Analysis of Murine *Hox* Gene Regulation and Creation of an Insertional Mutant Displaying Polydactyly

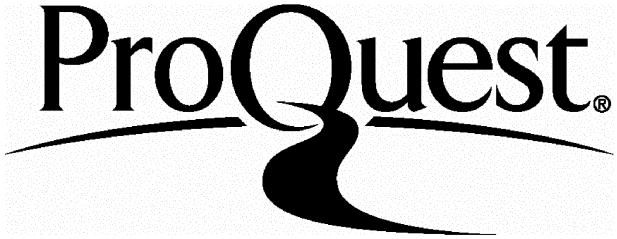
James Alexander Sharpe

Thesis for the degree of Doctor of Philosophy

December 1996

Division of Developmental Neurobiology
National Institute for Medical Research
The Ridgeway, Mill Hill,
London NW7 1AA.

University College London,
Gower Street,
London WC1E 6BT.



ABSTRACT

The vertebrate Hox genes are essential for embryo development, and are thought to specify positional information along the anteroposterior (A-P) axis. In this study, murine transgenic technology was used to further our understanding of how Hox genes are regulated, and to test a hypothesis which might explain their clustered organisation.

Two regions of the *Hoxb* complex were studied: the *Hoxb-9* region, and the intergenic stretch between *Hoxb-5* and *Hoxb-4*. In the first study, constructs were created which contained regulatory elements from near the *Hoxb-9* gene coupled to the reporter gene LacZ. These allowed identification of two regulatory regions which appear to be important for normal expression of the *Hoxb-9* gene.

Previous work on the regulation of the *Hoxb-5* gene had suggested that two regulatory elements which are probably important for *Hoxb-5* may also interact with the *Hoxb-4* promoter. In the second study, a double-reporter system was developed, with which it was possible to monitor the expression from two different promoters in the same construct. Constructs were created spanning the *Hoxb-5-Hoxb-4* region, in which the LacZ gene was inserted into the *Hoxb-5* coding region, and the human placental alkaline phosphatase (PLAP) gene was inserted into the *Hoxb-4* coding region. Using this construct, and two versions in which the regulatory regions had been deleted, it was possible to show that certain elements in this stretch of DNA may be able to activate both promoters.

During the course of these transgenic experiments, an insertional mutant was created which displayed preaxial polydactyly. In the last part of this study, the mutant strain was characterised with regard to its developmental and skeletal phenotype, the chromosomal localisation of the transgene, expression of the gene *Sonic hedgehog* in the developing limb bud, and expression of the transgenic PLAP reporter construct. Using these data the strain was compared to previously known polydactylous mutants from the hemimelia-luxate group.

ACKNOWLEDGEMENTS

Firstly, I wish to thank my external supervisor Dr. Robb Krumlauf, for his continuous support and for giving me the opportunity to work on the various projects I have been involved with. Secondly, I must thank all the members of the Krumlauf laboratory over the last four years: I was lucky enough to have not one teacher but ten. I tried to pester them all equally, and for the fact that they put up with it I am eternally grateful. Extra especial thanks go to Linda McNaughton who managed to teach me whole-mount *in-situ* hybridisation, Heather Marshall for help with all matters murine, Kamala Maruthuainer for tirelessly helping with sequencing, and Alastair Morrison for his general all-round entertainment value. Most importantly, I wish to thank my mentor Stefan Nonchev, who was forced to live through every step of my project, gel-by-gel, and without whom it would all have been impossible. Finally, I am indebted to my parents for their endless encouragement and support.

This research was carried out at the National Institute for Medical Research, and was financially supported by the Medical Research Council.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Introduction

Chapter 3: Regulation of the *Hoxb-9* gene

Chapter 4: Regulatory interactions between adjacent *Hox* genes

Chapter 5: Analysis of a new polydactylous mouse mutant

Chapter 6: Discussion

LIST OF TABLES

Chapter 2: Materials and methods

ABBREVIATIONS

AER	apical ectodermal ridge
A-P	anteroposterior
ANT-C	*Antennapedia* complex
BMP	bone morphogenetic protein
bp	base pair
BX-C	*Bithorax* complex
CNS	central nervous system
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
DEPC	diethyl pyrocarbonate
dGTP	deoxyguanosine 5'-triphosphate
DIG	digoxygenin
DMF	dimethylformamide
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dpc	days post coitum
DTT	dithiothreitol
dTTP	deoxythymidine 5'-triphosphate
EDTA	diaminoethanetetra-acetic acid
EGTA	ethylene glycol-bis-N,N,N',N'-tetraacetic acid
F	Farads
HOM-C	homeotic complex
hr.	hour
ISH	*in-situ* hybridisation
kb	kilobase
min.	minute
MOPS	3[-N-morpholino] propanesulphonic acid
NBT	4-nitro blue tetrazolium chloride
OD_{550}	optical density at a wavelength of 550nm

PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PI	positional information
PZ	progress zone
r	rhombomere
RT	room temperature
RA	retinoic acid
RAR	retinoic acid receptor
RARE	retinoic acid response element
RNA	ribonucleic acid
RXR	retinoid X receptor
SDS	sodium dodecyl sulphate
sec.	second
SLS	sodium lauryl sulphate
SPF	specific pathogen free
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	Tris[hydroxymethyl]methylamine
UV	ultraviolet
V	volts
W	Ohms
X-Gal	5-Bromo-4-chloro-3-indolyl-b-D-galactopyranoside
ZPA	zone of polarising activity

CHAPTER 1

INTRODUCTION

Due to the nature of the research I have performed for my PhD, this introduction describes two related topics. Sections 1 to 5 deal with *Hox* genes and the possible reasons for their special clustered organisation, and sections 6 to 9 deal with vertebrate limb development.

PART I

1.1 Background

All multicellular life on earth, depends on developmental programs which are encoded in networks of interacting genes. In trying to understand how these programs are constructed, and how millions of cells are coordinated during morphogenesis, a number of questions have arisen concerning the principles on which these systems are built: To what extent do different regions of an embryo act independently from the rest, following their own independent program? In other words, to what extent are patterning mechanisms local or global? How widespread an effect can a single gene have on development? How directly are genetic elements correlated to structural ones? One important contribution to answering these questions stems from the work of Bateson (Bateson, 1894). He coined the term *homeosis* to describe when "something has been changed into the likeness of something else". He was not studying mutant strains, only mutated individuals, however during the 1920's and early 30's, the first homeotic mutants were discovered in *Drosophila*, and these ultimately were the starting point for the current interest in *Hox* genes.

The mutants *bithorax* (*bx*), *spineless-aristapedia* (ss^a) and *proboscipedia* (*pb*) discovered by Bridges and Balkaschina (Lewis, 1994) all showed a transformation of one body part into another: *bx* has a second thoracic segment in place of a third, ss^a has a tarsus in place of an antenna, and *pb* has a tarsus in place of the proboscis. The interest in homeotic transformations lay in the demonstration that a single genetic element was controlling the coordinate patterning of a distinct region of the embryo. Unlike previously known mutations which resulted in a disruption or loss of patterning, these mutants simply copied a patterning process from one region to another, such that the misplaced structure was nevertheless perfectly constructed. For the first time, single genetic

elements had been found which related two things together in a very direct way: a correctly formed, complex, physical structure, and its position within the body of the organism. As such these discoveries were an important part of our gradual unravelling of the questions mentioned above.

1.2 The Bithorax and Antennapedia complexes

Subsequent to the discovery of the first three homeotic mutants, many more were generated which could cause transformations in almost any segment along the anteroposterior (A-P) axis. All these mutants were found to lie in just two genetic loci named the *Bithorax*-complex (BX-C) and the *Antennapedia*-complex (ANT-C) (mapped by (Lewis, 1978) and (Kaufman *et al.*, 1980), and collectively known as the homeotic complex or HOM-C (Akam, 1989)). Despite a high number of mutations found in each complex, there are only three homeotic genes in the BX-C and five in the ANT-C, and most of the mutations are now known to disrupt regulatory elements, not the protein-coding regions. *In-situ* analysis of the genes has shown that they are expressed shortly after segmentation has occurred, and that each one is expressed in a different but overlapping set of segments (Akam, 1987). In this way most segments express a unique combination of homeotic genes, and it is this combination which tells the cells in each segment which developmental program to run. For example, the segment which will grow antennae expresses *Sex combs reduced*, whereas the segments which will develop wings express *Antennapedia* and *Ultrabithorax* (Fig. 1.1). Thus the function of homeotic genes is described as giving segments their identities. The reason for homeotic transformations can easily be explained as a misregulation of homeotic genes: If a mutation causes the gene *Ultrabithorax* not to be expressed in the third thoracic segment, then its homeotic code will be the same as the second thoracic segment, and the fly will develop two pairs of wings instead of one (Ingham, 1985).

Although different segments do express different combinations of homeotic genes, the original idea that a genuine "combinatorial code" is the sole determining factor in segmental identity is now known to be an oversimplification. Certain of the homeotic genes are known to exert *phenotypic suppression* over others, such that their effect on segment identity is unchanged by the presence of the others (Gonzalez-Reyes and Morata, 1990; Duboule and Morata, 1994), and the spatial pattern of expression within a segment is also known to be important (Peifer *et al.*, 1987), especially in the abdominal segments which, despite expressing the same combination of homeotic genes, are not identical.

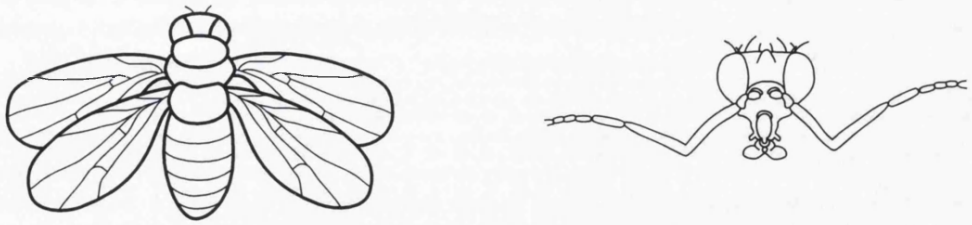
Fig. 1.1 The *Drosophila* homeotic genes.

(a) A combination of mutations which results in the absence of the gene *Ultrabithorax* in the third thoracic segment, causes a *homeotic* transformation which leads to a duplication of segment T2, and therefore the development of two pairs of wings instead of one. A similar type of mutation in the ANT-C (named *antennapedia*) can lead to the development of legs in the position where antennae should grow.

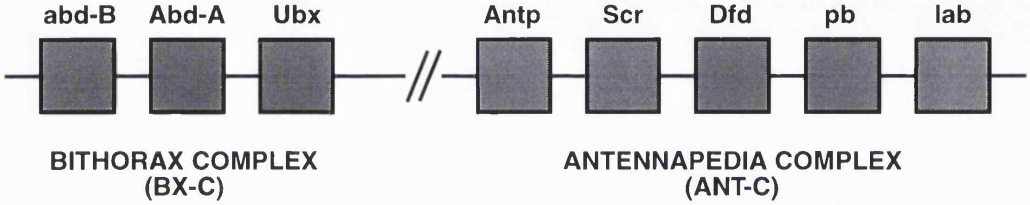
(b) The *Drosophila homeotic* genes are organised into two clusters which lie on the same chromosome.

(c) The expression domains of *homeotic* genes within the *Drosophila* embryo are arranged in such a way that most segments (or parasegments) express a different combination. In this way the genes are said to give the segments their identities. (The shading indicates how strongly the genes are expressed in each segment.)

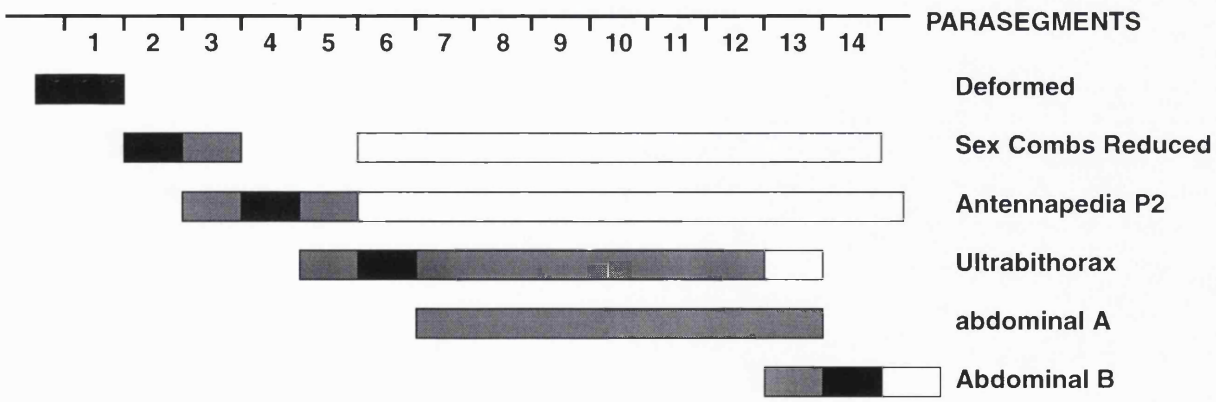
(a)



(b)


abd-B
Abd-A
Ubx
Antp
Scr
Dfd
pb
lab
BITHORAX COMPLEX
(BX-C)
ANTENNAPEDIA COMPLEX
(ANT-C)

(c)


PARASEGMENTS
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Deformed
Sex Combs Reduced
Antennapedia P2
Ultrabithorax
abdominal A
Abdominal B

As might be expected for genes with switch-like functions, they encode transcription factors, and can therefore activate or repress many other genes (Muller *et al.*, 1988; Scott *et al.*, 1989). They all have similar sequence motifs and are believed to have arisen by gene duplication. The region of greatest similarity is called the homeobox (McGinnis *et al.*, 1984) and encodes an 80 amino acid stretch of the protein called the homeodomain which is a helix-turn-helix motif used for binding to DNA (Otting *et al.*, 1988; Gehring *et al.*, 1990; Otting *et al.*, 1990).

1.3 The *Hox* genes

Using the homeobox as a probe to screen against genomic and cDNA libraries, it was found that all animal species tested so far contain homeobox genes (McGinnis *et al.*, 1984), and that *Drosophila* itself contains many more than those located in the BX-C and ANT-C (for example *engrailed* and *even-skipped*, (Fjose *et al.*, 1985; Kuner *et al.*, 1985; MacDonald *et al.*, 1986)). Most interestingly of all, it was found that in vertebrates the homeobox genes most closely related to the homeotic genes were also clustered together into large complexes. These were named the *Hox* genes. The relationship between the four vertebrate *Hox* complexes and the two *Drosophila* ones is shown in Fig. 1.2a, and strongly suggests that the common ancestor of vertebrates and arthropods possessed a single complex, which became duplicated in vertebrates to give four copies, and which split in half in *Drosophila* (Akam, 1989). Other insect species contain only one complex (Beeman *et al.*, 1989), so the split in *Drosophila* (which has left the two halves on the same chromosome) is believed to be relatively recent, and the vertebrate complexes are considered to be a closer paradigm of the ancestral complex.

The four vertebrate clusters are named *Hoxa* to *Hoxd* (Scott, 1992), and in mice and humans each occurs on a different chromosome. The ancestral cluster may have contained as many as 13 genes, but within each present-day cluster a number of these have been lost. Nevertheless *paralogous* relationships exist between genes which derived from a common ancestor gene, for example *Hoxa-4*, *b-4*, *c-4* and *d-4* are all more similar to each other than to any of the other genes. Paralogous groups are either named after their *Drosophila* homologue, or their numerical order, so that *a-4* etc. all belong to the *Deformed* group, or group 4.

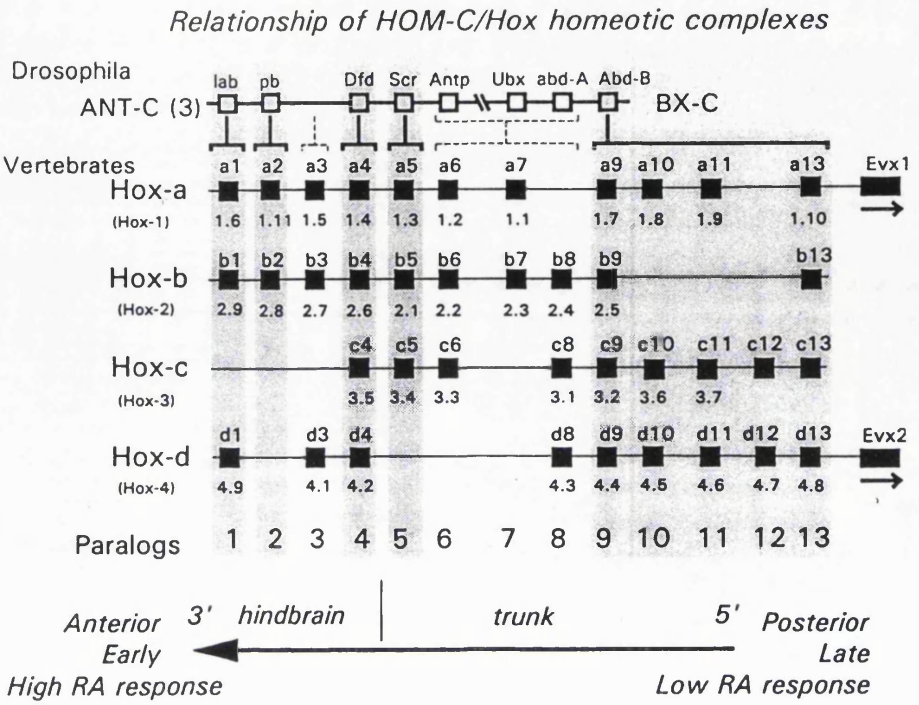
The expression patterns of the vertebrate *Hox* genes appears to be similar to those of the *Drosophila* homeotic genes: in overlapping domains along the A-P axis. *In-situ* hybridisation analysis (Duboule and Dolle, 1989; Graham *et al.*, 1989; Wilkinson *et al.*, 1989b) showed that they

Fig. 1.2 The Vertebrate *Hox* genes.

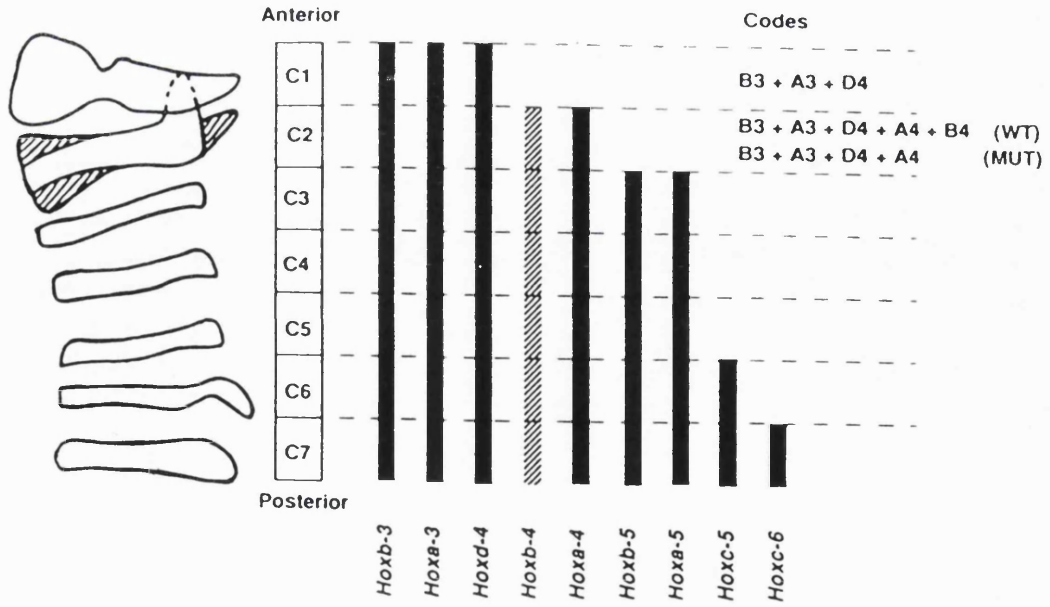
(a) An alignment of the vertebrate *Hox* clusters with each other and with their *Drosophila* homologues, reveals the paralogous groups. (The numbers under the genes are the old nomenclature.) This alignment is based firstly on sequence similarity, and only when such similarities are ambiguous, on chromosomal position. The shaded blocks contain genes with recognisable homologous relationships.

(b) In a similar manner to *Drosophila*, the vertebrate *Hox* genes appear to provide a combinatorial code for specification of certain segmental structures. In this Figure, the difference between the presence and absence of *Hoxb-4* is shown as an alteration of the normal shape of the second cervical vertebra, which correlates with the anterior-most position of the wildtype *Hoxb-4* expression pattern, (from Ramirez-Solis *et al.*, 1993).

(a)


Relationship of HOM-C/Hox homeotic complexes
Drosophila
ANT-C (3)
lab pb Dfd Scr Antp Ubx abd-A Abd-B
BX-C
Vertebrates
Hox-a
(Hox-1)
a1 a2 a3 a4 a5 a6 a7 a9 a10 a11 a13 Evx1
1.6 1.11 1.5 1.4 1.3 1.2 1.1 1.7 1.8 1.9 1.10
Hox-b
(Hox-2)
b1 b2 b3 b4 b5 b6 b7 b8 b9 b13
2.9 2.8 2.7 2.6 2.1 2.2 2.3 2.4 2.5
Hox-c
(Hox-3)
c4 c5 c6 c8 c9 c10 c11 c12 c13
3.5 3.4 3.3 3.1 3.2 3.6 3.7
Hox-d
(Hox-4)
d1 d3 d4 d8 d9 d10 d11 d12 d13 Evx2
4.9 4.1 4.2 4.3 4.4 4.5 4.6 4.7 4.8
Paralogs 1 2 3 4 5 6 7 8 9 10 11 12 13
Anterior 3' hindbrain
trunk
5' Posterior
Early
Late
High RA response
Low RA response

(b)


Anterior
C1
C2
C3
C4
C5
C6
C7
Posterior
Codes
B3 + A3 + D4
B3 + A3 + D4 + A4 + B4 (WT)
B3 + A3 + D4 + A4 (MUT)
Hoxb-3
Hoxa-3
Hoxd-4
Hoxb-4
Hoxa-4
Hoxb-5
Hoxa-5
Hoxc-5
Hoxc-6

are expressed in many tissues during embryogenesis and as a rule display sharp boundaries of expression at the anterior limit of these domains; also paralogous genes show similar boundaries along the A-P axis (Gaunt *et al.*, 1989; Hunt *et al.*, 1991). Despite the lack of large-scale segmentation during vertebrate embryogenesis, a number of tissues do go through a segmented stage: somites, vertebrae, ribs, muscles and nerves. The embryonic hindbrain also passes through a phase in which it consists of 8 repeated morphological bulges called rhombomeres (Lumsden and Keynes, 1989). During this phase the anterior boundaries of expression of *Hox* genes in the neural tube coincide with specific rhombomeric boundaries in a very similar way to homeotic gene expression boundaries in the segments of *Drosophila* (Wilkinson *et al.*, 1989b; Hunt *et al.*, 1991; reviewed in Keynes and Krumlauf, 1994). It is therefore extremely likely that *Hox* genes are an ancient mechanism for specifying positional information along the A-P axis, which is conserved across most species of animal (reviewed in McGinnis and Krumlauf, 1992).

Due to the more complex nature of mammalian development compared with *Drosophila,* natural mutant strains which correspond to mutations in *Hox* genes have rarely been found, so to test their function, targeted disruption has been a major approach. In this technique (developed by Capecchi (1989)) homologous recombination is used to disrupt a gene in pluripotent embryonic stem cells (ES cells) which are then inserted into blastocysts. Chimeric mice develop, some of which carry the mutation in their germ line cells and these mice are founders for a null mutant strain. All *Hox* genes disrupted so far, produce developmental alterations in the anterior region of their expression domains, and this supports the hypothesis that vertebrate *Hox* genes have similar functions to those in *Drosophila.* The first two disruptions performed (Chisaka and Capecchi, 1991; Lufkin *et al.*, 1991), *Hoxa-1* and *Hoxa-3*, showed regionally-restricted loss of neural crest-derived structures. A number of *Hox* null mutants have displayed anterior homeotic transformations in which the morphology of certain vertebrae has partially adopted the appearance of the adjacent anterior vertebra: (Le Mouellic *et al.*, 1992; Condie and Capecchi, 1993; Dolle *et al.*, 1993; Gendron-Maguire *et al.*, 1993; Ramirez-Solis *et al.*, 1993) for *Hoxc-8*, *d-3*, *d-13*, *a-2* and *b-4* respectively (Fig. 1.2b). This result is similar to the effect of null mutations in *Drosophila*, however a couple of cases have displayed posterior transformations (*Hoxa-5* and *a-11*) (Jeannottee *et al.*, 1993; Small and Potter, 1993). The theory that the four complexes arose by duplication from an ancestral cluster, coupled with the similarity of expression domains within each paralogous group, suggests that some degree of functional overlap or redundancy may be occurring between the genes, and this may explain the subtlety and complexity of some of the phenotypes observed. In agreement

with this proposal is the observation that disrupting *Hoxa-11* or *Hoxd-11* independently causes relatively minor phenotypes, whereas the double-mutant displays an almost complete loss of the ulna and radius (Davis *et al.*, 1995).

The belief that the *Hox* genes perform essentially the same role across the animal kingdom (providing positional information to cells along the A-P axis) has led to the proposal that possessing a *Hox* cluster is the key requirement for any organism to be classified as an animal (Slack *et al.*, 1993). The embryonic stage when *Hox* genes are expressed (shortly after neuralation) has been named the *phylotypic stage* and the pattern of *Hox* genes the *zootype*.

1.4 Organisation of *Hox* genes and colinearity

Another feature of this gene family which has been conserved between vertebrates and insects, is the phenomenon termed *spatial colinearity* (Graham *et al.*, 1989). The expression patterns of *Hox* genes in mice, *Xenopus* and *Drosophila* display a correlation between the position of a gene within the cluster and the position of the expression domain along the A-P axis of the embryo. All *Hox* genes are transcribed in the same direction with respect to each other (except the *Deformed* gene in *Drosophila*), so the clusters can be described as having a 5' and a 3' end. Genes at the 3' end are always expressed more anteriorly than those at the 5' end (Fig. 1.3). In mice the expression domains for most of the genes extend posteriorly into the tip of the tail, and it is only the anterior boundary that shows colinearity. This feature is particularly well demonstrated in the hindbrain: *Hoxb-2* is expressed up to the third rhombomere (r3), but not in the second, with a sharp boundary at the r2/3 junction, and the subsequent genes *b-3*, *b-4* and *b-5* are expressed up to r5, r7 and the hindbrain/spinal cord junction. In this case the shift in boundary position is always two rhombomeres (Fig. 1.3b). The exception to this rule in the *Hoxb* cluster, is *Hoxb-1* which is only expressed strongly in r4 (Murphy and Hill, 1989, Wilkinson *et al.*, 1989b).

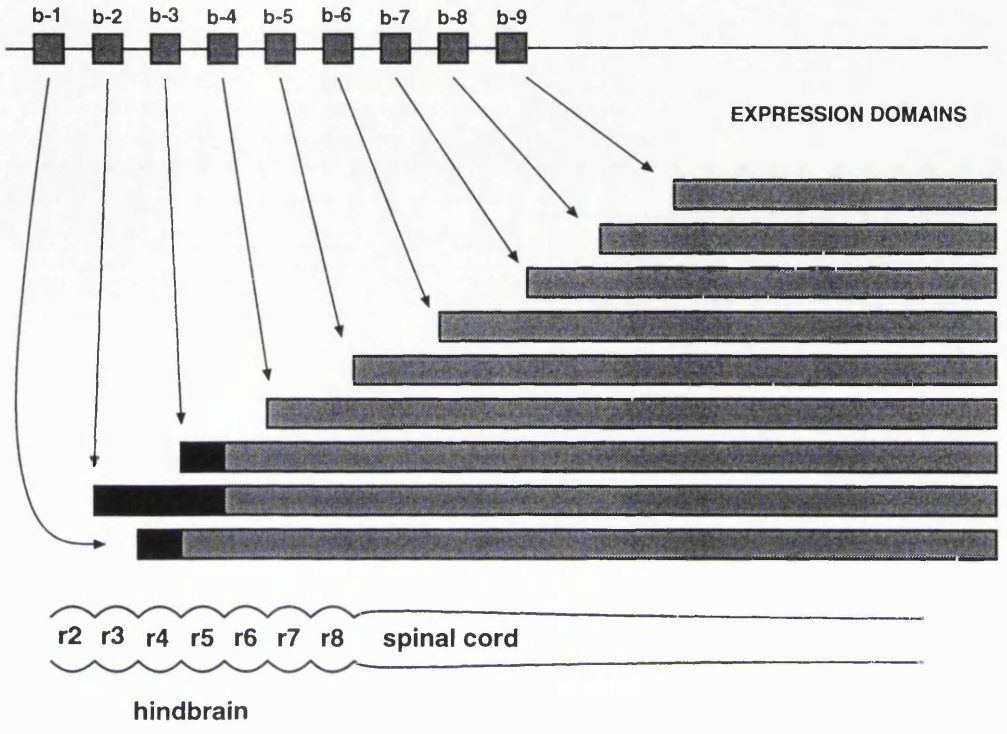
In *Drosophila* it is not only the expression patterns of the homeotic genes which display spatial colinearity. In fact the earliest realisation of colinearity was by Lewis (1978) before molecular biology had allowed visualisation of gene expression patterns. He mapped mutations onto the BX-C, and discovered the correlation between their positions in the complex and the segments they affected. Although there are only three protein-coding genes in this complex, there are 9 different mutation zones, each of which correspond to a regulatory region, and all of which fit the colinearity rule (Fig. 1.3a). More recently it has been found that transcripts are generated from these regulatory

Fig. 1.3 Colinearity in *Drosophila* and vertebrates.

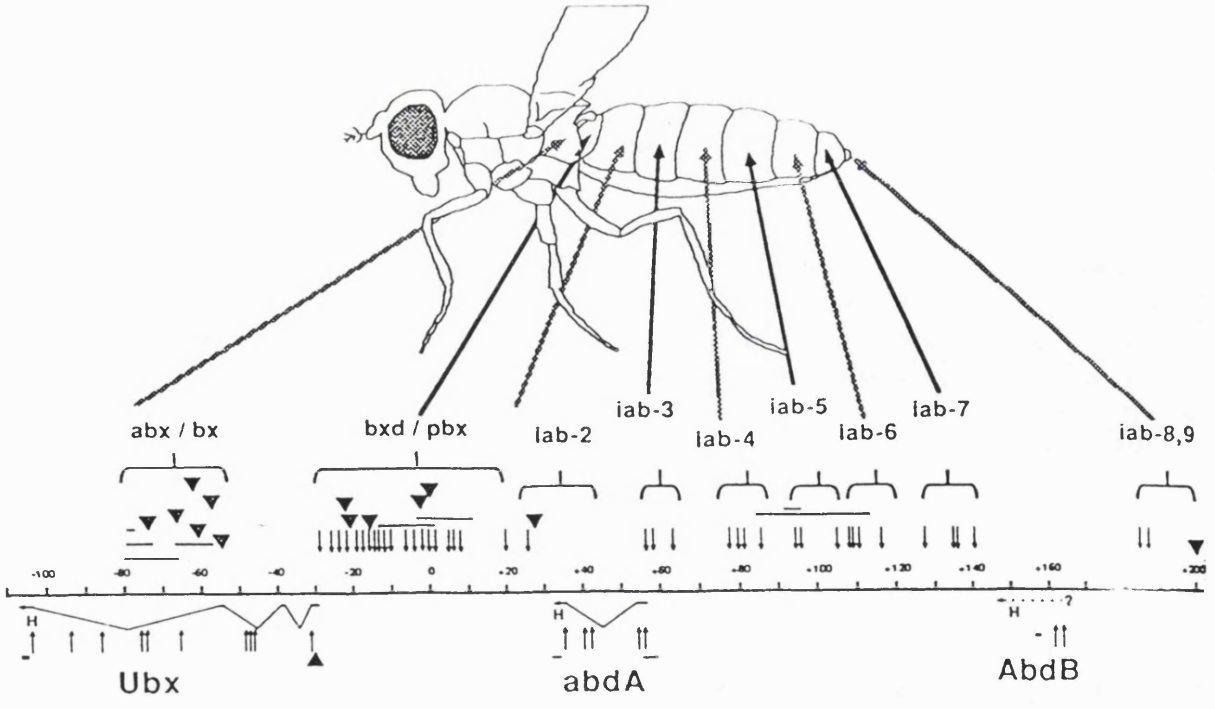
(a) The *Hoxb* genes show colinearity in their expression patterns along the mouse neural tube. More 3' genes are expressed up to a more anterior position, with the distance between adjacent boundaries being two rhombomeres in the hindbrain, but less distinct in the spinal cord.

(b) Although there are only three *homeotic* genes in the *Drosophila* Bithorax-Complex, there are nine regulatory regions, and they are active in a colinear pattern, each one displaying its primary importance in a different segment, (Lipshitz *et al.*, 1987; Sanchez-Herrero and Akam, 1989).

(a)


b-1
b-2
b-3
b-4
b-5
b-6
b-7
b-8
b-9
EXPRESSION DOMAINS
r2 r3 r4 r5 r6 r7 r8
spinal cord
hindbrain

(b)


abx / bx
bxd / pbx
lab-2
iab-3
lab-4
iab-5
lab-6
lab-7
iab-8,9
Ubx
abdA
AbdB

regions, and that the spatial distribution of these is also colinear. Lipshitz *et al.*, (1987) studied transcripts from *bxd*, a regulatory region for the *Ubx* gene, and Sanchez-Herrero and Akam (1989) studied the *iab* (*infra-abdominal*) region which regulates the *abd-A* and *Abd-B* genes). The function of these transcripts is not known, and despite their clear colinearity, there appears to be no strong relationship between function and expression pattern: The segment most strongly affected in a particular mutant, contains in the wildtype only very low levels (if any) of the transcripts from that regulatory region. High levels are found in all segments posterior to the strongly affected one.

In addition to spatial colinearity, vertebrate *Hox* genes display *temporal colinearity* (Dolle *et al.*, 1989). This means that in individual cells the most 3' gene is activated first, and the rest are sequentially activated in the order they occur along the chromosome. In the *Hoxb* cluster this is particularly well demonstrated, with *Hoxb-1* switching on first, then *Hoxb-2*, then *Hoxb-3* etc. A third type of colinearity has also been described, which is a graded sensitivity to activation by retinoic acid (Simeone *et al.*, 1990; Boncinelli *et al.*, 1991). In cell culture experiments undifferentiated human ES cells can be stimulated to activate *Hox* genes, by exposing them to RA. At high concentrations all genes are activated in the normal temporal order, but at lower concentrations only the more 3' genes become active. However, it seems likely that this phenomenon and temporal colinearity are different manifestations of the same mechanism. If *Hox* activation is seen as a progressive process in which *Hoxb-2* cannot be active until *Hoxb-1* is on, and *Hoxb-3* cannot switch on until *Hoxb-2* is on etc. then RA may simply be a general activating signal, which at low levels fails to push the activation process all the way through the complex.

In fact, temporal and spatial colinearity may also be different ways of describing the same thing. The patterns of *Hox* gene activation in mouse embryos show that the earliest gene (*Hoxb-1*) is first expressed in a posterior domain which then spreads anteriorly as development proceeds (Murphy *et al.*, 1989; Frohman *et al.*, 1990; Murphy and Hill, 1991). The next gene (*Hoxb-2*) is activated in the posterior domain shortly after *Hoxb-1*, and similarly spreads forward during embryogenesis. Simply because it was activated after *Hoxb-1* its anterior boundary of expression is posterior to that of *Hoxb-1*. At a certain stage, the forward spreading of domains ceases, and each successive gene has a more posterior boundary because it was activated later than its 3' neighbour (except for *Hoxb-2* which ends up more anteriorly than *Hoxb-1*). Monitoring the expression within any individual cell, would demonstrate a successive activation of *Hox* genes, which could theoretically be a cell-autonomous process (although initiation of the process must be coordinated along the axis). The spatial colinearity may therefore be a consequence of the temporal colinearity.

However, this is not the case in *Drosophila* where all the homeotic genes are activated at the same time.

1.5 Regulation of *Hox* genes

Before the discovery of colinearity and the analysis of the β-globin complexes, it was generally believed that the position of a gene on a chromosome, relative to other genes, was not critical to its function, and indeed for the majority of genes this view is still held. As result, chromosomal rearrangements during evolution cause a "mixing" of the genome so that genes which at one time were next to each other become separated. This is considered possible because the regulatory regions of most genes are small, modular and near the promoter and protein-coding region so the unit can operate independently of its neighbours. The maintenance of *Hox* genes in their complexes over such a long evolutionary time period (probably over 800 million years) suggests that some aspect of their function would be impaired if the genes were split up. The feature of colinearity, in which there is a remarkable coordination of expression between the genes, suggests that this critical aspect is gene regulation, and consequently one of the main goals for the study of *Hox* gene regulation is to find a mechanism which explains why colinearity exists.

The initial proposals for both *Drosophila* and mouse, were that prior to *Hox* gene activation, the complex is "closed" in a state of tightly packed heterochromatin, and that it is progressively "opened" starting from the 3' end (Peifer *et al.*, 1987; Dolle *et al.*, 1989). This *accessibility model* attempted to link colinearity with the initiation of *Hox* genes. However, experiments to uncover local *cis*-acting regulatory elements have been very successful and appear to dispense with the need for clustering during initiation (as described in the next three sections). The accessibility model was therefore modified to explain the subsequent *maintenance* of expression domains rather than their initiation (Gaunt and Singh, 1990; Paro, 1990), as described in sections **1.5.4** and **1.5.5**. Recently however, the viewpoint appears to be changing again. It is now recognised that the function of chromatin-based mechanisms is probably different between *Drosophila* and mice, and that in mammals these systems may well be involved in *Hox* gene initiation, as originally proposed (van der Hoeven *et al.*, 1996; van der Lugt *et al.*, 1996).

1.5.1 Initiation of homeotic gene expression in *Drosophila*

There are two major differences between the *Drosophila* homeotic genes and their vertebrate homologues: The complex has split into two halves, and there is no observable temporal colinearity. Since it is known that the split is relatively recent, and that the complex appears to be rigorously maintained in all other species tested, it is expected that extreme selective pressures on *Drosophila* (to speed up embryogenesis), are responsible for the primitive, more colinear state being abandoned. Despite this, the BX-C does show spatial colinearity, and more is known about the factors which regulate homeotic genes than any vertebrate *Hox* genes.

Formation of segments in *Drosophila* can be split into three stages, each with its own set of genes (Nusslein-Volhard and Wieschaus, 1980; Ingham, 1988). First, the *gap* genes define five broad, adjacent domains along the A-P axis. Second, the *pair-rule* genes use positional information from the gap genes to establish the 14 segment positions. Third, the *segment polarity* genes differentiate between the posterior and anterior sides of each segment, thereby specifying the morphological segment boundaries. The homeotic genes, which must be expressed in segment-specific patterns, use positional information from both the gap genes and the pair-rule genes to achieve this. Cis-acting regulatory elements (CREs) have been found to contain binding sites for these upstream factors which exert both positive and negative effects over gene regulation (Qian *et al.*, 1989; Riley *et al.*, 1991; Muller and Bienz, 1992; Shimell *et al.*, 1994). If this was the only information necessary for correct regulation, it would suggest that the genes no longer need to be next to each other, and indeed experiments which move parts of the cluster to new chromosomal locations indicate that this is the case (Struhl, 1984). However, as mentioned above, *Drosophila* is unusual with respect to colinearity and *Hox* gene clustering, so perhaps it has evolved new ways to regulate these genes, by linking them into a set of pre-existing genes (the segmentation genes) which can provide an alternative source of positional information.

1.5.2 *Cis*-acting regulatory elements of vertebrate *Hox* genes

The first step in attempting to understand *Hox* gene regulation has usually been analysis of *cis*-acting elements. Transgenic mouse embryos are created using DNA constructs in which local enhancer regions are tested for their ability to activate a promoter fused to the bacterial *LacZ* reporter

gene. A CRE's pattern of activity can then be visualised by treating the fixed embryo with a stain which detects the presence of the β-galactosidase enzyme (the gene product of LacZ). With a number of *Hox* genes (eg. *Hoxb-4* (Whiting *et al.*, 1991), *Hoxa-7* (Puschel *et al.*, 1991), *Hoxb-1* (Marshall *et al.*, 1994)) it has been possible to find local elements which when used together recreate the complete normal gene expression pattern. These results were initially unexpected, as they seem to dispense with the need for gene clustering (transgenic constructs insert randomly throughout the genome). However, a major drawback of these experiments is that the endogenous gene is still intact within the *Hox* complex, and many *Hox* genes are thought to autoregulate themselves (Popperl and Featherstone, 1992; Popperl *et al.*, 1995). This means that the correct regulation of the transgenic construct may reflect cross-regulation from the endogenous gene. The experiment which might most clearly resolve this would be a "knock-out" (targeted disruption as explained in section **1.3**) followed by a rescue (attempting to revert the phenotype back to a wildtype morphology by inserting a functional copy of the gene by transgenesis). But even this experiment might be complicated by the possibility that paralogous genes can cross-regulate each other (eg. in a knock-out of *Hoxb-4*, the *Hoxd-4* gene which has a similar anterior boundary of expression, might be able to cross-regulate a transgenic *Hoxb-4* rescue construct).

Despite the success in reconstructing many of the *Hox* gene patterns, it is actually hard to determine whether these patterns are correct at a biologically relevant level of detail. It is therefore still possible that the clustering is required for a more refined regulation. Additionally, for some genes it has never been possible to find local CREs which direct the normal pattern (eg. *Hoxb-7*, (Vogels *et al.*, 1993)).

Although these enhancer analysis experiments have not been successful in explaining colinearity, they have highlighted another intriguing question. It has been shown by many experiments from the Krumlauf lab that the enhancer-promoter interactions which appear to drive *Hox* gene expression, are very non-specific. This means that, for example, the neural CRE for *Hoxb-4* is able to activate the promoters of many different genes: *Hoxb-5*, *Hoxb-4*, *Hoxb-1*, *b-globin* and *hsp68*, and conversely, the promoter of *Hoxb-4* can be activated by CREs from *Hoxb-5*, *Hoxb-4* and *Hoxb-1*. The question then arises: In a condensed cluster of 9 genes in a stretch of 80kb, how do the right CRE's control the right promoters? It is known that this situation of low specificity is distinctly different to that found in some other cases. For example the promoters of the two *Drosophila* genes *gooseberry* and *gooseberry neuro* (*gsb* and *gsbn*) are separated by a 10kb region which contains their respective enhancers (Li and Noll, 1994). Despite the proximity of

elements in this small region, each of the two enhancers only activates one promoter. It has been shown that this strong specificity is *not* simply the result of biased competition, but is instead due to an incompatibility between elements which should not interact: even when the *gsb* enhancer is placed directly next to the *gsbn* promoter in the absence of other elements, transcriptional activation does not occur (and the same is true for the *gsbn* enhancer on the *gsb* promoter). The lack of this type of enhancer-promoter exclusivity within the *Hox* clusters suggests that extra mechanisms must exist.

1.5.3 Upstream factors involved in *Hox* gene initiation

Although the *Hox* genes themselves were found due to their similarity to the *Drosophila* homologues, the same approach has not been successful in identifying vertebrate upstream regulators. As mentioned above, it is likely that vertebrate *Hox* clusters more closely represent the primitive state, and retain more of the original regulatory mechanisms (which may include the reason for colinearity), so that although vertebrate homologues for a number of the pair-rule genes which are involved in patterning processes have been found (*En-1* and *En-2* are homologues of *engrailed* (Davidson *et al.*, 1988), and the *Pax* genes are homologues of *paired* (Gruss and Walther, 1992)) they do not appear to give a similar input to *Hox* regulation. Only one such gene is thought to be a *Hox* gene regulator, and although its *Drosophila* homologue is the gap gene *Krüppel*, its role in mouse is probably more similar to a pair-rule gene.

The *Krox20* gene encodes a zinc-finger transcription factor, which binds to specific DNA sequences *in vitro* (Chavrier *et al.*, 1988; Chavrier *et al.*, 1990). It is expressed in the developing hindbrain, before rhombomere formation, in two stripes which later correspond to r3 and r5 (Wilkinson *et al.*, 1989a). *Hoxb-2* is initially expressed uniformly throughout the neural tube and hindbrain up to the r2/3 boundary, but is then upregulated to higher levels in rhombomeres 3, 4 and 5 (Sham *et al.*, 1993). This upregulation is mediated by two regulatory elements near the *Hoxb-2* gene, which independently control upregulation in r4 and r3/5. The r3/5 enhancer contains three binding sites for the *Krox20* gene, and when these are used to drive *LacZ* in transgenic experiments, expression is found specifically in r3 and r5. Moreover, if a second transgenic construct is introduced which expresses the *Krox20* protein ectopically throughout the neural tube, then the *LacZ* reporter is also expressed in the extended pattern, providing strong evidence for a direct interaction (Sham *et al.*, 1993).

There is evidence for direct *Hox* regulation by another type of DNA-binding protein: the nuclear retinoid receptor family. Since the experiments of Nieuwkoop in 1952 (Nieuwkoop, 1952) it has been proposed that an anterior fate is the default state for axial tissue during A-P patterning, and that the signal responsible for creating the gradient of positional information is a *posteriorising* one. There is much evidence that retinoic acid (RA) could be such a molecule. When applied to developing *Xenopus* embryos it causes A-P transformations of the central nervous system, in which anterior structures often fail to develop due to excessive posteriorisation (Durston *et al.*, 1989; Sive *et al.*, 1990; Papalopulu *et al.*, 1991a; Ruiz i Altaba and Jessell, 1991). In mouse, RA can transform rhombomeres 2 and 3 into an r4/5 identity (Marshall *et al.*, 1992), and can cause posterior transformations of vertebrae along the entire body axis (Kessel and Gruss, 1991). RA seems to play a similar role in A-P patterning of the limb bud, which is described in section **1.7.2.** Also, endogenous RA has been found in slight spatial gradients in the chick limb bud and whole *Xenopus* embryos (Thaller and Eichele, 1987; Durston *et al.*, 1989).

Since the *Hox* genes are implicated in providing cells with positional information, it is possible that they are direct targets for an A-P patterning mechanism. Ectopic RA has been shown to induce *Hox* gene expression in cell culture experiments in a colinear fashion (Simeone *et al.*, 1990; Papalopulu *et al.*, 1991b), and to alter *Hox* gene expression patterns in various embryonic tissues (Kessel and Gruss, 1991; Morriss-Kay *et al.*, 1991; Papalopulu *et al.*, 1991b; Conlon and Rossant, 1992; Marshall *et al.*, 1992). In the majority of cases, the result is an anterior shift in expression domains, as would be expected for a posteriorising signal.

RA acts as the ligand for a family of proteins called nuclear retinoid receptors, which in turn bind to target DNA sequences named RAREs (retinoic acid response elements) to produce transcriptional activation of target genes. There are six members: three retinoic acid receptors (RARα, RARβ and RARγ), and three retinoid X receptors (RXRα, RXRβ and RXRγ) (Chambon *et al.*, 1991; Mangelsdorf *et al.*, 1992). In support of the connection between RA and *Hox* regulation, many RAREs have been found in *Hox* regulatory elements: in *Hoxa-1* (Langston and Gudas, 1992), *Hoxd-4* (Popperl and Featherstone, 1993), and *Hoxb-1* (Marshall *et al.*, 1994; Studer *et al.*, 1994), and correct functioning of the regulatory sequences in transgenic experiments has been shown to depend on the RARE. Point mutations are enough to abolish transcriptional activation (or repression in the case of the *Hoxb-1* 5' RARE). Moreover, these small RARE transgenic constructs respond to ectopic RA in the same way as the endogenous genes - an anterior shift in expression domain.

There is still uncertainty about the specificity of interactions between RAREs and their binding proteins, but the important point with respect to *Hox* gene organisation is that they do appear to be important, and yet do not explain colinearity. If RAREs were only found at one end of the complex, it could suggest that proximity of a *Hox* gene to the RA-responsive region would control its regulation. However, as RAREs are found towards the middle of complex as well, it makes this theory much less likely. If RAREs are found associated with many more genes, they will not seem to provide the long-range links which could prevent the complex from drifting apart during evolution.

1.5.4 Maintenance of homeotic expression patterns

In addition to the homeotic genes themselves, another class of mutants were found to produce homeotic transformations. The first member identified, *Polycomb* (*Pc*), has the most striking effect: in heterozygotes the second and third pairs of legs are transformed into the morphology of the first pair, and in homozygotes all segments take on the identity of the eighth abdominal segment. Lewis (1978) interpreted the cause for these changes as being expression of BX-C genes throughout the entire embryo, and therefore suggested that *Pc* was a homeotic repressor. Since then, many more members of the *Polycomb*-Group (*Pc*-G) have been found, all of which are related by function only. *Extra sex combs* (*esc*) was found to have very similar phenotypes to *Pc* (Struhl, 1981), and similar effects on BX-C gene expression (Struhl and Akam, 1985), and some of the weaker members were shown to create strong phenotypes when combined in double mutants (Jurgens, 1985).

Although expression of the homeotic genes in these mutants is derepressed at germ band stage, at earlier stages it appears completely normal (Struhl and Akam, 1985; Kuziora and McGinnis, 1988), and this is because initiation of homeotic genes is controlled by the segmentation genes. The function of the *Pc*-G is to maintain their spatially restricted domains of expression (Paro, 1990). This is necessary for two reasons: homeotic gene expression outlasts that of the segmentation genes which provided their original positional cues, and it must be maintained through a number of cell divisions. To perform this function it is believed that the proteins encoded by the Pc-G are involved in controlling chromatin structure. The protein sequences of *Polycomb* and *Posterior Sex Combs* are similar to proteins encoded by the genes *Su(var)205* and *Su(z)2*, which are suppressers of position-effect variegation and transvection respectively (Wu *et al.*, 1989;

Henikoff, 1990; Brunk *et al.*, 1991; Paro and Hogness, 1991;). Both processes are mediated through modification of chromatin structure, and the protein encoded by *Su(var)205* is HP1, a nonhistone heterochromatin-associated protein.

The idea of homeotic repression through modification of chromatin structure fits the observations perfectly. The mechanism must be a "passive" one, not dictating expression patterns to the homeotics (as this is performed by the segmentation genes), but reading their states after the domains are established. In agreement with this, the *Pc* gene is not expressed in a spatially restricted domain, it occurs uniformly throughout the embryo (Paro and Hogness, 1991). A chromatin-based mechanism could use the presence of transcriptional repressors as nucleation sites for progressive heterochromatinisation involving proteins from the Pc-G (Zhang and Bienz, 1992). This "closed" chromatin state could then be maintained through cell divisions in a similar manner to CpG methylation. A reduction or loss of members of the Pc-G would weaken or remove this inherited repression (Paro, 1990).

To further test this hypothesis, Orlando and Paro (1993) analysed the binding distribution of Pc protein to the BX-C, from *Drosophila* SL-2 cells. In these cells the genes *Ubx* and *abd-A* are repressed whereas *Abd-B* is active. The distribution of Pc protein was found to cover the entire complex except for the coding region of *Abd-B*, demonstrating a complete correlation between repressed regions and Pc binding. However, it has since been demonstrated that there is a discrete element responsible for *Pc*-repression in the *Ubx* gene called the PRE (*Polycomb* response element) (Chan *et al.*, 1994). The even distribution of Pc protein seen over the *Ubx* gene may therefore be the result of heterochromatin spreading along the DNA from the PRE acting as a nucleation site.

Immunohistochemical analysis of Pc protein distribution in polytene chromosomes, and transgenic experiments using the *Antp* promoter to drive LacZ, both show that this chromatin-based mechanism acts on the ANT-C as well as the BX-C, and at approximately 60 other loci in the genome (Zink and Paro, 1989; Zink *et al.*, 1991). Also there is a second family of less-well characterised genes called the *trithorax*-Group, which appear to have the reverse function to the *Pc*-G, ie. they are proteins which maintain an "open" chromatin state, to keep target genes active (Reuter *et al.*, 1990; Kennison, 1993). Whereas the derepression caused by *Pc*-G mutants causes posteriorisation of segments, *trx*-G mutants cause anteriorisation, due to unchecked repression of the *homeotics*. Combining *trx* mutants with *Pc* mutants causes a cancellation of effects, such that double mutant embryos survive much further through development than either of the single mutants

(Ingham, 1983). In agreement with *trx*'s putative role in chromatin control is its structural similarity with *Drosophila Suvar(3)7*, another heterochromatin-associated protein involved in position-effect variegation.

It has been proposed that this new form of the accessibility model (Peifer *et al.*, 1987; Dolle *et al.*, 1989), in which chromatin structures affect maintenance rather than initiation (Gaunt and Singh, 1990) may be the reason for colinearity. Due to the order of the elements, the complex would always be split into only one 3' repressed region and one 5' active region. These regions would be larger than if the elements were in any other order, and may consequently be more stable. There is no real evidence to support this argument, and the observation that *Abd-B* can be active while Pc-protein complexes are repressing *Ubx* and *abd-A* (Orlando and Paro, 1993), goes against the colinearity rule, so despite good evidence in *Drosophila* that this regulatory mechanism is important, it does not seem to explain colinearity.

1.5.5 Maintenance of vertebrate *Hox* gene expression

Although the mechanism of *Hox* gene initiation is not highly conserved between *Drosophila* and mouse, the mechanism of maintenance may be. The murine proto-oncogene *bmi-1* displays extensive sequence similarity with the *Drosophila* Pc-G gene *Posterior sex combs* (*Psc*) (Brunk *et al.*, 1991; van Lohuizen *et al.*, 1991), and when mutated causes posterior transformation of vertebrae along the entire skeletal axis (van der Lugt *et al.*, 1994). When the protein is ectopically overexpressed in transgenic mice the opposite set of transformations occurs, in which vertebrae adopt more anterior morphologies, and the anterior boundary of the expression domain of *Hoxc-5* shifts posteriorly (Alkema *et al.*, 1995). All these effects are, in principle, identical to those of the Pc-G genes in *Drosophila*.

A murine homologue of the *trithorax* gene has also been found, *Mll*, which when mutated displays haploinsufficiency and in heterozygotes produces similar effects in mice to those of *trithorax* in *Drosophila*, ie: the opposite effect of *bmi-1* (Yu *et al.*, 1995). Anterior transformations are seen (as well as posterior ones), and expression of various *Hox* genes is either shifted posteriorly or absent at 10.5dpc.

Few other murine homologues of Pc-G or *trx*-G genes have been found or studied (the Mel-18 protein is 70% identical to the Bmi-1 protein (Goebl, 1992), and the murine protein M33 can functionally substitute for Pc protein in *Polycomb* null *Drosophila* (Pearce *et al.*, 1992; Muller

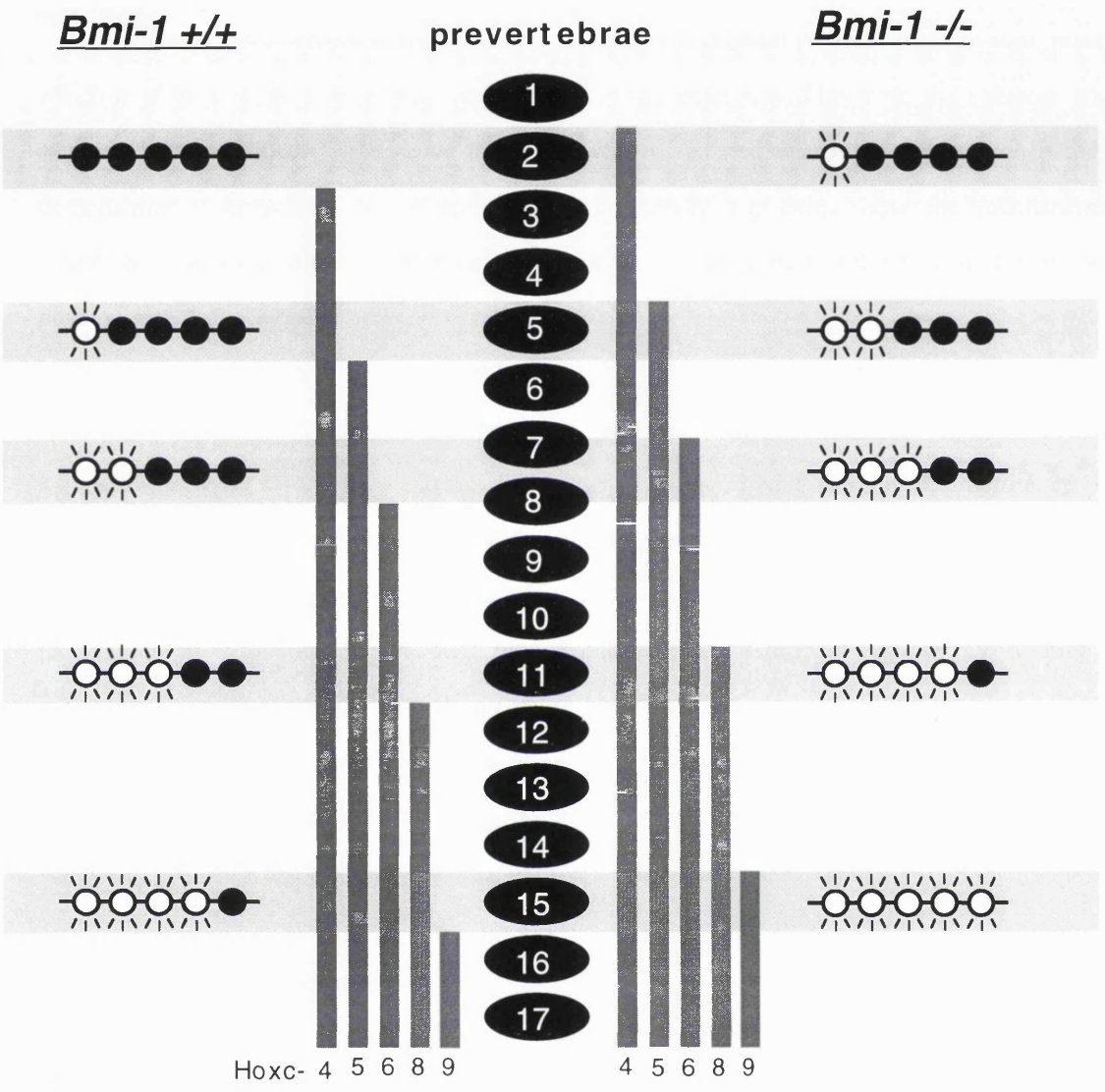
et al., 1995). However, the recent data from van der Lugt *et al.* on changes in *Hox* gene expression in the *bmi-1* mutant mice (van der Lugt *et al.*, 1996) provide the strongest evidence yet of a colinear regulatory mechanism. They report that the anterior boundary of many *Hox* genes in the mesoderm shifts anteriorly by one somite. If chromatin-associated repressors like *bmi-1* acted on each of the *Hox* genes independently, through local Polycomb-response elements, then the expected result for a loss of BMI-1 protein would be derepression of all *Hox* genes containing the relevant PREs. For example, *Hoxc-9* usually has a boundary between somites 15 and 16. The uniform loss of BMI-1 protein in the null mutant, would therefore be expected to cause a uniform derepression of *Hoxc-9*, such that somites 1 to 15 would all express *Hoxc-9* to some extent. The fact that this does not happen is very significant. Derepression of *Hoxc-9* is restricted to the one somite anterior to its usual boundary, and this is the same for all the affected genes. This aspect alone, immediately suggests that the effect of BMI-1 protein on repression of a *Hox* gene depends on more than just whether that gene is active or inactive.

If we consider individual cells along the A-P axis, it appears that the BMI-1-mediated repression of *Hox* genes takes positional information into account, and this can only be done by monitoring the complex as a whole, not solitary *Hox* genes: for example *Hoxc-6* can only tell a cell whether it is anterior or posterior to somite 8. The expression domains of all the following genes, *Hoxc-4, 5, 6, 8* and *9*, are affected in the mutant, however, due to their different anterior boundaries and the one-somite shift, only *one* gene is affected in any individual cell (Fig. 1.4). For example, in somite 15 of a wildtype embryo at 12.5dpc, *Hoxc-4*, *5*, *6* and *8* are on, and *c-9* is off. In the mutant only the expression of *c-9* is altered. Similarly, in somite 11, *c-4*, *5* and *6* are normally on, and *c-8* and *c-9* are off. This time only *c-8* is altered in the mutant. This occurs throughout the A-P axis, and consequently throughout the complex, such that in somite 2 all the genes mentioned are normally off, and the mutation only affects expression of *c-4*. *Hoxc-5*, *6*, *8* and *9* are unaffected. So if we reconsider the *Hox* complex as being spilt into just two domains: a 5' active region and a 3' repressed region, then the gene to be affected in a particular cell is always the one just 3' of the putative chromatin junction.

In other words, the effect of removing BMI-1 protein is to shift this junction in the 5' direction to reveal just one more gene for transcription. The observation that this effect is the same irrespective of the junction's position, indicates a complex-wide uniformity in the way BMI-1 influences *Hox* gene regulation, and strongly suggests that the position of the junction depends on a balance of opposing forces. It is expected that these opposing forces correspond to Pc-G and trx-G

Fig. 1.4 The effects of removing the murine Polycomb-related gene *Bmi-1* on expression of the *Hoxc* genes.

This figure shows the activity state of five genes from the *Hoxc*-complex for cells at different positions along the A-P axis (with reference to the prevertebrae). Inactive genes are represented as filled black circles, while active ones are open. The extent of the expression domain for each *Hox* gene is shown as a vertical grey bar. When the gene *Bmi-1* is mutated in mice, instead of a uniform derepression of the *Hox* genes, each affected gene is only derepressed in the somite anterior to its normal expression domain. This means that only certain somites display an altered *Hox* code (these are indicated by the horizontal shaded bands), while the rest remain unchanged (for example somites 1, 3, 4, 6, etc.) This strongly suggests that the BMI-1 protein is involved in a regulatory mechanism which takes into account the A-P position of each cell, and this information could be provided by the *Hox* genes themselves. The only uniform event along the whole A-P axis is a shifting of the boundary between active and inactive genes in a 3' direction, and this suggests that the regulatory mechanism involved does not treat each gene as an autonomous unit, but rather treats the cluster as a whole.



The *Hoxc* Complex

proteins respectively, which could bind cooperatively to DNA and to each other creating large multimeric complexes which spread along the chromosome to keep it "open" or "closed". The directionality of the forces could simply be the result of tethering opening factors to the 5' end and closing factors to the 3' end.

Here at last, is a putative mechanism controlling the position of the chromatin-state junction which would not work if the genes were ordered in any other way. However, the evidence which suggests this mechanism, does so by interfering with its internal workings, and as such is able to demonstrate a dependence on colinearity without actually explaining how the mechanism is controlled. It therefore probably represents only one half of the colinear link between chromosomal position and A-P position.

How the mechanism is controlled is very unclear. The homologous mechanism in *Drosophila* is clearly a maintenance system which takes its cue from the state of HOM-C expression dictated by the segmentation genes, but in mouse, *bmi-1* is expressed before the *Hox* genes, and more importantly the earliest analysed *Hox* gene in the *bmi-1* mutants is already affected while its domain is being established (at 9.5dpc). It therefore seems very possible that the chromatin mechanism in vertebrates is involved in initiation of expression. Its use in *Drosophila* for maintenance could reflect either a widespread function found also in vertebrates, or an adaptation to cope with the new rapid method of HOM-C initiation. A colinear mechanism for initiation of *Hox* expression would provide a more direct correlation between the chromosomal position and A-P patterning, and probably a stronger selection force for its maintenance through evolution.

1.5.6 Other clues to colinearity

There is one piece of evidence for another unusual aspect of *Hox* gene regulation which may relate to colinearity. In the human *Hoxc* complex a master promoter has been found which produces primary transcripts containing at least three adjacent *Hox* genes: *Hoxc-5, 6* and *8* (Boncinelli *et al.*, 1991). The transcript is alternatively spliced to produce mature mRNAs encoding any of the three genes. Each of the genes also has its own proximal promoter, and it is known that the proteins encoded by mRNAs from the master promoter are truncated versions of those encoded by the proximal promoters. It is believed that the truncated proteins contain the normal DNA-binding activity, but lack the motifs for transcriptional activation. Coupled with the fact that their expression domains are probably different but overlapping with that of the full protein (Cho *et al.*,

1988), it is possible that they function by competing with the full protein for DNA target sites, thereby acting as a repressor.

This unusual situation could relate to colinearity, because it is the only time when the linear order of the genes is preserved beyond the level of DNA, into the RNA. In theory, this means that proteins which control alternative splicing have access to the information of the relative order of the genes. How this information would be read, or how it would affect alternative splicing has not even been speculated, but it is a formal possibility.

PART II

1.6 The stages of limb morphogenesis

Limb development can be described as having four stages (Cohn and Tickle, 1996): initiation of the limb bud, specification of pattern, differentiation of tissues, and growth of the miniature limb to adult size. Due to the nature of the mutant studied in this PhD, my description of limb development will concentrate on the issues of pattern specification that are relevant to the digits.

The earliest structure of the vertebrate limb bud is a bulging of the ectodermal layer on the lateral side of the body wall, enclosing a pocket of mesenchymal cells. This initial bulging is formed not as a result of increased cell proliferation of the prospective limb bud mesenchyme, but rather a decrease in proliferation of tissue on either side of the buds (Searls and Janners, 1971). The protrusion then grows and extends into a slightly-flattened tube with proliferation rates now highest in the tip (Hornbruch and Wolpert, 1970). The region of mesenchyme which performs this growth is known as the *progress zone* (Summerbell *et al.*, 1973). As it extends away from the body, the cells left behind begin to differentiate such that fate-specification proceeds in a proximal to distal direction.

After a few days of development (depending on the species) the tip of the bud flattens out further to form a plate. Condensation of cartilaginous elements proceeds through the bud (from proximal to distal) and into the plate such that it produces the stylopod, the zeugopod and then through a series of complex bifurcations, the wrist or ankle bones and the digits. (Gruneberg, 1963; Hinchliffe, 1977; Hinchliffe and Johnson, 1980; Murray, 1989). The remaining stages of

development involve terminal differentiation of many more cells types, the growth of connective tissues, muscle and the circulatory system into the limb, and a large increase in overall size. However, these subsequent tissues use the early skeletal elements as a scaffold to direct their own development (Hinchliffe and Johnson, 1980), so by this stage the most critical events affecting overall limb morphology have already occurred.

1.7 Signalling regions of the limb bud

1.7.1 The AER

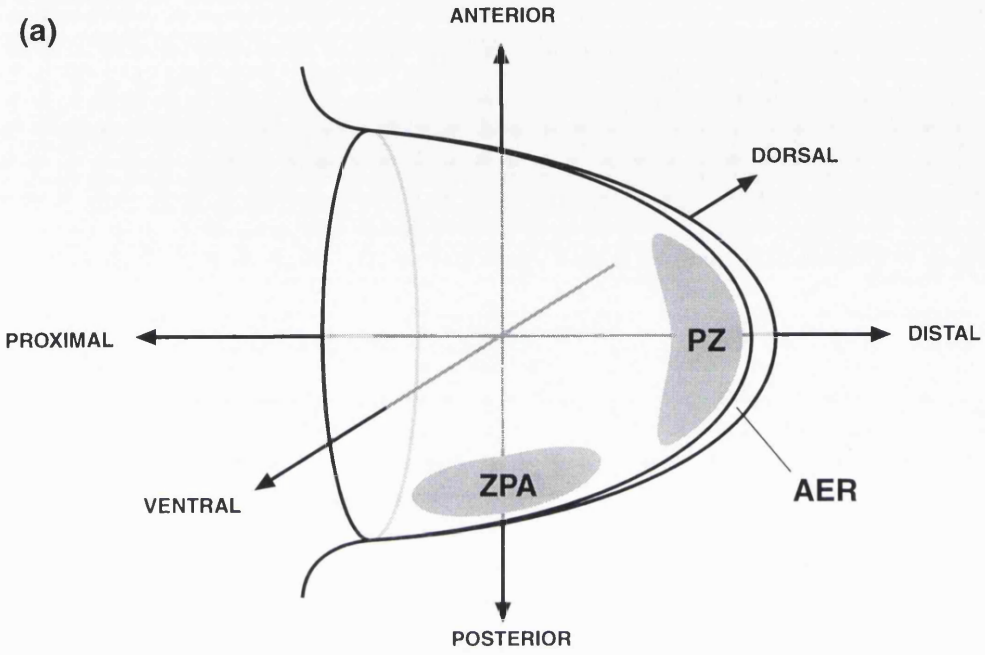
Overlying the mesenchyme cells of the PZ is a strip of the ectodermal layer which is thickened to form a ridge, known as the apical ectodermal ridge (AER). It is orientated along the anteroposterior axis and is present from the earliest stages of limb bud outgrowth (Fig. 1.5a). It is composed of pseudostratified elongated cells that are closely packed and linked by extensive gap junctions (Fallon and Kelley 1977) and is therefore rigid compared to the surrounding epithelium and this probably allows it to maintain the flattened plate structure of the distal limb bud. The results of many experiments (mostly performed in the chick) suggested that its major role is to maintain and regulate the proliferative state of the underlying PZ and thereby control the growth and extension of the limb bud, especially in the proximodistal direction (although a more detailed understanding of its other functions is described in section **1.8.3**). Removing it at any stage during development prevents further outgrowth of the limb bud, resulting in truncated limbs (Saunders, 1948; Summerbell, 1974). Although signals from the AER are essential to continued outgrowth and therefore continued progression of proximodistal differentiation, they do not impart P-D positional information themselves. In experiments recombining limb buds with AERs of different ages or different species it is always the age or species of the PZ that determines the structures formed (Zwilling 1959), (see section **1.8.4** on P-D patterning mechanisms).

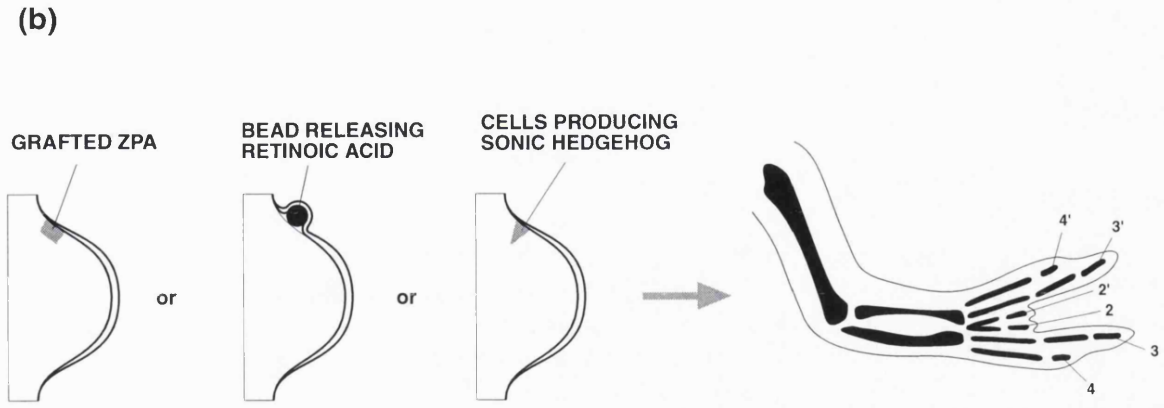
It has also been known for many years that as well as transmitting a signal to the PZ, the AER is maintained by a signal from the mesenchyme, originally hypothetically named AEMF (apical ectodermal maintenance factor), and this reciprocal interaction between mesenchyme and ectoderm was first proposed in detail as the Saunders-Zwilling hypothesis (Zwilling, 1961). The mesenchyme-to-ectoderm signal is in fact strong enough to initiate an ectopic AER in lateral ectoderm that would not normally produce a limb, and thereby cause an ectopic limb (Kieny, 1968).

Fig. 1.5 Organisation of the early vertebrate limb-bud.

(a) The limb-bud is composed of two cell types: an endothelial layer of ectodermal cells surrounding the loosely-packed mesenchyme. The apical ectodermal ridge (AER) is a thickened strip of the ectoderm which is orientated in an A-P orientation along the most distal end of the bud. Within the mesenchyme are two identifiable regions: the progress zone (PZ) which lies just under the distal ectoderm, and the zone of polarising activity (ZPA) which lies in the posterior-most region, just proximal to the PZ. As the bud extends, all three regions move in the distal direction maintaining their spatial relationships with each other through molecular communication involving the products of genes such as *Shh* and *Fgf-4*.

(b) Mirror-image duplications of the chick wingbud can be produced by grafting either an ectopic ZPA, or a bead soaked in retinoic acid, or cells expressing *Shh* into the anterior region.


(a)
ANTERIOR
DORSAL
PROXIMAL
PZ
DISTAL
ZPA
AER
VENTRAL
POSTERIOR


(b)
GRAFTED ZPA
or
BEAD RELEASING
RETINOIC ACID
or
CELLS PRODUCING
SONIC HEDGEHOG
4'
3'
2'
2
3
4

There are a number of putative intercellular signalling molecules expressed in the AER which may be responsible for this signal to the PZ: bone morphogenetic proteins (which are members of the TGF-b superfamily) BMP-2 (Lyons *et al.*, 1990) and BMP-4 (Jones *et al.*, 1991) and fibroblast growth factors: FGF-2 (Fallon *et al.*, 1994), FGF-4 (Niswander and Martin, 1992) and FGF-8 (Crossley and Martin, 1995; Mahmood *et al.*, 1995; Crossley *et al.*, 1996). So far, the strongest evidence supporting any of these candidates is for FGF-4: In short-term organ cultures it can maintain distal outgrowth of mouse limb buds that have been stripped of their AER (Niswander *et al.*, 1993), and outgrowth of chick wing buds whose AERs have been removed can be rescued by application of beads soaked in FGF-4 protein. However, expression of FGF-4 is not evenly spread throughout the AER: it is more concentrated in the posterior region, and both FGF-2 and FGF-8 which do display homogeneous distributions may also perform the PZ-maintenance function, as similar experimental results have been reported for them. The application of BMP-2 produces the opposite effect, suggesting that both mitogenic and inhibitory signals are used by the AER to control the PZ (Niswander and Martin, 1993).

In addition to the task of maintaining the proliferative state of the PZ, some of these proteins, in particular FGF-4 and FGF-2 are involved in interactions with the zone of polarising activity (ZPA), and this is the probable reason for the non-even distribution of FGF-4 in the AER.

1.7.2 The ZPA

The polarising region, or ZPA (zone of polarising activity) was first discovered in the chick by Saunders and Gasseling (1968). They found that when a small region of the posterior limb bud was grafted into the anterior region of a host limb bud, a drastic but organised change in development was seen. Instead of developing the normal 3-digit pattern, which displays a digit order of *234* (from anterior to posterior) they created limbs which had 6 digits in the order *432234* (Fig. 1.5b). In addition to the normal three digits, a mirror-image duplication of the wing had occurred. This immediately suggested an important principle by which anteroposterior patterning of the limb was occurring - that a signalling centre was asymmetrically located on one side of the limb, whose signal was transmitted across the limb field (directly or indirectly), and that cells within the field were thus given the positional information they required to differentiate into the correct tissue-type (Tickle *et al.*, 1975). Analysis of the manipulated wings showed that the extra digits were

derived from host tissue, and that the transplanted cells were therefore acting as the source for a signal.

Cells of the ZPA cannot be distinguished from the surrounding mesenchyme by histology, but their distribution has been mapped by cutting out different regions and testing their activity in the digit-inducing assay (eg. Hinchliffe and Sansom, (1985), Honig and Summerbell, (1985)). From these studies it is known that the activity is first present in cells of the lateral body wall mesenchyme long before the limb bud develops. During limb bud extension it is consistently found in the patch of mesenchyme cells just proximal to the PZ on the posterior side, despite the fact that this region is continually moving distally following the PZ. This means that either polarising activity is only transiently expressed in cells which have just left the posterior side of the PZ, or that a stable cell population of ZPA cells is continuously being displaced distally. The polarising activity disappears when the AER regresses.

Initial attempts to understand the nature of the ZPA signal involved many grafting experiments. Grafting different numbers of cells (Tickle, 1981) or exposing host tissue to donor cells for varying lengths of time (Smith, 1980) showed that both of these factors effect the strength of the duplications in a graded fashion, suggesting a dose-dependent response to the hypothetical signal. It was also found that whether placed anteriorly or distally the donor cells always induced more posterior characteristics (digits) closer to themselves, and more anterior ones further away (as is the case in normal development).

Retinoic acid (RA) was discovered to mimic the ZPA activity when applied to the anterior margin of the limb bud (Tickle *et al.*, 1982) and in these experiments was found to be distributed in a gradient. Also, although retinol, the biosynthetic precursor of RA, is evenly distributed across the limb bud, endogenous RA appears to be distributed in a slight gradient across the A-P axis, with posterior tissue possessing a 2.6-fold higher concentration than anterior tissue (Thaller and Eichele, 1988). Until recently, these data made RA the strongest candidate for the hypothetical ZPA morphogen, as it fitted-in perfectly with the gradient hypothesis of Tickle *et al.*, 1975 (see section **2.4.2**).

However, molecular analysis has discovered a protein with at least as strong a claim to be the key ZPA signalling molecule as RA. *Sonic hedgehog* (*Shh*) was cloned by homology to the *Drosophila* gene *hedgehog*, which was known to be involved in cell signalling during segmentation (Echelard *et al.*, 1993; Krauss *et al.*, 1993; Riddle *et al.*, 1993). It is expressed specifically in the ZPA cells of the limb bud as well as other sites of signalling activity in the embryo such as the

notochord (where it induces the floorplate in the ventral neural tube), but it is not expressed in the early flank mesoderm cells which display polarising activity in grafting assays (Hornbruch and Wolpert, 1991). Its expression in key signalling tissues (Roelink *et al.*, 1994; Lopez-Martinez *et al.*, 1995), as well as its induction activity in *in*-vitro experiments (Fan *et al.*, 1995; Roelink *et al.*, 1995) provide overwhelming evidence that this protein is indeed a diffusable morphogen. Transferring *Shh*-expressing cells into the anterior edge of a developing limb bud recreates the same effect as transferring ZPA cells or RA-soaked beads: a mirror-image duplicated wing is induced. And it is now known that the earliest detectable effect of RA-soaked beads on the anterior mesenchyme is to induce *Shh* (Riddle *et al.*, 1993), therefore bringing into question whether RA is actually a morphogen, or is important in a different way.

1.7.3 Interactions between the ZPA and the AER

Tickle (1981) showed that close contact of donor ZPA with host AER is important for a strong induction, suggesting that some interaction between the two tissues may be occurring during normal development. It is now known that a second function of the FGF signals from the AER (in addition to controlling proliferation in the PZ) is to maintain the ZPA. When attempting to rescue AER-removed limb buds using beads soaked in FGF-4, beads placed at the most anterior part of mesoderm can maintain outgrowth but fail to recreate normal patterning. However, when Niswander *et al.* (1993) added a second bead in a more posterior position (near where the ZPA should be), the limbs were now patterned correctly, because the ZPA was maintained. Additionally, if cells from the ZPA region are tested from a limb bud whose AER has been removed, they will not display polarising activity in grafting assays, but if after removing the AER an FGF-4-soaked bead is placed on the ZPA mesoderm, the cells retain their polarising activity (Vogel and Tickle, 1993). In agreement with this, it is likely that *Shh* is a downstream target of FGF-4, as removal of AER from mesenchymal cells results in a down-regulation of *Shh* transcripts (Laufer *et al.*, 1994).

This positive influence of FGF-4 on *Shh* expression, is now believed to be part of a reciprocal feedback loop. The normal expression of *Fgf-4* in the AER is concentrated posteriorly. When *Shh*-expressing virus (Laufer *et al.*, 1994) or *Shh*-expressing cells (Niswander *et al.*, 1994) are inserted into the anterior wing bud margin proximally to the AER, it induces expansion of the morphological AER and expression of *Fgf-4* in a strip which is outside its normal expression domain. These data, in conjunction with the spatial and temporal correlation seen between *Fgf-4* and

Shh expression in the normal situation, suggest very strongly that *Shh* acts upstream of *Fgf-4* as well as downstream. This feedback loop between the ZPA and AER, through *Shh* and *Fgf-4* is proposed to be the mechanism by which A-P pattern formation and P-D growth control are integrated into one process. In other words, the cross-regulation ensures that the two processes occur at the correct rate with respect to one another.

Laufer *et al.* (1994) have determined that a third function of the FGF signals from the AER is to confer competence of the underlying mesoderm to respond to *Shh* from the ZPA. Replication-competent *Shh*-expressing virus was used as a means to ectopically express *Shh* in anterior regions of the limb bud, and *Bmp-2* and *Hoxd* genes were used as downstream markers of *Shh* activity. If the injection was far from the overlying AER, no induction of *Bmp-2*, *Hoxd-11* or *Hoxd-13* occurred (Fig. 1.6). Also, if the injection was at a position which would usually lead to duplication, but the anterior half of the AER had been surgically removed, then no induction of the putative targets occurred, and under the influence of the remaining posterior half of the AER the limb bud developed almost normally. Strengthening the evidence that the important missing signal is FGF-4, was the observation that anterior induction of *Bmp-2* and the *Hoxd* genes in this experiment was restored if an FGF-4-soaked bead was placed where the anterior AER had been removed from.

1.8 Patterning the limb

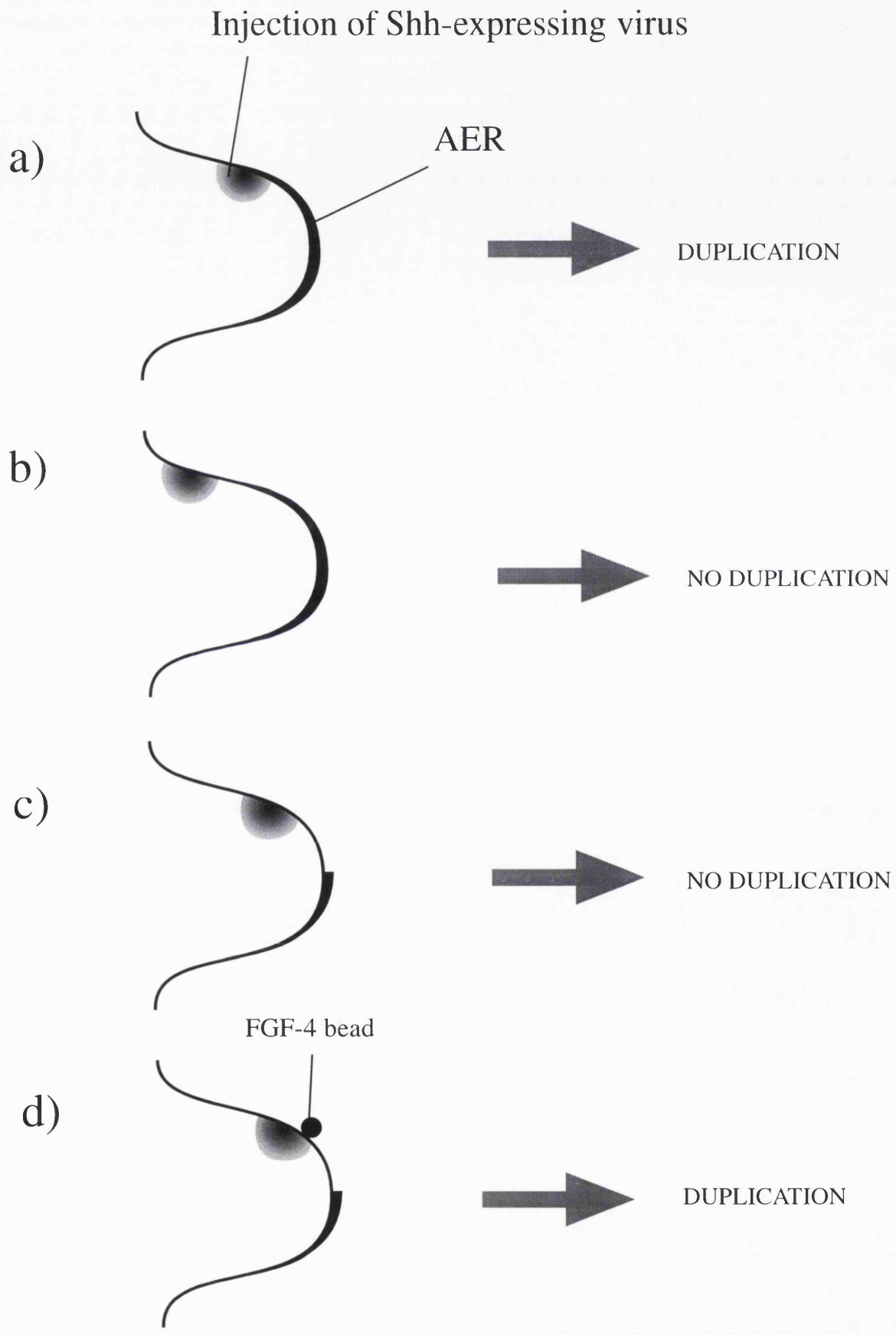
Originally, theories explaining A-P and P-D patterning considered the two systems to be completely independent (Summerbell *et al.*, 1973; Tickle *et al.*, 1975). Because of our current knowledge of the feed-back interactions between AER and ZPA (described above) and information about *Hox* gene involvement in both processes (described below in sections **1.8.3** and **1.8.4**), we now view the two systems as being connected and possibly inter-dependant. However, the principles by which the two axes are patterned can still be considered separately. A number of genes and gene families have been implicated in specific patterning processes, and these will briefly be reviewed here, along with some theoretical work on the patterning principles involved.

1.8.1 Dorsoventral patterning

Since the grafting experiments of (Pautou and Kieny, 1973) and (MacCabe *et al.*, 1974) it has been known that D-V patterning is mediated by signals from the ectoderm acting on cells in the

Fig. 1.6 FGF-4 protein from the AER stimulates competence in the underlying mesenchyme to respond to the *Shh* signal from the ZPA.

If replication-competent virus expressing *Shh* is injected into the chick wingbud, just under the anterior edge of the AER (a), it will induce mirror-image expression of the *Hoxd* genes, and a mirror-image duplication of digits. If this injection is made in a more proximal position, where there is no overlying AER (b), then the mesenchymal cells do not respond. If the injection is made in the same position as the first experiment, but the anterior part of the AER is cut away (c), then still no duplication occurs, because the mesenchymal cells are not recieving the *Fgf-4* signal. If however this anterior part of the AER is replaced by a bead soaked in FGF-4 protein (d), then a duplication will occur.


Injection of Shh-expressing virus
a)
AER
DUPLICATION
b)
NO DUPLICATION
c)
NO DUPLICATION
FGF-4 bead
d)
DUPLICATION

progress zone. More recently it has been shown that *Wnt7a* is a critical signal emanating from the dorsal ectoderm, whose absence results in limbs which have a ventral phenotype on both surfaces (Parr and McMahon, 1995), and that it activates the gene *Lmx1* in the dorsal mesoderm (Riddle *et al.*, 1995; Vogel *et al.*, 1995). *Lmx1* is related to the *Drosophila* gene *apterous*, which intriguingly is also expressed in the dorsal compartment of *Drosophila* wing discs. Whether other signals are involved, for example a ventralising signal, is not known.

1.8.2 Anteroposterior patterning - theories

Of the three limb axes, the A-P axis is the one most relevant to this thesis. Much work has been performed on it, for the following reasons: Firstly, it is the axis which distinguishes the digits and therefore is probably involved in specifying their differences (as well as those of the carpals or tarsals). Interest in digit specification is derived from a common view of the tetrapod limb as an archetypal example of adaptive radiation. Secondly, it seems that the molecules and mechanisms involved are similar to those which pattern the primary A-P axis of the whole embryo. In fact, some schemes (Duboule, 1992) consider the limb A-P axis to be a lateral extension of the primary axis. There are two main similarities: the response to RA, and a temporally and spatially colinear activation of *Hox* genes. (The important signalling molecule *Shh* does not fit into this common mechanism, as it is expressed evenly along the entire length of the notochord and floorplate and is a component of the D-V patterning system (Roelink *et al.*, 1994)). Thirdly, when the expression patterns of the *Hoxd* genes were first discovered, it seemed that they would fit perfectly into existing "positional information" (PI) models of A-P patterning, as representing the different morphogen threshold states, and consequently much attention was put on them. As explained later, subsequent studies have proved these original theories to be inadequte.

The ZPA is considered to be the signalling centre responsible for A-P patterning (see section **1.7.2**). Its localisation to the posterior margin of the limb bud is critical for this function and an experiment by (Charite *et al.*, 1994) has suggested that *Hox* gene expression along the primary axis is responsible for this localisation. In mouse lateral plate mesoderm *Hoxb-8* is usually expressed up to an anterior boundary adjacent to where the forelimb ZPA will form. When this expression is extended anteriorly (under the transgenic control of the *RARβ2* promoter) so that it now lies next to the anterior part of the limb, a second ZPA is induced in this anterior region and a

mirror-image duplicated limb develops, suggesting the involvement of *Hoxb-8* in this specification. Presumably a different *Hox* gene is responsible for the same function in the hindlimbs.

Informative as this is, the real debate which has intrigued scientists for so long is how the ZPA exerts its influence over the limb bud field - how it dictates positional information - and since the mid-70s two alternative theories have been competing to explain this. Although variations exist, the two extremes are described as follows:

1) The gradient hypothesis, (Tickle *et al.*, 1975) centres on the principle of a single substance which diffuses from the posterior ZPA across the limb bud to the anterior margin (Fig. 1.7a). Due to metabolism of the substance (either throughout the field, or only at the anterior side) the diffusion creates a concentration gradient and each position along the A-P axis experiences a different strength of this signal. Wolpert's original idea (Wolpert, 1969) of positional information (PI) is used in this model: positions are represented as thresholds in sensitivity of receptors to a morphogen, and consequently each position can be encoded by a gene (for the corresponding receptor). Thus the system can be thought of as a two-step process: global specification of positional information, followed by local interpretation.

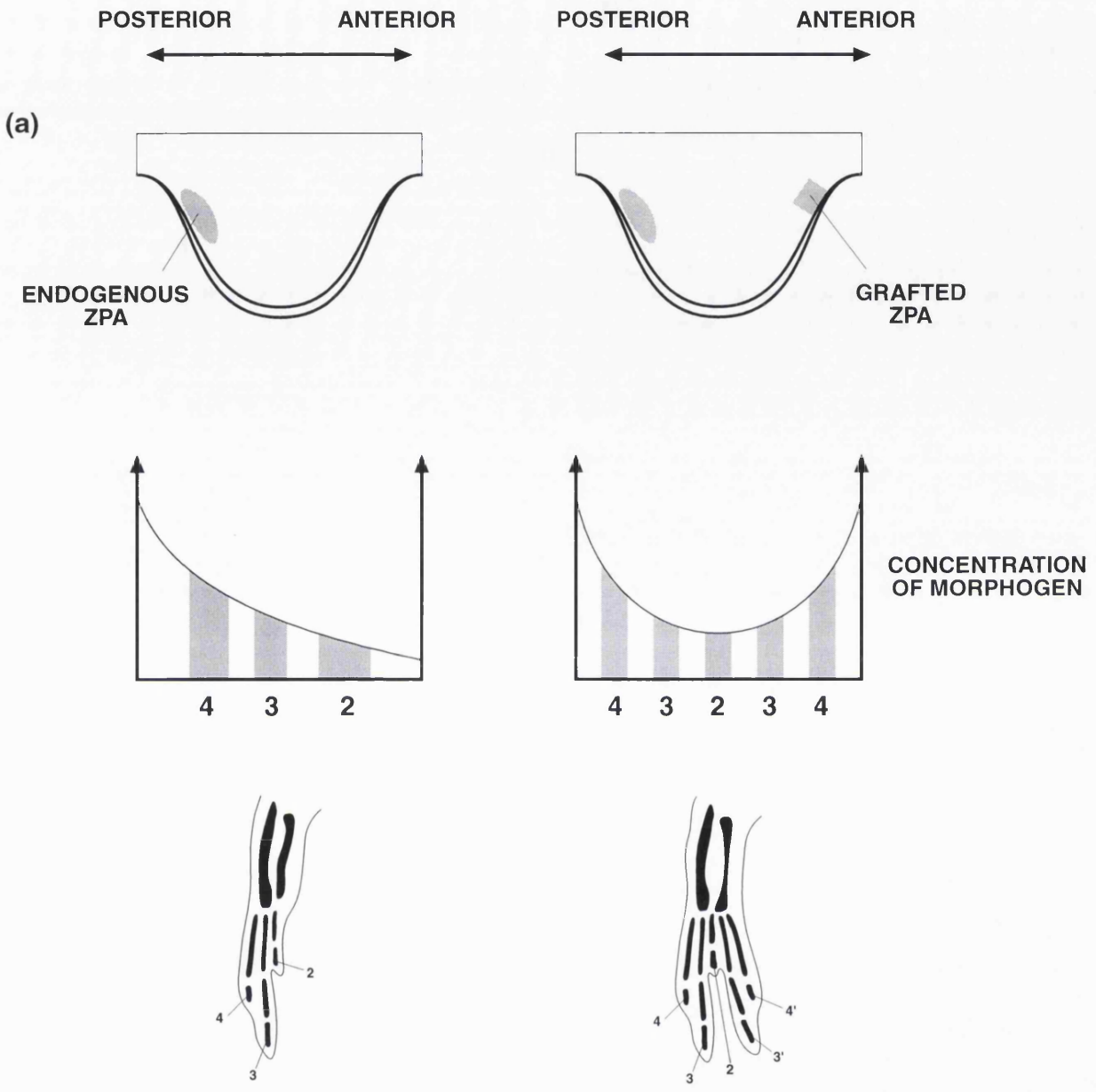
2) The polar coordinate model (PCM), (French *et al.*, 1976; Bryant *et al.*, 1981) stipulates a coordinate system in which the limb is considered as a cone extending away from the body wall such that the P-D axis is represented along its length, and the transverse axes (A-P and D-V) are represented across a cross-section of the cone (Fig. 1.7b). Positional values are then given to the circumference of the cross-section (which represents the ectoderm or outer layer of mesoderm), and patterning information is assumed to derive from this ring. The most important feature of the model is that it operates by *intercalation*, and therefore requires only short-range communication between cells. If positions around the circumference are numbered from 1 to 12, then cells always act to regenerate a continuous sequence of positional values. For example, if region 3 and region 7 are grafted together, then the regions 4, 5 and 6 will regenerate in between. (It is assumed in the model that the shortest route is used, so that the alternative intercalation of 2-1-12-11-10-9-8 is not chosen.)

Both models use the ZPA - the gradient model uses it as the source of the morphogen, and the PCM as a reference point for the coordinates - and both models have been useful for understanding the concepts behind patterning and successful in explaining certain experimental data. The gradient hypothesis correctly predicts the number, order and identity of digits created when the grafted ZPA is placed at different distances from the endogenous one, using the idea of a double-peaked morphogen gradient (Summerbell, 1979; Tickle *et al.*, 1975). The PCM was originally

Fig. 1.7 Two models to explain the patterning of vertebrate limbs.

(a) In the gradient hypothesis (Tickle *et al.*, 1975) spatial positional values are taken as threshold responses to the concentration gradient of a morphogen. The first case shows the proposed wildtype situation, in which the gradient is is a single slope eminating from a single ZPA. This causes every point along the A-P axis to have a unique positional value, and these values are then interpreted into the relevent cell-types. This means that genes for cartilage differentiation would have to distinguish between a number of different concentration ranges (shaded and non-shaded regions in the diagram). The result of posessing a ZPA at both ends of the axis would be a symmetrical gradient in which each positional value is allocated twice in a mirror-image pattern.

(b) The coordinate system used in the PCM (the polar coordinate model, Bryant *et al.*, 1981) is represented by a cone, in which the proximo-distal axis is labelled with the letters A to D, and the antero-posterior and dorso-ventral axes are combined into a single radial coordinate, numbered 1 to 12.


POSTERIOR
ANTERIOR
POSTERIOR
ANTERIOR
(a)
ENDOGENOUS ZPA
GRAFTED ZPA
4 3 2
4 3 2 3 4
CONCENTRATION OF MORPHOGEN
2
4
3
4
4'
3'
3
2
(b)
12
11
10
9
8
7
6
D
C
B
A
12
1
2
3
4
5
6
7
8
9
10
11

developed as the result of experiments designed to understand limb regeneration in insects, and its greatest success in vertebrates has also been in predicting the orientation and number of supernumerary limbs in grafting experiments.

However, neither are now considered to be adequate explanations. Although the general principles of the gradient hypothesis still gather support from experimental data (mostly through the work on RA and *Shh*), a couple of reasons suggest it is not a full explanation. Firstly, it does not explain the results of the grafting experiments which are correctly predicted by the PCM. Secondly, in its original form, the concept of precise threshold responses is used to explain how the digit pattern is established. This predicts that very thin or very thick digit bones should form in many of the ZPA-grafting experiments, and this has never been seen. It is recognised now that the real mechanism is more regulative, and involves more local cell-cell communication than proposed in the original global PI hypothesis.

The PCM, while successfully explaining many observations on the basis of local cell interactions, nevertheless does not agree with other experimental data. For example, the "shortest intercalation rule" stated that regenerating tissue would use the shortest route of positional values, and this is not always the case (Maden and Turner, 1978; Wallace and Watson, 1979). To account for these problems Meinhardt (1983) proposed a boundary model, in which key signalling molecules are proposed to emanate from the boundaries between two regions of differing positional value. This model integrates both diffusable morphogens and a degree of local intercalation and is significantly more successful in its predictions than the PCM (Maden, 1983). However, it has yet to be proven at a molecular level, and currently we are left with only one measurable signal centre and only one proven signalling molecule.

1.8.3 Anteroposterior patterning - molecular interpretation

The most important group of genes to be considered in the question of A-P patterning is the *Hox* gene family. Initial interest came from the report of nested expression domains of the 5' *Hoxd* genes (*Hoxd-9* to *Hoxd-13*) centred on the ZPA (Dolle *et al.*, 1989). *Hoxd-13* had the smallest domain extending not far from the ZPA itself, and each subsequent gene (sequentially along the *Hox* cluster) had a slightly larger domain with a slightly more anterior boundary of expression. This mirrors very closely the patterns of *Hoxb* genes found along the primary axis (see section **1.3**). This discovery fitted well with data from (Simeone *et al.*, 1990) which demonstrated that in cell culture

Hox genes are induced by RA in a colinear manner (see section **1.4**), ie. the paralogous groups are activated in sequential order, with *Hoxb-1* coming on first, followed by *Hoxb-2*, then *Hoxb-3* etc. It seemed that this property of differential sensitivity to RA, could easily explain activation of different *Hox* genes at different distances from a morphogen source, and that a neat link had been found between ZPA activity and discreet zones of patterning across the A-P axis. And because there were found to be precisely 5 different *Hoxd* genes in this region, which were expressed in 5 *Hox*-code zones, it prompted Tabin (1992) to write the article "Why we have (only) five fingers per hand: *Hox* genes and the evolution of paired limbs". It was proposed that the 5 zones correspond precisely to the 5 different digit types.

Early experimental data did not contradict this idea. It was discovered that the polydactylous chicken mutant *Talpid*, which lacks morphological distinctions between digits, does not display the normal nested patterns of *Hoxd* expression, but instead has an even distribution of all 5 genes across the entire A-P axis, thereby giving the same identity to each digit (Coelho *et al.*, 1992; Izpisua-Belmonte *et al.*, 1992). Chick limb bud manipulations that led to digit duplications were found to create mirror-image inductions of nested *Hoxd* genes in the anterior region (Izpisua-Belmonte *et al.*, 1991), and only a complete induction of all 5 genes would create a complete duplication (Izpisua-Belmonte *et al.*, 1992). A more direct approach was taken by Morgan *et al.*, (1992), who injected *Hoxd-11* expressing virus into the developing hindlimb bud, and thereby caused the first digit to devlop a digit-II morphology. This was the expected result because the *Hox* codes for digit I and digit II were thought to be 9+10, and 9+10+11 respectively, and therefore adding *Hoxd-11* to the digit I code would transform it to the digit II code.

However it has subsequently turned out that the nested domains discovered in the early analysis are only half the story, and the patterns before and after that embryonic stage paint a much more complicated picture in which *Hoxd* genes appear to be important in both A-P and P-D patterning. One of the most important pieces of evidence for this view was the targeted disruption of the *Hoxd-13* gene in mice (Dolle *et al.*, 1993). The presence of an extra carpal and a supernumery digit indicate a disruption in A-P patterning, but the absence of the second phalange in digits II and V indicates a disruption of the P-D patterning as well. Also the fact that digits II and V were severely affected while digits I,III and IV were not, strongly disputes both the notion of simple discrete *Hox*-code zones along the A-P axis, and the idea of a single gradient patterning system. Instead Dolle *et al.* (1993) propose that the main function of *Hox* genes is to regulate growth rates, and in this way control how much mesenchymal tissue is available in a given region for the

subsequent condensation process. Exactly how the *Hox* genes themselves are regulated is still unclear (section **1.5**) but Dolle *et al.* propose that their sequential, temporal colinear activation is critical to this mechanism of limb patterning, and that positional values arise because the overlapping expression domains impart information about their position in a "temporal referential", as opposed to a strictly spatial one. From this viewpoint they describe the *Hoxd-13* mutant phenotype as displaying a localised heterochrony, ie. the regions of the limb which are affected (reduced) are the ones usually last to develop, and this causes the limb to appear retarded. This model of A-P patterning is very similar to that proposed by Summerbell et al. (1973) for the P-D axis, in which positional values are determined by the time at which cells leave the PZ.

1.8.4 Proximodistal patterning

From experiments performed by (Summerbell *et al.*, 1973) it was suggested that the progress zone (PZ), which remains in an undifferentiated state, is involved in a timing mechanism which provides cells with information about their position along the proximodistal axis, ie. cells which leave the PZ earlier will adopt proximal fates and those leaving later more distal fates. This is suggested from the autonomous nature in which it functions in grafting experiments: Transplanting a young limb bud tip onto an old limb bud results in duplication of skeletal elements along the proximodistal axis, whereas an old-to-young graft results in missing elements.

When considering which genes may be involved in specifying (or recording) the different P-D positions as the limb extends, the *Hox* genes again are the most important candidates to date. Although patterns of *Hoxd* gene expression go through a stage (phase II) when they are nested along the A-P axis (approximately 10.5 dpc in the mouse), there is a time before this (phase I) when the early-activated genes (*d-9*, *d-10* and *d-11*) are arranged in zones along the P-D axis (Morgan and Tabin, 1994) suggesting their involvement in P-D axis patterning. Direct evidence for this involvement has come from a double targeted disruption experiment in which both *Hoxa-11* and *Hoxd-11* were mutated (Davis *et al.*, 1995). Due to redundancy between paralogous *Hox* genes (section **1.5.2**) the phenotypes of homozygotes for either single mutation were weak (Small and Potter, 1993; Davis *et al.*, 1994). However, the double homozygotes have a very dramatic phenotype in which the entire radius and ulna are absent. They proposed a patterning scheme in which each of the five 5' *Hox* paralogous groups is responsible for a different section of the P-D

axis (so that the sequence from group 9 to group 13 corresponds to: the scapula, the stylopod, the zeugopod, the wrist or ankle, and the digits).

In addition to phases I and II (described above) there is also a later stage (phase III), when expression domains twist distally and anteriorly such that they are again arranged along the distal P-D and may be involved in P-D patterning of the digits (see effects of *Hoxd-13* null mutant above). Sordino *et al.* (1995) believe that this final twist of *Hoxd* expression domains reflects an evolutionary change that was essential in the adaptation of fish fins to tetrapod limbs: Fish fins do not display this twist and do not possess a differentiated skeletal autopod. Their fin rays develop from a dermal skeleton which tetrapods have lost. These data support the original hypothesis by Shubin & Alberch (1986) that evolution of the tetrapod autopod was accomplished by a skewing of the teleost fin primary axis in a preaxial (or anterior) direction.

The recent discovery that the human type II synpolydactyly is caused by a mutation in *HOXD13* is further evidence of *Hox* involvement in this last phase of P-D and A-P patterning in the digits (Akarsu *et al.*, 1996; Muragaki *et al.*, 1996). Whereas heterozygotes display polydactyly and syndactyly (considered an alteration in A-P patterning), homozygotes display a reduction in digit length which is due to the metacarpals developing a smaller, rounder phenotype typical of carpals (which is an alteration in P-D patterning). The proposed explanation is that during this 3rd phase of *Hox* expression, when autopod specification is occuring, the distal expression domain of *Hoxa-13* includes the phalanges, metacarpals and carpals, whereas the similar domain of *Hoxd-13* only covers the phalanges and metacarpals. Consequently removal of the *Hoxd-13* expression domain results in the metacarpals experiencing the same *Hox* code as the carpals. This would suggest that the role of *Hoxd-13* is to specify long digit bones, again consistent with the idea it controls tissue growth rates.

Interestingly, the four different pedigrees studied (all with the same type of synpolydactyly) were created by four different mutations, but all involving the insertion of extra alanine residues into an alanine stretch found in the amino-terminal of the protein (7, 8, 9 or 10 extra residues were all observed). The fish *Hoxd-13* gene has no alanine stretch, the chicken *Hoxd-13* (known as *CHoxd-13*) has a stretch of 9 alanines, and the human *HOXD-13* has 15 alanines, and it has been suggested by Muragaki *et al.* that a progressive insertion of alanines during the evolution from teleosts to mammals accompanied the introduction of the phase III twisting of *Hoxd* expression domains, and that both processes were involved in the specialisation of the autopod. They quote data from (Han and Manley, 1993a; Han and Manley, 1993b) that similar alanine-rich stretches have been

associated with transcriptional repressive activity in the *Drosophila* proteins *even-skipped*, *engrailed* and *Krüppel*. Consequently altering the length of the alanine stretch in *Hoxd-13* proteins during evolution or mutation may alter its regulative function in a subtle, quantitative manner which would be consistent with slight changes in growth rates, as occurs in the human synpolydactyly (and which might even explain an increased growth in the distal posterior limb bud margin of tetrapods which could be responsible for creating the skewed primary limb axis (Sordino *et al.*, 1995)).

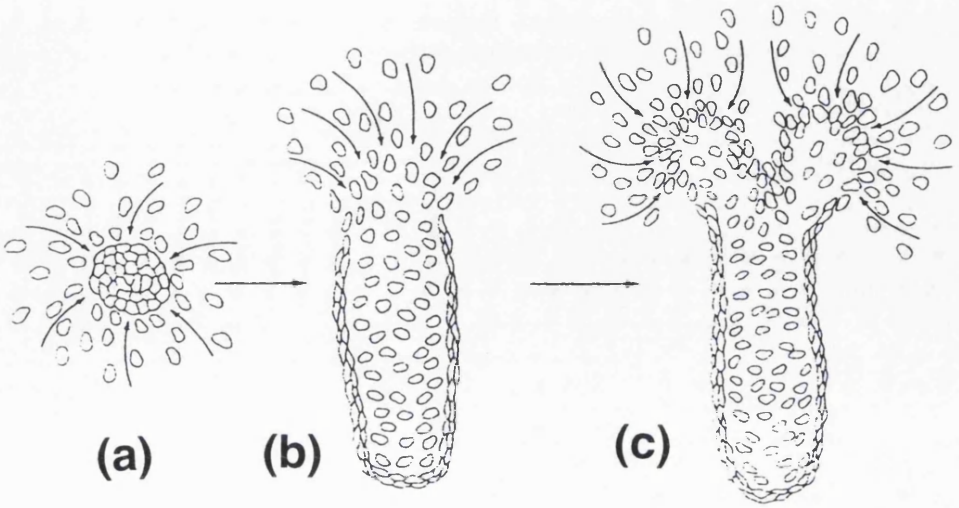
1.8.5 Condensations Patterns

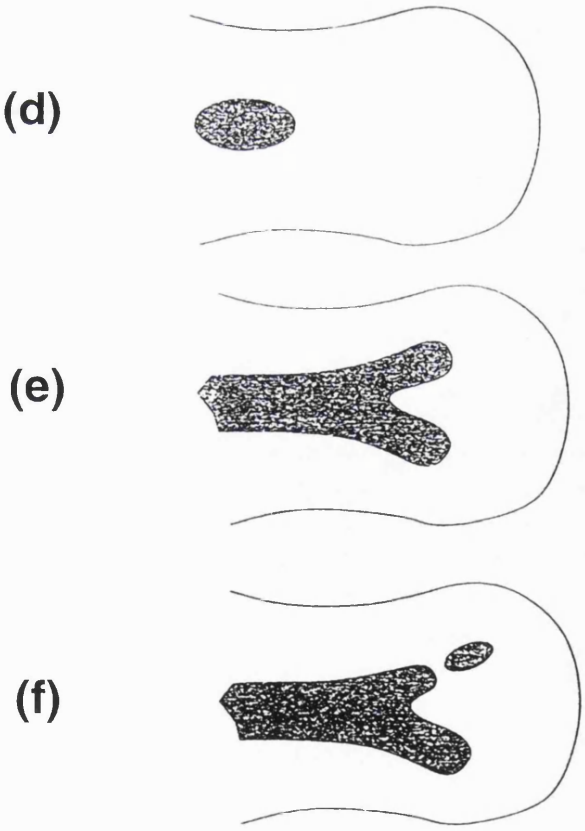
I have discussed how positional information may be generated along the A-P and P-D axes of the limb, but unlike the original PI models of Wolpert (1969) this molecular information is not simply translated into a map of different cell types which defines the structure of the developing limb, instead it is known that local cell communication and interaction is important. Experimental and theoretical work, primarily by Hinchliffe (1977, 1983), Oster (1983, 1985), Shubin & Alberch (1986) and Murray (1983, 1989) has developed the view that the organisation and pattern of the skeletal elements is the result of a progressive condensation of mesenchyme cells into cell aggregations, which extend in length in a proximal-to-distal direction through the loosely-packed mesenchyme (following the movement of the progress zone) by recruiting the surrounding cells into the tightly-packed conformation (Fig. 1.8). In addition to simple extension of these condensations, two other processes can occur: (1) the extension can undergo a *branching bifurcation*, in which the growing tip splits into two, resulting in a Y-shaped structure, (2) a constriction and separation can occur within the condensation itself, leading to a *segmental condensation*. In addition to these processes, new unconnected cell aggregates can form *de novo*, called *focal condensations*. Combinations of these events can generate any branching pattern seen in nature. (Hinchliffe studied these patterns in many different species using the uptake of $^{35}SO_4$ into chondroitin sulphate as the earliest marker of condensation sites; visualised by autoradiography.)

It is the progressive extension of these condensations through the developing limb and their various bifurcations which is modulated by the A-P and P-D information contained in the *Hox* gene expression patterns. It is expected that a default condensation program exists (determined by a gene network involving short-range signalling, cell adhesive molecules, etc.) which in absence of other signals would create a particular branching pattern, but which contains certain parameters (many of which relate to proliferation rates) that can be modified by the different *Hox* gene combinations. In

Fig. 1.8 Condensation patterns in the developing vertebrate limb.

A focal condensation event occurs when loosely-packed cells of the mesenchyme are recruited into a tightly-packed organisation (a) and (d). Subsequent elongation then proceeds in a directional manner, due to the recruitment of new cells being restricted to one end of the condensation. This is achieved by a network of peripheral fibroblasts which surround the structure (b). Continued growth of the condensation can result in a *branching bifurcation* (c) and (e) or a *segmental bifurcation* in which a local constriction causes the structure to "bud" into two sections (f).


(a)
(b)
(c)


(d)
(e)
(f)

this way, the axis-patterning mechanisms discussed earlier can control both the organisation of the condensation pattern and the shape of individual bones, without having to specify exactly which cell type goes where. This allows for greater developmental and evolutionary flexibility, (ie. the system is very regulative) since small changes due to mutations or physical perturbations are less likely to have deleterious effects on the overall skeletal morphology. Experiments by Oster et al (1983) indicate that the cross-sectional area and shape of the limb bud may also modulate this default condensation program.

1.9 Polydactylous mouse mutants

Many mutant strains of mouse and chicken plus a few strains of other species (guinea pig, rabbit, cow and cat) have been found in which the number or organisation of digits in the limb is abnormal (Gruneberg, 1963; Johnson, 1980). Polydactyly is the most common disorder, resulting in supernumerary digits being formed, but ectrodactyly (a reduction in digit number) and syndactyly (fusion of digits) and combinations of these are also seen. Polydactylous mouse mutants can be divided into pre- and post-axial types, depending on whether the extra digits are considered to be anterior or posterior to the normal digits. Preaxial polydactyly appears to be more common (there are at least 15 strains displaying preaxial polydactyly, as opposed to only 2 specifically showing postaxial polydactyly), and it can be divided into three classes:

(1) Mutants which display abnormalities in other parts of the body in addition to the limb. For example, *Myelencephalic blebs* (*my*, on Chr.3, Gruneberg 1952) or *Tail-short* (*Ts*, on Chr.11, Doel 1961).

(2) Mutants in which only the digits are affected. For example, *Holt's polydactyly* (*py*, on Chr.1, Holt 1945).

(3) The hemimelia-luxate group, which is most relevant to this thesis and is discussed in the next section.

1.9.1 The Hemimelia-luxate mutants

This is the largest of the three groups listed above, and holds particular interest due to the similarity of characteristics seen in all members (Forsthoefel, 1958; Forsthoefel, 1962; Gruneberg, 1963; Johson, 1980). Three descriptive terms have been used to name most of these mutants:

hemimelic, luxate or luxoid, and all refer to the reduction or absence of the tibia or radius. In chronological order of first publication they are:

lx	(5)	luxate	Carter 1951
lu	(9)	luxoid	Green 1955
lst	(2)	Strong's luxoid	Strong & Hardy 1956
Dh	(1)	Dominant hemimelia	Searle 1964
Xt	(13)	Extra toes	Johnson 1967
Hx	(5)	Hemimelic-extra toe	Kalter 1980
Xpl	(X)	X-linked polydactyly	Sweet & Lane 1980

The mutant polydactyly Nagoya (Hayasaka et al., 1980) is not included in this list because it is believed to be allelic to *Xt*.

The common features of this group are:

(1) A reduction or absence of the tibia or radius often resulting in the corresponding fibula or ulna being bowed to compensate. Sometimes the fibula or ulna may also be reduced. Rarely a duplication of these bones may occur (in *Dh* and *lst*).

(2) Preaxial polydactyly, which has often been described as a mirror-image duplication. However, this is not exclusively the case, and occasionally ectrodactyly occurs instead.

(3) The earliest observable phenotype is seen before autopod condensations occur (at 11.5 dpc), and is a bulging of the limb bud on the anterior side where the later digit defects will delevop. It appears that this over-growth of mesenchymal cells could be intrinsic or induced by an enlarged AER.

(4) None of the mutants are recessive, however there is an incomplete penetrance of the phenotype on any genetic background (from 30% to 90%), and a single strain usually displays a wide range severities. For example, a homozygous luxoid hindlimb can develop anything from 7 to 4 digits on the same genetic background.

(5) In general there appears to be a gradation of sensitivity to the mutations along two spatial axes: (a) distal elements (digits) are affected more readily than proximal ones (the

hemimelia), as seen in the difference between heterozygotes and homozygotes in *lst*, *lx*, *lu* and *Xt*, and (b) hindlimbs are more readily affected than forelimbs (in *lx*, *lu*, *lst*, *Xt*, *Hx*).

From this early embryological analysis, predictions can be made regarding the underlying similarity between these mutants. The observation of mirror-image-like duplications obviously suggests involvement of the A-P-patterning system and the possibility of an anteriorly duplicated ZPA. However, there is quite clearly an asymmetry across the A-P axis of the limb, not only in structure, but also in variability: The posterior side always appears unaffected (and therefore invariant), whereas the anterior side varies greatly. There are two possible sources of this variability:

(1) The fact that mutations in 7 different loci all produce a similar phenotype suggests that the disrupted mechanism is complex, and may require a large number of genes not simply to construct a patterning process but to stabilise it.

(2) The variability may derive from the subsequent condensation program. The cross-sectional area and shape of the limb bud are known to affect this process, and it is likely that certain shapes (like that of the wildtype limb bud) intrinsically induce more regular branching patterns than others. This would allow the possibility that the mutant genes are involved in controlling mesenchyme proliferation rates in the anterior limb bud, as suspected by the bulges seen at 11.5dpc, and that this early abnormality is essentially regular.

Despite the similarities mentioned, there are also differences. They all differ slightly in comparative strength of the phenotype, but only the qualitative differences will be described here. The first mutant discovered, *lx*, only displays hindlimb abnormalities. The second two, *lu* and *lst*, are the most similar, both displaying only hindlimb defects in the heterozygotes, but both forelimb and hindlimb defects in homozygotes. The next mutant, *Dh*, has a similar heterozygous phenotype (although ectrodactyly is more common) but a lethal homozygous phenotype which interestingly does not affect the forelimbs. The next two discovered, *Xt* and *Hx*, are similar to each other in affecting all four limbs in the heterozygous state, but whereas *Hx* displays hemimelia as well as polydactyly, *Xt* only displays polydactyly. The homozygotes of both these mutants are very severe and die during embryogenesis, but it is only in this state that hemimelia is seen in the *Xt* mutant. *Xt* appears to be slightly different from the other hemimelia-luxate mutants, as all digits look the same and consequently it is hard to determine whether the polydactyly is strictly preaxial or not. The final mutant *Xpl* is interesting in that hemizygotes (males) and homozygotes (females) have

indistinguishable limb phenotypes which only affect the hindlimbs. (*Xpl*, *Dh* and *lx* are the only three which never affect the forelimbs.)

It is also relevant to this thesis that the *Hx* locus is linked very closely with *Hm* (Hammer-toe) and thought possibly to be allelic to it, despite the significant difference in phenotypes. *Hm* mice display webbing between digits 2, 3, 4 and 5, which during limb formation causes a pronounced flexing of the second phalange of these digits. Two observations indicate a potential similarity between these phenotypes: (1) in both cases hindlimbs are more severely affected than forelimbs (although this is a common feature in the hemimelia-luxate mutants), (2) the *Hx* phenotype is thought to be a consequence of disrupted patterns of programmed cell death (Knudsen and Kochar, 1981), and a reduction in cell death in the interdigital zones could also be responsible for the webbing in *Hm* mutants.

1.9.2 Molecular analysis of the hemimelia-luxate mutants

Despite advances in molecular techniques, the responsible gene has been identified in only one of the hemimelia-luxate mutants, and this is the less typical *Xt*. It had been proposed that *Xt* might be the murine equivalent of the human *Greig Cephalopolysyndactyly Syndrome* (*GCPS*), based on morphological and gene mapping studies (Winter & Huson, 1988), and also that GLI3 may be the responsible gene in humans (Vorkamp et al., 1991). When *Gli3* transcripts were analysed in *Xt* embryos, it was found that heterozygotes expressed only half the wildtype level, and homozygotes expressed none.

Gli3 is one of a family of 3 vertebrate genes (*Gli1-3*) which are *Krüppel*-related zinc finger genes (transcription factors) and thought to be involved in embryonic development and tissue-specific differentiation (Ruppert et al., 1988). They are most similar to the *Drosophila* segment polarity gene *cubitus interuptus Dominant*, which is thought to regulate other segment polarity genes such as *wingless* (*wg*) and *gooseberry* (*gsb*), (and this has lead by extension to the idea that *Gli* genes may regulate the vertebrate homologues of *wg* and *gsb*, the *Wnt* and *Pax* genes).

This discovery originally suggested that two of the other hemimelia-luxate mutants might have disruptions in *Gli1* and *Gli2*. However, despite extensive screening, and the initially encouraging result that *Gli2* is located in the same region of chromosome 1 as *Dh*, it is now believed that neither of these two genes is involved in the mutants (pers. comm. D. Hughes).

Interestingly the same group who studied the *Gli3* expression in *Xt* mice, also created a new recessive allele of the same locus, by random transgene insertion (Pohl et al., 1990, named *add* - anterior digit-pattern deformity). Although the phenotype displays fundamental similarities with the previously described mutants (preaxial polydactyly), it is unique in only affecting the forelimbs.

Recently another luxoid-like mutant has been created, named *Rim4* (Recombination induced mutant 4), (Masuya *et al.*, 1995). This displays all the typical characteristics of the hemimelia-luxate group, but maps to a different chromosome (chr. 6), and thus brings to 8 the number of different loci which display this common phenotype.

1.9.3 *Sonic hedgehog* in the hemimelia-luxate mutants

One possible cause of the polydactyly in these mutants was considered to be anterior duplication of the ZPA, so involvement of *Shh* has been recently studied. When mapped by Chang *et al.* (1994) they discovered it to be close to the loci of *Hm/Hx* and *lx* on chromosome 5. However, they could detect no changes to the gene in *Hm* or *Hx* mice. More recently the expression patterns of *Shh* in four mutants (*lst*, *Rim4*, *Hx* and *Xt*) was analysed (Chan *et al.*, 1995; Masuya *et al.*, 1995) and in every case expression was found in a small domain on the anterior side of the limb bud, in addition to its normal site in the ZPA. They also found the expected anterior expression of *Hoxd-11*, thought to be downstream of *Shh* in the A-P patterning mechanism (see section **1.8.3**). This is strong evidence that the polydactyly in all these cases can be considered as the result of a misplaced anterior ZPA which alters the positional information in this region. The altered information (probably carried in the expression of *Hox* genes) may affect proliferation rates of the undifferentiated mesenchyme, thereby causing a change in tissue geometry. Both the altered geometry and the altered *Hox* code itself may then influence the parameters of the condensation program, to create the abnormal skeletal phenotype.

It is not proven that all the hemimelia-luxate mutants posses an anterior ZPA, but it currently seems very likely. This raises the following intriguing question: Why would mutations in 8 different genes all perform this quite specific event? One common proposal (Johnson, 1980; Masuya *et al.*, 1995) is that ZPA induction is a symmetrical process which "attempts" to create a ZPA at both ends of the limb bud, but is prevented from doing so at the anterior side due to the presence of repressive signals or the absence of cooperative ones. This view would treat *Hoxb-8* as a permissive signal rather than an instructive one (see section **1.8.2**). Johnson also suggested that the

increased symmetry of fish fins compared to tetrapods might indicate that a double ZPA was the original, primitive state, but it is now known that fish fins are also patterned by a single posterior ZPA (Krauss *et al.*, 1993; Sordino *et al.*, 1995).

However, the idea of these genes acting as repressors is more attractive than the converse, because if they are activators it would appear highly coincidental that they all display gain-of-function mutations which cause ectopic expression in the same place. Consequently it is considered more likely that most of them are repressors whose normal expression domain includes the anterior limb region. This could allow all the cases to be explained by a predisposition of only one gene (*Shh*) to be active in that region. In agreement with this are the early findings of (Buscher *et al.*, 1996), who believe that the GLI3 protein, encoded by the *Xt* gene, is a repressor of *Shh*, and is expressed throughout the limb-bud except for the position of the normal ZPA.

The other possibility is that the effect of these genes on ZPA regulation is indirect, for example, ZPA establishment may involve a community effect (Gurdon *et al.*, 1993). Their direct function could then be in regulating any number of different tissue characteristics such as cell adhesion, proliferation rates, cell-cell signals, etc., all of which could influence critical threshold parameters in the ZPA community effect.

1.10 The aims of this project

Finding enhancers which recreate the endogenous expression patterns, has proved successful for many of the *Hoxb* genes. Jenny Whiting (in the lab) had demonstrated that a 17kb region of DNA encompassing the *Hoxb-9* gene was able to direct the majority of the wildtype pattern, but was lacking the late neural tube expression. The first project in my PhD was to refine this enhancer analysis, and to extend the tested region to see if local elements could account for the whole pattern.

Whiting et al. (Whiting *et al.*, 1991), discovered that the main local regulatory elements for the *Hoxb-4* gene, were located within the intron and 3' of the transcription unit. Together they were able to recreate all of the basic *Hoxb-4* pattern in LacZ-reporter constructs, with one element being responsible for the correct anterior boundary of expression in the neural tube, and a second directing the boundary for the somites. Expecting a similar set of enhancers to exist for the regulation of *Hoxb-5*, Stefan Nonchev in the lab, performed a similar series of transgenic experiments, in which different regions around the *Hoxb-5* gene were tested. He defined two further regions, which

appeared to be responsible for mesoderm and neural expression respectively. These regions are in the intergenic DNA between *Hoxb-5* and *Hoxb-4*, and in further transgenic experiments were shown to be able to drive the *Hoxb-4* promoter at least as well as the *Hoxb-5* one. Thus arose several questions: Since these enhancers are near to both the *Hoxb-4* and *Hoxb-5* promoters, which one do they normally control? Could they be important to both? If they should only act on one promoter, how is interaction with the other one prevented? Answers to these questions could be relevant to the tightly-clustered organisation of the complex, because if enhancers need to be "shared" between two adjacent promoters (or possibly between more than two genes) then this could be a reason why the complex cannot split-up during evolution.

I attempted to answer these questions by developing a double-reporter system, in which both LacZ and PLAP (human placental alkaline phosphatase) are used to monitor the expression from two different promoters. Large constructs were made containing both the *Hoxb-4* and *Hoxb-5* genes. LacZ was inserted after the *Hoxb-5* gene, and PLAP after *Hoxb-4*, and then enhancer regions were removed to see if their loss affected both genes.

In the course of testing the PLAP reporter by making transgenic mice with regulatory regions of *Hoxb-1*, an insertional mutant was created which displays preaxial polydactyly typical of the hemimelia-luxate group. In this part of my PhD research I sought to characterise this new strain morphologically, compare it to the previously described *hemimelia-luxate* mutants, determine whether it represents identification of a new gene or is a new allele of an old mutant, and study the expression of *Shh* in the embryonic limbs. In addition, due to the fact that the transgene contained the reporter gene PLAP (human placental alkaline phosphatase) I have been able to gather unique expression data not available for the classical mutants described above. For the purpose of this thesis, the mutant strain will be named *sasquatch* (*sas*).

CHAPTER 2

MATERIALS AND METHODS

2.1 Standard solutions

Acrylamide solution (40%)	380g acrylamide (sequencing grade), 20g N,N'-(40%) methylenebisacrylamide, distilled water to 600ml. The solution was heated to 37°C to dissolve compounds, and volume the made up to 1l with dH_2O. The solution was filter sterilised and stored in the dark at RT.
Acrylamide/Urea solution (6%)	75ml of 40% acrylamide solution, 25ml 10x TBE, 230g Urea made up to 500ml with dH_2O, filter sterilised and stored at 4°C.
PLAP Staining Buffer	100mM Tris pH 8.5, 100mM NaCl, 50mM $MgCl_2$.
Alkaline SDS	0.2M NaOH, 1%SDS (w/v).
Ampicillin	Dissolved at 50µg/ml in dH_2O, filter sterilised and stored at -20°C. Used in media and agar at a final concentration of 50ml/ml.
Avertin (100%)	10g 2,2,2-Tribromoethanol dissolved in 10ml tertiary amyl alcohol. Stored in dark at 4°C.
BCIP solution (X-phosphate solution)	50mg/ml 5-bromo-4-cloro-3-indolyl-phosphate in DMF.
"Blue Juice" (x10)	50% glycerol (v/v), 20mM Tris.Cl, 20mM EDTA (pH 8.2), 0.1% Bromophenol blue.

Denaturing Solution	0.5M NaOH, 1.5M NaCl.
EDTA, 500mM	pH adjusted to 8.0 and autoclaved.
X-Gal stock solution	40mg/ml X-Gal dissolved in DMF, stored at 4°C.
X-Gal staining solution	5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6.3H_2O$, 2mM $MgCl_2$, 0.01% (w/v) sodium deoxycholate, 0.02% NP40, 1mg/ml X-Gal (from stock), stored at 4°C.
Hybridisation buffer (for whole-mount ISH)	50% deionised formamide, 5x SSC, 2% blocking reagent (Boehringer #1096176), 0.1% Triton X-100, 0.1% Chaps, 50µg/ml Heparin (grade 1-A, Sigma H-3393), 10mg t-RNA (Type VI Sigma R6625), 5mM EDTA.
Injection buffer	10mM Tris.Cl (pH 7.6), 0.1mM EDTA.
KTBT	50mM Tris.Cl (pH 7.5), 150mM NaCl, 10mM Kcl, 1% Triton X-100.
L-agar	1% (w/v) bacto tryptone, 0.5% bacto yeast extract, 0.5% NaCl, 1.5% bacto-agar.
L-broth	As L-agar but without bacto-agar.
Ligation buffer (x10)	200mM Tris.Cl (pH 7.6), 50mM $MgCl_2$, 50mM DTT.
Lysis buffer (for maxi-prep)	50mM glucose, 25mM Tris.Cl (pH 8.0), 10mM EDTA.
Lysis buffer	100mM Tris.Cl (pH 8.5), 5mM EDTA, 0.2% SDS, 200mM

(for tail DNA prep)	NaCl, 100μg/ml Proteinase K.
NBT solution	75mg/ml NBT in 70% DMF.
dNTP solution (x10)	250μM each of dATP, dCTP, dGTP and dTTP.
Neutralising solution	0.5M Tris.Cl (pH 7.6), 3M NaCl.
New Wash	50% ethanol, 0.1M NaCl, 10mM Tris.Cl (pH 7.5), 1mM EDTA.
NZY+broth	10g NZ amine (casein hydrolysate), 5g yeast extract, 5g NaCl in 1l dH_2O, autoclaved. Just prior to use add (per 100ml): 1.25ml 1M $MgCl_2$, 1.25ml 1M $MgSO_4$ and 1ml 2M filter-sterilised glucose.
PBS (x10)	1.3M NaCl, 70mM Na_2HPO_4, 30mM NaH_2PO_4, pH adjusted to 7.0 and autoclaved.
PCR salt solution (x10)	100mM Tris.Cl (pH 8.3), 15mM $MgCl_2$, 500mM Kcl.
SOB-Mg	20g Bacto-tryptone, 5g Bacto yeast extract, 0.5g NaCl, 0.19g KCl in 1l of dH_2O. Adjust pH to 7.0 with NaOH, autoclaved.
SOB	To 1l SOB-Mg add 10ml of: 1M $MgCl_2$, 1M $MgSO_4$.
SOC	To 1l SOB add 10ml 2M glucose (filter strilised).
SSC (x20)	3M NaCl, 0.3M sodium citrate pH adjusted to 7.0.
STET	8% sucrose (w/v), 0.5% Triton X-100, 50mM EDTA, 10mM Tris.Cl (pH8.0).

Sodium acetate, 3M	pH adjusted to 5.2 with glacial acetic acid, then autoclaved.
TfbI	30mM potassium acetate, 100mM RbCl, 10mM $CaCl_2.6H_2O$, 50mM $MnCl_2.H_2O$, 15% glycerol. pH to 5.8 with acetic acid. Store at 4°C.
TfbII	10mM MOPS, 75mM $CaCl_2.6H_2O$, 10mM RbCl, 15% glycerol. pH to 6.5 with KOH. Store at 4°C.
Tris.Cl, 1M	pH adjusted using HCl (to 7.6, 8.0 or 9.0), then autoclaved.
Tris-acetate (TAE) x10	40mM Tris-acetate, 1mM EDTA.
Tris-borate (TBE) x10	900mM Tris, 900mM boric acid, 20mM EDTA (pH8.2).
TE	10mM Tris.Cl (pH 8.0), 1mM EDTA (pH 8.0). Autoclaved.

2.2 Cloning and DNA manipulation

2.2.1 Production and transformation of competent cells

Due to the relative difficulty of transforming large DNA plasmids (more than 20kb) into bacterial cells, a number of different transformation techniques were used during the cloning steps. For small plasmids either the "same-day" protocol or the Tfb protocol were used. For larger plasmids either electrotransformation was performed, or Epicurian Coli XL2-Blue ultracompetent cells were used, bought from Statagene.

Same-day protocol

A fresh culture of DH5a F' strain *E.coli* cells grown from a single colony was sub-cultured into fresh L-broth and then grown until the OD_{550} of the culture was between 0.2 and 0.4. The cells were then centrifuged at 4°C for 10 min. at 4,000rpm. The supernatant was removed and the pellet

resuspended in 1/10th of the original culture volume in ice cold 50mM $CaCl_2$. The centrifugation and resuspension were repeated using only 1/25th original volume of $CaCl_2$. The cell suspension was then stored on ice until required.

Tfb protocol

A fresh culture of DH5a F' strain *E.coli* cells grown from a single colony was sub-cultured into fresh L-broth and then grown until the OD_{550} of the culture was between 0.45 and 0.5. The culture was chilled on ice for 5 min. and the cells were pelleted by centrifugation 2,500 rpm for 10 min. at 4°C. The supernatant was discarded and the pellet was resuspended in 200ml of TfbI, placed on ice for 5 min. before further centrifugation at 2,000 rpm for 10min. at 4°C. The cells were resuspended in 20ml of TfbI, stored on ice for 15 min. before separating into 200ml aliquots in 0.5ml microfuge tubes and snap-freezing in liqid nitrogen. The tubes were stored at -70°C.

Preparation of cells for electroporation

A single colony of DH5a F' strain *E.coli* cells was used to inoculate 10ml of SOB-Mg in a 50ml conical flask. This culture was incubated overnight at 37°C with shaking. 5ml of the culture was used to inoculate 500ml of pre-warmed SOB-Mg which was incubated at 37°C with shaking until the OD_{550} reached 0.75. The cells were centrifuged at 5,000rpm for 10 min. at 0°, the supernatant removed and the pellet resuspended into 10ml ice-cold 10% glycerol. When fully mixed a further 390ml of 10% glycerol was mixed in. The centrifugation and resuspension in 10% glycerol was repeated twice. The pellet was resuspended in 0.5ml 10% glycerol and the volume adjusted to give a OD_{550}. The cell suspension was then divided into 100ml aliquots and snap-frozen in liquid nitrogen.

Heat-shock transformation of competent cells

The cells were thawed on ice and either 50ml or 100ml dispensed into a prechilled 15ml Falcon 2059 polypropylene tube. For the XL2 cells, 1.7ml of b-mercaptomethanol was added to 100ml of the cells which were left on ice for 10 min. with occasional swirling. 1ml of the DNA

solution was added to each aliquot, and left on ice for 10-30 min. with occasional swirling. The NZY+broth was prewarmed to 42°C. The tubes were heat pulsed in the 42°C waterbath for just 30 sec., then back on ice for 2 min. and then the NZY+broth was added and the tubes incubated at 37°C for 1 hr. with gentle shaking. The cells were pelleted at 3,000rpm for 10 min. before being plated out in ~200ml remaining SOC onto selective agar.

Electrotransformation

The cells were thawed on ice, before 20ml was mixed with 1ml of each DNA solution and transferred to a pre-cooled electroporation chamber (with a 1mm gap between electode plates). The cells were shocked under the following conditions: 1.8KV, 25mF, 200W. 1ml of pre-warmed (37°C) SOC was immediately added to the cells in the chamber. The cells were then incubated in polypropylene tubes at 37°C for 1 hr. with gentle shaking (250rpm), before being pelleted at 3,000rpm for 10 min. and then plated out onto selective agar in ~200ml of remaining SOC.

2.2.2 Agarose gel electrophoresis

Restriction enzymes were used according to the manufacturers instructions in the presence of the appropriate buffer supplied. Restriction digests were analysed by electrophoresis through agarose gels which varied in concentration between 0.5% and 2% (w/w) of agarose powder in TAE buffer, depending on the size of the DNA fragments to be analysed. To visualise the DNA bands, ethidium bromide, which fluoresces strongly under ultra-violet light and intercalates with DNA, was added to the gels at a concentration of 0.5mg/ml. The same TAE solution was used as the running buffer during electrophoresis, and DNA samples were loaded after being mixed with "blue-juice" which contains glycerol (to ease loading of the sample into the well), and Bromophenol blue (to aid visualisation of the DNA sample during the loading). After electrophoresis the size and mass of DNA fragments was determined by comparison with DNA markers of known molecular weight or mass which were run alongside the samples. The marker used for fragment size determination was a 1kb ladder, and for quantitative analysis a *HindIII* digest of the lambda DNA was used, both obtained from BRL.

2.2.3 Purification of DNA from agarose gels

A small slice of gel containing the required DNA fragments was cut out over the UV-transilluminator using a blunt spatula. The slice was weighed and dissolved in 2.5 volumes (w/v) of 6M sodium iodide at 55°C. 5ml of Glassmilk (Geneclean kit for Stratech) was mixed into the solution and then left on ice for 10 min. The solution was centrifuged for 5 sec. at 14,000rpm in a benchtop microfuge, the supernatant removed, and the pellet gently resuspended in 200ml New Wash. Centrifugation and washing was performed three times, then the Glassmilk was resuspended in 20ml of distilled water and incubated for 20 min. at 55°C. After a final centrifugation for 1 min. at 14,000rpm the DNA solution was recovered as the supernatant.

2.2.4 Blunt-ending and ligation of DNA fragments

If fragments were to be ligated which had been digested with different restriction enzymes, the single-stranded overhangs were removed using 0.1U of Klenow per µl in a solution of 10mM Tris (pH 7.4), 5mM MgCl2, and 1mM of each of the four dNTPs, which was incubated at room temperature for 30 min. The vector and insert fragments were then mixed in a molar ratio of 1:3, and incubated in 1x ligation buffer and 1U of T4 DNA ligase in a total volume of 10µl. This was left for a few hours or over night at room temperature, and subsequently used to transform competent cells.

2.2.5 Mini-preparations of plasmid DNA

Single colonies of Amp-resistant bacteria were picked to inoculate 5ml cultures of L-broth containing 50 µl /ml of ampicillin, and were then incubated overnight at 37°C with shaking. 1.5ml of this was then removed to a standard Eppendorf tube, and centrifuged at 14,000rpm for 1 min. The supernatant was removed and the pellet resuspended in 200µl of STET, which contained 1mg/ml of lysosyme. The tubes were then placed in a boiling water bath for 50 sec. and centrifuged at 14,000rpm for 15min. The pellets were removed and discarded using toothpicks. 80µl of 5M ammonium acetate and 600µl of ethanol were added and the tubes centrifuged at 4dC for 15min. at 13,000rpm. The supernatant was then removed, the pellets rinsed in 70% ethanol, and then dried for

5 min. using a vacuum "Speedvac". The pellets were then dissolved in 50µl of TE. About 5µl was then used for each subsequent restriction digest.

2.2.6 Maxi-preparations of plasmid DNA

The following method was used to make large-scale preparations of highly purified plasmid DNA. 1ml of the culture grown for mini-preps was used to inoculate 500ml of L-broth containing ampicillin. This was incubated on a gyratory shaker (set to 250rpm) overnight at 37°C. The following day 1ml of culture was removed and added to 500µl of sterile glycerol in an Eppendorf tube. This mixture could then be stored indefinitely at -70°C as a glycerol stock. The remaining was split between two Sorvall 500ml centrifuge tubes and spun at 5,000rpm for 10 min. 50ml of alkaline SDS was then added and the mixture shaken vigorously, to lyse the cells, and then placed on ice for 5 min. 37.5ml of 3M potassium acetate (pH 4.8) was added and mixed and the tubes were placed on ice for a further 10 min. This liquid was centrifuged at 8,000rpm for 20 min. and the supernatant passed through 2 layers of sterile gauze to remove traces of the precipitate. The clean supernatant was mixed with 48ml of chilled isopropanol, incubated at -20dC for 30min. and the DNA pelleted by centrifugation at 8,000rpm for 10min. This pellet was then redissolved in 3ml of TE and the total volume measured. To this solution caesium chloride ($CsCl_2$) was added to a final concentration of 1.15g/ml, and the volume was remeasured. 80µl of 10mg/ml ethidium bromide were added per ml of the DNA solution, and this was made up to 5ml using a solution of $CsCl_2$ and ethidium bromide at the same concentrations. Any extra precipitation was then removed by centrifugation at 3,000rpm for 5min. and the remaining solution transferred to a ultra-centrifuge tube which was heat-sealed.

2.2.7 Manual DNA sequence determination

Manual DNA sequencing was performed using the Sequenase version 2.0 Sequencing Kit (from USB). This method is based on the dideoxynucleotide system developed by Sanger et. al., and involves two steps: labelling with radioactive nucleotides, and termination of the chain-extension reaction. The DNA solutions were denatured by adding 2ml of 2M NaOH to 18ml of the sample containing 3-5mg of DNA and leaving at room temperature for 5 min. The DNA was precipitated by adding 8ml of 5m sodium acetate (pH 7.5), mixing, then adding 100ml of 95%

ethanol and incubating at -20°C for 15 min. The DNA was pelleted by centrifugation at 12,000rpm for 20 min. The pellet was then washed in 70% ethanol, air dried, and dissolved in 7ml of dH_2O. The template was annealed to the primer by adding 2ml of the reaction buffer and 1ml of the primer at a concentration of 0.5pmol/ml. The mixture was incubated 65°C for 2 min. and was then allowed to cool slowly before being placed on ice. The labelling reaction was performed by adding 2ml of the labelling mix, 1ml of 0.1M DTT, 0.5ml [a-^{35}S]dATP and 2ml of diluted Sequenase version 2.0 T7 DNA polymerase to the annealed template and primer. This reaction was incubated for 2-5 min. at room temperature before transferring 3.5ml of the mix to each of four tubes containing 2.5ml of the A, C, G or T termination mixes, containing the appropriate dideoxynucleotide, which had been prewarmed to 37°C. These reactions were usually performed in a mini-microtitre plate. The termination reactions were incubated at 37°C for 5 min. before adding 5 ml of Stop Solution to each tube. The samples were then denatured for 2 min. at 90°C before electrophoresis through a polyacrylamide gel.

Glass sequencing plates were washed with detergent, rinsed well with tap water and finally with dH_2O. They were then rinsed with 70% ethanol and left to air-dry. The surface of one plate was treated with dimethyldichlorosilane solution (BDH #33164) to prevent the gel sticking tightly to both plates. Plastic spacers (0.4mm) positioned along the sides separated the two plates which were bound together with masking tape and bulldog clips. The polyacrylamide gel was prepared by mixing 60ml of 6% acrylamide/urea solution with 120ml of 10% ammonium persulphate solution to 120ml TEMED. The mixture was immediately drawn into a 50ml syringe and slowly poured between the sequencing plates in a continuous stream to avoid trapping air bubbles. The flat side of a shark's tooth comb was then inserted approximately 0.5cm into the solution and the gel allowed to polymerise for 1 hr. Electrophoresis was carried out at a constant voltage of 1800V. The gel was then fixed with 10% methanol/10% acetic acid for 30 min. before transferring onto a sheet of Whatman 3MM paper and drying under a vacuum, using a Bio-Rad slab gel drier, for 1 hour. The labelled DNA fragments were visualised by autoradiography using Kodak X-OMAT AR film.

2.3 Construct building

All constructs which were used in transgenic analysis, were created by recombination of clones which already existed in the lab. (plasmids and cosmids). For the analysis of *Hoxb-1,* construct b9-A had been created and tested by Jenny Whiting. Constructs b9-B, C and D, were created from

restriction fragments from cosmid cos54, which was cloned and mapped in the lab about five years ago. These fragments were ligated into the plasmid #1084, which contains the *Hoxb-4* promoter and the LacZ reporter to create constructs containing a 5.5kb HindIII fragment, a 2.9kb SalI-HindIII fragment, and a 1.9kb HindIII fragment, respectively. (Restriction sites can be seen in the Fig. 3.1 in chapter 3). They were each tested for orientation (so that they are in the correct 5'-3' direction with respect to transcription), and fragments for microinjection were removed using NotI.

The outline of how constructs were built for the analysis of *Hoxb-5* and *Hoxb-4* is described in section **4.3**, so only technical details of lesser importance are described here. Figure 4.5 includes all the relevant constructs. The PLAP gene was kindly provided by C. Cepko in the plasmid pDAP. From this it was removed as a 1.9kb SalI fragment. In the first construct, b4AP, the PLAP SalI fragment was ligated into the SalI site of plasmid #968. This plasmid has a pPolyIII vector with the SalI site destroyed, and it contains a 17kb stretch surrounding *Hoxb-4*, from the 5' ClaI site to an EcoRI site 1.7kb 3' of region A. b4ΔN was created by performing a partial digest of #968 with NcoI, selecting the largest single-cut band of the resulting fragments, filling-in the site with klenow, religating, and then screening for absence of the 3' site. This was then transferred to the pGP1f vector (kindly provided by GenPharm), by opening both plasmids with NotI, and ligating the 17kb *Hoxb-4* fragment into pGP1f. The digestion of pGP1f by NotI caused the removal of its own polylinker, so the polylinker from pPolyIII was now in its place. The regions A, B and C were removed from GPb4 to create b4Δ1, by digesting with SalI and NcoI, filling-in the overhangs, and religating. This ligation of the blunt-ended SalI and NcoI sites recreated a SalI site in the same reading-frame as before (with respect to the *Hoxb*-4 coding sequence), which was necessary for the subsequent insertion of APpA3 (described below). b4Δ1 was then modified to remove the 3' XhoI site (using the same partial digest strategy described above for NcoI). To create a SalI fragment which contained PLAP and the polyadenylation (pA) signal from SV40, the original SalI fragment from pDAP was ligated into the SalI site of pGEMT, which contains the pA sequence. To include the pA within a SalI fragment, first the 3' SalI site was destroyed (as described above for NcoI), and then a SalI linker was ligated into the SwaI site. APpA3 was then ligated into the recreated SalI site of b4Δ2 to create [ED]b4.

To add fragments including *Hoxb-5*, a modification of an existing construct was necessary. Construct #1029, made by Stefan Nonchev in the lab, contains a stretch around *Hoxb-5* with the LacZ gene inserted into the BamHI site within the first exon. The construct extends from a BglII site 5' of the gene, to the KpnI site in between regions E and D, and therefore includes region E. To

join this to the *Hoxb-4* constructs, region E had to be removed, and this was achieved by digesting #1029 into three ClaI fragments and then ligating the 5.1kb and 8.5 kb fragments together (which represent the vector, and the 3' region of *Hoxb-5* respectively). The resulting construct, b5(BC), was in the vector pSal, which has two SalI sites flanking its polylinker. The *Hoxb-5* section of DNA could therefore be removed using SalI and ligated into the unique XhoI site in the 5' polylinker of either [ED]b4 or b4Δ2 (only the latter of which succeeded). Deletions of DG[ED] were produced by digesting with KpnI and religating (as described in section **4.3**).

The construct b1-A was generated from one made by Alex Gould in the lab (2.9RVΔATG), in which a 108bp deletion was made which removed the start codon of the *Hoxb-1* gene, and a EagI site had been inserted. b1-A was made by blunt-ending the 1.9kb SalI fragment which contains PLAP without the pA signal, and ligating it into the EagI site of 2.9RVΔATG.

2.4 Production of transgenic mice

Throughout these experiments (CBA x C57Bl10)F_1 mice were used as embryo donors, stud males, pseudopregnant females, vasectomised males and mature females for breeding. All techniques performed on animals were licensed under the Animals (Scientific Procedures) Act 1986, license no. PIL 80/00831.

2.4.1 Transgenic media

The most commonly used embryo culture media is M16, which is very similar to Whitten's medium and is bicarbonate buffered. However, fertilised eggs do nor readily continue development beyond the late two-cell stage *in vitro*. To overcome this, T6 media was used and the components are given in Table 1. This medium was incubated at 37°C in 5% CO_2, 95% air during use. For collecting embryos, and for experiments in which the embryos are handled for prolonged periods outside the incubator (e.g. microinjection), HEPES buffer is added in place of the bicarbonate in order to maintain the correct pH. This medium with HEPES supplement is named M2 and the components are shown in Table 2. Stock E was adjusted to pH 7.4 with NaOH. The stock solutions for both M2 and T6 could be stored for 3 months at -20°C. To make up 50ml of M2 the following were mixed: 5ml stock A, 0.8ml stock B, 0.5ml stock C, 0.5ml stock D, 4.2ml stock E, 39ml sterile distilled water and 200mg bovine serum albumin (BSA). To make up 50ml of T6 the

TABLE 1: Preparation of T6 Media

STOCK A (10 x conc.)	**COMPONENT**	**g/100ml**
	NaCl	4.721
	KCl	0.110
	$MgCl_2.6H_2O$	0.100
	$NaH_2PO_4.2H_2O$	0.061
	Sodium lactate (60%)	3.4 ml
	Glucose	1.000
	Penicillin G	0.060
	Streptomycin sulphate	0.050
STOCK B (10 x conc)	**COMPONENT**	**g/100ml**
	$NaHCO_3$	2.100
	Phenol Red	0.010
STOCK C (100 x conc)	**COMPONENT**	**g/10ml**
	Sodium pyruvate	0.029
STOCK D (100 x conc)	**COMPONENT**	**g/10ml**
	$CaCl_2.2H_2O$	0.260

TABLE 2: Preparation of M2 Media

STOCK A (10 x conc.)	**COMPONENT**	g/100ml
	NaCl	5.534
	KCl	0.356
	KH_2PO_4	0.162
	$MgSO_4.7H_2O$	0.293
	Sodium lactate(60% syrup)	3.4 ml
	Glucose	1.000
	Penicillin G	0.060
	Streptomycin sulphate	0.050
STOCK B (10 x conc.)	**COMPONENT**	**g/100ml**
	$NaHCO_3$	2.101
	Phenol Red	0.010
STOCK C (100 x conc.)	**COMPONENT**	**g/10ml**
	Sodium pyruvate	0.036
STOCK D (100 x conc.)	**COMPONENT**	**g/10ml**
	$CaCl_2.2H_2O$	0.252
STOCK E (10 x conc.)	**COMPONENT**	**g/100ml**
	HEPES	5.958
	Phenol Red	0.010

following were mixed: 5ml stock A, 5ml stock B, 0.5ml stock C, 0.5ml stock D, 39ml sterile distilled water and 200mg BSA. The solutions were then filter-sterilised before aliquoting into sterile containers. Once prepared, M2 and T6 are stable at 4°C for up to two weeks.

After about two years of using these media, it was decided to try a commercial product instead, as the quality of T6 and M2 was found to vary considerably from month to month. The substitute for T6 used was Whitten's medium, and for M2 the media names KSOM was used. Although the quality of these two products was found never to be as high as the best T6 or M2, it was very consistent and proved to be satisfactory.

2.4.2 Preparation of DNA for microinjection

Linear DNA fragments were prepared using Glassmilk (see section **2.2.3**) and redissolving the DNA in injection buffer instead of dH_2O. The concentration was determined by running an aliquot of the solution through an agarose gel next to the *HindIII* DNA marker. The concentration was then adjusted to 1ng/ml, and the final solution was cleaned by centrifugation through a Spin-X column (Costar #8162).

2.4.3 Preparation of egg donors by superovulation

Naturally ovulating females will typically produce 6-10 eggs, so to reduce the number of females needed for each experiment the females used for this purpose were superovulated. This results in 20-30 eggs being produced from each 4-week old animal. The females, which are adjusted to a light period of 5am-7pm were given an intraperitoneal injection with 5IU of pregnant mare's serum (PMS) at about 3pm, 3 days before the eggs are to be recovered. A second injection of 5IU human chorionic gonadotrophin (hCG) was given at about 1pm, the day before the experiment. Both hormones were obtained from Intervet Laboratories, as Folligon and Chorulon respectively. Following hCG injection, each female was placed in a cage with a stud F_1 male. The males used were between 2 and 5 months old. The number of viable, fertilised eggs obtained from a mating with a male older than 5 months was significantly reduced, so the stud males were replaced every 3 months. Copulation plugs were checked the following morning and these females removed for oviduct dissection. The oviducts were dissected into M2 medium and the eggs released by opening them with forceps. Cumulus cells were removed from the zygotes by adding hyaluronidase

to the M2 to a final concentration of 300ml/ml, and leaving the cells for 5 min. The eggs were then rinsed three times in M2, before being transferred to T6 and stored in the 37°C incubator (with 5% CO_2). The drop of T6 was prevented from evaporating by a covering of parafin oil.

2.4.4 Pseudopregnant recipients

Pseudopregnant female mice, between 6 and 8 weeks of age, were prepared by mating females in natural oestrus with vasectomised males. Vasectomised males were prepared in the SPF facilty of the NIMR, and pseudoprenant females were ordered from the unit as required. The females were anaesthetised at 0.5dpc by an intraperitoneal injection of 0.35ml 2.5% avertin, or 0.35ml Hyp/Hyp. Microinjected embryos, at the one- or two-cell stage, were then transferred through the infundibulum into the oviducts, using a heat-polished pulled-capillary needle. Between 10 and 15 embryos were transferred into each oviduct.

2.5 Analysis of transgenic mice

2.5.1 Transgenic detection by PCR

Tail biopsies were taken from mice at 3 weeks of age, or yolk sac tissue was retained following dissection of embryos. The tissue was placed in 100ml of tail lysis buffer and incubated overnight at 55°C. After phenol/chloroform extraction (to remove protein) 1µl of the DNA solution was added to: 2µl PCR salt solution, 2µl 10x dNTP solution, 0.1µl *Taq* polymerase (at 5U/µl) and 13µl of dH_2O. The samples were denatured for 3 min. at 94°C, followed by 30 cycles of denaturation at 94°C (20 sec.), annealing at 57°C (1 min.), and primer extension at 72°C (2 min.). A final extension cycle was performed at 72°C for 3 min. The resulting PCR products were loaded onto a 2% TAE agarose gel for electrophoresis (30 min. at 100 volts). Two control primers from the myogenin gene were used as a positive control for the PCR reaction in each case (MGP1 and MGP2). Primers from the *Hoxb-4* promoter and the PLAP gene were used for all the constructs involving the double reporter strategy described in chapter 4. The *Hoxb-4* primer was also used in conjunction with a LacZ primer to test for all the constructs described in chapter 3 (testing regions of *Hoxb-9*). The construct which created the *sasquatch* mutant, was made from the 5' region of *Hoxb-1*, so a primer from the *Hoxb-1* promoter was used in conjunction with the PLAP primer. Two primers were

made to the PLAP gene, which create PCR fragments of 400bp and 300bp when used in conjunction with the *Hoxb-4* primer. The second one allows tripple PCRs to be performed, so that when a LacZ transgenic line is bred into a PLAP transgenic line, the presence of both constructs can be tested, along with the myogenin control.

MGP1 - CCAAGTTGGTGTCAAAAGCC

MGP2 - CTCTCTGCTTTAAGGAGTCAG

Hoxb-1 - AGCTTCAGCTCTGTGACATACTGCCG

Hoxb-4 - GGAAAACCGAGTCAGGGGTC

LacZ - TAGATGGGCGCATCGTAACCGTGCAT

PLAP1 - AGCTGTCACCGTAGACACC

PLAP2 - TCCAGAAGTCCGGGTTCTCC

2.5.2 Assay for b-Galactosidase

Embryos were dissected from pregnant females and washed in PBS and stained for b-galactosidase activity as follows: 10-20 min. in 4% paraformaldehyde at 4°C, followed by 5 washes in PBS for 30 min. each. Then the embryos were incubated in staining solution in the dark at room temperature. The strength of expression varied greatly between different constructs, and the length of incubation time therefore ranged from a few hours to a couple of days. If levels were very low, incubation was performed at 37°C.

2.5.3 Assay for transgenic alkaline phosphatase

Embryos being processed for the presence of PLAP, had usually already been stained for b-gal activity (as the high temperature required for the PLAP staining procedure destroys the b-gal activity). However, it was very important that the X-gal staining solution was thoroughly washed out of the embryos, using 5 incubations in PBS on a shaker. The embryos were then heated to 65°C for 30 min. to heat-inactivate the endogenous alkaline phosphatases, allowed to cool for 20min. before being transferred to staining buffer containing 240µg/ml of levamisol, and incubated for 1 hour at room temperature. The BCIP and NBT stock solutions were then added (to a concentration of 100µg/ml and 1mg/ml respectively) and the embryos further incubated in the dark, at room

temperature. The colour reaction was usually complete within 1-2 hours, at which point it was stopped by adding a solution of 50mM EDTA at a pH of 5.0, and then refixing the embryos in 4% paraformaldehyde.

2.5.4 Histological studies

Embryos to be sectioned by cryostat, were equilibrated in 30% sucrose overnight at 4°C. They were then placed in thin plastic moulds and embedded in OCT compound, which was frozen by placing on 'dry ice'. Sections were cut in the cryostat at about -26°C, generally with a thickness of 8 or 10μg. Staining for PLAP was then performed in a narrow slide-container.

Hindbrain flatmount preparations were prepared by isolating the hindbrain, disecting away the branchial arches and surrounding mesoderm tissue, openeing along the roof plate and flattening the tissue between a slide and a coverslip in 100% glycerol.

2.5.5 Skeletal analysis

The skin and internal organs were removed from young adult mice, and the remaining soft tissues dissolved in 2% potassium hydroxide for two days. The skeletons were then stained for 24 hours in 1% potassium hydroxide and 75μg/ml alizarin red S. When a strong colouration had been achieved, the skeletons were destained for a week in 20% glycerol, 1% potassium hydroxide, changing the solution daily. Remaining loose tissue (fat and tendons) were removed, then the skeletons transferred to 20% glycerol, 20% ethanol overnight, and 50% glycerol, 50% ethanol the following day for further clearing, indefinite storage and photography.

2.5.6 Whole-mount *in-situ* hybridisation

Synthesis of probe

The following synthesis reaction was incubated overnight at 37°C: 10μl of dH_2O, 4μl of 5x transcription buffer, 2μl of 0.1M DTT, 2μl of 10x DIG nucleotide mix, 1.5μl of the linearised DNA template (at 1μg/μl), 0.5μl of Rnasin ribonuclease Inhibitor, and 1.5μl of SP6, T7 or T3 RNA polymerase (depending on the promoter used). After synthesis, 1μl was removed and run on

an agarose gel to estimate the amount of RNA. 2µl of DNAse was added and the reaction incubated at 37°C for 15min. The RNA was then precipitated by adding 50µl of dH_2O, 1µl of glycerol, 25µl of 10M ammonium acetate and 200µl of ethanol, mixing well, leaving on dry ice for 30min. and then centrifuging at 13,000rpm for 20min. at 4°C. The pellet was then washed in 70% and dried in a 'SpeedVac' vacuum drier for 3min. The pellet ws then re-dissolved in 5µl DEPC water and store for use at -20°C.

Treatment of embryos

The embryos were dissected in PBS, fixed overnight at 4°C in 4% paraformaldehyde and rinsed twice in PBT at 4°C for 5min. They were then taken through a graded methanol series: 10min. at each of 25%, 50%, 75%, 100%, in PBT, and then back down through the same series. Protein was destroyed by treating embryos with 10µg/ml of proteinase K in PBT, for 5-10 min. at room temperature. The embryos were then rinsed well in PBT, refixed for 30min. in 4% paraformaldehyde at 4°C, rinsed again in PBT, and then placed in the prehybridisation mix overnight with gentle rocking at 62°C.

The hybridisation process was performed under the same conditions as prehybridisation, with the addition of 2µl of probe to the mix. It was allowed to proceed for 2-3 days.

After the hybridisation the embryos are taken through a series of washes: twice in 2xSSC + 0.1% Chaps for 40min. at 62°C, once in 0.2xSSC + 0.1% Chaps for 30min. also at 62°C, once in KTBT at room temperature, and once in 20% lamb serum in KTBT for 4 hours at 4°C with gentle rocking. The embryos are then incubated in a 1:1000 dilution of DIG-AP-labelled antibody in the same mixture, for 2-3 days at 4°C with gentle rocking. The embryos were then washed five times in KTBT for 30min. at room temperature and the final wash left overinght at 4°C. To stain the embryos they were rinsed in "alkaline phosphatase buffer" for 5 min. at room temperature, and then in 10ml of the buffer containing 25µl of the BCIP solution and 35µl of the NBT solution. This reaction is light-sensitive, and was therefore carried out in the dark, until a suitable strength of staining had been achieved (generally 1-2 hours). To stop the reaction the embryos were rinsed well in KTBT, twice for 30min., and then refixed in 4% paraformaldehyde for 2 hours.

CHAPTER 3

Regulation of the *Hoxb-9* gene

One of the most productive approaches in the attempt to understand *Hox* gene regulation, is using transgenic analysis to map local enhancer elements. DNA fragments from the vicinity of the gene are connected to a promoter which directs transcription of a reporter gene. A reporter gene, is one encoding a protein whose distribution in the embryo can easily be visualised by a histochemical staining reaction. In this part of the study, the bacterial reporter gene LacZ is used, which encodes the protein β-galactosidase. This enzyme catalyses the cleavage of galactose rings from many different compounds, and the most common substrate used for these experiments is X-gal (5-Bromo-4-chloro-3-indolyl-b-D-galactopyranoside), which changes from colourless to blue. In this way, the blue staining pattern shows the spatial pattern of activity encoded by any enhancers in the tested DNA segment.

When a transgenic construct is injected into a mouse zygotic pronucleus, it can integrate almost anywhere, so it may come under the control of other regulatory sequences. For this reason, a number of different insertion sites must always be analysed. Any aspect of the patterns which is common to all embryos can then be attributed to the transgenic construct itself, and not the "position effects" of random integration sites.

Before leaving the lab, Jenny Whiting initiated a transgenic study of *Hoxb-9*. She created a line using construct b9-A (from Fig. 3.1). This is a 12kb EcoRI fragment which extends 5' and 3' of the gene itself, into which the LacZ gene was inserted (at the SalI restriction site). In this construct the normal *Hoxb-9* promoter was used to drive expression of the LacZ gene. The large size of the construct was chosen to maximise the chance that all important regulatory sequences would be included. A time-course of embryonic development was made (Fig. 3.1) to examine expression patterns of the LacZ, and it was found that the majority of the normal *Hoxb-9* pattern was created. The major deficiency, was a lack of correct neural tube expression in the later stages of development.

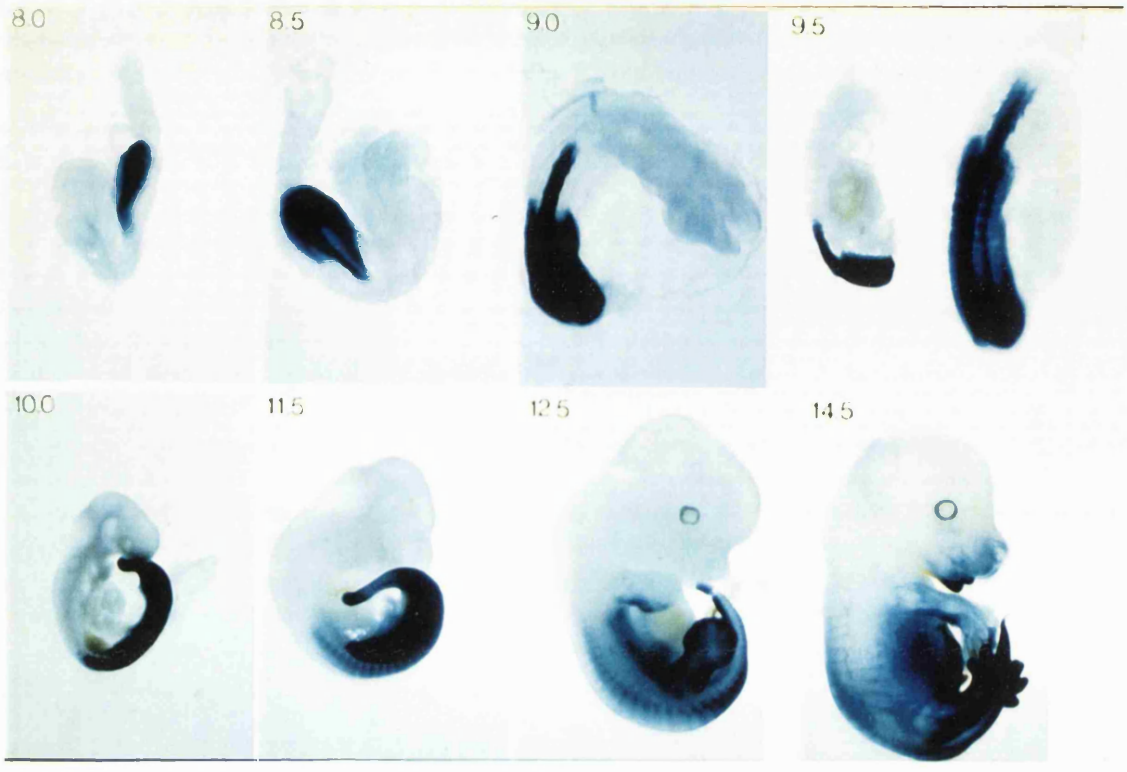
The aim of this part of the study, was to perform similar experiments to determine whether the important enhancer activity from the 12kb region could be mapped to a smaller section. In addition, I attempted to see if regulatory sequences responsible for the late neural expression could

Fig. 3.1 Transgenic constructs for studying *Hoxb-9*.

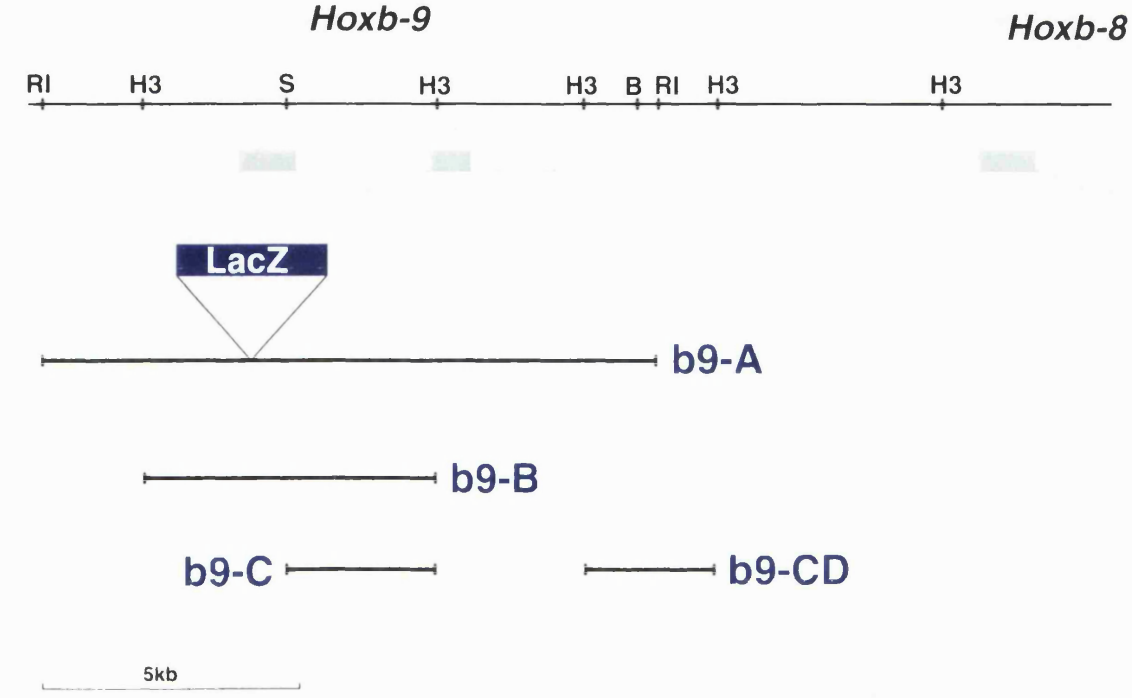
(a) A time-course was made (by J. Whiting) of the original 12kb construct which surrounds the *Hoxb-9* gene. From 8.0dpc to 9.5dpc, the transgenic embryos express LacZ in the same pattern as the endogenous gene. After 10.0dpc expression in the limbs and mesoderm continues, whereas neural expression fades away. Already by 11.5dpc it is almost completely absent, and this remains the case at 12.5dpc and 14.5dpc.

(b) The four constructs described in this chapter are labelled A to D. b9-A is the original 12kb frament tested by J. Whiting. b9-B and b9-C are deletions of this which concentrate on the intron, and b9-D includes the region just 3' to construct b9-A.

A


8.0
8.5
9.0
9.5
10.0
11.5
12.5
14.5

B


Hoxb-9
Hoxb-8
RI
H3
S
H3
H3
B
RI
H3
H3
LacZ
b9-A
b9-B
b9-C
b9-CD
5kb

be found outside this tested region. All images of *Hox* protein expression patterns (revealed by antibody staining) were provided by Alex Gould.

3.1 A 5.5kb HindIII fragment can recreate most of the pattern from the 12kb construct

The first subsection tested was the 5.5kb HindIII fragment from within the 12kb region already tested (Fig. 3.1). Instead of inserting the LacZ gene into the SalI site, which could disrupt the spacing of elements on either side of the first exon, the fragment was joined to a standard reporter construct which contains the promoter from the *Hoxb-4* gene attached to the LacZ gene. The resulting construct was named b9-B.

The construct was injected into mouse zygotes and 7 embryos were created which expressed the LacZ gene. All 7 displayed the same expression pattern at different intensities. The embryos were harvested at 10.5dpc, which is the most critical stage for comparing with the previous construct. The pattern was very similar, with the exception that the staining in the posterior region of the limb bud was stronger. Expression in the dorsal root ganglia (DGRs) was up to a similar A-P boundary to that of the wildtype gene (Fig. 3.2a,b), but as with the construct b9-A, expression in the neural tube faded out too posteriorly. This is clearly seen by the fact that neural expression of *Hoxb-9* protein (Fig. 3.2f) extends more anteriorly than DRG expression, whereas in the transgenic embryos it is DRG expression which extends more anteriorly.

3.2 The intron can also recreate most of the pattern

Due to the convenient SalI restriction site, the fragment tested in construct b9-B could easily be subdivided into two halves. The 3' half contains the intron of *Hoxb-9*, and since a number of other *Hoxb* genes have been found to contain mesodermal enhancers within the intron, this half was tested first (construct b9-C). Out of four different integration sites (3 transient embryos and 1 line) three displayed a consistent pattern of LacZ expression (the fourth showed ectopic expression in the head).

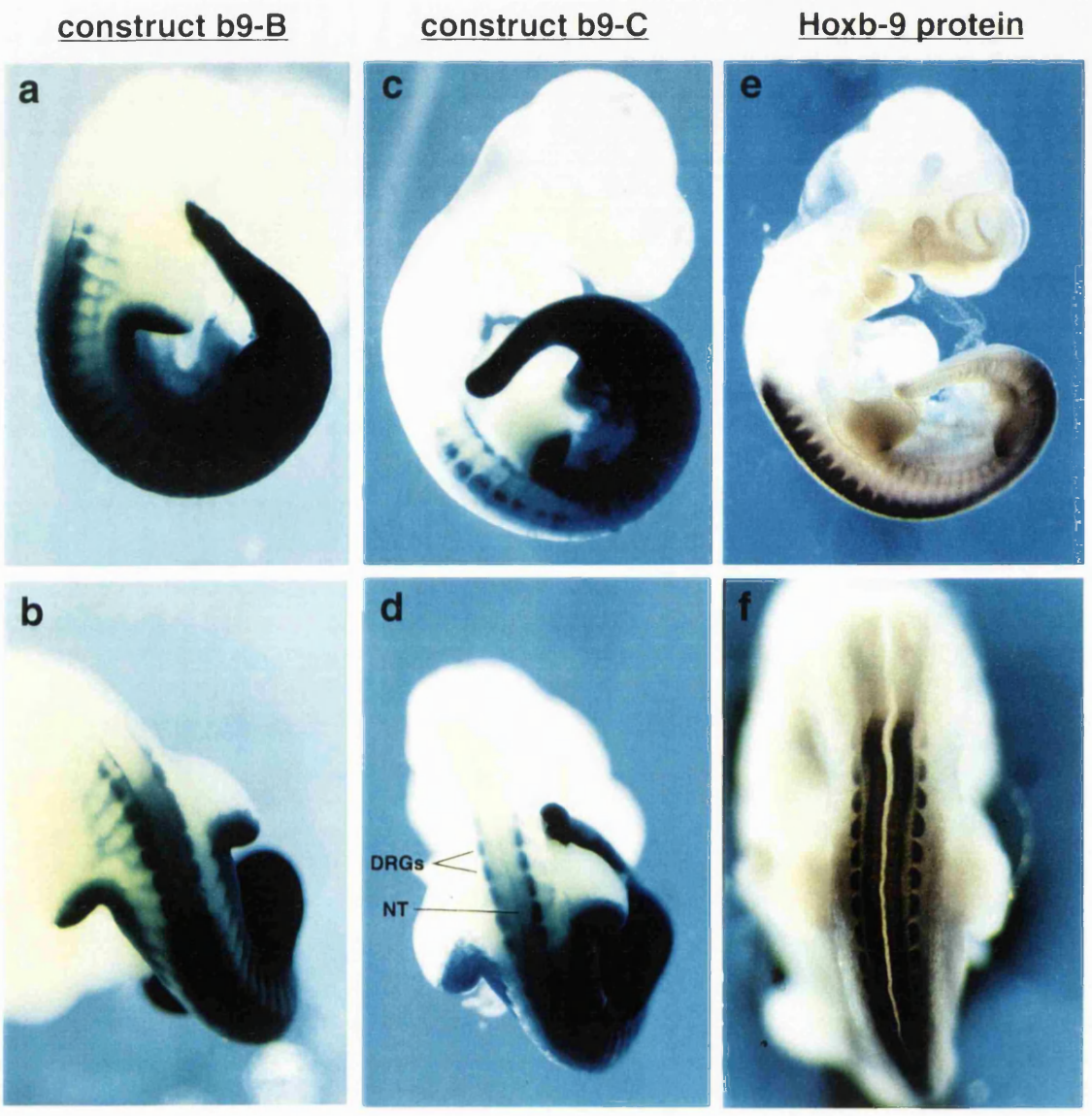
A comparison of b9-C with b9-B is shown in Fig. 3.2. Again, very little of the pattern was lost by removing the 5' half of b9-B, and the basic pattern of b9-A remained. The only difference seen, was a slight further posteriorisation of the neural expression - the DRG expression again extends further anteriorly than the neural expression (Fig. 3.2d).

Fig. 3.2 Expression patterns at 10.5dpc of the intron-containing constructs.

(a,b) Lateral and dorsal views of expression pattern of construct b9-B at 10.5dpc. Expression can clearly be seen in the hindlimb-bud, the posterior part of the forelimb-bud, neural and mesodermal tissue of the tail region, somites posterior to the forelimb-bud, the dorsal root ganglia (DRGs) and the neural tube. Note that expression in the DRGs extends more anteriorly than expression in the neural tube (NT).

(c,d) Lateral and dorsal views of the expression pattern of construct b9-C at 10.5dpc. The expression pattern is almost identical to that of construct b9-B, except that the neural expression appears even more posterior compared to the DRGs.

(e,f) Lateral and dorsal views of the *Hoxb-9* protein pattern as revealed by antibody staining (by Alex Gould).


construct b9-B
construct b9-C
Hoxb-9 protein
a
b
c
d
e
f
DRGs
NT

Since the basic mesodermal and DRG expression pattern had now been localised to the intron, a time-course of this line was made (Fig. 3.3). When compared to the pattern of *Hoxb-9* protein as visualised by antibody staining, the pattern of the transgene appears to start correctly (at the 9.5dpc stage) but gradually diverge over the next 3 days. This divergence is seen as an over-expression in the posterior third of the forelimb-bud, and a loss of neural expression. The limb bud discrepancy is very similar in all three constructs, but appears to occur slightly earlier in the shorter constructs (being visible just before 10.5dpc in b9-C, but only by late 10.5dpc in b9-A).

3.3 A neural enhancer lies just 3' of the 12kb EcoRI fragment

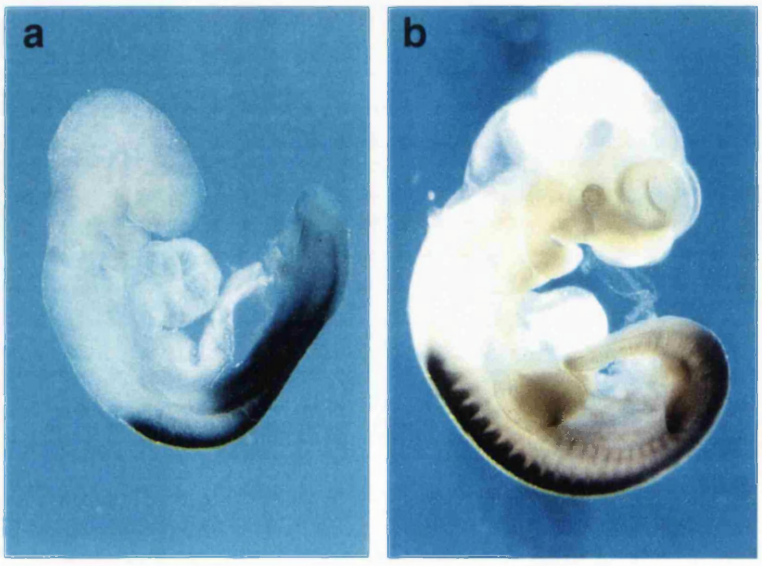
As a first extension of the 12kb region to search for the *Hoxb-9* neural enhancer, the HindIII fragment which overlaps with its 3' end was used to create construct b9-D (Fig. 3.1). Eight LacZ-expressing embryos were created, all of which displayed a similar pattern of expression at different intensities. In embryos of 10.5dpc and 11.5dpc, LacZ expression was strong in the neural tube - a clear distinction from constructs A, B and C (Fig. 3.4). At 10.5dpc the expression in the DRGs appeared to be similar to the previous constructs, but neural expression was now extending up to the same boundary (Fig. 3.4f). At 11.5dpc, neural expression in b9-C had almost completely disappeared (Fig. 3.4c and d), whereas in b9-D it was still maintained up to a very sharply defined boundary (Fig. 3.4g and h).

Despite the discovery of an obvious neural element with a very sharp boundary of expression, there is a discrepancy between it and the pattern of endogenous *Hoxb-9* protein revealed by *Hoxb-9* antibody staining (Fig. 3.4i and j). At 10.5dpc the protein is expressed up to a slightly more anterior boundary which is 2-3 DRGs beyond the anterior margin of the forelimb. The anterior boundary of the reporter gene is approximately level with anterior margin of the forelimb. This difference could be because extra regulatory information is normally provided to the endogenous gene from the Hox complex, while the local neural enhancer tested in various positions within the genome only gives approximate positional information. It does however, give a very sharp boundary of expression.

Fig. 3.3 Time-course of the expression pattern of construct b9-C.

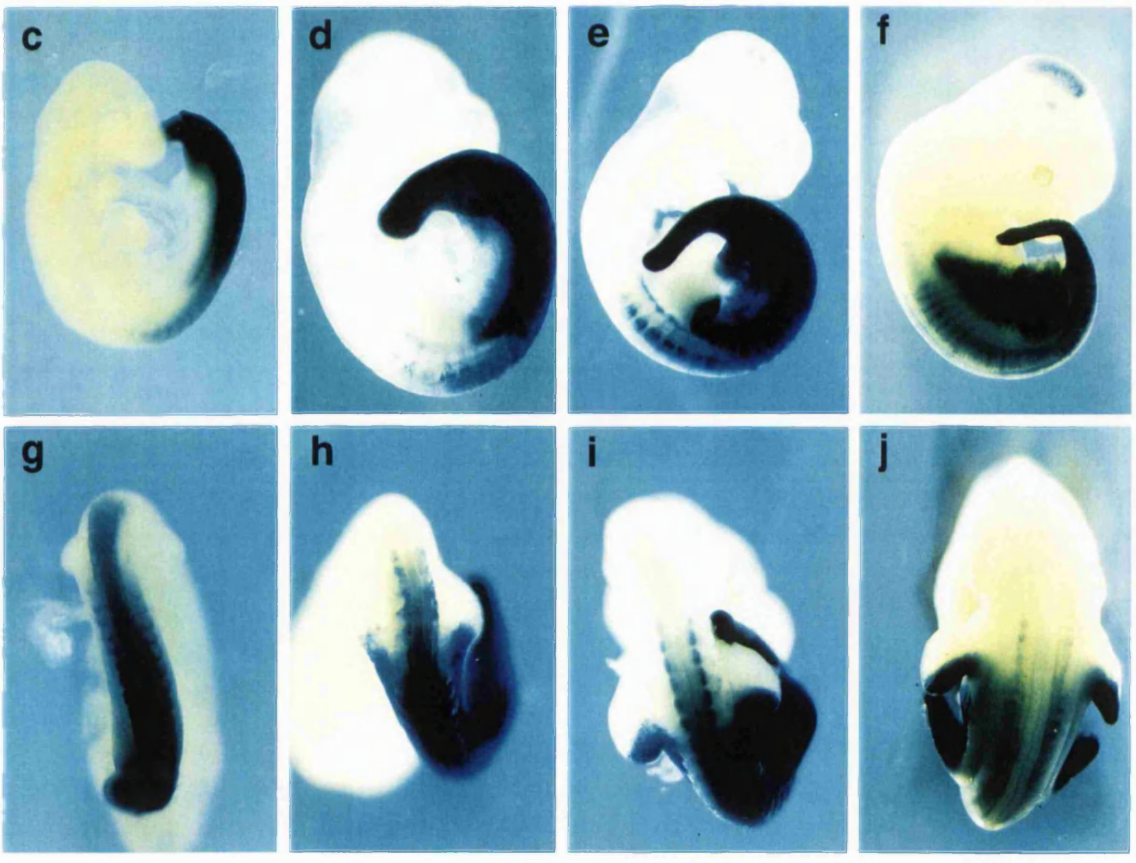
Panels (a,b) show lateral views of the *Hoxb-9* protein pattern at 9.5dpc and 10.5dpc (by Alex Gould). Panels (c-f) show the lateral views of embryos expressing the construct b9-C, at stages 9.5dpc, 10.0dpc, 10.5dpc and 12.5dpc respectively (and (g-j) show the dorsal views of the same embryos). At 9.5dpc this expression pattern is almost identical to the endogenous protein, but by 10.0dpc it is already very different - the neural expression having regressed to the level of the forelimb-bud. By 12.5dpc neural expression is only seen in the tail region. Also noticeable is the strong staining in the forelimb-bud, which is not reflected in antibody staining for the protein.

Hoxb-9 protein



9.5 10.5

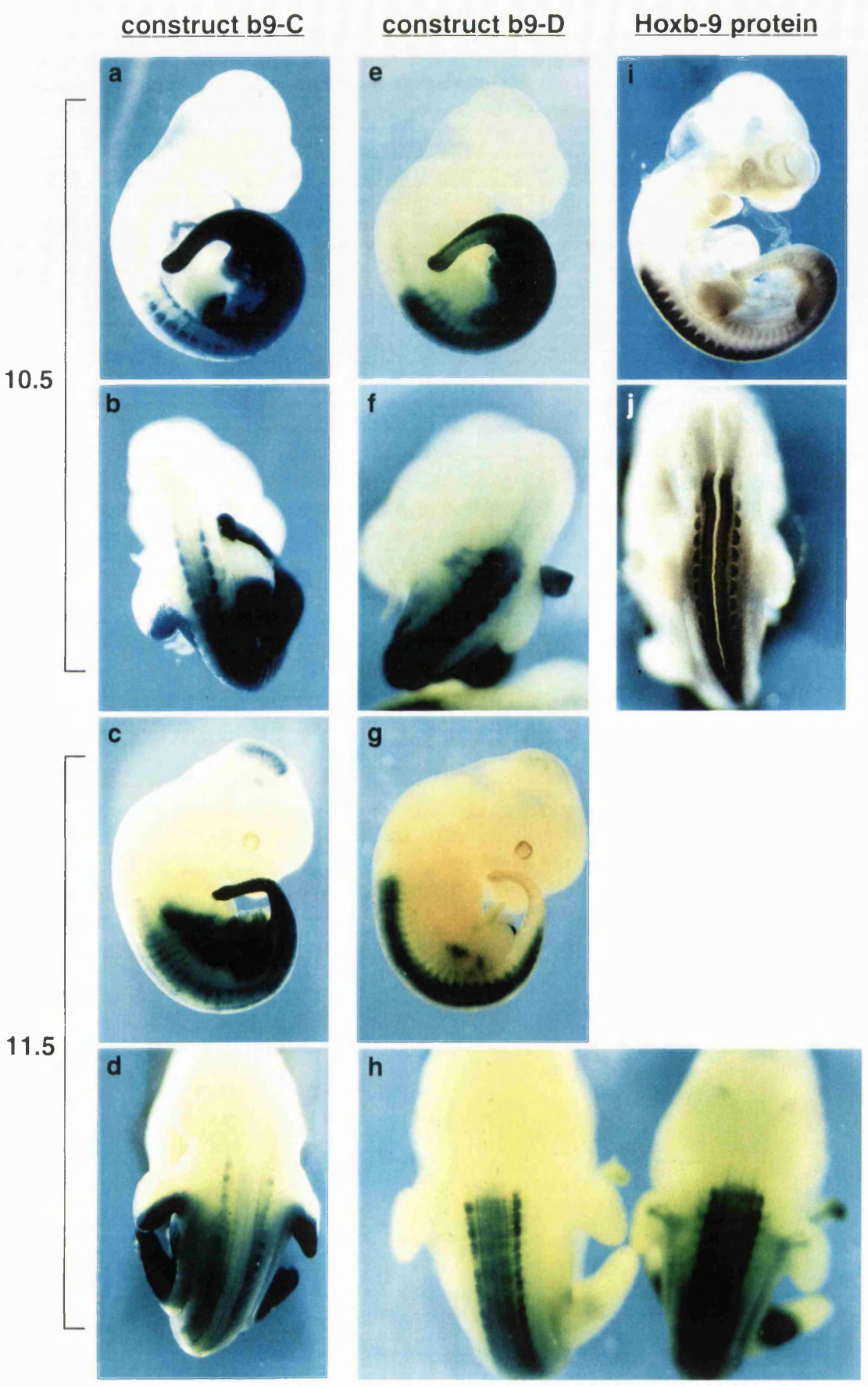
construct b9-C



9.5 10.0 10.5 12.5

Fig. 3.4 Comparison of the neural element with the intron.

Panels (a-d) show expression patterns for the intron construct (b9-C), and panels (e-h) show expression patterns for the neural element (b9-D). Whereas neural expression from the intron enhancer has regressed more posteriorly than the DGRs by 10.5dpc (a,b), in construct b9-D it is still strong, with a clear boundary near the anterior margin of the forelimb-bud (e,f). However the protein pattern at the same stage appears to have a more anterior boundary (i,j). At 11.5dpc comparison of the two constructs (c,d,g,h) shows almost complimentary patterns, in which b9-C gives limb, tail, lateral and somitic mesoderm expression, and b9-D gives clear neural expression up to a very sharp boundary.


construct b9-C
construct b9-D
Hoxb-9 protein
a
e
i
10.5
b
f
j
c
g
11.5
d
h

3.4 Expression of *Hoxb-9* in the forelimb

At first glance, the forelimb pattern seen from construct D appeared more similar to the *Hoxb-9* protein pattern than constructs A to C. In both the antibody staining and construct D there appears to be virtually no staining. In contrast, constructs A, B and C, all gave expression in the posterior third of the forelimb. However, on more detailed examination, the protein pattern does contain a very small patch of expression in the posterior edge of the forelimb (Fig. 3.4i), and it may be the case that the transgenic expression seen in A to C, is a more sensitive reflection of this normal expression domain. Considereing that this same feature was seen in all three constructs which contain sequences very close to the *Hoxb-9* promoter, it is very unlikely that this strong, regular domain reflects an enhancer for a more distant gene. Furthermore, this pattern is also seen with the chicken *Hoxb-9* gene, indicating that it is conserved (Pers. comm. L. McNaughton).

3.5 Summary

It has been shown that the majority of what is thought to be the normal expression pattern for the *Hoxb-9* gene, can be controlled from an enhancer (or cluster of enhancers) which lies within the 2.9kb intron. This basic pattern appears to be complete for the mesodermal tissues, and in this respect is similar to Hoxb-4 which also contains a mesodermal enhancer within its intron. However, the late neural expression is not directed from this sequence, and it has been shown that the neural enhancer closest to the *Hoxb-9* gene lies about 6kb 3' from the promoter. Although in transgenic constructs the anterior boundary of expression driven from this enhancer is not in exactly the correct position along the A-P axis, it is close enough to suggest that this is the important neural element for *Hoxb-9*. There are no neural elements closer, or more similar in pattern to the *Hoxb-9* gene, and the extra information required for its normal positioning may be derived from a global effect of the intact Hox complex. This 3' positioning of a neural enhancer is also reflected in the Hoxb-4 and probably Hoxb-5 genes (see next chapter) which also possess a 3' neural enhancer. It may therefore be an indication of the original gene duplication events which must have created the first Hox complex hundred's of millions of years ago. This is discussed in more detail in the discussion.

CHAPTER 4

Regulatory interactions between adjacent Hox genes

The clustered organisation of the Hox genes, has led to two main questions: (1) Why do the genes need to be close to each other? (2) Does the proximity of promoters and enhancers from adjacent genes cause regulatory interactions to occur between the genes? In this part of the study I attempted experiments which relate to both of these questions.

Whiting et al. (Whiting *et al.*, 1991), discovered that the main local regulatory elements for the *Hoxb-4* gene, were located in three regions labelled A, B and C (Fig. 4.1). In transgenic experiments it was found that constructs containing region A, plus the *Hoxb-4* promoter and the LacZ gene, caused expression of β-galactosidase in the neural tube, from the posterior tip up to the boundary between rhombomeres 6 and 7 (r6/7). This is the same anterior boundary as the endogenous *Hoxb-4* gene, and demonstrates that within region A is a spatially-specific neural enhancer for this gene. Similarly, in region C an enhancer was found which caused expression up to the correct anterior boundary in the somitic mesoderm (as well as some non-*Hoxb-4*-specific neural expression). Expecting a similar set of enhancers to exist for the regulation of *Hoxb-5*, Stefan Nonchev in the lab, performed a series of transgenic experiments, in which different regions around the *Hoxb-5* gene were tested. He defined two further regions, labelled D and E (Fig. 4.1), which appeared to be responsible for mesoderm and neural expression respectively. These regions are in the intergenic DNA between *Hoxb-5* and *Hoxb-4*, and in further transgenic experiments it was shown that they could activate the *Hoxb-4* promoter at least as well as the *Hoxb-5* one. Thus arose the questions: Since these enhancers are near to both the *Hoxb-4* and *Hoxb-5* promoters, which one do they normally control? Could they be important to both? If they should only act on one promoter, how is interaction with the other one prevented?

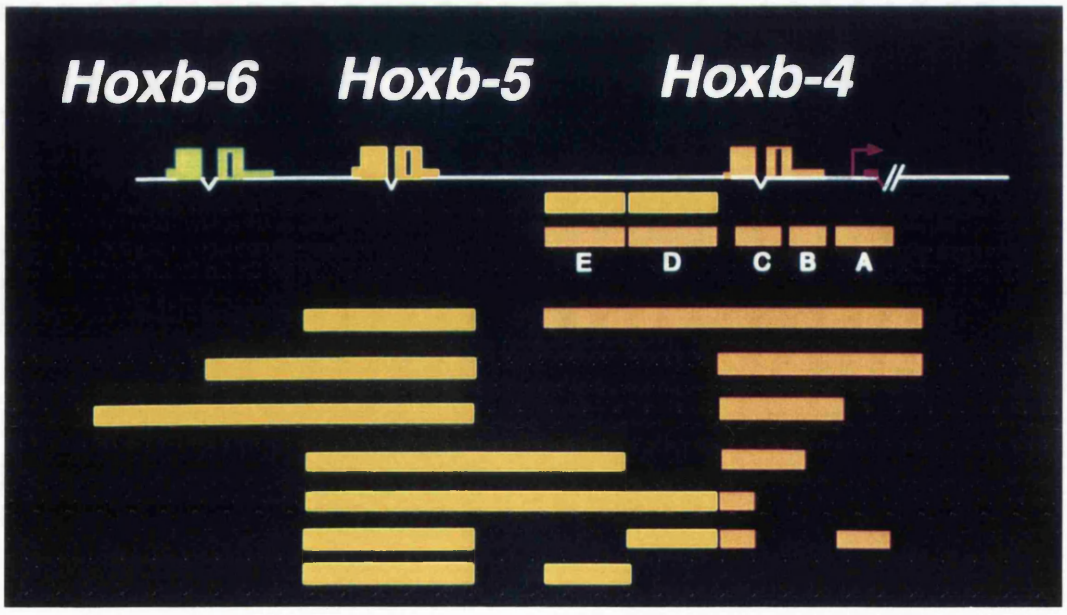
The experiments described in this chapter were an attempt to prove whether any enhancers between adjacent genes can operate on both genes at once. The approach was to create large "double-reporter" constructs, which consisted of a wildtype stretch of DNA containing two Hox genes, each with a different reporter gene to monitor their expression. This approach was to be used for two gene pairs: *Hoxb-4/b-5*, and *Hoxb-1/b-2*. The second reporter gene used in each case (in conjunction with LacZ) was the human placental alkaline phosphatase gene (PLAP). After initial tests of the PLAP reporter in *Hoxb-1* and Hoxb-4, the latter pair (Hoxb-5/b-4) was chosen to

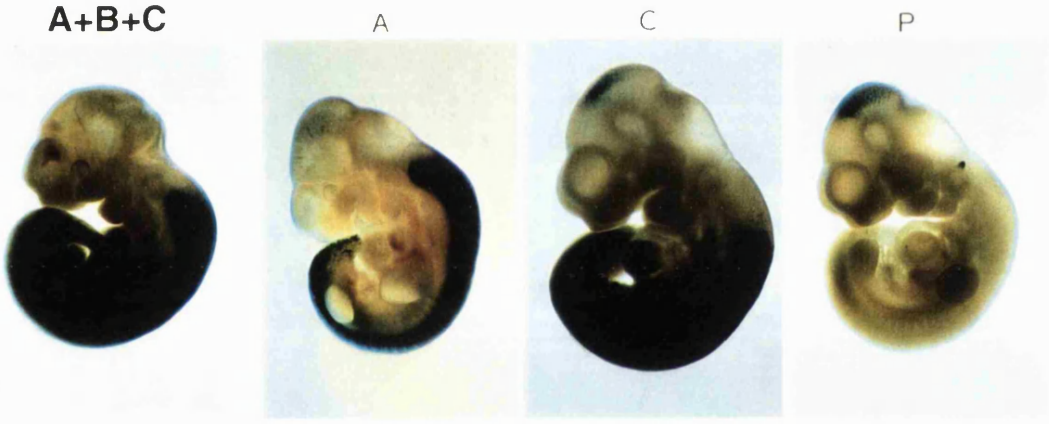
Fig. 4.1 Previous analysis on the regulation of *Hoxb-4* and *Hoxb-5*.

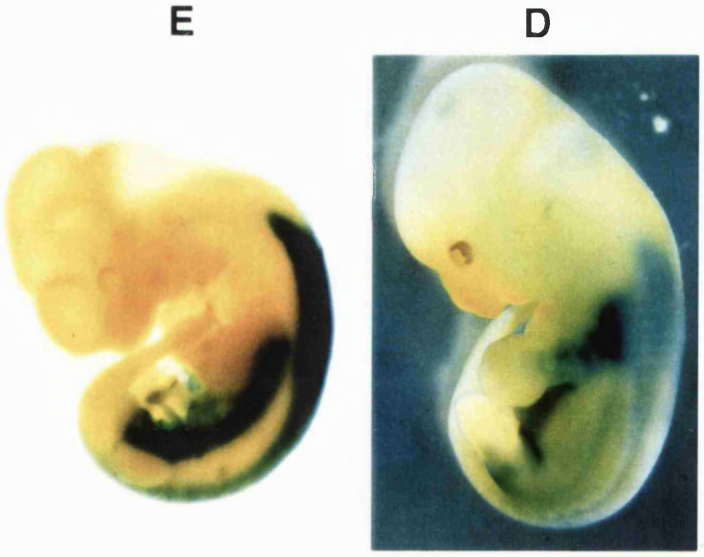
The top panel shows the organisation of the *Hoxb* complex for the *Hoxb-6,5* and *4* genes, and below them the stretches of DNA which have been tested for their regulatory effects. The attempt to find the important regulatory elements for *Hoxb-5* is shown by the yellow bars (performed by Stefan Nonchev). Extensions were first made in a 5' direction, and then in a 3' direction. Only when regions E and D were included was the correct regulation found, and the two last constructs tested them independantly on a minimal *Hoxb-5* region. The orange bars show the equivalent constructs used to define the region A, B and C for *Hoxb-4* (performed by J. Whiting).

Representative examples of the activity found for the regions A, C, D and E are shown below. In the central strip of images, the right-most embryo shows the pattern for activity of the *Hoxb-4* promoter alone. The expression seen in the midbrain, is a well-characterised misregulation which occurs from this promoter. The panels labelled A and C show independantly the effects of region A and C (notice the midbrain expression again occuring in the second of these two constructs). The A+B+C construct (second down of the orange bars) recreates almost the entire *Hob-4* pattern.

The bottom two photos show the neural activity of region E, and the mesodermal (plus weak neural) activity of region D, both on the *Hob-5* promoter (the bottom two yellow bars).


Hoxb-6
Hoxb-5
Hoxb-4
E
D
C
B
A


A+B+C
A
C
P


E
D

continue the shared enhancer analysis, while the *Hoxb-1* line was used for a different study described in chapter 5.

The basic experiment was as follows: First analyse the complete construct with both regulatory regions intact (regions D and E between Hoxb-5 and Hoxb-4). Then compare it with deleted constructs in which either element has been removed. If expression in a particular tissue is lost from the patterns of both reporter genes, then the deleted region must be responsible for this aspect of regulation for both promoters. All images of *Hox* protein expression patterns (revealed by antibody staining) were provided by Alex Gould.

4.1 Testing PLAP in the *Hoxb-1* gene

The first step was to determine whether PLAP was a suitable reporter gene for these experiments. This involved testing it on its own, and in conjunction with a second construct which uses the LacZ gene. The first test performed used construct b1-A (Fig. 4.2). From previous experiments in the lab, a 108bp deletion had been made to remove the translation-start codon from a 7.5kb EcoRV fragment which encompasses the *Hoxb-1* gene. In place of the deletion, an EagI restriction site had been engineered and this was used to insert a copy of the PLAP gene. The resultant 9.5kb construct was used to create a transgenic line (JS-4).

This line was used to optimise the staining protocol for the alkaline phosphatase. The original protocol used, was essentially the same as that used for the colour-reaction in a DIG-labelled *in-situ* hybridisation protocol. The main differences being that the embryos had to be heated to 65°C for 30 min. before staining, in order to inactivate the endogenous embryonic alkaline phosphatases (PLAP is a very heat-stable protein). The only modification which was found to improve signal-to-background contrast was incubating the embryos in levamisol for at least one hour before adding the BCIP and NBT (instead of adding it at the same time), see section **2.5.3**. The expression of PLAP was found very clearly in rhombomere 4 (r4) Fig. 4.2.

This transgenic line was then crossed with the line named ML-19, which contains a construct in which the LacZ gene is controlled by an r3/5 enhancer from *Hoxb-2*. This was chosen so that in the resultant double-transgenic embryos, cells expressing the different reporter genes would be adjacent to each other. The β-gal protein is denatured by high temperatures so analysing the embryos consisted of staining for β-gal first, then heating the embryos to 65°C for 30 min., and

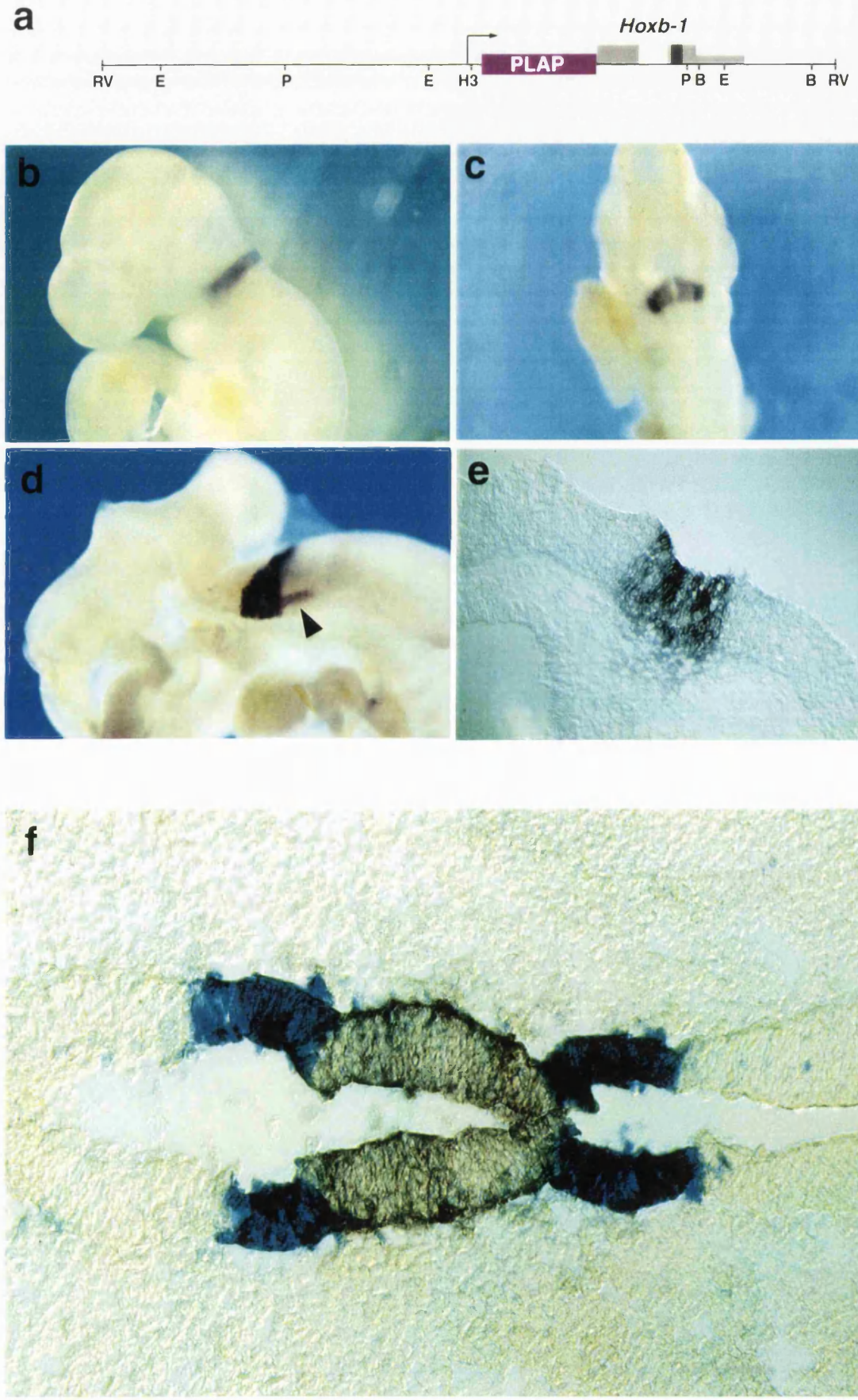
Fig. 4.2 Expression pattern of the rhombomere-4 enhancer from *Hoxb-1*.

(a) The construct b1-A, which is 9kb long, consists of a 7kb EcoRV fragment encompassing the *Hoxb-1* gene, with the PLAP gene inserted at the beginning of the first exon.

RV = EcoRV, E = EcoRI, P = PstI, H3 = HindIII, B = BamHI.

(b-e) Expression of the PLAP reporter gene in rhombomere 4 (r4). (b) and (c) show lateral and dorsal views of whole-mount stained embryos at 9.5dpc in which r4 can be seen as a strong pink stripe, and (d) shows the inside of a 11.5dpc embryo which was sagitally-bisected. The extension of *Hoxb-1*-positive cell bodies which move from r4 into r5 can be seen (black arrowhead). (e) shows a cryostat coronal section in which r4 can be seen as the group of strongly stained cells displaying sharp boundaries with the adjacent rhombomeres.

(f) A coronal section through the hindbrain of a 9.5dpc embryo which is transgenic for both the b1-A construct, and a second construct which directs expression of LacZ in rhombomeres 3 and 5. Anterior is to the left.


a
Hoxb-1
PLAP
RV
E
P
E
H3
P
B
E
B
RV
b
c
d
e
f

then staining for PLAP. In these tests the double-staining procedure was shown to be suitable for the remaining experiments (Fig. 4.2e).

For the remainder of this chapter analysis of potentially-shared enhancers was performed on the *Hoxb-4/b-5* pair, and not the *Hoxb-1/b-2* pair. However, the *Hoxb-1*-PLAP line was found to display polydactyly, and was the subject of a study described in chapter 5.

4.2 Testing PLAP in the *Hoxb-4* gene

In a similar test to that described above, the PLAP gene was also tested in *Hoxb-4* constructs. Unlike *Hoxb-1*, there were no constructs available in which the translation-start codon had been deleted. Previous constructs using LacZ as a reporter, had used a SalI restriction site which lies 36bp inside the coding region. In such a fusion the LacZ gene was in-frame with the initial *Hoxb-4* coding region and a functional β-gal protein was produced. A large 12kb fragment surrounding *Hoxb-4* was chosen as a test construct, which had been shown to produce almost all of the normal *Hoxb-4* expression pattern (Whiting *et al.*, 1991). The 5' sequence of the PLAP gene was not known, but the PLAP coding region was enclosed within a 1.9kb SalI fragment, so it was decided to ligate this fragment into the SalI site. The construct generated (named b4AP) could then be quickly tested, and the 5' region of the SalI fragment sequenced to determine whether the fusion was in-frame.

The transgenic embryos from b4AP displayed the expected expression pattern for the *Hoxb-4* gene (Fig. 4.3). The construct was sequenced from a primer in the *Hoxb-4* promoter, using the dideoxy-nucleotide method developed by Sanger et al. Surprisingly, the PLAP gene was not in-frame with the *Hoxb-4* coding-region (Fig. 4.4). According to Kozak (Kozak, 1989), translation is rarely initiated from an ATG codon if it is not the most 5' one. This is because the initial ribosome sturcture is thought to bind to the mRNA molecule at the 5' cap site, and then scan in the 3' direction until reaching a translation start codon (which will be the most 5' one). The proportion of ribosomes which slide past this first ATG (a process known as "leaky scanning") is thought to be very low unless it does not reside in a Kozak consensus sequence. The *Hoxb-4* ATG codon does not possess a Kozak consensus sequence, and this may be the reason for successful translation of the PLAP gene. Any translation which is initiated at the *Hoxb-4* ATG will produce a short 54 amino acid polypeptide, which is presumably non-functional (see Fig. 4.4). The fact that translation of PLAP is occuring from its own ATG may be extremely beneficial, as the first 17 amino acids of

Fig. 4.3 Expression pattern of the *Hoxb-4* regulatory regions.

(a) Embryos expressing the construct b4AP stained for PLAP, at stages 9.5dpc, 10.0dpc and 10.5dpc. The overall pattern is the same as that seen for the equivalent LacZ construct, and mirrors the normal *Hoxb-4* expression pattern very closely.

(b) A coronal section through the hindbrain of a 10.5dpc embryo expressing b4AP. The anterior boundary of expression is very close to the r6/7 junction. (ov = otic vesicle).

(c) Close-up of a 10.5dpc whole-mount stained embryo, showing the strong staining of the vagus nerve (open arrowhead).

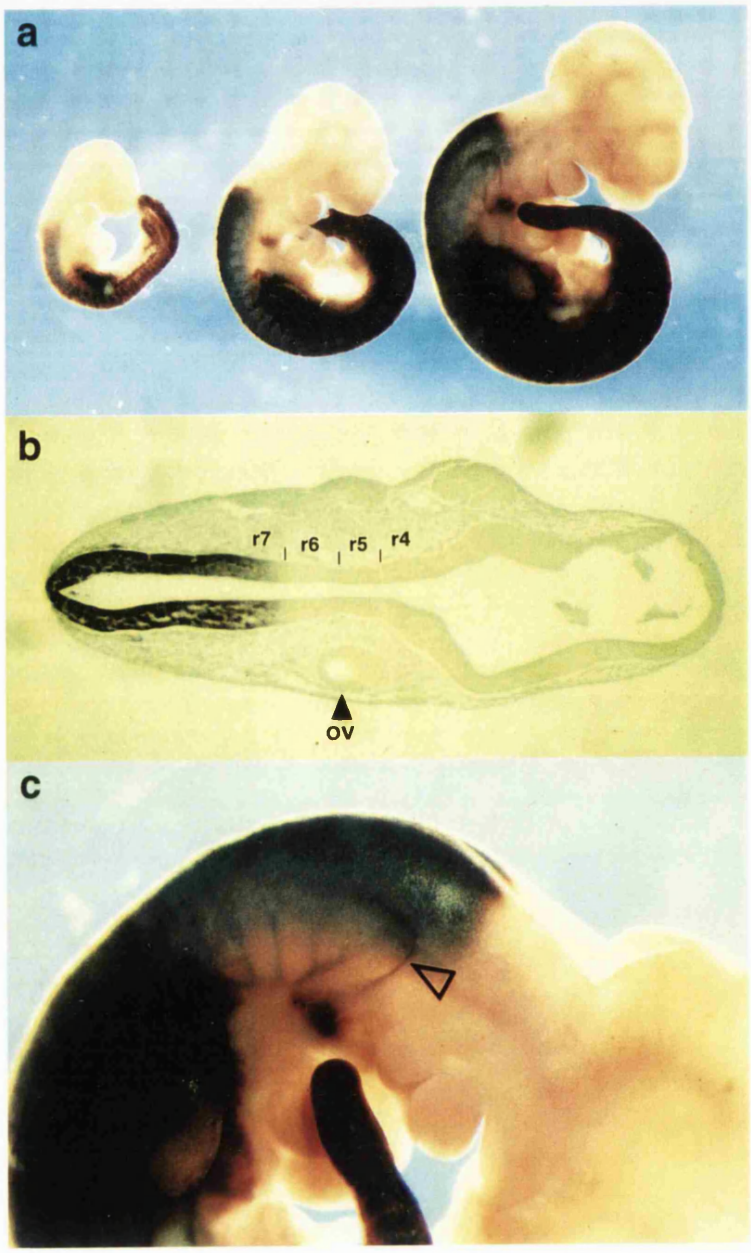

a
b
r7
r6
r5
r4
ov
c

Fig. 4.4 DNA sequence of the PLAP insertion into the *Hoxb-4* coding region.

The DNA sequence is shown in bold starting just before the ATG of the *Hoxb-4* gene, and ending after the first 25 triplets of the PLAP coding sequence. Above it is a translation of the *Hoxb-4* protein sequence, and below it is the translated PLAP protein sequence. The junction between the endogenous *Hoxb-4* sequence and the reporter sequence is indicated as the SalI site. The predicted (presumably non-functional) polypeptide which would be created from translation starting at the normal ATG, continues beyond this SalI site and terminates within the coding region for the PLAP protein.

(Start of *Hoxb-4* translation)

MET Ala MET Ser Ser Phe Leu Ile Asn Ser Asn

AAATTA ATG GCT ATG AGT TCC TTT TTG ATC AAC TCA AAC

Tyr Val Asp Pro Ser Ser Asp Pro Leu Val Thr Ala Ala

ATA GTC GAC CCG AGC TCG GAT CCA CTA GTA ACG GCC GCC

SalI

Ser Val Leu Glu Phe Val Pro Arg His Cys Pro Ala Ala

AGT GTG CTG GAA TTC GTC CCT CGC CAC TGT CCT GCT GCC

Leu Gln Thr Cys Trp Gly Pro Ala Cys Cys Cys Cys Cys

CTC CAG ACA TGC TGG GGC CCT GCA TGC TGC TGC TGC TGC

MET Leu Leu Leu Leu

(Start of PLAP translation)

(End of truncated *Hoxb-4* translation)

Cys Cys Trp Ala STOP

TGC TGC TGG GCC TGA GG CTA CAG CTC TCC CTG GGC ATC

Leu Leu Leu Gly Leu Arg Leu Gln Leu Ser Leu Gly Ile

SIGNAL PEPTIDE

ATC CCA GTT GAG GAG GAG AAC ... remainder of PLAP ORF.

Ile Pro Val Glu Glu Glu Asn ...

the protein are a signal peptide which directs it to the cell membrane, and this function may be impaired if a significant number of extra amino acids are fused onto the amino terminal.

The cell-membrane localisation of the PLAP protein may be of benefit to its use as a reporter protein, because along the length of long, narrow axonal projections the amount of cytoplasm is very small, whereas the amount of membrane is relatively high. In agreement with this hypothesis projection of the vagus neurons was seen even more clearly in the transgenic embryos than in previous cases where LacZ had been used as the reporter (Fig. 4.3c).

4.3 Building constructs for the analysis of *Hoxb-4* and *Hoxb-5*

In order to build the large double-reporter constructs for this study a low-copy pBR-based plasmid was used, called pGP1f (created by GenPharm). It had specifically been created for large-size constructs, and had been successfully used with 80kb inserts (pers. comm. M. Rubock). The strategy followed in designing constructs was based on the idea of "cassettes". This was done for two reasons: Firstly, the longer a stretch of DNA is, the more chance it has of containing any particular restriction site. Unique sites are important for construct-building as they allow DNA segments to be cut-out at specific positions or ligated together in predetermined orientations. Secondly, flexibilty is useful in a construct-building strategy, in case a useful modification is desired after a number of steps have already been carried out, or if a particular step (usually a ligation) proves difficult to achieve. All the constructs used in this part of the study are shown in Fig. 4.5.

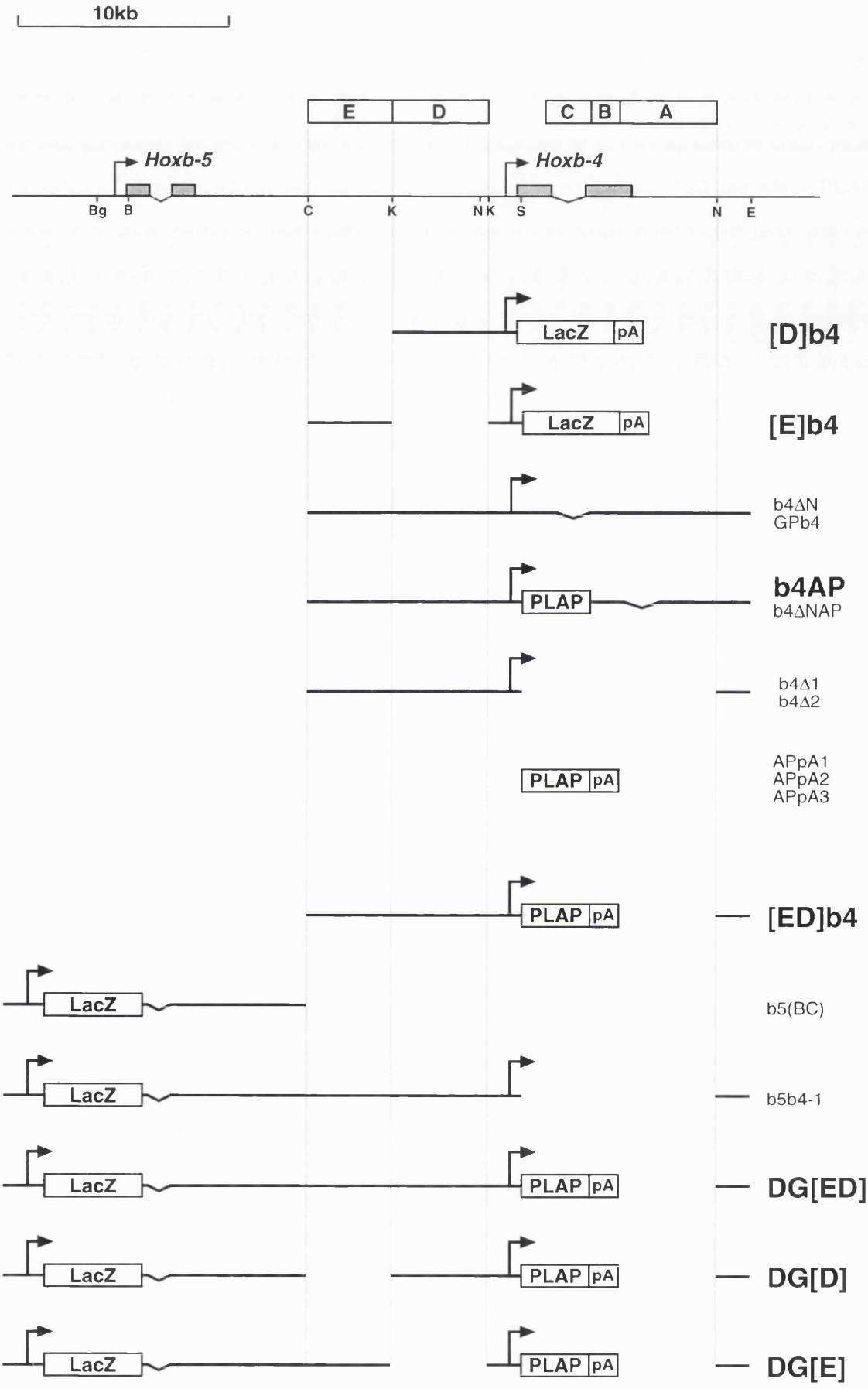
In order to give concise names to the constructs from this study, a simple nomenclature was adopted. Letters in square brackets refer to the defined regulatory regions (eg. [ED]), whereas those enclosed in rounded brackets refer to restrictions sites which define the boundaries of the DNA segment (eg. (BC) refers to a *BglII* to *ClaI* fragment). ΔN refers to a destroyed NcoI site, AP to alkaline phosphatase, and pA to the poly-adenylation signal from the SV40 virus. All constructs containing the *Hoxb-5* or *Hoxb-4* promoter possess 'b5' or 'b4' in their name except for the large double-reporter constructs which are named 'DG', which stands for *double gene*.

The first constructs used were built by Jenny Whiting, to test regions D and E on the *Hoxb-4* promoter - [D]b4 and [E]b4. The first PLAP construct tested (described above) was b4AP, created simply by inserting the PLAP gene into a SalI site in the coding region of *Hoxb-4*. The 5' end of region A was defined by an NcoI site, but there was also one near the *Hoxb-4* promoter. In order to be able to remove (and replace) the regions C, B and A, the 3' NcoI site had to be

Fig. 4.5 DNA constructs for transgenic analysis of *Hoxb-5* and *Hoxb-4*.

The organisation of the *Hob-5/Hoxb-4* region is shown as the top line, above which the regions A to E are indicated. The constructs which were injected for this thesis are named in bold. The remaining constructs were intermediary steps used in the construct-building strategy. The promoters are indicated as black arrows, the LacZ, PLAP and pA sequences as boxes, and the *Hox*-complex DNA as thick black lines.

Bg = BglII, B = BamHI, C = ClaI, K = KpnI, S = SalI, N = NcoI, E = EcoRI.


10kb
E
D
C
B
A
Hoxb-5
Hoxb-4
Bg
B
C
K
N
K
S
N
E
LacZ
pA
[D]b4
LacZ
pA
[E]b4
b4ΔN
GPb4
PLAP
b4AP
b4ΔNAP
b4Δ1
b4Δ2
PLAP
pA
APpA1
APpA2
APpA3
PLAP
pA
[ED]b4
LacZ
b5(BC)
LacZ
b5b4-1
LacZ
PLAP
pA
DG[ED]
LacZ
PLAP
pA
DG[D]
LacZ
PLAP
pA
DG[E]

destroyed, which created b4ΔN. This was then inserted into the pGP1f plasmid to create GPb4. From this, two PLAP versions were created. In b4ΔNAP, PLAP was simply inserted into the SalI site used previously for b4AP. For the second version, a number of modifications were made: First, the region CBA was removed and the construct re-ligated to recreate the SalI site where PLAP would be inserted (b4Δ1); second a polyadenylation signal was added to the PLAP gene, and this was engineered into a SalI fragment through the three steps of APpA1 to APpA3; third, in order to create a unique XhoI site at the 5' end of the construct so that the *Hoxb-5* regions could be included the 5' XhoI site from the polylinker was destroyed (b4Δ2); and finally the APpA3 SalI fragment was ligated into b4Δ2, to create [ED]b4.

To create the double-reporter constructs, first a construct from Stefan Nonchev was modified to generate a BglII to ClaI fragment (b5(BC)). To join the *Hoxb-5* and *Hoxb-4* sections together, there were two options. The most obvious was to insert b5(BC) into [ED]b4, but unfortunately this proved impossible. The second (which succeeded) was to first insert b5(BC) into b4Δ2, creating (b5b4-1), and then re-insert the APpA3 section. In this full-length construct, DG[ED], there were three KpnI sites which defined the boundaries of regions E and D. The deletions were therefore made by cutting with this enzyme, and and then re-ligating the same mixture of fragments. This produced the two deletions DG[D] and DG[E].

4.4 Analysis of regions D and E on the *Hoxb-4* promoter

In order to interpret results from the double-reporter experiments, it was first necessary to further test the two regions D and E on a single promoter, and for this purpose the *Hoxb-4* promoter was chosen. Analysis of D and E together was performed using the alkaline phosphatase construct, [ED]b4. Analysis of D and E independently was performed using the two previously created constructs #1191 and #1180, which are here named [D]b4 and [E]b4.

4.4.1 Region D+E

The construct [ED]b4 was created to test the combined regulatory influences that regions D and E could have on the *Hoxb-4* promoter, and to see whether the resultant patterns were more similar to *Hoxb-4* or *Hoxb-5*. It consisted of a complete stretch of DNA including E, D and the *Hoxb-4* promoter, such that the wildtype spacing of enhancers relative to the promoter was

maintained. Two transgenic lines were generated which both displayed the same expression pattern. A short time-course of one of these lines (JS-10) is shown in Fig. 4.6.

Panels (a-d) show the expression patterns of the *Hoxb-5* and *Hoxb-4* proteins, revealed by antibody staining. The pattern of the transgenic construct at the same stage (10.5dpc) displayed anterior boundaries of expression similar to *Hoxb-5* (panels e-g). The neural boundary was at a similar distance posterior to the otic vesicle (compare the transgene in (e) and (g), with *Hoxb-5* (a) and *Hoxb-4* (d)). Expression in the somitic mesoderm extended to the two somites anterior to the forelimb bud (f), which is the same as that seen for the *Hoxb-5* gene (b). Whereas the *Hoxb-4* protein can clearly be seen in three anterior somites (panel c). However, one site of expression was seen in the transgenics which is more similar to the *Hoxb-4* pattern - strong staining in the forelimb-bud. Although *Hoxb-5* protein can be seen in the forelimb, it is concentrated in a small anterior patch, and is not as strong as *Hoxb-4* which is expressed throughout the bud (compare panels (b) and (c)). It is therefore possible that a forelimb enhancer exists within this region which is more important for *Hoxb-4* than *Hoxb-5*.

Panels (h-j) show the expression pattern at 11.5dpc and 12.5dpc. During this stage expression in both fore- and hindlimbs remained strong. The pattern in the somites remains unchanged, but in the neural tube it appears to extend anteriorly (beyond the cervical flecture by 12.5dpc), and the vagus nerve becomes positive.

4.4.2 Region D

Transgenic embryos were made from the construct [D]b4, which consists of region D, the *Hoxb-4* promoter and the LacZ gene. Ten transient transgenic embryos were created which expressed β-gal, all of which displayed the same pattern of expression at different intensities. This basic pattern is shown in Fig. 4.7a-c. (The β-gal staining in the midbrain region is a well-characterised mis-regulation from the *Hoxb-4* promoter, which occurs in many of these constructs.)

Staining in the neural tube was the same as that seen for the combined D+E construct, displaying a *Hoxb-5* pattern. Expression in the somites, although less well-defined, also appeared unchanged. Interestingly, the forelimb expression, while generally strong, was weakest in the anterior region. This is the site of strongest limb expression for the *Hoxb-5* protein, thereby reinforcing the idea that this limb enhancer is more important for *Hoxb-4*, and additionally demonstrating that it resides within region D.

Fig. 4.6 Expression pattern of region E plus D.

Panels (a-d) show the expression patterns of the *Hoxb-5* and *Hoxb-4* proteins, revealed by antibody staining. The black semi-circles indicate those positively-staining somites which are anterior to the forelimb bud. The black arrowhead indicates the position of the otic vesicle. Panels (e-g) show the extent of the PLAP staining from the construct [ED]b4 in the neural tube and the somites at 10.5dpc. The two black semi-circles in (f) show that the boundary of expression in the somites is more similar to *Hoxb-5* than *Hoxb-4*. Panels (h-i) show expression patterns from the same construct at 11.5 dpc and 12.5dpc. The open arrowhead indicates the positively-staining vagus nerve.

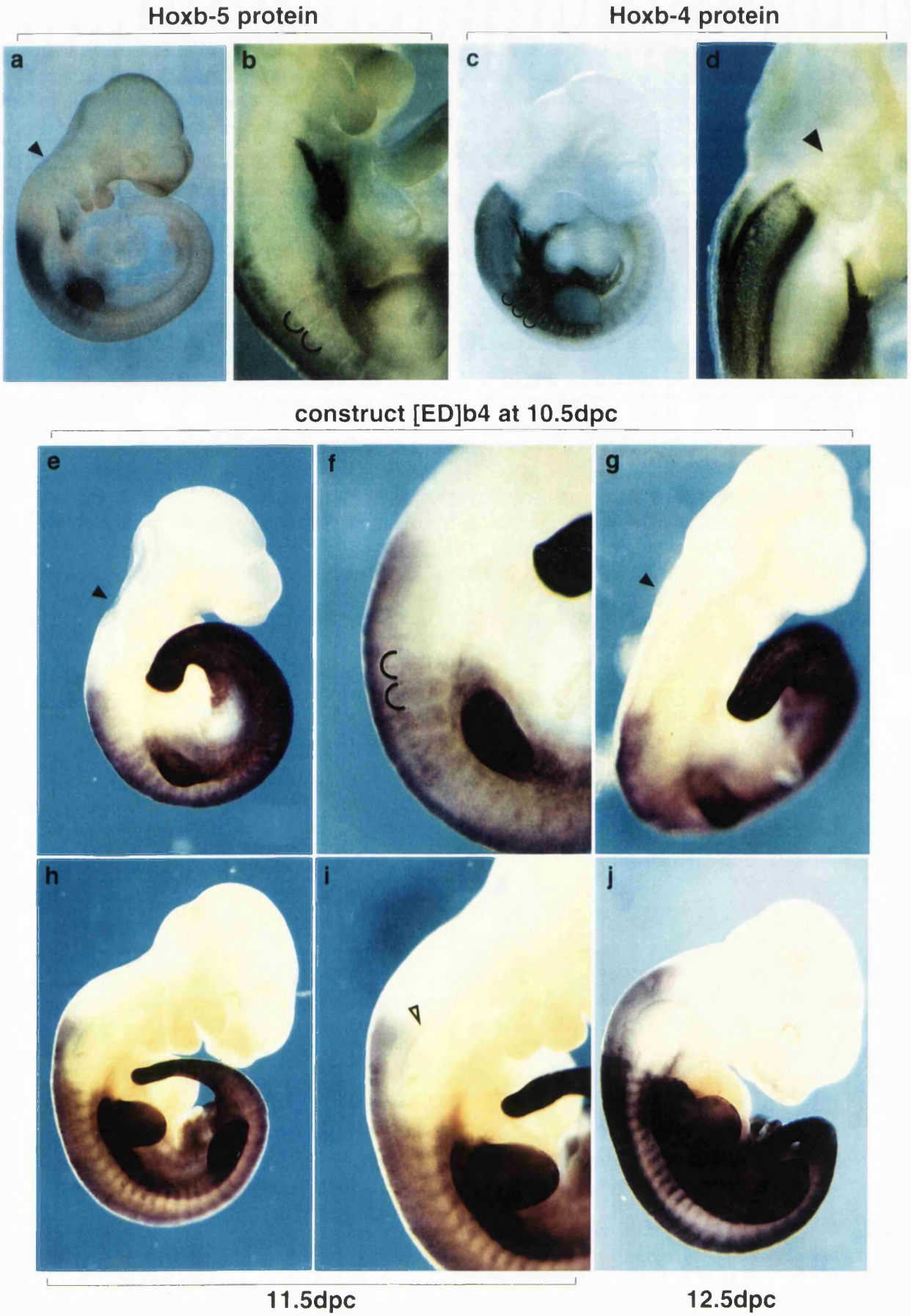
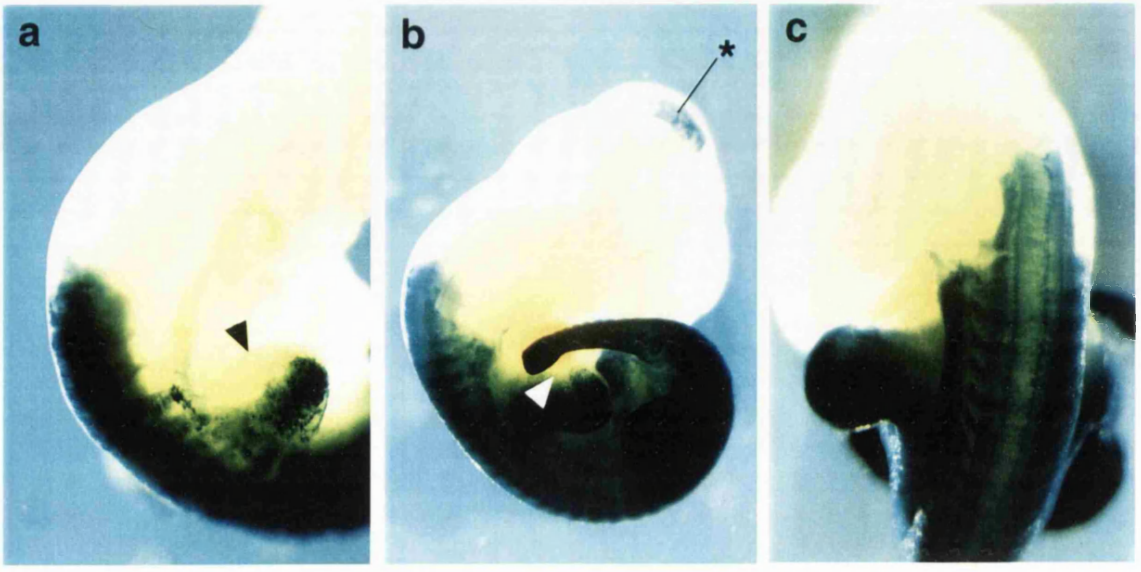

Hoxb-5 protein
Hoxb-4 protein
a
b
c
d
construct [ED]b4 at 10.5dpc
e
f
g
h
i
j
11.5dpc
12.5dpc

Fig. 4.7 Expression patterns of regions E and D separately.

Panels (a-c) show the pattern of construct [D]b4 in transient transgenic embryos of ages 10.0dpc and 11.0dpc. Staining of the b-galactosidase protein can be seen in the neural tube, somites, lateral mesodermand the limbbud. The anterior part of the forelim-bud, which is negative, is indicated by the black and white arrowheads. The staining in the midbrain (indicated by the asterisk) is a well-characterised ectopic regulation driven by the *Hoxb-4* promoter.

Panels (d-f) show the expression pattern of region E, from the construct [E]b4. (d) and (e) show two embryos from the JS-5 line, and (f) shows one from the JS-6 line, which expresses in a weaker subdomain. The mesonephric expression is indicated by *mn*.

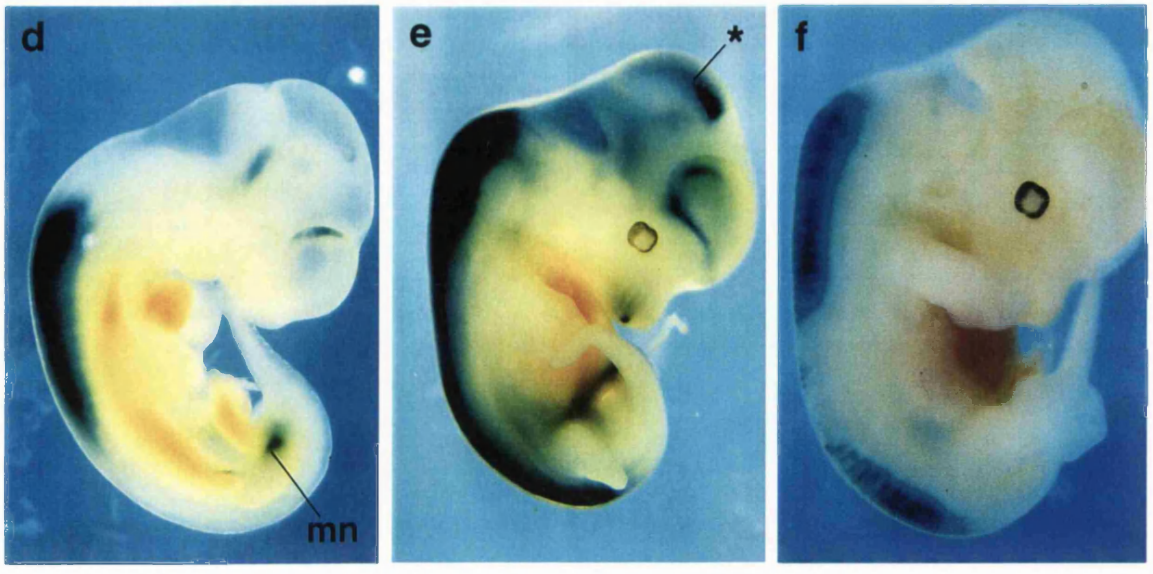
[D]b4



10.0dpc

11.0dpc

[E]b4



10.5dpc

12.5dpc

4.4.3 Region E

Region E was tested using the construct [E]b4, which was made by ligating region E to the *Hoxb-4* promoter with LacZ. Three expressing transgenic lines were created, all of which displayed the same staining patterns. Two of these lines (JS-5 and JS-6) are shown in Fig. 4.7d-f.

Panels (d) and (e) show 10.5dpc and 12.5dpc embryos from JS-5. From these, region E was seen to display strong and specific neural activity. The anterior boundaries were again unchanged from construct [ED]b4APpA, suggesting that neural enhancers with a *Hoxb-5* anterior boundary probably exist in both regions D and E. However, the absence of any somite, lateral mesoderm or limb-bud expression demonstrates that unlike D, region E is fairly tissue-specific. The only other site of expression regularly seen was the mesonephric ducts (seen in panel d). Panel (f) shows a 12.5dpc embryo from JS-6, which was much weaker, but still displayed a pattern which resides purely in the neural tube, and within the limits of the pattern shown by JS-5.

4.5 Promoter-specific expression from shared enhancers

Construct DG[ED] was generated as described in section **4.3**. It was designed as a complete wildtype stretch of DNA which runs from the *Hoxb-5* promoter to the *Hoxb-4* promoter, including all the regulatory sequences in between (region D and E), with the LacZ gene inserted just after the *Hoxb-5* promoter, and the PLAP gene inserted just after the *Hoxb-4* gene. In this experiment it was hoped to determine whether regions D and E could regulate both promoters in one construct, and to see if this would result in the same pattern for each.

To analyse this construct, three transient embryos and three transgenic lines were generated. All transgenic embryos displayed staining for both the LacZ and the PLAP reporter genes, however the LacZ staining was often extremely weak. This was considered to be an intrinsic aspect of the interaction between the LacZ gene and the *Hoxb-5* promoter, and since the pattern observed was invariant, it was not considered to be a problem.

Fig. 4.8 shows the patterns of expression for the *Hoxb-5* protein and the DG[ED] construct. Panels (a) to (c) show that at 9.5dpc expression from both promoters was very similar to the *Hoxb-5* protein. In all three cases expression was seen in the limb-bud, in mesoderm extending posteriorly to the tip of the tail, and up to the same anterior boundary in the neural tube, and in the somites (which at this age, is adjacent to the anterior limb-bud margin). There was also noticeably

Fig. 4.8 Expression pattern from the construct DG[ED].

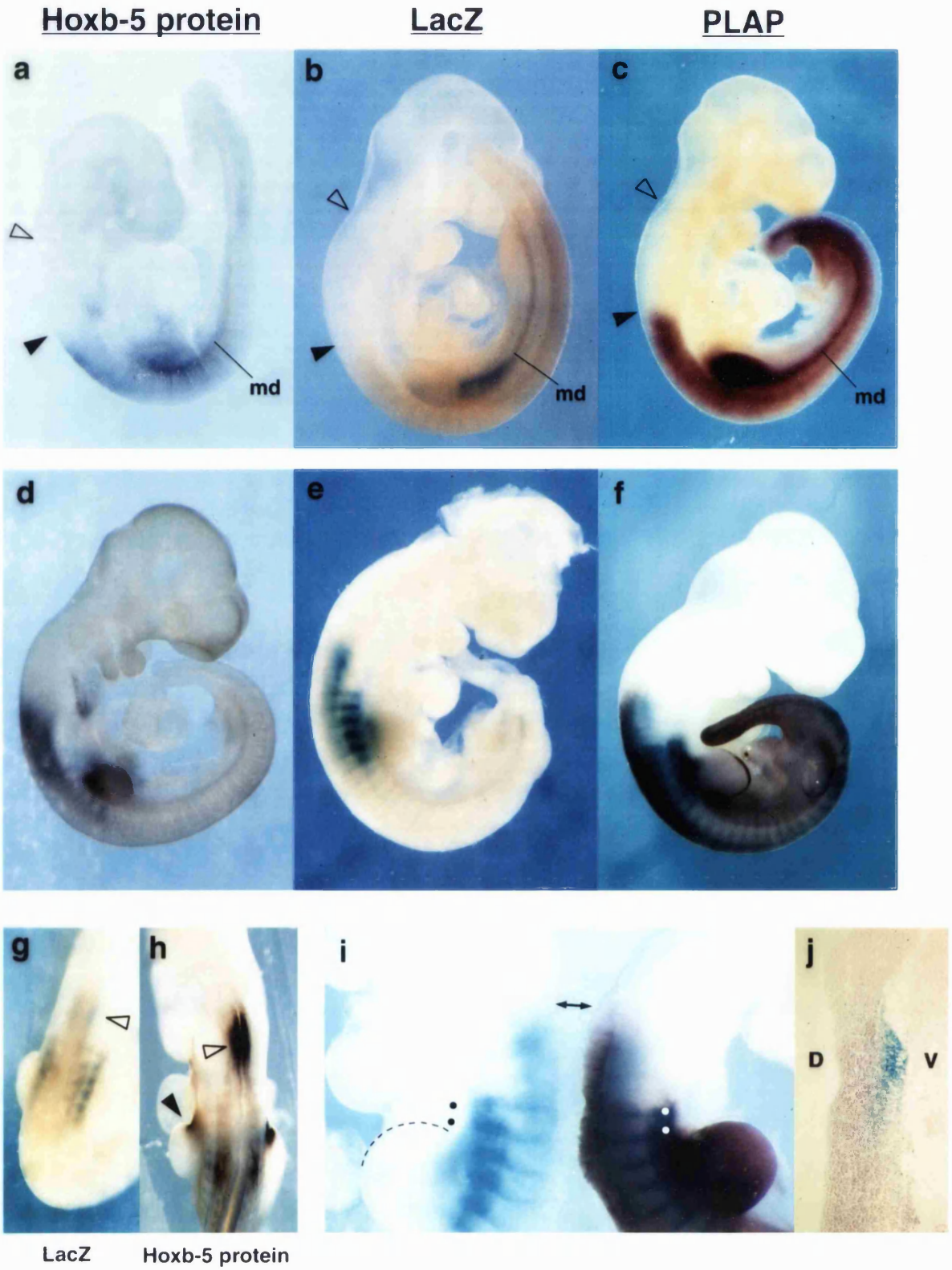
(a-c) 9.5dpc embryos showing the pattern of the wildtype *Hoxb-5* protein, compared to the LacZ pattern and the PLAP pattern from the construct DG[ED]. LacZ patterns are derived from the *Hoxb-5* promoter, and PLAP patterns are driven from the *Hoxb-4* promoter. It can be seen that expression from both promoters is very similar to the *Hoxb-5* protein. Open arrowheads indicate the position of the otic vesicle, and closed arrowheads show the anterior extent of expression in the neural tube. Although the protein staining and the LacZ staining both lack sharp boundaries, their most anterior limit is very similar to that for the PLAP staining. It is also clearly seen that expression from both promoters extends posteriorly to the tail, and is strong in the mesonephric ducts (*md*).

(d-f) Three embryos of age 10.5dpc, which were stained in the same way as (a-c). At this later stage, the LacZ pattern from the *Hoxb-5* promoter is more similar to the endogenous *Hoxb-5* promoter than the PLAP pattern from the *Hoxb-4* promoter. The first two fade posteriorly in both the neural tube and the somites, whereas the PLAP expression is strong into the tail. Nevertheless, the anterior boundaries for both reporters are still almost identical, and reflect the *Hoxb-5* protein boundary.

Panels (g) and (h) show that the rapid posterior fading of LacZ expression in the neural tube, such that only a small A-P zone displays high levels, is reflected by the *Hoxb-5* protein expression pattern (both are indicated by open arrowheads). The closed arrowhead shows the small patch of *Hoxb-5* expression in the anterior of the forelimb-bud, which is also reflected by the LacZ pattern (see expression in the forelimb-bud in panels (e) and (I), as compared to the PLAP pattern (f) and (I)). This pattern is complementary to the expression of PLAP from the [ED]b4 construct shown in Fig..4.7a,b.

Panel (i) shows a back-to-back comparison of two halves from one 10.5dpc embryo expressing the DG[ED] construct. The whole embryo was stained for LacZ, then sectioned into two two halves, and the right half stained for PLAP. When the two sides were aligned, it was clear that the anterior boundaries of both neural tissue (double-arrow), and the somites are the same. The two somites anterior to the forelimb-buds are indicated by black and white dots.

(j) A sagital section through a 10.5dpc transgenic embryo, which demonstrates that double-positive cells can be identified. Although the D-V distribution of the b-gal does not reflect the normal regulation of either the *Hoxb-5* or *Hoxb-4* proteins, (the LacZ pattern is mostly ventral, whereas the PLAP pattern is more uniform) there is a zone of overlap in the middle of the neural tube, where double-positive cells can be seen.


Hoxb-5 protein
LacZ
PLAP
a
b
c
md
md
md
d
e
f
g
h
i
j
D
V
LacZ
Hoxb-5 protein

high expression in the mesonephric ducts. There appeared to be very little promoter-specificity in the response to regions D and E, at this stage.

Just 24 hours later, the LacZ and PLAP patterns looked significantly different. The anterior boundaries of expression in both the somites and the neural tube were still the same, and they were still at the level of *Hoxb-5* boundaries (as opposed to *Hoxb-4*). However, the posterior boundaries in the same two tissues were very different, and more reminiscent of their normal promoter activity. The LacZ pattern from the *Hoxb-5* promoter faded out rapidly in the posterior direction, such that only 6 or 7 somites were positive, and the neural expression faded out at a similar level of the A-P axis. This posterior fading is reflected in the *Hoxb-5* protein pattern (panel d). Although D-V patterning of the neural tube appeared to be quite different between the LacZ reporter and the *Hoxb-5* protein, the A-P patterning was strikingly similar. Direct comparison of dorsal views (panels (g) and (h)) show that in both cases neural expression was strongest in just a small patch anterior to the somitic expression. In contrast, the PLAP pattern from the *Hoxb-4* promoter did not fade out posteriorly in either the neural or mesodermal tissues. Antibody staining for the *Hoxb-4* protein shows a similar trait (Fig. 4.5c), although not quite as pronounced as the PLAP staining.

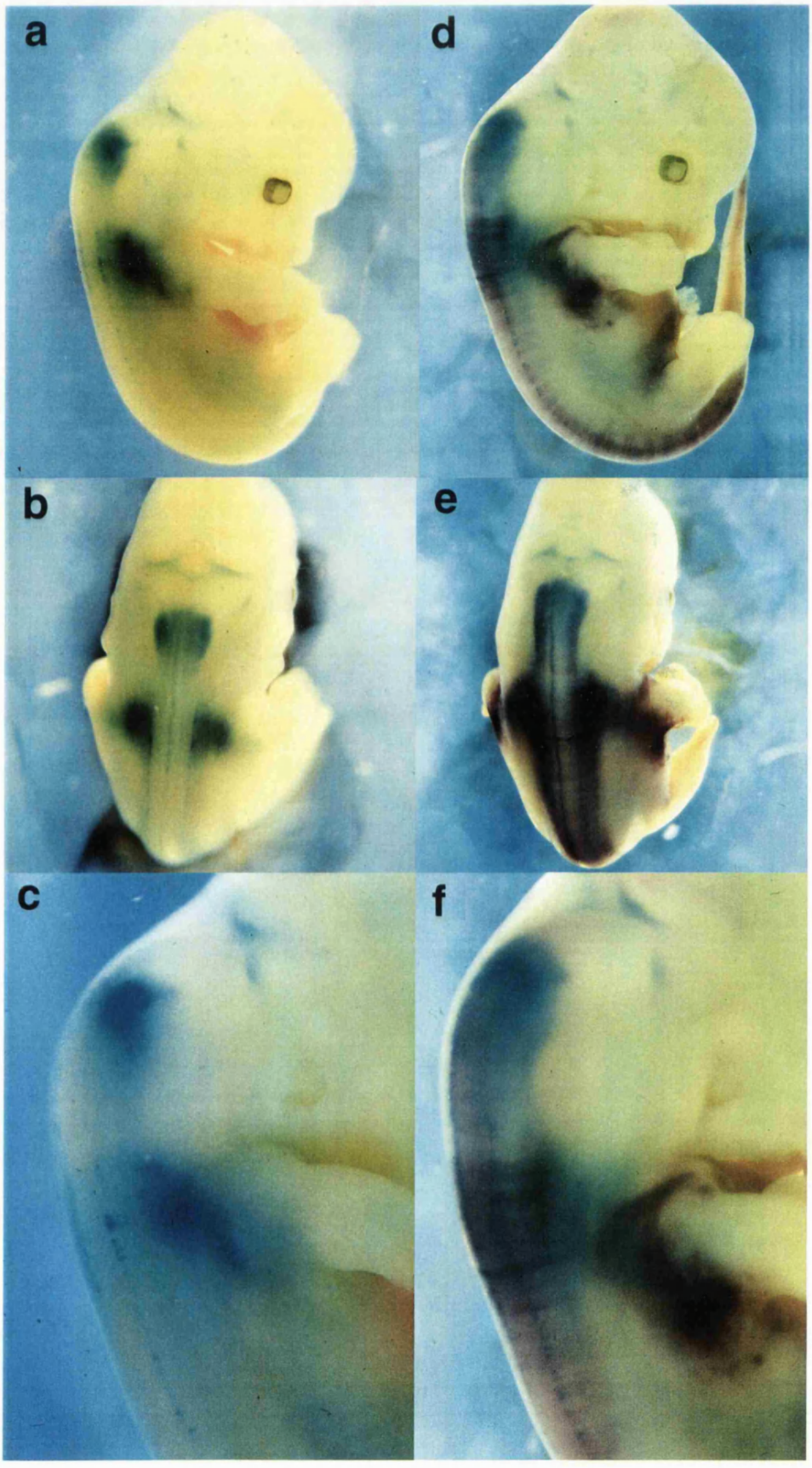
The other tissue in which differences could be found is the limb-bud. LacZ staining was weak in the limb-bud and appeared slightly stronger in the anterior, proximal region, which again reflects the pattern seen for *Hoxb-5* protein. PLAP staining was strong throughout the limb-bud and therefore reflects the normal *Hoxb-4* pattern. The anterior pattern seen for the LacZ reporter and the *Hoxb-5* protein is complementary to the expression of PLAP in the construct [ED]b4, which is strong throughout the forelimb-bud *except* for the anterior patch.

From these results we can conclude that the responses to D and E from the two promoters display some characteristics in common, and some which are different. The common aspects are the A-P positioning of anterior boundaries of expression for neural and somitic tissue, so this information seems to be enhancer-specific. The differences in response - the posterior fading and limb-bud expression - are therefore due to promoter-specific integration of regulatory information. These experiments also demonstrate that double-positive cells can be found in the neural tube, which are transcribing from both the *Hoxb-4* and *Hoxb-5* promoters (panel j).

This construct was also tested in 12.5dpc embryos (Fig. 4.9). The general features described above, were found to be maintained at this stage. The anterior boundaries of expression for the two reporters was the same in the neural tube and the post-somitic mesoderm, and corresponded to the *Hoxb-5* pattern. The posterior expression was significantly different, such that

Fig. 4.9 Expression at 12.5dpc from construct DG[ED].

The DG[ED] construct was tested in 12.5dpc embryos. This Fig.ure shows a single embryo photographed first after the LacZ staining (a-c), and then after the subsequent staining for PLAP (d-f). As at earlier time-points, expression from the *Hoxb-5* promoter is much more limited than that from the *Hoxb-4* promoter. In the neural tube it is still restricted to only a small A-P zone near the base of the hindbrain, and a small patch of tissue, dorsal to the forelimbs, which is probably derived from the few somites which were positive for LacZ at 10.5dpc. By contrast, the PLAP expression extends through the neural tube, and additional mesoderm derivatives, posteriorly into the tail. The only significant change from earlier time-points is that LacZ expression in the neural tube appears to span the whole D-V axis, whereas at 10.5dpc, it is only present in the ventral half.


a
d
b
e
c
f

LacZ was still only expressed in a small anterior patch of the neural tube and small region of mesoderm adjacent to the forelimb, whereas the PLAP expression extended posteriorly to the tail.

4.6 Deleting D or E from the double-reporter construct

In order to test whether elements in either D or E could be working on both the *Hoxb-5* and *Hoxb-4* promoter, the two final constructs were tested. In DG[D] region E was removed from the complete DG[ED] construct, and in DG[E] region D was removed.

Three transgenic lines were generated using the DG[D] construct (JS-15,16 and 17) however, the first two showed no expression of LacZ. Due to the generally weak expression of LacZ seen throughout these experiments, it was assumed that the JS-17 line was likely to reflect a genuine regulatory event within this construct (especially due to the pattern seen). Embryos from JS-15 and JS-17 are shown in Fig. 4.10. Panels (a) and (b) show the LacZ pattern, which is a purely somitic one. It was identical to the LacZ pattern seen from construct DG[ED] with the neural expression removed. Combined with the fact that this pattern is the same as the somitic expression of *Hoxb-5* protein, this is evidence for a *Hoxb-5* somite-enhancer in region D. Conversely, the fact that adding region E back to this construct (to create DG[ED]) regenerated the normal *Hoxb-5* neural pattern is evidence for a *Hoxb-5* neural-enhancer in region E.

The PLAP expression from DG[D] was basically unchanged from the previous construct (panels (c) and (d)). This confirms the results from section **4.2.2** that enhancers for at least three tissues reside in region D: neural tube, somites and forelimb. As described in the previous section these enhancers are specifying *Hoxb-5* A-P patterning for the first two tissues and *Hoxb-4* patterning for the limb-bud.

Two transgenic lines were generated from the construct DG[E], and their expression pattern is shown in panels (e) and (f). No staining from the LacZ gene was observed in either of the lines, but significant PLAP staining was found. In both lines PLAP was expressed in the neural tube, and in one of them (JS-14) it was also seen in the lateral mesoderm, between the two limb-buds. The neural expression had a less sharp anterior boundary than seen in the previous constructs, and its general position appeared to be more posterior than normal, extending only slightly more anterior than the forelimb-bud.

The fact that no somite expression was seen from either of the reporter genes, strongly suggests that the important somite enhancer is in region D. The fuzziness of the neural expression

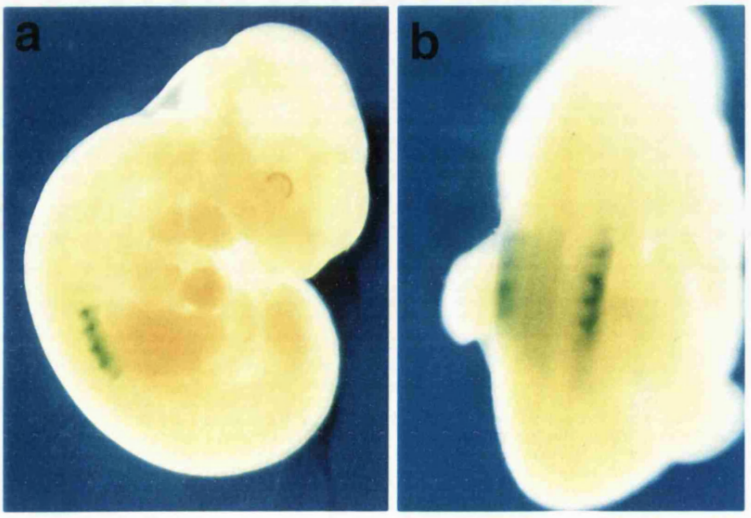
Fig. 4.10 Expression patterns from the two deletion constructs [D]b4 and [E]b4.

(a,b) LacZ expression in a 10.5dpc embryo from the DG[D] construct shows that when region E is deleted from the full double-labelled construct (DG[ED]), the neural expression from the *Hoxb-5* promoter disappears, but the somitic expression remains in exactly the same pattern as before (compare with Fig. 4.8e,g).

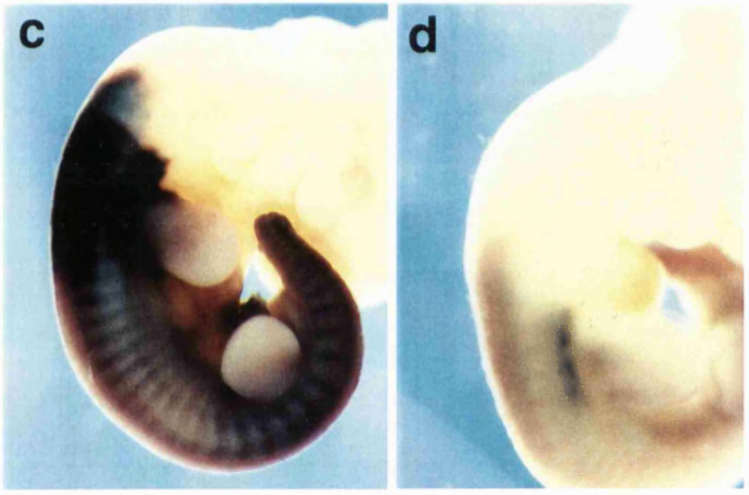
(c,d) When this same construct is tested for PLAP expression (from the *Hoxb-4* promoter), the pattern revealed is unchanged from DG[ED] (compare with Fig. 4.8f and Fig. 4.9d), which suggests that region E is not important for the regulation of *Hoxb-4*. The ages of the two embryos shown are 11.5dpc and 10.5dpc, and in both cases the blue LacZ-positive somites can be seen 'underneath' the purple PLAP-positive region.

(e,f) In the construct DG[E], it is region D which as been removed. No staining for LacZ was found in these embryos, but the PLAP pattern (driven from the *Hoxb-4* promoter) was striking, in that there was absolutely no somite staining, despite a clear maintenance of the neural expression. This sugests that the neural enhancer within region E may take part in *polar competition* for promoters (see text, section **4.7**). Also clearly visible is a stripe of lateral mesoderm expression between the two limb-buds.

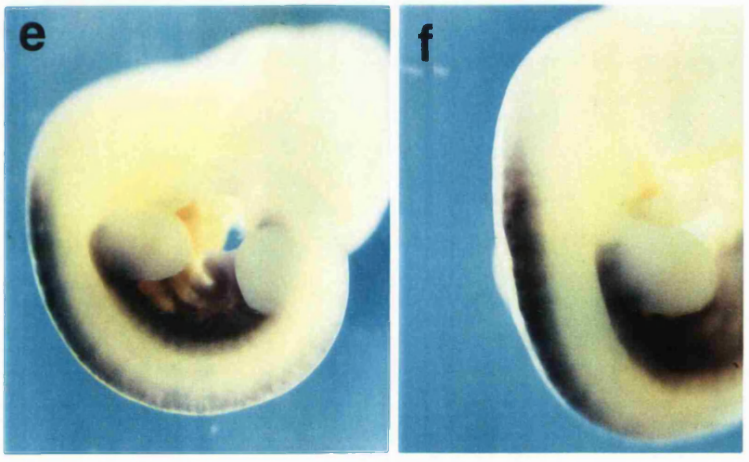
LacZ expression from DG[D]



PLAP expression from DG[D]



PLAP expression from DG[E]



boundary suggests that the enhancer in region E may not play a role in determining precise A-P positional information, but may instead be more of a tissue-specific element. Why there is no neural expression from the *Hoxb-5* promoter is unknown. It could either be due to competition between the two promoters for the same enhancer, or simply that the LacZ expression was too weak to detect.

4.7 Summary

The two regions D and E have been analysed in their capacity to regulate either the *Hoxb-4* promoter on its own, or both the *Hoxb-4* and *Hoxb-5* promoters in large double-reporter constructs. This was performed to determine whether there are any sites of expression from both promoters, which depend on one of the regulatory regions. Neural tissue is no longer a strong candidate for this phenomenon, for the following reasons: Region E appears to be a neural, tissue-specific enhancer. It generates neural patterns from the *Hoxb-4* promoter in construct [E]b4 and DG[E], and its removal from DG[ED] leads to a loss of neural expression from the *Hoxb-5* promoter (in construct DG[D]). However, neural expression is also clearly present in region D, since both [D]b4 and DG[D] display a *Hoxb-5* neural pattern from the *Hoxb-4* promoter.

Despite the fact that both regions contain neural elements, these results have shown that one of these enhancers can work on both promoters. The region E neural enhancer causes neural expression from the *Hoxb-5* promoter in b5[E] and DG[ED], and from the *Hoxb-4* promoter in [E]b4 and DG[E]. The fact that in this last construct it appears unable to drive expression from both promoters simultaneously, is evidence that this enhancer may specifically be "non-sharable" between two genes. This type of "polar competition" may be one of the explanations for how these non-specific enhancers can be restricted to the correct promoter *in-vivo*.

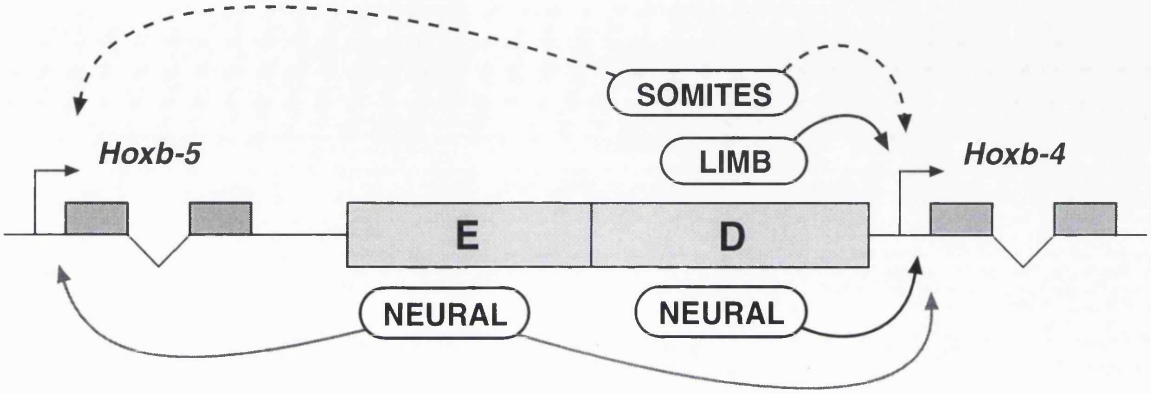
A candidate for a shared element should be confined to only one regulatory region, but be active on both promoters. This means that the limb enhancer is also ruled-out because it only works on the *Hoxb-4* promoter. However, somitic tissue is a good candidate, because the same anterior boundary of expression in the somites is found for both the *Hoxb-4* and *Hoxb-5* promoters (constructs [ED]b4, [D]b4, DG[ED] and DG[D]), and yet no somitic activity can be found without region D.

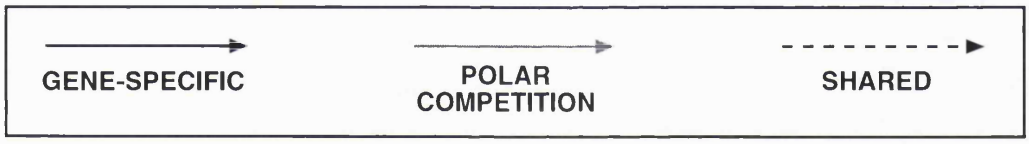
The resulting picture of regulatory interactions between the two genes is shown in Fig. 4.11. The region D neural and limb-bud enhancers are strictly operating on the *Hoxb-4* promoter. The

region E neural element would probably normally interact with the *Hoxb-5* promoter, but is also able to work on the *Hoxb-4* promoter. Only the somite enhancer appears to be conclusively operating on both promoters. Why such an enhancer, which directs a *Hoxb-5*-like pattern would be necessary for the *Hoxb-4* gene is unknown, and this is discussed in further detail in chapter 6.

Fig. 4.11 Summary of regulatory effects in regions E and D.

Diagram showing the basic structure of the *Hoxb-5/Hoxb-4* region with the four identifiable regulatory effects from regions E and D. The three different interaction types (distinguished by black, grey or dashed arrows) are described in more detail in the discussion (section **6.2.2**).


SOMITES
LIMB
Hoxb-5
Hoxb-4
E
D
NEURAL
NEURAL


GENE-SPECIFIC
POLAR
COMPETITION
SHARED

CHAPTER 5

Analysis of a new polydactylous mouse mutant

In the course of studying potentially shared enhancers in the Hox complex, I created a transgenic line (JS-4) which contained a construct containing the *Hoxb-1* gene linked to the PLAP reporter (named b1-A, see section **4.1**). It was found that a large proportion of the transgenic mice in this line (which were each tested for presence of the transgene by PCR) displayed polydactyly. This chapter describes analysis of the mutants in relation to: morphology of adult and embryonic phenotypes, chromosomal localisation of the transgene, and embryonic expression patterns of the transgenic reporter gene and the endogenous *Shh* gene.

5.1 Anatomical terms used in the phenotypic description

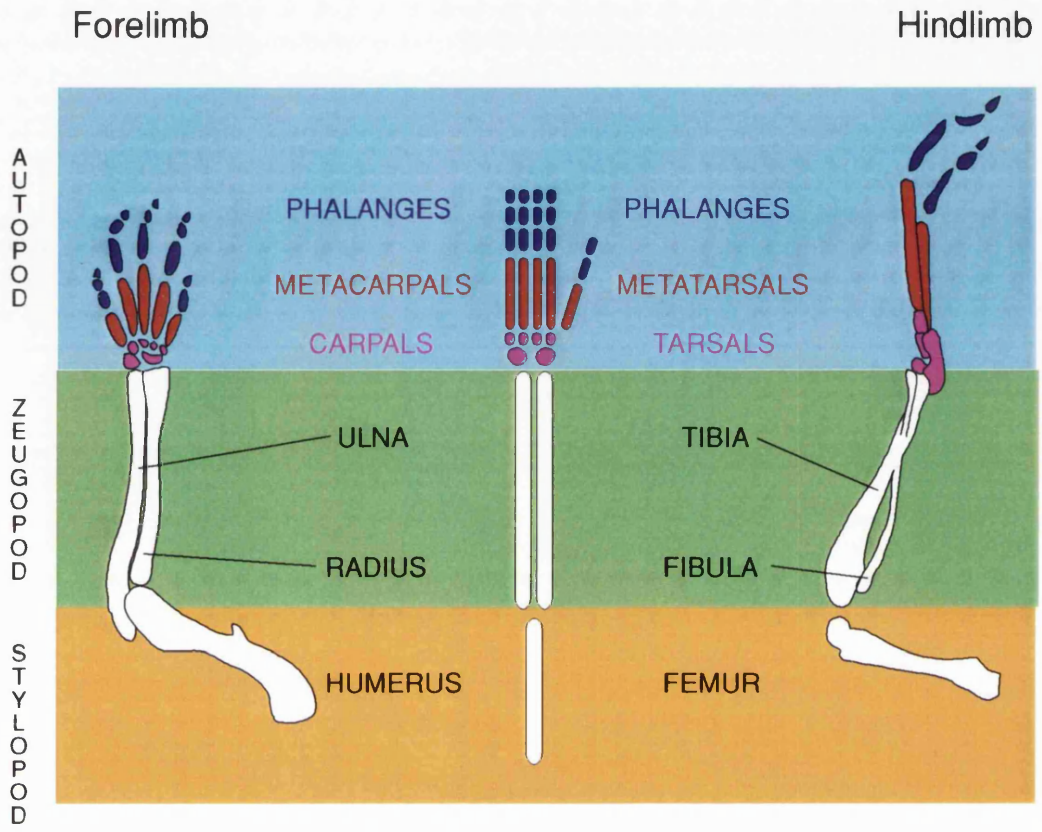
Different authors have used different names for the wrist (carpus) and ankle (tarsus) bones of mice, and this may partly due to the fact that even within the single species *Mus musculus* different strains have different adult bone morphologies (Forsthoefel, 1958). For example, the C57BL/10 strain has two bones in the carpus and two in the tarsus which in most other strains are fused into one. In tetrapod limbs in general (Hinchliffe & Johnson, 1980; Hinchliffe & Griffiths, 1983), these bones are considered as belonging to three rows: the distal row which articulates with the metacarpals or metatarsals (ie. the first elements of the digits), and the central and proximal rows. The distal and central bones are often simply numbered in an anterior to posterior direction (d1-d5 and c1-c4), whereas the proximal bones are given specific names: radiale, intermedium, ulnare and pisiform. However, in the evolution of mammals a significant reduction in the number of bones has occurred, and due to the intrinsic problems of recognising homologous bones, different schemes have described the same mouse bones as belonging to different rows (eg. Forsthefel names three particular carpus bones as: the ulnare, the intermedium/radiale, and d2, whereas Hinchliffe & Griffiths name the same three as c4, c1 and c2/3, thereby indicating both different rows and different fusions).

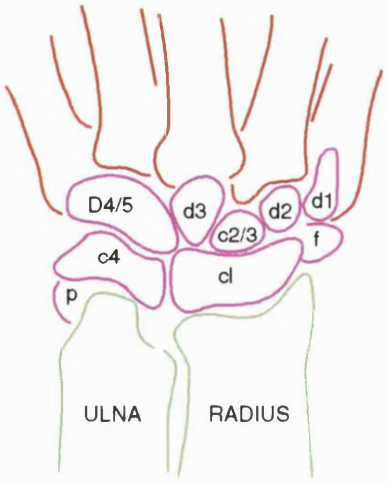
In all subsequent descriptions I have used the nomenclature shown in Fig 5.1. All carpus bones (carpals) are named as either distal or central (d1, d2, d3, d4/5, c1, c2/3, c4). The bones of the tarsus appear to be less evolutionary variable, and more specialised, and so are given specific

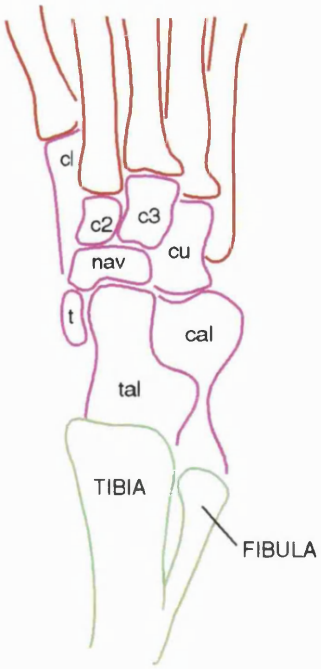
Fig. 5.1 Anatomical terminology of the limb skeleton.

The tetrapod limb is divided into three parts: the autopod (cyan), the zeugopod (green) and the stylopod (orange). Within the autopod are three further divisions: the phalanges (blue), the metacarpals or metatarsals (red), and the carpals or tarsals (magenta).

The mouse carpals are: the *distal carpals* (d), *central carpals* (c), the *pisiforme* (p) and the *falciforme* (f). The mouse tarsals are: the *cuneiformes* (c), the *cuboideum* (cu), the *naviculare* (nav), the *tibiale* (t), the *talus* (tal), and the *calcaneus* (cal).


Forelimb
Hindlimb
AUTOPOD
ZEUGOPOD
STYLOPOD
PHALANGES
METACARPALS
CARPALS
PHALANGES
METATARSALS
TARSALS
ULNA
RADIUS
HUMERUS
TIBIA
FIBULA
FEMUR


D4/5
d3
d2
d1
c2/3
f
c4
cl
p
ULNA
RADIUS


cl
c2
c3
nav
cu
t
cal
tal
TIBIA
FIBULA

names: talus, calcaneus, naviculare, cuboideum, and the cuneiformes numbered from 1 to 3. (The cuneiformes are the distal tarsals, therefore providing the articulation sites for the metatarsals, and the cuboideum can probably be considered as a fusion that contains cuneiforme 4.) The digits are composed of one proximal metacarpal or metatarsal and three phalanges, except for the most anterior digit which only contains two phalanges. The phalanges are numbered P1 to P3, with P1 being the most distal. Traditionally this digit is called the pollex in the forelimb and the hallux in the hindlimb, but in this thesis they will both be referred to as the thumb. Numbering of carpals and digits starts at the anterior side, however, carpals are numbered using arabic numerals, whereas by convention, digits are numbered using roman numerals.

5.2 Generation and analysis of heterozygous mutants

The original founder mouse (a female) of this transgenic line was an F_2 generation from (CBA x C57Bl10)F_1 parents, and she did not display any limb abnormalities. When bred to a wildtype F_1 male, she produced 26 offspring from three litters, of which 4 females and 3 males were transgenic (by PCR). Of these, 1 female (No.669) and 2 males (No.683 and 689) displayed hindlimb polydactyly, and the phenotype of No.669 is shown in Fig. 5.2.

Further heterozygous mice were generated by mating individuals from the F_3 generation to wildtypes. Out of 93 mice, 49 displayed limb abnormalities, and when PCR analysis was carried out to test for presence of the transgene, 13 of the normal mice were shown to be transgenic. Normally this would indicate a phenotypic penetrance of 80%, however, one of the affected mice (No.2646) did not give a positive PCR result, bringing into question whether the mutation was actually caused by the transgene insertion. Data presented in later sections of this chapter provide good evidence that the transgene is indeed the cause of the mutation, and the contradictory data just described can probably be explained in one of two ways: (1) To analyse the hundreds of mice used in transgenic experiments, each mouse is tailed and ear-marked. In this experiment the adjacent mouse (no. 2625) was one of the 13 which tested positive but displayed no phenotype. It is possible that these two DNA samples were accidentally swapped. (2) The transmission rate for the original founder female was 27% (7 positives out of 26 offspring), which is probably due to mosaicism, ie. not all cells of a founder mouse will inherit the transgene. However, her heterozygous offspring should not be mosaic and should therefore have a transmission rate of 50%. Out of a fairly high sample size of 93 mice, the transmission rate as measured by PCR was 66% (61 mice out of 93).

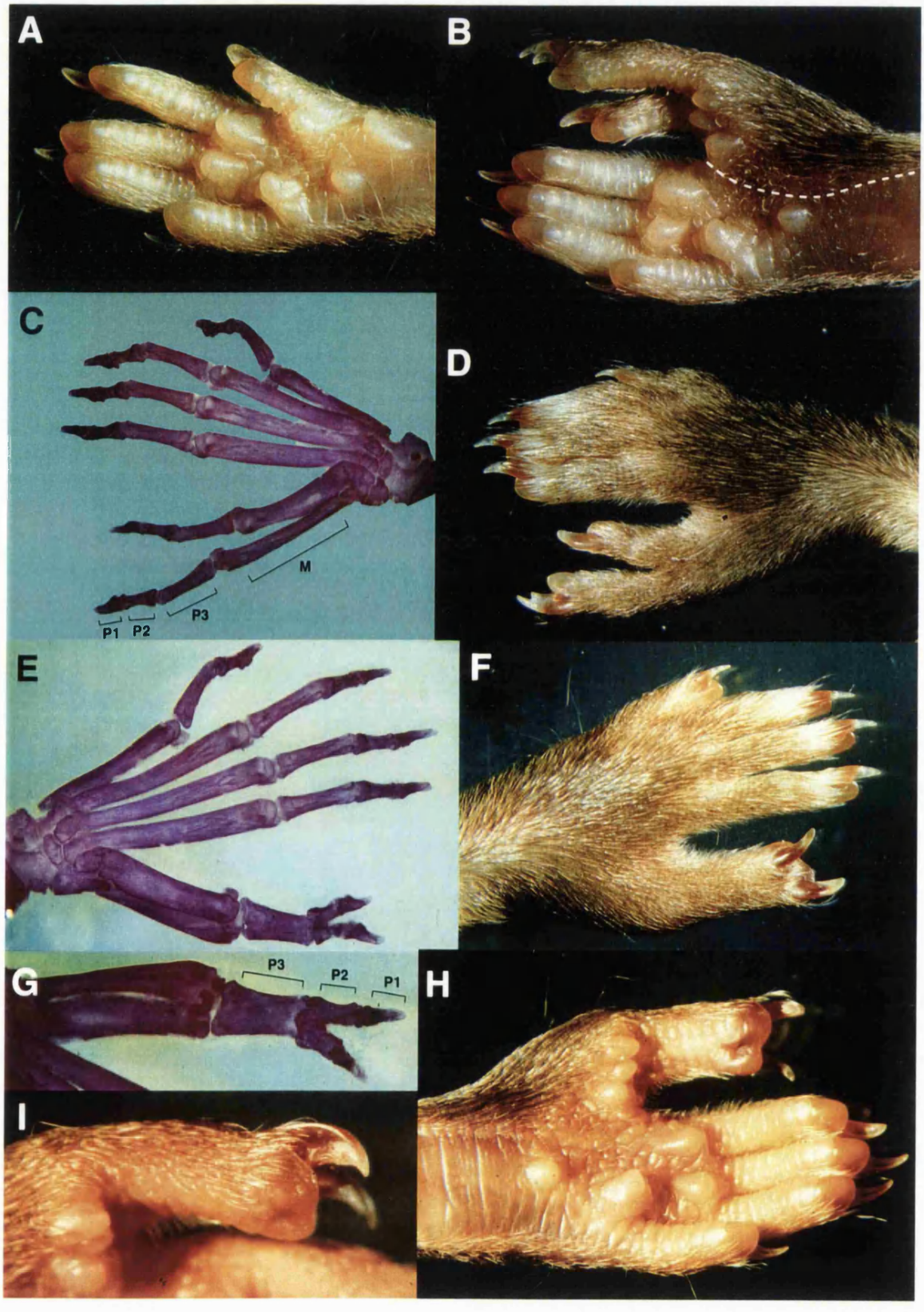
Fig. 5.2 External and skeletal limb phenotype of mouse #669.

(a,b) Comparison of the ventral right hindfoot of a wildtype and a mutant mouse. The dashed line indicates the extent of abnormality in the mutant: the positions of pads below it are normal. As seen in (a) the ventral surfaces of the foot never display thick hair-growth, but in the mutant hair is seen above the dashed line.

(c,d) Skeletal and external phenotype of right hindfoot from mouse #669 (dorsal view). Two triphalangeal digits are seen in the place of the normal thumb. (P1,2 and 3 are the phalanges; M is the metatarsal).

(e,f) Skeletal and external phenotype of left hindfoot from mouse #669 (dorsal view).

(g,h,i) Ventral views of left hindfoot from mouse #669. P3 is totally fused, P2 is partially fused, and P1 is completely duplicated.


A
B
C
M
P1
P2
P3
D
E
F
G
P3
P2
P1
H
I

names: talus, calcaneus, naviculare, cuboideum, and the cuneiformes numbered from 1 to 3. (The cuneiformes are the distal tarsals, therefore providing the articulation sites for the metatarsals, and the cuboideum can probably be considered as a fusion that contains cuneiforme 4.) The digits are composed of one proximal metacarpal or metatarsal and three phalanges, except for the most anterior digit which only contains two phalanges. The phalanges are numbered P1 to P3, with P1 being the most distal. Traditionally this digit is called the pollex in the forelimb and the hallux in the hindlimb, but in this thesis they will both be referred to as the thumb. Numbering of carpals and digits starts at the anterior side, however, carpals are numbered using arabic numerals, whereas by convention, digits are numbered using roman numerals.

5.2 Generation and analysis of heterozygous mutants

The original founder mouse (a female) of this transgenic line was an F_2 generation from (CBA x C57Bl10)F_1 parents, and she did not display any limb abnormalities. When bred to a wildtype F_1 male, she produced 26 offspring from three litters, of which 4 females and 3 males were transgenic (by PCR). Of these, 1 female (No.669) and 2 males (No.683 and 689) displayed hindlimb polydactyly, and the phenotype of No.669 is shown in Fig. 5.2.

Further heterozygous mice were generated by mating individuals from the F_3 generation to wildtypes. Out of 93 mice, 49 displayed limb abnormalities, and when PCR analysis was carried out to test for presence of the transgene, 13 of the normal mice were shown to be transgenic. Normally this would indicate a phenotypic penetrance of 80%, however, one of the affected mice (No.2646) did not give a positive PCR result, bringing into question whether the mutation was actually caused by the transgene insertion. Data presented in later sections of this chapter provide good evidence that the transgene is indeed the cause of the mutation, and the contradictory data just described can probably be explained in one of two ways: (1) To analyse the hundreds of mice used in transgenic experiments, each mouse is tailed and ear-marked. In this experiment the adjacent mouse (no. 2625) was one of the 13 which tested positive but displayed no phenotype. It is possible that these two DNA samples were accidentally swapped. (2) The transmission rate for the original founder female was 27% (7 positives out of 26 offspring), which is probably due to mosaicism, ie. not all cells of a founder mouse will inherit the transgene. However, her heterozygous offspring should not be mosaic and should therefore have a transmission rate of 50%. Out of a fairly high sample size of 93 mice, the transmission rate as measured by PCR was 66% (61 mice out of 93).

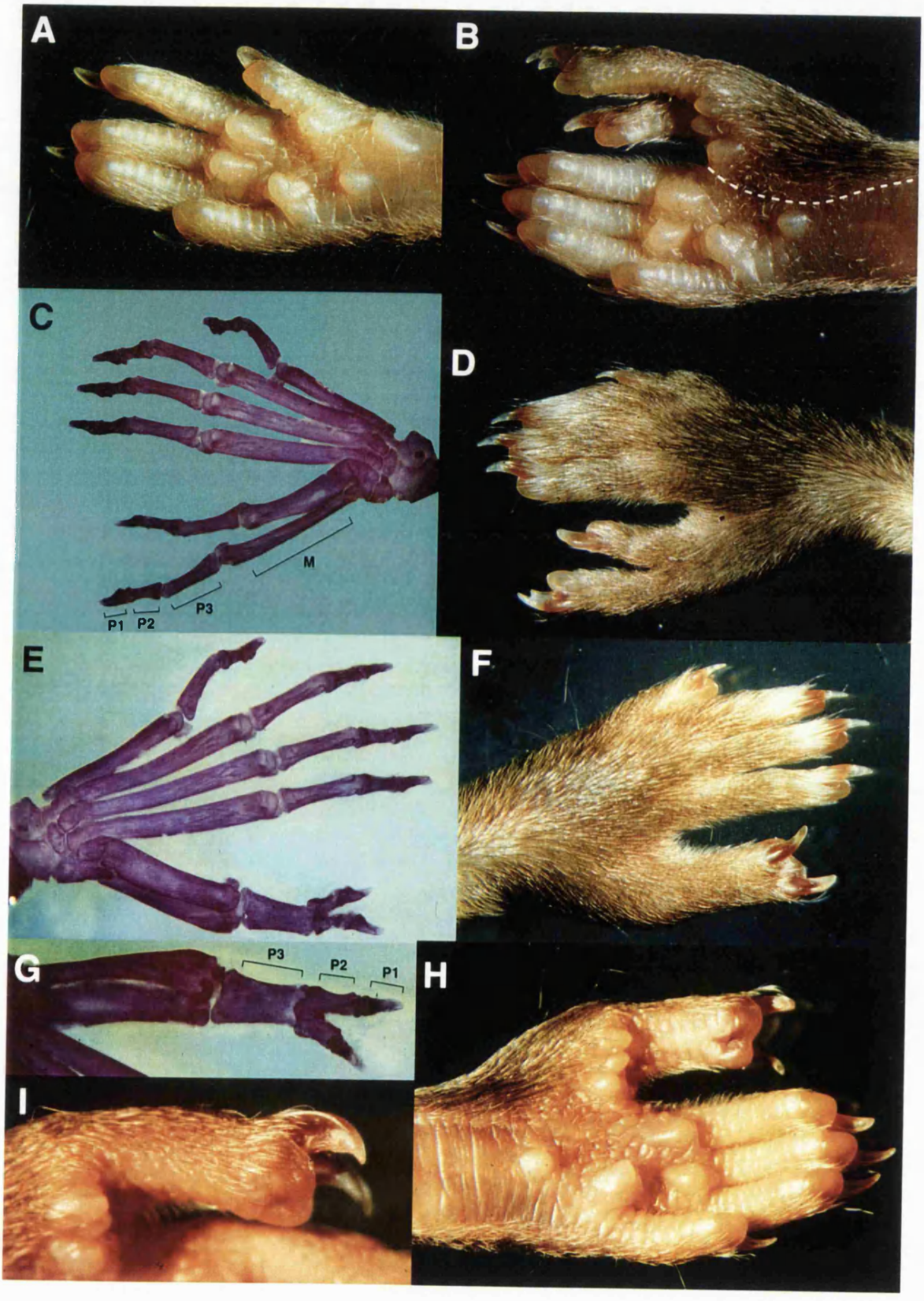
Fig. 5.2 External and skeletal limb phenotype of mouse #669.

(a,b) Comparison of the ventral right hindfoot of a wildtype and a mutant mouse. The dashed line indicates the extent of abnormality in the mutant: the positions of pads below it are normal. As seen in (a) the ventral surfaces of the foot never display thick hair-growth, but in the mutant hair is seen above the dashed line.

(c,d) Skeletal and external phenotype of right hindfoot from mouse #669 (dorsal view). Two triphalangeal digits are seen in the place of the normal thumb. (P1,2 and 3 are the phalanges; M is the metatarsal).

(e,f) Skeletal and external phenotype of left hindfoot from mouse #669 (dorsal view).

(g,h,i) Ventral views of left hindfoot from mouse #669. P3 is totally fused, P2 is partially fused, and P1 is completely duplicated.


A
B
C
M
P3
P2
P1
D
E
F
G
P3
P2
P1
H
I

This may represent a low-frequency systematic error in the PCR analysis, which would equally explain the "false-negative" as well as an unknown number of "false-positives". However, if this mutant can be considered a member of the hememelia-luxate group, it would be expected to have an incomplete penetrance (see section **1.9.1**), so the percentage of PCR positive mice would be expected to be higher than the percentage of polydactylous mice (which is 53%). This suggests that the PCR analysis is probably accurate, and that human error is to blame.

5.2.1 The adult heterozygote phenotype

Fig. 5.2 shows the hindfeet of female No.669, and the right hindfoot (panels a-d) displays a typical heterozygote phenotype. Triphalangy of the original thumb has occurred (ie. there are three phalangeal bones instead of two), and also an additional triphalangeal digit has developed preaxially. The fact that the supernumerary digit is preaxial is most clearly seen from the external phenotype: not only do the size and shape of digits II, III, IV and V appear normal, but the associated pads (toughened, raised areas of skin) also maintain their wildtype patterning (compare Fig. 5.1a with 5.1b). The skeletal preparation also shows that the digit which is in place of the thumb, despite being triphalangeal has nevertheless retained some thumb-like characteristics: it is shorter than digits II-IV, and it protrudes away from them at a slight angle.

The left hindfoot (Fig. 5.2e-i) externally appears less affected than the right hindfoot, as if the thumb has developed thicker and longer than usual with two claws. The skeleton however, reveals a similar structure to the right hindfoot, with the difference that the third phalange of the two anterior digits is fused into one bone. The metatarsals and more distal phalanges of these digits are however, clearly not fused.

The phenotype of heterozygotes is extremely variable, even within a litter. It is essentially composed of two effects: production of up to two extra digits and triphalangy of the thumb, but these extra skeletal elements are often fused to their neighbours or only partially formed. For example one of the weakest phenotypes can be described as a partial double thumb (Fig. 5.3a,b). In this situation proximal elements appear completely normal, and only distal elements are duplicated. Alternatively, an almost complete extra digit can be formed but is fused to its neighbour at the level of metatarsal (Fig. 5.3c) or proximal phalange (Fig. 5.2g). Duplication always occurs before triphalangy of the thumb, and consequently the weakest phenotype observed is a double-thumb.

Fig. 5.3 Variation in the heterozygous limb phenotype.

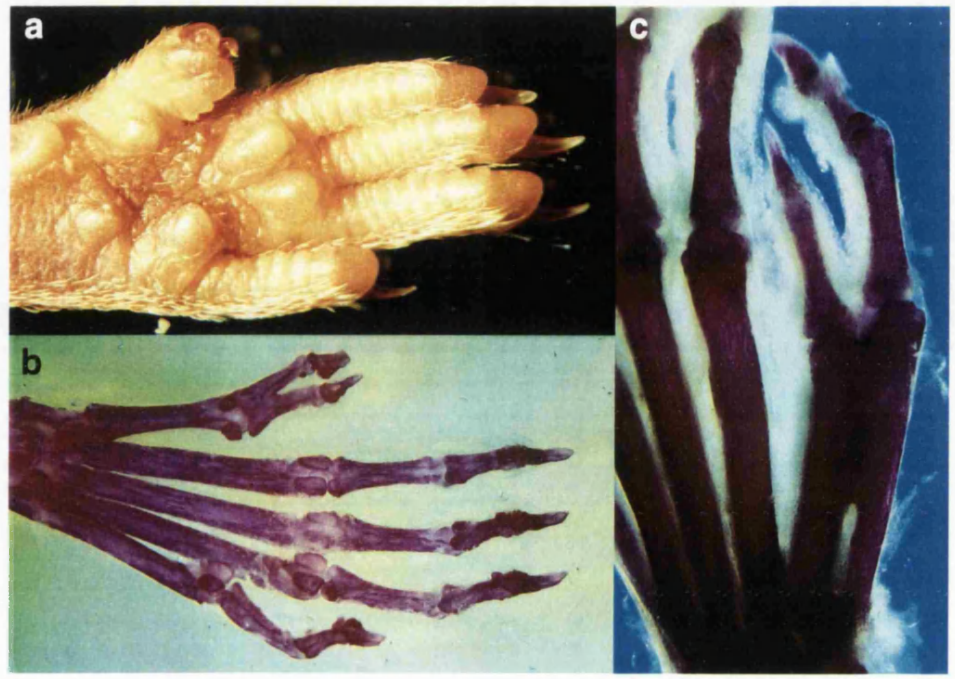
(a,b) Ventral external and skeletal views of the rare 'partial double thumb' phenotype, in which the thumb metatarsal is normal, but the second phalange is partially duplicated, and the first phalange is completely duplicated.

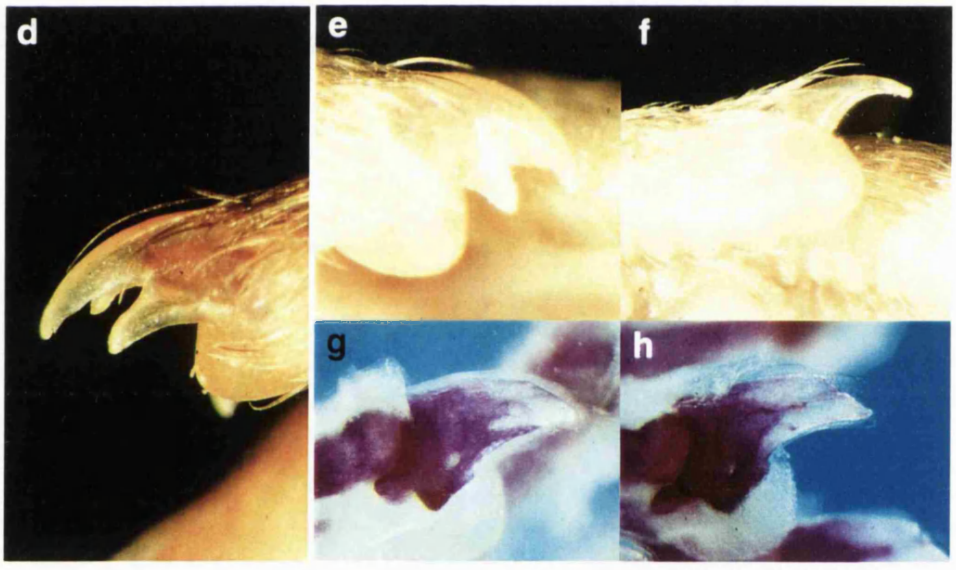
(c) Dorsal skeletal view of an extra triphalangeal digit which is slightly fused at the metatarsal level to the thumb. In this case the thumb has remained biphalangeal.

(d-h) In a few cases, duplicated claws were seen on the ectopic digits, (d) and (e). A normal claw is shown in (f).

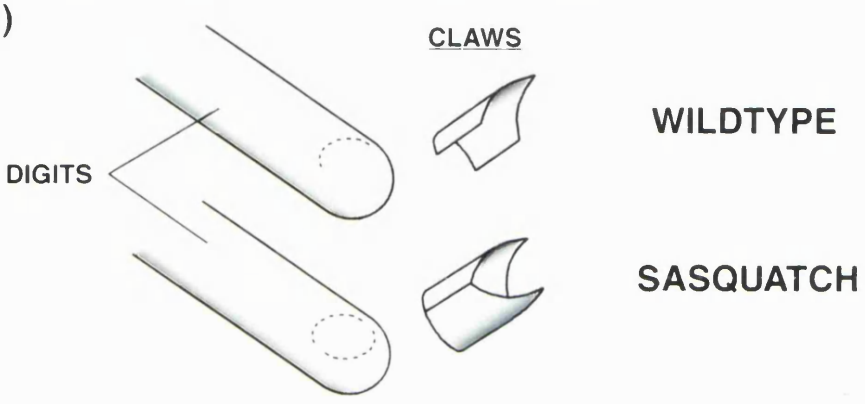
(g) and (h) display the first phalange for the digits in (e) and (f), and show that the abnormal claw structure was not reflected in an abnormal skeletal structure.

(i) Diagram showing how the duplicated claws may be the result of a mirror-image duplication along the proximodistal axis, resulting in a tube-like structure.





(i)



Fusion of extra digits can occur with the thumb or with a neighbouring extra digit, but digits II-V are never fused.

Another abnormality is occasionally observed in these mice: an abnormal claw on the some of the ectopic digits (Fig. 5.3d,e). It appears as a double-claw structure, with a second claw growing underneath (ventral) to the natural one. However, the orientation of the transverse curvature of the ectopic claw is reversed with respect to the normal claw, suggesting that a mirror-image duplication along the proximodistal axis has occurred. In this situation the lateral edges of the two claws fuse, creating a tube-like structure. The normally-orientated claw still curves distally and convexly, and this causes the duplicated claw to curve also distally and therefore concavely. Comparison of the first phalanges from digits displaying normal and abnormal claws revealed no significant difference (Fig. 5.3e-h), suggesting that this abnormality is not the knock-on consequence of skeletal changes.

5.2.2 Development of polydactylous limbs

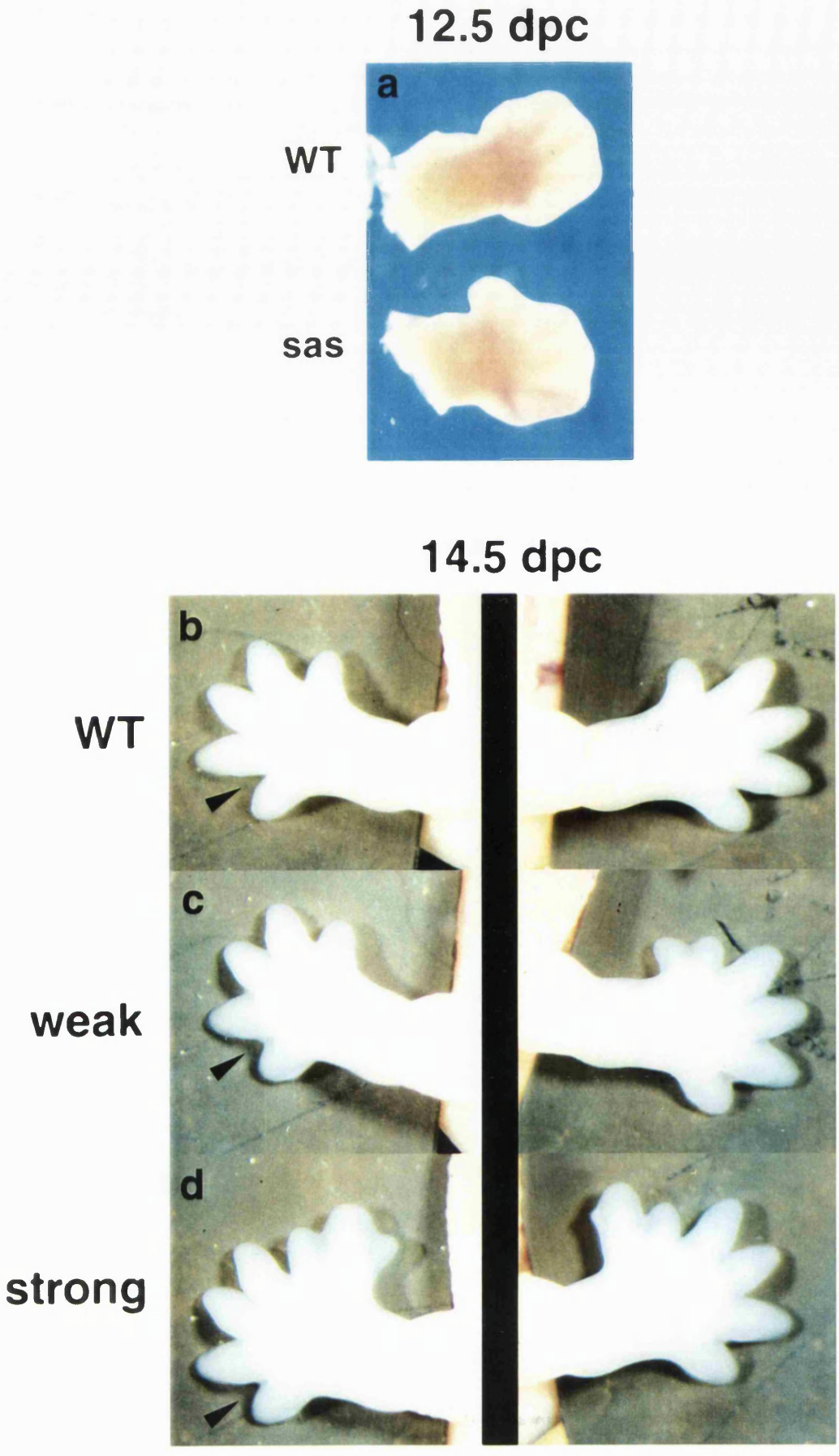
The earliest sign of abnormality is seen at 11.75dpc as a very slight swelling on the anterior side of the limb bud. By 12.5dpc this is more clearly seen as an obvious localised outgrowth of tissue in the region where the thumb would normally develop from (Fig. 5.4a). By 14.5dpc, the digits can be morphologically distinguished from each other, and the polydactyly and its severity can be seen. The three pairs of hindlimbs in Fig. 5.3b-d, are from one wildtype and two heterozygous embryos from the same litter. The phenotype variability can clearly be seen. The embryo in (c) has no obvious defect in its left hindlimb and only a probable double-thumb in its right hindlimb. The left hindlimb in (d) would develop at least one extra triphalangeal digit with a partially duplicated extra digit anteriorly to it, and possibly a triphalangeal thumb. (The large extra digit in this case may in fact condense into two very fused digits, as it appears to be very thick and slightly squared at the distal tip.) The corresponding right hindlimb shows a similar but slightly reduced phenotype, with only a tiny piece of tissue anteriorly to the extra digit.

Although the left hindlimb in panel (c) is normal with respect to the digit pattern, it displays a subtle phenotype which is present in all mutant developing limbs. The interdigital gaps appear to be slightly more webbed than normal (see arrowheads). This is suggestive of reduced cell death, which may indicate a similarity with other hemimelia-luxate mutants (see discussion, section **1.9.1**).

Fig. 5.4 Development of heterozygous *sasquatch* limbs.

(a) View of a wildtype and mutant limb at 12.5dpc.

(b-d) Left and right hindlimbs of three 14.5dpc embryos which show the wildtype, weak and strong *sasquatch* phenotype. There is a correlation between the severity of phenotype of limbs on the left and right sides of each embryo. The black arrow indicates the absence or presence of slight webbing between the posterior-most digits.


12.5 dpc
a
WT
sas
14.5 dpc
b
WT
c
weak
d
strong

5.2.3 Statistics of heterozygous phenotype

Not all polydactylous animals were studied by skeletal preparation, so in order to determine a rough idea of variability of phenotype, and whether any differences exist between the sexes or between the left and right sides, a system was devised for scoring external phenotypes.

Since in heterozygotes digits II-V were never affected, the notation describing each mutant foot relates only to digits produced in place of the thumb. Each digit was described as short, medium or long, and was given a score of 1, 2 or 3 respectively. These were later reliably analysed to be biphalangeal, short triphalangeal or long triphalangeal. If two or three adjacent digits appeared to be significantly fused (or only partially duplicated), the score of the combined digits was reduced by a third. In this way a range of scores was obtained, in which 1 represents a wildtype foot, and 4 represents the most common phenotype of one large extra triphalangeal digit (in addition to the thumb).

The graphs in Fig. 5.5 give an idea of the variabilty found. There is an obvious correlation between the left and right sides of an individual animal, although a few cases did occur with a large difference between the sides. Also, there is no significant bias towards one side or the other - 15 are more severely affected on the left foot, 17 on the right foot, and 17 are equally affected on both feet. The female mice have a higher average phenotypic score than males (3.81 as opposed to 2.81), which is >99% significant using the Wilkoxon Rank Sum Test. They also appear to have slightly less variation than the males.

Interestingly, the distribution of phenotype severity appears to be bimodal, with values 1 and 4 being much more common than 2 or 3 (Fig. 5.5). Although the scoring system devised is quite arbitrary, and may contain some non-linearities, it is nevertheless the case that those phenotypes represented by the values 2 and 3 (which includes all "double-thumbs") is very rare. It therefore seems that a feedback mechanism is directing all *slightly* abnormal limb-buds either back towards normality, or towards a "typical" polydactyly of severity 4.

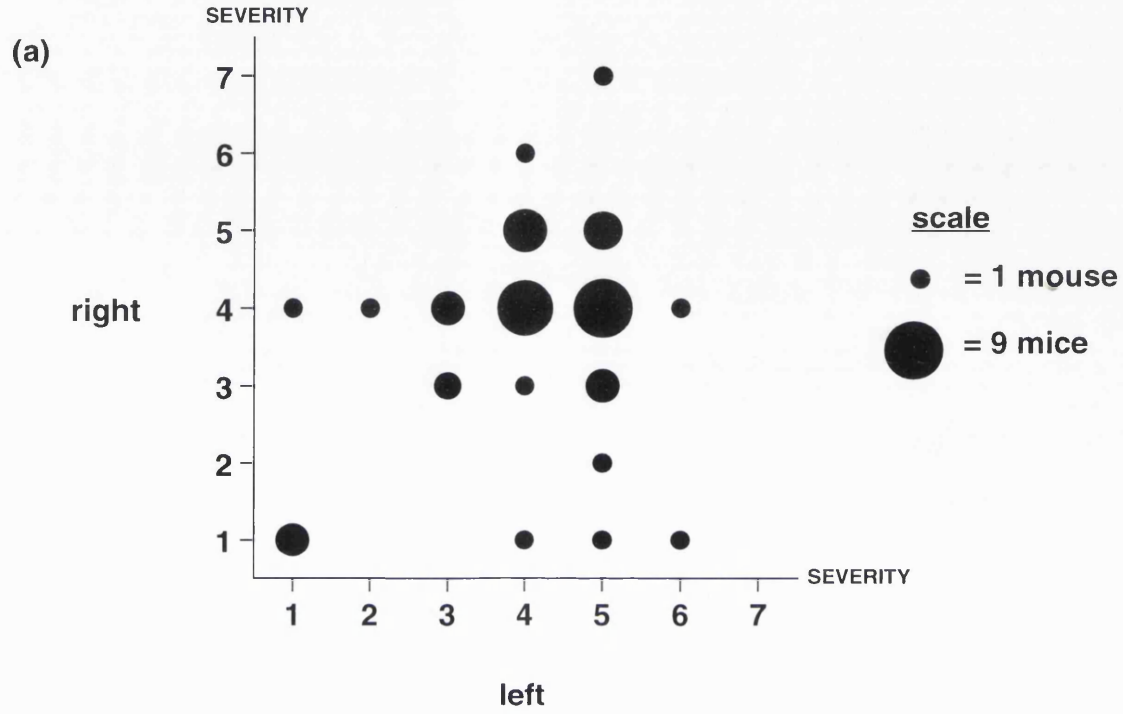
5.3 Phenotype of homozygous mice

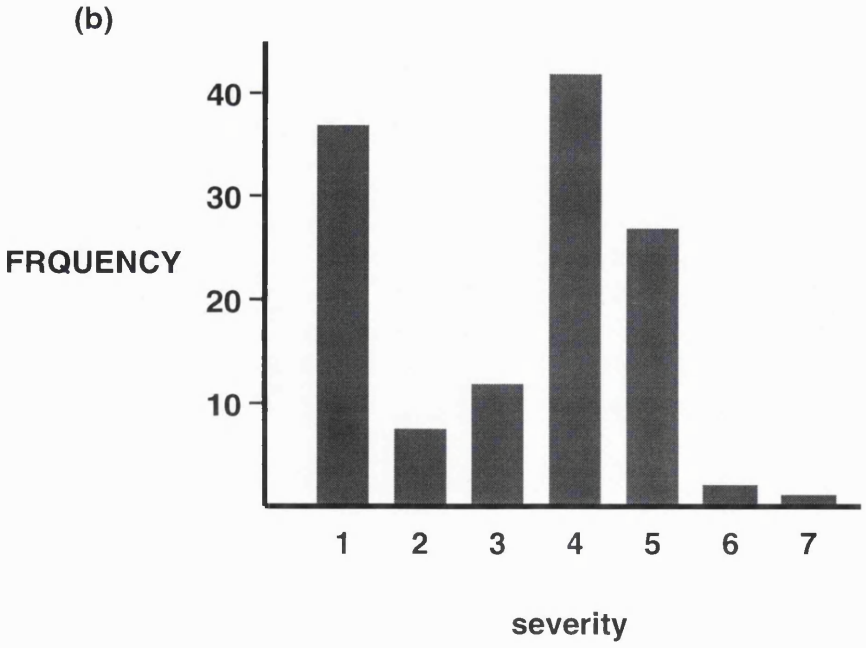
When heterozygous mice were bred together a new, more severe phenotype was observed. These mice were considered to be homozygous for two reasons: (1) The heterozygous phenotype though variable, remained within the limits described above. Certain features of the new phenotype

Fig. 5.5 Statistics of the heterozygous phenotype.

(a) A plot of phenotype severity for left limbs against right limbs of 62 mice positive by PCR for the b1-A transgene. The area of each dot is proportional to the number of mice with that phenotype. Although a few mice occured with only one limb strongly affected, in general there was a correlation between the two sides. (A value of 1 = wildtype).

(b) A histogram showing the frequency of the 7 different severity categories. Severities of 6 or 7 are rare which suggests that an upper limit is being approached. The lower limit for severity is a completely wildtype foot (value 1), however, the intermediate values of 2 or 3, are much rarer than either a value of 1 or the common phenotype of 4 or 5, thus showing that the distribution appears to be bimodal (discussed in section **5.2.3**).


(a)
SEVERITY
7
6
5
4
3
2
1
right
SEVERITY
1
2
3
4
5
6
7
left
scale
= 1 mouse
= 9 mice


(b)
40
30
20
10
FRQUENCY
1
2
3
4
5
6
7
severity

were never seen in mice known to be heterozygous. (2) The new phenotype occurred in less than 25% of the offspring from double heterozygote matings. Unequivocally determining whether these mice were homozygous was only possible in one case, because PCR or Southern blotting are not quantitatively reliable enough, and the majority of the severely affected mice could not mate. However, one female of the strong phenotype did produce a litter, and all the offspring were shown to be transgenic.

5.3.1 The digits

The most obvious change from the heterozygous phenotype was the presence of polydactyly in the forelimbs (although as mentioned above, this does not prove that all homozygotes display this feature). The skeleton of a wildtype forelimb possesses five digits (4 triphalangeal and one biphalangeal), but externally the thumb is reduced to a small pad without a claw (Fig. 5.6a). In homozygote mutants, polydactyly of the forelimb can produce anything from 5 to 7 triphalangeal digits (or partial digits) with claws (Fig. 5.6b). Fig. 5.6c-e show external and skeletal views of the same forelimb. In this case the thumb has remained biphalangeal, and an extra triphalangeal digit has developed preaxially to it. In addition, a small pad of tissue has developed attached laterally to the ectopic digit tip. Its morphology appears consistent with its distal position (although it is not growing a claw), and it possesses no skeletal component. Another feature seen for the first time in homozygotes is a posterior extension of the abnormal zone so that it included digit II. In heterozygotes digits II-V are always unaffected, whereas one homozygote mouse was found in which the original digit II is split at the level of the 3rd phalange, into two digits (Fig. 5.7c).

The hindlimbs of the strong phenotype are consistently more affected than those from heterozygotes. Biphalangeal digits are never seen, the typical pattern is 7 triphalangeal digits. As with heterozygotes, fusions and partial digits are common, especially in the metatarsal region. A typical homozygous hindlimb is shown in Fig. 5.6j, which displays fusion of the 2nd and 3rd metacarpals. Panels (f-i) show external and skeletal views of another hindlimb in which a very abnormal metatarsal has developed, as well as digit fusion of the 2nd and 3rd phalanges of the first two digits. The unusual shape of the large metatarsal (h) demonstrates the difficulty of understanding how these abnormal condensations are controlled. This bone could either be thought of as a single, straight metatarsal which grew two extra processes: one distally and anteriorly (labelled A), and one towards the cuneiformes and posteriorly (labelled B), or alternatively as two

Fig. 5.6 Limb phenotype of homozygous mice.

(a) A wildtype forelimb paw, which posesses four external digits (with claws) despite having five skeletal digits (see Fig. 5.7b). The thumb digit is reduced to a pad.

(b) A right forelimb (from the dorsal side) which posesses six digits. The anterior two digits (upper-most in the photo) are the two extra ones.

(c-e) External and skeletal views of one left forelimb which posesses two extra digits (on left-hand side of photo), and has an extra pad of tissue attached to the anterior side of the ectopic digit (e).

(f-i) External and skeletal views of a left hindlimb which displays seven digits, and an extremely distorted metatarsal. The labels A and B indicate two lateral projections from the metatarsal (described in detail in the text) and the dashed line indicates the shape of the B extension which has not yet completely ossified. In panel (i) it can be seen that although the metatarsals and P1 phalanges are separate, the P2 and P3 phalanges are fused.

(j) The skeleton of the right hindlimb from the same mouse as in (f-i).


a
b
c
d
e
f
g
j
h
B
A
i

very curved metatarsals which became fused in the middle. Interestingly, the second, incompletely ossified extension (B) appears to be attempting to articulate with the second cuneiforme despite the obstacle of the metatarsal from digit IV.

It is a common feature of these limbs that the second metatarsal exhibits signs of duplication or fusion. A small protrusion on the posterior side of the proximal end, similar to (B) but much smaller, often occurs irrespective of whether there is a duplication at the distal end (see Fig. 5.7 panels (g) and (i), white arrowheads).

5.3.2 The carpus and tarsus

Fig. 5.7 shows the left autopod and right tarsus of two typical homozygotes and one wildtype mouse (named WT, Hom1 and Hom2). In the carpus from Hom1, d2 is fused with c2/3, the falciforme is absent, and an ectopic distal carpal-like bone is present at the anterior base of the abnormal metatarsal. Polydactyly in this case is thus an abnormality mostly restricted to the digits. In Hom2, the degree of polydactyly is greater - a total of 7 digits distally, created from 6 metacarpals (due to a split digit) - but alterations in the carpus are only slightly increased. The whole carpus is slightly compressed along the proximodistal axis, and this causes the bones c1 and d4/5 to be pushed apart from each other. c2/3 is enlarged into a very abnormal shape, and like Hom1 there is an additional distal carpal-like bone at the base of the two anterior-most metacarpals.

In wildtype hindfeet, unlike the forelimb situation, there is a clear correspondence between the metatarsals and the tarsals with which they articulate: metatarsals I, II and III associate closely with cuneiformes 1, 2 and 3, and metatarsals IV and V associate with the cuboideum. The junctions between these bones occur at different levels along the proximodistal axis for each pair (see panel h) and each of the distal tarsals appears quite different in size. The common feature of the homozygous phenotype is that the cuneiforems, of which there are now 4 instead of 3, become similar in size and shape, resulting in a uniformity across the AP axis (metacarpal articulation sites are in a fairly straight row). The naviculare which usually sits proximally to only cu2 and cu3, is fused with the tibiale, and is now adjacent to all 4 of the cuneiformes. The cuboideum is not significantly altered.

The overall appearance of the homozygous tarsus is shorter and broader than the wildtype, and this is the combined result of smaller cuneiformes, a slightly shorter talus, and more distal elements across the AP axis.

Fig. 5.7 The carpus and tarsus of homozygous mice.

Dorsal, skeletal views of the carpus and tarsus of two homozygous mice and one wildtype (described in detail in the text).

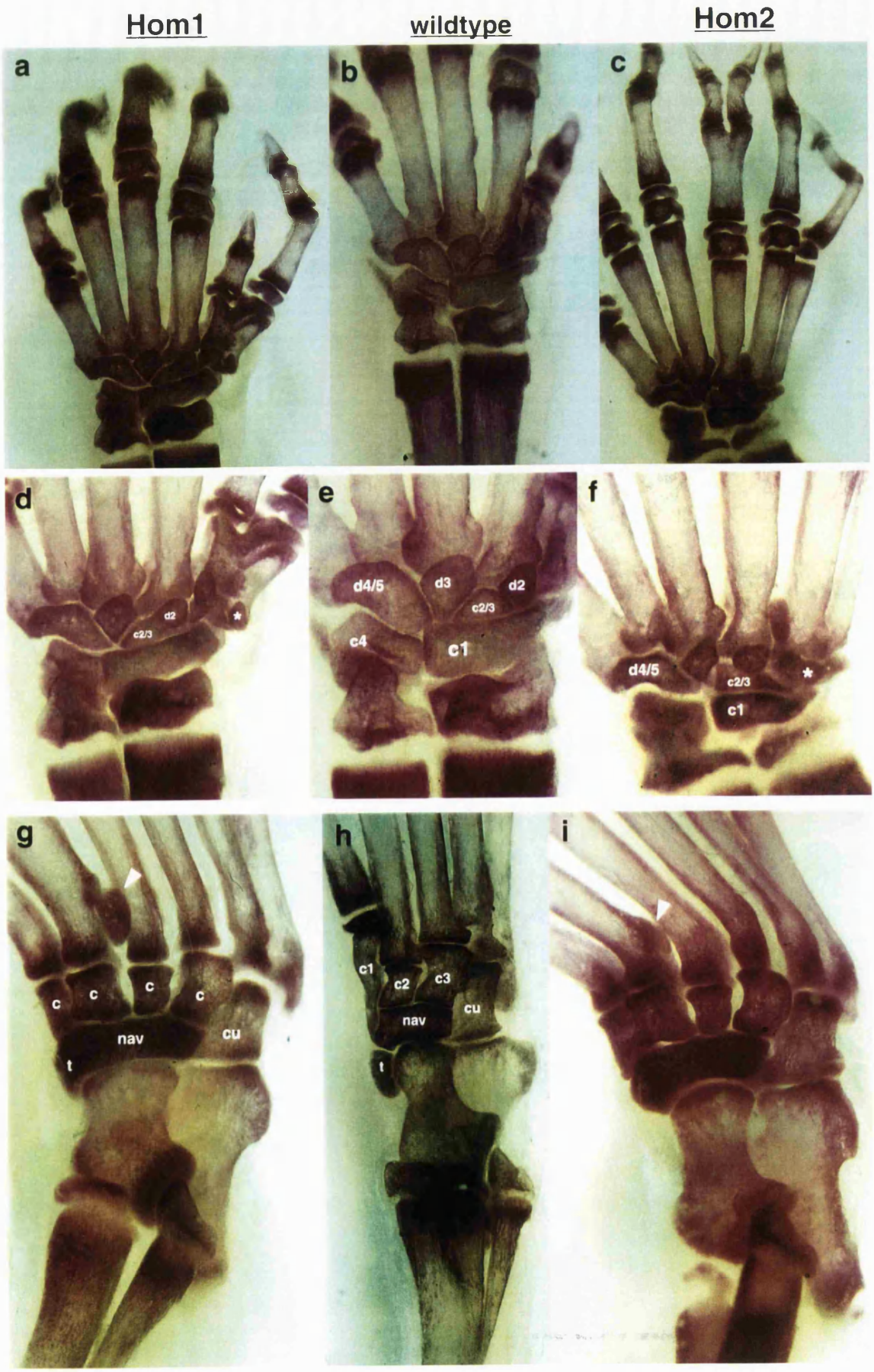
(a-c) Left forelimbs.

(d-f) Close-up of forelimbs.

d2, d3, d4/5 = distal carpals; c1, c2/3, c4 = central carpals; * = ectopic bone.

(g-i) Right hindlimbs. The white arrowheads indicate the lateral/posterior extension common on the proximal end of mutant metatarsals (see end of section **5.3.1**).

c1, c2, c3 = specific wildtype cuneiformes; c = unspecific mutant cuneiforme; nav = naviculare; cu = cuboideum; t = tibiale.


Hom1
wildtype
Hom2
a
b
c
d
d2
c2/3
*
e
d4/5
d3
d2
c2/3
c4
c1
f
d4/5
c2/3
*
c1
g
c
c
c
c
nav
cu
t
h
c1
c2
c3
nav
cu
t
i

5.3.3 The zeugopod

In addition to abnormalities in the autopod, homozygotes show alterations of the zeugopod, which is much more significant in the hindlimb than the forelimb. Fig. 5.8b shows the left forelimb of a homozygote (top) compared to a wildtype (both aged 20 days). The scapula and humerus are not significantly different, but the zeugopod of the mutant is shortened. As with all skeletal features of this mutant, this shortening varies even between the left and right zeugopod of the same animal (Fig. 5.8c). Both elements of one zeugopod are similarly affected in all animals studied, but often the radius is very slightly shorter than the ulna resulting in a slight inward twisting of the autopod.

Fig. 5.8d shows the right hindlimbs of a homozygote (top) and a wildtype (both aged 20 days). In wildtype mice the tibia is a large bone and the fibula is thin and fused to it, whereas in the mutants these two bones are more distinct and occasionally completely unfused. Compared to the wildtype, the mutant tibia is shorter, thinner at the proximal end, and bent dorsally, and the mutant fibula is thicker and follows a similar path. One consequence of these abnormalities is that the distal ends of the two bones are in different positions relative to each other - whereas they are usually located anteriorly and posteriorly (tibia and fibula respectively) in the mutant they are dorsal and ventral. This means that the plane of the autopod may be twisted through an angle of up to 90° (as seen in Fig. 5.8d). The severity of this distortion can affect the ability of these animals to walk.

5.4 Comparison of the *sasquatch* phenotype with hemimelia-luxate mutants

The phenotype of the *sasquatch* mutation displays all the characteristics in the limb phenotype typical of the hemimelia-luxate group: preaxial polydactyly, reduction and abnormalities of the zeugopod, over-growth of mesenchyme tissue early in development, extreme variability of the phenotype, and a gradation of severity which affects distal elements more readily than proximal elements, and hindlimbs more readily than forelimbs. However, the most important conclusion from a comparison of limb phenotypes is that *sasquatch* is clearly not identical to any of them. Additionally, no abnormalities could be found in the rest of the skeleton (Fig. 5.8a).

Since *sas* may be a new allele of an old mutant, it is useful to determine which of these it is most similar to. The most useful feature for this distinction, is the observation that forelimb polydactyly occurs only in homozygous mice. This is clearly not the case for *Dh*, *lx*, and *Xpl* which never display forelimb polydactyly, nor is it seen in *Hx* or *Xt* which can exhibit forelimb polydactyly

Fig. 5.8 General skeletal morphology of homozygous mice.

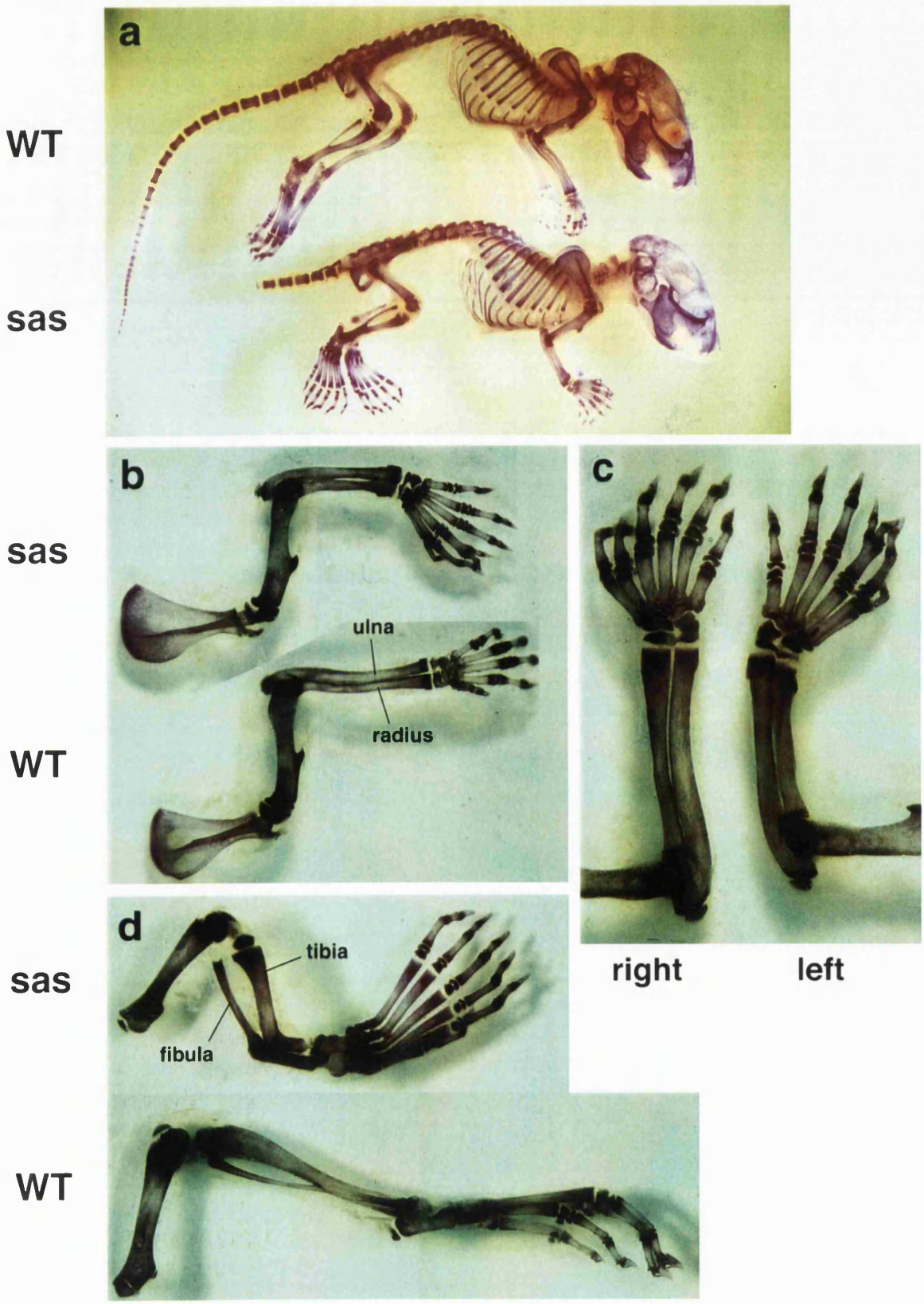
Skeletal elements of homozygous and wildtype mice:

(a) Whole skeleton.

(b) Comparison between a homozygous forelimb (top) and a wildtype forelimb. Although both the ulna and radius are shorter in the mutant, the radius has been reduced more than the ulna, which may be the cause of the slight twisting of the paw. This fits in with the general observation that defects are concentrated on the anterior side of the limb.

(c) Comparison between the left and right forelimbs of a homozygous mouse, which shows that variation of this phenotype can occur within individuals.

(d) Comparison between a homozygous hindlimb (top) and a wildtype forelimb. The distortions of the tibia and fibula have resulted in their distal ends being skewed from their normal orientation, such that the tarsus and foot as a whole is now twisted through 90°.


a
WT
sas
b
sas
ulna
radius
WT
c
right
left
d
sas
tibia
fibula
WT

in heterozygotes, but this is the situation found in *lst*, *lu* and *Rim4*. Due to the variable nature of the phenotype and the underlying similarities between these three mutants, a more precise idea of which is most similar to sasquatch is hard to obtain. However, there are differences which may be relevant: *lst* displays more severe polydactyly than *sas* or *lu*, but *lu* sometimes displays ectrodactyly, which has not been seen in any *sas*, *lst* or *Rim4* mutants. (The published description of *Rim4* is not as detailed as for *lst* or *lu*.)

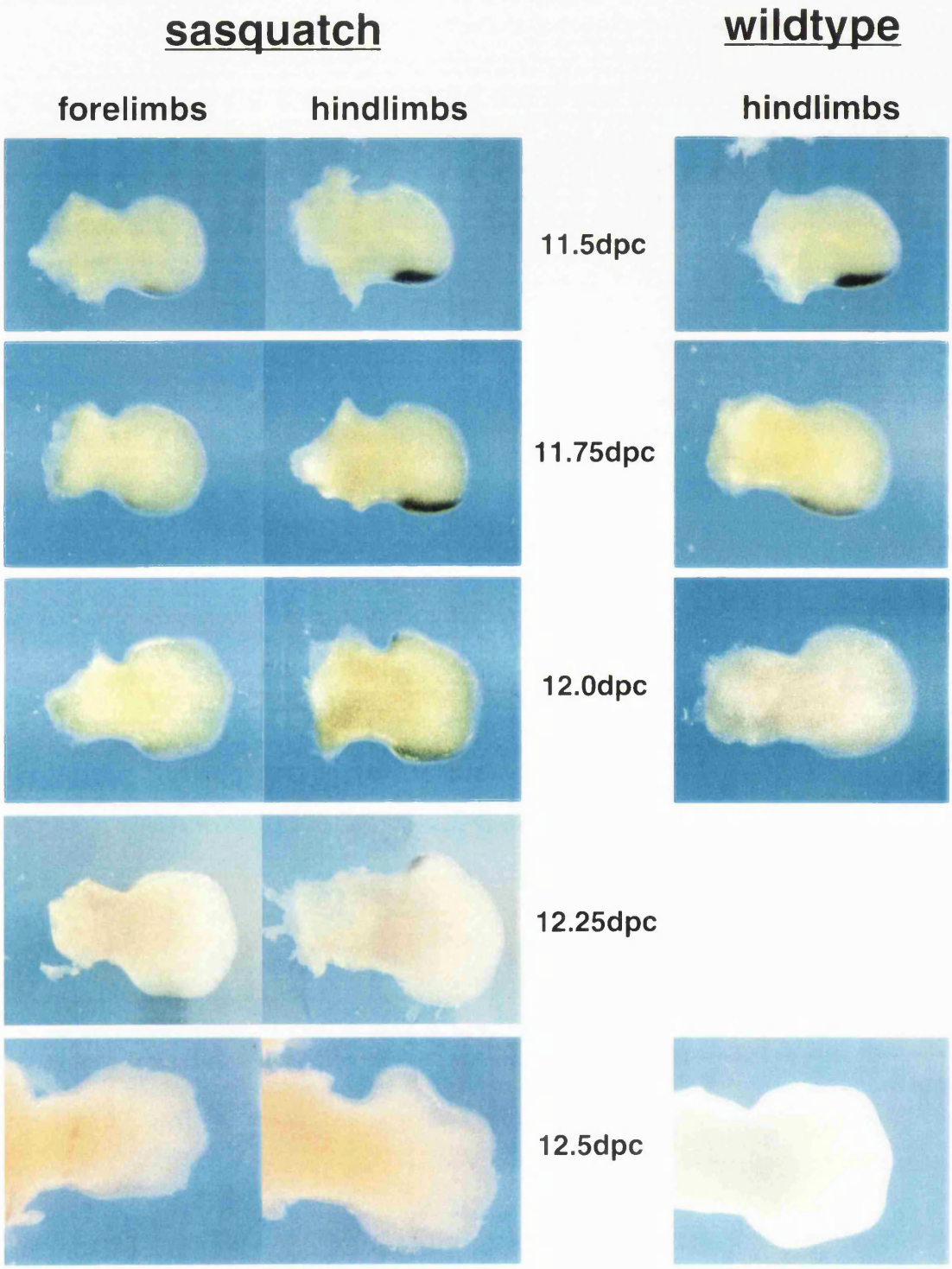
The only other phenotypic character of *sas* which may give a clue to potential similarities, is the slightly webbed interdigital gaps. If this does indicate a reduction in programmed cell death, it could suggest a connection with the *Hm/Hx* mutants, as it has been proposed that both of the phenotypes from this locus may be due to a lack of cell death (see section **1.9.1**). However, the similarity of phenotypes within the hemimelia-luxate group, is probably due to common underlying mechanisms, and the cell death program may therefore be similarly affected in all cases. The presence of slight webbing during embryogenesis, as seen in *sas*, has not been reported for any of the other mutants.

5.5 Expression of *Shh* during limb development

Shh is known to be a good molecular marker for the ZPA, and is probably the signalling molecule responsible for its function. Whole-mount *in-situ* analysis of *Shh* has been performed for four of the previously described hemimelia-luxate mutants (*lst*, *Rim4*, *Hx*, *Xt*) and in each case an anterior ectopic domain of expression was found (see introduction, section **1.9.3**). The same analysis was performed for *sasquatch* mutant limbs of varying ages (a cDNA of mouse *Shh* was kindly provided by A. McMahon). At 11.5dpc, when the phenotypic bulge is only just visible, the expression of *Shh* appears completely normal (Fig. 5.9). Over the next 24 hours *Shh* expression in the ZPA gradually fades away, and it has completely disappeared by 12.5dpc. During this period the ectopic bulge grows, and expression of *Shh* is seen in the proximal half. At about 12.0dpc expression can be seen in both the normal ZPA and the ectopic bulge, and by 12.25dpc it can be seen only in the bulge. At 12.5dpc expression of *Shh* has completely ceased. The ectopic *Shh* expression is therefore clearly a later event than the normal ZPA expression, with both its initiation and its down-regulation occurring after that of the ZPA. Indeed, at the stage when normal *Shh* expression is strongest, expression of the ectopic *Shh* cannot be found. These data therefore reinforce the hypothesis that *sasquatch* is a member of the hemimelia-luxate group.

Fig. 5.9 Expression of *Shh* in developing limb buds.

Whole-mount *in-situ* hybridisations of a probe for *Shh*. Quater day time-points were taken from the time before abnormalities are seen, until expression of *Shh* is no longer visible. The pattern in *sasquatch* and wildtype hindlimbs appeared the same at 11.5dpc and 11.75dpc, with expression only in the ZPA. At 12.0dpc, the anterior phenotypic bulge was present in the mutant, and a small patch of *Shh* expression was seen within it. By 12.25dpc, *Shh* expression in the ZPA had disappeared in both the mutant and wildtype limb-buds, but was still visible in the ectopic anterior bulge. By 12.5dpc expression of *Shh* was absent from all limb-buds tested.


sasquatch
wildtype
forelimbs
hindlimbs
hindlimbs
11.5dpc
11.75dpc
12.0dpc
12.25dpc
12.5dpc

5.6 Chromosomal localisation of the transgene

From the results described above *sas* could either represent a new locus, or could be a new allele of an old mutant. Since the genes responsible for most of the hemimelia-luxate mutants are not cloned, the quickest way to determine whether *sas* was located in a site already associated with polydactyly, was to perform fluorescent *in-situ* hybridisation (FISH) on metaphase chromosome spreads. The 9kb b9-A construct was used as the probe, and cells were taken from the spleen of heterozygous animals. The FISH was performed by Margret Fox from the Department of Genetics at UCL.

As can be seen in Fig 5.10, only one chromosome displayed hybridisation, and from the G-banding pattern this can be identified as chromosome 5. However, there are always two sites of hybridisation on each chromatid, which appear approximately 5 to 10 cM apart. This is an unusual result for random integration of DNA. It suggests two scenarios: (1) Integration occurred at one site, but the middle of the transgene array then became one of the breakpoints for a chromosomal inversion. The inversion therefore carried half of the copies to a new site. (2) The two sites represent two independent insertion events. These two alternatives should be distinguishable by whether recombination through breeding can separate the two hybridisation sites. An inversion would prevent such a recombination event from occurring.

Although there is little data to reliably translate physical chromosomal maps onto genetic maps (Lyon and Searle, 1989), it is clear that the approximate region indicated by this FISH analysis includes two of the hemimelia-luxate group, *Hx* and *lx* (Fig. 5.11). The fact that there are two sites of hybridisation and two polydactylous loci is however, likely to be a coincidence. The gene *rl* is one of the few for which FISH data is available, and it lies approximately in the middle of the B band, and the inversion In2Rk is known to include the *W* gene and to have a proximal boundary in the middle of band D. The more proximal of the two integration sites is in the B band, and is therefore probably close to *rl*, which would locate it about 10 cM away from *Hx* and 17 cM from *lx*. The distal site is within band C, which could locate it close to *lx* or *Hx*. It therefore seems likely that only one of the transgene sites is responsible for the phenotype. Interestingly, out of the hemimelia-luxate group neither of these mutants are as similar in phenotype to *sas* as *lst*, *lu* or *Rim4*, however, the potentially reduced cell death in the interdigital regions may indicate a link with *Hx* (see intro section **1.9.1**).

Fig. 5.10 FISH analysis of *sasquatch* chromosomes.

Two representative chrmosome spreads which show the localisation of the transgene (in green) on chromosome 5. The homologous chromosome 5 is indicated by a black arrow. In all cases two spots were seen on each of the sister chromatids (four spots in all for each chromosome). From this a very rough idea of the position of the transgenes was obtained.

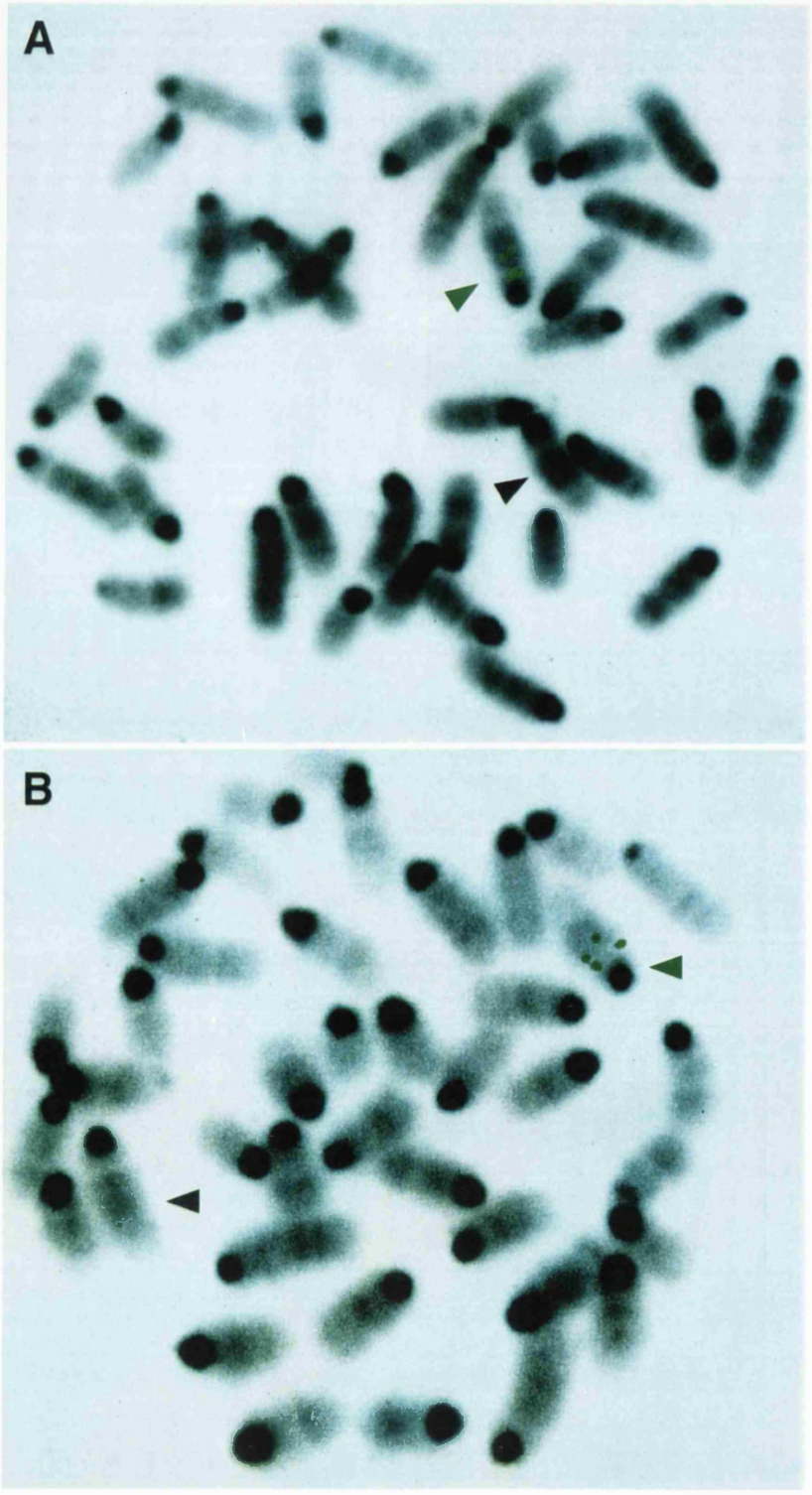
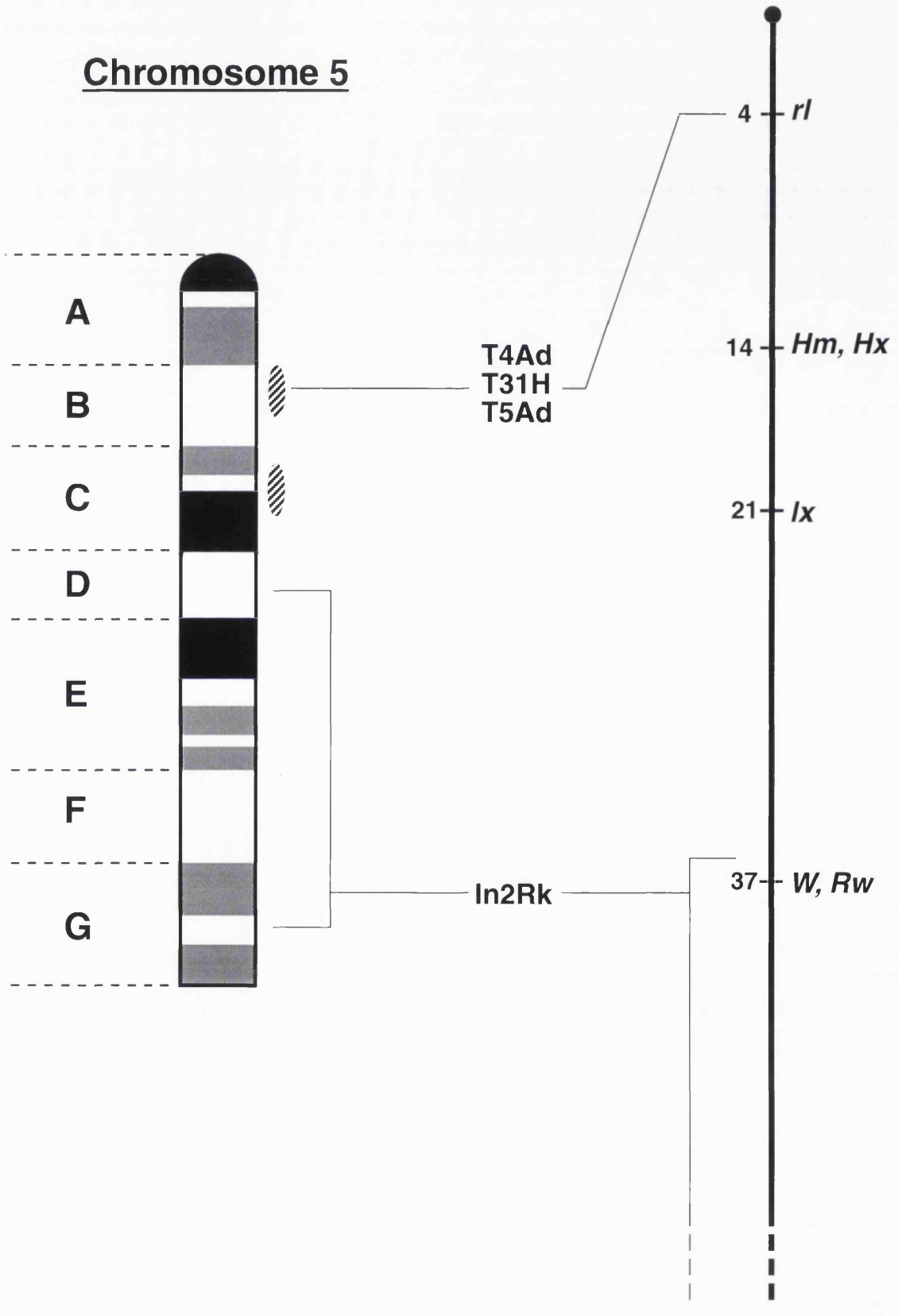

A
B

Fig. 5.11 Physical map of chromosome 5.

On the left is the G-band pattern for chromosome 5, and to the right of this, indicated as patterned ovals, are the very approximate positions of the two hybridisation sites for the transgenic insertions. On the right is part of a physical map, with only the relevent sites marked, and its physical relationship with the G-band pattern as ascertained by previous FISH analysis and observation of the inversion In2Rk (Lyon and Searle, 1989).


Chromosome 5
A
B
C
D
E
F
G
T4Ad
T31H
T5Ad
In2Rk
4 rl
14 Hm, Hx
21 Ix
37 W, Rw

Another gene which is located in this region of chromosome 5 and is involved in limb development is *Shh* (Chang *et al.*, 1994). It is therefore possible that the mutation is a direct alteration of *Shh* regulation. For example, adding an enhancer which directs expression in the anterior part of the limb bud (or blocking a repressor) would produce exactly the results seen. Data presented in the next section partially support this idea, however, it must be remembered that the general features of *sas* fit perfectly into the hemimelia-luxate phenotype, and none of these mutations affect *Shh* in this direct way.

5.7 Expression of the reporter gene during development

5.7.1 Expression in the limb

Mutant embryos were analysed for expression of the PLAP reporter gene using the method described in section **2.5.3**. As expected, expression of PLAP was found in rhombomere 4 at 9.5dpc and 10.5dpc (see section **4.1**). However, it was also seen in the limb bud during development (Fig. 5.12). At 11.5dpc the pattern of PLAP expression was very similar to that of *Shh*. However, whereas *Shh* was weaker in the forelimb compared to the hindlimb, for PLAP the situation was reversed. Since the forelimb at this stage is developmentally ahead of the hindlimb, it showed that whereas *Shh* was starting to down-regulate, PLAP expression was increasing.

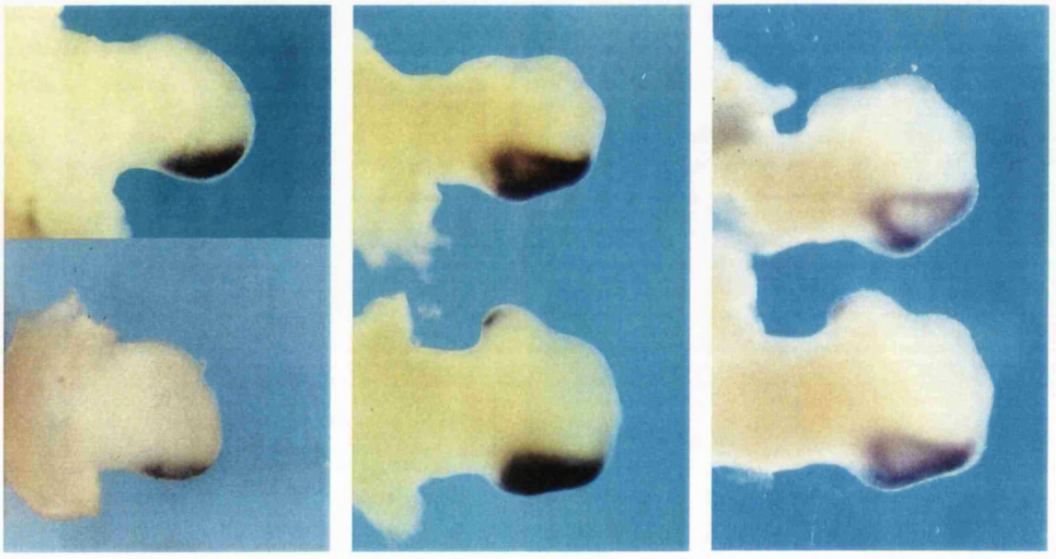
12 hours later PLAP displayed another resemblance to *Shh*: at 12.0dpc an anterior ectopic patch of expression was seen in the proximal half of the ectopic bulge, which is very similar to *Shh* in the mutants. However, at 12.5dpc the resemblance was lost, because whereas *Shh* was completely absent at this stage, the PLAP expression was present and had expanded into a new pattern. It was still weakly seen all over the posterior region, but was now more strongly seen in the two posterior condensing digits, and the adjacent margin of the bud.

These data show that the PLAP reporter construct has integrated into a locus which is expressed in the developing limb, and also provide strong evidence that it is the cause of the phenotype. They also show that the mutated locus contains a cis-acting regulatory element which initially produces a *Shh*-like pattern, but which becomes more extensive than the ZPA later on. Since the expression patterns of *Hx* and *lx* are not known, it is possible that the pattern displayed by the PLAP reporter, is in fact the normal expression pattern for one of them.

Fig. 5.12 Expression of PLAP in the developing limb bud.

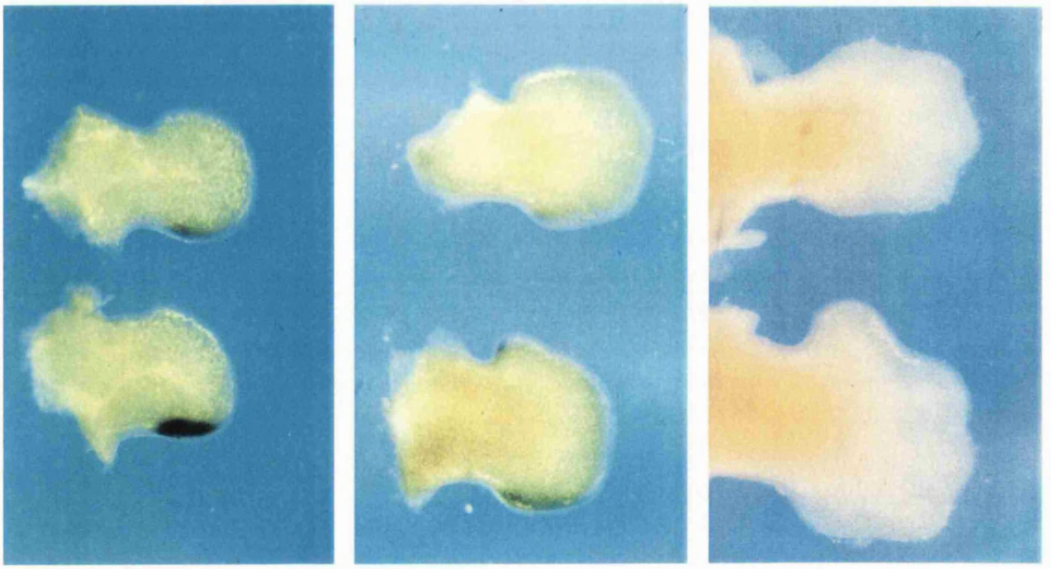
The top panels show expression of the PLAP reporter for heterozygous embryos at stages 11.5dpc, 12.0dpc and 12.5dpc. Each panel shows the forelimb-bud above the hindlimb-bud. The bottom panels show the expression of *Shh* in forelimb and hindlimb-buds at the same stages.

transgenic AP



11.5 dpc 12.0 dpc 12.5 dpc

Sonic hedgehog



11.5 dpc 12.0 dpc 12.5 dpc

Because both the reporter construct and the *in-situ* hybridisation protocol use alkaline phosphatase and BCIP+NBT to produce the colour reaction, it is theoretically possible that the staining for the *in-situ* probe of *Shh* would reveal the transgenic PLAP activity. As a control, the hindbrain staining for these two techniques was compared. In Fig. 5.13a-c, is shown the rhombomere 4-specific staining of the PLAP gene, compared to the floorplate and notochord staining of the *Shh* gene, proving that the reporter gene did not interfere with the *in-situ* protocol.

5.7.2 Expression in the rest of the embryo

There were two other sites of expression seen during embryogenesis. The first was at about 9.5dpc, where it was observed in the flanking region of the embryo anteriorly to the forelimb bud (Fig. 5.13d). It gradually disappeared over the next 24 hours, and was therefore not contiguous with the ectopic anterior expression seen later (which arises *de novo*). This region does not seem to correspond to the earliest region in chick which displays polarising activity (Hornbruch and Wolpert, 1991). The second site was at about 14.4dpc, in the region of the primitive nose (Fig. 5.13i,j). This site does not correspond to the normal *Shh* expression which is observed at 12.5dpc in the very nearby structures of the developing whisker barrels (Fig. 5.13h).

5.8 Summary

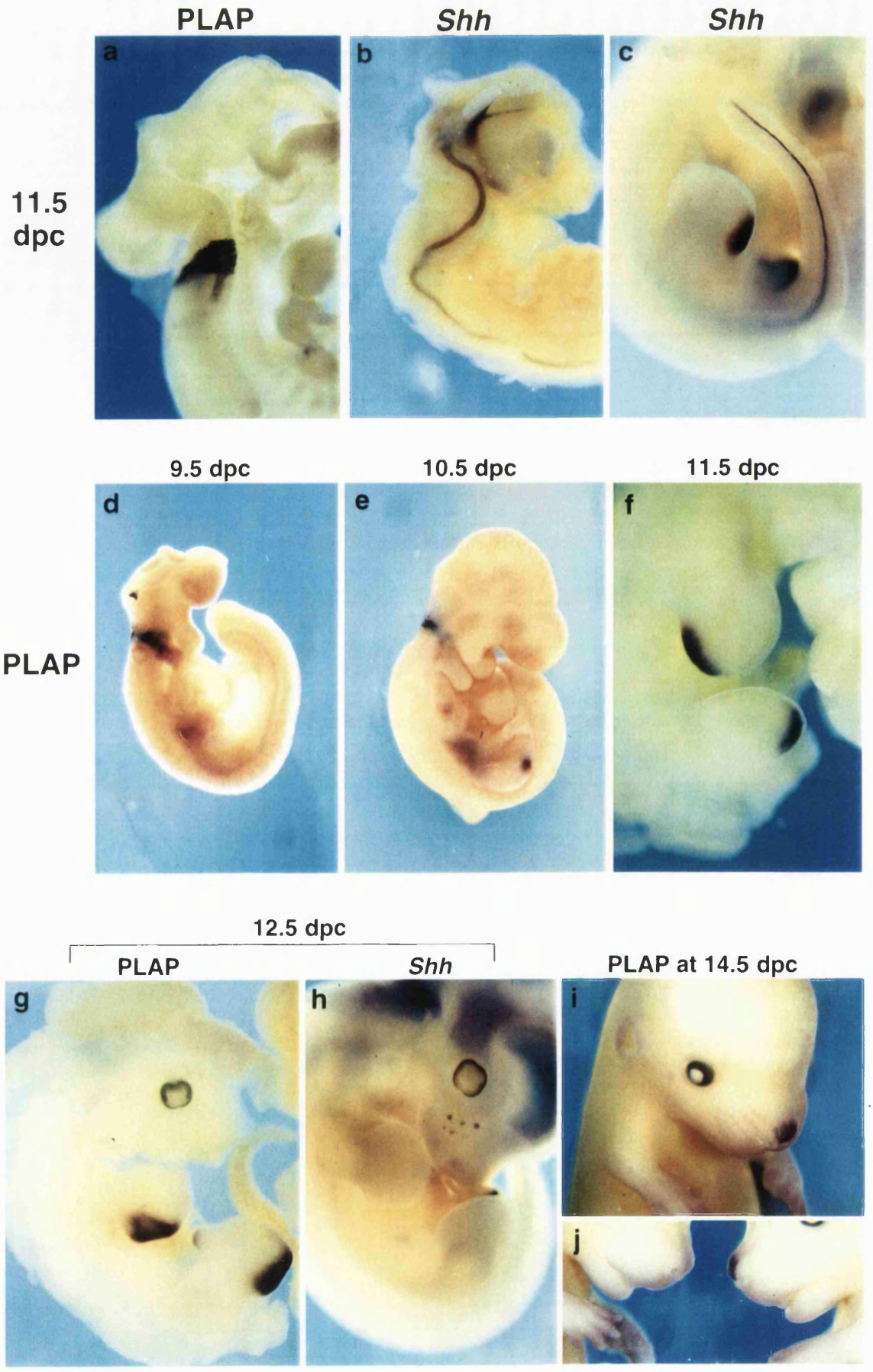
A transgenic line was created which displayed preaxial polydactyly. The observation that PLAP from the transgene was expressed in a *Shh*-like pattern in the developing limb-buds, strongly suggested that this mutation was caused by the transgene insertion. The mutant phenotype was found to be similar, but not identical, to any of the hemimelia-luxate strains. FISH analysis showed that the transgene was inserted at two loci near each other, in the proximal region of chromosome 5. They are near to three loci known to be involved in limb development: *Hx*, *lx* and *Shh*. Whole-mount *in-situ* analysis of limb-buds from 11.5dpc to 12.5 dpc shows that *Shh* is ectopically expressed in a small anterior patch, corresponding to the region where phenotypic defects occurred. This is a phenomenon seen in four of the other members of the hemimelia-luxate group.

Fig. 5.13 Other sites of PLAP and *Shh* expression.

(a-c) Control staining of embryos at 11.5dpc, which show that PLAP-staining in the neural tube is specific to r4, and the *Shh in-situ* hybridisation only picks up the floorplate/notochord expression (b) and not r4. Both embryos in (a) and (b) have been sagitally-bisected.

(d-f) Expression of the PLAP reporter gene from 9.5dpc to 11.5dpc, shows that the earliest site of activity outside r4 is in a small patch anterior to the forelimb-bud. However, this fades away, so that by 11.5dpc it has disappeared, and this is before the ectopic limb-bud expression has begun (at about 12.0dpc, see Fig. 5.9).

Expression of PLAP is not seen in the whisker barrels at 12.5dpc (g) which is another site of *Shh* expression (h), however it is seen in the developing nasal region at 14.5dpc (i,j). (The left-hand embryo in (j) is a wildtype.)


PLAP
Shh
Shh
11.5 dpc
a
b
c
9.5 dpc
10.5 dpc
11.5 dpc
PLAP
d
e
f
12.5 dpc
PLAP
Shh
PLAP at 14.5 dpc
g
h
i
j

CHAPTER 6

DISCUSSION

The intriguing question concerning Hox gene regulation is: How does it relate to the clustered organisation of the Hox complex? This regulation can be broken down into three different types: (1) Regulation at the strictly local level, which would allow genes to act completely independently from each other. (2) Interaction of enhancers and promoters from adjacent or nearby genes. (3) Global regulation imposed by the Hox complex. The last two levels both provide reasons for maintenance of the cluster through evolution, and for varying degrees of coordination of Hox gene regulation. Experiments described in chapters 3 and 4, have provided extra data with which to consider gene regulation at all of these three levels.

6.1 Strictly local regulation of Hox genes

Despite the widespread agreement that there must be more to Hox gene regulation than a collection of strictly local enhancers, nevertheless local enhancers exist and have so far been the most productive area of Hox regulation research. This is mostly because they are easier to study with current experimental approaches (mostly transgenic technology).

6.1.1 Did Hox gene duplications preserve the organisation of local enhancers?

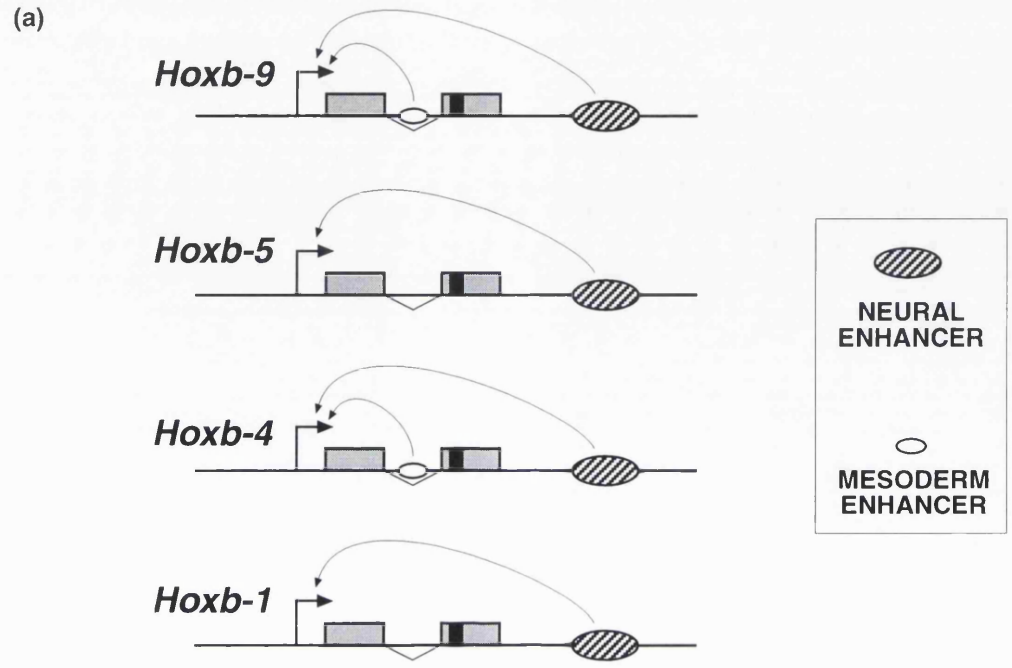
It is generally assumed that the original Hox complex was the result of tandem duplication of an ancestral Hox gene, because in addition to their high sequence similarity, all Hox genes are transcribed in the same direction, and all possess an intron in the same position of the coding region (McGinnis and Krumlauf, 1992). If a particular organisation of regulatory elements already existed around this primitive gene, it may also have been duplicated. Any complex-wide mechanisms of regulation would only have evolved subsequent to this event, but the duplicated enhancers which controlled the original gene may have been retained.

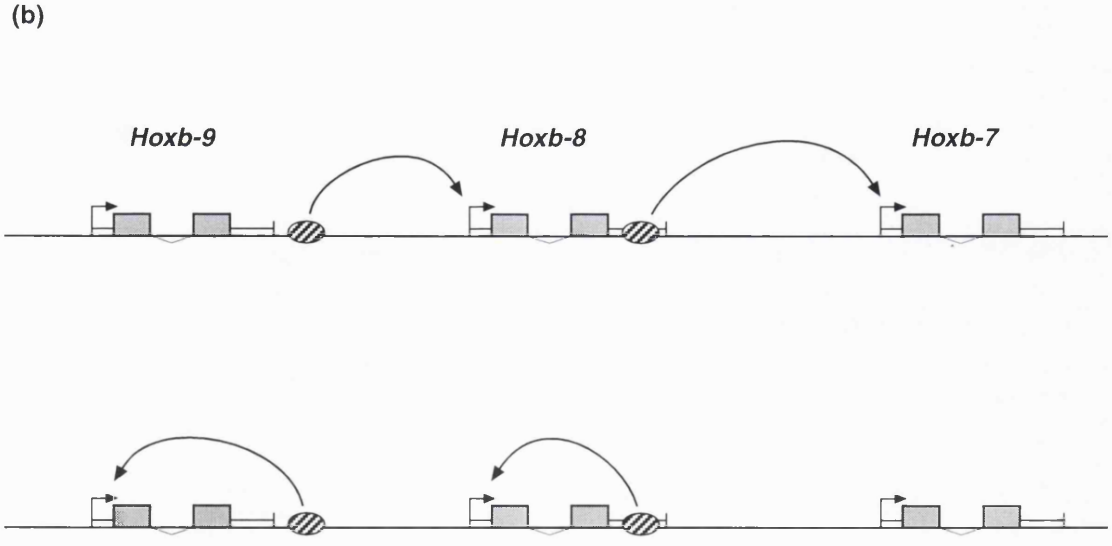
For this reason it is very interesting that the most likely candidates for the *Hoxb-9* and *Hoxb-5* neural enhancers have been found 3' of their respective transcription units. This mirrors very closely the situations already found for two other Hox genes (Fig. 6.1a). The element which

Fig. 6.1 The organisation of regulatory elements near *Hoxb* genes.

(a) Data from this thesis sugget that there are now four genes from the *Hoxb*-complex which posess a neural enhancer 3' to their coding sequence. Of these, two probably contain a mesoderm enhancer within their intron (*Hoxb-9* and *Hoxb-4*).

(b) Work by Charite *et al.* (1995) and Vogels *et al.* (1993), suggested that enhancers for the genes *Hoxb-8* and *Hoxb-7* were each located near to their 5' neighbour, as depicted in the top senario. Data from this thesis suggests that this allocation may be incorrect, and that both enhancers described may be more important for the gene they are closest to.


(a)
Hoxb-9
Hoxb-5
Hoxb-4
Hoxb-1
NEURAL ENHANCER
MESODERM ENHANCER


(b)
Hoxb-9
Hoxb-8
Hoxb-7

drives a *Hoxb-4* neural pattern is in region A (Fig. 4.1), and the early neural expression for *Hoxb-1* is also found 3' to the transcription unit (Marshall *et al.*, 1994). There are other neural enhancers for the *Hoxb-1* gene, but these are involved in the r4-specific upregulation, and are likely to represent an elaboration which arose after the complex had evolved (Studer *et al.*, 1994).

Similarly, it is interesting that the enhancer located within the *Hoxb-9* intron is able to recreate all of the normal *Hoxb-9* pattern except for the late neural expression. The situation is almost identical for *Hoxb-4*, whose region C, also located within the intron, can recreate the whole *Hoxb-4* pattern except for the correct neural expression (and a small region of the lung). From experiments described in chapter 4, and previous work by Stefan Nonchev, it seems that a mesoderm enhancer for the *Hoxb-5* gene is present 3' of the transcription unit, however this does not rule out the possibility of another one within the intron. All the *Hoxb-5*-LacZ experiments have produced quite weak staining, and this may have been the reason why the early constructs tested by Stefan Nonchev showed no enhancer activity very close to the gene. Also there is suggestion of an important enhancer within the intron of the *Xenopus Hoxb-8* gene (Bittner *et al.*, 1993), however this was not recognised in the study by Charite *et al.* for the mouse homologue (Charite *et al.*, 1995). Apart from these cases there is little evidence for mesoderm-specific enhancers within Hox introns.

In summary, it is hard to claim that there is strong evidence for a conserved organisation of local enhancers around each Hox gene. The true situation may well be a mixture, in which the original elements have been retained for some genes (eg *Hoxb-9* and *Hoxb-4*), but lost, modified or displaced for others, and that some of these originally single-gene elements may have come to regulate more than one gene.

6.1.2 Are regulatory elements in the 5' end of the Hox-complex gene-specific?

Another way to consider whether regulatory elements are strictly local, is to see if they produce a spatial pattern which is similar to the normal expression domain of a particular gene. This has been the key evidence in the case of the 3' genes, for believing that local enhancers are operating. The r4-specific enhancer discovered near the *Hoxb-1* gene, and the r3/5 enhancer near *Hoxb-2*, reflect the known patterns of upregulation of the respective genes so closely that it would be hard to believe that they do not directly regulate these genes. However, as mentioned before it

appears that these rhombomere-specific patterns of the 3' Hox genes may not represent the underlying regulatory mechanism believed to be the cause of Hox gene clustering.

Determining and comparing the expression patterns of more 5' Hox genes which are not expressed in the rhombencephalon, is much harder due to the lack of anatomical landmarks in sections of the neural tube and the somites (Vogels *et al.*, 1990). However, the general rule that more 5' genes are expressed up to a more posterior boundary of expression appears to hold true in most cases (Krumlauf, 1994). It certainly appears to be the case for *Hoxb-9* and *Hoxb-8* at the ages around 10.5dpc to 12.5dpc. Charite *et al.* (Charite *et al.*, 1995) have proposed that the neural enhancer located within the HindIII fragment described in chapter 3, is a *Hoxb-8* element, despite their admission that it creates an expression pattern significantly more posterior than the endogenous *Hoxb-8* gene. The ability to compare the element in this study with the pattern of *Hoxb-9* protein, has suggested that this enhancer is instead involved in *Hoxb-9* regulation. Its location, although equidistant from the two promoters, is closer to the *Hoxb-9* transcription unit, and its expression pattern is much more similar to *Hoxb-9* than *Hoxb-8* (whose anterior boundary is near the junction between the spinal cord and the hindbrain at this stage).

Interestingly when this element was tested on the *Hoxb-8* promoter (Charite *et al.*, 1995) it produced a pattern much more extensive than that found in this study (presumably due to promoter specificity which is discussed in section **6.2.3**). Expression was seen in somites, limb-buds and other mesodermal tissue, as well as up to a slightly more anterior boundary in the neural tube than when tested on the *Hoxb-4* promoter (in this study). This more anterior neural expression actually corresponds very closely to the *Hoxb-9* pattern revealed by antibody staining, confirming even further that this element is more likely to be a *Hoxb-9* enhancer.

The idea proposed by Charite et al. that *Hoxb-8* regulatory elements are located close to the 5' adjacent gene (*Hoxb-9*), was also proposed for *Hoxb-7* (by the same group). Vogels *et al.* (Vogels *et al.*, 1993) concluded that the most important elements for *Hoxb-7* are in the 3' untranslated region of *Hoxb-8*, despite the fact that the expression patterns created were again too posterior. These expression boundaries were however, two prevertebrae *anterior* of the boundaries shown for the subsequent study on putative *Hoxb-8* enhancers (near *Hoxb-9*). This suggests that the enhancers found were in fact more likely to be regulating *Hoxb-8*.

The work described in chapter 3, which points to the HindIII neural enhancer being the only one which could drive a *Hoxb-9*-like pattern, has a knock-on effect for the allocation of enhancers to promoters for the two neighbouring genes. If enhancers are assigned to promoters by comparing

their patterns of activity with those of their neighbouring genes, then the element previously allocated to regulate *Hoxb-8* now regulates *Hoxb-9*, and the suggested element for *Hoxb-7* now regulates *Hoxb-8* (Fig. 6.1b). This reallocation also makes enhancers regulate the promoter that they are closer to, which seems to make sense. However, the difficulty revealed here of being certain about which enhancer is regulating which promoter, also leaves wide open the possibility that elements do not have strict one-to-one relationships with promoters, and instead "flip" back and forth between many genes, and this is discussed in section **6.2**.

6.1.3 Possible mechanisms of the gene-specific activity in region D

Four basic activities could be attributed to the enhancers within regions D and E: neural tube, somites and limb-bud expression from region D, and neural expression from region E. It is not possible to determine whether the neural element in region E is gene-specific *in vivo*, although the results from chapter 4 show that it *is* able to work on both promoters in transgenic experiments (although not simultaneously). The somitic element appears to be the best candidate for a shared enhancer, and in this respect is not gene-specific. The limb enhancer in region D acted in a very promoter-specific way. In all cases when D was tested on the *Hoxb-4* promoter, limb expression was seen, and in all cases when it was in the same construct as the *Hoxb-5* promoter, LacZ expression was never found in the strong, posterior limb pattern.

Results for the region D neural activity, were essentially the same as the limb element, but its interpretation is slightly complicated by the fact that region E also displayed neural activity. This means that in the full double-labelled constructs (in which both promoters show neural expression) it is not formally possible to know which enhancer is working on which promoter. However, the idea that the enhancers would reach over each other to interact with a promoter which is further away than their neighbour seems highly unlikely, and when region D was tested on its own, it was unable to drive neural expression from the *Hoxb-5* promoter.

This type of promoter-specific interaction of a large regulatory fragment like region D, can be thought of within either a *strictly local* or a *shared* context. To be the result of a strictly-local situation, in which all enhancers can only operate on a single gene, region D would have to contain elements for *Hoxb-4* which are completely separable from those which can work on *Hoxb-5*. The prevention of *Hoxb-4* elements from working on the *Hoxb-5* promoter could then be achieved in three different ways: boundary elements, enhancer/promoter incompatibility, or polar competition.

Boundary elements are DNA sequences which can prevent regulatory elements on one side from interacting with promoters on the other (Galloni *et al.*, 1993; Gyurkovics *et al.*, 1990). Alternatively, promoter/enhancer pairs which are not meant to interact could be incompatible through the conformation of the "adaptor" proteins which mediate these reactions (Li and Noll, 1994). Although both of these are possible, boundary elements have never been found in mice, and most regulatory elements within the Hoxb complex are able to work on a number of different promoters, and therefore appear to function as general enhancers.

Polar competition describes the idea that although an element *can* work on two different promoters, it is prevented from operating on one of them due to competition from another enhancer. (It can also be thought of in the reverse situation, in which two promoters compete for one enhancer.) This type of system for specifying enhancer-promoter interactions, could result either in a strict one-to-one relationship between enhancers and promoters, or if the competition was not complete (ie. not *polar*) it could result in the *sharing* of elements. In this case the term *promoter specificty* would not need to refer to the prevention of certain enhancer-promoter interactions, but instead to the fact that the same enhancer might produce different regulatory effects on different promoters. This distinction is discussed further in section **6.2.3**.

6.2 Interactions between regulatory elements

As an alternative mechanism to strictly local enhancers, it is possible that regulatory elements might be important for more than one promoter. In the most simple case (which is discussed in chapter 4) enhancers may control two adjacent genes, however the real situation may be far more complex. It is possible that instead of a small number of enhancers for each gene, most regions of the complex possess some degree of regulatory influence, and that only when combined together will they correctly control the 9 different promoters within the Hoxb-complex. Experiments performed by Deschamps (Charite *et al.*, 1995) suggest that every subsection of DNA around the *Hoxb-8* gene displays some spatially-specific regulation. Some of these patterns are similar to *Hoxb-8*, and some are not. These results allow two alternative conclusions: either spatially-specific patterns can be directed by spurious activity of "unimportant" pieces of DNA, or enhancers of varying strength are thickly scattered along the Hox complex and work together to achieve their function. The fact that even the non-*Hoxb-8* patterns, were nevertheless Hox-like suggests this latter situation.

6.2.1 Is the HindIII neural element shared between *Hoxb-9* and *Hoxb-8*?

The study of *Hoxb-8* regulation by Charite *et al.* was described in section **6.1.2**. Our attempt to assign a number of enhancers in a one-to-one relationship to their neighbouring promoters (*Hoxb-9* to *Hoxb-7*), was rendered non-trivial by the fact that none of the enhancer patterns perfectly match any of the gene patterns. Instead of forcing these relationships based on pattern similarity and element proximity, a more accurate view of the situation may be that correct regulation is achieved only through the interaction of many enhancers. This interaction could be competition, or it could be modulation of each other's activity through physical interaction.

Sharp boundaries of expression from reporter genes, such as those seen with construct b9-D at 11.5dpc (Fig. 3.4h), might initially suggest that the enhancer responsible is acting with a high degree of accuracy, and therefore does not require cooperation with other *cis*-acting regions. However, it is quite probable that sharp boundaries are created and maintained not by extremely accurate, cell-autonomous reading of positional information, but by cell-cell communication and local regulation of Hox expression (Wilkinson, 1993).

In an attempt to determine whether the HindIII neural element is important for both *Hoxb-9* and *Hoxb-8*, a collaboration has been initiated with Deschamps' laboratory. PLAP has been inserted into the original 12kb fragment of *Hoxb-9*, at the site of the LacZ gene. This construct will then be joined to a fragment containing *Hoxb-8* labelled with LacZ, plus its flanking regions. If initial constructs drive the same boundary of expression in the neural tube for both the *Hoxb-9* and *Hoxb-8* genes, then extra 3' sequences will be added back in the hope of creating a difference between the patterns. In both constructs the neural element will be deleted to determine whether it is contributing to the regulation of both genes.

6.2.2 Three different modes of regulatory interaction between *Hoxb-5* and *Hoxb-4*

Determining whether a single enhancer functions *in vivo* on two genes is not easy. Frasch *et al.* (1995) were able to identify at least one enhancer between the genes *Hoxa-1* and *Hoxa-2*, which drives a subset of the expression pattern for both genes (rhombomere 2), but were unable to determine whether it was important for one or the other, or both of the genes. Van der Hoeven *et al.* (van der Hoeven *et al.*, 1996) performed experiments in which a copy of *Hoxd-11* was inserted next

to the *Hoxd-13* gene, and concluded that in this artificial situation, a forearm-specific enhancer was acting on both genes. However, this conclusion was based on the assumption that a forearm enhancer was located within the transgene construct, despite the fact that forearm expression was never observed when the transgene inserted randomly into the genome.

The most direct way to study a potentially shared element would be to mutate it *in vivo* and analyse the resultant gene expression patterns. But even in this experiment, the fact that cross-regulation occurs between Hox genes (Faiella *et al.*, 1994; and unpublished results from the Krumlauf lab.) could make it impossible to know whether the effect on a neighbouring gene was the indirect result of *trans*-regulation from the single gene which was genuinely affected by the mutated *cis*-acting element. The approach chosen in this study, could also potentially suffer from one experimental drawback: multiple-copy transgenics. But as described in section **6.2.4**, this issue may not be a problem in these particular experiments, and it will not be discussed in this section.

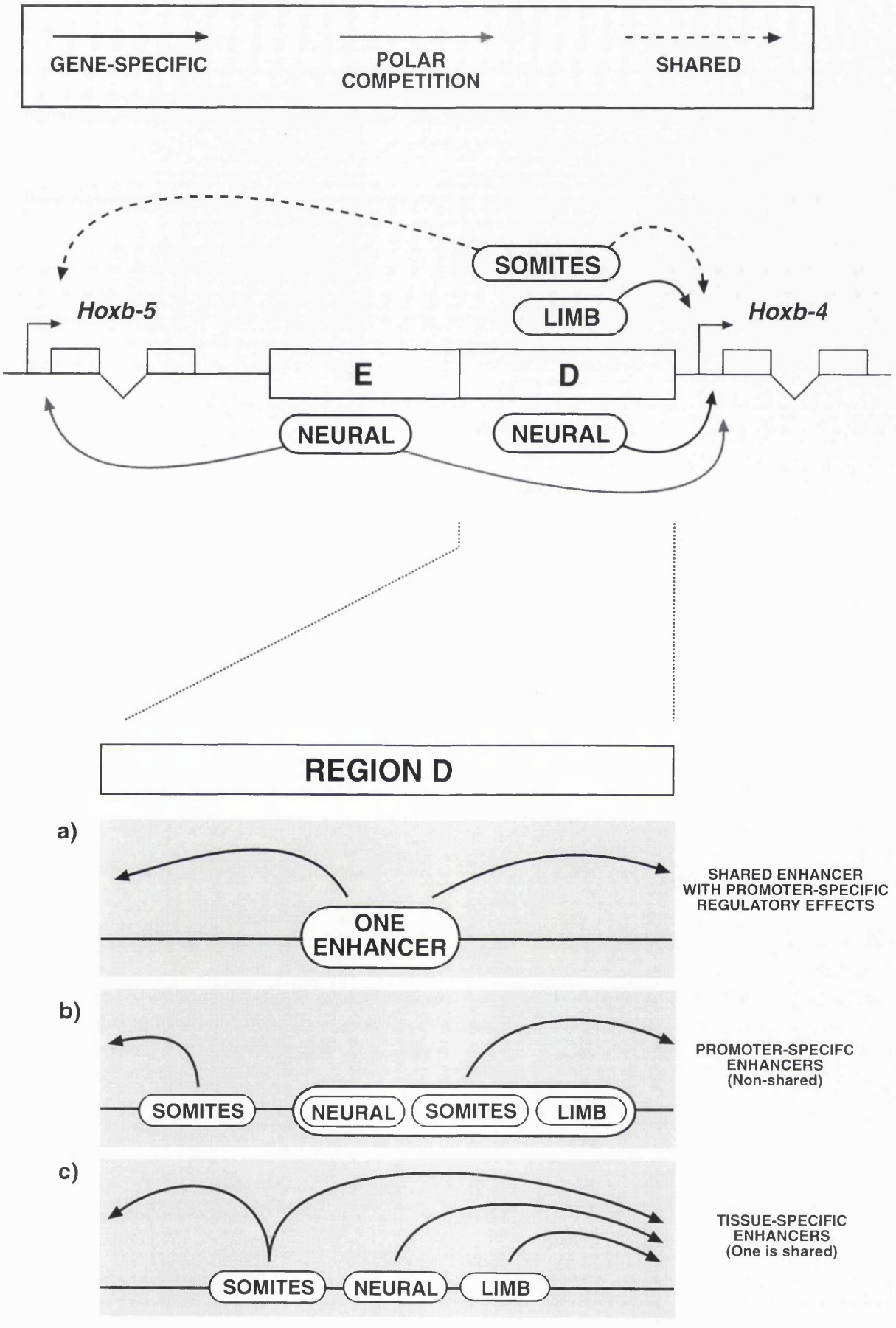
The results presented in chapter 4, show that within the DE region are examples of three different types of regulatory elements: gene-specific activity (described in section **6.1.3**), polar competition, and sharing (Fig. 6.2). The limb and neural elements in region D appear to be unable to work on the *Hoxb-5* promoter, but the neural element in region E is different. In the full construct (DG[ED]), it activated neural expression from the *Hoxb-5* promoter (because the region D neural element cannot activate this promoter), and in constructs b4[E] and DG[E] it activated the *Hoxb-4* promoter. However, when given the choice of which promoter to work on (in the construct DG[E]) it only activated *Hoxb-4*. This particular element therefore appears to display polar competition. The discovery of this phenomenon strengthens the subsequent belief that some enhancers (namely the region D somite element) may be able to compete in a less biased way, and thereby operate on both promoters, by flipping back-and-forth between them. This has been suggested as the mechanism of controlling multiple genes in the β-globin locus (Wijgerde *et al.*, 1995).

The third type of regulatory element found within region D is the somite-specific activity. In all the constructs which contain region D, somite expression is observed from both promoters (if both are present), and in all the constructs which do not contain region D, no somite expression is found from either promoter. In this respect the two promoters act identically. In a section of DNA as large as region D (4.5kb) there is room for a number of functionally-independent enhancers, so these findings are not conclusive proof of a single element working on two different promoters (a criticism equally applicable to the claim of van der Hoeven *et al.*). However, the anterior boundary of expression is identical from the two promoters, and this strengthens the hypothesis. At the very

Fig. 6.2 The regulatory effects of regions E and D.

Regulatory activities within regions E and D can be differentiated into three types: the limb and neural activity from region D on *Hoxb-4* is *gene-specific*; the neural activity from region E can work on either the *Hoxb-5* promoter *or* the *Hoxb-4* promoter, but not on both, so it displays *polar competition*; and the somitic enhancer appears to be able to work on both promoters at the same time, so may represent a *shared* enhancer.

There are many ways in which the complex activity from region D could occur, and three general alternatives are shown: (a) The entirety of the pattern possible from region D is defined by a single *trans*-acting factor or complex, and the different subsets of this which occur from the two promoters are selected by promoter-specific information. (b) Enhancers within D are divided into two completely independant groups, each of which can only interact with one of the promoters. (c) Enhancers are tissue-specific (or spatially-specific) but not necessarily promoter-specific, such that enhancers for tissues which require both genes will be shared between both promoters.


GENE-SPECIFIC
POLAR COMPETITION
SHARED
SOMITES
LIMB
Hoxb-5
Hoxb-4
E
D
NEURAL
NEURAL
REGION D
a)
ONE ENHANCER
SHARED ENHANCER WITH PROMOTER-SPECIFIC REGULATORY EFFECTS
b)
SOMITES
NEURAL
SOMITES
LIMB
PROMOTER-SPECIFC ENHANCERS (Non-shared)
c)
SOMITES
NEURAL
LIMB
TISSUE-SPECIFIC ENHANCERS (One is shared)

least, a strong position has been found from which to refine the analysis. Smaller deletions within region D can be made, with the same goal of finding the loss of somite expression from both reporters.

6.2.3 Promoter specificity

If there is a single somite-specific enhancer within region D, then it is having a different regulatory effect on the *Hoxb-4* promoter compared to *Hoxb-5* (Fig. 6.2). The LacZ expression from *Hoxb-5* was always restricted to about 7 somites, whereas the PLAP expression continued posteriorly into the tail. Two experimental problems could be considered as the explanation for this discrepancy: (1) the PLAP protein is stronger or degraded more slowly than the β-gal protein, (2) the arrangements of reporters with their respective promoters causes a difference in strength of expression (as the LacZ staining was always generally weaker than the PLAP). However, comparison with the protein expression patterns for the two wildtype genes (Fig. 4.6) suggests that in fact, this difference reflects the normal situation. It appears that Hox gene somite expression can occur in either of these two modes: the 6-8 somite pattern (for example, *Hoxb-5*, *Hoxd-4* (L. McNaughton, Pers. comm.), *Hoxc-6* (Duboule, 1994)), or the continuous pattern (for example: *Hoxb-4*, *Hoxb-9*). It is not clear what the purpose or mechanism of this difference may be, but it does not correlate with the paralogous group or cluster the gene is from.

It is therefore possible that in the case of *Hoxb-4* and *Hoxb-5*, this difference is controlled by promoter-specific information. No two Hox promoters that have been sequenced are identical, and it is known that within the *Hoxb-4* promoter are a number of different regions which each influence different aspects of regulation (Gutman 1995). The correct regulation of Hox genes could therefore be achieved through the integration of enhancer-specific and promoter-specific information. This kind of promoter-specificity, which fits very well with the idea of shared enhancers, is quite distinct from the other type, in which promoters are specific for which enhancers they can interact with (which is called enhancer/promoter incompatibility, described above).

If promoter-specificity is an important component of Hox gene regulation, this could have important consequences for experiments on enhancer analysis. The classical definition of an enhancer (that it can work on heterologous promoters, in both orientations), may be particularly inappropriate for the Hox complex, because it precludes any regulatory element whose activity may be modulated by a promoter. It now seems more clear that even the difference between Hox

promoters has an effect not only on the strength of reporter expression, but also the spatial regulation. For example, the neural enhancer described in chapter 3 produced a more posterior boundary of expression on the *Hoxb-4* promoter, than it did on the *Hoxb-8* promoter in the study by Charite *et al.* (Charite *et al.*, 1995).

Considering this problem, how can we chose which promoter to use in these experiments? A Hox promoter is preferable to a non-Hox promoter, as it will interact with the enhancer in a more wildtype manner, but some promoters, like *Hoxb-5*, are simply not very strong. The ideal solution would be to test every enhancer on a range of different promoters, but in lieu of that, it would be useful to have an idea of what types of influence promoters can have *versus* enhancers. In the case of the *Hoxb-9* neural enhancer, the two different promoters caused a shift in anterior boundary. Understanding the situation in region D depends on information which we do not have: how many different enhancers are responsible for the three different activities. If the neural and somitic activities are controlled by separate elements, and there is only one somite element, then the different promoters are not causing a shift in anterior boundary, but instead causing the difference between a 6-somite pattern and the extended pattern. This would suggest that the main activity of an enhancer which cannot be modulated by its promoter is *tissue-specificity*. Alternatively, if the neural and somitic expression is controlled by a single element, then to a certain extent the promoter *is* able to determine tissue-specificity, from the fact that the *Hoxb-5* promoter inhibits expression in the neural tube. Obtaining a more detailed description of the elements within region D, will therefore be of enormous help in determining which of these alternatives is occuring.

In general there is a lot of evidence suggesting that enhancers are tissue-specific, but this does not mean that one enhancer is restricted to one tissue-type. In fact, in the Hoxb genes there appear to be two main enhancer types: those which drive expression exclusively in the neural tube (eg. regions A and E), and those which are active in many tissues (eg. regions C and D). This may reflect an underlying difference in the way different tissues are patterned, or it could be simply a consequence of the way the patterning system evolved.

6.2.4 Are multiple-copy integrations no problem?

A potential problem with experiments which attempt to prove that one element is operating on two promoters, is that in multiple-copy transgenics there will be more than one copy of the enhancer. Therefore, in theory, some copies of the enhancer could be working on the first promoter,

and the others could be working on the second (Fig. 6.3a). Enhancers from one copy of the construct could also work on promoters from the adjacent copy. In this case an enhancer designed to work only on one promoter could be recorded as a shared element. One solution to this problem would be to generate single-copy transgenics. However, there is evidence from the results presented in this thesis, that in the case of elements in region D and E, this may not be a problem.

The neural element within region E is strongly capable of driving expression from both the *Hoxb-5* and the *Hoxb-4* promoter. But in construct DG[E], in which it is given the option to work on both, it only works on *Hoxb-4*. The chances of a single-copy integration event are at a maximum of 10% (using very low-concentration DNA), so the chance that both transgenic lines created with this construct contained single-copy integrations is a maximum of 1%, and probably much lower. It suggests that in this case, every copy of the transgene behaved the same way, and regulatory interactions between copies did not occur (Fig. 6.3b).

From the following observations we can assume that this neural element is normally designed to work on only one promoter: (1) When given the option of working on two it only chooses one. (2) It is the closest neural element to *Hoxb-5* and gives a *Hoxb-5* pattern. (3) There is another neural element between it and the *Hoxb-4* promoter (in region D), which has a higher likelihood of working on *Hoxb-4* than it does. From this assumption we can extrapolate that all enhancers which are supposed to work on only one promoter, are *able* to display this preference in the double-reporter assay. Consequently, any enhancer which does not display such a preference, must not be restricted in this way. In other words, if the somite enhancer normally works on only one promoter (due to domain boundaries, or a preference for one promoter over the other) then this behaviour would be allowed in the double-reporter assay, because all wildtype sequences between the two genes are present. The fact that it is *not* observed is therefore strong evidence for its *in vivo* ability to work on both genes.

6.3 Global interactions

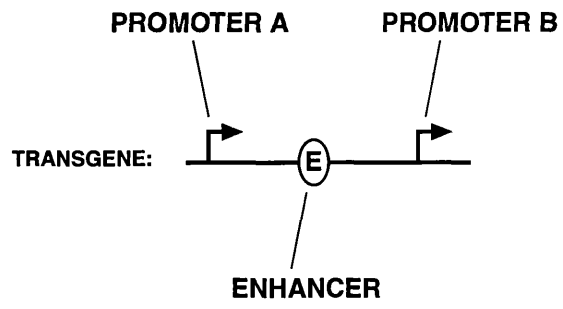
Complex-wide mechanisms which could be operating in the Hox cluster have been described in the introduction, and often involve the ideas of progressive chromatin state-changes. As explained in section **1.5.5** these ideas were first linked to *initiation* of Hox expression, then to *maintenance* (as in *Drosophila*), and more recently are again being connected with initiation (van der Hoeven *et al.*, 1996; van der Lugt *et al.*, 1996).

Fig. 6.3 Muliple-copy tandem arrays from integration of transgenic constructs.

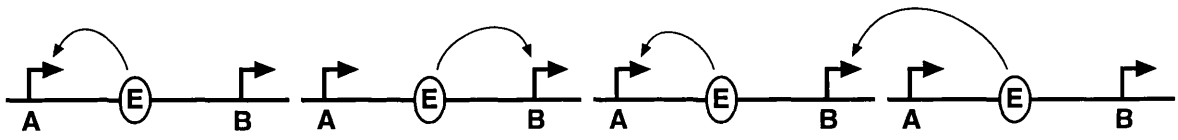
(a) If a hypothetical construct with two promoters (A and B) and one enhancer is integrated into a tandom array, then a number of different regulatory interactions could occur. In the scenario shown, the first copy of the enhancer activates promoter A, and the second copy activates B. In the third transgene, A is activated by its own enhancer, and B is activated by the enhancer from the adjacent transgene.

(b) Despite the several types of interaction that could occur, and the fact that region E is *able* to activate the *Hoxb-5* promoter, in construct DG[E] only the *Hoxb-4* promoter is active.

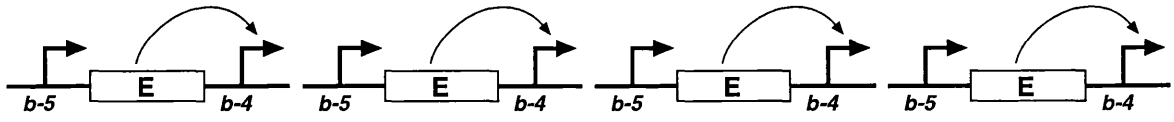
(a)



TANDEM ARRAY:



(b)



The data presented in this study cannot contribute to our knowledge of this level, apart from confirming the common observation that completely perfect Hox regulation is never achieved from transgenic constructs which only contain parts of the cluster, and that this missing information may well come from a larger-scale mechanism. It appears that enhancer sharing and regulatory interactions are likely to be important aspects of Hox regulation, and may even contribute to the evolutionary stability of the complex through evolution, but understanding of more global mechanisms is more likely to result from analysis of the vertebrate Pc-G and Trx-G genes, and chromatin-related phenomena.

6.4 Future experiments

The most important experiment remaining from the study of shared enhancers, is to more precisely localise the somitic enhancer within region D, and to then repeat the double-reporter experiment with only this smaller region deleted. If this is possible, and the same result is found (that somitic expression is removed from both promoters) then proof will have been found for a shared enhancer. Other experiments which are also planned, are constructs in which regions A, B and C are included in their normal position. Since the anterior boundary of expression for these elements are more anterior than those for region D and E, the first question will be whether the *Hoxb-4* promoter is regulated with these more anterior boundaries while *Hoxb-5* retains the pattern driven from regions D and E. If this is the case, then further experiments can determine whether the prevention of regions A, B and C from acting on *Hoxb-5* is due to boundary elements, or polar competition. As described in section **6.2.1** a collaboration for similar experiments on the genes *Hoxb-9* and *Hoxb-8* has also been initiated with the Deschamps laboratory.

PART II

6.5 Is the *sasquatch* mutation caused by the transgene?

Apart from a single discrepancy which was probably the result of human error, all mice found with the *sasquatch* phenotype showed a positive PCR result for presence of the transgenic construct. However, even more suggestive than that, expression of PLAP was found in the region

where the defect occurs. The first question to ask is: Could the phenotype be the consequence of activity from the transgene itself? (As opposed to disruption of the target locus.) It is unlikely that the alkaline phosphatase protein is triggering a change in pattern formation, since it has been used in numerous previous experiments described in this thesis with no detectable problem. Most of these constructs have directed expression in the limb buds as well as many other parts of the embryo. The only other potentially active gene within the construct is *Hoxb-1*. As described extensively in the introduction, Hox genes are extremely important in the developing limb, so theoretically ectopic expression of *Hoxb-1* could disrupt patterning. But this seems extremely unlikely, as the first 33 codons of *Hoxb-1*, including the translation-start site, have been removed and the open-reading frame is 3' of the PLAP gene (at least 2kb away from the mRNA cap site).

The next question is: Why is the phenotype in the limb? All previous constructs using the same region of DNA from the *Hoxb-1* gene to drive LacZ (by Heather Marshall and Michele Studer in the lab.) have never produced expression in the limbs. This ectopic site of expression is therefore due either to local regulatory influences, or to a spurious "limb-responsive enhancer" created by chance during the integration process. We cannot assess this distinction directly, but if our assumptions about the previous question are correct, then the site of integration must contain a gene involved in limb development, and this would suggest involvement of a limb enhancer. Statistically this is the most likely explanation.

6.6 The connection with *Sonic*

There are three separate effects that we can measure in *sasquatch*: (1) the phenotype, (2) the ectopic expression of *Shh*, and (3) the expression of PLAP in a *Shh*-like pattern. They are likely to be intimately connected, but can be considered separately.

The phenotype is extremely similar to those produced by mutants of the hemimelia-luxate group, but it is not identical and therefore probably represents a new allele or a new gene. All these mutants are characterised by an abnormal degree of mesenchymal growth in the anterior region of the limb-bud at stages 11.5dpc to 12.5dpc. The published accounts of *Shh* expression in *lst*, *Rim4*, *Hx* and *Xt* have described an ectopic region very similar to that found in this study (Chan *et al.*, 1995; Masuya *et al.*, 1995). However it is hard to prove whether the primary effect of these mutations is direct misregulation of *Shh* which then causes the phenotype, or whether this ectopic expression is a downstream or secondary side-effect.

It has been suggested that the ectopic *Shh* expression is due to the disruption of a repression event (Chan *et al.*, 1995). If this is the case, *sas* could be another mutation which performs this disruption. Why the default state would be induction of *Shh* in both the anterior and posterior of the limb is unclear - analysis of fish has shown that despite possessing more symmetrical fin-bids than tetrapods they still only exhibit a single posterior ZPA (Krauss *et al.*, 1993; Sordino *et al.*, 1995).

The expression of PLAP is very suggestive of involvement of *Shh*, however the pattern at 12.5dpc, in which staining is seen concentrated in the posterior two condensing digits and the adjacent limb-bud margin, is hard to explain. Either it is a technical side-effect due to the produrance of the PLAP protein, which may be a result of its extreme stability. Alternatively, if it is a true reflection of a functional enhancer, the transgene must have inserted near a gene which is involved in digit specification. If this were the case then a spurious ectopic activation of the gene in the anterior of the limb-bud, could cause re-specification of that region into tissue with a posterior character. Despite this uncertainty, the correlations between PLAP expression and *Shh* expression are very strong. In the mutant hindlimbs of heterozygotes PLAP is expressed both in a posterior domain which looks initially like *Shh*, and then slightly later in the anterior patch which also reflects *Shh*. However, in the forelimbs which are not mutant, both PLAP and *Shh* are only expressed in the posterior region (Fig. 6.4). These correlations, of PLAP with *Shh*, and of anterior expression of both with the phenotype, make it most likely that *sas* has integrated near a gene which is either downstream or upstream of *Shh*, or is *Shh* itself.

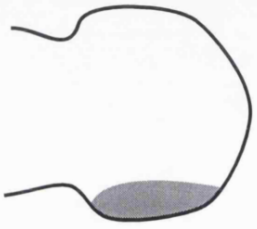
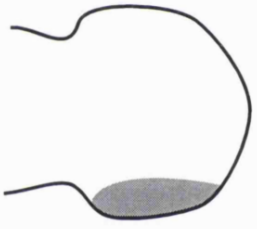
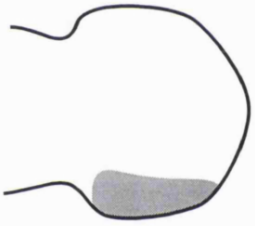
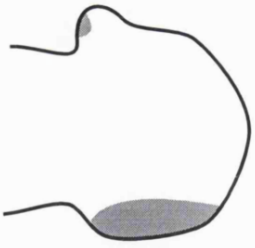
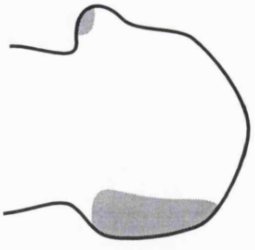
6.6.1 Has the transgene integrated into a gene downstream of *Shh*?

This hypothesis fits with two of the pieces of data, the early correlation of PLAP expression with *Shh*, and the later non-*Shh*-like pattern in the two posterior digits. A limb gene directly responding to the *Shh* signal would initially display a pattern which mirrored it, but could subsequently be modified for its own patterning function. If the gene was used to distinguish digits 4 and 5 from the rest, it could use the *Shh* signal to initiate its pattern in the correct general region, and then a condensation-specific enhancer to further restrict this to the presumptive digits.

The obvious problem with this theory is that it would not directly explain the mis-regulation of *Shh*. It is quite possible for a gene to be both downstream and upstream of *Shh* (as is thought to be the case for *Fgf-4* (Laufer *et al.*, 1994)), but this is most likely to occur as a mutually-dependant feedback loop, such that each partner acts as a positive influence on the other. In such a situation,

Fig. 6.4 The correlation between PLAP and *Shh* expression in sasquatch mice.

In a *sasquatch* forelimb the pattern of both *Shh* and PLAP is restricted to the ZPA, and the limb develops normally. In a *sasquatch* hindlimb the pattern of both genes includes the ZPA and the small anterior patch in the abnormal bulge, and the limb becomes polydactylous. These correlations strongly suggest that the transgene is responsible for the mutant phenotype, and that it is closely associated (directly or indirectly) with the *Shh* gene.

	Shh	PLAP	phenotype
Wildtype			normal
sasquatch forelimb			normal
sasquatch hindlimb			polydactyly

ectopic expression of *Shh* would be the result of ectopic expression of the other gene, and we would then have to explain why this other gene was only activated in this small anterior patch.

In trying to explain why a number of different mutations all appear to cause ectopic *Shh* expression in the same location, we can consider two alternatives. Either they are mutants in positive regulatory genes, which all by sheer coincidence have become activated in the same tissue. Or alternatively, *Shh* has a predisposition for being active in that location, and the mutant genes were repressors which have become deactivated. The latter theory has two points in its favour. First, it removes the idea of coincidence that the many positive regulatory genes which are not normally active in the anterior region, all became active in the same place by chance. Instead, the normal expression domains of the different genes can be extremely variable, as long as they include the anterior patch. The prevelence of this site for *Shh* misregulation can then be entirely due to the regulatory predisposition of a single gene (the *Shh* gene itself). Second, it allows all (or most) of the mutations to be null mutations, which are always more likely than gain-of-function. So far, only *Xt* is known to be a loss-of-function, and only *Dh* is confidently considered to be a gain-of-function (pers. comm. D. Hughes), but evidence is emerging that the GLI3 protein encoded by the *Xt* gene is indeed a repressor of *Shh*, and that its normal expression pattern in the limb is complementary to *Shh* (Buscher *et al.*, 1996).

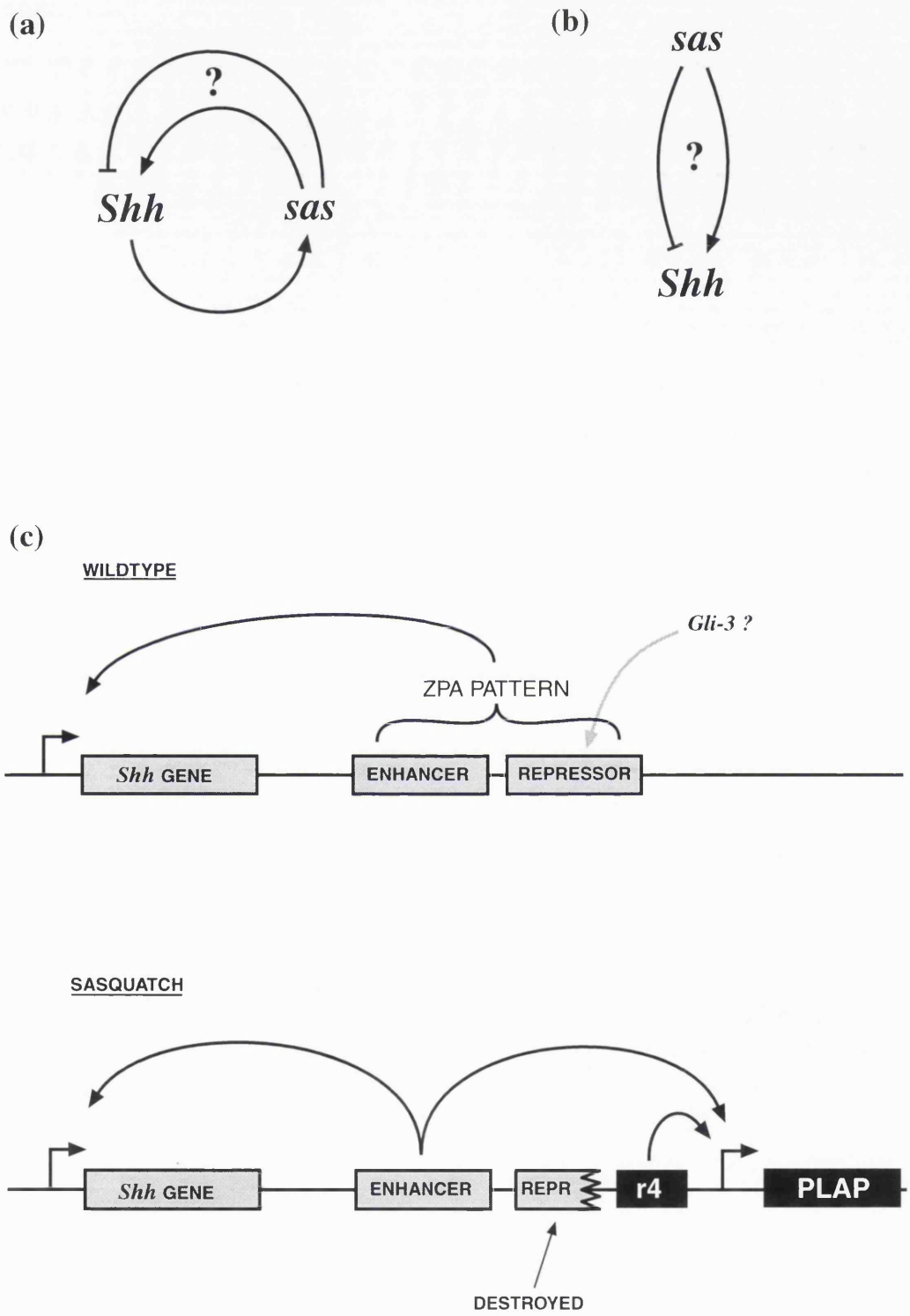
If the genes responsible for these mutants (including *sas*) are more likely to be repressors, then they are very unlikely to be in mutually-dependant feedback loops. This in itself, suggests that *sas* is unlikely to be a downstream target of *Shh*, but the closely-matched patterns of PLAP and *Shh* are also good evidence that *sas* is unlikely to be a repressor of *Shh* (Fig. 6.5a).

6.6.2 Has the transgene integrated into a gene upstream of *Shh*?

The arguments in favour of this hypothesis are very similar to the previous case, except for the important advantage that the gene responsible does not have to be both upstream and downstream (Fig. 6.5b). If the conclusion described above is true - that hemimelia-luxate mutants are most likely to be repressors of *Shh* - then this would also argue against *sas* being an upstream gene. The correlated patterns of *Shh* and PLAP suggest induction not repression. It is therefore theoretically possible that the PLAP reporter has picked up activity of a *Shh*-activating gene, and that this has by coincidence become activated in the anterior patch.

Fig. 6.5 Possible interactions between *Shh* and the transgene.

As explained in sections **6.6.1** and **6.6.2**, if it is considered that *sas* is a separate gene which interacts with *Shh* in *trans*, then there are a number of difficulties in understanding the nature of this interaction (a,b). The theory that the transgene is interacting with *Shh* in *cis*, is easier to postulate. In the most obvious scenario (c), integration of the transgene (black boxes) disrupts the normal ZPA regulation of *Shh* (maybe by destroying a repressor which responds to the GLI-3 protein) thereby derepressing *Shh* in the anterior patch, and simulataneously causing the PLAP reporter to pick-up the activity of the new, altered *Shh* pattern of regulation.


(a)
?
Shh
sas
(b)
sas
?
Shh
(c)
WILDTYPE
Gli-3 ?
ZPA PATTERN
Shh GENE
ENHANCER
REPRESSOR
SASQUATCH
Shh GENE
ENHANCER
REPR
r4
PLAP
DESTROYED

6.6.3 Has the transgene integrated into *Shh*?

Both of the previous theories had the advantage of explaining the late pattern of PLAP expression. This was because the gene considered (either upstream or downstream) is unknown and the late PLAP expression could therefore simply reflect its normal activity. This last theory must consider the late pattern as being either an artifact, or a more sensitive reading of the endogenous *Shh* pattern. Despite this, it appears to the most likely explanation.

The thoughts which have led to the proposal that hemimelia-luxate mutants are generally repressors, also suggested that *Shh* must have a predisposition for activity in that anterior patch where it is so commonly mis-expressed. These two ideas suggest that a *cis*-acting repressive element may occur near the *Shh* gene, through which the genes of the hemimelia-luxate mutants act directly or indirectly, to prevent ectopic *Shh* expression (similar to the repressor element in the *Hoxb-1* (Studer *et al.*, 1994)). If the transgene were to damage this element, by inserting into the *Shh* locus, it could cause all three effects seen (Fig. 6.5c). The non-functional repressor would cause *Shh* to be activated in the anterior patch. This would in effect cause a phenocopy of hemimelia-luxate mutants, but only in the limb. Lastly, the PLAP reporter gene would be able to pick-up regulatory influences from the *Shh* locus, and mirror the expression of *Shh*. The fact that PLAP staining was not found in the floorplate or notochord suggests that the limb regulatory sequences may be distant from the remaining *Shh* enhancers. Interestingly, this situation would be a demonstration of an enhancer being shared between two promoters, and one in which there is only a single copy of the putative enhancer, although as in the case of van der Hoeven *et al* (1996) it is a completely artificial situation.

6.7 Recent and future work

The discovery of two integration sites raised questions about the nature of the mutation. As mentioned before, the expression pattern of PLAP indicated the involvement of the transgene in the phenotype-causing mutation. Nevertheless, it was possible that the two sites represented the ends of a chromosomal inversion, induced by the recombination event. Recent FISH analysis from R. Hill's laboratory, has shown that the segment of DNA between the two integration sites is unlikely to be inverted, as the relative positions of *lx* and *msx-1* between the two sites is unaltered. It is possible therefore that the two integration events were independent, and that their proximity is due to

coincidence. If this is the case, then probably only one of these sites is responsible for the mutation (although both sites may contribute to PLAP expression in rhombomere 4), and it should be possible through breeding experiments to separate them. This work is being initiated in R. Hill's laboratory.

The recent FISH analysis has also shown that the proximal of the two integration sites is located close to the *Shh* locus - estimated at 1000kb by interphase double-labelled FISH. This distance seems to be very large for *cis*-effects, but it is only a rough estimate, and if it turns out to be close enough to affect the regulation of the *Shh* gene, it would fit very well with the theories described above. The fact that *Shh*-like expression was only seen in the limb, had already raised the possibility that the limb enhancers (and repressors) are some distance from the remaining regulatory elements.

The next phase of work to be performed on *sasquatch*, will include pulsed-field gel analysis to determine the distance between the *sas* integration and the *Shh* locus. This work would be helped if the two integration sites can be separated, and a breeding program is being set-up to attempt this. Further proof that the transgene is directly responsible for the phenotype will come from testing the prediction that *sasquatch* homozygotes should display anterior PLAP activity in forelimbs as well as hindlimbs. Also, whole-mount *in-situ* hybridisation analysis will be performed using *Fgf-4* and *Gli-3* as probes. In past analysis of ectopic *Shh* duplications, the pattern of *Fgf-4* has always also been changed (Laufer *et al.*, 1994), and there have been recent suggestions that the *Gli-3* gene (which is mutated in the *Xt* mutant) is a direct repressor of *Shh*, whose normal expression domain covers all of the limb-bud except for the ZPA (Buscher *et al.*, 1996). If *sas* causes a change to *Shh* directly, we would expect that *Gli-3* expression is unaltered in these mutants.

BIBLIOGRAPHY

Akam, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101,** 1-22.

Akam, M. (1989). Hox and HOM: homologous gene clusters in insects and vertebrates. *Cell* **57,** 347-349.

Akarsu, A. N., Stoilov, I., Yilmaz, E., Sayli, B. S., and Sarfarazi, M. (1996). Genomic structure of *HOXD13* gene: a nine polyalanine duplication causes synpolydactyly in two unrelated families. *Hum. Mol. Genet.* **5,** 945-952.

Alkema, M., van der Lugt, N. M. T., Bobeldijk, R. C., Berns, A., and van Lohuizen, M. (1995). Transformation of axial skelton due to overexpression of *bmi-1* in transgenic mice. *Nature* **374,** 724-727.

Bateson, W. (1894). "Materials for the study of variation." MacMillan & Co.,

Beeman, R., Stuart, J., Haas, M., and Denell, R. (1989). Genetic analysis of the homeotic gene complex (HOM-C) in the beetle *Tribolium castaneum*. *Devl Biol.* **133,** 196-209.

Bittner, D., De Robertis, E., and Cho, K. (1993). Characterization of the *Xenopus Hox2.4* gene and identification of control elements in its intron. *Developmental Dynamics* **196,** 11-24.

Boncinelli, E., Simeone, A., Acampora, D., and Mavilio, F. (1991). HOX gene activation by retinoic acid. *TIG* **7,** 329-334.

Brunk, B., Martin, E., and Adler, P. (1991). *Drosophila* genes *Posterior Sex Combs* and *Supressor two of zeste* encode proteins with homology to the murine *bmi-1* oncogene. *Nature* **353,** 351-353.

Bryant, S. V., French, V., and Bryant, P. J. (1981). Distal Regeneration and Symmetry. *Science* **212,** 993-1002.

Buscher, D., Bosse, B., and Ruther, U. (1996). Evidence for genetic control of *Shh* by *Gli3* in mouse limb development. *In* "Mouse Molecular Genetics", pp. 10, Cold Spring Harbor Laboratory.

Capecchi, M. (1989). The new mouse genetics: altering the genome by gene targeting. *TIG* **5,** 70-76.

Chambon, P., Zelent, A., Petkovich, M., Mendelsohn, C., Leroy, P., Krust, A., Kastner, P., and Brand, N. (1991). The family of retinoic acid nuclear receptors. *In* "Retinoids: 10 years on" (J. H. Saurat, Ed.), pp. 10-27. Karger, Basel.

Chan, C.-S., Rastelli, L., and Pirrotta, V. (1994). A *Polycomb* response element in the *Ubx* gene that determines an epigenetically inherited state of repression. *EMBO J.* **13,** 2553-2564.

Chan, D. C., Laufer, E., Tabin, C., and Leder, P. (1995). Polydactylous limbs in Strong's Luxoid mice result from ectopic polarizing activity. *Development* **121,** 1971-8.

Chang, D. T., Lopez, A., von Kessler, D. P., Chiang, C., Simandl, B. K., Zhao, R., Seldin, M. F., Fallon, J. F., and Beachy, P. A. (1994). Products, genetic linkage and limb patterning activity of a murine hedgehog gene. *Development* **120,** 3339-53.

Charite, J., de Graaf, W., Shen, S., and Deschamps, J. (1994). Ectopic expression of *Hoxb-8* causes duplication of the ZPA in the forelimb and homeotic transformation of axial structures. *Cell* **78,** 589-601.

Charite, J., de Graaf, W., Vogels, R., Meijlink, F., and Deschamps, J. (1995). Regulation of the *Hoxb-8* gene: synergism between multimerized *cis*-acting elements increases responsiveness to positional information. *Developmental Biology* **171,** 294-305.

Chavrier, P., Vesque, C., Galliot, B., Vigneron, M., Dolle, P., Duboule, D., and Charnay, P. (1990). The segment-specific gene *Krox-20* encodes a transcription factor with binding sites in the promoter of the *Hox 1.4* gene. *EMBO J.* **9,** 1209-1218.

Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R., and Charnay, P. (1988). A gene encoding a protein with zinc fingers is activated during Go/G1 transition in cultured cells. *EMBO J.* **7,** 29-35.

Chisaka, O., and Capecchi, M. (1991). Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene {Ihox1.5}. *Nature* **350,** 473-479.

Cho, K., Goetz, J., Wright, C., Fritz, A., Hardwicke, J., and De Robertis, E. (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.

Coelho, C. N., Upholt, W. B., and Kosher, R. A. (1992). Role of the chicken homeobox-containing genes GHox-4.6 and GHox-8 in the specification of positional identities during the development of normal and polydactylous chick limb buds. *Development* **115,** 629-37.

Cohn, M. J., and Tickle, C. (1996). Limbs: a model for pattern formation within the vertebrate body plan. *TIG* **12,** 253-257.

Condie, B., and Capecchi, M. (1993). Mice homozygous for a targeted disruption of *Hoxd-3*(*Hox-4.1*) exhibit anterior transformations of the first and second cervical vertebrae, the atlas and axis. *Development* **119,** 579-595.

Conlon, R. A., and Rossant, J. (1992). Exogenous retinoic acid rapidly induces anterior ectopic expression of murine *Hox-2* genes *in vivo*. *Development* **116,** 357-368.

Crossley, P. H., and Martin, G. R. (1995). The mouse Fgf-8 gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development* **121,** 439-451.

Crossley, P. H., Minowada, G., MacArthur, C. A., and Martin, G. R. (1996). Roles for FGF8 in the induction, initiation, and maintenance of chick limb development. *Cell* **84,** 127-36.

Davidson, D., Graham, E., Sime, C., and Hill, R. (1988). A gene with sequence similarity to Drosophila engrailed is expressed during the development of the neural tube and vertebrae in the mouse. *Development* **104,** 305-316.

Davis, A., Witte, D., Hsieh, L., Potter, S., and Capecchi, M. (1995). Absence of radius and ulna in mice lacking *Hoxa-11* and *Hoxd-11*. *Nature* **375,** 791-795.

Davis, S., Gale, N. W., Aldrich, T. H., Maisonpierre, P. C., Lhotak, V., Pawson, T., Goldfarb, M., and Yancopoulos, G. D. (1994). Ligands for EPH-related receptor tyrosine kinases that require membrane attachment or clustering for activity. *Science* **266,** 816-819.

Dolle, P., Dierich, A., LeMeur, M., Schimmang, T., Schuhbaur, B., Chambon, P., and Duboule, D. (1993). Disruption of the *Hoxd-13* gene induces localized heterochrony leading to mice with neotenic limbs. *Cell* **75,** 431-441.

Dolle, P., Izpisua-Belmonte, J. C., Falkenstein, H., Renucci, A., and Duboule, D. (1989). Coordinate expression of the murine Hox-5 complex homoeobox-containing genes during limb pattern formation. *Nature* **342,** 767-772.

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 of homeobox genes." IRL Press, Oxford.

Duboule, D., and Dolle, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *EMBO J.* **8,** 1497-1505.

Duboule, D., and Morata, G. (1994). Colinearity and functional heiarchies among genes of the homeotic complexes. *TIG* **10,** 358-364.

Durston, A., Timmermans, J., Hage, W., Hendriks, H., de Vries, N., Heideveld, M., and Nieuwkoop, P. (1989). Retinoic acid causes an anteroposterior transformation in the developing central nervous system. *Nature* **340,** 140-144.

Echelard, Y., Epstein, D., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. A., and McMahon, A. P. (1993). *Sonic hedgehog*, a Member of a Family of Putative Signalling Molecules, Is implicated in the Regulation of CNS Polarity. *Cell* **75,** 1417-1430.

Faiella, A., Zappavigna, V., Mavilio, F., and Boncinelli, E. (1994). Inhibition of retinoic acid-induced activation of 3' human *HOXB* genes by antisense oligonucleotides affects sequential activation of genes located upstream in the four *HOX* clusters. *Proc. Natl. Acad. Sci. USA* **91,** 5335-5339.

Fallon, J. F., Lopez, A., Ros, M. A., Savage, M. P., Olwin, B. B., and Simandl, B. K. (1994). FGF-2: apical ectodermal ridge growth signal for chick limb development. *Science* **264,** 104-7.

Fan, C.-M., Porter, J., Chiang, C., Chang, D. T., Beachy, P., and Tessier-Lavigne, M. (1995). Long-range sclerotome induction by Sonic hedgehog: direct role of the amino-terminal cleavage product and modulation by the cyclic AMP signalling pathway. *Cell* **81,** 457-465.

Fjose, A., McGinnis, W. J., and Gehring, W. J. (1985). Isolation of a homeobox containing gene from the engrailed region of *Drosophila* and the spatial distribution of its transcripts. *Nature* **313,** 284-289.

Forsthoefel, P. F. (1958). The skeletal effects of the luxoid gene in the mouse, including its interactions with the luxate gene. *Journal of Morphology* **102,** 247-287.

Forsthoefel, P. F. (1962). Genetics and manifold effects of Stong's Luxoid gene in the mouse, including its interactions with Green's Luxoid and Carter's Luxate genes. *Journal of Morphology* **110,** 391-420.

Frasch, M., Chen, X., and Lufkin, T. (1995). Evolutionary-conserved enhancers direct region-specific expression of the murine*Hoxa-1* and *Hoxa-2* loci in both mice and *Drosophila*. *Development* **121,** 957-974.

French, V., Bryant, P. J., and Bryant, S. V. (1976). Pattern regulation in epimorphic fields. *Science* **193,** 969-981.

Frohman, M. A., Boyle, M., and Martin, G. R. (1990). Isolation of the mouse *Hox-2.9* gene; analysis of embryonic expression suggests that positional information along the anterior-posterior axis is specified by mesoderm. *Development* **110,** 589-607.

Galloni, M., Gyurkovics, H., Schedl, P., and Karch, F. (1993). The bluetail transposon: evidence for independent *cis*-regulatory domains and domain boundaries in the bithorax complex. *EMBO J.* **12,** 1087-1097.

Gaunt, S. J., Krumlauf, R., and Duboule, D. (1989). Mouse homeo-genes within a subfamily, Hox-1.4, -2.6 and -5.1, display similar anteroposterior domains of expression in the embryo, but show stage- and tissue-dependent differences in their regulation. *Development.* **107,** 131-141.

Gaunt, S. J., and Singh, P. B. (1990). Homeogene expression patterns and chromosomal imprinting. *Trends. Genet.* **6,** 208-212.

Gehring, W., Muller, M., Affolter, M., Percival-Smith, A., Billeter, M., Qian, Y., Otting, G., and Wuthrich, K. (1990). The structure of the homeodomain and its functional implications. *TIG* **6,** 323-329.

Gendron-Maguire, M., Mallo, M., Zhang, M., and Gridley, T. (1993). *Hoxa-2* mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. *Cell* **75,** 1317-1331.

Goebl, M. G. (1992). The *bmi-1* and *mel-18* gene products define a new family of DNA-binding proteins involved in cell proliferation and tumorigenesis. *Cell* **66,** 623.

Gonzalez-Reyes, A., and Morata, G. (1990). The developmental effect of overexpressing a *Ubx* product in *Drosophila* embryos is dependent on its interactions with other homeotic products. *Cell* **61,** 515-522.

Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The murine and *Drosophila* homeobox clusters have common features of organisation and expression. *Cell* **57,** 367-378.

Gruneberg. (1963). "The Pathology of Development." Blackwell Scientific Publications, Oxford.

Gruss, P., and Walther, C. (1992). *Pax* in development. *Cell* **69,** 719-722.

Gurdon, J. B., Lemaire, P., and Kato, K. (1993). Community effects and related phenomena in development. *Cell* **75,** 831-834.

Gyurkovics, H., Gausz, J., Kummer, J., and Karch, F. (1990). A new homeotic mutation in the Drosophila bithorax complex removes a boundary separating two domains of regulation. *EMBO J.* **9,** 2579-2585.

Han, K., and Manley, J. L. (1993a). Functional domains of the *Drosophila* Engrailed protein. *EMBO J.* **12,** 2723-2733.

Han, K., and Manley, J. L. (1993b). Transcriptional repression by the *Drosophila* Even-skipped protein: definition of a minimal repression domain. *Genes & Dev.* **7,** 491-503.

Hayasaka, I., Nakatsuka, T., Fujii, T., Naruse, I. (1980). Polydactyly Nagoya *Pdn*: a new mutant gene in the mouse. *Experimental animals* **29,** 391-395.

Henikoff, S. (1990). Position-effect variegation after 60 years. *TIG* **6,** 422-426.

Hinchliffe. (1977). "Vertebrate limb and somite morphogenesis." Cambridge University Press, Cambridge.

Hinchliffe, and Johnson. (1980). "The Development of the Vertebrate Limb." Clarendon Press, Oxford.

Hinchliffe, J. R., and Griffiths, P. J., The pechondrogenic patterns tetrapod limb development and their phylogenetic significance. *Development and evolution.* (ed. Goodwin, B. C., Holder, N., Wylie, C. C.) Cambridge University Press.

Hinchliffe, J. R., and Sansom, A. (1985). The distribution of the polarizing zone (ZPA) in the legbud of the chick embryo. *J. Ebryol. exp. Morph.* **86,** 169-175.

Honig, L. S., and Summerbell, D. (1985). Maps of strength of positional signalling activity in the developing chick wing bud. *J. Embryol. exp. Morph.* **87,** 163-174.

Hornbruch, A., and Wolpert, L. (1970). Cell divisions in the early growth and morphogenesis of the chick limb. *Nature* **226,** 764-766.

Hornbruch, A., and Wolpert, L. (1991). The spatial and temporal distribution of polarising activity in the flank of the pre-limb-bud stages in the chick embryo. *Development* **111,** 725-731.

Hunt, P., Gulisano, M., Cook, M., Sham, M., Faiella, A., Wilkinson, D., Boncinelli, E., and Krumlauf, R. (1991). A distinct *Hox* code for the branchial region of the head. *Nature* **353,** 861-864.

Ingham, P. (1985). The regulation of the bithorax complex. *TIG* **1,** 112-116.

Ingham, P. (1988). The molecular genetics of embryonic pattern formation in *Drosophila. Nature* **335,** 25-34.

Ingham, P. W. (1983). Differential expression of *bithorax complex* genes in the absence of the *extra sex combs* and *trithorax* genes. *Nature* **306,** 591-593.

Izpisua-Belmonte, J.-C., Ede, D. A., Tickle, C., and Duboule, D. (1992). The mis-expression of posterior *Hox-4* genes in *talpid* (*ta3*) mutant wings correlates with the absence of anteroposterior polarity. *Development* **114,** 959-963.

Izpisua-Belmonte, J.-C., Tickle, C., Dolle, P., Wolpert, L., and Duboule, D. (1991). Expression of homeobox *Hox-4* genes and the specification of position in chick wing development. *Nature* **350,** 585-589.

Jeannottee, L., Lemieux, m., Charron, J., Poirier, F., and Robertson, E. (1993). Specification of axial identity in the mouse: role of the *Hoxa-5*(*Hox1.3*) gene. *Genes Dev.* **7,** 2085-2096.

Johnson, D. R. (1986). The Genetics of the Skeleton. Clarendon Press. Oxford.

Jones, C. M., Lyons, K., and Hogan, B. (1991). Involvement of *Bone Morphogenetic Protein-4* (*BMP-4*) and *Vgr-1* in morphogenesis and neurogenesis in the mouse. *Development* **111,** 531-542.

Jurgens, G. (1985). A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature* **316,** 153-155.

Kaufman, T. C., Lewis, R., and Wakimoto, B. (1980). Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: the homeotic gene complex in polytene chromosomal interval 84A,B. *Genetics.* **94,** 115-133.

Kennison, J. A. (1993). Transcriptional activation of *Drosophila* homeotic genes from distant regulatory elements. *TIG* **9,** 75-79.

Kessel, M., and Gruss, P. (1991). Homeotic transformations of murine prevertebrae and concommitant alteration of *Hox* codes induced by retinoic acid. *Cell* **67,** 89-104.

Keynes, R., and Krumlauf, R. (1994). *Hox* genes and regionalization of the nervous system. *Ann. Rev. Neurosci.* **17,** 109-132.

Kieny, M. (1968). Variation de la capacite inductrice du mesoderm et de la competence de l'ectoderm au cours de l'induction primaire du bourgeon de membre chez l'embryon de poulet. *Arch. Anat. Microsc. Morphol. Exp.* **57,** 401-418.

Knudsen, T. B., and Kochar, D. M. (1981). The role of morphogenetic cell death during abnormal limb-bud outgrowth in mice heterozygous for the dominant mutation *Hemimelia-extra toe* (*Hm*). *J. Embryol. exp. Morph.* **65,** 289-307.

Kozak, M. (1989). The Scanning Model for Translation: An Update. *Journal of Cell Biology* **108,** 229-241.

Krauss, S., Concordet, J.-P., and Ingham, P. W. (1993). A Functionally Conserved Homolog of the Drosophila Segment Polarity Gene *hh* Is Expressed in Tissues with Polarizing Activity in Zebrafish Embryos. *Cell* **75,** 1431-1444.

Krumlauf, R. (1994). Hox genes in vertebrate development. *Cell* **78,** 191-201.

Kuner, J., Nakanishi, M., Ali, Z., Drees, B., Gustavson, E., Theis, J., Kauver, L., Kornberg, T., and O'Farrell, P. (1985). Molecular cloning of *engrailed*, a gene involved in the development of pattern in *Drosophila melanogaster*. *Cell* **42,** 309-316.

Kuziora, M. A., and McGinnis, W. (1988). Different transcripts of the *Drosophila Abd-B* gene correlate with distinct genetic sub-functions. *EMBO J.* **7,** 3233-3244.

Langston, A. W., and Gudas, L. J. (1992). Identification of a retinoic acid responsive enhancer 3' of the murine homeobox gene {IHox-1.6}. *Mech. Dev.* **38,** 217-228.

Laufer, E., Nelson, C. E., Johnson, R. L., Morgan, B. A., and Tabin, C. (1994). Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. *Cell* **79,** 993-1003.

Le Mouellic, H., Lallemand, Y., and Brulet, P. (1992). Homeosis in the mouse induced by a null mutation in the homeo-gene *Hox-3.1*. *Cell* **69,** 251-264.

Lewis, E. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276,** 565-570.

Lewis, E. B. (1994). Homeosis: the first 100 years. *TIG* **10,** 341-343.

Li, X., and Noll, M. (1994). Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the *Drosophila* embryo. *Junk* **13(2),** 400-406.

Lipshitz, H. D., Peattie, D. A., and Hogness, D. S. (1987). Novel transcripts from the *Ultrabithorax* domain of the bithorax complex. *Genes Dev.* **1,** 307-322.

Lopez-Martinez, A., Chang, D. T., Chiang, C., Porter, J. A., Ros, M. A., Simandl, B. K., Beachy, P. A., and Fallon, J. F. (1995). Limb-patterning activity and restricted posterior localisation of the amino-terminal product of Sonic hedgehog cleavage. *Current Biology* **5,** 791-796.

Lufkin, T., Dierich, A., LeMeur, M., Mark, M., and Chambon, P. (1991). Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression. *Cell* **66,** 1105-1119.

Lumsden, A., and Keynes, R. (1989). Segmental patterns of neuronal development in the chick hindbrain. *Nature* **337,** 424-428.

Lyon, M., and Searle, A. (1989). "Genetic variants and strains of the laboratory mouse." Oxford University Press.

Lyons, K. M., Pelton, R. W., and Hogan, B. L. M. (1990). Organogenesis and pattern formation in the mouse: RNA distribution patterns suggest a role for *Bone Morphogenic Protein-2A* (*BMP-2A*). *Development* **109,** 833-844.

MacCabe, J. A., Errick, J., and Saunders, J. W. (1974). Ectodermal control of the dorsoventral axis in the leg bud of the chick embryo. *Dev. Biol.* **39,** 69-82.

MacDonald, P., Ingham, P., and Struhl, G. (1986). Isolation, structure and expression of even-skipped: a second pair-rule gene of *Drosophila* containing a homeobox. *Cell* **47,** 721-734.

Maden, M. (1983). A test of the predictions of the boundary model regarding supernumerary limb structure. *J. Embryol. exp. Morph.* **76,** 147-155.

Maden, M., and Turner, R. N. (1978). Supernumerary limbs in the axolotl. *Nature* **273,** 232-235.

Mahmood, R., Bresnick, J., Hornbruch, A., Mahony, C., Morton, N., Colquhoun, K., Martin, P., Lumsden, A., Dickson, C., and Mason, I. (1995). A role for FGF-8 in the initiation and maintenance of vertebrate limb bud outgrowth. *Current Biology* **5,** 797-806.

Mangelsdorf, D. J., Borgmeyer, U., Heyman, R. A., Yang Zhou, J., Ong, E. S., Oro, A. E., Kakizuka, A., and Evans, R. M. (1992). Characterization of three RXR genes that mediate the action of 9-cis retinoic acid. *Genes Dev.* **6,** 329-344.

Marshall, H., Nonchev, S., Sham, M. H., Muchamore, I., Lumsden, A., and Krumlauf, R. (1992). Retinoic acid alters hindbrain Hox code and induces transformation of rhombomeres 2/3 into a 4/5 identity. *Nature* **360,** 737-741.

Marshall, H., Studer, M., Popperl, H., Aparicio, S., Kuroiwa, A., Brenner, S., and Krumlauf, R. (1994). A conserved retinoic acid response element required for early expression of the homeobox gene *Hoxb-1*. *Nature* **370,** 567-571.

Masuya, H., Sagai, T., Wakana, S., Moriwaki, K., and Shiroishi, T. (1995). A duplicated zone of polarizing activity in polydactylous mouse mutants. *Genes & Dev.* **9,** 1645-1653.

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A., and Gehring, W. (1984). A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell* **37,** 403-408.

McGinnis, W., and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68,** 283-302.

Meinhardt, H. (1983). A boundary model for pattern formation in vertebrate limbs. *J. Embryol. exp. Morph.* **76,** 115-137.

Morgan, B., Izpisua-Belmonte, J.-C., Duboule, D., and Tabin, C. (1992). Targeted misexpression of *Hox-4.6* in the avian limb bud causes apparent homeotic transformations. *Nature* **358,** 236-239.

Morgan, B. A., and Tabin, C. (1994). Hox genes and growth: early and late roles in limb bud morphogenesis. *Dev Suppl* , 181-6.

Murray, J. D. (1989). Mathematical Biology. Springer-Verlag

Murray, J. D., Oster, G. F., Harris, A. K. (1983). A mechanical model for morphogenesis. *J. Math. Biol.* **17,** 125-129.

Morriss-Kay, G. M., Murphy, P., Hill, R. E., and Davidson, D. R. (1991). Effects of retinoic acid excess on expression of *Hox-2.9* and *Krox-20* and on morphological segmentation in the hindbrain of mouse embryos. *EMBO J.* **10,** 2985-2995.

Muller, J., and Bienz, M. (1992). Sharp anterior boundary of homeotic gene expression conferred by the fushi tarazu protein. *Junk* **11(10),** 3653-3661.

Muller, J., Gaunt, S., and Lawrence, P. A. (1995). Function of the Polycomb protein is conserved in mice and flies. *Development* **121,** 2847-2852.

Muller, M., Affolter, M., Leupin, W., Otting, G., Wuthrich, K., and Gehring, W. J. (1988). Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *Junk* **7,** 4299-4304.

Muragaki, Y., Mundlos, S., Upton, J., and Olsen, B. R. (1996). Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. *Science* **272,** 548-51.

Murphy, P., Davidson, D. R., and Hill, R. E. (1989). Segment-specific expression of a homeobox-containing gene in the mouse hindbrain. *Nature* **341,** 156-159.

Murphy, P., and Hill, R. E. (1991). Expression of the mouse *labial*-like homeobox-containing genes, *Hox 2.9* and *Hox 1.6*, during segmentation of the hindbrain. *Development* **111,** 61-74.

Nieuwkoop, P. D. (1952). Activation and organisation of the central nervous system in amphibians. *J. Exp. Zool.* **120,** 1-108.

Niswander, L., and Martin, G. R. (1992). *Fgf-4* expression during gastrulation, myogenesis, limb and tooth development in the mouse. *Developement* **114,** 755-768.

Niswander, L., and Martin, G. R. (1993). FGF-4 and BMP-2 have opposite effects on limb growth. *Nature* **361,** 68-71.

Niswander, L., Tickle, C., Vogel, A., Booth, I., and Martin, G. R. (1993). FGF-4 replaces the apical ectodermal ridge and directs outgrowth and patterning of the limb. *Cell* **75,** 579-87.

Nusslein-Volhard, C., and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287,** 795-801.

Orlando, V., and Paro, R. (1993). Mapping polycomb-repressed domains in the Bithorax complex using *in vivo* formaldehyde cross-linked chromain. *Cell* **75,** 1187-1198.

Oster, G. F., Murray, J. D., Harris, A. K. (1983). Mechanical aspects of mesenchymal morphogenesis. *J. Embryol. exp. Morph.* **78,** 83-125.

Oster, G. F., Murray, J. D., Maini, P. K. (1985). A model for chondrogenic condensations in the developing limb: the role of extracellular matrix and cell tractions. *J. Embryol. exp. Morph* **89,** 93-112.

Otting, G., Qian, Y. Q., Billeter, M., Muller, M., Affolter, M., Gehring, W. J., and Wuthrich, K. (1990). Protein--DNA contacts in the structure of a homeodomain--DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *Junk* **9,** 3085-3092.

Otting, G., Qian, Y. Q., Muller, M., Affolter, M., Gehring, W., and Wuthrich, K. (1988). Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. *Junk* **7,** 4305-4309.

Papalopulu, N., Clarke, J., Bradley, L., Wilkinson, D., Krumlauf, R., and Holder, N. (1991a). Retinoic acid causes abnormal development and segmental patterning of the anterior hindbrain in *Xenopus* embryos. *Development* **113,** 1145-1159.

Papalopulu, N., Lovell-Badge, R., and Krumlauf, R. (1991b). The expression of murine *Hox-2* genes is dependent on the differentiation pathway and displays collinear sensitivity to retinoic acid in F9 cells and *Xenopus* embryos. *Nucl. Acids Res.* **19,** 5497-5506.

Paro, R. (1990). Imprinting a determined state into the chromatin of *Drosophila. TIG* **6,** 416-421.

Paro, R., and Hogness, D. (1991). The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila. Proc. Natl. Acad. Sci. USA* **88,** 263-267.

Parr, B. A., and McMahon, A. P. (1995). Dorsalizing signal Wnt-7a required for normal polarity of D-V and A-P axes of mouse limb. *Nature* **374,** 350-3.

Patel, K., Nittenberg, R., D, D. S., Irving, C., Burt, D., Wilkinson, D. G., and Tickle, C. (1996). Expression and regulation of Cek-8, a cell to cell signalling receptor in developing chick limb buds. *Development* **122,** 1147-55.

Pautou, M. P., and Kieny, M. (1973). Interaction ecto-mesodermique dans l'establissement de la polarite dorso-ventrale du pied de l'embryon du poulet. *C. R. Acad. Sci. Paris Serie D* **277,** 1225-1228.

Pearce, J., Singh, P., and Gaunt, S. (1992). The mouse has a *Polycomb*-like chromobox gene. *Development* **114,** 921-929.

Peifer, M., Karch, F., and Bender, W. (1987). The bithorax complex: control of segment identity. *Genes & Dev.* **1,** 891-898.

Popperl, H., Bienz, M., Studer, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S., and Krumlauf, R. (1995). Segmental expression of *Hoxb-1* is controlled by a highly conserved autoregulatory loop dependent upon *exd/Pbx*. *Cell* **81,** 1031-1042.

Popperl, H., and Featherstone, M. S. (1992). An autoregulatory element of the murine *Hox-4.2* gene. *EMBO J.* **11,** 3673-3680.

Popperl, H., and Featherstone, M. S. (1993). Identification of a retinoic acid repsonse element upstream of the murine *Hox-4.2* gene. *Mol. Cell. Biol.* **13,** 257-265.

Puschel, A., Balling, R., and Gruss, P. (1991). Separate elements cause lineage restriction and specify boundaries of *Hox-1.1* expression. *Development* **112,** 279-288.

Qian, Y. Q., Billeter, M., Otting, G., Muller, M., Gehring, W. J., and Wuthrich, K. (1989). The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* **59,** 573-580.

Ramirez-Solis, R., Zheng, H., Whiting, J., Krumlauf, R., and Bradley, A. (1993). *Hox-B4* (*Hox-2.6*) mutant mice show homeotic transformation of cervical vertebra and defects in the closure of the sternal rudiments. *Cell* **73,** 279-294.

Reuter, G., Giarre, M., Farah, J., Gausz, J., Spierer, A., and Spierer, P. (1990). Dependence of position-effect variegation in *Drosophila* on dose of a gene encoding an unusual zinc-finger protein. *Nature* **344,** 219-223.

Riddle, R. D., Ensini, M., Nelson, C., Tsuchida, T., Jessell, T. M., and Tabin, C. (1995). Induction of the LIM homeobox gene Lmx1 by WNT7a establishes dorsoventral pattern in the vertebrate limb. *Cell* **83,** 631-40.

Riddle, R. D., Johnson, R. L., Laufer, E., and Tabin, C. (1993). *Sonic hedgehog* mediates the polarizing activity of the ZPA. *Cell* **75,** 1401-1416.

Riley, G., Jorgensen, E., Baker, R., and Garber, R. (1991). Positive and negative control of the Antennapedia promoter P2. *Development Supplement* **1,** 177-185.

Roelink, H., Augsburger, A., Heemskerk, J., Korzh, V., Norlin, S., Ruiz i Altaba, A., Tanabe, Y., Placzek, M., Edlund, T., Jessell, M., and Dodd, J. (1994). Floor Plate and Motor Neuron Induction by *vhh-1*, a Vertebrate homolog of *hedgehog* Expressed by the Notochord. *Cell* **76,** 761-775.

Roelink, H., Porter, J. A., Chiang, C., Tanabe, Y., Chang, D. T., Beachy, P. A., and Jessell, T. M. (1995). Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of Sonic hedgehog auttoproteolysis. *Cell* **81,** 445-455.

Ruiz i Altaba, A., and Jessell, T. (1991). Retinoic acid modifies the pattern of cell differentiation in the central nervous system of neurula stage *Xenopus* embryos. *Development* **112,** 945-958.

Sanchez-Herrero, E., and Akam, M. (1989). Spatially ordered transcription of regulatory DNA in the bithorax complex of *Drosophila*. *Development* **107,** 321-329.

Saunders, J. (1948). The proximo-distal sequence of origin of the parts of the chick wing and the role of the ectoderm. *J. Exp. Zool.* **108,** 363-404.

Saunders, J. W. and Gasseling, M. T. (1968). Ectodermal-mesenchymal interactions in the origin of limb asymmetry. *Epithelial-mesenchymal interactions* (ed. R. Fleischmajor and R. F. Billingham), pp. 78-97. Williams & Wilkins, Baltimore.

Scott, M. P. (1992). Vertebrate Homeobox Gene Nomenclature. *Cell* **71,** 551-553.

Scott, M. P., Tamkun, J. W., and Hartzell, G. W., 3d. (1989). The structure and function of the homeodomain. *Biochim. Biophys. Acta.* **989,** 25-48.

Searls, R. L., and Janners, M. Y. (1971). The initiation of limb bud outgrowth in the embryonic chick. *Dev. Biol.* **24,** 198-213.

Sham, M. H., Vesque, C., Nonchev, S., Marshall, H., Frain, M., Das Gupta, R., Whiting, J., Wilkinson, D., Charnay, P., and Krumlauf, R. (1993). The zinc finger gene *Krox20* regulates *HoxB2* (*Hox2.8*) during hindbrain segmentation. *Cell* **72,** 183-196.

Shimell, M. J., Simon, J., Bender, W., and O'Connor, M. B. (1994). Enhancer point mutation result in a homeotic transformation in *Drosophila*. *Science* **264,** 968-971.

Shubin, N., Alberch, P. (1986). A morphogenetic approach to the origin and basic organisation of the tetrapod limb. *Evolutionary biology* **20,** 319-387. Plenum. New York.

Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli, E., and Mavilio, F. (1990). Sequential activation of HOX2 homeobox genes by retinoic acid in human embryonal carcinoma cells. *Nature* **346,** 763-766.

Sive, H., Draper, B., Harland, R., and Weintraub, H. (1990). Identification of a retinoic acid-sensitive period during primary axis formation in *Xenopus laevis*. *Genes Dev.* **4,** 932-942.

Slack, J. M. W., Holland, P. W. H., and Graham, C. F. (1993). The zootype and the phylotypic stage. *Nature* **361,** 490-492.

Small, K. S., and Potter, S. (1993). Homeotic transformations and limb defects in *Hoxa-11* mutant mice. *Genes Dev.* **7,** 2318-2328.

Smith, J. C. (1980). The time required for positional siganlling in the chick wing bud. *J. Embryol. exp. Morph.* **60,** 321-328.

Sordino, D., van der Hoeven, F., and Duboule, D. (1995). *Hox* gene expression in teleost fins and the origin of vertebrate digits. *Nature* **375,** 678-681.

Struhl, G. (1981). A gene product required for the correct initiation of segmental determination in *Drosophila*. *Nature* **293,** 36-41.

Struhl, G. (1984). Splitting the bithorax complex of Drosophila. *Nature* **308,** 454-457.

Struhl, G., and Akam, M. (1985). Altered distributions of *Ultrabithorax* transcripts in *extra sex combs* mutant embryos of *Drosophila*. *EMBO J.* **4,** 3259-3264.

Studer, M., Popperl, H., Marshall.H., Kuroiwa, A., and Krumlauf, R. (1994). Role of a conserved retinoic acid response element in rhombomere restriction of *Hoxb-1*. *Science* **265,** 1728-1732.

Summerbell, D. (1974). A quantitative analysis of the effect of excision of the AER from the chick limb-bud. *J. Embryol. exp. Morph.* **32,** 651-660.

Summerbell, D. (1979). The zone of polarizing activity: evidence for a role in normal chick limb morphogenesis. *J. Embryol. exp. Morph.* **50,** 217-233.

Summerbell, D., Lewis, J. H., and Wolpert, L. (1973). Positional information in chick limb morphogenesis. *Nature* **244,** 492-496.

Tabin, C. (1992). Why we have (only) five fingers per hand: Hox genes and the evolution of paired limbs. *Development* **116,** 289-296.

Thaller, C., and Eichele, G. (1987). Identification and spatial distribution of retinoids in the developing chick limb bud. *Nature* **327,** 625-628.

Thaller, C., and Eichele, G. (1988). Characterisation of retinoid metabolism in the developing chick limb bud. *Development* **103,** 473-483.

Tickle, C. (1981). The number of polarizing region cells required to specify additional digits in the developing chick wing bud. *Nature* **289,** 295-298.

Tickle, C., Alberts, B., Wolpert, L., and Lee, J. (1982). Local application of retinoic acid to the limb bud mimics the action of the polarising region. *Nature* **296,** 564-566.

Tickle, C., Summerbell, D., and Wolpert, L. (1975). Positional signalling and specification of digits in chick limb morphogenesis. *Nature* **254,** 199-202.

van der Hoeven, F., Zakany, J., and Duboule, D. (1996). Gene transpositions in the *HoxD* Complex reveal a hierarchy of regulatory controls. *Cell* **85,** 1025-1035.

van der Lugt, N. M. T., Alkema, M., Berns, A., and Deschamps, J. (1996). The *Polycomb*-group homolog Bmi-1 is a regulator of murine *Hox* gene expression. *Mech. Dev.* **58,** 153-164.

van der Lugt, N. M. T., Domen, J., Linders, K., van Room, M., Robanus-Maandag, E., te Riele, H., van der Valk, M., Deschamps, J., Sofroniew, M., van Lohuizen, M., and Berns, A. (1994). Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the *Bmi-1* proto-oncogene. *Genes Dev.* **8,** 757-769.

van Lohuizen, M., Frasch, M., Wientjens, E., and Berns, A. (1991). Sequence similarity between the mammalian *bmi-1* proto-oncogene and the *Drosophila* regulatory genes *Psc* and *Su(z)2*. *Nature* **353,** 353-355.

Vogel, A., Rodriguez, C., Warnken, W., and Izpisua Belmonte, J. C. (1995). Dorsal cell fate specified by chick Lmx1 during vertebrate limb development. *Nature* **378,** 716-20.

Vogel, A., and Tickle, C. (1993). FGF-4 maintains polarizing activity of posterior limb buds cells in vivo and in vitro. *Development* **119,** 199-206.

Vogels, R., Charite, J., de Graaff, W., and Deschamps, J. (1993). Proximal *cis*-acting elements cooperate to set *Hox-b7* (*Hox-2.3*) expression boundaries in transgenic mice. *Development* **118,** 71-82.

Vogels, R., De Graaff, W., and Deschamps, J. (1990). Expression of the murine homeobox-containing gene *Hox-2.3* suggests multiple time-dependent and tissue-specific roles during development. *Development* **110,** 1159-1168.

Wallace, H., and Watson, A. (1979). Duplicated axolotl regenerates. *J. Embryol. exp. Morph.* **49,** 243-258.

Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W. J., Stott, D., and Allemann, R. K. (1991). Multiple spatially specific enhancers are required to reconstruct the pattern of *Hox-2.6* gene expression. *Genes Dev.* **5,** 2048-2059.

Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcriptional complex stability and chromatin dynamics *in vivo*. *Nature* **377,** 209-213.

Wilkinson, D. G. (1993). Molecular mechanisms of segmental patterning in the vertebrate hindbrain and neural crest. *Bioessays.* **15,** 499-505.

Wilkinson, D. G., Bhatt, S., Chavrier, P., Bravo, R., and Charnay, P. (1989a). Segment-specific expression of a zinc finger gene in the developing nervous system of the mouse. *Nature* **337,** 461-464.

Wilkinson, D. G., Bhatt, S., Cook, M., Boncinelli, E., and Krumlauf, R. (1989b). Segmental expression of Hox-2 homeobox-containing genes in the developing mouse hindbrain. *Nature* **341,** 405-409.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *J. Theoret. Biol.* **25,** 1-47.

Wu, C.-t., Jones, R. S., Lasko, P. F., and Gelbart, W. (1989). Homeosis and the interaction of zeste and white in *Drosophila*. *Mol. Gen. Genet.* **218,** 559-564.

Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. J., and Korsmeyer, S. J. (1995). Altered *Hox* expression and segmental identity in *Mll*-mutant mice. *Nature* **378,** 505-508.

Zhang, C. C., and Bienz, M. (1992). Segmental determination in *Drosophila* confered by *hunchback* (*hb*), a repressor of the homeotic gene *Ultrabithorax* (*Ubx*). *Proc. Natl. Sci. Acad. USA* **89,** 7511-7515.

Zink, B., Engstrom, Y., Gehring, W., and Paro, R. (1991). Direct interaction of the *Polycomb* protein with *Antennapedia* regulatory sequences in the polytene chromosomes of *Drosophila melanogaster*. *EMBO J.* **10,** 153-162.

Zink, B., and Paro, R. (1989). *In vivo* binding pattern of a trans-regulator of homeotic genes in *Drosophila melanogaster*. *Nature* **337,** 468-471.

Zwilling, E. (1961). Limb morphogenesis. *Advaces in Morphology* **1,** 301-330.