Isolation and Characterisation of Ovine Homeobox Genes in Wool Follicles

Guy Rex Sander

B.Sc.(Hons.)

A Thesis submitted for the degree of

Doctor of Philosophy in the University of Adelaide

Department of Animal Sciences

September, 2000.

Contents

Chapter 2 : Materials and Methods

Chapter 3 : Isolation of Partial Ovine Homeobox cDNAs

Chapter 4 : Characterisation of Partial Ovine Homeobox cDNAs

Chapter 5 : Isolation and Characterisation of Ovine *Barx2*

Chapter 6 : Isolation of Ovine *Hoxc-13*

Chapter 7 : Characterisation of Ovine *Hoxc-13*

Chapter 8 : Concluding Discussion

Thesis Summary

This thesis describes the screening of the sheep wool follicle for expression of *Antennapedia* - like homeobox genes by RT-PCR and the isolation and characterisation of the *Hoxc-13* gene and a novel homeobox gene, *Barx2*.

In the RT-PCR screen of the wool follicle, nine known homeobox cDNAs, *Hoxa-4*, *Hoxb-2*, *Hoxb-6*, *Hoxb-8*, *Hoxb-9*, *Hoxc-4*, *Hoxc-5*, *Hoxc-9*, *Hoxc-13* and the novel homeobox cDNA, *Barx2* were isolated.

Northern blot analysis detected transcripts for *Hoxc-13* (2.6 kb) and *Barx2* (2.0 kb). Expression was confirmed by the more sensitive method of RNA protection for all of the homeobox genes except *Hoxc-4*, *Hoxc-5* and *Hoxa-4*.

In situ hybridisation analysis was performed using cRNA probes synthesised from the homeobox of *Hoxb-6, Hoxc-5*, *Hoxc-9*, *Hoxc-13* and *Barx2. Hoxc-9* and *Hoxc-5* were expressed asymmetrically in the cortex of the lower follicle shaft. *Hoxc-13* expression was located in the upper follicle bulb and lower shaft. A signal was not detected for *Hoxb-6* or *Barx2*.

The 3' cDNA ends for both *Hoxc-13* and *Barx2* were isolated using 3' RACE and sequenced. Using longer gene-specific probes, the initial Northern blot and RNA protection results for *Hoxc-13* and *Barx2* were confirmed. The *Barx2* and *Hoxc-13* genes were isolated from an EMBL3 lambda sheep genomic library and the sequence of *Hoxc-13* is reported for the first time. The Hoxc-13 protein contains a predicted 330 residues.

The expression of *Barx2* and *Hoxc-13* was analysed during the embryonic development of sheep skin and in the adult skin by *in situ* hybridisation using gene-specific probes. Both *Barx2* and *Hoxc-13* exhibit a spatial and differential regulated expression pattern during sheep embryonic skin development.

A *K14 / Hoxc-13* transgene was constructed and the production of transgenic mice was attempted. Of 65 mice produced no transgenic mice were obtained, possibly due to embryonic lethality of the transgene. The results suggest that both *Barx2* and *Hoxc-13* are important in hair development and may be involved in the regulation of cellular adhesion molecules and the hair keratin genes respectively.

Declaration

This thesis contains no material which has been accepted for the award of a degree or diploma in any University and in my belief, contains no material which has been published by another person, except where due reference is given.

NAME: Guy Rex Sander COURSE: Biochemistry, Ph.D.

I give consent to this copy of my thesis, when deposited in the University library, being availably for loan and photocopying.

SIGNED: DATE: 20/6/01

Acknowledgements

Foremostly, I would like to thank GOD for his omnipresent love and his providing hand while undertaking this Ph.D. May all the glory go to you.

I wish to thank Dr Barry Powell for his professional supervision that was always available any time of the day, leadership, impartation of scientific knowledge, positive criticism, encouragement throughout the course of this work but mostly your friendship.

I wish to thank Dr Simon Bawden for his tireless contribution to teaching most of the molecular biology methods required to complete this Ph.D., sharing of novel ideas, supervision, critical reading of this thesis and your friendship in the laboratory.

I wish to thank Emeritus Professor George Rogers for the opportunity to undertake this Ph.D. with its first year in his laboratory at the Department of Biochemistry and the latter three at the Department of Animal Sciences. Thanks for the stories over coffee break and the honour of working in your world class laboratory. I would also like to thank Dr John Brooker and Dr Simon Maddocks and Professor Phil Hynd who were all head of the Department of Animal Sciences for the opportunity to complete the last years of this work at the Department of Animal Sciences.

I thank the wool biology research staff, Mrs Toni Nesci for help with *in situ* hybridisation and molecular biology techniques and shared life experiences, the good, the bad and the ugly, Clive M^{c}laughlan for his technical help which was always available and his comical comments during coffee breaks, Dr Stephanie Dunn, who had an ear to listen when I spoke and for her friendship and Dr Rebecca Keough for her advice and help with computer software at the start of this Ph.D.

I thank all the research staff at both departments who helped in small but crucial ways in every day experiments that led to the completion of this Ph.D, especially Natasha Penno for help with animal handling and tissue culture, Nicole Wilson for processing sheep tissue and every day cheerfulness, Greg Nattress for sharing knowledge in molecular biology techniques, and Edna Bates for teaching cell culture techniques.

I thank my Father (Ian Sander) and Mother (Freda Sander) for their support, motivation, love and wisdom in each year of this Ph.D. Without you it would never have started.

I would like to thank my wife Trena for her constant motivational support, friendship, patience, understanding and most of all love.

During the course of this work I was supported by an ARC small grant scholarship and financial assistance from the CRC for Premium Quality Wool.

Abbreviations

AER	: apical ectodermal ridge
bp	: base pair
BSA	: bovine serum albumin
Ci	: curie
cRNA	: complementary RNA
cDNA	: complementary DNA
dATP	: deoxyadenosine triphosphate
dCTP	: deoxycytidine triphosphate
dNTP	: deoxynucleoside triphosphate
DNA	: deoxyribonucleic acid
E. coli	: *Escherichia coli*
HOX	: homeobox
kb	: kilobase pair
M_r	: molecular weight
mRNA	: messenger RNA
ORS	: outer root sheath
IRS	: inner root sheath
PAGE	: polyacrylamide gel electrophoresis
PCR	: polymerase chain reaction
RACE	: rapid amplification of cDNA ends
RNA	: ribonucleic acid
RPM	: revolutions per minute
RP-1	: reverse phase chromatography
RT	: reverse transcriptase
RT-PCR	: reverse transcription polymerase chain reaction
rUTP	: ribouridine triphosphate
SDS	: sodium dodecyl sulphate
tRNA	: transfer RNA
ZPA	: Zone of polarizing activity

Chapter 1

Introduction

Chapter 1 : Introduction

The work presented in this thesis focuses on expressed homeobox genes in the wool follicle and their expression pattern in the developing wool follicle. Chapter one is an introductory chapter and thus contains several sections that give background information on wool follicle development, wool keratin expression in the follicle, the definition and classification of the homeobox genes and lastly a short review of homeobox gene expression in the developing skin and follicle. The aims of the thesis are then presented at the end of this Chapter.

1.1 Wool Follicle Development

Human hair and wool follicle development begins at approximately 40 days gestational age and undergoes continual developmental changes through-out the life span of the organism. There are three main stages of hair growth once the follicle is formed (Fig. 1.1). The active phase of growth and differentiation, anagen, is followed by a short transitional phase, catagen, in which hair growth ceases and both the dermal papilla and the club formed at the end of the hair shaft regress towards the epidermis. Telogen, a resting phase, succeeds catagen and during this phase the club end of the hair is anchored by cortical cells, which penetrate the epithelial cap. The next round of hair growth is thought to be initiated by the interaction of the stem cells found in the bulge region with adjacent dermal papilla cells (Cotsarelis, 1990).

The overview of the developmental plan of a hair follicle is similar between species however individual species can have unique differences. For example the mouse has a pattern of hair growth where follicles cycle synchronously with a short anagen phase lasting for approximately two weeks. In comparison, the sheep has an anagen phase of years and neighboring follicles can be at different stages of the follicle developmental cycle.

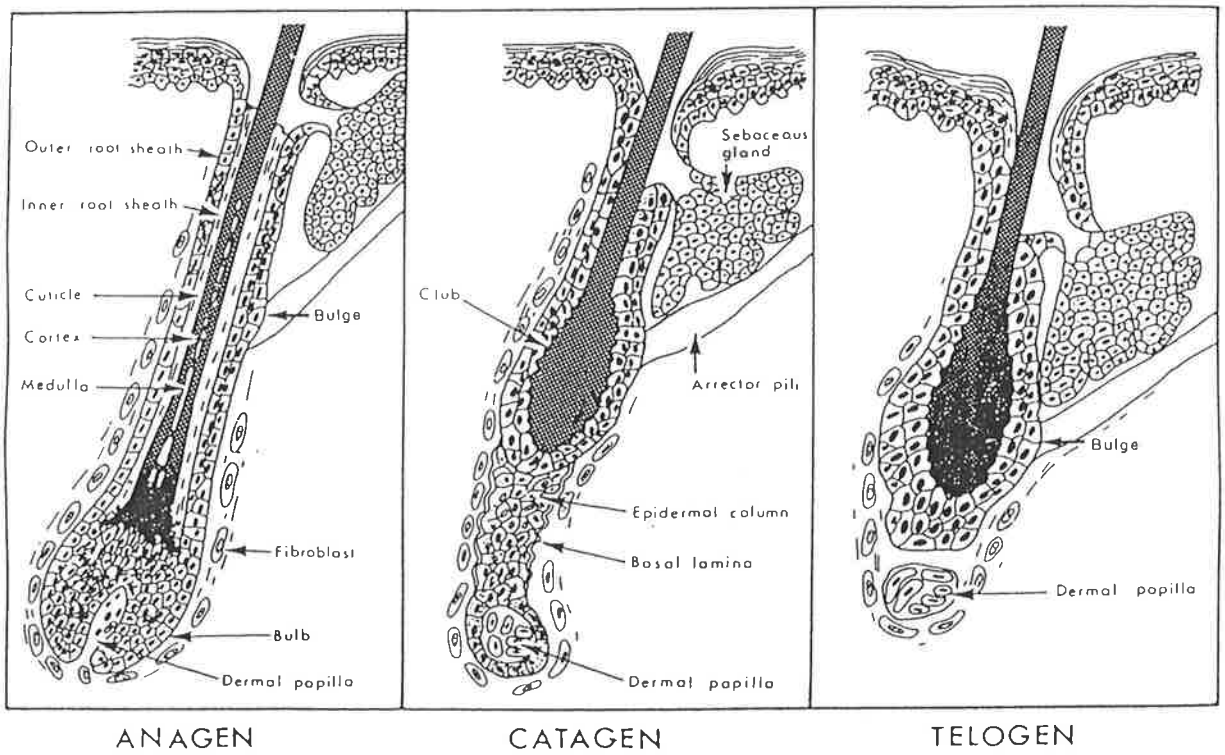


Figure 1.1 : Diagrammatic representation of the Hair Cycle. The dermal papilla persists in each stage. (Reprinted from Ebling et al., 1988).

1.2 Hair Morphogenesis

At the beginning of hair follicle morphogenesis, a cluster of mesenchymal-dermal cells aggregate into a condensed ball beneath the epidermis. A signal is then thought to be sent from the mesenchyme to the overlying epidermis where a localised thickening of the basal layer forms the hair peg (Hardy and Lyne, 1956). Epidermal cells at the leading edge of the peg begin to proliferate, forming a plug that moves in a proximal direction into the dermis. This action pushes the original cluster of dermal cells downwards and in the process forms a tubular shaft. The plug flattens before it engulfs the mesenchymal-dermal cells as a papilla that persists during all three stages of the hair follicle cycle. Unknown signals are then sent from the dermal papilla cells to the germinative bulb cells that line the dermal papilla stimulating them to divide rapidly. The bulb cells proliferate continuously at a high rate and can generate daughter cells within 12 hours for mice and 30-44 hours in sheep. As the daughter cells migrate towards the hair shaft in a stream like fashion (Hardy, 1992), their structure and function changes in a spatial and temporal manner. Terminal differentiation transforms the migrating cells into specific types of follicle cells which form the concentric cell layers of the adult follicle (Hardy, 1992; see Fig. 1.2). Cell types that originate from the germinative cells of the follicle bulb are the hair cuticle cells, cortical cells (orthocortical and paracortical), the medulla cells (not present in all fibres), the three different types of cells that make up the three inner root sheath (IRS) layers, the Henle, Huxley and IRS cuticle and the companion layer. The companion layer is located between the IRS and the ORS and there is both histological and biochemical evidence to support the notion that the companion layer is not part of either the outer or inner root sheath (Rothnagel and Roop, 1995). The mature hair fibre consists of at least two cell types, the flattened overlapping scale cells of the cuticle and the spindle shaped cortical cells that are 100 μM long and 5-10 μM wide. Some hairs have a central medulla which are interspersed with relatively large air spaces. The outer root sheath cells (ORS), form the peripheral layer around the outside of the follicle they are an autonomous layer and are not derived from the follicle bulb. The IRS cells move into the sloughing zone, cytoplasmic organelles are

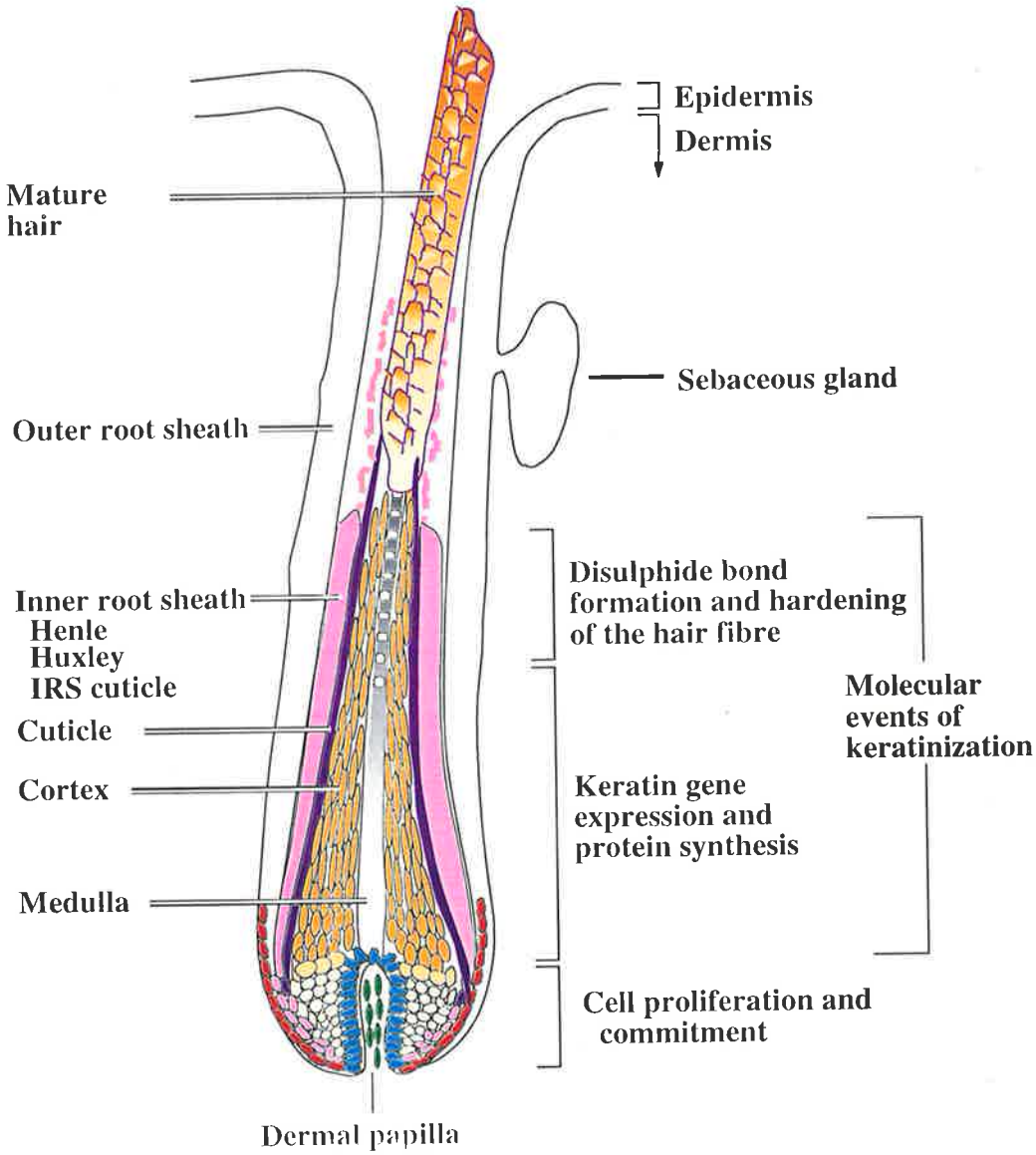


Figure 1.2 : Schematic representation of a longitudinal view of a hair follicle. The different cell types of the follicle bulb and hair shaft are shown. The zone of proliferation and differentiation is positioned at the bottom of the bulb. Directly above this zone keratin gene expression begins.
(Reprinted from Chuong, 1998).

degraded and the mature hair fibre emerges.

1.3 Hair Keratin Gene Expression

As the follicle bulb cells divide and migrate towards the epidermis they undergo terminal differentiation. A marked feature of terminal differentiation of the hair keratinocytes is the turning on of keratin gene families which produce keratin intermediate filament proteins and keratin associated proteins that cross-link via disulphide bonds to give the hair fibre its hardened character (for review see Powell and Rogers, 1994; Powell and Rogers, 1997). There are estimated to be 50-100 different hair keratin proteins residing in the hair fibre (Gillespie, 1991; Powell and Rogers, 1997). They can be divided into two major groups of proteins, the keratin intermediate filament proteins (IF), containing at least 40 components, and the keratin associated proteins (KAP), containing at least 10 families (Powell and Rogers, 1994). These two groups are present in roughly equal amounts in wool.

The hair IF keratin protein group can be subdivided into 2 classes, type I and type II that span the molecular weight range of 40-70 Kd. The type I keratins, in general, are of smaller size with acidic isoelectric points when compared to the neutral-basic type II proteins. The type I IF genes contain six introns and are approximately 4-5 kb in size and the type II IF genes contain eight introns and are approximately 7-9 kb in size (Wilson et al., 1988; Kaytes et al., 1991; Powell et al., 1992; Powell and Beltrame, 1994).

Keratin IF are heteropolymers requiring a type I and a type II chain for IF assembly (Hatzfield and Franke, 1985). They have the ability to assemble themselves without the aid of additional proteins. Type I and type II intermediate filament genes are coordinately expressed and 4-5 genes of each type expressed in the hair cortex and form filaments of 8-10 nM in diameter in the hair cortical cells (Powell and Rogers, 1994). The basic molecular structure for each type of IF protein is similar. A central α-helical core of 310 amino acids is divided into four sections 1A, 1B, 2A and 2B by three linker regions L1, L12 and L2. Heptad repeats of hydrophobic residues are found throughout the α-helical segment which enables the IF's to form coiled-coil subunits (for review see Steinert and Roop, 1988). The

core is flanked by non-helical sequences (amino (N)-terminal and carboxy (C)-terminal domains) which vary in length and in sequence. They play a role in the lateral and end-to-end cross-linking interactions necessary to pack or stabilise the coiled-coil subunits into the filamentous structure (Steinert et al., 1985; Steinert and Roop, 1988).

During terminal differentiation of the cortical cells, the heteropolymer IFs cross-link with an interfilamentous matrix that is composed of the keratin associated or matrix proteins. KAPs comprise several protein families which can be subdivided into the cysteine-rich and the glycine/tyrosine rich groups based on their amino acid composition (For reviews see Gillespie, 1990; Powell and Rogers, 1994). The cysteine and HGT proteins of the cortex are relatively low in molecular weight (6,000-9,000 Daltons) and help to form the tightly bound IF aggregates.

The expression of keratin IF proteins in hair follicles has been shown by using antibodies (Kemp and Rogers, 1970; Lynch et al., 1986; French and Hewish, 1986; Heid et al., 1988 a,b). These immunocytochemical studies lacked the ability to distinguish between individual components of either the hair type I or type II IF families due to their conserved sequences. In order to overcome this problem Powell et al., (1991), synthesised unique keratin probes representative of individual members of each family. *In situ* hybridisation of these probes to consecutive longitudinal sections of the follicle revealed the stage-specific expression of transcriptional products of individual members (Fig. 1.3). The IF keratin genes are first to be activated and expressed in the lower follicle bulb but not in the mitotic regions. They are subsequently expressed in all cortical cell types. *K1.15* is the first to be expressed in the follicle bulb cells (Lel Whitbread, Phd Thesis). *K2.12* is transcriptionally activated in cells in the upper part of the follicle bulb followed by the progressive expression of *K2.9*, *K2.10* (upper part of the bulb) and *K2.11* respectively in the lower shaft. The expression patterns of the keratin-associated proteins have also been studied by *in situ* hybridisation. The HGT keratin proteins (KAP6 family and KAP7 and KAP8 proteins) are the next group to be expressed in one side of the cortex and soon after the cysteine genes (*KAP1-4* families) are transcriptionally activated in the cells of the complementary side. The UHS (cysteine rich KAPs) gene expression pattern is restricted to one side of the cortex unlike the HGT and HS proteins which are expressed in both sides

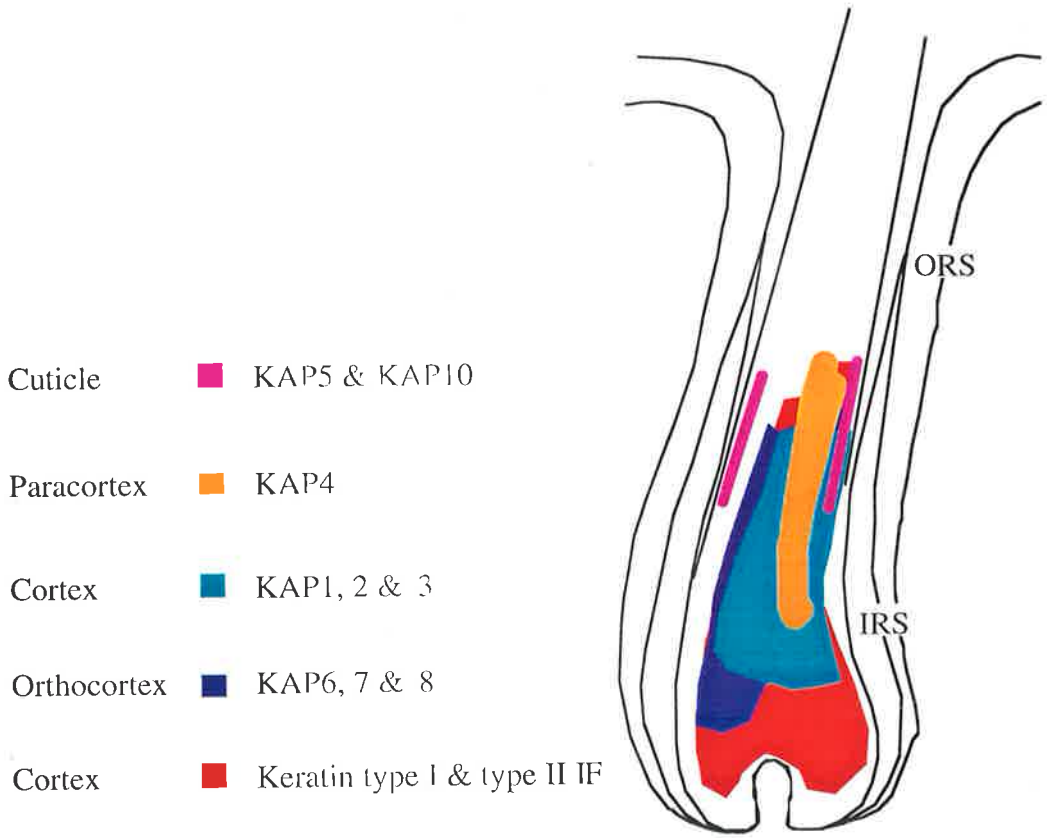


Figure. 1.3 : Cartoon of the hair follicle summarizing regional and temporal differences in keratin gene expression. Reprinted from Chuong, 1998.

further up the follicle (Powell et al., 1992). The KAP5 family is expressed specifically in the cuticle and is transcriptionally activated late in follicle differentiation (Mac Kinnon et al., 1990).

1.4 Control of Hair Keratin Gene Transcription

Most work to date has concentrated mainly on the isolation and expression patterns of keratin proteins in the follicle and epithelial tissues. Little is known about the gene sequences and transcription factors that control and regulate the expression of the keratin gene families. The keratin genes are usually clustered in families usually 5-10 kb apart. The type II IF locus has been predicted to be 500-600 kb in size containing 20-30 genes (Powell and Beltrame, 1994; Yoon et al., 1994). This is based on data obtained from yeast artificial chromosomes (YACS) (Yoon et al., 1994) and overlapping cosmid clones (Powell and Rogers., 1994).

The 5'-flanking promoter regions of the keratin gene families have been sequenced to ascertain if there are common sequence motifs that are capable of binding regulatory proteins. Promoter sequence analysis has revealed several putative, conserved sequence motifs, also found in epithelial IF genes, which may be involved in the regulation of hair keratin gene expression (Powell et al., 1991, 1992; see Fig. 1.4). A palindromic sequence named HK1, 5'-CTTTGAAG-3', resides in all hair keratin gene promoters including both IF and KAP's and is a candidate for regulating their expression (Powell et al., 1992). In the K2.10 gene promoter, HK1 is located between 180-240 bp upstream of the transcription start site and is not present in the proximal promoter regions of any epidermal keratin IF genes. Interestingly, lymphoid enhancing factor 1 (LEF-1) which recognises the same core consensus sequence 5' - CTTTGA / TA / T - 3' is up-regulated in a highly restricted pattern just before the formation of the underlying mesenchymal condensates and commitment of the overlying ectodermal cells to invaginate and become hair follicles (Zhou et al., 1995). LEF-1 is able to bind to the promoters of *K2.9* and *K2.10* through the HK-1 motif and there is preliminary evidence to suggest that that the LEF-1 binding site is an enhancer element of the K2.10 promoter (Dunn et al., 1998). Knockout mice have shown that LEF-1 is not absolutely essential for initiation of follicle morphogenesis as some follicles in the

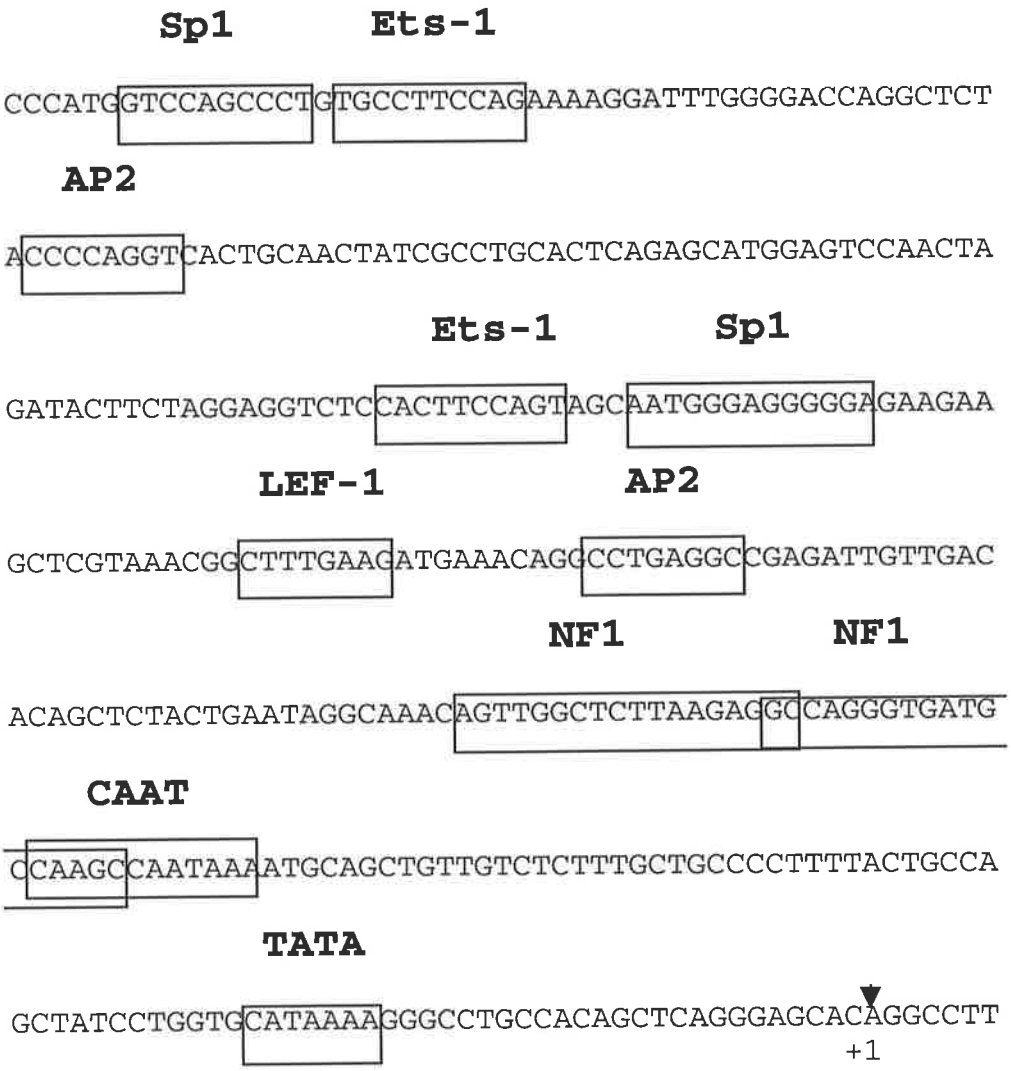


Figure 1.4 : Putative regulatory motifs in the promoter region of a wool cortical keratin IF type II gene (sheep *K2.10*). The location of the transcription start site is indicated at +1bp. (Reprinted from Chuong, 1998).

knock out mice developed until approximately day 17 gestational age and then appeared to be arrested (Van Genderen et al., 1994). Expression of LEF-1, driven by the *K1.14* epidermal basal cell promoter results in abnormal follicle development such as defects in the angle that the follicle develops and remarkable abnormal hair growth in the gums (Zhou et al., 1995). These results imply that LEF-1 has a major role in follicle development at the embryonic level.

The consensus sequences recognised by the two common transcription factors AP1 and AP2 have been found in a range of IF type II genes, *K2.9*, *K2.10*, *K2.13*, *K2.14* (Powell et al., 1992). AP2 is thought to be involved in general keratin expression (Byrne and Fuchs, 1993; Leask et al., 1990, 1991; Snape et al., 1991). Two motifs have been found in epidermal keratin IF genes (Leask et al., 1990: Snape et al., 1990). They are both similar to the AP2 motif and are involved in DNA binding interactions but do not have a tissue specific expression. FOS, one of the proteins of the AP1 complex has also been found in differentiating hair follicle keratinocytes (Fisher et al., 1991) which further supports the evidence that AP1 and AP2 are involved in the regulation of keratin gene expression.

Ten putative regulatory motifs were found within 700 bp of the transcriptional start site in the promoter region of a wool cortical keratin IF type II gene (Powell et al., 1991, Dunn et al., 1998). In addition to the CAAT and TATA motifs, various potential binding sites for AP2, Sp1, Ets-1, NF1 and Lef1 were detected (Fig. 1.4). Other motifs included the HS-1 and KTF-1 box, the latter originally identified in a Xenopus IF type I gene acting as a general activator of embryonic keratin transcription (Snape et al., 1990) and an AARCCAAA box (R is a purine) identified by Blessing et al., 1987.

No documentation has thus far reported the consensus sequences, containing a central TAAT motif, recognised by the Hox gene products in the hair keratin gene promoters.

1.5 Definition of the Homeobox

Major advances in our understanding of developmental control mechanisms have come from studying genes which cause developmental anomalies. In particular, genes

which affect the patterning of segments in the developing embryo of *Drosophila melanogaster* are prime examples. These are the coordinate genes, segmentation genes and homeotic genes. The coordinate genes are expressed maternally during oogenesis and establish the anterior-posterior and dorsal-ventral axes in the embryo. Segmentation genes affect the number and polarity of the body segments and can be subdivided into i) gap genes ii) pair-rule genes and iii) segment polarity genes. Homeotic genes act more specifically and effect changes in structures associated with individual segments as well as changing the repetitive nature of similar segments to become individually unique. For example the *Bithorax* mutation, the first homeotic mutation discovered (Lewis, 1978), generates four-winged flies by transforming the third thoracic segment from which halteres develop towards a second thoracic segment from which wings develop (Bridges and Morgan, 1923) and the *Antennapedia* (*Antp*) homeotic gene partially transforms the antennae on the head of the fly into legs (Le Calvez, 1948).

In the *Drosophila* genome, eight different homeobox containing genes, also called homeotic selector genes are clustered together on chromosome three forming a complex named HOM-C (Homeotic-Complex). HOM-C consists of the two separate complexes ANT-C (Antennapedia-C) and the BX-C (Bithorax-C) which are thought to have arisen from the single division of an ancestral HOM-C cluster (Beeman, 1989). Homeobox gene expression is present at all levels of development and determines either segmental identity, regionalization or cell identity.

McGinnis et al, (1984) demonstrated through cross hybridisation of the last exon of the homeotic *Ultrabithorax* gene (found in BX-C) with the homeotic genes *Antennapedia* and *fushi tarazu* (belonging to the drosophila embryonic segment-polarity gene class) the presence of a 180-183 bp DNA sequence common to these genes. This segment of DNA was subsequently named the 'homeobox'. The homeobox occupies only a small proportion of the coding region of the gene and encodes the 60-61 amino acid protein domain, the 'homeodomain' (Fig. 1.5). The three homeoboxes of the *Antennapedia, fushi tarazu* and *Ultrabithorax* genes are highly conserved and share a 75-77% nucleotide sequence identity. When comparisons are made at the amino acid sequence level the sequence homology increases suggesting that some amino acids within the homeodomain are conserved between

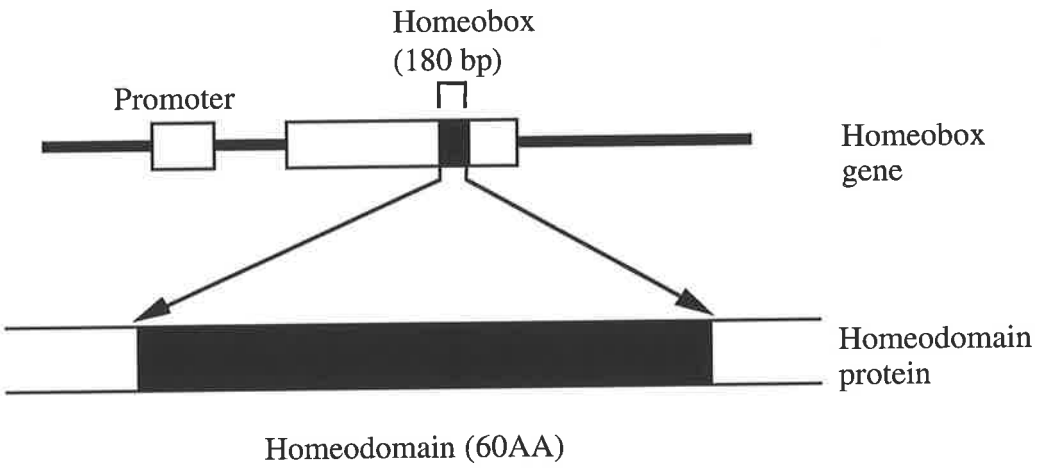


Figure 1.5 : Schematic representation of a Hox gene and location of homeobox. The Homeobox consists of a 180 bp DNA sequence usually located towards the 3' end of the gene. It encodes part of the protein (60 amino acids) known as the homeodomain

classes and essential for proper homeobox function.

Hundreds of Hox (homeobox) genes have been progressively isolated from every phylum over the previous 15 years by cross homology / hybridisation experiments and RT-PCR technology (for review see Carrasco et al., 1984; Duboule, 1994). Degenerate primers are designed from uniquely conserved amino acid regions found in most homeobox gene classes and used to amplify partial cDNA or genomic DNA fragments. Within the mouse genome more than 70 Hox genes have been isolated of which more than 30 have been extensively characterised (Akam, 1989; Kessel and Gruss, 1990) and more than 50 HOX genes have been isolated from Homo sapiens.

The vertebrate Hox genes are found clustered like their *Drosophila* counterparts (Fig. 1.6). In mouse and humans there are four clusters of Hox genes located on separate chromosomes (human chromosomes 11, 6, 15 and 2) each containing 13 genes that can be arranged into paralogous groups by amino acid sequence similarity and position within each cluster. Vertebrate homologues are found at the same relative position within their clusters. Each cluster spans approximately 100 kb and individual genes occupy on average 10 kb of cluster space (Ruddle et el., 1994). The clusters occurred perhaps by a series of chromosomal duplications and expansion of a putative ancestral primordial gene cluster (Schughart et al., 1989). It has been suggested that duplication of clusters has provided the developmental capacity to evolve increasingly complex body plans (Kappen et al., 1989; Ruddle et al., 1994).

Expression studies have shown that the chromosomal order of the Hox genes is the same as their respective expression domains in the anterior-posterior axis of an organism. This has been termed the colinearity rule such that genes expressed most anteriorly are found at the 3' end of the cluster and genes expressed towards the posterior are located at the 5' end. Unlike the *Drosophila* counter-parts where expression is contained within discrete sections, the expression domains of vertebrate HOX genes overlap and form a graded overlapping pattern. Most of the HOX genes are expressed in the posterior parts of the vertebrate body and only some are expressed in the anterior regions (Hunt and Krumlauf, 1992).

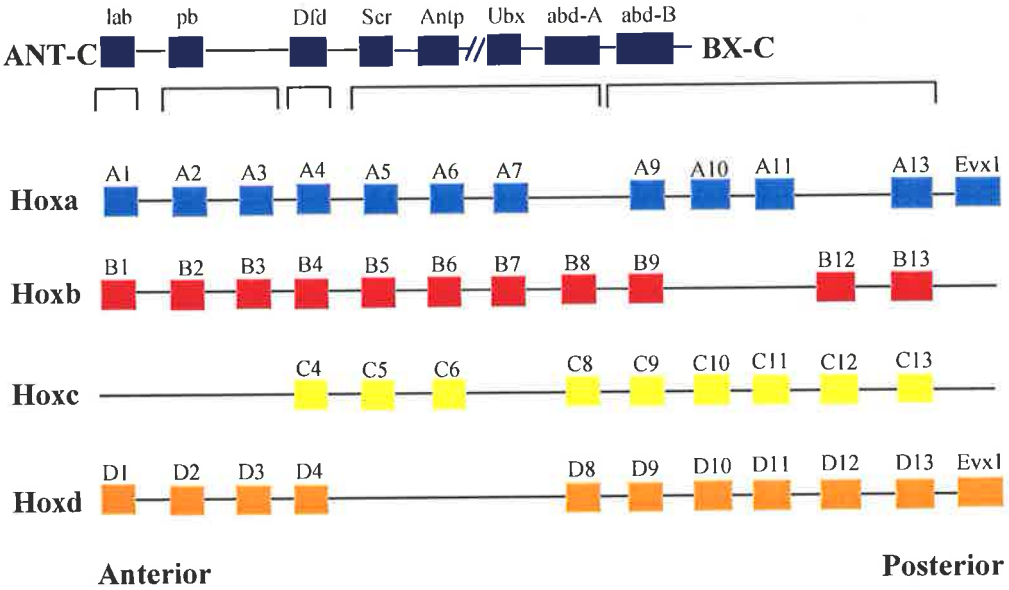


Figure 1.6 : Organisation of the *Drosophila melanogaster* HOM-C cluster and the vertebrate *HOX* clusters. The organisation of the four mammalian HOX clusters is shown in the lower part of the figure and above them is a representative of the *Drosophila* ANT-C and BX-C clusters.
(Reprinted from Duboule, 1994).

1.6 Classification of the Homeobox Genes

Difficulties arose with the random nature of the naming and classification system as more and more homeobox genes were isolated. A nomenclature format was devised by Scott (1992) to simplify the rather complex task of naming each new gene and has been adopted by the current author. The following principles apply for the mammalian *Antp*-type cluster genes and are currently adhered to by the scientific community in large. Firstly, each cluster has been given an alphabetical letter, A, B, C, or D. Secondly, each paralogous gene within the clusters are given a number from one to thirteen with the most 5' gene being designated thirteen and the most 3' paralogous gene designated one. Thirdly, human cluster genes are typed in uppercase, for example HOXC9 and the homologous mouse genes are typed in lower case, for example Hoxc-9. Furthermore, for the mouse homeobox genes, a dash is placed between the letter which specifies the cluster and the number which species the resident paralogous genes. Lastly, the name Hox is reserved for all *Antp*-type genes.

Other mammalian homeobox genes not found in the clusters form different smaller classes depending on their sequences or association with other conserved motifs (Herr et al., 1988). They are named by a three-letter abbreviation other than the term Hox. Figure 1.7 lists different classes of homeobox genes encoding conserved motifs outside of the homeodomain. The motifs are found mostly upstream of the common homeodomain but can be found downstream as seen in the NK-2 and PBC classes.

1.7 Homeobox Genes are Translated into Transcription Factors

In the normal development of an organism beginning from the embryo and continuing through adult life, complex processes such as embryogenesis and organogenesis are set into motion by molecular mechanisms only now beginning to be elucidated. These developmental processes are controlled by the coordinated expression of genes in a temporal and spatial dependent manner. The development of unique gene expression patterns in morphologically identical cells is due to the action of a hierarchy of regulatory genes, many of which encode transcription factors.

Homeobox containing genes regulate gene transcription by coding for gene regulatory proteins acting as transcription factors. Transcription factors can activate or

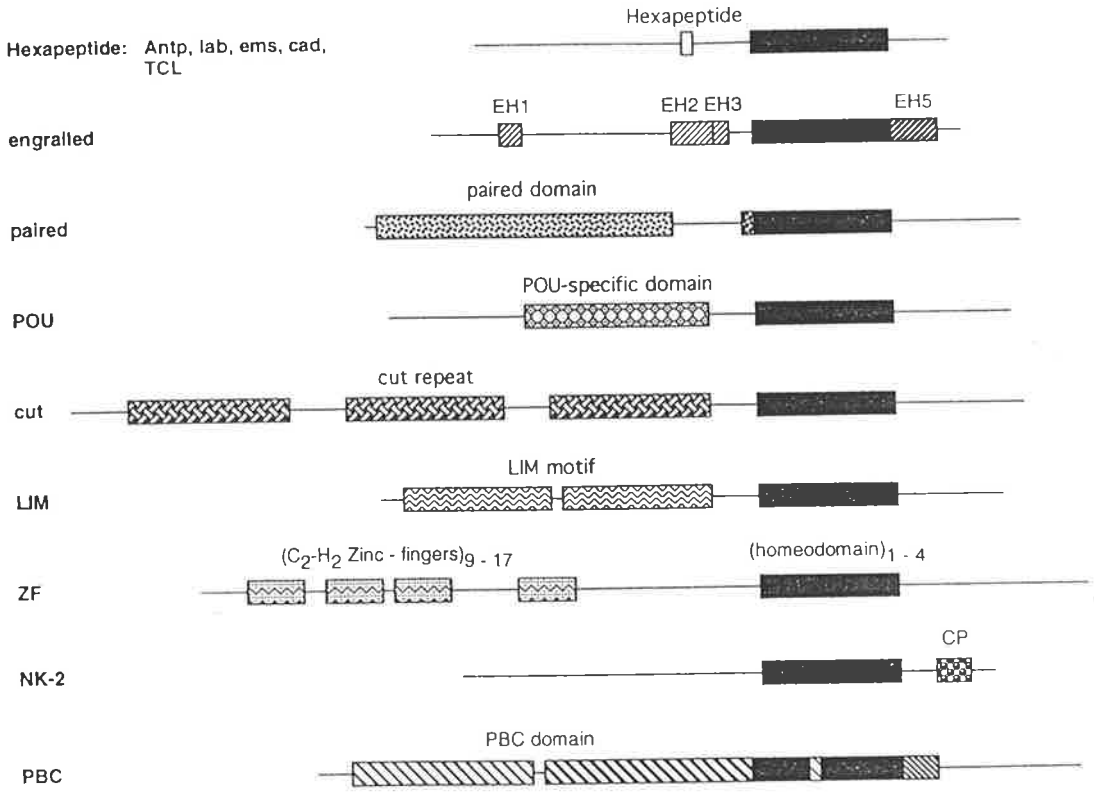


Figure 1.7 : Schematic representation of classes of homeobox genes encoding conserved motifs outside of the homeodomain. On the left are names given to the different classes. The black box represents the homeodomain and the other boxes represent conserved motifs specific to individual classes. (Reprinted from Duboule, 1994).

repress the level of transcription of a gene by binding to cis-regulatory regions such as promoters, matrix-attached regions, enhancers and silencers in single or multi-component complexes. *In vivo*, transcriptional activation was first demonstrated for the homeodomain protein *bicoid* which interacts with its target gene *hunchback* (Struhl et al., 1989; Driever et al., 1989). The *fushi tarazu* protein also acts as a trancriptional factor which dircctly interacts with the *fushi tarazu* autoregulatory enhancer (Schier and Gehring, 1992).

1.8 The Homeodomain

Homeobox proteins bind to DNA *in vitro* (Fainsod et al., 1987; Odenwald et al., 1989; Sasaki et al., 1990; Laughon, 1991) through interaction of the DNA with the homeodomain (Fig. 1.5). The homeodomain structure resembles the yeast MAT alpha 1 and alpha 2 type mating proteins whose helix-turn-helix conformation enables DNA binding (Gehring et al., 1990). The homeodomain contains three defined helices, of which helices II and III form a helix-turn-helix structure similar to the yeast repressor proteins (Fig. 1.8). The C terminal end of helix III forms a separate helix four in the Antp homeodomain. The 'turn' resides between the helices II and III and consists of three invariant amino acids.

DNA binding studies have revealed critical residues in the homeodomain required for sequence-specific contacts to the phosphates on the DNA backbone (Kissinger et al., 1990; Otting et al., 1990; Wolberger et al., 1991). The main residues providing base-specific contacts are found in helix III (Fig. 1.8), which sits in the major groove of the DNA (Otting et al., 1990). Position 6, 9 and 10 of helix III and 13 of helix IV are critical residues responsible for sequence-specific DNA contacts (Treisman et al., 1989). Residues at positions 3, 5 and 7 at the amino-terminus of the homeodomain and parts of helix II (which contact the sugar phosphate backbone), also contribute to DNA contacts (Duboule, 1994). Of the 60 amino acid residues of the homeodomain, 19 are important in contributing to DNA binding by binding to either the minor or major grooves or to the sugar phosphate backbone (Fig. 1.8).

Some motifs found outside the homeodomain as seen in the paired and POU families also confer separate DNA binding properties. For example, the paired-box family of genes contains a 182 amino acid domain found upstream of the homeodomain (Frigerio et al.,

Figure 1.8: Amino acid consensus sequence based on 346 homeodomanin sequences compiled by Duboule, 1994. Schematic representation of intramolecular and protein-DNAcontacts made by individual amino acids of the homeodomains of Antp, engrailed and MATα2. Not all of these contacts are made by each of the homeodomains. The first amino acid of the homeodomain is indicated by the number 1 located above the sequence. The four helices are boxed.

m = Residues that make contact in the minor groove of DNA.
# = Residues that contact the sugar-phosphate backbone.
B = Residues that make contact in the major groove of DNA.
H = residues that contribute to the hydrophobic core that is responsible for the tertiary structure of the homeodomain. (Reprinted from Duboule, 1994).

1986) and the POU class of genes contains an additional conserved domain which varies in length from 147 to 156 amino acids and both motifs enable DNA binding properties separate to the homeobox (Ruvkun and Finney, 1991).

1.9 Sequences Recognised by the Homeodomains

In cell culture, homeodomain containing proteins have been shown to regulate transcription through interactions with specific DNA recognition sequences (Zappavigna et al.,1991; Arcioni et al., 1992; Jones et al., 1992; Popperl and Feathersone, 1992). These have been identified for several homeodomain classes including the POU domain (Sturm and Herr, 1988), the paired (*prd*) domain (Adams et al., 1992) and the eve homeodomain (Hoey et al., 1988). The sequences are usually a group of 10-12 base pairs centered on a TAAT motif (Scott and Goldsmith, 1993). For example, the consensus palindromic sequence GTTAATNATTAAC is recognised by the sequence specific DNA-binding protein, HNF-1 (hepatic nuclear factor-1) originally identified from rat liver. HNF-1 is necessary for the transcription of many genes in the liver such as albumin, alpha-1-antitrypsin and surface antigens of HBV virus (reviewed in Mendel and Crabtree, 1991; Rey-Campos and Yaniv, 1992). The consensus sequence ATTTTAATTGT is recognised by the *Fushi tarazu* protein (Treisman et al., 1989), *engrailed* protein (Desplan et al., 1988) and the *Xenopus* homeodomain-containing protein Hoxc-6 (Cho et al., 1988). Another consensus sequence G/A/CTAATG is recognised by Hoxa-5, Hoxa-7, Hoxb-5 (Odenwald, 1989; Gross and Gruss, 1994). The protein product of the *Ultrabithorax* (*UBX*) homeotic gene has been found to bind to specific DNA sequences containing the consensus sequence 5'TAATAATAATAATAATAA 3' and to repeats of the hexanucleotide TAATCG (Beachy et al., 1988). Some homeodomains can bind to two classes of sites in which binding is dependent on critical amino acids in the homeodomain. The eve homeodomain recognises the TAATTG sequence and also TCAGCACCG (Hoey and Levine, 1988; Hoey et al., 1988) and the *prd* homeodomain recognises TAATCG and TTTGACGT (Treisman et al., 1989). For the *prd* homeodomain, recognition of the sequence TAATCG is dependent on the amino acid at position 50, located in the third helix and recognition of TTTGACGT is dependent on the amino acid at position 41. A perfect reverse complement of (C /

G)TTAATTG, was shown to be an optimal binding sequence for the homeodomain of Msx-1 protein (Catron et al., 1993) and was also located in the murine *Msx-1* promoter which implies autoregulation capabilities.

1.10 Regulation of Homeobox Gene Function

The regulation of homeobox gene function involves the coordinated action of growth factors and cytokines including post-translational modification of the protein product itself through phosphorylation and glycosylation (reviewed by Scott and Goldsmith, 1993). Hox genes can also directly regulate themselves either by autoregulation or alloregulation (Fainsod et al., 1987; Jones et al., 1992).

As the DNA binding specificity is similar for the Hox proteins *in vitro* (Hayashi and Scott, 1990) they are unlikely to achieve their specificity of action *in vivo* only on the basis of DNA-protein interaction. Experimental results to date demonstrate that many of the proteins can bind to similar sites, and that at least some of the proteins can bind to more than one type of site (Zappavigna, 1994). It has been shown that biological activity *in vivo* and *in vitro* is not always dependent on homeodomain DNA binding (Schier and Gehring, 1993). Protein-protein interactions between different Hox proteins and with other components of the transcriptional machinery could be critical factors in the regulation of Hox functional specificity (McGinnis and Krumlauf, 1992). In cell culture, human HOXD9 and the immediate upstream HOXD10 protein activate transcription of the *HOXD9* promoter by binding to a 100bp upstream element termed the Hox cross-talk region (HCR) (Zappavigna et al., 1994). HOXD8 does not bind to the HCR but antagonises the activating function of HOXD9 / HOXD10 as shown in a cotransfection assay (Zappavigna, 1991). The HOXD8 transcriptional inhibitory function is DNA-binding independent and is not affected by deletion of the homeodomain helix II / III region (which forms the helix-turn-helix DNA binding structure) but does require the helix 1 region of the homeodomain and an additional, effector domain located at the protein amino-terminal end.

Another type of protein-protein interaction is exemplified by the POU transcription factors (see Stern et al., 1989 and Verrijzer et al., 1992) which form dimers to increase their DNA-binding specificities (Manak and Scott, 1993) much like the prokaryotic helix-turn-

helix proteins (Pabo and Sauer, 1984; Ptashne 1986). Dimers are formed for several homeodomain proteins and in each case dimerization is mediated by sequences outside the homeodomain (Goutte and Johnson 1988; Sauer et al., 1988). The best understood example is the MATα2 homeodomain proteins which repress transcription by binding to DNA as either a MATα2-MATα2 or a MATα2-MATα1 dimer.

1.11 Hox Gene Promoter Organisation

Mammalian homeobox gene promoters are not well characterised but the present data shows that in general, most Hox genes have minimal promoters that do not have TATA boxes but rely on other initiators instead. Recent studies have revealed that organisation can differ between different Hox gene promoters. For example TATA boxes are found in the *HOXC5* (Arcioni et al., 1992) and *HOXD4* promoters (Cianetti et al., 1990) but not in the *Hoxa-4* (Galiot et al., 1989) or *Hoxc-4* promoter (Geada et al., 1992). *Hoxa-4* and *Hoxc-5* contain potential Sp1 binding sites that can mediate the initiation of transcription at multiple closely spaced sites. Other Hox genes, *Hoxc-6* (Coletta et al., 1991) and *Hoxb-4* (Gutman et al., 1994) also lack the TATA sequence motif and may depend on initiators such as Sp1. The promoter of the mouse *Hoxb-4* gene contains multiple positive and negative regulatory elements and requires HTF (Hox transcription factor), a previously unknown factor, to bind to the sequence GCCATTGG (+148 to +155) for expression (Gutman et al., 1994).

Recently, the promoter of the murine *Msx-1* gene was analysed for potential autoregulatory cis-acting elements. The *Msx-1* promoter also lacks a TATA-like sequence but has the recognition elements for AP-1, AP-2, AP-3, Sp1, a possible binding site for RAR : RXR (Retinoic acid receptor) and a number of TCF-1 (thyroid cell factor) consensus motifs (Kuzuoka etal., 1994).

1.12 Hox Genes, Patterning and Cell Fate

For more than twenty years, the *Drosophila* embryo has been a powerful model in studying the role of homeotic genes in the patterning of segments and segmental structures. These genes were first discovered by their mutations which cause homeotic tansformations

(Lewis, 1978). Over the last decade, a vast body of work has also accumulated that shows Hox genes act to interpret positional information and control vertebrate morphogenesis (review in Holland and Hogan, 1988; Deschamps and Meijlink, 1992; McGinnis and Krumlauf, 1992). It is the combinatorial and overlapping expression patterns of Hox genes in the anterior-posterior, proximal-distal and ventral-dorsal axes which specifies the positioning and development of the body structures (Graham, 1993; Hunt and Krumlauf, 1992; McGinnis and Krumlauf, 1992; Ros et al., 1994). This action seems to be conserved for the vertebrates and invertebrates and hence is a universal mechanism of development. However, how this activity relates to the generation of anatomical variation remains to be answered.

Genetic knockout experiments have lead to developmental abnormalities demonstrating a direct causitive role of the homeobox genes in differentiation and cellular lineage specification during embryogenesis (Chisaka et al., 1992; Chisaka and Capecchi, 1991; Joyner et al., 1991). For example, the knock-out of Hoxa-2 in mice results in homeotic transformation of structures of the second branchial arch into structures typical of the first arch (Rijli et al., 1993).

Vertebrate homeo proteins have also been identified as transcription factors required for expression of lineage-specific genes. Pit-1 (a member of the POU gene family) is associated with expression of several genes in the pituitary (Ingraham et al., 1988); Oct-2 (also a member of the POU gene family) is necessary for immunoglobin expression by B lymphocytes (Muller et al., 1988) and TTF-1 (thyroid transcription factor-1) necessary for the expression of thyroid-specific genes (Guazzi et al., 1990). In the developing limb bud, transcripts of *Msx-1* and *Msx-2* become restricted to the distal mesenchyme and apical ridge (Tickle and Eichele, 1994). The genes mentioned above do not belong to the *Antennapedia* class Hox genes and are referred to as divergent homeobox genes. The cluster Hox genes provide positional information for limb development along the anterior-posterior axis but it may be that the divergent genes are required for control of cell type-specific gene expression.

1.13 Hox Genes and the Developing Limb

The developing limb has also been exploited for the study of pattern formation and is proving invaluable in understanding how Hox genes work (review in Gardiner and Bryant., 1996; Tickle, 1996). The limb first develops as a bud of mesenchymal cells originating from the lateral plate mesoderm (Chevallier et al., 1977) and is covered by the overlying ectoderm. The growth of the bud is then controlled by the apical ectodermal ridge (AER; Fig. 1.9), a psuedo-stratified epithelium which is itself induced and maintained by the underlying mesoderm. A population of mesoderm cells located underneath the epithelium is called the progess zone and is itself controlled by the AER. As the limb develops in the three axial directions, a high cell mitotic rate can be detected in the zone of polarizing activity (ZPA) which is found at the posterior edge of the limb bud and signals to nearby mesenchyme (Tickle and Eichelle, 1994).

In the developing limb, the expression of Hox genes in a temporal and spatially dependent manner has been well documented for the vertebrates (Ros et al., 1994; Tickle and Eichele, 1994 and Tabin, 1995). The Hox genes are sequentially expressed in each axial direction and this is dictated by their chromosomal order. This is in accordance with the temporal colinearity rule such that genes located at the 3' end of the clusters are expressed before genes at the 5' end and tend to have a more rostral boundary of expression. Recently, attention has been focused on the *Abd B* class genes of the Hox D cluster which are located at the extreme 5' end and comprise the paralogous groups 9 to 13 (see Fig. 1.6). The *Abd B* class genes are expressed in succesively more postero-distal areas at early stages of limb development (Dolle et al., 1989; See Fig. 1.10). Genes located at the 3' end of the *Abd B* class (ie. paralogous groups 9 and 10) are widely expressed during budding and are transcribed in more proximal areas of the growing limb (Izpisua et al., 1991). Interestingly, genes of the HOX B complex are not expressed during the development of the limb.

The expression patterns of the *Abd B* genes overlap in the antero-posterior axis of the limb along which each digit develops. The zone of polarising activity is thought to control antero-posterior patterning and maintain the apical ectodermal ridge. Hence it is has been proposed that the ZPA orders the overlapping expression domains of the Hox genes which in turn define territories that will become digits (Izpisua et al., 1991; Nohno, 1991).

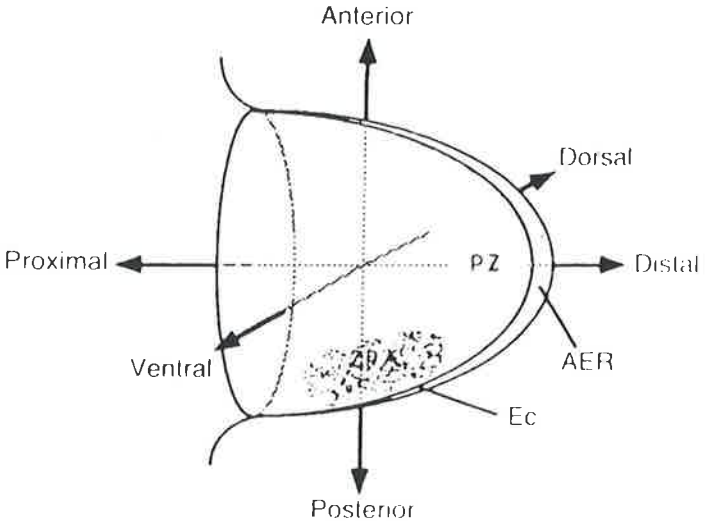


Figure 1.9 : Schematic representation of the developing limb bud showing the apical ectodermal ridge (AER) and the progress zone (PZ) located underneath the AER. Ec = ectodermal layer. The three developmental axis are also shown. (Reprinted from Duboule, 1992).

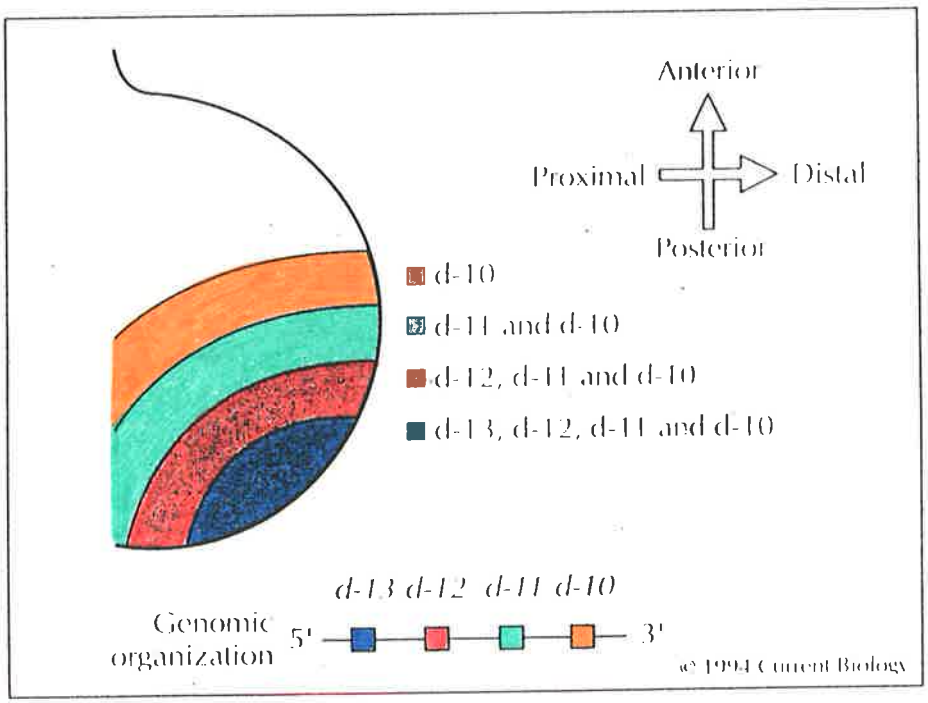


Figure 1.10 : Schematic representation of the limb bud. The location of the *Abd B* class genes in the posterior-distal region are shown. The bottom diagram shows the genomic organisation of the *Abd B* class genes. (Reprinted from Graham, 1994)

1.14 Hox Expression in the Skin

Reports of homeobox gene expression in skin are progressively increasing (see Scott and Goldsmith, 1993 and Widelitz, et al., 1997 for reviews). Differential homeobox gene expression in the skin was initially analysed by RT-PCR or RNA protection. *In situ* hybridisation and immunolocalisation analysis have also been used to locate their cell-type expression in the skin. Both Hox genes and divergent homeobox genes have been shown to be expressed in the dermis, epidermis and follicle. The following paragraphs summarise the results of studies that have analysed homeobox gene expression in the skin with a focus on the epidermis and dermis. Expression in the hair follicle is detailed in the next Section (Section 1.15).

One of the earliest reports to show that homeobox genes are expressed in keratinocytes used RT-PCR (Thomas et al., 1989). A homeobox gene of the *Antennapedia* class, *HOXA7*, homologous to the mouse *Hoxa-7* gene, was isolated from a cultured adult human keratinocyte library using RT-PCR. Soon after this report was published several other groups reported differential homeobox gene expression in the skin. The differential expression of all the Hox B genes, except *Hoxb-1*, in the skin of murine fetus was demonstrated by RNA protection (Detmer et al., 1993). Expression for each gene varied considerably at different fetal developmental stages. Additionally, *Hoxa-4*, *-b-4*, *-c-4* and *-d-4* paralogous genes were shown to be expressed in mouse fetal skin with the amount of transcript for each gene being relatively constant from the earliest gestational day examined (day 16) to the day of birth (day 19) (Detmer, 1993). *HOXC4* expression in adult human skin was discovered by RT-PCR and RNA *in situ* hybridisation (Rieger et al., 1994). *HOXC4* expression in the epidermal suprabasal layers was similar to the differentiation markers, keratins 1, 10 and 11 suggesting a possible role for homeobox genes in the regulation of epidermal keratins. A review by Scott and Goldsmith, (1993), reported the expression of *HOXA1*, *HOXA4*,*HOXA6*, *HOXC2* and *HOXC3* genes in human fetal melanocytes. A divergent homeobox gene, *Brn-5*, a member of the POU family was reported to be expressed in the skin (Andersen et al., 1993a).

Gradually, more reports have located homeobox gene expression in the skin using *in*

situ hybridisation. These reports show that homeobox genes exhibit differential spatio-temporal expression patterns. In the mouse, two homeobox genes, *Otx1* and *Otx2*, which are related to the *orthodenticle Drosophila* gene, are both expressed in nasal epithelium at E12.5, (Kanzler et al., 1994). *Otx1* was expressed in both the neurodermal and ectodermal epithelia including corneal epithelium and upper lip epidermis. *Otx2* expression was restricted to the nasal epithelium and its associated glands. In the same study, *Hoxc-8* was detected initially in ventral thoracic skin at E14.5. During the period of E16.5 to E18.5 expression was detected in dorsal epidermal cells with strongest expression in the thoracic ventral epidermis (Kanzler et al., 1994). Expression was restricted in the dorsal region to the lumbar and caudal skin along the cephalo-caudal axis and from the throat to the caudal tip in the ventral region (Kanzler et al., 1994). In the same report, the Hox genes *Hoxd-9* and *Hoxd-11* were shown to exhibit spatial and differential expression in developing caudal mouse skin (Kanzler et al., 1994). *Hoxd-9*and *Hoxd-11* were first detected in the epidermis at E14.5 and reached their highest level of expression at E16.5. At E18.5 both *Hoxd-9*and *Hoxd-11* were expressed in the epidermis. *Hoxd-9* was uniformly expressed in both the basal and suprabasal epidermal layers while *Hoxd-11* was preferentially localized in the epidermal basal layer. At both two and four days after birth, *Hoxd-9*and *Hoxd-11* were expressed mostly in the epidermal basal layer. Another cluster gene, the *Hoxb-6* gene was shown to have a uniform expression pattern in the developing murine epidermis by *in situ* hybridisation (Mathews et al., 1993).

In recent years *Msx-1* and *Msx-2* which belong to the non-cluster homeobox class *msh* (muscle segment homeobox) have been reported to be expressed in skin (Chuong et al., 1996; Stelnicki et al., 1997; Jiang et al., 1999). These two genes are expressed at almost all sites of epithelial-mesenchymal interactions and may play a role in organ formation controlled by these interactions (reviewed in Davidson, 1995 and Chuong, 1998). A recent detailed report describes the differential expression patterns of HOXA4, HOXA5, HOXA7, HOXC4, HOXB4, and HOXB7 in human fetal and adult skin (Stelnicki et al., 1998). These genes are all expressed in the epidermis during human embryonic skin development and although expression levels differ at any one embryonic time point, they show similar temporal and spatial expression patterns to each other in the epidermis. Their

expression is restricted to the epidermal basal layer during early skin development, spreads to the suprabasal layers during mid to late second trimester development and is downregulated and restricted to the epidermal suprabasal layers in newborn and adult skin (Stelnicki et al., 1998).

1.15 Expression of Hox Genes in the Hair Follicle

Reports of homeobox gene expression in the follicle are progressively increasing. The location of homeobox gene expression in the follicle are summarised in Table 1.

Table 1: Summary of homeobox gene expression patterns in mouse and human follicles.

The known expression patterns of homeobox genes reported at the beginning of this project are denoted by an '*'.

Hox		Location of Expression	Reference
Skn1i	*	Epidermis and hair cortex	Anderson et al., (1993b)
Skn1a	*	Epidermis and hair cortex	Anderson et al., (1993b)
Hoxb-2	*	Murine hair placode	Whiting et al., (1991)
Hoxb-4		Follicle placodes (promoter - tg study)	Detmer et al., (1993)
Hoxb-6		Vibrissae	Mathews et al., (1993)
Hoxc-4		Upper ORS and epidermal suprabasal layer of adult follicle	Rieger et al., (1994)
Hoxc-8	*	Epithelial cells of peg, dermal papilla fibroblast and epidermis	Bieberich et al., (1991) Kanzler et al., (1994)
Hoxc-13		Bulb, cortex, companion layer of ORS	Godwin and Capecchi (1998)
Hoxd-9		ORS, bulb, epidermis and epithelial peg	Kanzler et al., (1994)
Hoxd-11		ORS, bulb and epithelial peg	Kanzler et al., (1994)
Hoxd-13		Matrix cells of stage 7 caudal hair	Kanzler et al., (1994)
HOXA4		Epidermal layers of developing placode.	Stelnicki et al., (1998)
HOXA5		Inner root sheath	Stelnicki et al., (1998)
HOXA7		Epidermal layers of developing placode.	Stelnicki et al., (1998)
HOXC4		Epidermal layers of developing placode. Dermal fibroblasts at mid-second trimester, absent in adult.	Stelnicki et al., (1998)
Alx4		Dermal papilla of hair and vibrissa	Hudson et al., (1998)
Dlx3		Matrix cells of bulb and epidermis	Morasso et al., (1996)
MOX-1		Matrix of developing placode, epidermis	Stelnicki et al., (1997)
Msx1		Matrix of developing placode, epidermis	Reginelli et al., (1995)
Msx2		Matrix of developing placode, epidermis	Reginelli et al., (1995)
Otx1		Epithelial peg and upper lip epidermis	Kanzler et al., (1994)

One of the first reports of homeobox expression in the follicle was by Whiting et al., (1991) who showed *Hoxb-2* expression in the murine hair placode. In the same year, Bieberich et al., (1991) studied Hox gene expression in transgenic mice. The transgene used in this study consisted of the *Hoxc-8* promoter driving the *E. coli* Galactosidase reporter gene. Their data suggests that *Hoxc-8* is expressed in the dermal papilla of posterior ventral and posterior dorsal pelage follicles. The authors cautioned that the expression may not be indicative of the endogenous gene activity which may be regulated by 3' flanking or more distant 5' flanking DNA sequences not included in the transgene. They used reverse transcriptase-PCR analysis to show that *Hoxc-8* transgene expression was greater in posterior skin samples. In another report, *Hoxc*-8 expression in the dermal papilla was confirmed by in situ hybridisation by Kanzler et al., (1994). Expression was also detected in the epithelial component of the developing placodes located in the dorsal region at E18.5.

Expression of the two homeobox genes, *Skn-1a* and *Skn-1i* which are closely related to *Oct-2*, a member of the POU family of transcriptional activators, was detected in the cortical cells of mouse anagen hair follicles (Andersen et al., 1993b). The expression of these two genes was restricted to a subset of follicles. This may be due to stage-specific expression during cyclic mouse hair growth. These genes are divergent homeobox genes and may contribute in part to the development of the cortical cell lineage.

Msx1 and *msx2* which belong to the non-cluster homeobox class *msh* (muscle segment homeobox) are expressed in the epithelial matrix of the hair follicle from the placode stage onward (Reginelli et al., 1995; Noveen et al., 1995).

Four homeobox genes, *Hoxd-9*, *Hoxd-11*, and *Hoxd-13* (*Antennapedia*-type) and *Otx1* (*orthodenticle*) exhibit differential expression in the mouse integument (Kanzler et al., 1994). *Otx1* transcripts were present in both nasal and facial integuments, including nasal glands and epithelial cells forming the hair vibrissa follicles at E15.5. *Hoxd*-9 and *Hoxd-11* expression was located in caudal skin from 14.5 days of gestation. At E18.5 both *Hoxd-9* and *Hoxd-11* expression was detected in the hair bud epithelial cells (stage 3 of hair formation). Two days after birth, *Hoxd-9* and *Hoxd-11* were expressed in all the epithelial follicular cells, however, four days after birth they were expressed mainly in the uppermost outer root sheath and bulb of stage 8 follicles. Interestingly, *Hoxc-13* expression was

detected in the hair matrix 2 days after birth but was not detected in the follicle four days after birth (Kanzler et al., 1994).

HOXA4, *HOXA5*, *HOXA7*, *HOXC4* and *HOXB7* exhibit temporal and spatial expression patterns within the follicles of human fetal skin (Stelnicki et al., 1998). *HOXA4*, *HOXA5* and *HOXC4* are expressed at week 17 however *HOXC4* has a weaker signal. A signal for *HOXB7* was strong at 21 weeks and from 21 weeks to four months, *HOXA4*, *HOXA7* and *HOXC4* are expressed in the cortex while *HOXA5* was expressed in the inner root sheath. *HOXA4*, *HOXA5* and *HOXA7* were expressed in newborn follicles and extremely weak HOX gene expression was seen in adult hair follicles (Stelnicki et al., 1998).

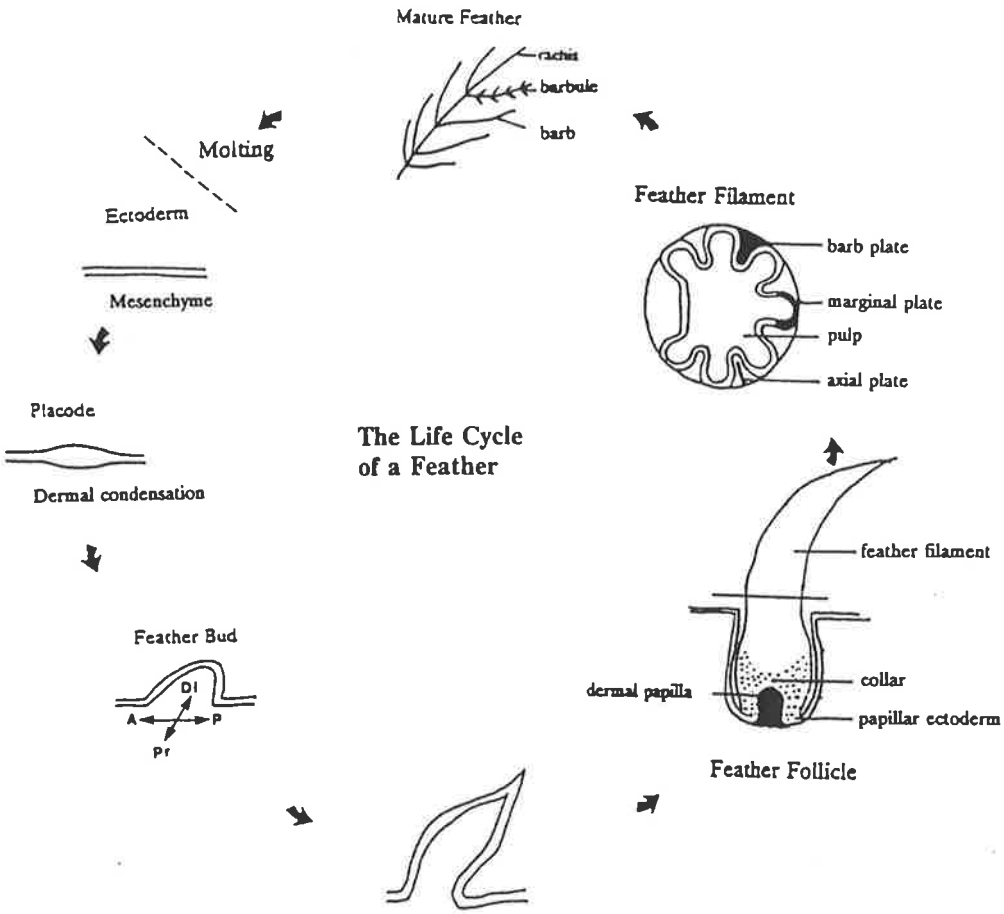
Hoxc-13, a member of the *Abd-B* class was recently reported to be expressed in the bulb, cortex and companion layer of the outer root sheath of vibrissae and hair follicles throughout the body (Godwin and Capecchi, 1998). Expression was first detected in vibrissae at E13.5 and was first detected in pelage hair follicles at E15.5 (tylotrichs).

1.16 Hox Expression in Hair-Related Appendages

The chick feather bud is an analog of the developing hair follicle (Fig. 1.11 A). Over the last ten years expression of several homeobox genes have been reported in the feather follicle namely Hoxc-6, Hoxd-9, *Msx*-1, *Msx-2* and *Hoxc-8* (Chuong et al., 1993, Noveen et al., 1995, Chuong 1996, Kanzler et al., 1997). Using antibodies to Hox-c6 and Hoxd-4, Chuong et al., (1993) showed that there is a homeoprotein gradient within the mesoderm of developing feathers, and that the expression pattern is position-specific. They hypothesized that Hox codes, derived from the combined expression pattern of homeoproteins, determine the phenotypes and orientation of skin appendages. Furthermore, another Hox gene, *Hoxc-8*, has been proposed to play a role in the specification of dorsal skin morphogenesis and is expressed at the first stages of feather formation in the epidermis and dermis (Kanzler et al., 1997). Two genes originally identified on the basis of homeobox sequence homology to the *Drosophila msh* gene (muscle segment homeobox gene), *Msx-1* and *Msx-2*, are both expressed in the epithelial placode of the feather follicle (Noveen et al., 1995).

The morphogenesis and development of the tooth also has a similar plan to the hair

A .



B .

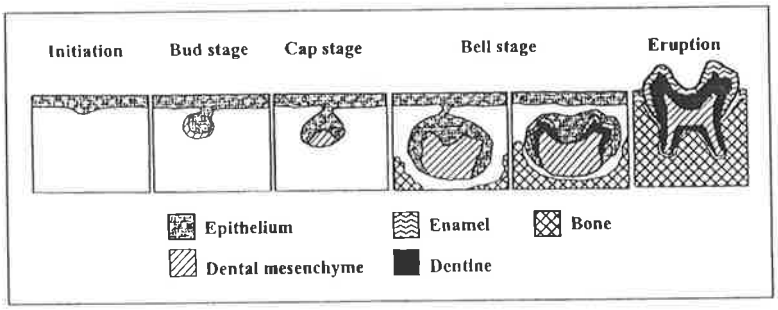


Figure 1.11 :

A. Schematic representation of the life cycle of the feather. The structure of the feather follicle is the analog to the hair follicle. (Reprinted from Chuong, 1993).

B. Schematic representation of the different stages of tooth development. The development and developmental structures of the tooth are similar to those of the hair follicle. For example, the dental papilla of the tooth appears very similar in structure and composition to the hair dermal papilla. (Reprinted from Chuong, 1998)

follicle (Fig. 1.11 B). Tooth development is initiated in the oral epithelium that thickens at discrete sites and invaginates to form the dental lamina. Over the last eight years many homeobox genes have been implicated in the molecular basis of patterning of teeth by various researchers (MacKenzie et al., 1992, Jowett et al., 1993, Vainio et al., 1993, Weiss, 1994, Nieminen et al., 1995, Tureckova et al., 1995, Vastardis et al., 1996; review in Chuong, 1998, Karg et al., 1997). The divergent homeobox genes *Msx-1* and *Msx-2* have been shown to be expressed in the dental placodes of mouse oral epithelium, suggesting a role in the specification of tooth position (MacKenzie et al., 1992). Interestingly, growth factors in the BMP- and FGF-families are expressed in dental epithelium during initiation of tooth development and have been shown to upregulate the expression of Msx-1 and Msx-2, (Thesleff and Sahlberg, 1996). *Msx-1* knockout mice exhibit multiple craniofacial anomalies including complete tooth agenesis (Kurisu et al., 1997). The homeobox genes *dlx1*, *dlx2*, *dlx3* and *dlx5* belonging to the *Distal-less* (*Dlx*) gene family, have also been detected in murine developing tooth germ by RT-PCR (reviewed in Chuong, 1998, Price et al., 1998, Lezot et al., 2000). Recently Pitx2, a Bicoid-type homeobox was reported to be expressed in the tooth (Hjalt et al., 2000).

1.17 Aims of the Research Project

A most important step in the study of hair growth is the identification of the controlling genes. To date, very little is known about the molecular mechanisms that control keratin gene expression in the wool follicle. It has been well established that homeobox genes have a direct causitive role in organogenesis and patterning of the early embryo and are pivotal in the differentiation of many cell types. Furthermore, they are focal points in cell commitment and terminal differentiation, acting as transcription factors of particular genes or sets of genes. In the hair follicle bulb, the commitment of the proliferating cells to specific fates with distinctive keratin gene expression programs could be regulated by homeobox genes. Discovering and correlating homeobox gene expression during the hair cycle would be an important contribution to the understanding of the control of hair growth.

Prior to the initiation of this project, only a few homeobox genes had been reported in the follicle (see Table 1). Hence, the first aim of this project was to screen the anagen

wool follicle for expressed homeobox genes by degenerate-primed RT-PCR based on amplification of the180 bp homeobox region.

The second aim of the project was to isolate and characterise novel homeobox genes identified in the RT-PCR screen. In this objective the corresponding genes would be isolated from a sheep genomic library and sequenced.

The third aim was to investigate the expression of homeobox genes in hair follicles by RNA *in situ* hybridisation and compare with hair cell type and keratin gene expression patterns.

The last aim was to undertake a functional analysis of a selected homeobox gene in the hair follicle by ectopic expression in transgenic mice.

Chapter 2

Materials and Methods

2.0 Materials

2.1.0 Bacterial strains

Three *E.coli* K12 strains were used to propogate recombinant DNA. The genotypes of each strain are

ED8799 : *hsdS, metB7, supE(glnV) 44, supF (tyrT) 58, lacZ_M15,* r_k^- m_k^- (a gift from Dr C.S. Bawden, Department of Animal Sciences).

XL1Blue : *recA1,endA1,gyrA96, thi-1, hsdR17, supE44, relA1, lac,* [F' *proAB,* $lacI^q$, *ZΔM15* Tn*10*(tet^r)] (Bullock, 1987).

DH5 : *supE44ΔlacU169 (Ø80 lacZ* ΔM15)*hsdR17 recA1 endA1 gyrA96 thi-1 relA1* (Hanahan 1983).

2.1.1 Phagemid Strains

Recombinant DNA strands were cloned and subcloned into the following plasmid strains pGEM-3Zf(+), pGEM-5Zf(+), pGEM-7Zf(+), pGEM-11Zf(+): purchased from Promega Corporation. pBluescript II KS (+): purchased from Stratagene.

2.1.2 Construct

K1.14 / B- Globin: A gift from Associate Professor E. Fuchs.

2.1.3 Enzymes

Klenow fragment (*E.coli* DNA polymerase I), T4 Polynucleotide kinase, T4 bacteriophage DNA ligase, Taq polymerase and RNase inhibitor were purchased from Bresatec Ltd. Restriction enzymes were purchased from either Pharmacia, New England Biolabs or

Boehringer Mannheim.

Ribonuclease T1 (RNase T1) and Ribonuclease A (RNase A) were purchased from Sigma Chemical Company.

T4 DNA polymerase, Calf Intestinal Phosphatase (CIP) and Proteinase K were supplied by Boehringer Mannheim.

SP6 and T7 RNA polymerase and T4 DNA ligase were purchased from Promega Corporation.

AMV Reverse transcriptase was purchased from Pharmacia.

MMuLV Reverse transcriptase was purchased from GIBCO-BRL.

2.1.4 Buffers and Solutions

TE - 10mM Tris-HCl pH 7.5 or 8.0, 0.1mM EDTA

SSC - 150mM NaCl, 20mM sodium citrate, 1.0 mM EDTA, pH 7.0

SSPE - 150mM NaCl, 20mM sodium dihydrogen phosphate, 1.0 mM EDTA, pH 7.4

TAE - 40mM Tris-acetate pH 8.2, 20mM Na acetate, 1mM EDTA pH 7.4

TBE - 130mM Tris, 50mM Boric acid, 2.5mM EDTA pH 8.3

PBS - 136mM NaCl, 2.6mM KCl, 1.5mM K_3PO_4, 8.0mM Na_3PO_4, pH 7.3

Denhardt's - 0.02% (w/v) polyvinyl pyrollidone, 0.02% (w/v) BSA, 0.02% (w/v) Ficoll.

5 X Agarose loading buffer - 0.1% Bromophenol Blue, 40% Glycerol, 5mM EDTA, 50mM Tris-HCL (pH8.0)

2 X Sequencing Loading Buffer - 95% formamide, 10mM NaOH, 20mM EDTA, 0.03% bromophenol blue, 0.03% xylene cyanol.

1 X RNA Formamide Loading Buffer -95% formamide, 20mM EDTA, 0.03% bromophenol blue, 0.03% xylene cyanol.

2.1.5 Bacterial Growth Media

All media were prepared with water purified through the Milli-Q Reverse Osmosis purification system and sterilized by autoclaving.

Preparation of media was the same as that described in Sambrook et al (1989).

Luria Broth and 2 x YT media were used to propagate *E.coli* ED8799, XL1 Blue and DH5-_.

SOC medium and Luria Broth were used for the recovery of cells after electroporation.

Agar plates were made by adding 1.5% Bacto-agar (Difco) to Luria Broth. The antibiotic Ampicillin (Sigma) was added at 50µg / ml once the temperature of the autoclaved solution cooled to approxiamately 40^{o}C.

2.1.6 Radiochemicals

All radiochemicals were purchased from Geneworks (Adelaide, South Australia).

[α -^{32}P] - ATP (specific activity, 3000 Ci / mmole), 2. [α -^{32}P] - dATP (specific activity, 3000 Ci / mmole), 3. [α -^{32}P] - dCTP (specific activity, 3000 Ci / mmole), 4. [α -^{32}P] - rUTP (specific activity, 3000 Ci / mmole), 5. [α -^{33}P] - rUTP (specific activity, 3000 Ci / mmole), 6. [α -^{35}S]dATP (specific activity, 1000 - 1500 Ci / mmole).

2.1.7 Molecular Biology Kits

MEGA prime kit (DNA probe labelling) and cyclone sequencing kits were purchased from Geneworks, Pty, Ltd.

Superscript II kit for cDNA synthesis was purchased from GIBCO-BRL.

Excel DNA sequencing kit was purchased from Astral Scientific.

2.1.8 Oligonucleotides

Crude and reverse phase column (RP-1) purified oligonucleotides used for sequencing and PCR were purchased from DNA Express or Bresatec Ltd. Primers are 5' to 3' as written.

1. C1724 (18MER) : GAA CTG GAA AAG GAG TTC
2. C24 HINT (26MER) : TA GTC GAC CAG CCA GCA GGT TCA TCA
3. C17 HINT (26MER) : TA GTC GAC CAG ACA GGT TGG ACT TGG
4. 3'MCS (24MER) : CCG GAT CCC TGC AGG AAT TCG TCG
5. 3'MCS (18MER) : CCG GAT CCC TGC AGG AAT
6. 5' SALI C1724 (26MER) : TA GTC GAC GAA CTG GAA AAG GAG TTC
7. P1 (18MER) : GGA GGT GGG TCT GAA TGG
8. P2 (18MER) : CCC AGA AAC TCC TAC CAC
9. P3 (18MER) : TCA GTG CTG GGG AGG TAT
10. P4 (18MER) : CCT GCA TTC CTC TTT CTG
11. P5 (18MER) : GCG TGG TAG GAG TTT CTG
12. P6 (18MER) : CCC ATA CCT CCC CAG CAC
13. P7 (18MER) : GCC TGG CCT GGG GAC TCG
14. P8 (18MER) : GGG CAG GAG GGA GGG AGC
15. P9 (18MER) : TGT GCC TTT GTA TCC ATT
16. P10 (26MER) : TA GAA TTC TTC CCC GCA CCA ATA GTT
17. P11 (27MER) : TA GAA TTC TGG TTC TTA AAG GTG GAC A
18. P8SEQ (18MER) : GGG CAG GAG GGA GGG AGC
19. PUC16 (26MER) : TA GAA TTC GGA TCT CAA CAG CGG TAA
20. PUC68 (26MER) : TA GAA TTC CCA ATG ATG AGC ACT TTT
21. P12 (18MER) : CGT TCA TTC TCG AGC CCC

22. PUC16V2 (18MER) RP-1 : GGA TCT CAA CAG CGG TAA

23. PUC68V2 (22MER) RP-1 : ATC GCC AAT GAT GAG CAC TTTT

24. P14 (22MER) RP-1 : TAC GCG GCG ACA TAA ATA GAC G

25. PRT (24MER) RP-1 : TCC ATT CAG ACC CAC CTC CAA AGC

26. P277 (25mer) RP-1 : CTC CCT GCA TTC CTC TTT CTG GTT A

27. P151 (25MER) RP-1 : GTC AGG TAG AGT GAA GGT GAG GCG T

28. PUC 10 (25MER) RP-1 : CGA ACT GGA TCT CAA CAG CGG TAA G

29. P56 (25MER) RP-1 : GAG AGG TTT GTG GTA GCC GAG ATG C

31. (21MER) : AGA CGT ACA TAA GGC TCT CCG

32. Lambda gt11 forward primer (24MER) : GGT GGC GAC GAC TCC TGG AGC CCG

33. Lambda gt11 reverse primer (24MER): TTG ACA CCA GAC CAA CTG GTA ATG

34. T3 (23MER) : GCA ATT AAC CCT CAC TAA AGG GA

35. C13 272 (20MER) : GCC TCC ATT TCT CTC CTT AA

36. C13 330 (21MER) : CGG TGT GAT GAT GTT GTG TCT

37. C13 P439 (21MER) : AGT CTT GCT GCT CGG TGT CGG

38. C13p489 (20MER) : CGG GGC TCC CTA CTC CTT GC

39. C13474 (21MER) : GAA CCC CAG GAG ACC CAA CCA

40. P982 (21MER) : TTC TTA ATG AAG TGT ACC CGC

41. C13 16 (21MER) : GTG GGG AGG GCG GTG GAA ACA

42. P1033 (24MER) : AAA CAC CAG GAG GAG GAG GGG AGG

43. C13 in 282 (21MER) : TGT CTC ACA ACG TGA ACC TGC

44. C13 IN3 (20MER) : CAC CTC TGG AAG TCT CCC TT

45. P899 (20MER) : CTC TTA CTT GCC CGC CCT CA

2.2 General Chemicals

Ajax Chemical Co

Ammonium persulphate, Chloroform, Diethyl ether, Isoamyl alcohol, Propan-2-ol.

BDH

Acetic acid, Bromophenol blue, Butan-1-ol, Chloroform, Dimethyl sulphoxide (DMSO), Formaldehyde, Formamide, Isopropanol, Magnesium Chloride, Polyethylene glycol (6000), Hydrochloric acid, Iso-amyl alcohol, Sodium dihydrogen orthophosphate, Di-sodium hydrogen orthophosphate, Sucrose, Urea.

Bio - Rad

Zetaprobe GT membrane.

Boehringer Mannheim

Bovine Serum Albumin (BSA), EDTA (Ethylenediaminetetracetic acid), glycogen.

F. S. E Scientific

Ethanol (crude 95%).

Gibco - BRL

Agarose low melting temperature.

May and Baker

Dimethyl formamide, Glacial acetic acid, Glycerol, Orthophosphoric acid, Sodium chloride.

Merck

Guanidine Isothyocyanate, HEPES ([4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid].

Miles Laboratories, Inc.

Opitmal cutting temperature embedding compound.

NEN Dupont

Gene-Screen membrane.

Pharmacia

Dextran sulphate, Ficoll 400,Sephadex G-50.

Progen

X-gal.

Sigma

Acrylamide, Ampicillin, rATP, Bis-acrylamide, Amberlite MB-1, DNA (salmon testes), DTT (dithiothreitol), Eosin, Ethidium Bromide, Haematoxylin, IPTG (isopropylthiogalactosidase), B- Mercaptoethanol, Mops, PIPES (1,4- piperazinebis(ethanesulphonic acid), Salmon sperm DNA, Sarkosyl, Sodium acetate, SDS (Sodium dodecyl Sulphate), TEMED (N,N,N,N'-tetramethylethylenediamine), TESPA (3-aminopropyltriethoxysilane), Trizma base, Tween20, Urea, Yeast tRNA.

Tokyo Kasei

Xylene cyanol.

2.3 Methods

2.3.1 General DNA Methods

Phenol extraction was performed as described by Maniatis et al., 1993. DNA was purified and concentrated by ethanol precipitation by the method of Zeugin and Hartley (1985).

2.3.1.1 Agarose Gel Electrophoresis

DNA fragments were size fractionated on agarose gel electrophoresis as described in Sambrook et al.,(1989). Electrophoresis was carried out in horizontal gel running apparatus connected to a Pharmacia power supply. Gels of 55mm x 75mm were used for analyses of restriction digests, PCR products and vector DNA. 1% agarose 1 X TAE gels were run at 100 mA in 1 X TAE buffer. 0.7% agarose TAE gels used in southern analyses were 20 mm x 25 mm in size and were run at 25 mA over night in 1 X TAE buffer. Gels were stained in 0.1% ethidium bromide for 5 minutes and analysed under short wave UV light. Samples were loaded in glycerol loading buffer containing 5 % glycerol, 0.025 % bromophenol blue and 0.025 % xylene cyanol

Sizes of unknown DNA fragments were determined by comparison with marker DNA fragments in HpaII-digested pUC19 DNA, EcoRI-digested bacteriophage SPP-1 DNA and HindIII-digested lambda phage DNA co-electrophoresed with the test DNA samples. All markers were purchased from Geneworks Pty. Ltd.

2.3.1.2 Polyacrylamide Gel Electrophoresis

DNA fragments of less than 700 base pairs were size fractionated on 5-20% (w/v) acrylamide/bisacrylamide gels. Oligonucleotides were electrophoresed using a 20% non-denaturing polyacrylamide gel. Sequencing reactions were analysed using a 0.2 mm thick 5-6% denaturing polyacrylamide gel containing 8M Urea. The electrophoresis was performed at 50 Watts using TBE as the running buffer. The gel was then transferred to 3MM paper and dried using a Bio Rad gel drier and then autoradiographed at room temperature for one to two days depending on the intensity of the signal.

The acrylamide mix was dissolved in 1X TBE and polymerised by addition of 0.1% (w/v) APS and 0.1% (v/v) TEMED. Gels were allowed to polymerise for at least 2 hours and then pre-electrophoresed in 1X TBE buffer for 30 minutes.

2.3.1.3 Plasmid Mini Prep DNA Isolation

To analyse clones containing possible recombinant DNA molecules the method of Ish-Horowicz and Burke (1981) was followed.

2.3.1.4 Large Scale Plasmid DNA Isolation.

Plasmid DNA was prepared by using Qiagen mini or midi kits following the manufacturers protocol.

2.3.2 PCR

DNA was amplified using according to the method of Saiki et al., (1988). PCR was performed in thin-walled polypropylene PCR tubes (Lab Supplies) and in a reaction volume of 25 µl, containing 2.5 µl of 10 X PCR buffer (500 mM KCl, 100 mM Tris.Cl [pH 8.3], 1% gelatin), 1.0 µl of a 5 mM dNTP mix, 0.5 mM-3.0 mM $MgCl_2$, 200 ng of each primer, 100 pg of DNA, 1 unit of Taq polymerase and water to a final volume of 25 µl. The reaction was then overlayed with 15 µl of mineral oil. Optimal $MgCl_2$ concentrations were determined for each primer pair by varying the $MgCl_2$ concentrations until a specific product was obtained in maximal amounts.

2.3.2.1 PCR Cycling Parameters

PCR was performed in a Perkin-Elmer robocycler. All PCRs began with a hot start by heating the reaction to 94°C for 5 minutes and then holding at 75°C before addition of Taq polymerase to decrease non-specific binding of primers to the DNA template. Each PCR cycle consisted of three steps; a denaturation step to denature double-stranded DNA (usually 95°C for 30 seconds), an annealing step to allow the primers to hybridise to the DNA template and an extension step at 72°C to allow the Taq polymerase to synthesise double-stranded DNA. The annealing temperature of each PCR was determined by the

primer with the lowest Tm (melting temperature) calculated using Oligo 4 software default parameters. The length of time of the extension step was determined by the length of the DNA template. Approximately thirty seconds was added to the extension time for each kilo base of DNA.

2.3.2.2 Primer Design

Primers for PCR were designed using Oligo-4 and DNAStar software which excludes primers that form secondary structures and primer dimers. PCR primers ranged in length from 18 to 25 nucleotides and were designed to contain at least a 50% GC content. Primers did not contain stretches of purines or pyrimidines greater in length than three. Annealing temperatures of upper and lower primers were matched as close as possible. Primers were synthesised without a 5' phosphate attached. In order to clone PCR products, primers were phosphorylated before addition to the PCR. Primers of an RP-1 (Reverse Phase column) purity were found to be adequate for standard PCR. To obtain 100% full-length primers for primer extension analysis RP-1 oligonucleotides were purified through a 20% polyacrylamide non-denaturing gel. 1µg of primer was loaded in Formamide loading buffer and size fractionated by electrophoresis. Primers were excised from the gel and eluted over night in 300 µl of 1X TE at 37°C followed by ethanol precipitation.

2.3.3 DNA Sequencing

Double stranded DNA templates were sequenced using the dideoxy chain termination method (Sanger et al., 1980; Messing et al., 1981).

2.3.3.1 Isothermal DNA Sequencing

The T4 Isothermal sequencing kit (USB) was used following the manufacturer's conditions.

2.3.3.2 Cycle Sequencing

PCR-amplified DNA and recombinant DNA were sequenced using the Cyclone Sequencing kit (Geneworks) and the Excel Sequencing kit (Epicentre) following the manufacturer's protocol.

2.3.3.3 Sequencing Primers

Sequencing primers were designed using Oligo-4 software and were purified to an RP-1 (reverse phase) standard. Crude primers were purified through a CL6B spin dialysis column before use.

2.3.4 Southern Transfer

PCR fragments, vector and genomic DNA fragments were transferred to Bio-rad Zetaprobe membrane by the modified method of Southern (1977). DNA was size fractionated in a 1% agarose gel, ethidium bromide stained and photographed. DNA greater than 1 kb was nicked by soaking the top of the gel in 0.1 M HCl for 10 minutes. DNA was transferred to Bio-Rad Zetaprobe membrane under vaccuum for two hours in 0.4 M NaOH as described by Reed and Mann (1985). DNA was cross-linked to Zetaprobe membrane by Ultra Violet light using a Bio-Rad Ultra Violet Stratalinker. The membrane was allowed to dry at 60^{o}C for 30 minutes.

2.3.4.1 Hybridisation of Radiolabelled Probes

Prehybridisation consisted of adding 10 mls of Denhardts - hybridisation solution to the membrane in a Hybaid bottle. The membrane was incubated using a hybaid incubator at 42^{o}C for 2 hours and then dispensed. Labelled probe was mixed with 10 mls of hybridisation solution and then added to the Hybaid bottle and incubated over night at 42^{o}C. Membranes were washed at a stringency of 2 X SSC / 0.1 % SDS at room temperature for 30 minutes, then 2 X SSC / 0.1 % SDS at 65^{o}C for 30 minutes and finally at 0.1 X SSC / 0.1 % SDS at 65^{o}C for 30 minutes. Radioactive signal was then detected by autoradiography or use of a phosphor-image screen.

2.3.5 Recombinant DNA Methods

2.3.5.1 Restriction Enzyme Digestion

Cleavage of DNA was carried out in a reaction volume of100 µl containing 5 units of restriction enzyme, 1 X manufacturers buffer and incubated at 37°C or as stipulated by the manufacturer for at least 2 hours.

2.3.5.2 End Filling of DNA fragments

To create blunt ends, DNA containing 5' overhangs were end-filled using either Klenow fragment or T4 DNA polymerase as described in Maniatis et al., (1989), following the manufacturers conditions.

2.3.5.3 DNA Fragment Isolation from Agarose Gels

PCR-amplified DNA, restriction fragments and vectors were purified from agarose using one of the following:

1) 1% low melting point (LMP) agarose extraction procedure.
DNA was fractionated on 1% LMP by gel electrophoresis, size-selected by comparison to known molecular weight markers and excised from the gel. The gel slice was heated in an eppendorf at 68°C in 1 X TE buffer for 15 minutes, phenol (pH 8.0) extracted and ethanol precipitated.

2) Isolation from (TAE) agarose gels using a 'Geneclean' kit obtained from Bresatec and used according to the manufacturers conditions.

3) Elution from 1% agarose gel slices or 5 - 20% polyacrylamide gel slices. The gel slices were left in 1 X TE / 0.1% SDS solution overnight followed by ethanol precipitation.

4) Excised LMP agarose (containing DNA fragments) was enzymatically cleaved into smaller sugar molecules using Agarase (Life Technologies) following the manufacturers conditions. Followed by a phenol extraction and then ethanol precipitated.

2.3.5.4 Vector DNA Preparation

Vector DNA was linearised with the appropriate restriction enzyme(s), phenol extracted and ethanol precipitated. Following Klenow (required for blunt end cloning only) and CIP treatment, the vector was again phenol extracted and ethanol precipitated. DNA was then fractionated by (1 X TAE) agarose gel electrophoresis. Uncut vector DNA was separated from linear DNA by extraction from low melting point agarose. Purified DNA was checked for contaminating, undigested DNA by fractionation via (1 X TAE) agarose gel electrophoresis and, if required, extracted a second time.

2.3.5.5 Dephosphorylation of vector DNA

Calf Intestinal Phosphatase was used to dephosphorylate the 5' phosphate group according to Maniatis et al., 1992.

2.3.5.6 Insert DNA Preparation

Insert DNA was excised by restrictiion digestion and then phenol extracted and ethanol precipitated. Following Klenow treatment (required for blunt end cloning only) the insert was again phenol extracted and ethanol precipitated. Insert DNA was separated from plasmid DNA or non-specific PCR fragments by extraction from low melting point agarose. Inserts smaller than 500 base pairs were eluted from 5% (20:1 acrylamide:bis-acrylamide) non-denaturing acrylamide. PCR-amplified DNA products were treated with Proteinase K (50 µg / ml) for 30 minutes at 37°C to denature Taq DNA polymerase.

2.3.5.7 Ligation of Vector and Insert DNA Molecules

Vector DNA (20 ng – 100 ng) was ligated to insert DNA in a 10 µl volume containing 1 X ligase buffer (Promega), 10% PEG 6000, 2 units of T4 bacteriophage DNA ligase (Promega) and sterile MQ water. The concentration of insert DNA depended on its size and in keeping with a 1:1 and 1:3 vector to insert molar ratio. The reactions were performed at 18°C overnight in a PTC-100 thermocycler (Bresatec). 1 unit of an appropriate restriction enzyme was added to the reaction to decrease the amount of vector to

vector religation. This step could only be administered where the enzyme did not cut the insert internally.

2.3.6 Transformation of Bacterial strains

2.3.6.1 Calcium Chloride Competent Cells

Cells (XL1 Blue, ED8799, DH5α) were made competent and transformed with plasmid DNA using the $CaCl_2$ method as described by Dagert and Ehrlich (1979). Half of a ligation reaction mix was added to the cells before the heat shock step. Cells were resuspended in 1ml of LB and grown for 30 minutes, rotating at 37°C before plating onto LB agar plates containing an antibiotic to select for transformants.

2.3.6.2 Preparation of Competent Cells For Electroporation

Bacterial cells were propagated in 500 mls of LB until the $OD_{600} = 0.9$. Cells were spun at 3,500 rpm in a JA-20 fixed angle rotor and the supernatant discarded. Residual salt was removed by resuspending the pellet thoroughly in sterile 10% glycerol at 4°C. The cells were then harvested by spinning at 3,500 rpm at 4°C. The pellet was resuspended in 2 mls of sterile 10% glycerol at 4°C and aliquoted into 80µl lots, frozen with dry ice and stored at -80°C.

2.3.6.3 Electroporation

Ligation reaction mixes were phenol extracted, ethanol precipitated and resuspended in 10 µl of Milli-Q water. 1 µl was then mixed with 40 µl of electrocompetent cells and left on ice for one minute. Cells were then placed into a 0.2 ml Bio-Rad electroporation cuvette and pulsed using a Bio-Rad electroporator set at the following parameters: Resistance: 200 Ohms, Capacitance: 25 µfds, Set Volts: 250 volts. Cells were quickly resuspended in 1ml of SOC medium and grown for 30 minutes at 37°C.

2.3.6.4 Glycerol stocks

Glycerol stocks of transformed bacteria were made by adding 400 µl of an overnight culture to 400 µl of 80% glycerol and frozen at -80°C.

2.3.7 RNA Methods

2.3.7.1 Total RNA Isolation

Total RNA was prepared from sheep wool follicles and outer root sheath cell lines following the method described by Chomczynski and Sacchi (1987).

2.3.7.2 mRNA Isolation

mRNA was isolated from total RNA using the Poly-A Tract system from Promega corporation following the manufacturer's protocol. The final eluate was ethanol precipitated using sodium acetate and 10 μg of Glycogen.

2.3.7.3 Northern Transfer

Sheep follicle RNA was fractionated through a 1% agarose gel containing formaldehyde as described by Hansen et al., (1989). RNA was transferred to zeta probe membrane in 0.4 M NaOH under vacuum for 2 hours.

2.3.7.4 Ribo-probe synthesis

Cloned fragments were transcribed with SP6, T3 or T7 RNA polymerase according to the method of Krieg and Melton (1987). cRNA probes were labelled to high specific activity by incorporation of α- ^{33}P- rUTP.

2.3.7.5 RNA protection analyses

RNA protection analyses was performed as described by Krieg and Melton (1987). To ascertain probe stability and efficiency of the RNase treatments during overnight hybridisation, probe transcripts were incubated with yeast tRNA. Post-hybridisation, these reactions were treated with RNase A / T1 or left untreated and fractionated alongside experimental samples.

2.3.7.6 Primer extension analysis

Primer extension analysis was performed as described by McKnight et al., (1981).

2.3.7.7 *In situ* Hybridisation Analysis

Radio-labelled riboprobes were hybridised to tissue sections *in situ* following the method of Cox et al., (1984) with the modifications of Powell and Rogers (1990). Tissues were stained using Heamotoxylin (Nakanishi et al., 1975) or SACPIC (Nixon et al., 1993).

2.3.8 Gene Isolation Methods

2.3.8.1 cDNA Synthesis for 3' R.A.C.E

cDNA was synthesised using sterile PCR pipettes, pipette tips, eppendorfs and stock solutions kept in a separate room from PCR analyses areas to prevent contamination from amplified PCR products. Approximately 1.0 µg of wool follicle mRNA was mixed with 2.0 µl of 10 X PCR amplification buffer, 2.0 µl of each dNTP (to give a final concentration of 1 mM each), 0.25 µl of ribonuclease inhibitor (40,000 units / µl; Promega), 1.0 µl 50 mM $MgCl_2$, 1.0 µl of 0.1 M DTT, 100 ng of the 3' Multiple Cloning Site-oligo dT (3' MCS dT) primer 5' - CCG GAT CCC **TGC AG**G AAT TCG TCG (dT)17 YX - 3' (Y = A, G or C and X = A, C, G or T), 2.0 µl of AMV reverse transcriptase (25U / µl) and MQ water to a final volume of 20 µl. The cDNA reaction was vortex mixed, heated at 42°C for 30 minutes and then heated at 50°C for 30 minutes to over come premature termination of the enzyme. AMV reverse transcriptase was inactivated by heating at 95°C for 5 minutes. The cDNA reaction mix was diluted to 100 µl in 1 X TE. Only 1.0 µl of the final cDNA reaction mix was used for each subsequent 3' RACE PCR.

To ensure that high quality full length first strand cDNA was being synthesised, cDNA was radioactively labelled by adding 1 µl of α– ^{32}P- dCTP (3000 Ci / mmol) to the reverse transcription reaction. Radioactively labelled cDNA was size-separated on an alkaline denaturing gel (Maniatis et al., 1989) and detected using phosphor-image technology.

2.3.8.2 3' RACE-PCR Conditions for *Hoxc-13*

To obtain all 3' coding and non-coding sequence located downstream of the homeobox, the technique of 3' RACE described by Frohman et al., (1988) was employed. One twentieth of the cDNA reaction was used as template for the 3' RACE reaction. Oligo 4 software was used to design a 5' species-specific oligonucleotide from the homeobox of *Hoxc-13* using the sequence from clone C24 (Fig. 6.1A). The oligonucleotide sequence was 5' - TA GTC GAC CAG CCA GCA GGT TCA TCA - 3' and was named C24Hint. A SalI restriction endonuclease site was incorporated at the 5' end which enabled the 3' RACE-PCR product to be cloned into the SalI site of pBSc (KS) by sticky end ligation. SalI did not cleave the 3'RACE-PCR product internally (data not shown). The sequence of the 3' primer used in all 3'RACE-PCR reactions was 5' - CCG GAT CCC TGC AGG AAT TCG TCG - 3' and named multiple cloning site primer (MCS). MCS contained complementary sequence to the 3' MCS dT oligonucleotide (excluding the poly T tail) and hybridised to the multiple cloning site region. Control PCRs were included to determine whether PCR products were obtained from follicle cDNA or genomic DNA

A magnesium concentration of 1.5 mM was empirically determined and the following cycling parameters used in each PCR;

1. 94°C (45secs), 72°C (2 minutes)

The PCR was held at 72°C while the TAQ polymerase and the 5' primer C24HINT was added. This was followed by:

2. Four cycles of 94°C (45secs), 56°C (2 minutes), 72°C (2 minutes)
3. The PCR was held at 72°C while the 3'MCS primer was added
4. 35 cycles of 94°C (45secs), 63°C (2 minutes), 72°C (2 minutes)
5. A final extension step of 72°C (7 minutes).

2.3.8.3 3' RACE-PCR Conditions for Ovine *Barx2*

The isolation of the 3' cDNA end for ovine *Barx2* was achieved by 3' RACE-PCR following the method of Frohman (1988). To ensure that *Barx2* mRNA transcripts were present, the RNA sample described in Section 3.2.1 was used.

A magnesium concentration of 1.5 mM was used with the following cycling parameters in each PCR.

1. 94°C (45 seconds), 72°C (2 minutes).

Each PCR was held at 72°C while the TAQ polymerase and the 5' primer, C17Hint were added.

2. 4 cycles of 94°C (45 seconds), 56°C (2 minutes), 72°C (2 minutes)
3. The PCR was held at 72°C while the 3'MCS primer was added
4. 35 cycles of 94°C (45 seconds), 63°C (2 minutes), 72°C (2 minutes)
5. A final extension step of 72°C (7 minutes)

2.3.8.4 RNA Preparation for 5' RACE

The 'New 5' R.A.C.E.' protocol as described by Frohman, 1992, was used to amplify a partial fragment of the 5' end of *Hoxc-13* with the following changes. The procedure requires a small RNA oligonucleotide that is not expressed in the tissue under study to be ligated to the 5' end of wool follicle Poly A (+) RNA. A 110 base fragment from the pUC19 base vector (nucleotides 3054-3164 genbank Accession No: M77789) was cloned into the SacI site of the pBSc KS vector and the clone named pUC19-110. A 110 bp RNA oligonucleotide in the sense orientation was synthesised from clone pUC19-110 using the T7 RNA polymerase, purified by elution from a 5% polyacrylamide gel slice and then ligated to approximately 1 µg of wool follicle poly A(+) RNA.

2.3.8.5 cDNA Synthesis for 5' R.A.C.E

cDNA used in the 5' R.A.C.E technique was synthesised using the components of the Supescript II kit following the manufacturer's protocol with the following changes

1) Inclusion of the 10 X PCR buffer during the hybridisation step of primer to RNA

2) Hybridisation cycling parameters were: 80°C (5 minutes), 42°C (120 minutes), 55°C (10 minutes) and 75°C (10 minutes).

2.3.8.6 5' New R.A.C.E

A gene-specific oligonucleotide named PRT (nucleotides 1216-1238; see Fig. 6.10 for reference), designed from the *Hoxc-13* 3' non-coding region was used as the primer to synthesise the first strand of cDNA. PCR was performed using the 5' primer, pUC16, designed to hybridise to the 110 bp pUC19 RNA oligonucleotide and a nested *Hoxc-13* primer named P4 which binds upstream from the PRT primer (Fig. 6.5 B). Cycling parameters used in the first PCR were

1. 98°C (5'), 72°C (40'), add one unit of Taq polymerase (hot start)
2. 5 cycles [94°C (30"), 42°C (2'), 72°(3')] Add pUC16
3. 35 cycles [94°C (30"), 45°C + 0.2°C / cycle (1'), 72°(3')]
3. 72°C (15')

200 ng of each primer and a magnesium concentration of 1.5 mM were used in each PCR. One twentieth of the PCR reaction was used as a template for a second nested PCR using a magnesium concentration of 1.5 mM (Fig. 6.5 B). 200 ng each of a second set of nested primers, named pUC68 and P14 (nucleotides 1014-1031; Fig. 6.10) were used. Cycling parameters used in the nested PCR were

1. 98°C (5'), 72°C (40'), add one unit of Taq polymerase
2. 35 cycles [94°C (30"), 45°C + 0.2°C / cycle (1'), 72°(3')]
3. 72°C (15')

2.3.8.7 cDNA library Screening

A lambda ZAP cDNA library, enriched for ovine follicle cDNA's (Clontech) was screened for full-length homeobox cDNA's using gene-specific probes. Individual clones were isolated following the method described in Maniatis et al., (1989). To ascertain the size of individual clones, inserts were PCR-amplified using primers designed to bind to the flanking regions of the lambda zap vector arms.

2.3.8.8 Lambda Genomic Library Screening

To isolate a genomic clone containing a full-length homeobox gene, a lambda EMBL3 genomic DNA library (Clontech) was screened at a high density (approximately 30,000 plaques per 15 cm filter) by the plaque hybridisation method described by Benton and Davies (1977). Individual clones were isolated following the method described in Maniatis et al., (1989).

Chapter 3

Isolation of Partial Ovine homeobox cDNAs

Chapter 3 : Isolation of Partial Ovine Homeobox cDNAs

3.1 Introduction

The wool follicle contains at least 10 different cell types that make up the ORS, IRS, cuticle, cortex, dermal papilla and in some sheep the medulla. The factors that determine the fate of each cell type in the follicle are unknown. One group of genes that are master regulatory genes involved in organogenesis and the determination of cell fate are the homeobox (Hox) genes. The expression of several Hox genes in the hair follicle has recently been documented in the literature (see section 1.15) and this group of genes may be involved in determining the cell fate of hair follicle keratinocytes.

Once hair follicle keratinocytes in the bulb are set on a predetermined developmental pathway, several families of hair keratin differentiation-specific genes are activated in complex spatial and temporal patterns of expression (for review see, Powell and Rogers, 1997). How these genes are regulated is not well understood, however, because of their ability to regulate differentiation-specific genes the Hox genes are potential candidates for the regulation of the hair keratin genes.

Prior to the initiation of this project no one had undertaken a dedicated search for Hox genes in the hair follicle and only four Hox genes were reported to be expressed in the hair follicle. Two were *Antennapedia*-type Hox genes, *Hoxb-2*, expressed in the murine hair follicle placode (Whiting et al., 1991) and *Hoxc-8* expressed in the dermal papilla of the

anagen hair follicle (Bieberich et al., 1991). The other two Hox genes were members of the POU family, Skn-1a and Skn-1I, expressed in the cortical cells of mouse anagen hair follicles (Andersen et al., 1993). Because vertebrates contain at least 38 *Antennapedia*-type Hox genes it was logical to first screen the wool follicle for the expression of more of these genes.

This chapter details the isolation, sequencing and classification of ten partial cDNAs representative of three homeobox classes expressed in the wool follicle.

Results

3.2 Isolation of Ovine Wool Follicle Partial Hox cDNAs

3.2.1 Amplification of Ovine Wool Follicle Partial Hox cDNAs

Partial homeobox cDNAs from the wool follicle were amplified using RT-PCR using degenerate oligonucleotides designed to conserved amino acid regions of the *Antennapedia* homeobox gene class (Fig. 3.1). Two conserved amino acid regions were found by sequence comparison of previously reported vertebrate *Antennapedia*-like sequences (Burglin et al., 1989). The degenerate primers to the conserved ELEKEF (helix 1) and WFQNRR (helix 3) motifs each had a 128 - fold degeneracy (Fig. 3.1). The length of DNA that includes the two conserved homeobox motifs is 117bp. The two degenerate primers used to amplify this 117bp region each contain eight extra nucleotides at the 5' end to aid in cloning. The nucleotides added to the 5' end of the primers resulted in a PCR product of 133bp in length.

Figure 3.1 : Conserved Motifs of Homeobox Genes.

Homeodomains from different species share sequence similarities. *Abd-B* (Karch et al., 1985; Regulski et al., 1985), *Antp* (Garber et al., 1983; Scott et al., 1983), *Dfd* (Hazelrigg et al., 1983) and *Lab* (Mlodzik et al., 1988) are *Drosophila* genes and *Hoxa-4* (Duboule et al., 1986) and *Hoxa-9* (Rubin et al., 1987) are mouse genes. The ELEKEF and WFQNRR motifs boxed are perfectly conserved and were used to design degenerate primers. The degenerate primers used to amplify partial homeobox cDNAs by RT-PCR were;

P1 = 5'- TA GAATTC GA(A/G) (C/T)TX GA(A/G) AA(A/G) GA(A/G) TT-3'

(from the ELEKEF motif) and

P2 = 5'- TA GAATTC CC(G/T) XC(G/T) (A/G)TT (T/C)TG (A/G)AA CCA-3'

(from the WFQNRR motif).

Both primers have a 128-fold degeneracy. The size of a PCR product expected using these primers is 133 bp; 16 bp contain the introduced EcoRI cloning sites and 117 bp contain the ELEKEF and WFQNRR terminal motifs. 1000 ng of each primer was used for PCRs. An EcoR1 site was engineered into the 5' end of both oligonucleotides so that RT-PCR products could be cloned using sticky end ligation. The positions of the α-helices in the homeodomain are represented by the shaded rectangular boxes. X = A, C, G or T.

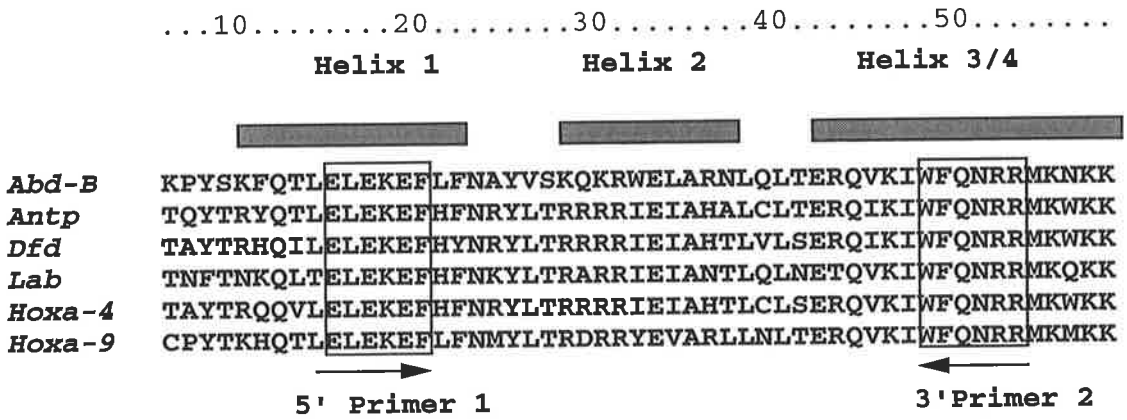

...10.....20.....30.....40.....50.....
Helix 1
Helix 2
Helix 3/4
Abd-B KPYSKFQTLELEKEFLFNAYVSKQKRWELARNLQLTERQVKIWFQNRRMKNKK
Antp TQYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKK
Dfd TAYTRHQILELEKEFHYNRYLTRRRRIEIAHTLVLSERQIKIWFQNRRMKWKK
Lab TNFTNKQLTELEKEFHFNKYLTRARRIEIANTLQLNETQVKIWFQNRRMKQKK
Hoxa-4 TAYTRQQVLELEKEFHFNRYLTRRRRIEIAHTLCLSERQVKIWFQNRRMKWKK
Hoxa-9 CPYTKHQTLELEKEFLFNMYLTRDRRYEVARLLNLTERQVKIWFQNRRMKMKK
5' Primer 1
3'Primer 2

Fig 3.2 : Synthesis of cDNA and Amplification of the Homeobox cDNAs.

A. cDNA Template

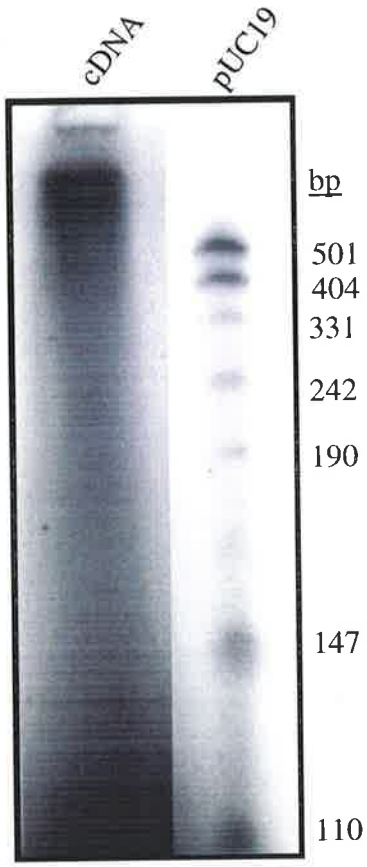
The size of the cDNA for RT-PCR was checked by incorporation of α^{32}P-dCTP during synthesis. Single-stranded, heat denatured cDNA (95^{o}C, 5 minutes) was size-separated through a denaturing 6% polyacrylamide gel (19:1). The size of the cDNA transcripts were compared to radiolabelled HpaII-digested pUC19 markers on the right.

B. Amplification of the Homeobox of Antennapedia Class Genes by RT-PCR

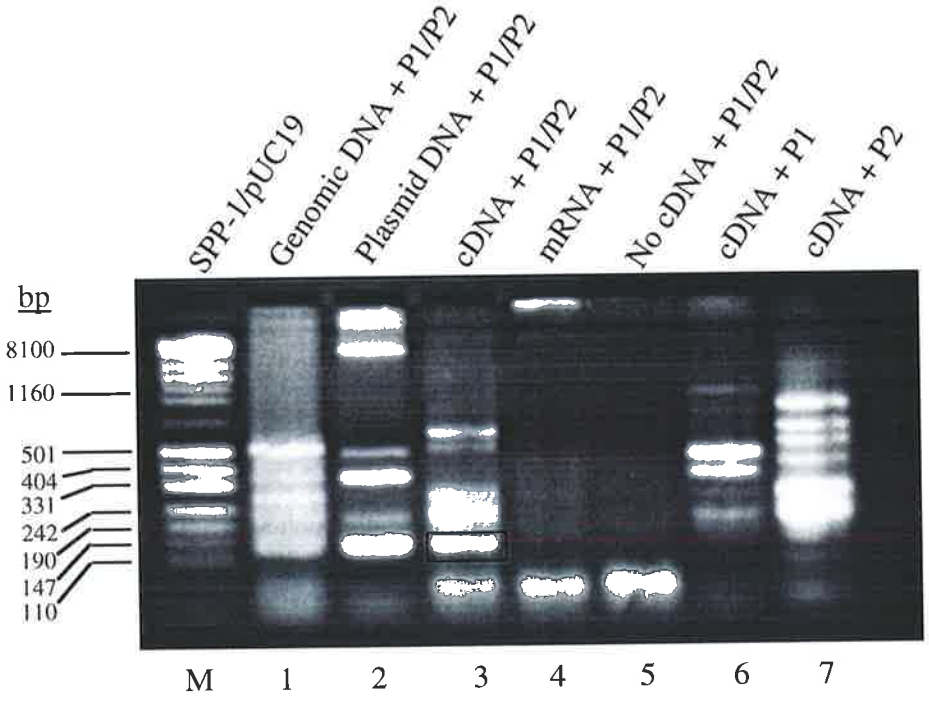
Partial homeobox cDNAs were amplified by RT-PCR. PCR products were size-separated on a 2% agarose gel, stained with ethidium bromide and detected under UV light. The PCRs were performed using cDNA, genomic DNA, mRNA or plasmid control DNA in combination with the primers indicated above the gel (lanes 1-7). The plasmid control contains a homeodomain that includes ELEKEF and WFQNRR motifs. A 133 bp PCR product is boxed.

P1 = 5' degenerate primer designed from the ELEKEF motif of the Antennapedia homeobox. P2 = 3' degenerate primer designed from the WFQNRR motif of the Antennapedia homeobox. Size of SPP-1 and pUC19 markers are shown on the left of the gel (M).

A


cDNA
pUC19
bp
501
404
331
242
190
147
110

B


SPP-1/pUC19
Genomic DNA + P1/P2
Plasmid DNA + P1/P2
cDNA + P1/P2
mRNA + P1/P2
No cDNA + P1/P2
cDNA + P1
cDNA + P2
bp
8100
1160
501
404
331
242
190
147
110
M
1
2
3
4
5
6
7

Total wool follicle RNA was obtained from adult sheep that had been infused with the amino acid L-cysteine for seven days (Fratini et al., 1994). Poly A(+) RNA was isolated and used to synthesise cDNA greater than 500 bases (Fig. 3.2 A). The majority of the cDNA was longer than 500 bases.

PCR conditions were optimized for the degenerate primer pair using a control plasmid that contained a homeodomain and the ELEKEF and WFQNRR sequence motifs (a gift from Paul Moretti, Hanson Centre for Cancer Research). RT-PCR cycling parameters were;

3 cycles of;

1. 94°C for 45 seconds,
2. 45°C for one minute and
3. 72°C for one minute followed by a further 37 cycles of;
4. 94°C for 45 seconds,
5. 56°C for one minute and
6. 72°C for one minute.

To ensure the amplified products were of full length, a final extension cycle consisted of one cycle at 56°C for two minutes and then 72°C for seven minutes. Negative controls were included to identify any contamination of amplified products from the control plasmid or genomic DNA. PCRs consisted of the following templates and primers (Fig. 3.2 B);

Positive controls;

1. Genomic DNA with both primers

2. Plasmid DNA with both primers

PCR experiment;

3. cDNA with both primers

Negative Controls;

4. mRNA (no reverse transcriptase) with both primers.

5. no cDNA with both primers

6. cDNA with primer 1

7. cDNA with primer 2

Control reaction 4 ensured that any amplified RT-PCR product was not due to contamination of genomic DNA.

3.2.2 RT-PCR Result

A 133 bp product was amplified from cDNA using the two degenerate primers ELEKEF (P1) and WFQNRR (P2)(Fig. 3.2 B; lane 3). In the same PCR reaction, several amplified products were produced which were not of the expected size of 133 bp suggesting that the primers were hybridising non-specifically. However, the 133 bp product was absent from the control PCRs which contained the cDNA template and either primer one or two (Fig. 3.2 B; lane 6 and 7) confirming that the 133 bp product in lane 3 was amplified by both primers and was not a spurious product obtained by the non-specific binding of either primer.

A 133 bp product and multiple non-specific products were amplified from genomic DNA (Fig. 3.2 B; lane 1) and plasmid DNA (Fig. 3.2 B; lane 2). No products were detected in the PCR which contained mRNA that had not been treated with reverse transcriptase, verifying that the 133 bp product obtained by RT-PCR was not due to genomic DNA or plasmid contamination.

3.3 Cloning and Sequencing of Partial Homeobox cDNAs

An EcoR1 restriction site was designed into the 5' end of the two degenerate oligonucleotides ELEKEF and WFQNRR enabling the 133 bp PCR product to be cleaved with EcoR1 and sticky end-ligated into an EcoR1-digested pGEM 7Zf (+) vector. The recombinant DNA was then transformed into XL-1 blue by electroporation. Approximately 300 colonies were obtained, of which fifty were screened. Twenty of these contained an insert of 133 bps (data not shown). The twenty recombinant clones were manually sequenced in both directions.

3.4 Classification of Wool Follicle Partial cDNA Clones

To verify that the inserts contained a partial homeobox, a 'Fasta' sequence search was performed on DNA sequence using the data base Genbank through ANGIS (Australian National Genomic Information Service). Some of the clones isolated were found to be duplicates. Ten different genes were represented by the 20 clones and were grouped into three homeobox gene classes. Three genes (*Hoxb-9*, *Hoxc-9* and *Hoxc-13*) belonged to the *Abd-b*

Fig 3.3 : Classification of the Nucleotide and Amino Acid Sequences for the 21 Ovine RT-PCR Clones.

The sequence of the clones and published murine sequences have been compared to human homeobox sequences and the sequence identity is expressed as a percentage (% Id) at the right of each clone. Values given for the sequence identity were calculated using the sequences within the boxed region. Sequence within the degenerate oligonucleotide regions were omitted from calculations. Dashes represent sequence similarity to the sequence, shown in full, typically the human sequence, except for *Hoxc-9* where it is mouse. Ovine *Hoxc-9* was compared to murine sequence due to the unavailability of human sequence. O = Ovine (Ovis aries), H = human (Homo sapiens), M = mouse (Mus musculus). All homeobox sequences and their references can be found in Duboule (1994).

Hoxa-4

Clone #	Species	Amino Acid Sequence Comparison
21	O	----------
	H	ELEKEFHFNRYLTRRRRIEIAHTLCLSERQVKIWFQNRR
	M	----------

Clone #	Species	DNA Sequence Comparison	% Id
		10 20 30 40 50 60 70 80 90 100 110 117	
21	O	---T----A- -------[--- ---T--C--- ---------- ----T----- ---------- ---------- ----C----- ---------- ----------]------A-- T--TC--	95
Published	O	---------- -A------[--- ---T--C--- ---------- ----T----- ---------- ---------- ----C----- ----GC---- ----------]---------- ---A--G	90
	H	GAGCTGGAGA AGGAGTTCCA CTTCAATCGC TACCTGACCC GGCGGCGCCG CATCGAGATC GCCCACACGC TCTGTTTGTC TGAGCGCCAG GTCAAGATCT GGTTTCAGAA CCGGAGA	
	M	--A------- ----A--[--- ---T--C--- ---------- ---------- ---------- ---------- ----C----- G---------- ----------]---------- -------	95

Hoxb-2

Clone #	Species	Amino Acid Sequence Comparison
23	O	-F----------
28	O	----------
	H	ELEKEFHFNKYLCRPRRVEIAALLDLTERQVKVWFQNRR
	M	----------

Clone #	Species	DNA Sequence Comparison	% Id
		10 20 30 40 50 60 70 80 90 100 110 117	
23	O	--GT-C---- ----G-[--- ---------- ---------- ----G--A-- ---G------ ---------- ---------- ---------- --G-----G-]----C--A-- ---C-	94
28	O	--------A- ----G-[--- ------C--- AC-TGTGT-- ----G--A-- ---G------ ---------- ---------- ---------- --G-----G-]----C----- T--A--G	84
	H	GAACTGGAGA AGGAATTCCA CTTTAATAAG TACCTGTGCC GGCCACGCCG CGTCGAGATC GCGGCCTTGC TGGACCTCAC CGAAAGGCAG GTCAAAGTCT GGTTTCAGAA CCGGCGC	
	M	--G------- ----G-[--- ---C------ ---------- ----G--T-- ---------- --T------- ---------- ---------- ----------]----C----- ---A---	95

Hoxb-6

Clone #	Species	Amino Acid Sequence Comparison
T1C4	O	----------
36	O	----------
16	O	----------G---------S-
22	O	----------S-
	H	ELEKEFHYNRYLTRRRRIEIAHALCLTERQIKIWFQNRR
	M	----------

Clone #	Species	DNA Sequence Comparison	% Id
		10 20 30 40 50 60 70 80 90 100 110 117	
T1C4	O	---------- ----A----- ---------- ---------A ---------- ---------- --------G- -C-----C-- A--A--A--- --------C- ----T----- ---CA-	90
36	O	---------- ----A----- ---------- ---------A ---------- ---------- --------G- -C-----C-- A--A--A--- --------C- ----T----- ---CA-G	90
16	O	--AT------ ---------- ---------- ---------A ---------- ---------- --------G- -C-----C-- A--AG-A--- --------C- ---------- -A-CA-G	89
22	O	-----A--A- -A-------- ---------- ---------- ---------- ---T------ ---------- ---------- ---AC----- --------C- -------A-- -A-CA-G	95
	H	GAGCTGGAGA AGGAGTTTCA CTACAATCGC TACCTGACGC GGCGGCGGCG CATCGAGATC GCGCACGCCC TGTGCCTGAC GGAGAGGCAG ATCAAGATAT GGTTCCAGAA CCGACGC	
	M	----------A----- ---T------ --------C- -C-------- ---A------ ---------- ---------- C---C----- --------T- ----T----- ---G---	91

Hoxb-8

Clone #	Species	Amino Acid Sequence Comparison
14	O	-F----------
31	O	----------
	H	ELEKEFLFNPYLTRKRRIEVSHALGLTERQVKIWFQNRR
	M	----------

Clone #	Species	DNA Sequence Comparison	% Id
		10 20 30 40 50 60 70 80 90 100 110 117	
14	O	--AT-T--A- ----A----- ---------- ---C------ -C-------- G--------- -----T--A- ---------- ---------- --------T- -------A-- ---CC--	93
31	O	---T------ ---------- ---------- ---------- -C-------- G--------- -----T--A- ---------- ---------- --------T- ----T----- ---CC--	94
	H	GAGCTGGAGA AGGAGTTCCT ATTTAATCCC TATCTGACTC GTAAGCGGCG AATCGAGGTA TCGCACGCCC TGGGACTGAC AGAGAGACAG GTCAAAATCT GGTTCCAGAA CCGGAGG	
	M	---------- ---------- ---------- ---------- -C------A- G--------- --------G- ---------- ---------- ---------- ---------- T-----A	95

Hoxb-9

Clone #	Species	Amino Acid Sequence Comparison
35	O	----------
RT1C3	O	----------
	H	ELEKEFLFNMYLTRDRRHEVARLLNLSERQVKIWFQNRR
	M	----------

Clone #	Species	DNA Sequence Comparison	% Id
		10 20 30 40 50 60 70 80 90 100 110 117	
35	O	--AT-G---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----C--A-- ---AA--	100
RT1C3	O	--AT-G---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----C--A-- ---AA--	100
	H	GAGCTAGAGA AGGAGTTTCT GTTCAATATG TACCTCACCA GGGACCGTAG GCACGAAGTG GCCAGACTCC TCAATCTGAG TGAGAGACAA GTCAAAATCT GGTTTCAGAA CCGGCGG	
	M	---------- -------C-- ---T------ ---------- ---------- ---------- ---------- ---------- ---------- ---------- -------	98

Hoxc-4

Clone #	Species	DNA Sequence Comparison
C25	O	----------
	H	ELEKEFHYNRYLTRRRRIEIAHSLCLSERQIKIWFQNRR
	M	----------

Clone #	Species	DNA Sequence Comparison	% Id
		10 20 30 40 50 60 70 80 90 100 110 117	
C25	O	---C-T---- -G---------- ------T--- ---------- ----A----- ---------- ---------- ---------- ---------- ---------- -------G-- ---G--G	98
	H	GAATTAGAGA AAGAGTTTCA TTACAACCGC TACCTGACCC GAAGGAGAAG GATCGAGATC GCCCACTCGC TGTGCCTCTC TGAGAGGCAG ATCAAAATCT GGTTCCAAAA CCCTCGC	
	M	---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------T	100

Hoxc-5

Clone # Species Amino Acid Sequence Comparison

```
20        O   ----------------------------------------S-
          H   ELEKEFHFNRYLTRRRRIEIANNLCLNERQIKIWFQNRR
          M   ----------------------------------------
```

Clone # Species DNA Sequence Comparison % Id

```
                    10        20        30        40        50        60        70        80        90       100       110   117
20        O   --GT-G---- -G-------- ---------- --------A- ---------- ---------- ---------- ---------- C--------- ---------- -------A-- ---C---  98
          H   GAACTCGAGA AAGAATTCCA CTTTAACCGC TACCTCACTC GCCGCAGGCG CATAGAGATC GCCAACAACT TGTGTCTCAA TGAGAGACAG ATCAAGATCT GGTTCCAGAA CCGACGG
          M   Unavailable
```

Hoxc-9

Clone # Species Amino Acid Sequence Comparison

```
30        O   -F--------------------------------------
18        O   ----------------------------------------
          M   ELEKEFLFNMYLTRDRRYEVARVLNLTERQVKIWFQNRR
          H   ----------------------------------------
```

Clone # Species DNA Sequence Comparison %ID

```
                    10        20        30        40        50        60        70        80        90       100       110   117
30        O   --GT-T--A- --------- ---------- ---------- ---------- ---------- -----G---- ----C----- C--------- ---------- -------A-- ---A---  96
18        O   --G--A---- --------- ---------- ---------- ---------- ---------- -----G---- ----C----- C--------- ---------- ----C----- ---CC--  96
Published O   ---------- --------- ---------- ---------- ---------- ---------- -----G---- ----C----- C--------- ---------- ----T----- T------  97
          M   GAACTGGAGA AGGAGTTTCT CTTCAATATG TATTTAACCA GGGACCGTCG GTACGAGGTG GCCCGTGTTC TCAATCTCAC TGAGCGGCAG GTCAAAATCT GGTTTCAGAA CCGGAGG
          H   Unavailable
```

Hoxc-13

Clone No Species Amino Acid Sequence Comparison

```
13        O       -----F-----------------------------------
24        O       -----F---R-------------------------------
26        O       -----F-----------------------------------
34        O       -----F-----------------------------------
          H       ELEKEYAASKFITKEKRRRISATTNLSERQVTIWFQNRR
          M       ----------------------------------------
```

Clone No Species DNA Sequence Comparison

% Id

```
                  10        20        30        40        50        60        70        80        90       100       110   117
13        O       ---T-G---- ----G-T--- A--C------ ---------- ---------- ---------- --G--T---- -A-------- G--------- --C------- -------A-- T--TAGG  91
24        O       --AT-G--A- ----G-T--- A--C----G- ---------- ---------- ---------- --G--T---- -A-------- G--------- --C------- -------A-- ---ACGG  90
26        O       ---T-G--A- ----G-T--- A--C------ ---------- ---------- ---------- --G--T---- -A-------- G--------- --C------- -------A-- ---ACGG  91
34        O       ---------- -A--G-T--- A--C------ ---------- ---------- ---------- --G--T---- -A-------- G--------- --C------- -------A-- ---GCGG  91
          H       GAGCTAGAGA AGGAATACGC GGCTAGCAAG TTCATCACCA AAGAGAAGCG CCGGCGCATC TCCGCCACCA CGAACCTCTC TGAGCGCCAG GTAACCATCT GGTTCCAGAA CCG
          M       ---------------G---- A--C-----A ---------- ---------- ---------- --A------- ---------- G-----T--- --C------- ---------- T--       91
```

Fig 3.4 : *Barx2* Sequence Comparisons.

A. Comparison of the homeodomain of clone 17 with the related proteins, Dbx (murine; Lu et al., 1994) and Cnox3 (*Chlorohydra viridissima;* Schummer et al., 1992). The underlined regions at the ends represent the locations of the degenerate oligonucleotide primers and have not been included in the % identity calculation. Excluding the degenerate primer region, the homeodomain encoded by clone 17 shared a 67% sequence identity with the homeodomain encoded by *Dbx* (Lu et al., 1994) and 63% with the homeodomain encoded by *Cnox3* (Schummer et al., 1992).

B. Amino acid sequence comparison of the ovine Barx2 PCR product and the murine Barx2 homeobox (Jones et al., 1997). Ovine Barx2 homeodomain sequence between the two degenerate primer regions ELEKEF and WFQNRR has 100% sequence identity with the murine Barx2 homeodomain. The homeobox of ovine *Barx2* has 94% identity with the mouse *Barx2* homeobox between the two degenerate primer regions. Dashes represent conserved amino acids. O = Ovine (Ovis aries), H = human (Homo sapiens), M = mouse (Mus musculus).

A

		% Identity
Dbx	----T------I-K---KK--SK---KDS--------R-	67
Clone 17	ELEKEFQKQKYLSTPDRLDLAQSLGLTQLQVKTWFQNLR	
cnox3	---R--NNR------Q-TN--DR---N-T-----Y--R-	63

B

Barx2

Clone #	Species	Amino Acid Sequence Comparison
17	O Barx2	E---E----------F--L-
	M Barx2	GLEKKFQKQKYLSTPDRLDLAQSLGLTQLQVKTWYQNRR
	H Barx2	Unavailable

Clone #	Species	DNA Sequence Comparison	% ID
		10 20 30 40 50 60 70 80 90 100 110 117	
17	O Barx2	-AA--G--A- --G-G----- ---------- --C-----C- ---------- ---------- ---------- ---------- ---A-----G --------C- ---T---A-- C--GC--	94
	M Barx2	GGCCTAGAGA AGAAATTCCA GAAGCAGAAG TATTTGTCTA CCCCAGACAG GTTGGACTTG GCCCAGTCTC TGGGACTCAC TCAGCTGCAA GTGAAGACTT GGTATCAGAA TCGCAGG	
	H Barx2	Unavailable	

class, six (*Hoxa-4*, *Hoxb-2*, *Hoxb-6*, *Hoxb-8*, *Hoxc-4* and *Hoxc-5*) belonged to the *Antp* class (Fig. 3.3) and one novel gene, *Barx2* belonged to the *Bar* class (Fig. 3.4).

The homeodomain of one clone, clone 17, was quite different from anything in the literature or databases (Fig. 3.4A). Given the normal finding of high identity between homeodomains of equivalent genes from different species, the sequence of this clone appeared to represent a novel homeobox gene. Clone 17 was named *fox* (follicle hox). Three years later while this work was in progress the murine *Barx2* cDNA was isolated and the full cDNA sequence published (Jones et al., 1997). The fox sequence had 100% sequence identity with the mouse *Barx2* sequence and it was apparent that fox was the sheep orthologue to *Barx2* (Fig. 3.4 B). The fox gene was renamed ovine *Barx2*.

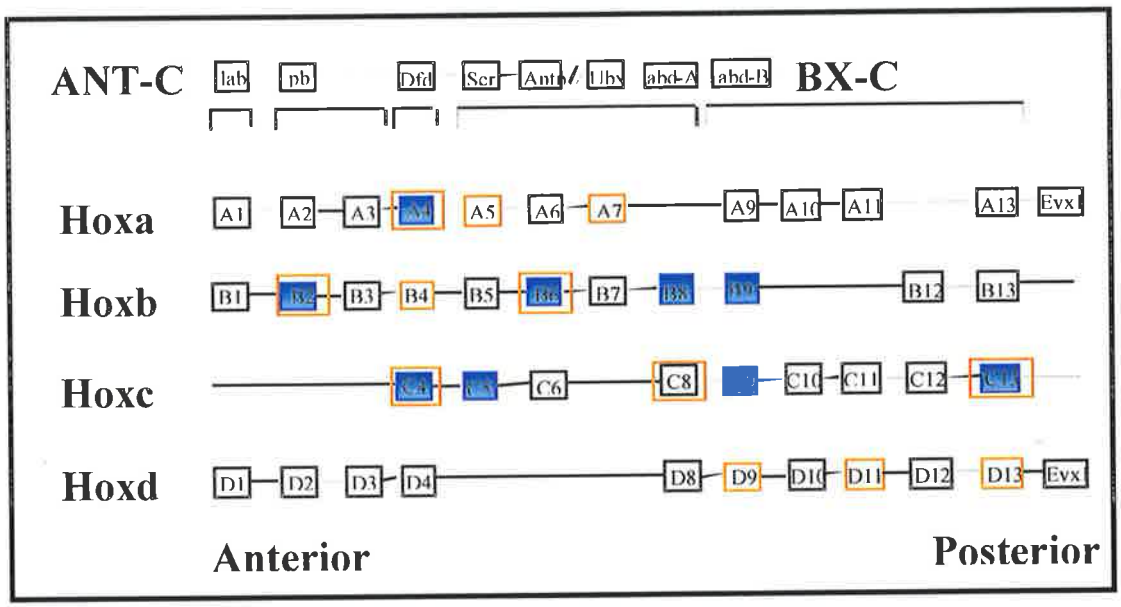
The clones representing the genes *Hoxa-4*, *Hoxb-2*, *Hoxb-8*, *Hoxb-9*, *Hoxc-4*, *Hoxc-8* and *Hoxc-9* contained 100% amino acid sequence identity to the human orthologues (Fig. 3.3; listing of genes are in alphabetical order). Only *Hoxb-6* (clone 16) and *Hoxc-13* (clone 24) contained single amino acid differences in the region between the degenerate primers, both due to a single nucleotide difference (Fig. 3.3). In both clone 16 and clone 24 a guanine base replaced an adenine base resulting in arginine to glycine and lysine to arginine substitutions, respectively. Interestingly, each of the clones representing the ovine *Hoxc-13* gene contained a phenylalanine residue (F) as expected in the ELEKEF degenerate primer region but the human and mouse orthologues have a tyrosine (Y) at the same position.

To compare species-specific nucleotide changes within the homeobox, DNA sequence where the degenerate primers hybridised during PCR (36 bp in length) was excluded, leaving 81 of a possible 117 nucleotides for sequence comparisons. Values of sequence identity between human and sheep do not necessarily reflect values for the entire homeobox. DNA sequence identity of the 21 ovine Hox clones to their respective human orthologue ranged between 84-100% and the results are listed to the right of the sequencing data in figure 3.3.

3.5 Discussion

This chapter has described the use of the RT-PCR screening method to detect expressed *Antennapedia*-like homeobox genes in the sheep wool follicle. Partial cDNA clones representative of nine *Antennapedia* homeobox genes, namely *Hoxa-4*, *Hoxb-2*, *Hoxb-6*, *Hoxb-8*, *Hoxb-9*, *Hoxc-4*, *Hoxc-5*, *Hoxc-9* and *Hoxc-13* and a novel gene, ovine *Barx2* were isolated and sequenced in both directions. When the Hox genes were initially isolated, no data was published for the detection of any of the genes in the follicle except *Hoxb-2* (Whiting et al., 1991). Over the last several years, several groups have reported gene expression of *HOXA4*, *Hoxb-2*, *Hoxb-6*, *Hoxc-4*, *HOXC4* and *Hoxc-13* in follicles (see Fig. 3.5 and Section 1.15) but as yet, no one has studied the expression of *Hoxb-8*, *Hoxb-9*, *Hoxc-5*, *Hoxc-9* and *Barx2* during follicle development.

Fig. 3.5 : A diagram of the four mammalian HOX clusters and a representative of the *Drosophila* ANT-C and BX-C clusters. Homeobox genes boxed in blue represent those isolated in the RT-PCR screen. Homeobox genes boxed in tan represent genes that have been reported to be expressed in the follicle. (Modified from Duboule, 1994).



During the cloning of the partial homeobox cDNAs (133 bp RT-PCR product amplified with the degenerate primers, WFQNRR and ELEKEF), more than 200 colonies with potential Hox gene inserts were not analysed. Only 20 clones were sequenced of which ten different Hox genes were located. This suggests that there may be other Hox genes, apart from the genes mentioned so far, that are expressed in the follicle that have yet to be isolated. Furthermore, Hox genes containing an EcoR1 site within the homeobox may not

have been detected because their shorter insert length would make detection impossible by the cracking methodology used to detect recombinant DNA. This problem could be overcome by blunt end cloning of the RT-PCR product.

Upon analysis of sequence data, the 5' and 3' ends of all of the cloned partial ovine cDNAs were highly divergent to published human sequence data. This was attributed to the use of degenerate primers in the RT-PCR. Interestingly the primers did not require a perfect match with the target template in order to produce an amplified product. For example, clones 13, 24, 26 and 34 encode the same ovine *Hoxc-13* homeodomain but have differences in their DNA sequence within the ELEKEF motif. The low annealing temperature of 45°C in the first 3 cycles of the RT-PCR no doubt allowed for these differences in primer-template hybridisation from the population of degenerate primers. DNA sequence differences between clones of the same gene (ie. ovine *Hoxb-6* clone 16, contains a guanine at position 85 and ovine *Hoxb-6* clone 22, contains a cytosine at position 85) may be due to nucleotides being misincorporated in the early steps of PCR amplification. Hence, to discern and overcome consensus sequence errors due to PCR artefacts a number of clones representative of the Hox gene must be sequenced.

The species-specific amino acid substitution of lysine to arginine in clone 24 (*Hoxc-13*) was discarded once sequence data had been obtained from genomic DNA (Chapter 6) and it was concluded to be a PCR artefact. All four isolates of the *Hoxc-13* gene also contained the amino acid phenylalanine (F) at position 20 of the homeodomain. Although the amino acid difference was in the ELEKEF degenerate primer area, sequencing data obtained from 5'

RACE and genomic clones confirmed that phenylalanine is a species-specific amino acid change (Chapter 5).

Ovine *Barx2* was a novel gene when first isolated during the course of this thesis but during its characterisation the cDNA sequence of the murine orthologue was published by Jones et al., (1997). The ovine *Barx2* gene is the second vertebrate gene in the *Bar* family. The partial ovine *Barx2* homeodomain shared 67% sequence identity with the homeodomain encoded by *Dbx* and a 63% sequence identity with the homeodomain encoded by *Cnox3*. These genes are members of the *Bar* family which have a tyrosine instead of the usual phenylalanine at position 49 of the homeodomain. All metazoan homeobox genes studied to date contain a phenylalanine at that position. Clone 17 contains a phenylalanine at position 49 located within the WFQNRR degenerate primer region. Note that DNA sequence within the degenerate primer region is ambiguous and requires further confirmation by sequence analysis of a full-length cDNA clone or 3' RACE product and is discussed in Chapter 4.

In summary, this chapter has reported the isolation, sequencing and classification of clones containing partial cDNAs representative of nine different *Antennapedia*- type homeobox genes and one novel gene named ovine *Barx2*.

Chapter 4

Characterisation of Partial Ovine homeobox cDNAs

Chapter 4 : Characterisation of Partial Ovine homeobox cDNAs

4.1 Introduction

Ten partial homeobox cDNAs were isolated from the wool follicle by RT-PCR (Chapter 3) and were classified into the 3 classes, *Abd-B*, *Antp* and *Bar*. A literature search confirmed that none of the archetypal genes of the three classes had been investigated in relation to hair follicle development and differentiation.

To characterise the expression of these cDNAs in the wool follicle the techniques of Northern blot analysis, RNA protection and *in situ* hybridisation were employed. Characterisation was preliminary because the probes used in the expression analyses were synthesised from the homeobox region. It was hoped that the preliminary characterisation would acquire enough information to enable the choice of one or two homeobox genes for further detailed study.

The aim of the work described in this chapter was to investigate the expression of the ten homeobox cDNAs in the wool follicle.

Results

4.2 Preliminary Northern Analysis

The possibility that the wool follicle expressed more than one type of homeobox gene was strongly supported by a preliminary experiment in which adult wool follicle poly A(+) RNA was analysed for expressed homeobox genes by Northern blot analysis. The Northern blot was probed with 2 degenerate homeobox oligonucleotides designed to detect adjacent segments at the end of helix 3 of the homeodomain (Fig. 4.1). Five poly A(+) RNA transcripts of differing sizes were detected, 0.8 kb, 2.0 kb, 2.4 kb, 3.0 kb, and 5.0 kb, suggesting that several homeobox genes may be involved in hair growth (B. Powell, personal communication).

4.3 Analysis of Expression of Hox Genes in the Wool Follicle.

4.3.1 Northern Analysis of Wool Follicle Hox Genes

Expression of eight homeobox genes *Hoxc-13*, *Barx2*, *Hoxc-5*, *Hoxc-9*, *Hoxb-2*, *Hoxb-8*, *Hoxb-9* and *Hoxa-4* in sheep skin was analysed by Northern blot (Fig. 4.2). Ovine *Hoxc-13* and ovine *Barx2* gave relatively strong hybridisation signals, with bands at 2.6 kb and 2.0 kb respectively indicating that they are probably more highly expressed than the other isolated Hox genes in the sheep wool follicle. Six of the eight probes failed to give a detectable hybridisation signal despite the membranes being washed at low stringency and exposed for two days to a phosphor - image screen.

Figure 4.1 : Northern Transfer Hybridisation Analysis of Sheep Wool Follicle Poly A(+) RNA.

A Northern blot containing 5µg of wool follicle poly A(+) RNA was probed with two degenerate oligonucleotides designed to the end of helix 3 of the Antennapedia-type homeodomain (B. Powell, unpublished). The sequences of the oligonucleotides used were;

1. HB-1: 5' - AAA/G ATX TGG TTT/C CAA/G AAC/T A/CGX A/CG 3' and
2. HOM-1: 5' - GCA/G/T TAC TTC/T GTA/G TTC/T TTC/T CTC/T - 3'

where X = A, G, C or T. Oligonucleotides were a gift of Dr Bill Kalionis, Department of Obstetrics and Gynaecology, Flinders University, SA and were designed by Burglin et al, (1989). Oligonucleotide labelling and hybridisation conditions were the same as those described by Burglin et al, (1989).

The size (kb) of five putative homeobox RNA transcripts are shown on the left of the figure and were determined by size comparison with an RNA ladder (not shown).

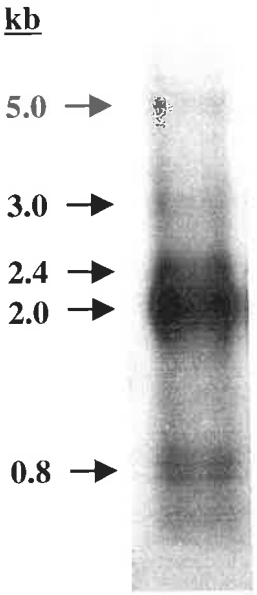
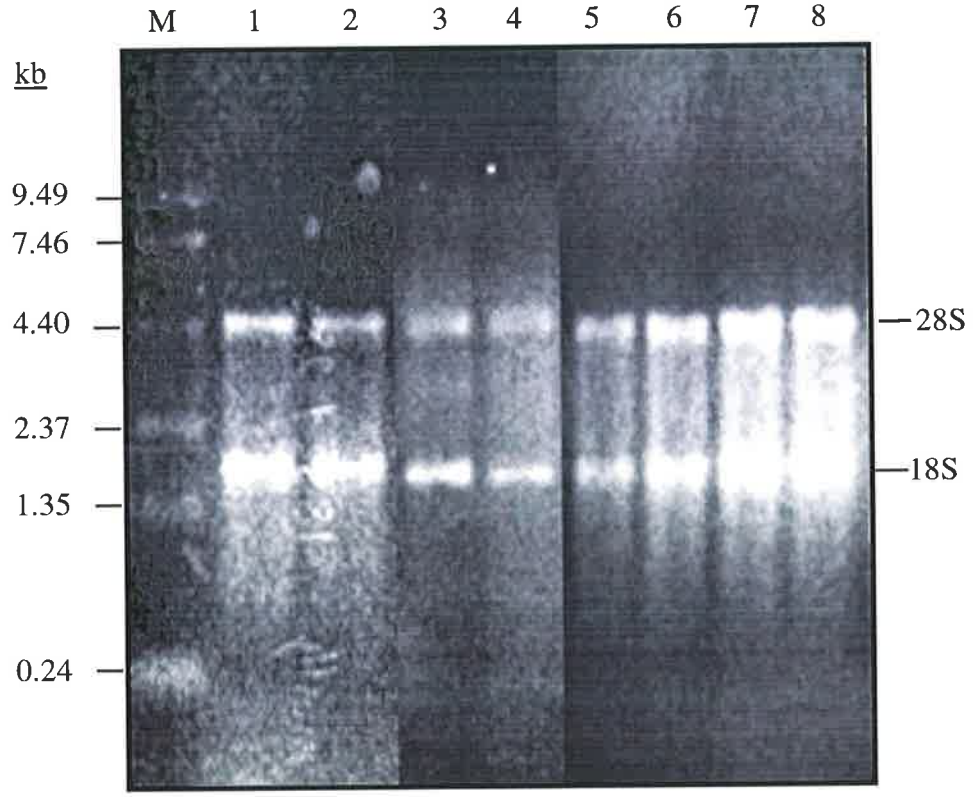

kb
5.0
3.0
2.4
2.0
0.8

Figure 4.2 : Northern Analysis of Homeobox Gene Expression in the Wool Follicle.

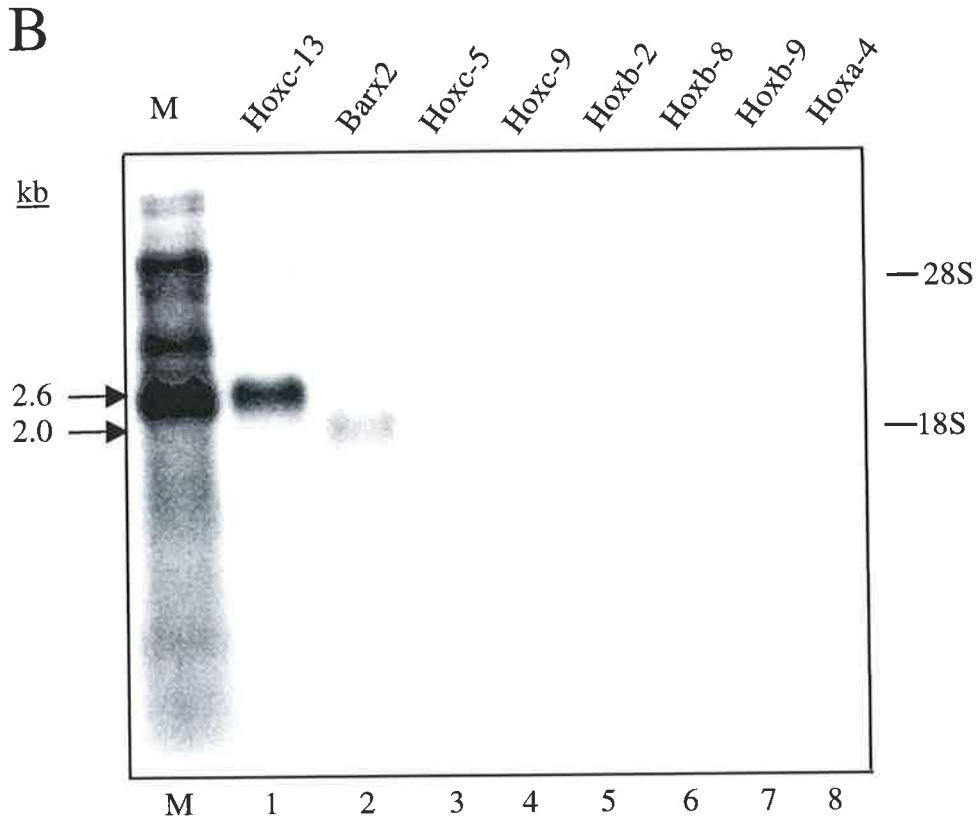
A. Poly A(+) RNA (approximately 5 μg / lane; lanes 1-8) isolated from wool follicles of a Corriedale sheep was size fractionated in a 1% agarose formaldehyde gel, transferred to Zetaprobe GT membrane and stained with Ethidium Bromide. The location of the 28S rRNA and 18S rRNA bands are indicated to the right of the figure. Sizes of the BRL RNA fragments (M) are indicated on the left of the figure in kilobases.

B. Strips from the Northern blot were probed separately with α- ^{32}P- dCTP radio-labelled fragments synthesised from the homeobox regions of ovine *Hoxc-13* (clone24), *Barx2* (clone17), *Hoxc-5* (clone 20), *Hoxc-9* (clone 18), *Hoxb-2* (clone28), *Hoxb-8* (clone 31), *Hoxb-9* (clone 35) and *Hoxa-4* (clone 21). Filters were washed at low stringency: 2.0 X SSPE, 1% SDS at 65^{O}C for one hour and the signal detected by exposing for two days to a phosphorimage screen. A mRNA species detected by the *Hoxc-13* probe (lane1) was sized at 2.6-kb while one detected by the *Barx2* probe (lane2) was sized at 2.0-kb (*Barx2*). Sizes were estimated by comparison with a BRL RNA ladder (M). No hybridisation signal resulted using the ovine *Hoxc-5, c-9 b-2, b-8, b-9* and *a-4* probes. The location of the 28S rRNA and 18S rRNA bands are indicated to the right of the figure.

A


M
1
2
3
4
5
6
7
8
kb
9.49
7.46
4.40
2.37
1.35
0.24
28S
18S

B


M
Hoxc-13
Barx2
Hoxc-5
Hoxc-9
Hoxb-2
Hoxb-8
Hoxb-9
Hoxa-4
kb
2.6
2.0
28S
18S
M
1
2
3
4
5
6
7
8

4.3.2 RNA protection assay

Given that only two Hox gene transcripts, *Hoxc-13* and *Barx2* were detected by Northern analysis, the more sensitive RNA protection assay was used to confirm that the ovine *Hoxb-9*, *Hoxb-8*, *Hoxb-6*, *Hoxb-2*, *Hoxa-4*, *Hoxc-5* and *Hoxc-9* homeobox genes amplified by the RT-PCR technique were also expressed in the wool follicle. The RNA protection assay gave protected products ranging in size from 93 -111 bp and differing in abundance for all of the homeobox genes studied, except for *Hoxa-4* and *Hoxc-5* (Fig. 4.3 A & C). *Hoxb-2* and *Hoxb-8* genes were likely the third and fourth most abundantly expressed respectively (Fig. 4.3 A).

4.3.3 RNA *In situ* Hybridisation Analysis

To detect the location of Hox RNA transcripts in the wool follicle a preliminary RNA *in situ* hybridisation analysis was performed using riboprobes to the cloned homeobox regions of *Hoxb-6*, *Hoxc-5*, *Hoxc-9*, *Hoxc-13* and *Barx2*.

In the *in situ* hybridisation analysis, riboprobes were hybridised to whole skin sections excised from the mid-flank region of Corriedale sheep. Control slides in which tissues were hybridized with a sense RNA probe for each Hox class investigated were negative.

The antisense probe derived from *Hoxc-13* (clone 24), detected mRNA transcripts in a region of follicle bulb cells adjacent and surrounding the apex of the dermal papilla (Fig. 4.4). Expression was strongest in bulb cells located above the dermal papilla, tapered off in the lower part of the follicle shaft and was not seen in the upper shaft. Expression was not seen in the follicle outer root sheath, dermal papilla, inner root sheath or dermis.

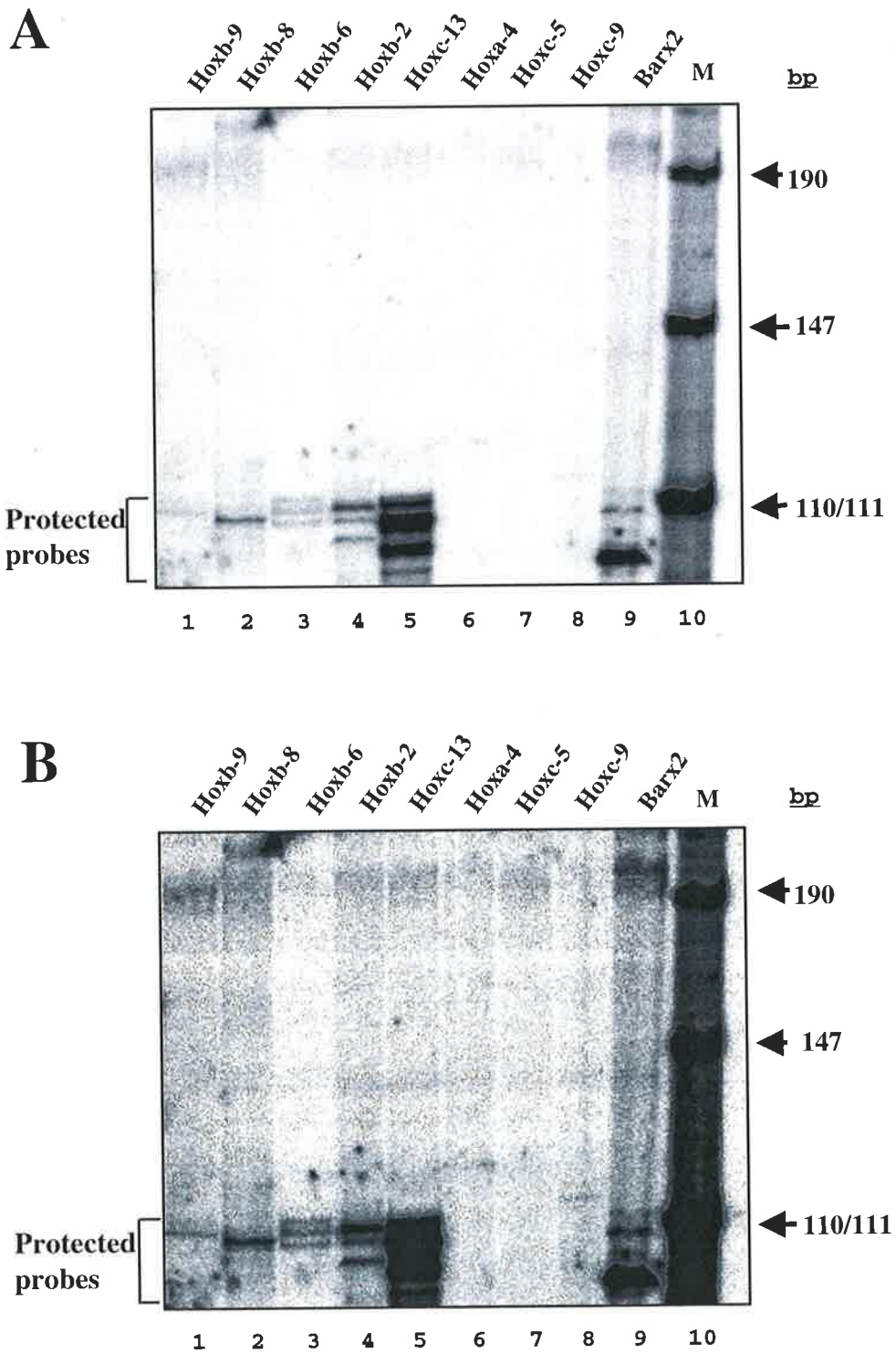
Figure 4.3 : RNA Protection Analysis of *Hox* Gene Expression in Sheep Wool Follicle mRNA.

Approximately 1.0 µg of poly A(+) RNA isolated from adult Corriedale sheep wool follicles was hybridised with α– ^{32}P- rUTP radio-labelled anti-sense RNA probes to the homeodomain. Clones represent the three classes of *Antennapedia Hox* genes identified in the original RT-PCR screen for expressed sheep wool follicle *Hox* genes. RNase-treated hybrids were fractionated in a 6 % polyacrylamide gel and exposed to a phosphorimage screen for 48 hours.

A. Lanes 1-9: hybridisation products for Homeobox coding probes as indicated above the lanes. Protected products were detected for ovine *Hoxb-9, b-8, b-6, b-2, c-13* and *Barx2*. The size (bases) of α– ^{32}P- dCTP end-labelled HpaII-digested pUC19 markers in lane 10 are indicated on the right.

B. Computer image enhancement of A to show *Hoxb-9*, *b-8*, *b-6* more clearly.

C. To detect transcripts for *Hoxc-5 and c-9* the RNA protection assay was repeated with freshly prepared adult Corriedale wool follicle total RNA. *Hoxa-4* was not included. 100µg of total RNA was hybridised with α– ^{32}P- rUTP radio-labelled anti-sense cRNA probes to *Hoxc-9* and *Hoxc-5*. A protected product was detected for *Hoxc-9* (lane 6) but not for *Hoxc-5* (lane 9). Lanes 2, 5 and 8 show probe controls with 10µg yeast tRNA and no RNase treatment. Lanes 1, 4 and 7 show yeast controls treated with RNase. *Barx2* (lane 1, 2 and 3) was included as a positive control.


A
Hoxb-9
Hoxb-8
Hoxb-6
Hoxb-2
Hoxc-13
Hoxa-4
Hoxc-5
Hoxc-9
Barx2
M
bp
190
147
110/111
Protected probes
1 2 3 4 5 6 7 8 9 10
B
Hoxb-9
Hoxb-8
Hoxb-6
Hoxb-2
Hoxc-13
Hoxa-4
Hoxc-5
Hoxc-9
Barx2
M
bp
190
147
110/111
Protected probes
1 2 3 4 5 6 7 8 9 10

C

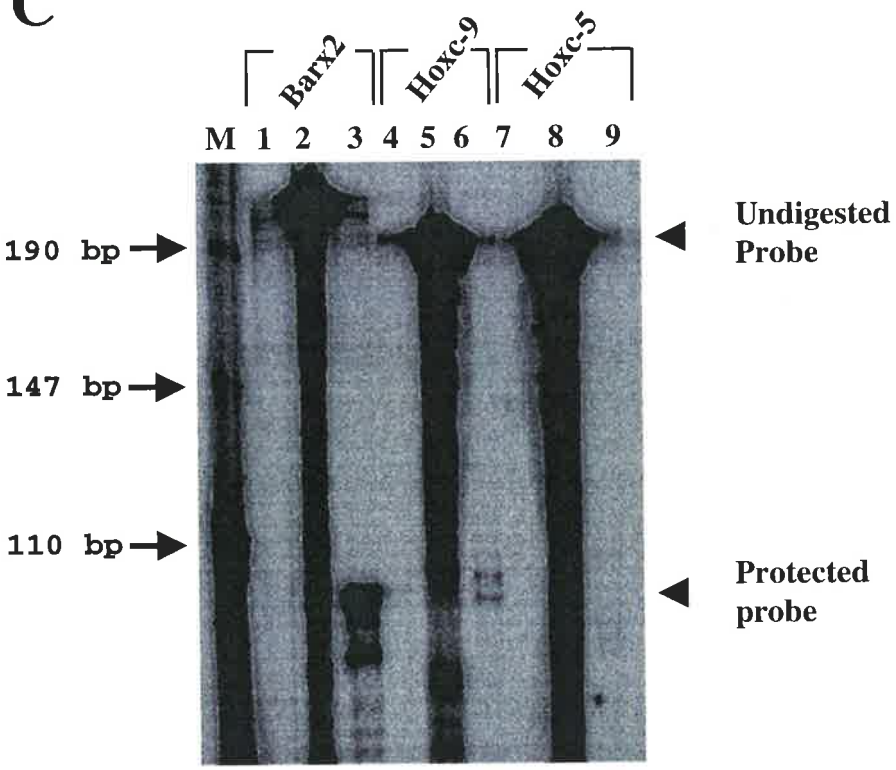
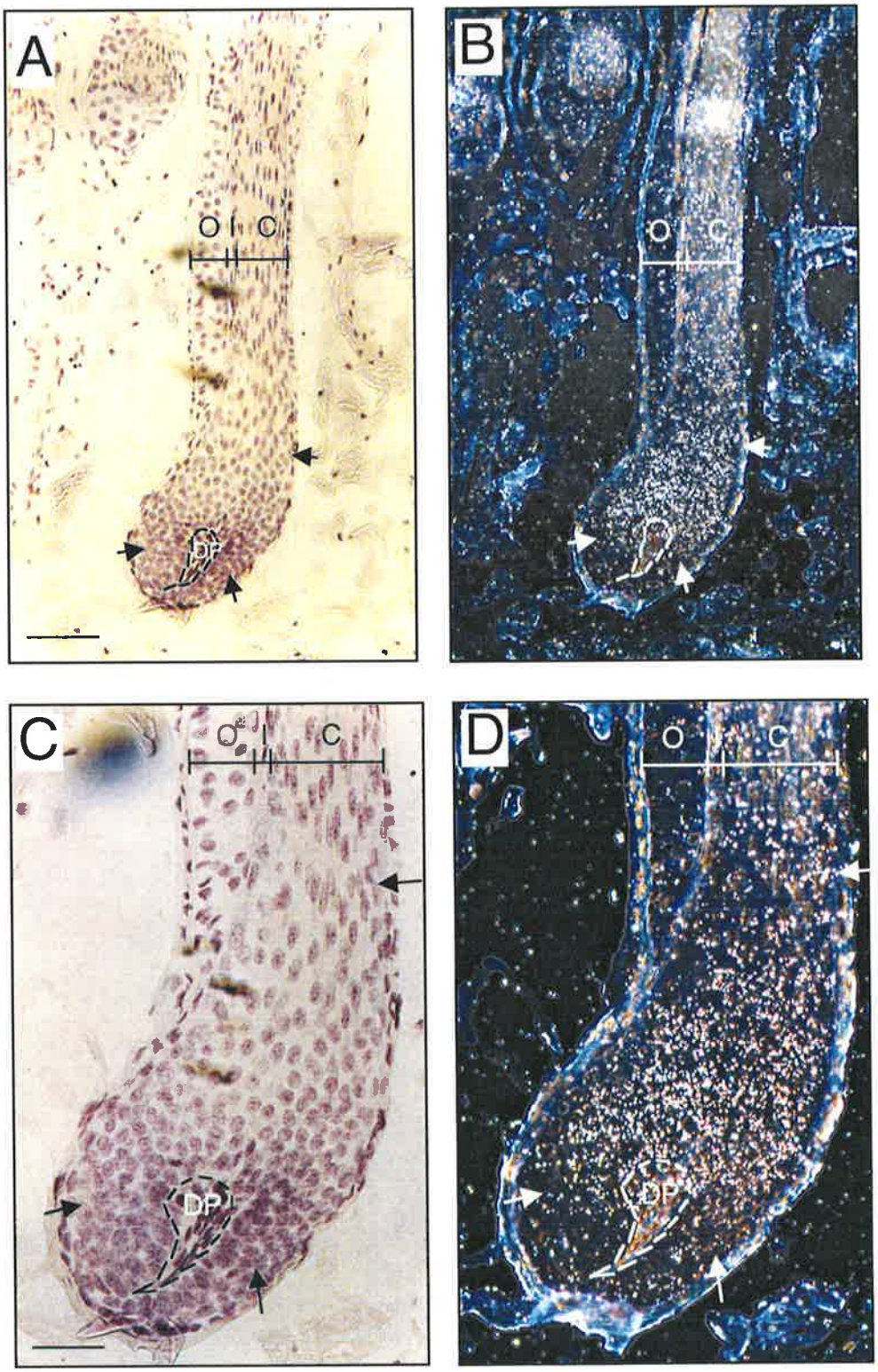

Barx2
Hoxc-9
Hoxc-5
M 1 2 3 4 5 6 7 8 9
190 bp
147 bp
110 bp
Undigested Probe
Protected probe

Figure 4.4 : Ovine *Hoxc-13 In situ* Hybridisation Analysis of a Corriedale Sheep Wool Follicle.

Brightfield (A and C) and darkfield (B and D) photographs of a Corriedale skin section hybridised with an ovine *Hoxc-13* α- ^{33}P- rUTP radio-labelled anti-sense cRNA probe synthesised from the homeobox region. Final wash stringency was 0.1 X SSC at 65^{o}C for 30 minutes. The emulsion (Ilford L4) was exposed to the hybridised tissue sections for 17 days at 4^{o}C. Arrow's point to the hybridisation signal of ovine *Hoxc-13*. Note that it commences in the middle of the bulb, surrounds the upper part of the dermal papilla and tapers off in the lowest region of the follicle shaft. Expression was not detected in the remainder of the shaft, dermal papilla or outer root sheath. Sections were stained with haemotoxylin. C and D are enlargements of A and B, focusing on the bulb. C = Hair shaft cortex and cuticle, DP = Dermal papilla, O = Outer root sheath, I = Inner root sheath. Bar in panel A = 35 μM and C = 18 μM.


A
O
C
DP
B
O
C
C
O
C
DP
D
O
C
DP

A riboprobe prepared from the *Hoxc-9* (clone 18) gene gave a hybridisation signal in a restricted asymmetric pattern in the cortex of the lower hair shaft (Fig. 4.5). Expression commenced in the lower shaft some 10 cells above the dermal papilla. In the left follicle in Figure 4.5 B, expression is restricted to one side of the shaft and continues a short distance towards the upper shaft. The upper shaft cannot be seen because the follicle was sectioned at an oblique angle. In Figure 4.5 C & D two follicles have been cut transversely through the lower shaft. In both follicles, *Hoxc-9* is expressed on one side of the cortex. The signal is in a region of cortical cells located furthest away from the ORS thickening. To determine if *Hoxc-9* was expressed in the upper shaft transverse sections of the upper shaft were analysed. Signal was not detected in the upper shaft (data not shown).

Hoxc-5 (clone 20) was expressed in a restricted asymmetric pattern in the cortex of the lower hair shaft in the same region where *Hoxc-9* was detected (Fig. 4.6 B). *Hoxc-5* expression appears to be on both sides of the left follicle in Figure 4.6 B however the follicle was only cut through the side of the follicle shaft that contained *Hoxc-5* expression. A transverse section of the lower follicle shaft revealed that the *Hoxc-5* signal was restricted to one side of the cortex (Fig. 4.6 D). Like *Hoxc-9*, signal is in a region of cortical cells located furthest away from the ORS thickening.

The ovine *Barx2* and ovine *Hoxb-6* Hox box antisense probes did not detect mRNA transcripts in any region of the skin.

Figure 4.5 : *In situ* Hybridisation Analysis of Ovine *Hoxc-9*.

Brightfield (A and C) and darkfield (B and D) photographs of Corriedale sheep pelage follicles hybridised *in situ* with α- ^{33}P- rUTP radio-labelled cRNA probe for ovine *Hoxc-9* (clone 18) synthesised from the homeobox region. Final wash stringency was 0.1 X SSC at 65°C for 30 minutes. The emulsion was exposed to hybridised tissue sections for 17 days at 4°C. B is a darkfield photograph of the section in A. A and B show oblique follicle sections, C and D show transverse sections. The arrowheads point to an asymmetric labelling of follicle cortical cells in the lower region of the follicle shaft. No expression was seen in the follicle bulb, outer root sheath, dermal papilla or inner root sheath. Note that the signal in D is in a region of cortical cells located furthest away from the outer root sheath thickening. Sections were stained with haemotoxylin. C = Cortex, DP = Dermal papilla, O = Outer root sheath, I = Inner root sheath. Bar in panel A = 35 μM and C = 18 μM.

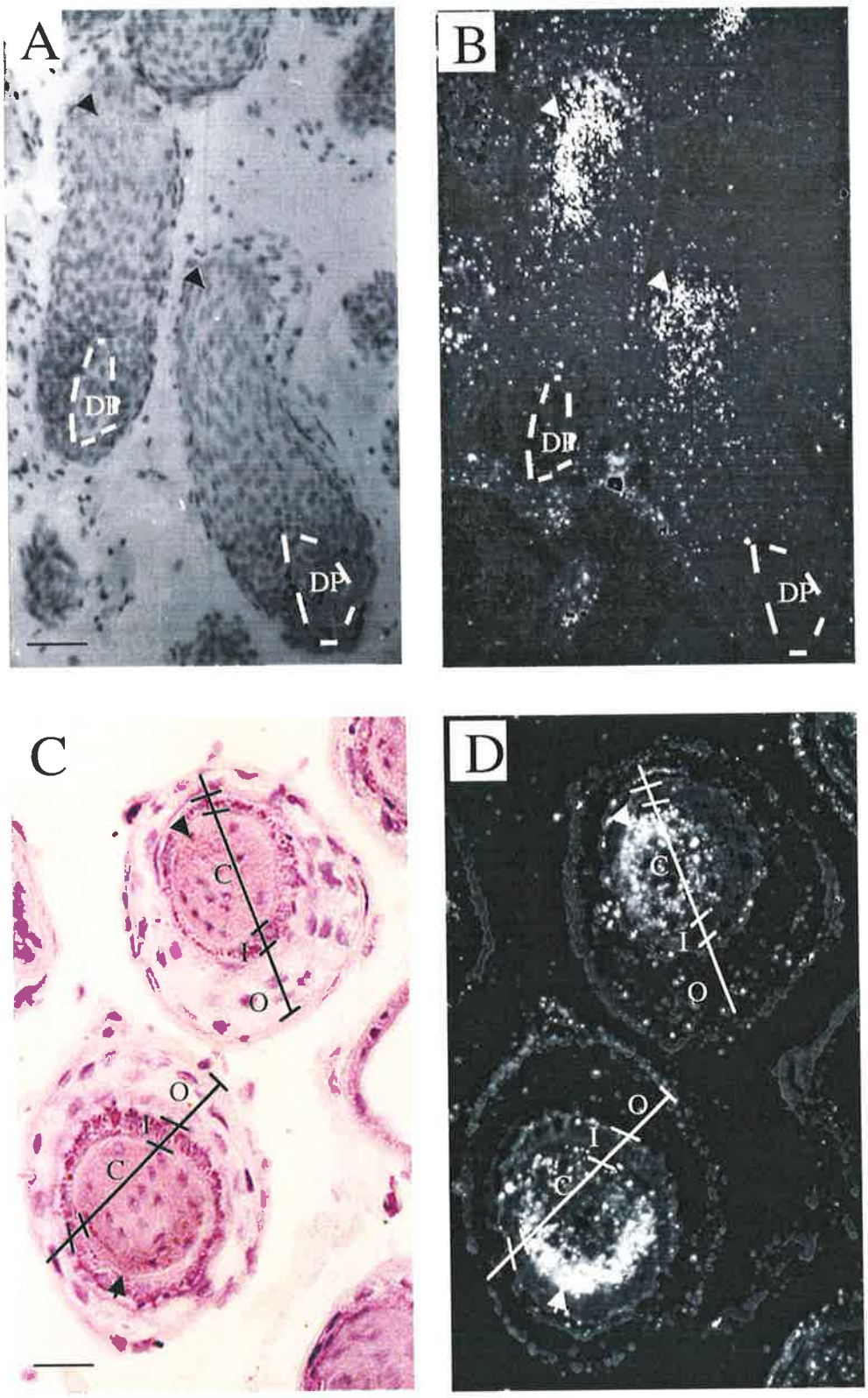
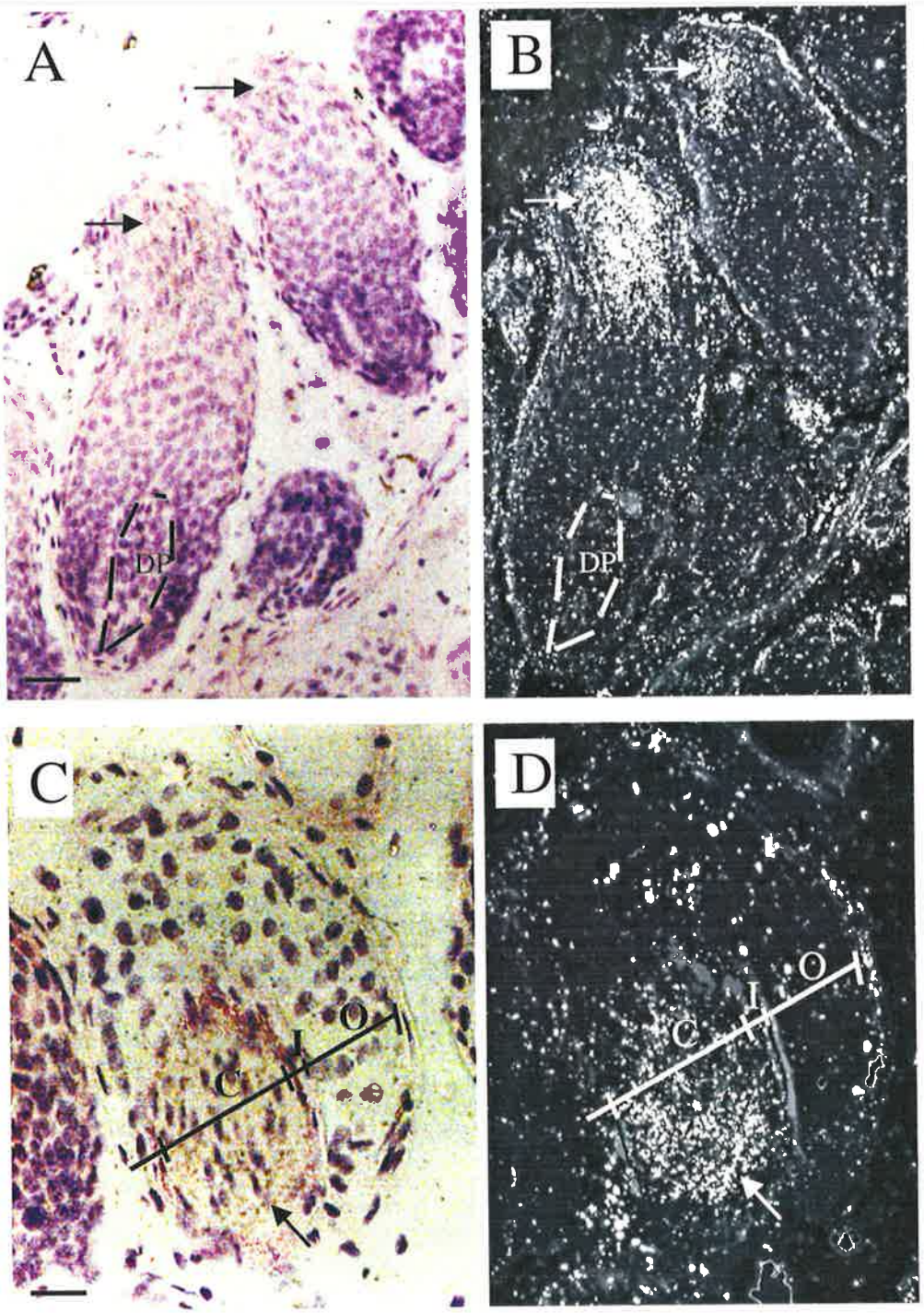

A
DP
DP
B
DP
DP
C
C
I
O
O
I
C
D
C
I
O
O
I
C

Figure 4.6 : *In situ* Hybridisation Analysis of Ovine *Hoxc-5*.

Brightfield (A and C) and darkfield (B and D) photographs of Corriedale sheep pelage follicles hybridised *in situ* with α- ^{33}P- radio-labelled cRNA probe for ovine *Hoxc-5* (clone 20) synthesised from the homeobox region. Final wash stringency was 0.1 X SSC at 65^{o}C for 30 minutes. The emulsion was exposed to the hybridised tissue sections for 17 days at 4^{o}C. B is a darkfield photograph of the section in A. A and B show oblique follicle sections, C and D show transverse sections. The arrows point to an asymmetric labelling of follicle cortical cells in the lower region of the follicle shaft. Again, no expression was seen in the follicle bulb, outer root sheath, dermal papilla or inner root sheath and the signal in D is in a region of cortical cells located furthest away from the outer root sheath thickening. Sections were stained with haemotoxylin. C = Cortex, DP = Dermal papilla, O = Outer root sheath, I = Inner root sheath. Bar in panel A = 35 μM and C = 15 μM.


A
DP
B
DP
C
O
I
C
D
O
I
C

4.4 Discussion

Expression of the partial cDNA clones representative of nine *Antennapedia* homeobox genes namely *Hoxa-4*, *Hoxb-2*, *Hoxb-6*, *Hoxb-8*, *Hoxb-9*, *Hoxc-4*, *Hoxc-5*, *Hoxc-9* and *Hoxc-13* and a novel gene, ovine *Barx2* was characterised by Northern blot, RNA protection or *in situ* hybridisation.

4.4.1 Northern Blot Analysis

The length of the mRNA transcripts for *Hoxc-13* and *Barx2* were determined by Northern analysis to be 2.6 kb and 2.0 kb respectively. Interestingly, similar sized transcripts of approximately 2.4 kb and 2.0 kb were detected in a Northern blot using 2 degenerate homeobox oligonucleotides designed to detect adjacent segments at the end of helix 3 of the homeodomain (Fig. 4.1). Transcripts of 5.0 kb, 3.0 kb and 0.8 kb were also detected suggesting that at least three more homeobox genes can be detected by Northern blot.

No signal was detected for the ovine *Hoxa-4*, *c-5*, *c-9*, *b-2*, *b-8*, *b-9* genes by Northern blot. Given that a faint signal was detected for the ovine *Hoxc-9*, *b-2*, *b-8*, *b-9* genes in the more sensitive RNA protection analysis a signal may be detected by Northern blot if more poly A(+) RNA (eg. 15 μg) is used in combination with a longer probe and longer exposure times.

Ovine *Hoxb-6* was not analysed by Northern blot, however given that *Hoxb-2* appears to be expressed 2-3 times higher than *Hoxb-6,* as determined by RNA protection analysis (see Fig. 4.3) and was not itself detected by Northern blot, it could be assumed that a Northern signal for ovine *Hoxb-6* would not have been detected.

4.4.2 RNA Protection

The more sensitive method of RNA protection assay was used to confirm that nine of the ten homeobox genes detected by RT-PCR were expressed in the wool follicle (Fig. 4.3). Of the nine genes analysed, a signal was detected for seven of them, *Hoxb-9*, *Hoxb-8*, *Hoxb-6*, *Hoxb-2*, *Hoxc-13*, *Hoxc-9* and *Barx2*. The *Hoxc-13* and *Barx2* probes gave the strongest signals confirming the Northern blot data that these genes are the most abundantly expressed.

The intensity of bands (protected product) was different for each probe and may reflect the levels of mRNA transcript being synthesised for each Hox gene in the wool follicle. Although it is assumed that equal amounts of follicle poly A(+) RNA template and radioactively labelled probe were used for each RNA protection, it should be noted that no internal standard was used for quantitative analysis, hence comparison of relative expression levels is only an estimate.

In the RNA protection assay, protected double-stranded RNA products ranging in size from approximately 93 to 111 bases were seen in several lanes. This size-variation can be attributed to the riboprobes having non-homologous sequence at the 3' and 5' ends, an inherent problem when using degenerate primers to amplify RT-PCR products. Different-sized products will be generated depending on the ability of the 3' and 5' ends of the labelled riboprobe to hybridize fully or partially to follicle RNA. Unstable hybridisation may expose single stranded transcript ends to the degrading action of ribonuclease-A and T1, while double stranded RNA products are protected.

A signal for *Hoxa-4* and *Hoxc-5* was not detected by RNA protection. Five years after the initial RNA protection analysis was complete (Fig. 4.3), HOXA4, the human orthologue of ovine *Hoxa-4* was reported to be expressed in the cortex of human hair

follicles (Stelnicki et al., 1998). HOXA4 is expressed very strongly in developing human follicles of the fetus but has a very weak expression level in adult follicles (Stelnicki et al., 1998). Therefore the expression levels of *Hoxa-4* and *Hoxc-5* may be assumed to be very low in adult sheep follicles and may only be detected by RT-PCR or *in situ* hybridisation analysis, techniques that can both detect expression at the single cell level. Failure to detect *Hoxa-4* and *Hoxc-5* by Northern blot or RNA protection is probably due to insufficient sensitivity of those methods (under the conditions used). Using longer probes and a larger quantity of RNA in the RNA protection and Northern analysis may overcome the detection problem.

4.4.3 *In Situ* Hybridisation

In choosing which of the ten Hox genes should be analysed by *in* situ hybridisation, *Hoxc-13* and *Barx2* were chosen because of the strong signal obtained in the Northern blot. *Hoxc-9* was chosen because an RNA protection experiment had detected its expression in the wool follicle (Fig. 4.3 C). Expression of *Hoxc-5* was not detected by Northern blot or RNA protection analyses however it was hoped that the more sensitive method of *in situ* hybridisation could detect the location of its expression in the wool follicle. *Hoxb-6* was randomly chosen from the remaining genes that had not been analysed by RNA protection. It should be noted that RNA protection data for the genes *Hoxb-4*, *Hoxb-9*, *Hoxb-8*, *Hoxb-6*, *Hoxb-2* and *Hoxa-4* was not obtained before the initiation of the *in situ* hybridisation analysis experiments.

In hindsight, the *Hoxb-2* and *Hoxb-8* genes could have been analysed by *in situ* hybridisation as they were determined to be the third and fourth most abundantly expressed genes by RNA protection (Fig. 4.3 A). However, at that stage a decision had been made to

focus on the two most abundantly expressed genes in the wool follicle as determined by Northern blot, *Hoxc-13* and *Barx2*.

The *in situ* hybridisation experiments were performed using ribo-probes encompassing the homeobox. Due to the high sequence identity in the homeobox shared by paralogous cluster genes the results reported thus far with probes prepared from segments of the homeobox are preliminary and should be repeated using gene-specific probes.

The location of *Hoxc-13* expression in the follicle bulb surrounding the dermal papilla suggests that it may be involved in the transcriptional regulation of the keratin IF and/or KAP genes expressed in the cortex of the wool fibre (Powell et al., 1992). Further discussion on the expression of *Hoxc-13* is found in Chapter 7.

Hoxc-5 and *Hoxc-9* were expressed in an asymmetric pattern in the cortex of the lower shaft. There are two cortical cell types, the orthocortical and paracortical cells that are bilaterally distributed in fine wool follicles (Mercer, 1961). Most follicles have a kinked morphology beginning at approximately the middle of the bulb, useful in determining where cell types are located. In kinked follicles the ORS is thicker on the concave side of the follicle. The orthocortex is located next to the ORS thickening and the paracortex on the side of least thickening. The expression pattern of ovine *Hoxc-5* seems to correlate with the paracortical cell type because it is located in the side of the cortex adjacent to the outer root sheath showing least thickening. Expression is remarkably similar to that seen for the cysteine regulated KAP4 gene family in the paracortex (Fratini et al., 1994). The cysteine rich proteins encoded by the KAP4 family are elevated in paracortical cells of sheep wool when circulating levels of L-cysteine are elevated (Fratini et al., 1994). To show that *Hoxc-5* is expressed in the same region as the KAP4 family,

serial skin sections from normal or a L-cysteine infused sheep could be probed with a KAP4 gene and *Hoxc-5* gene by *in situ* hybridisation and expression patterns compared.

The RNA used to isolate the ten Hox genes by RT-PCR was isolated from sheep that had been kept on a basal diet for four weeks and then infused for seven days with L-cysteine (Fratini et al., 1994). The L-cysteine infused RNA was chosen because it was readily available from a previous experiment undertaken in this laboratory. Cysteine infusion increases the abundance of KAP4 mRNAs that encode a family of cysteine-rich proteins in the paracortical cells of sheep wool, but does not affect the expression of other wool keratin gene families (Fratini et al., 1994). Northern blot and RNA protection analysis of the ten Hox genes was performed using RNA isolated from a different origin, namely a grazing Corriedale sheep. The expression levels of *Hoxc-5* and *Hoxc-9* may change with fluctuating levels of circulating L-cysteine. This may explain why no signal was detected for *Hoxc-5* by Northern blot and RNA protection analysis and only a faint signal detected for ovine *Hoxc-9* in the RNA protection analysis. It is possible that the ovine *Hoxa-4*, *Hoxb-9*, *Hoxb-8*, *Hoxb-6* and *Hoxb-2* genes may also be regulated by circulating levels of L-cysteine given that they also were screened from the same RNA sample used for RT-PCR. An RNA protection or RT-PCR quantitation of *Hoxc-5* and *Hoxc-9* gene expression using gene specific probes and follicle RNA isolated from sheep with and without infusion of L-cysteine should reveal whether the levels of these two Hox genes are regulated by L-cysteine.

The lack of a detectable signal for the ovine *Hoxb-6* and ovine *Barx2 in situ* hybridization's maybe due to the degradation of the radio-labelled cRNA riboprobe and *in vivo* mRNA transcripts by ribonuclease during the *in situ* procedure. This may explain why *Barx2* was detected by Northern blot and RNA protection analyses but not by the *in situ*

hybridisation experiment. An *in situ* hybridisation was repeated for *Barx2* with a gene-specific probe and the location of its expression is discussed in detail in Chapter 5.

Hoxc-4 was detected by RT-PCR but was not analysed by Northern blot, RNA protection or *in situ* hybridisation due to its initial wrong classification as *Hoxc-5*. The human orthologue *HOXC-4* was expressed in the suprabasal layers of the ORS in the infundibulum of human hair follicles and in the suprabasal layers of the epidermis suggesting that it may have a role in the differentiation of keratinocytes (Rieger et al., 1994). Fibroblasts and lymphocytes in the perivascular dermis also expressed *HOXC-4* (Rieger et al., 1994).

In summary, this chapter has reported the characterisation of clones containing partial cDNAs from the wool follicle representative of nine different *Antennapedia*-type homeobox genes and one novel gene named ovine *Barx2*. *Hoxc-13* was the only gene to be detected by *in situ* hybridisation, Northern blot and RNA protection analysis confirming that this gene is relatively highly expressed in the wool follicle. *Barx2* was detected by Northern blot and RNA protection however a signal was not detected by *in situ* hybridisation and possible reasons for this were discussed. *Hoxc-5* and *Hoxc-9* were detected by *in situ* hybridisation and are expressed asymmetrically in the cortex, probably the paracortex, in a similar location to the KAP4 genes that respond to L-cysteine infusion.

The methods employed in this chapter were used to identify one or two genes of interest to be focussed on in more detail in future experiments. The selection of the genes chosen was based on either one of two criteria: 1. the location of their expression in the wool follicle or 2. sequence novelty. *Hoxc-13* was chosen because it was expressed in the follicle bulb where keratinocytes are beginning to differentiate and *Barx2* was chosen because it was a novel gene.

Chapter 5

Isolation and Characterisation of Ovine *Barx2*

(Part of the work presented in this Chapter has been accepted for publication in the Journal of Investigative Dermatology. A copy is shown in Appendix 3)

Chapter 5 : Isolation and Characterisation of Ovine *Barx2*

5.1 Introduction

Characterisation of ten RT-PCR amplified *Hox* genes in the adult wool follicle by a combination of Northern blot, RNA protection and cRNA *in situ* hybridisation analyses enabled the following two genes to be chosen for further characterisation, *Hoxc13* because it is expressed in differentiating cells and *Barx2,* because it was a novel gene at the time of its isolation.

In Chapter 4, expression of *Barx2* was detected by Northern blot and RNA protection analysis but not by *in situ* hybridisation, possibly due to the degradation of the radio-labelled cRNA riboprobe and *in vivo* mRNA transcripts by ribonuclease during the *in situ* procedure. Analysis of *Barx2* expression in the adult wool follicle was initially performed using a probe containing part of the homeobox (*Fox* probe). Due to the high sequence conservation in the homeobox region between homeobox genes, the *Fox* probe may have crossed hybridized with other homeobox genes and yielded a false result in both the Northern blot and RNA protection assay. To overcome this potential problem, a longer gene-specific probe downstream of the homeobox was used to detect *Barx2* expression during embryonic wool follicle development by *in situ* hybridisation and adult follicles by Northern blot. To obtain a gene-specific probe, the technique of 3' RACE-PCR was used to isolate the 3' cDNA end of ovine *Barx2* from wool follicle RNA. A fragment from the *Barx2* 3' coding region was then subcloned and used to probe vertical skin sections taken from a sheep developmental time course. The RNA protection analysis results obtained in Chapter 4 were also confirmed with a gene-specific, 3' non-coding probe.

During the course of fully characterising ovine *Barx2*, the cDNA sequence of the murine orthologue and its expression in embryonic mice up to the age of E12.5 was published (Jones et al., 1997). However, *Barx2* expression during embryonic development of skin and follicles, the latter initiating at approximately E13.25, was not reported.

To elucidate the functional role of *Barx2* within the wool follicle, isolation of the full gene was attempted. A lambda sheep genomic clone containing the homeobox and 3' end (and possibly the 5' end) of *Barx2* was isolated and characterised. While further experiments required to determine the function of Barx2 were not possible due to time

constraints, this Chapter details isolation of the 3' end of *Barx2* and characterisation of its expression during embryonic and adult wool follicle development.

Results

5.2 Isolation of The Ovine *Barx2* 3' cDNA End

5.2.1 3' RACE-PCR

The isolation of the 3' cDNA end for ovine *Barx2* was achieved by 3' RACE-PCR following the method of Frohman (1988). To ensure that *Barx2* mRNA transcripts were present in the RNA, the same sample of RNA used to amplify the ten homeobox genes by RT-PCR (described in Section 3.2.1) was used to synthesise the cDNA used in the 3'RACE-PCR. The RNA was obtained from wool follicles of an adult sheep that had been infused with the amino acid L-cysteine for seven days (Fratini et al., 1994).

An oligonucleotide specific for *Barx2* was designed using Oligo-4 software (Geneworks) from the *Barx2* homeobox (clone C17; Fig. 5.1 A). The sequence of this oligonucleotide, was 5' - TA GTC GAC CAG ACA GGT TGG ACT TGG - 3' and contained a SalI restriction endonuclease site at the 5' end, because the recognition site occurs at a low frequency in genomic DNA. SalI did not cleave the 3'RACE-PCR product internally. Six PCRs were performed and contained the following templates and primer(s);

1. cDNA with C17Hint and 3'MCS primers (PCR #1)
2. mRNA with C17Hint and 3'MCS primers (PCR #2)
3. cDNA with the 3'MCS primer (PCR #3)
4. cDNA with the C17Hint primer (PCR #4)

Control reaction 2 identified whether amplified products were the result of genomic DNA contamination. PCR conditions are given in Section 2.3.8.3.

Multiple PCR products ranging in size from 200 bp to 1.1 kb were amplified from PCR #1 containing cDNA with both primers C17Hint and 3'MCS (Fig. 5.1 B, lane 1). No amplified products were seen in the control PCRs #2 to #4 ie. those containing mRNA with C17Hint and 3'MCS primers (PCR #2), cDNA with the 3'MCS primer (PCR #3), cDNA with the C17Hint primer (PCR #4)(Fig. 5.1 B, lanes 2, 3 and 4 respectively). As *Barx2* was a novel gene, the length of the 3' end was unknown so the largest amplified product that contained the *Barx2* homeobox was cloned.

Figure 5.1 : *Barx2* 3' R.A.C.E.

A. 5' Primer used in the Amplification of the *Barx2* 3' End

The partial sequence of the *Barx2* homeobox, obtained from clone 17 (Fig. 3.4) is shown. The 5' primer used in the amplification of the *Barx2* 3' end is indicated by the arrow and is named C17Hint. Note that the C17Hint primer does not overlap the degenerate ELEKEF motif. The arrow shows the direction of DNA synthesis.

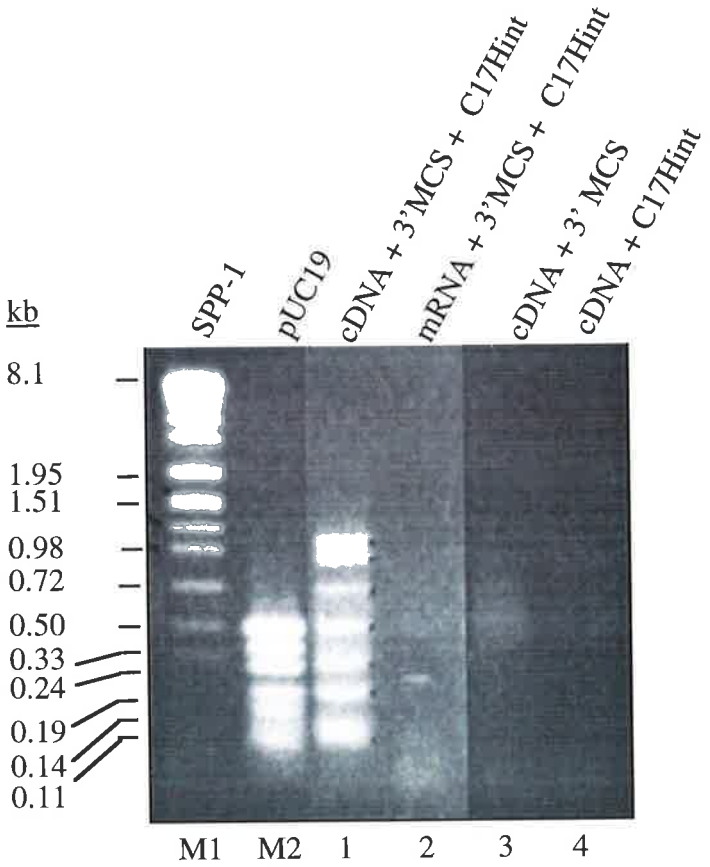
B. Amplification of the *Barx2* 3' cDNA End

The 3' cDNA end of *Barx2* was amplified using the 3'RACE-PCR technique. PCR products were size-separated on a 1% agarose gel and then stained with ethidium bromide and detected under UV light. PCRs were performed using wool follicle cDNA or mRNA in combination with the primers indicated above the gel. Note that there are multiple PCR products in lane 1 (the largest being 1.1 kb) and no products in lanes 2-4.

A

```
 E  L  E  K  E  F  Q  K  Q  K  Y  L  S  T  P  D  R  L  D  L      20
GAACTGGAAAAGGAGTTCCAGAAGCAGAAGTACTTGTCCACCCCAGACAGGTTGGACTTG     60
                                                  C17Hint
 A  Q  S  L  G  L  T  Q  L  Q  V  K  T  W  Y  Q  N  R  R        39
GCCCAGTCTCTGGGACTCACTCAACTGCAGGTGAAGACCTGGTACCAGAACCGCAGG       117
```

B



5.2.3 Southern Blot Analysis and Cloning of the 3' RACE Product

Southern blot analysis was used to determine which 3' RACE-PCR amplified products contained the homeobox and 3' cDNA end of *Barx2*. A blot containing one twentieth of the 3' RACE-PCR products was probed with radio-labelled fragments synthesised from the homeobox of *Barx2* (clone C17; see Appendix 1 A for plasmid map). The largest amplified PCR product that hybridised to the homeobox probe as detected by Southern blot analysis was 1.1 kb in length (Fig. 5.2 A, B). A 0.93 kb product also hybridised to the probe. The 1.1 kb DNA fragment was purified from the amplified 3' RACE-PCR products by extraction from low melting point agarose, cleaved with SalI, ligated into pBluescript KS (+) SalI vector and bacterial transformants screened for recombinant DNA. A clone named C17B1 contained an insert of 1.1 kb and was characterised further (see Appendix 1 B for plasmid map).

5.3 Confirmation of the Identity of the 3' cDNA End of *Barx2* (Clone C17B1)

5.3.1 Sequencing of the 3' RACE Product

To verify that clone C17B1 contained the homeobox of *Barx2,* nucleotide sequence for approximately 200 bp of the 5' end of the clone was determined and compared to sequence of the *Barx2* homeobox (see Fig. 3.4 B for sequence). Excluding the degenerate primer region, the homeobox of these clones shared 100% identity (data not shown).

Before completely sequencing the remainder of the RACE-PCR product, fragments from different regions of clone C17B1 were isolated and used to probe a wool follicle Northern blot and a genomic DNA Southern blot to confirm that clone C17B1 was not a hybrid of the *Barx2* homeobox and another cDNA.

5.3.2 Confirmation of the 3' cDNA End by Northern Blot

Northern blot analysis was used to verify that clone C17B1 truly represented the full 3' cDNA end of *Barx2*. A 163 bp fragment in close proximity to the 3' end of clone C17B1 was excised with BsteII and DraI, cloned and named C17B1-NC (NC = non-coding; see Fig. 5.5 for location of the probe and Appendix 1 D for plasmid map of clone C17B1-NC).

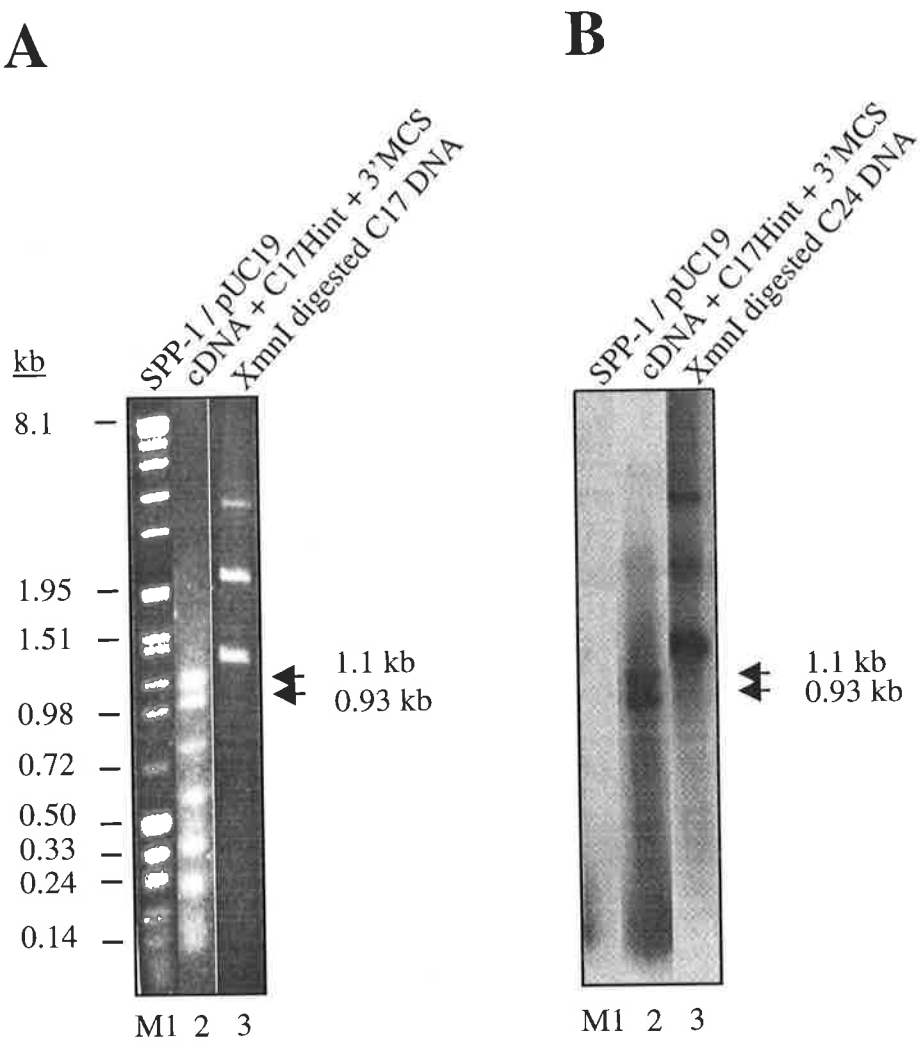
Figure 5.2 : Detection of the *Barx2* 3' End by Southern Blot.

A. *Barx2* 3' RACE-PCR

The 3' RACE-PCR amplified products were fractionated on a 1% agarose gel. Lane M1, 500ng of SPP-1 and pUC19 markers. Lane 2, ovine *Barx2* 3' RACE-PCR products amplified from cDNA with the C17Hint and 3'MCS primers. Lane 3, XmnI partially digested plasmid DNA from clone C17 which contains the homeobox of *Barx2*. The top band in lane 3 at 3.3 kb is linearised plasmid DNA from clone C17 which hybridised to the probe. Clone C17 contained two XmnI sites resulting in the second band at 2.1 kb and the third band at 1.2 kb. The 1.2 kb DNA fragment also contained the homeobox and hybridised to the probe. Marker sizes (kb) are indicated on the left side of the figure.

B. Southern Blot Analysis of *Barx2* 3' RACE-PCR Products

3' RACE-PCR products were analysed by Southern blot. 3' RACE-PCR products were transferred to Zetaprobe membrane and probed with an α-^{32}P-dCTP radio-labelled EcoR1-resected homeobox fragment from clone C17 (*Barx2*). The largest PCR amplified products of 0.93 kb and 1.1 kb hybridised to the probe and are indicated by the arrows. The 1.1 kb fragment was cloned into the SalI site of pBSc KS (+).


A
B
SPP-1 / pUC19
cDNA + C17Hint + 3'MCS
XmnI digested C17 DNA
SPP-1 / pUC19
cDNA + C17Hint + 3'MCS
XmnI digested C24 DNA
kb
8.1
1.95
1.51
0.98
0.72
0.50
0.33
0.24
0.14
1.1 kb
0.93 kb
1.1 kb
0.93 kb
M1 2 3
M1 2 3

The fragment was prepared, radioactively labelled and used to probe wool follicle poly A(+) RNA. A 2.0 kb mRNA transcript was detected, the same size detected by the probe encompassing the ovine *Barx2* homeobox (Fig. 5.3).

The same size transcript was also detected using a 186 bp probe DNA fragment from the coding region of *Barx2*. This fragment was amplified by PCR using the two primers P7 and P11 (see Fig. 5.5 and Fig. 5.6 for location of both primers). The amplified fragment was cloned into the SmaI site of pGEM-7Zf (+) and named C17B1-COD (Appendix 1C).

5.3.3 Confirmation of the *Barx2* 3' cDNA End by Southern Blot

A genomic Southern blot was performed to further confirm that the cloned fragments representing the coding (C17B1-COD; Appendix 1 C) and 3' non-coding (C17B1-NC; Appendix 1 D) regions of ovine *Barx2* were representative of the same gene. A sheep genomic DNA blot was probed separately with radio-labelled fragments synthesised from each clone. A band at the size of 8.5 kb appeared for each probing (Fig. 5.4) suggesting that the separately cloned fragments from the coding and non-coding regions of ovine *Barx2* originated from the same gene.

5.4 Nucleotide Sequencing of the 3'cDNA End of *Barx2* (Clone C17B1)

To obtain the sequence of the 3' cDNA end of *Barx2* a combination of subcloning, progressive deletion and oligonucleotide-primed sequencing were employed (Fig. 5.5). DNA was sequenced in both directions to ensure that sequence information was unambiguous.

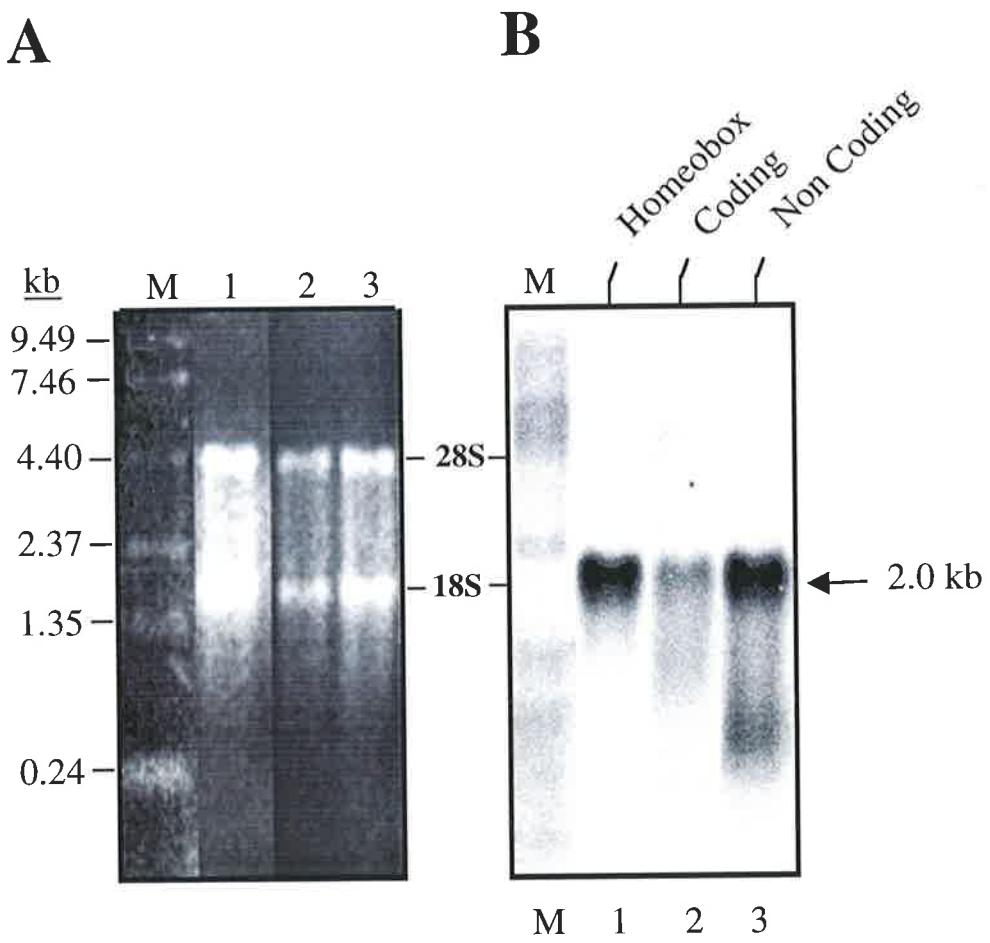
The nucleotide sequence of clone C17B1 together with the deduced amino acid sequence is shown in Fig. 5.6. The 3' cDNA end of *Barx2* is 1,104 base pairs long and contains a partial open reading frame of 133 amino acids (including the homeobox). The carboxyl end of Barx2 from the end of the homeodomain to the stop codon, contains 87 amino acid residues. The 3' non-coding sequence contains 702 nucleotides and two consensus polyadenylation signals (AATAAA) located at positions 928 and 1,085 (Fig. 5.6).

Figure 5.3 : Northern Transfer Hybridisation Analysis of Sheep Wool Follicle Poly- A(+) RNA.

A. Poly- A(+) RNA was prepared using the poly-A-tract kit (Promega) from 1mg of total RNA isolated from Corriedale sheep wool follicles. Approximately 5 µg samples were fractionated in a 1 % agarose formaldehyde gel then transferred to Zetaprobe GT membrane (lanes 1-3). The size (kb) of the marker bands are indicated to the left of the figure. The location of the 28S and 18S rRNA bands are indicated. M = Gibco BRL RNA ladder.

B. Northern blot containing approximately 5 µg of poly- A(+) RNA. Strip filters were separately probed with α-^{32}P-dCTP radio-labelled fragments prepared from the homeobox (clone C17, lane 1), coding (clone C17B1-COD, lane 2) and 3' non-coding regions (clone C17B1-NC, lane 3). Filters were then washed in 2 x SSC / 0.1% SDS / 65^{o}C for 30 minutes and exposed to a phosphor-image screen for 24 hours. The size (kb) of the hybridisation product is indicated to the right of the figure. The location of the 28S and 18S rRNA bands are indicated. Lane one (M) = Gibco BRL RNA ladder.

C. Schematic showing the position of each probe within the 3' end of *Barx2*. Lined box represents the homeobox (1), shaded box represents the coding region (2) and the open box represents the 3' non-coding region (3) of *Barx2*. AAAAA represents the poly- A tail located at the 3'end.


A
B
Homeobox
Coding
Non Coding
kb
M 1 2 3
9.49
7.46
4.40
2.37
1.35
0.24
28S
18S
M
2.0 kb
M 1 2 3

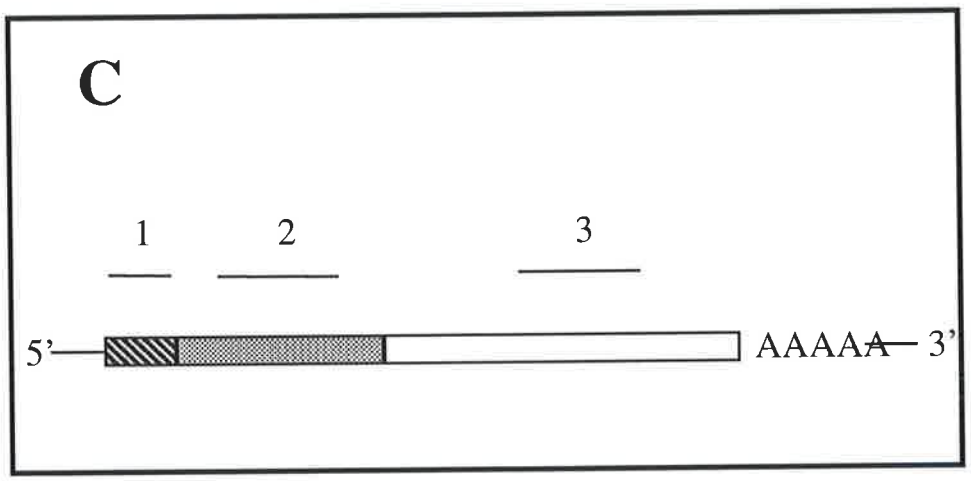

C
1
2
3
5'
AAAAA
3'

Figure 5.4 : Southern Transfer Hybridisation Analysis of EcoRI-digested Sheep Genomic DNA.

Sheep genomic DNA was digested with EcoRI and then 10 µg samples fractionated in multiple lanes in a 1 % agarose (1 X TAE) gel (lanes 2-8) then transferred to Zetaprobe GT membrane. Strip filters were separately probed with radio-labelled fragments prepared from the 3' non-coding (clone C17B1-NC, see Appendix 1 D, lanes 2-5) and coding regions (clone C17B1-COD; see Appendix 1 C, lanes 6-8) of ovine *Barx2*. Filters were washed at 2 x SSC / 0.1% SDS at 65^{o}C for 30 minutes and exposed overnight to a phosphor-image screen. The size of the hybridisation product is indicated to the right of the figure by the arrow. The arrowhead points to Satellite DNA that has hybridised non-specifically to the probe. Lane one (M) = SPP-1 and lambda Hi-Lo markers, sizes (kb) are indicated on the left.

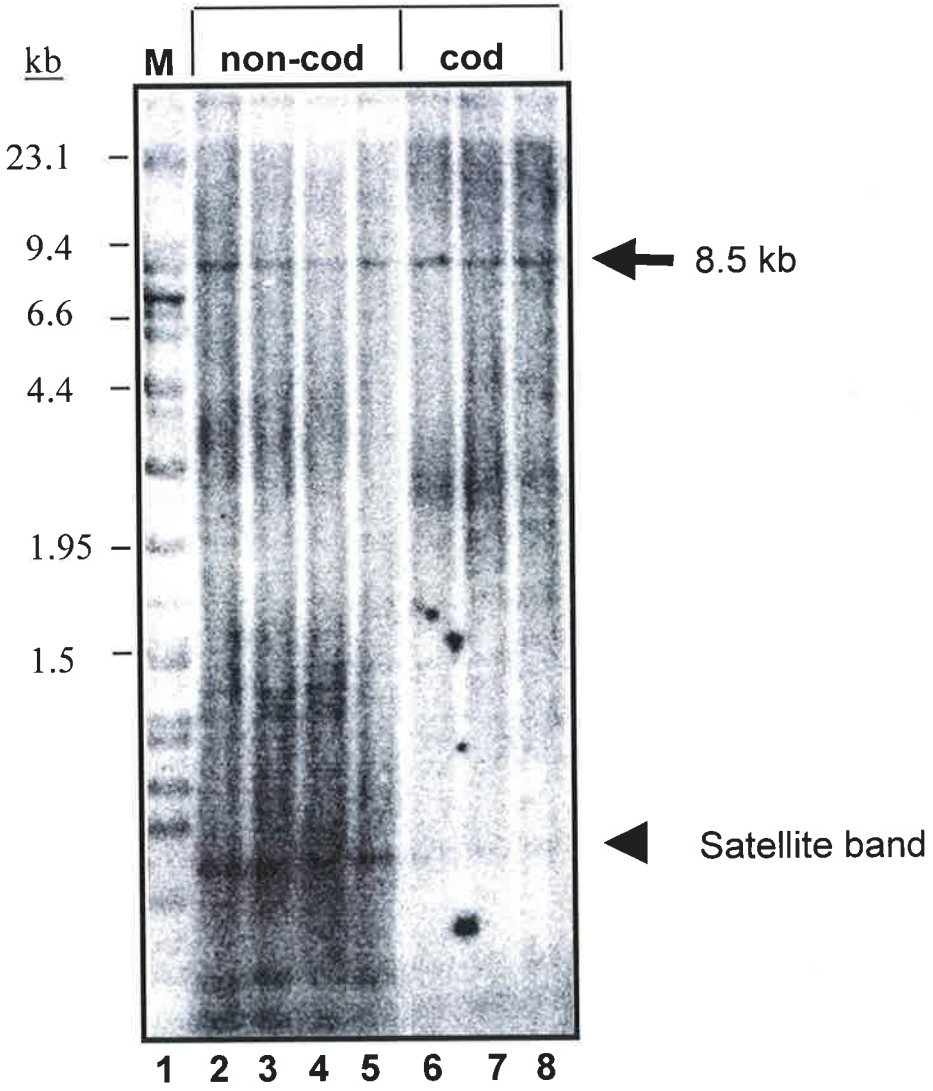

kb
M
non-cod
cod
23.1
9.4
6.6
4.4
1.95
1.5
8.5 kb
Satellite band
1 2 3 4 5 6 7 8

Figure 5.5 : Schematic Showing the Restriction Map of the *Barx2* 3' End and the Sequencing Strategy.

Restriction sites used for subcloning of the 3' non-coding fragment are shown. A 163 bp DraI / BstEII fragment was subcloned into the SmaI site of pBSc KS and sequenced. Smaller fragments of the 3' cDNA end of *Barx2* were prepared by progressive deletion of DNA from clone C17B1 using the Erase-a-base kit (Promega). Deleted fragments were sub-cloned into the SmaI site of the pGEM- 7Zf (+) base vector. To extend the sequence data, oligonucleotides were designed to the sequence obtained from the 5' and 3' ends of the subclones. Names of sequencing primers are given at the beginning of the arrows. Arrows indicate the direction and extent of the sequencing reactions. DNA from clone C17B1 was used as the template for all cycle sequencing reactions.

The lined box represents the homeobox, the shaded box represents the coding region and the open box represents the 3' non-coding region. AAA depicts the poly A tail. Line represents the polylinker region of pBSc KS. RSP = Reverse sequencing primer, USP = Universal sequencing primer.

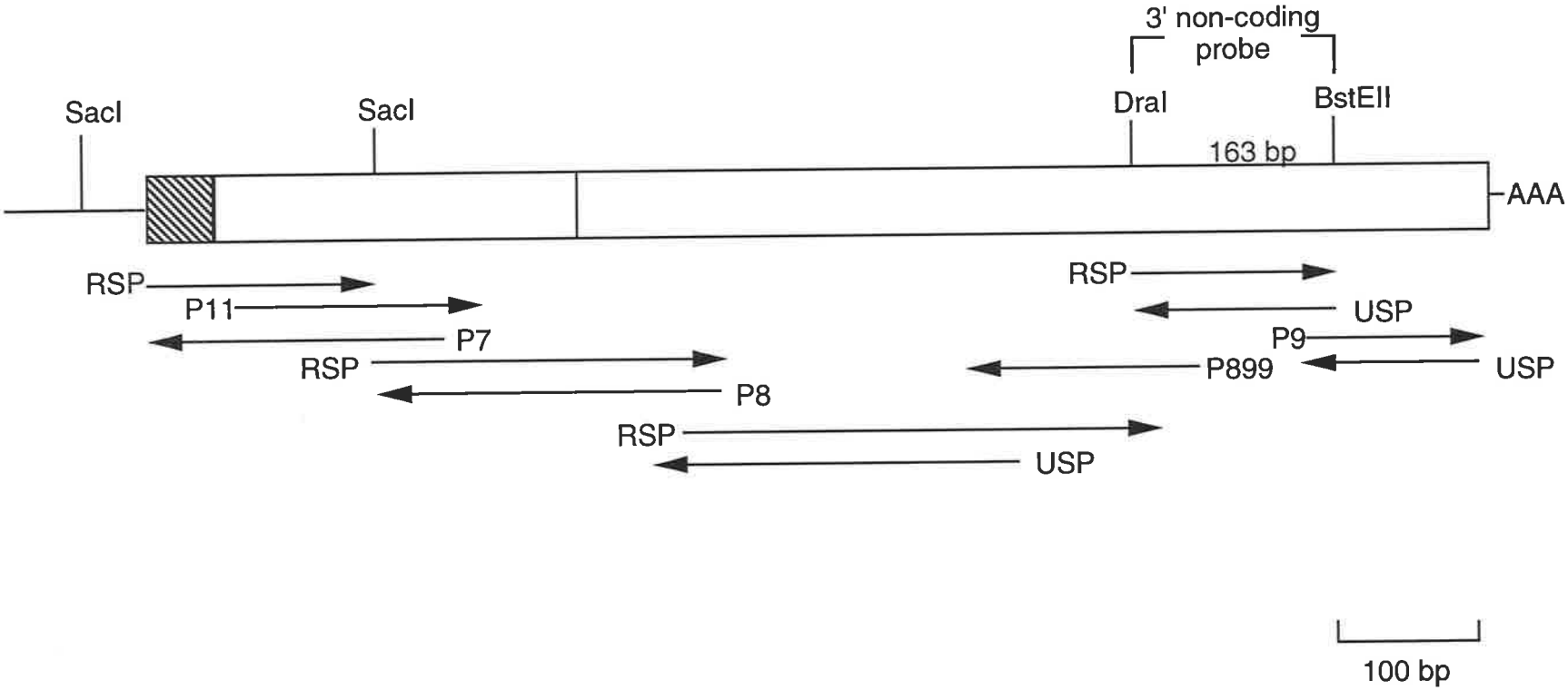
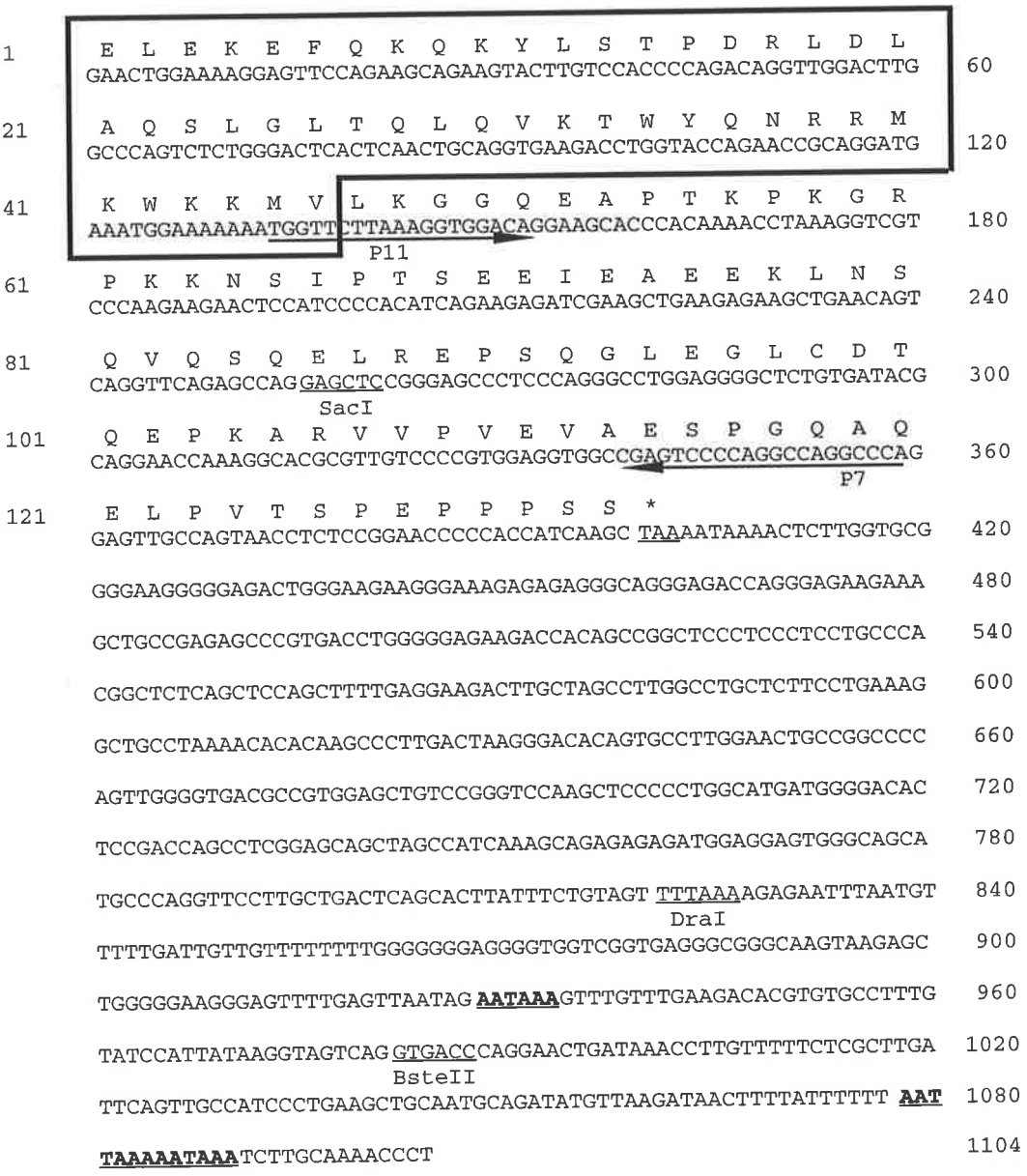

3' non-coding
probe
SacI
SacI
DraI
BstEII
163 bp
AAA
RSP
P11
P7
RSP
P8
RSP
USP
RSP
USP
P9
P899
USP
100 bp

Figure 5.6 : The Nucleotide Sequence and Predicted Amino Acid of the 3' End of *Barx2* (clone C17B1).

The 3' cDNA end of *Barx2* was sequenced and the resultant nucleotide and predicted amino acid sequence are shown. C17B1 is 1,104 bases in length and contains a partial open reading frame of 133 amino acids. The homeodomain is boxed. Restriction sites used to sub-clone fragments from the 3' end are denoted. The arrows denote the location of the P7 and P11 primers used to amplify the 3' coding probe. The proposed polyadenylation signals are typed in bold and underlined. The stop codon (TAA) is underlined and denoted by an asterisk (nucleotide # 401). Numbers on the left of the figure refer to amino acid position and numbers on the right refer to nucleotide position.


1
E L E K E F Q K Q K Y L S T P D R L D L
GAACTGGAAAAGGAGTTCCAGAAGCAGAAGTACTTGTCCACCCCAGACAGGTTGGACTTG 60
21
A Q S L G L T Q L Q V K T W Y Q N R R M
GCCCAGTCTCTGGGACTCACTCAACTGCAGGTGAAGACCTGGTACCAGAACCGCAGGATG 120
41
K W K K M V L K G G Q E A P T K P K G R
AAATGGAAAAAAATGGTTCTTAAAGGTGGACAGGAAGCACCCACAAAACCTAAAGGTCGT 180
P11
61
P K K N S I P T S E E I E A E E K L N S
CCCAAGAAGAACTCCATCCCCACATCAGAAGAGATCGAAGCTGAAGAGAAGCTGAACAGT 240
81
Q V Q S Q E L R E P S Q G L E G L C D T
CAGGTTCAGAGCCAGGAGCTCCGGGAGCCCTCCCAGGGCCTGGAGGGGCTCTGTGATACG 300
SacI
101
Q E P K A R V V P V E V A E S P G Q A Q
CAGGAACCAAAGGCACGCGTTGTCCCCGTGGAGGTGGCCGAGTCCCCAGGCCAGGCCCAG 360
P7
121
E L P V T S P E P P P S S *
GAGTTGCCAGTAACCTCTCCGGAACCCCCACCATCAAGC TAAAATAAAACTCTTGGTGCG 420
GGGAAGGGGGAGACTGGGAAGAAGGGAAAGAGAGAGGGCAGGGAGACCAGGGAGAAGAAA 480
GCTGCCGAGAGCCCGTGACCTGGGGGAGAAGACCACAGCCGGCTCCCTCCCTCCTGCCCA 540
CGGCTCTCAGCTCCAGCTTTTGAGGAAGACTTGCTAGCCTTGGCCTGCTCTTCCTGAAAG 600
GCTGCCTAAAACACACAAGCCCTTGACTAAGGGACACAGTGCCTTGGAACTGCCGGCCCC 660
AGTTGGGGTGACGCCGTGGAGCTGTCCGGGTCCAAGCTCCCCCTGGCATGATGGGGACAC 720
TCCGACCAGCCTCGGAGCAGCTAGCCATCAAAGCAGAGAGAGATGGAGGAGTGGGCAGCA 780
TGCCCAGGTTCCTTGCTGACTCAGCACTTATTTCTGTAGT TTTAAAAGAGAATTTAATGT 840
DraI
TTTTGATTGTTGTTTTTTTTGGGGGGGAGGGGTGGTCGGTGAGGGCGGGCAAGTAAGAGC 900
TGGGGGAAGGGAGTTTTGAGTTAATAG AATAAAGTTTGTTTGAAGACACGTGTGCCTTTG 960
TATCCATTATAAGGTAGTCAG GTGACCCAGGAACTGATAAACCTTGTTTTTCTCGCTTGA 1020
BsteII
TTCAGTTGCCATCCCTGAAGCTGCAATGCAGATATGTTAAGATAACTTTTATTTTTT AAT 1080
TAAAAATAAATCTTGCAAAACCCT 1104

5.5 Sequence Analysis of Ovine *Barx2*

Excluding the glutamic acid (E) residue at the N-terminal end of the ovine Barx2 homeodomain sequence, which may be a result of degenerate primer mismatch during RT-PCR amplification, the partial amino acid sequence of the ovine Barx2 homeodomain is identical to both the human Barx2 homeodomain (Accession number: AF031924) and the mouse Barx2 homeodomain (Jones et al., 1997), and 82.5% similar to the mouse Barx1 homeodomain (Tissier-Seta et al., 1995). As found in other members of the *Bar* family, the ovine *Barx2* gene contains the two atypical residues, threonine and tyrosine within helix 3 at position 47 and 49 respectively (Fig. 5.7).

Downstream from the homeodomain, a leucine residue, two glycine residues, a 17 amino acid region (PTKPKGRPKKNSIPTSE) and three nearby glutamic acid residues are common to all four proteins. The carboxyl end of Barx2 shared an amino acid sequence identity of 84% with the mouse Barx2 carboxyl end and 63% with the mouse Barx1 carboxyl end. The ovine *Barx2* 3' non-coding region shares a 60.2% nucleotide sequence identity with the mouse *Barx2* 3' non-coding region (Fig. 5.8).

5.6 Polyadenylation Signals of *Barx2*

Barx2 contains four motifs for the polyadenylation signal (AATAAA), at positions 402, 928, 1077 and 1085 in the 3' non-coding sequence (Fig. 5.6). As only one band was detected in the Northern blot (Fig. 5.3) using the homeobox and coding probes, the polyadenylation signal at position 402 was determined to be non-functional because a shorter transcript of approximately 1,417 bases was not detected. It is possible that the Northern blot was not able to detect the small size difference (149/157bp) between the predicted transcripts produced by the polyadenylation signals at positions 928 and 1077/1085 due to insufficient size-separation of the RNA. However, two products were detected in the Southern blot of the 3' RACE products consistent with active polyA signals at these sites (see Fig 5.2). It is not known which of the polyadenylation signals at positions 1077 and 1085 are functional however, both are near the end of the mRNA and tandem polyadenylation signals are not uncommon.

Figure 5.7 : Comparison of Vertebrate Barx proteins.

The partial sequence of the ovine Barx2 homeodomain and C-terminal end are compared to human Barx2 (Accession number: AF031924), mouse Barx1 (Tissier-Seta et al., 1995) and mouse Barx2 (Jones et al., 1997). The partial homeodomain is boxed. Dashes indicate sequence identity with ovine Barx2. Note that excluding the glutamic acid (E) residue at the N-terminal end of the ovine sequence, which may be a result of degenerate primer mismatch during RT-PCR amplification, the Barx2 homeodomain from each species is conserved. Residues that are shared with each protein are typed in bold. Downstream from the homeodomain, a leucine residue, two glycine residues, a 17 amino acid region and three nearby glutamic acid residues are common to all four proteins. An extra leucine residue present in the comparable region of Barx1, downstream of the homeodomain is shown beneath the aligned sequences. The carboxy-terminal end of each protein is indicated by an asterisk.

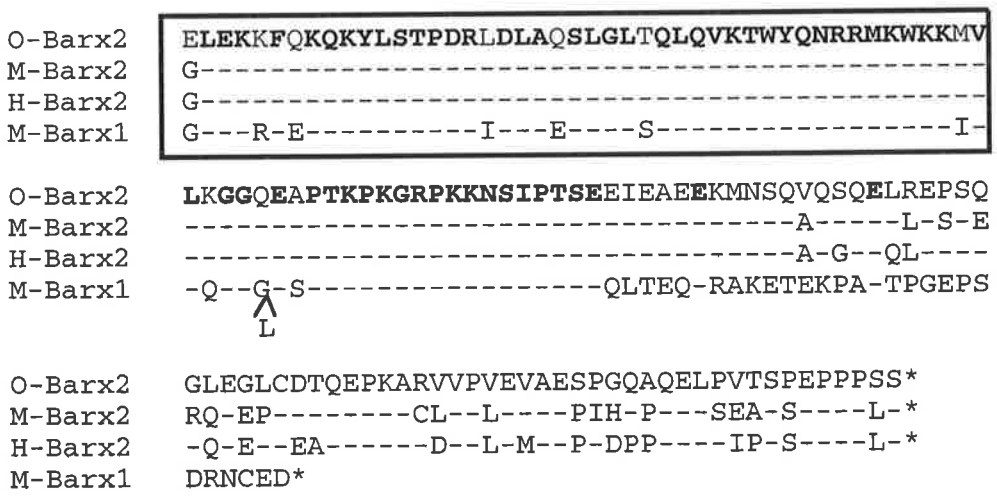
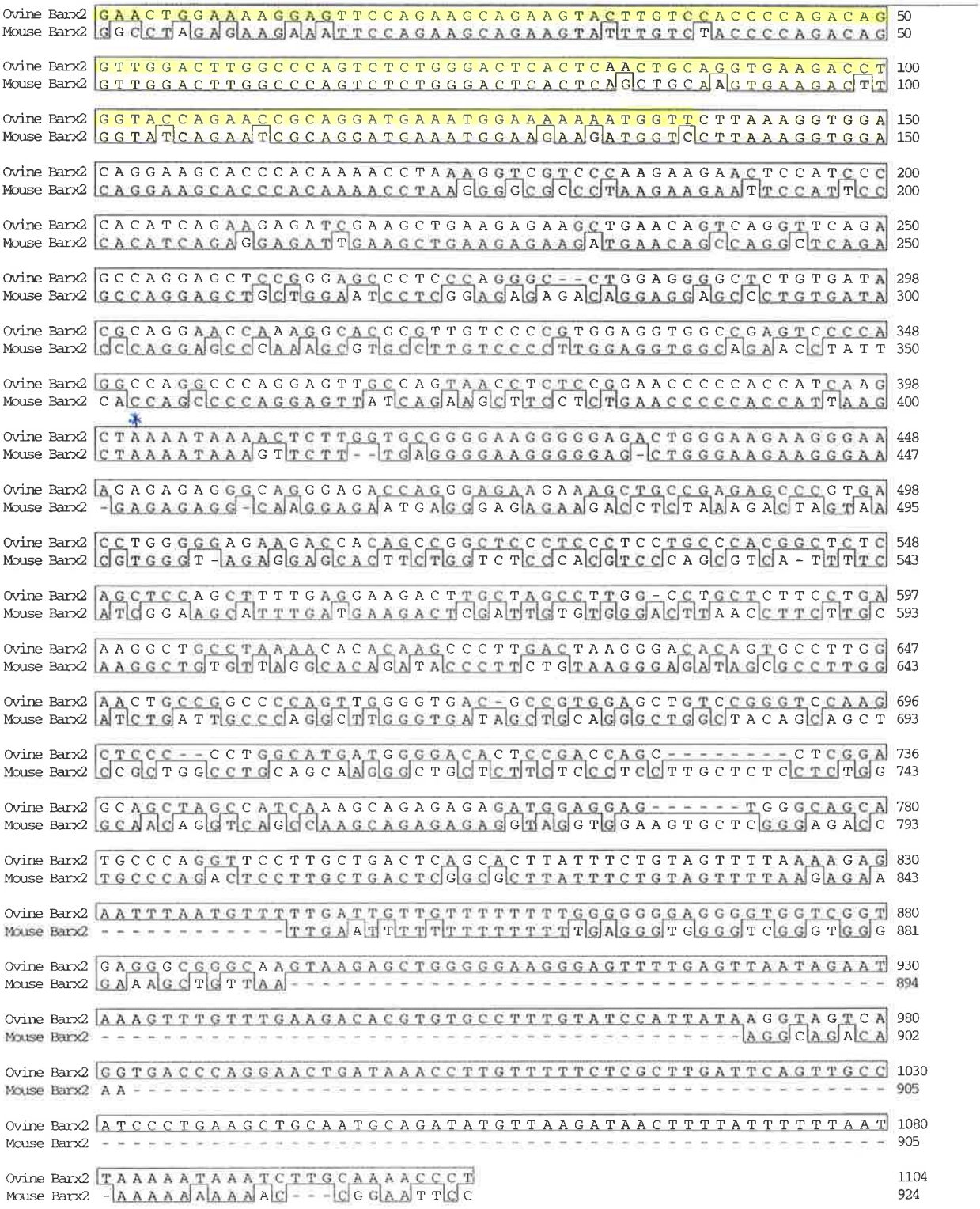

O-Barx2 ELEKKFQKQKYLSTPDRLDLAQSLGLTQLQVKTWYQNRRMKWKKMV
M-Barx2 G----------
H-Barx2 G----------
M-Barx1 G---R-E----------I---E----S----------I-
O-Barx2 LKGGQEAPTKPKGRPKKNSIPTSEEIEAEEKMNSQVQSQELREPSQ
M-Barx2 ----------A-----L-S-E
H-Barx2 ----------A-G--QL----
M-Barx1 -Q--G-S----------QLTEQ-RAKETEKPA-TPGEPS
L
O-Barx2 GLEGLCDTQEPKARVVPVEVAESPGQAQELPVTSPEPPPSS*
M-Barx2 RQ-EP--------CL--L----PIH-P---SEA-S----L-*
H-Barx2 -Q-E--EA------D--L-M--P-DPP----IP-S----L-*
M-Barx1 DRNCED*

Figure 5.8 : Comparison of Mouse and Sheep *Barx2* 3' end.

DNA sequence from the 3' end of sheep *Barx2* (top sequence) and mouse *Barx2* (bottom sequence) have been aligned using DNASTAR, Inc, software. The stop codon TAA is depicted by an asterix. The homeobox is highlighted in yellow. Boxed nucleotides indicate sequence identity with ovine *Barx2*.



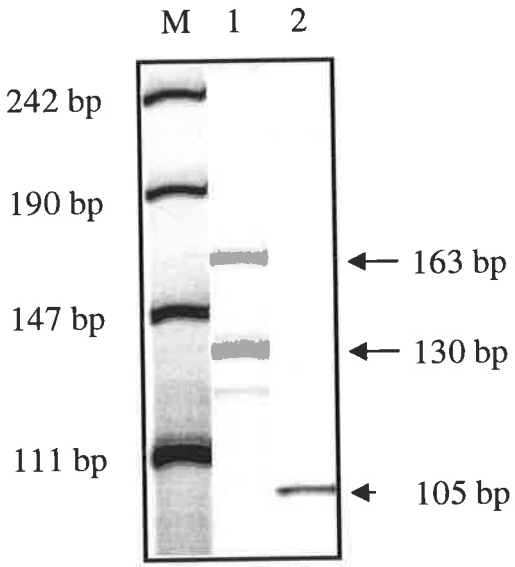
Decoration 'Decoration #2': Box residues that match Ovine Barx2 exactly.

Figure 5.9 : RNA Protection Analysis of *Barx2* Gene Transcripts in Total Wool Follicle RNA from a Corriedale Sheep.

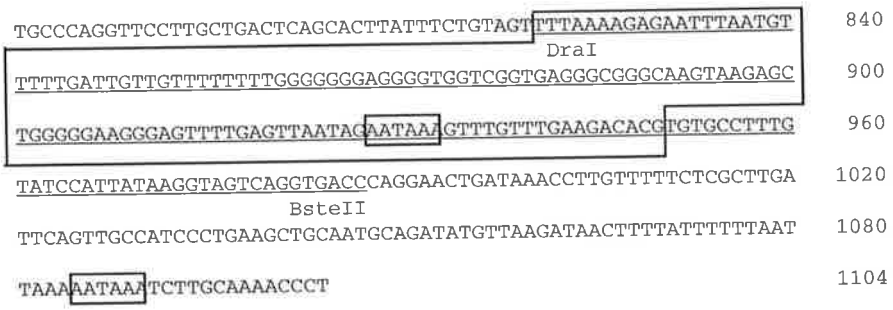
A. Ten µg of total RNA isolated from Corriedale wool follicles was analysed by RNA protection for the presence of *Barx2* gene transcripts using the 3' non-coding antisense probe (clone C17B1-NC; lane 1) which had been radiolabelled with α– ^{32}P- rUTP. The 3' non-coding fragment contains the first polyadenylation signal of *Barx2*. The homeobox probe (clone C17, lane 2) was used as a control and yielded a protected product of 105 bases (arrowhead). The products of the protection reactions were fractionated in a denaturing 6% (19:1) urea-PAGE system and the gel exposed to a phosphor screen for 48 hours. The position of the 160 base and 130 base protected *Barx2* RNA are indicated on the right side of the figure. The signal for both protected RNAs is equally intense. M = pUC19 markers, sizes of which are given to the left of the figure.

B. DNA sequence of the 3' end of ovine *Barx2* between nucleotides 780 – 1104 (taken from Fig. 5.6). The 163 bp DNA fragment used as a probe in the RNA protection in A. above, is underlined. The predicted 130 bp protected fragment is boxed. Two polyadenylation signals located at positions 928 and 1085 are also boxed. The location of the restriction enzymes DraI and BsteII, used to excise the 163 bp 3' non-coding fragment are shown.

A


M
1
2
242 bp
190 bp
147 bp
111 bp
163 bp
130 bp
105 bp

B


TGCCCAGGTTCCTTGCTGACTCAGCACTTATTTCTGTAGTTTTAAAAGAGAATTTAATGT
DraI
840
TTTTGATTGTTGTTTTTTTTGGGGGGGAGGGGTGGTCGGTGAGGGCGGGCAAGTAAGAGC
900
TGGGGGAAGGGAGTTTTGAGTTAATAGAATAAAGTTTGTTTGAAGACACGTGTGCCTTTG
960
TATCCATTATAAGGTAGTCAGGTGACCCAGGAACTGATAAACCTTGTTTTTCTCGCTTGA
1020
BsteII
TTCAGTTGCCATCCCTGAAGCTGCAATGCAGATATGTTAAGATAACTTTTATTTTTTAAT
1080
TAAAAATAAATCTTGCAAAACCCT
1104

To determine if the polyadenylation signal located at 928 was functional, an RNA protection was performed using Corriedale wool follicle RNA and a 3' non-coding probe (Fig 5.9 B) which encompassed the polyadenylation site. The RNA protection assay revealed two protected products of 163 bases (full-length protected probe) and approximately 130 bases (partially protected probe; lane 1, Fig. 5.9 A). The 160 base protected product was the result of the probe hybridizing to a RNA transcript which is produced when the second polyadenylation signal at position 1085 is utilized by *Barx2*. The 130 base protected product is produced when the first polyadenylation signal at 928 is utilized by *Barx2*. Signals from the protected products were found to be of an equal intensity.

5.7 Isolation of an Ovine *Barx2* Genomic Lambda Clone

Lambda clones containing the ovine *Barx2* gene were isolated from a λ-GT11 sheep genomic DNA library. Approximately 5 genome equivalents of clones were plated onto TB agar plates and screened with a 3' non-coding fragment (clone C17B1-NC). Two clones named 2A and 2B hybridised to the probe after a final wash at a stringency of 0.1 X SSC at 65^{o}C and were purified by four rounds of screening.

5.8 Analysis of the *Barx2* Genomic Clone 2A

Clone 2A contained a 16.0 kb SalI insert (see Fig 5.10 A, lane 4). The DNA was then cleaved with restriction endonucleases to determine the location of the restriction enzyme sites. The position of the 3' end of *Barx2* within clone 2A was located by Southern blot using a *Barx2* non-coding probe (clone C17B1-NC, Fig. 5.10 B) and was found to lie within a 4.2 kb HindIII fragment. The exact position within the 4.2 kb HindIII fragment is not known. A detailed restriction map of clone 2A and the location of the 4.2 kb HindIII fragment that contains the *Barx2* non-coding region are shown in Fig. 5.10 C.

5.9 Species Conservation of *Barx2*

To determine if the coding region of the *Barx2* gene is conserved between species a 3' coding fragment isolated from *Barx2* (clone C17B1-COD; see Appendix 1 C for plasmid

Figure 5.10 : Restriction Map of the *Barx2* Genomic Lambda Clone 2A.

A. DNA from the genomic clone 2A was digested with single and multiple restriction endonucleases as indicated above each lane. Cleaved fragments were size-separated through a 1% agarose gel by gel electrophoresis and the gel stained with EtBr. The size of the fragments were determined by comparison with SPP-1 and Lambda Hi-Lo markers (lane M). Marker sizes (kb) are shown on the right.

B. Southern blot analysis of clone 2A. Fragments from **A.** were transferred to Zetaprobe GT membrane and probed with a 163 bp *Barx2* non-coding probe from clone C17B1-NC (see Appendix 1 D for clone reference). Bands indicate which fragments hybridised to the probe. Hybridised filters were washed in 0.1 X SSC, 0.1% SDS at 65^{o}C for 30 minutes and exposed to an autoradiograph for 1 hour. Size markers are shown (in kb).

C. Restriction map of the *Barx2* genomic clone 2A. The *Barx2* 3' non-coding region was located within the 4.2 kb HindIII fragment. The location of SalI (S), EcoR1 (E) and HindIII (H) sites are denoted.

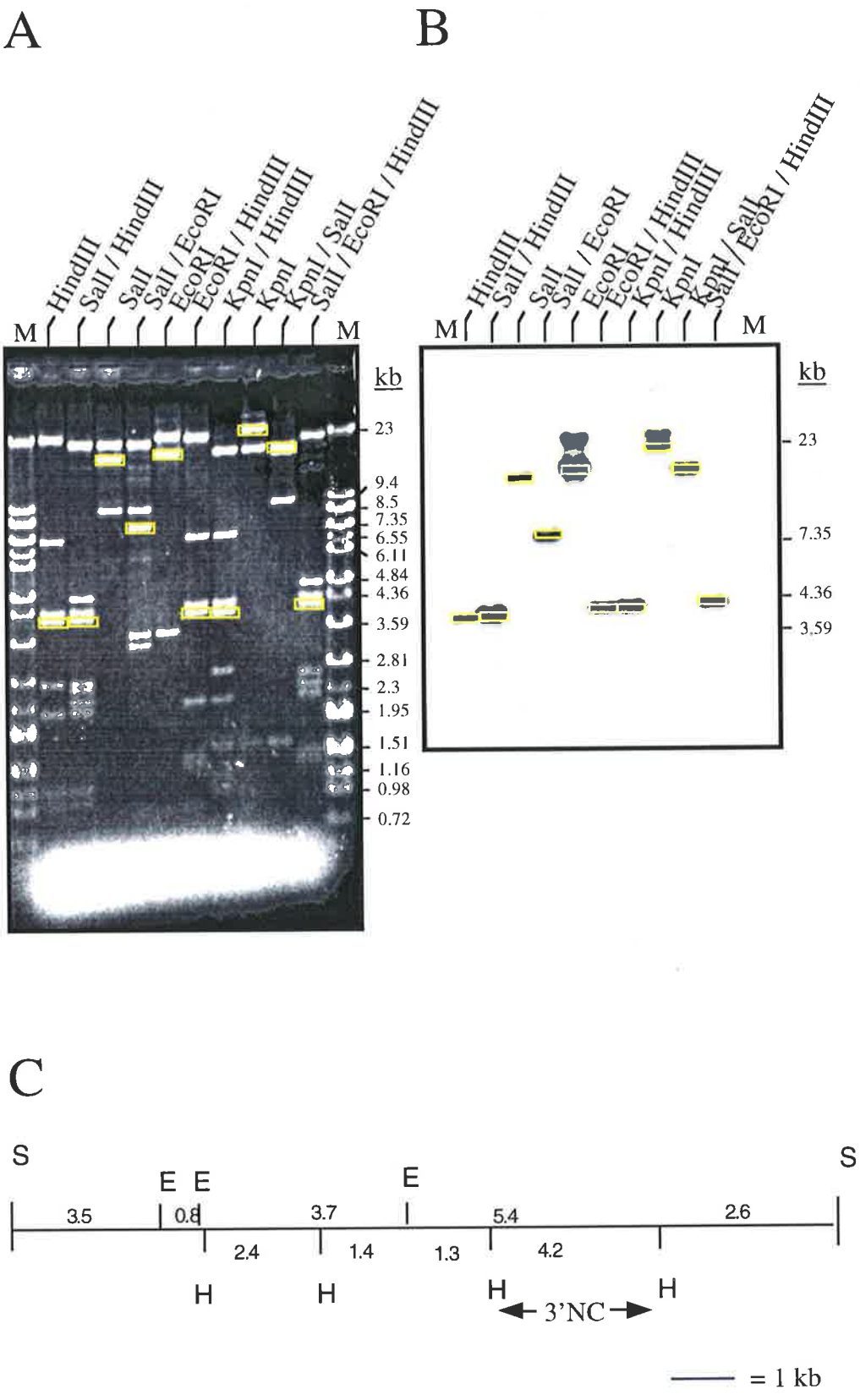
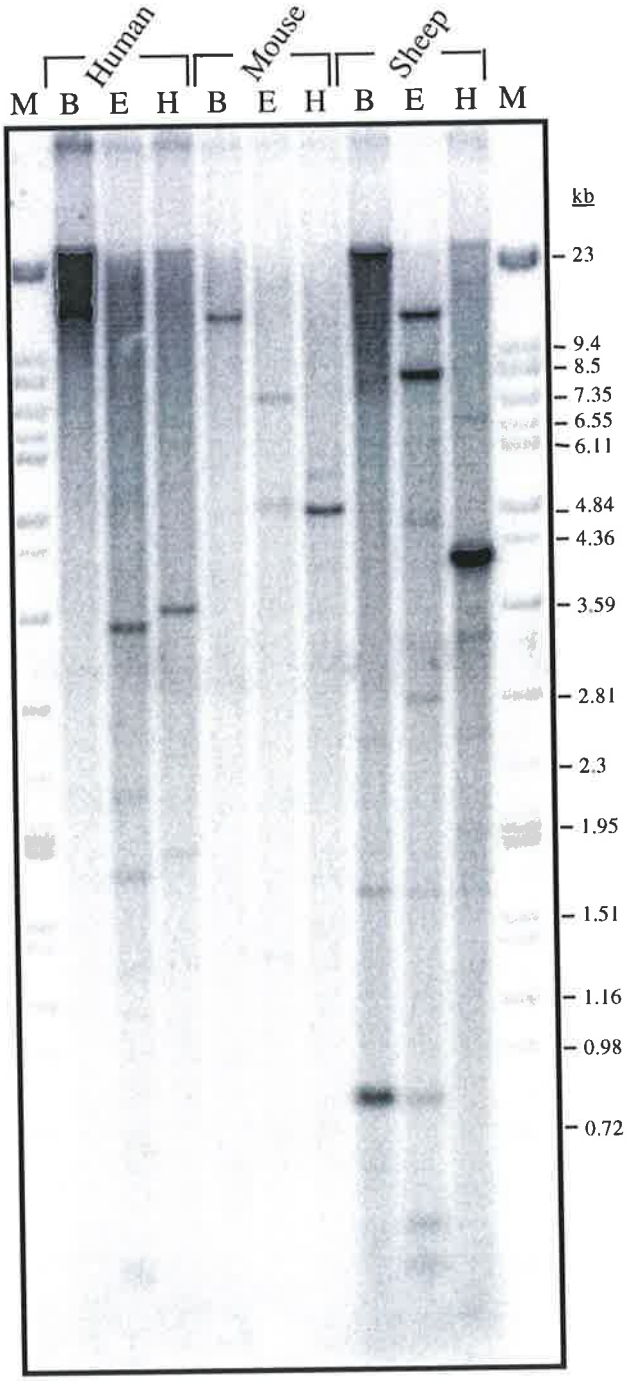

A
B
M
HindIII
SalI / HindIII
SalI
SalI / EcoRI
EcoRI
EcoRI / HindIII
KpnI / HindIII
KpnI
KpnI / SalI
SalI / EcoRI / HindIII
M
kb
23
9.4
8.5
7.35
6.55
6.11
4.84
4.36
3.59
2.81
2.3
1.95
1.51
1.16
0.98
0.72
C
S
E E
E
S
3.5
0.8
3.7
5.4
2.6
2.4
1.4
1.3
4.2
H
H
H
H
3'NC
= 1 kb

Figure 5.11 : Hybridisation Analysis of a Genomic Zoo Blot.

Mouse, human and sheep total genomic DNA (10 µg / lane) were digested with BamHI (B), EcoR1 (E), and HindIII (H), size-separated on a 1% agarose gel, transferred to Zetaprobe GT membrane and probed with the resected *Barx2* coding region fragment (clone C17B1-COD, see Appendix 1 C). The filter was washed in 2.0 X SSC, 0.1% SDS at 65°C for 30 minutes and exposed to an autoradiograph for 21 days at -80°C. M = Size markers are shown in kilobase pairs at the right of the figure.


Human
Mouse
Sheep
M B E H B E H B E H M
kb
23
9.4
8.5
7.35
6.55
6.11
4.84
4.36
3.59
2.81
2.3
1.95
1.51
1.16
0.98
0.72

map) was used to probe a Southern blot containing sheep, human and mouse genomic DNA cleaved separately with the enzymes BamHI, EcoR1 and HindIII (Fig. 5.11). The probe hybridised most strongly to sheep DNA but also bound to mouse and human DNA. The sizes of the hybridised fragments are given in table 1.

Table 1. Sizes of the mouse, human and sheep genomic DNA fragments that hybridised to the ovine *Barx2* coding probe.

Enzyme/ Species	Human	Mouse	Sheep
BamHI	14.5 kb	15 kb	>20 kb
EcoR1	7.4 kb	3.5 kb	14.5 kb and 8.5 kb
HindIII	4.9 kb	3.6 kb	4.2 kb

5.10 Characterisation of *Barx2* Expression by *In Situ* Hybridisation

5.10.1 Expression of *Barx2* in Embryonic Sheep skin

Barx2 expression during the development of embryonic sheep skin was analysed by *in situ* hybridisation using a gene-specific probe isolated from the *Barx2* 3' coding region (clone C17B1-COD). The probe was hybridised *in situ* to skin from the mid-flank and snout regions at the embryonic time points of E40, E51, E65, E75, E80, E88, E97 and E135 (mid-flank region) and E71, E81, E88, E97 and E135 (snout region). *Barx2* exhibited a differentially- and spatially - restricted expression pattern during both epidermal and follicle development. In the pre-follicle stage *Barx2* was expressed predominantly in the ectoderm which is 1-2 cells in thickness but was not expressed in the underlying dermis (Fig. 5.12 A and B). In the initial stage of follicle development, when localised thickening was first evident at discrete sites in the ectoderm, *Barx2* expression was present throughout the basal and suprabasal layers of the ectoderm but was specifically down-regulated in the ectodermal-derived cells of the incipient follicle placode and absent from the mesenchymal cell aggregate located directly beneath the placode (Fig. 5.12 C and D).

As the solid plug of epidermal cells moved downward into the dermis *Barx2* expression commenced in the inner epithelial cells of the placode and was present in the epidermis overlying the placode (Fig. 5.13 A-D). At this stage, occasional outer cuboidal

Figure 5.12 : *In situ* Hybridisation Analysis of *Barx2* Expression in Mid - flank Embryonic Skin.

Brightfield (A and C) and darkfield (B and D) photographs of mid - flank embryonic skin at E51 (A and B) and E65 (C and D), hybridised with a *Barx2* cRNA coding probe (clone C17B1-COD, see Appendix 1 C for plasmid reference). Sections were treated with a final wash stringency of 0.1 X SSPE at 65°C for 30 minutes and exposed for 10 days at 4°C. Arrow points to the follicle placode where down regulation of *Barx2* is evident. The border between the epidermis and the dermis is indicated by the dashed line. Bar = 25 µm.

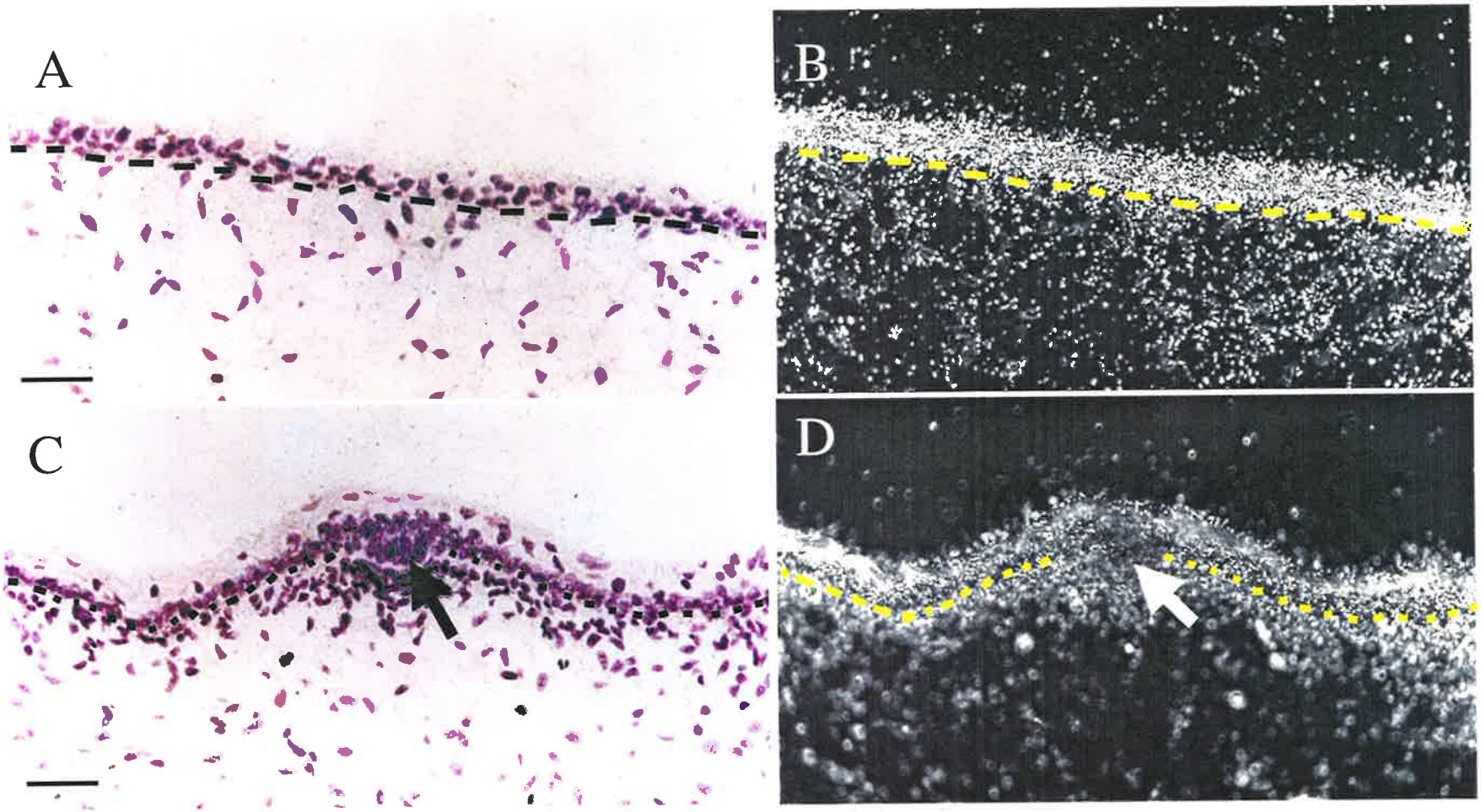
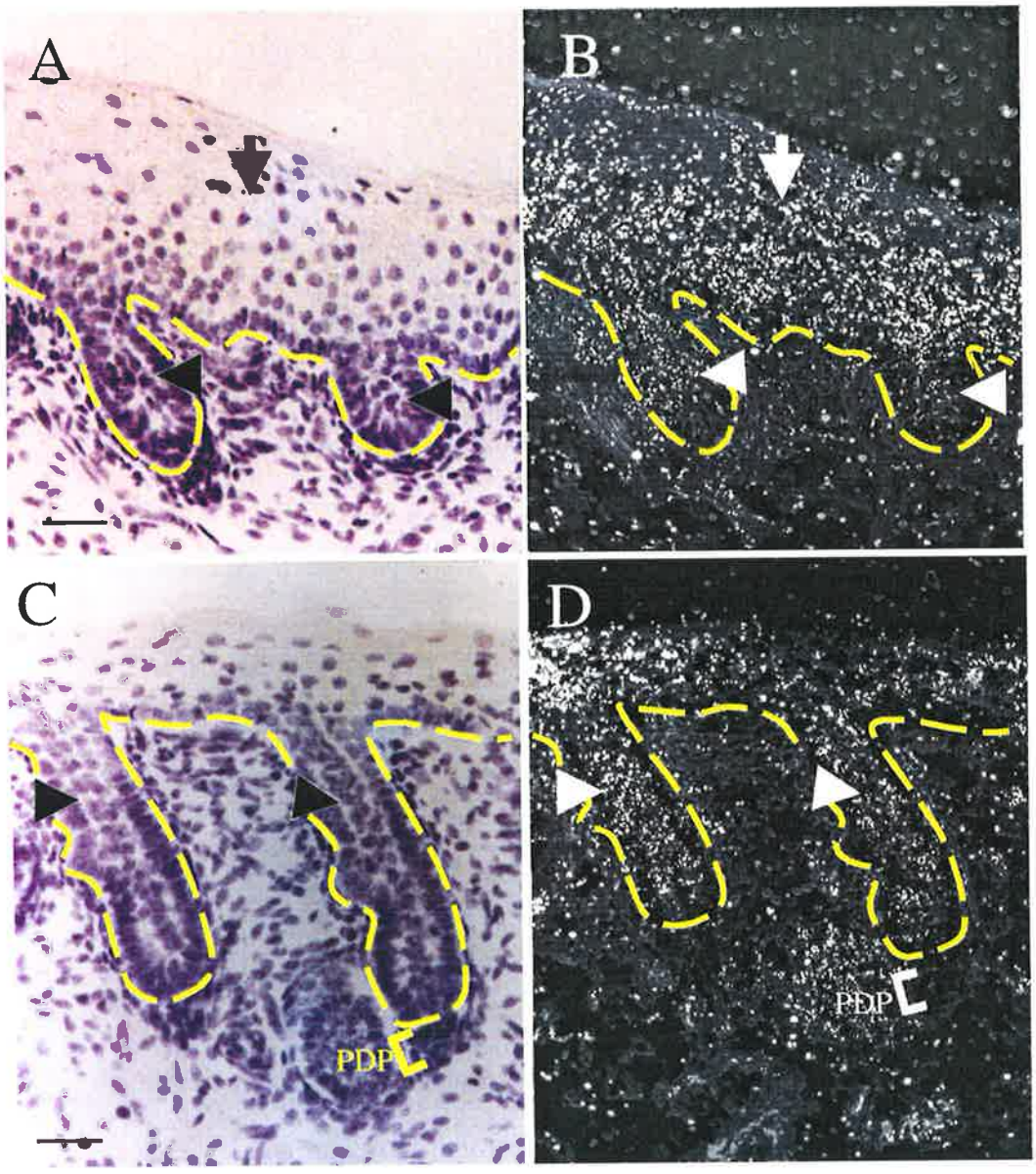

A
B
C
D

Figure 5.13 : *In situ* Hybridisation Analysis of *Barx2* Expression in Embryonic Skin.

Brightfield (A and C) and darkfield (Band D) photographs of sheep embryonic skin at E72. Probe, wash conditions and exposure time were as for the skin sections in Figure 5.12. Arrows indicate *Barx2* expression in the epidermis and arrowheads indicate *Barx2* expression in the epithelial component of the placode. The basement membrane separating the epidermis and the dermis is indicated by the dashed line. PDP = Predermal papilla. Bar in A and C = 20 μm.


A
B
C
D
PDP
PDP

cells in the upper part of the placode expressed *Barx2* whereas at more advanced stages, as the placode moved deep into the dermis *Barx2* was expressed in most cells of the placode (Fig. 5.14 A and B). *Barx2* expression was also detected in the bulge which is the putative location of hair follicle stem cells (Cotsarelis et al., 1990) but tapered off towards the leading edge of the placode where epithelial cells formed the hair plug (Fig. 5.14 A and B). Expression was absent from the adjacent pre-dermal papilla.

In more advanced embryonic follicles, *Barx2* was expressed along the entire length of the outer root sheath, from the middle of the bulb to the epidermis where the fibre emerged (Fig. 5.14 C and D). *Barx2* was not expressed in the follicle bulb, dermal papilla, inner root sheath or hair shaft cortex at any stage of development.

5.10.2 Expression of *Barx2* in Embryonic Vibrissae

Expression of ovine *Barx2* RNA was studied in vibrissae (whisker) follicles at the embryonic time points of E71, E81, E88, E97 and E135. *Barx2* was expressed in the vibrissae follicle in regions comparable to the pelage follicle. For example, as the solid plug of epidermal cells moved downward into the dermis, *Barx2* expression was detected in the inner epithelial cells of the placode and cuboidal shaped cells that lined the placode (Fig. 5.15 A and B; c.f. Fig. 5.14 A and B). Expression was also absent from cells at the leading edge of the placode, adjacent mesenchymal cells of the pre-dermal papilla and the dermis (Fig. 5.15 A and B).

In more advanced vibrissae follicles where the bulb and dermal papilla had developed, *Barx2* was expressed in the entire length of the ORS from the middle of the bulb to the epidermis (Fig. 5.16 A and B). In mature vibrissae follicles, a few cells within the blood sinus expressed ovine *Barx2* (Fig. 5.16 D arrowheads).

5.10.3 Expression of *Barx2* in Adult Sheep skin

To examine the cellular distribution of *Barx2* expression in adult sheep skin, *in situ* hybridisation analyses were conducted on adult Corriedale sheep skin taken from the mid-flank region. Expression was predominantly located along the entire length of the follicle outer root sheath but tapered off in the suprabasal layer of the epidermis located next to the

Figure 5.14 : *In situ* Hybridisation Analysis of *Barx2* Expression in Embryonic Pelage Skin.

Brightfield (A and C) and darkfield (B and D) photographs of the developing placode in embryonic skin at E81 (A and B) and E135 (C and D). Probe, wash conditions and exposure time were as for the skin sections in Figure 5.12. Arrows in A and B indicate *Barx2* expression in the epithelial component of the placode and bulge. Note there is no *Barx2* expression in the cells at the bottom of the placode and predermal papilla. Arrowheads in Figure D indicate *Barx2* expression in the outer root sheath of the wool follicle. The basement membrane separating the epidermis and the dermis is indicated by the dashed line. B = Bulge, DP = Dermal papilla, PDP = Predermal papilla. Bar in A = 20 μm and bar in C = 105 μm.

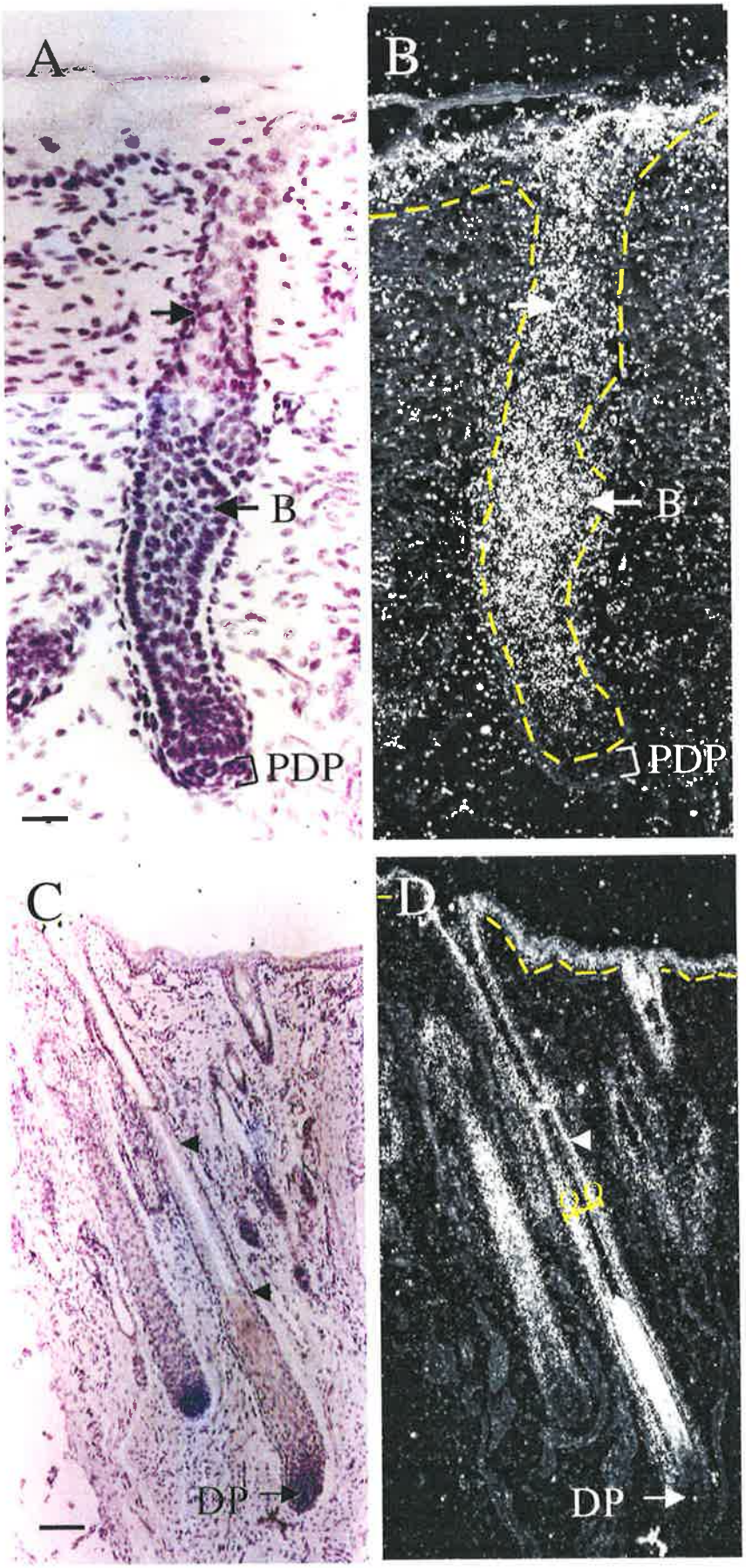

A
B
PDP
B
B
PDP
C
DP
D
DP

Figure 5.15 : *In situ* Hybridisation Analysis of *Barx2* Expression in Embryonic Vibrissa.

Brightfield (A) and darkfield (B) photographs of vibrissae from E71 skin. Probe, wash conditions and exposure time were as for the skin sections in Figure 5.12. *Barx2* is expressed in the inner cells of the developing placode. Note that the epidermis is not present above the placode. PDP = Predermal papilla. Bar in A = 40 μm.

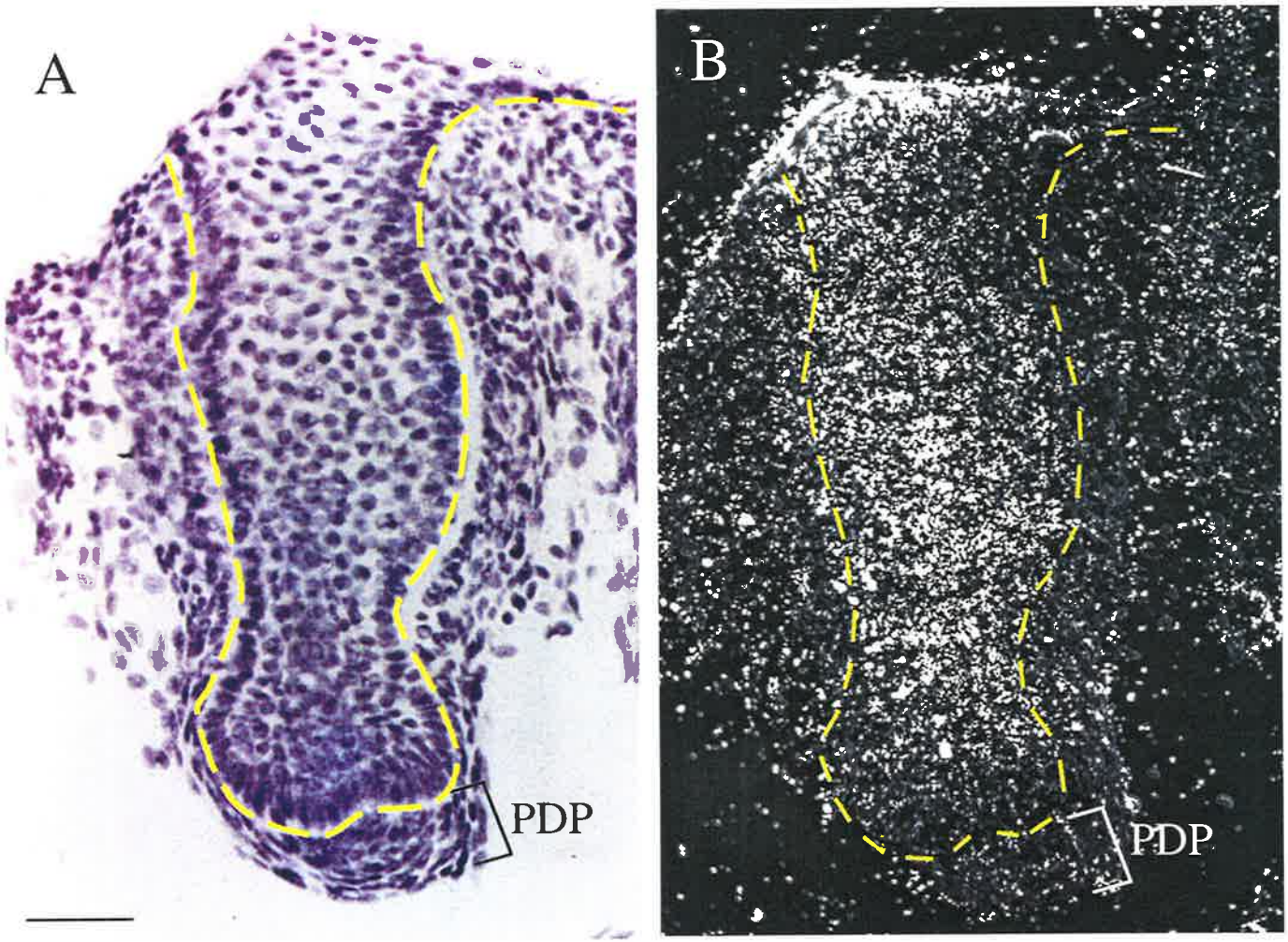
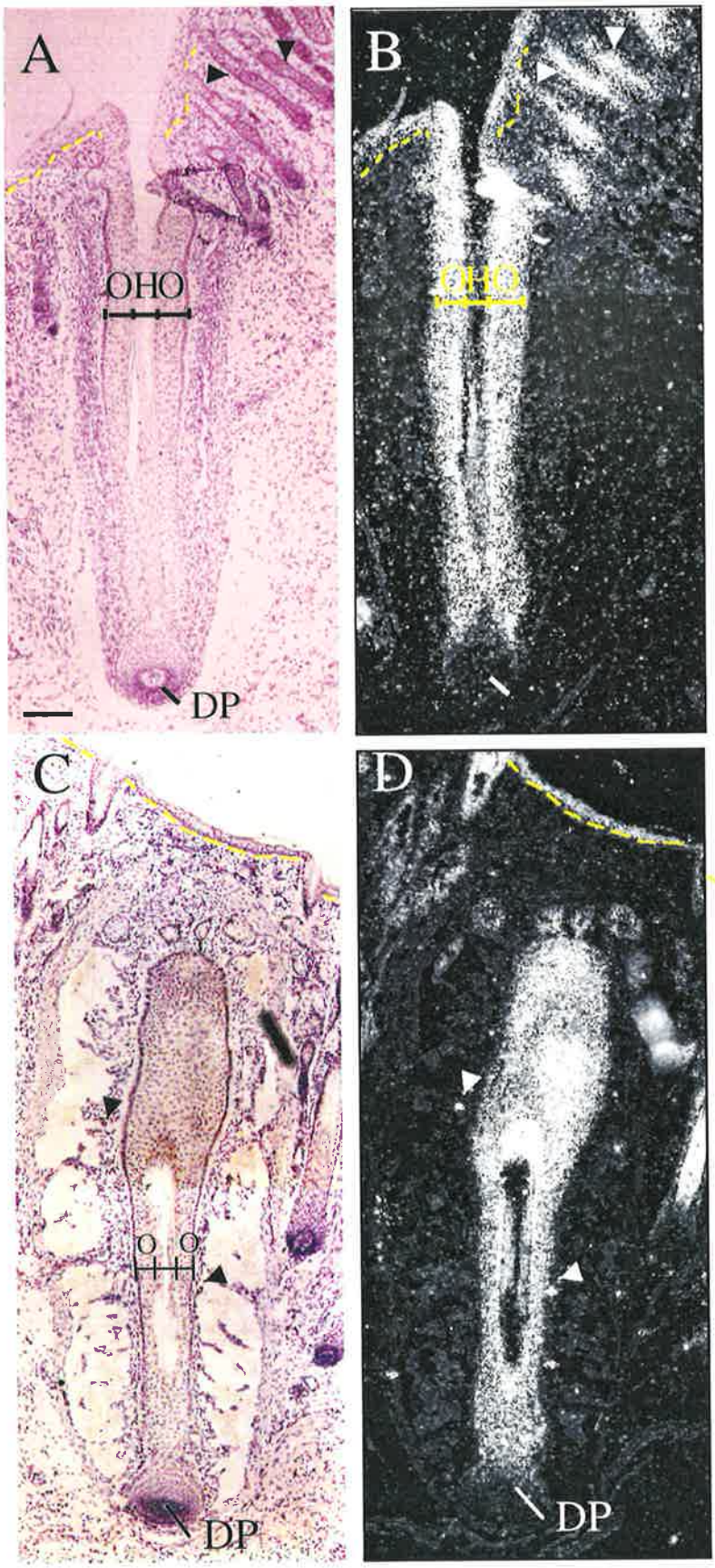

A
PDP
B
PDP

Figure 5.16 : *In situ* Hybridisation Analysis of Ovine *Barx2* Expression in Embryonic Vibrissa.

Brightfield (A and C) and darkfield (B and D) photographs of embryonic sheep skin from the snout region at E97 (A and B) and E135 (C and D) skin containing a vibrissa. Probe, wash conditions and exposure time were as for the skin sections in Figure 5.12. *Barx2* is expressed in the vibrissa outer root sheath. Arrowheads in C and D indicate expression in a few cells in the blood sinus. Note that the upper shaft is not present in A or C because the follicle was cut at a slightly oblique angle. O = Outer root sheath, DP = dermal papilla and H = Hair shaft. Bar (all panels) = 100 μm.


A
OHO
DP
B
OHO
C
O
O
DP
D
DP

protruding fibre (Fig. 5.17 A and B; also see Fig. 5.16 A, B). A few cells of the upper dermis expressed *Barx2* resulting in a scattered pattern throughout the dermis (Fig. 5.17 B). Dermal cells expressing *Barx2* appeared to have bi-lobed nuclei possibly undergoing cell division (Fig. 5.17 C, arrow heads) and some contained multi-lobed nuclei, typical of leukocytes (Fig. 5.17 D arrows).

As the ectoderm developed from the pre-follicle to the adult stage, *Barx2* expression appeared to be gradually down-regulated in the epidermis; compare embryonic (Fig. 5.12 A, B) to pre-birth (Fig. 5.18 A, B) and adult (Fig. 5.18 C, D). In addition to this gradual decline, regional variation in *Barx2* expression was noted. Generally, *Barx2* expression levels were higher in thicker, more cellularised epidermis (compare Fig. 5.13 B with Fig. 5.13 D) and at sites of follicle openings. *Barx2* was weakly expressed in the basal layer of the sebaceous gland and down-regulated in sebocytes located at the centre of the gland (Fig. 5.18 E, F).

5.11 Screening of Sheep Tissues for *Barx2* Expression

Expression of *Barx2* was studied in a range of sheep organs by RNA protection (Fig. 5.19 A and B). *Barx2* was expressed in sheep oesophagus and tongue and weakly in the thymus but was not expressed in skeletal muscle, brain, heart, kidney, spleen, liver, intestine, small intestine or lung.

Barx2 expression was also located by *in situ* hybridisation in the basal and suprabasal epithelial cell layers of the sheep tongue (Fig. 5.20). Expression was not detected in the upper epidermis or the filiform papilla. The location of *Barx2* expression in the sheep oesophagus and thymus were not determined by *in situ* hybridisation.

5.12 Discussion

5.12.1 Characterisation of *Barx2*

Ovine *Barx2* was placed as the second vertebrate member of the *Bar* family and encodes a predicted 670 amino acid protein. This is104 amino acids larger than the homologous mouse Barx2, a 566 amino acid protein. The homeodomain of ovine *Barx2* contains two atypical residues for the *Bar* family, threonine and tyrosine at positions 47 and

Figure 5.17 : *Barx2* Expression in Adult Sheep Skin.

Brightfield (A and C) and darkfield (B and D) photographs of adult sheep skin. Probe, wash conditions and exposure time were as for the skin sections in Figure 5.12. Arrowheads in A and B indicate cells in the upper dermis expressing *Barx2*. Arrow in A and B indicates *Barx2* expression in the outer root sheath. C and D are photographs of cells from the upper dermis (taken at a higher magnification). Arrowheads in C indicate cells that may be dividing and expressing *Barx2*. Arrows in D indicate cells with a multi-lobed appearance typical of leukocytes. The basement membrane separating the epidermis and the dermis is indicated by the dashed line. Bar in A = 100 μm and bar in C and D = 10 μm.

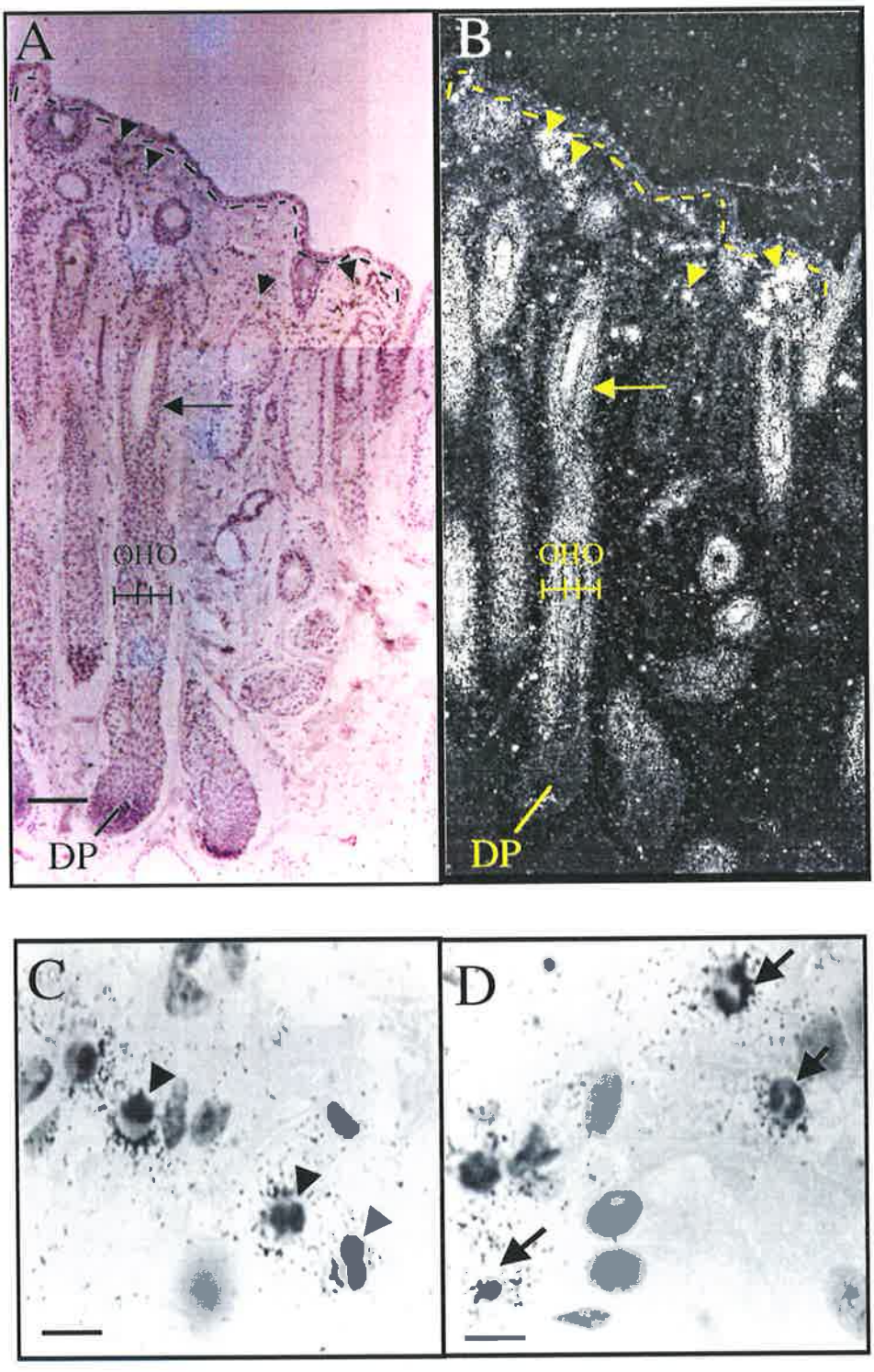

A
B
OHO
OHO
DP
DP
C
D

Figure 5.18 : *Barx2* Expression in Sheep Epidermis and Sebaceous Gland.

Brightfield (A, C and E) and darkfield (B, D and E) photographs of longitudinal sheep skin sections. Probe, wash conditions and exposure time were as for the skin sections in Figure 5.12. *Barx2* is expressed in the epidermal basal and suprabasal layers several days prior to birth at E135 (A and B). *Barx2* is weakly expressed in the adult epidermis (C and D). Note the intense expression of *Barx2* in a cell in the upper dermis just below the basal epidermal layer (arrow). *Barx2* expression is primarily associated with the basal layer of the sebaceous gland (E and F). Bar in A, C and E = 15 μm

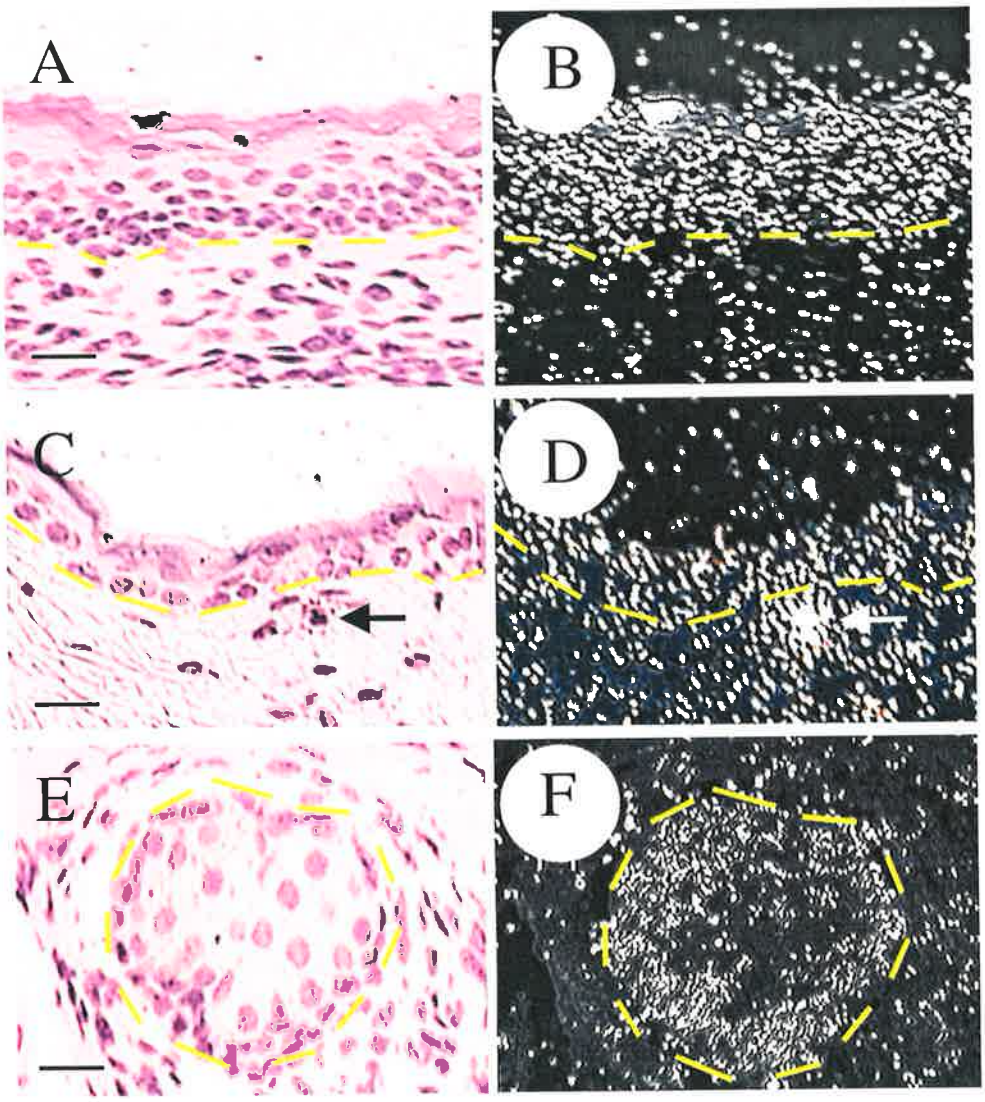

A
B
C
D
E
F

Figure 5.19 : RNA Protection Analysis of Ovine *Barx2* Gene Expression in Total RNA from Several Tissues of a Merino Sheep.

Total RNA was isolated from an adult Merino sheep **A.** oesophagus, spleen, thymus, small intestine, tongue and wool follicle **B.** wool follicle, spleen, thymus, brain, kidney, heart, intestine, small intestine and lung. 10 µg of RNA from these tissue samples was analysed by RNA protection for the presence of *Barx2* expression using the C17B1-COD antisense probe (see Appendix 1 C) which had been radio-labelled with α– ^{32}P-rUTP. The products of the protection reactions were fractionated in a denaturing 6% (19:1) urea-PAGE system and the gel exposed to a phosphor-image screen for 72 hours. The probe controls (yeast RNA +/- RNase) were co-fractionated with test samples on the gel. The position of the 223 base protected *Barx2* RNA detected in the wool follicle, oesophagus and tongue samples is indicated on the right side of the figure by the arrow. Some full-length undigested probe of 293 bases is present in each lane and is indicated by the arrowhead.

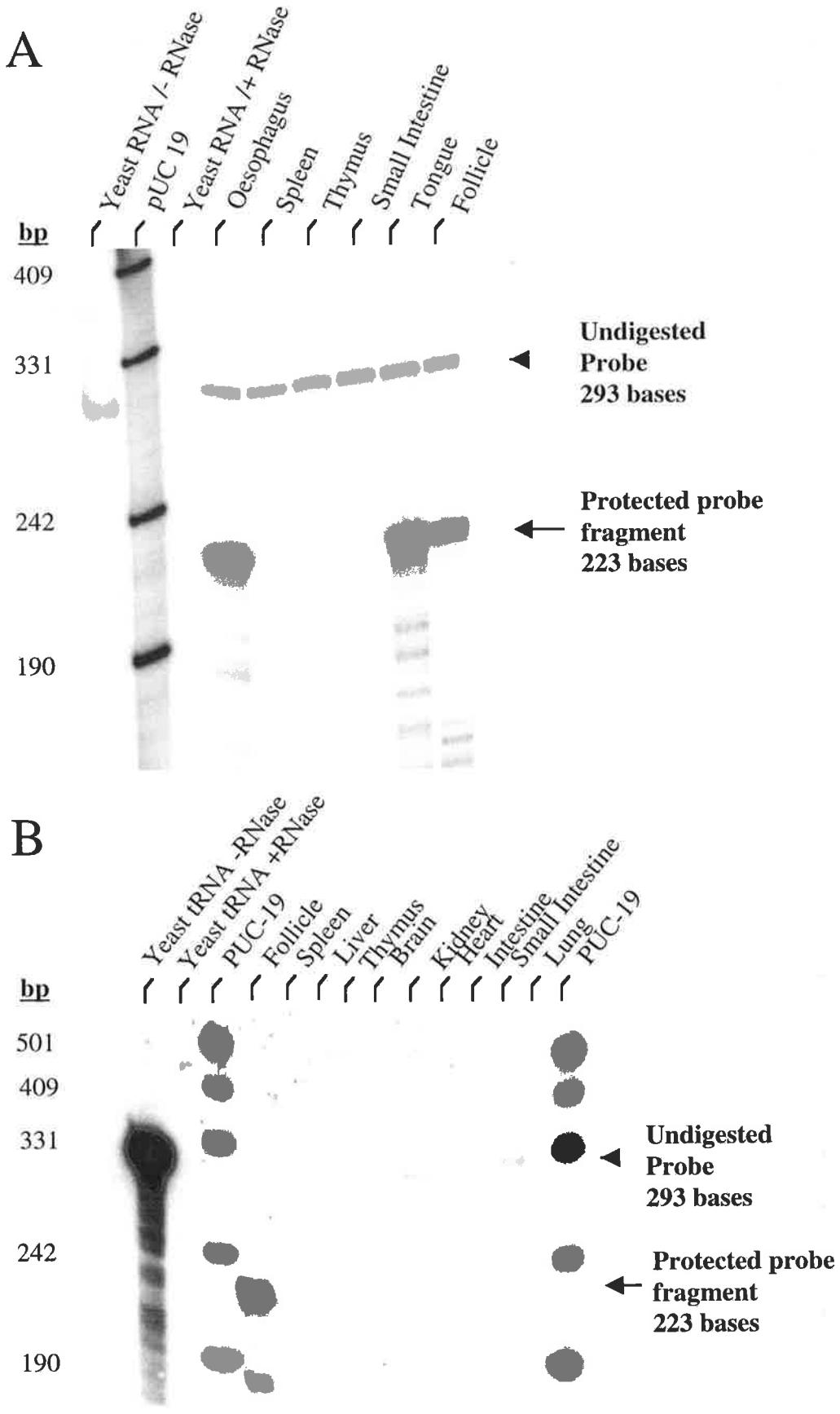
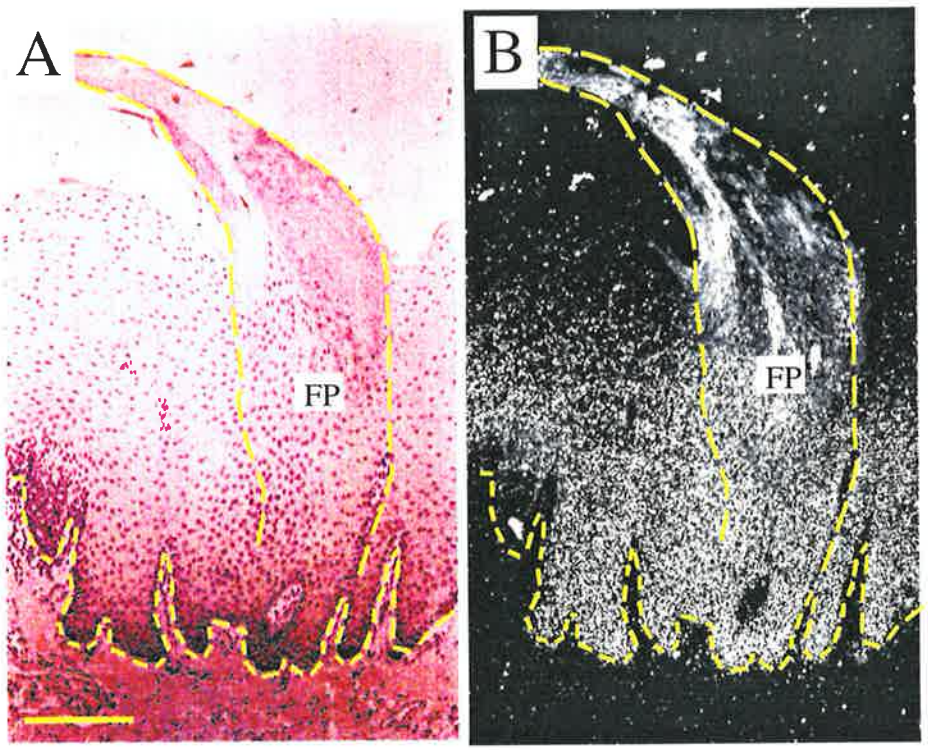

A
Yeast RNA /- RNase
pUC 19
Yeast RNA /+ RNase
Oesophagus
Spleen
Thymus
Small Intestine
Tongue
Follicle
bp
409
331
242
190
Undigested
Probe
293 bases
Protected probe
fragment
223 bases
B
Yeast tRNA -RNase
Yeast tRNA +RNase
PUC-19
Follicle
Spleen
Liver
Thymus
Brain
Kidney
Heart
Intestine
Small Intestine
Lung
PUC-19
bp
501
409
331
242
190
Undigested
Probe
293 bases
Protected probe
fragment
223 bases

Figure 5.20 : *In situ* Hybridisation Analysis of Ovine *Barx2* Expression in Sheep Tongue.

Brightfield (A) and darkfield (B) photographs of adult sheep tongue. Probe, wash conditions and exposure time were as for the skin sections in Fig. 5.12. *Barx2* is expressed in the basal and suprabasal layers of the sheep tongue. The basement membrane separating the epidermis and the dermis is indicated by the dashed line. FP = filiform papilla.
Bar = 95 μM


A
FP
B
FP

49, respectively. Most homeodomains contain an isoleucine residue at position 47 and a phenylalanine at position 49. These residues, including a glutamine at position 50 residing in the helix 3 region, have been shown to influence the recognition of DNA target sequences (Burglin et al., 1992). The *Barx2* homeodomain is conserved between human, mouse and sheep and shares a high degree of sequence identity (87%) with the mouse Barx1 homeodomain (Tissier-Seta et al., 1995). The ovine Barx2 homeodomain also shares a 65% sequence identity with Cnox3 (*C. viridissima*, Schummer et al., 1992), and 58% with BarH1 (*D. melanogaster*, Higashijima et al., 1992), BarH2 (*D. melanogaster,* Higashijima et al., 1992) and Om(1D) (Tanda and Corces, 1991).

The carboxyl end of ovine Barx2 contained a string of 17 amino acid residues located downstream of the homeodomain that may be unique for the Barx1 and Barx2 proteins and are possibly involved in protein-protein interactions due to their basic nature and close proximity to the homeodomain (Jones et al., 1997). The mouse *Barx2* protein contains a putative leucine zipper which may mediate protein - DNA interactions and a poly-alanine tract at the N terminal end that are not found in *Barx1* (Jones et al., 1997).

In Northern blot analyses, *Barx2* coding and 3' non-coding probes detected a single 2.0 kb mRNA transcript in wool follicle RNA. A *Barx2* mRNA transcript of 1.7 kb was previously reported in total RNA isolated from E12.5 mouse embryos (Jones et al., 1997). The predicted, smaller ovine *Barx2* RNA transcript of 1.83 kb was not detected and presumably migrates too closely to the 2.0 kb transcript to be resolved. However, two products of 1.1 kb and 0.93 kb both containing the *Barx2* homeobox were amplified by 3' RACE-PCR, and RNA protection confirmed that both polyadenylation signals were functional producing two equally abundant, different-sized RNA transcripts. The second ovine *Barx2* polyadenylation signal at position 1085 (Fig. 5.6) was not reported in mouse *Barx2* by Jones et al., (1997).

The *Barx2* gene is unique in the sheep, human and mouse genomes as single bands were detected using a *Barx2* coding probe in a Southern blot of human, mouse and sheep genomic DNA digested with BamHI, EcoR1 or HindIII. In what appears to be a contradiction, two bands at 8.5 kb and 14.5 kb were detected for the EcoR1 digested sheep genomic DNA shown in Figure 5.11. However, in a separate Southern blot analysis of

EcoR1 digested sheep genomic DNA (Fig. 5.4), only the 8.5 kb band was detected from 3 different sheep. This suggests that the 14.5 kb band is probably the result of a polymorphism.

A genomic clone which hybridizes to the *Barx2* 3' coding region was isolated and partially mapped for the first time as no other map from another species has been published to date. The location of the 3' end of *Barx2* was mapped but the location of the 5' end was not attempted because at that stage it was decided to focus on *Hoxc-13* characterisation that had been preceding in parallel up to this stage (see chapter 6). Hence the location of intron and exon boundaries were not found.

5.12.2 *Barx2* Function

Cell adhesion molecules (CAMs) regulate cell to cell contacts and are possibly involved in the regulation of cellular proliferation. To date, several CAMs have been reported to be expressed in the follicle (reviewed by Muller-Rover and Paus, 1998 and Stenn et al., 1994). E-cadherin and P-cadherin, for example, are expressed in either complementary or overlapping patterns in the follicle depending on developmental stage and location. Little is known about how cell adhesion molecules are regulated, although there is increasing evidence that one of the types of target genes regulated by *Hox* genes are those that encode cell adhesion molecules and several examples of this now exist (reviewed by Chuong, 1993; Lawrence et al., 1996 and Jones et al., 1997). One pertinent example involves *Hoxb-9*, possibly expressed in the wool follicle (Chapter 3), which enhances transcriptional activation of N-CAM *in vitro* (Jones et al., 1992).

Evidence to date suggests that Barx2 may regulate the transcription of cell adhesion molecules expressed during mouse embryogenesis. During mouse embryogenesis the expression pattern of Barx2 overlaps the expression pattern of the cell adhesion molecules L1, Ng-CAM and other neural CAMs expressed in the mantle layer (Jones et al., 1997; Krushel et al., 1993; Moscoso et al., 1995). Furthermore, Barx2 activates and represses L1 gene activity *in vitro* and may also serve a dual role as activator and repressor of the genes for the neural cell adhesion molecule Ng-CAM (Jones et al., 1997). Mouse Barx2 is predominantly expressed in tissues under-going epithelio-mesenchymal transformations

such as the lung buds and mesenchyme of the forelimb in embryonic mice, between E10.5 and E12.5 (Jones et al., 1997). It is also expressed in the ectodermal infoldings that surround the eye and maxillo-nasal groove and in the first branchial pouch. However, its expression pattern in skin at the later embryonic times of E13 to E19 when hair follicles are developing in mice has not been reported to date.

In embryonic sheep skin, *Barx2* is expressed in a temporally and spatially restricted manner in the epidermis, pelage and vibrissae follicles. The pattern of *Barx2* expression in embryonic sheep skin overlaps the expression pattern of the cell adhesion molecules E-cadherin reported in embryonic mouse skin (Hardy and Vielkind, 1996). The developmental stages of epidermis and follicles in mouse and sheep are very similar but occur at different embryonic time points. For example, placodes first develop at approximately E14 in mice and at E51 in sheep skin. Before the formation of the placode, ovine *Barx2*, murine E-Cad and murine P-Cad are expressed in the single-layered epidermis. Once the epidermal suprabasal layer and placodes begin to form, the expression pattern of ovine *Barx2* appears most similar to the expression pattern of murine E-cadherin whereas murine P-cadherin is expressed in different regions of the developing follicle and epidermis (Hardy and Vielkind., 1996). For example, both ovine *Barx2* and murine E-cadherin are expressed in the inner epithelial cells of the developing placode and the cells derived from the epithelial basal layer that line the placode. Furthermore, as the placode moves deeper into the dermis, both ovine *Barx2* and murine E-cadherin are absent from cells at the leading edge of the follicle placode and adjacent mesenchymal cells which form the pre-dermal papilla.

Some variation between *Barx2* and E-cadherin does exist during the development of the epidermis. *Barx2* is intermittently expressed throughout the epidermal suprabasal layer and is upregulated within regions of epithelial thickenings and epidermis located directly above developing hair and vibrissae follicles. In contrast, murine E-cadherin appears to be expressed evenly throughout the epidermal suprabasal layers and is moderately expressed in the epidermal basal layer at each stage of embryonic skin development (Hardy and Vielkind, 1996).

By analogy with Barx2 activity on L1 and Ng CAM (Jones et al., 1997) and that ovine *Barx2* is expressed in a temporally- and spatially- dependent manner during the formation of the epidermis and follicles which compares with the expression pattern of murine E-cadherin and murine P-cadherin, Barx2 may be involved in the regulation of these molecules at different developmental stages.

Cell adhesion is also an important part of the immune response and a number of families of cell adhesion molecules have been identified in leukocytes, including integrins, selectins and ICAMS (Butcher and Picker, 1996; Kansas et al., 1996; Lawrence et al., 1996). Ovine *Barx2* was expressed in cells scattered throughout the dermis and blood sinus of the vibrissae follicle that morphologically appeared to be immune cells because of their distinctive lobular nuclei. *Hox* genes have been reported to be expressed in hematopoietic cells lines including B and T cell lines, erythroid cell lines and myelomonocytic cell lines (reviewed in Lawrence et al., 1996). Given the association of *Barx2* with the L1 cell adhesion molecule (Jones et al., 1997) *Barx2* may also be involved in the regulation of CAMs expressed in these cell types.

Chapter 6

Isolation of Ovine *Hoxc-13*

Chapter 6 : Isolation of Ovine *Hoxc-13*

6.1 Introduction

Characterisation of all ten homeobox genes described in Chapter 3 was unachievable within the time frame of the PhD and it was decided to concentrate on two genes selected on the basis of two qualities i) sequence novelty and ii) location of expression. The selection and characterisation of one of them, the ovine *Barx2* gene was described in Chapter 5.

The second gene, ovine *Hoxc-13* was chosen because no information regarding its gene structure or functional role during the embryonic development of skin or any other organ was known. Only the sequence of the homeobox domain and its chromosomal location had previously been reported in human (Acampora, 1989). Secondly, its expression in the follicle bulb (described in Chapter 4) together with the knowledge that homeobox genes are master regulatory molecules indicated that it may be a candidate gene for the direct or indirect regulation of the wool keratin genes which are expressed in the upper follicle bulb and lower hair shaft (Powell and Rogers, 1997; Rogers et al., 1997).

To investigate whether Hoxc-13 might regulate the wool keratin genes at a transcriptional level, the entire *Hoxc-13* gene was isolated, translated and used in functional studies detailed in the next Chapter.

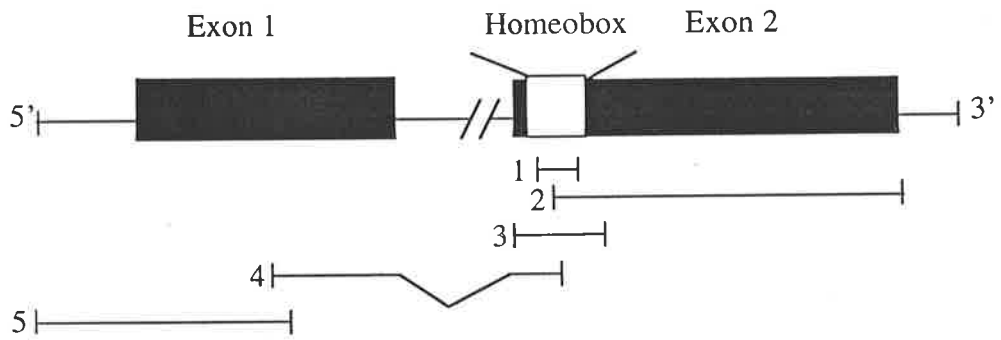
This chapter reports the results of several strategies used in isolating *Hoxc-13* in which the 3' cDNA end was first isolated using the 3' RACE technique (Frohman, 1988) and the remaining 5' sequence was then isolated using a combination of the 5'RACE technique (Frohman, 1994) and the screening of in-house sheep genomic and sheep cDNA libraries.

This chapter also details the complete gene structure of the *Hoxc-13* gene for the first time, and subsequent analysis with DNA and protein sequences of similar *Abd-B* homeobox genes.

Results

6.2 Diagram Summarising How the Sequence of *Hoxc-13* was Obtained

The following is a diagrammatic representation showing the steps involved in obtaining the full ovine *Hoxc-13* gene coding and flanking regions.



1. RT-PCR clone amplified from the homeobox region using degenerate primers (Chapter 3)
2. 3' RACE product
3. 5' RACE product
4. cDNA clone isolated from in-house sheep cDNA library
5. Genomic DNA clone (9A) isolated from an in-house sheep genomic DNA library

Diagram is not drawn to scale. Lines indicate the regions of sequence obtained from each clone and PCR products.

6.2.1 3' RACE-PCR Results

cDNA for the 3' RACE-PCR amplification of the ovine *Hoxc-13* 3' end was the same batch used to isolate the 3' end of *Barx2* as described in Section 5.2.1. Six PCRs were performed with the following template and primers:

1. cDNA with C24Hint (see Fig. 6.1 A) and MCS primers (PCR #1)
2. mRNA with C24Hint and MCS primers (PCR #2)
3. cDNA with the MCS primer (PCR #3)
4. cDNA with the C24Hint primer (PCR #4)
5. cDNA with no primers (PCR #5)
6. no cDNA with both primers (PCR #6).

PCR conditions and cycling parameters are given in Chapter two, Methods, Section 2.3.8.2. Multiple PCR products ranging from 200 bp to 1.4 kb were amplified from PCR #1 containing wool follicle cDNA and both primers (Fig. 6.1 B, lane 1). Apart from a faint smear at approximately 500 bp in lanes 2 and 3, no amplified products were seen in the control PCRs. It was not known if any of the PCR products were indicative of *Hoxc-13*.

6.2.2 Southern Blot Analysis and Cloning of 3' RACE Products

Southern blot analysis was used to determine if any of the 3' RACE-PCR amplified products contained the homeobox of *Hoxc-13*. A blot containing 3' RACE-PCR amplified products was prepared and probed with the ovine *Hoxc-13* homeobox probe (Fig. 6.2 B). The amplified 3' RACE-PCR products detected were at 1.56 kb and 0.36 kb. The length of the 3' end of ovine *Hoxc-13* could not be estimated from any orthologous *Hoxc*-13 genes due to the lack of published sequence data. Hence, the 1.56 kb 3' RACE-PCR product was isolated and cloned (clone C24B1; see Fig. 6.8 for clone reference and Appendix 1 B) because it had the highest probability of containing the full 3' cDNA end, being of almost equal size to the 3' end of *Hoxa-10*, another member of the *Abd-B* class.

Figure 6.1 : Ovine *Hoxc-13* 3' RACE.

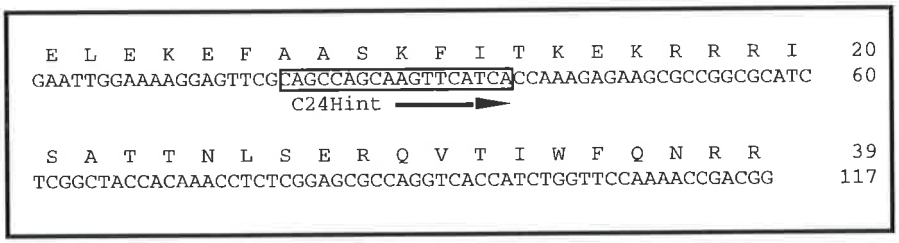
A. 5' Primer Used in the Amplification of the *Hoxc-13* 3' End

The partial sequence of the *Hoxc-13* homeobox, obtained from clone 24 (Fig. 3.3) is shown. The 5' primer used in the amplification of the *Hoxc-13* 3' end is boxed and is named C24Hint. The arrow shows the direction of DNA synthesis.

B. Amplification of the Ovine *Hoxc-13* 3' RACE-PCR

The 3' cDNA end of ovine *Hoxc-13* was amplified using the 3'RACE-PCR technique (Frohman, 1988). PCR products were size-separated on a 1% agarose gel and then stained with ethidium bromide and detected under UV light. PCRs were performed using wool follicle cDNA or mRNA as template in combination with the primers indicated above the gel (P1 = C24Hint, P2 = MCS-24). Note that there are multiple PCR products in lane 1 (the largest being 1.56 kb) and a smear at approximately 500 bp in lanes 2 and 3. The lower smeared band in lane 2 represents unused oligonucleotides. Molecular size of the SPP-1 and pUC19 markers are shown on the left of the gel (kb).

A


E L E K E F A A S K F I T K E K R R R I 20
GAATTGGAAAAGGAGTTCGCAGCCAGCAAGTTCATCACCAAAGAGAAGCGCCGGCGCATC 60
C24Hint
S A T T N L S E R Q V T I W F Q N R R 39
TCGGCTACCACAAACCTCTCGGAGCGCCAGGTCACCATCTGGTTCCAAAACCGACGG 117

B

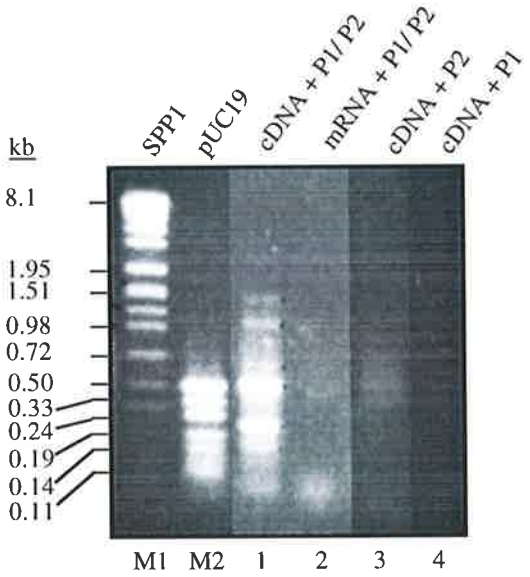

kb
SPP1
pUC19
cDNA + P1/P2
mRNA + P1/P2
cDNA + P2
cDNA + P1
8.1
1.95
1.51
0.98
0.72
0.50
0.33
0.24
0.19
0.14
0.11
M1 M2 1 2 3 4

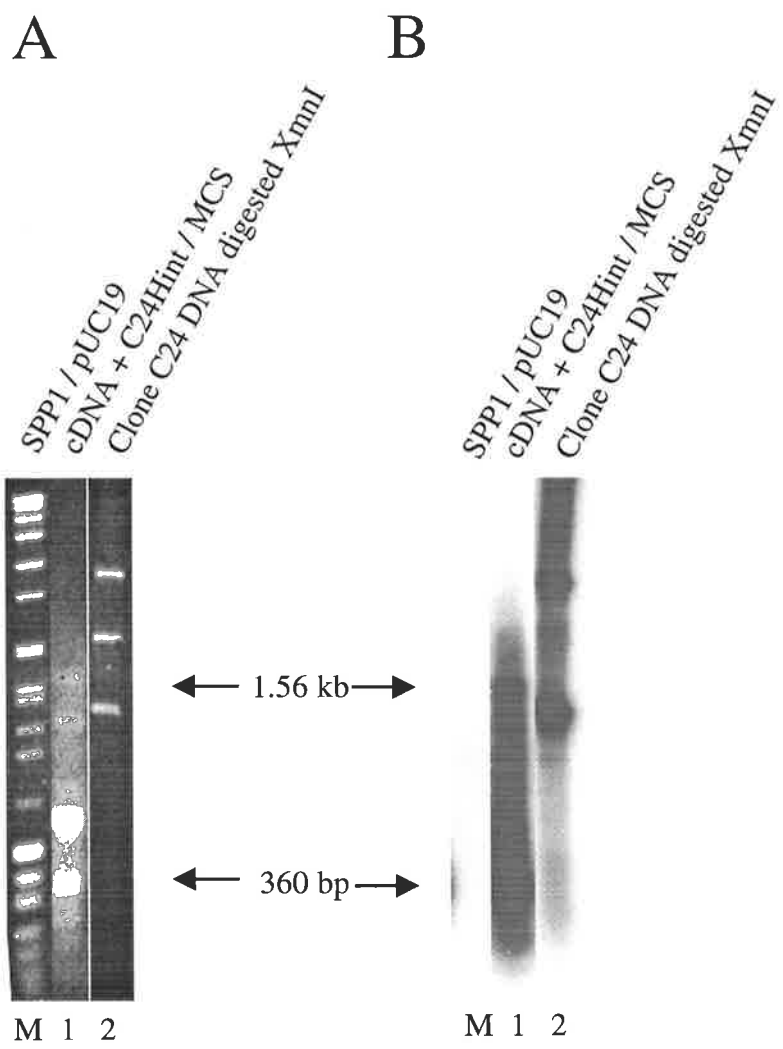
Figure 6.2 : Detection of the *Hoxc-13* 3' End by Southern Blot.

A. *Hoxc-13* 3' RACE-PCR

The 3' RACE-PCR amplified products were fractionated on a 1% agarose gel. Lane 1, 500 ng each of SPP-1 and pUC19 markers. Lane 2, *Hoxc-13* 3' RACE-PCR products amplified from cDNA with the C24Hint and 3'MCS primers. Lane 3, control DNA containing 1pg of XmnI partially digested plasmid DNA from clone C24 that contains the homeobox of *Hoxc-13*. The top band in lane 3 at 3.3 kb is linearised plasmid DNA from clone C24 that hybridised to the probe. Clone C24 contained two XmnI sites resulting in the second band at 2.1 kb (vector only) and the third band at 1.2 kb that contained part of the vector and the homeobox that hybridised to the probe.

B. Southern Blot Analysis of *Hoxc-13* 3' RACE-PCR Products

3' RACE-PCR products were analysed by Southern blot. 3' RACE-PCR products were transferred to Zetaprobe membrane and probed with an α-^{32}P-dCTP radio-labelled EcoR1-resected homeobox fragment from clone C24. The hybridised membrane was washed at a final stringency of 0.1 X SSC, 0.1% SDS at 65^{o}C for 1 hour prior to autoradiography (5 hours at room temperature). The largest PCR amplified product of 1.56 kb hybridised to the probe and is indicated by the arrow. A product at 360 bp also hybridised to the probe. Unfortunately, a shorter autoradiographic exposure of this result was not produced.


A
B
SPP1 / pUC19
cDNA + C24Hint / MCS
Clone C24 DNA digested XmnI
1.56 kb
360 bp
M 1 2
M 1 2

6.3 Confirming the Identity of the 3' cDNA End of Ovine *Hoxc-13* (Clone C24B1)

6.3.1 Confirmation of the 3' cDNA End by Sequencing

To verify that clone C24B1 contained the ovine *Hoxc-13* homeobox, approximately 200 bp of the 5' end was sequenced and compared to the ovine *Hoxc-13* homeobox sequence previously obtained from clone C24 (Chapter 3). Excluding the degenerate primer region, the homeobox regions were identical (data not shown).

Before completely sequencing the remaining 3' RACE-PCR product, steps were taken to prove that all of clone C24B1 was representative of *Hoxc-13* and not a hybrid between *Hoxc-13* and a non-specific PCR product. To prove this, fragments from different regions of clone C24B1 were isolated and used to probe a wool follicle Northern blot and a genomic DNA Southern blot.

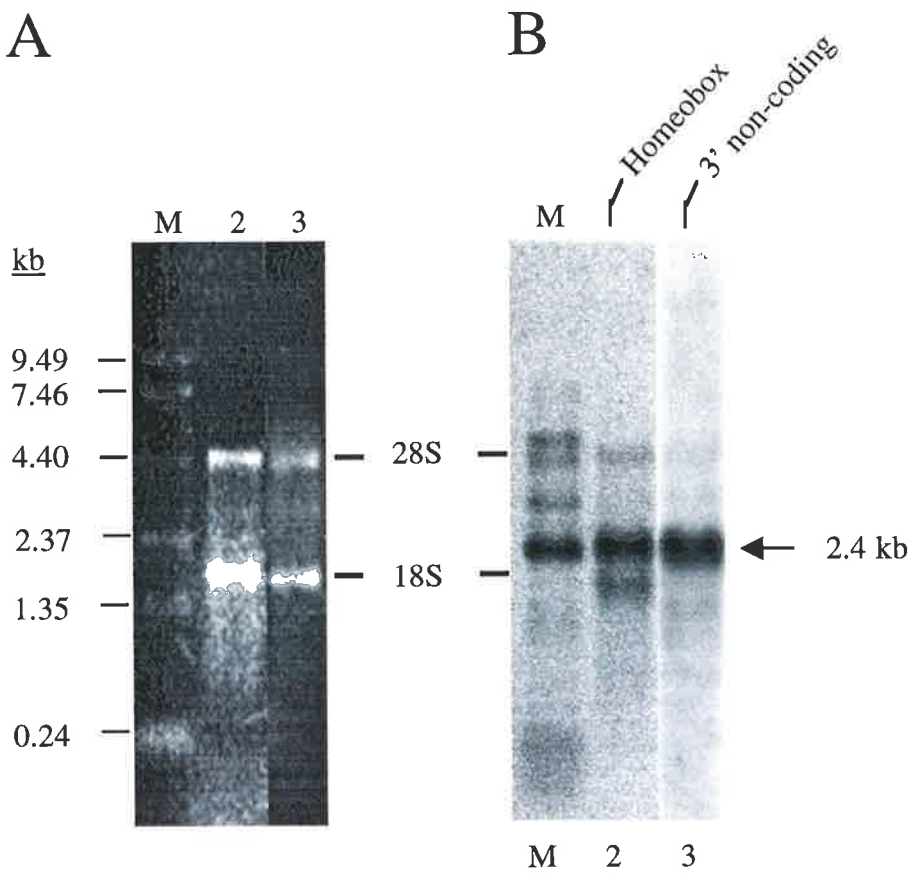
6.3.2 Confirmation of the *Hoxc-13* 3' cDNA End by Northern Blot

A Northern blot containing wool follicle poly A(+) RNA was probed with a fragment isolated from the 3' end of clone C24B1 to confirm that it detects the same size RNA transcript detected by the *Hoxc-13* homeobox probe (Fig. 6.3). To locate a fragment at the 3' end, a restriction enzyme map of clone C24B1 was obtained by digestion with single and multiple restriction endonucleases (data not shown). A recognition site for each of the restriction enzymes PvuII and SmaI was mapped to the 3' end of clone C24B1 downstream from the homeobox (Fig. 6.8). The 186 bp PvuII / SmaI 3' non-coding fragment was isolated and directionally cloned into PvuII / SmaI digested pBSc KS(+) and named C24B1-NC. The 186 bp fragment was resected, radioactively labelled with α–^{32}P-dCTP and used to probe a wool follicle poly A(+) RNA Northern blot. A 2.4 kb mRNA transcript was detected (Fig. 6.3, lane 3), the same size detected by the ovine *Hoxc-13* homeobox probe (clone C24; Fig. 6.3 lane 2). This suggested that the fragments from clones C24 and C24B1-NC were contained within the same mRNA molecule.

Figure 6.3 : Northern Blot Analysis of Sheep Wool Follicle Poly A(+) RNA with *Hoxc-13* Probes.

A. Approximately 5 µg of poly A(+) RNA was prepared from sheep wool follicles, size-fractionated in a 1 % agarose formaldehyde gel and then transferred to Zetaprobe GT membrane. The size (kb) of marker bands in lane M are shown to the left of the figure. The location of the 28S and 18S RNA bands are indicated. M = BRL RNA ladder.

B. Strip filters from the blot were separately probed with radio-labelled fragments prepared from the homeobox (lane 2) and 3' non-coding (lane 3) regions of *Hoxc-13* (clones C24 and C24B1-NC respectively, see Appendix 2 for clone reference). Filters were washed at a final stringency of 2 x SSC / 0.1% SDS / 65°C for 30 minutes. The size (kb) of the hybridisation product indicated to the right of the figure was determined by comparison to a BRL RNA ladder in lane M. Signals were detected by an over-night exposure of the probed membrane to a phosphor-imaging screen. The position of the 28S rRNA and 18S rRNA bands which have hybridised non-specifically to the homeobox probe are indicated to the left of the figure. The position of the major hybridisation at 2.4 kb is marked by an arrow at the right of the figure.


A
B
M
2
3
kb
9.49
7.46
4.40
2.37
1.35
0.24
28S
18S
Homeobox
3' non-coding
2.4 kb

6.3.3 Confirmation of the 3' cDNA End by Southern Blot

A sheep genomic blot was probed separately with radio-labelled fragments synthesised from the homeodomain (clone C24) and 3' non-coding (clone C24B1-NC) regions of ovine *Hoxc-13*. A single 6.4 kb band appeared for each probing (Fig. 6.4 B) suggesting that each region of the 1.56 kb RT-PCR product was contained within one 6.4 kb genomic EcoRI DNA fragment.

6.4 Isolation of the *Hoxc-13* 5' End

6.4.1 5' New R.A.C.E

The length of the *Hoxc-13* 5' cDNA end yet to be obtained was calculated to be between 600-850 nucleotides by subtracting the length of the 3' RACE product (1,482 nucleotides; see later) from the full length mRNA transcript (2,400 nucleotides) and assuming a poly A tail of between 50-300 bases.

Isolation of the *Hoxc-13* 5' cDNA end was attempted using the 5' New RACE technique as described by Frohman (1994). A fragment at approximately 800 bp, not large enough to be the full 5' end of *Hoxc13* was amplified in the first PCR, containing the primer pair, Puc#16 and P4 (Fig. 6.5 B, lane 2). Given the location of both primers, the 800 bp product was 400 bp too short. It was assumed that the full length 5' end may have been amplified but was not detectable by agarose gel electrophoresis and ethidium bromide staining.

One twentieth of the first PCR reaction was then used as a template for a second nested PCR which amplified three major products at approximately 1.3 kb, 1.0 kb and 250 bp (B1, B2 and B3 respectively; Fig. 6.5 B, lane 3). Some faint products at 0.5 kb, 1.15 kb and a smear ranging from 1.4 kb to 3 kb were also detected.

Amplified products from the first and second nested PCR were analysed by Southern blot hybridisation using a probe that encompassed the *Hoxc-13* homeobox (resected from clone C24, Appendix 2A). The 250 bp band hybridised very strongly with

Figure 6.4 : Southern Blot of EcoRI - Digested Sheep Genomic DNA with *Hoxc-13* Probes.

A. Sheep genomic DNA was digested with EcoRI. 10 µg samples were fractionated in a 1 % agarose (1 X TAE) gel, stained with Ethidium Bromide, photographed and then transferred to Zetaprobe GT membrane.

B. Strip filters from the Southern blot in (A) were separately probed with radio-labelled fragments prepared from the homeobox (lane 2) and 3' non-coding (lane 3) regions of *Hoxc-13* (ie. resected inserts from clones C24 and C24B1-NC respectively; see Appendix 2 for probe reference). Filters were washed at a final stringency of 2 x SSC / 0.1% SDS / 65^{o}C for 30 minutes. Signals were detected by an over-night exposure of the probed membrane to a phosphor-imaging screen. The size (kb) of the hybridisation product (detected using phosphor-image technology) indicated to the right of the figures was determined by size comparison to SPP-1 and lambda Hi-Lo markers in lane M.

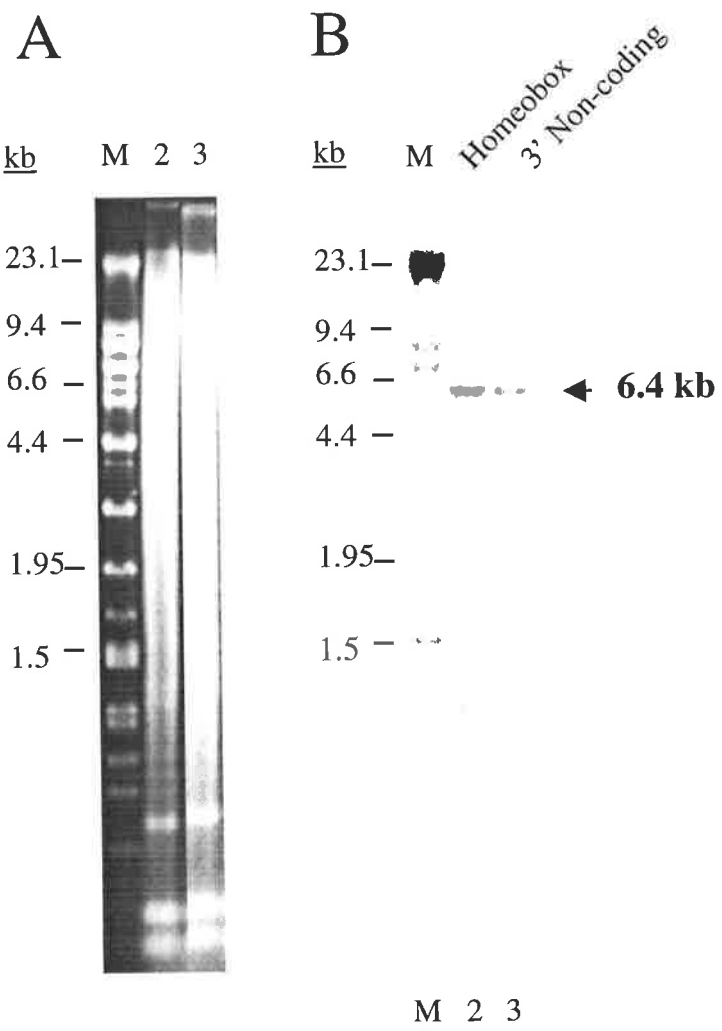

A
B
kb
M 2 3
23.1
9.4
6.6
4.4
1.95
1.5
kb
M
Homeobox
3' Non-coding
23.1
9.4
6.6
4.4
1.95
1.5
6.4 kb
M 2 3

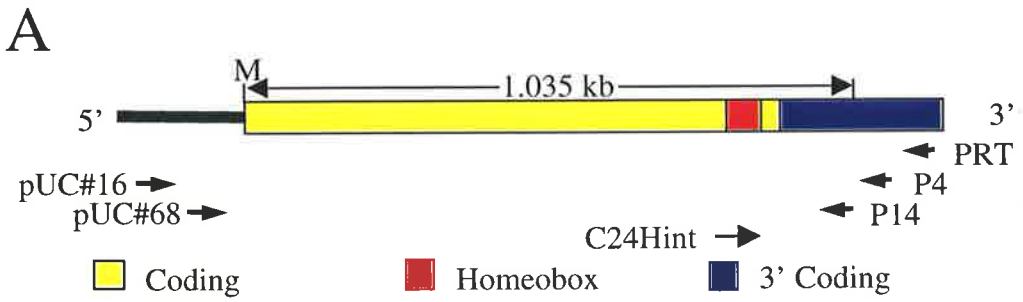
Figure 6.5 : Amplification of the *Hoxc-13* 5' end by 5' New RACE.

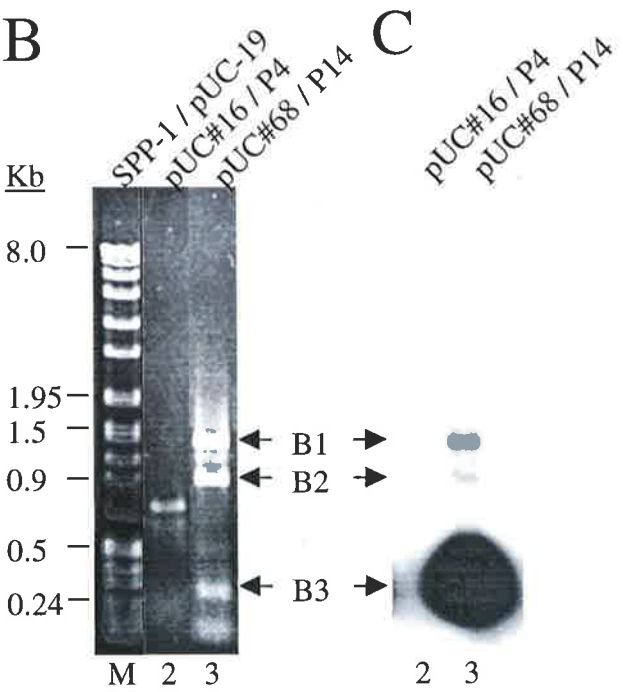
A. Schematic of the 5' New RACE procedure. Arrows indicate the location and direction of the 5' New RACE oligonucleotides that are named. The black bar at the 5' end depicts the ligated RNA oligonucleotide.

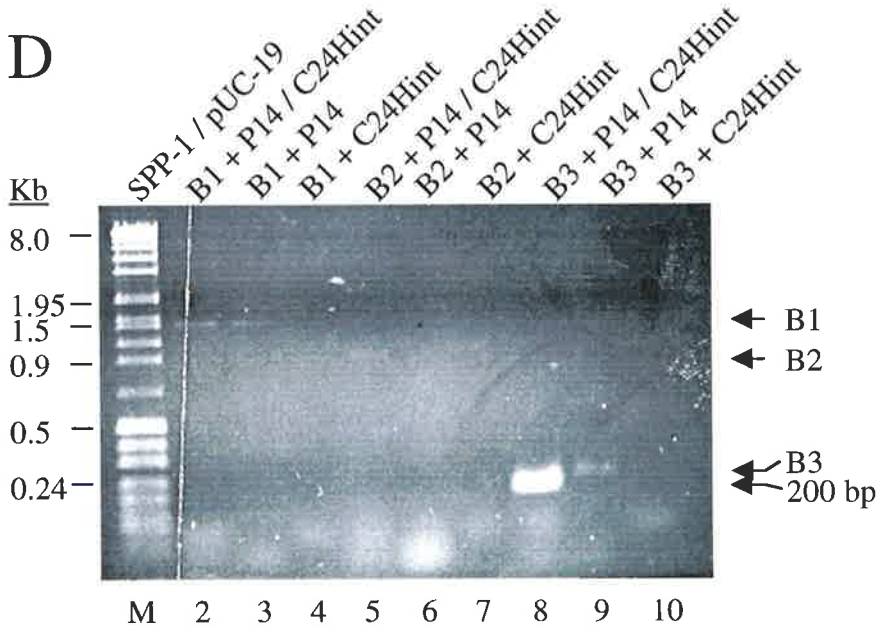
B. 5' New RACE products of the first (lane 2) and nested (lane 3) PCR's were size separated through a 1% agarose gel. Synthesis of cDNA and PCR cycling parameters are noted in Chapter 2. The size of various markers are shown on the left of the figure.

C. Southern blot analysis of 5' New RACE products depicted in **B**. PCR products were transferred to Zetaprobe (GT) membrane and probed with an a-32P-dCTP radio-labelled EcoR1-resected homeobox fragment from clone C24. The hybridised membrane was washed at a final stringency of 0.1 X SSC, 0.1% SDS at 65°C for 1 hour prior to autoradiography (2 hours at room temperature). PCR amplified products that hybridised to the probe are indicated by the arrows. A 250 bp product (B3) yielded a strong hybridisation signal. Products at 1.3 kb (B1) and 1.0 kb (B2) hybridised weakly.

D. Internal PCR check. Photograph of amplified PCR products size separated through a 1% agarose gel. The primers P14 and C24Hint (Lanes 2, 5 and 8) amplify a 200 bp fragment that is detected only in lane 8 suggesting that the templates, B1 and B2 are not derived from *Hoxc-13*. DNA template. Primers used for each reaction are depicted above each lane. PCR cycling parameters were as follows: 98°C (5'), 72°C (3') 20 cycles of [94°C (30"), 45°C (30" + 0.2°C / cycle), 72°C (3')] followed by a final extension step of 72°C (15'). A final $MgCl_2$ concentration of 1.5mM was used in each PCR. B1 = 1.3 kb product, B2 = 1.0 kb product and B3 = 250 bp product. The size of various markers are shown on the left of the figure.


A
M
1.035 kb
5'
3'
PRT
P4
P14
pUC#16
pUC#68
C24Hint
Coding
Homeobox
3' Coding


B
C
SPP-1 / pUC-19
pUC#16 / P4
pUC#68 / P14
Kb
8.0
1.95
1.5
0.9
0.5
0.24
B1
B2
B3
M 2 3
2 3


D
SPP-1 / pUC-19
B1 + P14 / C24Hint
B1 + P14
B1 + C24Hint
B2 + P14 / C24Hint
B2 + P14
B2 + C24Hint
B3 + P14 / C24Hint
B3 + P14
B3 + C24Hint
Kb
8.0
1.95
1.5
0.9
0.5
0.24
B1
B2
B3
200 bp
M 2 3 4 5 6 7 8 9 10

the homeobox probe and the 1.3 and 1.0 kb bands hybridised very weakly. The fainter products did not hybridize with the probe at all (Fig. 6.5 C, lane 3).

All three PCR products detected in the Southern blot analysis were isolated, purified and tested further to ensure that they were representative of *Hoxc-13*. Each of the three PCR products were used as templates in another PCR designed to amplify an internal 200 bp fragment using the primers C24Hint and P14 (Fig. 6.5 A and D). The 200 bp internal fragment amplified only from the 250 bp nested PCR product, B3 (see Fig. 6.5 D, lane 8), but not from the 1.3 kb (B1) or 1.0 kb (B2) templates suggesting that they were not representative of *Hoxc-13*.

Each band was then sequenced directly by cycle sequencing using the P14 gene-specific primer as the sequencing oligonucleotide. The sequencing data revealed that the 250 bp product contained *Hoxc-13* sequence but the 1.3 kb and 1.0 kb PCR products did not (data not shown).

Due to the limited progress with the 5' New RACE-PCR technique, a decision was made to isolate the *Hoxc-13* gene by another strategy.

6.4.2 Detection and Analysis of Genomic *Hoxc-13* Clones

Genomic clones containing all or part of the *Hoxc-13* gene were isolated from a λGT11 sheep genomic DNA library. Approximately 5 genome equivalents of clones were plated onto TB agar plates and screened with the 186 bp PvuII / SmaI 3' non-coding probe (Clone C24B1-NC; see Appendix 2 and Fig. 6.8 for location of probe). Four positive clones were detected which were fully purified by four rounds of screening. Two of the clones, 5A and 9A contained inserts of 15 kb and 12 kb respectively. Both clones were cleaved with a series of single and double restriction enzyme digests (Fig. 6.6, A and C) and the restriction enzyme sites mapped (see Fig. 6.7 C). The position of the homeobox in both clones was mapped by Southern blot analysis (see Fig. 6.6, B and D). A detailed restriction map of both clones and the location of the homeobox is shown in figure 6.7 C. Clones were in the same orientation in λ, with 5' ends closest to the left arm of the

Figure 6.6 : Southern Blot Analysis of the *Hoxc-13* Lambda Clones 9A and 5A.

A. DNA from genomic clone 5A was digested with single and multiple restriction endonucleases as indicated above each lane. Cleaved fragments were size-separated through a 1% agarose gel by gel electrophoresis and the gel stained with ethidium bromide. The size of the fragments were determined by comparison with SPP-1 and Lambda Hi-Lo markers (lane M). Note that the DNA is partially digested by the KpnI enzyme (lanes 8 and 9). Yellow rectangles indicate bands that hybridised to the probe.

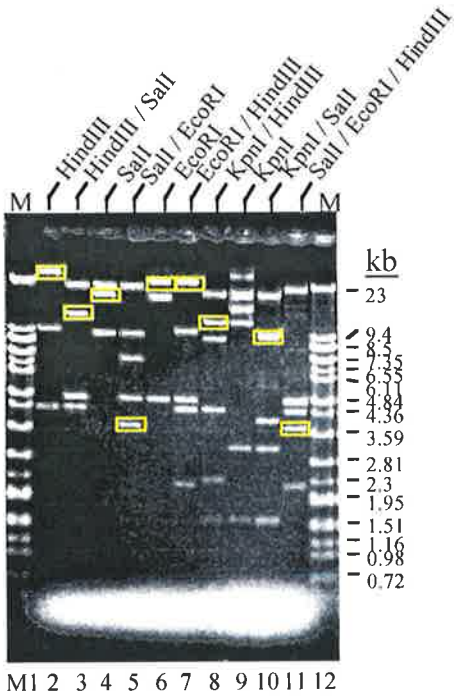
B. Southern blot analysis of clone 5A. Fragments from **A.** were transferred to Zetaprobe membrane and probed separately with a homeobox probe (clone C24; see Fig. 6.5 for clone reference). Bands indicate fragments of DNA that hybridised to both probes. Multiple bands can be detected in lanes 6, 8 and 9 due to the partially digested DNA. Hybridisations were treated with a final wash in 0.1 X SSC, 0.1% SDS at 65^{o}C for 30 minutes and exposed to an autoradiograph for 1 hour. Size markers are shown (in kb).

C. DNA from genomic clone 9A was digested with single and multiple restriction endonucleases as indicated above each lane. Cleaved fragments were size-separated through a 1% agarose gel by gel electrophoresis and the gel stained with ethidium bromide. The size of the fragments were determined by comparison with SPP-1 and Lambda Hi-Lo markers (lane M). Note that the DNA is partially digested by the KpnI enzyme (lane 9). White rectangles indicate bands that hybridised to the probe.

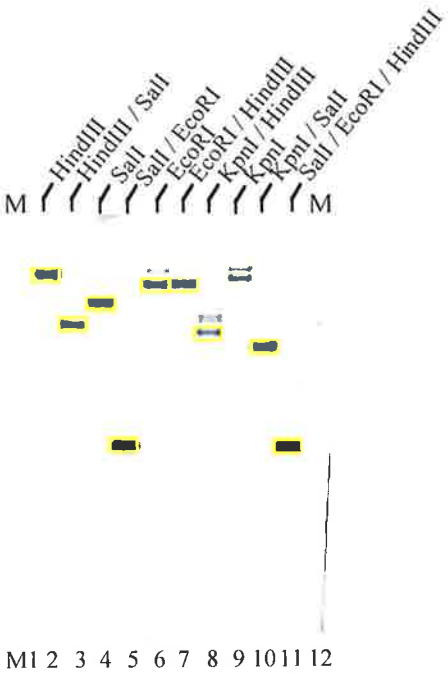
D. Southern blot analysis of clone 9A. Fragments from **C.** were transferred to Zetaprobe membrane and probed with a homeobox probe (clone24). Bands indicate fragments of DNA that hybridised to the probe. Hybridisations were treated with a final wash in 0.1 X SSC, 0.1% SDS at 65^{o}C for 30 minutes and exposed to an autoradiograph for 1 hour. Size markers are shown (in kb). Multiple bands can be detected in lane 9 due to the partially digested DNA.

Clone 5A

A

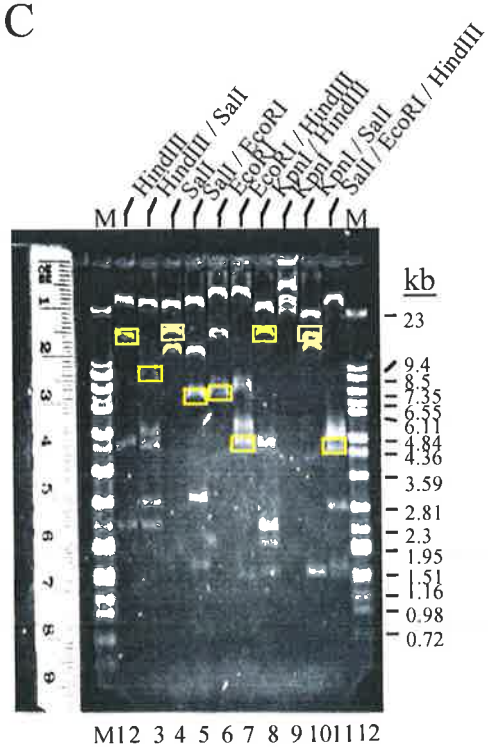


B



Clone 9A

C



D

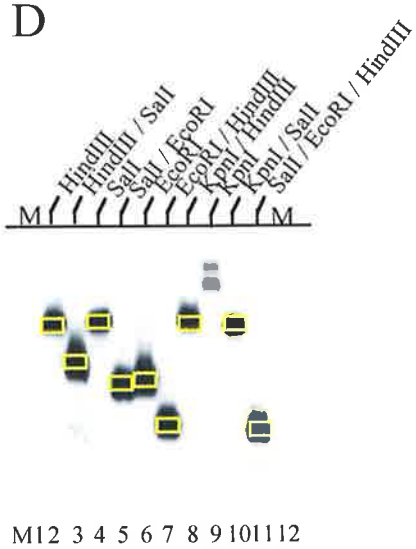


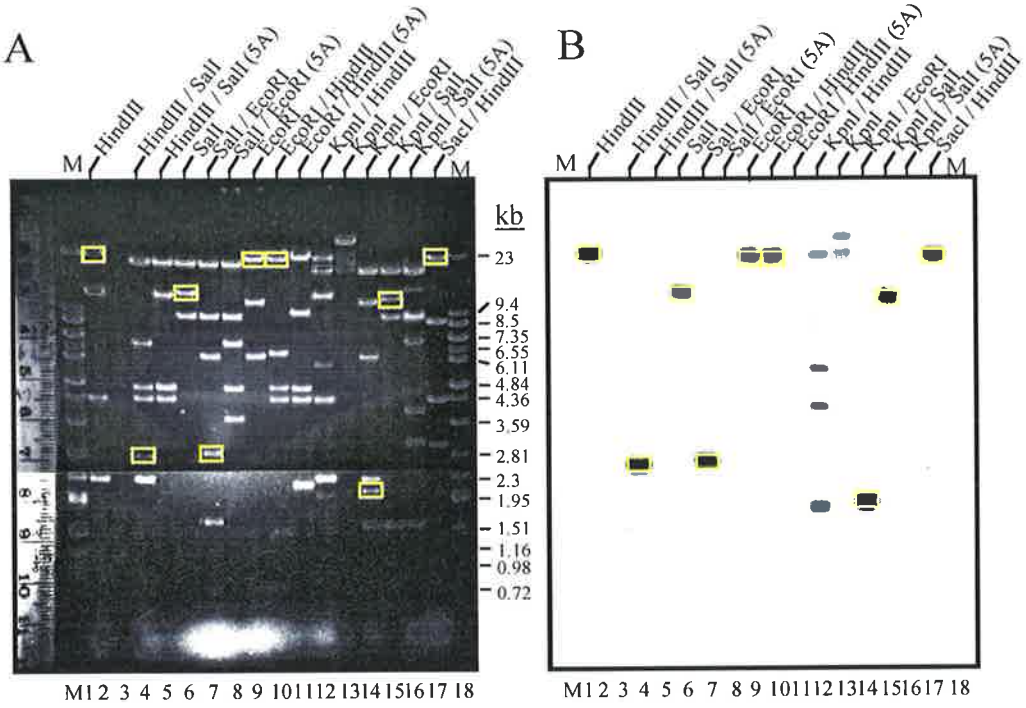
Figure 6.7 : Southern Blot Analysis of the *Hoxc-13* Lambda Clone 9A and Restriction Map of Clones 5A and 9A.

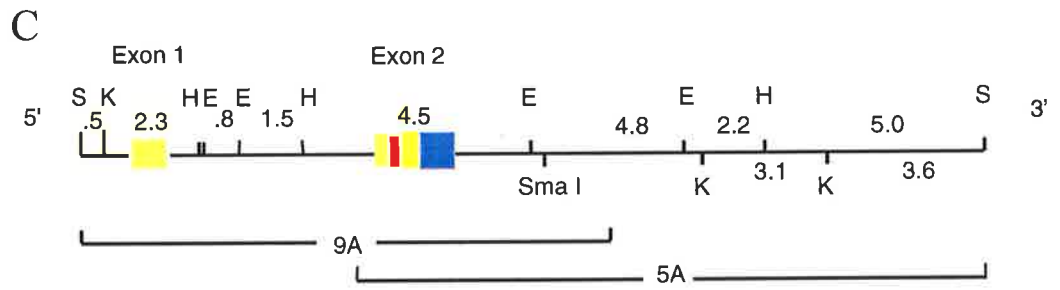
A. DNA from genomic clone 9A was digested with single and multiple restriction endonucleases as indicated above each lane. Cleaved fragments were size-separated through a 1% agarose gel by gel electrophoresis and the gel stained with ethidium bromide. The size of the fragments were determined by comparison with SPP-1 and Lambda Hi-Lo markers (lane M). Lanes 5, 8, 11 and 16 contain DNA from clone 5A so that banding patterns between clone 9A and clone 5A could be compared. Note that the DNA is partially digested by the KpnI enzyme in lanes 12 and 13. Yellow rectangles indicate digested DNA that hybridised to the probe excluding the KpnI digests of lanes 12 and 13.

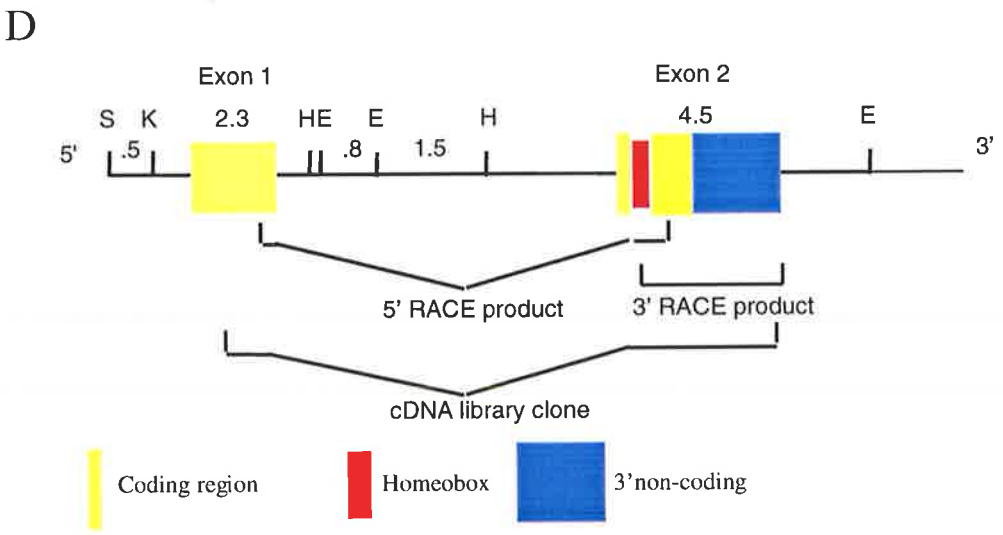
B. Southern blot analysis of clones 9A and 5A. Fragments from **A.** were transferred to Zetaprobe GT membrane and probed with a 450 bp PCR product amplified from a partial *Hoxc-13* cDNA clone that contained part of exon 1 (C13cDNA-450: see Section 6.4.3 for probe reference). Yellow rectangles indicate the comparative fragment of DNA in panel A that hybridised with the probe. Note that clone 5A (lanes 5, 8, 11 and16) does not hybridize to the coding probe and therefore does not contain the first exon. Hybridisations were treated with a final wash in 0.1 X SSC, 0.1% SDS at 65^{o}C for 30 minutes and exposed to an autoradiograph for 1 hour. Size markers are shown (in kb).

C. Restriction map of the genomic clones 5A and 9A. The location of the SalI, EcoRI and HindIII, SmaI and KpnI sites are denoted. Exon 1 and exon 2 are indicated by the colored rectangles. The homeobox is indicated by the small red rectangle. The homeobox, exon 2 and the 3' non-coding region are located within the 4.5 kb HindIII / EcoRI fragment. The length of the overlapping clones 9A and 5A are represented by the brackets underneath the map. E = EcoRI, H = HindIII, K = KpnI and S = SalI.

D. Restriction map of the genomic clone 9A. The length and relationship of the 3' RACE, 5' RACE and cDNA library clones are indicated by the brackets underneath the plasmid map of clone 9A.


A
B
M
HindIII
HindIII / SalI
HindIII / SalI (5A)
SalI
SalI / EcoRI
SalI / EcoRI (5A)
EcoRI
EcoRI / HindIII
EcoRI / HindIII (5A)
KpnI / HindIII
KpnI
KpnI / EcoRI
KpnI / SalI
KpnI / SalI (5A)
SacI / HindIII
kb
23
9.4
8.5
7.35
6.55
6.11
4.84
4.36
3.59
2.81
2.3
1.95
1.51
1.16
0.98
0.72
M1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18


C
Exon 1
Exon 2
5'
S K
.5
2.3
HE E
.8
1.5
H
4.5
E
4.8
E
2.2
H
3.1
K
5.0
S
3'
Sma I
3.6
9A
5A


D
Exon 1
Exon 2
S K
.5
2.3
HE E
.8
1.5
H
4.5
E
5'
3'
5' RACE product
3' RACE product
cDNA library clone
Coding region
Homeobox
3'non-coding

λ vector. Clone 5A contained 800bp of DNA upstream of the homeobox to the SalI site of the λgt11 left arm (Fig. 6.7 C). Given that the calculated length of the 5' end of *Hoxc-13* was approximately 1.4 kb, clone 5A did not contain the complete 5' end. Clone 9A contained 7.7 kb of DNA upstream of the homeobox and was used to obtain the remaining sequence of *Hoxc-13*.

6.4.3 Screening of a Wool Follicle cDNA Library for *Hoxc-13*

At the same time that the sheep genomic library was screened, an amplified sheep follicle cDNA λGT11 library (Clontech) was screened with a radio-labelled *Hoxc-13* 3' non-coding fragment (clone C24B1-NC). Seven plaques hybridised to the probe after a final wash stringency of 0.1 X SSC at 65^{o}C for 30 minutes and were purified by four rounds of screening.

To determine the insert size of each purified cDNA clone, the 5' end was amplified by PCR using the primers, C13In3 (nucleotides 713-734; Fig. 6.10) and forward or reverse primers which bind to the λGT11 vector arms. A 450 bp PCR product was amplified from each cDNA clone and named C13cDNA-450 (data not shown). C13cDNA-450 was not cloned but was cycle sequenced in both directions using both the C13In3 primer and λGT11 reverse primer.

The cDNA clones contained a partial cDNA of approximately 1.95 kb. The cDNA clones were 450 bp short of containing the full *Hoxc-13* cDNA as determined by subtracting the insert length of the cDNA clones from the mRNA transcript of 2.4 kb detected by Northern blot.

6.5 Location of the Intron / Exon Splice Site

The position of the intron / exon splice site was located by direct comparison of the DNA sequence from the cDNA clone and the genomic clone 9A (data not shown). DNA sequence differed 43 nucleotides upstream from the homeobox suggesting an intron / exon boundary. Furthermore, the intron / exon consensus 3' splice sequence of 'CCAG' was located at this point. *Abd-B* genes usually contain only one intron located a short distance (40-50 nucleotides) upstream from the homeobox.

6.6 Location of the First Exon in the Genomic Clone 9A

To determine if either of the genomic clones 5A or 9A contained the first exon, a Southern blot prepared from the clones DNA, cleaved with multiple restriction enzymes (Fig. 6.7 A) was probed with the C13cDNA-450 PCR product which hybridizes to the first exon (Fig. 6.7 B). The probe hybridised to clone 9A but not to clone 5A (Fig. 6.7 B; lanes 5, 8, 11 and 16) which suggested that clone 5A did not contain the first exon. The first exon was mapped to the 2.8 kb SalI / HindIII fragment located at the 5' end of the lambda genomic clone 9A (Fig. 6.7 C). A map showing the relationship between the λ genomic and cDNA clones and the 5' and 3' RACE products is given in Figure 6.7 D.

6.7 Nucleotide Sequencing of *Hoxc-13*

To obtain the full sequence of *Hoxc-13* a combination of subcloning, progressive deletion (Erase-a-base kit) and oligonucleotide primed sequencing were employed (Fig. 6.8 and Fig. 6.9). The 3' RACE clone, C24B1, was used as the template for all cycle sequencing reactions which used oligonucleotides designed to the 3' end of *Hoxc-13* (Fig. 6.8). To sequence the 5' end of *Hoxc-13*, the 2.8kb SalI / HindIII fragment from genomic clone 9A was sub-cloned into a SalI / HindIII- digested pBSc (KS) vector and used as the template for sequencing the 5' end (Fig. 6.9).

Figure 6.8 : Sequencing Strategy for the 3' cDNA End of *Hoxc-13*.

Schematic showing the restriction map of clone C24B1 and the sequencing strategy. Restriction sites used for subcloning are shown. Two restriction sites for each of the restriction enzymes SacI and EcoRV were located in clone C24B1, one in the polycloning site and one in the 3' end of *Hoxc-13*. These enzymes were used to synthesise two subclones. A 790 bp EcoRV fragment was cloned into the SmaI site of pGEM - 7Zf (+) and the remaining linearised vector (3.0 kb) containing 690 bp of the *Hoxc-13* 3' end was religated and transformed. The sub-clones were named C24B1-790 and C24B1-690 respectively. In the same way two more subclones, containing an 820 bp and a 630 bp insert, were made using SacI. The clones were named C24B1-820 and C24B1-630. Smaller sized subclones of approximately 300 bp in length were prepared by progressive deletion of DNA from clone C24B1 using the Erase-a-base kit protocol (Promega). Deletion fragments were cloned into pGEM - 7Zf (+) and the insert size verified by restriction digestion of plasmid mini-prep DNA. Arrows indicate the direction and extent of the sequencing reactions. The lined rectangle represents the homeobox, the dotted rectangle represents the 3' coding region and the open rectangle represents the 3' non-coding region. AAAAA depicts the poly A tail. T7, T3 and RSP (reverse sequencing primer) indicates the beginning of sequence stretches obtained from the EcoRV or SacI ends of the subclones. P1 - P7 and P56 are the sequencing primers designed using Oligo 4 software as the *Hoxc-13* DNA sequence was accumulated.

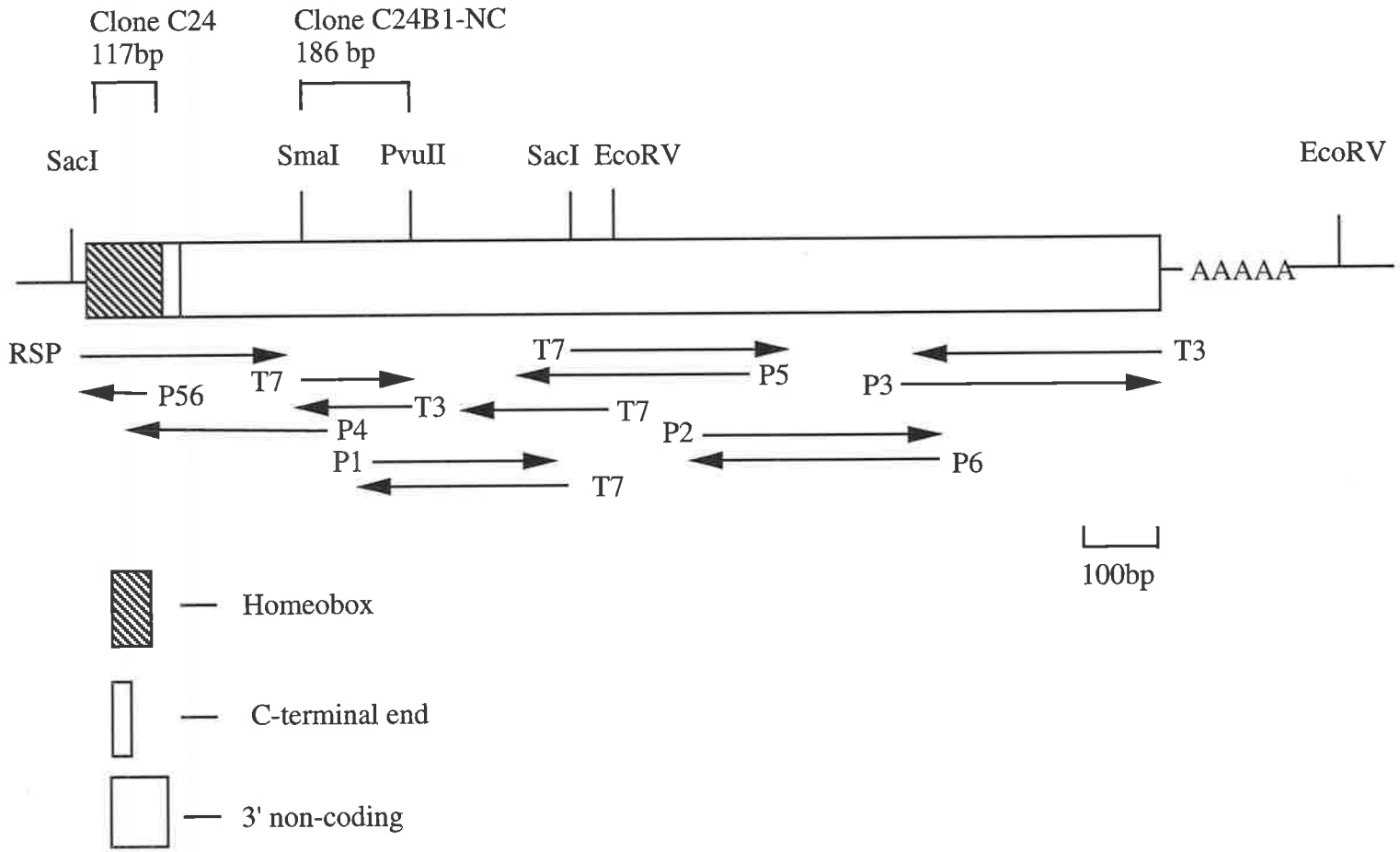
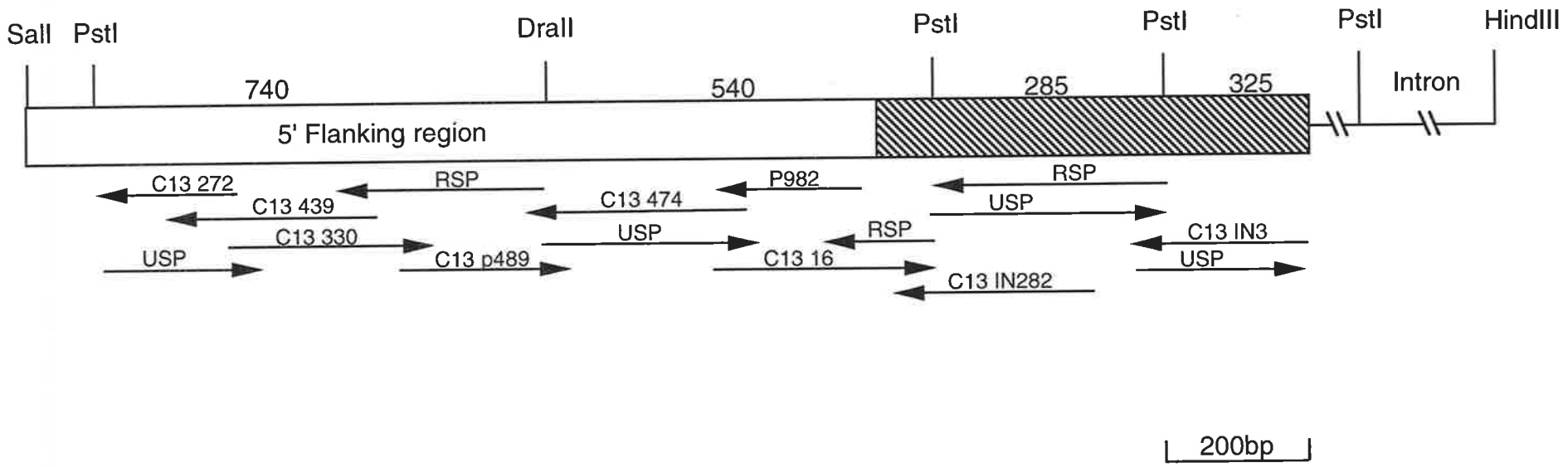

Clone C24
117bp
Clone C24B1-NC
186 bp
SacI
SmaI
PvuII
SacI
EcoRV
EcoRV
AAAAA
RSP
T7
P56
T7
T3
P4
T7
P5
T7
P1
T7
P3
T3
P2
P6
100bp
Homeobox
C-terminal end
3' non-coding

Figure 6.9 : Sequencing Strategy for the 5' cDNA End of *Hoxc-13*.

Schematic showing the restriction map of the 2.8 kb SalI / HindIII subclone of the λ sheep genomic clone 9A and the sequencing strategy. Restriction sites used for further subcloning and sequencing are shown. The following fragments were subcloned into pBSc (KS) by sticky end ligation; 740 bp PstI / DraII, 540 bp DraII / PstI, 285 bp PstI / PstI and 325 bp PstI / PstI. Each subclone was sequenced in both directions to ensure that sequence information was unambiguous. Arrows indicate the direction and extent of the sequencing reactions. The shaded box represents the 5' coding region (611 bp) and the open box represents the 5' flanking region (1189 bp).


SalI
PstI
DraII
PstI
PstI
PstI
HindIII
740
540
285
325
Intron
5' Flanking region
C13 272
RSP
P982
RSP
C13 439
C13 474
USP
C13 330
USP
RSP
C13 IN3
USP
C13 p489
C13 16
USP
C13 IN282
200bp

The nucleotide sequence and the deduced amino acid sequence of *Hoxc-13* are shown in figure 6.10. *Hoxc-13* encodes a predicted 330 amino acid protein with the homeodomain located close to the C terminal end.

Sequence data revealed that the 2.8 kb SalI / HindIII clone contained the entire 5' end of *Hoxc-13* including the initiation codon (ATG), and 1.2 kb of 5' flanking sequence. The 3' non-coding region is 1,321 base in length of which the last 16 are adenine residues suggesting the beginning of the poly A tail. Sequence obtained from the 3' RACE product revealed that *Hoxc-13* did not contain a consensus polyadenylation signal. A similar motif, AAGAAA was present 19 bases upstream from the beginning of the poly A tail.

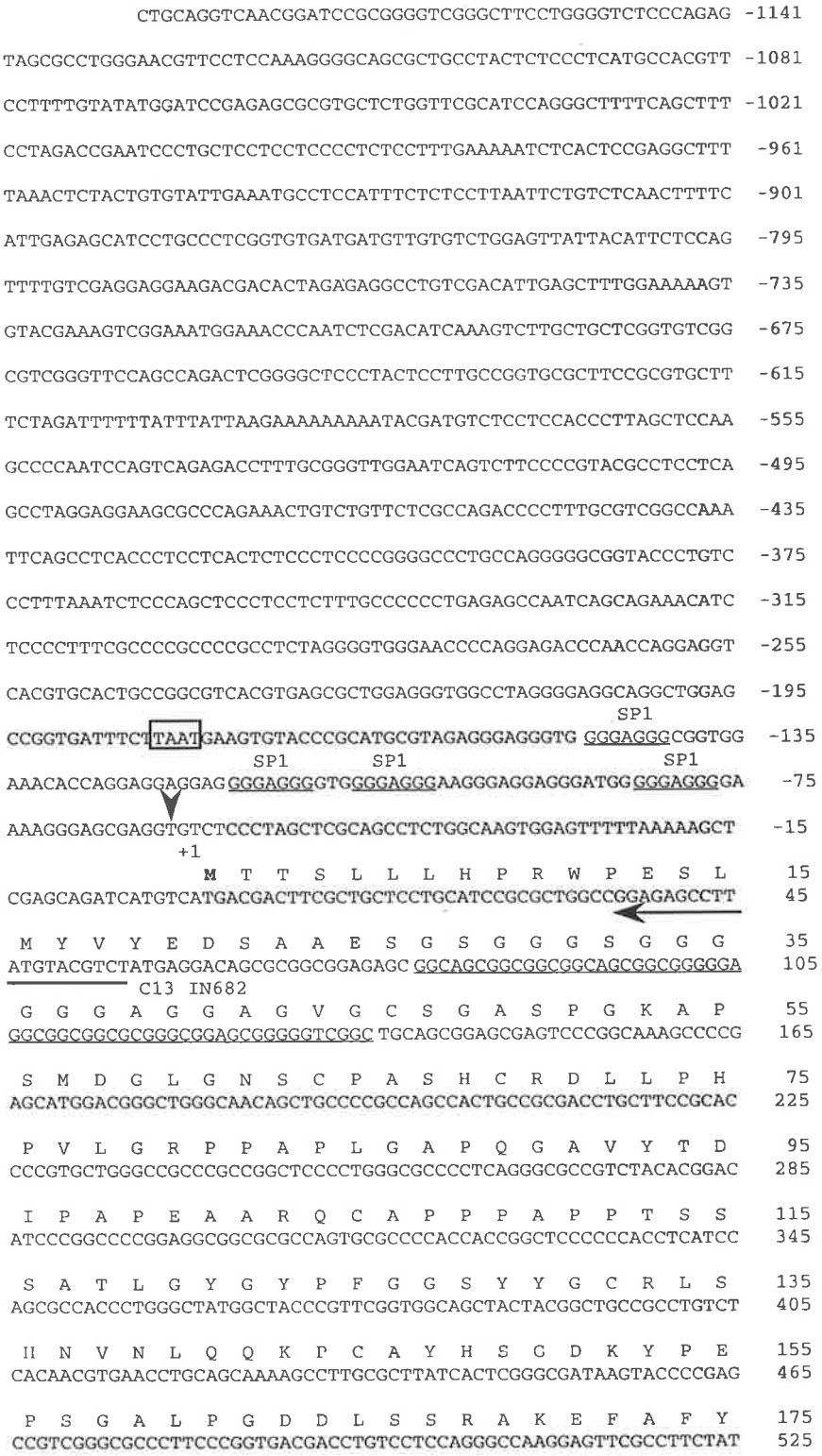
6.8 Predicted Secondary Structure and Surface Probability of Hoxc-13

The secondary structure of sheep Hoxc-13 was predicted using the Chou-Fasman method (Fig. 6.11 A, Mac Vector 6.0 Software). This method predicts that the Hoxc-13 protein consists of many turns and some regions that may form a helix structure. The homeodomain region, known to have a helix-turn-helix conformation was predicted at the C-terminal end.

Using Mac Vector 6.0 software, regions of the protein that are most likely to lie on the surface were predicted (Fig. 6.11 B). This prediction was based on knowledge of which amino acids are more likely to be found on the surface of proteins of known structure. The majority of the amino acids at the Hoxc-13 N-terminal end (amino acids 1-250) were predicted to be internal. The homeodomain at the C-terminal end was predicted to be found on the surface of the protein. This is in accordance with the known ability of homeodomain proteins to bind DNA and act as transcription factors.

Figure 6.10 : The Nucleotide Sequence and Predicted Amino Acid Sequence of *Hoxc-13*.

Hoxc-13 was sequenced as outlined in Figure 6.8 and Figure 6.9 and the resultant nucleotide and predicted amino acid sequence (single letter code) are shown. *Hoxc-13* mRNA is 2,361 bases in length plus the poly- A tail and contains an open reading frame of 330 amino acids. The homeodomain is boxed. The proposed polyadenylation signal is in bold and underlined at nucleotides 2261-2266. The stop codon (TGA) is denoted by an asterisk. Nucleotide numbering begins at the first nucleotide (adenine) of the start codon ATG that has been designated +1. Numbering of amino acids starts from the methionine residue shown in bold. The full intronic sequence is not shown. A glycine -rich region is underlined at the amino end (amino acids 27-45). The transcriptional start point is indicated by the arrow-head. The primer (C13 In682) used to determine the transcriptional start point (see Fig. 6.11) is indicated by an arrow between nucleotides 35–55. Arrows indicate the location of primers (PRT, P4 and P14) used during the 5' RACE technique. A region rich in G and A nucleotides which consists of multiple putative SP1 binding motifs is underlined in the 5' flanking region (nucleotides -127 to - 67). A possible homeobox binding motif (TAAT) is enclosed by the rectangle (nucleotides -182 to -179).


CTGCAGGTCAACGGATCCGCGGGGTCGGGCTTCCTGGGGTCTCCCAGAG -1141
TAGCGCCTGGGAACGTTCCTCCAAAGGGGCAGCGCTGCCTACTCTCCCTCATGCCACGTT -1081
CCTTTTGTATATGGATCCGAGAGCGCGTGCTCTGGTTCGCATCCAGGGCTTTTCAGCTTT -1021
CCTAGACCGAATCCCTGCTCCTCCTCCCCTCTCCTTTGAAAAATCTCACTCCGAGGCTTT -961
TAAACTCTACTGTGTATTGAAATGCCTCCATTTCTCTCCTTAATTCTGTCTCAACTTTTC -901
ATTGAGAGCATCCTGCCCTCGGTGTGATGATGTTGTGTCTGGAGTTATTACATTCTCCAG -795
TTTTGTCGAGGAGGAAGACGACACTAGAGAGGCCTGTCGACATTGAGCTTTGGAAAAAGT -735
GTACGAAAGTCGGAAATGGAAACCCAATCTCGACATCAAAGTCTTGCTGCTCGGTGTCGG -675
CGTCGGGTTCCAGCCAGACTCGGGGCTCCCTACTCCTTGCCGGTGCGCTTCCGCGTGCTT -615
TCTAGATTTTTTATTTATTAAGAAAAAAAAAATACGATGTCTCCTCCACCCTTAGCTCCAA -555
GCCCCAATCCAGTCAGAGACCTTTGCGGGTTGGAATCAGTCTTCCCCGTACGCCTCCTCA -495
GCCTAGGAGGAAGCGCCCAGAAACTGTCTGTTCTCGCCAGACCCCTTTGCGTCGGCCAAA -435
TTCAGCCTCACCCTCCTCACTCTCCCTCCCCGGGGCCCTGCCAGGGGGCGGTACCCTGTC -375
CCTTTAAATCTCCCAGCTCCCTCCTCTTTGCCCCCCTGAGAGCCAATCAGCAGAAACATC -315
TCCCCTTTCGCCCCGCCCCGCCTCTAGGGGTGGGAACCCCAGGAGACCCAACCAGGAGGT -255
CACGTGCACTGCCGGCGTCACGTGAGCGCTGGAGGGTGGCCTAGGGGAGGCAGGCTGGAG -195
SP1
CCGGTGATTTCTTAATGAAGTGTACCCGCATGCGTAGAGGGAGGGTG GGGAGGGCGGTGG -135
SP1 SP1 SP1
AAACACCAGGAGGAGGAG GGGAGGGGTGGGGAGGGAAGGGAGGAGGGATGG GGGAGGGGA -75
AAAGGGAGCGAGGTGTCTCCCTAGCTCGCAGCCTCTGGCAAGTGGAGTTTTTAAAAAGCT -15
+1
M T T S L L L H P R W P E S L 15
CGAGCAGATCATGTCATGACGACTTCGCTGCTCCTGCATCCGCGCTGGCCGGAGAGCCTT 45
M Y V Y E D S A A E S G S G G G S G G G 35
ATGTACGTCTATGAGGACAGCGCGGCGGAGAGC GGCAGCGGCGGCGGCAGCGGCGGGGGA 105
C13 IN682
G G G A G G A G V G C S G A S P G K A P 55
GGCGGCGGCGCGGGCGGAGCGGGGGTCGGC TGCAGCGGAGCGAGTCCCGGCAAAGCCCCG 165
S M D G L G N S C P A S H C R D L L P H 75
AGCATGGACGGGCTGGGCAACAGCTGCCCCGCCAGCCACTGCCGCGACCTGCTTCCGCAC 225
P V L G R P P A P L G A P Q G A V Y T D 95
CCCGTGCTGGGCCGCCCGCCGGCTCCCCTGGGCGCCCCTCAGGGCGCCGTCTACACGGAC 285
I P A P E A A R Q C A P P P A P P T S S 115
ATCCCGGCCCCGGAGGCGGCGCGCCAGTGCGCCCCACCACCGGCTCCCCCCACCTCATCC 345
S A T L G Y G Y P F G G S Y Y G C R L S 135
AGCGCCACCCTGGGCTATGGCTACCCGTTCGGTGGCAGCTACTACGGCTGCCGCCTGTCT 405
H N V N L Q Q K P C A Y H S C D K Y P E 155
CACAACGTGAACCTGCAGCAAAAGCCTTGCGCTTATCACTCGGGCGATAAGTACCCCGAG 465
P S G A L P G D D L S S R A K E F A F Y 175
CCGTCGGGCGCCCTTCCCGGTGACGACCTGTCCTCCAGGGCCAAGGAGTTCGCCTTCTAT 525

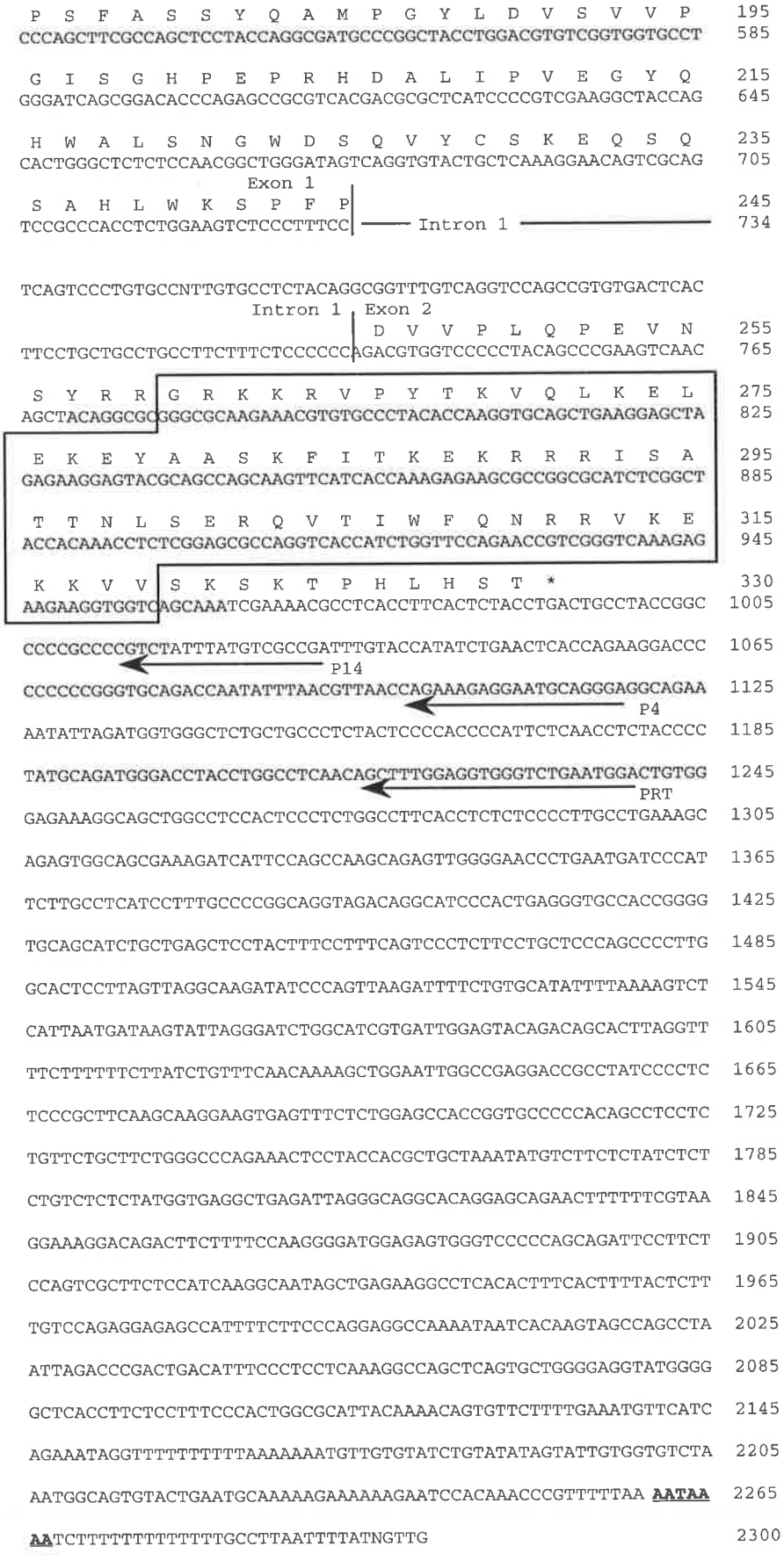
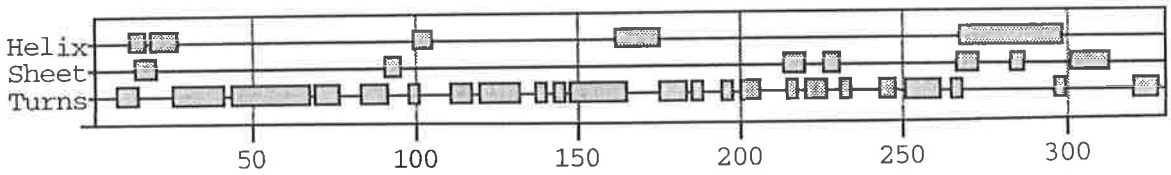

P S F A S S Y Q A M P G Y L D V S V V P 195
CCCAGCTTCGCCAGCTCCTACCAGGCGATGCCCGGCTACCTGGACGTGTCGGTGGTGCCT 585
G I S G H P E P R H D A L I P V E G Y Q 215
GGGATCAGCGGACACCCAGAGCCGCGTCACGACGCGCTCATCCCCGTCGAAGGCTACCAG 645
H W A L S N G W D S Q V Y C S K E Q S Q 235
CACTGGGCTCTCTCCAACGGCTGGGATAGTCAGGTGTACTGCTCAAAGGAACAGTCGCAG 705
Exon 1
S A H L W K S P F P 245
TCCGCCCACCTCTGGAAGTCTCCCTTTCC Intron 1 734
TCAGTCCCTGTGCCNTTGTGCCTCTACAGGCGGTTTGTCAGGTCCAGCCGTGTGACTCAC
Intron 1 Exon 2
D V V P L Q P E V N 255
TTCCTGCTGCCTGCCTTCTTTCTCCCCCCAGACGTGGTCCCCCTACAGCCCGAAGTCAAC 765
S Y R R G R K K R V P Y T K V Q L K E L 275
AGCTACAGGCGCGGGCGCAAGAAACGTGTGCCCTACACCAAGGTGCAGCTGAAGGAGCTA 825
E K E Y A A S K F I T K E K R R R I S A 295
GAGAAGGAGTACGCAGCCAGCAAGTTCATCACCAAAGAGAAGCGCCGGCGCATCTCGGCT 885
T T N L S E R Q V T I W F Q N R R V K E 315
ACCACAAACCTCTCGGAGCGCCAGGTCACCATCTGGTTCCAGAACCGTCGGGTCAAAGAG 945
K K V V S K S K T P H L H S T * 330
AAGAAGGTGGTCAGCAAATCGAAAACGCCTCACCTTCACTCTACCTGACTGCCTACCGGC 1005
CCCCGCCCCGTCTATTTATGTCGCCGATTTGTACCATATCTGAACTCACCAGAAGGACCC 1065
P14
CCCCCCGGGTGCAGACCAATATTTAACGTTAACCAGAAAGAGGAATGCAGGGAGGCAGAA 1125
P4
AATATTAGATGGTGGGCTCTGCTGCCCTCTACTCCCCACCCCATTCTCAACCTCTACCCC 1185
TATGCAGATGGGACCTACCTGGCCTCAACAGCTTTGGAGGTGGGTCTGAATGGACTGTGG 1245
PRT
GAGAAAGGCAGCTGGCCTCCACTCCCTCTGGCCTTCACCTCTCTCCCCTTGCCTGAAAGC 1305
AGAGTGGCAGCGAAAGATCATTCCAGCCAAGCAGAGTTGGGGAACCCTGAATGATCCCAT 1365
TCTTGCCTCATCCTTTGCCCCGGCAGGTAGACAGGCATCCCACTGAGGGTGCCACCGGGG 1425
TGCAGCATCTGCTGAGCTCCTACTTTCCTTTCAGTCCCTCTTCCTGCTCCCAGCCCCTTG 1485
GCACTCCTTAGTTAGGCAAGATATCCCAGTTAAGATTTTCTGTGCATATTTTAAAAGTCT 1545
CATTAATGATAAGTATTAGGGATCTGGCATCGTGATTGGAGTACAGACAGCACTTAGGTT 1605
TTCTTTTTTCTTATCTGTTTCAACAAAAGCTGGAATTGGCCGAGGACCGCCTATCCCCTC 1665
TCCCGCTTCAAGCAAGGAAGTGAGTTTCTCTGGAGCCACCGGTGCCCCCACAGCCTCCTC 1725
TGTTCTGCTTCTGGGCCCAGAAACTCCTACCACGCTGCTAAATATGTCTTCTCTATCTCT 1785
CTGTCTCTCTATGGTGAGGCTGAGATTAGGGCAGGCACAGGAGCAGAACTTTTTTCGTAA 1845
GGAAAGGACAGACTTCTTTTCCAAGGGGATGGAGAGTGGGTCCCCCAGCAGATTCCTTCT 1905
CCAGTCGCTTCTCCATCAAGGCAATAGCTGAGAAGGCCTCACACTTTCACTTTTACTCTT 1965
TGTCCAGAGGAGAGCCATTTTCTTCCCAGGAGGCCAAAATAATCACAAGTAGCCAGCCTA 2025
ATTAGACCCGACTGACATTTCCCTCCTCAAAGGCCAGCTCAGTGCTGGGGAGGTATGGGG 2085
GCTCACCTTCTCCTTTCCCACTGGCGCATTACAAAACAGTGTTCTTTTGAAATGTTCATC 2145
AGAAATAGGTTTTTTTTTTAAAAAAATGTTGTGTATCTGTATATAGTATTGTGGTGTCTA 2205
AATGGCAGTGTACTGAATGCAAAAAGAAAAAAGAATCCACAAACCCGTTTTTAA **AATAA** 2265
AATCTTTTTTTTTTTTTTGCCTTAATTTTATNGTTG 2300

Fig 6.11 : Predicted Secondary Structure and Surface Probability of Hoxc-13

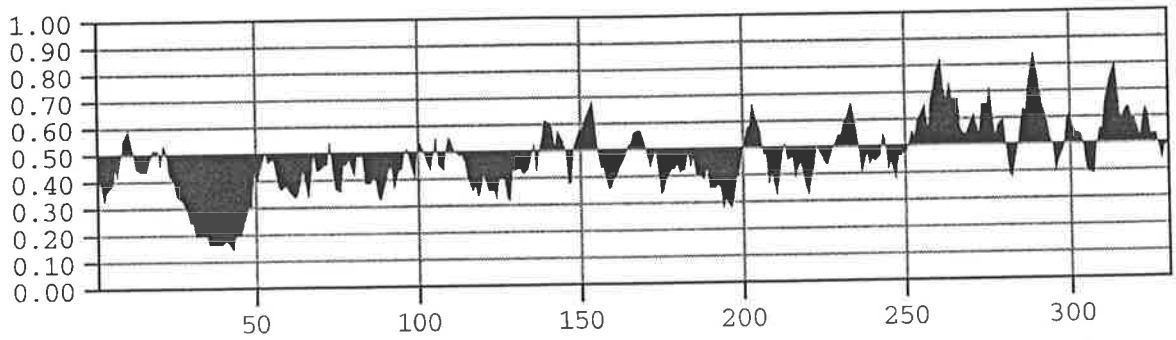
A. The predicted secondary structure of Hoxc-13 was predicted by Mac Vector 6.0 software (Oxford Molecular Group) that used the method of Chou-Fasman. Note that the homeodomain region located at the C-terminal end is predicted to have a helix-turn-helix structure in accordance with its DNA binding ability.

B. Using Mac Vector 6.0 software, the regions of Hoxc-13 most likely to lie on the surface were predicted. Note that the region of the homeodomain is predicted to reside on the surface.

A Secondary Structure



B Surface Probability



6.9 Identification of the *Hoxc-13* Transcriptional Start Site

The start of transcription of the gene was determined by primer extension analysis (Fig. 6.12). A gene-specific oligonucleotide (C13IN682; Fig. 6.10, Nucleotides 35 – 55) located in the first exon produced a strong extension on sheep wool follicle mRNA. One extension product was detected using both total RNA and poly A (+) RNA resulting in a predicted 5' non-coding region of 62 bp.

6.10 Species Conservation of *Hoxc-13*

To determine if the coding region of the *Hoxc-13* gene is conserved between species a Southern blot containing sheep, human and mouse genomic DNA cleaved separately with the enzymes BamHI, EcoR1 and HindIII (Fig. 6.13 A) was probed with a *Hoxc-13* 3' coding fragment (Fig. 6.13 B; see Appendix 2 C for plasmid map of clone C13-COD). The probe hybridised most strongly to sheep DNA but also bound to mouse and human DNA. The sizes of the hybridised fragments are given in table 1.

Table 1. Sizes of the mouse, human and sheep genomic DNA fragments that hybridised to the ovine *Hoxc-13* coding probe.

Enzyme / Species	Human	Mouse	Sheep
BamHI	6.5 kb	7.4 kb	7.0 kb
EcoR1	2.1 kb	8.9 kb	3.5 kb
HindIII	10 kb	3.5 kb	6.0 kb

Figure 6.12 : Mapping of the Ovine *Hoxc-13* Transcription Start Point.

The transcription start point of ovine *Hoxc-13* was located by a 5' extension reaction using the oligonucleotide C13IN682 (see Fig. 6.10 for location of C13IN682) and 10 µg of total RNA (lane 1 and 3) or 0.5 µg of mRNA(lane 2 and 4) isolated from Corriedale wool follicles. An oligonucleotide, P150B (Powell et al, 1992) detects a 74 base extension product from the hair keratin gene K2.9 and was used as a control in separate 5' extension reactions (lanes 1 and 2). Size of the 5' extension products were estimated from a DNA sequence ladder at the right of the figure. C13IN682 detected a 117 base extension product (lanes 3 and 4; indicated by the arrow on the left of the figure) resulting in a 5' non-coding sequence of 62 bases.

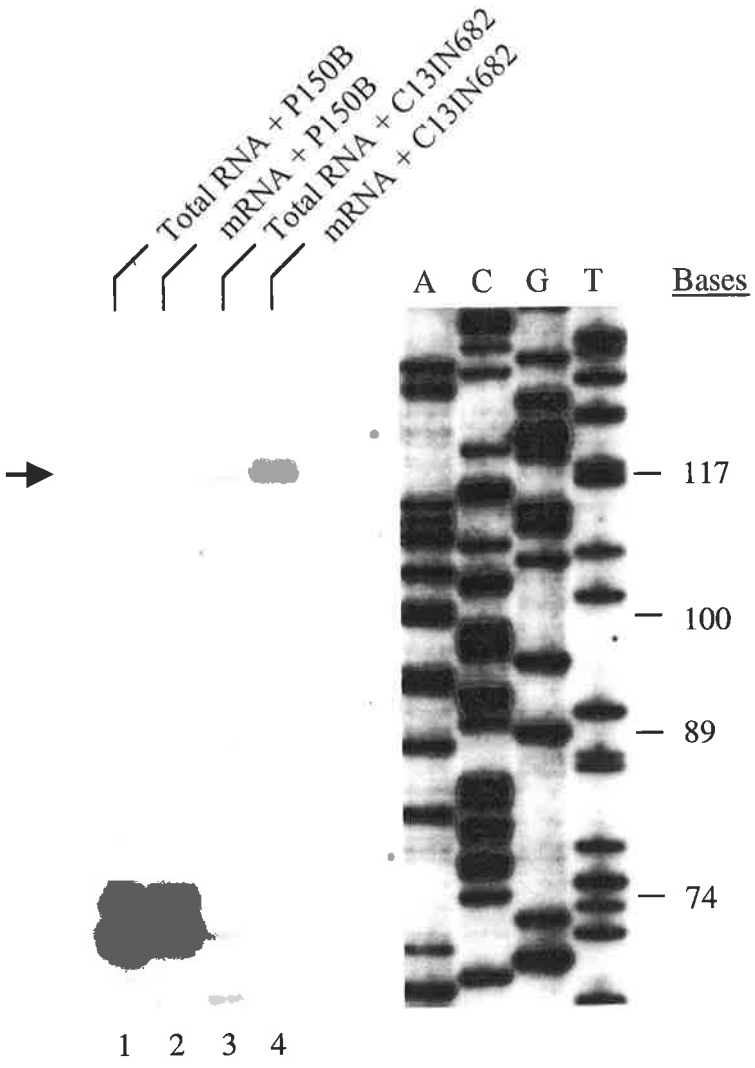
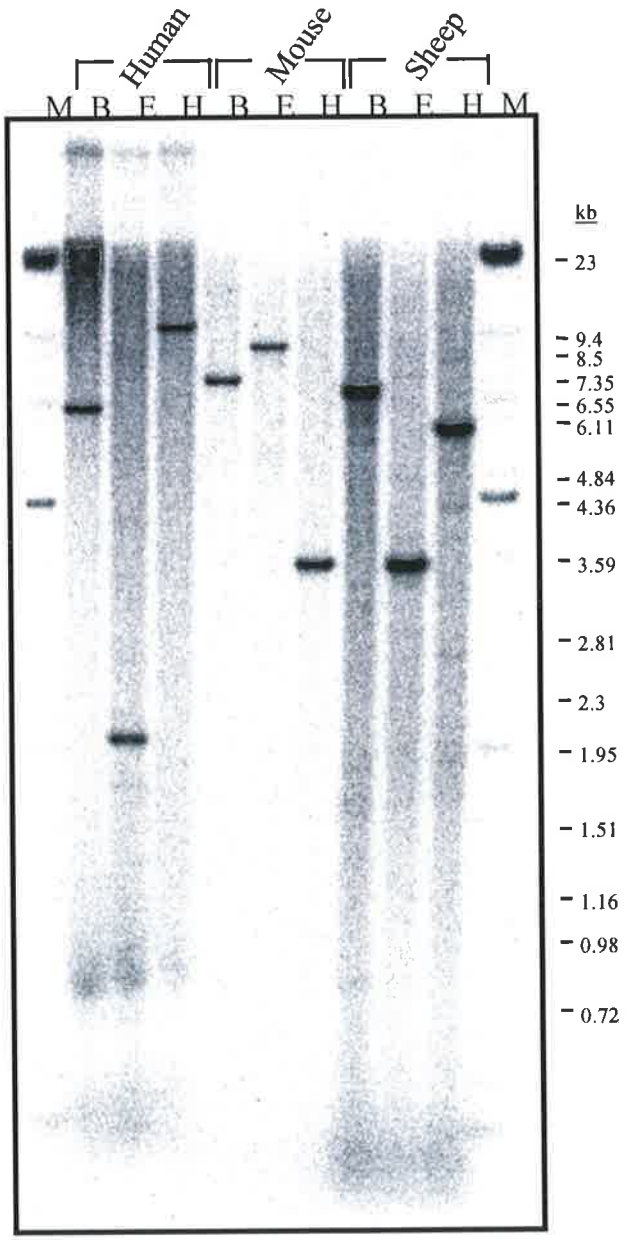

Total RNA + P150B
mRNA + P150B
Total RNA + C13IN682
mRNA + C13IN682
A C G T
Bases
117
100
89
74
1 2 3 4

Figure 6.13 : Hybridisation Analysis of a Genomic Zoo Blot with Ovine *Hoxc-13* Probes

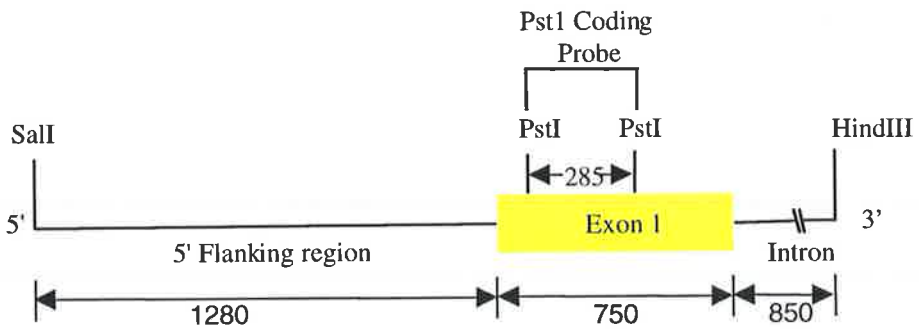
A. Human, mouse, and sheep total genomic DNA (10 µg / lane) were digested with BamHI (B), EcoR1 (E), and HindIII (H), size-separated on a 1% agarose gel, transferred to Zetaprobe GT membrane and probed with the resected *Hoxc-13* coding region fragment (clone C13-COD, see Appendix 2 C for plasmid map). The filter was washed in 2.0 X SSC, 0.1% SDS at 65^{o}C for 30 minutes and exposed to an autoradiograph for 21 days at -80^{o}C. M = Size markers are shown in kilobase pairs at the right of the figure.

B. Schematic of the 2.8 kb SalI / HindIII *Hoxc-13* fragment from the genomic lambda clone 9A showing the location of the 285 bp PstI coding fragment (located in exon 1) used as a radio-labelled probe in the Southern blot analysis in A.

A


Human
Mouse
Sheep
M B E H B E H B E H M
kb
23
9.4
8.5
7.35
6.55
6.11
4.84
4.36
3.59
2.81
2.3
1.95
1.51
1.16
0.98
0.72

B


Pst1 Coding
Probe
SalI
PstI
PstI
HindIII
285
5'
Exon 1
3'
5' Flanking region
Intron
1280
750
850

6.11 Discussion

This chapter described the isolation and sequencing of ovine *Hoxc-13*. This represents the first complete *Hoxc-13* sequence. Initially, 1.4 kb of the 3' cDNA end of *Hoxc-13* was amplified by 3' RACE-PCR and sequenced. The sequence downstream from the homeobox could not be compared to the human or mouse sequence because there is no published sequence information outside of the homeodomain for *Hoxc-13* orthologues. Without comparative sequence data it was necessary to ensure that the amplified 3' RACE-PCR product was representative of the *Hoxc-13* gene. Southern blot and Northern blot analysis confirmed that the 3' non-coding fragment isolated from the 1.4 kb 3' RACE-PCR product was from *Hoxc-13* mRNA. The Northern blot, probed separately with the homeobox and 3' non-coding probes confirmed that *Hoxc-13* is transcribed into a 2.4 kb mRNA species.

The acquisition of the remaining 800 bp of the *Hoxc-13* 5' cDNA end was difficult. Failure to obtain this sequence by 5' New RACE was possibly due to the inherent high GC content within the 5' end of the *Hoxc-13* coding region. Since GC- rich regions form tertiary structures that have a high melting temperature. These structures may cause DNA polymerase and reverse transcriptase to pause as they are replicating the RNA or DNA template strand and, consequently, the full-length coding region is not synthesised and clones containing partial cDNA result. The recent availability of thermostable reverse transcriptases, such as Tth (Astral Scientific) may reduce such problems in future.

Several attempts were made in the isolation of the 5' end of *Hoxc-13* with each attempt contributing more sequence. The 'new 5' RACE' technique (Frohman, 1994) extended the sequence upstream from the homeobox by only 80 bp. Clones isolated from an in-house sheep follicle cDNA library extended the sequence by another 450 bp and enabled the first exon to be mapped in the sheep genomic clone, 9A. Finally, the remaining 5' end and 1.2 kb of flanking sequence was obtained using clone 9A.

An ovine *Hoxc-13* coding probe hybridised to genomic DNA, digested with BamHI, EcoR1 or HindIII, from human and mouse suggesting that the gene is likely

conserved in these species. *Hoxc-13* encodes a predicted 330 amino acid protein similar to other *Abd-B*-related genes. The N-terminal end contains a glycine-rich domain between residues 27 and 48. Stretches of homopolymeric amino acids (HPAA) are a feature commonly found in homeobox genes (reviewed in Duboule, 1994). Stretches of glycine HPAA have been reported in a number of homeobox genes *including Da BarH1*, *Dfd*, *Df Ubx*, *Md Ubx*, *Ubx*, *inv*, *lab*, *otd*, *s59/NK-1*, *x Hoxb-5*, *h HOXA10*, *h HOXB3*, h *Hoxc-6*, *h HOXD3*, *h HOXD9*, *AmphiHox3*, *m Evx-2*, *ATBF1*, *TCL-3*, *h Oct-2*, *m Oct-2*, *m SCIP*, *r SCIP*, *m Brn-1*, *m Brn-2*, *m Hoxb-3*, *m Hoxc-6*, *m Hoxd-9*, *m Hoxd-11*, *Tes-1*, *m TTF-1*, *r TTF-1* (all references can be found in Duboule, 1994). The function of homopolymeric amino acid regions is not clear. Some homopolymeric amino acid regions are thought to be involved in rapid protein degradation in the PEST hypothesis (Rogers et al., 1986) while others (acidic regions, glutamine-rich regions and proline-rich regions) have been shown to function as transcriptional activators (Ptashne, 1988; Struhl, 1989).

The ovine *Hoxc-13* promoter lacks the TATA box however upstream of the *Hoxc-13* transcription start site are several SP1 motifs within a short distance of each other. Most cluster homeobox genes do not contain a TATA box but rely on other initiators such as SP1. For example both *Hoxa-4* (Galiot et al., 1989) and *Hoxc-4* (Geada et al., 1992) lack a TATA box and transcription is initiated by SP1. The transcription start point was located 16 nucleotides downstream from the last predicted SP1 motif or 62 nucleotides upstream of the translation start codon.

The gene structure of *Hoxc-13* is similar to other reported genes belonging to the *Abd-B* class. It possesses all of the hallmark amino acids in the homeodomain that characterise the 13 group paralogues, including a threonine at position 46 and a serine at position 41. *Hoxc-13* does not contain a Hexapeptide motif upstream of the homeodomain which distinguishes *Abd-B* genes from the other cluster genes. It also does not contain a conserved tryptophan residue (W) at position -6 preceding the homeodomain or a PY motif which are distinctive for all of the *Abd-B* genes (Izpisua-Belmonte et al., (1991) and the reason for these differences is not known.

Like the paralogous genes *Hoxa-13*, *Hoxb-13* and *Hoxd-13,* the ovine *Hoxc-13* gene contained only one intron located a short distance upstream from the homeobox. The homeodomain of Hoxc-13 was located at the carboxyl end in the second exon and was separated by a short region of eleven amino acids from the stop codon, a trait shared by the *Abd-B* type genes. For example, Hoxa-10 contains only 14 amino acid residues and Hoxb-13 contains 9 amino acid residues downstream of the homeodomain (Benson et al., 1995; Zeltser et al., 1996). The *Hoxc-13* 3' cDNA end did not contain the consensus polyadenylation signal AATAAA. A similar motif of AAGAAA known to work as a polyadenylation signal (Sheets et al., 1990) was present 19 bases upstream from the beginning of the poly A tail. However, sequence obtained from the *Hoxc-13* lambda genomic clone indicated a thymidine at that position, the typical nucleotide in the polyadenylation signal. This discrepancy may therefore be an artifact due to the misincorporation of the guanidine base by Taq polymerase in the early cycles of the 3' RACE-PCR. It could however, reflect a real difference because the cDNA and genomic libraries were prepared from material isolated from different sheep breeds, namely Corriedale (RT-PCR clones) and Merino (λ genomic clones).

The secondary structure of the Hoxc-13 protein, predictedby the Chou-Fasman method, consists of many turns that presumably internalise most of the regions of the protein. The homeodomain is the main region that is predicted to lie on the surface of the molecule and this is in accordance with the known ability of the homeodomain to make contact with DNA.

Chapter 7

Characterisation of Ovine *Hoxc-13*

Chapter 7 : Characterisation of Ovine *Hoxc-13*

7.1 Introduction

The involvement of *Hoxc-13* in follicle function was discovered in the initial RT-PCR screen for homeobox genes expressed in the wool follicle (see Chapters 3 and 4). Preliminary *in situ* localization using a homeobox gene probe suggested *Hoxc-13* was expressed in the follicle bulb and lower shaft. This indicated a potential functional role for *Hoxc-13* in the regulation of the structural proteins of the wool fibre as hair keratin gene expression begins in the upper bulb (Powell, 1992).

Using a gene-specific probe located outside of the homeobox region, the work presented here set out to confirm and extend the *in situ* data presented in Chapter 4. *Hoxc-13* expression in adult tongue and during embryonic pelage and vibrissa follicle development was studied by *in situ* hybridisation. *Hoxc-13* expression in the developing placode and embryonic skin was reported as a novel finding. Various adult sheep tissues were also screened for *Hoxc-13* expression by RNA protection using a gene-specific probe. Lastly, to investigate the hypothesis that *Hox*c-13 was involved in the regulation of keratin gene expression, the promoter of the intermediate-filament type I gene, *K14* was used to drive ectopic expression of *Hox*c-13 in transgenic mice.

7.2 Characterisation of *Hoxc-13* Expression by *In Situ* Hybridisation

7.2.1 *Hoxc-13* Expression in Embryonic Pelage Follicles

To investigate if *Hoxc-13* was differentially expressed during embryonic sheep pelage wool follicle development, the *in situ* hybridisation technique was used to locate its expression in Merino sheep skin, from the mid-flank region, at the embryonic time points of E51, E65, E80, E99 and E135. A gene-specific probe which hybridizes to the 3' non-coding region of the *Hoxc-13* gene (clone C13-NC, Appendix 2 D) was isolated from the 3' RACE clone, C24B1 (Appendix 2 B) and used in the *in situ* hybridisation analyses of *Hoxc-13* expression.

Hoxc-13 was expressed in a spatially and differentially restricted manner during both epidermal and follicle development. In the pre-follicle stage, at E51, *Hoxc-13* was expressed predominantly in the ectoderm which was 1-2 cells in thickness but was not expressed in the underlying dermis (Fig. 7.1 A, B).

At the initial stage of primary follicle development, at approximately E65 in sheep, when localised thickening was first evident at discrete sites in the ectoderm, *Hoxc-13* expression was present in the epithelial component of the placode (Fig. 7.1 D, E). At E65 and E80, expression was located throughout the basal and suprabasal layers of the ectoderm and appeared to be present in the mesenchymal cells (dermal condensate) directly beneath the epidermal basal layer (Fig. 7.1 D, E, G and H). However, at E65, *Hoxc-13* expression was absent from the mesenchymal cell aggregate that is destined to become the predermal papilla.

At E80, as the solid plug of epidermal cells moved downward into the dermis, *Hoxc-13* expression was detected mostly in the inner epithelial cells of the placode and at a reduced level in the outer cuboidal cells, (Fig. 7.1 G, H). Expression was absent from the adjacent pre-dermal papilla. No expression was detected in the periderm.

At E99 and E135, follicles were at a more advanced developmental stage (stage 7 and 8 respectively) and were producing a keratinised wool fibre. At both time points, *Hoxc-13* expression was detected in the upper bulb and lower shaft (Fig. 7.2 and Fig. 7.3). At E135, (but not at E99) faint expression of *Hoxc-13* was detected in the outer root sheath suprabasal layer (Fig. 7.3).

Figure 7.1 : *In situ* Hybridisation Analysis of *Hoxc-13* Expression in Embryonic Skin and Developing Pelage Follicles.

Brightfield (A, D and G) and darkfield (B, C, E, F, H, and I) photographs of sheep embryonic skin at E51 (A - C), E65 (D - F) and E80 (G - I), hybridised with an α-^{33}P-rUTP radio-labelled *Hoxc-13* cRNA non-coding probe synthesised from clone C13-NC (see Appendix 2 for plasmid reference). Sections were treated with a final wash stringency of 0.1 X SSPE at 65^{o}C for 30 minutes, exposed for 10 days at 4^{o}C and stained in Haemotoxylin.

Hoxc-13 is expressed in the epidermis at each time point. Arrow in D and E points to the follicle mesenchymal aggregate where *Hoxc-13* expression is not evident. At E80, *Hoxc-13* is expressed in the epithelial component of the placode and is absent in the predermal papilla (G and H).

The follicle placode, the predermal papilla and the basement membrane separating the epidermis and the dermis are indicated by the dashed line. Panels C, F and I represent darkfield photographs of hybridised sense probe to sheep embryonic skin sections at E51, E65 and E80 respectively. PDP = Predermal papilla. MA = Mesenchymal aggregate. Bar in each panel = 20 μm.

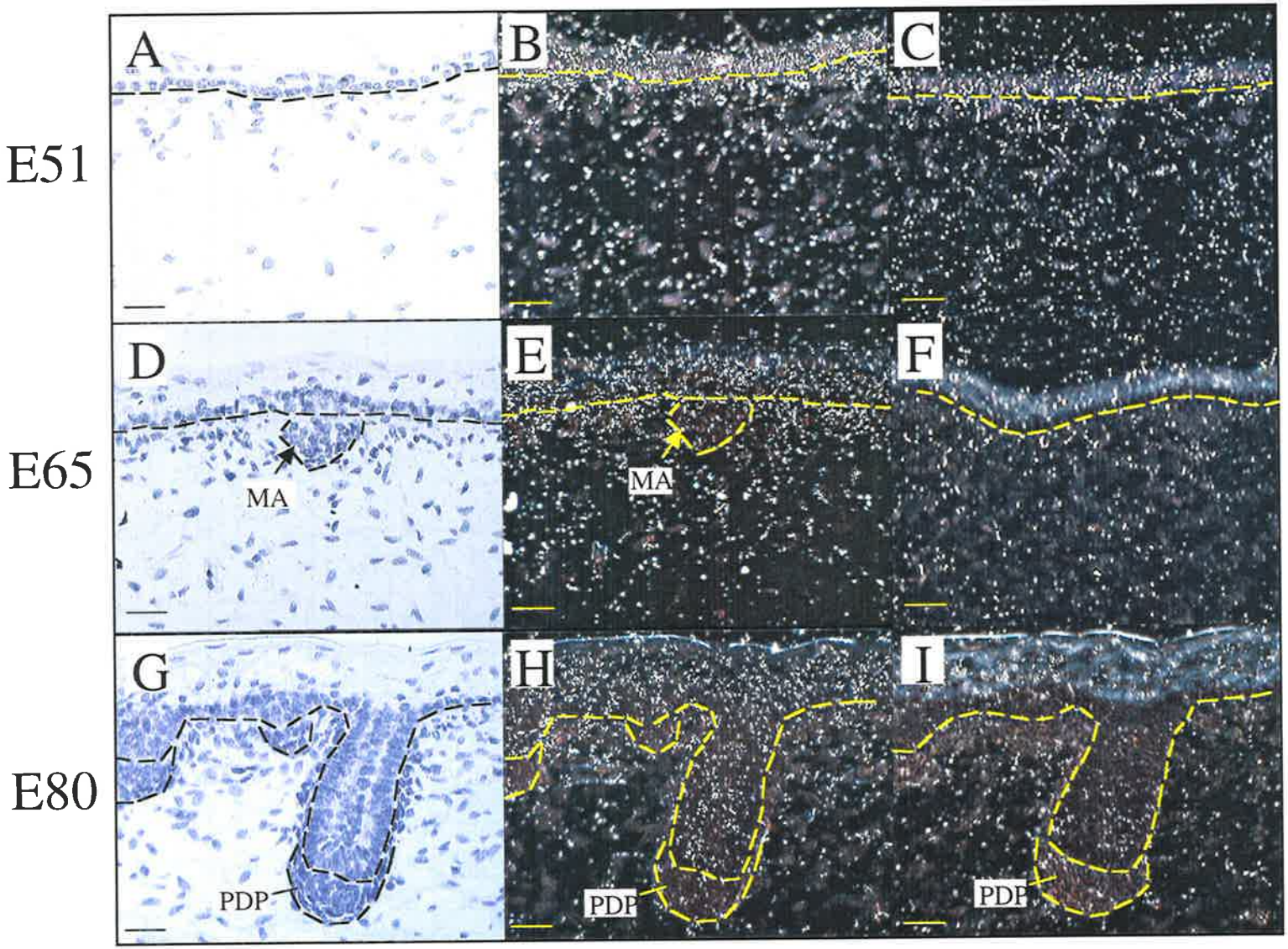

E51
E65
E80
A
B
C
D
E
F
G
H
I
MA
MA
PDP
PDP
PDP

Figure 7.2 : *In situ* Hybridisation Analysis of *Hoxc-13* Expression in E99 Sheep Skin.

Brightfield (A, C and E) and darkfield (B, D and F) photographs of an E99 sheep skin section. Probe, wash conditions, exposure time and staining were as for the skin sections in figure 7.1. Panels C and D are enlargements of the follicle outlined in panels A and B respectively. Panels B and D show *Hoxc-13* expression in the wool follicle bulb and lower shaft. Expression was strongest in the differentiating keratinocytes above the dermal papilla and lower shaft. *Hoxc-13* signal was not detected in the outer root sheath, epidermis or dermis. Panels E and F represent brightfield and darkfield photographs of hybridised sense probe. Note the high non-specific and random signal given by the sense probe in panel F. The follicle, dermal papilla and the basement membrane separating the epidermis and the dermis are indicated by the dashed line.

o = Outer root sheath, ci = Cortex and inner root sheath, DP = Dermal papilla. Bars in each panel = 20 μm.

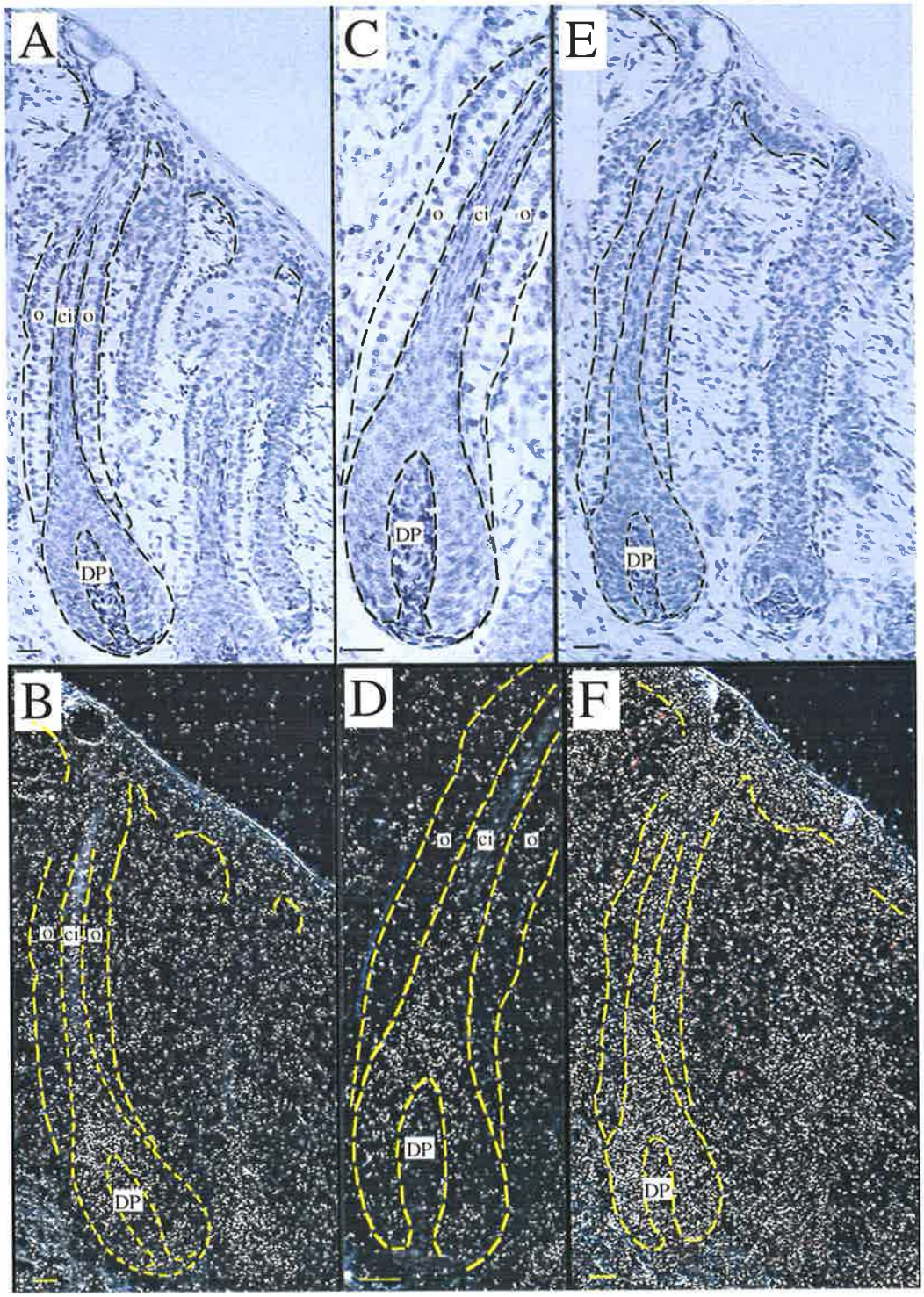
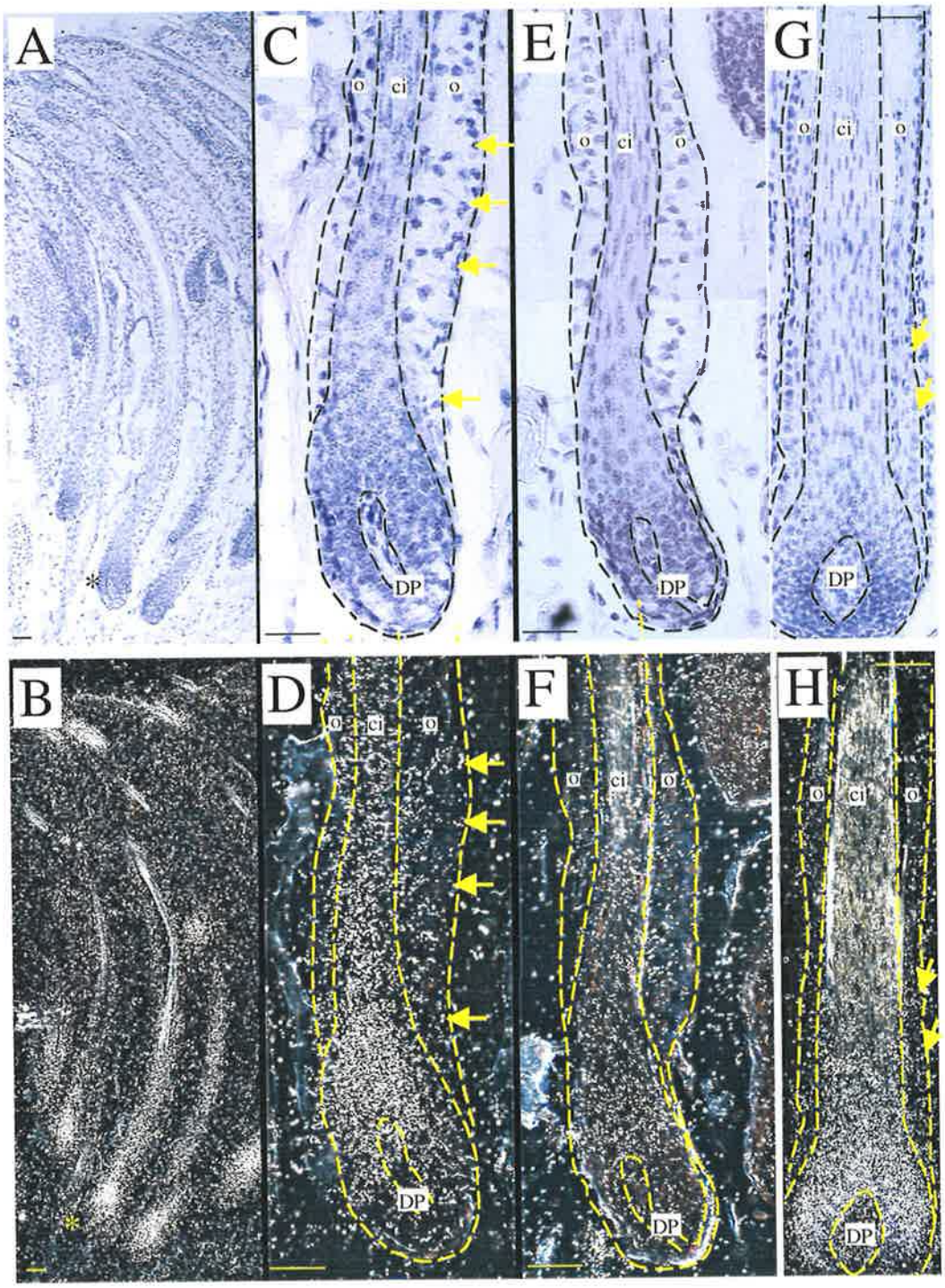

A
C
E
o
ci
o
DP
B
D
F

Figure 7.3 : *In situ* Hybridisation Analysis of *Hoxc-13* Expression in E135 Sheep Skin.

Brightfield (A, C, E and G) and darkfield (B, D, F and H) photographs of E135 sheep skin from the mid-flank region (A – F) and snout (G and H). The pelage follicle in panels C and D is an enlargement of the follicle denoted by an (*) in panel A. Probe, wash conditions, exposure time and staining were as for the skin sections in figure 7.1. *Hoxc-13* is expressed in the follicle bulb, lower shaft and outer root sheath suprabasal layer (arrows, panel C and D). Expression was strongest in the keratinocytes above and surrounding the dermal papilla and in the lower shaft. No *Hoxc-13* signal was detected in the epidermis. Brightfield and darkfield photographs of hybridised sense probe to sheep E135 skin section (mid-flank region) are shown in panels E and F. Some background signal is present in the suprabulbar region however this is well below the signal obtained for the antisense probe. The follicle outer root sheath, cortex/inner root sheath and dermal papilla are outlined by the dashed line. It should be noted that there is some artificial fluorescence in the cortex and inner root sheath in panel H. o = Outer root sheath, ci = Cortex and inner root sheath, DP = Dermal papilla. Bars in each panel = 20 μm.


A
C
E
G
B
D
F
H
o
ci
o
DP
*

At E99, *Hoxc-13* expression was not detected in the epidermal basal or suprabasal layers (Fig. 7.2. A, B). However, at E135, low expression was detected in the epidermal basal and suprabasal layers of some sections but not in others (compare Fig 7.3 A, B with Fig. 7.4 A, B). This conflicting result may due to varying levels of *Hoxc-13* expression in different sheep at E135 or secondly, and more likely, the RNA may have been degraded in some tissue sections but not in others. *Hoxc-13* was not expressed in the epidermis in adult sheep (Fig. 7.4. D, E). Interestingly, in E135 and adult skin, unidentified cells scattered in the upper dermis expressed *Hoxc-13* (Fig. 7.4 B and E).

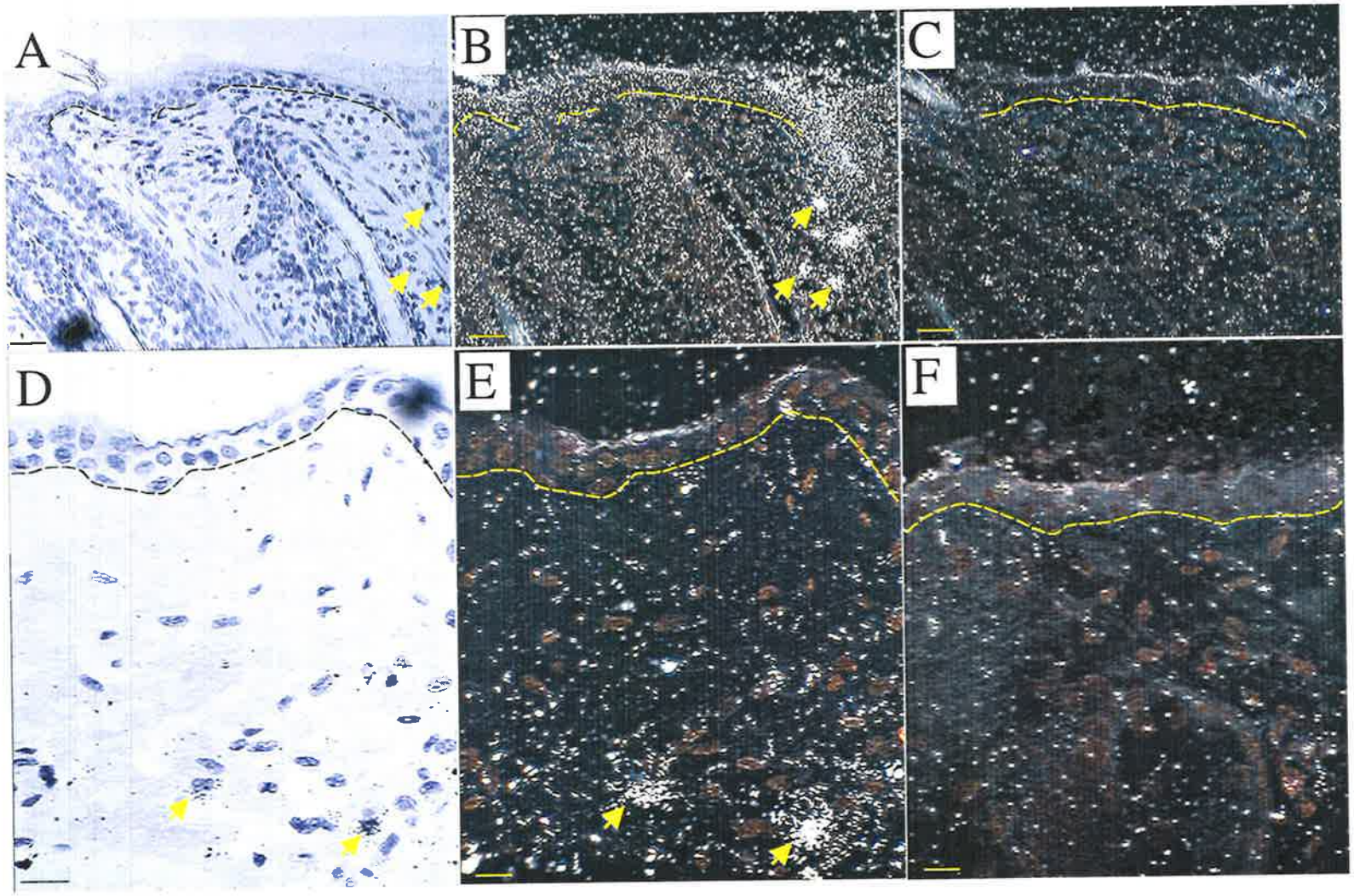
7.2.2 *Hoxc-13* Expression in Embryonic Vibrissa

Cluster homeobox gene expression in the developing embryo follows what has been termed the spatial colinearity rule such that genes located further towards the 5' end of the cluster show increasingly posteriorly-restricted expression. Due to the location of *Hoxc-13* at the 5' end of the cluster, expression would be predicted by the spatial colinearity rule to be restricted to follicles located at the posterior end (caudal) of the sheep. To test if *Hoxc-13* expression was expressed in anterior follicles of the sheep, vibrissa (whisker) follicles from the snout region, one of the most anterior parts of the sheep were studied. These follicles are several times larger than pelage follicles and enabled demarcation of expression boundaries to be distinguished more clearly. Mature vibrissa are also anatomically different to pelage follicles in that they contain a blood sinus that surrounds the outer root sheath and they produce a fibre that has a tactile function.

Hoxc-13 expression was studied in vibrissa follicles at the embryonic time points of E81, E88, and E135. Vibrissa follicles begin development at an earlier time point than pelage follicles and hence a keratin fibre can be detected before pelage follicles at E81. *Hoxc-13* expression was located in vibrissa follicles at all time points studied and its expression pattern was strikingly similar to the expression pattern seen in pelage follicles at comparable developmental stages. At each time point studied, *Hoxc-13* was expressed in the proliferating and differentiating cells of the bulb and lower shaft, with strongest expression in the keratinocytes located directly above and surrounding the dermal papilla

Figure 7.4 : *In situ* Hybridisation Analysis of *Hoxc-13* Expression in Embryonic Skin.

Brightfield (A and D) and darkfield (B, C, E and F) photographs of sheep embryonic skin section at E135 (A - C), and adult skin section (D - F). Probe, wash conditions, exposure time and staining were as for the skin sections in figure 7.1. *Hoxc-13* is expressed in the epidermis at E135 but not in adult epidermis. Arrows in panels A, B, D and E point to dermal cells expressing *Hoxc-13*. The basement membrane separating the epidermis and the dermis is indicated by the dashed line. Panels C and F represent darkfield photographs of sense probe hybridised to embryonic skin (E135) and adult respectively. Bar in A - C = 20 μm; D - F = 10 μm.


A
B
C
D
E
F

and in the lower shaft (Fig. 7.5 and Fig. 7.6). Expression was not detected in the proliferating cells located at the bottom of the follicle bulb or in the dermal papilla or outer dermal sheath in any vibrissa at any of the time points studied.

At E81, E88 and E135, expression was detected in the suprabasal outer root sheath layer, including the cells directly adjacent to the inner root sheath also known as the companion layer (Fig. 7.5 and Fig. 7.6). Expression appeared slightly stronger in the companion layer than in the suprabasal layer. Furthermore, at E135 *Hoxc-13* expression appeared to be present in the outer root sheath basal layer for the first time but at a reduced level compared to the suprabasal layer. At E135, signal was detected in the inner root sheath but this may be due to signal emanating from the hybridised, ^{33}P-rUTP radio-labelled *Hoxc-13* probe in the bulb and cortex and was therefore inconclusive at this stage. Analysis by non-radioactive *in situ* hybridisation, a technique that does not have the problem of signal scatter, may be able to determine the status of *Hoxc-13* expression in the inner root sheath. Interestingly, at E135, *Hoxc-13* expression was detected in a few cells of the blood sinus and dermis, however, these cells were not further identified (Fig. 7.6 A, B).

7.2.3 *Hoxc-13* Expression in Adult Pelage Wool Follicles

Hoxc-13 expression was identified in follicles of different sheep breeds that differ substantially in fibre diameter and medullation (Fig. 7.7). Merino, Romney, and Tukidale sheep follicles all expressed *Hoxc-13* in the rapidly dividing and differentiating keratinocytes above and surrounding the dermal papilla and the matrix cells lining the dermal papilla (Fig. 7.7). In each breed, *Hoxc-13* expression was strongest in a region of cells located above the dermal papilla and tapered off in the lower part of the follicle shaft. *Hoxc-13* expression was reduced in the cortical cells of the lower shaft and was reduced even more in the upper shaft. Tukidale sheep expressed *Hoxc-13* in the medulla of the lower shaft but further up the follicle shaft expression was reduced in the medulla (Fig. 7.7 E, F). In each breed, expression was detected in the outer root sheath suprabasal layer including the companion layer (Fig. 7.7 D). Expression was not detected in the dermal papilla, upper cortex or IRS of the adult follicle. Furthermore, expression was not detected in the epidermis or dermis of the adult skin (data not shown).

Figure 7.5 : *In situ* Hybridisation Analysis of *Hoxc-13* Expression in E81 and E88 Vibrissae Follicles.

Brightfield (A, C and E) and darkfield (B, D and F) photographs of Corriedale sheep E81 (A and B) and E88 (C - F) skin sections from the snout. Probe, wash conditions, exposure time and staining were as for the skin sections in figure 7.1. Panels E and F are enlargements of the area enclosed in the rectangle in panel C. Note the hybridisation signal of *Hoxc-13* in the middle of the bulb, surrounding the dermal papilla and in the lower follicle shaft. *Hoxc-13* is expressed in the suprabasal layers of the outer root sheath. Both follicles are cut at an oblique angle, hence the top half of the hair shaft is not shown. Some autofluorescence is detected in the inner root sheath. DP = Dermal papilla, c = Cortex, i = Inner root sheath, o = Outer root sheath. Bar in A, B = 20 μm; C, D = 12 μm; E, F = 8 μm.

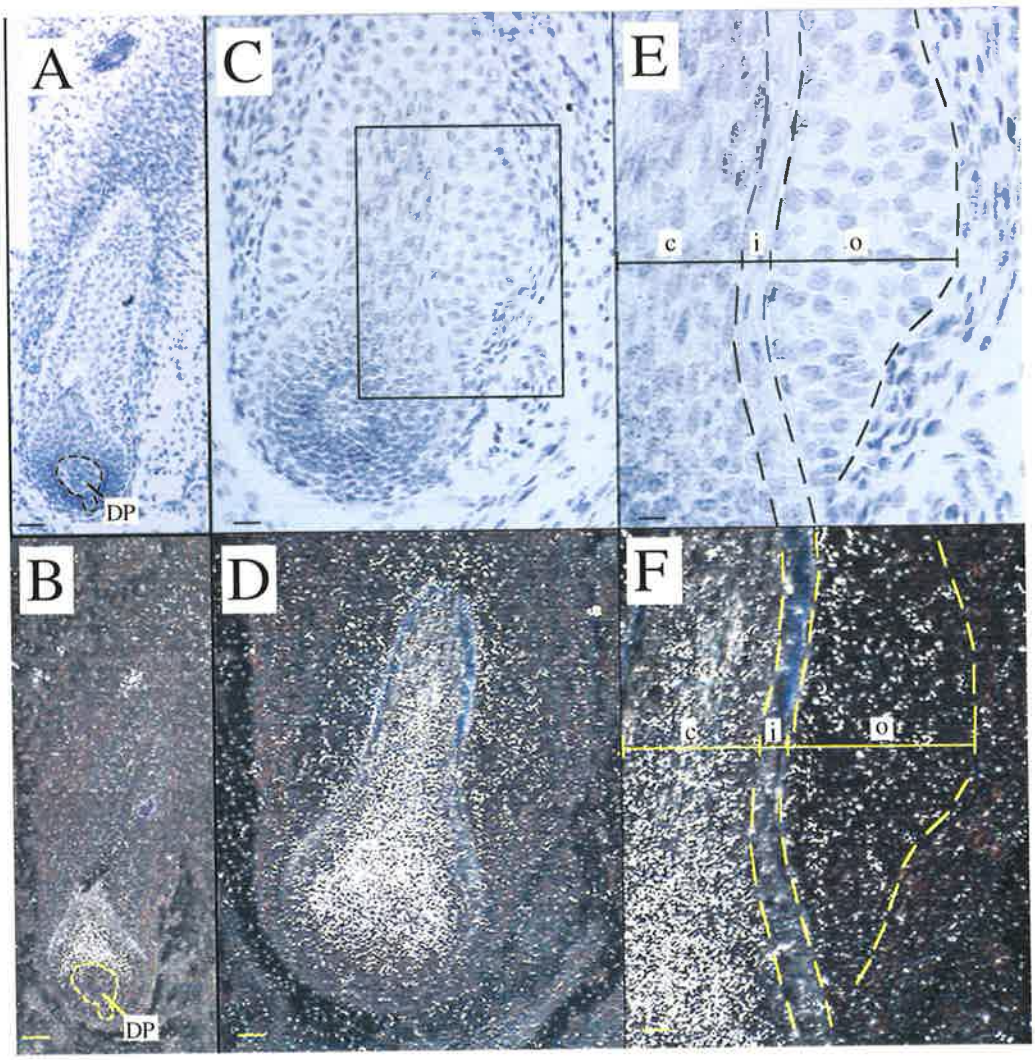

A
C
E
DP
c
i
o
B
D
F
DP
c
i
o

Figure 7.6 : *In situ* Hybridisation Analysis of *Hoxc-13* in an E135 Vibrissa Follicle.

Brightfield (A and C) and darkfield (B and D) photographs of a Corriedale vibrissa follicle at E135. Probe, wash conditions, exposure time and staining were as for the skin sections in figure 7.1. Panels C and D are enlargements of the boxed area in panel A. *Hoxc-13* is strongly expressed in the middle of the bulb, surrounding the dermal papilla and in the lower part of the follicle shaft. *Hoxc-13* is expressed in the basal and suprabasal layers of the outer root sheath. Arrows in panels A and B point to cells in the blood sinus expressing ovine *Hoxc-13*. The follicle is cut at an oblique angle, hence the top half of the hair shaft is not shown. bs = Blood sinus, DP = Dermal papilla, c = Cortex, i = Inner root sheath, o = Outer root sheath. Bar in A, B = 80 μm; C, D = 50 μm.

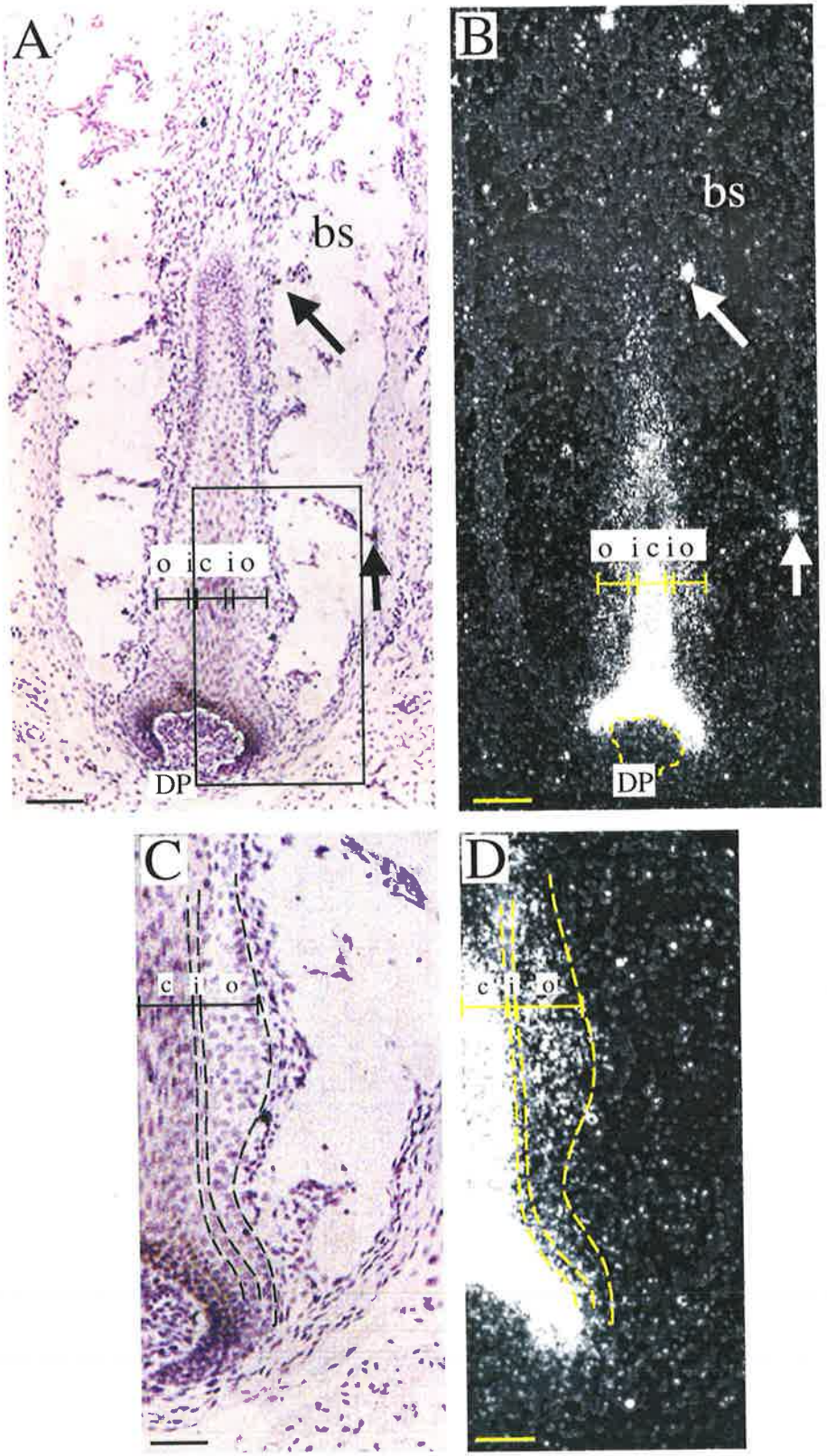
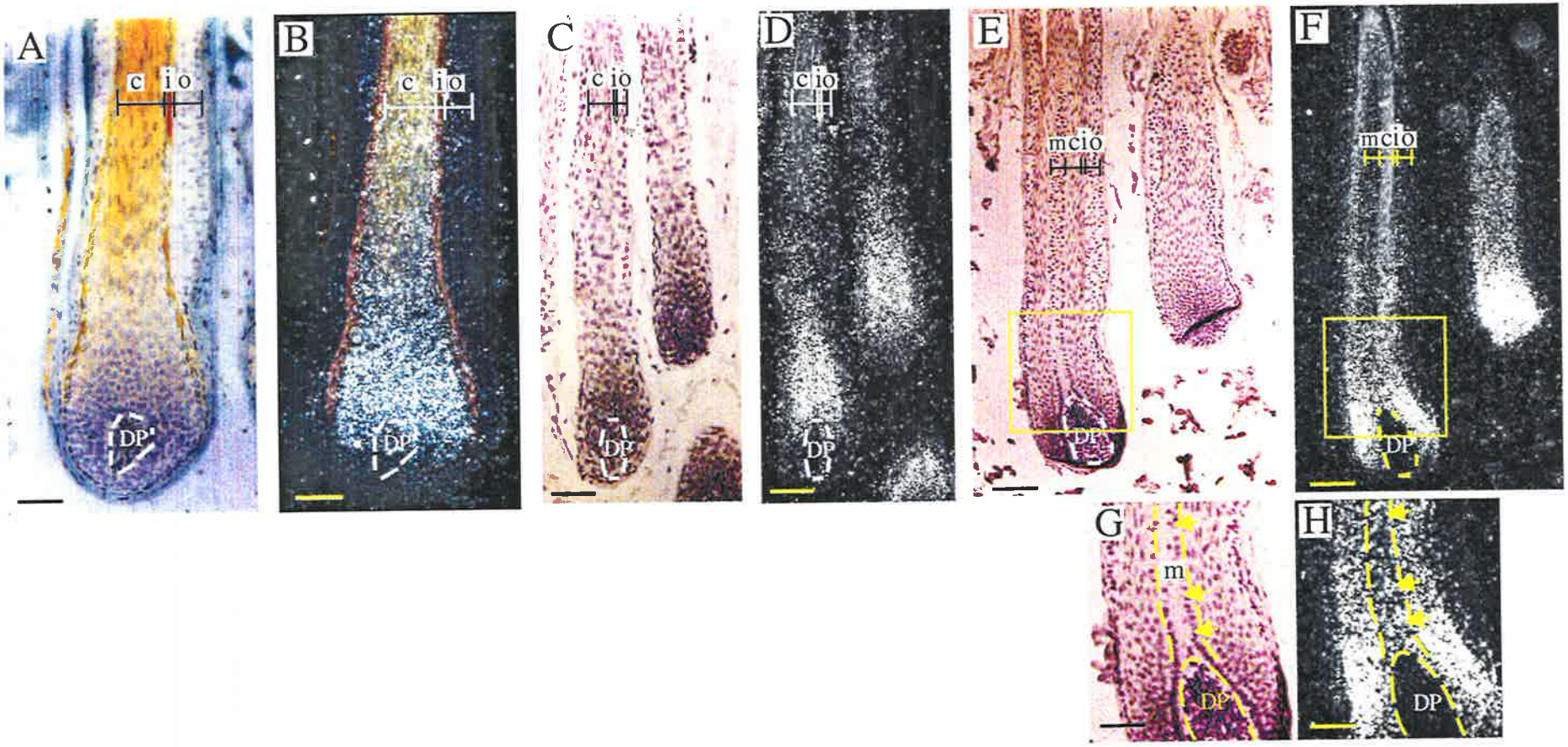

A
bs
o i c i o
DP
B
bs
o i c i o
DP
C
c i o
D
c i o

Figure 7.7 : *In situ* Hybridisation Analysis of Ovine *Hoxc-13* Expression in Adult Romney, Merino and Tukidale Follicles.

Brightfield (A, C, E and G) and darkfield (B, D, F and H) photographs of adult Romney (A and B), Merino (C and D) and Tukidale (E and F) skin sections. Probe, wash conditions, exposure time and staining were as for the skin sections in figure 7.1. *Hoxc-13* is expressed in the middle of the bulb, surrounding the dermal papilla and in the lower part of the follicle shaft including the medulla in Tukidale follicles (arrows in panels G and H). Panels G and H are magnifications of the boxed region in panels E and F. Signal is strongest in the differentiating keratinocytes above the dermal papilla. *Hoxc-13* is expressed in the outer root sheath of each breed. Signal is not detected in the inner root sheath or dermal papilla. Note, in B, D and F there is some autofluorescence in the follicle shaft. Note, in E and F the bottom of the follicle bulb has folded up during sectioning. c = Cortex, i = Inner root sheath, o = outer root sheath, m = Medulla. Bar in A, B = 40 μm, bar in C, D = 30 μm, bar in E, F = 60 μm, bar in G, F = 40 μm.


A
c
i
o
DP
B
c
i
o
DP
C
c
i
o
DP
D
c
i
o
DP
E
m
c
i
o
DP
F
m
c
i
o
DP
G
m
DP
H
DP

To identify whether *Hoxc-13* expression overlapped with the zone of cell proliferation in the lower bulb, expression of *Hoxc-13* was compared to *Histone H3* in consecutive skin sections from a Corriedale sheep (Fig. 7.8). *Histone H3* was a marker of cellular proliferation (Chou et al., 1990) and was expressed in the lower to middle region of the bulb where the mitotically active bulb cells are located. *Hoxc-13* expression appeared to overlap some of the mitotically active cells located in the middle of the wool follicle bulb surrounding the tip of the dermal papilla but did not extend into the lower regions of the bulb where *histone H3* was strongly expressed (compare Fig. 7.8 A, B and E, F).

As cells move up through the lower hair shaft a host of hair-keratin proteins are expressed (Powell and Rogers, 1994). *K2.10*, a hair intermediate filament type II gene, was one of the keratins detected in the lower shaft (Powell et al., 1992). Cells in the lower region of the shaft which express *K2.10* also express *Hoxc-13* (compare Fig. 7.8 A, B and C, D). Thus, expression of *Hoxc-13* appears to overlap the proliferation marker, *histone H3*, at its lower boundary and the differentiation marker *K2.10,* at its upper boundary.

7.3 Screening of Sheep Tissues for *Hoxc-13* Expression

Having established that *Hoxc-13* was expressed in both pelage and vibrissa follicles, various hair related and non-hair related tissues were screened for expression.
The tongue contains structures known as filiform papillae that are present on the dorsal side of the tongue. The filiform papilla contains specialised compartments that undergo esophageal, skin and hair-types of differentiation (Dhouailly et al., 1989). The tongue is thought to resemble the hair follicle due its expression of hair-specific keratin genes such as the hair keratin intermediate filament genes *mHa3* (*K1.2*), *mHb3* (and *mHb4*) which are expressed in the posterior compartment of the filiform papillae (Tobiasch et al., 1992; Winter et al., 1994). Furthermore, the antibody AE14, which was specific for late hair cortex differentiation reacts in the spine region of tongue filiform papillae (Dhouailly et al., 1989).

RNA *in situ* hybridisation was used to locate *Hoxc-13* expression in the sheep tongue. *Hoxc-13* expression was located in the base of the filiform papilla and was restricted to the hair-like compartment (Fig. 7.9). Expression did not appear to extend into

Figure 7.8 : *In situ* Hybridisation Analysis of *Hoxc-13*, *K2.10* and *Histone H3* Expression in Consecutive Corriedale Follicle Sections.

Brightfield (A, C and E) and darkfield (B, D and F) photographs of Corriedale skin sections hybridised to gene-specific *Hoxc-13* (A and B), *K2.10* (Powell et al., 1992; C and D) and *histoneH3* (Chou et al., 1990; E and F) cRNA probes. Probe, wash conditions, exposure time and staining were as for the skin sections in figure 7.1. Brackets in A and B locate the extent of hybridisation signal of *Hoxc-13*, from the middle of the bulb, to the lower shaft. Brackets in C and D indicate the extent of *K2.10* expression in the cortex which overlaps the *Hoxc-13* signal in the lower shaft. Brackets in E and F indicate the extent of *histone H3* expression in the mitotically active cells of the bulb. *Histone H3* signal in the middle of the bulb appears to overlap the lower expression boundary of *Hoxc-13*. The dermal papilla is outlined by dashed lines. The follicles are cut at an oblique angle, hence the top half of the hair shaft is not shown. c = Cortex, i = Inner root sheath and o = outer root sheath. Bar in A - E = 40 μm.

Hoxc-13

K2.10

Histone, H3

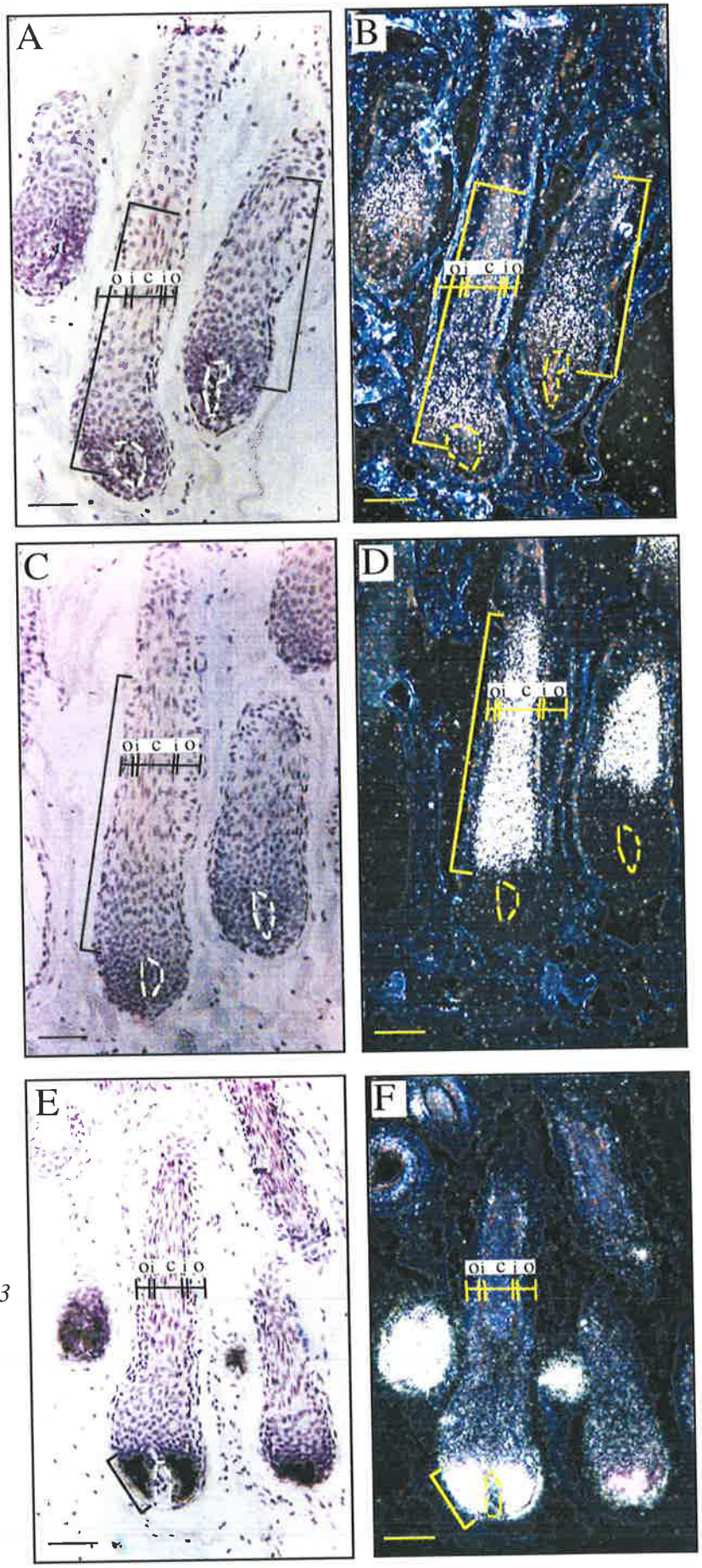
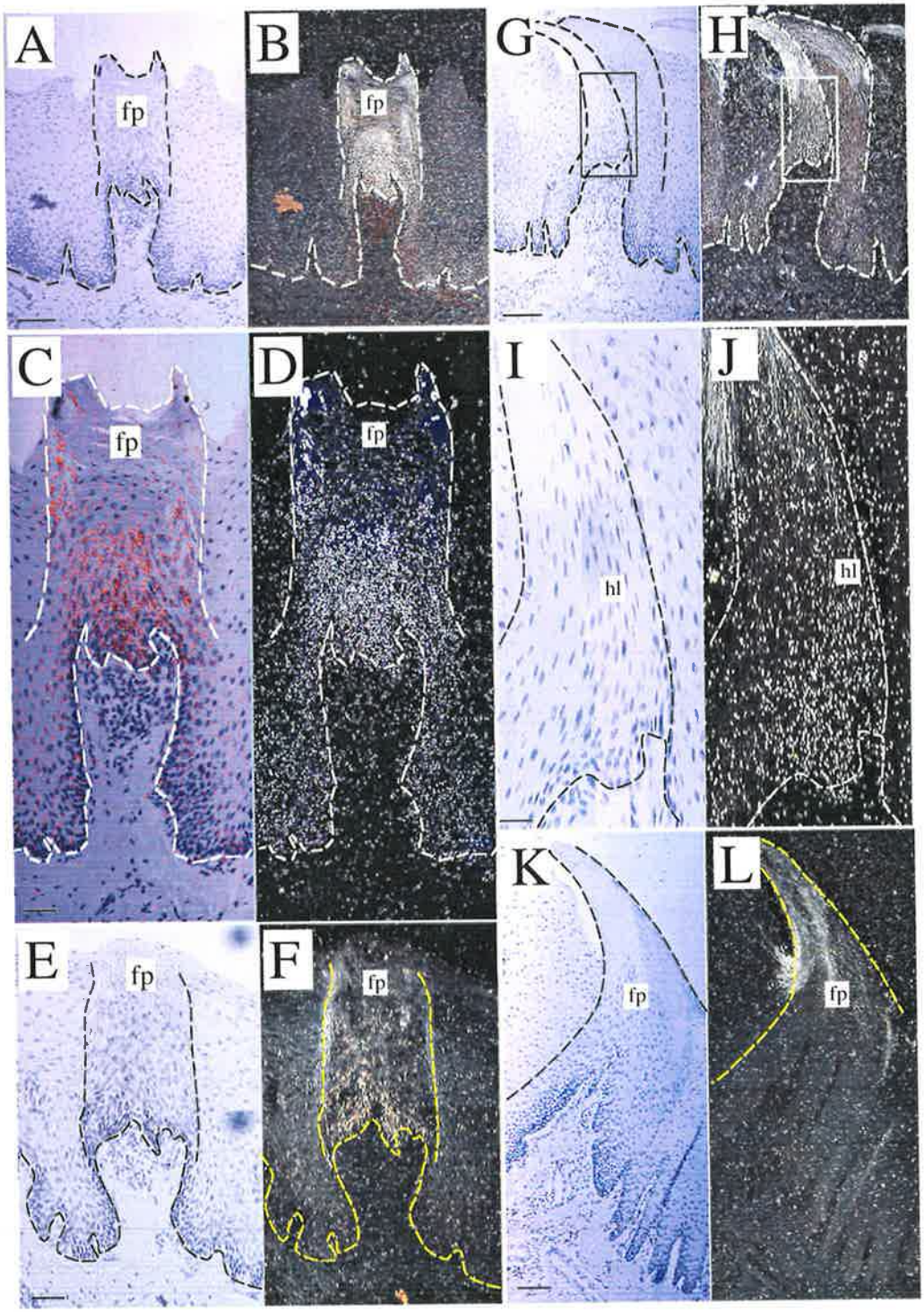


Figure 7.9 : *In situ* Hybridisation Analysis of *Hoxc-13* expression in Adult Sheep Tongue.

Brightfield (A, C, E, G, I and K) and darkfield (B, D, F, H, J and L) photographs of adult sheep tongue sections. Probe, wash conditions, exposure time and staining were as for the skin sections in figure 7.1. *Hoxc-13* is expressed in the base of the filiform papilla. Panels A - F and G - L show the filiform papilla from different view points, approximately 90 degrees horizontally rotated. In panels A – D and G – J, *Hoxc-13* is expressed in the base of the filiform papilla but more specifically in the hair-like compartment. Panels C and D are enlargements of the tongue filiform papilla shown in panels A and B respectively. The black grains in panel C were enhanced and converted to red for visual clarity using Adobe Photoshop. Panels I and J are enlargements of the boxed region shown in panels G and H respectively. Brightfield and darkfield photographs of hybridised sense probe to sheep tongue sections are shown in panels E, F, K and L. Note that there is some autofluorescence in panels B, H, F and L. The basal, suprabasal layer and filiform papilla are indicated by the dashed line. fp = Filiform papilla and hl = Hair-like compartment. The bar in panels A and G = 55 μm; C = 30μm; E = 45μm; I = 20μm; K = 40μm.


A
fp
B
fp
G
H
C
fp
D
fp
I
hl
J
hl
E
fp
F
fp
K
fp
L
fp

the keratinised spine that protrudes out of the surface of the tongue. Expression was not detected in the dermal papillae or suprabasal or basal epithelial layers of the tongue.

Hoxc-13 expression in various sheep tissues was analysed by RNA protection. *Hoxc-13* RNA transcripts were detected in adult wool follicle RNA but not in spleen, liver, thymus, brain, kidney, heart, large intestine, small intestine or lung (Fig. 7.10).

7.4 Functional Characterisation of *Hoxc-13*

In the adult, *Hoxc-13* expression appeared to be restricted to the hair follicle and tongue, both of which express hair-type keratins. The expression of *Hoxc-13* in the follicle bulb and in the tongue filiform papillae base, regions where the keratin genes are first expressed, strongly suggested that *Hox*c-13 was involved in the regulation of the hair keratin genes. To test this hypothesis, an attempt to create transgenic mice that ectopically express *Hox*c-13 was undertaken. We hoped that the ectopic expression of *Hox*c-13 would cause an alteration in the normal mouse phenotype and more importantly direct the expression of hair keratin genes in tissues where they were not normally expressed.

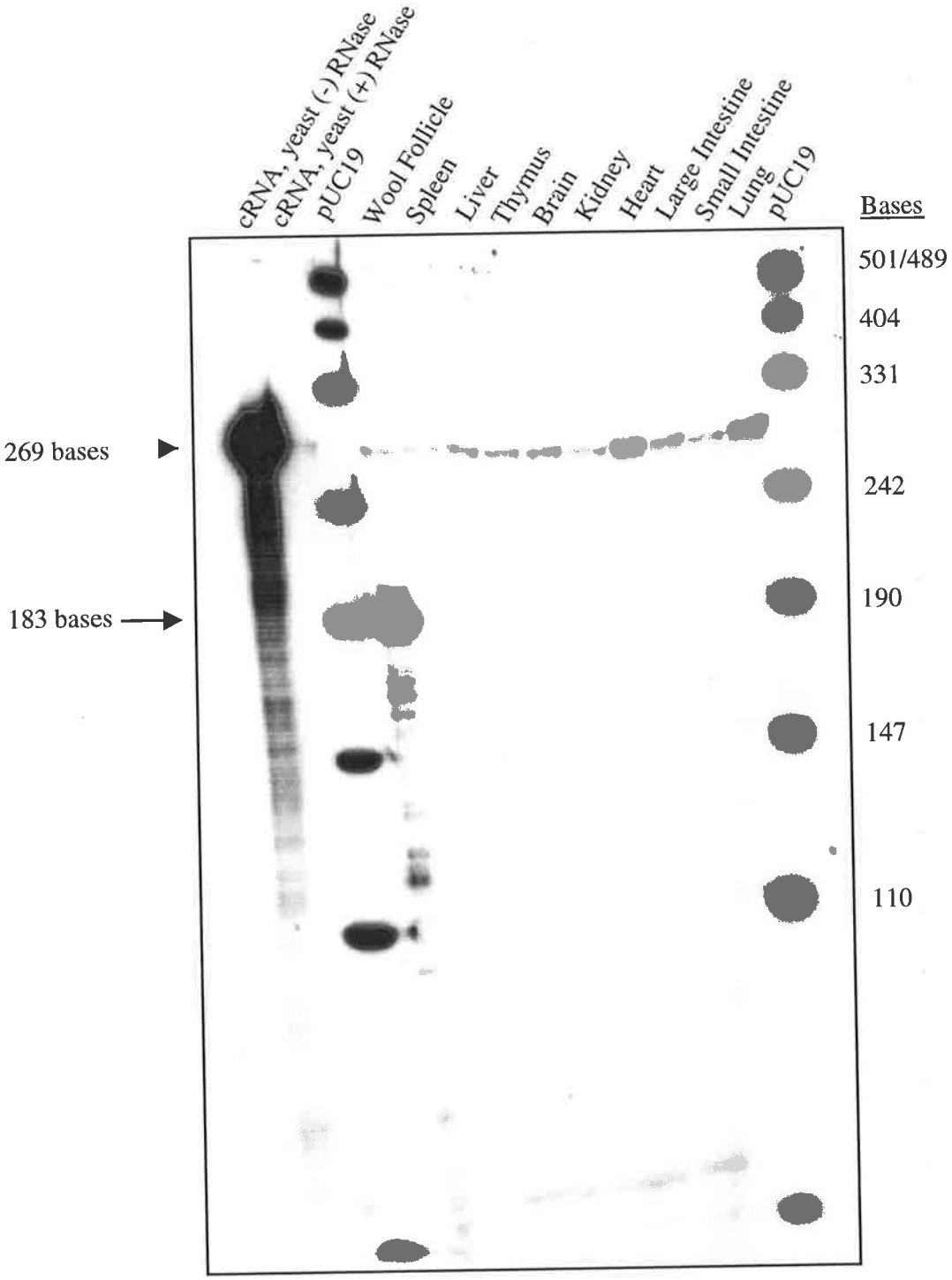
7.4.1 Construction of the *K14 / Hoxc-13* Transgene

To gain insight into the possible role of the *Hoxc-13* gene in follicle development, an attempt was made to ectopically express sheep *Hoxc-13* in mouse basal keratinocytes using the human keratin K14 promoter. The intermediate filament type I gene, *K14*, was expressed in the epidermal basal layer (Vassar et al., 1989), a layer that does not normally express *Hoxc-13* in the adult sheep (see Fig 7.4). Hence, the *K14* promoter was ideal to drive the ectopic expression of *Hoxc-13* to the basal layer of the epidermis. The K14 promoter has previously been used to drive the ectopic expression of TGF-α and TNF-α in transgenic mice (Vassar and Fuchs, 1991; Cheng et al., 1992). Furthermore, the K5 promoter, the intermediate filament partner of K14 has also been used to drive the ectopic expression of another homeobox gene, *Dlx* (Morasso et al., 1996).

The first step in producing transgenic mice was to construct a *K14 / Hoxc-13* transgene (detailed in Fig. 7.11, Fig. 7.12 and Fig. 7.13). In the making of the transgene, the intention was to isolate the *Hoxc-13* gene from the genomic clone 9A and insert it

Figure 7.10 : RNA Protection Analysis of Ovine *Hoxc-13* gene Expression in Total RNA from Various Adult Sheep Tissues.

Ten μg of total RNA isolated from adult wool follicle, spleen, liver, thymus, brain, kidney, heart, large intestine, small intestine and lung tissue samples taken from a Corriedale sheep were analysed by RNA protection for the presence of ovine *Hoxc-13* gene transcripts using the 183 bp non-coding antisense probe derived from clone C24B1-NC (Appendix 2) that had been radiolabelled with α– ^{32}P- rUTP. The products of the protection reactions were fractionated in a denaturing 6% (19:1) urea-PAGE system and the gel exposed to a phosphor screen for 36 hours. The probe controls (yeast RNA +/- RNase) were co-fractionated with test samples on the gel. The position of the 183 base protected ovine *Hoxc-13* RNA detected in the wool follicle sample is indicated by the arrow on the left side of the figure. No protected products of this size were detected in any of the other tissue samples. In each lane residual undigested probe that contains insert and polylinker sequence is indicated by the small arrow at 269 bases. The sizes of the end radio-labelled α-^{32}P-dCTP HpaII-digested pUC19 fragments are indicated on the right side of the image. It should be noted that the location of tissue biopsies with respect to the whole organ was not recorded.


cRNA, yeast (-) RNase
cRNA, yeast (+) RNase
pUC19
Wool Follicle
Spleen
Liver
Thymus
Brain
Kidney
Heart
Large Intestine
Small Intestine
Lung
pUC19
Bases
501/489
404
331
242
190
147
110
269 bases
183 bases

downstream of the K14 promoter. The first step in the production of the *K14 / Hoxc-13* transgene was to subclone the entire *Hoxc-13* gene (from the genomic clone 9A) into the base vector, pGem 11Zf (Promega). An attempt to subclone the entire gene, contained within a 10.5 kb SalI / EcoRI fragment of clone 9A (Fig. 7.11 A), proved unsuccessful because no fragments were being inserted into the vector. At this stage it was assumed that the length of the DNA fragment (10.5 kb) was too long for complete replication by the bacterial strains used (ED8799 and XL1 Blue) or secondly, *Hox*c-13 was toxic to the bacterial strains.

To avoid the problem of length, the *Hoxc-13* gene was shortened by 2.3 kb by omitting part of the 5.3 kb intron during subcloning of the *Hoxc-13* gene. This was achieved by joining two genomic fragments 1) a 1.6 kb XhoI / HindIII fragment, that contained the 5' end of *Hoxc-13* and 2) a 4.5 kb HindIII / EcoRI fragment, containing the 3' end of *Hoxc-13* (see Fig. 7.11 A for the location of both fragments in the genomic clone 9A). Initially, a 2.8 kb SalI / HindIII fragment located at the 5' end of clone 9A, that contained the 1.6 kb XhoI / HindIII fragment, was subcloned into the vector pBSc SK (Fig. 7.11 B; clone C13 S/H 2.8). Likewise, the 4.5 kb HindIII / EcoRI fragment located at the 3' end of clone 9A was subcloned into the vector pBSc SK (Fig. 7.11 C; clone C13 H/E 4.5). Both clones were digested with restriction enzymes to ensure that the correct fragments had been cloned (Fig. 7.11 D, E). A XhoI, SalI and HindIII digest of clone C13 S/H 2.8 produced DNA fragments of the predicted sizes of 3.0 kb (pBSc vector), 1.6 kb (exon 1 and part of the 5' end of the intron) and 1.2 kb (5' flanking sequence). A HindIII and EcoRI digest of clone C13 H/E 4.5 produced predicted DNA fragments of 3.0 kb (pBSc vector) and 4.5 kb (3' end of *Hoxc-13*).

The 1.6 kb XhoI / HindIII fragment was resected from clone C13 S/H 2.8 and joined to the 4.5 kb HindIII / EcoRI fragment, resected from clone C13 H/E 4.5, by *in vitro* ligation. Both fragments joined through their common HindIII site forming a 6.1 kb fragment (Fig. 7.12 A) that contained 12 nucleotides of the 5' untranslated region, the translation start codon, both exons, 3.03 kb of the intron, the 3' non-coding region (1.3 kb) and 0.8 kb of the 3' untranslated region. Four distinct products were detected in the ligation reaction, three of which were not required. The 9.0 kb fragment was formed by the 4.5 kb

Figure 7.11 : Subcloning Fragments of the Genomic Clone 9A.

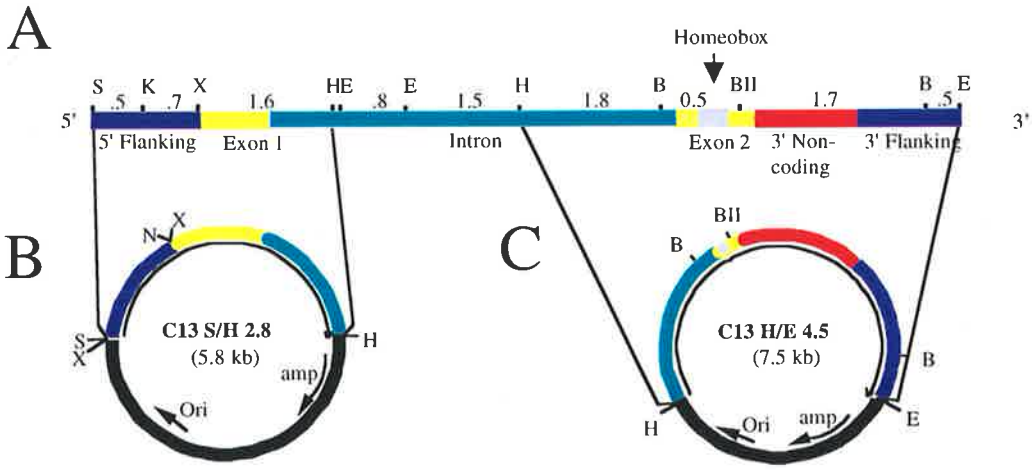
A. Partial schematic of the lambda clone 9A containing the ovine *Hoxc-13* gene.

B. Plasmid map of clone C13 S / H 2.8. A 2.8 kb SalI / HindIII fragment from lambda clone 9A containing 5' flanking region, 5' untranslated region, the first exon and part of the first intron was subcloned into the SalI / HindIII sites of pBSc (KS).

C. Plasmid map of clone C13 H / E 4.5. A 4.5 kb HindIII / EcoRI fragment from lambda clone 9A containing the 3' end of the intron, exon 2 and 3' non-coding region and 3' flanking region was directionally cloned into the HindIII and EcoRI sites of pBSc (KS).

D. Restriction digest of clone C13 S/H 2.8. Triple-digested (H = HindIII, S = SalI and X = XhoI) DNA was size-separated through a 1 X TAE 1 % agarose gel. Predicted fragments of 3.0 kb (vector pBSc KS), 1.6 kb (XhoI / HindIII fragment) and 1.2 kb (SalI / XhoI fragent) were detected.

E. Restriction digest of clone C13 H/E 4.5. Double-digested (E = EcoRI, H = HindIII) DNA was size-separated through a 1 X TAE 1 % agarose gel. Predicted fragments of 3.0 kb (vector pBSc KS) and 4.5 kb (HindIII, EcoRI fragment).


A
Homeobox
S .5 K .7 X 1.6 HE .8 E 1.5 H 1.8 B 0.5 BII 1.7 B .5 E
5'
3'
5' Flanking
Exon 1
Intron
Exon 2
3' Non-coding
3' Flanking
B
N
X
S
X
H
C13 S/H 2.8
(5.8 kb)
amp
Ori
C
BII
B
C13 H/E 4.5
(7.5 kb)
B
H
Ori
amp
E

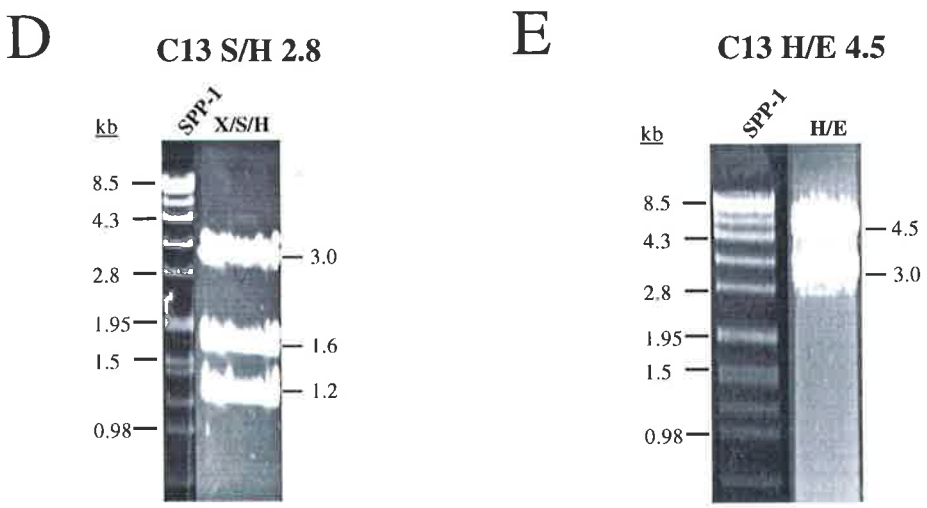

D
C13 S/H 2.8
kb
SPP-1
X/S/H
8.5
4.3
2.8
1.95
1.5
0.98
3.0
1.6
1.2
E
C13 H/E 4.5
kb
SPP-1
H/E
8.5
4.3
2.8
1.95
1.5
0.98
4.5
3.0

Figure 7.12 : *In vitro* Ligation of the 5' and 3' ends of *Hoxc-13*.

A. Products of a ligation reaction between the two fragments 1) 1.6 kb XhoI / HindIII fragment resected from clone C13 S/H 2.8 (Fig. 7.11 B) and 2) 4.5 kb HindIII / EcoRI fragment resected from clone C13 H/E 4.5 (Fig. 7.11 C) were size-separated through a 1 X TAE 1 % agarose gel (lane 2). Both fragments joined through their common HindIII site and the subsequent 6.1 kb fragment is indicated by the arrowhead on the right of the gel. Excess fragments in the ligation reaction can be seen in lane 2 at 9.0 kb (4.5 kb HindIII / EcoRI fragment joined to itself), at 3.2 kb (1.6 kb XhoI / HindIII fragment joined to itself) and at 1.6 kb (1.6 kb XhoI / HindIII fragment).

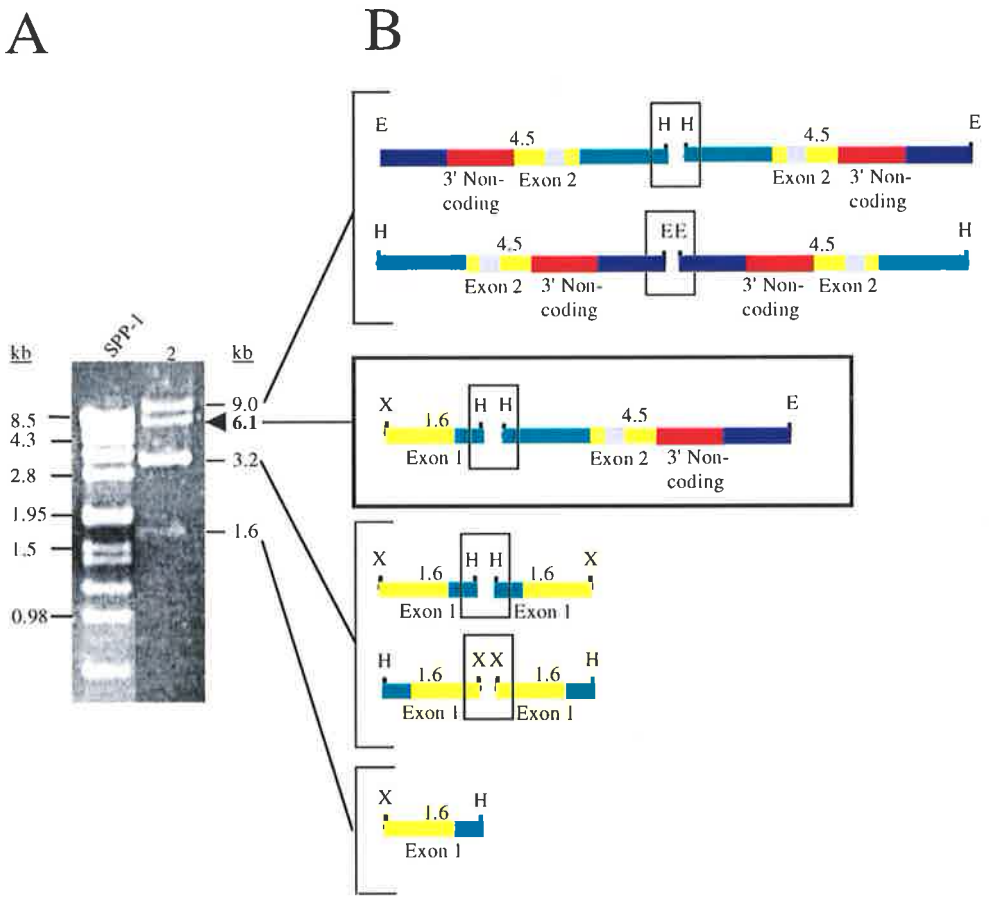
B. Cartoon showing the possible combinations of the two fragments in the ligation reaction of panel A. The 6.1 kb transgene is boxed.

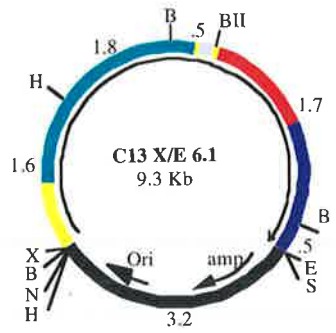
C. Plasmid map of clone C13 X/E 6.1 containing the shortened *Hoxc-13* gene (6.1 kb). Location of restriction enzyme sites are indicated.

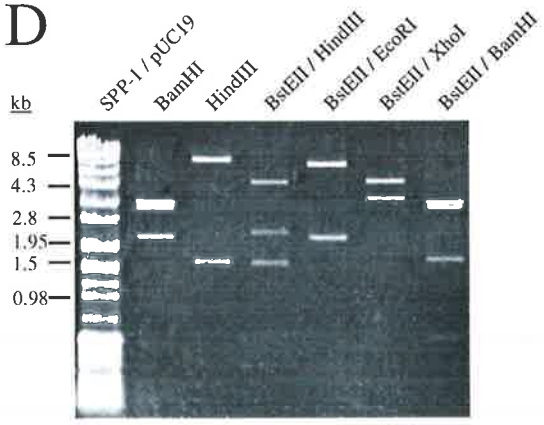
D. Single and multiple restriction digests of clone C13 X/E 6.1. Cleaved DNA was size-separated through a 1 X TAE 1 % agarose gel. Restriction enzymes for each reaction are indicated above each lane. Predicted fragments of 1.6 kb and 7.7 kb were detected for the HindIII digest verifying that the correct fragment was cloned.

B = BamHI, BII = BstEII, E = EcoRI, H = HindIII, K = KpnI, N = NsiI, S = SalI,

X = XhoI.


A
B
kb
SPP-1
2
kb
8.5
4.3
2.8
1.95
1.5
0.98
9.0
6.1
3.2
1.6
E
4.5
H H
4.5
E
3' Non-
coding
Exon 2
Exon 2
3' Non-
coding
H
4.5
EE
4.5
H
Exon 2
3' Non-
coding
3' Non-
coding
Exon 2
X
1.6
H H
4.5
E
Exon 1
Exon 2
3' Non-
coding
X
1.6
H H
1.6
X
Exon 1
Exon 1
H
1.6
XX
1.6
H
Exon 1
Exon 1
X
1.6
H
Exon 1


C
B
.5
BII
1.8
H
1.7
1.6
C13 X/E 6.1
9.3 Kb
B
.5
E
S
X
B
N
H
Ori
amp
3.2


D
SPP-1 / pUC19
BamHI
HindIII
BstEII / HindIII
BstEII / EcoRI
BstEII / XhoI
BstEII / BamHI
kb
8.5
4.3
2.8
1.95
1.5
0.98

HindIII / EcoRI fragment ligating to itself at both the HindIII and EcoRI sites (Fig 7.12 B). A band at 1.6 kb represented excess 1.6 kb XhoI / HindIII fragment that was not involved in the ligation reaction. Some of the 1.6 kb XhoI / HindIII fragment ligated to itself resulting in a band at 3.2 kb.

The 6.1 kb fragment was excised and purified from low melting point agarose and directionally cloned into the XhoI and EcoRI sites of the vector pGem 11Zf (clone C13 X/E 6.1; Fig. 7.12 C, D). Successful cloning of the 6.1 kb fragment was verified by digestion of the DNA of clone C13 X/E with restriction enzymes that yielded predicted DNA fragments (Fig. 7.12 D). For example, BamHI digestion of the plasmid DNA yielded the correct fragments of 2.2 kb, 3.4 kb and 3.7 kb.

To complete the production of the *K14 / Hoxc-13* transgene, the shortened 6.1 kb *Hoxc-13* gene fragment was inserted downstream of the *K14* promoter. Initial attempts to subclone the shortened 6.1 kb *Hoxc-13* gene fragment into the XhoI site of the K14 plasmid (see Fig. 7.13 A, for plasmid map) by standard cloning techniques failed. This may have been due to the combined length of the *K14 promoter / β - globin* intron (2.7 kb) and the *Hoxc-13* gene fragment (6.1 kb) being too large for stable bacterial replication. This problem was overcome by using the method of '*in vitro* ligation' which enabled the transgene to be produced without the need for stable bacterial replication of the full-length transgene. Using this technique, the 2.7 kb *K14 promoter / β - globin* intron fragment was joined to the shortened 6.1 kb *Hoxc-13* gene fragment (Fig. 7.13 D). The resultant *K14 / β - globin / Hoxc-13* transgene (8.8 kb) was excised from a low melting point agarose gel and purified. To ensure that the 8.8 kb *in vitro* ligated product was isolated, the DNA of the purified transgene was size-separated through a 1% agarose gel (Fig. 7.13 E) and secondly, digested with the restriction enzyme XhoI to yield predicted DNA fragments of 2.7 kb and 6.1 kb (Fig. 7.13 F).

Furthermore, the purified transgene DNA was sequenced using the primer C13IN682, the same primer used to determine the transcriptional start point (see Fig. 6.10 for location of primer). The primer, C13IN682, binds downstream of both the XhoI ligation site and the *Hoxc-13* translation start codon. Sequence data that spanned the XhoI ligation

Figure 7.13 : Construction of the *K14 / Hoxc-13* Transgene.

A. Plasmid map of clone K14. Clone K14 contains the mouse *K14* promoter and human *β-globin* intron fragment (Vassar et al., 1989). The cassette was a gift from Professor Elaine Fuchs.

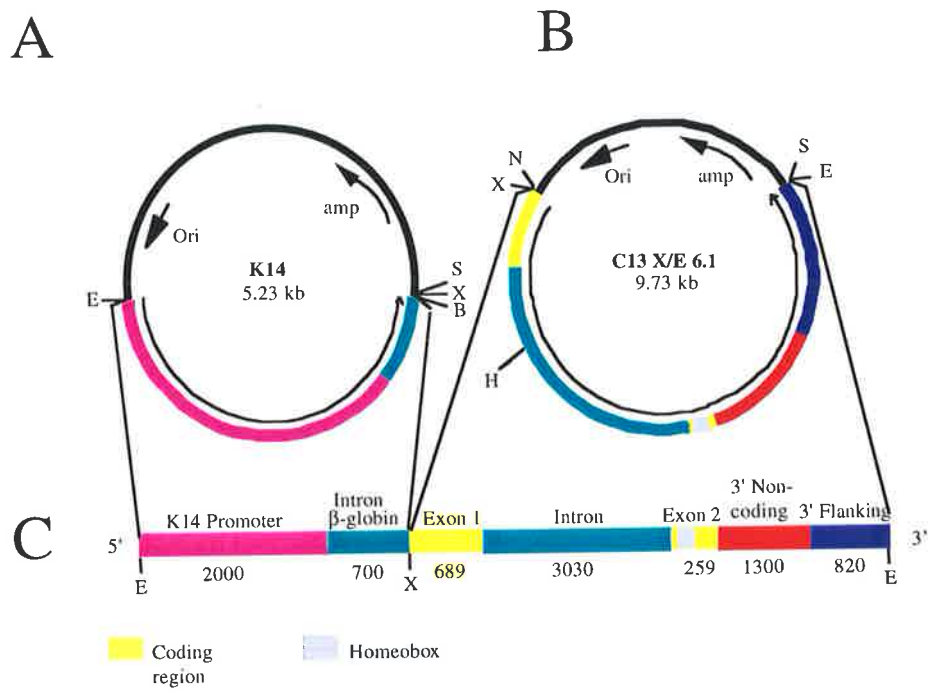
B. Plasmid map of clone C13 X/E 6.1. Location of restriction enzyme sites are indicated.

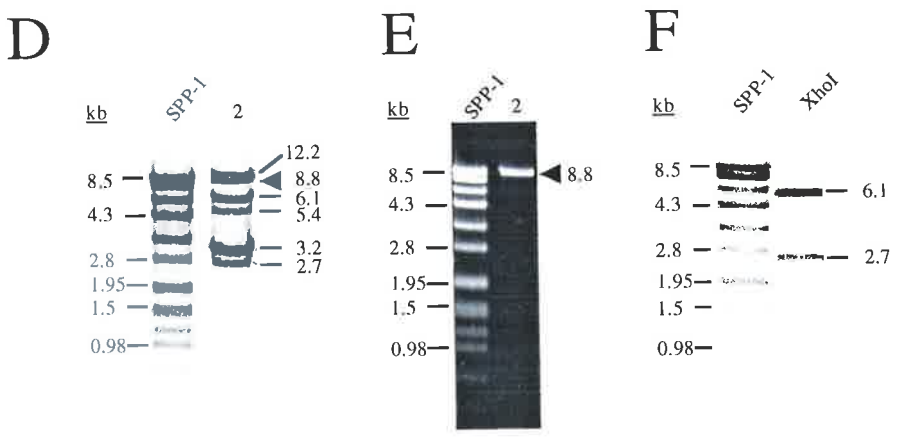
C. Schematic of the *K14 / Hoxc-13* transgene (not drawn to scale). The transgene was created in the following steps: 1) Clone C13 X/E 6.1 was cleaved with NsiI then end-filled using Klenow fragment to reduce further ligation at this site. 2) Clone K14 was linearised with SfiI then end-filled using Klenow fragment to reduce further ligation at this site. 3) Both linearised fragments were pooled, cleaved with XhoI, and then ligated *in vitro* through the common XhoI site. 4) The pGem 11Zf vector arms were excised with Eco R1 and the resultant 8.8 kb *K14 / Hoxc-13* fragment was purified from low melting point agarose.

D. Photograph of a 1% agarose gel containing, in lane 2, products of the *in vitro* ligation reaction described in **C.** The position of the *K14 / Hoxc-13* transgene (8.8 kb) is indicated by the arrowhead at the right of the figure. Excess fragments that did not participate in the ligation reaction can be seen in lane 2 at 2.7 kb (K14 promoter), at 3.2 kb (vector-11Zf) and at 6.1 kb (Hoxc-13 X/E fragment). Two bands at 12.2 kb and 5.4 kb were produced by ligation of the 6.1 kb C13 X/E fragment and 2.7 kb K14 promoter to themselves respectively through the XhoI site.

E. Photograph of a 1% agarose gel containing, in lane 2, the purified 8.8 kb *K14 / Hoxc-13* transgene that was injected into mouse blastocysts in an attempt to produce transgenic mice.

F. XhoI restriction digest of the 8.8 kb *K14 / Hoxc-13* transgene DNA. Cleaved DNA was size-separated through a 1 X TAE 1 % agarose gel. Two predicted fragments of 6.1 kb and 2.7 kb were detected. In relevant panels E = EcoRI, H = HindIII, N = NsiI, S = SfiI, X = XhoI. B = BamHI.


A
B
C
Ori
amp
K14
5.23 kb
E
S
X
B
N
X
Ori
amp
S
E
C13 X/E 6.1
9.73 kb
H
5'
K14 Promoter
Intron
β-globin
Exon 1
Intron
Exon 2
3' Non-
coding
3' Flanking
3'
E
2000
700
X
689
3030
259
1300
820
E
Coding
region
Homeobox


D
kb
SPP-1
2
8.5
4.3
2.8
1.95
1.5
0.98
12.2
8.8
6.1
5.4
3.2
2.7
E
kb
SPP-1
2
8.5
4.3
2.8
1.95
1.5
0.98
8.8
F
kb
SPP-1
XhoI
8.5
4.3
2.8
1.95
1.5
0.98
6.1
2.7

site confirmed that both the *K14 promoter / β - globin* intron fragment and the *Hoxc-13* gene fragment had joined at the Xho I site (data not shown).

7.4.2 Production of Transgenic Mice

Transgenic mice production was contracted out to the laboratory of Dr Alan Fawcett and performed interstate at CSIRO, Division of Animal Production, Prospect, Sydney, New South Wales. The purified *K14 / Hoxc-13* 8.8 kb transgene, prepared as detailed in Section 7.4.1, was resuspended in 10mM NaCl, 1mM Tris and 0.1mM EDTA and injected into mouse blastocysts at a concentration of 5μg / ml. The transgene was initially tested for possible toxic effects to embryos by examining progression to the blastula stage after microinjection. The transgene was microinjected into the pronucleus of 76 embryos that were cultured *in vitro* for several days at 38°C / 5% CO_2 until they developed into a blastula. An acceptable blastocyst survival rate of 64% was obtained (data not shown).

To produce transgenic mice, the transgene was microinjected into the pronucleus of 552 embryos that were transferred to 27 pseudo-pregnant females. An overall pupping rate of 12.6% was obtained resulting in 67 pups being born. It should be noted that attempts in making transgenic mice with other investigators transgenes before and after the attempt with the *K14 / Hoxc-13*10 kb transgene yielded higher pupping rates ie. before (20.5% and 24.9%) and after (19.7%, 25.5% and 44.5%).

7.4.3 Detection of Transgenic Mice

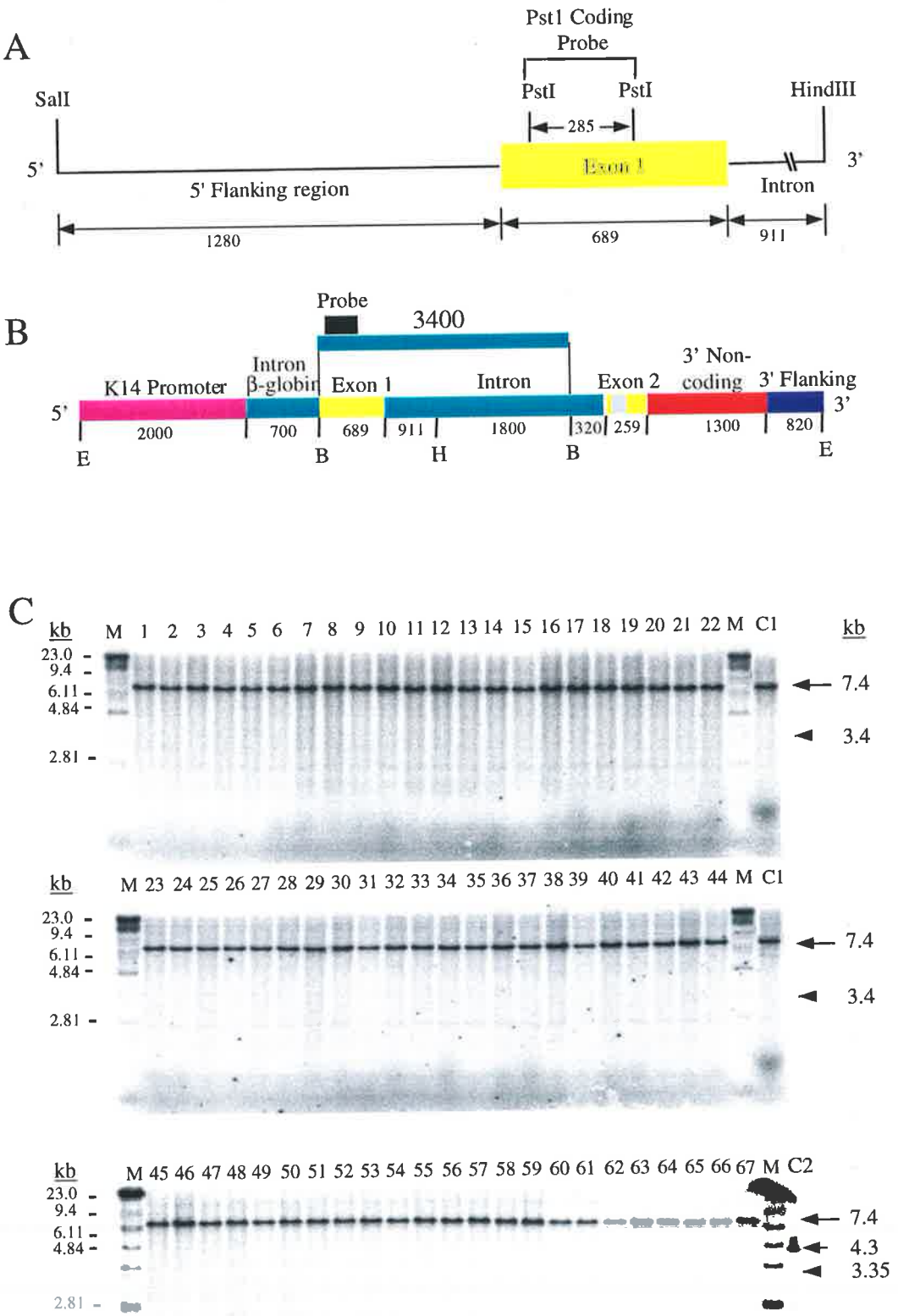
Southern blot analysis was used to test if pups carried the *K14 / Hoxc-13* transgene. Genomic DNA was isolated from the tails of 67 pups born, digested with BamHI and the Southern blot was probed with a 0.28 kb PstI *Hoxc-13* coding fragment (Fig. 7.14 A). A 7.4 kb BamHI fragment from the endogenous mouse *Hoxc-13* gene, predicted from the Southern blot of mouse genomic DNA shown in Chapter 6 (Fig. 6.13), was detected in each lane of the Southern blot (Fig. 7.14 C). However, the predicted 3.4 kb BamHI fragment, located only in the transgene (Fig. 7.14 B), was not detected in any lane. Thus, none of the pups carried the *K14 / B-Globin / Hoxc-13* construct.

Figure 7.14 : Southern Blot Analysis of Mice to Detect the *K14 / Hoxc-13* Transgene.

A. Schematic of the 2.8 kb SalI / HindIII fragment from clone 9A showing the location of the 285 bp PstI coding fragment (located in exon 1) used as a radio-labelled probe in the Southern blot analysis (panel C) of mouse genomic DNA to detect the transgene. See Appendix 2 C for the plasmid map of clone C13-COD that contains the 285 bp PstI coding fragment. Schematic is not drawn to scale.

B. Schematic of the *K14 / Hoxc-13* transgene showing the location of the restriction enzymes and size of various regions of the transgene. The location of the 3.4 kb BamHI fragment, present only in the transgene, is indicated by the green bar. The location of the probe (285 bp PstI coding fragment) that binds to the 3.4 kb BamHI fragment is indicated by the black bar. Note that 2.2 kb of the intron between the BamHI sites was deleted from the transgene and is denoted by the '//'. Schematic is not drawn to scale.

C. Genomic DNA isolated from the tails of 67 mouse pups was cleaved with BamHI and analysed by Southern blot to test if the *K14 / Hoxc-13* transgene was present. A 7.4 kb BamHI fragment from the endogenous *Hoxc-13* gene was detected in each lane however the 3.4 kb BamHI fragment from the transgene was not detected indicating that none of the pups were transgenic. M = Lambda Hi-Lo markers. C1 = 10 µg of non-transgenic mouse genomic DNA cleaved with BamHI. C2 = *K14 / Hoxc-13* transgene cleaved with HindIII. A predicted fragment of 4.3 kb is detected by the PstI coding probe.


A
Pst1 Coding Probe
PstI
PstI
HindIII
285
SalI
5'
3'
Exon 1
5' Flanking region
Intron
1280
689
911
B
Probe
3400
Intron β-globin
K14 Promoter
Exon 1
Intron
Exon 2
3' Non-coding
3' Flanking
5'
3'
2000
700
689
911
1800
320
259
1300
820
E
B
H
B
E
C
kb
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 M C1
23.0
9.4
6.11
4.84
2.81
7.4
3.4
kb
M 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 M C1
23.0
9.4
6.11
4.84
2.81
7.4
3.4
kb
M 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 M C2
23.0
9.4
6.11
4.84
2.81
7.4
4.3
3.35

7.5 Discussion

This Chapter has described the analysis of *Hoxc-13* expression in sheep embryonic and adult wool follicles by *in situ* hybridisation and secondly, an attempt to produce transgenic mice to gain insight into the function of *Hox*c-13.

7.5.1 : *Hoxc-13* Expression in Sheep skin and Comparison with other *Hox* Genes

During embryonic development, *Hoxc-13* expression varies differentially and spatially in both the epidermis and developing placode. In the epidermis, ovine *Hoxc-13* was expressed evenly thoughout the basal and suprabasal layers of embryonic skin at E51, E65, E80, E99 and E135 but was not expressed in adult epidermis. Expression of other homeobox genes in the developing epidermis has been reported by other investigators (Mathews et al., 1993; Kanzler et al., 1994; Chuong et al., 1996; Stelnicki et al., 1997; Stelnicki et al., 1998; Jiang et al., 1999). Two *Abd*-B genes, *Hoxd-9* and *Hoxd-11*, show some similar expression patterns in embryonic skin at comparable time points as *Hoxc-13* although their expression was restricted to the mouse caudal skin (Kanzler et al., 1994; Chuong, 1998). Expression of both genes was first detected in the epidermis at E14.5 and reached a peak level at E16.5. At E18.5, *Hoxd-9* was expressed evenly throughout the epidermis whereas *Hoxd-11* was predominant in the epidermal basal layer. At two and four days after birth both *Hoxd-9* and *Hoxd-11* were expressed predominantly in the epidermal basal layer. *Hoxd-13*, another member of the *Abd-B* class and in the same paralogous group as *Hoxc-13*, was not expressed in the developing murine epidermis during the same time points.

The homeobox genes *HOX*A4, *HOX*A5, *HOX*A7, *HOX*C4, *HOX*B4, and *HOX*B7 are also expressed in the epidermis during human embryonic skin development (Stelnicki et al., 1998). Although expression levels differ at any one embryonic time point, these genes show similar temporal and spatial expression patterns to each other in the epidermis. Their expression was restricted to the epidermal basal layer during early skin development, spread to the suprabasal layers during mid to late second trimester development and was then downregulated and restricted to the epidermal suprabasal layers in newborn and adult skin.

This was in contrast to ovine *Hoxc-13* which was expressed uniformly throughout embryonic epidermis but was not expressed in the adult epidermis (Stelnicki et al., 1998).

Hoxc-13 expression was also detected transiently in the dermal condensate at initial stages of follicle placode development at E65 and E81. Interestingly, in the chick embryo, the paralogous gene, *Hoxd-13*, is expressed in dermal cells during plantar reticula morphogenesis however it was not expressed once the plantar reticula were formed (Kanzler et al., 1997, Chuong, 1998). Furthermore, *Hoxd-13* was expressed in the perichondrial mesenchyme and plantar skin in the foot of chick (E10.5) and mouse (E17.5) embryos, (Kanzler et al., 1997, Chuong, 1998). Lastly, ovine *Hoxc-13* expression, like the human homeobox genes, could not be detected in the initial stages of follicle placode development in the mesenchymal aggregate. This suggests that *Hoxc-13* is not involved in follicle initiation.

Hoxc-13 was expressed in the epidermal-derived cells of the developing placode, a similarity that echoes the expression of the homeobox genes *Hoxc-8*, *Hoxd-9* and *Hoxd-11* in the mouse (Kanzler, et al., 1994). Like *Hoxc-13*, all three genes were expressed in the epithelial cells of the developing placode. Varying slightly, *Hoxc-8* expression was also detected in the dermal papilla, a region that does not express *Hoxc-13*. Two days after birth, *Hoxd-9* and *Hoxd-11* continued to be expressed in all the epithelial cells of the placode. At this time, *Hoxd-13* expression was confined to the hair matrix epithelial cells where presumably *Hoxc-13* is expressed at this time point in mice. Four days after birth, both *Hoxd-9* and *Hoxd-11* were expressed in the uppermost outer root sheath and the lower bulb, a region that appears to overlap the proximal limits of *Hox-c13* expression. At this point in time, *Hoxd-9*, *Hoxd-11* and *Hoxd-13* were not detected in the hair matrix cells above the dermal papilla where *Hoxc-13* is predominantly expressed. Hence, *Hoxc-13* appears to have a unique expression pattern in the developing follicle amongst the known *Hox* genes reported to date.

In sheep, once vibrissa and pelage follicles had developed to a stage where a fibre was produced, *Hoxc-13* expression was detected in the follicle bulb and lower shaft. A clear comparison can be made between the timing of the formation of the fibre in developing wool follicles and the transcriptional activation of *Hoxc-13* in the bulb. Production of the

keratinised fibre begins at approximately E99 for sheep pelage follicles and at approximately E81 in vibrissae that are at an F4 stage of development (Hardy and Lyne, 1956) the same time that *Hoxc-13* was expressed in the follicle bulb and lower shaft. Little is known about when the hair keratin genes are transcriptionally activated in embryonic sheep follicles but many keratin gene families are activated during the formation of the fibre. Transcription factors such as LEF-1, Sp1, AP2-like and NF1-like proteins have been shown to bind to the promoter of the intermediate filament gene K2.10 (Dunn et al., 1998). The expression data for *Hoxc-13* suggests that it could be involved directly or indirectly in the regulation of the differentiation-specific genes of the wool fibre, as early as E81 in whisker and E99 in sheep pelage follicles.

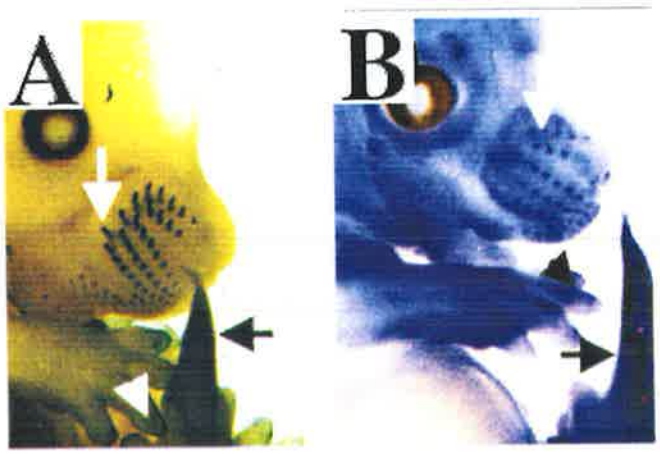
In adult follicles, the extent of *Hoxc-13* expression in the bulb and lower shaft was compared with the expression of *histoneH3* (Chou et al., 1990) a marker of cellular proliferation and the keratin type II intermediate filament gene K2.10, a marker of cellular differentiation (Powell et al., 1992). *Hoxc-13* expression was detected in some of the proliferating cells surrounding the apex of the dermal papilla which also express *histoneH3*. Strongest *Hoxc-13* expression was seen in cells of the upper bulb and lower shaft. In the lower follicle shaft, expression was detected in differentiating cells that express *K2.10*. Hence, *Hoxc-13* expression encompasses the region where keratinocytes are either proliferating or differentiating. This suggests that *Hox*c-13 may play a role in the transitionary state between proliferation and differentiation for hair keratinocytes located in the bulb. This is further discussed in the concluding Discussion in the following Chapter.

During the course of this work, *Hox*c-13 expression was reported in the mouse hair follicle (Godwin and Capecchi, 1998). In their experiments, Godwin and Capecchi, (1998) inserted lacZ into the homeodomain so that a non-functional protein was synthesised whose location could be detected by enzymatic assay. The knockout mice developed follicles that showed that *Hox*c-13 was not required for follicle initiation or development. However, the follicles produced hair fibres that were brittle and broke off at the skin surface resulting in a hair-less phenotype suggesting that *Hoxc-13* plays an important role in hair fibre formation.

Godwin and Capecchi (1998) reported *Hoxc-13* expression, at both the RNA and protein level, in similar regions of the follicle and tongue as reported here. They found

Hoxc-13 protein in the hair bulb and companion layer, the layer of the outer root sheath next to the inner root sheath, but not in the epidermis or other outer root sheath layers. In the present study, ovine *Hoxc-13* expression was detected in the companion layer of sheep pelage and vibrissa follicles and in the basal and suprabasal cells of the outer root sheath layer, the developing epidermis and placode. The difference in *Hoxc-13* expression between sheep and mouse may be species-specific. It should be noted that in sheep, it is not known if the *Hoxc-13* mRNA is translated in the outer root sheath suprabasal layer or during the development of the epidermis and placode. The expression of *Hoxc-13* mRNA in these regions was not reported in the knockout mouse. However, on closer inspection of the *Hoxc-13* RNA *in situ* hybridisation analysis by Godwin and Capecchi (1998) there seems to be a mRNA signal in the E13.5 skin of the forelimbs and head region of wild type mouse embryos (Fig. 7.15).

Fig. 7.15 : Photographs of E13.5 mouse embryos. The mouse embryo in panel A contains a non-functional Hoxc-13 protein. Hoxc-13 protein expression was monitored by β-galctosidase staining of heterozygous *Hoxc-13*^{lacZ} embryos (panel A) and mRNA expression by *in situ* hybridisation in wild-type embryos (panel B). Note the blue staining in panel B that indicates mRNA expression, in the skin of the forelimbs, shoulders, neck and face. White arrow indicates expression in the vibrissae. Black arrow indicates expression in the tail and black arrowhead indicates faint expression in the nails. Reprinted from Godwin and Capecchi (1998).



The authors dismiss this expression, claiming a low signal to noise ratio, yet the forelimb signal is stronger than that accepted as real around the vibrissa. The data presented here support *Hoxc-13* expression in embryonic epidermis. Because they focussed on protein detection, Godwin and Capecchi (1998) may have overlooked potential expression of *Hoxc-13* at the RNA level during the development of the epidermis, placode and in the outer root sheath suprabasal layer. It is possible that *Hoxc-13* is transcribed in these regions but not translated. This result emphasizes the need to analyse the expression of homeobox genes at both the RNA and protein level. Immunolocalisation analysis of *Hox*c-13 in embryonic sheep skin sections may determine if the protein is expressed in these regions. Unfortunately, a specific antibody, was not available during the course of this thesis for such a study.

In the present study, *Hoxc-13* expression was also detected in the base of the filiform papilla of the tongue. Filiform papilla contain a hair-like compartment which express hair-type keratin proteins (Dhouailly et al., 1989; Tobiasch et al., 1992; Winter et al., 1994). *Hoxc-13* was not expressed in other tissues surveyed such as the large intestine, small intestine, liver, lung, skeletal muscle, brain, thymus, heart, kidney, esophagus or spleen. Although not an exhaustive survey, the finding that *Hoxc-13* was expressed in only the tongue and the hair follicle suggests that it might have been "recruited" to these anatomically similar structures for a similar role. If *Hoxc-13* does regulate the expression of hair keratin proteins, those keratins expressed in both the follicle and tongue are prime candidates.

Expression of *Hoxc-13* in both the tongue and vibrissae, two of the most anterior locations of the sheep, breaks the rule of spatial colinearity for the first time (Duboule, 1998). The rule states that *Hox* genes located at the 5' end of the cluster are transcribed in more caudal areas than the *Hox* genes located at the 3' end. Therefore expression of the *Abd-b* type genes that are located at the 5' end of the cluster, would be expected to be restricted to follicles located at the posterior region of the sheep (reviewed in Duboule, 1994). The expression of other cluster homeobox genes such as *Hoxd-9*, *Hoxd-11* and *Hoxd-13*, which also belong to the *Abdominal B* group appear to follow this rule. Their

expression is limited to the posterior regions and hence are confined to epidermis and follicles located in the most caudal regions (Kanzler, 1994).

The expression of *Hoxc-13* in unidentified cells of the dermis and vibrissae blood sinus also breaks the spatial colinearity rule (Fig. 7.3 E and F). These cells may be fibroblastic or hematopoietic, however conclusive identification requires the use of molecular markers. *Hox* gene expression in hematopoietic cells has been reported (Lawrence et al., 1996, Zimmerman and Rich, 1997) but a clear role for *Hox* gene function in these cells has not been determined. Interestingly, *HOXC4* which also exhibits spatial and temporal changes during human embryonic skin and follicle development is expressed in fibroblasts and occasional lymphocytes in the perivascular dermis (Rieger et al., 1994).

7.5.2 : *K14 / Hoxc-13* Transgene Production.

To investigate the role of *Hoxc*-13 we attempted to produce transgenic mice expressing a *K14 / Hoxc-13* transgene. The K14 promoter was chosen because it is expressed in the epidermal basal layer (Vassar et al., 1989) a region that does not express *Hoxc-13* in the adult epidermis. To investigate if *Hoxc-13* regulates, directly or indirectly, the hair keratin genes, an ideal cell line was one that was similar to hair keratinocytes but that did not express hair keratins. Epidermal keratinocytes are functionally and morphologically similar to hair keratinocytes and do not express hair-specific keratins. They were therefore a logical choice for the ectopic expression of *Hoxc-13*.

The construction of the *K14 / Hoxc-13* transgene proved difficult and attempts to transform plasmids containing the full-length transgene into the bacterial strains ED8799 and XL1 Blue proved unsuccessful. The method of '*in vitro*' ligation was used to overcome problems of stable bacterial replication. Interestingly, several months after the attempt to produce transgenic mice, the full-length transgene was cloned using the vector PBR322 and the bacterial strain ABLE K (Stratagene) which lowers the plasmid copy number to 10 per bacterium to avoid toxicity problems.

Unfortunately, of 67 mice born, there were no transgenics and a second attempt could not be undertaken due to budget and PhD time limitations. It should be noted that transgenic mice were produced with various transgenes from different researchers by the

microinjection facility at CSIRO, Division of Animal Production, in Sydney. A pupping rate of approximately 20% before and after the microinjection of the *K14 / Hoxc-13* transgene was obtained, suggesting that there were no routine technical problems. However the lower pupping rate obtained with the *K14 / Hoxc-13* transgene suggests that cannibalisation by the mother may have occurred. Cannibalisation is known to occur if pups have an abnormal phenotype. In the present study, it was not known whether abnormal transgenic mice were born which may have been cannibalized by the mother because they were first checked two days after birth. Another possibility that cannot be discounted is that there may have been resorption of abnormal transgenic fetuses during pregnancy. In a repeat experiment, pups at various time points could be taken prior to birth and tested by RT-PCR or Southern Blot to see if they carry the *K14 / Hoxc-13* transgene.

Interestingly, transgenic mice ectopically expressing a similar type of transgene *K5 / Dlx3* in the epidermal basal cell layer lead to an abnormal epidermal phenotype and perinatal lethality (Morasso *et al*, 1996). In that case the expression of the homeobox gene, *Dlx3*, was driven by the keratin *K5* promoter that targets the same tissues as the *K14* promoter. This implies that the *K14 / Hoxc-13* transgene may cause perinatal death.

Chapter 8

Concluding Discussion

Chapter 8: Concluding Discussion

8.1 Introduction

This thesis has described the results obtained from screening the sheep wool follicle for expressed *Antennapedia* -like homeobox genes and secondly the isolation and characterisation of two genes, *Hoxc-13*, related to the *Drosophila* homeotic gene *Abdominal-B* and *Barx2*, that belongs to the *Bar* class.

Prior to the initiation of this project no one had undertaken a dedicated search for homeobox genes in the hair follicle and only four homeobox genes were known to be expressed in the hair follicle. Two were *Antennapedia*-type homeobox genes, *Hoxb-2*, expressed in the murine hair follicle placode (Whiting et al., 1991) and *Hoxc-8* expressed in the dermal papilla of the anagen hair follicle (Bieberich et al., 1991). The other two homeobox genes were members of the POU family, *Skn-1a* and *Skn-1I*, expressed in the cortical cells of mouse anagen hair follicles (Andersen et al., 1993b).

Reports of homeobox expression during skin and follicle development have increased over recent years. Several reports have speculated on the function of the homeobox genes in the follicle, however, no study has explained why there are so many homeobox genes activated in a temporally and spatially restricted manner during both embryonic skin and follicle development. Furthermore, the genes that are targeted by the homeoproteins in the skin are mostly unknown.

The main aim of this Chapter is to summarise the thesis results, relate them to the wider fields of follicle and homeobox gene biology and discuss future directions of study from the work presented.

8.2 Isolation and Characterisation of Homeobox cDNA Clones from the Wool Follicle

The first aim of the thesis was to isolate homeobox cDNA clones from the anagen wool follicle and characterise their expression pattern in the wool follicle. Given that two *Antennapedia*-type homeobox genes had been reported before this project had started and that vertebrates were known to contain at least 38 *Antennapedia*-type homeobox genes, it was logical to first screen the wool follicle for the expression of more of these genes. In the RT-PCR screen of the wool follicle using degenerate primers to the ends of the *Antennapedia* -type homeobox, nine homeobox genes, *Hoxa-4*, *Hoxb-2*, *Hoxb-6*, *Hoxb-8*, *Hoxb-9*, *Hoxc-4*, *Hoxc-5*, *Hoxc-9*, *Hoxc-13* and a novel homeobox gene *Barx2*, which belonged to the *Bar* family were isolated. During the cloning of the partial homeobox cDNAs, more than 200 colonies with potential Hox gene inserts were not analysed. Only 20 clones were sequenced of which ten different Hox genes were located. This suggests that there may be other Hox genes, apart from the genes mentioned so far, that are expressed in the follicle that have yet to be isolated. A question arises then, that if ten genes were found from sequencing only 20 clones, how many more genes would be found by sequencing the 200 clones isolated? Another point to consider is that although the degenerate primers used in the RT-PCR screen were designed initially to detect *Antennapedia*-type genes, they were able to amplify the *Barx2* gene, a gene that does not reside within the Hox cluster. This implies that divergent homeobox genes can be amplified using the degenerate primers described, as long as the conserved motifs of WFQNRR and ELEKEF reside in their homeodomain.

Expression of the ten isolated homeobox genes in the wool follicle was analysed by a combination of Northern blot and RNA protection. Northern blot analysis detected transcripts for only two of the eight genes studied, ovine *Hoxc-13*, which produced a 2.6 kb

RNA transcript and ovine *Barx2* which produced a 2.1 kb RNA transcript. Using the more sensitive method of RNA protection, expression was confirmed for all of the homeobox genes except *Hoxc-5* and *Hoxa-4*. During the course of this project, *HOXA4* was reported by others to be expressed in very low levels in the cortex of follicles (Stelnicki et al, 1998). It should be noted that *Hoxc-4* was not analysed by RNA protection in this thesis due to an oversight during the initial analysis.

Over the last few years, several groups have reported the expression of *Hoxb-2*, *Hoxb-6*, *Hoxc-4*, *HOXA4*, *HOXC4* and *Hoxc-13* in follicles (Table 1). To date no one has studied the expression of *Hoxb-8*, *Hoxb-9*, *Hoxc-5*, *Hoxc-9* and *Barx2* during follicle development. Both *Hoxb-8* and *Hoxb-9* expression have been detected by RNA protection analysis of whole skin extracts from E12 and E17 fetal murine skin (Detmer et al., 1993). It is not known if these genes are expressed in the follicle, dermis or epidermis as the location of their expression was not analysed by *in situ* hybridisation or immunolocalisation. In the present study, the possibility that these genes were amplified from a dermal or epidermal origin cannot be ruled out. This may have occurred if a few cells from the dermis and/or epidermis were carried over during the wool follicle isolation procedure resulting in a mixed pool of RNA, used in the RT-PCR screen, from follicle, dermal and epidermal origin.

Preliminary expression data of ovine *Hoxb-8*, *Hoxb-9*, *Hoxc-5* and *Hoxc-9* in adult sheep wool follicles including RNA protection, Northern and *in situ* hybridisation analysis (*Hoxc-5* and *Hoxc-9*) was obtained using cRNA probes that encompassed their homeobox. Non-specific hybridisation of these probes to other homeobox genes may occur as a result of the high percentage of sequence identity shared between the homeobox genes within this region. Therefore, these results require verification using cRNA probes that encompass

a region of the gene outside the homeobox. Usually, the 3' non-coding DNA of a gene contains regions that are gene-specific.

The preliminary *in situ* hybridisation analysis of ovine *Hoxc-9* and *Hoxc-5* revealed that they were asymmetrically expressed in the cortex, possibly the paracortex, of the lower follicle shaft where the keratinocytes begin their terminal differentiation programme. Interestingly, the cysteine rich proteins encoded by the KAP4 family are elevated in paracortical cells of sheep wool when circulating levels of L-cysteine are elevated (Fratini et al., 1994). The expression pattern suggests that *Hoxc-9* and *Hoxc-5* may play some role in the regulation of the keratin genes expressed in the paracortex. These results require further verification by *in situ* hybridisation using *Hoxc-9* and *Hoxc-5* gene-specific probes located outside the homeobox. Co-localisation of these homeobox genes with a member of the KAP4 family could be achieved by two colour *in situ* hybridisation using digoxygenin and fluorescein labeled riboprobes. Further investigation is required to determine if the KAP4 proteins are upregulated directly or indirectly by the homeoproteins Hoxc-9 and Hoxc-5. A super-shift assay in which hybridisation of the Hoxc-9 and Hoxc-5 proteins with the KAP4 promoter may show that the Hoxc-9 and Hoxc-5 proteins can bind directly to the KAP4 promoter. Co-transfection of a construct that drives the expression of Hoxc-9 or Hoxc-5 proteins with a KAP4 promoter / LacZ reporter construct in a keratinocyte cell line may show if the Hoxc-9 and Hoxc-5 proteins can transcriptionally activate the KAP4 proteins. It should be noted that one of the difficulties in this approach is that there are no immortalised hair keratinocyte cell lines available. In another approach, a transient co-transfection assay as reported by Andersen et al, (1993b) could be used. In their transient co-transfection assay, a CMV promoter was used to drive the ectopic expression of the two POU homeobox genes, Skn-1a and Skn1-i in HeLa cells. Skn-1a and Skn1-i

proteins then bound to the other co-transfected construct that contained the promoter of the human cytokeratin 10 gene linked to a luciferase reporter. Both Skn-1a and Skn1-I activated transcription of the human cytokeratin 10 gene that was detected by a luciferase reporter. Therefore, co-transfection of a construct containing the CMV promoter that drives the expression of Hoxc-9 or Hoxc-5 proteins with a KAP4 promoter / LacZ reporter construct in the HeLa cell line may show if the Hoxc-9 and Hoxc-5 proteins can transcriptionally activate the KAP4 proteins.

8.3 *Barx2*

It was impossible within the three-year period of this PhD to isolate, sequence and characterise all of the homeobox genes amplified in the RT-PCR screen (Chapter 3). *Barx2* was chosen because it was a novel gene at the time of its isolation.

Initially, 3' RACE was used to amplify the 3' cDNA end of ovine *Barx2* that was isolated and sequenced. To obtain the full ovine *Barx2* gene for functional studies, an EMBL3 lambda sheep genomic library was screened and a clone that may contain the full *Barx2* gene was isolated. The sheep genomic clone was isolated in 1994 and three years later while characterising *Barx2*, the murine orthologue was isolated by another research group and the cDNA sequence published (Jones et al, 1997).

Barx2 was most similar to the *Bar* family of *Hox* genes. The Barx2 homeodomain is identical between sheep, mouse and human and differs by one residue from chicken Barx2b. As found in other members of the *Bar* family, the ovine *Barx2* homeobox contains the two atypical residues, threonine and tyrosine within helix 3. This region may influence DNA binding specificity (Jones etal.,1997). Downstream of the homeodomain the Barx2

proteins are perfectly conserved for a further 35 amino acids including a segment of 17 amino acids, including a number of basic residues, also conserved with Barx1. This 17-amino acid basic region resulted in a 2-fold decrease in binding to homeodomain binding site sequences from the Ng-CAM and L1 genes and a 6.5-fold increase in binding to the homeodomain binding site from the N-CAM promoter (Edelman et al., 2000). Furthermore, a 24-residue N-terminal region of Barx2 was essential for repression of N-CAM promoter activity. When this region was absent, Barx2 activated the N-CAM promoter. A 63-residue C-terminal domain was required for this activation (Edelman et al., 2000). During hair follicle development N-CAM is expressed in the epithelial cells of the hair placode (stage 2), outer cells of the placode destined to become the outer root sheath (stage 3 and 4), and in the dermal papillae and dermal cells during stages 4 to 8 (Vielkind et al., 1995). Hence, one of the functions of Barx2 during follicle development may be to regulate and more specifically to repress N-CAM activity during the development of the follicle.

The sheep wool follicle and mouse *Barx2* transcripts are of similar size. In addition, hybridisation studies of human, mouse and sheep genomic DNA using a *Barx2* coding probe detected single bands in several genomic digests, indicating that *Barx2* may be encoded by a single gene. This finding was also confirmed for the mouse using two different *Barx2* probes (Jones, et al., 1997). In the present study, a partial map of a genomic clone containing the 3' end of *Barx2* was reported. The human BARX2 gene was recently cloned, consists of four exons ranging in size from 85 to 1099 bp, and is located on human chromosome 11q25 (Hjalt and Murray (1999), Krasner, et al., 2000). Interestingly, this chromosomal location is within the minimal deletion region for Jacobsen

syndrome, a syndrome including craniosynostosis and other developmental abnormalities (Krasner, et al., 2000).

Barx2 was abundantly expressed in other keratinizing epithelia, namely oesophagus and tongue, and weakly expressed in thymus. In the tongue, *Barx2* was expressed in the basal and suprabasal cells of the lower half of the tongue epithelium but not in the upper epithelium or filiform papilla. Expression was not detected in several other non keratinising tissue namely, adult spleen, skeletal muscle, brain, heart, kidney, liver and lung. Although expression was not detected in the adult sheep spleen, *Barx2* was reported to be strongly expressed in the adult mouse spleen (Jones et al., 1997). The reason for this difference between the two species is unclear. A subsequent analysis of *Barx2* expression in the spleen from another sheep using RT-PCR or RNA protection analysis will confirm the initial result.

To explore the role of *Barx2* in the follicle, *Barx2* expression in wool follicle development was examined from initiation through to formation of the mature, fibre-producing follicle. *Barx2* exhibited a spatially restricted expression pattern during both epidermal and follicular development. *Barx2* was uniformly expressed in the embryonic ectoderm but was transiently downregulated during the initiation of follicle morphogenesis. Subsequently, *Barx2* was expressed throughout the epithelial component of the developing follicle except for a small group of cells at the leading edge of the follicle placode. Those *Barx2*-negative cells are destined to form the follicle bulb and are the progenitors of the inner root sheath and hair shaft. In adult follicles, *Barx2* was expressed throughout the outer root sheath but not in the inner root sheath or hair shaft, or in dermal cells associated with the follicle.

During the course of this project, mouse *Barx2* was isolated in a screen for factors regulating the gene for the cell adhesion molecule, L1 (Jones *et al*, 1997). Mouse *Barx2* appears to be able to both activate and repress the *L1* promoter. In mouse embryogenesis *Barx2* is expressed in the central and peripheral nervous system and the ectodermal lining of craniofacial tissues, regions which express L1 and another cell adhesion molecule, Ng-CAM (Jones *et al*, 1997). *Barx2* expression was studied in the developing embryo at E10.5, E11.5 and E12.5 but was not reported at later developmental time points. Hence, *Barx2* expression was not reported in follicle development which, for pelage hairs, begins at E13.25 in the shoulders.

Interestingly, the pattern of *Barx2* expression in follicle morphogenesis is strikingly similar to that of the cell adhesion molecule, E-cadherin. E-cadherin and P-cadherin, the subject of several studies in the follicle, are expressed in complementary or partially overlapping patterns, depending on developmental stage and location (Hirai *et al*, 1989; Fujita *et al*, 1992; Hardy and Vielkind, 1996; Muller-Rover *et al*, 1999). E-cadherin is primarily located in the suprabasal layers and to a lesser extent in the basal layers, whereas P-cadherin is found in the proliferative, basal layers. Thus, in developing hair placodes, E-cadherin is abundant in the upper portion and reduced in the epithelial cells at the leading edge and P-cadherin shows the opposite pattern. Advanced follicles exhibit high levels of E-cadherin in the outer root sheath and a lesser amount in the differentiating hair matrix cells; in contrast, P-cadherin is restricted to a low level in the basal outer root sheath but is abundantly expressed in proliferating hair matrix cells.

Two other intriguing features of *Barx2* expression in follicle morphogenesis have been uncovered in this study. In the midst of a generalised expression in the embryonic ectoderm and the follicle placode, *Barx2* is transiently downregulated in the ectoderm at

follicle initiation and is not expressed in epithelial cells at the leading edge of the placode. At both sites dynamic cell movement occurs as the cells move down into the dermis, a process that requires plasticity in cellular adhesion. Exclusion of *Barx2* expression from these regions may signify that the cellular processes controlled by *Barx2* in follicle morphogenesis are likely to be involved in maintaining a more stable tissue architecture rather than a dynamic one. Given the finding of *Barx2* expression in tongue and oesophagus, it is possible that *Barx2* may have a general function in controlling adhesive processes in keratinising epithelia.

Lastly, *Barx2* expression was also detected in dermal cells that looked morphologically similar to leukocytes. Faint expression of *Barx2* was also detected in the sheep thymus, an organ that produces lymphocytes. Furthermore, *Barx2* expression was also detected in the spleen of mice (Jones et al., 1997) an organ that stimulates lymphocytes in the same manner as in lymph nodes. Other *Hox* genes have been reported to be expressed in hematopoietic cells lines including B and T cell lines, erythroid cell lines and myelomonocytic cell lines (reviewed in Lawrence et al., 1996). *Barx2* may play some role in regulating the cell adhesion molecules of these cells.

8.4 *Hoxc-13*

Ovine *Hoxc-13* gene, a member of the *Abdominal-B* class, was chosen because its expression in the follicle bulb correlated to the initiation of hair keratin gene expression in the hair cortex. A combination of RT-PCR, 3' and 5' RACE and genomic cloning was used to obtain a complete, predicted cDNA sequence for *Hoxc-13* and over 1 kb of promoter sequence. As for all *Abd-B*-related genes, *Hoxc-13* does not have the conserved pentapeptide upstream of the homeobox (Duboule, 1994). It contains two exons separated

by an intron of 5.3 kb that is located a short distance upstream from the homeobox, a similarity in other *Abd-B*-related genes, for example *Hoxb-13* (Zeltser et al., 1996). The homeodomain is in the second exon and the protein terminates soon after the C-terminal end of the homeobox. The protein encoded by the *Hoxc-13* gene contained 330 residues, similar in size to other *Abd-B* related genes that range between 320 and 340 residues (see Duboule, 1994).

Expression of *Hoxc-13* during embryonic pelage and vibrissae follicle development was studied by *in situ* hybridisation analysis. *Hoxc-13* expression was detected in the epidermis before the initiation of follicle development at E51 and continued in the epidermis during each stage of follicle develop until at least E135 although the levels gradually declined and were negligible in the adult epidermis.

Hoxc-13 was expressed in both the dividing and differentiating cells of the embryonic pelage and vibrissae bulb. *Hoxc-13* expression initiated in the bulb of both pelage and vibrissae follicles during the formation of the wool fibre when most likely the hair keratin gene families are activated. Interestingly, when Hoxc-13 is deleted from the mouse genome, both follicles and brittle hair fibres are produced (Godwin and Capecchi, 1998). Hence Hoxc-13 is not the initiating signal for follicle development and is not solely required for production of a keratinised fibre. However, the fact that the knockout fibres were brittle suggests that Hoxc-13 may be required for the activation and expression of specific keratins as similar phenotypes have been produced by the overexpression of keratin intermediate filament genes (Powell and Rogers, 1990). Brittle hairs resulting in a hair-loss phenotype were produced in mice that overexpress a keratin intermediate filament protein in the cortex. The brittleness was attributed to an imbalance in the ratio of keratin intermediate filament to keratin associated proteins. Hence, when *Hoxc-13* expression was

ablated, an imbalance of keratin intermediate filament or keratin associated proteins may have lead to the production of brittle fibres.

Is *Hoxc-13* required for the expression of a particular type of keratin protein? One would assume that the same keratin protein would be expressed in all of the morphologically different and independent cells where the Hoxc-13 protein is expressed such as the hair, tongue and nails. The possibility arises that an unknown secondary factor may be required in addition to Hoxc-13 for expression of the target genes. It should be noted that at this stage it is not known whether the Hoxc-13 protein is expressed in the scattered cells of the upper dermis. Furthermore, it is not known what genes Hoxc-13 regulates and it may be that Hoxc-13 regulates other transcription factors or growth factors that in turn regulate keratin genes. Comparison of the hair IF and KAP protein composition of normal mouse follicles with mouse follicles that do not contain the functional Hoxc-13 protein (ie. knockout mice), by two dimensional protein gel, may give insights into which type of keratin genes are regulated, directly or indirectly, by Hoxc-13.

To examine the possibility that Hoxc-13 directly regulates keratin IF and/or KAP genes, a transient co-transfection assay as reported by Andersen et al, (1993b) could be used. In this approach, a construct containing the CMV promoter that drives the expression of the Hoxc-13 protein would be co-transfected into the HeLa cell line with another construct containing the promoter of an intermediate filament type I or II gene linked to a LacZ reporter. Expression of the lacZ gene is detected by staining for β-Galactosidase. The intermediate filament genes would be the logical choice to first study as their expression initiates in the upper bulb and overlaps *Hoxc-13*. Three ovine genes that are prime candidates are *K1.2*, *K2.10* or *K2.12* because their expression domains overlap ovine

Hoxc-13 in the bulb. *K2.10* (mouse *Hb4*, human *Hb3*) expression initiates in the lower shaft and extends well into the upper wool shaft (Powell et al., 1992). *K2.12* (human *Hb5*) expression parallels *Hoxc-13* even closer than *K2.10*. *K2.12* expression begins immediately above a small population of matrix cells at the base of the hair bulb and the trichocytes lining the dermal papilla and extends upward through the matrix and ends in the lower part of the cortex of the hair shaft (Powell and Rogers, 1997, Rogers et al., 1997). *K1.2* (mouse and human *Ha3*) is expressed in hair shaft cortex, above the dermal papilla and also in the filiform papilla of tongue (Winter et al., 1994).

Another role that *Hoxc-13* may have is to control the proliferation or differentiation status of these cells. Although it is not known if the cells of the upper dermis expressing *Hoxc-13* are proliferating or differentiating, cells expressing *Hoxc-13* in the tongue, follicle bulb and outer root sheath are moving between a transitionary state of proliferation and differentiation. Furthermore, homeobox genes have been previously shown to control the proliferation and differentiation status of hematopoietic cells (Deguchi, et al., 1992; Sauvageau, et al., 1995; Helgason, et al., 1996).

To try and elucidate the functional role of Hoxc-13 in the follicle, a construct was made with the intention of ectopically expressing Hoxc-13 under the control of the K14 promoter. The K14 promoter has been shown to be functional in transgenic studies undertaken by others (Vassar and Fuchs, 1991; Cheng et al., 1992). K14 was chosen to drive the expression of Hoxc-13 in cells of the adult epidermal basal layer, a region that we thought, at the time of transgene construction, did not express Hoxc-13. Initial *in situ* hybridisation analysis of *Hoxc-13* mRNA in adult sheep skin did indeed reveal an absence of *Hoxc-13* expression in the epidermal basal and suprabasal layers. *In situ* hybridisation analysis, performed after the attempt to make a transgenic mouse, revealed *Hoxc-13*

mRNA expression also in the embryonic epidermis and developing placode. However, Godwin and Capecchi (1998) showed that the Hoxc-13 protein was not expressed in the embryonic epidermis and postnatal epidermis. Therefore, in hindsight, it would be better to drive the ectopic expression of Hoxc-13 in a tissue where Hoxc-13 has not been expressed in at any time point.

8.5 Final Conclusion

It is difficult to predict the total number of homeobox genes that are involved in the development of the follicle. The number of reported homeobox genes expressed in the hair follicle has increased in recent years and it would appear that there are many more yet to be characterised as only a few of the currently known homeobox genes have been studied in the follicle (see Introduction, Table. 1). This is probably due to the vast numbers of homeobox genes that have been isolated and new ones continually being added to the list each year. Current cDNA array chip technology, that is able to determine the expression profile of thousands of expressed genes in any tissue, could be used to determine the total number of homeobox genes expressed in the follicle in a short period of time. Furthermore, this technology could show the differential expression of all of the homeobox genes during each stage of embryonic follicle development.

An important point to consider is that most reports on homeobox gene expression have focussed mainly at the RNA level but have not shown if the homeodomain protein is expressed in the same tissues. It may be that some homeobox genes are transcribed but not translated. A pertinent example was shown in this thesis where *Hoxc-13* appeared to be transcribed during the development of the embryonic mouse epidermis but the Hoxc-13 protein was not expressed (Godwin and Capecchi et al., 1998). *Hoxc-13* RNA was

expressed in the sheep embryonic epidermis and placode however it is not known if the protein was expressed. Immunolocalisation of Hoxc-13 in sheep skin using an antibody should be performed in order to verify Hoxc-13 expression in these regions.

Homeobox genes are known to be essential for the patterning and development of many segmental structures in both invertebrates and vertebrates (Gilbert, 1994) but questions that still remain to be addressed in the development of the epidermis and follicle are;

1) Why are there so many homeobox genes expressed during the development of the epidermis and follicle and what function do they individually have? Is there a redundancy aspect?
2) Do homeodomain proteins function independently or cooperatively with other homeodomain proteins or unknown secondary factors to determine the fate of every follicle cell type?
3) If the homeodomain proteins act as transcriptional regulators what genes do they target in the follicle?

The role of homeobox genes in the development of embryonic epidermis and follicle and the mechanism of action is not fully understood. Stelnicki et al., 1998 propose that HOX proteins function as gene activators to upregulate cellular proliferation in the early developing epidermis and then act as repressers that downregulate cellular proliferation in the outer layers of the mature skin. The mode of action may be I) as independent DNA binding proteins or II) in conjunction with a variable set of partner proteins (including PBX) to activate and repress cellular proliferation depending on the tissue they are expressed in.

In conclusion the target genes of the transcription factors encoded by the homeotic genes are largely unknown, but have been well implicated in controlling the expression of differentiation-specific genes and adhesion molecules on the membrane.

Appendix

Appendix 1

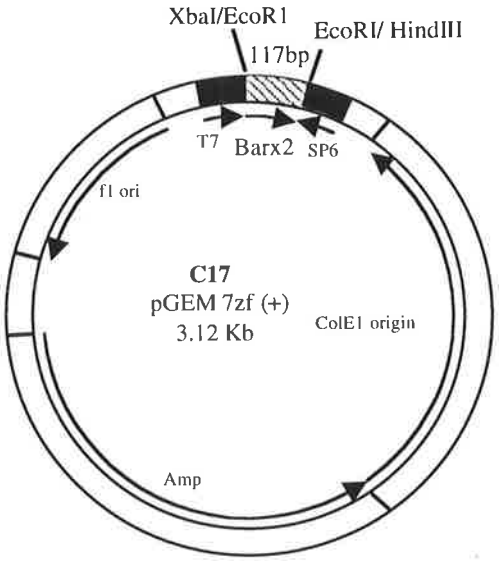
A. Plasmid map of clone C17. Clone C17 contains the partial homeobox of *Barx2*(117bp). Clone C17 was created by inserting the EcoRI cleaved117bp *Barx2* homeobox amplified by RT-PCR using the degenerate primers, WFQNRR and ELEKEF, into the EcoRI site of pGEM 7Zf (+).

B. Plasmid map of clone C17B1. A 1.1 kb 3'RACE product was klenow treated and then inserted into the SmaI site of pBSc KS.

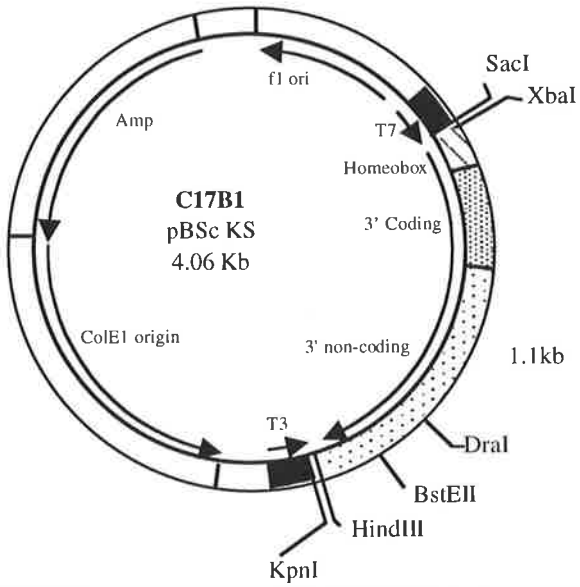
C. Plasmid map of clone C17B1-COD. A 223 bp coding fragment from clone C17B1 was PCR - amplified using the primers P7 and P11 (see figure 5.6 for location of the primers) and inserted into the SmaI site of pGEM 7Zf (+).

D. Plasmid map of clone C17B1-NC. A 163 bp BsteII / DraI fragment from clone C17B1 was excised with the enzymes BsteII and DraI, klenow treated and inserted into the SmaI site of the base vector pBSc KS.

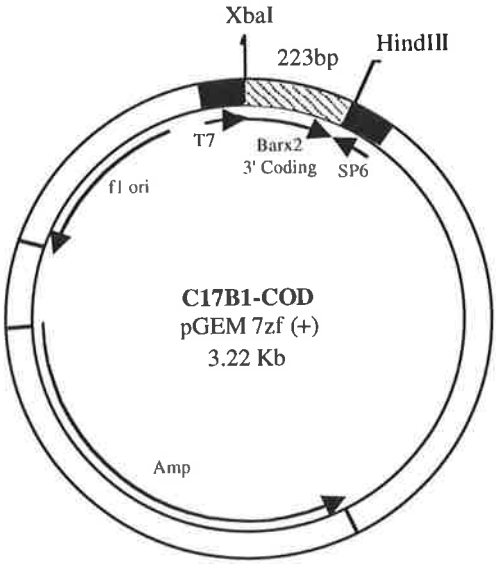
A


XbaI/EcoR1
EcoRI/ HindIII
117bp
T7
Barx2
SP6
f1 ori
C17
pGEM 7zf (+)
3.12 Kb
ColE1 origin
Amp

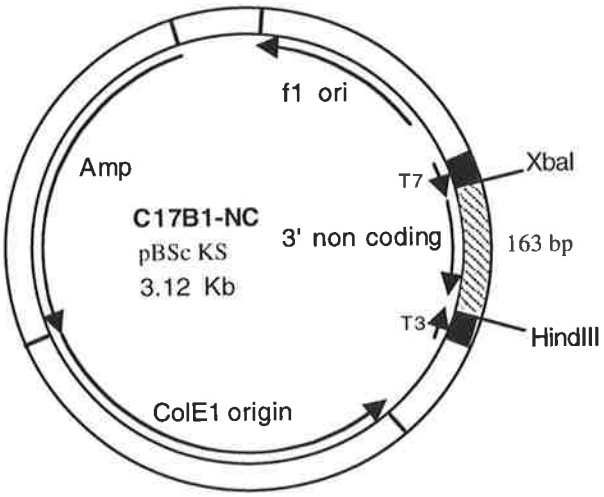
B


f1 ori
SacI
XbaI
Amp
T7
Homeobox
C17B1
pBSc KS
4.06 Kb
3' Coding
ColE1 origin
3' non-coding
1.1kb
T3
DraI
BstEII
HindIII
KpnI

C


XbaI
223bp
HindIII
T7
Barx2
3' Coding
SP6
f1 ori
C17B1-COD
pGEM 7zf (+)
3.22 Kb
Amp

D


f1 ori
Amp
XbaI
T7
C17B1-NC
pBSc KS
3.12 Kb
3' non coding
163 bp
T3
HindIII
ColE1 origin

Appendix 2

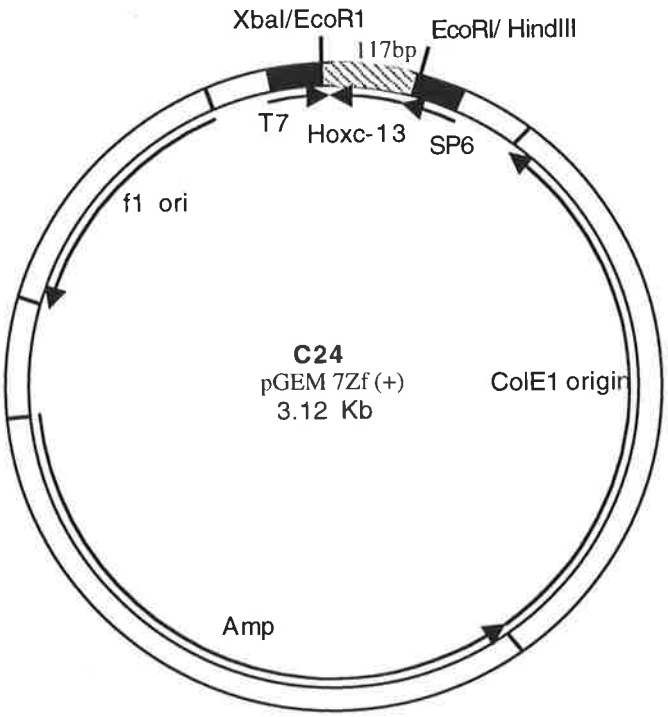
A. Plasmid map of clone C24. Clone C24 contains the partial homeobox of *Hoxc-13* (117bp). Clone C24 was created by inserting the EcoRI cleaved117bp *Hoxc-13* homeobox amplified by RT-PCR using the degenerate primers, WFQNRR and ELEKEF, into the EcoRI site of pGEM 7Zf (+).

B. Plasmid map of clone C24B1. A 1.4 kb 3'RACE product was klenow treated and then inserted into the SmaI site of pBSc KS.

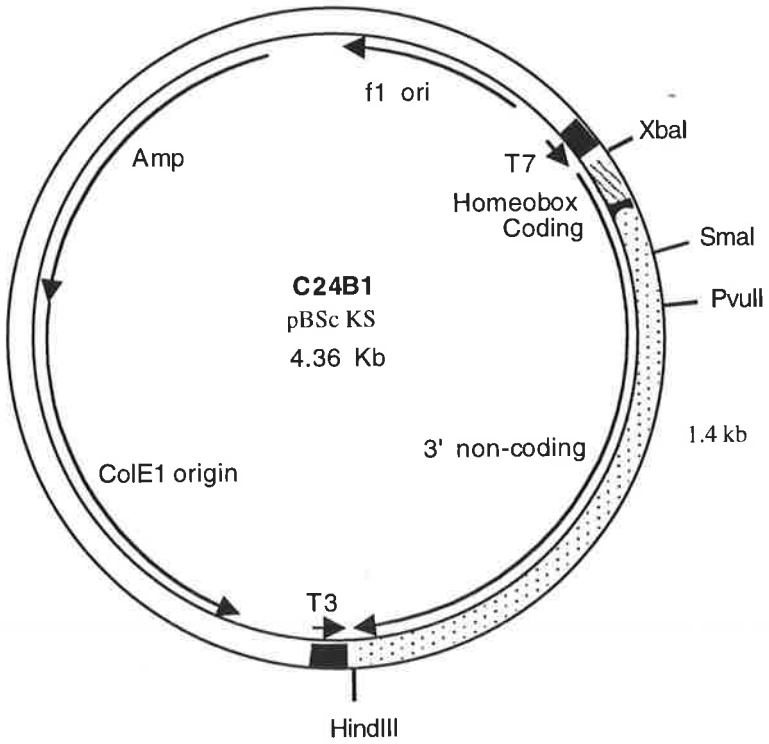
C. Plasmid map of clone C13B1-COD. A 285 bp Pst I coding fragment was cleaved from the 2.8 kb SalI / HindIII fragment of the lambda clone 9A (see Fig. 6.7 C and Fig. 7.14 A) and inserted into the PstI site of the base vector pBSc KS.

D. Plasmid map of clone C13-NC. A 183 bp SmaI / PvuII 3' non-coding fragment was cleaved from clone C24B1 and inserted into the base vector pBSc KS.

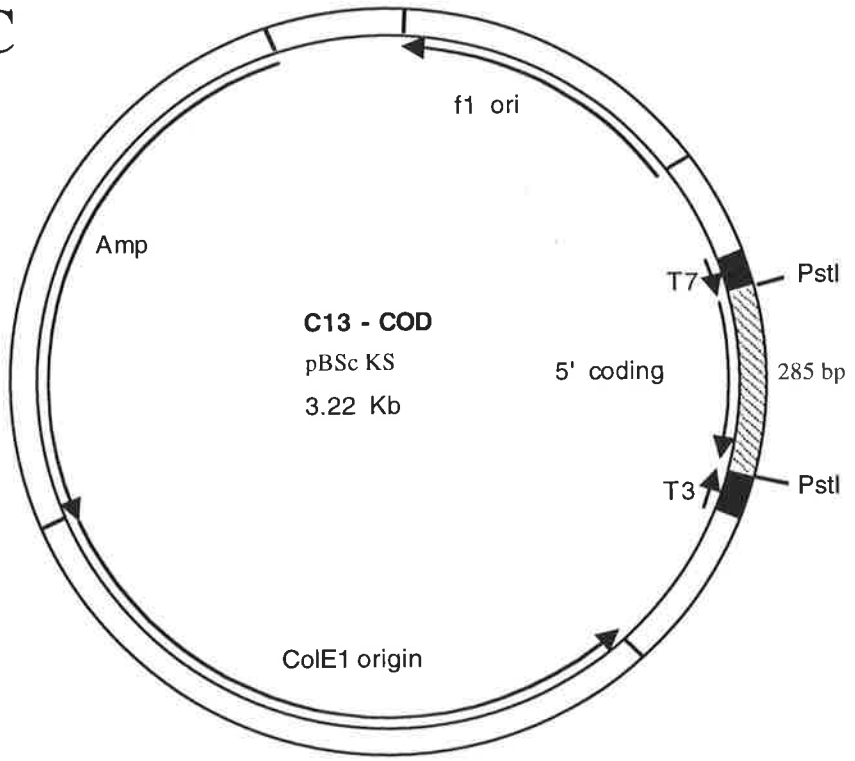
A


XbaI/EcoR1
EcoRI/ HindIII
117bp
T7
Hoxc-13
SP6
f1 ori
C24
pGEM 7Zf (+)
3.12 Kb
ColE1 origin
Amp

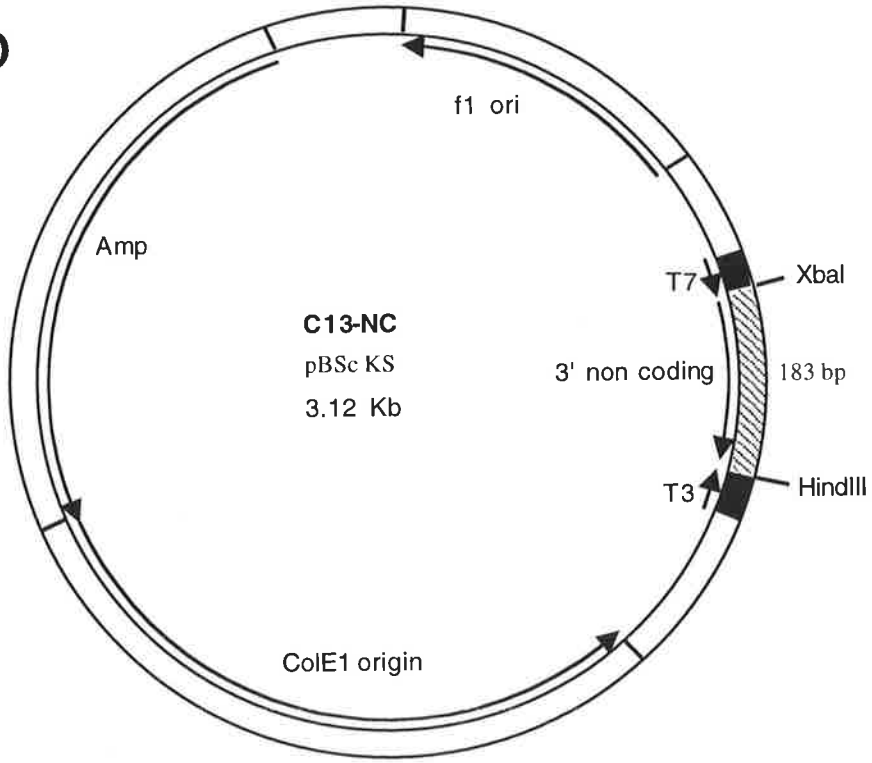
B


f1 ori
Amp
XbaI
T7
Homeobox
Coding
SmaI
PvuII
C24B1
pBSc KS
4.36 Kb
3' non-coding
1.4 kb
ColE1 origin
T3
HindIII

C


f1 ori
Amp
T7
PstI
C13 - COD
pBSc KS
3.22 Kb
5' coding
285 bp
T3
PstI
ColE1 origin

D


f1 ori
Amp
T7
XbaI
C13-NC
pBSc KS
3.12 Kb
3' non coding
183 bp
T3
HindIII
ColE1 origin

Appendix 3

Publication in the Journal of Investigative Dermatology.

SHORT COMMUNICATION

Expression of the Homeobox Gene, *Barx2*, in Wool Follicle Development

Guy Sander,[1] C. Simon Bawden, Philip I. Hynd, Antonietta Nesci, George Rogers, and Barry C. Powell[1]
Department of Animal Science, University of Adelaide, Waite Campus, Glen Osmond, South Australia

We have cloned ovine *Barx2*, a member of the *Bar* class of homeobox genes, and present the first description of *Barx2* expression in wool follicle development. *Barx2* is uniformly expressed in the embryonic ectoderm but is transiently downregulated during the initiation of follicle morphogenesis. Subsequently, *Barx2* is expressed throughout the epithelial component of the developing follicle except for a small group of cells at the leading edge of the follicle placode. These *Barx2*-negative cells are destined to form the follicle bulb and are the progenitors of the inner root sheath and hair shaft. In adult follicles, *Barx2* is expressed throughout the outer root sheath but not in the inner root sheath or hair shaft, or in dermal cells associated with the follicle. The pattern of *Barx2* expression in follicle morphogenesis is similar to that of the cell adhesion molecule E-cadherin, a similarity that echoes *Barx2* coexpression with the L1 cell adhesion molecule in other tissues during mouse embryogenesis. *Barx2* is also expressed in tongue and esophagus, two other keratinizing tissues, and we speculate that *Barx2* may have a general function in controlling adhesive processes in keratinizing epithelia. ***Key words: Barx/cadherin/cell adhesion/follicle/homeobox/leukocyte/molecule/outer root sheath. J Invest Dermatol 115:000–000, 2000***

Hair follicle morphogenesis initiates in response to signaling between embryonic ectoderm and mesoderm, and the subsequent extensive morphologic changes that occur involve dynamic interactions between epithelial cells and between epithelial and mesenchymal cells. Cell–cell interactions are mediated by cell adhesion molecules, and expression of a number of these, principally the cadherins and integrins, has been identified during follicle morphogenesis (for a review see Muller-Rover and Paus, 1998). E-cadherin and p-cadherin, the subject of several studies in the follicle, are expressed in complementary or partially overlapping patterns, depending on developmental stage and location (Hirai *et al*, 1989; Fujita *et al*, 1992; Hardy and Vielkind, 1996; Muller-Rover *et al*, 1999). E-cadherin is primarily located in the suprabasal layers and to a lesser extent in the basal layers, whereas p-cadherin is found in the proliferative basal layers. Thus, in developing hair placodes, E-cadherin is abundant in the upper portion and reduced in the epithelial cells at the leading edge, and p-cadherin shows the opposite pattern. Advanced follicles exhibit high levels of E-cadherin in the outer root sheath and a lesser amount in the differentiating hair matrix cells; in contrast, p-cadherin is restricted to a low level in the basal outer root sheath but is abundantly expressed in proliferating hair matrix cells. In a pioneering experiment, Hirai *et al* (1989) demonstrated the importance of these molecules in follicle morphogenesis using antibodies to block both cadherins in cultured skin explants, causing dramatic changes in follicle development.

Another group of molecules that have been implicated in the development of follicles are the homeobox genes. In recent years, a number of homeobox genes have been shown to be differentially and spatially expressed during embryonic development of the epidermis and follicle: msx-1 and msx-2 (MacKenzie *et al*, 1991; 1992; Noveen *et al*, 1995), Skn-1A and Skn-1I (Andersen *et al*, 1993), Hox2.2 (Mathews *et al*, 1993), Otx-1, Otx-2, Hoxc-8, d-9, d-11, and d-13 (Kanzler *et al*, 1994), HOXC4 (Rieger *et al*, 1994), Dlx3 (Morasso *et al*, 1996), Hoxc-13 (Godwin and Capecchi, 1998), HOXA4, HOXA5, HOXA7, HOXC4, HOXB4, and HOXB7 (Stelnicki *et al*, 1998). Although little is known about which molecules are regulated by the products of these genes, other homeobox genes have been implicated in controlling expression of cell adhesion molecules (Cillo *et al*, 1996; Wang *et al*, 1996; Jones *et al*, 1997; Edelman and Jones, 1998; King *et al*, 1998; Lincecum *et al*, 1998; Meech *et al*, 1999; Nguyen Ba-Charvet *et al*, 1999; Siegler and Jia, 1999). Recently, a vertebrate homolog of the *Drosophila Bar* homeobox gene, mouse *Barx2*, was isolated in a screen for factors regulating the gene for the cell adhesion molecule L1 (Jones *et al*, 1997). *Barx2* appears to be able to both activate and repress the *L1* promoter. In mouse embryogenesis *Barx2* is expressed in the central and peripheral nervous system and the ectodermal lining of craniofacial tissues, regions that express L1 and another cell adhesion molecule, Ng-CAM (Jones *et al*, 1997). In this study we have cloned an ovine *Barx2* cDNA from the follicle and present the first description of its expression in follicle development in a pattern that is strikingly similar to that of E-cadherin.

MATERIALS AND METHODS

Cloning of ovine *Barx2* Using reverse transcriptase polymerase phase reaction (RT-PCR), a partial ovine *Barx2* cDNA clone of 117 bp was isolated from adult wool follicle cDNA. Degenerate primers to the

Manuscript received January 14, 2000; revised May 22, 2000; accepted for publication July 17, 2000.

Reprint requests to: Dr. Barry Powell, Child Health Research Institute, Women's and Children's Hospital, 72 King William Road, North Adelaide, South Australia, 5006. Email: bpowell@medicine.adelaide.edu.au

[1]Present address: Child Health Research Institute, Women's and Children's Hospital, 72 King William Road, North Adelaide, South Australia, 5006.

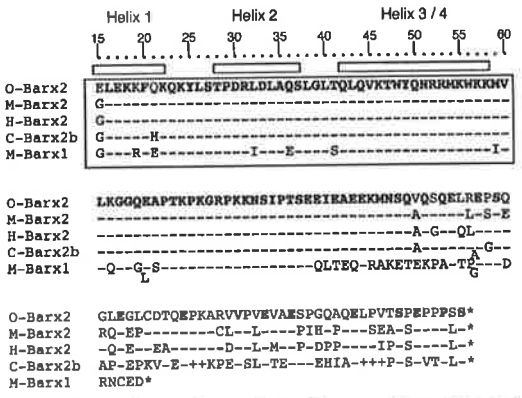


Figure 1. Comparison of vertebrate Barx proteins. The partial sequence of the ovine *Barx2* homeodomain and C-terminal domain are compared with chicken Barx2b (Smith and Tabin, 1999), human Barx2 (accession number AF031924), mouse Barx1 (Tissier-Seta *et al*, 1995), and mouse Barx2 (Jones *et al*, 1997). The partial homeodomain is boxed. Dashes indicate sequence identity with ovine Barx2. Residues that are shared with each Barx2 protein are typed in bold. The carboxy-terminal end of each protein is indicated by an asterisk. To maximize alignment, "extra" leucine and glycine residues in mouse *Barx1* and an "extra" alanine residue in chick Barx2b relative to ovine *Barx2* are shown below the sequence comparisons. A + indicates that a space has been added between residues to maximize alignment. Note that the glutamic acid (E) residue at the N-terminal end of the ovine sequence may be a result of degenerate primer mismatch during RT-PCR amplification.

conserved regions (ELEKEF and WFQNRR) of the *Drosophila* Antennapedia homeodomain were as described by Burglin *et al* (1989). The partial homeobox clone was sequenced in both directions by cycle sequencing and identified as *Barx2* by database sequence searches using ANGIS (Australian National Genomic Information Retrieval System).

A subsequent partial ovine *Barx2* cDNA clone of 1104 bp (accession number AF265552) representing the 3′ end of Barx2 was isolated using 3′ PCR-RACE (Frohman *et al*, 1988). The oligonucleotide 5′–CCG GATCCCTGCAGGAATTCG TCG$(dT)_{17}$YX–3′ (Y = A, G, or C, and X = A, C, G, or T) was used to prime the first strand cDNA. The upper primer (5′–TAGTCGACCAGACAGGTTGGACTTGG–3′) was designed from the partial *Barx2* homeobox using Oligo-4 software (Bresatec, SA). The oligonucleotide contained a SalI restriction endonuclease site for cloning into pBSc KS (Stratagene). The sequence of the lower primer was 5′–CCGGATCCCTGCAGGAATTCGTCG–3′. A magnesium concentration of 1.5 mM was used in the PCR with the following parameters: (i) one cycle at 94°C (45 s), 72°C (2 min); (ii) while the PCR was held at 72°C the TAQ polymerase and the upper primer were added; (iii) four cycles of 94°C (45 s), 56°C (2 min), 72°C (2 min); (iv) the PCR was then held at 72°C while the lower primer was added; (v) 35 cycles of 94°C (45 s), 63°C (2 min), 72°C (2 min); (vi) a final extension step of 72°C (7 min) and then held at 4°C (overnight).

Once the 3′ cDNA end was cloned, a 223 bp *Barx2* 3′ coding fragment adjacent to the homeobox was amplified by PCR using the upper primer 5′–TGGTTCTTAAAGGTGGACA–3′ and lower primer 5′–CTGGGC-CTGGCCTGGGGACTCG–3′ and was cloned into a pGEM-7Zf SmaI vector (*Barx2* 3′ coding). This *Barx2* 3′ coding fragment was used as a gene-specific probe for the northern, RNA protection, and *in situ* hybridization experiments.

Northern blot analysis Total RNA was isolated from adult Corriedale sheep wool follicles and poly-A(+) RNA was prepared (Poly-A-tract kit System IV, Promega). Approximately 3 μg samples of poly-A(+) RNA were fractionated in a 1% agarose formaldehyde gel and then transferred to Zetaprobe GT membrane (Bio-Rad) and probed with the 223 bp *Barx2* 3′ coding fragment labeled with α-^{32}P-dCTP. Filters were washed in 2 × sodium citrate/chloride buffer/0.1% sodium dodecyl sulfate at 65°C for 30 min and exposed to a phosphor-image screen for 24 h.

RNA protection Ten micrograms of total RNA isolated from adult sheep esophagus, spleen, thymus, tongue, and wool follicles was analyzed

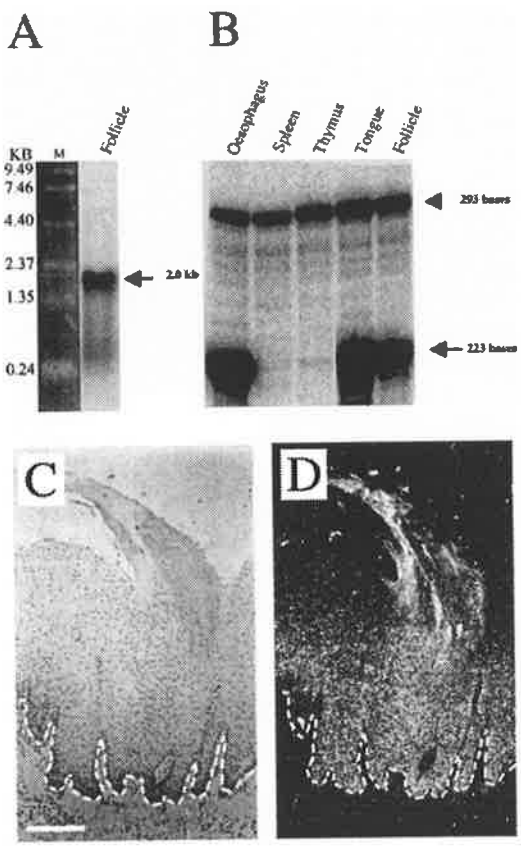


Figure 2. Analysis of *Barx2* expression in adult sheep wool follicles. (*A*) A transcript of approximately 2 kb was identified by northern blot. M = BRL RNA ladder. (*B*) RNA protection analysis of adult sheep esophagus, spleen, thymus, tongue, and wool follicle with an ovine *Barx2* 3′ coding probe. The position of protected *Barx2* RNA is indicated by the *arrow*. Some full-length undigested probe of 293 bases is present in each lane (*arrowhead*). (*C*, *D*) *Barx2* expression in vertical sections of adult sheep tongue. Note expression in the lower half of the tongue epithelium. Part (*C*) is photographed in bright field and (*D*) in dark field. Dashes mark the underlying basal layer. *Scale bar*: 150 μm.

by RNA protection for *Barx2* expression using an α-^{32}P-rUTP labeled *Barx2* 3′ coding antisense probe. The products of the protection reactions were fractionated in a denaturing 6% (19:1) urea-polyacrylamide gel electrophoresis system and the gel was exposed to a phosphor-image screen for 72 h. Probe controls (with yeast RNA, ± RNase) were cofractionated (data not shown).

***In situ* hybridization** Vertical skin sections from a time-course of wool follicle development and from adult sheep tongue were hybridized with an α-^{32}P-rUTP-labeled cRNA probe, prepared from the 223 bp *Barx2* 3′ coding fragment, as described previously (Powell and Rogers, 1990). Hybridized sections were treated with a final wash stringency of 0.1 × SSPE at 65°C for 30 min and exposed for 10 d at 4°C. Sections were stained with hematoxlin and photographed under bright-field and dark-field illumination.

RESULTS AND DISCUSSION

We have cloned a partial ovine *Barx2* cDNA from the wool follicle by RT-PCR (**Fig 1**). The Barx2 homeodomain is identical in sheep, mouse, and human and differs by one residue from chicken Barx2b. As found in other members of the *Bar* family, the ovine

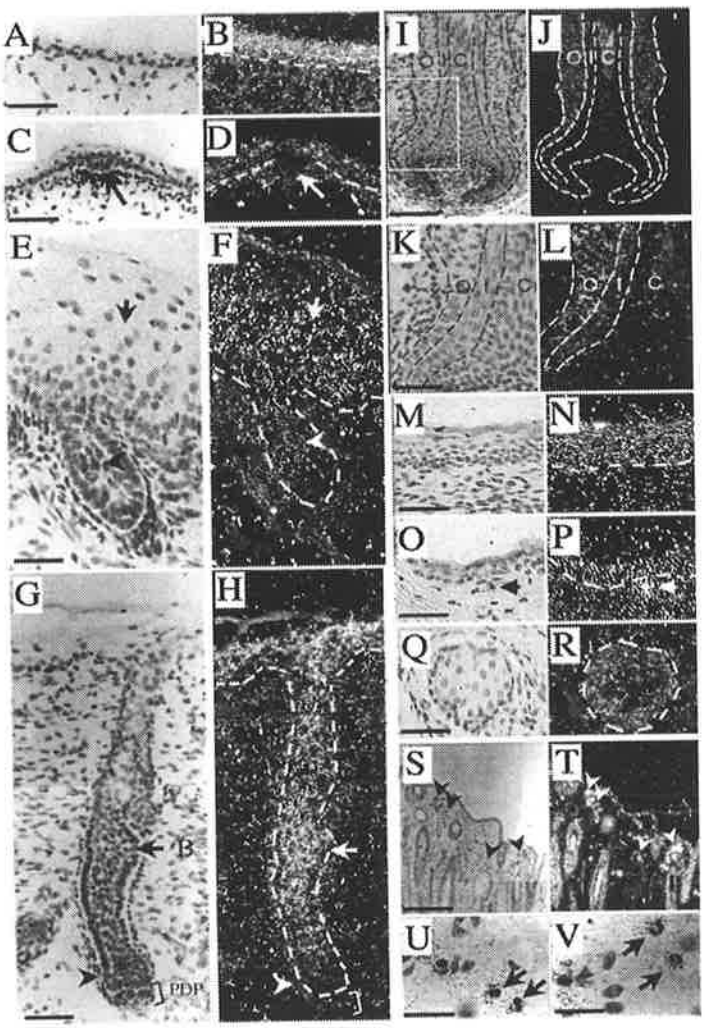


Figure 3. *In situ* hybridization analysis of *Barx2* expression during embryonic follicle development. (*A–H*) *Barx2* expression in the developing embryonic epidermis and follicle. (*I–L*) *Barx2* is expressed in the outer root sheath of a vibrissa follicle. (*K*, *L*) Higher magnification of the boxed section in part (*I*) showing expression of *Barx2* in the outer root sheath. (*M*, *N*) *Barx2* expression in the epidermis several days prior to birth. (*O*, *P*) *Barx2* is weakly expressed in adult epidermis. Note the intense expression of *Barx2* in a cell in the upper dermis (*arrow*) just below the epidermal basal layer. (*Q*, *R*) *Barx2* expression is primarily associated with the basal layer of the sebaceous gland. (*S*, *T*) *Barx2* is abundantly expressed in cells scattered throughout the upper adult dermis (*arrowheads*) and in the outer root sheath of pelage follicles. The artificial fluorescence in the cortex is due to its high refractivity. (*U*, *V*) Higher magnification of dermal cells showing *Barx2* expression (black grains). In all panels, the basal layer is indicated by the dashed line. *Arrows* and *arrowheads* in all figures point to features described in the text. The following stages of follicle development follow Hardy (1992): (*A*, *B*), stage 0; (*C*, *D*), stage 1; (*E*, *F*), stage 2; (*G*, *H*), stage 3; (*I–L*), stage 8; (*M*, *N*), E135 skin; (*O–V*), adult skin of old sheep. Note that parts (*S*) and (*T*) show partial sections of stage 8 follicles. PDP, predermal papilla; B, bulge; O, outer root sheath; C, hair cortex and cuticle; I, inner root sheath. *Scale bar.* (*A*) 40 μm; (*C*) 55 μm; (*E*) 30 μm; (*G*) 40 μm; (*I*) 70 μm; (*K*, *M*, *O*) 35 μm; (*Q*) 25 μm; (*S*) 210 μm; (*U*) 15 μm.

Barx2 homeobox contains the two atypical residues threonine and tyrosine within helix 3. Downstream of the homeodomain the Barx2 proteins are perfectly conserved for a further 35 amino acids including a segment of 17 amino acids also conserved with Barx1. In the remaining carboxy terminal region, 15 amino acids found at various positions are conserved between the Barx2 proteins; eight of these are glutamic acid residues, such that the carboxyl end has a markedly acidic nature.

Northern analysis revealed a *Barx2* mRNA of approximately 2 kb in sheep wool follicle RNA (**Fig 2*A***), of similar size to the mouse *Barx2* transcript (Jones *et al*, 1997). *Barx2* was abundantly expressed in other keratinizing epithelia, namely esophagus and tongue, and weakly expressed in thymus (**Fig 2*B***); in the tongue *Barx2* was expressed in the basal and suprabasal cells of the lower half of the tongue epithelium but not in the upper epithelium or filiform papilla (**Fig 2*C, D***). Expression was not detected in adult spleen (**Fig 2*B***), skeletal muscle, brain, heart, kidney, liver, and lung (data not shown).

To explore the role of *Barx2* in the follicle, we examined *Barx2* expression in wool follicle development from initiation through to formation of the mature, fiber-producing follicle (**Fig 3**). *Barx2* exhibited a spatially restricted expression pattern during both epidermal and follicular development. Before follicle initiation, *Barx2* was expressed throughout the embryonic ectoderm (**Fig 3*A, B***). In the initial stage of follicle development, morphologically identified by epithelial thickening, *Barx2* was *transiently* down-regulated in the ectodermal-derived cells of the incipient follicle placode (**Fig 3*C, D***; *arrow*). As the follicle placode advanced down into the dermis, *Barx2* was expressed in the inner epithelial cells of the placode (**Fig 3*E, F***; *arrowhead*) and was present in the epidermis overlying the placode (**Fig 3*E, F***; *arrow*). This transient down-regulation in expression suggests that not only is *Barx2* not involved in follicle initiation but that its continued expression might be counter-productive.

At more advanced stages the level of *Barx2* expression appeared to increase in the follicle placode (compare **Fig 3*E, F, G, H***) and most placode cells expressed *Barx2*, including cells in the bulge (**Fig 3*G, H***; *arrow*), the putative location of hair follicle stem cells (Cotsarelis *et al*, 1990). A notable distinction was the absence of *Barx2* expression from the epithelial cells at the leading edge of the follicle placode, however (**Fig 3*G, H***; *arrowhead*). These epithelial cells, which are rapidly dividing, are destined to form the follicle

bulb and are the progenitors of the inner root sheath and hair shaft. In late-stage pelage and vibrissae follicles *Barx2* was expressed along the entire length of the outer root sheath, abundantly in the suprabasal cells and to a lesser extent in the basal cells (**Fig 3*I–L***). At the junction of the outer root sheath and the epidermis Barx2 expression was markedly reduced. Expression was not detected in the dermal papilla, follicle bulb, inner root sheath, or hair shaft at any stage of development.

As the ectoderm developed from the prefollicle to the adult stage, *Barx2* expression appeared to be gradually downregulated in the basal and suprabasal layers of the epidermis [compare, for example, embryonic (**Fig 3*A, B***) to prebirth (**Fig 3M, *N***) and adult (**Fig 3*O, P***)]. In addition to this gradual decline, regional variation in epidermal *Barx2* expression was noted. Generally, *Barx2* expression levels were higher in thicker, more cellularized epidermis and at sites of follicle openings (**Fig 3*E, F***). We also noted that *Barx2* was weakly expressed in the basal layer of the sebaceous gland and downregulated in differentiated sebocytes located at the center of the gland (**Fig 3*Q, R***).

In adult skin, several other cell types showed intense *Barx2* expression. These included some cells of the blood sinus in vibrissae follicles (data not shown) and scattered cells in the upper dermis (**Fig 3*S–V***). The latter cells appeared to have bi-lobed or multi-lobed nuclei, typical of leukocytes.

We have shown that the homeobox gene, *Barx2*, is expressed in diverse epithelial tissues and have focused on its expression in follicle development. *Barx2* is expressed in a pattern that is strikingly similar to E-cadherin in embryonic ectoderm, placode development, and the adult outer root sheath (Hirai *et al*, 1989; Fujita *et al*, 1992; Hardy and Vielkind, 1996). We did not detect *Barx2* expression in differentiating hair matrix cells or the inner root sheath, however, where isolated E-cadherin immunoreactivity is reported. In general, the similarity in expression between *Barx2* and E-cadherin echoes *Barx2* coexpression with another cell adhesion molecule, L1, in other tissues during mouse embryogenesis (Jones *et al*, 1997) and we speculate that *Barx2* may regulate changes in cell adhesion during follicle morphogenesis.

Two other intriguing features of *Barx2* expression in follicle morphogenesis have been uncovered in this study. In the midst of a generalized expression in the embryonic ectoderm and the follicle placode, *Barx2* is transiently downregulated in the ectoderm at follicle initiation and is not expressed in epithelial cells at the leading edge of the placode. At both sites dynamic cell movement occurs as the cells move down into the dermis, a process that requires plasticity in cellular adhesion. Exclusion of *Barx2* expression from these regions may signify that the cellular processes controlled by *Barx2* in follicle morphogenesis are likely to be involved in maintaining a more stable tissue architecture rather than a dynamic one. Given the finding of *Barx2* expression in tongue and esophagus, it is possible that *Barx2* may have a general function in controlling adhesive processes in keratinizing epithelia.

We thank Nicole Wilson for kindly providing the embryonic sheep skin sections. This project was funded by grants from the Australian Research Council and the Cooperative Research Center for Premium Quality Wool.

REFERENCES

Andersen B, Schonemann MD, Flynn SE, Pearse RVd, Singh H, Rosenfeld MG: Skn-1a and Skn-1i: two functionally distinct October-2-related factors expressed in epidermis [published erratum appears in *Science* 262 (5139):1499, 1993]. *Science* 260:78–82, 1993

Burglin TR, Finney M, Coulson A, Ruvkun G: Caenorhabditis elegans has scores of homeobox-containing genes. *Nature* 341:239–243, 1989

Cillo C, Cantile M, Mortarini R, Barba P, Parmiani G, Anichini A: Differential patterns of HOX gene expression are associated with specific integrin and ICAM profiles in clonal populations isolated from a single human melanoma metastasis. *Int J Cancer* 66:692–697, 1996

Cotsarelis G, Sun T-T, Lavker RM: Label-retaining cells reside in the bulge area of the pilosebaceous unit: implications for follicular stem cells, hair cycle and skin carcinogenesis. *Cell* 61:1329–1337, 1990

Edelman GM, Jones FS: Gene regulation of cell adhesion: a key step in neural morphogenesis. *Brain Res Brain Res Rev* 26:337–352, 1998

Frohman MA, Dush MK, Martin GR: Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* 85:8998–9002, 1988

Fujita M, Furukawa F, Fujii K, Horiguchi Y, Takeichi M, Imamura S: Expression of cadherin cell adhesion molecules during human skin development: morphogenesis of epidermis, hair follicles and eccrine sweat ducts. *Arch Dermatol Res* 284:159–166, 1992

Godwin AR, Capecchi MR: Hoxc13 mutant mice lack external hair. *Genes Dev* 12:11–20, 1998

Hardy MH: The secret life of the hair follicle. *Trends Genet* 8:55–61, 1992

Hardy MH, Vielkind U: Changing patterns of cell adhesion molecules during mouse pelage hair follicle development. 1. Follicle morphogenesis in wild-type mice. *Acta Anat* 157:169–182, 1996

Hirai Y, Nose A, Kobayashi S, Takeichi M: Expression and role of E- and P-cadherin adhesion molecules in embryonic histogenesis. II. Skin morphogenesis. *Development* 105:271–277, 1989

Jones FS, Kioussi C, Copertino DW, Kallunki P, Holst BD, Edelman GM: Barx2, a new homeobox gene of the Bar class, is expressed in neural and craniofacial structures during development. *Proc Natl Acad Sci USA* 94:2632–2637, 1997

Kanzler B, Viallet JP, Le Mouellic H, Boncinelli E, Duboule D, Dhouailly D: Differential expression of two different homeobox gene families during mouse tegument morphogenesis. *Int J Dev Biol* 38:633–640, 1994

King MW, Ndiema M, Neff AW: Anterior structural defects by misexpression of Xgbx-2 in early Xenopus embryos are associated with altered expression of cell adhesion molecules. *Dev Dyn* 212:563–579, 1998

Lincecum JM, Fannon A, Song K, Wang Y, Sassoon DA: Msh homeobox genes regulate cadherin-mediated cell adhesion and cell-cell sorting. *J Cell Biochem* 70:22–28, 1998

MacKenzie A, Leeming GL, Jowett AK, Ferguson MW, Sharpe PT: The homeobox gene Hox 7.1 has specific regional and temporal expression patterns during early murine craniofacial embryogenesis, especially tooth development *in vivo* and *in vitro*. *Development* 111:269–285, 1991

MacKenzie A, Ferguson MW, Sharpe PT: Expression patterns of the homeobox gene, Hox-8, in the mouse embryo suggest a role in specifying tooth initiation and shape. *Development* 115:403–420, 1992

Mathews CH, Detmer K, Lawrence HJ, Largman C: Expression of the Hox 2.2 homeobox gene in murine embryonic epidermis. *Differentiation* 52:177–184, 1993

Meech R, Kallunki P, Edelman GM, Jones FS: A binding site for homeodomain and Pax proteins is necessary for L1 cell adhesion molecule gene expression by Pax-6 and bone morphogenetic proteins. *Proc Natl Acad Sci USA* 96:2420–2425, 1999

Morasso MI, Markova NG, Sargent TD: Regulation of epidermal differentiation by a Distal-less homeodomain gene. *J Cell Biol* 135:1879–1887, 1996

Muller-Rover S, Paus R. Topobiology of the hair follicle: adhesion molecules as morphoregulatory signals during hair follicle morphogenesis. In: Chuong C-M, eds, *Molecular Basis of Epithelial Appendage Morphogenesis*. R.G. Landes, Austin, TX, 1998, pp 283–314

Muller-Rover S, Tokura Y, Welker P, Furukawa F, Wakita H, Takigawa M, Paus R: E- and P-cadherin expression during murine hair follicle morphogenesis and cycling. *Exp Dermatol* 8:237–246, 1999

Nguyen Ba-Charvet KT, von Boxberg Y, Godement P: The mouse homeodomain protein OTX2 regulates NCAM promoter activity. *Brain Res Mol Brain Res* 67:292–295, 1999

Noveen A, Jiang TX, Ting-Berreth SA, Chuong CM: Homeobox genes Msx-1 and Msx-2 are associated with induction and growth of skin appendages. *J Invest Dermatol* 104:711–719, 1995

Powell BC, Rogers GE: Cyclic hair-loss and regrowth in transgenic mice overexpressing an intermediate filament gene. *EMBO J* 9:1485–1493, 1990

Rieger E, Bijl JJ, van Oostveen JW, *et al*: Expression of the homeobox gene HOXC4 in keratinocytes of normal skin and epithelial skin tumors is correlated with differentiation. *J Invest Dermatol* 103:341–346, 1994

Siegler MV, Jia XX: Engrailed negatively regulates the expression of cell adhesion molecules connectin and neuroglian in embryonic Drosophila nervous system. *Neuron* 22:265–276, 1999

Smith DM, Tabin CJ: Chick barx2b, a marker for myogenic cells also expressed in branchial arches and neural structures. *Mech Dev* 80:203–206, 1999

Stelnicki EJ, Komuves LG, Kwong AO, *et al*: HOX homeobox genes exhibit spatial and temporal changes in expression during human skin development. *J Invest Dermatol* 110:110–115, 1998

Tissier-Seta JP, Mucchielli ML, Mark M, Mattei MG, Goridis C, Brunet JF: Barx1, a new mouse homeodomain transcription factor expressed in cranio-facial ectomesenchyme and the stomach. *Mech Dev* 51:3–15, 1995

Wang Y, Jones FS, Krushel LA, Edelman GM: Embryonic expression patterns of the neural cell adhesion molecule gene are regulated by homeodomain binding sites. *Proc Natl Acad Sci USA* 93:1892–1896, 1996

Bibliography

Acampora, D., M. D'Esposito, A. Faiella, M. Pannese, E. Migliaccio, F. Morelli, A. Stornaiuolo, V. Nigro, A. Simeone, and E. Boncinelli. 1989. The human HOX gene family. *Nucleic Acids Res.* 17:10385-10402.

Adams, M.D., M. Dubnick, A.R. Kerlavage, R. Moreno, and J.M. Kelley. 1992. Sequence identification of 235 human brain genes. *Nature.* 355:632-634.

Akam, M. 1989. Hox and HOM: homologous gene clusters in insects and vertebrates. *Cell.* 57:347-349.

Andersen, B., M.D. Schonemann, R.D. Pearse, K. Jenne, J. Sugarman, and M.G. Rosenfeld. 1993a. Brn-5 is a divergent POU domain factor highly expressed in layer IV of the neocortex. *J Biol Chem.* 268:23390-23398.

Andersen, B., M.D. Schonemann, S.E. Flynn, R.V. Pearse, H. Singh, and M.G. Rosenfeld. 1993b. Skn-1a and Skn-1i: Two functionally distinct Oct-2-related factors expressed in epidermis. *Science.* 260:78-81.

Arcioni, L., A. Simeone, S. Guazzi, V. Zappavigna, E. Boncinelli, and F. Mavilio. 1992. The upstream region of the human homeobox gene *HOX3D* is a target for regulation by retinoic acid and HOX homeoproteins. *EMBO J.* 11:265-277.

Bates, E.J., P.I. Hynd, N.M. Penn, and M.J. Nancarrow. 1997. Serum-free culture of wool follicles: effects of nutrients, growth factors and hormones. *Br J Dermatol.* 137:498-505.

Beachy, P.A., M.A. Krasnow, E.R. Gavis, and D.S. Hogness. 1988. An *Ultrabithorax* protein binds sequences near its own and the *Antennapedia* P1 promoters. *Cell.* 55:1069-1081.

Beeman, R.W., J.J. Stuart, M.S. Haas, and R.E. Denell. 1989. Genetic analysis of the homeotic gene complex (HOM-C) in the beetle *Tribolium castaneum. Dev Biol.* 133:196-209.

Benson, G.V., T.H. Nguyen, and R.L. Maas. 1995. The expression pattern of the murine Hoxa-10 gene and the sequence recognition of its homeodomain reveal specific properties of Abdominal B-like genes. *Mol Cell Biol.* 15:1591-1601.

Benton, W.D., and R.W. Davis. 1977. Screening lambda gt recombinant clones by hybridization to single plaques *in situ. Science.* 196:180-182.

Bieberich, C.J., F.H. Ruddle, and K.S. Stenn. 1991. Differential expression of the Hox-3.1 gene in adult mouse skin. *PNAS.* 642:346-354.

Blessing, M., H. Zentgraf, and J.L. Jorcano. 1987. Differentially expressed bovine cytokeratin genes. Analysis of gene linkage and evolutionary conservation of 5'-upstream sequences. *EMBO J.* 6:567-575.

Boulikas, T. 1993. Homeodomain protein binding sites, inverted repeats, and nuclear matrix attachment regions along the human *beta-globin* gene complex. *J Cell Biochem.* 52:23-36.

Bridges, C.B., and T.H. Morgan. 1923. The third chromosome group of mutant characters of *Drosophila melanogaster. Carnegie Inst. Wash. Pub. No. 327, 93.*

Bullock, W.E., and S.D. Wright. 1987. Role of the adherence-promoting receptors, CR3, LFA-1, and p150,95, in binding of Histoplasma capsulatum by human macrophages. *J Exp Med.* 165:195-210.

Burglin, T.R., and G. Ruvkun. 1992. New motif in *PBX* genes [letter]. *Nat Genet.* 1:319-320.

Burglin, T.R., M. Finney, A. Coulson, and G. Ruvkun. 1989. Caenorhabditis elegans has scores of homoeobox-containing genes. *Nature.* 341:239-243.

Butcher, E.C., and L.J. Picker. 1996. Lymphocyte homing and homeostasis. *Science.* 272:60-66.

Byrne, C., and E. Fuchs. 1993. Probing keratinocyte and differentiation specificity of the human *K5* promoter *in vitro* and in transgenic mice. *Mol Cel Biol.* 13:3176-3190.

Carrasco, A.E., W. McGinnis, W.J. Gehring, and E.M. De Robertis. 1984. Cloning of a *Xenopus laevis* gene expressed during early embryogenesis that codes for a peptide region homologous to *Drosophila* homeotic genes. *Cell.* 37:409-414.

Catron, K.M., N. Iler, and C. Abate. 1993. Nucleotides flanking a conserved TAAT core dictate the DNA binding specificity of three murine homeodomain proteins. *Mol Cell Biol.* 13:2354-2365.

Cheng, J., K. Turksen, Q.-C. Yu, H. Schreiber, M. Teng, and E. Fuchs. 1992. Cachexia and graft-vs-host disease-type skin changes in keratin promoter-driven TNFα transgenic mice. *Genes Dev.* 6:1444-1456.

Chevallier, A., M. Kieny, A. Mauger, and P. Sengel. 1977. Vertebrate limb and somite morphogenesis. Cambridge: Cambridge University Press. 421-432.

Chisaka, O., and M.R. Capecchi. 1991. Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene *hox-1.5. Nature.* 350:473-479.

Chisaka, O., T.S. Musci, and M.R. Capecchi. 1992. Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene *Hox-1.6. Nature.* 355:516-520.

Cho, K.W., J. Goetz, C.V. Wright, A. Fritz, J. Hardwicke, and R.E. De. 1988. Differential utilization of the same reading frame in a *Xenopus* homeobox gene encodes two related proteins sharing the same DNA-binding specificity. *Embo J.* 7:2139-2149.

Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. *Anal Biochem.* 162:156-159.

Chou, M.Y., A.L. Chang, J. McBride, B. Donoff, G.T. Gallagher, and D.T. Wong. 1990. A rapid method to determine proliferation patterns of normal and malignant tissues by *H3* mRNA in situ hybridization. *Am J Pathol.* 136:729-733.

Chuong, C.M. 1998. Molecular basis of epithelial appendage morphogenesis. R. G. Landes Company, Austin. 1-444 .

Chuong, C.-M., G. Oliver, S.A. Ting, B.G. Jegalian, H.M. Chen, and E.M.D. Robertis. 1990. Gradients of homeoproteins in developing feather buds. *Develop.* 110:1021-1030.

Chuong, C.M., R.B. Widelitz, and T.X. Jiang. 1993. Adhesion molecules and homeoproteins in the phenotypic determination of skin appendages. *J Invest Dermatol.* 101:10S-15S.

Chuong, C.M., R.B. Widelitz, S. Ting-Berreth, and T.X. Jiang. 1996. Early events during avian skin appendage regeneration: dependence on epithelial-mesenchymal interaction and order of molecular reappearance. *J Invest Dermatol.* 107:639-646.

Cianetti, L., A. Dicristofaro, V. Zappavigna, L. Bottero, G. Boccoli, U. Testa, G. Russo, E. Boncinelli, and C. Peschle. 1990. Molecular mechanisms underlying the expression of the Human *Hox-5.1*-gene. *Nucl Acids Res.* 18:4361-4368.

Cillo, C., M. Cantile, R. Mortarini, P. Barba, G. Parmiani, and A. Anichini. 1996. Differential patterns of HOX gene expression are associated with specific integrin and ICAM profiles in clonal populations isolated from a single human melanoma metastasis. *Int J Cancer.* 66:692-697.

Coletta, P.L., S.M. Shimeld, C. Chaudhuri, U. Muller, J.P. Clarke, and P.T. Sharpe. 1991. Characterisation of the murine *Hox-3.3* gene and its promoter. *Mech Dev.* 35:129-142.

Conway, J.F., and D.A.D. Parry. 1988. Intermediate filament structure: 3. Analysis of sequence homologies. *Int. J. Biol. Macromol.* 10:79-98.

Cotsarelis, G., T.-T. Sun, and R.M. Lavker. 1990. Label-retaining cells reside in the bulge area of the pilosebaceous unit: implications for follicular stem cells, hair cycle and skin carcinogenesis. *Cell.* 61:1329-1337.

Coulombe, P.A., R. Kopan, and E. Fuchs. 1989. Expression of keratin K14 in the epidermis and hair follicle: insights into complex programs of differentiation. *J Cell Biol.* 109:2295-2312.

Cox, K.H., D.V. DeLeon, L.M. Angerer, and R.C. Angerer. 1984. Detection of mRNAs in sea urchin embryos by *in situ* hybridization using asymmetric RNA probes. *Dev Biol.* 101:485-502.

Dagert, M., and S.D. Ehrlich. 1979. Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene.* 6:23-28.

Davidson, D. 1995. The function and evolution of *Msx* genes: Pointers and paradoxes. *Trends Genet.* 11:405-411.

Deguchi, Y., A. Kirschenbaum, and J.H. Kehrl. 1992. A diverged homeobox gene is involved in the proliferation and lineage commitment of human hematopoietic progenitors and highly expressed in acute myelogenous leukemia. *Blood.* 79:2841-2848.

Deschamps, J., and F. Meijlink. 1992. Mammalian homeobox genes in normal development and neoplasia. *Crit Rev Oncog.* 3:117-173.

Desplan, C., J. Theis, and P.H. O'Farrell. 1988. The sequence specificity of homeodomain-DNA interaction. *Cell.* 54:1081-1090.

Detmer, K., H.J. Lawrence, and C. Largman. 1993. Expression of class I homeobox genes in fetal and adult murine skin. *J Invest Dermat.* 101:517-522.

Dhouailly, D., C. Xu, M. Manabe, A. Schermer, and T.-T. Sun. 1989. Expression of hair-related keratins in a soft epithelium: subpopulations of human and mouse dorsal tongue keratinocytes express keratin markers for hair-, skin-, and esophageal-types of differentiation. *Expl Cell Res.* 181:141-158.

Dollé, P., B.J. Izpisüa, H. Falkenstein, A. Renucci, and D. Duboule. 1989. Coordinate expression of the murine *Hox-5* complex homoeobox-containing genes during limb pattern formation. *Nature*. 342:767-772.

Driever, W., and C. Nusslein-Volhard. 1989. The bicoid protein is a positive regulator of hunchback transcription in the early Drosopihila embryo. *Nature*. 337:138-143.

Duboule, D. 1992. The vertebrate limb: a model system to study the Hox/HOM gene network during development and evolution. *Bioessays*. 14:375-384.

Duboule, D. 1994. Guidebook to the homeobox genes. Oxford University Press Inc, New York, Geneva. 1-284.

Duboule, D. 1998. *Hox* is in the hair: A break in colinearity? *Genes Dev*. 12:1-4.

Duboule, D., A. Baron, P. Mahl, and B. Galliot. 1986. A new homeo-box is present in overlapping cosmid clones which define the mouse *Hox-1* locus. *Embo J*. 5:1973-1980.

Duboule, D., and G. Morata. 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. *Trend Gen*. 10:358-364.

Dunn, S.M., R.A. Keough, G.E. Rogers, and B.C. Powell. 1998. Regulation of a hair follicle keratin intermediate filament gene promoter. *J Cell Sci*. 111:3487-3496.

Ebling, F.J. 1988. The hair cycle and its regulation. *Clin Dermatol*. 6:67-73.

Edelman, G.M., and F.S. Jones. 1998. Gene regulation of cell adhesion: a key step in neural morphogenesis. *Brain Res Brain Res Rev*. 26:337-352.

Edelman D.B., Meech R, and F.S. Jones. 2000. The homeodomain protein Barx2 contains activator and repressor domains and interacts with members of the CREB family. *J Biol Chem*. 28:331737-1745.

Fainsod, A., L.D. Bogarad, T. Ruusala, M. Lubin, D.M. Crothers, and F.H. Ruddle. 1987. The homeo domain of a murine protein binds 5' to its own homeo box. *PNAS*. 83:9532-9536.

Fisher, C., M.R. Byers, M.J. Iadrola, and E.A. Powers. 1991. Patterns of epithelial expression of Fos protein suggest an important role in the transition from viable to cornified cell during keratinization. *Develop*. 111:253-258.

Frain, M., G. Swart, O. Moaci, A. Nicosia, S. Stampfli, R. Frank, and R. Cortese. 1989. The liver-specific transcription factor LFB1 contains a highly diverged homeobox DNA binding domain. *Cell*. 59:145-157.

Fratini, A., B.C. Powell, P.I. Hynd, R.A. Keough, and G.E. Rogers. 1994. Dietary cysteine regulates the levels of mRNAs encoding a family of cysteine-rich proteins of wool. *J Invest Dermatol*. 102:178-185.

French, P.W., and D.R. Hewish. 1986. Localization of low-sulphur keratin proteins in the wool follicle using monoclonal antibodies. *J Cell Biol*. 102:1412-1418.

Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation gene *paired* and the *Drosophila PRD* gene set as part of a gene network. *Cell*. 47:735-746.

Frohman, M.A. 1994. On beyond classic RACE (Rapid Amplidication of cDNA Ends). *PCR Meth Appl.* 4:S40-S58.

Frohman, M.A., M.K. Dush, and G.R. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *PNAS.* 85:8998-9002.

Galliot, B., P. Dollé, M. Vigneron, M.S. Featherstone, A. Baron, and D. Duboule. 1989. The mouse *Hox-1.4* gene: primary structure, evidence for promoter activity and expression during development. *Develop.* 107:343-359.

Garber, R.L., A. Kuroiwa, and W.J. Gehring. 1983. Genomic and cDNA clones of the homeotic locus *Antennapedia* in *Drosophila. Embo J.* 2:2027-2036.

Gardiner, D.M., and S.V. Bryant. 1996. Molecular mechanisms in the control of limb regeneration - the role of homeobox genes. *Int J Devel Biol.* 40:797-805.

Geada, A.M., S.J. Gaunt, M. Azzawi, S.M. Shimeld, J. Pearce, and P.T. Sharpe. 1992. Sequence and embryonic expression of the murine *Hox-3.5* gene. *Develop.* 116:497-506.

Gehring, W.J., M. Muller, M. Affolter, S.A. Percival, M. Billeter, Y.Q. Qian, G. Otting, and K. Wuthrich. 1990. The structure of the homeodomain and its functional implications. *Trends Genet.* 6:323-329.

Gilbert, S.F. 1994. Developmental Biology. Sinauer.

Gillespie, J.M. 1990. The proteins of hair and other hard α-keratins. *In* Cellular and Molecular Biology of Intermediate Filaments. R.D. Goldman and P.M. Steinert, editors. Plenum Press, New York. 95-128.

Gillespie, J.M. 1991. The structural proteins of hair: isolation, characterization and regulation of biosynthesis. *In* Physiology, Biochemistry and Molecular Biology of the Skin. Vol. I. L.A. Goldsmith, editor. Oxford Uni Press, Oxford. 625-659.

Godwin, A.R., and M.R. Capecchi. 1998. Hoxc13 mutant mice lack external hair. *Genes Dev.* 12:11-20.

Goutte, C., and A.D. Johnson. 1988. A1 protein alters the DNA binding specificity of α-2 repressor. *Cell.* 52:875-882.

Graham, A. 1994. The *Hox* code out on a limb. *Curr Biol.* 4:1135-1137.

Graham, A., I. Heyman, and A. Lumsden. 1993. Even-numbered rhombomeres control the apoptotic elimination of neural crest cells from odd-numbered rhombomeres in the chick hindbrain. *Develop.* 119:233-245.

Gross, M.K., and P. Gruss. 1994. Functional analysis of mouse Hoxa-7 in Saccharomyces cerevisiae: sequences outside the homeodomain base contact zone influence binding and activation. *Mol Cell Biol.* 14:238-254.

Guazzi, S., M. Price, M. De Felice, G. Damante, M. Mattei, and R. Di Lauro. 1990. Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. *EMBO J.* 9:3631-3639.

Gutman, A., J. Gilthorpe, and P.G.J. Rigby. 1994. Multiple positive and negative regulatory elements in the promoter of the mouse homeobox gene *Hoxb-4*. *Mol Cell Biol.* 14:8143-8154.

Hanahan, D., and M. Meselson. 1983. Plasmid screening at high colony density. *Methods Enzymol.* 100:333-342.

Hansen, A.J., B.K. May, and L.A. Elferink. 1989. Sequence of a chicken phenobarbital-inducible cytochrome P450 cDNA: regulation of two P450 mRNAs transcribed from different genes [published erratum appears in DNA 1990 May;9(4):301]. *DNA.* 8:179-191.

Hardy, M.H. 1952. The histochemistry of hair follicles in the mouse. *Am J Anat.* 90:285-337.

Hardy, M.H. 1992. The secret life of the hair follicle. *Trend Gen.* 8:55-62.

Hardy, M.H., and A.G. Lyne. 1956. The pre-natal development of wool follicles in merino sheep. *Australian J Biol Sci.* 6:423-441.

Hardy, M.H., and U. Vielkind. 1996. Changing patterns of cell adhesion molecules during mouse pelage hair follicle development. 1. Follicle morphogenesis in wild-type mice. *Acta Anat.* 157:169-182.

Hatzfield, M., and W. Franke. 1985. Pair formation and promiscuity of cytokeratins: Formation *in vitro* of heterotypic complexes and intermediate filaments by homologous and heterologous recombinations of purified polypeptides. *J. Cell. Biol.* 101:1826-1841.

Hayashi, S., and M.P. Scott. 1990. What determines the specificity of action of *Drosophila* homeodomain proteins? :published erratum appears in Cell 1991 Mar 8;64 :following 1046:. *Cell.* 63:883-894.

Hazelrigg, T., and T.C. Kaufman. 1983. Revertants of dominant mutations associated with the *Antennapedia* gene complex of *Drosophila melanogaster*. *Cytol Gen.* 105:581-600.

Heid, H.W., I. Moll, and W.W. Franke. 1988a. Patterns of expression of trichocytic and epithelial cytokeratins in mammalian tissues. II. Concomittant and mutually exclusive synthesis of trichocytic and epithelial cytokeratins in diverse human and bovine tissues (hair follicle, nail bed and matrix, lingual papilla and thymic reticulum). *Differentiation.* 37:215-230.

Heid, H.W., I. Moll, and W.W. Franke. 1988b. Patterns of expression of trichocytic and epithelial cytokeratins in mammalian tissues. I. Human and bovine hair follicles. *Differentiation.* 37:137-157.

Helgason, C.D., G. Sauvageau, H.J. Lawrence, C. Largman, and R.K. Humphries. 1996. Overexpression of HOXB4 enhances the hematopoietic potential of embryonic stem cells differentiated in vitro. *Blood.* 87:2740-2749.

Herr, W., R.A. Sturm, R.G. Clerc, L.M. Corcoran, D. Baltimore, P.A. Sharp, H.A. Ingraham, M.G. Rosenfeld, M. Finney, G. Ruvkun, and H.R. Horvitz. 1988. The POU domain: a large conserved region in the mammalian *pit-1*, *oct-1*, *oct-2*, and *Caenorhabditis elegans unc-86* gene products. *Genes Dev.* 2:1513-1516.

Higashijima, S., T. Michiue, Y. Emori, and K. Saigo. 1992. Subtype determination of *Drosophila* embryonic external sensory organs by redundant homeobox genes *BarH1* and *BarH2*. *Genes Dev.* 6:1005-1018.

Hjalt, T.A., and J.C. Murray. 1999. The human *BARX2* gene: genomic structure, chromosomal localization, and single nucleotide polymorphisms. *Genomics.* 62:456-459.

Hjalt, T.A., E.V. Semina, B.A. Amendt, and J.C. Murray. 2000. The Pitx2 protein in mouse development. *Dev Dyn.* 18:195-200.

Hoey, T., and M. Levine. 1988. Divergent homeobox proteins recognize similar DNA sequences in *Drosophila*. *Nature.* 332:858-861.

Hoey, T., R. Warrior, J. Manak, and M. Levine. 1988. DNA-binding activities of the *Drosophila melanogaster* even-skipped protein are mediated by its homeodomain and influenced by protein context. *Mol Cell Biol.* 8:4598-4607.

Holland, P.W.H., and B.L.M. Hogan. 1988. Expression of homeobox genes during mouse development: a review. *Genes Dev.* 2:773-782.

Hunt, P., and R. Krumlauf. 1992. HOX codes and positional specification in vertebrate embryonic axes. *Ann Rev Cell Biol.* 8:227-256.

Hunter, T., and M. Karin. 1992. The regulation of transcription by phosphorylation. *Cell.* 70:375-385.

Ingraham, H.A., R.P. Chen, H.J. Mangalam, H.P. Elsholtz, S.E. Flynn, C.R. Lin, D.M. Simmons, L. Swanson, and M.G. Rosenfeld. 1988. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. *Cell.* 55:519-529.

Izpisua-Belmonte, J.C., C. Tickle, P. Dolle, L. Wolpert, and D. Duboule. 1991. Expression of the homeobox *Hox-4* genes and the specification of position in chick wing development. *Nature.* 350:585-589.

Izpisua-Belmonte, J.C., H. Falkenstein, P. Dolle, A. Renucci, and D. Duboule. 1991. Murine genes related to the *Drosophila Abd-B* homeotic genes are sequentially expressed during development of the posterior part of the body. *EMBO J.* 10:2279-2289.

Jiang, T.X., Y.H. Liu, R.B. Widelitz, R.K. Kundu, R.E. Maxson, and C.M. Chuong. 1999. Epidermal dysplasia and abnormal hair follicles in transgenic mice overexpressing homeobox gene *MSX-2*. *J Invest Dermatol.* 113:230-237.

Jones, F.S., B.D. Holst, O. Minowa, E.M. De Robertis, and G.M. Edelman. 1993. Binding and transcriptional activation of the promoter for the neural cell adhesion molecule by HoxC6 (Hox-3.3). *PNAS.* 90:6557-6561.

Jones, F.S., C. Kioussi, D.W. Copertino, P. Kallunki, B.D. Holst, and G.M. Edelman. 1997. *Barx2*, a new homeobox gene of the *Bar* class, is expressed in neural and craniofacial structures during development. *PNAS.* 94:2632-2637.

Jones, F.S., E.A. Prediger, D.A. Bittner, E.M. De Robertis, and G.M. Edelman. 1992. Cell adhesion molecules as targets for *Hox* genes: neural cell adhesion molecule promoter activity is modulated by cotransfection with Hox-2.5 and -2.4. *PNAS.* 89:2086-2090.

Jones, F.S., G. Chalepakis, P. Gruss, and G.M. Edelman. 1992. Activation of the cytotactin promoter by the homeobox-containing gene *Evx-1*. *PNAS*. 89:2091-2095.

Jowett, A.K., S. Vainio, M.W. Ferguson, P.T. Sharpe, and I. Thesleff. 1993. Epithelial-mesenchymal interactions are required for *msx 1* and *msx 2* gene expression in the developing murine molar tooth. *Develop.* 117:461-470.

Joyner, A.L., K. Herrup, B.A. Auerbach, C.A. Davis, and J. Rossant. 1991. Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the *En-2* homeobox. *Science*. 251:1239-1243.

Kansas, G.S. 1996. Selectins and their ligands: current concepts and controversies. *Blood*. 88:3259-3287.

Kanzler, B., F. Prin, J. Thelu, and D. Dhouailly. 1997. *CHOXC-8* and *CHOXD-13* expression in embryonic chick skin and cutaneous appendage specification. *Dev Dyn*. 10:274-287.

Kanzler, B., J.P. Viallet, H.L. Mouellic, E. Boncinelli, D. Duboule, and D. Dhouailly. 1994. Differential expression of two different homeobox gene families during mouse tegument morphogenesis. *Int J Devel Biol*. 38:633-640.

Kappen, C., K. Schughart, and F.H. Ruddle. 1989. Two steps in the evolution of *Antennapedia*-class vertebrate homeobox genes. *PNAS*. 86:5459-5463.

Kappen, C., K. Schughart, and F.H. Ruddle. 1993. Early evolutionary origin of major homeodomain sequence classes. *Genomics*. 18:54-70.

Karch, F., B. Weiffenbach, M. Peifer, W. Bender, I. Duncan, S. Celniker, M. Crosby, and E.B. Lewis. 1985. The abdominal region of the bithorax complex. *Cell*. 43:81-96.

Karg, H., E.H. Burger, D.M. Lyaruu, A.L. Bronckers, and J.H. Woltgens. 1997. Spatiotemporal expression of the homeobox gene *S8* during mouse tooth development. *Arch Oral Biol*. 42:625-631.

Kaytes, P., A.R. McNab, T.J. Rea, V. Groppi, T.T. Kawabe, A.E. Buhl, A.P. Bertlino, N.T. Hatzenbuhler, and G. Vogeli. 1991. Hair-specific keratins: Characterization and expression of a mouse type I keratin gene. *J Invest Dermat*. 97:835-842.

Kemp, D.J., and G.E. Rogers. 1970. Immunological and immunofluorescent studies on keratins of the hair follicle. *J Cell Sci*. 7:273-283.

Keough, R., B. Powell, and G. Rogers. 1995. Targeted expression of SV40 T antigen in the hair follicle of transgenic mice produces an aberrant hair phenotype. *J Cell Sci*. 108:957-966.

Kessel, M., and P. Gruss. 1990. Murine developmental control genes. *Science*. 249:374-379.

King, M.W., M. Ndiema, and A.W. Neff. 1998. Anterior structural defects by misexpression of Xgbx-2 in early Xenopus embryos are associated with altered expression of cell adhesion molecules. *Dev Dyn*. 212:563-579.

Kissinger, C.R., B. Liu, E. Martin-Blanco, T.B. Kornberg, and C.O. Pabo. 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: A framework for understanding homeodomain-DNA interactions. *Cell*. 63:579-590.

Krasner, D., Wallace L, Thiagalingam A, Jones C, Lengauer C, Minahan L, Ma Y, Kalikin L, Feinberg AP, Jabs EW, Tunnacliffe A, Baylin SB, Ball DW, and N. BD. 2000. Cloning and chromosomal localization of the human *BARX2* homeobox protein gene. *Gene.* 250:171-180.

Krieg, P.A., and D.A. Melton. 1987. *In vitro* RNA synthesis with SP6 RNA polymerase. *Methods Enzymol.* 155:397-415.

Krumlauf, R. 1993. Hox genes and pattern formation in the branchial region of the vertebrate head. *Trends Genet.* 9:106-112.

Krushel, L.A., and D. Van Der Kooy. 1993. Pattern formation in the developing mammalian forebrain: selective adhesion of early but not late postmitotic cortical and striatal neurons within forebrain reaggregate cultures. *Dev Biol.* 158:145-162.

Kurisu, K., and M.J. Tabata. 1997. Human genes for dental anomalies. *Oral Dis.* 3:223-228.

Kuzuoka, M., T. Takahashi, C. Guron, and R. Raghow. 1994. Murine homeobox-containing gene, *Msx-1*: analysis of genomic organization, promoter structure, and potential autoregulatory cis- acting elements. *Genomics.* 21:85-91.

Laughon, A. 1991. DNA binding specificity of homeodomains. *Biochem.* 30:11357-11367.

Lawrence, H.J., G. Sauvageau, R.K. Humphries, and C. Largman. 1996. The role of HOX homeobox genes in normal and leukemic hematopoiesis. *Stem Cells.* 14:281-291.

Le Calvez, J. 1948. SSAr: Mutation *Aristapedia,* heterozygote dominante, homozygote lethal chez *Drosophila melanogaster. Bull. Biol. France. Belg.* 82:97-113.

Leask, A., C. Byrne, and E. Fuchs. 1991. Transcription factor AP2 and its role in epidermal-specific gene expression. *PNAS.* 88:7948-7952.

Leask, A., M. Rosenberg, R. Vassar, and E. Fuchs. 1990. Regulation of a human epidermal keratin gene: sequences and nuclear factors involved in keratinocyte-specific transcription. *Genes and Develop.* 4:1985-1998.

Lendahl, U., L.B. Zimmerman, and R.D.G. McKay. 1990. CNS stem cells express a new class of intermediate filament protein. *Cell.* 60:585-595.

Lewis, E.B. 1978. A gene complex controlling segmentation in *Drosophila. Nature.* 276:565-570.

Lezot, F., B. Thomas, D. Hotton, N. Forest, S. Orestes-Cardoso, B. Robert, P. Sharpe, and A. Bersdal. 2000. Biomineralization, life-time of odontogenic cells and differential expression ofthe two homeobox genes *MSX-1* and *DLX-2* in transgenic mice. *J Bone Miner Res.* 15:430-431

Lincecum, J.M., A. Fannon, K. Song, Y. Wang, and D.A. Sassoon. 1998. *Msh* homeobox genes regulate cadherin-mediated cell adhesion and cell- cell sorting. *J Cell Biochem.* 70:22-28.

Lu, S., T.L. Wise, and F.H. Ruddle. 1994. Mouse homeobox gene *Dbx*: sequence, gene structure and expression pattern during mid-gestation. *Mech Dev.* 47:187-195.

Lynch, M.H., W.M. O'Guin, C. Hardy, L. Mak, and T.-T. Sun. 1986. Acidic and basic nail/hair ("hard") keratins: their colocalization in upper cortical and cuticle cells of the human hair follicle and their relationship to "soft" keratins. *J Cell Biol.* 103:2593-2606.

Mackenzie, A., M.W.J. Ferguson, and P.T. Sharpe. 1992. Expression patterns of the homeobox gene, *Hox-8*, in the mouse embryo suggest a role in specifying tooth initiation and shape. *Develop.* 115:403-420.

MacKinnon, P.J., B.C. Powell, and G.E. Rogers. 1990. Structure and expression of genes for a class of cysteine-rich proteins of the cuticle layers of differentiating wool and hair follicles. *J Cell Biol.* 111:2587-2600.

Manak, J.R., and M.P. Scott. 1993. Able assistants for homeodomain proteins. *Curr Biol.* 3:318-320.

Mathews, C.H., K. Detmer, H.J. Lawrence, and C. Largman. 1993. Expression of the *Hox 2.2* homeobox gene in murine embryonic epidermis. *Differentiation.* 52:177-184.

McGinnis, W., and R. Krumlauf. 1992. Homeobox genes and axial patterning. *Cell.* 68:283-302.

McGinnis, W., M. Levine, E. Hafen, A. Kuroiwa, and W.J. Gehring. 1984. A conserved DNA sequence found in homeotic genes of the *Drosophila Antennapedia* and *Bithorax* complexes. *Nature.* 308:428-433.

McKnight, S.L., E.R. Gavis, R. Kingsbury, and R. Axel. 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. *Cell.* 25:385-398.

Meech, R., P. Kallunki, G.M. Edelman, and F.S. Jones. 1999. A binding site for homeodomain and Pax proteins is necessary for L1 cell adhesion molecule gene expression by Pax-6 and bone morphogenetic proteins. *PNAS.* 96: 2420-2425.

Mendel, D.B., and G.R. Crabtree. 1991. HNF-1, a member of a novel class of dimerizing homeodomain proteins. *J Biol Chem.* 266:677-680.

Mercer, E.H. 1961. Keratin and Keratinization. Pergamon Press, New York.

Messing, J., R. Crea, and P.H. Seeburg. 1981. A system for shotgun DNA sequencing. *Nucleic Acids Res.* 9:309-321.

Mizobuchi, M., and L.A. Frohman. 1992. A rapid and simple method for labeling short DNA fragments using Taq polymerase. *Biotechniques.* 12:350-354.

Mlodzik, M., A. Fjose, and W.J. Gehring. 1988. Molecular structure and spatial expression of a homeobox gene from the labial region of the *Antennapedia*-complex. *Embo J.* 7:2569-2578.

Monaghan, A.P., D.R. Davidson, C. Sime, E. Graham, R. Baldock, S.S. Bhattacharya, and R.E. Hill. 1991. The *Msh*-like homeobox genes define domains in the developing vertebrate eye. *Develop.* 112:1053-1061.

Morasso, M.I., N.G. Markova, and T.D. Sargent. 1996. Regulation of epidermal differentiation by a *Distal-less* homeodomain gene. *J Cell Biol.* 135:1879-1887.

Moscoso, L.M., and J.R. Sanes. 1995. Expression of four immunoglobulin superfamily adhesion molecules (L1, Nr-CAM/Bravo, neurofascin/ABGP, and N-CAM) in the developing mouse spinal cord. *J Comp Neurol.* 352:321-334.

Muller-Immergluck, M.M., S. Ruppert, W. Schaffner, and P. Matthias. 1988. A cloned transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature.* 336:544-551.

Muller-Rover, S., E.J. Peters, V.A. Botchkarev, A. Panteleyev, and R. Paus. 1998. Distinct patterns of NCAM expression are associated with defined stages of murine hair follicle morphogenesis and regression. *J Histochem Cytochem.* 46:1401-1410.

Nakanishi, H. 1975. Experimental gut anastomoses and their revascularization. *Aust N Z J Surg.* 45:309-314.

Nguyen Ba-Charvet, K.T.V., Y. Von Boxberg, and P. Godement. 1999. The mouse homeodomain protein OTX2 regulates *NCAM* promoter activity. *Brain Res Mol Brain Res.* 67:292-295.

Nieminen, P., S. Arte, S. Pirinen, L. Peltonen, and I. Thesleff. 1995. Gene defect in hypodontia: Exclusion of *MSX1* and *MSX2* as candidate genes. *Human Genetics.* 96:305-308.

Nixon, A.J. 1993. A method for determining the activity state of hair follicles. *Biotech Histochem.* 68:316-325.

Nohno, T., S. Noji, E. Koyama, K. Ohyama, F. Myokai, A. Kuroiwa, T. Saito, and S. Taniguchi. 1991. Involvement of the *Chox-4* chicken homeobox genes in determination of anteroposterior axial polarity during limb development. *Cell.* 64:1197-1205.

Noveen, A., T.X. Jiang, S.A. Ting-Berreth, and C.M. Chuong. 1995. Homeobox genes *Msx-1* and *Msx-2* are associated with induction and growth of skin appendages. *J Invest Dermatol.* 104:711-719.

Odenwald, W.F., J. Garbern, H. Arnheiter, L.E. Tournier, and R.A. Lazzarini. 1989. The Hox-1.3 homeobox protein is a sequence-specific DNA-binding phosphoprotein: published erratum appears in Genes Dev 1989 Aug;3(8):1267:. *Genes Dev.* 3:158-172.

Otting, G., Y.Q. Qian, M. Billeter, M. Muller, M. Affolter, W.J. Gehring, and K. Wuthrich. 1990. Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* 9:3085-3092.

Pabo, C.O., and R.T. Sauer. 1984. Protein-DNA recognition. *Ann Rev Biochem.* 53:293-303.

Popperl, H., and M.S. Featherstone. 1992. An autoregulatory element of the murine *Hox-4.2* gene. *EMBO J.* 11:3673-3680.

Powell, B., L.A. Crocker, and G.E. Rogers. 1992. Hair follicle differentiation: expression, structure and evolutionary conservation of the hair type II keratin intermediate filament gene family. *Develop.* 114:417-434.

Powell, B.C., A. Nesci, and G.E. Rogers. 1991. Regulation of keratin gene expression in hair follicle differentiation. *Ann. N. Y. Acad. Sci.* 642:1-20.

Powell, B.C., and G.E. Rogers. 1990. Hard keratin IF and associated proteins. *In* Cellular and Molecular Biology of Intermediate Filaments. R.D. Goldman and P.M. Steinert, editors. Plenum Press, New York. 267-300.

Powell, B.C., and G.E. Rogers. 1994. Differentiation in hard keratin tissues: hair and related structures. *In* Keratinocyte Handbook. I. Leigh, F. Watt, and E.B. Lane, editors. Cambridge University Press. 401-436.

Powell, B.C., and G.E. Rogers. 1997. The role of keratin proteins and their genes in the growth, structure and properties of hair. *Exs.* 78:59-148.

Powell, B.C., and J.S. Beltrame. 1994. Characterisation of a hair (wool) keratin intermediate filament gene domain. *J Invest Dermat.* 102:171-177.

Powell, B.C., J. Arthur, and A. Nesci. 1995. Characterization of a gene encoding a cysteine-rich keratin associated protein synthesized late in rabbit hair follicle differentiation. *Differentiation.*

Price, J.A., D.W. Bowden, J.T. Wright, M.J. Pettenati, and T.C. Hart. 1998. Identification of a mutation in DLX3 associated with tricho-dento- osseous (TDO) syndrome. *Hum Mol Genet.* 7:563-569.

Ptashne, M. 1986. A genetic switch. P. Alto, editor. Cell/Blackwell Scientific.

Ptashne, M. 1988. How eukaryotic transcriptional activators work. *Nature.* 335:683-689.

Reed, K.C., and D.A. Mann. 1985. Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res.* 13:7207-7221.

Reginelli, A.D., Y.Q. Wang, D. Sassoon, and K. Muneoka. 1995. Digit tip regeneration correlates with regions of Msx1 (Hox 7) expression in fetal and newborn mice. *Develop.* 121:1065-1076.

Regulski, M., K. Harding, R. Kostriken, F. Karch, M. Levine, and W. McGinnis. 1985. Homeobox genes of the *Antennapedia* and *bithorax* complexes of *Drosophila. Cell.* 43:71-80.

Rey-Campos, J., and M. Yaniv. 1992. Regulation of albumin gene expression. In: Genetic intervention in diseases with unknown etiology. Elsevier Science Publishers B.V. 1-14 .

Rieger, E., J.J. Bijl, J.W.V. Oostveen, H.P. Soyer, C.B.M. Oudejans, N.M. Jiwa, J.M.M. Walboomers, and C.J.L. Meijer. 1994. Expression of the homeobox gene *Hox C4* in keratinocytes of normal skin and epithelial skin tumors is correlated with differentiation. *J Invest Dermat.* 103:341-346.

Rijli, F.M., M. Mark, S. Lakkaraju, A. Dierich, P. Dolle, and P. Chambon. 1993. A homeotic transformation is generated in the rostral branchial region of the head by disruption of *Hoxa-2*, which acts as a selector gene. *Cell.* 75:1333-1349.

Robert, B., D. Sassoon, B. Jacq, W. Gehring, and M. Buckingham. 1989. *Hox-7*, a mouse homeobox gene with a novel pattern of expression during embryogenesis. *EMBO J.* 8:91-100.

Rogers, B.L., and G.F. Saunders. 1986. Transcriptional enhancers play a major role in gene expression. *Bioessays.* 4:62-65.

Rogers, M.A., L. Langbein, S. Praetzel, I. Moll, T. Krieg, H. Winter, and J. Schweizer. 1997. Sequences and differential expression of three novel human type-II hair keratins. *Differentiation*. 61:187-194.

Ros, M.A., G.E. Lyons, S. Mackem, and J.F. Fallons. 1994. Recombinant limbs as a model to study homeobox gene regulation during limb development. *Dev Biol*. 166:59-72.

Rothnagel, J.A. and D.R. Roop. 1995. Hair follicle companion layer: reacquainting an old friend. *J Invest Dermat*. 104(5 Suppl):42S-43S.

Rubin, M.R., W. King, L.E. Toth, I.S. Sawczuk, M.S. Levine, P. D'Eustachio, and M.C. Nguyen-Huu. 1987. Murine *Hox-1.7* homeo-box gene: cloning, chromosomal location, and expression [published erratum appears in Mol Cell Biol 1988 Dec;8(12):5593]. *Mol Cell Biol*. 7:3836-3841.

Ruddle, F.H., J.L. Bartels, K.L. Bentley, C. Kappen, M.T. Murtha, and J.W. Pendleton. 1994. Evolution of the Hox genes. *Ann Rev Genet*. 28:423-442.

Ruvkun, G., and M. Finney. 1991. Regulation of transcription and cell identity by POU domain proteins. *Cell*. 64:475-478.

Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. 239:487-490.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Moleuclular cloning. A laboratory manual. Second edition. Cold spring harbor laboratory press.

Sanger, F., A.R. Coulson, B.G. Barrell, A.J. Smith, and B.A. Roe. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J Mol Biol*. 143:161-178.

Sasaki, H., E. Yokoyama, and A. Kuroiwa. 1990. Specific DNA binding of the two chicken *Deformed* family homeodomain proteins, Chox-1.4 and Chox-a. *Nucleic Acids Res*. 18:1739-1747.

Sauer, R.T., D.L. Smith, and A.D. Johnson. 1988. Flexibility of the yeast alpha 2 repressor enables it to occupy the ends of its operator, leaving the centre free. *Genes Develop*. 2:807-816.

Sauvageau, G., U. Thorsteinsdottir, C.J. Eaves, H.J. Lawrence, C. Largman, P.M. Lansdorp, and R.K. Humphries. 1995. Overexpression of HOXB4 in hematopoietic cells causes the selective expansion of more primitive populations *in vitro* and *in vivo*. *Genes and Develop*. 9:1753-1765.

Schier, A.F., and W.J. Gehring. 1992. Direct homeodomain-DNA interaction in the autoregulation of the *Fushi tarazu* gene. *Nature*. 356:804-814.

Schier, A.F., and W.J. Gehring. 1993. Functional specificity of the homeodomain protein fushi tarazu: the role of DNA-binding specificity *in vivo*. *PNAS*. 90:1450-1454.

Schubert, F.R., S.K. Nieselt, and P. Gruss. 1993. The *Antennapedia*-type homeobox genes have evolved from three precursors separated early in metazoan evolution. *PNAS*. 90:143-147.

Schughart, K., D. Pravtcheva, M.S. Newman, L.W. Hunihan, Z.L. Jiang, and F.H. Ruddle. 1989. Isolation and regional localization of the murine homeobox-containing gene *Hox-3.3* to mouse chromosome region 15E. *Genomics*. 5:76-83.

Schummer, M., I. Scheurlen, C. Schaller, and B. Galliot. 1992. HOM/HOX homeobox genes are present in hydra (*Chlorohydra viridissima*) and are differentially expressed during regeneration. *Embo J.* 11:1815-1823.

Scott, G.A., and L.A. Goldsmith. 1993. Homeobox genes and skin development: a review. *J Invest Dermatol.* 101:3-8.

Scott, M.P. 1992. Vertebrate homeobox gene nomenclature [letter]. *Cell.* 71:551-553.

Scott, M.P., A.J. Weiner, T.I. Hazelrigg, B.A. Polisky, V. Pirrotta, F. Scalenghe, and T.C. Kaufman. 1983. The molecular organization of the *Antennapedia* locus of *Drosophila*. *Cell.* 35:763-776.

Sheets, M.D., S.C. Ogg, and M.P. Wickens. 1990. Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracty and efficiency of cleavage and polyadenylation *invitro*. *Nucleic Acids Res.* 18:5799-5805.

Siegler, M.V., and X.X. Jia. 1999. Engrailed negatively regulates the expression of cell adhesion molecules connectin and neuroglian in embryonic *Drosophila* nervous system. *Neuron.* 22: 265-276.

Snape, A.M., E.A. Jonas, and T.D. Sargent. 1990. KTF-1, a transcriptional activator of *Xenopus* embryonic keratin expression. *Develop.* 109:157-165.

Snape, A.M., R.S. Winning, and T.D. Sargent. 1991. Transcription factor AP-2 is tissue specific in *xenopus* and is closely related or identical to keratin transcription factor 1 (KTF-1). *Develop.* 113:283-293.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol.* 98:503-517.

Steinert, P.M., and D.R. Roop. 1988. Molecular and cellular biology of intermediate filaments. *Ann Rev Biochem.* 57:593-625.

Steinert, P.M., D.A.D. Parry, W.I. Idler, L.D. Johnson, A.C. Steven, and D.R. Roop. 1985. Amino acid sequences of mouse and human epidermal type II keratins of Mr 67,000 provide a basis for the structural and functional diversity of the end domains of keratin intermediate filament subunits. *J Biol Chem.* 260:7412-7249.

Stelnicki, E.J., L.G. Komuves, A.O. Kwong, D. Holmes, P. Klein, S. Rozenfeld, H.J. Lawrence, N.S. Adzick, M. Harrison, and C. Largman. 1998. HOX homeobox genes exhibit spatial and temporal changes in expression during human skin development. *J Invest Dermatol.* 110:110-115.

Stelnicki, E.J., L.G. Komuves, D. Holmes, W. Clavin, M.R. Harrison, N.S. Adzick, and C. Largman. 1997. The human homeobox genes *MSX-1*, *MSX-2*, and *MOX-1* are differentially expressed in the dermis and epidermis in fetal and adult skin. *Differentiation.* 62:33-41.

Stenn, K.S., S.M. Prouty, and M. Seiberg. 1994. Molecules of the cycling hair follicle - A tabulated review. *J Dermatol Sci.* 7 Suppl:S109-124.

Stern, S., M. Tanaka, and W. Herr. 1989. The Oct-1 homoeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. *Nature*. 341:624-630.

Struhl, G., K. Struhl, and P.M. Macdonald. 1989. The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell*. 57:1259-1273.

Sturm, R.A., and W. Herr. 1988. The POU domain is a bipartite DNA-binding structure. *Nature*. 336:601-604.

Tabin, C. 1995. The initiation of the limb bud: Growth factors, Hox genes and Retinoids. *Cell*. 80:671-674.

Tanda, S., and V.G. Corces. 1991. Retrotransposon-induced overexpression of a homeobox gene causes defects in eye morphogenesis in Drosophila. *Embo J*. 10:407-417.

Thesleff, I., and C. Sahlberg. 1996. Growth factors as inductive signals regulating tooth morphogenesis [Review]. *Seminars in Cell & Dev Biol* . 7:185-193.

Thomas, N.E., R. Polakowska, K.E. Rogers, E. Kelly, and L.A. Goldsmith. 1989. Homeobox containing gene found in adult human keratinocyte library. *Clin Res*. 37:595A.

Tickle, C. 1996. Vertebrate limb development [Review]. *Sem Cell Dev Biol* . 7:137-143.

Tickle, C., and G. Eichele. 1994. Vertebrate limb development. *Ann Rev Cell Biol*. 10:121-152.

Tissier-Seta, J.P., M.L. Mucchielli, M. Mark, M.G. Mattei, C. Goridis, and J.F. Brunet. 1995. Barx1, a new mouse homeodomain transcription factor expressed in cranio- facial ectomesenchyme and the stomach. *Mech Dev*. 51:3-15.

Tobiasch, E., H. Winter, and J. Schweizer. 1992. Structural features and sites of expression of a new 65kD and 48kD hair-related keratin pair associated with a special type of parakeratotic epithelial differentiation. *Differentiation*. 50:163-178.

Treisman, J., P. Gönczy, M. Vashishtha, E. Harris, and C. Desplan. 1989. A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell*. 59:553-562.

Tureckova, J., C. Sahlberg, T. Aberg, J.V. Ruch, I. Thesleff, and R. Peterkova. 1995. Comparison of expression of the *msx-1*, *msx-2*, *BMP-2* and *BMP-4* genes in the mouse upper diastemal and molar tooth primordia. *Int J Dev Biol*. 39:459-468.

Vainio, S., I. Karavanova, A. Jowett, and I. Thesleff. 1993. Identification of BMP-4 as a signal mediating secondary induction between epithelial and mesenchymal tissues during early tooth development. *Cell*. 75:45-58.

Van Genderen, C., R.M. Okamura, I. Farinas, K.G. Quo, T.G. Parslow, L. Brahn, and R. Grosschedl. 1994. Development of several organs that require inductive epithelial-mesenchymal interactions is imparied in LEF-1 deficient mice. *Genes*. 8:2691-2703.

Vassar, R., and E. Fuchs. 1991. Transgenic mice provide new insights into the role of TGF-α during epidermal development and differentiation. *Genes and Develop*. 5:714-727.

Vassar, R., M. Rosenberg, S. Ross, A. Tyner, and E. Fuchs. 1989. Tissue-specific expression and differentiation-specific expression of a human *K14* keratin gene in transgenic mice. *PNAS*. 86:1563-1567.

Vastardis, H., N. Karimbux, S.W. Guthua, J.G. Seidman, and C.E. Seidman. 1996. A human msx1 homeodomain missense mutation causes selective tooth agenesis. *Nature Genetics*. 13:417-421.

Verrijzer, C.P., J.A. Van Oosterhout, and P. Van Der Vliet. 1992. The Oct-1 POU domain mediates interactions between Oct-1 and other POU proteins. *Mol Cell Biol*. 12:542-551.

Vielkind, U., M.K. Sebzda, I.R. Gibson, and M.H. Hardy. 1995. Dynamics of Merkel cell patterns in developing hair follicles in the dorsal skin of mice, demonstrated by a monoclonal antibody to mouse keratin 8. *Acta Anat*. 152:93-109.

Wang, Y., F.S. Jones, L.A. Krushel, and G.M. Edelman. 1996. Embryonic expression patterns of the neural cell adhesion molecule gene are regulated by homeodomain binding sites. *PNAS*. 93:1892-1896.

Weiss, K.M., J. Bollekens, F.H. Ruddle, and K. Takashita. 1994. *Distal-less* and other homeobox genes in the development of the dentition. *J Exp Zool*. 270:273-284.

Whitbread, L.A. 1992. A study of sheep epithelial intermediate filament gene expression. Adelaide. PhD Thesis.

Whiting, J., H. Marshall, M. Cook, R. Krumlauf, P.W. Rigby, D. Stott, and R.K. Allemann. 1991. Multiple spatially specific enhancers are required to reconstruct the pattern of *Hox-2.6* gene expression. *Genes Dev*. 5:2048-2059.

Widelitz, R.B., T.X. Jiang, A. Noveen, S.A. Ting-Berreth, E. Yin, H.S. Jung, and C.M. Chuong. 1997. Molecular histology in skin appendage morphogenesis. *Microsc Res Tech*. 38:452-465.

Wilson, B.W., K.J. Edwards, M.J. Sleigh, C.R. Byrne, and K.A. Ward. 1988. Complete sequence of a type I-microfibrillar keratin gene. *Gene*. 73:21-31.

Winter, H., M. Rentrop, R. Nischt, and J. Schweizer. 1990. Tissue specific expression of murine keratin K13 in internal stratified squamous epithelia and its aberrant expression during two-stage mouse skin carcinogenesis is associated with the methylation state of a distinct CpG site in the remote 5'-flanking region of the gene. *Differentiation*. 43:105-114.

Winter, H., P. Siry, E. Tobiasch, and J. Schweizer. 1994. Sequence and expression of murine type I hair keratins *mHa2* and *mHa3*. *Exp Cell Res*. 212:190-200.

Wolberger, C., A.K. Vershon, B. Liu, A.D. Johnson, and C.O. Pabo. 1991. Crystal structure of a MAT alpha-2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell*. 67:517-528.

Yoon, S.-J., J. LeBlanc-Straceski, D. Ward, K. Krauter, and R. Kucherlapati. 1994. Organization of the human keratin type II gene cluster at 12q13. *Genomics*. 24:502-508.

Zappavigna, V., A. Renucci, J.C. Izpisua-Belmonte, G. Urier, C. Peschle, and D. Duboule. 1991. *HOX4* genes encode transcription factors with potential auto- and cross-regulatory capacities. *EMBO J*. 10:4177-4187.

Zappavigna, V., D. Sartori, and F. Mavilio. 1994. Specificty of Hox protein function depends on DNA protein-protein interactions, both mediated by the homeodomain. *Genes and Develop*. 8:732-744.

Zeltser, L., C. Desplan, and N. Heintz. 1996. *Hoxb-13*: a new Hox gene in a distant region of the *HOXB* cluster maintains colinearity. *Develop.* 122:2475-2484.

Zhou, P., C. Byrne, J. Jacobs, and E. Fuchs. 1995. Lymphoid enhancer factor 1 directs hair follicle patterning and epithelial cell fate. *Genes Dev.* 9:700-713.

Zimmermann, F., and I.N. Rich. 1997. Mammalian homeobox B6 expression can be correlated with erythropoietin production sites and erythropoiesis during development, but not with hematopoietic or nonhematopoietic stem cell populations. *Blood.* 89:2723-2735.