GENETIC AND PHENOTYPIC ANALYSIS OF DANFORTH'S SHORT-TAIL MUTATION IN THE MOUSE

JANE ALFRED

PhD

UNIVERSITY OF EDINBURGH

1996

Declaration

I declare

a) that this thesis is composed by myself and

b) that the work is my own, except where otherwise stated

March 1996.

Abstract

The semi-dominant mutation Danforth's short-tail (*Sd*), arose spontaneously in the 1940's and affects the posterior development of the mouse during embryogenesis. In homozygotes (*SdSd*), the mutation is characterised by severe malformation of the posterior vertebral column and the urogenital organs. The notochord in these mice disintegrates from day 10 of development, resulting in the aberrant expression of genes which determine dorso-ventral patterning in the sclerotome and neural tube. In addition to neural tube defects, homozygotes characteristically have bilateral renal agenesis resulting in death soon after birth. The development of other posterior organs such as the external genitalia, bladder and hindgut are also affected by the mutation. Heterozygotes have a highly variable phenotype, where tail length can vary independently of the severity of the kidney phenotype, and where unilateral renal agenesis, or the reduction in size of one or both kidneys, is commonly seen.

The principle aim of this project was to finely map the proximal region of mouse chromosome 2, in order to identify markers genetically linked to the *Sd* gene. To this end, an interspecific backcross in which *Sd* was segregating was established, and the backcross progeny haplotyped using a panel of simple sequence repeat markers. Potential candidate genes, such as *Pax8* and vimentin, were excluded by backcross mapping and markers flanking *Sd* were identified and utilised in the screening of YAC libraries. YACs isolated with these markers were used to generate a YAC contig in the region of *Sd*.

Since tail length in the backcross varies considerably, the backcross was used to identify regions of the genome affecting the expressivity of *Sd*. In a collaboration with B. Taylor at the Jackson Laboratory, and with P. Keightly at the University of Edinburgh, several regions of the genome have been

found which show association with more or less severely affected tails. QTL analysis has been applied to these data to confirm the presence of modifiers acting on the *Sd* mutation at these loci.

Using an *in vitro* organ culture assay, the development of embryonic kidney rudiments from *Sd* mice has been studied. Utilising fluorescent-tagged antibodies against laminin and calbindin, laser-scanning confocal microscopy was used to make a detailed analysis of the development of *Sd* kidneys *in vitro*. From this work we conclude that mutant ureteric buds are able to induce the metanephric mesenchyme to condense and undergo epithelialisation but are unable to branch, possibly because they lack the receptor to transduce the mesenchymal-derived signal, the signal itself is aberrant or a component of the downstream transduction pathway is affected by the mutation.

Acknowledgements

First thanks must go to my supervisors, Ian Jackson and Cathy Abbott for their constant guidance, support and enthusiasm. They must have known that taking on a non-geneticist would have meant being subjected to endless questioning, and between them they have saved me hours of searching through books and papers by their ever-willingness to explain things to me just one more time. They have patiently initiated me into the world of genetics and for that I owe them my heartfelt thanks. Thanks also to the lab: to Julia, Doreen, Peter, Ruth and Marina, and a special thanks to Siobhan for proof reading my thesis when I'm sure she had much better things to do, and when Ian and Cathy had already read it a thousand times. Thanks must go to Stefan Larsson for taking the time to teach me the intricacies of the kidney culture assay, and for the advice, good conversation and bad jokes during our collaboration Thanks also to Prof. Kaufmann at the University of Edinburgh, for the discussions on the development and anatomy of the Danforth short-tail mouse, and to Vince Ranaldi for his work at the Biomedical Research Facility.

There are many people to thank at the Human Genetics Unit who made the writing of this thesis possible when it felt anything but. Thanks to Jem, Becky, Donald, Georgia and Viv for their support, company and advice. Thanks to Sandy who performed a miracle with my figures given the least possible time to do so. Thanks also to Andreas Schedl and Beate Nurnberger for their comments on certain sections of this thesis, and to Kelly Rance at the University of Edinburgh, and to Andrew Carrothers at the Human Genetics Unit, for their help and patience with the statistics.

Final thanks go to my parents and sisters, for the years of love, encouragement and support, and for always being there; and to my friends from home (Jess, Claire, Steph, Rich and John) who, over the past months, have shown me what friendship is really about.

ABBREVIATIONS

<	less than
>	greater than
µCi	microcuries
µM	micromolar
µl	microlitres
A	adenine
ABS	absorbance
APC	adenomatous polyposis coli gene
ATP	adenosine triphosphate
BAC	bacterial artificial chromosome
bp	base pair
BSA	bovine serum albumin
C	cytosine (in chapter 2)
C	*Mus musculus castaneus* allele (in chapters 3 to 5)
cAMP	cyclic adenosine monophosphate
cDNA	complementary deoxyribonucleic acid
cm	centimetre
cM	centiMorgan
CO_2	carbon dioxide
CTAB	hexadecyltrimethylammonium bromide
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dH_20	distilled water
DNA	deoxyribonucleic acid
DTAB	dodecyltrimethylammonium bromide
DTT	dithiothreitol
dTTP	deoxythymine triphosphate

E.Coli	Escherischia Coli
EDTA	ethylenediamine tetra-acetic acid disodium salt
EST	expressed sequence tag
EtBr	ethidium bromide
FITC	fluorescein isothiocyanate
g	gram
G	guanosine
HGF	hepatocyte growth factor
ICRF	Imperial Cancer Research Fund
Il1rn	interleukin 1 receptor antagonist gene
IMEM	improved modified eagles medium
kb	kilobase
l	litre
L	litter
LHS	left hand side
LMP	low melting point
LOD	logarithm of odds
m	milli
M	molar (in chapter 2)
M	*Mus musculus* CBA allele (in chapters 3 to 5)
M.m.	*Mus musculus*
Mb	megabase
$MgCl_2$	magnesium chloride
mi	murine micropthalmia gene
Min-1	multiple intestinal neoplasia gene
MITF	human micropthalmia gene
MIT	Massachusettes Institute of Technology
Mom-1	Modifier of *Min-1* gene
MRC	Medical Research Council
$NaBH_4$	sodium borohydride
NaCl	sodium chloride

NaOH	sodium hydroxide
OH	hydroxy
PCR	polymerase chain reaction
PDBu	beta-phorbol-12, 13-dibutyrate
PKC	protein kinase C
PMSF	phenylmethylsulphonyl fluoride
Pr	pair
QTL	quantitative trait locus/loci
RFLP	restriction fragment length polymorphism
RHS	right hand side
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
RT-PCR	reverse transcriptase polymerase chain reaction
Sd	Danforth's short tail
SDS	sodium dodecyl sulphate
SEM	standard error of the mean
SSR	simple sequence repeat
T	thymine
TE	10mM Tris.HCl, 1mM EDTA
TEMED	N,N,N,N'-tetramethylethylenediamine
T_m	melting temperature
V	volts
vim	vimentin gene
WT	wild type
WT1	Wilm's Tumour 1 gene
w/v	weight / volume
YAC	yeast artificial chromosome

CONTENTS

List of figures

List of tables

Chapter 1: Introduction

1.0 Introduction

This chapter serves as an introduction to the use of the mouse in genetics and to the Danforth short-tail mutation. At the beginning of each results chapter, detailed introductions are given to the areas of research relevant to the techniques used and the data presented.

1.1 The role of the mouse in genetics

The mouse is a vital model organism for the Human Genome Project and a substantial emphasis has been placed on the genetic and physical mapping of the mouse genome world-wide (Brown, 1992). Historically, the mouse has been the mammal of choice for genetic analysis for a number of reasons: because of its short gestation period and large litter sizes, the availability of inbred and wild-derived strains, and because of the ability to perform controlled matings in the mouse (Avner *et al.*, 1988; Copeland *et al.*, 1993). In addition to these inherent advantages, the comparison of mouse and human genetic maps reveals large number of chromosomal regions where both gene order and genetic content are conserved between the two genomes (Copeland *et al.*, 1993). These conserved linkage groups enable the identification of mouse mutations which may be homologous to human genetic disorders. In this way, mouse mutations provide powerful models for many human genetic diseases, including reproductive disorders (Koopman *et al.*, 1991; Pivnik *et al.*, 1991), cancers (Moser *et al.*, 1990), and polygenic disorders such as diabetes (Todd *et al.*, 1991) and epilepsy (Rise *et al.*, 1991).

In recent years, the advances made in the field of transgenic and embryonic stem cell technology have made it possible to generate new loss-of-function and gain-of-function mouse mutations, through recombinant DNA techniques such as gene targeting, as well as to express ectopically biologically diverse genes in a tissue specific manner (Capecchi 1989; Joyner, 1991). This

makes the mouse a highly manipulable organism to use in the identification of new genes and in the analysis of gene function. Information generated in this way can also be used in the analysis of human homologues.

Mouse mutations, whether spontaneously arisen, induced in radiation or chemical mutagenesis programmes, or created by recombinant DNA technology, provide a vital means of access to the mouse genome and offer insights into the biological function of genes in the absence of their identity. The advances made in mouse genetic mapping, such as the provision of high resolution interspecific backcrosses (The European Backcross Collaborative Group, 1994) and polymorphic DNA markers (Dietrich *et al.*, 1994) to the mapping community, have greatly facilitated the cloning and identification of genes responsible for mutant phenotypes. Recently it has become possible to introduce yeast artificial chromosomes into the mouse germ line (Schedl *et al.*, 1993), generating substantial potential for genome analysis and for the development of new mouse models of human disease. These technological advances, along with the inherent advantageous properties of the mouse, make this organism an ideal experimental animal to use in the analysis of both the mouse and human genomes.

The Danforth short-tail mouse mutation was investigated in this project in order to identify the gene responsible for the mutation. Identification of this gene may provide new information as to the developmental processes required for the correct morphogenesis of the neural tube and kidney.

1.2 Danforth's short tail (*Sd*) mutation

Danforth's short-tail mutation was first identified by Professor C. H. Danforth, among the descendants of his posterior duplication stock, a strain of mice in which 20% of the offspring exhibit duplications of the pelvic region, urogenital system and hind limbs (Danforth, 1930; Dunn *et al.*, 1940). The mutant mice

which arose from this strain did not exhibit the posterior duplication phenotype but were found to have a novel mutation which segregated as a semi-dominant, and which affected the development of the spine, tail and urogenital system. This mutation was given the symbol *Sd* (for short-Danforth) (Dunn *et al.*, 1940).

1.2.1 Phenotype of Danforth's short-tail mouse

Heterozygous phenotype

From the earliest analysis, the phenotype of *Sd* heterozygotes was found to be highly variable and dependent on genetic background (Dunn *et al.*, 1940; Glueckson-Schoenheimer, 1943; Dunn and Glueckson-Schoenheimer, 1945; Gruneberg, 1953).

In heterozygous neonates, skeletal defects are found concentrated in the posterior region of the spine, with the caudal and sacral vertebrae frequently fused together or malformed. In severely affected mice, the lower lumbar vertebrae can also be affected (Glueckson-Schoenheimer, 1943; Gruneberg, 1953). One study found vertebral components, such as vertebral bodies and intervertebral discs, to be absent in 14% of the heterozygous mice examined, with the affected vertebrae concentrated in the sacral part of the spinal column (Dietrich *et al.*, 1993a). In *Sd*+ mice, the bodies of all vertebrae are reported to be reduced in the dorso-ventral direction, with this reduction most marked in the cervical and first few thoracic vertebrae (Gruneberg, 1953), indicating that although the posterior half of the vertebral column is the region most severely affected by this mutation, the anterior vertebral column does not remain unaltered in the presence of *Sd*. An early finding of interest was that *Sd*+ heterozygotes lack the odontoid process (the dens epistrophei), a conserved structure which enables mammals to articulate the head from side to side. Its loss was attributed to the degeneration of the notochord during embryogenesis (Gruneberg, 1953).

When *Sd* heterozygotes (*Sd*+) were examined during embryogenesis, it was found that the development of the tail progressed normally up to day 10 of gestation, after which progressive degeneration of the tail somites occurred in a proximal direction (Glueckson-Schoenheimer, 1945). This degeneration was characterised by the appearance of blebs and haematomas at the tip of the degenerating tail bud, followed by a gradual decrease in the diameter and length of the tail, often resulting in a contorted, boneless filament at the distal end of the tail which persisted to adulthood (Glueckson-Schoenheimer, 1945).

Notochord development in heterozygotes is normal up to day 10 of gestation. In serial sections through day 11 *Sd*+ embryos, the notochord becomes discontinuous along the antero-posterior axis, and by mid-gestation has collapsed and disappeared throughout the length of the animal (Glueckson-Schoenheimer, 1945; Gruneberg, 1953; Center *et al.*, 1982; Boloventa and Dodd, 1991). This loss of the notochord during embryogenesis was proposed to be the primary defect in *Sd* mice (Gruneberg, 1953; Theiler, 1959). More recently it has been found that the floor plate is also absent during the development of the posterior vertebral column in *Sd*+ mice, and it has been suggested that this loss of floor plate is intimately associated with the loss of the notochord during *Sd* embryogenesis (Placzek *et al.* 1991; Boloventa and Dodd, 1991).

In *Sd*+ neonates, abnormalities in the urogenital system can range from a complete absence of both kidneys and both ureters, through all degrees of malformation, to an apparently normal phenotype (Glueckson-Schoenheimer, 1943). Reduction in the size of one or both kidneys has been found to be the most common effect of the mutation, bilateral renal agenesis being less common. The urogenital malformations in heterozygotes affect postnatal viability in those mice more severely affected by the mutation, phenotypic severity being dependent on genetic background (Dunn *et al.*,

1940; Dunn and Glueckson-Schoenheimer, 1945). Where one kidney is absent or malformed, the kidney originating from the left side of the body is found to be the most frequently affected (Glueckson-Schoenheimer, 1943; Mesrobian and Sulik, 1992).

During embryogenesis, the development of the urogenital system in heterozygous mice is normal up day 10.5 when the ureters bud off from the mesonephric ducts (Glueckson-Schoenheimer, 1945). Abnormalities in day 10 embryos are first seen when the ureteric bud appears to be morphologically abnormal compared to wild type littermates. By day 11, the ureteric bud is either inhibited or retarded in development and in severely affected mice, the bud can fail to branch altogether (Glueckson-Waelsch and Rota, 1963).

Homozygous phenotype

Homozygote neonates have a striking spectrum of developmental abnormalities including a complete absence of tail, bilateral renal agenesis, absence of external urinary and anal openings, and spina bifida (Dunn *et al.*, 1940; Glueckson-Schoenheimer, 1943; Glueckson-Schoenheimer, 1945). These mice survive to term and compete to suckle, but die 18-20 hours after birth due to autotoxicity (Dunn *et al.*, 1940).

In the absence of a vertebral column in the posterior region of the embryo, the spinal cord develops to a normal length but often ends in a cyst in the sacral region of the spine, beneath the dorsal epidermis. Due to the absence of lower lumbar, caudal, and sacral vertebrae, homozygous neonates are sometimes paralysed, with luxated hind limbs (Dunn *et al.*, 1940; Glueckson-Schoenheimer, 1943). The cervical vertebrae in these mice show antero-posterior reductions, including the loss of the dens axis. Sagitally split vertebrae are found in the lower thoracic and upper lumbar skeleton, with the vertebral body absent from the vertebral anlagen. The terminal vertebrae are

occasionally represented by amorphous lateral cartilage condensations. When split vertebrae extend into the thoracic vertebral column, the ribs of affected mice are disconnected from the vertebrae or completely absent. These observations indicate that it is the ventral aspect of the vertebral skeleton which is most severely affected by this mutation (Dunn *et al.*, 1940; Glueckson-Schoenheimer, 1943; Gruneberg, 1958; Dietrich *et al.*, 1993a).

The rectum in homozygous mice fails to develop since the intestine rarely continues beyond the colon. Failure in the development of the anus, the external genital papillae and urethral openings are also features of homozygote mice. On the rare occasion when kidneys are present, they are always abnormal in size and structure, and often displaced. Development of the bladder is frequently affected by this mutation, and is rarely of normal size in homozygous mice. Occasionally, malformation of the genital duct has been observed in female homozygotes (Dunn *et al.*, 1940; Glueckson-Schoenheimer, 1943).

Homozygous embryos are phenotypically distinguishable from wild type and heterozygous littermates from day 9 of gestation, due to the constriction of the tail bud at the base of the tail. The constricted tail is shorter and thinner than the normal tail and is characterised by the presence of haematomas at the distal tip of the tail bud. In sections of the tail bud at nine days, somitic tail structures are abnormal and the notochord substantially fragmented (Glueckson-Schoenheimer, 1945; Gruneberg, 1958; Theiler, 1959). By mid-gestation the notochord is entirely absent from the antero-posterior axis of homozygous embryos, as is the floor plate (Placzek *et al.* 1991; Boloventa and Dodd, 1991).

By day 10 of gestation, the mesonephric ducts are often abnormal at their posterior tips, and fail to reach beyond the point where the uereteric bud branches off towards the metanephric blastema (Glueckson-Schoenheimer,

1945; Glueckson-Waelsch and Rota, 1963). Failure of the posterior part of the mesonephric duct to develop properly was proposed to be responsible for the abnormalities in bladder development (Glueckson-Schoenheimer, 1945). However, the normal appearance of the gonads and their ducts indicates that the mesonephroi, and the anterior part of the mesonephric ducts, are morphologically normal. Nephrogenic tissue found at the tip of homozygote ureteric buds has been described as failing to undergo differentiation. Failure of the urogenital sinus and rectum to develop separately after day 10 of gestation, can result in the persistence of a cloaca (Glueckson-Schoenheimer, 1945; Glueckson-Waelsch and Rota, 1963).

1.2.2 Vertebral development in *Sd* mice

Components of the vertebral column are derived from the paraxial mesoderm, a tissue laid down during gastrulation which flanks the notochord and neuroectoderm. The axial skeleton of vertebrates is composed of a metameric series of vertebral bodies and intervertebral discs. Both of these structures are derived from the somites. The sclerotome, the part of the somite which gives rise to skeletal structures, is generated as a consequence of the differentiation of the somites into the mesenchymal sclerotome and the epithelial dermomyotome. Following de-epitheliasation of the somites, medial sclerotome cells migrate and proliferate towards the notochord to form the anlagen of vertebral bodies and intervertebral discs, whereas the lateral sclerotome forms the pedicles and laminae of the neural arches and ribs (Christ and Wilting, 1992). The somite also produces a second tissue, the dermomyotome, which subdivides into the dermotome and myotome to form dermis and skeletal muscle. With respect to their position, somite cells have to acquire a dorsolateral identity to become dermomyotome, whereas a ventromedial identity restricts them to a sclerotome lineage. Thus the dorsoventral differentiation of the somite is the crucial step in the formation of the axial skeleton as a whole.

A substantial body of experimental evidence has accumulated which indicates that the dorso-ventral specification of the neural tube and somites is strongly influenced by the notochord (Yamada *et al.*, 1991). It has been found in notochord ectomised amphibian neuralae and chick embryos, that in the absence of the notochord, the myotome expands and the sclerotome either degenerates or forms an unsegmented cartilage tube around the spinal cord, indicating that the notochord is required for the development of the sclerotome (Youn and Malacinski, 1980; Teillet and Le Douarin; 1983). The notochord is also responsible for inducing the development of the floor plate, a highly specialised region of the neural tube, since by grafting a piece of notochord to the dorsal midline of the developing chick neural tube, the ectopic induction of a floor plate can be induced (Yamada *et al.*, 1991). Evidence indicates that the floor plate plays a vital role in establishing cell identity along the dorso-ventral axis of the neural tube, as well as guiding commisural axon growth to, and across, the ventral midline of the spinal cord (Jessel *et al.*, 1989). In the absence of a floor plate, commisural axons are found to have aberrant trajectories, and the number of motor neurones which form are decreased or completely absent (Yamada *et al.*, 1991).

The development of the notochord and perinotochordal sheath in homozygous and heterozygous *Sd* mice has been studied (Center *et al.*, 1982). The perinotochordal sheath is a basement membrane which encloses the notochord as a thin ring in the anterior half of the animal. Towards the posterior region, the notochord remains attached to the ventral surface of the neural tube and the perinotochordal sheath becomes continuous with the basement membrane of the neural tube (Center *et al.*, 1982). The perinotochordal sheath was found in this study to be normal in the anterior half of day 10 *Sd*+ embryos, but reduced in the mid-trunk to posterior half of the animal, where the notochord was disorganised and malformed. In homozygote mice, the anterior notochordal complex was abnormal in shape and situated closer to the ventral surface of the neural tube than normal. By

day 12.5 of development, the sheath was reduced to a poorly formed, anterior structure with no basement membrane in *Sd+* mice. In the mid-trunk to posterior region of these embryos, there was little to no indication of a defined notochord or sheath. In the 12.5d homozygous embryos, very little evidence of the notochordal complex was found at any level (Center *et al.*, 1982).

Notochordal cells in 11-11.5 day *SdSd* embryos have ultrastructural features which differ in several respects from those of wild type (WT) littermates. The cells of *SdSd* notochords lack surface protrusions, and are deficient in cytoplasmic filaments seen in normal cells (Wilson *et al.*, 1982). Also found was a complete lack of glycogen in these cells, a carbohydrate which is present in abundance in WT notochordal cells in the form of granules. This lack of glycogen in mutant notochordal cells was thought to reflect either an earlier than normal transition from a state of proliferation to a state of differentiation, or an inability to form sufficient amounts of this energy source during early differentiation (Wilson *et al.*, 1982).

A number of studies have been performed to investigate the effect of the *Sd* mutation on the dorsal-ventral patterning of the neural tube and somites during development. These studies have examined genes which have highly localised domains of expression during neural development, to investigate whether their expression patterns are altered in *Sd* mice (Dietrich *et al.*, 1993a; Koseki *et al.*, 1993a; Phelps and Dressler, 1993). In addition, work has also been carried out to analyse the affect that the mutation has on neuronal differentiation and axonal guidance in the developing neural tube, processes which normally require the inductive activity of the floor plate and notochord (Boloventa and Dodd, 1991).

Members of the paired box family of transcription factors, the *Pax* genes, are expressed in distinct spatiotemporal patterns during embryogenesis. Several

of these genes are expressed in overlapping dorsoventral patterns within the developing neural tube (Deutsch and Gruss, 1991). The expression domains of a number of these genes have been examined in *Sd* mice (Dietrich *et al.*, 1993a; Phelps and Dressler, 1993).

Pax1 is the earliest marker for the ventral region of the somite. At day 9.5 it is expressed in the entire sclerotome, an expression pattern which later becomes restricted to the ventral region of the sclerotome, the anlagen of the intervertebral discs (Deutsch *et al.*, 1988). The *Pax1* mutation undulated (*un*) affects mainly the development of vertebral bodies and intervertebral discs (Balling *et al.*, 1988), indicating that *Pax1* is required for the correct patterning and development of these ventromedial structures. When homozygous undulated mice are crossed with *Sd+* mice, the resulting double mutants have a more severe vertebral phenotype than is seen in either mutation alone, indicating that the mutations are not allelic, as predicted by their map positions, and that *Pax1* activity is affected by the degeneration of the notochord (Koseki *et al.*, 1993a).

In contrast to *Pax1*, the expression of *Pax3* is restricted to the dorsal aspect of both somite and neural tube (Goulding *et al.*, 1991, 1993). The mouse *Pax3* mutation splotch demonstrates the importance of this gene for the development of dorsal structures, as the dorsally emerging neural crest cells are defective in these mice, frequently causing failure in neural tube closure. Notochord excision in chick embryos results in the ventral extension of the *Pax3* expression domain in the neural tube (Goulding *et al.*, 1993), indicating that the notochord is responsible for maintaining the dorsally-defined expression of *Pax3*, either directly or through the release of ventralizing signals, such as sonic hedgehog (Bumcrot and McMahon, 1995). Thus *Pax3* serves as an appropriate dorsolateral marker for analysis of the interaction of the notochord with the somite and spinal cord (Dietrich *et al.*, 1993a).

Analysis of *Pax1* and *Pax3* expression showed that the expression of both genes is altered as a consequence of the degeneration of the notochord in heterozygous and homozygous *Sd* embryos (Dietrich *et al.*, 1993a). The use of the notochord-specific marker *T*, showed that the notochord degenerated in homozygotes from day 9.5 to day 11 of development. In the anterior region of homozygous embryos, despite the absence of the notochord, *Pax3* and *Pax1* expression were found to be unaltered, and induction of the floor plate was evident (Dietrich *et al.*, 1993a). Since transplantation experiments have shown that the floor plate is able to mimic the ventralising function of the notochord (Yamada *et al.*, 1991), it was proposed that in *Sd* homozygotes, the floor plate may similarly substitute for the notochord function in establishing a ventral identity in the neural tube. Alternatively, after sufficient development of the somite and neural tube, the degeneration of the notochord may no longer be able to affect these tissues (Dietrich *et al.*, 1993a). At the posterior end of homozygous embryos, *Pax1* expression failed to be detected, whilst expression of *Pax3* was found to be present throughout the neural tube and somites, instead of being restricted to the dorsal aspect of these structures as in WT embryos. Since the expression of *Pax3* is known to be restricted by the notochord, via the expression of ventralizing genes such as *Pax1*, to the dorsal half of the neural tube, this result indicated a number of conclusions (Dietrich *et al.*, 1993a): that the absence of notochord from the developing posterior half of the embryo results in the failure of floor plate induction, and that in the absence of floor plate and notochord, the developing somites and neural tube cannot acquire a ventral identity, as indicated by the loss of *Pax1* expression (Dietrich *et al.*, 1993a; Bumcrot and McMahon, 1995). Consequently, the neural tube and somites develop with a dorsalized phenotype, brought about by the unrestricted expression of dorsalizing genes, such as *Pax3*. This is reflected in the phenotype where ventral structures such as the vertebral bodies and intervertebral discs are among the skeletal elements most severely affected by the mutation. The phenotype is less severe in the anterior portion of the

vertebral column because the notochord functions sufficiently to induce floor plate development before degenerating. The floor plate is then able to induce the expression of ventralizing genes in the anterior neural tube and somites (Dietrich *et al.*, 1993a; Boloventa and Dodd, 1991).

The expression pattern of *Pax2* has also been examined in heterozygote and homozygote *Sd* mice (Phelps and Dressler, 1993). In the spinal cord, *Pax2* is expressed at the onset of neuronal differentiation as cells migrate from the ventricular zone (Nornes *et al.*, 1990). *Pax2* expressing cells are initially detected in the ventral two-thirds of the alar plate and the dorsal half of the basal plates (Nornes *et al.*, 1990). Although normal patterns of *Pax2* expression were observed in the thoracic and lumbar regions of the spinal cord of *SdSd* mice, the sacral and caudal regions showed ectopic *Pax2* expression in the ventral half of the basal plate and through the ventral midline. In the absence of floor plate or notochord, this ectopic expression of *Pax2* was proposed to inhibit the development of motor neurones in *SdSd* mice. The development of motor neurones is greatly reduced in heterozygote and homozygote mice (Boloventa and Dodd 1991), in the same way that they are absent from chicks in which the notochord has been surgically removed (Yamada *et al.*, 1991). As in *Sd* mice, this not only causes reduction and ectopic development of motor neurones, but also the loss of floor plate and changes in the expression pattern of ventrally expressed genes. Thus, alterations in the expression domain of *Pax2* may be involved in the reduction of motor neurone development in these mice (Phelps and Dressler, 1993).

The loss of the notochord and floor plate in *Sd+* and *SdSd* mice, not only affects motor neurone development, but also the growth and trajectory of commisural axons (Boloventa and Dodd, 1991). Evidence exists which suggests that the floor plate may contribute to axon guidance, and it has been proposed that it may be the source of diffusible chemoattractants which

orientate the growth of spinal commisural axons (Tessier-Lavigne *et al.*, 1988). In *SdSd* mice, commisural growth axons project normally in the ventral direction, but then fail to project ventromedially to the ventral midline of the neural tube in parts of the spinal cord where the floor plate is absent (Boloventa and Dodd, 1991).

Although these studies have provided valuable insights into the developmental pathways which have been affected by the mutation, it is still not known why the notochord in *Sd+* and *SdSd* mice degenerates during embryogenesis.

1.2.3. Kidney development in *Sd* mice

Embryological studies of mice heterozygous and homozygous for *Sd* have revealed abnormalities in the formation, growth and differentiation of the ureteric bud and metanephric mesenchyme (Glueckson-Schoenheimer, 1945; Glueckson-Waelsch and Rota, 1963).

The interdependence of the two main component tissues of nephrogenesis, the ureteric bud and the metanephric mesenchyme, was first demonstrated by Grobstein (1955, 1956). Tubule formation in the mesenchyme was shown to be dependent upon an inductive influence provided by the ureteric bud, while branching of the ureteric bud was found to be dependent on the metanephric mesenchyme. Dorsal spinal cord was also found to be a more potent *in vitro* inducer of mesenchymal differentiation than ureteric tissue. A more detailed account of the development of the kidney, and the genes involved in these inductive interactions, is given in the introduction to chapter 6.

An early study was undertaken to investigate the role of the *Sd* mutation in the inductive interactions between the metanephric mesenchyme and ureter during kidney development, using an *in vitro* organ culture assay in which

embryonic kidneys were cultured in clots prepared from chicken plasma (Glueckson-Waelsch and Rota, 1963). Kidney rudiments from 11-12 day heterozygote and homozygote *Sd* mice were either cultured intact, or the ureteric bud and metanephric mesenchyme were enzymatically separated and cultured in combination with wild type/ mutant ureteric bud, metanephric mesenchyme or spinal cord. Homozygous rudiments cultured intact were found to be distorted and retarded in development. The ureteric branches failed to branch properly, and although some tubule formation was seen where ureteric branching had occurred, it was considerably retarded in development. (Glueckson-Waelsch and Rota, 1963).

When the metanephric mesenchyme from homozygote rudiments was separated from the ureteric bud and cultured with ureters from wild type (WT) embryos, only 29% of cultures showed tubule differentiation. However, since only 50% of normal mesenchymes underwent differentiation in the presence of WT ureters, rudiment development was not consistent in this assay used. Mutant mesenchyme produced less tubule differentiation in combination with WT ureter than normal mesenchyme. When mutant mesenchyme was cultured with WT dorsal spinal cord, most differentiated well. However, when normal mesenchyme was combined with ureters from 11-12 day homozygotes, although 50% of cultures showed mesenchymal differentiation, the response was limited, since tubulogenesis occurred only where ureteric branch tips had developed (Glueckson-Waelsch and Rota, 1963).

These data suggested that the *Sd* mutation does not completely suppress either the inductive effect, nor the reacting ability of the ureter and metanephric mesenchyme. Although differentiation was abnormal it was not possible to conclude whether both tissues were retarded in development, or whether the metanephric mesenchyme was retarded as a result of the decreased responsiveness of the ureteric bud to branching stimuli

(Glueckson-Waelsch and Rota, 1963). The findings produced by this study are discussed further in chapter 6, section 6.2.2.

The expression of the transcription factor *Pax2* in *Sd* rudiments has also been investigated (Phelps and Dressler, 1993). As detailed in section 6.2.2, *Pax2* is expressed during kidney organogenesis in the pro- and mesonephric tubuli and the ureteric bud (Dressler *et al.*, 1990). Following nephrogenic induction, *Pax2* is present in the mesenchymal condensations and its early epithelial derivatives, the comma and S-shaped bodies. As epithelial differentiation continues and kidney tubules are formed, *Pax2* expression declines, indicating a role for *Pax2* in the conversion of mesenchyme to epithelium during kidney organogenesis (Dressler and Douglass, 1992). In day 12 homozygous embryos, a metanephros could not be identified and by day 13, the mesenchymal cells surrounding the posterior mesonephric duct did not express *Pax2*. On the occasions when the ureteric bud branched from the mesonephric duct, *Pax2* expression was seen, but the cells surrounding the bud failed to express *Pax2* (Phelps and Dressler, 1993). This work suggests that in the absence of ureteric branching, the signalling cascade responsible for the mesenchymal to epithelial differentiation of the metanephric mesenchyme is not activated in homozygous *Sd* mice. Normal *Pax2* expression in the mesonephroi and Wolffian duct of homozygotes indicates that these structures develop normally, supported by the fact that gonadal development is unaffected by the *Sd* mutation (Gluecksohn-Schoenheimer, 1943).

1.3 Analysis of Danforth's short-tail mutation

Danforth's short-tail mutation was originally assigned to the fifth linkage group in the mouse, where it was linked to pallid and agouti, two phenotypically distinguishable mutations on the same chromosome (Borger, 1950). Later it was mapped to the proximal end of mouse chromosome 2

(Beechey and Searle, 1980), where several cloned genes of interest have been mapped (Siracusa and Abbott, 1994). At the onset of this project, the paired box gene *Pax8* was considered to be a strong candidate for the mutation, owing to its map position in the proximal region of mouse chromosome 2 and to its expression in both the developing spinal cord and kidney (Plachov *et al.*, 1990). One of the primary aims of this project was to generate a high resolution map in the region of *Sd* to assess *Pax8*'s candidacy for the mutation. In order to do this, an inter-subspecific backcross in which *Sd* was segregating was generated. A review of the approaches used to map genes in the mouse, and the results of the mapping data, are given in chapter 3.

Since *Pax8* was excluded as a candidate gene for *Sd* on the basis of the map that was generated in this study, a positional cloning exercise was initiated with the aim of isolating the gene responsible for the mutation (chapter 4). A review of the strategies which can be used to positionally clone genes is given in the introduction to this chapter.

As mentioned in section 1.2.1, *Sd* has a highly variable phenotype. Whilst generating the backcross it was noted that tail length varied between F1 and backcross generations. As a result of this finding, a genome wide analysis of backcross progeny was performed, to identify quantitative trait loci responsible for modifying tail-length in the presence of *Sd*. A review of quantitative trait loci mapping is given in the introduction to chapter 5, where the results of this work are detailed.

An *in vitro* organ culture assay was used to investigate the development and branching morphogenesis of embryonic kidneys from heterozygous and homozygous *Sd* mice (chapter 6). A review of kidney organogenesis, and of the genes involved in renal induction and morphogenesis, is given in the introduction to this chapter.

Danforth's short-tail mutation is a semi-dominant developmental mutation which affects both the development of the neural tube and tail, and the differentiation of the urogenital system (Gluecksohn-Schoenheimer, 1943; Gluecksohn-Schoenheimer, 1944). From its wide-ranging phenotype it is clear that the correct expression and biological function of this gene are vital for the normal development of two diverse and highly patterned tissues during embryogenesis. It is possible that the identification of the *Sd* gene may lead to a better understanding of renal disorders in humans, such as the dominantly inherited form of renal adysplasia which results in unilateral and bilateral renal agenesis (McPherson *et al.*, 1987). In this way the identification of the *Sd* gene may offer considerable insights into the genetic basis of renal and neural tube disorders in both mice and humans.

Chapter 2: Materials and methods

2.1 Maintenance of mice stocks

Mice were kept at the Biomedical Research Facility at the Western General Hospital where their care was supervised by V. Ranaldi. The animal work carried out in this project was authorised by the Home Office, licence number 60/01412 (Experimental Modification of the Mouse Genome).

2.1.1 Origins and strains of mice

Danforth's short tail mice (*Sd*) were obtained from Harwell where they were maintained as a linkage stock on an outbred background. On arrival at the Biomedical Research Facility, the *Sd* mice were bred with a *M. musculus* inbred laboratory strain derived from CBA/Ca and were subsequently maintained on this background. For the backcross, *M. musculus castaneus* mice, of the strain CAST/Ei, were obtained from Harlan Olac.

2.1.2 Backcross protocol

One mating was set up to produce F1 offspring (*Sd* x *M.m.castaneus*). Heterozygous *Sd* male and female F1 offspring were mated back to CBA/Ca mice to produce the backcross progeny. Thirteen (F1 x CBA/Ca) mating pairs were set up in a ratio of one male to one female.

Backcross progeny were killed three days after weaning and all were dissected for the removal of their spleens and kidneys. Tail lengths were measured and recorded, along with kidney size and any visible abnormalities. Spleens and tail tips were snap frozen in liquid nitrogen and stored at -70°c until required for the preparation of DNA.

2.2 Kidney culture assays

2.2.1 Media and growth factors used in kidney culture assays

Ham's F10 solution

Ham's F10 solution was purchased in powder form from Gibco BRL and prepared according to the manufacturer's instructions.

Improved Modified Eagles Medium, IMEM

Customised IMEM was made available to us as a kind gift from P. Ekblom and was produced by Gibco, BRL. 1 litre of IMEM was prepared with 25mM $NaHCO_3$ (Analar, BDH) and 20mM Hepes (Flow Labs) and the pH adjusted to 7.6 using NaOH. The medium was then stored in 10ml aliquots at 4°C. Directly before use the medium was supplemented with 50µg/ml sterile, iron-poor apo-Transferrin (Sigma) and 300µg/ml sterile glutamine (Gibco, BRL).

IF base

Phosphate Buffered Saline (PBS, Oxoid) supplemented with 1% Bovine Serum Albumin (BSA, Boehringer Mannheim) and 0.01% sodium azide, pH7.3

Mounting solution

Equal volumes of non-fluorescent glycerine (Merck) and PBS were mixed with 2.5% 1, 4, diazabicyclo-(2,2,2)octane (Sigma).

Hepatocyte Growth Factor (HGF)

5µg human recombinant HGF (R&D Systems) was diluted in IMEM and 200ng/ml aliquots were stored at -20°C. Directly before use, each aliquot was diluted with 2.5ml IMEM and used in culture at a concentration of 80ng/ml.

Phorbol dibutyrate (PDBu)

PDBu (Sigma) was diluted directly before use in IMEM, from a 1µM stock stored at -20°C.

2.2.2 Antibodies used for immunohistochemistry

Primary antibodies

Mouse monoclonal anti-calbindin D antibody (Sigma) was diluted in IF base and used at a final dilution of 1:200.

Rabbit polyclonal anti-mouse laminin antibody (Euro-Diagnostica) was diluted in IF base and used at a final dilution of 1:400

Secondary antibodies

Anti-mouse, host horse, Texas Red-conjugated antibody (Vector) was diluted in IF base and used at a final dilution of 1:100.

Anti-rabbit, host goat, fluorescein isothiocyanate (FITC)-conjugated antibody (Vector) was diluted in IF base and used at a final dilution of 1:100.

2.2.3 Isolation of embryos

Male and female heterozygote Sd mice were mated and the day of the vaginal plug was designated as day 0.5. On day 12.5 of gestation, pregnant females were killed by cervical dislocation and the embryos dissected intact into ice cold PBS (Oxoid). Each embryo was phenotyped according to its tail length and appearance as being homozygous for *Sd* (*SdSd*), heterozygous (*Sd+*) or wild type (+/+).

2.2.4 Dissection of embryonic kidneys

Embryonic kidneys were dissected out into ice cold Ham's F10 solution, in sterile glass Petri dishes, using 21 gauge hypodermic needles (Becton Dickinson). In brief, kidneys were located by making transverse cuts just anterior and posterior to the hind limb buds. Holding the hind limb region dorsal side up, a saggital cut was made through the neural tube to separate the embryo into left and right halves, and the limb bud was removed from each half. The kidney rudiment was located by following the Wolffian duct until the metanephric blastema could be seen. On dissection, as much undifferentiated tissue from around the metanephric mesenchyme was removed as possible and the kidneys were transferred, individually by siliconised glass pipette, to a 96 well microtitre plate (Costar) where they were stored in IMEM at 4°C until needed. From the time of the sagittal section through the neural tube, it was noted from which side of the embryo (left or right) each kidney was dissected from.

2.2.5 Embryonic kidney culture assay

The organ culturing method used was based upon that described by Saxen, 1987.

Kidney rudiments were cultured for 24-60 hours on 1 μm nuclepore filters (Costar) which had been placed on stainless steel Trowell grids. Trowell grids were placed in Petri dishes to which supplemented IMEM was added almost to the top of the grid (usually 2.5ml), in this way kidneys were cultured in a film of medium generated by surface tension. Kidneys rudiments were cultured in a humidified incubator in 7% CO_2, at 37°C, for varying lengths of time.

Where kidney rudiments were cultured in the presence of growth factors, the growth factors were diluted in IMEM and added to the culture medium directly before incubation at 37°C. For each kidney rudiment cultured, its left or right pair was cultured without growth factor as a control.

2.2.6 Immunohistochemical staining of cultured kidney rudiments

Cultured kidneys were fixed overnight at 4°C in methanol and washed the following day in IF base, at room temperature, for 30 minutes. They were then exposed to primary antibodies for 4 hours at 37°C, followed by a 3 hour wash at room temperature in IF base, followed by overnight incubation with the secondary antibodies at 4°C. Before mounting, the kidney cultures were given a final wash in IF base at room temperature for 3 hours.

Kidney cultures were mounted on modified glass slides. To avoid crushing the cultures with a cover slip, a small piece of glass was glued to either end of a glass slide (Chance Propper Ltd.) and the kidney cultures placed in the space between the two glass pieces. In this way, the cover slip rested on the glass pieces and not the cultures themselves. Cultures were mounted using a mountant containing an anti-fading reagent (see section 2.2.1), so that the fluorescent signal of the secondary antibodies would be protected during laser-scanning confocal microscopy. To prevent the mountant from drying

out, slide edges were sealed using a vulcanized rubber solution (Stahlgruber, Rema Tip Top solution).

2.2.7 Confocal imaging of cultured embryonic kidneys

The fluorescently stained kidney cultures were optically sectioned using an Ortholux II microscope (Leitz) equipped for confocal microscopy with a MRC 600, Biorad scanhead. Fluorescently labelled cells were excited with an argon laser and selected filters were used to allow for the simultaneous detection of FITC and Texas Red signals. For the HGF experiments, a plan 10 objective was used on the microscope and for the PDBu experiments, a 6.3 objective was used.

2.3 Yeast cell culture and DNA isolation

2.3.1 Media and additives

AHC broth and agar

AHC is a rich selective medium which lacks uracil and tryptophan. It was used for selective growth of YAC recombinants prior to the isolation of YAC DNA and the production of plugs.

1.7g yeast nitrogen base (without amino acids and ammonium sulphate) (Difco), 5g ammonium sulphate and 10g casein hydrolysate (low salt) were added to 1 litre distilled water and the pH adjusted to 5.8. 17-20g of Bacto agar (Difco) were added to each litre of broth to make AHC agar. After autoclaving and cooling, 50ml filter sterilised 40% glucose and 10ml of 2mg/ml adenine sulphate were added to the broth or agar.

YPD broth and agar

YPD medium was used for growing yeast cultures.

20g bactopeptone (Difco), 5g yeast extract (Difco) and 5g sodium chloride were added to 1L distilled water. 1.5% agar was added at this stage for YPD agar. After autoclaving and cooling, 50ml sterile 40% glucose was added to the broth or agar.

2.3.2 Yeast strains

Saccharomyces cerevisiae strains YP148 (Jones *et al*, 1989) and AB1380 (Link and Olson, 1991) were used as size markers for pulse field gel electrophoresis. Table 2.3 lists the sizes of the chromosomes of the two strains.

AB972		YP148	
Chr. Number	**Size (kb)**	**Chr. Number**	**Size (kb)**
IV	1640	XII	2500
VII, XV	1105	IV	1500
XVI	955	XV	1125
XIII	930	VII proximal	1035
II	830	XVI	1000
XIV	790	XIII	940
X	750	II	830
XI	690	XIV	790
V, VIII	585	X	750
IX	445	XI	680
III	350	V	600
VI	285	VIII	550
I	240	IX	440
		III	350
		VI	270
		I	210
		VII distal	90

Table 2.3: Chromosomal sizes of yeast size markers
Size of marker fragments is given in kilobases (kb).

2.3.3 Isolation of DNA from yeast

Two methods were used for the isolation of yeast DNA.

A rapid method to extract high quality genomic DNA (Gustinich *et al*, 1991) was used to prepare small quantities of yeast DNA. Briefly, a 5ml yeast culture (10^8 cells/ml, grown in AHC) was spun down and washed twice in TE (10mM Tris.HCl pH8.0; 1mM EDTA), and resuspended in 300µl of zymolyase buffer (1M sorbitol, 0.1M citrate pH7, 0.06M EDTA, 0.1mM β-mercaptoethanol (Sigma), 2mg/ml zymolyase [ICN Biomedicals Inc.]). Following incubation with the zymolyase buffer for 30 minutes at 35°C, the cells were lysed with 600µl of lysis buffer (8% dodecyltrimethylammonium bromide [DTAB, Sigma], 1.5M NaCl, 100mM Tris pH8.6, 50mM EDTA) at 68°C for five minutes. 900µl chloroform was then added to the spheroblasts and the mixture centrifuged at 14,000rpm for 10 minutes. 750µl of the upper aqueous layer was removed and 900µl water and 5µl 10mg/ml RNA-ase A (Boehringer Mannheim) added to it and incubated at 68°C for 10 minutes. 100µl of hexadecyltrimethylammonium bromide (CTAB, Sigma), from a 5% stock in 0.4M NaCl, was added to the mixture and the DNA precipitated by centrifugation at 10,000rpm for 5 minutes. The pellet was resuspended in 300µl of 1.2M NaCl and 750µl ethanol added to it. Following centrifugation at 10,000 rpm for 10 minutes the pellet was washed in 70% ethanol and the pellet resuspended in 300µl TE at 55°C for 10 minutes.

For larger scale preparation of genomic yeast DNA the Nucleon II Extraction Kit was used (Scotlab). In brief, a yeast culture was grown to 3-4 x 10^9 cells/ ml in AHC. The cells were lysed with SCEM buffer (1M sorbitol, 10mM EDTA, 0.1M sodium citrate, 18mM β-mercaptoethanol and 4 units zymolyase) and incubated at 37°C for 2 hours. Spheroblasts were centrifuged and treated with 45µl 50µg/ml RNA-ase A (37°C, 30 minutes) and 1.5ml 5M sodium perchlorate (15 minutes at room temperature). 5.5ml of ice cold chloroform and 800µl of Nucleon silica suspension were added to the tube and after centrifugation (3250 rpm for 3 minutes) the DNA-containing phase above the

silica suspension was transferred to a fresh tube. The DNA was precipitated with ethanol and washed in 70% ethanol and the pellet resuspended in 250-500µl TE.

2.3.4 Preparation of agarose plugs

This method allows for the extraction of intact yeast chromosomes, including yeast artificial chromosomes (YACs), from cells. It was adapted from Maule (1994).

A 100ml culture was grown from a single, PCR-screened colony (see section 2.10.5). The culture was grown in selecting AHC medium at 30°C for 48 hours or until the cells reached approximately 10^8 cells/ml. The culture was chilled on ice for 15 minutes prior to centrifugation at 4,000 rpm for 10 minutes. The supernatant was discarded and the cell pellet washed twice in 50ml chilled 50mM EDTA, pH7.5. After washing, the cell pellet was resuspended in 3ml 50mM EDTA pH7.5 and warmed to 37°C. Cell walls were disrupted by the addition of 1.2ml freshly prepared cell wall digestion solution (2ml SCE, pH8.0 [1M sorbitol; 0.1M sodium citrate; 60mM EDTA]; 100µl β-mercaptoethanol; 2mg/ml zymolyase). 6ml of 1% low melting point agarose (Gibco BRL) in 0.125M EDTA, pH7.5, at 50°C was added to this. After mixing thoroughly, 100µl aliquots of the solution were quickly dispensed into plug moulds and left to cool on ice. Once set, the plugs were ejected into 25ml ETM solution (0.45M EDTA; 10mM Tris.HCl, pH8.0; 7.5% β-mercaptoethanol) and incubated at 37°C in a waterbath for 24 hours. The ETM was replaced with 20ml 1% NDS solution (0.45M EDTA; 10mM Tris.HCl, pH9.0; 1% SDS) containing Proteinase K (Boehringer Mannheim) at 1mg/ml, and the plugs incubated at 55°C for 24 hours. This step was repeated with fresh 1% NDS solution. Finally, the plugs were stored in 0.5M EDTA at 4°C.

2.3.5 YAC libraries

To isolate YACs the Genethon Mouse YAC screening service was used. This service uses a PCR-based assay to screen three major libraries: the ICRF

YAC library (Larin *et al*, 1991), the St. Mary's Hospital Library (Chartier *et al*, 1992) and the Research Genetics Library (Kusumi *et al*, 1993).

1mM of each primer was lyophilised and sent to Genethon to screen the above three libraries with. All positive clones were returned in the form of agar stabs which were then plated out onto AHC plates and grown at 30°C for 48 hours. Swabs were taken of colonies from each plate and resuspended in water. After lysing at 100°C for 10 minutes, the lysates were PCR-screened (section 2.10.5) for the presence of the marker with which the YAC had been isolated. Single colonies were also picked and PCR-assayed (section 2.10.5) to check for the presence of markers.

2.4 DNA isolation from murine tissues

2.4.1 DNA extraction from spleens

After dissection, spleens were snap frozen in liquid nitrogen and stored at -70°C until required. DNA was extracted from spleens using a high salt method.

Spleens were homogenised in 2mls TNES (50mM Tris.HCl, pH7.5; 400mM NaCl; 100mM EDTA; 0.5% SDS) and incubated over night, at 55°C, with 100µg/ml Proteinase K. The resulting solution was split in two and an equal volume of 2.6M NaCl added to each tube, to adjust its molarity to 1.5M. The tube was briefly shaken and centrifuged at 12,000g for 5 minutes. The supernatant was decanted to a clean tube and the precipitate discarded. An equal volume of ethanol was added and the tube inverted to allow the DNA to be spooled out using a sealed glass pipette. Once dry, the DNA was dissolved in 500µl TE (10mM Tris.HCl, pH8.0; 1mM EDTA) or water.

If necessary, DNA was further cleaned by phenol/ chloroform extraction. An equal volume of phenol was added to the tube and mixed, followed by centrifugation at 3000rpm for 10 minutes. The upper aqueous layer was removed and the process repeated, first with chloroform, then with a 24:1

solution of chloroform: isoamyl alcohol. The DNA was precipitated by adding two and a half volumes of ice cold ethanol and mixing thoroughly. The DNA was spooled out and resuspended as before.

2.4.2 DNA extraction from tail tips

This method was adapted from Laird *et al*, 1991. The lysis buffer has been adjusted to allow restriction digestion or PCR-amplification of the DNA without prior organic solvent extraction.

Once the tail tip had been obtained, it was immediately placed into 0.5ml of lysis buffer (100mM Tris.HCl, pH8.5; 5mM EDTA; 0.2% SDS; 200mM NaCl; 100µg/ml Proteinase K) and incubated overnight at 55°C. The tubes were then briefly vortexed before being spun in an Eppendorf centrifuge at 14000rpm for 15 minutes. The supernatant was decanted to a fresh tube and 0.5ml of isopropanol added to it to precipitate the DNA. After inverting the tube, the DNA was spooled out using a disposable yellow tip and, after dipping in 80% isopropanol, was dissolved overnight in 200µl TE.

2.5 DNA electrophoresis

2.5.1 Electrophoresis solutions

20 X TBE

1M Tris.HCl, pH8.0; 20mM EDTA; 1M boric acid, pH8.3.

20 X TAE

0.8M Tris.HCl, pH 8.0; 20mM EDTA; 0.4M acetic acid.

10 X DNA Loading Buffer

20% ficoll, 100mM EDTA, orange G.

6% Denaturing Polyacrylamide

Pre-prepared acrylamide/ bis acrylamide (19:1) sequencing solution (Severn Biotech Ltd.) was used, containing 6% w/v acrylamide, 7M urea and 1XTBE.

60mls of this solution was used to pour each gel. 75µl of N,N,N,N'-tetramethylethylenediamine (TEMED, Sigma) and 0.5ml of freshly made 10% ammonium persulphate were added to the acrylamide directly before use.

8% Non-Denaturing Polyacrylamide

50mls of solution was made to pour each gel. 10mls of 40% 29:1 acrylamide: bis-acrylamide solution (EasiGel, Scotlab) was diluted in 30mls of water and 10mls 5 X TBE. 700µl 10% ammonium persulphate and 30µl of N,N,N,N'-tetramethylethylenediamine (TEMED, Sigma) were added to the acrylamide immediately before pouring the gel.

HydroLink (Mutation Detection Enhancement - MDE) Gel

To make 50mls of solution: 25mls MDE gel solution (AT Biochem) was added to 3 mls 10 X TBE and 22mls water. 300µl 10% ammonium persulphate and 30µl of N,N,N,N'-tetramethylethylenediamine (TEMED, Sigma) were added to the gel solution immediately before pouring the gel.

Size markers used in Agarose Gel Electrophoresis

The size markers used were either λ DNA digested with *Hind*III (Boehringer Mannheim) or ϕX174 digested with *Hae*III (Boehringer Mannheim), depending on the size of the DNA fragments being resolved. Marker sizes are shown in table 2.5.

2.5.2 Agarose gel electrophoresis

DNA molecules were separated according to their size on horizontal agarose (SeaKem LE, FMC Bioproducts) gels. The percentage of agarose used to make the gel depended on the size range of the DNA molecules being resolved. Restricted genomic DNA was commonly run on large, 300ml 0.8% agarose gels, whereas smaller fragments, such as PCR products, were run on 1-2% agarose gels. PCR fragments of 300bp or smaller were run out on 4% NuSieve (FMC, Bioproducts) gels. All agarose gels were made from and run in either 1 X TBE or 1 X TAE. To visualise the DNA, ethidium bromide

was added to agarose gels at a concentration of 1μg/ ml of buffer. 1/10 of the sample volume of 10X loading buffer was added to the DNA prior to loading the sample on the gel.

Mini gels were run in Northumbria Biologicals Ltd. gel boxes or Electro-4 Hybaid gel tanks. Gels made with and run in 1 X TBE were run at 100 volts, and gels made with and run in 1 x TAE were run at 50 volts. Time of running depended on the degree of separation needed.

DNA fragments were visualised on a UV transilluminator at 305nm and photographed using a video copy processor (Mitsubishi).

ϕx*Hae*III (bp)	λ*Hind*III (bp)
1353	23130
1078	9416
872	6682
603	4361
310	2322
281	2027
271	546
234	125
194	
118	
72	

Table 2.5: DNA size markers ϕx*Hae*III and λ*Hind*III.

2.5.3 Purification of DNA from agarose

DNA fragments were run in low melting point agarose (Ultrapure LMP agarose, Gibco BRL) in 1 X TAE. Gels were viewed on a UV transilluminator at 305nm and the required fragment rapidly excised using a sterile scalpel

blade. Care was taken to ensure that a minimum of agarose was excised with the required band of DNA.

DNA was isolated from agarose using Geneclean (Bio 101 Inc.) according to the manufacturer's instructions.

DNA to be used as hybridisation probes was treated as follows. After running DNA on a 1% LMP gel, the required fragment was excised and weighed. 1-3X its weight in water was added to it, the mixture boiled for 2 minutes and 20µl aliquots made. Each aliquot, containing 25-50ng DNA, was stored at -20°C and then used directly in labelling reactions.

2.5.4 Polyacrylamide gel electrophoresis

Denaturing polyacrylamide/urea gels were used to resolve DNA sequencing reactions (see section 2.5.1) and end-labelled PCR products (see section 2.10.7).

Glass plates were prepared by washing thoroughly in detergent and distilled water, followed by a final rinse with 100% ethanol. To prevent the gel from adhering to the front plate, it was coated with dimethyldichlorosilane (Analar, BDH) and left to dry. Spacers were placed at the sides of the back plate and the front plate was placed on top. The sides and bottom of the plates were sealed together with tape and clamps used to hold them in place. Freshly prepared polyacrylamide was poured between the plates and a sharks-tooth comb was inserted at the top of the gel. The gel was allowed to dry for at least an hour at room temperature before running.

Gels were run in 1XTBE on a vertical slab gel apparatus (Gibco, BRL) at 55 volts for 2-6 hours. Tape at the bottom of the plates was removed prior to running and wells were flushed out with buffer before loading.

After electrophoresis, tape at the sides of the plates was removed and the gel blotted onto 3MM Whatman paper, covered with Saran wrap (Dow Chemical

Company) and dried at 80°c on a Drygel Slab Drier (Hoefer) for one hour. Once dry, the Saran wrap was removed and the gel exposed to Biomax sequencing film (Kodak) or to a phosphorimage screen, overnight at room temperature.

Non-denaturing 8% acrylamide gels were prepared as in section 2.5.1 and were run in The Sturdier vertical slab gel apparatus (Model SE 400, Hoefer Scientific Instruments) using 1.5mm spacers and comb. 8% acrylamide gels were used to separate small PCR fragments with a size difference of 5-15bp. 2-10µl of PCR sample was mixed (10:1) with loading buffer and run at 250-300 volts for 2-3 hours, depending on the size of the fragments, in 1 x TBE.

DNA electrophoresed on 8% acrylamide was visualised using silver staining. On separating the glass plates, the gel was left on the back plate and washed twice in a solution of 10% ethanol and 0.5% acetic acid. It was then incubated in 500mls 1g/L silver nitrate (Analar, BDH) for 10 minutes at room temperature, followed by two washes in distilled water. The silver stain was developed by incubating the gel with a staining solution (0.375M NaOH, 2.6mM $NaBH_4$, 0.15% formaldehyde) for 20 minutes. Gels were photographed using normal light and a video copy processor (Mitsubishi).

2.5.5 HydroLink gel electrophoresis

Non-denaturing HydroLink Mutation Detection Enhancement gels (MDE, AT Biochem Inc) were used to detect conformational differences in double stranded molecules (heteroduplexes). These gels were used to map loci where restriction fragment length polymorphisms (RFLPs) could not be easily found.

PCR products were denatured at 90°C for 10 minutes and allowed to cool slowly, thus promoting the formation of annealed heteroduplexes, before being run out on a MDE gel. 1µl loading buffer was added to 10µl sample prior to loading. Gels were poured in the vertical slab Sturdier apparatus and

run at room temperature for approximately 16 hours at 40 volts in 0.6 X TBE. DNA was visualised by ethidium bromide staining.

2.5.6 Pulse field gel electrophoresis

Agarose plugs were washed once in TE and equilibrated in the running buffer (0.5 X TBE) over three hours. A 1% agarose gel was made up in 0.5 X TBE and the plugs inserted into the wells. The wells were sealed with 1% agarose and the gel run in pre-cooled 0.5 X TBE at 12°C. Gels were run using the Chef-DR II pulse field tank and control module (Biorad) and run times and conditions were varied depending on the size resolution required.

After electrophoresis, DNA was visualised by staining the gel in running buffer containing ethidium bromide at 5μg/ml for 40 minutes.

2.6 Transfer of DNA to membranes

2.6.1 Southern transfer protocol

DNA was transferred from gels to nylon membranes by capillary blotting. This method was adapted from Southern (1975).

Prior to blotting, gels were photographed with a ruler to allow for future sizing of DNA fragments. Pulse field gels underwent an additional pre-blotting step where they were UV-irradiated at 305nm for 5 minutes. This was done to aid the transfer of large yeast chromosomes and YACs.

The DNA was denatured by incubating the gel in denature solution (0.5M NaOH, 1.5M NaCl) for 45 minutes. The gel was then neutralised with neutralising solution (1M Tris.HCl, 2M NaCl, pH5.5) for an additional 45 minutes. A large strip of 3MM filter paper (Whatman) was soaked in 20 X SSC (3M NaCl, 0.3M $Na_3C_6H_5O_7.2H_2O$, pH7.0) and placed on a board. The ends of the paper were placed in a reservoir of 20 X SSC, generating a wick, and the gel was placed on top of the wick and a piece of nylon membrane (N-bond, Amersham), cut to fit and pre-soaked in 2 X SSC, was placed directly

onto the gel. Air bubbles were removed and two pieces of 3MM blotting paper, also pre-soaked in 2 X SSC, were placed on top of the membrane. Any exposed wick was screened off with Saran wrap (Dow Chemical Company) and a weighted-down stack of paper towels placed on top.

The length of blotting time varied according to the size of fragments which were being transferred. Pulse field gels were blotted for two days, genomic digests and smaller fragments overnight.

After blotting the membranes were rinsed briefly in 2 X SSC and DNA was bound to the filter by exposing it to 1200μJoules of UV irradiation (Stratagene's UV Stratalinker). Nylon membranes were stored in Saran wrap at room temperature. Gels were also restained after blotting with ethidium bromide, to check that all the DNA had been transferred.

2.7 Enzymatic manipulation of DNA

2.7.1 Restriction enzyme digestion of genomic DNA

Digestion of DNA with restriction endonucleases was carried out in the appropriate buffer at the recommended temperature. Commonly, 5μg of DNA was digested in 50μl volume in 10 units of enzyme for 2-3 hours. High molecular weight DNA was often digested overnight in the presence of 5mM spermidine (Sigma) and the amount of restriction enzyme used reduced accordingly. If two different enzymes were used, both of which required the same buffer, the digests were carried out simultaneously. Otherwise, digestion with the enzyme requiring the lower salt buffer concentration was carried out first. The salt concentration was then altered by addition of appropriate amounts of sodium chloride before the second digestion was carried out. When necessary, reactions were terminated by heating to 68°C or 80°C for 15 minutes, according to the heat sensitivity of the enzyme (NEB Catalogue, 1992). Restriction digests which were to be run on gels were

terminated by the addition of 1/10th of their volume of "stop mix" (100mM EDTA, pH8.0; 20% Ficoll and Orange G).

2.7.2 Restriction digests of agarose plugs

Agarose plugs, used for pulse field analysis, were pre-treated with phenylmethylsulfonyl fluoride (PMSF, Sigma) prior to restriction. PMSF is a potent inhibitor of Proteinase K and was used to inactivate residual Proteinase K present in the plugs as a result of their preparation (see section 2.2.4). Directly prior to use a 20mg/ml solution of PMSF was prepared, at 55°C, in propan-2-ol. The protocol followed was that from Maule, 1994.

Plugs were soaked in a large excess of sterile TE at room temperature for 10 minutes. They were then transferred to a solution of TE containing 40µg/ml PMSF. The plugs were incubated in this solution for 1 hour at 50°C with one change of solution after 30 minutes. The plugs were then soaked in 1 X restriction enzyme buffer for 2 hours at room temperature, with frequent inverting.

For restriction enzyme digestion, each plug was placed in a 1.5ml microfuge tube (Treff, AG) and put on ice. To it was added 0.1% Triton X100 (Sigma) and 200µg/ml Bovine Serum Albumin (BSA, Boehringer Mannheim). 20 units of restriction enzyme was then added and the volume adjusted to 200µl. Digestion was carried out overnight at the appropriate temperature. Following digestion, the tube was cooled on ice for 15 minutes, to facilitate the handling of the plugs, and the liquid removed with a fine tip pastette. The plugs were washed once in ice cold TE and the reaction stopped by placing the plugs in 200µl stop buffer (0.5 X TBE, 10mM EDTA, 2mg/ml Orange G) for 20 minutes before loading the plugs onto the gel.

2.7.3 Restriction enzyme digestion of PCR products

Usually, 20µls of PCR product was digested in a volume of 30µls with 1-2 units of restriction enzyme. The appropriate amount of 10X concentrated

restriction buffer was added to the digestion mix. When using enzymes which required a low salt buffer, the PCR product was cleaned using Wizard PCR Preps DNA Purification System (Promega) following the manufacturer's instructions. If necessary, this was preceded by a gel purification step (see section 2.4.3).

2.7.4 Restriction of plasmid pBR322

To generate probes which would recognise the right and left arms of the YAC vector, pYAC4, a double digest of the plasmid pBR322 was carried out using the restriction enzymes, *Pvu*II and *Bam*HI. Since both enzymes require the same salt buffer, the digestions were carried out simultaneously. 2µg plasmid DNA was restricted in 30µl volume with 5 units of each enzyme, at 37°C for 3 hours. The digests were electrophoresed on a 1% LMP gel (section 2.5.3) and two bands excised from the gel for use as hybridisation probes. The band used as the right arm probe was 1.4 kb in size, and that used for the left arm probe was 2.6kb in size.

2.8 Radiolabelling of DNA

2.8.1 Random Priming of DNA probes

Probes were labelled with [^{32}P]dCTP by random priming from hexadeoxyribonucleotides using the Klenow fragment of *E. Coli* DNA Polymerase I (Feinberg and Vogelstein, 1983, 1984).

25-50ng DNA (either in the form of a gel slice or in solution) in a volume of 20µl was denatured by incubation at 100°C for 10 minutes. If the DNA was in solution, the reaction was cooled rapidly on ice, or if a gel derived probe was used, the reaction was cooled at 37°C for 3 minutes. A Random Priming kit (Boerhinger Mannheim) was used to label the probe. 3µl of 10X reaction buffer, 1µl each of dATP, dTTP, dGTP (0.5mM), 2 units of Klenow enzyme and 30µCi [^{32}P]dCTP (Amersham) were added to the denatured probe, which was then incubated at 37°C for at least 45 minutes. At the end of the

incubation the amount of [^{32}P]dCTP label incorporated into the probe was assayed for using trichloroacetic acid precipitation (Sambrook *et al*, 1989). Once a sufficiently high level of incorporation had been achieved (at least 50%), unincorporated nucleotides were removed from the probe using a Nick Column (Pharmacia). On removal of the storage buffer, the column was equilibrated in 800µl TNE (10mM Tris.HCl, pH8.0; 1mM EDTA, pH8.0; 0.2M NaCl). TNE was also added to the probe, to bring its volume to 100µl, before it was added to the column with a further 300µl of TNE. The probe was then eluted with a further 400µl TNE. 500µg of sonicated salmon sperm DNA was added to the probe and the mix denatured together at 100°C for 10 minutes. The probe was chilled on ice for 10 minutes before being added to a Hybaid hybridisation bottle.

2.8.2 Preannealing of repetitive sequences

Probes containing repeat sequences were annealed, after labelling, to repeat-containing DNA. This was done to prevent the repetitive part of the probe from taking part in the hybridisation reaction. This procedure was adapted from Sealey *et al*, 1985.

After the probe was labelled, it was denatured at 100°C for 10 minutes in the presence of 1mg sonicated mouse DNA. The two were then incubated together in a pre-warmed lead pot at 65°C for 30 minutes, before being added to the hybridisation mixture.

2.8.3 End labelling of DNA oligonucleotides

Oligonucleotides were labelled by the transfer of a ^{32}P-labelled phosphate onto the terminal 5'-OH group.

30ng oligonucleotide was labelled in a 20µl volume which contained 1X Polynucleotide kinase buffer (5mM Tris.HCl, pH8; 1mM $MgCl_2$, 0.5mM DTT, Boehringer Mannheim), 10 units of polynucleotide kinase (Boehringer Mannheim) and 30µCi [^{32}P]γATP (Amersham). Reactions were incubated at

37°C for 30-40 minutes and added to a hybridisation bottle containing an oligo hybridisation solution.

2.9 Hybridisation protocols

2.9.1 Hybridisation solutions

Hybridisation mix

6X SSC, 10% dextran sulphate, 0.1% sodium pyrophosphate, 0.5% sodium dodecyl sulphate, 4X Denhardts solution (from 100X Denhardts solution: 2% BSA; 2% polyvinylpyrrolidone; 2% ficoll; 1mM EDTA; stored at 4°C).

Oligo hybridisation mix

5X SSC, 0.05% BSA, 0.05% ficoll, 0.1% sodium dodecyl sulphate, 0.05% polyvinylpyrrolidone, 0.1% sodium pyrophosphate.

2.9.2 Pre-hybridisation protocols

When an end-labelled oligonucleotide was used, the pre-hybridisation of filters was carried out in oligo hybridisation mix at 5°C below the T_m of the oligonucleotide (section 2.10.1). For random primed probes, pre-hybridisation took place at 68°C in hybridisation mix, except for the hybridisation of YAC end fragments (section 2.10.6) onto YAC, or mouse genomic, DNA filters, when a hybridisation temperature of 65°C was used. Filters were placed between two gauze sheets, soaked in 2X SSC and placed in a hybridisation bottle (Hybaid). Air bubbles were removed from the gauze and 10ml of hybridisation mix added to the small bottles and 20ml to the larger ones. Bottles were placed in a Hybaid hybridisation oven on a rotating spindle for at least one hour before the addition of labelled probes. When salmon sperm was used to block the probe, a 100μg/ml solution was denatured and added to the hydridisation mix.

2.9.3 Hybridisation and washing protocols

End-labelled oligonucleotides hybridisations were carried out for 4-16 hours at the appropriate hybridisation temperature, usually 45-48°C. Filters were removed from the hybridisation bottles and given three, five minute washes

using a solution, at the same temperature at which the hybridisation was carried out in, of 4X SSC, 0.1% SDS.

Random primed hybridisations were incubated at 68°C or 65°C for 16 hours. Filters were washed first in their bottles using a wash solution of 2X SSC, 0.1% SDS for 20 minutes at 68°C (65°C for YAC end fragment hybridisations). They were then removed from the bottles and washed with increasing stringency with 1X SSC, 0.1% SDS and finally in 0.1 X SSC, 0.1% SDS at 68°C, or 65°C.

2.9.4 Detection of hybridisation signal

Autoradiography The filters were placed in a light-safe cassette with a signal enhancing screen. They were then exposed to X-OMAT x-ray film (Kodak) for a length of time dependent on the amount of radiolabelled probe left bound to the filter (several minutes to several days). Filters hybridised to ^{32}P-labelled probes were exposed at -70°C, those labelled with ^{35}S were exposed at room temperature. Stratagene Glogos II luminescent markers were used for alignment. The film was developed on an automatic x-ray film processor RGII (Fuji).

Phosphorimaging Alternatively, the filters were exposed to a phosphor screen (Molecular Dynamics) for periods ranging from a few hours to several days. The screen is scanned on a PhosphorImager (Molecular Dynamics), where a laser beam converts the radioactive signal into a digital image, with variations in the pixel value proportional to the amount of radioactive signal present. The grey-scale image was adjusted as desired and the image printed out on a laser printer (Hewlett and Packard).

2.9.5 Removal of radioactive probes from filters

Filters may be used several times with different hybridisation probes, so it was sometimes necessary to remove radioactive probes from filters. This was done in one of two ways. Firstly, the filters could be washed in 0.4M NaOH at

50°C for 30 minutes, to denature the probe DNA strand from the surface of the filter. The neutral pH was then restored by a 30 minute wash in a solution of 0.1x SSC, 0.1% SDS and 0.2M Tris.HCl, pH 7.5, also at 50°C. Alternatively the filters could be subjected to one very stringent wash, in 0.1% SDS at 100°C for 5 minutes.

The filters were then exposed to x-ray film overnight to check that the radioactive probe had been removed.

2.10 Amplification of DNA by the Polymerase Chain Reaction (PCR)

PCR is a technique used for the amplification of a specific DNA sequence (Saiki *et al*, 1988; Mullis and Faloona, 1987). The specificity is provided by oligonucleotide primers complementary to the 5' ends of the two strands of the sequence to be amplified. These primers anneal to the template DNA strand and direct amplification by a thermostable polymerase. The PCR reaction is a series of cycles, each consisting of three steps. The first is a short high temperature step (92-94°C) to denature the template. This is followed by a reduction in temperature to allow the primers to anneal to the template (~50-65°C depending on the primers used), then finally the temperature is increased to 72°C to permit extension from the primers by the polymerase. Multiple cycles of these three steps result in exponential amplification of the desired sequence.

2.10.1 PCR conditions

Commonly 100ng of template genomic DNA was amplified in a 50µl total volume, with 0.2µl (1 unit) Amplitaq (Perkin Elmer Cetus) in 1x reaction buffer (10mM Tris.HCl, pH 8.3, 50mM KCl, 1.5mM $MgCl_2$, 0.01%(w/v) gelatin) and 0.1% Triton X-100 (Sigma). The reaction employed 300ng of each primer and 200µM each dNTP (Advanced Biotechnologies). The reaction was overlaid with 40µl mineral oil (Sigma).

PCR programmes were run on an Hybaid Omnigene machine. The basic programme was 30 cycles of 3 different steps:

a) Denaturation: A short step at 92-94°C. Usually 3 minutes in the first cycle and 30 seconds thereafter.

b) Annealing: The annealing temperature is critical in determining the success of the PCR reaction and is dependent on the structure of the primer. The annealing temperature of the primer was always 5°C lower than the melting temperature (T_m) of the primer, which was determined using the equation:

$$T_m \text{ at 1M } Na^+ \text{ concentration} = 4(G+C) + 2(A+T)$$

In a number of cases conditions were modified in order to optimize amplification. If, at first, no product was produced then the annealing temperature was dropped by 2°C until a product was amplified. The $MgCl_2$ concentration was also modified if amplification failed. Where non-specific amplification occurred, the annealing temperature was increased by 2°C until a product of a specific size was seen.

c) Extension: The extension time is also critical in determining the success of the PCR, since larger products need longer extension times (approximately 1 minute per kb of sequence). Thirty cycles of 72°C was generally used but the number of cycles were increased where amplification after 30 cycles was insufficient. If the size of the product was unknown, longer extension times were used with an increase after 10 - 20 cycles.

After the reaction, 5-10µl of the PCR reaction was visualised by electrophoresis on 1-4% agarose gels (see section 2.4.2).

2.10.2 Oligonucleotide Synthesis

Oligonucleotides were synthesised as ammonium stocks on an Applied Biosystems 381A oligonucleotide synthesiser.

Oligonucleotides were precipitated from ammonium stocks by the addition of 1/10th volume of 3M sodium acetate and 2.5 volumes of ethanol, followed by one hour incubation at -20°C. The precipitated DNA was pelleted by centrifugation at 4°C, washed twice in 80% ethanol and resuspended in 200µl TE. The concentration of the DNA was determined by measuring the optical density of the solution, where the absorbance at 260nm (ABS_{260}) = 1 at 25µg/ml single stranded DNA.

2.10.3 Oligonucleotide primer design

Primers were designed using the programme Oligo4 (Hybaid). Primers were designed from available sequence information and were between 18 and 25 oligonucleotides in length depending on the DNA source on which they would be used. They were chosen to be stable oligonucleotides with a T_m between 45°C and 65°C.

2.10.4 Oligonucleotide sequence

The primer sequences, conditions of amplification, size of the amplified fragment and the source of sequence information is given in table 2.10.4 (at the end of the chapter).

2.10.5 PCR from yeast colonies

This method was adapted from Taylor (1991).

Individual yeast colonies were picked with a sterile wooden toothpick and resuspended in 50µl of dH_2O. The resuspension was boiled for 10 minutes, spun down and 3µl of supernantant used in a 25µl volume PCR reaction, as detailed in 2.9.1.

2.10.6 Inverse PCR

Adapted from Arveiler and Porteous, 1991.

This method was used to isolate fragments from the ends of YACs. 2µg YAC DNA was digested with a range of enzymes in a volume of 20µl (see section 2.6.1). For right arm end fragments, YAC DNA was digested individually with *Pst*I, *Rsa*I, *Eco*RV and *Nde*I, and for left arm end fragments *Taq*I, *Rsa*I, *Eco*RI and *Sac*I were used (all enzymes from Boehringer Mannheim). The digestion products were cleaned using Geneclean (Bio 101), resuspended in 20µl H_2O and ligated under dilute conditions, to generate monomeric circles.

Ligation reactions were carried out in a 10µl volume which contained: 0.1 unit of T4 DNA ligase (Boehringer Mannheim), 1µl of the digested, cleaned DNA and 1µl of 10X ligase buffer (Boehringer Mannheim, 50mM Tris.HCl, pH7.4; 100mM $MgCl_2$; 10mM DTT; 1mM spermidine; 1mM ATP; 100g/ml BSA).

The reaction was incubated overnight at 16°C, after which the enzyme was heat inactivated by incubation at 68°C for 20 minutes. 1µl of this ligation reaction was used in a 50µl volume PCR reaction as described in section 2.9.1.

PCR products from inverse PCR reactions were run out on 1% agarose gel. Each checking gel was Southern blotted (section 2.5.1) and the filter probed with internal oligos; the right arm oligo used was 770 and the left arm one was 768 (see table 2.10.4). This was done to verify which inverse PCR products were derived from YAC inserts and which had arisen as a result of non-specific amplification. Where 2 or more bands arose following inverse PCR, the PCR reaction was run on a 1% LMP gel (section 2.5.3) and the upper band was excised for use as an end fragment probe.

2.10.7 PCR using end-labelled oligonucleotides

End-labelled oligonucleotides were used for PCR amplification when the species-specific size polymorphisms in PCR products were too small to be resolved on 4% NuSieve or 8% acrylamide gels.

One of the PCR primers was end-labelled in a 20µl volume labelling reaction, as in section 2.8.3. The PCR was set up as detailed in section 2.10.1, with the addition of 1µl of the end-labelled oligonucleotide. Before loading onto the gel, 5µl of the PCR reaction was mixed with 5µl of Stop solution (95% deionised formamide, 20mM EDTA, 0.05% bromophenol blue, 0.05% xylene cycinol FF) and denatured at 90°C for 10 minutes, then chilled on ice. 4µl of the sample was then loaded onto a 6% denaturing polyacrylamide gel (see section 2.5.4) for electrophoresis.

2.11 DNA sequencing

2.11.1 Sequencing of PCR products

This method has been modified from Winship, 1989.

In order to directly sequence PCR products, the required product was purified from a gel in order to separate it from primer dimers and spurious PCR products. A substantial amount of DNA was required in the reaction itself, so in some cases, the product of several PCR reactions were pooled and concentrated into a smaller volume by ethanol precipitation. The resulting product was gel purified by running it on a 1% LMP agarose gel and the required band was excised. The DNA was recovered from the gel slice as described in section 2.5.3. Following resuspension, the DNA concentration was checked by running 1µl of the sample on an agarose gel, next to a known amount of size marker. Approximately 200ng DNA was used per sequencing reaction.

The sequencing reactions were carried out using a Sequenase Version 2.0 kit (Amersham). Firstly, the DNA was mixed with 300ng sequencing primer, 1.25µl DMSO and 2.5µl 5x Sequenase reaction buffer and made up to a total volume of 12.5µl with water. This mix was denatured by heating it to 100°C for 3 minutes. Following denaturation, the tube was immediately removed to liquid nitrogen or dry ice, to snap freeze the DNA mix and prevent the strands from re-annealing.

A labelling mix was made up from the kit containing 1.75µl enzyme dilution buffer, 1µl DTT, 0.4µl dGTP labelling mix and 0.25µl Sequenase enzyme per reaction. To this was added 1.6µl water and 0.5µl of 10µCi/µl [α-^{35}S]-dATP, and the tube kept on ice. One tube of DNA mix was thawed quickly, then 5.5µl of the labelling mix was immediately added and mixed by pipetting. 3.5µl of this was then applied to the side of 4 Eppendorf tubes, each one containing 2.5µl of one of the 4 termination mixes containing 10% DMSO (G,A,T and C). The 4 tubes were then briefly spun to combine the labelling and termination mixtures simultaneously. The time from thawing the DNA, to mixing with the termination mix, was kept to approximately 1 minute.

Termination tubes were transferred to a 37°C water bath and incubated for 5 minutes. Reactions were stopped by the addition of 4µl stop mix (95% formamide, 20mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF) and were stored at -20°C for up to one week. Prior to loading, the sequencing reactions were heated to 80°C for 2 minutes to denature the DNA and 2.5µl of each sample was electrophoresed on a 6% denaturing polyacrylamide gel (section 2.5.4).

2.12 Statistics used in backcross analysis

2.12.1 Calculation of genetic distance

Genetic distances, measured in centiMorgans (cM), were calculated using the following equation:

Distance in centiMorgans = (R/ N) X 100

where R = the number of recombinants

and N = the number of animals haplotyped at that locus

2.12.2 Calculation of standard errors

In order to calculate the standard error for genetic distances, the following equation was used and the result multiplied by 100 for use with cM distances:

$$SE = \sqrt{\frac{pq}{n}}$$

where: p = recombination frequency

q = 1 - p

n = no. of animals

Table 2.10.4: PCR primer sequences and reaction conditions

Locus	Primer Sequence	Reaction conditions	Reference
D2Mit1	5' CAC TTG ACA CCA ACC CCT G 3' 5' CAT TCT ATA CTG TTC CGA AAG AAA CC 3'	94°C for 3 min x1 94°C for 30 secs] 55°C for 1 min] x30 72°C for 1 min]	Dietrich *et al* (1994) Copeland *et al* (1993) Whitehead/MIT Center Database release (1995)
D2Mit31	5' ATT TGA AAA GAC TGG ACT CCT CC 3' 5' TGA CTG GGA ATA ACC TCT CTG C 3'	As above	As above
D2Mit119	5' AGG TCC AGG TTC AAT GCA AG 3' 5' AGA GAC AGC TTC GTG TGC AA 3''	As above	As above
D2Mit149	5' ATA TCA TAT ATA GTA GAG AAG CGT GCT G 3' 5' TCA TTA GAC TGG AAA AAA GTTT GC 3'	As above	As above
D2Mit118	5' TGG CTT CCA CAT GCA CAC 3' 5' GCA TAT GTT TGA CCA GTT TTT TAT G 3'	As above	As above

Table 2.10.4: PCR primer sequences and reaction conditions

Locus	Primer Sequence	Reaction conditions	Reference
D2Mit267	5' TGT CTA AGG GAT AGG CCA CG 3' 5' AGC ACA TAT GCT CGC ACT TG 3'	94°C for 3 min x1 94°C for 30 secs] 55°C for 1 min] x30 72°C for 1 min]	Dietrich *et al* (1994) Copeland *et al* (1993) Whitehead/MIT Center Database release (1995)
D2Mit 362	5' ATG CAG AGC ACT TAA ATT TCA TAG G 3' 5' AAC CTT TCT CTG GTA CAA ATT GC 3'	As above	As above
D2Mit464	5' GAA TAG GGG AGT TAG AGG GGG 3' 5' GAA ACT CTT TTT TTT TCT TCA GTG TG 3'	As above	As above
D2Mit178	5' TTA TGG TGG TGG TGA TGG TG 3' 5' ACT GAC TGA GAT GGC TTG ATC A 3'	As above	As above
D2Mit80	5' TAG CCT ACA GAG TGG ACA GCC 3' 5' CTAGGCTTTATGTAGCCTCTTTGC 3'	As above	As above

Table 2.10.4: PCR primer sequences and reaction conditions

Locus	Primer Sequence	Reaction conditions	Reference
D2Mit6	5' AAC AAA CAA ACC CCT TGC CC 3' 5' CTC TAA CAC AGC CCC AGG TG 3'	94°C for 3 min x1 94°C for 30 secs] 55°C for 1 min] x30 72°C for 1 min]	A Dietrich *et al* (1994) Copeland *et al* (1993) Whitehead/MIT Center Database release (1995)
D2Mit292	5' GAA GGT GTG GGT GCA TTT G 3' 5' TGA GAT TGT TCT CTG ACC TCC A 3'	As above	As above
D2Mit79	5' TAG AGG AAG CAA GCC ACA CA 3' 5' GAC ATG TGA CAT GAA TGC TGC 3'	As above	As above
D2Mit7	5' AAG GCA AGC ATT CTG CCA CT 3' 5' CTC CCG GCA AGA ACT GTT TT 3'	As above	As above
D10Mit80	5' CAA AAA AAA CCC TGA TTC TAC CA 3' 5' GTG TGC ATA TGG CAG TAA CTT TG 3'	As above	As above

Table 2.10.4: PCR primer sequences and reaction conditions

Locus	Primer Sequence	Reaction conditions	Reference
D10Mit105	5' AAT CTG GTC ACA CAG ATT TGA CC 3' 5' GTC ATC CCT GTG TAG TGT CAA CA 3'	94°C for 3 min x1 94°C for 30 secs] 55°C for 1 min] x30 72°C for 1 min]	Dietrich *et al* (1994) Copeland *et al* (1993) Whitehead/MIT Center Database release (1995)
D10Mit21	5' CAT CGA TGG AGC TGT GAA GA 3' 5' TAA GAC TCC CCC GGG AAG 3'	As above	As above
D10Mit31	5' CAT AAG GAG CAC AGG CAT GA 3' 5' CCC TGT ACG TGC ATG CTG TA 3'	As above	As above
D10Mit121	5' GAT GAA CAT GAC AGA AGG TAG CC 3' 5' TGG CCT CCA TGT GTA CAT ACA 3'	As above	As above
D10Mit164	5' CAA TTC CTG GAT AGC CTG GA 3' 5' CTC CTT ATG CAC CCC TCA AA 3'	As above	As above

Table 2.10.4: PCR primer sequences and reaction conditions

Locus	Primer Sequence	Reaction conditions	Reference
Pax8	5' GTC ATG GGG ACA TGT GGA AG 3' 5' CCT AGG CTG TGA TTG TAG GGC 3'	95°C for 2 min x1 92°C for 1min] 54°C for 1 min 30 secs] x33 72°C for 1 min 30 secs]	Plachov *et al* (1990)
Il1rn	5' GGC CTG TAA TAA TCA CCA AC 3' 5' CAG TTT AAG GCA TCA CAC AA 3'	95°C for 2 min x1 92°C for 1min] 55°C for 1 min] x33 72°C for 1 min 30 secs]	Zahedi *et al* (1991)
nodal	5' CTC CAC AAT CAT GTC CTT GTG 3' 5' GGC GAG TGT CCT AAC CCT GTG 3'	95°C for 2 min x1 92°C for 1min] 50°C for 1 min 30 secs] x30 72°C for 2 min]	Zhou *et al* (1993)
Inverse PCR Left arm Right arm	 5' GAA TGG ATC CAC AGG ACG GG 3' 5' GCC AAG TTG GTT TAA GGC GC 3' 5' GGA AGA ACG AAG GAA GGA GC 3' 5' GCC CGA TCT CAA GAT TAC G 3'	94°C for 3 min x1 92°C for 45 sec] 60°C for 1 min] x17 72°C for 2 min 30 sec] 92°C for 45 sec } 60°C for 1 min } x17 72°C for 3 min } 72°C for 10 min x1	Arveiler & Porteous (1991)

Chapter 3: Chromosome 2 mapping in the region of *Sd*

3.1 Introduction

In mouse genetics, the search for a gene responsible for a mutation with an unknown biochemical basis begins with the establishment of a detailed linkage map. This approach is also used in human genetics, but mapping genes in the mouse is simpler than mapping genes in humans because controlled matings between different strains of mice can be used. Mouse genetic mapping serves two distinct purposes. Primarily, mapping allows a locus responsible for a mutation to be placed onto a map relative to other chromosomal markers, either probes derived from functional genes or anonymous DNA markers. In this way, known genes can be mapped with respect to mutant loci. If a gene fails to recombine with such a locus, and if it has an expression pattern or known biochemical function which correlates with the mutation's aetiology, such a gene can be considered as a "candidate" gene for the mutation. Several genes have successfully been cloned in this way, including the α_1 subunit of the inhibitory glycine receptor, responsible for the mouse mutation *spasmodic* (Ryan *et al.*, 1994) and *tyrosinase-related protein*, the gene responsible for the mouse mutation *brown* (Jackson, 1988). The drawback to this strategy is that cloning by the candidate gene approach still depends on partial functional information about the disorder being available, allowing for "educated" guesses to be made as to the identity of the gene responsible. To some extent, high resolution mapping can reduce the guesswork involved in the candidate gene approach since the resolving power of a cross increases with its size. In this way, genes which appear to be ideal candidates for a mutation can be excluded if they recombine outwith of a defined interval, the size of which decreases with the number of animals mapped.

The other main purpose that mouse genetic mapping serves is in the development of physical maps. Physical maps become vital when mutant genes cannot be identified by the candidate gene approach, thus necessitating the initiation of a positional cloning exercise (Collins, 1995).

Markers that have been finely mapped and ordered via backcross, or intercross, mapping serve as a scaffold onto which overlapping YAC or cosmid clones can be placed and positioned relative to each other and the mutant locus.

Mapping in the mouse is not without its drawbacks. In situ hybridisation and somatic cell hybrids, which have been powerful tools in the mapping of human disorders, have been used less successfully in mouse genetic mapping. In the case of *in situ* hybridisation, this has been due to the acrocentric nature of mouse chromosomes making it harder to cytologically discriminate between individual chromosomes. In addition, somatic cell hybrids that carry single mouse chromosomes, or chromosomes with deletions or translocations, are not as widely available in mouse genetics as they are in human genetics (Copeland *et al.*, 1993). For this reason, mouse genetics relies heavily on backcross and intercross mapping for the assignment of markers to loci of interest. In response to this emphasis, new types of genetic markers have been developed to allow for the rapid generation of high resolution linkage maps. Mouse mapping traditionally relied on the identification of restriction fragment length variants (RFLVs) which allowed published genes to be mapped using inherent polymorphisms between strains or species of mice. A more rapid and recent means of mapping, has been to utilise simple sequence repeat variants (usually di-, tri- or tetranucleotide repeats), which have a high frequency of polymorphic length variation between strains and species of mouse. These are PCR-amplifiable and have been generated in large numbers, now covering most of the mouse genome (Love *et al.*, 1990; Dietrich *et al.*, 1994).

In addition to the rapid generation of new markers, a substantial improvement in the power of backcross and intercross mapping came with the introduction of interspecific crosses (Avner *et al.*, 1988). Until relatively recently, mouse gene assignments relied on crosses between laboratory

strains of mouse or on recombinant inbred strains (Taylor, 1989). These crosses were, however, limited by the low degree of allelic variation between the strains, a problem which was overcome with the introduction of interspecific crosses. Interspecific crosses exploit the genetic diversity between wild species of mouse (such as *Mus spretus*) and laboratory strains. For this reason, most genes and DNA fragments are variant in an interspecific cross and can be ordered relative to other genes and markers in a single cross. There are, however, disadvantages in using *Mus spretus* as the wild mouse species in an interspecific cross, the main one being that the F1 males are infertile, thus preventing the study of male meiosis and slowing down the accumulation of sufficient backcross progeny to allow for high resolution mapping. It is for these reasons that we chose to use the *Mus musculus* sub-species, *Mus musculus castaneus*, in place of *Mus spretus* in setting up the *Sd* backcross, since both male and female F1 mice are fertile. Using *M.m.castaneus* has allowed us to rapidly generate sufficient numbers for mapping *Sd* in a highly polymorphic backcross.

3.2 *Sd* backcross

Mice heterozygote for Danforth's short tail (*Sd*) mutation were obtained from Harwell and were maintained by the Biomedical Research Facility, at the Western General Hospital. An *Sd* stock was established by crossing the *Sd* mice to an inbred laboratory strain of mouse derived from CBA/Ca, for two generations prior to setting up the backcross. *Sd* mice were then mated with inbred *M.m.castaneus* CAST/Ei mice to produce the F1 generation. Thirteen F1 heterozygous *Sd* mice, both male and female, were set up in breeding pairs with CBA mice to produce the backcross progeny. All the backcross offspring were killed at 24 days of age except in a small number of cases, when backcross offspring were killed at a later age due to errors in the identification of pregnant females.

The backcross generated 336 offspring: 156 females and 180 males. *Sd* is fully penetrant in this backcross with a ratio of 164 *Sd* mice to 163 wild type mice, showing the expected 50:50 ratio of a fully penetrant genetic trait. The first nine offspring of the backcross were not included in this calculation as the wild type mice from these litters were mistakenly killed. For a summary of the breeding protocol used in this backcross, see figure 3.2.

3.3 Phenotypic records of backcross progeny

A detailed phenotypic record was kept of each backcross animal, both wild type and *Sd*. Kidneys were dissected and their size and appearance recorded (kidney length was measured in millimetres). It was noted whether there was a difference in size between the two kidneys, and if so, from which side of the animal (left or right) the larger/ smaller kidney came from. Tail length was measured and any abnormalities in the tails' appearance were noted. If the overall size of the animal was different from its littermates, and on the few occasions when paralysis of the hind limbs was evident, this was noted along with the overall condition of the animal. For a summary of results of the phenotypic analysis of the backcross progeny, see table 3.3.1 at the end of the chapter.

Bred onto a CBA background for two generations, the *Sd* mice used in this study and produced by the backcross did not show the high rate (70%) of postnatal mortality observed in the original *Sd* mice (Dunn *et al.*, 1940), suggesting that the phenotype of the mutation on this genetic background is not as severe as it has been on others (Dunn *et al.*, 1940). Data presented in table 3.3.1 indicates that the increased viability of heterozygous mice, in this study, is likely to be due to the decreased severity of the kidney phenotype on this genetic background, as only 2.4% of the backcross progeny had one kidney.

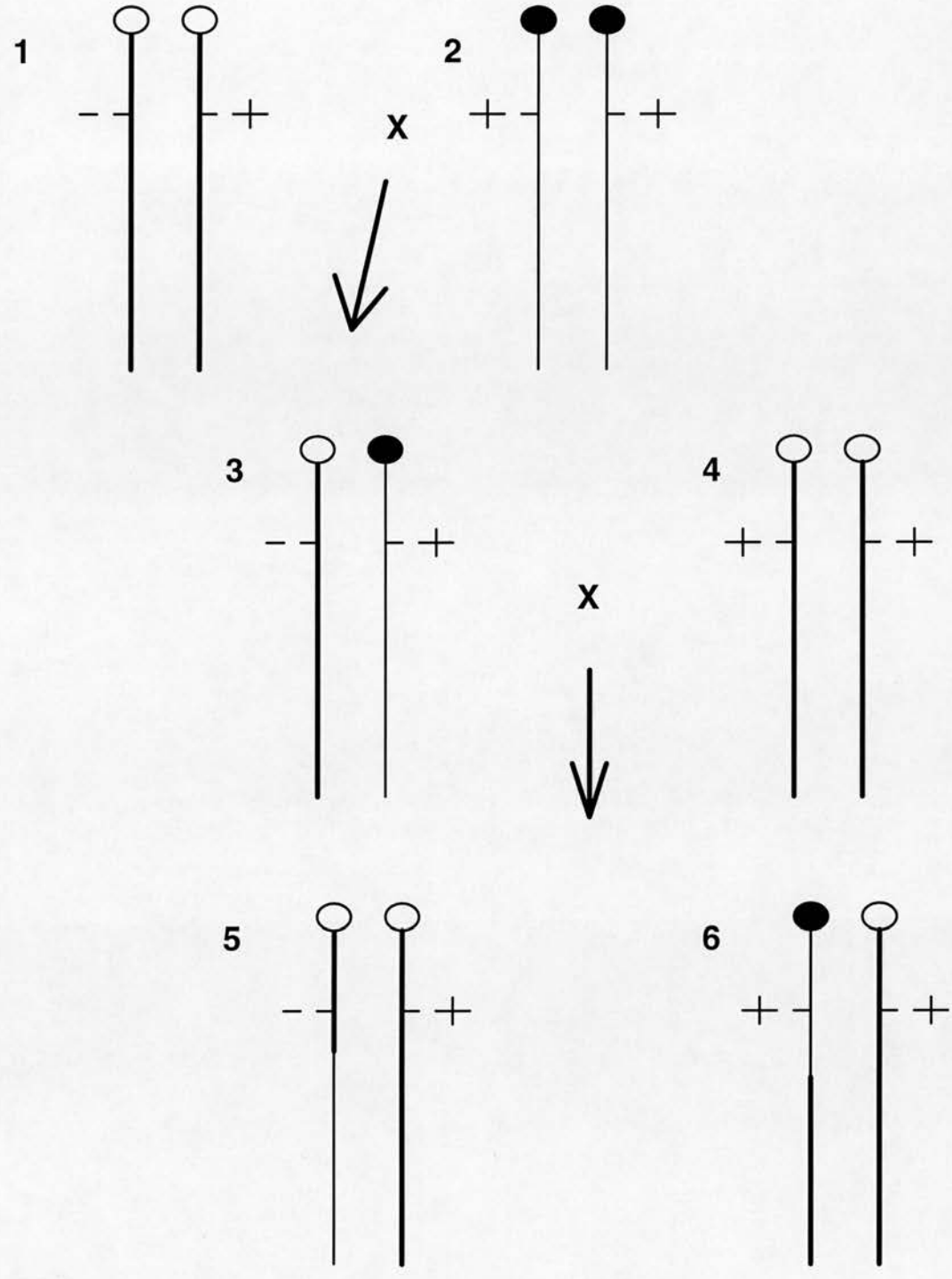

Figure 3.2 Breeding protocol for *Sd* backcross
The first pair of chromosomes, **1**, represent the *Sd* allele (denoted with a minus sign) on a CBA background (clear centromere, bold line). *Sd* mice were crossed with CAST/Ei mice (**2**), to produce the F1 generation (**3**) (the *M.m.castaneus* alleles are depicted by a filled in centromere and fine line). Mice heterozygous for *Sd* in the F1 generation were then backcrossed to CBA mice (**4**), to produce heterozygous *Sd* (**5**) and wild type (**6**) backcross progeny. Allele pairs **5** and **6** are depicted as recombinant chromosomes.

In order to facilitate the search for genes modifying the expressivity of *Sd*, records were made of the tail length and kidney phenotype of a number of the F1 generation. For a summary of this data see table 3.3.2.

Pair No.	Sex	Phenotype	Tail length	Kidney phenotype
Pr4	female	*Sd+*	4.5cm	normal
Pr4	male	WT	9cm	normal
Pr5	female	WT	9.5	normal
Pr5	male	*Sd+*	5.5cm	normal
Pr6	female	WT	9.5cm	normal
Pr6	male	*Sd+*	6cm	small
Pr7	male	*Sd+*	4.5cm	normal
Pr7	female	WT	9.5cm	normal
Pr9	female	*Sd+*	5cm	small
Pr9	male	WT	9.5cm	normal
Pr13	female	*Sd+*	3.3cm	normal
Pr13	male	WT	9.5cm	normal

Table 3.3.2 Phenotypic analysis of F1 breeding pairs

Sd+ indicates heterozygous *Sd* mice, WT indicates wild type mice. Tail lengths are given in centimetres. Small kidneys were those of less than 1cm in length.

3.4 Mapping *Sd* backcross

The first 100 mice of the backcross were typed for the following markers: *D2Mit1, D2Mit31, Pax8, D2Mit79* and *D2Mit7*. The D2Mit markers were typed by PCR using the conditions shown in table 2.10.4. *Pax8* (figure 3.4.2) was typed by amplifying the 3' end of the gene using the PCR primers shown in table 2.10.4, and resolving the PCR product on a HydroLink gel using the running conditions given in section 2.5.5. The haplotyping of these mice is shown in figure 3.4.1. This initial data established *D2Mit31* and *D2Mit7* as markers which flanked *Sd*.

Mice subsequent to the first 100 offspring were initially typed using *D2Mit31* and *D2Mit7*. Mice that were recombinant in this interval were typed for *Pax8*. Those mice which were recombinant between *D2Mit31* and *Pax8* were mapped more finely using the following panel of markers: *D2Mit119, D2Mit149* (vimentin), *D2Mit118, D2Mit267, D2Mit464, D2Mit362, D2Mit178, D2Mit80, D2Mit6, Il1rn* and *D2Mit292* (see table 3.4.1). Once *Pax8* had been excluded as a candidate gene, the last 56 animals (280-336) were initially typed for just *D2Mit31* and *D2Mit7* and the recombinant mice finely mapped with the markers in table 3.4.1. Since it proved difficult to find a variance between CBA and CAST/Ei DNA to use in the mapping of *Pax8*, the majority of animals were already mapped for *D2Mit31* and *D2Mit7* before the 3' variant was found which could be typed using HydroLink analysis. For this reason, once the backcross analysis had excluded *Pax8* as a candidate gene, it was no longer mapped due to the greater effort involved than with the simpler PCR-able D2Mit markers.

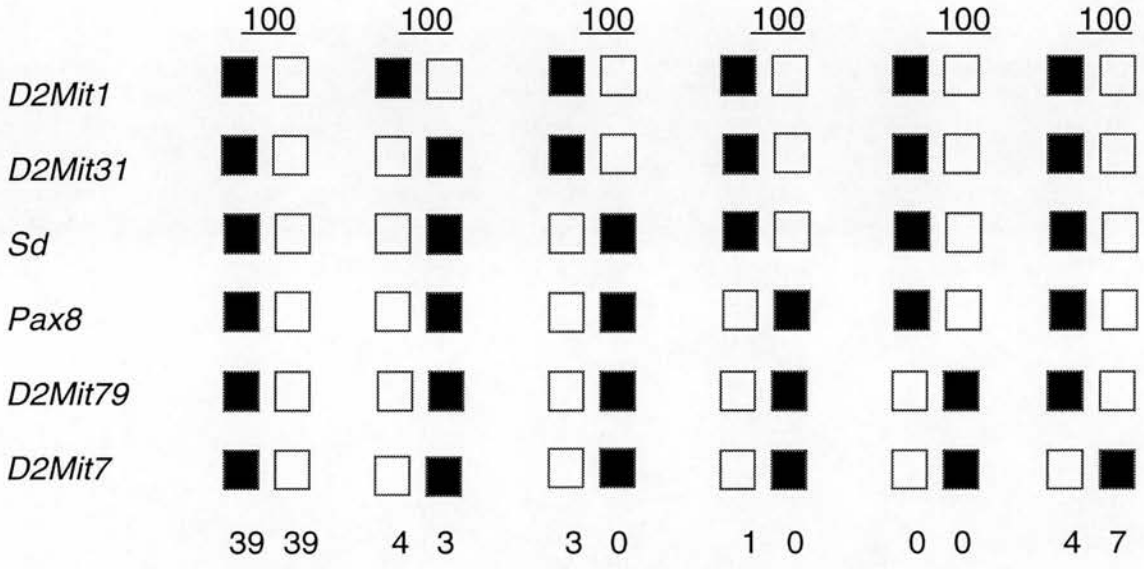

Figure 3.4.1 Haplotype analysis of the first 100 mice of the backcross

The filled in black box (■) denotes the *M.m.castaneus* allele and the unfilled box (❑ denotes the *Sd* allele. The figures at the end of each column indicate the number of mice which were recombinant between the markers shown. The underlined figure at the top of each set of columns indicates the number of backcross animals typed.

Mit Marker	Distance between markers (cM)	CBA-derived allele (bp)	*Sd*-derived allele (bp)	CAST/Ei -derived allele (bp)	Gel used to resolve alleles
D2Mit1	0	120	120	140	4% NuSieve
D2Mit31	7 +/- 2.5	136	136	96	4% NuSieve
*D2Mit119**	0	148	148	142	HydroLink
*D2Mit149**	0.6 +/- 0.4	197	197	209	8% acrylamide
*D2Mit118**	0	123	123	72	4% NuSieve
*D2Mit267**	1.5 +/- 0.7	133	133	119	4% NuSieve
*D2Mit464**	0.6 +/- 0.4	126	140	134	8% acrylamide
*D2Mit362**	0	115	115	147	4% NuSieve
D2Mit178	N/D	125	127	127	6% acrylamide
*D2Mit80**	0.9 +/- 0.5	180	180	186	8% acrylamide
*D2Mit6**	0	124	124	105	8% acrylamide
*D2Mit292**	0.6 +/- 0.4	140	140	122	4% NuSieve
D2Mit79	0	142	142	118	4% NuSieve
D2Mit7	11 +/-3.1	214	214	256	4% NuSieve

Table 3.4.1 D2Mit marker panel used to type *Sd* backcross

D2Mit1 was used as the most proximal marker on chromosome 2 according to its map position from the MIT database. Recombination distances for the remaining markers were calculated from the *Sd* backcross and the cM value given indicates how far each marker maps from the marker directly above it. Recombination distances between *D2Mit1* and *D2Mit31*, *D2Mit292* and *D2Mit79*, and *D2Mit79* and *D2Mit7* were calculated from the first 100 recombinants of the backcross. Remaining recombination distances were calculated from the whole backcross. Standard error and cM values were calculated as in section 2.12. Markers denoted with an asterix (*) were mapped to a panel of mice which were non-recombinant between the markers *D2Mit31* and *D2Mit7.* Two markers, *D2Mit464* and *D2Mit178*, amplified differently sized *Sd* and CBA alleles. It was not possible to map *D2Mit178* onto the backcross.

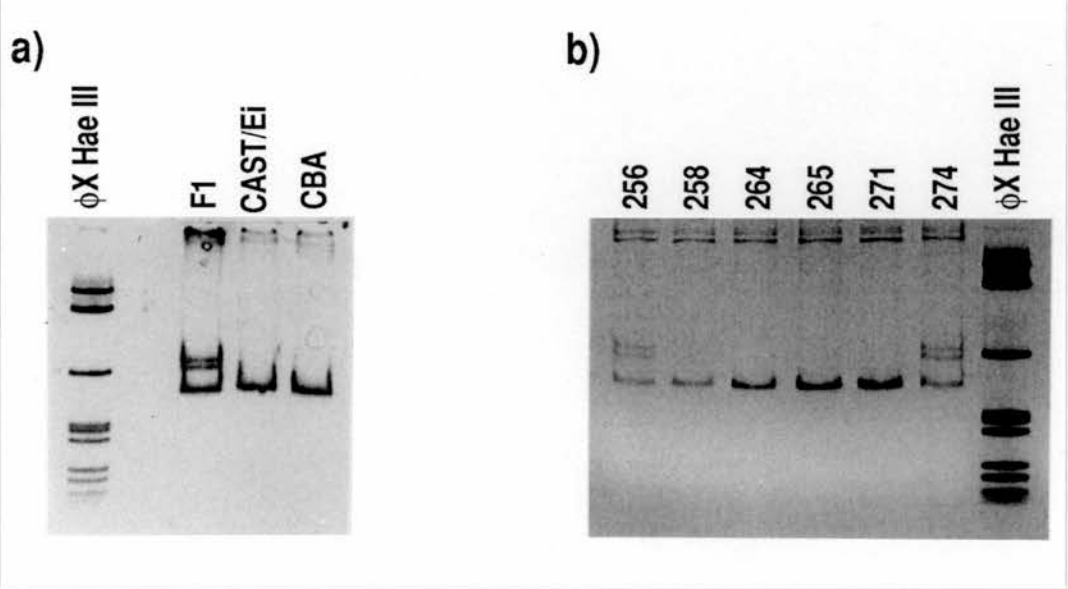

Figure 3.4.2 *Pax8* haplotyping of backcross

Pax8 was amplified by PCR (table 2.10.4) and the product resolved on a HydroLink gel using conditions given in section 2.5.5. Figure a) shows the CBA, CAST/Ei and F1 DNA controls, and figure b) DNA from the backcross. The upper pair of bands show the CBA-CAST/Ei heteroduplex.

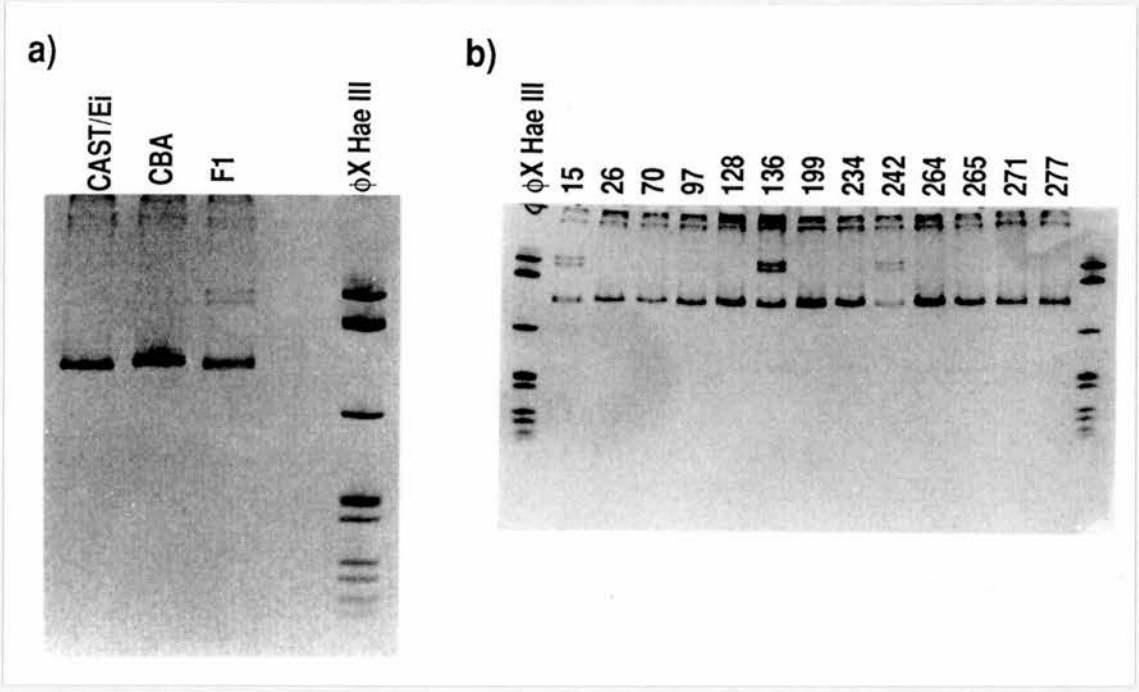

Figure 3.4.3 *Il1rn* haplotyping of backcross

Il1rn was amplified by PCR (table 2.10.4) and the product resolved on a HydroLink gel using conditions given in section 2.5.5. Figure a) shows the CBA, CAST/Ei and F1 DNA controls, and figure b) DNA from the recombinant panel of backcross mice. The upper pair of bands show the CBA-CAST/Ei heteroduplex.

To ensure that the markers used for the haplotyping of the recombinant mice were mapping to the correct region of chromosome 2, those marked with an asterix in table 3.3.1 were mapped to a panel of mice which were non-recombinant between the markers *D2Mit31* and *D2Mit7*. All markers showed concordance with the haplotype derived from the mapping of these two markers. *Il1rn* (figure 3.4.3) was typed in a similar way to *Pax8*, by PCR-amplifying the 3' untranslated end of the gene and resolving the PCR product on a HydroLink gel, using the running conditions given in section 2.5.5. As with *Pax8*, this allowed size-retarded CBA-CAST/Ei heteroduplexes to be identified, which could be used for backcross haplotyping (see section 2.5.5).

Of the markers mapped onto the recombinant panel of mice, only *D2Mit178* could not be haplotyped, since the products produced from the *Sd* and CAST/Ei allele were identical in size (table 3.4.1). *D2Mit464* also amplified differently sized products from the CBA and *Sd* alleles. These markers may therefore define a congenic interval flanking *Sd* which is not derived from the CBA genome, but from the unknown strain on which *Sd* arose.

For a summary of the mapping data see table 3.4.2 and figure 3.4.4. Mapping data for the whole backcross is given in figure 3.4.4. Table 3.4.2 gives the individual haplotypings of the mice in the recombinant panel which were finely mapped with the markers between *D2Mit31* and *D2Mit7*.

From figure 3.4.4 it can be seen that segregation distortion has occurred between two sets of markers, *D2Mit118* and *D2Mit267*, where 5 out of 5 recombinations have been from the CAST/Ei allele to the CBA allele; and between *D2Mit362* and *D2Mit80*, where 3 out of 3 recombinations have been from the CAST/Ei allele to the CBA allele. It is not possible from this number of mice, however, to be certain that the minor level of distortion seen here is a real effect.

No.	*D2Mit 31*	*D2Mit 119*	*D2Mit 149*	*D2Mit 118*	*D2Mit 267*	*D2Mit 464*	*D2Mit 362*	*Sd*	*D2Mit 80*	*D2Mit 6*	*IL1rn*	*D2Mit 292*	*Pax8*	*D2Mit 7*
295	C	N/D	N/D	C	C	C	C	C	C	C	N/D	C	N/D	M
271	C	N/D	N/D	C	C	C	C	C	M	M	M	M	M	M
234	C	C	C	C	C	C	C	C	M	M	M	M	M	M
26	C	C	N/D	C	C	C	C	M	M	M	M	M	M	M
142	C	N/D	C	C	C	M	M	M	M	N/D	N/D	M	M	M
265	C	N/D	N/D	C	M	M	M	M	M	M	M	M	M	M
264	C	N/D	N/D	C	M	M	M	M	M	M	M	M	M	M
319	C	N/D	N/D	C	M	M	M	M	M	M	N/D	M	N/D	M
97	C	C	C	C	M	M	M	M	M	M	M	M	M	M
70	C	C	C	C	M	M	M	M	M	M	M	M	M	M
199	C	C	M	M	N/D	N/D	N/D	M	M	M	M	N/D	M	M
306	M	N/D	N/D	M	M	M	M	M	M	M	N/D	N/D	N/D	C
330	M	N/D	N/D	N/D	M	M	M	M	M	M	N/D	M	N/D	C
318	M	N/D	N/D	M	M	M	M	M	M	M	N/D	M	N/D	C
277	M	N/D	N/D	N/D	M	M	M	M	M	M	M	C	N/D	C
15	M	N/D	N/D	M	M	M	M	M	N/D	M	C	C	C	C
136	M	N/D	M	M	M	C	C	C	C	C	C	C	C	C
242	M	M	C	C	C	C	C	C	C	C	C	C	C	C

Table 3.4.2 Fine mapping of backcross progeny recombinant between *D2Mit31* and *D2Mit7*
Mice denoted by M are mice which are homozygous for the CBA-sized allele, and those denoted by C are heterozygous for the CAST/Ei and CBA-sized alleles at the marker shown. Not Done (N/D) indicates markers which were not typed. Markers were not typed when flanking markers had already established where a recombination had taken place (eg. mouse 199) or where the locus was outwith the region directly flanking *Sd* (eg. mouse 295).

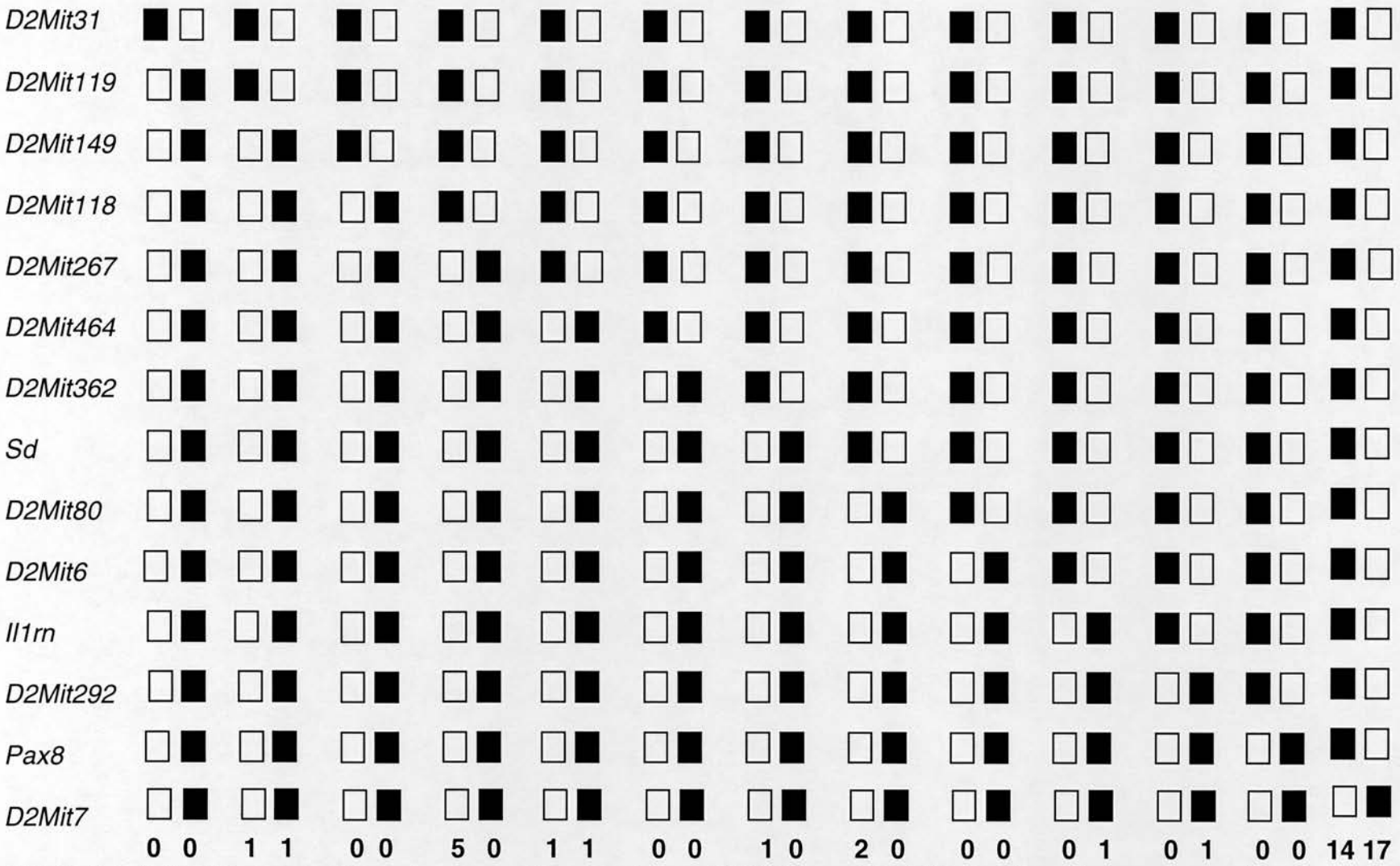

Figure 3.4.4: Haplotypings of the mice recombinant between *D2Mit31* and *D2Mit7* (continued on next page)

Figure 3.4.4 (continued)

Haplotype data for markers *D2Mit31* to *D2Mit6* was derived from the whole backcross (n = 336), by typing a panel of mice recombinant between the markers *D2Mit31* and *D2Mit7*. In the same way, haplotype data for markers *Il1rn, D2Mit292, Pax8* and *D2Mit7* was derived from the first 280 progeny of the backcross.
(■) denotes CAST/Ei allele, (□) denotes the CBA allele.

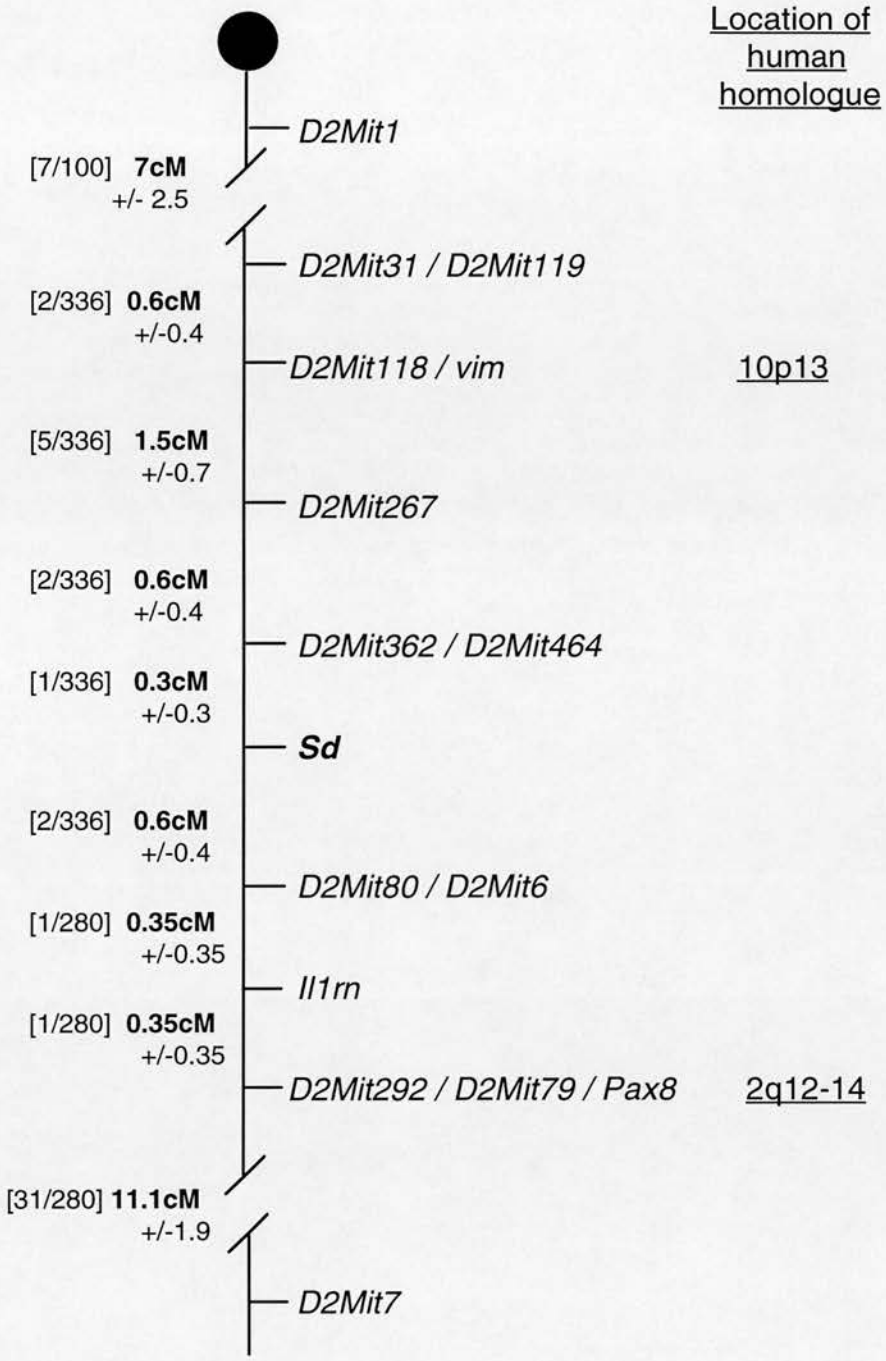

Figure 3.5.1 Genetic map of mouse chromosome 2, in the region of *Sd*
Mapped loci are listed to the right of the chromosome, distance between loci in centiMorgans (cM) to the left. Figures in parentheses show the number of mice recombinant between loci, over the number of mice haplotyped. cM distances and standard errors were calculated as in section 2.12. Regions of conserved synteny were obtained from Siracusa and Abbott, 1994.

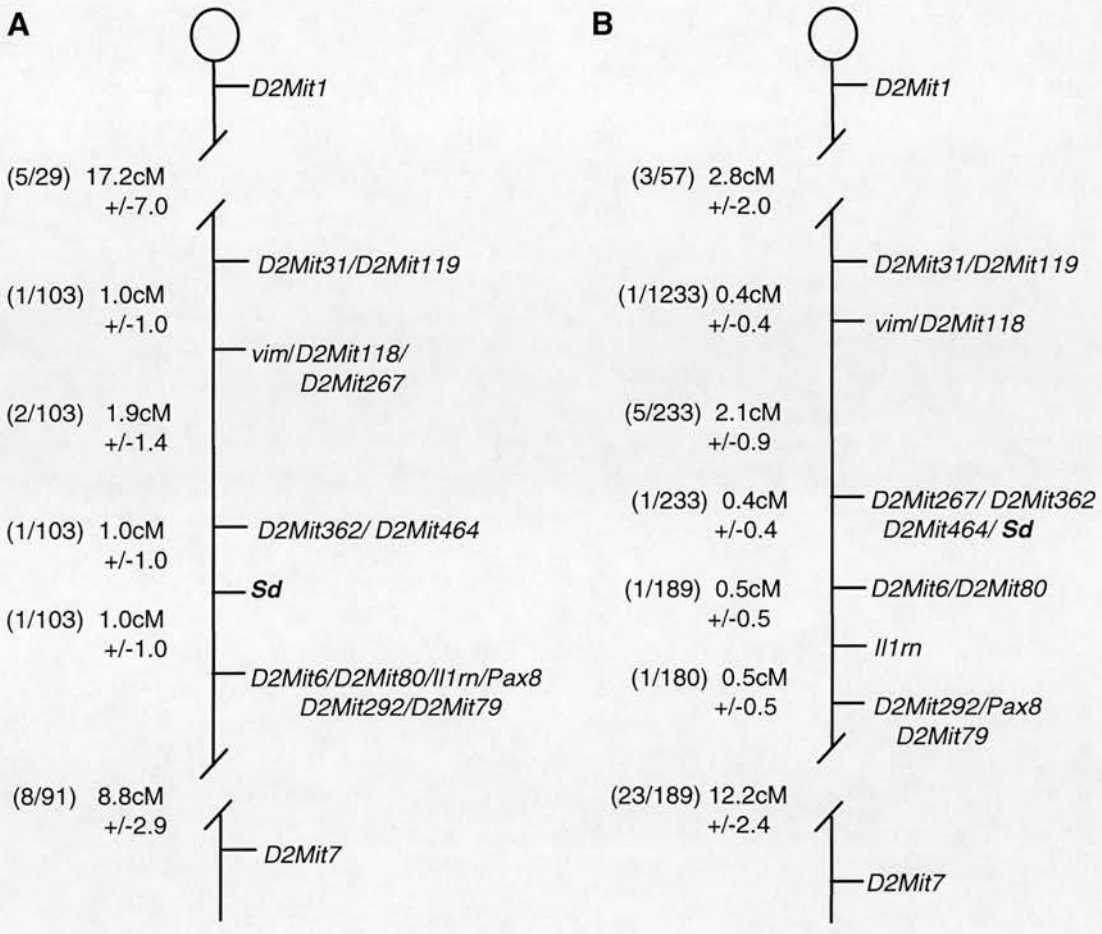

Figure 3.5.2: Genetic maps of mouse chromosome 2 generated from male and female meioses

Map **A** was generated from progeny arising from female F1 mice, and map **B** from the offspring of the male F1 mice. The figures in parentheses indicate the number of mice recombinant in each interval over the total mapped. There were 233 backcross progeny from the male meioses, and 103 from the female meioses. In the first 100 offspring, there were 29 mice from the female, and 71 mice from the male meioses. In the first 280 backcross mice, there were 189 male-meioses and 91 female-meioses offspring. CentiMorgan distances (cM) and standard errors were calculated as in section 2.12.

Genetic maps in the region of *Sd* generated from male and female meioses, are given in figure 3.5.2, whilst an integrated map using mapping data from the whole backcross is shown in figure 3.5.1. Mammalian recombination frequencies are generally higher in females (Roderick and Hillyard, 1989), and this proves to be true for the *Sd* backcross, since the size of the region between *D2Mit1* and *D2Mit7* is greater in the female map (30.9cM), than in the male map (18.9cM). However in map A (figure 3.5.2) a cluster of linked loci can be seen (*D2Mit6, D2Mit80, Pax8, Il1rn, D2Mit79* and *D2Mit292*), between which recombination in the female meioses has not occurred. In the male map B, some of these loci can be ordered due to increased recombination in this region.

3.5 Discussion

At the outset of this project, the gene *Pax8* was a very strong candidate for the Danforth's short tail mutation, since it was known to map to the region and is expressed in both the kidney and the neural tube during development (Plachov *et al.*, 1990). Other members of the *Pax* gene family had previously been found to be responsible for a number of semi-dominant developmental mutations, e.g. *Pax6* was found to be mutated in small eye mice (Hill *et al.*, 1991) and in humans with aniridia (Jordan *et al.*, 1992); similarly, *Pax3* was found to be mutated in the mouse mutation splotch (Epstein *et al.*, 1991), and in humans with Waardenburg's syndrome Type I (Tassabehji *et al.*, 1992)). These findings gave weight to the possibility that *Pax8* could have been responsible for *Sd*. Before it was possible to map *Pax8* onto the *Sd* backcross considerable analysis of the gene was carried out, including sequencing of the gene from cDNA prepared from homozygous and heterozygous mice, and *in situ* analysis of *Pax8* expression in *Sd* embryos. We were unable exclude *Pax8*'s candidacy, however, until the gene was mapped onto the backcross and shown to be outwith the region directly flanking *Sd* (figure 3.5.1), confirming the findings of Koseki *et al.* (1993b).

Using the backcross, we were also able to assess other candidate genes in the region. *D2Mit149* is a marker derived from a repeat within the vimentin gene, which is expressed in both the kidney (Franke *et al.*, 1982) and central nervous system (Cochard and Paulin, 1984) during mouse development, and as such was a potential candidate for *Sd*. Backcross analysis, however, showed that this gene maps 2.7cM proximal to *Sd* and is therefore not responsible for the mutation (figure 3.5.1). Finally, interleukin 1 receptor antagonist, *Il1rn*, has also been excluded by the mapping presented here as a candidate gene for the mutation.

The resulting genetic map of the region around *Sd* (figure 3.5.1) shows that *D2Mit362* and *D2Mit464* are the closest flanking markers proximal to *Sd*, and that *D2Mit6* and *D2Mit80* are the closest flanking markers distal to *Sd*. By using these flanking markers to screen YAC libraries, a positional cloning exercise was initiated in order to isolate the gene responsible for *Sd*.

No.	Sex	Parentage	Phenotype	Tail Length	Tail Type	Kidney No.	Abnorm. kidneys
1	Male	Pr2L1	*Sd+*	1.5cm	E	1	LHS
2	Male	Pr2L1	*Sd+*	0.8cm	F	2	RHS>LHS
3	Female	Pr2L1	*Sd+*	3cm	B	1	LHS
4	Male	Pr3L1	*Sd+*	1cm	C	1	LHS
5	Male	Pr3L1	*Sd+*	1.8cm	D	1	LHS
6	Female	Pr3L1	*Sd+*	2.7cm	D	2 small	RHS>LHS
7	Female	Pr3L1	*Sd+*	3.5cm	E	2	RHS>LHS
8	Female	Pr3L1	*Sd+*	0.8cm	C	2	LHS>RHS
9	Female	Pr2L2	*Sd+*	4.4cm	B.	2 small	
10	Male	Pr2L2	WT	8cm	A	2	
11	Male	Pr2L2	WT	8cm	A	2	
12	Male	Pr2L2	WT	7.5cm	A	2	
13	Female	Pr3L2	WT	7.1cm	A	2	
14	Female	Pr3L2	WT	7.5cm	A	2	
15	Male	Pr3L2	*Sd+*	1.5cm	B	2 small	
16	Male	Pr3L2	*Sd+*	1cm	C	2 small	
17	Male	Pr3L2	*Sd+*	2.5cm	B	2	
18	Male	Pr3L2	*Sd+*	1.3cm	D	2 small	
19	Male	Pr3L2	WT	7cm	A	2	
20	Male	Pr3L2	WT	7cm	A	2	
21	Female	Pr4L1	*Sd+*	0.3cm	C	2 small	
22	Female	Pr4L1	WT	7cm	A	2	
23	Female	Pr4L1	WT	6.5cm	A	2	
24	Female	Pr4L1	*Sd+*	1.3cm	D	2	
25	Male	Pr4L1	WT	6.4cm	A	2	
26	Male	Pr4L1	*Sd+*	2.1cm	B	2	
27	Male	Pr4L1	WT	7cm	A	2	
28	Male	Pr4L1	WT	7cm	A	2	
29	Male	Pr3L3	*Sd+*	2.3cm	F	2	RHS>LHS
30	Male	Pr3L3	*Sd+*	1.7cm	B	2	RHS>LHS*
31	Male	Pr3L3	*Sd+*	1cm	C	2	
32	Male	Pr3L3	WT	7cm	A	2	
33	Female	Pr3L3	*Sd+*	1.5cm	D	2	RHS>LHS
34	Female	Pr3L3	WT	6cm	A	2	
35	Female	Pr3L3	WT	6cm	A	2 small	
36	Male	Pr2L3	*Sd+**	0.5cm	C	2	RHS>LHS
37	Male	Pr2L3	*Sd+*	0.8cm	C	2 small	
38	Male	Pr2L3	WT	6.5cm	A	2	
39	Female	Pr2L3	*Sd+*	2.5cm	E	2	
40	Female	Pr2L3	*Sd+*	3cm	B	2	RHS>LHS*
41	Female	Pr2L3	WT	6cm	A	2	
42	Female	Pr2L3	WT	7.1cm	A	2	

Table 3.3.1 Phenotypic analysis of backcross progeny

No.	Sex	Paren-tage	Pheno-type	Tail Length	Tail Type	Kidney No.	Abnorm. kidneys
43	Female	Pr2L3	WT	7.1cm	A	2	
44	Female	Pr3L4	*Sd+*	2cm	B	2	RHS> LHS
45	Female	Pr3L4	*Sd+*	1.7cm	B	2	
46	Female	Pr3L4	WT	6.5cm	A	2	
47	Male	Pr3L4	*Sd+*	2.3cm	B	2 small	
48	Male	Pr3L4	*Sd+*	1.2cm	E	2	RHS> LHS
49	Male	Pr3L4	WT	6.5cm	A	2	
50	Male	Pr7L1	*Sd+*	2.4cm	B	2	
51	Male	Pr7L1	WT	7.5cm	A	2	
52	Female	Pr7L1	WT	7cm	A	2	
53	Male	Pr7L1	WT	7cm	A	2	
54	Male	Pr5L1	WT	7.2cm	A	2	
55	Male	Pr5L1	*Sd+*	2.7cm	E	2	RHS> LHS
56	Male	Pr5L1	*Sd+*	1cm	D	2	RHS> LHS
57	Male	Pr5L1	*Sd+*	1.6cm	D	2	RHS> LHS
58	Male	Pr5L1	WT	7.5cm	A	2	
59	Male	Pr5L1	WT	7cm	A	2	
60	Male	Pr5L1	WT	6.5cm	A	2	
61	Female	Pr5L1	*Sd+*	0.5cm	C	2	RHS>LHS
62	Male	Pr2L4	*Sd+*	1.5cm	B	2	RHS> LHS
63	Male	Pr2L4	WT	6.5cm	A	2	
64	Male	Pr2L4	WT	6.7cm	A	2	
65	Male	Pr2L4	WT	7.1cm	A	2	
66	Male	Pr2L4	WT	6.5cm	A	2	
67	Female	Pr2L4	WT	6.7cm	A	2	
68	Female	Pr2L4	*Sd+*	1cm	C	2 small	
69	Female	Pr6L1	WT	6.6cm	A	2	
70	Female	Pr6L1	*Sd+*	1.7cm	B	2	RHS>LHS
71	Female	Pr6L1	WT	7.5cm	A	2	
72	Female	Pr6L1	WT	7cm	A	2	
73	Male	Pr6L1	WT	7cm	A	2	
74	Male	Pr9L1	*Sd+*	1.5cm	E	2 small	LHS*.
75	Male	Pr9L1	*Sd+*	1.5cm	B	2 small	RHS>LHS*
76	Male	Pr9L1	*Sd+*	3cm	B	2	
77	Male	Pr9L1	WT	6.2cm	A	2	
78	Male	Pr9L1	WT	6cm	A	2	
79	Female	Pr9L1	*Sd+*	2cm	E	2 small	RHS> LHS*
80	Female	Pr9L1	WT	6.5cm	A	2	
81	Male	Pr8L1	*Sd+*	0.9cm	C	2	LHS>RHS
82	Male	Pr8L1	*Sd+*	0.4cm	C	2 small	
83	Male	Pr8L1	WT	7.5cm	A	2	
84	Male	Pr8L1	WT	7cm	A	2	

Table 3.3.1 Phenotypic analysis of backcross progeny (continued)

No.	Sex	Paren-tage	Pheno-type	Tail Length	Tail Type	Kidney No.	Abnorm. kidneys
85	Female	Pr8L1	*Sd+*	2.8cm	E	2	
86	Female	Pr8L1	*Sd+*	1cm	D	2	
87	Female	Pr8L1	*Sd+*	2.5cm	E	2 small	RHS>LHS
88	Female	Pr8L1	WT	7.1cm	A	2	
89	Female	Pr8L1	WT	6.7cm	A	2	
90	Male	Pr4L2	*Sd+*	1.5cm	B	2 small	
91	Male	Pr4L2	*Sd+*	1.6cm	E	2	RHS> LHS
92	Male	Pr4L2	WT	7cm	A	2	
93	Male	Pr4L2	WT	7cm	A	2	
94	Female	Pr4L2	*Sd+*	1cm	C	2 small	RHS>LHS
95	Female	Pr4L2	*Sd+*	3cm	E	2 small	
96	Female	Pr4L2	*Sd+*	1.2cm	D	2	RHS>LHS*
97	Female	Pr5L2	*Sd+*	2cm	E	2	RHS>LHS
98	Male	Pr5L2	*Sd+*	2cm	E	2	
99	Female	Pr5L2	*Sd+*	2.3cm	E	2 small	RHS>LHS
100	Female	Pr5L2	WT	6.4cm	A	2	
101	Female	Pr5L2	WT	6.5cm	A	2	
102	Male	Pr5L2	*Sd+*	2.7cm	B	2	
103	Male	Pr5L2	*Sd+*	3.5cm	B	2	
104	Male	Pr5L2	*Sd+*	1.4cm	E	2	
105	Male	Pr5L2	WT	7cm	A	2	
106	Male	Pr7L2	WT	7.2cm	A		
107	Female	Pr7L2	*Sd+*	No tail		2	RHS>LHS
108	Female	Pr6L2	*Sd+*	2cm	B	2 small	RHS>LHS
109	Male	Pr6L2	WT	6.2cm	A	2	
110	Male	Pr6L2	WT	6cm	A	2	
111	Male	Pr2L5	*Sd+*	1.5cm	B	2	
112	Male	Pr2L5	WT	6.8cm	A		
113	Female	Pr2L5	*Sd+*	1.5cm	C	2 small	
114	Female	Pr2L5	*Sd+*	1.7cm	B	2 small	
115	Female	Pr2L5	*Sd+*	0.7cm	C	2	RHS>LHS
116	Female	Pr2L5	WT	6.5cm	A	2	
117	Female	Pr8L2	*Sd+*	2.5cm	E	2 small	
118	Female	Pr8L2	WT	7.7cm	A	2	
119	Female	Pr8L2	*Sd+*	1.2cm	C	2 small	
120	Male	Pr8L2	WT	7.4cm	A	2	
121	Male	Pr8L2	*Sd+**	0.1cm	C	2	RHS>LHS
122	Female	Pr9L2	WT	7.5cm	A	2	
123	Female	Pr9L2	WT	7.3cm	A		
124	Female	Pr9L2	WT	7cm	A		

Table 3.3.1 Phenotypic analysis of backcross progeny (continued)

No.	Sex	Paren-tage	Pheno-type	Tail Length	Tail Type	Kidney No.	Abnorm. kidneys
125	Female	Pr9L2	*Sd+*	1.5cm	B	2 small	RHS>LHS
126	Male	Pr9L2	WT	7cm	A	2	
127	Male	Pr9L2	*Sd+*	2.5cm	B	1	RHS only*
128	Male	Pr9L2	*Sd+*	1cm	C	2 small	
129	Male	Pr7L3	WT	7.1cm	A	2	
130	Male	Pr7L3	*Sd+*	3cm	B	2	RHS>LHS
131	Male	Pr7L3	*Sd+*	2cm	B	2	
132	Female	Pr7L3	WT	7.5cm	A	2	
133	Female	Pr7L3	WT	7.5cm	A	2	
134	Female	Pr7L3	WT	7cm	A	2	
135	Female	Pr7L3	*Sd+*	3.5cm	E	2	
136	Female	Pr4L3	WT	7.7cm	A	2	
137	Male	Pr4L3	WT	8cm	A	2	
138	Male	Pr4L3	WT	7cm	A	2	
139	Male	Pr4L3	WT	7.5cm	A	2	
140	Male	Pr4L3	WT	7.7cm	A	2	
141	Male	Pr4L3	*Sd+*	2.5cm	B	2 small	
142	Male	Pr4L3	*Sd+*	3cm	E	2	RHS>LHS
143	Female	Pr3L5	WT	8cm	A	2	RHS>LHS
144	Female	Pr3L5	WT	7.9cm	A	2	
145	Female	Pr3L5	*Sd+*	3.5cm	B	2	RHS>LHS
146	Male	Pr6L3	WT	6.5cm	A	2	
147	Male	Pr6L3	*Sd+*	2.5cm	D	2 small	RHS>LHS
148	Female	Pr6L3	WT	6.7cm	A	2	
149	Female	Pr6L3	*Sd+*	2.1cm	D	2	RHS>LHS*
150	Male	Pr5L3	*Sd+*	2.5cm	F	2	
151	Male	Pr5L3	*Sd+*	2.5cm	D	2	RHS>LHS*
152	Male	Pr5L3	*Sd+*	2cm	C	2	RHS>LHS*
153	Female	Pr5L3	WT	7cm	A	2	
154	Female	Pr5L3	WT	6.8cm	A	2	
155	Female	Pr5L3	*Sd+*	4.5cm	F	2	RHS>LHS
156	Female	Pr5L3	*Sd+*	1.5cm	C	2	RHS>LHS
157	Female	Pr7L4	*Sd+*	2.3cm	B	2	RHS>LHS
158	Female	Pr7L4	*Sd+*	0.5cm	C	2	RHS>LHS
159	Male	Pr7L4	*Sd+*	1cm	C	2 small	
160	Male	Pr7L4	WT	6.5cm	A	2	
161	Male	Pr7L4	*Sd+*	3cm	B	2 small	RHS>LHS
162	Male	Pr7L4	*Sd+*	1.2cm	C	1.smal	LHS *
163	Male	Pr7L4	*Sd+*	2.5cm	B	2	
164	Female	Pr9L3	*Sd+*	2.5cm	E	2	RHS>LHS
165	Male	Pr9L3	WT	6.5cm	A	2	
166	Male	Pr9L3	WT	6.8cm	A	2	

Table 3.3.1 Phenotypic analysis of backcross progeny (continued)

No.	Sex	Paren-tage	Pheno-type	Tail Length	Tail Type	Kidney No.	Abnorm. kidneys
167	Male	Pr9L3	WT	7cm	A	2	
168	Male	Pr9L3	*Sd+*	2.5cm	B	2	RHS>LHS
169	Male	Pr9L3	*Sd+*	2cm	B	2 small	
188	Female	Pr5L4	*Sd+*	3.5cm	B	2	
189	Female	Pr5L4	*Sd+*	3.5cm	E	2	RHS>LHS
190	Male	Pr5L4	WT	7.5cm	A	2	
191	Male	Pr5L4	WT	8cm	A	2	
192	Male	Pr5L4	*Sd+*	1.5cm	B	2	RHS >LHS
193	Male	Pr3L7	*Sd+*	3.5cm	B	2	RHS>LHS
194	Male	Pr3L7	*Sd+*	3cm	B	2	
195	Male	Pr3L7	*Sd+*	2.3cm	E	2	
196	Male	Pr3L7	*Sd+*	2cm	B	2	RHS>LHS
197	Female	Pr3L7	WT	7cm	A	2	
198	Female	Pr3L7	WT	7cm	A	2	
199	Female	Pr3L7	*Sd+*	1.5cm	C	2 small	
200	Male	Pr9L4	WT	7.5cm	A	2	
201	Male	Pr9L4	WT	7.5cm	A	2	
202	Male	Pr9L4	*Sd+*	1cm	B	2 small	
203	Female	Pr9L4	*Sd+*	0.9cm	B	2 small	
204	Female	Pr9L4	*Sd+*	0.2cm	C	2 small	RHS>LHS
205	Female	Pr7L5	*Sd+*	4.5cm	F	2	RHS>LHS
206	Male	Pr7L5	WT	8.5cm	A	2	
207	Male	Pr7L5	WT	8.5cm	A	2	
208	Male	Pr7L5	WT	8.5cm	A	2	
209	Male	Pr7L5	*Sd+*	3cm	D	2	
210	Male	Pr7L5	*Sd+*	3cm	D	2	
211	Male	Pr7L5	*Sd+*	3cm	E	2	LHS>RHS
212	Male	Pr6L5	WT	8cm	A	2	
213	Male	Pr6L5	WT	8cm	A	2	
214	Male	Pr6L5	*Sd+*	4cm	E	2	
215	Female	Pr6L5	WT	7.5cm	A	2	
216	Female	Pr6L5	WT	8cm	A	2	
217	Female	Pr6L5	WT	7.5cm	A	2	
218	Female	Pr6L5	*Sd+*	2.5cm	E	2	RHS>LHS
219	Female	Pr6L5	*Sd+*	1.5cm	B	2 small	
220	Female	Pr6L5	*Sd+*	no tail		2	RHS>LHS
221	Male	Pr13L1	WT	7cm	A	2	
222	Male	Pr13L1	*Sd+*	2cm	F	2	
223	Female	Pr13L1	WT	6.5cm	A	2	
224	Female	Pr13L1	*Sd+*	3cm	B	2 small	
225	Female	Pr13L1	*Sd+*	1cm	E	2	

Table 3.3.1 Phenotypic analysis of backcross progeny (continued)

No.	Sex	Paren-tage	Pheno-type	Tail Length	Tail Type	Kidney No.	Abnorm. kidneys
226	Female	Pr9L5	WT	6cm	A	2	
227	Female	Pr9L5	WT	7.5cm	A	2	
228	Female	Pr9L5	WT	7cm	A	2 small	
229	Female	Pr9L5	WT	7.5cm	A	2	
230	Female	Pr9L5	*Sd+*	1.5cm	B	2	RHS>LHS
231	Male	Pr9L5	WT	7cm	A	2	LHS>RHS
232	Female	Pr6L6	WT	6.5cm	A	2	
233	Female	Pr6L6	WT	6.5cm	A	2	RHS>LHS
234	Female	Pr6L6	WT	7m	A	2	
235	Male	Pr6L6	WT	6.5cm	A	2	
236	Male	Pr6L6	*Sd+*	2cm	B	2 small	RHS>LHS
237	Male	Pr6L6	*Sd+*	1cm	D.	2	
238	Female	Pr4L5	WT	7.5cm	A	2	
239	Female	Pr4L5	*Sd+*	2cm	E	2 small	
240	Female	Pr4L5	*Sd+*	3cm	B	2 small	RHS>LHS
241	Male	Pr4L5	WT	7.5cm	A	2	
242	Male	Pr4L5	WT	7.5cm	A	2	
243	Male	Pr4L5	*Sd+*	1.5cm	E.	2	
244	Male	Pr4L5	*Sd+*	0.4cm	C	2	LHS*
245	Male	Pr4L5	*Sd+*	0.4cm	C	1	LHS only
246	Male	Pr13L2	*Sd+*	1.5cm	E	2	
247	Female	Pr13L2	WT	6.5cm	A	2	
248	Female	Pr13L2	WT	5.5cm	A	2	
249	Female	Pr13L2	WT	5.5cm	A	2	
250	Female	Pr13L2	*Sd+*	1.5cm	B	2	RHS>LHS
251	Female	Pr7L6	*Sd+*	1.3cm	D	2	
252	Male	Pr7L6	WT	7cm	A	2	
253	Male	Pr7L6	*Sd+*	1.8cm	B	2 small	LHS*.
254	Female	Pr5L5	WT	7cm	A	2	RHS>LHS
255	Female	Pr5L5	WT	7cm	A	2	
256	Male	Pr5L5	WT	7.3cm	A	2	
257	Male	Pr5L5	*Sd+*	2cm	E.	2	
258	Male	Pr6L7	*Sd+*	0.9cm	B	2 small	
259	Female	Pr6L7	WT	6.5cm	A	2	
260	Female	Pr6L7	WT	7cm	A	2	
261	Female	Pr6L7	WT	5.5cm	G	2 small	
262	Female	Pr6L7	*Sd+*	2.5cm	E.	2 small	
263	Female	Pr6L7	*Sd+*	3cm	F.	2	
264	Female	Pr6L7	*Sd+*	2cm	E.	2	
265	Female	Pr6L7	*Sd+*	2cm	B	2 small	
266	Female	Pr9L6	WT	7.2cm	A	2	

Table 3.3.1 Phenotypic analysis of backcross progeny (continued)

No.	Sex	Paren-tage	Pheno-type	Tail Length	Tail Type	Kidney No	Abnorm. kidneys
267	Female	Pr9L6	WT	7cm	A	2	LHS>RHS
268	Female	Pr9L6	WT	7.5cm	A	2	
269	Female	Pr4L6	WT	7cm	A	2	
270	Female	Pr4L6	*Sd+*	0.5cm	E	2.small	
271	Male	Pr4L6	WT	6.5cm	A	2	
272	Male	Pr4L6	WT	8cm	A	2	
273	Male	Pr4L6	WT	7.5cm	A	2	
274	Male	Pr4L6	WT	7.5cm	A	2	
275	Male	Pr4L6	*Sd+*	1cm	B	2	
276	Male	Pr4L6	*Sd+*	0.2cm	C	2	
281	Female	Pr5L6	WT	7cm	A	2	
282	Female	Pr5L6	WT	7cm	A	2	
283	Female	Pr5L6	*Sd+*	1.5cm	E	2	RHS>LHS
284	Male	Pr5L6	*Sd+*	1.2cm	C	1	RHS, big
285	Male	Pr6L8	WT	6.5cm	A	2	
286	Male	Pr6L8	WT	7.2cm	A	2	
287	Male	Pr6L8	*Sd+*	2.5cm	B	2	RHS>LHS
288	Female	Pr6L8	WT	7cm	A	2	RHS>LHS
289	Male	Pr13L4	WT	7cm	A	2	
290	Female	Pr13L4	WT	6.5cm	A	2	
291	Female	Pr13L4	WT	6.3cm	A	2	
292	Female	Pr13L4	*Sd+*	3cm	E	2	RHS>LHS
293	Male	Pr4L7	WT	6.7cm	A	2	
294	Male	Pr4L7	WT	6.5cm	A	2	
295	Female	Pr4L7	WT	6.7cm	A	2	
296	Female	Pr4L7	WT	7cm	A	2	
297	Female	Pr4L7	*Sd+*	3cm	E	2 small	RHS>LHS
298	Female	Pr4L7	*Sd+*	3.5cm	D	2 small	
299	Female	Pr4L7	*Sd+*	1cm	F	2	RHS>LHS*
300	Female	Pr6L9	WT	5cm	G	2	
301	Female	Pr6L9	*Sd+*	2.5cm	D	2	RHS>LHS
302	Male	Pr6L9	*Sd+*	0.5cm	C	2	RHS>LHS
303	Female	Pr4L8	WT	6cm	A	2	
304	Female	Pr4L8	WT	6cm	A	2	RHS>LHS
305	Male	Pr4L8	WT	6.3cm	A	2	
306	Female	Pr4L8	*Sd+*	1.5cm	B	2	RHS>LHS
307	Female	Pr4L8	*Sd+*	1.8cm	B	2 small	RHS>LHS
312	Male	Pr13L6	WT	6.5cm	A	2	
313	Male	Pr13L6	*Sd+*	0.6cm	E	2 small	RHS>LHS
314	Female	Pr13L6	WT	7cm	A	2	
315	Female	Pr13L6	*Sd+*	1.1cm	D	2 small	
316	Female	Pr13L7	WT	6.5cm	A	2	

Table 3.3.1 Phenotypic analysis of backcross progeny (continued)

No.	Sex	Paren-tage	Pheno-type	Tail Length	Tail Type	Kidney No.	Abnorm. kidneys
317	Female	Pr13L7	*Sd+*	1.5cm	E	2	
318	Male	Pr13L7	*Sd+*	0.8cm	E	2	
319	Male	Pr13L7	*Sd+*	1.2cm	D.	2	RHS>LHS
320	Male	Pr13L8	*Sd+*	0.5cm	C	2	
321	Female	Pr13L8	*Sd+*	2cm	B	2 small	
322	Female	Pr13L8	*Sd+*	2.8cm	D	2	
323	Female	Pr13L9	*Sd+*	3.2cm	F	2 small	
324	Male	Pr13L9	WT	7cm	A	2	
325	Male	Pr13L9	WT	6.5CM	A	2	
326	Male	Pr13L9	WT	7cm	A	2	
327	Male	Pr13L9	WT	6.5cm	A	2	
328	Male	Pr13L9	*Sd+*	2.5cm	E	2	RHS>LHS
329	Female	Pr13L9	*Sd+*	0.5cm	C	2	RHS*>LHS
330	Female	Pr13L10	*Sd+*	1cm	C	2 small	
331	Male	Pr13L10	WT	5.8cm	A	2	
332	Male	Pr13L10	WT	6.2cm	A	2	
333	Male	Pr13L10	WT	6.1cm	A	2	
334	Male	Pr13L10	*Sd+*	0.4cm	F	2 small	
335	Male	Pr13L10	*Sd+*	1.2cm	F	2	RHS>LHS
336	Male	Pr13L10	*Sd+*	1.4cm	E	2 small	

Table 3.3.1 Phenotypic analysis of backcross progeny (continued)

Parentage of each mouse was recorded from the breeding pair (Pr) and the number of the litter (L) from which the mouse came. Tail length is given in centimetres (cm).

The type of tail each mouse had was classed into one of 7 classes (**A** - **G**). **A** represents a normal tail; **B**, a tail which is normal up to a point of constriction, after which no bone is present but filament and skin are. **C** represents tails which have no bone, only skin and filament. These were usually very small. **D** represents tails which are similar to B but which have kinked filament at the end of them, and **E** represents tails in which the bone itself was crooked. These tails often had constrictions in them, followed by kinked filment and skin. **F** represents tails which were straight and appeared normal except for their length and **G** represents tails from wild type mice which appeared smaller than average.

The length of each kidney was measured (data not shown) and any differences in size or appearance noted. When both kidneys were smaller than average, 'small' has been typed next to the number of kidneys present. The final column details abnormalities in the size or appearance of the kidneys. An asterix (*) denotes kidneys which were abnormal in appearance or very much smaller in size than the kidney from the other side of the body. When the kidneys from an animal were not of equal size, the symbol > was used to show from what side of the body the larger kidney arose from.

Chapter 4: Positional cloning - assembly of a YAC contig in the region of *Sd*

4.1 Introduction

Through the identification of disease causing genes, progress has been made in the dissection of many mammalian developmental and biological pathways. In humans, disease genes have generally been identified using two strategies: functional cloning and positional cloning.

4.1.1 Cloning strategies

Functional cloning relies on pre-existing information regarding the basic biological defect of a disease being available. Strategies such as purification of a protein using immunoprecipitation, followed by amino acid sequence determination, allow a gene to be identified on the basis of prior knowledge about its product. As the name suggests, genes can also be cloned when information regarding gene function is available. In this way, the genes for a number of transmembrane receptor proteins (Lin *et al.*, 1992; Matthews and Vale, 1991) and metabolic disorders, such as severe combined immunodeficiency (Valerio *et al.*, 1985), have successfully been cloned. However, since there are a large number of human disorders for which no functional or biochemical information exists, the number of genes which can be cloned using this strategy is considerably limited.

Positional cloning is a more recent strategy and consequently fewer genes have been cloned using this approach. Whereas functional cloning depends on application of biological information, positional cloning is initiated when a disease locus is mapped to a chromosomal location. Such a strategy was originally called "reverse genetics" but was renamed "positional cloning" (Collins, 1992) to indicate that a gene could be isolated on the basis of its genetic and physical location in the genome, without prior knowledge of its biological function. This strategy is frequently more time consuming than a candidate gene (see chapter 3) or a functional cloning approach, since a chromosomal walk from flanking

genetic markers is required, followed by analysis of transcripts in a defined interval. Nevertheless, a number of human disease genes have successfully been positionally cloned, such as the gene responsible for cystic fibrosis (Riordan *et al.*, 1989; Rommens *et al.*, 1989).

In the study of human disorders, positional cloning begins with a collection of pedigrees in which a disease gene is segregating. These pedigrees are analysed with polymorphic markers until evidence of linkage to a marker is established. Additional fine mapping can then be employed to narrow down the region, since by decreasing the size of the interval and the number of candidate transcripts requiring analysis, the effort required to find the correct one can be reduced. However, the number of informative meioses in available pedigrees limits the extent to which a region of linkage can be finely mapped. Because of this, fine mapping is often limited to candidate intervals of around 1cM, approximately 1 million basepairs (Collins, 1992). Depending on the gene density of a chromosomal region, this could result in a long and time consuming search through many transcripts before identifying the correct one. One of the advantages of positionally cloning genes in the mouse is that the problem of a limited number of informative meioses is overcome by producing crosses with large numbers of progeny. This increases the number of meioses for fine mapping and reduces the effort required to find the correct gene, since a smaller physical interval can be defined.

An alternative means of cloning new genes in mouse and man is to employ both strategies in a single study. This combined approach has been termed the "positional candidate strategy" (Ballabio, 1993). This entails defining a candidate interval through linkage analysis and then screening computer databases for genes or cDNA's which have already been isolated, and assessing them as candidates on the basis of their mapped location and biological function. This differs from a straight forward candidate gene

approach in that weaker candidate genes come into consideration as the mapping of an interval becomes finer.

An alternative means of accessing novel genes is to clone disease genes in the mouse. A recent comparative mapping study has shown that of 2616 loci mapped in the mouse, 917 have mapped human homologues (Copeland *et al.*, 1993). These loci cover 101 segments of conserved linkage homology. Because of this, it is possible to use genes identified from mouse mutations as candidate genes for human disorders which have a similar aetiology; for example, the cloning of the mouse microphthalmia (*mi*) gene (Hodgkinson *et al.*, 1993) led to the discovery of the human homologue, MITF (Tachibana *et al.*, 1994) which was found to be responsible for Waardenburg syndrome type II (Tassabehji *et al.*, 1994). Similarly, identification of a type VII myosin gene responsible for the mouse mutation *shaker-1* (Gibson *et al.*, 1995), led to the discovery that mutations in the human homologue were responsible for Usher syndrome type 1B (Weil *et al.*, 1995). Mapping genes in the mouse can also facilitate the cloning of human genetic disease loci in the absence of appropriate mouse models, since it is often easier to map a new gene in the mouse and then predict its location in humans, than it is to map the gene directly in humans.

4.1.2 Yeast artificial chromosomes

Yeast artificial chromosomes (YACs) have been an essential tool in the positional cloning approach, as they have considerably facilitated the generation of physical maps from genetic linkage data. YACs contain the basic functional units of yeast chromosomes: centromeres, telomeres and autonomous replication sequences (Burke *et al.*, 1987). When these elements are combined within a cloning vector, they allow large linear DNA molecules (from 50kb to 2Mb) to be transformed into yeast and maintained there as an artificial chromosome. Traditional cloning systems, such as plasmid and cosmid vectors, are well suited to the analysis of tightly packed

genomes, but are unsuitable for the larger and more complex mammalian genome, since their cloning capacity is small. The development of YACs has therefore permitted large contiguous genomic regions to be clonally propagated in a single vector, allowing the analysis of genes to take place within their immediate genomic context.

There are many advantages of using such a high capacity cloning system. Using genetically linked flanking markers, YACs can be isolated from a number of widely available mouse or human YAC libraries (see section 2.3.5). YAC end fragments can be generated in a number of ways, for example by vectorette (Riley *et al.*, 1990) or inverse PCR (Arveiler and Porteous, 1991) (see section 2.10.5), for use as new markers in backcross or pedigree analysis and chromosomal walking. The large insert capacity of YAC vectors means that complete genetic units can be analysed in a single vector. This becomes important when the search for a mutation within the coding regions of a gene is unsuccessful. Promoter and enhancer elements can then be analysed for mutations within the same, or an overlapping, clone without extensive screening required to find a new set of clones to cover potentially large regions. A powerful means of employing YACs in the search for murine developmental genes is to reintroduce YACs into the germline. In this way it is possible to assay for complementation of a mutant phenotype (Schedl *et al.*, 1993). However, such an analysis is only successful in loss of function mutations, either a recessive mutation or a dominant haploinsufficient mutation. Genes for gain of function mutations cannot be identified using this technique, unless the YAC library itself has been made from mutant DNA.

The use of YACs is not without its drawbacks. Despite the generation of YAC libraries in recombination deficient strains of yeast (Chartier *et al.*, 1992), YAC clones are still prone to instability and internal deletions. High rates of chimerism are common and YACs require careful analysis to ensure that the

insert is of a single origin. A considerable problem is the time consuming task of isolating all the transcripts within a large genomic region, in order to identify the correct one. This has been mitigated to some extent by the development of new techniques, such as novel exon trap vectors (Nehls *et al.*, 1994), however, the effort required to proceed from YAC analysis to the isolation of a candidate transcript is still substantial.

4.2. YAC analysis

The markers *D2Mit267*, *D2Mit362* and *D2Mit80* were sent to the Genethon YAC screening service. Of the YACs that were returned, only those isolated with *D2Mit362* have been analysed in detail.

4.2.1 PCR analysis of *D2Mit362* YACs

15 YACs isolated with *D2Mit362* were returned from the Genethon YAC screen. From the 15 YACs analysed, 8 were positive for *D2Mit362* by PCR (section 2.10.5). Positive single colonies identified by PCR were used to prepare single colony streaks for the preparation of DNA (section 2.3.3) and agarose plugs (section 2.3.4). Subsequently it was not possible to identify positive colonies in single colony streaks for YACs 362Lh/1 and 362Ln/2, and they were considered to be too unstable for further analysis. For a summary of data from the *D2Mit362* isolated YACs, see table 4.2.1

A PCR-screen for the presence of the two markers which flank the *Sd* region, *D2Mit80* and *D2Mit267*, on the 8 YACs isolated with *D2Mit362*, showed that these markers were not present in these YACs.

Name	Library	Positive for *D2Mit362*	Size (kb)
362Ln/1	Research Genetics	√	590
362Ln/2	Research Genetics	X	N/D
362Lh/1	ICRF/ Lehrach	X	N/D
362Lh/2	ICRF/ Lehrach	√	620
362Lh/3	ICRF/ Lehrach	√	510
362Lh/4	ICRF/ Lehrach	√	?
362Br/1	St. Mary's	√	?
362Br/2	St. Mary's	√	?

Table 4.2.1 YACs isolated with the marker *D2Mit362*

A question mark indicates those YACs which did not show a single clone on pulse field analysis (figure 4.2.3b). The size of YACs 362Ln/2 and 362Lh/1 were not ascertained as they proved to be unstable at an early stage in YAC analysis. N/D indicates not done.

4.2.2 Pulse field analysis of *D2Mit362* YACs

Agarose plugs prepared from YACs 362Ln/1, 362Lh/2, 362Lh/3, 362Lh/4, 362Br/1 and 362Br/2 were analysed by pulse field gel electrophoresis (section 2.5.6). In figure 4.2.2, 362Lh/2, 362Lh/3 and 362Br/1 YACs are clearly visible as distinct chromosomes, whilst YACs 362Ln/1, 362Br/2 and 362Lh/4 are not resolved from yeast chromosomes under these running conditions.

4.2.3 Southern blot analysis of *D2Mit362* YACs

Southern blots were made of 362 YAC pulse field gels as detailed in section 2.6.1. A *D2Mit362* primer was end-labelled (section 2.8.3) and hybridised (section 2.9.3) to a 362 YAC pulse field blot. Figure 4.2.3a shows *D2Mit362* to be present in the YACs 362Ln/1, 362Lh/2, 362Lh/3, 362Lh/4, 362Br/1 and 362Br/2.

Probes generated from the right and left arms of the YAC vector (section 2.7.4) were radiolabelled (section 2.8.1) and hybridised (section 2.9.3) to a Southern blot made from the pulse field gel shown in figure 4.2.2. The right arm probe (figure 4.2.3b) shows a hybridisation pattern consistent with the presence of a single YAC clone in YACs 362Ln/1, 362Lh/2 and 362Lh/3. YACs 362Lh/4, 362Br/1 and 362Br/2 show a hybridisation pattern which indicates the presence of either two YACs in a single clone, or a secondary YAC with a deleted insert. Hybridisation of the probe to the larger YAC in both 362Br/1 and 362Br/2 is considerably weaker than its hybridisation to the smaller one. Figure 4.2.3a shows that, of the two YACs present in 362Br/1 and 362Br/2, only the smaller YAC is picked up by the *D2Mit362* probe. From this data it would appear that only the smaller YACs present in 362Br/1 and 362Br/2 are positive for the *D2Mit362* region of mouse chromosome 2. This could be due to the larger YACs not being of chromosome 2 origin or

because the *D2Mit362* probe hybridises to the larger YACs too weakly to give a visible hybridisation signal, in the way that the right arm probe gives a weaker signal in figure 4.2.3b. It is of interest that in YAC 362Lh/4, it is the larger YAC that is positive for the *D2Mit362* marker, thus the smaller YAC is most likely a derived deletion product. Hybridisation of the left arm probe to these YACs gives an identical result (figure 4.2.3c), supporting the conclusion that 362Ln/1, 362Lh/2 and 362Lh/3 contain a single YAC clone, whilst 362Lh/4, 362Br/1 and 362Br/2 contain either 2 YACs in a single clone or an additional, deleted YAC.

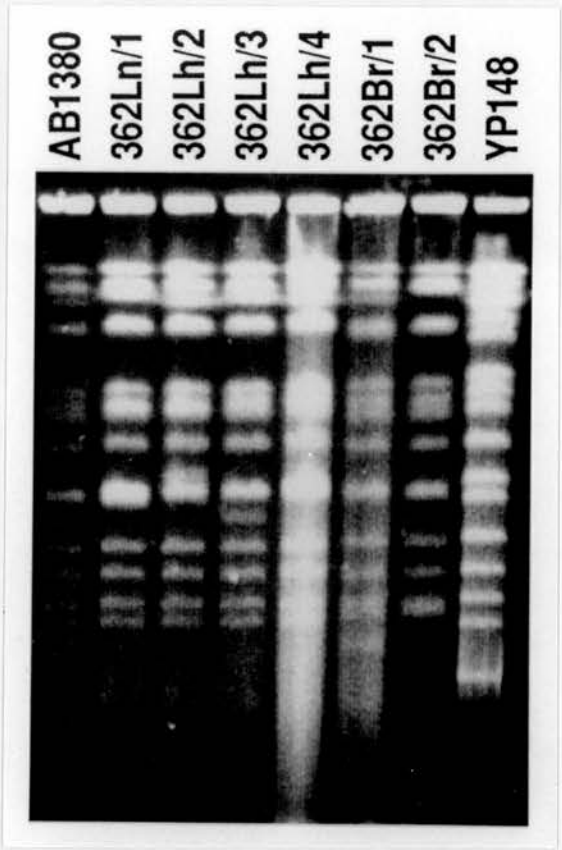

Figure 4.2.2 Resolution of *D2Mit362*-isolated YACs by pulse field gel electrophoresis

Agarose plugs were electrophoresed in 0.5 X TBE at 12°C, at 200 volts for 24 hours. The pulse rate was ramped from 50-90 seconds. Chromosome sizes for marker strains AB1380 and YP148 are given in table 2.3.

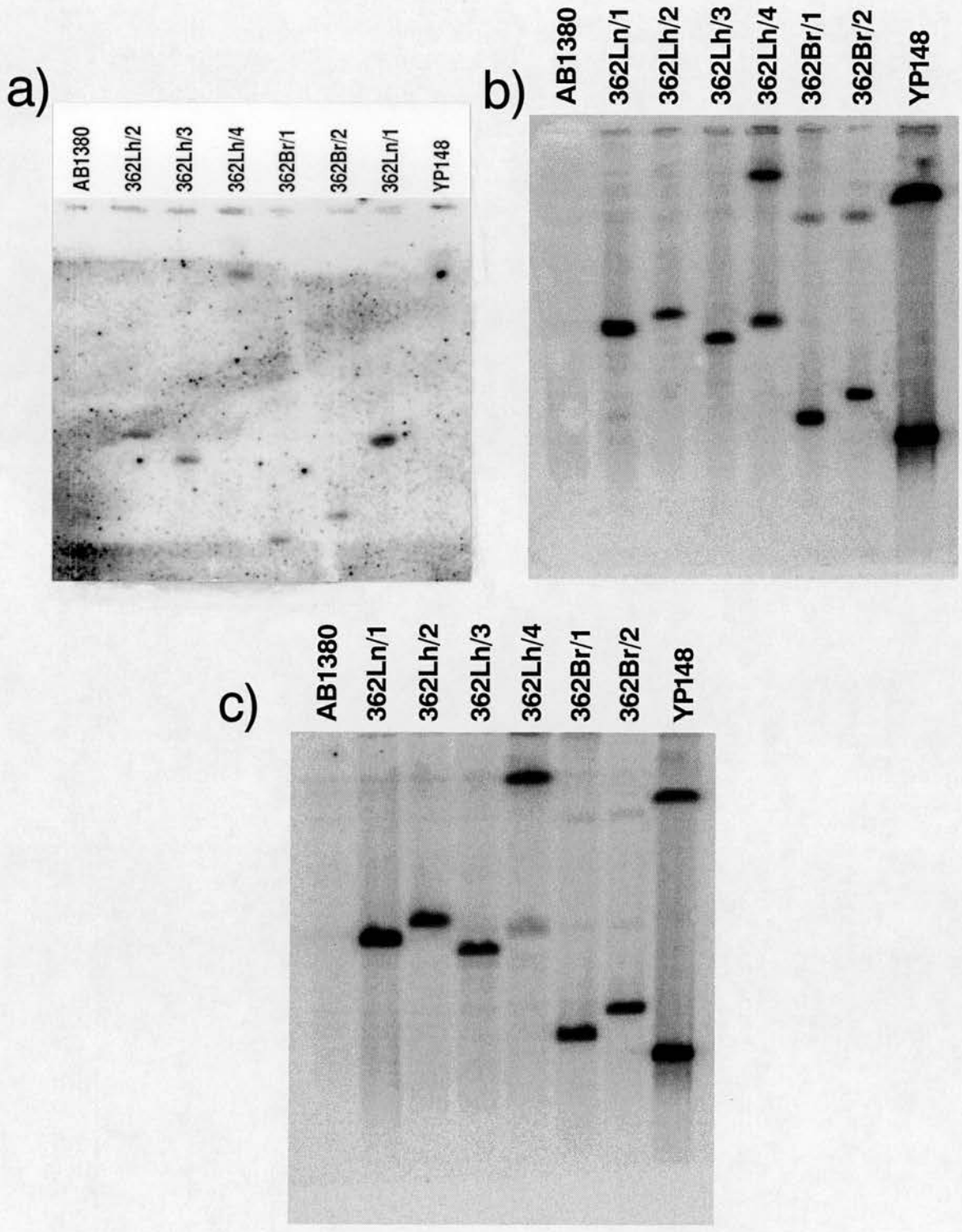

Figure 4.2.3 Southern blot analysis of the *D2Mit362*-isolated YACs

The filter used in figures *a, b* and *c* was prepared from the pulse field gel shown in figure 4.2.2, as in section 2.6.1. *Figure a* shows the hybridisation of the *D2Mit362* marker to the YACs 362Ln/1 to 362Br/2. *Figure b* shows the hybridisation of the right arm probe of the YAC vector, and *figure c* that of the left arm probe of the vector, to the *D2Mit362*-isolated YACs.

4.3 YAC end fragment analysis

4.3.1 Generation of YAC end fragments

End fragments from the mouse DNA inserts of the 5 of the 362 YACs were isolated using a method adapted from Arveiler and Porteous, 1991 (section 2.10.6). A summary of the end fragments that were generated is given in table 4.3.1.

4.3.2 Southern blot analysis of end fragments

Yeast genomic DNA was prepared (section 2.3.3) from YACs 362Ln/1, 362Lh/2, 362Lh/3, 362Lh/4, 362Br/1and 362Br/2 and from 4 control YACs: 160, I80D2, I84A11 and I82E2. Following restriction with *Eco*RI (section 2.7.1), the digests were electrophoresed (section 2.5.2) and Southern blots prepared from the gels (section 2.6.1).

In order to assemble a contig of the 362 YACs, each end fragment in table 4.3.1 was radiolabelled (section 2.8.1) and hybridised (section 2.9.3) to the Southern blots prepared from the *Eco*RI digests of YAC DNA. After radiolabelling, all end fragments were preannealed to mouse genomic DNA to compete out hybridisation of the probe to murine repetitive sequences in the YAC insert (section 2.8.2).

YAC	Left arm probe	Size of left arm probe	Southern blot variant	Right arm probe	Size of right arm probe	Southern blot variant
362Ln/1	Ln1LAT	~600bp	*Hind*III RFLV	Ln1RAR	~900bp	*Taq*I RFLV
362Lh/2	Lh2LAT	~750bp	Not found	Lh2RAP	~600bp	Not found
362Lh/3	Lh3LAT	~800bp	*Eco*RI RFLV	RA probe not generated	N/D	N/D
362Br/1	Br1LAT	~950bp	N/D	Br1RAR	~700bp	N/D
362Br/2	Br2LAT	~400bp	N/D	Br2RAR	~600bp	N/D

Table 4.3.1 End fragments generated from 362 YACs

For left arm and right arm probes, the last letter of the name indicates the restriction enzyme digest from which the probe was generated. T=*Taq*I, R= *Rsa*I and P=*Pst*I. End fragments were not generated for the YAC 362Lh/4. Not found has been used where a restriction fragment length variant (RFLV) was not found, and N/D indicates where a procedure was not done.

When hybridised to the *Eco*RI blots, the right arm end fragment from YAC 362Ln/1 (Ln1RAR) gave two sets of hybridisation signal (figure 4.3.2a). The upper signals, A, appear in all tracks, and are due to the vector sequence in the end fragment hybridising to the vector DNA in the YACs. The second set of signals, B, are derived from the YACs' insert, as this hybridisation signal is not present in the control YACs, indicating that it is not as a result of the end fragment hybridising to vector or yeast DNA. This result shows that the end fragment Ln1RAR hybridises to the YAC it was generated from, 362Ln/1, and to a similarly sized fragment in YACs 362Br/2, 362Lh/2 and 362Lh/3 indicating that the right arm of 362Ln/1 is present in these YACs. Ln1RAR does not hybridise to 362Br/1 or 362Lh/4, and is unlikely therefore to overlap with these YACs.

The left arm end fragment generated from 362Ln/1 (Ln1LAT) does not give a clear hybridisation signal (figure 4.3.2b). Bands appear to be present in all the lanes and could be due to the vector sequence in the end fragment recognising the vector DNA in the YACs, or a repeat sequence in the end fragment which is present in mouse or yeast DNA. The bands that appear in 362Br/1 and 362Br/2 do not appear in 362Ln/1, and so are unlikely to be insert derived. It is not possible to conclude from these data with which YACs the left arm of 362Ln/1 overlaps.

The right arm end fragment from YAC 362Lh/2 (Lh2RAP) gives two hybridisation signals when used as a probe on the *Eco*RI blots (figure 4.3.2c). Signal A is in all the lanes, including the control YACs, and is due to the vector sequence in the end fragment hybridising back to the YACs. Signal B shows the end fragment hybridising to the YAC from which it was derived (362Lh/2) and to an identically sized fragment in YAC 362Ln/1. This hybridisation signal does not arise in any of the other YACs, indicating that the right arm of YAC 362Lh/2 is present in only 362Ln/1.

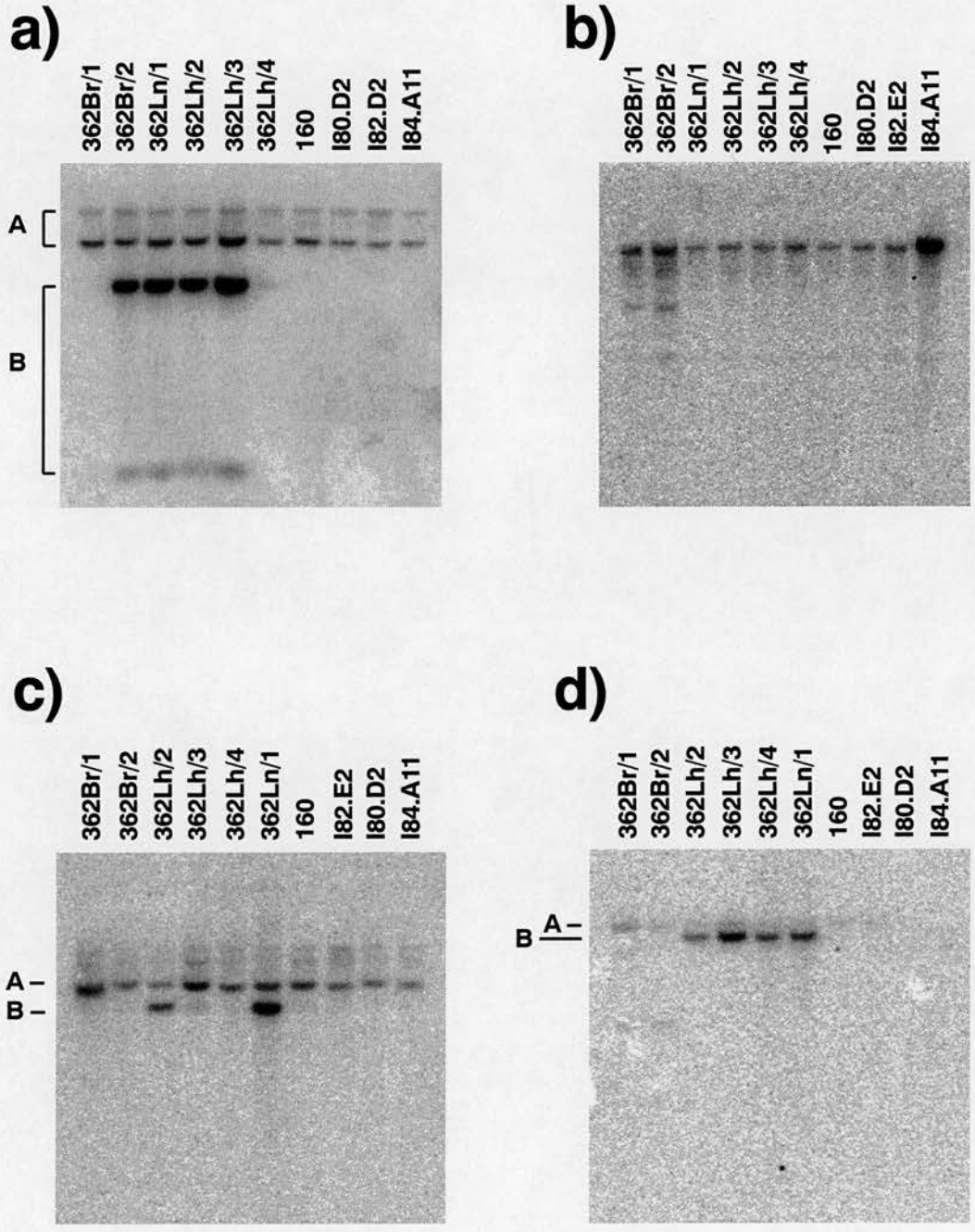

Figure 4.3.2. Southern blot analysis of YAC end fragments *(a-d)*

End fragment probes were hybridised to Southern blots prepared from *Eco*R1 digests of YAC DNA. Hybridisations with the following end fragments are shown: a*)* Ln1RAR; b) Ln1LAT; c) Lh2RAP; and d) Lh2LAT.

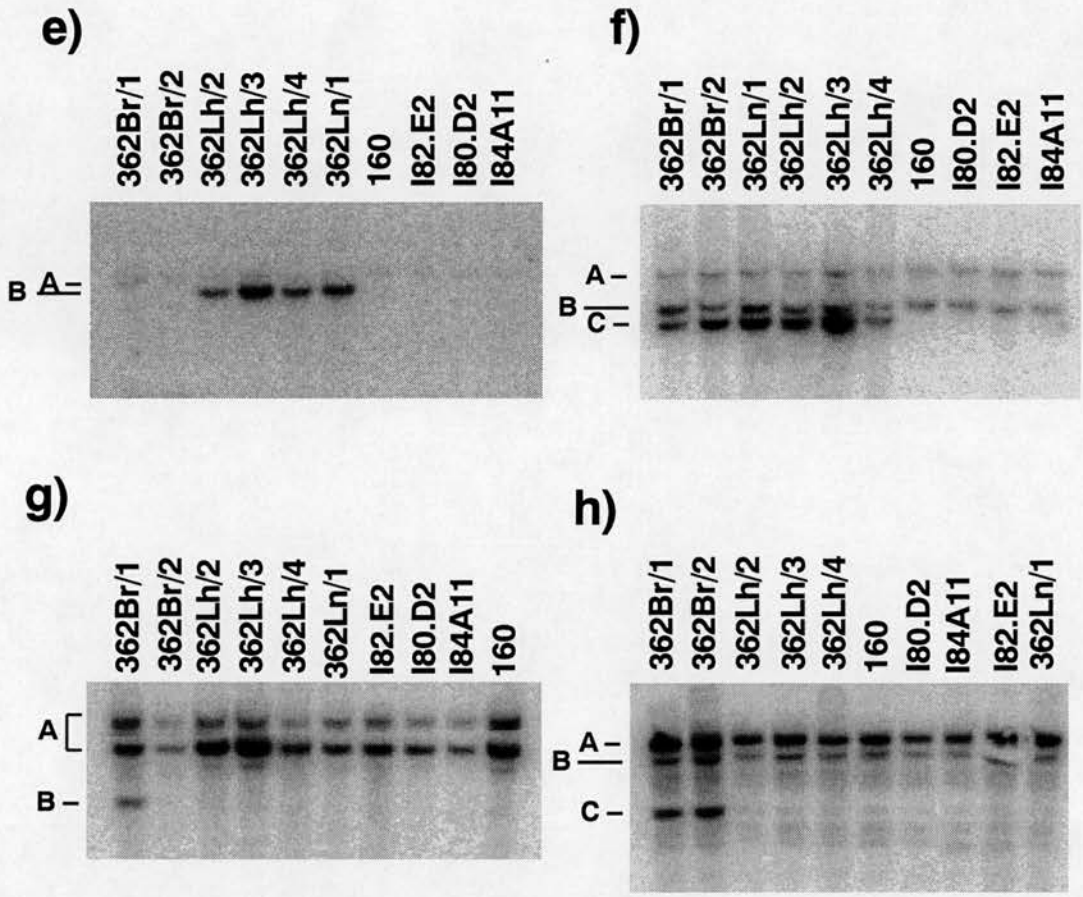

Figure 4.3.2. Southern blot analysis of YAC end fragments *(e-f)*

End fragment probes were hybridised to Southern blots prepared from *Eco*R1 digests of YAC DNA. Hybridisations with the following end fragments are shown: e*)* Lh3LAT; f) Br1RAR; g) Br1LAT; and h) Br2LAT.

The hybridisation signal from the left arm end fragment of 362Lh/2 (Lh2LAT) also gives two hybridisation signals (figure 4.3.2d). Signal A appears faintly in all the lanes, including the control YACs and is probably the vector derived signal. Signal B shows Lh2LAT hybridising back to the YAC from which it was derived, 362Lh/2 and hybridising to similarly sized fragments in YACs 362Lh/3, 362Lh/4 and 362Ln/1. It is therefore possible that the left arm of YAC 362Lh2 overlaps with these YACs and not with YACs 362Br/1 and 362Br/2.

The left arm end fragment generated from YAC 362Lh/3 (Lh3LAT) gives two hybridisation signals when hybridised to the *Eco*RI blots of YAC DNA (figure 4.3.2e) which appear to be almost identical to the Lh2LAT hybridisation pattern in figure 4.3.2d. The fainter signal, A, is present in all the lanes and is likely to be the vector derived signal. The second, more intense signal, B shows Lh3LAT hybridising back to the YAC from which it was generated, 362Lh/3, and to similarly sized fragments in YACs 362Lh/2, 362Lh/4 and 362Ln/1. This hybridisation signal is not present in the control YACs or in 362Br/1 or 362Br/2. It is possible, therefore, that the left arm of 362Lh/3 is present in 362Lh/2, 362Lh/4 and 362Ln/1, or that the end fragment is repeat-derived, and that the same repeat is present in end fragment Lh2LAT. An alternative explanation for these results could be that Lh3LAT and Lh2LAT have been derived from the same YAC insert sequence. Because of these ambiguities, it is not possible to conclude which YACs Lh2LAT and LH3LAT definitely overlap with. Right arm end fragments for 362Lh/3 could not be generated using the inverse PCR method.

When hybridised to the YAC *Eco*RI blots (figure 4.3.2f), the right arm end fragment of YAC 362Br/1 (Br1RAR) gave 3 hybridisation signals. Signals A and B, present in all the lanes, are likely to be either vector or repeat derived. Signal C is of a stronger intensity and is not present in the lanes containing control YAC DNA. This signal shows Br1RAR hybridising back to the YAC it

was generated from, 362Br/1, and hybridising to a similarly sized band in YACs 362Ln/1, 362Lh/2, 362Lh/3, 362Lh/4 and 362Br/2. This suggests that the right arm of this YAC is present in all the 362 YACs.

The left arm end fragment generated from YAC 362Br1 (Br1LAT) gives a number of bands when hybridised to the *Eco*RI blots of YAC DNA (figure 4.3.2g). Two sets of bands, A, are common to all tracks. A single band, B, can be seen where the probe has hybridised back to the YAC from which it was derived. Br1LAT does not hybridise to a similarly sized band in the other lanes, indicating that the left arm of 362Br/1 is not present in any of the 362 YACs. Either this end of the YAC is at the extreme end of the contig, or this part of the YAC is not derived from the *D2Mit362* region of mouse chromosome 2 and is therefore chimeric.

The right arm end fragment generated from 362Br/2 (Br2RAR) hybridised to multiple fragments when used as a probe on the YAC *Eco*RI blot. When the end fragment was remade from a *Eco*RV restriction digest of YAC 362Br/2 DNA, the new probe gave a similarly unclear result (data not shown). It is not possible to conclude from this result whether the right arm end of 362Br/2 is present in the other 362 YACs, or whether it is chimeric.

The left arm end fragment made from 362Br/2 (Br2LAT) gives 3 classes of hybridisation signals when hybridised to the *Eco*RI blot of YAC DNA (figure 4.3.2h). Two of these signals, A and B, are present in all the YACs, including the controls and are likely to be vector derived. Signal C shows Br2LAT hybridising back to the YAC it was generated from, 362Br/2, and to a similarly sized fragment in YAC 362Br/1. This suggests that this end of 362Br/2 is present in YAC 362Br/1, but not in any of the other 362 YACs.

Figure 4.3.3 gives a summary of the YAC contig that was assembled on the basis of this data.

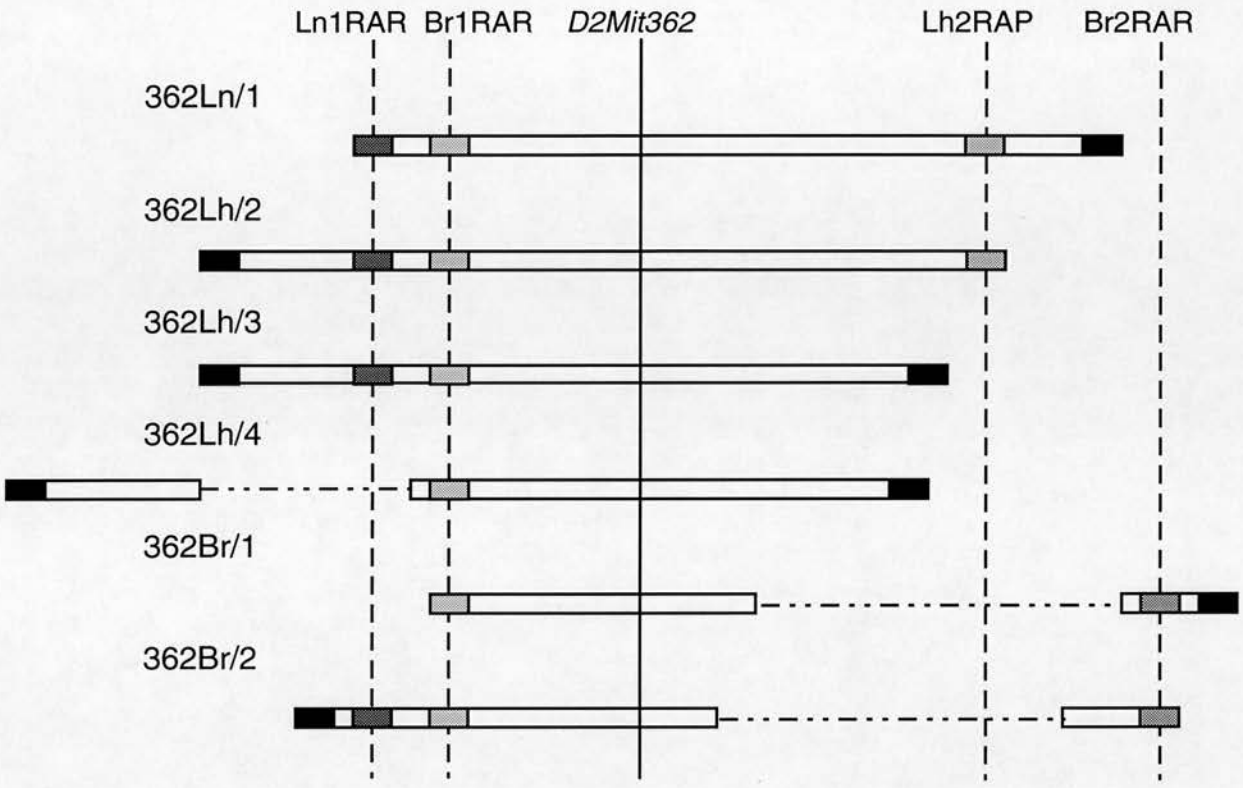

Figure 4.3.3 Contig of 362 YACs generated from end fragment analysis

The presence of the marker *D2Mit362* is indicated by a line running through all the YACs. The presence of end fragments in other YACs is indicated by a broken line. Filled in YAC ends (▬) are those which could not be placed onto the contig. Hatched and shaded YAC ends are those which are present in other YACs. The presence of end fragment Ln1RAR is represented by a dark hatched box (▩), end fragment Br1RAR is represented by a light hatched box (▩), Lh2RAP by a stippled box (▩) and Br2RAR by a shaded box (▩). Internal deletions are indicated by a broken line in YACs 362Lh/4, 362Br/2 and 362Br/1.

4.3.3 Discussion of contig assembly

The contig was assembled using results from those end fragments which clearly hybridised to other 362 YACs, primarily Ln1RAR and Lh2RAP, and were the least likely to be chimeric. Ln1RAR, the end fragment derived from the right arm of 362Ln/1, is present in 3 YACs: 362Lh/2, 362Lh/3 and 362Br/2. Since this end fragment is not present in 362Lh/4, it is possible that the insert region of 362Lh/4 which would have contained the Ln1RAR sequence has been deleted (see figure 4.2.3b). Since hybridisation with Lh2RAP clearly shows the right arm end of 362Lh/2 to be present in 362Ln/1, it is unlikely that the left arm end of this YAC could also be present in 362Ln/1, due to 362Lh/2 being larger in size. It is therefore doubtful that Ln1RAR is at the extreme end of the contig, but in the absence of clear data from the end fragments Lh2LAT and Lh3LAT, it is not possible to conclude which end fragment marks the left hand contig boundary.

The other extreme of the contig can be defined by the result that shows that the right arm of 362Lh/2 is present in only 362Ln/1. In the absence of data from the left arm end fragment of this YAC (Ln1LAT), this result suggests that the left arm of 362Ln/1 is at the end of the contig, however, it is not possible to tell from this result whether it is chimeric or not. The left arm end fragment from 362Br/1 (Br1LAT), which is not present in the other 362 YACs, could be at the end of the contig, but since it is likely that this YAC has either a chimeric or deleted inset (see figure 4.2.3b), it would not be a very informative marker to use in further analysis. The presence of deletions in YACs 362Br/1, 362Br/2 and 362Lh/4 are proposed on the basis of the pulse field results in section 4.2, which show the presence of additional clones in each YAC. However, there are alternative contig arrangements of the YACs that could be derived from this data. To ensure that the correct end fragments were used in a chromosomal walk, those thought to be at the extremes of the contig were mapped onto the backcross.

4.4 Analysis of backcross recombinants

As a result of the Southern blot analysis, the end fragments thought to be either non-chimeric or at the ends of the contig were Ln1RAR, Ln1LAT and Lh2RAP. Lh3LAT was also included in this analysis to gauge whether it was at a more extreme point in the contig than Ln1RAR. In order to map these end fragments onto the *Sd* backcross, RFLVs between *M.m.castaneus* and CBA DNA were identified.

DNA from *M.m.castaneus* and CBA DNA was digested with the restriction enzymes: *Bgl*II, *Eco*RV, *Eco*RI, *Pst*I, *Bam*HI, *Hind*III and *Taq*I (section 2.7.1). The digests were electrophoresed (section 2.5.2) and Southern blotted (section 2.6.1).

End fragments Ln1RAR, Ln1LAT, Lh2RAP and Lh3LAT were radiolabelled (section 2.8.1), preannealed to genomic mouse DNA, to minimise hybridisation of the probe to murine repeat sequences (section 2.8.2) and hybridised to the CBA/ *M.m.castaneus* polymorphic detection blots (polyblots).

4.4.1 Backcross analysis with end fragment Ln1RAR

When used as a probe on the CBA/ *M.m.castaneus* DNA polyblot, the right arm end fragment generated from YAC 362Ln/1 (Ln1RAR), gave a hybridisation signal that showed a *Taq*I RFLV between CBA and *M.m.castaneus* DNA (figure 4.4.1a). The difference in size between the CBA and *M.m.castaneus* bands in the *Bam*HI lane is most likely due to incomplete restriction giving rise to partially digested fragments. The remaining lanes show that the probe hybridises to the identically sized fragments in CBA and *M.m.castaneus* DNA when restricted with *Bgl*II, *Eco*RV, *Eco*RI, *Pst*I and *Hind*III.

Genomic DNA was prepared (section 2.4.1) from the mice recombinant between *D2Mit118* and *D2Mit80* and restricted with *Taq*I (section 2.7.1). The digests were electrophoresed and Southern blots prepared (section 2.6.1). Ln1RAR was radiolabelled as before and hybridised to the *Taq*I blot of the recombinant mice. Figure 4.4.1b shows that this probe gave two signals when hybridised to the *Taq*I blot. Signal A is the CBA-specific fragment detected by the probe, and can be seen in all the lanes, except for the *M.m.castaneus* DNA control lane. Since the backcross progeny were generated by backcrossing to CBA, it is expected that all the backcross offspring be positive for this band. Signal B is a band of smaller size and is present in the *M.m.castaneus* DNA control, and in mice 26, 136, 234, 271 and the non-recombinant DNA control. These mice, heterozygous for the *M.m.castaneus* and CBA alleles, were haplotyped at this locus (as in the chapter 3) as being C, and those homozygous for the CBA allele, were haplotyped as being M (table 4.4.2).

To ensure that the end fragment Ln1RAR was derived from the *D2Mit362* region of mouse chromosome 2, genomic DNA from the first 25 animals of the backcross was restricted with *Taq*I. A Southern blot was prepared of the digests and probed with Ln1RAR (figure 4.4.1c). The backcross offspring on this blot were haplotyped as above. From these results, it was possible to conclude that the right arm of 362Ln/1 (Ln1RAR) was chromosome 2 derived (table 4.4.1), and that it mapped two recombinants distal to *D2Mit267* and was non-recombinant with *D2Mit362* on the *Sd* backcross (figure 4.4.5).

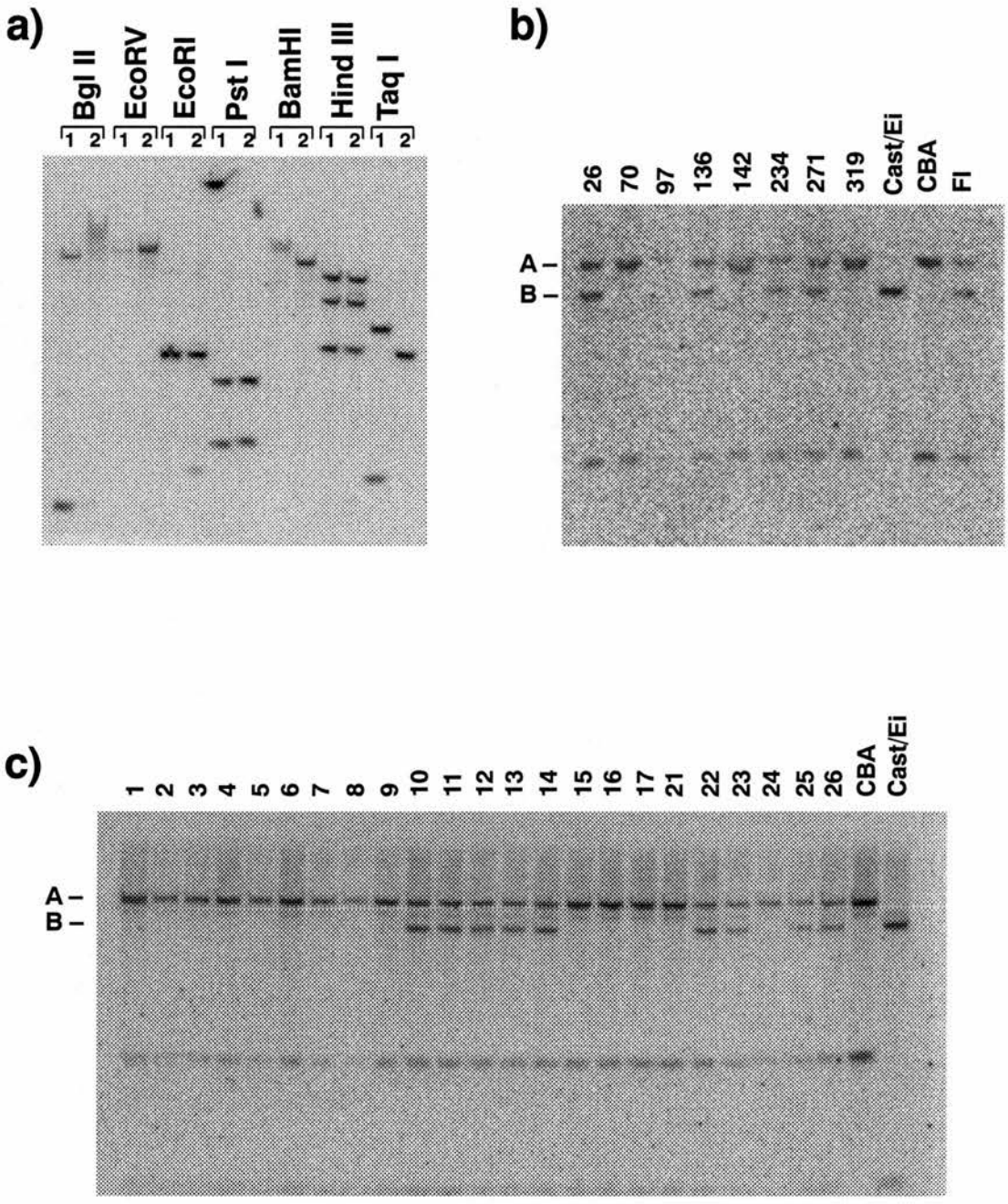

Figure 4.4.1 Mapping of YAC end fragment Ln1RAR onto *Sd* backcross

1 denotes CBA DNA, **2** denotes CAST/Ei DNA in figure *a*. The last 2 tracks show the *Taq*1 RFLV used to map Ln1RAR. The mapping of Ln1RAR onto the recombinant panel of mice is shown in *b*, and onto the panel of non-recombinant mice in *c*.

4.4.2 Backcross analysis with end fragment Ln1LAT

When the left arm end fragment of YAC 362Ln/1 (Ln1LAT) was hybridised to the polyblot shown in figure 4.4.1a, the probe hybridised to differently sized *M.m.castaneus* and CBA *Hind*III fragments. This RFLV was used to map the end fragment onto the backcross.

Genomic DNA from the recombinant panel of mice, and from the first 15 mice of the *Sd* backcross, was restricted with *Hind*III. The digests were electrophoresed and Southern blotted. Ln1LAT was radiolabelled and hybridised to the *Hind*III blot (figure 4.4.2). Lanes 1, 2 and 4 do not give a hybridisation signal with this probe as these DNAs were not fully digested. The remaining lanes show two hybridisation signals. Signal A is the smaller band and is present in all the lanes except for the *M.m.castaneus* DNA control and lanes 1, 2, and 4. This is the CBA-specific band and it is present in all the backcross offspring. Signal B, the larger band, is present in the *M.m.castaneus* control lane and is the *M.m.castaneus*-specific band. This band is present in mice 10, 11, 12, 13, 14, 136, 234 and 271. As in chapter 3, these mice were haplotyped as being C at this locus and the mice with only band A were haplotyped as being M. Since the typing of the first 15 mice at this locus was in concordance with their haplotyping at other backcross markers (table 4.4.1), it was concluded that this end of YAC 362Ln/1 was derived from the proximal region of chromosome 2, indicating that the YAC is non-chimeric at both ends. Since there is no data from this end fragment to show that it has sequence in common with the other YACs, it is not known on the basis of this analysis whether 362Ln/1 contains any internal deletions.

Analysis of the recombinant panel of mice (table 4.4.2) positioned Ln1LAT on the backcross map as being non-recombinant with *Sd* (figure 4.4.5 and 4.5), crossing one recombinant between marker *D2Mit362* and *Sd*. Because

Ln1LAT provided a new marker which was non-recombinant with *Sd*, the end fragment was sequenced (section 2.11) in order to generate PCR primers which could be used for the screening of YAC libraries.

4.4.3 Backcross analysis with end fragment Lh3LAT

The left arm of YAC 362Lh/3 (Lh3LAT) was used as a probe on the polyblot shown in figure 4.4.1, and hybridised to differently sized *M.m.castaneus* and CBA *Eco*RI fragments. This RFLV was used to map the end fragment onto the backcross.

Genomic DNA from the recombinant panel of mice and four control DNA's from the backcross were restricted with *Eco*RI. The digests were electrophoresed and Southern blots prepared from the gels. When Lh3LAT was hybridised to the *Eco*RI blot two bands could be identified (figure 4.4.3). Band A, the smaller CBA-specific band was present in all the lanes except for the *M.m.castaneus* DNA control lane; and band B, the larger *M.m.castaneus*-specific band, was present in the *M.m.castaneus* DNA control and lanes 26, 136, 234, 271, and control DNAs 10 and 27. The background smear in each lane is likely to be due to the probe hybridising to a murine repeat sequence. Since the haplotyping of the control mice at this locus was concordant with their haplotyping at other backcross markers (table 4.4.1), we concluded that this end of YAC 362Lh/3 was derived from the proximal region of chromosome 2. Haplotype analysis of the recombinant mice (table 4.4.2) placed this marker 2 recombinants distal to *D2Mit267*, showing that it was non-recombinant with *D2Mit362* and the right arm end fragment of YAC 362Ln/1, Ln1RAR, and one recombinant away from *Sd* (figure 4.4.5 and 4.5).

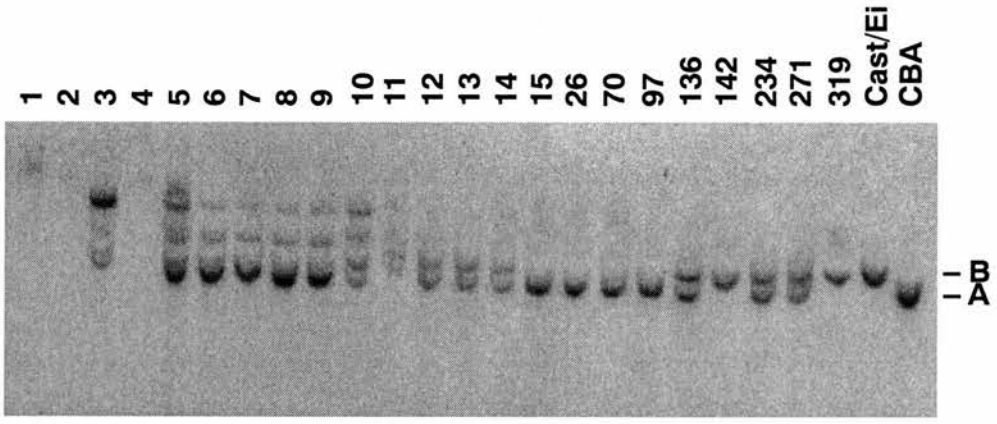

Figure 4.4.2 Mapping of YAC end fragment Ln1LAT onto *Sd* backcross
Ln1LAT was mapped onto the backcross using a *Hind*III RFLV. Tracks 1-15 show DNA from the non-recombinant panel of mice, and the remaining tracks show DNA from the recombinant mice, with CBA and CAST/Ei DNA controls. DNA samples were digested with *Hind*III.

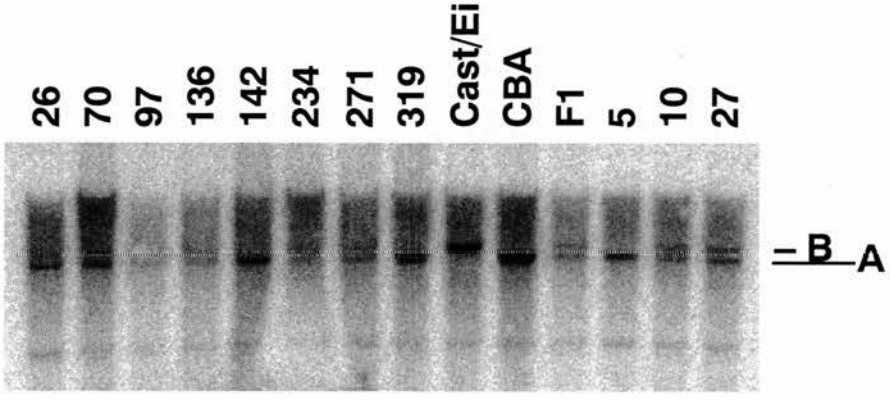

Figure 4.4.3 Mapping of YAC end fragment Lh3LAT onto *Sd* backcross
Lh3LAT was mapped onto the backcross using a *Eco*RI RFLV. The first 8 tracks show DNA from the recombinant panel of mice. The remaining tracks show CBA, CAST/Ei, F1 and non-recombinant DNA controls. DNA samples were digested with *Eco*RI.

4.4.4 Backcross analysis of end fragment Lh2RAP

Despite the clean hybridisation signal that this probe gave when hybridised to YAC DNA, it was not possible to use this end fragment as a probe on mouse genomic DNA blots, as it contained a murine repeat sequence. Pre-annealing the probe to mouse genomic DNA to compete out recognition of repeat sequences did not work sufficiently well to make the probe usable. To overcome this problem, the end fragment was sequenced (section 2.11.1) in order to generate PCR primers (section 2.10.3).

The primers that were generated (table 2.10.4) were used to PCR amplify (section 2.10) the end fragment sequence from CBA and *M.m.castaneus* DNA. PCR products were analysed for size differences between the two sub-species of DNA. Since no size differences were found, the PCR products were restricted with a range of frequent cutter enzymes (section 2.7.3) and the digestion fragments electrophoresed on 4% NuSieve (section 2.5.2) and 8% acrylamide gels (section 2.5.4). However, no polymorphisms were identified by which this end fragment could be mapped onto the backcross. Since data suggests that this marker could be at the extreme end of the contig in the direction of *Sd*, work is in progress to find a means of mapping this marker.

No.	*D2Mit 31*	*D2Mit 464*	Lh3 LAT	Ln1 RAR	*D2Mit 362*	*Sd*	Ln1 LAT	*D2Mit 80*	*Pax8*
1	M	M		M		M			M
2	M			M		M			M
3	M			M		M			M
4	M			M		M			M
5	M		M	M		M			M
6	M			M		M	M		M
7	M			M		M	M		M
8	M			M		M	M		M
9	M			M		M	M		M
10	C		C	C		C	C		C
11	C			C		C	C		C
12	C		C	C		C	C		C
13	C			C		C	C		C
14	C			C		C	C		C
16	M			M		M			M
17	M			M		M			M
21	M			M		M			M
22	C			C		C			C
23	C			C		C			C
24	M			M		M			M
25	C			C		C			C

Table 4.4.1 Genetic mapping of YAC derived markers to chromosome 2.

End fragment markers, Lh3LAT, Ln1RAR and Ln1LAT, derived from YACs isolated with the *D2Mit362* marker, were mapped onto a panel of mice from the *Sd* backcross which were non-recombinant between the markers *D2Mit31* and *Pax8*. This was done to ensure that the markers were derived from the *D2Mit362* region of mouse chromosome 2. Boxes which are lightly shaded denote loci which were not typed. Mice scored with an M were homozygous for the CBA-sized allele, those scored with a C were heterozygous for the CBA and *M.m.castaneus* sized alleles.

No.	*D2Mit 31*	*D2Mit 464*	Lh3 LAT	Ln1 RAR	*D2Mit 362*	Ln1 LAT	*Sd*	*D2Mit 80*	*D2Mit 7*
295	C	C	N/D	N/D	C		C	C	M
271	C	C	C	C	C	C	C	M	M
234	C	C	C	C	C	C	C	M	M
26	C	C	C	C	C	M	M	M	M
142	C	M	M	M	M	M	M	M	M
265	C	M	N/D	N/D	M	N/D	M	M	M
264	C	M	N/D	N/D	M	N/D	M	M	M
319	C	M	M	M	M	M	M	M	M
97	C	M	M	M	M	M	M	M	M
70	C	M	M	M	M	M	M	M	M
199	C	N/D	N/D	N/D	N/D	N/D	M	M	M
306	M	M	N/D	N/D	M	N/D	M	M	C
330	M	M	N/D	N/D	M	N/D	M	M	C
318	M	M	N/D	N/D	M	N/D	M	M	C
277	M	M	N/D	N/D	M	N/D	M	M	C
15	M	M	N/D	M	M	M	M	N/D	C
136	M	C	C	C	C	C	C	C	C
242	M	C	N/D	N/D	C	N/D	C	C	C

Table 4.4.2 Genetic mapping of YAC derived markers onto recombinant panel of mice.

End fragment markers, Lh3LAT, Ln1RAR and Ln1LAT, derived from YACs isolated with the *D2Mit362* marker, were mapped onto a panel of mice from the *Sd* backcross which were recombinant between the markers *D2Mit31* and *Pax8*. The mice were scored as being M or C as for table 4.3.1. Loci not haplotyped are indicated (N/D, not done). Only one end fragment, Ln1LAT, crosses a recombination point, in mouse no. 26, indicating that this end of YAC 362Ln/1 is segregating with *Sd*.

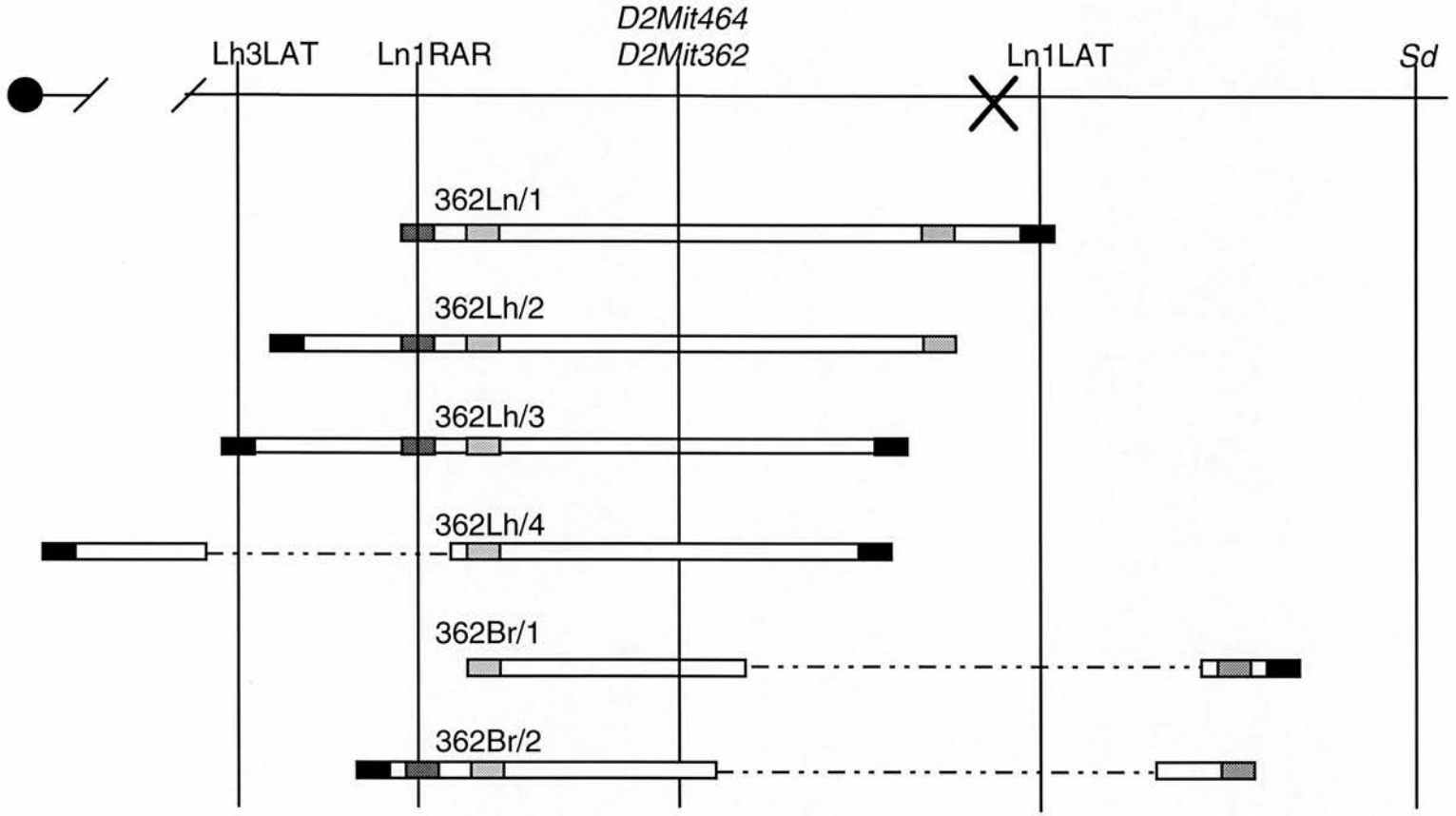

Figure 4.4.5 Orientation of YAC contig with respect to *Sd.*
End fragments are shaded as in figure 4.3.3. The line at the top of the diagram represents chromosome 2 with the centromere to the left handside of the page. The positions of the end fragments Lh3LAT, Ln1RAR and Ln1LAT with respect to the markers *D2Mit362, D2Mit464* and *Sd* are shown on this line. The cross to the left handside of Ln1LAT represents the recombinant that the end fragment crosses, placing it on the *Sd* side of the recombination point.

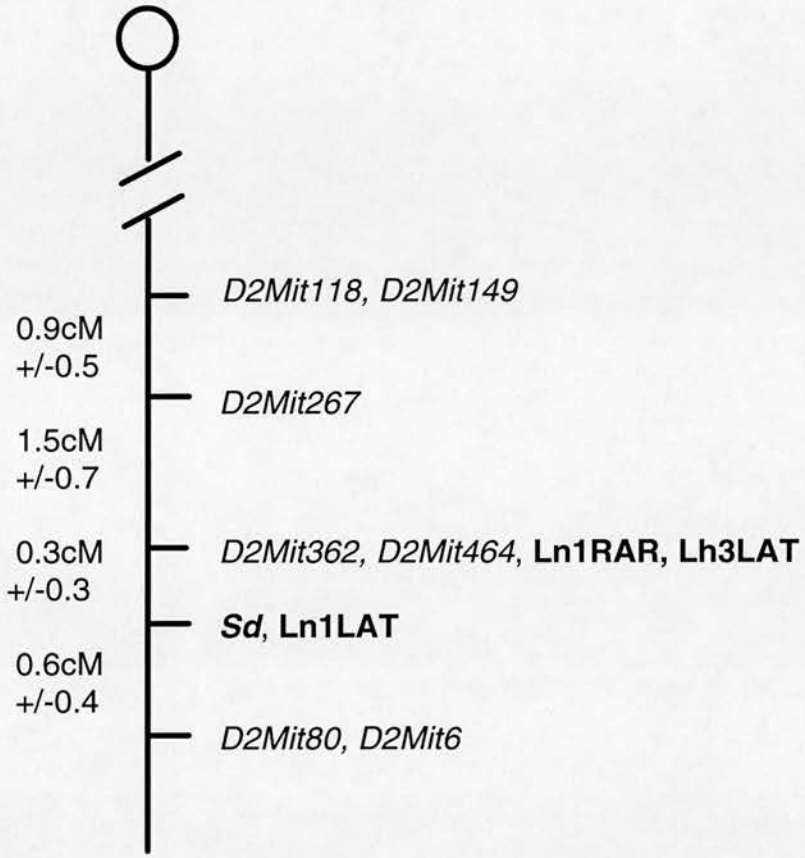

Figure 4.5 Genetic map of proximal mouse chromosome 2, showing the map position of new end fragment markers relative to *Sd*.

End fragment markers were derived from YACs isolated with the *D2Mit362* marker and are shown in bold type (Ln1RAR, Ln1LAT and Lh3LAT). As can be seen from this map, Ln1LAT is non-recombinant with *Sd*. Genetic distance, in centiMorgans (cM), between the markers is given on the left hand side of the figure. CentiMorgan and standard error values were calculated as in section 2.12.

4.4.5 Discussion of end fragment mapping

By mapping the end fragments Ln1RAR, Ln1LAT and Lh3LAT onto the *Sd* backcross, it was possible to orientate the YAC contig with respect to *Sd* (figure 4.4.5). Ln1LAT crosses a single recombinant between the marker *D2Mit362* and *Sd* and thus provides the project with a new marker which is non-recombinant, and therefore closely physically linked with *Sd*. With the aim of isolating new YACs which will cross the recombinants on the *D2Mit80* side of *Sd*, PCR primers generated from Ln1LAT will be used in the rescreening of YAC libraries. Once a contig is established in which YAC ends are crossing recombinants on both sides of *Sd*, it will be possible to identify YACs in the contig which contain the *Sd* gene. The isolation of the gene is then feasible through the use of techniques such as exon trapping (Nehls *et al.*, 1994), cDNA selection (Lovett *et al.*, 1991) and candidate transcript analysis. A detailed discussion of the approaches which can be taken to isolate candidate gene transcripts is given in chapter 7.

Chapter 5: Mapping tail-length modifying loci in *Sd* mice

5.1 Introduction

Mutations in many genes which follow monogenic, Mendelian patterns of inheritance allow the genotype of a particular locus to be unambiguously inferred from the phenotype of an individual. These alleles, which give rise to major phenotypic effects, have been called macromutations (Tanksley, 1993).

Whilst macromutations are common in the laboratory, they are rarer in natural populations since many are deleterious to survival and are removed through natural selection. The incidence of macromutations sustained in natural populations is therefore low. Nevertheless, phenotypic and genetic variation still exists in these populations. Whilst genetic variation is occurring at discrete loci, phenotypic variation is continuous for many traits, controlled by allelic variation at several genetic loci, each contributing small effects to an overall phenotype. Traits, whose phenotypic variation is continuous and determined by the segregation of multiple loci, have been referred to as complex or quantitative traits, and their mode of inheritance as polygenic (Lander and Botstein, 1989; Tanksley, 1993). The identification of genes contributing to a quantitative phenotype is problematic in that the relationship between genotype and phenotype breaks down, since the same phenotype can result from different genotypic combinations.

As in other species, macromutations in the human genome are deleterious to survival, because of this many human diseases for which the genetic basis has been established have been rare in the human population as a whole. Since most common traits of medical relevance do not follow simple Mendelian monogenic inheritance, it has been harder to identify the genes responsible for common human complex traits such as susceptibility to heart disease, epilepsy, hypertension, diabetes, obesity and cancer (Lander and Schork, 1994; Copeland *et al.*, 1993). Disorders such as these have a high incidence in western populations, and the need for a means of analysing the

complex nature of their aetiology has given recent prominence to the field of quantitative trait loci mapping.

Although it is possible to analyse polygenic traits in human populations, the inability to control human matings and non-genetic sources of variation, such as environmental conditions, makes such analysis difficult and complex. For this reason, the mouse is an ideal experimental model to use in the genetic dissection of polygenic traits. The ability to carry out controlled matings and standardise environmental conditions remove sources of variation which are not possible to control with humans (Avner, 1994). Mouse strains also show striking variation in their susceptibility to a number of murine disorders which are models for polygenic human diseases, such as epilepsy (Rise *et al.*, 1991) and type 1 and type 2 diabetes (de Guoyon *et al.*, 1993; Todd *et al.*, 1991; Bahary *et al.*, 1990).

A good example of interstrain variation is the incidence of intestinal adenomas in multiple intestinal neoplasia (*Min*) mice (Dietrich *et al.*, 1993b). *Min* mice carry a mutation in the *Apc* gene, the mouse homologue of the human *APC* gene, responsible for the Mendelian-inherited colon cancer, adenomatous polyposis coli. It has been found that within some human APC pedigrees, family members who have inherited the same *APC* mutation, show a striking variation in the number of colonic polyps and other neoplasias that they develop. Since different laboratory strains of mice carrying the *Min* mutation also showed variation in the numbers of colonic tumours they developed, it was possible to utilise them to investigate whether the variation in mice was due to differences in genetic backgrounds. Through the mapping of backcrosses set up between the most severely affected strain (C57Bl/6J) and the least severely affected strain (Akr) of *Min* mice, a locus on chromosome 4 was found that strongly modified tumour multiplicity, called *Mom-1* (for Modifier of *Min*-1). *Mom-1* was shown to contribute ~50% of the genetic variance responsible for intestinal neoplasia

in *Min* mice. The recent identification of a mouse candidate gene for *Mom-1*, the secretory phospholipase A2 gene (MacPhee *et al.*, 1995) indicates how mouse models of complex human traits can enable the identification of modifier loci. Alternative methods for this type of analysis are discussed further in the chapter.

5.1.1 Markers used in quantitative trait loci mapping

The principles of quantitative trait loci (QTL) mapping were first established by Sax (1923), when he reported an association of seed size in beans (a quantitatively inherited character) with seed-coat pigmentation (a monogenic trait). This finding gave rise to the idea that the segregation of a single gene marker could be used to detect and estimate the effect of a linked polygene. Thoday (1961) proposed that if single gene markers were scattered throughout the genome, it should be possible to map and characterise all of the polygenes affecting a character. However, a considerable problem in the use of such marker genes was encountered when a marker gene itself had a larger effect on the quantitative character than the linked polygene, preventing its detection or an over/under estimation of its effect. For example, in plants a common monogenic marker is dwarfism. It would not be possible to use a dwarfism marker to study the polygenic inheritance of plant height since the effect of the marker gene on plant height would be much greater than that of an individual polygene. Other problems associated with the early use of marker genes were epistasis (whereby one gene interferes with the phenotypic expression of another, non-allelic gene) and lack of polymorphism in natural populations (Tanksley, 1993).

An advancement in quantitative trait loci mapping came with the discovery that allelic forms of enzymes (isoenzymes) could be electrophoretically distinguished from one another (Hunter and Markert, 1957). Such polymorphic markers allowed polygenes to be genetically mapped with greater independence from the phenotypic changes incurred with the use of

marker genes. This change in the nature of the marker used to map polygenes considerably increased the success of QTL mapping studies.

With the introduction of DNA-based genetic markers (Botstein *et al.*, 1980), it became possible to survey genetic variation directly at the DNA level. The use of restriction fragment length polymorphisms (RFLPs) and simple sequence length polymorphisms (SSLPs), greatly increased the level of polymorphic variation which could be studied. With the creation of high-density genetic linkage maps, the inheritance of markers spanning an entire genome can now be simultaneously followed to identify chromosomal regions linked to a varying trait. As with isoenzymes, polymorphism for DNA markers usually has no phenotypic affect. Today molecular linkage maps covering the entire genome are available for quantitative trait studies in a number of species including human (Weissenbach *et al.*, 1992), mice (Copeland *et al.*, 1993) and rats (Serikawa *et al.*, 1992).

5.1.2 Strategies for detecting polygenes

A variety of statistical techniques have been proposed for estimating the minimum number of genetic loci contributing to the difference in a quantitative trait between two inbred lines grown in a common environment (Castle, 1921; Wright, 1968). Wright's method assumes that, with respect to all relevant loci, one line is fixed with alleles increasing the character of interest, and the other line is fixed with alleles decreasing it. This is most likely when the divergence of the populations was caused by sustained artificial selection on the trait of interest or on some character(s) highly correlated with it genetically. Estimates of the minimum number of genes influencing such a trait are obtained by comparing the phenotypic means and variances of the character in the F_1, F_2 and backcross populations. The requirement of inbred lines, however, can produce complications through inbreeding depression on the mean and developmental instability associated with inbred lines. More recent interpretations of Wright's formulae (Lande,

1981) have shown that his method can still be applied to crosses between heterogenous (or wild) populations.

The use of marker loci to detect polygenes relies on the assumption that linkage disequilibrium (the non-random association of alleles) exists between alleles at the marker locus and alleles of the polygenes. Although linkage disequilibrium can be caused by factors such as selection and genetic drift, its predominant cause, in crosses between different strains, is the physical linkage of loci. The use of linkage disequilibrium can be very powerful in populations derived from controlled matings, such as backcross or F2/F3 populations. QTL mapping in these populations is more powerful than in human pedigrees because genetic noise can be reduced through the use of recombinant inbred and congenic strains of mice. The use of experimental crosses with large numbers of offspring, also greatly increases the power of the statistical tests which can be used to analyse linkage data (Lander and Schork, 1994).

Statistical procedures used to determine whether a polygene is linked to a marker gene share a basic principle, that is to group a population into genotypic classes based on the genotype at the marker locus and then use correlative statistics to determine whether the individuals of one genotype differ significantly compared with individuals of other genotypes with respect to the trait being measured. If the phenotypes differ significantly, it can be interpreted that a gene (or genes) affecting the trait is linked to the marker locus used to haplotype the population. This procedure can be repeated for additional marker loci throughout the genome to detect as many QTL as possible. This type of analysis does not determine whether an effect detected with a marker is due to one or more linked genes, but defines a chromosomal region that has a significant effect on a quantitative trait.

There are three approaches which can be used in the mapping of a QTL: single point analysis (Edwards *et al.*, 1987), interval mapping (Lander and Botstein, 1989) and the analysis of phenotypic extremes (Lander and Botstein, 1989; Taylor *et al.*, 1994).

Single point analysis

Single point analysis employs one marker at a time in the detection of a QTL and does not, therefore, require a complete molecular marker linkage map (Tanksley, 1993). The problem with the use of this type of analysis is that the magnitude of the QTL effect revealed by the marker is the actual allelic effect devalued by the map distance to the marker. Since this analysis does not take into account recombination between the marker and QTL, it can often result in misclassification of QTL genotype.

Interval mapping

With the availability of high density molecular linkage maps covering entire genomes, some of the problems associated with single point analysis have been overcome with the use of interval mapping (Lander and Botstein, 1989). Interval mapping uses phenotypic and genetic marker information to estimate the probable genotype and the most likely QTL effect at every point in the genome, by means of a maximum-likelihood linkage analysis. Maximum likelihood linkage analysis uses known parameters (such as phenotype) along with unknown, estimated parameters (such as the location of the trait-causing locus) to produce a model that best explains the data being analysed. By using linked markers it is possible to compensate for recombination between the markers and the QTL, increasing the probability of QTL detection. The maximum benefit of interval analysis versus point analysis is found when linked markers are 20-35cM apart, since there are likely to be more crossovers between the marker and the QTL which interval analysis can compensate for.

Analysis of phenotypic extremes

A cost efficient and rapid means of detecting QTL is to perform marker analysis on only those individuals at the extreme tails of a phenotype's spectrum i.e. on those individuals which are phenotypically the most or the least affected by modifying QTL. In a segregating backcross, this analysis would be possible by measuring the trait in each of the backcross progeny and then subdividing the population into two phenotypic extremes. If the allele frequency at any marker locus differs significantly between the two sub-populations, it could be inferred that a QTL controlling the measured trait is located near that marker.

Recently, Taylor *et al.* (1994) introduced a method with the potential to further reduce the effort to map trait-modifying loci, termed phenotypic pooling. Phenotypic pooling refers to the pooling of DNA from the progeny of a segregating population, on the basis of similar phenotype, prior to marker analysis. Using this approach makes it unnecessary to individually type each animal in a phenotypic sub-population with genome-wide markers, since the marker analysis can be carried out using a single sample of pooled DNA, made up of equal amounts of DNA from each animal within a phenotypic sub-class. In the absence of linkage, such an analysis would show a 50:50 ratio of each allele at any given marker. Markers that are linked to a QTL show an enrichment for a particular allele on amplification of the pooled DNA by PCR. This enrichment can be visualised when the PCR product is electrophoresed and one allele-derived band is of greater intensity than the other. With the use of careful controls, it is possible to gauge from this technique which chromosomal regions show linkage to QTL for a single trait. Although this method was originally developed to facilitate the mapping of single locus mutations, findings presented in this thesis support the proposal that this method can also be used in the rapid mapping of quantitative trait loci.

One of the main disadvantages in the analysis of phenotypic extremes, and in the use of phenotypic pooling, is that more segregating individuals must be analysed for the quantitative phenotype, in order to collect enough individuals in the phenotypic extremes. Due to differences in breeding, this is not possible for every organism, but is one of the advantages of using mice for this type of analysis. It is also not possible to use this analysis to map more than one quantitative character, since individuals with extreme phenotypes for a particular trait are not likely to represent the extremes for another.

5.1.3 Gene action

Gene action is a term used to describe the ways in which two alleles at a genetic locus can interact to produce the phenotype of an individual (Tanksley, 1993). In classical genetics, the phenotypes of macromutations' alleles are usually described as being either dominant or recessive. However, phenotypes are seldom absolutely recessive or dominant in loci underlying quantitative variation. The gene action for quantitative trait alleles can range from complete dominance to complete recessiveness, to overdominance, where offspring exceed the phenotype of either parent, to anything in between. In order to describe such continuous gene action, quantitative geneticists have devised a number of statistical parameters. The dominance/additivity statistic (d/a) is commonly used, and describes the degree to which an offspring, genotypically heterozygote for a QTL, resembles the parental homozygotes. Values assigned to the d/a statistic describe whether a QTL has a dominant, recessive or over/underdominant gene action. Additive gene action is used to describe a heterozygote QTL which is phenotypically intermediate between the parental homozygotes, and dominant gene action describes an individual more phenotypically similar to one parent than the other. In this way, similarity of phenotype can be used to infer similarity to parental genotype, with reference to the genotype of a known marker.

5.1.4 Variation in the phenotype of *Sd*

The two phenotypic manifestations of the *Sd* mutation, the shortening in tail length and the urogenital anomalies, are strongly influenced by genetic background (Dunn *et al.*, 1940; Gluecksohn-Schoenheimer, 1943). For example, breeding *Sd* mice onto a BALB/C background increases the expression and dominance of the *Sd* mutation to the point where the affect of *Sd* approaches that of a lethal dominant (Dunn *et al.*, 1940). In a recent study, *Sd*+ mice (on a RSV/Le background) were outcrossed to NMRI and C57Bl/6 mice. When the number of vertebra in each *Sd*+ mouse were counted, those on a NMRI background developed 30-37 vertebrae, whilst those on the C57Bl/6 background developed 37-42, displaying a less severe phenotype than in the NMRI-crossed mice (Dietrich *et al.*, 1993a). In early selection experiments, the two phenotypes were found to vary independently of each other (Dunn, 1942), as was found in this backcross. However, on some genetic backgrounds the severity of the spinal phenotype and has been found to correlate with that of the renal phenotype (Mesrobian and Sulik, 1992).

5.2 Analysis of tail length variation in the *Sd* backcross

As described in chapter 3, tail lengths of all the backcross offspring were measured and recorded (see table 3.3.1). In 24 day old *Sd* progeny, tail lengths varied from 0cm to 4.5cm. In the 24 day old wild type (WT) progeny, tail lengths varied from 5cm to 8.5cm. The distribution of tail length variation for both populations is shown in figure 5.2. This figure shows that the longest *Sd* tail is smaller than the shortest WT tail and, therefore, that the two tail length distributions do not overlap. It was noted at this time that the tail lengths of the *Sd* backcross progeny were, on average, smaller in length than those of their (CAST/Ei X CBA) *Sd* parents (see table 3.3.2) (taking into account the differences in age). This led to the hypothesis that the *M.m.castaneus* genome could be carrying quantitative trait loci which were acting to increase tail length in *Sd* mice.

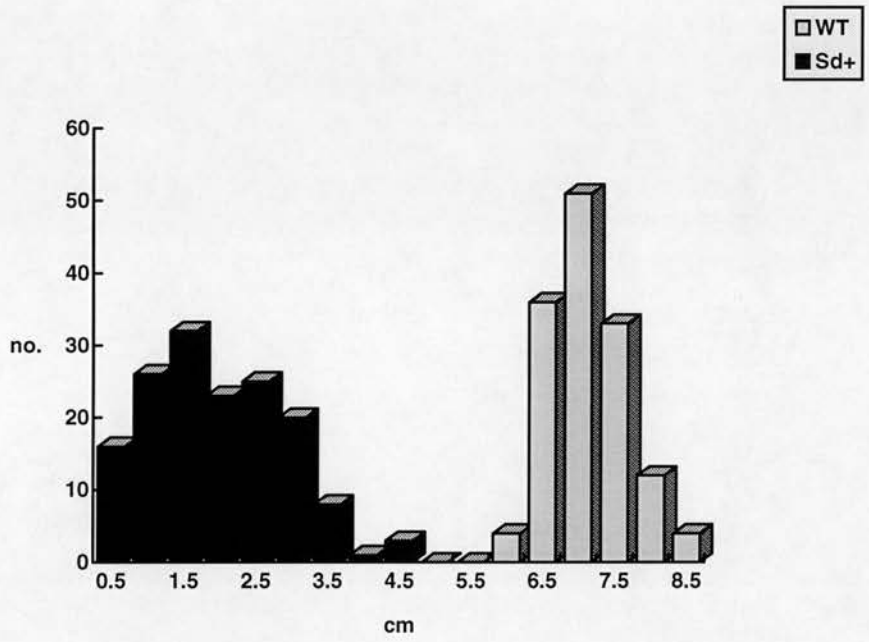

Figure 5.2: Tail length distribution of wild type and *Sd* mice in the backcross.

The number of mice with a given tail length is shown on the Y axis and tail length is shown in centimetres (cm) on the X axis. The dark bars show tail lengths in the *Sd* progeny, and the light bars show tail lengths in the wild type (WT) progeny of the backcross.

In order to ascertain whether the variance of tail length in the *Sd* progeny was significantly different to that of the WT progeny, statistical analysis was carried using the F test (appendix D). This showed that difference in tail length variation between *Sd* and WT mice was highly significant ($p<0.001$).

5.3 Phenotypic pooling of *Sd* backcross progeny

The work in this section of the chapter was carried out in collaboration with S. Phillips and B. Taylor at the Jackson Laboratory, using the phenotypic pooling technique previously described (Taylor *et al.*, 1994).

From each extreme of the tail length distribution, 30 *Sd* mice with the shortest, and 30 with the longest tails were picked for phenotypic pooling. 500ng of DNA from each mouse was used to make up the pools for marker analysis: pool A was made up of DNA from mice with the shortest tails and pool B was made up of DNA from those with the longest. WT mice were not used in this analysis (table 5.3.1).

Based on the hypothesis that the *M.m.castaneus* genome is carrying QTL which increase tail length in *Sd* mice, if a marker is physically linked to a tail-length modifying locus, one would expect to see an excess of the CAST/Ei allele in the long-tail pool and a reduction in the short-tail pool. Likewise, excess of the CBA allele in the short tail pool and a reduction in the long-tail pool would also be expected if this hypothesis were true, since non-linked loci would normally show a 50:50 distribution of alleles. An enrichment for the CAST/Ei allele in the long tail pool at a given marker implies that the majority of long-tailed *Sd* mice at that locus are positive for the trait-modifying genotype, and vice versa for the CBA allele in the short-tail pool. Results showing the enrichment of alleles from the relevant pooled sample would give credence to the hypothesis that the *M.m.castaneus* genome carries modifiers which favour long tails, and would identify chromosomal regions requiring more detailed analysis for the identification of QTL.

Pool A		Pool B	
Mouse No.	Tail length (cm)	Mouse No.	Tail length (cm)
107	0	85	2.8
220	0	322	2.8
121	0.1	3	3.0
204	0.2	40	3.0
276	0.2	76	3.0
21	0.3	95	3.0
82	0.4	130	3.0
244	0.4	142	3.0
245	0.4	161	3.0
36	0.5	194	3.0
61	0.5	209	3.0
158	0.5	210	3.0
270	0.5	211	3.0
302	0.5	224	3.0
320	0.5	240	3.0
313	0.6	263	3.0
115	0.7	292	3.0
8	0.8	297	3.0
37	0.8	7	3.5
318	0.8	103	3.5
81	0.9	135	3.5
203	0.9	145	3.5
258	0.9	188	3.5
4	1.0	189	3.5
56	1.0	193	3.5
86	1.0	298	3.5
128	1.0	214	4.0
237	1.0	9	4.4
299	1.0	155	4.5
2	1.5	205	4.5

Table 5.3.1 *Sd* mice used to generate phenotypic pools A and B

Pool A was generated from *Sd* mice with the smallest tails in the backcross and Pool B was generated from *Sd* mice with the longest tails in the backcross. In Pool A, tail lengths ranged from 0-1cm, and in Pool B tail lengths ranged from 2.8-4.5cm.

5.3.1 Marker analysis of phenotypically pooled DNA

Phenotypic pools, A and B, were used in a genome-wide marker analysis to identify chromosomal regions showing linkage to tail length modifying genes. Sixty five simple sequence repeat (SSR) markers, which can be amplified by PCR, were utilised in this analysis (figure 5.3.1). SSR markers that amplified *M.m.castaneus* and *M.m.musculus* alleles equally well from N1 hybrid DNA, and which gave strong, well-separated bands in agarose gels, were chosen for use in this study. They were also selected such that any autosomal locus would be no more than 20cM from the nearest marker (the MIT Genome Centre map was used to ascertain both the genetic length of the chromosomes and the positions of SSR markers along them). A control sample of N1 hybrid DNA was used in the PCR reactions to ensure that differences in the allele amplification were not artefactual. The results of the marker analysis are given in table 5.3.2 and figure 5.3.1.

From figure 5.3.1 and table 5.3.2 it can be seen that a number of markers show skewing from the 50:50 random distribution of alleles. An indication of the sensitivity of this technique can be seen in the results for chromosome 2, where a deficiency of the CAST/Ei product, in both phenotypic pools, was found in the markers which flank *Sd* (*D2Mit1*, at a distance of 10cM proximal to *Sd*, and *D2Mit7*, approximately 13cM distal to *Sd*). Since *Sd* is present on the *M.m.musculus* allele and is the phenotype that has been selected for in all the mice, one would expect to see a deficiency of the *M.m.castaneus* allele in this region of chromosome 2 if the markers were detecting linkage from the phenotypic pools. This confirmed to us that the pooling technique could detect linkage to a mutant allele over a 20cM distance. However, some regions, such as chromosome 7, proved difficult to analyse due to lack of polymorphism between the mouse strains.

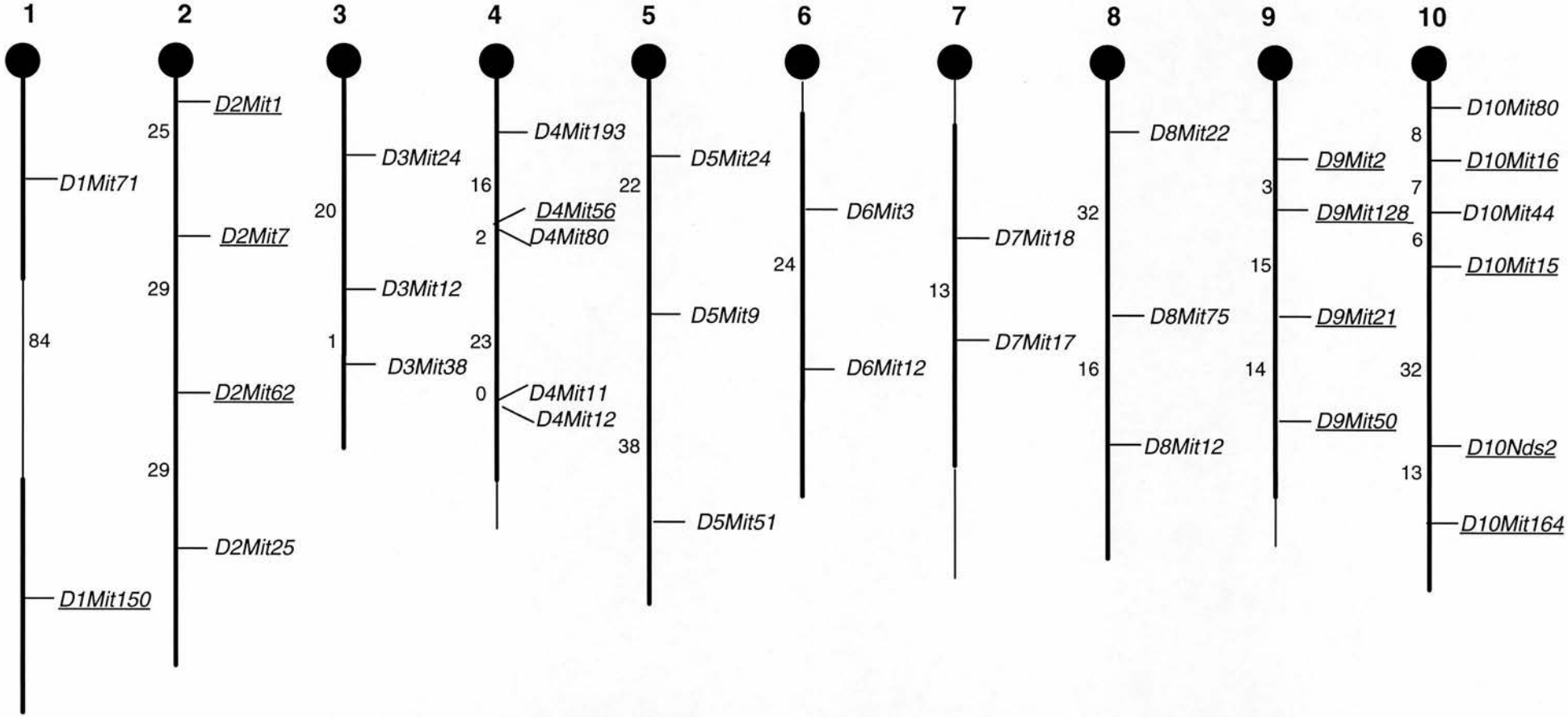

Figure 5.3.1 The chromosomal location of markers selected for pooled-SSR analysis, mouse chromosomes 1-10.
The marker map was supplied by S. Phillips and B. Taylor (personal communication), marker positions were obtained from the microsatellite map distributed by the MIT Genome Centre. Thickened lines indicate regions within 20cM of a SSR marker. Numbers to the left of each chromosome (to the right for chr.1) indicate the genetic distance between markers in centiMorgans. Underlined markers are those which showed linkage, by allele enrichment, in one or both of the phenotypic pools.

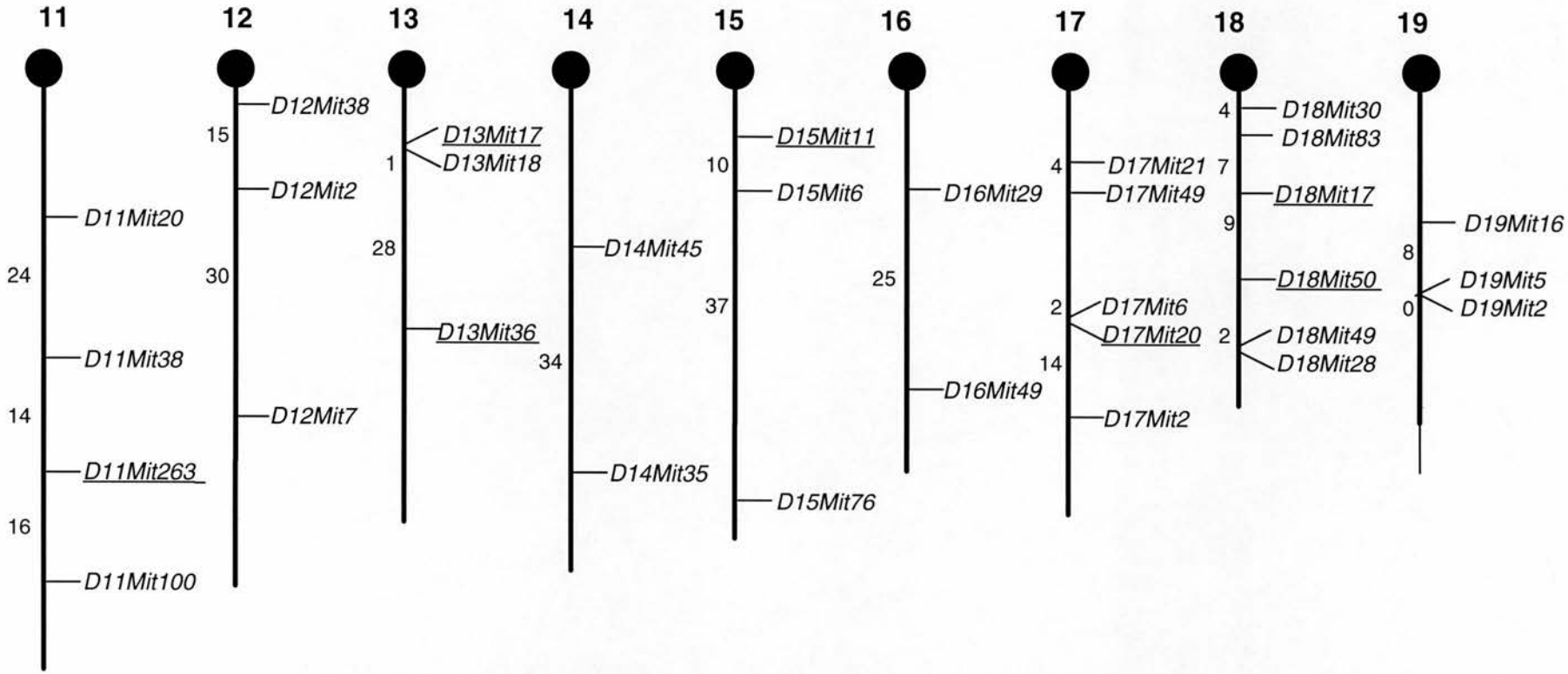

Figure 5.3.1 (continued) The chromosomal location of markers selected for pooled-SSR analysis, mouse chromosomes 11-19
The marker map was supplied by S. Phillips and B. Taylor (personal communication), marker positions were obtained from the microsatellite map distributed by the MIT Genome Centre. Thickened lines indicate regions within 20cM of a SSR marker. Numbers to the left of each chromosome indicate the genetic distance between markers in centiMorgans. Underlined markers are those which showed linkage, by allele enrichment, in one or both of the phenotypic pools.

MARKER	Size of CBA allele	Size of CAST/Ei allele	RESULTS Pool A = short tail pool, Pool B = long tail pool
D1Mit71	154	184	A = B = N1
D1Mit150	126	154	Less CAST and slightly more CBA in pool A
D2Mit1	120	140	No CAST present in either pool
D2Mit7	214	256	Almost no CAST in pool A or pool B
D2Mit62	146	160	More CBA in pool B, pool A gave same results as N1 DNA control
D2Mit25	116	138	A = B = N1
D3Mit24	130	168	A = B = N1
D3Mit12	154	118	N/D
D3Mit38	120	142	A = B = N1
D4Mit193	129	161	A = B = N1
D4Mit56	154	134	Less CAST in pool B, more CAST in pool A
D4Mit80	158	136	A = B = N1
D4Mit11	144	170	A = B = N1
D4Mit12	168	190	A = B = N1
D5Mit4	144	164	Slightly more CAST in pool A
D5Mit9	268	301	A = B = N1
D5Mit51	142	166	A = B = N1
D6Mit3	308	263	A = B = N1
D6Mit12	123	147	A = B = N1
D7Mit18	120	109	A = B = N1
D7Mit17	162	144	A = B = N1
D8Mit22	152	142	A = B = N1
D8Mit75	134	170	A = B = N1

Table 5.3.2 Results of SSR genome-wide marker analysis of phenotypically pooled DNA

MARKER	Size of CBA allele	Size of CAST/Ei allele	RESULTS (Pool A = short tail pool, Pool B = long tail pool)
D8Mit12	118	126	A = B = N1
D9Mit2	182	159	More CBA in pool B, less CAST in pool A, reciprocal skew
D9Mit128	143	157	N/D
D9Mit21	185	205	More CBA in pool B, less CAST in pool A, reciprocal skew
D9Mit50	146	134	Less CAST in pool A, results for pool B as for N1 DNA control
D10Mit80	164	158	A = B = N1
D10Mit16	152	134	More CAST, less CBA pool A; long pool almost same as N1 control
D10Mit44	160	172	A = B = N1
D10Mit15	180	135	Less CAST in pool A and slightly more in pool B
D10Nds2	142	123	Less CAST pool A, pool B results as for N1 control
D10Mit164	130	108	Less CAST pool A, pool B results as for N1 control
D11Mit20	150	126	A = B = N1
D11Mit38	112	128	A = B = N1
D11Mit263	119	130	Less CAST in pool A, pool B results as for N1 control
D11Mit100	158	170	A = B = N1
D12Mit38	148	172	A = B = N1
D12Mit2	132	178	A = B = N1
D12Mit7	123	130	A = B = N1
D13Mit17	154	136	Slightly less CAST pool A, pool B results as for N1 control
D13Mit18	190	214	A = B = N1
D14Mit45	106	140	A = B = N1
D14Mit35	222	244	A = B = N1
D15Mit11	106	126	Slightly more CBA in pool B

Table 5.3.2 Results of SSR genome-wide marker analysis of phenotypically pooled DNA (continued)

MARKER	Size of CBA allele	Size of CAST/Ei allele	RESULTS (Pool A = short tail pool, Pool B = long tail pool)
D15Mit11	106	126	Slightly more CBA pool B
D15Mit6	128	102	A = B = N1
D15Mit76	254	222	A = B = N1
D16Mit29	148	172	A = B = N1
D16Mit49	222	218	A = B = N1
D17Mit21	122	105	A = B = N1
D17Mit49	224	270	A = B = N1
D17Mit6	192	192	N/D
D17Mit20	170	192	Slightly more CAST pool A, slightly more CBA pool B, slight skew
D17Mit2	230	250	A = B = N1
D18Mit30	152	176	A = B = N1
D18Mit83	115	135	A = B = N1
D18Mit17	188	201	Slightly more CAST pool A, pool B results as for N1 control
D18Mit50	162	175	Less CAST pool A, pool B results as for N1 control
D18Mit49	113	99	N/D
D18Mit128	134	120	A = B = N1
D19Mit16	118	146	A = B = N1
D19Mit5	213	193	A = B = N1
D19Mit2	185	163	A = B = N1

Table 5.3.2 Results of SSR genome-wide marker analysis of phenotypically pooled DNA (continued)
The results section of the table indicates which markers showed enrichment for the CAST/Ei-sized allele (CAST) or the CBA-sized allele (CBA) in either phenotypic pool. A = B = N1 has been used to indicate where allele amplification from either pool was the same as that from the N1 control. N/D = not done. Allele sizes are given in basepairs.

Where isolated regions displayed allele enrichment in only one of the phenotypic pools, evidence of linkage could not be assumed since these results could not be explained by the hypothesis and could be due to chance. An example of this can be seen at the marker *D4Mit56*, where less CAST/Ei product was seen in the long-tail pool and more in the short-tail pool. Because the remaining markers on chromosome 4 showed no skewing in allele amplification, it is unlikely that the result for *D4Mit56* is significant. Stronger proof of linkage arises where evidence of allele enrichment can be seen with 2 or more markers across a chromosomal region. Results for chromosomes 9, 10 and 18 provide the strongest proof of linkage to tail-length modifying genes, since allele enrichment in one or both pools has been detected with more than one marker. In chromosome 9, linkage to all 4 markers in either one or both phenotypic pools was seen (table 5.3.2), and in chromosome 10, only *D10Mit44* and *D10Nds2* failed to show linkage out of the 6 markers analysed (table 5.3.2). In chromosome 18, evidence of linkage is less well founded, since only 2 markers, *D18Mit17* and *D18Mit50* showed distortion in the distribution of CAST/Ei and CBA alleles.

5.3.2 Discussion

This technique, designed for the detection of single mutant loci, allowed for the rapid and efficient identification of three regions which could be positive for tail-length modifying loci on chromosomes 9, 10 and 18. To confirm these findings, it was necessary to singly haplotype the individual mice in each phenotypic pool with a panel of closely spaced markers on each chromosome. This was done to increase evidence of linkage and to identify markers more closely linked to putative QTL for use in the analysis of the entire backcross.

5.4 Chromosome 10 marker analysis

In order to analyse more closely the evidence of linkage on chromosome 10, a panel of 8 SSR markers spaced between 10-20cM apart (table 5.4.1) were selected and used to individually haplotype by PCR, the *Sd* progeny comprising the two phenotypic pools. The results of this analysis are shown in tables 5.4.2 (short-tail pool) and 5.4.3 (long-tail pool). The inclusion of the marker nodal in these tables is discussed later in the chapter.

Marker	Distance from centromere (cM)	CBA-sized allele (bp)	CAST/Ei-sized allele (bp)
D10Mit80	2.2	144	172
D10Mit105	13.4	145	182
D10Mit31	18.4	154	134
D10Mit21	39.4	132	160
D10Mit121	60.7	190	217
D10Mit164	68.5	130	108
D10Mit145	73.0	141	101

Table 5.4.1 Chromosome 10 markers used for the haplotyping of the long and short-tailed *Sd* mice from the backcross.

The distance of each marker from the top of chromosome 10 is given in centiMorgans (cM) and was obtained from the microsatellite map distributed by the MIT Genome Centre. Allele sizes are given in basepairs (bp).

No.	Sex	Tail length (cm)	*D10 Mit 80*	*D10 Mit 105*	nodal	*D10 Mit 31*	*D10 Mit 21*	*D10 Mit 121*	*D10 Mit 164*	*D10 Mit 145*
107	F	0	M	M	M	C	C	C	C	C
220	F	0	M	M	M	M	M	M	C	C
121	M	0.1	M	M	M	M	M	M	M	M
204	F	0.2	C	C	C	C	C	M	M	M
276	M	0.2	M	M	M	M	M	C	C	C
21	F	0.3	M	M	M	M	C	C	C	C
82	M	0.4	C	C	M	M	M	C	C	C
244	M	0.4	M	M	M	M	M	C	C	C
245	M	0.4	M	M	M	M	M	M	M	C
36	M	0.5	M	M	M	M	M	M	M	M
61	F	0.5	C	C	C	C	C	C	C	M
158	F	0.5	M	M	M	M	M	M	M	M
270	F	0.5	M	M	M	M	M	M	C	C
302	M	0.5	C	C	C	C	C	C	C	C
320	M	0.5	M	M	M	M	M	M	M	M
313	M	0.6	C	C	C	M	M	M	M	M
115	F	0.7	M	M	M	M	M	M	M	M
8	F	0.8	M	M	M	M	M	M	M	M
37	M	0.8	M	M	M	M	M	M	M	M
318	M	0.8	C	C	C	C	C	M	M	M
81	M	0.9	M	M	M	M	M	M	M	M
203	F	0.9	M	M	M	M	M	C	C	C
258	M	0.9	M	M	M	M	M	M	M	M
4	M	1.0	M	M	M	M	M	M	M	M
56	M	1.0	C	C	C	C	C	C	C	C
86	F	1.0	C	C	C	C	C	M	M	M
128	M	1.0	M	M	M	M	M	M	M	M
237	M	1.0	C	C	C	C	C	C	C	M
299	F	1.0	M	M	M	M	M	M	M	M
2	M	1.5	M	M	M	M	M	C	C	C
% CBA allele			70%	70%	73%	73%	67%	63%	56%	60%

Table 5.4.2 Chromosome 10 haplotyping of mice in the short-tail pool
F stands for female and M for male in column 2. Tail lengths are in centimetres (cm). Mice haplotyped as C were heterozygous for the CAST/Ei and CBA alleles. Mice haplotyped as M were homozygous for the CBA allele. The final row indicates the percentage of mice at each marker which were genotyped as being homozygous for the CBA allele (M). Markers are given in a centromeric (left) to telomeric (right) order. The map positions of the MIT markers were obtained from the MIT Genome Centre map. The relative position of nodal was obtained from the Jackson Laboratory's Encyclopaedia of the Mouse.

No	Sex	Tail length (cm)	*D10 Mit 80*	*D10 Mit 105*	nodal	*D10 Mit 31*	*D10 Mit 21*	*D10 Mit 121*	*D10 Mit 164*	*D10 Mit 145*
85	F	2.8	C	C	C	M	M	M	M	M
322	F	2.8	C	C	C	C	C	C	C	C
3	F	3.0	C	C	C	C	C	C	C	C
40	F	3.0	C	C	C	C	M	M	M	M
76	M	3.0	C	C	C	C	C	C	C	C
95	F	3.0	C	C	C	M	M	M	M	M
130	M	3.0	C	C	C	C	C	C	C	C
142	M	3.0	M	C	C	C	C	C	C	C
161	M	3.0	C	C	C	M	M	M	M	M
194	M	3.0	M	M	M	M	M	M	M	M
209	M	3.0	M	C	C	C	C	C	C	C
210	M	3.0	C	C	C	C	C	M	M	M
211	M	3.0	C	C	C	C	C	C	C	C
224	F	3.0	C	C	M	M	M	C	C	C
240	F	3.0	C	C	M	M	M	M	M	M
263	F	3.0	C	C	M	M	M	M	M	M
292	F	3.0	C	C	C	C	C	C	C	C
297	F	3.0	M	C	C	C	C	C	C	C
7	F	3.5	M	M	M	M	M	M	M	M
103	M	3.5	C	C	C	C	C	M	M	M
135	F	3.5	M	C	C	C	C	C	C	C
145	F	3.5	C	C	C	C	C	C	C	C
188	F	3.5	M	M	M	M	M	M	M	M
189	F	3.5	C	C	C	C	C	M	M	M
193	M	3.5	M	C	C	C	C	C	C	M
298	F	3.5	M	C	C	C	C	C	C	C
214	M	4.0	C	C	C	C	C	C	C	C
9	F	4.4	C	C	C	C	C	C	C	C
155	F	4.5	C	C	C	C	C	C	C	C
205	F	4.5	C	C	C	C	C	C	C	C
% CAST/Ei allele			66%	87%	80%	70%	66%	60%	60%	56%

Table 5.4.3 Chromosome 10 haplotyping of mice in the long-tail pool
F stands for female and M for male in column 2. Tail lengths are in centimetres (cm). Mice haplotyped as C were heterozygous for the CAST/Ei and CBA alleles. Mice haplotyped as M were homozygous for the CBA allele. The final row indicates the percentage of mice at each marker which were genotyped as being heterozygous for the CAST/Ei and CBA alleles (C). Markers are given in a centromeric (left) to telomeric (right) order. The map positions of the MIT markers were obtained from the MIT Genome Centre map. The relative position of nodal was obtained from the Jackson Laboratory's Encyclopaedia of the Mouse.

From table 5.4.2, it can be seen that over 50% of the mice in the short-tail pool are genotypically CBA when haplotyped using the chromosome 10 markers. *D10Mit80* and *D10Mit105* show the CBA genotype in 70% of the mice, with the skew in allele distribution peaking at 73% at *D10Mit31*. The number of mice with the CBA genotype decreases to 67% at *D10Mit21*. The incidence of the CBA allele at markers more telomeric (at *D10Mit121, D10Mit164* and *D10Mit145*) ceases to follow a genotypic pattern indicative of linkage to a QTL. From these preliminary results, it is possible to infer that the markers *D10Mit80* to *D10Mit31* define a region of chromosome 10 which is in linkage disequilibrium with tail-length modifying QTL.

The chromosome 10 haplotyping results from the mice in the long-tail pool (table 5.4.3) shows that linkage to the CAST/Ei allele increases from 66% at *D10Mit80* to a peak at *D10Mit105*, where 87% of the mice are of the CAST/Ei genotype. Linkage reduces slightly to 70% at *D10Mit31*, and then the incidence of the CAST/Ei allele decreases the further down the chromosome the markers are situated, with *D10Mit21*, *D10Mit121, D10Mit164* and *D10Mit145* showing similar slight skews in allele distribution as those seen in the short-tail pool (tables 5.4.2 and 5.4.3). These results support the data produced from marker analysis of the short-tail pool, suggesting the presence of a tail-length modifying QTL between the markers *D10Mit80* and *D10Mit31* on chromosome 10.

Using the chromosome 10 mapping data produced from the long and short-tailed mice, the search for genes mapping in the region of *D10Mit105* to *D10Mit31* identified nodal as a potential candidate gene on the basis of its map position and biological function. Nodal belongs to the TGFβ superfamily of genes and encodes a signalling molecule thought to be essential for mesoderm formation and the organisation of axial structures in early mouse development (Zhou *et al.*, 1993). Mapping data obtained from the Jackson Laboratory's Encyclopaedia of the Mouse positions nodal 14.5cM distal to

D10Mit105 and 6.5cM proximal to *D10Mit31*. Since nodal maps to the appropriate region of chromosome 10, and has a biological function which could be used to explain the variation in the expressivity of *Sd*, primers were designed (table 2.10.4) to amplify over an intron, the size of which varied between *M.m.musculus* and *M.m.castaneus* DNA. Using this variant, mice in the long and short-tail pools were genotyped at nodal (see tables 5.4.2 and 5.4.3).

The results of mapping nodal onto the panel of long and short-tailed mice are shown in tables 5.4.2 and 5.4.3. In the long-tail pool (table 5.4.3), nodal was genotyped as being CAST/Ei in 80% of the mice. This shows a decrease in the number of mice of the CAST/Ei genotype between *D10Mit105* and nodal. In the short-tail pool, nodal was genotyped as being CBA in 73% of the mice (table 5.4.2), showing an increase in the number of mice of this genotype between *D10Mit105* and nodal. These results show that nodal maps within the region of linkage disequilibrium defined by the markers *D10Mit105* and *D10Mit31*. Data from the long-tail pool, showing a decrease in the number of mice with the CAST/Ei genotype between *D10Mit105* and nodal, could be due to recombination occurring between the marker and QTL. However, it was not possible to statistically confirm the presence of a tail-length modifying gene on this region of chromosome 10 by using data from such a small number of mice.

5.5 Chromosome 10 backcross analysis

In order to have sufficient data to statistically identify a tail-length modifying locus on chromosome 10, backcross progeny were haplotyped with the markers *D10Mit80, D10Mit105, D10Mit31* and nodal. Only those mice which were 24 days old when assayed were used in this analysis (300 out of the 336 backcross progeny). Figure 5.5.1 shows the results of chromosome 10 haplotyping of the backcross.

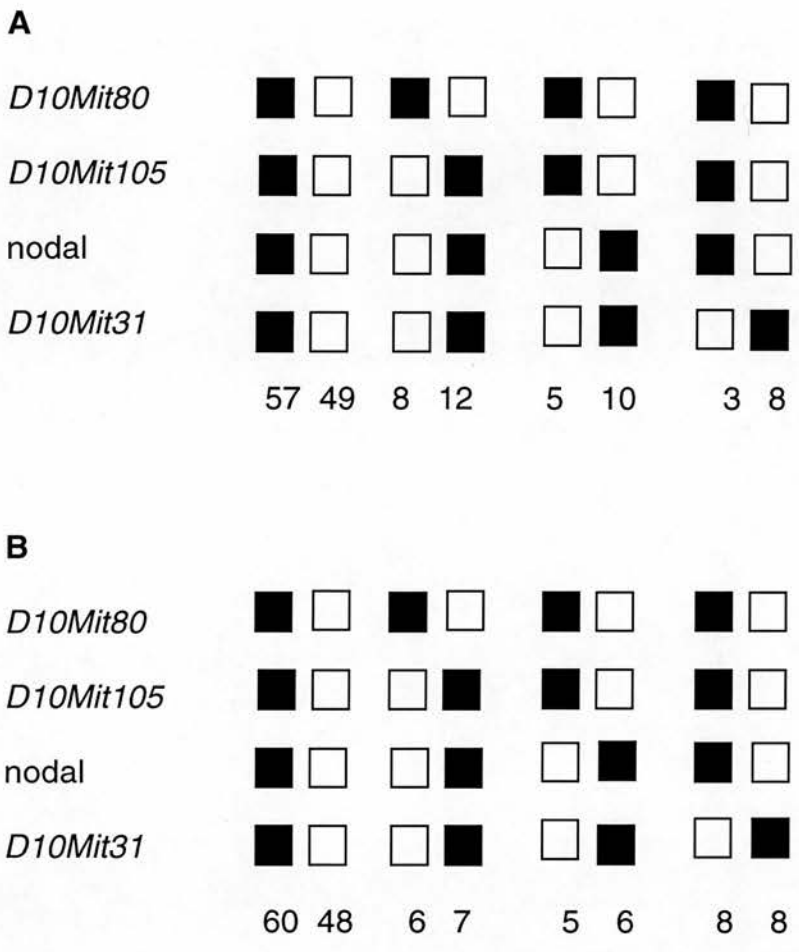

Figure 5.5.1 Chromosome 10 haplotyping of the *Sd* backcross

152 *Sd* mice (figure **A**) and 148 wild type (figure **B**) were haplotyped. The number at the base of each column indicates the number of animals recombinant between the markers shown. The filled in box (■) denotes the CAST/Ei allele and the unfilled box (□) denotes the CBA allele.

5.5.1 Statistical analysis of chromosome 10 mapping data

The statistical analysis in this section of the thesis was carried out in collaboration with K. Rance and P. Keightley at the Institute of Cell and Animal Population Biology, University of Edinburgh.

The statistical method used to analyse the chromosome 10 mapping data was based on that of Haley and Knott (1992), where a regression analysis is used in place of a maximum likelihood method (the two methods have been found to give comparable results). A statistical programme based on this method which took into account variation arising from sex, parity, litter number and litter size was devised. Sex and parity were found to be insignificant in the analysis of variance and were therefore excluded from the regression analysis. The chromosome 10 mapping results from backcross *Sd* and wild type mice were analysed as two separate data sets.

In the regression analysis employed, the putative QTL is positioned at 1cM intervals along the chromosome and the statistical model is fitted at each point. The test statistic is calculated in $\log_{10}$ and is comparable with a LOD score (Haley and Knott, 1992). The likelihood profiles which are generated show a peak in the test statistic where a QTL is most likely to be located on the chromosome. A drop in the test statistic of more than 2 defines a 95% confidence interval for the presence of a QTL in that region.

Figures 5.5.2 and 5.5.3 show the likelihood profiles generated from the chromosome 10 mapping data of the *Sd* and WT progeny of the backcross (respectively). Figure 5.5.3 shows that there is no evidence of linkage to a tail-length modifying gene in the WT progeny, as the test statistic fails to reach over a value of 2 at most positions along the chromosome. There is also no peak in the test statistic to indicate linkage to a modifying gene over a particular region of chromosome 10. Conversely, figure 5.3.2 shows that the test statistic peaks significantly at a number of points along the

chromosome in the likelihood profile of the *Sd* backcross progeny. The test statistic peaks at a value of 13.37, equivalent to a LOD score of the same value, 12cM distal to *D10Mit80*. This peak defines the region of maximum linkage to a tail length modifying gene on chromosome 10. Either side of this peak, between 7.5 to 22cM away from *D10Mit80*, the 95% confidence interval spans a 14.5cM stretch of chromosome 10. This model suggests, therefore, that a tail length modifying gene in *Sd* progeny could be found within this region. Since *D10Mit105* maps 9.2cM (+/-2.3) distal to *D10Mit80* in the *Sd+* population, it maps within this region, close to the peak of maximum linkage. However, the candidate tail-length modifying gene nodal, maps 23cM (+/-3.8) from *D10Mit80*, and is outwith the region of chromosome 10 defined by the 95% confidence interval. Nodal is unlikely, therefore, to be responsible for tail length variation in the *Sd* progeny of the backcross. Nevertheless, this analysis has shown that a tail-length modifying gene is most likely to be present on the region of chromosome 10 between the markers *D10Mit80* and *D10Mit105*. Calculating the percentage of the mean (appendix E), indicates that this locus could be responsible for at least 21.2% of the genetic variance affecting tail lengths in *Sd* mice.

5.5.2 Discussion

These data indicate that a tail length modifying locus is present on chromosome 10, approximately 12cM distal to the marker *D10Mit80*, and in the vicinity of *D10Mit105*. The location of a tail length modifying gene to this region of chromosome 10 excludes the gene nodal as a candidate modifier. There are other genes in this region of the chromosome, such as the interferon gamma receptor gene (Kozak *et al.*, 1990), which are not as strong candidates as nodal, but which may warrant further investigation. It is likely that this gene is not the only locus affecting tail length in the presence of *Sd*, since markers on chromosomes 9 and 18 also showed linkage to the variable phenotype, and contribution of this locus to overall tail length variation may be only 21.2%.

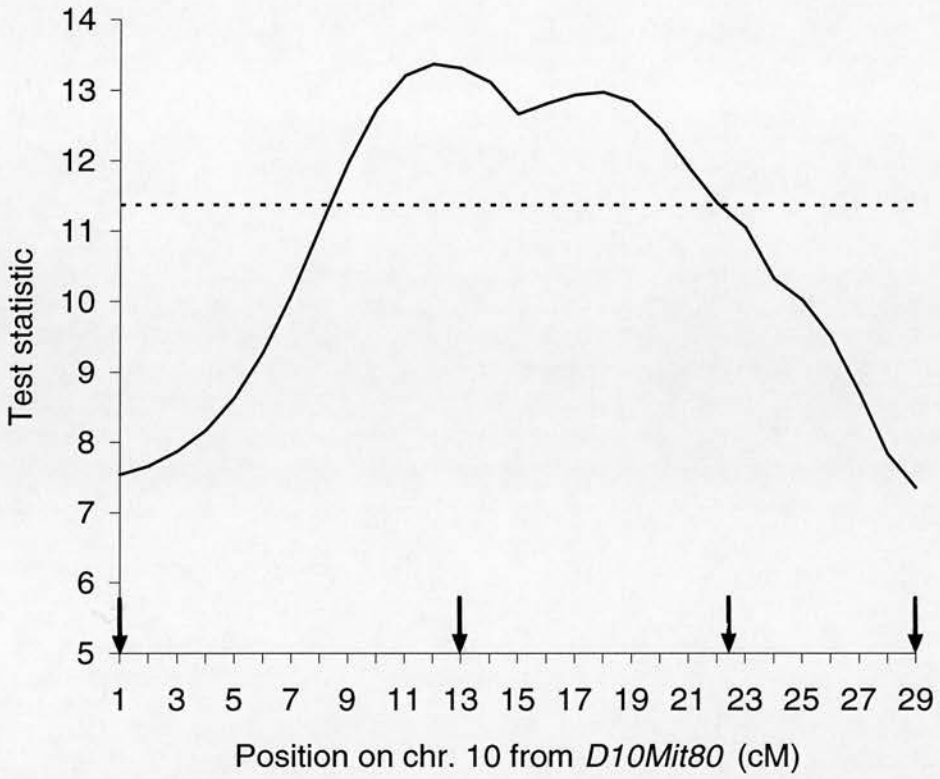

Figure 5.5.2 Regression analysis of chromosome 10 mapping data (*Sd* progeny)

The likelihood profile was produced using regression methods to fit a model for a single QTL at 1cM intervals along the chromosome (solid line). The height of the curve (X axis) was calculated in $\log_{10}$. Distance in centiMorgans (cM) is shown on the Y axis. The 95% confidence interval, defined by a drop of 2 in the test statistic, is indicated by the broken line. With markers placed in a telomeric (left) to centromeric (right) order: *D10Mit80* is positioned where the X and Y axes intersect; *D10Mit105* is positioned 13.2cM (+/-2.7) from *D10Mit80*, nodal is 9.9cM (+/-2.4) from *D10Mit105*, and *D10Mit31* is 7.2cM (+/-2.1) from nodal. Arrows indicate the approximate position of these markers. Map distances were calculated from the *Sd* progeny of the backcross (see figure 5.5.1), and cM and standard error values as in section 2.12. The data used to generate this likelihood profile is detailed in appendix A.

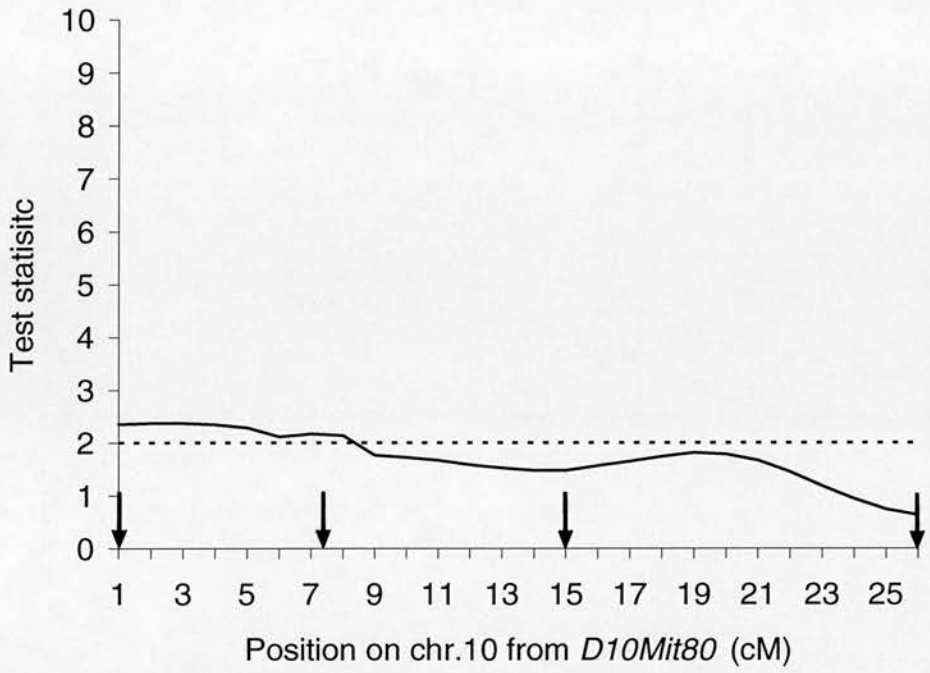

Figure 5.5.3 Regression analysis of chromosome 10 mapping data (WT progeny)

The likelihood profile was produced as in figure 5.3.2 using mapping data from the wild type (WT) progeny of the backcross (solid line). The height of the curve (X axis) was calculated in $\log_{10}$. Distance in centiMorgans (cM) is shown on the Y axis. The 95% confidence interval is indicated by the broken line. With markers placed in a telomeric (left) to centromeric (right) order: *D10Mit80* is positioned where the X and Y axes intersect; *D10Mit105* is positioned 7.4 (+/-2.2) from *D10Mit80*, nodal is 7.4 (+/-2.2) from *D10Mit105*, and *D10Mit31* is 10.8 (+/-2.5) from nodal. Arrows indicate the approximate position of these markers. Map distances were calculated from the WT progeny of the backcross (see figure 5.5.1), and cM and standard error values as in section 2.12. The data used to generate this likelihood profile is detailed in appendix B.

Chapter 6: *In vitro* development of embryonic kidneys from *Sd* mice

6.1 Introduction

There are three spatially and temporally discrete embryonic excretory organs: the pronephros, the mesonephros and the metanephros (Saxen, 1987). The pronephros is the most primitive of these organs and is functional in embryogenesis only in some of the lower vertebrates. The mesonephros, located caudal to the pronephros, is functional during embryogenesis and throughout life in higher fishes and amphibians. The metanephros, located most caudally, serves as the permanent kidney of the amniotes. The pronephros, mesonephros and metanephros consist of paired organs derived from the intermediate mesoderm located immediately lateral to the somites. Pronephric tubules empty into the primary nephric ducts, which are adopted by the mesonephros after degeneration of the pronephric tubules to become the mesonephric ducts. The mesonephric ducts empty into the cloaca and, at a later stage in development, are induced by the metanephric mesenchyme to develop into outgrowths called the ureteric buds. The ureteric buds then invade the metanephric blastema to set in motion the cascade of inductive interactions which give rise to the mature metanephric kidney.

6.1.1 kidney organogenesis

Differentiation of the metanephric blastema and of the ureteric bud is dependent on an inductive event (or events) that takes place when the ureteric bud encounters the metanephric blastema. The exact nature of this event is unknown, but the induction between the ureteric bud and metanephric blastema is known to be reciprocal (Grobstein, 1955).

On invasion of the metanephric mesenchyme by the ureteric bud, the bud induces the metanephric mesenchyme to undergo nephrogenic differentiation. For this, mesenchymal stem cells condense into small aggregates whilst others differentiate into stromal cells which remain located in the medulla. Each aggregate undergoes a mesenchymal to epithelial

conversion and a sequence of histologically definable growth stages. The cells in the condensed aggregate proliferate as part of a mitotic stage, after which mitotic activity decreases and the condensate assumes a comma shape. Cells farthest from the tip of a ureteric branch become elongated and funnel shaped, resulting in the formation of a slit which develops into the glomerular crevice. While glomerulogenesis is taking place, a second slit forms at the distal pole of the developing nephron giving rise to an S-shaped body (Saxen, 1987). This pole of the developing nephron fuses with the tip of a collecting duct and forms the distal tubule of the mature nephron, whilst the proximal pole gives rise to the glomerulus and proximal tubule. The metanephric blastema differentiates into all of the tubular structures of the adult nephron with the exception of the collecting duct system.

On encountering the metanephric mesenchyme, the ureteric bud is induced to grow and bifurcate. In humans and mice, the first branching of the bud in the mesenchymal blastema is dichotomous and symmetric, creating two branches that grow in opposite directions at an angle of 180° (Saxen, 1987). As each new ureteric branch is capped by a mesenchymal condensate, serial, dichotomous branching is induced giving rise to a network of branches which form the collecting duct system of the metanephros. Whilst the distal tips of the ureteric bud divide and grow to the periphery of the metanephric mesenchyme, the proximal ends give rise to the ureter and renal pelvis.

6.1.2 Molecular and genetic events in kidney induction

Although the actual signalling molecules responsible for induction remain elusive, a number of growth factors and receptors are known to be expressed in the kidney at the time of induction; and although the nature of the primary mesenchymal induction signal is still unknown, an understanding of the molecular mechanisms underlying the induction of the ureteric bud may be closer at hand.

The proto-oncogene c-*ret* is a receptor tyrosine kinase expressed in the developing kidney at the tips of the branching ureteric bud (Pachnis *et al.*, 1993). Mice with a truncated form of Ret protein which fails in its biological function, are born with multiple defects of the urogenital system (Schuchardt *et al.*, 1994). Phenotypes in these mice range from renal agenesis to small, severely dysplastic kidney rudiments, abnormalities which occur primarily because the ureteric bud fails to proliferate and branch. Surrounding the tip of the growth retarded ureteric buds, nephrogenic tissue composed of glomeruli and tubules is found, indicating that the differentiation potential of the mesenchyme is not affected by the loss of Ret. The c-*ret* gene product may therefore be responsible for transducing the signals from the mesenchyme that induce growth and branching of the ureteric bud. Since an endogenous ligand for c-*ret* has yet to be identified, the nature of the mesenchymal-derived signal remains unknown.

The expression patterns of the proto-oncogene c-*met* and its ligand, hepatocyte growth factor (HGF) are indicative of their involvement in kidney induction and branching morphogenesis (Sonnenberg *et al.*, 1993). In the developing kidney, c-met is expressed and transcribed in both the ureter epithelial and mesenchymal cells. HGF is synthesised by the surrounding mesenchyme and is known to stimulate branching and tubulogenic morphogenesis in Madine-Darby canine kidney (MDCK) cells (Montesano *et al.*, 1991; Santos and Nigram, 1993). Co-culturing mouse 11.5d kidney rudiments with MDCK cells seeded in Type I collagen, induces the cells to undergo a branching morphogenesis which can be partially inhibited by monospecific antiserum against HGF. This suggests that the action of soluble branching morphogens released by the embryonic kidney can be partly blocked by inhibiting the binding of HGF. Inhibition of the HGF/ c-met interaction with anti-HGF antibodies, also reduces the growth of kidney rudiments in culture (Santos *et al.*, 1994). Nevertheless, mice lacking HGF fail to survive due to placental and liver defects (Schmidt *et al.*, 1995; Uehara

et al., 1995). Abnormalities in the development of excretory system were not reported, suggesting that HGF is probably not the only branching morphogen in the developing kidney.

Mice lacking the Wilms' tumour suppresser gene, *WT1*, have recently been generated (Kreidberg *et al.*, 1993). This gene was initially identified by a loss of heterozygosity in human embryonic kidney tumours and is expressed in a variety of cell types in the urogenital system (Armstrong *et al.*, 1992). The *WT1* knockout mouse shows a complete absence of ureteric bud outgrowth despite the fact that *WT1* is normally expressed in the uninduced mesenchyme, and not in the ureter epithelium or nephric duct. The *WT1* mutant mesenchyme is unable to respond to induction when co-cultured with wild type inducer, such as spinal cord, but instead undergoes apoptosis (Kreidberg *et al.*, 1993). This suggests that *WT1* is important in early kidney development, and is required for both the primary induction signal of the ureteric bud and in the differentiation events of the metanephric blastema.

WT1 is likely to play an important role in the regulation of *Pax2*, a transcription factor which is activated during early kidney morphogenesis and down regulated as the tubular epithelium matures (Dressler and Douglass, 1992). *Pax2* is expressed in condensing metanephric mesenchyme and in early epithelial structures derived from the mesenchyme. Repression of *Pax2* by antisense oligonucleotides inhibits condensation of the mesenchyme and its subsequent conversion to epithelium (Rothenpieler and Dressler, 1993); whilst persistent and ectopic expression of this gene in transgenic mice, results in renal abnormalities similar to those found in human nephrotic syndromes (Dressler *et al.*, 1993). The properly timed activation and down-regulation of *Pax2* is therefore essential for normal kidney development. Persistent expression of *Pax2* in the epithelial components of human Wilms' tumours (Dressler and Douglass, 1992) suggested that negative regulators of the gene were absent in these tissues. Expression analysis of WT1

revealed that its protein levels were markedly increased in individual precursor cells of the glomerular epithelium during periods of *Pax2* down-regulation in these cells. Binding studies and CAT assays demonstrated that WT1 is able to bind to three high affinity sites in the 5' untranslated region of *Pax2* to repress its transcription (Ryan *et al.*, 1995). These data indicate that *WT1* acts upstream of *Pax2* in the metanephric induction pathway.

Evidence exists to suggest that the signalling glycoprotein Wnt-4 is essential for normal kidney development (Stark *et al.*, 1994). *Wnt-4* expression is first seen in 11.5 day embryos where the first pre-tubular condensates form. It is later expressed in the comma shaped bodies, and is then restricted to the region in the S-shaped body where fusion of the tubular epithelium to the collecting duct occurs. Once this fusion is complete, *Wnt-4* expression is lost. Mice lacking *Wnt-4* activity develop severely growth retarded kidneys with morphologically undifferentiated mesenchyme. Pretubular aggregates were absent in the majority of 15 day embryos examined, although ureteric branching was found to be unaffected (Stark *et al.*, 1994). When the expression of kidney markers in *Wnt-4* null mice was examined, the initial expression of *WT1* in undifferentiated mesenchyme was found to be unchanged, indicating that the early activation of *WT1* remains unaffected by the loss of *Wnt-4*. *Pax2* and *N-myc* expression patterns were also normal to day 14.5 of gestation, but in embryonic rudiments where the mesenchyme had failed to undergo tubulogenesis, the expression of these genes was absent. These data suggested that the mesenchymal cells fated to form tubules were induced but failed to convert to epithelial structures in the absence of *Wnt-4*. *Wnt-4* was also found to be necessary for the activation and/or maintenance of *Pax8* expression. *Pax 8* is a transcription factor which is expressed in condensing metanephric mesenchyme and which could be involved in the regulation of other genes essential for tubule formation (Plachov *et al.*, 1990). It is likely therefore that the development of the ureteric bud, from 12.5 days of development, is independent of both *Wnt-4*

expression and tubule formation, as supported by the normal expression of c-*ret* and c-*ros*, genes which are markers for the differentiation and development of the ureteric bud. These data indicate that although the early expression of *WT1* and *Pax2* is induced in the absence of *Wnt-4*, the expression of this gene is vital for the mesenchymal to epithelial conversion of the metanephric mesenchyme.

There are many other genes known to be active in kidney morphogenesis. Cell proliferation and differentiation are intimately linked to ensure that the correct numbers of cells and cell types develop accurately. The insulin-like growth factors (IGFs) and their receptors (IGFrs) mediate proliferation and morphogenesis of the kidney epithelium (Patterson and Dressler, 1994). IGF-2 is expressed in undifferentiated metanephric mesenchyme and human Wilms' tumour (Yun *et al.*, 1993) and its expression is possibly controlled by *WT1* (Drummond *et al.*, 1992). Mice lacking *IGF-2* are dwarfed in development (Dechiara *et al.*, 1990), as are those lacking *IGF-1* (Powell Braxton *et al.*, 1993). Both IGFs bind to IGF1r, which is expressed in the ureteric bud and its branches (Wada *et al.*, 1993). However, since double mutants lacking both the receptor and the *IGF-1* gene develop histologically normal, but proportionately smaller, kidneys (Liu *et al.*, 1993a), it is likely that the IGFs do not act primarily as renal morphogens but as mitogens, stimulating cell proliferation during kidney development.

Mutations exist for a number of extracellular matrix components, proteins which provide the scaffold for the developing kidney epithelium. Collagen IV, which is expressed in the basement membranes of a developing kidney, is mutated in the human disorder Alport's syndrome, resulting in nephritis which leads to eventual renal failure, and sensorineural hearing loss (Barker *et al.*, 1990). With the exception of electron microscopic abnormalities in the basement membrane of affected individuals, renal morphology is normal and kidney failure develops only in later life, indicating that this matrix protein is

not required for kidney development but for postnatal renal function. Mutant mice with null alleles for tenascin (Saga *et al.*, 1992) and vimentin (Colucci-Guoyon *et al.*, 1994) develop normally without any renal histological abnormalities, indicating considerable functional redundancy in the biological activity of these proteins. An insertional mutation in collagen I, which is expressed in undifferentiated mesenchyme prior to induction, does result in a lethal phenotype, although the kidneys from such mutants develop normally when cultured *in vitro* (Kratochwil *et al.*, 1986).

Recently, the biological activities of sulphated proteoglycan have been found to be essential for normal kidney development (Davies *et al.* 1995). Some sulphated proteoglycans are biologically active through their core proteins, whilst the others function through their sulphated glycosaminoglycan (GAG) sidechains (Kjellen and Lindahl, 1991). An early extracellular marker of induction is the proteoglycan, syndecan1 (Vainio *et al.*, 1989). Syndecan 1 upregulation is reported to be necessary for mesenchymal condensation, the first morphological event in nephrogenesis (Vainio *et al.*, 1992). This syndecan 1-induced condensation can be severely impaired *in vitro* by heparin and heparan sulphate (Platt *et al.*, 1990). The timing of syndecan 1 elevation would suggest it has a role in nephrogenesis that is essential for proper mesenchymal condensation and differentiation. Impairment of sulphated proteoglycan function would therefore be expected to affect the development of nephrogenic condensates. However, *in vitro* cultivation of embryonic kidney rudiments with sodium chlorate (a specific inhibitor of GAG sulphation), and with the GAG-degrading enzymes heparitinase and chondroitinase ABC, showed that these inhibitors of GAG sidechain sulphation, did not prevent nephron development (Davies *et al.*, 1995). Unexpectedly, the inhibitors affected the growth and branching of the ureteric bud, in a reversible manner. Interestingly, the *in vitro* affect of sodium chlorate was partially overcome by two agents: the protein kinase C activator, phorbol 12-myristate acetate (PMA), and HGF. Whilst PMA was

able to restore the initial branching morphogenesis of the ureteric bud without improving its growth in the presence of sodium chlorate, HGF was able to restore bud growth without stimulating branching morphogenesis. This strongly indicates the presence of two independent pathways which control the growth and the branching morphogenesis of ureteric buds.

6.2 Analysis of kidney development in *Sd* mice

Renal abnormalities commonly seen in *Sd* mice are thought to occur primarily through the failure of the ureteric bud to branch and ramify through the metanephric blastema (Gluecksohn-Waelsch and Rota, 1963). To investigate more closely the development of *Sd* kidneys, an *in vitro* organ culture assay was used (section 2.2.3-2.2.5) to culture kidney rudiments from 12.5 day homozygous (*SdSd*), heterozygous (*Sd+*) and wild type (WT) embryos (section 2.2.3-2.2.4). Kidney rudiments were cultured in IMEM alone (section 2.2.1), or in the presence of HGF or a protein kinase C activator, beta-phorbol-12, 13-dibutyrate (section 2.2.1). Fluorescently-tagged secondary antibodies were used in this analysis, so that laser-scanning confocal microscopy could be employed to section optically through intact rudiments, making histological analysis quicker and simpler than in the use of paraffin embedded sections. Work presented in this chapter of the thesis was carried out in collaboration with S. Larsson at the MRC Human Genetics Unit.

6.2.1 *In vitro* development of kidney rudiments

Figure 6.2.1 shows an image generated by laser scanning confocal microscopy (section 2.2.7) of wild type (WT) kidney rudiments originating from the left side (LHS) and right side (RHS) of the body, which have been cultured for 48 hours (section 2.2.5), and then immunohistochemically stained (sections 2.2.2 and 2.2.6). The red signal is generated by the conjugation of a fluorescently-tagged (Texas Red) secondary antibody to a mouse monoclonal anti-calbindin antibody (section 2.2.2). Calbindin is a

calcium binding protein which is expressed in the ureteric bud during the early stages of collecting duct morphogenesis, but is not expressed in the metanephric mesenchyme (Liu *et al.*, 1993b). For this reason, anti-calbindin antibody is a good marker for the ureteric epithelium. The green signal is generated by the conjugation of a fluorescently-tagged (FITC) secondary antibody (section 2.2.2) to a rabbit polyclonal, anti-mouse-laminin antibody (section 2.2.2). The laminin glycoproteins are a major component of the epithelial basement membrane (Abrahamson and St. John, 1993), and their expression is confined to the basement membrane as the metanephric mesenchyme aggregates and epithelialises. Laminin is not greatly expressed in undifferentiated metanephric mesenchyme (Saxen, 1987), but its expression intensifies in epithelium generated from condensed mesenchyme and the ureteric bud. These antibodies were used throughout the analysis of the *in vitro* development of WT and *Sd* rudiments.

Figure 6.2.1 shows the considerable development of WT kidney rudiments following 48 hours in culture. When dissected, by day 12.5 of gestation kidney rudiments have undergone the first stage in branching morphogenesis, giving the ureteric bud a characteristic T-shape surrounded by undifferentiated mesenchyme. All development subsequent to this occurs in culture. Figure 6.2.1, shows how the ureteric branches have grown and ramified through the metanephric blastema. Laminin-stained condensates have epithelialised and differentiated into comma and S-shaped bodies. Some nephrogenic tubules have reached an advanced stage of maturation and have become fused to the epithelia of the collecting duct. Both kidneys show an approximately equal number of branching events on either side of the initial T-shaped bud. In addition, both rudiments are at a similar stage of development.

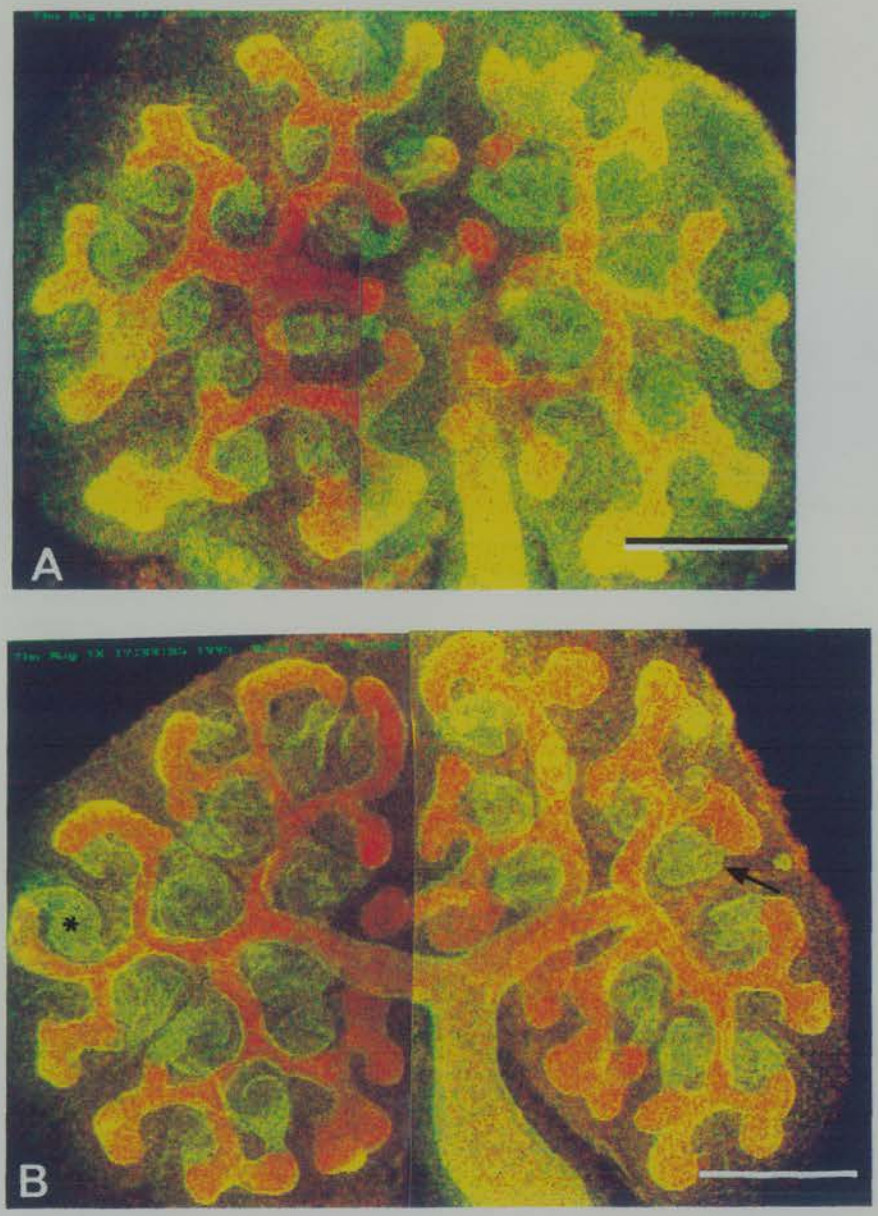

Figure 6.2.1 *In vitro* development of WT kidney rudiments

Rudiment A (from the RHS of the body) and rudiment B (from the LHS of the body) were cultured for 48 hours. The images were generated by laser-scanning confocal microscopy by the simultaneous collection of FITC (green) and Texas Red signals against laminin and calbindin, respectively. Nephrogenic condensates are shown in green and the collecting duct system in red. On figure B, the asterix denotes the development of a comma-shaped body and the arrow, that of a S-shaped body. Magnification was 6.3X and the scale bar represents 250 microns.

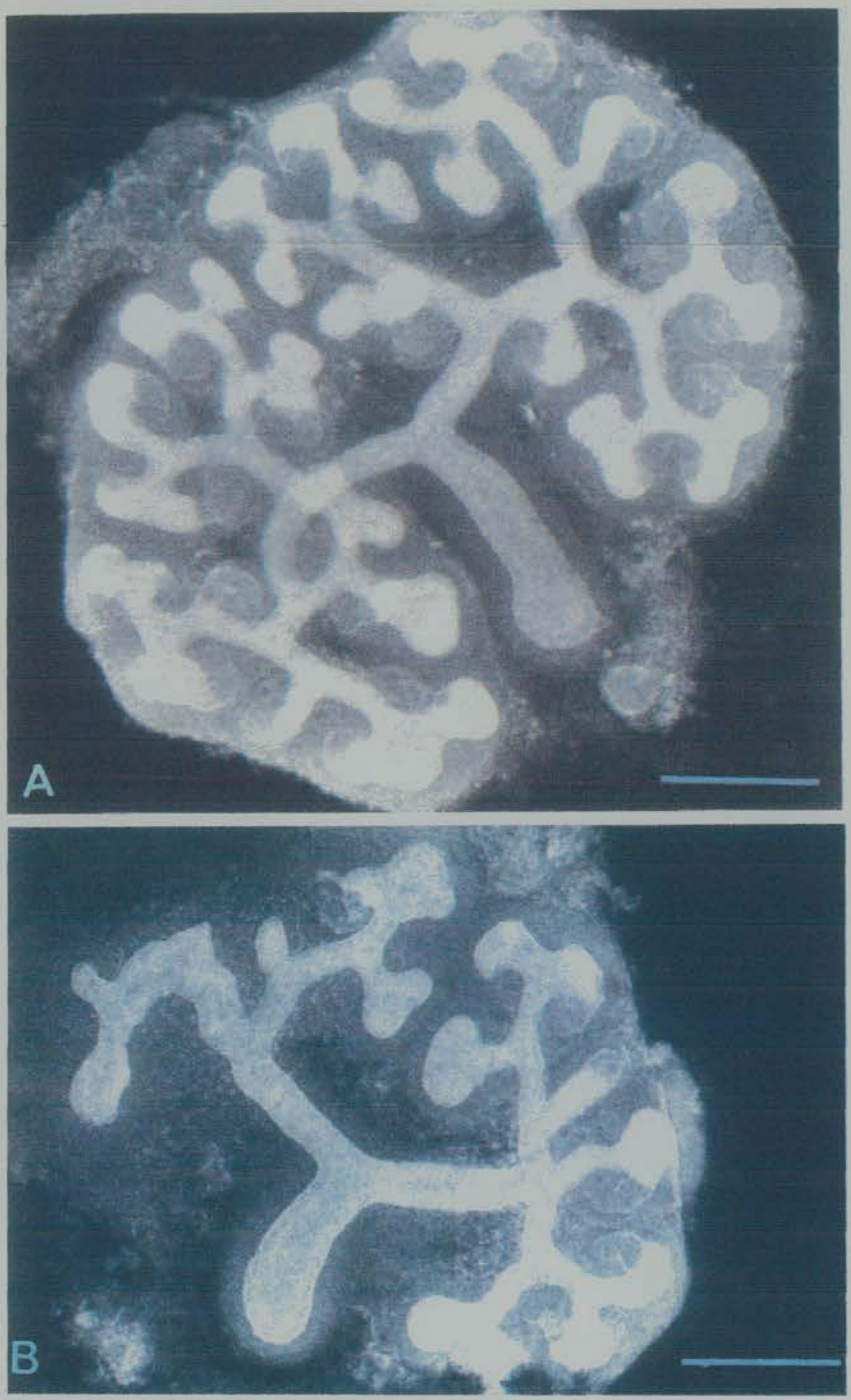

Figure 6.2.2 *In vitro* development of *Sd+* kidney rudiments
Rudiment A (from the RHS of the body) and rudiment B (from the LHS of the body) were cultured for 48 hours. The image was generated by laser-scanning confocal microscopy as in figure 6.2.1. Magnification is 6.3X and the scale bar represents 250 microns.

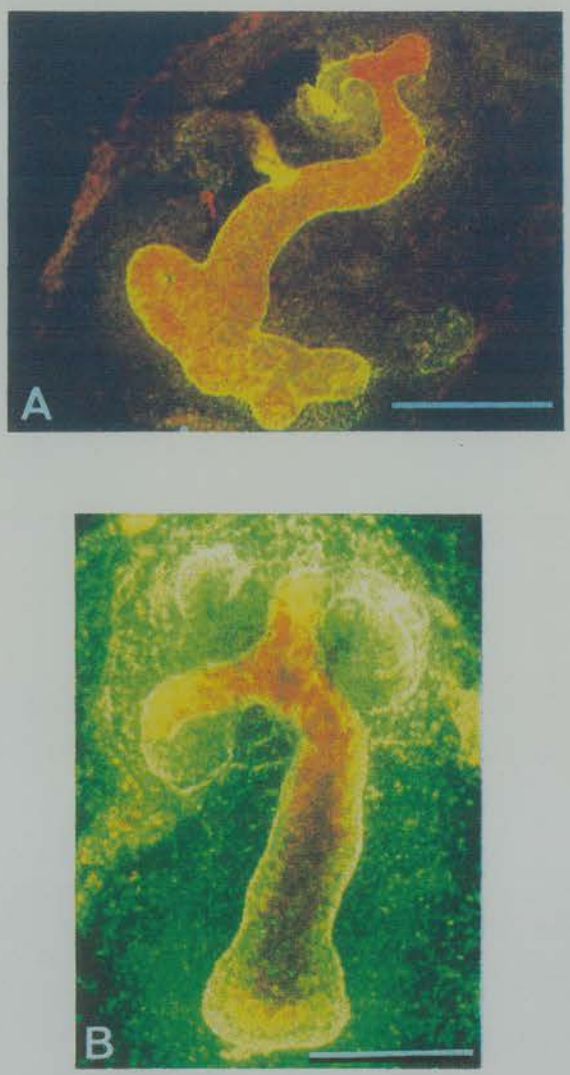

Figure 6.2.3 *In vitro* development of *SdSd* kidney rudiments
Rudiment A (from the RHS of the body) and rudiment B (from the LHS of the body) were cultured for 48 hours. The image was generated by laser scanning confocal microscopy as in figure 6.2.1. Magnification is 6.3X in figure A (scale bar represents 250 microns), and 10X in figure B (scale bar represents158 microns)

Figure 6.2.2 shows the *in vitro* development of kidney rudiments from a heterozygous (*Sd+*) Danforth short-tail mouse after 48 hours in culture. The image was generated from the anti-laminin and anti-calbindin fluorescent signals, and shows condensate development and the branching pattern of the ureteric bud in mutant mice. Figure 6.2.2 is an example of a phenomenon that was commonly seen in *Sd+* rudiment cultures: the retarded development of the LHS kidney in comparison to that of the RHS kidney, where both organs have been cultured under the same conditions and have originated from the same mouse. This phenomenon has been previously reported (Glueckson-Schoenheimer, 1943).

Figure 6.2.2 shows that the RHS kidney (A) has a WT pattern of branching morphogenesis, whilst branching in the LHS kidney (B) is poorly developed and dysmorphic. This variation in the development of the RHS and LHS kidneys supports the phenotypic data generated from the backcross (detailed in table 3.3.1), where it was found that of the 73 *Sd+* mice with one larger kidney, in 96% it was the RHS kidney which was the larger of the pair. This variation is reflected in the development of heterozygote rudiments in culture.

Despite the decrease in ureteric branching, the metanephric mesenchyme in *Sd+* rudiments is still able to condense and form comma and S-shaped bodies, indicating that mesenchymal induction and differentiation remain unaffected by the mutation. It was noted, however, that in *Sd+* rudiments with dysmorphic ureteric branching, fewer condensates were able to form than in rudiments obtained from WT littermates (see figures 6.2.1, 6.4.2 and 6.4.3), and that their progress through the different stages of nephrogenesis (section 6.1.1) was retarded.

Figure 6.2.3 shows the *in vitro* development, after 48 hours in culture, of kidney rudiments dissected from a 12.5d homozygous (*SdSd*) embryo.

Kidney rudiments from *SdSd* embryos were difficult to dissect as the ureteric bud was retarded in development by day 12.5. The characteristic T shape of the ureteric bud, which indicates that the initial branching of the bud has occurred, was frequently absent in *SdSd* mice. Ureteric buds were often dissected in the form of a single-tip bud, surrounded by undifferentiated mesenchyme, and were usually smaller in size than rudiments dissected from WT and *Sd+* littermates.

The RHS rudiment (figure 6.2.3a) shows a dysmorphic ureteric bud that has failed to undergo the initial, dichotomous branching step needed to form a T-shaped bud. Laminin expression is visible in the ureteric bud epithelium, but is not present in the metanephric mesenchyme immediately adjacent to the ureteric bud, showing the absence of differentiating condensates. In the LHS kidney (figure 6.2.3b), the ureteric bud has undergone the primary branching step to form a T-shaped ureteric bud. Laminin staining shows the presence of two maturing condensates on either side of the branch, which are expressing laminin and which have therefore epithelialised.

Unexpectedly, the condensates in figure 6.2.3b are at an advanced, S-shaped, stage of development and have fused to the tips of the ureteric ducts. These data strongly indicate that when branching of the ureteric bud occurs, mutant metanephric mesenchyme is able to respond to inductive signals provided by the ureteric bud, to undergo condensation and nephrogenesis. The number of nephrons able to form appears to be a consequence of branching: the less well branched the ureteric bud is, the fewer the number of nephrogenic condensates which are able to form.

6.2.2 Discussion

Data presented here strongly suggests that the *Sd* mutation primarily affects the branching capacity of the ureteric bud, since when ureteric branching has occurred, the mutant mesenchyme appears able to undergo proper

tubulogenic differentiation (figure 6.2.3b). This does not concur with the findings of Gluecksohn-Waelsch and Rota (1963), which were unable to confirm the responsive capability of the metanephric mesenchyme in *Sd* mice. The retardation of mutant metanephric responsiveness in this study could have been due to the nature of the assay used. Kidney organ culturing was in its infancy when this work was carried out, and the use of plasma clots may have had adverse affects on the *in vitro* development of rudiments. The fact that only 53% of their control cultures developed suggests that it was not an ideal *in vitro* system. In addition, the use of undefined horse serum introduced unquantified substances into the culture media, the affects of which would have been unknown. In contrast, this study used defined media.

6.3. Analysis of the *in vitro* effects of HGF

Hepatocyte growth factor (HGF) is synthesised by the metanephric mesenchyme, and is the ligand for the c-*met* tyrosine kinase receptor, a receptor which is expressed in the epithelium of the ureteric bud and which is associated with the formation and maintenance of epithelial lumenal structures (Tsarfarty *et al.*, 1994). Madine-Darby canine kidney cells (MDCK) undergo tubulogenesis and branching morphogenesis when grown in Type I collagen in the presence of HGF (Santos and Nigram, 1993). Since HGF is able to induce branching morphogenesis in MDCK cells, and in cell lines derived from murine inner medullary collecting duct epithelial cells (Cantley *et al.*, 1994), *Sd* rudiments were cultured with HGF to investigate whether it could induce improved branching morphogenesis in *Sd* kidneys.

6.3.1 *Sd* kidney development in the presence of HGF

HGF (section 2.2.1) was used in the culturing of 12.5 day embryonic kidney rudiments. Culture times varied from 48-54 hours, to a maximum culture time of 62 hours. Kidney rudiments were immunohistochemically stained (section

2.2.6) using the fluorescent secondary antibodies described in section 6.2.1 Whole mount rudiments were histologically analysed using laser-scanning confocal microscopy (section 2.2.7). In all culture assays, it was noted which side of the body each kidney originated from. Confocal microscopy images were generated from the anti-calbindin (Texas Red) fluorescent signal only, to facilitate the counting of branches.

From figure 6.3.1, it can be seen that the presence of HGF did not affect the growth or development of wild type (WT) rudiments (figure 6.3.1), following 54 hours in culture, since the development of the kidney originating from the LHS of the body, which was cultured in the presence of HGF, is comparable to that of the control kidney.

Figure 6.3.2 shows kidney rudiments from a heterozygous (*Sd+*) mouse cultured in the presence and absence of HGF for 48 hours. The RHS control kidney has a WT pattern of branching morphogenesis, with the ureteric system well developed. When the image was generated from the FITC (anti-laminin) fluorescent signal, mature condensates surrounding the tips of the collecting ducts were visible (data not shown). The LHS kidney shows a dysmorphic pattern of branching that is less well developed than that of the RHS kidney, despite the presence of HGF in the culture media. Since HGF failed to induce branching morphogenesis in other *Sd+* rudiment cultures (see table 6.3.2), it was concluded that under these culture conditions, and at a concentration of 80ng/ml, HGF was unable to induce improved branching morphogenesis in *Sd+* rudiments.

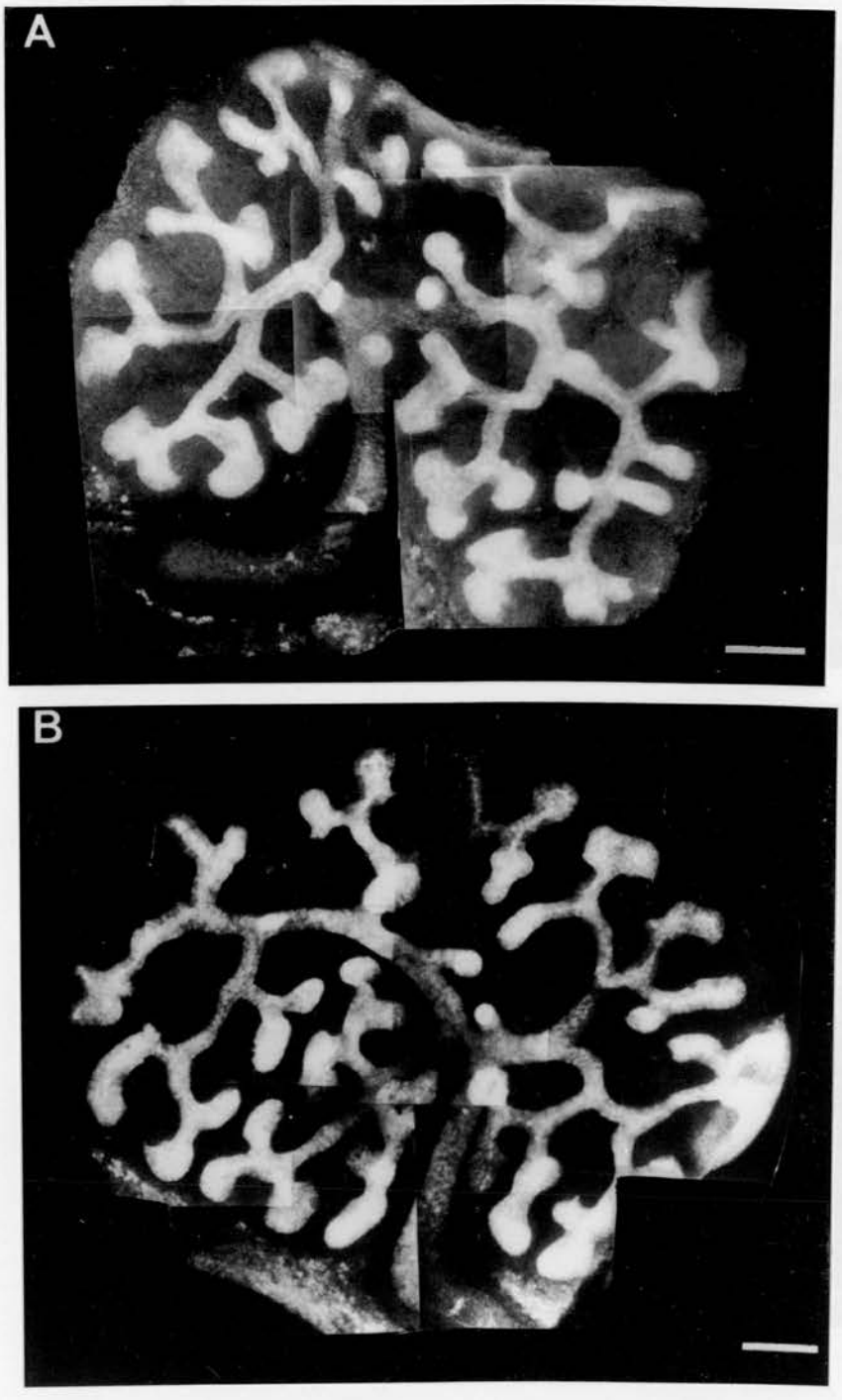

Figure 6.3.1 *In vitro* affect of HGF on the development of WT kidneys
Rudiment A (from the RHS of the body) and rudiment B (from the LHS of the body) were cultured for 54 hours. The image was generated by laser-scanning confocal microscopy, from the Texas Red signal against calbindin. HGF was used at a concentration of 80ng/ml. Magnification is 10X and the scale bar represents 158 microns.

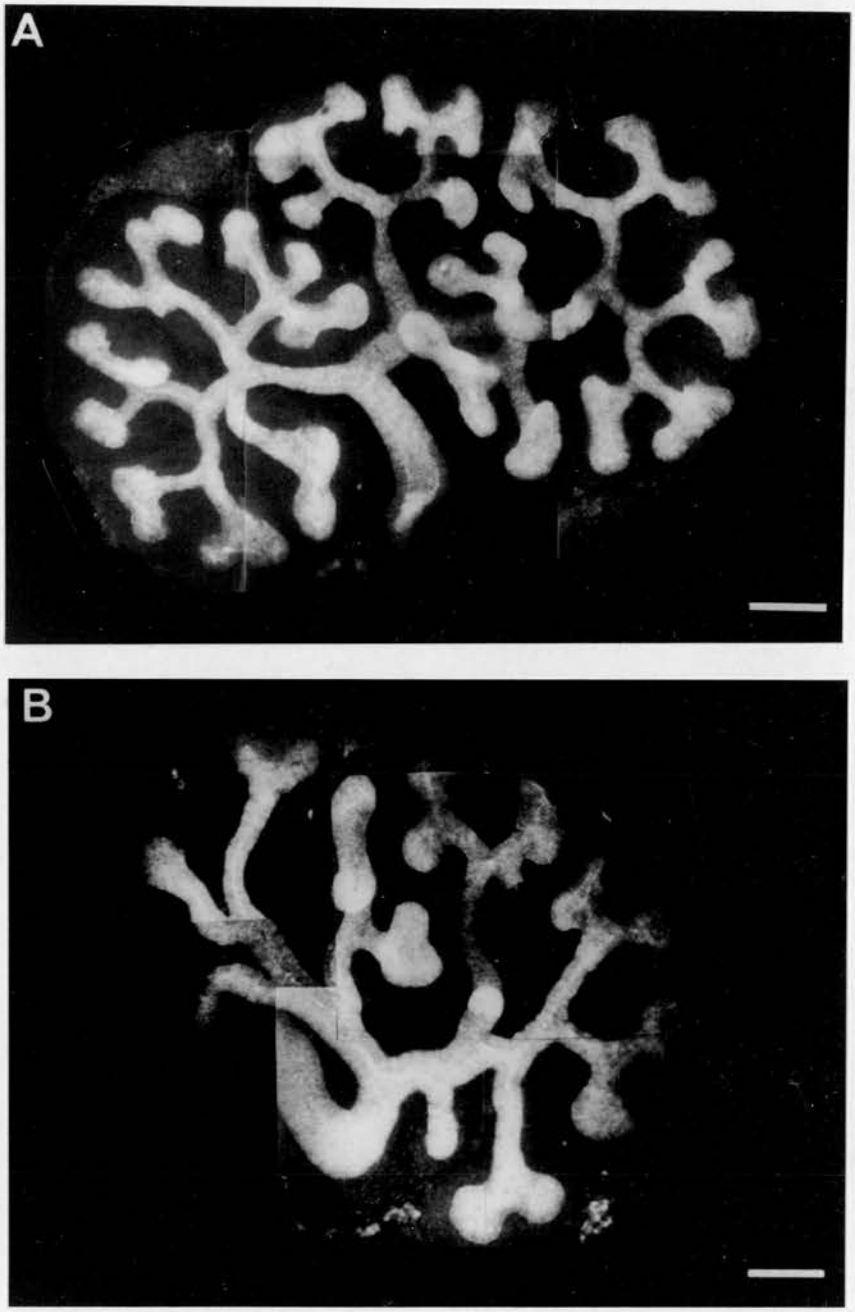

Figure 6.3.2 *In vitro* affect of HGF on the development of *Sd+* kidneys
Rudiment A (from the RHS of the body) and rudiment B (from the LHS of the body) were cultured for 48 hours. The image was generated by laser-scanning confocal microscopy as in figure 6.3.1. HGF was used at a concentration of 80ng/ml. Magnification is 10X and the scale bar represents 158 microns.

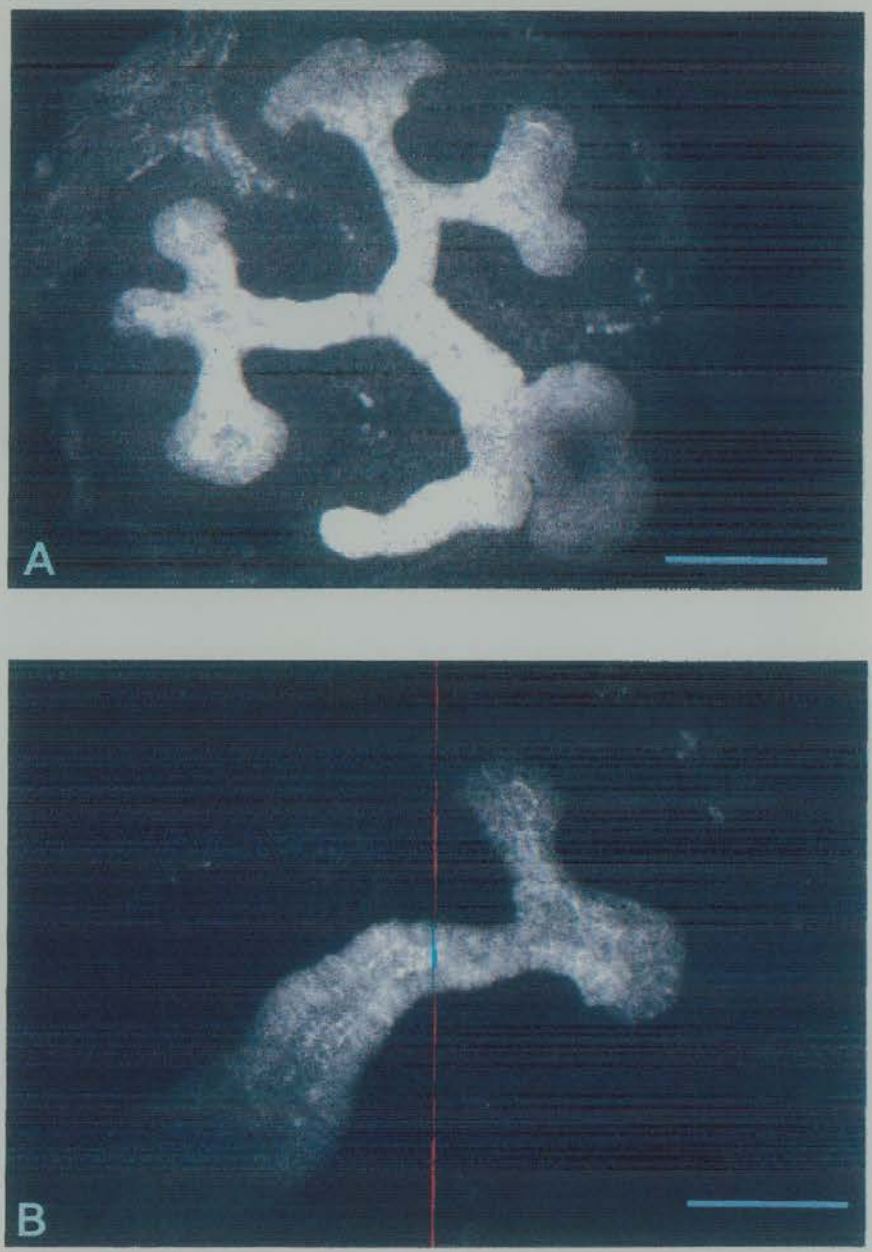

Figure 6.3.3 *In vitro* affect of HGF on the development of *SdSd* kidneys
Rudiment A (from the RHS of the body) and rudiment B (from the LHS of the body) were cultured for 48 hours. The image was generated by laser-scanning confocal microscopy as in figure 6.3.1. HGF was used at a concentration of 80ng/ml. Magnification is 10X and the scale bar represents 158 microns.

Figure 6.3.3 shows that HGF is not able to rescue the branching phenotype of *SdSd* kidney rudiments following 48 hours in culture, since despite the presence of HGF, the LHS kidney is still not able to proceed beyond the first step in branching morphogenesis. In comparison, the RHS rudiment has proceeded through 3 more branching events in the absence of HGF, indicating that *SdSd* kidneys have some residual branching ability which is retarded in development.

Culturing *Sd+* and *SdSd* rudiments in the presence of HGF for 54 and 62 hours produced similar results (confocal microscopy data for this is not shown), in that HGF was unable to induce improved branching morphogenesis in *Sd* kidneys, under the culture conditions used.

6.3.2 Branching analysis

In order to perform a preliminary analysis of the *in vitro* effect of HGF on branching morphogenesis, rudiments from WT, *Sd+* and *SdSd* mice were examined following 48 hours in culture. The number of branches which had formed were counted.

	WT cont.	**WT HGF**	***Sd+* cont.**	***Sd+* HGF**	***SdSd* cont.**	***SdSd* HGF**
Mean no. of branches	32	26	18	14	5	3
No. of rudiments	4	2	4	3	3	3
SEM	6.3	16	2.75	4.3	1.45	0.7

Table 6.3.2 Branching analysis of wild type (WT), heterozygous (*Sd+*) and homozygous (*SdSd*) rudiments.

HGF indicates those rudiments cultured with 80ng/ml HGF; cont. indicates control rudiments cultured without HGF. SEM (standard error of the mean) was calculated using the statistical analysis program on Excel 5.0 (Microsoft). Rudiments were cultured for 48 hours.

The Mann Whitney U Test (appendix C) was used to compare the mean number of branches from rudiments cultured in HGF, with the mean number of branches in the control rudiments. Probability values generated by this test indicate the degree to which samples from an identical population would be expected to overlap. For the WT rudiments, p=0.314; for the *Sd+* rudiments, p=0.35 and for the *SdSd* rudiments, p=0.1. These probability values indicate that there is no significant difference in the number of branches which form when the rudiments are cultured with or without HGF. Similar findings were observed in rudiments that were cultured for 54 and 62 hours in the presence of HGF (data not shown).

6.3.3 Discussion

From these data it was concluded that HGF was unable to induce improved branching morphogenesis in *in vitro* cultured *Sd* kidney rudiments. It was not investigated whether the use of HGF at other concentrations would have been affective, but since HGF has been reported to be biologically active at concentrations ranging from 30-100ng/ml in kidney culture assays (Davies *et al.*, 1995), a concentration of 80ng/ml should have been sufficient to produce a biological response. Nevertheless, it is not possible to definitively exclude HGF efficacy in the absence of positive controls. Future experiments could include a control to assay for the biological activity of HGF at different concentrations by BrdU incorporation, since HGF has known mitogenic activity. Preliminary data presented here suggests that *Sd* could affect a component of the HGF/c-met signalling pathway, preventing the ureteric epithelium from responding to the inductive signal provided by HGF. Alternatively, since HGF is not the primary branching morphogen of the mature kidney (Uehara *et al.*), *Sd* may affect an alternative signalling pathway.

6.4. Analysis of the *in vitro* effects of protein kinase C activation

It has been proposed that the activity of HGF, in inducing tubulogenesis and branching morphogenesis in MDCK cells, can be modulated by several different protein phosphorylation mechanisms, including those mediated by tyrosine kinases, protein kinase C and protein kinase A (Santos *et al.*, 1993). Protein kinase C activation has also been found to induce branching morphogenesis in sodium chlorate-inhibited rudiments (Davies *et al.*, 1995). To investigate whether protein kinase C (PKC) activation could induce branching morphogenesis in *Sd* kidneys, rudiments were cultured in the presence of beta-phorbol-12, 13-dibutyrate (PDBu), a phorbol ester known to activate PKC (Chiu *et al.*, 1987).

6.4.1 *In vitro* culture of kidney rudiments in the presence of PDBu

Figure 6.4.1 shows WT rudiments cultured for 48 hours in the presence of 10nM (B), 20nM (C) and 200nM (D) PDBu. Rudiment A was cultured in the absence of PDBu as a control. Rudiments A and B originated from the RHS of the body, and rudiments C and D from the LHS. PDBu appears not to affect the *in vitro* development of WT kidneys at these concentrations, over a 48 hour culture period, as rudiments B ,C, and D have developed to the same degree as that of the control rudiment, A.

Figure 6.4.2 shows heterozygous (*Sd+*) rudiments cultured in the presence and absence of PDBu for 48 hours. Rudiments A and B , cultured in the absence of PDBu, are an example of the variability seen in the renal phenotype of *Sd+* mice. The kidney dissected from the RHS of the body has developed poorly in comparison to that of the WT rudiments (figure 6.4.1). The branching pattern is not as extensive in this rudiment as it is in the WT rudiments, and branching appears to have failed at a number of ureteric tips. The rudiment dissected from the LHS of the body (figure 6.4.2B) has a more severely affected branching pattern. Branching appears retarded, and the branch on the right hand side of the T has failed to undergo a second

branching event, which has occurred normally on the left hand side of the T. The failure in proper branching at this point gives the ureteric branches a dysmorphic appearance. Where branching morphogenesis has occurred in rudiments A and B, condensates undergoing nephrogenesis are visible. It can be seen from these rudiments that the number of condensates which form is dependent on the number of duct tips available for the induction of mesenchymal differentiation.

Heterozygous (*Sd+*) rudiments C and D, in figure 6.4.2, were cultured in the presence of 20nM PDBu for 48 hours. Rudiment C was dissected from the RHS of the body and has a branching pattern which is indicative of a failure in ureteric branching. Culturing rudiment C with 20nM PDBu has not altered the rudiment's inability to branch normally. Rudiment D, dissected from the LHS of the body, is severely retarded in its development, and has not progressed much beyond the initial branching event which forms the immature ureteric bud's characteristic T-shape. Culturing the rudiment with 20nM PDBu has failed to rescue the retarded and dysmorphic development of this kidney. It was concluded from these data that PDBu, at a concentration of 20nM and over a culture period of 48 hours, was not able to improve the branching morphogenesis of *Sd+* kidneys *in vitro*.

Figure 6.4.3 shows heterozygous (*Sd+*) rudiments cultured in the presence of 100nM (rudiments A and B) and 200nM (C and D) PDBu, for 48 hours. Rudiment A, dissected from the RHS of the body, has a pattern of branching morphogenesis which is retarded in development but not severely dysmorphic. Culturing it in the presence of 100nM PDBu could be responsible for this improvement in ureteric branching, however, this affect could also be attributable to the variability of phenotype seen in rudiments originating from the RHS of the body. The fact that rudiment B, dissected from the LHS of the body, still has a severely retarded and dysmorphic pattern of branching morphogenesis, despite having also been cultured in

100nM PDBu, suggests that the latter is so. If an improvement in branching morphogenesis was due to the biological activity of PDBu, it would be evident in both RHS and LHS rudiments. Since no such improvement was evident, it was concluded that PDBu was not able to induce normal branching morphogenesis in *Sd+* kidneys at a concentration of 100nM and over a culture period of 48 hours.

Rudiments C and D (figure 6.4.3) were cultured in the presence of 200nM PDBu for 48 hours. Rudiment C, which originates from the RHS of the body, has an almost WT pattern of ureteric branching, with many condensates fused to collecting duct epithelia. Rudiment D, originating from the LHS of the body, has failed to undergo branching morphogenesis on the LHS branch of the T. The presence of 200nM PDBu in the culture was not sufficient to induce normal ureteric branching in this part of the rudiment. For this reason, it was concluded that PDBu, at a concentration of 200nM and over a culture time of 48 hours, was not able to induce normal ureteric branching in *Sd+* kidneys.

In figure 6.4.4, homozygous (*SdSd*) rudiments A (RHS of origin) and B (LHS of origin) were cultured in the absence of PDBu for 48 hours. Rudiment A has not progressed beyond the primary branching stage which divides the ureteric bud into a T. Nevertheless, nephrogenic condensates are visible at either end of the bud, confirming that the mutant metanephric mesenchyme is able to respond to ureteric bud-derived inductive signals, but only when branching of the bud has occurred. Rudiment B has undergone further ureteric branching than A, and mesenchymal condensates are visible where branching morphogenesis has occurred.

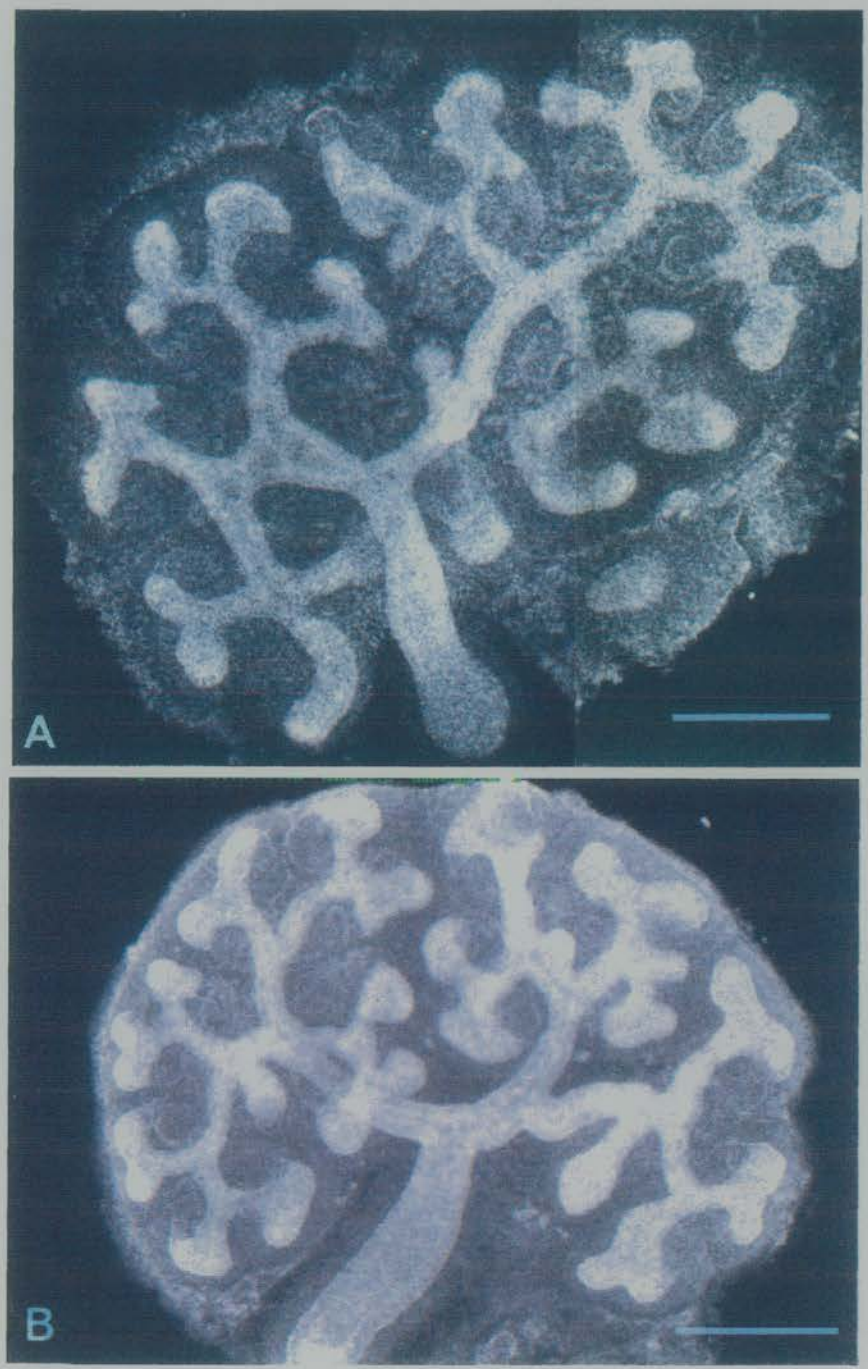

Figure 6.4.1 *In vitro* affect of PDBu on the development of WT kidneys

The figure is continued over page.

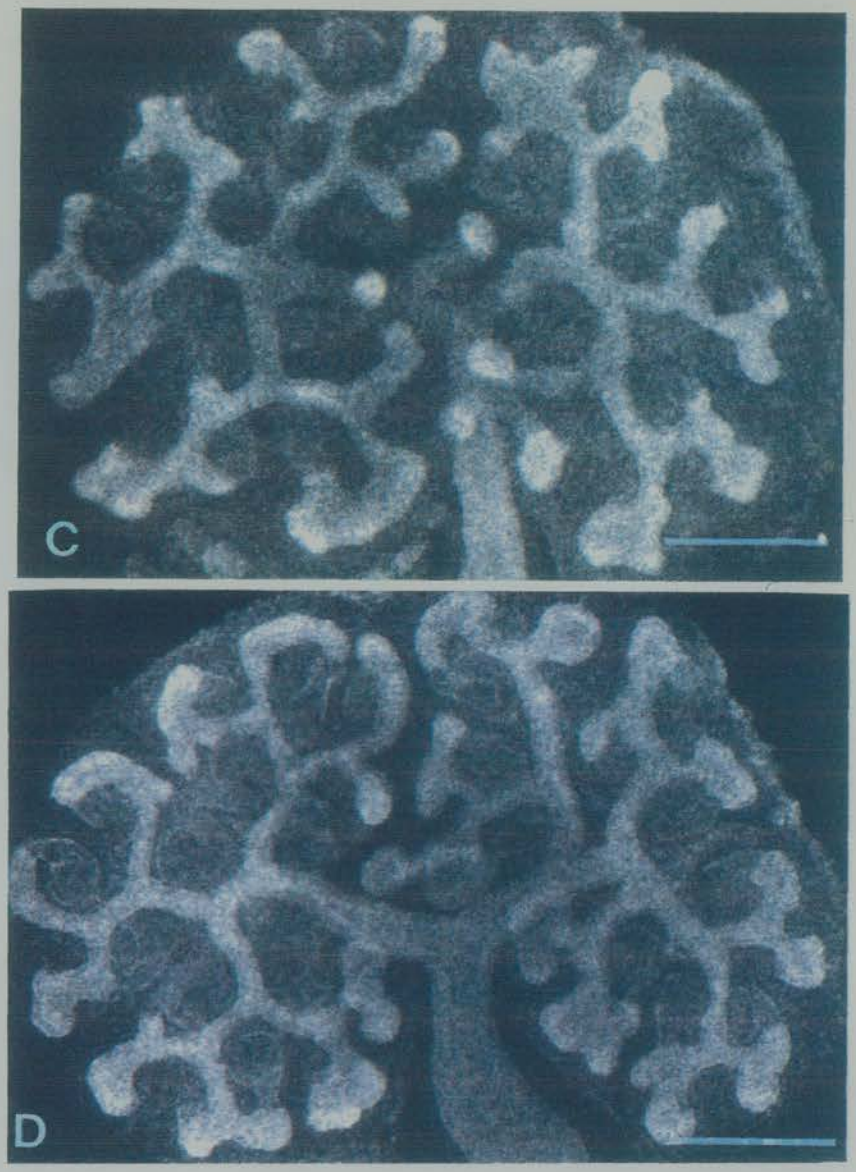

Figure 6.4.1 *In vitro* affect of PDBu on the development of WT kidneys (continued)

Rudiments A and B originated from the RHS of the body. Rudiment A was cultured without PDBu and rudiment B was cultured with 10nM PDBu for 48 hours. Rudiments C and D originated from the LHS of the body. Rudiment C was cultured with 20nM, and rudiment D with 200nM PDBu for 48 hours. The images were generated from FITC (anti-laminin) and Texas Red (anti-calbindin) fluorescent signals using laser-scanning confocal microscopy. Magnification is 6.3X and the scale bar represents 250 microns.

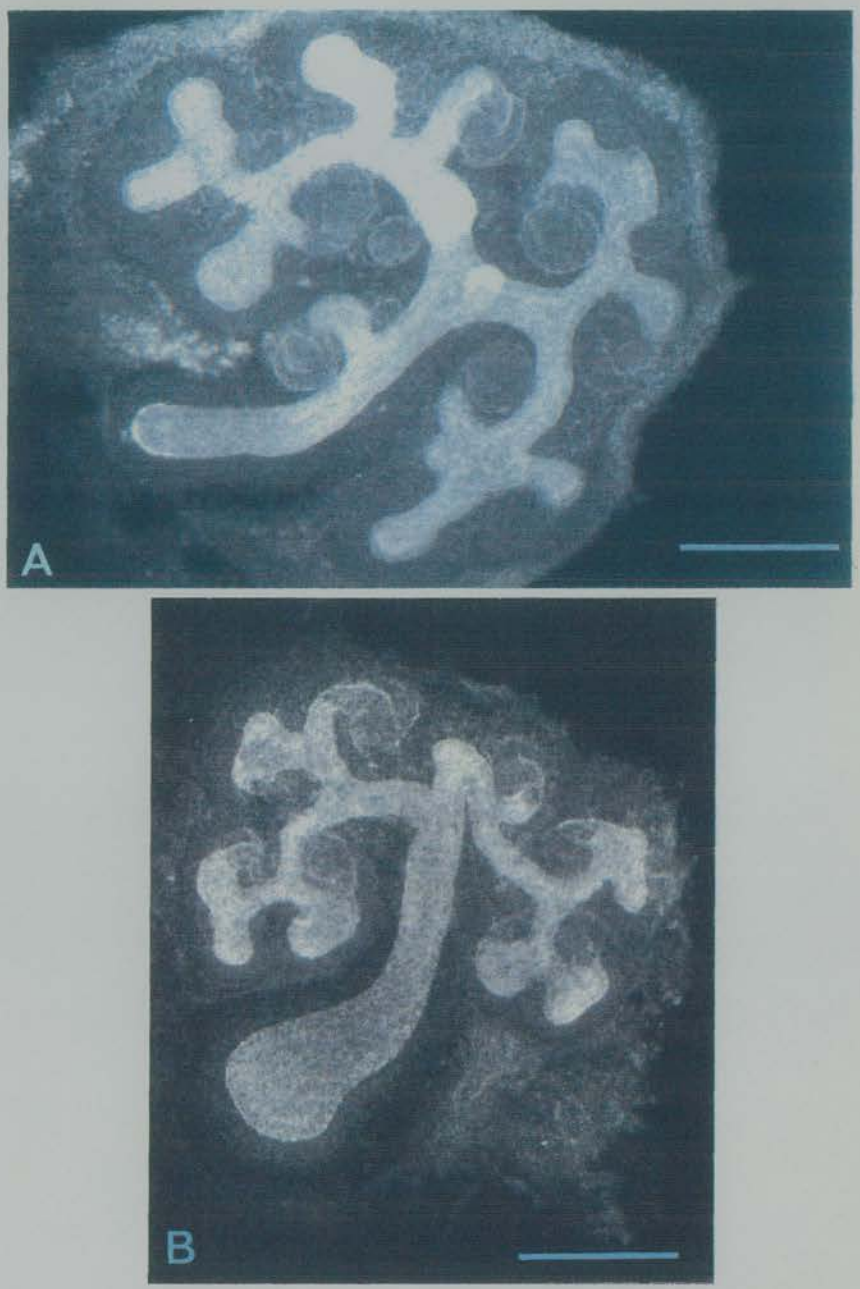

Figure 6.4.2 *In vitro* affect of PDBu on the development of *Sd+* kidneys

(Continued over page)

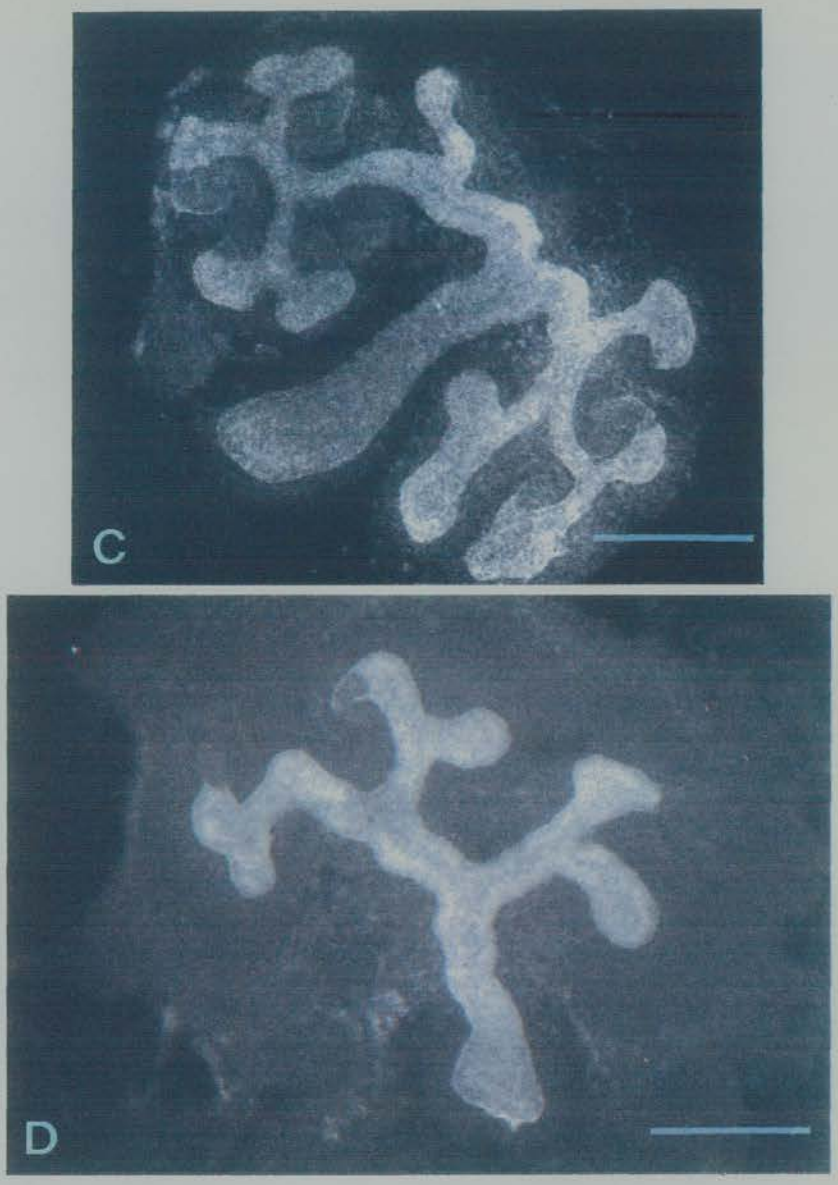

Figure 6.4.2 *In vitro* affect of PDBu on the development of *Sd+* kidneys
Rudiments A (from the RHS) and B (from the LHS of the body) were cultured without PDBu, rudiments C (from the RHS) and D (from the LHS) were cultured in the presence of 20nM PDBu, for 48 hours. The images were generated by laser-scanning confocal microscopy as in figure 6.4.1. Magnification is 6.3X and the scale bar represents 250 microns.

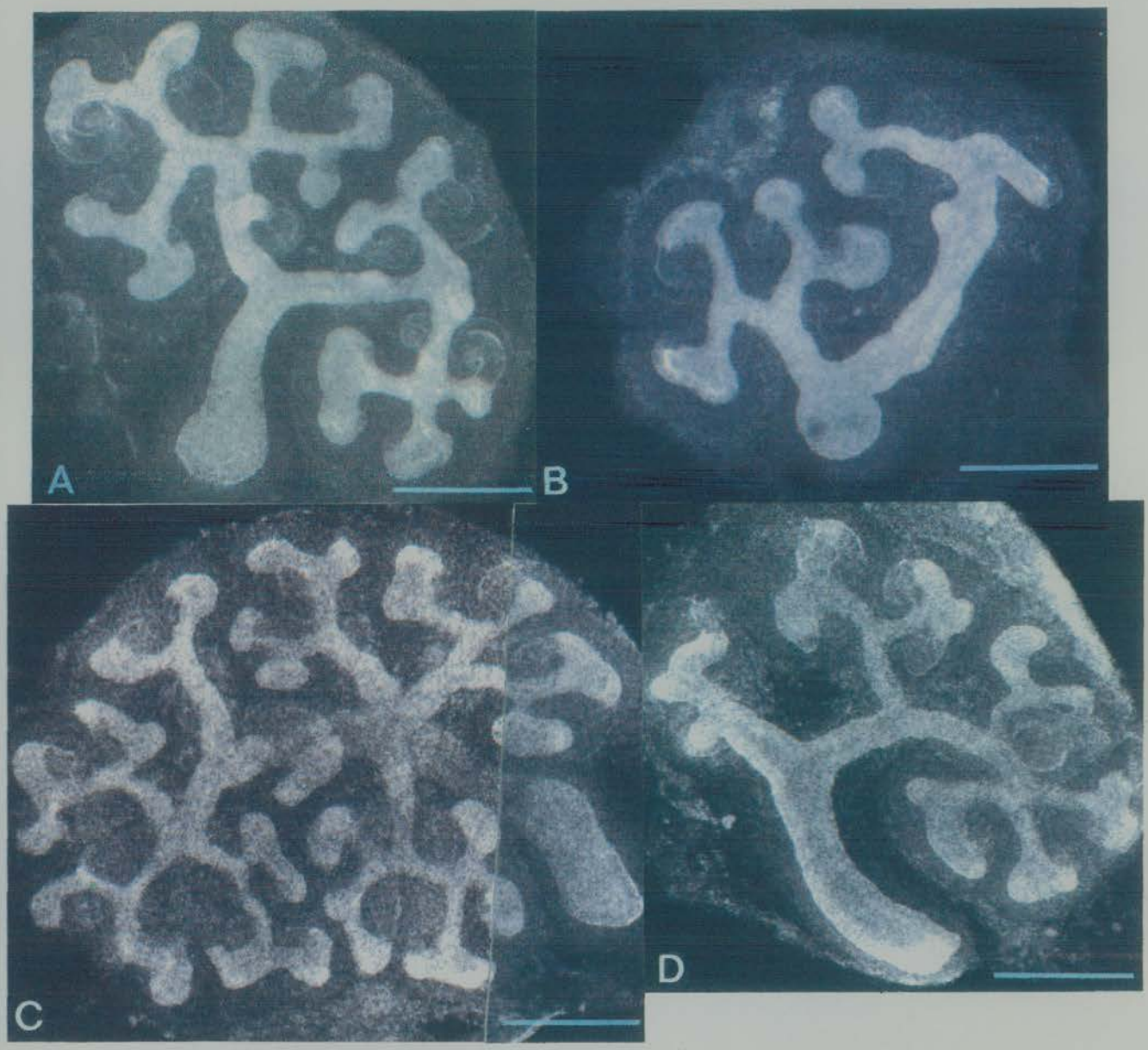

Figure 6.4.3 *In vitro* affect of PDBu on the development of *Sd+* kidneys
Rudiments A (from the RHS) and B (from the LHS of the body) were cultured in the presence of 100nM PDBu, rudiments C (from the RHS) and D (from the LHS of the body) were cultured in the presence of 200nM PDBu, for 48 hours. The images were generated by laser-scanning confocal microscopy as in figure 6.4.1. Magnification is 6.3X and the scale bar represents 250 microns.

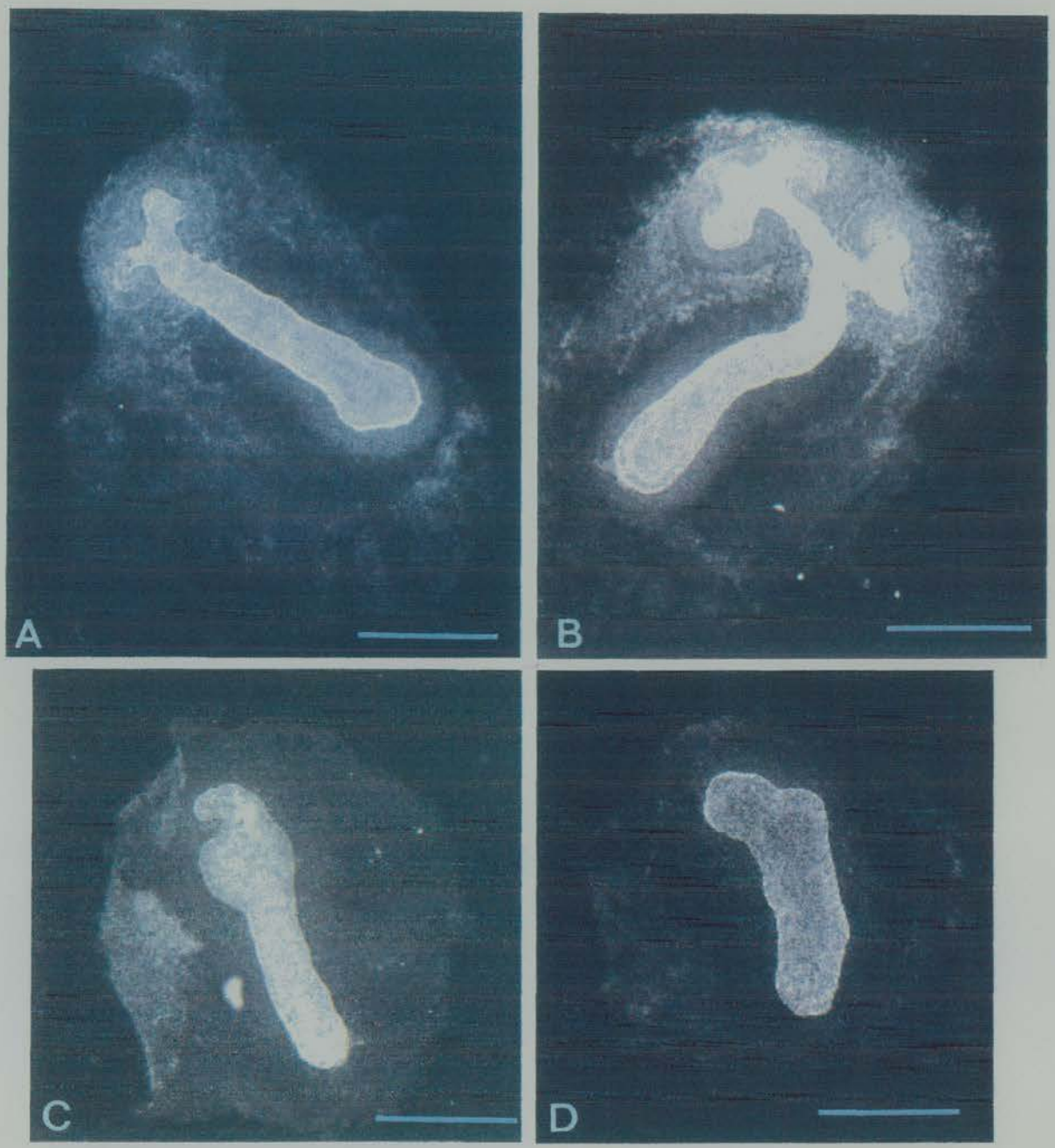

Figure 6.4.4 PDBu affect on the *in vitro* development of *SdSd* kidneys
Rudiments A and B were cultured without PDBu. Rudiment C was cultured with 20nM PDBu and rudiment D with 100nM PDBu, for 48 hours. The images were generated by laser-scanning confocal microscopy, as in figure 6.4.3. Magnification is 6.3X and the scale bar represents 250 microns.

Rudiments C (RHS of origin) and D (LHS of origin), from homozygous mutants, were cultured in the presence of 20nM and 100nM PDBu, respectively, for 48 hours. From the severely retarded development of both rudiments, and from the similarity they bear to the control rudiments A and B, it is evident that PDBu is not able to induce branching morphogenesis.

6.4.2 Discussion

Protein kinase C (PKC) activation, by PDBu, was not able to induce branching morphogenesis in *Sd* rudiments, in the concentration range (10-200nM) used. Due to the inhibitory affect which phorbol esters are known to have on PKC, over extended periods of time, it cannot be ruled out that PDBu was acting to inhibit PKC over the 48 hour culture period. Culturing WT rudiments, for 48 hours in different concentrations of PDBu (figure 6.4.1), appeared not to affect WT development, this does not exclude the possibility, however, that PDBu failed to induce branching morphogenesis because it was inhibiting, and not activating, PKC over the culture period used.

The inhibition and activation of PKC has been reported to have different affects on renal morphogenesis depending on the biological assay used. In collagen type I-seeded MDCK cells, the branching activity of HGF is reported to be enhanced by inhibitors of PKC, suggesting that PKC activity may negatively modulate HGF-induced branching and tubulogenic processes in this system (Santos *et al.*, 1993). Because *Sd* rudiments were cultured in the absence of an appropriate positive control, such as the use of sodium chlorate branch-inhibited rudiments, it was not possible to define the degree to which PDBu was activating or inhibiting PKC. If it was inhibiting PKC, and if these two systems were considered to be analogous, one might expect to see an increase in ureteric branching. Recent data does exist to suggest, however, that the MDCK system and the rudiment culture assay are not strictly comparable systems. In the development of wild type rudiments cultured in the presence of sodium chlorate, PKC activation by phorbol 12-

myristate acetate (PMA) was reported to partially restore the branching ability of chlorate-inhibited ureteric bud (Davies *et al.*, 1995). The development of PMA-induced swellings, indicative of branch initiation, that were seen in these cultures were not evident in *Sd* rudiments cultured in the presence of PDBu, despite the fact that both chlorate-inhibited rudiments and *Sd* kidneys fail to develop primarily through the failure of the ureteric bud to branch normally. Thus, simple conclusions cannot be drawn from either system as to why PDBu failed to induce branching morphogenesis in *Sd* kidneys. The *Sd* mutation may affect a signal transduction pathway which does not result in the activation of PKC, alternatively, it could affect the mesenchymal derived signal itself or the ureteric receptor upon which the signal acts.

Of interest was the finding that mice heterozygous for the *Sd* mutation share a common phenotypic feature, where the kidney originating from the left side of the body was frequently more severely affected by the mutation than the kidney originating from the right side. From table 3.3.1 it can be seen that only 5 wild type mice from the backcross had a RHS kidney which was bigger than the LHS one, compared to 73 heterozygous mice. This indicates that the degree of asymmetry seen in the renal development of heterozygous *Sd* mice is higher than that seen in the wild type population. This feature has been previously reported (Glueckson-Schoenheimer, 1943; Mesrobian and Sulik, 1992).

Whilst the action of modifying genes could explain renal phenotypic variation between individuals in a backcross population, such a genetic action would not explain the asymmetrical development of kidneys arising from a single mouse which are genotypically identical. However, left-right asymmetries are common in vertebrates. The consistent handedness of left-right asymmetric organs relative to the dorsoventral and anteroposterior axis is evolutionarily conserved. For example, the heart tube loops to the right in all chordates

(Yost, 1995). A number of mouse mutations exist which alter the development of left-right asymmetric organs, such as *iv* (Hummel and Chapman, 1959), in which 50% of the offspring are phenotypically situs invertus, and *inv* (Yokoyama *et al.*, 1993), in which 100% of the offspring are situs invertus, with respect to the left-right axis. These mutations strongly suggest that left-right asymmetry has an underlying genetic basis. Recently, a number of RNAs, including those for sonic hedgehog and acitvin, were found to be asymmetrically distributed along the left-right axis in chick embryos, during gastrulation and neurula stages (Levin *et al.*, 1995), strongly implicating these genes in the establishment of left-right asymmetry. It is therefore possible that the asymmetry seen in *Sd* mice, in the development of their kidneys, could be due to alterations in gene expression and RNA distribution brought about by the loss of the *Sd* gene product. The possibility that *Sd* might be a gain of function mutation can also not be excluded, and as such, its misexpression or inappropriate biological activity could also underlie renal left-right asymmetry. Until the *Sd* gene itself is identified the question of this asymmetry in *Sd* mice will be difficult to investigate further.

In conclusion, this study has shown that mutant metanephric mesenchyme is able to respond to ureteric branching by undergoing condensation and nephrogenesis. This suggests that whilst we do not know the biological function of *Sd*, it is likely that the ability of the mesenchyme to respond to ureteric inductive signals is unaltered by the mutation, and that upon branching, the ureteric bud is still able to induce the mesenchyme to respond. Hepatocyte growth factor and the protein kinase C activator, beta-phorbol-12, 13-dibutyrate, have been shown to be unable to induce *in vitro* branching morphogenesis in mutant kidney rudiments.

Chapter 7: Future Work

Since a discussion of results has been presented at the end of each results section, this chapter will focus on plans for future work.

The primary objective is to isolate the gene for *Sd* using a positional cloning strategy. The rate-limiting step in positional cloning is the identification and characterisation of genes in a genomic region, thus, it is important that the region of linkage has been well defined by mapping, and a number of strategies are available for the isolation of candidate transcripts. Since the physical size of the region in the vicinity of *Sd* remains unknown, it is possible that an additional backcross may need to be set up if this region is large (>1Mb). If additional backcrosses are needed, the mouse species of choice for the backcross could be the *Mus musculus* sub-species *M.m.molossinus*. The advantage of using *M.m.molossinus* over *Mus spretus* would be that F1 mice of both sexes are fertile. This allows the quicker accumulation of informative recombinants. In addition, it is thought that recombinants in a backcross can cluster in a particular region, thus by using a different strain or subspecies of mouse it may be possible to avoid clusters in the same regions. Furthermore, markers that were not informative in one strain may be informative in another.

Genes can be isolated from a specific region of mammalian genomic DNA by a number of methods. These include the identification of CpG islands as potential sign posts for the ends of 50% of transcription units (Lavia *et al.*, 1987) and the use of "zoo blots" to detect cross-species conservation of genomic sequences (Rommens *et al.*, 1989). Random cloning of genomic segments, followed by sequence analysis, can also be used where mapping studies have defined a candidate interval (Harshman *et al.*, 1995). More technically sophisticated methods include the use of exon trapping techniques, to capture expressed sequences from large genomic regions (Duyk *et al.*, 1990), and the screening of cDNA libraries with complex probes (Lovett *et al.*, 1991). The degree of success in a positional cloning strategy

may depend upon the combined use of a number of these techniques, since each method has inherent problems (Harshman *et al.*, 1995). For example, exon trapping increases the likelihood of isolating genes composed of a number of exons, whilst decreasing the chances of identifying those composed of a few exons, and, depending on the vector, is unable to trap genes with only one or two exons. Nevertheless, exon trapping alone has proved to be efficient in the identification of novel genes, as in the cloning of the obese (*ob*) gene in the mouse (Zhang *et al.*, 1994). It is likely that a number of these strategies will be used in the *Sd* project.

As *Sd* is flanked on only one side by close polymorphic markers, it will be necessary to delineate close recombinants on the other side in order to define a candidate interval. To do this, YAC libraries will be screened with the YAC end fragment marker, Ln1LAT. This marker is non-recombinant with *Sd* in the backcross making it the closest marker to *Sd* that has been isolated to date. YACs identified with this marker will be used to generate end fragments, which will in turn be mapped onto the backcross. Once a region of the contig has been identified which crosses recombinants either side of *Sd*, it will be possible to determine which YAC (or YACs) contain *Sd*. Markers generated from this YAC can be used in the screening of mouse P1 (Pierce *et al.,* 1992) or BAC (bacterial artificial chromosome) (Shizuya *et al.*, 1992) libraries. The advantage of using these cloning vectors over YACs is that they have a reasonable insert size (an average of 150kb for BACs and 85kb for P1s), without the problems of insert instability and chimerism common to YACs.

Once an interval has been established, YAC or P1 clones positive for *Sd* will be analysed by restriction digestion to identify rare cutter sites (such as *Not*1) which are indicative of the presence of CpG islands, and to generate physical maps of the inserts. YAC and P1 clones will also be subjected to exon trap analysis, for this a recently described phage vector λGET (Nehls *et*

al., 1994) will be used. This vector accommodates 6.5-19kb of DNA and allows inserts to be automatically subcloned as multi-copy plasmids containing splice donor and acceptor sites which flank the genomic insert. Transfection of COS cells with this plasmid allows exons within the coding region to be *in vivo* processed, or "trapped". In addition, cDNA selection can be used to enrich for cDNAs which are encoded in the region of *Sd*. This can be done by hybridising a library of cDNA's (for example from a mouse foetal kidney library) to an immobilised YAC or P1 clone (Lovett *et al.*, 1991). It is also possible to carry out cDNA selection without first purifying YAC DNA from total yeast DNA, which makes the technique simpler and faster to use (Parimoo *et al.*, 1993). Once candidate transcripts have been isolated and sequenced, database searches will identify transcripts which have already been isolated and cloned. Some of these may be genes whose expression and function is well established. Matches will also be expected to ESTs (expressed sequence tags), which are the result of large scale random cDNA sequencing projects. Expression patterns for ESTs are not usually known. The expression of ESTs and novel transcripts during development could be studied using *in situ* and RT-PCR techniques.

The size of the *Sd* transcript will determine the approach to mutation analysis. In the event of the transcript being particularly large, direct sequencing of the region would not be practical. An alternative approach would be to PCR-amplify smaller fragments from the region (1-1.5kb) and use these in the detection of mismatched heteroduplexes of mutant versus wild type DNA. Since DNA from the original strain on which *Sd* arose is not available for the wild type control, a number of different wild type DNAs would need to be used in its place. Heteroduplexes formed between wild type and mutant cDNA, identify regions of interest for direct analysis by sequencing. Alterations in the mutant sequence will need to be analysed in a number of inbred strains to ensure that a polymorphism is not mistaken for the mutation. In the absence of any obvious structural changes to the coding

sequence, such as a frame shift causing a premature stop codon, transgenic methods could be used to confirm the identification of the mutation. If the introduction of an *Sd* transgene into mutant mice were to rescue the phenotype, it would indicate that *Sd* is a loss of function mutation which arises through haploinsufficiency. If the phenotype fails to be rescued it could indicate one of two possibilities: i) that the mutant transcript identified is not *Sd*, or ii) that *Sd* is a gain of function mutation.

Based on the present knowledge of the mutation, it is difficult to predict the nature or biological function of the *Sd* gene. The semi-dominant mode of the inheritance of *Sd* suggests that it could be a loss of function mutation. If this were so, one would expect to find the gene normally expressed in the developing kidney and notochord/neural tube, where the loss of its function would most severely impair development. In these systems, *Sd* could act as a regulator of transcription or as a component of an intracellular signalling pathway. Interestingly, no other alleles of *Sd* have been identified, and despite its easily recognisable phenotype, neither chemical nor radiation-induced mutagenesis programs have identified a mutation with a similar phenotype, although this could be because tail mutations are commonly generated and not always tested, or because Sd represents a "mild" allele, the complete loss of which could be lethal in heterozygotes. The possibility nevertheless exists that *Sd* may not be a loss of function, but a gain of function mutation. Gain of function mutations are less common in mice than loss of function mutations, but several are known to exist. One such mutation is disorganization (*Ds*), which causes a wide range of sporadic developmental anomalies including supernumerary limbs, body wall defects and abnormalities of the viscera (Crosby *et al.*, 1992). The *Ds* locus exhibits low pleiotropism, low penetrance and variable expressivity. Although the locus has yet to be cloned, it was found to be a gain of function mutation when trisomic fetuses had the same phenotype as heterozygote and homozygote *Ds* mice (Crosby *et al.*, 1992). The question then arises as to

what mechanism could *Sd* be a gain of function mutation, when it is inherited in a semi-dominant manner? A semi-dominant pattern of inheritance suggests that reduced gene dosage is responsible for the mutant phenotype, a mechanism which is usually associated with loss of function mutations. Increased gene dosage, ectopic or temporally altered mRNA expression, increased or constitutive protein activity, and novel protein functions are mechanisms which are all associated with gain of function mutations (Wilkie, 1994). It is possible that the *Sd* gene product assumes a novel function in mutant mice which alters the transcriptional control or expression of other genes. A secondary effect such as this may be responsible for a semi-dominant phenotype, since in the presence of one unaffected allele, there may be sufficient wild type activity to allow the partial functioning of genes downstream of *Sd*. However, alterations in protein function are often deleterious to survival.

A key objective in cloning the *Sd* gene is to investigate its involvement in human neural tube and renal disorders. Towards this end, a collaboration has been set up with Phillip Cox at the pathology department of the Hammersmith Hospital, London. Dr. Cox has assembled a database of foetal post mortems which have occurred over the past ten years. There are numerous incidences of renal malformations occurring concurrently with neural tube defects, and tissue samples would be available for each case in the database. This offers a valuable opportunity for future work on the *Sd* gene, to define a possible role in humans. Providing tissue samples which are paraffin embedded have been carefully prepared, DNA could be extracted from them to investigate whether *Sd* is altered in those cases with neural tube and renal defects. Tissue samples which are fresh or which have been stored in liquid nitrogen may also become available for analysis.

A more immediate plan for future work, in addition to the positional cloning project, is to further investigate the *in vitro* development of *Sd* kidneys using

the organ culture assay. It would be interesting to repeat the tissue combination assays performed by Gluecksohn-Waelsch and Rota (1963) (section 1.2), using defined media and laser scanning confocal microscopy. By enzymatically separating mutant ureteric bud from its surrounding metanephric mesenchyme, and culturing each tissue in combination with wild type and mutant inducers, it may be possible to clarify which aspect of the reciprocal induction process is affected by the mutation. An analysis of ureteric markers in *Sd* kidneys, such as c-*ret* and c-*ros*, would also be of interest, to investigate whether the mutation affects their expression. In addition, a transgenic mouse has recently been generated at the Human Genetics Unit which expresses the Lac-Z gene under the control of a WT1 promoter (A.Schedl, personal communication). By crossing *Sd* mice with the WT1-LacZ mouse it would be possible to make a detailed analysis of any alterations in WT1 expression brought about by the mutation. Since WT1 is known to be active in both kidney development and the development of the myotome (Miyagawa *et al.*, 1994), it is likely that its expression will be altered in the *Sd* mouse. Such a study could provide valuable information concerning the way in which the *Sd* and *WT1* genes interact during development, and ultimately it may be possible to extend this type of analysis to identify other gene acting downstream of *Sd*.

REFERENCES

Abrahamson, D.R., St.John, P.L. (1993) Laminin distribution in developing glomerular basement membranes. *Kidney International.* **43,** 73-78.

Armstrong, J.F., Pritchard-Jones, K., Bickmore, W.A., Hastie, N.D., Bard, J.B.L. (1992) The expresssion of the Wilm's tumour gene, *WT1*, in the developing mammalian embryo. *Mechanisms of Development.* **40,** 85-97.

Arveiler, B., Porteous, D.J. (1991) Amplification of end fragments of YAC recombinants by inverse-polymerase chain reaction. *Technique.* **3,** 24-28.

Avner, P. (1994) Quantity and quality: polygenic analysis in the mouse. *Nature Genetics.* **7,** 3-4.

Avner, P., Amar, L., Dandolo, L., Guenet, J. (1988) Genetic analysis of the mouse using interspecific crosses. *Trends in Genetics.* **4,** 18-23.

Bahary, N., Leibel, R.L., Joseph, L., Friedman, J.F. (1990) Molecular mapping of the mouse db mutation. *Proceedings of the National Academy of Sciences, USA.* **87,** 8642-8646.

Ballabio, A. (1993) The rise and fall of positional cloning. *Nature Genetics.* **3,** 277-279.

Balling, R., Deutsch, U., Gruss, P. (1988) *undulated*, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of *Pax1*. *Cell.* **55,** 531-535.

Barker, D.F., Hostikka, S.L., Zhou, J., Chow, L.T., Oliphant, A.R., Gerken, S.C., Gregory, M.C., Skolnik, M.H., Atkin, C.L., Tryggvason, K. (1990) Identification of mutations in the COL4A5 collagen gene in Alport's Syndrome. *Science.* **248,** 1224-1227.

Beechey, C.V., Searle, A.G. (1980) Chromosome 2: distance of *Sd* from centromere. *Mouse News Letter.* **63,** 17

Boloventa, P., Dodd, J. (1991) Pertubation of neuronal differentiation and axon guidance in the spinal cord of mouse embryos lacking a floor plate: analysis of Danforth's short-tail mutation. *Development.* **113,** 625-639.

Borger, R. (1950) Order of genes in the fifth linkage group of the house mouse. *Nature.* **166,** 697-698.

Botstein, D., White, R.L., Skolnick, M., Davis, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics.* **32,** 314-331.

Brown, S.D. (1992) The Mouse Genome Project and human genetics. A report from the 5th International Mouse Genome Mapping Workshop, Lunteren, Holland. *Genomics.* **13,** 490-492.

Bumcrot, D.A., McMahon, A.P. (1995) Sonic signals somites. *Current Biology.* **5,** 612-614.

Burke, D.T., Carle, G.F., Olson, M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science.* **236,** 806-814.

Cantley, L.G., Barros, E.J.G., Gandhi, M., Rauchman, M., Nigram, S.K. (1994) Regulation of mitogenesis, motogenesis, and tubulogenesis by hepatocyte growth factor in renal collecting duct cells. *American Journal of Physiology.* **36,** F271-F280.

Capecchi, M.R. (1989) Altering the genome by homologous recombination. *Science.* **244,** 1288-1292.

Castle, W.E. (1921) An improved method of estimating the number of genetic factors concerned in cases of blenidng inheritance. *Science.* **54**, 223-225.

Center, E.M., Spigelman, S.S., Wilson, D.B. (1982) Perinotochordal sheath of heterozygous and homozygous Danforth's short-tail mice. *Journal of Heredity.* **73,** 299-300.

Chartier, F.L., Keer, J.T., Sutcliffe, M.J., Henriques, D.A., Mileham, P., Brown, S.D.M. (1992) Construction of a mouse yeast artificial chromosome library in a recombination-deficient strain of yeast. *Nature Genetics.* **1,** 132-136.

Chiu, P.J.S., Tetzloff, G., Chatterjee, M., Sybertz, E.J. (1987) Phorbol 12,13-dibutyrate, an activator of protein kinase C, and vascular smooth muscle contraction and calcium fluxes. *Federation Proceedings.* **46**, 551-555.

Christ, B., Wilting, J. (1992) From somites to vertebral column. *Annals of Anatomy.* **174,** 23-32.

Cochard, P., Paulin, D. (1984) Initial expression of neurofilaments and vimentin in the central and peripheral nervous system of the mouse embryo *in vivo. Journal of Neuroscience.* **4,** 2080-2094.

Collins, F.S. (1992) Positional cloning: Let's not call it reverse anymore. *Nature Genetics.* 3-6.

Collins, F.S. (1995) Positional cloning moves from perditional to traditional. *Nature Genetics.* **9,** 347-350.

Colucci-Guyon, E., Portier, M., Dunia, I., Paulin, D., Pournin, S., Babinet, C. (1994) Mice lacking vimentin develop and reproduce without an obvious phenotype. *Cell.* **79,** 679-694.

Copeland, N.G., Jenkins, N.A., Gilbert, D.J., Eppig, J.T., Maltais, L.J., Miller, J.C., Dietrich, W.F., Weaver, A., Lincoln, S.E., Steen, R.G., Stein, L.D., Nadeau, J.H., Lander, E.S. (1993) A genetic linkage map of the mouse: current applications and future prospects. *Science.* **262,** 57-66.

Crosby, J.L., Varnum, D.S., Washburn, L.L., Nadeau, J.H. (1992) Disorganization is a completely dominant gain of function mouse mutation causing sporadic developmental defects. *Mechanisms of Development.* **37,** 121-126.

Danforth, C.H. (1930) Developmental anomalies in a special strain of mice. *American Journal of Anatomy.* **45,** 275-287.

Davies, J., Lyon, M., Gallagher, J., Garrod, D. (1995) Sulphated proteoglycan is required for collecting duct growth and branching but not nephron formation during kidney development. *Development.* **121,** 1507-1517.

De Gouyon, B., Melanitou, E., Richard, M.F., Requarth, M., Hahn, I.H., Guenet, JL., Demenais, F., Julier, C., Lathrop, G.M., Boitard, C., Avner, P. (1993) Genetic analysis of diabetes and insulitits in an interspecific cross of the nonobese diabetic mouse with *Mus spretus*. *Proceedings of the National Academy of Science, USA.* **90,** 1877-1881.

Dechiara, T.M., Efstradis, A., Robertson, E.J. (1990) A growth-deficient phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targetting. *Nature.* **345,** 78-80.

Deutsch, U., Dressler, G.R., Gruss, P. (1988) Pax-1, a member of a paired box homologous gene family, is expressed in segmented structures during development. *Cell.* **53,** 617-625.

Deutsch, U., Gruss, P. (1991) Murine paired domain proteins as regulatory factors of embryonic development. *Seminars in Developmental Biology.* **2,** 413-424.

Dietrich, S., Schubert, F.R., Gruss, P. (1993a) Altered Pax gene expression in murine notochord mutants: the notochord is required to initiate and maintain ventral identity in the somite. *Mechanisms of Development.* **44,** 189-207.

Dietrich, W.F., Lander, E.S., Smith, J.S., Moser, A.R., Gould, K.A., Luongo, C., Borenstein, N., Dove, W. (1993b) Genetic identification of *Mom-1*, a major modifier locus affecting *Min*-induced intestinal neoplasia in the mouse. *Cell.* **75,** 631-639.

Dietrich, W.F., Miller, J.C., Steen, R.G., Merchant, M., Damron, D., Nahf, R., Gross, A., Joyce, D.C., Wessel, M., Dredge, R.D., Marquis, A., Stein, L.D., Goodman, N., Page, D.C., Lander, E.S. (1994) A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Nature Genetics.* **7,** 220-245.

Dressler, G.R., Deutsch, U., Chowdhury, K., Norner, H.O., Gruss, P. (1990) *Pax2*, a new murine paired-box containing gene and its expression in the developing excretory system. *Development.* **109**, 787-795.

Dressler, G.R., Douglass, E.C. (1992) Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor. *Proceedings of the National Academy of Sciences, USA.* **89,** 1179-1183.

Dressler, G.R., Wilkinson, J.E., Rothenpieler, U.W., Patterson, L.T., Williams Simons, L., Westphal, H. (1993) Deregulation of Pax-2 expression in transgenic mice generates severe kidney abnormalities. *Nature.* **362,** 65-67.

Drummond, I.A., Madden, S.L., Rowher-Nutter, P., Bell, G.I., Sukhatme, V.P., Rauscher, F.J.I. (1992) Repression of the insulin-like growth factor II gene by the Wilm's tumour suppressor WT1. *Science.* **257,** 674-678.

Dunn, L.C., Gluecksohn-Schoenheimer, S., Bryson, V. (1940) A new mutation in the mouse affecting spinal column and urogenital system. *Journal of Heredity.* **31,** 343-348.

Dunn, L.C., Gluecksohn-Schoenheimer, S. (1945) Dominance modification and physiological effects of genes. *Proceedings of the National Academy of Sciences, USA.* **31,** 82-84.

Duyk, G.M., Kim, S.W., Myers, R.M., Cox, D.R. (1990) Exon trapping - a genetic screen to identify candidate transcribed sequences in cloned mammalian DNA. *Proceedings of the National Academy of Sciences, USA.* **87**, 8995-8999.

Edwards, M.D., Stuber, C.W., Wendel, J.F. (1987) Molecular marker-facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distribution and types of gene action. *Genetics.* **116,** 113-125.

Epstein, D.J., Vekemans, M., Gros, P. (1991) Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell.* **67,** 767-774.

European Backcross Collaborative Group. (1994) Towards high resolution maps of the mouse and human genomes - a facility for ordering markers to 0.1 cM resolution. *Human Molecular Genetics.* **3,** 621-627.

Feinberg, A.P., Vogelstein, A. (1983) A technique for radiolabelling DNA restriction fragments to high specific activity. *Annals of Biochemistry.* **132,** 6-13.

Feinberg, A.P., Vogelstein, A. (1984) A technique for radiolabelling DNA fragments to high specific activity. *Annals of Biochemistry.* **137,** 266-267.

Franke, W.W., Grund, C., Kuhn, C., Jackson, B.W., Illmensee, K. (1982) Formation of cytoskeletal elements during mouse embryogenesis. III. Primary mesenchymal cells and the first appearance of vimentin filaments. *Differentiation.* **23,** 43-59.

Gibson, F., Walsh, J., Mburu, P., Varela, A., Brown, K.A., Antonio, M., Beisel, K.W., Steel, K.P., Brown, S.D.M. (1995) A type VII myosin encoded by the mouse deafness gene *shaker-1*. *Nature.* **374,** 62-64.

Gluecksohn-Schoenheimer, S. (1943) The morphological manifestations of a dominant mutation in mice affecting tail and urogenital system. *Genetics.* **28,** 341-348.

Gluecksohn-Schoenheimer, S. (1945) The embryonic development of mutants - the *Sd* strain in mice. *Genetics.* **30,** 29-38.

Gluecksohn-Waelsch, S., Rota, T.R. (1963) Development in organ tissue culture of kidney rudiments from mutant mouse embryos. *Developmental Biology.* **7,** 432-444.

Goulding, M.D., Chalepakis, G., Deutsch, U., Erselius, J., Gruss, P. (1991) Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO Journal.* **10,** 1135-1147.

Goulding, M.D., Lumsden, A., Gruss, P. (1993) Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord. *Development.* **117,** 1001-1016.

Grobstein, C. (1955) The inductive interactions in the development of the mouse metanephros. *Journal of Experimental Zoology.* **130,** 319-340.

Grobstein, C. (1956) Trans-filter induction of tubules in mouse metanephric mesenchyme. *Experimental Cell Research.* **10**, 424-440.

Gruneberg, H. (1953) Genetical studies on the skeleton of the mouse VI. Danforth's short-tail. *Journal of Genetics.* **51,** 317-326.

Gruneberg, H. (1958) Genetical studies on the skeleton of the mouse XXII. The development of Danforth's short-tail. *Journal of Embryology and Experimental Morphology.* **6**, 124-148.

Gustinich, S., Caninci, P., Schneider, C. (1991) A rapid method to extract high-quality genomic DNA from yeast. *Technique.* **3,** 76-77.

Haley, C.S., Knott, S.A. (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity.* **69,** 315-324.

Harshmann, K., Bell, R., Rosenthal, J., Katcher, H., Miki, Y., Swenson, J., Gholami, Z., Fry, C., Ding, W., Dayananth, P., Eddington, K., Norris, F.H., Bristow, P.K., Phelps, R., Hattier, T., Stone, S., Shaffer, D., Bayer, S., Hussey, C., Tran, T., Richardson, K., Dehoff, B., Lai, M., Rosteck, P.R. Jr., Skolnick, M.H., Shattuck-Eidens, D., Kamb, A. (1995). Comparison of the positional cloning methods used to isolate the BRCA1 gene. *Human Molecular Genetics.* **4**, 1259-1266.

Hill, R.E., Favor, J., Hogan, B.L., Ton, C.C., Saunders, G.F., Hanson, I.M., Prosser, J., Jordan, T., Hastie, N.D., van Heyningen, V. (1991) Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature.* **354,** 522-525.

Hodgkinson, C.A., Moore, K.J., Nakayama, A., Steingrimsson, E., Copeland, N.G., Jenkins, N.A., Arnheiter, H. (1993) Mutations at the mouse micropthalmia locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein. *Cell.* **74,** 395-404.

Hummel, K.P., Chapman, D.B. (1959) Visceral inversion and associated anomalies in the mouse. *Journal of Heredity.* **50**, 9-23.

Hunter, R.L., Markert, C.L. (1957) Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. *Science.* **125,** 1294-1295.

Jackson, I.J. (1988) A cDNA encoding tyrosinase-related protein maps to the brown locus in mouse. *Proceedings of the National Academy of Sciences, USA.* **85,** 4392-4396.

Jessel, T.M., Boloventa, P., Placzek, M., Tessier-Lavigne, M., Dodd, J. (1989) Polarity and patterning in the neural tube: the origin and function of the floor plate. *Ciba Foundation Symposium.* **144,** 255-280.

Jones, C., Janson, M., Nordenskjold, M. (1989) Separation of yeast chromosomes in the megabase range suitable for size markers for pulse field gel electrophoresis. *Technique.* **1,** 90-95.

Jordan, T., Hanson, I., Zaletayev, D., Hodgson, S., Prosser, J., Seawright, A., Hastie, N., van Heyningen, V. (1992) The human PAX6 gene is mutated in two patients with aniridia. *Nature Genetics.* **1,** 328-332.

Joyner, A.L. (1991) Gene targetting and gene trap screens using embryonic stem cells: New approaches to mammalian development. *Bioessays.* **13,** 649-656.

Kitchin, I.C. (1949) The effects of notochordectomy in *amblystoma mexicanum*. *Journal of Experimental Zoology.* **112,** 393-415.

Kjellen, L., Lindahl, U. (1995) Proteoglycans: structure and functions. *Annual Review of Biochemistry.* **60,** 443-475.

Koopman, P., Gubbay, J., Vibian, N., Goodfellow, P., Lovell-Badge, R. (1991) Male development of chromosomally female mice transgenic for *sry*. *Nature.* **351,** 377-383.

Koseki, H., Wallin, J., Wilting, J., Mizutani, Y., Kispert, A., Ebensperger, C., Herrmann, B.G., Christ, B., Balling, R. (1993a) A role for Pax-1 as a mediator of notochordal signals during the dorsoventral specification of vertebrae. *Development.* **119,** 649-660.

Koseki, H., Zachgo, J., Mizutani, Y., Simons-Chazottes, D., Guenet, J., Balling, R., Gossler, A. (1993b) Fine genetic mapping of the proximal part of mouse chromosome 2 excludes *Pax8* as a candidate for Danforth's short-tail (*Sd*). *Mammalian Genome.* **4,** 324-327.

Kozak, C.A., Peyser, M., Krall, M., Mariano, T.M., Kumar, C.S., Pestka, S., Mock, B.A. (1990) Molecular genetic markers spanning chromosome 10. *Genomics.* **8,** 519-524.

Kratochwil, K., Dziadek, M., Lohler, J., Harbers, K., Jaenisch, R. (1986) Normal epithelial branching morphogenesis in the absence of collagen I. *Developmental Biology.* **117,** 596-606.

Kriedberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., Jaenisch, R. (1993) *WT-1* is required for early kidney development. *Cell.* **74,** 679-691.

Kusumi, K., Smith, J.S., Segre, J.A., Koos, D.S., Lander, E.S. (1993) Construction of a large-insert yeast artificial chromosome library of the mouse genome. *Mammalian Genome.* **4,** 391-392.

Lande, R. (1981) The minimum number of genes contributing to quantitative variation between and within populations. *Genetics.* **99**, 541-553.

Lander, E.S., Botstein, D. (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics.* **121,** 185-199.

Lander, E.S., Schork, N.J. (1994) Genetic dissection of complex traits. *Science.* **265,** 2037-2048.

Larin, Z., Monaco, A.P., Lehrach, H. (1991) Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proceedings of the National Academy of Sciences, USA.* **88,** 4123-4127.

Lavia, P., Macleod, D., Bird, A. (1987) Coincident start sites for divergent transcripts at a randomly selected CpG-rich island of mouse. *EMBO Journal.* **6**, 2773-2779.

Levin, M., Johnson, R.L., Stern, C.D., Kuehn, M., Tabin, C. (1995) A molecular pathway determining left-right assymmetry in chick embryogenesis. *Cell.* **82**, 803-814.

Lin, H.Y., Wang, X., Ng-Eaton, E., Weinberg, R.A., Lodish, H.L. (1992) Expression cloning of the TGF-beta type II receptor, a functional transmembrane serine/threonine kinase. *Cell.* **68,** 775-785.

Link, A.J., Olson, M.V. (1991) Physical map of the *Saccharomyces cerevisiae* genome at 110kb resolution. *Genetics.* **127,** 681-698.

Liu, J.P., Barker, J., Perkins, E.J., Robertson, E.J., Efstradis, A. (1993a) Mice carrying null mutations of the genes encoding insulin-like growth factor (Igf-1) and type 1 IGF receptor (Igf1r). *Cell.* **75,** 59-72.

Liu, L., Dunn, T., Christiakos, S., Hanson-Paintin, O., Bourdeau, J.E. (1993b) Calbindin-D28K gene expression in the developing mouse kidney. *Kidney International.* **44,** 322-330.

Love, J.M., Knight, A.M., McAleer, M., Todd, J.A. (1990) Towards construction of a high resolution map of the mouse genome using PCR-analysed microsatellites. *Nucleic Acids Research.* **18,** 4123-4130.

Lovett, M., Kere, J., Hinton, L.M. (1991) Direct selection: A method for the isolation of cDNAs encoded by large genomic regions. *Proceedings of the National Academy of Sciences, USA.* **88,** 9628-9632.

MacPhee, M., Chepenik, K.P., Liddell, R.A., Nelson, K.K., Siracusa, L.D., Buchberg, A.M. (1995) The secretory phospholipase A2 gene is a candidate for the *Mom1* locus, a major modifier of Apc^{Min}-induced intestinal neoplasia. *Cell.* **81,** 957-966.

Mann, H.B., Whitney, D.R. (1947) On a test of whether one or two random variables is stochastically larger than the other. *Annals of Mathematics and Statistics.* **18**, 50-60.

Matthews, L.S., Vale, W.W. (1991) Expression cloning of an activin receptor, a predicted transmembrane serine kinase. *Cell.* **65,** 973-982.

Maule J. C. (1994) Electrophoretic karyotype analysis: pulsed field gel electrophoresis. In: *Chromosome Analysis Protocols* (Ed. Gosden J. R.). Humana Press, pp. 221-52.

McPherson, E., Carey, J., Kramer, A., Hall, J.G., Pauli, R.M., Schimke, R.N., Tasin, M.H. (1987) Dominantly inherited renal adysplasia. *American Journal of Medical Genetics.* **26,** 863-872.

Mesrobian, H.G.J., Sulik, K.K. (1992) Characterization of the upper urinary tract anatomy in the Danforth spontaneous murine mutation. *Journal of Urology.* **148,** 752-755.

Miyagawa, K., Kent, J., Schedl, A., van Heyningen, V., Hastie, N.D. (1994) WT1- a case of disrupted development. *Journal of Cell Science.* **S18**, 1-5.

Montesano, R., Matsumoto, K., Nakamuta, T., Orci, L. (1991) Identification of a fibroblast-derived morphogen as hepatocyte growth factor. *Cell.* **67,** 901-908.

Moser, A.R., Pitot, H. C., Dove, W.F. (1990) A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science.* 247: 322-324.

Mullis, K.B., Faloona, F.A. (1987) Specific synthesis of DNA *in vitro* via a polymerase catalysed chain reaction. *Methods of Enzymology.* **155,** 335-350.

Nehls, M., Pfiefer, D., Boehm, T. (1994) Exon amplification from complete libraries of genomic DNA using a novel phage vector with automatic plasmid excision facility: application to the mouse neurofibromatosis-1 locus. *Oncogene.* **9**, 2169-2175.

Nornes, H.O., Dressler, G.R., Knapik, E.W., Deutsch, U., Gruss, P. (1990) Spatially and temporally restricted expression of *Pax2* during murine neurogenesis. *Development.* **109,** 797-809.

Pachnis, V., Mankoo, B., Costantini, F. (1993) Expression of the c-*ret* proto-oncogene during mouse embryogenesis. *Development.* **119,** 1005-1017.

Parimoo, S., Kolluri, R., Weissman, S.M. (1993). cDNA selection from total yeast DNA containing YACs. *Nucleic Acids Research.* **21**, 4422-4423.

Patterson, L.T., Dressler, G.R. (1994) The regulation of kidney development: new insights from an old model. *Current Opinions in Genetics and Development.* **4**, 696-702.

Phelps, D.E., Dressler, G.R. (1993) Aberrant expression of *Pax-2* in Danforth's short tail (*Sd*) mice. *Developmental Biology.* **157,** 251-258.

Pierce, J.C., Sauer, B., Sternberg, N. (1992) A positive selection vector for cloning high molecular weight DNA by the bacteriophage P1 system: Improved cloning efficiency. *Proceedings of the National Academy of Sciences, USA.* **89**, 2056-2060.

Pivnick, E.K., Opp, J.S., Muller, U., Simpson, J.L., Wachtel, S., Bishop, C.E. (1991) Correlation between mutations in the human SRY gene and 46, XY pure gonadal dysgenesis. *American Journal of Human Genetics.* **49,** 396

Plachov, D., Chowdhury, K., Walther, C., Simon, D., Guenet, J., Gruss, P. (1990) *Pax-8*, a murine paired box gene expressed in the developing excretory system and thyroid gland. *Development.* **110,** 643-651.

Placzek, M., Yamada, T., Tessier Lavigne, M., Jessell, T., Dodd, J. (1991) Control of dorsoventral pattern in vertebrate neural development: induction and polarizing properties of the floor plate. *Development.* **Suppl 2,** 105-122.

Platt, J.L., Trescony, P., Lindmann, B., Oegama, T. (1990) Heparin and heparin sulphate delimit nephron formation in foetal metanephric kidneys. *Developmental Biology.* **139,** 338-348.

Powell-Braxton, L., Hollingshead, P., Warburton, C., Dowd, M., Pitts-Meek, S., Dalton, D., Gillet, N., Stewart, T.A. (1993) IGF-1 is required for normal embryonic growth in mice. *Genes and Development.* **7,** 2609-2617.

Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J.C., Markham, A.F. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosomes. *Nucleic Acids Research.* **18,** 2887-2830.

Riordan, J.R., Rommens, J.M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J., Drumm, M.L., Iannizzi, M., Collins, F.S., Tsui, L. (1989) Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science.* **245,** 1066-1073.

Rise, M.L., Frankel, W.N., Coffin, J.M., Seyfried, T.N. (1991) Genes for epilepsy mapped in the mouse. *Science.* **253,** 669-673.

Roderick, T.H., Hillyard, A.L. (1989) Differences in recombination due to sex in mice. *Mouse News Letter.* **85,** 87-89.

Rommens, J.M., Iannuzzi, M., Kerem, B., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J.R., Tsui, L., Collins, F.S. (1989) Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science.* **245,** 1059-1069.

Rothenpieler, U.W., Dressler, G.R. (1993) Pax2 is required for mesenchyme to epithelium conversion during kidney development. *Development.* **119,** 711-720.

Ryan, S.G., Buckwalter, M.S., Lynch, J.W., Handford, C.A., Segura, L., Shiang, R., Wasmuth, J.J., Camper, S.A., Schofield, P., O'Connell, P. (1994) A missense mutation in the gene encoding the alpha-1 subunit of the inhibitory glycine receptor in the spasmodic mouse. *Nature Genetics.* **7,** 131-135.

Ryan, G., Steele-Perkins, V., Morris, J.F., Rauscher, F.J.I., Dressler, G.R. (1995) Repression of *Pax2* by *WT-1* during normal kidney development. *Development.* **121,** 867-875.

Saga, Y., Yagi, T., Ikawa, Y., Sakakura, Y., Aizawa, S. (1992) Mice develop normally without tenascin. *Genes and Development.* **6,** 1821-1831.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Schart, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science.* **239,** 487-491.

Sambrook J, Fritsch EF, Maniatis T. (1989) Molecular Cloning: *A Laboratory Manual.* Second ed. Cold Spring Harbour Laboratory Press.

Santos, O.F.P., Moura, L.A., Rosen, E.M., Nigam, S.K. (1993) Modulation of HGF-induced tubulogenesis and branching by multiple phosphorylation mechanisms. *Developmental Biology.* **159,** 535-548.

Santos, O.F.P., Nigam, S.K. (1993) HGF-induced tubulogenesis and branching of epithelial cells is modulated by extracellular matrix and TGF-beta. *Developmental Biology.* **160,** 293-302.

Santos, O.F.P., Barros, E.J.G., Yang, X., Matsumoto, K., Nakamura, T., Park, M., Nigram, S.K. (1994) Involvement of hepatocyte growth factor in kidney development. *Developmental Biology.* **159,** 535-548.

Sax, K. (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris. Genetics.* **8,** 552-560.

Saxen L. (1987) *Organogenesis of the kidney.* (Eds. Barlow P. W., Green P. B., and White C. C.). First ed. Cambridge University Press.

Schedl, A., Larin, Z., Montoliu, L., Thies, E., Kelsey, G., Lehrach, H., Schutz, G. (1993) A method for the generation of YAC transgenic mice by pronuclear microinjection. *Nucleic Acids Research.* **21,** 4783-4787.

Schmidt, C., Bladt, F., Goedecke, S., Brinkman, V., Zschiesche, W., Sharpe, M., Gheradi, E., Birchmeier, C. (1995) Scatter factor / hepatocyte growth factor is essential for liver development. *Nature.* **373,** 699-702.

Schuchardt, A., D'Agati, V., Larsson-Bloomberg, L., Costantini, F., Pachnis, V. (1994) Defects in the kidney and enteric system of mice lacking the tyrosinase receptor Ret. *Nature.* **367,** 380-383.

Sealey, P.G., Whittaker, P.A., Southern, E.M. (1985) Removal of repeat sequences from hybridisation probes. *Nucleic Acids Research.* **13,** 1905-1912.

Serikawa, T., Kuramoto, T., Hilbert, P., Mori, M., Yamada, J., Dubay, C.G., Lindpainter, K., Ganten, D., Guenet, J., Lathrop, G.M., Beckmann, J.S. (1992) Rat gene mapping using PCR-analyzed microsatellites. *Genetics.* **131,** 701-721.

Shizuya, H., Birren, B., Kim, U-J., Mancino, V., Slepak, T., Tachiri, Y., Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F factor based vector. *Proceedings of the National Academy of Sciences, USA.* **89**, 8794-8797.

Siracusa, L., Abbott, C.M. (1994) Mouse chromosome 2. *Mammalian Genome.* **5,** S22-S39.

Sonnenberg, E., Meyer, D., Weidner, K.M., Birchmeier, C. (1993) Scatter factor/hepatocyte growth factor and its receptor, the c-met tyrosine kinase, can mediate a signal exchange between mesenchyme and epithelia during mouse development. *Journal of Cell Biology.* **123,** 223-235.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology.* **98,** 503-517.

Stark, K., Vainio, S., Vassileva, G., McMahon, A.P. (1994) Epithelial transformation of metanephric mesenchyme in the developing kidney regulated by *Wnt-4. Nature.* **372,** 679-683.

Tachibana, M., Perez-Jurado, L., Nakayama, A., Hodgkinson, C.A., Li, X., Schneider, M., Miki, T., Fex, J., Francke, U., Arnheiter, H. (1994) Cloning of MITF, the human homolog of the mouse micropthalmia gene and assignment to chromosome 3p14.1-p12.3. *Human Molecular Genetics.* **3,** 553-557.

Tanksley, S.D. (1993) Mapping polygenes. *Annual Review of Genetics.* **27,** 205-233.

Tassabehji, M., Read, A.P., Newton, V.E., Harris, R., Balling, R., Gruss, P., Strachan, T. (1992) Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature.* **355,** 635-636.

Tassabehji, M., Newton, V., Read, A.P. (1994) Waardenburg syndrome type 2 caused by mutations in the human micropthalmia (MITF) gene. *Nature Genetics.* **8,** 251-255.

Taylor B. A. (1989) *Origins of Inbred Mice.* (Ed. H. C. Morse). Academic Press, pp. 423-40.

Taylor G.R. (1991) Polymerase chain reaction: basic principles and automation. In: *PCR. A Practical Approach* (Eds. McPherson M.J., Quirke P., Taylor G.R) Oxford University Press. pp.1-14.

Taylor, B.A., Navin, A., Phillips, S.J. (1994) PCR-amplification of simple sequence repeat variants from pooled DNA samples for rapidly mapping new mutations of the mouse. *Genomics.* **21,** 626-632.

Teillet, M.A., Le Douarin, N.M. (1983) Consequences of neural tube and notochord excision on the development of the peripheral nervous system in the chick embryo. *Developmental Biology.* **98,** 192-211.

Tessier-Lavigne, M., Placzek, M., Lumsden, A., Dodd, J., Jessell, T.M. (1988) Chemotropic guidance of developing axons in the mammalian central nervous system. *Nature.* **336,** 775-778.

Theiler, K. (1959) Anatomy and development of the truncate (boneless) mutation in the mouse. *American Journal of Anatomy.* **104,** 319-343.

Thoday, J.M. (1961) Location of polygenes. *Nature.* **191,** 368-370.

Todd, J.A., Aitman, T.J., Cornall, R.J., Ghosh, S., Hall, J.R., Hearne, C.M., Knight, A.M., Love, J.M., McAleer, M.A., Prins, J.B., Lathrop, M., Peterson, L., Wicker, L. (1991) Genetic analysis of autoimmune type 1 diabetes mellitus in mice. *Nature.* **351,** 542-547.

Tsarfaty, I., Rong, S., Resau, J.H., Rulong, S., Pinto da Silva, P., Vande Woude, G.F. (1994) The met proto-oncogene mesenchymal to epithelial cell conversion. *Science.* **263,** 98-101.

Uehara, Y., Minowa, O., Mori, C., Shiota, K., Kuno, J., Noda, T., Kitamura, N. (1995) Placental defect and embryonic lethality in mice lacking hepatocyte growth factor / scatter factor. *Nature.* **373,** 702-705.

Vainio, S., Lehtonen, E., Jalkanen, M., Bernfield, M., Saxen, L. (1989) Epithelial-mesenchymal interactions regulate the stage-specific expression of cell surface proteoglycan, syndecan, in the developing kidney. *Developmental Biology.* **134,** 283-391.

Vainio, S., Jalkanen, M., Bernfield, M., Saxen, L. (1992) Transient expression of syndecan in mesenchymal cell aggregates of the embryonic kidney. *Developmental Biology.* **152,** 221-232.

Valerio, D., Duyvesteyn, M.G.C., Dekker, B.M.M., Weeda, G., Berkvens, T.M., van der Voorn, L., van Ormondt, H., van der Eb, A.J. (1985) Adenosine deaminase: characterization and expression of a gene with a remarkable promoter. *EMBO Journal.* **4,** 437-443.

Wada, J., Liu, Z.Z., Alvares, K., Kumar, A., Wallner, E., Makino, H., Kanwar, Y.S. (1993) Cloning of cDNA for the alpha subunit of mouse insulin-like growth factor I receptor and the role of the receptor in metanephric development. *Proceedings of the National Academy of Sciences, USA.* **90,** 10360-10364.

Weil, D., Blanchard, S., Kaplan, J., Guilford, P., Gibson, F., Walsh, J., Mburu, P., Varela, A., Levilliers, J., Weston, M.D., Kelley, P.M., Kimberling, W.J., Wagenaar, M., Levi-Acobas, F., Larget-Pier, D., Munnich, A., Steel, K.P., Brown, S.D.M., Petit, C. (1995) Defective myosin VIIA gene responsible for Usher syndrome type 1B. *Nature.* **374,** 60-61.

Weissenbach, J., Gyapay, G., Dib, C., Vignal.A., Morisette, J., Millasseau, P., Vaysseix, G., Lathrop, M. (1992) A second-generation linkage map of the human genome. *Nature.* **359,** 794-801.

Wilson, D.B., Finta, L.A., Center, E.M., Paavola, L.G. (1982) An electron microscopic analysis of notochordal and mesenchymal cell abnormalities in embryos of Danforth's short-tail *(Sd)* mice. *Virchows Archives [Cell Pathology].* **39,** 101-110.

Winship, P.R. (1989) An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucleic Acids Research.* **19,** 1266-1226.

Wilkie, A.O.R. (1994) The molecular basis of genetic dominance. *Journal of Medical Genetics.* **31**, 89-98.

Wright, S. (1968) The genetics of quantitative variability. In: *Evolution and the Genetics of Populations: Genetic and Biometric Foundations.* University of Chicago Press. pp.373-420.

Yamada, T., Placzek, M., Tanaka, H., Dodd, J., Jessell, T.M. (1991) Control of cell pattern in the developing nervous system: polarizing activity of the floor plate and notochord. *Cell.* **64,** 635-647.

Yokoyama, T., Copeland, N.G., Jenkins, N.A., Montgomery, C.A., Elder, F.F., Overbeek, P.A. (1993). Reversal of left-right assymmetry: a situs inversus mutation. *Science.* **260**, 679-682.

Yost, H.J. (1995). Vertebrate left-right development. *Cell.* **82**, 689-692.

Youn, B.W., Malacinski, G. (1981) Axial structure development in ultraviolet-irradiated (notochord defective) amphibian embryos. *Developmental Biology.* **83,** 339-352.

Yun, K., Molenaar, A.J., Fiddler, A.M., Mark, A.J., Eccles, M.R., Becroft, D.M., Reeve, A.E. (1993) Insulin-like growth factor II messenger ribonucleic acid expression in Wilm's tumour, nephrogenic rest, and kidney. *Laboratory Investigation.* **69,** 603-615.

Zahedi, K., Seldin, M.F., Rits, M., Ezekowitz, R.A.B., Whitehead, A.S. (1991) Mouse IL-1 receptor antagonist protein. *Journal of Immunology.* **146,** 4228-4223.

Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., Friedman, J.M. (1994) Positional cloning of the mouse *obese* gene and its human homologue. *Nature.* **372**, 425-432.

Zhou, X., Sasaki, H., Lowe, L., Hogan, B.L.M., Kuehn, M.R. (1993) Nodal is a novel TGF-beta-like gene expressed in the mouse node during gastrulation. *Nature.* **361,** 543-545.

APPENDICES

Appendix A

Data for maximum likelihood analysis of chromosome 10 mapping data from *Sd+* backcross progeny. N=152

Additive value	Standard error	Domi-nance	Standard error	RSS_{red}	RSS_{full}	Test statistic
-0.0815	0.3133	-0.178	0.286	96.97	108.7	7.53801
-0.0820	0.3193	-0.105	0.309	96.79	108.7	7.66066
-0.0825	0.3269	-0.008	0.335	96.49	108.7	7.86558
-0.0829	0.3365	0.119	0.363	96.04	108.7	8.17416
-0.0833	0.3482	0.280	0.393	95.37	108.7	8.6363
-0.0836	0.3620	0.474	0.421	0.421	108.7	9.27619
-0.0836	0.3767	0.689	0.444	93.28	108.7	10.099
-0.0833	0.3903	0.897	0.460	91.96	108.7	11.0398
-0.0827	0.4002	1.059	0.463	90.67	108.7	11.9724
-0.0818	0.4043	1.144	0.454	89.64	108.7	12.7266
-0.0808	0.4024	1.144	0.435	89.00	108.7	13.1996
-0.0799	0.3961	1.078	0.410	88.77	108.7	13.3704
-0.0791	0.3873	0.973	0.382	88.85	108.7	13.311
-0.0785	0.3778	0.854	0.353	89.12	108.7	13.1107
-0.0816	0.4158	0.229	0.325	89.73	108.7	12.6604
-0.0820	0.4208	0.339	0.368	89.53	108.7	12.8077
-0.0825	0.4261	0.465	0.418	89.37	108.7	12.9257
-0.0830	0.4303	0.587	0.472	89.31	108.7	12.9701
-0.0834	0.4309	0.654	0.520	89.50	108.7	12.8298
-0.0836	0.4252	0.601	0.546	90.02	108.7	12.4474
-0.0833	0.4139	0.416	0.539	90.77	108.7	11.8997
-0.0827	0.4011	0.176	0.501	91.47	108.7	11.3925
-0.0820	0.3908	0.036	0.451	91.95	108.7	11.047
-0.0828	0.3635	0.629	0.400	92.97	108.7	10.3188
-0.0834	0.3638	0.741	0.470	93.39	108.7	10.0212
-0.0840	0.3611	0.808	0.555	94.13	108.7	9.50023
-0.0846	0.3524	0.725	0.645	95.26	108.7	8.71248
-0.0847	0.3363	0.373	0.706	96.53	108.7	7.83822
-0.0841	0.3189	0.172	0.700	97.24	108.7	7.35446

Appendix B

Data for maximum likelihood analysis of chromosome 10 mapping data from WT backcross progeny. N=148

Additive value	Standard error	Domi-nance	Standard error	RSS_{red}	RSS_{full}	Test statistic
-0.0873	0.0458	-0.240	0.233	28.50	29.56	2.34722
-0.0864	0.0460	-0.269	0.269	28.49	29.56	2.36977
-0.0854	0.0463	-0.296	0.309	28.49	29.56	2.36977
-0.0850	0.0466	-0.307	0.342	28.50	29.56	2.34722
-0.0856	0.0468	-0.283	0.354	28.53	29.56	2.27959
-0.0873	0.0467	-0.228	0.336	28.56	29.56	2.21204
-0.0891	0.0465	-0.171	0.300	28.58	29.56	2.16705
-0.0905	0.0462	-0.130	0.260	28.59	29.56	2.14456
-0.0839	0.0467	0.094	0.237	28.76	29.56	1.7635
-0.0841	0.0468	0.110	0.271	28.78	29.56	1.71882
-0.0835	0.0470	0.0470	0.109	0.304	29.56	1.67417
-0.0820	0.0470	0.078	0.324	28.84	29.56	1.58496
-0.0797	0.0468	0.021	0.320	28.87	29.56	1.51813
-0.0778	0.0465	-0.040	0.296	28.89	29.56	1.47362
-0.0765	0.0462	-0.084	0.262	28.89	29.56	1.47362
-0.0572	0.0460	0.221	0.185	28.85	29.56	1.56267
-0.0590	0.0462	0.263	0.205	28.81	29.56	1.65185
-0.0609	0.0464	0.308	0.227	28.77	29.56	1.74116
-0.0625	0.0466	0.352	0.249	28.74	29.56	1.80821
-0.0632	0.0468	0.382	0.270	28.75	29.56	1.78585
-0.0624	0.0469	0.383	0.283	28.80	29.56	1.67417
-0.0598	0.0468	0.346	0.284	28.90	29.56	1.45137
-0.0561	0.0466	0.280	0.275	29.02	29.56	1.18504
-0.0522	0.0463	0.204	0.256	29.13	29.56	0.941864
-0.0489	0.0461	0.135	0.235	29.22	29.56	0.743584
-0.0464	0.0458	0.080	0.213	29.27	29.56	0.633693

Appendix C

Mann Whitney U Test (Mann and Whitney, 1947)

The following formulae were used:

$$U_1 = n_1 n_2 + [n_1 (n1 + 1) / 2] - R_1$$

Where: n_1 = the number of cases in the smaller group

n_2 = the number of cases in the larger group

R_1 = the sum of the ranks assigned to the group of size n_1

$$U_2 = n_1 n_2 - U_1$$

The smallest of U_1 and U_2 was denoted by U. The significance level associated with U was looked up in the table appropriate for the values of n_1 and n_2.

Appendix D

F Test Calculation

The F test used in chapter 5 was calculated by A. Carrothers at the MRC Human Genetics Unit Edinburgh. The Minitab statistical program (Minitab Inc.) was used to calculate standard deviations.

Values used in the calculation of the F test were as follows:

Phenotype	Number	Mean tail length (cm)	Standard deviation
Sd	147	1.881	0.9755
WT	144	6.975	0.6091
All	291	2.678	2.6779

The data is given in rows. Only 291 mice were used for this calculation as the backcross was not complete when this analysis was performed.

The significance value was calculated by squaring the ratio of the standard deviations for the WT and *Sd*+ mice. From this calculation the value of 2.56 was obtained. With 146 and 143 degrees of freedom this gave a probability value of $p < 0.001$.

Appendix E

The percentage contribution which the chromosome 10 modifier loci has on overall tail length variation was assessed by calculating the effect that the loci has on the mean tail length.

% of the mean = additive effect at maximum point / mean for *Sd* tail lengths

= (0.3961 / 1.87) x 100 = 21.1%

% of the mean = 21.1%

[see appendix A for additive effect at maximum point]