad for their support both emotionally and financially in my entire student career. Finally, a special **THANK YOU!!!** to my beautiful and loving wife, Jenny Weng Hua Lu, who has stood by me every step of the way and deserves all the credits. I would like to dedicate this thesis to my family.

Chapter 1

The ETS family of transcription factors

1.1 General introduction

All multicellular organisms undergo complex patterns of development. Development is a balanced combination of several processes, including cell proliferation, cell death, growth, morphogenesis, and cellular differentiation. In eukaryotes, control of gene expression is extremely complex - cells must respond to changes in their environment by turning on and off appropriate sets of genes. The organization of cells into tissues that have different functions may also require that a gene respond in one way to a given signal in one tissue and another way in a different tissue. In addition, throughout the many cell divisions that transform the fertilized egg into the mature adult, certain cells must become committed to a particular form and function. For development to proceed in an orderly, organized manner, many different genes must be activated or inactivated, and specific regulatory patterns have to be established.

The tightly controlled gene expression of eukaryotic cells involves a specialized group of proteins called transcription factors. Transcription factors function via their activation domains, their repression domains or their DNA binding domains and may interact either directly with the general transcription machinery or through co-activators or co-repressors (Wasylyk *et al.*, 1992; Graves and Petersen, 1998). The ETS family is a class of transcription factors that plays important roles in transcriptional regulation of gene expression.

1.2 The ETS family of transcription factors

The first ETS family member was originally discovered in the avian transforming retrovirus, E26. This virus encodes a tripartite protein ($p135^{gag\text{-}myb\text{-}ets}$) that fused a v-*ets* (E26 transformation specific or E-twenty-six specific) sequence with the v-*gag* gene and the v-*myb* oncogene (Leprince *et al.*, 1983; Nunn *et al.*, 1983). The E26 virus is capable of causing both myeloblastosis, through v-myb, and erythroblastosis through v-ets, although the v-ets part of the fusion protein also contributes to the development of myeloblastosis (Golay *et al.*, 1988). It has been demonstrated that a retrovirus containing the avian *gag-myb-ets* sequences induces erythroid and myeloid leukemia in nude mice (NFS/N strain) (Yuan *et al.*, 1989). *In vitro*, E26 can transform uncommitted bipotent hematopoietic precursor stem cells (Moscovici *et al.*, 1983), as well as erythroid and myelomonocytic precursor cells (Radke *et al.*, 1982; Moscovici *et al.*, 1983; Nunn and Hunter, 1989).

ETS proteins are found in human, mouse, chicken, frog, sea urchin and fruit fly (Dittmer and Nordheim, 1998; Graves and Petersen, 1998). Although there are already more than 50 family members identified in species ranging from sea urchin to human, this family is still growing in number. However, *ets* genes have not been identified in the budding yeast *Saccharomyces cerevisiae*, and it is also not known if ets proteins are present in plants, fungi, or any protozoan (Graves and Petersen, 1998). Since *Ets* genes have been identified in both vertebrates and invertebrates, it is thought that duplication of genes occurred during early metazoan evolution, and that the vast number of *ets* genes, discovered in the above-mentioned species, arose from a single ancestral gene (Degnan *et al.*, 1993; Shenk and Steele, 1993; Graves and Petersen, 1998).

1.2.1 The ETS domain is highly conserved among the ETS family members

The most noticeable feature of the ETS protein family is the presence of a highly conserved DNA-binding domain of approximately 85 amino acids, termed the 'ETS domain' (Watson *et al.*, 1988; Karim *et al.*, 1990; Gutman and Wasylyk, 1991; Macleod *et al.*, 1992; Seth *et al.*, 1992; Wasylyk *et al.*, 1993; Graves and Petersen, 1998). This domain recognizes a purine-rich core sequence, GGAA/T, in a sequence-specific manner, in the promoters and enhancers of various target genes (Macleod *et al.*, 1992; Janknecht and Nordheim, 1993; Wasylyk *et al.*, 1993; Werner *et al.*, 1995; Kodandapani *et al.*, 1996). The amino acid

residue positions important both for the correct protein folding and for DNA contact are highly conserved among the ETS family members. Studies of the secondary structure of a number of ETS proteins, such as hFLI1, mEts1, hETS1 and hPU.1, revealed high degree of conservation of amino acid residues within the α helices and β strands of the ETS domain (Donaldson *et al.*, 1994; Liang *et al.*, 1994; Werner *et al.*, 1995; Donaldson *et al.*, 1996; Kodandapani., 1996). Most interestingly, highly conserved amino acid residues in the secondary structure elements of PU.1 are found to contact DNA in the PU.1-DNA complex (Kodandapani *et al.*, 1996; Pio *et al.*, 1996).

Phylogenetic analysis using the amino acid sequence of the ETS domains from 49 *ETS* genes suggests that the ETS family can be grouped into 9 subfamilies, namely ELF, ELG, ELK, ERF, ERG, ETS, PEA3, SPI, and YAN. Since a number of subfamilies contain *ETS* genes from both invertebrates and vertebrates, it is suggested that *ETS* family genes are duplicated before the evolution of veterbrates (Graves and Petersen, 1998).

1.2.1.1 The secondary and tertiary structures of the ETS domain

The sequence-specific DNA-binding activities of the ETS domain has been demonstrated by deletion analysis, however, other domains in the ETS proteins may affect the interaction between the ETS domain and its target DNA motif (Laudet *et al.*, 1999). Determination of the tertiary structures of the ETS domain for a number of ETS proteins has greatly facilitated the understanding of the interaction between the ETS protein and its DNA target sites.

The secondary structures of the ETS domain for murine Ets1 and human FLI1 were initially established by NMR spectroscopy (Donaldson *et al.*, 1994; Liang *et al.*, 1994). Both of the ETS domains contain three α-helices (H) and four β-strands (S) arranged in the order H1-S1-S2-H2-H3-S3-S4. Interestingly, similar secondary structures are also observed in the DNA-binding domains of a number of proteins that share common tertiary structures (Brennan, 1993). This includes the *E. coli* catabolite activator protein (CAP) (Schultz *et al.*, 1991), *E. coli* biston repressor (birR) (Wilson *et al.*, 1992), hepatocyte nuclear factor (HFN-3γ) (Clark *et al.*, 1993), yeast and *Drosophila* heat shock factors (HSF) (Harrison *et al.*, 1994; Vuister *et al.*, 1994; Vuister *et al.*, 1994), and the globular domain of histone (H5) (Ramakrishnan *et al.*, 1993). These DNA-binding domains are found to form "winged

helix-turn-helix" (wHTH or "winged helix") tertiary structures, and the α-helices and β-strands within the wHTH are physically involved in DNA contacts as determined by crystallographic and NMR studies (Clark *et al.*, 1993). Ogata *et al.* (1992) showed that the conserved tryptophans in the all-helical *myb* domain contribute to the formation of a similar tertiary structure. This is also true of the wHTH domain of the yeast HSF protein in which the tryptophans are either preserved or conservatively substituted in two α-helices and one β-strand (Harrison *et al.*, 1994). Besides the similar ordering of the α-helices and the β-sheet in the ETS domain of murine Ets1, three tryptophans are also found in two helices (H1 and H2) and a β-sheet (S1) in this domain, which suggests that the ETS domain will form a wHTH-like tertiary structure (Donaldson *et al.*, 1994).

High-resolution crystallographic analysis of the PU.1-DNA complex (Kodandapani *et al.*, 1996; Pio *et al.*, 1996) demonstrated that the helix H3 and the turn in front of H3 mediate base-specific DNA contacts to the GGA core. In particular, the two invariant arginine residues in H3 directly contact the two guanine residues. The other α-helices and β-strands are involved in DNA backbone contacts. A study of the Ets1-DNA structure showed a similar binding complex (Werner *et al.*, 1995, 1997 a & b).

1.2.1.2 Sequence-specific DNA-binding properties of the ETS domain

Although the amino acid residues in helix H3 are either highly preserved or conservatively substituted among the ETS family members, it is noteworthy that residues in other α-helices and β-strands, involved in DNA contacts, appear to be more variable among the ETS proteins. Therefore, ETS proteins may selectively recognize different ETS-binding sites that are flanked by different DNA bases, resulting in sequence-specific DNA binding (Graves and Petersen, 1998).

Mouse Ets1 protein has the highest affinity toward an ETS binding site with a 5′-ACCGGAACG-3′ sequence (Nye *et al.*, 1992), which apparently also has high affinity for a number of other ETS proteins (Gunther and Graves, 1994; Graves *et al.*, 1996). The nucleotide flanking the 3′ end of the GGA core is found to be important in determining which ETS factor(s) will bind to this site. An adenine (A) at this nucleotide position is invariant for the recognition by mElf1 (John *et al.*, 1996), de74 (Urness and Thummel,

1990), and mFli1 (Mao *et al.*, 1994). However, mEts1 (Nye *et al.*, 1992), mSap1 (Shore and Sharroks, 1995), mGabpα, mEr71, mEr81 (Brown and McKnight, 1992), mElk1 (Shore and Sharroks, 1995), mPu.1, mSpiB (Ray-Gallet *et al.*, 1995), and hELF3 (Tymms *et al.*, 1997) can bind to a GGA core with either a 3′ flanking thymine (T) or an adenine (A). When a single amino acid within the cEts1 ETS domain is changed to match the comparable residue in Elf1, the mutant Ets1 protein can no longer recognize a GGAT core (Bosselut *et al.*, 1993). Interestingly, a novel ETS transcription factor, hPDEF, can only bind with high affinity to the GGAT-containing oligonucleotide derived from the PSA promoter, but not to the GGAA-containing oligonucleotide (Oettgen *et al.*, 2000). In addition, the nucleotides flanking the GGAA/T core also confer DNA-binding specificity. For example, PU.1 and Ets1 do not recognize and bind to the same ETS-binding site in the immunoglobulin heavy-chain enhancer. Instead, each protein binds to its own preferred ETS-binding site and both Ets proteins are required for the regulation of *IgH* gene expression (Nelsen *et al.*, 1993). In another case, an ETS-binding site within the enhancer of the *IL2 receptor* is found to downregulate IL2 receptor expression, and this site is recognized by Elf1 but not Ets1 (John *et al.*, 1996). Most interestingly is the ability of PU.1 to recognize a 5′-AGA-3′ DNA core element within the immunoglobulin J chain enhancer while other ETS proteins prefer a 5′-GGA-3′ DNA core (Shin and Koshland, 1993). Finally, the number of nucleotides of the ETS-binding site is also important in sequence-specific recognition of the ETS proteins. For example, PU.1 and SpiB prefer sites composed of more than 15 bp. In contrast, the Fli1 recognition site is only 13 bp long whereas other family members prefer 9-11 bp long sites. Therefore, variations of the nucleotide residues within or flanking the core DNA-binding sites will influence the binding affinity and specificity of ETS proteins (Graves and Petersen, 1998).

1.2.2 A highly conserved PNT domain is present in a subset of the ETS family members

Besides the ETS domain, another highly conserved domain is also found among some of the *ETS* factors, within the ETS, ELG, ERG, and YAN groups (Graves and Petersen, 1998). This domain was originally called the "A" domain (Papas *et al.*, 1989) but later termed the pointed domain (or PNT domain) after the *Drosophila ets* family member, *pointed* (*pnt*) (Klambt, 1993). The presence of the PNT domain in only a subset of the ETS family members makes its evolutionary origin unclear. It is not known if it was lost along the

evolutionary process, or if certain ETS family members gained this domain after early gene duplication of the family. Interestingly, all *ETS* subfamily genes encoding the PNT domain have their ETS domain located 3′ of the PNT domain. In contrast, the genes of the ELK group have their ETS domains located at the 5′ end and do not have a PNT domain. It has been proposed that a chromosomal rearrangement of the gene locus encoding the founder of the ELK group resulted in the loss of the PNT domain, however, the actual mechanisms involving the loss of this domain in some ETS protein are unknown (Graves and Petersen, 1998).

A NMR spectroscopy study of the secondary structure of the mouse Ets1 PNT domain revealed five α-helices within the domain, four of which contain highly conserved amino acid residues. The fifth α-helix, extending toward a MAP kinase site in the amino-terminal region of Ets1, is less conserved among the ETS proteins (Graves and Petersen, 1998; Slupsky *et al.*, 1998).

The PNT domain is a subclass of the SAM, SPM or SEP domain family found in a variety of different classes of proteins including Polycomb proteins and Eph receptors (Schultz *et al.*, 1997; Kyba and Brock, 1998; Stapleton *et al.*, 1999). A crystallization study of the SAM domain in the Eph receptor, and protein-protein interaction studies of Polycomb proteins, have shown the direct involvement of the SAM domain in both homo- and heterodimerization of proteins. In addition, it was suggested that the SAM domain may be involved in the formation of higher order protein complexes (Schultz *et al.*, 1997; Stapleton *et al.*, 1999).

The PNT domain of TEL has already been shown to act as a self-association domain, however, it is not known if other PNT domain-containing ETS proteins can also associate with themselves (Carroll *et al.*, 1996; Golub *et al.*, 1996; McLean *et al.*, 1996; Jousset *et al.*, 1997; Graves and Petersen, 1998). It has also been suggested that the PNT domain of some ETS proteins may associate with other regulatory proteins to form a functional protein complex (Graves and Petersen, 1998).

1.2.3 Protein-protein interactions involving the ETS family members

ETS transcription factors can function through the ETS binding site in the promoters and enhancers of various downstream target genes. In addition to the DNA-binding activity, the ETS domain is also involved in protein-protein interactions with numerous other transcription factors. Therefore, protein-protein interaction is thought to be another critical factor in regulating the activity of ETS proteins. For instance, c-Jun interacts directly with the carboxyl-terminal portion of ERM, which contains the ETS domain, resulting in increased transcriptional activity of ERM *in vitro* (Nakae *et al.*, 1995). Several other ETS proteins, including Ets1, Elf1, PU.1 and Fli1, have been shown to interact with the basic domain of the Jun transcription factor through the ETS domain *in vitro* and in activated human T cells. This interaction induces the transcriptional activity of enhancer elements containing adjacent ETS and AP1 binding sites (Bassuk and Leiden, 1995). ETS1 can also associate with PAX-5 via the ETS domain to form a ternary complex (Fitzsimmons *et al.*, 1996; Wheat *et al.*, 1999).

The PNT domain is also involved in protein-protein interactions. In addition to the well documented observations of TEL self-association via the PNT domain (Carroll *et al.*, 1996; Golub *et al.*, 1996; McLean *et al.*, 1996; Jousset *et al.*, 1997), other ETS family members are also found to associate with themselves and with other ETS proteins. For example, the ETS transcription factor human TEL2 is capable of self-association, presumably via its PNT domain, and it can also heterodimerize with TEL via the PNT domain of TEL (Potter *et al.*, 2000). ERG and its isoforms are also known to form homodimers or heterodimers with other ETS proteins, and these protein-protein interactions are mediated via the PNT and ETS domains (Carrere *et al.*, 1998). ETS2 can associate with ERG via both the ETS domain and the PNT domain (Basuyaux *et al.*, 1997). The PNT and ETS domains of TEL are both required to interact with FLI1, and this protein-protein interaction can inactivate a FLI1-specific promoter (Kwiatkowski *et al.*, 1998).

In vitro studies often show that ETS proteins on their own are weak transcriptional regulators of downstream target gene promoters. However, the presence of protein-protein interacting partner(s) will result in either enhanced or repressed transcriptional activity of the ETS proteins (Sharrocks *et al.*, 1997). It has been shown that Ets1 has enhanced binding affinity toward a low affinity ETS binding site in the human *T-cell receptor (TCR) β-chain*

promoter when Ets1 functions cooperatively with CREB and AML1 (Halle *et al.*, 1997). In another case, the DNA-binding affinity of an ETS family member, GABPα, is dramatically improved when it is complexed with its partner protein, GABPβ. Two GABPα/β heterodimers can bind to two tandem repeats of the ETS-binding site in a downstream target gene promoter and become transcriptionally active (Thompson *et al.*, 1991; Batchelor *et al.*, 1998). The same GABP complex is found to bind to the two closely positioned tandem repeats of the ETS-binding site in the *TN-C* promoter and hence stimulates the promoter activity (Shirasaki *et al.*, 1999). Other ETS proteins also need the presence of other transcription factors to maximize the transcriptional activity of their respective target genes through an ETS binding site and adjacent DNA sequences (Wasylyk *et al.*, 1989; Gutman and Wasylyk, 1990; Dalton and Treisman, 1992; Graves and Petersen, 1998). For example, a factor(s) binding to the PEA3 site of the polyoma virus enhancer and collagenase promoter, acts synergistically with AP1 to express maximal transcription activity (Wasylyk *et al.*, 1989). Ets1 and Ets2 also cooperate synergistically with c-Fos and c-Jun in the regulation of the polyoma virus enhancer (Wasylyk *et al.*, 1990). Ets1 and GABP can activate the *HTLV1 long terminal repeat* (Gegonne *et al.*, 1993) and the *CD18* promoter (Rosmarin *et al.*, 1998) respectively, in synergy with Sp1, and ELF3 activates the *TGM3* and *profillagrin* promoters by cooperating with Sp1 and AP1, respectively (Andreoli *et al.*, 1997). The interaction between PU.1 and Pip (PU.1-interacting protein or NF-EM5) facilitates the activation of *immunoglobulin light-chain* gene enhancers (Judde and Max, 1992; Pongubala *et al.*, 1992). A conserved B box shared within the ELK group of ETS proteins is involved in the formation of a ternary complex with the serum response factor (SRF) on the serum response element (SRE) in the *c-fos* promoter (Shaw *et al.*, 1989; Hill *et al.*, 1993). Shore and Sharrocks (1994) showed that the B box can interact with SRF, and this interaction can be abolished by introducing a single amino acid substitution in the B box (Ling *et al.*, 1997).

Protein-protein interaction partnerships are often restricted to a single ETS protein or group, but in other cases partnerships may be available to multiple ETS proteins. For example, protein complexes of Pax-5 and either Fli1, Ets1, or GABPα are identified in B cell nuclear extracts. Mutation analyses demonstrated that the ETS domain found in Net, Elk1, Ets1, and Fli1, but not SAP1, is capable of forming a ternary complex with Pax-5 (Fitzsimmons *et al.*, 1996). Ets1, ERG, PU.1 and ERM are also shown to be capable of interacting with AP1

transcription factors in the regulation of target gene expression (Wasylyk *et al.*, 1990; Nakae *et al.*, 1995; Buttice *et al.*, 1996; Logan *et al.*, 1996; Zhang *et al.*, 1999).

ETS transcription factors are also found in higher order complexes. For example, the optimal activation of the *HIV-1* enhancer requires the presence of a multiprotein complex including LEF1, Ets1, TFE3, NFκB and Sp1 (Sheridan *et al.* 1995). In another case, a ternary complex of Ets1 and CBF assembled on the *T cell receptor α subunit* enhancer is found to be stabilized in the presence of ATF2/CREB and LEF1 (Giese *et al.*, 1995; Mayall *et al.*, 1997). Thomas *et al.*, (1997) have also shown that the activation of the *GM-CSF* promoter in activated T cells by either or both NFκB and AP1 is significantly increased in the presence of ETS1. The maximal transcription activity requires the presence of all three factors, and therefore suggests that these proteins may be part of a higher order transcriptional complex.

Numerous partnerships between ETS and other proteins have been identified, and these protein partnerships appear to enhance the promoter specificity of ETS transcription factors. However, further studies are required to elucidate the protein interaction domains used by the ETS proteins.

1.2.4 Transcriptional regulation by the ETS family members

Transcription factors are important for the regulation of gene expression. Activation and repression are two basic properties of transcription factors that may require the presence of cofactors in certain cases to achieve optimal transcription activity for the target gene promoter. Although all ETS proteins contain a highly conserved ETS domain, sequence divergence outside of the ETS domain among the family members is evident and these differences may be critical in the selection of cofactors. Studies have shown that there are both transcriptional activators and repressors in the ETS family of transcription factors.

1.2.4.1 Transcriptional activators of the ETS family

Most ETS factors have the ability to transactivate from binding sites in promoters and enhancers of various downstream target genes. Although ETS proteins do not have a

common transactivation domain (TAD) among the family members, the transcriptional activation domains in ETS proteins do contain features of a classical transactivation domain, including a richness in acidic amino acid residues and certain other residues such as proline and glutamine. PU.1 contains two TADs, one rich in acidic residues, the other rich in glutamine residues (Hagemeier *et al.*, 1993; Shin and Koshland, 1993; Kominato *et al.*, 1995; Klemsz and Maki, 1996). In addition, it has been shown that closely related family members share sequence similarity in their transactivation domains. For example, TADs rich in acidic residues are found to be conserved within the PEA3 group (ERM, ER81 and PEA3) (Laget *et al.*, 1996) and within the ETS group (Ets1 and Ets2) (Schneikert *et al.*, 1992). Similarly, the TAD is found to be conserved within the ELK group (Elk1 and SAP1) (Bhattacharya *et al.*, 1993; Marais *et al.*, 1993; Janknecht *et al.*, 1994) and within the ERG group (Erg and Fli1) (Rao *et al.*, 1993; Siddique *et al.*, 1993). Therefore, the sequence divergence observed amongst the TADs could explain the functional specificities of each individual ETS protein (Graves and Petersen, 1998).

Unlike most of the ETS transcriptional activators, GABPα does not appear to have a TAD. Instead, GABPα needs to heterodimerize with its protein-partner GABPβ to become transcriptionally active (Gugneja *et al.*, 1995; Gugneja *et al.*, 1996).

1.2.4.2 Transcriptional repressors of the ETS family

A number of ETS factors have been shown to possess repressor-like activity. For example, ERF, Net and yan can act as repressors of gene expression (Lai and Rubin, 1992; Giovane *et al.*, 1994; Sgouras *et al.*, 1995; Day *et al.*, 1998). Promoters negatively regulated by the binding of ETS family members, such as cEts1, mElk3 and mPE1 (or PEP1) have also been reported (Goldberg *et al.*, 1994; Chen and Boxer, 1995; Nozaki *et al.*, 1996; Bidder *et al.*, 2000).

Several studies demonstrated that certain ETS proteins can repress target gene expression by either competing with the transactivator for the ETS binding site in the target gene promoter or by recruiting a corepressor to the target gene promoter for maximal repressing activity (Graves and Petersen, 1998). For example, *Drosophila* yan (Lai and Rubin, 1992; O'Neill *et al.*, 1994; Rebay and Rubin, 1995), *Drosophila* e74b (Fletcher and Thummel, 1995) and ERF (Sgouras *et al.*, 1995) have been shown to repress target gene promoters activated by

other ETS proteins. Also, TEL has been shown to interact with mSin3A (Fenrick *et al.*, 1999) and Fli1 (Kwiatkowski *et al.*, 1998) to repress target gene promoter activity.

Interestingly, ERG and ETS1 have been shown to act as both transactivators and repressors under different circumstances (Suzuki *et al.*, 1995; Buttice *et al.*, 1996). Similarly, ELF3 (ESX/ESE1/JEN/ERT) is shown to not only transactivate the *WAP* (Thomas *et al.*, 2000), *PSMA*, *SPRR2A*, *Endo A* (*keratin 8*) (Oettgen *et al.*, 1997; Kas *et al.*, 2000), *TGM3* (Andreoli *et al.*, 1997), *TGFβRII* (Choi *et al.*, 1998) and *ERBB2/HER2/neu* (Chang *et al.*, 1997) promoters, but to repress the *K4* (*keratin 4*) (Brembeck *et al.*, 2000) and *PSA* promoters (Kas *et al.*, 2000). Like TEL, the repressing activity of ELF3 on the *K4* promoter is not due to a direct binding of ELF3 to an ETS binding site but rather to the recuitment of other corepressors to the promoter (Brembeck *et al.*, 2000).

1.2.5 Signal transduction involving the ETS family members

Phosphorylation plays an important role in signal transduction pathways and can affect the transcriptional activity of certain transcription factors. A number of ETS factors are shown to be targets of MAP kinase signaling pathways. For example, the transcriptional activity of Elk1 and SAP1 can be increased due to the phosphorylation of the C box region by the MAP kinase ERK under growth factor stimulation (Hipskind *et al.*, 1994; Treisman, 1994; Hill and Treisman, 1995; Janknecht *et al.*, 1995; Price *et al.*, 1995; Whitmarsh *et al.*, 1995). In addition, the DNA binding activity of Elk1 is also shown to be increased due to the phosphorylation of Elk1 by the MAP kinase ERK (Kortenjann *et al.*, 1994; Gille *et al.*, 1995; Gille *et al.*, 1996; Shore *et al.*, 1996). Similarly, a *C. elegans* ETS protein, Lin1, of the ELK group is also regulated by the MAP kinase-signaling pathway (Beitel *et al.*, 1995). However, the same extracellular signal has little or no effect upon another TCF protein, SAP2/NET (Giovane *et al.*, 1994; Lopez *et al.*, 1994; Price *et al.*, 1995; Maira *et al.*, 1996), which suggests a different regulatory pathway or mechanism is involved in the regulation of SAP2/NET.

The MAP kinase-signaling pathway also regulates other ETS proteins. For example, *Drosophila* pnt p2 can promote *Drosophila* photoreceptor differentiation as a result of Ras/MAPK signaling (O'Neill *et al.*, 1994), however, this differentiation can be repressed by another *Drosophila* ets protein, yan, that is also regulated by the same pathway (Rebay

and Rubin, 1995; Treier *et al.*, 1995). Furthermore, the transcriptional activities of Ets1 and Ets2 are also shown to be activated by the Ras/MAP kinase pathway (Rabault *et al.*, 1996; Yang *et al.*, 1996). This pathway also appears to be involved in the regulation of other ETS proteins such as ERM, GABPα, ER81, and ERF (Sgouras *et al.*, 1995; Flory *et al.*, 1996; Janknecht, 1996; Janknecht *et al.*, 1996).

1.2.6 Autoinhibition of the ETS family members

A number of ETS proteins, such as SAP1 (Dalton and Treisman, 1992), Elk1 (Rao and Reddy, 1992; Janknecht *et al.*, 1994), Net (Giovane *et al.*, 1994; Lopez *et al.*, 1994; Price *et al.*, 1995; Maira *et al.*, 1996), ERM (Laget *et al.*, 1996), Ets1 and Ets2 (Hagman and Grosschedl, 1992; Lim *et al.*, 1992; Wasylyk *et al.*, 1992; Fisher *et al.*, 1994; Petersen *et al.*, 1995; Jonsen *et al.*, 1996), are shown to contain sequences that inhibit the DNA binding activities of their ETS domains. In the case of ERM, Elk1, Ets1 and Ets2, interactions between the TAD and ETS domains within each individual protein down-regulate their transcriptional activities (Schneikert *et al.*, 1992; Chumakov *et al.*, 1993; Janknecht *et al.*, 1994; Laget *et al.*, 1996). For some other ETS proteins, such as SAP1a (Dalton and Treisman, 1992), Elk1 (Rao and Reddy, 1992; Janknecht *et al.*, 1994) and ERM (Nakae *et al.*, 1995), protein-protein interactions are required to release this DNA-binding inhibition. In addition, phosphorylation is also found to influence the autoinhibition of certain ETS family members. For example, phosphorylation of Elk1 promotes ternary complex formation with SRF, and this leads to increased transcriptional activity of the c-fos promoter (Gille *et al.*, 1995). In contrast, Rabault and Ghysdael (1994) demonstrated that the DNA binding activity of Ets1 is inhibited by the phosphorylation of the inhibitory region of Ets1.

1.2.7 The expression of the ETS family members

The ETS transcription factors are a large family implicated in the control of cellular proliferation, differentiation and tumorigenesis. Mammalian ETS proteins display distinct but overlapping patterns of expression. Expression of Ets2, Gabpα, ER81, ERM, Tel and NERF are widely distributed, but expression of Spi1/PU.1, SpiB, Elf1, Fli1, and TEL2 are predominantly found in the hematopoietic lineages. In the case of PEA3, ER71 and Elk1, expression is limited to testis and brain. Interestingly, expression of certain *ETS* genes (Ets1

and Erg) is regulated during development (Crepieux *et al.*, 1994; Oettgen *et al.*, 1996; Wang *et al.*, 1997; Potter *et al.*, 2000).

In spite of their wide distribution, many *ETS* genes are shown to play important roles in regulating the development and differentiation of the hematopoietic lineages. This is strongly supported by *in vivo* transgenic and gene-targeting experiments in the mouse (Sharrocks *et al.*, 1997; Graves and Petersen, 1998; Dittmer and Nordheim, 1998). In addition, ETS DNA binding sites are located in the promoters and enhancers of various genes involved in the regulation of the hematopoietic system (Ghysdael and Boureux, 1997; Bassuk and Leiden, 1997). However, little is known about the function of ETS proteins in epithelial cells.

ELF3 (*ESX/ESE1/JEN/ERT*) was identified by several groups as the first member of the *ETS* transcription factor gene family whose expression is strictly limited to cells of epithelial origin (Andreoli *et al.*, 1997; Chang *et al.*, 1997; Choi *et al.*, 1998; Oettgen *et al.*, 1997; Tymms *et al.*, 1997), and its expression is found in almost all epithelial cells examined. *EHF/ESE3* (Bochert *et al.*, 1998; Kleinbaum *et al.*, 1999; Kas *et al.*, 2000) and *PDEF/PSE* (Oettgen *et al.*, 2000; Yamada *et al.*, 2000) together with *ELF5*, which will be described in this thesis, are three novel epithelial-specific *ETS* genes that may play important roles in epithelial cell proliferation, differentiation and tumorigenesis. *EHF/ESE3* is expressed mainly in epithelial cells of glandular organs such as the salivary gland, pancreas, prostate and mammary gland. The expression of *PDEF/PSE* is restricted to the digestive organs and the accessory glands of reproductive organs.

1.2.8 Biological function of the ETS family members

Cellular proliferation, differentiation and tumorigenesis are accompanied by changes in gene expression, and these changes are regulated by transcription factors. Transcription factors are capable of regulating target gene expression by direct DNA-protein interaction, therefore, their DNA-binding affinity and specificity are critical factors in their regulating mechanism. However, the availability of cofactors may either promote or interfere with this type of control mechanism via protein-protein interaction and the presence of an autoinhibitory sequence will add an additional layer of regulation. Therefore, disruption of any of these regulatory steps could lead to deregulation of target gene expression. ETS

proteins often display overlapping expression patterns, and many ETS factors have been shown to have high affinity toward a common DNA-binding site in the promoter of a target gene *in vitro*. However, ETS proteins exhibit distinct biological functions as demonstrated by *in vivo* studies (Graves and Petersen, 1998).

1.2.8.1 Retroviral insertion involving the ETS family members

Early *Ets* gene function analyses involved studies of retroviral infections in mice, and it was found that upregulation of different *Ets* genes due to retroviral insertion resulted in erythroleukemias. In 95% of Friend erythroid tumors induced by spleen focus-forming virus (SFFV), retroviruses were found integrated upstream of the *Spi1* gene resulting in its aberrant expression (Morea-Gachelin *et al.*, 1988; Paul *et al.* 1989; Ben-David *et al.*, 1991). In 75% of erythroleukemias resulting from a Friend murine leukemia virus (F-MuLV) infection, a retrovirus was found upstream of the *Fli1* gene causing overexpression of the *Fli1* gene (Ben-David *et al.*, 1990; Ben-David *et al.*, 1991). However, one interesting exception involves the Moloney murine leukemia virus (MoMuLV). This virus integrates upstream of the rat *c-Ets1* locus (Bear *et al.*, 1989), but its integration does not significantly upregulate the expression of the *c-Ets1* gene (Bellacosta *et al.*, 1994).

1.2.8.2 Overexpression transgenics involving the ETS family members

Overexpression of Ets2, Pu.1, and Fli1 in mice have resulted in distinctive phenotypes. *Ets2* transgenic mice exhibited skeletal/bone abnormalities (Sumarsono *et al.*, 1996) which are similar to those seen in human Down syndrome (trisomy 21) patients. Interestingly, the human *ETS2* gene is located on chromosome 21 and therefore it is proposed that overexpression of *ETS2* may contribute to the skeletal/bone abnormalities in those patients (Sumarsono *et al.*, 1996). Pu.1 is required in lymphoid and myeloid cell differentiation and overexpression of Pu.1 has caused erythroleukemia in mice (Moreau-Gachelin *et al.*, 1996). Study of Fli1-overexpressing mice has revealed a severe renal disease due to a defective immune system that resulted in renal failure. These mice also had an increased proliferation rate of the splenic B cells, and these cells survived longer than their normal counterpart in the presence of mitogen. These results suggested that the normal B cell function in the Fli1-overexpressing mice was disrupted (Zhang *et al.*, 1995).

Overexpression studies of *Ets* genes in mice suggest distinctive roles of different Ets proteins, however, care should be taken in interpreting the results from these kinds of experiments. It is debatable whether the true function of a gene can be reflected by the outcome of such studies, because it is difficult to precisely mimic the aberrant expression state of a particular gene in disease (Graves and Petersen, 1998).

1.2.8.3 *Drosophila* and *Caenorhabditis elegans* ets family members

There are seven known *Drosophila ets* genes, *e74*, *pnt*, *yan*, *elg*, *ets3*, *ets4 and ets6*, and one *Caenorhabditis elegans ets* gene, *lin1* (Graves and Petersen, 1998). Based upon the sequence similarity in the ETS domain, *Drosophila* e74, pnt, yan, elg and ets6 fall into the ELF, ETS, YAN, ELG and ERG groups, respectively (Graves and Petersen, 1998). *Drosophila* ets4 is most closely related to hPDEF/mPse (Kas *et al.*, 2000; Oettgen *et al.*, 2000; Yamada *et al.*, 2000), and *C. elegans* lin1 falls into the ELK group. *Drosophila* ets3 does not have a complete ETS domain sequence, and therefore it is difficult to predict which group it belongs to.

Mutation analyses have demonstrated important biological functions of four *Drosophila ets* genes (*e74*, *pnt*, *elg*, and *yan*). *Drosophila* e74 is shown to be essential for ecdysone-induced pupation and metamorphosis (Burtis *et al.*, 1990; Fletcher and Thummel, 1995; Fletcher *et al.*, 1995). *Drosophila* pnt was shown to be necessary in the development of midline glial cells and photoreceptor R7 cells (Brunner *et al.*, 1994; Krasnow, 1996). Mutations in the *pnt* gene lead to alterations in the central nervous system due to abnormal interactions between glial cells and migrating midline neuronal cells, resulting in larval lethality (Klambt, 1993). Specific disruption of the *pnt p2* transcript additionally results in abnormal wings in heterozygous flies (Scholz *et al.*, 1993). Studies of loss-of-function mutations in the *Drosophila elg* gene demonstrated that elg is required for both egg chamber patterning and embryonic patterning (Schulz *et al.*, 1993; Gajewski and Schulz, 1995). Mutations in the *yan* gene disrupted proper differentiation of the photoreceptor cells (Lai and Rubin, 1992). The *C. elegans ets* gene, *lin1*, was shown to be important in the vulva development due to its negative regulating effect (Beitel *et al.*, 1995).

1.2.8.4 Gene disruptions of the mouse *Ets* family members

Gene targeting technology is a useful tool for studying the biological functions of the Ets transcription factors in mice. *Pu.1*, *Ets1*, *Fli1*, *Pea3*, *Tel*, *SpiB*, *Ets2*, *Er81*, and *Elf3* mutant mice displayed distinct phenotypes, indicating the functional specificity of these murine *Ets* genes.

1.2.8.4.1 Gene disruptions in the SPI group

Spi1/Pu.1 is highly expressed in B lymphocytic, granulocytic and monocytic cells (Klemsz *et al.*, 1990; Galson *et al.*, 1993; Hromas *et al.*, 1993; Chen *et al.*, 1995). Scott *et al.* (1994a) demonstrated that *Pu.1*$^{-/-}$ embryos died at day 16.5 of gestation due to the absence of normal progenitors for B and T lymphocytes, monocytes, or granulocytes. In addition, Pu.1 was shown to be necessary for the terminal differentiation of the myeloid lineage (Olson *et al.*, 1995). Interestingly, McKercher *et al.* (1996) produced live *Pu.1*$^{-/-}$ mice with no mature B and T lymphocytes, macrophages, and neutrophils, and the mice died as a result. However, antibiotic treatments have restored normal development of a small number of T cells in these mice. Therefore, these experiments demonstrate that Pu.1 is only required for the differentiation of the myeloid and lymphoid cell lineages.

SpiB is most closely related to Pu.1, and its expression is lymphoid-specific in the mouse embryo and adult (Chen et al., 1995; Su *et al.*, 1996). *SpiB*$^{-/-}$ mice have mature and functional T cells, however, these mice have functional defects in mature B cells. In addition, *SpiB*$^{-/-}$ mice fail to respond to T-dependent antigen stimulation. Therefore, SpiB is necessary for the B cells to respond and survive upon antigen stimulation mediated by the B cell receptor (Su *et al.*, 1997).

1.2.8.4.2 Gene disruptions in the ETS group

Ets1 is expressed in thymus, astrocytes, heart, lung, gut and spleen (Cher. 1985; Ghysdael *et al.*, 1986; Wernert *et al.*, 1992; Kola *et al.*, 1993; Maroulakou *et al.*, 1994; Fleischman *et al.*, 1995). An *Ets1*$^{-/-}$*RAG2*$^{-/-}$ mouse model was generated for studying the Ets1 function in T and B cells (Bories *et al.*, 1995; Muthusamy *et al.*, 1995). The studies show that Ets1 is necessary to maintain a healthy population of T and B cells, and loss-of-function mutation in

Ets1 gene leads to abnormal T cell apoptosis and B cell differentiation. In addition, Barton *et al.* (1998) showed that the development and function of mature natural killer (NK) cells in *Ets1*$^{-/-}$ mice was severely disrupted. The *Ets1*$^{-/-}$ NK cells fail to respond to exogenous IL2, IL12, or IL15 *in vitro*. A subsequent study showed that Ets1 is also required for the development of NK1.1^{+} T (NK T) cells (Walunas *et al.*, 2000). Taken together, these results identify Ets1 as an essential regulator for the development of both the NK and NK T cell lineages.

Ets2 is expressed ubiquitously in all adult organs and in many cell types in developing mouse embryos (Kola *et al.*, 1993; Maroulakou *et al.*, 1994). The Ets2 protein is most closely related to Ets1. Disruption of the DNA binding domain of Ets2 protein results in embryonic lethality around 8.5 days of embryonic development (Yamamoto *et al.*, 1998). *Ets2*$^{-/-}$ embryos are retarded due to growth arrest in the ectoplacental cone which may reflect a deficiency in matrix metalloproteinase-9 (MMP9, gelatinase B) expression and consequent failure to break down the extracellular matrix surrounding the embryo preventing the trophoblastic cells of the embryo from migrating. Therefore, Ets2 is important for placental development. Although *Ets2*$^{-/-}$ embryos were rescued by aggregation with tetraploid mouse embryos, *Ets2*$^{-/-}$ adult mice still displayed abnormal hair and hair follicle development. In addition, Ets2 is found to be necessary in regulating growth factor responsive gene expression of MMP13 (collagenase3) and MMP3 (stromelysin1) in fibroblasts (Yamamoto *et al.*, 1998). Interestingly, female *Ets2*$^{+/-}$ mice with mammary tumors induced by the *PyMT* oncogene have their tumor size halved in comparison with the wildtype controls. In addition, the tumor tissues from *PyMT/Ets2*$^{+/-}$ mice were more differentiated than tumor tissues from *PyMT/Ets2*$^{+/+}$ mice. It appears that Ets2 is not playing a role in the tumor growth but rather a role in regulating its progression (Neznanov *et al.*, 1999).

1.2.8.4.3 Gene disruptions in the PEA3 group

The formation of direct and indirect connections between proprioceptive sensory and motor neurons (MNs) during development is important for the control of motor behavior (Brown, 1981; Zelena, 1994). Lin *et al.* (1998) suggested that ER81 and PEA3 may be important in establishing these connections since both genes are expressed during this process but turned off soon after the formation of sensory-motor connections. *Er81*$^{-/-}$ mice have a defect in motor coordination due to disrupted connections between proprioceptive afferents and MNs,

and these mice died at 3-5 weeks of age. These results suggest that Er81 is essential for the establishment of a functional sensory-motor connection in the developing spinal cord (Arber *et al.*, 2000).

PEA3 is closely related to ER81, however, *Pea3*$^{-/-}$ mice are viable with no distinct phenotype. Interestingly, *Pea3*$^{-/-}$ male mice fail to impregnate female mice despite normal spermatogenesis, spermiogenesis, sperm maturation and sexual behavior. In addition, *Pea3*$^{-/-}$ sperm are shown to fertilize eggs normally *in vitro*. It is suggested that Pea3 may play a role in sensory and motor neurons related to reproduction in mice (Laing *et al.*, 2000). Therefore, Pea3 appears to have a distinct function from Er81.

1.2.8.4.4 Gene disruption of *Tel*

The *TEL* (*ETV6*) gene is known to be involved in chromosomal translocations in both myeloid and lymphoid leukemias in human (Golub *et al.*, 1994; Buijs *et al.*, 1995; Golub *et al.*, 1995; Papadopoulos *et al.*, 1995; Romana *et al.*, 1995a; Romana *et al.*, 1995b; Shurtleff *et al.*, 1995; Golub *et al.*, 1996; McLean *et al.*, 1996; Raynaud *et al.*, 1996a; Raynaud *et al.*, 1996b). Loss-of-function mutation of the *Tel* gene results in death of *Tel*$^{-/-}$ embryos between 10.5-11.5 days of embryonic development. The vitelline vessels of the *Tel*$^{-/-}$ yolk sacs are either missing or regressed. In addition, neural cell apoptosis is evident in the *Tel*$^{-/-}$ embryo proper. Therefore, Tel is necessary for angiogenesis in the yolk sac, and the survival of mesenchymal and neural cells in the embryo (Wang *et al.*, 1997).

1.2.8.4.5 Gene disruption of *Fli1*

The human *FLI1* gene is also involved in chromosomal translocations resulting in Ewing's sarcoma (Delattre *et al.*, 1992; May *et al.*, 1993). *Fli1* is preferentially expressed in hematopoietic cells, and its expression is seen in thymus, ovary, bone marrow, spleen and heart (Ben-David *et al.*, 1991; Watson *et al.*, 1992; Klemsz *et al.*, 1993). The disruption of the gene encoding Fli1 was performed independently by two laboratories. Melet *et al.* (1996) disrupted exon II of the *Fli1* gene in mice, however, a truncated but functional Fli1 protein was produced. This homozygous mutation in mice resulted in a reduction of thymocytes and hence a smaller thymus. In contrast, the *Fli1*$^{-/-}$ mice created by Spyropoulos *et al.* (2000) and Hart *et al.* (2000) resulted in embryonic lethality. *Fli1*$^{-/-}$ embryos exhibited

severe hemorrhage at 11.0 days of embryonic development that led to death at embryonic days 12.5. This lethality was found to be associated with the disruption of the columnar neuroepithelium and its adjacent basement membrane in *Fli1*$^{-/-}$ embryos where the hemorrhaging sites were located. Interestingly, these tissues in *Fli1*$^{+/+}$ embryos are found to express high levels of *Fli1* mRNA, suggesting a specific role of Fli1 in maintaining the integrity of these tissues (Spyropoulos *et al.*, 2000). In addition, megakaryocyte progenitors from *Fli1*$^{-/-}$ fetal livers fail to differentiate (Hart *et al.*, 2000; Spyropoulos *et al.*, 2000). Hart *et al.* (2000) show that the *FLI1* gene is deleted in 14 human patients with Jacobsen or Paris-Trousseau Syndrome which display dysmegakaryopoiesis. Therefore, a similar defect to that observed in *Fli1*$^{-/-}$ mouse embryos also supports a Fli1-specific function in megakaryopoiesis. Taken together, these results demonstrate that Fli1 plays an important role in the regulation of hematopoiesis and hemostasis (Hart *et al.*, 2000; Spyropoulos *et al.*, 2000).

1.2.8.4.6 Gene disruption of *Elf3*

ELF3 (ESX/ESE1/JEN/ERT) is a novel epithelial-specific ETS transcription factor, and its expression is found in almost all epithelial cells examined (Andreoli *et al.*, 1997; Chang *et al.*, 1997; Choi *et al.*, 1998; Oettgen *et al.*, 1997; Tymms *et al.*, 1997). Approximately 40% of *Elf3*$^{-/-}$ embryos die around embryonic day 11.5-12.5. In addition, histological and immunohistochemical analyses of *Elf3*$^{-/-}$ mice revealed epithelial cell defects in the small intestine, uterus, prostate and seminal vesicles. The results demonstrate that Elf3 plays an important role during embryonic development, and in ductal morphogenesis and differentiation of epithelial cells of the small intestine, uterus, prostate and seminal vesicles in the adult (Ms. A.Y.N. Ng, personal communication).

1.2.8.5 ETS transcription factors in human cancer

Chromosomal translocation can change specific gene expression patterns and the resulting protein structure. Chromosomal translocations involving ETS family members are frequently associated with human leukemias and other human cancers. Fusion proteins that contain domains of ETS proteins have been identified in certain types of leukemias and in Ewing tumors (Dittmer and Nordheim, 1998; Graves and Petersen, 1998).

In Ewing tumors, a fusion protein is produced with an ETS domain fused to the N-terminal portion of the EWS oncoprotein. In the majority of the Ewing's sarcomas, the ETS domain is derived from FLI1 (Delattre *et al.*, 1992; May *et al.*, 1993) or ERG (Sorensen *et al.*, 1994), representing 85% and 10% of the cases, respectively. ER81 (or ETV1), E1AF (or ETV4) and FEV account for the remaining cases (Jeon *et al.*, 1995; Kaneko *et al.*, 1996; Peter *et al.*, 1997). It has been shown that EWS-Fli1 and EWS-Erg fusion proteins have both a sequence-specific DNA binding domain and a transcriptional activation domain, derived from the Ets protein and the EWS protein, respectively (Bailly *et al.*, 1994). These fusion proteins are regarded as oncogenic proteins due to their capabilities of transforming NIH3T3 cells (May *et al.*, 1993) and inhibiting apoptosis in NIH3T3 and Ewing's sarcoma cells (Yi *et al.*, 1997). In addition, the anchorage-independent growth and tumorigenicity of Ewing's sarcoma cells in mice can be reversed by introducing antisense *EWS-Fli1 or EWS-Erg* (Ouchida *et al.*, 1995; Kovar *et al.*, 1996; Tanaka *et al.*, 1997). These results strongly support a causative role for EWS-ETS fusion proteins in the development of Ewing tumors (Dittmer and Nordheim, 1998).

In certain types of human leukemia, fusion proteins are found involving the ETS transcription factors, TEL (or ETV6) or ERG. Chromosomal translocations involving TEL and PDGFRβ result in chronic myelomonocytic leukemia (Golub *et al.*, 1994), while translocations between TEL and AML1 lead to acute lymphoblastic leukemia (Golub *et al.*, 1995). TEL is also fused to the Janus tyrosine kinase 2 (JAK2) resulting in pre-B acute lymphoid leukemia and atypical chronic myelogenous leukemia (Peeters *et al.*, 1997). In addition, chromosomal translocations between TEL and MN1 can produce fusion proteins that are associated with myeloid and lymphoid leukemias (Buijs *et al.*, 1995). In the case of ERG, a fusion protein between ERG and FUS/TLS is found in acute myeloid leukemia (Ichikawa *et al.*, 1994; Panagopoulos *et al.*, 1994).

ETS proteins, such as PEA3, ER81, ERM (Monte *et al.*, 1995; Chen *et al.*, 1996) and ELF3 (Chang *et al.*, 1997; Tymms *et al.*, 1997), are overexpressed in certain breast cancer cells. Interestingly, ER81 is only overexpressed in breast cancers that lack estrogen and progesterone receptors (Monte *et al.*, 1995), whereas *ELF3* overexpression is only found in breast cancers that are *ERBB2/HER2/neu* positive (Chang *et al.*, 1997). It was also shown that recombinant PEA3 (Benz *et al.*, 1997) and ELF3 (Chang *et al.*, 1997) can transactivate

the *ERBB2/HER2/neu* promoter *in vitro*. Therefore, certain types of breast tumors may be the result of overexpression of certain ETS family members (Dittmer and Nordheim, 1998).

1.3 Perspective of this thesis

Growth factors, hormones, adhesion molecules, and extracellular matrix molecules are all required for regulating epithelial cell proliferation and differentiation. In particular, cross talk between adjacent epithelial and mesenchymal cells is extremely important for developing glandular organs. However, the cellular signals and transcriptional regulators that are involved in these processes are largely unknown, and very few of the known transcription factors are restricted to epithelial cells (Kas *et al.*, 2000).

The most common solid tumors in humans are carcinomas that arise from the transformation of epithelial cells (Birchmeier *et al.*, 1995). Transformed breast epithelial cells, for example, have been shown to express the ETS family members GABPα, PEA3, ELF1, ETS1 and ELK1 (Scott *et al.*, 1994b; Delannoy-Courdent *et al.*, 1996), but expression of these ETS family members is not restricted to epithelial cells. ELF3 (Tymms *et al.*, 1997) is a novel ETS transcription factor that plays a role in mammary gland remodeling, in the early differentiation of the ductal epithelium and during terminal differentiation of the epidermis (Andreoli *et al.*, 1997; Oettgen *et al.*, 1997; Neve *et al.*, 1998). In addition, ELF3 is found to activate endogenous TGFβRII expression in Hs578t human breast cancer cells, raising the possibility that ELF3 is a critical regulator of TGFβ signaling in certain cancer cells (Choi *et al.*, 1998; Chang *et al.*, 2000). Interestingly, Brembeck *et al.*, (2000) found that ELF3 is a transcriptional repressor in regulating genes involved in early squamous epithelial cell differentiation. Furthermore, knockout experiments in mice have demonstrated that ELF3 plays an important role in embryonic development and in the ductal morphogenesis and differentiation of the epithelial cells of the small intestine, uterus, prostate and seminal vesicles in the adult (Ms. A.Y.N. Ng, personal communication). However, experimental evidence strongly supports the existence of other as yet unidentified ETS transcription factor(s) that have important functions in epithelial cells (Fujimura *et al.*, 1994; Welte *et al.*, 1994; Bradford *et al.*, 1995; Roberson *et al.*, 1995; Fischer *et al.*, 1996; Gambarotta *et al.*, 1996; Lee *et al.*, 1996; Olsen *et al.*, 1997; Zhang *et al.*, 1997b; Choi *et al.*, 1998; Xing *et al.*, 2000). Therefore, it is important to identify any additional ETS factors that are

epithelial-specific. The aim of this study was to clone and characterize a novel *ETS* family member, *ELF5/Elf5*, and to determine its biological function(s) using gene targeting technology.

Chapter 2

Materials and Methods

2.1 General laboratory reagents

Unless indicated otherwise, all laboratory reagents were analytic grade and supplied by Ajax Chemicals, Australia; BDH Chemicals, UK; Bio-Rad Laboratories, USA; Boehringer Mannheim, Germany; Progen Industries, USA; Promega Corporation, USA; or Sigma Aldrich, USA.

DNA restriction endonucleases and modifying enzymes were supplied by Boehringer Mannheim, Germany; GIBCO BRL, UK; New England Biolabs, USA; or Promega Corporation, USA.

Solutions were prepared using Millipore MilliQ H2O (Millipore Corporation, USA) using the protocols of Sambrook *et al.*, (1989) unless stated otherwise (see Appendix A for composition of general laboratory reagents).

2.2 DNA manipulation

2.2.1 Restriction endonuclease digestion

DNA was digested using appropriate restriction endonucleases under conditions recommended by the manufacturer (Promega, Madison, WI, USA).

2.2.2 DNA analysis by gel electrophoresis

Agarose gels of varying concentrations were used for analyses such as sizing, separating and purifying of DNA fragments. Ethidium bromide (0.2 μg/ml) was added to the gel for the visualization of DNA under ultraviolet (UV) light (302 nm).

2.2.3 Purification of DNA fragments from agarose gel

Specific DNA fragments were excised from agarose gels under UV illumination (302 nm), and purified using the QIAEX II Gel Extraction Kit, according to the manufacturer's instructions (Qiagen, Chatsworth, CA, USA).

2.2.4 Library screening

The murine *Elf5* cDNA was isolated from an adult lung cDNA library in λZAPII (Stratagene, La Jolla, CA, USA) following hybridization screening with a cDNA probe containing the ETS domain region of human *ELF3*. A human *ELF5* cDNA fragment was isolated from a SUPERSCRIPT™ human lung cDNA library (GIBCO BRL, Paisley, UK) following hybridization screening with a cDNA probe containing the coding sequence of mouse *Elf5*. Murine *Elf5* genomic clones were isolated from a 129SvJ λFIXII genomic library (Stratagene) using m*Elf5* cDNA probes.

The XL-1 blue MRF' (Stratagene), Y1090r⁻ (GIBCO BRL) and XL-1 blue MRA (P2) (Stratagene) strains of *E. coli* were used as hosts for bacteriophages λgt11 (GIBCO BRL), λZAPII (Stratagene) and λFIXII (Stratagene), respectively. The host bacteria were incubated in 50 mls of Luria-Bertani (LB) medium, supplemented with 0.2% w/v maltose and 10 mM $MgSO_4$, at 37°C with shaking at 200 rpm (4300; Innova, New Brunswick Scientific, New Jersey, USA) overnight. The bacteria were harvested and resuspended in 10 mM $MgSO_4$ to OD_{600} = 0.5. Bacteriophage (1 x 10^6 plaque forming units) (pfu) were incubated with host bacteria at 37°C for 15 minutes to allow the bacteria to adsorb to plating cells before being plated on 150-mm LB agar plates (LB medium containing 1.5% w/v agar) with 0.7% w/v top agar. The plates were inverted and incubated at 37°C for approximately 6-8 hours for plaque formation.

The bacteriophage DNA was transferred and fixed onto duplicate Hybond-N nylon membrane discs (Amersham Pharmacia, Buckinghamshire, UK). The bacteriophage DNA was denatured and neutralized in 1.5 M NaCl/0.5 M NaOH for two minutes (once) and 1.5 M NaCl/0.5 M Tris(hydroxymethyl)methylamine-HCl (Tris-HCl, pH 8.0) for five minutes (twice), respectively. The DNA was crosslinked onto the membranes by exposure to UV for five minutes.

The membranes, containing the bacteriophage DNA, were prehybridized in 50 mls of a solution containing 5× SSC, 5× Denhardt's, 0.5% w/v SDS and 200 μg/ml denatured herring sperm DNA (Boehringer Mannheim, Mannheim, Germany) in a rotating oven (XTRON HI2002, Bartelt Instruments, Heidelberg West, VIC, Australia) at 50°C for 2 hours. The membranes were then hybridized with the appropriate denatured cDNA probe (see section 2.2.12), [α-^{32}P] dCTP-labelled (10 mCi/ml; Amersham Pharmacia), in 50 mls of fresh solution as above at 50°C overnight. Membranes were washed to a final stringency of 0.2× SSC/0.1% w/v SDS at 50°C. Positive clones were identified by autoradiography and purified by two more rounds of screening as before. A single plaque from each positive clone was isolated using a sterile glass Pasteur pipette, and stored in 1 ml of SM buffer containing a drop of chloroform at 4°C.

The cDNA inserts carried by the λZAPII vectors were finally recovered in pBluescript plasmids by *in vivo* excision using the ExAssist/SOLR system, according to the manufacturer's instructions (Stratagene).

2.2.5 Rapid amplification of cDNA ends (RACE)

Additional cDNA sequence was obtained by reverse transcriptase-polymerase chain reaction (RT-PCR) using a Marathon cDNA synthesis kit and RACE (Rapid Amplification of cDNA Ends), according to the manufacturer's instructions (Clontech, Palo Alto, CA, USA). Adaptor-ligated cDNA libraries were constructed using 1 μg of either day 14 murine placental or human placental Poly(A)$^+$ mRNA. RACE PCR amplification was carried out using the adaptor and nested adaptor primers (AP1 and AP2), together with *ELF5/Elf5* gene-specific and nested gene-specific primers. The *ELF5/Elf5* gene-specific PCR products

were cloned into pGEM-T vector (Promega). All cDNA sequences were confirmed by sequencing both strands at least once using the vector-based T7 and SP6 primers.

Murine *Elf5* primers

5′-RACE gene-specific primer 1: 5′-GCCAGTCTTGGTCTCTTCAGCATC-3′
5′-RACE nested gene-specific primer 2: 5′-AGGAGATGCAGTTGGCATCAAGCT-3′
3′-RACE gene-specific primer 1: 5′-AGCCAGTGTTATGGGTGCTG-3′
3′-RACE nested gene-specific primer 2: 5′-ACAGTCACTTGATCCACGGCCAATCC-3′

Human *ELF5* primers

5′-RACE gene-specific primer 1: 5′-GCCTTTCGAATGTCTATTGCAATCTG-3′
5′-RACE nested gene-specific primer 2: 5′-GAGCTTGATGCCTGGAGCAG-3′

2.2.6 DNA ligation

Ligation

Specific DNA fragments were ligated, using T4 DNA ligase, into a linearized vector backbone, that has been treated with calf intestinal alkaline phosphatase under conditions recommended by the manufacturer (Promega) to remove 5′-terminal phosphate moieties. Compatible vector and insert ends were created either by restriction endonucleases or by creating blunt ends with Klenow DNA polymerase, under conditions recommended by the manufacturer (Promega). Varying molar ratios of insert:vector were used to achieve optimal ligation efficiency. The amount of DNA insert to be used was calculated by the following formula:

$$\frac{\text{ng of vector} \times \text{insert size (kb)}}{\text{vector size (kb)}} \times \text{insert:vector molar ratio} = \text{ng of insert}$$

Bacterial transformation and transformant screening

Following ligation, plasmid vectors containing putative DNA inserts were transformed into JM109 competent cells under conditions recommended by the manufacturer (Promega). These cells were plated onto LB agar plates containing isopropylthio-β-D-galactosidase (IPTG, 0.5 mM), 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal, 80 μg/ml) and ampicillin (100 μg/ml), and incubated at 37°C overnight. Blue/white color selection using IPTG/X-Gal (Promega) facilitated the identification of putative recombinant bacterial colonies (white colonies) that contained DNA insert. Following DNA preparation, appropriate restriction endonuclease digestions were used to characterize the transformants.

2.2.7 DNA sequencing

Sequencing

DNA sequencing was performed in-house by the sequencing facility (Wellcome Trust Sequencing Centre, Monash University, Australia) using an ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (with AmpliTaq® DNA Polymerase, FS; PERKIN ELMER, Norwalk, Connecticut, USA) and an ABI 373A DNA sequencer (PERKIN ELMER), according to the manufacturer's instructions.

Sequence analysis

Analysis and alignment of nucleotide and protein sequences was performed using the sequence analysis software package by the Genetics Computer Group (GCG; Altschul *et al.*, 1990, 1997): (1) BLASTN, BLASTP and FASTA - homology searches against the GenBank, EMBL, PIR, PDB and SWISS-PROT nucleotide and protein databases; (2) BESTFIT and PILEUP - alignment of multiple ETS protein sequences; and (3) MOTIFS - scanning of consensus phosphorylation motif sites.

The phylogenetic tree of the ETS domain was produced by maximum likelihood analysis. The alignment of ETS domains was analyzed using the JTT-F substitution model (Jones *et*

al., 1992) and local bootstrap values were estimated for all internal branches, both by using PROTML in Q mode, followed by a second run in R mode (Adachi and Hasegawa, 1996).

2.2.8 Sequence Tagged Site (STS) content mapping

The following sequence-specific primers for the human *ELF5* gene were used for PCR.

Forward primer : 5'-GGGTGGCAGGAAGACAAGCTATGA-3'
Reverse primer : 5'-CCAATTAAGTCCCAGCTTGATGGC-3'

The PCR reactions contained 5 µl of Opti-Primer™ 10× buffer #3 (100 mM Tris-HCl pH 8.3, 35 mM $MgCl_2$, 250 mM KCl), 1 µl of Master Mix 50× buffer (20 mM Tris-HCl pH 8.0, 250 nM EDTA) (Opti-Primer™ PCR Optimization Kit, Stratagene), 50 ng of template DNA, 0.2 µg of each primer, 1 µl of 10 mM dNTPs and 0.25 unit of Taq DNA polymerase (Promega) in a total volume of 50 µl. PCR parameters were an initial denaturation step at 94°C for 1 min, followed by 30 cycles of 94°C (1 min), 65°C (1 min), 72°C (1 min). For the Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre), PCR reactions were performed separately for each of the individual hybrids. The PCR results from the 93 hybrids were submitted to the Radiation Hybrid Mapping server at the Whitehead Institute/MIT Center for Genome Research (http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl). The STS content mapping experiment was performed in duplicate and included PCR reactions with no DNA, total human DNA and total hamster DNA as controls.

2.2.9 Small-scale plasmid preparation

A modified alkaline lysis/polyethyleneglycol (PEG) precipitation method (Tartof and Hobbs, 1987) was used for small-scale plasmid DNA preparation. 1.5 mls of TB (Terrific Broth), supplemented with ampicillin (50 µg/ml), was inoculated with a single bacterial colony containing a required DNA insert, and then incubated at 37°C with shaking overnight. Bacteria were harvested and resuspended in 200 µl of Solution I (1 mM EDTA pH 8.0, 50 mM Tris/HCl pH 8.0). The cell suspension was mixed with 200 µl of freshly prepared Solution II (0.2 M NaOH, 1% w/v SDS) and allowed to lyse at room temperature

for 5 minutes. Bacterial cell debris, genomic DNA and proteins were precipitated by the addition of 200 µl ice-cold Solution III (3 M potassium acetate pH 5.5), and removed by centrifugation at 13,000 rpm (1-15; SIGMA, Osterode am Harz, Germany) for 10 minutes. The supernatant containing the plasmid DNA was incubated with RNase A (50 µg/ml; Boehringer Mannheim) at 37°C for 30 minutes. Plasmid DNA was extracted repeatedly with 400 µl of chloroform until no debris was visible at the interface. An equal volume of isopropanol was added to precipitate the plasmid DNA, which was pelleted by centrifugation at 13,000 rpm (1-15; SIGMA) for 10 minutes. The DNA pellet was washed with 500 µl of 70% v/v ethanol, air-dried and resuspended in 50 µl of sterile MilliQ water.

The plasmid DNA was further purified with a PEG_{6000} precipitation step if the DNA was to be sequenced. The plasmid DNA was re-precipitated with 30 µl of 20% w/v PEG_{6000} (in 2.5 M NaCl) on ice for 1 hour followed by centrifugation at 13,000 rpm (1-15; SIGMA) for 5 minutes at 4°C. The DNA pellet was washed with 70% v/v ethanol, air-dried and resuspended in 20 µl of sterile MilliQ water.

2.2.10 Large-scale plasmid preparation

A modified alkaline lysis/cesium chloride method (Sambrook *et al.*, 1997) was used for large-scale plasmid DNA preparation. 200 mls of TB (Terrific Broth), supplemented with ampicillin (50 µg/ml), was inoculated with a single bacterial colony containing the required DNA insert, and then incubated at 37°C with shaking overnight. Bacteria were harvested and resuspended in 5 mls of Solution I (25 mM Tris/HCl pH 8.0, 10 mM EDTA pH 8.0, 50 mM glucose, stored at 4°C), supplemented with 2 mg/ml lysozyme and 50 µg/ml RNase A. The cell suspension was mixed with 10 mls of freshly prepared Solution II (0.2 M NaOH, 1% w/v SDS) and allowed to lyse on ice for 10 minutes. Bacterial cell debris, genomic DNA and proteins were precipitated by the addition of 7.5 mls ice-cold Solution III (5 M potassium acetate, 11.5% v/v glacial acetic acid) and the mixture was incubated on ice for 20 minutes. The precipitate was removed by centrifugation at 10,000 rpm (J2-21 M/E; Beckman, Fullerton, CA, USA) for 15 minutes at 4°C, and the supernatant was filtered through gauze. The supernatant was extracted with 20 mls of Tris buffer saturated phenol/chloroform (1:1) followed by 20 mls of chloroform. The aqueous phase was separated from extraction solvent by centrifugation at 5,000 rpm (J2-21 M/E; Beckman) for

10 minutes. Plasmid DNA was precipitated by the addition of 11.25 mls of isopropanol and pelleted by centrifugation at 10,000 rpm (J2-21 M/E; Beckman) for 10 minutes. The DNA pellet was washed with 70% v/v ethanol, air-dried and resuspended in 2.5 mls of sterile MilliQ water. Cesium chloride (CsCl, 1 g/ml) and ethidium bromide (0.2 mg/ml) were added to the plasmid solution. The plasmid solution was transferred to a heat sealable TL100 ultracentrifuge tube and centrifuged in a Beckman ultracentrifuge (TL100; Beckman, Fullerton, CA, USA) at 100,000 rpm at 20°C for 16 hours. Approximately 1 ml of plasmid DNA was collected with a syringe and needle (18G). An equal volume of water-saturated butanol was added and removed repeatedly to extract the ethidium bromide. Plasmid DNA was then precipitated by the addition of 10 mls of 80% v/v ethanol and pelleted by centrifugation at 10,000 rpm (J2-21 M/E; Beckman) for 10 minutes. The DNA pellet was washed with 70% v/v ethanol, air-dried and resuspended in 0.5-1 ml of 1× TE (pH 8.0).

2.2.11 Large-scale bacteriophage DNA preparation

A modified method of Sambrook *et al.* (1997) was used for large-scale bacteriophage DNA preparation. 50-100 µl of the bacteriophage suspension (~ 10^5 pfu) and 100 µl of MRA(P2) host-bacteria were mixed and incubated at 37°C for 20 minutes, and then added to 0.7% w/v top agarose before being plated onto a LB agarose plate (150-mm plate, containing 1.5% w/v bottom agarose). The plate was inverted and incubated at 37°C for approximately 6-8 hours until plaques were almost confluent. Bacteriophage were eluted into 10 mls of SM buffer at room temperature with gentle shaking for 5 hours. Bacterial debris was removed by centrifugation at 9,000 rpm (J2-21 M/E; Beckman) for 10 minutes at 4°C. The supernatant was treated with RNase A (1 µg/ml) and DNase I (1 µg/ml) at 37°C for 30 minutes. Bacteriophage particles were precipitated by incubating the supernatant with an equal volume of a solution containing 20% w/v PEG_{6000} and 2 M NaCl (in SM buffer) at 0°C (iced-water) for 1 hour, and then pelleted by centrifugation at 10,000 rpm (J2-21 M/E; Beckman) for 20 minutes at 4°C. The pellet was resuspended in 1 ml of SM buffer. The residual debris was removed by centrifugation at 9,000 rpm (1-15; SIGMA) for 2 minutes at 4°C. The supernatant was incubated with 10 µl of 10% w/v SDS and 10 µl of 0.5 M EDTA (pH 8.0) at 68°C for 15 minutes. The mixture was then extracted once with Tris buffer saturated phenol, once with Tris buffer saturated phenol/chloroform (1:1), and once with chloroform. The aqueous phase was separated from extraction solvent by centrifugation at

5,000 rpm (1-15; SIGMA) for 10 minutes at 4°C. Bacteriophage DNA was precipitated by the addition of an equal volume of isopropanol, and stored at -70°C for 20 minutes. The DNA was pelleted by centrifugation at 13,000 rpm (1-15; SIGMA) for 15 minutes at 4°C. The pellet was washed with 70% ethanol, air-dried and resuspended in 50-100 µl of 1× TE (pH 8.0).

2.2.12 Random primed labelling of double stranded DNA

Double stranded DNA probes were radio-labelled with [α-^{32}P] dCTP (10 mCi/ml; Amersham Pharmacia) using a random primed DNA labelling kit, according to the manufacturer's instructions (Boehringer Mannheim). The labelled probes were purified by column chromatography using a 2 ml column of G-50 Sephadex beads (Amersham Pharmacia) in 1× TE (pH 8.0). The radio-activity of the labelled probe (1 µl) was measured using a Liquid Scintillation Analyzer (1900 TR; Canberra Packard, Mt. Waverley, Victoria, Australia) in 5 mls of Hydrofluor Scintillation fluid (Pational Diagnostics, New Jersey, USA). 2 x 10^6 cpm of probe per ml of hybridization buffer was denatured by boiling for 5 minutes prior to being added to the hybridization mix.

2.2.13 End-labelling of single stranded oligonucleotides

Single stranded oligonucleotides were radio-labelled with [γ-^{32}P] dATP (10 mCi/ml; Amersham Pharmacia). 100 ng of oligonucleotide was incubated with 2 µl of [γ-^{32}P] dATP and 1 µl of T4 polynucleotide kinase (Promega) in 1× kinase buffer (Promega) in a total volume of 10 µl at 37°C for 30 minutes prior to being added to the hybridization mix.

2.2.14 Dot blot analysis

Dot blot analysis was performed to characterize the *mElf5* genomic fragments. 100 ng of bacteriophage DNA was spotted onto the Hybond-N nylon membrane (Amersham Pharmacia). The membrane was then denatured in 1.5 M NaCl/0.5 M NaOH for five minutes (once) and neutralized in 1.5 M NaCl/0.5 M Tris-HCl (pH 8.0) for five minutes (twice). The DNA was crosslinked onto the membrane by exposing to UV for five minutes. The membrane was prehybridized in 3 mls of a solution containing 5× SSC, 5× Denhardt's,

1% w/v SDS and 100 μg/ml denatured herring sperm DNA (Boehringer Mannheim) at 42°C for 1-2 hours, and then hybridized in a rotating oven (Bartelt Instruments) with the appropriate [γ-^{32}P] dATP labelled oligonucleotide in 3 mls of fresh solution (as above) at 42°C, overnight. The membrane was washed twice in 1× SSC/0.1% w/v SDS for 10 minutes at 42°C, and then exposed to a Fuji phosphorimage screen and visualized using a phosphorimage analyzer (FLA-2000; FUJIFILM, Tokyo, Japan).

2.2.15 Genomic DNA isolation

Isolation of genomic DNA from mouse embryonic stem (ES) cells

A modified method of Laird *et al.* (1991) was used for isolating the genomic DNA from mouse ES cells. Cells were treated with 500 μl of lysis buffer (100 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 0.2% w/v SDS, 200 mM NaCl, 100 μg/ml proteinase K) at 55°C for 3 hours. The genomic DNA was then precipitated by the addition of 1 ml of isopropanol. The spooled DNA was washed with 70% v/v ethanol, air-dried (30-60 minutes) and resuspended in 200 μl of 0.25× TE pH 8.0.

Isolation of genomic DNA from mouse tails

Mouse tail (~5 mm) was lysed in 700 μl of tail lysis buffer (1% w/v SDS, 0.1 M NaCl, 0.1 M EDTA, 0.05 M Tris-HCl pH 8.0, 300 μg/ml Proteinase K) at 55°C overnight, and then extracted once with Tris buffer saturated phenol/chloroform (1:1) and once with chloroform. The aqueous phase was separated from the extraction solvent by centrifugation at 13,000rpm (1-15; SIGMA) for 5 minutes, and then treated with 1 μl of 10 mg/ml RNase A for 30 minutes at 37°C. The genomic DNA was precipitated by the addition of 1 ml of 100% v/v ethanol, spooled out and washed with 70% v/v ethanol, air-dried and resuspended in 200 μl of 1× TE pH 8.0.

Isolation of genomic DNA from postimplantation embryos or yolk sacs

Mouse postimplantation embryo or yolk sac was lysed in 100 µl of lysis buffer (0.45% v/v NP-40, 0.45% v/v Tween-20, 100 µg/ml proteinase K) at 55°C for 30-60 minutes, and then boiled for 10 minutes to inactivate the proteinase K.

Isolation of genomic DNA from murine blastocysts

Murine blastocysts were cultured in ES medium for 4 days (see section 2.5.1), and then lysed in 20 µl of lysis buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl, 2 mM $MgCl_2$, 0.45% v/v NP-40, 0.45% v/v Tween-20, 60 µg/ml proteinase K) at 55°C for 1 hour, followed by boiling for 10 minutes to inactivate the proteinase K (Takai *et al.*, 2000). Approximately 8 µl of the lysate was subjected to PCR amplification.

2.2.16 Genotyping

Genotyping by Southern blot analysis

Approximately 10 µg of mouse genomic DNA was used to determine the genotype by Southern blot analysis. The *Hind*III digested genomic DNA was separated on a 0.8% w/v agarose gel, and then transferred to GeneScreen Plus nylon membrane (NEN™ Life Science Products, Boston, MA, USA) using 0.4 M NaOH as transfer and fixation buffer, according to the manufacturer's instructions. Post-blotting, the membrane was air-dried at room temperature for 1 hour. The membrane was prehybridized in 10 mls of a solution containing 10% w/v dextran sulphate, 1% w/v SDS, 1 M NaCl and 100 µg/ml of denatured herring sperm DNA, at 65°C for 2 hours, and then hybridized with denatured [α-^{32}P] dCTP-labelled DNA probe (see section 5.2.6) in 10 mls of freshly prepared solution (as above) at 65°C, overnight. The membrane was washed twice with 2× SSC/0.1% w/v SDS at 65°C for 30 minutes, and then exposed to a Fuji phosphorimage screen and visualized using a phosphorimage analyzer (FLA-2000; FUJIFILM).

Genotyping by Polymerase Chain Reaction (PCR)

100 ng of mouse genomic DNA was used to determine the genotypes of mouse tails and embryos by PCR. The PCR reaction mix contained 1 µl of Master Mix 50× buffer (20 mM Tris-HCl pH 8.0, 250 nM EDTA pH 8.0; Opti-Prime™ PCR Optimization Kit, Stratagene), 5 µl of Opti-Prime 10× buffer #2 (100 mM Tris-HCl pH 8.3, 15 mM $MgCl_2$, 750 mM KCl; Opti-Prime™ PCR Optimization Kit, Stratagene), 1 µl of 10 mM dNTPs, 1.25 units of Taq DNA polymerase (Promega), 10 pmol of each oligonucleotide primer (P_S, P_{AS1}, P_{NEO}; see section 5.2.7), in a 50 µl final volume. The amplification protocol comprised an initial incubation at 94°C for 5 minutes, followed by 45 cycles of 94°C for 1 minute, 50°C for 1 minute, and 72°C for 1 minute (Takai *et al.*, 2000). The PCR reaction was completed with a polymerization cycle of 15 minutes at 72°C. PCR products were analyzed by agarose gel electrophoresis, and then transferred and fixed onto a GeneScreen Plus nylon membrane using 0.4 M NaOH. An internal oligonucleotide (P_{AS2}; see section 6.2.1), labelled with [γ-^{32}P] dATP, was used as a hybridization probe (see section 2.1.14) to determine the PCR product specificity.

P_S : 5′-GCACACCCAGAATTGAAGATTCC-3′

P_{AS1} : 5′-CCTTCACTGCACGTGGACTG-3′

P_{NEO} : 5′-ATTCGCCAATGACAAGACGC-3′

P_{AS2} : 5′-TTGGACAGGCCTGATGGGTC-3′

2.3 RNA manipulation

Solutions associated with RNA manipulations were treated with 0.1% v/v diethylpyrocarbonate (DEPC) at room temperature, overnight, and then autoclaved prior to use.

2.3.1 Total RNA isolation

A modified method of Chomczynski and Sacchi (1987) was used for isolating total RNA from tissues or cells. Frozen tissue or a cell pellet was homogenized in 10 mls of ice-cold solution D (4 M guanidinium thiocyanate, 25 mM sodium citrate pH 7.0, 0.5% w/v sodium

lauryl sarkosinate, 720 μM β-mercaptoethanol) at room temperature for 1 minute. The resulting tissue/cell suspension was mixed with 0.1 volume of 2 M sodium acetate pH 4.0, followed by 1 volume of water-saturated phenol and 0.2 volume of ice-cold chloroform/isoamylalcohol (49:1), and then left on ice for 20 minutes. The aqueous phase was separated from the extraction solvent by centrifugation at 5,000 rpm (J2-21 M/E; Beckman) at 4°C for 20 minutes, and then extracted again with 1 volume of water-saturated phenol and 0.2 volume of chloroform/isoamylalcohol (49:1). The RNA in the aqueous phase was precipitated with an equal volume of isopropanol and stored at -20°C overnight. The RNA was pelleted by centrifugation at 5,000 rpm (J2-21 M/E; Beckman) at 4°C for 15 minutes, and then washed twice with ice-cold 75% v/v ethanol, air-dried and resuspended in 50-100 μl of DEPC-treated MilliQ water. The concentration and quality of the RNA were determined by UV spectrophotometry at OD_{260nm} and OD_{280nm} (Lambda Bio20; PERKIN ELMER).

2.3.2 Poly(A)$^+$ mRNA isolation

Poly(A)$^+$ mRNA was extracted from tissues or cells using oligo-dT cellulose (Boehringer Mannheim). Frozen tissue or a cell pellet was homogenized in 25 mls of RNA extraction buffer (20 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 0.1 M NaCl, 0.5% w/v SDS, 10 mg proteinase K) for 1 minute, and then incubated at 55°C for 30 minutes. After cooling to room temperature, the tissue/cell suspension was mixed with 2 mls of 5 M NaCl and 2.5 mls of oligo-dT cellulose, which was resuspended in RNA binding buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 0.5 M NaCl, 0.1% w/v SDS). The mixture was incubated at room temperature with shaking for 2 hours. The RNA-bound oligo-dT cellulose was washed twice with 10 mls of RNA binding buffer, and once with 10 mls of RNA wash buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 0.15 M NaCl, 0.1% w/v SDS), and retrieved by centrifugation at 3,000 rpm (J2-21 M/E; Beckman) for 3 minutes. poly(A)$^+$ mRNA was eluted from oligo-dT cellulose by incubating with 2 mls of RNA elution buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 0.1% SDS) at 60°C for 5 minutes. This elution step was repeated with 1.5 mls of RNA elution buffer. The eluants were combined and extracted once with 3 mls of water-saturated phenol/chloroform (1:1), and once with 3 mls of chloroform. The aqueous phase was separated by centrifugation at 5,000 rpm (J2-21 M/E; Beckman) for 5 minutes. The poly(A)$^+$ mRNA was precipitated by incubating with 8 mls of

100% ethanol and 400 µl of 3 M sodium acetate (pH 5.2) at -20°C overnight, and pelleted by centrifugation at 10,000 rpm (J2-21 M/E; Beckman) for 40-60 minutes at 4°C. The poly(A)$^+$ mRNA pellet was washed with 70% v/v ethanol, air-dried and resuspended in 200 µl of 1× TE (pH 8.0) and 20 units of RNase inhibitor (rRNasin®; Promega). A 10 µl aliquot of the sample was diluted to 400 µl with TE for the determination of poly(A)$^+$ mRNA concentration and quality spectrophotometrically at OD_{260nm} and OD_{280nm} (PERKIN ELMER). The remaining 190 µl of poly(A)$^+$ mRNA was mixed with 550 µl of 100% v/v ethanol and 19 µl of 3 M sodium acetate (pH 5.2) and stored at -70°C. The concentration of poly(A)$^+$ mRNA was calculated by following formula:

Total amount of poly(A)$^+$ mRNA (µg) in 760 µl = OD_{260nm} x 304

2.3.3 RNA gel electrophoresis

Total RNA (20 µg) or poly(A)$^+$ mRNA (3 µg) was denatured in 10 µl of RNA loading buffer (20 mM 3-(N-Morpholino)propanesulfonic acid (MOPS), 1 mM EDTA pH 8.0, 5 mM sodium acetate pH 5.2, 50% v/v formamide, 0.63% v/v formaldehyde) at 65°C for 5 minutes, followed by fractionation on a 1% w/v agarose gel, containing 1× MOPS (20 mM MOPS, 1 mM EDTA pH 8.0, 5 mM sodiuum acetate pH 5.2), 0.63% v/v Formaldehyde, 1 µg of ethidium bromide, in 1× MOPS electrophoresis buffer. RNA was transferred to Hybond-C Extra nitrocellulose membrane (Amersham Pharmacia) using 20× SSC as the transferring buffer, according to the manufacturer's instructions, and then fixed onto the membrane by baking at 80°C in vacuum for 2 hours.

2.3.4 Northern blot analysis

RNA-bound membranes were prehybridized in 10 mls of a solution containing 50% v/v deionized formamide, 5× SSC, 1× Denhardt's, 300 µg/ml denatured herring sperm DNA, at 42°C for 2 hours, and then hybridized with denatured [α-^{32}P] dCTP-labelled DNA probe (see section 4.2.1-4) in 10 mls of a solution containing 50% deionised formamide, 1% w/v SDS, 1 M NaCl, 10% w/v dextran sulphate, 100 µg/ml denatured herring sperm DNA, at 42°C overnight. The membrane was washed twice with 2× SSC/0.1% w/v SDS, and twice with 0.2× SSC/0.1% w/v SDS for 30 minutes each at 65°C. The membrane was exposed to a

Fuji phosphorimage screen overnight and visualized using a phosphorimage analyzer (FLA-2000, FUJIFILM).

2.3.5 RNase Protection Assay

ELF5 mRNA abundance in total RNA from human cell lines was determined as described previously (Tymms, 1995). Anti-sense RNA probes for human *ELF5* and *GAPDH* transcribed from linearized plasmid vectors, generated full-length probes of 388 bp and 216 bp, respectively. The protected products generated by hybridization and RNase digestion are 298 bp for *ELF5* and 160 bp for *GAPDH*.

2.4 Protein manipulation

2.4.1 Cell lines and culture

Monkey COS7 cells were grown in Dulbecco's modified Eagle's medium (DMEM; GIBCO BRL) supplemented with 10% v/v foetal calf serum, 100 U/ml penicillin and 100 µg/ml streptomycin (GIBCO BRL), and maintained in a humidified incubator in 5% CO_2 at 37°C.

2.4.2 Plasmids

The pHis6-Elf5 expression vector was constructed as follows: The murine *Elf5* cDNA was amplified using PCR oligonucleotide primers (5'-CGGGATCCTTGGACTCCGTAACCC-ATAGC-3' and 5'-GCAGATCTCAGAGTTTCTCTTCCTGCC-3'). The forward primer contains a *Bam*HI restriction site (underlined) followed by 21 nucleotides of the murine *Elf5* coding sequence. The reverse primer contains a *Bgl*II restriction site (underlined) followed by 19 nucleotides complementary to the last 20 nucleotides of the *Elf5* coding sequence. The PCR fragment was cloned into the pGEM-T vector (Promega), and a *Bam*HI-*Sac*I restriction fragment containing the *Elf5* coding sequence was then sub-cloned into the *Bam*HI-*Sac*I sites of the pQE30 (Qiagen) bacterial expression vector resulting in a N-terminal fusion of *Elf5* protein to six histidine residues (His-Tag).

The *Elf5* mammalian expression construct (pBOSElf5s) contains the full length mouse *Elf5* cDNA blunt cloned into the blunted *Xba*I site of pEFBOS (Mizushima and Nagata, 1990). Expression from pEFBOS is driven by the elongation factor-1 promoter. The *Elf5* anti-sense expression construct is similar, but with reverse orientation of the *Elf5* cDNA. p5×poly was made by cloning multimerized polyomavirus enhancer oligonucleotides into the *Bam*HI site of pBLCAT2.

2.4.3 Electrophoretic mobility shift assays

Recombinant Elf5 and Ets1 proteins were produced as 6×His-tag fusions in *E. coli* using the QIAexpress expression system (Qiagen). Overnight cultures were diluted 1/10 in LB broth and grown for 1 hour at 37°C. Expression of recombinant proteins was induced by addition of 0.1 mM IPTG and culture of cells for two hours. Cells were harvested and sonicated in lysis buffer (6 M guanidine, 20 mM Tris-HCl, 50 mM NaCl, pH 8.0), and cell debris was removed by centrifugation. One ml of metal His-affinity resin (Clontech) was incubated with the supernatant for 30 minutes, collected, washed in wash buffer (8 M urea, 20 mM Tris-HCl, 50 mM NaCl, pH 8.0), and resuspended in renaturation buffer (20 mM Tris-HCl, 50 mM NaCl, 3 mM dithiothreitol (DTT), pH 8.0). Proteins were eluted from the beads in renaturation buffer supplemented with 100 mM imidazole. The purity and integrity of recombinant proteins was confirmed by denaturing SDS-polyacrylamide gel electrophoresis (PAGE, Coomassie blue staining).

DNA binding experiments with recombinant proteins were performed using electrophoretic mobility shift assays (EMSA), as previously described (Thomas *et al.*, 1995; Thomas *et al.*, 1997). Briefly, purified double stranded oligonucleotides were labelled with [γ-^{32}P] dATP and T4 polynucleotide kinase (Promega). The oligonucleotide probe (1 ng) was incubated for 10 minutes with approximately 20 ng purified Elf5/Ets1 protein in DNA binding buffer (1 mM EDTA, 10 mM Tris-HCl pH 8.0, 50 mM NaCl, 3 mM DTT, 1 mg/ml BSA, 500 ng/ml poly-d(I-C)d(I-C), 500 ng/ml poly dI-dC, 200 ng/ml sheared salmon sperm DNA), ± 100 ng unlabelled competitor oligonucleotide, in 10 µl final volume. Assays were run through non-denaturing, 7% w/v acrylamide (29 acrylamide: 1 bis-acrylamide; BIO-RAD), 0.5× TBE gels at 4°C.

Oligonucleotide sequences (shown in double stranded conformation)

E74	5′-gatcATAACCGGAAGTAACT-3′
	3′-TATTGGCCTTCATTGActag-5′
E74m1	5′-gatcATAACCAGAAGTAACT-3′
	3′-TATTGGTCTTCATTGActag-5′
GMETS	5′-gatcCACAGAGGAAATGATT-3′
	3′-GTGTCTCCTTTACTAActag-5′
MSV	5′-gatcGAGAGCGGAAGCGCGC-3′
	3′-CTCTCGCCTTCGCGCGctag-5′
ERBB2	5′-gatcGCTTGAGGAAGTATAA-3′
	3′-CGAACTCCTTCATATTctag-5′

2.4.4 Transfection of COS cells and chloramphenicol acetyltransferase assays

COS7 cells were transfected with 5 μg chloramphenicol acetyltransferase (CAT) reporter plasmid and 10 μg of expression constructs by electroporation. Subconfluent cells were trypsinized, washed and resuspended in growth medium (as described in section 2.3.1) supplemented with 20 mM HEPES, at 5×10^6 cells/ml. 300 μl of cells were mixed with 20 μg DNA in 0.4 cm gap electroporation cuvettes and pulsed at 180 V and 960 μF (Biorad Gene Pulsar; Hercules, CA, USA). Cells were re-plated into 10 cm Petri dishes, and harvested 48 hours later. Cell lysates were processed for CAT assays as previously described (Thomas *et al.*, 1995). The basal expression of p5×poly was arbitrarily assigned the value of '100', and other raw data was normalized to this value. Means and standard error of the means (s.e.m) were generated from four replicates of each experiment. Data was subject to statistical analysis using unpaired two-tailed *t*-tests, where resultant *p* values less than 0.05 were considered significant.

2.5 Generation of the *Elf5* knockout mouse

2.5.1 Generation of the *Elf5* knockout targeting construct

Generation of the *Elf5* knockout targeting construct is described in detail in section 5.2.5.

2.5.2 J1 Embryonic Stem (ES) cell culture

J1 ES cells of strain 129SvJ were maintained on a monolayer of neomycin-resistant embryonic fibroblasts in ES medium, comprising DMEM containing 4 mM L-glutamine and 25 mM HEPES (GIBCO BRL) supplemented with 1,000 U/ml leukemia inhibitory factor, 15% fetal calf serum, 1 mM sodium pyruvate, 0.1 mM DMEM non-essential amino acid solution, penicillin (100 U/ml), streptomycin (100 μg/ml), and 90 μM β-mercaptoethanol (GIBCO BRL). Cells were grown in tissue culture dishes (Falcon; Becton Dickson Labware, Franklin Lakes, NJ, USA) pre-coated with 0.2% w/v gelatin (in 1× PBS pH 7.4), and split every two days by dislodging the cells using 0.25% Trypsin/EDTA (GIBCO BRL).

2.5.3 Electroporation of J1 ES cells

J1 ES cells (2 x 10^7 cells; passage 16) were trypsinized and washed once with 10 mls of electroporation buffer (1× Hanks (GIBCO BRL), 20 mM HEPES (GIBCO BRL), 0.11 mM β-mercaptoethanol, 1mM NaOH) before being resuspended in 0.8 ml of electroporation buffer. The cells were transferred into an electroporation cuvette (4 cm; BIO-RAD) and incubated with 33 μg of linearized targeting construct DNA at room temperature for 5-10 minutes. Electroporation was carried out at 240 V and 500 μF using a BIO-RAD Gene Pulser. After delivering the electric pulse, cells were left at room temperature for 5-10 minutes before being plated on petri dishes that were pre-coated with neomycin-resistant embryonic fibroblasts.

2.5.4 G418 and GANC selection of J1 ES cells

Selection began 24 hours after the electroporation using 300 μg/ml of geneticin (G418; GIBCO BRL) and 2 μM of gancyclovir (GANC; Syntex Australia Limited, North Sydney, NSW, Australia). One dish was not supplemented and hence was used as a negative control. After 10 days of selection, with medium being changed daily, 336 surviving colonies were picked and cultured individually in a 24-well plate (Falcon) pre-coated with neomycin-resistant embryonic fibroblasts. When 60% confluent, two thirds of the cells were resuspended in 2× ES cell freezing medium (60% v/v ES medium, 20% v/v FCS, 20% v/v

DMSO), and left at -80°C overnight, followed by long term storage in liquid nitrogen. The remaining one third of cells was re-plated on pre-coated gelatinized plates without embryonic fibroblasts, and grown to confluency without antibiotic selection for genotyping purposes (see sections 2.1.15-16).

2.5.5 Generation of *Elf5*$^{+/-}$ and *Elf5*$^{-/-}$ mice

The correctly targeted *Elf5*$^{+/-}$ J1 ES clones, #1 and #130, were expanded for microinjection into the blastocoelic cavity of blastocysts derived from the wildtype C57BL/6J females (Papaioannou and Johnson, 1993). The blastocysts were then transferred into the uterine horns of CBA × C57Bl/6J F1 female recipients that were three days pseudopregnant. The resulting high percentage male chimaeras were mated with wildtype C57Bl/6J females to generate *Elf5*$^{+/-}$ mice. *Elf5*$^{+/-}$ mice were then intercrossed to generate *Elf5*$^{-/-}$ mice.

2.6 Phenotypic analysis of *Elf5*$^{+/-}$ and *Elf5*$^{-/-}$ mice

2.6.1 *In vitro* blastocyst outgrowth assays

Blastocysts were obtained from *Elf5*$^{+/-}$ intercrosses. The morning of the day on which a vaginal plug was detected was designated E0.5. Blastocysts were collected by flushing the uterine horns with HEPES-buffered medium 2 (M2) (Sigma, St Louis, MO, USA), and then individually cultured in poly-L-lysine-coated 96 well plates containing ES medium without leukemia inhibitory factor, in 5% CO_2 at 37°C (Shen-Li *et al.*, 2000; Takai *et al.*, 2000). Blastocyst outgrowths were inspected and photographed daily using a LEICA DMIRB microscope (5× and 20× objectives; Nussloch, Germany) and a PENTAX MZ-5N camera. After 4 days in culture, genomic DNA was isolated and genotyped as described in sections 2.1.15-16.

2.6.2 Mammary gland whole mounts

Inguinal mammary glands were dissected and spread onto an ESCO Superfrost Plus slide (Biolab Scientific, Gymea, NSW, Australia). The mammary tissue was fixed in Carnoy's fixative (6 parts of 100% v/v ethanol, 3 parts of chloroform, 1 part of glacial acetic acid) at

room temperature for 2-4 hours, and then washed in 70%, 35%, 17.5%, 8.75% and 4.375% v/v ethanol for 15 minutes each. After rinsing in MilliQ water for 5 minutes, the tissue was stained in carmine alum (0.2% carmine (SIGMA), 0.5% $AlK(SO_4)_2$) at room temperature overnight, followed by washing in 70%, 95% and 100% v/v ethanol for 15 minutes each. The mammary tissue was then cleared in histolene (Fornine, Riverstone, NSW, Australia) and mounted with 100% v/v glycerol in a Petri dish. Stained mammary tissue was photographed using a LEICA MZ6 microscope (0.63× and 4.0× objectives) and a PENTAX MZ-5N camera.

2.6.3 Histological analysis of mammary glands

Inguinal mammary glands were dissected and fixed in Bouin's fixative (75% v/v picric acid, 10% v/v formaldehyde, 5% v/v glacial acetic acid) at room temperature overnight. After dehydration with increasing concentrations of ethanol, the mammary tissue was cleared with histosol, embedded in paraffin wax and sectioned (10 μm) using a paraffin microtome (LEICA). The ribbons of tissue sections were floated in a 50°C water bath before being mounted onto ESCO Superfrost Plus slides (Biolab Scientific). Tissue sections were deparaffinized at 37°C overnight and then treated twice with histolene (Fornine) for 5 minutes each. The slides were treated with 100% (twice) and 70% v/v ethanol (once) for 5 minutes each, and then cleared in running tap water. After deparaffinization and rehydration, mammary tissue sections were treated with Harris's haematoxylin (AMBER Scientific, Belmont, WA, Australia) (5 minutes), acid-alcohol (1-2 dips to differentiate cell nuclei), Scott's tap water (AMBER Scientific) (1 minute), and then counterstained with eosin (AMBER Scientific) (5 minutes). Slides were washed in running tap water between each step. The slides were then dehydrated with 70% v/v ethanol for 2 minutes (once), 100% v/v ethanol for 2 minutes (twice), and histosol for 2 minutes (twice), before being mounted using the permanent DePex mounting media (BDH Chemicals). The mammary tissue sections were photographed using a LEICA DMR microscope (5× and 40× objectives).

2.7 Animal Work

2.7.1 General husbandry and maintenance

All mice were housed in windowless rooms with controlled temperature (22 ± 2°C) on a 12 hour light (8:00 AM to 8:00 PM) and dark cycle. Pups were weaned from the mothers at 21 days of age, ear tagged and 0.5 cm of their tail cut for genotyping purposes (see sections 2.1.15-16).

2.7.2 Collection of mouse embryos and tissues

Mouse blastocysts were obtained as described in section 2.5.1. Postimplantation embryos were obtained from intercrossed *Elf5*$^{+/-}$ females. The morning of the day on which a vaginal plug was detected was designated E0.5. The pregnant female was killed by cervical dislocation. The embryos and their yolk sac were dissected in ice-cold 1× PBS under the dissecting microscope, and genotyped as described in sections 2.1.15-16. For RNA expression studies, tissues were dissected from embryos, neonates and adult mice that were decapitated, and then snap-frozen in liquid nitrogen.

Chapter 3

Cloning and characterization of a novel ETS transcription factor, ELF5

3.1 Introduction

The ETS transcription factors are a large family implicated in the control of cellular proliferation, differentiation and tumorigenesis. The first member of the *ETS* gene family, *v-ets*, was discovered upon studying the avian erythroblastosis virus E26 in the early 1980s (Leprince *et al.*, 1983; Nunn *et al.*, 1983). Subsequently, numerous cellular *ETS* genes were identified in species ranging from sea urchin to human, to form this still growing transcription factor family.

The ETS family of transcription factors share a highly conserved DNA binding domain, which covers approximately 85 amino acids, termed the 'ETS domain' (Watson *et al.*, 1988; Karim *et al.*, 1990; Gutman and Wasylyk, 1990; Macleod *et al.*, 1992; Seth *et al.*, 1992; Wasylyk *et al.*, 1993; Graves and Petersen, 1998). The ETS domain recognizes and binds to purine rich GGA(A/T) core motifs in the promoters and enhancers of various target genes (Macleod *et al.*, 1992; Janknecht and Nordheim, 1993; Wasylyk *et al.*, 1993; Werner *et al.*, 1995; Kodandapani *et al.*, 1996). The ETS family does not maintain overall similarity outside of the ETS domain, but can be grouped into subfamilies based upon variation within the ETS domain, and also by the arrangement and presence of other protein domains (Janknecht and Nordheim, 1993; Wasylyk *et al.*, 1993).

ETS factors also have important developmental roles and many ETS factors have been implicated in the control of cellular proliferation and tumorigenesis (see chapter 1). *ETS1,*

ETS2, ERG2 and *PU.1* are proto-oncogenes with mitogenic and transforming activity when overexpressed in fibroblasts (Seth *et al.*, 1989; Seth and Papas, 1990; Hart *et al.*, 1995; Moreau-Gachelin *et al.*, 1996). In addition, chromosomal translocations involving ETS family members are associated with different human cancers. Also, ETS1, ETS2 and ERG regulate the expression of metalloproteinase genes, such as stromelysin and collagenase (Wasylyk *et al.*, 1991; Buttice and Kurkinen, 1993; Buttice *et al.*, 1996), which are important for extracellular matrix degradation concomitant with tumor vascularization (angiogenesis) and metastasis.

The ETS factors are almost all expressed in hematopoietic lineages (Bhat *et al.*, 1989; Bhat *et al.*, 1990; Kola *et al.*, 1993), and indeed appear to function predominantly in these cells and their related neoplasms. However, the most common solid tumors in humans are carcinomas that arise from the transformation of epithelial cells. Transformed breast epithelial cells, for example, have been shown to express ETS family members GABPα, PEA3, ELF1, ETS1 and ELK1 (Scott *et al.*, 1994b; Delannoy-Courdent *et al.*, 1996), but expression of these ETS family members is not restricted to epithelial cells. One ETS family member, ELF3/ESX/ESE-1/JEN/ERT, has recently emerged with epithelial and epithelial cancer-specific expression (Andreoli *et al.*, 1997; Chang *et al.*, 1997; Choi *et al.*, 1998; Oettgen *et al.*, 1997; Tymms *et al.*, 1997). Given the extensive involvement of ETS factors in development and tumorigenesis, it becomes important to identify any additional ETS genes that are epithelial-specific, especially those that may be involved in epithelially derived cancers.

This chapter describes the cloning and initial characterization of a novel ETS family member, ELF5 (E74-Like-Factor 5). In a comparison of the ETS domains, ELF5 displays the strongest similarity to the ELF/E74 subfamily of ETS factors, and is most closely related to ELF3 and ESE3. Other than a conserved pointed domain, ELF5 has little similarity to ETS family members outside of the ETS domain. ELF5 functions as a transcription factor with similar sequence-specific DNA binding characteristics to other ETS family members.

3.2 Results

3.2.1 Isolation of the mouse *Elf5* cDNA

The murine *Elf5* cDNA was isolated from an adult mouse lung cDNA library following screening with a cDNA probe containing the ETS domain region of human *ELF3*. Compilation of sequence data revealed a 1437 bp sequence with a maximum open reading frame (ORF) of 759 bp, predicted to encode a 253 amino acid protein of approximately 30 kDa (Figure 3.1 a). An upstream, in-frame stop codon suggests that this ORF represents the full-length coding sequence of *Elf5*. No evidence was found for alternative splicing within the ORF. Additionally, 91 bp of 5′ and 696 bp of 3′ sequence was obtained by reverse transcriptase polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE), using embryonic day 14 mouse placental RNA. Sequence analysis revealed two discrete polyadenylation signals present in the 3′ untranslated region (UTR). The first of these (at 1391 bp) appears to be an overlapping poly-A recognition signal, AATTAA and ATTAAAA, similar to that identified in the gene for C4b-bindng protein by Kristensen *et al.* (1987). The second is a consensus polyadenylation signal, AATAAA, at 2181 bp. These polyadenylation signals are found close to the 3′ terminus of the original clone and the 3′ RACE product, respectively, suggesting that these represent poly-A signals for two separate mRNA products. Thus, the two predicted *Elf5* cDNAs are 1528 bp and 2224 bp long (Figure 3.1 a). Northern blot analysis, using the *Elf5* coding sequence as a probe, confirmed the presence of two predominant *Elf5* transcripts in day 14 mouse placental tissue, *Elf5-a* and *Elf5-b*, of approximately 2.5 kb and 1.5 kb, respectively. Only *Elf5-a* was identified using a 3′ UTR fragment from between the polyadenylation signals as a probe (Figure 3.1 a & b), indicating that the transcripts differ in their 3′ UTR sequences. Interestingly, the sequence found in *Elf5-a*, but not *Elf5-b*, contains multiple ATTTA motifs that are associated with rapid mRNA turnover (Savant-Bhonsale and Cleveland, 1992; Akashi *et al.*, 1994) (Figure 3.1 a).

3.2.2 Isolation of the human *ELF5* cDNA

A partial 1 kb human *ELF5* cDNA fragment was isolated from a human fetal lung cDNA library following screening with a cDNA probe containing the coding sequence of mouse

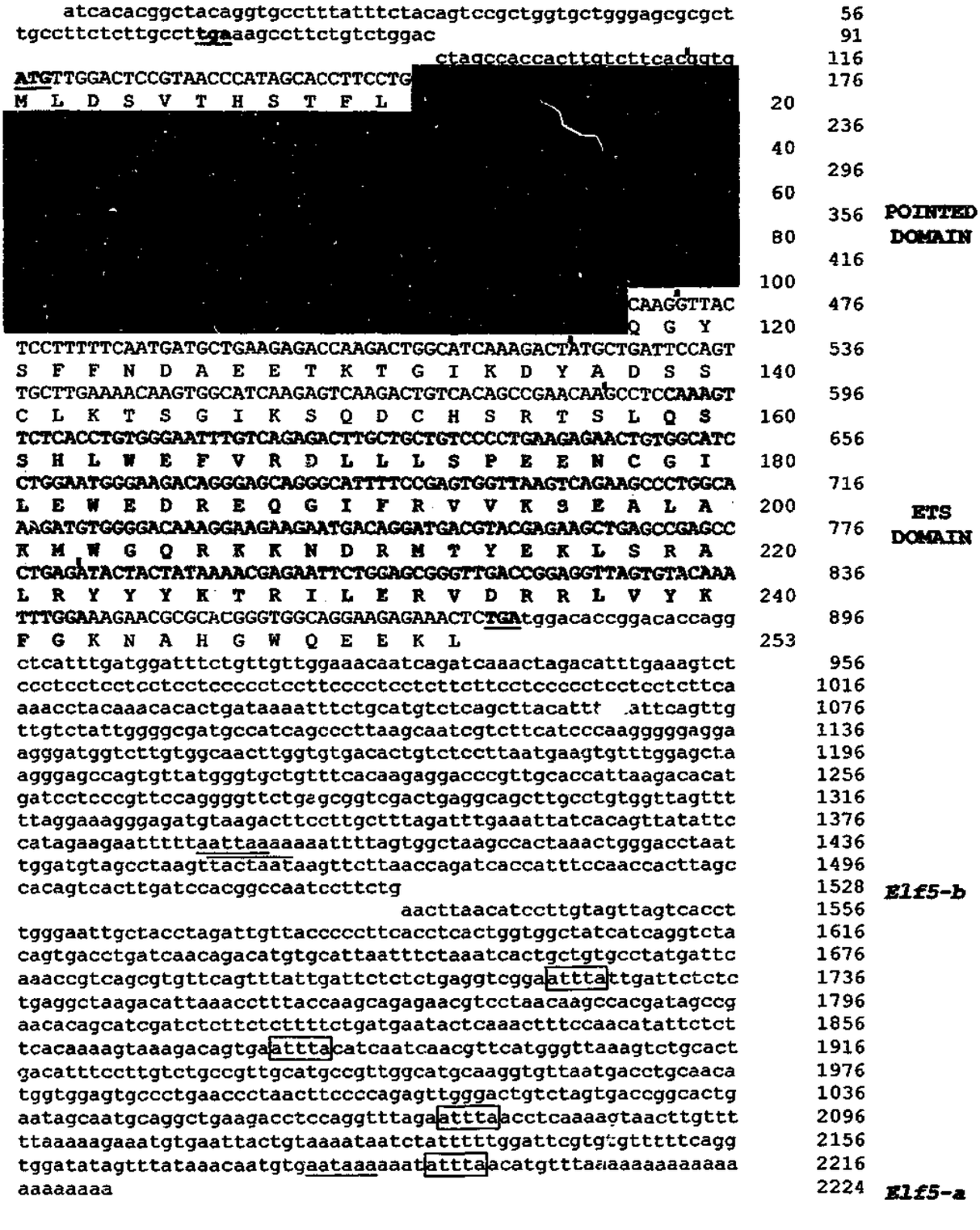


Figure 3.1 Mouse *Elf5* cDNA. **(a)** Complete nucleotide sequence and predicted amino acid sequence of mouse *Elf5*. *Elf5-a* and *Elf5-b* share the same coding sequence. *Elf5-a* has additional 696 bp of 3'-UTR sequence as indicated. The open reading frame (ORF) is shown in one-letter code, with the initiating start and stop codons underlined. A stop codon, in the same reading frame as the ORF, but 5 prime to the initiating codon, is also underlined. The Pointed domain is shaded in blue and the Ets domain is shaded in yellow. The putative polyadenylation signals are underlined. The ATTTA motifs associated with rapid mRNA turnover are boxed. The known exon-intron boundaries are indicated with red arrows.

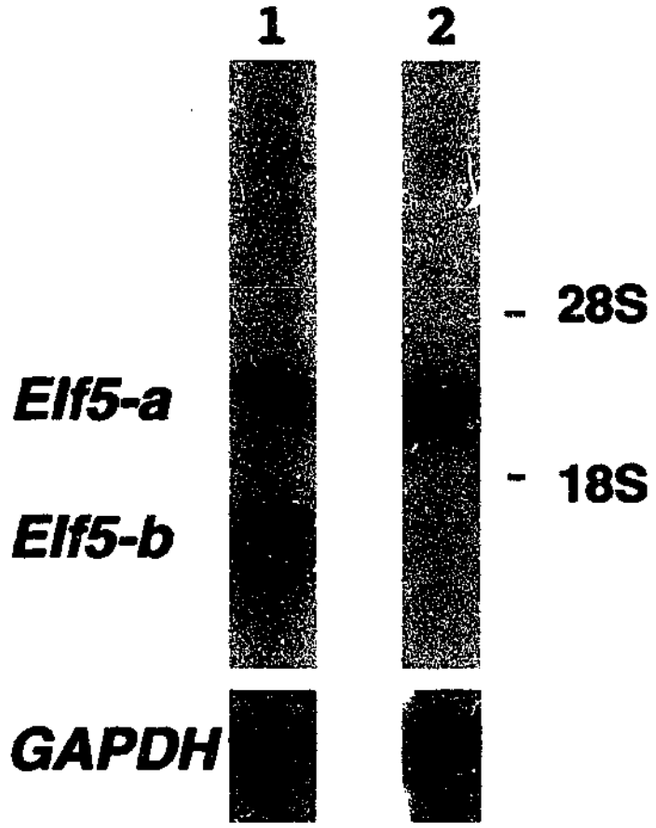


Figure 3.1 Mouse *Elf5* mRNA. **(b)** Northern blot analysis of murine *Elf5* expression in day 14 mouse placenta: lane 1, probed with random-prime-labelled 940 bp *Sty*I murine *Elf5* cDNA fragment (probe 1); lane 2, probed with random-prime-labelled murine *Elf5* 696 bp 3′-RACE PCR product (probe 2). Size markers are indicated. Both lanes were also probed with *GAPDH* cDNA (lower panels).

Elf5. Sequence analysis showed that it started with a partial ORF encoding 57 amino acids prior to the TGA stop codon. Fifty five of these amino acids were identical to the mouse Elf5 protein sequence. An ATTTA motif that is associated with rapid mRNA turnover (Savant-Bhonsale and Cleveland, 1992; Akashi *et al.*, 1994) was detected in the 3′ UTR, but no polyadenylation signal was detected in the 3′ UTR sequences indicating that the clone did not contain the 3′ end. RT-PCR and RACE using human placental RNA was then used to obtain the full coding sequence of human *ELF5*. Two alternatively spliced forms, *ELF5-a* and *ELF5-b*, were identified. Sequence analysis revealed an open reading frame of 255 amino acids and a predicted molecular mass of approximately 30 kDa for ELF5-a (Figure 3.2 a) and an open reading frame encoding a 160 amino acid protein with a predicted molecular mass of approximately 20 kDa for ELF5-b (Figure 3.2 b). Compared to *ELF5-a, ELF5-b* is missing a 285 bp spliced DNA fragment. The spliced junctions of the alternative transcript coincide with the exon-intron boundaries of the human *ELF5* genomic structure, and the spliced DNA fragment encompasses two consecutive exons (Figure 3.2 a & b). This alternative splicing event has caused an amino acid residue substitution (amino acid #41, D→A) at the splice junction in the predicted human ELF5-b protein sequence (Figure 3.2 a & b). *ELF5-b* also has an additional 24 bp extension at the 5′ UTR, that may be important in the regulation of *ELF5-b* transcription. Human ELF5-a contains an additional two amino acid insertion compared to mouse Elf5 (Figure 3.2 c).

Recently, a third human *ELF5* isoform, *ESE-2a*, together with *ELF5-a* (also named *ESE-2b*) was isolated from a human adult prostate cDNA library (Oettgen *et al.*, 1999). Sequence analysis revealed an open reading frame encoding a 265 amino acid protein with a predicted molecular mass of 31 kDa for ESE-2a (Figure 3.2 d). The 5′-end of *ELF5-a/ESE-2b* is divergent from *ESE-2a* due to the presence of an alternative 5′ exon, indicating the presence of different transcription start sites possibly derived from different promoters (Figure 3.2 e). As a result of this alternative splicing, the ESE-2a isoform has a 10 amino acid extension at the amino terminus in comparison with ELF5-a/ESE-2b (Figure 3.2 a & c & d). Several in-frame stop codons upstream of the ATG initiation condon are found in *ESE-2a*, but no in-frame stop codons are detected upstream of the ATG initiation codon in *ELF5-a/ESE-2b*. The 3′ UTR sequences of both cDNAs are the same and contain four classical polyadenylation signals, two ATTTA motifs and a highly repetitive sequence element (an Alu sequence) (Figure 3.2 a-c).

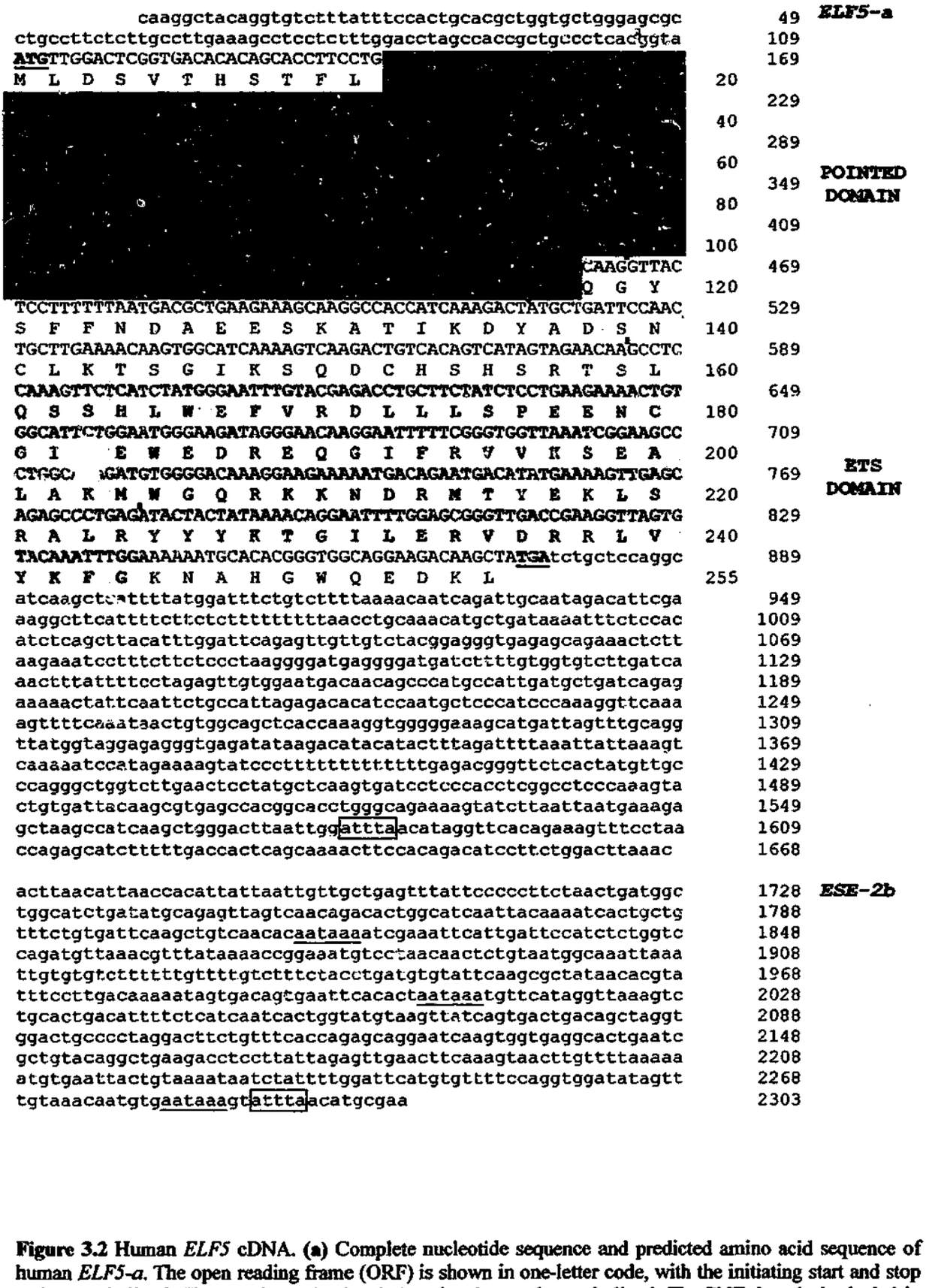


Figure 3.2 Human *ELF5* cDNA. **(a)** Complete nucleotide sequence and predicted amino acid sequence of human *ELF5-a*. The open reading frame (ORF) is shown in one-letter code, with the initiating start and stop codons underlined. The putative polyadenylation signals are also underlined. The PNT domain is shaded in blue and the ETS domain is shaded in yellow. The ATTTA motifs associated with rapid mRNA turnover are boxed. The known exon-intron boundaries are indicated with red arrows. The additional 3′-UTR sequence identified in *ESE-2b* (Oettgen *et al.*, 2000) is marked and separated by a blank line.

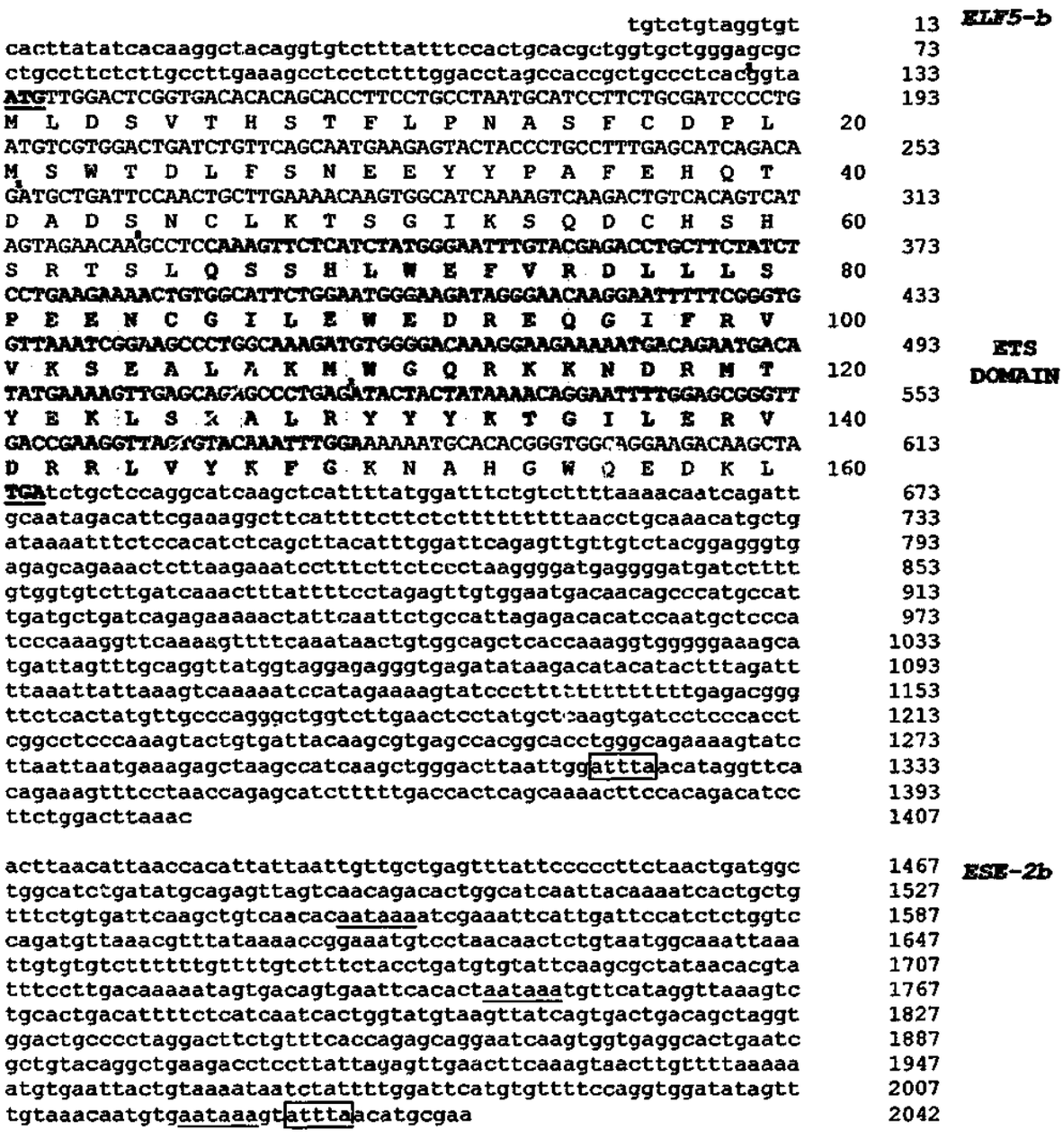


Figure 3.2 Human *ELF5* cDNA. **(b)** Complete nucleotide sequence and predicted amino acid sequence of human *ELF5-b*. The open reading frame (ORF) is shown in one-letter code, with the initiating start and stop codons underlined. The putative polyadenylation signals are also underlined. The ETS domain is shaded in yellow. The ATTTA motifs associated with rapid mRNA turnover are boxed. The known exon-intron boundaries are indicated with red arrows. The additional 3′-UTR sequence identified in *ESE-2b* (Oettgen *et al.*, 2000) is marked and separated by a blank line.

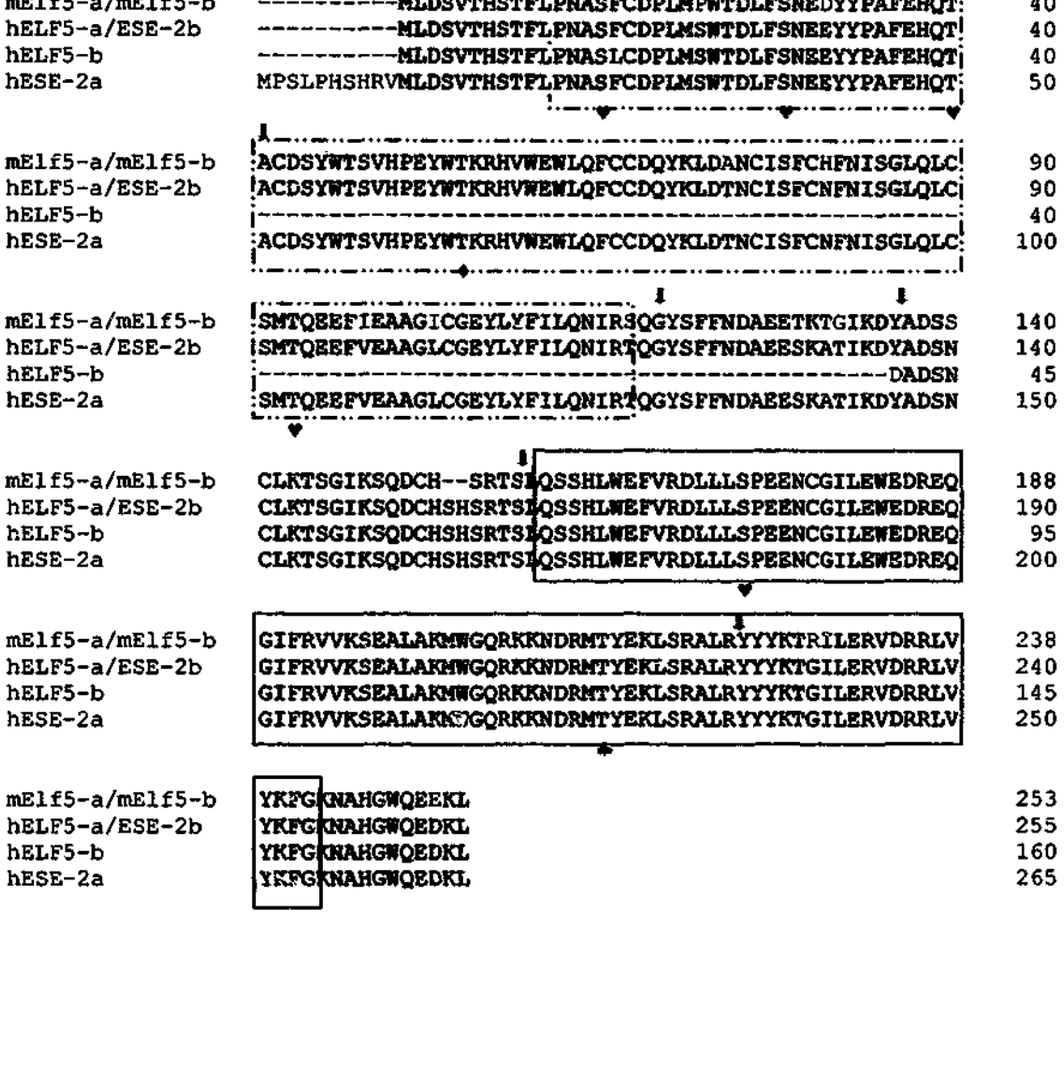


Figure 3.2 Human *ELF5* gene. **(c)** Comparison of human ELF5 and mouse Elf5 ORFs. The open reading frame (ORF) is shown in one-letter code. Amino acids that are present in at least two proteins are shaded in yellow. The ETS domain is boxed with a solid line and the pointed domain with a dashed line. Putative phosphorylation sites, conserved between the two species are marked: ♥ as CKII (casein kinase II), ♦ as PKC (protein kinase C) and ♣ as TyP (tyrosine kinase). The known exon-intron boundaries are indicated with red arrows.

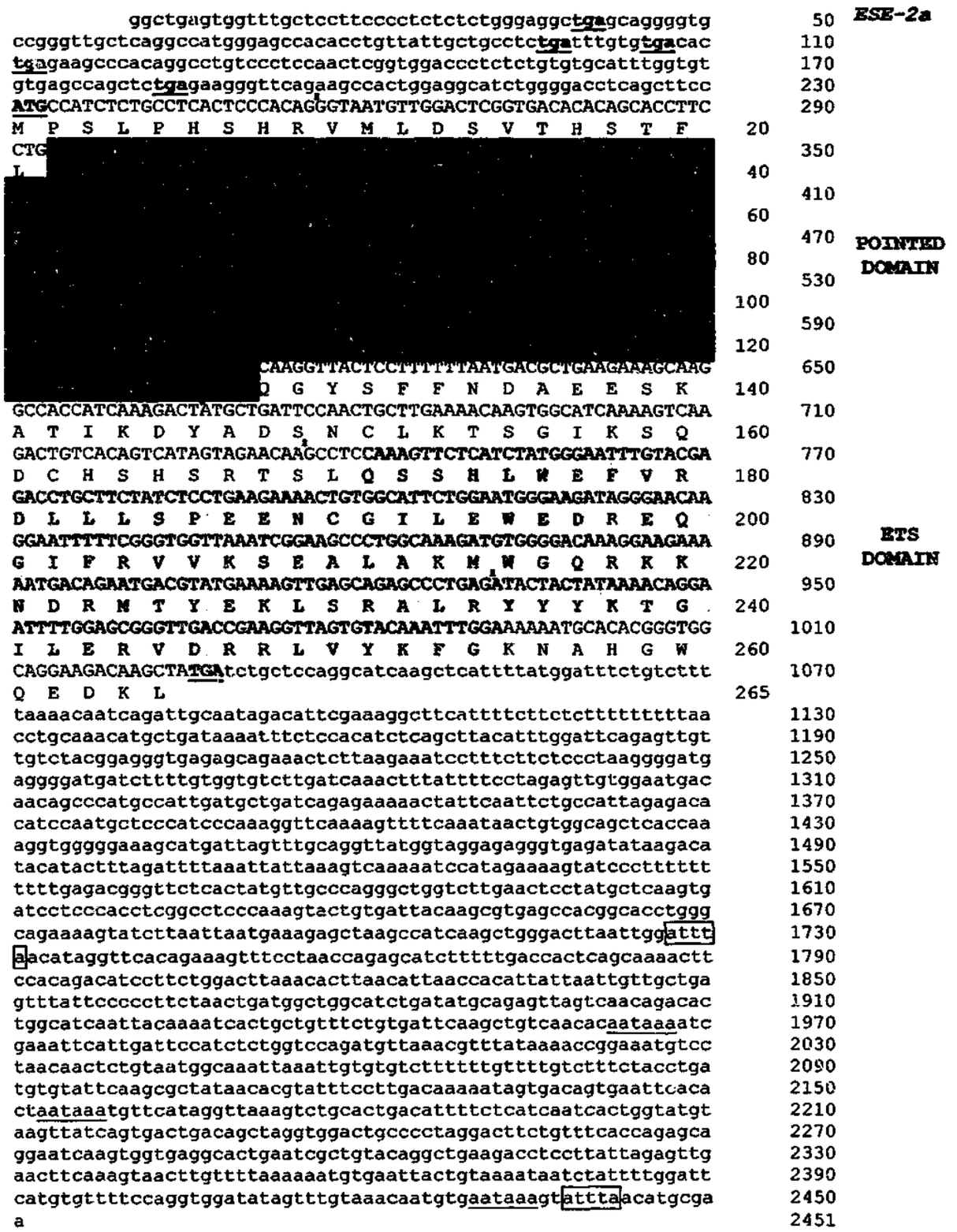


Figure 3.2 Human *ELF5* cDNA. **(d)** Complete nucleotide sequence and predicted amino acid sequence of human *ESE-2a*. The open reading frame (ORF) is shown in one-letter code, with the initiating start and stop codons underlined. The stop codons, in the same reading frame as the ORF, but 5-prime to the initiating codon, are also underlined. The PNT domain is shaded in blue and the ETS domain is shaded in yellow. The putative polyadenylation signals are underlined. The ATTTA motifs associated with rapid mRNA turnover are boxed. The known exon-intron boundaries are indicated with red arrows.

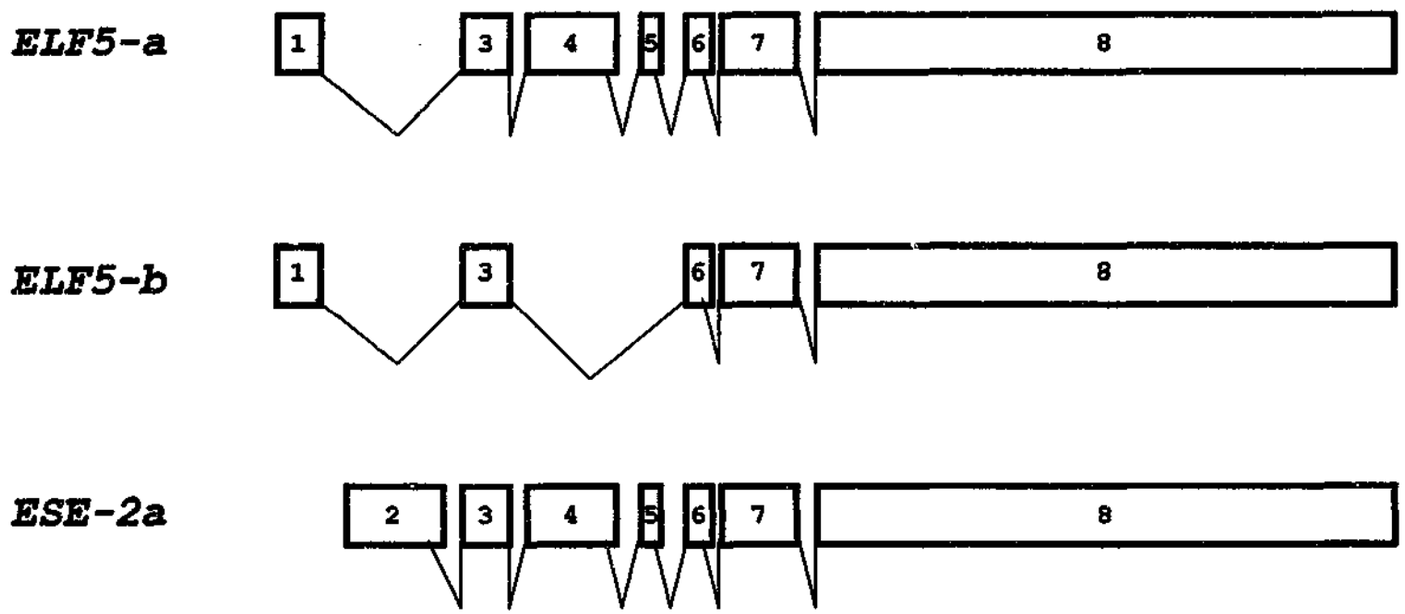


Figure 3.2 Human *ELF5* gene. **(e)** Three isoforms of human *ELF5* gene arise from the utilization of alternative exons.

3.2.3 Characterization of human ELF5 and mouse Elf5 protein sequences

The predicted amino acid sequences of human ELF5-a and mouse Elf5 are highly conserved, with approximately 95% overall identity (Figure 3.2 c). Only a single amino acid substitution was observed within the putative ETS domains of human ELF5-a and mouse Elf5, with most of the other amino acid residue changes being conservative substitutions (8/13), suggesting that the two proteins are homologs (i.e. having an inferred common ancestry). However, two additional isoforms are detected in human tissues. The human isoform ELF5-b was isolated from placenta. ELF5-b has a 95 amino acid deletion that removes most of the pointed domain, but it is otherwise identical to ELF5-a (Figure 3.2 c). The third human ELF5 isoform, ESE-2a, was detected in prostate. ESE-2a has a 10 amino acid extension at the amino terminus in comparison with ELF5-a/ESE-2b, but is also otherwise identical to ELF5-a (Figure 3.2 c). The murine counterparts for human ELF5-b and ESE-2a were not detected.

3.2.3.1 The ETS domain in the ELF5 protein is highly conserved

Transcription factors contain DNA-binding domains that recognize specific DNA motifs/targets, and this sequence-specific DNA binding will result in transcriptional regulation of the target gene promoter. Interestingly, many transcription factors share similar structural motifs for DNA binding, and this sequence conservation in the DNA-binding domains has been used to define and group transcription factors into families. The ETS family of transcription factors is shown to share a highly conserved DNA-binding domain of approximately 85 amino acids, termed the 'ETS domain' (Watson *et al.*, 1988; Karim *et al.*, 1990; Gutman and Wasylyk, 1990; Macleod *et al.*, 1992; Seth *et al.*, 1992; Wasylyk *et al.*, 1993; Graves and Petersen, 1998).

ELF5 has a highly conserved DNA-binding domain situated at the carboxyl terminus of the protein. It is highly similar to that of human ELF3, with 65% amino acid identity. However, this domain is only moderately similar to that of other ETS family members. It shares 51% amino acid identity with human NERF, 50% with human MEF, *Drosophila* ets4 and e74, and 49% with human ELF1. Sequence identity with other family members is in the range of 46% to 31% (Figure 3.3 a). Recent studies have shown that human ESE3 is closely related

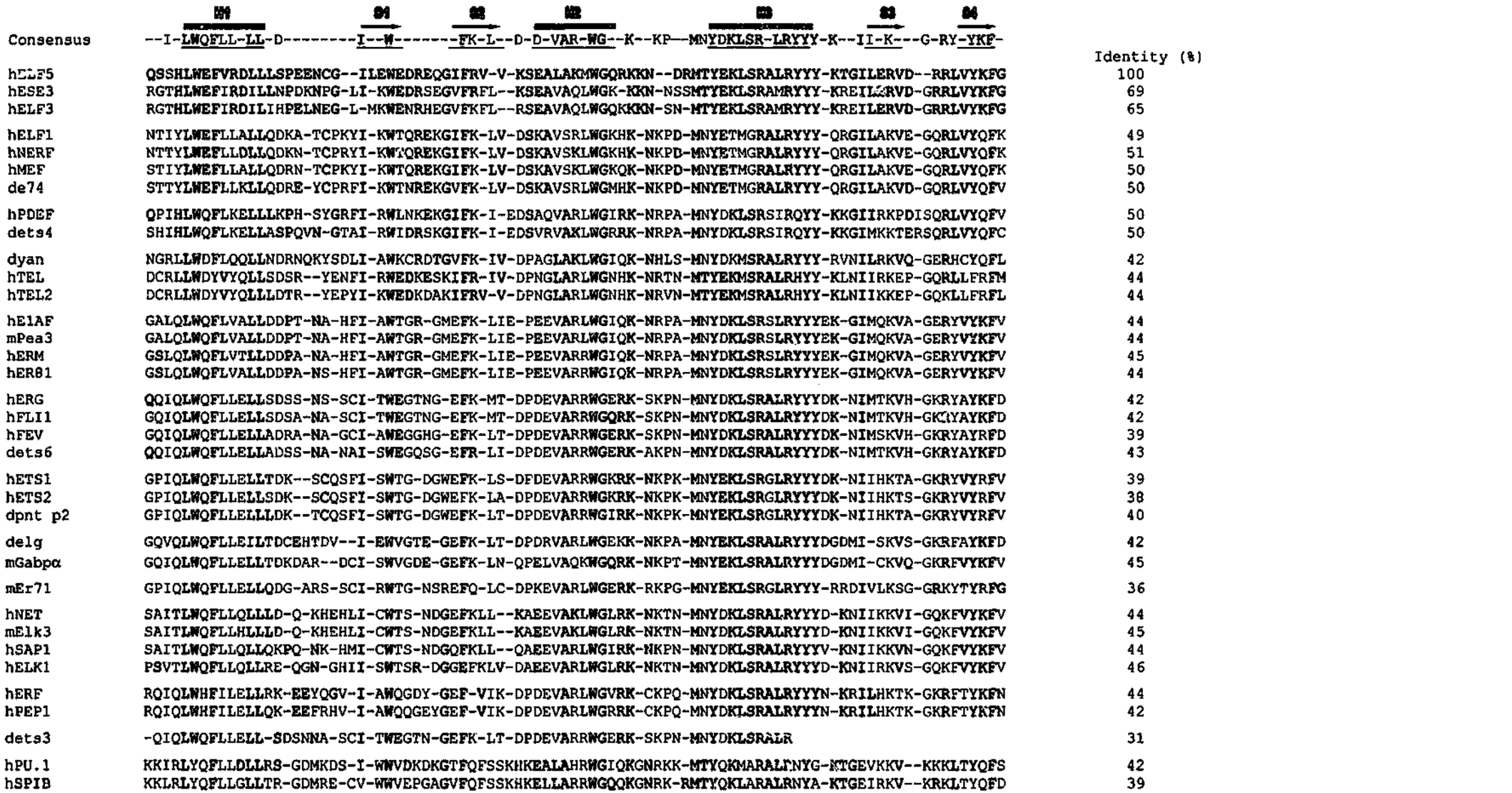


Figure 3.3 The ETS domain of ELF5. **(a)** Comparison of the ETS domain of human ELF5 with those of known members of the ETS gene family. The alignment was generated by using CLUSTAL W (Thompson *et al.*, 1994) with the default settings, and the result was subsequently adjusted manually. The ETS factors examined are labelled on the left, where 'h' denotes human, 'm' mouse and 'd' *Drosophila*. The ETS consensus sequence is a list of the amino acids most often conserved between ETS family members. Shading denotes amino acid identity with human ELF5, and the percent identity of each ETS domain is indicated on the right. Tentative α-helices (H) and β-sheets (S) as predicted by NMR and x-ray crystallography for some of the ETS domains are indicated above the consensus sequence by blue boxes and red arrows, respectively.

to human ELF3 and human ELF5, with 84% and 69% amino acid identity, respectively, within the ETS domain (Kas *et al.*, 2000). The human PDEF has also been show to be quite closely related to human ELF5 within the ETS domain, with 50% amino acid identity (Oettgen *et al.*, 2000).

The amino acids that are highly conserved amongst ETS family members (Janknecht and Nordheim, 1993) are also invariant in ELF5 (29/48) (Figure 3.3 a). In particular, all but one of the amino acid residues in the three α-helices and four β-strands, identified in the ETS domains of mEts1 and hFLI1 (Donaldson *et al.*, 1994; Liang *et al.*, 1994), are either invariant or conservative substitutions.

Based on ETS domain similarities, a recent phylogenetic analysis (Graves and Petersen, 1998) proposed the grouping of ETS factors into subfamilies, one of which is the ELF (E74-Like-Factor) subfamily. The ELF subfamily includes *Drosophila* e74, human ELF1 and NERF. We generated a phylogenetic tree, including ELF5 and ELF3, by maximum likelihood analysis of the ETS domains (Figure 3.3 b). It shows that the human ELF5 and mouse Elf5 sequences group most closely with the human ELF3 and mouse Elf3 sequences, and that both ELF3 and ELF5 are most closely related to *Drosophila* ets4, e74 and human ELF1 and NERF. Thus, *Drosophila* ets4, and human/mouse ELF3 and ELF5 may also fall into the ELF subfamily of ETS factors. It is for this reason that ELF5 (E74-Like-Factor 5) is so named. Our analysis on the phylogenetic relationship of hELF3, mElf3, dets4, hELF1, hNERF, de74, hTEL, dyan, hERM, mEr81, mPea3, hELK1, mErp and hSAP1 sequences was identical to that published by Tymms *et al.*, (1997). However, the phylogenetic relationship among the remaining 12 sequences (ELF5 excluded) was different. These differences occur at the basal bipartitions in the data, probably due to the use of a more rigorous phylogenetic method than that used previously. The same phylogenetic relationship among the ETS domains in dyan, hTEL, de74, hELF1, hNERF and dets4 was reported by Graves and Petersen (1998), who obtained a majority-rule consensus of 1,000 neighbor-joining trees generated by bootstrap analysis.

The phylogeny depicted in Figure 3.3 b shows the unrooted relationship among 28 ETS domains. Any attempt to infer its root would lead to the conclusion that there has been substantial variation in the substitution rate. Because variation in this rate may affect the chances of recovering the true phylogenetic tree (Hillis *et al.*, 1994), we remain cautious

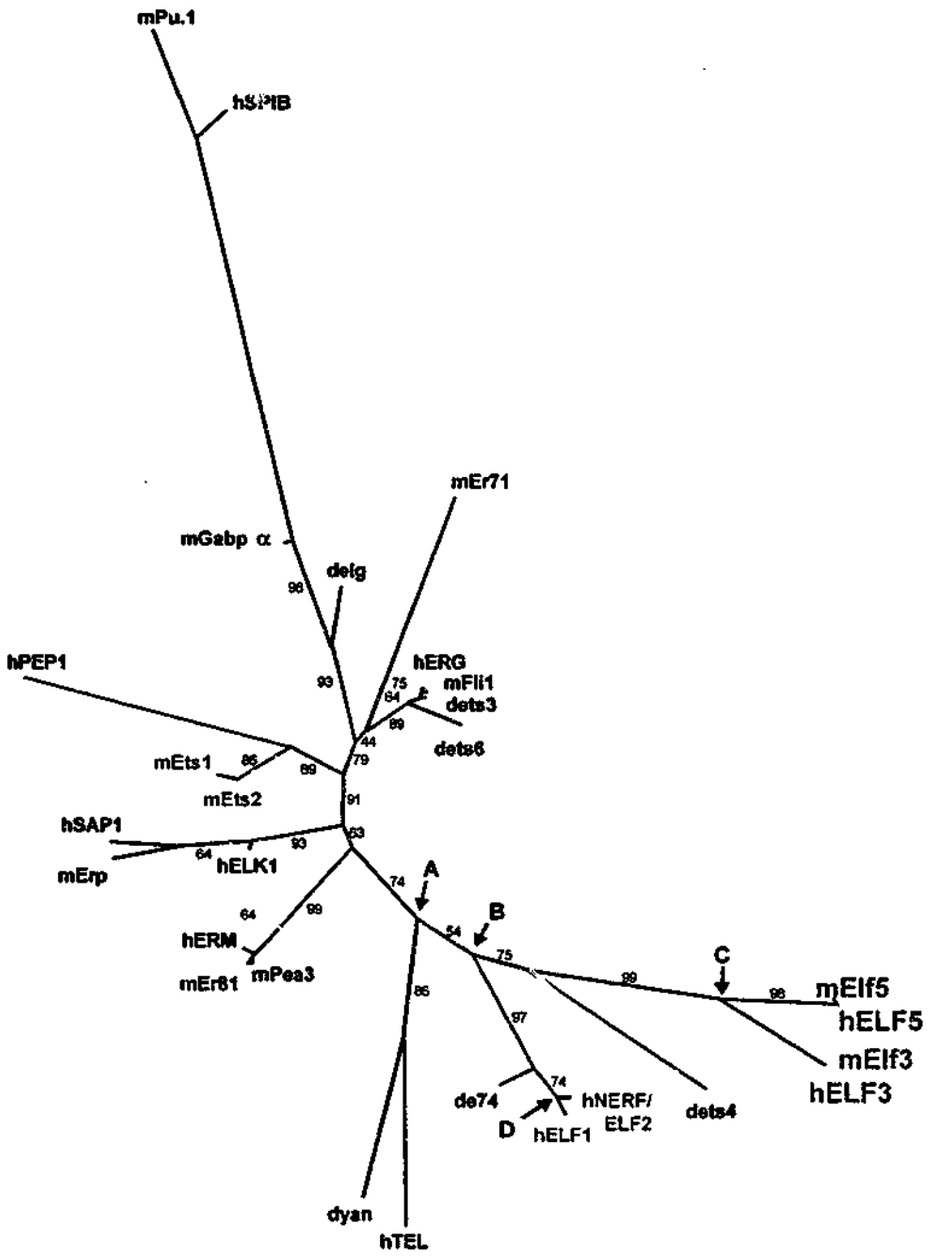


Figure 3.3 The ETS domain of ELF5. **(b)** Phylogenetic tree of the ETS domain produced by maximum likelihood analysis. The ETS domain alignment was analyzed using the JTT-F substitution model (Jones *et al.*, 1992) and local bootstrap values were estimated for all internal branches, both by using PROTML in Q mode followed by a second run in R mode (Adachi and Hasegawa, 1996). An underlying assumption of the phylogenetic analysis is that the amino acid content does not vary significantly among the sequences. This assumption was not assessed because tools for doing so are still under development (LSJ, unpublished work). Therefore, the tree may be the result of both historical and compositional components. The four points at which gene duplications have been inferred are marked A, B, C and D.

about the phylogeny as a whole. However, if we assume that the root is somewhere near the centre of the tree, we may focus on the evolution of ETS domains in dyan, hTEL, de74, hELF1, hNERF, dets4, hELF3, mElf3, hELF5 and mElf5. Given that these ETS domains occur in three taxa, we can infer at least four gene duplications (A, B, C & D in Figure 3.3 b). The first two gene duplications (A & B) involve both protostome and deuterostome taxa (*Drosophila* and mammals, respectively) and therefore must have occurred before the origin of the deuterostome lineage 550-750 Myr ago (Doolittle *et al.*, 1996). A third gene duplication (C) involves two mammalian lineages (Rodenta and Primata) and therefore must have occurred 115–129 Myr ago (Easteal and Herbert, 1997). A fourth gene duplication (D) involves human sequences only and therefore we cannot infer its age. Consequently, we conclude that hELF5 and mElf5 (like hELF3 and mElf3) are orthologous gene products (ie. their origin can be traced back to a a speciation event) and that ELF3 and ELF5 are paralogous gene products (ie. their origin can be traced back to a gene duplication) (Moritz and Hillis, 1996). Interestingly, ELF5 displays considerable N-terminal sequence similarity (outside the ETS domain) with ETS1, ETS2, ERG, FLI1, TEL and Gabpα (Figure 3.4) which will be described in section 3.2.3.2.

Graves and Petersen (1998) produced a dendrogram based on sequence similarity within the ETS domains. This type of analysis grouped the 49 known ETS family members into nine subfamilies. Since then, a number of novel epithelial-specific ETS factors have been identified, including ELF3 (Andreoli *et al.*, 1997; Chang *et al.*, 1997; Choi *et al.*, 1998; Oettgen *et al.*, 1997; Tymms *et al.*, 1997), ELF5, PDEF (Oettgen *et al.*, 2000, Yamada *et al.*, 2000) and ESE3 (Kas *et al.*, 2000, Bochert *et al.*, 1998). We have extended the dendrogram to incorporate these new members (Figure 3.3 c). Three of the epithelial-specific ETS transcription factors (ELF5, ESE3 and ELF3) are grouped together close to the ELF subfamily. Therefore, we suggest that ELF5, ESE3 and ELF3 fall into the ELF subfamily. The other epithelial-specific ETS factor, PDEF, is most closely related to *Drosophila* ets4 and appears more distantly related to ELF5, ESE3 and ELF3, than ELF1, NERF, MEF and *Drosophila* e74.

3.2.3.2 The Pointed domain in the ELF5 protein is conserved

The PNT domain present in many of the ETS family members, such as pointed p2, TEL, yan, elg, Gabpα, ETS1, ETS2, ERG, FLI1, ELF3, ESE3 and PDEF, also appears to be

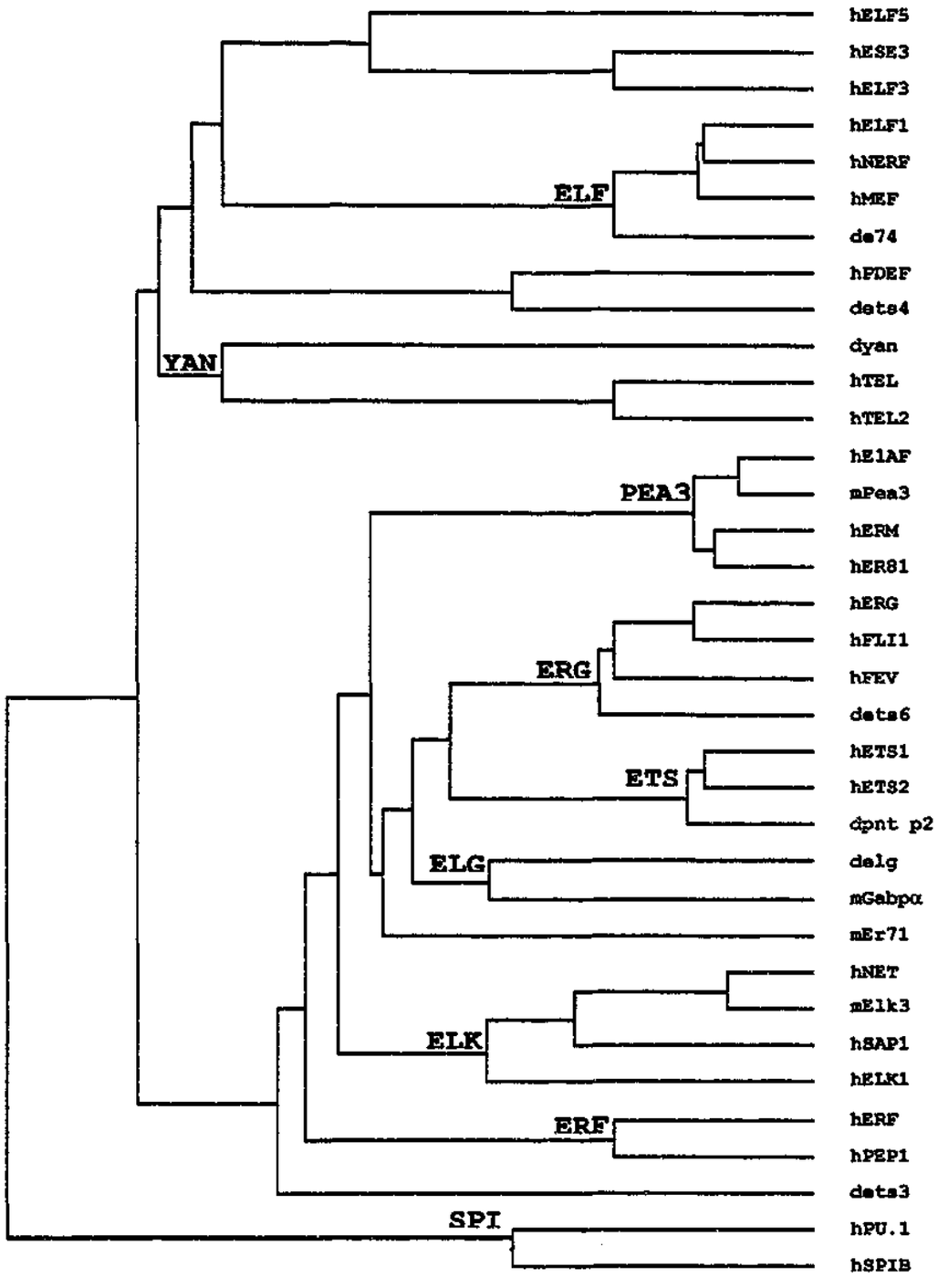


Figure 3.3 The ETS domain of ELF5. **(c)** Dendrogram showing the ETS domain family tree, based on amino acid sequence similarity using CLUSTAL W method. The names for the subfamily suggested by Graves and Petersen (1998) are indicated.

present in ELF5. Although the hELF5 and mElf5 pointed domains display 93% amino acid identity, the similarities to the PNT domains of other ETS family members is relatively low. For example, the PNT domains of hESE3 and mEhf show only 42% and 41% amino acid identity to the PNT domain of ELF5, respectively. The identity between hELF5 and hELF3 is 22% whereas the identity between hELF5 and mElf3 within the PNT domain is 26%. The amino acid residue identity to other ETS family members is in the range of 18% to 25% (Figure 3.4 a). Although this similarity is relatively weak, of those amino acid residues conserved between ETS family members, 22/60 are also present in ELF5. Indeed, within the four α-helices (H2-H5), 25 out of 31 amino acid residues are either invariant or conservative substitutions (Figure 3.4 a). Therefore, it is likely that the ELF5 PNT domain is involved in similar protein-protein interactions as those described for other ETS family members. However, a mitogen activated protein (MAP) kinase site, involved in enhancing transactivation (Yang *et al.*, 1996), present at the extreme amino end of the pointed domains of ETS1, ETS2 and pnt p2, does not appear to be present in ELF5.

A dendrogram generated based upon amino acid sequence identity in the PNT domain changes the groupings slightly. Instead of having hESE3/mEhf and hElf3/mElf3 grouped closely together, hELF5/mElf5 is grouped with hESE3/mEhf (Figure 3.4 b). However, the general grouping of ELF5, ESE3 and ELF3 is retained and coincides with the grouping based on the similarity within the ETS domain (Figure 3.3 c & 3.4 b).

3.2.4 Mouse Elf5 is capable of binding to DNA in a sequence-specific manner

All ETS proteins bind to similar DNA sequences because of the high conservation within the ETS domain. A combination of genetic and biochemical findings have provided an overall picture of this consensus binding sequence (reviewed by Graves and Petersen, 1998). Short DNA sequences, of approximately 9-15 bp, that contain a GGA core were identified in the promoters and enhancers of various genes, and were shown to have high affinity for ETS proteins. A number of ETS proteins were shown to bind an ETS-binding site that normally binds mouse Ets1 (Nye *et al.*, 1992; Gunther and Graves, 1994; Graves *et al.*, 1996). However, a number of ETS DNA-binding studies have shown that different ETS proteins can discriminate and select ETS binding sites that deviate from the consensus sequence (Nelsen *et al.*, 1993; Shin and Koshland, 1993; Ray-Gallet *et al.*, 1995; John *et al.*, 1996).

a

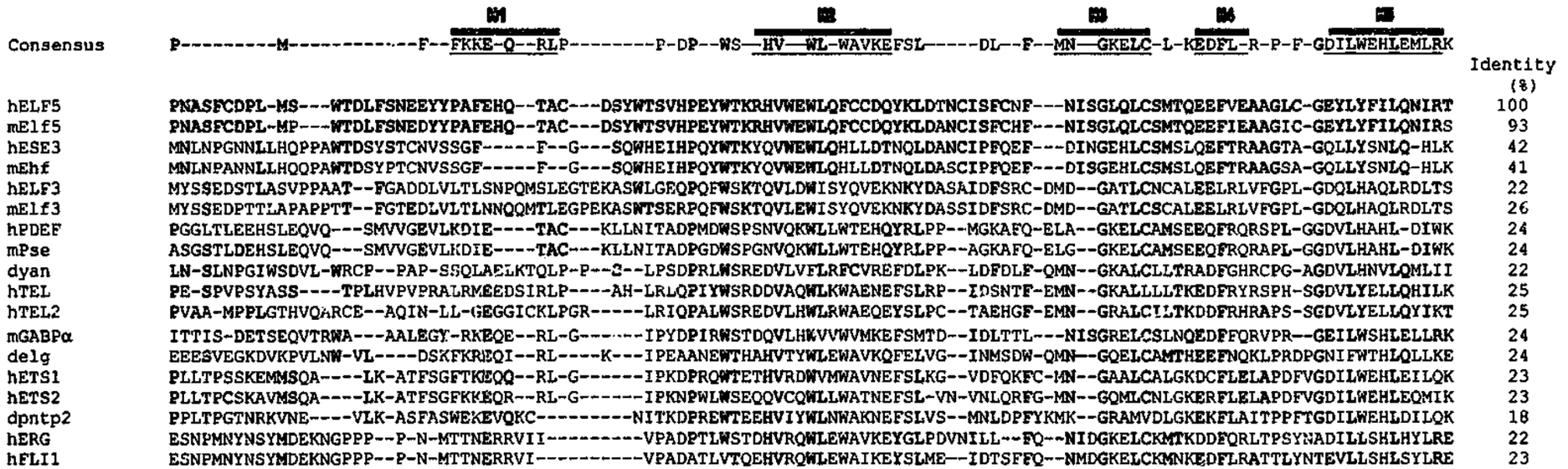


Figure 3.4 The PNT domain of ELF5. **(a)** Comparison of the Pointed domain of human and mouse ELF5 with those of known members of the ETS gene family. The alignment was generated by using CLUSTAL W (Thompson *et al.*, 1994) with the default settings, and the result was subsequently adjusted manually. The ETS factors examined are labelled on the left, where 'h' denotes human, 'm' mouse and 'd' *Drosophila*. The Pointed domain consensus sequence is a list of the amino acids most often conserved between ETS family members (Graves and Petersen, 1998). Shading denotes amino acid identity with human ELF5, and the percent identity of each pointed domain is indicated on the right. Tentative α-helices (H) predicted by NMR are indicated above the consensus sequence by blue boxes.

b

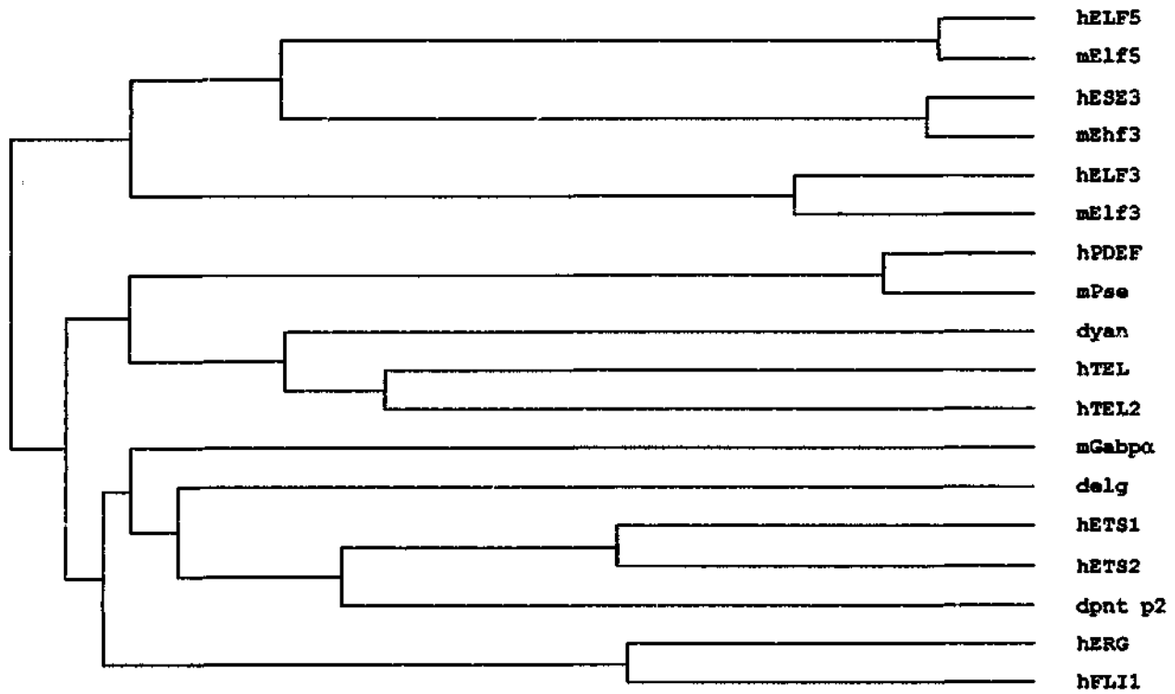


Figure 3.4 The PNT domain of ELF5. **(b)** Dendrogram showing the Pointed domain family tree, based on amino acid sequence similarity using CLUSTAL W method.

Therefore, ETS proteins may acquire their specific biological functions due to their differential DNA-binding affinity toward different targets.

The hallmark of ETS factors ability to bind DNA sites containing a GGAA core in a sequence-specific manner is shared by ELF5. A recombinant mouse Elf5 HIS-tag protein of approximately 30 kDa, expressed in *E. coli* and purified by metal-affinity chromatography (Figure 3.5 a, lane 4), displayed strong binding to a consensus ETS binding site, as analyzed by electrophoretic mobility shift assay (EMSA) (Figure 3.5 b). Mouse Elf5 bound the E74 oligonucleotide (containing a GGAA core) (lane 1), but not to the E74m1 oligonucleotide (containing a mutated AGAA core) (lane 2). The first G residue of the core has been demonstrated to be a physical point of DNA contact for mouse Ets1, and consequently is essential for DNA binding (Fisher *et al.*, 1991; Nye *et al.*, 1992). Thus, mouse Elf5 displayed sequence-specific binding to a consensus ETS binding site and that binding was disrupted by a mutation known to similarly affect other ETS family members. These results were confirmed through competition analysis. The Elf5-E74 complex (lane 3) was efficiently competed by the addition of a 100-fold excess of unlabelled E74 (lane 4), but not by E74m1 (lane 5).

Elf5 also displayed sequence-specific binding to different ETS binding sequences, and did so with differential affinity (Figure 3.5b). The Elf5-E74 complex (lane 3) was competed off by the addition of oligonucleotides containing ETS sites from the *GM-CSF* promoter (lane 6), Erb-B2 promoter (lane 7) and *moloney sarcoma virus* (*MSV*) *long terminal repeat* (*LTR*) (lane 8). The relative ability of Elf5 to bind these sequences occurred in the order: E74 > ErbB2 > MSV > GM-CSF. Elf5 binding did not appear to be competed off by an oligonucleotide containing a consensus AP1 binding site (lane 9). Ets1 binding to E74 was used as a positive control (Lane 10).

3.2.5 Mouse Elf5 acts as a transcriptional activator

Transcription factors can activate or repress target gene expression either directly or via cofactors. Studies have shown that there are both transcriptional activators and repressors in the ETS family of transcription factors. It was therefore of interest to evaluate the ability of mouse Elf5 to function as a transcriptional activator or repressor.

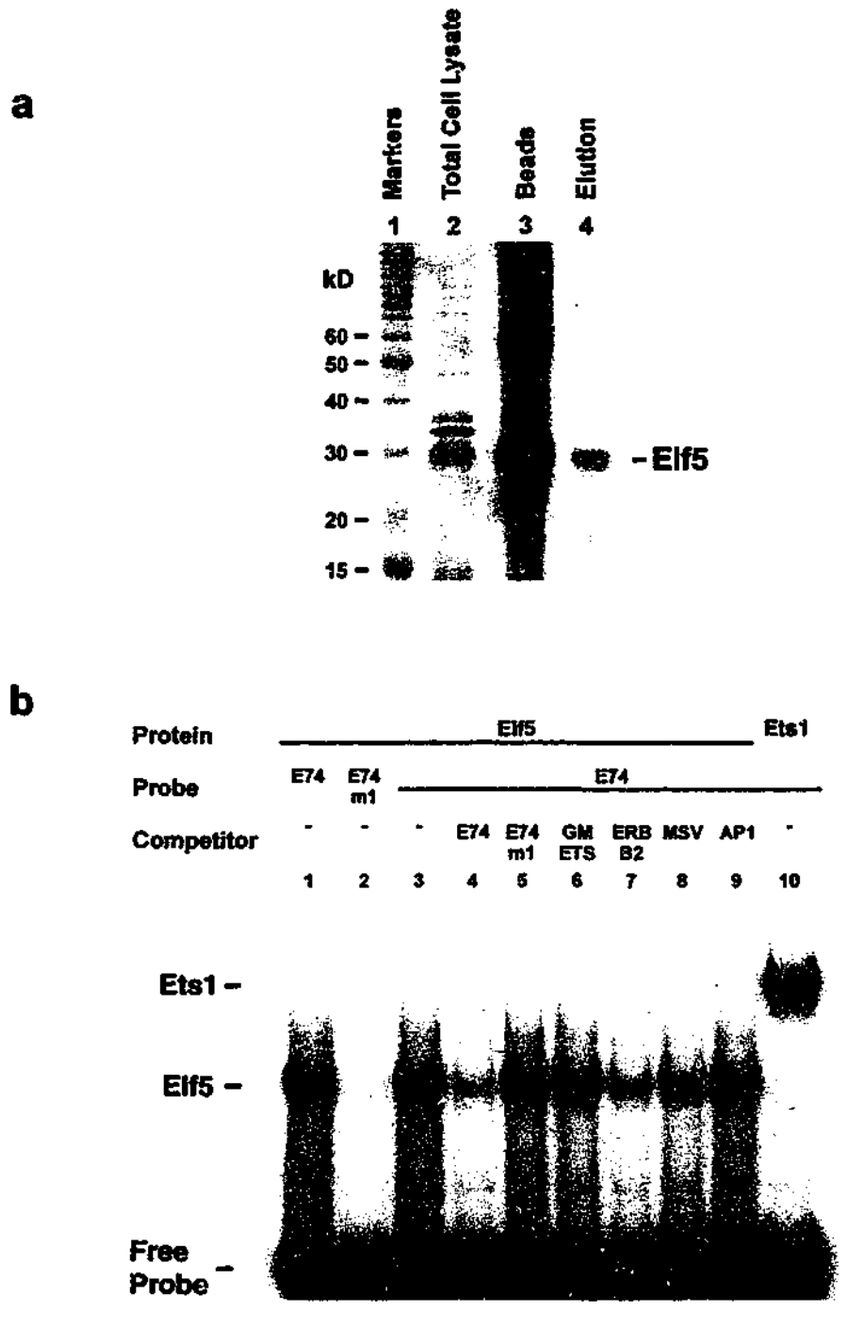


Figure 3.5 Elf5 can recognize ETS-binding motif. **(a)** His-tagged Elf5 recombinant protein, present in *E. coli* lysates (lane 2), was purified by metal-affinity chromatography to approximately 90% (lane 3) and eluted with imadazole (lane 4). **(b)** Elf5 binds to consensus ETS binding sequences. Specific DNA binding of Elf5 was analysed by electrophoretic mobility shift assay (EMSA), using labelled double-stranded oligonucleotides as probes. E74 contains a consensus binding site for ETS family members (lane 1). E74m1 is a mutant oligonucleotide based on E74, but with the core GGAA replaced by AGAA (lane 2). Binding to other consensus ETS sites was analyzed by the ability of a 100-fold excess of unlabelled double-stranded oligonucleotide to compete with E74 for Elf5 binding. GMETS contains an ETS binding site from the human GM-CSF promoter (lane 6). ERBB2 contains an ETS binding site from the human *erbB2*/HER2 promoter (lane 7). MSV contains an ETS binding site present in the long terminal repeat of the Moloney sarcoma virus (lane 8). AP1 contains a consensus AP1 binding site used as a negative control. Elf5-DNA complexes are marked. Binding of ETS1 to E74 was used as a positive control (lane 10).

A reporter construct, containing chloramphenicol acetyl-transferase (CAT) driven by a minimal TK promoter in addition to multiple ETS/AP1 binding sites (from the polyomavirus enhancer), was co-transfected into COS cells together with the mouse Elf5 expression construct (Figure 3.6). Analysis of CAT activities revealed that Elf5 expression resulted in an average five-fold transactivation of the reporter. Furthermore, this transactivation was inhibited by addition of an anti-sense *Elf5* mRNA expression vector, indicating that Elf5 transactivation was due specifically to the product translated from the sense construct.

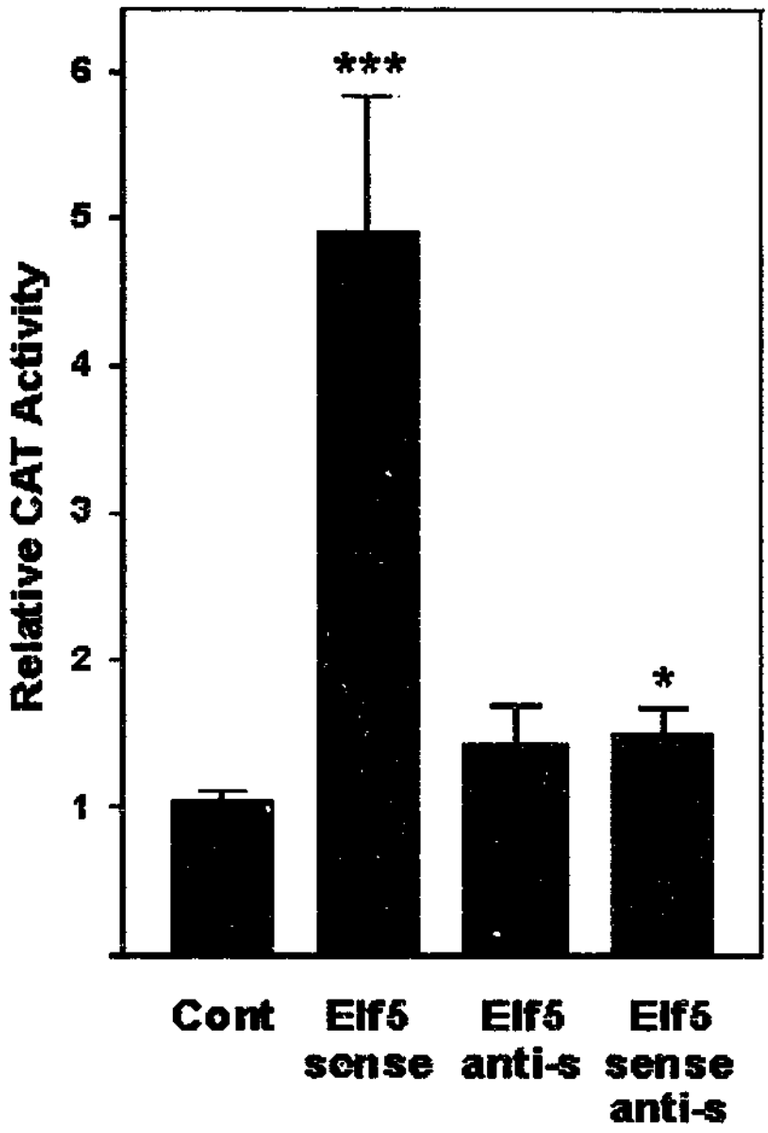


Figure 3.6 Transactivation by Elf5. COS cells were co-transfected with CAT reporter and Elf5 expression constructs. Transcription of the CAT gene was driven by the thymidine kinase (tk) minimal promoter with five copies of the polyomavirus enhancer inserted upstream (p5Xpoly). The polyomavirus enhancer contains adjacent ETS and AP1 binding sites. The Elf5 sense construct (pBOSElf5s) was designed to express Elf5 protein, and the *Elf5* anti-sense construct (pBOSElf5as) to produce anti-sense transcripts. In the absence of expression construct, the equivalent amount of base vector (pEFBOS) was co-transfected. COS cells were processed for CAT assays and the results of at least 4 replicates are shown as the mean with standard error of the mean (SEM) bars. Statistically significant results are indicated by asterisks. A single asterisk indicates moderate significance ($0.05>p>0.01$) and triple asterisks indicate very high significance ($p<0.001$).

3.3 Discussion

Two predominant mouse *Elf5* mRNA species (*Elf5-a* and *Elf5-b*), identified in a mouse placenta cDNA library, appear to arise from utilization of two distinct polyadenylation sites. The larger of these two transcripts, *Elf5-a*, incorporates multiple A/T-rich signals that have been implicated in processes of increased mRNA turnover (Savant-Bhonsale and Cleveland, 1992; Akashi *et al.*, 1994). Two human *ELF5* isoforms were also identified in a human placenta cDNA library. Although the cDNAs appear to share the same 5′ and 3′ UTRs, they differ in their ORFs due to an alternative splicing event in *ELF5-b*. Consequently, *ELF5-b* is 285 bp shorter than *ELF5-a* in the coding region. A third human *ELF5* isoform, *ESE-2a*, was identified by another group in a human prostate cDNA library (Oettgen *et al.*, 1999). The predicted protein sequence of ESE-2a has a 10 amino acid extension at the amino terminus, but whether these 10 amino acids encode any specific function is unknown since no similarity to any known protein domain or other proteins can be detected. Interestingly, the murine counterparts for human ELF5-b and ESE-2a were not detected.

The ETS domains are composed of three α-helices (H) and four β-strands (S) arranged in the order of H1-S1-S2-H2-H3-S3-S4. Similar tertiary structures of the ETS domains are found in hFLI1, mEts1, hETS1 and hPU.1 (Donaldson *et al.*, 1994; Liang *et al.*, 1994; Werner *et al.*, 1995; Donaldson *et al.*, 1996; Kodandapani *et al.*, 1996; Werner *et al.*, 1997a & b). The amino acid residues within the α-helices and β-strands of the ETS domain are highly conserved among the ETS family members. These highly conserved residues including three tryptophan residues are thought to be structurally critical for DNA binding (Karim *et al.*, 1990; Gutman and Wasylyk *et al.*, 1990; Wang *et al.*, 1992; Wasylyk *et al.*, 1992).

Amino acids that are highly conserved amongst ETS family members (Janknecht and Nordheim, 1993) are also invariant in ELF5 (29/48), including the three tryptophan residues (Figure 3.3a). In particular, all but one amino acid residue in the three α-helices and four β-strands are either invariant or conservative substitutes. The two arginines in H3, that are invariant in the ETS family and provide base-specific contacts to the two guanine residues of the GGA core, are also invariant in ELF5. Furthermore, two key lysine residues, one in the loop between S3 and S4, the other within the turn between H2 and H3, provide contacts

with the phosphate backbone. The first residue is invariant in ELF5, and the second is found to be a conservative substitution. Therefore, we expect ELF5 to recognize and bind to the consensus GGA core like all other ETS family members. However, outside of these highly conserved residues, ELF5 diverges considerably and it is these divergent regions of the ETS domain that are thought to provide the sequence-specific DNA-binding specificity of each individual ETS family member (Graves and Petersen, 1998).

Although human ELF5 exhibits the highest degree of amino acid sequence identity to human ELF3 (65%) and ESE3 (69%) in the ETS domain, the relative degree of identity in this region is not as high as that observed between other ETS subfamily members, where the identity within the ETS domain is 80-90% (Dalton and Treisman, 1992; Price *et al.*, 1995; Oettgen *et al.*, 1996; Kas *et al.*, 2000). Even within its own subgroup, the ELF5 ETS binding domain is the most divergent, with ELF3 and ESE3 sharing 86% identity. In addition to its function in DNA-binding, the ETS domain is also involved in protein-protein interactions with numerous other transcription factors, which is thought to confer another layer of regulation on the activity of ETS proteins. For instance, ERM binds Jun via the ETS domain and the C-terminal extension (Nakae *et al.*, 1995). Several other ETS proteins, including Ets1, Elf1, PU.1 and Fli1, have been shown to interact with the basic domain of the Jun transcription factor through the ETS domain *in vitro* and in activated human T cells. This interaction induces the transcriptional activity of enhancer elements containing adjacent ETS and AP1 binding sites (Bassuk and Leiden, 1995). ETS1 can also associate with PAX-5 via the ETS domain to form a ternary complex (Fitzsimmons *et al.*, 1996; Wheat *et al.*, 1999). Therefore, the relatively low degree of sequence identity of the ELF5 ETS domain to other ETS family members may determine differences in the protein-protein interactions of ELF5.

The PNT domain is also implicated in protein-protein interactions. Some ETS proteins are capable of forming homo- or heterodimers or possibly higher order complexes via their PNT domains (Carroll *et al.*, 1996; Golub *et al.*, 1996; McLean *et al.*, 1996; Jousset *et al.*, 1997; Schultz *et al.*, 1997; Graves and Petersen, 1998; Kyba *et al.*, 1998; Stapleton *et al.*, 1999). For instance, ETS2 can associate with ERG via both the ETS domain and the PNT domain (Basuyaux *et al.*, 1997). ERG and its isoforms are also known to form homodimers or heterodimers with other ETS proteins, and these protein-protein interactions are mediated via the PNT and ETS domains (Carrere *et al.*, 1998). The PNT and ETS domains of TEL are both required to interact with FLI1, and this protein-protein interaction can inactivate a

FLI1-specific promoter (Kwiatkowski *et al.*, 1998). The ETS transcription factor human TEL2 is capable of self-association, presumably via its PNT domain, and it can also heterodimerize with TEL via the PNT domain of TEL (Potter *et al.*, 2000).

The amino acid sequence similarity in the PNT domain of human ELF5/mouse Elf5 is relatively weak compared to other ETS family members, but 22/60 of the highly conserved residues are invariant in ELF5. Within the four α-helices (H2-H5) that are most highly conserved in the PNT domain, 25 out of 31 amino acid residues are either invariant or conservative substitutions (Figure 3.4 a). Therefore, we predict the ELF5 PNT domain to be involved in similar protein-protein interactions as described for other ETS family members. Presumably, analogous to the ETS domain, this relatively low degree of sequence identity within the ELF5 PNT domain determines specificity for ELF5's protein-protein interactions.

Notably, human ELF5-b has a 95 amino acid deletion that removes the four α-helices (H2-H5) in the PNT domain. The function of this human isoform is not known yet, but one can speculate that it may have an opposing effect on ELF5-a and/or ESE-2a. It is possible that the ETS domain in ELF5-b can compete with other human ELF5 isoforms for the ETS binding sites in the promoters and enhancers of various target genes and hence quench or repress their transcriptional regulatory activities. Alternatively, it can associate with other ELF5 isoforms or other transcription factors and regulate downstream target genes in a manner different from the whole protein.

Phylogenetic analysis based upon the amino acid sequence identity in the ETS domain indicates that human ELF5 displays closest similarity to human ELF3 and ESE3, and these ETS proteins are also closely related to the members of the ELF subfamily (ELF1, NERF, MEF and e74). Therefore, we propose that ELF5, ELF3 and ESE3 form part of an extended ELF subfamily within the ETS family of transcription factors. It is noteworthy that the PNT domain is not found in the ELF subfamily members, ELF1, NERF, MEF and e74, and this might indicate that hELF5/mElf5, hELF3/mElf3 and hESE3/mEhf represent a distinctly separate subfamily of ETS molecules. However, regardless of whether the ETS domain or the PNT domain is used to make the alignment, ETS genes containing a PNT domain cluster basically to the same group (Figure 3.3 c & 3.4 b).

Human *ELF5*, *ELF3* and *ESE3* are all expressed in a similar subset of epithelial tissues (Kas *et al.*, 2000). Interestingly, expression of the SPI subfamily of the ETS factors is restricted to the immune system. In the case of ERG subfamily, *ERG*, *FLI1/ERG-B*, and *FEV* are all involved in chromosomal translocations resulting in Ewing's sarcoma (Ida *et al.*, 1995; Peter *et al.*, 1997). These examples possibly represent a functional similarity among the closely related ETS subfamily members.

The recombinant mouse Elf5 protein is capable of binding to several ETS oligonucleotides, containing a GGAA core motif, in a sequence-specific manner (Figure 3.5b). Recombinant mouse Elf5 also displays the ability to transactivate a minimal TK promoter containing consensus ETS binding sites (Figure 3.6). In addition, recombinant human ELF5-a is capable of transactivating the *whey acidic protein* (*WAP*) promoter (Thomas *et al.*, 2000) in mouse mammary epithelial cells (HC-11 cells).

Interestingly, binding of Elf5 to the E74 oligonucleotide is efficiently competed by an oligonucleotide containing an ETS binding site from the *ERBB2* promoter (Figure 3.5 b). This site has previously been shown to bind and be transactivated by recombinant PEA3 (Benz *et al.*, 1997) and ELF3 (ESX) (Chang *et al.*, 1997). PEA3 has also been shown to suppress *ERBB2* overexpression and inhibit tumorigenesis via the ETS binding site in the *ERBB2* promoter (Xing *et al.*, 2000), suggesting that it is other (unidentified) transactivating factor(s), that recognize the same regulatory element in the *ERBB2* promoter, contributing to *ERBB2* expression (Xing *et al.*, 2000). *ERBB2* is expressed in epithelial cells of organs such as breast, intestine, kidney and ovary (Gullick *et al.*, 1987; Kokai *et al.*, 1987; Mori *et al.*, 1989; Press *et al.*, 1990). Overexpression of ERBB2 is found in at least 40% of all non-invasive intraductal breast carcinomas, and its overexpression is also implicated in the genesis of a number of human tumors, including 20-30% of all breast carcinomas, linking ERBB2 to an increased propensity of tumors to invade and metastasize resulting in a poor prognosis for the patient (Singleton and Strickler, 1992; Tripathy and Benz, 1994). The expression pattern of *ELF5* matches closely that of *ErbB2,* with the exception that *ELF5* does not appear to be expressed in the intestine. Thus, ELF5 would appear to be another candidate for regulating the *ERBB2* promoter. The transcriptional activity of ELF5 upon the *ERBB2* promoter, if any, is yet to be determined.

It has been suggested that one of the human ELF5 isoforms, ESE-2a, has a negative regulatory domain within the first 42 amino acids at the amino-terminal end of the protein (Oettgen *et al.*, 1999). *In vitro* DNA-binding studies showed that ESE-2a does not bind to the consensus *Drosophila* e74 ETS binding site, whereas a truncated ESE-2a$_{\Delta 42}$ protein does. The truncated ESE-2a$_{\Delta 42}$ protein was shown to bind to ETS sites in the promoters of several salivary gland and prostate-specific genes (*CRISP-1*, *CRISP-3*, *PSP* and *MP6*) and the prostate and mucous gland-specific *PSP94* gene. In addition, it also binds to the ETS sites in the *SPRR2A* and *Endo-A* genes (Oettgen *et al.*, 1999). A number of ETS proteins, such as SAP1 (Dalton and Treisman, 1992), Elk1 (Rao and Reddy, 1992; Janknecht *et al.*, 1994), Net (Giovane *et al.*, 1994; Lopez *et al.*, 1994; Price *et al.*, 1995; Maira *et al.*, 1996), ERM (Laget *et al.*, 1996), Ets1 and Ets2 (Hagman and Grosschedl, 1992; Lim *et al.*, 1992; Wasylyk *et al.*, 1992; Fisher *et al.*, 1994; Petersen *et al.*, 1995; Jonsen *et al.*, 1996), are also shown to contain sequences that inhibit DNA binding activities of their ETS domains.

Despite its inability to bind to the consensus ETS binding sites, ESE-2a is capable of transactivating the *SPRR2A*, *PSP* and *PSA* promoters *in vitro* (Oettgen *et al.*, 1999). SPRR2A is expressed in the cornified cell envelope in the skin, and is associated with terminal differentiation of keratinocytes (Hohl *et al.*, 1995). PSP is found to be the most abundantly expressed protein in the parotid gland (Ann *et al.*, 1997; Laursen and Hjorth, 1997; Laursen *et al.*, 1998; Svendsen *et al.*, 1998), while PSA has been used widely as a diagnostic indicator for prostate cancer (Ablin, 1997). The transactivating ability of ESE-2a on these promoters suggests that posttranslational modifications, such as phosphorylation, or protein-protein interactions may be involved in overcoming the intramolecular inhibition of the DNA binding activities of ESE-2a (Oettgen *et al.*, 1999).

Chapter 4

ELF5 is an epithelial-specific ETS transcription factor

4.1 Introduction

Gene expressions is a tightly controlled and regulated process during the development of a single–cell embryo into a complicated organism containing a network of different organs and cell types. Transcription factors are the key regulators of the coordinated gene expression during the development (Walker *et al.*, 1983; Gehring, 1987; Levine and Hoey, 1988; Scholer, 1991; Biggs *et al.*, 1992; Dressler and Douglass, 1992; Hoch *et al.*, 1992). Therefore, the expression patterns of transcription factors may provide clues about their *in vivo* biological functions and targets (Kola *et al.*, 1993).

In spite of their wide distribution, many *ETS* genes play important roles in regulating the development and differentiation of the hematopoietic lineages. This is strongly supported by the *in vivo* transgenic and gene-targeting experiments in the mouse (Sharrocks *et al.*, 1997; Graves and Petersen, 1998; Dittmer and Nordheim, 1998). In addition, ETS DNA binding sites are located in the promoters and enhancers of various genes involved in the regulation of the hematopoietic system (Ghysdael and Boureux, 1997; Bassuk and Leiden, 1997). Interestingly, many epithelial-specific genes are regulated by ETS factors, including *Prolactin* (Howard and Maurer, 1994; Bradford *et al.*, 1995), *α–GSU* (Roberson *et al.*, 1995), *MASPIN* (Zou *et al.*, 1994; Zhang *et al.*, 1997b), cytokeratin *EndoA* (Hashido *et al.*, 1991; Fujimura *et al.*, 1994), *SPRR2A* (Oettgen *et al.*, 1997), *Aminopeptidase N* (*APN*) (Olsen *et al.*, 1997), *TRANSGLUTAMINASE 3* (Lee *et al.*, 1996) and *MET* (Gambarotta *et al.*, 1996). However, very few *ETS*-related genes are known to be expressed preferentially in epithelial cells. *ELF3* (*ESX/ESE-1/JEN/ERT*) (Andreoli *et al.*, 1997; Chang *et al.*, 1997;

Choi *et al.*, 1998; Oettgen *et al.*, 1997; Tymms *et al.*, 1997) was the first *ETS* factor identified to have epithelial- and epithelial-cancer-specific expression.

This chapter describes the expression pattern and human chromosomal localization of a novel *ETS* family member, *ELF5* (*E74-Like-Factor 5*). *ELF5* is the second *ETS* transcription factor reported to be expressed exclusively in a subset of epithelial cells, and *ELF5* expression appears to be lost in some epithelial cancer cell lines derived from tissues normally expressing *ELF5*. In addition, the human *ELF5* gene is localized to chromosome 11p13-15, a region that frequently undergoes loss of heterozygosity (LOH) in many types of human cancer, such as breast, lung, kidney and prostate.

4.2 Results

4.2.1 Mouse *Elf5* expression is restricted to tissues rich in epithelial cells

Poly(A)$^+$ mRNA derived from various mouse tissues was analyzed by Northern blot hybridization using the murine *Elf5* cDNA as a probe. A *GAPDH* probe was then used to control for RNA loading.

Analysis of *Elf5* expression in adult mouse tissues revealed that *Elf5*, like mouse *Elf3* (Tymms *et al.*, 1997), has a restricted expression pattern. Expression of two *Elf5* transcripts, *Elf5-a* (2.5 kb) and *Elf5-b* (1.5 kb), are observed in lung (Lu), kidney (Ki), stomach (St), ovary (Ov), tongue (To), bladder (Bl), day 2 pregnant mammary glands (2Ma), day 10 pregnant mammary glands (10Ma), caput epididymus (CA), corpus epididymus (CO) and seminal vesicles (SV), but no expression was observed in liver (Li), heart (He), small intestine (Sm), spleen (Sp), thymus (Th), pancreas (Pa), skeletal muscle (Sk), colon (Co), day 2 pregnant fat (2Fa) or day 10 pregnant fat (10Fa) (Figure 4.1 a). Stromal tissue from day2 (2Fa) and day 10 (10Fa) pregnant mice was used as a control for mammary expression, since the mammary gland contains much stromal tissue. A single transcript was observed in brain (arrow – approximately 2.1 kb), but of a different size to either of the two *Elf5* transcripts in other organs. However, its appearance was variable in repeated experiments with adult tissues (data not shown) and at different developmental stages (Figure 4.1 b & c). Pending further analysis, it appears likely that this brain-specific transcript is derived from cross hybridization to another highly expressed *Ets* family member, rather than *Elf5*.

4.2.2 Mouse *Elf5* expression is restricted to tissues rich in epithelial cells during development

Some ETS members are implicated in developmental processes. Therefore it was of interest to characterize the expression pattern of *Elf5* during mouse development. We examined the expression of *Elf5* in the neonatal mouse (Figure 4.1 b) and during embryogenesis on days 16, 17 and 19 (Figure 4.1 c), and observed a similar expression pattern compared to that of the adult. However, at day 16 of embryogenesis, low levels of *Elf5* expression were detected in the brain (regular sized transcripts) and small intestine, in addition to the expression pattern observed in the adult.

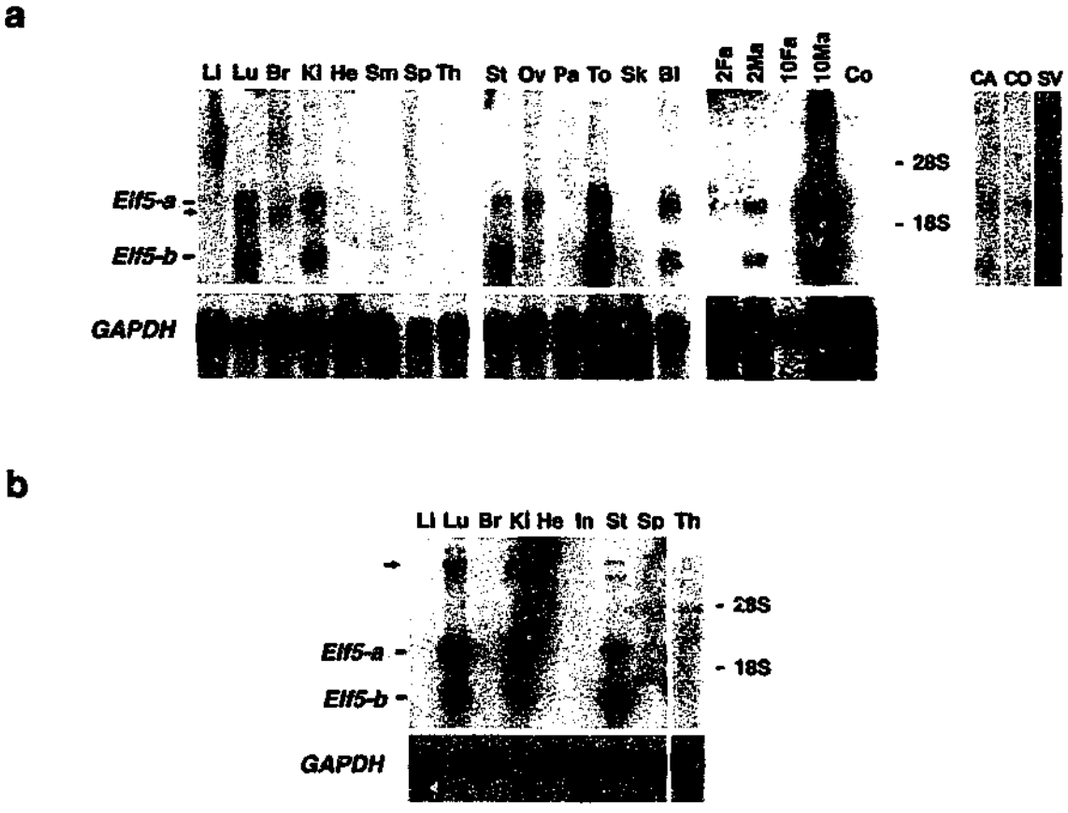


Figure 4.1 *Elf5* expression in mouse tissues. Position of size markers are indicated. *Elf5-a* and *Elf5-b* transcripts are indicated. **(a)** *Elf5* expression is restricted to adult mouse tissues rich in epithelial cells. Northern blots probed with murine *Elf5* cDNA (top panels) and GAPDH cDNA (lower panels). Abbreviations; Li: liver; Lu: lung; Br: brain; Ki: kidney; He: heart; Sm: small intestine; Sp: spleen; Th: thymus; St: stomach; Ov: ovary; Pa: pancreas; To: tongue; Sk: skeletal muscle; Bl: bladder; 2Fa: day 2 pregnant fat; 2Ma: day 2 pregnant mammary gland; 10Fa: day 10 pregnant fat; 10Ma: day 10 pregnant mammary gland; Co: colon; CA: caput epididymus; CO: corpus epididymus; SV: seminal vesicles. Arrow indicates position of brain-specific transcript (see text). **(b)** *Elf5* expression is restricted to neonate mouse tissues rich in epithelial cells. Northern analysis as above, but using RNA from day 1 neonate mouse tissues. Additional abbreviation; In: intestine. Arrow indicates position of large transcript (see text).

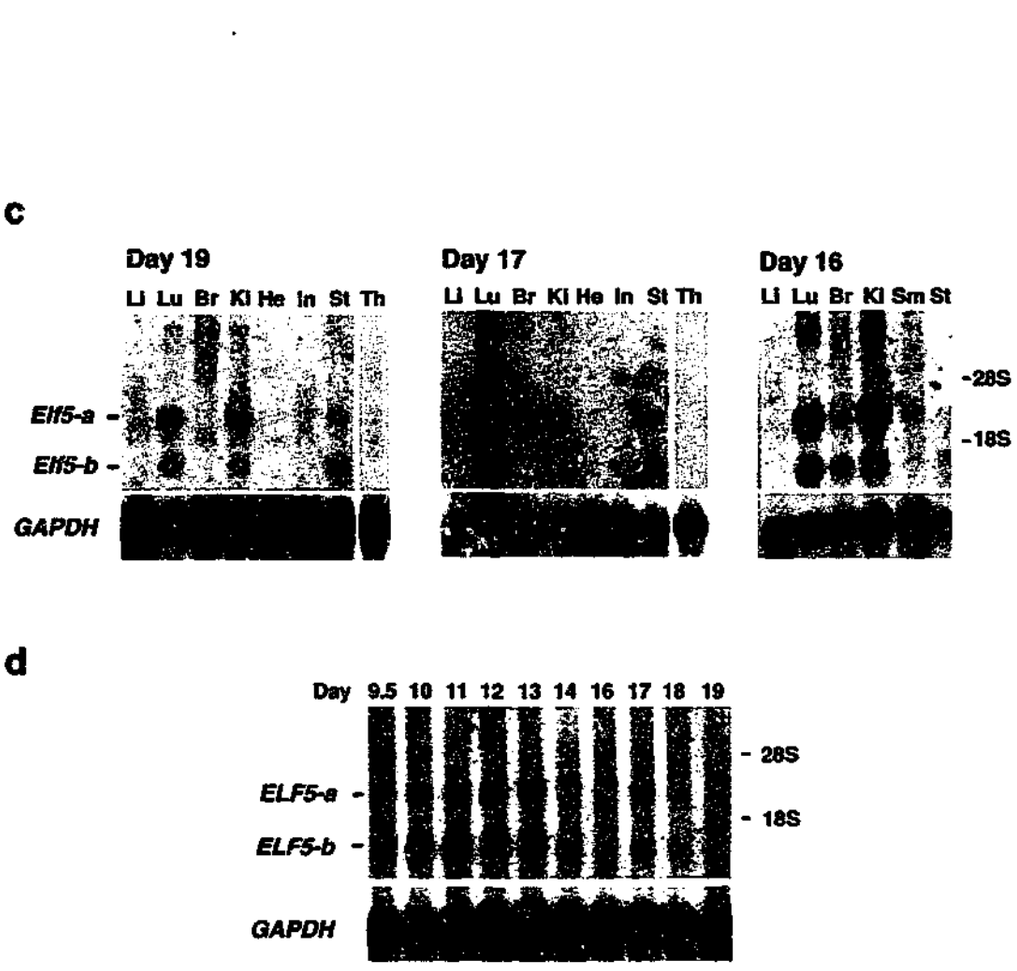


Figure 4.1 *Elf5* expression in mouse tissues. Position of size markers are indicated. *Elf5-a* and *Elf5-b* transcripts are indicated. (**c**) *Elf5* expression is restricted to embryonic tissues rich in epithelial cells. Northern analysis as in **a**, but using RNA from day 16, 17 and 19 embryonic tissues. (**d**) *Elf5* expression in placenta during mouse embryogenesis. Northern analysis as in **a**, but using RNA from day 9.5 to day 19 placental tissues as indicated. Additional abbreviation; Pl: placenta.

Placental expression of *Elf5* displays an interesting pattern during embryogenesis (Figure 4.1 d). Both transcripts are increasingly expressed from day 9.5 to day 13 before an overall decrease is observed from day 14 to day 19, although some expression is observed at day 17.

4.2.3 The two *Elf5* isoforms, *Elf5-a* and *Elf5-b*, are differenially expressed in different tissues

The two predominant *Elf5* mRNA transcripts are observed in variable ratios in different tissues, suggesting that the polyadenylation sites may be utilized differentially, or that the two transcripts are subject to differential degradation. *Elf5-a* is expressed more strongly in neonatal and embryonic lung and kidney (Figures 4.1 b & c), and adult ovary (Figure 4.1 a), compared to *Elf5-b*. In contrast, *Elf5-b* was stronger in adult tongue (Figure 4.1 a), and in all developmental stages of the stomach (Figures 4.1 a-c), compared to *Elf5-a*. However, since these two transcripts would appear to differ only in their 3′ UTR sequences, they are likely to produce the same translation product. In some RNA samples an additional larger (>10 kb) transcript is variably observed. Despite the fact that the RNA samples were poly(A)$^+$ selected, these bands could represent unspliced *Elf5* transcripts or genomic contamination.

4.2.4 Human *ELF5* expression is restricted to tissues rich in epithelial cells

Expression of *ELF5* in adult human organs was analyzed by Northern blot of poly(A)$^+$ mRNA probed with the human *ELF5* cDNA (Figure 4.2 a), a *β-ACTIN* cDNA probe was used to control for RNA loading. A single transcript of approximately 2.5 kb is strongly expressed in kidney (Ki) and prostate (Pr). However, much longer exposures of blots (data not shown) demonstrated barely detectable expression of *ELF5* in placenta (Pl) and lung (Lu). Furthermore, *ELF5* was cloned from human lung and placenta cDNA libraries, confirming that it is expressed in these tissues, albeit at very low levels. Although this expression pattern resembles that observed in the mouse, it is interesting that the human *ELF5* gene does not appear to utilize alternative polyadenylation sites, having only a single transcript.

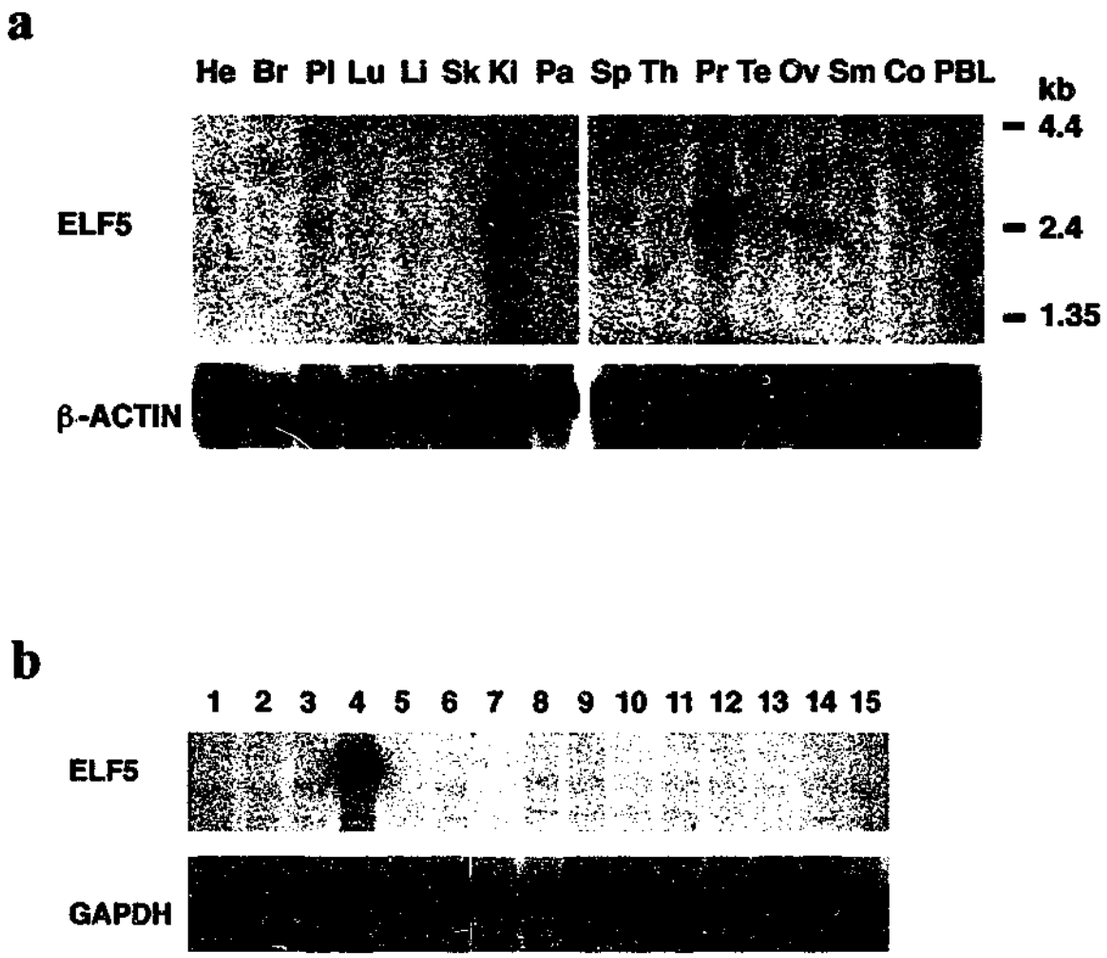


Figure 4.2 *ELF5* expression in human tissues and cell lines. **(a)** *ELF5* expression is restricted to adult human tissues rich in epithelial cells. Northern blots (Clontech) probed with human *ELF5* cDNA (top panels) and β-ACTIN cDNA (lower panels). The ELF5 transcript is indicated. Position of size markers are indicated. Abbreviations; He: heart; Br: brain; Pl: placenta; Lu: lung; Li: liver; Sk: skeletal muscle; Ki: kidney; Pa: pancreas; Sp: spleen; Th: thymus; Pr: prostate; Te: testis; Ov: ovary; Sm: small intestine; Co: colon mucosa; PBL: peripheral blood lymphocytes. **(b)** *ELF5* expression is restricted to human cancer cell lines of epithelial origin. RNase protection analysis of *ELF5* and GAPDH in cell lines; 1: CaOv-3 (ovarian carcinoma); 2: BT-549 (breast ductal carcinoma); 3: ZR- 75- 1 (breast carcinoma); 4: T47D (breast ductal carcinoma, progesterone sensitive); 5: 786- O (renal adenocarcinoma); 6: SK-HEP-1 (liver adenocarcinoma); 7: A549 (lung adenocarcinoma); 8: CCL32SK (primary fibroblast); 9: MEL28 (melanoma); 10: WISH (amnion carcinoma); 11: Jurkat (T cell leukemia); 12: DU145 (prostate carcinoma); 13: PC3 (prostate carcinoma); 14: HEC-1 (endometrium carcinoma); 15: K562 (erythroid leukemia).

4.2.5 Human *ELF5* expression is not detectable in a subset of epithelial cancer cells

Given the restricted expression pattern of *ELF5* in tissues commonly giving rise to carcinomas, it was of interest to examine *ELF5* expression in human cancers. A panel of cancer cell lines derived from carcinomas of the ovary (CaOv-3), breast (BT-549, ZR-75-1, T47D), kidney (786-O), liver (SK-HEP-1), lung (A549), amnion (WISH), prostate (DU145, PC3) and endometrium (HEC-1), in addition to melanoma (MEL28), T cell leukaemia (Jurkat) and erythroid leukaemia (K562) cell lines, were analyzed for *ELF5* expression by RNAse protection assay (Figure 4.2 b). A primary fibroblast cell line (CCL32SK) was also included as a sample of non-transformed cells. Of all these cell lines, only T47D, a progesterone sensitive ductal breast carcinoma, was observed to express *ELF5*.

A further expression study was performed on eight breast (MCF-7, T47D, PMC42, ADR, MDA436, MDA435, BT549 and ACC15) and six prostate (LNCaP, C4P36, C4-2, C4-2BP, TSUP11 and PC3P10) cancer cell lines by Northern blot analysis. In addition to T47D, MCF-7 (breast cancer cell line) and C4-2BP (prostate cancer cell line, LNCaP subline) are also found to express lower levels of *ELF5*. The remaining eleven breast and prostate cancer cell lines did not express detectable *ELF5* (Figure 4.2 c).

4.2.6 Southern blot analysis of the human *ELF5* genomic locus in epithelial cancer cells

To evaluate the possibility that lack of *ELF5* expression in carcinoma was due to genomic alterations, a panel of breast and lung carcinoma derived cell lines were analyzed by Southern blot (Figure 4.2 d). *ELF5* gene dosage was compared to that present in DNA from normal human blood (based on the 6.5 kb *BglII* fragment) and controlled by hybridization with a *β-ACTIN* cDNA probe. These results are summarized in the lower panel, where "2" represents a normal allele complement. No evidence was found for allelic loss or gene rearrangement in two of the breast carcinoma cell lines that do not express *ELF5* (BT-549 – lane 2, ZR-75-1 – lane 3). However, of nine lung carcinoma cell lines, evidence for loss of an *ELF5* allele was observed in two (NCI-H358 – lane 8, NCI-H441 – lane 11). Hybridization with an *ELF3* cDNA probe, which is localized to the long arm of chromosome 1 (Tymms *et al.*, 1997), helped to confirm the specific loss of *ELF5* alleles (data not shown, Dr. R.S. Thomas, personal communication). Two other lung carcinoma lines (SK-LU-1 – lane 10, NCI-H661 – lane 13) displayed multiple fragments (shaded

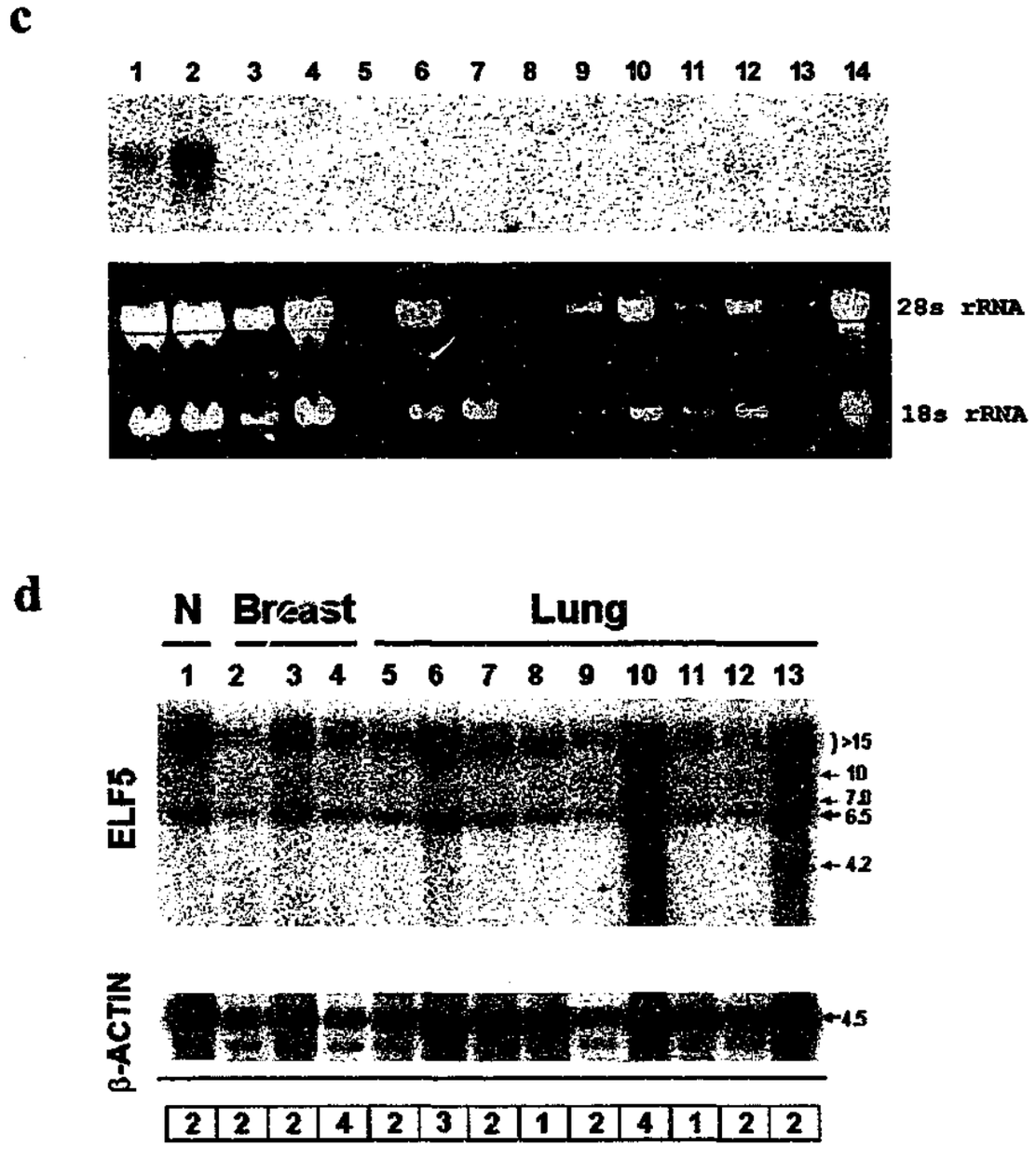


Figure 4.2 (c) *ELF5* expression is restricted to human cancer cell lines of epithelial origin (Northern blot analysis). Northern blots probed with human *ELF5* cDNA (top panels), 28s rRNA and 18s rRNA are used as loading controls (lower panels). The ELF5 transcript is indicated. Lane 1-8: human breast cancer cell lines MCF-7, T47D, PMC42, ADR, MDA436, MDA435, BT549 and Acc15; lane 9-14: human prostate cancer cell lines LNCaP, C4P36, C4-2, C4-2BP, TSUP11 and PC3P10. (d) Southern analysis of *ELF5* in *BglII* digested genomic DNA from human cancer cell lines; 1: normal blood; 2: BT-549 (ductal breast carcinoma); 3: ZR-75-1 (breast carcinoma); 4:T47D (ductal breast carcinoma); 5: NCI-H1299 (large cell lung carcinoma); 6: NCI-H187 (small cell lung carcinoma); 7: NCI-H322 (bronchioalveolar carcinoma); 8: NCI-H358 (bronchioalveolar carcinoma); 9: NCI-H522 (lung adenocarcinoma); 10: SK-LU-1 (lung adenocarcinoma); 11: NCI-H441 (bronchioalveolar carcinoma); 12: NCI-H460 (large cell lung carcinoma); 13: NCI-H661 (large cell lung carcinoma).

arrows) in addition to those observed in normal DNA (solid arrows), possibly indicating that at least one *ELF5* allele has been rearranged in these lines. Confirmation of rearrangement, rather than restriction fragment length polymorphism (RFLP), was made by additional restriction digests (data not shown, Dr. R.S. Thomas, personal communication). Some cell lines appeared to have amplification or additional copies of the *ELF5* gene. One of these, T47D (lane 4), expresses *ELF5*, and another, SK-LU-1 (lane 10), appears to have rearranged alleles.

4.2.7 Human *ELF5* is localized to chromosomal 11p13-15, a region implicated in cancer

TUBE NOS	CELL LINES	PCR RESULT	TUBE NOS	CELL LINES	PCR RESULT	TUBE NOS	CELL LINES	PCR RESULT	TUBE NOS	CELL LINES	PCR RESULT
1	4G6	-	24	4S3	-	49	4P11	-	75	4DD5	-
2	4B3	-	25	4I4	-	51	4K5	-	76	4S12	+
3	4E.6	-	26	4N3	-	53	4H9	+	77	4F6	-
4	4F13	+	27	4O5	-	54	4J9	+	78	4L3	-
5	4P9	-	28	4L6	-	55	4M5	-	79	4S10	-
6	4E.2	+	29	4T11	-	56	4F7	+	80	4R2	+
7	4Z11	+	31	4R1	-	57	4Y4	-	81	4V2	+
8	4Z9	-	32	4C3	-	58	4CC8	-	82	4J2	-
9	4BB6	-	33	4M4	-	59	4Z6	-	83	4R3	+
10	4D1	-	34	4W1	+	60	4K7	-	84	4V7	-
11	4E.4	+	35	4G7	+	62	4R10	-	85	4G5	+
12	4Y9	-	36	4L4	-	63	4Z12	-	86	4V8	-
13	4BB1	-	37	4N7	-	64	4AA5	-	87	4R5	-
15	4N6	-	38	4I1	-	65	4T3	-	88	4BB8	+
16	4T4	-	39	4R6	-	66	4D7	-	90	4S6	-
17	4U1	+	41	4N5	-	67	4Q4	-	91	4K12	-
18	4N12	-	42	4G11	-	68	4P2	-	92	4V3	+
19	4AA7	-	43	4H12	-	70	4K9	+	93	4E.11	-
20	4C11	-	45	4G1	+	71	4A4	+	94	A23	-
21	4T10	-	46	4DD8	-	72	4Y8	-	95	HFL	+
22	4DD2	-	47	4H1	-	73	4Q2	-			
23	4Z5	-	48	4J5	-	74	4H8	-			

Table 4.1 Human chromosomal localization of *ELF5* by PCR using gene-specific primers and the Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre). '+' represents a positive PCR result while '-' represents a negative PCR result.

The human chromosomal localization of *ELF5* was performed by PCR, using gene-specific primers and the Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre). With these primers, a single product of the expected size (234 bp) was amplified from total human DNA. The PCR reactions were then performed separately for each of the individual hybrids. The amplification results (Table 4.1) from the 93 hybrids were submitted to the Radiation Hybrid Mapping server at Whitehead Institute/MIT Center for Genome Research for analysis. The result, provided by the Radiation Hybrid Mapping server, demonstrated that *ELF5* is localized to human chromosome 11. The markers most tightly linked to *ELF5* are D11S3990 (6.5 cR) and D11S3998 (15.9 cR) (lod score > 3.0,. These markers are located in the region 11p13-15 (Figure 4.3 a) which frequently undergoes loss of heterozygosity (LOH) in several types of carcinoma (Iizuka *et al.*, 1995; Baffa *et al.*, 1996; Hirose *et al.*, 1996; Wilson *et al.*, 1996; Dahiya *et al.*, 1997; Kawana *et al.*, 1997; Lichy *et al.*, 1998).

The human chromosomal localization of *ELF5* by fluoresence *in situ* hybridization (FISH), using the *ELF5* cDNA as a probe, also assigned human *ELF5* gene to chromosome 11p13 (Figure 4.3 b, Dr. Erica Woollatt, personal communication). The genomic localization and gene symbol was assigned on 27th February 1998 (GDB Accession ID:9835139).

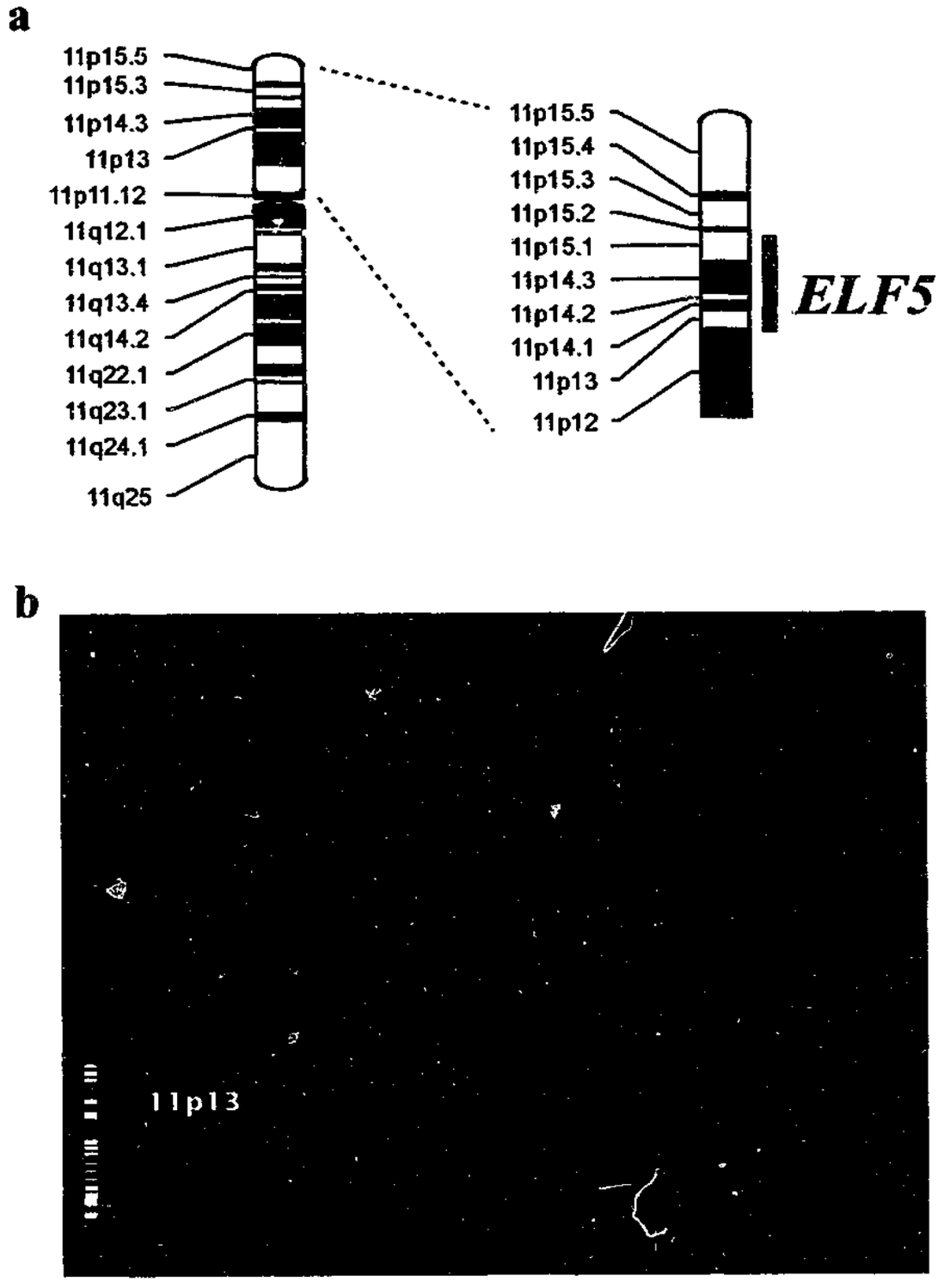


Figure 4.3 Chromosomal localization of the human *ELF5* gene. **(a)** Polymerase Chain Reaction (PCR). Human chromosomal localization of ELF5 was performed by PCR using gene-specific primers and the Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre). Diagram based upon PCR results showing localization of *ELF5* within chromosome 11. **(b)** Fluorescence *in situ* hybridizations (FISH). Computer enhanced image of metaphase preparations, showing *in situ* hybridization of biotinylated probe for *ELF5*. Hybridization sites are indicated with arrows. The FISH analysis was performed by Dr. Woollatt.

4.3 Discussion

Expression of *Elf5* in the adult mouse appears to be restricted to lung, kidney, stomach, ovary, tongue, bladder, epididymus, seminal vesicles and mammary gland. In the developmental stages examined, this expression pattern remains basically unchanged from day 16 of embryogenesis to adult (Figure 4.1 a-c). *Elf5* is also expressed in the placenta during embryogenesis (Figure 4.1 d). Two predominant mouse *Elf5* mRNA species (*Elf5-a* and *Elf5-b*), identified in a mouse placenta cDNA library, appear to arise from the utilization of two distinct polyadenylation sites. The larger of these two transcripts, *Elf5-a*, incorporates multiple A/T rich signals that have been implicated in processes of increased mRNA turnover (Savant-Bhonsale and Cleveland, 1992; Akashi *et al.*, 1994). Perhaps these signals are responsible for the variability of *Elf5-a*:*Elf5-b* ratios in different tissues, particularly the relatively lower proportion of *Elf5-a* observed in stomach.

Expression of *ELF5* in human tissues appears to be even more restricted than that for the mouse, with expression observed only in kidney and prostate (which was not examined in the mouse) (Figure 4.2 a). Inconsistencies between the mouse and human data involve the relatively low *ELF5* expression observed in human lung, ovary and placenta compared to that of mouse. *Elf5* expression in the mouse placenta decreases with advancing developmental time and is completely absent just prior to birth. This may explain the absence of *ELF5* in human placenta by Northern blot since the RNA was presumably obtained from a term placenta. Similarly, we cannot discount the possibility that *ELF5* expression varies with age, or even perhaps is variable within different parts of the lung and ovary, and that these regions were not present in our analysis.

In contrast, human ELF5 expression was found in lung and placenta in two other studies (Oettgen *et al.*, 1999; Kas *et al.*, 2000). In addition to the above tissues, salivary gland, mammary gland, trachea, fetal kidney and fetal lung were also found to express human *ELF5*. Therefore, *ELF5* expression may be limited to glandular epithelium since these tissues are all enriched with this cell type.

That three *ELF5* isoforms exist is not apparent by Northern blot. These isoforms must therefore exist as mRNA species of similar size. Oettgen *et al* (1999) demonstrated differential expression of the isoforms *ESE-2a* and *ESE-2b* in a limited set of tissues.

Interestingly, human prostate expresses both *ELF5-a* (*ESE-2b*) and *ESE-2a*, while kidney only expresses *ESE-2a* (Oettgen *et al.*, 1999).

It is noteworthy that the coding sequence differences observed in cDNA cloning experiments were not detectable by Northern blot analysis. This may due to other possible differences in the 5′ UTR and 3′ UTR sequences resulting in mRNA transcripts of similar size. Further expression studies targeting the coding sequence differences should clarify the expression pattern of all three human ELF5 isoforms.

Most *ETS* genes appear to be functional, or are at least expressed, during hematopoiesis (Bhat *et al.*, 1989; Bhat *et al.*, 1990; Kola *et al.*, 1993). However, human *ELF3* and mouse *Elf3* appear to be examples of epithelial-specific ETS genes (Andreoli *et al.*, 1997; Chang *et al.*, 1997; Oettgen *et al.*, 1997; Tymms *et al.*, 1997). h*ELF5* and m*Elf5*, similar to h*ELF3* and m*Elf3*, are not expressed in hematopoietic compartments, such as the thymus and spleen, but are expressed in organs such as lung, kidney, stomach and prostate. Together with the recently isolated h*ESE3*/m*Ehf* and h*PDEF*/m*Pse*, these novel transcription factors represent a new subclass of ETS family that have their expression restricted to epithelial-type organs, and consequently might perform very specialized functions in these organs. SPI is the only other ETS subfamily that has a highly restricted expression pattern, being found solely in the immune system.

Of these epithelial-specific ETS factors, h*PDEF*/m*Pse* has the most restricted expression pattern (Oettgen *et al.*, 2000; Yamada *et al.*, 2000). In contrast, h*ELF3*/m*Elf3* exhibits a widespread expression pattern in different sub-types of epithelia, including the simple epithelia of the gastrointestinal tract, the glandular epithelia of the mammary gland, uterus and prostate, and the stratified epithelia of the tongue (Andreoli *et al.*, 1997; Chang *et al.*, 1997; Oettgen *et al.*, 1997; Tymms *et al.*, 1997; Ms. A. Ng personal communication). The expression patterns of *ELF5* and *ESE3* are remarkably similar and restricted to a subset of epithelial cells, particularly the glandular epithelium (Oettgen *et al.*, 1999; Kas *et al.*, 2000). It is interesting to note that *ELF5, ELF3* and *ESE3* are all expressed in a similar subset of epithelial tissues, and also share considerable sequence similarity. However, several differences between *ELF5, ELF3* and *ESE3* expression were observed. *ELF3* is expressed strongly in the small intestine and liver, whereas *ELF5* and *ESE3* do not appear to be

expressed in these organs at all. Both *ELF3* and *ESE3* are expressed in colon and pancreas, but *ELF5* is not.

ELF3 has been demonstrated to be expressed in an epithelial-specific manner, both in normal and transformed breast epithelium (Chang *et al.*, 1997) and in lung carcinoma cell lines (Tymms *et al.*, 1997). We have examined *ELF5* expression in a panel of human carcinoma cell lines by RNAse protection assay and found expression in only one of eleven tested (T47D - a ductal carcinoma of the breast) (Figure 4.2 b). A further study of eight breast and six prostate cancer cell lines by Northern blot analysis found low levels of *ELF5* expression in two more human cancer cell lines, MCF-7 (breast cancer cell line) and C4-2BP (prostate cancer cell line, LNCaP subline) (Figure 4.2 c). It is interesting to note that most cell lines derived from lung, renal, prostate, ovarian and breast cancers lacked *ELF5* expression, even though the organs themselves express this gene. Furthermore, none of the other epithelial-derived carcinoma cell lines, including HeLa (cervical carcinoma), A431 (vulvar carcinoma) and C-33A (squamous carcinoma of the cervix) expressed *ELF5* (Oettgen *et al.*, 1999). An analysis of the genomic DNA from some of the carcinoma cell lines was not able to clarify why *ELF5* expression was not observed in two of the three breast carcinoma lines, but four of nine lung carcinoma lines displayed evidence for either loss of an allele or gene rearrangement. Interestingly, the rearrangements observed appeared to be identical in both independently derived cell lines, suggesting a common mechanism for loss of function. All lines appear to contain at least one normal *ELF5* allele, but sequence analysis may reveal smaller mutations that adversely influence expression of *ELF5*.

The human *ELF5* gene is localized to chromosome 11p13-15, a region that frequently undergoes loss of heterozygosity (LOH) in many types of cancer. Examples include ductal breast carcinoma (Lichy *et al.*, 1998), lung carcinoma (Iizuka *et al.*, 1995), rhabdoid tumor of the kidney (Hirose *et al.*, 1996), prostate carcinoma (Dahiya *et al.*, 1997; Kawana *et al.*, 1997), gastric carcinoma (Baffa *et al.*, 1996), and ovarian carcinoma (Wilson *et al.*, 1996). *ELF5* was independently assigned to the WAGR syndrome (wilms tumor, aniridia, genito-urinary anomalies and mental retardation) deletion region (Gawin *et al.*, 1995) on chromosome 11p14.1 (Kas *et al.*, 2000). This region is believed to harbour several tumor suppressor genes (Iizuka *et al.*, 1995; Zenklusen *et al.*, 1995; Baffa *et al.*, 1996; Feinberg, 1996; Ichikawa *et al.*, 1996; Coleman *et al.*, 1997; Gao *et al.*, 1997), based upon both LOH data and the ability to inhibit the tumorigenicity in chemically-induced murine squamous

cell carcinomas upon introduction of human chromosome 11 (Zenklusen *et al.*, 1995). It is interesting to note that other ETS genes implicated in tumorigenesis are thought to be dominant oncogenes, activated by translocation, amplification or viral insertion, rather than recessive tumor suppressors. Superficially, ELF5 appears to function similarly to other ETS family members (DNA binding and transactivation), but it is intriguing that *ELF5* expression appears to be lost in many cancer cell lines, a subset of which were found to have lost an allele or have rearrangement of the *ELF5* gene. A preliminary examination of *ELF5* expression indicates that *ELF5* is not detectable in a number of primary breast carcinomas, although it is strongly expressed in adjacent normal epithelium (data not shown, Dr. D. Venter, personal communication). Thus, it may be possible for an ETS gene, such as *ELF5*, to have tumor suppressor properties and to be lost in certain cancers. Indeed, it appears that ETS1 may function to suppress tumorigenicity of colon cancer cells, whereas it is usually considered to be an oncogene (Suzuki *et al.*, 1995).

Recently, Xing *et al.*, 2000 demonstrated that overexpression of an ETS protein, PEA3, in *ERBB2/HER2/neu* overexpressing cancer cells suppressed the cell growth and tumor development both *in vitro* and *in vivo*. However, this suppression effect was not observed in cancer cells with a low level of *ERBB2/HER2/neu* expression. In addition, it was shown that the *ERBB2/HER2/neu* promoter activity is downregulated by PEA3 via the ETS binding site within the promoter. It is interesting to observe *ELF5* expression in two human breast cancer cell lines, MCF-7 and T47D, that have low levels of *ERBB2/HER2/neu* (Chang *et al.*, 1997; Daly *et al.*, 1997), while *ELF5* expression is completely lost in the *ERBB2/HER2/neu* overexpressing breast cancer cell line, ZR-75-1(Chang *et al.*, 1997), suggesting a putative tumor suppressing activity of ELF5. Recombinant human ELF3 (ESX) protein has been shown to display high affinity for the ETS binding site in the *ERBB2/HER2/neu* promoter, and human ELF3 (ESX) is capable of transactivating this promoter in COS cells (Chang *et al.*, 1997). Recombinant mouse Elf5 protein has also been shown to display high affinity for the ETS binding site in the *ERBB2/HER2/neu* promoter (Figure 3.5 b), although it is not yet known if ELF5 is capable of transactivation. In addition, it has been shown that ELF3 (ESE1) and ELF5 (ESE2a) have opposite regulating activities upon the *PSA* and *PSMA* gene promoters. ELF3 (ESE1) activates the *PSMA* but represses the *PSA* gene promoter, while ELF5 (ESE2a) activates the *PSA* but has no effect on the *PSMA* gene promoter (Oettgen *et al.*, 1999). Therefore, ELF3 (ESE1) and ELF5 (ESE2a) may have different biological functions in the prostate epithelium. Similarly, *Drosophila* ets proteins pointed and yan also

regulate the same genes in the photoreceptor cells with opposite effects (O'Neil *et al.*, 1994; Rebay and Rubin, 1995). Therefore, one can speculate that while ELF3 transactivates the *ERBB2/HER2/neu* promoter, ELF5 may have an opposing effect, and that loss of *ELF5* expression could lead to the overexpression of *ERBB2/HER2/neu*.

Several circumstantial lines of evidence also suggest that there are unidentified ETS family members with tumor suppressor properties. An ETS element in the *MASPIN* promoter appears to be active in normal mammary and prostate epithelial cells, but inactive in tumor cells (Zhang *et al.*, 1997a; Zhang *et al.*, 1997b). MASPIN is a tumor-suppressing serpin expressed in normal breast and prostate epithelium, where *ELF5, ELF3* and *ESE3* are also expressed. Also, a recent report implicating ELF3 (ERT) in the positive regulation of *TGF-β RII* transcription (Choi *et al.*, 1998) notes that a poor TGF responsiveness, observed in many tumors and thought to contribute to malignant transformation, could be caused by defects in TGF receptor expression. Potentially, defects in *TGF-β RII* transcription could be caused by lack of the ETS factors that normally promote their expression. However, the transcriptional activity of ELF5 upon *MASPIN* or the *TGF-β RII* promoters is not known.

Although the study presented in this thesis has indicated loss of *ELF5* gene expression in some cancer cell lines that are derived from tissues that normally express *ELF5*, the cause of this loss of expression is still unknown. One can speculate that certain transcriptional repressors may suppress *ELF5* gene expression in these cells. Alternatively, the *ELF5* gene promoter may be silenced by mechanisms such as methylation and/or mutation. Detailed characterization of the *ELF5* gene promoter in the future may be useful for the identification of putative transcriptional repressors or methylation/mutation sites of this promoter. In addition, chromosomal rearrangement and mutations in the *ELF5* gene locus may also result in loss of gene expression. Therefore, sequence comparison between normal and cancerous *ELF5* gene loci may identify these defects. However, we can not rule out the possiblity of a combination of some or all of above mentioned causes exerting their effect on *ELF5* expression in cancer cell lines.

Some *ETS* factors are found clustered together on a particular chromosome. For instance, *ETS1* and *FLI1* co-localize to chromosome 11q23.24 whereas *ETS2* and *ERG* co-localize to chromosome 21q22. The process that resulted in the close linkage of these *ETS* genes is still

debatable. One suggestion is that the closely linked *ETS* and *ERG* genes resulted from the duplication of an ancestral *ets* gene at the same chromosomal locus, which was then duplicated to another part of the genome during evolution. However, others suggest that ancestral *ets* gene was duplicated onto another chromosome, then through a rearrangement process, these copies were brought together at the same locus before being further duplicated to another part of the genome (Laudet *et al.*, 1999). Human *ESE3* also appears to be colocalized with human *ELF5* on chromosome 11p14.1 (Kas *et al.*, 2000). Besides their highly similar expression patterns, these two ETS factors also share a high percentage of protein sequence similarity. In addition, both ELF5 and ESE3 are able to transactivate the same epithelial target genes, PSA and SPRR2A, *in vitro* (Oettgen *et al.*, 1999; Kas *et al.*, 2000). Therefore, it will be interesting to elucidate whether these two closely related ETS transcription factors also share functional similarity.

Chapter 5

Characterization of the mouse and human Elf5 genomic loci and Generation of mouse Elf5 knockout ES cells and mice

5.1 Introduction

Gene targeting technology has enabled the investigation of the biological functions of individual genes in the mouse genome. The phenotypes presented by mice with gene-specific mutations are extremely useful and informative in determining the functional specificity of a gene. In certain cases, genetic manipulation of the mouse can be used as an effective tool in modeling human cancers. As described in Chapter 1, nine Ets family members in mouse: Ets1 (Bories *et al.*, 1995; Muthusamy *et al.*, 1995; Barton *et al.*, 1998; Walunas *et al.*, 2000), Ets2 (Yamamoto *et al.*, 1998), Pu.1 (Scott *et al.*, 1994a; McKercher *et al.*, 1996), SpiB (Su *et al.*, 1997), Pea3 (J. Hassell unpublished results, cited in review by Graves and Petersen, 1998), Er81 (Arber *et al.*, 2000), Tel (Wang *et al.*, 1997), Fli1 (Melet *et al.* 1996; Hart *et al.*, 2000; Spyropoulos *et al.*, 2000) and Elf3 (Ms. A.Y.N. Ng unpublished results, personal communication), have been disrupted, and the phenotypes of the mutant mice are distinct. These experiments strongly suggest that each *Ets* gene has distinct biological functions within the mouse genome, although many family members seem to have similar biochemical functions and overlapping expression patterns. ELF5/Elf5 is a novel ETS transcription factor with a restricted epithelial-specific expression pattern. Given its transcriptional activity upon certain epithelial gene promoters, it is important for us to establish the biological function of the *Elf5* gene in the mouse, which may then provide clues to the role(s) played by the *ELF5* gene in human biology and human cancer.

Studies of embryonic stem (ES) cell lines, derived from the inner cell mass of developing blastocysts (Evans and Kaufman, 1981; Martin, 1981), has facilitated the understanding of mammalian developmental biology. These totipotent ES cells are easily maintained under defined culture conditions and are able to generate cells of all lineages, including the germline, after being introduced into host blastocysts (Bradley *et al.*, 1984; Robertson, 1987). Mice with specific genetic mutations can be generated by introducing the required mutations into murine ES cells by homologous recombination (Thomas and Capecchi, 1987; Capecchi, 1989). In addition, ES cells can be differentiated to generate certain cell lineages under appropriate conditions *in vitro* (Martin, 1981; Doetschman *et al.*, 1985; Robertson, 1987; Risau *et al.*, 1988; Wiles and Keller, 1991; Miller-Hance *et al.*, 1993; Rohwedel *et al.*, 1994; Bain *et al.*, 1995). Therefore, ES cell lines are ideal for studying gene functions both *in vitro* and *in vivo*.

Gene targeting technology involves the recombination between DNA sequences at a specific genomic locus and newly introduced DNA sequences that are highly similar (Thomas and Capecchi, 1987). This technology can be applied to a single cell and/or a whole animal. In particular, mice with specifically-designed genetic alterations have proven to be useful in elucidating the functions of individual components of complex biological systems (Thomas and Capecchi, 1987; Capecchi, 1989). Therefore, gene targeting has become one of the most reliable methods for studying the biological functions of murine genes. Two gene targeting strategies, conventional and conditional, are commonly used. However, we will only concentrate on the conventional strategy used in this study.

The first step in a conventional gene targeting experiment involves the generation of a targeting vector. This vector must contain sequences homologous to the genomic locus to be targeted. Modifications in the central portion of the targeting vector, between the homologous sequences, can be tolerated (Hasty and Bradley, 1993). It is also extremely important to use isogenic genomic DNA for the construction of the targeting vector because polymorphisms are present in DNA derived from different genetic backgrounds (te Riele *et al.*, 1992) which may affect the efficiency of homologous recombination. In addition, positive and negative selection markers that are controlled by their own promoters and poly(A) signals, should also be included in the targeting vector to improve the selection of correctly targeted events.

Antibiotic resistance genes, such as the neomycin resistance gene (*neo*) and the hygromycin B phosphotransferase gene (*hph*), are frequently used as positive selection markers. Cells carrying these markers are resistant to G418 and Hygromycin B, respectively. Additionally, a negative selection marker, such as the herpes simplex virus thymidine kinase gene (*HSV-tk*), is often used to make cells sensitive to gancyclovir (GANC) or 1-[2-deoxy, 2-fluoro-β-D-arabinofuranosyl]-5 iodouracil (FIAU) (Thomas *et al.*, 1986; Mansour *et al.*, 1988; Hasty and Bradley, 1993). The positive selection marker is inserted into or replaces a desired exonic genomic fragment which results in the disruption, deletion and/or modification of the functional domain(s) in the targeted allele after the homologous recombination event. In contrast, the negative selection marker is inserted at the 5'- or 3'-end of the genomic sequence or into the plasmid backbone in the targeting vector. In this way, the correctly targeted cells are resistant to both G418/Hygromycin B and GANC/FIAU. In the case of random integration, cells will carry both the positive and negative selection markers in the genomic DNA, and hence become resistant to G418/Hygromycin B but sensitive to GANC or FIAU. The presence of GANC/FIAU will eventually cause death in cells carrying random integrations (Mansour *et al.*, 1988).

The targeting vector may also contain a reporter gene to facilitate gene expression studies (Mansour *et al.*, 1990). For instance, an *E. coli β-galactosidase* gene (*LacZ*) may be inserted in-frame with the gene to be disrupted so that the endogenous gene controls *LacZ* expression. The expression of *LacZ* can be easily detected as blue staining in the presence of the chromogenic substrate, 5-bromo-4-chloro-3-indolyl-*β*-D-galactosidase (X-gal) (Horwitz *et al.*, 1964). Inclusion of a nuclear localization signal (*NLS*) in front of and in-frame with the *LacZ* gene has further improved the detection of the temporal and tissue-specific expression pattern of the endogenous gene in the cell nuclei (St-Onge *et al.*, 1997; Arber *et al.*, 2000).

After electroporation of the targeting vector into ES cells followed by selection and screening to verify the homologous recombination event, the correctly targeted ES cell clone is then microinjected into blastocysts to obtain germline chimaeras (Papaioannou and Johnson, 1993). XY bearing ES cell lines are most commonly used, because it is desirable to produce male chimaeras. Males are produced when the injected ES cells colonize sufficient portions of the sex-determining tissues in the developing embryo (Pascoe *et al.*, 1992;

Papaioannou and Johnson, 1993). The coat colour is a convenient way to detect chimaerism since mice with different genetic backgrounds often have different coat colours. The proportion of the ES cell-derived coat colour represents the tissue contribution of the injected ES cells in the chimaera. If the coat colour from ES cells constitutes more than 50-60% of the chimaera, the chance of the mutation being in the germline is very high (Papaioannou and Johnson, 1993). A high percentage chimaera is then mated with wildtype mice to generate heterozygous mice with the desired mutation in the target gene. Intercrossing the heterozygous mice will then generate mice that are homozygous for the desired gene mutation.

This chapter describes the characterization of the genomic structure of the murine and human *Elf5* gene loci and the generation of ES cells and mice with a targeted mutation in the *Elf5* gene using the conventional gene targeting approach.

5.2 Results

5.2.1 Isolation of the murine *Elf5* genomic clones

Murine *Elf5* genomic clones were isolated from a 129SvJ λFIXII genomic library (Stratagene) using either of two m*Elf5* cDNA probes. The first probe, corresponding to nucleotides 181 to 1128 of the mouse *Elf5* cDNA (Figure 3.1a), hybridized to two genomic clones, C5 and C7. The second probe, corresponding to nucleotides 121 to 512 of the mouse *Elf5* cDNA (Figure 3.1a), hybridized to another two genomic clones, C2 and C4.

5.2.2 Characterization of the murine *Elf5* genomic clones

A series of oligonucleotides (Table 5.1), derived from the murine *Elf5* cDNA sequence, was hybridized to all four genomic clones to map their relative positions in the *Elf5* gene locus (Figure 5.1). Genomic clones C5 and C7 appeared to be identical. Both clones hybridized to oligonucleotides P4-P6 and P8-P17, but not to oligonucleotides P1-P3 and P18, indicating that both clones contain most of the murine *Elf5* coding sequence. However, the most 5'-end of the murine *Elf5* cDNA sequence, including the translation start codon, appeared to be missing in both the genomic clones. Genomic clone C4 also does not appear to contain the most 5'-end of the murine *Elf5* cDNA sequence (including the translation start codon). In contrast, genomic clone C2 hybridized to oligonucleotides P2-P4 but not to oligonucleotide P1 indicating that C2 extended further to the 5'-end of the murine *Elf5* coding sequence, however, the 5'-UTR sequence appears to be missing in this clone. Therefore, genomic clone C2 appeared to be partially overlapping with the other three genomic clones C4, C5 and C7 (Figure 5.1). Only C2 and C5 were characterized further for genomic mapping.

A single DNA insert of ~14.5 kb and ~16.5 kb in length for C2 and C5 respectively was released by *Not*I restriction enzyme digest. A number of *Elf5* gene-specific oligonucleotides hybridized to both DNA inserts confirming their *Elf5* genomic sequence content (Figure 5.2). The λ phage DNA from both clones was also characterized further by additional restriction enzyme digests and oligonucleotide hybridization analysis.

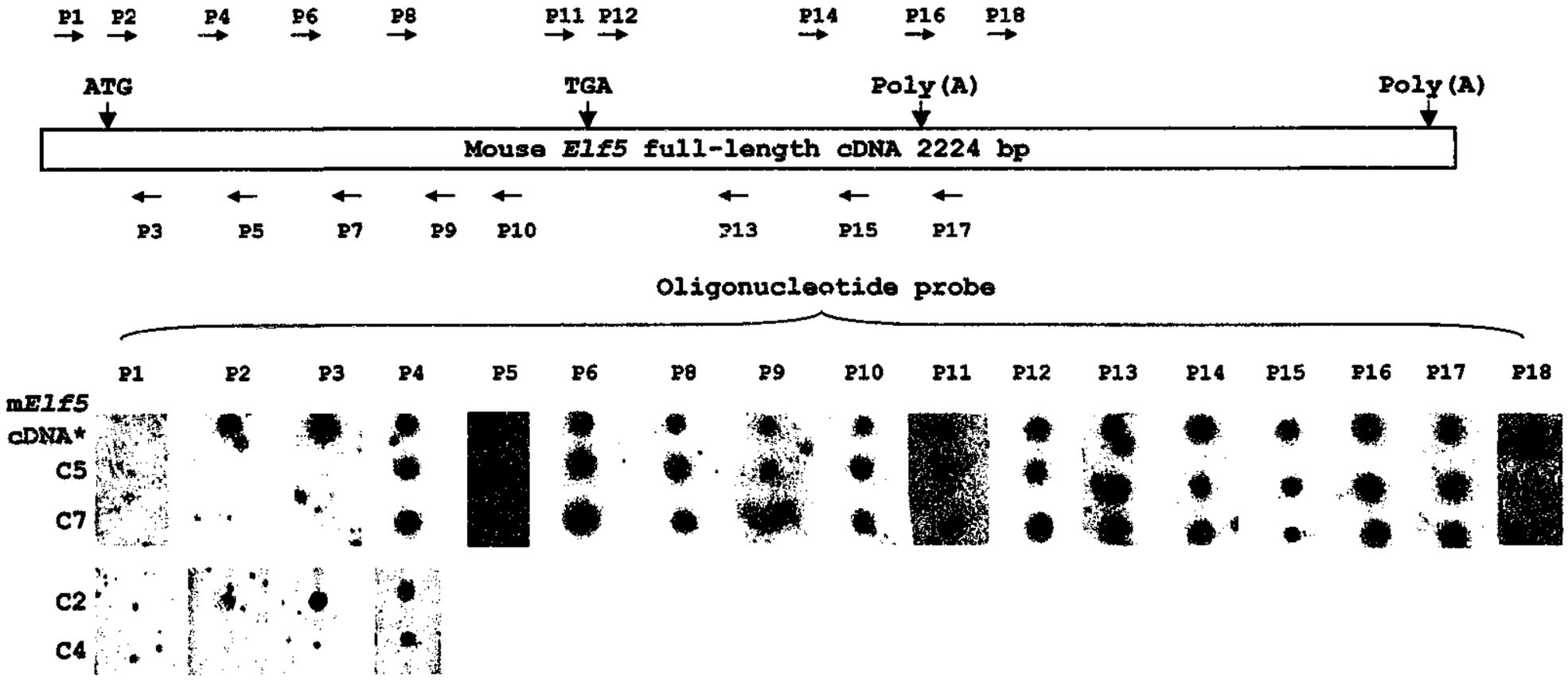


Figure 5.1 Characterization of murine *Elf5* genomic clones by dot blot analysis. Genomic DNA and *Elf5* cDNA fragments were immobilized onto the nitrocellulose membrane and probed with *Elf5* gene-specific oligonucleotides (Table 5.1) labelled with [γ-^{32}P] dATP. The *Elf5* cDNA fragment used corresponds to nucleotides 42 to 1526 of the mouse *Elf5* full-length cDNA sequence (Figure 3.1 a), and therefore does not contain oligonucleotide P1. Genomic clones C2 and C4 were only hybridized with oligonucleotide P1-4. Genomic clones C4, C5 and C7 appeared to have similar 5'ends, whilst C2 had more 5' sequences. A schematic representation of the oligonuceotide positions relative to the mouse *Elf5* full-length cDNA is also shown.* m*Elf5* cDNA used for dot blot analysis only contained oligonucleotides P2-18.

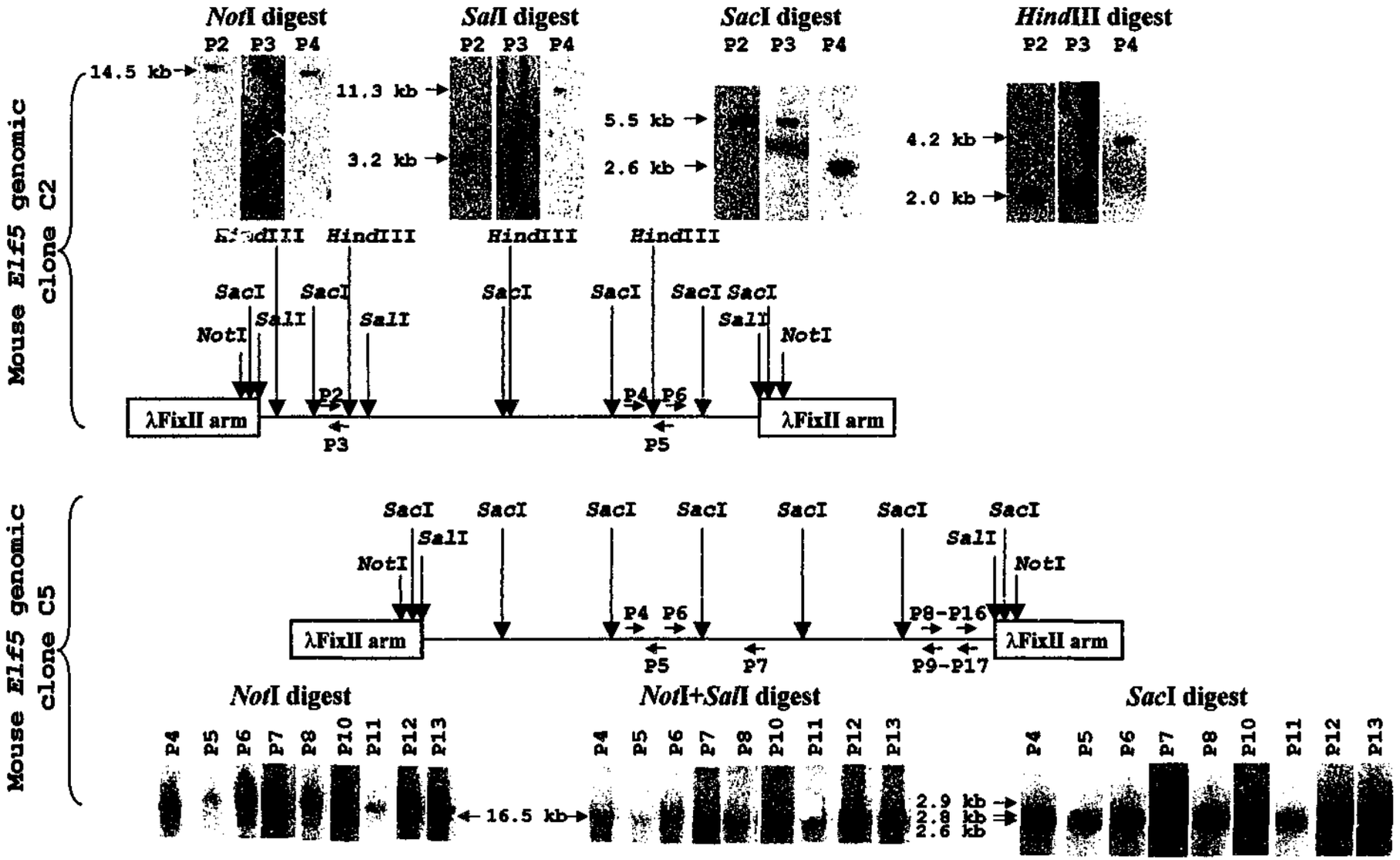


Figure 5.2 Characterization of murine *Elf5* genomic clones by restriction enzyme digestion and Southern blot analysis. Genomic fragments are flanked by λFixII vector arms. Restriction enzyme sites for two overlapping genomic clones, C2 and C5, are shown. *Elf5* gene-specific oligonucleotides (Table 5.1) labelled with [γ-^{32}P] dATP were hybridized to genomic DNA fragments digested by restriction enzymes as indicated. Only a number of representative Southern blots are shown.

Primer name	Starting nucleotide position in *Elf5* cDNA#	Oligonucleotide sequence
P1	43	5' GCT GGG AGC GCG CTT GCC TT 3'
P2	121	5' TGG ACT CCG TAA CCC ATA GCA CCT 3'
P3	200 *	5' GCT GAA CAG ATC GGT CCA AGG CAT CA 3'
P4	257	5' ATC AGT GCA CCC TGA ATA CTG GAC 3'
P5	351 *	5' AGG AGA TGC AGT TGG CAT CAA GCT 3'
P6	402	5' GAG TTC ATT GAG GCA GCC GGC AT 3'
P7	512 *	5' GCC AGT CTT GGT CTC TTC AGC ATC 3'
P8	558	5' AAG AGT CAA GAC TGT CAC AGC CGA 3'
P9	664 *	5' CAT TCC AGG ATG CCA CAG TT 3'
P10	761 *	5' CTC GTA CGT CAT CCT GTC ATT C 3'
P11	816	5' GAC CGG AGG TTA GTG TAC AA 3'
P12	888	5' GAC ACC AGG CTC ATT TGA TGG A 3'
P13	1113 *	5' ATT GCT TAA GGG CTG ATG GCA TCG 3'
P14	1201	5' AGC CAG TGT TAT GGG TGC TG 3'
P15	1301 *	5' AAG CTG CCT CAG TCG ACC GC 3'
P16	1362	5' ATC ACA GTT ATA TTC CAT AG 3'
P17	1455 *	5' GTA ACT TTC GCT ACA TCC AA 3'
P18	1498	5' ACA GTC ACT TGA TCC ACG GCC AAT CC 3'

Table 5.1 Oligonucleotide sequences and their relative starting positions in the mouse *Elf5* cDNA (GeneBank Accession: AF049702). * represents an antisense oligonucleotide sequence, # murine *Elf5* cDNA sequence shown in Figure 3.1a.

Genomic clone C5 was digested with *Sac*I, *Not*I and *Sal*I. Both *Not*I and *Not*I+*Sal*I digests released a single DNA insert of ~16.5 kb in length indicating that no additional *Sal*I site was present in C5 (Figure 5.2). Three *Sac*I digested fragments were subcloned into pBluescript. These included (a) an ~2.6 kb fragment hybridizing to oligonucleotides P4-P6; (b) an ~2.9 kb fragment hybridizing to oligonucleotide P7; and (c) an ~2.8 kb fragment hybridizing to oligonucleotides P8-P17 (Figure 5.2). Two additional *Sac*I fragments of ~3.2 kb and ~2.1 kb did not hybridize to any of the *Elf5* gene-specific oligonucleotides, and therefore probably represent intronic *Elf5* sequences. Since the total insert size is ~16.5 kb, there must be an additional ~2.9 kb *Sac*I fragment masked by the 2.9 kb hybridizing fragment.

Genomic clone C2 was digested with *Sal*I, *Sac*I and *Hind*III. *Sal*I released two DNA fragments of ~3.2 kb and ~11.3 kb in length. The ~3.2 kb *Sal*I fragment hybridized to oligonucleotides P2-P3, whereas the ~11.3 kb *Sal*I fragment hybridized to oligonucleotide P4 (Figure 5.2). *Sac*I digestion released five DNA fragments. These included (a) an ~2.6 kb fragment hybridizing to oligonucleotide P4; (b) an ~5.5 kb fragment hybridizing to oligonucleotides P2-P3 (Figure 5.2); and (c) three *Sac*I fragments of ~1.5 kb, ~1.7 kb and

~3.2 kb in length, that did not hybridize to any of the *Elf5* gene-specific oligonucleotides, and may therefore represent intronic *Elf5* sequences (data not shown). Finally, the *Hind*III digest released three DNA fragments of ~2 kb, ~4.2 kb and ~4.6 kb in length. The ~2 kb *Hind*III fragment hybridized to P2-P3, and the ~4.2 kb fragment hybridized to P4 (Figure 5.2). The ~4.6 kb fragment did not hybridize to any of the *Elf5* gene-specific oligonucleotides and may therefore also represent intronic *Elf5* sequences (data not shown). Since the ~3.2 kb *Sal*I and ~5.5 kb *Sac*I DNA fragments contained the most 5'-end of the murine *Elf5* gene exonic sequence, they were subcloned into pBluescript for use in subsequent targeting vector construction.

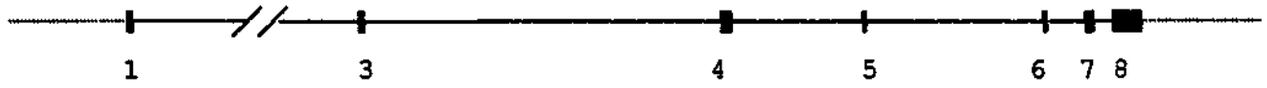
5.2.3 The partial genomic structure of the murine *Elf5* gene

Given that C2 and C5 partially overlap and contain most of the murine *Elf5* mRNA sequence (Figure 5.2), the DNA fragments from these two genomic clones were subcloned and then sequenced to determine the exon-intron structure of the murine *Elf5* gene locus. The genomic structure of the murine *Elf5* gene locus was only partially determined due to incomplete representation of the *Elf5* genomic sequence in clones C2 and C5.

At least seven exons and six introns were identified in ~21.4 kb of genomic sequence (Figure 5.3). The murine counterpart of human *ELF5* exon 2 was not found. The exon-intron boundaries of the *Elf5* gene were obtained by comparing the genomic sequence with the known *Elf5* cDNA sequence. The sequences of exon-intron boundaries and the approximate exon/intron sizes are shown in Figure 5.3. All exon-intron boundaries conform to the Breathnach (GT/AG) rule (Breathnach *et al.*, 1978), and all splice donor and acceptor site sequences are similar to the Mount consensus sequences (Mount, 1982).

The exon/intron organization of the murine *Elf5* gene is shown in Figure 5.3. The genomic DNA fragment containing the 5'-UTR was not isolated. It is assumed that the first exon will contain this sequence and its exact size is unknown. The ATG translation initiation codon and the TGA translation termination codon are found in exon 3 and exon 8, respectively, while exon 8 also contains part of the 3'-UTR. The genomic DNA fragment containing the remainder of the 3'-UTR was not isolated. The conserved PNT domain spans exons 3 and 4, whilst the ETS domain is encoded by exons 7 and 8.

Murine *Elf5* gene



Exon (size) bp	Consensus* match	Sequence at exon-intron boundary 5′ splice donor..(intron size, kb)..3′ splice acceptor			Consensus* match	Nucleotide# position
1 (>112)			..(	)..		
			..(> 2.0	)..CCTCTGATATTCCAG/G	13/16	112/113
3 (125)	8/9	CAG/GTGTGT..	(~ 9.0	)..TGTGCTCTGGTGCAG/C	11/16	237/238
4 (234)	8/9	AAG/GTCAGT	(~ 3.2	)..TCCTTTCTTTCCCAG/G	16/16	471/472
5 (51)	7/9	ACT/GTAAGT..	(~ 4.6	)..GTCCTTCTGTTTCAG/A	13/16	522/523
6 (63)	7/9	CAA/GTAAGG..	(~ 1.0	)..CTCCTTCCCCCAAAG/G	15/16	585/586
7 (196)	8/9	GAG/GTAAGT..	(~ 0.5	)..TGCTTGCTTTCACAG/A	13/16	781/782
8 (>695)						
Consensus@		MAG/GTRAGT..	(	)..YYYYYYYYYYNYAG/G		

* Fraction of nucleotides conserved relative to Mount consensus sequence (Mount, 1982)
\# Relative to murine *Elf5* cDNA sequence in Figure 3.1a
@ M = A or C; R = A or G; Y = C or T; N = A or C or G or T

Figure 5.3 Genomic organization of the murine *Elf5* gene. The partial genomic structure of murine *Elf5* is shown. Exons are represented as numbered solid boxes. Murine *Elf5* gene exon/intron sizes, sequences of exon-intron boundaries and their relative genomic positions are shown. Sequences of splice donor and splice acceptor sites are shown, along with their match to consensus sequences. All exon-intron boundaries conform to the Breathnach (GT/AG) rule (underlined; Breathnach *et al.*, 1978).

5.2.4 The partial genomic structure of the human *ELF5* gene

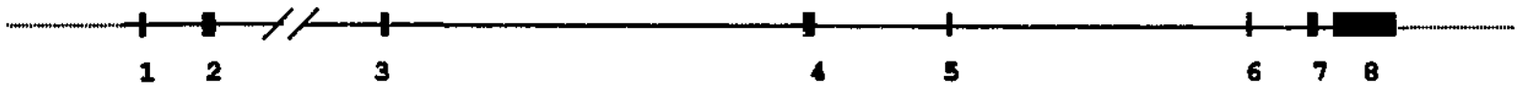
A 'working draft' sequence of human chromosome 11 is available in the GenBank database, and the human *ELF5* cDNA sequence is found to match sequence fragments in clone RP4-594L9 (GeneBank Accession: AL137224). The genomic structure of the human *ELF5* gene was obtained by comparing the genomic sequence with the known *ELF5* cDNA sequences. The sequences of exon-intron boundaries and the approximate exon/intron sizes are shown in Figure 5.4, however, some of the intron sizes can not be determined accurately due to incomplete genomic sequence presented in this 'working draft'. All exon-intron boundaries conform to the Breathnach (GT/AG) rule (Breathnach *et al.*, 1978), and all splice donor and acceptor site sequences are similar to the Mount consensus sequences (Mount, 1982).

The exon/intron organization of the human *ELF5* gene is shown in Figure 5.4. Eight exons and seven introns were identified in more than 34.4 kb of genomic sequence. Exon 1 encodes the 5'-UTR for *ELF5-a* and *ELF5-b* (see Figure 3.2 a & b), whereas exon 2 encodes an alternative 5'-UTR and 10 additional amino acids for *ESE-2a* (see Figure 3.2c). The ATG translation initiation codons are found in exon 2 and exon 3 corresponding to isoforms *ESE-2a* and *ELF5-a/ELF5-b*, respectively, and the TGA translation termination codon is located in exon 8 which also contains the 3'-UTR. The conserved PNT domain spans exons 3 and 4, whilst the ETS domain is encoded by exons 7 and 8. Interestingly, exons 4 and 5 are spliced out in *ELF5-b* (Figure 3.2 b & c & e).

5.2.5 Generation of the murine *Elf5* gene knockout targeting construct

The main objective of this study is to determine the biological role played by Elf5. One convenient way to achieve this goal is to completely disrupt its *in vivo* function(s) and study the consequent phenotype(s) due to this inactivation. Although the murine counterpart of human *ELF5-a* is the only isoform found in tissues examined, murine isoforms corresponding to human *ELF5-b* and *ESE-2a* may exist in a tissue-specific manner. Therefore, it is important to eliminate all possible functional proteins derived from these isoforms.

Human *ELF5* gene



Exon (size) bp	Consensus* match	Sequence at exon-intron boundary: 5′ splice donor..(intron size, kb)..3′ splice acceptor	Consensus* match	Nucleotide position
1 (>115)	7/9	CAC/GTAAGC..(~ 1.500)..		105# /
2 (>255)	7/9	CAG/GTACCT..(> 5.590)..CTCCTGCTGTTGCAG/G	14/16	256$ / 106#
3 (125)	8/9	CAG/GTGCGT..(11.917)..TCTCCTTTGGTGCAG/C	13/16	230# / 231#
4 (234)	8/9	AAG/GTCAGT..(3.419)..CTTTTTCTTTCCCAG/G	16/16	464# / 465#
5 (51)	7/9	ACT/GTAAGT..(7.509)..ATTATTCTATTTCAG/A	12/16	515# / 516#
6 (69)	8/9	CAA/GTAAGT..(1.463)..TCCCCTTTTCAAAAG/G	14/16	584# / 585#
7 (196)	8/9	GAG/GTAAGT..(0.457)..TGCTTGCTTTTTCAG/A	13/16	780# / 781#
8 (1518)				
Consensus@		MAG/GTRAGT..( )..YYYYYYYYYYYNYAG/G		

* Fraction of nucleotides conserved relative to Mount consensus sequence (Mount, 1982)
\# Relative to human *ELF5a* cDNA sequence in Figure 3.2a
$ Relative to human *ESE2a* cDNA sequence in Figure 3.2d
@ M = A or C; R = A or G; Y = C or T; N = A or C or G or T

Figure 5.4 Genomic organization of the human *ELF5* gene. The partial genomic structure of human *ELF5* is shown. Exons are represented as numbered solid boxes. Human *ELF5* gene exon/intron sizes, sequences of exon-intron boundaries and their relative genomic positions are shown. Sequences of splice donor and splice acceptor sites are shown, along with their match to consensus sequences. All exon-intron boundaries conform to the Breathnach (GT/AG) rule (underlined; Breathnach *et al.*, 1978).

We used a conventional gene targeting strategy in designing the murine *Elf5* knockout targeting construct. A positive selection marker, the neomycin resistance gene (*neo*), as well as a promoterless *E. coli LacZ* gene with a nuclear localization signal (*NLS-LacZ*), were inserted into exon 3, 84 bp downstream of the ATG translation initiation codon. A negative selection marker, the herpes simplex virus thymidine kinase gene (*HSV-tk*), was inserted at the extreme 3′-end of the 3′ homologous arm. A detailed outline of the cloning strategy for the murine *Elf5* gene knockout targeting construct is illustrated in Figure 5.5 a-c.

5.2.5.1 Generation of a positive selection cassette for the knockout targeting construct

The *NLS-LacZ* reporter gene cassette was a gift from Dr Seung Seng Tan (Howard Florey Cancer Institute, Victoria, Australia). The ~3.3 kb *NLS-LacZ* reporter cassette is in a pBluescript vector between the *Cla*I and *Xho*I restriction sites (*Elf5*#1 vector, Figure 5.5a). An ~1.1 kb neomycin resistance gene (*neo*) DNA fragment between *Xho*I and *Sal*I restriction sites was inserted into the unique *Xho*I restriction site in the *Elf5*#1 vector in the reverse orientation to the *NLS-LacZ* reporter cassette. The resulting vector (*Elf5*#2 vector) therefore contained a positive selection marker and was ready for insertion of the homologous arms (Figure 5.5a).

5.2.5.2 Generation and cloning of the 5′ homologous arm for the knockout targeting construct

The ~3.2 kb *Sal*I fragment from genomic clone C2 was subcloned into pBluescript vector, and then used as a template to amplify the genomic sequence between oligonucleotides T7, derived from the RNA promoter in the pBluescript vector, and P3 (Table 5.1), an *Elf5* gene-specific oligonucleotide. This PCR product contained ~2.1 kb intronic sequence, located at the 5′-end of the PCR fragment, and the first 84 bp of the *Elf5* gene coding sequence, located at the 3′-end of the PCR fragment. The blunted PCR fragment was digested with *Not*I to create a 5′-overhang at the 5′ end, and then cloned into a pBluescript vector cut with *Not*I and *EcoR*V. The resulting vector was digested again with *Not*I and *Xho*I to release the *Elf5* genomic DNA fragment, and the 5′-protruding end created by *Xho*I was

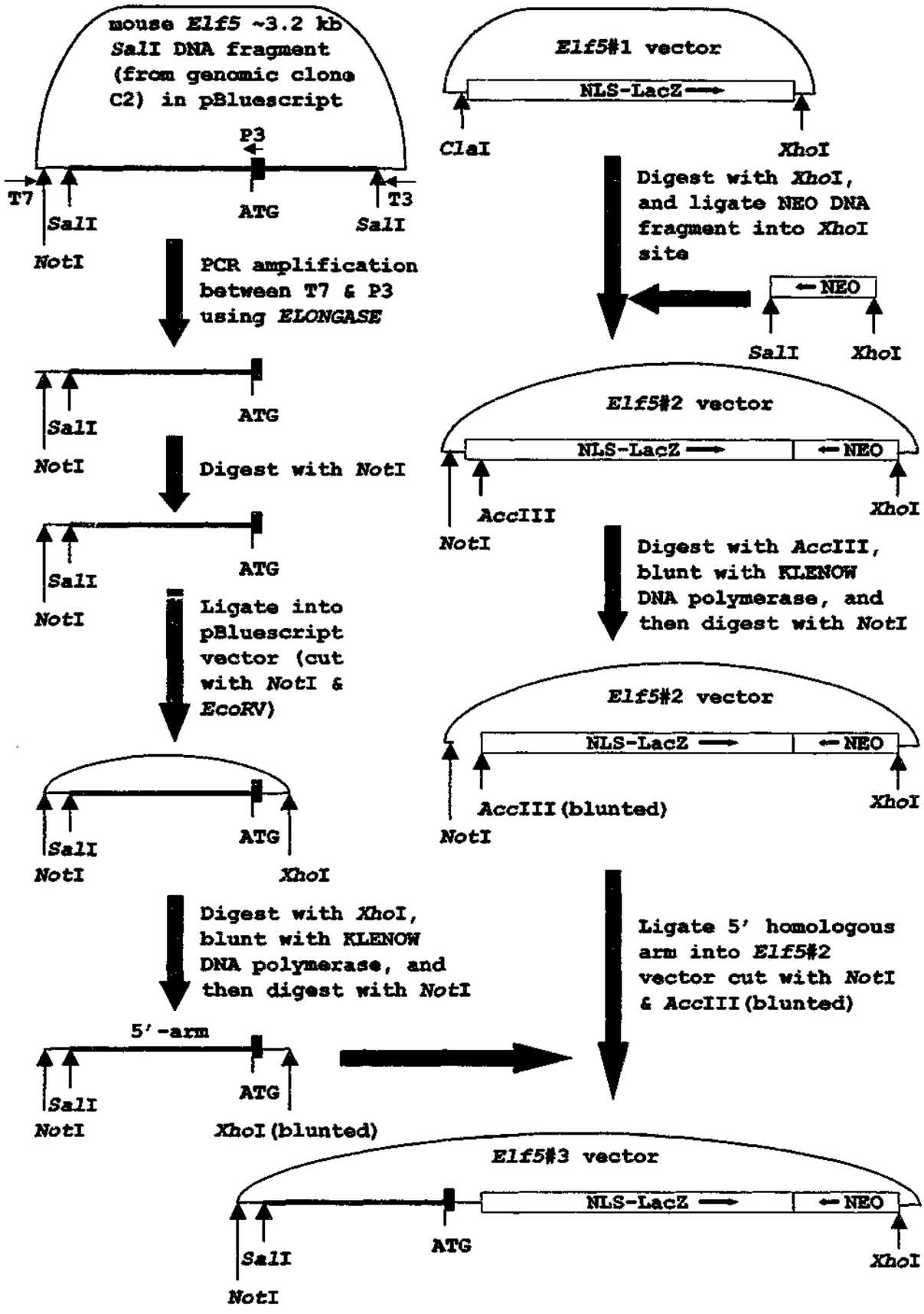


Figure 5.5 Generation of the murine *Elf5* gene knockout targeting construct. **(a)** A schematic representation of the generation and cloning of the positive selection cassette and the 5' homologous arm for the knockout targeting construct. An ~1.1 kb pMCI*neo* is used as a positive selection marker and an ~2.2 kb *Elf5* genomic DNA fragment is amplified as a 5' homologous arm. An ~3.3 kb *NLS-LacZ* reporter gene cassette is cloned in-frame of the *Elf5* gene translation initiation codon.

converted to a blunt-end terminus with dNTPs and *Klenow DNA polymerase*. This DNA fragment was then inserted into the *Elf5*#2 vector as a 5' homologous arm (Figure 5.5a).

In order to have the promoterless *NLS-LacZ* reporter gene expression controlled by the endogenous *Elf5* gene promoter, it must be placed in-frame and in the same orientation as the endogenous *Elf5* gene. An unique *Acc*III restriction site was found in the *NLS* sequence in the *Elf5*#2 vector. This site is located upstream of the sequence shown to be critical for nuclear localization of the protein (Kalderon *et al.*, 1984; Figure 5.5a). The *Elf5*#2 vector was digested with *Not*I and *Acc*III. The *Acc*III restriction site was converted to a blunt-end terminus with dNTPs and *Klenow DNA polymerase*. The 5' homologous arm was then inserted into the vector in the same orientation as the *NLS-LacZ* reporter gene cassette. This placed the *NLS-LacZ* fusion gene in-frame of the *Elf5* gene (*Elf5*#3 vector, Figure 5.5 a & c).

5.2.5.3 Generation and cloning of the 3' homologous arm for the knockout targeting construct

An ~1.85 kb *Elf5* genomic DNA fragment was excised using restriction enzymes *Xba*I and *BamH*I from the ~5.5 kb *Sac*I genomic subclone, derived from genomic clone C2. This fragment (from intron 3 of the murine *Elf5* gene locus) was subcloned into the pBluescript vector. An ~1.65 kb *Sal*I fragment derived from this construct was excised and inserted in a forward orientation into a compatible and unique *Xho*I restriction site in *Elf5*#3 vector to form the 3' homologous arm. The resulting *Elf5*#4 vector contained a positive *neo* selection marker and a *NLS-LacZ* reporter gene sequence between two arms with sequences homologous to the murine *Elf5* genomic sequence (Figure 5.5 b).

5.2.5.4 Cloning of the negative selection marker into the knockout targeting construct

An ~1.86 kb herpes simplex virus thymidine kinase gene (*HSV-tk*) DNA fragment was end-filled with dNTPs and *Klenow DNA polymerase*. This fragment was then inserted into the blunted unique *Kpn*I restriction site in the *Elf5*#4 vector in a reverse orientation to the murine *Elf5/NLS-LacZ* reporter gene cassette. The resulting vector (*Elf5*#5 vector) contained an ~2.2 kb 5' murine *Elf5* gene-specific arm, an ~3.3 kb in-frame *NLS-LacZ* fusion gene, an

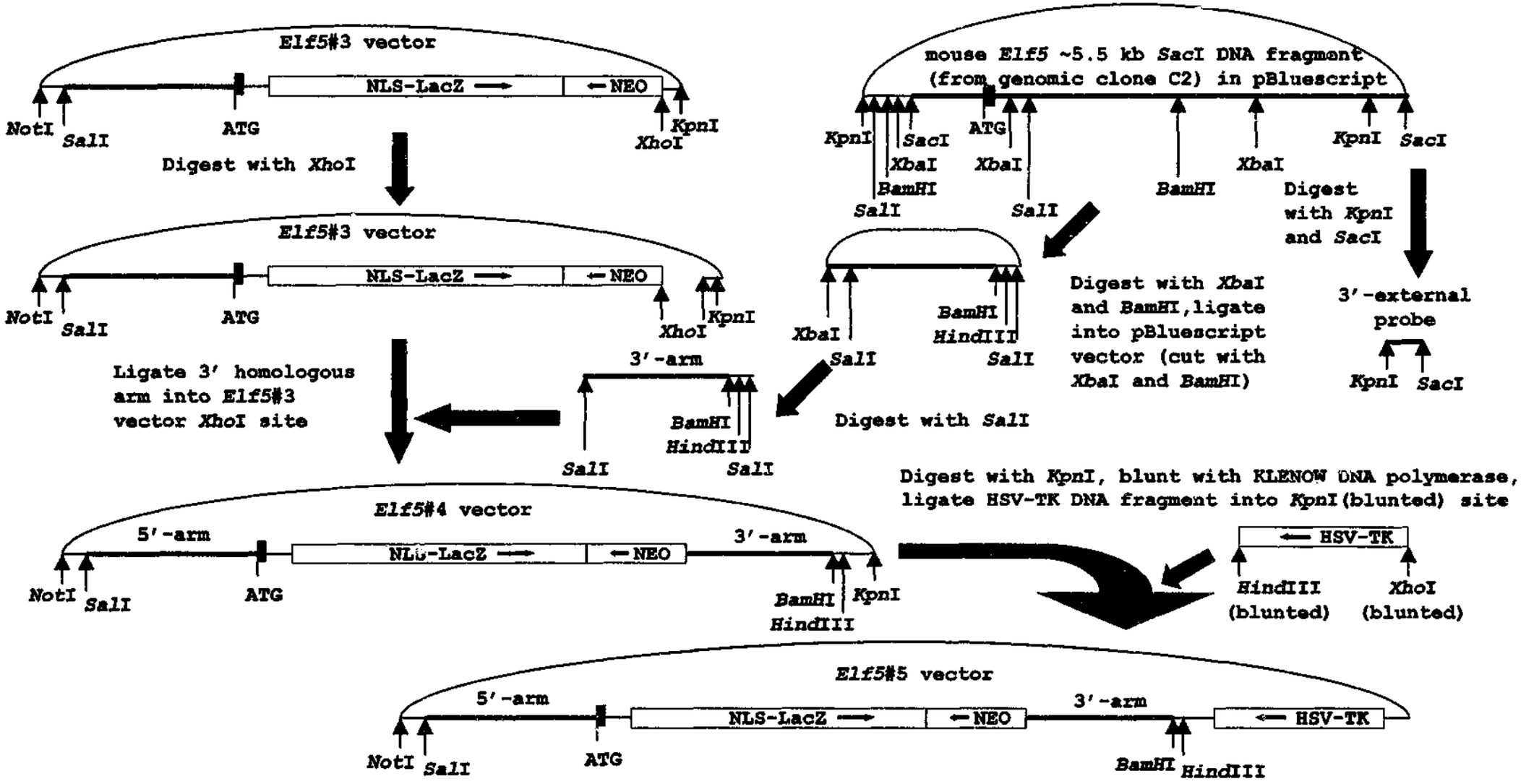


Figure 5.5 Generation of the murine *Elf5* gene knockout targeting construct. **(b)** A schematic representation of the generation and cloning of the 3' homologous arm and the negative selection cassette. An ~1.65 kb *Elf5* genomic DNA fragment is used as a 3' homologous arm and an ~1.86 kb *HSV-TK* fragment is used as a negative selection marker. An ~400 bp DNA fragment between *KpnI* and *SacI* will be used as a 3'-external probe for the identification of the correctly targeted ES cell clones by Southern blot analysis.

~1.1 kb positive (*neo*) selection marker, an ~1.65 kb 3′ murine *Elf5* gene-specific arm, and an ~1.86 kb negative (*HSV-tk*) selection marker. Both the positive and negative selection markers were in reverse orientation relative to *Elf5/NLS-LacZ* (Figure 5.5 b & c). Sequencing confirmed that *NLS-LacZ* was in-frame with *Elf5* (Figure 5.5 c).

5.2.6 Generation of the *Elf5*$^{+/-}$ ES cells

The final *Elf5* gene knockout construct (*Elf5*#5 vector) was linearized with *Sac*II and electroporated into isogenic 129SvJ J1 ES cells. Three hundred and thirty nine surviving ES cell clones were picked after 12 days of double selection in G418 and GANC. Genomic DNA was isolated from these ES cell clones and subjected to Southern blot analysis to identify correctly targeted *Elf5*$^{+/-}$ ES cell clones.

An ~400 bp DNA fragment, located at the extreme 3′-end of the ~5.5 kb *Sac*I DNA fragment (derived from genomic clone C2), was used as a 3′-external probe (Figure 5.5 b & Figure 5.6 a). *Hind*III restriction enzyme sites in the murine *Elf5* gene locus were used for the screening purposes (Figure 5.6 a). The 3′-external probe should recognize an ~4.6 kb *Hind*III DNA fragment in the wildtype allele. In contrast, an ~8.6 kb novel *Hind*III DNA fragment should be generated by the targeted allele. Eleven ES cell clones (#1, #55, #79, #95, #130, #202, #213, #274, #293, #301 and #325) were identified as having undergone homologous recombination. A representative Southern blot is shown in Figure 5.6 b. An expected ~8.6 kb *Hind*III DNA fragment was also detected in all positive clones using a probe for the neomycin resistance gene (*neo*) (Figure 5.6 b).

5.2.7 Generation of the *Elf5*$^{+/-}$ and *Elf5*$^{-/-}$ mice

Elf5$^{+/-}$ ES cell clones #1 and #130 were microinjected into C57Bl/6J murine blastocysts, which were then transferred into the uterine horns of F1 (CBA×C57Bl/6J) pseudopregnant female mice. The resulting male chimaeras, exhibiting more than 90% agouti coat colour due to the 129SvJ genetic background, were mated with *Elf5*$^{+/+}$ C57Bl/6J female mice. Genomic DNA from the pups was isolated and subjected to Southern blot analysis for the identification of the *Elf5*$^{+/-}$ mice.

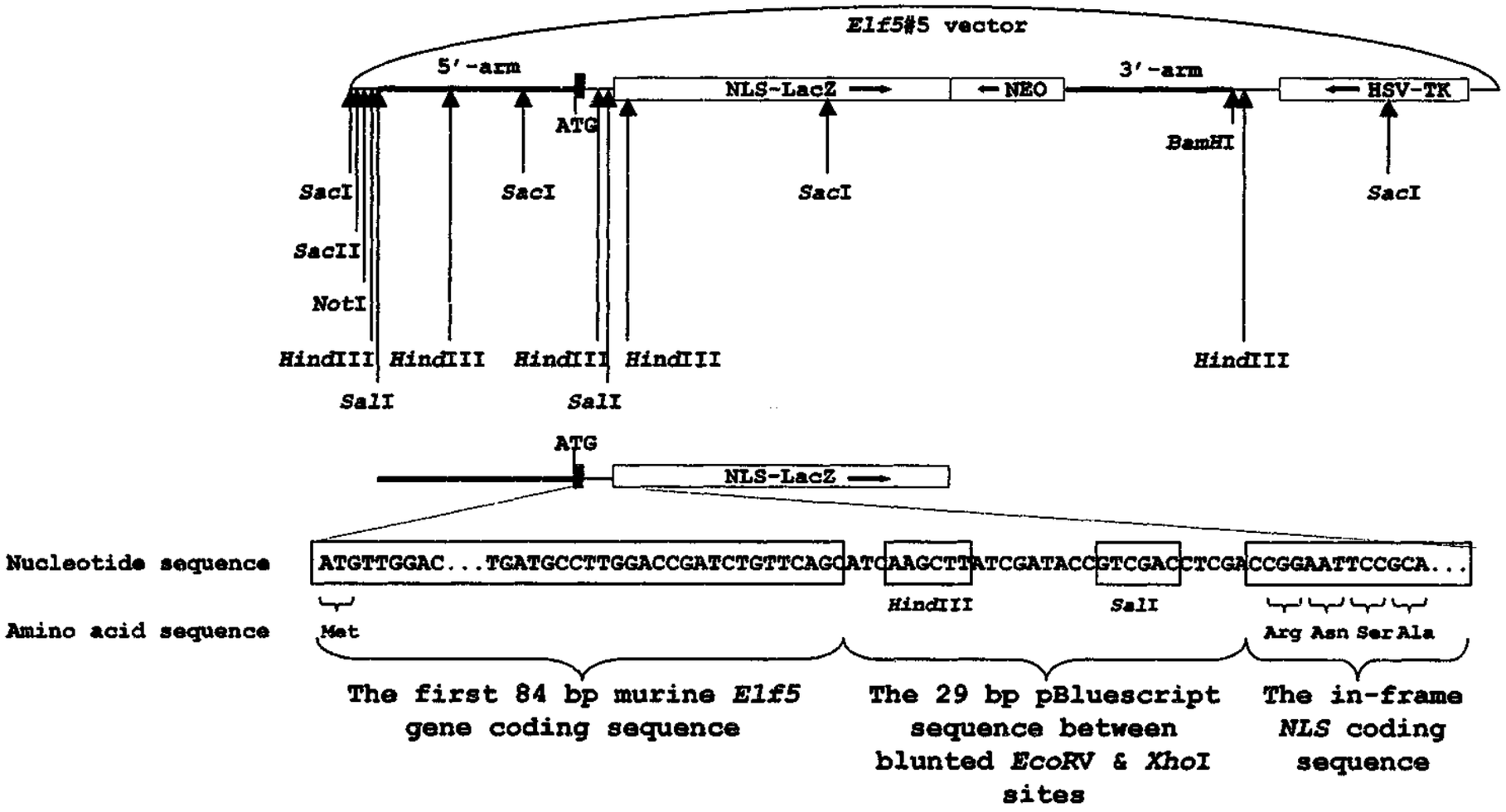


Figure 5.5 Generation of the murine *Elf5* gene knockout targeting construct. (c) A schematic representation of the final gene knockout targeting construct including an ~2.2 kb 5' homologous arm; an ~3.3 kb *NLS-LacZ* reporter cassette; an ~1.1 kb *pMCI-neo* positive selection marker; an ~1.65 kb 3' homologous arm; and an ~1.86 kb *HSV-TK* negative selection marker. The requisite restriction enzyme sites are shown. The *NLS-LacZ* reporter cassette is in-frame of the *Elf5* gene as shown by sequencing through the junction between the 5' homologous arm and the *NLS-LacZ* reporter cassette. *Sac*II is a unique restriction enzyme site that can be used to linearize the construct before electroporating it into the ES cells.

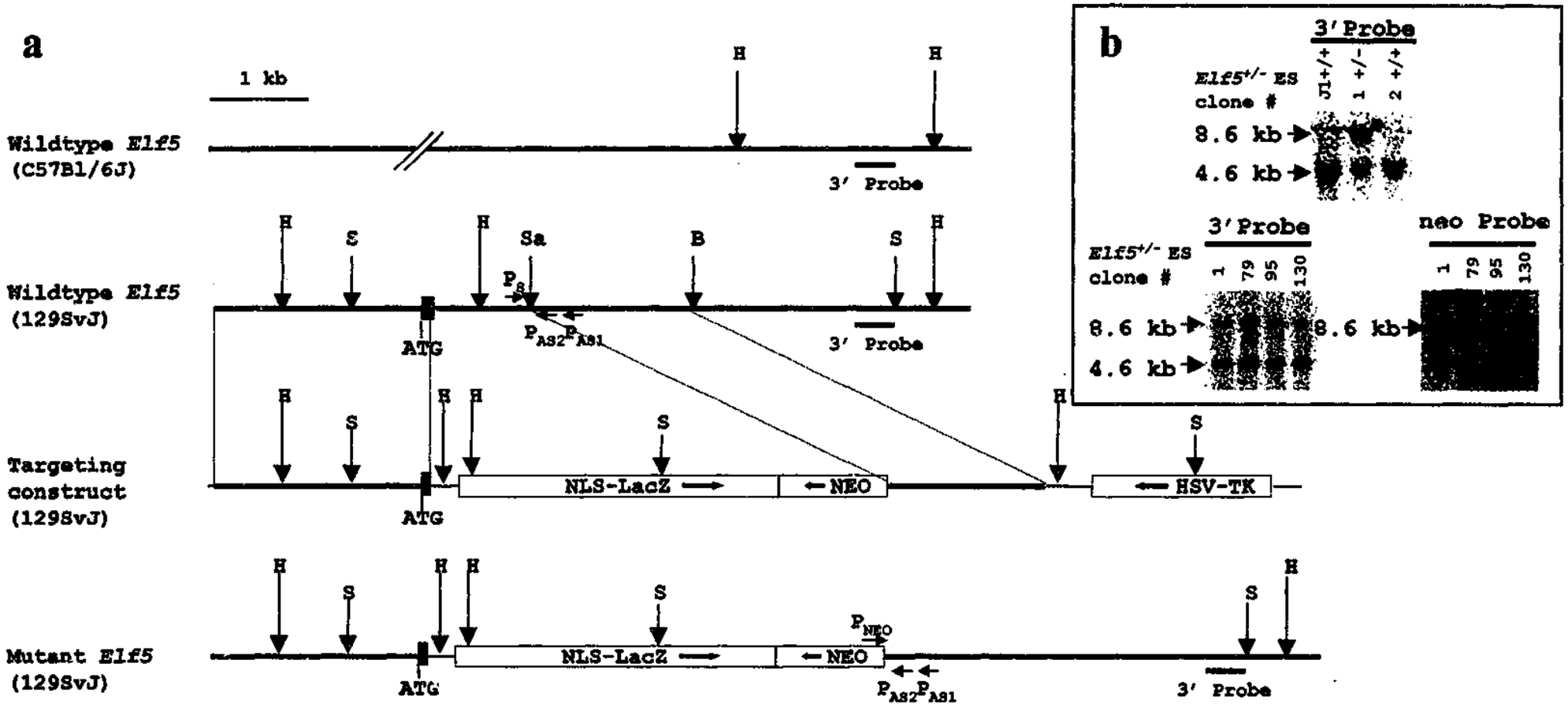


Figure 5.6 Gene targeting of the murine *Elf5* locus. **(a)** A schematic representation of the targeting strategy. The homologous recombination event replaces a portion of exon 3 and intron 3 with a *NLS-LacZ*-neo cassette. A 400 bp 3'-external probe should hybridize to a 4.6 kb 129SVJ wildtype and a 8.6 kb 129SVJ targeted *Hind*III DNA fragment. An additional *Hind*III restriction site is found in C57Bl/6J wildtype *Elf5* gene locus which will give a 2.0 kb wildtype *Hind*III DNA fragment. A 169 bp PCR fragment between P_S and P_{AS1} can be amplified from the wildtype *Elf5* gene locus, and a 237 bp PCR fragment between P_{NEO} and P_{AS1} can be amplified from the targeted *Elf5* gene locus. P_{AS2} can be used to verify the specificity of the PCR products. H: *Hind*III; S: *Sac*I; Sa: *Sal*I; B: *Bam*HI. **(b)** Generation of targeted ES cells. Southern blot analyses of only a representative number of surviving ES cell clones is shown. G418-resistant and Gancyclovir-sensitive ES cell clones were screened by Southern blot analysis with a 3'-external probe using *Hind*III digest. All ES cell clones giving a 4.6 kb wildtype and a 8.6 kb targeted *Hind*III DNA fragment were then screened by Southern blot analysis with a *neo* probe. The correctly targeted ES cell clones have a 8.6 kb targeted *Hind*III DNA fragment whilst a random integration event should give a 5.9 kb *Hind*III DNA fragment (data not shown).

The targeted allele was transmitted successfully into the germline for both *Elf5*$^{+/-}$ ES cell clones. Southern analysis revealed a polymorphism in the wildtype *Elf5* gene locus (Figure 5.6 a & Figure 5.7 a). Southern blot analysis using the 3′-external probe identified a single ~4.6 kb *Hind*III fragment in the 129SvJ J1 ES cell DNA, and a single ~2.0 kb *Hind*III fragment in C57Bl/6J mouse DNA (Figure 5.7 a). Therefore, the *Elf5*$^{+/+}$ progeny resulting from the chimaera-C57Bl/6J cross carried wildtype *Hind*III DNA fragments of ~4.6 and ~2.0 kb, arising from 129SvJ and C57Bl/6J genetic backgrounds, respectively. In contrast, the *Elf5*$^{+/-}$ littermates carried the ~8.6 kb 129SvJ targeted allele and the ~2.0 kb C57Bl/6J wildtype allele. This polymorphism was not apparent in 129Svter (Figure 5.7a).

Elf5$^{+/-}$ ES cell clone	No. of litters	Total no. of pups	*Elf5*$^{+/+}$		*Elf5*$^{+/-}$		*Elf5*$^{-/-}$	
#130	77	513	99	80	182	152	0	0
		Total number	179		334		0	
		Ratio	1	:	1.87	:	0	
#1	8	43	8	5	13	17	0	0
		Total number	13		30		0	
		Ratio	1	:	2.31	:	0	
		Expected Mendelian ratio	1	:	2	:	1	

Table 5.2 Genotypes of the progeny from the *Elf5*$^{+/-}$ intercrosses determined by Southern blot analysis and/or PCR.

Intercrossing *Elf5*$^{+/-}$ mice should generate progeny of three different genotypes: *Elf5*$^{+/+}$, *Elf5*$^{+/-}$ and *Elf5*$^{-/-}$. The genotypes of the mice were determined either by Southern blot analysis or by PCR. By Southern blot, we expect the *Elf5*$^{+/+}$ mice to harbour two wildtype *Hind*III DNA fragments of ~2.0 kb; the *Elf5*$^{+/-}$ mice to have the ~8.6 kb targeted and the ~2.0 kb wildtype bands; and the *Elf5*$^{-/-}$ mice to have the ~8.6 kb targeted bands (Figure 5.6 a & Figure 5.7 b). The PCR screening strategy uses two *Elf5* gene-specific primers, P_S and P_{AS1}, and a *neo* gene-specific primer, P_{NEO}. The PCR amplifies a single 169 bp DNA fragment between P_S and P_{AS1} in *Elf5*$^{+/+}$ mice, and a single 237 bp DNA fragment between P_{NEO} and P_{AS1} in *Elf5*$^{-/-}$ mice, whilst both DNA fragments will be amplified in the *Elf5*$^{+/-}$ mice (Figure 5.6 a & Figure 5.7 b). Five hundred and thirteen three-week-old mice derived

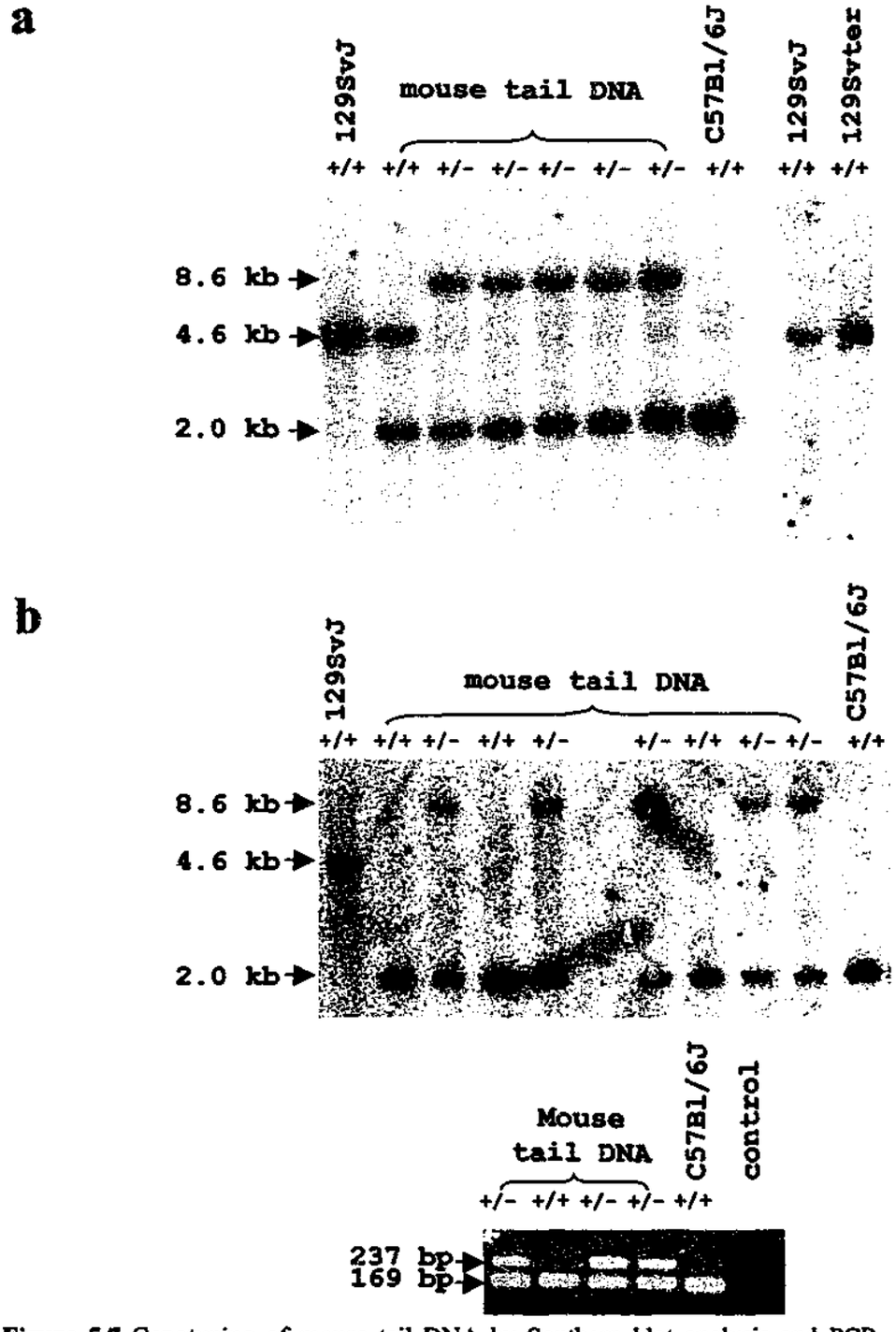


Figure 5.7 Genotyping of mouse tail DNA by Southern blot analysis and PCR. Genomic DNA was extracted from three week old mouse tails. Southern blot analysis using *Hind*III digested genomic tail DNA and the murine *Elf5* 3'external probe detected 4.6 kb (129SvJ and 129Svter) and 2.0 kb (C57Bl/6J) wildtype bands, and an 8.6 kb mutant band. **(a)** Germ line transmission of the targeted *Elf5* allele. Southern blot analysis of *Hind*III digested genomic DNA fragments as described above. *Elf5*$^{+/-}$ mice are identified by the presence of 8.6 kb targeted (129SvJ) and 2.0 kb wildtype (C57Bl/6J) bands. *Elf5*$^{+/+}$ mice are identified by the presence of 4.6 kb wildtype (129SvJ) and 2.0 kb wildtype (C57Bl/6J) bands. **(b)** Genotyping of tail DNA from the progeny of a *Elf5*$^{+/-}$ intercross. Southern blot analysis (upper panel) of *Hind*III digested genomic DNA fragments as described above. *Elf5*$^{+/-}$ mice are identified by the presence of 8.6 kb targeted (129SvJ) and 2.0 kb wildtype (C57Bl/6J) bands. *Elf5*$^{+/+}$ mice are identified by the presence of only the 2.0 kb wildtype C57Bl/6J band. No *Elf5*$^{-/-}$ mutant mice (identified by the presence of only the 8.6 kb targeted 129SvJ band) were detected. PCR analysis (lower panel) detects a 169 bp wildtype band between P_S and P_{ANI}, and a 237 mutant band between P_{NEO} and P_{ANI}. A no template H_2O control is used to detect contamination. *Elf5*$^{+/-}$ mice are identified by the presence of both the 169 bp and 237 bp bands. *Elf5*$^{+/+}$ mice are identified by the presence of only the 169 bp band. No *Elf5*$^{-/-}$ mutant mice (identified by the presence of only the 237 bp band) were detected.

from the ES cell clone #130 were genotyped. Ninety nine male and 80 female mice were *Elf5*$^{+/+}$ and 182 male and 152 female mice were *Elf5*$^{+/-}$. Both screening methods failed to detect any *Elf5*$^{-/-}$ mice. Therefore, the genotype distribution of *Elf5*$^{+/+}$: *Elf5*$^{+/-}$: *Elf5*$^{-/-}$ was 1 : 1.87 : 0 (Table 5.2). For the 43 three-week-old mice derived from the ES cell clone #1, 8 male and 5 female mice were *Elf5*$^{+/+}$ and 13 male and 17 female mice were *Elf5*$^{+/-}$. The genotype distribution of *Elf5*$^{+/+}$: *Elf5*$^{+/-}$: *Elf5*$^{-/-}$ was 1 : 2.31 : 0 (Table 5.2), therefore the *Elf5*$^{-/-}$ genotype is embryonic lethal.

5.2.8 The targeted *Elf5* gene allele is a null allele

The homozygously deficient *ELF5* mice die *in utero*, therefore obtaining *Elf5*$^{-/-}$ tissue for analysis to confirm that the targeted *Elf5* allele is truly non-functional is impossible. However, it was reasoned that the heterozygous allele should produce half the amount of *Elf5* transcript. Expression of *Elf5* in adult mouse organs from both *Elf5*$^{+/+}$ and *Elf5*$^{+/-}$ mice was analyzed by Northern blot of poly(A)$^+$ mRNA probed with the mouse *Elf5* cDNA (Figure 5.8). A *GAPDH* cDNA probe was used to control for RNA loading. Expression of two *Elf5* transcripts, *Elf5-a* (2.5 kb) and *Elf5-b* (1.5 kb), are observed in day 18.5 pregnant mammary glands, day 1 lactating mammary glands and seminal vesicles from both *Elf5*$^{+/+}$ and *Elf5*$^{+/-}$ mice. However, the *Elf5* mRNA levels in *Elf5*$^{+/-}$ day 18.5 pregnant mammary glands and seminal vesicles are reduced to ~50% of the corresponding *Elf5*$^{+/+}$ organs, and the *Elf5* mRNA level in *Elf5*$^{+/-}$ day 1 lactating mammary glands is reduced to ~25% of the corresponding *Elf5*$^{+/+}$ mammary glands. No additional mRNA transcript was observed in these *Elf5*$^{+/-}$ organs.

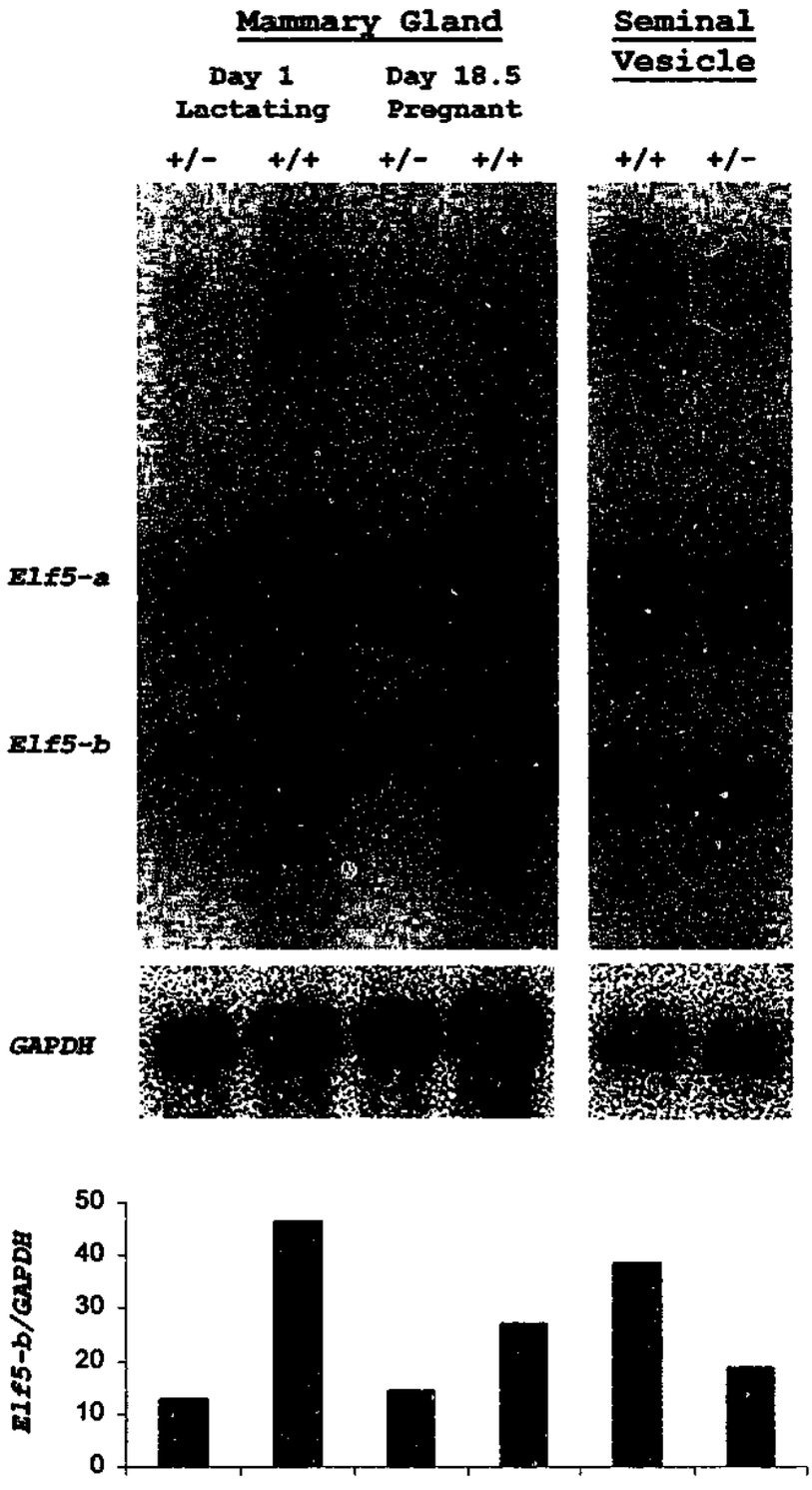


Figure 5.8 *Elf5* expression in $Elf5^{+/-}$ and $Elf5^{+/+}$ tissues. *Elf5-a* and *Elf5-b* transcripts are indicated. Northern blots probed with murine *Elf5* cDNA (top panels) and *GAPDH* cDNA (lower panels). Relative expression levels of *Elf5-b/GAPDH* are indicated at the bottom.

5.3 Discussion

The genomic organization of the human *ELF5* gene was determined by comparing a 'working draft' sequence of human chromosome 11 with the known *ELF5* cDNA sequences. The human *ELF5* gene contains eight exons and seven introns spanning more than 34.4 kb of genomic sequence (Figure 5.4).

Two murine *Elf5* genomic clones, C2 and C5, were found to overlap and contain the known *Elf5* gene coding sequence (Figure 5.2). Both clones were extensively characterized by restriction mapping and sequencing analysis to determine the genomic organization of the *Elf5* gene (Figure 5.2 and Figure 5.3). The murine *Elf5* gene contains at least seven exons and six introns spanning more than 20.4 kb of genomic sequence. We were unable to find the murine counterpart of the human *ELF5* gene exon 2. Library screening and RACE PCR using cDNA derived from mouse lung and placenta failed to detect the existence of this murine exon (data not shown).

Approximately 3% of the surviving ES cell clones (11 out of 339) were correctly targeted $Elf5^{+/-}$ ES cell lines. Interestingly, numerous attempts to generate $Elf5^{-/-}$ ES cell lines were unsuccessful with up to 10 mg/ml G418 (data not shown). Large numbers of $Elf5^{+/-}$ ES cells survived whilst the $Elf5^{+/+}$ embryonic fibroblast feeder-layer diminished. Since growth on embryonic fibroblasts is important for retaining the totipotential capacity of the ES cells (Bradley *et al.*, 1984), higher concentrations of G418 could not be used. However, alternative options are available for generating the $Elf5^{-/-}$ ES cells. For example, the $Elf5^{+/-}$ ES cells can be retargeted with a targeting vector containing a different positive selection cassette. Alternatively, $Elf5^{-/-}$ ES cells can be isolated from the $Elf5^{-/-}$ blastocyst.

The targeted *Elf5* gene locus was successfully transmitted through the germline by injecting two independent $Elf5^{+/-}$ ES cell lines, #1 and #130. $Elf5^{+/-}$ mice were generated by mating the male chimaeras with C57Bl/6J $Elf5^{+/+}$ female mice. Intercrossing the $Elf5^{+/-}$ mice should generate offspring of three different genotypes: $Elf5^{+/+}$, $Elf5^{+/-}$ and $Elf5^{-/-}$. The genotype distribution of 513 mice derived from ES cell clone #130 was found to be 1 : 1.87 : 0 for $Elf5^{+/+}$, $Elf5^{+/-}$, and $Elf5^{-/-}$ respectively (Table 5.2). Similarly, the genotype distribution of 43 mice derived from ES cell clone #1 was found to be 1 : 2.31 : 0 (Table 5.2). The expected Mendelian ratio is 1 : 2 : 1 for $Elf5^{+/+}$: $Elf5^{+/-}$: $Elf5^{-/-}$ respectively. No $Elf5^{-/-}$ mice were

produced indicating that Elf5 is essential during murine embryogenesis. The *Elf5*$^{-/-}$ embryos derived from the ES cell clone #130 were found dead during an early embryonic stage. The initial characterization of this phenotype will be described and discussed in Chapter 6.

The *Elf5*$^{+/-}$ mice derived from both *Elf5*$^{+/-}$ ES cell clones #1 and #130 are viable. However, a severe mammary gland developmental defect is found in the pregnant and lactating *Elf5*$^{+/-}$ female mice. This phenotype will be described and discussed in Chapter 7.

Northern blot analysis demonstrates that the targeted *Elf5* gene allele in *Elf5*$^{+/-}$ mice is inactivated and consequently the *Elf5* mRNA levels are reduced by ~50% in the organs examined (Figure 5.8). In addition, no additional mRNA transcript is observed in these *Elf5*$^{+/-}$ organs indicating that the targeted *Elf5* gene allele is incapable of producing any functional Elf5 protein. Therefore, any phenotypes observed in *Elf5*$^{+/-}$ and *Elf5*$^{-/-}$ mice/embryos are the result of the inactivation of the *Elf5* gene. Interestingly, the *Elf5* mRNA level in *Elf5*$^{+/-}$ day 1 lactating mammary glands is reduced to ~25% of the corresponding *Elf5*$^{+/+}$ mammary glands. One can speculate that the loss of one copy of the functional *Elf5* gene in the developing mammary gland may affect the development and integrity of the mammary epithelial cells, and hence the level of *Elf5* expression in these cells.

Chapter 6

Elf5 is essential during mouse early embryogenesis

6.1 Introduction

Many ETS factors have been shown to have important functions during embryogenesis (as discussed in Chapter 1), however, *in vivo* gene disruption experiments have established distinct roles played by different Ets family members at different embryonic stages (Scott *et al.*, 1994; Yamamoto *et al.*, 1998; Hart *et al.*, 2000; Spyropoulos *et al.*, 2000). *Elf5* gene disruption has resulted in embryonic lethality. It is of interest to try to establish the timing and the cause of this phenotype. This chapter describes the initial characterization of this lethal phenotype. It was found that lethality occurred around the implantation period of mouse embryogenesis.

Mouse early embryogenesis

Mouse embryogenesis spans approximately 19 days. A fertilized mammalian egg develops into a blastocyst after six cell divisions (Collins and Fleming, 1995; Kaufman, 1999; Kaufman and Bard, 1999). The blastocyst consists of an inner totipotent and nonpolarized cell mass (ICM), sitting in the blastocoelic cavity, and an outer differentiated and polarized trophectoderm (TE). The next step of embryogenesis involves the implantation of the blastocyst in the endometrial lining of the uterus. The "hatching" of the embryo from the zona pellucida facilitates contact between the trophectoderm and the endometrial cells, and this contact leads to the normal decidual response (or decidual reaction) surrrounding the implanting embryo in the uterus (Bell, 1985).

Just before hatching (~E4.5), the totipotent ICM cells lining the blastocoelic surface, which have been exposed to the blastocoelic fluid, delaminate/differentiate to form primary or primitive endoderm (Dziadek, 1979), while the remaining ICM cells maintain their pluripotential properties and form the embryonic epiblast. Embryonic stem (ES) cells can be derived from the ICM cells at this stage if blastocysts are maintained under defined culture conditions (Evans and Kaufman, 1981).

After hatching, the trophectoderm cells begin to transform into giant cells during implantation. This process starts at the abembryonic pole and eventually reaches the ICM region. In addition, the cells at the polar region of the trophectoderm, overlying the inner cell mass, contribute to the formation of the ectoplacental cone during implantation (Gardner and Papaioannou, 1975; Copp, 1978, 1979).

6.2 Results

6.2.1 Genotyping of prenatal and neonatal mice from *Elf5*$^{+/-}$ intercrosses

Given that *Elf5*$^{-/-}$ animals were not detected among 513 offspring (at three weeks of age) from heterozygote intercrosses (Table 6.1), we investigated the timing of the presumptive neonatal and/or embryonic mortality in *Elf5*$^{-/-}$ mice by analyzing the genotypes of neonates and embryos at various stages of embryogenesis.

Embryonic stage	No. of litters	Total no. of mice	Genotypes		
			+/+	+/-	-/-
E3.5	9	58	14	26	18
E7.5	4	40	15	25	0
E8.5	3	18	6	12	0
E9.5	4	36	8	28	0
E10.5	6	46	17	29	0
E11.5	7	45	14	31	0
E13.5	1	6	2	4	0
E14.5	1	6	2	4	0
E18.5	4	33	12	21	0
Neonate	4	24	7	17	0
Adult	77	513	179	334	0

Table 6.1 Genotypes of the embryos, neonates and adults from the *Elf5*$^{+/-}$ intercrosses as determined by Southern blot analysis and/or PCR. The *Elf5*$^{+/-}$ mice used for intercrosses were derived from the *Elf5*$^{+/-}$ ES cell clone #130. All animals were of 129SvJ and C57Bl/6J mixed genetic background. Tail or yolk sac DNA was used for genotyping.

Elf5$^{-/-}$ animals were not detected among neonates and embryos at E13.5, E14.5 and E18.5 by Southern blot analysis (as described in Chapter 5). PCR analysis of postimplantation embryos dissected from the decidua at E7.5-11.5 (with the morning of vaginal plug detection corresponding to E0.5) showed that none of these embryos were *Elf5*$^{-/-}$ (Table 6.1, PCR analysis as described in Chapter 5). However, empty decidua were often observed at E7.5-11.5 in heterozygote intercrosses (data not shown). These results suggested that the lethality of *Elf5*$^{-/-}$ embryos might occur before E7.5.

We isolated blastocysts from intercrossed *Elf5*$^{+/-}$ females at E3.5. PCR analysis detected *Elf5*$^{-/-}$ blastocysts in a ratio expected for a Mendelian distribution (Table 6.1), suggesting

that *Elf5*$^{-/-}$ embryos die between E3.5 and E7.5. The specificity of the PCR products was confirmed by the hybridization of an oligonucleotide, P_{AS2} (Figure 6.1), located within the expected PCR fragment (see Chapter 5 Figure 5.6 a).

6.2.2 *In vitro* blastocyst outgrowth assay

To assess the growth capability of *Elf5*$^{-/-}$ embryos, we collected E3.5 blastocysts from heterozygous intercrosses and cultured them individually *in vitro* for four days in cES medium on poly-L-lysine coated petri dishes (Figure 6.2 a & b; two examples are shown). The genotype of each individual blastocyst was then determined by PCR. All blastocysts appeared normal morphologically and hatched from the zona pellucida. Following hatching, *Elf5*$^{+/+}$ blastocysts attached to the culture dish and developed outgrowths. In contrast, *Elf5*$^{-/-}$ blastocysts did not attach to the culture dish and did not develop outgrowths. Instead, these blastocysts retained their orginal appearance after 48-72 hours in culture, however, their size increased dramatically. Trophoblast giant cells could not be found in these *Elf5*$^{-/-}$ blastocyst cultures. Some of the *Elf5*$^{-/-}$ blastocysts also started to degenerate after 72 hours in culture. After 96 hours in culture, all of the blastocysts underwent severe degeneration.

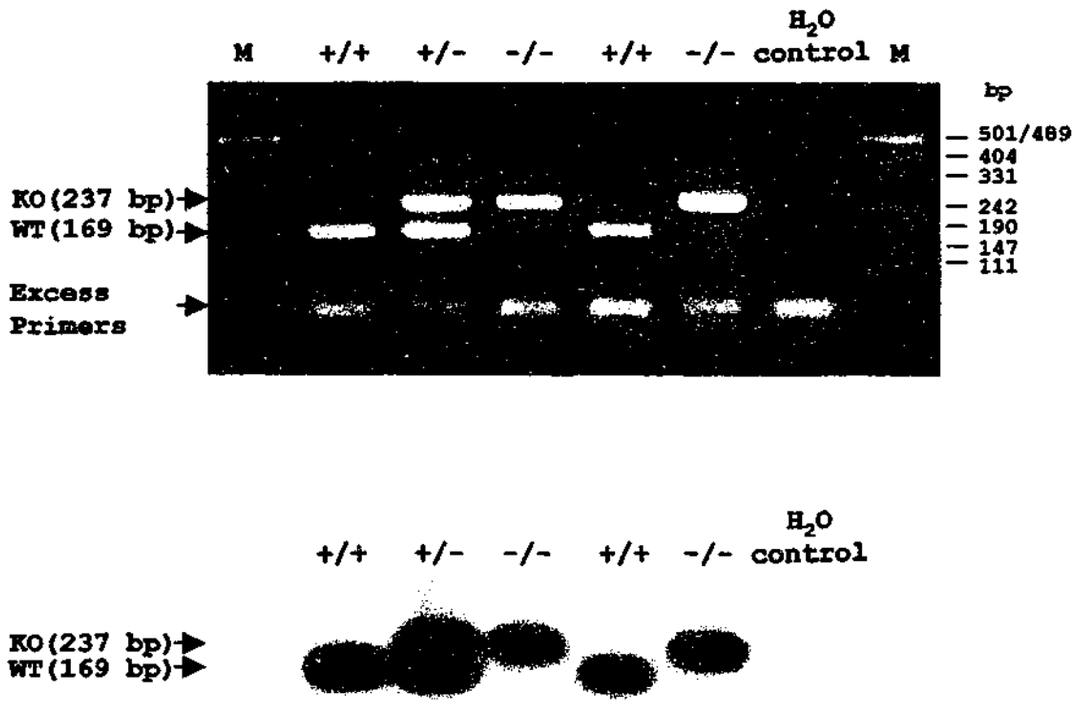


Figure 6.1 Genotyping of blastocysts by PCR and Southern blot analysis. Molecular weight marker (M) is shown. PCR analysis (top panel) detects a 169 bp wildtype (WT) band between P_S and P_{AS1}, and a 237 mutant band (KO) between P_{NEO} and P_{AS1}. A no template H_2O control is shown. *Elf5*$^{+/-}$ blastocysts are identified by the presence of both the 169 bp and 237 bp bands. *Elf5*$^{+/+}$ blastocysts are identified by the presence of only the 169 bp band. *Elf5*$^{-/-}$ blastocysts are identified by the presence of only the 237 bp band. The specificity of the PCR products was determined by Southern blot analysis (lower panel) probed with P_{AS2} labelled with [γ-^{32}P] dATP.

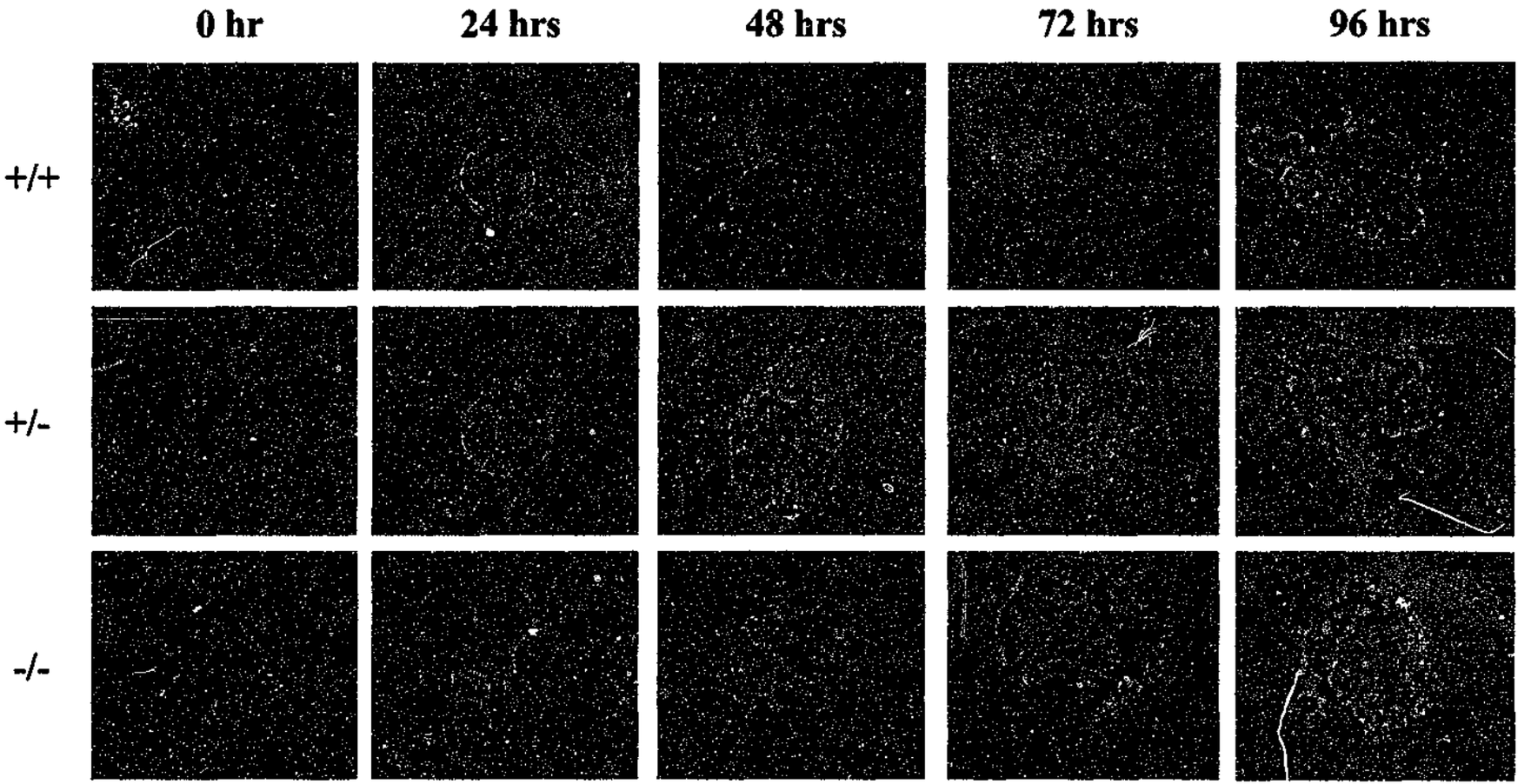


Figure 6.2 Defective growth of *Elf5*$^{-/-}$ blastocysts *in vitro*. **(a)** Example #1. Blastocysts were removed at E3.5 from intercrossed *Elf5*$^{+/-}$ female and cultured for four days in poly-L-lysine-coated petri dishes in ES medium, in 5% CO_2 at 37°C. All pictures are 400×. The length of culturing time is indicated at the top, and the genotype is indicated on the left. Only one example is shown for each genotype. All blastocysts were hatched from zona pellucida. All *Elf5*$^{+/+}$ blastocysts developed outgrowths. In contrast, *Elf5*$^{-/-}$ blastocysts retained their blastocyst appearance after 72 hours in culture, and no obvious trophoblast giant cells were found around the ICM.

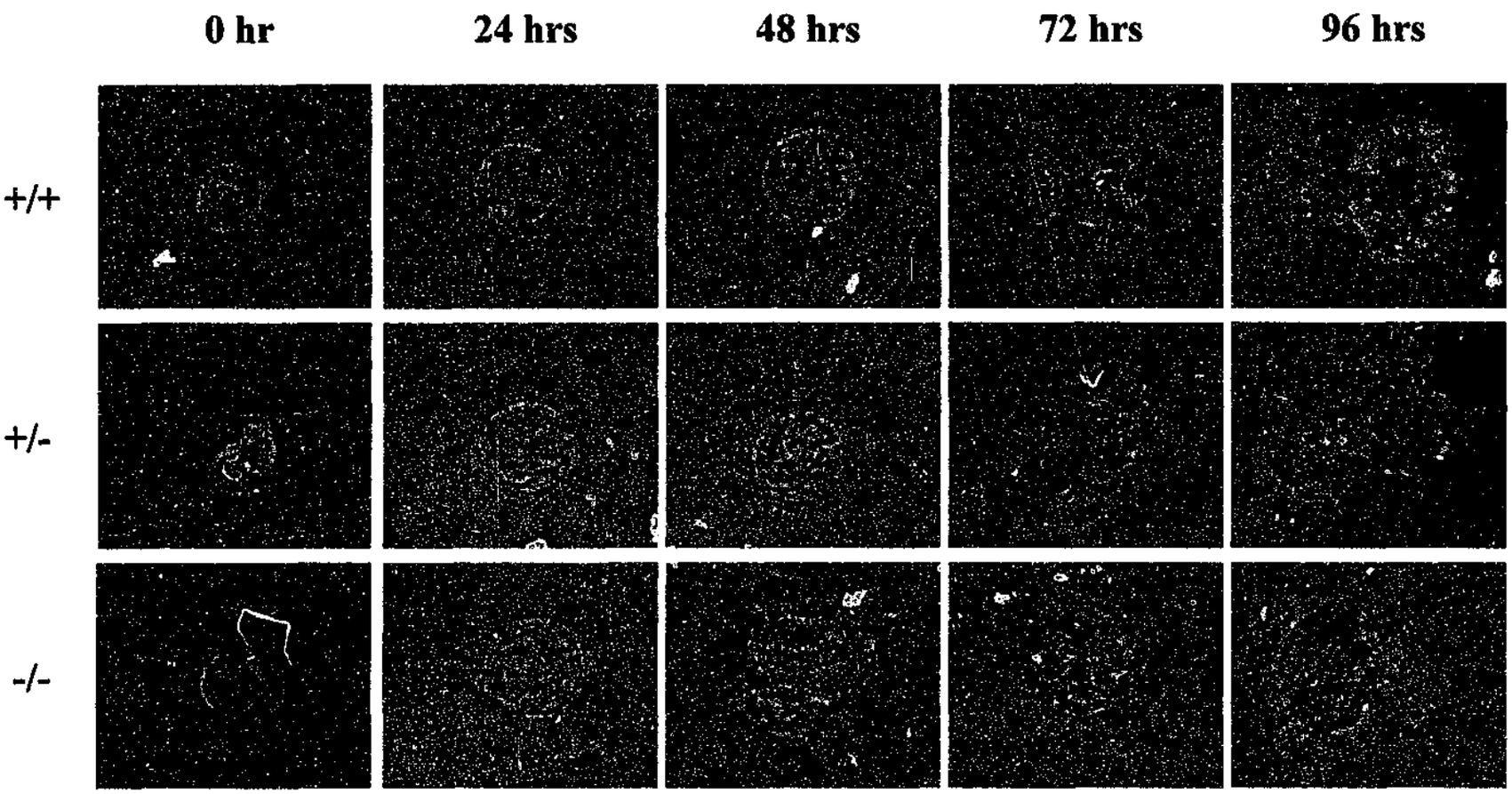


Figure 6.2 Defective growth of *Elf5*$^{-/-}$ blastocysts *in vitro*. **(b)** Example #2. Blastocysts were removed at E3.5 from intercrossed *Elf5*$^{+/-}$ female and cultured for four days in poly-L-lysine-coated petri dishes in ES medium, in 5% CO_2 at 37°C. All pictures are 400×. The length of culturing time is indicated at the top, and the genotype is indicated on the left. Only one example is shown for each genotype. All blastocysts were hatched from zona pellucida. All *Elf5*$^{+/+}$ blastocysts developed outgrowths. In contrast, *Elf5*$^{-/-}$ blastocysts retained their blastocyst appearance after 48-72 hours in culture, and no obvious trophoblast giant cells were found around the ICM.

6.3 Discussion

Genotyping results of mice derived from two independent *Elf5*$^{+/-}$ ES clones demonstrated that Elf5 is essential during mouse embryogenesis (Table 5.2). We investigated the time of death by genotyping neonates and embryos from the *Elf5*$^{+/-}$ intercrosses. No *Elf5*$^{-/-}$ animals were found from embryonic day 7.5 to birth, however, *Elf5*$^{-/-}$ embryos were identified in the expected Mendelian ratio at the blastocyst stage (E3.5) (Table 6.1). These results indicate that *Elf5*$^{-/-}$ embryos die between E3.5-7.5, during the period of embryonic implantation. Of nine Ets family members that have been disrupted in mice, *Pu.1*$^{-/-}$, *Tel*$^{-/-}$, *Fli1*$^{-/-}$ and *Elf3*$^{-/-}$ mice were embryonic lethal around E16.5, E10.5-11.5, E11-12.5 and E11.5-12.5, respectively (Scott *et al.*, 1994; Wang *et al.*, 1997; Hart *et al.*, 2000; Spyropoulos *et al.*, 2000; Ms. A.Y.N. Ng unpublished results, personal communication). In contrast, *Ets2*$^{-/-}$ embryos died before E8.5 due to growth arrest of the ectoplacental cone. In addition, break down in the vascular interactions between the embryo and the maternal circulation also contributes to this embryonic lethality. The disruption of Ets2 resulted in apoptosis of the embryonic ectoderm (Yamamoto *et al.*, 1998). Elf5 appears to be another Ets factor involved in early mouse embryogenesis.

We investigated the growth potential of *Elf5*$^{-/-}$ blastocysts *in vitro*. Blastocysts (E3.5) from heterozygous matings were individually cultured for four days (Figure 6.2 a & b). All blastocysts appeared to be normal morphologically and hatched from the zona pellucida. Unlike *Elf5*$^{+/+}$ blastocysts, *Elf5*$^{-/-}$ blastocysts failed to develop outgrowths. The Elf5-deficient embryos retained their blastocyst appearance up to 72 hours in culture. During this time, the only perceptible change was a dramatic increase in size. In addition, no obvious trophoblast giant cells formed around *Elf5*$^{-/-}$ embryos, which was consistent with the fact that the *Elf5*$^{-/-}$ blastocysts did not attach to the culture dish. Trophectoderm possesses typical epithelial adhesive and junctional cell contacts (Collins and Fleming, 1995). Failure of giant trophoblast formation would certainly lead to embryonic death. We are still yet to determine the status of *Elf5* gene expression in this cell type.

The ICM of *Elf5*$^{+/+}$ blastocysts expanded quickly during culture whilst in contrast, the ICM of *Elf5*$^{-/-}$ blastocysts appeared to be defective in cell proliferation (Figure 6.2 a & b). Some of the *Elf5*$^{-/-}$ blastocysts also started to degenerate after 72 hours in culture. Significantly, numerous attempts to generate *Elf5*$^{-/-}$ ES cells by increasing the concentration of G418 were

unsuccessful (data not shown). Taken together, the results from these *in vitro* experiments suggested that Elf5 is required for ES cell proliferation and differentiation.

We do not know yet if cellular proliferation and differentiation in the trophectoderm and the ICM were blocked due to this Elf5-deficiency. Histological and immunohistochemical analyses of the *Elf5*$^{-/-}$ embryos at the time of embryonic implantation will be necessary to clarify the nature of these embryonic defects at the cellular and molecular levels. Although *Ets2*$^{-/-}$ embryos die around a similar time of mouse embryogenesis, Elf5 and Ets2 play distinctive biological roles since the presence of one Ets factor cannot compensate for the deficiency of the other.

Epithelial cell proliferation and differentiation is tightly controlled/regulated during mammalian development. Misregulation of these processes results in many diseases such as cancer. Studies of the development of the mouse early embryo may provide insight into the regulation of epithelial cell proliferation and differentiation. In addition, the identification of the regulators and the regulating pathways of these processes may be invaluble in future medical research (Collins and Fleming, 1995).

Chapter 7

Elf5 is essential for mouse mammary gland development

7.1 Introduction

A number of epithelial-specific ETS transcription factors (ELF3, ELF5, and ESE3) are expressed in the mammary gland, suggesting potential functions of these proteins in mammary epithelium (Chang *et al.*, 1997; Neve *et al.*, 1998; Oettgen *et al.*, 1999; Kas *et al.*, 2000; also discussed in Chapter 4). We have shown that ELF3 and ELF5 can transactivate a pregnancy- and lactation-associated milk protein (WAP) gene promoter *in vitro* (Thomas *et al.*, 2000). Others have shown that some ETS proteins may be involved in breast cancer (Chang *et al.*, 1997; Atlas *et al.*, 2000; Xing *et al.*, 2000). This chapter describes a severe mammary gland developmental defect observed in pregnant and lactating $Elf5^{+/-}$ females. The offspring from those females died as a result.

In order to understand and place into context the mammary gland developmental defect observed in the $Elf5^{+/-}$ females, what follows is a brief description of the three temporal stages of mammary gland growth and development and some of the genes implicated in this process. The discussion is limited to the genes that when disrupted, demonstrate a mammary gland phenotype similar to that observed in the $Elf5^{+/-}$ females.

7.1.1 The development of the mouse mammary gland

The mammary gland undergoes development at three distinct stages in the life of the mouse; *in utero*, at puberty, and during pregnancy. Before mice reach puberty, there is only a small mammary ductal tree, embedded in a mammary fat pad, around a nipple. The mammary ductal tree can elongate and branch quickly through the entire mammary fat pad under the

stimulation of ovarian hormones. Pregnanacy hormones will then promote alveolar proliferation, and the alveolar epithelial cells from the resulting lobuloalveolar structures will eventually differentiate into secretory epithelial cells at parturition (Robinson *et al.*, 2000). The essential regulatory hormonal factors involved during puberty are estrogen, adrenocorticoid and growth hormone. Estrogen, progesterone and placental lactogen and/or prolactin are essential during pregnancy (Nandi, 1958; Neville and Daniel, 1987). This hormonally-induced massive growth during pregnancy never fully regresses after estrus or weaning (Vonderhaar, 1988).

At the start of puberty (about four weeks of age), proliferation of terminal end buds (TEB) at the end of the growing mammary ducts leads to elongation and branching of the mammary tree (Daniel and Silberstein, 1987). The TEB contains both body cells and cap cells that are the precursors of the mammary epithelial cells and the myoepithelial cells, respectively (Humphreys *et al.*, 1996). The TEBs will eventually disappear when the mammary ductal tree penetrates the entire mammary fat pad. During the extension of the mammary ductal tree, cellular proliferation and death is tightly controlled to achieve ductal morphogenesis (Bresciani, 1968; Humphreys *et al.*, 1996). A single layer of luminal epithelial cells is found lining the internal wall of the ducts whereas the myoepithelial cells surround the primary ducts (Hennighausen and Robinson, 1998).

During pregnancy, alveolar epithelial cells differentiate to become secretory cells, resulting in a number of milk proteins being synthesized under a specific regulating system. WDNM1 is the first milk protein produced at ten days of pregnancy, and it is followed by the production of β-casein and whey acidic protein (WAP) at twelve days and fifteen days of pregnancy, respectively. Therefore, these proteins are useful in determining the differentiation state of the mammary epithelium (Robinson *et al.*, 1995).

7.1.2 Genes that are involved in the development of the mouse mammary gland

7.1.2.1 *Prolactin receptor (PRLR)*

Prolactin is the major ligand for the prolactin receptor (PRLR). Most of the six to eight week-old $PRLR^{+/-}$ females exhibited lactation failure due to the defective development of the mammary gland after their first pregnancy, whilst $PRLR^{-/-}$ females were sterile (Ormandy *et*

al., 1997; Bole-Feysot *et al.*, 1998). Only 14% of the pups born to $PRLR^{+/-}$ females survived. A correlation between the extent of mammary gland development and the lactation capability was observed in $PRLR^{+/-}$ females by histological analysis. The non-lactating $PRLR^{+/-}$ mothers had severely underdeveloped mammary glands in comparison to the lactating $PRLR^{+/+}$ mothers. A second pregnancy in $PRLR^{+/-}$ mothers had improved the overall pup survival rate to 96%. In addition, aged $PRLR^{+/-}$ mothers (at 20 weeks of age) had an overall pup survival rate of 76% after their first pregnancy, and the overall survival rate of the pups from these females after their second pregnancy increased to 90%. These results suggested that one copy of the functional *PRLR* gene might be insufficient for mammary gland epithelial cell proliferation during pregnancy and the postpartum period (Ormandy *et al.*, 1997; Bole-Feysot *et al.*, 1998), and that an alternative pathway in lobuloalveolar proliferation and differentiation may compensate for the deficiency of PRLR in the subsequent pregancies.

7.1.2.2 *Stat5a* and *Stat5b*

Binding of PRL to its receptor PRLR leads to the activation of the JAK2, Fyn, and MAP kinase systems, which in turn, stimulates cell growth (Das and Vonderhaar, 1995; Lebrun *et al.*, 1995a, b; Bole-Feysot *et al.*, 1998). Stat proteins are involved in PRLR signaling and PRL-induced mammary development (Darnell *et al.*, 1994; Horseman and Yu-Lee, 1994; Wakao *et al.*, 1994; Liu *et al.*, 1997; Bole-Feysot *et al.*, 1998). Stat5 protein is shown to be activated by tyrosine phosphorylation in response to PRL stimulation, and this phosphorylation may be important in mammary gland development and milk protein synthesis (Gouilleux *et al.*, 1994; Wakao *et al.*, 1994). Stat5a and Stat5b are encoded by two genes with 96% similarity (Liu *et al.*, 1995). Although $Stat5a^{-/-}$ females are fertile, these mice had defects in mammary lobuloalveolar proliferation and terminal alveolar differentiation resulting in lactation failure (Liu *et al.*, 1997). In contrast, although the fertility of $Stat5b^{-/-}$ females was greatly reduced, those mice that maintained their pregnancy, delivered normal sized litters and were able to lactate, indicating that Stat5b is not required for normal mammary development and function (Udy *et al.*, 1997; Teglund *et al.*, 1998). In nonfunctional mammary tissues of $Stat5a^{-/-}$ females, Stat5b activity and *WAP* gene expression was diminished (Liu *et al.*, 1997). A partially functional gland is observed in $Stat5a^{-/-}$ females after three pregnancies and extensive stimulation by suckling. Notably, Stat5b activity and *WAP* gene expression in the mammary glands of these females was

restored after parturition. These results suggested that Stat5b could substitute for Stat5a (Liu *et al.*, 1998). Teglund *et al.*, (1998) showed that pregnant mice lacking both Stat5a and Stat5b failed to reach term, however, their mammary gland development during puberty was normal.

7.1.2.3 *Galanin*

Galanin, a neuropeptide, is expressed in the lactotrophs of the rodent pituitary and hypothalamus where PRL is also expressed. *Galanin*$^{-/-}$ females are fertile, but these mice fail to respond to pregnancy hormones and are unable to develop mature and functional lactating mammary glands. Loss of Galanin results in reduced *PRL* expression in lactotrophs, and failed proliferation of the lactotrophs (Wynick *et al.*, 1998). Therefore, Galanin may have a regulatory role in PRL-mediated mammary gland development.

7.1.2.4 *Osteoprotegerin (OPGL)* and *receptor activator of NFκB (RANK)*

The TNF family member, osteoprotegerin-ligand (OPGL), and its receptor, RANK (receptor activator of NFκB), are involved in osteoclast-associated bone metabolism. Interestingly, despite normal mammary gland development before and during early pregnancy, *OPGL*$^{-/-}$ and *RANK*$^{-/-}$ females failed to develop functional lactating mammary glands at parturition due to defects in terminal alveolar differentiation. However, local administration of rOPGL restored mammary gland functionality in *OPGL*$^{-/-}$ females, but not in *RANK*$^{-/-}$ females (Fata *et al.*, 2000). Therefore, this ligand-receptor relationship is also required for the late pregnancy-associated mammary gland development.

7.1.2.5 *Cyclin D1* and *cyclin E*

Cyclins and their associated cyclin-dependent kinases (CDKs) are critical factors in regulating the mammalian cell cycle (Sherr, 1994). In particular, the G1 phase cyclins, cyclin Ds and cyclin E, can reverse the cell cycle blockage effect of the *Rb* tumor suppressor gene in response to extracellular growth signals (Weinberg, 1995; Sherr, 1996). Interestingly, although *cyclin D1*$^{-/-}$ mice were growth retarded and had neurological abnormalities, these mice were viable and fertile (Fantl *et al.*, 1995; Sicinski *et al.*, 1995). In addition, the pregnant *cyclin D1*$^{-/-}$ females failed to undergo lobuloalveolar development in

their mammary glands resulting in lactation failure. (Fantl *et al.*, 1995; Sicinski *et al.*, 1995; Fantl *et al.*, 1999). The mammary alveoli buds failed to proliferate and differentiate in response to normal levels of pregnancy hormones. Mammary epithelium transplantation experiments eliminated any extrinsic cause for the defect in the *cyclin D1*$^{-/-}$ mammary gland (Fantl *et al.*, 1999). Interestingly, the mammary developmental defect was rescued by replacing the *cyclin D1* gene with *cyclin E* (Geng *et al.*, 1999). Transplantation of *cyclin D1*$^{-/-}$ mammary epithelium harbouring a cyclin E knock-in into wildtype females resulted in full lobuloalveolar development.

7.1.2.6 *LAR*

The LAR receptor-like protein tyrosine phosphatase is thought to be involved in the transduction of extracellular signals. Interestingly, *LAR*$^{-/-}$ females failed to lactate at parturition due to terminal mammary alveoli differentiation failure (Schaapveld *et al.*, 1997). Therefore, in response to pregnancy hormone stimulation, LAR may be responsible for the activation of the JAK-STAT or other signaling pathways in mammary gland development (Hennighausen and Robinson, 1998).

7.1.2.7 *BRCA1*

Familial breast and ovarian cancer patients frequently carry mutations of *BRCA1*, a tumor suppressor gene (Casey, 1997; Hill *et al.*, 1997). Marquis *et al.*, (1995) demonstrated that mouse mammary epithelial cells expressed high levels of *Brca1* during pregnancy but lower levels during lactation. *Brca1*$^{-/-}$ females were fertile, but the mammary ductal trees of these mice were incompletely formed and did not fully penetrate the mammary fat pad during pregnancy. Although the terminal mammary alveolar differentiation in *Brca1*$^{-/-}$ females was normal and these mice were capable of lactation, ~20% of alveolar epithelial cells were apotosing. In addition, these *Brca1*$^{-/-}$ females developed mammary gland tumors that were exacerbated by the loss of p53 (Xu *et al.*, 1999).

7.1.2.8 *Plasminogen (Plg)*

Plasminogen (Plg) is implicated in tissue remodeling (Sternlicht and Werb, 1999). Lund *et al.*, (2000) demonstrated that the mammary glands from *Plg*$^{-/-}$ females had delayed ductal

branching that led to smaller mammary glands, and consequently ~28% of the females failed to lactate at parturition. In addition, the mammary gland involution in $Plg^{-/-}$ females was abnormal resulting in the accumulation of fibrotic stroma. Interestingly, 7% of the $Plg^{+/-}$ females also displayed a similar mammary phenotype but to a lesser extent (Lund *et al.*, 2000).

7.2 Results

Elf5$^{+/-}$ females, derived from *Elf5*$^{+/-}$ ES cell clone #130, were fertile and give birth to litters of morphologically normal pups whose numbers were comparable to those of *Elf5*$^{+/+}$ females. However, despite normal nursing and mothering characteristics displayed by *Elf5*$^{+/-}$ females, most of their first litter pups died within 24 hours of birth, and virtually the entire litter had perished by 48 hours. All pups were observed to attach to the nipple and suckle, but newborns of *Elf5*$^{+/-}$ females died of starvation and dehydration. Lethality of the newborns was independent of their genotype (*Elf5*$^{+/+}$ or *Elf5*$^{+/-}$). Examination of one day old pups revealed that they lacked milk in their stomachs. These results suggested that *Elf5*$^{+/-}$ females might have a mammary gland defect.

No. of Elf5$^{+/-}$ mothers	First pregnancy			Subsequent pregnancies	
	All died	Some died	All survived	Death of some or entire litter	Always died
19	√			N/A	N/A
1			√	N/A	N/A
7	√				√
15	√			√	
3		√		√	
4			√	√	

Table 7.1 *Elf5*$^{+/-}$ mothers, derived from *Elf5*$^{+/-}$ ES cell clone #130, failed to keep their pups. N/A: not applicable.

We studied 49 *Elf5*$^{+/-}$ mothers. Of these, only eight supported some or all of their first litter pups. Nineteen *Elf5*$^{+/-}$ females were culled after the death of their first litter. Seven *Elf5*$^{+/-}$ females failed to support any of their pups after multiple pregnancies. In the subsequent pregnancies of eighteen *Elf5*$^{+/-}$ females that lost some or all of their first litter pups, this lactation problem seemed to be less severe, although unpredictably, part of, or entire litters still died occasionally. Of five *Elf5*$^{+/-}$ females that supported all of their first litter pups, one

was culled and the remaining four all experienced problems with lactation in subsequent pregnancies, resulting in death of the newborn pups.

Two *Elf5*$^{+/-}$ females, derived from *Elf5*$^{+/-}$ ES cell clone #1, were also observed for their lactating abilities. Both females failed to lactate after their first pregnancy, and pup death due to starvation and dehydration was also observed in some of their subsequent pregnancies.

Examination of whole mount mammary glands from age-matched virgin *Elf5*$^{+/+}$ and *Elf5*$^{+/-}$ females demonstrated that elongation and extension of the mammary ductal tree, as well as ductal side branching were comparable (Figures 7.1-7.3). Therefore, we hypothesized that loss of one allele of the *Elf5* gene might affect mammary gland development during pregnancy.

In *Elf5*$^{+/+}$ females at 18.5 days of pregnancy, proliferation of mammary ductal epithelium and sprouting of alveolar buds was induced. Proliferation and terminal differentiation of the alveolar buds followed, resulting in fully developed lobuloalveolar structures (Figures 7.4-7.6 a-d). In *Elf5*$^{+/-}$ females, increased ductal side branching and formation of the initial alveolar buds also occurred normally (Figures 7.4-7.6 e and f), indicating that loss of one copy of the *Elf5* gene did not disrupt the proliferation of ductal epithelium and sprouting of alveolar buds. However, differentiation and expansion of the alveolar buds into mature lobuloalveolar mammary tissues was severely impaired in *Elf5*$^{+/-}$ females (Figures 7.4-7.6 e-h). Although some proliferation and terminal differentiation of the alveolar buds in very localized regions was observed in some mammary glands in a small number of pregnant *Elf5*$^{+/-}$ females (Figure 7.4 h), these glands were in no circumstance comparable to the fully developed wildtype glands (Figures 7.4-7.6 a-d).

The mammary gland defects in pregnant *Elf5*$^{+/-}$ females remained evident during lactation. Full mammary growth was attained in lactating *Elf5*$^{+/+}$ females one day postpartum. The alveoli subunits expanded and filled with milk indicating that a functional secretory state of the mammary gland was reached (Figures 7.7-7.9 a-d). In contrast, the alveolar buds in *Elf5*$^{+/-}$ mammary glands did not proliferate and differentiate, and milk was not produced in these defective glands (Figures 7.7-7.9 e-h). Dilation of primary ducts was also impaired in some *Elf5*$^{+/-}$ females (Figures 7.7-7.8 g).

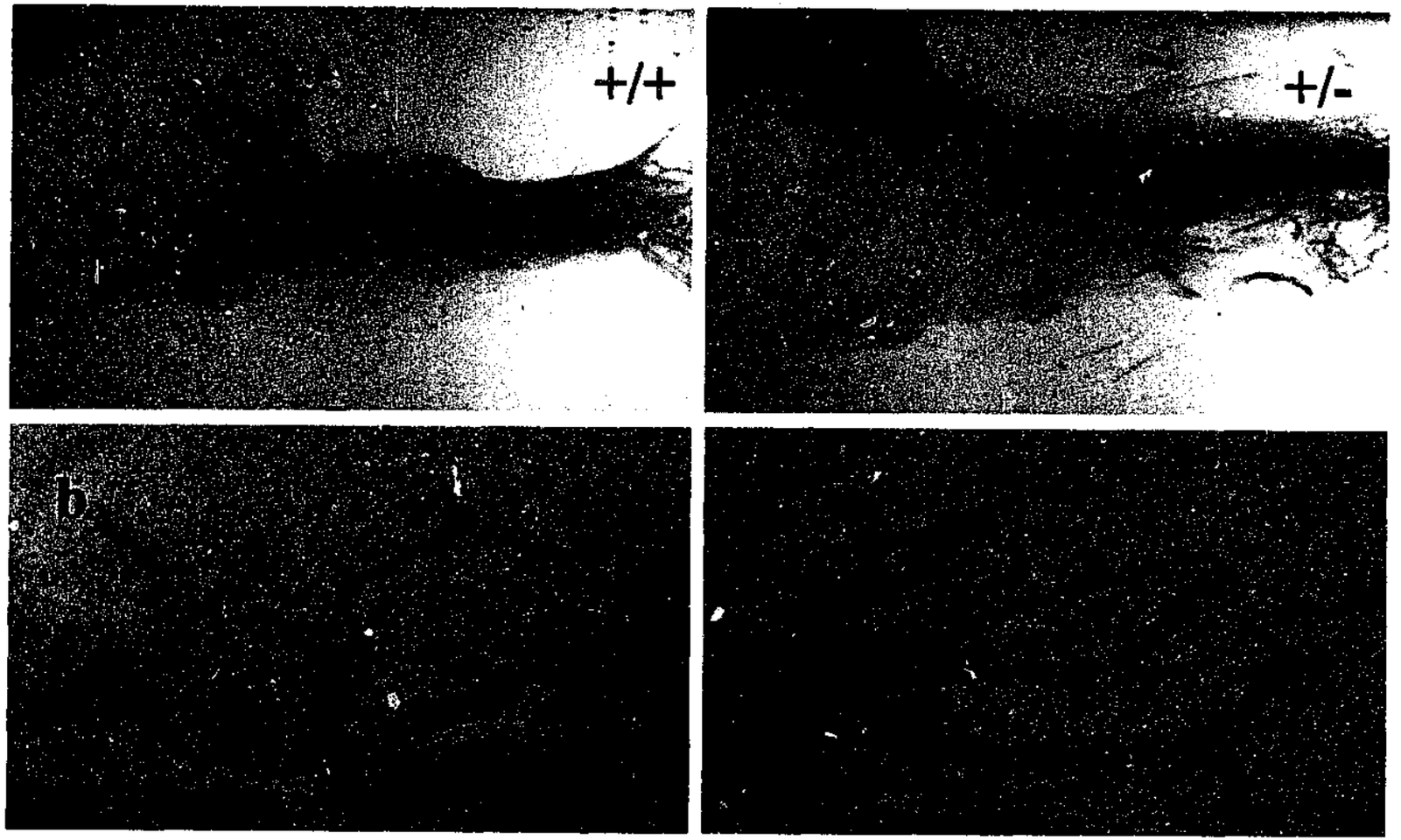


Figure 7.1 Mammary gland development is normal in virgin *Elf5*$^{+/-}$ females (example #1). Mammary ductal elongation and extension was comparable in age-matched *Elf5*$^{+/+}$ and *Elf5*$^{+/-}$ females (48 days old). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (c-d) virgin females. Magnifications: a & c ×18.9, b & d ×120. The whole mounts of inguinal mammary glands were stained with carmine red.

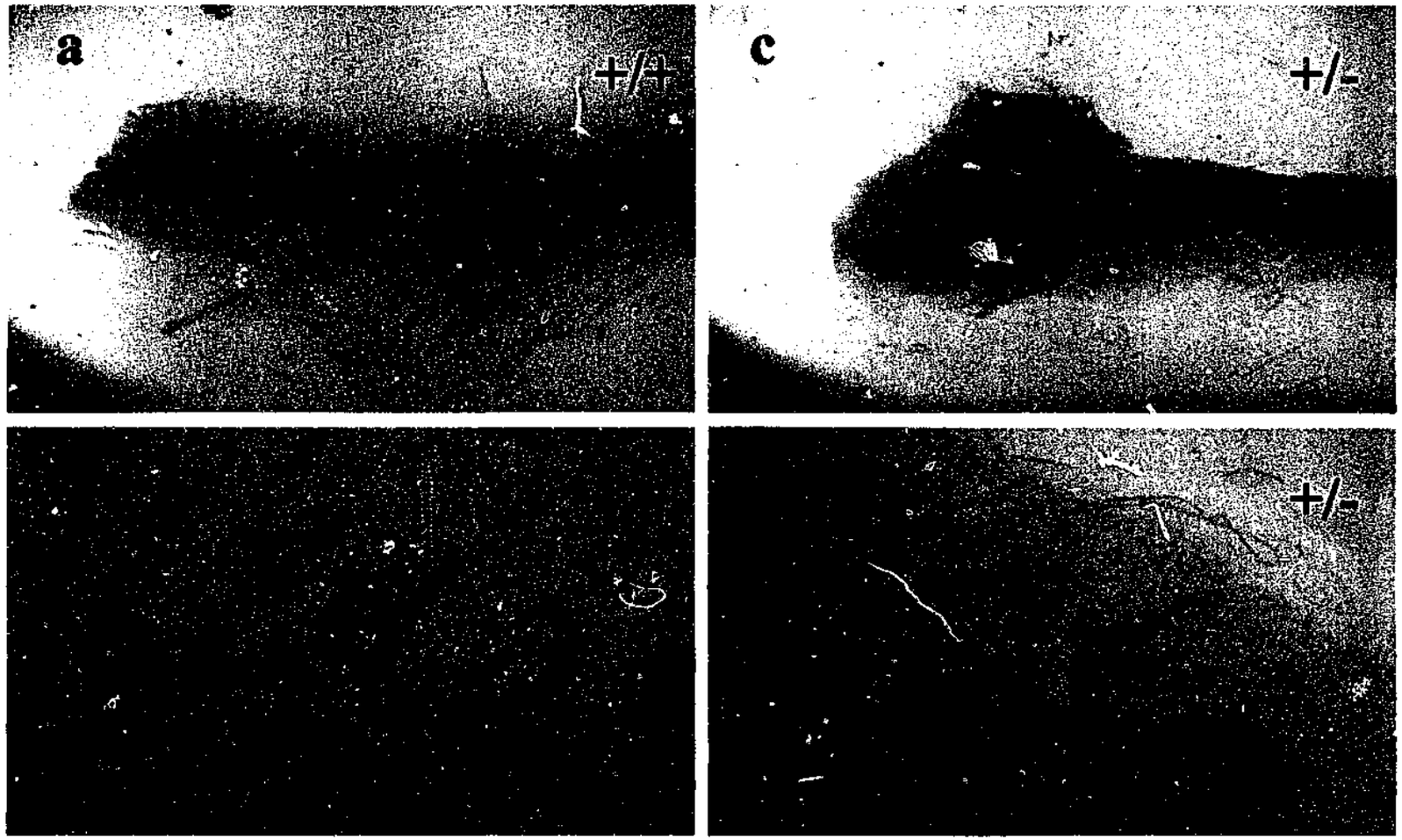


Figure 7.2 Mammary gland development is normal in virgin *Elf5*$^{+/-}$ females (example #2). Mammary ductal elongation and branching was comparable in age-matched *Elf5*$^{+/+}$ and *Elf5*$^{+/-}$ females (131 days old). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (c-d) virgin females. Magnifications: a & c ×18.9, b & d ×120. The whole mounts of inguinal mammary glands were stained with carmine red.

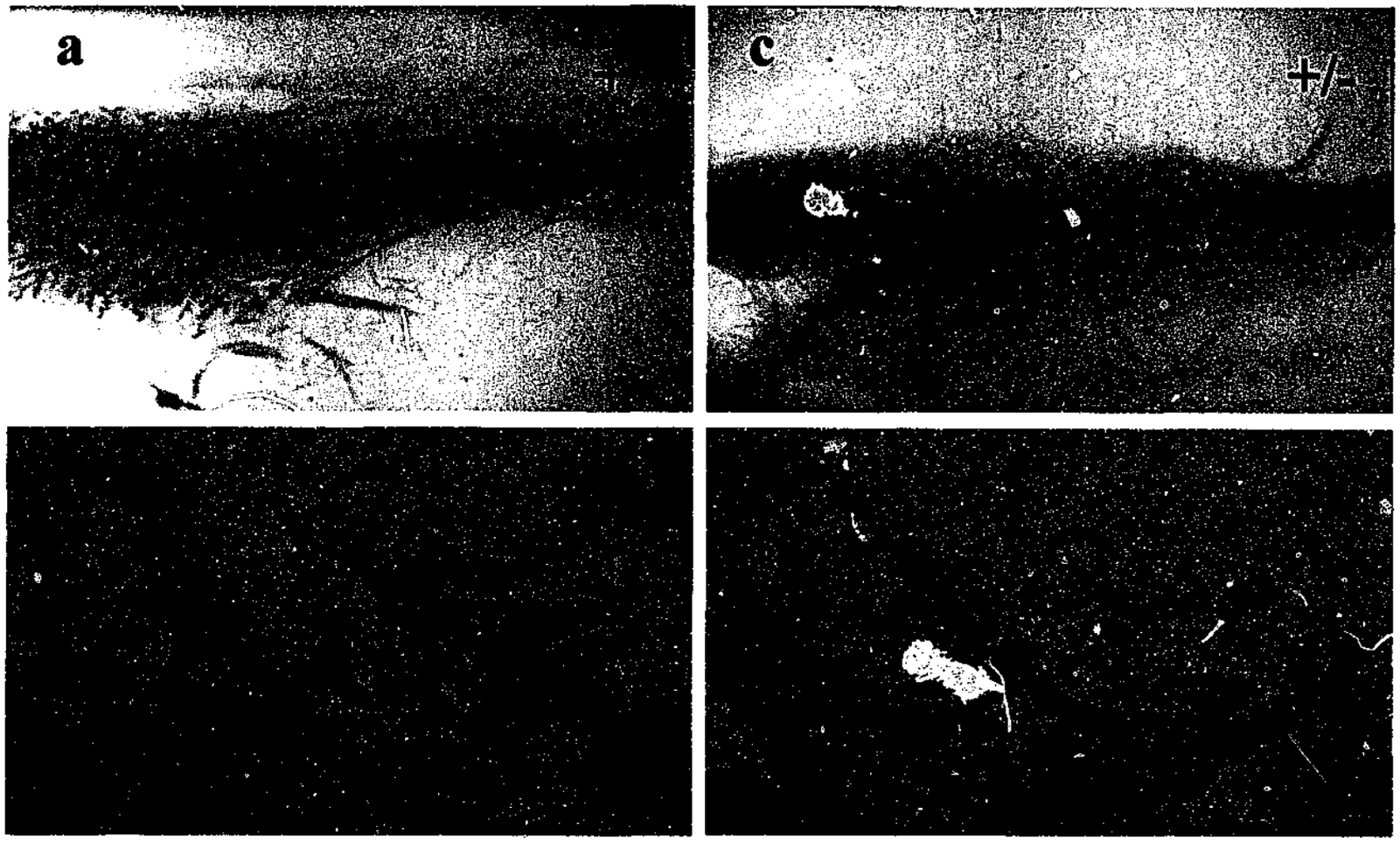


Figure 7.3 Mammary gland development is normal in virgin *Elf5*$^{+/-}$ females (example #3). Mammary ductal elongation and branching was comparable in age-matched *Elf5*$^{+/+}$ and *Elf5*$^{+/-}$ females (176 days old). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (c-d) virgin females. Magnifications: a & c ×18.9, b & d ×120. The whole mounts of inguinal mammary glands were stained with carmine red.

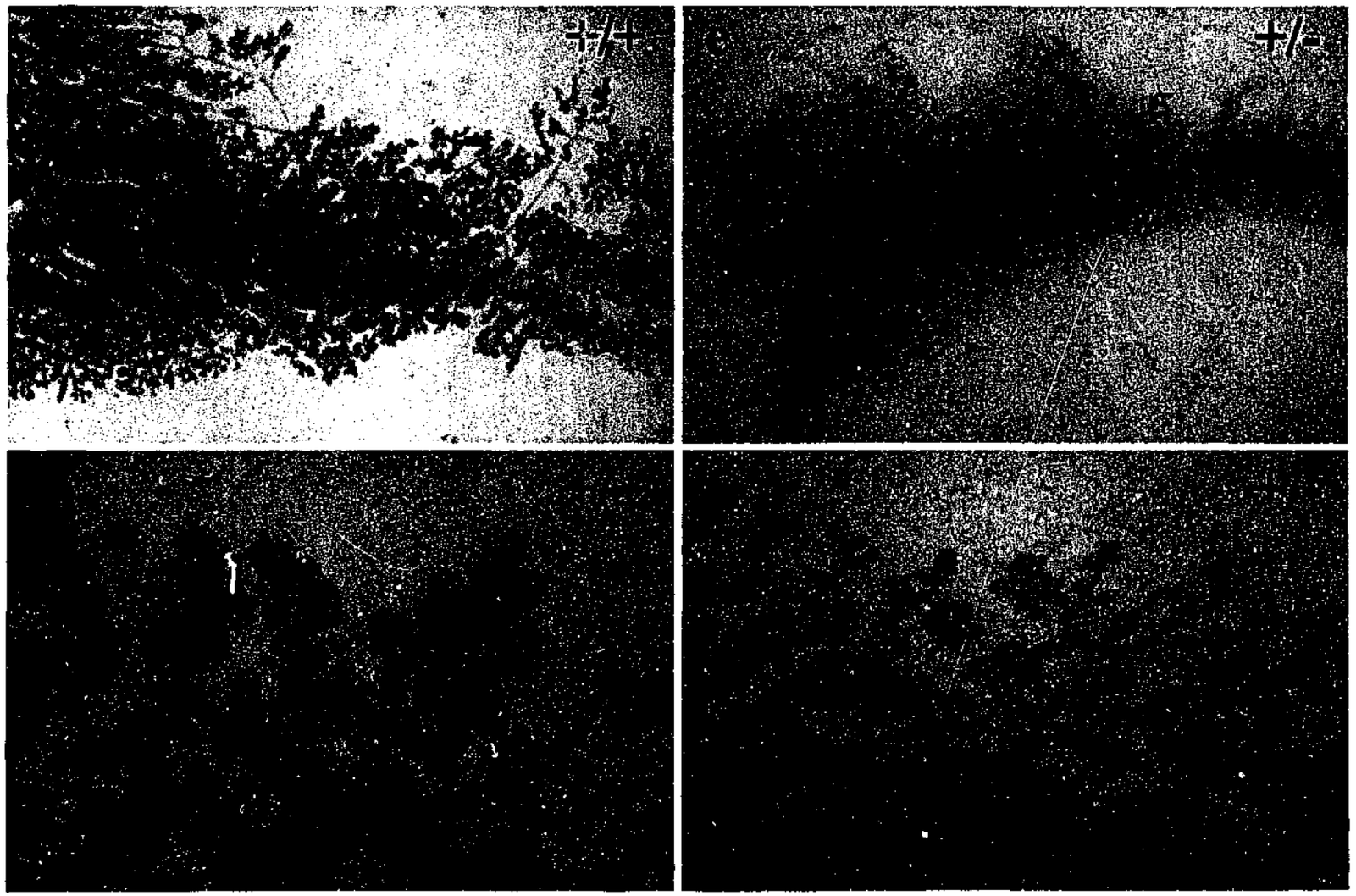


Figure 7.4 Mammary gland development is defective in pregnant *Elf5*$^{+/-}$ females (example #1). Mammary alveolar proliferation and terminal differentiation was severely impaired in an 18.5 day pregnant *Elf5*$^{+/-}$ female (118 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (121 days old, a-d). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (e-f) females during their first pregnancy. Magnifications: a & e ×18.9, b & f ×120. The whole mounts of inguinal mammary glands were stained with carmine red. Part of gland was removed for histological analysis. (Continued on next page)

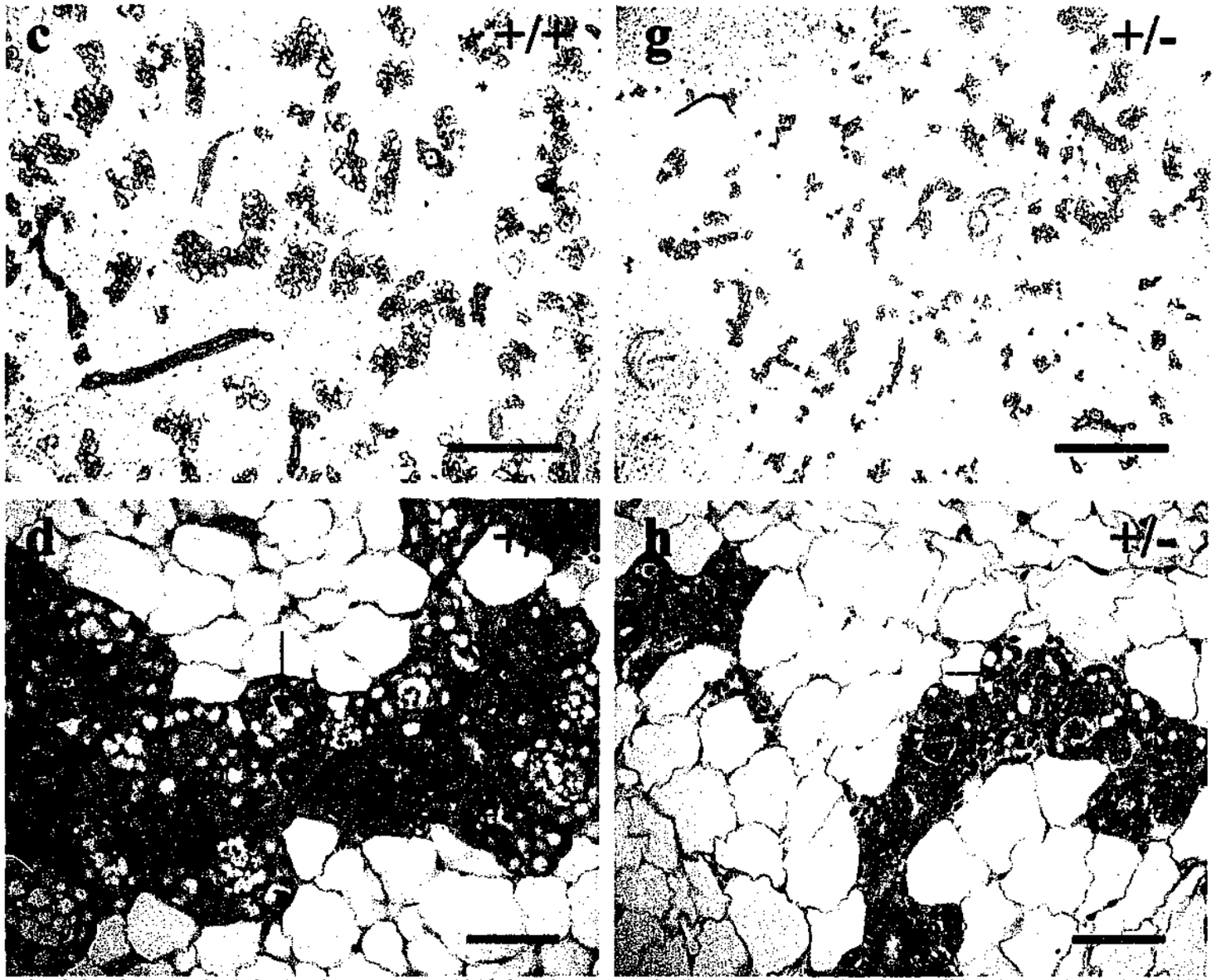


Figure 7.4 Mammary gland development is defective in pregnant *Elf5*$^{+/-}$ females (example #1). Mammary alveolar proliferation and terminal differentiation was severely impaired in an 18.5 day pregnant *Elf5*$^{+/-}$ female (118 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (121 days old, a-d). Histological analyses of mammary tissue of *Elf5*$^{+/+}$ (c-d) and *Elf5*$^{+/-}$ (g-h) females during their first pregnancy. Arrows point to the alveoli. The bar denotes 0.5 mm (c & g) or 0.05 mm (d & h). The paraffin histological sections (c-d & g-h) of inguinal mammary glands were stained with hematoxylin and eosin. (Continued from previous page)

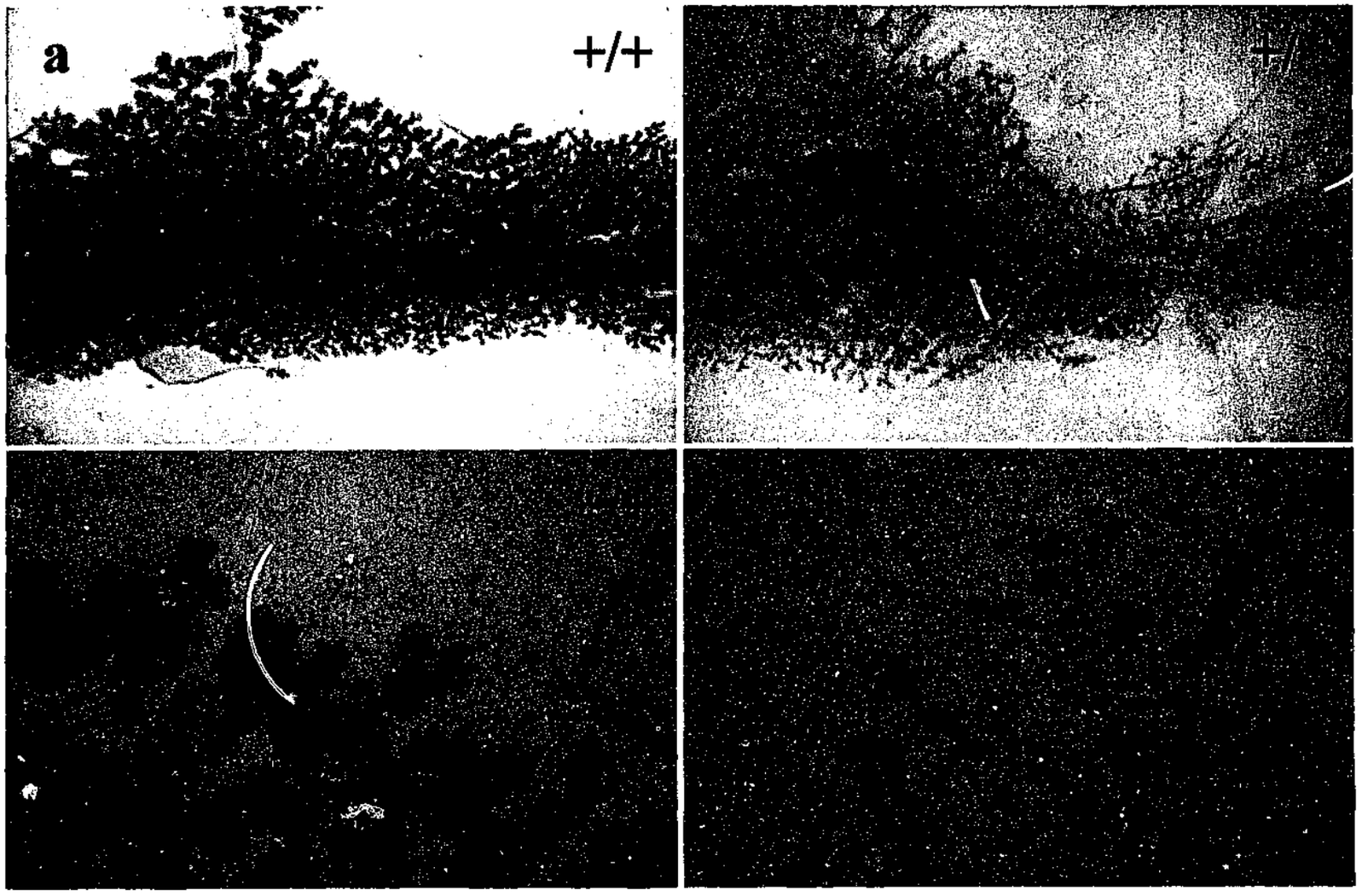


Figure 7.5 Mammary gland development is defective in pregnant *Elf5*$^{+/-}$ females (example #2). Mammary alveolar proliferation and terminal differentiation was severely impaired in an 18.5 day pregnant *Elf5*$^{+/-}$ female (116 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (136 days old, a-d). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (e-f) females during their first pregnancy. Magnifications: a & e ×18.9, b & f ×120. The whole mounts of inguinal mammary glands were stained with carmine red. Part of gland was removed for histological analysis. (Continued on next page)

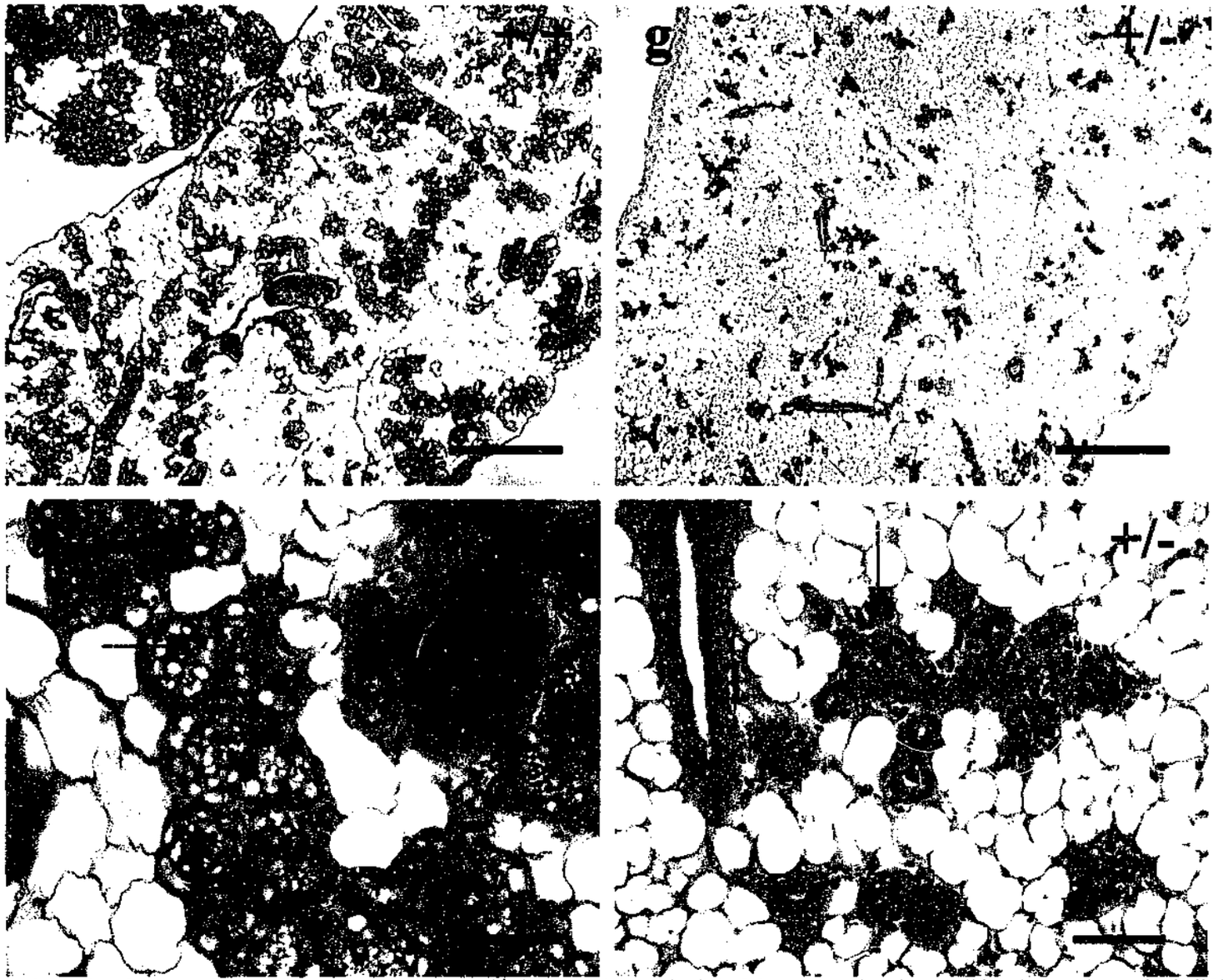


Figure 7.5 Mammary gland development is defective in pregnant *Elf5*$^{+/-}$ females (example #2). Mammary alveolar proliferation and terminal differentiation was severely impaired in an 18.5 days pregnant *Elf5*$^{+/-}$ female (116 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (136 days old, a-d). Histological analyses of mammary tissue of *Elf5*$^{+/+}$ (c-d) and *Elf5*$^{+/-}$ (g-h) females during their first pregnancy. Arrows point to the alveoli. The bar denotes 0.5 mm (c & g) or 0.05 mm (d & h). The paraffin histological sections (c-d & g-h) of inguinal mammary glands were stained with hematoxylin and eosin. (Continued from previous page)

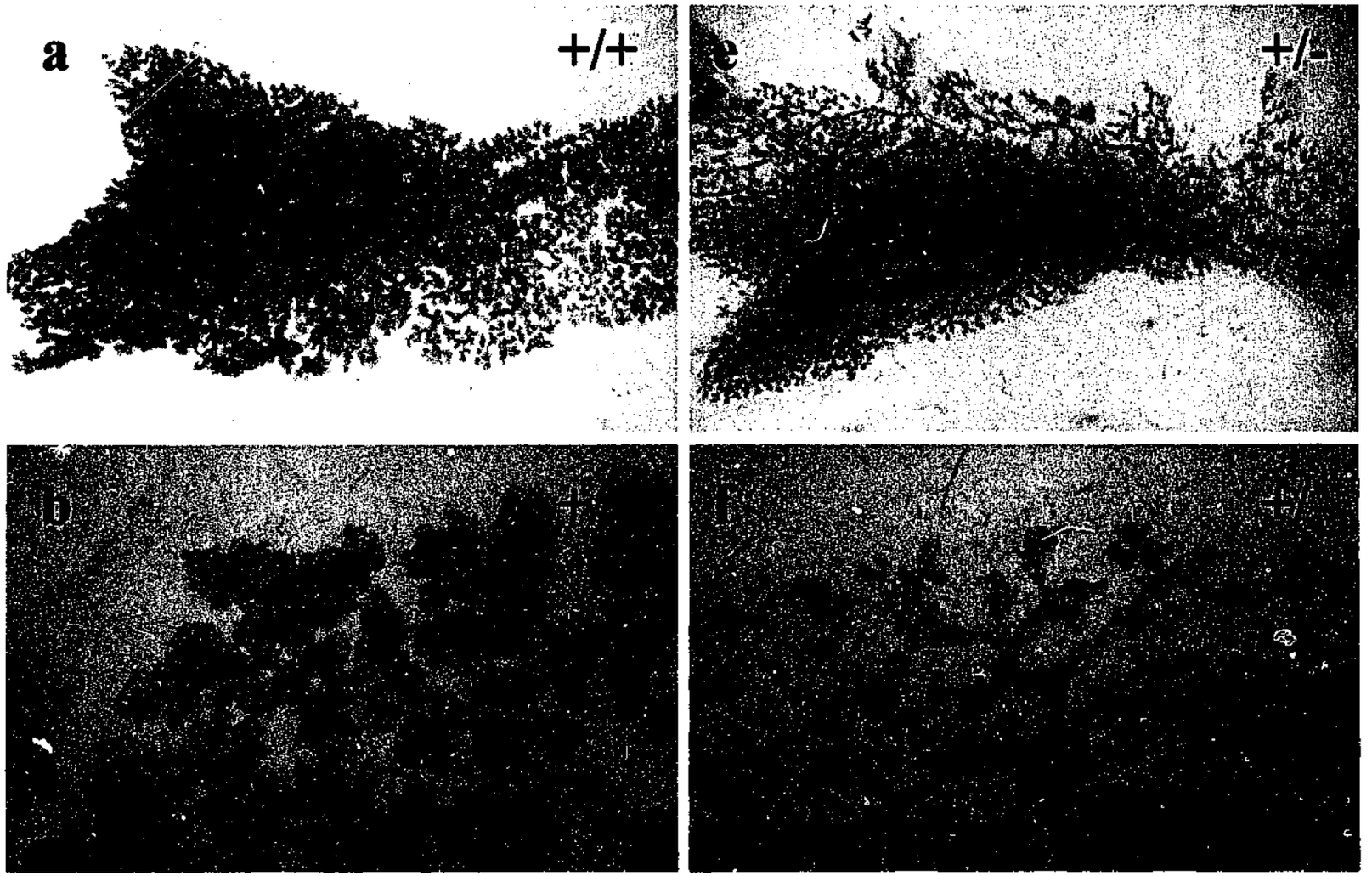


Figure 7.6 Mammary gland development is defective in pregnant *Elf5*$^{+/-}$ females (example #3). Mammary alveolar proliferation and terminal differentiation was severely impaired in an 18.5 day pregnant *Elf5*$^{+/-}$ female (119 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (119 days old, a-d). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (e-f) females during their first pregnancy. Magnifications: a & e ×18.9, b & f ×120. The whole mounts of inguinal mammary glands were stained with carmine red. Part of gland was removed for histological analysis. (Continued on next page)

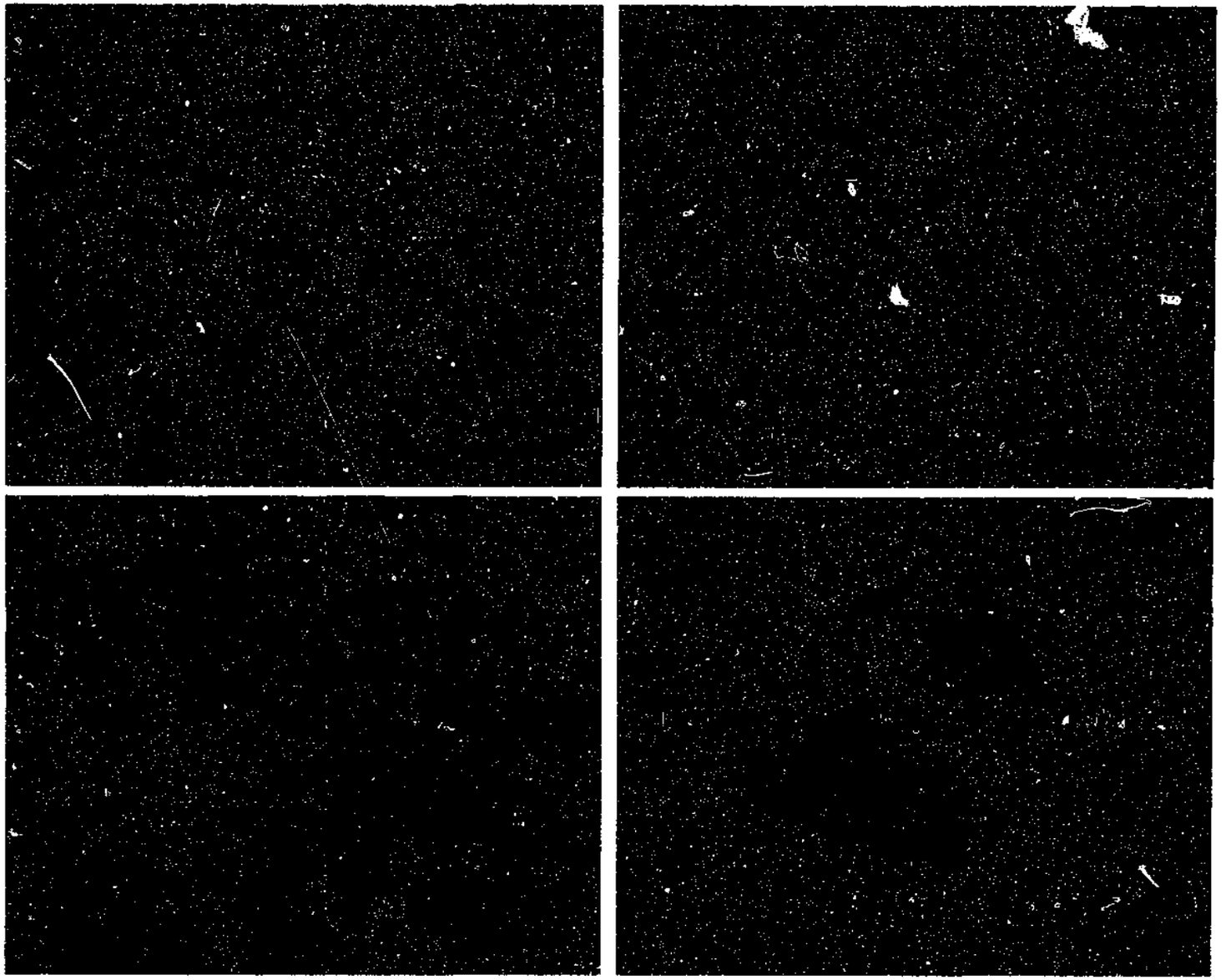


Figure 7.6 Mammary gland development is defective in pregnant *Elf5*$^{+/-}$ females (example #3). Mammary alveolar proliferation and terminal differentiation was severely impaired in an 18.5 day pregnant *Elf5*$^{+/-}$ female (119 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (119 days old, a-d). Histological analyses of mammary tissue of *Elf5*$^{+/+}$ (c-d) and *Elf5*$^{+/-}$ (g-h) females during their first pregnancy. Arrows point to the alveoli. The bar denotes 0.5 mm (c & g) or 0.05 mm (d & h). The paraffin histological sections (c-d & g-h) of inguinal mammary glands were stained with hematoxylin and eosin. (Continued from previous page)

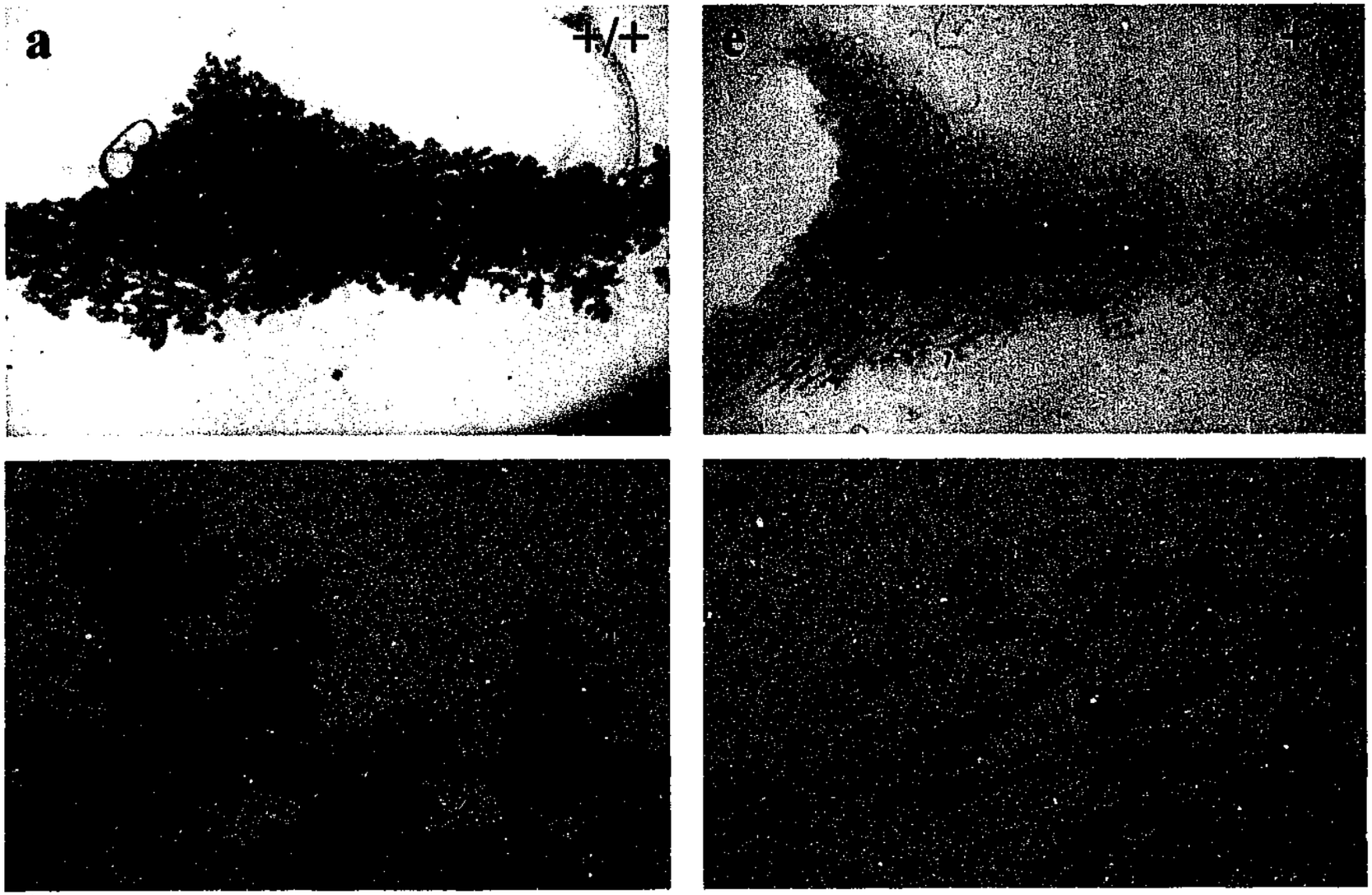


Figure 7.7 Mammary gland development is defective in lactating *Elf5*$^{+/-}$ females (example #1). Mammary alveolar proliferation and terminal differentiation was severely impaired in an one day postpartum *Elf5*$^{+/-}$ female (127 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (127 days old, a-d). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (e-f) females after their first pregnancy. Magnifications: a & e ×18.9, b & f ×120. The whole mounts of inguinal mammary glands were stained with carmine red. Part of gland was removed for histological analysis. (Continued on next page)

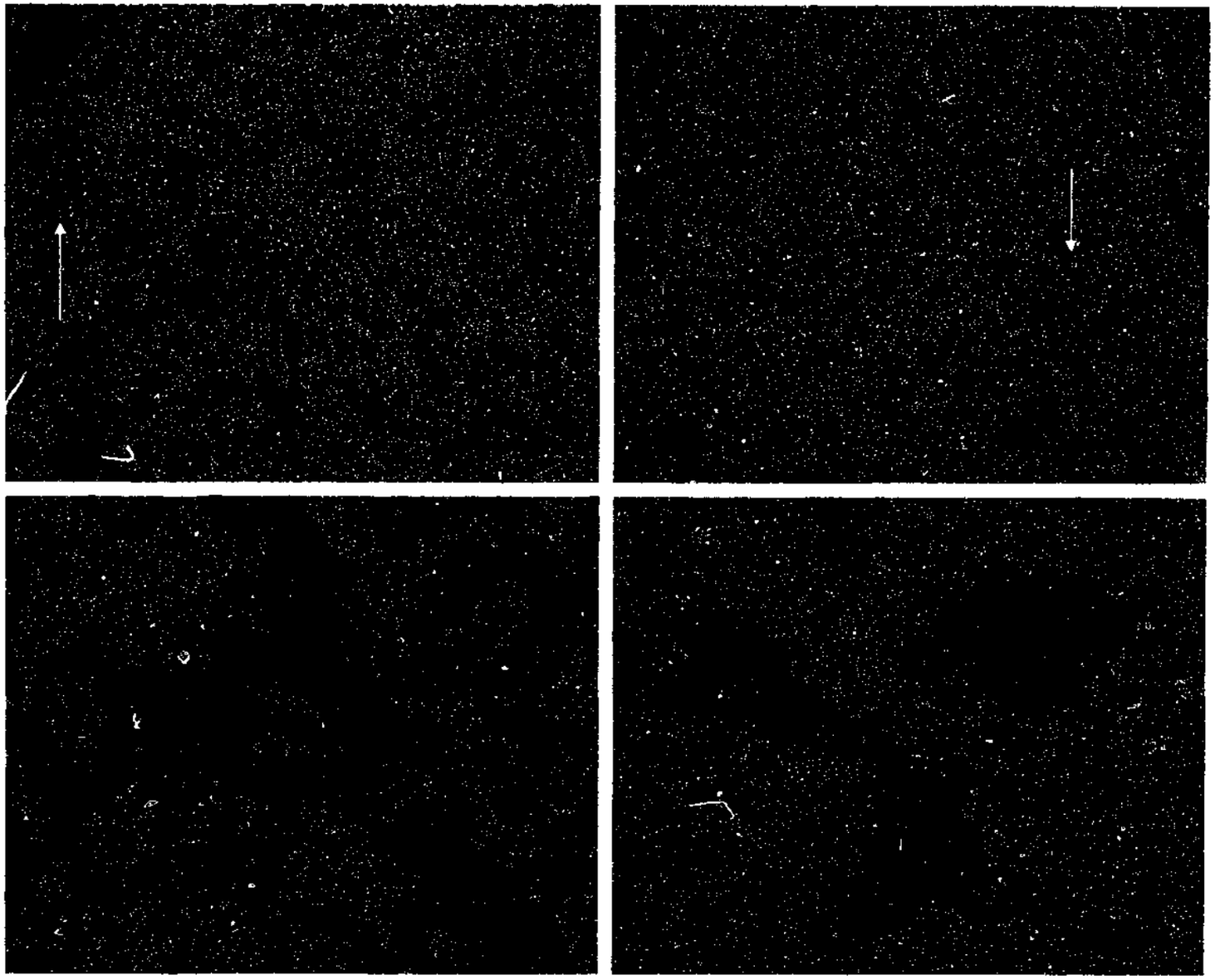


Figure 7.7 Mammary gland development is defective in lactating *Elf5*$^{+/-}$ females (example #1). Mammary alveolar proliferation and terminal differentiation was severely impaired in an one day postpartum *Elf5*$^{+/-}$ female(127 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (127 days old, a-d). Histological analyses of mammary tissue of *Elf5*$^{+/+}$ (c-d) and *Elf5*$^{+/-}$ (g-h) females after their first pregnancy. White arrows point to the ducts, and black arrows point to the alveoli. The bar denotes 0.5 mm (c & g) or 0.05 mm (d & h). The paraffin histological sections (c-d & g-h) of inguinal mammary glands were stained with hematoxylin and eosin. (Continued from previous page)

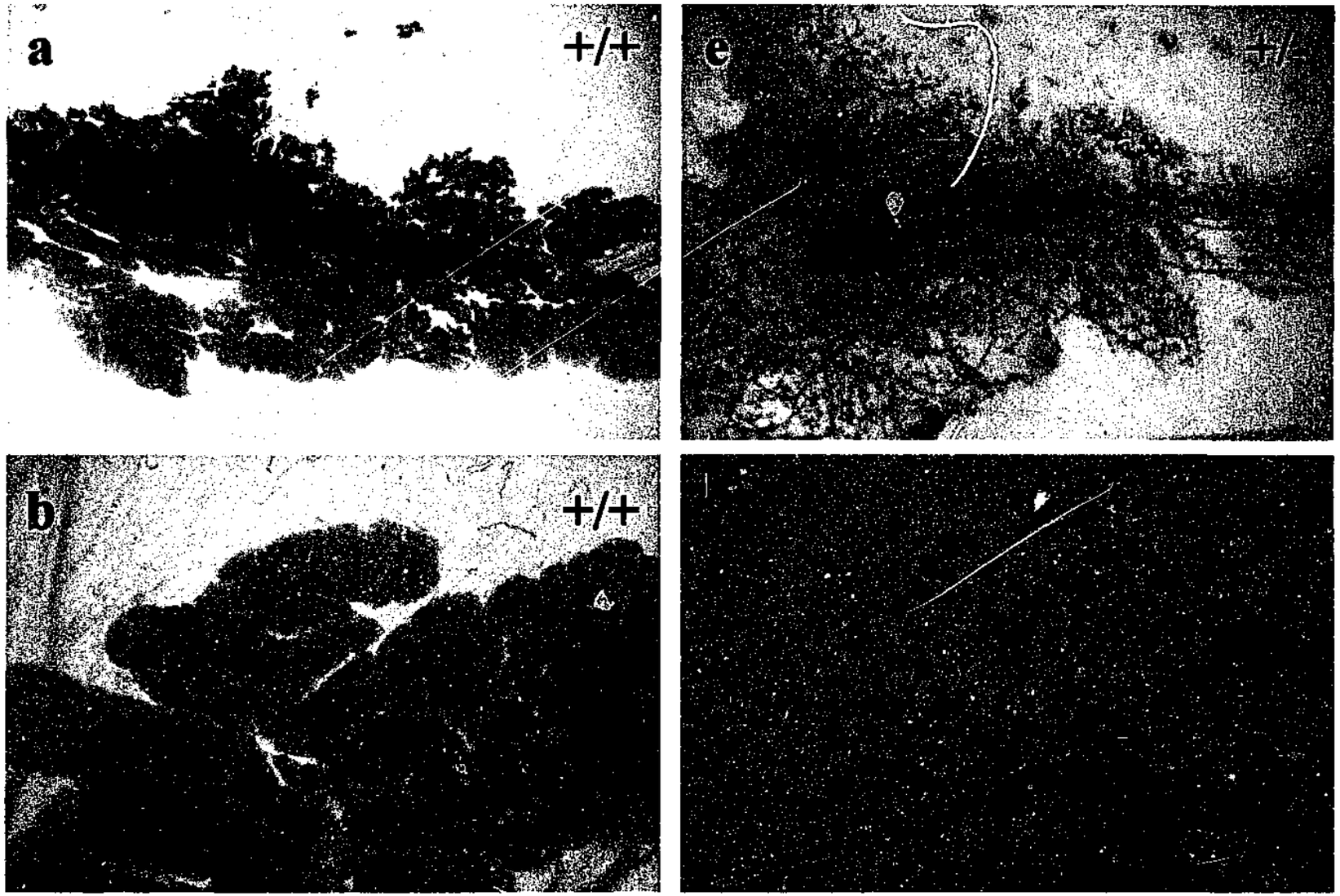


Figure 7.8 Mammary gland development is defective in lactating *Elf5*$^{+/-}$ females (example #2). Mammary alveolar proliferation and terminal differentiation was severely impaired in an one day postpartum *Elf5*$^{+/-}$ female (169 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (137 days old, a-d). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (e-f) females after their first pregnancy. Magnifications: a & e ×18.9, b & f ×120. The whole mounts of inguinal mammary glands were stained with carmine red. Part of gland was removed for histological analysis. (Continued on next page)

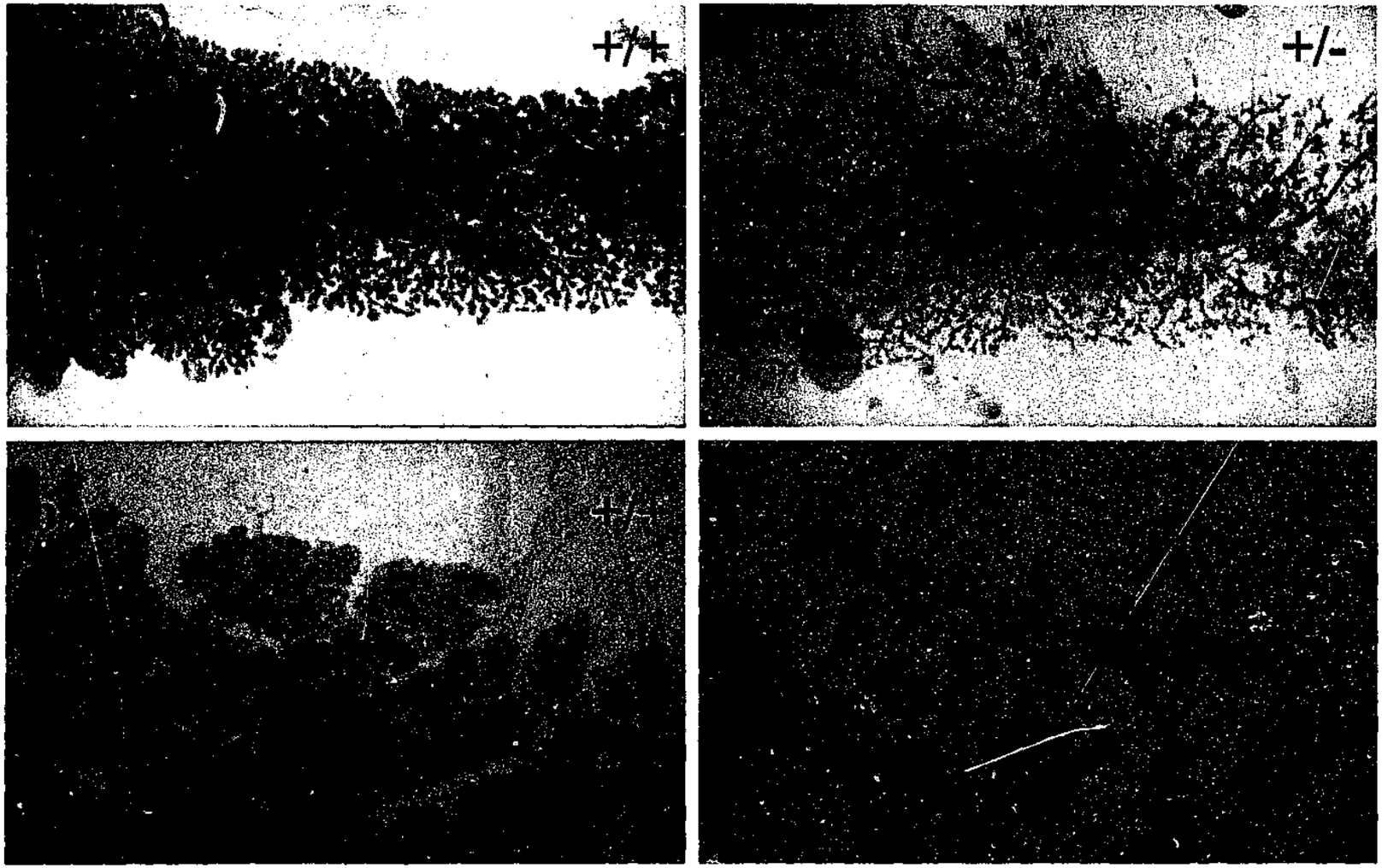


Figure 7.9 Mammary gland development is defective in lactating *Elf5*$^{+/-}$ females (example #3). Mammary alveolar proliferation and terminal differentiation was severely impaired in an one day postpartum *Elf5*$^{+/-}$ female (137 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (130 days old, a-d). Whole mount analyses of mammary tissue of *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (e-f) females after their first pregnancy. Magnifications: a & e ×18.9, b & f ×120. The whole mounts of inguinal mammary glands were stained with carmine red. Part of gland was removed for histological analysis. (Continued on next page)

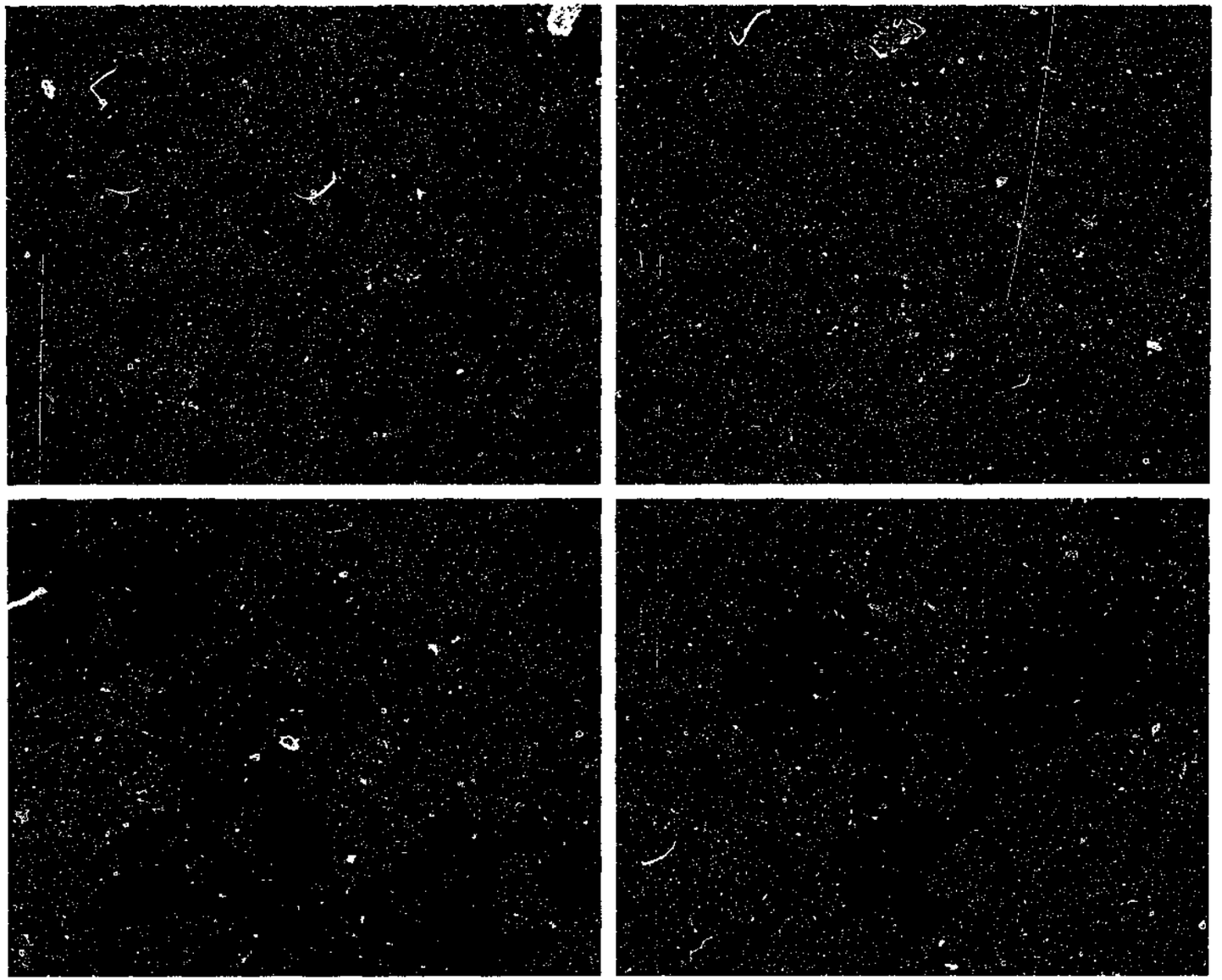


Figure 7.9 Mammary gland development is defective in lactating *Elf5*$^{+/-}$ females (example #3). Mammary alveolar proliferation and terminal differentiation was severely impaired in an one day postpartum *Elf5*$^{+/-}$ female (137 days old, e-h) in comparison to an *Elf5*$^{+/+}$ female (130 days old, a-d). Histological analyses of mammary tissue of *Elf5*$^{+/+}$ (c-d) and *Elf5*$^{+/-}$ (g-h) females after their first pregnancy. Black arrows point to the alveoli. The bar denotes 0.5 mm (c & g) or 0.05 mm (d & h). The paraffin histological sections (c-d & g-h) of inguinal mammary glands were stained with hematoxylin and eosin. (Continued from previous page)

The lactational defect observed persisted in the subsequent pregnancies in *Elf5*$^{+/-}$ females. Whole mount and histological examination of the 18.5 day pregnant *Elf5*$^{+/-}$ mammary gland during a second pregnancy revealed some alveolar development (Figure 7.10 i-l). However, this development was still under-rated in comparison to the wildtype glands (Figures 7.4-7.6 a-d). The *Elf5*$^{+/-}$ gland at one day postpartum after the second pregnancy did not undergo massive proliferation and differentiation. The alveolar structures were distended whilst the primary duct failed to dilate (compare Figure 7.10 a-d with e-h).

The *Elf5* targeted allele was backcrossed to a pure 129Svter genetic background, and six *Elf5*$^{+/-}$ females were mated with either *Elf5*$^{+/+}$ or *Elf5*$^{+/-}$ males. All six females failed to keep any of their pups after multiple pregnancies (up to four pregnancies).

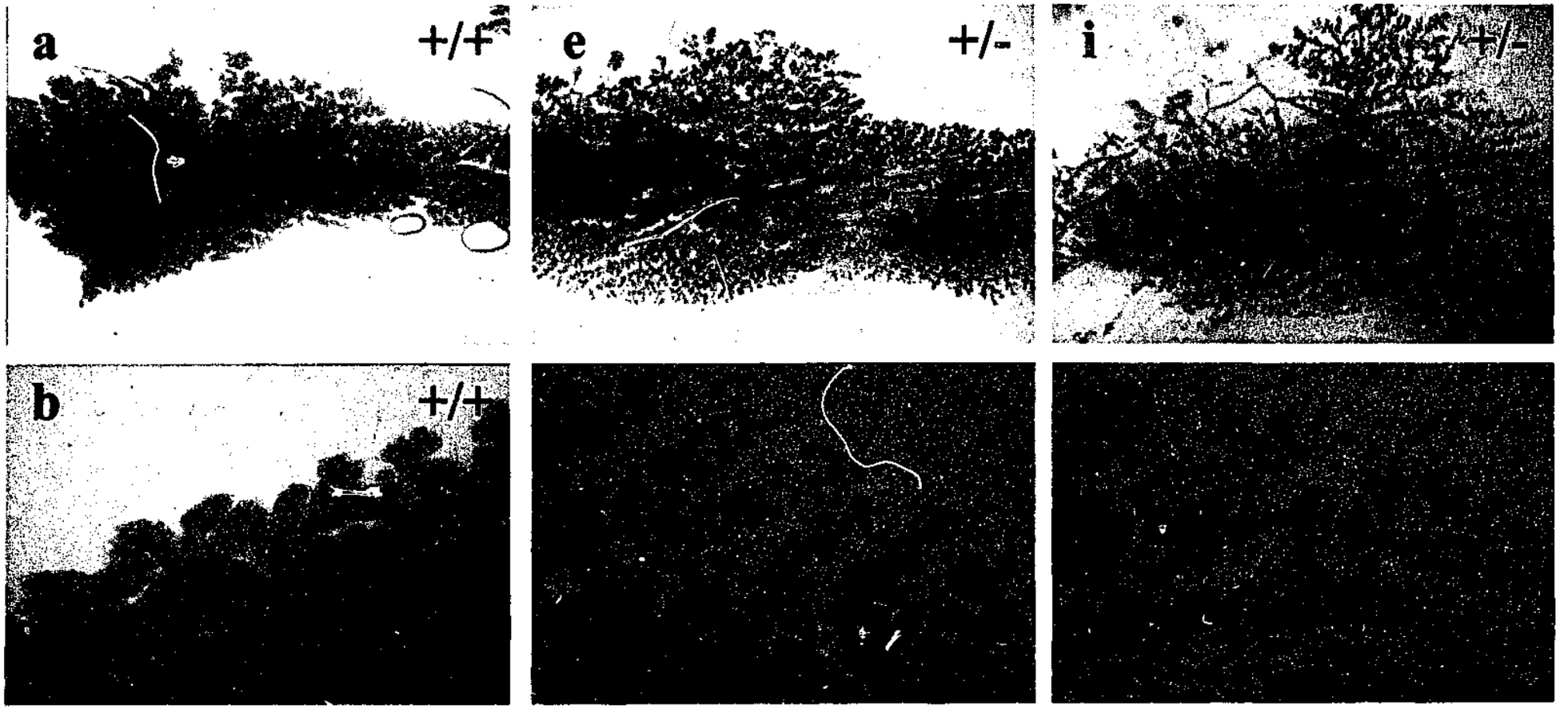
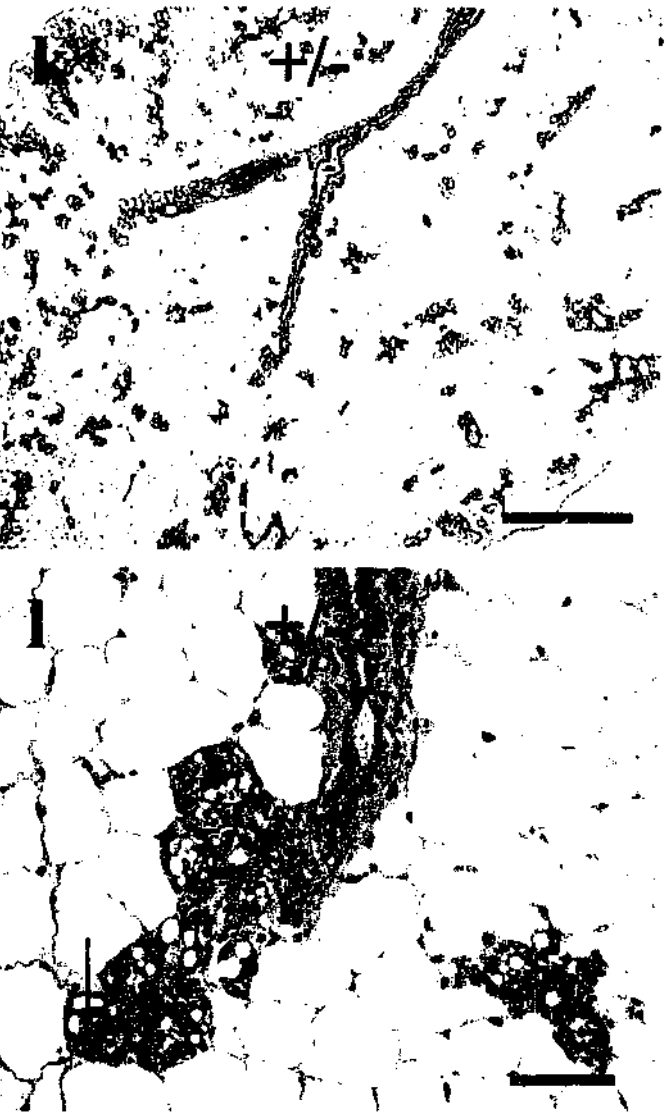


Figure 7.10 Mammary gland development is defective in pregnant and lactating *Elf5*$^{+/-}$ females (second pregnancy). Mammary alveolar proliferation and terminal differentiation was severely impaired in an one day postpartum *Elf5*$^{+/-}$ female (140 days old, e-h) and in an 18.5 day pregnant *Elf5*$^{+/-}$ female (150 days old, i-l) whilst normal mammary development was observed in an one day postpartum *Elf5*$^{+/+}$ female (140 days old, a-d). Whole mount analyses of mammary tissue of an *Elf5*$^{+/-}$ (i-j) female during the second pregnancy, and *Elf5*$^{+/+}$ (a-b) and *Elf5*$^{+/-}$ (e-f) females after their second pregnancies. Magnifications: a & e & i ×15.8, b & f & j ×100. The whole mounts of inguinal mammary glands were stained with carmine red. Part of gland was removed for histological analysis. (Continued on next page)



ancy). Mammary alveolar proliferation and terminal egnant *Elf5*$^{+/-}$ female (150 days old, i-l) whilst normal of mammary tissue of an *Elf5*$^{+/-}$ (i-j) female during the ducts, and black arrows point to the alveoli. The bar mammary glands were stained with hematoxylin and

7.3 Discussion

The mammary ductal tree can elongate and branch quickly through the entire mammary fat pad under the stimulation of ovarian hormones. Pregnanacy hormones will then promote alveolar proliferation, and the alveolar epithelial cells from the resulting lobuloalveolar structures will eventually differentiate into secretory epithelial cells at parturition.

Elf5$^{+/-}$ females displayed normal ductal elongation and outgrowth during puberty, and normal side branching and sprouting of alveolar buds during pregnancy, indicating that the loss of one *Elf5* allele did not disrupt mammary development at these stages (Figures 7.1-7.6). However, the presence of one mutated *Elf5* allele resulted in impaired proliferation and terminal differentiation of the alveolar buds at late pregnancy and lactation (Figures 7.4-7.10). We are still yet to determine the precise timing for the onset of this phenotype, but our results demonstrate that mammary epithelial cell proliferation and differentiation during pregnancy and the postpartum period depends on a threshold level of Elf5, and that this level is unobtainable with just one functional *Elf5* allele. Death of the pups born to *Elf5*$^{+/-}$ females was due to the consequences of inappropriate mammary alveolar growth and a failure in maternal milk production.

Various factors have been implicated at distinct stages in the proliferation, differentiation and involution of mammary epithelial structures. The study of various animal models in which the gene of interest has either been overexpressed or deleted has made it possible to assign each candidate gene to a specific stage of mammary gland development. It should be noted however, that with the exceptions of *PRLR*$^{+/-}$ and *Plg*$^{+/-}$ females (Ormandy *et al.*, 1997; Lund *et al.*, 2000), no other gene identified as being involved in mammary gland development and function, displayed a phenotype in the heterozygous state.

To gain some insight into where *Elf5* may fit into this scheme, it is instructive to compare the mammary gland phenotype observed in *Elf5*$^{+/-}$ females with the phenotypes observed in other models of mammary gland development and function. Similarities may reveal the existence of common pathways.

The *Elf5*$^{+/-}$ mammary development resembles some aspects of the phenotypes observed in mice with deficiencies in LAR, Galanin, Stat5a/Stat5b, PRLR and OPGL/RANK (see

Introduction - section 7.1.2). However, the mammary defect observed in *Elf5*$^{+/-}$ females most closely resembles that seen in *cyclin D1*$^{-/-}$ females (Fantl *et al.*, 1995; Sicinski *et al.*, 1995). Interestingly, it is suggested that growth factor-mediated transcriptional regulation of the *cyclin D1* gene may involve ETS proteins (Albanese *et al.*, 1995). MAP kinases are shown to be activated by EGF (Pelech and Sanghara, 1992) and several ETS proteins are shown to be activated by the MAP kinase signaling pathway (Janknecht *et al.*, 1993; Marais *et al.*, 1993; Wasylyk *et al.*, 1993; O'Neill *et al.*, 1994). Also, it is found that overexpression of p41^{MAPK} or c-Ets-2 can activate *cyclin D1* promoter activity, which is downregulated by dominant-negative p41^{MAPK}. In addition, ETS binding sites were identified within the proximal *cyclin D1* promoter (Albanese *et al.*, 1995). Therefore, we propose that Elf5 is a candidate for the transcriptional regulation of the *cyclin D1* promoter, especially since *Ets2*$^{-/-}$ and *Ets2*$^{+/-}$ females do not exhibit a similar mammary defect to that observed in *cyclin D1*$^{-/-}$ (Fantl *et al.*, 1995; Sicinski *et al.*, 1995) and *Elf5*$^{+/-}$ females.

Members of the ETS transcription factor family are also implicated in the transcriptional regulation of the PRL promoter (Bradford *et al.*, 1995; Howard and Maurer, 1995; Bradford *et al.*, 1996; Bradford *et al.*, 1997; Day *et al.*, 1998). Indeed, the cooperative interactions between Pit-1 and ETS family members appears particularly important in establishing pituitary lactotroph-specific *PRL* gene expression. Interestingly, the *Elf5*$^{+/-}$ mammary defect resembled that seen in *PRLR*$^{+/-}$ females. In *Elf5*$^{+/-}$ females, mammary development was insufficient to deliver a lactating gland after the first, and some of the subsequent pregnancies. Similarly, mammary glands of *PRLR*$^{+/-}$ females displayed severely compromised alveolar proliferation and failed to lactate after their first pregnancy (Ormandy *et al.*, 1997). Since disruptions in either *Elf5* or *PRLR* produce a similar phenotype, it is likely that *Elf5* and *PRLR* participate in the same pathway.

Although the mammary defect observed in *Elf5*$^{+/-}$ females was different to that seen in mice with the *Brca1* gene deleted in mammary epithelium (Xu *et al.*, 1999), an ETS protein has been implicated as a critical regulator of the *BRCA1* promoter (Atlas *et al.*, 2000). Notably, null mutations of *Brca1* in mice resulted in early embryonic lethality (E6.5-9) (Gowen *et al.*, 1996; Hakem *et al.*, 1997; Ludwig *et al.*, 1997), suggesting another possible link between *Elf5* and *Brca1*.

Lactational performance in *Elf5*$^{+/-}$ females on a 129SvJ-C57Bl/6J mixed genetic background correlated with the degree of mammary gland development. The variable phenotype observed is probably due to a genetic modifier present in one of the genetic backgrounds. *Elf5*$^{+/-}$ females that were backcrossed to a pure 129Svter genetic background were incapable of lactation after multiple pregnancies, suggesting that the mammary alveolar developmental defect was 100% penetrant in the 129Svter strain. The histological examination of these defective mammary glands is currently underway.

Mammary gland development during puberty, pregnancy, and lactation is coordinated precisely by the actions of the female endocrine hormones upon the mammary gland epithelium and stroma (Daniel and Silberstein, 1987). It was shown that stromal estrogen expression mediates mammary epithelial mitogenesis (Cunha *et al.*, 1997) whilst stromal PR is required for mammary ductal outgrowth (Humphreys *et al.*, 1997). Wiesen *et al.*, (1999) also showed that mammary ductal growth and branching is mediated by stromal EGFR. In addition, OT expression in the myoepithelial cells, surrounding the alveoli, is required for milk ejection (Young *et al.*, 1996; Nishimori *et al.*, 1996; Wagner *et al.*, 1997). Although Elf5 is apparently epithelial-specific, we are yet to determine its cellular distribution in the mammary gland. We cannot rule out the possibility that the mammary alveolar proliferation and terminal differentiation defect in *Elf5*$^{+/-}$ females may be induced by insufficient expression of Elf5 in the surrounding stromal environment. If the mammary defect in strain 129Svter is 100% penetrant, *Elf5*$^{+/-}$ mammary alveolar epithelial cell proliferation and differentiation could be studied by transplanting the *Elf5*$^{+/-}$ glands into the fat pads of *Elf5*$^{+/+}$ females, which are devoid of mammary epithelium.

Chapter 8

Final discussion

8.1 ELF5 is a novel ETS transcription factor

ETS factors have important developmental roles and many ETS factors have also been implicated in the control of cellular proliferation and tumorigenesis (see Chapter 1). Most ETS factors identified to date appear to function predominantly in hematopoietic lineages. However, a subclass of ETS factors that are epithelial-specific has emerged suggesting novel epithelial functions for the ETS transcription factor family.

ELF3, the first epithelial-specific ETS transcription factor described, was found to have important *in vivo* biological functions during embryonic development, and in the adult where it is required for the morphogenesis and differentiation of epithelial cells of the small intestine, uterus, prostate and seminal vesicles (Ms. A.Y.N. Ng, personal communication). Subsequently, the identification of a number of other epithelial-specific ETS transcription factors has been described. (Bochert *et al.*, 1998; Zhou *et al.*, 1998; Kleinbaum *et al.*, 1999; Oettgen *et al.*, 1999; Kas *et al.*, 2000; Oettgen *et al.*, 2000; Yamada *et al.*, 2000)

We have identified a novel ETS family member, ELF5 (Zhou *et al.*, 1998). Sequence analysis of the gene revealed the presence of two putative functional domains, the ETS and PNT domains, that are highly conserved among ETS family members (see Chapter 3). Variation in the sequences of these domains between ELF5 and other ETS factors has prompted us to suggest that these differences may confer functional specificity to the ELF5 protein. *In vitro* studies have demonstrated that Elf5 can function through its ETS domain as a transcriptional activator (see Chapter 3). In addition, we and others have also identified a

number of putative downstream targets for ELF5, such as *WAP* (Thomas *et al.*, 2000), *SPRR2A*, *PSP* and *PSA* (Oettgen *et al.*, 1999).

The expression pattern of *ELF5/Elf5* appears to be highly restricted to tissues rich in epithelial cells (see Chapter 4). Cell line studies also demonstrated expression of *ELF5/Elf5* in epithelial cells only (see Chapter 4; Oettgen *et al.*, 1999). Therefore, ELF5 appears to be another epithelial-specific ETS transcription factor.

8.2 *In vivo* biological functions of Elf5

Interestingly, hELF5/mElf5 share a high degree of sequence similarity in the ETS and PNT domains with other epithelial-specific ETS factors, such as hELF3/mElf3, hESE3/mEhf, and hPDEF(hPSE)/mPse (see Chapter 3). In addition, these *ETS* family members exhibit partially overlapping expression patterns. Moreover, *in vitro* studies have demonstrated that these ETS factors can also function through the same ETS binding sites in the promoters of putative downstream target genes (Oettgen *et al.*, 1999; Kas *et al.*, 2000; Thomas *et al.*, 2000). Therefore, questions are raised about the functional specificity of these ETS factors.

Knockout mouse models are commonly used to study the biological functions of a gene of interest. Studies on the *Elf3*$^{-/-}$ mouse have established specific functions for Elf3 during embryonic development and in the adult (Ms. A.Y.N. Ng, personal communication). Importantly, these results showed that other epithelial-specific ETS family members could not substitute for Elf3-specific functions in the affected tissues.

In order to elucidate the biological functions of Elf5, we have generated an *Elf5*$^{-/-}$ mouse model. Knockout experiments have demonstrated a crucial role for Elf5 during early embryogenesis (see Chapter 6). Null mutations of the *Elf5* gene resulted in embryonic lethality around E3.5-7.5. *In vitro* blastocyst outgrowth assays revealed defects in the cellular proliferation and differentiation of *Elf5*$^{-/-}$ blastocysts. These blastocysts failed to outgrow and attach to the culture dish, suggesting that *Elf5*$^{-/-}$ embryos fail to implant. Although the status of *Elf5* expression in the early mouse embryo and its corresponding extraembryonic tissues (before E9.5) is unknown, *Elf5* was found to be expressed in the placenta throughout E9.5-18 indicating a potential function for Elf5 in this tissue (see Chapter 4). The placenta is derived from the trophoblast (Kaufman, 1999; Kaufman and

Bard, 1999), and the trophectoderm of the blastocyst is important for the exchange of metabolites between the mother and the embryo (Collins and Fleming, 1995). The early lethality of *Elf5*$^{-/-}$ embryos may be due to the loss of Elf5 function in the embryonic trophoblast. However, we can not rule out possible defects in the inner cell mass since these cells appeared to degenerate *in vitro* earlier than their wildtype and heterozygous counterparts. Numerous attempts to generate *Elf5*$^{-/-}$ ES cells by increasing the concentration of G418 were unsuccessful, suggesting that these cells had a growth disadvantage. This supported the *in vitro* observation of a prematurely degenerated inner cell mass in *Elf5*$^{-/-}$ embryos (see Chapter 6). However, we are yet to determine the expression pattern of the *Elf5* gene at this early stage (E3.5-7.5) and thus pinpoint the exact loss of function defect.

A defect involving alveolar proliferation and differentiation was observed in the mammary glands of pregnant and lactating *Elf5*$^{+/-}$ females. This defect resulted in lethality of newborn pups (see Chapter 7), indicating that a functional lactating mammary gland requires a threshold level of Elf5 that is unobtainable with just one functional gene allele. This phenotype resembled closely that seen in *PRLR*$^{+/-}$ females (Ormandy *et al.*, 1997), however, *PRLR*$^{+/-}$ females appeared to recover from this mammary defect in their subsequent pregnancies. Since disruptions in either *Elf5* or *PRLR* produce a similar phenotype, it is likely that *Elf5* and *PRLR* participate in the same pathway. A number of studies have shown that ETS factors are potentially involved in the transcriptional regulation of the PRL promoter (Bradford *et al.*, 1995; Howard and Maurer, 1995; Bradford *et al.*, 1996; Bradford *et al.*, 1997; Day *et al.*, 1998). At first sight, the fact that the mammary developmental defect of *PRL*$^{-/-}$ females occurs during puberty (Horseman *et al.*, 1997), may seem to rule out any Elf5 involvement in this pathway, since the Elf5-induced defect occurs just prior to and during lactation. However, we can not rule out more severe and earlier mammary developmental defects in mice that completely lack *Elf5* expression. In addition, we cannot rule out a possible alternative pathway for this mammary developmental period which may compensate for the deficiency of either *Elf5* or *PRLR*. Albanese *et al.*, (1995) suggested that transcriptional regulation of the *cyclin D1* gene might involve ETS proteins. Interestingly, a similar mammary developmental defect was also observed in pregnant and lactating *cyclin D1*$^{-/-}$ females. Taken together, these results suggest a possible functional relationship between Elf5 and PRL/PRLR and/or cyclin D1. It will be necessary to examine the expression levels of cyclin D1 and critical factors, such as PRL, PRLR, Stat5a and Stat5b in

the PRL/PRLR signaling pathway in pregnant and lactating *Elf5*$^{+/-}$ mammary glands to investigate this possibility.

ETS binding sites have been identified in both PRL and PRLR promoters (Hu *et al.*, 1999; Jacob *et al.*, 1999). This may suggest that ELF5 is a potential regulator of the PRL/PRLR signaling pathway. The activities of STAT5 proteins, as downstream targets of PRL signaling, may also be affected as a result of ELF5 deficiency. Therefore, studies of the regulation of the PRL/PRLR promoters and the activities of STAT5 proteins in response to ELF5 may further strengthen the hypothesis that ELF5 is involved in PRL/PRLR signaling in mammary gland development. Several ETS proteins are shown to be activated by the MAP kinase signaling pathway (Janknecht *et al.*, 1993; Marais *et al.*, 1993; Wasylyk *et al.*, 1993; O'Neill *et al.*, 1994), and it was shown that the regulation of the *cyclin D1* promoter may involve ETS proteins in response to growth factor-mediated activation of the MAP kinase signaling pathway (Albanese *et al.*, 1995). Therefore, future *in vitro* growth factor-associated studies of the transcriptional activity of the *cyclin D1* promoter in the presence of ELF5 may shed additional light. In addition, expression levels of milk proteins should also be investigated in *Elf5*$^{+/-}$ mammary glands during pregnancy and lactation. Each of the milk proteins, such as WDNM1, β-casein and whey acidic protein (WAP), are expressed at a different stage of pregnancy-associated mammary gland development, and hence represent different mammary alveolar differentiation states. Thomas *et al.*, (2000) have shown that ELF5 can upregulate the *WAP* promoter activity via the ETS binding site. Therefore, the *in vivo* mammary gland developmental defect in the *Elf5*$^{+/-}$ pregnant female may coincide with WAP-associated alveolar differentiation.

The human *ELF5* gene is localized to chromosome 11p13-15 (see Chapter 4), a region that frequently undergoes loss of heterozygosity (LOH) in many types of cancer. Examples include ductal breast carcinoma (Lichy *et al.*, 1998), lung carcinoma (Iizuka *et al.*, 1995), rhabdoid tumor of the kidney (Hirose *et al.*, 1996), prostate carcinoma (Dahiya *et al.*, 1997; Kawana *et al.*, 1997), gastric carcinoma (Baffa *et al.*, 1996), ovarian carcinoma (Wilson *et al.*, 1996), and the WAGR syndrome (wilms tumor, aniridia, genito-urinary anomalies and mental retardation) (Gawin *et al.*, 1995). This region is believed to harbour several tumor suppressor genes (Iizuka *et al.*, 1995; Zenklusen *et al.*, 1995; Baffa *et al.*, 1996; Feinberg *et al.*, 1996; Ichikawa *et al.*, 1996; Coleman *et al.*, 1997; Gao *et al.*, 1997), based upon both LOH data and the ability to inhibit the tumorigenicity in chemically-induced murine

squamous cell carcinomas, upon introduction of human chromosome 11 (Zenklusen *et al.*, 1995). It is intriguing that *ELF5* expression appears to be lost in many cancer cell lines, a subset of which were found to have lost an allele or to have a rearrangement of the *ELF5* gene. Preliminary examination of *ELF5* expression indicates that *ELF5* is not detectable in a number of primary breast carcinomas, although it is strongly expressed in adjacent normal epithelium (data not shown, Dr. D. Venter, personal communication). Thus, it may be possible for an *ETS* gene, such as *ELF5*, to have tumor suppressor properties and to be lost in certain cancers. Detailed characterization of the *ELF5* gene, including its promoter, in the future may be useful for the identification of putative transcriptional repressors of the promoter. In addition, possible methylation/mutation sites throughout the whole gene or gene rearrangements that are responsible for silencing the ELF5-specific functions may be identified by screening the *ELF5* gene locus in a wide range of cancers. ETS factors have been proposed to have tumor suppressor properties and to be critical regulators of tumor suppressor genes (including *Brca1*) (Suzuki *et al.*, 1995; Zhang *et al.*, 1997a; Zhang *et al.*, 1997b; Choi *et al.*, 1998; Atlas et al., 2000; Xing *et al.*, 2000). Interestingly, pregnant and lactating *Brca1*$^{-/-}$ females exhibited a mammary developmental defect (Xu *et al.*, 1999). In addition, null mutations of *Brca1* in mice resulted in early embryonic lethality (E6.5-9) (Gowen *et al.*, 1996; Hakem *et al.*, 1997; Ludwig *et al.*, 1997), suggesting a possible functional relationship between Elf5 and Brca1. Therefore, it will be necessary to examine the expression status of *Brca1* in pregnant and lactating *Elf5*$^{+/-}$ mammary glands. A GABPα/β heterodimer is capable of regulating *BRCA1* promoter activity via the ETS binding site in the promoter (Atlas *et al.*, 2000). This transcriptional regulation property may be shared by ELF5 and can be studied in a similar system *in vitro*.

Tubulogenesis and branching morphogenesis are developmental processes common to the formation of many organs, particularly lung, trachea, salivary gland, mammary gland, pancreas, prostate, and kidney. *ELF5/Elf5* expression has been established in most of these organs. However, early embryonic lethality of the *Elf5*$^{-/-}$ animals has hindered our study of the biological consequences of null mutations in the *Elf5* gene in these organs. In addition, loss of Elf5 biological function in other tissues may also be masked by this embryonic lethality. *Ets2*$^{-/-}$ embryos died during early embryogenesis due to defects in trophoblast function, however, this lethality was rescued by introducing normal extraembryonic tissues to the mutant embryo (Yamamoto *et al.*, 1998). *Elf5*$^{-/-}$ embryos also displayed trophoblast defects that may be rescued by aggregation with tetraploid mouse embryos. Alternatively, a

conditional *Elf5*$^{-/-}$ mouse model can be generated that is specific to a developmental process. Any phenotype resulting from the specific mutation can then be compared with the existing knowledge in the field and addressed by additional *in vitro* and *in vivo* studies.

The conditional gene knockout technology has allowed the study of gene functions specific to a particular tissue or a developmental stage. A LoxP/CRE system could be used to generate a mammary-specific *Elf5* knockout mouse model via the expression of a *CRE* transgene controlled by a mammary–specific gene promoter. Ideally, we would cross our targeted *Elf5* mice with mice expressing *CRE* under a variety of mammary gland specific promoters to knockout *Elf5* at critical stages of development of the mammary gland. Similarly, in light of the mammary phenotype observed in *Elf5*$^{+/-}$ pregnant and lactating females, it appears that the *WAP* gene promoter, turned on at approximately 15 days of pregnancy, is a suitable candidate for turning off *Elf5* expression during this critical period of mammary gland development. However, a more detailed examination of the *Elf5*$^{+/-}$ mammary gland is required to determine the timing of onset of this phenotype. If mammary defects are observed during earlier stages of pregnancy, promoters of other milk proteins, such as WDNM1 and β–casein, may be used for driving *CRE* transgene expression in the conditional knockout. In addition, Elf5 may also be essential during the prepuberal and puberal stages of mammary gland development. Therefore, genes, such as *ER* and *PR*, that regulate these developmental stages may be used to control *CRE* transgene expression. The *mouse mammary tumor virus-long terminal repeat* (*MMTV-LTR*) was used successfully to drivie *CRE* transgene expression in the conditional mutation of *Brca1* in the mouse mammary gland (Xu *et al.*, 1999). This promoter also appears to be active in other epithelial tissues, including prostate, salivary gland, kidney, lung, pancreas, seminal vesicle, epididymis, and testis (Tsubura *et al.*, 1981; Imai *et al.*, 1983). Therefore, conditional mutation of *Elf5* using *MMTV-LTR* driven *CRE* may facilitate studies of the development of other *Elf5*-expressing tissues.

Deficiency in tumor suppressor genes, such as *p53* (Donehower *et al.*, 1992; Harvey *et al.*, 1993), *Pten* (Di Cristofano *et al.*, 1998) and *Brca1* (Xu *et al.*, 1999), has led to the development of tumors in mice. The potential tumor suppressing activity of Elf5 can be tested by aging the *Elf5*$^{+/-}$ and *Elf5*$^{-/-}$ mice (either rescued by aggregation with tetraploid mouse embryos or generated as tissues-specific conditional knockout) to determine their susceptibility to spontaneous tumor formation. If these mice do form tumors, they can be

tested further by challenging with carcinogenic agents for accelerated tumor formation. Alternatively, mutations of known tumor suppressor genes can be introduced into these mice and examined for accelerated tumor formation.

8.3 Conclusions

The objective of this study was to isolate and characterize a novel epithelial-specific ETS transcription factor. We have identified a novel ETS factor, ELF5, that is expressed in tissues rich in epithelial cells. Gene targeting experiments demonstrated an important role for Elf5 in the proliferation and differentiation of mouse mammary alveolar epithelial cells during pregnancy and lactation. Mammary gland development during these stages is extremely sensitive to the level of Elf5 expression. The loss of one functional allele leads to complete developmental arrest of the mammary gland. In addition, we have shown that Elf5 has crucial cellular proliferation and differentiation functions during early mouse embryogenesis. This function may be related to the implantation of the mouse embryo, where the loss of function leads to embryonic lethality. This study has also identified possible functional relationships of ELF5 with a hormone signaling pathway (PRL/PRLR), a cell cycle regulator (CYCLIN D1), and a breast cancer suppressor protein (BRCA1). The *Elf5*$^{+/-}$ mouse model generated in this study will make an invaluable contribution towards understanding the function of ETS transcription factors in mammalian development, and possibly in human disease.

Appendix A Abbreviations

α–GSU	Gonadotropin-releasing hormone α–subunit
AML1	Acute myeloid leukemia 1
APN	Aminopeptidase N
ATF2	Activating transcription factor 2
birR	Biston repressor
BRCA1	Breast cancer 1
CAP	Catabolite activator protein
CAT	Chloramphenicol acetyl-transferase
CBF	Core binding factor
CDKs	Cyclin-dependent kinases
CRE	Cre recombinase
CREB	Cyclic amp response element-binding protein
CRISP-1	Cysteine-rich secretory protein 1
CRISP-3	Cysteine-rich secretory protein 3
DNA	Deoxyribonucleic acid
E1AF	E1A factor, same as ETV4
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EHF	Ets homologous factor, same as ESE3
ELF1	E74-like factor 1
ELF3	E74-like factor 3, same as ESX/ESE1/JEN/ERT
ELF5	E74-Like Factor 5, same as ESE2
EMSA	Electrophoretic mobility shift assay
Endo A	Extra-endodermal cytoskeletal protein A
ER81	Ets translocation variant gene 1, same as ETV1
ERF	Ets2 repressor factor
ERG	Ets-related gene
ERG-B	Ets-related gene B, same as FLI1
ERK	Extracellular-regulated kinase
ERM	Ets-related molecule, ets translocation variant gene 5
Erp	Ets-related protein
ES	Embryonic stem
ESE1	Epithelium-specific Ets transcription factor 1, same as ELF3/ESX/JEN/ERT
ESE2	Epithelium-specific Ets transcription factor 2, same as ELF5
ESE3	Epithelium-specific Ets transcription factor 3, same as EHF3
ERT	Ets-related transcription factor, same as ELF3/ESE1/ESX/JEN
ESX	Epithelial-restricted with serine box, same as ELF3/ESE1/JEN/ERT
ETS	E26 transformation specific or E-twenty-six specific
ETV1	Ets translocation variant gene 1, same as ER81
ETV3	Ets translocation variant gene 3, same as PE1
ETV4	Ets translocation variant gene 4, same as E1AF
ETV6	Ets translocation variant gene 6, same as TEL
EWS	Ewing's sarcoma
FEV	Fifth Ewing variant
FIAU	1-[2-deoxy, 2-fluoro-β-D-arabinofuranosyl]-5 iodouracil
FLI1	Friend leukaemia integration 1, same as ERG-B
F-MuLV	Friend murine leukemia virus
FUS/TLS	Fusion/t(12;16) malignant liposarcoma
GABPα/β	GA-binding protein transcription factor alpha subunit/beta subunit
GANC	Gancyclovir
GAPDH	Glyceraldehyde phosphate dehydrogenase
GM-CSF	Granulocyte-macrophage colony-stimulating factor
HFN-3γ	Hepatocyte nuclear factor-3γ
HIV-1	Human immunodeficiency virus type 1
hph	Hygromycin B phosphotransferase
HSF	Heat shock factors
HSV-tk	Herpes simplex virus thymidine kinase

HTLV1	Human T cell lymphotropic virus type 1
ICM	Inner cell mass
IgH	Immunoglobulin heavy-chain
IL12	Interleukin 12
IL15	Interleukin 15
IL2	Interleukin 2
JAK2	Janus tyrosine kinase 2
K4	Keratin 4
LacZ	β-galactosidase gene
LEF1	Lymphoid enhancer-binding factor 1
LOH	Loss of heterozygosity
LTR	Long terminal repeat
MAP	Mitogen-activated protein
MEF	Myeloid elf-1-like factor
MMP13	Matrix metalloproteinase-13 or collagenase3
MMP3	Matrix metalloproteinase-3 or stromelysin1
MMP9	Matrix metalloproteinase-9 or gelatinase B
MMTV-LTR	Mouse mammary tumor virus-long terminal repeat
MN1	Meningioma nuclear factor 1
MNs	Motor neurons
MoMuLV	Moloney murine leukemia virus
MP6	Mouse proline rich protein
MSV	Moloney sarcoma virus
neo	Neomycin resistance gene
NERF	New Ets-related factor
NET	New ets factor
NFκB	Nuclear factor of kappa light polypeptide gene enhancer in B-cells
NK T	Natural killer T cell
NK	Natural killer cell
NLS	Nuclear localization signal
NMR	Nuclear magnetic resonance
OPGL	Osteoprotegerin-ligand
ORF	Open reading frame
OT	Oxytocin
PAX-5	Paired box gene 5
PDEF	Prostate-derived Ets factor, same as PSE
PDGFRβ	Platelet-derived growth factor receptor β
Pea3	Polyomavirus enhancer activator 3
Pip	PU.1-interacting protein
Plg	Plasminogen
PNT	Pointed
PRL	Prolactin
PRLR	Prolactin receptor
PSA	Prostate-specific antigen
PSE	Prostate-specific Ets, same as PDEF
PSMA	Prostate-specific membrane antigen
PSP	Parotid secretory protein
PSP94	Prostate secretory protein 94
PTEN	Phosphatase and tensin homolog
PyMT	Polyoma middle T oncogene
RACE	Rapid amplification of cDNA ends
RAG2	Recombination-activating gene 2
RANK	Receptor activator of NFκB
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase-polymerase chain reaction
SAM	Sterile alpha motif
SAP1	SRF accessory protein 1
SAP2	SRF accessory protein 2
SFFV	Spleen focus-forming virus
SpiB	Spleen focus forming virus (sffv) proviral integration B

SPRR2A	Small proline-rich protein 2A
SRE	Serum response element
SRF	Serum response factor
STAT5	Signal transducers and activators of transcription 5
TAD	Transactivation domain
TCR	T-cell receptor
TE	Trophectoderm
TEB	Terminal end buds
TFE3	Transcription factor for immunoglobulin heavy-chain enhancer 3
TGFβRII	Transforming growth factor, beta receptor II
TGM3	Transglutaminase 3
TK	Thymidine kinase
TN-C	Tenascin-C
UTR	Untranslated region
UV	Ultraviolet
WAGR syndrome	Wilms tumor, aniridia, genito-urinary anomalies and mental retardation
WAP	Whey acidic protein
wHTH	Winged helix-turn-helix or winged helix

Appendix B Common solutions

Acid alcohol
1% (v/v) HCl in 70% ethanol

Bouin's fixative
75% (v/v) picric acid, 10% (v/v) formaldehyde, 5% (v/v) glacial acetic acid

100× Denhardt's solution
2% (w/v) BSA, 2% (w/v) Ficoll, 2% (w/v) PVP

DNA loading dye
0.42% (w/v) bromophenol blue, 67% (w/v) sucrose, 50% (v/v) glycerol

Eosin
1.25% (w/v) eosin, 0.625% (w/v) potassium dichromate, 12.5% (v/v) picric acid, 12.5% (v/v) ethanol, 0.0625% (v/v) glacial acetic acid

Harris's heamatoxylin
0.5% (w/v) haematoxylin, 5% (v/v) ethahol, 10% (w/v) potassium aluminium sulphate, 0.25% (w/v) red mercuric oxide, 4% (v/v) glacial acetic acid

Luria Bertani broth (LB)
1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH7.0

1× PBS
0.8% (w/v) NaCl, 0.02% (w/v) KC1, 0.145% (w/v) Na_2HPO_4, 0.02% (w/v) KH_2PO_4, pH7.4

RNase A
10 mg/ml RNase A, 10 mM Tris-HCl pH7.5, 15 mM NaCl

RNA loading buffer
0.25% (v/v) bromophenol blue, 0.25% (v/v) xylene cyanol FF, 50% (v/v) glycerol, 1 mM EDTA pH 8.0

Scott's tap water
0.35% (w/v) sodium bicarbonate, 2% (w/v) magnesium sulphate, 1 thymol crystal

SM Buffer
0.58% (w/v) NaCl, 0.2% (w/v) $MgSO_4.7H_20$, 50 mM Tris-HCl, pH 7.5

SOC Medium
2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 20 mM glucose, 20 mM $MgCl_2$, 20 mM $MgSO_4$

20× SSC
3 M NaCl, 0.3 M sodium citrate , pH 7.0

1× TAE
40 mM Tris-Base, 1 mM EDTA, pH 8.0

Terrific Broth (TB broth)
1.2% (w/v) tryptone, 2.4% (w/v) yeast extract, 0.4% (v/v) glycerol, supplement with 17 mM KH_2PO_4 and 72 mM K_2HPO_4 prior to use

1× TE
10 mM Tris-HCl, 1 mM EDTA, pH 8.0

Appendix C List of equipment

Agarose gel electrophoresis units	Mini-sub horizontal slab, BIO-RAD MPH, International Biotechnologies, Inc.
Centrifuges	1-15, SIGMA Biofuge stratos, Heraeus J2-21 M/E, Beckman TL100 Ultracentrifuge, Beckman
Electroporation apparatus	Gene Pulsar, BIO-RAD
Homogenizer	Ika-ultra-turrax T25
Hybridization ovens	XTRON HI2002, Bartelt Instuments
Incubator	SANYO
Liquid scintillation analyzer	1900TR, Canberra Parkard
Microscope	DMIRB, Leica DMR, Leica MZ6, Leica
MilliQ biocel	Millipore
Parrafin embedding centre	EG1160, Leica
Paraffin microtome	RM2135, Leica
Paraffin processor	TP1020, Leica
PCR machines	GeneAmp PCR System 2400 and 9600, Perkin Elmer
Phosphorimager	FLA-2000, Fujifilm
Shaking incubator	Innova 4300, New Brunswick Scientific
UV illuminator	GelDoc 1000, BIO-RAD
UV/VIS spectrophotometer	Lambda Bio20, Perkin Elmer

Appendix D List of suppliers

American Tissue Cell Cultures (ATCC)
Rockville, MD, USA

Amersham Pharmacia Biotech.
Buckinghamshire, UK

AMRAD
Del Mar, CA, USA

Bartelt Instruments
Heidelberg West, VIC, Australia

BDH Biochemicals
Poole, UK

Beckman
Fullerton, CA, USA

Beckton Dickinson Labware
Franklin Lakes, NJ, USA

BIO-RAD
Hercules, CA, USA

Biolab Scientific
Gymea, NSW, Australia

Boehringer Mannheim
Mannheim, Germany

Bresatec
Austin, TX, USA

Canberra Packard
Mt. Waverley, VIC, Australia

Clontech
Palo Alto, CA, USA

Difco
Michigan, USA

FUJIFILM
Tokyo, Japan

GIBCO BRL
Paisley, UK

Hoefer
San Francisco, CA, USA

Kodak Scientific Imaging Film
Rochester, NY, USA

ICN Biochemicals
Aurora, Ohio, USA

Leica
Nussloch, Germany

Life Technologies Inc.
Paisley, UK

NEN-Life Science Products
Boston, MA, USA

Perkin Elmer
Norwalk, Connecticut, USA

Pierce
Rockford, IL, USA

Progen Industries
Rockford, IL, USA

Promega
Madison, WI, USA

Qiagen
Chatsworth, CA, USA

Sigma
St Louis, MO, USA

Stratagene
La Jolla, CA, USA

Syntex Australia Limited
North Sydney, NSW, Australia

Appendix E List of Publications and Conference Presentations

Part of the work in this thesis is related and published in the following articles:

Zhou J, Ng AYN, Tymms M, Jermiin LS, Seth AK, Thomas RS and Kola I (1998). A novel transcription factor, *ELF5*, belongs to the *ELF* subfamily of *ETS* genes and maps to human chromosome 11p13-15, a region subject to LOH and rearrangement in human carcinoma cell lines. *Oncogene* 17: 2719-32.

Thomas RS, Ng AN, **Zhou J**, Tymms MJ, Doppler W, and Kola I (2000). The Elf group of Ets-related transcription factors, ELF3 and ELF5. *Adv. Exp. Med. Biol.* 480: 123-8.

Tymms MJ, Ng AYN, Thomas RS, Schutte BC, **Zhou J**, Eyre HJ, Sutherland GR, Seth A, Rosenberg M, Papas T, Debouck C and Kola I (1997). A novel epithelial-expressed *ETS* gene, *ELF3*: human and murine cDNA sequences, murine genomic organisation, human mapping to 1q32.2 and expression in tissues and cancer. *Oncogene* 15: 2449-62.

Part of the work in this thesis was presented as posters at the following conferences:

Zhou J, Lapinskas E, Peeters J, Venter D, Hammacher A, Pritchard M, and Kola I (2000). Characterization of mouse *Elf5* gene by homologous recombination. 12th Lorne Cancer Conference, Lorne, Victoria, Australia, February 10-13th.

Zhou J, Lapinskas EJ, Peeters J, Venter D, Hammacher A, Pritchard M, and Kola I (1999). *ELF5*: A candidate tumor suppressor gene?. Third PeterMac Symposium, Melbourne, Victoria, Australia November 7-10th.

Zhou J, Venter D, Thomas RS, Pritchard M, and Kola I (1999). *Elf5* belongs to the ELF subfamily of *ETS* genes and maps to human chromosome 11p13-15, a region subject to LOH and rearrangement in human carcinoma cell lines. 11th Lorne Cancer Conference, Lorne, Victoria, Australia, February 11-14th.

Zhou J, Ng ANY, Thomas RS, Tymms MJ, and Kola I (1998). A novel *ETS* transcription factor *ELF4** is related to the epithelium-specific *ETS* factor *ELF3*. 10th Lorne Cancer Conference, Lorne, Victoria, Australia, February 12-15th.

* The novel epithelial-specific *ETS* transcription factor, *ELF5*, was originally named as *ELF4* by us. It was found out later that 'ELF4' has been reserved as a gene name for *MEF*, another novel *ETS* transcription factor.

Bibliography

Ablin RJ (1997). A retrospective and prospective overview of prostate-specific antigen. *J. Cancer Res. Clin. Oncol.* 123 (11-12): 583-94.

Adachi J, and Hasegawa M (1996). MOLPHY Version 2.3: Programs for molecular phylogenetics based on maximum likelihood. Computer Science Monographs. The Institute of Statistical Mathematics, Tokyo.

Akashi M, Shaw G, Hachiya M, Elstner E, Suzuki G, and Koeffler P (1994). Number and location of AUUUA motifs: role in regulating transiently expressed RNAs. *Blood* 83 (11): 3182-7.

Albanese C, Johnson J, Watanabe G, Eklund N, Vu D, Arnold A, and Pestell RG (1995). Transforming p21ras mutants and c-Ets-2 activate the cyclin D1 promoter through distinguishable regions. *J. Biol. Chem.* 270 (40): 23589-97.

Andreoli JM, Jang SI, Chung E, Coticchia CM, Steinert PM, and Markova NG (1997). The expression of a novel, epithelium-specific ets transcription factor is restricted to the most differentiated layers in the epidermis. *Nucleic Acids Res.* 25 (21): 4287-95.

Ann DK, Lin HH, and Kousvelari E (1997). Regulation of salivary-gland-specific gene expression. *Crit. Rev. Oral. Biol. Med.* 8 (3): 244-52.

Arber S, Ladle DR, Lin JH, Frank E, and Jessell TM (2000). ETS gene Er81 controls the formation of functional connections between group Ia sensory afferents and motor neurons. *Cell* 101 (5): 485-98.

Atlas E, Stramwasser M, Whiskin K, and Mueller CR (2000). GA-binding protein alpha/beta is a critical regulator of the BRCA1 promoter. *Oncogene* 19 (15): 1933-40.

Baffa R, Negrini M, Mandes B, Rugge M, Ranzani GN, Hirohashi S, and Croce CM (1996). Loss of heterozygosity for chromosome 11 in adenocarcinoma of the stomach. *Cancer Res.* 56 (2): 268-72.

Bailly RA, Bosselut R, Zucman J, Cormier F, Delattre O, Roussel M, Thomas G, and Ghysdael J (1994). 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.* 14 (5): 3230-41.

Bain G, Kitchens D, Yao M, Huettner JE, and Gottlieb DI (1995). Embryonic stem cells express neuronal properties *in vitro*. *Dev. Biol.* 168 (2): 342-57.

Barton K, Muthusamy N, Fischer C, Ting CN, Walunas TL, Lanier LL, and Leiden JM (1998). The Ets-1 transcription factor is required for the development of natural killer cells in mice. *Immunity* 9 (4): 555-63.

Bassuk AG, and Leiden JM (1995). A direct physical association between ETS and AP-1 transcription factors in normal human T cells. *Immunity* 3 (2): 223-37.

Bassuk AG, and Leiden JM (1997). The role of Ets transcription factors in the development and function of the mammalian immune system. *Adv. Immunol.* 64: 65-104.

Basuyaux JP, Ferreira E, Stehelin D, and Buttice G (1997). The Ets transcription factors interact with each other and with the c-Fos/c-Jun complex via distinct protein domains in a DNA-dependent and -independent manner. *J. Biol. Chem.* 272 (42): 26188-95.

Batchelor AH, Piper DE, de la Brousse FC, McKnight SL, and Wolberger C (1998). The structure of GABPalpha/beta: an ETS domain-ankyrin repeat heterodimer bound to DNA. *Science* 279 (5353): 1037-41.

Bear SE, Bellacosa A, Lazo PA, Jenkins NA, Copeland NG, Hanson C, Levan G, and Tsichlis PN (1989). Provirus insertion in Tpl-1, an Ets-1-related oncogene, is associated with tumor progression in Moloney murine leukemia virus-induced rat thymic lymphomas. *Proc. Natl. Acad. Sci. U S A* 86 (19): 7495-9.

Beitel GJ, Tuck S, Greenwald I, and Horvitz HR (1995). The Caenorhabditis elegans gene lin-1 encodes an ETS-domain protein and defines a branch of the vulval induction pathway. *Genes Dev.* 9 (24): 3149-62.

Bell SC (1985). Comparative aspects of decidualization in rodents and human: cell types, secreted products and associated function. In: Implantation of the Human Embryo. (eds. Edwards RG, Purdy JM, and Steptoe PC), pp. 71-122. Academic Press, London, UK.

Bellacosa A, Datta K, Bear SE, Patriotis C, Lazo PA, Copeland NG, Jenkins NA, and Tsichlis PN (1994). Effects of provirus integration in the Tpl-1/Ets-1 locus on Moloney murine leukemia virus-induced rat T-cell lymphomas: levels of expression, polyadenylation, transcriptional initiation, and differential splicing of the Ets-1 mRNA. *J. Virol.* 68 (4): 2320-30.

Ben-David Y, Giddens EB, and Bernstein A (1990). Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells induced by Friend murine leukemia virus. *Proc. Natl. Acad. Sci. U S A* 87 (4): 1332-6.

Ben-David Y, Giddens EB, Letwin K, and Bernstein A (1991). Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. *Genes Dev.* 5 (6): 908-18.

Benz CC, O'Hagan RC, Richter B, Scott GK, Chang CH, Xiong X, Chew K, Ljung BM, Edgerton S, Thor A, and Hassell JA (1997). HER2/Neu and the Ets transcription activator PEA3 are coordinately upregulated in human breast cancer. *Oncogene* 15 (13): 1513-25.

Bhat NK, Komschlies KL, Fujiwara S, Fisher RJ, Mathieson BJ, Gregorio TA, Young HA, Kasik JW, Ozato K, and Papas TS (1989). Expression of ets genes in mouse thymocyte subsets and T cells. *J. Immunol.* 142 (2): 672-8.

Bhat NK, Thompson CB, Lindsten T, June CH, Fujiwara S, Koizumi S, Fisher RJ, and Papas TS (1990). Reciprocal expression of human ETS1 and ETS2 genes during T-cell activation: regulatory role for the protooncogene ETS1. *Proc. Natl. Acad. Sci. U S A* 87 (10): 3723-7.

Bhattacharya G, Lee L, Reddy ES, and Rao VN (1993). Transcriptional activation domains of elk-1, delta elk-1 and SAP-1 proteins. *Oncogene.* 1993 Dec;8(12):3459-64.

Bidder M, Loewy AP, Latifi T, Newberry EP, Ferguson G, Willis DM, and Towler DA (2000). Ets domain transcription factor PE1 suppresses human interstitial collagenase promoter activity by antagonizing protein-DNA interactions at a critical AP1 element. *Biochemistry* 39 (30): 8917-28.

Biggs J, Murphy EV, and Israel MA (1992). A human Id-like helix-loop-helix protein expressed during early development. *Proc. Natl. Acad. Sci. U S A* 89 (4): 1512-6.

Birchmeier C, Meyer D, and Riethmacher D (1995). Factors controlling growth, motility, and morphogenesis of normal and malignant epithelial cel . *Int. Rev. Cytol.* 160: 221-66.

Bochert MA, Kleinbaum LA, Sun LY, and Burton FH (1998). Molecular cloning and expression of Ehf, a new member of the ets transcription factor/oncoprotein gene family. *Biochem. Biophys. Res. Commun.* 246 (1): 176-81.

Bole-Feysot C, Goffin V, Edery M, Binart N, and Kelly PA (1998). Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. *Endocr. Rev.* 19 (3): 225-68.

Bories JC, Willerford DM, Grevin D, Davidson L, Camus A, Martin P, Stehelin D, and Alt FW (1995). Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the Ets-1 proto-oncogene. *Nature* 377 (6550): 635-8.

Bosselut R, Levin J, Adjadj E, and Ghysdael J (1993). A single amino-acid substitution in the Ets domain alters core DNA binding specificity of Ets1 to that of the related transcription factors Elf1 and E74. *Nucleic Acids Res.* 21 (22): 5184-91.

Bradford AP, Conrad KE, Wasylyk C, Wasylyk B, and Gutierrez-Hartmann A (1995). Functional interaction of c-Ets-1 and GHF-1/Pit-1 mediates Ras activation of pituitary-specific gene expression: mapping of the essential c-Ets-1 domain. *Mol. Cell Biol.* 15 (5): 2849-57.

Bradford AP, Conrad KE, Tran PH, Ostrowski MC, and Gutierrez-Hartmann A (1996). GHF-1/Pit-1 functions as a cell-specific integrator of Ras signaling by targeting the Ras pathway to a composite Ets-1/GHF-1 response element. *J. Biol. Chem.* 271 (40): 24639-48.

Bradford AP, Wasylyk C, Wasylyk B, and Gutierrez-Hartmann A (1997). Interaction of Ets-1 and the POU-homeodomain protein GHF-1/Pit-1 reconstitutes pituitary-specific gene expression. *Mol. Cell Biol.* 17 (3): 1065-74.

Bradley A, Evans M, Kaufman MH, and Roberson E (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* 309 (5965): 255-6.

Breathnach R, Benoist C, O'Hare K, Gannon F, and Chambon P (1978). Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc. Natl. Acad. Sci. U S A* 75 (10): 4853-7.

Brembeck FH, Opitz OG, Libermann TA, and Rustgi AK (2000). Dual function of the epithelial

specific ets transcription factor, ELF3, in modulating differentiation. *Oncogene* 19 (15): 1941-9.

Brennan RG (1993). The winged-helix DNA-binding motif: another helix-turn-helix takeoff. *Cell* 74 (5): 773-6.

Bresciani F (1968). Topography of DNA synthesis in the mammary gland of the C3H mouse and its control by ovarian hormones. *Cell Tissue Kinet.* 1: 51-63.

Brown AG (1981). Organization in the spinal cord. pp. 154-214. Springer, New York, USA.

Brown TA, and McKnight SL (1992). Specificities of protein-protein and protein-DNA interaction of GABP alpha and two newly defined ets-related proteins. *Genes Dev.* 6 (12B): 2502-12.

Brunner D, Ducker K, Oellers N, Hafen E, Scholz H, and Klambt C (1994). The ETS domain protein pointed-P2 is a target of MAP kinase in the sevenless signal transduction pathway. *Nature* 370 (6488): 386-9.

Buijs A, Sherr S, van Baal S, van Bezouw S, van der Plas D, Geurts van Kessel A, Riegman P, Lekanne Deprez R, Zwarthoff E, Hagemeijer A, et al. (1995). Translocation (12;22) (p13;q11) in myeloproliferative disorders results in fusion of the ETS-like TEL gene on 12p13 to the MN1 gene on 22q11. *Oncogene* 10 (8): 1511-9.

Burtis KC, Thummel CS, Jones CW, Karim FD, and Hogness DS (1990). The Drosophila 74EF early puff contains E74, a complex ecdysone-inducible gene that encodes two ets-related proteins. *Cell* 61 (1): 85-99.

Buttice G, and Kurkinen M (1993). A polyomavirus enhancer A-binding protein-3 site and Ets-2 protein have a major role in the 12-O-tetradecanoylphorbol-13-acetate response of the human stromelysin gene. *J. Biol. Chem.* 268 (10): 7196-204.

Buttice G, Duterque-Coquillaud M, Basuyaux JP, Carrere S, Kurkinen M, and Stehelin D (1996). Erg, an Ets-family member, differentially regulates human collagenase1 (MMP1) and stromelysin1 (MMP3) gene expression by physically interacting with the Fos/Jun complex. *Oncogene* 13 (11): 2297-306.

Capecchi MR (1989). Altering the genome by homologous recombination. *Science* 244 (4910): 1288-92.

Carrere S, Verger A, Flourens A, Stehelin D, and Duterque-Coquillaud M (1998). Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complexes via two distinct domains. *Oncogene* 16 (25): 3261-8.

Carroll M, Tomasson MH, Barker GF, Golub TR, and Gilliland DG (1996). The TEL/platelet-derived growth factor beta receptor (PDGF beta R) fusion in chronic myelomonocytic leukemia is a transforming protein that self-associates and activates PDGF beta R kinase-dependent signaling pathways. *Proc. Natl. Acad. Sci. U S A* 93 (25): 14845-50.

Casey G (1997). The BRCA1 and BRCA2 breast cancer genes. *Curr. Opin. Oncol.* 9 (1): 88-93.

Chang CH, Scott GK, Kuo WL, Xiong X, Suzdaltseva Y, Park JW, Sayre P, Erny K, Collins C, Gray JW, and Benz CC (1997). ESX: a structurally unique Ets overexpressed early during human breast tumorigenesis. *Oncogene* 14 (13): 1617-22.

Chang J, Lee C, Hahm KB, Yi Y, Choi SG, and Kim SJ (2000). Over-expression of ERT(ESX/ESE-1/ELF3), an ets-related transcription factor, induces endogenous TGF-beta type II receptor expression and restores the TGF-beta signaling pathway in Hs578t human breast cancer cells. *Oncogene* 19 (1): 151-4.

Chen HM, and Boxer LM (1995). Pi 1 binding sites are negative regulators of bcl-2 expression in pre-B cells. *Mol. Cell Biol.* 15 (7): 3840-7.

Chen HM, Zhang P, Voso MT, Hohaus S, Gonzalez DA, Glass CK, Zhang DE, and Tenen DG (1995). Neutrophils and monocytes express high levels of PU.1 (Spi-1) but not Spi-B. *Blood* 85 (10): 2918-28.

Chen JH (1985). The proto-oncogene c-ets is preferentially expressed in lymphoid cells. *Mol. Cell Biol.* 5 (11): 2993-3000.

Chen JH, Vercamer C, Li Z, Paulin D, Vandenbunder B, and Stehelin D (1996). PEA3 transactivates vimentin promoter in mammary epithelial and tumor cells. *Oncogene* 13 (8): 1667-75.

Choi SG, Yi Y, Kim YS, Kato M, Chang J, Chung HW, Hahm KB, Yang HK, Rhee HH, Bang YJ, and Kim SJ (1998). A novel ets-related transcription factor, ERT/ESX/ESE-1, regulates expression of the transforming growth factor-beta type II receptor. *J. Biol. Chem.* 273 (1): 110-7.

Chomczynski P, and Sacchi N (1987). Single-step method of RNA isolation by acid guanidinium

thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162 (1): 156-9.

Chumakov AM, Chen DL, Chumakova EA, and Koeffler HP (1993). Localization of the c-ets-2 transactivation domain. *J. Virol.* 67 (4): 2421-5.

Clark KL, Halay ED, Lai E, and Burley SK (1993). Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* 364 (6436): 412-20.

Coleman WB, Esch GL, Borchert KM, McCullough KD, Reid LH, Weissman BE, Smith GJ, Grisham JW (1997). Localization of a putative liver tumor suppressor locus to a 950-kb region of human 11p11.2-p12 using rat liver tumor microcell hybrid cell lines. *Mol. Carcinog.* 19 (4): 267-72.

Collins JE, and Fleming TP (1995). Epithelial differentiation in the mouse preimplantation embryo: making adhesive cell contacts for the first time. *Trends Biochem. Sci.* 20 (8): 307-12.

Copp AJ (1978). Interaction between inner cell mass and trophectoderm of the mouse blastocyst. I. A study of cellular proliferation. *J. Embryol. Exp. Morphol.* 48: 109-125.

Copp AJ (1979). Interaction between inner cell mass and trophectoderm of the mouse blastocyst. II. The fate of the polar trophectoderm. *J. Embryol. Exp. Morphol.* 51: 109-120.

Crepieux P, Coll J, and Stehelin D (1994). The Ets family of proteins: weak modulators of gene expression in quest for transcriptional partners. *Crit. Rev. Oncog.* 5 (6): 615-38.

Cunha GR, Young P, Hom YK, Cooke PS, Taylor JA, and Lubahn DB (1997). Elucidation of a role for stromal steroid hormone receptors in mammary gland growth and development using tissue recombinants. *J. Mammary Gland Biol. Neoplasia* 2 (4): 393-402.

Dahiya R, McCarville J, Lee C, Hu W, Kaur G, Carroll P, and Deng G (1997). Deletion of chromosome 11p15, p12, q22, q23-24 loci in human prostate cancer. *Int. J. Cancer* 72 (2): 283-8.

Dalton S, and Treisman R (1992). Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. *Cell* 68 (3): 597-612.

Daly JM, Jannot CB, Beerli RR, Graus-Porta D, Maurer FG, and Hynes NE (1997). Neu differentiation factor induces ErbB2 down-regulation and apoptosis of ErbB2-overexpressing breast tumor cells. *Cancer Res.* 57 (17): 3804-11.

Daniel CW, and Silberstein GB (1987). Postnatal development on the rodent mammary gland. In The mammary gland: Development, regulation, and function. (eds. Neville MC, and Daniel CW), pp. 3-36. Plenum Press, New York, USA.

Darnell JE Jr, Kerr IM, and Stark GR (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 264 (5164): 1415-21.

Das R, and Vonderhaar BK (1995). Transduction of prolactin's (PRL) growth signal through both long and short forms of the PRL receptor. *Mol. Endocrinol.* 9 (12): 1750-9.

Day RN, Liu J, Sundmark V, Kawecki M, Berry D, and Elsholtz HP (1998). Selective inhibition of prolactin gene transcription by the ETS-2 repressor factor. *J. Biol. Chem.* 273 (48): 31909-15.

Degnan BM, Degnan SM, Naganuma T, and Morse DE (1993). The ets multigene family is conserved throughout the Metazoa. *Nucleic Acids Res.* 21 (15): 3479-84.

Delannoy-Courdent A, Fauquette W, Dong-Le Bourhis XF, Boilly B, Vandenbunder B, and Desbiens X (1996). Expression of c-ets-1 and uPA genes is associated with mammary epithelial cell tubulogenesis or neoplastic scattering. *Int. J. Dev. Biol.* 40 (6): 1097-108.

Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G, et al. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumors. *Nature* 359 (6391): 162-5.

Di Cristofano A, Pesce B, Cordon-Cardo C, and Pandolfi PP (1998). Pten is essential for embryonic development and tumor suppression. *Nat. Genet.* 19 (4): 348-55.

Dittmer J, and Nordheim A (1998). Ets transcription factors and human disease. *Biochim. Biophys. Acta.* 1377 (2): F1-11.

Doetschman TC, Eistetter H, Katz M, Schmidt W, and Kemler R (1985). The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morphol.* 87: 27-45.

Donaldson LW, Petersen JM, Graves BJ, and McIntosh LP (1994). Secondary structure of the

ETS domain places murine Ets-1 in the superfamily of winged helix-turn-helix DNA-binding proteins. *Biochemistry* 33 (46): 13509-16.

Donaldson LW, Petersen JM, Graves BJ, and McIntosh LP (1996). Solution structure of the ETS domain from murine Ets-1: a winged helix-turn-helix DNA bindingmotif. *EMBO J.* 15 (1): 125-34.

Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CA Jr, Butel JS, and Bradley A (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumors. *Nature* 356 (6366): 215-21.

Doolittle RF, Feng DF, Tsang S, Cho G, and Little E (1996). Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* 271 (5248): 470-7.

Dressler GR, and Douglass EC (1992). Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor. *Proc. Natl. Acad. Sci. U S A* 89 (4): 1179-83.

Dziadek M (1979). Cell differentiation in isolated inner cell masses of mouse blastocysts *in vitro*: onset of specific gene expression. *J. Embryol. Exp. Morphol.* 53: 367-79.

Easteal S, and Herbert G (1997). Molecular evidence from the nuclear genome for the time frame of human evolution. *J. Mol. Evol.*44 Suppl 1: S121-32

Evans MJ, and Kaufman MH (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292 (5819): 154-6.

Fantl V, Stamp G, Andrews A, Rosewell I, and Dickson C (1995). Mice lacking cyclin D1 are small and show defects in eye and mammary gland development. *Genes Dev.* 9 (19): 2364-72.

Fantl V, Edwards PA, Steel JH, Vonderhaar BK, and Dickson C (1999). Impaired mammary gland development in Cyl-1(-/-) mice during pregnancy and lactation is epithelial cell autonomous. *Dev. Biol.* 212 (1): 1-11.

Fata JE, Kong YY, Li J, Sasaki T, Irie-Sasaki J, Moorehead RA, Elliott R, Scully S, Voura EB, Lacey DL, Boyle WJ, Khokha R, and Penninger JM (2000). The osteoclast differentiation factor osteoprotegerin-ligand is essential for mammary gland development. *Cell* 103 (1): 41-50.

Feinberg AP (1996). Multiple genetic abnormalities of 11p15 in Wilms' tumor. *Med. Pediatr. Oncol.* 27 (5): 484-9.

Fenrick R, Amann JM, Lutterbach B, Wang L, Westendorf JJ, Downing JR, and Hiebert SW (1999). Both TEL and AML-1 contribute repression domains to the t(12;21) fusion protein. *Mol. Cell Biol.* 19 (10): 6566-74.

Fischer DF, Gibbs S, van De Putte P, and Backendorf C (1996). Interdependent transcription control elements regulate the expression of the SPRR2A gene during keratinocyte terminal differentiation. *Mol. Cell Biol.* 16 (10): 5365-74.

Fisher RJ, Koizumi S, Kondoh A, Mariano JM, Mavrothalassitis G, Bhat NK, and Papas TS (1991). Human ETS1 oncoprotein. Purification, isoforms, -SH modification, and DNA sequence-specific binding. *J. Biol. Chem.* 267 (25): 17957-65.

Fisher RJ, Fivash M, Casas-Finet J, Erickson JW, Kondoh A, Bladen SV, Fisher C, Watson DK, and Papas T (1994). Real-time DNA binding measurements of the ETS1 recombinant oncoproteins reveal significant kinetic differences between the p42 and p51 isoforms. *Protein Sci.* 3 (2): 257-66.

Fitzsimmons D, Hodsdon W, Wheat W, Maira SM, Wasylyk B, and Hagman J (1996). Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev.* 10 (17): 2198-211.

Fleischman LF, Holtzclaw L, Russell JT, Mavrothalassitis G, and Fisher RJ (1995). ets-1 in astrocytes: expression and transmitter-evoked phosphorylation. *Mol. Cell Biol.* 15 (2): 925-31.

Fletcher JC, and Thummel CS (1995). The Drosophila E74 gene is required for the proper stage- and tissue-specific transcription of ecdysone-regulated genes at the onset of metamorphosis. *Development* 121 (5): 1411-21.

Fletcher JC, Burtis KC, Hogness DS, and Thummel CS (1995). The Drosophila E74 gene is required for metamorphosis and plays a role in the polytene chromosome puffing response to ecdysone. *Development* 121 (5): 1455-65.

Flory E, Hoffmeyer A, Smola U, Rapp UR, and Bruder JT (1996). Raf-1 kinase targets GA-binding protein in transcriptional regulation of the human immunodeficiency virus type 1 promoter. *J. Virol.* 70 (4): 2260-8.

Fujimura Y, Yamamoto H, Hamazato F, and Nozaki M (1994). One of two Ets-binding sites in the cytokeratin EndoA enhancer is essential for enhancer activity and binds to Ets-2 related proteins. *Nucleic Acids Res.* 22 (4): 613-8.

Gajewski KM, and Schulz RA (1995). Requirement of the ETS domain transcription factor D-ELG for egg chamber patterning and development during Drosophila oogenesis. *Oncogene* 11 (6): 1033-40.

Galson DL, Hensold JO, Bishop TR, Schalling M, D'Andrea AD, Jones C, Auron PE, and Housman DE (1993). Mouse beta-globin DNA-binding protein B1 is identical to a proto-oncogene, the transcription factor Spi-1/PU.1, and is restricted in expression to hematopoietic cells and the testis. *Mol. Cell Biol.* 13 (5): 2929-41.

Gambarotta G, Boccaccio C, Giordano S, Ando M, Stella MC, and Comoglio PM (1996). Ets up-regulates MET transcription. *Oncogene* 13 (9): 1911-7.

Gao AC, Lou W, Dong JT, and Isaacs JT (1997). CD44 is a metastasis suppressor gene for prostatic cancer located on human chromosome 11p13. *Cancer Res.* 57 (5): 846-9.

Gardner RL, and Papaioannou VE (1975). Differentiation in the trophectoderm and inner cell mass. In: The early Development of Mammals. (eds. Balls M, and Wild AE), pp. 107-32. Cambridge University Press, Cambridge, UK.

Gawin B, Klamt B, Konig A, Thate C, Le Paslier D, Chumakov I, Bhogal R, Zehetner G, Bruns G, and Gessler M (1995). An integrated YAC clone contig for the WAGR region on human chromosome 11p13-p14.1. *Genomics* 30 (1): 37-45.

Gegonne A, Bosselut R, Bailly RA, and Ghysdael J (1993). Synergistic activation of the HTLV1 LTR Ets-responsive region by transcription factors Ets1 and Sp1. *EMBO J.* 12 (3): 1169-78.

Gehring WJ (1987). Homeo boxes in the study of development. *Science* 236 (4806): 1245-52.

Geng Y, Whoriskey W, Park MY, Bronson RT, Medema RH, Li T, Weinberg RA, and Sicinski P (1999). Rescue of cyclin D1 deficiency by knockin cyclin E. *Cell* 97 (6): 767-77.

Ghysdael J, Gegonne A, Pognonec P, Dernis D, Leprince D, and Stehelin D (1986). Identification and preferential expression in thymic and bursal lymphocytes of a c-ets oncogene-encoded Mr 54,000 cytoplasmic protein. *Proc. Natl. Acad. Sci. U S A* 83 (6): 1714-8.

Ghysdael J, and Boureux A (1997). The Ets family of transcriptional regulators. In: Oncogenes as Transcriptional Regulators. (eds. Yaniv M, and Ghysdael J), Vol. 1, pp.22. Birkhauser Verlag, Basel, Switzerland.

Giese K, Kingsley C, Kirshner JR, and Grosschedl R (1995). Assembly and function of a TCR alpha enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. *Genes Dev.* 9 (8): 995-1008.

Gille H, Kortenjann J, Thomae O, Moomaw C, Slaughter C, Cobb MH, and Shaw PE (1995). ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and transactivation. *EMBO J.* 14 (5): 951-62.

Gille H, Kortenjann J, Strahl T, and Shaw PE (1996). Phosphorylation-dependent formation of a quaternary complex at the c-fos SRE. *Mol. Cell Biol.* 16 (3): 1094-102.

Giovane A, Pintzas A, Maira SM, Sobieszczuk P, and Wasylyk B (1994). Net, a new ets transcription factor that is activated by Ras. *Genes Dev.*8 (13): 1502-13.

Golay J, Introna M, and Graf T (1988). A single point mutation in the v-ets oncogene affects both erythroid and myelomonocytic cell differentiation. *Cell* 55 (6): 1147-58.

Goldberg Y, Treier M, Ghysdael J, and Bohmann D (1994). Repression of AP-1-stimulated transcription by c-Ets-1. *J. Biol. Chem.* 269 (24): 16566-73.

Golub TR, Barker GF, Lovett M, and Gilliland DG (1994). Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell* 77 (2): 307-16.

Golub TR, Barker GF, Bohlander SK, Hiebert SW, Ward DC, Bray-Ward P, Morgan E, Raimondi SC, Rowley JD, and Gilliland DG (1995). Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. U S A* 92 (11): 4917-21.

Golub TR, Goga A, Barker GF, Afar DE, McLaughlin J, Bohlander SK, Rowley JD, Witte ON, and Gilliland DG (1996). Oligomerization of the ABL tyrosine kinase by the Ets protein TEL in human leukemia. *Mol. Cell Biol.* 16 (8): 4107-16.

Gouilleux F, Wakao H, Mundt M, and Groner B (1994). Prolactin induces phosphorylation of Tyr694 of Stat5 (MGF), a prerequisite for DNA binding and induction of transcription. *EMBO J.* 13 (18): 4361-9.

Gowen LC, Johnson BL, Latour AM, Sulik KK, and Koller BH (1996). Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nat. Genet.* 12 (2): 191-4.

Graves BJ and Petersen JM (1998). Specificity within the ets family of transcription factors. *Adv. Cancer Res.* 75: 1-55.

Graves BJ, Gillespie ME, and McIntosh LP (1996). DNA binding by the ETS domain. *Nature* 384 (6607): 322.

Gugneja S, Virbasius JV, and Scarpulla RC (1995). Four structurally distinct, non-DNA-binding subunits of human nuclear respiratory factor 2 share a conserved transcriptional activation domain. *Mol. Cell Biol.* 15 (1): 102-11.

Gugneja S, Virbasius CM, and Scarpulla RC (1996). Nuclear respiratory factors 1 and 2 utilize similar glutamine-containing clusters of hydrophobic residues to activate transcription. *Mol. Cell Biol.*16 (10): 5708-16.

Gullick WJ, Berger MS, Bennett PL, Rothbard JB, and Waterfield MD (1987). Expression of the c-erbB-2 protein in normal and transformed cells. *Int. J. Cancer* 40 (2): 246-54.

Gunther CV, and Graves BJ (1994). Identification of ETS domain proteins in murine T lymphocytes that interact with the Moloney murine leukemia virus enhancer. *Mol. Cell Biol.* 14 (11): 7569-80.

Gutman A, and Wasylyk B (1990). The collagenase gene promoter contains a TPA and oncogene-responsive unit encompassing the PEA3 and AP-1 binding sites. *EMBO J.* 9 (7): 2241-6.

Hagemeier C, Bannister AJ, Cook A, and Kouzarides T (1993). The activation domain of transcription factor PU.1 binds the retinoblastoma (RB) protein and the transcription factor TFIID *in vitro*: RB shows sequence similarity to TFIID and TFIIB. *Proc. Natl. Acad. Sci. U S A* 90 (4): 1580-4.

Hagman J, and Grosschedl R (1992). An inhibitory carboxyl-terminal domain in Ets-1 and Ets-2 mediates differential binding of ETS family factors to promoter sequences of the mb-1 gene. *Proc. Natl. Acad. Sci. U S A* 89 (19): 8889-93.

Hakem R, de la Pompa JL, Elia A, Potter J, and Mak TW (1997). Partial rescue of Brca1 (5-6) early embryonic lethality by p53 or p21 null mutation. *Nat. Genet.* 16 (3): 298-302.

Halle JP, Haus-Seuffert P, Woltering C, Stelzer G, and Meisterernst M (1997). A conserved tissue-specific structure at a human T-cell receptor beta-chain core promoter. *Mol. Cell Biol.* 17 (8): 4220-9.

Harrison CJ, Bohm AA, and Nelson HC (1994). Crystal structure of the DNA binding domain of the heat shock transcription factor. *Science* 263 (5144): 224-7.

Hart A, Melet F, Grossfeld P, Chien K, Jones C, Tunnacliffe A, Favier R, and Bernstein A (2000). Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia. *Immunity* 13 (2): 167-77.

Hart AH, Corrick CM, Tymms MJ, Hertzog PJ, and Kola I (1995). Human ERG is a proto-oncogene with mitogenic and transforming activity. *Oncogene* 10 (7): 1423-30.

Harvey M, McArthur MJ, Montgomery CA Jr, Butel JS, Bradley A, and Donehower LA (1993). Spontaneous and carcinogen-induced tumorigenesis in p53-deficient mice. *Nat. Genet.* 5 (3): 225-9.

Hashido K, Morita T, Matsushiro A, and Nozaki M (1991). Gene expression of cytokeratin endo A and endo B during embryogenesis and in adult tissues of mouse. *Exp. Cell Res.* 192 (1): 203-12.

Hasty P, and Bradley A (1993). Gene targeting vectors for mammalian cells. In: Gene targeting: A practical approach. (ed. Joyner AL), pp. 1-31. IRL Press, Oxford.

Hennighausen L, and Robinson GW (1998). Think globally, act locally: the making of a mouse mammary gland. *Genes Dev.* 12 (4): 449-55.

Hill AD, Doyle JM, McDermott EW, and O'Higgins NJ (1997). Hereditary breast cancer. *Br. J. Surg.* 84 (10): 1334-9.

Hill CS, Marais R, John S, Wynne J, Dalton S, and Treisman R (1993). Functional analysis of a growth factor-responsive transcription factor complex. *Cell* 73 (2): 395-406.

Hill CS, and Treisman R (1995). Differential activation of c-fos promoter elements by serum, lysophosphatidic acid, G proteins and polypeptide growth factors. *EMBO J.* 14 (20): 5037-47.

Hillis DM, Huelsenbeck JP, and Swofford DL (1994). Hobgoblin of phylogenetics? *Nature* 369 (6479): 363-4.

Hipskind RA, Buscher D, Nordheim A, and Baccarini M (1994). Ras/MAP kinase-dependent and -independent signaling pathways target distinct ternary complex factors. *Genes Dev.* 8 (15): 1803-16.

Hirose M, Yamada T, Toyosaka A, Hirose T, Kagami S, Abe T, and Kuroda Y (1996). Rhabdoid

tumor of the kidney: a report of two cases with respective tumor markers and a specific chromosomal abnormality, del(11p13). *Med. Pediatr. Oncol.* 27 (3): 174-8.

Hoch M, Gerwin N, Taubert H, and Jackle H (1992). Competition for overlapping sites in the regulatory region of the Drosophila gene Kruppel. *Science* 256 (5053): 94-7.

Hohl D, de Viragh PA, Amiguet-Barras F, Gibbs S, Backendorf C, and Huber M (1995). The small proline-rich proteins constitute a multigene family of differentially regulated cornified cell envelope precursor proteins. *J. Invest. Dermatol.* 104 (6): 902-9.

Horseman ND, and Yu-Lee LY (1994). Transcriptional regulation by the helix bundle peptide hormones: growth hormone, prolactin, and hematopoietic cytokines. *Endocr. Rev.* 15 (5): 627-49.

Horseman ND, Zhao W, Montecino-Rodriguez E, Tanaka M, Nakashima K, Engle SJ, Smith F, Markoff E, and Dorshkind K (1997). Defective mammopoiesis, but normal hematopoiesis, in mice with a targeted disruption of the prolactin gene. *EMBO J.* 16 (23): 6926-35.

Horwitz JP, Chua J, Curby RJ, Tomson AJ, DaRooge MA, Fisher BE, Mauricio J, and Klundt I (1964). Substrates for cytochemical demonstration of enzyme activity. I. Some substituted 3-indolyl-β-D-glycopyranosides. *J. Med. Chem.* 7: 574.

Howard PW, and Maurer RA (1994). Thyrotropin releasing hormone stimulates transient phosphorylation of the tissue-specific transcription factor, Pit-1. *J. Biol. Chem.* 269 (46): 28662-9.

Howard PW, and Maurer RA (1995). A composite Ets/Pit-1 binding site in the prolactin gene can mediate transcriptional responses to multiple signal transduction pathways. *J. Biol. Chem.* 270 (36): 20930-6.

Hromas R, Orazi A, Neiman RS, Maki R, Van Beveran C, Moore J, and Klemsz M (1993). Hematopoietic lineage- and stage-restricted expression of the ETS oncogene family member PU.1. *Blood* 82 (10): 2998-3004.

Hu ZZ, Zhuang L, Meng J, Leondires M, and Dufau ML (1999). The human prolactin receptor gene structure and alternative promoter utilization: the generic promoter hPIII and a novel human promoter hP(N). *J. Clin. Endocrinol. Metab.* 84 (3): 1153-6.

Humphreys RC, Krajewska M, Krnacik S, Jaeger R, Weiher H, Krajewski S, Reed JC, and Rosen JM (1996). Apoptosis in the terminal endbud of the murine mammary gland: a mechanism of ductal morphogenesis. *Development* 122 (12): 4013-22.

Humphreys RC, Lydon J, O'Malley BW, and Rosen JM (1997). Mammary gland development is mediated by both stromal and epithelial progesterone receptors. *Mol. Endocrinol.* 11 (6): 801-11.

Ichikawa H, Shimizu K, Hayashi Y, and Ohki M (1994). An RNA-binding protein gene, TLS/FUS, is fused to ERG in human myeloid leukemia with t(16;21) chromosomal translocation. *Cancer Res.* 54 (11): 2865-8.

Ichikawa T, Nihei N, Kuramochi H, Kawana Y, Killary AM, Rinker-Schaeffer CW, Barett JC, Isaacs JT, Kugoh H, Oshimura M, and Shimazaki J (1996). Metastasis suppressor genes for prostate cancer. *Prostate Suppl.* 6: 31-5.

Ida K, Kobayashi S, Taki T, Hanada R, Bessho F, Yamamori S, Sugimoto T, Ohki M, and Hayashi Y (1995). EWS-FLI-1 and EWS-ERG chimeric mRNAs in Ewing's sarcoma and primitive neuroectodermal tumor. *Int. J. Cancer* 63 (4): 500-4.

Iizuka M, Sugiyama Y, Shiraishi M, and Jones C (1995). Allelic losses in human chromosome 11 in lung cancers. *Genes Chromosomes Cancer* 13 (1): 40-6.

Imai S, Morimoto J, Tsubura Y, Iwai Y, Okumoto M, Takamori Y, Tsubura A, and Hilgers J (1983). Tissue and organ distribution of mammary tumor virus antigens in low and high mammary cancer strain mice. *Eur. J. Cancer Clin. Oncol.* 19 (7): 1011-9.

Jacob KK, Wininger E, DiMinni K, and Stanley FM (1999). The EGF response element in the prolactin promoter. *Mol. Cell. Endocrinol.* 152 (1-2): 137-45.

Janknecht R, Ernst WH, Pingoud V, and Nordheim A (1993). Activation of ternary complex factor Elk-1 by MAP kinases. *EMBO J.* 12 (13): 5097-104.

Janknecht R, and Nordheim A (1993). Gene regulation by Ets proteins. *Biochim. Biophys. Acta.*1155 (3): 346-56.

Janknecht R, Zinck R, Ernst WH, and Nordheim A (1994). Functional dissection of the transcription factor Elk-1. *Oncogene* 9 (4): 1273-8.

Janknecht R, Ernst WH, and Nordheim A (1995). SAP1a is a nuclear target of signaling cascades involving ERKs. *Oncogene* 10 (6): 1209-16.

Janknecht R (1996). Analysis of the ERK-stimulated ETS transcription factor ER81. *Mol. Cell Biol.* 16 (4): 1550-6.

Janknecht R, Monte D, Baert JL, and de Launoit Y (1996). The ETS-related transcription factor ERM is a nuclear target of signaling cascades involving MAPK and PKA. *Oncogene* 13 (8): 1745-54.

Jeon IS, Davis JN, Braun BS, Sublett JE, Roussel MF, Denny CT, and Shapiro DN (1995). A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. *Oncogene* 10 (6): 1229-34.

John S, Marais R, Child R, Light Y, and Leonard WJ (1996). Importance of low affinity Elf-1 sites in the regulation of lymphoid-specific inducible gene expression. *J. Exp. Med.* 183 (3): 743-50.

Jones DT, Taylor WR, and Thornton JM (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8 (3): 275-82.

Jonsen MD, Petersen JM, Xu QP, and Graves BJ (1996). Characterization of the cooperative function of inhibitory sequences in Ets-1. *Mol. Cell Biol.* 16 (5): 2065-73.

Jousset C, Carron C, Boureux A, Quang CT, Oury C, Dusanter-Fourt I, Charon M, Levin J, Bernard O, and Ghysdael J (1997). A domain of TEL conserved in a subset of ETS proteins defines a specific oligomerization interface essential to the mitogenic properties of the TEL-PDGFR beta oncoprotein. *EMBO J.* 16 (1): 69-82.

Judde JG, and Max EE (1992). Characterization of the human immunoglobulin kappa gene 3' enhancer: functional importance of three motifs that demonstrate B-cell-specific *in vivo* footprints. *Mol. Cell Biol.* 12 (11): 5206-16.

Kalderon D, Roberts BL, Richardson WD, and Smith AE (1984). A short amino acid sequence able to specify nuclear location. *Cell* 39 (3 Pt 2): 499-509.

Kaneko Y, Yoshida K, Handa M, Toyoda Y, Nishihira H, Tanaka Y, Sasaki Y, Ishida S, Higashino F and Fujinaga K (1996). Fusion of an ETS-family gene, EIAF, to EWS by t(17;22)(q12;q12) chromosome translocation in an undifferentiated sarcoma of infancy. *Genes Chromosomes Cancer* 15 (2): 115-21.

Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki RA, Gunther CV, Nye JA, et al. (1990). The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes Dev.* 4 (9): 1451-3.

Kas K, Finger E, Grall F, Gu X, Akbarali Y, Boltax J, Weiss A, Oettgen P, Kapeller R, and Libermann TA (2000). ESE-3, a novel member of an epithelium-specific ets transcription factor subfamily, demonstrates different target gene specificity from ESE-1. *J. Biol. Chem.* 275 (4): 2986-98.

Kaufman MH (1999). The Atlas of Mouse Development. Revised ed., Academic Press, CA, USA.

Kaufman MH, and Bard JBL (1999). The Anatomical Basis of Mouse Development. Academic Press, CA, USA.

Kawana Y, Komiya A, Ueda T, Nihei N, Kuramochi H, Suzuki H, Yatani R, Imai T, Dong JT, Imai T, Yoshie O, Barrett JC, Isaacs JT, Shimazaki J, Ito H, and Ichikawa (1997). Location of KAl1 on the short arm of human chromosome 11 and frequency of allelic loss in advanced human prostate cancer. *Prostate* 32 (3): 205-13.

Klambt C (1993). The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* 117 (1): 163-76.

Kleinbaum LA, Duggan C, Ferreira E, Coffey GP, Buttice G, and Burton FH (1999). Human chromosomal localization, tissue/tumor expression, and regulatory function of the ets family gene EHF. *Biochem. Biophys. Res. Commun.* 264 (1): 119-26.

Klemsz MJ, McKercher SR, Celada A, Van Beveren C, and Maki RA (1990). The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. *Cell* 61 (1): 113-24.

Klemsz MJ, Maki RA, Papayannopoulou T, Moore J, and Hromas R (1993). Characterization of the ets oncogene family member, fli-1. *J. Biol. Chem.* 268 (8): 5769-73.

Klemsz MJ, and Maki RA (1996). Activation of transcription by PU.1 requires both acidic and glutamine domains. *Mol. Cell Biol.* 16 (1): 390-7.

Kodandapani R, Pio F, Ni CZ, Piccialli G, Klemsz M, McKercher S, Maki RA, and Ely KR (1996). A new pattern for helix-turn-helix recognition revealed

by the PU.1 ETS-domain-DNA complex. *Nature* 380 (6573): 456-60.

Kokai Y, Cohen JA, Drebin JA, and Greene MI (1987). Stage- and tissue-specific expression of the neu oncogene in rat development. *Proc. Natl. Acad. Sci. U S A* 84 (23): 8498-501.

Kola I, Brookes S, Green AR, Garber R, Tymms M, Papas TS, and Seth A (1993). The Ets1 transcription factor is widely expressed during murine embryo development and is associated with mesodermal cells involved in morphogenetic processes such as organ formation. *Proc. Natl. Acad. Sci. U S A* 90 (16): 7588-92.

Kominato Y, Galson D, Waterman WR, Webb AC, and Auron PE (1995). Monocyte expression of the human prointerleukin 1 beta gene (IL1B) is dependent on promoter sequences which bind the hematopoietic transcription factor Spi-1/PU.1. *Mol. Cell Biol.* 15 (1): 59-68.

Kortenjann M, Thomae O, and Shaw PE (1994). Inhibition of v-raf-dependent c-fos expression and transformation by a kinase-defective mutant of the mitogen-activated protein kinase Erk2. *Mol. Cell Biol.* 14 (7): 4815-24.

Kovar H, Aryee DN, Jug G, Henockl C, Schemper M, Delattre O, Thomas G, and Gadner H (1996). EWS/FLI-1 antagonists induce growth inhibition of Ewing tumor cells *in vitro*. *Cell Growth Differ.* 7 (4): 429-37.

Kristensen T, Ogata RT, Chung LP, Reid KB, and Tack BF (1987). cDNA structure of murine C4b-binding protein, a regulatory component of the serum complement system. *Biochemistry* 26 (15): 4668-74.

Kwiatkowski BA, Bastian LS, Bauer TR Jr, Tsai S, Zielinska-Kwiatkowska AG, and Hickstein DD (1998). The ets family member Tel binds to the Fli-1 oncoprotein and inhibits its transcriptional activity. *J. Biol. Chem.* 273 (28): 17525-30.

Kyba M, and Brock HW (1998). The SAM domain of polyhomeotic, RAE28, and scm mediates specific interactions through conserved residues. *Dev. Genet.* 22 (1): 74-84.

Labosky PA, Winnier GE, Jetton TL, Hargett L, Ryan AK, Rosenfeld MG, Parlow AF, and Hogan BL (1997). The winged helix gene, Mf3, is required for normal development of the diencephalon and midbrain, postnatal growth and the milk-ejection reflex. *Development* 124 (7): 1263-74.

Laget MP, Defossez PA, Albagli O, Baert JL, Dewitte F, Stehelin D, and de Launoit Y (1996). Two functionally distinct domains responsible for transactivation by the Ets family member ERM. *Oncogene* 12 (6): 1325-36.

Lai ZC, and Rubin GM (1992). Negative control of photoreceptor development in Drosophila by the product of the yan gene, an ETS domain protein. *Cell* 70 (4): 609-20.

Laing MA, Coonrod S, Hinton BT, Downie JW, Tozer R, Rudnicki MA, and Hassell JA (2000). Male sexual dysfunction in mice bearing targeted mutant alleles of the *PEA3 ets* gene. *Mol. Cell. Biol.* 20 (24): 9337-45.

Laird PW, Zijderveld A, Linders K, Rudnicki MA, Jaenisch R, and Berns A (1991). Simplified mammalian DNA isolation procedure. *Nucleic Acids Res.* 19 (15): 4293.

Laudet V, Hanni C, Stehelin D, and Duterque-Coquillaud M (1999). Molecular phylogeny of the ETS gene family. *Oncogene* 18 (6) : 1351-9.

Laursen J, and Hjorth JP (1997). A cassette for high-level expression in the mouse salivary glands. *Gene* 198 (1-2): 367-72.

Laursen J, Krogh-Pedersen H, Dagnaes-Hansen F, and Hjorth JP (1998). The main regulatory region in the murine PSP gene is a parotid gland enhancer. *Transgenic Res.* 7 (6): 413-20.

Lebrun JJ, Ali S, Goffin V, Ullrich A, and Kelly PA (1995a). A single phosphotyrosine residue of the prolactin receptor is responsible for activation of gene transcription. *Proc. Natl. Acad. Sci. U S A* 92 (9): 4031-5.

Lebrun JJ, Ali S, Ullrich A, and Kelly PA (1995b). Proline-rich sequence-mediated Jak2 association to the prolactin receptor is required but not sufficient for signal transduction. *J. Biol. Chem.* 270 (18): 10664-70.

Lee JH, Jang SI, Yang JM, Markova NG, and Steinert PM (1996). The proximal promoter of the human transglutaminase 3 gene. Stratified squamous epithelial-specific expression in cultured cells is mediated by binding of Sp1 and ets transcription factors to a proximal promoter element. *J. Biol. Chem.* 271 (8): 4561-8.

Leprince D, Gegonne A, Coll J, de Taisne C, Schneeberger A, Lagrou C, and Stehelin D (1983). A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* 306 (5941): 395-7.

Levine M, and Hoey T (1988). Homeobox proteins as sequence-specific transcription factors. *Cell* 55 (4): 537-40.

Liang H, Olejniczak ET, Mao X, Nettesheim DG, Yu L, Thompson CB, and Fesik SW (1994). The secondary structure of the ets domain of human Fli-1 resembles that of the helix-turn-helix DNA-binding motif of the Escherichia coli catabolite gene activator protein. *Proc. Natl. Acad. Sci. U S A* 91 (24): 11655-9.

Lichy JH, Zavar M, Tsai MM, O'Leary TJ, and Taubenberger (1998). Loss of heterozygosity on chromosome 11p15 during histological progression in microdissected ductal carcinoma of the breast. *Am. J. Pathol.* 153 (1): 271-8.

Lim F, Kraut N, Framptom J, and Graf T (1992). DNA binding by c-Ets-1, but not v-Ets, is repressed by an intramolecular mechanism. *EMBO J.* 11 (2): 643-52.

Lin JH, Saito T, Anderson DJ, Lance-Jones C, Jessell TM, and Arber S (1998). Functionally related motor neuron pool and muscle sensory afferent subtypes defined by coordinate ETS gene expression. *Cell* 95 (3): 393-407.

Ling Y, Lakey JH, Roberts CE, and Sharrocks AD (1997). Molecular characterization of the B-box protein-protein interaction motif of the ETS-domain transcription factor Elk-1. *EMBO J.* 16 (9): 2431-40.

Liu X, Robinson GW, Gouilleux F, Groner B, and Hennighausen L (1995). Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in prolactin signal transduction in mouse mammary tissue. *Proc. Natl. Acad. Sci. U S A* 92 (19): 8831-5.

Liu X, Robinson GW, Wagner KU, Garrett L, Wynshaw-Boris A, and Hennighausen L (1997). Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes Dev.* 11 (2): 179-86.

Liu X, Gallego MI, Smith GH, Robinson GW, and Hennighausen L (1998). Functional release of Stat5a-null mammary tissue through the activation of compensating signals including Stat5b. *Cell Growth Differ.* 9 (9): 795-803.

Logan SK, Garabedian MJ, Campbell CE, and Werb Z (1996). Synergistic transcriptional activation of the tissue inhibitor of metalloproteinases-1 promoter via functional interaction of AP-1 and Ets-1 transcription factors. *J. Biol. Chem.* 271 (2): 774-82.

Lopez M, Oettgen P, Akbarali Y, Dendorfer U, and Libermann TA (1994). ERr', a new member of the ets transcription factor/oncoprotein family: cloning, characterization, and differential expression during B-lymphocyte development. *Mol. Cell Biol.* 14 (5): 3292-309.

Ludwig T, Chapman DL, Papaioannou VE, and Efstratiadis A (1997). Targeted mutations of breast cancer susceptibility gene homologs in mice: lethal phenotypes of Brca1, Brca2, Brca1/Brca2, Brca1/p53, and Brca2/p53 nullizygous embryos. *Genes Dev.* 11 (10): 1226-41.

Lund LR, Bjorn SF, Sternlicht MD, Nielsen BS, Solberg H, Usher PA, Osterby R, Christensen IJ, Stephens RW, Bugge TH, Dano K, and Werb Z (2000). Lactational competence and involution of the mouse mammary gland require plasminogen. *Development* 127 (20): 4481-92.

Lydon JP, DeMayo FJ, Funk CR, Mani SK, Hughes AR, Montgomery CA Jr, Shyamala G, Conneely OM, and O'Malley BW (1995). Mice lacking progesterone receptor exhibit pleiotropic reproductive abnormalities. *Genes Dev.* 9 (18): 2266-78.

Macleod K, Leprince D, and Stehelin D (1992). The ets gene family. *Trends Biochem. Sci.* 17 (7): 251-6.

Maira SM, Wurtz JM, and Wasylyk B (1996). Net (ERP/SAP2) one of the Ras-inducible TCFs, has a novel inhibitory domain with resemblance to the helix-loop-helix motif. *EMBO J.* 15 (21): 5849-65.

Mansour SL, Thomas KR, and Capecchi MR (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* 336 (6197): 348-52.

Mansour SL, Thomas KR, Deng CX, and Capecchi MR (1990). Introduction of a lacZ reporter gene into the mouse int-2 locus by homologous recombination. *Proc. Natl. Acad. Sci. U S A* 87 (19): 7688-92.

Mao X, Miesfeldt S, Yang H, Leiden JM, and Thompson CB (1994). The FLI-1 and chimeric EWS-FLI-1 oncoproteins display similar DNA binding specificities. *J. Biol. Chem.* 269 (27): 18216-22.

Marais R, Wynne J, and Treisman R (1993). The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell* 73 (2): 381-93.

Maroulakou IG, Papas TS, and Green JE (1994). Differential expression of ets-1 and ets-2 proto-

oncogenes during murine embryogenesis. *Oncogene* 9 (6): 1551-65.

Marquis ST, Rajan JV, Wynshaw-Boris A, Xu J, Yin GY, Abel KJ, Weber BL, and Chodosh LA (1995). The developmental pattern of Brca1 expression implies a role in differentiation of the breast and other tissues. *Nat. Genet.* 11 (1): 17-26.

Martin GR (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U S A* 78 (12): 7634-8.

May WA, Gishizky ML, Lessnick SL, Lunsford LB, Lewis BC, Delattre O, Zucman J, Thomas G, and Denny CT (1993). Ewing sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation. *Proc. Natl. Acad. Sci. U S A* 90 (12): 5752-6.

Mayall TP, Sheridan PL, Montminy MR, and Jones KA (1997). Distinct roles for P-CREB and LEF-1 in TCR alpha enhancer assembly and activation on chromatin templates *in vitro*. *Genes Dev.* 11 (7): 887-99.

McKercher SR, Torbett BE, Anderson KL, Henkel GW, Vestal DJ, Baribault H, Klemsz M, Feeney AJ, Wu GE, Paige CJ, and Maki RA (1996). Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J.* 15 (20): 5647-58.

McLean TW, Ringold S, Neuberg D, Stegmaier K, Tantravahi R, Ritz J, Koeffler HP, Takeuchi S, Janssen JW, Seriu T, Bartram CR, Sallan SE, Gilliland DG, and Golub TR (1996). TEL/AML-1 dimerizes and is associated with a favorable outcome in childhood acute lymphoblastic leukemia. *Blood* 88 (11): 4252-8.

Melet F, Motro B, Rossi DJ, Zhang L, and Bernstein A (1996). Generation of a novel Fli-1 protein by gene targeting leads to a defect in thymus development and a delay in Friend virus-induced erythroleukemia. *Mol. Cell Biol.* 16 (6): 2708-18.

Miller-Hance WC, LaCorbiere M, Fuller SJ, Evans SM, Lyons G, Schmidt C, Robbins J, and Chien KR (1993). *In vitro* chamber specification during embryonic stem cell cardiogenesis. Expression of the ventricular myosin light chain-2 gene is independent of heart tube formation. *J. Biol. Chem.* 268 (33): 25244-52.

Monte D, Baert JL, Laget MP, Defossez PA, Coutte L, Pelczar H, A geli I, Dewitte F, Stehelin D, and De Launoit Y (1995). Transcription factors of the PEA3 group in mammary cancer. *Ann. Endocrinol. (Paris)* 56 (5): 547-51.

Moreau-Gachelin F, Tavitian A, and Tambourin P (1988). Spi-1 is a putative oncogene in virally induced murine erythroleukaemias. *Nature.* 331 (6153): 277-80.

Moreau-Gachelin F, Wendling F, Molina T, Denis N, Titeux M, Grimber G, Briand P, Vainchenker W, and Tavitian A (1996). Spi-1/PU.1 transgenic mice develop multistep erythroleukemias. *Mol. Cell Biol.* 16 (5): 2453-63.

Mori S, Akiyama T, Yamada Y, Morishita Y, Sugawara I, Toyoshima K, and Yamamoto T (1989). C-erbB-2 gene product, a membrane protein commonly expressed on human fetal epithelial cells. *Lab. Invest.* 61 (1): 93-7.

Moritz C, and Hillis DM (1996). In Molecular systematics, 2nd ed. (eds. Hillis DM, Moritz C, and Mable BK). Sinauer, Sunderland, USA.

Moscovici MG, Jurdic P, Samarut J, Gazzolo L, Mura CV, and Moscovici C (1983). Characterization of the hemopoietic target cells for the avian leukemia virus E26. *Virology* 129 (1): 65-78.

Mount SM (1982). A catalogue of splice junction sequences. *Nucleic Acids Res.* 10 (2): 459-72.

Muthusamy N, Barton K, and Leiden JM (1995). Defective activation and survival of T cells lacking the Ets-1 transcription factor. *Nature* 377 (6550): 639-42.

Nakae K, Nakajima K, Inazawa J, Kitaoka T, and Hirano T (1995). ERM, a PEA3 subfamily of Ets transcription factors, can cooperate with c-Jun. *J. Biol. Chem.* 270 (40): 23795-800.

Nandi S (1958). Endocrine control of mammary gland development and function in the C3H/He Crgl mouse. *J. Natl. Cancer Inst.* 21: 1039-63.

Nelsen B, Tian G, Erman B, Gregoire J, Maki R, Graves B, and Sen R (1993). Regulation of lymphoid-specific immunoglobulin mu heavy chain gene enhancer by ETS-domain proteins. *Science* 261 (5117): 82-6.

Neve R, Chang CH, Scott GK, Wong A, Friis RR, Hynes NE, and Benz CC (1998). The epithelium-specific ets transcription factor ESX is associated with mammary gland development and involution. *FASEB J.* 12 (14): 1541-50.

Neville MC, and Daniel CW (1987). The mammary gland. Development, regulation, and function. In Plenum Press, New York, USA.

Neznanov N, Man AK, Yamamoto H, Hauser CA, Cardiff RD, and Oshima RG (1999). A single targeted Ets2 allele restricts development of mammary tumors in transgenic mice. *Cancer Res.* 59 (17): 4242-6.

Nishimori K, Young LJ, Guo Q, Wang Z, Insel TR, and Matzuk MM (1996). Oxytocin is required for nursing but is not essential for parturition or reproductive behavior. *Proc. Natl. Acad. Sci. U S A* 93 (21): 11699-704.

Nozaki M, Onishi Y, Kanno N, Ono Y, and Fujimura Y (1996). Molecular cloning of Elk-3, a new member of the Ets family expressed during mouse embryogenesis and analysis of its transcriptional repression activity. *DNA Cell Biol.* 15 (10): 855-62.

Nunn MF, Seeburg PH, Moscovici C, and Duesberg PH (1983). Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature* 306 (5941): 391-5.

Nunn MF, and Hunter T (1989). The ets sequence is required for induction of erythroblastosis in chickens by avian retrovirus E26. *J. Virol.* 63 (1): 398-402.

Nye JA, Petersen JM, Gunther CV, Jonsen MD, and Graves BJ (1992). Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev.* 6 (6): 975-90.

O'Neill EM, Rebay I, Tjian R. and Rubin GM (1994). The activities of two Ets-related transcription factors required for Drosophila eye development are modulated by the Ras/MAPK pathway. *Cell* 78 (1): 137-47.

Oettgen P, Akbarali Y, Boltax J, Best J, Kunsch C, and Libermann TA (1996). Characterization of NERF, a novel transcription factor related to the Ets factor ELF-1. *Mol. Cell Biol.* 16 (9): 5091-106.

Oettgen P, Alani RM, Barcinski MA, Brown L, Akbarali Y, Boltax J, Kunsch C, Munger K, and Libermann TA (1997). Isolation and characterization of a novel epithelium-specific transcription factor, ESE-1, a member of the ets family. *Mol. Cell Biol.* 17 (8): 4419-33.

Oettgen P, Kas K, Dube A, Gu X, Grall F, Thamrongsak U, Akbarali Y, Finger E, Boltax J, Endress G, Munger K, Kunsch C, and Libermann TA (1999). Characterization of ESE-2, a novel ESE-1-related Ets transcription factor that is restricted to glandular epithelium and differentiated keratinocytes. *J. Biol. Chem.* 274 (41): 29439-52.

Oettgen P, Finger E, Sun Z, Akbarali Y, Thamrongsak U, Boltax J, Grall F, Dube A, Weiss A, Brown L, Quinn G, Kas K, Endress G, Kunsch C, and Libermann TA (2000). PDEF, a novel prostate epithelium-specific ets transcription factor, interacts with the androgen receptor and activates prostate-specific antigen gene expression. *J. Biol. Chem.* 275 (2): 1216-25.

Ogata K, Hojo H, Aimoto S, Nakai T, Nakamura H, Sarai A, Ishii S, and Nishimura Y (1992). Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl. Acad. Sci. U S A* 89 (14): 6428-32.

Olsen J, Kokholm K, Troelsen JT, and Laustsen L (1997). An enhancer with cell-type dependent activity is located between the myeloid and epithelial aminopeptidase N (CD 13) promoters. *Biochem. J.* 322 (Pt 3): 899-908.

Olson MC, Scott EW, Hack AA, Su GH, Tenen DG, Singh H, and Simon MC (1995). PU. 1 is not essential for early myeloid gene expression but is required for terminal myeloid differentiation. *Immunity* 3 (6): 703-14.

Ormandy CJ, Camus A, Barra J, Damotte D, Lucas B, Buteau H, Edery M, Brousse N, Babinet C, Binart N, and Kelly PA (1997). Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. *Genes Dev.* 11 (2): 167-78.

Ouchida M, Ohno T, Fujimura Y, Rao VN, and Reddy ES (1995). Loss of tumorigenicity of Ewing's sarcoma cells expressing antisense RNA to EWS-fusion transcripts. *Oncogene* 11 (6): 1049-54.

Panagopoulos I, Aman P, Fioretos T, Hoglund M, Johansson B, Mandahl N, Heim S, Behrendtz M, and Mitelman F (1994). Fusion of the FUS gene with ERG in acute myeloid leukemia with t(16;21)(p11;q22). *Genes Chromosomes Cancer* 11 (4): 256-62.

Papadopoulos P, Ridge SA, Boucher CA, Stocking C, and Wiedemann LM (1995). The novel activation of ABL by fusion to an ets-related gene, TEL. *Cancer Res.* 55 (1): 34-8.

Papaioannou V, and Johnson R (1993). Production of chimeras and genetically defined offspring from targeted ES cells. In: Gene Targeting - A Practical

Approach. (ed. Joyner AL). Oxford University Press, New York, USA.

Papas TS, Fisher RJ, Bhat N, Fujiwara S, Watson DK, Lautenberger J, Seth A, Chen ZQ, Burdett L, Pribyl L, et al. (1989). The ets family of genes: molecular biology and functional implications. *Curr. Top. Microbiol. Immunol.* 149: 143-7.

Pascoe WS, Kemler R, and Wood SA (1992). Genes and functions: trapping and targeting in embryonic stem cells. *Biochim. Biophys. Acta.* 1114 (2-3): 209-21.

Paul R, Schuetze S, Kozak SL, and Kabat D (1989). A common site for immortalizing proviral integrations in Friend erythroleukemia: molecular cloning and characterization. *J. Virol.* 63 (11): 4958-61.

Peeters P, Raynaud SD, Cools J, Wlodarska I, Grosgeorge J, Philip P, Monpoux F, Van Rompaey L, Baens M, Van den Berghe H, and Marynen P (1997). Fusion of TEL, the ETS-variant gene 6 (ETV6), to the receptor-associated kinase JAK2 as a result of t(9;12) in a lymphoid and t(9;15;12) in a myeloid leukemia. *Blood* 90 (7): 2535-40.

Pelech SL, and Sanghera JS (1992). MAP kinases: charting the regulatory pathways. *Science* 257 (5075): 1355-6.

Peter M, Couturier J, Pacquement H, Michon J, Thomas G, Magdelenat H, and Delattre O (1997). A new member of the ETS family fused to EWS in Ewing tumors. *Oncogene* 14 (10): 1159-64.

Petersen JM, Skalicky JJ, Donaldson LW, McIntosh LP, Alber T, and Graves BJ (1995). Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix. *Science* 269 (5232): 1866-9.

Pio F, Kodandapani R, Ni CZ, Shepard W, Klemsz M, McKercher SR, Maki RA, and Ely KR (1996). New insights on DNA recognition by ets proteins from the crystal structure of the PU.1 ETS domain-DNA complex. *J. Biol. Chem.* 271 (38): 23329-37.

Pongubala JM, Nagulapalli S, Klemsz MJ, McKercher SR, Maki RA, and Atchison ML (1992). PU.1 recruits a second nuclear factor to a site important for immunoglobulin kappa 3' enhancer activity. *Mol. Cell Biol.* 12 (1): 368-78.

Potter MD, Buijs A, Kreider B, van Rompaey L, and Grosveld GC (2000). Identification and characterization of a new human ETS-family transcription factor, TEL2, that is expressed in hematopoietic tissues and can associate with TEL1/ETV6. *Blood* 95 (11): 3341-8.

Press MF, Cordon-Cardo C, and Slamon DJ (1990). Expression of the HER-2/neu proto-oncogene in normal human adult and fetal tissues. *Oncogene* 5 (7): 953-62.

Price MA, Rogers AE, and Treisman R (1995). Comparative analysis of the ternary complex factors Elk-1, SAP-1a and SAP-2 (ERP/NET). *EMBO J.* 14 (11): 2589-601.

Rabault B, and Ghysdael J (1994). Calcium-induced phosphorylation of ETS1 inhibits its specific DNA binding activity. *J. Biol. Chem.* 269 (45): 28143-51.

Rabault B, Roussel MF, Quang CT, and Ghysdael J (1996). Phosphorylation of Ets1 regulates the complementation of a CSF-1 receptor impaired in mitogenesis. *Oncogene* 13 (4): 877-81.

Radke K, Beug H, Kornfeld S, and Graf T (1982). Transformation of both erythroid and myeloid cells by E26, an avian leukemia virus that containsthe myb gene. *Cell* 31 (3 Pt 2): 643-53.

Ramakrishnan V, Finch JT, Graziano V, Lee PL, and Sweet RM (1993). Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* 362 (6417): 219-23.

Rao VN, and Reddy ES (1992). elk-1 domains responsible for autonomous DNA binding, SRE:SRF interaction and negative regulation of DNA binding. *Oncogene* 7 (11): 2335-40.

Rao VN, Ohno T, Prasad DD, Bhattacharya G, and Reddy ES (1993). Analysis of the DNA-binding and transcriptional activation functions of human Fli-1 protein. *Oncogene* 8 (8): 2167-73.

Ray-Gallet D, Mao C, Tavitian A, and Moreau-Gachelin F (1995). DNA binding specificities of Spi-1/PU.1 and Spi-B transcription factors and identification of a Spi-1/Spi-B binding site in the c-fes/c-fps promoter. *Oncogene* 11 (2): 303-13.

Raynaud SD, Baens M, Grosgeorge J, Rodgers K, Reid CD, Dainton M, Dyer M, Fuzibet JG, Gratecos N, Taillan B, Ayraud N, and Marynen P (1996a). Fluorescence in situ hybridization analysis of t(3; 12)(q26; p13): a recurring chromosomal abnormality involving the TEL gene (ETV6) in myelodysplastic syndromes. *Blood* 88 (2): 682-9.

Raynaud S, Cave H, Baens M, Bastard C, Cacheux V, Grosgeorge J, Guidal-Giroux C, Guo C, Vilmer E, Marynen P, and Grandchamp B (1996b). The 12;21 translocation involving TEL and deletion of

the other TEL allele: two frequently associated alterations found in childhood acute lymphoblastic leukemia. *Blood* 87 (7): 2891-9.

Rebay I, and Rubin GM (1995). Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* 81 (6): 857-66.

Risau W, Sariola H, Zerwes HG, Sasse J, Ekblom P, Kemler R, and Doetschman T (1988). Vasculogenesis and angiogenesis in embryonic-stem-cell-derived embryoid bodies. *Development* 102 (3): 471-8.

Roberson MS, Misra-Press A, Laurance ME, Stork PJ, and Maurer RA (1995). A role for mitogen-activated protein kinase in mediating activation of the glycoprotein hormone alpha-subunit promoter by gonadotropin-releasing hormone. *Mol. Cell Biol.* 15 (7): 3531-9.

Robertson E (1987). Teratocarcinomas and embryonic stem cells a practical approach. Practical Approach Series. pp. 71-112. IRL Press, Oxford.

Robinson GW, McKnight RA, Smith GH, and Hennighausen L (1995). Mammary epithelial cells undergo secretory differentiation in cycling virgins but require pregnancy for the establishment of terminal differentiation. *Development* 121 (7): 2079-90.

Robinson GW, Hennighausen L, and Johnson PF (2000). Side-branching in the mammary gland: the progesterone-Wnt connection. *Genes Dev.* 14 (8): 889-94.

Rohwedel J, Maltsev V, Bober E, Arnold HH, Hescheler J, and Wobus AM (1994). Muscle cell differentiation of embryonic stem cells reflects myogenesis *in vivo*: developmentally regulated expression of myogenic determination genes and functional expression of ionic currents. *Dev. Biol.* 164 (1): 87-101.

Romana SP, Mauchauffe M, Le Coniat M, Chumakov I, Le Paslier D, Berger R, and Bernard OA (1995a). The t(12;21) of acute lymphoblastic leukemia results in a tel-AML1 gene fusion. *Blood* 85 (12): 3662-70.

Romana SP, Poirel H, Leconiat M, Flexor MA, Mauchauffe M, Jonveaux P, Macintyre EA, Berger R, and Bernard OA (1995b). High frequency of t(12;21) in childhood B-lineage acute lymphoblastic leukemia. *Blood* 86 (11): 4263-9.

Rosmarin AG, Luo M, Caprio DG, Shang J, and Simkevich CP (1998). Sp1 cooperates with the ets transcription factor, GABP, to activate the CD18 (beta2 leukocyte integrin) promoter. *J. Biol. Chem.* 273 (21): 13097-103.

Samakovlis C, Hacohen N, Manning G, Sutherland DC, Guillemin K, and Krasnow MA (1996). Development of the Drosophila tracheal system occurs by a series of morphologically distinct but genetically coupled branching events. *Development* 122 (5): 1395-407.

Sambrook J, Fritsch EF, and Maniatis T (1997). Molecular cloning: A laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press, USA.

Savant-Bhonsale S, and Cleveland DW (1992). Evidence for instability of mRNAs containing AUUUA motifs mediated through translation-dependent assembly of a > 20S degradation complex. *Genes Dev.* 6 (10): 1927-39.

Schaapveld RQ, Schepens JT, Robinson GW, Attema J, Oerlemans FT, Fransen JA, Streuli M, Wieringa B, Hennighausen L, and Hendriks WJ (1997). Impaired mammary gland development and function in mice lacking LAR receptor-like tyrosine phosphatase activity. *Dev. Biol.* 188 (1): 134-46.

Schneikert J, Lutz Y, and Wasylyk B (1992). Two independent activation domains in c-Ets-1 and c-Ets-2 located in non-conserved sequences of the ets gene family. *Oncogene* 7 (2): 249-56.

Scholer HR (1991). Octamania: the POU factors in murine development. *Trends Genet.* 7 (10): 323-9.

Scholz H, Deatrick J, Klaes A, and Klambt C (1993). Genetic dissection of pointed, a Drosophila gene encoding two ETS-related proteins. *Genetics* 135 (2): 455-68.

Schultz J, Ponting CP, Hofmann K, and Bork P (1997). SAM as a protein interaction domain involved in developmental regulation. *Protein Sci.* 6 (1): 249-53.

Schultz SC, Shields GC, and Steitz TA (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* 253 (5023): 1001-7.

Schulz RA, Hogue DA, and The SM (1993). Characterization of lethal alleles of D-elg, an ets proto-oncogene related gene with multiple functions in Drosophila development. *Oncogene* 8 (12): 3369-74.

Scott EW, Simon MC, Anastasi J, and Singh H (1994a). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* 265 (5178): 1573-7.

Scott GK, Daniel JC, Xiong X, Maki RA, Kabat D, and Benz CC (1994b). Binding of an ETS-related protein within the DNase I hypersensitive site of the HER2/neu promoter in human breast cancer cells. *J. Biol. Chem.* 269 (31): 19848-58.

Seth A, Watson DK, Blair DG, and Papas TS (1989). c-ets-2 protooncogene has mitogenic and oncogenic activity. *Proc. Natl. Acad. Sci. U S A* 86 (20): 7833-7.

Seth A, and Papas TS (1990). The c-ets-1 proto-oncogene has oncogenic activity and is positively autoregulated. *Oncogene* 5 (12): 1761-7.

Seth A, Ascione R, Fisher RJ, Mavrothalassitis GJ, Bhat NK, and Papas TS (1992). The ets gene family. *Cell Growth Differ.*3 (5): 327-34.

Sgouras DN, Athanasiou MA, Beal GJ Jr, Fisher RJ, Blair DG, and Mavrothalassitis GJ (1995). ERF: an ETS domain protein with strong transcriptional repressor activity, can suppress ets-associated tumorigenesis and is regulated by phosphorylation during cell cycle and mitogenic stimulation. *EMBO J.* 14 (19): 4781-93.

Sharrocks AD, Brown AL, Ling Y, and Yates PR (1997). The ETS-domain transcription factor family. *Int. J. Biochem. Cell Biol.* 29 (12): 1371-87.

Shaw PE, Schroter H, and Nordheim A (1989). The ability of a ternary complex to form over the serum response element correlates with serum inducibility of the human c-fos promoter. *Cell* 56 (4): 563-72.

Shenk MA, and Steele RE (1993). A molecular snapshot of the metazoan 'Eve'. *Trends Biochem. Sci.* 18 (12): 459-63.

Shen-Li H, O'Hagan RC, Hou H Jr, Horner JW 2nd, Lee HW, and DePinho RA (2000). Essential role for Max in early embryonic growth and development. *Genes Dev.* 14 (1): 17-22.

Sheridan PL, Sheline CT, Cannon K, Voz ML, Pazin MJ, Kadonaga JT, and Jones KA (1995). Activation of the HIV-1 enhancer by the LEF-1 HMG protein on nucleosome-assembled DNA *in vitro*. *Genes Dev.* 9 (17): 2090-104.

Sherr CJ (1994). G1 phase progression: cycling on cue. *Cell* 79 (4): 551-5.

Sherr CJ (1996). Cancer cell cycles. *Science* 274 (5293): 1672-7.

Shin MK, and Koshland ME (1993). Ets-related protein PU.1 regulates expression of the immunoglobulin J-chain gene through a novel Ets-binding element. *Genes Dev.* 7 (10): 2006-15.

Shirasaki F, Makhluf HA, LeRoy C, Watson DK, and Trojanowska M (1999). Ets transcription factors cooperate with Sp1 to activate the human tenascin-C promoter. *Oncogene* 18 (54): 7755-64.

Shore P, and Sharrocks AD (1994). The transcription factors Elk-1 and serum response factor interact by direct protein-protein contacts mediated by a short region of Elk-1. *Mol. Cell Biol.* 14 (5): 3283-91.

Shore P, and Sharrocks AD (1995). The ETS-domain transcription factors Elk-1 and SAP-1 exhibit differential DNA binding specificities. *Nucleic Acids Res.* 23 (22): 4698-706.

Shore P, Whitmarsh AJ, Bhaskaran R, Davis RJ, Waltho JP, and Sharrocks AD (1996). Determinants of DNA-binding specificity of ETS-domain transcription factors. *Mol. Cell Biol.* 16 (7): 3338-49.

Shurtleff SA, Buijs A, Behm FG, Rubnitz JE, Raimondi SC, Hancock ML, Chan GC, Pui CH, Grosveld G, and Downing JR (1995). TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup of patients with an excellent prognosis. *Leukemia* 9 (12): 1985-9.

Sicinski P, Donaher JL, Parker SB, Li T, Fazeli A, Gardner H, Haslam SZ, Bronson RT, Elledge SJ, and Weingerg RA (1995). Cyclin D1 provides a link between development and oncogenesis in the retina and breast. *Cell* 82 (4): 621-30.

Siddique HR, Rao VN, Lee L, and Reddy ES. Characterization of the DNA binding and transcriptional activation domains of the erg protein. *Oncogene* 8 (7): 1751-5.

Singleton TP, and Strickler JG (1992). Clinical and pathologic significance of the c-erbB-2 (HER-2/neu) oncogene. *Pathol. Annu.* 27 Pt 1: 165-90.

Slupsky CM, Gentile LN, Donaldson LW, Mackereth CD, Seidel JJ, Graves BJ, and McIntosh LP (1998). Structure of the Ets-1 pointed domain and mitogen-activated protein kinase phosphorylation site. *Proc. Natl. Acad. Sci. U S A* 95 (21): 12129-34.

Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ, and Denny CT (1994). A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat. Genet.* 6 (2): 146-51.

Spyropoulos DD, Pharr PN, Lavenburg KR, Jackers P, Papas TS, Ogawa M, and Watson DK (2000). Hemorrhage, impaired hematopoiesis, and lethality in mouse embryos carrying a targeted disruption of the Fli1 transcription factor. *Mol. Cell Biol.* 20 (15): 5643-52.

Stapleton D, Balan I, Pawson T, and Sicheri F (1999). The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nat. Struct. Biol.* 6 (1): 44-9.

Sternlicht MD, and Werb Z (1999). ECM proteinases. In Guidebook to the Extracellular Matrix, Anchor, and Adhesion Proteins. (eds. Kreis T, and Vale R), pp.503-63. Oxford University Press, Oxford.

St-Onge L, Sosa-Pineda B, Chowdhury K, Mansouri A, and Gruss P (1997). Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. *Nature* 387 (6631): 406-9.

Su GH, Ip HS, Cobb BS, Lu MM, Chen HM, and Simon MC (1996). The Ets protein Spi-B is expressed exclusively in B cells and T cells during development. *J. Exp. Med.* 184 (1): 203-14.

Su GH, Chen HM, Muthusamy N, Garrett-Sinha LA, Baunoch D, Tenen DG, and Simon MC (1997). Defective B cell receptor-mediated responses in mice lacking the Ets protein, Spi-B. *EMBO J.*16 (23): 7118-29.

Sumarsono SH, Wilson TJ, Tymms MJ, Venter DJ, Corrick CM, Kola R, Lahoud MH, Papas TS, Seth A, and Kola I (1996). Down's syndrome-like skeletal abnormalities in Ets2 transgenic mice. *Nature* 379 (6565): 534-7.

Suzuki H, Romano-Spica V, Papas TS, and Bhat NK (1995). ETS1 suppresses tumorigenicity of human colon cancer cells. *Proc. Natl. Acad. Sci. U S A* 92 (10): 4442-6.

Svendsen P, Laursen J, Krogh-Pedersen H, and Hjorth JP (1998). Novel salivary gland specific binding elements located in the PSP proximal enhancer core. *Nucleic Acids Res.* 26 (11): 2761-70.

Takai H, Tominaga K, Motoyama N, Minamishima YA, Nagahama H, Tsukiyama T, Ikeda K, Nakayama K, Nakanishi M, and Nakayama Ki (2000). Aberrant cell cycle checkpoint function and early embryonic death in Chk1(-/-) mice. *Genes Dev.* 14 (12): 1439-47.

Tanaka K, Iwakuma T, Harimaya K, Sato H, and Iwamoto Y (1997). EWS-Fli1 antisense oligodeoxynucleotide inhibits proliferation of human Ewing's sarcoma and primitive neuroectodermal tumor cells. *J. Clin. Invest.* 99 (2): 239-47.

te Riele H, Maandag ER, and Berns A (1992). Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc. Natl. Acad. Sci. U S A* 89 (11): 5128-32.

Teglund S, McKay C, Schuetz E, van Deursen JM, Stravopodis D, Wang D, Brown M, Bodner S, Grosveld G, and Ihle JN (1998). Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. *Cell* 93 (5): 841-50.

Thomas KR, Folger KR, ans Capecchi MR (1986). High frequency targeting of genes to specific sites in the mammalian genome. *Cell* 44 (3): 419-28.

Thomas KR, and Capecchi MR (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 51 (3): 503-12.

Thomas RS, Tymms MJ, Seth A, Shannon MF, and Kola I (1995). ETS1 transactivates the human GM-CSF promoter in Jurkat T cells stimulated with PMA and ionomycin. *Oncogene* 11 (10): 2135-43.

Thomas RS, Tymms MJ, McKinlay LH, Shannon MF, Seth A, and Kola I (1997). ETS1, NFkappaB and AP1 synergistically transactivate the human GM-CSF promoter. *Oncogene* 14 (23): 2845-55.

Thomas RS, Ng AN, Zhou J, Tymms MJ, Doppler W, and Kola I (2000). The Elf group of Ets-related transcription factors. ELF3 and ELF5. *Adv. Exp. Med. Biol.* 480: 123-8.

Thompson CC, Brown TA, and McKnight SL (1991). Convergence of Ets- and notch-related structural motifs in a heteromeric DNA binding complex. *Science* 253 (5021): 762-8.

Thompson JD, Higgins DG, and Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22 (22): 4673-80.

Treier M, Bohmann D, and Mlodzik M (1995). JUN cooperates with the ETS domain protein pointed to induce photoreceptor R7 fate in the Drosophila eye. *Cell* 83 (5): 753-60.

Treisman R (1994). Ternary complex factors: growth factor regulated transcriptional activators. *Curr. Opin. Genet. Dev.* 4 (1): 96-101.

Tripathy D, and Benz C (1994). Growth factors and their receptors. *Hematol. Oncol. Clin. North. Am.* 8 (1): 29-50.

Tsubura Y, Imai S, Morimoto J, and Hilgers J (1981). Strain difference in the expression of mammary tumor virus antigen in the male genital organs of mice during aging. *Gann.* 72 (3): 424-9.

Tymms MJ (1995). Quantitative measurement of mRNA using the RNase protection assay. *Methods Mol. Biol.* 37: 31-46.

Tymms MJ, Ng AY, Thomas RS, Schutte BC, Zhou J, Eyre HJ, Sutherland GR, Seth A, Rosenberg M, Papas T, Debouck C, and Kola I (1997). A novel epithelial-expressed ETS gene, ELF3: human and murine cDNA sequences, murine genomic organization human mapping to 1q32.2 and expression in tissues and cancer. *Oncogene* 15 (20): 2449-62.

Udy GB, Towers RP, Snell RG, Wilkins RJ, Park SH, Ram PA, Waxman DJ, and Davey HW (1997). Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. *Proc. Natl. Acad. Sci. U S A* 94 (14): 239-44.

Urness LD, and Thummel CS (1990). Molecular interactions within the ecdysone regulatory hierarchy: DNA binding properties of the Drosophila ecdysone-inducible E74A protein. *Cell* 63 (1): 47-61.

Vonderhaar B (1988). Regulation of development of the normal mammary gland by hormones and growth factors, In Breast cancer: Cellular and molecular biology. (eds. Lippman M, and Dickson R), pp.252-66. Kluwer Academic Publishers, Boston, USA.

Vuister GW, Kim SJ, Orosz A, Marquardt J, Wu C, and Bax A (1994). Solution structure of the DNA-binding domain of Drosophila heat shock transcription factor. *Nat. Struct. Biol.* 1 (9): 605-14.

Vuister GW, Kim SJ, Wu C, and Bax A (1994). NMR evidence for similarities between the DNA-binding regions of Drosophila melanogaster heat shock factor and the helix-turn-helix and HNF-3/forkhead families of transcription factors. *Biochemistry* 33 (1): 10-6.

Wagner KU, Young WS 3rd, Liu X, Ginns EI, Li M, Furth PA, and Hennighausen L (1997). Oxytocin and milk removal are required for post-partum mammary-gland development. *Genes Funct.* 1 (4): 233-44.

Wakao H, Gouilleux F, and Groner B (1994). Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *EMBO J.* 13 (9): 2182-91.

Walker MD, Edlund T, Boulet AM, and Rutter WJ (1983). Cell-specific expression controlled by the 5'-flanking region of insulin and chymotrypsin genes. *Nature* 306 (5943): 557-61.

Walunas TL, Wang B, Wang CR, and Leiden JM (2000). Cutting edge: the Ets1 transcription factor is required for the development of NK T cells in mice. *J. Immunol.* 164 (6): 2857-60.

Wang CY, Petryniak B, Ho IC, Thompson CB, and Leiden JM (1992). Evolutionarily conserved Ets family members display distinct DNA binding specificities. *J. Exp. Med.* 175 (5): 1391-9.

Wang LC, Kuo F, Fujiwara Y, Gilliland DG, Golub TR, and Orkin SH (1997). Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. *EMBO J.* 16 (14): 4374-83.

Wasylyk B, Wasylyk C, Flores P, Begue A, Leprince D, and Stehelin D (1990). The c-ets proto-oncogenes encode transcription factors that cooperate with c-Fos and c-Jun for transcriptional activation. *Nature* 346 (6280): 191-3.

Wasylyk B, Hahn SL, and Giovane A (1993). The Ets family of transcription factors. *Eur. J. Biochem.* 211 (1-2): 7-18.

Wasylyk C, Flores P, Gutman A, and Wasylyk B (1989). PEA3 is a nuclear target for transcription activation by non-nuclear oncogenes. *EMBO J.* 8 (11): 3371-8.

Wasylyk C, Gutman A, Nicholson R, and Wasylyk B (1991). The c-Ets oncoprotein activates the stromelysin promoter through the same elements as several non-nuclear oncoproteins. *EMBO J.* 10 (5): 1127-34.

Wasylyk C., Kerckaert JP, and Wasylyk B (1992). A novel modulator domain of Ets transcription factors. *Genes Dev.* 6 (6): 965-74.

Watson DK, McWilliams MJ, Lapis P, Lautenberger JA, Schweinfest CW, and Papas TS (1988). Mammalian ets-1 and ets-2 genes encode highly conserved proteins. *Proc. Natl. Acad. Sci. U S A* 85 (21): 7862-6.

Watson DK, Smyth FE, Thompson DM, Cheng JQ, Testa JR, Papas TS, and Seth A (1992). The

ERGB/Fli-1 gene: isolation and characterization of a new member of the family of human ETS transcription factors. *Cell Growth Differ.* 3 (10): 705-13.

Weinberg RA (1995). The retinoblastoma protein and cell cycle control. *Cell* 81 (3): 323-30.

Welte T, Garimorth K, Philipp S, Jennewein P, Huck C, Cato AC, and Doppler W (1994). Involvement of Ets-related proteins in hormone-independent mammary cell-specific gene expression. *Eur. J. Biochem.* 223 (3): 997-1006.

Werner MH, Clore M, Fisher CL, Fisher RJ, Trinh L, Shiloach J, and Gronenborn AM (1995). The solution structure of the human ETS1-DNA complex reveals a novel mode of binding andtrue side chain intercalation. *Cell* 83 (5): 761-71.

Werner MH, Gronenborn AM, and Clore GM (1997a). ETS1-DNA binding and intercalation: correction. *Science* 276 (5321): 1957.

Werner MH, Clore GM, Fisher CL, Fisher RJ, Trinh L, Shiloach J, and Gronenborn AM (1997b). Correction of the NMR structure of the ETS1/DNA complex. *J. Biomol. NMR* 10 (4): 317-28.

Wernert N, Raes MB, Lassalle P, Dehouck MP, Gosselin B, Vandenbunder B, and Stehelin D (1992). c-ets1 proto-oncogene is a transcription factor expressed in endothelial cells during tumor vascularization and other forms of angiogenesis in humans. *Am. J. Pathol.* 140 (1): 119-27.

Wheat W, Fitzsimmons D, Lennox H, Krautkramer SR, Gentile LN, McIntosh LP, and Hagman J (1999). The highly conserved beta-hairpin of the paired DNA-binding domain is required for assembly of Pax-Ets ternary complexes. *Mol. Cell. Biol.* 19 (3): 2231-41.

Whitmarsh AJ, Shore P, Sharrocks AD, and Davis RJ (1995). Integration of MAP kinase signal transduction pathways at the serum response element. *Science* 269 (5222): 403-7.

Wiesen JF, Young P, Werb Z, and Cunha GR (1999). Signaling through the stromal epidermal growth factor receptor is necessary for mammary ductal development. *Development* 126 (2): 335-44.

Wiles MV, and Keller G (1991). Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture. *Development* 111 (2): 259-67.

Wilson AP, Dent M, Pejovic T, Hubbold L, and Radford H (1996). Characterisation of seven human ovarian tumor cell lines. *Br. J. Cancer* 74 (5): 722-7.
Wilson KP, Shewchuk LM, Brennan RG, Otsuka AJ, and Matthews BW (1992). Escherichia coli biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc. Natl. Acad. Sci. U S A* 89 (19): 9257-61.

Wynick D, Small CJ, Bacon A, Holmes FE, Norman M, Ormandy CJ, Kilic E, Kerr NC, Ghatei M, Talamantes F, Bloom SR, and Pachnis V (1998). Galanin regulates prolactin release and lactotroph proliferation. *Proc. Natl. Acad. Sci. U S A* 95 (21): 12671-6.

Xing X, Wang SC, Xia W, Zou Y, Shao R, Kwong KY, Yu Z, Zhang S, Miller S, Huang L, and Hung MC (2000). The ets protein PEA3 suppresses HER-2/neu overexpression and inhibits tumorigenesis. *Nat. Med.* 6 (2): 189-95.

Xu X, Wagner KU, Larson D, Weaver Z, Li C, Ried T, Hennighausen L, Wynshaw-Boris A, and Deng CX (1999). Conditional mutation of *Brca1* in mammary epithelial cells results in blunted ductal morphogenesis and tumor formation. *Nat. Genet.* 22 (1): 37-43.

Yamada N, Tamai Y, Miyamoto H, and Nozaki M (2000). Cloning and expression of the mouse Pse gene encoding a novel Ets family member. *Gene* 241 (2): 267-74.

Yamamoto H, Flannery ML, Kupriyanov S, Pearce J, McKercher SR, Henkel GW, Maki RA, Werb Z, and Oshima RG (1998). Defective trophoblast function in mice with a targeted mutation of Ets2. *Genes Dev.* 12 (9): 1315-26.

Yang BS, Hauser CA, Henkel G, Colman MS, Van Beveren C, Stacey KJ, Hume DA, Maki RA, and Ostrowski MC (1996). Ras-mediated phosphorylation of a conserved threonine residue enhances the transactivation activities of c-Ets1 and c-Ets2. *Mol. Cell Biol.* 16 (2): 538-47.

Yi H, Fujimura Y, Ouchida M, Prasad DD, Rao VN, and Reddy ES (1997). Inhibition of apoptosis by normal and aberrant Fli-1 and erg proteins involved in human solid tumors and leukemias. *Oncogene* 14 (11): 1259-68.

Young WS 3rd, Shepard E, Amico J, Hennighausen L, Wagner KU, LaMarca ME, McKinney C, and Ginns EI (1996). Deficiency in mouse oxytocin prevents milk ejection, but not fertility or parturition. *J. Neuroendocrinol.* 8 (11): 847-53

Yuan CC, Kan N, Dunn KJ, Papas TS, and Blair DG (1989). Properties of a murine retroviral recombinant of avian acute leukemia virus E26: a murine fibroblast assay for v-ets function. *J Virol.* 63 (1): 205-15.

Zelena J (1994). Nerves and mechanoreceptors - the role of innervation in the development and maintenance of mammalian mechanoreceptors. pp. 1-137. Chapman and Hall, New York.

Zenklusen JC, Oshimura M, Barrett JC, and Conti CJ (1995). Human chromosome 11 inhibits tumorigenicity of a murine squamous cell carcinoma cell line. *Genes Chromosomes Cancer* 13 (1): 47-53.

Zhang L, Eddy A, Teng YT, Fritzler M, Kluppel M, Melet F, and Bernstein A (1995). An immunological renal disease in transgenic mice that overexpress Fli-1, a member of the ets family of transcription factor genes. *Mol. Cell Biol.* 15 (12): 6961-70.

Zhang M, Maass N, Magit D, and Sager R (1997a). Transactivation through Ets and Ap1 transcription sites determines the expression of the tumor-suppressing gene maspin. *Cell Growth Differ.* 8 (2): 179-86.

Zhang M, Magit D, and Sager R (1997b). Expression of maspin in prostate cells is regulated by a positive ets element and a negative hormonal responsive element site recognized by androgen receptor. *Proc. Natl. Acad. Sci. U S A* 94 (11): 5673-8.

Zhang P, Behre G, Pan J, Iwama A, Wara-Aswapati N, Radomska HS, Auron PE, Tenen DG, and Sun Z (1999). Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc. Natl. Acad. Sci. U S A* 96 (15): 8705-10.

Zhou J, NG AY, Tymms MJ, Jermiin LS, Seth AK, Thomas RS, and Kola I (1998). A novel transcription factor, ELF5, belongs to the ELF subfamily of ETS genes and maps to human chromosome 11p13-15, a region subject to LOH and rearrangement in human carcinoma cell lines. *Oncogene* 17 (21): 2719-32.

Zou Z, Anisowicz A, Hendrix MJ, Thor A, Neveu M, Sheng S, Rafidi K, Seftor E, and Sager R (1994). Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells. *Science* 263 (5146): 526-9.

15.

The Elf Group of Ets-Related Transcription Factors
ELF3 and ELF5

Ross S. Thomas[2], Annie N.Y. Ng[1], Jiong Zhou[1], Martin J. Tymms[3], Wolfgang Doppler[4] and Ismail Kola[1]

[1] Centre for Functional Genomics and Human Disease, Monash University, Melbourne, Australia: [2] Transcription Laboratory, Imperial Cancer Research Fund, London, UK: [3] National Vision Research Institute of Australia, Melbourne, Australia: [4] Institut für Medizinische Chemie und Biochemie, Innsbruck, Austria

Key words ELF3, ELF5, transcription factors

1. INTRODUCTION

We provide an initial characterisation of two new mammalian ETS transcription factors, ELF3 and ELF5. Expression of ELF3 and ELF5 appears to be restricted to the epithelial cells of multiple organs, and we are examining their role in normal mammary differentiation and function, and in mammary neoplasia. We show evidence that both ELF3 and ELF5 are able to positively regulate transcription of the whey acidic protein (WAP) promoter in mammary epithelial cells, independently of hormone treatment.

1.1 Background

One of the major milk proteins produced by mammary epithelial cells during pregnancy and lactation is whey acidic protein (WAP). WAP expression is very low until mid-gestation, after which time the combined actions of the glucocorticoid and prolactin pathways produce a several thousand-fold increase in transcription. Transcription is also partly dependent upon the binding of factors to the mammary-cell activating factor (MAF) site[1,2], a conserved ETS-like element present in the WAP proximal promoter and in the regulatory regions of other mammary specific genes.

The ETS family of transcription factors regulate gene expression during normal biological processes such as haemopoiesis, angiogenesis and cartilage/skeletal development, and aberrant forms have also been implicated with various types of neoplasia. Despite the large variety of ETS factors, functions in epithelial cells, such as those in the mammary gland, have not been examined extensively. Overexpression of PEA3 is, however, associated with aggressive subsets of breast carcinoma that also overexpress erbB2[3]. PEA3 does not appear to be involved in normal mammary function, but may be required for early mammary development.

We have recently cloned two new ETS transcription factors, ELF3[4] and ELF5[5]. The HUGO Nomenclature Committee has upheld the ELF naming convention despite ELF3 being cloned and differently named by other groups (ESX[6], ESE-1[7], jen[8], ERT[9]).

2. EXPRESSION OF ELF3 AND ELF5

We examined the expression of ELF3 and ELF5 in mouse organs by Northern blot analysis of poly A^+ RNA. The expression patterns of these two genes were similar, but both differ significantly to the expression of other ETS family members. Most other ETS factors are expressed strongly in haemopoietic compartments, but ELF3 and ELF5 levels were undetectable in adult thymus or spleen (Figs 1a and 2a), nor in peripheral lymphocytes of leukaemic cell lines (data not shown). Strong expression, however, was observed in a subset of organs that contain secretory epithelial cells, such as the lung, stomach, prostate and mammary gland.

The mammary gland undergoes distinct phases of growth, differentiation and regression; from virginal mice to pregnancy, parturition and weaning. Examination of expression levels through these stages revealed that ELF3 is present in virginal mammary gland and during early pregnancy, when cells are proliferating, but drops in late pregnancy when epithelial cells undergo terminal differentiation and begin to produce milk components (Fig 1b). Expression remains low during lactation, but reappears upon weaning, when most of the epithelial cells in the mammary gland are undergoing apoptosis. Another group's study made similar findings[10]. *In situ* expression analysis in human mammary gland confirms that ELF3 is expressed specifically in the epithelial cells of the ductules and lobular structures (Fig 1c).

One possible explanation for the pattern of ELF3 expression is that it is associated with a small proportion of epithelial stem cells, rather than the differentiated functional cells, and that these cells survive apoptosis to re-colonise the mammary gland during the next pregnancy.

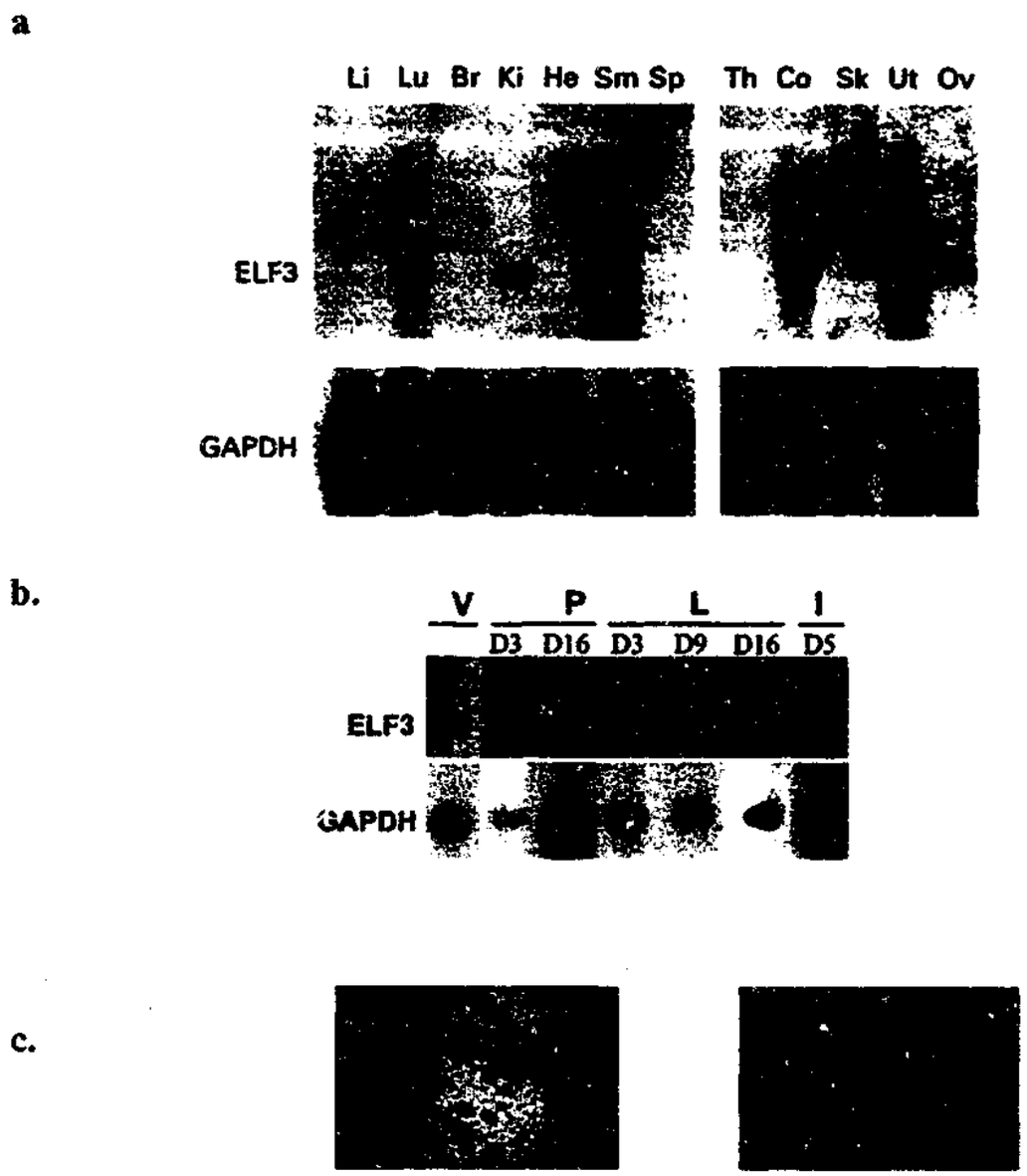
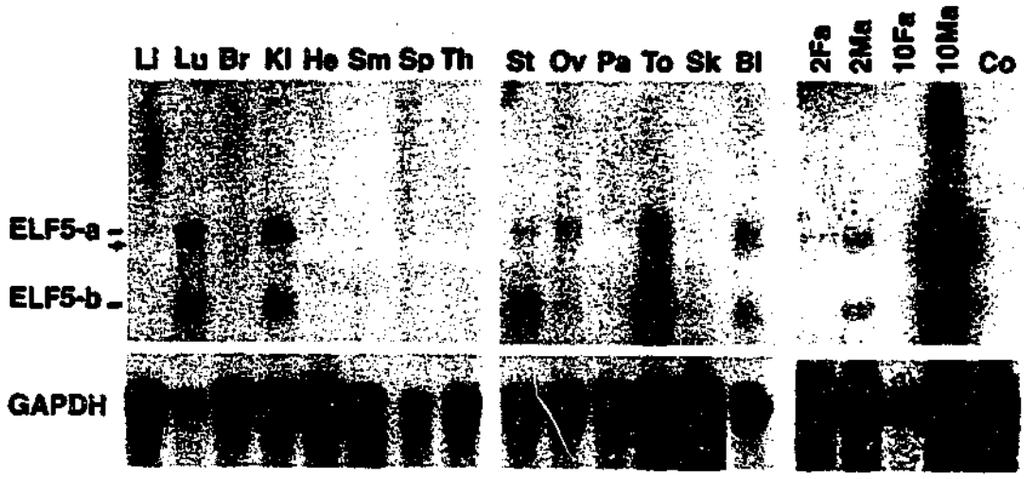


Fig. 1. Expression of ELF3. **a.** Northern blot of mouse organs probed with labelled ELF3 and GAPDH cDNAs. Li; liver, Lu; lung, Br; brain, Ki kidney, He; heart, Sm; small intestine, Sp; spleen, Th; thymus, Co; colon, Sk; skin, Ut; uterus, Ov; ovary. **b.** Northern blot of mouse mammary gland. V; virgin, P; pregnant, L; lactating, I; involuting. **c.** Human mammary gland section probed with labelled ELF3 cDNA (light field and exposure).

ELF5 is also expressed specifically in human mammary gland epithelial cells (Fig 2b), and, in two time points, mammary ELF5 expression is increased during pregnancy in the mouse (Fig 2a), suggesting that the temporal pattern of ELF5 expression may be very different to that of ELF3.

a.



b.

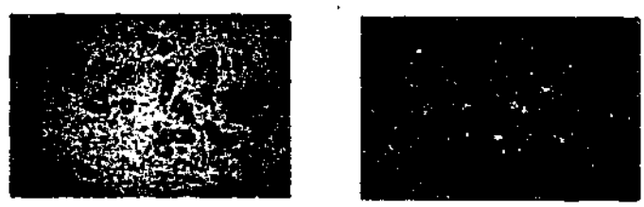


Fig. 2. Expression of ELF5. **a.** Northern blot of mouse organs probed with labelled ELF5 and GAPDH cDNAs. As above, and additional organs St; stomach, Pa; pancreas, To; tongue, Bl; bladder, 2Fa; fat, 2Ma; mammary gland day2 pregnancy, 10Fa; fat, 10Ma; mammary gland day 10 pregnancy. **b.** Human mammary gland section probed with labelled ELF5 cDNA (light field and exposure).

3. FUNCTION OF ELF3 AND ELF5 IN MAMMARY GLAND

Expression of the WAP gene has been studied extensively as it is highly specific to mammary epithelial cells and is induced several thousand fold during late pregnancy and lactation. The WAP promoter responds to lactogenic hormones, but also contains other important transcription factor elements that are essential for expression. *In vitro* studies indicate that the MAF site is critical for maximal WAP expression, independently of hormone treatment[1]. Mouse transgenic studies indicate that the site is important for expression of WAP during pregnancy, but not critical during lactation, when the promoter is most active and under the control of hormones[2].

The MAF site is similar to others described as binding sites for the ETS family of transcription factors. We have investigated the ability of ELF3 and

ELF5 to control WAP transcription through the MAF site (herein called ETS/1) and another adjacent ETS site (ETS/2).

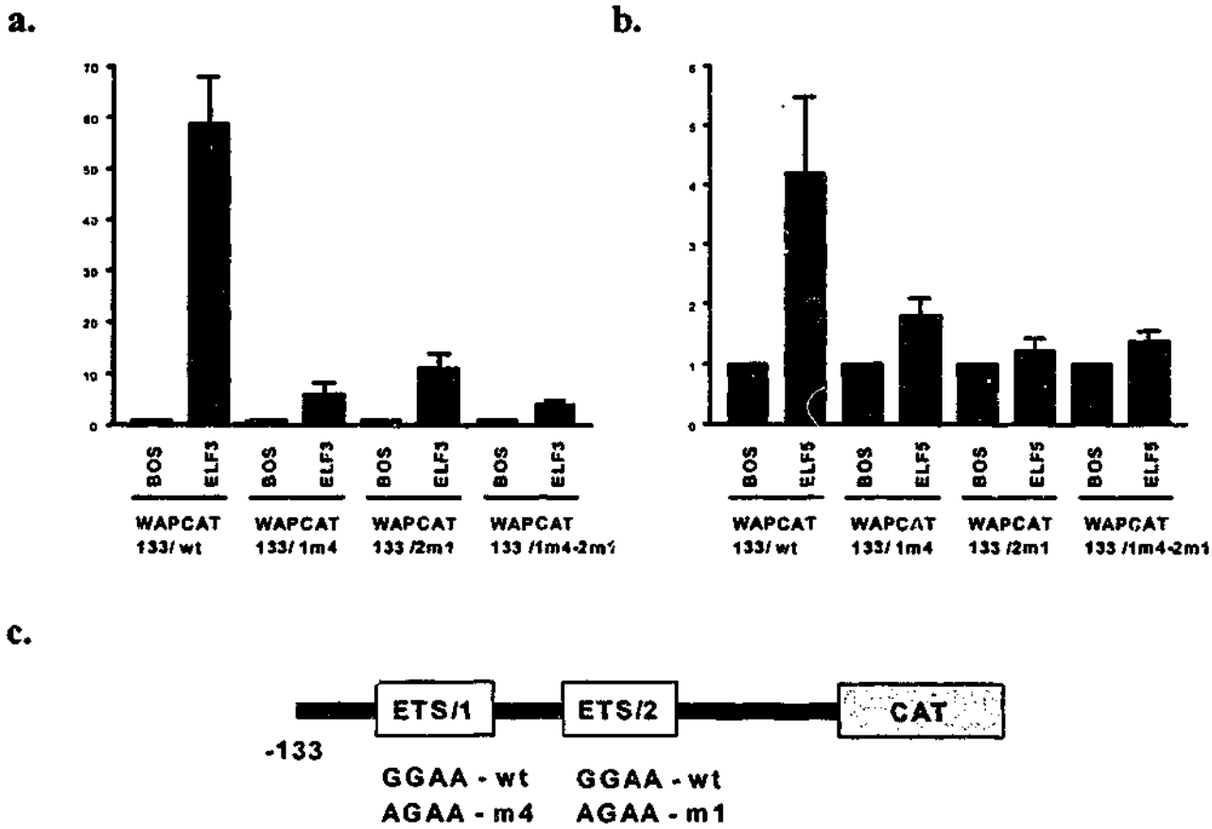


Figure 3. Transactivation of the WAP promoter by ELF3 and ELF5. **a.** Co-transfection of an ELF3 expression vector with WAP promoter-reporter constructs in HC-11 cells. Transcriptional activity is measured by CAT assay. BOS is an empty expression vector control. **b.** As above, but with ELF5. **c.** Schematic of the 133 bp WAP promoter attached to the CAT reporter gene. ETS/1 (MAF) and ETS/2 are binding sites for ELFs. Mutations m4 and m1 of the binding sites are indicated.

Recombinant ELF3 and ELF5 proteins are able to strongly interact with the ETS/1 site, and also more weakly with the ETS/2 site, in electrophoretic mobility shift assay (data not shown). Consistent with the binding, we found that both ELF3 (fig 3a) and ELF5 (fig 3b) were able to transactivate a WAP promoter construct in a mammary epithelial cell line, although ELF3 was much more potent than ELF5. Both factors were able to function independently of lactogenic hormones. A longer promoter construct, containing more distal hormone response elements, behaved similarly (data not shown).

Point mutations (fig 3c) of either ETS/1 or ETS/2 alone were enough to disable most transactivation, indicating the ELF3 and ELF5 operate through these sites. Both mutations together had only a slightly greater effect, suggesting that the ETS/1 and ETS/2 sites work cooperatively, both being required for maximal promoter activity.

4. CONCLUSION

Expression of mammary-specific genes is under the control of both hormone dependent and independent mechanisms, neither of which may be totally mammary-specific. It seems that a complex interplay of transcription factors may be responsible for mammary expression. Our challenge is to determine whether ELF3 and/or ELF5 are important for the expression of mammary-specific genes *in vivo*, and whether they cooperate with other stimuli, such as prolactin. Because the mammary gland is not a static tissue, but can continually undergo cycles of growth, differentiation and function, it will be interesting to examine whether ELF3 and ELF5 operate cooperatively or differentially during these different phases.

REFERENCES

1. Welte T, Garimorth K, Philipp S, Jennewein P, Huck C, Cato ACB and Doppler W, 1994, *Eur. J. Biochem.* **223**: 997-1006.
2. McKnight RA, Spencer M, Dittmer J, Brady JN, Wall RJ and Hennighausen L, 1995, *Mol. Endo.* **9**: 717-24.
3. Benz CC, O'Hagan RC, Richter B, Scott GK, Chang CH, Xiong X, Chew K, Ljung BM, Edgerton S, Thor A and Hassell JA, 1997, *Oncogene* **15**: 1513-25.
4. Tymms MJ, Ng AY, Thomas RS, Schutte BC, Zhou J, Eyre HJ, Sutherland GR, Seth A, Rosenberg M, Papas T, Debouck C, Kola I, 1997, *Oncogene* **15**: 2449-62.
5. Zhou J, NG AY, Tymms MJ, Jermiin LS, Seth AK, Thomas RS, Kola I, 1998 *Oncogene* **17**: 2719-32.
6. Chang CH, Scott GK, Kuo WL, Xiong X, Suzdaltseva Y, Park JW, Sayre P, Erny K, Collins C, Gray JW, Benz CC, 1997, *Oncogene* **14**: 1617-22.
7. Oettgen P, Alani RM, Barcinski MA, Brown L, Akbarali Y, Boltax J, Kunsch C, Munger K, Libermann TA, 1997, *Mol. Cell. Biol.* **17**: 4419-33.
8. Andreoli JM, Jang SI, Chung E, Coticchia CM, Steinert PM, Markova NG, 1997, *Nuc. Acids Res.* **25**: 4287-95.
9. Choi SG, Yi Y, Kim YS, Kato M, Chang J, Chung HW, Hahm KB, Yang HK, Rhee HH, Bang YJ, Kim SJ, 1998, *J. Biol. Chem.* **273**: 110-7.
10. Neve R, Chang, CH, Scott, GK, Wong, A, Friis RR, Hynes, NE and Benz CC, 1998, Faseb J. 12;1541-50.

Oncogene (1998) 17, 2719–2732

http://www.stockton-press.co.uk/onc

A novel transcription factor, ELF5, belongs to the ELF subfamily of ETS genes and maps to human chromosome 11p13–15, a region subject to LOH and rearrangement in human carcinoma cell lines

Jiong Zhou[1], Annie YN Ng[1], Martin J Tymms[2], Lars S Jermiin[3], Arun K Seth[4], Ross S Thomas[1] and Ismail Kola[1]

[1]*Molecular Genetics and Development Group, Institute of Reproduction and Development, Monash University, Melbourne, Victoria;* [2]*National Vision Research Institute of Australia, Melbourne, Victoria;* [3]*The John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia;* [4]*Department of Pathology, University of Toronto/Women's College Hospital, Toronto, Ontario, Canada*

The ETS transcription factors are a large family implicated in the control of cellular proliferation and tumorigenesis. In addition, chromosomal translocations involving ETS family members are associated with a range of different human cancers. Given the extensive involvement of ETS factors in tumorigenesis, it becomes important to identify any additional ETS genes that may also play oncogenic roles. We identify a novel gene, *ELF5*, that appears to belong to the ELF (E74-like-factor) subfamily of the ETS transcription factor family, based upon similarity within the 'ETS domain'. *ELF5* displays a similar, but more restricted, expression pattern to that of the newly isolated epithelium-specific ETS gene, *ELF3*. Unlike most other ETS family members, *ELF5* is not expressed in hematopoietic compartments, but is restricted to organs such as lung, stomach, kidney, prostate, bladder and mammary gland. *ELF5* is localized to human chromosome 11p13–15, a region that frequently undergoes loss of heterozygosity (LOH) in several types of carcinoma, including those of breast, kidney and prostate. We find that *ELF5* expression is not detectable in a number of carcinoma cell lines, some of which display loss or rearrangement of an *ELF5* allele. Similar to other ETS family members, ELF5 displays specific binding to DNA sequences containing a GGAA-core. In addition, ELF5 is able to transactivate through these ETS sequences, present upstream from a minimal promoter. Our data suggest that ELF5 may play roles in mammary, lung, prostate and/or kidney function, and possibly also in tumorigenesis.

Keywords: transcription factor; epithelial cells; cancer; erbB2; transcription; LOH

Introduction

The ETS family of transcription factors share a highly conserved DNA binding domain, termed the 'ETS domain', first identified in the *gag-myb-ets* fusion protein of avian leukemia virus E26 (Nunn *et al.*, 1983; Watson *et al.*, 1988; Karim *et al.*, 1990; Gutman and Wasylyk, 1991; Seth *et al.*, 1992). The ETS domain recognizes and binds to purine rich GGA(A/T) core motifs in the promoters and enhancers of various target genes (Macleod *et al.*, 1992; Wasylyk *et al.*, 1993; Janknecht and Nordheim, 1993; Werner *et al.*, 1995; Kodandapani *et al.*, 1996). The ETS family does not maintain overall similarity outside of the ETS domain, but can be grouped into subfamilies based upon variation within the ETS domain, and also by the arrangement and presence of other domains, such as those involved in transactivation and sites of phosphorylation (Lautenberger *et al.*, 1992; Wasylyk *et al.*, 1993; Janknecht and Nordheim, 1993). Over 30 ETS gene family members have been identified in species ranging from sea urchin to human.

Many ETS factors have been implicated in the control of cellular proliferation and tumorigenesis (Seth *et al.*, 1992; Macleod *et al.*, 1992; Wasylyk *et al.*, 1993; Janknecht and Nordheim, 1993; Scott *et al.*, 1994a; Muthusamy *et al.*, 1995). *ETS1*, *ETS2*, *ERG2* and *PU.1* are proto-oncogenes with mitogenic and transforming activity when overexpressed in fibroblasts (Seth *et al.*, 1989; Seth and Papas, 1990; Hart *et al.*, 1995; Moreau-Gachelin *et al.*, 1996). In addition, chromosomal translocations involving ETS family members are associated with different human cancers. *ERG* and *ERGB/FLI1* are fused to the *EWS* gene in t(21;22) and t(11;22) translocations, respectively, in Ewing's sarcoma and other primitive neuroectodermal tumors (Sorensen *et al.*, 1994; Ida *et al.*, 1995). *FEV* is fused to *EWS* in a subset of Ewing's tumors in t(2;22) translocations (Peter *et al.*, 1997). *TEL* is fused to the platelet-derived growth factor receptor beta (PDGFRβ) gene in t(5;12) translocations of chronic myelomonocytic leukemia, and to the acute myeloid leukemia 1 (AML1) transcription factor gene in t(12;21) translocations of acute lymphoblastic leukemia (Golub *et al.*, 1994, 1995). Fusion of *TEL* to the receptor-associated kinase JAK2 results in early pre-B acute lymphoid leukemia in t(9;12), and in atypical chronic myelogenous leukemia in t(9;15;12) (Peeters *et al.*, 1997). Expression of *Spi1* and *Fli1* can be activated by position specific integration of the Friend murine leukemia virus in murine erythroleukemias (Ben-David *et al.*, 1991). Also, ETS1, ETS2 and ERG regulate the expression of metalloproteinase genes, such as stromelysin and collagenase (Buttice and Kurkinen, 1993; Buttice *et al.*, 1996; Wasylyk *et al.*, 1991), which are important for extracellular matrix degradation concomitant with tumor vascularization (angiogenesis) and metastasis.

Correspondence: I Kola
Received 18 March 1998, revised 4 June 1998; accepted 9 June 1998

ETS factors also have important developmental roles. *Pointed P2* and *yan* play critical roles in *Drosophila* eye development (O'Neill *et al.*, 1994). *ETS2* is involved in skeletal/cartilage development (Sumarsono *et al.*, 1996). *PU.1* null mutation results in hematopoietic abnormalities (McKercher *et al.*, 1996), and *ETS1* is involved in transactivation of genes required for T cell function (Muthusamy *et al.*, 1995; Sun *et al.*, 1995; Thomas *et al.*, 1995, 1997) and angiogenesis (Wasylyk *et al.*, 1991; Vandenhunder *et al.*, 1994; Wernert *et al.*, 1992).

The ETS factors are almost all expressed in hematopoietic lineages (Bhat *et al.*, 1989, 1990; Kola *et al.*, 1993), and indeed appear to function predominantly in these cells and their related neoplasms. However, the most common solid tumors in humans are carcinomas which arise from the transformation of epithelial cells. Transformed breast epithelial cells, for example, have been shown to express ETS family members GABPα, PEA3, ELF1, ETS1 and ELK1 (Scott *et al.*, 1994b; Delannoy-Courdent *et al.*, 1996), but expression of these ETS family members is not restricted to epithelial cells. One ETS family member, ELF3/ESX/ESE-1/ERT, has recently emerged with epithelial and epithelial-cancer specific expression (Tymms *et al.*, 1997; Chang *et al.*, 1997; Choi *et al.*, 1998; Oettgen *et al.*, 1997). Given the extensive involvement of ETS factors in tumorigenesis, it becomes important to identify any additional ETS genes that may also play oncogenic roles, especially those that may be involved in epithelially derived cancers.

Here we report the cloning and initial characterization of a novel ETS family member, ELF5, that displays a similar, but more restricted, expression pattern to that of ELF3. In a comparison of ETS domains, ELF5 displays strongest similarity to the ELF/E74 subfamily of ETS factors, and is most closely related to ELF3. Other than a conserved pointed domain, however, ELF5 has little similarity to ETS family members outside of the ETS domain. ELF5 functions as a transcription factor with similar sequence-specific DNA binding characteristics to other ETS family members.

Results

Isolation of mouse and human ELF5 cDNAs

The murine *Elf5* cDNA was isolated from an adult mouse lung cDNA library following screening with a cDNA probe containing the ETS domain of human *ELF3*. Amalgamation of sequence data revealed a 1437 bp sequence with a maximum open reading frame (ORF) of 759 bp, predicted to encode a 253 amino acid protein of approximately 31 kD (Figure 1a). An upstream, in-frame stop codon suggests that this ORF represents the full-length coding sequence of *ELF5*. No evidence was found for alternative splicing within the ORF. Additional 91 bp of 5′, and 696 bp of 3′ sequences were obtained by reverse transcriptase polymerase chain reaction (PCR) and rapid amplification of cDNA ends (RACE), using day 14 mouse placental RNA. Sequence analysis revealed two discrete polyadenylation signals present in the 3′ untranslated region (UTR). The first of these (at 1391 bp) appears to be an overlapping poly(A)$^+$ recognition signal, AATTAA and ATTAAAA, similar to that identified in the gene for C4b-binding protein by Kristensen *et al.* (1987). The second is a consensus polyadenylation signal, AATAAA, at 2181 bp. These polyadenylation signals are found close to the 3′ termination of the original clone and the 3′ RACE product, respectively, suggesting that these represent polyA signals for two separate mRNA products. Thus, the two predicted *Elf5* cDNAs are 2224 bp and 1528 bp long. Northern blot analysis, using the *Elf5* coding sequence as a probe, confirmed the presence of two predominant *Elf5* transcripts in placental tissue, *Elf5-a* and *Elf5-b*, of approximately 2.5 kb and 1.5 kb respectively. Only *Elf5-a* was identified using a 3′ UTR fragment from between the polyadenylation signals as a probe (Figure 1b), indicating that the transcripts differ in 3′ UTR sequences. Interestingly, the sequence found in *Elf5-a*, but not *Elf5-b*, contains multiple ATTTA motifs that are associated with rapid mRNA turnover (Savant-Bhonsale and Cleveland, 1992; Akashi *et al.*, 1994).

A human *ELF5* cDNA fragment was isolated from a human lung cDNA library following screening with a cDNA probe containing the coding sequence of mouse *Elf5*. The full coding sequence of human *ELF5* was then obtained by reverse transcriptase PCR and RACE using human placental RNA. Unlike the mouse, no evidence was found for the utilization of multiple polyA$^+$ signals in human *ELF5*, which is supported by the observation of a single transcript in Northern analysis of human tissues (see Figure 5a). Analysis revealed that the *ELF5* sequence is predicted to encode a 255 residue amino acid protein.

Comparison of human and mouse ELF5 amino acid sequences

The predicted amino acid sequences of human and mouse ELF5 are highly conserved, with approximately 95% identity (Figure 2a). Only a single amino acid substitution was observed within the putative ETS domain of human and mouse ELF5, and most of the other differing amino acid residues in the full-length sequences are conservative substitutions (8/13), suggesting that the two proteins are homologs (i.e. having an inferred common ancestry). Interestingly, human ELF5 does, however, contain an additional two amino acid insertion compared to mouse Elf5. In addition to the ETS domain, other features appear to be conserved between these two sequences. These include a putative 'pointed' domain (Seth *et al.*, 1992; Lautenberger *et al.*, 1992) and several consensus casein kinase II (CKII) (Pinna, 1990), protein kinase C (PKC) (Kishimoto *et al.*, 1985; Woodget *et al.*, 1986) and tyrosine kinase (Patschinsky *et al.*, 1982; Hunter, 1982; Cooper *et al.*, 1984) phosphorylation sites. The ETS and pointed domains are discussed below, with reference to other ETS family members, but the significance of consensus phosphorylation sites in ELF5 have not yet been determined.

The ETS domain found within all members of the ETS family is responsible for sequence-specific DNA binding (Seth *et al.*, 1992; Lautenberger *et al.*, 1992; Wasylyk *et al.*, 1993). The putative ETS domain of

human/mouse ELF5, situated at the carboxyl terminal of the protein, is highly similar to that of human/mouse ELF3, with amino acid identity being 67%. However, this domain is only moderately similar to that of other ETS family members, with the highest amino acid identity being 49% to human NERF, 48% to *Drosophila* ETS4 and E74A, and 46% to human ELF1 and ELK1 (Figure 2b). Sequence identity to other family members is in the range of 44–36%. However, amino acids highly conserved amongst ETS family members (Janknecht and Nordheim, 1993) are well conserved in ELF5 (23/38). Some of these highly conserved residues, such as the three tryptophan residues in the carboxyl half of the ETS domain, have been demonstrated to be structurally critical for DNA binding of other ETS family members (Wang *et al*., 1992; Wasylyk *et al*., 1992).

Based on ETS domain similarities, a recent phylogenetic analysis (Graves and Petersen, 1998) has proposed the grouping of ETS factors into subfamilies, one of which is the ELF (E74-like-factor) subfamily. The ELF subfamily includes *Drosophila* E74A, human ELF1 and NERF. We generated a phylogenetic tree, including ELF5 and recently isolated ELF3, by maximum likelihood analysis of the ETS domain (Figure 2c). It shows that the human and mouse ELF5 sequences group most closely with the human and mouse ELF3 sequences, and that both ELF3 and ELF5 are most closely related to *Drosophila* ETS4, E74A and human ELF1 and NERF within the ETS family. Thus, *Drosophila* ETS4, and human/mouse ELF3 and ELF5 may also fall into the ELF subfamily of ETS factors. It is for this reason that ELF5 is so named. The phylogenetic relationship of the mELF3, dETS4, hELF1, hNERF, dE74A, hTEL, dYAN, hERM, mER81, mPEA3, hELK1, mERP and hSAP1 sequences is identical to that published recently (Tymms *et al*., 1997) whereas the phylogenetic relationship among the remaining 12 sequences (ELF5 excluded) is different. These differences occur at the basal bipartitions in the data, probably due to the use of a more rigorous phylogenetic method than that used previously. The same phylogenetic relationship among the ETS domains in dYAN, hTEL, dE74A, hELF1, hNERF and dETS4 was reported by Graves and Petersen (1998), who obtained a majority-rule consensus of 1000 neighbor-joining trees generated by bootstrap analysis.

a

```
    atcacacggctacaggtgcctttatttctacagtccgctggtgctgggagcgcgct      56
tgccttctcttgccttgaaagccttctgtctggac                             91
                                  ctagccaccacttgtcttcacggtg     116

ATGTTGGACTCCGTAACCCATAGCACCTTCCTGCCCAACGCATCCTTCTGTGACCCCCTG    176

ATGCCTTGGACCGATCTGTTCAGCAATGAAGACTACTACCCTGCCTTTGAGCATCAGACA    236

GCCTGTGATTCCTACTGGACATCAGTGCACCCTGAATACTGGACCAAGCGCCATGTCTC.    296

GAATGGCTCCAATTCTGCTGTGACCAGTACAAC TTGATGCCAACTGCATCTCCTTCTGT    356

CACTTCAACATCAGCGGCCTGCAGCTCTGC GCATGACGCAGGAGGAGTTCATTGAGGCA    416

GCCGGCATCTGTGGGGAGTACCTGTAC TCATTCTCCAGAACATTCGCTCGCAAGGTTAC    476

TCCTTTTTCAATGATGCTGAAGAGACCAAGACTGGCATCAAAGACTATGCTGATTCCAGT    536

TGCTTGAAAACAAGTGGCATCAAGAGTCAAGACTGTCACAGCCGAACAAGCCTCCAAAGT    596

TCTCACCTGTGGGAATTTGTCAGAGACTTGCTGCTGTCCCCTGAAGAGAACTGTGGCATC    656

CTGGAATGGGAAGACAGGGAGCAGGGCATTTTCCGAGTGGTTAAGTCAGAAGCCCTGGCA    716

AAGATGTGGGGACAAAGCAAGAAGAATGACACGATGACCTACC AAGCTGAGCCGAGCC    776

CTGAGATACTACTATAAAACGAGAATTCTGGAGCGGGTTGACCGGAGGTTAGTGTACAAA    836

TTTGGAAAGAACGCGCACGGGTGGCAGGAAGAGAAACTCTGAtggacaccggacaccagg    896

ctcatttgatggatttcttattgttggaaacaatcagatcaaactagacatttgaaagtct    956
cctcttcttcttcttcttccctcttcccttcttcttcttctcccctcttcttcttca      1016
aaacctacaaacacactgataaaatttctgcatgtctcagcttacatttgaattcagttg    1076
ttgtctattggtggcgatgccatcagcctttaagcaatcgtcttcatcccaaggggaggga   1136
agggatggtctttgtggcaacttggtgtgacactgtctccttaatgaagtgtttggagcta   1196
agggagccagtgttacgggtgctgtttcacaagaggaccctgtgcaccattaagacacat    1256
gatcttcccgttccaggggttctgagcggtcgactgaggcagcttgccttgtggttagttt   1316
ttaggaaagggagatgtaagacttccttgctttagatttgaaattatcacagttatattc    1376
catagaagaattttttaattaaaaaaattttagtggctaagccactaaacfgggacctaat   1436
tggatgtagcctaagttactaataagttcttaaccagatcaccatttccaaccacttagc    1496
cacagtcacttgatccacggccaatccttctg                                1528
                                  aacttaacatccttgtagttagtcacct   1556
tgggaattgctacctagattgttacccccttcacctcactggtggctatcatcaggtcta    1616
cagtgacctgatcaacagacatgtgcattaatttctaaatcactgctgtgcctatgattc    1676
aaaccgtcagcgtgttcagtttattgattc tctgcggtcggatttctgttgattctctc    1736
tgaggctaagacattaaacctttaccaagcagagaacgtcctaacaagccacgatagccg    1796
aacacagcatcgatctctctcttttctgatgaatactcaaactttccaacatatttctct    1856
tcacaaaagtaaagacagtgattttatcaatcaacgttcatgggttaaagtctgcact      1916
gacatttccttgtctgccgttgcatgccgttggcatgcaaggtgttaatgacctgcaaca    1976
tggtggagtgccctgaaccctaacttcccagagttgggactgtctagtgaccggcactg     2036
aatagcaatgcaggctgaagcctccagggtttagatttcacctcaaaagtaacttgttt     2096
ttaaaaagaaatgtgaattactgtaaaataatctattttggattcgtgtgtttttcagg     2156
tggatatagtttataaacaatgtggatgaaaaatatttcatgtttaaaaaaaaaaaaa      2216
aaaaaaaa                                                        2224
```

b

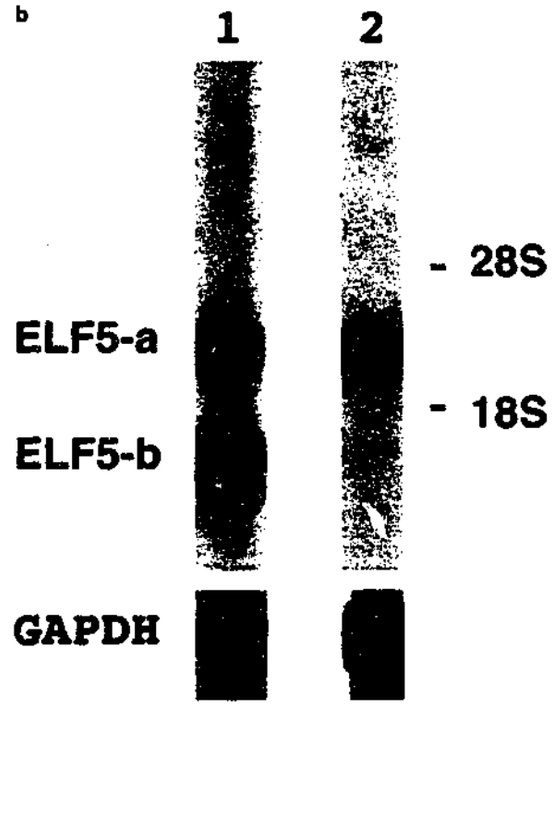
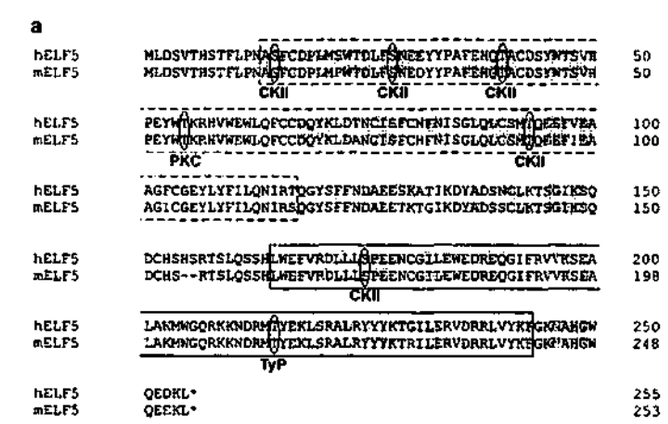


Figure 1 Murine *Elf5* cDNA sequence and relationship to mRNA transcripts. (**a**) The nucleotide sequence of murine *Elf5* is shown. Breaks in the sequence indicate the source of sequence data; the central region (92–1528) was sequenced from lambda clones, and 5 prime and 3 prime were added from sequencing of RACE PCR products. Numbering of the nucleotides, starting with the most 5-prime sequences obtained, are indicated on the right. The open reading frame (ORF) is shown in capital letters, with the initiating start and stop codons underlined. A stop codon, in the same reading frame as the ORF, but 5 prime to the initiating codon, is also underlined. The ETS domain is indicated in a shaded box. Putative polyadenylation signals are underlined with dashed lines. A/T rich tracts in the 3 prime untranslated region are boxed. (**b**) Northern blot analysis of day 14 mouse placenta; lane 1, probed with random-prime-labeled 940 bp *Sty*I murine *Elf5* cDNA fragment (probe 1); lane 2, probed with random-prime-labeled murine *Elf5* 696 bp 3′-RACE PCR product (probe 2). Positions of 28S and 18S markers are indicated. Both lanes were also probed with GAPDH cDNA (lower panels)

The phylogeny in Figure 2c shows the unrooted relationship among 28 ETS domains. Any attempt to infer its root would lead to the conclusion that there has been substantial variation in the substitution rate. Because variation in this rate may affect the chances of recovering the true phylogenetic tree (Hillis *et al.*, 1994), we remain cautious about the phylogeny as a whole. However, if we assume that the root is somewhere near the center of the tree, we may focus on the evolution of ETS domains in dYAN, hTEL, dE74A, hELF1, hNERF, dETS4, hELF3, mELF3, hELF5 and mELF5. Given that these ETS domains occur in three taxa, we can infer at least four gene duplications (A, B, C and D in Figure 2c). The first two gene duplications (A and B) involves both protostome and deuterostome taxa (*Drosophila* and mammals, respectively) and therefore must have occurred before the origin of the deuterostome lineage 550–750 Myr ago (Doolittle *et al.*, 1996). A third gene duplication (C) involves two mammalian lineages (Rodenta and Primata) and therefore must have occurred 115–129 Myr ago (Easteal and Herbert, 1997). A fourth gene duplication (D) involves human sequences only and therefore we can not infer its age. Consequently, we conclude that hELF5 and mELF5 (like hELF3 and mELF3) are orthologous gene products (i.e. their origin can be traced back to a speciation event) and that ELF3 and ELF5 are paralogous gene products (i.e. their origin can be traced back to a gene duplication) (for more details on orthology and paralogy, see Moritz and Hillis, 1996).

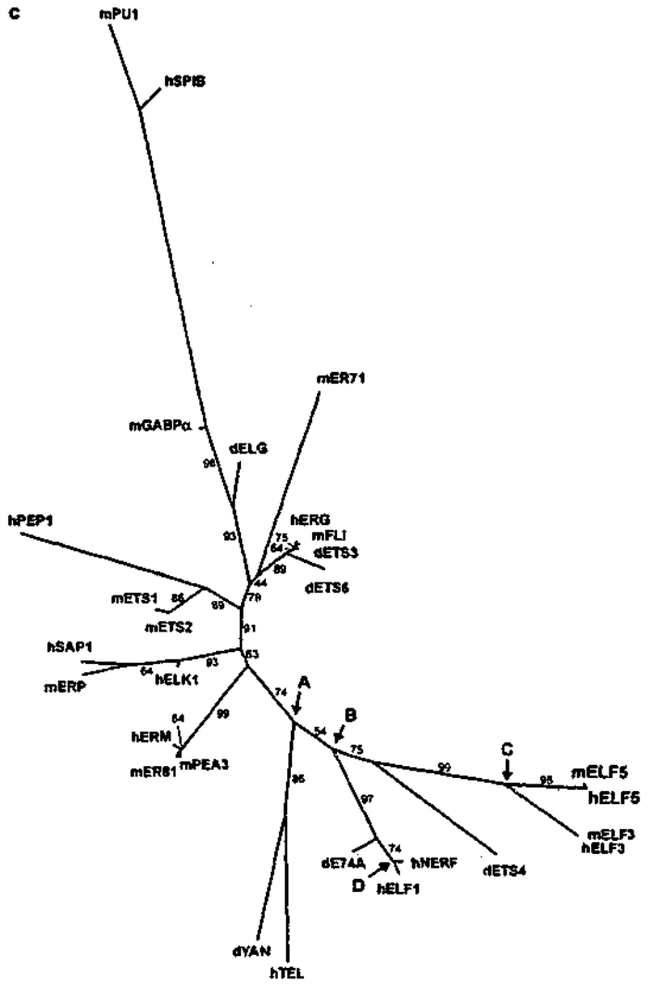
The large degree of sequence divergence among ETS domains (Figure 2b) is paralled by considerable amounts of sequences variation outside of this domain. For instance, similarity between the non-ETS domain ELF5 and ELF3 sequences is very low, with PEST (Tymms *et al.*, 1997), serine rich and high mobility group (HMG) (Chang *et al.*, 1997), and A/T hook (Oettgen *et al.*, 1997) domains reported in ELF3 not apparent in ELF5. However, the pointed domain present in many of the ETS family, such as *pointed*, TEL, *yan*, ETS1 and ELF3, also appears to be present in ELF5. The pointed domain present in many ETS factors displays weak similarity to helix-loop-helix domains, responsible for protein-protein interactions (Littlewood and Evan, 1994). The hELF5 pointed domain displays 26% amino acid residue identity to that of ERG, and 23% to 19% amongst a range of other ETS family members, including hELF3, hTEL, hGABPα, hETS1, hETS2, dYAN and dPOINTEDP2



b

		IDENTITY (%)
hELF5	LWEFVRDLLLSPE-ENCGILEWEDREQGIFRVV--KSEALAKMWGQRK-KNDRMTYEKLSRALRYYYKTGILERVD--RRLVYKF	100
mELF5	LWEFVRDLLLSPE-ENCGILEWEDREQGIFRVV--KSEALAKMWGQRK-KNDRMTYEKLSRALRYYYKTRILERVD--RRLVYKF	98
hELF3	LWEFIRDILIHPE-LNEGLMKWLNRHEGVFKFL--RSEAVAQLWGQKK-KNSNMTYEKLSRAMRYYYKREILERVD-GRRLVYKF	67
mELF3	LWEFIRDILIHPE-LNEGLMKWENRHEGVFKFL--RSEAVAQLWGQKK-KNSNMTYEKLSRAMRYYYKREILERVD-GRRLVYKF	67
hNERF	LWEFLLDLLQDKN-TCPRYIKWTQREKGIFKLV--DSKAVSKLWGKHK-NKPDMNYETMGRALRYYYQRGILAKVE-GQRLVYQF	49
dETS4	LWQFLKELLASPQ-VNGTAIRWIDRSKGIFKIE--DSVRVAKLWGRRK-NRPAMNYDKLSRSIRQYYKKGIMKKTERSQRLVYQF	48
dE74A	LWEFLLKLLQDRE-YCPRFIKWTNREKGVFKLV--DSKAVSRLWGMHK-NKPDMNYETMGRALRYYYQRGILAKVD-GQRLVYHF	48
hELF1	LWEFLLALLQDKA-TCPKYIKWTQREKGIFKLV--DSKAVSRLWGKHK-NKPDMNYETMGRALRYYYQRGILAKVE-GQRLVYQF	46
hELK1	LWQFLLQLLREQ--GNGHIISWTSRDGGEFKLV--DAEEVARLWGLRK-NKTNMNYDKLSRALRYYYDKNIIRKVS-GQKFVYKF	46
hTEL	LWDYVYQLLSDS--RYENFIRWEDKESKIFRIV--DPNGLARLWGNHK-NRTNMTYEKMSRALRHYYKLNIIRKEP-GQRLLFRF	44
hERM	LWQFLVTLLDDP--ANAHFIAWTGR-GMEFKLI--EPEEVARRWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYKF	44
mER81	LWQFLVALLDDP--SNSHFIAWTGR-GMEFKLI--EPEEVARRWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYKF	44
mPEA3	LWQFLVALLDDP--TNAHFIAWTGR-GMEFKLI--EPEEVARLWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYKF	44
mGABPα	LWQFLLELLTDK--DARDCISWVGDEG-EFKLN--QPELVAQKWGQRK-NKPTMNYEKLSRALRYYYDGDMICKVQ-GKRFVYKF	44
mERP	LWQFLLHLLLDQ--KHEHLICWTSNDG-EFKLL--KAEEVAKLWGLRK-NKTNMNYDKLSRALRYYYDKNIIKKVI-GQKFVYKF	42
dETS6	LWQFLLELLADS--SNANAISWEGQSG-EFRLI--DPDEVARRWGERK-AKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	42
mPU1	LYQFLLDLLRSG--DMKDSIWWVDKDKGTFQFSSKHKEALAHRWGIQKQNP ITYQKMARALRNYGKTGEVKKVK--KKLTYQF	42
hPE1	LWHFILELLQKE--EFRHVIAWQQGEYGEFVIK--DPDEVARLWGRRK APQMNYDKLSRALRYYYNKRILHKTK-GKRFTYKF	42
hSAP1	LWQFLLQLLQKP--QNKHMICWTSNDG-QFKLL--QAEEVARLWGIRK-NKPNMNYDKLSRALRYYYVKNIIKKVN-GQKFVYKF	42
hSPIB	LYQFLLGLLTRG--DMRECVWWVEPGAGVFQFSSKHKELLARRWGQQKGNRKRMTYQKLARALRNYAKTGEIRKVK--RKLTYQF	42
dYAN	LWDFLQQLLNDRNQKYSDLIAWKCRDTGVFKIV--DPAGLAKLWGIQK-NHLSMNYDKMSRALRYYYRVNILRKVQ-GERHCYQF	41
hERG	LWQFLLELLSDS--SNSSCITWEGTNG-EFKMT--DPDEVARRWGERK-SKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	41
mFLI1	LWQFLLELLSDS--ANASCITWEGTNG-EFKMT--DPDEVARRWGERK-SKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	41
dELG	LWQFLLEILTDC--EHTDVIEWVGTEG-EFKLT--DPDRVARLWGEKK-NKPAMNYEKLSRALRYYYDGDMISKVS-GKRFAYKF	40
dETS3	LWQFLLELLSDS--NNASCITWEGTNG-EFKLT--DPDEVARRWGERK-SKPNMNYDKLSRALR----------	39
mETS1	LWQFLLELLTDK--SCQSFISWTGDGW-EFKLS--DPDEVARRWGKRK-NKPKMNYEKLSRGLRYYYDKNIIHKTA-GKRYVYRF	37
mETS2	LWQFLLELLSDK--SCQSTISWTGDGW-EFKLA--DPDEVARRWGKRK-NKPKMNYEKLSRGLRYYYDKNIIHKTS-GKRYVYRF	37
mER71	LWQFLLKLLQDG--ARSSCIRWTGNSR-EFQLC--DPKEVARLWGERK-RKPGMNYEKLSRGLRYYYRRDIVLKSG-GRKYTYRF	36
consensus	LWQFLL LL D I W FK VAR WG K P MNY KLSR LRYYY I K GR Y F	

(Figure 2d). Although this similarity is relatively weak, of those amino acid residues conserved between ETS family members, 28/63 are also present in ELF5. A mitogen activated protein (MAP) kinase site, present at the extreme amino end of the pointed domain of ETS1, ETS2 and *pointedP2*, and involved in enhancing transactivation (Yang *et al.*, 1996), does not appear to be present in ELF5.



d

```
                                                                                                  Identity
                                                                                                    (%)
hELF5       A-SFCDPL-MSWTDLFSNEEYYPAFEHQTACDSYWTSVHPEYWTKRHVWEWLQFCCDQYKLDT-NCIS-FCNFNISGLQLCSMTQEEFVEAAG-FCGEYLYFILQNIRT   100
mELF5       A-SFCDPL-MFWTDLFSNEDYYPAFEHQTACDSYWTSVHPEYWTKRHVWEKLQFCCDQYKLDA-NCIS-FCHFNISGLQLCSMTQEEFIEAAG-ICGEYLYFILQNIRS    93
hERG        E-SNPMNYN-SYMDEK-NGPPPPNMTTNERRVIIVPA-DPTLWSTDHVRQWLEWAVKEYGLPDVN-ILLFQ--NIDGKELCKMTKDDFQRLTPSYNADILLSHLHYLRE    26
hELF3       A-SVP-PAATFGADDLVLTLSNPQMSLEGTEKASWLGEQPQFWSKTQVLDWISYQVEKNKYDA-SAID-FSRCDMDGATLCNCALEELRLVFG-PLGDQLHAQLRDLTS    23
hTEL        PESPV-PSYASSTPLHVPVPRALRMEEDSIRLPAHLRLQPIYWSRDDVAQWLKWAENEFSLRG-IDSNTFE-MN--GKALLLLTKEDPRYRSP-HSGDVLYELLGHIKQ    22
hGABPα      ITTISDETSEQVTRWAAALEGYRK-EQERLGIPY----DPIQMSTDQVLHWVVWVMKEFSMTDIDLTTL----NISGRELCANTHEEFNQKLPRDPGNIFWTHLQLLKE    22
hETS1       PLLTPSSKEMMSQALKATFSGFTK-EQQRLGIPI----DPRQWTETKVRDWVMWAVNEFSLKGVDFQK-FC-MN--GAALCALGKDCFLELAPDFVGDILWEHLEILQK    22
hETS2       PLLTPSSKAVMSQALKATFSGFKK-EQRRLGIPK----NPWLWSEQQVCQWLLWATNEFSLVNVNLQR-FG-MN--GQMLCNLGKERFLELAPDFVGDILWEHLEQMIK    21
dYAN        LNSLN-PGIWSDVLWRCPPAPSSQLAELKTQLPPSLPSDPRLWSREDVLVFLRFCVREFDLPK-LDFDLFQ-MN--GKALCLLTRADFGHRCP-GAGDVLHNVLQMLII    20
dPOINTEDP2  PPLTPGTNRKVNEVLKASFASWEK-EVQKCNITK----DPREWTEEHVIYWLNWAKNEFSLVSMNLDP-FYKMK--GKAMVDLGKEKFLAITPPFTGDILWEHLDILQK    19
consensus   P S      MS       F  FKK E QRL IP     DP  WS  HV  WL WAVKEFSL   NLI  F  MN  GKELC L KEDFLER P F GDILWEHLE LRK
```

Figure 2 (**a**) Comparison of human and mouse ORFs. Amino acid sequences present in both human and mouse ELF5 are shaded. The ETS domain is boxed with a solid line and the pointed domain with a dashed line. Putative phosphorylation sites, conserved between the two species are circled and labeled as CKII (casein kinase II), PKC (protein kinase C) or TyP (tyrosine kinase) substrates. (**b**) Comparison of the ETS domain of human and mouse ELF5 with those of known members of the ETS gene family. The alignment was generated using CLUSTAL W (Thompson *et al.*, 1994) with the default settings, and the result was subsequently adjusted manually. The ETS factors examined are labeled on the left and include hELF3, mELF3, hNERF, dETS4, dE74A, hELF1, hELK1, hTEL, hERM, mER81, mPEA3, mGABPα, mERP, dETS6, mPU1, hPE1, hSAP1, hSPIB, dYAN, hERG, mFLI1, dELG, dETS3, mETS1, mETS2, mER71, where 'h' denotes human, 'm' mouse and 'd' *Drosophila*. The ETS consensus sequence is a list of the amino acids most often conserved between ETS family members. Shading denotes amino acid identity with human ELF5, and the percent identity of each ETS domain is indicated on the right. (**c**) Phylogenetic tree of the ETS domain produced by maximum likelihood analysis. The alignment in Figure 2b was analysed using the JTT-F substitution model (Jones *et al.*, 1992) and local bootstrap values were estimated for all internal branches, both by using PROTML in Q mode followed by a second run in R mode (Adachi and Hasegawa, 1996). An underlying assumption of the phylogenetic analysis is that the amino acid content does not vary significantly among the sequences. This assumption was not assessed because tools for doing so are still under development (LSJ, unpublished work). Therefore, the tree may be the result of both historical and compositional components. The four points at which gene duplications have been inferred are marked A, B, C and D. (**d**) Comparison of the pointed domain of human and mouse ELF5 with those of other members of the ETS family. The ETS factors examined are labeled on the left and include hERG, hELF3, hTEL, hGABPα, hETS1, hETS2, dYAN and dPOINTEDP2. Other labels and conventions are as described for Figure 2b

Human chromosomal mapping of ELF5

We performed human chromosomal localization of *ELF5* by PCR, using gene specific primers and the Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre). With these primers, a single product of the expected size (234 bp) was amplified from total human DNA. The PCR reactions were then performed separately for each of the individual hybrids. The amplification results (data not shown) from the 93 hybrids were submitted to the Radiation Hybrid Mapping server at Whitehead Institute/MIT Center for Genome Research for analysis. The result provided by Radiation Hybrid Mapping server demonstrated that ELF5 is localized to chromosome 11. The markers most tightly linked to ELF5 are D11S3990 (6.5cR) and D11S3998 (15.9cR) (lod score > 3.0), and these markers are located in the region of 11p13–15 (Figure 3). This chromosomal region frequently undergoes loss of heterozygosity (LOH) in several types of carcinoma (Baffa *et al*., 1996; Dahiya *et al*., 1997; Hirose *et al*., 1996; Iizuka *et al*., 1995; Kawana *et al*., 1997; Lichy *et al*., 1998; Wilson *et al*., 1996).

Expression pattern of Elf5 in mouse tissues

Poly(A)$^+$ mRNA material derived from various mouse tissues were analysed by Northern blot hybridization using the murine *Elf5* cDNA as a probe. A GAPDH probe was then used to control for RNA loading.

Analysis of *Elf5* expression in adult mouse tissues revealed that *Elf5*, like mouse *ELF3* (Tymms *et al*., 1997), has a restricted expression pattern. Expression of two *Elf5* transcripts, *Elf5-a* (2.5 kb) and *Elf5-b* (1.5 kb), were observed in lung (Lu), kidney (Ki), stomach (St), ovary (Ov), tongue (To), bladder (Bl), and day 2 pregnant (2 Ma) and day 10 pregnant (10 Ma) mammary glands, but no expression was observed in liver (Li), heart (He), small intestine (Sm), spleen (Sp), thymus (Th), pancreas (Pa), skeletal muscle (Sk), colon (Co) or fat (2 Fa and 10 Fa) (Figure 4a). Fat from day 2 (2 Fa) and day 10 (10 Fa) pregnant mice was used as a control for mammary expression, since the mammary gland contains much fat tissue. A single transcript was observed in brain (arrow – approximately 2.1 kb), but of a different size to either of the two *Elf5* transcripts in other organs. However, its appearance was variable in repeated experiments with adult tissues (data not shown) and at different developmental stages (Figures 2b and c). Pending further analysis, it appears likely that this brain specific transcript is derived from cross hybridization to another highly expressed ETS family member, rather than *Elf5*.

Some ETS members are implicated in developmental processes. Therefore it is of interest to characterize the expression pattern of *Elf5* during mouse development. We examined the expression of *Elf5* in the neonatal mouse (Figure 4b) and during embryogenesis on days 19, 17 and 16 (Figure 4c), and observed a similar expression pattern compared to that of the adult.

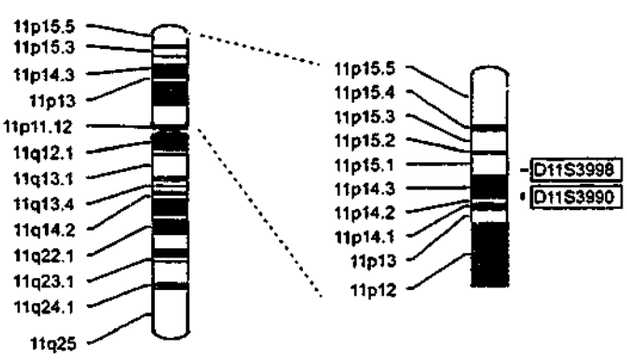


Figure 3 Chromosomal localization of human *ELF5*. Human chromosomal localization of ELF5 was performed by PCR using gene specific primers and the Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre). Diagram based upon PCR results (data not shown) showing localization of *ELF5* within chromosome 11, with respect to adjacent marker obtained from mapping data (see text)

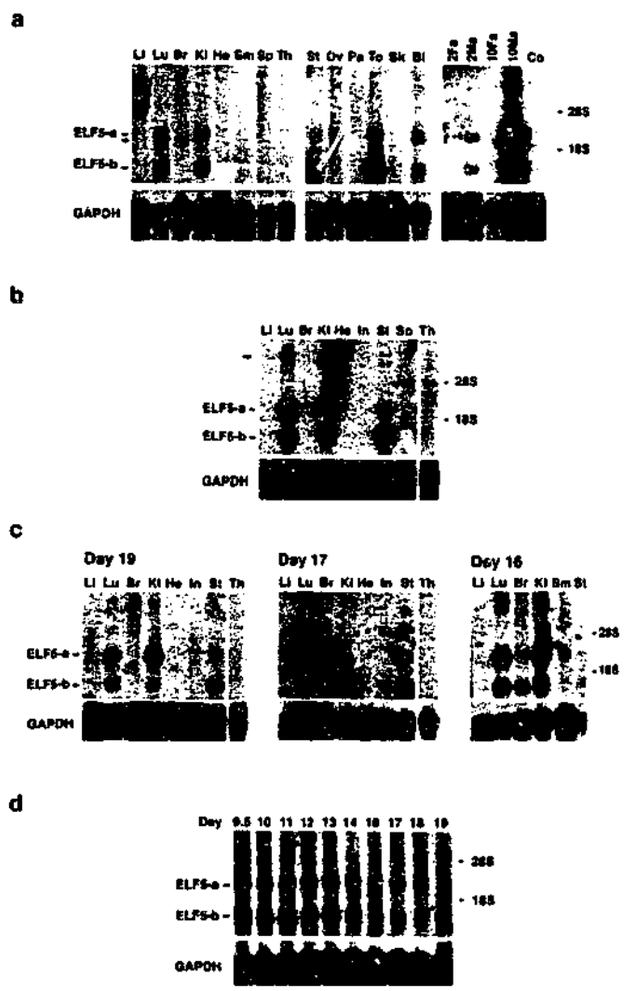


Figure 4 *Elf5* expression in mouse tissues. Positions of 28S and 18S markers are indicated. *Elf5a* and *Elf5b* transcripts are indicated. (**a**) Northern analysis of adult mouse tissues probed with murine *Elf5* cDNA (top panels) and GAPDH cDNA (lower panels). Abbreviations: Li: liver; Lu: lung; Br: brain; Ki: kidney; He: heart; Sm: small intestine; Sp: spleen; Th: thymus; St: stomach; Ov: ovary; Pa: pancreas; To: tongue; Sk: skeletal muscle; Bl: bladder; 2Fa: day 2 pregnant fat; 2 Ma: day 2 pregnant mammary gland; 10 Fa: day 10 pregnant fat; 10 Ma: day 10 pregnant mammary gland; Co: colon. Arrow indicates position of brain specific transcript (see text). (**b**) Northern analysis as above, but using RNA from day 1 neonate mouse tissues. Additional abbreviation: In: intestine. Arrow indicates position of large transcript (see text). (**c**) Northern analysis as above, but using RNA from day 16, 17 and 19 embryonic tissues. (**d**) Northern analysis as above, but using RNA from day 9.5 to day 19 placental tissues as indicated

However, at day 16 stage of embryogenesis low levels of *Elf5* expression were detected in brain (regular sized transcripts) and small intestine, in addition to the expression pattern observed in the adult.

Placental expression of *Elf5* displayed an interesting pattern during stages of embryogenesis (Figure 4d). Both transcripts were increasingly expressed from day 9.5 to day 13 before an overall decrease observed from day 14 to day 19, although some expression was observed at day 17.

The two predominant *Elf5* mRNA transcripts were observed in variable ratios in different tissues, suggesting that polyadenylation sites may be utilized differentially, or the two transcripts are subject to differential degradation. *Elf5-a* was expressed more strongly in neonatal and embryonic lung and kidney (Figures 4b and c), and adult ovary (Figure 4a), compared to *Elf5-b*. Conversely, *Elf5-b* was stronger in adult tongue (Figure 4a), and in all developmental stages of stomach (Figures 4a, b and c), compared to *Elf5-a*. However, since these two transcripts would appear to differ only in 3′ UTR sequences, they are likely to produce the same translation product. In some RNA samples a further large (>10 kb) transcript was variably observed. Despite the fact that the RNA samples were poly(A)$^+$ selected, these bands could represent unspliced *Elf5* transcripts or genomic contamination.

Expression pattern of ELF5 in human tissues and cancer cell lines

Expression of *ELF5* in adult human organs was also analysed by Northern blot of poly(A)$^+$ mRNA probed with the human ELF5 cDNA (Figure 5a). A single transcript of approximately 2.5 kb was strongly expressed in kidney (Ki) and prostate (Pr). However, much longer exposures of blots (data not shown) demonstrated just detectable expression of *ELF5* in placenta (Pl) and lung (Lu). Further, *ELF5* was cloned from human lung and placenta cDNA libraries, confirming that it is expressed in these tissues, albeit probably at very low levels. Although this expression pattern closely resembles that observed in the mouse, it is interesting that the human *ELF5* gene does not appear to utilize alternative polyadenylation sites, resulting in only a single transcript.

Given the restricted expression pattern of *ELF5* in tissues commonly giving rise to carcinomas, it is of interest to examine *ELF5* expression in human cancers. A panel of cancer cell lines, including carcinomas of the ovary (CaOv-3), breast (BT-549, ZR-75-1, T47D), kidney (786-O), liver (SK-HEP-1), lung (A549), amnion (WISH), prostate (DU145, PC3) and endometrium (HEC-1), and melanoma (MEL28), T-cell leukemia (Jurkat) and erythroid leukemia (K562), were analysed for *ELF5* expression by RNAse protection assay (Figure 5b). A primary fibroblast cell line (CCL32SK) was also included as a sample of non-transformed cells. Of all these cell lines only T47D, a progesterone sensitive ductal breast carcinoma, was observed to express *ELF5*.

To evaluate the possibility that lack of *ELF5* expression in carcinoma was due to genomic alterations, a panel of breast and lung carcinoma derived cell lines were analysed by Southern blot (Figure 5c). *ELF5* gene dosage was compared to that present in DNA from normal human blood (based on the 6.5 kb *Bgl*II fragment) and controlled by hybridization with a *β-actin* cDNA probe. These results are summarized in the

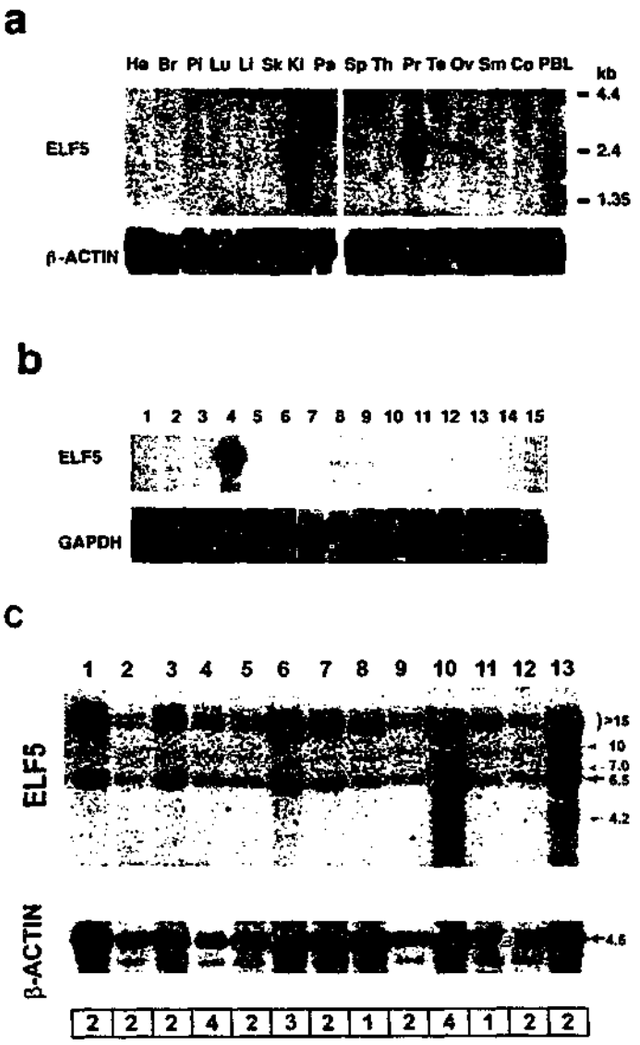


Figure 5 *ELF5* expression in human tissues and cell lines. (**a**) Northern analysis of adult human tissues probed with human *ELF5* cDNA (top panels) and β-Actin cDNA (lower panels). The single *ELF5* transcript is indicated. Other labels and conventions are as for Figure 4. Abbreviations: He: heart; Br: brain; Pl: placenta; Lu: lung; Li: liver; Sk: skeletal muscle; Ki: kidney; Pa: pancreas; Sp: spleen; Th: thymus; Pr: prostate; Te: testis; Ov: ovary; Sm: small intestine; Co: colon mucosa; PBL: peripheral blood lymphocytes. (**b**) RNAse protection analysis of *ELF5* and GAPDH in cell lines: 1: CaOv-3 (ovarian carcinoma); 2: BT-549 (ductal breast carcinoma); 3: ZR-75-1 (breast carcinoma); 4: T47D (ductal breast carcinoma, progesterone sensitive); 5: 786-O (renal adenocarcinoma); 6: SK-HEP-1 (liver adenocarcinoma); 7: A549 (lung adenocarcinoma); 8: CCL32SK (primary fibroblast); 9: MEL28 (melanoma); 10: WISH (amnion carcinoma); 11: Jurkat (T cell leukemia); 12: DU145 (prostate carcinoma); 13: PC3 (prostate carcinoma); 14: HEC-1 (endometrium carcinoma); 15: K562 (erythroid leukemia). (**c**) Southern analysis of *ELF5* in *Bgl*II digested genomic DNA from cell lines: 1: normal blood; 2: BT-549 (ductal breast carcinoma); 3: ZR-75-1 (breast carcinoma); 4: T47D (ductal breast carcinoma); 5: NCI-H1299 (large cell lung carcinoma); 6: NCI-H187 (small cell lung carcinoma); 7: NCI-H322 (bronchioalveolar carcinoma); 8: NCI-H358 (bronchioalveolar carcinoma); 9: NCI-H522 (lung adenocarcinoma); 10: SK-LU-1 (lung adenocarcinoma); 11: NCI-H441 (bronchioalveolar carcinoma); 12: NCI-H460 (large cell lung carcinoma); 13: NCI-H661 (large cell lung carcinoma)

lower panel, where '2' represents a normal allele complement. No evidence was found for allelic loss or gene rearrangement in the two breast carcinoma cell lines that did not express *ELF5* (BT-549 – lane 2, ZR-75-1 – lane 3). However, of nine lung carcinoma cell lines, evidence for loss of an *ELF5* allele was observed in two (NCI-H358 – lane 8, NCI-H441 – lane 11). Hybridization with an ELF3 cDNA probe, which is localized to the long arm of chromosome 1 (Tymms *et al*, 1997), helped to confirm the specific loss of *ELF5* alleles (data not shown). Two other lung carcinoma lines (SK-LU-1 – lane 10, NCI-H661 – lane 13) displayed hybridization with multiple fragments (shaded arrows) in addition to those observed in normal DNA (solid arrows), possibly indicating that at least one *ELF5* allele has been rearranged in these lines. Confirmation of rearrangement, rather than restriction fragment length polymorphism (RFLP), was made by additional restriction digests (data not shown). Some cell lines appeared to have amplification or additional copies of the *ELF5* gene. One of these, T47D (lane 4), was the only cell line demonstrated to express ELF5, and another, SK-LU-1 (lane 10), appeared to have rearranged alleles.

Sequence-specific binding of Elf5 to DNA sequences containing consensus ETS sites

Although ELF5 displays similarity to the consensus ETS domain, characterizing it as an ETS family member, this sequence is still quite divergent from most other ETS family members. The hallmark of ETS factors to bind DNA sites containing a GGAA-core in a sequence-specific manner is however shared by ELF5, demonstrating an additional functional similarity to the ETS family. A recombinant Elf5 HIS-tag protein of approximately 29 kD, expressed in *E. coli* and purified by metal-affinity chromatography (Figure 6a, lane 4), displayed strong binding to consensus ETS binding sites, as analysed by electrophoretic mobility shift assay (EMSA) (Figure 6b). Elf5 bound the E74 oligonucleotide (containing a GGAA-core) (lane 1), but not to the E74m1 oligonucleotide (which had been mutated to an AGAA-core) (lane 2). The first G-residue of the core has been demonstrated to be a physical point of DNA contact for ETS1, and consequently essential for DNA binding (Fisher *et al*., 1991; Nye *et al*., 1992). Thus, Elf5 displays sequence specific binding to a consensus ETS binding site, binding that is disrupted by a mutation known to similarly affect other ETS family members. These results were confirmed through competition analysis. The Elf5-E74 complex (lane 3) was efficiently competed by the addition of a 100-fold excess of unlabeled E74 (lane 4), but not by E74m1 (lane 5).

Elf5 also displayed sequence specific binding to different consensus ETS binding sequences, and did so with differential affinity (Figure 6b). Competition of the Elf5-E74 complex (lane 3) was achieved by consensus ETS sites from the GM-CSF promoter (lane 6), *erb*-B2 promoter (lane 7) and moloney sarcoma virus (MSV) long terminal repeat (LTR) (lane 8). The relative ability of Elf5 to bind these sequences occurred in the order: E74> erbB2>MSV>GM-CSF. Elf5 did not appear to be

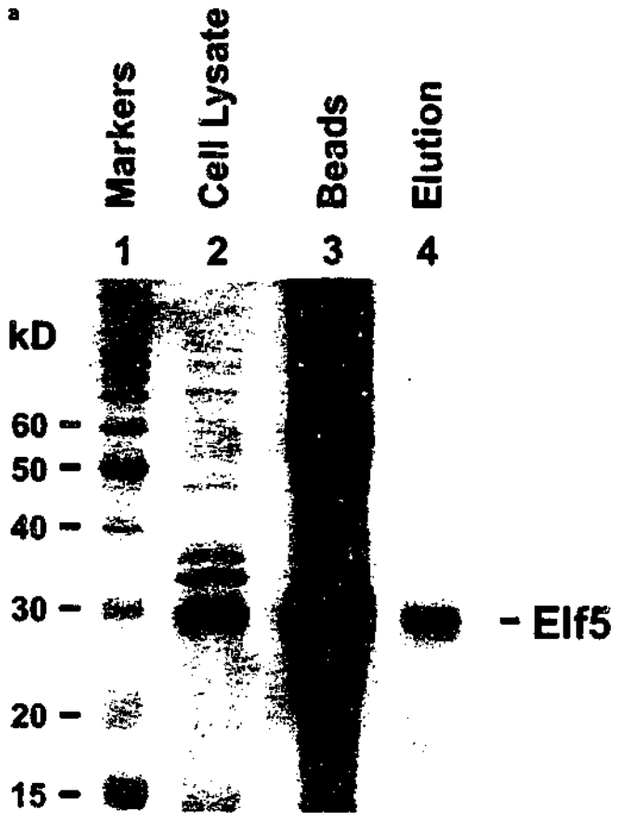


Figure 6 Elf5 binds to consensus ETS binding sequences. (a) His-tagged Elf5 recombinant protein, present in *E. coli* lysates (lane 2), was purified by metal-affinity chromatography to approximately 90% (lane 3) and eluted with imadazole (lane 4). (b) Specific DNA binding of Elf5 was analysed by electrophoretic mobility shift assay (EMSA), using labeled double-stranded oligonucleotides as probes. E74 contains a consensus binding site for ETS family members (lane 1). E74ml is a mutant oligonucleotide based on E74, but with the core GGAA replaced by AGAA (lane 2). Binding to other consensus ETS sites was analysed by the ability of a 100-fold excess of unlabeled double-stranded oligonucleotide to compete with E74 for Elf5 binding. GMETS contains an ETS binding site from the human GM-CSF promoter (lane 6). ERBB2 contains an ETS binding site from the human *erbB2*/HER2 promoter (lane 7). MSV contains an ETS binding site present in the long terminal repeat of the Moloney sarcoma virus (lane 8). API contains a consensus API binding site used as a negative control. Elf5-DNA complexes are marked. Binding of ETS1 to E74 was used as a positive control (lane 10)

competed at all by an oligonucleotide containing a consensus AP1 binding site (lane 9). ETS1 binding to E74 was used as a positive control (lane 10).

Mouse ELF5 acts as a transcriptional activator

In addition to DNA binding, another characteristic of most ETS factors is their ability to transactivate from binding sites in promoters and enhancers. An exception is *ERF*, which displays strong repressor-like activity (Sgouras *et al*., 1995). It is therefore of interest to evaluate the ability of Elf5 to function as a transcriptional activator or repressor. This is especially pertinent to Elf5 given its small size (29 kD), since absence of a transactivation domain could potentially result in competitive inhibition of transactivation by other ETS family members.

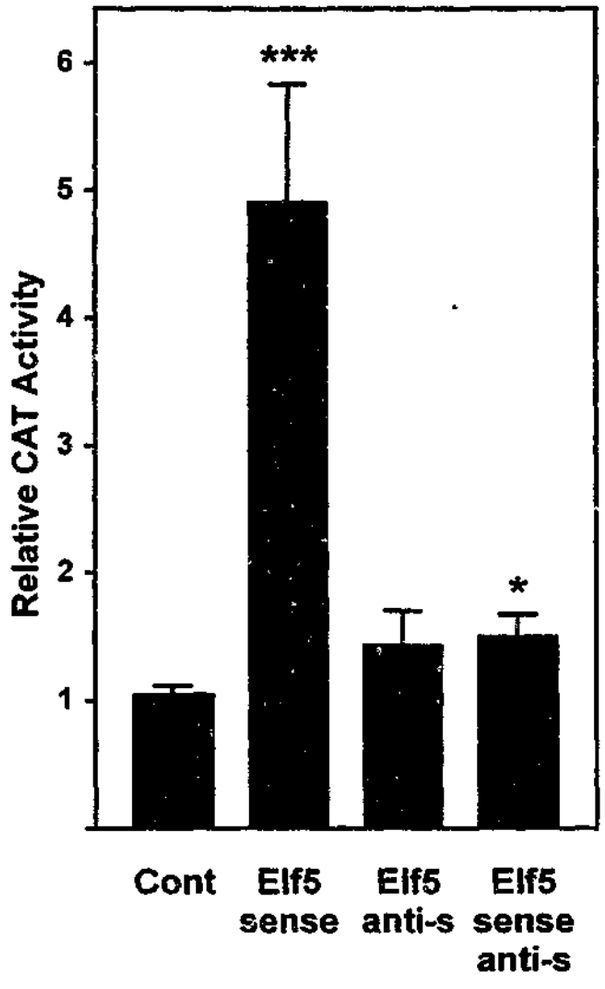


Figure 7 Transactivation by Elf5. COS cells were co-transfected with CAT reporter and Elf5 expression constructs. Transcription of the CAT gene was driven by the thymidine kinase (tk) minimal promoter with five copies of the polyomavirus enhancer inserted upstream (p5Xpoly). The polyomavirus enhancer contains adjacent ETS and AP1 binding sites. The Elf5 sense construct (pBOSElf5s) was designed to express Elf5 protein, and the *Elf5* anti-s construct (pBOSElf5as) to produce anti-sense transcripts. In the absence of expression construct the equivalent amount of base vector (pEFBOS) was co-transfected. COS cells were processed for CAT assays and the results of at least four replicates are shown as the mean with standard error of the mean (s.e.m.) bars. Statistically significant results are indicated by asterisks. A single asterisk indicates moderate significance ($0.05 > P > 0.01$) and triple asterisks indicate very high significance ($P < 0.001$)

A reporter construct, containing the chloramphenicol acetyl-transferase (CAT) driven by a minimal TK promoter and multiple ETS/AP1 binding sites (from the polyomavirus enhancer), was co-transfected into COS cells together with an Elf5 expression construct (Figure 7). Analysis of CAT activities revealed that Elf5 expression resulted in an average fivefold transactivation of the reporter. Further, this transactivation was inhibited by addition of an anti-sense *Elf5* mRNA expression vector, indicating that Elf5 transactivation was due specifically to the product translated from the sense construct.

Discussion

The ETS transcription factors comprise a burgeoning family of proteins with similar sequence-specific DNA binding properties. Much of the ETS research has focused upon functions in hematopoietic cells, but very little is known about the functions of ETS factors in other cell types. We have isolated a novel gene, *ELF5*, that belongs to the ETS transcription factor family. Unlike most other ETS family members (Bhat *et al*., 1989, 1990; Kola *et al*., 1993), *ELF5* is not expressed in hematopoietic compartments, but is restricted to organs such as lung, stomach, kidney, prostate, bladder and mammary gland.

Two predominant *Elf5* mRNA species in the mouse (*Elf5-a* and *Elf5-b*) appear to arise from utilization of two distinct polyadenylation sites. The larger of these two transcripts, *Elf5-a*, incorporates multiple A/T rich signals that have been implicated in processes of increased mRNA turnover (Savant-Bhonsale and Cleveland, 1992; Akashi *et al*., 1994). Perhaps these signals are responsible for the variability of *Elf5-a*:*Elf5-b* ratios in different tissues, particularly the relatively lower proportion of *Elf5-a* observed in stomach. Only a single *ELF5* transcript was observed in human tissues, of approximately the same size as *Elf5-a* in mouse. We are currently sequencing the full-length human *ELF5* cDNA to investigate whether one of the polyadenylation signals, present in mouse, has been lost. Differences in the polyadenylation sites present in mouse and human genes has been reported for *ETS2* (Watson *et al*, 1990), but it is difficult to speculate why such species differences might exist. The cDNA ORFs predict a mouse Elf5 protein of 253 amino acids and a human ELF5 protein of 255 amino acids. These proteins are highly homologous, with a conserved ETS domain, pointed domain and multiple putative phosphorylation sites.

The ETS domain of this new ETS family member displays highest similarity to ELF3 (67%) and moderate similarity to other ETS family members. Phylogenetic analysis indicates that both ELF3 and ELF5 display closest similarity to members of the ELF sub-family (ELF1, NERF, E74), amongst ETS factors. Therefore, we propose that ELF5 and ELF3 (and *Drosophila* ETS4) form part of an extended ELF subfamily within the ETS family of transcription factors.

Expression of *Elf5* in the adult mouse appears to be restricted to lung, kidney, stomach, ovary, tongue, bladder and mammary gland. In the developmental stages examined, this expression pattern remains

basically unchanged from day 16 of embryogenesis to adult. *Elf5* was also found to be expressed in the placenta during embryogenesis. Expression of *ELF5* in human tissues appears to be even more restricted than that for the mouse, with expression observed only in kidney and prostate (which was not examined in the mouse). Inconsistencies between the mouse and human data involve the relatively low *ELF5* expression observed in human lung, ovary and placenta compared to that of mouse. We have observed that *Elf5* expression in the mouse placenta varies with developmental time points, therefore perhaps explaining the variability in levels observed with the human placental data. Similarly, we can not discount the possibility that *ELF5* expression varies with age, or even perhaps within different parts, of the lung and ovary, not present in our human analysis. We are currently investigating the cellular distribution of *ELF5* by *in situ* analysis. Expression of ELF5 was not examined in human stomach, tongue, bladder or mammary gland.

Most ETS genes appear to be functional, or are at least expressed, during hematopoiesis (Bhat *et al*., 1989, 1990; Kola *et al*., 1993). However, *ELF3* is an example of an epithelial-specific ETS gene (Tymms *et al*., 1997; Oettgen *et al*., 1997; Chang *et al*., 1997). *ELF5*, similarly to *ELF3*, is not expressed in hematopoietic compartments, such as the thymus and spleen, but is expressed in organs such as lung, kidney, stomach and prostate. Several differences between *ELF3* and *ELF5* expression were however observed. *ELF3* is expressed strongly in small intestine, colon and liver, whereas *ELF5* does not appear to be expressed in these organs at all. Conversely, *ELF5* appears to be expressed, albeit weakly, in ovary, whereas *ELF3* is not. Even so, *ELF5* displays a strikingly similar expression pattern to that of its closest relative, *ELF3*, which is perhaps related to a gene duplication event. *ELF3* has been demonstrated to be expressed in an epithelial-specific manner, both in normal and transformed breast epithelium (Chang *et al*., 1997) and in lung carcinoma cell lines (Tymms *et al*., 1997). We have examined *ELF5* expression in a panel of human carcinomas and found expression in only one of 11, T47D a ductal carcinoma of the breast. Even given the limited expression of *ELF5* amongst the epithelial-derived carcinomas examined, it appears likely that ELF5, like ELF3, is also an epithelial-specific transcription factor. It is interesting to note that most cell lines derived from lung, renal, prostate, ovarian and breast cancers lacked *ELF5* expression, even though the organs themselves express this gene. Analysis of genomic DNA from carcinoma cell lines was not able to clarify why *ELF5* expression was not observed in two of the three breast carcinoma lines, but four of nine lung carcinoma lines displayed evidence for either loss of an allele or gene rearrangement. Interestingly the rearrangements observed appear to be identical in both independently derived cell lines, suggesting a common loss of function. All lines appeared to contain at least one normal *ELF5* allele, but sequence analysis may reveal smaller mutations that adversely influence expression of *ELF5* (the possibility of which is supported by loss of an allele and gene rearrangement found in some cases).

We have shown that recombinant mouse Elf5 protein is capable of binding to several ETS consensus oligonucleotides, containing a GGAA core motif, in a sequence-specific manner. In addition, Elf5 also displayed the ability to transactivate a promoter containing consensus ETS binding sites. Although these are characteristic of many other ETS factors, it is interesting that ELF5 is capable of DNA binding and transactivation even though it is only approximately half the size of most other ETS family members. The presence of a pointed domain and multiple potential phosphorylation sites also suggest that the economy of ELF5 does not sacrifice regulatory regions, although we are yet to demonstrate that these are functional.

Interestingly, binding of Elf5 to the E74 oligonucleotide was efficiently competed by an oligonucleotide containing an ETS binding site from the *erbB2* promoter. This site has previously been shown to bind and be transactivated by recombinant PEA3 (Benz *et al*., 1997) and ELF3 (ESX) (Chang *et al*., 1997), but ELF5 would appear to be another ETS family member that could contribute to *erbB2* expression. *erbB2* is expressed in the epithelial cells of organs such as the breast, intestine, kidney and ovary (Gullick *et al*., 1987; Kokai *et al*., 1987; Mori *et al*., 1989; Press *et al*., 1990), and its overexpression is thought to play a role in human carcinoma progression (Singleton and Strickler, 1992; Tripathy and Benz, 1994). The expression pattern of *ELF5* matches closely that of *erbB2*, with the exception that *ELF5* does not appear to be expressed in the intestine. ETS factors have also been implicated in the regulation of other genes important to epithelial cell function and/or transformation, including c-*met* (Gambarotta *et al*., 1996), transforming growth factor-β type II receptor (TGF-β RII) (Choi *et al*., 1998), transglutaminase-3 (Lee *et al*., 1996), SPRR2A (Fischer *et al*., 1996) and whey acidic protein (WAP) (Welte *et al*., 1994), but identification of the specific ETS factors involved has not yet been made.

Finally, we have localized *ELF5* to human chromosome 11p13–15, a region which undergoes loss of heterozygosity (LOH) in many types of cancer, such as ductal breast carcinoma (Lichy *et al*., 1998), lung carcinoma (Iizuka *et al*., 1995), rhabdoid tumor of the kidney (Hirose *et al*., 1996), prostate carcinoma (Kawana *et al*., 1997; Dahiya *et al*., 1997), gastric carcinoma (Baffa *et al*., 1996), and ovarian carcinoma (Wilson *et al*., 1996). This region is believed to harbor several tumor suppressor genes (Baffa *et al*., 1996; Coleman *et al*., 1997; Feinberg, 1996; Gao *et al*., 1997; Ichikawa *et al*., 1996; Iizuka *et al*., 1995; Zenklusen *et al*., 1995), based upon both LOH data and the ability to inhibit the tumorigenicity in chemically induced murine squamous cell carcinomas upon introduction of human chromosome 11 (Zenklusen *et al*., 1995). It is interesting to note that other ETS genes implicated in tumorigenesis are thought to be dominant oncogenes, activated by translocation, amplification or viral insertion, rather than recessive tumor suppressors. Superficially, ELF5 appears to function similarly to other ETS family members (DNA binding and transactivation), but it is intriguing that *ELF5* expression appears to be lost in many cancer cell lines, a subset of which were found to have lost an allele or have rearrangement of the *ELF5* gene. A preliminary examination of *ELF5* expression indicates

that ELF5 is not detectable in a number of primary breast carcinomas, although it is strongly expressed in adjacent normal epithelium (data not shown). Thus, it may be possible for an ETS gene, such as *ELF5*, to have tumor suppressor properties and to be lost in certain cancers. Indeed, it appears that ETS1 may function to suppress tumorigenicity of colon cancer cells, whereas it is usually considered to be an oncogene (Suzuki *et al*., 1995). Several circumstantial lines of evidence suggest that there are unidentified ETS family members with tumor suppressor properties. An ETS element in the maspin promoter appears to be active in normal mammary and prostate epithelial cells, but inactive in tumor cells (Zhang *et al*., 1997a, b). Maspin is a tumor-suppressing serpin expressed in normal breast and prostate epithelium. A recent report implicating ELF3 (ERT) in the positive regulation of TGF-β RII transcription (Choi *et al*., 1998) notes that a poor TGF responsiveness, observed in many tumors and thought to contribute to malignant transformation, could be caused by defects in TGF receptor expression. Potentially, defects in maspin and TGF-β RII transcription could be caused by lack of the ETS factors that normally promote their expression. We are currently investigating the possibility that *ELF5*, involved in LOH of chromosome 11p13–15 in carcinomas, regulates the transcription of tumor suppressing genes.

In summary, we have isolated a novel ETS gene of the ELF sub-family which we have named *ELF5*. ELF5 displays high ETS domain similarity and similar expression pattern to that of the recently isolated epithelial-specific ETS gene, *ELF3* (*ESX*, *ESE-1*, *ERT*). The chromosomal localization of *ELF5*, and limited expression and gene rearrangement in human carcinoma cell lines, make this an interesting gene for further study in the field of human cancer.

Materials and methods

Isolation and characterization of full-length murine Elf5 *cDNA*

The murine *Elf5* cDNA was isolated from an adult lung cDNA library in Lambda ZAPII (Stratagene) following screening with a cDNA probe containing the ETS domain region of human *ELF3*. Additional 5′ sequence and 3′ sequence were obtained by RT-PCR using a Marathon cDNA synthesis Kit (Clontech) and RACE (Rapid Amplification of cDNA Ends) using day 14 murine placental Poly(A)$^+$ RNA. The murine *Elf5*-specific PCR products were cloned into pGEM-T vector (Promega Corp., Madison, WI, USA). All cDNA sequences were confirmed by sequencing both strands at least once. 5′-RACE gene-specific primer 1: 5′-GCCAGTCTTGGTCTCTTCAGCATC-3′; 5′-RACE nested-gene-specific primer 2: 5′-AGGAGATGCAGTTGGCATCAAGCT-3′; 3′-RACE gene-specific primer 1: 5′-AGCCAGTGTTATGGGTGCTG-3′; 3′-RACE nested-gene-specific primer 2: 5′-ACAGTCACTTGATCCACGGCCAATCC-3′.

Isolation of human ELF5 coding sequence

A human *ELF5* cDNA fragment was isolated from a human lung cDNA library (GIBCO BRL) following screening with a cDNA probe containing the coding sequence of mouse *Elf5*. The coding sequence was then obtained by RT-PCR using a Marathon cDNA synthesis Kit (Clontech) and RACE (Rapid Amplification of cDNA Ends) using human placental Poly(A)$^+$ RNA. The human *ELF5*-specific PCR products were cloned into pGEM-T vector (Promega Corp., Madison, WI, USA). All cDNA sequences were confirmed by sequencing both strands at least once (data not shown).

STS content mapping

The following sequence specific primers for human ELF5 were used for PCR. Forward primer: 5′-CCTGTGACTCATACTGGACATC-3′; Reverse primer: 5′-CTTGTGTGCCGATGTTCTGG-3′. The PCR reactions were performed in Opti-Primer™ 10 × buffer #3 (100 mM Tris-HCl pH 8.3, 35 mM $MgCl_2$, 250 mM KCl) with 1 μl of Master Mix 50 × buffer (20 mM Tris-HCl pH 8.0, 250 nM EDTA) (Opti-Primer™ PCR Optimization Kit, Stratagene), 50 ng of template DNA, 0.2 μg of each primer, 1 μl of 10 mM dNTPs and 0.25 U of Taq DNA polymerase in a total volume of 50 μl. PCR parameters were an initial denaturation step at 94°C for 1 min, followed by 30 cycles of 94°C (1 min), 60°C (1 min), 72°C (1 min). For Genebridge 4 Radiation Hybrid DNA panel (UK HGMP Resource Centre), PCR reactions were performed separately for each of the individual hybrids. The PCR results from the 93 hybrids were submitted to the Radiation Hybrid Mapping server at Whitehead Institute/MIT Center for Genome Research (http://www.genome.wi.mit.edu/cgi bin/contig/rhmapper.pl). The STS content mapping experiment was performed in duplicate and included PCR reactions with no DNA, total human DNA and total hamster DNA as controls.

Southern and Northern blot analysis

Northern analysis of *ELF5* expression in human adult organs was performed with commercially available blots containing 2 μg of Poly(A)$^+$ RNA (Clontech). For other Northern blots Poly(A)$^+$ mRNA was isolated by a modification of Gonda *et al*. (1992). Genomic DNA was isolated by standard techniques (Sambrook *et al*., 1997). Random-primed probes using a 898 bp human *ELF5* cDNA fragment and a 940 bp *StyI* mouse *Elf5* cDNA fragment were generated and Southern/Northern hybridizations performed using standard procedures. Blots were re-probed with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) or *β-actin* cDNAs to verify RNA/DNA loading.

RNAse protection analysis

ELF5 mRNA abundance in total RNA from human cell lines was determined as described previously (Tymms, 1995). Anti-sense RNA probes for human *ELF5* and GAPDH transcribed from linearized plasmid vectors generated full-length probes of 388 bp and 216 bp, respectively. The protected products generated by hybridization and RNAse digestion are 298 bp for *ELF5* and 160 bp for GAPDH.

Cell lines and culture

Monkey COS7 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, 100 U/ml penicillin and 100 μg/ml streptomycin, and maintained in a humidified incubator at 5% CO_2 and 37°C.

Plasmids

pHis6-Elf5 expression vector was made as follows: The murine *Elf5* cDNA was amplified using PCR oligonucleotide primers (5′-CGGGATCCTTGGACTCCGTAACCCATAGC-3′ and 5′-GCAGATCTCAGAGTTTCTCTTCC-

TGCC-3') containing a *Bam*HI restriction site followed by 21 nucleotides of the murine *Elf5* coding sequence and a *Bgl*II restriction site followed by 19 nucleotides complementary to the last 20 nucleotides of the *Elf5* coding sequence. The PCR fragment was cloned into the pGEM-T vector (Promega Corp., Madison, WI, USA), the *Bam*HI-*Sac*I restriction fragment with the *Elf5* coding sequence was then cloned into the *Bam*HI·*Sac*I sites of the pQE30 (Qiagen, Inc. Chatsworth, CA, USA) bacterial expression vector resulting in a N-terminal fusion of *Elf5* protein to six histidine residues (His-Tag).

The *Elf5* mammalian expression construct (pBOSElf5s) contains the full mouse *Elf5* cDNA blunt cloned into the T4 polymerase blunted *Xba*I site of pEFBOS (Mizushima and Nagata, 1990). Expression from pEFBOS is driven by the elongation factor-1 promoter. The *Elf5* anti-sense expression construct is similar, but with reverse orientation of the *Elf5* cDNA. p5Xpoly was made by cloning multimerized polyomavirus enhancer oligonucleotides into the *Bam*HI site of pBLCAT2.

Electrophoretic mobility shift assays (EMSA)

Purified recombinant *Elf5* and Ets1 proteins were produced as 6XHis-tag fusions in *E. coli* using the QIAexpress expression system (Qiagen). Overnight cultures were diluted 1/10 in LB broth and grown for 1 h at 37°C. Expression of recombinant proteins were induced by addition of 0.1 mM IPTG and culture of cells for 2 h. Cells were harvested and sonicated in lysis buffer (6 M guanidine, 20 mM Tris-HCl, 50 mM NaCl, pH 8.0), and cell debris removed by centrifugation. One ml of metal His-affinity resin was incubated with supernatants for 30 min, collected, washed in wash buffer (8 M urea, 20 mM Tris-HCl, 50 mM NaCl, pH 8.0), and resuspended in renaturation buffer (20 mM Tris-HCl, 50 mM NaCl, 3 mM dithiothreitol (DTT), pH 8.0). Proteins were eluted from the beads in renaturation buffer supplemented with 100 mM imidazole. Purification and integrity of recombinant proteins were confirmed by denaturing SDS-polyacrylamide gel electrophoresis (PAGE).

DNA binding experiments with recombinant proteins were performed using EMSA, as previously described (Thomas *et al*., 1995, 1997). Briefly, purified double stranded oligonucleotides were labeled with γ-^{32}P dATP and T4 polynucleotide kinase. Oligonucleotide probe (1 ng) was incubated for 10 min with approximately 20 ng purified *Elf5*/*Ets1* protein in DNA binding buffer (1 mM EDTA, 10 mM Tris-HCl pH 8.0, 50 mM NaCl, 3 mM DTT, 1 mg/ml BSA, 500 ng/ml poly-d(I-C)d(I-C), 500 ng/ml poly dI-dC, 200 ng/ml sheared salmon sperm DNA), ± 100 ng unlabelled competitor oligonucleotides, in 10 µl final volume. Assays were run through non-denaturing, 7% acrylamide (29 acrylamide:1 bis-acrylamide), 0.5 × TBE gels at 4°C.

Oligonucleotide sequences (shown in double stranded conformation)

E74	5'-gatcATAACCGGAAGTAACT-3'
	3'-TATTGGCCTTCATTGActag-5'
E74m1	5'-gatcATAACCAGAAGTAACT-3'
	3'-TATTGGTCTTCATTGActag-5'
GMETS	5'-gatcCACAGAGGAAATGATT-3'
	3'-GTGTCTCCTTTACTAActag-5'
MSV	5'-gatcGAGAGCGGAAGCGCGC-3'
	3'-CTCTCGCCTTCGCGCGctag-5'
ERBB2	5'-gatcGCTTGAGGAAGTATAA-3'
	3'-CGAACTCCTTCATATTctag-5'

Transfection of COS cells and chloramphenicol acetyltransferase (CAT) assays

COS7 cells were transfected with 5 µg CAT reporter plasmid and 10 µg of expression constructs by electroporation. Subconfluent cells were trypsinized, washed and resuspended in growth medium (as above) supplemented with 20 mM HEPES, at 5×10^6 cells/ml. 300 µl of cells were mixed with 20 µg DNA in 0.4 cm gap electroporation cuvettes and pulsed at 180 V and 960 µF (Biorad Gene Pulsar). Cells were re-plated into 10 cm petri dishes, harvested 48 h later. Cell lysates were assayed for protein concentration, and 2 µg was processed for CAT assays as previously described (Thomas *et al*., 1995). The basal expression of p5Xpoly was arbitrarily assigned the value of '100', and other raw data was normalized to this value. Means, and standard error of the means (s.e.m) were generated from four replicates of each experiment. Data was subject to statistical analysis using unpaired two-tailed *t*-tests, with resultant *p* values less than 0.05 considered significant.

Acknowledgements

The authors wish to thank Mr Daniel Lazner for kindly providing embryonic mouse tissues, and Dr Sika Ristevski for kindly providing day 14 mouse placental cDNA library, and Mr Peter Griffiths for expert technical assistance. We also thank the UK HGMP Resource Centre for allowing us to use the Genebridge 4 Radiation Hybrid DNA panel. This work was partially funded by grants from the National Health and Medical Research Council of Australia and the Anti-Cancer Council of Victoria. Genomic localization and gene symbol assigned 27 February 1998 (GDB Accession ID: 9835139). GeneBank Accessions: Human *ELF5*: AF049703; murine *Elf5*: AF049702.

References

Adachi J and Hasegawa M. (1996). *MOLPHY Version 2.3: Programs for molecular phylogenetics based on maximum likelihood*. Computer Science Monographs. The Institute of Statistical Mathematics. Tokyo.

Akashi M, Shaw G, Hachiya M, Elstner E, Suzuki G and Koeffler P. (1994). *Blood*, **83**, 3182–3187.

Baffa R, Negrini M, Mandes B, Rugge M, Ranzani GN, Hirohashi S and Croce CM. (1996). *Cancer Res.*, **56**, 268–272.

Ben-David Y, Giddens EB, Letwin K and Bernstein A. (1991). *Genes. Dev.*, **5**, 908–918.

Benz CC, O'Hagan RC, Richter B, Scott GK, Chang CH, Xiong X, Chew K, Ljung BM, Edgerton S, Thor A and Hassell JA. (1997). *Oncogene*, **15**, 1513–1525.

Bhat NK, Komschlies KL, Fujiwara S, Fisher RJ, Mathieson BJ, Gregorio TA, Young HA, Kasik JW, Ozato K and Papas TS. (1989). *J. Immunol.*, **142**, 672–678.

Bhat NK, Thompson CB, Lindsten T, June CH, Fujiwara S, Koizumi S, Fisher RJ and Papas TS. (1990). *Proc. Natl. Acad. Sci. USA*, **87**, 3723–3727.

Buttice G and Kurkinen M. (1993). *J. Biol. Chem.*, **268**, 7196–7204.

Buttice G, Duterque-Coquillaud M, Basuyaux JP, Carrere S, Kurkinen M and Stehelin D. (1996). *Oncogene*, **13**, 2297–2306.

Chang CH, Scott GK, Kuo WL, Xiong X, Suzdaltseva Y, Park JW, Sayre P, Erny K, Collins C, Gray JW and Benz CC. (1997). *Oncogene*, **14**, 1617–1622.

Choi SG, Yi Y, Kim YS, Kato M, Chang J, Chung HW, Hahm KB, Yang HK, Rhee HH, Bang YJ and Kim SJ. (1998). *J. Biol. Chem.*, **273**, 110–117.

Coleman WB, Esch GL, Borchert KM, McCullough KD, Reid LH, Weissman BE, Smith GJ and Grisham JW. (1997). *Mol. Carcinog.*, **19**, 267–272.

Cooper JA, Esch FS, Taylor SS and Hunter T. (1984). *J. Biol. Chem.*, **259**, 7835–7841.

Dahiya R, McCarville J, Lee C, Hu W, Kaur G, Carroll P and Deng G. (1997). *Int. J. Cancer*, **72**, 283–288.

Delannoy-Courdent A, Fauquette W, Dong-Le Bourhis XF, Boilly B, Vandenbunder B and Desbiens X. (1996). *Int. J. Dev. Biol.*, **40**, 1097–1108.

Doolittle RF, Feng D-F, Tsang S, Cho G and Little E. (1996). *Science*, **271**, 470–477.

Easteal S and Herbert G. (1997). *J. Molec. Evol.*, **44** (Suppl. 1), S121–S132.

Feinberg AP. (1996). *Med. Pediatr. Oncol.*, **27**, 484–489.

Fisher RJ, Mavrothalassitis G, Kondoh A and Papas TS. (1991). *Oncogene*, **6**, 2249–2254.

Fischer DF, Gibbs S, van De Putte P and Backendorf C. (1996). *Mol. Cell Biol.*, **16**, 5365–5374.

Gambarotta G, Boccaccio C, Giordano S, Ando M, Stella MC and Comoglio PM. (1996). *Oncogene*, **13**, 1911–1917.

Gao AC, Lou W, Dong JT and Isaacs JT. (1997). *Cancer Res.*, **57**, 846–849.

Gullick WJ, Berger MS, Benett PL, Rothbard JB and Waterfield MD. (1987). *Int. J. Cancer*, **40**, 246–254.

Golub TR, Barker GF, Bohlander SK, Hiebert SW, Ward DC, Bray-Ward P, Morgan E, Raimondi SC, Rowley JD and Gilliland DG. (1995). *Proc. Natl. Acad. Sci. USA*, **92**, 4917–4921.

Golub TR, Barker GF, Lovett M and Gilliland DG. (1994). *Cell*, **77**, 307–316.

Gonda TJ, Sheiness DK and Bishop JM. (1992). *Mol. Cell Biol.*, **2**, 617–624.

Graves BL and Petersen JM. (1998). *Adv. Cancer Res.*, in press.

Gutman A and Wasylyk B. (1991). *Trends Genet.*, **7**, 49–54.

Hart AH, Corrick CM, Tymms MJ, Hertzog PJ and Kola I. (1995). *Oncogene*, **10**, 1423–1430.

Hillis DM, Huelsenbeck JP and Swofford DL. (1994). *Nature*, **369**, 363–364.

Hirose M, Yamada T, Toyosaka A, Hirose T, Kagami S, Abe T and Kuroda Y. (1996). *Med. Pediatr. Oncol.*, **27**, 174–178.

Hunter T. (1982). *J. Biol. Chem.*, **257**, 4843–4848.

Ichikawa T, Nihei N, Kuramochi H, Kawana Y, Killary AM, Rinker-Schaeffer CW, Barrett JC, Isaacs JT, Kugoh H, Oshimura M and Shimazaki J. (1996). *Prostate Suppl.*, **6**, 31–35.

Ida K, Kobayashi S, Taki T, Hanada R, Bessho F, Yamamori S, Sugimoto T, Ohki M and Hayashi Y. (1995). *Int. J. Cancer*, **63**, 500–504.

Iizuka M, Sugiyama Y, Shiraishi M, Jones C and Sekiya T. (1995). *Genes Chromosomes Cancer*, **13**, 40–46.

Janknecht R and Nordheim A. (1993). *Biochim. Biophys. Acta.*, **1155**, 346–356.

Jones DT, Taylor WR and Thornton JM. (1992). *Comput. Appl. Biosci.*, **8**, 275–282.

Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki RA, Gunther CV, Nye JA and Graves BJ. (1990). *Genes Dev.*, **4**, 1451–1453.

Kawana Y, Komiya A, Ueda T, Nihei N, Kuramochi H, Suzuki H, Yatani R, Imai T, Dong JT, Imai T, Yoshie O, Barrett JC, Isaacs JT, Shimazaki J, Ito H and Ichikawa T. (1997). *Prostate*, **32**, 205–213.

Kishimoto A, Nishiyama K, Nakanishi H, Uratsuji Y, Nomura H, Takeyama Y and Nishizuka Y. (1985). *J. Biol. Chem.*, **260**, 12492–12499.

Kodandapani R, Pio F, Ni CZ, Piccialli G, Klemsz M, McKercher S, Maki RA and Ely KR. (1996). *Nature*, **380**, 456–460.

Kokai Y, Cohen JA, Drebin JA and Greene MI. (1987). *Proc. Natl. Acad. Sci. USA*, **84**, 8498–8501.

Kola I, Brookes S, Green AR, Garber R, Tymms M, Papas TS and Seth A. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 7588–7592.

Kristensen T, Ogata RT, Chung LP, Reid KB and Tack BF. (1987). *Biochemistry*, **26**, 4668–4674.

Lautenberger JA, Burdett LA, Gunnell MA, Qi S, Watson DK, O'Brien SJ and Papas TS. (1992). *Oncogene*, **7**, 1713–1719.

Lee JH, Jang SI, Yang JM, Markova NG and Steinert PM. (1996). *J. Biol. Chem.*, **271**, 4561–4568.

Lichy JH, Zavar M, Tsai MM, O'Leary TJ and Taubenberger JK. (1998). *Am. J. Pathol.*, **153**, 271–278.

Littlewood TD and Evan GI. (1994). *Protein Profile*, **1**, 635–709.

Macleod K, Leprince D and Stehelin D. (1992). *Trends Biochem. Sci.*, **17**, 251–256.

Mizushima S and Nagata S. (1990). *Nucleic Acids Res.*, **18**, 5322.

McKercher SR, Torbett BE, Anderson KL, Henkel GW, Vestal DJ, Baribault H, Klemsz M, Feeney AJ, Wu GE, Paige CJ and Maki RA. (1996). *EMBO J.*, **15**, 5647–5658.

Moreau-Gachelin F, Wendling F, Molina T, Denis N, Titeux M, Grimber G, Briand P, Vainchenker W and Tavitian A. (1996). *Mol. Cell Biol.*, **16**, 2453–2463.

Mori S, Akiyama T, Yamada Y, Morishita Y, Sugawara I, Toyoshima K and Yamamoto T. (1989). *Lab Invest.*, **61**, 93–97.

Moritz C and Hillis DM. (1996). Pp 1–13. In *Molecular systematics*, 2nd ed., (eds Hillis DM, Moritz C and Mable BK). Sinauer, Sunderland, USA.

Muthusamy N, Barton K and Leiden JM. (1995). *Nature*, **377**, 639–642.

Nunn MF, Seeburg PH, Moscovici C and Duesberg PH. (1983). *Nature*, **306**, 391–395.

Nye JA, Petersen JM, Gunther CV, Jonsen MD and Graves BJ. (1992). *Genes. Dev.*, **6**, 975–990.

O'Neill EM, Rebay I, Tjian R and Rubin GM. (1994). *Cell*, **78**, 137–147.

Oettgen P, Alani RM, Barcinski MA, Brown L, Akbarali Y, Boltax J, Kunsch C, Munger K and Libermann TA. (1997). *Mol. Cell Biol.*, **17**, 4419–4433.

Patschinsky T, Hunter T, Esch FS, Cooper JA and Sefton BM. (1982). *Proc. Natl. Acad. Sci. USA*, **79**, 973–977.

Peeters P, Raynaud SD, Cools J, Wlodarska I, Grosgeorge J, Philip P, Monpoux F, Van Rompaey L, Baens M, Van den Berghe H and Marynen P. (1997). *Blood*, **90**, 2535–2540.

Peter M, Couturier J, Pacquement H, Michon J, Thomas G, Magdelenat H and Delattre O. (1997). *Oncogene*, **14**, 1159–1164.

Pinna LA. (1990). *Biochim. Biophys. Acta.*, **1054**, 267–284.

Press MF, Cordon-Cardo C and Slamon DJ. (1990). *Oncogene*, **5**, 953–962.

Sambrook J, Fritsch ER and Maniatis T. (1997). *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor, Laboratory Press.

Savant-Bhonsale S and Cleveland DW. (1992). *Genes. Dev.*, **6**, 1927–1939.

Scott EW, Simon MC, Anastasi J and Singh H. (1994a). *Science*, **265**, 1573–1577.

Scott GK, Daniel JC, Xiong X, Maki RA, Kabat D and Benz CC. (1994b). *J. Biol. Chem.*, **269**, 19848–19858.

Seth A and Papas TS. (1990). *Oncogene*, **5**, 1761–1767.

Seth A, Ascione R, Fisher RJ, Mavrothalassitis GJ, Bhat NK and Papas TS. (1992). *Cell Growth Differ.*, **3**, 327–334.

Seth A, Watson DK, Blair DG and Papas TS. (1989). *Proc. Natl. Acad. Sci. USA*, **86**, 7833–7837.

Sgouras DN, Athanasiou MA, Beal GJ Jr, Fisher RJ, Blair DG and Mavrothalassitis GJ. (1995). *EMBO J.*, **14**, 4781–4793.

Singleton TP and Strickler JG. (1992). *Pathol. Annu.*, **27**, 165–190.

Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ and Denny CT. (1994). *Nat. Genet.*, **6**, 146–151.

Sumarsono SH, Wilson TJ, Tymms MJ, Venter DJ, Corrick CM, Kola R, Lahoud MH, Papas TS, Seth A and Kola I. (1996). *Nature*, **379**, 534–537.

Sun W, Graves BJ and Speck NA. (1995). *J. Virol.*, **69**, 4941–4949.

Suzuki H, Romano-Spica V, Papas TS and Bhat NK. (1995). *Proc. Natl. Acad. Sci. USA*, **92**, 4442–4446.

Thomas RS, Tymms MJ, Mckinlay LH, Shannon MF, Seth A and Kola I. (1997). *Oncogene*, **14**, 2845–2855.

Thomas RS, Tymms MJ, Seth A, Shannon MF and Kola I. (1995). *Oncogene*, **11**, 2135–2143.

Thompson JD, Higgins DG and Gibson TJ. (1994). *Nucleic Acids Res.*, **22**, 4673–4680.

Tripathy D and Benz C. (1994). *Hematol. Oncol. Clin. North. Am.*, **8**, 29–50.

Tymms MJ. (1995). *Methods Mol. Biol.*, **37**, 31–46.

Tymms MJ, Ng AY, Thomas RS, Schutte BC, Zhou J, Eyre HJ, Sutherland GR, Seth A, Rosenberg M, Papas T, Debouck C and Kola I. (1997). *Oncogene*, **15**, 2449–2462.

Vandenbunder B, Wernert N, Queva C, Desbiens X and Stehelin D. (1994). *Folia. Biol. (Praha.)*, **40**, 301–313.

Wang CY, Petryniak B, Ho IC, Thompson CB and Leiden JM. (1992). *J. Exp. Med.*, **175**, 1391–1399.

Wasylyk B, Hahn SL and Giovane A. (1993). *Eur. J. Biochem.*, **211**, 7–18.

Wasylyk C, Gutman A, Nicholson R, and Wasylyk B. (1991). *EMBO. J.*, **10**, 1127–1134.

Wasylyk C, Kerckaert JP and Wasylyk B. (1992). *Genes. Dev.*, **6**, 965–974.

Watson DK, McWilliams MJ, Lapis P, Lautenberger JA, Schweinfest CW and Papas TS. (1988). *Proc. Natl. Acad. Sci. USA*, **85**, 7862–7866.

Watson DK, Mavrothalassitis GJ, Jorcyk CL, Smyth FE and Papas TS. (1990). *Oncogene*, **5**, 101–107.

Welte T, Garimorth K, Philipp S, Jennewein P, Huck C, Cato AC and Doppler W. (1994). *Eur. J. Biochem.*, **223**, 997–1006.

Werner MH, Clore M, Fisher CL, Fisher RJ, Trinh L, Shiloach J and Gronenborn AM. (1995). *Cell*, **83**, 761–771.

Wernert N, Raes MB, Lassalle P, Dehouck MP, Gosselin B, Vandenbunder B and Stehelin D. (1992). *Am. J. Pathol.*, **140**, 119–127.

Wilson AP, Dent M, Pejovic T, Hubbold L and Radford H. (1996). *Br. J. Cancer*, **74**, 722–727.

Woodget JR, Gould KL and Hunter T. (1986). *Eur. J. Biochem.*, **161**, 177–184.

Yang B-S, Hauser CA, Henkel G, Colman MS, Van Beveren C, Stacey KJ, Hume DA, Maki RA and Ostrowski C. (1996). *Mol. Cell. Biol.*, **16**, 538–547.

Zenklusen JC, Oshimura M, Barrett JC and Conti CJ. (1995). *Genes Chromosomes Cancer*, **13**, 47–53.

Zhang M, Maass N, Magit D and Sager R. (1997a). *Cell. Growth Differ.*, **8**, 179–186.

Zhang M, Magit D and Sager R. (1997b). *Procl. Natl. Acad. Sci. USA*, **94**, 5673–5678.

Oncogene (1997) 15, 2449–2462

A novel epithelial-expressed ETS gene, *ELF3*: human and murine cDNA sequences, murine genomic organization, human mapping to 1q32.2 and expression in tissues and cancer

Martin J Tymms[1], Annie YN Ng[1], Ross S Thomas[1], Brian C Schutte[2], Jiong Zhou[1], Helen J Eyre[3], Grant R Sutherland[3], Arun Seth[4], Martin Rosenberg[5], Takis Papas[6], Christine Debouck[5] and Ismail Kola[1]

[1]*Molecular Genetics and Development Group, Institute of Reproduction and Development, Monash University, Melbourne, Victoria, Australia 3168;* [2]*Department of Pediatrics, University of Iowa, Iowa City, Iowa 52242, USA;* [3]*Centre for Medical Genetics, Department of Cytogenetics and Molecular Genetics, Women's and Children's Hospital, Adelaide, South Australia, Australia 5006;* [4]*Department of Pathology, Faculty of Medicine and Faculty of Dentistry, University of Toronto and Women's College Hospital, Toronto, Ontario MDS 1A8, Canada M5S 1A8;* [5]*SmithKline Beecham Pharmaceuticals, PO Box 1539, King of Prussia, Pennsylvania 19406-0939, USA;* [6]*Center for Molecular and Structural Biology, Hollings Oncology Center, Medical University of South Carolina, 171 Ashley Avenue, Charleston, South Carolina 29425-2213, USA*

The ETS family of genes are implicated in cancers such as Ewings sarcoma, acute myeloid leukemia and chronic myelomonocytic leukemia. Further, they have important functions in embryonic development. Hence, identification and characterization of members of this family are important. We identify a novel ETS family member, *ELF3*, and report its human and murine cDNA sequences. The mouse cDNA has an alternatively spliced transcript with an extra 60 bp inserted. Hence we present the organization of the murine *Elf3* gene together with its exon/intron structure. This gene consists of 9 exons and 8 introns spanning 4.8 kb. ELF3 binds and transactivates ETS sequences and interestingly also shows the ability to bind a GGAT-like purine core, a preferential ETS1/ETS2 type binding site. The expression of *ELF3*, unlike most other ETS family members, is absent in hematopoietic cells and hematopoietic organs in humans and mice. Intriguingly, the gene is specifically expressed in cell lines of epithelial origin and in organs such as lung, stomach, intestine, kidney that have specialized epithelial cells. We localize the human gene to 1q32.2, a region that is amplified in epithelial tumors of the breast, lung and prostate. Finally, we show that *ELF3* expression is increased in a lung carcinoma and adenocarcinoma, as compared to normal tissue. *ELF3* is also expressed in cell lines derived from lung cancers. These results suggest that this novel ETS gene may be involved in lung tumorigenesis.

Keywords: transcription factor; lung cancer; breast cancer; ETS; ELF; prostate cancer

Introduction

The ETS family of transcription factors play a role in cancer and disease (Wasylyk *et al.*, 1993; Hromas and Klemsz, 1994). The v-*ets* oncogene was originally described as a part of the *gag-myb-ets* fusion gene expressed by the avian retrovirus E26 (Nunn *et al.*, 1983). The cellular homolog of v-*ets* is ETS1 which is a transcription factor that binds to a GGA(A/T) core sequence in the promoters and enhancers of various cellular genes and viruses. *ETS1*, *ETS2*, *ERG2* and PU-1 are proto-oncogenes which have cellular transforming and mitogenic properties when over-expressed (Seth *et al.*, 1989; Seth and Papas, 1990; Hart *et al.*, 1995; Moreau-Gachelin *et al.*, 1996). Translocations involving ETS family members are associated with cancers in humans. *ERG* and closely related homologue *ERGB/FLI-1* have been shown to be disrupted in t(21;22) and t(11;22) chromosomal translocations, respectively, that are diagnostic of Ewings sarcomas and other primitive neuroectodermal tumors (Sorensen *et al.*, 1994). These translocations involve fusion of ERG/FLI1 sequences to the EWS gene (Ida *et al.*, 1995). *ERG* is also involved in t(16;21) translocations associated with acute myeloid leukemia (Shimizu *et al.*, 1993). ETS family member *TEL* is involved in t(5;12) translocations in chronic myelomonocytic leukemia (Golub *et al.*, 1994) and t(12;21) translocations in acute lymphoblastic leukemia (Golub *et al.*, 1995).

The ETS family of transcription factors are conserved from Drosophila to humans. In Drosophila, ETS family members play important developmental roles: E74 plays a role in early gene regulation following ecdysone triggering (Burtis *et al.*, 1990; Fletcher *et al.*, 1995); *pointed P2* and *yan* play critical roles in eye development (O'Neill *et al.*, 1994); and *D-elg* plays a role in developmental programming (Schulz *et al.*, 1993). In humans and mice ETS proteins may function in a wide range of biological processes such as hematopoiesis (Scott *et al.*, 1994a; Muthusamy *et al.*, 1995), organogenesis and branching morphogenesis (Kola *et al.*, 1993), and skeletal/cartilage development (Sumarsono *et al.*, 1996). These proteins have also been implicated in regulation of gene expression appropriate for lymphoid-function (Bhat *et al.*, 1989, Ho *et al.*, 1990; Bhat and Thompson, 1990; Thomas *et al.*, 1995) and invasive processes such as tumor metastasis and extracellular matrix degradation (Wernert *et al.*, 1992).

One of the most common solid tumors in humans are carcinomas which arise from the transformation of epithelial cells. Transformed breast epithelial cells have been shown to express ETS family members *GABPα*, *PEA-3*, *ELF1* and *ELK1* (Scott *et al.*, 1994b), but expression of these members is not restricted to

Correspondence: I Kola
MJT and AYNN equally contributed to this paper
Received 24 April 1997; revised 14 July 1997; accepted 14 July 1997

epithelial cells. Normal epithelial organization is dependent upon specific cell adhesion involving E-cadherin, actin cytoskeletal-associated proteins, and integrins and the decreased expression of E-cadherin and integrins occur in breast and prostate carcinomas, and this decrease correlates with increased invasiveness (Oka *et al.*, 1993; Zutter *et al.*, 1993). Epithelial tubulogenesis is responsible for the intricate organization of parenchymal organs such as the lung and mammary gland where elongation and branching of epithelial tubules is critical for organ formation (Wu and Santoro, 1996). During development new epithelia usually arise from existing ones but mesenchymal cells are the source of new epithelial cells during organogenesis of the kidney, ovary and testis. Soluble factors produced by mesenchymal cells are thought to regulate the growth, motility and differentiation of epithelial cells *via* tyrosine kinase receptors. One critical interaction appears to be between HGF (hepatocyte growth factor or scatter factor) which is produced by mesenchymal cells and the tyrosine kinase receptor c-*met* which is expressed predominantly on epithelial cells (Weider *et al.*, 1996). Co-expression of c-*met* and its ligand HGF in NIH3T3 fibroblasts results in the formation of tumors in nude mice with epithelial characteristics (Tsarfaty *et al.*, 1994). A number of studies have implicated ETS family members in the regulation of genes important in epithelial cell function including *erb*-B2 (Scott *et al.*, 1994b), c-*met* (Gambarotta *et al.*, 1996) and mammary whey acidic protein (Welte *et al.*, 1994).

In this study we describe a new ETS family member *ELF3* which is expressed in epithelial cells. The human *ELF3* has been mapped to the 1q32.2 region which is frequently associated with epithelial cancers. ELF3, like other ETS family members, functions as a sequence specific DNA binding and transactivational factor. Further investigations may define an important role for this gene in epithelial cancers of the lung, breast and prostate.

Results

Isolation of a full length human and murine ELF3/Elf3 cDNA

A cDNA fragment encoding the ETS1 DNA binding domain was used to screen the Human Genome Sciences (HGS) cDNA expressed sequence tag (EST) database. A

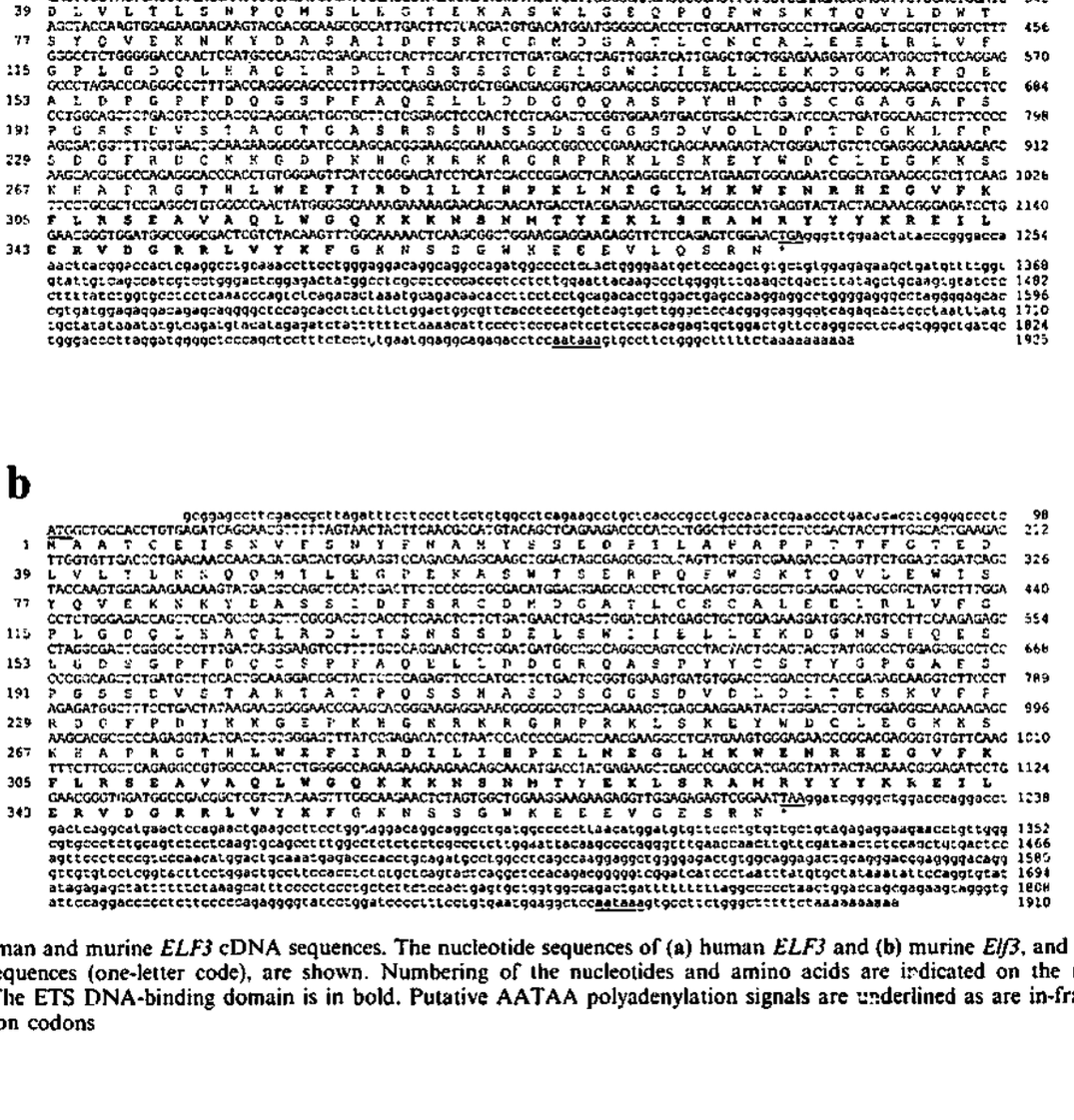


Figure 1 Human and murine *ELF3* cDNA sequences. The nucleotide sequences of (a) human *ELF3* and (b) murine *Elf3*, and their deduced amino acid sequences (one-letter code), are shown. Numbering of the nucleotides and amino acids are indicated on the right and left respectively. The ETS DNA-binding domain is in bold. Putative AATAA polyadenylation signals are underlined as are in-frame initiation and termination codons

homologous partial cDNA for *EPR-1* (*E*TS-related protein from *pr*ostate, HGS12160) was identified from a human prostate cancer library. Preliminary expression analysis by RNase protection showed that *EPR-1* was expressed highly in fetal lung, in addition to prostate cancer cell lines (data not shown). A full length human cDNA was subsequently obtained by a combination of screening a fetal lung lambda cDNA library and by 5'-RACE. The full length cDNA sequence is 1915 bp, and predicted to encode a protein of 371 amino acids (Figure 1a). Northern analysis of human *ELF3* mRNA shows the presence of minor larger-sized transcripts (Figure 5a). Analysis of a larger cDNA clone from a fetal lung lambda cDNA library has established that the minor transcripts contain additional 3' untranslated sequences which result from differential polyadenylation (data not shown).

The equivalent full-length murine cDNA was also isolated. This was achieved by screening an adult mouse lung cDNA library and by 5'-RACE. The full-length murine cDNA is 1900 bp (Figure 1b), which is predicted to encode a 371 amino acid protein which displays 89% identity and 93% similarity with the human amino acid sequence (Figure 2a). Upon advice from the Genome Database (GDB) nomenclature committee, this gene was assigned the gene symbol *ELF3* as of May 10 1996, the date of deposition of the localization on the GDB (Accession ID: 1230388). The human and mouse *ELF3* cDNAs encode an identical putative ETS DNA-binding domain (domain C) of 80 amino acids. The ETS domain of ELF3 is moderately homologous to that of other ETS family members, with the highest amino acid identity being 49%, to Drosophila E74A (Figure 2b). Of all of the human ETS family members, ELF3 had the highest identity with ELF1 (and NERF) within the ETS domain. However, it is interesting to note that, amongst the ETS domain sequences of various ETS family members, ELF3 is relatively divergent. The human ELF3 sequence also contains a consensus PEST domain (rich in proline, serine and threonine amino acids) and the murine amino acid sequence is highly homologous in this region (Figure 2a). Thus, this sequence may play a role in protein stability due to protease targeting.

A computer-generated phylogenetic tree, based on ETS domain homology, suggests that ELF3 may have diverged from other ETS family members, such as ETS1 and ELF1, at a fairly early stage in the evolution of ETS genes (Figure 2c). Perhaps significantly, ELF3 also displays considerable N-terminal homology (domain A) with ETS1, ETS2, ERG2, TEL and GABPα (Figure 2d).

Isolation of the entire murine genomic sequence

Many ETS genes are alternatively spliced to give mature mRNAs which encode protein variants. Sequencing of 5'-RACE PCR products, during the cloning of the murine cDNA, identified a clone containing a 60 bp in-frame insertion in the murine *Elf3* cDNA sequence. This insert could arise from differential splicing of the murine *Elf3* gene. To ascertain if this insertion arose from splicing and additionally to facilitate gene targeting experiments, we isolated and sequenced the full *Elf3* genomic sequence. Genomic clones were isolated from a 129SVJ genomic library. Two independent clones, λ3.24 and λ3.26, encompassing 30 kb of genomic sequence, span the entire *Elf3* gene including 10 kb of sequence 5'-prime to the transcriptional initiation site (Figure 3a). The 9 exons and 8 introns span 4.8 kb, with the largest intron (intron 7) only being 853 bp. Exon/intron sizes, sequences of exon/intron junctions and their genomic positions are shown in Figure 3b. All exon/intron boundaries conform to the Breatnach (GT/AG) rule (Breatnach *et al*., 1978) and all splice donor and acceptor sites conform to Mount's consensus sequence (Mount, 1982). Examination of the sequence of intron II suggests that the *Elf3* cDNA variant containing the extra 60 bp insertion (named *Elf3b*) results from the use of an alternate splice acceptor site 60 bp 5-prime of the exon 3 splice acceptor site of *Elf3a* (Figure 3a and b). This additional exonic sequence is predicted to contribute 20 more amino acids to the protein encoded by *Elf3b*.

The site of transcriptional initiation was determined by 5'-RACE PCR. A specific *Elf3* region was amplified (Figure 3c) and upon sequencing of several clones containing these PCR products, a dominant transcriptional initiation site was identified (Figure 3d). This dominant transcriptional initiation site was 51 bp 3' of a consensus TATA box and 127 bp 3' of a CAAT box (Figure 3d).

ELF3 is localized to human chromosome 1q32.2; a region implicated in epithelial cancers and encompassing the Van der Woude Syndrome (VWS) gene

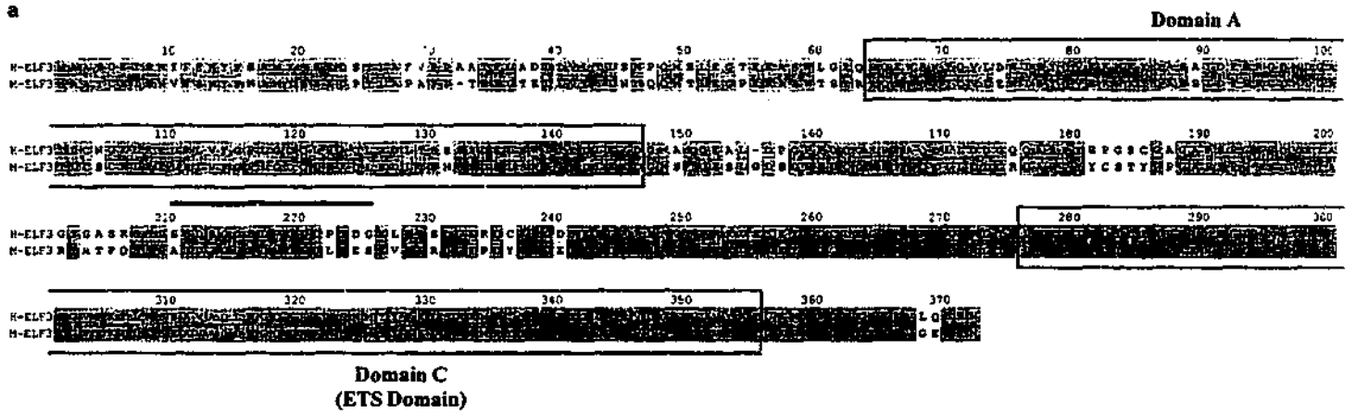
We performed human chromosomal localization of *ELF3* by FISH, using the *ELF3* cDNA as a probe. Examination of 20 metaphases from a normal human male showed 19 metaphases with signal on one or both chromatids of chromosome 1, in the region 1q32-1q41 (76% of the signal was at 1q32.2) (Figure 4a). There was a total of 6 non-specific background dots observed in these 20 metaphases. A similar result was obtained from hybridization of the probe to 15 metaphases from a second normal male (data not shown). This localization data was lodged with the GDB. The localization of *ELF3* is significant as it is to a region of chromosome 1 that encompasses the *VWS* gene, and also lies within a region implicated in breast cancer specifically and epithelial cancers in general (1q).

To assess if *ELF3* is a candidate for the *VWS* gene and to obtain finer resolution mapping of *ELF3* we carried out STS content mapping using a set of primers derived from the 3' UTR of *ELF3*. With these primers, a single product of the expected size (182 bp) was amplified from total human DNA (Figure 4b). The nucleotide sequence of this product was identical to the *ELF3* cDNA sequence (data not shown) demonstrating that the primers were specific for the *ELF3* locus. Three sets of genomic DNA templates were used for the STS content mapping experiments. The Coriell Mapping Panel 2 was used to verify the chromosomal assignment of *ELF3* and two radiation hybrid panels, Genebridge 4 and Stanford G3, were used to refine the chromosomal localization. When the *ELF3*-specific primers were used in PCR reactions with the Coriell Monochromosome Panel B, a product of the correct size was amplified in DNA pools A and B (Figure 4b). The only human chromosome contained in these two

pools was chromosome 1, verifying the assignment of *ELF3* to chromosome 1. The same primers were used to screen the radiation hybrid panels. The amplification results (data not shown) were submitted to their respective computer servers for analysis (see Materials and methods). For the Genebridge 4 Panel, the marker most tightly linked to *ELF3* was WI9272 at a distance of 2.4 cR (lod score > 3). A more precise genetic localization can be inferred by identifying the closest flanking markers, D1S412 and D1S504, which were also provided by the Radiation Hybrid Mapper server (Figure 4c). The relative positions of these markers on chromosome 1 are 242 cM and 253 cM from the top (Dib *et al.*, 1996). For the Stanford G3 Panel, the marker most tightly linked to *ELF3* was D1S306 at a distance of 34 cR (lod score > 4), which is 249 cM from the top of chromosome 1 (Dib *et al.*, 1996). *ELF3* is therefore located 14 cM proximal to the VWS gene, and hence we have excluded *ELF3* as being the gene involved in VWS. The markers found near *ELF3* map to the YAC contig WC1.18 (Hudson *et al.*, 1995) which also contains the cardiac troponin T (*TNNT2*) and the slow skeletal muscle troponin I (*TNNI1*) genes (B Schutte, manuscript in preparation). Both *TNNT2* and *TNNI1* map to 1q32 (Mesnard *et al.*, 1995; Eyre *et al.*, 1993), and support the cytogenetic localization of *ELF3* shown in Figure 4a.

Unlike most ETS family members ELF3 is not expressed in hematopoietic cells

Analysis of *ELF3* expression in adult human tissues by Northern blot to poly(A)$^+$RNA (Figure 5a) and in human fetal tissues by RNase protection analysis (Figure 5b) demonstrates that *ELF3* has a restricted expression pattern. Most apparent is the lack of expression in hematopoietic organs such as spleen and thymus, which are known to express many of the known ETS family of genes (Bhat *et al.*, 1997), and in testicular and ovarian tissues. Further, *ELF3* expression was not detected in peripheral blood lymphocytes (PBLs), reinforcing the idea that expression is absent/low in hematopoietic cells. Analysis of cell lines (Figure 5c) shows that *ELF3* expression is absent from all but those with an epithelial origin. Lung tissue, which displays *ELF3* expression contains a significant proportion of specialized epithelial cells, type I and type II pneumocytes. Also *ELF3* is highly expressed in fetal lung (Figure 5b), but at lower levels in adult lung (Figure 5a). It is undetectable in brain and heart of



b

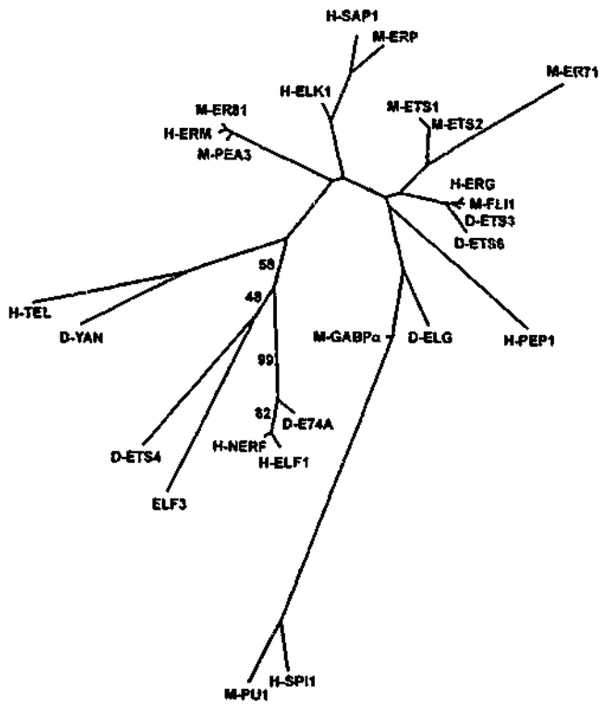
		Identity %
ELF3	PTALWEFIRDILIHP-ELNEGLMKDENRHEGVFKFL--RSEAVADLWGDKK-KNSNMTYEKLSRAMRYYYKREILERVD-GRKLVYKF	100
D-E74A	TTYLWEFLLKLLQDR-EYCPRFIKWTNREKGVFKLV--DSKAVSRLWGMHK-NKPDMNYETMGRALRYYYQRGILAKVD-GQRLVYHF	49
H-ELF1	TIYLWEFLLALLQDK-ATCPKYIKWTQREKGIFKLV--DSKAVSRLWGKHK-NKPDMNYETMGRALRYYYQRGILAKVE-GQRLVYQF	43
H-NERF	TTYLWEFLLDLLQDK-NTCPRYIKWTQREKGIFKLV--DSKAVSKLWGKHK-NKPDMNYETMGRALRYYYQRGILAKVE-GQRLVYQF	43
H-PEP1	QIQLWHFILELLQKE--EFRHVIAWQQGEYGEEVIK--DPDEVARLWGRRK-CKPQMNYDKLSRALRYYYNKRILHKTK-GKRFTYKF	40
D-ETS4	HIHLWQFLKELLASP-QVNGTAIRMIDRSKGIFKIE--DSVRVAKLWGRRK-NRPAMNYDKLSRSIRQYYKKGIMKKTERSQRLVYQF	40
D-YAN	GRLLWDFLQQLLNDRNQKYSDLIAWKCRDTGVFKIV--DPAGLAKLWGIQK-NHLSMNYDKMSRALRYYYRVNILRKVQ-GERHCYQF	39
H-TEL	CRLLWDYVYQLLSDS--RYENFIRWEDKESKIFRIV--DPNGLARLWGNHK-NRTNMTYEKMSRALRHYYKLNIIRKEP-GQRLLFRF	36
H-ELK1	SVTLWQFLLQLLREQ--GNGHIISWTSRDGGEFKLV--DAEEVARLWGLRK-NKTNMNYDKLSRALRYYYDKNIIRKVS-GQKFVYKF	43
M-ERP	AITLWQFLLHLLLDQ--KHEHLICWTSNDG-EFKLL--KAEEVAKLWGLRK-NKTNMNYDKLSRALRYYYDKNIIKKVI-GQKFVYKF	44
H-SAP1	AITLWQFLLQLLQKP--QNKHMICWTSNDG-QFKLD--QAEEVARLWGIRK-NKPNMNYDKLSRALRYYYVKNIIKKVN-GQKFVYKF	43
M-ER71	PIQLWQFLLKLLQDG--ARSSCIRWTG-NSREFQLC--DPKEVARLWGERK-RKPGMNYEKLSRGLRYYYRRDIVLKSG-GRKYTYRF	36
H-ERG	QIQLWQFLLELLSDS--SNSSCITWEG-TNGEFKMT--DPDEVARRWGERK-SKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	41
M-FLI1	QIQLWQFLLELLSDS--ANASCITWEG-TNGEFKMT--DPDEVARRWGERK-SKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	41
D-ETS6	QIQLWQFLLELLADS--SNANAISWEG-QSGEFRLI--DPDEVARRWGERK-AKPNMNYDKLSRALRYYYDKNIMTKVH-GKRYAYKF	40
M-ETS1	PIQLWQFLLELLTDK--SCQSFISWTG-DGWEFKLS--DPDEVARRWGKRK-NKPKMNYEKLSRGLRYYYDKNIIHKTA-GKRYVYRF	37
M-ETS2	PIQLWQFLLELLSDK--SCQSFISWTG-DGWEFKLA--DPDEVARRWGKRK-NKPKMNYEKLSRGLRYYYDKNIIHKTS-GKRYVYRF	37
M-GABPα	QIQLWQFLLELLTDK--DARDCISWVG-DGEEFKLN--QPELVAQKWGQRK-NKPTMNYEKLSRALRYYYDGDMICKVQ-GKRFVYKF	43
D-ELG	QVQLWQFLLELLTDC--EHTDVIEWVG-TGEEFKLT--DPDRVARLWGERK-NKPAMNYEKLSRALRYYYDGDMISKVS-GKRFAYKF	44
M-ER81	SLQLWQFLVALLDDP--SNSHFIAWTG-RGMEFKLI--EPEEVARRWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYKF	40
H-ERM	SLQLWQFLVTLLDDP--ANAHFIAWTG-RGMEFKLI--EPEEVARRWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYKF	40
M-PEA3	ALQLWQELVALLDDP--TNAHFIAWTG-RGMEFKLI--EPEEVARLWGIQK-NRPAMNYDKLSRSLRYYYEKGIMQKVA-GERYVYRF	41
D-ETS3	QIQLWQFLLELLSDS--NNASCITWEG-TNGEFKLT--DPDEVARRWGERK-SKPNMNYDKLSRALR----------	38
M-PU1	KIRLYQFLLDLERSG--DMKDSIWWVDKDKGTFQFSSKHKEALAHRWGIQKGNRKKMTYQKMARALRNYGKTGEVKKVK--KKLTYQF	32
H-SPIB	KLRLYQFLLGLLTRG--DMRECVWWVEPGAGVFQFSSKHKELLARRWGQQKGNRKRMTYQKLARALRNYAKTGEIRKVK--RKLTYQF	35

fetal tissues and seen at low levels in fetal thymus (Figure 5b). It may be that the expression in fetal thymus is not due to lymphoid cell expression but in some other cell-type.

A significant number of lung cancers, adenocarcinomas and large cell carcinomas are derived from respiratory epithelium. We therefore examined *ELF3* expression in some primary cancers and cell lines derived from lung cancers. Expression of *ELF3* was detected in a primary large cell carcinoma and a primary adenocarcinoma, and at a higher level than in age matched normal lung (Figure 5d, lanes P1, P2 and N). *ELF3* was also detected in adenocarcinoma cell line PL27 (Figure 5d, lane C3) and carcinoma A549 (Figure 5d, lane C5), but not in the small cell carcinoma RHSCC11 (Figure 5d, lane C1) or squamous cell carcinomas SK-MES2 and L162 (Figure 5d, lane C2 and C4). This pattern of expression suggests that transcription of *ELF3* may be differentially regulated in different lung cancer sub-types.

Analysis of *Elf3* expression in murine tissues demonstrates a similar pattern to that found in human tissues (Figure 6). *Elf3* expression in adult mice was observed in small intestine, colon, lung, kidney and uterus, but no expression was seen in spleen, thymus, brain, heart, skeletal muscle, or ovary (Figure 6a). However adult murine liver displayed no *Elf3* expression, whereas adult human liver displayed high *ELF3* expression. Further investigation will be necessary to determine if *ELF3* expression occurs in normal human liver or may have resulted from some other causes such as liver damage or cancer, especially

c



d

		Identity %
H-ELF3	PQFWSKTQVLDWISYQVEKNKYDASAIDFSRCDMDGATLCNCALEELRLVFGPLGDQLHAQLRDLTSSSSDELSWIIELLEKD	100
H-ETS1	PRQWTETHVRDWVMWAVNEFSLKG--VDFQKFCMNGAALCALGKDCFLELAPDF----------VGDILWEHLEILQKE	28
H-GABPα	PIQWSTDQVLHWVVWVMKEFSMTD--IDLTTLNISGRELCSLNQEDFFQRVPRG----------EILWSHLELLRKY	27
H-ETS2	PWLWSEQQVCQWLLWATNEFSLVN--VNLQRFGMNGQMLCNLGKERFLELAPDF----------VGDILWEHLEQMIKE	24
H-TEL	PIYWSRDDVAQWLKWAENEFSLRP--IDSNTFEMNGKALLLLTKEDFRYRSPH----------SGDVLYELLQHILKQ	24
H-ERG	PTLWSTDHVRQWLEWAVKEYGLPDV-NILLFQNIDGKELCKMTKDDFQRLTPSY----------NADILLSHLHYLRET	19

Figure 2 Comparison of human and murine ELF3/Elf3 with other ETS family members. (a) Alignment of the human (H-ELF3) and murine (M-ELF3) protein sequences. Gaps, indicated by dashes, were introduced to optimize alignment. Identical amino acids are dark-shaded and similar amino acids are light-shaded. Overall, identity and similarity between the human and murine counterparts are 89 and 93%, respectively. Numbering of amino acids is shown. Domain A and domain C (ETS domain) are boxed and labeled. The region overlined by a solid bar corresponds to a putative PEST sequence. (b) Alignment and comparison of the ETS domain (domain C) of human/mouse ELF3 with other ETS family members in different species. ETS proteins are indicated with a species prefix: D (Drosophila), H (human), M (mouse). Proteins examined include: E74A (Burtis *et al*., 1990), ELF1 (Leiden *et al*., 1992), NERF (Oettgen *et al*., 1996), PEP1 (Klemsz *et al*., 1994), ETS4 (Chen *et al*., 1992), YAN (Lai and Rubin, 1992), TEL (Golub *et al*., 1994), ELK1 (Rao *et al*., 1989), ERP (Lopez *et al*., 1994), SAP1 (Dalton and Treisman, 1992), ER71 (Brown and McKnight, 1992), ERG (Rao *et al*., 1987), FLI1 (Ben-David *et al*., 1991), ETS6 (Chen *et al*., 1992), ETS1 (Chen, 1990), ETS2 (Watson *et al*., 1988), GABPα (LaMarco *et al*., 1991), ELG (The *et al*., 1992), ER81 (Monte *et al*., 1995), ERM (Monte *et al*., 1994), PEA3 (Xin *et al*., 1992), ETS3 (Chen *et al*., 1992), PU1 (Klemsz *et al*., 1990), and SPIB (Ray *et al*., 1992). Amino acids with identity to ELF3 are shaded. Overall identity is given as a percentage. Alignments were produced with CLUSTAL W (Thompson *et al*., 1994). (c) Phylogenetic tree of ETS DNA binding domains produced using maximum likelihood analysis. The alignment shown in b was analysed using the JTT-F substitution model (Jones *et al*., 1992) and local bootstrap values were determined for all internal branches by PROTML using R-mode search (Adachi and Hasegawa, 1996). Only the bootstrap values for branches close to ELF3 are shown. (d) Alignment and comparison of the N-terminal sequence (domain A) of ELF3 with other ETS family members

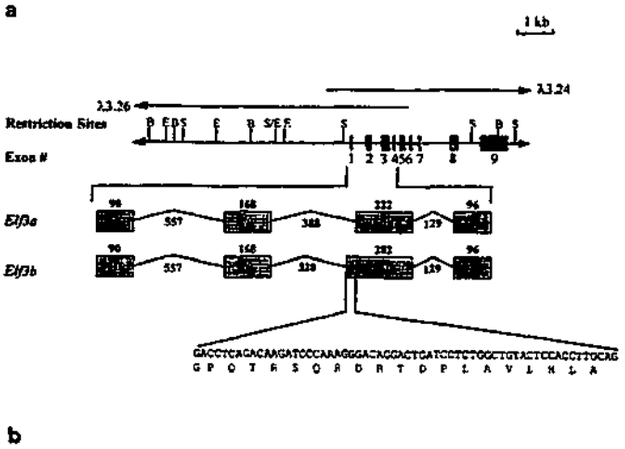
since the exact status/condition of the human RNA sample is unknown.

In the light that the organ expression patterns of some ETS genes vary during murine development from fetus to adult (Kola *et al*., 1993), we examined the expression of *Elf3* during embryogenesis on days 19 and 17, compared to that of the adult (Figure 6b and c). Essentially, the *Elf3* organ expression pattern observed in the adult reflects that observed in fetal organs, at least at the embrogenic stages examined.

ELF3 displays sequence specific binding to oligonucleotides containing consensus ETS sites

The ability of ELF3 to bind specifically to DNA was tested by electrophoretic mobility shift assay (EMSA). ELF3 bound strongly to the E74 oligonucleotide, containing a consensus ETS binding site (Figure 7a, lane 1). Binding to E74ml, a mutant E74 oligonucleotide, was virtually undetectable (lane 2). E74ml contains a GGAA→AGAA substitution in the core of the ETS recognition site, which we and others have previously shown to obliterate the binding of ETS factors (Wasylyk *et al*., 1990; Nye *et al*., 1992; Thomas *et al*., 1995, 1997). ELF3 also binds to a weaker ETS-binding site from the GM-CSF promoter (GM–lane 4), but not to an AGAA mutant of this oligonucleotide (GMml–lane 5). It has been suggested that a GGAT variation of the core ETS binding site is relatively selective for ETS1 and ETS2 binding, in that these proteins bind GGAT-containing sequences much more strongly than other ETS factors (Wang *et al*., 1992; Bosselut *et al*., 1993). Mutant E74 and GM oligonucleotides (E74m2 and GMm2 respectively) contain GGAA→GGAT substitutions. Interestingly, these still retain some binding activity for ELF3 (lanes 3 and 6 respectively). In particular, binding of ELF3 to GMm2 appears to be slightly stronger than that of the wild-type GM (lanes 4 and 6).

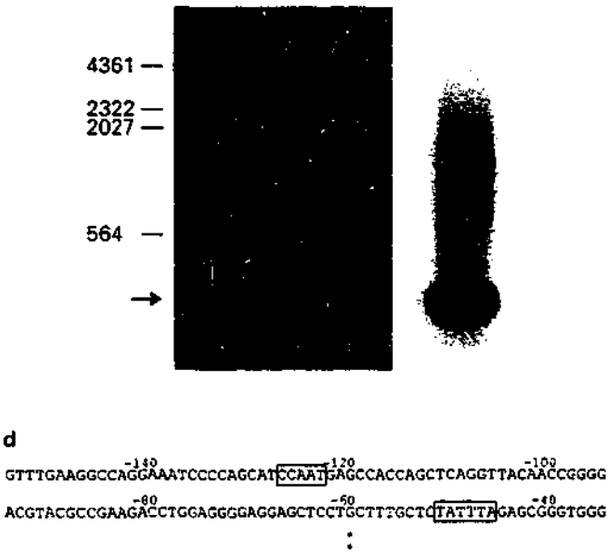
The ability of ELF3 to bind DNA specifically was confirmed by competition experiments, in which ELF3 binding to E74 was competed by addition of excesses of unlabeled oligonucleotides (Figure 7b). Binding of ELF3 to E74 (lane 1) was competed efficiently by 100-fold and 300-fold excesses of itself (lanes 2 and 3), and by E74m2 (lanes 6 and 7). Competition by E74ml was



b

Exon (size) bp	consensus match	sequence at exon-intron junction: 5' splice donor..(intron size, bp)..3' splice acceptor	consensus match	nucleotide position
1 (90)	5/9	TCG/GTACGG..(557)..GTCTCCTCTTCCCAG/C	14/15	90/91
2a (168)	9/9	CAG/GTGAGT..(388)..TACTCCACCTTCCAG/A	13/15	258/259
2b (168)	9/9	CAG/GTGAGT..(328)..CCTCGGGTTGTGTAG/G	12/15	258/259
3a (222)	5/9	TCA/GTAAGG..(129)..CCCTGCCTCCGCCAG/C	13/15	480/581
3b (282)	5/9	TCA/GTAAGG..(129)..CCCTGCCTCCGCCAG/C	13/15	480/581
4 (96)	5/9	TTG/GTGAAA..(188)..AACTTCTTTTCTCAG/A	13/15	576/577
5 (120)	8/9	CAA/GTAAGT..(148)..TTCTTCACCCCTTAG/G	15/15	696/697
6 (90)	8/9	GAG/GTGAGT..(142)..ACCCCCTTCCCATAG/A	14/15	793/794
7 (117)	7/9	ACG/GTAAGC..(853)..AATTTTCCCTTGTAG/C	13/15	902/903
8 (196)	8/9	GAG/GTGAGT..(545)..ACATCTTCCCAGCAG/G	13/15	1099/1100
9 (800)				
Consensus[c]		MAG/GTRAGT.....YYYYYYYYYYNYAG/G		

[a] Fraction of nucleotides conserved relative to Mount consensus sequence
[b] Relative to *Elf3* cDNA, Figure 1b
a *Elf3a* donor and acceptor splice sequence different from that of *Elf3b*
b *Elf3b* donor and acceptor splice sequence different from that of *Elf3a*
[c] M = A or C; R = A or G; Y = C or T; N = A or C or G or T



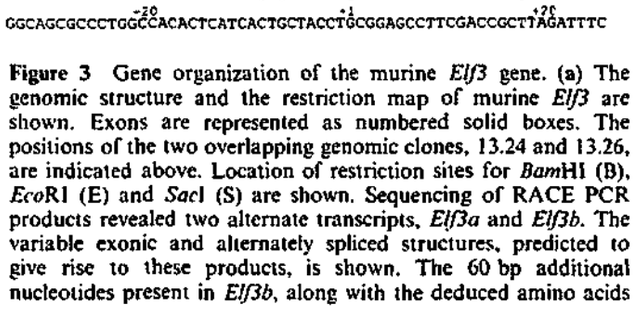


Figure 3 Gene organization of the murine *Elf3* gene. (**a**) The genomic structure and the restriction map of murine *Elf3* are shown. Exons are represented as numbered solid boxes. The positions of the two overlapping genomic clones, 13.24 and 13.26, are indicated above. Location of restriction sites for *Bam*HI (B), *Eco*RI (E) and *Sac*I (S) are shown. Sequencing of RACE PCR products revealed two alternate transcripts, *Elf3a* and *Elf3b*. The variable exonic and alternately spliced structures, predicted to give rise to these products, is shown. The 60 bp additional nucleotides present in *Elf3b*, along with the deduced amino acids (one-letter code), is shown. (**b**) Exon/intron sizes, sequences of exon-intron junctions and their relative genomic positions. Sequences of splice donor and splice acceptor sites are shown, along with their match to consensus sequences. All exon-intron junctions conform to the Breatnach (GT/AG) rule (underlined). (**c**) *Elf3* RACE PCR amplification. RACE products were analysed by agarose gel electrophoresis. The left panel (lanes 1–3) has been ethidium bromide stained, and the right panel probed with *Elf3* cDNA (lanes 4–6). Negative controls containing the anchor primer (AP1, lanes 2 and 5) or the *Elf3*-specific primer (P1, lanes 3 and 6) are shown. A major band representing *Elf3*-specific PCR product was obtained using both primers (lanes 1 and 4, position marked by an arrow). Although other bands were seen on the EtBr stained gel (lane 1), they were not specific since they failed to hybridize with *Elf3*-specific probe (lane 4). (**d**) Mapping of *Elf3* transcriptional initiation sites. The *Elf3* RACE PCR product was subcloned and 7 clones sequenced. The transcriptional initiation site predicted by each clone are marked by an asterisk (*) above the nucleotide sequence of *Elf3*. TATA and CAAT motifs are boxed

less effective both at 100 × excess and 300 × excess (lanes 4 and 5). Thus ELF3 displays sequence specific binding to different consensus ETS binding sequences, and does so with differential affinity. We have previously conducted similar experiments comparing the differential binding of ETS1 and ELF1 to GM and GMm2 (Thomas *et al*., 1995). This analysis confirmed that ETS1, but not ELF1 could bind to the GGAT-core modified oligonucleotide. Thus it appears that ELF3, like ETS1, is capable of binding GGAT more effectively than other ETS family members, such as ELF1, that show little or no binding to these sequences.

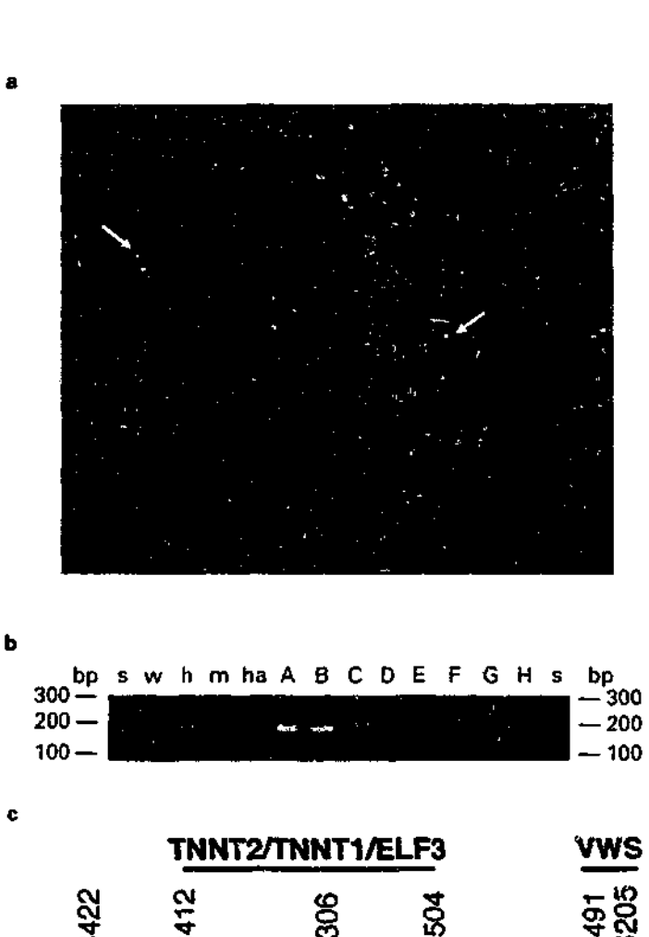


Figure 4 Genomic localization of human *ELF3*. (**a**) Fluorescence *in situ* hybridization (FISH). Computer enhanced image of metaphase preparations, showing *in situ* hybridization of biotinylated probe for *ELF3* hybridization sites are indicated with arrows. (**b**) STS content analysis of the 'Coriell Mapping Panel 2' with *ELF3*-specific primers. PCR products were separated by agarose gel electrophoresis, following amplification with *ELF3*-specific primers using the following genomic templates: no DNA control (w), total human DNA (h), total mouse DNA (m), total hamster DNA (ha) and DNA pools of monochromosome hybrids (A–H). Chromosomes present in pools are as follows: A (1,8,14,19,22,Y), B (1,2,9,15,20,X), C (2,3,8,10,16,21), D (3,4,9,11,14,17), E (5,10,12,15,18,19), F (4,5,6,13,20,22), G (6,7,11,12,16,Y), H (7,13,17,18,21,X). Each chromosome is represented in two pools. The size of DNA standards (s) are indicated. (**c**) Diagram showing localization of *ELF3* within chromosome one, with respect to adjacent markers obtained from mapping data (see text)

Elf3 transactivates the polyomavirus enhancer, containing ETS and AP1 binding motifs

COS cells were co-transfected with CAT reporter and ELF3 expression constructs to determine if ELF3 could function as a transcriptional activator (Figure 8). The CAT reporter gene driven by the thymidine kinase minimal promoter pTKmin was not responsive

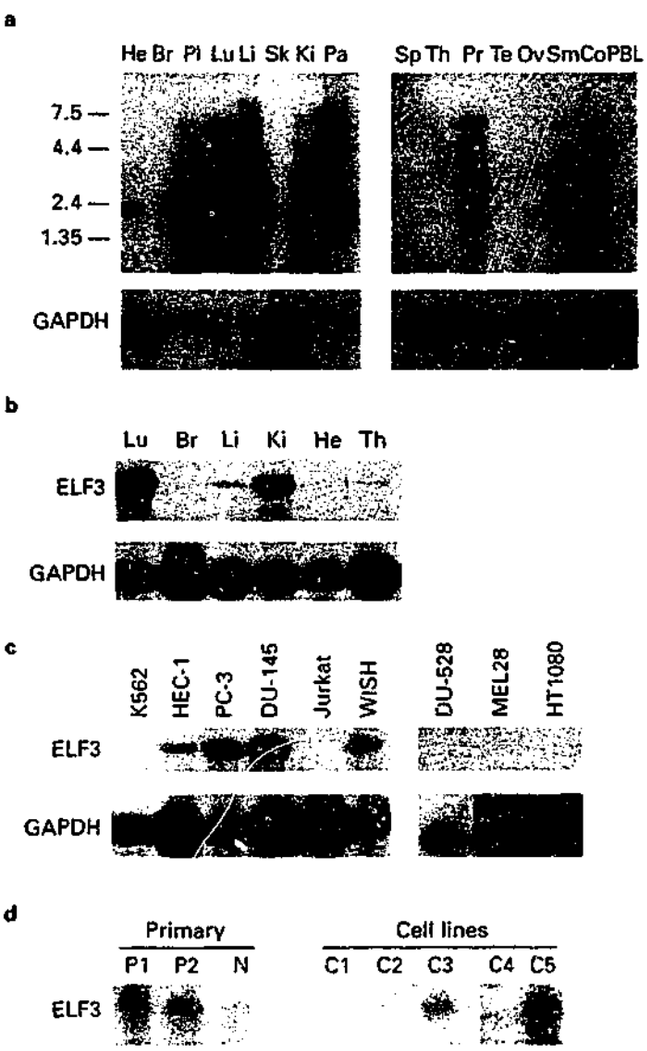


Figure 5 *ELF3* expression in human tissues and cell lines. (**a**) Northern analysis of adult human tissues probed with human *ELF3* cDNA (top panels) and GAPDH cDNA (lower panels). Abbreviations; He: heart; Br: brain; Pl: placenta; Lu: lung; Li: liver; Sk: skeletal muscle; Ki: kidney; Pa: pancreas; Sp: spleen; Th: thymus; Pr: prostate; Te: testis; Ov: ovary; Sm: small intestine; Co: colon mucosa; PBL: peripheral blood lymphocytes. Position and size of RNA markers are shown (kb) T. (**b**) RNase protection analysis of human fetal tissues. *ELF3* and GAPDH protected fragments are shown. (**c**) RNase protection analysis of cells lines; K562 (erythroid); HEC-1 (endometrial carcinoma); PC-3 (prostate carcinoma); DU-145 (prostate carcinoma); Jurkat (T-cell); WISH (amnion carcinoma); DU-528 (macrophage); MEL28 (melanoma), HT1080 (fibrosarcoma). (**d**) RNA protection analysis of lung cancers and cell lines; P1: primary large cell carcinoma; P2: primary adenocarcinoma; N: normal lung; C1: RHSCC11 (small-cell carcinoma); C2: SK-MES2 (squamous cell carcinoma); C3: PL27 (adenocarcinoma); C4: L162 (squamous cell carcinoma); C5: A549 (carcinoma)

to expression of the ETS family members ETS2 or ELF3. The CAT reporter gene driven by five copies of the polyomavirus enhancer upstream of the TK-minimal promoter (p5XpolyTK) had relatively higher basal expression than did pTKmin. Co-transfection with the ELF3 expression construct resulted in a further 2.5-fold transactivation of the p5XpolyTK reporter (this transactivation being statistically significant at $P<0.001$, as compared with the basal p5XpolyTK reporter). This transcription was approximately twofold greater than that achieved with an ETS2 expression construct in these cells (Figure 8, ETS2 transactivation is also statistically significant, compared with the basal p5XpolyTK reporter, $P<0.01$).

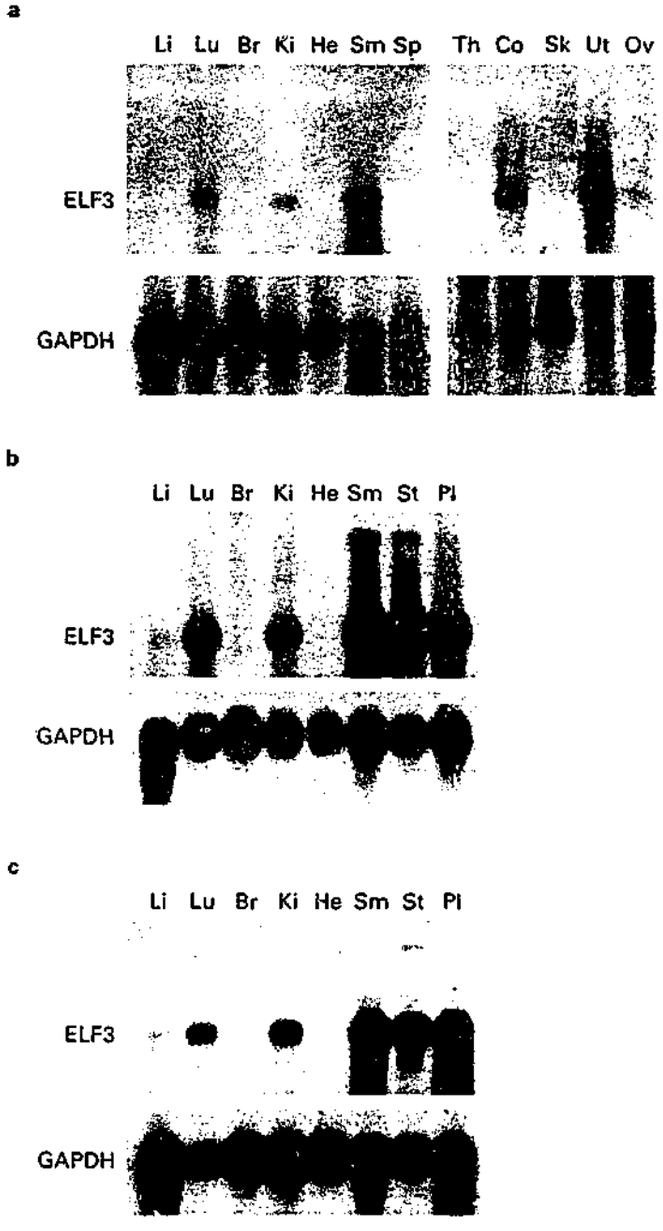


Figure 6 *Elf3* expression in mouse tissues. (a) Northern analysis of adult mouse tissues probed with mouse *Elf3* cDNA (top panels) and GAPDH cDNA (lower panels). Abbreviations; Li: liver; Lu: lung; Br: brain; Ki: kidney; He: heart; Sm: small intestine; Sp: spleen; Th: thymus; Co colon mucosa; Sk: skeletal muscle; Ut: uterus; Ov: ovary. (b) Northern analysis as above, but using RNA from day 19 embryonic tissues. Abbreviations, St: stomach; Pl: placenta. (c) Northern analysis as above, but using RNA from day 17 embryonic tissues

Discussion

ETS family transcription factors play an important role in normal cellular function and in disease. We have isolated a novel ETS gene (*ELF3*) from both human and mouse and additionally report the genomic organization of the murine gene. Interestingly, the mRNA expression of *ELF3*, unlike that of most ETS family members, was absent in hematopoietic cell lines, hematopoietic organs (such as thymus and spleen) and in peripheral blood lymphocytes. Intriguingly, *ELF3* expression is only detected in cell lines of epithelial origin. Further, the gene is expressed in organs such as the lung which contain a high proportion of specialized epithelial cells. These data suggest a role for *ELF3* in epithelial function.

It appears, however, that not all epithelial cells express detectable levels of *ELF3* mRNA, as organs such as ovary, thymus, spleen and brain have a component of specialized epithelial cells. It will be necessary to undertake detailed cellular localization studies to define specifically the cellular distribution of *ELF3*. However, the finding that *ELF3* is expressed specifically in cell lines of epithelial origin is significant. A number of relatively ubiquitously expressed transcription factors such as AP1 (Reddy *et al*., 1995) and AP2 (Hennig *et al*., 1996) have been shown to be important in regulating transcription of epithelial genes. Only a small number of the currently described transcription factors have been shown to be restricted to epithelial cells and most of these are restricted to only a subset of specialized epithelial cells. For example, Skn-1a and Skn-1i-2, two Oct-2-related factors, are restricted to keratinocytes (Anderson *et al*., 1993). TTF-1 (thyroid transcription factor-1) is a homeodomain transcription factor expressed in the thyroid, lung and parts of the brain (Ikeda *et al*., 1995), and HNF-1 (hepatocyte nuclear factor-1) is a variant homeodomain protein expressed in liver (Kuo *et al*., 1991). Thus, the identification of an ETS family member expressed specifically or predominantly in epithelial cells may be relevant to normal function and to tumorigenic processes in these cells. Identification of epithelial expressed genes with cellular regulatory functions is of particular interest to the study of tumorigenesis, since the majority of solid tumors are of epithelial origin. In this context, it is also significant that we have identified *ELF3* as being overexpressed in lung cancer cell lines and in primary lung adenocarcinomas and carcinomas.

The *ELF3* gene was mapped to human chromosome 1q32 by FISH and the assignment of *ELF3* to chromosome 1 was verified by STS content analysis with the Coriell Mapping Panel 2. This region of chromosome 1 contains the gene responsible for Van der Woude Syndrome, a dominant craniofacial disorder. However, when the location of the *ELF3* locus was refined with two radiation hybrid panels, it mapped between the markers D1S412 and D1S504 which are about 10 cM proximal to the Van der Woude critical region, ruling out *ELF3* as a candidate for the Van der Woude Syndrome gene.

The localization of *ELF3* is significant, and this together with the epithelial expression pattern described in this paper and the known oncogenic properties of members of the ETS family, make

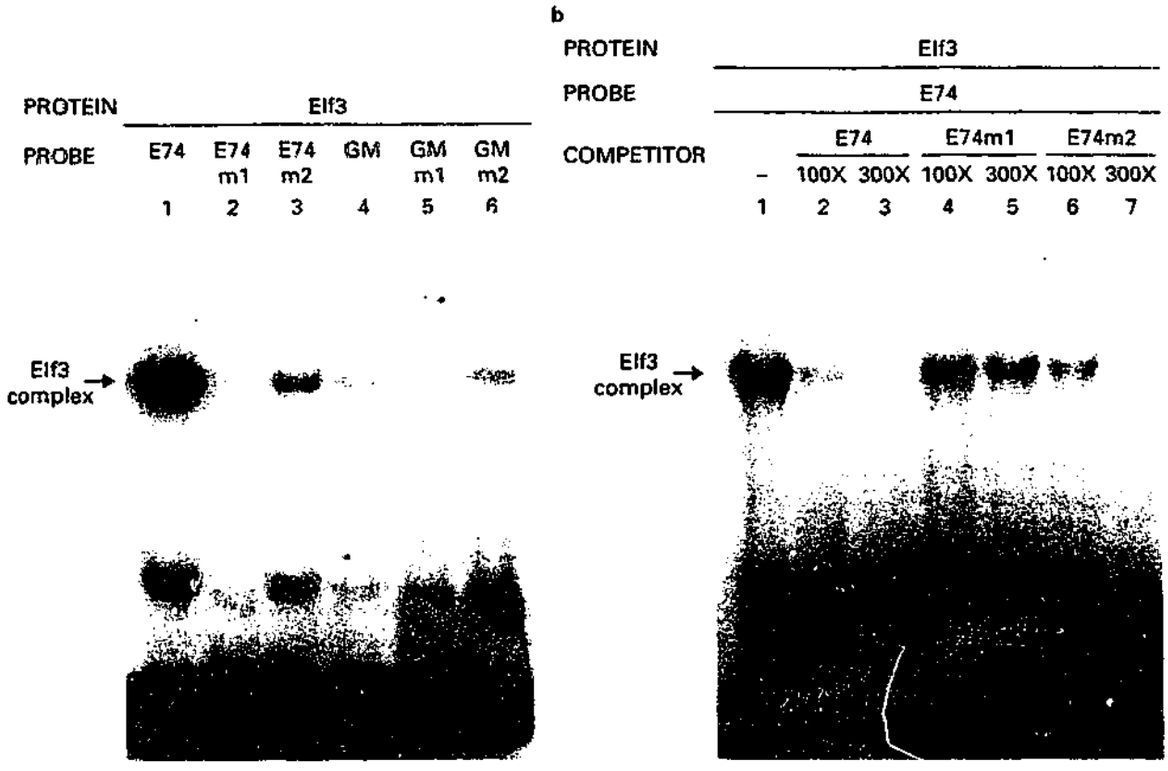


Figure 7 ELF3 binding to consensus ETS binding sequences. (a) Specific DNA binding of *E. coli* produced ELF3 was analysed by electrophoretic mobility shift assay (EMSA), using labeled double stranded oligonucleotides as probes. E74 contains a consensus binding site for the ETS family member E74A. E74m1 and E74m2 are mutant oligonucleotides based on E74, but with the core GGAA replaced with AGAA and GGAT, respectively. GM contains an ETS binding site from the human GM-CSF promoter. GMm1 and GMm2 contain the AGAA and GGAT mutations, respectively. ELF3-DNA complexes are marked. (b) Competition of ELF3 binding to the E74 probe by an excess of unlabeled oligonucleotide. Competitor oligonucleotides (E74, E74m1, E74m2) were added in either 100-fold (100×) or 300-foid (300×) excess, relative to the amount of probe

ELF3 a candidate gene for involvement in epithelial cancers. Cytogenetic markers involving 1q are the most frequently observed karyotypic changes in breast cancer. Loss of heterozygosity with allelic loss of q23-q32 has been commonly observed (Chen *et al.*, 1989). The common region, shown to be deleted in this study, was between the renin gene and marker AT3. This is the same region identified by our *ELF3* localization. On the other hand, other data associates over-expression of 1q localized gene(s) with cancer. Chromosomal abnormalities in 1q are common in a wide range of solid tumors including carcinomas (Brito-Babapulle and Atkin, 1981; Douglass *et al.*, 1985) and some leukemias (Oshimura *et al.*, 1976). Trisomy of 1q or a portion of 1q has been shown to be frequently present in carcinomas including breast carcinoma (Pandis *et al.*, 1992; Ried *et al.*, 1995), and large bowel adenocarcinomas (Reichmann *et al.*, 1984). This has led to the speculation that over-expression of a gene or genes between 1q24 and 1q32 contributes to tumor progression in carcinomas (Chromosome 1 workshop 1995). 1q32 is a fragile site and translocations involving 1q32 have been shown to be associated with bladder carcinoma (Barrios *et al.*, 1990) and ovarian cancers (Bello and Rey, 1990) and also some acute myelomonocytic leukemias (Yip *et al.*, 1991). Thus, it is possible to speculate, on the basis of these data, that aberant over-expression of *ELF3* in human cancers may result in the transcriptional upregulation of genes important in tumor progression, especially since we have shown *ELF3* over-expression in some cancers and cancer cell lines. For example ETS family members have been implicated in the transcriptional regulation of c-*met* (Gambarotta *et al.*, 1996) and *erb*-B2 (Scott *et al.*, 1994b), which are expressed in epithelial cells and are thought to play a role in carcinoma tumor progression.

Protein sequence homology and conservation between various ETS family members has proven useful in identifying functional domains. For this reason, the structure of the ELF3 protein, in relation to other ETS family members, is of interest and involves the following points. Firstly, the amino acid sequence of ELF3 is quite divergent from other ETS family members, and secondly, in addition to the ETS-DNA binding domain at the C-terminal end of the protein, ELF3 shows some homology with the N-terminal region of a number of ETS members. This N-terminal domain found in a number of ETS family members has been termed the 'A-domain' (Seth and Papas, 1990) and more recently it has been shown to be important for dimerization by TEL, and called the TEL domain (Jousset *et al.*, 1997). This domain may play a critical role in transactivation by specifying protein-protein interactions., Human ELF3 also has a consensus PEST sequence which is thought to be important for protein stability. Amongst the ETS family members, PU.1 also has this structural motif (Klemsz *et al.*, 1990), but its significance in protein stability will need to be assessed.

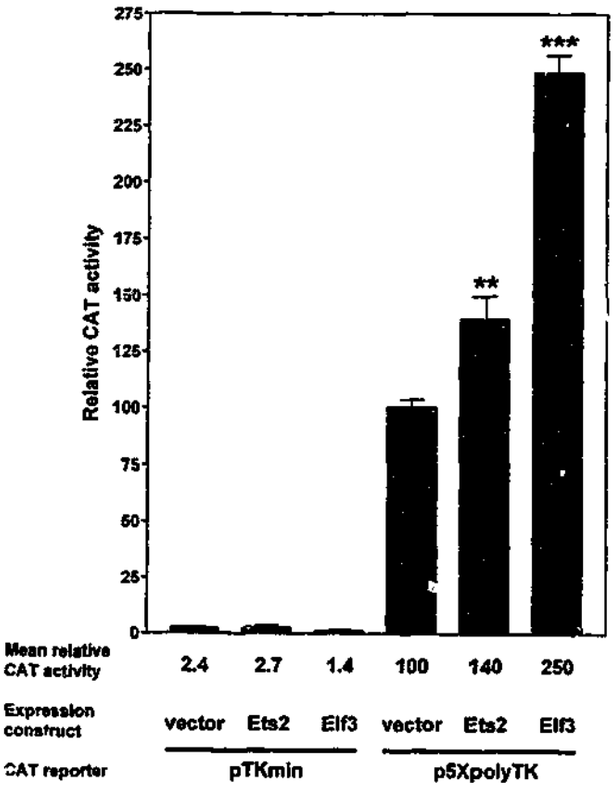


Figure 8 Transactivation by ELF3. COS cells were co-transfected with CAT reporter and transcription factor exspression constructs. Transcription of the CAT gene was driven by either pTKmin (minimal thymidine kinase promoter) or p5XpolyTK (five copies of the polyomavirus enhancer inserted upstream of TKmin). ETS factor expression constructs used were ETS2 (pBOSETS2) or ELF3 (pBOSELF3). In the absence of expression construct the equivalent amount of base vector (pEFBOS) was co-transfected. COS cells were processed for CAT assays and the results of at least four replicates are shown as the mean relative CAT activity with standard error of the mean (s.e.m.) bars. The mean relative CAT activity for each experiment is also given numerically. Statistically significant results are indicated by asterisks. Double asterisks indicate high significance ($P<0.01$) and triple asterisks indicate extremely high significance ($P<0.001$)

The ETS domain of ELF3 putatively involved in DNA binding shows an amino acid identity of only 32–49% to a number of ETS family members (Figure 2b). However the critical residues shown to be important for DNA contacts in the crystal structures appear to be conserved (Donaldson *et al.*, 1996; Kodandapani *et al.*, 1996). The closest ELF3 amino acid identity is observed to Drosophila E74A (49%) and human ELF1 (43%).

Using recombinant ELF3 we have shown that it binds to an ETS consensus oligonucleotide sequence containing GGAA and GGAT core nucleotides. Furthermore, we have shown that ELF3 can bind to a lower affinity binding site such as that from the human GM-CSF promoter, which is ETS1 regulated in T cells. (Thomas *et al.*, 1995; Thomas *et al.*, 1997) The binding of ELF3 to sequences containing a GGAT core sequence may be especially significant, since it has been suggested that this variation of the GGAA core may exclude or greatly reduce the binding of ETS family members other than ETS1 (Wang *et al.*, 1992). The ability of ETS1 to bind GGAT core sequences appears to be associated with a critical lysine residue (amino acid 388) in its DNA binding domain (Bosselut *et al.*, 1993), an amino acid missing in ETS family members, such as ELF1 and E74A but present in ELF3. In this study, GGAA-only core specificity could be induced in ETS1 by mutation of this lysine. Reciprocally, GGAT-permissive binding could be provoked in ELF1 by mutation of the corresponding site to a lysine. Several ETS family members that contain the GGAT binding-associated lysine residue, similarly to ETS1, can also bind a GGAT core, but do so only in particular circumstances. Hence ELK1 can bind an ETS site with a GGAT core, but only in association with adjacently bound serum response factor (SRF) (Rao and Reddy, 1992; Treisman *et al.*, 1992; Shore and Sharrocks, 1995). ELK1 related SAP1 binds weakly to GGAT core binding sites, but both ELK1 and SAP1 require ACC residues flanking one side of the core (positions −5, −4 and −3, in the ETS1 consensus) and prefer GT residues flanking the other side (positions +3 and +4) to achieve binding to GGAT containing sites. It is interesting to note that these specific flanking sequences, as present in the E74 oligonucleotide, confer very high affinity binding for ETS1 (Nye *et al.*, 1992; Woods *et al.*, 1992). FLI1 also contains the GGAT binding-associated lysine residue, but appears to bind GGAT sequences very rarely (1 in 22 binding sites tested) (Mao *et al.*, 1994), and again only in the context of the highest affinity flanking sequences. The situation of ETS1 binding to GGAT sequences is different, where a wide range of flanking sequences are permissible (Nye *et al.*, 1992; Woods *et al.*, 1992). Thus, ETS1 can bind GGAT core sequences, even when the flanking sequences specify low affinity binding, and other ETS family members appear to bind GGAT core sequences: (1) only in the context of required flanking sequences specifying high affinity; (2) only in conjunction with other factors; or (3) not at all. In the experiments presented here, ELF3 appears to display ETS1-like binding characteristics in contrast to other ETS family members such as ELF1. Thus, on a functional basis, ELF3 could also be classified in the ETS1 subfamily of ETS genes, although its DNA-binding domain sequence alone place it in the E74/ELF1 subfamily.

ETS family members containing the 'A domain' (Seth and Papas, 1990) such as ETS1, ETS2 and ERG2 are able to transactivate reporter constructs containing multimerized Ets-binding sites from the polyoma enhancer. ELF3 is able to transactivate a polyoma-TK minimal promoter construct (Figure 8). Even though the level of transactivation in COS cells is modest, it is statistically significant and greater than that achieved by ETS2. Furthermore, at this stage no data is available on phosphorylation of ELF3 which may influence transactivation. Recent data has shown that ETS1 and ETS2 require phosphorylation of a specific residue in the N-terminal domain by the MAP kinase pathway for full transactivational activity (Yang *et al.*, 1996). This phosphorylation site found in ETS1 and ETS2 is conserved with Drosophila ETS family member *pointed*, which plays a critical role in photoreceptor downstream of the tyrosine kinase receptor sevenless (O'Neill *et al.*, 1994). ELF3 does not appear to have the *pointed* phosphorylation domain but does contain a number of consensus phosphorylation sites for

other protein kinases which may regulate ELF3 function.

In conclusion, this manuscript reports murine and human sequences of a novel ETS gene that is specifically expressed in epithelial cell lines. It maps to a region that is implicated in epithelial cancers and thus future experimentation will be aimed at studying its functional role in normal epithelial development/function and in epithelial cancers such as lung, breast and prostate.

Materials and methods

Cell lines and culture

Monkey COS7 cells were grown in DMEM supplemented with 10% fetal calf serum, 100 U/ml penicillin and 100 μg/ml streptomycin, and maintained in a humidified incubator at 5% CO_2 and 37°C. Cell lines PL27 and L162 (Lukeis *et al.*, 1993) were obtained from Dr S Hasthorpe (Royal Childrens Hospital, Melbourne) and maintained as for COS cells. Other cell lines were obtained from ATCC.

Plasmids

The human *ETS1* cDNA, encoding the p42 isoform, was cut from pSG42 with *Bgl*II and blunt-ended with T4 polymerase. The mouse *ets2* cDNA was cut from pGEM7-ME2A with *Bam*HI and blunt-ended with T4 polymerase. The human *ELF1* cDNA was cut from pSK-ELF1 (the gift of Dr J Leiden, University of Chicago, Illinois 60637, USA) with *Eco*RI and *Xho*I, and blunt-ended with T4 polymerase. *ETS1*, *ETS2*, *ELF1* and *ELF3* cDNAs were ligated into *Xba*I cut, T4 polymerase blunted, pEFBOS mammalian expression vector (Mizushima and Nagata, 1990), to give pBOS-ETS1, pBOS-ETS2, pBOS-ELF1 and pBOS-ELF3, respectively. pX5polyTK was made by cloning multimerized polyoma enhancer oligonucleotides into the *Bam*HI site of pBLCAT2. pHis6-ELF3 expression vector was made as follows: The human *ELF3* cDNA was amplified using PCR oligonucleotide primers (5'-CGGGATCCGCTGCAACCTGTGAGATTAGC-3' and 5'-GCAGATCTCAGTTCCGACTCTGGAGAACC-3') containing a *Bam*HI restriction site which encodes codon 2 followed by 21 nucleotides of the human *ELF3* coding sequence and a *Bgl*II restriction site followed by 21 nucleotides complementary to the last 22 nucleotides of the *ELF3* coding sequence. The *Bam*HI–*Bgl*II restriction fragment was cloned into the *Bam*HI site of the pQE-30 (Qiagen, Inc. Chatsworth, CA) bacterial expression vector resulting in a N-terminal fusion of ELF3 protein to six histidine residues (His-Tag).

Electrophoretic mobility shift assays (EMSA)

Purified recombinant ELF3 protein was produced in *E. coli* using pHis6-ELF3. An overnight culture was diluted 1/10 in LB broth and grown for 1 h at 37°C. Recombinant protein expression was induced by addition of 0.1 mM IPTG and culture of cells for 2 h at 25°C. Cells were harvested and lysed by sonication in PBS and debris removed by centrifugation. One ml of Talon matrix (Clontech, Palo Alto, CA) was incubated with the supernatant for 30 min and the beads collected and washed with TEN buffer (20 mM Tris-HCl, 50 mM NaCl, pH 8.0). ELF3 was eluted from the beads in TEN buffer supplemented with 100 mM imidazole. Analysis of the preparation by standard denaturing SDS-polyacrylamide gel electrophoresis revealed that the preparation contained approximately 90% monomer ELF3, having the expected molecular weight of 42 kD.

Recombinant ETS1 and ELF1 were produced by transient transfection of the respective expression vectors into COS cells. Sub-confluent cells were trypsinized, washed in growth medium and resuspended in PBS with 20 mM HEPES at 5×10^6 cells/ml. 300 μl of cells were mixed with 20 μg DNA in 0.4 cm gap electroporation cuvettes and pulsed at 170 V and 500 μF (Biorad Gene Pulsar). Cells were re-plated into 10 cm petri dishes and harvested 24 h later. Nuclear lysates from COS cells were prepared as previously described (Thomas *et al.*, 1995). DNA binding experiments with nuclear lysates were performed using EMSA, as previously described (Thomas *et al.*, 1995). Briefly, purified double stranded oligonucleotides were labeled with α-^{32}P-dATP by Klenow fill-in reaction. 10 ng oligonucleotide probe was incubated for 10 min with approximately 0.1 μg purified ELF3 protein or 2 μg COS nuclear extract in DNA binding buffer (1 mM EDTA, 10 mM Tris-HCl pH 8.0, 50 mM NaCl, 5 mM $MgCl_2$, 3 mM dithiothreitol (DTT), 1 mg/ml BSA, 500 ng/ml poly-d(I-C).d(I-C), 500 ng/ml poly dI-dC, 200 ng/ml sheared salmon sperm DNA), ± 1 μg/3 μg unlabeled competitor oligonucleotides, in 10 μl final volume. Assays were run through non-denaturing, 7% acrylamide (29 acrylamide:1 bis-acrylamide), 0.5 × TBE gels at 4°C.

Oligonucleotide sequences used (shown in double stranded conformation):

E74	gatcATAACCGGAAGTAACT TATTGGCCTTCCATTGActag
E74m1	gatcATAACCAGAAGTAACT TATTGGTCTTCATTGActag
E74m2	gatcATAACCGGATGTAACT TATTGGCCTACATTGActag
GM	gatcCACAGAGGAAATGATT GTGTCTCCTTTACTAActag
GMm1	gatcCACAGAAGAAATGATT GTGTCTTCTTTACTAActag
GMm2	gatcCACAGAGGATATGATT GTGTCTCCTATACTAActag

Transfection of COS cells and chloramphenicol acetyltransferase (CAT) assays

COS7 cells were transfected with 5 μg CAT reporter plasmid and 10 μg of expression constructs by electroporation, as described above. Cells were harvested 48 h later and processed for CAT assays as previously described (Thomas *et al.*, 1995). All raw CAT data was normalized to the expression of p5XpolyTK, which was given the value of '100'. The mean and standard error of the mean (s.e.m.) were generated from four replicates of each experiment. Data was subject to statistical analysis using unpaired two-tailed *t*-tests, with resultant *P* values less than 0.05 considered significant.

Isolation and characterization of full-length human and murine ELF3/Elf3 cDNAs

An EST from prostate cDNA library was used to screen a fetal lung cDNA library (Clontech HL3022b). A number of cDNA clones were obtained but none encoded the full-length transcript. Additional 5-prime sequence was obtained by RT–PCR using a Marathon cDNA synthesis kit (Clontech) and RACE (Rapid Amplification of cDNA Ends) using human fetal lung RNA. The murine *Elf3* cDNA was isolated from an adult lung cDNA library in Lambda ZAPII (Clontech). Additional 5-prime sequence was obtained by RACE using murine adult lung RNA (see below). All cDNA sequences were confirmed by sequencing both strands at least once.

Determination of the murine Elf3 transcriptional initiation site

An adapter-ligated cDNA library was made with 2 μg of murine adult lung poly(A)$^+$ RNA using a Marathon cDNA synthesis kit, according to the manufacturers instructions (Clontech). RACE amplification was carried out using the supplied anchor primer AP1 together with the *Elf3*-specific primer P1 (5'-GGGTCAACAC-CAAGTCTTCAGTGCCAA-3') close to the 5' end of the cDNA. A second round of amplification was then carried out using the nested anchor primer AP2 and the nested *Elf3* primer P2 (5'-GCTGTACATGGCGTTGAAG-TAGTTAC-3'). PCR products were analysed on a 2.0% agarose gel, Southern blotted and probed with a 1.6 kb murine *Elf3* cDNA which contained the poly(A)$^+$ tail (obtained from library screening). The *Elf3*-specific PCR product was cloned into pGEM-T vector (Promega Corp., Madison, WI) and DNA sequencing carried out.

Isolation and characterization of the full murine Elf3 genomic sequence

Genomic DNA clones were isolated by screening a murine 129SVJ library in the lambda FIXII vector (Stratagene). Genomic sequences were subcloned for sequencing by either direct subcloning of restriction fragments into appropriate plasmid vectors or by cloning of PCR-generated fragments into pGEM-T plasmid vector. PCR reactions were performed in a Perkin Elmer 9600 thermocycler with 1 μL of phage stock in reaction buffer containing 1.5 mM $MgCl_2$ and 0.2 U/μL Taq DNA polymerase (Boehringer Mannheim, Indianapolis, IN) with a 30 cycle amplification of 94°C (30 s), 50°C (30 s), 72°C (from 1–8 min).

FISH

Human chromosomal localization by FISH was as previously described (Callen *et al*., 1990). A 900 bp *ELF3* cDNA probe was nick translated with biotin-^{14}-dATP and hybridized independently to metaphases from two normal males.

STS content mapping

PCR primers specific for *ELF3* were designed with the computer program Primer v0.5. The 5' nucleotide for the forward (5'-CTCCAGCACCTTCTTTCTGG-3') and reverse (5'-TGGAACAGTCCAGCACTCTG-3') primers were 1622 and 1803, respectively. The PCR reactions were performed in standard buffer (10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl) with 25 ng of genomic template DNA, 1 μM of each primer, 200 μM of each dNTP and 0.25 U of Taq DNA polymerase in a total volume of 10 μL. PCR parameters were an initial denaturation step at 94°C for 2 min, followed by 35 cycles of 94°C (30 s), 55°C (30 s), 72°C (30 s). For PCR experiments with the Coriell Mapping Panel 2 (Coriell Institute, Camden NJ), the DNA from the 24 monochromosome hybrids were pooled as described in Figure 4b in order to reduce the number of reactions. For PCR experiments with the Genebridge 4 and Stanford G3 Radiation Hybrid Panels (Research Genetics; Huntsville, AL), reactions were performed separately for each of the individual hybrids. The PCR results from the 93 hybrids for the Genebridge 4 panel were submitted to the Radiation Hybrid Mapper server at Whitehead Institute/MIT Center for Genome Research (http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl). The PCR results from the 82 hybrids for the Stanford G3 panel were submitted to the Radiation Hybrid Mapper server at the Stanford Human Genome Center (rhserver@toolkit.Stanford.edu). All STS content mapping experiments were performed in duplicate and included PCR reactions with no DNA, total human DNA and total mouse DNA as controls.

RNase protection analysis

ELF3 mRNA abundance in total RNA from human tissues and cell lines was determined as described previously (Tymms, 1995). Antisense RNA probes for human *ELF3* and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) transcribed from linearized plasmid vectors generated full-length probes of 451 bp and 216 bp respectively. The protected products generated by hybridization and RNase digestion are 373 bp for *ELF3* and 160 bp for GAPDH.

Northern analysis

Northern analysis of *ELF3* expression in human adult organs was performed with commercially available blots containing 2 μg of poly(A)$^+$RNA (Clontech). For other Northern blots poly(A)$^+$RNA was isolated by a modification of Gonda *et al*. (1992). Random-primed probes using a 700 bp *Bam*HI-*Bgl*II human *ELF3* cDNA fragment and a 1.6 kb murine *Elf3* fragment containing the poly(A)$^+$ tail (obtained from cDNA library screening) were generated and hybridizations performed using standard procedures. Blots were reprobed with GAPDH cDNA to verify RNA loading.

Note added in proof

After this manuscript was submitted a manuscript describing the human *ELF3* cDNA appeared (Chang *et al*, *Oncogene*, **14**, 1617–1622) with the gene name *ESX*.

Acknowledgements

The authors wish to thank Lars Jermiin for assistance in computer analysis and Margaret I Malik and Leigh H McKinlay for expert technical assistance. This work was supported in part by grants from the NIH P30-HD-27748 and 2P50-DE-09170 and the National Health and Medical Research Council of Australia. Full-length *ELF3* cDNA clone deposited with ATCC in September/October 1996. Genomic localization and gene symbol assigned 10th May 1996 (GDB Accession ID:1230388). GenBank Accessions: Human *ELF3*: AF 016295; murine *Elf3*: AF 016294.

References

Adachi J and Hasegawa M. (1996). MOLPHY Version 2.3: *Programs for molecular phylogenetics based on maximum likelihood, Computer Science Monographs*. The Institute of Statistical Mathematics, Tokyo.

Anderson B, Schonemann MD, Flynn SE, Pearse RV 2nd, Singh H and Rosenfeld MG. (1993). *Science*, **260**, 78–82.

Barrios L, Miro R, Caballin MR, Fuster C, Guedea F, Subias A and Egozcue J. (1990). *Cancer Genet. Cytogenet.*, **49**, 107–111.

Bello MJ and Rey JA. (1990). *Int. J. Cancer*, **45**, 50–54.

Ben-David Y, Giddens EB, Letwin K and Bernstein A. (1991). *Genes Dev.*, **5**, 908–918.

Bhat NK and Thompson CB. (1990). *Proc. Natl. Acad. Sci. USA*, **87**, 3723–3727.
Bhat NK, Fisher RJ, Jujiwara S, Ascione R and Papas TS. (1987). *Proc. Natl. Acad. Sci. USA*, **84**, 3161–3165.
Bhat NK, Komschlies KL, Fukiwara S, Fisher RJ, Mathieson BJ, Gregorio TA, Young HA, Kasik JW, Ozato K and Papas TS. (1989). *J. Immunol.*, **142**, 672–678.
Bosselut R, Levin J, Adjadj E and Ghysdael J. (1993). *Nucleic Acids Res.*, **21**, 5184–5191.
Breatnach R, Benoist C, O'Hare K, Gannon F and Chambon P. (1978). *Proc. Natl. Acad. Sci. USA*, **75**, 4853–4857.
Brito-Babapulle V and Atkin NB. (1981). *Cancer Genet. Cytogenet.*, **4**, 215–225.
Brown TA and McKnight SL. (1992). *Genes Dev.*, **6**, 2502–2512.
Burtis KC, Thummel CS, Jones CW, Karim FD and Hogness DS. (1990). *Cell*, **61**, 85–99.
Callen DF, Baker E, Eyre HJ, Chernos JE, Bell JA and Sutherland GR. (1990). *Ann. Genet.*, **33**, 219–221.
Chen J-H. (1990). *Oncogene Res.*, **5**, 277–285.
Chen LC, Dollbaum C and Smith HS. (1989). *Proc. Natl. Acad. Sci. USA*, **86**, 7204–7207.
Chen T, Bunting M, Karim FD and Thummel CS. (1992). *Dev. Biol.*, **151**, 176–191.
Dalton S and Treisman R. (1992). *Cell*, **68**, 597–612.
Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E *et al.* (1996). *Nature*, **380**, 152–154.
Donaldson LW, Peterson JM, Graves BJ and McIntosh LP. (1996). *EMBO J.*, **15**, 125–134.
Douglass EC, Green AA, Hayes FA, Etcubanas E, Horowitz I and Wilimas JA. (1985). *J. Natl. Cancer. Inst.*, **75**, 51–54.
Eyre HJ, Akkari PA, Meredith C, Wilton SD, Callen DC, Kedes L and Laing NG. (1993). *Cytogenet. Cell Genet.*, **62**, 181–182.
Fletcher JC, Burtis KC, Hogness DS and Thummel CS. (1995). *Development*, **121**, 1411–1421.
Gambarotta G, Boccaccio C, Giordano S, Ando M, Stella MC and Comoglio PM. (1996). *Oncogene*, **13**, 1911–1917.
Golub TR, Barker GF, Bohlander SK, Hiebert SW, Ward DC, Bray-Ward P, Morgan E, Raimondi SC, Rowley JD and Gilliland DG. (1995). *Proc. Natl. Acad. Sci. USA*, **92**, 4917–4921.
Golub TR, Barker GF, Lovett M and Gilliland DG. (1994). *Cell*, **77**, 307–316.
Gonda TJ, Sheiness DK and Bishop JM. (1992). *Mol. Cell Biol.*, **2**, 617–624.
Hart AH, Corrick CM, Tymms MJ, Hertzog PJ and Kola I. (1995). *Oncogene*, **10**, 1423–1430.
Hennig G, Lowrick O, Birchmeier W and Behrens J. (1996). *J. Biol. Chem.*, **271**, 595–602.
Ho IC, Bhat NK, Gottschalk LR, Lindsten T, Thompson CB, Papas TS and Leiden JM. (1990). *Science*, **250**, 814–818.
Hromas R and Klemsz M. (1994). *Int. J. Hematol.*, **59**, 257–265.
Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonim DK, Baptista R, Kruglyak L, Xu SH, *et al.* (1995). *Science*, **270**, 1945–1954.
Ida K, Kobayashi S, Taki T, Hanada R, Bessho F, Yamamori S, Sugimato T, Hoki M and Hayashi Y. (1995). *Int. J. Cancer*, **63**, 500–504.
Ikeda K, Clark JC, Shaw-White JR, Stahlman MT, Boutell CJ and Whitsett JA. (1995). *J. Biol. Chem.*, **270**, 8108–8114.
Jones DT, Taylor WR and Thornton TJ. (1992). *Computer Appl. Biosci.*, **8**, 275–282.
Jousset C, Carron C, Boureux A, Tran Quang C, Oury C, Dusanter-Fourt I, Charon M, Levin J, Bernard D and Ghysdael J. (1997). *EMBO J.*, **16**, 69–82.
Klemsz M, Hromas R, Raskind W, Bruno E and Hoffman R. (1994). *Genomics*, **20**, 291–294.
Klemsz M, McKercher SR, Celada A, Van Beveren C and Maki RA. (1990). *Cell*, **61**, 113–124.
Kodandapani R, Pio F, Ni C-Z, Piccialli G, Klemsz M, McKercher S, Maki RA and Ely KR. (1996). *Nature*, **380**, 456–459.
Kola I, Brookes S, Green AR, Garber R, Tymms MJ, Papas TS and Seth A. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 7588–7592.
Kuo CJ, Mendel BB, Hansen LP and Crabtree GR. (1991). *EMBO J.*, **10**, 2231–2236.
Lai Z-C and Rubin GM. (1992). *Cell*, **70**, 609–620.
LaMarco K, Thompson CC, Byers BP, Walton EM and McKnight SL. (1991). *Science*, **253**, 789–792.
Leiden JM, Wang C-Y, Petryniak B, Markovitz DM, Nabel GJ and Thompson CB. (1992). *J. Virol.*, **66**, 5890–5897.
Lukeis R, Ball D, Irving L, Garson OM and Hasthorpe S. (1993). *Genes Chromosom. Cancer*, **8**, 262–269.
Lopez MP, Oettgen P, Akbarali Y, Dendorfer U and Libermann TA. (1994). *Mol. Cell Biol.*, **14**, 3292–3309.
Mao X, Miesfeldt S, Yang H, Leiden JM and Thompson CB. (1994). *J. Biol. Chem.*, **269**, 18216–18222.
Mesnard L, logeart D, Taviaux S, Diriong S, Mercadier JJ and Samson F. (1995). *Circ. Res.*, **76**, 687–692.
Mizushima S and Nagata S. (1990). *Nucleic Acids Res.*, **18**, 5322.
Monte D, Baert JL, Defossez PA, de Launoit Y and Stehelin D. (1994). *Oncogene*, **9**, 1397–1406.
Monte D, Coutte L, Baert J-L, Angeli I, Stehelin D and de Launoit Y. (1995). *Oncogene*, **11**, 771–779.
Moreau-Gachelin F, Wendling F, Molina T, Denis N, Titenx M, Grimber G, Briand P, Vainchenker W and Tavitian A. (1996). *Mol. Cell Biol.*, **16**, 2453–2463.
Mount SM. (1982). *Nucleic Acids Res.*, **10**, 459–471.
Muthusamy N, Barton K and Leiden JM. (1995). *Nature*, **377**, 639–642.
Nunn MF, Seeburg PH, Moscovici C and Duesberg PH. (1983). *Nature*, **306**, 391–395.
Nye JA, Petersen JM, Gunther CV, Jonsen MD and Graves BJ. (1992). *Genes Dev.*, **6**, 975–990.
O'Neill EM, Rebay I, Tjian R and Rubin GM. (1994). *Cell*, **78**, 137–147.
Oettgen P, Akbarali Y, Boltax J, Best J, Kunsch C and Libermann TA. (1996). *Mol. Cell Biol.*, **16**, 5091–5106.
Oka HH, Shiozaki K, Kobayashi M, Inoue H, Tahara T, Kobayashi Y, Takatsuka T, Matsuyoshi N, Hirano S, Takeichi M *et al.* (1993). *Cancer Res.*, **53**, 5104–5109.
Oshimura M, Sonta S and Sandberg AA. (1976). *J. Natl. Cancer Inst.*, **56**, 183–184.
Pandis N, Heim S, Bardi G, Idvall I, Mandahl N and Mitelman F. (1992). *Cancer*, **5**, 235–238.
Rao VN and Reddy ES. (1992). *Oncogene*, **7**, 2335–2340.
Rao VN, Huebner K, Isobebe M, Ar-Rushdi A, Croce CM and Reddy ES. (1989). *Science*, **244**, 66–70.
Rao VN, Papas TS and Reddy ES. (1987). *Science*, **237**, 635–639.
Ray D, Bosselut R, Ghysdael J, Mattei MG, Tavitian A and Moreau GF. (1992). *Mol. Cell Biol.*, **12**, 4297–4304.
Reddy SPM, Chuu YJ, Lao PN, Donn J, Ann DK and Wu R. (1995). *J. Biol. Chem.*, **270**, 26451–26459.
Reichmann A, Martin P and Levin B. (1984). *Cancer Genet. Cytogenet.*, **12**, 295–301.
Ried T, Just KE, Holtgreve-Grez H, du Manoir S, Speicher MR, Schrock E, Latham C, Blegen H, Zetterberg A, Cremer T *et al.* (1995). *Cancer Res.*, **55**, 5415–5423.
Schulz RA, Hogue DA and The SM. (1993). *Oncogene*, **8**, 3369–3374.
Scott EW, Simon MC, Anastasi J and Singh H. (1994a). *Science*, **265**, 1573–1577.
Scott GK, Daniel JC, Xiong X, Maki RA, Kabat D and Benz CC. (1994b). *J. Biol. Chem.*, **269**, 19848–19858.

Seth A and Papas TS. (1990). *Oncogene*, **5**, 1761–1767.
Seth A, Watson DK, Blair DG and Papas TS. (1989). *Proc. Natl. Acad. Sci. USA*, **8**, 7833–7837.
Shimizu K, Ichikawa H, Tojo A, Kaneko Y, Maseki N, Hayashi Y, Ohira M, Asano S. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 10280–10284.
Shore P and Sharrocks AD. (1995). *Nucleic Acids Res.*, **23**, 4698–4706.
Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ and Denny CT. (1994). *Natl. Genet.*, **6**, 146–151.
Sumarsono SH, Wilson TJ, Tymms MJ, Venter DJ, Corrick CM, Kola R, Lahoud MH, Papas TS, Seth A and Kola I. (1996). *Nature*, **379**, 534–537.
The SM, Xie X, Smyth F, Papas TS, Watson DK and Schulz RA. (1992). *Oncogene*, **7**, 2471–2478.
Thomas RS, Tymms MJ, McKinlay LH, Shannon MF, Seth A and Kola I. (1997). *Oncogene*, **14** (in press).
Thomas RS, Tymms MJ, Seth A, Shannon MF and Kola I. (1995). *Oncogene*, **11**, 2135–2143.
Thompson JD, Higgins DG and Gibson TJ. (1994). *Nucl. Acids Res.*, **22**, 4673–4680.
Treisman R, Marais R and Wynne J. (1992). *EMBO J.*, **11**, 4631–4640.
Tsarfaty I, Rong S, Resau JH, Rulong S, da Silva PP and Vande Woude GF. (1994). *Science*, **263**, 98–101.
Tymms MJ. (1995). *Methods in Mol. Biol.*, **37**, 31–46.
Wang CY, Petryniak B, Ho I, Thompson CB and Leiden JM. (1992). *J. Exp. Med.*, **175**, 1391–1399.
Wasylyk B, Hahn SL and Giovane A. (1993). *Eur. J. Biochem.*, **211**, 7–18.
Wasylyk B, Wasylyk C, Flores P, Begue A, Leprince D and Stehelin D. (1990). *Nature*, **346**, 191–193.
Watson DK, McWilliams MJ, Lapis P, Lautenberger A, Schweinfest CW and Papas TS. (1988). *Proc. Natl. Acad. Sci. USA*, **85**, 7862–7866.
Weider KM, Di Cesare S, Sachs M, Brinkmann V, Behrens J and Brichmeier W. (1996). *Nature*, **384**, 173–176.
Welte T, Garimorth K, Philipp S, Jennewein P, Huck C, Cato AC and Doppler W. (1994). *Eur. J. Biochem.*, **223**, 997–1006.
Wernert N, Raes MB, Lassalle P, Dehouck MP, Gosselin B, Vandenbunder B and Stehelin D. (1992). *Am. J. Pathol.*, **140**, 119–127.
Woods DB, Ghysdael J and Owen MJ. (1992). *Nucleic Acids Res.*, **20**, 699–704.
Wu JE and Santoro SA. (1996). *Dev. Dyn.*, **206**, 169–181.
Xin JH, Cowie A, Lachance P and Hassell JA. (1992). *Genes Dev.*, **6**, 481–496.
Yang BS, Hauser Ca, Henkel G, Colman MS, Van Beveren C, Stacey KJ, Hume Da, Maki RA, Ostrowski MC. (1996). *Mol. Cell. Biol.*, **16**, 538–547.
Yip MY, Sharma P and White L. (1991). *Cancer Genet. Cytogenet.*, **51**, 235–238.
Zutter MM, Kringman HR and Santoro SA. (1993). *Am. J. Pathol.*, **142**, 1439–1448.