Male-Specific Transcripts From *Drosophila melanogaster.*

A thesis submitted for the degree of
Doctor of Philosophy at the University of Glasgow.

by

Diane T. Harbison

Institute of Genetics
University of Glasgow
Church Street
Glasgow
G11-5JS

May 1996.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

The research reported within this thesis is my own work except where otherwise stated and has not been submitted for any other degree.

Diane T. Harbison

For my Mum and Dad,
but most of all for Sean.

Acknowledgements.

I would like to thank all members of the fifth floor past and present: Colin and Peggy for keeping things going when I was otherwise detained; Ron, Alan, Simon and Andy for all their support, help and encourgement and Stephen and Audrey for their friendship both in and out of the lab. Also to all the other members of the department, especially Dave and Jen, for making this such a good place to work in.

Thanks to Tony Dornan for all his help and advice on *in situ* hybridisations and biotin probes. Thanks also to Douglas Armstrong for scanning my whole mounts into the confocal microsope and to Zong Sheng Wang for the polytene *in situs*. A very large thank you to all the media ladies without whom this would have been a lot harder and a lot less fun.

I am also extremely grateful to Professor Mary Bownes for her continued support and encouragement and to Kathleen, Anji and the rest of the Bownes group for their friendship and support.

I would like to thank Graeme Brown and Frank Johnstone from the photography department at The University of Edinburgh and all the Staff at the University Nursery for doing their best to tire out Sean.

Thanks to Fiona and to Janie for their continued friendship over all these years!

Thanks to Caroline for looking after Sean and I. An extremely large thank you to Colin for all his help with this thesis; for helping to draw the diagrams; proof-reading; keeping me supplied with coffee and chocolate and walking me home in the early hours of the morning.

By far the largest vote of thanks is for Mum, Dad, Karen and Suzanne for their continual support and encouragement. Most of all, I am indebted to Karen, for all her babysitting, which helped to make this task much easier.

This work was initially supported by the Medical Research Council. Thanks to the Carnegie Trust for assistance with fees.

Abbreviations.

Å	- Angstrom
ATP	- adenosine triphosphate
BSA	- bovine serum albumin
bp	- base pair
bis	- N, N'-methylenebisacrylamide
cDNA	- complementary DNA
Ci	- Curie
cm	- centimetre
cpm	- counts per minute
°C	- degree centigrade
DAB	- diaminobenzidine
DEPC	-diethylpyrocarbonate
DNA	- deoxyribonucleic acid
DNP	- deoxyribonucleoproteins
d(N6)	- random hexanucleotides
dNTP	- deoxyribonucleotide triphosphate
DTT	- dithiothreitol
EDTA	- ethylene diamine tetra-acetic acid
EthBr	- ethidium bromide
FITC	- fluorescein-5-isothiocyanate
g	- gram
g	- centrifugal force equivalent to gravitational acceleration
gK	- genomic Lambda DNA clone
GCG	- Genetics Computing Group (University of Wisconsin)
hr	- hour
IPTG	- isopropyl-b-D-thiogalactopyranoside
kb	- kilobase/kilobasepair
kD	- kilodaltons
l	- litre
L1, L2, L3	- 1st, 2nd, 3rd instar larvae
LMP	- low melting point
M	- molar
min	- minute
ml	- millilitre
mM	- millimolar

mRNA	- messenger ribonucleic acid
$NaPP_i$	- sodium pyrophosphate
ng	- nanogram
nm	- nanometre
nt	- nucleotide
OD	- optical density
pfu	- plaque-forming unit
pH	- [$-\log_{10}$ (molar concentration of H^+ ions)]
P[Z]	- P lacZ enhancer-trap element
RF	- RNase free
rpm	- revolutions per minute
rRNA	- ribosomal ribonucleic acid
RNP	- ribonucleoproteins
SDS	- sodium dodecylsulphate
TEMED	- N, N, N' N', -tetramethylenediamine
Tris	- tris (hydroxyl methyl) amino ethane
U	- unit
µCi	- microcurie
µg	- microgram
µl	- microlitre
UV	- ultraviolet
v	- volt
vol	- volume
wt	- wild type
X-gal	- 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Summary

Eight genomic DNA clones previously identified in a differential screen to identify male-specific transcripts (Russell, 1989) were selected for analysis. These recombinant phage appear to represent six genomic loci. Three of these genomic clones gK22, gK30 and gK31 were identified once while others (gK6 & gK27 and gK8 & gK10) represented genomic DNA phage which originated from the same genomic region. Clone gK17 appeared to be identical to clone gS4 characterised by Russell (1993) and was not analysed any further. The relationship between gK6 & gK27 and gK8 & gK10 was confirmed by cross-hybridisation and restriction mapping. The longest clone from each of the two classes (gK10 and gK27) were selected for further analysis.

The data obtained from Southern analysis and *in situ* hybridisation to polytene chromosomes suggests that three (gK10, gK22, & gK31) of the genomic DNA clones which were selected for further characterisation represent single copy genes located on the autosomes. Genomic DNA clone gK27 appears (from Southern and sequence analysis) to contain a duplication of the fragment which contains the putative male-specific transcript. None of them appear to be re-isolates of previously characterised male-sterile or male-lethal loci.

Some of the male-specific transcripts are expressed in the larval testis and in the case of two of the transcripts represented by gK27 and gK31, contain transcripts that are expressed in the adult testis. Restriction of the transcripts to the adult testis has been confirmed by northern analysis for two of the clones gK10 and gK27. In both cases, transcripts were detected only in males and not in females. Additionally, transcripts derived from clones gK10 and gK27 were not detected in *tudor* flies nor in female flies transformed into pseudomales by mutations in the *transformer*2 gene. These results suggest a germline-specific function for these transcripts.

Two P[*lac*Z] insertion lines recovered during a screen for spermatogenic mutatations (Castrillon *et al.*; 1993) mapped near the cytological locations of two of the genomic DNA clones gK27 (insertion line designated MS(3)90E) and gK22 (designated *pistachio*, (*pto*)). Insertions at these locations resulted in sterility (*pto*) and in sterility and semi-lethality (MS(3)90E) respectively.

cDNA clones corresponding to the male-specific regions of genomic DNA clones gK31 and gK27 were isolated and the DNA sequenced. The cDNA corresponding to genomic DNA clone gK31 (cK31) showed no apparent significant homology with other sequences in the SwissProt database, whilst cDNA corresponding to genomic DNA clone gK27 (cK27) showed significant homology to a group of proteins known as the HMG-box proteins. The HMG-box is a DNA binding domain that was first identified in a group of proteins that comprised the non-histone component of chromatin (The High Mobility Group or HMG proteins). As cK27 appeared from northerns not to represent a full length transcript, the two male-specific genomic regions from the original genomic DNA clone gK27 were sequenced. This analysis suggested that the presence of two male-specific regions in the gK27 genomic DNA clone had arisen due to a duplication of this region. cDNAs corresponding to the other two clones(gK10 and gK22) have been isolated but have not yet been characterised.

The four genomic clones selected for further characterisation in this study appear to represent a class of germ-line specific transcripts. It is possible that they may be involved in spermatogenesis.

Table of Contents

Chapter One Introduction

Introduction.

Spermatogenesis is the process of producing haploid sperm cells from predetermined diploid cells. Immense morphological changes are specified that form complex, highly specialised motile sperm from an undifferentiated germ cell. An essential component of spermatogenesis is the process of meiosis, ensuring the production of haploid sperm cells. Spermatogenesis is one of the fundamental developmental processes that occurs in higher eukaryotes. It is also one of the most elaborate and highly conserved processes that occurs during development. The few differences that are known to exist between insects and mammals relate more to differences in their regulation of spermatogenesis rather than to differences in the basic morphological features.

The formation of the male gamete from an undifferentiated germ cell is a process that begins with a reduction in the chromosome number and proceeds through various morphological stages converting the haploid spermatid into a fully differentiated spermatozoon. The composition of the mature sperm is highly specialised enabling the sperm to carry out its two major functions (transport of the sperm nucleus to the egg and subsequent fusion with the egg) with the maximum efficiency. As can be seen in Figure 1.1, the sperm can be divided into two components, the **head** consisting of the **nucleus** (N) and **acrosome** (A) and the **tail,** formed from the **flagella** (or **axial complex** AC) and a fused **mitochondrion** (also known as the **nebenkern** (N) or **mitochondrial derivative** (MD)).

As described in more detail later in this chapter, all stages of spermatocyte and spermatid development occur within the adult testis. Every 10 hours, an asymmetric division of a **stem cell** (S) at the apex of the testis results in the formation of a **gonial stem cell** (G) and a new stem cell (Figure 1.2). Once

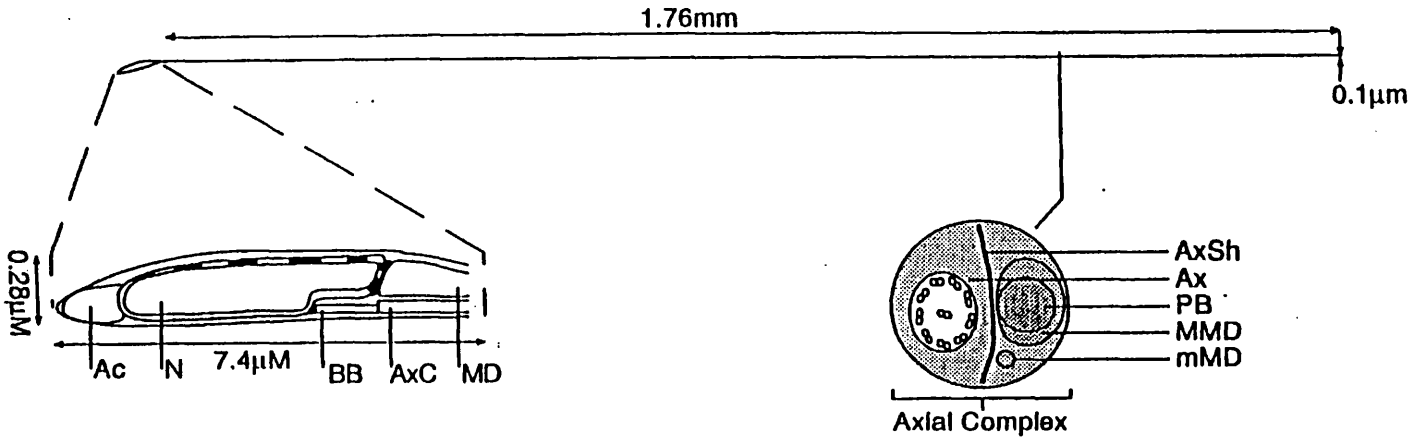


Figure 1.1 is a schematic representation of a mature sperm. The cross section through the axial complex shows the axoneme surrounded by the plasma membrane. The major mitochondrial derivative (MMD) is attached to the axonemal sheath (AxSh). The minor mitochondrial derivative (mMD) is reduced in size. The enlargement of the head region shows the attachment of the nucleus to the axial complex (AxC). The nucleus, (N) extends downwards to contact the mitochondrial derivative, (MD). The acrosome (Ac) contacts the plasma membrane at the top of the sperm and is also in contact with the nucleus (N). The areas of contact are indicated by the dark shading. Drawn from Kiefer (1970) and Lindsley and Tokuyasu (1980). PB - Paracrystalline Body, BB - Basal Body.

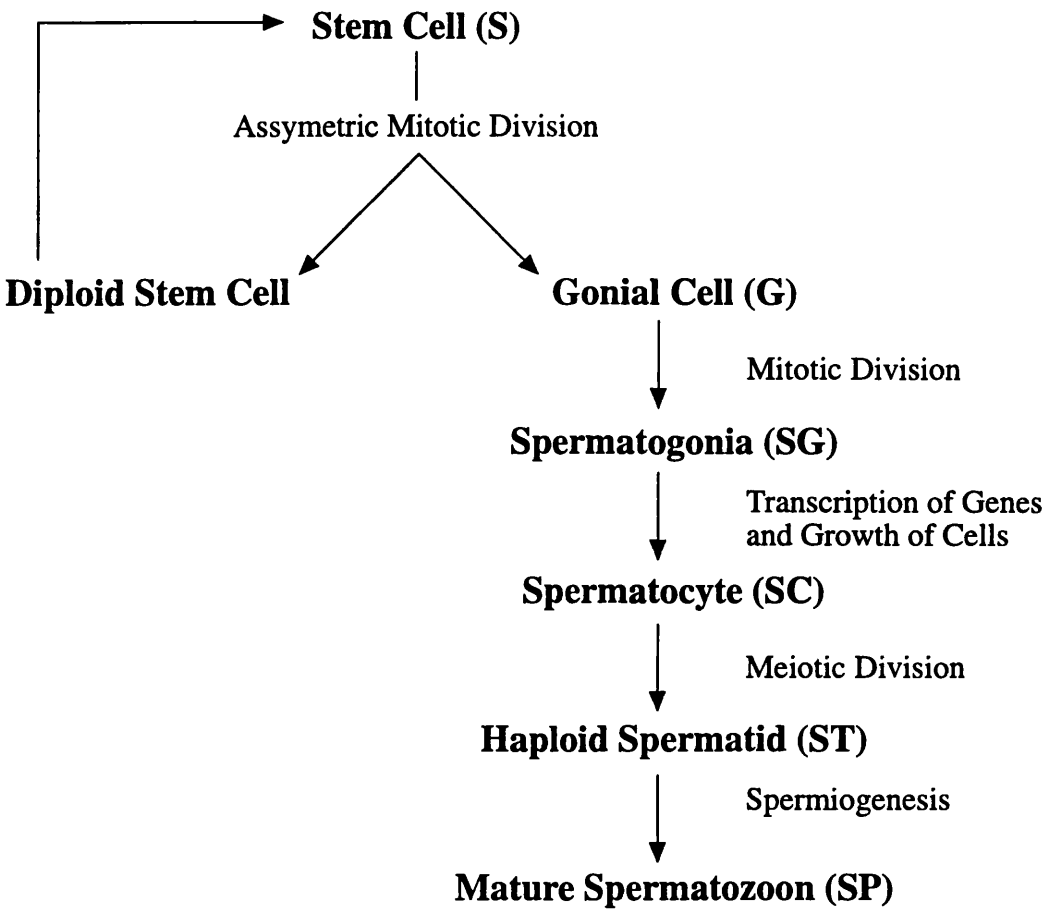


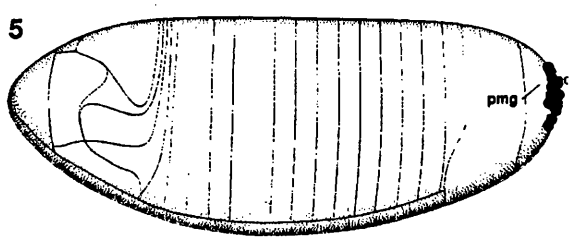
Figure 1.2. A schematic representation of the different processes and stages of *Drosophila* spermatogenesis.

formed the diploid gonial cells undergo several rounds of mitotic divisions resulting in the formation of 16 diploid **spermatogonia** (SG). These mitotic cell divisions are incomplete, leaving the cytoplasms of the spermatogonia continuous with each other. Spermatogonia then enter a period of growth in which transcription of all genes required for the later stages of spermatogenesis occurs. This results in the formation of large diploid **spermatocytes** (SC). Once expression of all the gene products required for maturation (a process known as **spermiogenesis**) has occurred, the 16 primary spermatocytes undergo meiosis resulting in the formation of 64 **haploid spermatids** (ST), which then proceed into **spermiogenesis**. As with the previous mitotic divisions, cytokinesis is incomplete, and consequently the spermatids remain connected throughout development until the final stages of spermiogenesis. Once this process of individualisation has occurred, the mature differentiated **Sperm** (or **Spermatozoa** (SP)) are released into the seminal vesicle where they are stored prior to ejaculation.

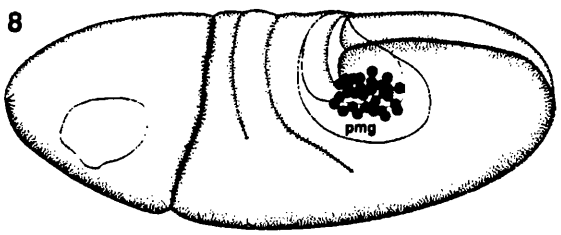
Development Of The Male Reproductive Tract.

The first 13 cleavages during *Drosophila* embryogenesis (stages 1-4) take place without cell division leading to the formation of a syncytium in the embryo. Just before stage 5 (see Figure 1.3a) the sequestration of nuclei in the posterior of the embryo occurs to form the pole cells. At stage 5, the remaining nuclei line up beneath the cell surface leading to the formation of the cellular blastoderm. Between stages 6-8, the process of gastrulation occurs, resulting in the establishment of the three germ layers (the ectoderm, the endoderm and the mesoderm). By stage 8 (Figure 1.3b) the pole cells have migrated into the interior of the embryo forming the germ cell line primordia, and stage 9, the process of germ band elongation has started with several furrows formed along the length of the embryo. The establishment of several organ primordia is also

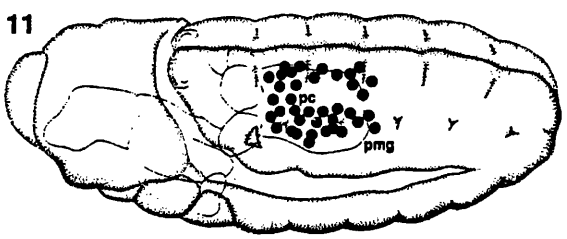
Figure 1.3: Development of the Male Reproductive Tract.



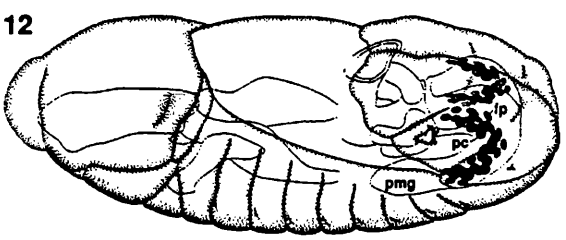
(a) Stage 5 Embryo.
Sequestration of nuclei from the syncytium to form the pole cells (pc). (germ cell primordia, pmg).



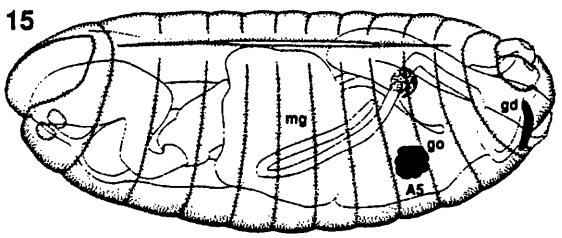
(b) Stage 8 Embryo.
Migration of pole cells to interior of embryo to form germ cell primordia (pmg).



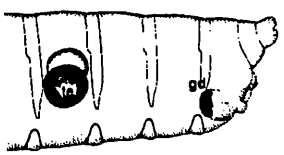
(c) Stage 11 Embryo.
Pole cells start to migrate laterally.



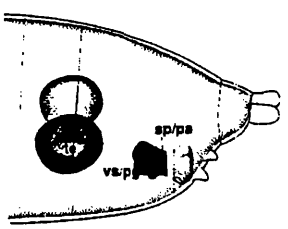
(d) Stage 12 Embryo.
Continued migration of the pole cells.



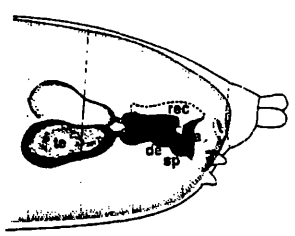
(e) Stage 15 Embryo.
Migration of gonad (go) toward segment A5 complete. Formation of the genital imaginal disc (gd).



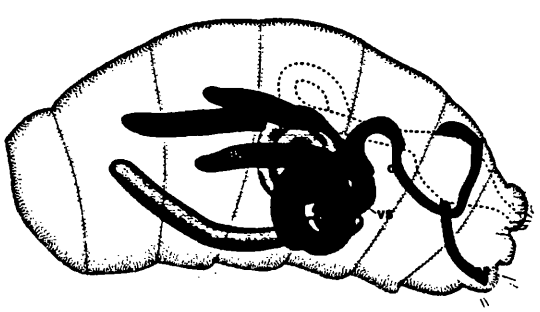
(f) Male 3rd Instar Larvae.
Increase in size of the male gonad. Formation of lumen in the genital disc. (testis, te).



(g) Early Pre-pupal Stage.
Increased growth of testis. Growth of the accessory structures, seminal vesicles (vs) and paragonia (pg) begins from the genital imaginal disc.



(h) Pupal Stage.
Growth and attachment of accessory structures from the genital disc to the testis (ejaculatory duct, de; sperm pump, sp).



(i) Adult Male.
All accessory structures formed. Pigmentation of the outer sheath cells has occurred and the testes have assumed their spiral shape. Adapted from Hartenstein, (1993).

evident. As development proceeds, the furrows laid down earlier disappear so that by stage 11 all but the cephalic furrow have faded. In addition, at this stage, parasegmental border formation occurs. At stage 12, the process of germ band retraction starts and the optic lobe invaginates. The pole cells begin to migrate out of the midgut during the 11th & 12th stages (Figure 1.3c and d). The organ primordia start to differentiate during stage 13, and during stage 14 invagination causes the formation of all of the structures in the head. This process continues during stage 15, with an expansion of the trunk segments. The pole cells at this time are located laterally in segment A5. The genital disc which is the primordia for some of the adult testicular structures, is formed between the anal plate and the denticle belt of segment A8 (Figure 1.3e) (Sonnenblick, 1950; Hartenstein, 1993). This expansion process continues during both stage 16 and 17 until the final larval morphology is assumed: the cuticle thickens and the denticle belts and other specialised structures are formed on the surface of the cuticle.

In larvae which are between 6-10 hours old the male gonad can be seen through the larval integument whilst the much smaller female gonad is difficult to see, even in third instar larvae. This difference in size is a consequence of the different number of cells contained in the male gonad compared with the female gonad: approximately 8-12 germ cells in the female versus 36-38 in the male. During the course of larval development, the male gonad continues to grow in size while the female gonad remains smaller (Figure 1.3f). At metamorphosis, the testis become attached to the vas deferens (Figure 1.3h). This interaction between the gonad and the vas deferens determines the final morphology of the adult testis (Figure 1.3i). Once attachment has occurred, the testis begins to elongate and assume a spiral shape (Bodenstein 1950). In addition, pigment cells from the gonad migrate towards the *vas deferens* producing the species-specific colour.

The genital imaginal disc is responsible for the formation of several structures in the adult testis including the genital ducts, the accessory glands and the external genitalia. The genital disc is known to contain three distinct primordia, one which is specific for the formation of the male structures, one specific for the female structures and the third responsible for formation of the analia (Nothiger *et al.*, 1977; Schupbach *et al.*, 1978). In the late larval period, the genital disc of both sexes forms a sac with an inner lumen. (Figure 1.3f). At the prepupal stage in the male, the posterior part of this structure goes on to form the external genitalia, while the anterior part of the disc forms the *vas deferens* which grows toward the gonad (Figure 1.3g). This growth and differentiation continues in the pupa (Figure 1.3h) giving rise to the genital duct primordia and the accessory structures. Figure 1.3i shows the final organisation of these structures in the adult male with the accessory structures attached to the testis (Bodenstein, 1950; Hartenstein, 1993).

Normal Spermatogenesis In *Drosophila*.

Spermatogenesis has already been initiated in the 1st instar larva (Figure 1.4). After hatching, the third instar larval testis (Figure 1.5) contains 36-38 spermatogonia (SG) with the first cysts of the primary spermatocytes (SC) observed 1 day later. The cysts of growing spermatocytes arise at the apex of the testis and are displaced by subsequent spermatogonia dividing to form new cysts of primary spermatocytes. The terminal cells (TC) seem to be the site at which attachment of the vas deferens occurs. As development proceeds, both spermatogonia and primary spermatocytes are seen to co-exist in the larval testis until just before pupation. Normally, the primary spermatocytes do not enter into meiosis until the white prepupal stage early in metamorphosis (see Figure 1.4). Since development of the soma and the gonad are not tightly linked, it is possible for meiotic and early postmeiotic spermatogenic stages to exist in the third instar larvae (Fuller, 1994).

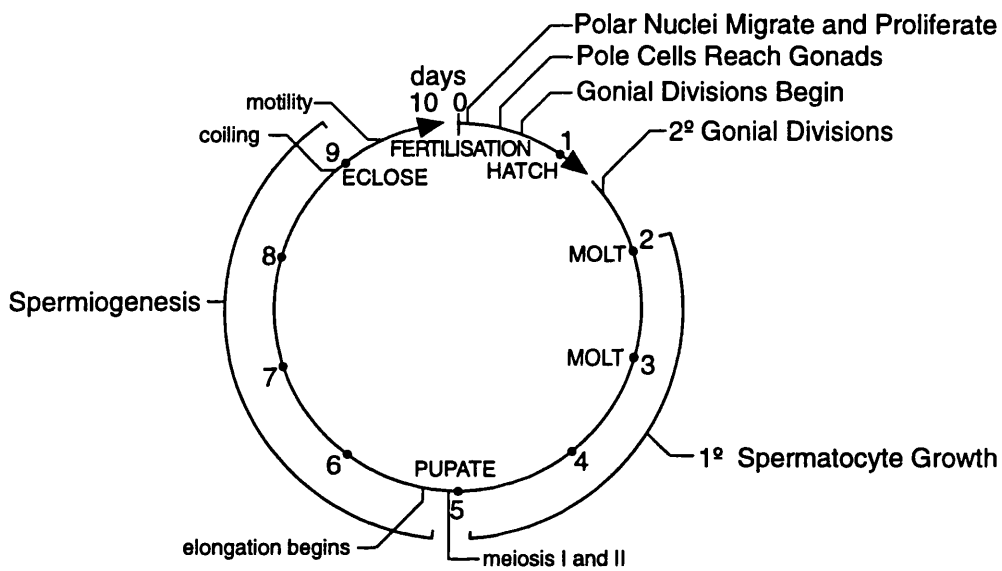


Figure 1.4: The spermatogenic cycle as it occurs in *Drosophila melanogaster* superimposed on the *Drosophila* life cycle (25°C). Adapted from Lindsley and Lifschtyz (1971).

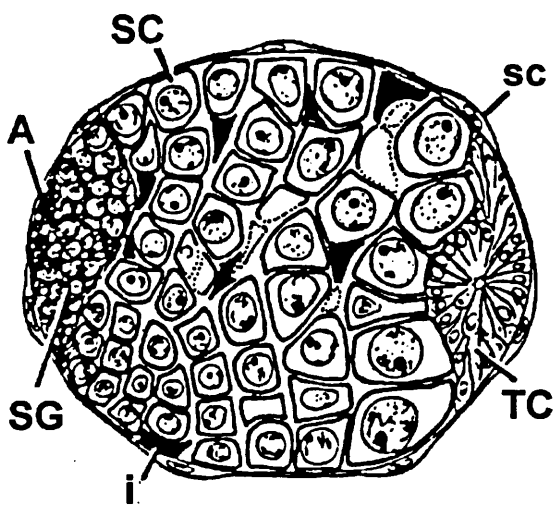


Figure 1.5: Schematic representation of male third instar larval testis. Apical cells (A) at the base, with spermatogonia (SG) forming behind them. The centre of the testis comprises primary spermatocytes (SC) interspersed with interstitial cells (i). The caudal half of the testis consists of terminal cells (TC). Squamous cells (sc) encircle the testis. Taken from Cooper (1950).

The adult testis (Figure 1.6) represents a 'steady state' system in which all stages of spermatogenesis are present. The spermatogenic process can be split into several stages:- spermatogonial growth and mitotic division; primary spermatocyte growth and male meiotic division; and spermiogenesis. Each of these stages will be considered in turn.

Spermatogonial Growth And Mitotic Division.

The tip of the *Drosophila* testis (Figure 1.7) consists of three types of cell: the **apical** cells (A) the **stem** cells (S) and the **cyst progenitor** cells (C). The apical cells form the hub of the testis, and are surrounded by the stem cells (normally 5-9), which in turn are enclosed by the cyst progenitor cells (Lindsley & Tokuyasu, 1980). At the onset of spermatogenesis (Figure 1.2), one of the stem cells (S) undergoes an asymmetric mitotic division to produce a new stem cell, which remains at the hub, and a gonial cell (G) which becomes displaced laterally away from the apex. The gonial cell (enclosed by two of the cyst progenitor cells) will found a cluster of spermatogonia that develop and divide in synchrony. In order to produce primary spermatocytes, the gonial cells undergo several mitotic divisions, the number of which is species-dependent (e.g. , 4 in *Drosophila melanogaster*; 3 in *Drosophila hydei).* This produces a cyst containing 16 primary spermatocytes (*D. melanogaster*) which are linked by cytoplasmic bridges due to incomplete cytokinesis (these bridges persist throughout the subsequent development and differentiation of the cyst).

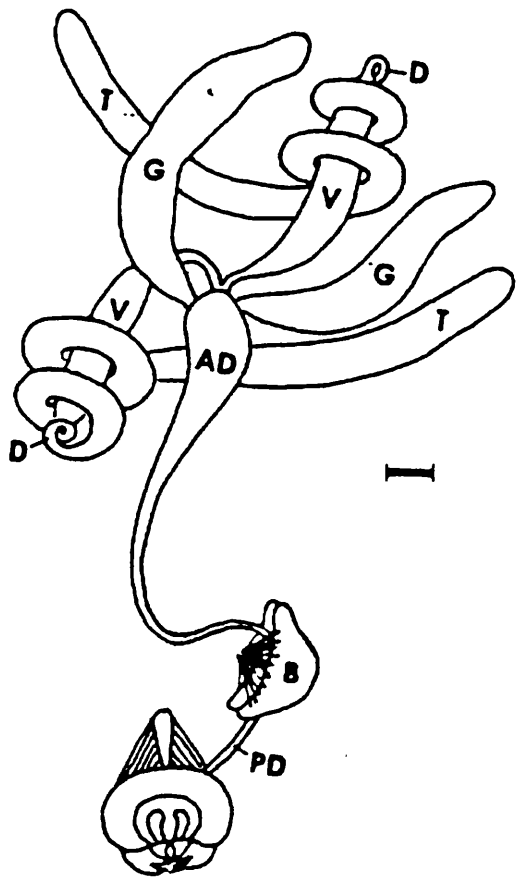


Figure 1.6: Schematic representation of the male adult reproductive tract. testis (T), testicular duct (D), seminal vesicle (V), accessory gland (G), anterior ejaculatory duct (AD), ejaculatory bulb (B), posterior ejaculatory duct (PD). Taken from Lindsley and Tokuyasu (1980).

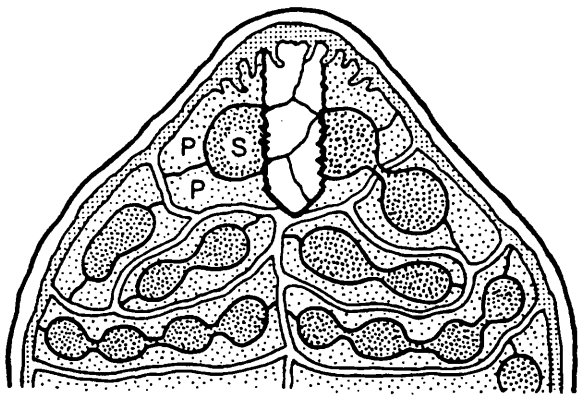


Figure 1.7a: Longitudinal section through the apex of the testis. The suggested process by which primary gonial cells (PC) are generated from stem cells (S) is shown. The hub is attached to the apex *via* extensions of the basal lamina. Six stem cells (S) surround the hub. Each stem cell (S) is surrounded by a pair of cyst progenitor cells (P).

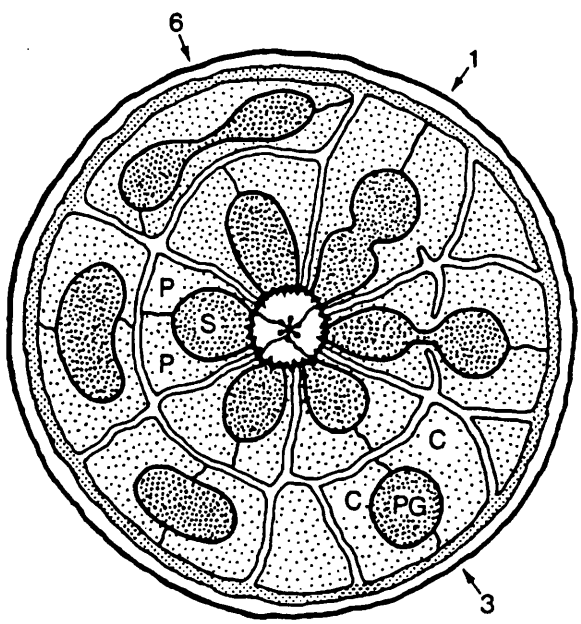


Figure 1.7b: Cross section through the testicular hub. Generation of primary gonial cells occurs in a clockwise direction (1 to 6). Division of the stem cell (S) to produce the primary gonial cell (PG) occurs at the same time as division of cyst progenitor cells (C). The cyst cells and stem cells undergo cytokinesis together ensuring that the primary gonial cell (PG) is enclosed by a pair of cyst progenitor cells (C). As the primary gonial cell undergoes the first gonial division, the stem cell divides again to generate another primary gonial cell. Taken from Lindsley and Tokuyasu (1980).

Primary Spermatocyte Growth And The Male Meiotic Division.

The primary spermatocytes now begin a new programme of development. This stage of spermatogenesis, which lasts for ~90 hr, is characterised by a large increase in cell volume (25-30%) (Lindsley & Lifschytz, 1971). It is during this stage that transcription of all gene products required for subsequent morphogenesis is thought to occur. The studies of Olivieri & Olivieri (1965) and Gould-Somero & Holland (1974) suggest that most, if not all, transcription ceases after this stage. Another important feature of primary spermatocytes is the appearance of the 'Y loops' on the Y chromosome, suggesting these chromosomes have become transcriptionally active. The implications of this will be discussed in the section dealing with the Y- chromosome.

During this time the primary spermatocyte itself undergoes morphological changes. At the onset, the primary spermatocyte is the same size as its gonial cell precursor, however during this time it develops through a series of stages before the formation of the mature primary spermatocyte, the largest spermatocyte in the testis.

In primary spermatocytes, the events described above take place during an elongated G_2 stage. The transition from this G_2 phase to the metaphase stage of the first meiotic division is characterised by various changes in the nuclear and chromosomal morphology. In *Drosophila* males, the stage at which the X and Y chromosomes separate (the reduction division) is in meiosis I. During both meiosis I and meiosis II, the mitochondria line up parallel to the nucleus ensuring their equal division during cytokinesis. Together meiosis I and II produce a cyst in which the 16 primary spermatocytes have divided to produce 64 haploid spermatids (*D. melanogaster*). As with previous divisions, cytokinesis is incomplete and the spermatids remain connected *via* cytoplasmic bridges.

completing meiosis the spermatids undergo extensive differentiation during the maturation process known as spermiogenesis.

Spermiogenesis.

The haploid spermatids now enter into the process of spermiogenesis. This maturation process may be divided into several intervals: pre-elongation, elongation, transition, post-elongation, individualisation and coiling. Together, they ensure the transformation of the haploid spermatid into a mature sperm. The process has been extensively discussed by Lindsley & Tokuyasu (1980) and Fuller (1994) and is summarised below.

1. Pre-Elongation.

During this stage, the structures of the spermatid components are altered in order to facilitate the morphological changes which occur later in spermiogenesis. It is characterised by movement of the mitochondria, which gather and fuse to form a spherical mass located to the side of the nucleus. This developmental period, often termed the 'onion' stage, results in the formation of a giant bipartite mitochondrion, known as the **nebenkern**. The nebenkern consists of layer upon layer of mitochondrial membranes in a tightly-wrapped structure which has the appearance of an onion in cross-section. Consequently the mitochondrion changes from a sphere to a cylinder and the axoneme grows from the basal body during elongation of the axial complex (Figure 1.8). In addition to nebenkern formation, two other morphological changes occur at this stage: the **axoneme** grows out from the basal body (located at the base of the nucleus) starting the development of the sperm tail, and fusion of the Golgi bodies occurs to form the **acroblast** (located opposite the basal body) at the apex of the nucleus (Figure 1.8).

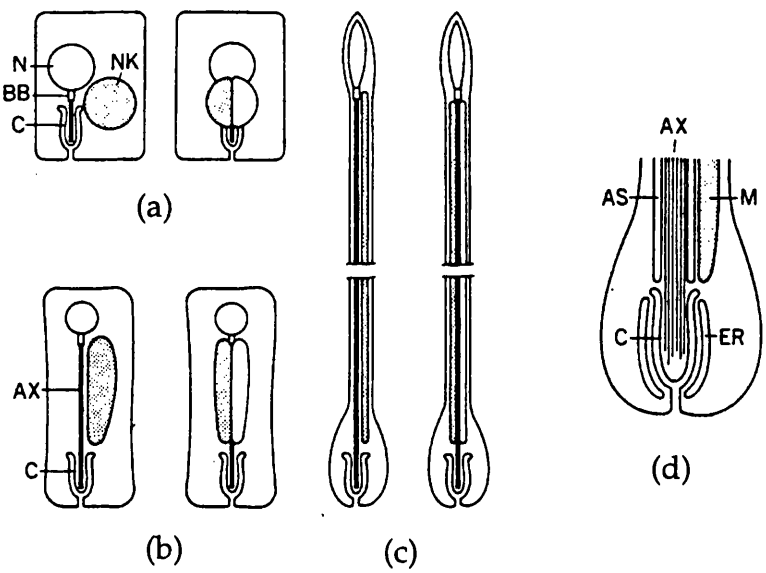


Figure 1.8: Changes which occur in the nucleus and cytoplasm during elongation. **(a)** shows a pre-elongation spermatid. Nucleus (N), basal body (BB), Cilium (C), nebenkern (NK) axoneme (AX), Axonemal Sheath (AS) and Mitochondrion (M). The bipartite structure of the nebenkern can be seen. **(b)** shows the elongation of the axoneme and the nebenkern, whilst in **(c)** the continued elongation of the axoneme and the nebenkern is shown and the nucleus has started to elongate, changing from a spherical shape **(a)** into a needle **(c)**. **(d)** shows in detail the extremity of the axoneme. This has not fully differentiated into the 9+2 structure and remains enclosed within the cilium. The cilium, in turn is enclosed within a sheath of endoplasmic reticulum (ER). Taken from Lindsley and Tokuyasu (1980).

2. Elongation.

At this point the spermatids are still disc shaped. During this phase their components elongate to the dimensions of a mature sperm and involves elongation of the axoneme, nebenkern and nucleus (Figure 1.8).

The axoneme.

The axoneme (the propelling force of the sperm) is derived from the basal body of the nucleus. It has the 9+2 structure characteristic of other flagellar structures derived from microtubules (Gibbons 1981). During elongation, its structure becomes more complex (Figure 1.9) with the addition of accessory structures. A projection of tubulin originates from the microtubule doublet, forming the accessory microtubule. This structure itself becomes more complex, with spokes forming between the outer doublets and the central region of the axoneme. The additional microtubules and other accessory structures are first formed on the side of the axoneme that is opposite the nebenkern (Kiefer, 1970; Williamson, 1976).

The Nebenkern.

Towards the end of the onion stage, the nebenkern layers begin to unfold. As the cell elongates, the mitochondria become more spear shaped and filled with cytoplasm containing ribosomes and other cytoplasmic constituents (these disappear as elongation proceeds). The elongation of the mitochondrial derivatives is thought to be due to the action of cytoplasmic microtubules. In addition to elongating, the mitochondrion also splits into two parts (Figure 1.8), the major and minor mitochondrial derivatives (Figure 1.14). During elongation, the major derivative remains larger than the minor derivative, which

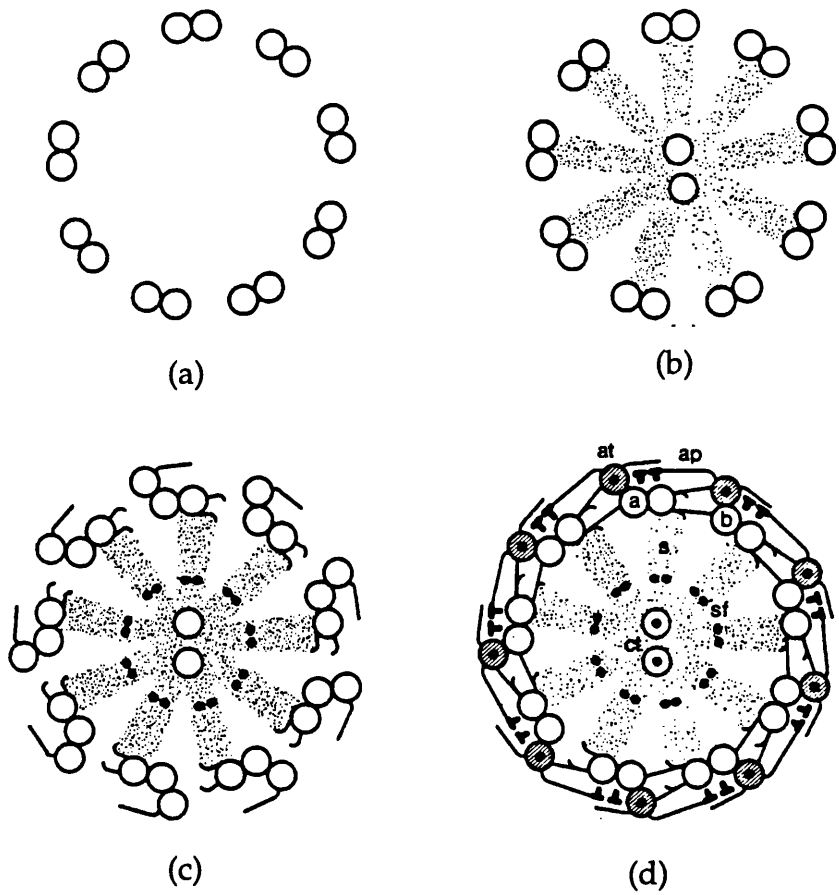


Figure 1.9: Development of the axoneme.
(a) shows the axoneme early in development where only the nine peripheral fibres can be seen. In later stages, shown in **(b)**, the central microtubule pair has developed and the spokes between the central microtubule pair and the peripheral microtubules become apparent. **(c)** shows the growth of the accessory microtubules from the B fibre and the development of the connecting fibres from the A fibre. **(d)** shows the structure of the complete axoneme. The accessory microtubules (at), the spokes (s) and secondary fibres (sf) can clearly be seen (Taken from Williamson, 1976).

gradually reduces in size. At maturation, the minor derivative has almost disappeared whilst the major derivative consists mostly of paracrystalline material (Figure 1.10). This process of degeneration of one of the mitochondrial derivatives in *D. melanogaster* differs from that in other dipteran species as *D. hydei,* where both mitochondrial derivatives are maintained. At the latest stages of elongation, deposition of densely staining granules results in the formation of the paracrystalline array in the major mitochondrial derivative. This deposition occurs at the boundary between the axonemal sheath and the mitochondrial derivative. The paracrystalline body has a very regular structure consisting of round crystals (with a diameter of 80Å) arranged in a hexagonal array. It has been postulated (Bairati & Perotti, 1970) that this arrangement is related to complex biochemical function. The crystalline protein structure is thought to arise as a consequence of reorganisation of the mitochondrial cristae proteins and respiratory complexes during spermatogenesis.

Nuclear Shaping And Elongation.

As the axial complex grows during elongation, the morphology of the nucleus alters from a sphere to a hook shaped structure (Figure 1.11). Concurrent with deposition of the paracrystalline array in the mitochondria, the nucleus becomes flattened and fenestrated (perforated). A concave depression forms at the flattened side, which is filled with perinuclear microtubules (Figure 1.11d). Microtubules are also laid down at the convex edge of the nucleus facilitating elongation. On the convex side of the nucleus, chromatin condensation occurs (Figure 1.11f). Budding off of nucleoplasm and nuclear membrane from the concave side results in a 200-fold reduction in nuclear volume yielding a nucleus which is 9μm in length and 0.3μm wide (Figure 1.11h - j).

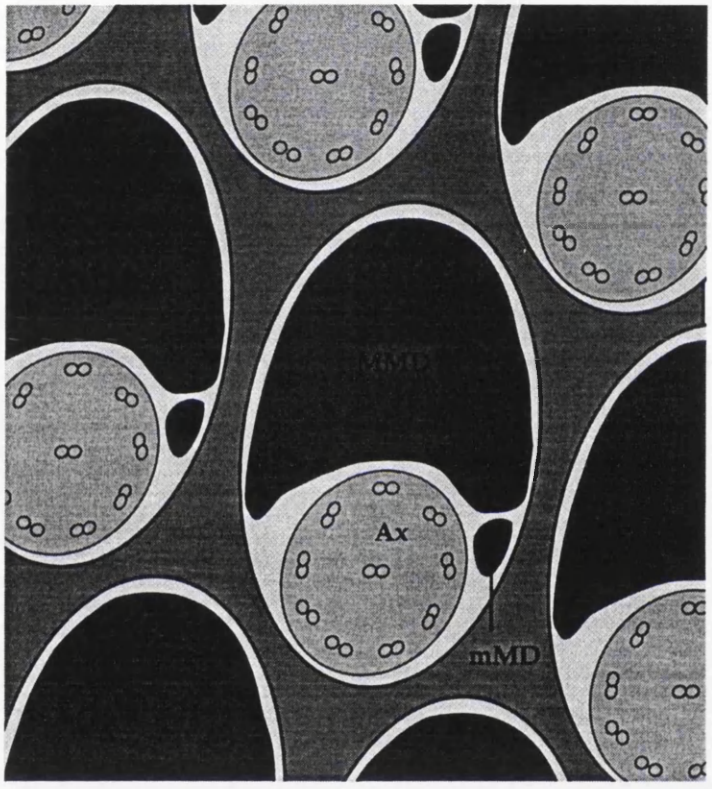


Figure 1.10: Schematic representation of a cross-section through a mature sperm bundle showing details of the axial complex. Major Mitochondrial Derivative (MMD) , minor Mitochondrial Derivative (mMD), Axoneme (Ax). Drawn from Kiefer (1970).

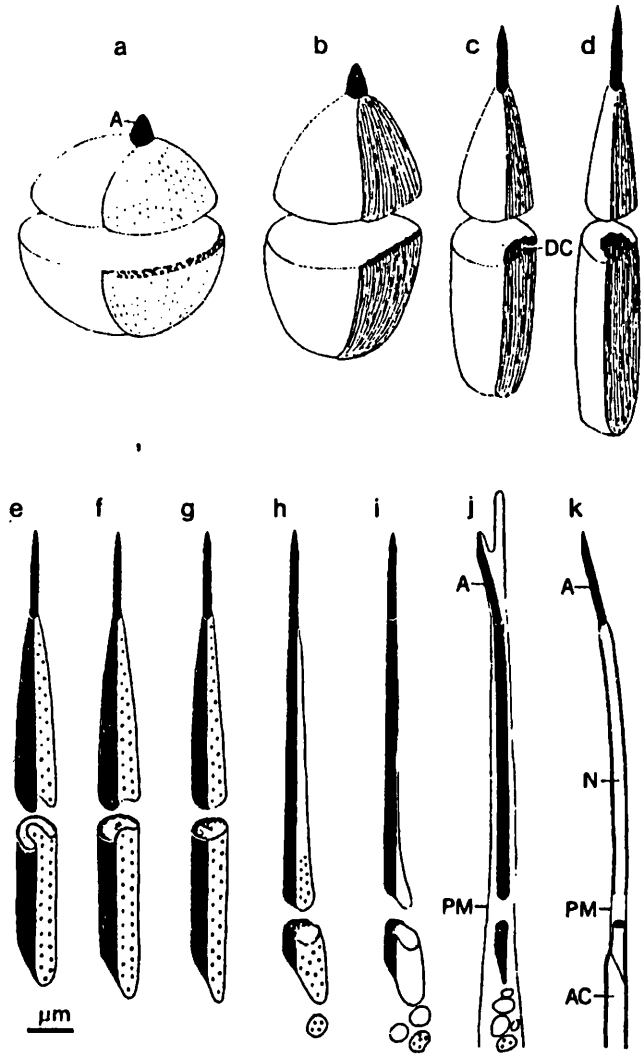


Figure 1.11 shows the developmental changes occurring in the nucleus during the processes of elongation, transition, post-elongation and individualisation.
Figure 1.11a shows a nucleus at the start of elongation. In **Figure 1.11b** deposition of microtubules occurs resulting in a depression. In **Figure 1.11c & d** the nucleus begins to be transformed, adopting a more concave-shape and growing in length. **Figure 1.11e** is a transition nucleus. The nucleus has reached almost full length and the microtubules begin to migrate out of the concave depression over to the convex side. (shaded part of nucleus). The alignment of the nucleus with respect to the axial complex also occurs at this stage. The basal end of the nucleus becomes more pointed towards the axoneme (compare **Figure 1.11d** to **Figure 1.11e**). **Figures 1.11f - j** show the nucleus in the post-elongation period. This is characterised by a reduction in nuclear volume. As the microtubules migrate out of the nucleus, the depth of the concave side is reduced. Chromatin condensation occurs on the non-fenestrated part of the nucleus. In addition, the nucleus reduces its volume by the formation of vesicles causing the loss of nuclear membrane. Initially, this process occurs at the fenestrated part (**Figure 1.11h**) followed by loss from the non-fenestrated part. The acrosome (A) becomes tilted with respect to the nuclear axis (**Figure 1.11j**). **Figure 1.11k** represents the nucleus after the process of individualisation. The microtubules which had surrounded the nucleus become lost (indicated by a loss of shading) leaving the acrosome (A), nucleus (N) and Axial complex (Ac) surrounded by a thin layer of plasma membrane (PM). Taken from Lindsley and Tokuyasu (1980).

3. Transition Period.

During this stage two processes occur, redistribution of the microtubules surrounding the nucleus and alignment of the nucleus to the axial complex.

Microtubule Redistribution.

As described, during nuclear elongation the nucleus become fenestrated and concave. This altered morphology is coincident with an approximate 2-fold increase in the number of microtubules on the concave side (Figure 1.12b). Additionally, the microtubules move out from this area in both directions resulting in the fragmentation of the endoplasmic reticulum. The endoplasmic reticulum which previously enclosed the nucleus, is now restricted to the paranuclear strip. This movement results in a single layer of microtubules covering the convex nuclear surface (Figure 1.12c).

Nuclear- Axial Complex Realignment.

The nucleus and the axial body are in contact at the base of the nucleus (where the basal body is located), although the orientation of the nucleus relative to the axial body is not fixed. The microtubules in the axial body and those of the nucleus are separate and discontinuous. However, once the nucleus changes its microtubule distribution, two events take place. First, there is movement of nuclear microtubules down to cover the junction between these two structures, resulting in the formation of a sheath of microtubules. Secondly, there is a 2-3 fold increase in the diameter of the anterior tips of the mitochondrial derivatives. This enlarged axial complex fits into the microtubular sheath, resulting in the alteration and compaction of both these structures. A posterior extension forms at the fenestrated side of the nucleus. The nuclear socket (containing the basal

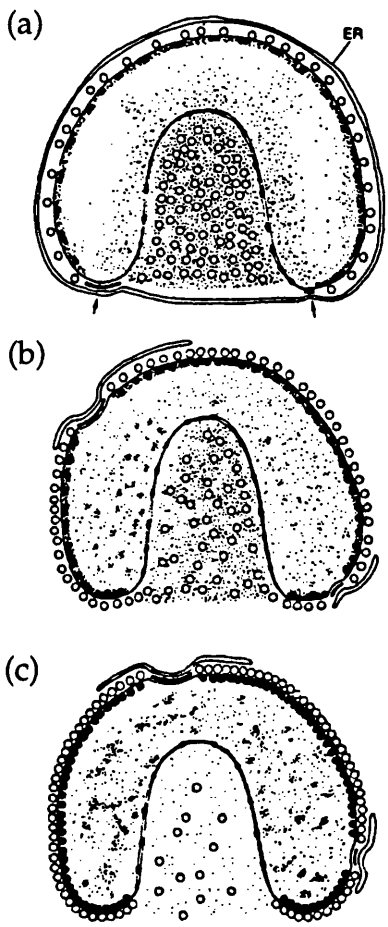


Figure 1.12 shows the progress of microtubule redistribution in a transition and post-elongation period nucleus.
Figure 1.12a shows a transition stage nucleus surrounded by the endoplasmic reticulum (ER). The paranuclear strips, located at the boundary between the concave and convex side of the nucleus, are indicated by arrows.
In Figure 1.12b, the movement of the paranuclear strips towards the convex side results in the movement of the microtubules (white circles) out of the concave depression. This movement results in disruption of the ER except for that which covers the paranuclear strips. Chromatin compaction (darker circles) has started to occur. In the post-elongation nucleus, the larger of the paranuclear strips stops at the top of the nucleus (c) whilst the smaller strip is located in a more lateral position. Chromatin compaction has increased, with its pattern of condensation following that of the microtubules. Taken from Lindsley and Tokuyasu (1980).

body) forms a groove. The anterior tip of the mitochondrial derivative becomes flattened and aligns with the basal body, restricting the rotational freedom of the axial complex and the nucleus. The nucleus, therefore, becomes constrained at two points: by the basal body and by the projections of the major mitochondrial derivative. In mature sperm, the nucleus and axial complex are continuous (Figure 1.1).

4. Post-Elongation.

Events that are required for tail formation are often initiated prior to post-elongation. During this time, events such as the development of the accessory structures continue until the highly ordered structures such as the spokes connecting the peripheral element can easily be distinguished.

During the preceding stages, the nuclear volume has been kept constant, but now the nuclear surface area is gradually reduced. This is achieved by reducing the depth of the concave side of the nucleus so that in cross-section the nucleus appears almost round (Figure 1.11g). As discussed in the section on nuclear shaping, chromatin condensation and budding of the nuclear membrane occurs (Figure 1.11f-j). Membrane is first removed from the fenestrated edge of the nucleus moving into the non-fenestrated portion. Its final shape depends upon the perinuclear microtubules. The **acrosome**, situated at the apex of the nucleus, initially becomes tilted with respect to the nuclear axis, forming a forked configuration at the head of the spermatid (Figure 1.11j). This organelle, derived from the **acroblast**, may be considered a modified lysosome. It contains enzymes which will facilitate the penetration of the egg by the sperm. When the sperm is in the vicinity of the egg, it undergoes a complex reaction resulting in the shedding of both the outer plasma membrane and the acrosome membrane.

This facilitates the release of the acrosome enzymes which assist the sperm penetrating the egg. The *Drosophila* acrosome is typical of that observed in other species. It has a dagger-like shape and seems to consist of a single compact substance containing no accessory structures. Different regions appear to exist within the acrosome; this zonation might reflect a different functional or biochemical requirement for each.

5. Individualisation And Coiling.

At the start of these stages, the spermatids are still connected *via* the cytoplasmic bridges. However, during the course of this process, the spermatids are stripped of the cytoplasmic bridges, excess cytoplasm and unnecessary organelles. The process begins at the head region of the spermatid bundle, with the expansion of the spermatid bundle to form the **cystic bulge** (CB) (Figure 1.13b). This structure moves caudally along the spermatid bundle, removing the excess structures and cytoplasm with a resulting complementary increase in its own volume. The leading edge of this structure, known as the **investment cone** (I), functions by surrounding each sperm tail and using its numerous microfilaments assists the cystic bulge in its progression along the spermatid bundle. The investment cones, starting in the head (Figure 1.13a), strip the spermatids of excess nuclear membrane, leaving the nucleus surrounded by a single layer of microtubules. Migration down the spermatids removes the perinuclear microfilaments, the paranuclear strips and dramatically reduces in size the minor mitochondrial derivative. The axonemal sheath furthest from the mitochondrial derivative is also removed, leaving only that part which separates the axoneme from the mitochondrial derivative. At the end of this process, the acrosome attaches to the plasma membrane and the nucleus attaches to both the acrosome and the axial complex. The sperm tail is divided by the remnants of the axonemal

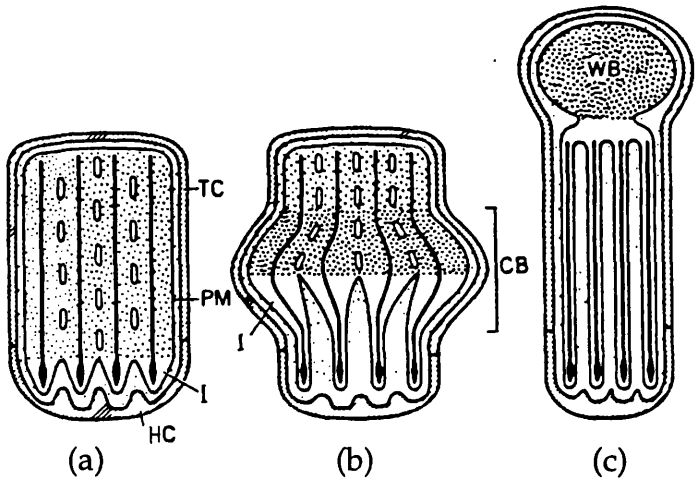


Figure 1.13 represents the process of individualisation. In **(a)**, the spermatids are still in the syncytium, they are enclosed with in a plasma membrane (PM) which in turn is enclosed in the head cyst (HC) and tail cyst (TC) cells. As the investment cones (I) progress from the head to the tail, there is a loss of the cytoplasmic organelles (represented by a loss of shading in **(a)**). In **(b)** the investment cone has moved down the spermatid bundle clearing the cytoplasm and leaving the spermatids in their own plasma membrane. The cystic bulge (CB) is formed and the organelles are collected in this region. Once individualisation is complete **(c)**, the spermatids are fully enclosed within their own cytoplasm, the cystic bulge has progressed along the length of the spermatid, removing all the organelles and forming the waste bag (WB) at the end of the axial complex. Taken from Lindsley and Tokuyasu (1980).

remnants of the axonemal sheath into two compartments, one containing the axoneme, the other the mitochondrial derivatives (Figures 1.1 & 1.10). The axoneme attaches to the nucleus at its anterior and to the plasma membrane at its posterior, whilst the mitochondrial derivatives attach to the remnants of the axonemal sheath. This arrangement of tail structures may facilitate sperm motility. After progressing along the spermatid bundle, the cystic bulge forms the **waste bag** (WB) at the bottom of the tail cyst (Figure 1.13c).

Coiling represents the final stage in the processes of spermiogenesis. The sperm heads within the head cyst cell (Figure 1.13a) become enclosed by terminal epithelial cells in the basal region of the testis. The coiling commences at the head of the sperm, continuing along the sperm tails resulting in their removal from the apical end of the testis. In the coiled area of the bundle, microtubules are formed. These move along the tails packaging the sperm into a highly ordered hexagonal array. Abnormal sperm whose structure does not favour the hexagonal packaging are removed by these microtubules into the waste bag. The microtubules depolymerise at the completion of this process. Normal sperm are released into the testicular lumen, leaving abnormal sperm in the waste bag. The terminal cells phagocytose the waste bags which are degraded by the action of lysosomes. The mature sperm progress along the testicular duct until they reach the seminal vesicle. Here the sperm are aligned, in both a parallel and anti-parallel manner forming huge bundles consisting of a few thousand mature sperm.

As can be seen from the above description, spermatogenesis represents an extremely complex morphological process. These events will obviously require the participation and interaction of a large number of genes. The following sections describe those genes that are known to have a role in spermatogenesis. Some of these genes are known to be Y-linked while the majority are located on the autosomes.

The Y Chromosome.

In *Drosophila*, the process of sex-determination is achieved by determining the ratio of X-chromosomes to Autosomes (X:A ratio) (Baker, 1989; Cline, 1993). The Y-chromosome in this case contains no genes that determine sex, contrasting with mammals where the Y chromosome carries the gene SRY (sex-determining region of Y). In mammals, the presence of the Y-chromosome dictates that the male pattern of development is followed, irrespective of the rest of the chromosomal composition (Koopman *et al.*, 1991). Although it has no role in sex-determination, there is an absolute requirement for the presence of a Y-chromosome during *Drosophila* **spermiogenesis**. Males which do not have a Y-chromosome (X:O) though indistinguishable externally from normal males, are sterile. In X:O males, normal spermatogenesis seems to occur until spermiogenesis. After this time, the sperm degenerate in the posterior region of the testis so that no motile sperm are found in the *vas deferens*. The defects in spermatogenesis for Y-deficient males appear to fall into two classes :- those affecting the mitochondrial derivatives and those which affect the axoneme (Kiefer 1966, Williamson 1976). In both cases, the relative orientation of these two structures becomes disturbed (Figure 1.14a), resulting in structures having different ratios of axonemal complexes to mitochondrial derivatives (Figure 1.14b). Formation of the paracrystalline array is often affected (Figure 1.14b). Normally, this structure occurs at the junction between the axoneme and the mitochondrial derivative. However, in X:O males, paracrystalline formation may initiate at several times resulting in more than one paracrystalline body (Figure 1.14b). Formation of the paracrystalline array may also occur in the minor mitochondrial derivative whereas in normal X:Y males it is confined to the major derivative. The axoneme, in contrast does not appear to be as severely affected by Y-deficiency as is the mitochondrial derivative. Axonemes formed in the absence of a Y-chromosome appear to

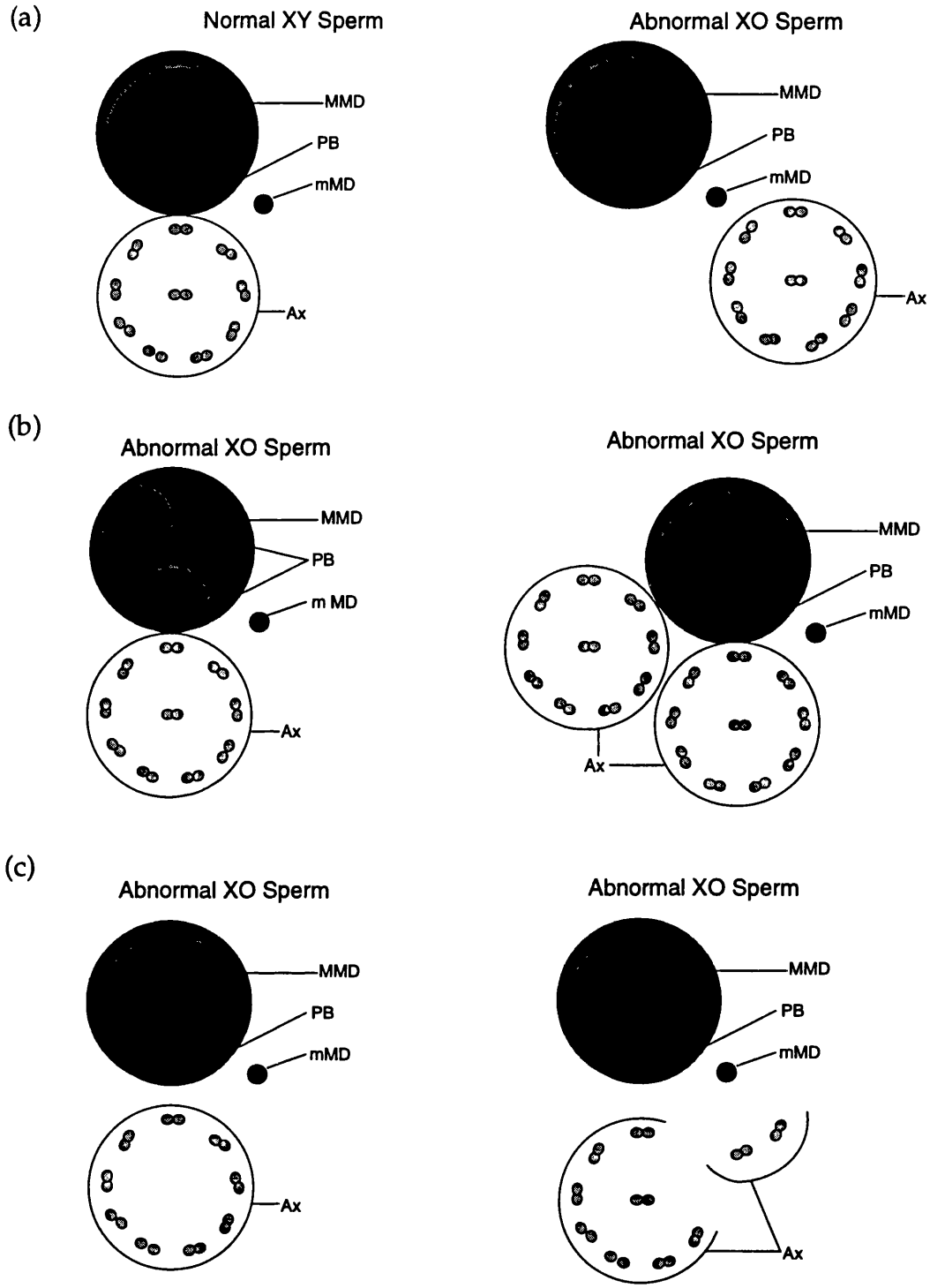


Figure 1.14: spermatogenic defects observed in XO flies. **(a)** represents a comparison of wild type sperm with that of X:O sperm. The orientation of the mitochondria and axoneme is disrupted. There is no attachment between the mitochondria and the axoneme in the mutant sperm. **(b)** shows an example of mitochondria containing two paracrystalline bodies and two axial fibre complexes. In **(c)** some of the defects observed in axonemes; those lacking the central pair, and those lacking the peripheral fibres. Major Mitochondrial Derivative (MMD), minor Mitochondrial Derivative (mMD), Axoneme (Ax), Paracrystalline Body (PB) . Drawn from Kiefer, (1970).

have all the correct components laid down in the right configuration. Abnormal axonemes which do occur appear to miss some structures such as the central tubule pair or peripheral tubules. (Figure 1.14c). These defective axonemes degenerate during the process of spermatid maturation (Kiefer, 1970). The axial fibre complex becomes disorganised, resulting in the loss of spokes and secondary fibres until only the central and accessory fibres remain. These are then all lost concurrently.

The Y-chromosome of *Drosophila* is a large heterochromatic chromosome and although it comprises 12% of the *Drosophila* male genome, relatively few Y-linked loci have been identified and localised (Figure 1.15a). Some of the genes that have been localised to the Y-chromosome are *suppresser of stellate Su(ste)* (Livak 1990), the *bobbed* (*bb*) gene family (Ritossa, 1976; Lindsley & Zimm, 1992) and various fertility factors (six in *D. melanogaster*). The *Su(ste)* locus is located between two fertility factors, kl-1 and kl-2, on the long arm of the Y-chromosome. It affects the abundance and splicing of the X-linked gene *stellate* (*ste*). Males that are defective for both *ste* and $Su(ste)^-$ are sterile and can easily be identified by the presence of star shaped crystals in the nuclei and cytoplasm of their primary spermatocytes. Their spermatids are also affected, resulting in formation of micronuclei and irregular nebenkern. In contrast, $ste^+/Su(ste)^-$ males have needle-like structures in their primary spermatocytes. These males are fertile, although non-disjunction of all chromosomes occurs. Meiotic drive is also seen to occur, with those sperm that do not contain the full chromosome complement recovered more often than those that do (Meiotic drive is a term used to describe the situation whereby there is an unequal distribution of alleles in the gametes of a heterozygote). An example of meiotic drive in *D. melanogaster* occurs at the *Segregation Distorter* (*SD*) locus. In males heterozygous for the mutant allele of the *SD* locus, very few sperm are recovered which carry SD^+.

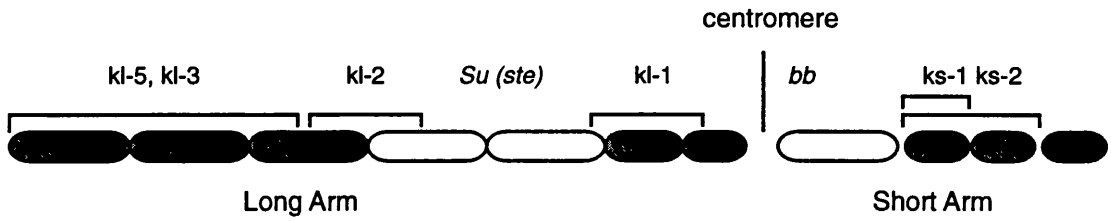


Figure 1.15a is a schematic representation of the Y chromosome of *Drosophila melanogaster*. the positions of the fertility factors are shown. Drawn from Hackstein, 1987.

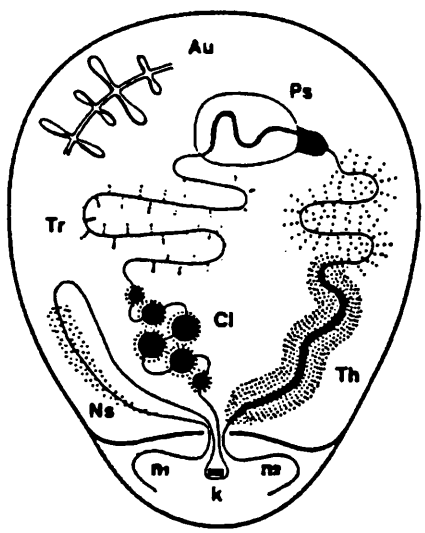


Figure 1.15b is a schematic representation of the structures formed in the primary spermatocyte nucleus of *D. hydeii*. nooses loop (Ns), clubs (Cl), threads (Th), tubular ribbons (Tr) and pseudonucleolus (Ps). Taken from Glatzer et al., 1984.

The *Su(ste)* locus has been analysed at the molecular level. It shares sequence homology with the *ste* locus on the X-chromosome. The *Su(ste)* locus exists as repeats of approximately 80 copies; although this figure is species dependent. The length of the repeat is between 2.6-2.8 kb. Deletion of the cluster results in the over-production of *ste* RNA.

One locus that is essential for viability is *bobbed* (*bb*), an array of ribosomal genes (situated proximal to ks-1 and ks-2). A second array of ribosomal genes is also present on the X-chromosome. Both chromosome clusters are thought to contain approximately 225 rRNA genes organised as a series of tandem arrays consisting of transcribed units separated by non-transcribed spacers. Three types of array are known to exist:- type I appears to be the functional copy of the gene, consisting of 11 kb of DNA, consisting of an external transcribed spacer followed by the structural gene for 18s rRNA, an internal transcribed spacer containing the sequences for 5s rRNA, followed by the 28s rRNA sequences interspersed with short transcribed spacers. The type II and III forms of the array have an additional insert in the structural gene for 28s rRNA; these appear to be transcriptionally inactive. All three types are randomly distributed throughout the gene cluster.

The *bb* loci on the X and Y chromosome differ in three respects:- 65% of the rDNA on the X is type I, these type I sequences are reduced in the Y-chromosome. Additionally, the 18s rRNA transcribed from the X-chromosome differs from that from the Y by at least 1 base substitution. The sequence of the 5' non-transcribed spacers also differ between the X and Y chromosome. There is no proven exchange of sequence between these loci. However, it has been postulated that one of the roles of the non-transcribed spacer is to allow chromosome pairing (McKee *et al.*, 1992).

In addition to these few genes and the fertility factors described below, the Y-chromosome contains retrotransposon like sequences (Huijser, 1988). These seem to consist of two classes:- short, tandem repeats restricted to the Y-chromosome (Y-specific) and repeated sequences several kb long which occur in other regions of the genome as well as on the Y (Y-associated) (Hennig *et al.*, 1988).

The Fertility Factors

The fertility factors are large complex genes that function exclusively in the process of spermatogenesis. They are responsible for the large lampbrush or 'Y loop' structures observed during the primary spermatocyte stage (Glatzer *et al.;* 1984). Loss of one of these fertility factors (*D. melanogaster* is thought to have 6, *D. hydei* between 7 and 16) appears to be just as detrimental as is loss of the whole Y-chromosome (Hess 1973). The fertility factors in *D. melanogaster* are organised into two complexes, KS (on the short arm of the Y) and KL (on the long arm). Saturation mutagenesis was used to define these factors using both X-rays (Brosseau, 1960) and Ethyl Methyl Sulphate (EMS) (Kennison, 1983).

The KS complex consists of two complementing factors, ks-1 and ks-2. Mutations affecting either ks-1 or ks-2 result in sterility and in both cases the structure of the axoneme is affected. In the case of ks-1 mutants, tubules are missing from the axoneme while in ks-2 mutants the alignment of the axoneme and the mitochondrial derivative is defective. The development of the major and minor mitochondrial derivatives with respect to the endoplasmic reticulum is also disrupted in ks-2 mutants.

The KL complex consists of four complementation groups:- kl-1, kl-2, kl-3 and kl-5. Again, defects in these factors results in sterility and disruption of the

axoneme. In addition to the above four, some investigators have postulated the existence of a fifth factor in KL, kl-4 (Brosseau, 1960; Williamson, 1972). However, kl-4 has not been localised to a single site or break-point, leading others (Kennison, 1983) to dispute its existence.

There has been considerable speculation as to the possible role for these fertility factors. It had been suggested (Williamson, 1976; Lindsley & Tokuyasu, 1980) that these genes do not function by coding for proteins specific to spermatogenesis; rather their function was thought to be in directing the correct morphogenesis of the sperm components, facilitating their assembly into mature sperm. However recent evidence (Goldstein *et al.*, 1982, Gepner & Hays, 1993; Rasmusson *et al.*, 1994) has suggested a more conventional protein coding role for these loci. Deletions of regions of the Y-chromosome corresponding to the fertility factors kl-5 or kl-2 leads to the lack of the outer dynein arms in the flagella of the axial complex. In addition, high molecular weight proteins (300-325 kDa) are absent from the testes of these mutant males. These polypeptides have the same molecular weights as the outer arm dyneins found in the unicellular alga, *Chlamydomonas reinhardtii* (Goldstein *et al.*, 1982). Recently, in *D. melanogaster* a family of seven dynein molecules (including one that is Y-specific) has been cloned using the polymerase chain reaction (PCR) (Rasmusson *et al.*, 1994). The primers were derived from sea urchin dynein sequences. The Y-specific dynein maps to the same cytogenetic region as kl-5. The predicted protein sequence of the Y-specific dynein shares a high degree of homology with other known dynein proteins (Gepner & Hays, 1993).

Nevertheless, the fertility factors do not have a conventional structure. Using EMS as a mutagen, it appears that Y-linked genes present a target that is approximately ten times the size of 'normal' loci (Kennison, 1983). In addition, complementation analysis suggests that these loci are physically large genes.

The frequency of isolation of male-sterile translocations that affect the same complementation group, and also the cytological distribution of non-complementing male-sterile breakpoints reinforces this suggestion (Kennison, 1981; Gatti & Pimpinelli, 1983; Bonaccorsi, 1988). Again, the presence of so much repeated DNA in this area (Y-specific and Y-associated) implies that the structure of these genes may be similar to those in other heterochromic regions such as *light* (*lt*) (Devlin *et al.*, 1983). This gene has regions of coding exons interspersed by introns containing repetitive DNA sequences.

As described earlier, the transcription of most 'spermatogenic' genes takes place during the primary spermatocyte stage. The fertility factors form large structures, 'Y-loops', in the nucleus of the primary spermatocyte. Most of the structural studies of the Y-loop have focused on *D. hydei* where five to six pairs of loops can be easily identified (Hennig, 1985, 1987; Hackstein, 1987). In *D. hydei* these loops have been further classified on the basis of their structures into 'nooses', 'tubular ribbons', 'clubs', 'pseudonucleoli' and 'threads' (Figure 1.15b). The 'nooses' loop of *D. hydei* is a single 260 kb transcription unit, consisting of middle and low repetitive Y-specific DNA sequences interrupted with Y-associated retroposon-like elements. The 'threads' and 'pseudonucleolus' structures are largely composed of retrotransposon-like sequences (Huijser 1988). In the case of *D. melanogaster*, three pairs of Y-loops, thought to correspond to kl-2, kl-3 and kl-5, have been observed. As with *D. hydei*, characterisation of the *D. melanogaster* loops at the cytological and molecular level has shown that they are also comprised largely of simple sequence repeats (Bonaccorsi *et al.*, 1990). Since these transcripts remain with the loops and do not appear in the cytoplasm and antibodies against both proteins and ribonulceoproteins (RNP's) are localised to the loops (Glatzer *et al.*, 1984) it is possible to see where the idea of the 'Y-loops' acting as some sort of scaffold or organisation centre originated.

Whilst these fertility factors obviously have an important function in the spermatogenic process, it seems likely from their mutant phenotypes that their role is more structural than regulatory (e.g. kl-5 possibly coding for the dynein in the axoneme). Consequently, other autosomal loci must exist that code for regulatory proteins as well as for other structural proteins.

Genetic Control Of Spermatogenesis.

As can be seen from the preceding sections, the process of spermatogenesis is extremely complex. It is further complicated by the fact that it does not occur by linear progression, with each structure having to be formed before the next can start. In spermatogenesis, each of the components seems to be derived from an independent pathway. Consequently, an early-acting defect that affects one structure may not affect the overall structure of the sperm until much later in the assembly process. For example, the mitochondrial derivative and the axoneme both form the axial complex. A defect in either one of these components does not affect the correct differentiation of the other. This poses problems for studying male-sterile mutations since many independent mutations often give rise to a similar phenotype when examined under the light microscope. Ultrastructural studies do not help in characterising when the defect occurred, they merely refine the defects obvious under light microscopy. Additionally, since many of the morphogenic events described depend upon the action of microtubules, defects in these proteins can affect the spermatogenic process.

Obviously, when studying complex processes the problem arises of how to determine which genes are actually specific for the processes. Disruption of normal cellular process such as transcription and translation will also cause 'spermatogenic defects' although obviously not specific to the spermatogenic

process. It appears that a large portion of *Drosophila* genes are involved in spermatogenesis. Using a variety of mutagens, is has been shown that ~15% as many genes can be mutated to give a male-sterile phenotype as can be mutated to give a lethal phenotype (Lindsley & Lifschytz, 1972; Lindsley & Tokuyasu, 1980). This figure appears to hold true for the X-chromosome and the autosomes. In addition, saturation mutagenesis of particular regions of the genome (Judd *et al.*, 1972) suggested that out of the 5000 salivary gland chromosome bands, 75% are represented by lethal mutations (this assumes 'one-gene-per-band'). Given the frequency of male-sterile mutations to lethal mutations (~15%) this gives ~500 loci which when mutated would give a male-sterile phenotype. In order to determine what proportion of these lethal mutations also resulted in male-sterility, temperature sensitive lethals (carried on the X) were raised under permissive conditions and their fertility examined (Lindsley & Tokuyasu, 1980). 16% of the mutations caused sterility under such conditions, whilst a further 18% caused sterility in adults shifted from the permissive to the non-permissive temperature. Assuming that what applies to the X also applies to the autosomes, and that sterility and lethality associated with these mutations is due to single gene defects, we can predict that as many as 1250 lethal complementation groups also play a role in male-fertility. This figure added to the 500 loci defined by sterility alone gives approximately 1750 loci required for male-fertility. However, this figure maybe an overestimate since out of the 500 loci representing male-sterile alleles some of these are likely to be weak alleles of vital genes.

Recent evidence regarding the organisation of genes in the *Drosophila* chromosomes suggests that these figures may have to be recalculated. Judd's theory of 'one-gene-per-band' has had to be revised. Recent studies by Kozlova (Kozlova *et al.*, 1994) have attempted to address this problem. Focusing again on the X-chromosome, they cloned the region from 91F-10A7, a region encompassing seven bands. Using saturation mutagenesis, twelve genes had

been previously localised to this region. They observed that within this region, the amount of DNA in each band varied from 4 kb to 195 kb. Thick bands and the finer inter bands seem to have different DNP packaging. Furthermore, the genetic content across this region was not constant, the finer bands contained between one and three genes whilst the thicker bands contained three genes and at least six transcribed DNA regions. Earlier work by Eberl (Eberl *et al.*, 1992) on region 17 of the polytene X-chromosome had identified 17 new genes by saturation mutagenesis in a region that comprises between 37 to 43 polytene bands. These studies led to the reclassification of the polytene bands into four different groups: polygenic bands contain moderately repeated genes such as histones and ribosomal RNA, oligogenic bands containing several genes in a band, simple bands containing only one gene per band and finally silent bands containing no genes or genes that only occupy a small part of that band. Moreover, as will be described later, there is evidence that the introns of some genes may contain the coding sequence of another unrelated gene (Kalderon & Rubin, 1988; Russell & Kaiser, 1993). Recent estimates suggest that the number of genes does not exceed the number of chromosome bands by more than a factor of 5 or 10. Consequently, the number of genes thought to have a spermatogenic function may be greater than the figure of 1750 derived using Judd's calculations and may in fact be between 8750 and 17500.

Whatever the number of independent genes, mutations affecting male-fertility can be classified into three groups: weak alleles of general metabolic genes, specific alleles of genes coding for vital functions which also have an additional spermatogenic function and those which represent spermatogenic loci.

1. General Metabolic Genes.

Alterations in many general metabolic pathways are likely to affect the process of gametogenesis. In *Drosophila*, both spermatogenesis and oogenesis require high levels of gene expression. During spermatogenesis, in the primary spermatocyte stage to produce all the gene products that will be required during the later stages of spermiogenesis, and in the female, to produce gene products that will support oocyte growth. Consequently, many of the temperature sensitive lethal alleles that have been isolated have an associated female-sterile as well as male-sterile phenotype. An example of a gene with such a pleiotropic phenotype is *ms(3)neo30*. This gene has been identified as a weak allele of *belle*, an essential gene coding for a helicase of the *vasa*/eIF4A type (Jones & Rawls, 1988). Another class includes genes such as *downy* and *brief* (Kiefer 1973; Lindsley & Zimm, 1992). Mutations in both these genes produce an abnormal bristle morphology and a male-sterile phenotype. In *downy* flies, the testis appears normal but spermatogenesis does not continue past meiosis. The formation of the nebenkern and the appearance of the 'Y-loops' are affected and the spermatids degenerate after a short elongation period. Neither of these genes has been characterised at the molecular level so their gene products are not known.

2. Genes With An additional Spermatogenic Function .

As the processes of mitosis and meiosis are essential for spermatogenesis, one would expect that mutations in genes involved in them would produce a sterile phenotype. Investigations into the effects of mitotic genes such as *abnormal spindle* (*asp*, Casel *et al.*, 1990; Gonzalez *et al.*, 1990) and *merry go- round* (*mgr*, Gonzalez *et al.*, 1988) on male mitosis show that for late-lethal or semi-lethal

alleles male mitosis I is affected. The *asp* gene product codes for a non-tubulin component of the microtubules which has been implicated in the formation of correct spindle structure or function. The defects that can be observed are variable sized nuclei (which may arise as a consequence of non-disjunction) and nebenkern of various sizes. Additionally, some of the spermatids may be multinucleate. The alignment of the mitochondrial derivative and the axoneme in the axial complex is disrupted. The mitochondria are abnormally sized and may align with more than one axoneme (Figure 1.14b). The process of individualisation is also affected with the abnormal sperm which form earlier in development remaining in syncytium. However, *asp* males are still capable of producing functional mature sperm with any abnormal sperm removed into the waste bag during the individualisation stage.

A secondary screen for male meiotic mutants was performed on a group of X-linked male-steriles (Lifschytz & Meyer, 1977). Selection was for the presence of multinucleate spermatids. None of the mutants isolated appeared to be deficient in the process of mitosis despite defective spindle formation in both meiosis I and II. The regulation of the meiotic cell cycle and process of spindle formation appears to be different from the control of mitosis since none of the gonial mitotic divisions were affected in these mutants. Consequently, although the same structural proteins (tubulins and actins) may be used by both these processes, mitosis and meiosis may act to regulate the action of these proteins independently. The gene *roughex* (*rux*) appears to modulate entry into the second meiotic division (Gonczy *et al.*, 1994). Mutations in this gene go through both meiosis I and II, but instead of stopping once these have taken place, the cell goes through an additional division. The *rux* gene product acts on Cyclin A during the premeiotic G_2 stage, modulating the activity of the Cyclin A-cdc2 kinase during meiosis I.

3. Spermatogenic Loci.

Since spermatogenesis involves such complex morphological changes, it is not altogether surprising that a large number of genes (between 8750-17500) have been postulated to act in this process. Some of these genes will be responsible for structural proteins such as sperm-tail proteins (Schafer *et al.*, 1993) and dynein in the axoneme, (Gepner & Hays, 1993; Neesen *et al.*, 1994) or responsible for the process of nebenkern and flagellar elongation and chromatin condensation. In addition, there appear to be testis-specific isoforms of some cellular proteins. Considering the vast changes that the spermatogonia undergo before becoming mature sperm and the vital role that microtubules have in this process (as well as in the processes of mitosis and meiosis) it is not surprising that mutations in these genes disrupt spermatogenesis. Microtubules achieve heterogeneity by expressing variants of several distinct tubulin genes. *Drosophila* is known to have several functional β-tubulin genes, including one which is testis-specific. In many organisms, there is a difference between those microtubules which form the cytoplasmic tubulin population and those comprising the axonemal structure of a flagellum or cilium. However, it appears for *Drosophila* that this is not the case. The β-tubulin subunit that is found in the axonemal complex also performs other roles during the spermatogenic process. Flies mutant in the β-tubulin gene produce no motile sperm and when examined at the ultrastructural level a range of defects is observed. The spermatid components are disorganised, aberrant axonemal structures are formed and the microtubules required for spindle formation also appear to be absent or malformed. Consequently, several meiotic defects are observed, movement of the chromosomes does not occur correctly or fails completely. In addition there is often a failure in cytokinesis. These observations have led to the suggestion that this tubulin has a role in meiosis, in the formation of cytoplasmic microtubules and microtubules of the axonemal complex. It is synthesised just

before the onset of meiosis, suggesting that it has no mitotic function and accounts for 80% of the β-tubulin found in the fly and 40% of the total tubulin.

Genes involved in cell cycle control are also expressed in the testis, such as the *Drosophila* homologues of cdc25, *string* and *twine*. *twine* is expressed in a different subset of cells from *string* (Alphey *et al.*, 1992 , Courtot *et al.*, 1992). Its role differs from that of *string* as it is believed to function only during the meiotic process and not during mitosis. However, a mitotic function can not totally be excluded as along with *string* it can be detected in the early embryo prior to cellularisation. *twine* expression is restricted to the germline of both males and females during the later stages of gametogenesis. In the male, this expression occurs in the growth phase of the cyst before meiosis is due to start.

The defect in the *twine* product that results in a male/female sterile phenotype is a nonsense mutation near to the active site in a region conserved in all known cdc25 homologues. Mutants in *twine* never enter into the first meiotic division. The rest of the spermatogenic processes described earlier still occur (despite the absence of meiosis) resulting in the formation of an apparently normal looking sperm. The structure of the nucleus is obviously affected; the head of the sperm does not form the rod shape presumably due to the bulk of the 4N nuclei.

The use of enhancer trap elements has facilitated the identification of testis-specific transcripts (Bownes 1990; Gonczy *et al.*, 1992; Castrillon, *et al.*, 1993). This approach overcomes the problem of identifying a male-sterile phenotype, which as discussed could arise due to mutations in genes other than those strictly involved with the spermatogenic process. Using this approach, Castrillon identified 83 recessive autosomal mutations. 58 of which represent insertions giving rise to a male-sterile phenotype (including some male-steriles which

had previously been identified). These insertions were classified into various groups depending on their phenotype. These were: proliferation phase defects, growth phase defects, meiotic entry defects, meiotic division defects, postmeiotic defects, behavioural defects and sperm transfer defects. Castrillon's insertion lines provide additional information towards elucidation of the genetic control of spermatogenesis. The list of genes now known to have an effect on spermatogenesis is rapidly increasing. Fuller (1994) describes 59 loci (defined by various mechanisms such as EMS & P-element mutagenesis and heterologous probing) in *Drosophila* which are known to affect the various stages and processes which occur in spermatogenesis. These defects range from those affecting germline survival or proliferation such as *chickadee* (*chic*), to meiotic division such as *twine*, those affecting the mitochondrion, *nebbish* (*neb*) and the axoneme, *whirligig* (*wrl*). In fact, for each of the stages described previously a mutation affecting that process has been isolated and in many cases the gene products are beginning to be characterised at the molecular level. However, Fuller's list is not exhaustive, for example it does not include the sperm tail protein gene family Mst(3)CPG identified by Schafer (1986). The Mst(3)CPG cluster of seven genes has been extensively characterised. cDNAs corresponding to members of this family have been sequenced and code for a short proteins containing CYS-PRO-GLY (CPG) repeats, such that 38% of the amino acids are cystine, 29% are glycine and 20% are proline. A gene fusion was constructed. The Mst(3)CPG coding sequence was subcloned into a transformation vector containing a truncated *E. coli lacZ* gene. Fusion proteins were produced which are detected in the testis. Antibodies against the fusion protein label the whole of post-elongation spermatids, suggesting that these proteins have a role in axoneme formation. It is possible that they may form the accessory structures formed late during axoneme development. A regulatory element responsible for spermatocyte-specific transcriptional and translational control has been identified (Kuhun *et al.*, 1988; Schafer *et al.*, 1990, Kempe *et al.*, 1993).

As described, the process of spermatogenesis is a sequential developmental programme resulting in the production of a mature sperm cell. In addition to 'spermatogenic' loci, this process also requires meiotic processes, which occur simultaneously with spermiogenesis to produce mature sperm from the diploid spermatocyte. These pathways are regulated by different genes and take place independently from one another. The genetic components of spermatogenesis can be split into two groups:- those that provide a structural function and those that regulate the structural genes. Many of the genes isolated that appear to function in the spermatogenic process represent weak alleles of vital or metabolic genes. Consequently 'Reverse Genetic' approaches that do not require the presence of a male-sterile phenotype may be useful in identifying 'true' spermatogenic loci.

In this laboratory, Russell and Kaiser performed a differential screen in order to isolate transcripts expressed in the male. As described in more detail in chapter 3, they succeeded in isolating 28 male-specific/elevated clones (Russell 1989). At least two of these clones have been shown to have a pattern of expression that is restricted to the testes.

The Mst 40 locus identified in this screen has an unusual genomic structure. It consists of 1.4 kb tandemly repeated unit that is situated in the β-heterochromatin at the base of 2L. This repetitive structure is conserved between all the *D. melanogaster* strains examined by Russell but is not found in other *Drosophila* species such as *D. virilus*, *D. simulans* or *D. yakuba*. Whilst this region of the chromosome was known to contain seven vital or semi-vital loci, none of these correspond to the Mst 40 locus. Northern blot analysis showed that transcripts are present in male third instar larvae, male adults and pupae. In addition, expression was also found in early embryonic stages (0-4 hr). Expression of this transcript was abolished in males carrying the grandchildless mutant *tudor*,

mutant *tudor,* a maternal effect gene, mutations in which result in progeny which are agametic. Since no expression is observed when there is no germline, this suggests that this transcript may be germline specific. Further evidence for a germline function for this clone was provided by data from *in situ* hybridisation to whole mount embryos and testis dissected from third instar larvae. Expression in the embryo is restricted to two regions of the early cleavage nucleus. During later embryonic development the signal is restricted to the pole cells (which as described earlier are the precursors of the gonad in both sexes). Hybridisation to the larval testis shows that the signal is restricted to the spermatocytes. Three cDNA clones corresponding to the genomic clone have been isolated and sequenced. An open reading frame (ORF) has been found in each of the cDNAs but they do not show any homology to previously characterised sequences. Sequence of a genomic fragment, in contrast, shows ~80% homology to the Long Terminal Repeat (LTR) of the retrotransposon HMS Beagle. This genomic fragment does not carry sequences in the LTR which are required for transposition and it is likely that this sequence is unable to transpose. Consequently, although this clone is expressed in the germline, it is not clear whether it has a protein coding function or whether in common with other repetitive DNA such as *Supressor of stellate Su(ste)* it has an unknown function (Russell & Kaiser, 1994).

The same screen also identified a group of clones (gS1, gS8 & gS9) one of which appears to encode a histone-like protein similar to H1/H5 and also to mammalian cysteine-rich protamines (Russell & Kaiser, 1993). These genes have a complex genomic organisation with some copies located on the Y-chromosome and a single copy on the autosome. The autosomal copy (S8) is located at 77F on the right arm of the third chromosome. It is located within the first intron of the gene encoding a regulatory subunit of the cAMP dependent protein kinase (Kalderon & Rubin, 1988). This copy translates as a polypeptide

with a homology to both a histone-like protein similar to H1/H5 and also to mammalian cysteine-rich protamines. The Y-associated sequences do not lie within any of the regions corresponding to the fertility factors. Northern blot analysis shows that transcripts are detected in male third instar larvae, male adults and pupae. As with Mst40, no expression is detected in agametic males. Preliminary data from *in situ* hybridisation to whole mount larval testis suggest that transcription is restricted to spermatocytes.

This work describes one approach, differential screening which has identified several loci containing male-specific transcripts. Although none of these transcripts have been shown to function in spermatogenesis, preliminary evidence from the genomic DNA clones characterised by Russell (Russell & Kaiser; 1993, 1994) suggests that this might be the case. The expression of transcripts from two of these genomic DNA clones has been shown by northern blot analysis and whole mount tissue *in situ* hybridisation to be restricted to the testis. One of these genomic DNA clones has copies present on both an autosome and the Y-chromosome. The Y-chromosome appears to function only during spermatogenesis. Consequently any genes located here are likely to have a spermatogenic function.

The genomic DNA clones characterised in this work were isolated by Russell (Russell, 1989) in the same screen. The initial characterisation (discussed in chapter 3) was to ensure that all of the genomic clones to be studied were not duplications and to isolate the region containing the male-specific transcript from the genomic DNA clones. These regions were then used as probes for further analysis. The hybridisation of these probes to polytene chromosomes will assist in the determination of whether these represent novel male-specific transcripts or reisolates of male-specific transcripts characterised by other laboratories (Schafer, 1986; DiBenedetto, 1987). In addition, as discussed in

chapter 5, the use of *in situ* hybridisation to whole mount tissues may suggest function in the absence of known mutations. Northern analysis using male and female flies and mutant strains readily available in *Drosophila* will be used to determine which tissues express the transcripts. Finally, the isolation of cDNA clones corresponding to each of the genomic DNA clones may assist in assigning a function. These clones were isolated in the same screen which produced Russell's two male-specific transcripts. It is likely (for reasons discussed more fully in chapter 3) that these will also represent transcripts which are restricted to the germline which may have a spermatogenic function.

Aims and Objectives.

1. To classify the differential clones into groups representing different genes.
2. To identify the male-specific region of each class of clone.
3. Determination of chromosomal localisation of the male-specific region.
4. Northern analysis to determine the size, abundance, sex- and tissue-specificity of the transcripts.
5. Whole mount *in situ* hybridisation to determine when and where the transcripts are expressed.
6. Screening cDNA libraries to identify cDNA clones corresponding to the original male-specific genomic fragment.
7. Sequence analysis of the resulting cDNA clones.

Chapter Two Materials and Methods

Materials.

Enzymes:

Restriction enzymes - Promega and BRL, T4 DNA ligase, *E.coli* DNA polymerase I, Klenow fragment *E.coli* DNA polymerase I - BRL, AMV reverse transcriptase - Pharmacia.

Reagents:

oligo(dT)cellulose type 7, oligo(dT)$_{12\text{-}18}$, dNTP's, RNAGUARD, pd(N_6), Sephadex G-50, Polyadenylic acid, SDS- Pharmacia.

biotin-16-dUTP, agarose - Boehringer Mannheim.

Streptavidin-Biotinylated Horseradish Peroxidase complex, Hybond-N and Hybond-C$^+$ membrane - Amersham.

Acrylamide, bis acrylamide, ammonium persulphate, TEMED - BioRad.

Anti-biotin FITC labelled antibody, BSA (Fraction V), Giemsa, Glycerol mountant, DAB, Antifoam A - Sigma.

Avidin/biotin blocking kit, secondary goat anti-rabbit-Texas Red labelled antibody, Avidin/FITC conjugate, Avidin/Texas Red conjugate, Fluoromount mountant -Vector.

Primary rabbit anti-*LacZ* antibody - Capell.

1kb ladder, RNA size ladder - BRL.

Sequenase sequencing kit (Version 2.0) - United States Biochemical.

DPX mountant, Triton X-100, Tween 20, Sarcosine -BDH.

Normal goat serum- Scottish Antibody Production Unit (SAPU).

Radioisotope - NEN

2.1 Bacterial Strains.

Strain	Genotype	Reference
NM621	(*hsd*R *mcr*A *mcr*B *sup*E44 *rec*D1009)	Whittaker *et al.*, (1983)
NM514	(*hsd*R *arg*H *gal*E *gal*X *str*A *lyc*b7)	N. Murray (1983)
XL-1 Blue	*rec*A$^-$ (*rec*A1 *lac end*A1 *gyr*A96 *thi hsd*R17 *sup*E44 *rel*A1 (F′ *pro*AB *lac*Iq *lacZ*ΔM15 Tn10 tet^r)	W.O Bullock *et al.*, (1987)
MC1061	*arad*139 Δ(*ara leu*)7697 Δ(*lac*)χ74 *gal*U *gal*K *hsd*R2 *str*A *mcr*A *mcr*B	Bio-Rad
DS941	(*rec*F143 *pro*A7 *str*31 *thr*1 *leu*6 *tsx*33 *mlt*12 *his* 4 *arg*E3 *lac*Iq *lacZ*ΔM15)	D. Sherratt (pers. comm.)

2.2 Plasmids and Lambda Vectors.

Plasmid	Genotype	Reference
YP-1	*Bam*HI-*Bgl*II fragment of *D.melanogaster* yolk protein 1 in pGem-1	Hung & Wensink, (1981)
pBRrp49	*Eco*RI/*Hind*III fragment of the *D. melanogaster* ribosomal protein 49 cloned in pBR322	O'Connell & Rosbash, (1984)
pBluescript	*amp*^{r}	Short *et al* (1988)
Mst(3)336	male-specific genomic lambda clone.	Di Benedetto *et al.*, (1987)
NM1149	λ*b*538 s*r*I λ3^o imm434 s*r*Iλ 4^o s*hnd*IIIλ 6^o srIλ 5^o. This is an insertion vector. cDNAs are cloned into the unique *Eco*RI site of the immunity region.	Murray, (1983)
EMBL3	s*b*hI λ1^o *b*189 KH54 *chi*C s*r*Iλ 4^o *nin*5 s*r*I λ5^o This is a replacement vector, which means that the 'internal stuffer' region is removed and replaced by genomic DNA. Inserts in this case are released by digestion with *Sal*I.	Frieschauf *et al.*, (1983)

2.3 *Drosophila* Strains.

All *Drosophila* stocks used were as described by Lindsley & Grell (1968) or Lindsley & Zimm (1992). The stocks were maintained at either 25°C or 18°C depending on genotype. A description of the strains used can be found in Table 2.1.

Description of crosses used to generate flies homozygous for the following mutations:

***transformer 2* (*tra2*):** (Baker & Ridge, 1980). *cn bw tra2^{B}/CyO* males were crossed with virgin *cn bw tra2^{B}/CyO* females. This cross generates females flies that have all their somatic tissue transformed into male tissue (pseudomales). They were mated at 25°C for three days after which time the parents were removed. Homozygous *tra2^{B}* flies were identified by their straight wings, the males identified by their bar-eyed phenotype.

***tudor* (*wc^{8}*):** (Boswell & Mahowold 1985). This is a maternal effect mutation resulting in flies lacking a germline. Homozygous F^{1} virgin females were collected from a *bn wc^{8} sp/CyO* stock. These were crossed with their homozygous brothers to produce an F_2 generation lacking pole cells.

2.4 *E. coli* and *Drosophila* Culture Media.

L-Broth: 10g Bacto-tryptone (Difco), 5g yeast extract (Difco), 5g NaCl, 1g glucose made up to 1 litre in distilled water and adjusted to pH 7.0 with NaOH.
0.7% (w/v) Top Agarose: 0.7g agarose added to 100ml of L-broth, containing 10mM $MgSO_4$ heated to dissolve the agarose and left to cool to 50°C before use.

Strain Genotype	Reference
Oregon R (wt)	Lindsley & Grell (1968)
Canton S (wt)	Lindsley & Grell (1968)
m56i (wt)	Schalet & Lefevre (1973)
cn bw tra2^{B}/CyO	Baker & Ridge (1980)
wc8 bw sp/CyO	Boswell & Mahowold (1985)
ms(90)E	Castrillon *et al.*, (1993)
pto	Castrillon *et al.*,(1993)

Table 2.1 Genotypes of *Drosophila* strains used in this study.

Glasgow fly food: 10g Bacto-agar (Difco), 15g sucrose, 30g glucose, 35g yeast, 15g maize meal, 10g wheatgerm, 30g treacle, 15g Soya flour, 0.1% (v/v) Nipagen and 0.5% (v/v) proprionic acid per litre of water.
Rich Larval media: 100g glucose, 100g yeast, 20g agar, 0.1% (v/v) Nipagen per litre of water.

Methods.

2.5 Buffers and Solutions.

Electrophoresis.

10x TBE buffer pH 8.3: 109g Tris, 55g boric acid, 9.3g $Na_2EDTA.2H_20$ made up to 1 litre in distilled water, the pH is 8.3.
50x TAE buffer: 242g Tris, 57.1 ml glacial acetic acid 100 ml 0.5M EDTA pH 8.0 made up to 1 litre with distilled water.
10x MOPS: 46.26g MOPS, 2.1g NaOAc, 3.72g EDTA made up to 950 ml with distilled water, pH to 7.0 with NaOH made up to 1 litre.
10x Agarose gel loading buffer: 0.25% (w/v) Bromophenol blue, 0.25% (w/v) xylene cyanol, 25% (w/v) Ficoll (400) made up in distilled water.
Formaldehyde loading buffer: 5ml glycerol, 4 ml 0.5M EDTA pH 8, 20mg bromophenol blue, 20mg xylene cyanol 5g Ficoll (400) made up to 20ml with distilled RF water.

DNA Manipulations.

Restriction buffers: 10X buffers supplied by BRL and Promega.
1xTE Buffer: 10mM Tris.HCl, 1mM EDTA pH 7.6.

Hybridisations.

100x Denhardts solution: 0.2mg/ml BSA (Fraction V), 0.2mg/ml Ficoll (400), 0.2mg/ml Polyvinylpyrolidone.

20xSSC: 3M NaCl, 300mM trisodium citrate pH 7.0.

20x SSPE: 3.6M NaCl, 200mM NaH_2PO_4, 20mM Na_2EDTA pH 7.4.

10% (w/v) SDS : 100g of SDS made up to 1 litre in distilled water.

Denaturing solution: 1.5M NaCl, 0.5M NaOH.

Neutralising solution: 1.5M NaCl, 0.5M Tris.HCl pH 7.2, 0.001M EDTA.

SSC Prehybridisation solution: 6xSSC (w/v), 0.1% (w/v) Polyvinylpyrolidone, 0.1% (w/v) Ficoll (400), 0.1% (w/v) BSA (Fraction V), 50mM $NaPO_4$ pH 6.8, 1% (w/v) SDS, 30µg/ml Poly A, 100µg/ml sonicated Herring sperm DNA, 0.005% (w/v) $NaPP_i$, 1mM EDTA pH 8.

Formamide prehybridisation solution: 50% (v/v) Formamide, 1% SDS (w/v), 0.005% (w/v) $NaPP_i$, 100µg/ml sonicated Herring sperm DNA, 15µg/ml Poly A, 10x Denhardts.

Biotin Prehybridisation solution: 0.6M NaCl, 50mM Sodium Phosphate pH 6.8, 1xDenhardts, 0.5mM $MgCl_2$.

Washing solutions:

Low stringency wash: 1xSSC, 1% (w/v) SDS.

High Stringency wash: 0.2xSSC, 1% (w/v) SDS.

Nucleic Acid preparation, purification and general purpose.

Phenol: All phenol used contained 0.1% (w/v) 8-hydroxyquinoline to prevent oxidation and was saturated with 1M Tris.HCl pH 8.0.

Birnboim Doly I: 50mM glucose, 25mM Tris.HCl pH 8.0, 10mM EDTA.

Birnboim Doly II: 0.2 N NaOH, 1% (w/v) SDS. Made fresh each time.

Birnboim Doly III: 5M KOAc pH 4.8. Mix 60ml of 5M KOAc, 11.5ml of glacial acetic acid and 28.5ml H_2O.

PEG solution: 20% (w/v) PEG MW 6000, 2.5M NaCl.

Phage buffer: 10mM Tris.Cl pH 7.5, 10mM $MgSO_4$.

10mg/ml RNase/DNase solution: DNase I and RNase A were mixed to give a 10mg/ml stock solution. Stored at -20°C.

RNA Denaturing solution: 4M guanidinium thiocyanate, 0.1M Tris.HCl pH 8.0, 10µl antifoam A, made up to 100ml with DEPC treated water. Added 0.1M β–mercaptoethanol immediately before use.

Homogenisation buffer: 0.03M Tris.HCl pH 8.0, 0.01M EDTA, 0.1M NaCl, 10mM β–mercaptoethanol.

Nuclear lysis buffer: 0.1M Tris.HCl pH 8.0, 0.1M EDTA, 0.1M NaCl, 0.5 mg/ml Proteinase K.

Grinding buffer: 5% (w/v) sucrose, 80mM NaCl, 0.1M Tris.HCl pH 8.0, 0.5% (w/v) SDS, 50mMEDTA

Binding buffer (BB) for Oligo (dT) cellulose: 0.5M NaCl, 10mM Tris.HCl pH 8.0, 1.0% (v/v) Sarcosine, 1mM EDTA. Made up to 500ml for 2x binding buffer, 1 litre for 1x binding buffer.

5x 1st strand cDNA buffer: 250mM Tris.HCl pH 8.3, 700mM KCl, 50mM $MgCl_2$, 50mM DDT.

4x Random Priming buffer: 250mM Tris.HCl, pH 8.0, 25mM $MgCl_2$, 100µM dNTPs, 1M Hepes pH 6.6, 27 A_{260} units/ml random hexanucleotides.

10x Nick Translation Buffer: 0.5M Tris.HCl pH 7.5, 0.1M $MgSO_4$, 1mM DDT, 500µg/ml BSA (Fraction V).

Spermidine: 1M stock solution, dissolve 2.54g in 10ml water. Working solution 0.2M, stored at -20°C.

Ampicillin: 50mg/ml stock solution in sterile distilled water. Working solution 50µg/ml. Stored at -20°C.

Tetracycline: 12.5mg/ml stock solution in absolute ethanol. Working solution 12.5µg/ml. Stored at -20°C.

X-gal: Dissolved in dimethylformamide to give a 20mg/ml stock solution, used at a final concentration of 20µg/ml. Stored at -20°C.

IPTG: Dissolved in sterile distilled water. Stored at -20°C as 20mg/ml stock solution. Used at a final concentration of 20µg/ml.

Polyacrylamide gel mix: For a 6% poly acrylamide gel, mix acrylamide stock solution (Acrylamide: bis 19: 1) with 7M urea and 1xTBE. Polymerisation was initiated by the addition of 1ml of 10% (w/v) ammonium persulfate followed by 20µl of TEMED to 150ml of 6% acrylamide/urea.

Ringers solution: 13.6g KCl, 2.7g NaCl, 0.33g $CaCl_2.H_2O$, 1.21g Tris.HCl. pH to 7.2 with 1M HCl made up to 1 litre.

4% (w/v) Paraformaldehyde: 40g Paraformaldehyde added to 800ml warm PBS, make up to 1L with PBS.

PBS: 8g NaCl, 0.2g KCl, 1.44g NaP_2HPO_4 and 0.24g KH_2PO_4 to 800ml H_2O. pH adjusted to 7.4 using HCl, made up to 1L.

PBT: PBS containing 1% (v/v) Tween 20.

PAT: 1% (v/v) Triton X-100, 1% (w/v) BSA (Fraction V) in PBS.

2.6 Nucleic Acid Isolation

2.6.1 Bacteriophage DNA isolation:

2.6.1.1 Plate lysate method.

The method for preparing small amounts of bacteriophage DNA was essentially as described in Sambrook *et al.*, (1989). A single isolated plaque was picked and put into 1ml of phage buffer containing a drop of chloroform. This was left at 4°C for several hours to allow the phage to diffuse out of the top agarose. Approximately 10^5 pfu phage obtained in this way were added to 150µl of NM621 plating cells and incubated on the bench for 20 min. 3ml of 0.7% (w/v) top agarose were added and this was poured onto an L-agarose plate. The plates were incubated at 37°C for 8 hours. The plates were then covered with 5ml phage buffer and left on a 'rock and roller' for 2 hours. The phage buffer was decanted into an oakridge tube and spun at 15k for 5 min. To 4ml of this solution was added 10µl of 10mg/ml RNase/DNase solution and this was incubated at room temperature for 30 min. 4ml of a 20% (w/v) PEG, 2.5M NaCl solution were added and the solution left on ice for 60 min. This was spun at 15K for 10 min in a JA15 rotor in a Beckman centrifuge to pellet the phage particles. The supernatant was removed, the tubes inverted to drain off the remaining PEG solution. The phage pellet was resuspended in 500µl of Phage buffer and transferred to an eppendorf tube. 5µl of 10% (w/v) SDS and 5µl of 0.5M EDTA were added and the solution incubated at 70°C for 15 min. 500µl of phenol were added and the solution mixed by inversion and then spun for 10 min in a eppendorf centrifuge. 450µl of the aqueous phase was removed to an eppendorf and 450µl of phenol/chloroform added, mixed by inversion and spun for 10 min. 400µl of the aqueous phase was removed and 400µl of chloroform added. After centrifugation, 350µl of aqueous phase was removed

and the phage DNA precipitated at room temperature for 30 min by the addition of 350µl of isopropanol. The solution was centrifuged for 30 min and the resulting pellet washed in 70% ethanol before being resuspended in 50ml TE pH 8.0. The yield obtained was typically 5-10µg of DNA.

2.6.1.2 Large-scale preparations.

This protocol was as described by Lehrach (1992). The phage were incubated with 150µl of NM621 plating cells as described for the plate lysis method. The plates were covered with 4ml phage buffer. 400ml of L-broth containing 10mM $MgSO_4$ was inoculated with 1ml of NM621 cells. These were grown at 37°C until their OD_{600} was 0.3. The phage buffer from the plate lysates was added and the cultures grown until lysis had occurred. After lysis, 2ml of chloroform were added to each culture and these were shaken for a further 5 min. After centrifugation for 20 min at 4K to remove cellular debris, NaCl (40g/L) and PEG 6000 (100g/L) were added. It was left on a 'rock and roller' until all the PEG had dissolved. The solution was left overnight at 4°C to precipitate the phage. In order to pellet the phage, the solution was spun at 10K for 20 min. The supernatant was removed, the centrifuge tube rotated 180° and then centrifuged for a further 10 min. The tubes were left inverted to ensure that all the PEG solution had been removed. The pellet was resuspended in 11ml phage buffer. This was mixed with 5ml chloroform and centrifuged at 15K for 10 min. 0.71g CsCl/ml were added to the phage suspension and this was mixed in a Falcon tube. This solution was loaded into a polyallomer tubes and spun in a Beckman ultracentrifuge at 49K for 16 hours in a Beckman Ti70 rotor at 20°C. The phage band was removed from the tube using an 18 gauge needle and a syringe. It was dialysed for an hour against a solution of 10mM NaCl, 50mM Tris.HCl pH 8.0 and 10mM $MgCl_2$. The phage solution was dialysed against three changes of this buffer. The phage were decanted into a centrifuge tube,

EDTA was added to give a final concentration of 20mM, Proteinase K to a final concentration of 50µg/ml and SDS to a final concentration of 0.5% (w/v). This was incubated at 65°C for 1 hour. The solution was phenol/chloroform extracted and dialysed against several changes of 1xTE. This usually gave a yield of approximately 200µg bacteriophage DNA.

2.6.2 Plasmid DNA isolation.

2.6.2.1 Large scale plasmid preparation.

Large scale preparations of plasmid DNA were prepared as described in Sambrook *et al.*, (1989) which is a modified version of the Birnboim and Doly (1979) alkaline lysis method. The plasmid containing bacteria were inoculated into L-broth containing the appropriate antibiotic. This was grown with shaking at 37°C. The culture was spun down and the bacterial pellet resuspended in 20ml solution I, to this suspension was added 40ml of freshly prepared solution II. The contents were mixed by inversion and incubated at room temperature for 5-10 min. 20ml of ice cold solution III were added and the solution shaken vigorously before incubating on ice for 10 min. The resulting white precipitate was removed by centrifugation at 4000 rpm for 15 min at 4°C. The supernatant was mixed with 0.6 vol of isopropanol and left at room temperature for 10 min. The DNA was pelleted by centrifugation at 5000 rpm in a JA10 rotor in a Beckman for 15 min at room temperature. The tubes were left inverted to ensure that all the supernatant had been removed. The pellet was washed with 70% ethanol, left to dry and resuspended in 10ml of TE pH 8.0. The plasmid DNA was purified using CsCl gradients. 1g CsCl was added per ml of DNA solution. In addition, 0.5ml of EthBr (10mg/ml) was added. The solution was transferred to a polyallomer tube and ultracentrifuged at 49K

for 18 hour in a Ti70 rotor at 20°C. The resulting plasmid band was removed from the tube, extracted with water saturated butanol to remove the EthBr and dialysed against 1xTE. Yields of 1-2mg were obtained.

2.6.2.2 Small scale plasmid preparation.

The above protocol was also used to produce midi and mini plasmid preparation. Mini-prep volumes are given in parenthesis. In this instance, 15ml (1.5ml) of the overnight culture was used, 1ml (80µl) of solution I, 2ml (80µl) of solution II, 1.5ml (150µl) of solution III and 6ml (600µl) isopropanol. The resulting plasmid DNA was resuspended in 200µl (40µl) of TE pH 8.0 containing RNase A. Yields of 50µg (5µg) were obtained.

In some cases, the Promega Wizard_ preparations were used to isolate small amounts of DNA (~20µg) for sequencing reactions. Procedure followed was as described by the manufacturer.

2.6.3 .1 Genomic DNA isolation.

Approximately 1g of flies were added to a mortar which contained liquid Nitrogen. Before the liquid Nitrogen evaporated completely, the flies were ground to a fine powder. Using a small paint brush (cooled in N_2) the powder was transferred from the mortar into a 15ml Wheaton homogeniser (on ice) containing 9 ml of ice cold Homogenisation buffer (HB). To this solution was added 500ml of 10% (v/v) Triton X-100. This solution was homogenised thoroughly and the resulting homogenate decanted through gauze into a sterile 30ml Corex tube on ice. The homogenate was spun immediately at 4K for 10 min at 4°C in a cooled rotor. The supernatant was decanted and the nuclear pellet resuspended in 1ml of ice cold HB (No Triton X-100). 5ml of Nuclear

lysis buffer and 200µl of 30% (v/v) Sarkosyl were added to this solution. This was mixed by swirling until lysis had occurred. The lysate was incubated overnight at 37°C. After centrifugation to remove the debris, the supernatant was decanted into a preweighed Falcon Tube and 1.25g CsCl per ml of homogenate was added. The solution was loaded into Polyallomer tubes which were filled with 1.25g/ml CsCl/dH_2O. The tubes were ultracentrifuged at 45K for 24 hours in Ti70 rotor at 25°C. Samples were collected by dripping the gradient through an 18 gauge needle at bottom of the centrifuge tube. 1.5 ml fractions were collected initially, and then 0.5 ml fractions once it appeared more viscous. The concentration of the DNA samples was crudely estimated using EthBr plates as described in Section 2.7. The best fractions were pooled and dialysed against TE. The yields of genomic DNA obtained were generally 100-200µg/gram of starting material.

2.6.3.2 Genomic DNA Preparation For Plasmid Rescue.

Approximately 100 flies were homogenised in 0.5ml of Grinding buffer using a 15ml Wheaton Homogeniser. The homogeniser was rinsed with an additional 0.5ml of grinding buffer and immediately placed at -20°C for 30 min. The sample was then incubated at 70°C for 30 min. To this solution was added KOAc to a final concentration of 160mM. This was then incubated on ice for 30 min. The resulting white precipitate was removed by centrifugation and the DNA solution extracted twice with an equal vol of phenol/chloroform followed by a chloroform extraction. The DNA was precipitated using 0.75 vol of isopropanol. It was washed with 70% EtOH before being resuspended in 0.5ml of TE containing 20µg of RNase A. DNA concentrations were estimated using EthBr plates as described previously. Approximately 1µg of genomic DNA was restricted with the appropriate enzyme, depending on the type of enhancer trap element used (Hamilton, 1991; Mlodzik, 1992). This was ligated in a total

volume of 200µl using 1U of Ligase. The reaction was left at room temperature overnight. It was then precipitated using 2 vol of EtOH and 1/10 vol of 3M NaOAc followed by incubation at -20°C for 30 min. After centrifugation the resulting pellet was washed with 70% EtOH and resuspended in 10µl sterile H2O. 5µl of the DNA solution were transformed into MC1061 competent cells using a Bio-Rad electroporator. DNA from the resulting transformants was prepared as described in Section 2.6.2.2.

2.6.4.1 Isolation of RNA.

This procedure is adapted from the protocols described in Chirgwin (1979) and Chomczynski & Sacchi (1987). 1g of flies were homogenised in 20ml of RNA Denaturing solution using a Kinematica Polytron. The homogenate was then transferred to a 30ml RF Corex tube or RF 30ml Falcon tube. To this was added 1/10 vol 2M Na OAC pH 4.0, an equal vol phenol and 1/5 vol chloroform. This solution was shaken vigorously before being incubated on ice for 15 min. The solution was spun in a bench top centrifuge for 15 min at 15,000 rpm. The clear upper phase was removed using an RF glass pipette. An equal volume of isopropanol was added and the RNA precipitated at -20°C for an hour. After this time, the solution was centrifuged for 10 min at 10,000 rpm. The supernatant was discarded and the RNA pellet resuspended in 5ml RNA denaturing solution. The RNA was again precipitated using an equal volume of isopropanol followed by incubation for an hour at -20°C. After centrifugation the RNA was resuspended in 5-10ml of DEPC treated water. The RNA was stored as an ethanol precipitate at -20°C . Yield was 1mg/gram flies.

2.6.4.2 Isolation of Poly A+ mRNA.

0.5g of oligo (dT) were equilibrated in 10ml of 1x binding buffer. (BB). This was left to swell for 1 hour at 4°C. A 5ml syringe was plugged with RF glass wool and filled up with oligo dT cellulose to give a bed volume of 1ml. The column was washed with 10ml of 0.1M NaOH and rinsed with several volumes of RF water until the pH of the effluent was less than pH 8.0. The column was washed with 10-20 volumes of 2xBB. The RNA was dissolved in 2xBB, heated to 65°C, cooled on ice and added to the column. The effluent was collected, reheated and re-applied to the column, this procedure was repeated so that the effluent was applied to the column three times. The column was now washed with 10-20 volumes of 1xBB. The bound RNA was eluted from the column using RF water that had been heated to 65°C (usually 3ml of water were used). An equal volume of 2xBB was added to the eluted RNA. The column was treated with NaOH as before and the whole procedure repeated. Once the RNA had been eluted, it was precipitated using 1/10 vol of NaOAc pH 5.2 and 2.5 vol of ethanol and left at -20°C overnight. Generally, yields of 20-50µg of poly(A)$^+$ mRNA/g of tissue were obtained.

2.7 Quantification Of Nucleic Acids.

In order to determine the concentration of DNA or RNA in a sample, readings were taken at wavelengths of 260nm. An OD_{260} =1 corresponds to ~50µg/ml for double-stranded DNA, ~40µg/ml for RNA and ~33µg/ml for oligonucleotides. In other instances, the concentration of DNA was estimated by spotting the sample and known standards onto the surface of a 1% (w/v) agarose gel containing EthBr (0.5µg/ml). The gel was photographed using short-wavelength UV illumination (254nm) and the concentration of the DNA

sample was estimated by comparing the intensity of fluorescence in the sample with those of known DNA concentration standards.

2.8 *In Vitro* DNA Manipulations .

All procedures for transforming *E.coli,* subcloning of fragments and restriction enzyme digests were carried out as described by Sambrook *et al.,* (1989).

2.8.1 Estimation of fragment sizes.

Unknown fragment sizes were calculated as described below. Migration distances of known molecular weight markers were measured from each gel. These were used to construct calibration curves by plotting the size of the fragment (in kb) on a logarithmic scale against the distance migrated. The resulting curve could then be used to approximate the sizes of other fragments on the gel by measuring the distance each fragment had migrated and interpolation.

2.9 Electrophoresis And Blotting.

2.9.1 DNA Electrophoresis And Southern Blots.

TBE agarose gels of 1% (w/v) (for plasmid and phage DNA) and 0.8% (w/v) (for genomic DNA) were used to separate the DNA on the basis of size. DNA was visualised by staining the gel in 1xTBE buffer containing 0.5µg/ml EthBr for 30 min, followed by destaining in 1xTBE containing 1mM MgSO4 for 30 min. Gels were photographed using a Polaroid camera loaded with 667- land film and fitted with a Kodak Wratten filter No. 23A. Markers used were either

Lambda cI875 DNA digested with *Hin*dIII and *Eco*RI or 1kb ladder. If required the DNA was transferred onto Hybond N membrane following the procedure of Southern (1975). After electrophoresis, the DNA was immersed in 0.25M HCl for 15 min on a shaking platform to depurinate the DNA (for Genomic Southerns), they were then soaked in Denaturing solution for 20 min, followed by soaking in two changes of Neutralising solution for 20 min each. It was left to transfer onto the membrane overnight. DNA was fixed to the membrane by baking at 80°C for two hours or by UV irradiation using a Stratalinker_ (Stratagene).

Hybridisation Conditions.

All hybridisations were initially carried out in plastic sealed bags and agitated in a shaking water bath at 65°C overnight. Latterly they were carried out in hybridisation tubes in a Hybaid oven at 65°C. The membrane was prehybridised in SSC prehybridisation solution for at least one hour before the addition of the probe. The probe was boiled for 10 min and snap cooled before being added to the filter and left to hybridise for at least 16h. Filters were washed in 2xSSC at room temperature followed by two washes for 15 min in low stringency wash at 65°C, and two washes for 15 min in the high stringency wash also at 65°C. The filters were blotted dry and covered in Saran Wrap_. Autoradiography of probed filters was carried out at -70°C, using intensifying screens and Fuji NIF RX X-ray film. Films were developed using a Kodak X-Omat film processor.

2.9.2 Reverse Northern.

This procedure was essentially as described in Fryxell & Meyerowitz, (1987). 5µg of genomic lambda DNA restricted with the appropriate enzymes were run in duplicate on a 1% (w/v) TBE agarose gel. The gel was blotted as described

above. The filter was cut in half, one half probed with the male cDNA probe, the other with the female cDNA probe. Filters were hybridised for a minimum of 8hr at 65°C in a shaking water bath. Washing conditions were as described above.

2.9.3 RNA Electrophoresis And Northern Blots.

Total RNA corresponding to 20µg of RNA in a volume of 5µl was loaded onto a 1.5% (w/v) MOPS/Formaldehyde denaturing gel. Before being loaded onto the gel the samples were denatured by the addition of 10µl of formamide, 2µl of 5xMOPS buffer, 3.5µl of formaldehyde, 1µl of EthBr (1mg/ml stock), and heated to 70°C for 5 min, cooled immediately on ice. Prior to loading 2µl of formaldehyde loading buffer was added to each sample (Sambrook *et al.*, 1989). The gels were run in 1xMOPS, with circulation from anode to cathode chambers in order to maintain a constant pH. The gels were photographed as described before. After photography, the gel was soaked in 0.05M NaOH for 20 min followed by soaking for 30 min in 20xSSC. The gels were transferred to Hybond C+. RNA was fixed to the membrane by baking for 2hr at 80°C.

Hybridisation conditions.

DNA/RNA hybridisations were carried out at 42°C in Formaldehyde prehybridisation solution, or in SSC prehybridisation solution at 65°C. Filters were pre-hybridised at 42°C or 65°C for at least 3hr before addition of the denatured radioactive probe and hybridised for a minimum of 16hr. After hybridisation, the filters were washed in 2xSSC at room temperature followed by two 20 min washes in low stringency wash solution at 65°C, followed by one 15 min wash in high stringency wash solution at 65°C. The filters were blotted dry and covered with Saran Wrap before autoradiography.

2.10 Labelling Of Nucleic Acids.

2.10.1 First Strand cDNA Probe.

In order to produce high specific activity 1st strand cDNA probes, 150µCi of 600 Ci/mmole of [α-^{32}P]dCTP was dried down in a siliconised microfuge tube, to this was added 4µl 5x 1st strand buffer, 1µl 80mM NaPPi, 30U RNAGUARD, 2µl Oligo (dT)$_{12-18}$, 1µg Poly A$^+$ RNA, 20U AMV reverse transcriptase and made up to a final volume of 20µl using RF water. The reaction was incubated at 42°C for 30 min, at which time 1µl of 10mM dCTP was added and the incubation continued for a further 30 min. Hydrolysis of the RNA was achieved by the addition of 1 volume of 0.6M NaOH, 20mM EDTA followed by incubation at 65°C for 30 min. Unincorporated nucleotides were separated using Sephadex G-50 columns (Sambrook *et al.*, 1989). Incorporation was assessed in a scintillation counter using Cherenkov counting on a small aliquot of the probe. Specific activities of ~1x10^8 cpm/µg RNA were normally obtained.

2.10.2 Random Priming.

The procedure was essentially as described in Feinberg & Vogelstein (1983). Between 25-50ng of plasmid in 12µl of water was boiled for 10 min and cooled on ice. To this was added 6µl 4x Random priming Buffer, 30µCi of 600 Ci/mmole of [α-^{32}P]dCTP and 2U of Klenow enzyme. The mixture was usually incubated for at least 12hr at room temperature. Probes were purified using Sephadex G-50 columns prepared in 1ml syringes. Incorporation of radioactive precursor was calculated using the scintillation counter and Cherenkov counting. For random primed probes specific activities of 10^8-10^9 cpm/µg were normally obtained.

2.10.3 Nick Translated Biotin Probes.

In order to produce biotin labelled probes for *in situ* hybridisation, 2.5µl of 10x Nick translation buffer, 2.5µl of dNTP mix (0.3mM of each base), 2µl of Bio-16-dUTP, 1µl of ^{32}P (trace), 0.5µg of DNA, 1.5µl of DNase I diluted (1:1000) made up to 25µl with distilled water and 10U of DNA Polymerase I were mixed in an eppendorf. This was incubated at room temperature for 1hr. The probe was precipitated by the addition of 1µl of 0.2M spermidine followed by incubation for 30 min on ice and spun down at 4°C for 30 min. Incorporation was roughly estimated by comparing the resuspended pellet with the supernatant. An incorporation of between 25-30% was found to produce optimal probes. Incorporation of too much biotin is thought to interfere with the hybridisation process due to the bulk of the biotin.

2.11 Screening Lambda cDNA Libraries.

Screening of lambda cDNA libraries was essentially as described in (Sambrook *et al.*,). Briefly, five to ten thousand recombinant phage were plated onto 10 by 10 cm L-agar plates or 100,000 recombinants on a 22.5 by 22.5 cm L-agar plate, using 10^8 host cells and 0.7% (w/v) L-agarose containing 10mM $MgSO_4$. The plates were incubated at 37°C for ~ 8hr or until the phage were just visible. Replica filters were lifted from each plate, denatured for 3 min, neutralised twice for 5 min and washed in 2xSSC for a minimum of 15 min. They were air dried before being baked at 80°C for two hours. The filters were prehybridised in plastic bags or hybridisation tubes containing 10ml of SSC prehybridisation solution for at least 2hr at 65°C. The filters were left to hybridise in hybridisation solution containing the probe for a minimum of 8hr before being washed in 2xSSC at room temperature, two 15 min low stringency washes

followed by two 15 min high stringency washes all at 65°C. They were dried before being wrapped in Saran Wrap and autoradiographed as described previously.

2.12 DNA Sequencing

2.12.1 Double Stranded DNA Sequencing.

Double stranded sequencing reactions using the dideoxy chain termination method (Sanger et al 1977) were carried out as described in the Sequenase manual using α^{35}S dATP as a label. Sequence specific primers were generated using the solid state method on an Applied Biosystems Inc. PCR-MATE 391 DNA Synthesiser, employing phosphoramidite chemistry. The primers were deprotected using ammonium hydroxide and dried down under vacuum. They were resuspended in water and their concentration determined by taking a reading of the optical density at 260nm, and made into a working solution that was 0.5 - 1.0pmol. Table 2.2 lists the primers used in this work.

2.12.2 Polyacrylamide gels.

Samples from sequencing reactions were run on 6% (w/v) acrylamide gels. Gels were left to polymerise overnight before use and were pre-run at 100W for 1hr in 1xTBE after which time, the gel temperature was normally 55°C. Before loading onto the gel, the samples were denatured for 2 min at 75°C and then cooled on ice. Gels were run for 2hr for a short run (150 bp) or for 5hr for a long run (300 bp) on constant power, and then dried for 1-2hr at 80°C onto Whatman 3MM paper under vacuum. Autoradiography was carried out without intensifying screens at room temperature.

Oligo Name	Sequence	Lengt
K31/T3/1	5' GGC CAA GAA TCC GTT CAG TGA TC 3'	23
K31/T3/2	5' CGC GAT GGC ATG TAC GG 3'	17
K31/T7/1	5' ATT TGC ATC GCG TTG TG 3'	17
K31/T7/2	5' CTG GTG CGA TTT GCA TC 3'	17
K31/T7/3	5' CGT CGA GTC GGA CAA GG 3'	17
K27/T7/1	5' TGC CTT CAA GGA GCA CC 3'	17
K27/T7/2	5' GTT GCG AGA TGT TGT TC 3'	17
K27/T3/1	5' GCA CTT GGC CTT TGA GAA GCA CG 3'	23
K27.1/T3/3	5' GTC AAT GCC TAC TGA ATA G 3'	19
K27.1/T7/3	5' TGC TTG GCT GTA CTC TCC A 3'	19
K27.2/T7/4	5' TGG CCC AAA GTT CCA TG 3'	17
K27.2/T3/4	5' ATT ATA TCA GCT GCA GGG 3'	18

Table 2.2 List of oligonucleotides, used in the sequencing experiments described in Chapter 6.

2.13 *In Situ* Hybridisations.

2.13.1 *In Situ* Hybridisation To Polytene Chromosomes.

The procedure for in situ hybridisation to polytene chromosomes was essentially as described by Pardue (1986). The larvae were grown on rich larval media prepared as described in Section 2.4. Probes used were subcloned fragments from EMBL3 genomic phage labelled with Bio-16-dUTP as described in Section 2.9.3. Hybridisation to the chromosome was detected using a Strepavidin-Biotinylated Horseradish Peroxidase complex and DAB. After hybridisation, the slides were stained with Giemsa and mounted using DPX. Slides were photographed using Kodak Ectachrome 160 tungsten film on a Leitz Vario-orthomat camera system.

2.13.2 *In Situ* Hybridisation To Whole Mount Testis.

The whole mount testis *in situ* protocol was an adaptation by Luke Alphey (pers. comm.) of the whole mount embryo protocol (Tautz, 1983). Testes from adults or third instar larvae were crudely dissected in Ringers solution and fixed in 4% (w/v) paraformaldehyde/1x PBS for 30 min at room temperature. This was followed by rinsing five times for 2 min each with PBT. After incubation for 5 min at room temperature in 100mg/ml Proteinase K, they were washed with PBT containing 2mg/ml glycine twice for 2 min each. Fixing and washing were as before. The testis were prehybridised in 500µl of biotin prehybridisation solution at 55°C for at least one hour. 25µl of boiled probe was added to the prehybridisation solution. The probes used were subcloned fragments from EMBL3 genomic phage or cDNAs corresponding to these genomic fragments labelled with biotin as described in Section 2.9.3. These were left at 55°C for a minimum of 8hr. After removal of the probe, the testis were washed with

hybridisation solution, followed by PBS for 2 min, PBT for 2 min and PBS for 2 min. In order to reduce the background signal, they were incubated with Avidin/Biotin blocking kit for at least 30 min. Slides were mounted in Vectashield mountant. The signal was detected using a Molecular Dynamics confocal microscope using molecular dynamics 'Image Space' software. The fluorescent antibodies were excited using a Krypton/Argon laser. The Texas Red was excited at 568nm and detected at 660nm, while FITC was excited at 488nm and detected at 560nm.

2.13.3 Detection Of *Lac*Z Staining.

Protocol devised by Douglas Armstrong (pers. comm.). Testis were dissected and fixed as described above. They were then washed in PBS, followed by PBT, and then PBS for 2 min each. After washing, the testis were rinsed in PAT and then incubated in PAT containing 3% (v/v) normal goat serum and 1:2000 dilution of the primary rabbit anti-*LacZ* antibody. These were incubated at room temperature overnight. The testis were then washed as before, followed by detection in PAT containing a 1:250 dilution of the secondary goat anti-rabbit-Texas Red labelled antibody. Again this was incubated at room temperature. Mounting and detection of signal was as described above.

Chapter Three Initial Characterisation of Eight Genomic DNA Clones

Introduction.

In *Drosophila*, several strategies may be used to identify genes involved in different developmental pathways. The classical approach has been to utilise mutagenic screens to identify genes whose mutant phenotype results in perturbation of the developmental pathway being studied. While this approach has been used successfully to identify genes involved in many developmental pathways such as body pattern specification (Nusslein-Vollard & Wieschaus, 1980) and genes controlling behaviour (Horyk & Sheppard, 1977), it has not proved particularly successful in the isolation of genes that are involved strictly in spermatogenesis. During a classical mutagenesis of the fly, mutants can easily be isolated that display a male and/or female sterile phenotype. As discussed in chapter 1 however, not all male-sterile mutations represent loci that function solely in the spermatogenic pathway. Male-sterile phenotypes could arise as a consequence of perturbations in normal cellular processes (Lindsley & Zimm, 1992), as a consequence of behavioural mutations such as *fruitless* (*fru*) (Gailey & Hall, 1989; Castrillon *et al.*, 1993) or be due to gross anatomical defects in the reproductive apparatus as in the mutants *ken and barbie* (*ken*) or *twig* (Castrillon *et al.*, 1993). As discussed in chapter 1, classical estimates of the number of genes that can be mutated to give a male-sterile phenotype range from 500 to more than 1750. Therefore, a large number of male-sterile mutants would have to be screened in order to identify true spermatogenic mutants (Lindsley & Lifschytz, 1972; Lifschytz, 1987). While there have been problems identifying 'true' spermatogenic mutants, this technique has not been totally unsuccessful. A structural gene for α-tubulin (α3t) was isolated in a classical mutagenesis using Ethyl Methane Sulphate (EMS) as a mutagen (Fuller 1986). When homozygous, this mutation causes a male-sterile phenotype. Additionally, all of the Fertility Factors (discussed in chapter 1) were identified *via* classical screening approaches, with X-rays and EMS as the mutagens (Brosseau, 1960; Kennison, 1983).

An alternative to the classical genetic method has been to use 'Reverse Genetic' approaches such as differential screening or enhancer-trapping to isolate genes in a particular developmental pathway. Enhancer-trapping has been used to attempt to identify transcription units which are expressed in a particular tissue of interest, in this case, identifying genes that are expressed in the testis. Such an approach allows the identification of testis-specific transcription units without the observation of a male-sterile phenotype. This approach has been facilitated by the development of 'enhancer-trapping' (O'Kane, 1987). These are P-elements which have been modified to contain the *E. coli* β-galactosidase gene (*lacZ*). These elements do not code for a functional transposase, but are able to jump if supplied with transposase *in trans*. The strain Δ2-3(99B) (Robertson *et al.*, 1988) for example, contains a modified P-element that is able to supply large amounts of transposase, but is unable to jump itself.

Enhancer-trap elements contains the *lacZ* gene under the control of a P element promoter which drives a low level of *lacZ* transcription. If upon transposition, the P-element inserts near a tissue specific enhancer, elevated expression of the *lacZ* construct within the tissue of interest can easily be detected using the chromogenic substrate X-gal. This allows selection of those lines which exhibit staining patterns in the tissue of interest. Since the new generation of enhancer-trap elements contain a plasmid origin of replication and an antibiotic resistance marker, the flanking DNA 5' and 3' of the insertion site can be isolated by cutting with a variety of restriction enzymes, ligating the genomic DNA and by transforming *E. coli*, a process known as plasmid rescue. (Perucho *et al.*, 1980, Steller & Pirrotta, 1986, Hamilton *et al.*, 1991) Several groups have used these enhancer-trap elements to investigate embryonic neural development (Bellen, 1989; Ethan, 1989). This strategy has enabled the selection of insertions by virtue of the staining patterns in the tissues of interest. Recently, this approach has been utilised by several groups to identify enhancers that are expressed

specifically in the adult testis (Bownes, 1990; Gonczy *et al.*, 1992; Castrillon *et al.*, 1993). The enhancer-trap screen by Castrillon *et al.* (1993), which is described in more detail in chapter 1, was extremely successful in isolating spermatogenic mutations. In the course of this screen, 89 recessive autosomal mutations were identified of which 58 (including previously identified loci) appeared to represent genes involved strictly in spermatogenesis.

Whilst this approach allows for the selection of lines in which enhancer elements direct tissue-specific expression, there are still some problems associated with this technique. The P-element may insert near an enhancer element which is located at some distance from the gene whose expression it drives. Therefore, although it may be possible to plasmid rescue the flanking DNA sequences, a considerable amount of work may be required to allow the identification of the gene regulated by that enhancer. While plasmid rescue has been used to isolate flanking sequences for some insertion lines, others do not contain the appropriate restriction sites in the genomic region close to the P-element insertion and hence the flanking genomic DNA is unable to be cloned except by the construction of a 'mini-library'. The insertion site may also be close to more than one enhancer element and give a false pattern of expression. Consequently, the gene that is isolated may have a pattern of expression that bears little relation to that seen in the enhancer-trap line.

Differential (+/-) hybridisation screening of genomic or cDNA libraries has been used successfully in various systems (St John & Davis, 1979; Taniguchi *et al.*, 1980) to identify tissue or stage-specific transcription units including those involved in spermatogenesis (Schafer, 1986; DiBenedetto, 1987). This technique relies on generating two different mRNA populations, one of which is enriched in the sequences of interest, the other, in which these transcripts are reduced or lacking. A differential screening approach can be used to identify genes whose

transcription is tissue-specific or is highly regulated in a spatial or temporal fashion. It does not, however, facilitate the detection of transcripts which show differential splicing. For example the sex determining gene *doublesex* would not have been isolated as it is expressed both in males and females. In common with most other molecular techniques, this approach does not give any insight into the actual function of the identified transcription unit. Some idea of function may be obtained by sequencing the genomic DNA contained within the clone; if homology is obtained to another group of cloned sequences, then some idea of function may be inferred. Additionally, the problem of assigning function can be resolved by using a differential screening approach in conjunction with classical techniques or other 'Reverse Genetic' approaches. In *Drosophila*, techniques such as site-selected mutagenesis (Kaiser & Goodwin, 1990; Milligan & Kaiser, 1993) using P or I transposable elements, or 'local jump' experiments using marked elements located near the gene of interest (Littleton, 1993) have been of great use in isolating mutations in genes cloned in this way.

Several factors must be considered when planning a differential screen: the availability of two mRNA populations, the abundance of the specific transcript and the type of library to be screened. The primary requirement for two different populations that can be used to isolate mRNA for cDNA probes is clearly not a problem when attempting to isolate sex-specific transcripts. However, the stage at which the mRNA is isolated has to be considered. While spermatogenic transcripts have been identified by other laboratories, (Schafer, 1986; DiBenedetto, 1987) neither group has identified transcripts expressed in the early embryonic stages. Since spermatogenesis has already been initiated in 1st instar larvae and many of the earliest events in sex determination (such as determination of germline sex) occur early in embryogenesis, Russell and Kaiser (1989) decided to screen cDNA and genomic libraries with 1st strand cDNA

probes made from mRNA derived from third instar larvae. This is the earliest stage at which sexual dimorphism becomes visible (Sonnenblick, 1950).

The relative abundance of the transcripts that are to be identified must also be considered. Transcripts which comprise greater than 1% down to 0.01% of the total mRNA population can be isolated using a differential screening approach, although the sensitivity of the screen may be enhanced by the use of subtracted probes (Timberlake, 1980; Zimmerman *et al.*, 1980; Sargent, 1987). Subtractive screening allows the detection of less abundant transcripts (less than 0.01%) and also reduces the problem of elimination of clones that contain a non-sex-specific gene in addition to a sex-specific transcript. On the assumption that some of these spermatogenic transcripts are likely to comprise 1% of the total mRNA population, and since several spermatogenic loci had already been isolated *via* a differential screening approach, (Schafer, 1986; DiBenedetto, 1987), the additional sensitivity obtained by using a subtracted probe was thought unnecessary by Russell (1989). The abundance of the transcripts not only affects the type of probe to be used (i.e. subtracted versus non-subtracted) but also the type of library to be screened (i.e. genomic library versus cDNA library). In a genomic library, those sequences that code for abundant transcripts are likely to be present at the same frequency as those encoding low abundance transcripts. Clearly, this is not the case for a cDNA library, where the frequency of a clone in the library is dependent upon its abundance in the total mRNA population. Another consideration that has to be taken into account when choosing between screening a cDNA or genomic library is the number of clones to be screened in order to have a 99% chance of identifying any desired sequence. This number is determined by insert size in the case of a genomic library, and by the abundance of the transcript in the case of a cDNA library. For a genomic library, using the equation given in Kaiser & Murray, (1985) the parameters which have to be taken into account are the size of the inserts in the library and the size of

the organism's genome. Given that the insert range of EMBL3 is 9-22 kb, 30,000 primary recombinants must be screened in order to have a 99% probability of isolating any individual sequence, assuming randomness. As this figure is dependent on genome size, an organism whose genome is significantly larger than *Drosophila* would require more clones to be screened. For screening a cDNA library, the equation given in Sargent (1987) suggests that for a transcript whose abundance is no less than 0.01% of the total mRNA population, (the limit that can be detected using non-subtracted probes) 50,000 clones would have to be screened to ensure a 99% probability of detecting the transcript of interest. For an organism such as *Drosophila*, the number of clones to be screened to identify a genomic clone (30,000) or a cDNA clone (50,000) are not sufficiently different to bias the selection of which one to screen.

Given that the number of clones that should be screened is not vastly different between the two types of libraries, the choice of which to screen might be determined by the type of insert contained in each library. Clones derived from a genomic library may contain sequences that are from more than one gene, consequently additional analysis would be required to identify which region of the clone contains the sequences of interest. In addition, as described previously, those clones containing sex-specific and non-sex-specific sequences on the same insert may not be detected unless a subtracted probe is used. The isolation of a genomic DNA clone facilitates studies on the genomic organisation and regulation of the sequences. Clones which have a spermatogenic function are transcribed early during the primary spermatocyte stage (as described in chapter 1), therefore some of these may be subject to translational control (Schafer, 1993; Kempe, 1994). Analysis of the sequence of these genomic DNA clones may facilitate the identification and characterisation of these sequences. In contrast, those derived from a cDNA library should represent a single gene or gene family. Sequences which represent highly labile RNA species may not

be isolated using a cDNA library, as their turnover might be too rapid. In addition, transcripts which do not have poly A tails will not be isolated from a cDNA library which was not generated by random priming. This problem will also be encountered when screening a genomic library with cDNA which has been primed using oligo dT.

In summary, differential screening may be performed by hybridising labelled cDNA (sex-specific or subtracted) to either a cDNA or genomic library. Using the considerations described earlier, Russell & Kaiser decided to screen both cDNA and genomic DNA libraries in an attempt to optimise identification of male-specific genes. However, due to screening problems with the cDNA libraries, male-specific clones were only identified from the genomic library screen (Russell, 1989).

Selection Of Male - Specific Genomic DNA Clones.

A *Drosophila* genomic DNA library (made by K. Kaiser, unpublished) was screened by Russell & Kaiser (Russell, 1989) for differential hybridisation to cDNAs prepared from male and female third instar larvae. The library was constructed in the lambda vector EMBL3 and contains 9-22 kb inserts derived from Oregon-R DNA partially digested with *Sau*3A. 30,000 primary recombinants were screened in duplicate at a density of 8,000 plaques per plate with labelled cDNA prepared from either male or female third instar larval poly(A)+ RNA. At the end of the screen, 20 clones were designated as putatively male-specific/differential. Dot blots of these 20 clones were screened by Russell (1989) using male and female, adult and larval cDNA probes to redefine their classification. These clones can be further classified into 3 groups based on their hybridisation patterns to adult and larval cDNA probes (Table 3.1).

Clone Classification	Number of Clones
male larval/adult specific	9
male larval/adult elevated	5
non-sex-specific	6

Table 3.1: Classification of putative male-specific/differential genomic clones

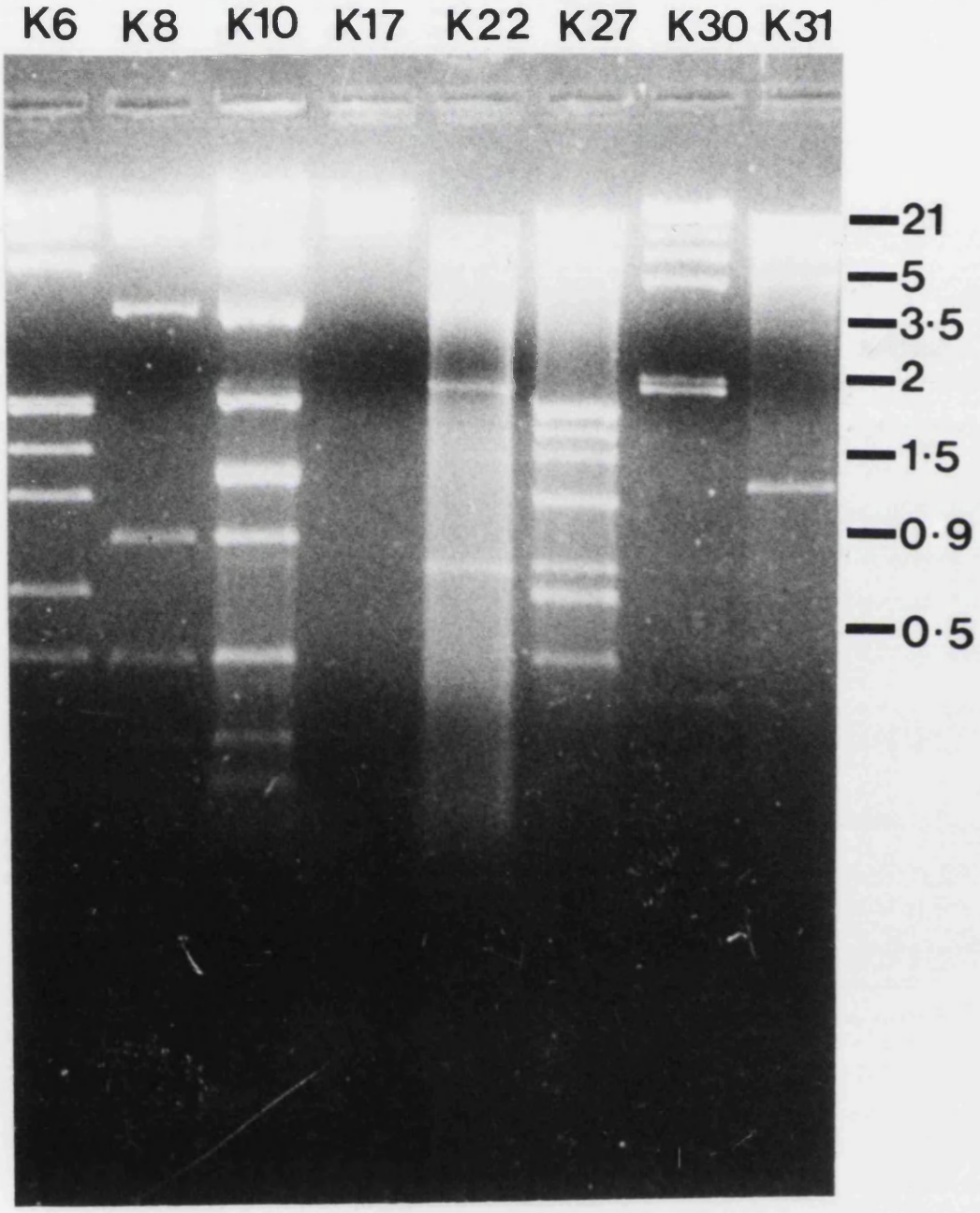
Results.

3.1 Preliminary Examination Of Eight Putative Male-Specific Clones:

Eight of the putative male-specific genomic DNA clones (gK6, gK8, gK10, gK17, gK22, gK27, gK30 and gK31) isolated by Russell were made available to me for further characterisation. Russell classified 5 as larval/adult male-specific (gK6, gK8, gK10, gK17 and gK22) and the remaining 3 (gK27, gK30 and gK31) as showing elevated expression in the male. DNA was prepared from all 8 genomic DNA clones. In order to estimate the insert size contained in each clone and to perform an initial comparison of the clones with each other, the clones were restricted with the enzymes *Sal*I and *Eco*RI (Figure 3.1). The *Sal*I sites are situated in the lambda polylinker and allow the inserts to be released from the lambda arms. The *Eco*RI sites in the vector arms have been removed by the construction of the library consequently, it is used in this case to cleave any internal sites in the insert (the insert may also contain sites for *Sal*I). Although there are some problems with this gel (for example the lack of vector arms in some tracks) the purpose of this gel was to quickly identify any clones which might be related. For this purpose, the gel was of sufficient quality since it is evident that several of the genomic DNA clones have similar restriction patterns. The lack of vector arms in some of the tracks might have been due to insufficient heating of the samples prior to loading. Consequently, the *cos* sites in the bacteriophage lambda arms may not have melted. This would result in these bands migrating aberrantly. In addition, the lack of vector arms might result from the incomplete restriction of the DNA. The results of these digests are summarised in Table 3.2.

From this preliminary analysis, it was evident that several of the genomic DNA clones had restriction fragments in common, therefore some of the genomic

Figure 3.1: EthBr stained 1% agarose gel containing *Eco*RI, *Sal*I double digests of male-specific clones selected for further analysis.


K6
K8
K10
K17
K22
K27
K30
K31
21
5
3·5
2
1·5
0·9
0·5

Clone Number	Size of *Eco*RI, *Sal*I fragments (kb)
K6	6.4, 1.8* 1.4, 1.15, 0.7, 0.45
K8	12, 2.8, 0.8, 0.45
K10	7.3, 2.8, 1.7, 1.25, 1.2, 0.8, 0.45, 0.25, 0.2
K17	No sites for these enzymes
K22	8.9, 3.3, 1.9, 0.8
K27	5.3, 1.8*, 1.6, 1.4, 1.15, 0.8, 0.7, 0.45
K30	5.8, 4.2, 2.0, 1.9, 0.4
K31	12.3, 4.2, 1.2

Table 3.2 shows the fragments produced by a *Sal*I, *Eco*RI double digest of all eight genomic clones. Fragment * is a doublet.

DNA clones may be derived from the same locus. gK6 and gK27 DNA share five bands of molecular weight 1.8, 1.4, 1.15, 0.7 and 0.45 kb, while genomic DNA clones gK8 and gK10 share three bands of 2.8, 0.8 and 0.45 kb (Figure 3.1). In order to examine further the relationship between genomic DNA clones gK6 and gK27, and genomic DNA clones gK8 and gK10, the genomic DNA clones were restricted with *Sal*I in combination with *Eco*RI, *Bam*HI and *Hin*dIII. Duplicate blots were prepared and hybridised to the two relevant genomic DNA clone (Figures 3.2 & 3.3) i.e. blots of gK6 and gK27 were probed with both gK6 and gK27, blots of gK8 and gK10 were probed with both gK8 and gK10.

gK6 and gK27.

The two genomic DNA clones gK6 & gK27 were digested with *Sal*I in combination with *Eco*RI, *Hin*dIII and *Bam*HI. The resulting EthBr stained gel is shown in Figure 3.2A. Duplicate Southern blots were probed with either gK6 (Figure 3.2B) or gK27 (Figure 3.2C). These results are summarised in Table 3.3. Genomic DNA clone gK27 was chosen for further characterisation as it is the larger of the two clones.

gK8 and gK10.

The same procedure was used to establish whether there was a relationship between genomic DNA clones gK10 and gK8 (see Figure 3.3A). Duplicate Southern blots were hybridised with either gK8 (Figure 3.3B) or gK10 (Figure 3.3C). These results are summarised in Table 3.4. As before, there appears to be considerable overlap between these two genomic DNA clones. Clone gK10 was selected for further analysis.

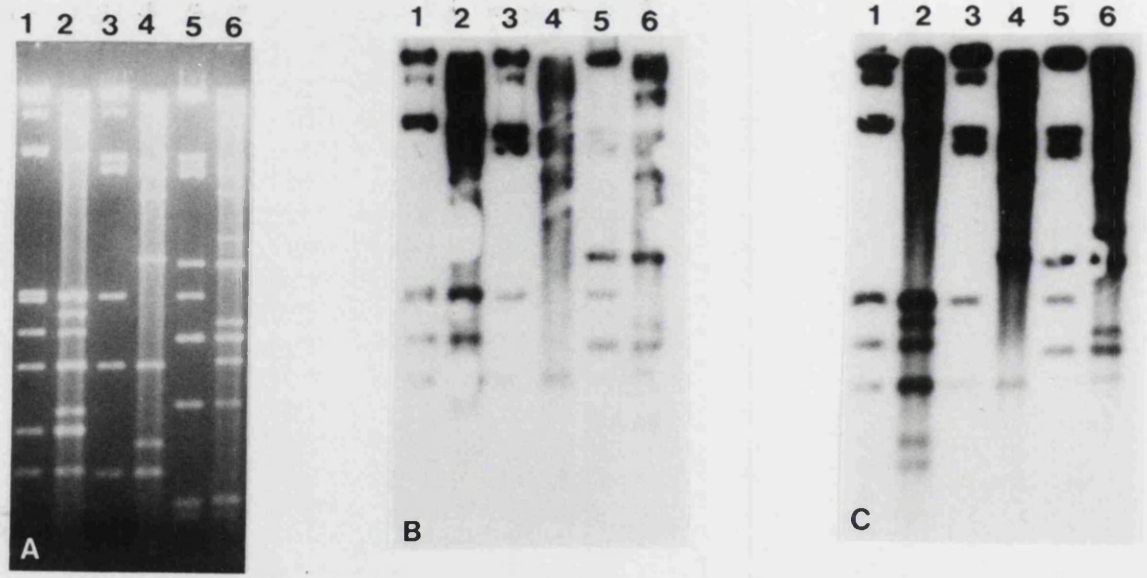


Figure 3.2a: EthBR stained 1% agarose gel. Lanes 1 & 2 are *Sal*I, *Eco*RI double digests of gK6 (lane 1) and gK27 (lane 2), lanes 3 & 4 are *Sal*I, *Bam*HI double digests of gK6 (lane 3) and gK27 (lane 4), lanes 5 & 6 are *Sal*I, *Hin*dIII double digests of gK6 (lane 5) and gK27(lane 6).

Figure 3.2b: Autoradiograph of Southern blot of gel in **Figure 3.2a** hybridised with gK6. The blot was washed to high stringency and exposed overnight.

Figure 3.2c: Autoradiograph of Southern blot of gel in **Figure 3.2a** hybridised with gK27. The blot was washed to high stringency and exposed overnight.

	K6	K27
SalI,EcoRI	6.4, **1.8, 1.4, 1.15, 0.7, 0.45,**	5.3, **1.8, 1.4, 1.15, 0.7, 0.45,**
SalI,BamHI	4.7, 3.9, 1.8, **1.08,**	4.2, 3.7, 3.1, 2.1, **1.08**, 0.3
SalI,HindIII	**2.1**, 1.8, **1.35, 0.9, 0.3, 0.25**	7, 3.8, 3.3, **2.1, 1.35, 0.9, 0.3, 0.25**

Table 3.3 shows hybridising fragments in gK27 and gK6. Common bands are highlighted in bold text. Unique bands, which hybridise to the reciprocal probe are indicated by plain text.

	K8	K10
SalI,EcoRI	12, **2.8, 0.8, 0.45**	7.3,**2.8, 0.8, 0.45**
SalI,BamHI	0.45	~12
SalI,HindIII	**All Bands**	**All Bands**

Table 3.4 shows hybridising fragments in gK8 and gK10. Common bands are highlighted in bold text. Unique bands, which hybridise to the reciprocal probe are indicated by plain text.

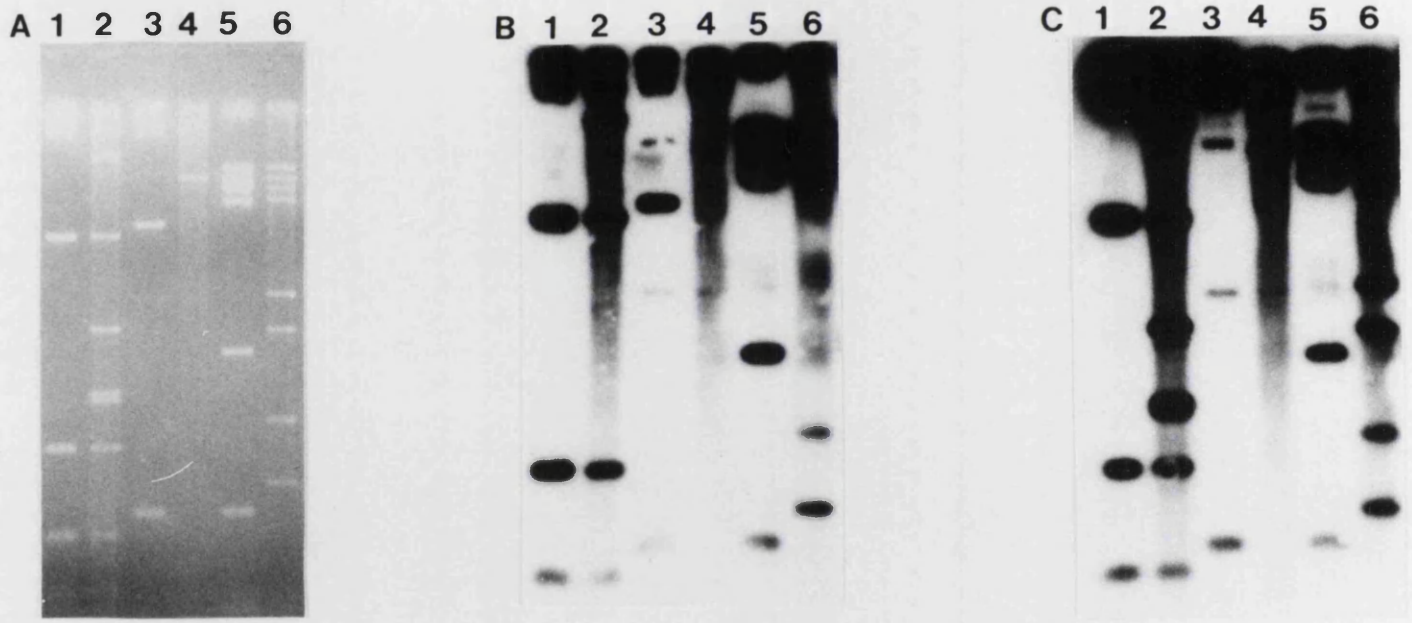


Figure 3.3a: EthBR stained 1% agarose gel. Lanes 1 & 2 are *Sal*I, *Eco*RI double digests of gK8 (lane 1) and gK10 (lane 2), lanes 3 & 4 are *Sal*I, *Bam*HI double digests of gK8 (lane 3) and gK10 (lane 4), lanes 5 & 6 are *Sal*I, *Hin*dIII double digests of gK8 (lane 5) and gK10 (lane 6).

Figure 3.3b: Autoradiograph of Southern blot of gel in **Figure 3.3a** hybridised with gK8. The blot was washed to high stringency and exposed overnight.

Figure 3.3c: Autoradiograph of Southern blot of gel in **Figure 3.3a** hybridised with gK10. The blot was washed to high stringency and exposed overnight.

gK17 and gS4.

Clone gK17 is unusual in that it contains no internal recognition sites for the enzymes *Eco*RI, *Bam*HI, *Sal*I and *Hin*dIII (Figure 3.1). In this respect, it is similar to clone gS4 previously characterised by Russell (1989). Digestion of gK17 and gS4 with either *Xho*I or *Sma*I releases a band of 1.4 kb, the intensity of which suggests that it has an increased molar ratio with respect to the other fragments released by the digest. In addition, a 1.8 kb fragment is released. Since the lambda vector has a cloning capacity for inserts in the range 9-22 kb, it seems likely that these digests release multiple bands of 1.4 kb. As the insert size is approximately 14 kb, it suggests that a tandem sequence of ~8 copies resides in this clone.

All the bands from *Xho*I and *Sma*I digests of both gS4 and gK17 genomic DNA clones hybridise when probed with gK17 and *vice versa*. (Figure 3.4 A,B& C). As described in chapter 1, the genomic DNA clone gS4 (now called Mst40) has been characterised by Russell (Russell & Kaiser, 1993). Since gK17 appeared to be related to this clone it was not characterised further.

On the basis of this preliminary analysis, clones gK6, gK8 and gK17 were not characterised further. This left five clones: gK10, gK22, gK27, gK30 and gK31 for further analysis.

Russell (1989) classified clone gK30 as male larval/adult elevated. 'Reverse northern' analysis shows that this clone contains no male-specific fragments, but two male-elevated fragments (Figure 3.5b). This contrasts with the other clones which contain at least one male-specific fragment. For this reason, this clone was not characterised further. Consequently the number of clones to be studied was reduced to four: gK10, gK22, gK27 and gK31.

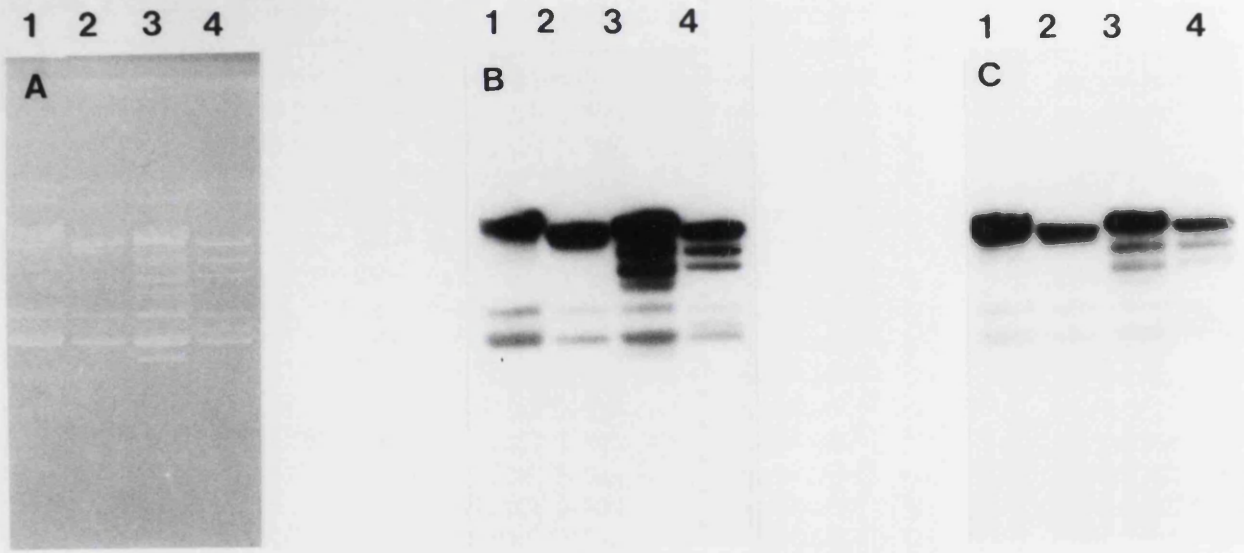
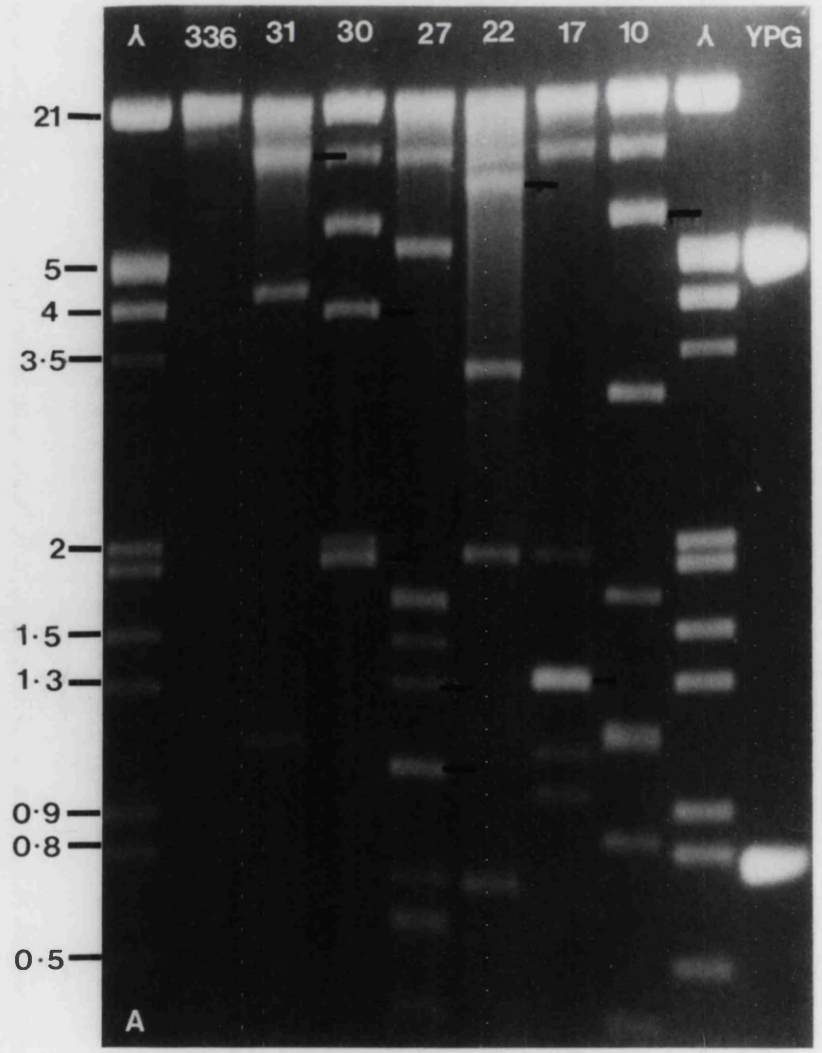
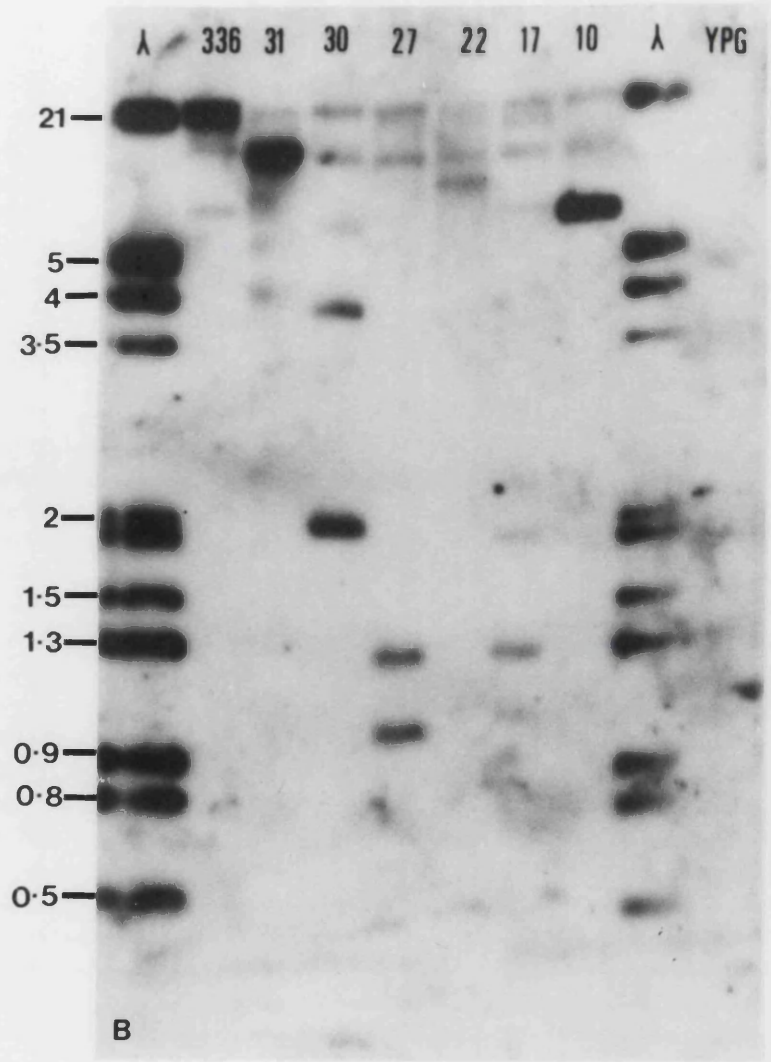


Figure 3.4a: EthBR stained 1% agarose gel. Lanes 1 & 2 are *Sal*I, *Xho*I double digests of gs4 (lane 1) and gK17 (lane 2), lanes 3 & 4 are *Sal*I, *Sma*I double digests of gS4(lane 3) and gK17 (lane 4),

Figure 3.4b Autoradiograph of Southern blot of gel in **Figure 3.4a** hybridised with gS4. The blot was washed to high stringency and exposed overnight.

Figure 3.4c: Autoradiograph of Southern blot of gel in **Figure 3.4a** hybridised with gK17. The blot was washed to high stringency and exposed overnight.


λ
336
31
30
27
22
17
10
λ
YPG
21
5
4
3·5
2
1·5
1·3
0·9
0·8
0·5
A


λ 336 31 30 27 22 17 10 λ YPG
21
5
4
3·5
2
1·5
1·3
0·9
0·8
0·5
B

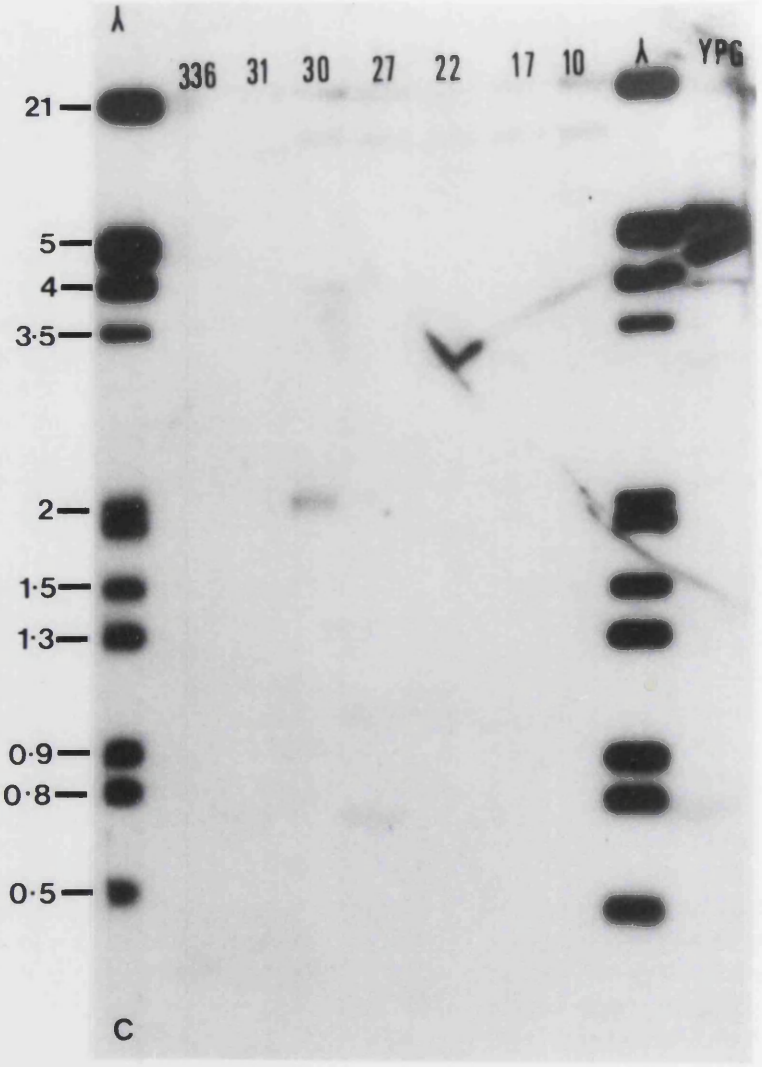

λ
336 31 30 27 22 17 10 λ YPG
21
5
4
3·5
2
1·5
1·3
0·9
0·8
0·5
C

Figure 3.5

Figure 3.5A: EthBr stained 1% agarose gel containing *Eco*RI/*Sal*I double digests of male-specific clones selected for further analysis (male-specific clone gK17 digested with *Sal* I/*Xho*I).

Figure 3.5B: Autoradiograph of a Southern blot of gel shown in **Figure 3.5A** hybridised with cDNA derived from adult males. Exposed for one week.

Figure 3.5C: Autoradiograph of Southern blot of duplicate gel shown in **Figure 3.5A** hybridised with cDNA derived from adult females. Exposed for one week.

3.2 Identification Of Male-Specific Fragments From Genomic DNA Clones.

To determine the male-specific region of each clone 'Reverse northern' analysis was performed (Fryxell & Meyerowitz, 1987). The lambda clones were double digested with *Sal*I and *Eco*RI (the exception was gK17 which was digested with *Sal*I and *Xho*I). The digested phage DNA were run in duplicate on a 1% agarose gel (See Figure 3.5A) and transferred to nylon filters. The filters were probed in parallel with cDNA probes prepared from male and female mRNA. The autoradiographs were compared to determine which fragments exhibited a male-specific pattern of hybridisation (Figure 3.5B and C). In order to confirm the sex-specificity of the probes, two controls were included on each filter. The female-specific control was a 0.8 kb *Bam*HI fragment of the yolk protein 1 gene in pGEM, designated YP1. This gene is expressed in the ovaries and fat bodies of females in response to the genes of the somatic sex determination hierarchy (Hung & Wensink, 1981). The male-specific control was the genomic DNA clone Mst(3)336 (DiBenedetto *et al.*, 1987). This was isolated from a screen for male-specific transcripts involved in spermatogenesis and is expressed in both male pupae and adults, but is not present in females. The results of the 'Reverse northern' analysis are summarised in Table 3.5. For each genomic DNA clone, at least one region hybridises strongly to the male cDNA probe (Figure 3.5B). These bands show little, or no hybridisation to the female-specific probe.

Table 3.5 shows the fragments produced by an *Eco*RI/*Sal*I digest of the genomic clones. Fragments which hybridise to the male and female cDNA probes are listed. Fragment * is a doublet.

Clone	*SalI/Eco*RI fragments (kb)	Male Specific Fragments	Female Specific Fragments
gK10	7.3,2.8, 1.7, 1.25, 1.2, 0.8, 0.45, 0.25	7.3	None
gK22	8.9, 3.3, 1.9, 0.8	8.9	None
gK27	5.3, 1.8*, 1.6, 1.4, 1.15, 0.8, 0.7, 0.45	1.4, 1.15	0.8
gK31	12.3, 4.2, 1.2	12.3	None

Discussion.

Russell's screening of a *Drosophila* genomic library was successful in identifying male-specific transcribed regions. Several of the clones isolated in this screen appear to have identified the same male-specific loci. Additionally, some of the clones characterised by Russell represented previously isolated male-specific genes identified by other laboratories, also using a differential screening approach (DiBenedetto *et al.*, 1987; Russell, 1989). This strongly supports the hypothesis that these clones represent genuine male-specific transcripts but also confirms that at this level of sensitivity, only a relatively small number of genes can be identified.

The use of third instar larval mRNA to prepare the probes has possibly biased the selection of clones towards those expressed in the gonad since this is one of the few sexually-dimorphic structures that can be observed in the larvae (Sonnenblick, 1950). The testis of the third instar larva is much larger than the female ovary (8-12 germ cells in the ovary compared with 36-38 in the testis) (Bodenstein, 1950) and consequently is likely to comprise approximately 1% of the total third instar larval mRNA. In addition, at the third instar larval stage the germ cells in the testis have begun to form primary spermatocytes (as discussed in chapter 1). This stage is characterised by a rapid increase in growth, resulting in a 25-30 fold increase in volume over a ~90 hour period (Lindsley & Lifschytz, 1972; Lindsley & Tokuyasu, 1980). Autoradiographic studies also suggest that most of the transcription of gene products required for later stages of spermatogenesis occur at this time (Olivieri & Olivieri, 1965; Gould-Somero & Holland, 1974). Characterisation of other genomic DNA clones from this screen have similarly been shown to have the expression of their transcripts restricted to the germline. (Russell & Kaiser, 1992; 1993).

The clones isolated in Russell's screen seem to represent moderately abundant transcripts. This abundance may give some insight into the possible function of these transcripts. Transcripts that have a high expression level are more likely to have a structural role, such as formation of sperm tail proteins (Schafer, 1993), than to perform a regulatory function. At the level of sensitivity used by differential screening, the number of genes that can be detected is close to saturation.

Whilst Russell's screen was successful, there are ways in which it could have been refined. As discussed previously, the use of subtracted probes would have allowed a more sensitive screening strategy. Due to problems with non-specific hybridisation, only the genomic library was screened successfully. It is therefore possible that there were male-specific clones which were not identified due to the close proximity of a non-sex-specific transcript on the same region of DNA. Additional levels of screening could have been achieved using cDNA probes derived from mutant fly strains such as *tudor* (*tud*) (Boswell & Mahowald, 1985) which lack a functional germline, or *transformer* (*tra*) (Baker & Ridge, 1980) where the female somatic tissue is transformed into male somatic tissue. This type of analysis would allow the further classification of the clones into those which are specific for the soma and those which are specific to the germline. The use of cDNA probes derived from different stages of the *Drosophila* life cycle could also have been used, again refining classification. However, a similar classification could be obtained by the use of RNA from the mutant flies and stages on northern blots.

The use of subtracted probes would also have reduced the possibility of re-isolating clones by identifying a class of much lower abundance transcripts than those identified by Schafer, (1986) and DiBenedetto, (1987).

An alternative approach could be to screen subtracted cDNA libraries (Palazzolo *et al.*, 1989). It was estimated that for an organism whose genome is the size of *Drosophila*, ~50,000 cDNA clones would have to be screened in order to identify sequences present at 0.01% of the total mRNA population. Whilst this number is not prohibitive, the use of a subtracted library would significantly reduce the number of clones that would have to be screened. Consequently, for an organism with a larger genome size, subtracted libraries would facilitate a differential screening procedure without a prohibitive number of plaques having to be screened. However, the use of subtracted libraries also has its problems (for example, the clones isolated in Palazzolo's screen all contained very small inserts). Therefore, in order to obtain the whole of the coding sequence, the positives from this screen were used as probes to rescreen other *Drosophila* cDNA libraries.

The initial characterisation of the genomic DNA clones provided a quick means of reducing the number of clones to be characterised. As can be seen from Figure 3.4, a gel which was possibly inadequate for publication could still provide sufficient information to allow elimination of the genomic DNA clone under study. gK17 had several features in common with gS4 (eg lack of restriction sites for common enzymes) and cross-hybridisation was merely used to confirm that the two genomic DNA clones were related. Consequently, the fact that insufficient DNA was loaded and that the fragments were not well separated did not call into question the results of the hybridisation. In this chapter, I attempted to ensure that all the genomic clones to be analysed were unique and contained male specific transcription units. Using the techniques described in this chapter, the number of clones to be characterised was reduced from eight to four. In addition, using 'Reverse northern' analysis, I have identified four unique genomic regions containing male-specific transcripts. In this respect the aims of this section of the work were realised. Their further characterisation will be discussed in detail in the subsequent chapters.

Chapter Four Genomic Structure and Organisation

Introduction.

The analysis of the organisation of the genomic DNA clones selected for characterisation was of interest for several reasons. It was necessary to determine whether these represent unique sequences in the genome or members of multi-gene families. As has been discussed in chapter 1, some of the genes that have been identified as putative 'spermatogenic' loci belong to multi-gene families. An example is provided by the Mst(3)CPG gene family first identified by Schafer (1986). This family is a small cluster of seven genes that appear to encode proteins that form the secondary fibres in the mature sperm tail. In addition, seven dynein genes have also been identified in *Drosophila* (Rasmusson *et al.*, 1994) of which one has been localised to the Y-chromosome (Gepner & Hays, 1993). Testis-specific isoforms of other genes have been identified, such as the testis-specific β-tubulin (*β2t*) or the cdc25 homologues *string* and *twine* (Alphey *et al.*, 1992; Courtot *et al.*, 1992). In addition, testis-specific gene families have also been identified in other *Drosophila* species such as *D. hydei* (Neesen *et al.*, 1994). It was therefore of interest to determine whether any of the male-specific fragments identified in chapter 3 represented single copy genes or members of a known or unknown gene family.

Since some 'spermatogenic' loci such as *stellate* consist of a tandemly-repeated sequence and two classes of genomic clones which had a repetitive structure had been isolated more than once during the screen (gS4/gK17, described in Russell and Kaiser 1994 and chapter 3, gS1/gS8/gS9, described in chapter 1 and Russell & Kaiser, 1993) it was of interest to determine whether any of my genomic DNA clones contained repeated elements.

In order to further characterise the cloned sequences, the cytological locations of each were determined. Since it was evident that the Y-chromosome has such a pivotal role in the process of spermiogenesis, it was of interest to determine whether any of the genomic clones identified in this screen represented Y linked loci. As has been described in more detail in chapter 1, there is an absolute requirement for the Y-chromosome in *Drosophila* in order for normal spermiogenesis to take place. Although very few genetic loci have been localised to the Y-chromosome, those such as the fertility factors, that have been located here seem to have an essential function during spermatogenesis. In addition, other Y-specific loci have been implicated in having a spermatogenic function. A more detailed description of these Y-linked genes is given in chapter 1. As described in chapter 1, genomic DNA clones gS1, gS8 and gS9 characterised by Russell (Russell & Kaiser, 1993) represent a class of sequences that are present on the Y-chromosome as well as on an autosome (Y-associated). The autosomal copy is at position 77F on the third chromosome where it resides within the first intron of a previously described *Drosophila* gene for the RI subunit of the cAMP-dependent protein kinase (Kalderon & Rubin, 1988). In addition, there appear to be five Y-linked copies of this sequence. Sequence analysis of the cDNAs corresponding to these genomic DNA clones suggests that the gene product resembles both members of the histone H1/H5 gene family and the cystine-rich protamines. Since protamines are smaller proteins than histones, they are thought to be involved in the compaction of the chromatin with in the sperm head.

As has been described in chapter 3, differential screens were used by several investigators (Schafer, 1986a; DiBenedetto, 1987) in an attempt to identify testis-specific transcription units. The screen performed by Schafer identified twelve genomic DNA clones which hybridised only to male cDNA and not embryonic nor female cDNA. Most of these represented genes whose expression was

restricted to the accessory gland (paragonia). However, one of these designated Mst(3)g19 (later renamed Mst(3)CPG) was found to have expression that was restricted to the germline (Schafer, 1986b). The screen by DiBenedetto identified six putative male-specific transcripts, one of these was found be expressed in the accessory gland while the five others were restricted to the germline. In addition, as discussed previously, the enhancer-trap screen of Castrillon (Castrillon *et al.*, 1993) identified 58 male-sterile loci (some of these represented previously identified male-sterile genes). It was therefore of interest to determine whether the genomic DNA clones described here represented novel 'spermatogenic' loci or if they represented previously identified genes or members of a gene family.

Since the chromosomal locations of these other spermatogenic loci was known, determining the cytological locations corresponding to these genomic DNA clones would provide a means of determining whether the latter represent previously identified loci. Other information that the knowledge of the chromosomal location would provide is the identification of enhancer-trap elements that are located in or near the region covered by the genomic DNA clone. As has been discussed, an obvious disadvantage of differential screening is that it does not generate mutations that can be used to give an insight into function. The identification of a closely-linked P-element or enhancer-trap element could be used to address this problem. If the element is located within the region of interest, it may give rise to a mutant phenotype. More severe mutations my be generated by the use of imprecise excisions. If the element does not actually lie within the region of interest, mutations can be created using 'local jumps' (Littleton *et al.*, 1993).

Results.

4.1 Subcloning Of The Male Specific Fragments.

Using the data from the 'Reverse northern' analysis discussed in chapter 3, the five male-specific fragments were subcloned into the plasmid pBluescript (Stratagene) for use as probes in subsequent analysis. Clone gK10 contained a single *Eco*RI/*Sal*I fragment of 7.3 kb which hybridised to the male-specific cDNA probe and not to the female cDNA probe. This male-specific transcript was subsequently found to be contained within a 5 kb *Hin*dIII fragment which was subcloned into the plasmid vector pBluescript (Stratagene). The resulting subclone containing the male-specific fragment was designated pmsfK10 (male-specific fragment gK10).

The genomic lambda clone gK22 also contained one male-specific fragment of 8.9 kb. This 8.9 *Eco*RI/*Sal*I fragment was subcloned into pBluescript and the recombinant plasmid was designated pmsfK22.

The two male-specific fragments of 1.4 and 1.15 kb contained in gK27 were sub-cloned as *Eco*RI fragments in pBluescript. The recombinant plasmid containing the 1.4 kb fragment was designated pmsfK27.1, whilst that containing the 1.15 kb fragment was designated pmsfK27.2.

Clone gK31 contained a single male-specific fragment of 12.3 kb. This fragment was subcloned as an *Eco*RI/*Sal*I fragment in pBluescript, pmsfK31.
These results are summarised in Table 4.1.

genomic DNA clone	recombinant plasmid	Size and nature of male-specific fragment (kb)
gK10	pmsfK10	5 kb *Hin*dIII
gK22	pmsfK22	8.9 kb *Eco*RI/*Sal*I
gK27	pmsfK27.1 & pmsfK27.2	1.4 & 1.14 kb *Eco*RI
gK31	pmsfK31	12.3 kb *Eco*RI/*Sal*I

Table 4.1: Table 4.1 shows the original genomic DNA clones with their corresponding subcloned male-specific fragment.

All five male-specific plasmids were used as probes in the subsequent characterisation described in this and the subsequent chapters.

4.2 Restriction Mapping Of The Genomic Lambda clones.

In order to determine the position of the male-specific fragments within the original genomic lambda clone, the phage DNA was restricted with a combination of the enzymes *Eco*RI, *Sal*I, and *Hin*dIII. *Sal*I cleaves the Lambda DNA to release the insert DNA from the Lambda arms (19.9 kb long arm and 8.8 kb short arm). The *Eco*RI sites present in the polylinkers are lost by cloning *Sau*3A genomic fragments into the *Bam*HI site. Consequently, any *Eco*RI fragments produced are as a result of internal *Eco*RI sites. The only *Hin*dIII sites present in the lambda arms is situated in the short arm which is cleaved into two fragments of 4.3 and 4.5 kb. This provided sufficient information to construct a restriction map for gK22 and gK31. For gK10 and gK27, which were much more complex phage, additional information was inferred from the smaller overlapping phage gK6 (for gK27) and gK8 (for gK10).

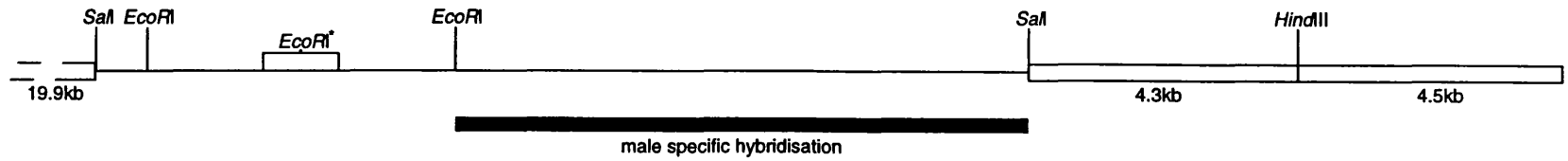
The resulting restriction maps are shown in Figure 4.1

Genomic Southerns.

4.3 Analysis Of Genomic Organisation Of Male And Female DNA.

In order to determine whether any of the genomic DNA clones were derived from the Y-chromosome, genomic DNA was prepared from males and females of three wild type *Drosophila* strains, Oregon-R, Canton-S and m56i. Since the only difference between the male and female flies is the presence or absence of

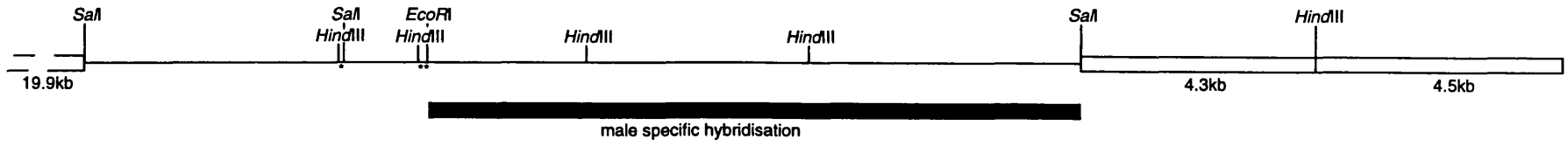
(a) Restriction map of genomic clone gK22



*Position of this site is ambiguous

1kb

(b) Restriction map of genomic clone gK31



*Sites separated by ~100bp. Order is unambiguous

**Sites separated by 200bp

1kb

Figure 4.1: Restriction maps of the genomic clones.
Figure 4.1a: Restriction map of clone gK22. Male-specific hybridisation is localised to a 8.9 kb *Eco*RI, *Sal*I fragment. There are no internal *Hind*III sites.
It is not possible to orientate the 1.9 and 3.3 kb *Eco*RI fragments.
Figure 4.1b: Restriction map of clone gK31. Male-specific hybridisation localised to a 12kb *Eco*RI fragment.

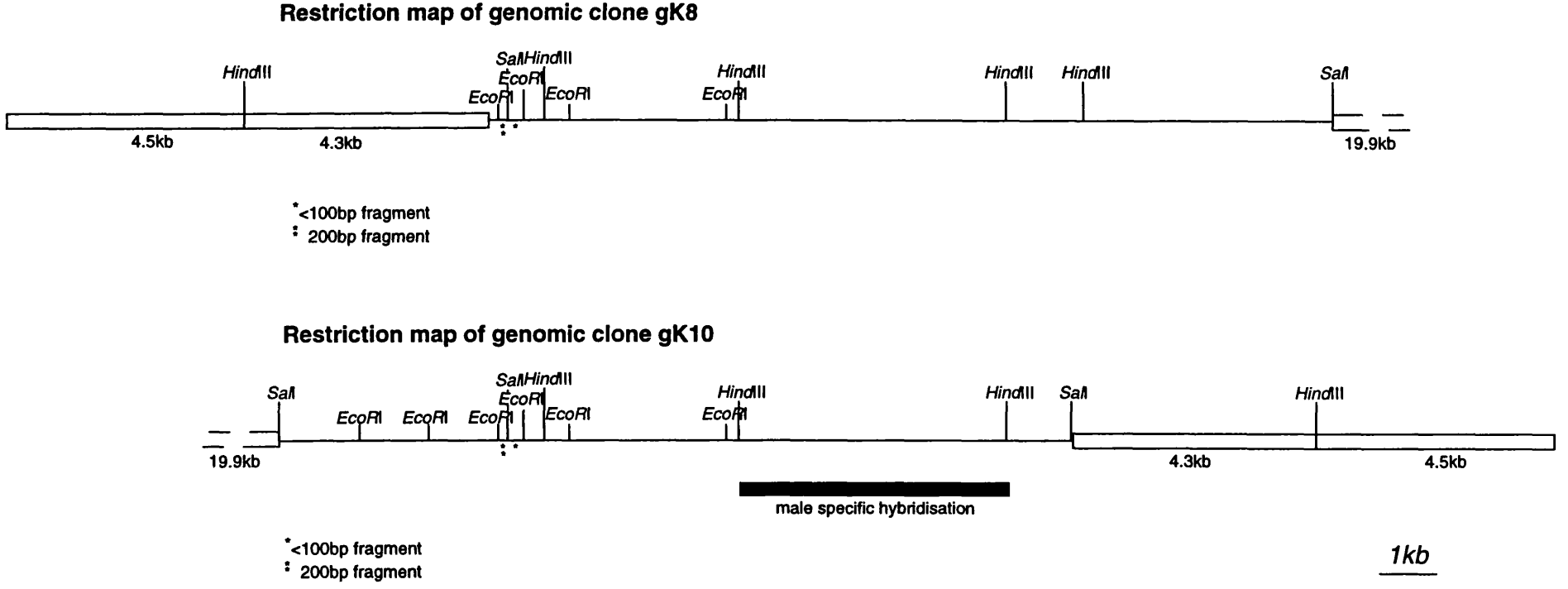


Figure 4.1c: Restriction map of clones gK8/gK10. Male-specific hybridsation localised to a 5 kb *Hind*III fragment of gK10. This fragment is wholly contained within clone gK8.

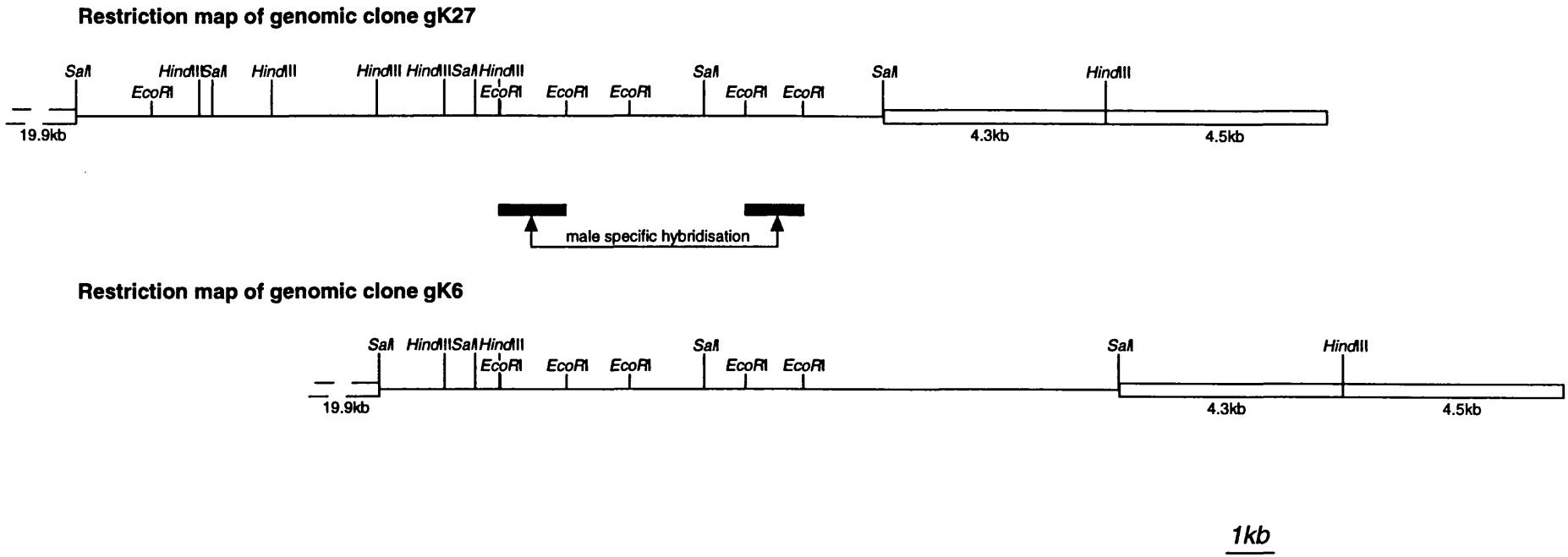


Figure 4.1d: Restriction map of clones gK6/gK27. Male-specific hybridisation localised to two *Eco*RI fragments of 1.4 and 1.15 kb. These fragments are not contiguous within the genome.

the Y-chromosome, any difference in the hybridisation pattern observed will be due to the Y-chromosome sequences.

Genomic DNA was restricted with *Hin*dIII for hybridisation with the pmsfK10 probe, *Eco*RI for the pmsfK27.1 and pmsfK27.2 probes and *Eco*RI/*Sal*I for the pmsfK22 and pmsfK31 probes.

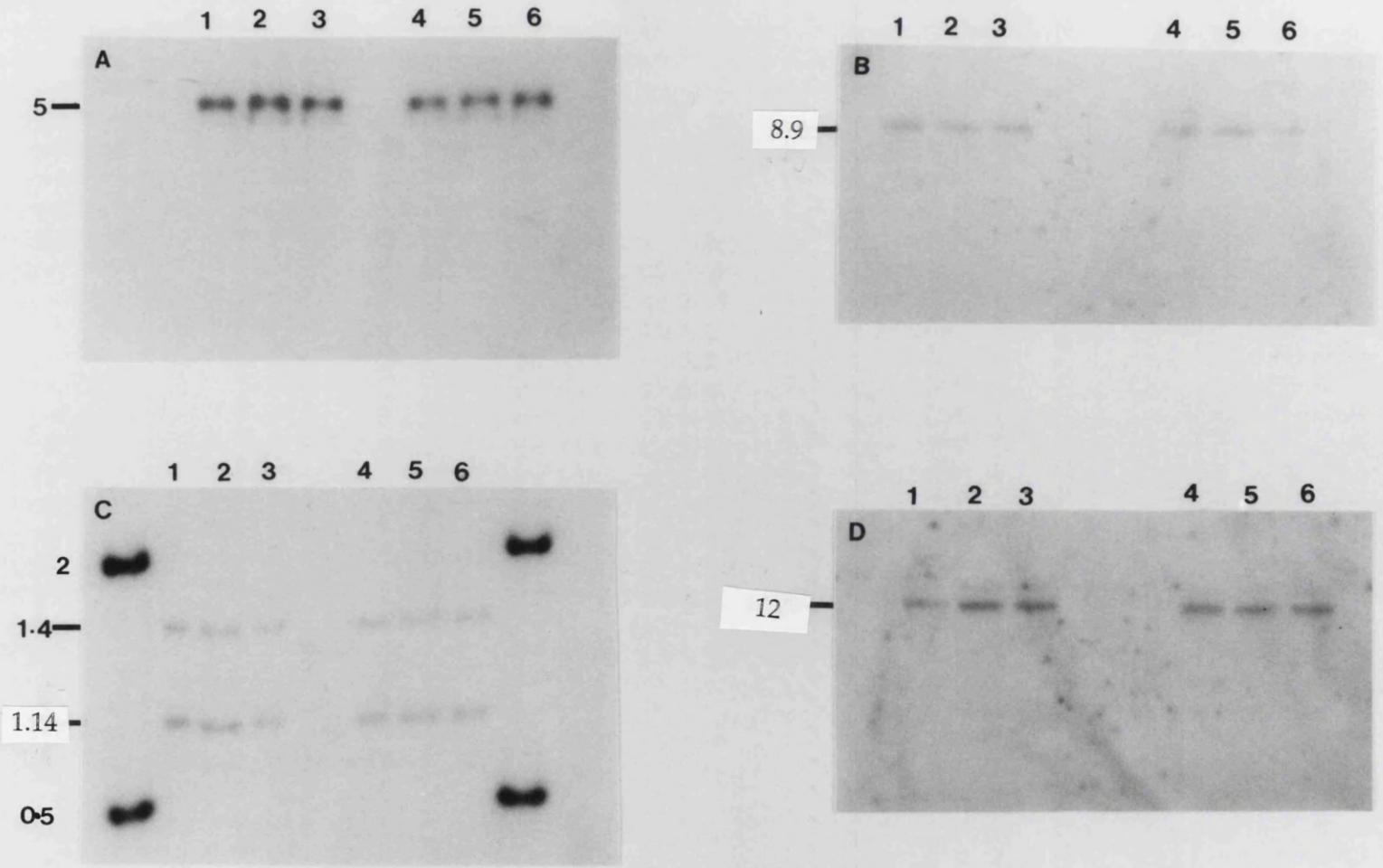
When male and female genomic DNA was restricted with *Hin*dIII and hybridised with pmsfK10 probe, (Figure 4.2A) a single band of 5 kb is observed in both males and females of all three wild type strains. If pmsfK22 is used as a probe to *Eco*RI/*Sal*I restricted genomic DNA, a single band of 8.9 kb is observed in all six tracks (Figure 4.2B). Figure 4.2C shows a mixed probe of pmsfK27.1 and pmsfK27.2 hybridised to genomic DNA restricted with *Eco*RI. Again there is no difference observed between males and females with two bands of 1.4 and 1.15 kb identified in all tracks (marker bands correspond to the 1.5 and 0.9 kb bands of lambda). Finally, when pmsfK31 is probed to genomic DNA restricted with *Eco*RI/*Sal*I, a single band of 12.3 kb is observed in all six tracks (Figure 4.2D). None of the genomic DNA clones show any evidence for polymorphism at these loci. From these results it was evident that none of these genomic DNA clones are derived from the Y-chromosome.

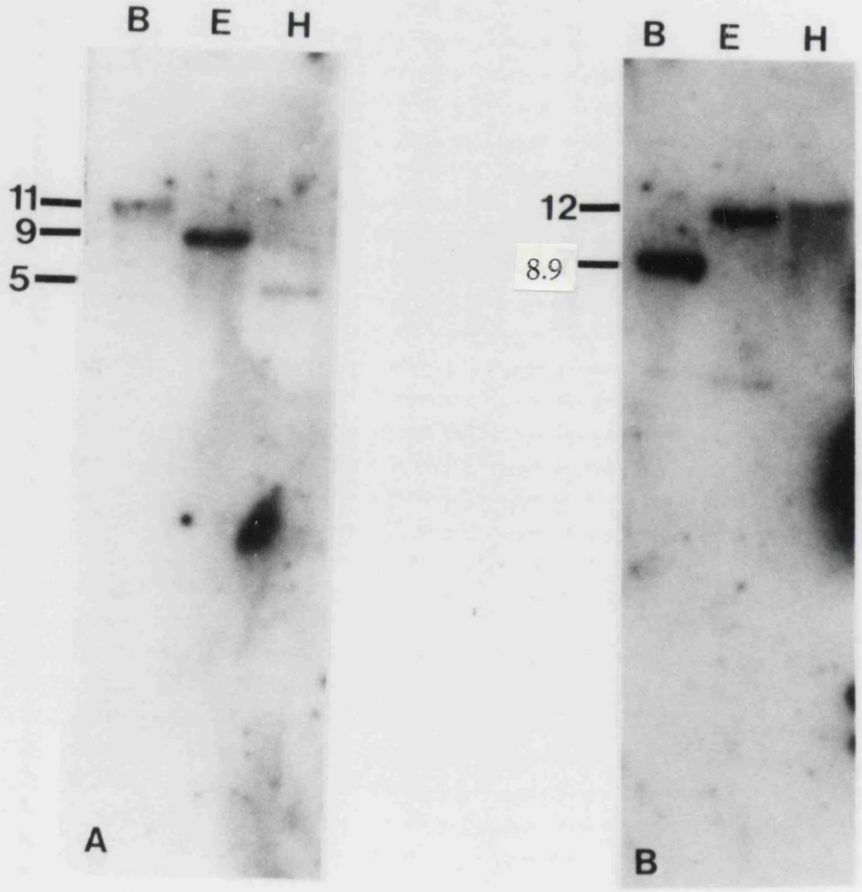
4.4 Genomic Southerns

The results obtained from Southerns of male and female genomic DNA suggested that all these clones represented single copy genes that were not present on the Y-chromosome. To confirm that these represented single copy genes and to establish whether the restriction maps deduced in Section 4.1 were correct, genomic DNA was prepared from wild-type Oregon-R flies and restricted with *Bam*HI, *Eco*RI and *Hin*dIII. All five male-specific fragments were used as probes. The results are shown in Figure 4.3.

Figure 4.2: Genomic Southerns with male and female DNA.
Tracks 1, 2 and 3 genomic DNA derived from males of three wild type *Drosophila* strains, Oregon-R, Canton-S and M56i. Tracks 4, 5 and 6, genomic DNA from females of Oregon-R, Canton-S and M56i.
Figure 4.2A: *Hin*dIII digest. Hybridised with pmsfK10.
Figure 4.2B: *Eco*RI/*Sal*I digest. Hybridised with pmsfK22.
Figure 4.2C: *Eco*RI digest. Hybridised with a mixed probe of pmsfK27.1 and pmsfK27.2. The marker bands in this panel represents end labelled lambda DNA (1.5 and 0.9 kb fragments).
Figure 4.2D: *Eco*RI/*Sal*I digest. Hybridised with pmsfK31.

All the filters were washed to high stringency and exposed for 48 hours.


1 2 3 4 5 6
A
5
1 2 3 4 5 6
B
8.9
1 2 3 4 5 6
C
2
1·4
1.14
0·5
1 2 3 4 5 6
D
12


B
E
H
11
9
5
A
B
E
H
12
8.9
B

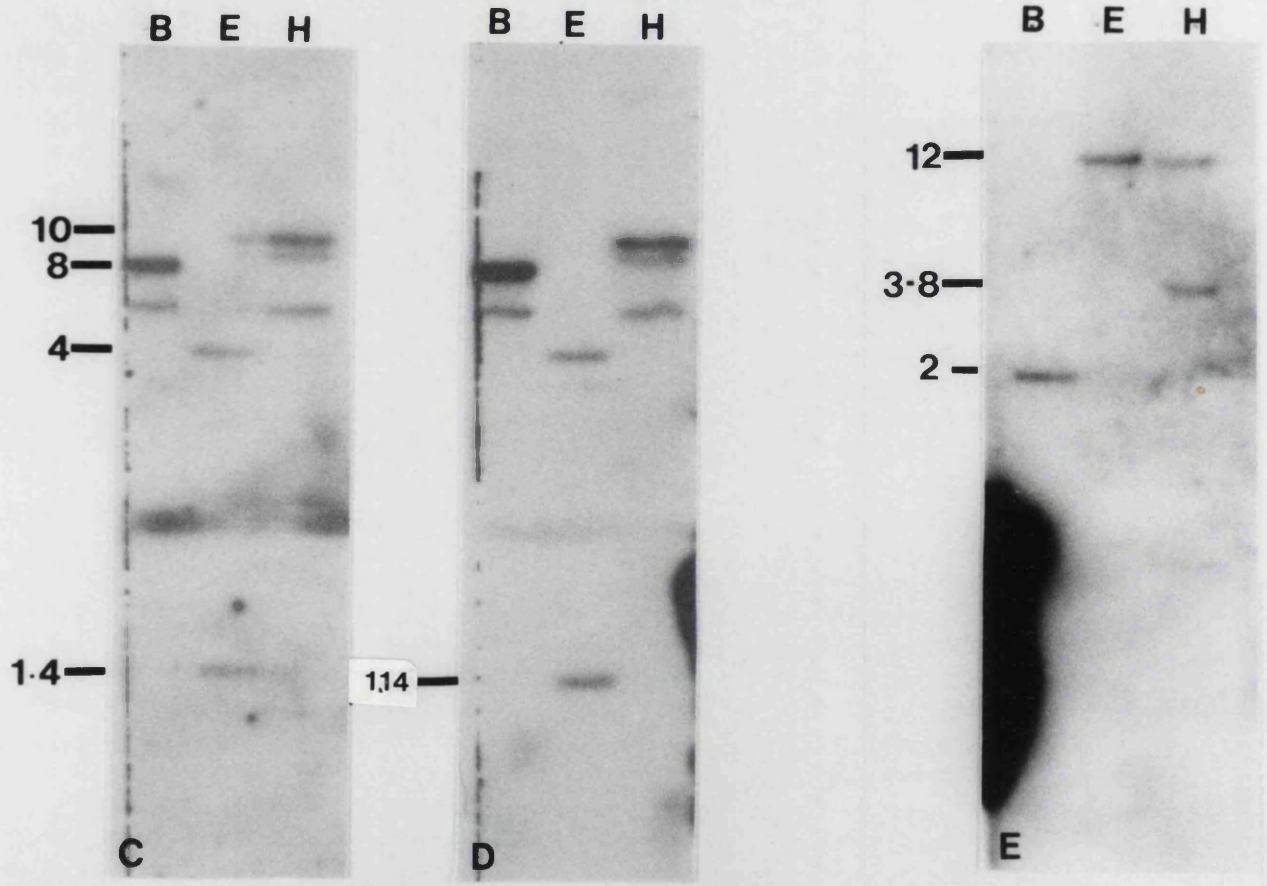

B E H
10
8
4
1·4
C
B E H
1.14
D
B E H
12
3·8
2
E

Figure 4.3: Genomic Southerns.
Genomic DNA prepared from Oregon-R and restricted with *Bam*HI (B) track 1, *Eco*RI (E) track 2 and *Hin*dIII (H) track 3. Southern blotted and hybridised with the male-specific fragments.
Figure 4.3A: pmsfk10.
Figure 4.3B: pmsfk22.
Figure 4.3C: pmsfk27.1.
Figure 4.3D: pmsfk27.2.
Figure 4.3E: pmsfk31.

All filters were washed to high stringencies and exposed for 48 hours.

Genomic DNA probed with pmsfK10 identifies a single band in each track. In the *Bam*HI digest, a fragment of 11 kb hybridises to the probe. (Figure 4.3A, track 1). In the *Eco*RI digest, a single band of 9 kb is observed (Figure 4.3A track 2), while in the *Hin*dIII digest, (Figure 4.3A track 3) a band of 5 kb is seen.

Figure 4.3B shows genomic DNA probed with pmsfK22, in track 1 the *Bam*HI digest, a single band of 8.9 kb hybridises to the probe. The *Eco*RI digest (Figure 4.3B track 2) also identifies a single band of 12 kb, whilst the *Hin*dIII digest gives a fragment of ~13 kb.

Clone gK27 appears to have a slightly more complex structure than that of the other clones characterised. Unlike all the others, this clone contains two male-specific bands. From sequence evidence (described in chapter 6) they appear to be duplications of a related sequence. Genomic Southerns using both pmsfK27.1 and pmsfK27.2 (Figure 4.3C and 4.3D) show that for the *Bam*HI digest a single band of 8 kb is observed with both probes (Figure 4.3C track 1 and Figure 4.3D track 1). In contrast, the *Eco*RI digest identifies two fragments. A fragment of 4kb with both probes (Figure 4.3C and D track 2)and a 1.4 kb fragment which is identified with the pmsfK27.1 probe, (Figure 4.3C track 2) whilst the pmsfK27.2 probe identifies a 1.15 kb fragment (Figure 4.3D track 2). For the *Hin*dIII digest, both probes identify the same fragment of 10 kb (Figure 4.3C track 3 and Figure 4.3D track 3). In addition to the strongly hybridising bands described above, other bands which hybridise less strongly to the probe were also observed. These bands were reproducible in genomic Southerns using all three probes derived from this region (pmsfK27.1, pmsfK27.2, pcK27). This implies that there are related sequences present in the genome. This will be discussed in more detail in chapter 6.

The same procedure was used to confirm the genomic organisation of gK31. As before, the DNA was restricted with *Bam*HI identifying a single band of 2 kb (Figure 4.3E track 1). The digest using *Eco*RI shows a single band of 12 kb (Figure 4.3E track 2). For the *Hin*dIII digest, two fragments were identified of 12 and 3.8 kb.

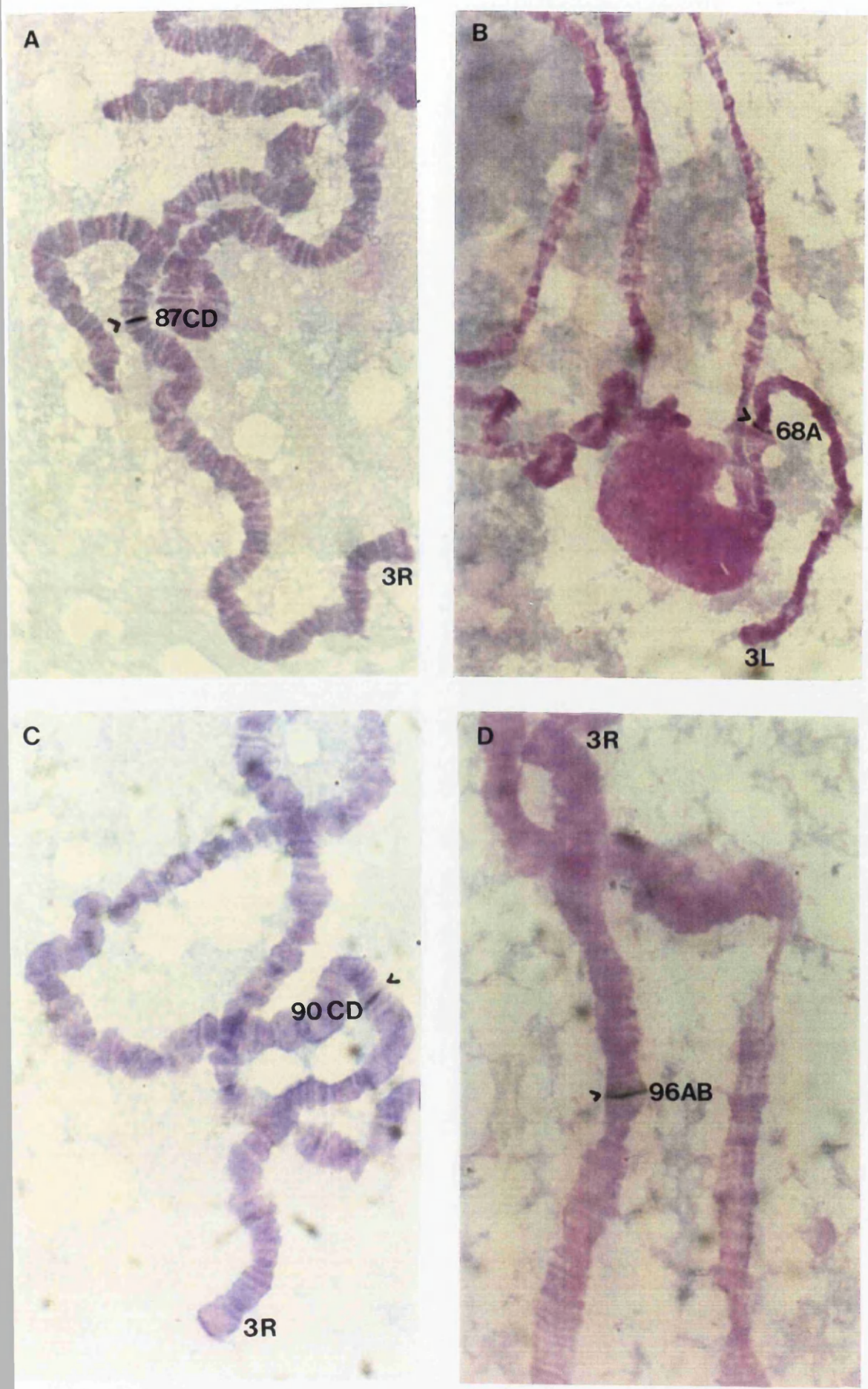
4.5 In Situ Hybridisation To Polytene Chromosomes.

As discussed previously, *Drosophila* are able to over-replicate their chromosomes in some tissues such as the salivary glands. This over replication enables the visualisation of the chromosomes, allowing the determination of the chromosomal location of the clone being characterised. The banding patterns of the polytene chromosomes are constant for the species studied, the bands observed can therefore serve as landmarks for locating various genomic features. The standard reference for *D. melanogaster* salivary gland chromosomes are the Bridges' maps. The photographic version of these maps produced by Lefevre (Ashburner and Novitski, 1976) was used to determine the cytological location of the male-specific fragments. The male-specific fragments are hybridised to the chromosome spreads and the signal detected as described in Section 2.13.1. This ensures that all the clones being characterised do not represent genes that have been isolated previously. As described above the knowledge of the cytological location of the clones provides additional information that can be used to assist in characterisation. The use of P-elements and enhancer-trap elements in order to assist in assigning function has been described in this chapter and also in chapter 3. The location of many P-elements and enhancer-trap elements are known. Consequently, if they are located within the same polytene band they can be used to generate mutations (Kaiser and Goodwin 1990; Littleton *et al.*, 1993). In addition, if the clone lies in a region of the genome that has been extensively characterised, various duplications and deletions of

that region of the chromosome may be available. These can facilitate studies on the effects of dosage of the gene on the development of the fly. All five of the male-specific fragments were labelled with biotin (as described in Section 2.10.3) and used to probe chromosome spreads from wild type Oregon-R.

As expected from previous results, all of the clones were located on the autosomes. The invariant banding patterns of the chromosomes in conjunction with chromosome specific areas of constrictions (areas of under-replication) can be used as markers to determine which chromosomal bands the probes hybridise to. In addition, the ends of the polytene chromosomes differ, allowing identification of the chromosome to which the probe is hybridising. The chromosomes also display areas of transcriptional activity represented by the appearance of 'puffs' in the polytene chromosomes. As a developmental profile of the puffing pattern of the chromosomes has been produced (Ashburner 1972) these may also be used as makers to assist in localisation of the probe. All of the male-specific fragments hybridise to single bands on the third polytene chromosome. Three of them, gK10, gK27 and gK31 are located on the right arm of chromosome three, while gK22 is located on the left arm. Figure 4.4 shows the *in situ* hybridisation of the male-specific fragments to wild type polytene chromosomes. The right arm of the third chromosome (Figure 4.4 A, C and D) has a flattened tip while that of the left arm (Figure 4.4B) is more rounded. The position of pmsfK22 was easily determined as it is situated in a very prominent landmark on the left arm of the third chromosome at 68A. The positions of the other fragments were determined as described previously by looking for landmark constrictions and puffs which were then used to assign chromosomal positions. The cytological locations of each of these clones are given in Table 4.2. At the time that the initial characterisation of these clones was taking place, no male-specific loci had been identified at these positions. However, recently an enhancer-trap screen for genes involved in

Figure 4.4: *In situ* hybridisation to polytene chromosomes spreads from Oregon-R.
Figure 4.4A: hybridised with biotinylated pmsfK10.
Figure 4.4B: hybridised with biotinylated pmsfK22.
Figure 4.4C: hybridised with biotinylated pmsfK27.1
Figure 4.4D: hybridised with biotinylated pmsfK31.


A
87CD
3R
B
68A
3L
C
90 CD
3R
D
3R
96AB

Name of plasmid	Cytological location
pmsfK10	87CD
pmsfK22	68A
pmsfK27.1	90CD
pmsfK27.2	90CD
pmsfK31	96AB

Table 4.2: The cytological locations of male-specific fragments from the genomic DNA clones.

spermatogenesis (described in chapter 1) produced P-element inserts into the same chromosomal locations as those of both pmsfK22 and pmsfK27 (Castrillon *et al.*, 1993). The characterisation of these enhancer-trap lines in relation to pmsfK22 and pmsfK27 will be discussed further in chapter 5.

Discussion.

All of the genomic fragments identified by 'Reverse northern' analysis represent (with the exception of pmsfK27.1 and pmsfK27.2) single copy genes. gK27 in contrast to the other clones, seems to consist of a duplication of the male-specific region. In addition, it is apparent from Southern analysis that related sequences may also be present elsewhere in the genome (but at the same cytological location). None of these additional fragments hybridise with an intensity that suggests that they may be the result of partial digestion or 'star' activity of the enzymes used. Fragments resulting from partial enzyme digestion would be expected to be larger than the proper fragment sizes, it can clearly be seen in Figure 4.3 C and D that this is not the case. The less intensely hybridising fragments are smaller than those that hybridise strongly to the probe. This supports the suggestion that these bands may be due to related sequences. None of these fragments represent Y-specific or Y-associated sequences. Additionally, no polymorphism at these loci between three wild type strains of *Drosophila* examined could be observed suggesting that these sequences are conserved between the different strains.

The aims of this chapter were to establish the genomic organisation of the genomic clones by restriction mapping and Southern blotting. *in situ* hybridisation to the polytene chromosomes was also used to determine if these represented previously identified mutations. The results obtained using genomic Southerns and *in situ* hybridisation to the polytene chromosomes

suggested that three of the genomic DNA clones (gK10, gK22 and gK31) are single copy autosomal genes. In the case of gK27, it is apparent that at least two copies of this sequence share the same cytological location. As can be seen from the Southern analysis, this locus has a very complex genomic organisation. The reasons for this complexity will be discussed in chapter 6.

When these results were obtained, no male-specific lethals or male-steriles had been identified at these chromosomal locations. However, toward the end of this work, possible enhancer-trap insertions within the cytological locations of two of the clones (gK22 and gK27) were generated by Castrillon (1993).

Introduction.

An important consideration in the process of spermatogenesis is the control of gene expression. As described in chapter 1, the transcription of all gene products required for the spermatogenic process is believed to occur during the primary spermatocyte stage, before meiosis has occurred. However, most of the morphogenic events that go in to fashioning the mature sperm take place in the post-meiotic spermatid. All the events of spermiogenesis (i.e. the translation and assembly of the components) take place without a requirement for the presence of DNA or RNA synthesis. Evidence that transcription is restricted to the primary spermatocyte stage has come from a number of sources: genetic studies have shown that spermatids that do not have a full chromosome complement are able to develop into mature sperm. Indeed, spermatids which only possess the fourth chromosome are able to produce mature sperm (Lindsley & Grell, 1969). These sperm are able to produce viable zygotes if they fertilise eggs which contain the chromosome constitution missing from the sperm. These defective sperm usually result from non-disjunction events as a result of mutations affecting meiosis I. After meiosis, the haploid sperm carrying the Y-chromosome will lack X-linked spermatogenic loci such as *stellate*, (*ste*). However, as has been discussed in chapter 1, incomplete cytokinesis is a feature of spermatogenesis, therefore spermatids develop side by side in a common cytoplasm. As the cytoplasmic environment of each spermatid is not independent, the RNA transcribed from each nucleus is translated in the shared cytoplasm. This syncytial development means that after meiosis the nuclear constitution may no longer make a significant contribution to the developmental process of the spermatid.

In addition, the autoradiographic studies of Olivieri & Olivieri (1965) and Gould-Somero & Holland (1974) provide further support for this theory. Olivieri & Olivieri set out to establish whether transcription during spermatogenesis was restricted to pre-meiotic stages in *Drosophila*. The incorporation of both [3H]uridine (to detect RNA synthesis) and [3H]thymine (to detect DNA synthesis) in the *Drosophila* testis was assessed using autoradiography. Labelled uridine could be detected in the nucleus 30 min after the initial injection. It was observed that eight hours after injection, the cytoplasm and the nucleolus were labelled while the nucleus no longer contained any label. The label was transmitted to the spermatid cytoplasm after four days and to mature sperm after six. No RNA synthesis could be detected in post-meiotic stages. DNA synthesis however, could be detected in spermatogonia 30 min after [3H]thymine injection. At eight hours post injection, this label was detected in spermatocytes and was transmitted through all the spermatogenic stages until it was detected in mature sperm in the seminal vesicles 10 days later.

Gould-Somero & Holland devised a method of culturing testes from late third instar larvae (containing spermatogonia and primary spermatocytes) or 1-2 day old pupae (which also contain elongating spermatids). Spermatogonia and the younger primary spermatocytes in the apex of the pupal testes were heavily labelled with [3H]uridine; whilst those older primary spermatocytes incorporated less label. Their results indicated that pre-meiotic cells were the only ones which accumulated label; post-meiotic cells were never labelled. To confirm that RNA synthesis only occurred during the primary spermatocyte stage, they studied the effects of inhibiting RNA synthesis on differentiation. Again, their results seem to confirm the findings Olivieri & Olivieri; namely that the RNA species required for spermiogenesis are synthesised during the primary spermatocyte stage.

As all the RNA products required for spermiogenesis are synthesised early during the spermatogenic process, mechanisms must exist which ensure the stability of the messages produced during the primary spermatocyte stage until spermatid maturation when they are required. During the autoradiographic studies described previously, the levels of [^{3}H]uridine incorporated during the primary spermatocyte stage persisted in the cytoplasm into the late stages of spermatid differentiation several days later. Studies on the testis-specific β-2tubulin (*β*2t) message provide preliminary evidence of a *cis*-acting sequence which may affect its stability. An 18 bp motif has been isolated in the 5' untranslated leader of the *β*2t message in both *D. melanogaster* and *D. hydei* (Michiels *et al.*, 1989). This motif is able to direct increased expression of the *lacZ* reporter gene in the adult testes (where post-meiotic stages are present) compared to the larval testis (which contains only spermatogonia and primary spermatocytes). Deletion of this region results in the same levels of expression of the reporter construct in both adult and larval testes.

In addition to message stability, there must be some form of translational control to ensure that these proteins are not produced at inappropriate times during the process of spermiogenesis. The mechanism for translational control has been extensively studied in the Mst(3)CPG family of genes. As discussed in chapter 1 , it has been postulated that this gene family is responsible for the production of sperm tail accessory proteins. The *cis*-acting sequences that act to control the testes-specific transcription and delay transcription until after meiosis have been identified (Schafer *et al.*, 1990). An element in the 5' leader sequence functions in directing testis-specific expression of a *lacZ* reporter construct in the larval testis and also of the resulting fusion protein in the elongated spermatid bundles.

This sequence, known as the translational control element (TCE), is conserved between all seven members of the Mst(3)CPG family. Deletion of this element leads to premature translation of the message in the larval testis. In addition, this element also seems to mediate elongation of the existing poly A^+ tail. This appears to be a common feature of messages which are transcriptionally regulated, the length of their poly A^+ tail increases when they become transcriptionally active.

Further support for a mechanism involving translational control is provided by the *janus* locus. This locus consists of two overlapping transcripts, *jan*A and *jan*B, both of which are expressed during spermatogenesis. *jan*A is expressed in both germline and somatic tissues of males and females; while *jan*B expression is restricted to the testis during spermatogenesis (Yanicostas & Lepesant, 1990). In common with other spermatogenic loci, *jan*B is transcribed during the primary spermatocyte stage. A sequence motif has been identified in the 5' region of the *jan*B message which appears to restrict transcription to the primary spermatocyte stage, and prevent translation until spermiogenesis. A comparison of this sequence with that obtained for the Mst(3)CPG TCE revealed that seven bases of the twelve base pair motif of the Mst(3)CPG TCE, are conserved in an element in the 5' leader of the *jan*B message. Additional residues outwith the TCE consensus sequence have also been conserved suggesting additional elements that may exert translational control. Although homology has been found between the putative control elements of these two genes, neither of them show any homology with the postulated control sequence of the testis-specific β–2tubulin (β2t) gene. The β2t gene, unlike *jan*B and the Mst(3)CPG family is not transcriptionally regulated, it is both transcribed and translated in the primary spermatocyte stages. Consequently, the TCE motif may represent a mechanism by which some of the gene products produced in the primary spermatocyte stage have their translation delayed until they are required during spermiogenesis.

The abundance of *Drosophila* single gene mutants provides a powerful mechanism whereby the expression pattern of a particular transcript can be studied. Mutants in the two processes of somatic sex determination and formation of the primordial germ cells, can be used to further elucidate whether expression of a transcript is restricted to the germline or to the soma.

The somatic sex determination hierarchy in *Drosophila* has been extensively characterised at both the genetic and molecular level (Baker, 1989; Cline, 1993). A hierarchy of four genes *Sex lethal* (*Sxl*), (Cline, 1978) *transformer* (*tra*) (Sturtevant, 1945), *transformer*2 (*tra2*) (Watanabe, 1975) and *intersex* (*ix*) (Baker & Belote, 1983) results in the differential splicing of the downstream 'binary switch' gene *doublesex* (*dsx*) (Hildreth, 1965) into the male (DSX^M) or female (DSX^F) form of the protein. Most of these genes are active only in the female, the male is the default state. The female specific SXL product is required for the correct splicing of *tra* to give a functional message. *tra* in conjunction with *tra*2 is required to direct the splicing of the *dsx* transcripts into the female form (dsx^F) of the message. In the male, the absence of functional SXL protein ensures that splicing of the *tra* message into the female form does not occur. *tra2* alone is unable to direct the female-specific splicing of the *dsx* transcript and the male protein (DSX^M)is produced by default. The female DSX^F protein, in conjunction with the *ix* product, acts to repress genes involved in male differentiation, allowing female differentiation to occur. In the male, the DSX^M protein represses female differentiation genes, allowing male development to occur. Mutations in genes of this hierarchy exhibit sex-specific lethality, sex-transformation or intersexuality.

Studying the effects of a mutation in either of the *transformer* (*tra*, *tra2*) genes can give an indication of whether the transcript is expressed in the soma or in the germline. Mutations in these genes result in the transformation of females

into pseudomales. These are females whose somatic tissue is transformed to male soma while their germline remains female. The absence of a transcript on northern blots using RNA derived from the pseudomales and its presence in RNA derived from normal males would suggests that a male germline is required for its correct expression.

Mutations in some of the genes involved in formation of the primordial germ cells (pole cells) exhibit a *grandchildless-knirps* phenotype i.e. sequestration of cells does not occur, therefore no polar granules are present and hence no pole cells are formed. In addition, in some mutants, abdominal segments are reduced or lacking (*knirps* phenotype). These are maternal effect mutants (i.e. mutations in these genes do not cause a defect in the mutant female but are observed in her progeny). The progeny of homozygous females are sterile due to their lack of pole cells. Accordingly, rudimentary ovaries (lacking oocytes) and testes (lacking sperm) are observed. The gene *tudor* (*tud*) is one of the seven genes in this class (Boswell & Mahowald, 1985). The effect of a mutation in this gene on the transcription of a clone can be used to determine whether a functional germline is required for its expression (i.e. transcripts that are specific to the germline will not be detected while those that are specific to the soma will).

in situ hybridisation to whole mount testes has been used by various investigators to study the expression pattern of testis-specific genes. In the case of both *twine* and *string*, expression is shown to be restricted to a limited subset of cells in the adult testis (Alphey *et al.*, 1992). *string* expression is restricted to the cells of the germinal proliferation centre at the apex of the testes, which is where the mitotic divisions of the germline stem cells occur. *twine* expression is restricted to the gyre of the testes where meiosis is thought to occur. Expression of this transcript can be detected in premeiotic cysts and in cysts which have completed meiosis I. No expression is detected in the

postmeiotic stages. The localisation of these transcripts supports the functions proposed for these homologues of cdc25. *string*, which is postulated to control the cell cycle during mitotic divisions is present in the germinal proliferation centre where mitosis of the stem cells occurs, whilst *twine* which is postulated to control the cell cycle during meiosis is localised with the premeiotic cysts. *In situ* hybridisation provides a useful mechanism for locating tissues where a transcript may be expressed. This in turn may provide an idea of function. As has been discussed for *string* and *twine*, the restricted pattern of *string* expression to an area of the testes where mitosis is known to occur, and the restriction of *twine* expression to the gyre where meiosis is thought to occur supports previous hypotheses regarding function.

In addition, the use of enhancer trap lines also facilitates the identification of enhancers that drive expression of a particular tissue-specific gene. Recent enhancer trap screens for spermatogenic loci (Castrillon *et al.*, 1993) isolated 58 male-sterile loci which are expressed in the testis. Using the chromogenic substrate X-gal, expression of a particular transcript in subsets of cells can be detected (Bownes, 1990; Gonczy *et al.*, 1992; Castrillon *et al.*, 1993). These lines provide a mechanism for attempting to identify genes involved in particular processes during spermatogenesis.

Chapter Five Expression Studies

Results.

5.1 Northern Analysis.

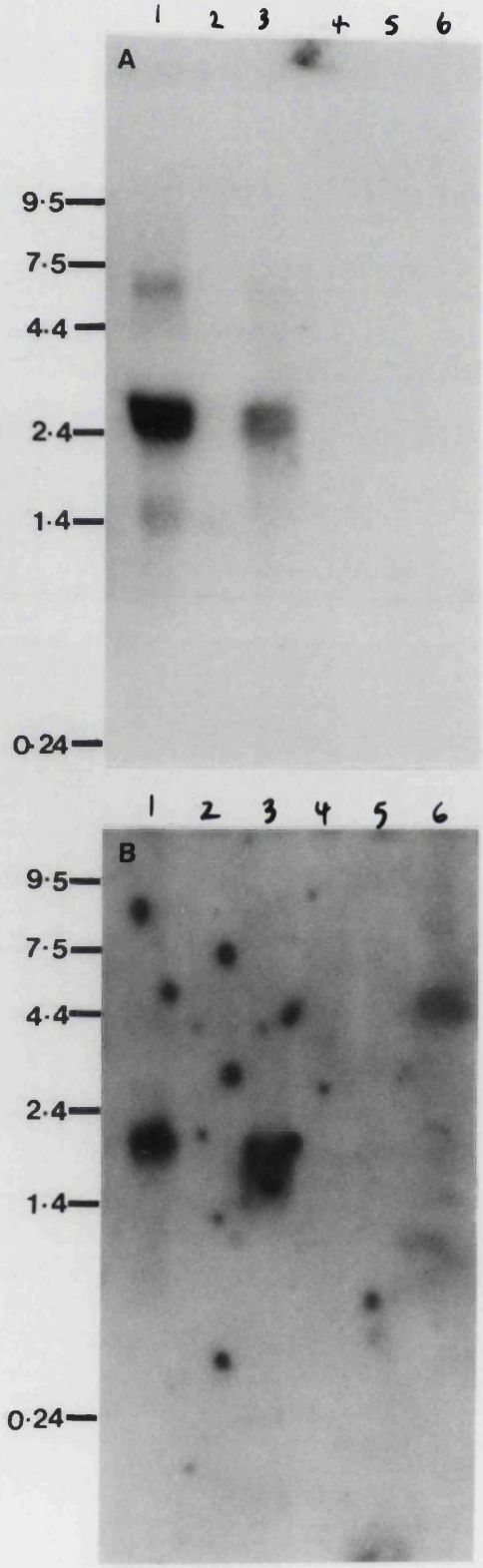
In order to determine whether the genomic fragments identified in chapter 3 represented true male-specific transcripts, RNA was prepared from wild type Oregon-R flies. To establish whether their transcripts were specific to the germline or to the soma, RNA was prepared from *tudor* (*tud*) and *transformer2* (*tra2*) flies. Details of the crosses used to generate these flies are given in Section 2.3. These blots were probed with each of the male-specific genomic fragments or a cDNA isolated using the genomic fragment as a probe (isolation of these cDNAs will be discussed in chapter 6).

Consistent with the hypothesis that some of these clones represent male-specific and not male-elevated transcripts (at the levels of detection used), when gK10 was used as a probe to total male and female RNA, three transcripts were detected. Two abundant transcripts of 2.7 and 2.6 kb are detected in the male and not in the female. (Figure 5.1 A). A less highly expressed transcript of ~6-7 kb is also observed. Comparison of expression in the chromosomal males that are mutant for *tra* and transformed pseudomales shows that the 2.7 and 2.6 kb transcripts are detected in the male and not in the transformed female. This indicates that the transcript may be germline-specific as this is the only tissue thought to differ between these flies. The studies on the agametic *tudor* flies adds further support to a pattern of expression that is restricted to the germline. These transcripts are lacking in both the male and female *tud* flies. Since these male *tud* flies differ primarily from normal males in not having a functional germline, the lack of these transcripts in these males again suggests that the expression of this transcript is restricted to the germline.

Figure 5.1: Northern analysis.
Track 1: Total mRNA from Oregon-R males. Track 2: Total mRNA from Oregon-R females. Track 3: Total mRNA from*tra* 2 males.
Track4: Total mRNA from *tud* males. Track5: Total mRNA from *tra* 2 females.
Track 6: Total mRNA from *tud* females.

Figure 5.1A: hybridised with pmsfK10.
Figure 5.1B: hybridised with pcK27.
RNA size marker is 0.24-9.5 kb RNA ladder from Gibco BRL.

Filters washed to high stringency and exposed for 3 days.


1 2 3 4 5 6
A
9·5
7·5
4·4
2·4
1·4
0·24
1 2 3 4 5 6
B
9·5
7·5
4·4
2·4
1·4
0·24

In the case of gK27, a cDNA (pcK27) corresponding to the genomic fragments was used as a probe. In common with the gK10 transcript, expression was again restricted to the male where a diffuse band of ~2.2 kb was seen. This transcript was lacking in the female. (Figure 5.1 B). Additionally this transcript could only be detected in chromosomal males mutant for *tra* and not in the pseudomale or any of the agametic flies. This again suggests that this transcript is restricted to the germline.

5.2 In Situ Hybridisation To Third Instar Larval Testis.

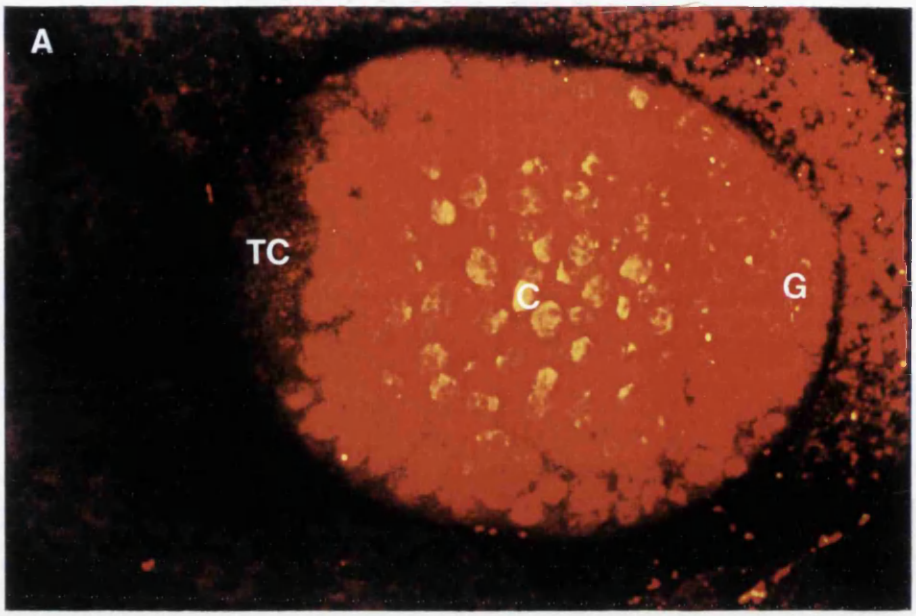
As the initial screen was performed using probes derived from the third instar larval stage, it would not be unreasonable to expect to detect expression of these transcripts in some larval tissues. As the results from the northern analysis suggested that these clones were specific to the germline, this was the tissue that was selected to look for tissue-specific expression. The gonads of the male third instar larvae can clearly be seen through the larval integument where they are encapsulated by the fat body. There is a definite gradation of cell types from the anterior (spermatogonia) to the posterior (primary spermatocytes) of the larval testis as shown in chapter 1, Figure 1.5. As discussed previously, most of the transcripts that are involved in spermatogenesis appear to be transcribed in the primary spermatocyte stage. Consequently, studying the expression of these transcripts in the larval testis, should indicate in which cell types they are expressed.

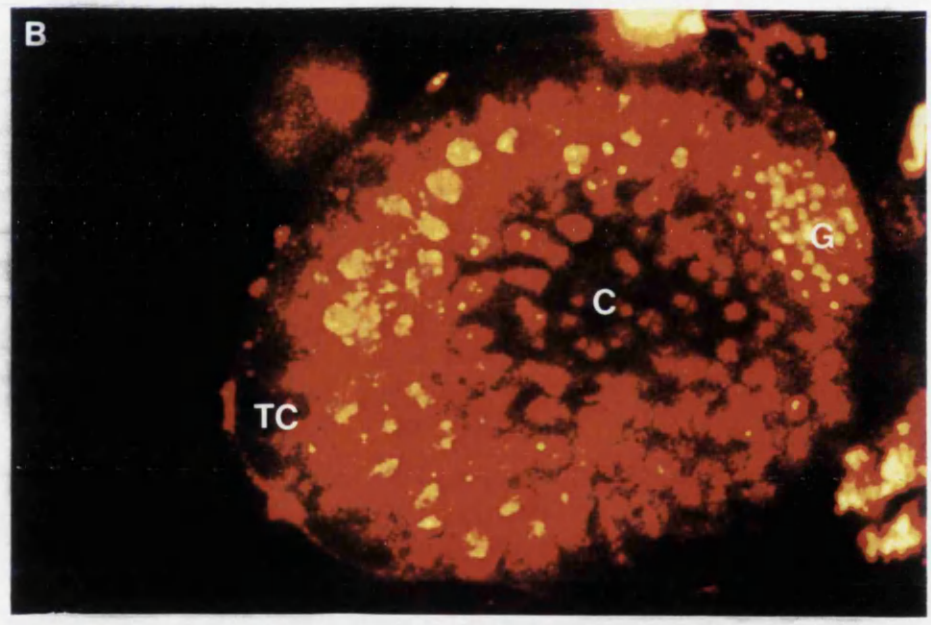
The testes were dissected and hybridised to biotin labelled probes as described in Section 2.13.2. The probes were either the original genomic fragment identified by 'Reverse northern' analysis or a cDNA corresponding to that genomic fragment.

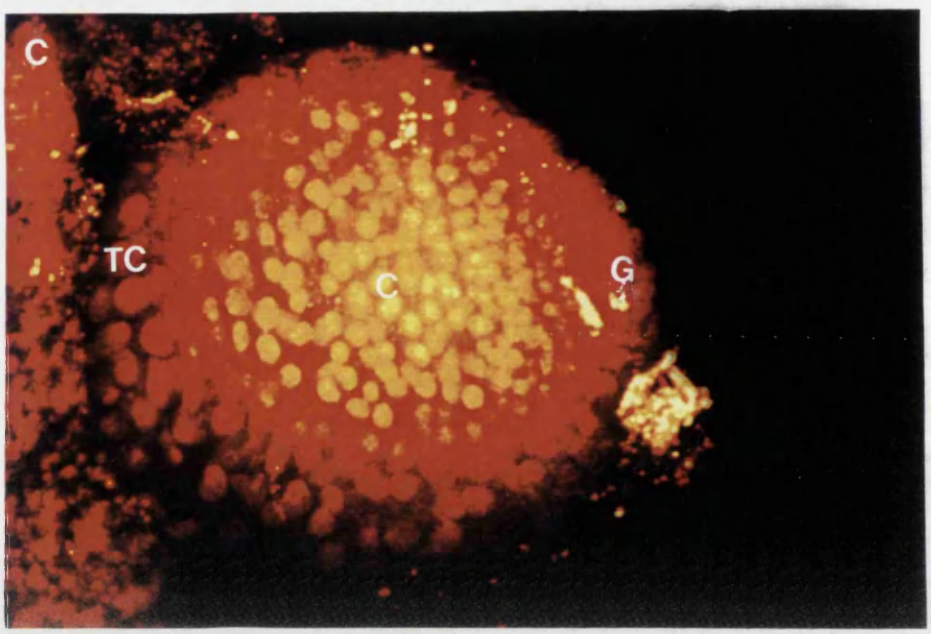
When hybridised to the larval testis plasmid pmsfK10 detects a high level of expression in the area corresponding to the primary spermatocyte (C). No expression is observed in the terminal cells (TC) of the testis (Figure 5.2 A).

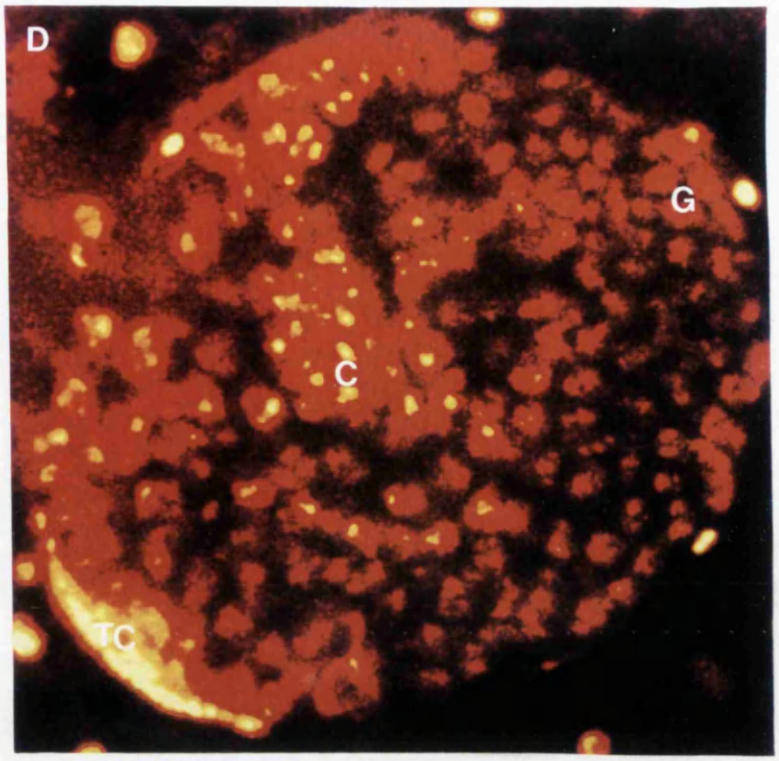
The pattern of hybridisation of pmsfK22 differs from that of pmsfK10. The expression of this transcript appears to be higher in the spermatogonia (G) than in the primary spermatocytes (C) (Figure 5.2 B). Levels of expression of this transcript appear to increase in older spermatocytes close to the terminal cells (TC). The testes in this case were probed with a cDNA corresponding to gK27 (pcK27). The expression of this transcript is highest in the primary spermatocytes (C). In common with gK10, expression is reduced in both spermatogonia (G) and terminal cells (TC) (Figure 5.2 C).

In order to ensure that the results I obtained were due to the male specific fragments, third instar larval testes were probed with positive and negative hybridisation controls. The positive hybridisation control was clone gS4 (also known as Mst40), which was isolated in the same screen and characterised by Russell (Russell & Kaiser, 1994). As described in chapter 1, this clone is expressed in the primary spermatocytes (C) and also in the terminal cells (TC). Additional staining is also observed in some cells in the fat body (FB) (Figure 5.2 D). The negative hybridisation control was the plasmid pBluescript (Stratagene) which was the vector used to subclone the male-specific fragments and cDNAs. In this case, no expression could be detected in any of the cells of the testes (Figure 5.2 E).


A
TC
C
G


B
G
C
TC


C
TC
C
G


D
G
C
TC

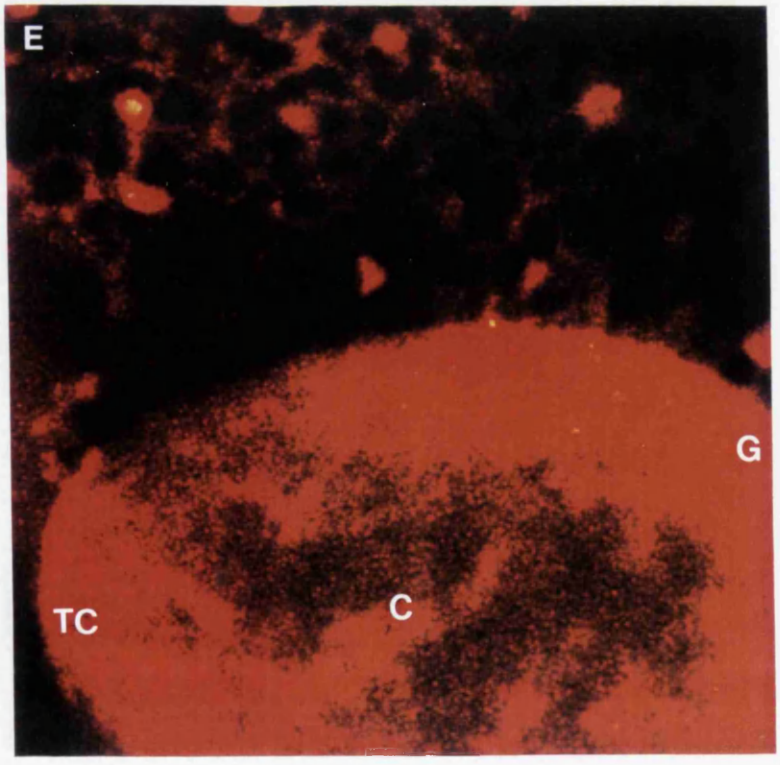

E
G
TC
C

Figure 5.2: *In situ* hybridisation to third instar larval testis.
Testis dissected from Oregon-R third instar larvae and hybridised with biotinylated probes as described in chapter Two section 2.10.3
Figure 5.2A: Hybridised with pmsfK10.
Figure 5.2B: Hybridised with pcK22.
Figure 5.2C: Hybridised with pmsfK27.
Figure 5.2D: Hybridised with gS4 (positive control).
Figure 5.2E: Hybridised with pBluescript (negative control).

5.3 *In Situ* Hybridisation To Adult Testes.

Since all the spermatogenic stages are present in the adult testes, it was of interest to see if transcripts could be detected in other spermatogenic stages. As discussed in chapter 1 and also in this chapter, it has been postulated that for most transcripts, expression is restricted to the primary spermatocyte stage until after spermiogenesis has occurred.

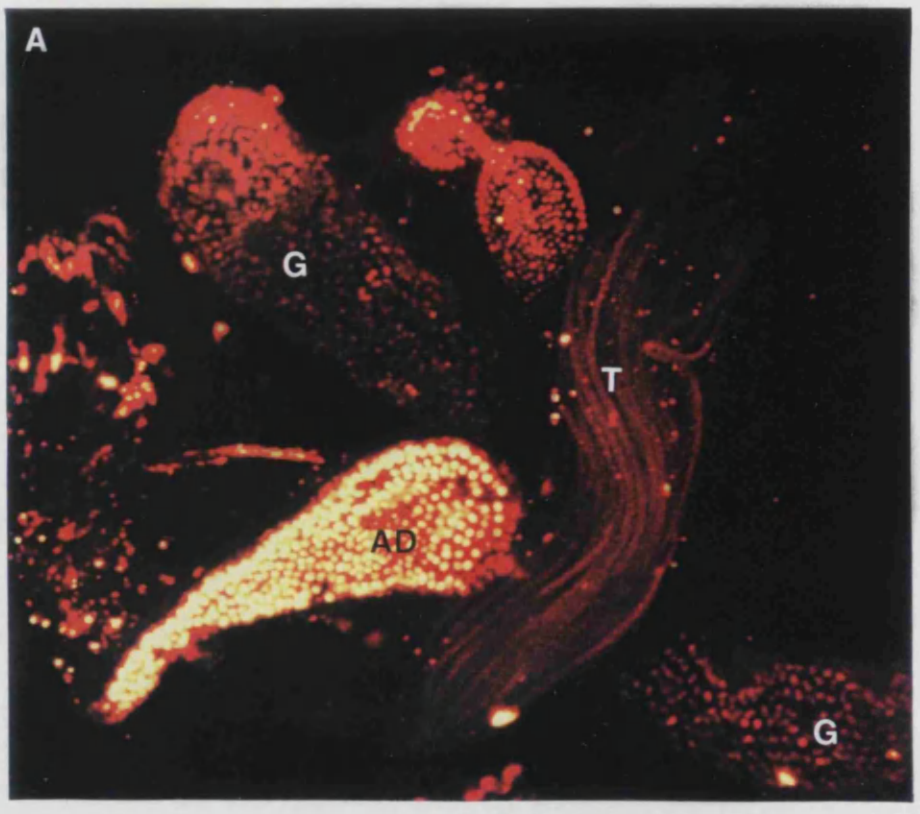
When hybridised with pcK27 (Figure 5.3 A), expression of this transcript was restricted to the seminal vesicle (AD) where the mature sperm are bundled prior to ejaculation. No expression was detected at the apex of the testes (T), nor in the testicular lumen after the gyre where post-meiotic elongating spermatids are found.

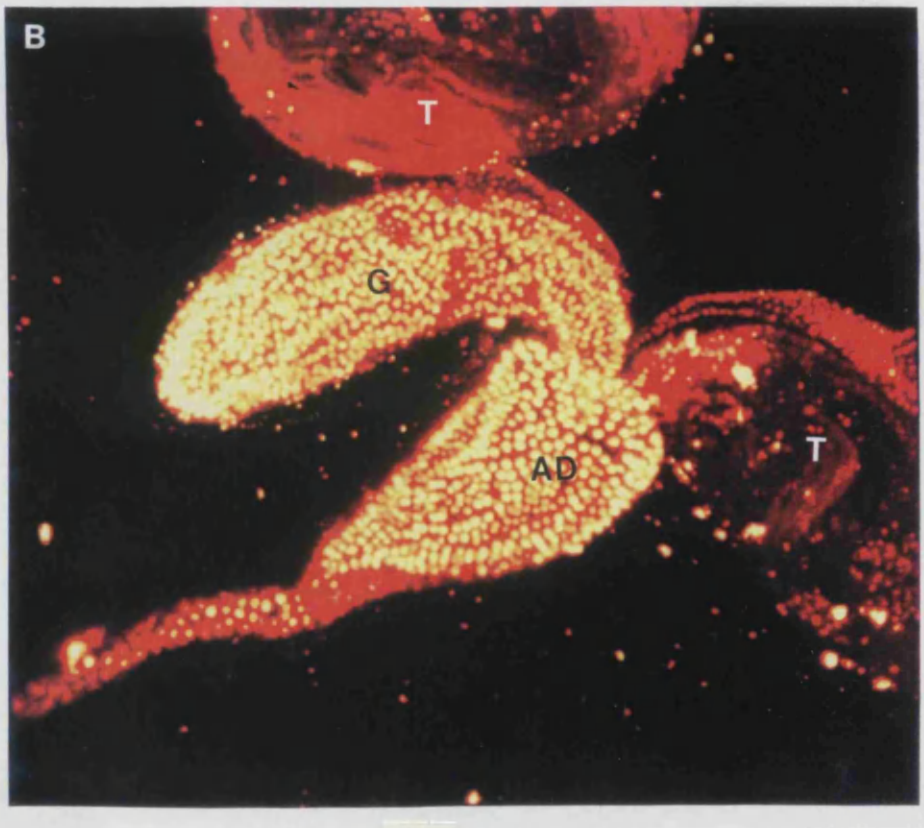
In the case of pcK31 (Figure 5.3 B), expression was observed both in the seminal vesicle (AD) and also in the accessory gland (G). The seminal vesicles as described previously are where the mature sperm are stored after individualisation and coiling. The accessory glands (G) are structures derived from the genital disc. This is where the peptides which comprise the seminal fluid are produced and stored. Again, no expression can be detected in the earlier stages of spermatogenesis.

5.4 Characterisation Of Enhancer-Trap Lines.

As discussed briefly in chapter 4, a recent enhancer-trap screen for spermatogenic loci (Castrillon *et al.*, 1993) produced P-element insertions near the chromosomal locations of two of the genomic clones. An insertion at 68A (the location of gK22) corresponds to a mutation Castrillon named *pistachio* (*pto*) and an insertion at 90E (the location of gK27) known as Ms(3)90E.

Figure 5.3: *in situ* hybridisation to Oregon-R adult testes.
Testes dissected out of Oregon-R males and hybridised with biotinylated probes.
Figure 5.3A: Hybridised with pcK27.
Figure 5.3B: Hybridised with pcK31.
testis (T), anterior ejaculatory duct (AD), accessory gland (G)


A
G
T
AD
G


B
T
G
AD
T

In order to investigate the relationship between these insertion lines and those genomic clones identified by this screen, genomic DNA was isolated from the wild type strain Ore-R and both of the enhancer trap lines. The resulting Southern blots were hybridised with pmsfK22 and pmsfK27 and also with the plasmid pBluescript (data not shown). The evidence from the Southerns is inconclusive but suggests that some genomic rearrangement may have occurred at these loci.

As has been described in chapter 3, enhancer-trap lines can be used to recover the flanking DNA 5' and 3' to the insertion site. The enhancer-trap line used in this screen was the P[Z] element (Mlodzik & Hiromi 1992). Genomic DNA was produced from each of the insertion stains, digested with *Xba*I and plasmid rescued as described in Section 2.6.3.2. In each case a 0.5 kb fragment was obtained. These fragments were subcloned into pBluescript (designated pb*pto* and pb*ms90*) and were used to probe chromosome squashes from the lines from which they were derived. Each of the clones hybridised to the same region of the polytene chromosome which corresponded to the original genomic clone. The results are shown in Figure 5.4. The polytene *in situ* hybridisations using the plasmid rescued fragments were kindly performed by Zong Sheng Wang.

5.5 Detection Of *Lac*Z Staining In Larval And Adult Testes.

The two insertion lines, *pistachio* (*pto*) and Ms(3)90E were generated in a P-element mutagenesis to isolate male-sterile mutations. The *pto* insertion produces a semi-sterile phenotype. The classification of semi-sterility was determined by fertility assays. Between 8-10 males homozygous for the insertion were crossed to a greater number of virgin females. If no progeny were produced, the insertion was designated sterile. Semi-sterile lines produced 10% fewer progeny than their heterozygous siblings. The insertion at 90E,

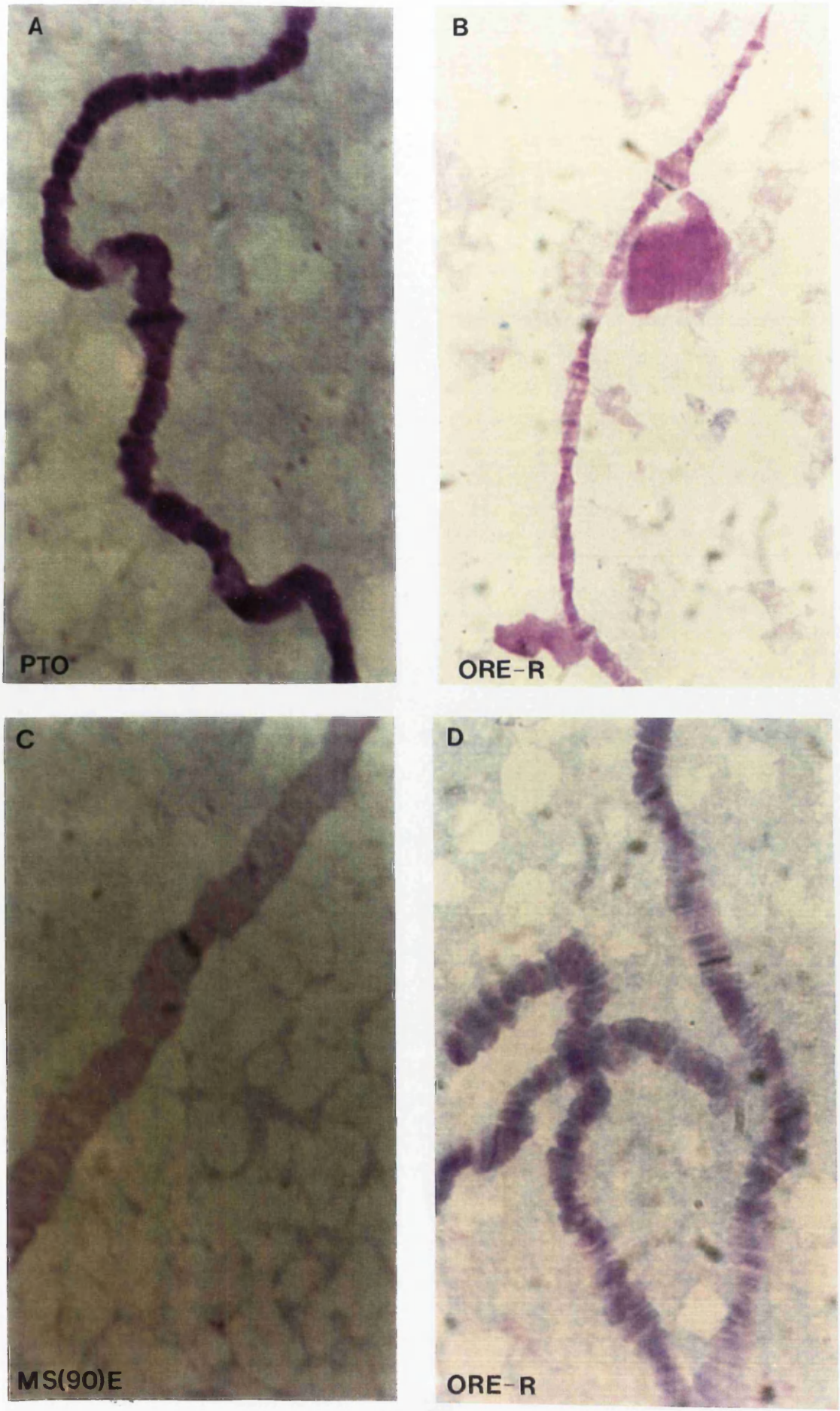
Figure 5.4: *in situ* hybridisation to polytene chromosomes using plasmid rescued fragments. DNA was plasmid recued from each of the enhancer trap lines. Each fragment was labelled with biotin and used to probe polytene chromosomes.

Figure 5.4A: *pto* insertion line polytene chromosome probed with pb*pto*.

Figure 5.4B: Oregon-R polytene chromosome probed with pmsfK22.

Figure 5.4C: Ms(3)90E insertion line polytene chromosome probed with pbms90.

Figure 5.4D: Oregon-R polytene chromosome probed with pmsfK27.1.


A
PTO
B
ORE-R
C
MS(90)E
D
ORE-R

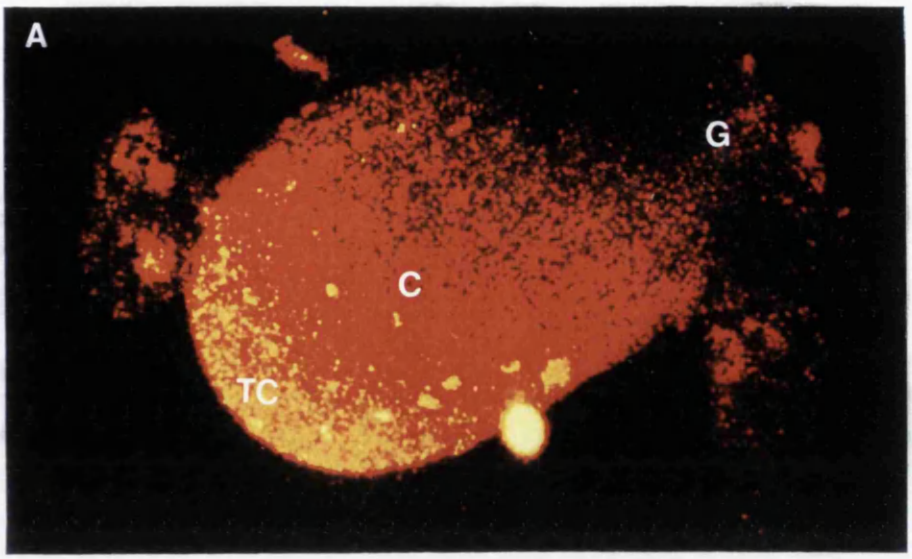
(Ms(3)90E) was designated as a semi-lethal mutation, homozygotes for the insertion were observed at less than 30% of the expected frequency. The determination of the staining pattern of the *lac*Z was performed as described in Section 2.13.2

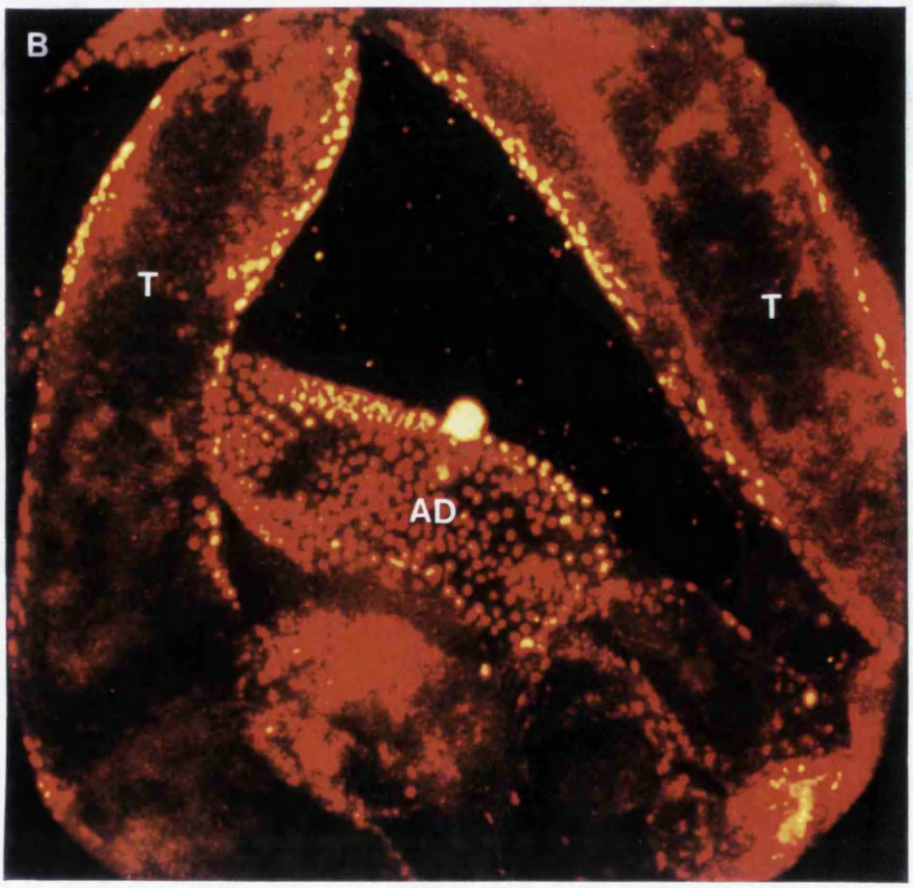
In both cases, no expression could be detected in the larval testes of homozygous or heterozygous insertion lines of Ms(3)90E (data not shown) or *pistachio* (*pto*), (Figure 5.5A).

Adult testis were dissected out of homozygous and heterozygous *pto* and Ms(3)90E flies and stained as before. In the case of *pto*, expression was restricted to the seminal vesicle in both homozygous (Figure 5.5 B) and heterozygous lines (data not shown). As described earlier (chapter 1, Section 5) the mature sperm at this stage are bundled prior to ejaculation. No *lac*Z staining in the testis of Ms(3)90E heterozygous flies was detected.

In addition, the testes dissected from both these lines exhibited gross structural abnormalities. In some instances, the apex of the testes formed bulbous structures, and the ejaculatory bulb was often enlarged. Testes which had one or both of these defects were often observed.

Figure 5.5: Detection of *lac*Z expression.
Adult and third instar larval testes were dissected from *pto* insertion line and *lac*Z expression detected as described in Chapter Two Section 2.13.2.
Figure 5.5A: larval testis dissected from *pto* flies.
Figure 5.5B: Adult testis dissected from *pto* flies.


A
G
C
TC


B
T
T
AD

Discussion.

In this chapter I attempted to confirm the restriction of the transcripts represented by the putative male-specific fragments using northern blotting and *in situ* analysis. Studies on the expression patterns of these male-specific transcripts suggests that some represent germline specific transcripts. A comparison of the transcripts in normal males and females shows that for gK10 and gK27, the transcripts are genuinely male-specific and not male-elevated (at the level of detection used). In addition, expression studies using the *tra* and *tud* mutations confirm the hypothesis that these represent germline specific transcripts. Attempts to reprobe the northern blots obtained for the gK10 and cK27 with gK22 and cK31 proved unsuccessful. From the results of the 'Reverse northern' in chapter 3 it can be predicted that these genomic fragments include male-specific sequences. However, it is impossible to infer that these may also represent transcripts whose expression is germline specific. The use of whole mount tissue *in situ* hybridisations should assist in determining whether there is a germline function for these clones. As has been described in chapter 1, the testis is one of the largest structures present in third instar larvae and at this stage transcription of gene products required for spermatogenesis has begun. Consequently, transcripts with a spermatogenic function are likely to be more prevalent.

Consistent with the observations of Olivieri & Olivieri and Gould-Somero & Holland, transcripts from the genomic DNA clones are detected during the primary spermatocyte stage (at least in the larval testis), the message is then detected in mature sperm for the two clones studied. However, as described in chapter 1, spermatocytes are also present within the adult testis (toward the apex) it is perhaps surprising that no signal was detected in this portion of the testis as well as in the mature sperm. This lack of signal in the primary

spermatocytes could be due to inadequate probe penetration. This part of the adult testis of *Drosophila* is enclosed within a thick sheath and several investigators increase incubation time with proteinase K to obtain a signal (Oliver *et al.*, 1993) In this case, the concentrations of proteinase K used appeared to be adequate to allow penetration of the probe without causing too much disruption to the morphology of the testis. The ability to obtain a signal with a probe using whole mount *in situ* hybridisation depends upon the ability of the probe to penetrate the tissue. If too low a concentration of proteinase K is used then no hybridisation will occur due to lack of probe penetration. Conversely, if too high a concentration of proteinase K is used then a signal may be obtained however the morphology of the tissue may be lost and therefore it will not be possible to determine exactly where the probe is hybridising. In this case, in order to ensure that the lack of staining in the apex of the testis was not due to insufficient probe penetration, a positive control such as *twine* (which hybridises to the apex of the testis) could be used (Alphey *et al.*, 1992). One of the clones, cK31 could not be detected in the whole mount larval testis. Since this technique had proven successful with the other clones (including gK22 which from Reverse northern data seems much less abundant) it seems likely that the explanation for the lack of signal resulted from problems with probe preparations rather than with the processing of the larval testis.

In the case of both gK10 and gK22, no signal could be obtained in the whole mount preparations of the adult testis. As discussed earlier, this might reflect a problem with probe penetration. The proteinase K used in the earlier experiments was titrated to determine the optimum time and concentration to ensure maximum penetration of the probe without loss of morphology. In the case of these two clones, a new batch of proteinase K was used. The concentrations and times used previously resulted in disintegration of the testis. No titration of the new batch of proteinase K was carried out. Others have

attempted to overcome the problem of probe penetration by the use of freeze substitution. In this procedure the dissected testis is sandwiched between two microscope slides. Slides are snap-frozen in liquid nitrogen and forced apart revealing the lumen of the testis (Lasko and Ashburner, 1990). This method was also tried but as there was insufficient time to optimise the conditions I decided to continue with the whole mount method which was giving more consistent results.

Using the approaches detailed in this chapter, I was able to demonstrate that the subcloned genomic fragments represent male-specific transcription units. Although not every technique was successful for each clone, sufficient information was obtained which confirmed the hypothesis that these are testis-specific clones which are likely to function during spermatogenesis.

Chapter Six Screening and Sequencing

Introduction.

As described in chapter 3, the function of a particular gene (represented by either a genomic or a cDNA clone) may be found by determining the sequence of that region of DNA. In this way, a comparison of the sequence obtained with others in various biological databases may give some insight into a possible biological function for that region if any homology is obtained. As with any approach, there are various factors that must be taken into consideration; for example the type of clone to be sequenced. Male-specific genomic fragments had already been identified which could have been sequenced directly. However, there are a number of benefits in obtaining cDNA clones corresponding to these genomic regions. The lack of introns (which in some cases may be a considerable size) in a cDNA clone means that any sequence obtained is more likely to represent coding sequence; this means that the complete sequence of the gene can be obtained much more efficiently. In addition, the lack of intronic sequences within a cDNA clone means that Open Reading Frames (ORF) can be assembled and predicted faster. Once the complete sequence of the cDNA has been obtained, this can be used to search databases of nucleic acids such as GenEmbl. If this search does not result in any significant homology, the sequence may be translated directly and ORF predicted. Those translations yielding long continuous ORFs may be used to search through a database of protein sequences such as SwissProt. Which of these searches will be successful is again dependent on several factors. If the sequence represents a gene that has previously been cloned in that organism, it is likely that a match will be obtained at the DNA level. Similarly, if the sequence is a new member of a gene family previously identified in that organism, it is likely that a DNA match will be obtained. If however the sequence has not been cloned, but has been isolated in another organism it is likely that homology will only be identified by the use of protein searches. The choice of which ORF

to search with can be assisted by the use of analysis programmes such as TESTCODE (GCG, Devereux *et al.*; 1984) which use codon preference data to predict the most likely ORF for a particular organism. The use of additional GCG programmes such as MOTIF assist in the prediction of a function for the clone by the identification of a common pattern within members of a gene family even if the whole sequence doesn't match, such as a conserved DNA binding domain (for example, a zinc finger) which may be the only common feature possessed by otherwise unrelated proteins.

An additional problem with direct sequencing of genomic fragments is that most of the fragments isolated by Reverse northern analysis were still large. In the case of gK31, the genomic fragment is around 12 kb. This fragment would, in theory, be large enough to contain more than one gene. Although the data from the Reverse northern suggests that this is not the case, it is still possible that an additional gene is present within this fragment that the Reverse northern is not sensitive enough to detect. Unless a DNA match is immediately identified the chances of finding any homology are reduced.

A drawback of differential screening as used in this study is that without mutations within a particular gene it is not possible to unequivocally assign a function to a fragment of cloned DNA. In *Drosophila*, there are two efficient mechanisms of obtaining mutations when cloned sequences are available whose functions are unknown. These are site-selected mutagenesis and 'local jumping'. Both of these techniques depend upon a P-element inserting in or near the gene of interest. ' local jumping' requires a P-element inserted near to the gene of interest. In the case of site-selected mutagenesis, sequence information on the gene to be mutated is required. Consequently sequencing a region of DNA may not only provide an idea of function but may also enable the introduction of P-element induced mutations into a particular locus.

Results.

6.1 Screening of cDNA Libraries.

For the reasons described, it was decided to screen libraries in an attempt to isolate the cDNAs corresponding to the male-specific genomic fragments. The four genomic clones initially identified by Russell (1989) are thought to have a role in spermatogenesis. Initially, a male larval cDNA library constructed by Russell (1989) in the lambda vector NM1149 was screened using each of the five genomic fragments as probes. During the primary screen 100,000 recombinant phage were plated out and screened as described in chapter 2. Four filters were lifted from the plate, these were screened with two clones. They were then stripped and reprobed with the remaining two clones.

Clone pmsfK10 was labelled with ^{32}P. The primary screen of the library identified approximately 50 strong positives which were present on both filters (data not shown). Although positives were identified during the subsequent secondary and tertiary screens, these hybridised with sequences in common with those in which the genomic fragments were cloned. This suggested that this library was contaminated with plasmid DNA. As these transcripts are expressed in the testis, I decided to screen a testis specific library (constructed in the vector lambda gt11 by Russell). Positives were identified which due to lack of time have not yet been resolved to single plaques. As described in chapter 5, pmsfK10 identified two abundant transcripts of 2.7 and 2.6 kb and a less highly abundant transcript of ~6-7 kb. Since the data from northern blots suggests that these are abundant transcripts it is perhaps surprising that cDNAs could not be obtained.

When ^{32}P labelled pmsfK22 was used as a probe, approximately 90 positives were identified, again during the subsequent screens these could not be purified to single plaques. This fragment was also used to screen the testis library. As before, positives have been obtained which have not been fully characterised. From the Reverse northern analysis, this clone appeared to be less highly expressed than the other three clones.

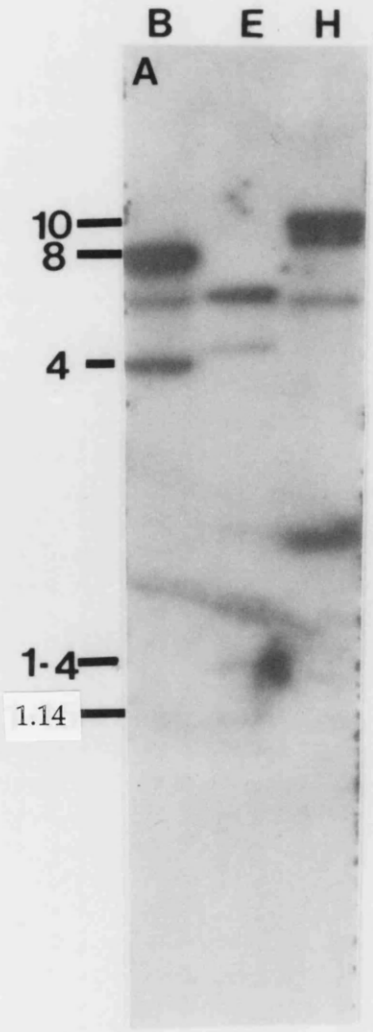
In the case of gK27, a mixture of pmsfK27.1 and pmsfK27.2 was used as a probe since there was evidence that these two plasmids cross-hybridised (data not shown). In addition, these two subclones produce a similar pattern of hybridisation on Southern blots of genomic DNA. In this case, approximately 90 positives were identified in the primary screen. Subsequent screens identified two positives, one corresponded to pBluescript, the other to sequences present in both pmsfK27.1 and pmsfK27.2 (data not shown). This 0.5 kb fragment was subcloned into pBluescript for use as a probe (see chapter 5) and as a template for sequencing. In order to ensure that this cDNA corresponded to the same region as both pmsfK27.1 and pmsfK27.2, a genomic Southern was probed with this fragment (designated pcK27) using genomic DNA derived from Oregon-R and restricted with *Bam*HI, *Eco*RI and *Hin*dIII as described in chapter 4. Using this fragment as a probe for the *Bam*HI digest a single band of 8 kb is observed (Figure 6.1A track 1), the *Eco*RI digest identifies fragments of 7, 1.4 and 1.15 kb (Figure 6.1A track 2). For the *Hin*dIII digest a band of 10 kb is seen (Figure 6.1A track 3). As described for pmsfK27.1 and pmsfK27.2 as well as the strongly hybridising bands described above, additional bands which hybridise less strongly to the probe were also observed. The cK27 cDNA clone identifies a 7 kb band in addition to those identified by the two genomic DNA fragments. The sizes of the other fragments identified by this probe corresponded to those identified when either pmsfK27.1 or pmsfK27.2 were used as a probe (see Figure 4.3). However, taken together, results suggest that this locus is even more

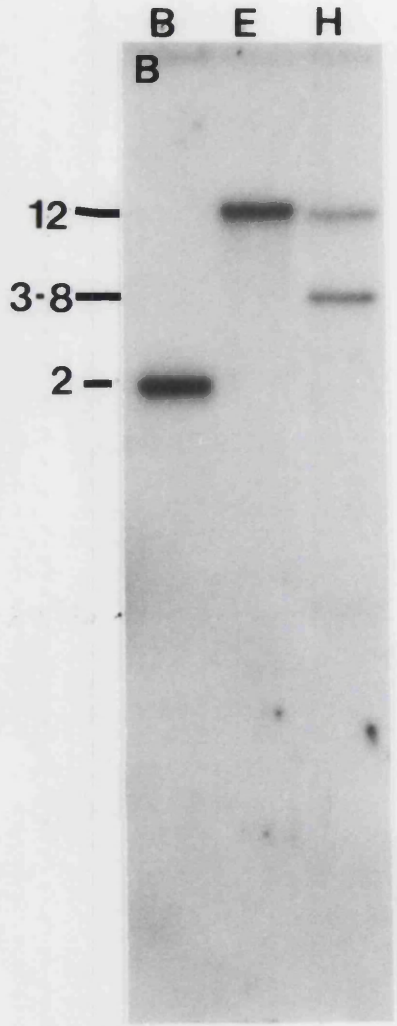
Figure 6.1 Genomic Southerns

Genomic DNA prepared from Oregon-R flies. Digested with *Bam*HI (Track 1, Figures 6.1 a & b), *Eco*RI (Track 2, Figures 6.1 a & b), *Hin*dIII (Track3, Figures 6.1 a &b).

Figure 6.1A shows the resulting Southern blot probed with cK27.

Figure 6.1B shows the resulting Southern blot probed with cK31.


B
E
H
A
10
8
4
1·4
1.14


B
E
H
B
12
3·8
2

complex than previously thought. A transcript of 2.2 kb had been detected on northern blots, suggesting that this clone was not full length. As I was screening the testis library to identify cDNA clones for pmsfK10 and pmsfK22, I decided to screen the library with pcK27 to isolate a full length clone.

When pmsfK31 was used as a probe approximately 140 positives were identified. Again a single positive was purified. A cDNA of 1.2 kb was identified which hybridised to the pmsfK31 probe. Since there is no northern data it is not known whether this represents a full length clone. This fragment was subcloned into pBluescript for use as a probe and for sequencing reactions and was designated pcK31. As with pcK27, genomic Southerns were probed with pcK31 and compared with that derived for pmsfK31 in chapter 4. Again Oregon-R DNA was restricted with *Bam*HI, *Eco*RI and *Hin*dIII. The probe identified a single fragment of 2kb in the *Bam*HI digest, (Figure 6.1B track 1). The digest using *Eco*RI shows a single band of 12 kb (Figure 6.1B track 2). For the *Hin*dIII digest, two fragments were identified of 12 and 3.8 kb (Figure 6.1B track 3). These fragments were also identified by pmsfK31 (see Figure 4.3,). In order to identify other cDNA clones corresponding to this region, the testis library was also screened with this cDNA.

6.2 Sequencing Of cDNA clones.

6.2.1 Sequencing of pcK27.

The fragment identified from the library screen was subcloned into the plasmid Bluescript. The T7 and T3 primers situated in the multiple cloning site were initially used to generate sequence. Subsequent primers were derived from

the sequence obtained from the T3 and T7 primers and are listed in table 2.2 . The complete sequence of the 481 bp cDNA obtained in the initial screen is shown in Figure 6.2. The cDNA also contains a good match with the consensus sequence for the *Drosophila* polyadenylation signal (AATAAA) separated by 43 nucleotides from a poly A tail consisting of 28 residues. When translated, the cDNA produced a long ORF in frame three producing a 113 amino acid polypeptide. This ORF was predicted using the codon preference facility of the IBI MacVector sequence analysis programme (Figure 6.3A). Using the program 'ProSearch' (Kolakowski *et al.*; 1992) it was shown that this cDNA showed considerable homology with a class of proteins known as the HMG-box containing proteins. These proteins show homology to the DNA binding domain (HMG-box) of the non-histone component of chromatin known as the High Mobility Group (HMG) of proteins. An alignment of cK27 with other HMG-box containing proteins is shown in Figure 6.4. cK27 seems to be a member of the class of proteins that are more similar to the HMG1/2 group rather than those that belong to the SRY/UBF transcription factor classes. A rooted dendrogram is shown in Figure 6.3B. This confirms the prediction that this protein is more similar to classical HMG1 proteins than to the HMG1 box containing transcription factors. Since cK27 appears to be most similar to the HMG1 group of proteins, a BESTFIT comparison was produced using the HMG1 box. This is shown in Figure 6.5. cK27 shows 29% identity and 47% similarity with the amino acid sequence of HMG1 box.

As the data from Northern blots suggested that this clone was not full length and the screening of the testis library had not yet yielded a longer cDNA. It was decided to sequence both the genomic fragments, pmsfgK27.1 and pmsfgK27.2. Data from the Southern blots described in Chapter Four (Figure 4.3) suggested that this clone, unlike the three other genomic clones, did not represent a single copy locus but consisted of a duplication. Since both K27

```
  1 AAT TCC ACC GGG CGC AAG AAC ATA AGA GCG GAG CAT CCC GAT TTT
        S   T   G   R   K   N   I   R   A   E   H   P   D   F

 46 AGT GTC CAA GAA GTG TCT GTG AAG GGC GGT GAG ATG TGG CGA GCC
    S   V   Q   E   V   S   V   K   G   G   E   M   W   R   A

 91 ATG GCC GAT GAG CAC AAG ATC GTG TGG CAG GAG TCG GCC AGC AAG
    M   A   D   E   H   K   I   V   W   Q   E   S   A   S   K

136 GCA ATG GCC GAG TAC AAG GAG AAG TTG GAG AAG TGG AAT GCC TTC
    A   M   A   E   Y   K   E   K   L   E   K   W   N   A   F

181 AAG GAG CAC CAG ACG GAG TCG TTT CCC CAT ATC TAT GAA GCT CCG
    K   E   H   Q   T   E   S   F   P   H   I   Y   E   A   P

226 TTG TCC TCT CGA TTC TCA AAA ACT AAC CAA AGG CCA ACC CTT TTT
    L   S   S   R   F   S   K   T   N   Q   R   P   T   L   F

271 GTG TAC GAC AGC AAG GAT GAA GCA ATG GCT CCG ATC TGC AGG ACG
    V   Y   D   S   K   D   E   A   M   A   P   I   C   R   T

316 TGC TTC TCA AAG GCC AAG TGC TTT CAC TAA
    C   F   S   K   A   K   C   F   H   *


346 ACGTTTGCTA AATTATTGTT TAGTCACAAA CAACCAGCAA AAGTTGCGAG


396 ATGTTGTTCA ATAAATTTAA ATATTTATAT TCAACTTGCT TACAACAAGT


446 GTGCCCAGAA AAAAAAAAAA AAAAAAAAAA AAAAAA
```

Figure 6.2 Complete sequence of the cDNA clone cK27.
The predicted ORF which starts at base 4 and ends at base 345 is shown. The concensus sequence for the polyadenylation signal is boxed. Vector derived sequence is underlined.

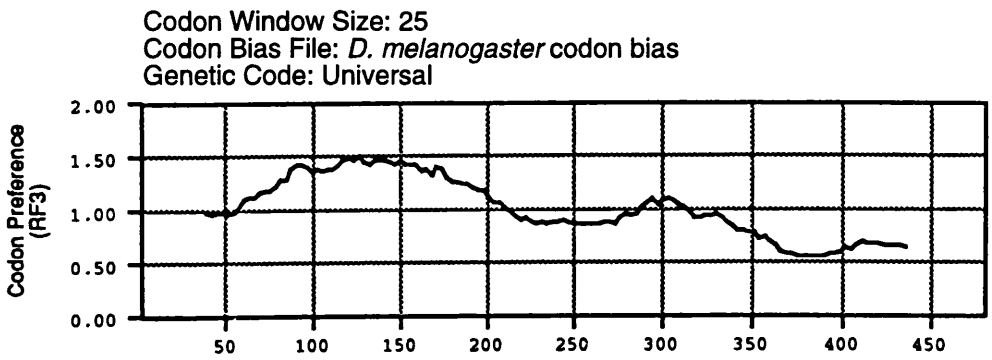


Figure 6.3a: Codon Preference of clone cK27
All three forward frames of the cDNA clone were examined with the MacVector sequence analysis software. Codon preference analysis using *Drosophila* codon bias gave the above prediction for frame three of the clone. Higher peaks (above 1.0) indicate likely coding regions.

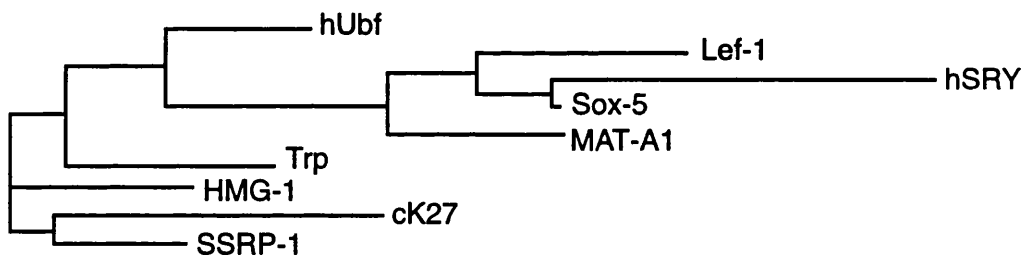


Figure 6.3b: Rooted dendrogram showing the relationship between cK27 and other HMG-1 box containing proteins (abbreviations of protein names as shown in Figure 6.4).

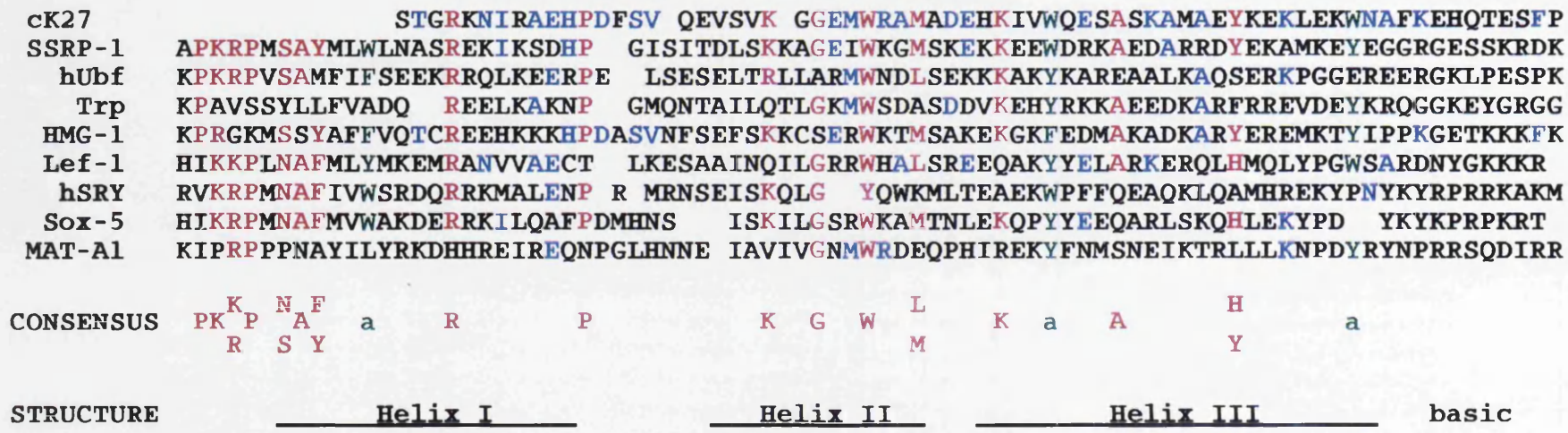

cK27 STGRKNIRAEHPDFSV QEVSVK GGEMWRAMADEHKIVWQESASKAMAEYKEKLEKWNAFKEHQTESFP
SSRP-1 APKRPMSAYMLWLNASREKIKSDHP GISITDLSKKAGEIWKGMSKEKKEEWDRKAEDARRDYEKAMKEYEGGRGESSKRDK
hUbf KPKRPVSAMFIFSEEKRRQLKEERPE LSESELTRLLARMWNDLSEKKKAKYKAREAALKAQSERKPGGEREERGKLPESPK
Trp KPAVSSYLLFVADQ REELKAKNP GMQNTAILQTLGKMWSDASDDVKEHYRKKAEEDKARFRREVDEYKRQGGKEYGRGG
HMG-1 KPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGETKKKFK
Lef-1 HIKKPLNAFMLYMKEMRANVVAECT LKESAAINQILGRRWHALSREEQAKYYELARKERQLHMQLYPGWSARDNYGKKKR
hSRY RVKRPMNAFIVWSRDQRRKMALENP R MRNSEISKQLG YQWKMLTEAEKWPFFQEAQKLQAMHREKYPNYKYRPRRKAKM
Sox-5 HIKRPMNAFMVWAKDERRKILQAFPDMHNS ISKILGSRWKAMTNLEKQPYYEEQARLSKQHLEKYPD YKYKPRPKRT
MAT-A1 KIPRPPPNAYILYRKDHHREIREQNPGLHNNE IAVIVGNMWRDEQPHIREKYFNMSNEIKTRLLLKNPDYRYNPRRSQDIRR
CONSENSUS PK P A a R P K G W L K a A H a
R S Y M Y
STRUCTURE Helix I Helix II Helix III basic

Figure 6.4 shows the alignment of cK27 amino acid sequence with various HMG-1 box containing proteins
The consensus sequence of the HMG box domain is shown below the alignment.
The predicted secondary structure is taken from the structural analysis of the B domain of HMG-1 (Weir *et al.*; 1993).

The proteins are as follows:-

SSRP-1: Protein involved In V(D)J recombination (Bruhn *et al.*; 1992).
hUBF : Human upstream binding factor (Jantzen *et al.*; 1990).
Trp : Trypanosome HMG-1 protein.
HMG-1 : Canonical HMG-1 box containing protein (Bianchi *et al.*; 1989).
Lef-1: Lymphoid enhancer binding factor-1 (Travis *et al.*; 1991).
hSRY: human testis determining factor (Sinclair *et al.*; 1990).
Sox-5 : Mouse SRY related protein expressed in the testis (Denny *et al.*; 1992).
Mat-A1 : Fungal mating type switch gene (Staben & Yanofsky 1990).

N: Identity to consensus sequence
N: Similarity to consensus sequence
N: Aromatic amino acids
N: Amino acids common to cK27 and other HMG-1 box proteins

```
BESTFIT    of: cK27   check: 9745  from: 1   to: 68
           to: HMG1   check: 4294  from: 1   to: 85


          Gap Weight:  3.000      Average Match:  0.540
      Length Weight:   0.100   Average Mismatch: -0.396

             Quality:   35.2            Length:      68
               Ratio:  0.542             Gaps:        2
 Percent Similarity: 47.692   Percent Identity: 29.231

 cK27 x HMG1

cK27  TGRKNIRAEHPDFSVQ..EVSVKGGEMWRAMADEHKIVWQESASKAMAEY.KEKLEKWNAFKEHQTES
      |.|.: : .||| ||.  |.| |.:|.|:.|....|  ::: |. ..| | :|.:..: : .. :|..
HMG1  TCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMMKTYIPPPKGETKK
```

Figure 6.5: BESTFIT comparison of the predicted translation of pcK27 with the HMG box of HMG1

subclones had been used as probes to identify this cDNA clone I decided to sequence both of these fragments in order to determine whether the complete sequence of the cK27 cDNA was contained within either of the two genomic regions. pmsfgK27.2 is identical with the 456 bases of cK27 that lie upstream of the poly A tail. A comparison of these two sequences is shown in Figure 6.6A. However, the sequence of neither pmsfgK27.1 nor pmsfgK27.2 provide any additional information regarding the sequence 5' of the start of the cK27 cDNA clone since both genomic fragments and the cDNA clone start at exactly the same sequence. The sequence of pmsfgK27.1 differed from that of both pmsfgK27.2 and cK27. Since the sequences derived from cK27 and pmsfgK27.2 are identical, the differences between the two genomic sequences are shown in Figure 6.6B. The two genomic fragments are relatively similar at the 5' end but diverge considerably after 175 bases.

6.2.2 Sequencing of cK31.

The cDNA clone derived from the region covered by pmsfgK31 in contrast, does not show homology with any sequences that are present in the databases. The sequence to date is shown in Figure 6.7 . At the 5' end (derived using the T7 primers) 416 base were obtained. At the 3' end 496 bases were obtained using the T3 primers (All primers are listed in Table 2.2). Approximately 300 bases remain to be sequenced. Although extensive analysis was carried out at the DNA and protein level, no significant homology was obtained. The program 'MOTIF' was used in order to attempt to identify any sequence motif that might give an indication of function. As before, the MacVector sequence analysis program and codon preference algorithms were used in an attempt to predict possible Open Reading Frames (ORFs). However in this instance no long ORFs could be identified. The lack of any homology may be solved by obtaining the remaining 300 bases of sequence and sequencing the cDNA clone on both strands.

```
  ck27 ATTCCACCGG GCGCAAGAAC ATAAGAGCGG AGCATCCCGA TTTTAGTGTC
gk27.2 ATTCCACCGG GCGCAAGAAC ATAAGAGCGG AGCATCCCGA TTTTAGTGTC

  ck27 CAAGAAGTGT CTGTGAAGGG CGGTGAGATG TGGCGAGCCA TGGCCGATGA
gk27.2 CAAGAAGTGT CTGTGAAGGG CGGTGAGATG TGGCGAGCCA TGGCCGATGA

  ck27 GCACAAGATC GTGTGGCAGG AGTCGGCCAG CAAGGCAATG GCCGAGTACA
gk27.2 GCACAAGATC GTGTGGCAGG AGTCGGCCAG CAAGGCAATG GCCGAGTACA

  ck27 AGGAGAAGTT GGAGAAGTGG AATGCCTTCA AAGGAGCACC AGACGGAGTC
gk27.2 AGGAGAAGTT GGAGAAGTGG AATGCCTTCA AAGGAGCACC AGACGGAGTC

  ck27 GTTTCCCCAT ATCTATGAAG CTCCGTTGTC CTCTCGATTC TCAAAAACTA
gk27.2 GTTTCCCCAT ATCTATGAAG CTCCGTTGTC CTCTCGATTC TCAAAAACTA

  ck27 ACCAAAGGCC AACCCTTTTT GTGTACGACA GCAAGGATGA AGCAATGGCT
gk27.2 ACCAAAGGCC AACCCTTTTT GTGTACGACA GCAAGGATGA AGCAATGGCT

  ck27 CCGATCTGCA GGACGTGCTT CTCAAAGGCC AAGTGCTTTC ACTAAACGTT
gk27.2 CCGATCTGCA GGACGTGCTT CTCAAAGGCC AAGTGCTTTC ACTAAACGTT

  ck27 TGCTAAATTA TTGTTTAGTC ACAAACAACC AGCAAAAGTT GCGAGATGTT
gk27.2 TGCTAAATTA TTGTTTAGTC ACAAACAACC AGCAAAAGTT GCGAGATGTT

  ck27 GTTCAATAAA TTTAAATATT TATATTCAAC TTGCTTACAA CAAGTGTGCC
gk27.2 GTTCAATAAA TTTAAATATT TATATTCAAC TTGCTTACAA CAAGTGTGCC

  ck27 CAGAAAAAAA AAAAAAAAAA AAAAAAAAAA A
gk27.2 CAGAAAtctt gttttttact ttcctttccc ttctctcttt cattccccac


gk27.2 caaaagccaa ttt
```

Figure 6.6 (a) Alignment of gK27.2 and cK27
This comparison shows that the two sequences are identical over their first 456 bases, diverging only at the point of addition of the poly A tail within the cDNA. Positions at which gK27.2 deviates from cK27 are shown in lower case.

```
gk27.2 GAATTCCACCGGGCGCAA.GAACATAA..GAGCGGAGCATCCCGATTTTA
gk27.1 GAATTCCACCGGGCGCAAcGA.CATAAaaaAGCGGAGCATCCCGATTTTA

gk27.2 GTGTCCAAGAAGTGTCTGTGAAGGGCGGTGAGATGTGGCGAGCCATGGCC
gk27.1 GTGTCCAAGAAGTGT.TGTGAAGGGCGGaGAGATGTGGC.AGCCATGGCC

gk27.2 GATGAGCACAAGATCGTGTGGCAGGAGTCGGCCAGCAAGGCAATGGCCGA
gk27.1 GATGAGgACAAGActGTGTGGCAGGAGTCGGCCA.CAAGGCAATGGCCGA

gk27.2 GTACAAGGAGAAGTTGGAGAAGTGGAATGCCTTCAAAGGAGCACCAGACG
gk27.1 GTACA.GGAGAAGTTGaAGgAGTGGAAaa.....
                                  GAP 2
gk27.2 GAGTCGTTTCCCCATATCTATGAAGC.....TCCGTTGTCCTCTCGATT.
                GAP 1
gk27.1 ......TTTCCCCAaggagcaccGCTTTTCGGACACACATGTATTTGTTC

gk27.2 CTCAAAAACTAACCAAAGGCCAACCCTTTTTGTGTACGACAGCAAGGATG
gk27.1 CTCAAAtACaAtCCAATGCCCCACCCTTTT.GTGTACGACACCAtGGATG

gk27.2 AAGCAATGGCTCCGATCTGCAGGACGTGCTTCTCAAAGGCCAAGTGCTTT
gk27.1 ActCgATGa

gk27.2 CACTAAACGTTTGCTAAATTATTGTTTAGTCACAAACAACCAGCAAAAGT


gk27.2 TGCGAGATGTTGTTCAATAAATTTAAATATTTATATTCAACTTGCTTACA


gk27.2 ACAAGTGTGCCCAGAAATCTTGT
```

Figure 6.6 (b) Alignment of gK27.2 and gK27.1
Sequences were aligned using IBI AssemblyLign Software. Positions at which gK27.1 deviates from gK27.2 are shown in lower case. Gap 1 (5bp) and Gap 2 (27bp) are indicated.

cK31t7 (5' end of clone)

```
  1 GGGAGCTGGA TAGGTCTCTA GCCTCCCAGT TTCGTGTTCG CAAAATTCCA
 51 TCTAAATTTC CGGCCCAAAC CTTTTGGAAA ATATATTTCG GTTAAAAATT
101 CAAAATTCCA GAGCTCAAAA ATGGCACTGA GTGTTAACCC TCCCGAGTAC
151 TGTTCCGTCC AATTTTTGGT GCGGTTAACA GCTGGTGCGA TTTGCATCSS
201 GTTGTGAAAA GAAGGGGCCG AGCCGGGCCC GAAGTGTTGA CCAAGTTCCC
251 CTGCGGCAAA CCCAATCTGC AGGCTCCGCC AAGAAGAAGA GGAAGATGGT
301 CAAGGCACAG TCCATGTGGC TGAACCCATT CTGTGATCCC GACGACACCG
351 CCTGTCCGTT CAATCCGCGC TTCGACGATA TCTACTACGT CGAGTCGGAC
401 AAGGCCAAGC GGAAGT
```

cK31t3 (3' end of clone, with PolyA tail)

```
  1 GCCGTGTGTG AAGCCCCGGA CGCCATATCC ATCTTTCTCC GAGTGCCGAC
 51 GTCTCAAGCC CGATGCCCCG CCACTGAAGG AGTGCAATTG TCTGGCGAAA
101 CCGCTGCTGT GTGAGATTTG GGCCGAGTTC CGACTACGAG CCATTGCCAA
151 GAAGTAGAGC CGTGGCCCTG ATTATATAAA GCGCTTTTAC GATTCTACAG
201 CATATACATT TTTTATTTGT ATGAGGAGAT TAAAATGTTT GCCCTGAGTG
251 CCAGCACGGA ATGCAGATGT GTGATCACTG AACGGATTCT TGGCCATATT
301 TGTCACGAGG CTCCAGAGCA GCGGATTGCA GTCACTATAG GTTTCGCTCG
351 CAGTCGGCAC CTCCACGTCC TGCGTCACGG GCATGTAAAT ACTTACTGGC
401 CACATTGTCC AGGGTGGGTC TGGAAATTCA TTTATAATTA CAATTCATAA
451 CTCAACCATT CCAAGTAAAA TGTCACATAC ACCAAAAAAA AAAAAA
```

Figure 6.6 shows the sequence of cK31 to date. The 3' end of the clone possesses a poly A tail of 13 residues. No consensus exists for a polyadenylation signal. The 5' end of this clone covers 496 bases whilst the 3' end covers 416 bases. Approximately 300 bases remain to be sequenced.

Discussion.

From northern analysis of clone pmsfgK10 both transcripts from the region covered by this genomic clone appeared to be fairly abundant, being detected in total RNA after an exposure of 16 hours. It was therefore surprising that I was unable to identify cDNA clones corresponding to these transcripts. In the case of gK22, there was no indication of abundance provided by northern analysis. However, data from the original Reverse northern analysis suggests that this transcript is present at lower levels than the other transcripts. When all of the five male-specific fragments were used as probes in library screens many positively hybridising plaques were identified. These positives often failed to rescreen.

The reason for the lack of cDNAs corresponding to two of the genomic fragments might have been due to the type of probes used. The whole plasmid (including the pBluescript vector) was used in making the probe, it is possible that those positives identified initially hybridised not to the insert DNA but to the pBluescript sequences. The subsequent screening of the testis cDNA library used probes that only contained the insert sequences. At the secondary and tertiary screening stage, an additional filter was lifted from each plate which was screened with pBluescript. This ensured that the positives plaques picked hybridised only to the insert sequences. The purification to homogeneity of those positives obtained from the testis library screen may assist in assigning function to these other genomic DNA clones.

Sequence of DNA derived from genomic clone gK27 begins at the same sequence as that of the cDNA clone cK27. The two fragments that were subcloned were the only male-specific regions identified by 'Reverse northern' analysis although it is possible that short regions of the male-specific transcript are contained within the flanking DNA which may be too short to be detected by the Reverse

northern. As northern blot analysis suggests that this clone is not full length and the open reading frame extends beyond the 500 base pair fragment identified by the library screen, the question arises as to the location of the 5' end of the cDNA. It is possible that the additional exons for the cDNA are contained outwith the original genomic DNA clone. The overlapping genomic DNA clone gK6 which was not analysed by the 'Reverse northern' may contain the additional sequences not found in gK27. In an attempt to identify the 5' region of this clone, 5' RACE (rapid amplification of cDNA ends) was also attempted. The initial results using this approach were not successful and insufficient time remained to optimise the conditions.

The data derived from chapter 4 in conjunction with the Southern and sequence analysis presented in this chapter confirm the complex nature of the genomic organisation in this region of the genome. It is apparent that the repeated HMG box represents a link between the genomic organisation and function of this transcript (several HMG box containing proteins have more than one HMG box, the sequences of which do not have to be identical. Landsman, D., & Bustin, M. (1993)). However, despite all the analysis on this region, its precise organisation is still unclear. A greater understanding of the nature of the complexity could be resolved by the characterisation of more genomic clones from this region (a suitable candidate, in the first instance, being the overlapping clone gK6). In addition, a longer cDNA clone containing 5' sequences not present in the previously characterised genomic fragments would be helpful in further analysis. Although the results from the Southerns were reproducible, further Southern analysis could be performed with probes derived from the fragments that flank the male-specific regions of the genomic clones. This will assist in determining the organisation of the locus.

The cDNA corresponding to gK27 identified a class of DNA binding proteins known as HMG-box containing proteins. These proteins all share a region of 80 amino acids known as the HMG-box and represent a diverse group of proteins implicated in a variety of cellular processes. For example, roles in the regulation of transcription, UBF (Zantzen *et al.*; 1990), and as tissue specific enhancer binding proteins such as LEF-1 and IRE-ABP (Travis *et al* .; 1991, Giese *et al.*; 1992). These proteins have also been implicated in the recombination events that take place during the formation of functional immunoglobulins (Shirakata *et al.*; 1991). In addition, several of these proteins have roles during the processes of sexual differentiation. The mating type switch genes of *S. pombe, N. crassa* and *P. anserina* (Kelly *et al.*; 1988; Staben & Yanofsky, 1990; Debuchey & Coppin, 1992) as well as the SRY gene (Sinclair *et al.*; 1990) contain HMG boxes. The number of HMG boxes contained by the various proteins differs, ranging from one in the case of the HMG1/2 proteins to as many as six in some forms of UBF. These proteins are known to bind specific structures identified in DNA molecules such as cruciform or kinked DNA (Bianchi *et al.*;1989), and also are able to bind to regions of single stranded DNA (Isackson *et al.*; 1979). An important attribute is their ability to bend DNA which may be important for function. It is conceivable that a HMG box containing protein functions by binding to specific regions of DNA causing them to bend, thus allowing other protein binding sites to be brought into closer proximity to each other. Two HMG box containing proteins, HMG-D and HMG-Z have been isolated in *Drosophila* (Ner *et al.*; 1993). Neither of these map to the cytological location of gK27.

One of the more interesting HMG-box containing proteins (for the purposes of this study) is represented by the Sox-5 gene (Denny *et al.*; 1992). This is an testis specific HMG-box containing protein isolated from the mouse using SRY as a probe. It is homologous to SRY only in the region of its HMG-box and

unlike SRY, is located on an autosome. The mRNA for this gene is expressed at elevated levels in the testis and the protein can be detected in post-meiotic germ cells. Restriction of this protein to the round spermatids (haploid cells that differentiate into mature sperm) in the mouse suggests a function for this protein. As discussed, an important feature of spermatogenesis is the transcription of gene products prior to spermiogenesis. Sox-5 may represent a testis specific transcription factor which acts to regulate the activity of spermatogenic transcripts. Further evidence to support this function has come from the identification of Sox-5 binding sites in the genomic regions which flank both rodent and human protamine genes. Protamines are histone like proteins which function in allowing a tighter packaging of the DNA in the sperm nucleus by virtue of their smaller size. As has been discussed, an important property of the HMG box proteins is their ability to bend DNA, facilitating activation of transcription by allowing remote DNA sites to be brought into close proximity allowing protein/protein interactions to occur.

Even although a full length cDNA was not obtained for cK27, a good idea of function was obtained from the small cDNA clone due to its homology with the HMG-Box family. It is therefore possible to assess more effectively which experiments to perform to further elucidate K27 function, for example identification of the DNA binding site of the K27 HMG-box. In the case of cK31, although more sequence was obtained for this clone, it was still not possible to assign a function. This further emphasises the importance of obtaining sequence, both to assist in determining function and when this does not prove possible, to use to isolate mutants.

Chapter Seven Discussion and Conclusions

Conclusions

This study describes the characterisation of four male-specific genomic clones identified *via* a differential screen for genes expressed specifically in the male. Apart from the obvious physical differences that are evident between males and females, there are also biochemical and behavioural differences which may arise as a result of differential gene expression during development. The most distinct differences are apparent in the processes of gametogenesis. As described in chapter 1, gametogenesis in the male (spermatogenesis) is an extremely complex process. Whilst the morphological changes that occur during the formation of a mature spermatozoon have been extensively documented, genes which act in the formation of specific structures or act to regulate these structural genes have not been readily identified. The number of genes which may function during spermatogenesis is thought to be large (perhaps as many as 17,500). This estimate also includes vital genes which when disrupted produce a spermatogenic defect. For example, most of the morphological changes observed during spermatogenesis are due to the action of microtubules and their associated proteins (Fuller, 1994). Not surprisingly, mutations in these genes often result in a spermatogenic defect, although these genes might not be considered 'true' spermatogenic loci. Consequently, mutations in 'true' spermatogenic loci are unlikely to result in a significantly different phenotype to those of housekeeping genes.

Advances in 'Reverse Genetic' techniques have provided a partial solution to this problem. A differential screening approach facilitates the identification of male-specific transcripts by highlighting the differences in gene expression during gametogenesis in the male and female. In addition, in this type of screen genes which have a housekeeping function, such as tubulins (mutation of which result in a spermatogenic phenotype) are not isolated. A similar

approach had already identified genes that were testis-specific (Schafer, 1986; DiBenedetto, 1987). On the basis of these results, a differential screen was performed by Russell and Kaiser in an attempt to identify male-specific transcripts expressed in third instar larvae.

As described in chapter 1, Russell characterised two clones (or groups of clones) out of 28 which may have a spermatogenic function. Consequently it was reasonable to expect that some of the other clones identified but not characterised by Russell would also have a role in spermatogenesis.

The four genomic clones characterised in this work seem to be good candidates for 'true' spermatogenic loci. Using Reverse northern analysis, all of the genomic clones studied contain (at this level of sensitivity) at least one fragment which is male-specific and not male-elevated (see figure 3.5).

Genomic clone gK10 identified a single male-specific fragment on a Reverse northern. From northern data it is evident that there are three male-specific transcripts originating from this genomic region. Evidence for a germline specific function for this clone is again provided by northern analysis. Studies from flies mutant in the processes of somatic sex determination (*tra*) and pole cell formation (*tud*) suggest that expression of these transcripts is restricted to the germline. Preliminary data from *in situ* hybridisation to larval testis confirms this finding. cDNA clones have been identified, but not fully characterised, for this region. It is not yet possible to infer a function for this gene. Once the fidelity of the cDNA clones has been confirmed and they have been sequenced it may be possible to assign function. In addition, it would be preferable to identify mutations at this locus, to determine whether spermatogenesis is affected. This could be achieved in several ways. The sequence data should provide information for the design of primers which could be used in site-

selected mutagenesis to obtain mutations in or near this locus. In addition, knowledge of the cytological location of this clone facilitates the identification of P-elements located in or near this region. Remobilisation of these elements may result in the generation of mutants at this locus.

Genomic clone gK22 also identifies a single male-specific fragment. As with gK10, *in situ* hybridisation suggests that this transcript is expressed in the larval testis. As described in chapter 5, autoradiographic studies (Olivieri & Olivieri 1965; Gould-Somero & Holland, 1974) suggest that most of the expression of genes required for the spermatogenic process takes place in the primary spermatocyte stage. However, in the mouse it appears that low levels of transcription of some genes takes place after this stage. The expression of gK22 however, is restricted to the spermatogonia and the older spermatocytes. In *Drosophila*, the processes of mitosis and meiosis have gene products in common, for example tubulin and *asp* (Causel *et al.*, 1990). As expression of gK22 is seen in primary spermatogonia which are about to enter mitosis and spermatocytes about to undergo meiosis, it is possible that the gene product of gK22 is a protein which is required for these processes in a similar function to that of *asp* (Causal *et al.*, 1990). Without northern analysis, it is not possible to say whether expression of this clone is restricted to the testis. cDNA clones have also been identified for the region corresponding to the genomic fragment however these have not been characterised or sequenced. A P-element enhancer trap insertion has been located in the region corresponding to the cytological location of gK22. This insertion gives rise to a male-sterile phenotype which fits with the hypothesis that this clone represents a spermatogenic locus.

The clone gK27 is the best characterised clone. In contrast to the other clones, this genomic region contains at least two male-specific fragments. A short cDNA clone corresponding to this locus has been identified. Northern analysis confirms that the transcript is restricted to the germline but implies that this

does not represent a full length cDNA. The sequence of this cDNA clone identifies a high degree of homology between this clone and genes encoding a group of DNA binding proteins. These proteins were originally identified as part of the non-histone component of chromatin and are known as High Mobility Group (HMG) proteins or HMG-box containing proteins (Landsman & Bustin, 1993). The members of this family contain a unique consensus sequence of regularly spaced aromatic amino acids known as the HMG-box. This is a diverse group of proteins whose members include the classical HMG-1 proteins, the sex determining region of Y (SRY) proteins (Sinclair *et al.*; 1990) and the transcription factor UBF (Jantzen *et al.*; 1990). Some HMG-box containing proteins are able to bind to non-B DNA without any requirement for sequence specificity (Hamada & Bustin, 1985; Bianchi *et al.*; 1989) , while others such as Sox-5 (Denny *et al.*; 1992) show sequence specificity.

Sox-5 is a testis specific HMG-box containing protein isolated from the mouse using SRY as a probe. The homology to SRY is confined to the region of its HMG-box. mRNA for this gene is expressed at elevated levels in the testis and the protein can be detected in post-meiotic germ cells. The restricted expression of Sox-5 to the round spermatids in the mouse may suggest a possible function for this protein. An important feature of spermatogenesis is the transcription of gene products prior to spermiogenesis. Therefore, Sox-5 may represent a testis specific transcription factor which acts to regulate the activity of spermatogenic transcripts. This is supported by the identification of Sox-5 binding sites in the genomic regions flanking both the rodent and human protamine genes. These are histone-like proteins which function to facilitate the tighter packaging of the DNA in the sperm nucleus. As discussed in chapter 6, an important property of the HMG-box proteins is their ability to bend DNA. Transcription may be activated by bringing into close proximity remote sites allowing protein/protein interactions to occur.

Two other HMG-box containing proteins have been identified in *Drosophila*, HMG-D and HMG-Z (Ner *et al.*; 1993) which are adjacent loci. Northern analysis shows that both genes encode two transcripts (0.65 and 1.2 kb for HMG-Z and 0.9 and 1.3 kb for HMG-D). HMG-Z is expressed in embryos and larval stages, while HMG-D is detected in embryos and is highly expressed in the adult female. No expression can be detected in the male (Ner et al.; 1993).

A role has recently been proposed for HMG-D (Ner & Travers, 1994). During the earliest stages of embryogenesis, development takes place in a syncytium (see chapter 1). The earliest nuclear divisions are very rapid, the nuclei are large and their chromosomes are undercondensed. As the nuclei migrate to the periphery of the embryo, the cycle times for nuclear division increase with a corresponding decrease in the nuclear volume. These early chromatin structures are found to be associated with HMG-D and not histone H1. Once the zygotic transcription of H1 occurs, the chromatin structures become more compact. This is correlated with the replacement of HMG-D by H1 suggesting that the organisation of the chromatin structures differs early in development.

As discussed earlier, HMG proteins are able to bend DNA. It is possible that during development HMG-D has a role in the formation of higher order chromatin structure. It may perform a similar function to the variants of H1 known as 'cleavage cycle 'histones found in *Xenopus* and the sea urchin. It has been postulated (Ner & Travers, 1994) that the less compact chromatin produced by binding HMG-D facilitates the rapid decondensation and condensation that is required during early development due to the short nuclear cycle times. The sequence of cK27 is more similar to the HMG-1 HMG-box than it is to either HMG-D or Sox-5. This suggests that it may function in a tissue-specific manner like HMG-1, with no sequence specificity, rather than as a sequence-specific, tissue-specific transcription factor like Sox-5. It is possible that in conjunction

with the histone/protamine switch there may be a switch from one type of HMG protein to another, again to assist in tight compaction of the chromatin. Since the cDNA is not full length it is not clear whether this protein contains more than one HMG-box. Once a full length cDNA has been isolated, it should be possible to determine whether this protein has a sequence specificity or whether it binds to abnormal DNA structures such as cruciforms and stem loop structures. This should assist in assigning a structural or regulatory function for this protein.

In an attempt to identify the 5' region of the HMG-box protein, the two genomic clones were sequenced. This analysis showed that one of the genomic clones (gK27.2) was identical to the cDNA (cK27) while the other (gK27.1) had diverged. In addition, towards the end of this study, 5' RACE was also attempted. Preliminary results with this technique were inconclusive and insufficient time remained to optimise the conditions to obtain the 5' end of the gene. *in situ* hybridisation of cK27 to both larval and adult testis shows that expression of this clone is restricted to the germline. In common with gK22, a P-element enhancer trap element was inserted near to the cytological location of this clone. Again, this insertion gave rise to a male-sterile phenotype.

Clone gK31 identifies a single male-specific fragment. A cDNA has been identified for this clone but as yet no significant homology has been identified. *in situ* hybridisation to the adult testis implies that the expression of this clone is restricted to the testis. As yet there is no evidence from northerns to confirm this. Again, the cytological location is known so mutants may be produced using 'local jumps' or site-selected mutagenesis using information derived from the sequence.

It is essential that the cDNAs corresponding to clones gK10, gK22 and cK27 are purified and sequenced. If the sequence identifies homology to known proteins in the database, then a possible function can be inferred (as for example in the case of cK27). If however the sequence of the cDNA does not identify any significant homology, the sequence can be used to design primers to use for site-selected or 'local jump' mutagenesis (Kaiser & Goodwin 1990 ; Littleton *et al.*, 1993). As has been described previously, the main disadvantage of a differential screening approach is that it only allows a prediction of function . Knowing the cytological location of a gene can also assist in obtaining mutations. As has been described in chapter 5, a P-element screen for genes involved in spermatogenesis produced insertions near two of the clones gK22 and gK27 (Castrillon *et al.*, 1993). Preliminary evidence suggests that these insertions may be within these clones. In the case of gK10 and gK31, knowledge of the cytological location of these clones means that fly lines containing mapped P-elements held at the fly stock centres can be examined to identify P-element insertions in or near these clones. The recent increase in the number of P-element insertion lines created makes this approach much more practical.

An additional advantage gained by the isolation of the cDNAs is the ability to produce fusion proteins which can be used to raise antibodies against that protein in the sperm. This has already been done successfully with the Mst(3)CPG group of proteins. In this case, no nucleotide or amino acid homologies have been identified using sequence analysis (Schafer *et al.*, 1993). However, the use of antibodies to the fusion protein indicated that it was one of the accessory proteins which coat the axoneme, suggesting a structural role for this gene.

Mechanisms by which the translation of some of the spermatogenic loci are controlled have been identified. A motif in the 5' untranslated region of both

the Mst(3)CPG gene family and the *jan*B locus known as the translational control element (TCE) has been identified (Kuhn *et al.*, 1988; Schafer *et al.*, 1993). This motif is thought to prevent inappropriate expression of gene products before they are required in spermiogenesis. An alternative control mechanism is that exhibited by the testis-specific β2-tubulin (Michiels *et al.*, 1989). In this case, a motif has been identified which appears to direct increased expression of the *lac*Z reporter gene in the adult testes in comparison to the larval gonads. Deletion of this region results in the same levels of expression of a reporter construct in both adult and larval testes. The identification of full length cDNA clones for these loci should assist in establishing whether there are additional control sequences or whether the two motifs described above represent the only control mechanisms utilised during spermatogenesis.

The identification of full length cDNAs for all of the male-specific fragments could also be used to rescue any mutations that are obtained in this region.

The use of a differential screening approach has been successful in identifying genes that are male-specific. At the levels of detection, all of the genomic DNA clones contained at least one male -specific fragment in the Reverse northern. This was confirmed for two of the clones by northern analysis. Further support for a male-specific function was provided by the whole mount *in situ* hybridisation, all of the male-specific fragments were expressed in either the larval gonad or adult testis. However, no other tissues were examined for expression. It is possible for the two clones gK22 and gK31 that their expression is not restricted to the germline and that there may be expression elsewhere in the fly. The use of third instar larval mRNA to make the probe has possibly biased the selection for those clones which are specific to the testis (i.e. those most likely to function during spermatogenesis). As described in chapter 1, the gonads are one of the largest structures in the male third instar larvae. In

addition, the gonad is largely composed of primary spermatocytes producing spermatogenic transcripts.

The aim of this study was to characterise a group of genomic DNA clones isolated by a differential screen to identify male-specific trancripts. Several different methods were used in this study. For all of the clones examined, results were obtained that suggested that these clones represent male-specific transcripts. Genomic clone gK27 represents the best characterised clone. In this case most of the objectives outlined in chapter 1 were achieved although identification of a full length cDNA clone and further sequence analysis would be required to complete the analysis of this genomic DNA clone . In the case of the other clones, a varying amount of success was achieved, cK31 requires northern analysis, larval whole mount *in situ* and the completion of the cDNA sequence (assuming that the cDNA clone identified was full length). The remaining two clones, cK10 and cK22, require slightly more analysis. In the case of both these clones, cDNAs have not yet been purified to homogeneity for sequence analysis. Northern analysis and whole mount *in situ* to adult testis are required for cK22, whilst gK10, requires an adult whole mount *in situ*.

Bibliography

Alphey, A., Jimenez, J., Cooper-White, H., Dawson, I., Nurse, P., & Glover, D. (1992). *twine*, a cdc25 homologue that functions in the male and female germline of *Drosophila*. Cell, **69**: 977-988.

Ashburner, M. (1972). In Results And Problems In Cell Differentiation, Vol 4, Developmental Studies On Giant Chromosomes. Beermann, W (Ed.) Springer-Verlag.

Baker, B. S., & Ridge, K. A. (1980). Sex And The Single Cell I: On The Action Of Major Loci Affecting Sex Determination In *Drosophila melanogaster*. Genetics, **94**: 383-423.

Baker, B. S.& Belote, J. M. (1983). Sex Determination and Dosage Compensation in *Drosophila melanogaster*. Ann. Rev. Genetics, **17**: 345-393.

Baker, B. S. (1989). Sex In Flies: The Splice Of Life. Nature, **340**: 521-524.

Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K., & Gehring, W. J. (1989). P-Element-Mediated Enhancer Detection, A Versatile Method To Study Development In *Drosophila*. Genes and Dev., **3**: 1288-1300.

Benton, W. D., & Davis, R. W. (1977). Screening Lambda gt Recombinant Clones By Hybridisation To Single Plaques *in situ*. Science, **196**: 180-182.

Bianchi, M. E., Falciola, L., Ferrari, S., & Lilley, D. M. J. (1992). The DNA Binding Site Of HMG1 Protein Is Composed Of Two Similar Segments (HMG boxes), Both Of Which Have Counterparts In Other Eukaryotic Regulatory Proteins. EMBO, **11**: 1055-1063.

Birnboim, H. C., & Doly, J. (1979). Rapid Alkaline Extraction Procedure For Screening Recombinant Plasmid DNA. NAR, **7**: 1513-1523.

Bodenstein, D. (1950). The Post Embryonic Development Of *Drosophila*. In The Biology of *Drosophila*. Demerec (Ed.). Hafner publishing.

Bonaccorsi, S., Pisano, C., Puoti, F., & Gatti, M. (1988). Y Chromosome Loops In *Drosophila melanogaster*. Genetics, **120**: 1015-1034.

Boswell, R. E., & Mahawold, A. P. (1985). *tudor*, A Gene Required For Assembly Of The Germ Plasm In *Drosophila melanogaster*. Cell, **43**: 97-104.

Bownes, M., & Nothiger, R. (1981). Sex Determining Genes And Vitellogenin Synthesis In *Drosophila melanogaster*. MGG., **182**: 222-228.

Bownes, M. (1990). Preferential Insertion Of P Elements Into Genes Expressed In The Germline Of *Drosophila melanogaster*. MGG., **222**: 475-460.

Bridges, M. A., Ercolani, L., Kong, X. F., & Nasrin, N. (1989). Identification Of A Core Motif That Is Recognised By Three Members Of The HMG Class Of Transcriptional Regulators: IRE-ABP, SRY, and TCF-1a. J. Cell. Bio, **48**: 129-135.

Brosseau, G.E. (1960). Genetic Analysis Of The Male-Fertility Factors On The Y-Chromosome Of *Drosophila melanogaster*. Genetics, **45**:257-274.

Bruhn, S.l., Pil, P. M., Essignmann, J.M.,Housemann, D.E.,& Lippard, J.S. (1992). Isolation And Charaterisation Of Human cDNA Clones Encoding A High Mobility Group Box Protein That Recognises Structural Distortions To DNA Caused By Binding Of The Anticancer Agent Cisplatin. PNAS, 89:2307-2311.

Bullock, W. O., Fernandez, J. M., & Short, J. M. (1987). XL1-Blue: A High Efficiency Plasmid *recA Escherichia coli* Transforming Strain With β Galactosidase Selection. Biotechniques, **5**: 376-379.

Castrillon, D., Gonczy, P., Alexander, S., Rawson, R., Eberhart, C. G., Viswanathan, S., DiNardo, S., & Wasserman, S. A. (1993). Toward A Molecular Genetic Analysis Of Spermatogenesis In *Drosophila melanogaster*: Characterisation Of Male-Sterile Mutants Generated By Single P Element Mutagenesis. Genetics, **135**: 489-505.

Causal, J., Gonzalez, C., Wandosell, F., Avila, J., & Ripoll, P. (1990). Abnormal Meiotic Spindles Cause A Cascade Of Defects During Spermatogenesis In *asp* Males Of *Drosophila*. Dev., **108**: 251-260.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979). Isolation Of Biologically Active RNA From Sources Enriched In Ribonuclease. Biochem., **24**: 5294-5299.

Chomczynski, P., & Sacchi, N. (1987). Single Step Method Of RNA Isolation By Acid Guanidinium Thiocyanate-Phenol-Chloroform Extraction. Anal. Biochem., **162**: 156-159.

Cline, T. W., (1978). Two closely linked mutations in *Drosophila melanogaster* that are lethal to the opposite sex and interact with *daughterless*. Genetics, **90**: 683-698.

Cline, T. W. (1993). The *Drosophila* Sex Determination Signal: How Do Flies Count To Two? TIG., **9**: 385-390.

Cooper, K. W. (1950). Normal Spermatogenesis In *Drosophila* In The Biology of *Drosophila*. Demerec (Ed.). Hafner publishing.

Courtot, C., Frankhauser, C., Simans,V., & Lehner, C., (1992). The *Drosophila cdc25* Homolog *twine* is Required For Meiosis. Development, **116**: 405-416.

Debuchy, R, & Coppin, E. (1992). The Mating Types Of *Podospora anserina* : Functional Analysis And Sequence Of The Fertilization Domains. MGG, **233**: 113-121.

Denny, P., Swift, S., Connor, F., & Ashworth, A. (1992). An SRY-Related Gene Expressed During Spermatogensis In The Mouse Encodes A Sequence-Specific DNA-Binding Protein. EMBO, **11**: 3705-3712.

Devereux, Haeberli & Smithies. (1984). A Comprehensive Set Of Sequence Analysis Programs For The Vax. NAR, **12**: 387-395.

Devlin, R. H., Bingham, B., & Wakimoto, B. T. (1983). The Organisation And Expression Of The *light* Gene, A Heterochromatic Gene Of *Drosophila melanogaster*. Genetics, 125: 129-140.

DiBenedetto, A. J., Lakich, D. M., Kruger, W. D., Belote, J. M., Baker, B. S., & Wolfner, M. F. (1987). Sequences Expressed Sex-Specifically In *Drosophila melanogaster* Adults. Dev. Biol., **119**: 242-251.

Eberl, D. F., Perkins, L. A., Engelstien, M., Hilliker, A. J., & Perrimon, N. (1992). Genetic And Developmental Analysis Of Polytene Section 17 Of The X Chromosome Of *Drosophila melanogaster*. Genetics, **130**: 569-583.

Ethan, B., Vaessin, H., Shepherds, S., Lee, K., McCall, K., Barbels, S., Ackerman, L., Carretto, R., Uemura, T., Grell, E., Jan, L. Y., & Jan, Y. N. (1989). Searching For Pattern And Mutation In The *Drosophila* Genome With P-LacZ Vectors. Genes and Dev., **3**: 1273-1287.

Fackenthal, J. D., Turner, F. R., & Raff, E. C. (1983). Tissue-Specific Microtubule Functions In *Drosophila* Spermatogenesis Require The β2-Tubulin Isotype-Specific Carboxy Terminus. Dev. Biol., **158**: 213-227.

Feinberg, A. P., & Vogelstein, B. (1983). A Technique For Radiolabelling DNA Restriction Fragments To High Specific Activity. Anal. Biochem., **132**: 6-13.

Frischauf, A., Lehrarch, H., Poustksa, A., & Murray, N. (1983). Lambda Replacement Vectors Carrying Polylinker Sequences. J. Mol. Biol., **170**: 827-842.

Fryxell, K. J., & Meyrowitz, E. M. (1987). An Opsin Gene That Is Expressed Only In The R7 Photoreceptor Of *Drosophila*. EMBO, **6**: 443-451.

Fuller, M. T. (1986). Genetic Analysis Of Spermatogenesis In *Drosophila*: The Role Of The Testis Specific β-Tubulin And Interacting Genes In Cellular Morphogenesis. Gall (Ed). Sym. Soc. Dev. Biol., **44**: 19-41.

Fuller, M. T. (1994). Spermatogenesis. The Development of *Drosophila*. Bate, M., & Martinez-Arias, A (Eds.). Cold Spring Harbour Press.

Gailey, D. A., & Hall, J. C. (1989). Behaviour And Cytogenetics Of Fruitless In *Drosophila melanogaster*: Different Courtship Defects Caused By Separate Closely Linked Lesions. Genetics, **121**: 773-785.

Gepner, J., & Hays, T. (1993). A Fertility Region On The Y Chromosome Of *Drosophila melanogaster* Encodes A Dynein Microtubule Motor. PNAS., **90**: 11132-11136.

Gibbons, I. R. (1981). Cilia And Flagella Of Eukaryotes. J. Cell Biol., **91**: 107-124.

Giese, K., Cox, J., & Grosschendl, R. (1992). The HMG Domain Of Lymphoid Enhancer Factor 1 Bends DNA And Facilitates Assembly Of Functional Nucleoprotein Structures. Cell, **69**: 185-195.

Goldstein, L. S. B., Hardy, R. W., & Lindsley, D. L. (1982). Structural Genes On The Y Chromosome Of *Drosophila melanogaster*. PNAS., **79**: 7405-7409.

Gonczy, P., Viswanathan, S., & DiNardo, S. (1992). Probing Spermatogenesis In *Drosophila* With P-Element Enhancer Detectors. Development, **114**: 89-98.

Gonczy, P., Thomas, B. J., & DiNardo, S. (1994). Roughex Is A Dose-Dependent Regulator Of The 2nd Meiotic Division During *Drosophila* Spermatogenesis. Cell **77**: 1015-1025.

Gonzalez, C. Casel, J., & Ripoll, P. (1989). Relationship Between Chromosome Content And Nuclear Diameter In Early Spermatids Of *Drosophila melanogaster*. Genet. Res. Camb., **54**: 205-212.

Gonzalez, C, Saunders, R. D. C., Casel, J., Molina, I., Carmena, C., Ripoll, P.,& Glover, D. M. (1990). Mutations At The *asp* Locus of *Drosophila* Lead To Multiple Free Centrosomes In Syncytial Embryos, But Restrict Centrosome Duplication In Larval Neuroblasts. J. Cell Sci., **89**:605-616.

Gould-Somero, M., & Holland, L. (1974). The Timing Of RNA Synthesis For Spermiogenesis In Organ Cultures Of *Drosophila melanogaster* Testis. Wilhelm Roux' s Archives, **174**: 133-148.

Grosschedl, R., Giese, K., & Pagel, J. (1994). HMG Domain Proteins: Architectural Elements In The Assembly Of Nucleoprotein Structures. TIG, **10**, 94-100.

Hackstein, J. H. P. (1987). Spermatogenesis In *Drosophila*. Hennig (Ed.) Springer-Verlag.

Hamilton, B. A., Palazzolo, M. T., Chang, J. H., Vijayraghavan, K., Mayeda, C. A., Whiteney, M. A., & Meyerowitz, E. M. (1991). Large-Scale Screen For Transposon Insertions Into Cloned Genes. PNAS., **88**: 2731-2735.

Hardy, R. W., Tokuyasu, K. T., & Lindsley, D. L. (1981). Analysis Of Spermatogenesis In *Drosophila melanogaster* Bearing Deletions For Y Chromosome Fertility Genes. Chromosoma, **83**: 593-617.

Hardy, R. W., Lindsley, D. L., Livak, K. J., Lewis, B., Siversten, A., Joslyn, G., Edwards, J., & Bonaccorsi, S. (1984). Cytogenetic Analysis Of A Segment Of The Y Chromosome Of *Drosophila melanogaster*. Genetics, **107**: 591-610.

Hartenstein, V. (1993). Atlas Of *Drosophila* Development. Supplement To The Development Of *Drosophila*. Bate, M., & Martinez-Arias, A. (Eds.). Cold Spring Harbour Press.

Hennig, W. (1985). Y Chromosome Function And Spermatogenesis In *Drosophila hydei*. Adv. Genet., **23**: 179-234.

Hennig, W. (1987). The Y Chromosomal Lampbrush Loops Of *Drosophila.* Hennig (Ed.). Springer-Verlag.

Hess, D. (1973). Local Structural Variations In The Y-Chromosome Of *Drosophila hydei* And Their Correlation To Genetic Activity. Cold Spring Harbour Symp. Quant. Biol., **38**: 663-672.

Hildreth, P. E. (1965). *doublesex* A Recessive Gene That Transforms Males And Females Of *Drosophila* Into Intersexes. Genetics, **51**: 659-678.

Homyk, T., & Sheppard, D. E. (1977). Behavioural Mutants Of *Drosophila melanogaster.* Isolation And Mapping Of Mutants Which Decrease Flight Ability. Genetics, **87**: 95-104.

Huijser, P., Kirchhoff, C., Lankenau, D.-K., & Hennig, W. (1988). Retrotransposon-Like Sequences Expressed In Y Chromosomal Lampbrush Loops Of *Drosophila hydei.* J. Mol. Biol, 204.

Hung, M. C., & Wensink, P. C. (1981). The Sequence Of The *Drosophila melanogaster* Gene For Yolk Protein 1. NAR, **9**: 6407-6419.

Isackson, P. J., Fishback, J. L., Bidney, D. L., & Reeck, G. R. (1979). Preferential Affinity Of High Mobility Group Non-Histone Chromatin Proteins For Single Stranded DNA. J. Bio. Chem., **254**: 5569-5572.

Jantzen, H. M., Admon, A., & Bell, S. P. (1990). Nucleolar Transcription Factor hUBF Contains A DNA-Binding Motif With Homology To HMG Proteins. Nature, **344**: 830-836.

Jones, W. K., and Rawl, J. M. (1988). Genetic And Molecular Mapping Of Chromosome Region 85A In *Drosophila melanogaster*. Genetics, **120**: 733-742.

Judd, B. H., Shen, M. W., & Kaufman, T. C. (1972). The Anatomy And Function Of A Segment Of The X Chromosome Of *Drosophila melanogaster*. Genetics, **71**: 139-156.

Kaiser, K., & Goodwin, S. F. (1990). "Site-Selected" Transposon Mutagenesis Of *Drosophila*. PNAS., **87**: 1686-1690.

Kaiser, K., & Murray, N. (1985). The Use Of Lambda Replacement Vectors In The Construction Of Replacement Genomic DNA Libraries. DNA cloning volume 1. Glover (Ed.).

Kalderon, D., & Rubin, G. M. (1988). Isolation And Characterisation Of *Drosophila* cAMP-Dependent Kinase Genes. Genes and Dev., **2**: 1539-1556.

Kelly, M., Burke, J., Smith, M., Klar, A., & Beach, D., (1988). Four Mating-Type Genes Control Sexual Differentiation In The Fission Yeast. EMBO, **7**: 1537-1547.

Kempe, E., Muhs, B., & Schaffer, M. (1994). Gene Regulation In *Drosophila* Spermatogensis: Analysis Of Protein Binding At The Translational Control Element TCE. Dev. Gen., **14**: 449-459.

Kemphues, K. J., Kaufman, T. C., Raff, R. A., & Raff, E. C. (1982). The Testis-Specific β-Tubulin Subunit Has Multiple Functions In Spermatogenesis. Cell, **31**: 655-670.

Kennison, J. (1983). Analysis Of Y-Linked Mutations To Male-Sterility In *Drosophila melanogaster*. Genetics, **103**: 219-234.

Kennison, J. (1991). The Genetic And Cytological Organisation Of The Y Chromosome Of *Drosophila melanogaster*. Genetics, **98**: 529-548.

Kennison, J. (1993). Analysis Of Y-Linked Mutations To Male Sterility In *Drosophila melanogaster*. Genetics, **103**: 219-234.

Kiefer, B. I. (1966). Ultrastructural Abnormalities In Developing Sperm Of X/O *Drosophila melanogaster*. Genetics, **54**: 1441-1452.

Kiefer, B. I. (1970). Development, Organisation, And Degeneration Of The *Drosophila* Sperm Flagellum. J. Cell. Sci., **6**: 177-194.

Kolakowski, L. F., Leunissen, J. A. M., & Smith, J. E. (1992). ProSearch : Fast Searching Of Protein Sequences With Regular Expression Patterns Related To Protein Structure And Function. Biotechniques, **13** : 919-921.

Koopman, R., Gubbay, J., Vivian, N., Goodfellow, P., & Lovell-Badge, R. (1991). Male Development Of Chromosomally Female Mice Transgenic For SRY. Nature, **35**: 117-121.

Kozlova, T. Y., Semeshin, V. F., Tretykova, I. V., Kokoza, E. B., Pirrotta, V., Grafodatskya, V. E., Belyaeva, E. S., & Zhimulev, I. F. (1994). Molecular And Genetic Characterisation Of The 10A1-2 Band And Ajoining Region In The *Drosophila melanogaster* Polytene X Chromosome. Genetics, **136**: 1063-1073.

Kuhn, R., Schafer, U., & Schafer, M. (1988). Cis-Acting Regions Sufficient For Spermatocyte-Specific Transcriptional And Spermatid-Specific Translational Control Of The *Drosophila melanogaster* Gene Mst(3)gl9. EMBO, **7**: 447-454.

Landsman, D., & Bustin, M. (1993). A Signature For The HMG-1 Box DNA-Binding Proteins. Bioessays, **15**: 539-546.

Lasko, P. F., & Ashburner, M. (1990). Posterior Localisation Of *vasa* Protein Correlates With, But Is Not Sufficient, For Pole Cell Development. Genes and Dev., **4**: 905-921.

Lefevre. G.Jr. (1976). In The Genetics and Biology of Drosophila. Ashburner & Novitski (Eds.). Academic Press.

Lehrach, H., & Frischauf, A. M. (1992). EMBL lab manual.

Lifschytz, E., & Meyer, G. F. (1977). Characterisation Of Male Meiotic-Sterile Mutations In *Drosophila melanogaster*. The Genetic Control Of Meiotic Divisions And Gametogenesis. Chromosoma, **64**: 371-392.

Lifschytz, E. (1987). The Developmental Programme Of Spermiogenesis In *Drosophila* : A Genetic Analysis. Int. Rev. Cytol, **109**: 211-258.

Lindsley D. L., & Grell E. H. (1968). Genetic Variations Of *Drosophila melanogaster.* Carnegie Insitution, Washington.

Lindsley, D. L., & Lifschytz, E. (1972). The Genetic Control Of Spermiogenesis In *Drosophila*. In Edinburgh Symposium On The Genetics Of The Spermatozoon. R. A. Betty & S. Gluecksohn-Waelsch (Eds.). Edinburgh. Bogtrykkeriet Forum.

Lindsley, D. L., & Tokuyasu, K. T. (1980). Spermatogenesis. In The Genetics And Biology Of *Drosophila*. Ashburner & Wright (Eds.). Academic Press.

Lindsley, D., & Zimm, G. G. (1992). The Genome Of *Drosophila melanogaster*. Academic press.

Littleton, J. T. Stern, M., Schuize, K., Perin, M., & Bellen H. J. (1993). Mutational Analysis Of *Drosophila* Synaptotagmin Demonstrates Its Essential Role In Ca2++ Activated Neurotransmitter Release. Cell, **74**: 1125-1134.

Livak, K. J. (1990). Detailed Structure Of The *Drosophila melanogaster Stellate* Genes And Their Transcripts. Genetics, **124**: 303-316.

Matthews, K. A., Miller, D. F. B., & Kaufman, T. C. (1990). Functional Implications Of The Unusual Spatial Distribution Of A Minor A Tubulin Isotype: A Common Thread Among Chordontonal Ligaments, Developing Muscle And Testes Cyst Cells. Dev. Biol., **137**: 45-61.

McKee, B. D., Habera, L., & Vrana, J. A. (1992). Evidence That Intergenic Spacer Repeats Of *Drosophila melanogaster* rRNA Genes Function As X-Y Pairing Sites In Male Meiosis, And A General Model For Achiasmatic Pairing. Genetics, **132**: 529-544.

Meyer, G. F. (1963). Die Funktionsstrukkturen Des Y-Chromosoms In Den Spermatocytenkernan Von *Drosophila hydei, D. neohydei, D. repleta* Und Einigen Anderan *Drosophila*-Arten. Chromosoma, **14**: 207-255.

Meyer, G. F. (1972). Influence Of The Y Chromosome On Fertility And Phenotype Of *Drosophila* Spermatozoa. In Edinburgh Symposium On The Genetics Of The Spermatozoon. R. A. Betty & S. Gluecksohn-Waelsch (Eds.). Bogtrykkeriet Forum.

Michiels, F., Gasch, A., Kaltschmidt, B., & Renkawitz-Pohl, R. (1989). A 146 bp Promoter Element Directs The Testis Specificity Of The *Drosophila* β2-Tubulin Gene. EMBO, **8**: 1559-1565.

Milligan, C. D., & Kaiser, K. (1993). 'Site-Selected' Mutagenesis Of A *Drosophila* Gene Using The I Factor Retrotransposon. NAR., **21**: 1323-1324.

Mlodzik, M. & Hiromi, Y. (1992). Enhancer Trap Method In *Drosophila*: Its Application to Neurobiology. Methods in Neurosciences **9**: 397-414.

Murray, N. E. (1983). Phage Lambda And Molecular Cloning. In Lambda II. R. W. Hendrix, J. W. Roberts, F. W. Stahl, & R. A. Weisberg (Eds.), Cold Spring Harbour Press.

Neesen, J., Bunemann, H., & Heinlein, U. A. O. (1994). The *Drosophila hydei* Gene Dhmst101(1) Encodes A Testis-Specific, Repetitive, Axoneme-Associated Protein With Differential Abundance In Y Chromosomal Deletion Mutant Flies. Dev. Biol, **162**: 414-425.

Ner, S. S., Churchill, M. E. A., Searles, M. A., & Travers, A. A. (1993). dHMG-Z, A Second HMG-1-Related Protein In *Drosophila melanogaster*. NAR, **21**: 4369-4371.

Nothiger, R., Dudendorfer, A., & Epper, F. (1977). Gyandromorphs Revel Two Separate Primordia For Male And Female Genitalia In *Drosophila melanogaster.* Wilhelm Rouxs Arch, **181**: 367-373.

Nusslein-Vollard, C., & Wieschaus, E. (1980). Mutations Affecting Segment Number And Polarity In *Drosophila.* Nature, **287**: 795-801.

O' Connell, P. O., & Rosbash, M. (1984). Sequence, Structure, And Codon Preference Of The *Drosophila* Ribososmal Protein 49 Gene. NAR, **12**: 5495-5513.

O'Kane, C. J., & Gehring, W. J. (1987). Detection *in situ* Of Genomic Regulatory Elements In *Drosophila.* PNAS., **84**: 9123-9127.

Oliver, B., Kim, Y.-J., & Baker, B. S. (1993). *Sex-lethal*, Master And Slave: A Hierarchy Of Germline-Sex Determination In *Drosophila.* Development, **119**: 897-908.

Olivieri, G., & Olivieri, A. (1965). Autoradiographic Study Of Nucleic Acid Synthesis During Spermatogenesis In *Drosophila melanogaster* Mutation Res. **2**: 366-380.

Palazzo, M. J., Hyde, D. R., Vijay-Raghavan, K., Mecklenlorgk, K., Benzer, S., & Meryerowitz (1989). Use Of A New Strategy To Isolate And Characterise 436 *Drosophila* cDNA Clones Corresponding To RNA's Expressed In The Head But Not In The Early Embryo. Neuron, **3**: 527-539.

Pardue, M. L. (1986). *In Situ* Hybridisation To DNA Of Chromosomes And Nuclei. In *Drosophila*: A Practical Approach. D. B. Roberts (Ed.) IRL Press.

Perucho, M., Hanahan, D., & Wigler, M. (1980). Isolation Of The Chicken Thymidine Kinase Gene By Plasmid Rescue. Nature, : 207-210

Rasmusson, K., Serr, M., Gepner, J., Gibbons, I., & Hays, T. S. (1994). A Family Of Dynein Genes In *Drosophila melanogaster*. Mol. Biol. Cell., **5**: 45-55.

Rigby, P. W. J., Dieckmann, M., Rhodes, C., & Berg, P. (1977). Labelling DNA To High Specific Activity *In Vitro* By Nick Translation With DNA Polymerase I. J. Mol. Biol., **113**: 237-251.

Ritossa, F. (1976). The *bobbed* Locus. In The Genetics And Biology Of *Drosophila*. Ashburner & Wright (Eds.). Academic press.

Robertson, H. M., Preston, C. R., Phillis, R. W., Johnston-Schlitz, D. M., Benz, W. K., & Engels, W. R. (1988) A Stable Genomic Source Of P-Element Transposase In *Drosophila melanogaster*. Genetics, **118**: 461-470.

Russell, S. (1989). Ph.D. Thesis, University of Glasgow.

Russell, S. R. H., & Kaiser, K. (1993). *Drosophila-melanogaster* Male Germ Line-Specific Transcripts With Autosomal And Y-Linked Genes. Genetics, **134**: 293-308.

Russell, S. R. H., & Kaiser, K. (1993). A *Drosophila* Chromosome 2L Repeat Is Expressed In The Germline. Chromasoma, **103**: 54-62.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular Cloning, A Laboratory Manual. Cold Spring Harbour Press.

Sargent, T. D. (1987). Isolation Of Differentially Expressed Genes. Methods In Enzymology. Guide To Molecular Cloning Techniques. Berger & Kimmel (Eds.) **152**: 423-432. Academic press.

Schafer, U. (1986a). Genes For Male-specific Transcripts In *Drosophila melanogaster*. MGG., **202**: 219-225.

Schafer, U. (1986b). The Regulation Of Male-Specific Transcripts By Sex Determining Genes In *Drosophila melanogaster* EMBO, **5**: 3579-3582.

Schafer, M., Kuhn, R., Bosse, F., & Schafer, U. (1990). A Conserved Element In The Leader Mediates Post-Meiotic Translation As Well As Cytoplasmic Polyadenylation Of A *Drosophila* Spermatocyte RNA. EMBO, **9**: 4519-4525.

Schafer, M., Borsch, D., Hulster, A., & Schafer, U. (1993). Expression Of A Gene Duplication Encoding Conserved Sperm Tail Proteins Is Transcriptionally Regulated In *Drosophila melanogaster*. Mol. Cell. Biol, **13**: 1708-1718.

Schupbach, T., Wieschaus, E., & Nothiger, R. (1978). The Embryonic Organisation Of The Genital Disc Studied In Genetic Mosaics Of *Drosophila melanogaster*. Wilhem Rouxs Arch, **185**: 249-270.

Short, J.M., Fernandez, J.M., Sorge, J.A., and Huse, W.D., (1988) λZAP: a bacteriophage λ expression vector with *in vivo* excision properties. Nucl. Acids Res. **16**, 7583-7600.

Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A., Lovell-Badge, R., & Goodfellow, P. (1990). A Gene From The Human Sex-Determining Region Encodes A Protein With Homology To A Conserved DNA-Binding Motif. Nature: 240-244.

Sonnenblick, B. P. (1950). The Early Embryogenesis Of *Drosophila*. In The Biology of *Drosophila*. Demerec (Ed.). Hafner publishing.

Southern, E. M. (1975). Detection Of Specific Sequences Among DNA Fragments Separated By Gel Electrophoresis. J. Mol. Biol., **98**: 503-517.

Staben, C., & Yanofsky (1990). *Neurospora crassa* : A Mating Type Region. PNAS., **87**: 4917-4921.

Steller, H., & Pirrotta, V. (1986). P-Transposons Controlled By The Heat Shock Promoter. Mol. Cell. Biol., **6**: 1520-1528.

St. John, T. P. S., & Davis, R. W. (1979). Isolation Of Glactose Inducible Sequences From *Saccaromyces cerevisiae* By Differential Plaque Hybridisation. Cell, **16**: 443.

Sturtevant, A. H. (1945). A Gene In *Drosophila melanogaster* That Transforms Females Into Males. Genetics, **30**: 297-299.

Taniguchi, T., Fujii-Kuriyama, Y., & Muramatsu, M. (1980). Molecular Cloning Of Human Interferon cDNA. PNAS., **77**: 4003-4006.

Tautz, D., & Pfeifle, C. (1989). A Non-Radioactive *in situ* Hybridisation Method For The Localisation Of Specific RNAs In *Drosophila* Reveals Translational Control Of The Segmentation Gene *hunchback*. Chromosoma, **98**: 81-85.

Timberlake, W. E. (1980). Developmental Gene Regulation in *Aspergillus nidulans*. Dev. Biol., **78**: 497.

Travis, A., Amsterdam, A., Belanger, C., & Grosschendel, R. (1991). LEF-1, A Gene Encoding A Lymphoid-Specific Protein, With A HMG Domain, Regulates T-Cell Receptor And Enhancer Function. Genes and Dev., **5**: 880-894.

Watanabe, T. K. (1975). A New Sex Transforming Gene On The Second Chromosome Of *Drosophila melanogaster*. Japn. J. Genet., **50**: 269-271.

Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R. C., Laue, E. D., & Thomas, J. O. (1993). Structure Of The HMG Box Motif In The β-Domain Of HMG1. EMBO, **12**: 1311-1319.

Wetering, M. v. d., Oosterwegal, M., Dooijes, D., & Clevers, H. (1991). Identification And Cloning Of TCF-1, A T Lymphocyte Specific Transcription Factor Containing A Sequence-Specific HMG box. EMBO, **10**: 123-132.

Wetering, M. v. d., & Cleavers, H. (1992). Sequence Specific Interactions Of The HMG Box Proteins TCF-1 and SRY Occurs Within The Minor Groove Of A Watson-Crick Double Helix. EMBO, **11**: 3039-3044.

Whittaker, P. A., Campbell, A. J. B., Southern, E. D., & Murray, N. E. (1988). Enhanced Recovery And Restriction Mapping Of DNA Fragments Cloned In A New Lambda Vector. NAR, **16**: 6725-6228.

Williamson, J. H. (1976). The Genetics Of The Y Chromosome. In The Genetics And Biology Of *Drosophila*. Ashburner & Wright (Eds.). Academic Press.

Yanicostas, C., & Lepesent, J. A. (1990). Transcriptional And Translational Cis -Regulatory Sequences Of The Spermatocytes-Specific *Drosophila janus*B Gene Are Located In The 3' Exonic Region Of The Overlapping *janus*A Gene. MGG, **224**: 450-458.

Zimmerman, C. R., Orr, W. C., Leclerc, R. F., & Bernard, E. C. (1980). Molecular Cloning And Selection Of Genes Regulated In *Apergillus nidulans*. Cell, **21**: 683.

Appendix I: Mapping of Genomic DNA Clone gK27.

Restriction data for λgK27.

DNA was isolated from genomic clone λgK27 and digested with *Eco*RI, *Sal*I and *Hin*dIII separated on 1% TBE agarose gels and Southern blotted. The filters were probed with pmsf27.1 and pmsf27.2, the male-specific fragments identified by Reverse northern analysis (chapter 3). An Ethidium Bromide stained agarose gel of the digests of gK27 genomic phage is shown in Figure A1.1.

These digests were used to construct the restriction maps of the clone shown Figure 4.1(d) in conjunction with data from other gels (data not shown). Data from the overlapping clone (gK6) was also used as these clones have been shown to be related (discussed in chapter 3). As can be seen from the gel photograph (Figure A1.1) and from the Southern analysis presented in chapter 4 and chapter 6, gK27 represents a complex genomic clone. Therefore, the map shown in Figure 4.1(d) represents one interpretation of the data, there are other interpretations that could produce a different map. The map could have been further resolved by using finer mapping techniques such as digesting with enzymes that were not found in the lambda arms, isolating specific fragments from the gel and restricting them with additional enzymes. An additional refinement could have been achieved by using a processive double stranded exonuclease (such as Bal 31), the resulting DNA is then restricted with the appropriate enzyme, the fragments are lost in the order in which the restriction site appears on the map.

However, as the map was constructed to facilitate the cloning of the male specific fragments the level of mapping used was sufficient.

The size of the restriction fragments in all of the digests for these clones are given in table A1.1 (gK27). The positions of the fragments on the gel are indicated by letters in the schematic representation of figure A1.1 (figure A1.2). Some of the fragments from table A1.1 are indicated on the map in figure A1.3. For the sake of clarity, only some of the fragments are indicated on the map.

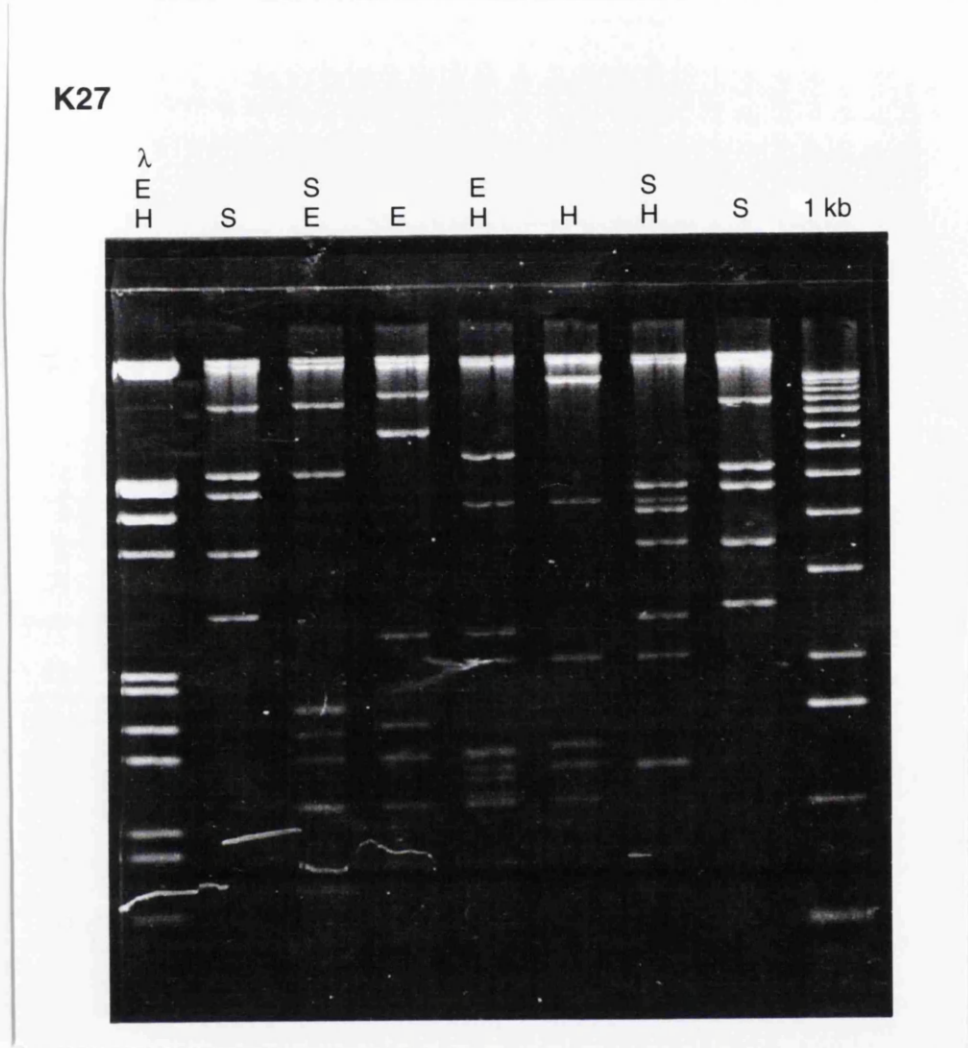


Figure A1.1 : A 1% TBE agarose gel stained with EthBr of genomic clone gK27. The enzymes used in each digest are indicated above the lane. The markers are 1 Kb ladder and Lambda DNA digested with *Eco*RI and *Hin*dIII.

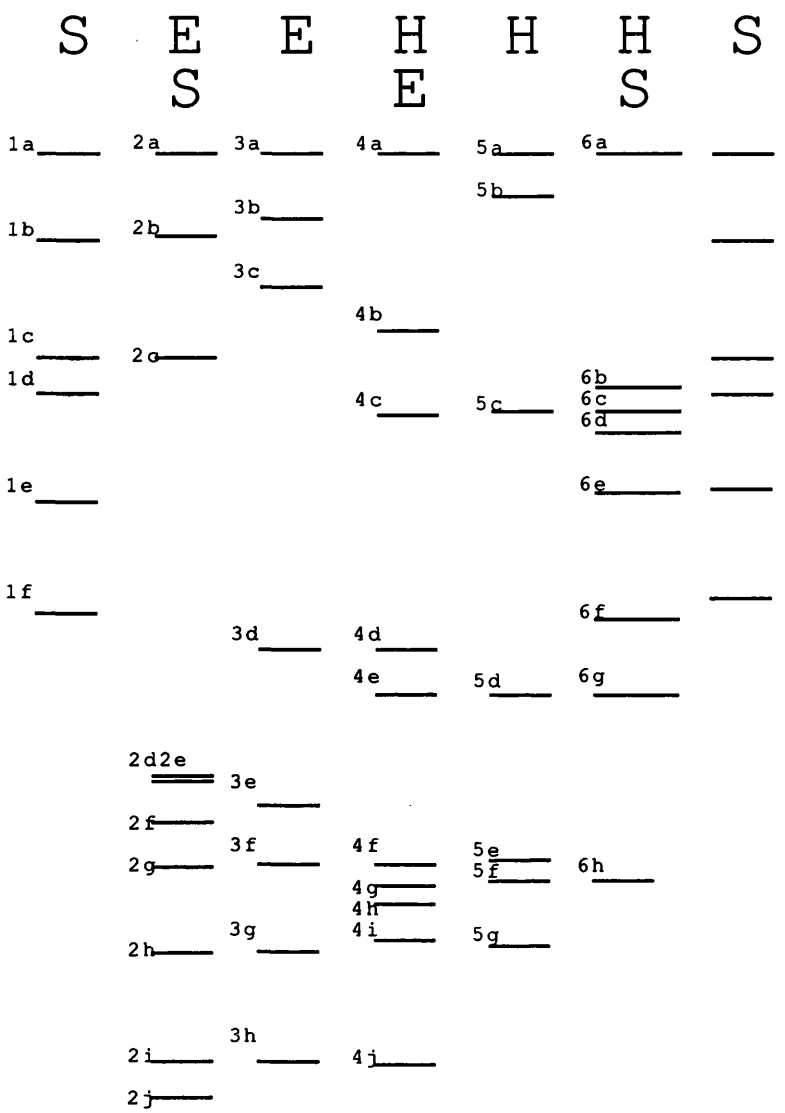


Figure A1.2
Schematic representation of Figure A1.1. The fragment sizes correspond to those quoted in Table A1.1.

λgK27

	1	2	3	4	5	6
	*Sal*I	*Sal*I/*Eco*RI	*Eco*RI	*Eco*RI/*Hin*dI II	*Hin*dIII	*Sal*I/*Hin*dIII
a	**19.9**	**19.9**	**19.9+**	**19.9+**	**19.9+**	**19.9**
b	**8.8**	**8.8**	**10**	**6**	**12**	4.7
c	5.2	5.3	6.8	4.5	**4.5**	**4.5**
d	4.7	1.8	2.3	2.3	2	4.3
e	3.5	1.8	1.65	2	1.5	3.5
f	2.6	1.6	1.4	1.4	1.3	2.4
g		1.4	1.15	1.3	1.1	2
h		1.15	0.8	1.25		1.3
i		0.8		1.15		
j		0.75		0.8		

Table A1.1: Table of the size of fragments (in Kb) produced by the various digests. Fragments that contain DNA from the lambda arms are indicated in bold type.

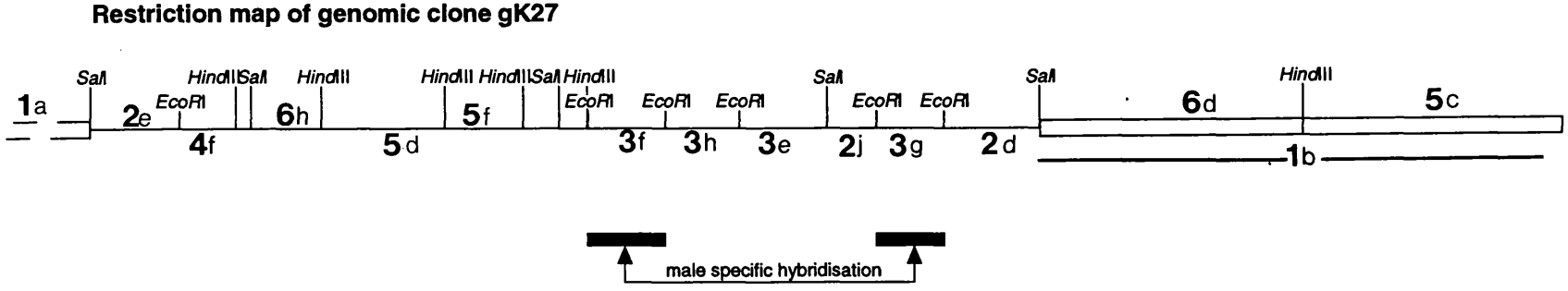


Figure A1.3: The restriction map of genomic clone gK27. The letters correspond to those in Table A1.1 and Figure A1.2.

Appendix Two: Sequence Analysis of K27 Clones.

The identification of the cDNA clone corresponding to the genomic clones pmsfK27.1 and pmsfK27.2 was described in chapter 6. Figure A2.1 shows an autoradiograph of a sequencing gel of cK27. In this figure, the poly[A]+ tail is clearly visible in the track corresponding to the sequencing reaction using dideoxyA. The junction between the BlueScript cloning vector (pBlueScript) and the cDNA clone (cK27) is indicated. The restriction site for the enzyme *Eco*RI which was used in the cloning of the cDNA is indicated. The sequence of the BlueScript vector is shown after the *Eco*RI site and the sequence of the 3' end of cK27 corresponds to that in Figure 6.2 in chapter 6.

Figure A2.2 shows an autoradiograph of the the 5′ ends of the two genomic clones, pmsfK27.1 and pmsfK27.2. In this autoradiograph, a comparison of the sequence shows that over this region, the two clones are identical. Again, the junction between the genomic clones and the pBlueScript vector is indicated and the *Eco*RI cloning site is also highlighted. The sequence on this autoradiograph corresponds to that of Figure 6.6 (b).

Figure A2.3 shows an autoradiograph of a sequencing gel of the internal sequences of the two genomic clones, pmsfK27.1 and pmsfK27.2. The sequence differences between the two genomic clones can clearly be seen. Two large regions of sequence divergence are indicated by GAP1 and GAP2 (these areas of divergence are also indicated in Figure 6.6(b)).

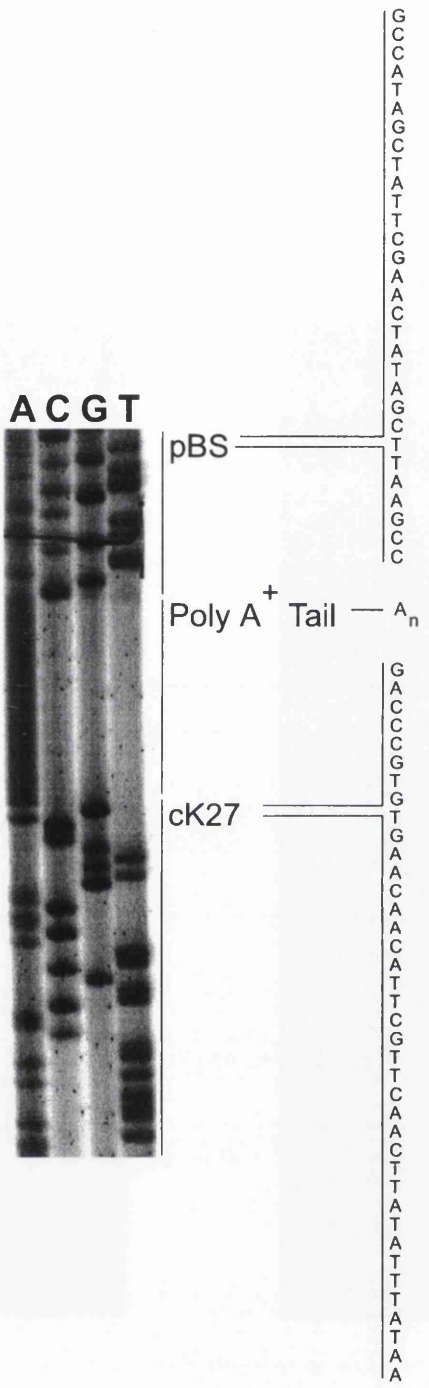


Figure A2.1
Autoradiograph of a sequencing gel of cK27. The sequence of cK27 corresponds to the 3'end of the cDNA clone described in Figure 6.2.

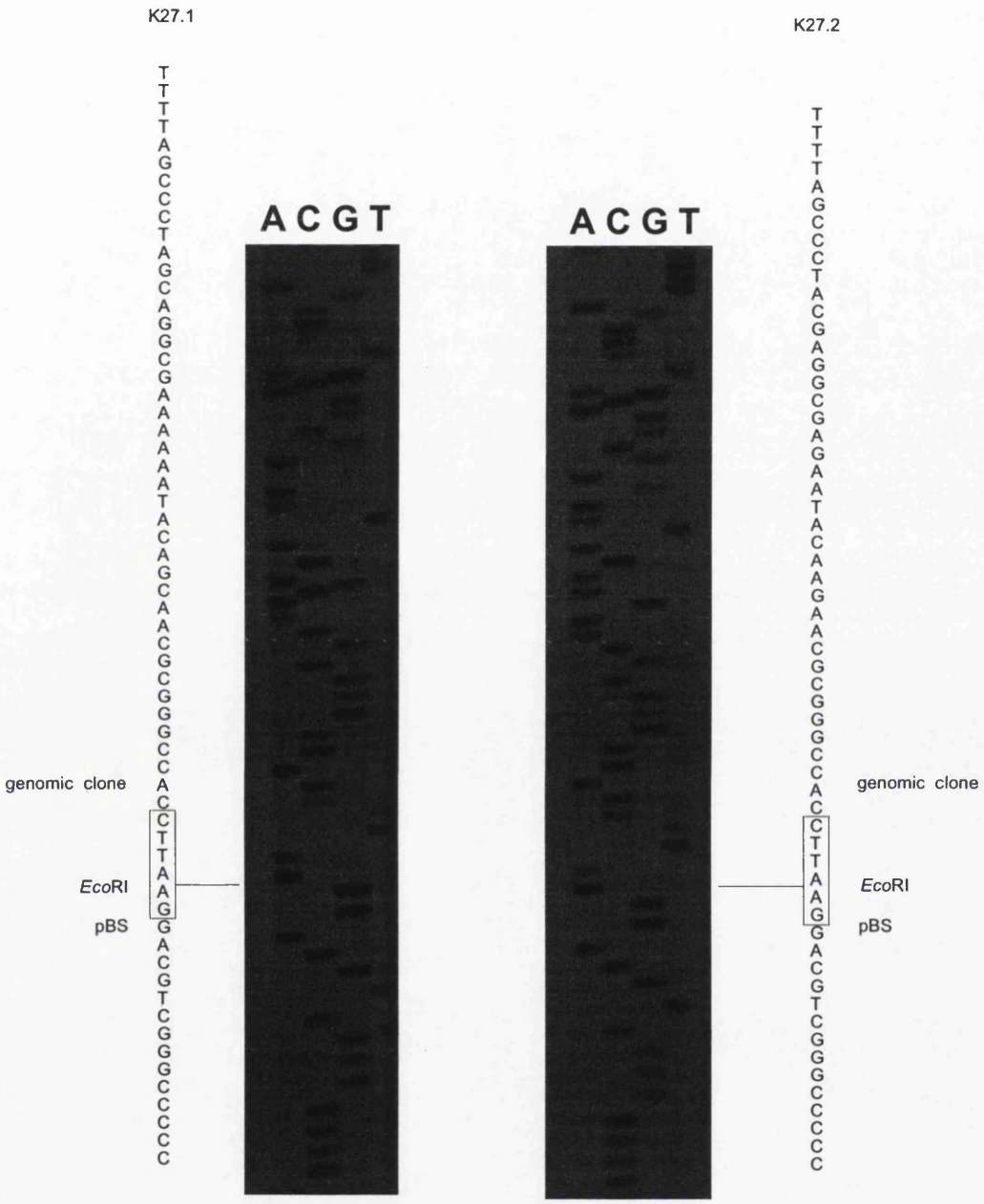


Figure A2.2:
Junction of the two genomic fragments (gK27.1 and gK27.2) with the pBlueScript vector. The *Eco*RI cloning site is indicated.

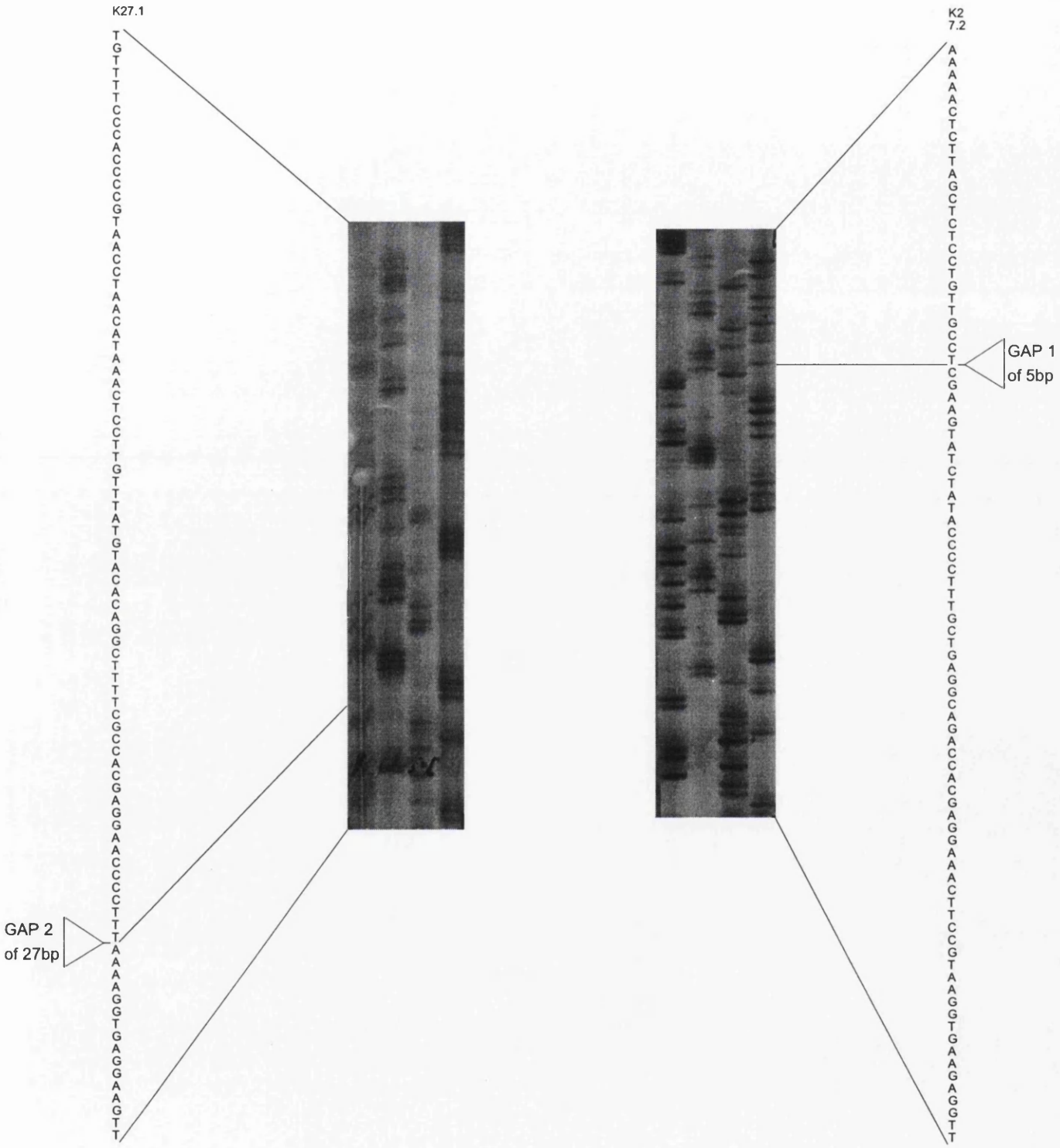


Figure A2.3:
Sequence of gK27.1 and gK27.2. The differences between the two clones are indicated. The sequence quoted corresponds to that given in Figure 6.6(b).