ISOLATION AND CHARACTERISATION OF THE *DROSOPHILA DROR2* GENE

A thesis presented in partial fulfilment of
the requirements for the degree of
Master of Science in Genetics
at Massey University, Palmerston North
New Zealand

Kathryn Joy Frith

1997

ABSTRACT

Receptor tyrosine kinases (RTKs) are a family of cell-surface receptors that have an important role in an array of cellular responses including cell migration, proliferation and differentiation (Fantl *et al.*, 1993). RTKs and their ligands are important components in the determination of cell fate through signalling pathways that are activated during both invertebrate and vertebrate development (Pawson and Bernstein, 1990).

The Ror subfamily of RTKs are thought to be important for the development of the nervous system as they are expressed highly in the nervous system in the developing embryo, but expression is minimal in adults. Three receptors in this subfamily have been identified, Ror1 and Ror2 from humans (Masiakowski and Carroll, 1992) and Dror from *Drosophila* (Wilson *et al.*, 1993). This thesis involved the isolation and characterisation of the fourth gene in this family *Dror2* from *Drosophila melanogaster*.

Degenerate oligonucleotides to conserved regions of the tyrosine kinase domain of RTKs were used to PCR amplify a 200 bp fragment from genomic DNA. A λ genomic library was screened with the labelled fragment in order to isolate the gene. The resulting clone was subcloned and sequenced to obtain the complete sequence of *Dror2*. The 3' end of the gene was determined by RT-PCR. The transcriptional start point was identified by using 5' RACE and sequencing of the amplification product. Expression of *Dror2* was examined using Northern Blot hybridisation and *in situ* hybridisation to whole mount embryos.

The 725 amino acid mature Dror2 protein comprises an extracellular domain containing the signal peptide, cysteine-rich region and kringle domain, a hydrophobic transmembrane domain and the intracellular domain containing a catalytic kinase domain. Three introns were identified, one in the middle of the cysteine-rich region and two flanking the kringle domain.

ACKNOWLEDGEMENTS

Firstly I would like to thank my supervisor Dr Max Scott for his wisdom, guidance and wealth of knowledge which was immeasurable. Thank you for your patience, especially when dealing with the many hurdles we encountered along the way.

A special thanks has to go to Helen whose continual support and constant laughter (at my expense usually) kept the lab a great place to work, even when things weren't going to plan. Thanks for the help with all the sequencing, editing and for your friendship. To Sheralee our faithful technician, thanks for all the demonstrations of techniques and for your readiness to help out whenever necessary. To everyone else who has passed through the lab; Andrea, Lewis, Xuelei, it's been fun.

I would like to thank all the people in MGU, especially Carolyn for all the advice and helpfulness, even when she was busy herself. Thanks also to BJ for the help getting my sequence presentable, I still don't like GcG! Thanks must go to George Ionas, whose advice about PCR was great and whose PCR machine was invaluable. It didn't always produce the results it was meant to, but we got there in the end. To Simon Greenwood, who helped with my "doomed" *in situ* hybridisations; thanks for trying.

Thank you also to my family who were always there for me, their love and support was endless. To my flatmates past and present, thanks for putting up with me, especially to Deb, Ruth and Jane for their friendship and advice. I have many other friends up and down the country to whom I owe much gratitude.

Finally, I must thank my fiancé Steve for everything he has done for me over the past three years; for your love, encouragement, guidance and great meals. I could not have done this without you.

ABBREVIATIONS

α	alpha
β	beta
Δ	delta
λ	lambda
aa	amino acids
bp	base pairs
BSA	bovine serum albumin
CIP	calf intestine phosphatase
CNS	central nervous system
cDNA	complementary DNA
°C	degrees Celsius
dNTP	deoxynucleotide triphosphate
DNase	deoxyribonuclease
DNA	deoxyribonucleic acid
DEPC	diethyl pyrocarbonate
DTT	dithiothreitol
dsDNA	double stranded DNA
f	fico
g	grams
kb	kilobase pairs
L	litre
mRNA	messenger RNA
μCi	micro Curies
μ	micro
m	milli
M	molar
n	nano

nm	nanometres
nt	nucleotide
PNS	peripheral nervous system
pfu	plaque forming units
poly A^+ RNA	polyadenylated RNA
PCR	polymerase chain reaction
RACE	rapid amplification of cDNA ends
RTK	receptor tyrosine kinase
RT-PCR	reverse transcriptase PCR
rpm	revolutions per minute
RNase	ribonuclease
RNA	ribonucleic acid
ssDNA	single stranded DNA
SDS	sodium dodecyl sulphate
UV	ultraviolet
U	units
UTR	untranslated region
V	volts
v/v	volume per volume
W	watts
w/v	weight per volume

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

1. INTRODUCTION

1.1 RECEPTOR TYROSINE KINASES (RTKS)

Receptor tyrosine kinases (RTKs) are a family of cell-surface receptors that possess an intrinsic, ligand sensitive protein tyrosine kinase activity (Yarden and Ullrich, 1988). RTKs and their ligands are important components in the determination of cell fate through signalling pathways that are activated during both invertebrate and vertebrate development (Pawson and Bernstein, 1990). The receptors have an important role in an array of cellular responses including cell migration, proliferation and differentiation (Fantl *et al.*, 1993).

RTKs are Type I transmembrane proteins that have their N-termini outside the cell and single membrane-spanning regions (Van der Geer *et al.*, 1994). There are several structural features that are conserved among all known RTKs. At the N-terminus is a signal peptide that targets the protein to the secretory pathway. This is followed by an extracellular domain of several hundred amino acids that contains a distinctive pattern of cysteine residues and often a characteristic array of structural motifs. This ligand-binding domain is the most distinguishing feature from other RTKs. The transmembrane domain follows, which consists of a stretch of hydrophobic residues that are followed by several basic residues functioning as a stop-transfer signal. A juxtamembrane region on the cytoplasmic side of the membrane precedes the catalytic domain. All RTKs kinases share a conserved catalytic domain that contains kinase residues that are involved in receptor function. The C-terminal tail is typically very hydrophilic and rich in small amino acids and appears to have varying functions (Yarden and Ullrich, 1988).

RTKs have been classed into 14 subfamilies depending on the structure of each RTK as seen in Figure 1.1. Members of a given subfamily share common structural features that are distinct from those found in other subfamilies (Fantl *et al.*, 1993). Each subfamily contains different combinations of recognisable sequence motifs in their ligand-binding domain but all catalytic domains are similar.

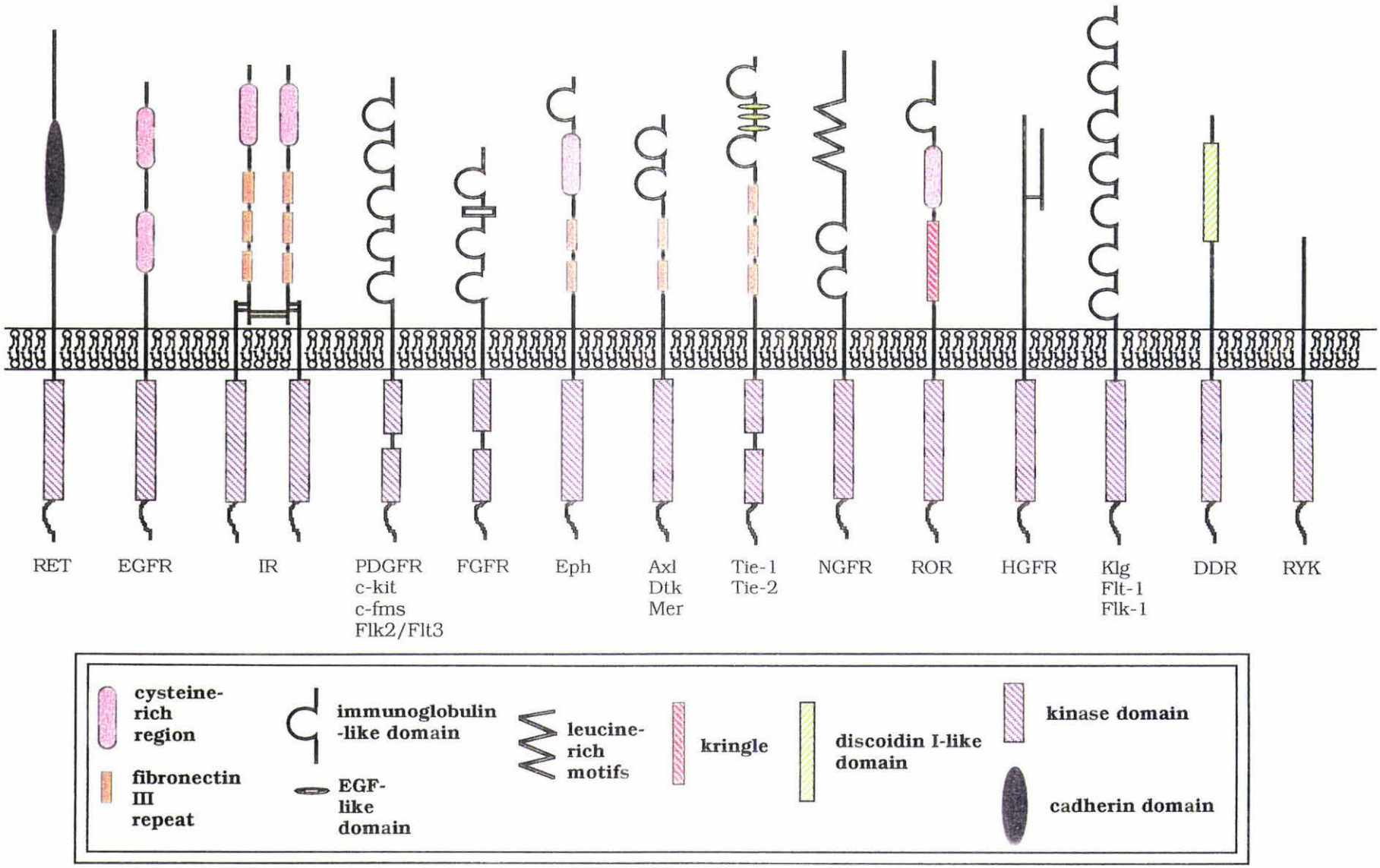
Despite the diversity of RTKs there is a great degree of commonality in the types of intracellular signalling pathways initiated by these proteins. In *Drosophila*, biochemical and molecular genetic analyses have shown that for all RTKs the binding of ligand to the extracellular domain activates the tyrosine kinase in the cytoplasmic domain. This leads to downstream activation of a number of common signalling molecules. The activation of signalling pathways involving these molecules leads to changes in gene expression and a change in the phenotypic state of the cell. A single type of RTK can elicit very different biological responses in different cell types (Fantl *et al.*, 1993). It is not clear how this is accomplished.

1.2 RTKs IN *DROSOPHILA*

The activation of similar signalling molecules has been confirmed through genetic analyses in two model systems, *Caenorhabditis elegans* and *Drosophila melanogaster*. These studies have proven that RTK molecules are important for the developmental specification of cell types and that RTKs in these organisms initiate signalling pathways that are very similar to those found in mammalian systems (Fantl *et al.*, 1993).

Figure 1. 1 Structure of the fourteen RTK subfamilies.

Abbreviations are as follows: EGFR: epidermal growth factor, IR: insulin receptor, PDGFR: platelet derived growth factor, *Flk*-2: foetal liver kinase-2, *Flt*-3: *fms*-like tyrosine kinase-3, FGFR: fibroblast growth factor receptor, NGFR: nerve growth factor receptor, HGFR: hepatocyte growth factor receptor, DDR: discoid domain receptor. Adapted from Van der Geer *et al.* (1994) by Dr Phil Crosier, University of Auckland.


RET
EGFR
IR
PDGFR
c-kit
c-fms
Flk2/Flt3
FGFR
Eph
Axl
Dtk
Mer
Tie-1
Tie-2
NGFR
ROR
HGFR
Klg
Flt-1
Flk-1
DDR
RYK
cysteine-rich region
fibronectin III repeat
immunoglobulin -like domain
EGF-like domain
leucine-rich motifs
kringle
discoidin I-like domain
kinase domain
cadherin domain

In *Drosophila*, the signalling cascades from three RTKs have been studied genetically. In the embryo, the torso (tor) RTK is present throughout the embryo and is required for the specification of terminal structures (Brunner *et al.*, 1994). In normal development an extracellular ligand activates tor at both poles of the embryo. There are no terminal structures in mutants that do not express tor. Tor gain-of-function mutants have a receptor that is active throughout the embryo, leading to the generalised expression of the tor target gene *tailless* (*tll*) (Perrimon and Desplan, 1994).

Another RTK important in development in *Drosophila* is the EGF receptor homolog named DER. DER has both maternal and zygotic functions. In the mother DER is required for the establishment of the dorsoventral polarity of the egg shell and the embryo (Schupbach, 1987). In the absence of a DER signal, a ventral instead of dorsal phenotype is induced (Pawson and Bernstein, 1990). In the zygote, the absence of DER activity at Stages 8 and 9 of embryonic development results in cells adopting an altered cell fate (Raz and Shilo, 1993).

In contrast to DER, the *sevenless* (*sev*) gene encodes a transmembrane RTK whose only known biological function is the specification of a single photoreceptor cell of the ommatidial clusters that comprise the *Drosophila* compound eye (Zipursky and Rubin, 1994). In the developing eye, specification of the R7 photoreceptor cell in each ommatidium depends on the local activation of the sev RTK in the R7 precursor cell by the membrane bound bride-of-sevenless (boss) protein on the neighbouring R8 cell (Hafen *et al.*, 1987).

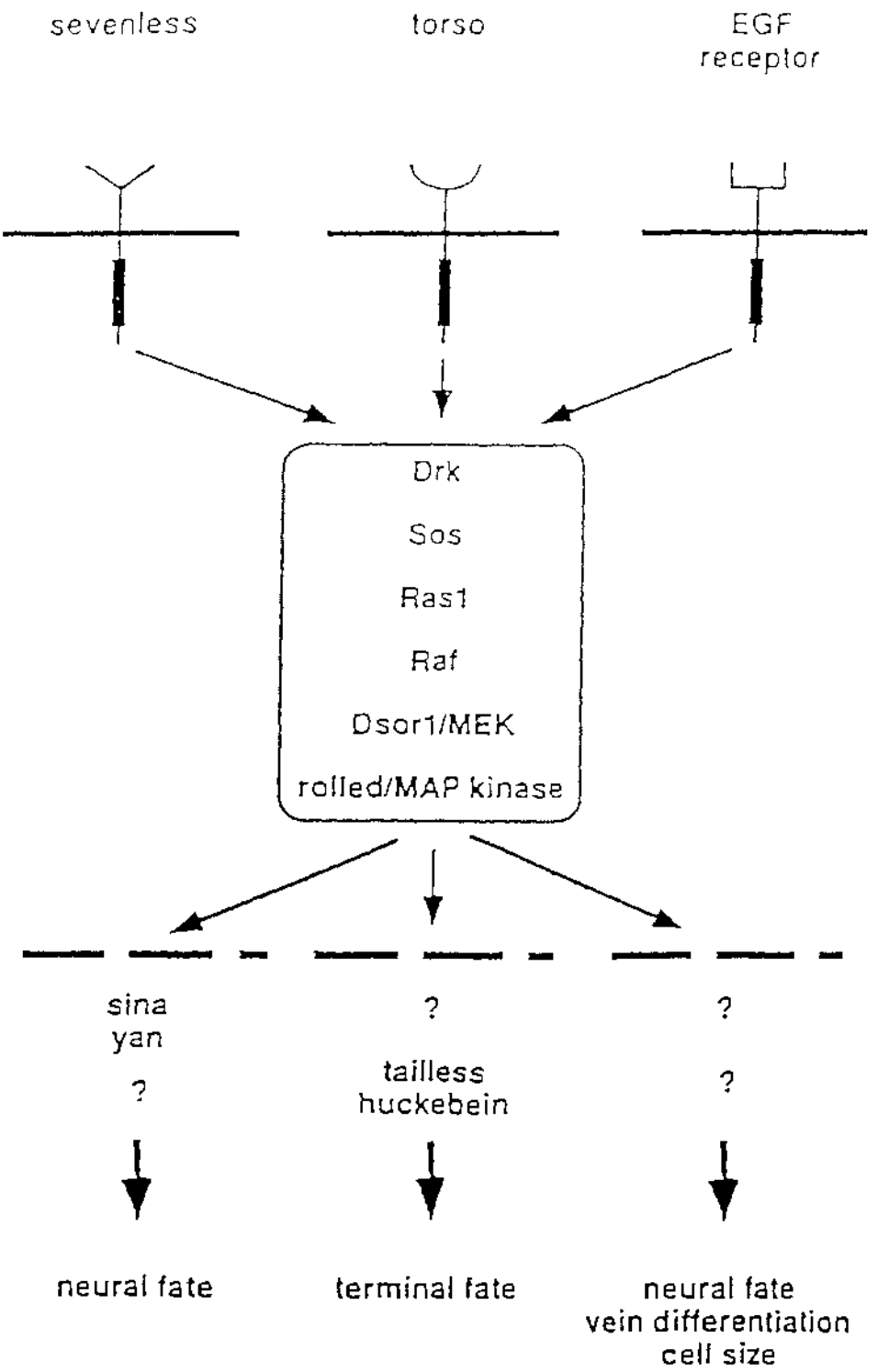
The sevenless, torso and DER RTKs are prototypes for tyrosine kinases that convey positional and developmental information by interpreting cues provided by nearby cells (Pawson and Bernstein, 1990). This type of local cellular interaction is likely to have vast significance in defining the developmental fate of cells in *Drosophila* embryogenesis.

These three RTKs have been shown to utilise a similar signalling cascade (Brunner *et al.*, 1994) as seen in Figure 1.2. Binding of a ligand to an RTK leads to activation of the cytoplasmic tyrosine kinase domain. This results in autophosphorylation of tyrosine residues allowing the receptor to interact with Drk. Following this, the Drk protein binds to the son-of-sevenless (sos) guanine nucleotide exchange factor activating the protein. Sos can now act to convert the inactive Ras-GDP to the active Ras-GTP form. The Ras guanine nucleotide binding protein has been identified as a critical determinant of cellular differentiation (Avruch *et al.*, 1994). The immediate target of Ras is Raf, a serine/threonine kinase that is the next member of the signal transduction cascade. Raf then binds and activates the MAP kinase activator MEK, which in turn activates MAP kinase. MAP kinase is the final component whose activation is both necessary and sufficient for signal transduction in the pathway.

From this point, transcription factors are phosphorylated by the pathway and subsequent genes are activated. In the torso pathway the transcription factor(s) involved is unknown, but both *tailless* (*tll*) and *huckebein* (*hkb*) are activated in the terminal regions in response to tor signalling (Perrimon and Desplan, 1994). In the sevenless pathway the expression of *sina* and *yan* are not dependent on sev activity, but may be required to mediate regulated transcription in response to sev activity (Brunner *et al.*, 1994). In the DER pathway neither the transcription factors nor genes involved in instructing dorsal follicle cell fate have yet been identified.

Figure 1. 2 Different receptors utilize a similar signalling cascade

Components common to the tor, sev and DER pathways are shown (Brunner *et al.*, 1994). See Section 1.2 for details.


sevenless
torso
EGF receptor
Drk
Sos
Ras1
Raf
Dsor1/MEK
rolled/MAP kinase
sina
yan
?
?
tailless
huckebein
?
?
neural fate
terminal fate
neural fate
vein differentiation
cell size

Along with the three major RTKs, a number of other *Drosophila* RTKs have been identified.

Dtrk is highly related to the trk family of mammalian neurotrophin receptors (Pulido *et al.*, 1992). The product of *Dtrk* is expressed in several areas of the developing nervous system. The gene is thought to promote cell adhesion that activates its tyrosine kinase activity. The trk family of receptors are discussed further in Section 1.3.3.

Dror encodes a gene that is homologous to the vertebrate Ror family of Trk-related RTKs (Wilson *et al.*, 1993). The neurotrophic receptor functions during the early nervous system development in flies. This family of receptors are discussed more fully in Section 1.3.4 and 1.4.

Dret is a homolog of the human proto-oncogene *ret*. It is transiently expressed in embryonic neuronal precursor cells including neuroblasts and CNS cells (Sugaya *et al.*, 1994). The structure of *Dret* and expression is highly related to that of human *ret*, so it is feasible that both genes are involved in the same developmental process in neuronal tissues.

The *Drosophila derailed* (*drl*) gene is involved in key aspects of neuronal pathway recognition (Callahan *et al.*, 1995). *Drl* is expressed in only a small subset of embryonic interneurons that share common pathways.

Lio encodes a putative RTK that is unusual as it contains a short extracellular domain and modified catalytic domain (Dura *et al.*, 1995). *Lio*, though not vital, is important for learning and memory in flies. Lio mutants have weak learning and memory capabilities with reduced shock reactivities and olfactory associative memory. Both *drl* and *lio* genes are homologous to the vertebrate gene *RYK*. The

function of *RYK* is unknown but evidence from *Drosophila* suggests the genes may be involved in a new, undetermined signal transduction pathway.

1.3 RTKs OF INTEREST

1.3.1 DTK

Dtk is a novel membrane bound developmental tyrosine kinase receptor that has been cloned recently in mice, zebra fish and humans (Crosier *et al.*, 1994a and b, Walshe *et al.*, 1995). Dtk is a well-conserved receptor tyrosine kinase in vertebrates, suggesting an important biological function. It is expressed in many tissues during embryogenesis and is strongly expressed in the adult brain and other tissues to a lesser extent. The 850 amino acid mature receptor protein comprises an extracellular domain with two immunoglobulin-like motifs and two fibronectin type III modules, a 25 amino acid transmembrane domain and a cytoplasmic region with a catalytic kinase domain (Crosier *et al.*, 1994a).

There are two repeating protein motifs within the domain. One of these contains two C-type immunoglobulin like motifs, the first domain has similar structure to a C1-like domain, while the second is more related to the C2-like domain. Also in the domain are two fibronectin type III motifs. Like Dtk, Axl (Ufo, Ark) also contains two immmunoglobulin domains followed by two fibronectin type III repeats in the extracellular domain (Crosier *et al.*, 1994a).

The mouse *Dtk* gene was first isolated from embryonic stem (ES) cells (Crosier *et al.*, 1994a). Using RNA isolated from differentiating ES cells, a number of RTKs were isolated that were potentially involved in haematopoietic development. Degenerate oligonucleotides derived from conserved domains within the catalytic kinase domain of protein tyrosine kinases were constructed and used in PCR

reactions to isolate the RTKs. A number of sequences from putative and known RTKs were found, one of which was designated *Dtk*.

In embryonic stem cells and embryonic bodies, *Dtk* is expressed almost uniformly from days 0-18. Within the embryonic tissues expression was detected in total RNA from a range of tissues including the brain, eye, thymus, lung, intestine, forelimb, hindlimb, and testis. There was limited expression in the heart and unfractionated liver. In the adult tissues the pattern of expression becomes restricted. Transcripts were greatest in the brain, oesophagus, bladder, testis and ovary. In the brain the expression of *Dtk* was more abundant in the adult than in embryonic tissue. There were also traces of *Dtk* in the lung, stomach and in the intestines (Crosier *et al.*, 1994a).

The human *Dtk* gene was isolated using the murine Dtk cDNA as a probe to screen a human brain cDNA library by Crosier *et al.* (1994b). This gene is identical to the human *Tyro3* gene cloned by Polvi *et al.* (1993). The protein has identical predicted structural features to those of the murine Dtk, and there is an 89 % amino acid identity between the two genes.

To analyse the expression of *Dtk* in humans, Northern blot hybridisation was carried out. *Dtk* was shown to be expressed in foetal brain and kidney, as well as low levels in foetal lung and heart (Crosier *et al.*, 1994b). In adult tissues *Dtk* expression was found in the brain, kidney, testis and ovary. Limited expression was seen in the other adult tissues with no expression in the peripheral blood leucocytes and liver. Thus, the pattern of expression was very similar to murine *Dtk*, with the notable exception that there is little expression in the adult kidney (Crosier *et al.*, 1994b).

1.3.2 EPH

The Eph subfamily is the largest of the RTKs with 13 receptors and 7 of their ligands identified in mammals to date (Muller *et al.*, 1996). The Eph subfamily of receptors comprise an extracellular domain that includes one immunoglobulin-like motif, two fibronectin type III repeats and a cysteine rich region, a hydrophobic region comprising the transmembrane domain and a cytoplasmic region with a catalytic kinase domain (Van der Geer *et al.*, 1994).

The first receptor isolated in the Eph subfamily was identified by Hirai *et al.* (1987) in a human genomic DNA library during a search for sequences homologous to the tyrosine kinase domain of *v-fps* (a viral oncogene). Following this, the remainder members of the family were identified from various vertebrate species including humans, mouse, rat, *Xenopus* and zebrafish (Pandey *et al.*, 1995).

Most of these receptors are specifically expressed in neurons, but the biological functions of many of the Eph-type RTKs are unknown. Ligand binding to an Eph RTK appears to lead to activation of a novel signal transduction pathway. For example, activation of the ectopically expressed receptors does not lead (as expected) to a biological response from the cell (Muller *et al.*, 1996).

Another unique feature of this subfamily is that only the membrane-bound forms of ligands are active, leading to speculation that these receptors play roles in axon guidance, neuronal bundling or angiogenesis (Stein *et al.*, 1996). More specifically it has been suggested that these receptors and ligands may play a significant role in forming topographic projections along the anteroposterior axis of the tectum (Tessier-Lavigne, 1995). The tectum is the structure in the brain that transmits the retinal image to the brain by rectifying the back-to-front, upside-down image. To do this, the anterior and posterior ganglion cells map to

the posterior and anterior tectum, and the dorsal and ventral ganglion cells map to the ventral and dorsal tectum (Orike and Pini. 1996). Studies performed on the chicken showed the presence of a retinotectal projection in which retinal ganglion cell axons project in a topographic manner onto the optic tectum, the target area. It is thought that complementary gradients of receptors in the retina and of corresponding ligands in the optic tectum are crucial for the formation of this projection (Muller *et al.*, 1996).

As well as this, expression of some Eph-related molecules in migratory cells suggests a further role of these receptors in the process of cell migration (Muller *et al.*, 1996). Thus it appears that Eph-type RTKs play important roles in maintaining and developing many cell and tissue types.

1.3.3 TRK

The Trk receptors belong to the Nerve Growth Factor (NGF) receptor subfamily of RTKs. The NGF receptors comprise an extracellular domain that includes a leucine rich region and two immunoglobulin-like motifs, a transmembrane domain and a cytoplasmic region with a catalytic kinase domain (Van der Geer *et al.*, 1994).

The original member of this family, proto-*trk* (Martin-Zanca *et al.*, 1986), was named after the human oncogenic product *TRK*, a cytosolic non-muscle tropomyosin receptor kinase fusion protein. This was renamed *TrkA* when two other members of this group were identified, *TrkB* (Klein *et al.*, 1990) and *TrkC* (Lamballe *et al.*, 1991). The Trk receptors share structure and sequence similarities and have been shown to bind similar ligands.

All three *Trk* genes are almost exclusively expressed in neurons and are the receptors for neurotrophins. Each neurotrophin activates specific receptors by binding to their extracellular domains. TrkA is a receptor for NGF the best

characterised neurotrophin, TrkB for BDNF and NT-4/5 and TrkC for NT-3 (Grimes *et al.*, 1993). These neurotrophins induce neurite outgrowth and rescue subpopulations of neurons from programmed cell death (Glass and Yancopoulos, 1993). In addition to the Trks, neurotrophins all bind to another cell surface receptor known as p75 or the low-affinity NGF receptor, whose physical role remains to be determined (Barbacid, 1993).

1.3.4 ROR

Ror1 and *Ror2* are two novel developmentally regulated receptor tyrosine kinases cloned from humans and rats (Masiakowski and Carroll, 1992). Both genes are highly expressed in the rat in the brain and body in early embryogenesis, suggesting a role in the development of the embryo. The Ror RTKs are thought to be involved in a network of regulatory interactions, including a possible role in signal transduction pathways and in cell-cell interactions. Ror1 and Ror2 contain one strongly homologous region to the Trk family of nerve growth factor receptors, the tyrosine kinase domain, but because of a unique extracellular region it appears they represent a new subfamily of RTKs (Van der Geer *et al.*, 1994).

The 937 (Ror1) and 943 (Ror2) amino acid mature receptor proteins comprise an extracellular domain with one immunoglobulin-like motif, a cysteine rich region and a kringle domain, a hydrophobic region comprising the transmembrane domain and a cytoplasmic region with a catalytic kinase domain (Masiakowski and Carroll, 1992). The Ror1 and Ror2 proteins are highly homologous, displaying an overall amino acid identity of 58 %.

A kringle domain is a cysteine-rich structure thought to participate in protein-protein interactions (Patthy *et al.*, 1984). Kringle domains have been identified in multiple copies in certain proteins involved in blood clotting, (Kanalas and Makker, 1991; Patthy, 1985), apolipoprotein(a) (McLean *et al.*, 1987) and

hepatocyte growth factor (Nakamura *et al.*, 1989; Seki *et al.*, 1991). Kringle domains have also been found in one muscle-specific RTK. The electric ray *Torpedo californica* possesses a kringle domain in the extracellular domain of the muscle-specific RTK (Jennings *et al.*, 1993). This RTK is thought to be part of the family of mammalian MuSKs (muscle-specific kinases), but is the only one to contain this kringle domain (Valenzuela *et al.*, 1995).

The cytoplasmic catalytic domains of Ror1 and Ror2 are very similar to the Trk kinase domain. For example, the kinase domains of Ror1 and Trk are 47 % identical (Masiakowski and Carroll, 1992).

Cloning of *Ror1* and *Ror2* in humans and rats was undertaken using PCR with degenerate oligonucleotide primers designed to conserved regions of Trk specific clusters of amino acids. Trk and TrkB sequences were eliminated by treating the PCR products with restriction enzymes expected to cut these genes. The remaining fragments were reamplified and cloned. These bacterial colonies were screened by PCR followed by direct sequencing, a cDNA library was screened with these fragments and the longest cDNA inserts sequenced.

To verify the enzymatic activity of Ror1 and Ror2, *in vitro* autophosphorylation of the immunoprecipitated kinase was undertaken. Data indicated that the protein kinase activity is specifically associated with the tyrosine kinase domain of Ror2, but results from Ror1 were inconclusive.

To analyse the expression of *Ror1* and *Ror2* in rats, Northern blot analysis was performed. Both *Ror1* and *Ror2* are highly expressed in the head and body during early embryonic development with the peak of expression at embryonic day 12. The level of expression drops drastically at embryonic day 16 and expression remains very low in adult rats. In contrast expression of *trkB* increases in the later stages of embryonic development (Klein *et al.*, 1990).

1.4 CLONING ROR IN *DROSOPHILA*

The *Drosophila* homologue of *Ror* (*Dror*) was cloned by Wilson *et al.* (1993) by screening a late third instar larval brain cDNA library using a PCR-based approach. Degenerate oligonucleotide primers designed to two conserved regions in the tyrosine kinase domain. Resulting clones were subcloned and sequenced, one of which was designated *Dror*.

Dror encodes a 685 amino acid mature receptor protein comprising an extracellular domain containing a cysteine rich region and a kringle domain, a hydrophobic transmembrane domain and a cytoplasmic region with a catalytic kinase domain (Wilson *et al.*, 1993). The Dror receptor demonstrates a 44 % identity with human Ror1, 40 % identity to human Ror2 and 32 % identity to human TrkB. Extensive sequence similarity between the Ror proteins and Dror is seen throughout the length. However, Dror lacks the immunoglobulin-like domain found in the extracellular domain of Ror1 and Ror2. Dror also lacks the extensive carboxyl terminus tail found in Ror1 and Ror2.

The tyrosine kinase domain of Dror shows significant similarity to the Ror1, Ror2 and Trk proteins of between 53 % and 61 %. A number of the shared amino acids in this domain are not found in many other RTKs suggesting the occurrence of a Trk-like superfamily of RTKs. Although *Dror* demonstrates the highest identity to Ror1 in the tyrosine kinase region (61 %), *Dror* does not share some of the features that distinguished the *Ror* genes from the *Trk* genes in this region.

Within the extracellular domain, Dror shares a 36 % identity with Ror1 and 34 % with Ror2, but no significant similarities to the Trk family. Like Ror1 and Ror2, Dror contains a cysteine-rich region and a kringle domain, but within the cysteine-rich region is a unique 55 amino acid lysine-rich insert. When these three proteins are aligned all 16 cysteines in this region in Ror 1 and Ror2 have

equivalent positions in the Dror protein. The immunoglobulin-like domain seen in Ror1 and Ror2 is not present in Dror. hence the cysteine rich region is positioned close to the amino terminus of Dror.

In situ hybridisation of whole mount embryos was performed to examine the expression of *Dror*. There was no expression of *Dror* in early embryos (Stage 1-10, 1-8 hours), but was seen at Stages 11 (extended germ-band stage) to Stage 15 (8-12 hours). Most of the neurons within the brain and ventral cord expressed *Dror*, as well as some cells in the head and trunk representing organs of the peripheral nervous system. No expression of *Dror* was seen outside the nervous system suggesting an important role of the Ror proteins in early neural differentiation and less important involvement in neuronal cell survival.

1.5 Aims of this Thesis

The original aim of the project was to clone the *Drosophila melanogaster* homologue of the human *Dtk* gene. This was to be accomplished by using the polymerase chain reaction (PCR) with degenerate oligonucleotides derived from conserved amino acid motifs within the tyrosine kinase domain.

Clones from four novel putative RTKs were isolated, but none were similar to Dtk. Thus, we were required to change the original aim of this thesis. The new aim was to characterise the gene for one of the novel RTKs, we subsequently called *Dror2*. The specific objectives were to first determine the sequence and structure of the *Dror2* gene. The latter included mapping intron-exon boundaries, transcription start point and poly A^+ site. The second objective was to determine the temporal and spatial pattern of expression of the *Dror2* gene. The former would be accomplished by using both RT-PCR and Northern blots on RNA isolated from various stages of *Drosophila* development. The pattern of expression would be determined by *in situ* hybridisation.

2. MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 OLIGONUCLEOTIDE PRIMERS

Oligonucleotide primers are listed in Table 2.1.

2.1.2 BACTERIAL STRAINS, λ CLONES AND PLASMIDS

Bacterial strains, λ clones and plasmids use are listed in Table 2.2.

2.1.3 ENZYMES

2.1.3.1 Restriction Enzymes

*Bam*HI, *Eco*RI, *Hind*III, *Pst*I, *Sac*II (*Sst*II), *Sal*I, *Xba*I, *Xho*I (All 20 U/μL).

*Asp*718, *Bcl*I, *Nco*I, *Not*I, *Pvu*II, *Rsa*I, *Sma*I (All 10 U/μL).

*Bsr*I (4 U/μL), *Spe*I (3 U/μL), *Sph*I (5 U/μL).

All from New England Biolabs (NEB) except *Asp*718, *Bcl*I and *Sma*I from Boehringer Manheim (BM), *Pvu*II and *Rsa*I from Life Technologies (BRL).

2.1.3.2 Restriction Buffers

*Bam*HI buffer, *Eco*RI buffer, NEB 2, 3, 4 and *Sal*I buffer (all from NEB).

Buffer A, B and H (BM).

React 1 and 6 (BRL).

Table 2.1 Detailed list of oligonucleotide primers.

All primers supplied by Life Technologies (BRL), except 9474/9476 supplied by Dr Phil Crosier, Auckland Medical School and pgd3'/pgd5', ZW3/ZW5 from Scott *et al.* (1993). Abbreviations are as follows; R = A + G, Y = C + T, M = A + C, K = G + T, S = C + G, W = A + T, H = A+ C + T, B = C + G + T, D = A + G + T, V = A + C + G, N = A + C + G + T.

Oligonucleotide	Sequence 5'-3'	Use
9474	ACCTTAAGWCCMCGGTAGGTGAASTR MCC	PCR
9476	ACGAATTCCACCGNGAYYTGGCNGCN MG	PCR
#1DrorA	ATGAACAAATACTCGGCATTTATAGT	PCR for Dror
#1DrorB	TTACATTTCTGGATTACTGGCCTTA	PCR for Dror
DO1	CACCGNGAYYTGGCNGCNMG	PCR
DO1R	CKRGCRGCCARRTCNCGRTG	PCR
DO2	GGNGCCATCCACTTSAYNGG	PCR
DO3	ACDCCRAANGMCCAVACRTC	PCR
DO4	GTBTGYGTBGCNGAYTTYGG	PCR
DO5	GTNYGCNTNGCYGAYTTCGG	PCR
D06	GATCAACGAGCACATGGCG	PCR
DrorA1	CATAGGCGCAGAGCATT	5' RACE
DrorB1	CTTCGCACACAGACCCGTAAG	SEQUENCING & 5' RACE
DrorB2	CGTTTCCAAGCAAAATCCTTGCGA	SEQUENCING
Dror2B3	CCTTAGGGGTGGTTTCCTGTC	SEQUENCING
DrorB4	CCTTGCCGCGCATTTAGCAACC	SEQUENCING
DrorB5	GGCGGTAGTTAATGTTACTTTAT	SEQUENCING
DrorC1	GCCGTGGTCACCTGCTCGTTCT	5' RACE

Oligonucleotide	Sequence 5'-3'	Use
DrorZ1	TGAGTCGGATGTGTGGGCATACG	3' RACE
Dtk2	CCNTTYATGAAGCAYGGHGA	PCR
Dtk3	CATRATYTCCCACATNGTVAC	PCR
Eph GSP1	TCTAAAAGCAATGGCTTCC	LIBRARY SCREEN
For2	TCATGCTGGAGGTTTCGGAGGTGG	SEQUENCING
For3	GATGGTCTCCGGTAACTGGTTCCT	SEQUENCING
GSP1	CTCAAGAGAACTTGTTG	5' RACE, LIBRARY SCREEN
GSP2	GGGATGAAGTCGTTCTCATCG	5' RACE & SEQUENCING
GSP2A	CATACACTATATCCCCGCGTGG	SEQUENCING
GSP3	CCGCCATGTGCTCGTTGATC	5' RACE
KFP1	CCTGGCGTATTCATGCGCTTCT	SEQUENCING
KFP2	CACAGCAATTGCGTGGCCCTTAT	SEQUENCING
KFP3	GCAGTGCTGCCAGCGCACTGAT	SEQUENCING
KFP4	AGACTTACAGCTCACTTCGGAC	SEQUENCING
Pgd3'	GTRTGNGCNCCRAARTARTC	PCR CONTROL
Pgd5'	AARATGGTNCAYAAYGGNAT	PCR CONTROL
Rev2	GGAGTCCGAAGTGAGCTGTAAGTC	SEQUENCING
rp49 GSP1	TGACCATCCGCCCAGCATACAG	3' RACE
rp49 GSP2	GATCGTGAAGAAGCGCACCAAG	3' RACE
rp49 GSP3	CAAATGTGTATTCCGACCACGTTAC	3' RACE
Trk	GATCAACGAGCATGGCG	PCR, 3' RACE, SEQUENCING
Trk2	GAATACGCCAGGTTGGGTGGAA	SEQUENCING
UP	GAGAGAATTCGGATCCTCTAGAG	3' RACE
UP-dT	GAGAGAATTCGGATCCTCTAGAGTTTT TTTTTTTTTTTTTTT	3' RACE
ZW3	GCNAARAARAARATHTAYCC	3' RACE CONTROL
ZW5	ARRTGRTTYTGCATNACRATC	3' RACE CONTROL

Table 2.2 Strains, λ clones and Plasmids used in this study.

Markus Noll is at the Institut fur Molekular biologie dër Universität, Zürich, Switzerland.

Simon Greenwood is at the University of Auckland, New Zealand.

Catherine Day is at Massey University, Palmerston North, New Zealand.

Strain, λ clone or plasmid	Relevant characteristics/Genotypes	Source/Reference
Bacterial Strains		
Escherichia coli		
DH5α	*sup*E44 F$^-$ϕ80*lacZ* Δ M15 (*lac*ZYA$^-$*arg*F) *deo*R *hsd*R17 ($r_k m_k^+$) *pho*A *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1	Hanahan, 1983
K802	*sup*E *hsd*R *gal met*B	Wood, 1966
LE392	*sup*E44 *sup*F58 *hsd*R514 *gal*K2 *gal*T22 *met*B1 *trp*55 *lac*Y1	Borck *et al.*, 1976
XL-1 Blue	*sup*E44 *hsd*R17 *rec*A1 *end*A1 *gyr*A46 *thi rel*A1 *lac*$^-$ F'[*pro*AB$^+$ *lac*I^q *lac*Z ΔM15 Tn*10* (*tet*^{r})]	Bullock *et al.*, 1987
λ Libraries		
λZAPII	*Drosophila* 4-8 hour embryonic cDNA library	Markus Noll
λEMBL4	*Drosophila* genomic library	Markus Noll
Plasmids		
pUC118	3.2 kb derived from pUC18 Amp^R	Messing, 1983
pGEM T-vector	3.0 kb derived from pGEM5ZF(+) Amp^R	Promega
pBluescript	3.0 kb derived from pUC19 Amp^R	Stratagene
pBS KS II (-)	3.0 kb Amp^R	Stratagene
pBS SK (-)	3.0 kb Amp^R	Stratagene
pBS SK (-)	3.0 kb Amp^R	Stratagene
pRP49	640 bp rp49 in pBS KS (-)	Simon Greenwood
pHkb	0.6 kb Hkb fragment in pBS SK (-)	Simon Greenwood
pRas	1.6 kb Ras fragment in pBS SK (-)	Simon Greenwood
pArm	2.5 kb Arm fragment in 9 kb vector (unknown)	Catherine Day
pKFB4	3.6 kb Dror2 fragment in pUC118	This study
pKFBX3	3.3 kb Dror2 fragment in pBS KS (-)	This study
pDror	2.0 kb Dror fragment in pGem	This study

2.1.3.3 Other enzymes and components

Bovine Serum Albumin (BSA)	100X	(NEB)
Calf Intestinal Phosphatase (CIP)	10 U/μL	(NEB)
DNA polymerase I (Klenow)	5 U/μL	(NEB)
DNase	10 μg/μL	(Sigma)
eLONGase enzyme	1 U/μL	(BRL)
pGEM ligase	50 ng/μL	(Promega)
Reverse transcriptase	200 U/μL	(BRL)
RNase (DNase free)	0.5 μg/μL	(BM)
RNase	10 ug/uL	(Sigma)
T4 DNA Ligase	200 U/μL	(BRL)
Taq DNA polymerase	5 U/μL	(NEB)
Tfl DNA polymerase	1 U/μL	(Epicentre technologies)

2.1.4 AGAROSE GEL DNA STANDARDS

The 1 kb DNA ladder (BRL) was used as the molecular marker for sizing small fragments of DNA on agarose gels. For sizing larger fragments another molecular marker was made by digesting Lambda DNA (BRL) with *Hind*III and *Sst*II. In this marker the 4.4 kb band was only visible when heated to 65 °C for 5 minutes.

The 0.24 - 9.5 kb RNA ladder (BRL) was used to estimate the size of single-stranded RNA on formaldehyde agarose gels.

2.1.5 MOLECULAR BIOLOGY BUFFERS AND SOLUTIONS

2.1.5.1 Buffers

APH (Aqueous prehybridisation/hybridisation) solution

SSC	5X
Denhardt solution	5X
SDS	1 % (w/v)

Add 100 µg/mL denatured salmon sperm DNA just before use.

Blocking Solution (DIG)

1 % Blocking reagent dissolved in Maleic acid buffer.

Denaturation solution

NaCl	1.5 M
NaOH	0.5 M

Denhardt solution (100X) (1 litre)

Ficoll 400	20 g
Polyvinylpyrrolidone (PVP)	20 g
Bovine serum albumin (BSA)	20 g

Detection buffer (DIG) pH 9.5

Tris-HCl	100 mM
NaCl	100 mM

Formaldehyde loading buffer

EDTA (pH 8.0)	1 mM
Bromophenol blue	0.25 % (w/v)
Xylene cyanol	0.25 % (w/v)
Glycerol	50 % (v/v)

FPH (Formamide prehybridisation/hybridisation) solution

SSC	5X
Denhardt solution	5X
Formamide	50 % (v/v)
SDS	1 % (w/v)

Add 100 µg/mL denatured salmon sperm DNA just before use.

Loading Dye for agarose gels (10X)

Ficoll 400	20 % (w/v)
NaEDTA (pH 8.0)	0.1 M
Bromophenol blue	0.25 % (w/v)

Maleic acid buffer (DIG) pH 7.5

Maleic acid	100 mM
NaCl	150 mM

MOPS running buffer (10X)

MOPS (pH 7.0)	0.4 M
Sodium acetate	0.1 M
EDTA	0.01 M

Neutralisation solution pH 7.0

NaCl	1.5 M
Tris-Cl	0.5 M

Sequencing TBE (Tris borate) buffer (10X)

Tris base	1.3 M
Boric acid	4.4 mM
EDTA $(Na_2)_2$ H_2O	25 mM

SM (Suspension Medium)

NaCl	100 mM
$MgSO_4 7H_2O$ (pH 7.5)	10 mM
Tris-HCl (pH 7.5)	50 mM

SSC (Sodium Standard Citrate) (20X) pH 7.0

NaCl	3 M
Sodium citrate	300 mM
Gelatin	0.01 % (w/v)

STE (Sodium Tris EDTA) Buffer

Tris-Cl	10 mM
NaCl	10 mM
EDTA (pH 8.0)	1 mM

TAE (Tris acetate) electrophoresis buffer (50X)

Tris base	2 M
Glacial acetic acid	5.7 % (w/v)
EDTA (pH 8.0)	50 mM

TBE (Tris borate) electrophoresis buffer (10X)

Tris base	900 mM
Boric acid	900 mM
EDTA (pH 8.0)	20 mM

TE (Tris EDTA) buffer (10/1)

Tris-HCl (pH 7.5)	10 mM
EDTA (pH 8.0)	1 mM

TNE (Tris EDTA NaCl) buffer (10X)

NaCl	1 M
Tris-HCl (pH 7.5)	100 mM
EDTA (pH 7.4)	10 mM

Washing buffer (DIG) pH 7.5

Maleic acid	100 mM
NaCl	150 mM
Tween 20	0.3 % (v/v)

2.1.5.2 Other chemicals

DEPC (Diethylpyrocarbonate) treatment of solutions (BDH)

Add 0.1 % (v/v) DEPC to solution, shake vigorously for 2 minutes, leave overnight in fumehood, autoclave.

Glycine (Sigma)

Stock solution 20 mg/mL, final concentration 20 µg/mL.

Lysozyme (from chicken egg white) (Sigma)

Stock solution 25 mg/mL, final concentration 3.5 mg/mL.

Proteinase K (from Tritirachium album) (Sigma)

Stock solution 50 mg/mL stored at -20 °C, final concentration 50 µg/mL.

Salmon sperm DNA (Sigma)

Stock solution 10 mg/mL made according to Sambrook *et al.* (1989), final concentration 10 µg/µL.

2.1.6 BACTERIAL GROWTH MEDIA

2.1.6.1 Liquid media

Luria Broth (LB)	(Miller, 1972)
NZCYM	(Sambrook *et al.*, 1989)
SOB medium	(Dower *et al.*, 1988)
SOC medium	(Dower *et al.*, 1988)
Terrific Broth (TB)	(Tartof and Hobbs, 1987)
λ broth	(Ausubel *et al.*, 1997)

2.1.6.2 Solid media

LB Agar	(Miller, 1972)
LB Agarose	(Ausubel *et al.*, 1997)
LB Top Agar	(Ausubel *et al.*, 1997)
Top Agarose (0.7%) (w/v)	(Ausubel *et al.*, 1997)
TB Top Agarose	(Ausubel *et al.*, 1997)
λ agar	(Ausubel *et al.*, 1997)

2.1.7 ANTIBIOTICS AND MEDIA SUPPLEMENTS

2.1.7.1 Antibiotics for selection

Ampicillin

Stock solution 50 mg/mL stored at -20 °C, final concentration 50 μg/mL.

Tetracycline

Stock solution 12 mg/mL dissolved in 70 % ethanol, final concentration 12 μg/mL.

2.1.7.2 LB Agar supplements for selection

IPTG (Isopropyl-I-thio-β-D-galactoside)

Stock solution 23 mg/mL, final concentration 23 μg/mL.

X-gal (5-Bromo-4-chloro-3-indoyl-β-D-galactoside)

Stock solution 20 mg/mL dissolved in N,N dimethyl formamide. final concentration 20 μg/mL.

2.1.7.3 LB Media supplements

Overnight cultures of LE392, and K802 were grown in LB broth containing 10 mM $MgSO_4$ and 0.2 % (w/v) Maltose.

2.1.8 FLY MEDIA

Cornmeal Agar : 6.4 g agar (Davis), 56 g cornmeal (Healtheries) and 19.2 g dried yeast (NZ Food Industries) were stirred together with 600 mL water. The mixture was then boiled for 2 minutes. Seventy seven g sugar (Chelsea), 2 g p-hydroxy-benzoic acid methyl ester (Sigma) and 20 mL of 95 % ethanol were added and mixed thoroughly. The mixture was then poured into vials, bottles or petri dishes as required.

2.1.9 SEQUENCE ANALYSIS AND ALIGNMENTS

Analysis was carried out using the BLAST program from the National Center for Biotechnology Information (NCBI), FASTA and BLITZ programs from the European Bioinformatics Institute (EBI). Alignments were performed using the DNAsis program from Hitachi Software Engineering Genetics Systems Group, the ClustalW program from the European Molecular Biology Laboratory (EMBL), Seq Vu 1.1 from the Galvan Institute of Medical Research and GcG Winconsin Package Version 9.0 from Genetics Computer Group, Madison.

2.2 METHODS

2.2.1 PREPARATION OF GENOMIC DNA

To make genomic DNA, Sambrook *et al.* (1989) (Protocol 47) was followed. Approximately 500 adult flies were used in the method.

2.2.2 PREPARATION OF RNA

2.2.2.1 Isolation of Total RNA

Total RNA was originally prepared from Sambrook *et al.* (1989) (Protocol 49). Approximately 200 adult flies were ground to obtain the required amount of RNA. Subsequently Total RNA was isolated using the TRIzol reagent (Total RNA Isolation reagent - BRL). This reagent works via a monophasic mixture of phenol and quanidine isothiocyanate, reducing potential RNase denaturants. Ten to one hundred mg of flies were homogenised in TRIzol and RNA purified according to the manufacturers protocol. Resulting RNA was quantified using a spectrophotometer (Section 2.2.9.2) at 260 nm.

To collect embryos, approximately 100 flies were placed in a laying chamber on cornmeal agar food (Section 2.1.8) with yeast paste in the centre of the plate. After laying overnight at 25 °C, the flies were put onto new plates and then left for an hour before the plates were changed. This allowed for the release of any eggs the females were storing. The flies were then placed on new plates and allowed to lay for the specified time required. The embryos were collected promptly after this period had elapsed and homogenised immediately. The 3rd instar larvae and pupae were collected directly from bottles set up with cornmeal agar food, as were the adult flies.

2.2.2.2 Isolation of poly A^+ mRNA

The Oligotex Direct mRNA kit (Qiagen) was utilised to isolate and purify poly A^+ mRNA from total RNA. The Oligotex Direct mRNA protocol was followed and approximately 12.5 μg poly A^+ RNA was expected to be yielded.

2.2.3 cDNA SYNTHESIS

cDNA was synthesised using 3 μM of a specific primer and 10 μg RNA that was heated to 70 °C for 10 minutes to denature the RNA. Each reaction was set up containing 1X 1st Strand Buffer, 10 mM dTT, 0.5 mM dNTPs, 16 U reverse transcriptase enzyme and sterile H_2O added up to 50 μL. The reaction was then incubated at 37 °C for 2 hours, extracted with phenol:chloroform, precipitated with ethanol (Section 2.2.4) and resuspended in 50 μL TE.

2.2.4 PHENOL:CHLOROFORM EXTRACTIONS AND ETHANOL PRECIPITATION

A 1:1 (v/v) mixture of phenol (water saturated, buffered) and chloroform was added to an equal volume of the DNA solution. The tubes were shaken for 5 minutes then centrifuged for 5 minutes at 13793 x g (13000 rpm) in a Heraeus benchtop centrifuge. The top layer was pipetted into a fresh tube and an equal volume of chloroform was added. This was shaken for 5 minutes, centrifuged for 5 minutes and the top layer pipetted into a fresh tube.

One tenth of the volume of 3 M sodium acetate is added as was two volumes (2X the new volume) of 100 % ethanol. This was mixed by inversion and placed at -20 °C for 2 hours or -70 °C for 30 minutes. The solution was centrifuged at maximum speed for 15 minutes, the supernatant decanted and 70 % ethanol added. This was centrifuged 5 minutes, the supernatant decanted and the pellet was vacuum dried for 5-10 minutes. The pellet was then resuspended in an appropriate volume of TE.

2.2.5 AGAROSE GEL ELECTROPHORESIS

The DNA was size fractionated on 1.0 % to 1.7 % (w/v) agarose (Seakem le agarose - FMC) gels in TBE buffer. One tenth volume of loading buffer was added to each sample prior to loading. The DNA was size fractionated in a BRL Horizon 58 Minigel box (BRL) at 70 - 90 Volts at room temperature for 1-1.5 hours. Staining of the gel was performed in 0.5 μg/mL Ethidium Bromide (from a stock solution of 10 mg/mL) for 20 minutes. The bands were then visualised on a Gel Documentation Alphalmager 2000 system.

DNA fragment sizes were determined using a *Hind*III/*Sst*II λ ladder run alongside the samples. Larger gels were photographed beside a ruler so the mobility of the DNA could be measured accurately. Molecular weights of the DNA fragments were obtained using a graph of the molecular weight markers versus the relative mobility on semi-log paper.

2.2.6 POLYMERASE CHAIN REACTION

Each reaction was set up containing 0.5 μM $MgCl_2$, 0.2 mM dNTPs, 1X Taq polymerase buffer and 5 U Taq polymerase (BRL). One reverse and one forward primer were added at a 0.5 μM concentration. Template DNA was diluted to approximately 1 ng/μL and sterile water was added to make up a final volume of 20 μL. Degenerate oligonucleotide primers derived from conserved amino acid motifs within the tyrosine kinase domain of RTKs were used as listed in Table 2.3.

Negative controls containing all the components except the template DNA, were included in all reactions. A positive control containing primers pgd3' and pgd5' yielded a product of 807 bp after amplification by PCR.

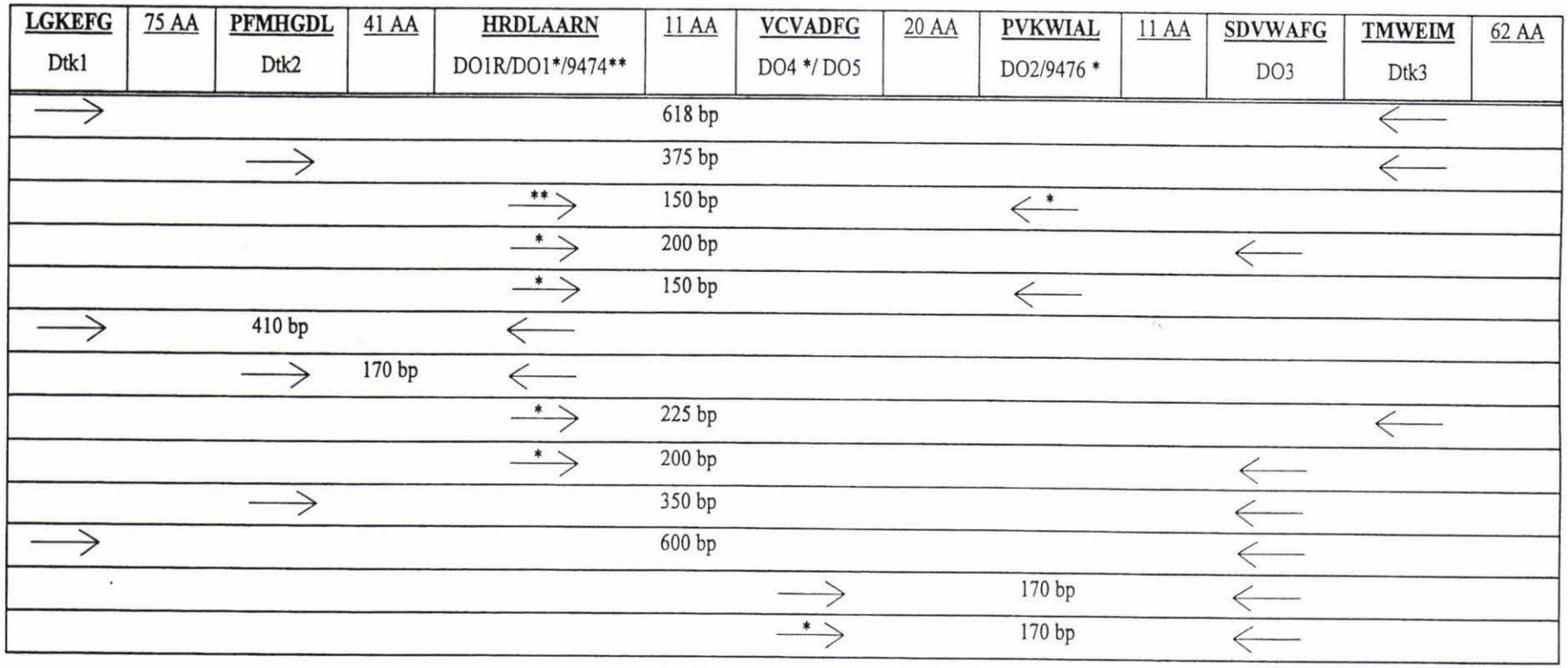
The samples were placed directly in a Perkin Elmer Gene Amp PCR System 9600 machine using the following cycling conditions.

Initial Denaturation	94 °C	1 minute	
Annealing	50 °C	1 minute	X 1
Extension	72 °C	1 minute	
Denaturation	94 °C	30 seconds	
Annealing	50 °C	45 seconds	X 30
Extension	72 °C	45 seconds	
Final extension	72 °C	5 minutes	
Store	4 °C		

The product was then size fractionated on a 1.7 % agarose gel in TBE buffer.

Table 2.3 Schematic diagram of primer combinations used in PCR strategy in relation to the kinase domain.

Underlined domains represent conserved regions between RTKs, from which the primers were selected. Arrows denote the primer direction, numbers denote size between the primers. Numbers underlined represent distance between each conserved region.



2.2.7 PURIFICATION OF PCR PRODUCTS

2.2.7.1 Wizard purification method

The Wizard PCR DNA Purification System (Promega) was used to purify fragments from low melting agarose gels. DNA was size fractionated on a 1.7 % Seaplaque agarose (FMC Bioproducts) gel in TAE buffer and the fragment excised. This was heated to 70 °C for 10 minutes, or until the gel was melted. The manufacturers protocol was then followed except for the addition of a 5 minute wait before adding the elution buffer (in this case TE). This step was added as it appeared the isopropanol needed time to evaporate, so not to carry isopropanol through into the eluted fragments. This purification worked sufficiently for larger fragments (300 bp - 3 kb), but the recovery of DNA dropped dramatically for anything smaller than 300 bp.

2.2.7.2 QIAquick purification method

Qiagen purification kits were used to purify DNA fragments obtained as this kit was efficient with both small and large fragments (100 bp - 10 kb). The other advantage this kit had over the Wizard preparation purification method was that gel extractions could be performed using either standard or low-melting gels in either TBE or TAE.

The Qiagen gel extraction kit was used to purify DNA fragments from agarose gels. The protocol was followed with an additional 1 minute centrifugation after the addition of the wash buffer to remove any residual buffer. The DNA was eluted in 30 µL of H_2O, then 3.3 µL of 10X TE was added to the eluted fragments.

If there was one amplification product after PCR, the Qiagen PCR purification kit was used to purify the fragment. This resulted in the removal of primers, nucleotides, polymerases and salts from PCR products.

2.2.8 SUBCLONING

2.2.8.1 pGEM T-vector

The pGEM T-vector (Promega) was used for subcloning of PCR products. The advantage of the pGEM plasmid over other plasmids is that pGEM contains a single 3' -T overhang at the insertion site. This improves the efficiency of the system because Taq amplification products have a 3' - A overhang. A 10 µL ligation was set up containing 1 µL T4 DNA Ligase 10X buffer, 1 µL pGEM T-vector, 1 µL T4 DNA Ligase and 7 µL insert DNA (between 6-9 ng for a 150 bp insert and 120 ng for a 2.0 kb insert). Ideally the DNA was at a 2-3 fold molar excess to increase the efficiency of the ligation. The ligation mixture was incubated at 18 °C overnight prior to transformation.

2.2.8.2 pUC118 ligations

Fragments from the phage were subcloned into pUC118 to facilitate gene sequencing. Prior to ligation, 6 µg of pUC118 was cut with *Bam*HI to linearise the vector. This was then treated with Calf Intestinal Phosphatase (CIP) by the addition of 10 U CIP and 5.5 µL CIP buffer to a 50 µL sample of linearised pUC118. This was incubated at 37 °C for 1 hour, followed by a 75 °C incubation for 10 minutes. The CIP treatment was necessary to remove 5' phosphates from the end so the vector could not religate onto itself. Following this a phenol:chloroform extraction and chloroform purification was undertaken as described in Section 2.2.4. The sample was ethanol precipitated and resuspended at 0.05 µg/µL.

After confirmation of DNA concentration on an analytical gel, the vector was ready for the ligation. The DNA to be inserted was cut with *Bam*HI and gel purified using the QIAquick gel extraction kit and quantified on an analytical gel. Upon establishing the concentration of the vector and insert, a ligation was prepared using 150 ng pUC118, 2 µL ligase buffer, 1 µL ligase, 500 ng of a 4.0 kb insert (2-3 fold molar excess) and H_2O up to 20 µL. This was incubated at 18 °C overnight.

2.2.9 DETERMINATION OF DNA CONCENTRATION

2.2.9.1 Determination by comparison with a Lambda ladder

A molecular weight ladder was size fractionated alongside an unknown amount of DNA. The DNA was quantified by comparison of the intensities of the ladder bands compared to the sampled DNA and an estimation made.

2.2.9.2 Spectrophotometric Determination

Concentration of the DNA was also determined by measuring the ultraviolet absorption of a sample of the DNA at 260 nm. From the analytical gel the dilution necessary to obtain an accurate measurement could be established. In calculating the concentration the parameters of $1E_{260}$ = 50 µg/mL DNA and $1E_{260}$ = 40 µg/mL RNA were used.

2.2.9.3 Fluorometer Determination

For measuring DNA samples between 100 ng/mL and 2000 ng/mL a dye solution was prepared fresh daily. This solution contained 1.0 µg/mL Hoechst 33258 solution and 1X TNE. The fluorometer was blanked using 2 mL of the working dye solution, and calibrated using 2 µL of a 100 µg/mL Calf Thymus DNA stock. Upon repeating the calibration of the fluorometer to check for reproducibility, the

DNA concentration was measured using 2 μL of sample DNA and 2 mL of working dye solution. If the reading was above the given parameters for accurate measurement of the DNA concentration (above 1000 on the fluorometer), the DNA was diluted as necessary.

2.2.10 TRANSFORMATIONS

2.2.10.1 DH5α cells

Transformations of DH5α were carried out using the Standard Transformation Protocol of Hanahan (1983). The protocol was followed exactly resulting in an average transformation efficiency of 1.0×10^8 colonies/μg pUC118 transformed.

LB Amp/IPTG/X-gal plates (Section 2.1.7.2) were used to utilise the blue/white screening of X-gal, the white colonies being the transformants. A control in which no DNA was added was also transformed to monitor contamination of the cells or components used in the protocol. For each transformation 2 ng of ligated DNA was added and 100 μL of each was spread onto one plate. To concentrate the rest of the cells the tubes of cells were centrifuged to pellet the remaining cells. Eight hundred μL of the supernatant was discarded and the remaining 100 μL was used to resuspend the pellet. This was then spread onto another plate. Ten pg of the pUC118 control plasmid was used to establish the transformation efficiency, 100 μL was spread onto a plate, as was 100 μL of a 1/10 dilution. The reason for this was to get an accurate quantification of the transformation efficiency.

2.2.10.2 Competent DH5α cells

To increase the efficiency of transformation, MAX Efficiency DH5α Competent Cells (BRL) were utilised. Instead of using the specified 100 μL of competent cells per transformation half this amount was used, as this amount was sufficient. Ten

pg of the supplied pUC19 control plasmid was used to establish transformation efficiency. Two ng of DNA from each ligation was used per transformation and the protocol was followed adjusting the amount of SOC according to the volume of cells used. Each transformation was plated as described in the above protocol and left overnight at 37 °C. The average efficiency of transformation was 1×10^9 transformants/μg pUC19.

2.2.11 PREPARATION OF PLASMID DNA

2.2.11.1 Alkaline lysis mini preparation method

The alkaline lysis mini preparation method described in Ausubel *et al.* (1997) (Unit 1.6) was followed and resulting DNA was resuspended in 30 μl. TE. After the DNA was isolated, a restriction digest was performed as described in Section 2.2.12.

2.2.11.2 Modified alkaline lysis method

The modified alkaline lysis and PEG precipitation procedure from Applied Biosystems Inc. (User Bulletin 18, October 1991) was followed to obtain pure enough DNA to sequence with their kit. The procedure was followed exactly and resulting DNA was quantified on a fluorometer (Section 2.2.9.3) and restriction digests performed to check purity of the DNA.

2.2.11.3 Quantum plasmid miniprep kit

The Quantum plasmid miniprep kit (Biorad) was followed to obtain pure DNA for automated sequencing. This method was efficient and the amount of DNA yielded was high.

2.2.12 RESTRICTION ENDONUCLEASE DIGESTS OF DNA

Restriction digests of DNA using restriction endonucleases were carried out according to manufacturers' guidelines using the following parameters:

- The DNA concentration was between 200 ng and 10 μg.
- Three units of enzyme were used for every μg of DNA and the final volume was at least 10X the volume of the enzyme, to keep glycerol concentration below one tenth.
- RNase (DNase free) was added at a final concentration of 0.5 ng/μL, to destroy any RNA in the digest.
- 10X BSA concentration was added to the digest if recommended for the enzyme.

The time of digest was usually 3 hours at 37 °C unless otherwise specified by the manufacturer. Digested DNA was size fractionated on an analytical gel to check the digest was complete.

2.2.13 SEQUENCING

2.2.13.1 Manual sequencing

DNA was sequenced using the Perkin Elmer AmpliCycle sequencing kit. Each reaction contained 100 fmol template, 1X cycling mix, 4 μM primer, 2 μCi α^{33}P-dCTP (Amersham Life Science) and H_20 to 6 μL. The samples were placed in a Perkin Elmer Gene Amp PCR 9600 machine preheated to 95 °C. Cycling conditions that were used is on the following page.

After thermal cycling, 4 μL of stop solution was added to each reaction, each reaction was denatured and loaded onto a 6 % polyacrylamide gel (7 M urea). The gel was prepared as specified in the AmpliCycle sequencing manual from a long

ranger gel mix (FMC Bioproducts). One times sequencing TBE was used as the running buffer and the gel was pre-run at 65 W for 30 minutes. Three µL of each reaction was loaded and electrophoresis was performed at 65 W for 90 - 120 minutes for a short run and 210 - 240 minutes for a long run. For the longer runs two copies of the sample were loaded, one at the start and one 120 minutes into the run.

Initial Denaturation	94 °C	1 minute	
Denaturation	94 °C	30 seconds	X 25
Annealing	65 °C	30 seconds	
Extension	72 °C	1 minute	
Final extension	72 °C	1 minute	
Cool	4 °C	Up to 45 minutes	

The gel was then dried onto 3MM paper at 80 °C and exposed against autoradiography film at room temperature from between 1 and 3 days. The resulting film was then manually checked and analysed.

2.2.13.2 Automated sequencing

Sequencing of the phage clone isolated was undertaken by the Waikato University sequencing facility and by the Massey University DNA analysis service. Primers were made to known sequences of the clone as well as using the pUC/M13 Forward and Reverse primers. Sufficient reactions were carried out until the entire *Dror2* gene had been sequenced from both directions. In all instances, double

stranded DNA was supplied at 200 ng/μL, PCR products at 20 ng/μL, λ DNA at 1 μg/μL and the primers at 0.8 pmol/μL. The sequences were run on a ABI prism 377 DNA sequencing machine and resulting sequences and electropherograms analysed.

2.2.14 LARGE SCALE DNA PREPARATIONS

2.2.14.1 Large scale $CsCl_2$ preparation of bacterial genomic DNA

This method was adapted from Sambrook *et al.* (1989) (Units 1.38, 1.42, 1.46) using a Caesium Chloride, Ethidium Bromide gradient. The final pellet was resuspended in 500 μL TE and quantified on a fluorometer and on an analytical gel to check the quality of DNA yielded. Restriction digests were then carried out on the DNA to check the purity of the DNA.

2.2.14.2 Qiagen Plasmid Maxi Preparation

The Qiagen Plasmid Maxi 500 kit was utilised in the large scale preparation of DNA. The protocol was followed according to manufacturers instructions. A yield of approximately 300 - 500 μg was achieved.

2.2.15 5' RACE

5' RACE (Frohman *et al.*, 1988) was attempted using the 5' RACE system for rapid amplification of cDNA ends instruction manual, Version 2.0 (BRL). The manufacturer protocols were followed exactly except that Taq polymerase was replaced by the eLONGase enzyme mix (BRL).

The samples were placed directly in a Perkin Elmer Gene Amp PCR System 9600 machine. The first round of PCR had the following cycling protocol:

Initial Denaturation	94 °C	45 seconds	
Annealing	55 °C	1 minute	X 1
Extension	68 °C	2 minutes	
Denaturation	94 °C	30 seconds	
Annealing	55 °C	45 seconds	X 30
Extension	68 °C	2 minutes	
Final extension	68 °C	5 minutes	
Store	4 °C		

The second round of PCR was identical, except the extension time of the 30 cycles was decreased to 1.5 minutes. Following the two rounds of synthesis 10 µL of each sample was size fractionated on a 1.7 % Agarose gel in TBE.

2.2.16 3' RACE

cDNAs for 3' RACE were made using the oligo (dT)-UP primer and using total RNA preparations of different stages of development. The synthesis of cDNA was identical to Section 2.2.3. A control PCR reaction was run to check the cDNA synthesis using two primers that amplified a specific region of rp49 DNA (O'Connell and Rosbash, 1984). In all reactions a negative control was added to check for contamination in any of the components.

The PCR reactions were carried out using the eLONGase enzyme (BRL). Two cocktails were made. Mix number 1 consisted of 400 nM Primer 1 and 2, 200 µM of each dNTP, template DNA, DEPC (Diethyl pyrocarbonate) treated H_2O up to 20 µL. Mix number 2 consisted of 4 µL Buffer A, 6 µL Buffer B (to give a final Mg^{2+}

of 1.6 mM), 2 U eLONGase enzyme, DEPC treated H_2O up to 30 µL. The two cocktails were then mixed thoroughly, to give a final volume of 50 µL per reaction.

The samples were placed directly in a Perkin Elmer Gene Amp PCR System 9600 machine. The following cycling conditions were used:

Initial Denaturation	94 °C	30 seconds	
Annealing	55 °C	1 minute	X 1
Extension	68 °C	3 minutes	
Denaturation	94 °C	30 seconds	
Annealing	55 °C	45 seconds	X 30
Extension	68 °C	3 minutes	
Final extension	68 °C	5 minutes	
Store	4 °C		

Between the two rounds of PCR, 5 µL of the reaction was diluted in 495 µL DEPC treated water, and 1 µL of each sample was used for the second round of PCR. The second round of PCR followed the same conditions as the first round, except the extension time was lowered to 2 minutes for the 30 cycles. A control 3' RACE reaction was carried out to check the conditions of the reactions using the rp49 primers again. Following the two rounds of synthesis 10 µL of each sample was size fractionated on a 1.7 % Agarose gel in TBE.

2.2.17 LIBRARY SCREENING

To calculate the titre of the λZAPII and λEMBL4 libraries, Ausubel *et al.* (1997) (Unit 1.11) was followed. A culture of *E. coli* (XL-1 for λZAPII, K802 for λEMBL4) was grown to saturation with supplements added as in Section 2.1.7.3. The library was plated out using serial dilutions and resulting plaques counted to quantitate the titre of each library.

Bacteriophage were plated onto agar plates at high density of 20,000 to 30,000 pfu per plate as described in Ausubel *et al.* (1997) (Unit 6.1), so the primary screen of the plaques could be carried out. All screening was done in duplicate (two filters per plate) to eliminate most false positives. The agar plates used were 150 mm in diameter, therefore 7 ml. of Top Agarose was needed per plate. Once the plaques had grown, they were transferred to nylon Hybond-N filters (Amersham Life Science), denatured, baked (see Section 2.2.18) and stored until needed. The filters were prehybridised, hybridised and washed using the same procedure as used in Southern Blots (see Section 2.2.18). The hybridisation was carried out in a glass dish in a 65 °C shaking water bath. The probe used was the subcloned 200 bp PCR fragment of Trk sequenced previously. This was done by linearising both inserts using *Sst*II for Trk and *Pst*I for Eph then performing a phenol:chloroform extraction and ethanol precipitation on each (Section 2.2.4). The probe DNA was added to a reaction containing GSP1 primer for Trk and Eph GSP1 primer for Eph and radioactively labelled (Section 2.2.18). The autoradiographs were left at -70 °C for between 1 and 7 days depending on the strength of the signal.

Upon obtaining positive plaques each region containing the positive plaques was excised as explained in Ausubel *et al.* (1997) (Unit 6.5). The phage titre was determined as described above. Seven hundred and fifty pfu were added per 82

mm diameter plate for the secondary screen of the positive clones. The secondary screen was carried out using the same conditions as to the primary screen. Well isolated positive plaques were excised and stored in 1 mL SM plus 1 drop of chloroform at 4 °C. These positive plaques were then titred again to determine the amount of phage/mL of SM.

2.2.17.1 λ DNA preparation number 1

λ phage DNA was originally prepared using an adapted plate lysate method from Sambrook *et al.* (1989). LB broth was supplemented with solutions in Section 2.1.7.3, inoculated with a single bacterial colony of K802 and incubated at 37 °C overnight. One hundred µL of this overnight culture was added to 100 µL of titred phage to give to give 10^5 pfu of phage per 88 mm LB agarose plate. The plates were left at 37 °C until confluent lysis occurred (usually about 8 hours). The plates were cooled to 4 °C, overlaid with 5 mL of SM buffer and left at 4 °C overnight. The buffer was collected and centrifuged for 10 minutes at 4 °C at 4500 x g (in a GSA rotor) to remove cellular debris. The resulting lysate was transferred to a fresh tube and stored at 4 °C until required.

DNase and RNase were added to the lysate at a final concentration of 1 µg/mL and incubated at 37 °C for 30 minutes. The phage were precipitated with an equal volume of PEG solution (20 % (w/v) PEG 6000, 2 M NaCl in SM buffer) and incubated on ice for at least 2 hours to precipitate the phage. The phage were pelleted at 8800 x g for 30 minutes at 4 °C, supernatant decanted off and the pellet resuspended in 0.5 mL of SM buffer. Added to this was 0.1 % SDS and 5 mM EDTA and the solution was incubated at 68 °C for 15 minutes.

A phenol:chloroform extraction was performed, the DNA was precipitated by the addition of an equal volume of isopropanol and incubated at -70 °C for 20 minutes.

The DNA was pelleted by a 5 minute centrifugation, washed in 70 % ethanol, dried under vacuum and resuspended in TE. Resulting DNA was quantified on a fluorometer and checked in restriction digests to analyse similarities between the λ DNA clones.

2.2.17.2 λ DNA preparation number 2

λ phage DNA was prepared using an adapted liquid lysate method from Ausubel *et al.* (1997). LB broth was supplemented with solutions in Section 2.2.7.3 inoculated with a single bacterial colony of LE392 and incubated at 37 °C overnight. One hundred µL of this culture was added to 100 µL of titred phage to give 10^6 phage/100 µL and incubated at 37 °C for 30 minutes. The phage mixture was transferred to 50 mL of NZCYM and vigourously shaken at 37 °C until confluent lysis occurred. Confluent lysis is indicated by the culture clearing and cell debris appearing in the mixture. A few drops of chloroform were added to lyse any remaining cells and the culture was shaken for another 10 minutes. The culture was harvested by centrifugation at 17,000 x g (in a GSA rotor) for 10 minutes, the lysate transferred to a fresh tube and stored at 4 °C until required.

DNase and RNase were added to the liquid lysate at a final concentration of 10 µg/mL and immediately incubated at 37 °C for 30 minutes. NaCl (to 0.5 M) and PEG 6000 (to 10 % (w/v)) were added and incubated on ice for 1 hour to precipitate the phage. The phage were pelleted at 4080 x g (in benchtop centrifuge) for 10 minutes at 4 °C, supernatant decanted off and the pellet resuspended in 1 mL of SM buffer. This solution was transferred into an eppendorf tube and centrifuged for 10 minutes at maximum speed (in a benchtop microcentrifuge) to pellet any remaining bacterial debris. The supernatant was transferred to a fresh eppendorf and proteinase K was added (to 0.1 mg/mL final) and incubated at 37 °C for 30 minutes.

A phenol:chloroform extraction was performed, vortexed for 20 minutes on an eppendorf platform and centrifuged for 5 minutes at maximum speed in a benchtop microcentrifuge. The phenol:chloroform extractions were repeated, followed by a chloroform extraction, the DNA was ethanol precipitated and resulting DNA resuspended in TE. Resulting DNA was then quantified on a fluorometer and checked in restriction digests to analyse similarities between the λ DNA clones.

2.2.17.3 Qiagen λ DNA midi preparations

The Mini preparation of λ DNA kit (Qiagen) was also used to prepare λ DNA. The preparation was followed exactly, except that after the addition of L3 buffer at Step 4 proteinase K was added to a final concentration of 0.2 mg/mL and incubated for 30 minutes at 55 °C. A 10 minute incubation at room temperature after the addition of isopropanol was added to facilitate precipitation of the DNA.

2.2.18 SOUTHERN BLOTS

The Southern Blotting protocol from Ausubel *et al.* (1997) (Unit 2.9.2) was followed exactly. A Sartolon Blotting Membrane (Sartorius) nylon filter was used. The Whatman 3MM filter paper wick method was performed instead of the sponge method, and the blot was left overnight. The following day the well positions were marked, orientation identified and the blot was rinsed in 2X SSC and allowed to dry. The blot was then sealed within 3MM filter paper and baked under vacuum for 2 hours at 80 °C. Following this the blot was wrapped in gladwrap and UV crosslinked for 1 minute with the DNA side down on the transilluminator and then stored at room temperature until needed.

Hybridisation was carried out using a radioactive probe as described in Ausubel *et al.* (1997) (Unit 2.10.2). Prehybridisation of the blot was at 65 °C in hybridisation tubes containing 50 mL of APH (Aqueous prehybrdisation/hybridisation) solution

for 3 hours in a rotating hybridisation oven. Labelling of the probe was performed by using random oligonucleotide-primed synthesis as described in Ausubel *et al.* (1997) (Unit 3.5.9). Each reaction contained 50 ng DNA, 50 μM 3 dNTPs (no dCTP), 1X DNA polymerase I (Klenow) buffer, 0.2 U/μL DNA polymerase I (Klenow), 0.4 μM primer, 50 μCi [α-^{32}P] dCTP 3000 Ci/mmol (Amersham Life Science) and H_2O up to 50 μL. The DNA labelling reaction was incubated for 1 hour at 37 °C before being purified using ProbeQuant G-50 Micro Columns (Pharmacia Biotech). These columns are designed for removing unincorporated nucleotides in the labelling reaction. The labelled probe was eluted from the G50 spin column, boiled for 5 minutes and subsequently placed on ice for 5 minutes. Following this, the labelled probe was added to between 5 and 10 mLs of APH solution in the hybridisation tube, then incubated at 65 °C in a rotating hybridisation oven overnight.

The following day the washes were undertaken as described in Ausubel *et al.* (1997) (Unit 2.10.3), these washes were altered slightly as shown below.

1. Two 5 minutes washes in 2X SSC/0.1 % (w/v) SDS at room temperature.
2. Two 5 minute washes in 1X SSC/0.1 % (w/v) SDS at room temperature.
3. Two 15 minute washes in 1X SSC/0.1 % (w/v) SDS at 65 °C (high stringency washes). The 1X SSC/0.1 % (w/v) was preheated to 65 °C before use.
4. Rinse in 2X SSC at room temperature, then blot off excess liquid and cover in gladwrap.

As all the Southern Blots undertaken had strong signals, the autoradiograph was left at - 80 °C for between 1 -16 hours, thawed out and developed according to Kodak.

2.2.19 NORTHERN BLOTS

The Northern Hybridisation protocol described in Ausubel *et al.* (1997) (Unit 4.9) was followed. Extra precautions were taken to prevent degradation due to contamination with RNase. The equipment used to run the gel, blotting equipment and hybridisation tubes were cleaned with DEPC treated water and sterilised for 6 hours at 180 °C. Gloves were used at all stages to prevent contamination with RNase present on hands.

The solutions making up the agarose gel were doubled as the gel apparatus used was a Biorad DNA subcell. Each sample contained 10 µg of total RNA or 2.5 µg of poly A^+ RNA and alongside the samples was 5 µg of the 0.24 - 9.5 kb RNA ladder (BRL). The gel was size fractionated at 100 V for 4 hours. The ladder was excised and stained overnight in 0.5 M Ammonium Acetate containing 0.5 µg/µL Ethidium Bromide, then destained. The gel was placed in deionised water for 5 minutes. The blot was set up as specified again using the Whatman 3MM filter paper wick method and left overnight to transfer. The membrane was rinsed, dried and cross-linked as described in Section 2.2.18.

2.2.19.1 Labelling a double stranded DNA radioactive probe

Hybridisation analysis was initially carried out using a double stranded DNA radioactive probe as described in Ausubel *et al.* (1997) (Unit 2.10.2.). Prehybridisation of the blot was at 42 °C in hybridisation tubes containing 50 mL of FPH (Formamide prehybridisation/hybridisation) solution for 3 hours in a rotating hybridisation oven.

Labelling of the probe was done using the Ready-To-Go DNA labelling kit (Pharmacia Biotech). Each labelling mix contained the 20 µL reaction mix, 50 ng

denatured DNA, 50 µCi [α-^{32}P] dCTP (Amersham Life Science) and H_2O up to 50 µL. The labelling mix was incubated for 1 hour at 37 °C before being purified using ProbeQuant G-50 Micro Columns (Pharmacia Biotech), as described in Section 2.2.18. Following this the labelled probe was added to between 5 and 10 mLs of EPH solution in the hybridisation tube and incubated 42 °C in a rotating hybridisation oven overnight.

The following day the washes were undertaken as described in Ausubel *et al.* (1997) (Unit 4.9.7). these washes were altered slightly as shown below.

1. Two 5 minutes washes in 2X SSC/0.1 % (w/v) SDS at room temperature.
2. Two 5 minute washes in 1X SSC/0.1 % (w/v) SDS at room temperature.
3. One 15 minute wash in 1X SSC/0.1 % (w/v) SDS at 65 °C (high stringency wash). The 1X SSC/0.1 % (w/v) was preheated to 65 °C before use.
4. One 15 minute wash in 0.5 SSC/0.1 % (w/v) SDS at 65 °C (high stringency wash). The 0.5X SSC/0.1 % (w/v) was preheated to 65 °C before use.
5. Rinse in 2X SSC at room temperature. then blot off excess liquid and cover in gladwrap.

Exposures were from between 1 day and 3 weeks at - 80 °C and developed according to Kodak.

2.2.19.2 Labelling a single stranded RNA probe using DIG

DIG labelled RNA probes were constructed as in Section 2.2.20.2 and hybridisation was performed at 58 °C in a rotating hybridisation oven. Washes were identical to Section 2.2.20.1. However, detection followed the protocol for detection of chemiluminescence detection described in the DIG System User's Guide for Filter Hybridisation (BM, 1995).

2.2.19.3 Labelling two single stranded DNA radioactive probes

To make single stranded DNA probes, plasmid KFB4 was digested separately with *Bam*HI and *Xba*I restriction enzymes. The digested DNA was phenol:chloroform extracted, ethanol precipitated and resuspended in an appropriate volume of TE. Two reactions were set up each containing 50 ng of cut DNA and 0.4 μM primer. The *Bam*HI digested DNA was radiolabelled with 50 μCi [α-^{32}P] dCTP using the KFP4 primer and the *Xba*I digested DNA radiolabelled using the KFP3 primer. Both these labelled probes were added to the same hybridisation reaction to increase specificity to the blot. All steps after the labelling step were identical to Section 2.2.19.1.

2.2.20 *IN SITU* HYBRIDISATION

2.2.20.1 *In situ* hybridisation using DIG labelled DNA probes

Whole mount *in situ* hybridisation to *Drosophila* embryos were carried out using a modified Tautz and Pfeifle (1989) protocol supplied by Dr Simon Greenwood, University of Auckland. Labelling of the probe was accomplished using the DIG High prime kit (BM).

The probes were checked by Southern Analysis by digesting pKFB4 with *Bam*HI/*Pst*I, and pHkb with *Sal*I/*Xba*I, size fractionated on an agarose gel and transferred to a nylon membrane. The blot was hybridised with the DIG labelled probe overnight at 42 °C in hybridisation buffer. Washes were carried out as described and then colorimetric detection with NBT and BCIP as described in the DIG System User's Guide for Filter Hybridisation (BM, 1995).

2.2.20.2 *In situ* hybridisation using DIG labelled RNA probes

In situ hybridisation using DIG labelled RNA probes were carried out using a modified Jiang *et al.* (1991) protocol supplied by Dr Simon Greenwood, University of Auckland. Labelling of the probe was undertaken using the DIG RNA Labelling Kit (BM). Prior to this, the DNA was cut using a restriction endonuclease, purified by phenol:chloroform extraction (Section 2.2.4), followed by an ethanol precipitation.

3. RESULTS

3.1 PREPARATION OF cDNA AND GENOMIC DNA

Three different DNA preparations were utilised as templates for PCR.

3.1.1 EMBRYO cDNA AND ADULT cDNA

Adult cDNA was synthesised from RNA isolated from adult flies by following Section 2.2.3, using 10 µg of RNA and the oligo (dT) primer (BRL). This method of RNA isolation worked well for adult flies, but not for any other stage of development, therefore a new protocol was utilised to isolate RNA from embryos. Embryos 3-6 hours old were collected and RNA isolated using TRIzol reagent following Section 2.2.2.1. Embryo cDNA was synthesised using the same procedure to adult cDNA synthesis.

3.1.2 GENOMIC DNA

Genomic DNA was prepared, following Section 2.2.1. The DNA was at a final concentration of 0.2 µg/µL.

3.2 AMPLIFICATION OF RTK FRAGMENTS FROM *DROSOPHILA*

3.2.1 PRIMARY PCR

In the primary PCR, combinations of primers (see Table 2.3) were amplified to yield products 150 - 620 bp as indicated. The primary PCR was size fractionated on a 1.7 % Agarose gel in TBE using 10µL of each sample as shown in Figure 3.1.

Figure 3. 1 Primary PCR product of a 200 bp fragment resulting from genomic DNA, embryo cDNA and adult cDNA templates using DO1 and DO3 oligonucleotide primers.

Lane 1: 1 kb ladder, lane 2: genomic DNA, lane 3: embryo cDNA, lane 4: adult cDNA, lane 5: no DNA control.

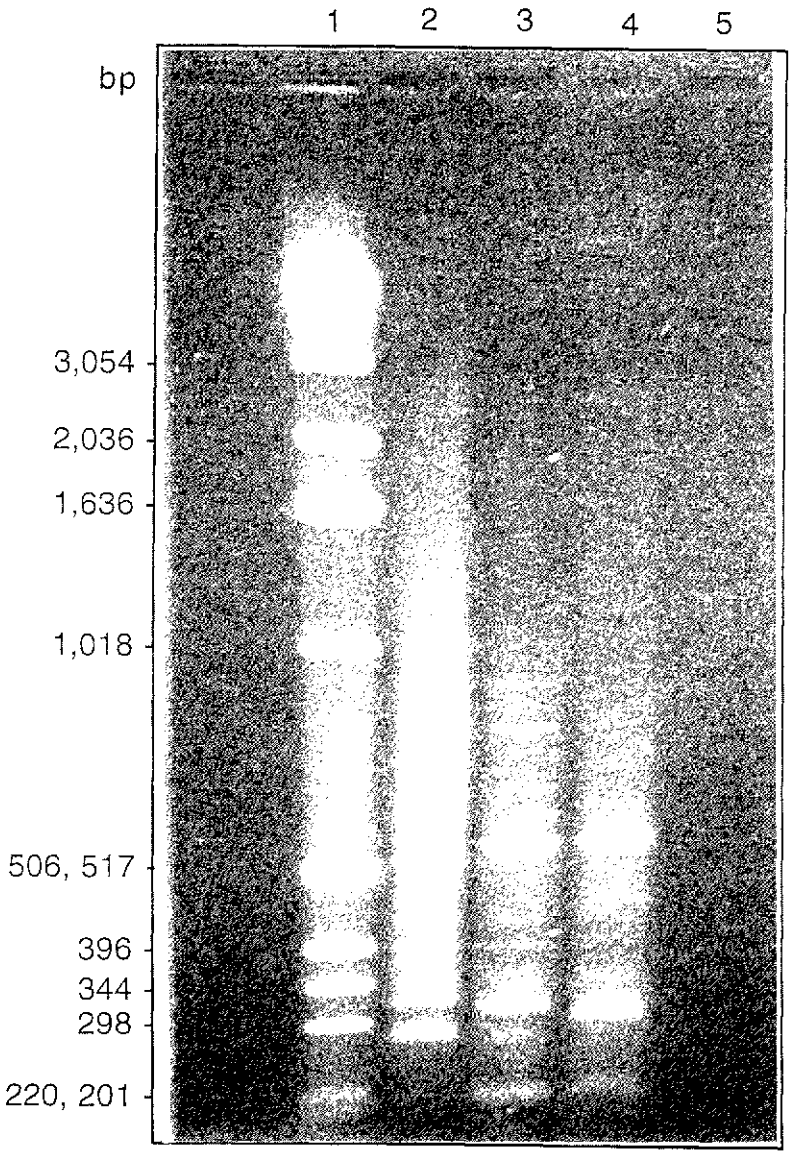
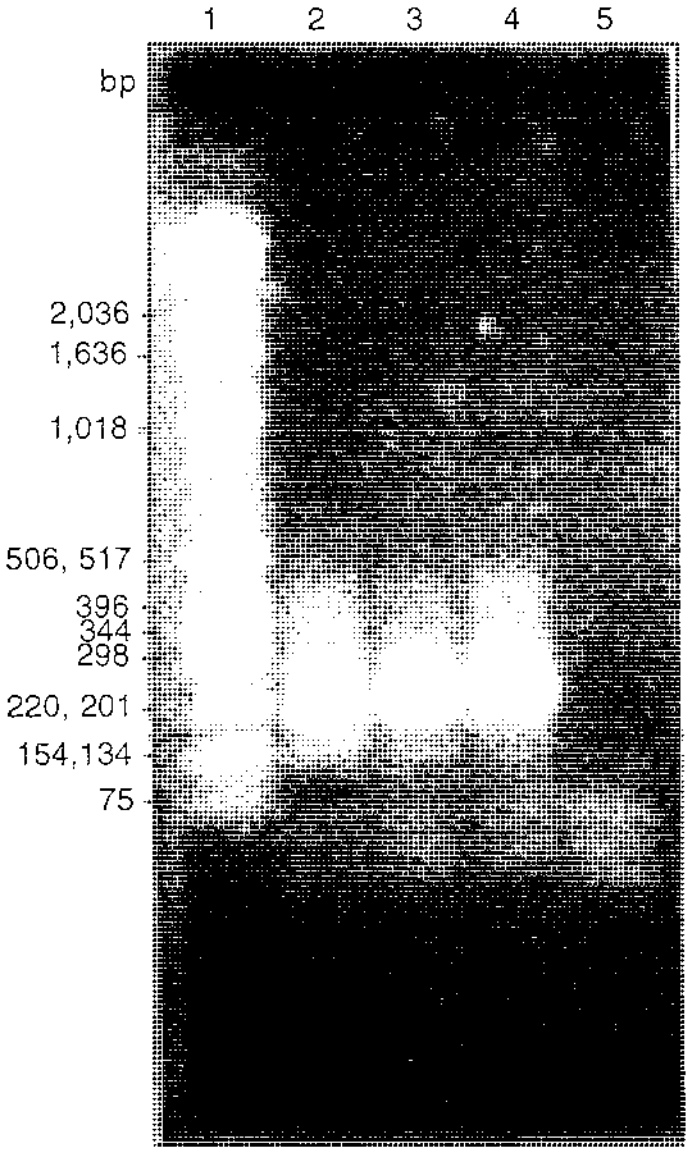

1
2
3
4
5
bp
3,054
2,036
1,636
1,018
506, 517
396
344
298
220, 201

Figure 3. 2 Secondary PCR product of a 200 bp fragment resulting from genomic DNA, embryo cDNA and adult cDNA templates using DO1/DO3.

Lane 1: 1 kb ladder, lane 2: reamplified genomic DNA, lane 3: reamplified embryo cDNA, lane 4: reamplified adult cDNA, lane 5: no DNA control.


1
2
3
4
5
bp
2,036
1,636
1,018
506, 517
396
344
298
220, 201
154,134
75

3.2.2 PURIFICATION OF PRIMARY PCR PRODUCTS

The Wizard PCR DNA Purification System (Section 2.2.7.1) was initially used to purify the PCR products obtained. After yielding no DNA from this method, it was discovered that recovery of DNA less than 300 bp in size was minimal. As the majority of fragments were less than 300 bp, a new purification method was needed.

The Qiagen purification kits (Section 2.2.7.2) were used to purify the PCR products. For PCR reactions that yielded one clean fragment, the PCR purification kit was used to remove primers, nucleotides, polymerases and salts. The resulting fragment was then reamplified as explained in Section 2.2.3. For PCR reactions that yielded more than one fragment, the required fragment was excised from the gel and purified using the gel extraction kit. The resulting purified fragment was then reamplified as described below.

3.2.3 SECONDARY PCR

After purifying the required fragments, a second round of PCR was performed. For each reaction 1 µL of purified template was added, using the same primers as for the primary PCR reaction. An aliquot (4-10 µL) of the secondary PCR products was resolved on a 1.7 % agarose gel as shown in Figure 3.2. The remainder of the samples were desalted (Section 2.2.7.2) and directly added to the ligation reactions (see below).

3.3 CLONING AND SEQUENCING OF RTK FRAGMENTS

3.3.1 CLONING FRAGMENTS INTO THE PGEM VECTOR

Upon obtaining fragments of interest, each fragment was ligated with the pGEM T-vector (Section 2.2.8.1) at 18 °C overnight. The ligated vectors were transformed into DH5α cells using the standard transformation protocol (Section 2.2.10.1). However, since only a few colonies were obtained, the transformation was repeated using MAX Efficiency DH5α Competent Cells (Section 2.2.10.2). Plasmid DNA was isolated from overnight growths of white colonies and analysed by digestion with restriction enzymes which flank the cloning site (usually *Sst*II and *Pst*I).

3.3.2 MANUAL SEQUENCING AND ANALYSIS OF SEQUENCES

In total, plasmid DNA was isolated from 133 clones. The subcloned fragments were obtained from a number of different PCR reactions which utilised a variety of primer combinations. For a detailed table of these clones see Appendix. All of the plasmid DNA from these clones was manually sequenced (Section 2.2.13.1) using the M13 forward primer supplied with the Amplicycle sequencing kit. Usually a short run was all that was required to obtain 250 bp clear sequence, but for longer fragments a long run was required to obtain approximately 500 bp of clear sequence.

The sequence of each clone was entered into the BLAST sequence similarity search program to determine if it corresponded to any known genes. If no similarity to any genes was found another run using the M13 reverse primer was undertaken, to obtain accurate sequence from sequencing in both directions. The improved sequence was then used for a BLAST search. If no similar genes were

identified, the amino acid sequence was deduced and entered in the FASTA similarity search. This search revealed the 50 closest protein sequences and from this evidence selected genes were followed further. A summary of the genes sequenced is shown in Table 3.1, the numbers represent the number of clones sequenced per gene.

In all, fragments of 4 possible new genes were isolated from three different template DNAs (see Figure 3.3). The possible Trk gene and Eph genes were chosen for further investigation.

3.3.3 LARGE SCALE PREPARATION OF CLONED FRAGMENT

A large scale preparation of the 200 bp Trk fragment in pGEM was performed resulting in 0.67 μg/μL of DNA by quantitating it on a spectrophotometer. The large scale preparation of the 200 bp Eph fragment yielded 2.50 μg/μL of DNA. The quality of DNA was examined by digesting 1 μg of each DNA with *Pst*I and *Sst*II (enzymes flanking the cloning site in pGEM) and checking it on an analytical gel.

3.3.4 PCR FOR cDNA CLONES OF TRK

Unlike Eph, the Trk fragment was only obtained from the PCR reaction which used genomic DNA as a template. Therefore, in order to determine if the putative Trk gene was expressed in either adults or embryos, PCR reactions were carried out using primers specific to the Trk gene and embryo or adult cDNA templates (Figure 3.4). Cycling times and temperatures were identical to Section 2.2.6, except the annealing temperature was raised to 52 °C. The results indicate that the putative Trk gene was expressed in both adults and embryos.

Table 3.1 Summary of genes obtained from sequencing cloned fragments.

Genes	Primers	Embryo cDNA	Adult cDNA	Genomic DNA	TOTALS
Non RTKs	DO1/DO3	19	16	26	61
	DO1/Dtk3	3	0	9	12
	DO1R/Dtk1	2	0	0	2
	DO3/Dtk2	0	0	2	2
	DO4/DO3	1	0	0	1
Drosophila TKs					
Dsrc28C	DO1/DO2	3	1	0	4
	DO1/Dtk3	3	0	0	3
	Dtk1/Dtk3	7	0	0	7
	DO1/DO3	2	0	0	2
hopscotch	DO1/DO3	0	2	0	2
Drosophila RTKs					
D-FGF	DO1/DO2	2	0	2	4
D-ror	DO1/DO3	0	1	0	1
D-ret	DO1/DO3	0	1	0	1
New genes					
FAK-like	DO1/DO3	2	4	1	7
	DO4/DO3	0	5	0	5
Srk-like	DO1/DO3	0	2	0	2
Trk-like	DO1/DO3	0	0	1	1
Eph-like	DO1/DO3	2	2	0	4
	DO4/DO3	0	2	0	2
Problem sequences (unable to sequence)	DO1/DO3	3	0	0	3
	DO1/Dtk3	1	0	0	1
	Dtk1/Dtk3	3	1	1	5
	DO4/DO3	1	0	0	1
TOTALS		54	37	42	133

Figure 3. 3 Template DNA used to isolate four new putative tyrosine kinase genes from *Drosophila*.

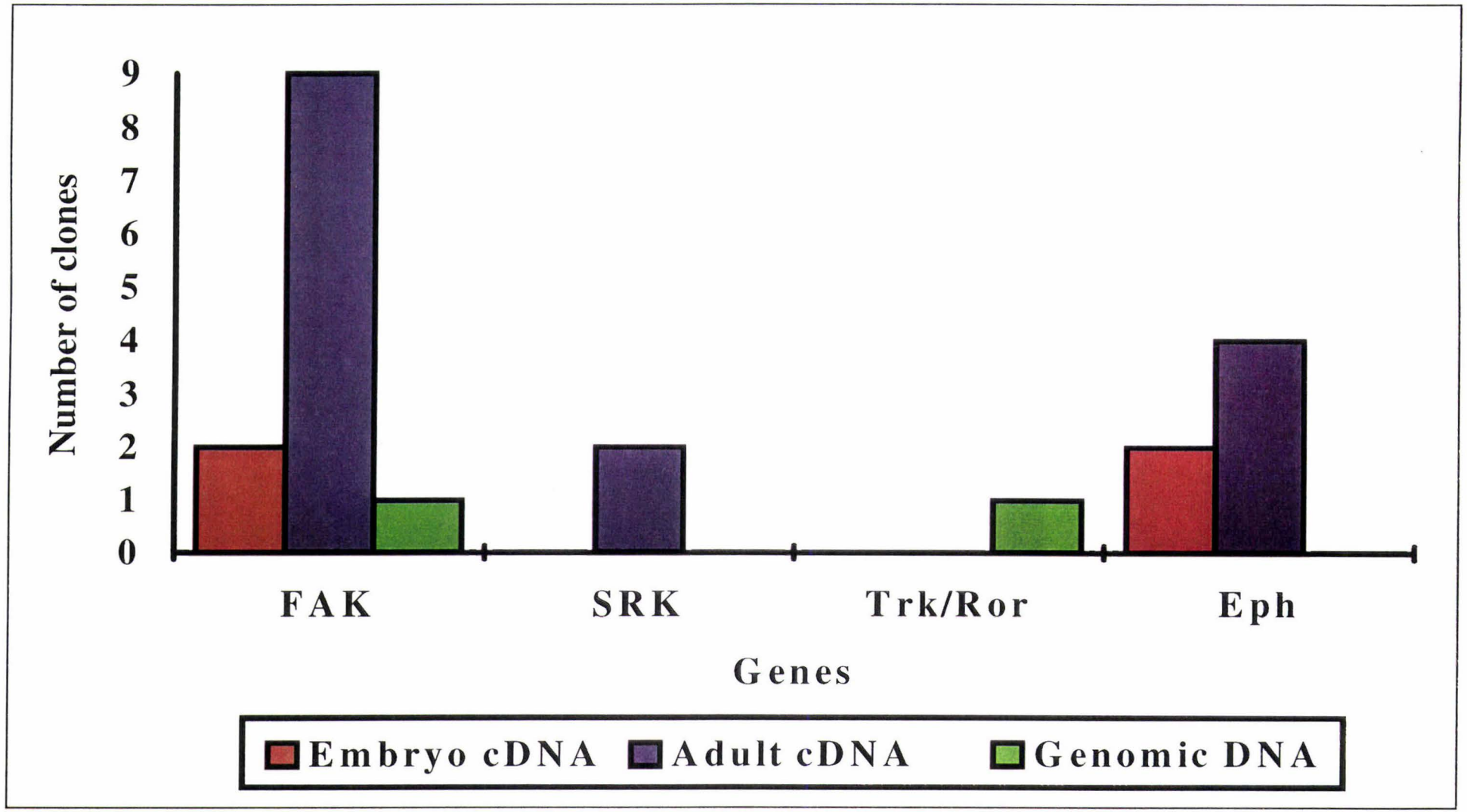
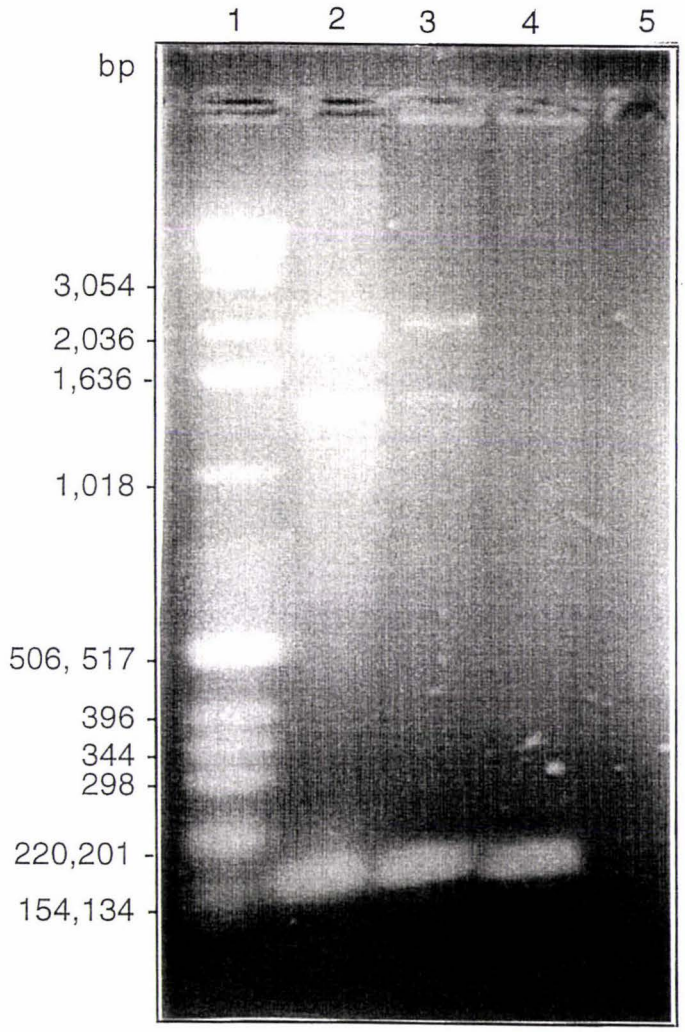

Number of clones
9
8
7
6
5
4
3
2
1
0
FAK
SRK
Trk/Ror
Eph
Genes
Embryo cDNA
Adult cDNA
Genomic DNA

Figure 3. 4 PCR products resulting from genomic DNA, embryo cDNA and adult cDNA templates using Trk/GSP1 primers.

Lane 1: 1 kb ladder, lane 2: genomic DNA, lane 3: embryo cDNA, lane 4: adult cDNA, lane 5: no DNA control.


1
2
3
4
5
bp
3,054
2,036
1,636
1,018
506, 517
396
344
298
220,201
154,134

3.4 LIBRARY SCREENING

3.4.1 CONSTRUCTING THE PROBES

To screen the cDNA and genomic libraries, probes were constructed using the sequenced Trk and Eph fragments (Section 2.2.17). A Southern Blot was performed to check for probe specificity, using the Eph and Trk fragments (in pGEM cut either side of the cloning site) as the DNA. The probes as hybridised strongly to their respective 200 bp fragments as expected.

3.4.2 SCREENING A cDNA LIBRARY

A *Drosophila* λZAPII embryonic 4-8 hour cDNA library was screened with Trk and Eph specific probes (Section 2.2.4). Four positive plaques were obtained from a primary screen of 100,000 plaques with the Trk probe. However, the secondary screen resulted in no positives, meaning all 4 were likely to be false positive plaques.

A primary screen of the λZAPII library using the Eph probe resulted in no positive plaques. At this point it was decided to continue with the cloning of the *Trk* gene and leave the cloning of the *Eph* gene.

3.4.3 SCREENING A LAMBDA GENOMIC LIBRARY

As the screening of the cDNA library was unsuccessful a λEMBL4 genomic library was screened with the Trk specific probe. A primary screen of 20,000 plaques yielded 21 positive plaques, from which 13 were picked and 6 of these titred. The secondary screen showed 1 of the initial 6 to be a false positive plaque. One distinct single plaque off each of the remaining 5 plates was picked, titred and stored in SM.

3.4.4 SMALL SCALE PHAGE DNA PREPARATIONS

Since the λEMBL4 genomic library had been amplified, it was necessary to determine if all five phage clones were identical. Small scale phage DNA preparations of each of the 5 clones had to be carried out to extract enough DNA to compare restriction digestion patterns. Initially a λ phage DNA preparation involving lysis off agarose plates was attempted (Section 2.2.17.1). Unfortunately, confluent lysis did not occur in the preparations attempted. However, a second method involving liquid lysis (Section 2.2.17.2) was more successful, although yields were poor.

3.4.5 RESTRICTION ANALYSIS OF THE ISOLATED CLONES

A restriction analysis was performed on the 5 isolated clones (referred to as A, B, C, D and E). Restriction digests containing 1 µg DNA were undertaken using *Eco*RI, *Bam*HI and *Sal*I alone, double digests of *Bam*HI/*Sal*I, *Sal*I/*Eco*RI and a triple digest of *Eco*RI/*Bam*HI/*Sal*I. This initially revealed that clones A and D were identical to B, but clones B, C and E may be different.

To determine which fragments in clones B, C and E contained the putative Trk gene, the phage DNAs were digested with various restriction enzymes and subjected to agarose gel electrophoresis (Figure 3.5 showing B and C). This gel was transferred to a nylon membrane and hybridised with a Trk specific probe (Section 2.2.4). Hybridisation occurred to one band in each lane indicating the fragment that contained the gene (Figure 3.6). Identical restriction digests hybridised to the Trk probe for clones B, C and E. These results indicated that the initial digests were misleading and that B and C were identical but that E was a mixture of two phages, one of which was identical to B. The fragment sizes for each digest performed on clone B can been seen in Table 3.2. Clone B was

selected for further analysis as the DNA preparation seemed to be of a high quality (as judged by restriction digests).

3.4.6 LARGE SCALE PREPARATION OF λ CLONE B

A large scale preparation of clone B was undertaken using the Qiagen preparations of λ DNA kit (Section 2.2.17.3). The DNA was quantitated on a fluorometer and restriction digests carried out to check the purity of the preparation.

As the restriction map of the phage had not yet been accurately calculated more restriction digests were carried out in an attempt to establish the map. Single digests of *Eco*RI, *Sac*II, *Sal*I and *Xba*I were performed as were double digests of *Bam*HI/*Xba*I, *Eco*RI/*Xba*I, *Sal*I/*Xba*I. Fragment sizes for each digest is in Table 3.2. The agarose gel showing the resulting fragments can be seen in Figure 3.7. From this a more accurate map could be calculated which can be seen in Figure 3.8.

3.5 SUBCLONING AND SEQUENCING DROR2

3.5.1 SUBCLONING INTO PUC118

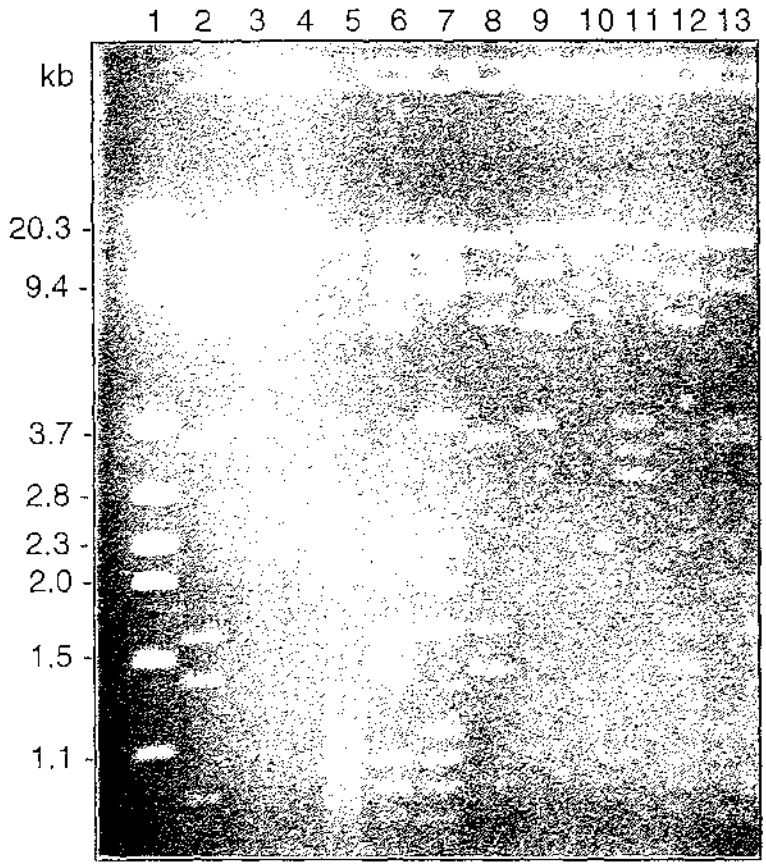
Southern analysis revealed that the 200 bp fragment of Trk was contained within the 3.6 kb *Bam*HI restriction fragment. To enable sequencing of this fragment the 3.6 kb fragment had to be subcloned into a plasmid vector containing the M13 Forward and Reverse primer sites. pUC118 contains these sites, therefore the 3.6 kb fragment was subcloned into the *Bam*HI cloning site (Section 2.2.8.2). Following the ligation of the vector and insert (referred to as plasmid KFB4) a transformation using competent cells (Section 2.2.10.2) and subsequent mini preparations (Section 2.2.11.1) were performed. A large scale preparation (Section 2.2.14.1) of one clone, plasmid KFB4 was carried out at 4.45 μg/μL.

Figure 3. 5 Restriction digestion profiles of λ clones B & C.

Lane 1: λ ladder, lanes 2-7: Clone B, lanes 8-13: Clone C, lanes 2 & 8: *Eco*RI digest, lanes 3 & 9: *Bam*HI digest, lanes 4 & 10: *Sal*I digest, lanes 5 & 11: *Bam*HI/*Sal*I double digest, lanes 6 & 12: *Eco*RI/*Sal*I double digest, lanes 7 & 13: *Bam*HI/*Eco*RI/*Sal*I triple digest.

Figure 3. 6 Autoradiograph of Southern blotted restriction digested DNA probed with 200 bp Ror/Trk radioactively labelled probe.

Lanes as in Figure 3.5, except lane 1 that was removed from the gel before blotting.


1 2 3 4 5 6 7 8 9 10 11 12 13
kb
20.3
9.4
3.7
2.8
2.3
2.0
1.5
1.1

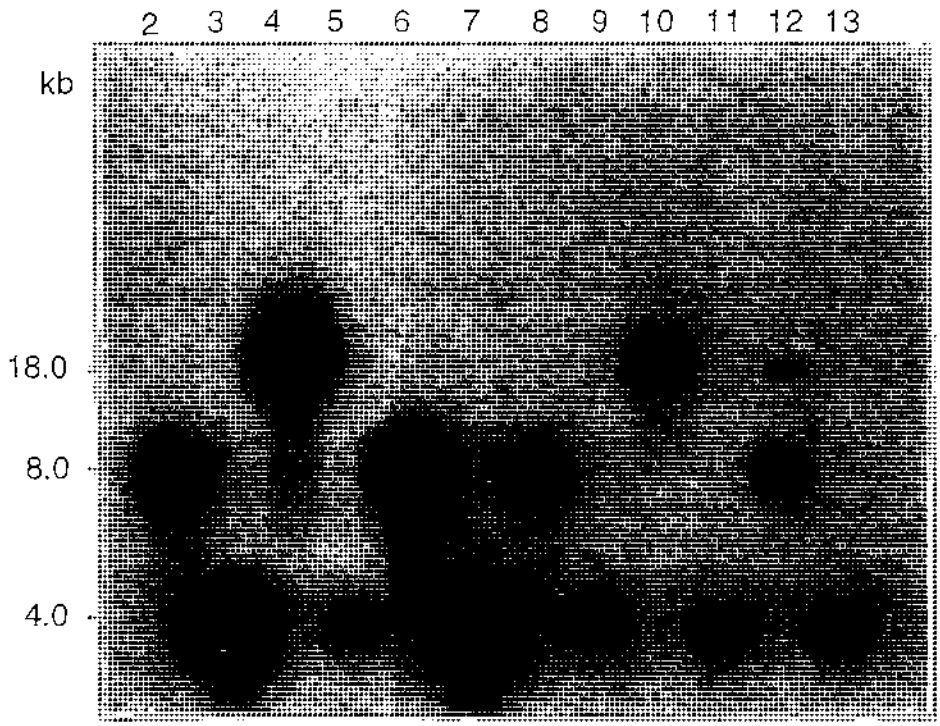

2 3 4 5 6 7 8 9 10 11 12 13
kb
18.0
8.0
4.0

Table 3.2 Restriction digest fragment sizes for clone B with *Eco*RI, *Bam*HI, *Sal*I, *Xba*I and *Sac*II.

Left arm of phage is 19.9 kb, right arm is 8.8 kb. The band that hybridised to the Trk probe is indicated with a asterisk (*).

Restriction enzyme	Fragment sizes (kb)
*Eco*RI	20.0, 8.8, 6.8*, 3.5, 1.4, 0.9, 0.8, 0.3, 0.3
*Bam*HI	19.0, 12.0, 6.6, 3.6*
*Sal*I	23.0, 18.0*
*Bam*HI/*Eco*RI/*Sal*I	20.8, 9.0, 3.6*, 2.0, 1.7, 1.2, 1.0, 0.9, 0.8, 0.3, 0.3
*Bam*HI/*Sal*I	20.5, 13.0, 4.3, 3.6*, 3.1
*Sal*I/*Eco*RI	20.0, 9.2, 6.8*, 3.6, 2.6, 1.7, 1.4, 1.0, 0.9, 0.8, 0.3, 0.3
*Xba*I	26.0, 13.0, 3.3
*Xba*I/*Bam*HI	20.0, 9.0, 3.6, 3.3, 3.3
*Xba*I/*Eco*RI	20.0, 8.9, 5.0, 2.1, 0.9, 0.8, 0.4, 0.3, 0.3
*Sac*II	24.9, 10.0, 7.6, 2.7

3.5.2 SUBCLONING INTO PBS

Plasmid KFB4 did not contain the 5' region of *Dror2*. One of the 3.3 kb *BamH*I/*Xba*I restriction fragments contained this region, so it was subcloned into pBS using the same procedures as Section 3.5.1 and named plasmid KFBX3. A Biorad mini preparation (Section 2.2.11.3) was carried out and DNA yielded at 0.75 μg μl^{-1}.

3.5.3 SEQUENCING PKFB4

The plasmid KFB4 was sequenced in both directions using an automated sequencer. Initially, pKFB4 was sequenced in from either end using the M13 forward and reverse primers and sequenced out from the 200 bp fragment already sequenced. From the sequences obtained, it was revealed that the gene isolated was closest to the Ror family of RTKs not the Trk family as previously indicated. This gene was nominated *Dror*2 as the first Ror gene in *Drosophila Dror* had already been isolated. Further sequencing reactions were carried out using primers derived from the ends of these sequences until an alignment of each strand was found, as can be seen in Figure 3.9.

3.5.4 SEQUENCING DROR2

Sequencing of plasmid KFB4 showed it to contain the 3' end of *Dror2* but not the 5' end, consequently primers had to be constructed to sequence through the phage to obtain the beginning of the gene. With the help of known sequence at the start of KFB4, primers were constructed and the phage sequenced. Following this, more primers could be constructed to sequence both strands of *Dror*2 including sequence upstream of the 5' UTR and downstream of the 3' UTR. Plasmid KFBX3 contained the 5' end of *Dror2* and was also sequenced to ensure accurate sequencing. A summary of the primer locations is shown in Figure 3.10.

Figure 3. 7 Restriction digestion profiles of λ clone B

Lane 1: λ ladder, lane 2: *Eco*RI digest, lane 3: *Sac*II digest, lane 4: *Sal*I digest, lane 5: *Xba*I digest, lane 6: *Bam*HI/*Xba*I double digest, lane 7: *Eco*RI/*Xba*I double digest.

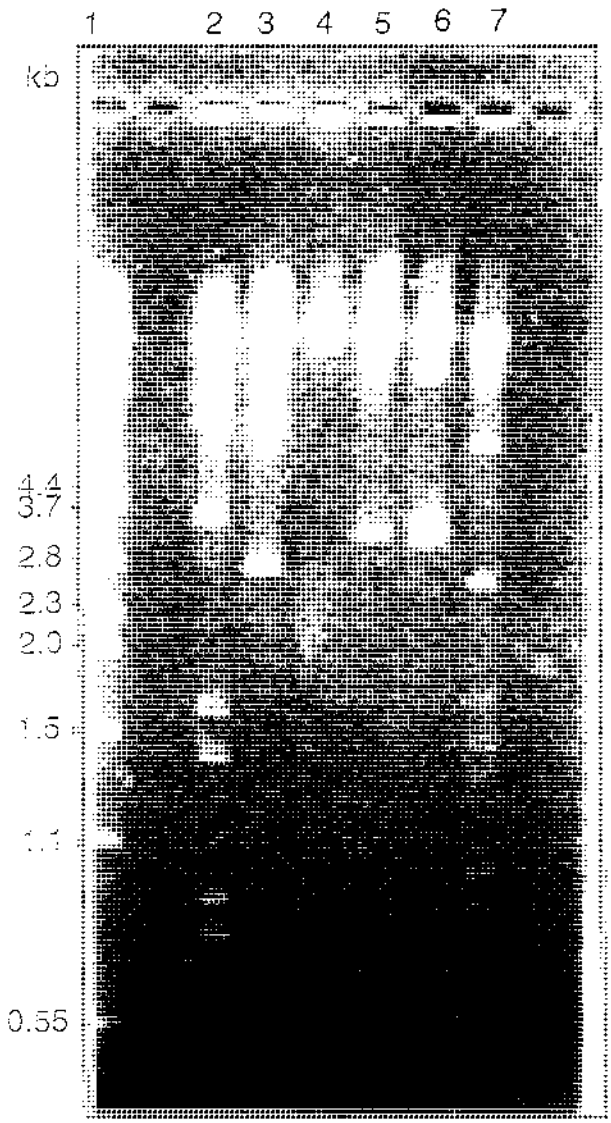

1
2
3
4
5
6
7
kb
4.4
3.7
2.8
2.3
2.0
1.5
0.55

Figure 3. 8 Restriction map of λ clone B isolated from a *Drosophila* genomic library that hybridised to the Trk fragment.

Arrow indicates position of *Dror*2 gene with arrowhead showing direction of *Dror*2. Digests undertaken did not identify exact placement of the two 0.3 kb and one 0.9 kb *Eco*RI fragments as shown by the double arrow. Not shown on the map is a 1.4 kb *Eco*RI site that is at one end of the map.

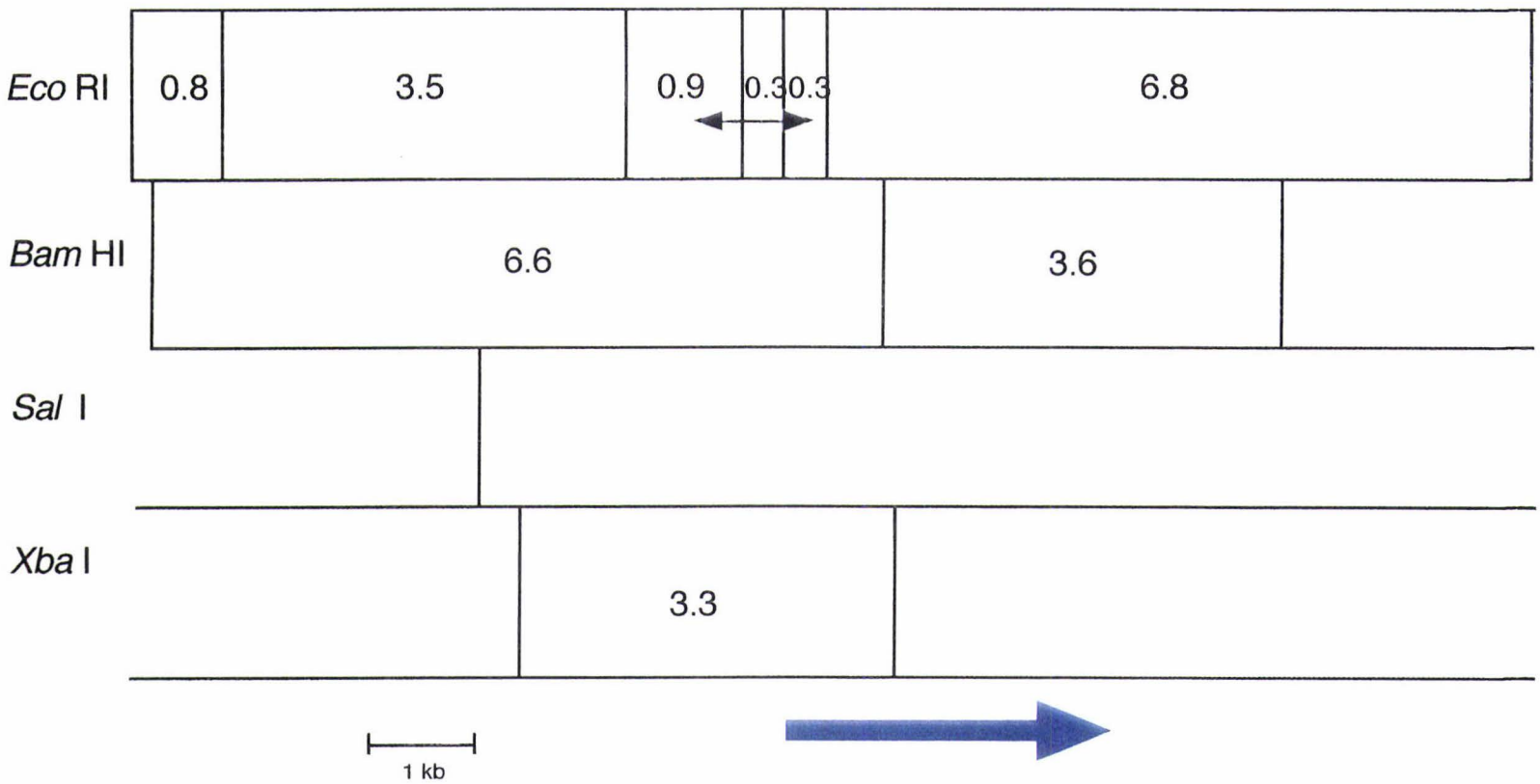

Eco RI
0.8
3.5
0.9
0.3
0.3
6.8
Bam HI
6.6
3.6
Sal I
Xba I
3.3
1 kb

Figure 3. 9 Nucleotide sequencing strategy for pKFB4 showing primer locations.

Arrows denote direction sequenced. Length of arrows indicate the approximate sequence obtained from each reaction.

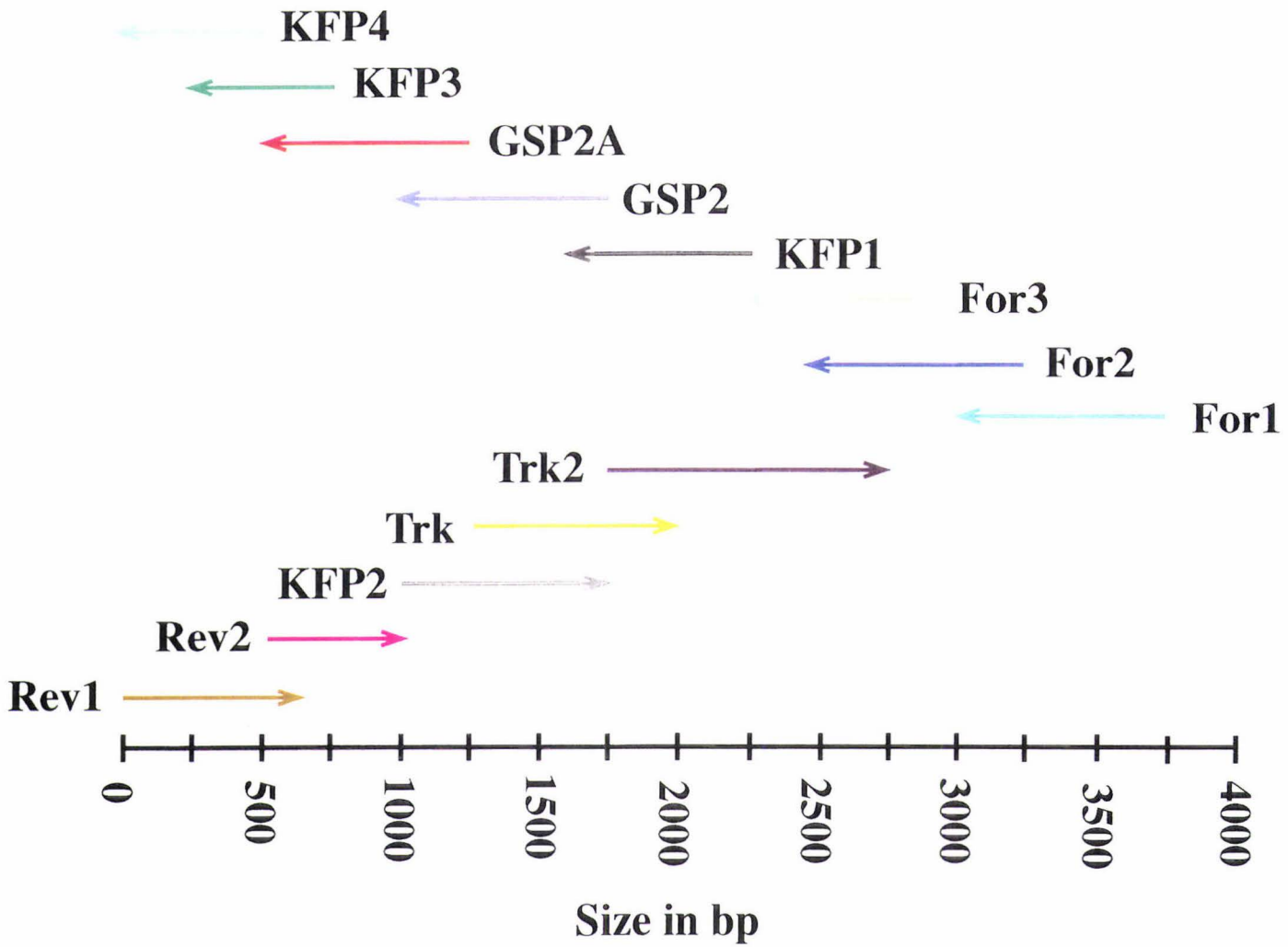
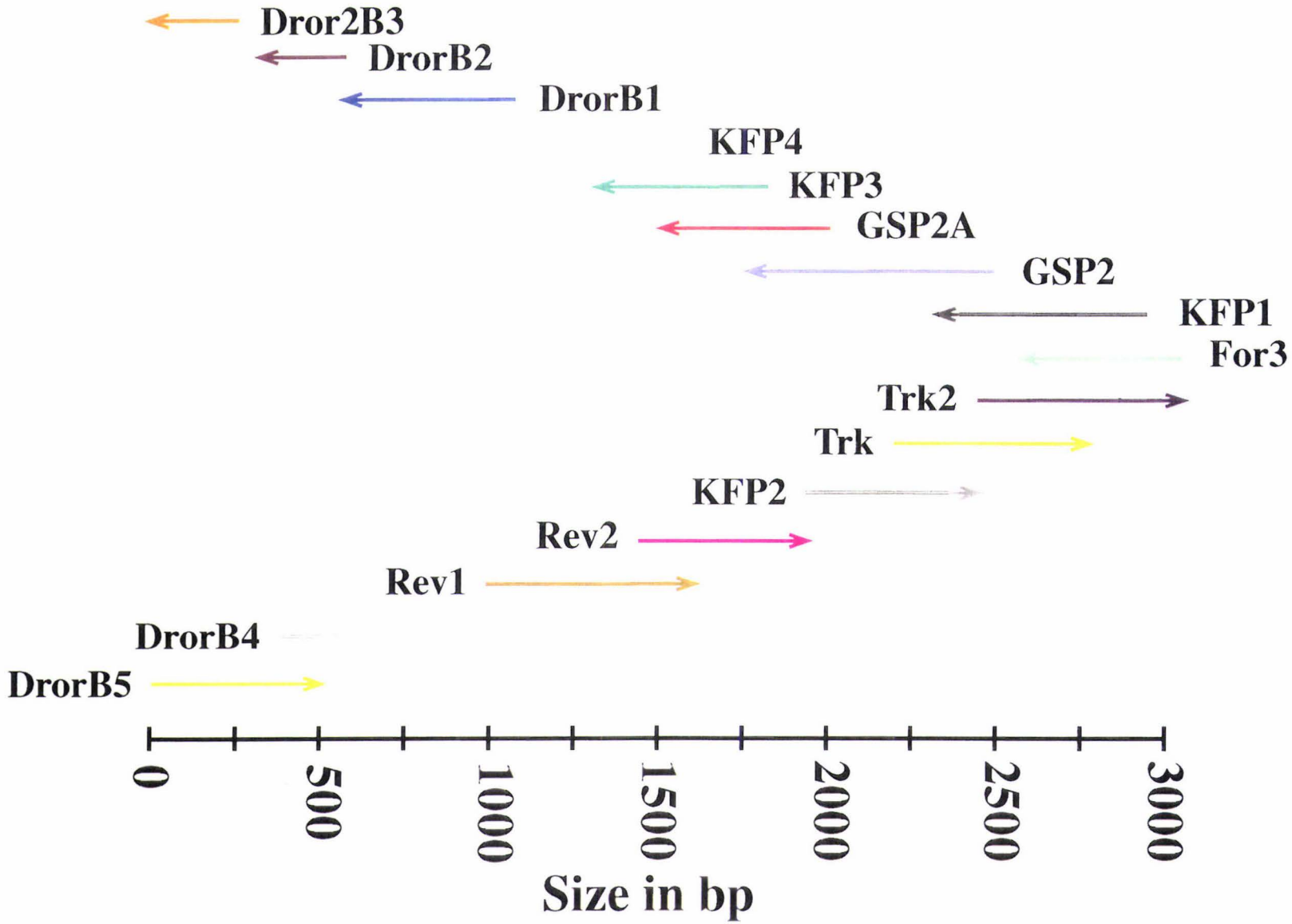

KFP4
KFP3
GSP2A
GSP2
KFP1
For3
For2
For1
Trk2
Trk
KFP2
Rev2
Rev1
0
500
1000
1500
2000
2500
3000
3500
4000
Size in bp

Figure 3. 10 Nucleotide sequencing strategy for *Dror2* showing primer locations.

Arrows denote direction sequenced. Length of arrows indicate the approximate sequence obtained from each reaction.


Dror2B3
DrorB2
DrorB1
KFP4
KFP3
GSP2A
GSP2
KFP1
For3
Trk2
Trk
KFP2
Rev2
Rev1
DrorB4
DrorB5
0
500
1000
1500
2000
2500
3000
Size in bp

3.5.4.1 Structural features of *Dror2*

The nucleotide sequence of *Dror2* and deduced amino acid sequence of the Dror2 protein are shown in Figure 3.11. The *Dror2* gene is 3094 bp in length and contains three introns, the first in the middle of the cysteine rich domain 62 nt in length, and two flanking the kringle domain the second being 66 nt and the third 57 nt in length. All three sequences at the exon-intron boundaries agree well with the consensus determined by Mount (1982).

The transcriptional start point (tsp) is 550 bp upstream of the translational start site. The tsp sequence AACAAAT matches the *Drosophila melanogaster* consensus sequence (Hultmark *et al.*, 1986) 4/7 nucleotides (nt), the agreeing nts are underlined. A putative TATA box is present 13 bp upstream of the translational start site. This sequence TATAGG shows a 4/6 nt agreement with the consensus sequence (Lewin, 1994).

Sequences upstream of the putative translational start site (the first in-frame ATG) correspond well to the *Drosophila* consensus for translational initiation (Cavener, 1987; Cavener and Ray, 1991). However, there is another in-frame ATG 51 bp downstream from the first in-frame ATG that could potentially act as the translational start site. A comparison between the two possible translational start sites (Table 3.3) shows that the first in-frame ATG matches the translational initiation consensus sequence (Cavener and Ray, 1991) eight out of twelve nucleotides compared to the second ATG that matches only three out of twelve. Therefore it is likely that the first in-frame ATG identified is the translational start site.

Following the initiating methionine is a 25 residue hydrophobic domain that may represent the signal peptide. Following the (-3, -1) rule of von Heijne (1983, 1986), the signal sequence cleavage site is most likely between the leucine and alanine amino acids (between residues 25 and 26).

This signal peptide (SP1 on the tables) was compared to that of the signal peptide (SP2) which follows the methionine encoded by the second in-frame ATG. Table 3.4 shows the two peptides, with their hydrophobic domains. In SP1 there are three charged amino acids in this region. This decreases the possibility of SP1 being the signal peptide as few charged amino acids are usually tolerated in the hydrophobic region of a signal peptide (von Heijne, 1985). Though the (-3, -1) rule of von Heijne (1983, 1986) is consistent for both these signal peptides, SP2 conforms better than SP1 to the consensus groups for residues -6 to +2 (Table 3.5).

The 3' untranslated sequence contains a putative polyadenylation signal (AATAAA) 28 nt upstream from the poly A^+ tail, consistent with Proudfoot and Brownlee (1976) and Birnstiel *et al.* (1985).

3.5.4.2 Organisation of the Dror2 protein

A schematic diagram of the organisation of the mature Dror2 protein seen in Figure 3.12. The 725 amino acid mature Dror2 protein comprises a 316 amino acid extracellular domain with a cysteine rich region and a kringle domain, a 26 amino acid transmembrane domain and 383 amino acid intracellular domain containing the catalytic kinase region.

Figure 3. 11 Nucleotide and amino acid sequence of Dror2.

Tsp indicated by bend arrow, the TATA box (green) and polyadenylation signal (blue) are underlined. The signal peptide (amino acids 1-23) is pink, the cysteine rich region (amino acids 68-202) is blue, the kringle domain (amino acids 219-297) is green, the transmembrane domain (amino acids 317-342) is purple and the tyrosine kinase domain (amino acids 441-711) is red.

```
  1 TTATTTCCTACGAATTCATACCATCAATTTTTTACGCTTATTCGAAGAGTTATAAGCTTA  60

 61 AATCATTTACAAAAAACCGAACCAATCTTCTTGCAGGGTCTGGATGCCATACCGGATGAG 120

121 AAAACGAACTAATTTTAATTAAGTAAAGTTTTCCGTTTAAGCTTAAGTGCAAAAGTGAAT 180

181 AAACGAAACGTGCAGTTCGGAAGTGATAGTTGAAATGGTCGTTTTTGGTTTTTCCTGTAG 240

241 TTCCTTACAGATTTTTTTTGCTGCCTCAGGTGTAAAGCTTCTTTCTCTATAGGGTCTTGA 300

301 ACAAATTCCCGATGAAGAGGTGAAGAAGGAGGAGAAGGACCCCAATGAGGACAGCGATGA 360

361 TCCCGCAATCGTGAAGAAGCGGGAGAAAGAGCGGAAGCTAAGGGAGAAACAGCTGAAGAA 420

421 GGAGGAAAAGCTCAGGGAGAAACAGCTGAAGGAGGAGGAGAAGCAGAAGAAAAATTAATA 480

481 TCAATAAAAACAAACATTTCGCCAGTGGAGCTCTCAAGCCGAGCACAGGACTTTCGGCCG 540

541 CCTTTGCCGCGCATTTAGCAACCCTGTGCTCGTGACAGGAAACCACCCCTAAGGGAAGAA 600

601 GTACGGGTGGCGAAACATCCCGTCGTCAGCTTTGTGTGGTATTCGGTGAAAAGGCGTCGT 660

661 CGAGCTGTGTGCGCGGAAATAAACAAATATTCTGCCGACTTACTTCGAGCTTACTTCTTA 720

721 AAATCTATTTTAAAATTCGGGTAGATTGAATATTGTGGTGTAAGGCGGTAAAATGCCTTA 780

781 GTTTATGAGTGCTCGCAAGGATTTTGCTTGGAAACGGTTCGCTAAACAAGGCATATACAT 840

841 AAAACACAGCCATGGCTGCCGGGCAATGGGTGGGGGTTGTTGAGCGGGTGCTCCGGGGAA 900
  1             M  A  A  G  Q  W  V  G  V  V  E  R  V  L  R  G  M  17

901 TGGTGCTGAAATGGGGGGCCAATTTGGCTGTCCTGGGGCTGTGCGTGTTTCTCTTTGCCA 960
 18   V  L  K  W  G  A  N  L  A  V  L  G  L  C  V  F  L  F  A  S  37

961 GCGCCACGCACGCGAACTCCCTGAACGCCATCGAGGAGCCCGCCACCCGGCGACACCACC 1020
 38   A  T  H  A  N  S  L  N  A  I  E  E  P  A  T  R  R  H  H  Q  57
```

```
1021  AGCGGCATCACGAGCGCGAGCGGGAGGAGAACGGCTACTGCGCTCCGTACAGCGGCAAGG  1080
  58    R  H  H  E  R  E  R  E  E  N  G  Y  C  A  P  Y  S  G  K  V    77

1081  TGTGCAAGGAATACCTCACCGGCCAGGTGTGGTACAGTCTGGAGGATCCCACTGGCGGGT  1140
  78    C  K  E  Y  L  T  G  Q  V  W  Y  S  L  E  D  P  T  G  G  W    97

1141  GGAAGAACGAGCAGGTGACCACGGCGCTCTGGGACGAGCTTATCTCCGATCTTACGGGTC  1200
  98    K  N  E  Q  V  T  T  A  L  W  D  E  L  I  S  D  L  T  G  L   117

1201  TGTGTCGCGAAGCAGCCGAGGTGAGTTTCGATTGATACATCTAGAGATCCCATCTCATCC  1260
 118    C  R  E  A  A  E                                             123

1261  CATTAATCTCTTTTCGCTACAGAAAATGCTCTGCGCCTATGCGTTTCCCAACTGCCACAT  1320
 124                        K  M  L  C  A  Y  A  F  P  N  C  H  M   136

1321  GGAGGGCGGTCGAGCGGTGAAGGCTCCTCTCTGCTTCGAGGATTGCCAGGCCACGCATCT  1380
 137    E  G  G  R  A  V  K  A  P  L  C  F  E  D  C  Q  A  T  H  L   156

1381  CCAGTTCTGCTACAACGACTGGGTGCTCATCGAGGAGAAGAAGGAGCGAAATATGTTCAT  1440
 157    Q  F  C  Y  N  D  W  V  L  I  E  E  K  K  E  R  N  M  F  I   176

1441  CAAGAGCCGCGGCCACTTCCGGCTACCCAACTGCTCCTCCTTGCCGCACTACAACGCTTC  1500
 177    K  S  R  G  H  F  R  L  P  N  C  S  S  L  P  H  Y  N  A  S   196

1501  CATGCGGCGACCCAACTGCTCCTACATCGGTCTCACCGAACTCAAGGAGTCCGAAGTGAG  1560
 197    M  R  R  P  N  C  S  Y  I  G  L  T  E  L  K  E  S  E  V  S   216

1561  CTGTAAGTCTTCATTATCTCAAACCCTCAAGAAGGTCTTTCTGATGCTAGAATCCCTTTT  1620

1621  CTTCACAGACGATTGCCGCAATGGAAACGGACGCTTCTACATGGGCACAATGAACGTGTC  1680
 217          Y  D  C  R  N  G  N  G  R  F  Y  M  G  T  M  N  V  S   234

1681  CAAGTCGGGCATTCCCTGCCAGCGCTGGGACACTCAGTACCCACACAAGCACTTCCAGCC  1740
 235    K  S  G  I  P  C  Q  R  W  D  T  Q  Y  P  H  K  H  F  Q  P   254

1741  ACCACTAGTCTTCCATCAGCTCCTGGAGGGCGAAAACTACTGCCGGAATGCTGGCGGCGA  1800
 255    P  L  V  F  H  Q  L  L  E  G  E  N  Y  C  R  N  A  G  G  E   274
```

```
1801  GGAGCCGCATCCCTGGTGCTACACTGTGGATGAATCAGTGCGCTGGCAGCACTGCGATAT  1860
 275   E  P  H  P  W  C  Y  T  V  D  E  S  V  R  W  Q  H  C  D  I   294

1861  ACCCATGTGTGTCGTAAGATCATAACATTTTTCCCTCTTTTACTTTTTGTAATTCACCCA  1920
 295   P  M  C                                                     297

1921  CTTCTGTTTTAGCCGGATTATGTGGACCCCAATGCTGGCGATTTGAACACGCCCATCAAG  1980
 298               P  D  Y  V  D  P  N  A  G  D  L  N  T  P  I  K   313

1981  ATGGAGAAGTTCTTCACGCCATCGATGATCTTTCTCTTGGCTGGAATAGGTTTCGTGGCC  2040
 314   M  E  K  F  F  T  P  S  M  I  F  L  L  A  G  I  G  F  V  A   333

2041  ATTGTGACCCTGCACTTGATGATATTGCTAGTCTATAAGTTGTCCAAGCACAAGGATTAC  2100
 334   I  V  T  L  H  L  M  I  L  L  V  Y  K  L  S  K  H  K  D  Y   353

2101  TCTCAGCCTGCGGGAGCAGCCACTGCCGAATGCAGTGTTTCCATGCGTGGAGGAGGAGAT  2160
 354   S  Q  P  A  G  A  A  T  A  E  C  S  V  S  M  R  G  G  G  D   373

2161  TGTGGCGGCAATCTGAACACCAGTAGAGAAACCCTCGGAGGCAATGGAAACACGAACACC  2220
 374   C  G  G  N  L  N  T  S  R  E  T  L  G  G  N  G  N  T  N  T   393

2221  TTGGCAAAATGGGGCACCATCAGGAGCACGGCCACAATACACAGCAATTGCGTGGCCCTT  2280
 394   L  A  K  W  G  T  I  R  S  T  A  T  I  H  S  N  C  V  A  L   413

2281  ACTACGGTGACCAATGTGTCTGATGCGAAGGGCACGAAACCGAATGCACGCCTGGAGAAG  2340
 414   T  T  V  T  N  V  S  D  A  K  G  T  K  P  N  A  R  L  E  K   433

2341  TTGGAGTACCCACGCGGGGATATAGTGTATGTGAGATCATTGGGTCAAGGAGCCTTCGGT  2400
 434   L  E  Y  P  R  G  D  I  V  Y  V  R  S  L  G  Q  G  A  F  G   453

2401  CGCGTCTTCCAGGCCAGGGCTCCTGGACTTGTTCCCGATCAGGAAGATCTACTAGTCGCT  2460
 454   R  V  F  Q  A  R  A  P  G  L  V  P  D  Q  E  D  L  L  V  A   473

2461  GTTAAGATGCTAAAGGACGACGCCAGCGACCAGATGCAGATGGATTTCGAGCGCGAGGCC  2520
 474   V  K  M  L  K  D  D  A  S  D  Q  M  Q  M  D  F  E  R  E  A   493

2521  TGTTTGCTGGCCGAGTTCGATCATCCCAATATCGTGAGGCTGCTGGGGGTGTGCGCCTTG  2580
 494   C  L  L  A  E  F  D  H  P  N  I  V  R  L  L  G  V  C  A  L   513
```

```
2581  GGCAGACCCATGTGCCTGCTCTTCGAGTACATGGCTCCTGGCGATCTAAGCGAGTTCTTG  2640
 514  G  R  P  M  C  L  L  F  E  Y  M  A  P  G  D  L  S  E  F  L    533

2641  CGCGCCTGCTCCCCATATGCCACACACCAGGCGCCGACACGGGATCGTCTGCAGTTGAAC  2700
 534  R  A  C  S  P  Y  A  T  H  Q  A  P  T  R  D  R  L  Q  L  N    553

2701  GAGCTACATCTGCTGCAGATGGCGGCCAACATTGCAGCGGGCATGCTGTATCTTTCGGAG  2760
 554  E  L  H  L  L  Q  M  A  A  N  I  A  A  G  M  L  Y  L  S  E    573

2761  AGAAAATTCGTCCACCGGGATTTGGCCACCAGGAATTGCCTGATCAACGAGCACATGGCG  2820
 574  R  K  F  V  H  R  D  L  A  T  R  N  C  L  I  N  E  H  M  A    593

2821  GTAAAGATCGCCGACTTTGGGCTCTCGCACAAGATCTATTTGCAGGACTATTACAAAGGC  2880
 594  V  K  I  A  D  F  G  L  S  H  K  I  Y  L  Q  D  Y  Y  K  G    613

2881  GATGAGAACGACTTCATCCCGATCCGCTGGATGCCGCTTGAGAGCATACTGTACAACAAG  2940
 614  D  E  N  D  F  I  P  I  R  W  M  P  L  E  S  I  L  Y  N  K    633

2941  TTCTCGCTTGAGTCGGATGTGTGGGCATACGGCATCTGTCTGTGGGAGGTCTTCTCCTTC  3000
 634  F  S  L  E  S  D  V  W  A  Y  G  I  C  L  W  E  V  F  S  F    653

3001  GCCTTGCAGCCCTACTTTGGGCTAACCCACGAGGAGGTGATCAAATACATCAAGGAGGGC  3060
 654  A  L  Q  P  Y  F  G  L  T  H  E  E  V  I  K  Y  I  K  E  G    673

3061  AACGTACTCGGCTGTCCGGACAACACGCCGCTCTCCGTCTACGCGCTGATGCGTCGCTGC  3120
 674  N  V  L  G  C  P  D  N  T  P  L  S  V  Y  A  L  M  R  R  C    693

3121  TGGAACCGCAAGCCCAGTGAGCGACCTGGCTTCGCCGAGATCAACCACTGCATCCAGCAC  3180
 694  W  N  R  K  P  S  E  R  P  G  F  A  E  I  N  H  C  I  Q  H    713

3181  AGCATCGCCGAGAGCGAGTGCAAGGCAATGCTCTAGGGGATTGCCGGAGAAGTGAATGAA  3240
 714  S  I  A  E  S  E  C  K  A  M  L  *                            725

3241  ATAATCCTTGTCGAACAGAACGATCGTCCGTTGTCCTTGGGGGCAAGCAATAGAAGCGCA  3300

3301  TGAATACGCCAGGTTGGGTGGAAAAACAATAGAAGGGAATGCCATCATCTTAACACTTGA  3360

3361  TAATAAACTTTTGGATCTTTAGTTGTGTATCTCAGCATTGCCGCTTTTTGTTTATGGCAT  3420
```

421 ACAAAATAGTTCTCAATTGCATAAAGGTGGCATTCCTTGACTATCTTTTTTGAGTCTCCC 3480

481 CAGCTCTTATGCATATTGAATTGAATTTTGCACAAGGTGAGATCAGAGGGGCAACTGGTT 3540

541 TGTGATTAAGTTAGTATAAATGACAACTCCATATCTTTAAGAGATAGCTCCGTTACCAAG 3600

601 TCTTTCATTCGATGCTTCCTCACATAACTTTCTTTAATCGATACAGAAACTATTTTGATA 3660

661 CACTTATACCCTCTGATCTCTCGATCTCGATCGAA 3695

Figure 3. 12 Schematic diagram of the organisation of the mature Dror2 protein.

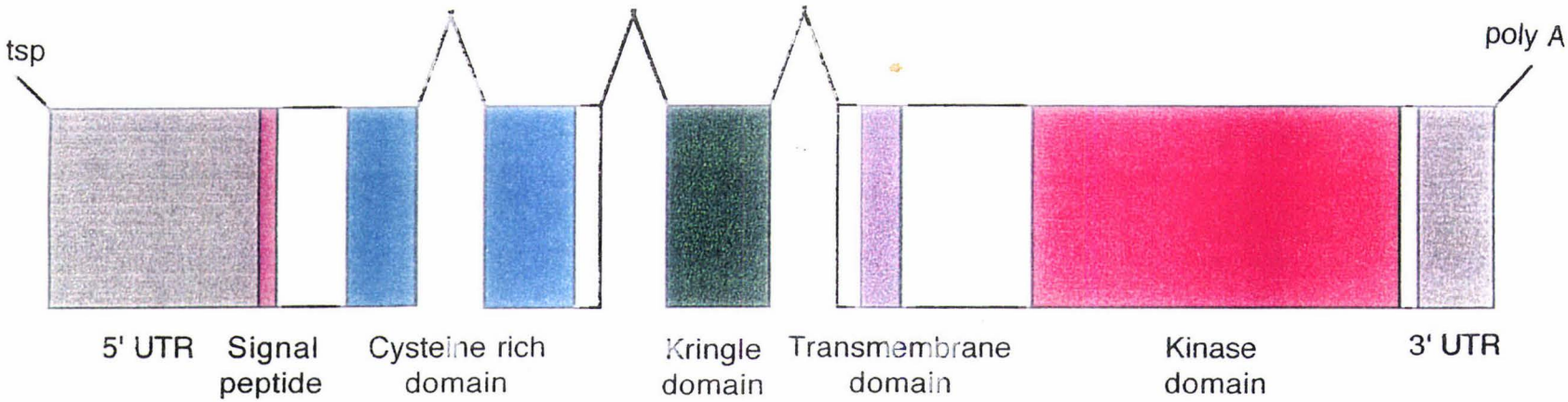

tsp
poly A
5' UTR
Signal peptide
Cysteine rich domain
Kringle domain
Transmembrane domain
Kinase domain
3' UTR

Table 3.3 Comparison between the two possible initiation consensus sequences of Dror2.

	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1		+4	+5
First ATG	a	a	a	c	a	c	a	g	c	c	ATG	g	c
Second ATG	g	c	t	c	c	g	g	g	g	a	ATG	g	t
consensus	a/c	a	c	c/a	a/g	c	c	a	a	a/c	ATG	a/g	c/a

Table 3.4 Sequences of the two possible signal peptides of Dror2

The M on each represents position -25 and the final amino acid represents -1.

SPs	Amino acids																								
SP1	M	A	A	G	Q	W	V	G	V	V	E	R	V	L	R	G	M	V	L	K	W	G	A	N	L
						Hydrophobic region of 14 amino acids																			
SP2	M	V	L	K	W	G	A	N	L	A	V	L	G	L	C	V	F	L	F	A	S	A	T	H	A
									Hydrophobic region of 12 aas																

Table 3.5 Comparison of the cleavage sites of the two possible signal peptides of Dror2

Position	Signal peptide 1		Signal peptide 2		Consensus group
	aa	Group	aa	Group	
-6	K	charged	A	neutral	hydrophobic
-5	W	aromatic	S	neutral	neutral
-4	G	Gly	A	neutral	neutral/ hydrophobic
-3	A	neutral	T	neutral	neutral
-2	N	hydrophobic	H	aromatic	hydrophobic/aromatic
-1	L	hydrophobic	A	neutral	neutral
−1	A	neutral	N	hydrophobic	polar
+2	V	hydrophobic	S	neutral	charged/ hydrophobic

3.5.4.3 Sequence comparisons

An alignment of the extracellular domains of all the Ror family of receptors is seen in Figures 3.13 and 3.14. Within the extracellular domain, the kringle domain (Figure 3.13) shares a 50 % identity with human Ror1, a 46 % identity with human Ror2 and a 42 % identity with Dror. In the cysteine rich domain (Figure 3.14) Dror2 shares a 26 % identity with Ror1 and a 23 % identity with Ror2 but only a 13 % identity to Dror. Dror contains a unique 55 amino acid lysine-rich insert in this region, without this there would be a 20 % identity with Dror2, very similar to Ror1 and Ror2. When these four proteins are aligned all 6 cysteine residues in the kringle domain and 10 residues in the cysteine-rich domain in Ror1, Ror2 and Dror have equivalent positions in the Dror2 protein.

An alignment of the kinase domains of Dror, Dror2, human Ror1 and Ror2 and mouse TrkB is shown in Figure 3.15. The tyrosine kinase domain of Dror2 shows a 45 % identity to Dror, a 44 % identity with Ror1, a 42 % identity with Ror2 and a 48 % identity to TrkB. A number of the shared amino acids in this domain are not found in many other RTKs suggesting the occurrence of a Trk-like superfamily of RTKs.

Contained in the tyrosine kinase domain is the YXXDYY sequence motif (amino acids 606-611) that corresponds to the autophosphorylation site of insulin receptors (Masiakowski and Carroll, 1992; Pearson and Kemp, 1991). Dror2 also contains two GXGXXG putative ATP-binding motifs found within the tyrosine kinase domain of Dtrk (Pulido *et al.*, 1992). The first is $G^{370}XGXXG/K^{406}$ and the second $G^{448}XGXXG/K^{473}$.

Figure 3. 13 Alignment of the kringle domain of Dror2 with the members of the Ror family of RTKs.

Alignments are of Ror1 and Ror2 from human and Dror and Dror2 from *Drosophila*. Purple blocks represent identical residues, asterisks represent the common cysteine residues (see text), dashes represent gaps for optimal alignment.

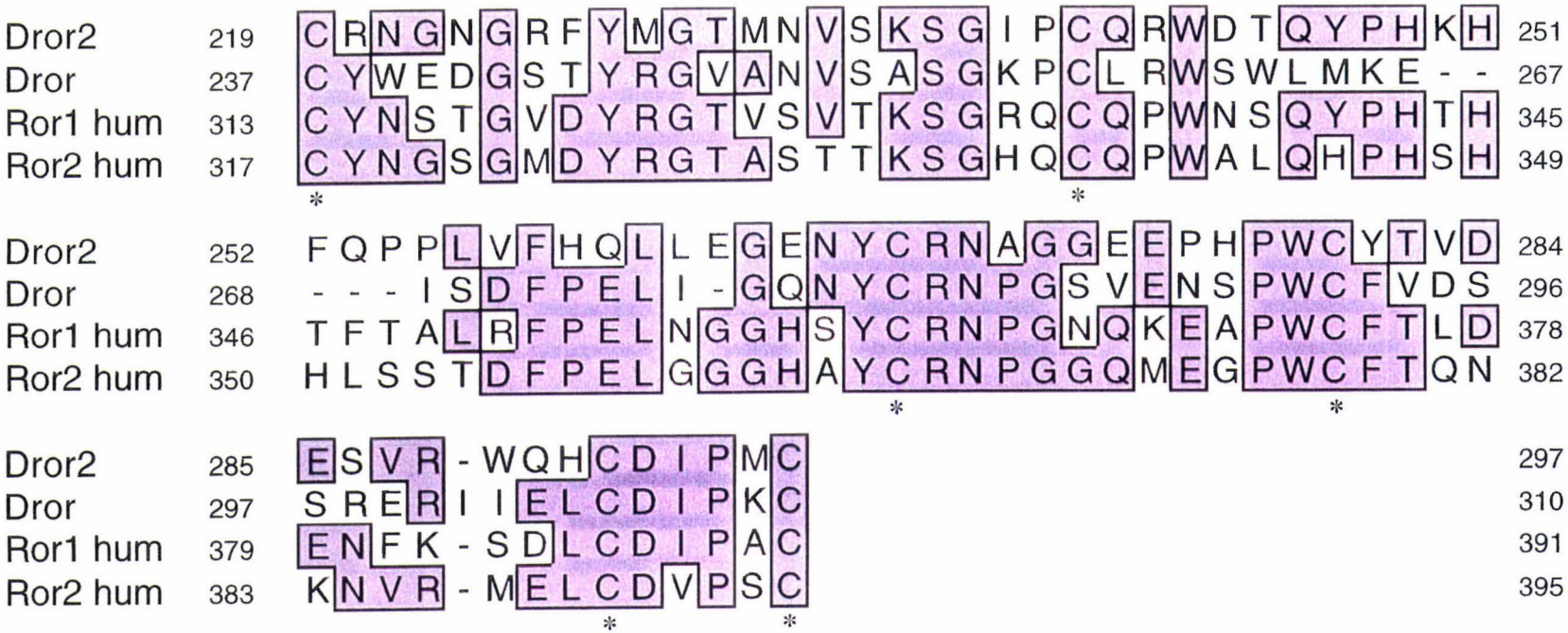
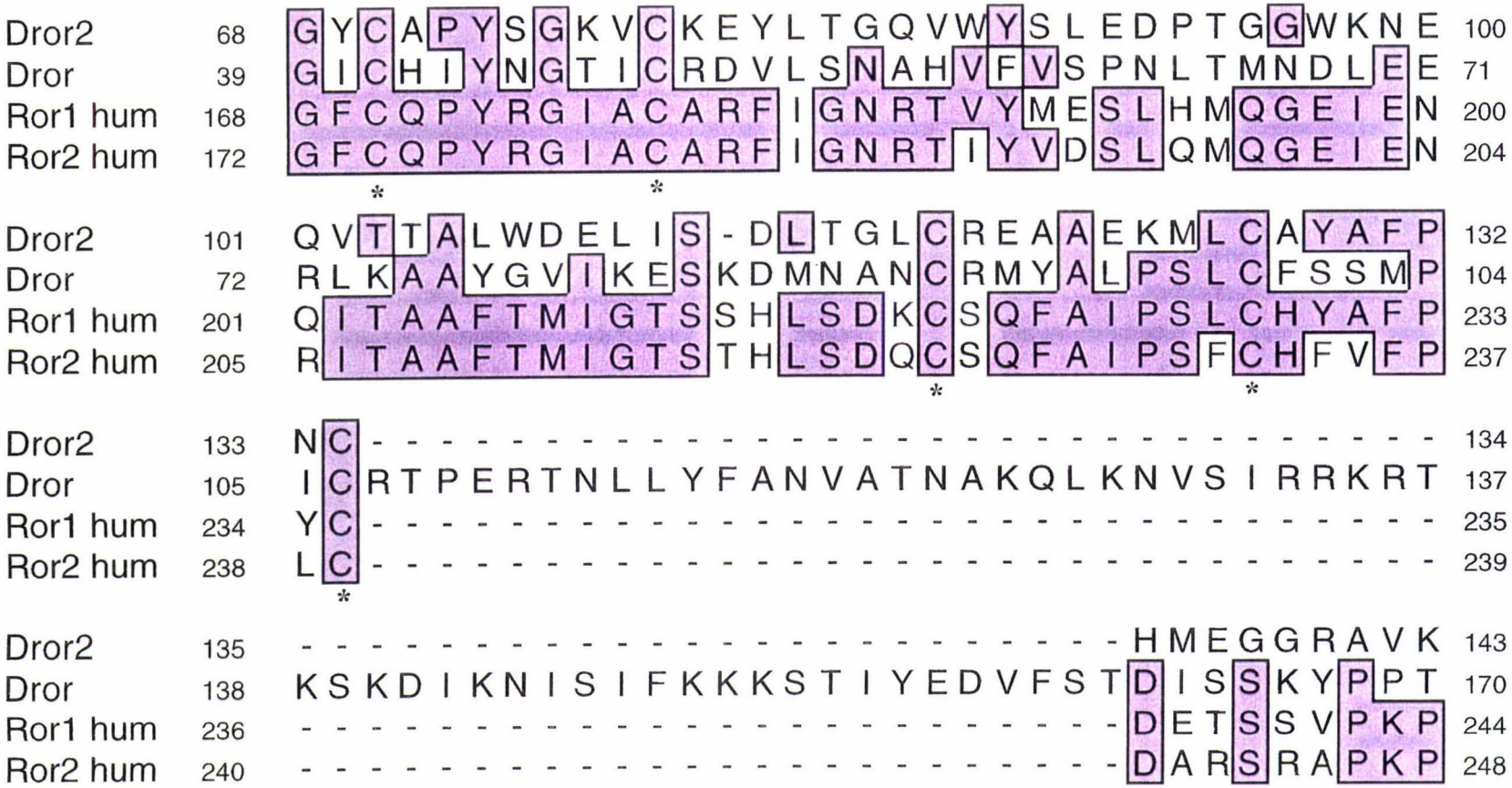

Dror2 219 CRNGNGRFYMGTMNVSKSGIPCQRWDTQYPHKH 251
Dror 237 CYWEDGSTYRGVANVSASGKPCLRWSWLMKE-- 267
Ror1 hum 313 CYNSTGVDYRGTVSVTKSGRQCQPWNSQYPHTH 345
Ror2 hum 317 CYNGSGMDYRGTASTTKSGHQCQPWALQHPHSH 349
Dror2 252 FQPPLVFHQLLEGENYCRNAGGEEPHPWCYTVD 284
Dror 268 ---ISDFPELI-GQNYCRNPGSVENSPWCFVDS 296
Ror1 hum 346 TFTALRFPELNGGHSYCRNPGNQKEAPWCFTLD 378
Ror2 hum 350 HLSSTDFPELGGGHAYCRNPGGQMEGPWCFTQN 382
Dror2 285 ESVR-WQHCDIPMC 297
Dror 297 SRERIIELCDIPKC 310
Ror1 hum 379 ENFK-SDLCDIPAC 391
Ror2 hum 383 KNVR-MELCDVPSC 395

Figure 3. 14 Alignment of the cysteine-rich domain of Dror2 with the members of the Ror family of RTKs

Alignments are of Ror1 and Ror2 from human and Dror and Dror2 from *Drosophila*. Purple blocks represent identical residues, asterisks represent the common cysteine residues (see text), dashes represent equal gaps for optimal alignment.


Dror2 68 GYCAPYSGKVCKEYLTGQVWYSLEDPTGGWKNE 100
Dror 39 GICHIYNGTICRDVLSNAHVFVSPNLTMNDLEE 71
Ror1 hum 168 GFCQPYRGIACARFIGNRTVYMESLHMQGEIEN 200
Ror2 hum 172 GFCQPYRGIACARFIGNRTIYVDSLQMQGEIEN 204
Dror2 101 QVTTALWDELIS-DLTGLCREAAEKMLCAYAFP 132
Dror 72 RLKAAYGVIKESKDMNANCRMYALPSLCFSSMP 104
Ror1 hum 201 QITAAFTMIGTSSHLSDKCSQFAIPSLCHYAFP 233
Ror2 hum 205 RITAAFTMIGTSTHLSDQCSQFAIPSFCHFVFP 237
Dror2 133 NC---------- 134
Dror 105 ICRTPERTNLLYFANVATNAKQLKNVSIRRKRT 137
Ror1 hum 234 YC---------- 235
Ror2 hum 238 LC---------- 239
Dror2 135 ----------HMEGGRAVK 143
Dror 138 KSKDIKNISIFKKKSTIYEDVFSTDISSKYPPT 170
Ror1 hum 236 ----------DETSSVPKP 244
Ror2 hum 240 ----------DARSRAPKP 248

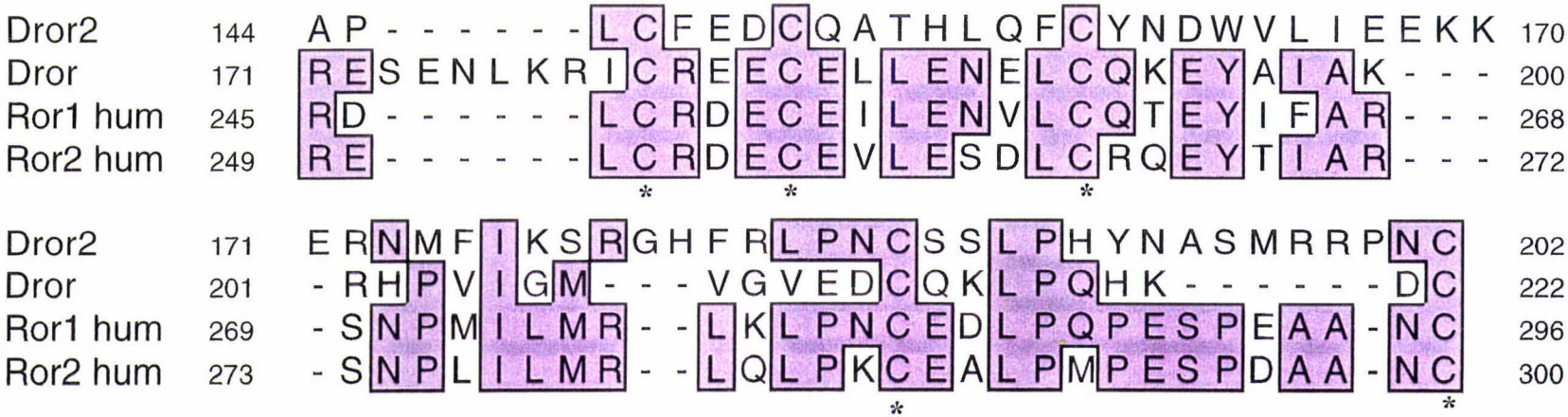
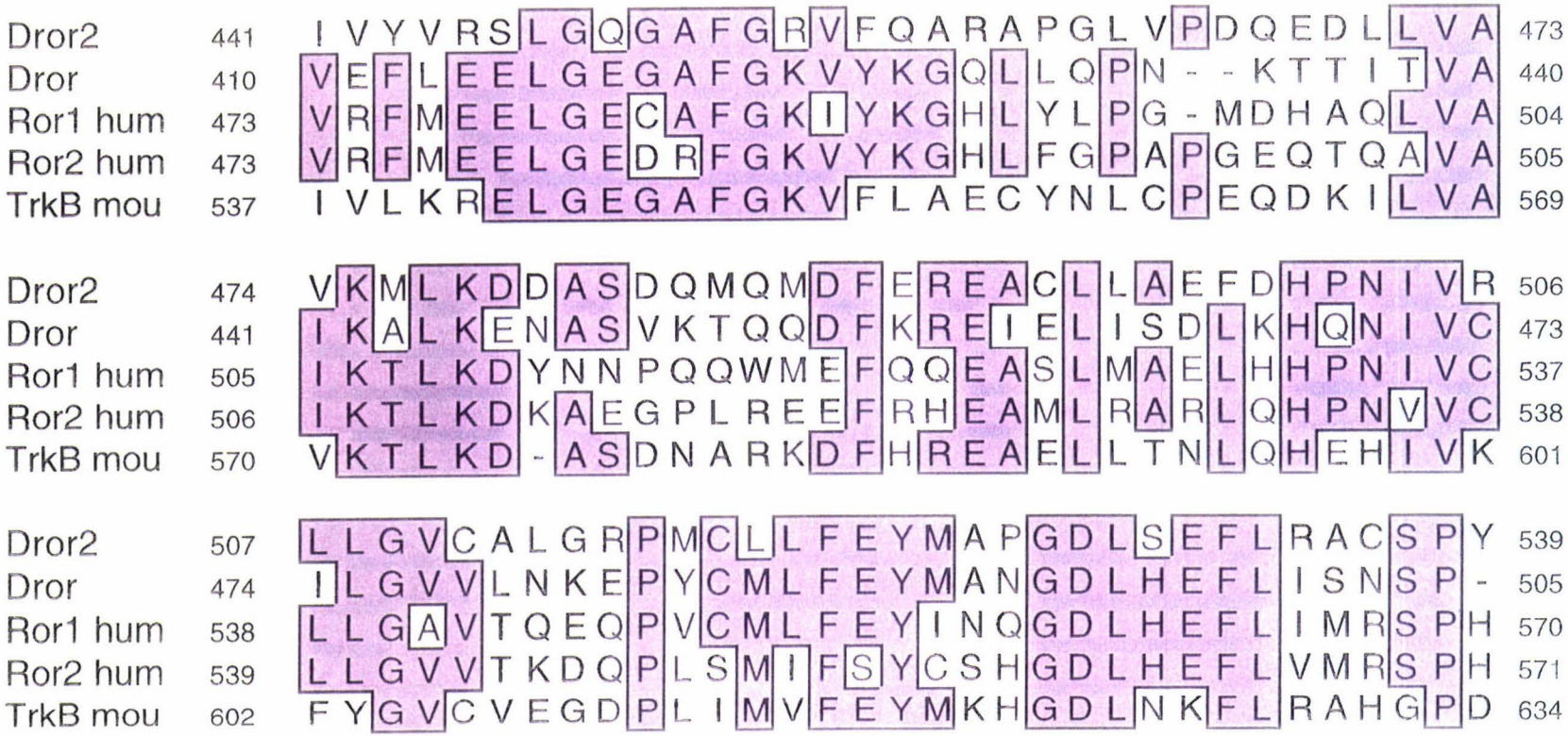
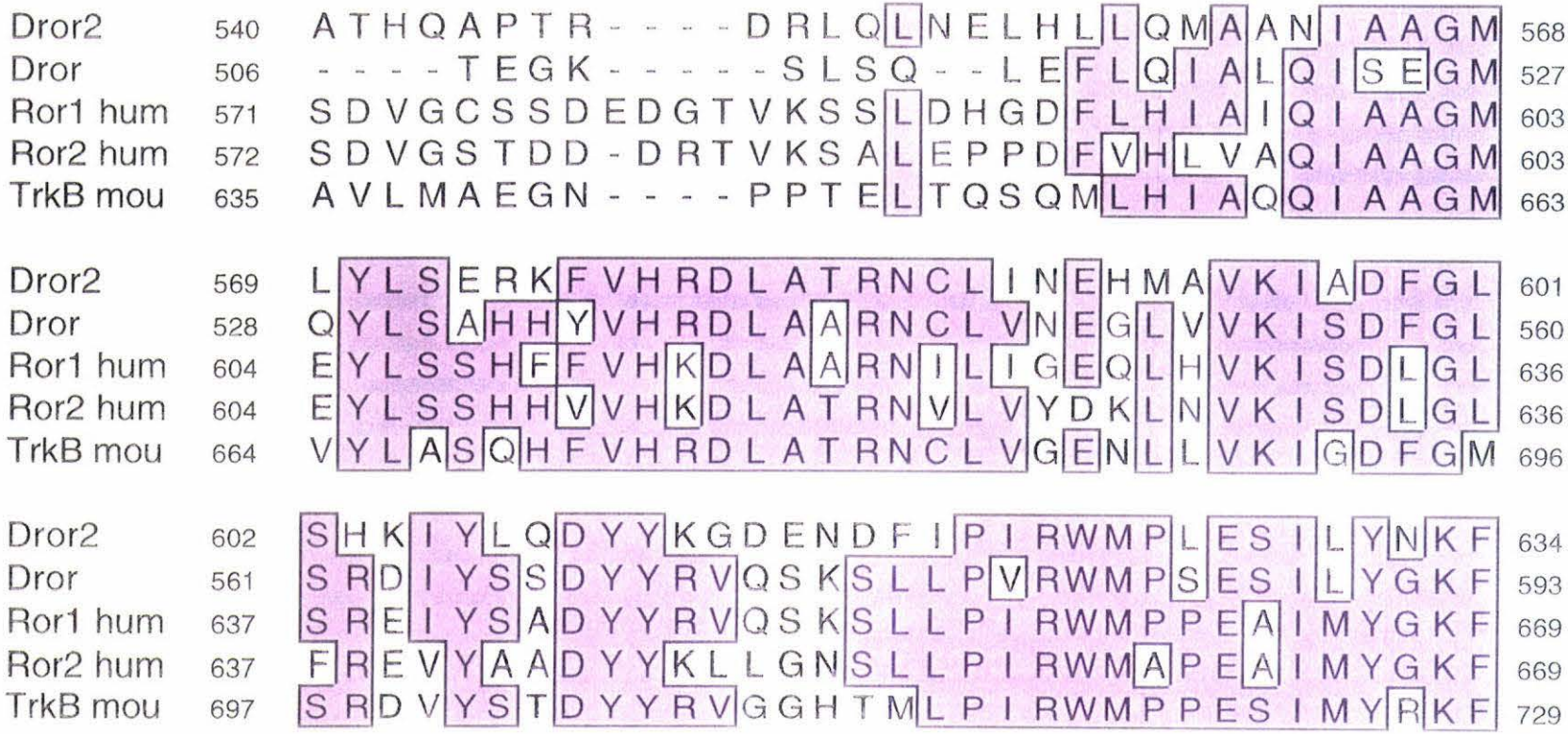
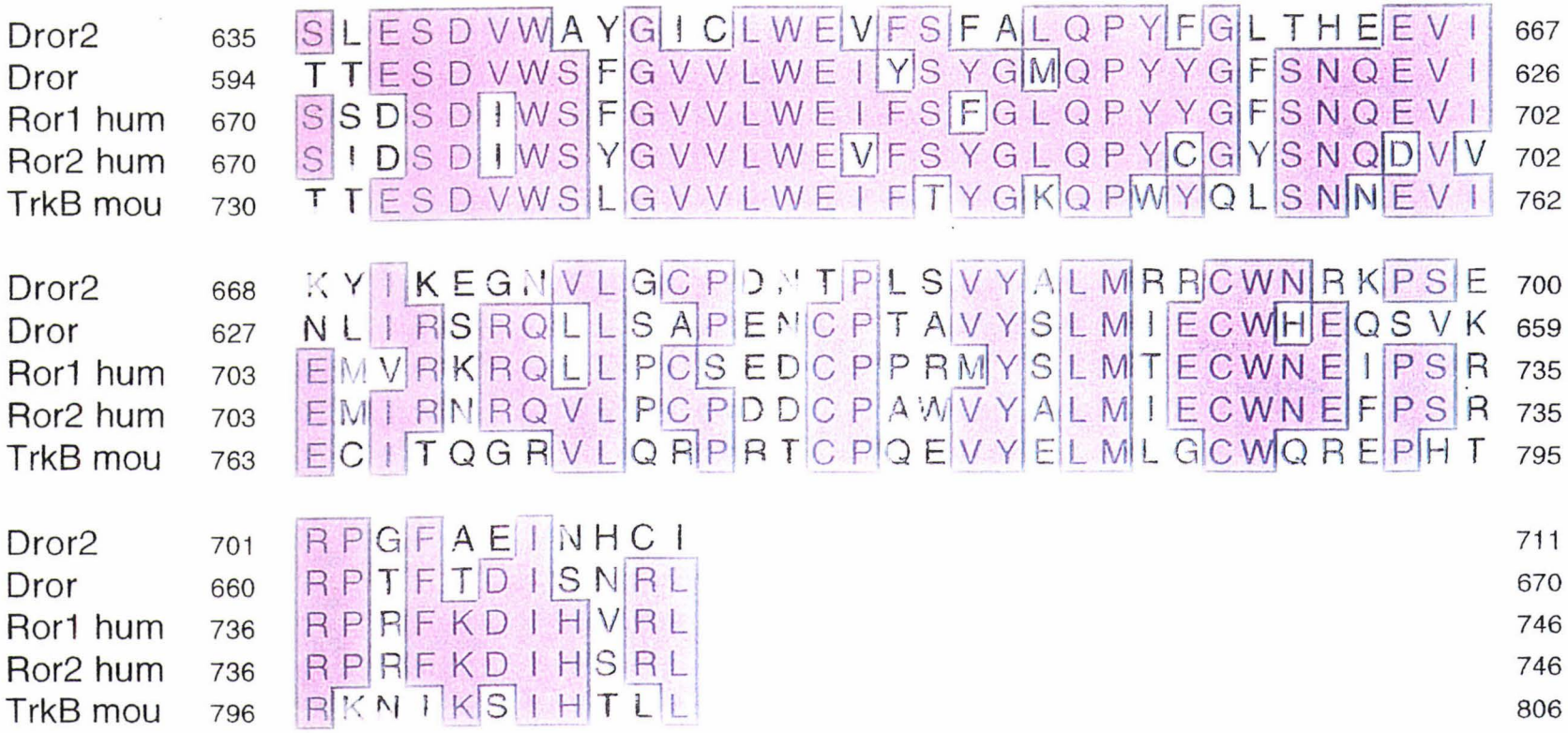

Dror2 144 AP------LCFEDCQATHLQFCYNDWVLIEEKK 170
Dror 171 RESENLKRICREECELLENELCQKEYAIAK--- 200
Ror1 hum 245 RD------LCRDECEILENVLCQTEYIFAR--- 268
Ror2 hum 249 RE------LCRDECEVLESDLCRQEYTIAR--- 272
Dror2 171 ERNMFIKSRGHFRLPNCSSLPHYNASMRRPNC 202
Dror 201 -RHPVIGM---VGVEDCQKLPQHK------DC 222
Ror1 hum 269 -SNPMILMR--LKLPNCEDLPQPESPEAA-NC 296
Ror2 hum 273 -SNPLILMR--LQLPKCEALPMPESPDAA-NC 300

Figure 3. 15 Alignment of kinase domains of Dror2 with other members of the Trk/Ror families of RTKs.

Alignments are of Ror1 and Ror2 from human, TrkB from mouse and Dror and Dror2 from *Drosophila*. Purple blocks represent identical residues, dashes represent gaps for optimal alignment.


Dror2 441 IVYVRSLGQGAFGRVFQARAPGLVPDQEDLLVA 473
Dror 410 VEFLEELGEGAFGKVYKGQLLQPN--KTTITVA 440
Ror1 hum 473 VRFMEELGECAFGKIYKGHLYLPG-MDHAQLVA 504
Ror2 hum 473 VRFMEELGEDRFGKVYKGHLFGPAPGEQTQAVA 505
TrkB mou 537 IVLKRELGEGAFGKVFLAECYNLCPEQDKILVA 569
Dror2 474 VKMLKDDASDQMQMDFEREACLLAEFDHPNIVR 506
Dror 441 IKALKENASVKTQQDFKREIELISDLKHQNIVC 473
Ror1 hum 505 IKTLKDYNNPQQWMEFQQEASLMAELHHPNIVC 537
Ror2 hum 506 IKTLKDKAEGPLREEFRHEAMLRARLQHPNVVC 538
TrkB mou 570 VKTLKD-ASDNARKDFHREAELLTNLQHEHIVK 601
Dror2 507 LLGVCALGRPMCLLFEYMAPGDLSEFLRACSPY 539
Dror 474 ILGVVLNKEPYCMLFEYMANGDLHEFLISNSP- 505
Ror1 hum 538 LLGAVTQEQPVCMLFEYINQGDLHEFLIMRSPH 570
Ror2 hum 539 LLGVVTKDQPLSMIFSYCSHGDLHEFLVMRSPH 571
TrkB mou 602 FYGVCVEGDPLIMVFEYMKHGDLNKFLRAHGPD 634


Dror2 540 ATHQAPTR----DRLQLNELHLLQMAANIAAGM 568
Dror 506 ----TEGK-----SLSQ--LEFLQIALQISEGM 527
Ror1 hum 571 SDVGCSSDEDGTVKSSLDHGDFLHIAIQIAAGM 603
Ror2 hum 572 SDVGSTDD-DRTVKSALEPPDFVHLVAQIAAGM 603
TrkB mou 635 AVLMAEGN----PPTELTQSQMLHIAQQIAAGM 663
Dror2 569 LYLSERKFVHRDLATRNCLINEHMAVKIADFGL 601
Dror 528 QYLSAHHYVHRDLAARNCLVNEGLVVKISDFGL 560
Ror1 hum 604 EYLSSHFFVHKDLAARNILIGEQLHVKISDLGL 636
Ror2 hum 604 EYLSSHHVVHKDLATRNVLVYDKLNVKISDLGL 636
TrkB mou 664 VYLASQHFVHRDLATRNCLVGENLLVKIGDFGM 696
Dror2 602 SHKIYLQDYYKGDENDFIPIRWMPLESILYNKF 634
Dror 561 SRDIYSSDYYRVQSKSLLPVRWMPSESILYGKF 593
Ror1 hum 637 SREIYSADYYRVQSKSLLPIRWMPPEAIMYGKF 669
Ror2 hum 637 FREVYAADYYKLLGNSLLPIRWMAPEAIMYGKF 669
TrkB mou 697 SRDVYSTDYYRVGGHTMLPIRWMPPESIMYRKF 729


Dror2 635 SLESDVWAYGICLWEVFSFALQPYFGLTHEEVI 667
Dror 594 TTESDVWSFGVVLWEIYSYGMQPYYGFSNQEVI 626
Ror1 hum 670 SSDSDIWSFGVVLWEIFSFGLQPYYGFSNQEVI 702
Ror2 hum 670 SIDSDIWSYGVVLWEVFSYGLQPYCGYSNQDVV 702
TrkB mou 730 TTESDVWSLGVVLWEIFTYGKQPWYQLSNNEVI 762
Dror2 668 KYIKEGNVLGCPDNTPLSVYALMRRCWNRKPSE 700
Dror 627 NLIRSRQLLSAPENCPTAVYSLMIECWHEQSVK 659
Ror1 hum 703 EMVRKRQLLPCSEDCPPRMYSLMTECWNEIPSR 735
Ror2 hum 703 EMIRNRQVLPCPDDCPAWVYALMIECWNEFPSR 735
TrkB mou 763 ECITQGRVLQRPRTCPQEVYELMLGCWQREPHT 795
Dror2 701 RPGFAEINHCI 711
Dror 660 RPTFTDISNRL 670
Ror1 hum 736 RPRFKDIHVRL 746
Ror2 hum 736 RPRFKDIHSRL 746
TrkB mou 796 RKNIKSIHTLL 806

3.6 3' RACE

cDNAs were synthesised (Section 2.2.3) from total RNA (Section 2.2.2.1) and poly A^+ RNA (Section 2.2.2.2) isolated from a range of developmental stages. These stages were 0-4 hour old embryos, 4-8 hour old embryos, 3rd instar larvae, pupae, female adult and male adult flies.

Taq polymerase was initially used in an attempt to isolate the 3' end of *Dror2*. cDNA was synthesised from both total RNA and poly A^+ RNA using either of three different primers: GSP2, oligo (dT) and oligo (dT)-*Bam*HI. Following this, the cDNA was checked using control primers. For cDNA synthesised by GSP2, Rev2 and GSP2A were used to amplify a known region in *Dror2*. This was not successful, so new cDNA was synthesised using the oligo (dT) primer and a control reaction using ZW3 and ZW5 primers was undertaken. ZW3 and ZW5 were chosen as they amplify a known section of the glucose-6-phosphate dehydrogenase (G6PD) gene containing a 125 bp intron (Scott *et al.*, 1993). The cDNA was amplified and the resulting fragment size compared to that of the genomic DNA to ensure no contaminations were present. These results were inconsistent with previous results, consequently another cDNA synthesis was attempted using the oligo (dT)-*Bam*HI primer. 3' RACE was attempted on this cDNA with the oligo (dT)-*Bam*HI and Trk primers for the first round of synthesis and oligo (dT)-*Bam*HI and DrorZ1 primers for the second round.

The results revealed a fragment the same size as the genomic DNA control in all the cDNA lanes including the no DNA control, demonstrating a genomic DNA contamination. All solutions in the PCR reaction were changed in an attempt to remove the contaminant. This was successful and it was decided to run cDNA synthesis controls and 3' RACE controls on all further reactions before attempting

the RACE with *Dror2*, in case the contaminant returned. It was also decided at this point to use the elONGase enzyme instead of Taq polymerase as it synthesises longer strands of DNA than Taq polymerase, hence the chances of success were greater.

Rp49 is expressed strongly in all stages of development (O'Connell and Rosbash, 1984), hence it was chosen for a positive control. A control PCR reaction using rp49 GSP1 and rp49 GSP3 (primers that amplified a specific region of rp49 DNA) was run to ensure the cDNA synthesis occurred. A second experiment was performed: a control of the 3' RACE using the cDNA synthesised from the oligo (dT)-UP primer to check the conditions of the reactions (Section 2.2.16). The first round of PCR contained the UP and rp49 GSP1 primers with 1 µL cDNA, the second round UP and rp49 GSP2 primers with 1 µL diluted cDNA. The resulting agarose gel is shown in Figure 3.16.

Once the cDNA synthesis and PCR conditions had been established, the cDNA synthesised with oligo (dT)-UP primer in the control reactions were used for 3' RACE of Dror2. The first round of PCR contained the UP and Trk primers, the second round contained the UP and DrorZ1 primers. The samples were size fractionated on a 1.7 % agarose gel as can be seen in Figure 3.17.

The 470 bp fragment amplified from 0-4 hour old embryo cDNA reaction was excised, gel purified (Section 2.2.7.2) and was subcloned into pGEM. A mini preparation of plasmid DNA was then sequenced using the M13 forward and reverse primers on the ABI sequencing machine. The sequence corresponded exactly with gene sequence to the poly A^+ site.

3.7 5' RACE

5' RACE was performed using 0-24 hour embryo, female adult and male adult RNA. The DrorA1 primer was used in synthesising the cDNA, DrorB1 and Abridged anchor primers (BRL) in the first round of PCR and DrorC1 and AUAP (BRL) primers in the second round.

A fragment of 900 bp was clearly visible in female adult, faintly in male adult and not seen at all in embryos (Figure 3.18). This fragment from female adult was excised and gel purified (Section 2.2.7.2). This was subcloned as in Section 3.6 and sequenced on the ABI sequencing machine.

Figure 3. 16 3' RACE of rp49

Lane 1: 1 kb ladder, lane 2: 0-4 hour old embryos, lane 3: 4-8 hour old embryos, lane 4: 3rd instar larvae, lane 5: pupae, lane 6: female adult, lane 7: male adult, lane 8: no DNA control.

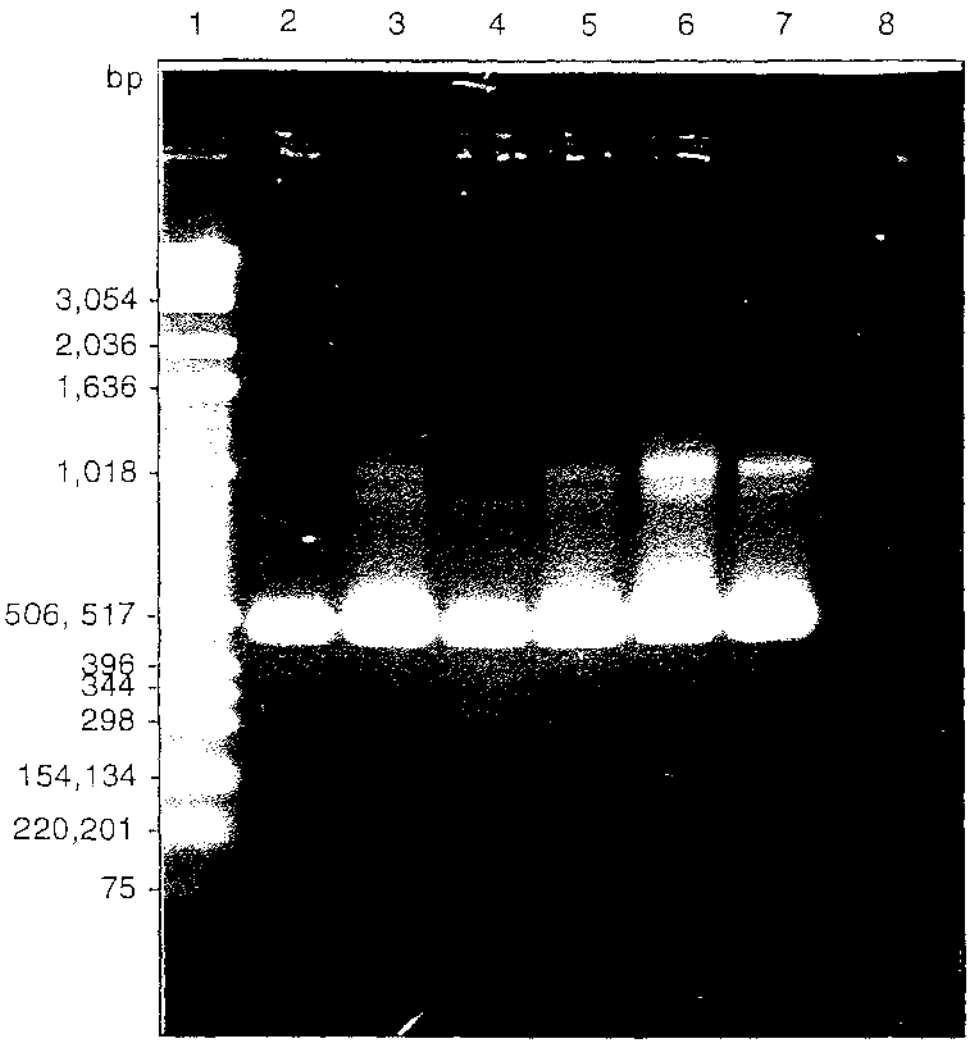

1
2
3
4
5
6
7
8
bp
3,054
2,036
1,636
1,018
506, 517
396
344
298
154,134
220,201
75

Figure 3. 17 3' RACE

Lane 1: 1 kb ladder, lane 2: 0-4 hour old embryos, lane 3: 4-8 hour old embryos, lane 4: 3rd instar larvae, lane 5: pupae, lane 6: female adult, lane 7: male adult, lane 8: no DNA control.

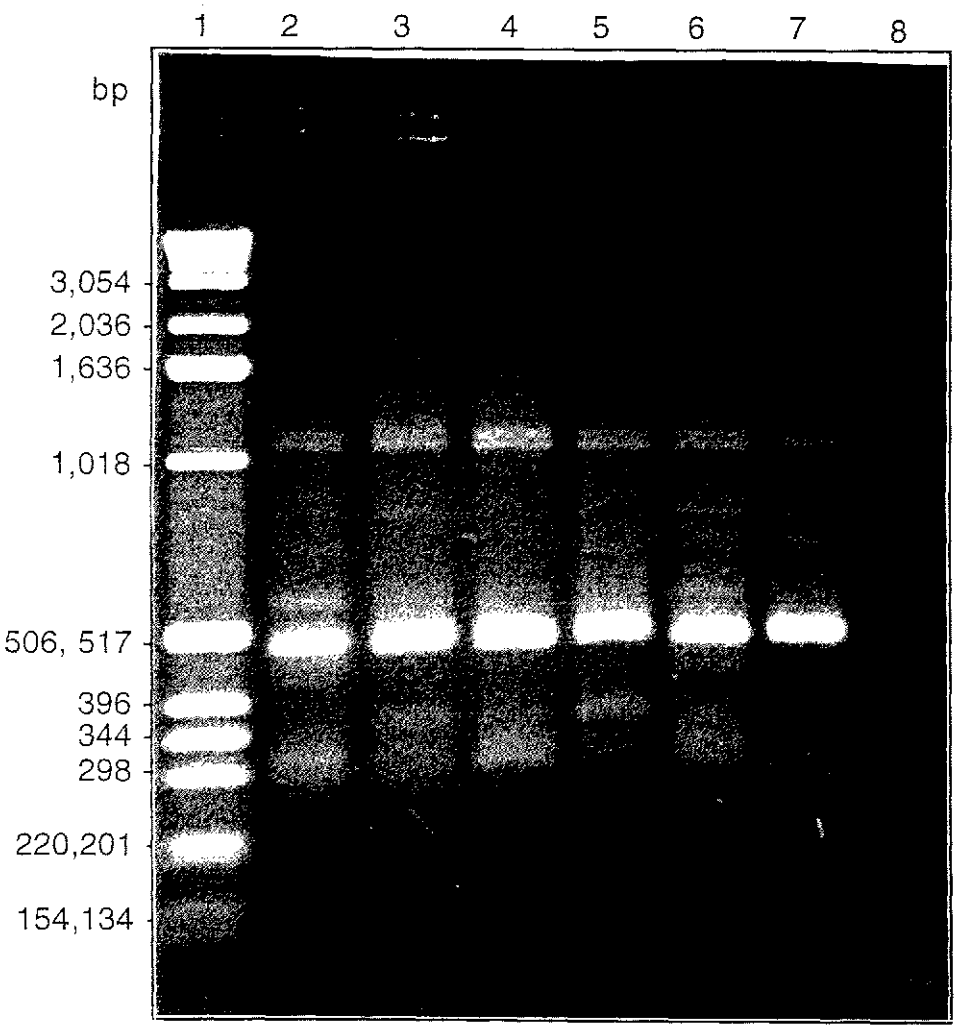
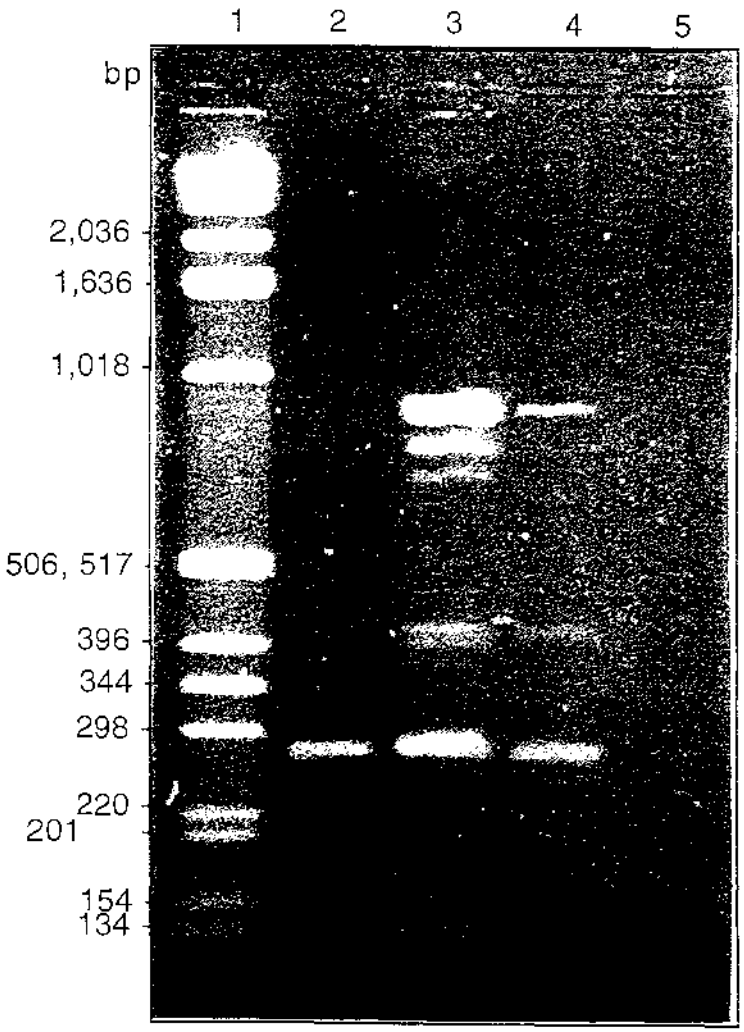

1
2
3
4
5
6
7
8
bp
3,054
2,036
1,636
1,018
506, 517
396
344
298
220,201
154,134

Figure 3. 18 5' RACE

Lane 1: 1 kb ladder, lane 2: 0-4 hour old embryos, lane 3: female adult, lane 4: male adult, lane 5: no DNA control.


1
2
3
4
5
bp
2,036
1,636
1,018
506, 517
396
344
298
220
201
154
134

3.8 EXPRESSION OF DROR2

3.8.1 NORTHERN BLOT HYBRIDISATION

Northern blot hybridisation (Section 2.2.19) was initially attempted using 10 μg of total RNA isolated from a range of developmental stages (Section 2.2.2.1). These stages were 0-4 hour old embryos, 4-8 hour old embryos, 3rd instar larvae, pupae, female adult and male adult flies. Plasmid KFB4 was digested in a double digest containing *Bam*HI and *Pst*I and resulting fragments were separated on a 1.0 % Seaplaque gel in TAE. The 1.6 kb band containing most of *Dror2* was excised, Qiagen purified and checked on a 1.0 % agarose gel (Section 2.2.7.2). This was then used as a probe for *Dror2*. Twenty five ng was radioactively labelled with ^{32}P (Section 2.2.19.1) and added to the hybridisation tube.

Upon obtaining no signal from the *Dror2* probe several control probes were tested. The *armadillo* (*arm*) gene encodes for a 3.2 kb mRNA and is expressed most during embryogenesis and pupation and least during larval development and in adults (Riggleman *et al.*, 1989). A 2.5 kb *Bam*HI fragment containing most of the *arm* gene was gel purified from pArm and labelled identically to the *Dror2* probe. However, again this yielded no signal, suggesting that the mRNAs were too small a proportion of the total RNA to be readily detected. Thus, it was decided to use poly A^+ RNA as it is enriched for mRNAs.

Poly A^+ RNA was isolated (Section 2.2.2.2) and approximately 2.5 μg of this was used per lane in the Northern Blot. *Ras*1 (D*mras*85D) is constantly expressed throughout the development of the fly (Mozer *et al.*, 1985) and encodes one major transcript of 2.0 kb (Brock, 1987). A 1.6 kb *Sal*I/*Eco*RI *Ras*1 fragment was purified from pRas and labelled identically to the *Dror2* probe. Unfortunately no

signal was detectable with this probe either, indicating that a more abundantly expressed control had to be tested.

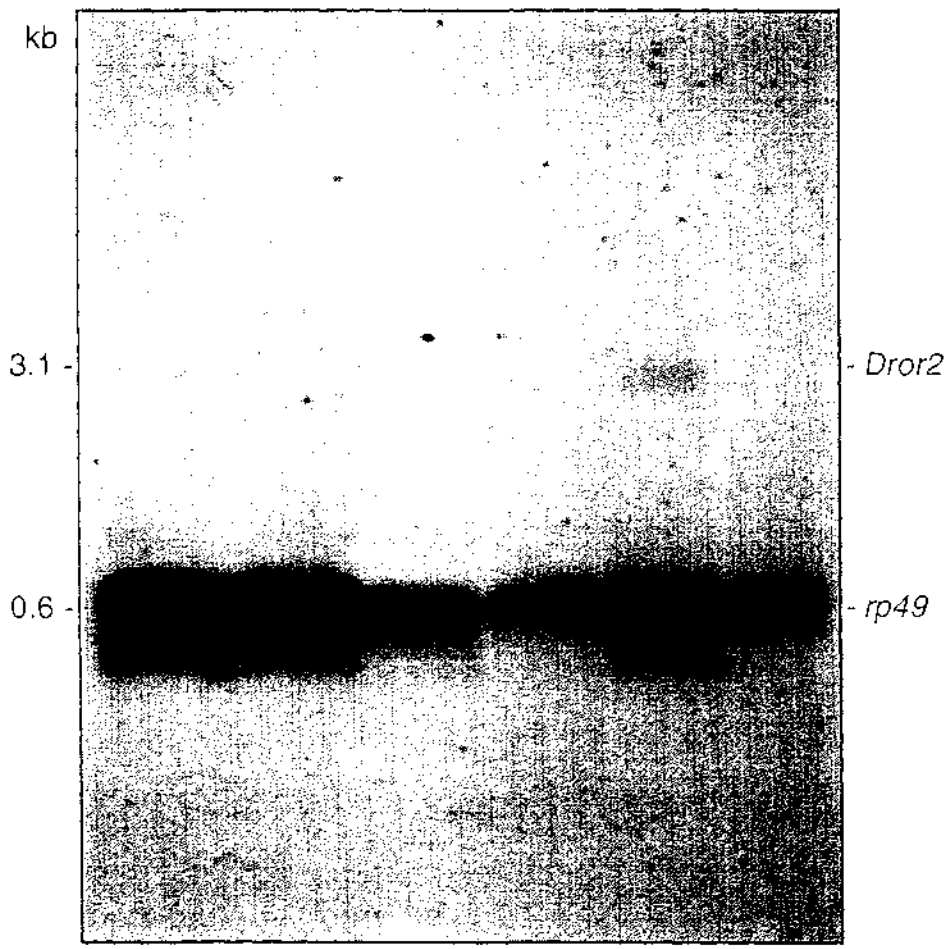
The *rp*49 gene encodes a ribosomal protein of 132 amino acids and contains a major transcript of 0.6 kb that is highly expressed in all stages of development (O'Connell and Rosbash, 1984). A 640 bp *Hind*III/*Eco*RI fragment containing *rp49* was purified from pRP49 and labelled as before. The *rp*49 probe was hybridised to the RNA on two Northern Blots, one containing total RNA and the other poly A^+ RNA. Signals were obtained from both blots, the total RNA signal faint, but strong signals were seen after an overnight exposure to the poly A^+ RNA blot.

The blot revealed that the amount of RNA was not constant for all stages of development, embryos and adult female were detected stronger than pupae, larvae and male adult (Figure 3.19). However, as there was some hybridisation in all lanes the blot was reprobed with the *Dror2* probe. No signal was detected after a 3 day exposure, but after a weeks exposure hybridisation was detected very weakly in the female adult lane as can be seen in Figure 3.19.

In an attempt to increase the sensitivity of the hybridisation, single stranded RNA DIG labelled probes were prepared (Section 2.2.19.2) for *Dror2* and *rp*49. No signal was obtained with either the *Dror2* probe or *rp*49 probe. The DIG System User's Guide for filter hybridisation booklet (BM, 1995) proposes this was probably due to RNase contamination of the DIG blocking solution. Thus, another hybridisation was performed using a mixture of two single stranded DNA probes (Section 2.2.19.3). However, this was also unsuccessful in obtaining a signal.

Figure 3. 19 Autoradiograph of poly A^+ RNA Northern Blotted with a 1.6 kb *Dror2* probe.

Lane 1: 0-4 hour old embryos, lane 2: 4-8 hour old embryos, lane 3: 3rd instar larvae, lane 4: pupae, lane 5: female adult, lane 6: male adult. Northern Blot had a one week exposure.


kb
3.1
0.6
Dror2
rp49

3.8.2 *IN SITU* HYBRIDISATION

3.8.2.1 Using DNA probes

In situ hybridisation was initially carried out using DIG labelled DNA probes (Section 2.2.20.1). The 1.6 kb *Dror2* probe was identical to the one made in Section 3.8.1 except that it was DIG labelled instead of radioactively labelled. Embryos 0-4 hours old and 4-8 hours old were collected and hybridised with the *Dror2* DIG probe, as described in Section 2.2.2.1. No hybridisation could be detected with either batch of embryos. To determine if the lack of hybridisation was due to a problem with the procedure, the embryos were hybridised with a control probe.

The *Drosophila* gene *Hucklebein* (*Hkb*) (Weigel *et al.*, 1990) is expressed in the caps of the poles of early embryos, and in the nervous system later in development. The *Hkb* probe was constructed by excising a 0.6 kb *SalI/XbaI* fragment from pHkb, purified and DIG labelled (Section 2.2.20.1). Again, no hybridisation was detected, so a Southern Blot to plasmid DNA was performed to check the DIG labelling of the probes. Both the probes hybridised strongly as expected. The *in situ* hybridisation was attempted again, but with no success. It was decided to use RNA probes to increase sensitivity of the hybridisation.

3.8.2.2 Using RNA probes

For preparation of *in situ* hybridisation to embryos, five probes were made. The antisense strand of *rp*49 was used as a control on 0-4 hour embryos. To construct the probe pRP49 was linearised by digesting the plasmid with *Eco*RI. The linear DNA was treated following Section 2.2.19.2 and probe synthesised *in vitro* using T3 RNA polymerase.

The sense and antisense strand of *Dror2* were DIG labelled by linearising the plasmids on either side of the cloning site. For the antisense probe pGEM containing the 1.6 kb fragment of *Dror2* was linearised with *Asp*718 and probes synthesised *in vitro* using T7 RNA polymerase. The sense strand was linearised using the same plasmid but digested with *Hind*III and the probe synthesised *in vitro* with T3 RNA polymerase.

The final probes made were the sense and antisense strand of *Dror*. The sense probe was a negative control, the antisense probe was expected to label the central nervous system of 7-11 hour and 10-16 hour embryos (Wilson *et al.*, 1993). A 2055 bp fragment containing the entire open reading frame of the *Dror* gene (Wilson *et al.*, 1993) was amplified by PCR using cDNA from 0-4 hour and 4-8 hour embryos. All solutions and conditions were identical to Section 2.2.15 except only one round of PCR was undertaken.

The resulting 2.0 kb fragment was excised, Qiagen purified, ligated into pGEM and transformed using the standard procedure (Sections 2.2.7.2, 2.2.8.1 and 2.2.10.1). Plasmid DNA was prepared by alkaline lysis (Section 2.2.11.1) and large scale plasmid preparations of clones containing the 2.0 kb insert were performed using the Qiagen kit (Section 2.2.14.2). An antisense probe of pDror was made by linearising the plasmid with *Not*I and the probe synthesised *in vitro* using T7 RNA polymerase. The sense strand was linearised using the same plasmid but digested with *Nco*I and the probe synthesised *in vitro* using T3 RNA polymerase.

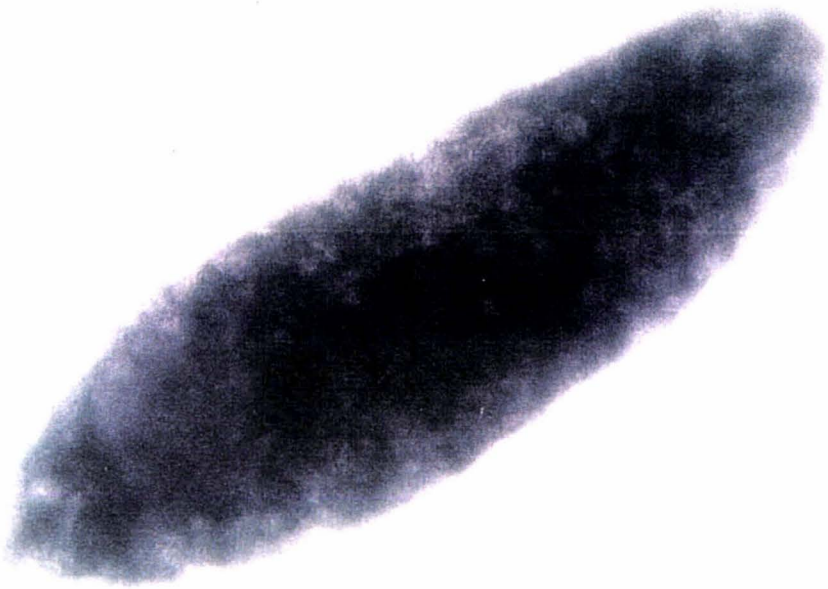
The *rp*49, *Dror* and *Dror2* probes were hybridised to 0-4 hour, 4-8 hour, 7-11 hour and 10-16 hour old embryos. Hybridisation of the *rp49* probe to the embryos was visualised clearly on 0-4 hour embryos as seen in Figure 3.20.

Hybridisation of the *Dror* control at all stages examined was minimal. Signals were detected in staining of the whole embryos but no staining was seen that resembled the results of Wilson *et al.* (1993). The *Dror2* antisense probe revealed hybridisation in the form of general staining from 0-4 hour old embryos (Figure 3.21a) and 4-8 hour old embryos (Figure 3.21b). More specific staining of the trachea is seen in 7-11 hour old embryos (Figure 3.21c) and 10-16 hour old embryos (Figure 3.21d). Some of the embryos at the later stages showed no specific staining (Figure 3.21e), indicating hybridisation was inconsistent. Unfortunately, the same specific staining was seen from the sense *Dror2* probe, suggesting the staining was due to non-specific hybridisation. At this point further *in situ* hybridisation work was abandoned due to time constraints.

Figure 3. 20 *In situ* hybridisation of a whole mount 0-4 hour old embryo using a DIG labelled rp49 RNA probe.

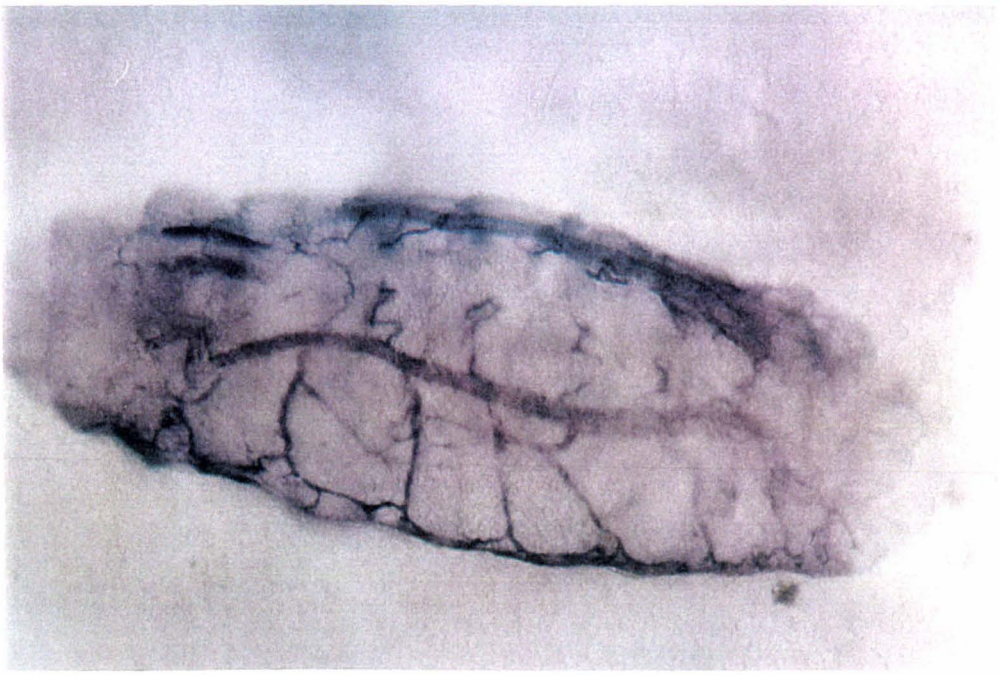
Figure 3. 21 *In situ* hybridisation of whole mount embryos using a DIG labelled Dror2 probe.

(a) 0-4 hour old embryo



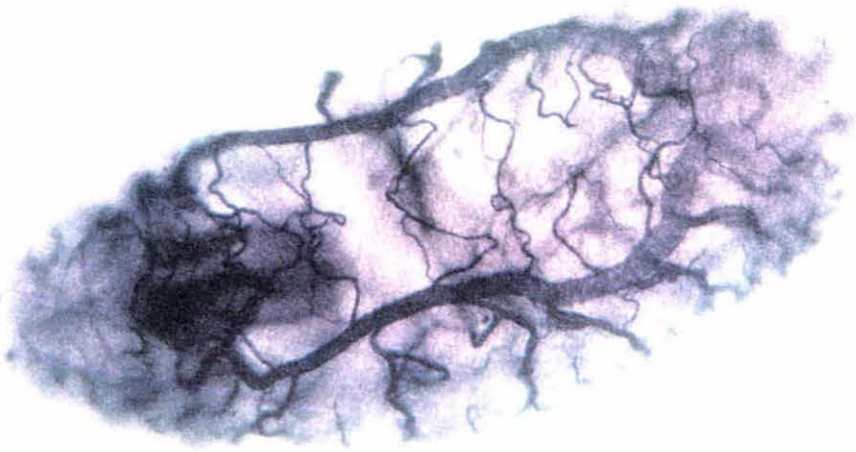
(b) 4-8 hour old embryo

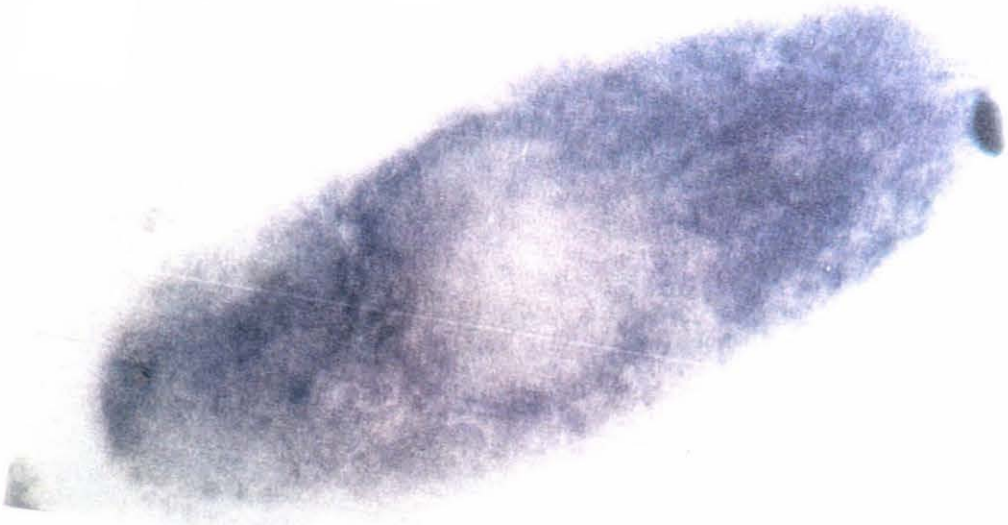
(c) 7-11 hour old embryo



(d) 10-16 hour old embryos. Shows strong hybridisation to the trachea of the embryo.

(e) 10-16 hour embryo. Shows general non-specific hybridisation.





4. DISCUSSION

Receptor tyrosine kinases (RTKs) are a family of cell-surface receptors that play an important role in the transfer of environmental cues essential for cell migration, proliferation and differentiation. *Drosophila melanogaster* possess a number of RTKs, the gene for one of which was isolated for this thesis.

4.1 DTK-LIKE GENES ARE ABSENT FROM *DROSOPHILA*

The original aim of this project was to clone the *Drosophila* homologue of the human *Dtk* gene. To do this PCR was applied using degenerate oligonucleotides derived from conserved amino acid motifs within the tyrosine kinase domain of RTKs. Primers were constructed to specific sequences in the *Dtk* genes as well as to conserved regions of the tyrosine kinase domain of RTKs. This was to ensure a number of RTKs apart from Dtk could be isolated in case the Dtk isolation was unsuccessful. No Dtk clones were isolated from sequencing all the clones from the primer combinations attempted. Thus, it appears that *Drosophila* does not contain a Dtk homologue, however it remains possible that the primers chosen simply did not amplify any Dtk-like genes due to sequence divergence. The project was changed to the isolation and cloning of a novel RTK in *Drosophila*.

4.2 NOVEL SEQUENCES IDENTIFIED

Novel sequences encoding four known genes were isolated as seen in Figure 3.3. The genes identified as the closest to that sequenced do not necessarily represent the actual gene isolated, as only a very small portion of the tyrosine kinase domain had been sequenced. The first gene sequenced focal adhesion kinase (FAK), is a widely expressed protein tyrosine kinase implicated in integrin mediated signal transduction pathways (Hanks *et al.*, 1992). The second gene sequenced was S-

related receptor kinase (SRK). This gene was first sequenced in *Brassica oleracea* and is analagous to the growth factor RTKs in animals (Stein *et al.*, 1991).

The third gene seems to be part of the ech/eph family of receptor tyrosine kinases. The similarity search pulled human eph (Hirai *et al.*, 1987) and eck (Ganju *et al.*, 1994) at the top of the homology, but as this family is the largest of the RTK subfamilies (see Section 1.3.2) it could have been any of them as none have been isolated in *Drosophila* to date. The fourth gene sequenced appeared to be the *Drosophila* homologue of Trk, possibly one of the three Trk genes found in vertebrates (see Section 1.3.3). The sequenced portion of the gene was the most homologous to trkB in mouse (Klein *et al.*, 1990). This gene was chosen to isolate and characterise for this thesis and it was revealed through further sequencing that it was part of the Ror subfamily of RTKs, not the Trk family as previously thought.

4.3 CHARACTERISATION OF THE *DROR2* GENE

4.3.1 INTRONS

There are three introns present in the *Dror2* gene, one in the middle of the cysteine rich domain and one on either side of the kringle domain. The kringle domain is not seen in other RTK subfamilies apart from the Ror subfamily and in one muscle-specific RTK, the electric ray *Torpedo californica* (Jennings *et al.*, 1993) (see Section 1.3.4). The positioning of the two introns either side of the kringle domain is possibly an example of exon shuffling. Exon shuffling is thought to occur because introns are hot spots for genetic recombination because of their length (Gilbert *et al.*, 1997). This phenomenon has been observed in eukaryotes and is thought to be important in gene evolution (Long *et al.*, 1995).

4.3.2 PROMOTER SEQUENCE

A putative TATA box was identified in the 5' UTR of Dror2, only 4/6 nt conformed to the consensus sequence questioning the validity of this identified box. This TATA box is only 13 bp upstream of the transcriptional start point (tsp), usually TATA boxes are 30 bp upstream of the tsp (Lewin 1994). For these reasons, it is possible that the tsp identified was not the actual tsp sequence. If the 5' RACE stopped transcribing at a strong stop point for the reverse transcriptase, an incorrect tsp would have been identified.

4.4 ANALYSIS OF DROR2 PROTEIN SEQUENCE

4.4.1 CONSERVED MOTIFS AND BINDING SITES

Within the extracellular domain of Dror2 are several conserved motifs, the cysteine rich domain and the kringle domain. The intracellular domain contains the conserved catalytic tyrosine domain seen in all RTKs. Within this domain are two putative ATP binding sites and tyrosine containing motifs that may interact with SH2 regions of cellular signalling molecules upon tyrosine phosphorylation. These motifs are not seen in Dror, but have been seen in Dtrk suggesting a similarity with the Trk subfamily of RTKs.

4.4.2 COMPARISONS TO OTHER GENES

Extensive sequence similarity between the Ror family and Dror2 are seen throughout the length of the proteins. However, Dror2 lacks the immunoglobulin-like domain found in the extracellular domain of Ror1 and Ror2. Dror2 also lacks the extensive carboxyl terminus tail found in Ror1 and Ror2. Dror also lacks these features suggesting an evolutionary change between the *Drosophila* and human counterparts.

The kinase domain is the most conserved domain between the Ror subfamily and TrkB with a number of the shared amino acids in this domain not found in many other RTKs. This suggests the occurrence of a Trk-like superfamily of RTKs.

Similarities between the Ror subfamily suggest that Ror1 and Ror2 are the mammalian homologues of Dror and Dror2. To date no ligands have been identified for the Ror subfamily of RTKs neither for *Torpedo* RTK (the muscle specific RTK).

4.5 *DNRK* GENE

While this thesis was in preparation, a paper was published by Oishi *et al.* (1997) reporting the cloning of a cDNA for a gene called *Dnrk* that is identical *Dror2*.

4.5.1 DNRK SEQUENCE DIFFERENCES

Upon comparing the nucleotide and amino acids in Dror2 and Dnrk it was revealed that there are 15 nucleotide differences resulting in a number of amino acid differences as seen in Table 4.1. The first amino acid difference is from a histidine in Dror2 to a tyrosine in Dnrk. Though this is not a conserved change, it is possibly a result of polymorphisms. The second and third differences are from a glycine in Dror2 to a valine in Dnrk which are conservative changes resulting from polymorphisms. The fourth amino acid difference is from a threonine in Dror2 to a methionine in Dnrk, not a conservative change but again possibly a result of polymorphisms.

The second to last nucleotide difference results in a frameshift of the coding sequence. In Dnrk, a C is missing from the sequence, which we believe to be a sequencing error. Both electropherograms of Dror2 clearly show a C (or G in the other direction), proving the presence of this extra C. This frameshift in turn, alters

all the remaining amino acids, in this case the first change is from an arginine in Dror2 to a glutamic acid in Dnrk.

A comparison between Dror2, Dnrk and other related RTKs was carried out (see Table 4.2) In Dror2, the amino acid glutamic acid at position 1 is a conservative change from aspartic acid, whereas Dnrk has an arginine at that position that is not a conservative change. At position 2, all receptors have an isoleucine except Dnrk that has a serine. At position 6, the isoleucine in Dror2 is a conservative change from leucine, whereas the serine in Dnrk is not. The serine at position 13 is seen in all compared receptors except Dror and Dnrk again showing a conserved residue. All these changes lead to the conclusion that the sequence of Dnrk is missing a C resulting in a frameshift of the sequence. Hence, the sequence of Dror2 is believed to be correct.

The rest of the nucleotide differences result in no changes of codons, as the change is in the third base, the wobble base. This base is usually degenerate, so no change in the corresponding codons occurs.

4.5.2 DNRK TRANSLATIONAL START SITE

The initiating methionine Oishi *et al.* (1997) identified is the second in-frame ATG identified (see Section 3.5.4.1). The first in-frame ATG was not identified by Oishi *et al.* (1997) because their cDNA clone begins 51 bp downstream of the putative translational start site.

Table 4.1 Nucleotide and amino acids changes between Dror2 and Dnrk.

Dror2		**Dnrk**		**Amino acid change**	
Position	Nucleotide	Position	Nucleotide	Dror2	Dnrk
152	C	104	T	-	-
483	C	435	T	H	Y
574	C	526	T	-	-
744	A	696	G	-	-
768	A	720	G	-	-
819	C	771	T	G	V
917	G	869	T	-	-
1062	T	1014	C	-	-
1160	G	1012	T	G	V
1172	C	1124	T	T	M
1875	G	1827	A	-	-
1908	T	1860	G	-	-
2064	G	2016	T	-	-
2115	C	2067	-	R	E
2172	C	2124	T	-	-

Table 4.2 Comparison between amino acid sequence of Dnrk, Dror2 and related RTKs after frameshift mutation.

Numbers refer to amino acids after the frameshift mutations. Dnrk, Dror2, Dror from *Drosophila*, Ror1 and Ror2 from humans and TrkB from mice. Asterisks (*) indicates a stop codon.

Position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Dror2	E	I	N	H	C	I	Q	H	S	I	A	E	S	E	C	K	A	M	L	*						
Dnrk	R	S	T	T	A	S	S	T	A	S	P	R	A	S	A	R	Q	C	F	R	G	L	P	E	K	*
Dror	D	I	S	N	R	L	K	T	W	H	E	G	H	F	K	A	S	N	P	E	M	*				
Ror1	D	I	H	V	R	L	K	T	W	E	G	L	S	S	H	T	S	S	T	T	P	S	G	G	N	A
Ror2	D	I	H	S	R	L	R	A	W	G	N	L	S	N	Y	N	S	S	A	Q	T	S	G	A	S	N
TrkB	N	I	H	T	L	L	Q	N	L	A	K	A	S	P	V	Y	L	G	*							

4.5.3 DNRK SIGNAL PEPTIDE

If the translational start site of Dnrk is incorrect, the signal peptide will also be wrong. A comparison between the two possible signal peptides of Dror2 (Table 3.3, Table 3.4) showed that SP2 was a better signal peptide compared to SP1. Thus, although the first in-frame ATG is a much better match to the consensus translation start site than the second, it is possible that translation may begin at the second ATG because SP2 is a better match to a consensus signal peptide.

4.5.4 EXPRESSION OF DNRK

4.5.4.1 Northern Blot Analysis

Oishi *et al.* (1997) performed Northern Blot analysis successfully from total RNA isolated from a range of developmental stages using radioactively labelled Dnrk (the whole ORF) as the probe. Expression was detected at high levels in 4-22 hour old embryos, larva and pupa. The highest level of expression was in the larva where restructuring of the nervous system is occurring. No expression was seen in adults, contradicting the results seen in Dror2 analysis. Northern Blot analysis of Dror2 showed expression only in female adults, but 3' RACE revealed some expression at all stages of development including adult flies. RT-PCR is very sensitive, but not quantitative under conditions employed. Thus, the Northern blots may give a better estimate of the relative abundance of Dror2 RNA during development. That I detected expression only in female adults may have been due to the amount of RNA loaded in that lane as indicated by the strong signal of rp49 (see Figure 3.18). Oishi *et al.* (1997) didn't show the rp49 hybridisation control in their results, but if the signal was not consistent in all stages (it was indicated that pupae showed less expression) then hybridisation may not have been detected in the adult stages.

4.5.4.2 *In situ* hybridisation to whole mount embryos

Oishi *et al.* (1997) performed *in situ* hybridisation successfully, no specific expression was observed in stages before germ band elongation. At Stage 10, weak expression was detected in the ventral area of the germ band, this expression became stronger at Stage 11. At this stage expression was limited to the layer of neural progenitor cells between the epidermal and mesodermal cell layers. This restricted expression was constant throughout the remainder of embryogenesis, solely in the neural cell lineage, resulting in expression in the brain and ventral nerve cord. At Stage 17, expression was also seen in cells corresponding to the peripheral nervous system.

These results revealed expression exclusively in the nervous system during embryogenesis. This was identical to Dror expression (Wilson *et al.*, 1993), strengthening the probability of a role in the nervous system in the developing embryo. Due to the delayed expression, the role of Dnrk and Dror may involve organisation of the differentiated cells of the nervous system, both in the peripheral and central cell lines. These receptors may also play a role in the later processes of neuronal cell survival as expression is seen throughout embryogenesis and subsequent developmental stages.

4.5.5 TYROSINE KINASE ACTIVITY

In vitro kinase analysis was performed on Dnrk to study the protein kinase activity, this showed that Dnrk did possess tyrosine kinase activity. Expression of HA-tagged cytoplasmic domains of Dnrk was examined to reveal which ATP-binding motif was required for the kinase activity. It was revealed that the proximal ATP-binding motif on its own is sufficient for the catalytic activity, but the distal motif may be required for full activity of Dnrk. *In situ* hybridisation to salivary gland

chromosomes revealed that Dnrk maps to the 49f region on the right arm of the second chromosome.

4.6 POSSIBLE FUNCTION OF ROR-TYPE RTKS

The Ror genes in mammals are expressed early during embryogenesis and expression decreases dramatically at later stages of embryonic development (Masiakowiski and Carroll, 1992). In *Drosophila*, both Dror and Dror2 (Dnrk) are specifically expressed in the developing nervous system. In contrast to this expression, trkB expression is low at the earlier stages of development and increases in the later stages of embryonic development (Klein *et al.*, 1990). This differing expression pattern may reveal differing roles for the Trk and Ror subfamilies of RTKs.

The muscle-specific RTK MuSK, related structurally to the Ror family, is specifically expressed in early myotomes and developing muscle and becomes localised to the neuromuscular junction upon muscular maturation (Valenzuela *et al.*, 1995). Recently it has been demonstrated that a targeted disruption of the *MuSK* gene in mice leads to profound defects in the formation of neuromuscular junctions (Glass *et al.*, 1996; DiChiara *et al.*, 1996). These discoveries reveal a critical role of MuSK in synapse formation at neuromuscular junctions and as this gene is related to the Ror family, it is conceivable that these RTKs may also play an important role in synapse formation in the nervous system.

4.7 FUTURE EXPERIMENTS

4.7.1 CONFIRMATION OF C-TERMINAL OF DROR2

As the TATA box identified did not match well to the consensus sequence, the C-terminal of Dror2 will need to be confirmed. To do this SI nuclease protection

and/or primer extension will be performed. The gene regulation of this promoter may also be examined..

4.7.2 EXPRESSION OF DROR2

If this project continues *in situ* hybridisation will be attempted again as Oishi *et al.* (1997) showed it can be done successfully. All techniques will be analysed and compared to procedures undertaken by Oishi *et al.* (1997).

4.7.3 ISOLATION OF MUTATIONS

A future goal may be to look at isolating mutations in the Dror2 gene. Dror -/- mutants develop normally (Wilson, personal communication). It is possible that Dror2 and Dror are functionally redundant as they have similar extracellular domains and are expressed in the same cells. If this is the case, Dror2 -/- mutants would also develop normally, but Dror -/-, Dror2 -/- double mutants may disrupt development of the nervous system.

The gene product of Dror2 may be required for development of certain neurons. Though a number of genes have been identified as important for the development of the nervous system, many other essential genes have yet to be isolated. For example, the *daughterless* gene is a proneural gene involved in the transcriptional regulation of the sensory organ formation in the PNS. Loss of function mutations in the *daughterless* gene in *Drosophila* leads to a decreased number of sensory organ precursors (Caudy *et al.*, 1988).

Another goal would be to identify factors that control tissue specific expression of Dror2. Promoter fusions would be constructed to reporter genes. Analysis would be carried out on these reporter fusions, for example promoter deletions, to identify DNA sequences needed for normal expression of neurons. Another way to

examine factors controlling expression would be to look at the expression of reporter constructs to *Drosophila* loss of function mutants. In particular, mutations in genes known to be required for the correct formation of the nervous system would be examined.

APPENDIX

Table of 133 *Drosophila* clones sequenced from a variety of primer combinations

CLONE NUMBER	DNA USED	PRIMER 1	PRIMER 2	EXPECTED SIZE	SIZE OBSERVED	SEQUENCE
1	Embryo cDNA	DO1	DO2	150bp	170bp	Dsrc28C
2	Embryo cDNA	DO1	DO2	150bp	170bp	FGF-R
3	Embryo cDNA	DO1	DO2	150bp	170bp	Dsrc28C
4	Adult cDNA	9474	9476	150bp	180bp	Dsrc28C
5	Genomic DNA	DO1	DO2	150bp	150bp	FGF-R
6	Genomic DNA	DO1	DO2	150bp	150bp	FGF-R
7	Embryo cDNA	DO1	DO2	150bp	150bp	Dsrc28C
8	Embryo cDNA	DO1	DO2	150bp	150bp	FGF-R
9	Embryo cDNA	DO1R	Dtk1	410bp	480bp	Nothing relevant
10	Embryo cDNA	DO1R	Dtk1	410bp	480bp	Same as 9
11	Embryo cDNA	DO1	Dtk3	225bp	225bp	Dsrc28C
12	Embryo cDNA	DO1	Dtk3	225bp	250bp	Nothing relevant
13	Embryo cDNA	Dtk1	Dtk3	618bp	225bp 410bp	Dsrc28C
14	Genomic DNA	DO1	Dtk3	225bp	300bp	Forward short nothing
15	Embryo cDNA	DO1	Dtk3	225bp	240bp	Dsrc28C
16	Embryo cDNA	DO1	Dtk3	225bp	225bp	Leave
17	Embryo cDNA	Dtk1	Dtk3	618bp	220bp 410bp	Dsrc28C
18	Embryo cDNA	Dtk1	Dtk3	618bp	220bp 410bp	Dsrc28C
19	Genomic DNA	DO1	Dtk3	225bp	310bp	Same as 14
20	Genomic DNA	DO1	Dtk3	225bp	250bp	Forward short nothing

CLONE NUMBER	DNA USED	PRIMER 1	PRIMER 2	EXPECTED SIZE	SIZE OBSERVED	SEQUENCE
21	Genomic DNA	DO1	Dtk3	225bp	250bp	Nothing F or R short
22	Genomic DNA	Dtk2	DO3	350bp	400bp	Nothing relevant
23	Genomic DNA	DO1	Dtk3	225bp	350bp	Same as 14
24	Genomic DNA	DO1	Dtk3	225bp	350bp	Same as 14
25	Genomic DNA	DO1	Dtk3	225bp	350bp	Same as 14
26	Genomic DNA	DO1	Dtk3	225bp	350bp	Same as 14
27	Genomic DNA	DO1	Dtk3	225bp	350bp	Same as 14
28	Genomic DNA	Dtk2	DO3	350bp	410bp	Dorsal
29	Embryo cDNA	DO1	Dtk3	225bp	225bp	Forward short nothing
30	Embryo cDNA	DO1	Dtk3	225bp	260bp	Leave
31	Embryo cDNA	DO1	Dtk3	225bp	260bp	Leave
32	Embryo cDNA	DO1	Dtk3	225bp	240bp	Nothing relevant
33	Embryo cDNA	DO1	Dtk3	225bp	240bp	Dsrc28C
34	Embryo cDNA	Dtk1	Dtk3	618bp	450bp	Leave
35	Embryo cDNA	Dtk1	Dtk3	618bp	450bp	Dsrc28C
36	Embryo cDNA	Dtk1	Dtk3	618bp	450bp	Dsrc28C
37	Embryo cDNA	Dtk1	Dtk3	618bp	225bp 450bp	Dsrc28C
38	Embryo cDNA	Dtk1	Dtk3	618bp	450bp	Dsrc28C
39	Embryo cDNA	DO1	DO3	200bp	240bp	Forward short nothing
40	Genomic DNA	DO1	DO3	200bp	240bp	Forward short nothing
41	Genomic DNA	DO1	DO3	200bp	240bp Faint	Forward short nothing
42	Genomic DNA	DO1	DO3	200bp	240bp Faint	Forward long nothing
43	Genomic DNA	DO1	DO3	200bp	240bp	Aligned AA: TRKA rat NGF
44	Genomic DNA	DO1	DO3	200bp	240bp	Nothing relevant
45	Genomic DNA	DO1	DO3	200bp	240bp Faint	Forward short nothing

CLONE NUMBER	DNA USED	PRIMER 1	PRIMER 2	EXPECTED SIZE	SIZE OBSERVED	SEQUENCE
46	Embryo cDNA	DO1	DO3	200bp	240bp	Leave
47	Embryo cDNA	DO1	DO3	200bp	240bp	Aligned AA: FAK
48	Embryo cDNA	DO1	DO3	200bp	240bp	Forward long nothing
49	Embryo cDNA	DO1	DO3	200bp	240bp	Dsrc28C
50	Embryo cDNA	DO1	DO3	200bp	240bp	Dsrc28C
51	Adult cDNA	DO1	DO3	200bp	240bp	Aligned AA: Srk gene
52	Adult cDNA	DO1	DO3	200bp	240bp Faint	Drosophila sequence leave
53	Adult cDNA	DO1	DO3	200bp	220bp	Transposon HB1
54	Genomic DNA	DO1	DO3	200bp	260bp	Forward short nothing
55	Genomic DNA	DO1	DO3	200bp	240bp	Drosophila sequence leave
56	Genomic DNA	DO1	DO3	200bp	260bp Faint	Forward short nothing
57	Embryo cDNA	DO1	DO3	200bp	240bp	Forward short nothing
58	Embryo cDNA	DO1	DO3	200bp	240bp	Aligned AA: Fak Mouse /Human
59	Embryo cDNA	DO1	DO3	200bp	260bp	Forward short nothing
60	Embryo cDNA	DO1	DO3	200bp	260bp	Aligned AA Eph/ech receptor
61	Embryo cDNA	DO1	DO3	200bp	280bp	Forward short nothing
62	Adult cDNA	DO1	DO3	200bp	260bp	Drosophila sequence leave
63	Adult cDNA	DO1	DO3	200bp	240bp	Copia like transposon
64	Adult cDNA	DO1	DO3	200bp	225bp	Hopscotch
65	Adult cDNA	DO1	DO3	200bp	260bp	Nothing relevant
66	Adult cDNA	DO1	DO3	200bp	255bp	Forward short nothing
67	Adult cDNA	DO1	DO3	200bp	255bp	Mouse focal adhesion kinase
68	Adult cDNA	DO1	DO3	200bp	260bp	Mouse focal adhesion kinase
69	Genomic DNA	DO1	DO3	200bp	240bp Faint	Leave
70	Genomic DNA	DO1	DO3	200bp	225bp	Mouse focal adhesion kinase

CLONE NUMBER	DNA USED	PRIMER 1	PRIMER 2	EXPECTED SIZE	SIZE OBSERVED	SEQUENCE
71	Genomic DNA	DO1	DO3	200bp	240bp	Forward short nothing
72	Genomic DNA	DO1	DO3	200bp	260bp	Nothing relevant
73	Genomic DNA	DO1	DO3	200bp	240bp	Drosophila sequence leave
74	Genomic DNA	DO1	DO3	200bp	260bp faint	Forward short nothing
75	Embryo cDNA	DO1	DO3	200bp	260bp faint	Forward short nothing
76	Embryo cDNA	DO1	DO3	200bp	260bp	Leave
77	Adult cDNA	DO1	DO3	200bp	260bp	Forward short nothing
78	Genomic DNA	DO1	DO3	200bp	240bp	Forward short nothing
79	Genomic DNA	DO1	DO3	200bp	240bp	Same as 78
80	Embryo cDNA	DO1	DO3	200bp	260bp	Forward short nothing
81	Embryo cDNA	DO1	DO3	200bp	260bp	Same as 80
82	Embryo cDNA	DO1	DO3	200bp	240bp	Mouse focal adhesion kinase
83	Adult cDNA	DO1	DO3	200bp	260bp	Nothing relevant
84	Adult cDNA	DO1	DO3	200bp	260bp	Dror
85	Adult cDNA	DO1	DO3	200bp	260bp	Mouse focal adhesion kinase
86	Adult cDNA	DO1	DO3	200bp	240bp	Mouse Focal adhesion kinase
87	Genomic DNA	DO1	DO3	200bp	Very faint	Forward short nothing
88	Embryo cDNA	DO1	DO3	200bp	Very faint	Nothing relevant
89	Adult cDNA	DO1	DO3	200bp	Very faint	D-ret
90	Adult cDNA	DO1	DO3	200bp	Very faint	hopscotch
91	Adult cDNA	DO1	DO3	200bp	Very faint	Forward short nothing
92	Genomic DNA	DO1	DO3	200bp	200bp	Forward short nothing
93	Genomic DNA	DO1	DO3	200bp	200bp	Nothing relevant
94	Genomic DNA	DO1	DO3	200bp	200bp	Forward short nothing
95	Genomic DNA	DO1	DO3	200bp	200bp	Drosophila sequence leave

CLONE NUMBER	DNA USED	PRIMER 1	PRIMER 2	EXPECTED SIZE	SIZE OBSERVED	SEQUENCE
96	Genomic DNA	DO1	DO3	200bp	200bp	Forward short nothing
97	Embryo cDNA	DO1	DO3	200bp	Very faint	Nothing relevant
98	Embryo cDNA	DO1	DO3	200bp	220bp	Leave unclear sequence
99	Embryo cDNA	DO1	DO3	200bp	220bp	Nothing relevant
100	Embryo cDNA	DO1	DO3	200bp	Very faint	Nothing relevant
101	Embryo cDNA	DO1	DO3	200bp	220bp	Forward short nothing
102	Adult cDNA	DO1	DO3	200bp	260bp	Nothing relevant
103	Adult cDNA	DO1	DO3	200bp	250bp	Srk gene
104	Adult cDNA	DO1	DO3	200bp	210bp	Forward short nothing
105	Adult cDNA	DO1	DO3	200bp	240bp	Nothing relevant
106	Adult cDNA	DO1	DO3	200bp	250bp	Leave unclear sequence
107	Adult cDNA	DO1	DO3	200bp	200bp	Nothing relevant
108	Genomic DNA	DO1	DO3	200bp	220bp	Nothing relevant
109	Genomic DNA	DO1	DO3	200bp	220bp	Nothing relevant
110	Genomic DNA	DO1	DO3	200bp	240bp	Nothing relevant
111	Genomic DNA	DO1	DO3	200bp	200bp	Nothing relevant
112	Genomic DNA	DO1	DO3	200bp	220bp	Forward short nothing
113	Genomic DNA	DO1	DO3	200bp	240bp Faint	Forward short nothing
114	Embryo cDNA	DO1	DO3	200bp	220bp	Ech/eph receptor
115	Embryo cDNA	DO1	DO3	200bp	220bp	Nothing relevant
116	Embryo cDNA	DO1	DO3	200bp	220bp Faint	Nothing relevant
117	Embryo cDNA	DO1	DO3	200bp	220bp Faint	Forward short nothing
118	Embryo cDNA	DO1	DO3	200bp	220bp Faint	Nothing relevant
119	Embryo cDNA	DO1	DO3	200bp	220bp	Nothing relevant
120	Adult cDNA	DO1	DO3	200bp	240bp	Nothing relevant

CLONE NUMBER	DNA USED	PRIMER 1	PRIMER 2	EXPECTED SIZE	SIZE OBSERVED	SEQUENCE
121	Adult cDNA	DO1	DO3	200bp	240bp	Forward short nothing
122	Adult cDNA	DO1	DO3	200bp	220bp	Nothing relevant
123	Adult cDNA	DO1	DO3	200bp	240bp	Eph/ech receptor
124	Adult cDNA	DO1	DO3	200bp	240bp Faint	Eph/ech receptor
125	Adult cDNA	DO4	DO3	170bp	100bp 80bp	Forward short nothing
126	Adult cDNA	DO4	DO3	170bp	170bp	Eph/ech receptor
127	Adult cDNA	DO4	DO3	170bp	175bp	Eph/ech receptor
128	Adult cDNA	DO4	DO3	170bp	170bp	Mouse focal adhesion kinase
129	Adult cDNA	DO4	DO3	170bp	170bp	Mouse focal adhesion kinase
130	Embryo cDNA	DO4	DO3	170bp	200bp	Leave
131	Adult cDNA	DO4	DO3	170bp	200bp	Mouse focal adhesion kinase
132	Embryo cDNA	DO4	DO3	170bp	200bp	Mouse focal adhesion kinase
133	Adult cDNA	DO4	DO3	170bp	200bp	Mouse focal adhesion kinase

5. REFERENCES

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, P.D., Seidman, J.G., Smith, J.A. and Struhl, K. (Eds) (1997) *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc.

Avruch, J., Zhang, X. and Kyriakis, J.M. (1994) Raf meets Ras: completing the framework of a signal transduction pathway. *TIBS* **19**: 279-283.

Barbacid, M. (1993) Nerve growth factor: a tale of two receptors. *Oncogene* **8**: 2033-2042.

Birnsteil, M.L., Busslinger, M. and Strub, K. (1985) Transcription termination and 3' processing: the end is in site. *Cell* **41**: 349-359.

Borck, K., Beggs, J.D., Brammer, W.J., Hopkins, A.S. and Murray, N.E. (1976) The construction *in vitro* of transducing derivatives of phage lambda. *Mol Gen Genet* **146**: 199-207.

Brock, H.W. (1987) Sequence and genomic structure of *ras* homologues Dm*ras*85D and Dm*ras*64B of *Drosophila melanogaster*. *Gene* **51**: 129-137.

Brunner, D., Oellers, N., Szadad, J., Biggs III, W.H., Zipursky, S.L. and Hafen, E. (1994) A gain-of-function mutation in *Drosophila* MAP kinase activates multiple receptor tyrosine kinase signalling pathways. *Cell* **76**: 875-888.

Bullock, W.O., Fernandez, J.M. and Short, J.M. (1987) XL-1 Blue: A high efficiency plasmid transforming *rec*A *Escherichia coli* with beta-galactosidase selection. *Biotechniques* **5**: 376-378.

Callahan, C.A., Muralidhar, M.G., Lundgren, S.E., Scully, A.L. and Thomas, J.B. (1995) Control of neuronal pathway selection by a *Drosophila* receptor protein-tyrosine kinase family member. *Nature* **376**: 171-174.

Caudy, M., Grell, E.H., Dambly-Chaudiere, C., Ghysen, A., Jan, L.Y. and Jan, Y.N. (1988) The maternal sex determination gene *daughterless* has zygotic activity necessary for formation of peripheral neurons in *Drosophila*. *Genes Dev* **2**: 843-852.

Cavener, D.R. (1987) Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucl Acids Res* **15**: 1353-1361.

Cavener, D.R. and Ray, S.C. (1991) Eukaryotic start and stop translation sites. *Nucl Acids Res* **19**: 3185-3192.

Crosier, P.S., Lewis, P.M., Hall, L.R., Vitas, M.R., Morris, C.M., Beier, D.R., Wood, C.R. and Crosier, K.E. (1994a) Isolation of a receptor tyrosine kinase (*Dtk*) from embryonic stem cells: structure, genetic mapping and analysis of expression. *Growth factors* **11**: 125-136.

Crosier, K.E., Hall, L.R., Lewis, P.M., Morris, C.M., Wood, C.R., Morris, J.C. and Crosier, P.S. (1994b) Isolation and characterization of the human *Dtk* receptor tyrosine kinase. *Growth factors* **11**: 137-144.

DiChiara, T.M., Bowen, D.C., Valenzuela, D.M., Simmons, M.V., Poueymirou, W.T., Thomas, S., Kinetz, E., Compton, D.L., Rojas, E., Park, J.S., Smith, C., DiStefano, P.S., Glass, D.J., Burden, S.J. and Yancopoulos, G.D. (1996) The receptor tyrosine kinase MuSK is required for neuromuscular junction formation *in vivo*. *Cell* **85**: 501-512.

Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nuc Acids Res* **16**: 6127-6145.

Dura, J-M., Taillebourg, E. and Preat, T. (1995) The *Drosophila* learning and memory gene *linotte* encodes a putative receptor tyrosine kinase homologous to the human RYK gene product. *FEBS Letters* **370**: 250-254.

Fantl, W.J., Johnson, D.E. and Williams, L.T. (1993) Signalling by receptor tyrosine kinases. *Annu Rev Biochem* **62**: 453-481.

Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* **85**: 8998-9002.

Ganju, P., Shigemoto, K., Brennan, J., Entwistle, A. and Reith, A.D. (1994) The Eck receptor tyrosine kinase is implicated in pattern formation during gastrulation, hindbrain, segmentation and limb development. *Oncogene* **9**: 1613-1624.

Gilbert, W., de Souza, S.J. and Long, M. (1997) Origin of genes. *Proc Natl Acad Sci USA* **94**: 7698-7703.

Glass, D.J. and Yancopoulos, G.D. (1993) The neurotrophins and their receptors. *Trends Cell Biol* **3**: 262-268.

Glass, D.J., Bowen, D.C., Stitt, T.N., Radziejewski, C., Bruno, J., Ryan, T.E., Gies, D.R., Shah, S., Mattson, K., Burden, S.J., DiStefano, P.S., Valenzuela, D.M., DiChiara, T.M. and Yancopoulos, G.D. (1996) Agrin acts via a MuSK receptor complex. *Cell* **85**: 513-523.

Grimes, M., Zhou, J., Holtzman, D. and Mobley, W.C. (1993) Neurotrophin signalling in the nervous system. *Seminars in the Neurosciences* **5**: 239-247.

Hafen, E., Basler, K., Edstroem, J.E. and Rubin, G.M. (1987) *Sevenless*, a cell-specific homeotic gene of *Drosophila,* encodes a putative transmembrane receptor with a tyrosine kinase domain. *Science* **236**: 55-63.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Bio.* **166**: 557-580.

Hanks, S.K., Calalb, M.B., Harper, C. and Patel, S.K. (1992) Focal adhesion protein-kinase phosphorylated in response to cell attachment to fibronectin. *Proc Natl Acad Sci USA* **89**: 8487-8491.

Hirai, H., Maru, Y., Hagiwara, K., Nishida, J. and Takaku, F. (1987) A novel putative tyrosine kinase receptor encoded by the *eph* gene. *Science* **238**: 1717-1720.

Hultmark, D. Klemenz, R. and Gehring, W.J. (1986) Translational and transcriptional control elements in the untranslated leader of the heat-shock gene *hsp22*. *Cell* **44**: 429-438.

Jennings, C.G.B., Dyer, S.M and Burden, S.J. (1993) Muscle-specific *trk*-related receptor with a kringle domain defines a distinct class of receptor tyrosine kinases. *Proc Natl Acad Sci USA* **90**: 2895-2899.

Jiang, J., Kosman, D., Ip, Y.T. and Levine, M. (1991) The dorsal morphogen regulates the mesoderm determinant twist in early *Drosophila* embryos. *Genes Dev* **10**: 1881-1891.

Kanalas, J.J. and Makker, S.P. (1991) Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor site for plasminogen. *J Bio Chem* **266**: 10825-10829.

Klein, R., Conway, D., Parada, L.F. and Barbacid, M. (1990) The *trkB* tyrosine kinase gene codes for a second neurogenic receptor that lacks the catalytic kinase domain. *Cell* **61**: 647-656.

Lamballe, F., Klein, R. and Barbacid, M. (1991) *trkC*, a new member of the *trk* family of tyrosine protein kinases, is a receptor for neurotrophin-3. *Cell* **66**: 967-979.

Lewin, B. (1994) *Genes V* Oxford University Press, New York.

Long, M., de Souza, S.J. and Gilbert, W. (1995) Evolution of the intron-exon structure of eukaryotic genes. *Curr Opin Genet Dev* **5**: 774-778.

Martin-Zanca D., Hughes, S.H. and Barbacid, M. (1986) A human oncogene formed by the fusion of truncated tropomyosin and protein tyrosine kinase sequences. *Nature* **319**: 743-748.

Masiakowski, P. and Carroll, R.D. (1992) A novel family of cell surface receptors with tyrosine kinase-like domain. *J Bio Chem* **267**: 26181-26190.

McLean, J.W., Tomlinson, J.E., Kuang, W-J, Eaton, D.L., Chen, E.Y., Gunther, M.F., Scanu, A.M. and Lawn, R.M. (1987) cDNA sequence of human apolipoprotein(a) is homologous to plasminogen. *Nature* **300**: 132-137.

Messing, J. (1983) New M13 vectors for cloning. *Methods Enzymol* **101**: 20-73.

Miller, J.H. (Ed) (1972) Experiments in molecular genetics. In: *Cold Spring Harbour Laboratory Press*. Cold Spring Harbour, New York.

Mount, S.M. (1982) A catalogue of splice junction sequences. *Nucl Acids Res* **10**: 459-472.

Mozer, B., Marlor, R., Parkhurst, S. and Corces, V. (1985) Characterisation and developmental expression of a *Drosophila ras* oncogene. *Mol Cell Biol* **5**: 885-889.

Muller B.K., Bonhoeffer, F. and Drescher, U. (1996) Novel gene families involved in neural pathfinding. *Curr Opin Genet Dev* **6**: 649-474.

Nakamura, T., Tsutomu, T., Hagiya, M., Seki, T., Shimomishi, M., Atsushi, S., Tashiro, K. and Shimizu, S. (1989) Molecular cloning and expression of human hepatocyte growth factor. *Nature* **342**: 440-443.

O'Connell, P. and Rosbash, M. (1984) Sequence, structure and codon preference of the *Drosophila* ribosomal protein 49 gene. *Nucl Acids Res* **12**: 5495-5514.

Oishi, I., Sugiyama, S., Liu, Z-J., Yamamura, H., Nishida, Y. and Minami, Y. (1997) A novel *Drosophila* receptor tyrosine kinase expressed specifically in the nervous system. *J Bio Chem* **272**: 11916-11923.

Orike, N. and Pini, A. (1996) Axon guidance: Following the Eph plan. *Curr Bio* **6**: 108-110.

Pandey, A., Lindberg, R.A. and Dixit, V.M. (1995) Receptor orphans find a family. *Curr Bio* **5**: 986-989.

Patthy, L. (1985) Evolution of the proteases of blood coagulation and fibrinolysis by assembly from modules. *Cell* **41**: 657-663.

Patthy, L., Trexler, M., Vali, Z., Banyzi, L. and Varadi, A. (1984) Kringles: modules specialized for protein binding. *FEBS Letters* **171**: 131-136.

Pawson, T. and Bernstein, A. (1990) Receptor tyrosine kinases: genetic evidence for their role in *Drosophila* and mouse development. *TIG* **6**: 350-356.

Pearson, R.B. and Kemp, B.E. (1991) Protein kinase phosphorylation site sequences and consensus specificity motifs: Tabulations. *Methods Enzymol* **200**: 62-81.

Perrimon, N. and Desplan, C. (1994) Signal transduction in the early *Drosophila* embryo: when genetics meets biochemistry. *TIBS* **19**: 509-512.

Polvi, A., Armstrong, E., Lai, C., Lemke, G., Huebner, K., Spitz, R.A., Nicholls, R.D. and Alitalo, K. (1993) The human TYRO3 gene and pseudogene are located in chromosome 15q14-q25. *Gene* **134**: 289-293.

Proudfoot, N.J. and Brownlee, G.G. (1976) 3' Non-coding region sequences in eukaryotic messenger RNA. *Nature* **263**: 211-214.

Pulido, D., Campuzano, S., Koda, T., Modolell, J. and Barbacid, M. (1992) *Dtrk*, a *Drosophila* gene related to the *trk* family of neurotrophin receptors, encodes a novel class of neural cell adhesion molecule. *EMBO J* **11**: 391-404.

Raz, E. and Shilo, B.Z. (1993) Establishment of ventral cell fates in the *Drosophila* embryonic ectoderm requires DER, the EGF receptor homolog. *Genes Dev* **7**: 1937-1948.

Riggleman, B., Wieschaus, E. and Schedl, P. (1989) Molecular analysis of the *armadillo* locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a *Drosophila* segment polarity gene. *Genes Dev* **3**: 96-113.

Sambrook, J., Fritch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Second Edition. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.

Schupbach, T. (1987) Germline and soma cooperate during oogenesis to establish the dorsoventral pattern of egg shell and embryo in *Drosophila melanogaster*. *Cell* **49**: 699-707.

Scott, M.J., Kriticou, D. and Robinson, A.S. (1993) Isolation of cDNAs encoding 6-phosphogluconate dehydrogenase and glucose-6-phosphate dehydrogenase from the mediterranean fruit fly *Ceratitis capitata*: correlating genetic and physical maps of chromosome 5. *Insect Molecular Biology* **1**: 213-222.

Seki, T., Hagiya, M., Shimonishi, M., Nakamura, T. and Shimizu, S. (1991) Organisation of the human hepatocyte growth factor-encoding gene. *Gene* **102**: 213-219.

Stein, E., Cerretti, D.P. and Daniel, T.O. (1996) Ligand activation of ELK receptor tyrosine kinase promotes its association with Grb10 and Grb2 in vascular endothelial cells. *J Bio Chem* **271**: 23588-23593.

Stein, J.C., Howlett, B., Boyes, D.C., Nasrallah, M.E. and Nasrallah, J.B. (1991) Molecular cloning of a putative receptor protein kinase gene encoded at the self-compatibility locus of *Brassica oleracea*. *Proc Natl Acad Sci USA* **88**: 8816-8820.

Sugaya, R., Ishimaru, S., Hosoya, T., Saigo, K. and Emori, Y. (1994) A *Drosophila* homolog of human proto-oncogene *ret* transiently expressed in embryonic neuronal precursor cells including neuroblasts and CNS cells. *Mech Dev* **45**: 139-145.

Tartof, K.D. and Hobbs, C.A. (1987) Improved media for growing plasmid and cosmid clones. *Bethesda Res Lab Focus* **9**: 12 -16.

Tautz, D. and Pfeifle, C. (1989) A non-radioactive *in situ* hybridisation method for the localisation of specific RNAs in *Drosophila* embryos reveals a translational control of the segmentation gene *hunchback*. *Chromosoma* **98**: 81-85.

Tessier-Lavigne, M. (1995) Eph receptor tyrosine kinases, axon repulsion, and the development of topographic maps. *Cell* **82**: 345-348.

Van der Geer, P., Hunter, T. and Lindberg, R.A. (1994) Receptor protein-tyrosine kinases and their signal transduction pathways. *Annu Rev Cell Biol* **10**: 251-337.

Valenzuela, D,M., Stitt, T.N., DiStefano, P.S., Rojas, E., Mattsson, K., Compton, D.L., Nunez, L., Park, J.S., Stark, J.L., Gies, D.R., Thomas, S., Le Beau, M.M., Fernald, A.A., Copeland, N.G., Jenkins, N.A., Burden, S.J., Glass, D.J. and Yancopoulos, G.D. (1995) Receptor tyrosine kinase specific for the skeletal muscle lineage: expression in embryonic muscle, at the neuromuscular junction and after injury. *Neuron* **15**: 573-584.

von Heijne, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. *Eur J Biochem* **133**: 17-21.

von Heijne, G. (1985) Signal sequences. The limits of variation. *J Mol Biol* **184**: 99-105.

von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucl Acids Res* **14**: 4683-4690.

Walshe, J.A., Crosier, K.E., Evans, C.E. and Crosier, P.S. (1995) Understanding involvement of the Dtk/Axl subfamily of receptor tyrosine kinases in haematopoiesis using the zebrafish as a model vertebrate. The Annual Meeting of the American Society of Haematology, Seattle. *Blood* **86**: 683a.

Weigel, D., Jurgens, G., Kligler, M. and Jackle, H. (1990) Two gap genes mediate maternal terminal pattern information in *Drosophila*. *Science* **248**: 495-498.

Wilson, C., Goberdhan, D.C.I. and Steller, H. (1993) *Dror*, a potential neurotrophic receptor gene, encodes a *Drosophila* homolog of the vertebrate *Ror* family of Trk-related receptor tyrosine kinases. *Proc Natl Acad Sci USA* **90**: 7109-7113.

Wood, W.B. (1966) Host specificity of DNA produced by *Escherichia coli*: Bacterial mutations affecting the restriction and modification of DNA. *J Mol Biol* **16**: 118-133.

Yarden, Y. and Ullrich, A.R. (1988) Growth factor receptor tyrosine kinases. *Ann Rev Biochem* **57**: 443-478.

Zipursky, S.L. and Rubin, G.M. (1994) Determination of neuronal cell fate: lessons from the R7 neuron of *Drosophila*. *Ann Rev Neurosci* **17**: 373-397.