Characterisation of the *Musca domestica* (Diptera) homologue of the *Drosophila melanogaster* segment polarity gene *patched*, and the cloning of a cDNA fragment, PD, from *Acheta domesticus*.

Graeme P. Daniels

PhD

The University of Edinburgh

1998

The candidate confirms that the work submitted is his own, and that appropriate credit has been given where reference has been made to the work of others.

Abstract.

The *Drosophila* segment polarity gene, *patched* (*ptc*), encodes a transmembrane protein that is the receptor for the hedgehog intercellular signalling molecule. The hedgehog pathway plays a part in many developmental processes, ranging from *Drosophila* segmentation and appendage patterning, to vertebrate neural tube patterning and limb development. *ptc* null mutations in *Drosophila* are embryonic lethal and display a segment polarity phenotype. Mutations in the human *ptc* gene, however, have been shown to be responsible for Gorlin's syndrome and a predisposition to nevoid basal cell carcinomas.

In this study, partial clones of *patched* homologues were isolated from the long germ band developing house fly, *Musca domestica* (Diptera), and the intermediate germ band developing house cricket, *Acheta domesticus* (Orthoptera). Sequence data obtained from the *Musca domestica* clones showed a high degree of similarity with the *Drosophila ptc* gene sequence, indicating that the *Musca domestica* clones contain real *ptc* homologues. The *Acheta ptc* fragment was cloned using degenerate PCR, and sequence data has shown it has a high degree of similarity to the comparable regions of other *ptc* homologues. Two phagemid clones were also isolated from an *Acheta* cDNA library using a strategy designed to isolate *Acheta ptc*. One of these clones, PD, was used to create RNA *in situ* hybridization probes, and its expression was examined during *Acheta* embryogenesis, although it was later shown that PD was not an *Acheta ptc* homologue.

Expression of the *Musca ptc* homologue was examined during early development using RNA *in situ* hybridisation, and immunohistochemistry. These studies have shown that the expression of *ptc* during *Musca domestica* development is very similar to *ptc* expression patterns seen during the early development of *Drosophila*, suggesting that *ptc* may be fulfilling a similar role in both species.

Abbreviations.

Gene Names.

antp	*antennapaedia*
aptc	*Acheta domesticus patched*
arm	*armadillo*
bic	*bicoid*
brr	*barrel*
cad	*caudal*
ci	*cubitus interruptus*
cos-2	*costal-2*
Dax	*divergent antennapaedia classclass homeobox*
Dfz-2	*Drosophila frizzled-2*
dpp	*decapentaplegic*
dsh	*dishevelled*
en	*engrailed*
eve	*even-skipped*
ftz	*fushi tarazu*
fu	*fused*
gt	*giant*
hb	*hunchback*
hh	*hedgehog*
hkb	*huckebein*
inv	*invected*
kni	*knirps*
Kr	*Kruppel*
mdptc	*Musca domestica patched*
nkd	*naked*
nos	*nanos*
odd	*odd-paired*
opa	*odd-skipped*
oro	*oroshigane*
pan	*pangolin*
porc	*porcupine*
prd	*paired*
ptc	*patched*
run	*runt*
shh	*sonic hedgehog*
slp	*sloppy-paired*
smo	*smoothened*
tll	*tailless*
tra-1	*transformer-1*
Ubx	*Ultrabithorax*
wg	*wingless*

Proteins.

Arm	armadillo
Bic	bicoid
Ci	cubitus interruptus
En	engrailed
Eve	even-skipped
Hb	hunchback
Hh	hedgehog
Nos	nanos
Prd	paired
Ptc	patched
Wg	wingless

Miscellaneous.

ALF	automated laser fluorescent (DNA sequencer)
bp	base pair
cDNA	complementary deoxyribonucleic acid
DIC	differential interference contrast
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP's	deoxyribonucleoside triphosphates
EDTA	ethylenediaminetetra-acetate
ICRF	Imperial Cancer Research Fund
Lef-1	lymphocyte enhancer binding factor
MOPS	2-(N-morpholino)ethanesulphonic acid
mRNA	messenger ribonucleic acid
OD	optical density
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PIPES	piperazine-N,N'-bis[2-ethanesulphonic acid]
PKA	protein kinase A
RNA	ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
UPGMA	unweighted pair group method

Acknowledgements.

Table of Contents.

1. Introduction.

The adult insect, or imago, is a complex organism which is based around a basic body plan, or *bauplan*, that is common to most insect species. This body plan consists of a head, three thoracic segments, and between eight and eleven abdominal segments. In most insects, the head bears a pair of compound eyes, a pair of antennae and the mouthparts, and each of the thoracic segments have a pair of legs on their ventral sides. In addition, Pterygote (winged) insects have a pair of wings, or modified wings, on the dorsal side of the meso and meta-thoracic segments. This highly patterned imago develops from an embryo which originally consists of a number of relatively undifferentiated epithelial cells. The study of insect development, first using the techniques of fragmentation, ablation (reviewed in Sander, 1976), and grafting (Lawrence, 1966; Locke, 1967; Stumpf, 1966), and more recently with genetic and molecular biology techniques has given us many insights into some of the basic concepts of pattern formation, such as positional information (French *et al.*, 1976; Wolpert, 1969), compartmentalisation (Martinez-Arias and Lawrence, 1985; Garcia-Bellido *et al.*, 1973), organising centers (Patel *et al.*, 1989a; Sander, 1976), and morphogens (Sander, 1976).

Much of the early work was performed on species such as the cricket, dragonfly, cockroach, leafhoppers, and beetles (reviewed in Sander, 1976). None of these species

has proved to be of use as a genetic system, hence the vast majority of information available about the genetic and molecular basis of insect pattern formation has come from studies of the fruitfly, *Drosophila melanogaster*.

With the advent of molecular biology it is now possible to analyse in other insects the homologues of the genes that *Drosophila* utilises in these processes. These studies will shed light on the variety of mechanisms of pattern formation and the evolutionary conservation of the genes involved in the generation of the adult insect.

This thesis will detail the characterisation of the homologues of a *D. melanogaster* gene, *patched* (*ptc*), in the house fly, *Musca domestica*, and the house cricket, *Acheta domesticus*. This introduction will, therefore, contain information comparing the development of *Musca domestica* and *Drosophila melanogaster*, and also detail the development of the short and intermediate germ band insects, of which *Acheta domesticus* is an example.

1.1. Insect early development.

Insect development begins during oogenesis, prior to fertilisation and the formation of an embryo. Maternal information that will direct early development is placed into the

developing oocytes during oogenesis. It is therefore necessary to understand the process

of oogenesis before moving onto embryonic development.

1.2. Insect oogenesis.

Insect ovaries can be divided into two major groups, panoistic or meroistic (Brandt,

1874), on the basis of their morphology, and whether there are nurse cells associated

with the developing oocytes. Nurse cells supply maternally derived macromolecules,

e.g., mRNA, to the oocyte prior to fertilisation, and are peculiar to meroistic oogenesis.

Even though the two types of ovary are structurally different, there are some

characteristics that are common to both. Insects have a pair of ovaries consisting of a

number of parallel ovarioles, each of which is anchored to the thoracic wall by the

terminal filament found at the anterior tip of the ovariole. The anterior of the ovary is

the called the germarium which contains the stem line oogonia (which produce the

oocytes). As the developing oocytes move posteriorly through the ovariole, they pass

from the germarium into the vitellarium (see Figure 1), where they increase in size and

finally mature, before being released into the oviducts.

1.2.1. Panoistic oogenesis.

Panoistic oogenesis is characterised by the absence of nurse cells associated with the developing oocyte (see Figure 1). This means that the oocyte nucleus is the main contributor of information to the mature egg (for review see Mahowald, 1972). The stem line oogonia are found in the anterior germarium, along with young oocytes in the early stages of meiosis. As the developing oocytes move into the posterior germarium they are in the later stages of meiosis, but are not yet completely surrounded by follicle cells. The oocytes increase their cytoplasmic volume and become entirely surrounded by follicle cells as they pass through to the vitellarium. Once in the vitellarium, the oocyte fills with yolk, and the oocyte nucleus moves to its characteristic position, which is typically mid-dorsal, as in *A. domesticus*. After vitellogenesis is complete, the follicle cells secrete the vitelline membrane and chorion after which the mature egg is ready to shed the follicle cells and pass into the oviduct.

1.2.2. Meroistic oogenesis.

There are two types of meroistic ovary; the polytrophic, and telotrophic, which are differentiated by the position and movement of the nurse cells throughout oogenesis. The nurse cells of the polytrophic meroistic ovary move out of the germarium with the

oocyte and travel along the length of the ovariole, whereas the nurse cells of the telotrophic meroistic ovary never leave the germarium and are associated with the developing oocytes via nutritive chords.

1.2.2.1. Polytrophic meroistic oogenesis.

The germarium of the polytrophic meroistic ovary can be subdivided into three regions, the anterior which consists of the stem line oogonia and the cystoblasts, the mid region where the daughter cells of the cystoblasts, or cystocytes, begin to enlarge, and the cell clusters start to become surrounded by pre-follicular cells, and the posterior where the oocyte differentiates and is positioned posteriorly to the nurse cells, and the cluster is fully surrounded by follicle cells. All of the cells in a developing egg chamber, excluding the follicular cells, are clonal, as they all descend from a single germarial cystoblast. The cystocytes undergo incomplete cytokinesis, and are all interlinked with one another via a system of ring canals, or fusomes. At the end of four mitotic divisions there are eight cystocytes with a single fusome connection, four with two connections, two with three connections, and two with four connections. The two cystocytes with four fusome connections are called pro-oocytes. One of these will develop into the oocyte and the other enters endomitosis to become a nurse cell. As the cluster proceeds through the germarium the nurse cells start to become highly polyploid (Mahowald, 1972), and once in the vitellarium, vitellogenesis occurs and yolk is deposited in the

oocyte. The nurse cells break down, and the follicle cells secrete both the vitelline membrane and the chorion around the oocyte, resulting in the formation of the mature egg. The follicle cells which surround the mature egg degenerate and are sloughed off as the egg leaves the ovariole and goes on to be fertilised in the oviduct or fertilisation chamber (reviewed by Sander, 1985).

1.2.2.2. Telotrophic meroistic oogenesis.

Telotrophic meroistic ovaries differ from polytrophic meroistic ovaries in the morphology of their germariums, and that the nurse cells never leave the germarium. Telotrophic stem line oogonia tend to cease dividing between the late larval and early adult stages, depending on the species, which results in there being a finite number of oocytes that each ovariole can produce. The descendants of the stem line oogonia are all surrounded by a sheath of somatic cells, which at more posterior levels, merges into the prefollicular cells. The germ cell descendants are divided into an anterior group which are determined to become nurse cells and a posterior group that are determined to become oocytes. The oocytes are connected to the nurse cells or tropharium by nutritive chords through which the maternally derived components are transferred. As the oocytes move posteriorly through the prefollicular tissue, they begin to enlarge by the accumulation of nurse cell derived material within them. Vitellogenesis is said to begin only once yolk particles start to accumulate within the oocyte. After vitellogenesis is

completed the follicle cells secrete the vitelline membrane and the chorion, before being

sloughed off prior to fertilisation.
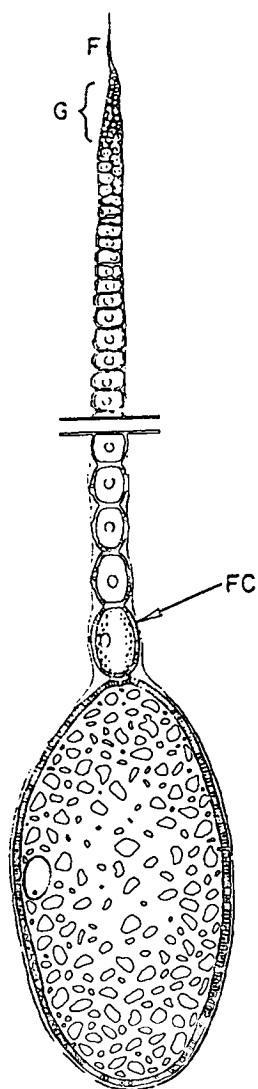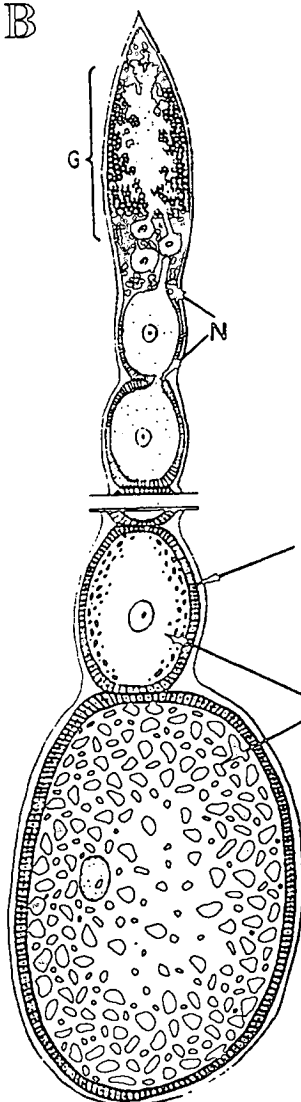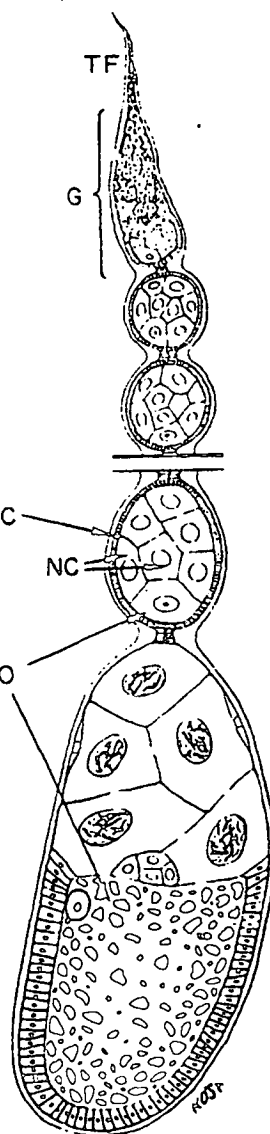
Figure 1. Diagrammatic representations of the three insect ovary types.

(A) The panoistic ovary. (B) The telotrophic meroistic ovary. (C) The polytrophic meroistic ovary. Labels: F, filament; G, germarium; FC, follicle cell; N, nutritive cord; O, oocyte; TF, terminal filament; NC, nurse cell. Figure adapted from Mahowald (1972).

1.3. Insect embryogenesis.

In the very early insect embryo, immediately following fertilisation, a number of nuclear divisions occur in the absence of cytokinesis which produce a syncitial blastoderm from which the embryonic primordia, or 'germ band' or 'germ anlage' cellularises. This general process of creating a syncitium from which the germ band cellularises is common to most insects. There are three basic modes of insect embryogenesis called long, short and intermediate germ band, which are distinguished by the amount of the body pattern that is specified in the early germ band (Krause, 1939; Sander, 1976).

1.3.1. Long germ band embryogenesis.

Since embryogenesis of the fruitfly, *Drosophila melanogaster*, has been extensively studied it will be used as the example for most of the description given here.

After fertilisation, the *Drosophila* embryo undergoes 14 rounds of nuclear divisions, in the absence of cytokinesis, to form a syncitium. The pattern of these nuclear divisions differ, even within the Diptera. In *Musca domestica*, the first twelve of the divisions are synchronous, but the thirteenth and fourteenth are not, as after the twelfth division the nuclei in the posterior lag behind those in the anterior (Sommer and Tautz, 1991a), and

the pattern of nuclear division in *Calliphora erythrocephala* is even more complex than in *Musca* (Lundquist, 1981). After the first seven zygotic nuclear divisions in *Drosophila melanogaster*, most of the syncitial nuclei migrate to the surface to form the syncitial blastoderm. Another seven nuclear divisions occur before the nuclei of the syncitial blastoderm begin to cellularise, forming the cellular blastoderm, although cellularisation is not completed until the onset of gastrulation. Those nuclei that do not migrate to the surface either remain within the yolk and become polyploid after three more divisions, or become incorporated into the posterior pole plasm. Once the cellular blastoderm has been formed (stage 5; Campos-Ortega and Hartenstein, 1985), the primordia of all of the body segments of the larval fly are present and gastrulation begins (see Figure 2).

During the first stage of gastrulation (stage 6), the cells fated to become mesoderm invaginate into the ventral furrow, which extends along most of the length of the embryo, and then spread laterally to generate a two layered structure. Germ band extension is initiated at this stage of development (stage 7), with cells at the posterior of the embryo moving in an antero-dorsal direction. Unlike the short and intermediate germ band species (see section 1.3.2.), the germ band of Dipteran (long germ band) species initially occupies the whole length of the egg, and it extends around the posterior tip of the egg, along the dorsal side (see Figure 2). Cells at both the posterior

and anterior of the ventral furrow invaginate as the endodermal primordia of the posterior and anterior midguts, respectively.

It is at stage 9, i.e., the extending germ band, of *Drosophila* embryogenesis that the first morphological signs of segmentation become visible as small thickenings in the mesoderm. Grooves appear in the ectoderm shortly after this (Turner and Mahowald, 1977), although they are slightly out of register with the periodic demarcations in the mesoderm. These periodic grooves demarcate, not the future segments, but parasegments consisting of the posterior compartment of one of the future segments and the anterior compartment of the next. The parasegmental boundaries are established by the expression of the pair rule genes *even-skipped* and *fushi tarazu*, and are essential for the correct segmental patterning of the embryo (see section 1.5.1.3.).

After stage 11, the germ band retracts back around the posterior tip of the egg (stage 12). It is during this stage that the parasegmental divisions of the ectoderm and mesoderm come into register with one another, and the segmental grooves that are evident in the later stages of the life cycle appear. Dorsal closure of the embryo and head involution characterise stage 14. After completion of head involution and gut growth, the fully developed embryo (stage 17) hatches into the first of three larval instars. Long germ band embryogenesis is remarkably fast, exemplified in *Drosophila melanogaster* where embryogenesis takes approximately 24 hours at 25° C.
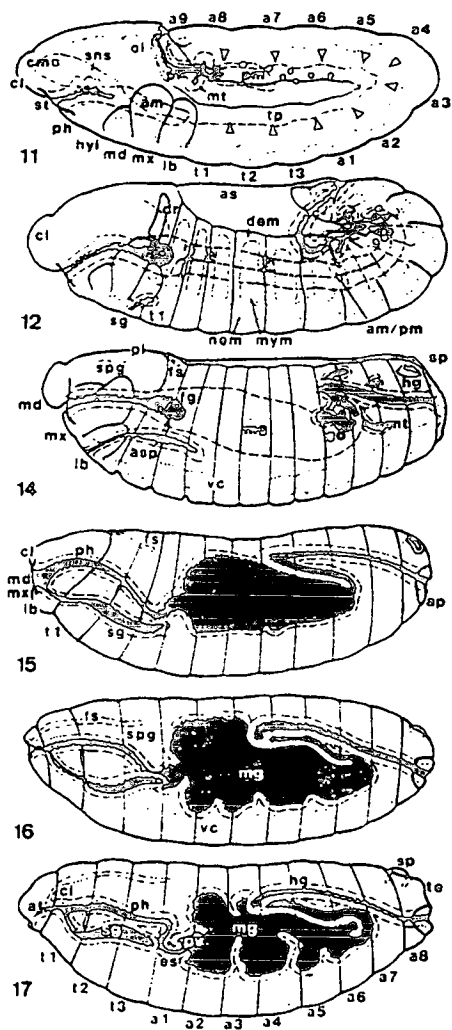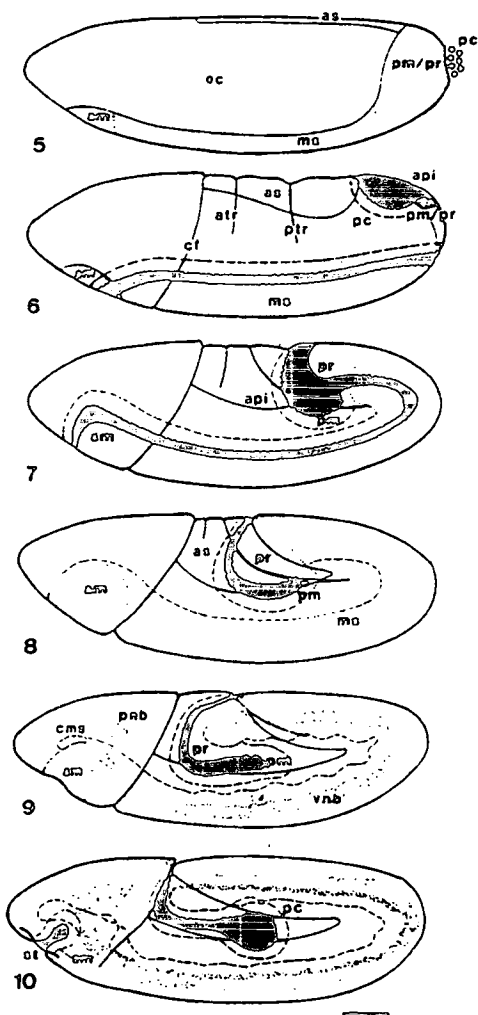
Figure 2. The stages of *Drosophila melanogaster* embryogenesis (taken from Campos-Ortega and Hartenstein, 1985).

Stage 5: The blastoderm. Stage 6: Gastrulation. Stages 7-11: Germ band elongation. Stage 11: Epidermal segmentation becomes evident. Stage 12: Germ band retraction. Stage 14: Dorsal closure. Stages 14-17: Head involution. Stage 17: The completed embryo. All the embryos are orientated with anterior to the left, and dorsal to the top. The scale bar represents 50 μm. Abbreviations: *a1-a10* - abdominal segments; *am* - anterior midgut; *ap* - anal plate; *api* - amnioproctodeal invagination; *as* - amnioserosa; *asp* -anterior spiracles; *at* - atrium; *atr* - anterior transverse furrow; *cf* - cephalic furrow; *cl* - clypeolabrum; *cms* - cephalic mesoderm; *dem* - dermomeres; *dr* - dorsal ridge; *ec* - ectoderm; *es* - oesophagus; *fg* - foregut; *fs* - frontal sack; *hg* - hindgut; hyl - hypopharyngeal lobe; *lb* - labial bud; *md* - mandibular bud; *mg* - midgut; *ms* - mesoderm; *mt* - malpighian tubules; *mx* - maxillary bud; *mym* - myomeres; *nem* - neuromeres; *ol* - optic lobe; *pc* - pole cells; *ph* - pharynx; *pl* - procephalic lobe; *pm* - posterior midgut; *pnb* - procephalic neuroblasts; *pr* - proctodeum; *ptr* - posterior transverse furrow; *pv* - proventriculus; *sg* - salivary glands; *sns* - stomatogastric nervous system; *sp* - posterior spiracles; *spg* - supraoesophageal ganglion; *st* - stomodeum; *t1-t3* - thoracic segments; *te* - telson; *tp* - tracheal pits; *vc* - ventral cord; *vnb* - ventral neuroblasts (All abbreviations are in alphabetical, not stage, order).

Drosophila melanogaster is a holometabolous insect, which means that it undergoes a

complete metamorphosis between the larval and adult stages of its life cycle. The larvae

contain both cells which make up the larval body, and cells that were 'set aside' during

embryogenesis as 'adult' cells, in the form of imaginal discs and histoblast nests (Bate

and Martinez-Arias, 1991; Cohen *et al.*, 1993; Fristrom and Fristrom, 1993; Nothinger,

1972).

The abdominal segments, A1-A8, are derived from the imaginal histoblast cells. There

are four histoblast nests per segment, each consisting of between six and fifteen cells.

These cells contribute to the larval epidermis, and proliferate during the pupal stages to

form the adult abdominal segments. The imaginal discs form the rest of the adult

epidermis. These are the labial, clypeo-labral, antennal, and eye discs which form the

adult head epidermis, the humeral, wing, and haltere discs which form the dorsal thorax,

the leg discs which form the ventral thorax, and the genital disc which forms the last

abdominal 'segment', A9. Imaginal disc primordia have been shown to be determined

early in embryogenesis (Lohs-Schardin *et al.*, 1979; Weischaus and Gehring, 1976a;

Weischaus and Gehring, 1976b), with the thoracic imaginal discs being recognisable as

distinct clusters of subepidermal cells midway through embryogenesis. Many of the

investigations into imaginal disc development have concentrated on the thoracic discs,

e.g. the leg, and wing discs (Bate and Martinez-Arias, 1991; Cohen *et al.*, 1993). Both

the dorsal (wing and haltere) and ventral (leg) thoracic discs have been shown to arise

from a single cluster of ventral cells spanning the parasegmental border (Cohen *et al.*,
1993), which divides them into an anterior and posterior compartment (Crick and
Lawrence, 1975). The two sets of discs then separate at around stage 13, and the dorsal
discs migrate away. During larval development the imaginal discs are invaginated
pouches of epithelial cells which evaginate to create the corresponding adult structures
during metamorphosis (see Cohen, 1993 for review).

1.3.2. Short and intermediate germ band embryogenesis.

The short and intermediate germ band modes of embryogenesis are morphologically
very similar (Sander, 1976). The oocyte is fertilised as it passes through the oviduct, the
maternal and paternal pronuclei fuse and synchronous nuclear divisions begin within the
yolk of the egg. These nuclei are surrounded by a small amount of cytoplasm and are
termed energids. After a number of divisions, e.g., six in *Acheta domesticus* (Sauer,
1966), the energids start to migrate to the periphery of the egg and first appear at the
posterior pole of the egg (reviewed in (Anderson, 1972). There are a number of
synchronous nuclear divisions that take place after energid emergence, e.g., in *Acheta*
domesticus there are three divisions after the emergence of the energids, which produce
a total of around 512 nuclei (Sauer, 1966, reviewed in Anderson, 1972). Some of the
energids then migrate to the posterior pole and cellularise, forming the germ band,
whilst the rest of the nuclei remain widely dispersed to eventually contribute towards

the extra-embryonic membranes, the amnion and serosa. The germ band, which initially

consists of the head (in short germ band insects) and thorax (in intermediate germ band

insects), then begins to elongate and sequentially form new segments by the

proliferation of the posterior blastema, or 'growth zone' and is now termed the extending

germ band. The germ band extends around the posterior tip of the egg, becoming

submerged in the yolk. This morphogenetic movement is known as anatrepsis and

results in the reversal of the antero-posterior polarity of the embryo with respect to that

of the egg (see Figure 3). During anatrepsis, segmentation of the germ band is

completed, and gastrulation begins (see Figure 4). Gastrulation typically begins with

the formation of the gastral groove along the ventral midline through which cells

invaginate to become the mesoderm. It is also during anatrepsis that the appendage

buds become evident in most short and intermediate germ band insects. After anatrepsis

there can be a developmental 'pause' before the embryos commence on another set of

morphogenetic movements known as katatrepsis. During katatrepsis the embryo moves

back around the posterior tip, restoring the original antero-posterior polarity of the

embryo with respect to that of the egg (see Figure 3). After katatrepsis, the embryo

continues to grow and the appendages continue to develop, dorsal closure occurs and

the embryo lays down the larval cuticle and pigments. Most short and intermediate

germ band insects are hemimetabolous, which means that the larvae resemble the final

adult form and do not undergo complete metamorphosis. Short germ embryogenesis is

much slower than long germ band embryogenesis, for example, *Acheta domesticus*

embryogenesis takes between two and three weeks, *Locusta migratoria* embryogenesis

takes approximately two weeks, and *Xiphidium ensiferum* embryogenesis takes five months.

A
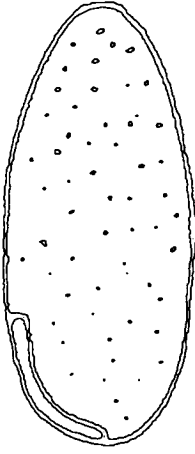
B
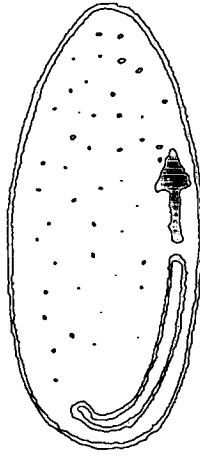
C
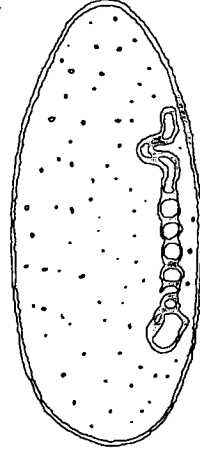
D
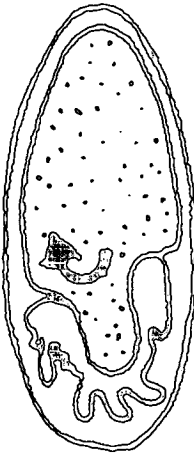
E

F
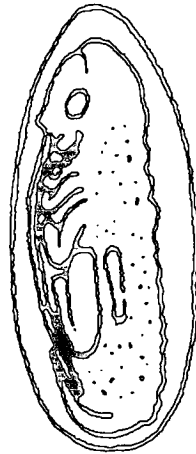
Figure 3. Morphogenetic movements during short germ embryogenesis.

(A) The embryonic primordia cellularises in the ventral posterior region of the egg. (B) Anatrepsis. As the embryo begins to elongate it moves around the posterior tip of the egg and becomes submerged in the yolk. The anterior-posterior orientation of the embryo becomes reversed with respect to that of the egg. (C) The embryo elongates and segments. (D) Katatrepsis. The embryo moves back around the posterior tip of the egg, regaining its original anterior-posterior orientation. (E) and (F) The embryo continues to develop, eventually occupying the entire egg. All drawings are orientated with the anterior at the top and ventral to the left. Figure adapted from Lauga (1969).
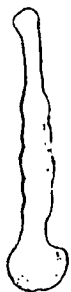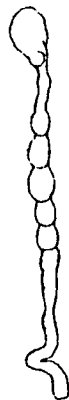
A   B   C   D E   F

G   H   I   J   K   L   M

Figure 4. The embryonic development of the cricket, *Acheta domesticus.*

(A) Stage 10. The early extending germ band. The first visible signs of segmentation become visible. (B) Stage 11. The extending germ band. The stomatodeum and antennal buds become evident. (C) Stage 12. The extending germ band. The antennal and limb buds become more pronounced. The embryo is now just over 1mm long. (D) Stage 13. The extending germ band. The embryo has finished anatrepsis, and the developing antennal and limb buds are very obvious. There is no visible segmentation in the abdomen of the embryo. The posterior tip of the embryo 'kinks'. (E) A lateral view of the stage 13 embryo. (F) Stage 14. The extending germ band. The posterior of the embryo begins to grow back towards the anterior of the embryo. Segmentation in the abdomen of the embryo becomes evident. (G) Stage 15. The extending germ band. (H) Stage 16. The fully extended germ band. The developing appendages are well advanced, and the cerci have become pronounced. (I) Stage 17. The fully extended germ band. Just prior to katatrepsis. (J) Stage 20. Mid katatrepsis. (K) Stage 23. Post katatrepsis. The appendages continue to grow. The pleuropodia appear laterally. (L) Stage 31. The embryo prior to hatching. (M) A dorsal view of the stage 31 embryo. All embryos are viewed ventrally unless indicated, and orientated with anterior to the top. Figure adapted from Lauga (1969).

1.4. Phylogenetic Correlations between Oogenesis and Embryogenesis.

The phylogeny of the Insecta is a matter of contention in some quarters, although there is a generally accepted scheme containing over 1 million extant species (Schwalm, 1988). Within this phylogeny there is a very particular distribution of ovary types, and types of embryogenesis (see Figure 5).

Examination of the distribution of ovary types and modes of embryogenesis has shown that insects that are thought to belong to the more ancient orders such as the Orthoptera, and Odonata have panoistic ovaries and exhibit either short or intermediate germ band embryogenesis, whereas in the more derived orders, such as the Diptera, the species are long germ band developing, and possess meroistic ovaries (King and Büning, 1985).

From the distribution and comparative morphology of ovaries (King and Büning, 1985), and types of embryogenesis (Tear et al, 1988) within the insects, it has been proposed that panoistic ovaries are ancestral to meroistic ovaries, and a mode of embryogenesis similar to short or intermediate germ band, where the anterior domain is specified and the posterior segments are added in a progressive manner, is ancestral in the all of the arthropods.

Investigation of orders such as the Coleoptera, Hemiptera, and Lepidoptera, which are thought to be phylogenetic intermediates between the ancient and derived orders, has

shown that they contain both species with panoistic, and species with meroistic ovaries, undergoing either short (or intermediate) germ band, or long germ band development, respectively (Sander, 1976). It has been noted, however, that there is no known example of an insect possessing panoistic ovaries and undergoing long germ band development, although there are species with meroistic ovaries that are intermediate germ band developing insects (Patel, 1994), suggesting that the possession of nurse cells is a prerequisite for long germ band embryogenesis, but it is not prohibitive of short or intermediate germ band development.

	Ovary Type	Germ Type

Hemimetabolous

Orthoptera
(grasshopper) — panoistic — short - intermediate

Holometabolous

Insecta

Coleoptera
(red flour beetle) — meroistic — short - long

Hymenoptera
(honey bee) — meroistic — short - long

Lepidoptera
(silk worm moth) — meroistic — long

Diptera
(fruit fly) — meroistic — long
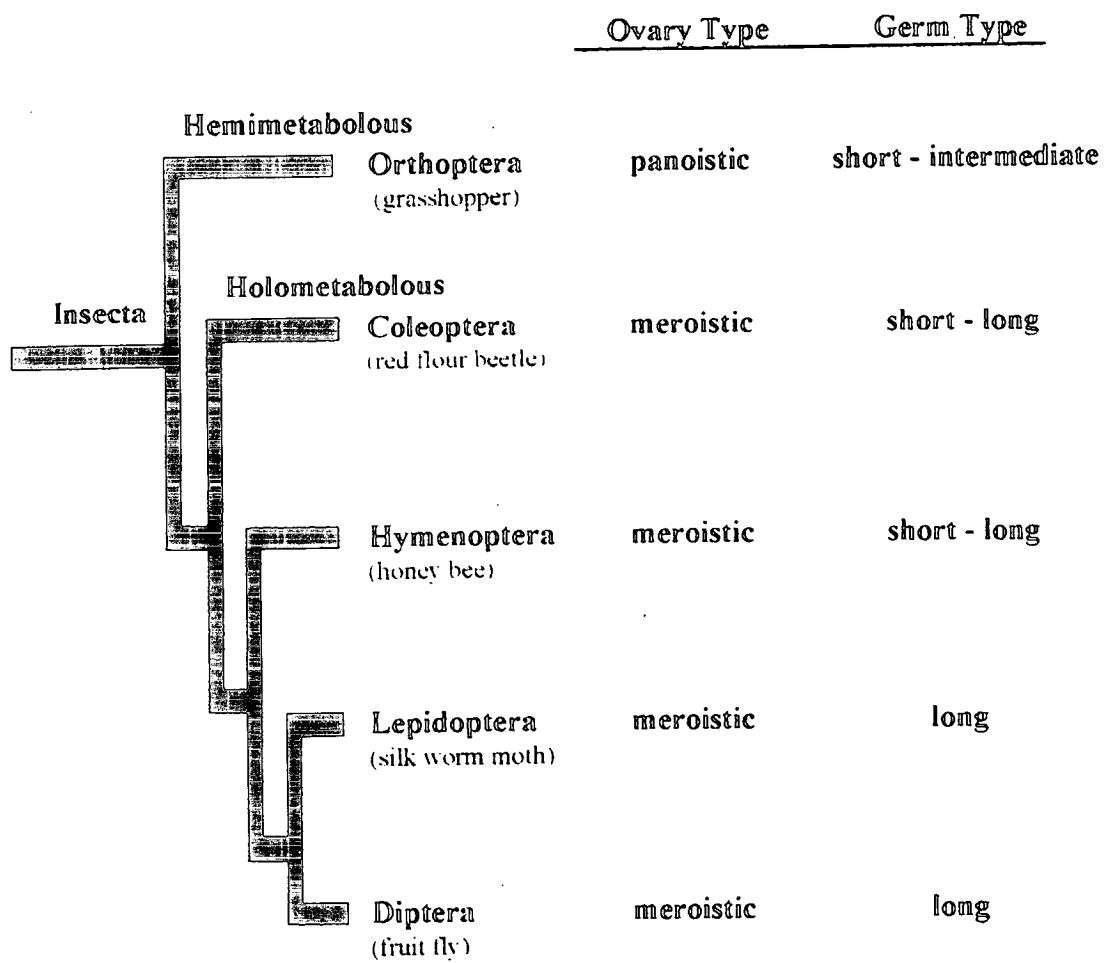
Figure 5. The phylogenetic relationship between selected insect orders.

The germ band types that are found within an order, and the ovary type possessed by a selected species within an order are shown. 'short - intermediate' indicates that both short and intermediate germ types are found within an order, 'short - long' indicates that all three germ types are found within the order. The phylogenetic relationships are based on Schwalm (1988). Figure adapted from Brown *et al* (1994).

1.5. The molecular basis of insect segmentation.

As has been described in sections 1.3.1. and 1.3.2., there a number of forms of insect

embryogenesis, as defined by morphological criteria, but it is unknown how similar

these processes are at the molecular level. Now that there is a large amount of

information available on the molecular basis of segmentation and axial patterning of the

Drosophila embryo, the question of how conserved these mechanisms are at the

molecular level can start to be answered using the techniques of homologous gene

cloning, immunohistochemistry, and *in situ* hybridisation. In the following sections I

will detail some of what is known about the molecular basis of segmentation in

Drosophila melanogaster, and ultimately describe the role that a segment polarity gene,

patched, plays in the processes of segmentation and limb development.

1.5.1. The molecular basis of segmentation in *Drosophila melanogaster*.

At the segmented germ band stage of *Drosophila melanogaster* embryogenesis, the

embryo consists of two non-segmental termini, the acron (at the anterior) and telson (at

the posterior), and fourteen trunk segments. The vast majority of molecular data

available on insect embryogenesis comes from work performed on *Drosophila*

melanogaster over the past 16 years. In the early 1980's, large scale genetic screens

were performed which identified many of the genes that are involved in the segmentation of the fly embryo (Nusslein-Volhard and Weischaus, 1980; Nusslein-Volhard *et al.*, 1984). These genes were placed into different groups on the basis of the disruption of the segmental pattern seen in the mutant larval cuticle. The 'gap gene' mutants were those in which a continuous stretch of segments were missing from the final pattern; the 'pair-rule gene' mutants were defined by deletions of a segmental amount of tissue in alternating segments; and the 'segment polarity gene' mutations were characterised by a fraction of the pattern of each segment being deleted and replaced by the mirror image of the remaining tissue.

In *Drosophila melanogaster*, segmentation is controlled by four groups of genes, which are known collectively as the 'segmentation genes'. During normal embryogenesis the segmentation genes are expressed in a strict spatial and temporal order, each group regulating its own expression and that of the proceeding group of genes in the hierarchy.

Temporally, the first group of segmentation genes to exert their effects are the co-ordinate genes. These are expressed solely from the maternal genome and their products are the first determinants to provide polarity to the oocyte and to divide the embryo into distinct regions, by defining limits of expression of the next group, the gap genes. These gap genes act to divide the embryo into large overlapping regions. The gap genes interact with one another and the co-ordinate genes to regulate both their own

expression, and that of the pair-rule genes, which are expressed in a periodic pattern of alternate segments. Finally, the pair rule genes regulate the expression of the segment polarity genes which are expressed in discrete regions within every segment of the embryo and organise the intra-segmental pattern (see Figure 6). In the following sections I will discuss what is known about the generation of the segmental pattern along the anteroposterior axis of the *Drosophila* embryo.

ANTERIOR   1 3 5 7 9 11 13 14   POSTERIOR

70%   10%
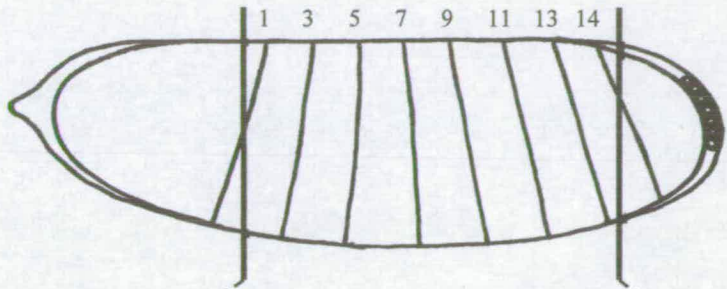
EGG LENGTH

SEGMENTS   MN   MX   LA   T1   T2   T3   A1   A2   A3   A4   A5   A6   A7   A8   A9

COMPARTMENTS P A P A P A P A P A P A P A P A P A P A P A P A P A P A P A

PARASEGMENTS   1   2   3   4   5   6   7   8   9   10   11   12   13   14

GAP GENE
EXPRESSION

Kruppel

hunchback   *knirps*

PAIR-RULE GENE
EXPRESSION

hairy

even-skipped

paired

fushi tarazu

SEGMENT POLARITY
GENE EXPRESSION

engrailed

wingless

patched
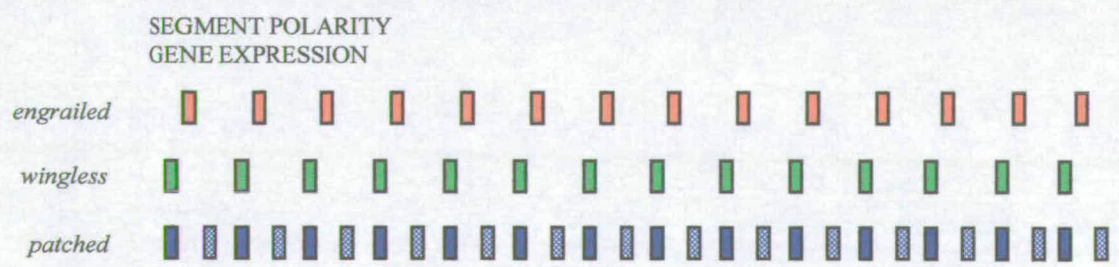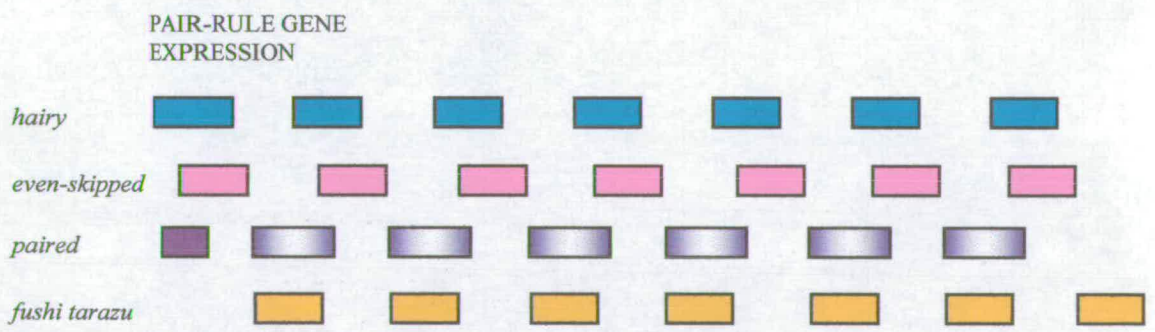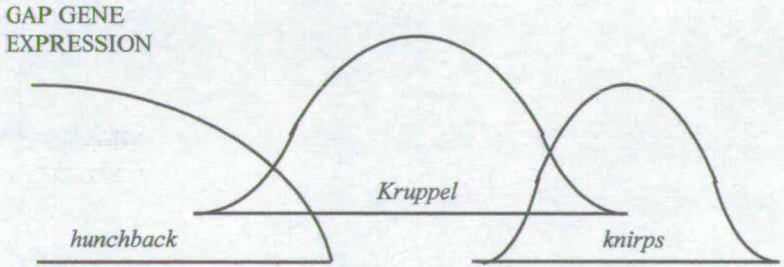
Figure 6. Segmentation gene expression in the *Drosophila* embryo.

From the top of the page:

A *Drosophila* embryo showing some of the parasegmental borders with the percentage egg length corresponding to the segmented trunk being shown underneath. Egg length (EL) – a measurement used to define anteroposterior position within the egg; anterior pole is 100%, posterior pole is 0%.

The spatial correlation of segments, segmental compartments, and parasegments. Abbreviations; MN, mandible; MX, maxillary; LA, labium; T1-3, thoracic segments; A1-9, abdominal segments; P, posterior segmental compartment; A, anterior segmental compartment.

Gap gene expression. The bell curves indicate distribution of protein within specific expression domains of *hunchback*, *Krüppel*, and *knirps*.

Pair-rule gene expression. The coloured boxes indicate expression domains of the pair-rule genes (green = *hairy*, purple = *even-skipped*, blue = *paired* (the gradation represents the maturation of the *prd* pattern into segmental stripes by loss of expression), orange = *fushi tarazu*).

Segment polarity gene expression. The coloured boxes represent expression domains of the segment polarity genes (*red* = engrailed, *green* = wingless).

Anterior is to the left. Spatial relationships between segments and expression domains are maintained vertically down the page. Figure adapted from Figures B1.1 and 3.4. The Making of a Fly (Lawrence, 1992), and Figure 4. (Akam 1987).

1.5.1.1. The Co-ordinate genes.

The co-ordinate genes act to specify the terminal regions of the embryo, the dorsoventral

axis, as well as the anteroposterior axis (for review see Nusslein-Volhard, 1991).

The co-ordinate genes *bicoid* (*bcd*), and *nanos* are the primary determinants of the

anteroposterior axis. The anteroposterior axis is determined by two separate systems,

the anterior, in which *bicoid* acts, and the posterior, in which *nanos* functions. Mutants

in the genes of both the anterior and posterior systems result in deletions of large

regions of the embryo, e.g., amorphic alleles of *bicoid* delete the head, gnathal, and

thoracic regions and the telson is duplicated in the 'anterior', whereas the abdomen is

missing in null mutants of *nanos*.

The *bicoid* gene encodes a homeobox protein, which is one of the few molecules that

has been shown to act as a diffusible morphogen, i.e., it confers positional information

by the creation of a concentration gradient that at particular levels, or between specific

thresholds, will direct specific cellular responses. The creation of this Bicoid protein

gradient is possible due to the syncitial nature of the early *Drosophila* embryo (Berleth,

1988). In brief, the *bicoid* transcript passes in from the nurse cells and is localised at the

anterior pole of the oocyte. These transcripts are translated shortly after the egg has

been laid, and Bicoid protein diffuses through the embryo creating an anteroposterior

diffusion gradient (Driever and Nusslein-Volhard, 1988a). This gradient of Bicoid

protein is then interpreted by the gap genes, e.g. expression of the gap gene *hunchback*

(*hb*) is activated in the anterior of the embryo above a threshold concentration of Bicoid

protein.

The posterior system is slightly different from the anterior system in that it removes a

repressor, the maternal transcript of the gap gene, *hunchback*, which then allows the

expression of the gap gene, *knirps* (*kni*), (Hulskamp *et al.*, 1989) (see section 1.5.1.2.).

The products of two genes, *nanos* (*nos*), and *pumilio* (*pum*), function together to repress

the translation of *hb* mRNA in the posterior of the embryo (Rivera-Pomar and Jackle,

1996).

1.5.1.2. The gap genes.

Gap gene expression is regulated by the co-ordinate genes of the anteroposterior system,

namely *bicoid* and *nanos*, and by, sometimes quite complex, cross-regulation by the gap

genes themselves (Jackle *et al.*, 1986). The gap genes function to refine the

anteroposterior positional information generated by *bicoid* and *nanos*, and to regulate

the expression of the pair-rule genes. All of the gap genes encode transcription factors,

proteins with DNA binding motifs such as zinc fingers, e.g., *hunchback* (Tautz *et al.*,

1987a), *krüppel (Kr)* (Rosenberg *et al.*, 1986), b-ZIP motifs, e.g., *giant (gt)* (Capovilla *et al.*, 1992), or steroid receptor like motifs, e.g., *knirps* (Nauber *et al.*, 1988). Because the gap genes are active while the embryo is still a syncitium, their gene products are able to gain direct access to neighbouring nuclei, thereby directly regulating the expression of their target genes without the need for complex signal transduction pathways that would be necessary in a cellular environment.

hunchback is unusual among the gap genes because there is a maternal, as well as a zygotic, component to its expression pattern (Rivera-Pomar and Jackle, 1996; Tautz, 1988a). Initially, within the egg, maternal expression of *hb* is uniform, but the translation of the maternal *hb* mRNA is antagonised by the action of the Nos and Pumilio proteins. Analysis of gap gene mutants has shown that the posterior limit of *hb* expression is also affected by *Kr* expression (Jackle *et al.*, 1986). This control has been demonstrated to be direct, with Kr protein binding to the upstream regulatory sequences of the *hb* gene, which represses transcription of *hb* (Treisman and Desplan, 1989).

In the anterior of the embryo, zygotic *hb* expression is activated by Bicoid protein in a concentration manner (Driever and Nusslein-Volhard, 1988b; Driever and Nusslein-Volhard, 1989; Driever and Nusslein-Volhard, 1988a; Struhl *et al.*, 1989), and is controlled by a synergistic interaction between the Bicoid and maternal Hunchback proteins (Simpson-Brose *et al.*, 1994). The combination of the anterior and posterior

mechanisms results in the expression domains of both maternal and zygotic *hb* being

similarly restricted to the anterior of the embryo, although the level of maternal

transcript is far lower than that of the zygotic transcript. Later, zygotic *hb* is also

expressed in a posterior domain which extends from 25-10% EL (Lehmann and

Nusslein-Volhard, 1987; Tautz *et al.*, 1987a).

Krüppel is first expressed in a band which extends from 60-50% EL, a region that

corresponds to the presumptive thorax and anterior abdomen, and which slightly

overlaps the posterior edge of the *hb* domain, with later expression in the posterior of

the embryo and anterior to the cephalic furrow (Knipple *et al.*, 1985). It has been shown

that *Kr* requires both *bcd* and *hb* for its activation, as in single mutants for both *bcd* and

both the maternal and zygotic components of *hb*, *Kr* is still expressed in a domain

similar to that seen in the wild type situation, but in a *bcd⁻hb⁻* double mutant *Kr*

expression is missing (Hulskamp *et al.*, 1990). An interesting twist to the control of *Kr*

regulation by *hb* is that at low level, such as that seen when only the maternal

component is present, *hb* is seems to activate *Kr* transcription, but at higher levels it has

a repressive affect (Gaul and Jackle, 1987; Jackle *et al.*, 1986). Furthermore *Kr* has

been shown to have Bicoid protein binding sites within its promoter region and that the

activation of *Kr* transcription is directly controlled by low concentrations of the Bicoid

protein (Hoch *et al.*, 1992; Hoch *et al.*, 1990; Hoch *et al.*, 1991). It is now thought that

Kr expression is activated by Hunchback and Bicoid, and is spatially restricted by the

action of the gap genes in adjacent expression domains (Rivera-Pomar and Jackle, 1996), for example, the posterior limit of *Kr* expression is defined by the repressive action of *knirps*, as shown by the posterior expansion of the *Kr* domain in *kni⁻* embryos (Hoch *et al.*, 1992).

The third of the genes, *knirps*, is expressed in band in the posterior of the embryo and is required for segmentation of the abdomen (Nauber *et al.*, 1988). The 900bp regulatory region of *knirps* has been cloned and characterised by deletion assay (Pankratz *et al.*, 1992; Pankratz *et al.*, 1989). It has been shown that *kni* expression is activated by the *caudal* gene product (Rivera-Pomar and Jackle, 1996). From the analysis of mutants for *hb*, *giant*, and another gap gene, *tailless*, which is expressed at both of the termini of the embryo (Pignoni *et al.*, 1990), and the identification of binding sites for both the *hb* and *tll* proteins within the *kni* regulatory region, it is evident that these three genes play a role in defining the expression domain of *kni* by its repression elsewhere in the embryo.

The gap genes represent the first level of zygotic interpretation of positional information laid down in the oocyte by the mother. The embryo is divided into broad domains by gap gene expression, which is then translated into a periodic segmental pattern. This is a two stage process, first, the gap gene products regulate the periodic expression, in alternate segmental primordia, of the pair rule genes, and second, the pair rule gene products regulate the segmental expression of the segment polarity genes.

1.5.1.3. The pair-rule genes.

The pair-rule genes *paired* (*prd*), *even-skipped*, *odd-skipped* (*odd*), *barrel* (*brr*), and

runt (*run*), were identified as part of the Nüsslein-Volhard and Weischaus screen for

genes involved in segmental patterning (Nusslein-Volhard and Weischaus, 1980). The

larval cuticles of pair-rule mutants show a remarkable phenotype in which alternate

segments are deleted, e.g., *eve* mutants in which the even numbered segments are

deleted (Nusslein-Volhard and Weischaus, 1980) and the parasegmental boundaries fail

to form (Martinez-Arias *et al.*, 1988; Martinez-Arias and White, 1988), or *prd* where the

posterior of the odd segments and the anterior of the even segments are missing

(Nusslein-Volhard and Weischaus, 1980).

Like the gap genes, the pair-rule genes encode proteins with DNA binding motifs, such

as homeo-domains, e.g., *fushi tarazu* (*ftz*) and *eve*, zinc-fingers, e.g., *odd* and *odd paired*

(*opa*), or helix-loop-helix domains, e.g., *hairy* (*h*). This suggests that they may directly

influence the transcription of their targets, i.e., some of the segment polarity genes.

The pair-rule genes have been subdivided into a primary group, which are directly

regulated by the gap (and co-ordinate) genes, and a secondary group which are regulated

by the primary pair-rule genes (for review see Ingham, 1988), although it is now thought that this is an over simplification of the situation (Ingham and Gergen, 1988; Carroll and Vavra, 1989; Yu and Pick, 1995).

hairy, *runt* (*run*), and *even-skipped* were classified as primary pair-rule genes since their expression patterns are not significantly altered by mutations in the other pair-rule genes, although it is now clear that there is a refinement of the expression patterns by interaction with other pair-rule gene products (Harding *et al*, 1986; Carroll and Vavra, 1989; Ingham and Gergen, 1988). These three genes are all expressed at the syncitial blastoderm stage in a uniform band which then resolves into a series of seven stripes in the region of the trunk (*hairy*: (Ingham *et al.*, 1985); *run*: (Gergen and Butler, 1988); *eve*: (Macdonald *et al.*, 1986). The stripes are spaced at two parasegment intervals, and can either be in exact register with the parasegments, e.g. the *eve* stripes which correspond to the even numbered parasegments, or can be out of register but still have the same spacing, e.g. *prd* (see Figure 6).

It has been shown both by genetic and molecular analysis that the stripes of the primary pair-rule genes are not globally regulated, but are regulated individually or in small groups (Goto *et al.*, 1989; Howard *et al.*, 1988; Warrior and Levine, 1990; Small, 1991; Stanojevic *et al*, 1989). The majority of the work has studied the regulation of *eve* stripe 2, which has been shown to be dependent on the activity of *bcd,* zygotic *hb, gt,*

and *Kr*. From analysis of *eve* expression in mutants for *bcd* and *hb,* in which *eve* stripe 2 is absent, it has been shown that they are both necessary for the activation of expression, whereas it has been shown by similar analysis in *gt⁻* and *Kr⁻* that they are both required for the repression of *eve* expression in the regions immediately adjacent to stripe 2 (Goto *et al.,* 1989; Small *et al.,* 1991; Warrior and Levine, 1990). Analysis of constructs which contain fragments of the 5' regulatory region of *eve* to drive *lacZ* expression, has demonstrated that there is a 480bp sequence approximately 1Kb upstream of the transcription start that is sufficient to get the specific expression of the second *eve* stripe 2 (Goto *et al.,* 1989), and it has now been found that there are overlapping activator and repressor binding sites that control the expression of this stripe (Frasch and Levine, 1987; Small *et al.,* 1993; Small *et al.,* 1991; Stanojevic *et al.,* 1989).

The secondary pair-rule genes such as *fushi tarazu, paired,* and *odd-paired* (*opa*) are regulated by the primary pair-rule genes. They seem to either be activated in a pair-rule pattern, or are repressed over a base of global activation. Much of the information available concerns the regulation of *ftz* which is initially expressed throughout the syncitial blastoderm, but expression is soon lost in the terminii, and then resolves into the characteristic pair-rule pattern (Hafen *et al.,* 1984). It has been shown by genetic analysis that *ftz* expression is regulated within the trunk region of the embryo by *cad* (Macdonald and Struhl, 1986; Mlodzik and Gehring, 1987), *run* (Carroll and Scott,

1986; Ingham and Gergen, 1988), *hairy* (Carroll and Scott, 1986; Howard and Ingham, 1986; Ish-Horowicz and Pinchin, 1987), and *eve* (Carroll and Scott, 1986; Ingham and Gergen, 1988). *cad* acts as an activator of *ftz* within its own expression domain, as in *cad⁻* mutants the posterior stripes of *ftz* are absent (Macdonald and Struhl, 1986). *run* is required for the maintainance of *ftz* expression after its initial activation by other factors (Carroll and Scott, 1986; Ingham and Gergen, 1988). *hairy* is a major factor in delimiting the correct expression of *ftz*; in *hairy* mutants *ftz* is expressed in an almost uniform manner, and in the background of uniform *hairy* expression there is no *ftz* expression, which suggests that *hairy* represses *ftz* (Carroll and Scott, 1986; Howard and Ingham, 1986; Ish-Horowicz and Pincin, 1987). The regulatory regions of *ftz* have been cloned and characterised (Dearolf *et al.*, 1989a; Dearolf *et al.*, 1989b; Hiromi and Gehring, 1987; Hiromi *et al.*, 1985) and consists of three separable domains, an autoregulative region, a neurogenic region which drives expression throughout the CNS later in development, and the 'zebra' element which is responsible for the pair-rule expression of *ftz* throughout the trunk region.

The result of this complex network of regulatory interactions is a pattern of pair-rule stripes that regulates expression of the segment polarity genes.

1.5.1.4. The segment polarity genes.

The segment polarity genes are a much more heterogeneous group than the other segmentation gene groups. They encode a variety of proteins, such as protein kinases, e.g., *fused* (*fu*), novel proteins, e.g., *patched* and *costal-2* (*cos2*), as well as those that contain DNA binding motifs, e.g., *engrailed*. Whereas the role of the pair-rule genes is to establish the periodic expression of the segment polarity genes, the segment polarity genes themselves are responsible for the maintenance and refining of the parasegmental boundaries and for the patterning of the parasegments. Two of the segment polarity genes, *engrailed* and *wingless* are critical for the maintenance of the parasegmental border (Martinez-Arias *et al.*, 1988).

The *engrailed* gene encodes a transcription factor which contains a homeobox (Fjose *et al.*, 1985; Poole *et al.*, 1985), a stretch of 60 amino acids that is known bind DNA (McGinnis *et al.*, 1984; Scott and Weiner, 1984). *engrailed* has been called a 'selector' gene as it is responsible for imparting the posterior cell state to those cells in which it is expressed. The expression of *en* in the *Drosophila* embryo has been described in great detail (DiNardo *et al.*, 1985; Ingham *et al.*, 1985; Kornberg *et al.*, 1985; Weir and Kornberg, 1985; Karr *et al.*, 1985). In brief, *en* transcription is first detected during nuclear cycle 14, throughout the trunk region of the blastoderm (DiNardo *et al.*, 1985; Weir and Kornberg, 1985). A stripe of *en*, the future stripe 2, forms at about 65% EL,

41

just posterior to the cephalic furrow, and another 13 are added in a roughly antero-posterior sequence, although stripe 12 appears slightly precociously. The stripes in the even numbered parasegments appear slightly prior to the stripes in the odd numbered parasegments, reflecting the pair-rule control of *en* expression. The stripes are initially one cell wide, and are separated by 2-3 cells which do not express *en*. As the germ band elongates, the *en* stripes expand to become 2-3 cells wide. The pattern of *en* at stage 11 consists of 14 stripes, corresponding to the posterior compartment of the trunk segments.

The segment polarity gene, *wingless*, is expressed at around the same time as *engrailed* (Baker, 1988a; Bejsovec and Martinez-Arias, 1991; Dougan and DiNardo, 1992). *wingless* is a member of the *Wnt* family of signalling molecules, and is the *Drosophila* homologue of the mouse *Int-1* proto-oncogene (Rijsewijk *et al*., 1987). *wg* transcript is detectable in the early blastoderm (Baker, 1987), and is initially found in the primordium of the foregut, in a ring surrounding the primordia of the hindgut and proctodeum, and several dorsal 'patches' (Baker, 1988a). Stripes begin to appear in the anterior of the embryo during the blastoderm stage, until at the extended germ band stage (stage 9) there is a stripe of cells that express *wg* anterior to the cells that express *en*. Each parasegmental boundary lies between, and depends upon, the stripes of *wg* and *en* expression (Perrimon and Mahowald, 1987).

The initial regulation of both *engrailed* and *wingless* is through the pair-rule genes, primarily *even-skipped* and *fushi-tarazu*, *paired* and *odd-paired* (reviewed in Fujioka and Jaynes, 1995; Ingham *et al.*, 1988; Ingham and Martinez-Arias, 1992; Mullen and DiNardo, 1995). *wg* expression is activated by *prd* and *opa* (Ingham *et al*, 1988; Hidalgo and Ingham, 1993), and *en* expression is regulated, directly or indirectly, by the pair-rule genes *sloppy-paired*, *runt*, *paired*, *odd-paired* (Fujioka and Jaynes, 1995), and *odd-skipped* (Mullen and DiNardo, 1995), as well as a number of segment polarity genes. There are two phases of *even-skipped* expression, that are regulated by separate enhancers (Fujioka and Jaynes, 1995). Only the early broad stripes of *eve* are important for initiating the expression of the *engrailed* stripes; the late expression is responsible for enhancing the expression of the *en* stripes in the odd numbered parasegments (Fujioka and Jaynes, 1995). In the even parasegments, *en* is expressed in those nuclei that express *ftz* but not *odd*, *eve* being responsible for repressing *odd* expression in the anterior cells of the *ftz* stripe (Manoukian and Krause, 1992; Fujioka and Jaynes, 1995). It is known that *eve* is a repressor of *engrailed* expression, but the odd numbered *en* stripes are initiated in the anterior cells of the early *eve* stripes. This is possible because of the overlap of the pair-rule gene, *paired*, and the anterior of the early *eve* stripe. At this position the levels of *eve* are not sufficient to repress the activation of *en* by *prd*. The expression of the pair-rule gene, *runt*, is also regulated by the early *eve* pattern, and is much more sensitive to repression by *eve* than is *prd*; this results in *en* being

expressed in those cells that are expressing *prd* but not *runt* (Manoukian and Krause,

1993). *sloppy-paired* is also repressed by *eve* in a concentration dependent manner,

rather like *runt*, and also acts to delimit the anterior border of *en* expression (Cadigan *et*

al., 1994b). This interaction between the regulatory actions of these pair-rule genes

generates the initial pattern of *en* and *wg* in the cellularising blastoderm. After the

activation of transcription of both *en* and *wg*, and the cellularisation of the blastoderm,

the other segment polarity genes play a greater role in the regulation of *en* and *wg*

expression.

After gastrulation, loss of expression of either *en* or *wg* results in the loss of expression

of the other in the embryonic epidermis (Martinez-Arias *et al.*, 1988). It has also been

shown that after gastrulation, and during germ band extension, the expression of *en* and

wg is delimited by the repressive actions of the segment polarity genes *naked* (*nkd*) and

patched, respectively (Martinez-Arias *et al.*, 1988).

1.5.1.4.1. *wingless* and the Wingless pathway.

wingless has been shown to be a member of the *Wnt* family which are known to function

as intercellular signalling molecules. Hence, the Wg protein can act as the signal from

the *wg* expressing cells to the *en* expressing cells, thereby maintaining expression of *en*.

Recently the product of *Drosophila frizzled 2* (*Dfz2*) was identified as the *wg* receptor in

cultured *Drosophila* cells (Bhanot *et al.*, 1996). Bhanot *et al.* (1996) have shown that

cultured *Drosophila* cells respond to Wg protein by increasing the levels of the

Armadillo (Arm) protein. The segment polarity gene *armadillo* (*arm*) encodes the

Drosophila homologue of the plakoglobin protein that when unphosphorylated can

associate with the adherens junctions. The product of the *zeste-white3* gene encodes a

serine-threonine kinase (Seigfreid *et al.*, 1990) that can phosphorylate the Arm protein,

releasing it from the adherens junction into the cytoplasm of the cell. The Arm protein

then relays the signal into the nucleus of the cell, and complexes with the homologue of

lymphocyte enhancer binding factor 1 (*Lef-1*), a murine transcription factor, encoded by

another segment polarity gene, *pangolin* (*pan*) (Brunner *et al.*, 1997), and effects

nuclear changes such as regulating *Ultrabithorax* (*Ubx*) expression (Riese *et al.*, 1997).

The products of two other segment polarity genes have been implicated in wingless

signalling on the basis of genetic epistasis experiments, *porcupine* (*porc*) (Siegfreid *et

al.*, 1994; Van den Heuvel *et al.*, 1993), and *dishevelled* (*dsh*) (Klingensmith *et al.*,

1994; Noordermer *et al.*, 1994; Siegfreid *et al.*, 1994). The product of *dsh* contains a

small GLGF repeat that is also found in proteins thought to be associated with cell

junctions (Klingensmith *et al.*, 1994; Theisen *et al.*, 1994), and relieves the inhibitory

effects of *zw-3* on the transduction of the Wg signal (Klingensmith *et al.*, 1989;

Perrimon and Mahowald, 1987). The product of *porc* has been shown to be required for

the secretion of the Wg protein (Siegfreid *et al.*, 1994; Van den Heuvel *et al.*, 1993), but

has recently been demonstrated to be more important in the autoregulatory control of *wg*

transcription than the paracrine signalling pathway described here (Manoukian *et al.*, 1995) (see Figure 7).

1.5.1.4.2. *engrailed, hedgehog* and the Hedgehog pathway.

The product of the *en* gene is a transcription factor, and therefore must indirectly regulate *wg* expression in the neighbouring cells. Another of the segment polarity genes, *hedgehog* (*hh*), is initially expressed in those cells that also express *en* (Lee *et al.*, 1992; Mohler and Vani, 1992; Tabata *et al.*, 1992), and has been shown to encode the signalling molecule from the *engrailed* expressing cells (Lee *et al.*, 1992; Tabata and Kornberg, 1994). Genetic data has shown that there is a correlation between the expression of *hh* and that of another segment polarity gene, *patched* (*ptc*) (Hidalgo and Ingham, 1990; Ingham *et al.*, 1991). Cloning of the *ptc* gene has shown that it encodes a novel 12 pass transmembrane protein that is normally expressed in those cells that flank the *hh* expressing cells which results in a pattern of expression that has two stripes per segment in the extended germ band embryo (Hooper and Scott, 1989; Nakano *et al*, 1989, Hidalgo and Ingham, 1990 and see Figure 6). Together these data suggested that the Ptc protein may function as the receptor of the Hh signal. It has recently been demonstrated that Hh protein binds to the Ptc protein which limits the distance over which Hh protein can diffuse through the segment, and triggers the transduction of the Hh signal into cells close (a few cell diameters) to the *Hh* expressing cells (Yen and

46

Struhl, 1996). The Ptc protein has also been shown to form a complex with the receptor like serpentine protein encoded by another segment polarity gene, *smoothened* (*smo*) (Alcedo, 1996; van den Heuvel and Ingham, 1996; Stone *et al.*, 1996; Alcedo and Noll, 1997). It is this Ptc/Smo complex that acts to relay the Hh signal into the interior of cells receiving the Hh signal.

The normal *ptc* expression pattern is described in detail in Chapter 4. Analysis of *ptc* expression suggests that it acts to delimit the expression of *wg*. It is clear, from the final expression pattern of ptc in which there are two stripes per segment, the most posterior of which fades at the late extended germ band stage, that regulation of *ptc* and *wg* is quite complex. When *hh* is expressed ubiquitously in the *Drosophila* embryo, both *ptc* and *wg* are ectopically expressed. These experiments show that *en* directly represses *ptc* expression, as ubiquitous *hh* expression drives *ptc* expression in all cells that do not express *en* (Ingham, 1993; Tabata and Kornberg, 1994). In these embryos, the *wg* expression domain expands to fill its entire competence group that is defined by the expression of the pair-rule gene *sloppy-paired* (Cadigan *et al.*, 1994b). These results are very similar to those found in *ptc⁻* mutant embryos (Hidalgo and Ingham, 1990; Ingham and Hidalgo, 1993; Martinez-Arias *et al.*, 1988), and taken together with data showing that the transcription of both *ptc* and *wg* are *hh* independent in *ptc⁻* mutant embryos, suggests that *ptc* suppresses its own transcription and represses that of *wg* (Ingham and Hidalgo, 1993; Ingham *et al.*, 1991). These results suggest that the Hh signal relieves

ptc repression, resulting in restricted expression of both *ptc* and *wg*. However, these results would also predict that *wg* would be expressed in two stripes, one either side of each *en* stripe. Given this, it would seem likely that *ptc* is expressed throughout the normal *wg* competence domain at a level that is undetectable by *in situ* hybridization, but which is sufficient to repress *wg* transcription in those cells that do not receive the Hh signal. It is, at present, unknown exactly how the posterior *ptc* stripe is regulated, and why there is no coincident *wg* expression.

A number of other genes have been implicated in the transduction of the Hh signal from the cell surface to the nucleus (see Figure 7). These include the segment polarity genes *fused* (Forbes, 1995; Ingham, 1993; Therond *et al.*, 1996), *Costal2* (Sisson *et al.*, 1997), *oroshigane* (*oro*) (Epps *et al.*, 1997), and *Cubitus interruptus* (*ci*) (Alexandre *et al.*, 1996), as well as the catalytic subunit of protein kinase A (PKA) (Li *et al.*, 1995; Pan and Rubin, 1995). Epistasis analysis has shown that the products of *fu*, *smo*, and *ci* are transcriptional activators of both *wg* and *ptc* (reviewed in Forbes, 1995), whereas Cos2 and PKA are repressors of of both *wg* and *ptc* transcription (Jiang and Struhl, 1995; Lepage *et al.*, 1995; Li *et al.*, 1995; Pan and Rubin, 1995; Preat *et al.*, 1991). This analysis of epistasis (Forbes, 1995) has suggested a particular ordering of genes within the pathway (see Figure 7); *fu*, which encodes a serine/threonine kinase (Preat *et al.*, 1991) is downstream of *ptc* and *smo*, *ci* is immediately upstream of the target genes *wg* and *ptc*, and *cos2*, which encodes a kinesin related protein forms a complex directly

with the Ci protein (Sisson *et al.*, 1997). *ci* encodes a protein that has strong homology to both the GLI family of proteins (Orenic *et al.*, 1990), and *the C. elegans* sex determining gene *transformer-1* (*tra-1*) (Zarkower and Hodgkin, 1992). Both the GLI proteins and Tra-1 have been shown to bind DNA in a sequence specific manner (Kinzler and Vogelstein, 1990; Zarkower and Hodgkin, 1993), and Ci has recently been demonstrated to be the transcription factor that directly controls transcription of both *wg* and *ptc*, and binds to the same DNA sequences as the GLI protein (Alexandre *et al.*, 1996; Von Ohlen *et al.*, 1997). The Ci protein has several PKA phosphorylation sites in its C-terminus, and from experiments in which PKA clones were induced in imaginal discs and could activate transcription of both *wg* and or *dpp* independently of Hh activity, but dominant active PKA could not suppress a *ptc* mutant phenotype it is thought that PKA may function in the hedgehog pathway by maintaining the effectors of the Hh signal in an inactive state.

Many of the segment polarity genes are expressed in the developing imaginal discs as well as in the embryo (Baker, 1988a; Brower, 1986; Phillips *et al.*, 1990; Whittle, 1990). The hedgehog signalling pathway has been shown to have a role in the patterning of the imaginal discs of *Drosophila* (Basler and Struhl, 1994; Capdevila *et al.*, 1994; Felsenfield and Kennison, 1995; Li *et al.*, 1995; Sanchez-Herrero *et al.*, 1996; Sanicola *et al.*, 1995; Tabata and Kornberg, 1994), and many of the experiments that have been used to dissect the pathway have been performed in the imaginal discs

(Alexandre *et al.*, 1996; Ingham and Fietz, 1995; Pan and Rubin, 1995; Sanchez-Herrero *et al.*, 1996; Tabata and Kornberg, 1994). One of the diffferences between the hedgehog pathway in the embryo and in the imaginal discs is the target genes. In the embryo, the major target gene is *wingless*, whereas in the wing imaginal discs, the major target gene is *dpp* (Ingham and Fietz, 1995; Pan and Rubin, 1995; Zecca *et al.*, 1995), which is responsible for patterning across the whole disc (Lecuit *et al.*, 1996).

Homologues of many of the genes identified as components of the Hedgehog pathway in *Drosophila*, including *hh*, *ptc*, *ci*, and *dpp* have now been found in several vertebrate species, the mouse (Goodrich *et al.*, 1996), chick (Marigo *et al.*, 1996) and fish (Concordet *et al.*, 1996). It has been demonstrated that the Hedgehog pathway is involved in the patterning of the limb, neural tube, and mid-line of these species, but not in trunk segmentation (reviewed in Fietz *et al.*, 1994; Ingham, 1995). It has also been shown that the spatial relationship between the homologues of *hh* and *ptc* have been conserved, and that they have the same functions in these species, i.e., the Hh protein is a signalling molecule and the Ptc protein is its receptor.

Partial *ptc* homologues have also been identified in several other insect species, indeed it was analysis of these sequences that facilitated the cloning of the mouse *ptc* homologue (Goodrich *et al*, in prep). There are also reports that putative *ptc*

50

homologues have been found during the *C. elegans* genome sequencing project,

although these sequences show most similarity at the level of hydropathy comparisons.
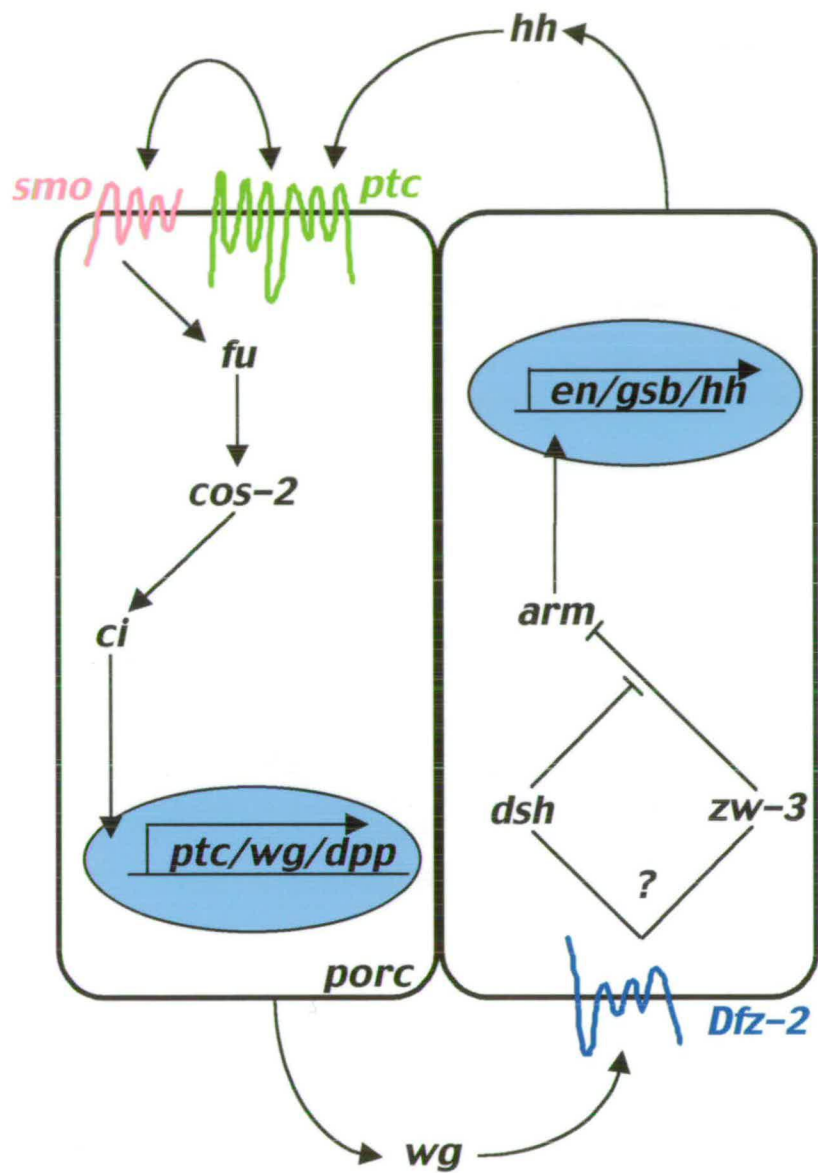
Figure 7. The Hedgehog and Wingless pathways.

A diagrammatic representation of the epigenetic interactions between genes known to play a role in intercellular signalling between *engrailed* and *wingless* expressing cells across the parasegmental boundary. Anterior is to the left, and the parasegmental border runs between the two cells. Abbreviations: hh - hedgehog; ptc - patched; smo - smoothened; fu - fused; cos-2 - costal 2; ci - cubitus interruptus; wg - wingless; dpp - decapentaplegic; Dfz-2 - frizzled 2; dsh - dishevelled; zw-3 - zeste-white 3 (shaggy); arm - armadillo; gsb - gooseberry; porc - porcupine (shown independently in the wingless expressing cell as it is more involved in the autocrine wingless signalling pathway).

Unfortunately, the battery of powerful genetic techniques that are used in *Drosophila*

are not available for the study of development in other insects. However, with the

advent of molecular biology techniques, it is possible to clone and study homologues of

the genes that have been shown to be important in *Drosophila*. It is generally assumed

that if a sequence motif is found to be conserved in the homologues of a gene in

different species, then it is likely to be important in the functioning of that protein. In

the absence of mutants in which to assess the function of a gene, less direct approaches

have to be employed. These include RNA *in situ* hybridisation, immunohistochemistry,

and immunocytochemistry to examine the temporal and spatial expression of the gene

transcripts and protein products, for comparison with expression in *Drosophila*.

Similarity between expression patterns indicates a possible conservation of function,

especially if it can be demonstrated that the expression patterns of other interacting

genes are also conserved (Concordet *et al.*, 1996; Marigo *et al.*, 1996). Expression

patterns are not, however, conclusive proof of a gene's function. Another way of

assaying function is to determine whether the gene, or its products, can rescue mutant

phenotypes in *Drosophila*. For example, it has been shown that a vertebrate homologue

of *hedgehog*, *sonic hedgehog* (*shh*), can rescue the *hh* mutant phenotype in the

Drosophila wing imaginal disc (Ingham and Fietz, 1995), and that *nanos* from various

Dipteran species can rescue *nanos* mutants in *Drosophila* (Binner and Sander, 1997).

1.7. The homologues of *Drosophila* segmentation genes in other Dipteran species.

Several homologues of *Drosophila melanogaster* segmentation genes have been cloned from other Dipteran species. These include a *hunchback* homologue from *Drosophila virilis* (Treier *et al.*, 1989), as well as smaller fragments from *Musca domestica*, *Calliphora vicina*, and *Psychoda cinerea* (Sommer *et al.*, 1989); small fragments of *Krüppel* from *M. domestica*, *Psychoda cinerea*, and *Sciara coprophila* (Sommer *et al.*, 1989); fragments of *bicoid*, *hunchback*, *Krüppel*, *knirps*, and *tailless* from *M. domestica* (Sommer and Tautz, 1991b); an *engrailed* homologue (Kassis *et al.*, 1986), and a *patched* homologue (Forbes, 1995), from *D. virilis*.

1.7.1. The Dipteran homologues of the co-ordinate gene, *nanos*.

Curtis *et al.* (1995) isolated homologues of *nanos* from *D. virilis*, *M. domestica*, and the midge, *Chironomus samoensis* (Curtis *et al.*, 1995). The predicted proteins showed overall similarities of 63%, 44%, and 30% to the *D. melanogaster* protein, respectively. This level of similarity rose to 97%, 89%, and 75% in the C-terminus of the proteins, a region that contains a novel zinc finger that may confer RNA binding properties to the protein. Examination of the expression pattern by *in situ* hybridisation revealed that

nanos RNA is localised in the posterior of the embryos in all the species studied. Use of a polyclonal antibody demonstrated that there was a posterior-to-anterior gradient of Nanos protein in each of the species, as in *D. melanogaster* (Ephrussi and Lehmann, 1992; Gavis and Lehmann, 1992). To assay conservation of Nanos protein function, they first attempted cytoplasmic transfer experiments in which cytoplasm from the posterior of an embryo from the donor species, *D. virilis*, *M. domestica*, or *C. samoensis*, was injected into the posterior of *nanos* mutant embryos from *D. melanogaster*. Rescue of the *nanos* mutant phenotype was strongest using cytoplasm from another Drosophilid, and weakest using the cytoplasm from the midge. To prove conclusively that it was *nanos* activity that rescued the *nanos* phenotype, P-element transformation was used to create transgenic *D. melanogaster* flies with *nanos* genomic DNA from *D. virilis*, *M. domestica*, and *C. samoensis*. These transgenic experiments proved that rescue of the *nanos* mutant phenotype was by *nanos* activity derived from the homologues of the other species, and not other cytoplasmic components carried over in the injections.

1.7.2. The segmentation gene homologues of *Drosophila virilis*.

Drosophila virilis is estimated to have diverged from *Drosophila melanogaster* about 60 million years ago (see Figure 8), which has been shown to have been sufficient time for non-essential sequences and elements of expression patterns to have diverged

(Blackman and Meselson, 1986; Kassis *et al.*, 1986; Tautz *et al.*, 1987b; Treier *et al.*, 1989).

Treier *et al.* (1989) sequenced 8.8Kb of genomic DNA from *Drosophila virilis*, which contain most of the coding region of a *hunchback* homologue, as well as several hundred base pairs of 5'UTR. They showed that the highest level of conservation is within the coding region of the gene, the two proteins having diverged by about 20%, with only one amino acid difference between the first zinc finger region, the second zinc finger region being identical in the two species. At the nucleotide level they showed that several of the sites that are known to be required for the binding of regulatory proteins are also conserved, including three Bcd binding sites (Driever and Nusslein-Volhard, 1989). The distribution of the Hb protein was examined using an antibody raised against the *D. melanogaster* Hb protein. They demonstrated that there are distinct differences in some aspects of the expression pattern such as a cap of Hb expression at the posterior pole at stage 14 in *D. melanogaster* that is completely absent *in D. virilis*, and a lack of dorsal-ventral modulation of the posterior stripe in *D. virilis* that is evident in *D. melanogaster*.

Kassis *et al.* (1986) examined the sequence conservation between the genomic sequence of an *engrailed* homologue from *D. virilis* with that of *D. melanogaster*, and showed that there was complete conservation of the C-terminal 30% of the protein which

contains the homeodomain, and an overall conservation of 71% over the rest of the protein (Kassis *et al.*, 1986). They did not present any expression data, although other groups have examined the distribution of the Engrailed (En) protein using a monoclonal antibody, mAb 4D9, in a range of insect species, e.g., *Schistocerca*, *Tribolium*, and have shown it to have a highly conserved pattern of expression in the posterior of the trunk segments (Brown *et al.*, 1994b; Fleig, 1990; Patel *et al.*, 1989a; Patel *et al.*, 1989b; Sommer and Tautz, 1991b).

Forbes (1995) cloned the *patched* homologue from *D. virilis*, and showed there was an overall homology across the entire sequence of 78% at the nucleotide level, and 80% at the amino acid level, although the degree of homology increased to 93% in the transmembrane domains. The expression pattern of *ptc* in *D. virilis* is almost indistinguishable from that of *ptc* in *D. melanogaster*, with a minor difference in the dorsal-ventral modulation of the posterior stripe in each segment in the two species. In *D. melanogaster* expression fades in the ventral part of the stripe, whereas it remains high in *D. virilis*, and an addition lateral patch of expression is seen *in D. virilis*, although no function has yet been ascribed to this patch.

Order	Suborder	Family	Species group	Species

Tipulidae

Nematocera

Chironomidae

C. samoensis

Diptera

Muscidae

M. domestica

Brachycera

Repleta

D. hydei

Virilis

D. virilis

Drosophilidae

Obscura

D. pseudoobscura

Melanogaster

D. melanogaster

250   200   150   100   50   present

Time of divergence (Myr)
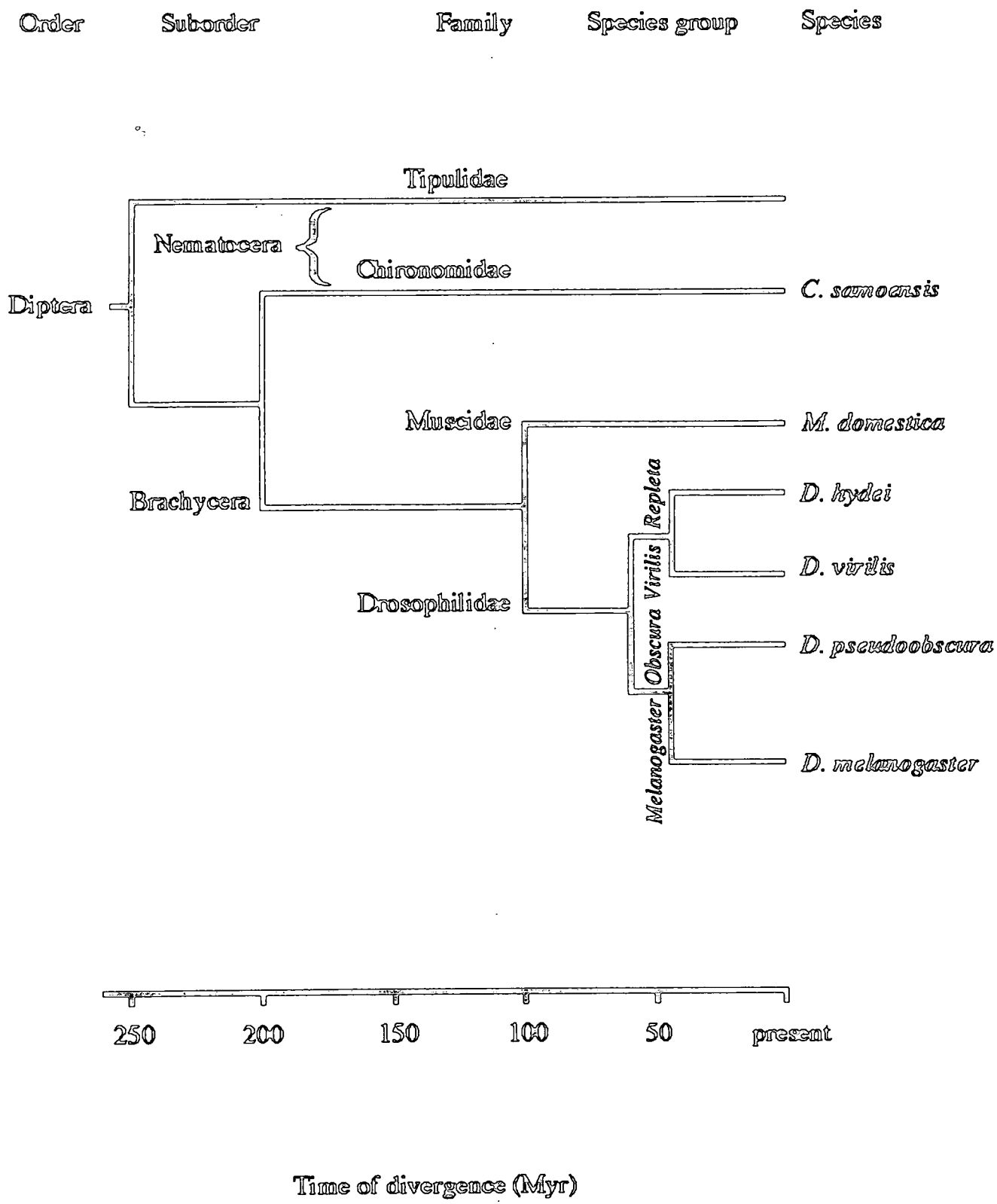
Figure 8. Partial phylogenetic tree of the relationship between selected Dipteran species.

Diagram modified from Curtis *et al* (1995). The times of divergence are in millions of years, and are approximate. This phylogeny was based on both the fossil record and immunological relatedness of larval serum proteins. The tree indicates that *D. melanogaster* and *D. virilis* diverged approximately 60 million years ago, whereas *M. domestica* diverged from these two species around 100 million years ago.

1.7.3. The segmentation gene homologues of *Musca domestica*.

Musca domestica is thought to have diverged from *Drosophila* at least 100 million years ago (Hennig, 1981), but exhibits the long germ band form of embryogenesis that is very similar to that seen in *Drosophila*.

Sommer and Tautz (1991) cloned fragments of *bicoid* (150bp), *hunchback* (345bp), *Krüppel* (243bp), *knirps* (530bp), and *tailless* (510bp) using the polymerase chain reaction (Sommer and Tautz, 1991b). Using these fragments to generate RNA probes, the expression patterns of these genes were analysed during *Musca* embryogenesis. In addition to these genes, the expression patterns of *hairy* and *engrailed* were examined using immunohistochemistry. The general pattern of the results they obtained showed that there is little difference in the expression patterns of the genes examined, and some, such as the *engrailed* pattern show no differences at all.

The zygotic expression of *hb* in *Musca* was very similar to that in *Drosophila*, transcript being detectable in the anterior half of the embryo at the blastoderm stage, before three stripes become visible. There is, however, slight heterochrony (or evolutionary change in developmental timing) between the development of the *hb* patterns in *Musca* and *Drosophila*, the most anterior stripe appearing relatively late in *Musca* in comparison to

its appearance in *Drosophila*. The *Musca* pattern then goes on to become a series of 13 irregularly spaced stripes at the beginning of gastrulation, which is unlike anything seen in any of the other segmentation genes, either in *Musca*, or *Drosophila*. The expression pattern of *tailless* in *Musca* shows some differences to that in *Drosophila*; an additional stripe forms in the anterior of the *Musca* embryo, and the posterior domain forms a stripe rather than a cap as in *Drosophila*. Sommer and Tautz argue that these differences are suggestive of differences in the regulation of *tailless* expression in the two species.

1.8. The homologues of *Drosophila* segmentation genes in non-Dipteran, long germ band species.

There have been few studies into the molecular basis of segmentation in non-Dipteran, long germ band insects. The expression patterns of *engrailed* (Fleig, 1990) and *even-skipped* (Binner and Sander, 1997) have been examined in the honeybee, *Apis mellifera* (Hymenoptera), using the monoclonal antibodies mAb 4D9 and mAb 2B8, respectively. In addition to these studies, even-skipped expression has also been examined in the long germ beetle, *Callosobruchus maculatus* (Coleoptera), again using the monoclonal antibody mAb 2B8 (Patel *et al.*, 1994b).

1.8.1. *engrailed* and *even-skipped* homologues in the honeybee, *Apis mellifera*.

There are several morphological differences between embryogenesis in *Apis mellifera* and *Drosophila melanogaster*. The morphological grooves appear during early gastrulation in the gnathal and thoracic regions in *Apis*, which is earlier than in *Drosophila*, and there is no involution of the *Apis* head segments (Fleig, 1990; Fleig and Sander, 1988; Krause, 1939). Examination of the *engrailed* expression pattern shows that the En protein is first detectable in the early gastrulation stages in stripes, one cell wide, in the first segmental grooves (Fleig, 1990). These stripes show an alternating pattern of intensities, which may reflect a pair-rule type regulation mechanism (Fleig, 1990). As gastrulation proceeds, stripes of En protein, one cell wide and of equal intensity, appear in an anterior to posterior sequence along the abdominal segments. These results show that *engrailed* is expressed in each of the metameric segments, but indicate that the mechanism of *en* regulation is somewhat different in the posterior of the honeybee embryo, due to the lack of a pair-rule like periodicity in the intensity of the stripes, and the strict anterior to posterior sequence in which they appear. Although these results indicate a change in regulation of *en* in the posterior of the honeybee embryo, it is possible that *en* transcription is controlled by the *Apis* homologues of the pair-rule genes; if the strength of activation and repression was equal in each of the parasegments, and there was no differences in the timing of *en* expression in each segment, then this regulation may not result in an alternating pattern of stripes.

even-skipped expression in the honeybee embryo shows a pair-rule phase, and then a later phase of segmental expression, as in *Drosophila* (Binner and Sander, 1997). There are six primary stripes that show a pair-rule type pattern of expression during gastrulation. The secondary, segmental, pattern of expression is generated by the loss of expression in the middle of the primary stripes, giving two narrow stripes (Binner and Sander, 1997). The appearance of the primary stripes in a pair-rule pattern, and then the resolving of a segmental pattern is reminiscent of *Drosophila eve* expression. However, the mechanism by which the secondary stripes are formed, by loss of expression in the primary stripes, is more like the generation of these stripes in the beetles (Patel *et al.*, 1994b), as in *Drosophila* the secondary stripes are formed by new *eve* expression within the interstripe region (Frasch and Levine, 1987).

1.9. Expression of *even-skipped* homologues in the Coleoptera.

Patel *et al.* (1994b) used the monoclonal antibody, mAb 2B8, to examine the expression of Eve protein during the embryonic development of three Coleopteran species, *Tribolium castaneum* (short germ), *Dermestes frischi* (intermediate germ) and *Callosobruchus maculatus* (long germ). It was shown that, in all three species, there are eight primary stripes of Eve, which are formed from a posterior region of expression by loss of expression in the interstripe regions. These primary stripes show a pair-rule

64

periodicity, before resolving into segmental stripes, again by loss of expression, this time from the cells in the middle of the *eve* stripes (Patel *et al.*, 1994b). These secondary stripes fade before the first appearance of morphological signs of segmentation. It was shown that these three species vary in the number of primary *eve* stripes that appear by the time the germ band cellularises; *Tribolium* has a single stripe at cellular condensation, another appearing at the onset of gastrulation, the next six appearing in an anterior to posterior sequence; *Dermestes* shows two stripes by cellular condensation, two more appear by the onset of gastrulation, and the final four appear during germ band elongation in an anterior to posterior sequence; *Callosobruchus* exhibits three stripes of *eve* expression by cellular condensation, three more appear by the onset of gastrulation, and the last two stripes appear in the posterior of the embryo during germ band elongation (Patel *et al.*, 1994b). Patel *et al.* (1994b) show that there is a pair-rule phase to the expression of *eve* homologues in these species, and they also show that the spatial relationship between *eve* and *en* is conserved between these beetles and *Drosophila*, by double labelling with the monoclonal antibodies mAb 2B8 (anti-Eve) and mAb 4D9 (anti-En/Inv) (Patel *et al.*, 1994b). Although *Callosobruchus* has been designated a long germ band insect, it is clear, from the expression of *eve* in its early embryo, that it is more closely related to *Tribolium* (short germ) than *Drosophila* (long germ), which is consistent with the accepted phylogenetic relationships between these insects. It is also clear from these experiments that it is possible to generate a pair-rule periodicity of expression in a cellular environment, but it has not been demonstrated that *eve* has a role in segmentation in these beetles, or whether the pair-

rule phase of expression of *eve* is necessary for any function. Patel *et al.* (1994) suggest

that the number of *eve* stripes that are present at the time of condensation of the syncitial

blastoderm is a more reliable indicator of germ type, than the traditional morphological

criteria.

1.9.1. Other segmentation gene homologues in the short germ band beetle, *Tribolium castaneum.*

The red-flour beetle, *Tribolium castaneum*, is one of the few insects that has any

potential as a genetic system (Beeman *et al.*, 1989; Sulston and Anderson, 1996). It is a

good example of a short germ band insect in which to study early development; it has

accessible embryos in which the techniques of RNA *in situ* hybridisation and

immunohistochemistry are well established, the number of mutations that are available

is increasing, and several of the homologues of the *Drosophila* segmentation genes have

been cloned (Brown *et al.*, 1994a; Nagy and Carroll, 1994; Sommer and Tautz, 1993;

Wolff *et al.*, 1995).

Using PCR primers designed to the zinc finger region of *Drosophila hb*, Sommer *et al.*
(1992) cloned a small fragment of the *Tribolium* homologue. Wolff *et al.* (1995) used
this fragment to screen both genomic and cDNA libraries and cloned the whole coding
region of the *Tribolium hb* homologue (Wolff *et al.*, 1995). Dot plot analysis showed
very little homology between the *Tribolium* and *Drosophila* sequences, although the
overall structure of the *Tribolium* gene indicated that it was the homologue of the
Drosophila hb gene. It was shown by RT-PCR, and RNA *in situ* hybridisation, that
there is a maternal phase of expression of *Tribolium hb* in the pre-cellular embryo, as in
Drosophila (see Figure 9). At this stage in development, all of the nuclei stain with
equal intensity, until just after the completion of the last nuclear division before
blastoderm when *hb* clears from the posterior terminus. This is reminiscent of the
regulation of *hb* in the posterior of *Drosophila* by Nos protein, suggesting that this
mechanism may have been conserved between the two species. The first zygotic
transcription of *Tribolium hb* is seen in the blastoderm, in an anterior cap and a large
posterior domain. Expression is also seen all of the serosal nuclei, and Hb protein is
seen in all of the serosal cells until very late in embryogenesis. From comparison with
the expression patterns of other segmentation gene homologues in *Tribolium*, it would
seem that it is the posterior domain of *hb* expression that is homologous to the anterior
expression domain in *Drosophila*, as it lies in the same region of the embryo that the

anterior pair-rule (Patel *et al.*, 1994b; Sommer and Tautz, 1993), and segment polarity

genes are expressed (Nagy and Carroll, 1994), and anteriorly to the *Tribolium Krüppel*

domain (Sommer and Tautz, 1993). Later, during germ band extension *hb* expression

becomes segmental, as in *Musca* (Sommer and Tautz, 1991b), and *Manduca sexta*

(Kraft and Jackle, 1994). This segmental phase of expression is not, however, seen in

Drosophila. A second posterior domain of *hb* expression is seen in *Tribolium*, and by

double labelling with a *hairy* probe was shown to overlap the sixth and seventh *hairy*

stripe, and is probably homologous to the posterior domain of expression seen in

Drosophila. Another posterior stripe of *hairy* appears in this posterior *hb* domain,

which is explained by *Tribolium* having one more abdominal segment than *Drosophila*.

There is also neuronal expression seen in *Tribolium* as in *Drosophila*, although it is

unknown whether the expression is seen in homologous neuroblasts in the two species.

1.9.1.2. *hairy and Krüppel.*

The coding region of a *Tribolium* homologue of the *Drosophila* pair-rule gene, *hairy*,

was cloned from a genomic library using a PCR fragment of *Tribolium hairy* as a probe.

Sequence comparison between *Tribolium hairy* and *Drosophila hairy* indicated that the

helix-loop-helix domains have been conserved, whereas stretches of sequence with

cryptic simplicity had diverged. *hairy* transcript is first detected in the blastoderm in

two circumferential stripes. These stripes shift posteriorly, and are only expressed

ventrally. *Kr* expression is seen in the posterior of the embryo at this stage, just prior to the appearance of a third *hairy* stripe. The domain of *Kr* expression moves anteriorly in the blastoderm and early germ band, preceding the appearance of *hairy* expression. A fourth *hairy* stripe appears by splitting of the third stripe, and the first and second stripe have faded in the early blastoderm, leaving two stripes in the early germ band, and expression becomes detectable in the mesoderm (see Figure 9). During germ band extension two *hairy* stripes are seen within the growth zone, and at the fully extended germ band stage *hairy* expression is seen in the proctodeum. Comparison of *hairy* and *en* expression indicates that the expression of *hairy* precedes that of *en*. Given that En expression, as detected by the monoclonal antibody mAb 4D9, is never detected in the growth zone, Sommer and Tautz argue that any *hairy-en* interactions must be taking place in the cellular environment of the extending germ band, and suggest that there has been conservation of the mechanisms of segmentation between the fly and the beetle.

1.9.1.3. *fushi tarazu.*

A cDNA library was screened with a probe containing the 3' end of the *Tribolium* homologue of *ftz*, which had been isolated during an earlier screen (Brown *et al.*, 1994a). The isolated cDNA clone contained a nearly full length *Tribolium ftz* fragment, which showed very little homology to the *Drosophila ftz* gene. Using DNA from beetles that have a partial deficiency of the *Tribolium* homeotic complex (Stuart *et al.*,

1991), they demonstrated by Southern hybridisation that the isolated cDNA clone maps into a region within the *sex combs reduced - Antennapedia* interval, which is in agreement with the genomic localisation of *Drosophila ftz*. RNA *in situ* hybridisation to the embryos of these homeotic deficient embryos failed to give any result, and the neuronal expression in embryos showed that a similar subset of neurons expressed this transcript in the beetle as *Drosophila* neurons express *ftz*, which together support the idea that the cDNA clone was the *Tribolium* homologue of *ftz*.

Expression of *Tribolium ftz* shows that it is first detectable in a broad band in the late blastoderm from 10-40% EL. The first *ftz* stripe appears by an increase in intensity in the broad domain in the region of the maxillary segment. Another six stripes arise during germ band extension, initially near the posterior of the germ band. These seven stripes show a pair-rule periodicity before weak expression becomes evident in the intervening (odd numbered) parasegments. Later in development there is expression in a subset of neurons in each segment.

Embryos carrying the homeotic deficiency show a homeotic like transformation, and not a pair-rule phenotype. Brown *et al.* (1994a) argue that this suggests that although the *Tribolium ftz* gene shows a pair-rule phase of expression, it plays no role in the segmentation of the *Tribolium* embryo.

70

Both genomic and cDNA clones of a *Tribolium en* homologue have been isolated

(Brown *et al.*, 1994b). The sequence and gene structure has shown that the *Tribolium*

en is more closely similar of *invected*. Brown *et al.* (1994b) suggest that the ancestral

gene in the common ancestor of holometabolous insects would have contained features

of both *inv* and *en*. The expression pattern of *Tribolium en* was visualised by

immunohistochemistry with the monoclonal antibody 4D9 (see Figure 9). En

expression was not detected in the blastoderm stage, a single stripe of En becoming

visible in the ventral posterior of the egg, in the position that the embryo will form, as

the germ band begins to cellularise. En protein accumulates exclusively in the nuclei of

the ectodermal cells at the posterior margin of each segment. The stripes of En appear

in a strict anterior to posterior sequence as the germ band elongates, each stripe

appearing prior to the first signs of morphological segmentation in that region of the

embryo. Stripes of En appear in the head of the *Tribolium* embryo throughout

embryogenesis, and the pattern of neurons in the brain that express En is similar in both

Tribolium and *Drosophila*. The later expression of En in the nervous system, in germ

band retracted embryos, is very similar to that which had previously been described in

other insects and crustaceans (Patel *et al.*, 1989a).

1.9.1.5. *wingless.*

A partial *Tribolium wg* homologue has been isolated from a cDNA library using a cDNA probe containing the third and fourth exons of the *Drosophila wg* gene (Nagy and Carroll, 1994). The predicted protein was calculated to be 78% similar (66% identical) to the corresponding region of the *Drosophila* Wg protein. All of the cysteines in the protein are conserved between the two species, and a hydrophilic region found in *Drosophila* Wg between residues 300 and 356, but is missing in the mouse Wnt-1 protein is also missing from the *Tribolium* Wg protein. Examination of the expression pattern by RNA *in situ* hybridisation shows that *wg* transcription was first detectable in the blastoderm in the posterior of the embryo, and then in the presumptive head lobes, a very similar pattern to that seen in *Drosophila*. The anterior domain of expression is anterior to the expression of the homologues of the pair-rule genes, indicating that this phase of *wg* expression is independent of pair-rule activity. These two domains of expression persist during embryogenesis. During germ band elongation, stripes of *wg* expression appear sequentially, with the first becoming apparent in the mandibular segment. The stripes are 2-3 cells wide and are separated by 6-8 cells which do not express *wg*. In the fully extended germ band there is a *wg* stripe in the three gnathal, three thoracic, and ten abdominal segments. These segmental *wg* stripes lie anteriorly adjacent to the *en* expressing cells, suggesting that the mechanism of segment polarity maintenance may be conserved between *Tribolium* and *Drosophila*.

The expression of *wg* becomes restricted to cells in the ventral part of each segment. In

the thorax most of these cells become the leg buds from which the adult legs develop.

In *Drosophila*, *wg* expression becomes restricted to anteroventral part of the leg

imaginal disc, which would be the corresponding cells to those in *Tribolium* that

become incorporated into the leg buds, suggesting that the role *wg* plays in patterning

the leg may have been conserved between the two species.

1.10. The segmentation gene homologues in the locust.

The Orthopterans, *Schistocerca gregaria* (locust), and *Schistocerca americana*

(grasshopper) are considered to be very similar to one another, and represent examples

of extreme short germ band insects. The homologues of the *Drosophila* genes, *eve*, *ftz*,

and *en* have been investigated in these species.

1.10.1. *even-skipped.*

Patel *et al.* (1992) cloned the *eve* homologue from the grasshopper, *Schistocerca*

americana, by screening cDNA libraries with a PCR fragment that had been amplified

using primers designed to the *eve* class of homeobox. They identified three regions of

homology between the predicted protein product and the *Drosophila* Eve protein. The

73

first conserved domain was the homeodomain itself, with 56 out of the 60 residues having been conserved. The region adjacent to the 3' end of the homeodomain had 17 out of 24 residues conserved, and another at the C-terminal of the protein had 7 out of 10 residues conserved. Using a monoclonal antibody 3C10, Eve protein is first detected in the grasshopper at the onset of gastrulation in the mesoderm in the posterior of the embryo (see Figure 9). During early germ band extension this mesodermal expression is maintained, until it fades at around 25% of development. At around 45% of development, Eve is detected in a subset of segmentally reiterated neurons, aCC, pCC, and RP2, segmentally reiterated cells in the dorsal mesoderm, and in a ring in around the anal pad (Patel et al., 1992). The early expression of Eve is reminiscent of the expression of the vertebrate eve homologues, Xhox3 (frog- Xenopus laevis), and Evx-1 (mouse). In both the mouse and frog, their eve homologues are expressed in the posterior mesoderm, and in the frog have been demonstrated to be involved in axial patterning (Ruiz i Altaba and Melton, 1989). The later expression in the neurons and around the anal pad is almost identical to that seen in Drosophila. Patel et al. (1992) suggest that the common ancestor to the vertebrates and arthropods had an eve-like gene which played a role in neurogenesis, and possibly axial patterning. Due to the lack of a pair-rule phase of expression in the grasshopper, Patel et al (1992) suggest that eve has been co-opted from this neuorgenic role into one of segmentation.

The homologue of *Drosophila ftz* has been characterised in the locust, *Schistocerca gregaria* (Dawes *et al.*, 1994). It has been named *Dax* (divergent Antennapedia class homeobox gene) because its homeodomain matches the Antp class homeodomain consensus sequence, contains the YPWM motif that is characteristic of the homeotic genes, and shows very little sequence homology to the *Drosophila ftz* gene. The *Drosophila* Ftz protein does not contain the YPWM motif, although it is found in the *Tribolium* Ftz protein, and the residues that flank this motif in *Tribolium* are also found in the Dax protein. Taking the similarities between the *Tribolium ftz* homologue and *Dax,* together with the expression pattern of *Dax* during neurogenesis (see below) Dawes *et al* (1994) conclude that *Dax* is the locust *ftz* homologue.

Expression of the Dax protein during embryogenesis was analysed using an anti-Dax antibody. Prior to cellularisation of the embryonic primordia, Dax is detected in a cresent of condensing nuclei in the dorso-ventral region of the egg. Post cellularisation, the Dax protein is found in the posterior of the embryo until it fades away at 23% of development. At around 15% of development Dax is found from the anterior of the gastral groove to the very posterior tip of the embryo. Dax protein then clears from the posterior tip of the embryo between 18 and 20% of development, before fading completely at around 23% of development. There is no pair-rule like phase of Dax

expression, and its expression fades 10-15 hours before *en* expression can be detected in a similar region of the embryo. As development proceeds, *Dax* expression is detected in the neurectoderm, on both sides of the midline. The expression domain of *Dax* shows a sharp anterior limit within the mandibular segment, just as *ftz* in *Drosophila*. Expression soon becomes detectable in a subset of the cells that delaminate from the neurectoderm. The first delaminated neuronal cells that express *Dax* are the MP2 (midline precursor-2) cells, which are also the same cells that express *ftz* in *Drosophila*. Around 20% of the cells in the anterior of each neuromere eventually express *Dax*, in a pattern that has striking similarities to the expression pattern of *ftz* in *Drosophila*.

1.10.3. *engrailed.*

Patel *et al.* (1989b) used the monoclonal antibody, mAb 4D9, to screen a cDNA library, and isolated several clones which contain *en* like homeobox sequence from the grasshopper, *Schistocerca americana*. From the sequence obtained from the homeobox, it was shown that they had cloned an *en* homologue, but could not state whether it was more closely related to the *Drosophila en* or *inv* gene. By Southern hybridisation it was demonstrated that the grasshopper only has a single *en* homologue (Patel *et al.*, 1989a; Patel *et al.*, 1989b). Using the monoclonal antibody, mAb 4D9, they followed *en* expression during embryogenesis in the grasshopper (see Figure 9). The En protein is first detected at 17% of development, after the onset of gastrulation, in a stripe in the

76

posterior of the first and second thoracic segments (Patel *et al.*, 1989a). The stripes

anterior to this region appear between 19 and 22% of development, whereas the stripes

in the posterior of the embryo appear in an anterior to posterior sequence, with the last

stripe in the tenth abdominal segment becoming evident at 31% of development. When

the stripes first arise they are 2-3 cells wide with rough borders, and separated by 7-8

cells. Slightly after the formation of an *en* stripe it widens to 5-6 cells, and the

interstripe increases in width to around 12-13 cells. The widening of the *en* stripes

seems to involve both cell division and recruitment of non-expressing cells (Patel *et al.*,

1989a). As the stripes widen, the anterior border of each stripe becomes sharper, and

Patel *et al.* (1989a) suggest that this anterior border demarcates the parasegmental

boundary. *en* is also expressed in cells in the posterior of the developing limb buds,

which corresponds well with the expression of *en* in the posterior compartment of the

imaginal discs in *Drosophila*.

D. melanogaster   *T. castaneum*   *S. gregaria*

follicle cells

bcd

nos

hb

A

E

I

nurse cells

h1 h2

hb

B

Kr

F

J

bcd

nos

embryo

Kr   *kni*

Kr

eve
in mesoderm

C

G

K

h stripes

en

h3 h4

en

eve
in mesoderm

en

D

H

h

L

thorax
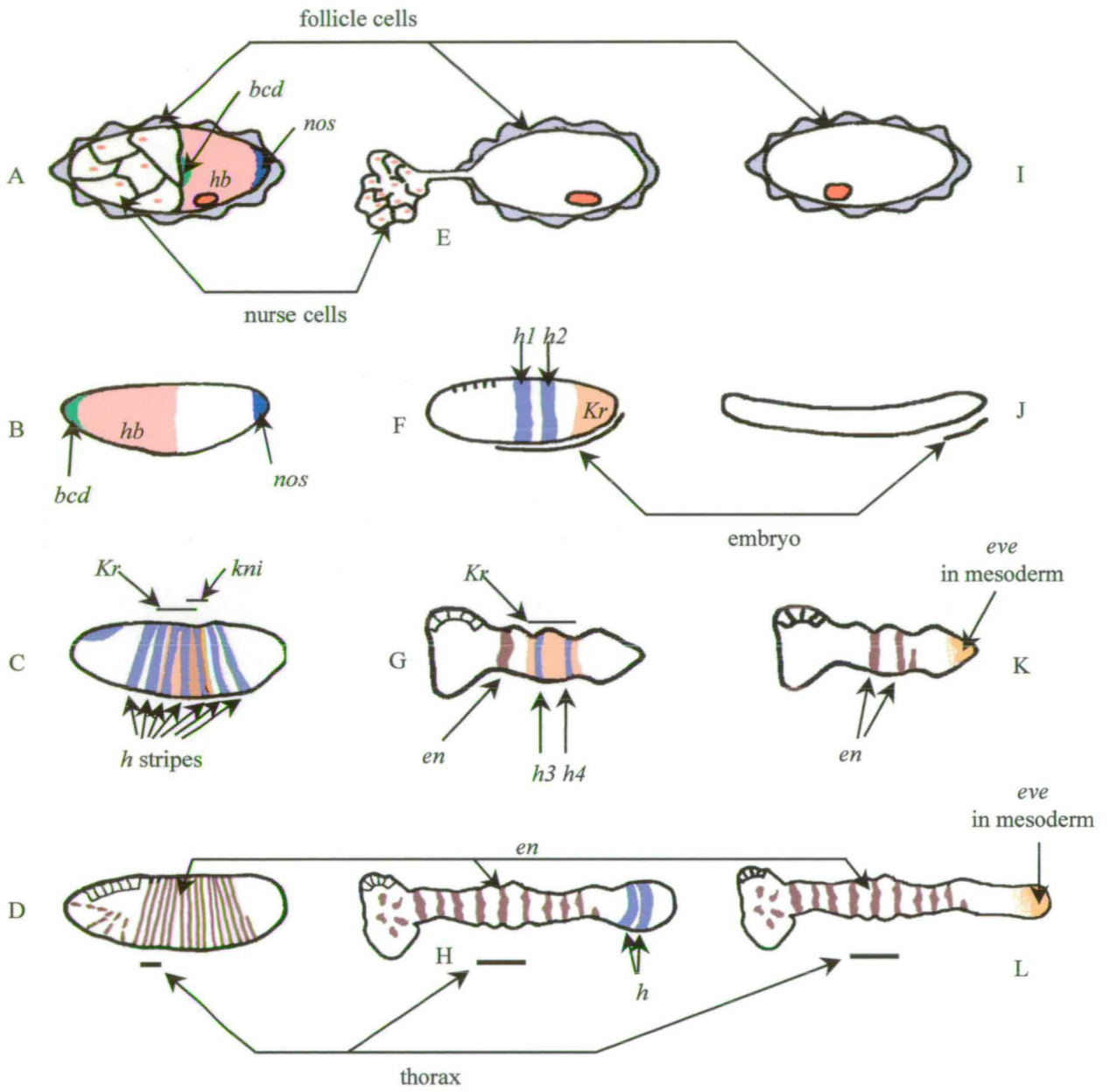
Figure 9. Representation of ovary types and oogenesis in *D. melanogaster*, *T. castaneum*, and *S. gregaria*.

Panels A, B, C, and D show *Drosophila* development, panels E, F, G, and H show *Tribolium* development, panels I, J, K, and L show *Schistocerca* development. Expression domains are all labelled on the figure. Panels A, E, and I show oocytes from the three species (A = *Drosophila*, meroistic ovary, long germ embryo; E = *Tribolium*, meroistic ovary, short or intermediate germ embryo; I = *Schistocerca*, panoistic ovary, short germ embryo). Panels B, C, F, and J are all syncitial stages, whereas panels D, G, H, K, and L depict cellular stages. Figure adapted from French, 1993.

1.5. Aims of this study.

Prior to this study, few of the genes involved in the segmentation of the *Drosophila* embryo had been cloned and characterised in other insects, and those studies that had been undertaken had concentrated on the gap and pair-rule genes. The results of these investigations suggested that at least some of the genes are involved in the segmentation of these other insects. From the results obtained from the analysis of pair-rule homologues in the locust (Patel *et al*, 1992; Dawes, 1994), it seemed that the pair-rule genes might only have a role in the segmentation of the embryos of higher insects. Of all the segmentation genes, it seemed that the segment polarity genes were the most likely to play a role in the segmentation of most insect embryos as they are the only level of the *Drosophila* hierarchy that are known to operate within a cellular environment. In accordance with this, it had been shown that the distribution of the En protein is remarkably similar to that seen in *Drosophila*, in a number of different insect species, such as the locust and *Tribolium* (see Figure 9).

Given this data it was decided that in order to further investigate the mechanisms controlling the different modes of insect embryogenesis, it would be valuable to study the segment polarity gene homologues in species showing short or intermediate germ band embryogenesis.

Expression of both the *engrailed* and *wingless* homologues from the short germ beetle, *Tribolium castaneum*, had been examined (Nagy and Carroll, 1994), and it had been shown that the spatial relationship between the two genes seen during *Drosophila* embryogenesis had been conserved. Because the rest of the genes involved in signalling across the parasegmental border, i.e., the components of the Hedgehog signal transduction pathway, had not been investigated in *Tribolium*, or any other short or intermediate germ band insect, it was not known whether the regulation of the homologues of *en* and *wg* had been conserved in short germ insects. It was, therefore, an aim of this study to investigate other genes known to be members of the Hedgehog signal transduction pathway in an intermediate germ band insect, the cricket, *Acheta domesticus* (Orthoptera).

The segment polarity gene, *ptc*, was chosen as the focus for this investigation, as it was already known that it plays an important role in the segmentation of the *Drosophila* embryo, being the putative receptor for the Hh signal.

Prior to this study, *ptc* had only been cloned from *Drosophila melanogaster* (Hooper and Scott, 1989; Nakano *et al.*, 1989), and *Drosophila virilis* (Forbes, 1995). It was an aim of this study to characterise the homologues of *ptc* from various insect species, in order to further our understanding of the evolution of the structure and function of this

gene. To facilitate this, homologues of *ptc* were cloned from the long germ insect,

Musca domestica (Diptera), and the intermediate germ insect, *Acheta domesticus*

(Orthoptera).

To investigate the role of *ptc* homologues in the development of insects exhibiting

different modes of embryogenesis, the pattern of its expression was examined during the

development of both *Musca*, and *Acheta*. As the Hedgehog pathway is known to play

roles in both segmentation and limb patterning in *Drosophila*, the expression of the *ptc*

homologues were examined during both processes in *Musca* and *Acheta*. As *Musca* is a

holometabolous insect, this involved examining expression during embryonic and larval

development. *Acheta*, in contrast, is a hemimetabolous insect, which means that its

appendages begin to develop during embryogenesis from limb buds. Comparisons of

the expression patterns of *ptc* in the segmenting germ bands of both insects would allow

ptc function to be assessed in the segmentation of a long, and intermediate germ band

insect. Examination of the expression of *ptc* in the imaginal discs of *Musca*, and the

limb buds of *Acheta* would allow for an indirect comparison of *ptc* function in the limb

development of a holometabolous, and a hemimetabolous insect.

Change in the regulation of the *ptc* homologues was thought to be a possible mechanism

for changes in the functioning of the Ptc protein. To investigate this, an attempt was

made to clone the regulatory regions of the *Musca ptc* gene, in order to see whether

these could drive the expression of a reporter gene, *LacZ*, in the same pattern as

endogenous *ptc* in *Drosophila melanogaster*.

2. Materials and Methods.

2.1. Molecular biology techniques.

2.1.1. General cloning techniques.

Standard cloning techniques such as minipreps of plasmid DNA, plasmid vector preparation (dephosphorylation and end-filling), DNA separation, southern transfers and hybridisations were performed as described in (Sambrook *et al*, 1989).

2.1.2. Restriction endonuclease digestion.

Restriction endonuclease digests were performed using the suppliers buffers and suggested enzyme concentrations (Pharmacia, Promega, Stratagene) in 10 µl volumes, and incubated at 37°C for 1-2 hours.

2.1.3. Ligations.

Ligation mixes were made as follows;

(a) for DNA with cohesive ends;

   0.1-0.5 µg digested vector DNA

   an approximately equimolar amount of insert DNA

   1 µl 10x ligation buffer (Promega)

   0.5 units T4 DNA ligase (Promega)

   Distilled water to 10 µl

(b) for blunt ended DNA;

   0.1-0.5 µg digested vector DNA

   approximately 3x molar amount of insert DNA

   1 µl 10 x ligation buffer (Promega)

   5 units T4 DNA ligase (Promega)

   Distilled water to 10 µl

These ligation mixes were incubated at 14°C for between 1 and 18 hours.

2.1.4. Preparation of competent cells.

Stocks of competent XL1-Blue (Stratagene, see appendix C.) were prepared using a protocol as described in (Inoue, 1990). XL1-Blue were streaked onto LB agar plates (see appendix A.), without antibiotic, and grown overnight at 37°C. Several large colonies were used to inoculate 250 ml of SOB medium (see appendix A.) in a 2 litre flask. This was incubated at 18°C in a shaking water bath (set at 250 rpm) until an optical density (OD) of 0.6 - 0.8 was achieved. The flask was put on ice for 10 minutes, after which time the culture was transferred to a 500 ml centrifuge tube and spun at 2,500 g for 10 minutes at 4°C. The pellet was immediately resuspended in 80 ml of ice cold TB buffer and incubated in an ice bath for 10 minutes. This was then spun at 2,500 g for another 10 minutes at 4°C. The supernatant was discarded , and the pellet resuspended in 20 ml of ice cold TB buffer (see appendix B.). DMSO was added, with gentle agitation, to a final concentration of 7 %. The bacteria were again incubated in an ice bath for 10 minutes, then aliquoted in 150 μl volumes into Eppendorf tubes, frozen in liquid nitrogen and stored at -70°C.

2.1.5. Transformations.

The competent bacteria were thawed on ice, and 50 μl used for each transformation. A maximum of 10 μl of a ligation, or 1-2 μl of plasmid DNA, were added to the bacteria

and these were mixed together by flicking the tube. This mix was left on ice for 30-45 minutes and then heat shocked at 42°C for 2 minutes. 1 ml of LB (see appendix A.) was added and the bacteria incubated at 37°C for 45-60 minutes. If the bacteria were being transformed with an identified plasmid then 20-100μl would be plated onto LB-ampicillin agar (see appendix A.) plates. If, however, the bacteria were being transformed with a ligation they would be spun down at 6,000 rpm in a bench top centrifuge, most of the supernatant removed and the pellet of bacteria resuspended in a volume of around 50μl and all of this plated on LB- ampicillin agar plates.

2.1.6. Gel purification of DNA bands.

DNA of less than 1Kb was purified from TAE agarose gels using the GeneClean II kit (BIO 101). The gels were run at 50-75 volts until the bands were sufficiently separated. The bands were visualised on a transilluminator on the low setting to minimise the UV light induced damage to the DNA. The appropriate bands were cut out of the gel and placed in 1.5 ml Eppendorf tubes. 500 μl of sodium iodide solution were added and the gel melted at 55°C. 5-8 μl of glassmilk (Bio 101) were added, depending on the amount of DNA in the band (see maunfacturers' instructions), and the tubes placed on ice for 5-10 minutes. The glassmilk was spun down in a bench top centrifuge and the pellet washed in new wash. This washing procedure was repeated 3 times. The pellet was air

dried, and the DNA eluted into either 20 μl distilled water or TE (see appendix B) by

incubating at 55°C for 3 minutes.

For larger sized DNA a different protocol was used. The bands were cut out of the gel

in the manner described above, but the gel was spun through glass wool columns and

the resulting liquid collected. The DNA was then ethanol/salt precipitated as described

in (Sambrook *et al.*, 1989).

2.1.7. ^{32}P-dCTP labelling of DNA probes.

DNA probes for Southern hybridisations and library screening were labelled using the

Multiprime DNA labelling kit (Amersham).

1. The tubes from the kit were thawed on ice.

2. 50-500 ng of DNA was dissolved in 10 μl of distilled water, and denatured by boiling

for 2-5 minutes, and then chilled on ice.

3. The following reaction mix was prepared;

    10 µl DNA

    10 µl 10 x Reaction Buffer (Amersham)

    5 µl Random Oligonucleotide Primer

    distilled water to 50 µl

    1-5 µl ^{32}P-dCTP (depending on the state of decay)

    2 µl Klenow DNA Polymerase

4. This mix was left at 37°C for more than 2 hours, or overnight at room temperature.

5. The efficiency of the labelling reaction was assessed by running <1µl of the mix on thin layer chromatography paper and exposing to X-OMAT film (Kodak) for 5 minutes. Unbound radio-nucleotide seperates from the labelled DNA and a comparison between the levels of the two resultant bands can be made.

6. Unincorporated radio-nucleotide was removed by spinning the mix through a small G-50 Sephadex column. See 2.1.7.1.

7. Before the radio-labelled DNA could be used in a hybridisation reaction it was heat denatured by boiling for 2-5 minutes.

2.1.7.1. Removal of unincorporated radionucleotide from a labelling reaction.

1. A hole was punched in the bottom of a 750 μl Eppendorf tube using a 19 gauge hypodermic needle, and the cap cut off.

2. The tube was partly (approx 1/3) filled with quite firmly packed glasswool and placed inside a 1.5 ml Eppendorf tube, the cap of which had been removed.

3. 50μl of G-50 Sephadex beads that had been washed and stored in (1 M)Tris.HCl (pH 7.4) were added to the small Eppendorf tube and briefly (15-20 seconds) spun in a bench top centrifuge. The liquid collected in the large Eppendorf tube was discarded. Three additions of G-50 Sephadex were normally sufficient to pack the glasswool with enough Sephadex beads.

4. The columns were briefly (15-20 seconds) spun another 3 times, discarding the liquid collected in the large Eppendorf tube each time. The columns were stored at 4°C until they were required.

5. To remove unincorporated radionucleotide from a labelling reaction, the mix was loaded into one of these columns and spun briefly in a bench top centrifuge. The labelled DNA is collected in the large Eppendorf tube, largely free (>90%) from unincorporated radionucleotide.

2.1.8. DNA library screening.

2.1.8.1 Bacteriophage λ libraries.

Approximately 200,000 phage, representing a 3- fold genome coverage, of an *M. domestica* λDASH II (Stratagene) genomic DNA library were plated using XL1-Blue MRA (Stratagene, see appendix C.) as the host bacteria, and transfered to Hybond N+ nylon filters (Amersham) as described in Sambrook *et al.*, (1989). The hybridisations were carried out essentially as described in Sambrook *et al.*, (1989), except that they were carried out in glass tubes, not plastic bags. λ DNA was prepared from clones containing appropriate inserts as described in 2.1.8.1.1.

2.1.8.1.1. Preparation of λDNA.

Liquid cultures of λ bacteriophage were set up in the following manner;

1. A single plaque was cored from a plate, added to 0.5ml of XL1-Blue MRA plating cells in a 50ml Falcon tube, and incubated for 15 minutes at room temperature.

2. 5 ml LB supplemented with 5 mM $CaCl_2$ was added to the tube which was then incubated at 37°C with vigorous shaking until lysis was observed. Lysis normally occurred between 3.5 and 4.5 hours.

3. Two drops of chloroform were added and the tube was shaken for a further 5 minutes.

4. The tube was then centrifuged at 3,000 rpm for 10 minutes to remove cellular debris. The supernatant was removed and either used to prepare λ DNA directly, or stored at 4°C and used as a stock for setting up new liquid cultures.

5. To prepare λ DNA, 5 ml lysate was mixed with 5 ml 20% PEG, 2.5M NaCl and placed on ice for 1 hour.

6. This was centrifuged at 4,000 rpm for 10 minutes.

7. The pellet was resuspended in 750 μl LB and transfered into a 1.5 ml Eppendorf tube.

8. 750μl DE-52 resin was added, and the tube inverted 20-30 times.

9. This was then centrifuged at 13,000 rpm for 5 minutes in a bench top centrifuge.

10. The supernatant was transfered into a new tube and centrifuged for another 5 minutes.

11. The supernatant was transfered into a new tube and 2 μl 5 mg/ml DNAse and 10 μl 10 mg/ml RNAse were added.

12. This was incubated at 37°C for 10 minutes.

13. For each ml of supernatant, 42.5 μl 10% SDS and 17.5 μl 0.4 mg/ml proteinase K were added.

14. This reaction was incubated at room temperature for 5 minutes.

15. 173 μl of 3 M potassium acetate was added and the reaction was incubated at 88°C for 20 minutes.

16. The tube was then put on ice for 10 minutes and then spun at 13,000 rpm in a bench top centrifuge.

17. The supernatant was removed and placed in a 2 ml Eppendorf tube.

18. An equal volume of isopropanol was added and the tube placed at -70°C for 10 minutes.

19. The DNA was pelleted by centrifuging at 13,000 rpm for 10 minutes in a bench top centrifuge, washed in 70% ethanol and redissolved in 50 µl TE.

2.1.8.2. λZAP II phagemid library.

Approximately 250,000 λ particles of an *A. domesticus* cDNA λZAP II (Stratagene) were plated using XL1-Blue MRF' (Stratagene, see appendix C.) as the host bacteria. These were transfered to Hybond N+ nylon filters and hybridised at low stringency as described in Sambrook et al., (1989), except that they were carried out in glass tubes rather than plastic bags. Excision of the phagemid was carried out as described in 2.1.8.2.1.

2.1.8.2.1. Phagemid excision

The λZAP II vector system (Stratagene) allows for the excision of the Bluescript

phagemid containing the insert of interest. This is accomplished by co-infecting the

λZAP II phage and a helper phage, ExAssist, into XL1-Blue MRF' cells, where the

phagemid is excised from the λZAP II vector. The phagemid is then selected for in

SOLR cells (see appendix C.) on LB-ampicillin agar plates and treated as a plasmid for

the remainder of the cloning. All of the necessary bacterial cells and stock solutions

were prepared following the manufacturers' instructions.

2.1.9. Preparation of total RNA from cricket embryos (modified from Sambrook, *et al.*, 1989).

1. Approximately 200 eggs were collected and put into a Dounce homogeniser.

2. 5 x the volume of lysis buffer was added and the eggs homogenised with 10-12

strokes.

Lysis buffer:

50 mM NaCl

50 mM Tris.Cl (pH 7.5)

5 mM EDTA (pH 8.0)

0.5 % SDS

200 µg/ml proteinase K

3. The homogenate was transferred into a heat sterilised 15 ml corex tube and sarcosyl was added to a final concentration of 1%. This was then spun at 10,000 rpm at room temperature in a Sorval HB4 rotor.

4. The homogenate was then drawn through a 16 gauge hypodermic needle in order to shear the DNA, and then transfered into a polypropylene tube containing 3 ml 5.7 M CsCl/0.01M EDTA (pH 7.5). This was then spun overnight in a Beckman SW41 swing-out rotor at 30,000 rpm.

5. Most of supernatant was aspirated off, the bottom of the tube cut off, and the pellet allowed to dry.

6. The pellet was resuspended in 100 µl of 0.2% SDS and transferred into an Eppendorf tube which was spun at 13,000 rpm in a bench top centrifuge.

7. The supernatant was aspirated off and kept. The pellet was resuspended in another 100 µl of 0.2% SDS, and spun again at 13,000 rpm in a bench top centrifuge.

8. The supernatant was again aspirated and the two SDS supernatants were pooled.

9. The RNA was precipitated by adding 20 µl sodium acetate (pH 5.2) and 2.5 volumes of ethanol, mixing thoroughly and placing at -20°C for at least an hour. The RNA was then pelleted by spinning in a bench top centrifuge for 10 minutes, washed in 80% ethanol, dried and resuspended in 100 µl TES (see appendix B.).

10. The guanadinium was removed by addition of 100 µl chloroform/phenol (4:1), vortexing and spinning for 5 minutes in a bench top centrifuge. The top (aqueous) phase, containing the RNA, was removed and placed into a new RNA grade tube.

11. The RNA concentration was then determined spectrophotometrically by measuring the OD at 260 nm.

12. An equal volume of chloroform:n-butanol (1:1) was added to the RNA solution, vortexed and spun at 13,000 rpm for 5 minutes in a bench top centrifuge.

13. The top (aqueous) phase was then removed to a new tube, and 1 µl glycerol, 20 µl of 3 M sodium acetate (pH 7.5) and 2.5 volumes of ethanol were added and was then placed at -20°C for 20 minutes.

14. The RNA was pelleted by spinning at 13,000 rpm in a bench top centrifuge and resuspended in 43 µl DEPC water.

15. 5µl medium salt buffer (restriction endonuclease medium salt buffer), 1 µl RNAse inhibitor, and 1µl RNAse free DNAse were added and this was incubated at 37°c for 15 minutes.

16. 100 µl phenol:chloroform (1:1) was added, vortexed and then spun in a bench top centrifuge at 13,000 rpm for 5 minutes.

17. The top phase was removed to a new tube, and sodium acetate was added to a final concentration of 0.3 M. 1 µl of glycogen and 2.5 volumes of ethanol were added and the tube placed at -70°C for 20 minutes, (or stored at -20°C for long term storage).

18. The RNA was pelleted by spinning in a bench top centrifuge at 13,000 rpm, washed in 80% ethanol, and resuspended in an appropriate volume of water.

19. The RNA concentration was then redetermined spectrophotometrically by measuring the OD at 260 nm.

This RNA was then used as the template for the synthesis of first strand cDNA using the Amersham kit.

2.1.10. DNA amplification using the polymerase chain reaction (PCR).

All PCR reactions were based around a general 'hot start' protocol (Mullis, 1991), using a Hybaid Thermal Reactor.

2.1.10.1. PCR reactions.

The PCR reactions were performed in 50 µl volumes, overlaid with approximately 20 µl of light mineral oil (Perkin Elmer, Sigma).

A general formula for the reaction mix is:

1 x Reaction Buffer (Boehringer Mannheim, Perkin Elmer)

200-250 µM dNTP's (25mM mix of dATP, dCTP, dGTP, dTTP)

0.5µM each primer

≈1µg template DNA

1.5-3.0 mM $MgCl_2$

0.2-0.5µl Taq (Boehringer Mannheim, Perkin Elmer)

A generalised 'hot start' PCR program is;

Hot Start	denature	94°C	7 minutes
		72°C	hold (add Taq)
35 cycles of	denature	94°C	30 seconds
	anneal	45-55°C	30 seconds
	extend	72°C	60-90 seconds
Final		72°C	10 minutes

2.1.11. Double stranded sequencing.

Three methods of sequencing double stranded plasmid DNA templates were used: manual radioactive dideoxy chain termination sequencing, automated fluorescent dideoxy chain termination sequencing, and automated fluorescent cycle sequencing. T3, T7, SK and KS sequencing primers were purchased from Stratagene, M13F, M13R were supplied with the sequencing kits, and all other primers were synthesised by the Imperial Cancer Research Fund (ICRF) facility at Clare Hall.

2.1.11.1. Preparation of DNA for double stranded sequencing.

Midipreps of plasmid DNA were performed using the Qiaprep 100 kit (Qiagen), which utilises a modified alkaline lysis protocol.

1. Single colonies were used to inoculate 50 ml of LB containing 100μg/ml ampicillin. Cultures were grown overnight at 37°C with shaking.

2. The bacteria were pelleted in Oakridge tubes by spinning at 6,000g for 10 minutes.

3. The pellet was resuspended by repeated pipetting in 4 ml buffer P1 (50mM Tris.HCl (pH 8.0), 10 mM EDTA (pH 8.0), 100 μg/ml RNAse A).

4. 4 ml of buffer P2 (200mM NaOH, 1% SDS) was added, mixed by inverting the tube several times, and incubated at room temperature for 5 minutes.

5. 4 ml of ice chilled buffer P3 (3.0M potassium acetate (pH 5.5))was added, mixed by inverting the tube several times, and incubated on ice for 15 minutes.

6. The tubes were then centrifuged at 4°C for 30 minutes at 30,000g to pellet the cellular debris.

7. Whilst the tubes were being centrifuged, the columns were equilibrated using 4 ml of buffer QBT (750 mM NaCl, 50 mM MOPS, 15% ethanol, 0.15% Triton X-100, pH 7.0).

8. The supernatant was added to the equilibrated column and allowed to flow through the resin.

9. The DNA bound to the resin in the column was then washed by two additions of 10 ml buffer QC (1.0 M NaCl, 50 mM MOPS, 15% ethanol, pH 7.0).

10. The DNA was then eluted with 5 ml of buffer QF (1.25 M NaCl, 50 mM Tris.HCl, 15% ethanol, pH 8.5)

11. The DNA was precipitated with 0.7 volumes of isopropanol and pelleted by centrifuging at 4°C for 15 minutes at 15,000g.

12. The supernatant was removed and the pellet washed in 5 ml of cold 70% ethanol.

13. The DNA was air dried and resuspended in 100μl of water. The concentration of the solution was determined spectrophotometrically by measuring its OD at 260 nm.

2.1.11.2. Manual sequencing.

Sequencing of double stranded plasmid DNA was performed manually using the Sequenase Version 2.0 kit (USB Corporation), which is based on the dideoxy chain termination sequencing method (Sanger, 1977). Template DNA was prepared using either the Qiagen midiprep system, or the alkaline lysis miniprep as described in Sambrook *et al.*, (1989).

1. 5 μg plasmid DNA in 10 μl of distilled water was denatured by adding 0.1 volumes of 2 M NaOH, 2 mM EDTA and incubating at 37°C for 30 minutes.

2. This was denatured by adding 0.1 volumes of 3 M sodium acetate (pH 5.5).

3. The DNA precipitated by adding 3 volumes of cold ethanol and placing at -70°C for 15 minutes or -20°C overnight.

4. The DNA was washed in 70% ethanol, air dried and was then ready for sequencing.

5. 2 µl of 5 x reaction buffer, 1 µl 0.5 pmol/µl primer, and 7 µl distilled water were added to the side of the tube, and were mixed with the DNA by briefly spinning in a bench top centrifuge. Annealing of the primer was achieved by incubating at 37°C for 30 minutes.

6. While annealing, 2.5µl of each of the termination mixes (ddA, ddC, ddG, ddT) were aliquoted into the wells of a 60 well Terasaki plate and kept at room temperature.

7. After the annealing reaction had finished, the tubes were chilled on ice, and the termination mixtures were warmed to 37°C.

8. The required amount of 5 x labelling mix was diluted with distilled water, and kept on ice until required.

9. The labelling reaction mix was prepared as the annealing reaction was nearing completion;

Labelling reaction mix:

1 μl 0.1 M DT

2 μl diluted labelling mix

0.1-0.5 μl ^{35}S-dATP (depending on the state of decay)

2 μl prediluted (1:8) Sequenase T7 DNA Polymerase

5.5 μl of the above mix was added to the ice-cold annealed DNA, and incubated at room temperature for 2 minutes.

10. 3.5 μl of the labelling reaction was added to each of the wells in the Terasaki plate containing the termination mixes, and incubated at 37°C for a maximum of 5 minutes.

11. The reactions were stopped by adding 4 μl of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF).

12. The samples were heat denatured at 75°C for 2 minutes before loading onto a 6% polyacrylamide sequencing gel.

2.1.11.2.1. Denaturing gel electrophoresis.

6% sequencing gels were made by diluting 50% Long Ranger Gel Solution (FMC

Bioproducts), or a 40% stock solution of acrylamide:bis-acrylamide (19:1), with 1 x

TBE (see appendix B.), 50% urea. The gels were run on an S2 sequencing apparatus

(Gibco BRL) using 1 x TBE running buffer.

2.1.11.3. Automated sequencing.

2.1.11.3.1. Fluorescent dideoxy sequencing.

Fluorescent dideoxy chain termination sequencing of plasmid DNA was performed

using the Autoread Sequencing kit and the Automated Laser Fluorescent (ALF) DNA

sequencer (Pharmacia Biotech). The template DNA was prepared using the Qiagen

midiprep system.

1. 5-10 µg of template DNA in 32 µl of distilled water was denatured by adding 8 µl of

2 M NaOH and incubating at room temperature for 10 minutes.

2. This was denatured by adding 7 µl 3 M sodium acetate (pH 4.8) and 4 µl of distilled

water.

3. The DNA was precipitated by adding 120 μl ethanol and placing on dry ice for 15 minutes.

4. The DNA was pelleted by spinning for 15 minutes in a bench top centrifuge, and then washed in 70% ethanol.

5. The DNA was then air dried and resuspended in 10 μl water.

6. Annealing of the primer to the template DNA was carried out by adding 2 μl of 2.1 pmol/μl fluorescent primer, and 2 μl annealing buffer to the resuspended DNA, and incubating at 65°C for 5 minutes, 37°C for 10 minutes, and room temperature for at least 10 minutes.

7. Whilst the annealing reaction was cooling, the T7 DNA polymerase was diluted with enzyme dilution buffer to a concentration of 4 units/μl, and 2.5 μl of the sequencing mixes (ddA, ddC, ddG, ddT) was aliquoted into the wells of a Terasaki plate and warmed to 37°C.

8. 1 μl extension buffer, 3 μl DMSO, and 2 μl diluted T7 DNA polymerase were added to the annealing reaction and mixed by pipetting.

9. 4.5 µl of this mix was added to each of the wells containing pre-warmed sequencing mixes, and incubated at 37°C for 5 minutes.

10. The reactions were stopped by adding 5µl stop solution, and denatured by heating to 95°C for 3 minutes before loading onto the sequencing gel.

2.1.11.3.1.1. Gel electrophoresis using the ALF DNA sequencer.

Gels were made and run on the ALF sequencer according to the manufacturers' instructions by G. Clark and A. Davies at the I.C.R.F. facility at Lincoln's Inn Fields.

2.1.11.3.2. Fluorescent cycle sequencing.

Fluorescent cycle sequencing using dye-labelled terminators was performed using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit and an ABI PRISM 377 DNA Sequencer (Perkin Elmer).

1. The sequencing reactions were set up as follows;

8.0 µl Terminator Ready Reaction Mix (with AmpliTaq DNA Polymerase FS)

3-5 µg template plasmid DNA

3.2 pmol primer

distilled water to 20 µl

2. This reaction was cycled in an MJ research PTC-200 thermal cycler using the

following program;

25 cycles of:	denature	96°C	30 seconds
	anneal	46-50°C	15 seconds
	extend	60°C	4 minutes

3. The PCR products were ethanol/salt precipitated by adding 2 µl 3M sodium acetate

(pH 5.5) and 50 µl ethanol into a 1.5 ml Eppendorf tube and then adding the sequencing

reactions to the same tube, and placing on ice for 10 minutes.

4. The DNA was pelleted by spinning in at 13,000 rpm in a bench top centrifuge for 20

minutes, washed in 250 µl 70% ethanol and dried in a speedvac.

5. The samples were then resuspended in 6-9 µl of loading buffer (deionised

formamide:25 mM EDTA (pH 8.0), 50 mg/ml Blue dextran (5:1)), and heated to 90°C

for 2 minutes before loading.

2.1.11.3.2.1. Gel electrophoresis using the ABI 377 sequencer.

The reactions were run on an ABI 377 sequencer using a 4.25% acrylamide:bis-

acrylamide (19:1) gel and 1 x TBE running buffer. Gels and buffers were made by G.

Clark and A. Davies at the I.C.R.F. facility at Lincoln's Inn Fields.

2.1.11.4. Sequence analysis.

All sequence analysis was performed using the GCG suite of programs (Wisconsin

Package Versions 7, 8, and 9, Genetics Computer Group (GCG), Madison, Wisc.).

2.2. Maintenance of animal stocks.

2.2.1. *Acheta domesticus.*

Acheta domesticus stocks were maintained at 26°C in large plastic boxes containing

scrunched-up paper towel, a cotton wool bunged water bottle and a petri dish of ground

cat food ('GoCat' - Carnation Pet Foods) with an occasional supplement of lettuce. Adult colonies were also given small pots of damp compost in which to lay their eggs. Laying pots that were for egg collection were left in the cage overnight, as shorter times resulted in considerably fewer eggs being laid.

2.2.2. *Musca domestica.*

Musca domestica stocks were kept at 25°C in densely populated muslin cages containing a damp cotton wool pack and were provided with a piece of red meat approximately 6 cm^3, which provided both food and a place for them to lay their eggs. The meat was usually left in the cage overnight, after which it was removed, and the eggs either removed, washed, aged and fixed or, if larvae were required, placed in a sandwich box containing scored larvae media.

2.2.3. *Drosophila melanogaster* and *Drosophila virilis.*

Drosophila stocks were kept at 25°C in bottles containing a medium of yeast, agar, flour, malt extract and molasses. Larvae were also collected and maintained on this medium. When collecting eggs, the flies were transferred into cages with a petri dish containing a medium of agar, sucrose and apple juice, with yeast paste smeared onto it,

and left at 25°C for the required period of time after which they were washed, aged and fixed.

2.3. Fixation of material for immunohistochemistry and *in situ* hybridisation.

2.3.1. Cricket embryos.

A. domesticus eggs were removed from the compost using fine forceps and a stereo dissecting microscope, and washed in 0.1% Triton X-100 in 1 x phosphate buffered saline (PBS, see appendix B.) in a glass embryo dish. Using two 19 gauge hypodermic needles attached to 1ml syringes, the anterior end of the egg was cut off and the embryo and yolk were squeezed out. Any adhering yolk was removed, and the extra-embryonic membranes, if present, were broken. When approximately 20 embryos had been dissected from their eggs, they were transferred into fixative (4% paraformaldehyde in 1 x PBS) and fixed for 30 - 40 minutes, depending on the developmental stage. They were then washed several times in 1 x PBS, and then dehydrated through a methanol series (30%, 50%, 70%, 95%, 100%) and stored at -20°C.

2.3.2. Fly embryos.

2.3.2.1. Dechoriomation.

The flies were provided with medium on which to lay their eggs. The eggs were removed by paintbrush and placed in a small wire mesh basket and washed several times in 0.1% Triton X-100 in 1 x PBS. The basket was placed on paper towel to remove the excess liquid. The wire mesh basket was then placed in a petri dish containing 25% sodium hypochlorite (Sigma) for 3 - 5 minutes, after which the embryos were washed 3-5 times in distilled water.

2.3.2.2. Fly embryos

Dechorionated embryos were fixed in a 1:1 mix of heptane : fix (4% paraformaldehyde in 1xPBS) for 20 minutes (*D. melanogaster* and *D. virilis*) or 30 minutes (*M. domestica*). The lower phase was removed and replaced with an equal volume of methanol and the tube vigorously shaken, causing the majority of the embryos to 'pop' out of their vitelline membranes. The embryos were removed from the lower phase and placed into a fresh Eppendorf tube and rinsed several times with 100% methanol. If the embryos were for use *in situ* hybridisations they could be stored at -20°C; embryos used for immunohistochemistry, however, were always used on the day of collection.

113

2.3.3. Fly larvae.

Third instar larvae were collected using a paintbrush and washed in 0.1% Triton X-100 in 1 x PBS. They were then placed in 1 x PBS in a glass embryo dish, and using two 19 gauge hypodermic needles attached to 1ml syringes, the anterior third of the larvae was removed and the posterior discarded. Using fine forceps, the anterior ends of the larvae were inverted and as much of the non-imaginal material removed as possible. The larvae 'heads' were then fixed in batches of 3 (*M. domestica*) or 5 (*D. melanogaster* and *D. virilis*) in 4% paraformaldehyde in 1 x PBS for 20 minutes (*D. melanogaster* and *D. virilis*) or 30 minutes (*M. domestica*) at 4°C. The larvae 'heads' were then washed several times in 1 x PBS and always used immediately for either *in situ* hybridisations or immunohistochemistry.

2.4. Whole mount immunohistochemistry (modified from Ingham and Martinez-Arias, 1986).

2.4.1. Fly embryos.

1. The fly embryos were rehydrated using 0.1% Triton X-100 in 1 x PBS.

2. The embryos were then resuspended in PAT.

PAT:

    1 x PBS

    0.1% Triton X-100

    1% bovine serum albumin

3. The embryos were then blocked by rolling in PAT at room temperature for at least 4 hours, typically 6-8 hours.

4. The embryos were then resuspended in PAT containing the appropriate concentration of primary antibody (5E10 1:5000 ; 4D9 1:4) and rolled overnight at 4°C.

5. The embryos were washed for 3 x 30 minutes in PBT containing 2% of the serum from the animal in which the secondary antibody was raised (PBTS).

PBT:

    1 x PBS

    1% Triton X-100

    0.1% bovine serum albumin

6. The embryos were then incubated for 2 hours at room temperature in PBT containing preabsorbed alkaline phosphatase conjugated secondary antibody at a 1:2000 dilution.

7. The embryos were washed for 10 minutes in PBT, and then twice for 10 minutes each in PTW.

PTW:

    1 x PBS

    0.1% Tween 20

8. The embryos were then rinsed twice for 2 minutes each in freshly prepared staining buffer.

staining buffer (50 ml):

    1.25 ml NaCl (4 M)

    2.5 ml $MgCl_2$ (1 M)

    5.0 ml Tris (1 M, pH 9.5)

    0.05 ml Tween 20

    41.2 ml distilled H_2O

9. The embryos were resuspended in 1ml of staining buffer with 4.5 µl nitro blue tetrazolium (NBT, Promega) and 3.5 µl 5-bromo-4-chloro-3-indolyl phosphate (BCIP, Promega), and the reactions were incubated in the dark until the colour had developed.

10. The reactions were stopped by washing the embryos 3 x 2 minutes in 1 x PTW, 10 mM EDTA.

11. The embryos were either dehydrated through a methanol series (30%, 50%, 70%, 95%, 100%) and stored at -20°C or mounted in glycerol and photographed on Ektachrome 64T (Kodak) using Zeiss DIC or bright field optics.

2.4.2. Fly imaginal discs.

Immunohistochemistry on imaginal discs was performed as previously described for fly embryos, with the following changes: step 5 was replaced by 3 x 1 hour washes in PBTS, step 6 was replaced by a 4 hour incubation at room temperature in the secondary antibody at 1:2000 in PBTS; step 7 was replaced by 3x1 hour washes in PBT. After the signal detection reactions had been stopped the larvae 'heads' were placed in 80% glycerol and the discs dissected away from the cuticle. The discs were then mounted in 100% glycerol and photographed as described for embryos.

2.5. Whole mount *in situ* hybridisation.

2.5.1. DIG labelled RNA probe preparation (modified from Tautz, 1989).

1. DIG labelled RNA probes were prepared by incubating the following reaction mix at 37°C for 2 hours;

Reaction Mix: 2 μl linearised DNA template (0.5 mg/ml)

2 μl Nucleotide Mix - 25μl DIG UTP (Boehringer Mannheim)

7.15μl 100 mM ATP (Boehringer Mannheim)

7.15μl 100 mM CTP (Boehringer Mannheim)

7.15μl 100 mM GTP (Boehringer Mannheim)

7.15μl 100 mM UTP (Boehringer Mannheim)

20 μl 10 mM Tris.HCl (pH 8.0)

2 μl Transcription Buffer (as supplied with the RNA polymerase)

0.5 μl RNAse Inhibitor (Boehringer Mannheim)

2 μl RNA polymerase (Boehringer Mannheim)

11.5 μl distilled H_2O

2. 2 μl of DNAse I was added and this was incubated at 37°C for a further 15 minutes.

3. The RNA was precipitated by adding 1 µl of tRNA (50mg/ml stock concentration), 1.25µl 8 M LiCl (Sigma), and 75 µl ethanol and placing at -20°C for several hours (usually overnight).

4. The RNA was pelleted by spinning at 13,000 rpm in a bench top centrifuge, washed in 70% ethanol and resuspended in 50 µl distilled H_2O.

This method was used to prepare the probes for hybridisation to the fly embryos, fly imaginal discs, and the *A. domesticus* embryos.

2.5.2. Fly embryos.

2.5.2.1. Pre-hybridisation.

1. The embryos were rehydrated through a methanol series (95%, 70%, 50%, 30%, 1 x PBS)

2. They were then re-fixed in 4% paraformaldehyde in 1 x PBS for 20 minutes.

3. The embryos were then rinsed 5 x 2 minutes in PTW.

4. Then 1 x 2 minutes in PTW:Hybridisation Buffer (1:1)

Hybridisation buffer (10 ml):

      5.0 ml deionised formamide

      2.5 ml 20 x SSC (see appendix B.)

      20 µl   tRNA (50 mg/ml)

      5 µl Heparin (100 mg/ml)

      2.45 ml distilled H_2O

5. The embryos were then prehybridised in hybridisation buffer at 55-60°C for at least 1 hour (usually 4-6 hours).

2.5.2.2. Hybridisation.

Embryos were hybridised overnight at 55-60°C in hybridisation buffer containing heat denatured RNA probe. The concentration for each probe had to be determined separately, but usually fell within the range of 1:200 and 1:1000.

2.5.2.3. Washes.

1. The embryos were washed at 65-70°C for 20 minutes in pre-heated hybridisation buffer.

2. They were then washed at room temperature for 20 minutes in PTW:Hybridisation Buffer (1:1).

3. Then 5 x 20 minutes at room temperature in PTW.

2.5.2.4. Signal detection.

1. The embryos were incubated at room temperature for 1-2 hours in PTW containing preabsorbed alkaline phosphatase conjugated anti-DIG secondary antibody (Boehringer Mannheim) at a final concentration of 1:2000.

2. They were then washed 4 x 20 minutes in PTW.

3. Next the embryos were rinsed twice in staining buffer (see immunohistochemistry).

4. The embryos were resuspended in staining buffer and 4.5 μl NBT and 3.5 μl BCIP were added.

5. The reactions were incubated in the dark until the colour developed and the reactions stopped by washing in 1 x PTW, 10 mM EDTA.

2.5.3. Fly imaginal discs and *A. domesticus* embryos.

2.5.3.1. Pre-hybridisation.

1. The embryos/discs were rehydrated through a methanol series (95%, 70%, 50%, 30%, 1 x PBS).

2. The material was then re-fixed in 4% paraformaldehyde in 1 x PBS for 20 minutes.

3. The material was then rinsed 3 x 2 minutes in PTW.

4. The material was then incubated at room temperature in 50 µg/ml proteinase K for between 2 and 4 minutes depending on the batch of proteinase K.

5. The proteinase K was then inactivated by rinsing 2 x 30 seconds in 1 x PTW, 2 mg/ml glycine.

6. The material was then rinsed 2 x 2 minutes in PTW.

7. The discs were re-fixed for 4 minutes in 4% paraformaldehyde in 1 x PBS, the cricket embryos were re-fixed for 7 minutes.

8. The material was then rinsed 5 x 2 minutes in PTW.

9. Then once in PTW:Hybridisation Buffer (1:1) for 2 minutes.

9. The pre-hybridisation was carried out in hybridisation buffer at 55-60°C for at least 1 hour (usually 4-6 hours).

2.5.3.2. Hybridisation, washes, and signal detection.

These were all carried out as described for fly embryos.

2.5.4. Microphotography of specimens.

This was performed as described for microphotography of immunohistochemistry specimens.

3. Molecular Cloning of the *M. domestica patched* homologue.

3.1. Introduction.

3.1.1. The *Drosophila melanogaster* segment polarity gene *patched*.

In 1980, Nüsslein-Volhard and Wieschaus published a paper detailing a mutation screen designed to identify loci involved in the generation of the segmental pattern of *Drosophila melanogaster*. By examining the cuticles of mutant larvae, fifteen loci were identified that produced abnormalities in either the number of segments present, or the polarity of those segments. These loci were classified into one of three groups, segment polarity, pair-rule or gap, depending on their mutant phenotypes.

One of the loci identified during this work was *patched* (*ptc*). On the basis of the associated mutant phenotype, *ptc* was classified as a segment polarity gene. Segment polarity mutants have the normal number of body segments, with part of the segment being deleted and replaced by a mirror image duplication of the rest of the segment. However, unlike the other segment polarity mutants found in the screen, *ptc* mutants have twice the number of segment boundaries, due to the middle region of the segment being deleted and replaced by a mirror image of the marginal regions, including the segmental boundary.

ptc was first cloned from *Drosophila melanogaster*, almost simultaneously, by two laboratories. Nakano *et al* (1989) predicted that the *ptc* locus contained a single open reading frame (ORF) of 3897 bp flanked by a 5' untranslated sequence of approximately 750 bp, and a 3' untranslated region of about 890 bp. This ORF was predicted to encode a protein consisting of 1299 amino acids with a molecular weight of 144 Kd.

Hooper and Scott (1989) identified a 6.3Kb transcript, and showed that the *ptc* gene consists of 6 or more exons covering at least 17.1 Kb of the genome. It was predicted that the transcription unit would consist of a single large ORF of 4152 bp, encoding a protein of 1286 amino acids with a molecular weight of approximately 143 Kd.

Although there was some discrepancy between the two laboratories as to the sizes of transcripts, predicted ORF's and protein sizes, both predicted that the *ptc* gene would encode a transmembrane protein of approximately 140 -145 Kd. The predicted protein showed little homology to any known protein, except for a stretch of 49 amino acids towards the 3' end, residues 968-1016 of the Hooper and Scott protein, which showed limited homology to the growth hormone, somatotrophin (Abdel-Meguid *et al.*, 1987). Within this region there are 16 identical residues and nine conservative amino acid substitutions. Using predicted hydropathy plots (Kyte and Doolittle, 1982), Nakano *et al* (1989) showed seven regions of hydrophobicity, of which four were large enough to span the membrane more than once. From the positioning of groups of charged

residues, which are thought to flank transmembrane domains, Nakano *et al* (1989) predicted ten transmembrane domains, two of which were long enough to enter and leave the membrane on the same side. Using this topology, and by placing the majority of the eight glycosylation sites extracellularly, both of the termini are situated intracellularly. This resulted in a predicted topology of two large extracellular loops that are hinged together by a smaller intracellular loop, which is reminiscent of the structure of bacterial ion channels (Nikaido and Saier, 1992). Hooper and Scott (1989) used several methods to determine the hydropathy of the *ptc* protein. Using a method that calculates hydropathy over 17 amino acid stretches (Klein *et al.*, 1985), it was predicted that the protein has 14 hydrophobic stretches, 12 of which were predicted to be transmembrane α-helices. A different method that allows for residue charge (Eisenberg *et al.*, 1984), predicted eight transmembrane α-helices, and yet another method which compensates for protein conformation (Argos and Rao, 1986), predicted only seven.

3.1.2. The *ptc* homologues.

3.1.2.1. Homologous gene cloning.

There are two classes of homologous genes (or homologues). Paralogous genes arise through the duplication of a common ancestral gene, whereas orthologous genes have diverged from the same ancestral sequences during the evolution of different lineages.

The cloning of a gene is often only the first step in the study of that gene. Usually, the structure and function of its product will be analysed, probably in the context of some complex biological process. Once this has been performed in one system, it raises a number of questions, such as whether the homologues of this gene are present in other organisms, what the function of these homologues might be in these species, and how these homologues, and the processes they are involved in, have evolved over time.

Cloning homologous genes utilises two techniques, PCR, and DNA hybridisation. Homologous genes are normally identifiable by similarity in their nucleic or amino acid sequence, and it is this similarity that is exploited in their cloning. Both PCR and hybridisation techniques rely on there being regions of DNA sequence of sufficient similarity that PCR primers, or nucleic acid probes, can be designed to bind to regions of the paralogues in the species of interest, or to its orthologues in another species.

The cloning and molecular characterisation, i.e., characterisation at the structural level, of homologous genes can provide a great deal of information. The degree of similarity, or rather the amount of change, between gene sequences will chart the evolution of a gene, and the creation of molecular phylogenies is often the best way to accurately deduce evolutionary relationships between species, and even phyla. For example, the sequences of the 5S RNA (Hori, 1975), and 18S RNA (Field et al., 1988), have been

used in the generation of a phylogeny that includes species from phyla spanning the entire animal kingdom.

The study of gene homologues can also be useful in the study of gene function. Those domains and motifs that are found to be conserved between homologous genes should be those that are important for gene function. Sequence data from gene homologues can hence, sometimes, provide clues as to gene function. For example, if a DNA binding motif is found to be conserved between homologues it would suggest that those genes may play some role in the regulation of themselves or other genes. Identification of these conserved regions is also important in the cloning of homologues from more divergent species, as it allows more specific PCR primers to be designed or DNA probes with greater cross-species reactivity to be created.

3.1.2.2. Cloning of the *ptc* gene homologues.

Since the cloning of *D. melanogaster ptc*, orthologues have been cloned from a number of other species. Using low stringency hybridisation, *ptc* has been cloned from a number of insect species; *Drosophila virilis* (Forbes, 1995), *Anopheles gambiae* (mosquito), *Precis coenia* (buckeye butterfly) and *Tribolium castaneum* (red flour beetle) (Goodrich *et al*, in prep).

128

Comparison of partial *ptc* sequence from the mosquito, butterfly and the beetle

(Goodrich *et al*, in prep) has identified regions of conservation and facilitated the design

of degenerate PCR primers which have recently been used in the cloning of the mouse

ptc homologue (Goodrich *et al.*, 1996). These same PCR primers have now been used

to clone the *ptc* homologues from both chick (Marigo *et al*, 1996), and zebrafish

(Concordet *et al.*, 1996).

As part of the *Caenorhabditis elegans* genome sequencing project, several predicted

transcription units with limited homology to *ptc* have been reported (Wilson *et al.*,

1994). The predicted nematode proteins show little homology with the other Ptc

proteins at the nucleotide or amino acid level, although hydropathy plots predict an

overall similarity in shape.

3.1.3. The degree of similarity between the *ptc* gene homologues.

The *ptc* homologues from *D. melanogaster* and *D. virilis,* show an overall similarity of

78% at the DNA level (Forbes, 1995), although there are two regions of sequence

divergence, one in the middle, and one at the 3'end (see Figure 10). At the amino acid

level, the level of similarity rises to 80% (or 83% including conservative substitutions).

There are, however, regions of the protein that exhibit much more similarity than others,

such as the two extracellular loops and the putative transmembrane domains. The *D.*

melanogaster protein has eight putative N-glycosylation sites (Asparagine-Any-Serine/Threonine). In comparison the *D. virilis* protein has ten N-glycosylation sites, although only seven of them are in the same position in the two proteins. In addition, the two proteins share eight conserved cysteine residues, which are important in the stabilization of protein 3D structure.

When this analysis of similarity was extended to the more recently cloned vertebrate homologues, the comparisons of the amino acid sequences showed a drop in the overall levels of similarity between all of the *ptc* proteins (see Table 1), with the greatest conservation being across the transmembrane domains. The predicted hydropathies suggest very similar topologies for all the proteins, with there being two extracellular loops hinged by a smaller intracellular loop.

The mouse *ptc* gene was the first vertebrate orthologue to be cloned (Goodrich *et al.*, 1996). The predicted single ORF encodes a putative protein of 1434 amino acids. The protein has an overall identity of 33% with the *D. melanogaster* protein, with only three of the potential N-glycosylation sites but eight cysteine residues being conserved. By using the method of Kyte and Doolittle (1982) for predicting hydropathy, the mouse protein was shown to be very similar in structure to the *D. melanogaster* protein.

The chick orthologue is predicted to encode a protein of 1442 amino acids (Marigo *et al.*, 1996) which is 86.2% identical to the mouse protein. This similarity drops off at the

130

two termini, being 70.6% at the amino terminus, and 67.7% at the carboxy terminus, but rises to 90% in the middle region. The chick *ptc* orthologue has an overall identity to the *Drosophila* protein of 33.4%, which is remarkably similar to that of the mouse protein.

The surprise came from the cloning of the zebrafish *ptc*. Concordet *et al* (1996) have cloned two *ptc* homologues. They have characterised the homologue called *ptc1* and have shown it to be 64% identical to both the mouse and the chick proteins, and 39% identical to the *D. melanogaster* protein. The single ORF encodes a protein of 1220 amino acids, which is considerably shorter than either of the other vertebrate homologues. Only one of the putative N-glycosylation sites seen in *D. melanogaster* is conserved in the fish, but all of the eight cysteine residues that are conserved between *Drosophila* and the mouse are also conserved in the chick and the fish.

3.1.4. Aims.

At the start of this PhD, *ptc* had only been cloned from *D. melanogaster* and *D. virilis*. Due to the generally high degree of similarity between the orthologues from these two species, it was reasoned that it would be necessary to clone the orthologues from species that are more distantly related to more accurately identify conserved domains. Hence, the primary aim of this project was to clone and characterise both the coding, and the regulatory regions of the *ptc* homologue(s) from the house fly, *Musca domestica*.

In characterising the *M. domestica ptc* homologue, *mdptc*, the conserved regions identified by comparison of the two *Drosophila ptc* seqences could be further dissected, not only to investigate the evolution of *ptc* within the Diptera, but in an attempt to design better strategies for cloning the vertebrate *ptc* homologues. The original screen for the *Musca domestica* homologues of *ptc* (*mdptc*) was performed in Dr Phil Ingham's laboratory by Dr Üwe Strahle, who isolated the two clones, pmusca2.6 and pmusca2.8, in order to determine identify conserved sequences that would be of use in the cloning of *ptc* homologues from the vertebrates.

```
                   1                                                   50
D. virilis         ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~MDRDs LpRvpdTHGd vVDeklfSDL
D. melanogaster    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~MDRDs LpRvpdTHGd vVDeklfSDL
M. domestica       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human              ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mouse              ~~~~~~~~~~ ~~~~~mAsAG naagalgRqa gggrrrrtGg phraapdrDy
chick              masaadalep esgsstAggG shpvraaRsa rgRrrrsgGt rraaapdrey
fish               ~~~~~~~~~~ ~~~~~~~mAs dprdpgpagg vfgdlppsyt rspppvnSDL

                   51                    1/2                     100
D. virilis         yIRtSwvDA. .QvALdQIdK GKARGnRtai yLRsVFQSHL etLGsSvQKH
D. melanogaster    yIRtSwvDA. .QvALdQIdK GKARGsRtai yLRsVFQSHL etLGsSvQKH
M. domestica       ~~~~~~~~~~ ~~~~~~~~~~ GKARGnRtsi yLRsVFQSHf esLGsSvQKH
human              ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mouse              LhRPSYCDAA ftFALeQIsK GKAtGrkaPl wLRakFQrlL fkLGcyiQKn
chick              LqRPSYCDAA ..FALeQIaK GrAtGrRaPl wLRakFQrlL fnLGcyiQKn
fish               LrRPSYChAA ..FALkQIsK GKAvGqkaPl wiRarFQafL fsLGchiQrH

                   101_____                     150
D. virilis         aGKVLFVaiL VLrAFCVGLK S..AQIhsKV hQLWiqEGGR lesELAYTQK
D. melanogaster    aGKVLFVaiL VLStFCVGLK S..AQIhsKV hQLWiqEGGg leaELAYTQK
M. domestica       aGKVLFVaiL VLStFCVGLK S..vQIhsKV hQLWiqEGGR letELAYTQK
human              ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mouse              cGKfLvVglL ifgAFaVGLK aftAnletnV eeLWvevGGR vsRELnYTrq
chick              cGKfLvVglL .ySAFaVGLr a..AnletnV eeLWvevGGR vsRELnYTrq
fish               cGKVLFiglL VfgAlsVGLr ..vAaIetdi ekLWveaGsR vskELrYTke

                   151                                             200
D. virilis         tIGEdEssTh QLiIQTahdp nAsVLhpqAL L..sHLElv kAtaVkVhMY
D. melanogaster    tIGEdEsaTh QLlIQTthdp nAsVLhpqAL L..aHLElv kAtaVkVhlY
M. domestica       tIGEsEssTh QLlIQTghdp nAsVLhpqAL L..THLElk kAtaVkihMf
human              ~~~~~~mFnp QLmIQTPkee GANVLTtEAL LQ..HLdsAL qASrVhVyMY
mouse              kIGEEamFnp QLmIQTPkee GANVLTtEAL LQfTHLdsAL qASrVhVyMY
chick              kIGEEamFnp QLmIQTPqed GtNVLTtEAL rQ..HLdsAL qASrVhVyMY
fish               kqGEEsvFTs QmlIQTPkqe GtNiLTqEAL L..lHLEaAL sASkVqVslY

                   201                                             250
D. virilis         dte..WgLrd mCnspttPsf EGhYYIEQIl khLiPCsIIT PLDCFWEGsq
D. melanogaster    dte..WgLrd mCnmpstPsf EGiYYIEQIl rhLiPCsIIT PLDCFWEGsq
M. domestica       dtd..WsLrd mCnspttPsf EGpYYIEQIl khLiPssIIT PLDCFWEGsq
human              nrQftWKLeh lCyKSGelIt EtG.YmdQII EyLyPClIIT PLDCFW..EG
mouse              nrQ..WKLeh lCyKSGelIt EtG.YmdQII EyLyPClIIT PLDCFWftEG
chick              nrQ..WKLeh lCyKSGelIt EaG.YmdQII EyLyPClIIT PLDCFW..EG
fish               gk..sWdLnk iCfKSGvPIi Env.mIErmI dkLfPCmIvT PLDCFW..EG

                   251        2/3                                  300
D. virilis         llgpetpvYi PGlnq.RLmW stLnPsavMq fmKqQMSDQK IsfdfDtvEq
D. melanogaster    llgpesavvi PGlnq.RLlW TtLnPasvMq ymKqkMSeeK IsfdfetvEq
M. domestica       llgpdfpvqi Psmde.RiTW stLnPsklMq smKKQMSDQt IpfdfDtiEq
human              aKLQSGTAYL lGKPPLRfTW TNfDPLEFlE ELKK...... INyqvDsWEE
mouse              aKLQSGTAYL lGKPPLR..W TNfDPLEFlE ELKK...... INyqvDsWEE
chick              aKLQSGTAYL lGKPPLq..W iNfDPLEFlE ELKK...... INyqvesWEE
fish               sKLQgGsAYL PGmPdiq..W mNLDPLklME ELsq...... ft.slegfrE

                   301                                             350
D. virilis         ymkrAaistG YMekP..CLN PqhPlCPdTA PNKNSrQ..P pDVgaiLSGG
D. melanogaster    ymkrAaiasG YMekP..CLN PlnPnCPdTA PNKNSTQ..P pDVgaiLSGG
M. domestica       ymkrAaissG YtekP..CLd PknlqCPeTA PNKNSdf..P LDVggiLtGG
human              MLnKAeVGHG YMdRP..CLN PADPDCPaTA PNKNSTfTKP LDmAlvLnGG
mouse              MLnKAeVGHG YMdRPftCLN PADPDCPaTA PNKNS..TKP LDVAlvLnGG
chick              MLnKAeVGHG YMdRP..CLN PADPDCPiTA PNKNS..TKP LDVAlvLSGG
fish               MLdKAqVGHa YMnRP..CLd PsDtDCPhsA PNKdpwQvp. .niAaeLqGG
```

```
              351                                                    400
D. virilis    CYGyaaKhMH  WpEqLIVGGa  qrNrsGhLKK  ..AkALQsvv  QLMTeKeMY.
D. melanogaster CYGyaaKhMH  WpEELIVGGa  krNrsGhLrK  ..AqALQsvv  QLMTeKeMY.
M. domestica  CYGfaaKYMH  WpEELIVGGv  qrNrtGhLkr  ..AkgiQTvv  QLMTeKelf.
human         ChGLSrKYMH  WQEELIVGGT  VKNstGkL..  VSAhALQTMF  QLMTPKQMYf
mouse         CqGLSrKYMH  WQEELIVGGT  VKNatGkLft  VSAhALQTMF  QLMTPKQMY.
chick         CYGLSrKYMH  WQEELiGGT   VKNssGkL..  VSAqALQTMF  QLMTPKQMY.
fish          ChGfSkKfMH  WQEELIlGer  VKdsqnaLq.  .SAeALQTMF  lLMsPKQlY.

              401                                                    450
D. virilis    .dqwqdhYkV  HHIgWtqqKA  AevLnAWQRN  FsreVeQllr  KQsriaanYD
D. melanogaster .dqwqdnYkV  HHlgWtqeKA  AevLnAWQRN  FsreVeQllr  KQsriatnYD
M. domestica  .dfwnenYkV  HHIsWtpeKA  AevLtAWQRN  FskeVenimt  gas.mskKYn
human         tEHFKgYeyV  sHINWNEDKA  AAILEAWQRt  yVEVVHQSVA  ..qNSTQK..
mouse         .EHFrgYdyV  sHINWNEDrA  AAILEAWQRt  yVEVVHQSVA  pftNSTQK..
chick         .EHFKgYeyV  sHINWNEDKA  AAILEAWQRm  yVEVVHQSVA  ..qNSTQK..
fish          .EHFKddYei  HdINWNEDKA  tAILEsWQRk  FVEVVHgSip  ..qNSssn..

              451                                                    500
D. virilis    iYVFSsaTLD  DIL..AkFSh  pSavsiviGv  a.tvlYAfcT  lLRWRDpVrg
D. melanogaster iYVFSsaaLD  DIL..AkFSh  pSalsiviGv  avtvlYAfcT  lLRWRDpVrg
M. domestica  VYVFSsaTLD  DIL..ekFSn  pkplsiliGv  iatvsYAfcT  liRWRDpVKg
human         VlsFtTTTLD  DILKftSFSD  VSIRVAsGY   LLMLAYAClT  MLRW.DCsKS
mouse         VlpFtTTTLD  DILK..SFSD  VSVIRVAsGY  LLMLAYAClT  MLRW.DCsKS
chick         VlsFtTTTLD  DILK..SFSD  VSVIRVAsGY  LLMLAYAClT  MLRW.DCaKS
fish          VYaFSTTTLn  DImK..SFSD  VnVIRVAgGY  LLMLAYACvT  MLRW.DCaKS

              501                                                    550
D. virilis    QssVGvA..G  VLLmcfStAA  GLGLCaLL..  GIvFNAAsTQ  VvPFLALGlG
D. melanogaster QssVGvA..G  VLLmcfStAA  GLGLsaLL..  GIvFNAAsTQ  VvPFLALGlG
M. domestica  QssVGvA..G  VLLigfStAA  GLGLCaiL..  GIvFNAAsTQ  VvPFLALGlG
human         QGAVGLA..G  VLLVALSVAA  GLGLCSLift  GISFNAATTQ  VLPFLALGvG
mouse         QGAVGLAftG  VLLVALSVAA  GLGLCSLi..  GISFNAATTQ  VLPFLALGvG
chick         QGAVGLA..G  VLLVALSVAA  GLGLCSLi..  GISFNAATTQ  VLPFLALGvG
fish          QGAVGLA..G  VLLVALSVAA  GLGLCSLL..  GlSFNAATTQ  VLPsLALGiG

              551                                                    600
D. virilis    VDhiFmLtaA  yaEsnr....  ....kEqTkl  iLKkvGpSi.  .LFsacStag
D. melanogaster VDhiFmLtaA  yaEsnr....  ....rEqTkl  iLKkvGpSi.  .LFsacStag
M. domestica  VDhiFmLtsA  yaEsnr....  ....kEqTky  iLKkvGpSi.  .LFsScSttg
human         VDDvFLLaHA  FsETgqNKRI  ..PFedRTGE  CLKRTGaSVA  LftTSiSNVt
mouse         VDDvFLLaHA  FsETgqNKRI  ftPFedRTGE  CLKRTGaSVA  L..TSiSNVt
chick         VDDvFLLaHA  FsETgqNKRI  ..PFedRTGE  CLKRTGaSVA  L..TSiSNVt
fish          VDDmFLLgHs  FtETrsN...  .iPFkERTGd  CLrRTGtSVA  L..TSvnNmi

              601                                                    650
D. virilis    sFFaAvfIPv  PALkvFcLQA  AiVmcFNlAa  ALL..vFPAm  iSLDLrRRta
D. melanogaster sFFaAAfIPv  PALkvFcLQA  AiVmcsNlAa  ALL..vFPAm  iSLDLrRRta
M. domestica  AFFaAAfIPa  PALkvFcLQA  giVmcFNlAa  ALL..vFPAm  iSLDLrRRta
human         AFFMAALIPI  PALRAFSLQA  AVVVVFNFAM  vLL..IFPAI  LSmDLyRRED
mouse         AFFMAALIPI  PALRAFSLQA  AVVVVFNFAM  vLLftIFPAI  LSmDLyRRED
chick         AFFMAALIPI  PALRAFSLQA  AVVVVFNFAM  vLL..IFPAI  LSmDLyRRED
fish          AFFMAALvPI  PALRAFSLQA  AVVVVFNFAM  ALL..IFPAI  LSLDLhRRED

              651                                                    700
D. virilis    gRaDIFc.CC  F..Piwkeqp  k..spqlamN  NngmRggrhl  kncnnnRTaq
D. melanogaster gRaDIFc.CC  F..Pvwkeqp  kVaPpvlplN  NnNgRgarhP  kscnnnRva.
M. domestica  gRaDIsc.CC  F..Piweeqp  vnpfQqiqnN  Nplkeftrqf  krvSdhtkes
human         rRLDftIFCC  FTSPCvSRVI  QVEPQAytdt  hDNTRYSpPP  PYsSHS..FA
mouse         rRLD..IFCC  FTSPCvSRVI  QVEPQAytep  hsNTRYSpPP  PYtSHSfTFA
chick         rRLD..IFCC  FTSPCvtRVI  QiEPQAyae.  NDNicYSsPP  PYsSHS..FA
fish          kRLD..IlCC  FySPCsSRVI  QiqPQelsda  NDNhqrapat  PtytgS.Tit
```

```
                         701                                                     750
      D. virilis         lqqQqqqQpa  qnpllE....  ..........  ..........  ..........
      D. melanogaster    .......lpa  qnpllE....  ..........  ..........  ..........
      M. domestica       lEkrrptspT  Vnagrq....  ..........  ..........  ..........
      human              HETQITMQsT  VQLRTEYftD  PHThVYYTTA  EPRSEISVQP  .....VTVTQ
      mouse              HEThITMQsT  VQLRTEY..D  PHThVYYTTA  EPRSEISVQP  .....VTVTQ
      chick              HETQITMQsT  VQLRTEY..D  PHTqaYYTTA  EPRSEISVQP  .....VTVTQ
      fish               tsThIT..tT  VQafTq..cD  aagqhivTil  pPtSqISttP  psmvlsTpTp


                         751                                                     800
      D. virilis         .......Prt  aggGvgssLp  a.........  .........f  sLarFAyryY
      D. melanogaster    .......qra  dipGSshsLa  S.........  .........f  sLatFAfqHY
      M. domestica       .......eEa  llSaaeknLp  S.........  .........f  sLakFAvKyY
      human              DTLSCQSPES  TSS..TRDLL  SQFSDSSLHC  LEPPCftTKW  tLSSFAEKHY
      mouse              DnLSCQSPES  TSStfTRDLL  SQFSDSSLHC  LEPPC..TKW  tLSSFAEKHY
      chick              DsLSCQSPES  aSS..TRDLL  SQFSDSSvHC  LEPPC..TKW  tLStFAEKHY
      fish               tTdpygSqvf  TtSsSTRDLL  aQveepkegr  ecvPlpffrW  nLSSFArekY


                         801         _____                        850
      D. virilis         tPFLmrswvK  flavmgFLca  vifSLY..ea  TalqDGLDii  DlVPkps..N
      D. melanogaster    tPFLmrswvK  fltvmgFLaa  LisSLY..As  TRlqDGLDii  DlVPkDs..N
      M. domestica       tPFLmKswiK  fiVItsFvGt  vifSLY..As  TklqDGLDLi  DlVPkDT..N
      human              APFLLKPkAK  vVVIfLFLGL  LgvSLYG..T  TRVrDGLDLT  DIVPReTRft
      mouse              APFLLKPkAK  vVVIlLFLGL  LgvSLYGftT  TRVrDGLDLT  DIVPReTR..
      chick              APFLLKPkAK  vVVIfLFLGL  LglSLYG..T  TRVrDGLDLT  DIVPRDTR..
      fish               APlLLKPetK  tVVvvvFvaL  LslSLYG..T  TmVhDGLyLT  DIVPRDTq..


                         851                                                     900
      D. virilis         EhkFldAQtr  lFgFYsMYaV  TQgnFeYPtq  QqLvcDyHea  FV..rVphVI
      D. melanogaster    EhkFldAQtr  lFgFYsMYaV  TQgnFeYPtq  QqLLrDyHdS  FV..rVphVI
      M. domestica       EYkFlnAQts  mFgFYsMYaV  TQgnFeYPNn  QrLLheyHea  FV..rVphVI
      human              EYdFIAAQFK  YFSFYNMYiV  TQka.DYPNi  QHLLYDLHrS  F..SnVkYVm
      mouse              EYdFIAAQFK  YFSFYNMYiV  TQka.DYPNi  QHLLYDLHkS  FftSnVkYVm
      chick              EYdFIAAQFK  YFSFYNMYiV  TQka.DYPNv  QHLLYeLHrS  F..SnVtYVl
      fish               EYeFItAQFK  YFSFYNMYlV  TmdgFDYars  QrqLlqLHna  F..nsVkYVv


                         901                                                     950
      D. virilis         KnDNGgLPDf  WLLL..FRDW  LsnLQriFDe  evrdGrlTke  NWYpNaSsDr
      D. melanogaster    KnDNGgLPDf  WLLL..FseW  LgnLQkiFDe  eyrdGrlTke  cWfpNaSsDa
      M. domestica       KnDNGgLPDf  WLsL..FRD~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
      human              leeNkqLPkM  WLftHYFRDW  LQGLQDAFDS  DWETGkImpn  N.YKNGSDDG
      mouse              leeNkqLPqM  WL..HYFRDW  LQGLQDAFDS  DWETGrImpn  N.YKNGSDDG
      chick              legdrqLPkM  WL..HYFRDW  LQGLQDAFDS  DWETGkITys  N.YKNGSDDa
      fish               KdgNhkLPrM  WL..HYFqDW  LkGLQatFDa  DWEaGkITyd  s.YrNGteDG


                         951                                                    1000
      D. virilis         iLAYK..LiV  QTGyvDnPvD  knlv..eTnR  LVnseGIINP  kAFYnYLSAW
      D. melanogaster    iLAYK..LiV  QTGhvDnPvD  kelv..lTnR  LVnsDGIINq  rAFYnYLSAW
      M. domestica       ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
      human              VLAYK..LLV  QTGSRDKPID  ISQLTKfTQR  LVDADGIINP  SAFYIYLTAW
      mouse              VLAYKftLLV  QTGSRDKPID  ISQLTK..QR  LVDADGIINP  SAFYIYLTAW
      chick              VLAYK..LLV  QTGnRaKPID  ISQLTK..QR  LVDADGIINP  nAFYIYLTAW
      fish               aLAYKp..Li  QTGSkkePfn  ySQLT..srR  LVDgDGlIpP  evFYIYLTvW


                         1001        3/4                                         1050
      D. virilis         atNDvfAYgA  SQgklyPePR  qyyhapneyD  ........lK  IPkslllvYA
      D. melanogaster    atNDvfAYgA  SQgklyPePR  qyfhqpneyD  ........lK  IPkslPlvYA
      M. domestica       ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~  ~~~~~~~~~~
      human              VSNDPVAYAA  SQANIrPH..  RPEWVHDKaD  YMPETRLRIf  tPAAEPiEYA
      mouse              VSNDPVAYAA  SQANIrPHft  RPEWVHDKaD  YMPETRLRI.  .PAAEPiEYA
      chick              VSNDPVAYAA  SQANIrPH..  RPEWVHDKaD  YMPETRLRI.  .PAAEPiEYA
      fish               VSNDPlgYAA  SQANfyPHPR  ..EWiHDKyD  ttgE.nLRI.  .PAAEPlEfA
```

135

```
                       1051                                                1100
D. virilis             QmPFYLhGLt DTSeiktlIg hiRdlsvkYe gF..GLpnYP sGiPFiFWEQ
D. melanogaster        QmPFYLhGLt DTSqiktlIg hiRdlsvkYe gF..GLpnYP sGiPFiFWEQ
M. domestica           ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human                  QFPFYLNGLR DTSDFVEAIE KVRtICsNYT S..LGLssYP NGYPFLFWEQ
mouse                  QFPFYLNGLR DTSDFVEAIE KVRvICnNYT SFtLGLssYP NGYPFLFWEQ
chick                  QFPFYLNGLR eTSDFVEAIE KVRaICnNYT S..LGiasYP NGYPFLFWEQ
fish                   QFPFYLNGLR qaSDFiEAIE sVRtICeef. .mrqGiknYP NGYPFLFWEQ

                       1101                                                1150
D. virilis             YmT..LRssL aLilacaLla alvlvsLlLL svWaA..vlV iftvLAslaq
D. melanogaster        YmT..LRssL amilacVLla alvlvsLlLL svWaA..vlV iLsvLAslaq
M. domestica           ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human                  YIftGLRHWL LLfISVVLAC TFLVCAvfLL NPWTAGIIVm vLAL..MTVE
mouse                  YI..sLRHWL LLSISVVLAC TFLVCAvfLL NPWTAGIIVm vLALftMTVE
chick                  YI..GLRHWL LLSISVVLAC TFLVCALfLL NPWTAGIIVV vLAL..MTVE
fish                   YI..GLRHWf LLSISVVLAC TFLVCAilLL NPWTAGvIVf iLpm..MTVE

                       1151                                                1200
D. virilis             iFGaMtLlGI KLSAi..PaV ILIlgVGmiV cFnVHIsLgF mTsvGnrqRR
D. melanogaster        iFGaMtLlGI KLSAi..PaV ILIlSVGmml cFnVlISLgF mTsvGnrqRR
M. domestica           ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human                  LFGMMGLIGI KLSAvftPVV ILIASVGIGV EFTVHvALaF LTAIsDkNRR
mouse                  LFGMMGLIGI KLSAv..PVV ILIASVGIGV EFTVHvALaF LTAIGDkNhR
chick                  LFGMMGLIGI KLSAv..PVV ILIASVGIGV EFTVHIALaF LTAIGDkNRR
fish                   LFGiMGLIGI KLSAi..PVV ILIASVGIGV EFTVHIALgF LTAIGDrNtR

                       1201                                                1250
D. virilis             vhLAmq..ls lgPlvhGmlt sgmaVfMLST ..SpFeFviR hFcwlLllvl
D. melanogaster        vqLsmq..ms lgPlvhGmlt sgvaVfMLST ..SpFeFviR hFcwlLlvvl
M. domestica           ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human                  AVLAlEH..M FAPVLDGAVS TLLGVLMLfT AGSdFDFIVR YFFAVLAILT
mouse                  AmLAlEHftM FAPVLDGAVS TLLGVLML.. AGSEFDFIVR YFFAVLAILT
chick                  AVLAlEH..M FAPVLDGAVS TLLGVLML.. AGSEFDFIVR YFFAVLAILT
fish                   savAmEH..M FAPViDGAiS TLLGVLML.. AGSEFDFImR YFFAVLAILT

                       1251                                                1300
D. virilis             cLGacNsLlv fPiLLSmlGP ..EaElvPle hpdRisTPSP lPMRsskRan
D. melanogaster        cvGacNsLlv fPiLLSmvGP ..EaElvPle hpdRisTPSP lPvRsskRsg
M. domestica           ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human                  iLGVLNGLVL LPVLwSFFGP Y..pEVsPAN GlNRLPTPSP Eft.PPP...
mouse                  vLGVLNGLVL LPVLLSFFGP ftcpEVsPAN GlNRLPTPSP E...PPP...
chick                  iLGVLNGLVL LPVLLSFFGP Y..pEVsPAc GrNRLPTPSP E...PPP...
fish                   lLGiLNGLVL LPVLLSlmGP ..paEVvPAN naNhLqsPSP EPM.PPPmnh

                       1301                                                1350
D. virilis             KSfVvngSrg ssrsSGSnnc hkahhyhkdV nindpsLtTI teEpQswkss
D. melanogaster        KSyVvqgSr. ssrgS....c qkshhhhhk. dlndpsLtTI teEpQswkss
M. domestica           ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human                  .SvVRFAmPP GhThSGSDSS DSEYSSQTTV SGlSEELr.. hYEAQQgaGg
mouse                  .SvVRFAvPP GhTNnGSDSS DSEYSSQTTV SGiSEELrft qYEAQQgaGg
chick                  .SiVRFAlPP GhTNnGSDSS DSEYSSQTTV SGiSEELh.. hYEAtQspGi
fish                   hgyyaghiPk ashqafSetS DSEYxSeT~~ ~~~~~~~~~~ ~~~~~~~~~~

                       1351                                                1400
D. virilis             nssiqmhndW sapptpppthS hhpHhlnnHy mnninnyQqr nereppHgaa
D. melanogaster        nssiqmpndW Tyqpreqrpa syaaPppay. .......... ......Hkaa
M. domestica           ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human                  PaHQViVEAf TTENPVFAhS TVVHPESRHh PPSnPkQQPH LdSGSLpPgr
mouse                  PaHQViVEA. .TENPVFArS TVVHPdSRHQ PPltPRQQPH LdSGSLsPgr
chick                  PvHQVvVEA. .TENPVFArS TVVqPESRHQ ..SsPRlQsn peaGtqqvwh
fish                   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

```
                        1401                                                    1450
D. virilis         sssncgyagp PPsyhkaaqP mpggygppel qsivvqpevt vetthSdsnt
D. melanogaster    aqqhhQhqgp Ptt...pPPP fptay.ppel qsivvqpevt vetthSdsnt
M. domestica       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human              Q..GQQPRrd PPRkGLwPPl YRftPRRDAF EISTEGHSGP SNraRwgprg
mouse              QftGQQPRrd PPREGLRPPP YR..PRRDAF EISTEGHSGP SNrdRSgprg
chick              Q..GrQPkqe .vREGLRPPP YR..PRRDAF EISTEGHSGP SNkdRlnhk.
fish               ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

                        1451                                                    1500
D. virilis         tkvtataNik velvtPGraV rSYnfts~~~ ~~~~~~~~~~ ~~~~~~~~~~
D. melanogaster    tkvtataNik velamPGraV rSYnfts~~~ ~~~~~~~~~~ ~~~~~~~~~~
M. domestica       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human              ArSHNpRNPt stAM..GsSV pgYCQPITTV TASASftVTV AVHPPpvPGP
mouse              ArSHNpRNPt stAMftGsSV pSYCQPITTV TASAS..VTV AVHPP..PGP
chick              AhSHNmRsPa fgAMgvpgS. .aYCQPITTV TASAS..VTV AVHPavh...
fish               ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

                        1501                                                    1550
D. virilis         ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
D. melanogaster    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
M. domestica       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human              GRNPRGGlCP G...YPETDH GlFEDPHVP. .FHVRCERRD SKVEVIELQD
mouse              GRNPRGGpCP GyEsYPETDH GvFEDPHVPf tFHVRCERRD SKVEVIELQD
chick              shNscrGsfP scEeYnEdDr GmFEDPHVP. .FnVRCERRn SKVEVIELQD
fish               ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

                        1551                                                    1600
D. virilis         ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
D. melanogaster    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
M. domestica       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
human              VftECEERFr GSSSN~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
mouse              V..ECEERPw GSSSN~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
chick              V..ECEERta GkiSexgwxl kgrggqrlkt vfpgtawkrt csdigrkgra
fish               ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

                        1601       1617
D. virilis         ~~~~~~~~~~ ~~~~~~~
D. melanogaster    ~~~~~~~~~~ ~~~~~~~
M. domestica       ~~~~~~~~~~ ~~~~~~~
human              ~~~~~~~~~~ ~~~~~~~
mouse              ~~~~~~~~~~ ~~~~~~~
chick              cikdnkvtgy wxphwyy
fish               ~~~~~~~~~~ ~~~~~~~
```

Figure 10. Alignment of the Ptc protein sequences.

This figure shows the alignment of the available Ptc protein sequences, indicating conserved motifs. The cysteine residues that are conserved between all of the species are highlighted. The putative N-glycosylation sites in the *D. melanogaster* sequence are in bold. The black lines above the sequence alignment indicate the position of the transmembrane domains as predicted by hydrophobicity plots. The numbers in bold above the sequence show the positions of the intron/exon boundaries deduced from the *Drosophila* sequence. The alignment and consensus were generated using the GCG suite of programs.

Table 1. Degree of similarity between the Patched protein sequences from various species.

The numbers indicate percentage identity, or, in parentheses, similarity, of the Ptc protein sequences from the mouse, chick, zebrafish, and *Drosophila melanogaster*.

species	D. melanogaster	Zebrafish	Chick	Mouse
Mouse	39 (59)	64 (79)	89 (94)	100 (100)
Chick	40 (60)	64 (79)	100 (100)	
Zebrafish	39 (62)	100 (100)		
D. melanogaster	100 (100)			

3.2. Results.

3.2.1. Library screen for the *M. domestica ptc* homologue.

Approximately 200,000 λphage of a *Musca domestica* genomic DNA library in λDASH

II (made and provided by Daniel Curtis and Javier Apfeld) were screened with a 356 bp

fragment from exon 3 of the *ptc* homologue from *D. virilis*. Assuming the *M. domestica*

genome consists of between approximately 1.2×10^9 (Curtis and Apfeld, personal

communication) and 1.6×10^9 base pairs (Hough-Evans *et al.*, 1980), and an average

insert size of 15Kb (Curtis and Apfeld, personal communication), then 200,000 phage

would represent 1.8 - 2.5 times genome coverage. Two phage clones were isolated, and

were analysed by Southern hybridisation. When digested with *EcoRI*, one of the phage

clones, ptc2-1, generated five fragments of between ≈600 bp and 8 Kb. Of these

fragments, two did not hybridise to the Southern probe, and the three remaining

fragments were subcloned into the *EcoRI* site of pBluescriptKS (Stratagene) for further

analysis.

The two plasmids indicated by the Southern analysis to be of most interest were

designated pmusca2.6 and pmusca2.8. The inserts are ≈400 bp, and ≈6.5 Kb

respectively, and together were found to contain the coding sequence spanning exons 2

and 3 of *mdptc,* as well as several kilobases of intron sequence (see Figures 11 and 12).

Exon 2 Exon 3 Exon 4 Exon 5 Exon 6

D. melanogaster
ptc

M. domestica
ptc

p2.6

Variable

p2.8

~6.5 Kb Intron Sequence
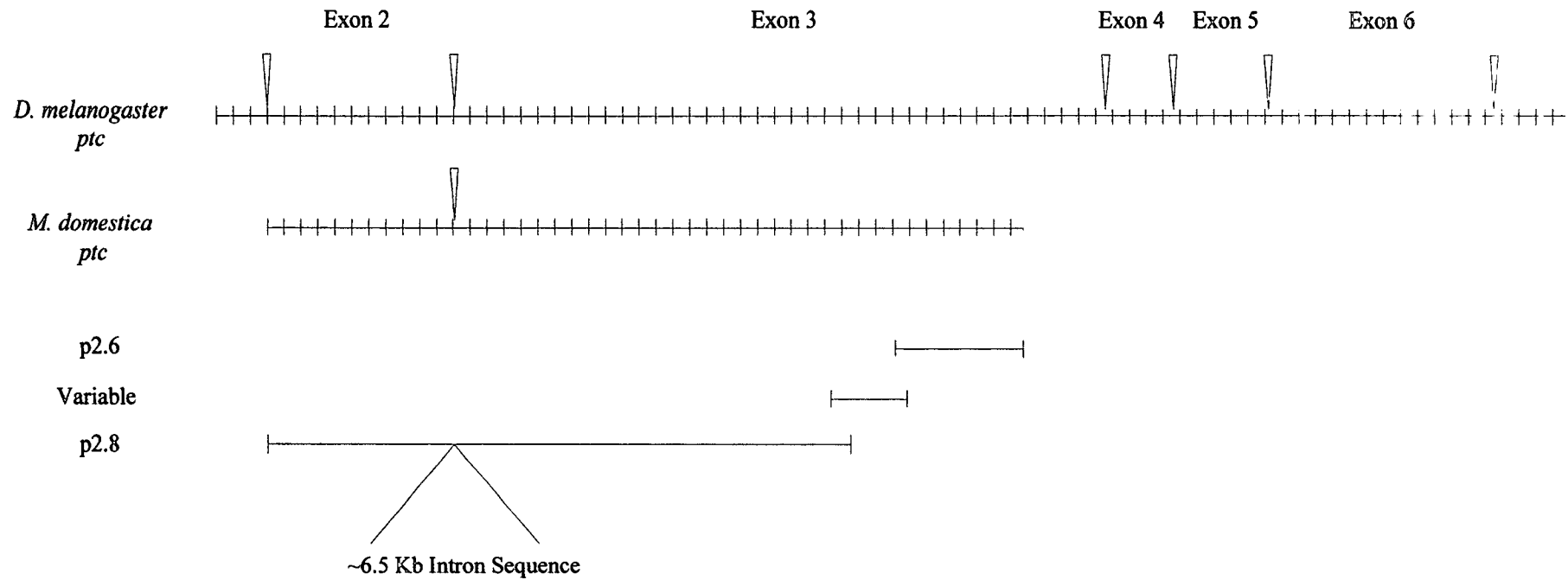
Figure 11. Diagrammatic representation of the genomic structure of *ptc*, and the position of clones.

Genomic structure of the *ptc* genes *from D. melanogaster* and *D. virilis*, showing the relative positioning of the *M .domestica* clones, pmusca2.6, pmusca2.8, and the PCR fragment covering a region that has been shown to be different in *D. melanogaster* and *D. virilis*. The open triangles show the position of the intron/exon boundaries in the *Drosophila ptc* genes.

3.2.2. Sequencing of pmusca2.8.

A restriction map of pmusca2.8 was generated using *BamHI, NotI, SacI, SacII, SmaI, XbaI, HindIII, KpnI, PstI, XhoI, EcoRV,* and *SalI* (see Figure 12). Five new deletion constructs, named pmusca2.8PH, pmusca2.8RJB, pmusca2.8RV, p2.8RVX (see Figure 12), and pmusca2.8XB, were made to aid in the sequencing of this plasmid. On sequencing these plasmids, using both ^{35}S dideoxy sequencing chemistry and fluorescent sequencing (see Chapter two.), it was found that the whole of *mdptc* exon 2, the 5' end of exon 3, the intervening intron, as well as at least 4 Kb of intron 5' to exon 2 are contained within pmusca2.8.

3.2.3. Sequencing of pmusca2.6.

Both strands of the pmusca2.6 insert were sequenced and shown to be coding sequence from *ptc* exon 3. The 3' end of exon 3 contained within pmusca2.6 fell within a region that has been demonstrated to be hypervariable in both *D. melanogaster* and *D. virilis* (Forbes, 1995), as well as the zebrafish and the mouse. Alignment of sequence data from both pmusca2.6 and pmusca2.8 to *ptc* sequence data from both *D. melanogaster* and *D. virilis*, indicated *mdptc* was shorter than both of the *Drosophila* sequences. The position at which the sequence from the two *Musca* plasmids joined is within one of the few regions of sequence in which the two *Drosophila* sequences were significantly divergent. Because of the variability in this region, and that the *Musca* sequence was

shorter than the *Drosophila* sequences it was proposed that there may be a gap in the exon 3 sequence data contained within pmusca2.6 and pmusca2.8.

3.2.4. Cloning of the variable region.

To confirm whether the *Musca* sequence was indeed shorter than the *Drosophila* sequences, or that there was sequence missing from the two *Musca* plasmids, PCR primers were designed to amplify over this putative 'gap' (see appendix D). A 225bp fragment was amplified from *M. domestica* genomic DNA, whose ends were identical to the ends of pmusca2.6 and pmusca2.8, indicating that it was the correct fragment. This fragment also contained 138 bp of sequence not found in either pmusca2.6 or pmusca2.8. As expected, this intervening sequence showed no homology to this region of *D. melanogaster* or *D. virilis*.

A

EcoRV   HindIII

intron   PstI

XhoI   XhoI

intron

BamHI

pmusca2.8

~9.4 Kb

EcoRV/XhoI

EcoRI   EcoRI

T3   T7

B

intron

XhoI   EcoRV

p2.8RVX

~4.5 Kb

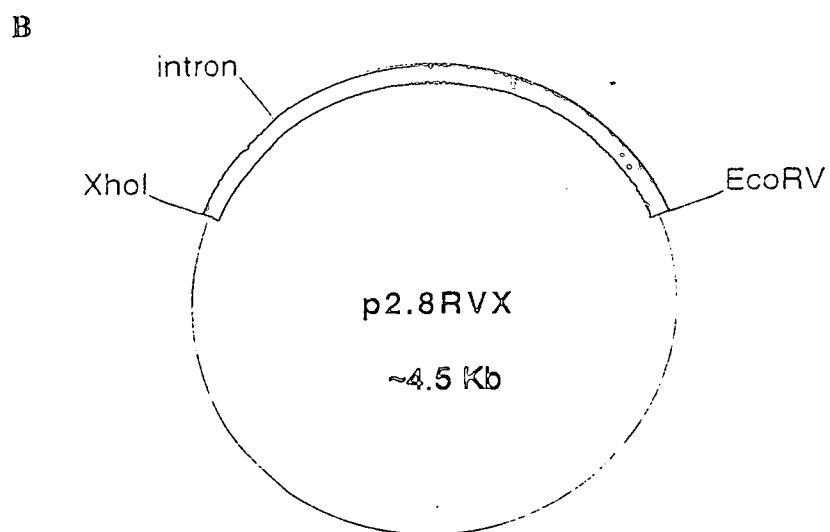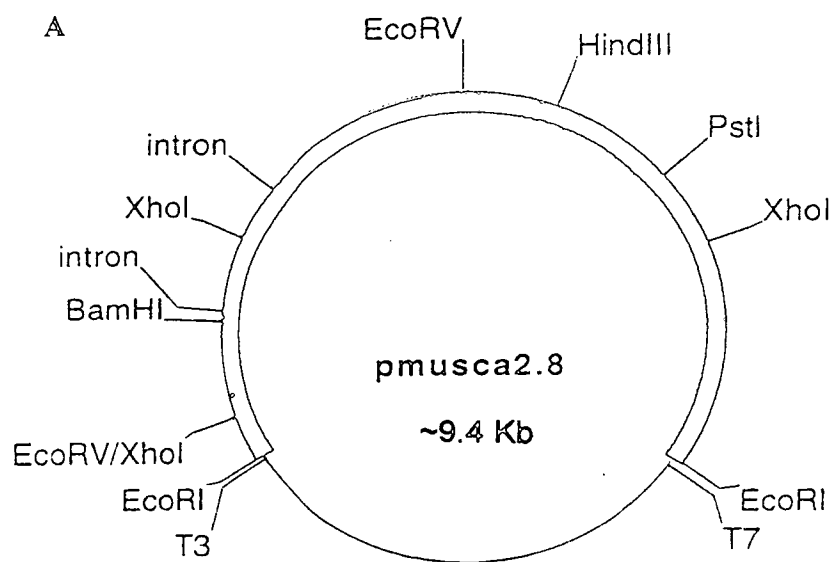*Figure 12.* The pmusca2.8 and p2.8RVX constructs.

(A) pmusca2.8, showing the restriction sites used in the construction of the deletion

constructs used to create RNA *in situ* hybridisation probes (see B), and for sequencing.

(B) p2.8RVX, an approximately 1.5 Kb EcoRV/XhoI fragment cloned into pBluescript

used to create an RNA *in situ* hybridisation probe used to visualise *ptc* expression in

Musca domestica.

3.2.5. *M. domestica* sequence analysis.

3.2.5.1. Exon 2.

Exon 2 of *mdptc* is 555 bp long, encoding 185 amino acids. When aligned to the second exon of the *D. melanogaster ptc* gene it is found to be 85.9% identical and 91.35% similar at the amino acid level (see Figure 13). The second exon encodes the first transmembrane domain, which at the amino acid level is 100% identical between the two species. In *D. melanogaster,* exon 2 also contains three of the cysteines that are conserved between the two *Drosophila* species, mouse, fish and chick. In *M. domestica* only two of these are conserved, the closest to the 5' end having been substituted by a serine.

3.2.5.2. Exon 3.

The third exon of *D. melanogaster* is 1926 bp long, and codes for 642 amino acids. 1677 bp of *mdptc* exon 3 was cloned from *M. domestica*, which suggests that there may be approximately 300 nucleotides missing from the 3' end of the *M. domestica* clone. The 559 amino acids of *mdptc* were aligned to the corresponding residues from *D. melanogaster.* The partial protein sequences were found to be 71.4% identical, and 82.6% similar. The *D. melanogaster* sequence that aligned to the *M. domestica* sequence encodes six transmembrane domains (domains 2-7), and contains another five

of the conserved cysteine residues, all of which are present in the *M. domestica* sequence. The levels of similarity between the individual transmembrane domains is given in Table 2. The degree of identity at the amino acid level with the transmembrane domains of *D. melanogaster* is between 52% for transmembrane domain 6 and 100% for transmembrane domain 1 whereas the levels of identity with those of *D. virilis* are between 52% for transmembrane domain 6 and 96% for transmembrane domain 5. The degree of similarity between the transmembrane domains is generally higher, between 77% and 100% for transmembrane domains 2 and 1 of *D. melanogaster* respectively, and 73% and 100% for transmembrane domains 6 and 5 of *D. virilis* respectively.

When the coding nucleotide sequence of *mdptc* is aligned to the corresponding region of *D. melanogaster ptc*, they show 69.3% similarity. The vast majority of the differences are in the third position of the triplet codons. Most of the amino acids are coded for by more than one triplet codon. This third base 'wobble' allows changes in the nucleotide sequence whilst maintaining the protein sequence due to the redundant nature of the genetic code.

Table 2. Similarity between the transmembrane domains in *mdptc*, and those in *D. melanogaster* and *D. virilis*.

The numbers indicate percentage identity, or, in parentheses, similarity of the transmembrane domains between the *ptc* gene of *M. domestica*, and those in the *ptc* genes of *D. melanogaster*, and *D. virilis*.

Domain	Size (amino acids)	Similarity to *D. melanogaster* (%)	Similarity to *D. virilis* (%)
Transmembrane 1	19	100 (100)	89.5 (89.5)
Transmembrane 2	22	63.6 (77.3)	66.7 (85.7)
Transmembrane 3	45	91.1 (97.8)	93.8 (97.9)
Transmembrane 4	27	85.2 (85.2)	81.5 (81.5)
Transmembrane 5	29	93.1 (96.6)	96.6 (100)
Transmembrane 6	23	52.2 (87)	52.2 (73.9)

```
    GGTAAAGCCCGCGGAAACCGTACATCCATTTACCTACGTTCAGTATTTCAATCACACTTC
  1 ---------+---------+---------+---------+---------+---------+  60
    CCATTTCGGGCGCCTTTGGCATGTAGGTAAATGGATGCAAGTCATAAAGTTAGTGTGAAG

    G  K  A  R  G  N  R  T  S  I  Y  L  R  S  V  F  Q  S  H  F  -


    GAATCTCTGGGCAGCTCGGTGCAAAAGCATGCCGGCAAGGTTCTATTTGTGGCAATTCTA
 61 ---------+---------+---------+---------+---------+---------+ 120
    CTTAGAGACCCGTCGAGCCACGTTTTCGTACGGCCGTTCCAAGATAAACACCGTTAAGAT

    E  S  L  G  S  S  V  Q  K  H  A  G  K  V  L  F  V  A  I  L  -


    GTATTAAGCACATTCTGCGTTGGCCTCAAAAGTGTGCAAATCCATTCGAAGGTCCACCAG
121 ---------+---------+---------+---------+---------+---------+ 180
    CATAATTCGTGTAAGACGCAACCGGAGTTTTCACACGTTTAGGTAAGCTTCCAGGTGGTC

    V  L  S  T  F  C  V  G  L  K  S  V  Q  I  H  S  K  V  H  Q  -


    CTTTGGATACAGGAGGGCGGACGACTCGAGACAGAGCTGGCGTACACACAGAAGACCATA
181 ---------+---------+---------+---------+---------+---------+ 240
    GAAACCTATGTCCTCCCGCCTGCTGAGCTCTGTCTCGACCGCATGTGTGTCTTCTGGTAT

    L  W  I  Q  E  G  G  R  L  E  T  E  L  A  Y  T  Q  K  T  I  -


    GGAGAGTCGGAGAGTTCAACACACCAATTGCTTATACAGACGGGTCACGATCCGAACGCA
241 ---------+---------+---------+---------+---------+---------+ 300
    CCTCTCAGCCTCTCAAGTTGTGTGGTTAACGAATATGTCTGCCCAGTGCTAGGCTTGCGT

    G  E  S  E  S  S  T  H  Q  L  L  I  Q  T  G  H  D  P  N  A  -


    TCGGTTCTTCATCCGCAAGCACTGCTGACGCATCTCGAGGTGCTAAAGAAAGCAACGGCC
301 ---------+---------+---------+---------+---------+---------+ 360
    AGCCAAGAAGTAGGCGTTCGTGACGACTGCGTAGAGCTCCACGATTTCTTTCGTTGCCGG

    S  V  L  H  P  Q  A  L  L  T  H  L  E  V  L  K  K  A  T  A  -


    GTTAAGATACACATGTTCGACACAGATTGGAGTCTTAGGGACATGTGCAATTCGCCAACT
361 ---------+---------+---------+---------+---------+---------+ 420
    CAATTCTATGTGTACAAGCTGTGTCTAACCTCAGAATCCCTGTACACGTTAAGCGGTTGA

    V  K  I  H  M  F  D  T  D  W  S  L  R  D  M  C  N  S  P  T  -


    ACTCCGAGTTTTGAAGGTCCTTATTACATAGAACAGATCCTAAAGCATCTTATTCCGAGT
421 ---------+---------+---------+---------+---------+---------+ 480
    TGAGGCTCAAAACTTCCAGGAATAATGTATCTTGTCTAGGATTTCGTAGAATAAGGCTCA

    T  P  S  F  E  G  P  Y  Y  I  E  Q  I  L  K  H  L  I  P  S  -


    TCCATAATAACGCCATTAGATTGTTTCTGGGAGGGGAGTCAGCTGCTGGGACCAGACTTT
481 ---------+---------+---------+---------+---------+---------+ 540
    AGGTATTATTGCGGTAATCTAACAAAGACCCTCCCCTCAGTCGACGACCCTGGTCTGAAA

    S  I  I  T  P  L  D  C  F  W  E  G  S  Q  L  L  G  P  D  F  -


                2/3
    CCCGTTCAAATTCCGAGTATGGACGAACGCATTACCTGGAGCACACTTAATCCCTCAAAA
541 ---------+---------+---------+---------+---------+---------+ 600
    GGGCAAGTTTAAGGCTCATACCTGCTTGCGTAATGGACCTCGTGTGAATTAGGGAGTTTT

    P  V  Q  I  P  S  M  D  E  R  I  T  W  S  T  L  N  P  S  K  -
```

```
        CTGATGCAGTCCATGAAGAAACAGATGAGTGATCAAACTATACCCTTTGACTTTGACACG
601     ---------+---------+---------+---------+---------+---------+ 660
        GACTACGTCAGGTACTTCTTTGTCTACTCACTAGTTTGATATGGGAAACTGAAACTGTGC

         L  M  Q  S  M  K  K  Q  M  S  D  Q  T  I  P  F  D  F  D  T   -


        ATCGAGCAGTACATGAAGCGTGCGGCAATTTCATCAGGTTACACCGAAAAGCCATGTCTG
661     ---------+---------+---------+---------+---------+---------+ 720
        TAGCTCGTCATGTACTTCGCACGCCGTTAAAGTAGTCCAATGTGGCTTTTCGGTACAGAC

         I  E  Q  Y  M  K  R  A  A  I  S  S  G  Y  T  E  K  P  C  L   -


        GATCCCAAAAACCTTCAGTGTCCCGAAACTGCACCGAACAAAAACAGTGATTTTCCCTTG
721     ---------+---------+---------+---------+---------+---------+ 780
        CTAGGGTTTTTGGAAGTCACAGGGCTTTGACGTGGCTTGTTTTTGTCACTAAAAGGGAAC

         D  P  K  N  L  Q  C  P  E  T  A  P  N  K  N  S  D  F  P  L   -


        GATGTTGGTGGAATACTGACTGGGGGCTGCTACGGATTTGCCGCCAAATACATGCATTGG
781     ---------+---------+---------+---------+---------+---------+ 840
        CTACAACCACCTTATGACTGACCCCCGACGATGCCTAAACGGCGGTTTATGTACGTAACC

         D  V  G  G  I  L  T  G  G  C  Y  G  F  A  A  K  Y  M  H  W   -


        CCGGAGGAGTTGATCGTGGGTGGTGTGCAACGAAATCGCACCGGCCACTTGAAAAGGGCC
841     ---------+---------+---------+---------+---------+---------+ 900
        GGCCTCCTCAACTAGCACCCACCACACGTTGCTTTAGCGTGGCCGGTGAACTTTTCCCGG

         P  E  E  L  I  V  G  G  V  Q  R  N  R  T  G  H  L  K  R  A   -


        AAGGGTATTCAGACTGTGGTGCAGCTGATGACGGAGAAGGAACTGTTTGACTTCTGGAAT
901     ---------+---------+---------+---------+---------+---------+ 960
        TTCCCATAAGTCTGACACCACGTCGACTACTGCCTCTTCCTTGACAAACTGAAGACCTTA

         K  G  I  Q  T  V  V  Q  L  M  T  E  K  E  L  F  D  F  W  N   -


        GAAAACTACAAAGTCCATCACATTAGCTGGACCCCGGAGAAGGCGGCAGAGGTCTTAACT
961     ---------+---------+---------+---------+---------+---------+ 1020
        CTTTTGATGTTTCAGGTAGTGTAATCGACCTGGGGCCTCTTCCGCCGTCTCCAGAATTGA

         E  N  Y  K  V  H  H  I  S  W  T  P  E  K  A  A  E  V  L  T   -


        GCCTGGCAACGAAATTTTTCCAAAGAAGTCGAAAACATTATGACCGGCGCTAGCATGTCC
1021    ---------+---------+---------+---------+---------+---------+ 1080
        CGGACCGTTGCTTTAAAAAGGTTTCTTCAGCTTTTGTAATACTGGCCGCGATCGTACAGG

         A  W  Q  R  N  F  S  K  E  V  E  N  I  M  T  G  A  S  M  S   -


        AAGAAGTACAACGTTTATGTGTTTTCGTCCGCCACTTTGGACGACATTTTAGAGAAGTTC
1081    ---------+---------+---------+---------+---------+---------+ 1140
        TTCTTCATGTTGCAAATACACAAAAGCAGGCGGTGAAACCTGCTGTAAAATCTCTTCAAG

         K  K  Y  N  V  Y  V  F  S  S  A  T  L  D  D  I  L  E  K  F   -


        TCGAACCCCAAACCTTTGAGTATTCTAATTGGAGTTATTGCTACCGTTAGCTATGCATTC
1141    ---------+---------+---------+---------+---------+---------+ 1200
        AGCTTGGGGTTTGGAAACTCATAAGATTAACCTCAATAACGATGGCAATCGATACGTAAG

         S  N  P  K  P  L  S  I  L  I  G  V  I  A  T  V  S  Y  A  F   -


        TGCACACTCATTCGTTGGCGAGACCCAGTGAAGGGACAGTCCAGCGTGGGAGTAGCTGGC
```

```
1201 ---------+---------+---------+---------+---------+---------+ 1260
     ACGTGTGAGTAAGCAACCGCTCTGGGTCACTTCCCTGTCAGGTCGCACCCTCATCGACCG

        C  T  L  I  R  W  R  D  P  V  K  G  Q  S  S  V  G  V  A  G  -
     GTATTATTGATTGGCTTTTCAACAGCTGCGGGATTGGGCTTGTGCGCCATATTGGGCATT
1261 ---------+---------+---------+---------+---------+---------+ 1320
     CATAATAACTAACCGAAAAGTTGTCGACGCCCTAACCCGAACACGCGGTATAACCCGTAA

        V  L  L  I  G  F  S  T  A  A  G  L  G  L  C  A  I  L  G  I  -

     GTCTTCAATGCGGCCAGCACACAAGTTGTGCCCTTTTTGGCATTGGGACTTGGAGTTGAT
1321 ---------+---------+---------+---------+---------+---------+ 1380
     CAGAAGTTACGCCGGTCGTGTGTTCAACACGGGAAAAACCGTAACCCTGAACCTCAACTA

        V  F  N  A  A  S  T  Q  V  V  P  F  L  A  L  G  L  G  V  D  -

     CATATATTCATGCTTACTTCTGCGTATGCAGAAAGCAATCGCAAGGAGCAAACTAAATAT
1381 ---------+---------+---------+---------+---------+---------+ 1440
     GTATATAAGTACGAATGAAGACGCATACGTCTTTCGTTAGCGTTCCTCGTTTGATTTATA

        H  I  F  M  L  T  S  A  Y  A  E  S  N  R  K  E  Q  T  K  Y  -

     ATTCTCAAGAAAGTGGGCCCAAGCATACTGTTTAGCTCTTGTTCTACTACAGGAGCGTTT
1441 ---------+---------+---------+---------+---------+---------+ 1500
     TAAGAGTTCTTTCACCCGGGTTCGTATGACAAATCGAGAACAAGATGATGTCCTCGCAAA

        I  L  K  K  V  G  P  S  I  L  F  S  S  C  S  T  T  G  A  F  -

     TTCGCGGCCGCCTTTATACCGGCGCCTGCCCTGAAAGTTTTCTGCCTTCAGGCGGGAATT
1501 ---------+---------+---------+---------+---------+---------+ 1560
     AAGCGCCGGCGGAAATATGGCCGCGGACGGGACTTTCAAAAGACGGAAGTCCGCCCTTAA

        F  A  A  A  F  I  P  A  P  A  L  K  V  F  C  L  Q  A  G  I  -

     GTTATGTGTTTCAATCTGGCTGCGGCCTTGCTGGTGTTCCCAGCCATGATATCGCTAGAT
1561 ---------+---------+---------+---------+---------+---------+ 1620
     CAATACACAAAGTTAGACCGACGCCGGAACGACCACAAGGGTCGGTACTATAGCGATCTA

        V  M  C  F  N  L  A  A  A  L  L  V  F  P  A  M  I  S  L  D  -

     CTGCGAAGACGGACTGCTGGCCGTGCGGATATTTCGTGTTGTTGTTTTCCCATATGGGAA
1621 ---------+---------+---------+---------+---------+---------+ 1680
     GACGCTTCTGCCTGACGACCGGCACGCCTATAAAGCACAACAACAAAAGGGTATACCCTT

        L  R  R  R  T  A  G  R  A  D  I  S  C  C  C  F  P  I  W  E  -

     GAACAACCAGTCAATCCGTTTCAACAGATCCAAAATAATAATCCCTTGAAGGAATTCACG
1681 ---------+---------+---------+---------+---------+---------+ 1740
     CTTGTTGGTCAGTTAGGCAAAGTTGTCTAGGTTTTATTATTAGGGAACTTCCTTAAGTGC

        E  Q  P  V  N  P  F  Q  Q  I  Q  N  N  N  P  L  K  E  F  T  -

     AGGCAGTTTAAGAGGGTCAGCGACCATACAAAAGAAAGTCTGGAAAAACGGCGACCGACG
1741 ---------+---------+---------+---------+---------+---------+ 1800
     TCCGTCAAATTCTCCCAGTCGCTGGTATGTTTTCTTTCAGACCTTTTTGCCGCTGGCTGC

        R  Q  F  K  R  V  S  D  H  T  K  E  S  L  E  K  R  R  P  T  -
```

```
      TCGCCTACAGTCAACGCTGGAAGACAAGAAGAAGCCTTGCTGTCTGCCGCAGAGAAAAAC
1801  ---------+---------+---------+---------+---------+---------+ 1860
      AGCGGATGTCAGTTGCGACCTTCTGTTCTTCTTCGGAACGACAGACGGCGTCTCTTTTTG

       S  P  T  V  N  A  G  R  Q  E  E  A  L  L  S  A  A  E  K  N   -


      TTACCTTCCTTTTCGCTTGCTAAATTTGCCGTGAAATACTATACGCCCTTCCTCATGAAG
1861  ---------+---------+---------+---------+---------+---------+ 1920
      AATGGAAGGAAAAGCGAACGATTTAAACGGCACTTTATGATATGCGGGAAGGAGTACTTC

       L  P  S  F  S  L  A  K  F  A  V  K  Y  Y  T  P  F  L  M  K   -

      AGCTGGATTAAATTCATTGTTATTACATCGTTTGTGGGTACGGTCATCTTTAGTCTGTAC
1921  ---------+---------+---------+---------+---------+---------+ 1980
      TCGACCTAATTTAAGTAACAATAATGTAGCAAACACCCATGCCAGTAGAAATCAGACATG

       S  W  I  K  F  I  V  I  T  S  F  V  G  T  V  I  F  S  L  Y   -


      GCCTCAACAAAGCTGCAAGATGGTTTGGATTTGATAGATTTGGTGCCCAAGGACACAAAC
1981  ---------+---------+---------+---------+---------+---------+ 2040
      CGGAGTTGTTTCGACGTTCTACCAAACCTAAACTATCTAAACCACGGGTTCCTGTGTTTG

       A  S  T  K  L  Q  D  G  L  D  L  I  D  L  V  P  K  D  T  N   -


      GAATACAAGTTCCTGAATGCTCAAACGTCGATGTTCGGTTTCTACAGCATGTACGCCGTC
2041  ---------+---------+---------+---------+---------+---------+ 2100
      CTTATGTTCAAGGACTTACGAGTTTGCAGCTACAAGCCAAAGATGTCGTACATGCGGCAG

       E  Y  K  F  L  N  A  Q  T  S  M  F  G  F  Y  S  M  Y  A  V   -


      ACGCAGGGAAACTTCGAGTATCCCAACAATCAACGTCTCCTGCACGAATACCACGAAGCC
2101  ---------+---------+---------+---------+---------+---------+ 2160
      TGCGTCCCTTTGAAGCTCATAGGGTTGTTAGTTGCAGAGGACGTGCTTATGGTGCTTCGG

       T  Q  G  N  F  E  Y  P  N  N  Q  R  L  L  H  E  Y  H  E  A   -


      TTTGTTCGTGTACCCCACGTTATTAAGAATGACAATGGTGGCCTACCCGACTTTTGGCTG
2161  ---------+---------+---------+---------+---------+---------+ 2220
      AAACAAGCACATGGGGTGCAATAATTCTTACTGTTACCACCGGATGGGCTGAAAACCGAC

       F  V  R  V  P  H  V  I  K  N  D  N  G  G  L  P  D  F  W  L   -


      TCTCTCTTTCGGGATTA
2221  ---------+------- 2237
      AGAGAGAAAGCCCTAAT

       S  L  F  R  D      -
```

153

Figure 13. Musca domestica patched sequence.

The figure shows the nucleotide sequence of exon 2, and partial sequence from exon 3 of *mdptc*, and the deduced amino acid sequence. 2/3 indicates the position between the two exons.

The clone pmusca2.8 does not contain exon 1 of *mdptc*. In an attempt to clone this

region, the *M. domestica* λDASH II library was rescreened using a 1.8 Kb *PstI* fragment

that was made from the 5' end of pmusca2.8. As the first intron in the *Drosophila ptc*

gene is known to be over 9 Kb (Forbes, 1995), and pmusca2.8 contained at least 4 Kb of

intron 1 sequence, this cloning strategy should also isolate upstream sequences. It

would then be possible to use the 5' regulatory regions of *mdptc* to drive the expression

of a *LacZ* construct in *D. melanogaster* to allow for the comparison of the expression

pattern of *ptc* generated by the endogenous and exogenous regulatory sequences. Any

differences between the normal wild type and the *Musca* driven *LacZ* patterns would

indicate functional divergence of either the regulatory sequences of the *ptc* gene itself,

or in its regulators such as *ci* (Alexandre *et al.*, 1996).

Four phage clones were isolated using this strategy and were analysed by Southern

blotting. Three probes were used in the Southern analysis of these phage clones. These

were the pmusca2.6 insert (to check for exon 3 coding sequence), a 400 bp exon 1

specific *BamHI/MluI* fragment from *D. melanogaster* (data not shown), and a 1.3 Kb 3'

intron1 *XhoI/EcoRI* fragment from pmusca2.8. This Southern analysis indicated one of

these clones, λ2, contained two *EcoRI* bands of approximately 19 Kb (Sλ2) and 21 Kb

(Lλ2) that were 5' to the most 3' intron sequence contained in pmusca2.8 (see Figure

14). This indicates that the restriction of λ2 by *EcoRI* was incomplete as λ phage arms can only accommodate inserts of up to around 20Kb.
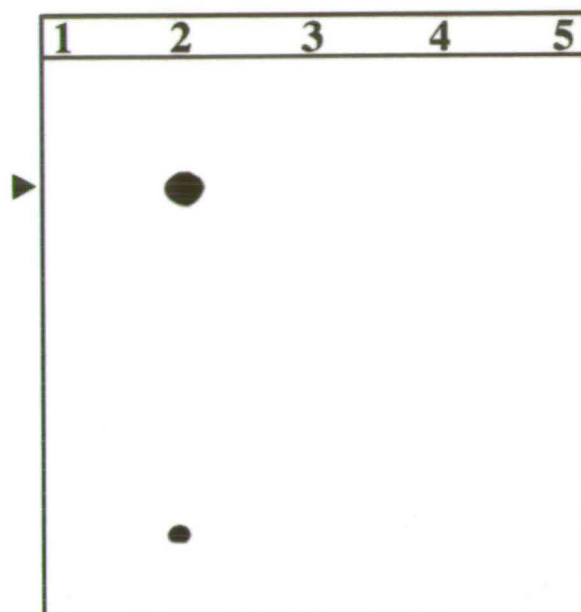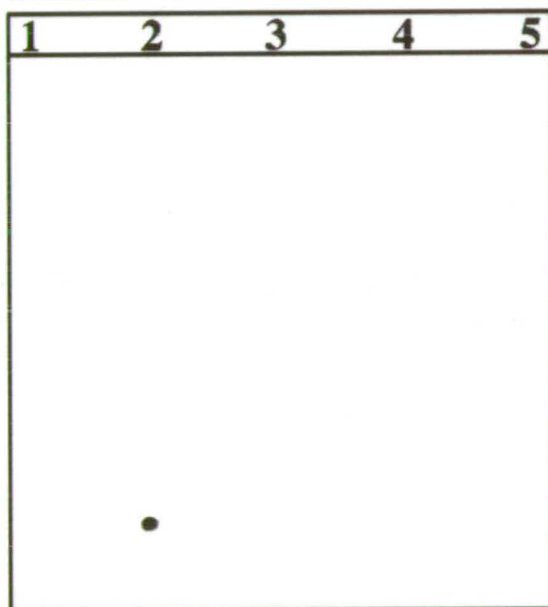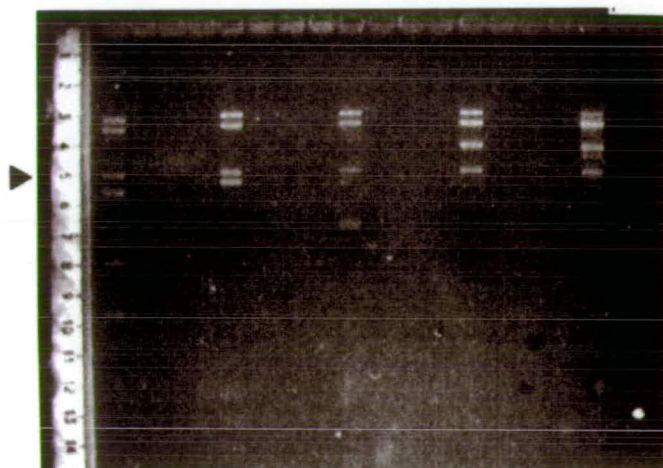
Figure 14. Southern hybridisation results from the screen for the 5' end of *mdptc*.

Lane 1 is a *HindIII* digested λ phage marker. Lanes 2, 3, 4, and 5, are *EcoRI* digests of four different λDNA clones isolated from a λZAP II *M. domestica* genomic DNA library by hybridisation using a 1.8Kb *PstI* fragment from the 5' end of pmusca2.8. (A) Southern hybridisation using a 1.3Kb 3' intron1 *XhoI/EcoRI* fragment from pmusca2.8. The top band in lane 2 (λ2) corresponds to a doublet at 19-21 Kb. (B) Southern hybridisation using pmusca2.6 to check for the presence of 3' coding sequence. The small band is constant in all of the Southern analyses of the λ2 DNA clone, regardless of the probe used. (C) Shows the agarose gel of the digested λDNA clones. These results indicate that the λ2 DNA clone contained 2 bands which may contain exon 1 and possibly 5' regulatory sequences from *mdptc*.

These *EcoRI* bands were purified, and attempts were made to subclone them into pBluescriptII KS. Unfortunately it was not possible to complete this during the term of my PhD.

A further aim of this study was to determine whether sites known to be important in the regulation of *ptc* in *Drosophila melanogaster* were also present in the upstream regulatory regions of *mptc*. *ci* encodes a transcription factor known to bind to the upstream regulatory sequence of *ptc* in *D. melanogaster*. The *ci* binding sites present in the 5' sequences of *D. melanogaster ptc* have been identified and sequenced (Alexandre *et al.*, 1996). There are at least three of these sites (5'- GACCACCCA) in the upstream region of *D. melanogaster ptc*, one in the forward orientation and two in the reverse. If *ci* plays the same role in *Musca* as it does in *Drosophila* then these sites should be under strong evolutionary constraint and be present in the upstream region of *mdptc*. Two PCR primers were designed so as to anneal to the *ci* sites in either orientation. Three PCR reactions were then performed using different combinations of the primers, two using a single primer and another using both primers, to allow amplification between the *ci* sites regardless of their orientation. Amplification from the $\lambda 2$ phage DNA indicates that there may be *ci* sites present in the upstream region of *mdptc*. Unfortunately, the number and orientation of the sites is impossible to determine accurately from the results obtained.

3.2.7. Phylogenetic analysis of the Ptc protein sequences.

3.2.7.1. Introduction.

Sequence data, especially at the amino acid level, can provide good raw data for the generation of phylogenies. This is because it is much less susceptible to subjective interpretation than traditional morphological criteria. The most common way of displaying the evolutionary relationships between different protein sequences is in the form of a tree. These trees are constructed from a distance matrix, which is generated by making pairwise comparisons of the given sequences (Farris, 1972 ; Feng and Doolittle, 1987; Fitch and Margoliash, 1967; Saitou and Nei, 1987). The resulting matrices can be translated into a tree whose branching order should reflect the evolutionary distances between the sequences under comparison.

A phylogenetic analysis of the available complete *ptc* sequences was performed. The analysis was performed using several different algorithms for calculating the evolutionary distances between the sequences.

160

3.2.7.2. The generation of the trees.

The distance matrices were created using the methods available on the GCG suite of programs. These were the Unweighted Pair Group Method by arithmetic Averaging (UPGMA) (Sneath and Sokal, 1973), and the Neighbor-Joining method of phylogenetic inference (Saitou and Nei, 1987).

The UPGMA calculations assume that the rate of mutation is constant and that the evolutionary time scale can be estimated by the number of amino acid substitutions, i.e., that the relationship between number of amino acid substitutions and evolutionary time scale is linear. The sequences are aligned to one another in a pair-wise manner, with those sequences that have already been aligned acting as a single 'sequence'. The sequences that are aligned to each other during each round are those that use the shortest distance to form the new cluster. The resulting tree is therefore rooted, with the root being defined by the two sequences with the most similarity. The distance between the cluster of sequences and the next sequence, or cluster of sequences, to be added to it is calculated using the equation;

distance (k,C) = [distance(k,a) x N(a) + distance (k,b) x N(b)] / (N(a) + N(b))

where C represents the new cluster being formed from clusters a and b. N(a) is the number of sequences in cluster a, N(b) is the number of sequences in cluster b, and k is another sequence.

The Neighbor-Joining method (Saitou and Nei, 1987; Studier and Keppler, 1988) aligns the sequences in a pair-wise manner by trying to minimise the total branch length of the entire tree. This generates an unrooted tree, as all of the possible pairings are considered during each round of comparison.

Several methods were used to correct for multiple substitutions at a single position. If there is no correction for this, the calculated distance between two distantly related sequences is likely to be an underestimate of the true distance.

3.2.7.2.1. Uncorrected distance.

This method does not compensate for multiple substitutions, and calculates the observed distance between two sequences. It does this by assigning fractional match scores to the residues at each position, i.e., for nucleotides N matched to A would score 0.25, and computing the similarity score using the equation;

$$S = \text{match score} / \text{number of positions scored}$$

where S is the similarity score, and the distance (D) is then 1-S.

3.2.7.2.2. Kimura protein distance.

This method is based on the relationship between observed residue substitutions and actual substitutions (Dayhoff *et al.*, 1972). Using only exact matches to calculate the similarity score;

$$S = \text{exact matches} / \text{number of positions scored}$$

$$D = 1\text{-}S$$

$$\text{distance} = -\ln(1\text{-}D - 0.2D(2))$$

Using this method can lead to overestimates of the true distance if the uncorrected distance exceeds 70 substitutions in 100 residues.

3.2.7.2.3. Jukes-Cantor distance.

This is another method based on the uncorrected distance (D), and can be used for both amino and nucleic acids. The distance is calculated using the equation;

$$distance = -b \ln(1 - D/b)$$

where the correction parameter b is 3/4 for nucleic acids, and 19/20 for amino acid sequences.

3.2.7.3. Description of the trees.

Using the methods described above, the possible phylogenies based on the available sequence of the *ptc* homologues were calculated. Two groups of sequence data were used to generate the phylogenies, one including the *C. elegans* sequence whose predicted protein showed the most similar hydropathy profile to the other *ptc* proteins, the other excluding it. The sequences were analysed in this way because the nematode sequence has not been shown to encode a true *ptc* homologue, as it also has homology to both the nematode *tra-2* and bacterial ion channels. Considering the two groups separately, the trees show similar results, depending on the methods used to generate them (see Figure 15 A-H and legends). There is a general pattern amongst all of the trees, in that the two Diptera species are very closely related, and are highly diverged from the vertebrate group. Among the vertebrate group the human and mouse are the most closely related, with the zebrafish being the most diverged. The phylogenies generated by either the UPGMA, or the neighbor-joining method, that include the *C. elegans* protein, place it as an out-group.
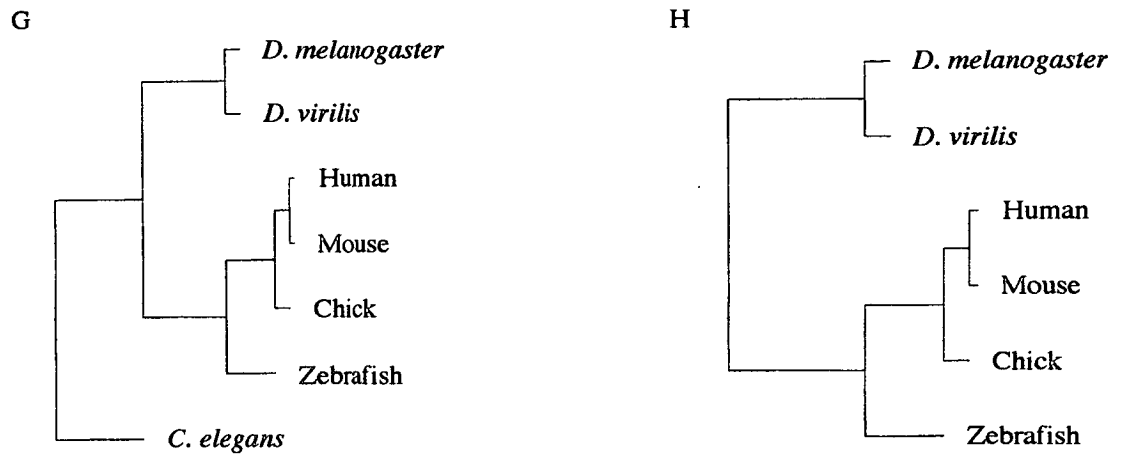
G

H

Fig 15. Phylogenetic relationship between the complete Ptc protein sequences.

(A, B, C, D, E, and F) All show phylograms generated using the neighbor-joining method.

(A, and B) Jukes Cantor distance correction.

(C, and D) Kimura protein distance correction.

(E, and F) Uncorrected distances.

(A, C, and E) Show phylograms with the inclusion of the *C. elegans* putative Ptc protein.

They all place the nematode as an outgroup from both the insects and the vertebrates, being slightly more closely related to the vertebrates than the insects.

(G, and H) Phylograms generated using the UPGMA method.

As the general pattern of the phylograms is the same regardless of which method of distance correction is used, only the uncorrected results are shown.

(G) Inclusion of the *C. elegans* putative Ptc protein does not affect the pattern of relationships seen between the flies and the vertebrates, as it is placed as an outgroup, approximately equally related to both the flies and the vertebrates.
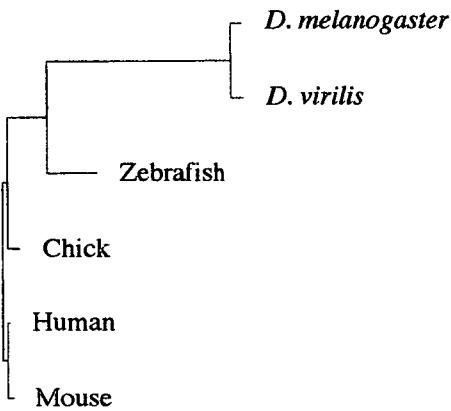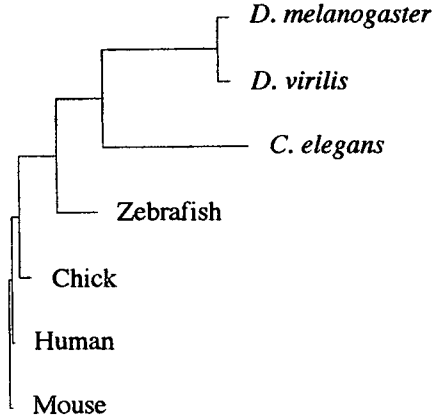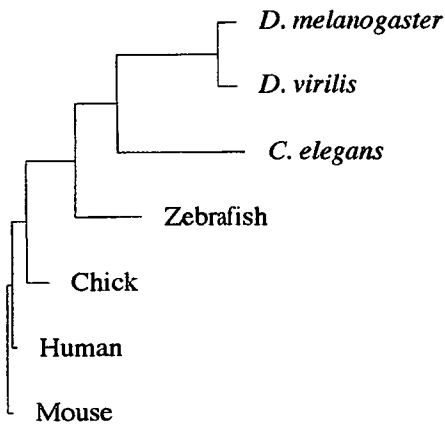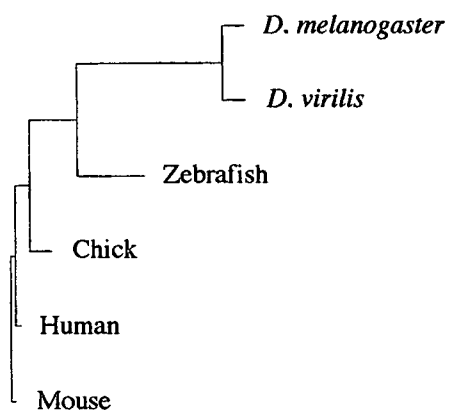
A
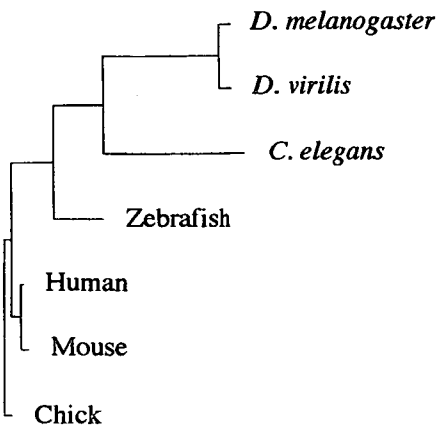
B
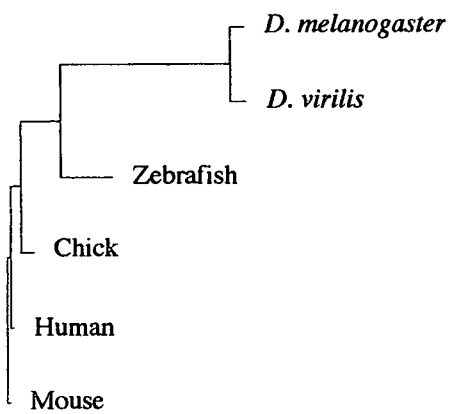
C

D

E

F

3.3. Discussion.

Until recently, the only *ptc* data came from two *Drosophila* species, *D. melanogaster* and *D. virilis*. These two species are thought to have diverged around 60 million years ago, which was considered to be long enough for non-essential sequence elements to have diverged (Forbes, 1995). The overall similarity between most genes that have been studied in both of these *Drosophila* species tends to be quite high, usually around 80% (Curtis *et al.*, 1995; Tillib *et al.*, 1995; Treier *et al.*, 1989). The similarity between the two *ptc* genes is in agreement with this, with there being 78% similarity between the two genes at the nucleotide level. This suggests that the assumption, that there has been sufficient time since the two species diverged for non-essential sequence elements to have diverged, is true. The level of similarity between the two *Drosophila ptc* genes is not, however, uniform along their entire length, with domains of much higher homology, such as the transmembrane domains, being separated by regions of little homology (Forbes, 1995). This compartmentalisation of similarity suggests that specific domains are important for *ptc* function. These domains have not yet been dissected to an extent where it is known what particular characteristics are the most important, i.e., is it the positioning of particular residues on the extracellular loops, the relative positioning of different elements of the protein, or indeed the overall structure of the entire protein that is the fundamental defining characteristic of the *ptc* homologues?

The partial cloning of *mdptc* is an important step in the study of the evolution of *patched* within the insects and its role in segmentation and appendage patterning (see Chapter one). *Musca* diverged from *Drosophila* at least 100 million years ago (Hennig, 1981), yet morphologically their early development is very similar. Given this degree of morphological similarity during the early development of *Musca* and *Drosophila,* it would seem likely that the molecular mechanisms involved in the development of these Dipteran species would be reasonably conserved. In particular, the specific domains that are necessary for *ptc* function in early development of the Diptera should have been conserved. These domains have already been grossly identified by the comparison of the two *Drosophila* sequences, but molecular analysis of *mdptc* should allow finer dissection of those domains. This will, of course, only be possible if *ptc* function is similar in the three Dipteran species. To investigate this, *mdptc* expression must be compared with the expression of the two *Drosophila* species. This analysis has been performed and the results are presented in Chapter four.

The genomic structure of the *mdptc* gene appears similar to that of the two *Drosophila* genes. The intron/exon boundary sequences that have been identified in all three species, are in similar positions within the gene. If it can be assumed that the *ptc* genes of the Diptera are very similar to one another then over half of the coding sequence of *mdptc* has been presented here.

Investigation of *mdptc* has predicted that it encodes a protein that is as similar to the *D. melanogaster* protein as is the *D. virilis* protein, at least over the two exons examined. The level of similarity between *mdptc* and the *Drosophila ptc* genes fluctuates in a similar manner to the *Drosophila ptc* genes themselves, the transmembrane domains showing greater similarity between the species than the intervening regions do. This result is not surprising since, from comparative embryology, it is known that embryogenesis is very similar in these Dipteran species, and that *ptc* has an important role in the transduction of the hedgehog signal during *Drosophila* segmentation and limb patterning (Hidalgo and Ingham, 1990; Johnson *et al.*, 1995; Phillips *et al.*, 1990). It is even less surprising now that it has been shown by hydrophobicity analysis that the overall structure of the Ptc protein has been conserved in some vertebrate species as well, and now that is known that the hedgehog pathway itself is conserved in a number of vertebrate species (Concordet *et al.*, 1996; Goodrich *et al.*, 1996; Marigo *et al.*, 1996).

The phylogenies generated using the complete Ptc protein sequences are in general agreement with phylogenies based on morphological criteria, as well as those based on molecular data, placing the insects together as one group and the vertebrates as another. Inclusion of the *C. elegans* sequence does not affect the relationships between the other Ptc sequences. The relationship between the nematode and the other species, indicated by the phylograms, suggests that the nematode lineage diverged before the invertebrate and vertebrate lineages diverged.

In conclusion, the results presented here show that *Musca domestica* has a *ptc* homologue which has been named *mdptc*. It has also been shown that there is a high level of similarity at both the nucleotide and amino acid level over the two exons that have been compared between *Musca domestica, Drosophila melanogaster,* and *Drosophila virilis*.

3.3.1. Future work.

Many questions arise from the results presented here. The first and most obvious question is 'what does the rest of *mdptc* look like?'. In order to be able to answer this question it is necessary to clone the remainder of the *mdptc* gene. It is possible that the 5' end of the gene is present in the λ2 phage DNA that has been isolated from a *M. domestica* genomic DNA library (see section 3.2.6). If this is the case, the λ2 phage DNA will contain coding sequence from both exons 1 and 2, as well as upstream regulatory sequences. Therefore, characterisation of the λ2 phage DNA is necessary. If the genomic structure of *mdptc* can be considered to be similar to that of *Drosophila ptc*, and the first intron is around 9 Kb, λ2 could contain up to 7 Kb upstream sequence. It is not likely that it would be possible to directly sequence the first exon of *mdptc*, which would be expected to be around 150 bp in length, from the λ2 clone because of the amount of flanking intron sequence contained in the clones. To aid in the

characterisation of exon 1 of *mdptc* , smaller restriction fragments of λ2 that contain

exon 1 of *mdptc* could be identified by Southern analysis, using exon 1 from *Drosophila*

ptc as a probe, and these should be easier to analyse. Another approach would be to use

Reverse Transcription - Polymerase Chain Reaction (RT-PCR), to amplify *mdptc* exon

1 which could then be used to screen another library for the 5' end of the gene, which

would hopefully be more amenable to analysis than the λ2 clone. The 3' end of the

gene would also require cloning. From sequence analysis of *ptc* from the various

species it is known that the 3' end of the gene is one of the regions that has little

homology between the species. It may be possible to amplify the missing 3' end with

RT-PCR, using a primer designed to the 3' end of exon 3 and a primer to the mRNA

poly-A tail. Cloning and sequencing of the remainder of *mdptc* would allow us to make

stronger arguments about the evolution of the gene within the insects.

Combining the molecular data presented here with similar data from the other Dipteran

species in which *ptc* has been cloned, shows that the *ptc* gene has been highly conserved

within this order of insects, but has its function?. The first step in investigating the

conservation of *ptc* function within the Diptera would be the examination of its

expression pattern in the three species. Comparative embryology has shown that the

early development of *D. melanogaster, D. virilis,* and *M. domestica* is remarkably

similar. Conservation of the elements in its expression pattern would therefore be

indicative of conservation of function. This kind of analysis has been performed

previously for other genes such as *nanos* (Curtis *et al.,* 1995), *hunchback* (Treier *et al.,*

1989; Wolff *et al.*, 1995), and *knirps* (Sommer and Tautz, 1991b). This analysis has been performed for the *ptc* homologues in both *D. virilis* (embryo expression: Forbes, PhD thesis; larval expression: this thesis), and now *Musca*, and the results will be presented in Chapter four.

Another aspect of *ptc* evolution within the Diptera would be the evolution of its regulatory mechanisms. It is known that *ptc* is part of the hedgehog pathway and is most probably the receptor for hedgehog itself. This pathway has been intensively studied in *D. melanogaster*, and it is known that both *cubitus interruptus* and *engrailed* both directly regulate the expression of *ptc* by binding to sites in the upstream sequences of the *ptc* gene (see Chapter one). Conservation of this pathway has been demonstrated between *D. melanogaster* and several vertebrate systems (Fietz *et al.*, 1994) (Ingham, 1995), so it does not seem unreasonable to assume that the hedgehog pathway would also be conserved in the housefly.

There would seem to be two ways in which *mdptc* regulation could be investigated. The first would be to show that there were binding sites for both *ci* and *en* in the upstream sequence of *mdptc*. There are three *ci* sites and two *en* sites within 600bp upstream of the transcription initiation site of *ptc* in *D. melanogaster* (Alexandre *et al.*, 1996). *engrailed* has been shown to be expressed in a large number of species using a monoclonal antibody, 4D9, that recognises a small epitope contained within the *en* homeodomain. Using this antibody, *en* expression has been described in *Musca*

(Sommer and Tautz, 1991b) (see Chapter four) and has been shown to be expressed in a pattern that is indistinguishable from the *Drosophila en* pattern. It would therefore seem probable that there would be *en* sites, and *ci*, sites within a similar region of the upstream sequence of *mdptc*. Indeed preliminary PCR data would seem to suggest that there are a number of *ci* sites present (data not shown). To unequivocally demonstrate the presence of such binding sites it would be necessary to sequence the upstream DNA of *mdptc* and identify the binding sites. Another method by which the regulation of *mdptc* could be investigated within the Diptera would be to use constructs containing the upstream region of *mdptc* driving *LacZ* expression and examining whether these can recreate the wild type *ptc* pattern of expression when introduced into *D. melanogaster*. However, even if the pattern of expression of *LacZ* generated by these constructs is similar to the endogenous *ptc* expression pattern caution must be employed when analysing the results. Jost *et al* (1995) have shown that the *D. hydei* orthologue of the *D. melanogaster ftz* could rescue the *ftz⁻* pair-rule phenotype in *D. melanogaster*, but not the embryonic lethality, even though *LacZ* expression patterns driven by the *D.hydei* upstream regulatory elements is identical to the wild type *D. melanogaster ftz* expression patterns (Jost *et al.*, 1995). However, if the *LacZ* pattern is not similar to that of *ptc*, it does not necessarily rule out the possibility that the *Musca* homologues of *ci* and *en* regulate *mdptc* expression in the housefly, only that *D. melanogaster ci* and *en* are sufficiently different from their *Musca* counterparts so as to be unable to function in their place. To address this possibility it would be necessary to characterise the *Musca ci* and *en* homologues. A third method would entail performing similar genetic analysis

of *mdptc* in segment polarity mutants in the housefly. Unfortunately at this time it is not possible to perform the same kind of genetic experiments that are possible in *D. melanogaster* in *Musca* as neither the mutants or the necessary transgenic techniques presently exist.

When considering the evolution of *ptc*, especially within the insect orders, it would be useful to know whether *ptc* homologues can be found in the lower orders, and if so whether it plays a similar function in the segmentation process. To start to address this question, a screen for *ptc* homologues was undertaken in the house cricket, *Acheta domesticus* (Orthoptera). The results of this screen are presented in Chapter five. It would also be useful to know whether the *C. elegans* transcription unit that has been reported to be a *ptc* homologue is indeed a true homologue, or whether it is more closely related to the nematode *tra-2* gene or the bacterial ion channels. Analysis of this transcription unit may be able to shed light on the evolution of the *ptc* gene within the invertebrates, and possibly even provide clues as to the origin of *ptc*.

4. Expression of the Dipteran homologues of *ptc*.

4.1. Introduction.

There are two distinct levels at which a gene can be conserved, structural and functional. Conservation of the DNA or protein sequence, i.e. structural similarity, can be demonstrated relatively easily by cloning and sequencing the gene, but structural similarity is not proof of functional equivalence. One way to demonstrate that a gene has been functionally conserved is to show that it can perform the same function as its homologues in a different system (Curtis *et al.*, 1995; Ingham and Fietz, 1995). This approach could, however, give misleading results, because it is possible for the sequence of orthologous genes which have been functionally conserved to have diverged sufficiently so they function in different species.

Another way of investigating functional equivalence, albeit indirectly, is to study patterns of gene expression. Comparison of the expression of homologous genes in different species with similar modes of development can potentially highlight aspects of that expression which are functionally important. This type of analysis has been used previously in the study of some of the segmentation genes. For example, expression of the gap gene *hunchback* (*hb*) has been compared in *D. melanogaster, D. virilis* and *Musca domestica*, and has been shown to have differences in some secondary pattern elements, i.e., those elements not directly associated with a role in segmentation, and

these have been assumed to be functionally unimportant (Sommer and Tautz, 1991b; Treier *et al.*, 1989).

Sommer and Tautz (1991b) partially cloned a number of the segmentation genes from *Musca domestica*, including *bicoid, hunchback, Krüppel, knirps, tailless, hairy,* and *engrailed*. They have shown some differences in the expression patterns of some of these genes with respect to the patterns in *D. melanogaster*. The most common variation was in the developmental timing of some of the pattern elements, which often appeared to be delayed in relation to the *Drosophila* patterns. *hunchback* expression showed the greatest divergence from that of *Drosophila*. As well as showing the typical *Drosophila* pattern elements, including the apparently redundant maternal expression (see Chapter One), *Musca hb* also showed a completely novel period of expression in 11-13 irregular stripes at the beginning of gastrulation. From the results of these investigations, Sommer and Tautz have concluded that the hierarchy of maternal, gap, pair-rule, and segment polarity genes elucidated in *D. melanogaster* has been conserved in *Musca domestica*.

ptc expression has previously been described throughout the early development of *D. melanogaster* (Nakano *et al*, 1989; Hooper and Scott, 1989; Phillips, 1990), and during the embryogenesis of *D. virilis* (Forbes, 1995), *Danio rerio* (Concordet *et al.*, 1996), *Gallus gallus* (Marigo *et al.*, 1996), and *Mus musculus* (Goodrich *et al.*, 1996).

4.1.1. *ptc* expression in *Drosophila* embryogenesis.

In *D. melanogaster,* it has been shown that *ptc* is expressed from early blastoderm

cellularisation (Hooper and Scott, 1989; Nakano *et al*, 1989; Forbes, 1995), and through

larval development (Phillips *et al.,* 1990). RNA *in situ* hybridisation, to both whole

embryos and sections, shows that the earliest detectable *ptc* transcripts are found at stage

5 (Campos-Ortega and Hartenstein, 1985), the cellularising blastoderm. At this stage,

the transcript is present at an almost uniform level across the entire surface of the

embryo, except for an antero-dorsal patch corresponding to the unsegmented acron, and

a posterior region surrounding the pole cells, in which *ptc* RNA is absent. At stage 6,

the beginning of gastrulation, raised levels of *ptc* transcription can be detected in stripes

2 cells wide. These stripes become more evident in an antero-posterior manner, until at

the beginning of stage 8 at least 11 stripes are clearly visible. This pattern continues to

resolve until at the end of stage 8, germ band extension, the transcripts are detectable in

a segmentally periodic pattern of 15 broad stripes. These stripes consist of an

ectodermal and mesodermal component. Expression in the ectodermal stripes is

strongest at the anterior edge and does not reach the dorsal edge of the embryo. The

ectodermal domains of expression occupy the posterior 75% of each parasegment,

whereas the mesodermal stripes are smaller and occupy the anterior third to half of each

metamere. The mesodermal stripes are out of register with the ectodermal stripes and

seem to fall in the gaps between them (Nakano *et al.,* 1989). At this stage there are also

two other broad domains of expression, one in the labrum (dorsal head), and the other in

the ectodermally derived cells of the hind gut/analia. The broad stripes of expression in the ectoderm become sharper and extend to the dorsal edge of the embryo, so at the end of stage 9 the stripes are exactly complementary to the *en* expressing cells. *ptc* expression is further modulated through stage 10. The broad bands begin to split into two stripes, each one cell wide, with *ptc* expression being lost between them. By stage 12, the two stripes are dorso-ventrally uniform, having gone through a series of regulatory intermediates, with expression in the anterior stripe being generally stronger than in the posterior. One of the stripes marks the anterior segment boundary, and the other is coincident with the *wg* expressing cells, which abut the cells that express *en,* i.e., they are the most posterior cells of each parasegment.

ptc expression during *D. virilis* embryogenesis is very similar to that in *D. melanogaster*, except for a small difference at stage 11 in the resolution of the posterior stripe (Forbes, 1995). In *D. melanogaster* the posterior narrow stripe becomes dorso-ventrally uniform, losing the strong ventral expression. The posterior stripe in *D. virilis* retains a high ventral expression, and in addition has a lateral patch of high expression that is not present in *D. melanogaster*. As the germ band begins to retract in *D. virilis*, the stripe becomes more weakly uniform along its dorsal-ventral axis, although expression remains high in the lateral patch. At the end of germ band contraction the lateral patch has faded and the expression patterns of *ptc* in *D. virilis* and *D. melanogaster* are indistinguishable (Forbes, 1995). No function has yet been ascribed to this lateral patch of *ptc* expression in *D. virilis*.

178

The aims of this study were to visualise the distribution of transcripts of *ptc* homologues, and Ptc protein distribution, using the techniques of RNA *in situ* hybridisation and immunohistochemistry, respectively, during the embryonic and larval development of *Musca domestica*, and the larval development of *D. virilis*. In addition, the expression of *en* during the embryonic and larval development of *Musca domestica* was also studied.

4.2. Results.

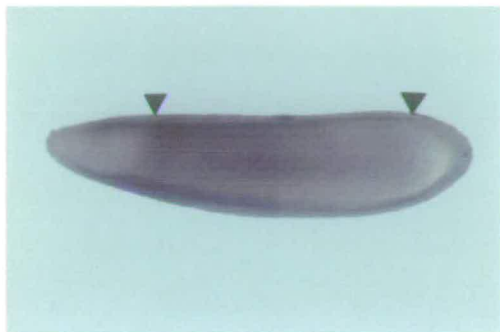### 4.2.1. Embryonic expression of *mdptc*.

Using RNA *in situ* hybridisation on whole mount *M. domestica* embryos, *mdptc* RNA was first detectable in the cellularising blastoderm from 10-90% EL (see Figure 16a). Expression at this stage appeared to be stronger in the anterior half of the embryo. Before the germ band began to extend, the large domain of expression resolves into a series of narrow stripes (see Figure 16b). This resolution appeared to occur in an antero-posterior sequence, so that just prior to gastrulation there are 13 stripes that extend along the entire dorso-ventral axis of the embryo. The two anterior most stripes were considerably stronger and more sharply defined than the stripes in the posterior (see Figure 16b).

179

The pattern evolved further as the germ band extended, so that towards the end of germ band extension *mdptc* was expressed in a pattern of twelve domains of expression in the segmented trunk region of the embryo, and in two other domains (see Figure 16c). The first of these two domains was found straddling the cephalic furrow. The expression anterior to the cephalic furrow was quite undefined and extended laterally into the posterior head region. The other domain was found in the posterior of the embryo, and again was quite undefined, appearing strongest around the edge of the hindgut invagination. If the expression pattern seen in *M. domestica* is directly comparable to that seen in *D. melanogaster*, it is probable that these cells are the ectodermally derived cells of the hindgut (Hooper and Scott, 1989). The domains of expression in the trunk region of the embryo show a higher level of expression in the posterior of each domain, and did not extend to the dorsal edge of the embryo.

At the end of germ band extension, just prior to germ band retraction, the domains of expression in the trunk have resolved into pairs of stripes (see Figure 16d). Neither stripe of a pair seemed to extend to the dorsal edge of the embryo, and they were slightly different from one another. The anterior stripe was more uniform along its extent, whereas the posterior stripe was more intense at its dorsal end, which is more lateral with respect to the embryo. At this stage there was also expression in the head region, the major domain being in the region of the stomodeum. As the germ band retracted, the stripes extended to completely encircle the embryo, and were now very narrow

180

bands, 1-2 cells wide. Expression in the posterior stripe of the pair faded and increased in the anterior stripe (with respect to a pair of stripes within a single segment; see Figure 6). The posterior stripe was not uniform along its length; expression in the region of the ventral midline was lower than in the rest of the stripe (see Figure 16e and f).
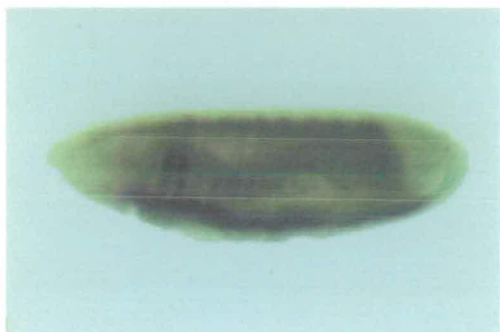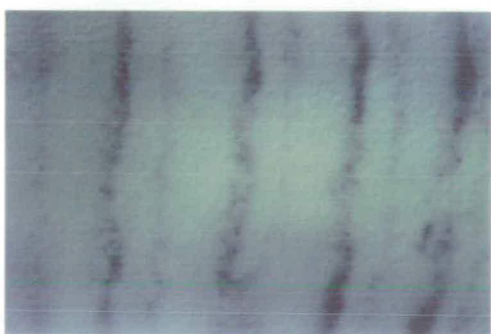
Figure 16. Expression of *mdptc* during *Musca* embryogenesis, as visualised by RNA *in situ* hybridisation with a probe generated from p2.8RVX.

(A) The blastoderm stage. *mdptc* is expressed uniformly between 10 and 90% EL (arrowheads demarcate limits of expression). (B) Prior to germ band extension the uniform expression seen in the early blastoderm resolved into a series of stripes within the original domain of expression. These stripes appeared in an anterior to posterior sequence. (C) The pattern of *mdptc* expression has resolved into twelve broad stripes in the segmented trunk region, as well as one in the anterior of the embryo, straddling the cephalic furrow (arrow), and one in the posterior of the embryo. (D) The fully extended germ band stage. The twelve broad stripes of *mdptc* expression have split, by loss of expression in the middle of the stripe, to give a pattern of two narrow stripes in each segment. (E) By the time the germ band has retracted the anterior of each of the stripes in a segment have faded, and the posterior stripe is a defined single row of intensely staining cells. (F) Shows a higher magnification ventral view of T4 - T7. The anterior stripe of the pair (arrowheads) has faded. Staining along the ventral midline (*VM*) is much reduced in both stripes of a pair. (A, C-F) Photographed using differential interference contrast (DIC) optics. (B) Photographed using bright field optics. Anterior is to the left, and dorsal to the top, except in panel (F) which is a ventral view (anterior to the left). Scale bar represents approximately 500 μm, but does not relate to panel F, which was a photographed at a higher magnification.

The expression pattern of *en* in *D. melanogaster* has been described in detail elsewhere

(DiNardo *et al*, 1985; Ingham *et al*, 1985; Kornberg *et al*,1985; Weir and Kornberg,

1985; Karr *et al*, 1989). *en* is often used as a marker of the posterior compartment of

the segment and the segment border, both in *Drosophila* and other species (DiNardo *et*

al., 1985; Karr *et al.*, 1989; Kornberg *et al.*, 1985; Patel *et al.*, 1989a; Patel *et al.*,

1989b). The monoclonal antibody, mAb 4D9, which recognises a 14 amino acid

epitope situated in the homeodomain of the *invected* protein, has been shown to have a

high degree of cross-reactivity and will recognise the *en* protein from a large number of

species across the taxa, e.g. *Drosophila*, grasshopper, some crustaceans, and vertebrates

(Patel *et al.*, 1989a; Patel *et al.*, 1989b). Using this antibody, the distribution of the *en*

protein was characterised throughout *M. domestica* embryogenesis.

During blastoderm cellularisation, 14 stripes of expression, 1-2 cells wide, appeared in

an anterior to posterior sequence. The most anterior stripe was just anterior to the

region of the future cephalic furrow. As the stripes first appeared, there was an

alternating modulation of intensity which soon disappeared (see Figure 17a). During

germ band extension, the stripes increased in width to become approximately 3 cells

wide (see Figure 17b, 17b (ii), and 17c). After germ band extension, several patches of

expression appeared in the head region of the embryo. By the end of germ band

retraction the expression of *en* protein in the head had become quite complex, and the

morphology of the stripes changed again. The two most anterior stripes remained quite wide, whereas the ten stripes posterior to these narrowed to 1-2 cells (see Figure 17d), and expression appeared in some cells of the CNS (data not shown).

These data show that the distribution pattern of the *en* protein is remarkably similar to that seen in *D. melanogaster*. Sommer and Tautz (1991) reported the same findings, although they only showed expression at the cellularising blastoderm and extended germ band stage, and not at the later stages.
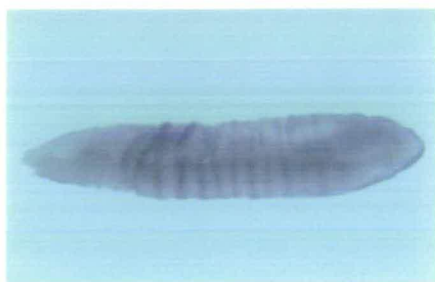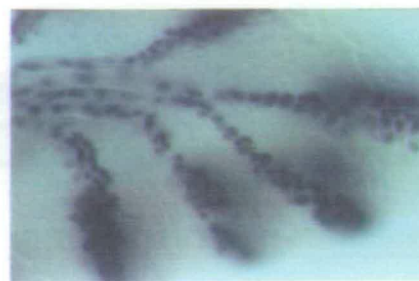
Figure 17. engrailed expression during *Musca* embryogenesis.

Embryos were stained using the monoclonal antibody mAb 4D9. (A) Expression in the blastoderm. Striped expression became apparent at this stage of *Musca* embryogenesis. The stripes arose in an anterior to posterior sequence. (B) As the germ band extended the stripes increased to three or four cells in width. (Bii) shows the 'streaming' of the nuclei as the germ band extended around the posterior tip of the embryo. (C) Prior to the retraction of the embryo, the stripes of *engrailed* expression became thinner, until at the fully retracted germ band stage (D) the stripes of *engrailed* expression are a single cell wide in the ectoderm of the segmented trunk region. All panels are orientated with anterior to the left, and dorsal to the top. Embryos were photographed using DIC optics. Scale bar represents approximately 500 μm, but does not relate to panel B(ii) which was photographed at a higher magnification.

4.2.3. Larval expression of Dipteran Ptc proteins.

A monoclonal antibody, mAb 5E10, was generated against the N-terminal end of the *D. melanogaster* Ptc protein by Dr. A. Taylor (I.C.R.F., London). This was used to investigate the expression of *ptc* protein in the thoracic, and the eye-antennal imaginal discs of third instar larvae of three dipteran species, *D. melanogaster, D. virilis*, and *M. domestica*. The antibody proved to work well in the third instar wing discs of all three species, and in the other imaginal discs in the two *Drosophila* species, but less well in *Musca*, in which RNA *in situ* was used to examine *mdptc* expression in the imaginal discs. The staining in all of the *M. domestica* discs was generally less intense than in either of the *Drosophila* species. This could be for a number of reasons; (a) the antibody may not have bound to the *Musca ptc* protein with as high an affinity as it did to the *Drosophila* proteins, (b) the fixation protocol used for the *Musca* discs was not optimised for this particular antibody, or (c) the protein was not expressed at as high a level in the *Musca* discs as it was in the *Drosophila* species. The antibody was also used to examine protein distribution in the embryos of the same species, although in *D. virilis* and *M. domestica* it proved impossible to reduce the level of background to allow accurate interpretation of the data.

It has previously been demonstrated in both the embryos and thoracic imaginal discs of *D. melanogaster* that *ptc* is expressed in the cells of the anterior compartment that abut the compartment boundary (Hidalgo and Ingham, 1990; Hooper and Scott, 1989;

Nakano *et al.*, 1989). A similar pattern of expression has been demonstrated in these discs in all of the Dipteran species examined in this study.

As can be seen from Figure 18, the general morphology of the third instar imaginal discs was very similar between the three species. Figure 18(c) shows the expression of *ptc* in a third instar wing disc from *D. melanogaster*. The stripe was several cell diameters in width and extended along the entire length of the compartment boundary. There was a quantitative difference in intensity of the staining, the stripe in the dorsal region of the disc was strong, whereas in the region ventral to the wing pouch staining was slightly weaker. The positioning and general shape of the stripe was the same in the wing discs of *D. virilis*, although the difference in stripe intensity was even less pronounced (see Figure 18 b). In the wing discs of *M. domestica*, the shape and positioning of the stripe was very similar (compare Figure 18 a(i) with Figure 18 a(ii) which shows a *M. domestica* wing disc stained with 4D9 to highlight the compartment boundary), and the difference in intensity was very pronounced, almost fading to nothing in the dorsal region of the disc (see Figure 18 a(i)). The other thoracic discs (leg and haltere) showed very similar patterns of expression in all the three species, with the *ptc* stripe generally bisecting the discs, presumably expressed in the cells of the anterior compartment next to the compartment boundary (see Figures 19 and 20).

The eye-antennal discs are morphologically very different from the thoracic imaginal discs. All three species demonstrated a similar pattern of *ptc* expression within the eye-

antennal discs (see Figures 19 and 20). Across the antennal region of the discs, *ptc* protein was expressed in a single band, through the middle of the antennal disc. This again, was most likely to be the cells in the anterior compartment next to the compartment boundary. In the eye region of the disc, the situation was different. The *ptc* protein is expressed in two stripes, one either side of the advancing morphogenetic furrow, and at a low level in the cells that are posterior to the furrow.

ptc expression was also examined in the thoracic, and eye-antennal discs of *M. domestica* (see Figure 20) by RNA *in situ* hybridisation. The observed patterns were very similar to the protein distribution, as would be expected for a non-secreted protein. The most striking difference between the patterns generated by the two methods of visualising *ptc* expression was that the stripe of transcript in the wing disc shows an even greater change in intensity. Staining in the dorsal part of the disc was still quite strong, although in the middle of the wing pouch the staining dropped to almost undetectable levels.
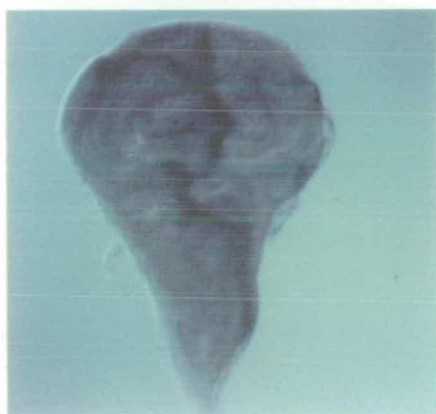
Figure 18. Ptc protein distribution in the third instar wing imaginal discs of *D. melanogaster, D. virilis,* and *M. domestica.*

(Ai, B, C) Distribution of Ptc protein in the third instar wing imaginal discs of *M. domestica* (Ai), *D. virilis* (B), and *D. melanogaster* (C) visualised by immunohistochemistry using the monoclonal antibody mAb 5E10. In all three species, *ptc* is expressed in a thin stripe on the anterior side of the compartment boundary. There is a modulation in the intensity of expression along the dorsal-ventral axis of expression; expression is almost absent in the ventral wing disc and the region of the future notum. The *D. melanogaster* wing disc shows the same variation in Ptc distribution as seen in the *Musca* disc. The variation is less pronounced in the wing disc of *D. virilis*, although this could be due to over staining this disc. (Aii) *engrailed* expression is seen throughout the posterior compartment in the third instar wing imaginal disc of *M. domestica*. Discs are orientated with anterior to the left and dorsal to the top. All discs were photographed using DIC optics; Scale bar represents approximately 100 μm in panels A(i) and A(ii), and 50 μm in panels B and C.

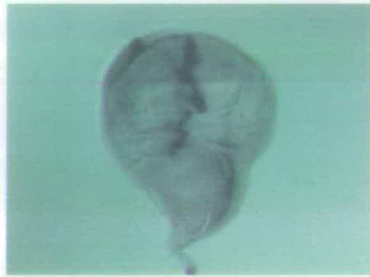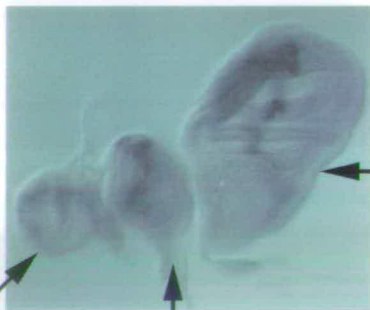*D. virilis*   *D. melanogaster*

A   B

C   D ← *MF*

E   F

G   H ← *wing disc*

leg disc

haltere disc

Figure 19. Ptc protein distribution in the imaginal discs of *D. melanogaster* and *D. virilis.*

Ptc protein distribution visualised by immunohistochemistry with the monoclonal antibody mAb 5E10. (A), (C), (E), and (G), Third instar imaginal discs from *D. virilis.* (B), (D), (F), (H), and (I), Third instar imaginal discs from *D. melanogaster.* (A), (B), (C) and (D), eye-antennal discs. Ptc protein is found in a narrow domain along both sides of the progressing morphogenetic furrow (*MF*). Expression in the antennal region of this disc is in a single stripe that bisects the disc, probably along the anterior-posterior compartment border (B). Expression in the antennal region of the *D. virilis* disc (A) is very weak compared to that seen in *D. melanogaster* (B). (E) and (F), third instar wing imaginal discs, compare with figure 18. (G) leg disc, and (H) thoracic discs (second and third thoracic leg discs and wing disc). Ptc protein is found in a single stripe that bisects each of these discs along the anterior-posterior axis, presumably along the compartment border. The discs were all photographed using DIC optics. Scale bar represents approximately 50 µm, but does not relate to panels C and D which were photographed at a higher magnification.

A   ← MF

B   ← *haltere disc*

                     ← *leg disc*
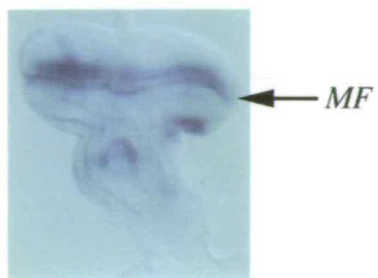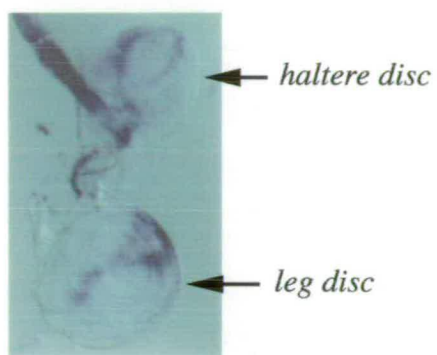
C

Figure 20. mdptc expression in the imaginal discs of *M. domestica.*

mdptc transcript was detected in the third instar imaginal discs of *M. domestica* by RNA *in situ* hybridisation with a probe generated from p2.8RVX. The expression patterns were similar to those seen with the monoclonal antibody mAb 5E10 in the wing discs of *M. domestica,* and in *D. melanogaster,* and *D. virilis.* (A) Expression in the eye-antennal disc. *mdptc* transcript was detected in two stripes, flanking the progressing morphogenetic furrow (*MF*), and weakly in some cells behind the furrow. (B) *mdptc* transcripts were seen in a stripe that bisects the leg discs. (C) The pattern of *mdptc* transcription in the wing disc was very similar to that in *D. melanogaster,* and *D. virilis.* The transcript was found localised along the anterior of the compartment border. The variation in intensity of staining along the dorsal-ventral axis is even more pronounced at the RNA level, than the protein level (see Figure 18). All the discs were photographed using DIC optics. Scale bar represents approximately 100 μm in panels A and C, and 125 μm in panel B.

4.3. Discussion.

As can be seen from the data presented here, and that reported previously, the early development of the three dipteran species, *D. melanogaster, D. virilis,* and *M. domestica,* is very similar (Curtis *et al.,* 1995; Sommer and Tautz, 1991b). Embryogenesis in all three species takes approximately 24 hours at 25°C. Each of the species has three larval instars, during which the imaginal discs proliferate and subsequently become patterned. This means that different developmental stages, such as the cellularising blastoderm, the extended germ band, and the larval instars, are directly comparable between the three species. This enables us to make direct comparisons of gene expression patterns in each of the species.

In conjunction with DNA sequence data from the different dipteran species, examination of expression patterns allows inferences to be made about the evolution of gene function across a large evolutionary timespan, in excess of 100 million years, in relatively similar developmental systems. The advantage of examining gene expression patterns in an experimental system like these Dipterans, is that components of the expression pattern that have been conserved through tens of millions of years can be assumed to have an important role in a particular process, i.e. the expression of *engrailed* in the cells defines them as the posterior lineage compartment of the segment, or the spatial relationship between the expression of *engrailed* and *wingless* for maintaining parasegmental borders (Martinez-Arias *et al.,* 1988). Differences may

become evident during such an analysis, but may not have any obvious functional significance, such as the lateral patch of *ptc* expression seen during the embryogenesis of *D. virilis* (Forbes, 1995). It is, however, important to note that expression patterns are not in themselves proof of function. For example in *D. melanogaster*, it is possible to drive expression of *ptc* uniformly throughout the embryo without causing any segmental defects (Sampedro and Guerrero, 1991). Had this pattern of expression been seen naturally occurring in another species it would have been assumed, on the basis of the expression pattern alone, that the *ptc* homologue in this species did not play a segment polarity role similar to that of *ptc* in *Drosophila*.

The data presented here show that the homologues of both *patched* and *engrailed* have very similar expression patterns in *D. melanogaster, D. virilis*, and *M. domestica*, during their early development.

It has been shown previously in a considerable number of organisms (Patel *et al.*, 1989a; Patel *et al.*, 1989b), that *en* can be used as a marker of the posterior of segments, and of segment borders. Given this and the morphological similarity between the species, no major differences would be expected in the expression of *en* between the species. Indeed, *en* expression was used in this study as an indicator of the position of the segment borders in the *Musca* embryos, and the antero-posterior compartment boundary in the imaginal discs (see Figure 18).

The embryonic expression of *mdptc* in *Musca* was shown not to differ significantly, spatially or temporally, from that of *ptc* in either *D. melanogaster* or *D. virilis*. The evolution of the final *mdptc* pattern goes through the same steps as the *Drosophila* pattern, from the uniform expression at the cellular blastoderm stage, the antero-posterior sequence of stripe formation, the dividing of a broad domain of expression into two thin stripes, and the final modulation of the stripes within a segment (see Figure 16f). This result lends support to the possibility that the regulation of *mdptc* in *Musca* may be similar to the regulatory mechanisms that control *ptc* expression in *D. melanogaster*. This fits well with recent data that has shown that the hedgehog pathway and the spatial relationship between its components have been conserved in species as far afield as the vertebrates (Concordet *et al.*, 1996; Goodrich *et al.*, 1996; Marigo *et al.*, 1996).

Expression of the *ptc* homologues in the third instar imaginal discs was also remarkably similar between the three species. All of the thoracic discs showed the same basic pattern of expression of *ptc* and its homologues when visualised using the 5E10 monoclonal antibody, and by RNA *in situ* hybridisation in *Musca*. The wing, leg, and haltere discs were bisected along the dorso-ventral axis by a thin stripe of *ptc* expression. This stripe is known to run along the anterior side of the compartment boundary in *D. melanogaster* separating the anterior and posterior compartments.

Comparison of the expression patterns of *en* and *mdptc* in *Musca* imaginal discs

suggests that this is also the case in the housefly.

There was an interesting quantitative difference of expression seen along the *ptc* stripe

in the third instar wing imaginal discs. The level of *ptc* expression was much lower in

the ventral part of the discs than in the dorsal region. It is intriguing that the degree of

difference of expression varies between the species, in *Musca* expression dropped to

virtually undetectable levels, whereas in *D. virilis* there seemed to be very little

alteration in the level of *ptc* expression between the dorsal and ventral regions of the

disc. One explanation for the variation in expression may come from the use of the

mAb 5E10 antibody. It is possible that the antibody had different binding affinities to

the Ptc proteins in the different species, or that the fixation protocols used for each

species was not optimised and hence, the difference was an artifact of the technique.

However, this seems unlikely, as the difference is seen within discs, and not between

discs, and the pattern is reproducible. This variation in expression was also detected in

the wing imaginal discs of *Musca* using RNA *in situ* hybridisation, which again suggests

that this is a real difference. Although this result may be explained, not in terms of the

regulation of *ptc* expression by members of the hedgehog pathway and its role in the

patterning of the wing disc, but in terms of the regulation of patterning along the discs'

dorso-ventral axis. Genetic analysis has shown that *D. melanogaster ptc* is expressed at

low levels across the whole of the anterior compartment (Phillips *et al.*, 1990). This

indicates that the control of *ptc* expression is much more complex than was first

thought. It has also been demonstrated that some components involved in dorsal-ventral patterning, such as *Notch,* have interactions with genes of the antero-posterior patterning system, such as *wingless* (Micchelli *et al.,* 1997; Neumann and Cohen, 1996), which in turn is known to control the expression of *engrailed* at some stages during embryogenesis. It is therefore possible that the modulation of expression of *ptc* along the antero-posterior boundary was a direct result of the expression of the genes of the dorsal-ventral system, and does not have a role to play in the antero-posterior patterning of the wing disc.

The pattern of expression of the *ptc* homologues in the antennal discs was not only very similar between the species, but also resembled that seen in the thoracic discs. The discs were again bisected by a thin stripe of *ptc* expression. The resemblance between the antennal *ptc* pattern and the pattern of the thoracic discs was not all that surprising if it is assumed that all of the ventral appendages of the adult fly have evolved from a common ancestral appendage.

It has been demonstrated that *hedgehog* (Ma *et al.,* 1993), *wingless,* and *patched* (Ma and Moses, 1995) are required for the progression of the morphogenetic furrow across the eye disc in *D. melanogaster. hh* is expressed posterior to the furrow, and the Hh protein signals across the furrow into the anterior cells to initiate the progression of the furrow (Ma *et al.,* 1993). The results presented here are consistent with this pattern of *hh* expression, as *ptc* is expressed in those cells anterior to the furrow, to receive the Hh signal. The continued expression of *ptc* in the cells posterior to the furrow, after the

furrow has passed, could be to ensure only transient expression of genes such as

hedgehog, in front of the furrow to control its ordered progression across the eye disc.

In conclusion, the data presented here shows that the expression of *mdptc* was

remarkably similar to that of *ptc* expression in the two *Drosophila* species during both

embryonic and larval development. This would indicate that the *Musca ptc* homologue

plays a homologous role to *Drosophila ptc* in the processes of segmentation and

appendage patterning, and that the regulatory mechanisms controlling the expression of

ptc and its homologues in the different Dipterans are probably also conserved.

4.3.1. Future Work.

The experiments that should follow on from the data presented here can be placed into

two categories; further characterisation of *mdptc* expression, and expression of the *ptc*

homologues in different species.

To enhance the characterisation of *mdptc* in the housefly, it would first be necessary to

complete the investigation into its expression during the larval development of *Musca*.

This would require examination of the expression of *mdptc* in the imaginal discs from

the earlier larval instars. However, although this is not impossible, it is technically very

difficult to obtain good preparations of the imaginal discs from earlier instars, due to

their small size. The embryonic pattern of *mdptc* presented in this study is one which

can be deduced from whole mount preparations of *Musca* embryos. It has been shown in *D. melanogaster* that embryonic *ptc* expression is segmental in both the ectoderm and the mesoderm, but the stripes of expression in the two germ layers are out of register with each other. To fully characterise the embryonic expression of *mdptc* it would therefore be necessary to examine thin sections through the embryo to be able to investigate expression in the different germ layers.

Although there is no direct evidence presented here for any spatial relationships between *mdptc* and *en* expression, nor any other segment polarity genes, it is important to investigate whether the spatial relationships between the segment polarity genes are conserved. In many species, it is possible to perform sequential or simultaneous RNA *in situ*, or immunohistochemical reactions, to directly analyse the spatial relationship between the domains of expression of different genes. I have attempted this sort of analysis, using both immmunohistochemistry and RNA *in situ* hybridisation to visualise *mdptc* expression, and immunohistochemistry to detect En protein. Unfortunately, I was unable to get this technique working well enough to be able to demonstrate unequivocally a spatial relationship between *mdptc* and En. The ability to examine the expression domains of multiple genes simultaneously would greatly enhance our understanding of the molecular basis of segmentation in *Musca domestica*, and the evolution of the mechanisms involved. It would therefore seem necessary to establish the protocols to perform these multiple stainings using either immunohistochemistry, RNA *in situ* hybridisation, or a combination of both techniques. It would then be

necessary to demonstrate that *mdptc* expression and *en* expression domains abut each

other during both embryonic and larval development, and examine *mdptc* expression in

relation to other segment polarity genes such as *hedgehog, cubitus interruptus,* and

wingless during early development. It would seem likely, given the level of

conservation of the hedgehog pathway, and the fundamental nature of segmentation and

appendage patterning, that the general spatial relationships between these genes in *D.*

melanogaster would be conserved.

This study showed that the general expression patterns of the dipteran *ptc* homologues

was highly conserved during the early development of the three species investigated

here. As already suggested in this thesis, this is not very surprising given the role of *ptc*

in the segmentation and patterning of appendages in *D. melanogaster*, and the

morphological similarities between *D. melanogaster*, *D. virilis*, and *M. domestica*.

Many of the lower orders of insect have very different modes of development from the

Diptera. Many are hemimetabolous, meaning that their larvae develop into adults

without metamorphosis, and develop their appendages, not as imaginal discs, but as

direct outgrowths during embryogenesis. These insects do not exhibit long germ band

embryogenesis, but rather short, or intermediate germ band embryogenesis. This means

that they do not form their segments simultaneously like *Drosophila*, but sequentially

from a growth zone at the posterior of a germ anlage (see Chapter One). Given these

differences in the early development of the insects, it would be interesting to determine

whether the lower insects have *ptc* homologues, that are used in segmentation and

appendage patterning. In an attempt to try to address these questions, partial clones

have been isolated from the house cricket, *Acheta domesticus* (Orthoptera), an

intermediate germ band, hemimetabolous insect, and the results are presented in Chapter

five.

5. Cloning and characterisation of a possible *ptc* homologue from the cricket, *Acheta domesticus*.

5.1. Introduction.

Much of the molecular data available about the development of short and intermediate germ band insects comes from the studies of two species; the locust, *Schistocerca gregaria* (or the very similar *Schistocerca americana*), and the red flour beetle, *Tribolium castaneum* (see Chapter One). These studies have concentrated on the homologues of the pair-rule genes, *hairy, fushi tarazu,* and *even-skipped*, the segment polarity gene, *engrailed*, and the homeotic genes (Brown *et al.*, 1994a; Brown *et al.*, 1994b; Dawes *et al.*, 1994; Patel *et al.*, 1992; Patel *et al.*, 1994b; Patel *et al.*, 1989a; Sommer and Tautz, 1993: Beeman, 1987; Beeman *et al*, 1989; Stuart *et al*, 1991; Stuart *et al*, 1993; Tear *et al*, 1990; Kelsh *et al*, 1993). However, there is not enough data to be able to fully understand the molecular basis of short germ embryogenesis, or draw conclusions about the molecular evolution of the different modes of insect embryogenesis. This situation is exacerbated by not knowing how representative the development of these species is of the short and intermediate modes of embryogenesis. These problems will only be clarified by the continued study of these species, and of the molecular basis of development in other short and intermediate germband insect species. One obvious system to investigate is the cricket, *Acheta domesticus*.

A. domesticus has long been a favoured insect system for traditional descriptive and experimental embryology (Sander, 1976). Experimental procedures such as ligature (Mahr, 1960), pinching (Vollmar, 1971), and irradiation (Kanellis, 1952; Sauer, 1962; Seidel, 1964) have all been used to study the cricket embryo (reviewed in Sander, 1976). The data obtained from such experiments have led to the formulation of several basic concepts of insect pattern formation, including the presence of ooplasmic determinants, and the totipotency of the syncitial nuclei. Unfortunately, the study of the development of *A. domesticus*, and of short and intermediate germ band insects in general, has been hampered by the inability to perform the sorts of genetic manipulations that are standard in *Drosophila* biology (see Chapter One). However, with the advent of molecular biology, it is now possible to investigate the molecular basis of *A. domesticus* development.

Very little molecular data is available from *A. domesticus*. A search through the GenEMBL sequence database reveals six entries, only three of which are complete codon sequence; diuretic hormone receptor (Reagan, 1996), apolipophorin-III (Smith *et al.*, 1994), and the 5S RNA gene (Cave *et al.*, 1987), and none of them are homologues of the *Drosophila* segmentation genes. Recently, however, there has been a report of the cloning of a *Distal-less* homologue, a gene involved in limb development in *Drosophila*, from the two-spotted cricket, *Gryllus bimaculatus* (Niwa *et al*, 1997).

ptc is an excellent candidate gene to examine as part of the study into the molecular basis of insect embryogenesis. It has been extensively studied in *D. melanogaster* and much is known about its genetic interactions with the other segmentation genes. The Ptc protein plays an important role in segmentation and pattern formation during the development of *D. melanogaster*, being a component of the hedgehog intercellular signaling pathway. This pathway has been found to be conserved between the insects (Diptera) and the vertebrates (mouse, chick and zebrafish), and the close spatial relationship between *ptc* and *hh* has been maintained in all of these species. The hedgehog pathway has been co-opted into many developmental processes in these species, including segmentation, limb patterning, and neural tube development. It would therefore be useful to study the components of the hedgehog pathway, including *ptc*, throughout the development of species across the animal kingdom, to shed light on the possible evolutionary relationships between the varied developmental processes that utilise the hedgehog pathway.

The aim of this part of the study was to investigate the cricket homologues of *ptc*. To facilitate this, a cDNA library was screened, and PCR used to obtain partial clones of a homologue of *ptc*. RNA *in situ* hybridisation was then used to visualise the distribution of the transcripts of one of the isolated clones during embryogenesis of *A. domesticus*, and the results of these studies are presented here.

5.2. Results.

5.2.1. Screening of the *A. domesticus* cDNA library.

Using the conditions previously described for the cloning of *sonic hedgehog* (Echelard *et al.*, 1993), a low stringency hybridisation screen of an *A. domesticus* cDNA λZAPII phagemid library (gift from Alex Kolodkin) was performed to isolate homologues of *ptc*.

A ^{32}P labeled 292bp *MluI / BamHI* fragment containing most of the coding region of exon 1 and 5' exon2 was derived from the 16C5 cDNA clone of *ptc* from *D. melanogaster* (originally isolated in Phil Ingham's laboratory), and used as a probe to screen approximately 250,000 clones. Two phagemid clones, PB and PD, 1.6Kb and 1.4Kb respectively, were isolated and partially sequenced from both ends, but revealed little sequence homology to *Drosophila ptc* (data not shown).

5.2.2. Degenerate PCR cloning of a partial cricket *ptc* homologue, *aptc*.

An alternative strategy to cloning homologues by hybridisation is to use degenerate PCR. This method is based on a standard PCR protocol, but utilises the fact that PCR primers do not have to be identical to the target sequence over their entire length in order to amplify specific products. It has been shown that for a primer of between 20

and 24bp the most important criterion is that the three most 3' bases of the primer are

100% identical, the degree of homology over the rest of the primer being less critical

(Sommer and Tautz, 1989). Degenerate PCR utilises primers where each 'primer' is a

pool of oligonucleotides that vary in base composition at one or more position, with the

premise that this sequence variation will compensate for changes in the target sequence

between species or genes within a gene family. Many genes have been cloned in this

way, including *bicoid*, *hunchback*, *Krüppel*, and *knirps* from *Musca domestica* (Sommer

and Tautz, 1991b), *ptc* from the zebrafish (Concordet *et al.*, 1996), and mouse

(Goodrich *et al.*, 1996), G-protein coupled receptors (Libert *et al.*, 1989), and the iron-

sulphur protein of succinate dehydrogenase from a variety of species including human,

rat, *Drosophila melanogaster*, *Arabadopsis thaliana*, and *Schizosaccharomyces pombe*

(Gould *et al.*, 1989).

A set of degenerate primers (the 'REV' primers, P4REV and P22 - Goodrich *et al*,

1996; see appendix D.) that have previously been shown to amplify a region of *ptc* of

between 330 and 350 bp from several vertebrate species, the mouse (Goodrich *et al*,

1996), the chick (Marigo *et al*, 1996), and the zebrafish (Concordet *et al*, 1996) as well

as *Drosophila*, were used to amplify the corresponding region from *A. domesticus*

cDNA. The REV primers amplify a region in *D. melanogaster* that spans the 3' end of

exon 2 and the 5' end of exon 3. This region contains three of the conserved cysteine

residues, and another is incorporated in the 5' primer, P4REV. A 345bp PCR fragment

was amplified from *Acheta* cDNA and cloned into pBluescriptKS II (Stratagene) (see

Figure 21). Sequencing the fragment, R_c10, indicated that it encoded a predicted protein sequence of 115 amino acids that is 58% identical and 75% similar to the *D. melanogaster* sequence, although it was three amino acids shorter than its *Drosophila* counterpart. All of the cysteine residues that are conserved in the REV fragments of the other species were also conserved in the *Acheta* R_c10 fragment. As might be expected, the alignment of the REV fragments from several species showed that R_c10 had greater similarity to the predicted protein sequences of the dipteran species, *D. melanogaster*, and *M. domestica*, than to either of the vertebrate sequences. On comparing the insect and vertebrate sequences, it is evident that the vertebrate REV fragments are shorter than the insect sequences (see Table 3). All of the insect sequences used show approximately the same degree of homology to the mouse sequence, *Acheta* REV being 45.4% identical to mouse REV whereas the *Musca* REV fragment is 45.9% identical to mouse REV. There was slightly more variation between the insects and the fish sequence, *Acheta* REV being 48.1% identical to zebrafish REV, whereas *Musca* REV was only 40.7% identical.

From the sequence data of *ptc* from the other insect species, *D. melanogaster, D. virilis,* and *M. domestica*, several areas of homology were identified and degenerate PCR primers were designed to these, in order to clone a larger region of the *Acheta ptc* homologue, *aptc*. The PREV4 primer was used in conjunction with cricket3'pcr (A3') (see appendix D.) to amplify a larger region of *aptc*. The expected product was a single band of around 1500bp, however at least three bands of less than 1Kb were amplified.

To establish whether any of these bands contained *aptc* sequences, they were analysed

by Southern hybridisation. Using the $R_c 10$ PCR fragment as a probe it was shown that

none of the products of the REV4/A3' PCR reaction contain the REV sequence (see

Figure 22).

Table 3. Degree of similarity between the REV PCR fragments from the *ptc* homologues of various species.

The numbers indicate identity, or, in parentheses, similarity, at the amino acid level of the REV PCR fragments of the *ptc* homologues from *Acheta domesticus, Drosophila melanogaster, Musca domestica, Danio rerio*, and *Mus musculus*.

	Length (aa)	A. domesticus	D. melanogaster	M. domestica	D. rerio	M. musculus
Length (aa)		115	118	115	108	109
A. domesticus	115	100 (100)	58.3 (74.8)	58.9 (75.0)	48.1 (70.4)	45.4 (69.4)
D.melanogaster	118	58.3 (74.8)	100 (100)	99.1 (100)	43.5 (66.7)	46.8 (66.1)
M. domestica	115	58.9 (75.0)	99.1 (100)	100 (100)	40.7 (62.0)	45.9 (66.1)
D. rerio	108	48.1 (70.4)	43.5 (66.7)	40.7 (62.0)	100 (100)	56.5 (76.8)
M. musculus	109	45.4 (69.4)	46.8 (66.1)	45.9 (66.1)	56.5 (76.8)	100 (100)

A

```
          TTAGATTGTTTTTGGGAAGGCTCGAAGTTGCTGGGTCCTGATTACCCTGTTCATATACCA
        1 ---------+---------+---------+---------+---------+---------+  60
          AATCTAACAAAAACCCTTCCGAGCTTCAACGACCCAGGACTAATGGGACAAGTATATGGT

a         L   D   C   F   W   E   G   S   K   L   L   G   P   D   Y   P   V   H   I   P   -

          ACATTAGGAGCGAAGGTGAAATGGACAAACCTCAACCCTTTGAAAATTGTAGAGGAAATG
       61 ---------+---------+---------+---------+---------+---------+  120
          TGTAATCCTCGCTTCCACTTTACCTGTTTGGAGTTGGGAAACTTTTAACATCTCCTTTAC

a         T   L   G   A   K   V   K   W   T   N   L   N   P   L   K   I   V   E   E   M   -

          AGAAATTTCGATTTCCACCATTTTCCCTTCGACACTTTGGAGGATTATATGAAAAGAGCA
      121 ---------+---------+---------+---------+---------+---------+  180
          TCTTTAAAGCTAAAGGTGGTAAAAGGGAAGCTGTGAAACCTCCTAATATACTTTTCTCGT

a         R   N   F   D   F   H   H   F   P   F   D   T   L   E   D   Y   M   K   R   A   -

          GGAATTAGTAGTGGGTACCAAGAAAAACCATGCCTCGATCCTAGTGATGTCGAATGCCCA
      181 ---------+---------+---------+---------+---------+---------+  240
          CCTTAATCATCACCCATGGTTCTTTTTGGTACGGAGCTAGGATCACTACAGCTTACGGGT

a         G   I   S   S   G   Y   Q   E   K   P   C   L   D   P   S   D   V   E   C   P   -

          GAAACTGCATCAAACAAGAAGGCTGGCCAGTCTCCAGATATTGGTGCTGAACTGACAGGA
      241 ---------+---------+---------+---------+---------+---------+  300
          CTTTGACGTAGTTTGTTCTTCCGACCGGTCAGAGGTCTATAACCACGACTTGACTGTCCT

a         E   T   A   S   N   K   K   A   G   Q   S   P   D   I   G   A   E   L   T   G   -

          GGGTGTTATGGATTTGCTGCCAAGTATATGCACTGGCCAGAGGAR
      301 ---------+---------+---------+---------+-----  345
          CCCACAATACCTAAACGACGGTTCATATACGTGACCGGTCTCCTY

a         G   C   Y   G   F   A   A   K   Y   M   H   W   P   E   E   -
```

B

```
            1                                                              50
fish        LDCFWEGSKL  qG..gsaylP  GmPd.iqWmN  LdPLklMEEl  sQ......ft
mouse       LDCFWEGaKL  qs..gtayll  GkPp.LrWTN  fdPLeflEEl  Kk......In
drosophila  LDCFWEGSqL  LGPEsAVviP  GLnQRLLWTt  LNPasVMqyM  KQKMSEEKIS
musca       LDCFWEGSqL  LGPEsAVviP  ...QRLLWTt  LNPasVMqyM  KQKMSEEKIS
cricket     LDCFWEGSKL  LGPdypVhIP  tLgakvkWTN  LNPLkivEEM  r...nfdfhh
Consensus   LDCFWEGSKL  LGPE-AV-IP  GLPQRLLWTN  LNPL-VMEEM  KQKMSEEKIS

            51                                                            100
fish        .slEgfreml  dkAqvGhaYM  nrPCLdPsDt  dCPhsAPNKd  pwQvPniaAe
mouse       yqvdswEeml  nkAevGhGYM  drPCLNPaDP  dCPaTAPNKN  STkPlDValv
drosophila  FDFETVFqYM  KRAAIGSGYM  EKPCLNPlnP  nCPDTAPNKN  STQPPDVGAi
musca       FDFETVEqYM  KRAAIaSGYM  EKPCLNPlnP  nCPDTAPNKN  STQPPDVGAi
cricket     FpFdTlEdYM  KRAgIsSGYq  EKPCLdPsDv  eCPeTAsNKk  agQsPDiGAe
Consensus   FDFETVE-YM  KRAAIGSGYM  EKPCLNP-DP  -CPDTAPNKN  STQPPDVGA-

            101         118
fish        LqGGChGfsk  KfMHWqEE
mouse       LnGGCqGlsr  KyMHWqEE
drosophila  LSGGCYGyAA  KhMHWPEE
musca       LSGGCYGyAA  KhMHWPEE
cricket     LtGGCYGfAA  KyMHWPEE
Consensus   LSGGCYG-AA  K-MHWPEE
```
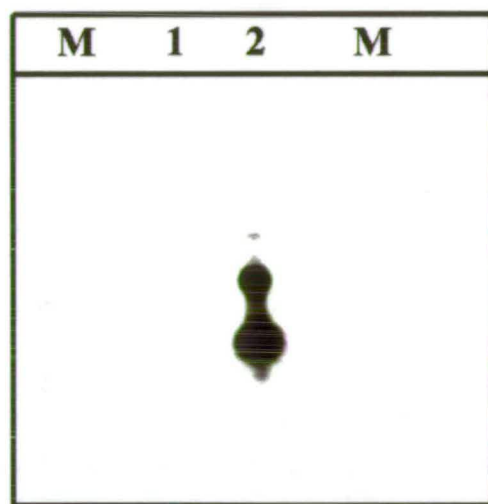
Figure 21. Rev PCR fragments.

(A) Nucleotide and deduced amino acid sequence of the Rev PCR fragment from *Acheta domesticus.* (B) Alignment of the Rev protein sequences from different species. Gaps are represented by dots, dashes represent unconserved residues, and conserved residues are given in capital letters in the consensus sequence.
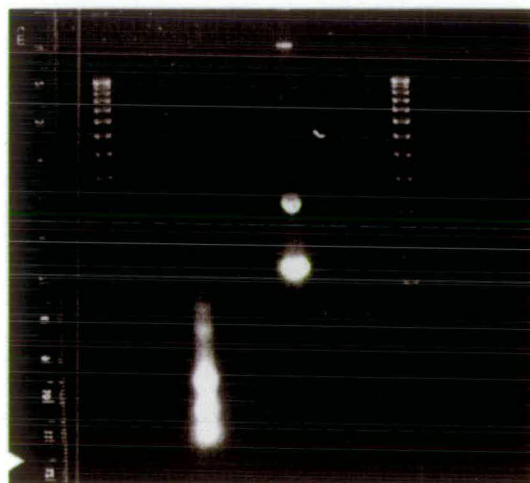
A

| M | 1 | 2 | M |

B

| M | 1 | 2 | M |

C

Figure 22. Southern hybridisation analysis of REV4/A3' PCR.

The R_c10 PCR fragment was used as a probe to determine whether the products of the PREV4/A3' PCR reaction contained the REV sequence. (A) Shows the autoradiograph after a 12 hour exposure. Lanes marked 'M' are size marker; lane 1, REV4/A3' PCR; lane 2, R_c10 positive control (R_c10 cloned into pBluescript II). The R_c10 probe hybridised to all of the PCR products as well as the positive control. (B) The same autoradiograph after a two hour exposure. There was no positive signal in lane 1. (C) Agarose gel showing the PCR products and R_c10 positive.

5.2.3. Identification of the PB and PD clone inserts.

The REV primer set was used to diagnose whether the two original clones from the library screen, PB and PD, contained the REV sequences. If the REV band were present in PCR reactions performed on the PB and PD phagemids, it would be indicative of one or both of them being *aptc*. PCR with the REV primer set resulted in the amplification of bands of the correct size from both PB and PD. These PCR products were cloned into the *EcoRI* site of pBluescriptKS II (Stratagene). Sequencing of the two products showed that the REV PCR product from the PB phagemid showed no sequence homology to the R_c10 fragment, or to any of the other known REV fragments. However, the REV PCR fragment generated from the PD phagemid is identical to the original R_c10 fragment, indicating that the PD phagemid contained sequence from the *aptc* gene. On this basis, PD was used to generate RNA probes for use in whole mount *in situ* reactions against *Acheta domesticus* embryos.

5.2.4. Embryonic expression.

There are two published developmental series for *Acheta domesticus*. Lauga (1969), published the first, based on externally visible events throughout development at 25°C. He assigns developmental stages based on morphology, and gives an absolute timing of embryogenesis. Edwards and Chen (1979) adapted Lauga's staging with particular reference to the development of the abdominal cerci. They state that the duration of

embryogenesis in *Acheta domesticus* is highly variable, (which agrees with my own observations), so rather than give an absolute time course of embryonic development in terms of hours, which can be inaccurate, they time the appearance of specific stages with respect to the percentage of the total mean length of embryogenesis. In the following results and discussion, Edwards and Chen (1979) staging has been used to define the embryonic stage of the specimens.

The PD phagemid generated a 1.4Kb RNA probe to examine the distribution of the PD transcript during *A. domesticus* embryogenesis. Both the sense and anti-sense strand probes were made by cutting the PD phagemid with *SacI* and transcribing with T7 RNA Polymerase, or by cutting with *KpnI* and transcribing with T3 RNA Polymerase, respectively.

No specific staining was seen at any of the stages examined when the control hybridisations were performed using the sense strand probe (data not shown). However, using the anti-sense strand probe, PD transcript was detected at around stage 15/16 (see Figure 23a), prior to the onset of katatrepsis (see Figure 6). This is considerably later than overt segmentation of the body region is first seen. When the transcript was first detected it was seen in each of the trunk segments present (see Figure 23a), but expression was qualitatively different in the abdominal and thoracic segments. In the thoracic segments, expression was seen in a thin stripe, 1-2 cells wide, posterior to the segment border, although there was a little diffuse staining evident in the posterior of

segments T1 and T2, i.e., anterior to the segment boundaries. The expression at the anterior of the segment, i.e. posterior to the segment border, was more intense than the expression in the posterior of the segments, although it should be noted that in all of the thoracic segments the staining was quite diffuse. In the abdominal segments, expression was primarily along the midline, and spread along the anterior border in the more anterior abdominal segments. There was also a quantitative difference between the expression in the different segments. Expression in the anterior of T1 was stronger than that in either of the other thoracic segments, and there was a gradient of expression down the segments of the abdomen. In A1, expression of PD was detected along the anterior border, although it did not reach the lateral edges, and in a wedge that extended from the anterior border, along the midline to approximately halfway through the segment. In the more posterior segments, expression was not detected along the anterior border, but was along the midline, although the expression became weaker in each more posterior segment until it was almost undetectable in the last segment of the specimen, A6. At stage 16/17, PD transcript was first detected in the antennae and some of the mouthparts. In the antennae expression was seen in a domain in the anterior proximal region, and was quite diffuse. Expression in the mouthparts was restricted to the mandibular and maxillary segments, and was seen in an anterior region and also at the distal tip (see Figure 23b). As development proceeded towards stage 18 the expression in the thoracic segments took on a wedge shape similar to that of the abdominal segments at stage 17 (see Figure 23c). Expression in the abdominal segments increased in width and intensity along the midline, slightly spreading along the segment borders

220

(see Figure 23c and d). Expression in the mouthparts intensified at the distal tips (see

Figure 23e), and appeared in a small domain slightly proximal to the tips. At this stage

there were several domains of expression that became evident in the labrum, and

dorsally in the very anterior tip of the head (see Figure 23e). At stage 18, expression

was detectable in the legs in two distinct spots on the distal tip of the last tarsal segment,

which were probably the claw precursors (see Figure 23c). In stage 20 embryos, the

latest developmental stage examined during this study, and characterised by the

completion of katatrepsis, distribution of the PD transcript did not change significantly

in the trunk segments. At this stage, however, three distinct domains of expression

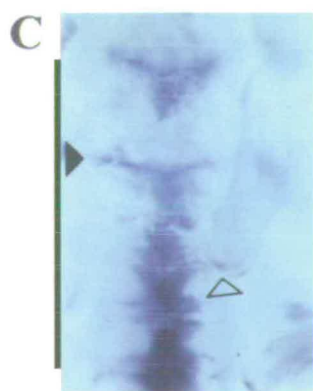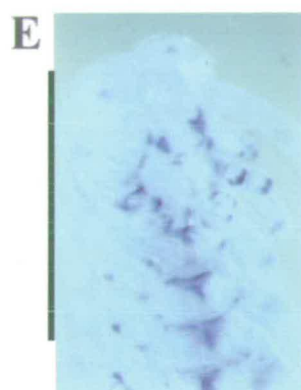became apparent in each of the coxa regions of the legs (see Figure 23f).

Figure 23. PD transcript distribution during *Acheta domesticus* embryogenesis.

Embryos were hybridised with a 1.4 Kb probe generated from the PD phagemid. (A) Dorsal view of a stage 15/16 embryo. PD transcript was first detected in the thoracic segments, and in the abdominal segments that were formed at this stage. The expression pattern was different in the thoracic and abdominal segments. In the thoracic segments, PD transcript was found in a thin stripe along the segment borders, whereas in the abdominal segments PD transcript was distributed in a wedge shape along the midline. (B) Lateral view of a stage 16/17. Transcript was detected in the anterior region of the proximal antenna. (C) Ventral view of a stage 18. Expression in the thoracic segments had expanded along the segment borders (arrow), and down the midline. In the abdominal segments the expression widened along the midline and began to spread along the segment borders. Expression was first detected in the developing legs, in two spots at the distal tip (tarsal precursors) (open arrow). (D) All of the body segments are formed, and PD transcript was detected in each. Expression was in a triangular domain in the thoracic segments, and along the midline in the abdominal segments. (E) Ventral view of a stage 19. PD transcript was evident in the head, and in the mouthparts. (F) Ventral view of a stage 20. The limbs were well developed, and PD transcript was detected in the tarsal precursors (open arrow) and in three domains at the proximal end of the limb (arrows). All of the embryos are orientated with anterior to the top. Scale bars represent 500 μm. Photography was performed using DIC optics.

## 5.2.5. Checking the identity of the PD clone insert.

In order to confirm that the PD clone contained a fragment of a cricket *ptc* homologue, and that the previously performed REV PCR had amplified the REV fragment from the PD clone and not from exogenous sources such as the R$_c$10 clone or REV product contamination, a series of vector anchored PCRs were performed. Using the T3 and T7 primers in the pBluescript phagemid vector in combination with the PREV4 and P22 primers, on the assumption that if the PD clone contained the REV sequence, PCRs using the PREV4/T3, and P22/T7 primer combinations would amplify products that spanned the entire 1.4 Kb insert (see Figure 24a). These PCRs were performed, along with single primer controls and a PREV4/P22 positive control, and the results are shown in Figure 24b. After subtraction of the T3 (~700 bp) and T7 (~1200 bp) single primed bands, there were no bands that when combined could have spanned the entire 1.4 Kb PD insert, although the PREV4/P22 positive control always gave a single band of the correct size (see Figure 24b). There were also bands in the P22/T3 (180 bp) and PREV4/T7 (680 bp) primer combination lanes, as well as the expected PREV4/T3 (530 bp) and P22/T7 (250 bp) primer combination lanes. These results suggest that the T3 and T7 primers could anneal to the PD clone in multiple positions in opposing orientations.

The PD phagemid was also resequenced using the T3, T7, PREV4, and P22 primers. The PREV4 and P22 primers repeatedly gave no readable sequence, whereas the T3

gave 473 bp of readable sequence, and the T7 gave 506 bp (data not shown). The T3

and T7 primed sequence was used in a BLASTX search, using the web based BLAST

programs available at http://www.ncbi.nlm.nih.gov/BLAST/. The T3 sequence

generated no positive hits, whereas the T7 sequence showed between 53 and 58%

similarity over short stretches (84 -127 bp) of nucleotide sequence of the major sperm

protein of several nematode species, *Pratylenchus scribneri*, *Pratylenchus penetrans*,

and *Asceris suum*.

In combination, these results showed that it is highly unlikely that the PD clone contains

sequence from the *Acheta* homologue of *ptc*, and the PREV4/P22 positive control is

probably amplifying from PCR product contamination rather than the R_c10 clone.
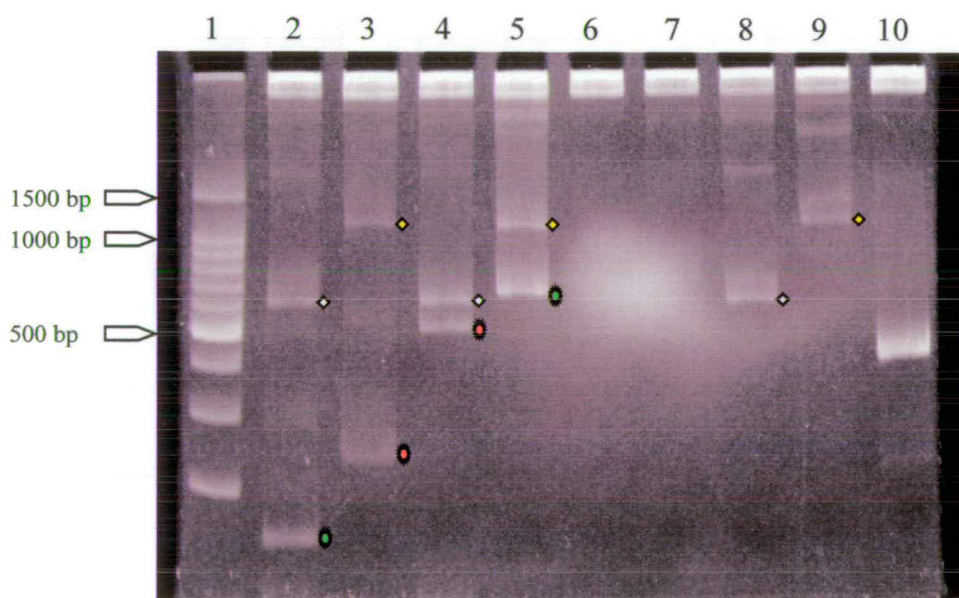
A

PREV4

REV

P22

T7

T3

PD phagemid
~ 4.3 Kb

B

1   2   3   4   5   6   7   8   9   10

1500 bp

1000 bp

500 bp
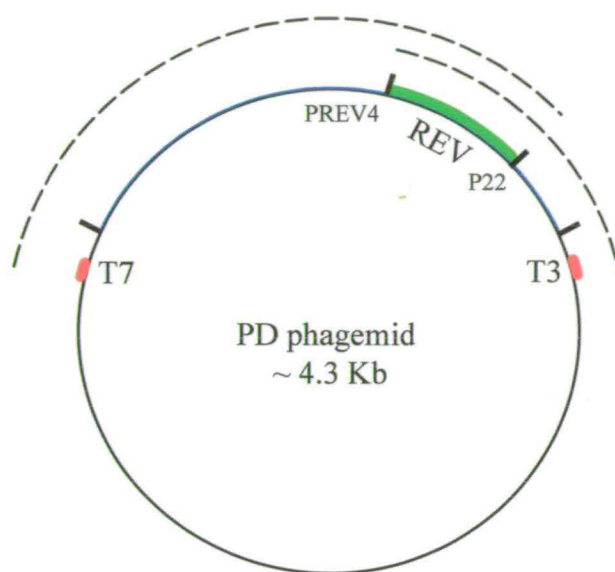
Figure 24. The Vector-Anchored PCR Experiment.

(A) The putative structure of the PD clone.
The figure shows the predicted structure of the PD clone, prior to the vector-anchored PCR experiments, and the two predicted PCR fragments generated by the vector-anchored PCR (dotted lines). The PD insert is shown in blue, the REV fragment in green, the T3 and T7 primer sites in red, and the PREV4 and P22 primer sites in black.

(B) The Results of the Vector-Anchored PCR.
Lane 1 is a 100bp marker; lane 2 = P22/T3; lane 3 = P22/T7; lane 4 = PREV4/T3; lane 5 = PREV4/T7; lane 6 = P22/P22; lane 7 = PREV4/PREV4; lane 8 = T3/T3; lane 9 = T7/T7; lane 10 = P22/PREV4.

The single primed bands are shown by white diamonds (T3/T3) and yellow diamonds (T7/T7).
The paired bands are shown by red stars (P22/T7 [~240 bp] and PREV4/T3 [~530 bp]), and green stars (P22/T3 [~180 bp] and PREV4/T7 [~680 bp]).

5.3. Discussion.

The results presented here show that the cricket, *Acheta domesticus,* has a *ptc*

homologue, a fragment of which was cloned using degenerate PCR, and was called

R_c10. Of the two phagemid clones, PB and PD, that were isolated from the cDNA

library, only PD was thought to contain a *ptc* homologue based on a diagnostic PCR

which amplifies the REV fragment. The PD clone was then used to create RNA probes

which were used in an attempt to examine the expression of the cricket *ptc* homologue

during embryogenesis. However, it was subsequently shown with further PCR

experiments (see Figure 24b), and partial sequence analysis (data not shown) that it is

unlikely that the PD clone contained cricket *ptc* sequence, and the expression data,

therefore, cannot be interpreted as the cricket *ptc* expression pattern. The PD insert

therefore remains unidentified because, although the BLASTX search generated hits

with several nematode sequences, it is probable that these sequences were pulled from

the database by chance, and this does not show that the PD phagemid contains these

sequences.

5.3.1. What is the PD clone?

Formally, there are two possible categories of answer to this question. The first is that

the PD clone does contain sequence from the cricket *ptc* homologue. Within this

category, there are several possibilities of what the PD insert could be;

228

(A) The PD insert is a fragment of the cricket *ptc* homologue.

The only data to support this hypothesis is the amplification of a band of approximately the correct size with the REV primers. However, the partial sequence data from the PD clone does not support the possibility of the PD insert being a 1.4 Kb cricket *ptc* fragment. It is much more likely that the band generated by the REV primers was not amplified from the PD clone, but from some exogenous source of REV sequence such as PCR product contamination, or the R_c10 clone. Given that none of the negative controls (template free) that were run with each of the PCR experiments amplified any product, the reagents and experimental setup were likely to be free from contaminating *ptc*-like sequences. These negative controls can not, however, rule out the possibility that the PD phagemid preparation might be contaminated with DNA containing REV-like sequences. From the data available it is highly unlikely that the PD clone contains a single *ptc* homologue fragment. This would, therefore, mean that the expression data can not be interpreted as the cricket *ptc* expression pattern.

(B) The PD insert is a product of a coligation event during the production of the library.

It is possible that the REV PCR data from the PD clone is real, but there are two (or more) unrelated sequences in the clone, one being a cricket *ptc* homologue, and the other unidentified. If the PD clone did indeed contain more than one fragment, it is possible that the expression data presented here could be the expression pattern of a

cricket *ptc* homologue, or the combined expression patterns of the genes from which the ligated fragments originated.

The library screening strategy, which utilised a *Drosophila ptc* probe that contained exon 1 and 5' exon 2 sequence, should have isolated cricket *ptc* homologue clones that contained extreme 5' coding sequence. The PD insert is approximately 1.4 Kb, so it is possible, if the cricket *ptc* homologue has a similar genomic organisation and is of a similar size to *Drosophila ptc*, that the PD clone contains 5' ptc sequence that is contiguous with the REV fragment, which in *Drosophila* spans 3' exon 2 and 5' exon 3 (see Figure 25). The partial sequence data is inconclusive with regard to this hypothesis. The T3 primed sequence does not generate any positive hits in a BLASTX search but, in this hypothesis, it might have been the extreme 5' end of the cricket *ptc* sequence, which is a region of sequence that has been shown to be quite different between species. However, if this hypothesis were correct, I would have expected the 473 bp of T3 primed sequence used in the BLASTX search to have resulted in positive hits with *ptc* sequence from other species. The T7 primed sequence showed very limited similarity to nematode major sperm protein in a BLASTX search, which suggests that a second fragment could be a major sperm protein homologue if the PD clone was a product of a coligation event. The vector anchored PCR experiments do not support this hypothesis, because if the PD insert contained the REV fragment, the vector anchored PCR would have generated a number of bands that when combined would have spanned the entire 1.4 Kb insert, independently of the rest of the insert sequence.

In conclusion, the available data does not support the hypothesis that the PD clone was

the result of a coligation event.

~500 bp sequence

~480 bp sequence

REV

T7

T3

MSP??

ptc
EXON 3

ptc
EXON 2

ptc
EXON1

Figure 25. One possible configuration of a coligated PD insert.

The diagram shows one possible configuration of a coligated PD insert. A *ptc* fragment is shown in green, and a putative genomic structure, based on that of *Drosophila ptc* indicated. Another sequence, possibly a major sperm protein homologue, is shown in red. The turquoise bars represent regions that have been sequenced, the T3 and T7 labels indicate orientation. The position of the REV fragment is indicated by the blue bar. This diagram shows how such a product may not have been recognised to contain *ptc* sequence, although it is highly unlikely that the 476 bp of T3 primed sequence would not have shown any homology to other *ptc* sequences, and generated *ptc* positive hits in the BLASTX search.

The other possibility is that the PD clone does not contain any sequence from the cricket *ptc* homologue.

(A) The PD insert contains REV-like, non-*ptc* related sequence.

It is possible that the PD insert contains sequence to which the REV primers can anneal, but is not related to *ptc*. The data from the vector anchored PCR experiments supports this hypothesis. The available sequence data does not provide any evidence that PD contains any sequence that is similar to the REV sequence, although it is possible that a REV-like sequence is present in the region of the insert that has not yet been sequenced. However, this hypothesis seems unlikely given the degree of similarity of the REV sequences from the *ptc* homologues from several different species, and the previously demonstrated specificity of the REV primers (Goodrich *et al*, 1996). If, however, this were the case, the expression pattern shown here would not be that of a cricket *ptc* homologue.

(B) The PD clone contains REV-like, *ptc* related sequence.

It is possible that the PD clone may contain a fragment of sequence from a gene that has arisen by duplication from a cricket *ptc* homologue, and since diverged so as to be unrecognisable as a *ptc* homologue by sequence alone. From the data available there is no way of validating this hypothesis, nor does it provide a framework within which the expression data presented here can be interpreted.

(C) The PD clone does not contain REV-like sequence.

From the available data, the most likely hypothesis is that the REV primers were amplifying from REV sequence containing contaminants of the PD preparation. The vector anchored PCR experiment, and the available sequence data are both supportive of this hypothesis.

These results raise the question; 'what is contaminating the PD preparation?'

If there had been contamination of the PD phagemid preparation with the R_c10 clone, the vector anchored PCR would have amplified a band of approximately 400 bp, which was not seen in any of the PCRs. If, however, the contamination was a PCR product which contained REV-like sequence, and the PD clone did not contain REV-like sequences, there would have been no bands generated in vector anchored PCRs. It follows from this logic that the most likely hypothesis is that the PD preparation is contaminated by a PCR product that contains REV-like sequence, but that the PD clone also contains sequences that the REV primers have been able to anneal to. This does not, however, prove conclusively that the PD clone contains a cricket *ptc* homologue. From the data available, I would suggest that the PD clone does contain REV-like sequence, but is probably not a *ptc* orthologue.

5.3.2. Interpretation of the PD expression pattern.

Given the data discussed in the previous sections, it has not been possible to demonstrate unequivocally that PD contains a fragment of a cricket *ptc* homologue. Therefore, one of the major questions to discuss here is; 'From the presented expression pattern, what processes could the PD clone be involved in?'

Before addressing this question, it is necessary to make one major assumption, which is that the PD insert is derived from a single transcription unit, (because if it was derived from more than one transcription unit it would be practically impossible to interpret the expression pattern).

The PD expression data presented here shows several points of note;

(1) The expression pattern of PD is complex and dynamic.

(2) PD is expressed in every trunk segment at some stage of embryogenesis.

(3) The expression pattern of PD is different in the thoracic and abdominal segments.

(4) PD is also expressed in the developing appendages.

Given this, it is possible that PD plays some role in the segmental patterning of the cricket embryo.

5.3.2.1. Segmental patterning in the cricket and locust.

The way that the body segments are formed in *Acheta* is very different to the segmentation of the *Drosophila* embryo. *Drosophila* is a long germ band insect and forms its segments by the almost simultaneous subdivision of the cellular blastoderm into repeating segmental units. These body segments only have rudimentary pattern information when they are first formed, i.e., the initial expression patterns of segment polarity genes, which requires enhancing and refining (see Chapter One).

Acheta domesticus has been classified as an intermediate germ band insect, because at cellularisation of the syncitial blastoderm the primordia of the head and the thoracic segments are already determined, and the abdominal segments are then generated sequentially from a posterior growth zone. Given this, it may be that the patterning of the thoracic segments is similar to the patterning of *Drosophila* segments (as they result from the subdivision of an already present field of cells, the relatively undifferentiated cellular blastoderm, rather like the segments of *Drosophila*). However, the segments of the abdomen may be patterned in a completely different way. It may be possible that, as the abdominal segments are formed, they have some degree of temporally regulated information imparted to them during the proliferation of the growth zone.

One of the few genes whose expression has been examined during *Acheta* embryogenesis is *engrailed*. Using the monoclonal antibody, 4D9, Patel (1994 and *pers*

communication), and I (data not shown) have shown that *en* is expressed in the

extending germ band in thin stripes in the posterior of each of the trunk segments.

These stripes appear in an anterior to posterior sequence down the abdomen just prior to

visible signs of segmentation in each developing segment. At the extending germ band

stage, *en* is also expressed in the posterior compartment of the developing appendages,

and later, in the fully extended germ band, is also expressed in a subset of neurons.

This data demonstrates that although the final pattern of *en* expression is very similar in

the cricket and *Drosophila* (see Chapter One), the way in which the pattern is generated

is quite different. This basis of this hypothesis is that in the cricket, and the locust, the

stripes of *en* appear without the initial pair-rule modulation seen in *Drosophila* (Patel,

1989). This data suggests that the regulation of *en* by the pair-rule genes seen in

Drosophila is not conserved in the short and intermediate germ insects (Patel, 1989;

Patel, 1993). This hypothesis is also supported by the finding that the locust

homologues of *eve* (Patel *et al*, 1992) and *ftz* (Dawes et al, 1994) do not have a phase of

pair-rule expression.

From the limited available information, it would seem likely that the interactions

between *en*, *wg*, *hh*, and *ptc* seen in *Drosophila*, and now several vertebrate species,

would be conserved in the cricket. The PD expression data presented here, however,

does not show the spatial relationship with *en*, seen in *Drosophila ptc*. Given the degree

of conservation of this spatial relationship seen in other species, the PD expression

pattern is more evidence in favour of PD not being the cricket *ptc* orthologue.

5.3.2.2. PD expression in the trunk segments.

The expression pattern of PD does not indicate that it fits into any of the categories of

segmentation genes i.e., the gap, pair-rule, or segment polarity genes.

In the thoracic segments, PD transcript is initially found in a pattern of stripes adjacent

to the segment borders, which matures into a more wedge-like domain expression (see

Figure 23). The early abdominal pattern shows segment-specific variation depending on

the relative ages of the segments. In Figure 23a, which shows a stage 18 embryo, it can

be seen that the pattern of expression of PD in the more posterior, and hence younger,

segments resembles a reduced version of the pattern seen in the more anterior (older)

segments. Later in development it can be seen that the pattern of expression in all of the

abdominal segments is spatially very similar. This would argue for there being a

temporally controlled activation of PD expression, and the same mechanism of

regulation of PD in each abdominal segment, which results in the development of the

mature, spatially restricted, pattern. This expression pattern suggests that in the thoracic

segments, PD may have a function at the segment border, but its major function could

be along the midline, as this component of the expression pattern is common to both the thoracic and abdominal segments.

Some of the midline expression is likely to be in the developing neuroblasts, although with the intensity and density of the staining observed, it was impossible to discern whether the expression was restricted to a specific subset of cells. It is, therefore, possible that PD may have a neurogenic function.

PD transcript was first detected in this study at stage 18, but it is possible that this is not the earliest stage at which it is expressed. Indeed from the evolution of the abdominal pattern it seems likely that it would be expressed at earlier stages, quite possibly from soon after cellularisation of the blastoderm. If PD is expressed earlier than stage 18, it would seem reasonable to assume that, as there is a considerable change in the morphology of the embryos during embryogenesis, the protocols used to visualise PD expression were not properly optimised for detection of the PD transcript in the early stages of embryogenesis.

5.3.2.3. PD expression in the developing cricket appendages.

In the cricket, the appendages develop directly from buds during embryogenesis and continue to develop through larval life, and these buds are thought to be directly comparable with the imaginal discs of holometabolous insects. From the expression of

PD seen in the developing appendages at the stages of embryogensis examined here, it is hard to see what process(es) PD may have a role in, although the pattern of expression in the developing cricket leg is somewhat similar to that of *aristaless* in *Drosophila* (Campbell *et al*, 1993). *aristaless* has previously been implicated in the establishment of the proximodistal axis in the developing appendages of *Drosophila*, and is expressed transiently in the thoracic segments in cells expressing both *wingless* and *decapentaplegic*, before reappearing later in the third instar imaginal discs, in the presumptive distal regions and the regions of the notum (in the wing disc), and the coxa (in the leg disc).

From examination of the expression pattern of PD during cricket embryogenesis it is impossible to unequivocally deduce the roles that it may be playing. It is again unfortunate that the cricket is not amenable to the same sorts of genetic manipulations that are commonplace in *Drosophila*, because it will be exceptionally difficult to unravel the roles of genes, such as PD, without them.

5.3.3. Future work.

The first goal would be to identify the PD insert. This would best be done by completely sequencing the PD insert. Analysis of the resulting sequence data and the predicted protein structure would identify the PD insert, either as a novel gene, or as a

homologue of a previously identified gene, and would provide a basis for the interpretation of the expression pattern. The expression series of PD during *Acheta* development would also require completion. This may shed some light on the functions PD has during cricket embryogenesis. Further characterisation of PD would depend on the results gained from these experiments.

The major aim of this part of the study was to clone and characterise the cricket homologue of *ptc*. If, as I have argued, PD is not the cricket *ptc* homologue, this is still to be done. The most efficient way to do this would be by screening cricket DNA libraries, either with *ptc* probes generated from other *ptc* homologues, or the cricket REV fragment. Cloning of the cricket *ptc* homologue would be a good starting point for investigating whether the Hedgehog pathway is conserved in the cricket, and how it is utilised during cricket development. This would be a long term project, and would use all of the techniques available for comparative molecular embryology. Having sequence data for cricket *ptc* would also allow the construction of molecular phylogenies based on the *ptc* homologues, although how much novel information this would afford is unclear.

6. Summary and conclusions.

The data presented in this thesis demonstrates that the house fly, *Musca domestica*, and the house cricket, *Acheta domesticus*, both possess homologues of the *Drosophila melanogaster* gene, *patched*, and that these homologues are expressed during the early development of both species.

6.1. *Musca domestica*.

Characterisation of the partial *Musca ptc* homologue, *mdptc*, indicated that it was very similar to the *Drosophila ptc* gene at the nucleotide, and amino acid levels across comparable regions. Homology at the amino acid level was greatest across the transmembrane domains, ranging from 52 to 100% identity, or 77 to 100% similarity, when conservative substitutions were allowed. When only the full Ptc protein sequences were used to generate a phylogeny, it was shown that the *Drosophila* species formed a group, as did the vertebrate species, as would be expected from traditional phylogenies based on morphology, 5S or 18S RNA (Hori, 1975), or cytochrome C (Dayhoff *et al.*, 1972) sequences. From the level of homology between the Mdptc protein sequence and that of the other Ptc protein sequences from other species, it would have been placed in the Dipteran group.

The expression of *mdptc* was examined using both RNA *in situ* hybridisation, to

visualise the distribution of *mdptc* transcript, and immunohistochemistry, to show the

protein distribution. In the *Musca* embryo, *mdptc* expression was investigated by RNA

in situ hybridisation, as the antibody, mAb 5E10, did not give clear signals in these

embryos. This is probably due to the antibody having been raised against the N-terminal

of the *D. melanogaster* Ptc protein which has been shown to be less well conserved than

other parts of the protein (Concordet *et al.*, 1996; Forbes, 1995; Goodrich *et al.*, 1996;

Marigo *et al.*, 1996).

The embryonic expression of *mdptc* followed a very similar pattern to that of *ptc* in

Drosophila. At the blastoderm stage, *mdptc* transcript was found between 10 and 90%

EL. By the time the cellular blastoderm had formed, *mdptc* expression had resolved into

a series of stripes throughout the future segmented trunk region of the embryo. As

embryogenesis proceeded the pattern of expression matured into the final pattern seen in

the retracted germ band stage, which was two thin stripes of expression in each trunk

segment, the anterior stripe of each pair being much less intense than the posterior.

Larval expression of *mdptc* was examined, both with RNA *in situ*, and

immunohistochemistry, and compared to the distribution of Ptc protein in the imaginal

discs of *D. melanogaster* and *D. virilis*. The pattern of expression in both *Musca* and *D.

virilis* was the same as seen in *D. melanogaster*. In the thoracic third instar imaginal

discs, expression was seen in a thin stripe which appeared to run along the anterior side

of the compartment border. This was concluded partly by comparison of the *mdptc* expression pattern in the wing disc, and that of *engrailed* at the same stage, which occupied the whole of the posterior compartment. It was already known that *ptc* was expressed along the anterior length of the compartment border in *D. melanogaster* (Phillips *et al.*, 1990). The morphology of the third instar wing discs of *Musca* and *Drosophila* are very similar. The similarity of the expression patterns of *ptc* in these two species reinforced the probability that *mdptc* expression in the wing discs was adjacent to the compartment border. In the eye disc the situation was slightly different; *mdptc* expression was seen on both sides of the morphogenetic furrow, as in the *Drosophila* species. *patched* has previously been shown to be a negative regulator of the morphogenetic furrow in the developing eye disc of *D. melanogaster* (Ma and Moses, 1995), and the expression patterns in the three Dipteran species suggest that this could be true in *Musca* and *D. virilis* as well. Given the morphological similarity between the early development of these three closely related Dipteran species, and the degree of conservation of the hedgehog pathway in diverse species, it was as predicted to find the expression patterns of their *ptc* homologues to be very similar, as it was thought that all three species would utilise the same developmental programs in the generation of the adult organism.

6.2. *Acheta domesticus.*

A PCR fragment was amplified from *Acheta* cDNA using the REV primers which amplify around 350 base pairs of *ptc* from *Drosophila*, and a number of vertebrate species. The *Acheta* fragment was shown to be homologous to the REV fragments from the other species. A phagemid containing a 1.4 Kb insert was isolated from an *Acheta* cDNA library, and was subsequently shown to contain a 345 bp sequence that was amplified using the REV primers. Comparison of the REV fragments from various species showed that the *Acheta* REV fragment was slightly more homologous to the Diptera (approximately 58% identity) than it was to the vertebrate sequences (approximately 45% identity). The level of homology between the insect species and the vertebrates are approximately equivalent (approximately 45% identity), and hence the *Acheta* REV fragment was concluded to be a fragment of the *Acheta ptc* homologue, *aptc*.

The expression pattern of *aptc* was examined using RNA *in situ* hybridisation. The antibody, mAb 5E10, did not give a clear signal in the cricket. The expression pattern of *aptc* did not suggest that it had a segment polarity type function during cricket embryogenesis. Expression of *aptc* was not detected in segments until after the first overt signs of segmentation were already visible in those segments. *aptc* was expressed in a wedge shape in the thoracic segments, primarily along the anterior segment border, but also in the cells in the posterior of the segment which would also be expressing

engrailed. This would suggest that the regulation of *aptc* is different to the regulation of *ptc* in *Drosophila*, as *engrailed* has been shown to directly repress *ptc* expression, and *ptc* is not expressed in cells that also express *engrailed*. *aptc* appeared in the abdominal segments in an anterior to posterior sequence, and was expressed primarily along the midline in a stripe that widened during embryogenesis. *Acheta* is an intermediate germ band insect, which means that the posterior segments are formed sequentially during embryogenesis, which could explain the temporal appearance of the abdominal domains of *aptc* expression. Although the domains in which *aptc* was expressed did not indicate a segment polarity type function for *aptc* during *Acheta* embryogenesis, the early appearance of *aptc* may indicate a role in early development.

Limb development in *Acheta* is morphologically different to that of *Drosophila*. *Acheta* is a hemimetabolous insect, and as such does not develop its limbs from imaginal discs, like *Drosophila*, but rather from limb buds, similar to the vertebrates. *aptc* is expressed in the developing limb buds of the *Acheta* embryo, but not in a pattern that would suggest it plays a role in imparting an anteroposterior polarity to the limbs as it does in the *Drosophila* wing disc (Ingham, 1995), and vertebrate limb buds (Marigo *et al.*, 1996). *aptc* is expressed in discrete domains in the developing cricket leg, at both the proximal and distal ends, and in a position in that would correspond to the distal tibia.

6.3 Discussion and conclusions.

The segment polarity gene, *patched*, is an integral part of the hedgehog intercellular

signalling pathway, which is known to be well conserved between very diverse species.

The structure of the predicted protein has been shown to be poorly conserved between

the insects and the vertebrates, with approximately 40% identity at the amino acid level,

although the topology of the protein, as predicted by hydropathy plots, is well

conserved.

ptc homologues have been isolated from three Dipteran species, *D. melanogaster*

(Hooper and Scott, 1989; Nakano *et al.*, 1989), *D. virilis* (Forbes, 1995), and *M.*

domestica (this thesis). These species have been diverging for the last 100 million

years, yet there is a very high degree of similarity in the mode of early development

exhibited by each, and seemingly the molecular mechanisms controlling it. The

expression of *ptc* homologues in the Diptera suggest that the mechanisms of early

development in long germ insects has been conserved.

The homologues of a few of the *Drosophila* segmentation genes have been investigation

in few other long germ band insects, such as *Apis mellifera*, and *Callosobruchus*

maculatus. These studies have indicated that genes such as *even-skipped* and *engrailed*

are expressed in these insects in a manner similar to their expression in *Drosophila*.

This has led to the conclusion that the molecular basis of long germ band embryogenesis

has been conserved throughout the insect species that exhibit this mode of development. In order to determine whether this is globally true for the long germ band insects, the mechanisms of segmentation will have to be investigated more thoroughly in a number of insect species that span all of the orders that exhibit long germ band embryogenesis.

The situation is, however, different in the short germ insects. Investigation of segmentation in the locust has shown that only *engrailed* is expressed in a pattern similar to that seen in Drosophila. The pattern of a single stripe in the posterior of each body segment is conserved, although it is generated in two different ways due to the manner in which the body segments are formed. In *Drosophila*, the segments are formed simultaneously when the syncitial blastoderm cellularises, whereas the locust embryo develops most of its body segments sequentially through the proliferation of a posterior growth zone. Examination of expression patterns of the homologues of the pair-rule genes, *even-skipped* (Patel *et al.*, 1992) and *fushi tarazu* (Dawes *et al.*, 1994)in the locust has revealed that there does not appear to be a pair-rule mechanism generating the segmental pattern in these insects. It is, of course, possible that the locust represents an extreme example of short germ embryogenesis. Indeed, Patel (1994b) showed that there is a pair-rule phase of expression in the short germ beetle, *Tribolium castaneum*, and has suggested that the number of pair-rule stripes present at the onset of gastrulation be used to define germ band types, rather than traditional morphological criteria.

Investigation of the *Acheta ptc* homologue can not answer the question of whether pair-rule patterning is a common phenomenon in insect segmentation. It does, however, indicate that the segment polarity mechanism that is necessary for the patterning of segments in *Drosophila* may not function in the same way in *Acheta* embryos. In *Drosophila* embryos, the segments are defined by the action of the pair-rule genes, and then patterned by the segment polarity genes. When these segments are first formed they have some pattern inherent from the patterns of pair-rule gene expression. The combined action of the products of *engrailed*, *wingless*, *patched*, and *hedgehog* genes, as well as others, are required to refine, and maintain, the pattern across the *Drosophila* segment. In *Acheta* embryos, only the thoracic segments are formed directly from the cellularisation of a syncitium, the abdominal segments are formed sequentially as in the locust. The pattern of *aptc* expression in the thoracic segments is not a typical segment polarity pattern, although it is expressed along the segment borders. The segmental expression of *aptc* in the abdominal segments is reminiscent of the expression of the vertebrate *ptc* homologues, where it is expressed in the paraxial mesoderm, the neurectoderm, and the presomitic mesoderm. Due to the way in which the segments are formed in *Acheta*, it is possible that intra-segmental pattern could be generated without the need of interaction between the pair-rule and segment polarity genes, with domains of segment polarity gene expression being defined temporally, rather than through spatial interactions.

Expression of *aptc* in the developing legs of the *Acheta* embryos is unlike the expression of any of the other homologues during appendage development. In Drosophila, and the vertebrates, the hedgehog pathway is involved in creating an anterior-posterior pattern in the developing appendages. The small discrete domains of *aptc* expression in the cricket leg suggest that *aptc* would be incapable of playing a role in creating this kind of anteroposterior pattern.

These data indicate that the common ancestor of the insects and the vertebrates had a *patched* like gene. This organism would most likely have been a limb-less invertebrate. Given this, and the expression pattern of *aptc* in the developing cricket, it is tempting to speculate that *ptc* function in this ancestral organism was to patterning the neural tube and somites, as in the present vertebrates. This would imply that *patched* (and the hedgehog pathway), in the insects, has been co-opted into a role in segmentation, early after the divergence of the insects, and again into a role in limb patterning in the Diptera, after they diverged away from the Orthoptera. The *ptc* homologues would then have been co-opted, independently, into a role in limb patterning in the vertebrates after the divergence of the vertebrate and invertebrate lineages. To confirm this it would be necessary to study the hedgehog pathway in a number of other organisms, including members of the more ancient insect orders such as the Odonata, and species from other arthropod classes such as the Crustacea, and Myriapoda.

Appendices.

A. Liquid media.

1. LB (Luria-Bertani) medium

Composition	Per litre
bacto-tryptone	10 g
bacto-yeast extract	5 g
NaCl	10 g
Deionised H_2O	950 ml
pH 7.0	

LB agar: for bacteria propogation and bottom plates add 15 g/l agar to LB medium
LB agarose: for top plates add 7g/l agarose to LB medium
LB ampicillin agar: as for LB agar plus 50 mg filter-sterilised ampicilin per litre.

2. SOB medium

Composition	Per litre
bacto-tryptone	20 g
bacto-yeast extract	5 g
NaCl	0.5 g
Deionised H_2O	950 ml
KCl (250 mM)	10 ml
$MgCl_2$ (2 mM)	5 ml
pH 7.0	

B. Buffers and reagents.

1. Tris Buffers

	Composition
TAE	0.04 M Tris-acetate 0.01 M EDTA
TBE	0.09 M Tris-phosphate 0.002 M EDTA
TE	10 mM Tris.Cl 1 mM EDTA
TES	10 mM Tris.Cl 1 mM EDTA 0.2 % SDS

2. PBS (Phosphate buffered saline)

Composition	Per litre
NaCl	8 g
KCl	0.2 g
Na_2PO_4	1.44 g
KH_2PO_4	0.24 g
pH 7.4	

3. SSC (20x)

Composition	Per litre
NaCl	175.3 g
sodium citrate	88.2 g
pH 7.0	

253

C. Bacterial strains.

Strain	Genotype
XLI-Blue	*supE44hsdR17recA1endA1gyrA46thi relA1lac⁻ F'[proAB⁻ lacIᵠlacZΔM15Tn10(t*
XL1-Blue MRA	*Δ(mcrA)183 Δ(mrcCB-hsdSMR-mrr)173endA1 supE44 thi-1 gyrA96 relA1 lacᶜ*
XL1-Blue MRF'	*Δ(mcrA)183 Δ(mrcCB-hsdSMR-mrr)173endA1 supE44 thi-1 gyrA96 relA1 lacᶜ* | *lacIᵠZΔM15* Tn*10* (Tetʳ]ᶜ
SOLR	*e14⁻(McrA⁻) Δ'(mcrCB-hsdSMR-mrr)171sbcC recB recJ uvrC::*Tn5 *(Kanʳ) lac* *thi-1 endA1 λᴿ [F' proABlacᵠZΔM15]ᶜ* Su⁻

D. PCR Primers.

(Cloning sequences are underlined)

1. REV Primers (Goodrich *et al.*, 1996).

P4REV: GGA CGA ATT CYT NGA NTG YTT YTG GGA

P22:   CAT ACC AGC CAA GCT TGT CIG GCC ART GCA T

2. Primers used to clone the *mdptc* hypervariable region.

M2.6:   5' GAC GGA TCC GCA AGC GAA AAG GAA GGT AAG

M2.8:   5' GAC CTC GAG TGT TGT TGT TTT CCC ATA TGG

3. Various cloning primers.

musca 3'PCRev: GGA GAA TTC AAT TAG AAT ACT TCA AAG GTT T

cricket 3' PCR (A3'): GGA CTC GAG NCC YTG NGT NAC NGC RTA CAT

Rev4 alternat: GGA GGA TCC YTN GAY TGY TTY GGG A

4. IUB codes.

A = Adenosine   R = A or G
C = Cytidine   Y = C or T
G = Guanosine   K = G or T
T = Thymidine   M = A or C

B = C,G, or T   S = G or C
D = A,G, or T   W = A or T
H = A,C, or T   N = aNy base
V = A,C, or G   I = Inositol

References.

Abdel-Meguid, S., Shieh, H., Smith, W., Dayringer, H., Violand, B., and Bentle, L. (1987). Three-dimensional structure of a genetically engineered variant of porcine growth hormone. Proc Natl Acad Sci USA *84*, 6434-6437.

Alcedo, J., and Noll, M. (1997). Hedgehog and its patched-smoothened receptor complex: a novel signalling mechanism at the cell surface. Biol Chem *378*, 583-590.

Alcedo J, A. M., Von Ohlen T, Noll M, Hooper JE. (1996). The Drosophila smoothened gene encodes a seven-pass membrane protein, a putative receptor for the hedgehog signal. Cell *86*, 221-232.

Alexandre, C., Jacinto, A., and Ingham, P. (1996). Transcriptional activation of hedgehog target genes in Drosophila is mediated directly by the cubitus interruptus protein, a member of the GLI family of zinc finger DNA-binding proteins. Genes Dev *10*, 2003-2013.

Anderson, D. (1972). The development of hemimetabolous insects. In Developmental Systems Insects: Counce, S.J., and Waddington, C.H., eds (Academic Press: London and New York).

Anderson, K., Bokla, L., and Nusslein-Volhard, C. (1985b). Establishment of dorsal-ventral polarity in the Drosophila embryo; the induction of polarity by the Toll gene product. Cell *42*, 791-798.

Anderson, K., Jurgens, G., and Nusslein-Volhard, C. (1985a). Establishment of dorsal-ventral polarity in the Drosophila embryo; genetic studies on the role of the Toll gene product. Cell *42*, 779-789.

Argos, P., and Rao, J. (1986). Prediction of protein structure. Methods Enzymol *130*, 185-207.

Baker, N. (1987). Molecular cloning of sequences from wingless, a segment polarity gene in Drosophila; the spatial distribution of a transcript in embryos. EMBO J *6*, 1765-1773.

Baker, N. E. (1988a). Embryonic and imaginal requirements for wingless a segment polarity gene in Drosophila. Dev Biol. *125.*, 96-108.

Basler, K., and Struhl, G. (1994). Compartment boundaries and the control of Drosophila limb pattern by hedgehog protein. Nature *368*, 208-214.

Bate, M., and Martinez-Arias, A. (1991). The embryonic origin of imaginal discs in Drosophila. Development. *112*, 755-761.

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

Bejsovec, A., and Martinez-Arias, A. (1991). Roles of Wingless in patterning the larval epidermis of Drosophila. Development *113*, 471-485.

Berleth, T. (1988). The role of localisation of bicoid RNA in organising the anterior pattern of the Drosophila embryo. EMBO J *7*, 1749-1756.

Bhanot, P., Brink, M., Samos, C., Hsieh, J., Wang, Y., Macke, J., Andrew, D., Nathans, J., and Nusse, R. (1996). A new member of the frizzled family from Drosophila functions as a Wingless receptor. Nature *382*, 225-230.

Binner, P., and Sander, K. (1997). Pair-rule patterning in the honeybee Apis mellifera; expression of even-skipped combines traits known from beetles and fruitfly. Dev Genes Evol. *206*, 447-454.

Blackman, R., and Meselson, M. (1986). Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the Drosophila hsp82 gene. J Mol Biol *188*, 499-515.

Brandt, A. (1874). Uber die Eirohren der Blatta (Periplaneta) orientalis. Mem.Acad

Imp Sci St Petersbourg. *21*, 1-30.

Bronner, G., and Jackle, H. (1991). Control and function of terminal gap gene

activity in the posterior pole of the Drosophila embryo. Mech Dev *35*, 205-211.

Brower, D. (1986). Engrailed gene expression in Drosophila imaginal discs. EMBO J

5, 2649-2656.

Brown, S., Hilgenfeld, R., and Denell, R. (1994a). The beetle Tribolium castaneum

has a fushi tarazu homolog expressed in stripes during segmentation. Dev Biol *91*,

12922-12926.

Brown, S., Patel, N., and Denell, R. E. (1994b). Embryonic expression of the single

Tribolium engrailed Homolog. Dev Genet., 07-18.

Brunner, E., Peter, O., Schweizer, L., and Basler, K. (1997). Pangolin encodes a Lef-

1 homologue that acts downstream of armadillo to transduce the wingless signal in

Drosophila. Nature *385*, 829-833.

Cadigan, K., Grossniklaus, U., and Gehring, W. (1994b). Localized expression of

sloppy paired protein maintains the polarity of Drosophila parasegments. Genes Dev

15, 899-913.

Campbell, G., Weaver, T., and Tomlinson, A. (1993). Axis specification in the developing Drosophila appendage: the role of wingless, decapentaplegic, and the homeobox containing gene aristaless. Cell 74, 1113-1123.

Campos-Ortega, J., and Hartenstein, V. (1985). The embryonic development of Drosophila melanogaster, (Berlin: Springer-Verlag).

Capdevila, J., Estrada, M., Sanchez-Herrero, E., and Guerrero, I. (1994). The Drosophila segment polarity gene patched interacts with DPP in wing development. EMBO J 13, 71-82.

Capovilla, M., Eldon, E., and Pirrotta, V. (1992). The giant gene of Drosophila encodes a b-ZIP DNA binding protein that regulates the expression of the other segmentation gap genes. Development 114, 99-112.

Carroll, S., and Scott, M. (1986). Zygotically active genes that affect the spatial expression of the fushi-tarazu segmentation gene during early Drosophila embryogenesis. Cell 45, 673-683.

Carroll, S., and Vavra, S. (1989). The zygotic control of Drosophila pair rule gene expression. Development 107, 673-683.

Cave, M., Benes, H., and Liarakos, C. (1987). Nucleotide sequence of two repeating units of the 5S rRNA gene from the house cricket Acheta domesticus. Gene *51*, 287-289.

Cohen, B., Simcox, A., and Cohen, S. (1993). Allocation of the thoracic imaginal primordia in the Drosophila embryo. Development *117*, 597-608.

Cohen, S. (1993). Imaginal disc development. *In*: The embryonic development of Drosophila melanogaster (vol II) Bate, M., Martinez-Arias, A., eds (Cold Sping Harbor laboratory Press: New York).

Concordet, J., Lewis, K., Moore, J., Goodrich, L., Johnson, R., Scott, M., and Ingham, P. (1996). Spatial regulation of a zebrafish patched homolog reflects the roles of sonic hedgehog and protein kinase A in neural tube and somite patterning. Development *122*, 2835-2846.

Couso, J., and Gonzalez-Gaitan, M. (1993). Embryonic limb development in Drosophila. TIGS *09*, 371-373.

Curtis, D., Apfeld, J., and Lehmann, R. (1995). nanos is an evolutionarily conserved organizer of anterior-posterior polarity. Development *121*, 1899-1910.

Dawes, R., Dawson, I., Falciani, F., Tear, G., and Akam, M. (1994). Dax, a locust Hox gene related to fushi-tarazu but showing no pair-rule expression. Development *120*, 1561-1572.

Dayhoff, M., Park, C., and McLaughlin, P. (1972). Building a phylogenetic tree: cytochrome c. *In*: Atlas of protein sequence and structure, ed M. Dayhoff. (Washington DC: National Biomedical Research Foundation,) pp. 7-16.

Dearolf, C., Topol, J., and Parker, C. (1989a). Transcriptional control of Drosophila fushi tarazu zebra stripe expression. Genes Dev *3*, 384-398.

Dearolf, C., Topol, J., and Parker, J. (1989b). The caudal gene product is a direct activator of fushi-tarazu transcription during Drosophila embryogenesis. Nature *341*, 340-343.

DiNardo, S., Kuner, J., Theis, J., and O, F., P (1985). Development of embryonic pattern in D. melanogaster as revealed by accumulation of the nuclear engrailed protein. Cell *43*, 59-69.

DiNardo, S., Sher, E., Heemskerk-Jongens, J., Kassis, J., and O'Farrell, P. (1988). Two-tiered regulation of spatially patterned engrailed gene expression during Drosophila embryogenesis. Nature *332*, 604-609.

Doe, C., Hiromi, Y., Gehring, W., and Goodman, C. (1988a). Expression and

function of the segmentation gene fushi tarazu during Drosophila neurogenesis.

Science *239*, 170-175.

Doe, C., Smouse, D., and Goodman, C. (1988b). Control of neuronal fate by the

Drosophila segmentation gene even-skipped. Nature *333*, 376-378.

Dougan, S., and DiNardo, S. (1992). Drosophila wingless generates cell type

diversity among engrailed expressing cells. Nature *360*, 347-350.

Driever, W., and Nusslein-Volhard, C. (1988b). The bicoid protein determines

position in the Drosophila embryo in a concentration dependent manner. Cell *54*, 95-

104.

Driever, W., and Nusslein-Volhard, C. (1989). The bicoid protein is a positive

regulator of hunchback transcription in the early Drosophila embryo. Nature *337*,

138-143.

Driever, W., and Nusslein-Volhard, C. (1988a). A gradient of bicoid protein in

Drosophila embryos. Cell *54*, 95-104.

Echelard, Y., Epstein, D., St-Jacques, B., Shen, L., Mohler, J., McMahon, J., and

Mcmahon, A. (1993). Sonic hedgehog, a member of a family of putative signaling

molecules,is implicated in the regulation of CNS polarity. Cell *75*, 1417-1430.

Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J Mol Biol *179*, 125-142.

Ephrussi, A., and Lehmann, R. (1992). Induction of germ cell formation by oskar. Nature *358*, 387-392.

Epps, J., Jones, J., and Tanda, S. (1997). oroshigane, a new segment polarity gene of Drosophila melanogaster, functions in hedgehog signal transduction. Genetics *45*, 1041-1052.

Farris, J. (1972). Estimating phylogenetic trees from distance matrices. American Nature *106*, 645-668.

Felsenfield, A., and Kennison, J. (1995). Positional signaling by hedgehog in Drosophila imaginal disc development. Development *121*, 1-10.

Feng, D., and Doolittle, R. (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J Mol Evol *25*, 351-360.

Field, K., Olsen, G., Lane, D., Giovannoni, S., Ghiselin, M., Raff, E., Pace, N., and Raff, R. (1988). Molecular phylogeny of the animal kingdom. Science *239*, 748-753.

Fietz, M., Concordet, J., Barbosa, R., Johnson, R., Krauss, S., McMahon, A., Tabin, C., and Ingham, P. (1994). The hedgehog gene family in Drosophila and vertebrate development. Development, 43-51.

Fietz, M., Jacinto, A., Taylor, A., Alexandre, C., and Ingham, P. (1995). Secretion of the amino-terminal fragment of the hedgehog protein is necessary and sufficient for hedgehog signaling in Drosophila. Curr Biol *5*, 643-650.

Fitch, W., and Margoliash, E. (1967). Construction of phylogenetic trees. Science *15*, 279-284.

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

Fleig, R. (1990). Engrailed expression and body segmentation in the honeybee Apis Mellifera. Wilhelm Roux Arch Dev Biol. *198*, 467-473.

Fleig, R., and Sander, K. (1988). Honeybee morphogenesis; embryonic cell movements that shape the larval body. Development *103*, 525-534.

Forbes, A. (1995). PhD Thesis, University of Oxford.

Frasch, M., and Levine, M. (1987). Complementary patterns of even skipped and Fushi tarazu expression involve their differential regulation by a common set of segmentation genes in Drosophila. Genes Dev *1*, 981-995.

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

Fristrom, D., and Fristrom, J. (1993). The metamorphic development of the adult epidermis. In The Development of Drosophila melanogaster: Bate, M, and Martinas-Arias, A,. (Cold Spring Harbor Laboratory Press, New York), pp. 843-897.

Fujioka, M., and Jaynes, J. (1995). Early even skipped stripes act as morphogenic gradients at the single cell level to establish engrailed expression. Development *121*, 4371-4382.

Gaul, U., and Jackle, H. (1987). Pole region-dependent repression of the Drosophila gap gene Kruppel by maternal gene products. Cell *51*, 549-555.

Gavis, E., and Lehmann, R. (1992). Localization of nanos RNA controls embryonic polarity. Cell *71*, 301-313.

Gergen, J., and ButlerBA (1988). Isolation of the Drosophila segmentation gene runt and analysis of its expression during embryogenesis. Genes Dev *2*, 1178-1193.

Goodrich, L., Johnson, R., Milenkovic, L., McMahon, J., and Scott, M. (1996). Conservation of the hedgehog/patched signaling pathway from flies to mice - Induction of a mouse patched gene by hedgehog. Genes Dev *10*, 301-312.

Goto, T., Macdonald, P., and Maniatis, T. (1989). Early and late periodic patterns of even-skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. Cell *57*.

Gould, S., Subramani, S., and Scheffler, I. (1989). Use of the DNA polymerase chain reaction for homology probing: isolation of partial cDNA or genomic clones encoding the iron-sulfur protein of succinate dehydrogenase from several species. Proc Natl Acad Sci USA *86*, 1934-1938.

Hafen, E., Kuroiwa, A., and Gehring, W. (1984). Spatial distribution of transcripts from the segmentation gene fushi tarazu during Drosophila embryonic development. Cell *37*, 833-841.

Hashimoto, C., Hudson, K., and Anderson, K. (1988). The toll gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell *52*, 269-279.

Hennig, W. (1981). Insect Phylogeny. Wiley, Chichester.

Hidalgo, A., and Ingham, P. (1990). Cell patterning in the Drosophila segment; spatial regulation of the segment polarity gene patched. Development *110*, 291-301.

Hiromi, Y., and Gehring, W. (1987). Regulation and function of the Drosophila segmentation gene fushi-tarazu. Cell *50*, 963-974.

Hiromi, Y., Kuroiwa, A., and Gehring, W. (1985). Control elements of the Drosophila segmentation gene fushi-tarazu. Cell *43*, 603-613.

Hoch, M., Gerwin, N., Taubert, H., and Jackle, H. (1992). Competition for overlapping sites in the regulatory region of the Drosophila gene Kruppel. Science *256*, 94-97.

Hoch, M., Schroder, C., Seifert, E., and Jackle, H. (1990). Cis-acting control elements for Kruppel expression in the Drosophila embryo. EMBO J *9*, 2587-2595.

Hoch, M., Seifert, E., and Jackle, H. (1991). Gene expression mediated cis-regulatory sequences of the Kruppel gene in response to the anterior morphogens bicoid and hunchback. EMBO J *10*, 2267-2278.

Hooper, J., and Scott, M. (1989). The Drosophila patched gene encodes a putative membrane protein required for segmental patterning. Cell *59*, 751-765.

Hori, H. (1975). Evolution of 5sRNA. J Mol Evol *7*, 75-86.

Hough-Evans, B., Jacobs-Lorenza, M., Cummings, M., Britten, R., and Davidson, E. (1980). Complexity of RNA in eggs of Drosophila melanogaster and Musca domestica. Genetics *95*, 81-94.

Howard, K., and Ingham, P. (1986). Regulatory interactions between the segmentation genes fushi , tarazu, hairy, and engrailed in the Drosophila blastoderm. Cell *44*, 949-957.

Howard, K., Ingham, P., and Rushlow, C. (1988). Region-specific alleles of the Drosophila segmentation gene hairy. Genes Dev *2*, 1037-1046.

Hulskamp, M., Pfiefle, C., and Tautz, D. (1990). A morphogenetic gradient of hunchback protein organises the expression of the gap genes Kruppel and knirps in the early Drosophila embryo. Nature *346*, 577-580.

Hulskamp, M., Schroder, C., Pfeifle, C., and Jackle, H. (1989). Posterior segmentation of the Drosophila embryo in the absence of a maternal posterior organiser gene. Nature *338*, 629-632.

Ingham, P. (1993). Localized hedgehog activity controls spatial limits of wingless transcription in the Drosophila embryo. Nature *366*, 560-562.

Ingham, P. (1988). The molecular genetics of embryonic pattern formation in Drosophila. Nature *335*, 25-33.

Ingham, P. (1995). Signaling by hedgehog family proteins in Drosophila and vertebrate development. Curr Biol *5*, 492-498.

Ingham, P., Baker, N., and Martinez- Arias, A. (1988). Regulation of segment polarity genes in the Drosophila blastoderm by fushi tarazu and even-skipped. Nature *331*, 73-75.

Ingham, P., and Fietz, M. (1995). Quantitative effects of hedgehog and decapentaplegic activity on the patterning of the Drosophila wing. Curr Biol *5*, 432-440.

Ingham, P., and Gergen, P. (1988). Interactions between the pair-rule genes runt, hairy, even skipped and fushi tarazu and the establishment of periodic pattern in the Drosophila embryo. Development *104*, 51-60.

Ingham, P., and Hidalgo, A. (1993). Regulation of wingless transcription in the Drosophila embryo. Development *117*, 283-291.

Ingham, P., Howard, K., and Ish-Horowicw, D. (1985). Transcription pattern of the Drosophila segmentation gene hairy. Nature *318*, 439-445.

Ingham, P., and Martinez-arias, A. (1992). Boundaries and fields in early embryos. Cell *68*, 221-235.

Ingham, P., Martinez-Arias, A., Lawrence, P., and Howard, K. (1985). Expression of engrailed in the parasegment of Drosophila. Nature *317*, 634-636.

Ingham, P., Taylor, A., and Nakiano, Y. (1991). Role of the Drosophila patched gene in positional signaling. Nature *353*, 184-187.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. Gene *96*, 23-28.

Ish-Horowicz, D., and Pincin, S. (1987). Pattern abnormalities induced by ectopic expression of the Drosophila gene hairy are associated with repression of ftz transcription. Cell *51*, 405-415.

Jackle, H., Tautz, D., Schuh, R., Seifert, E., and Lehmann, R. (1986). Cross regulatory interactions among the gap genes of Drosophila. Nature *324*, 668-670.

Jiang, J., and Struhl, G. (1995). Protein kinase-A and hedgehog signaling in Drosophila limb development. Cell *80*, 563-572.

Johnson, R., Grenier, J., and Scott, M. (1995). Patched overexpression alters wing disc size and pattern; transcriptional and post-transcriptional effects on hedgehog targets. Development *121*, 4161-4170.

Jost, W., Yu, Y., Pick, L., Preiss, A., and Maier, D. (1995). Structure and regulation of the fushi-tarazu gene from Drosophila hydei. Wilhelm Roux Arch Dev Bio *205*, 160-170.

Kanellis, A. (1952). Anlagenplan und regulationserscheinungen in der keimanlage des eies von Gryllus domesticus. Wilhelm Roux Arch. EntwMech. Org *145*, 417-461.

Karr, T., Ali, Z., Drees, B., and Kornberg, T. (1985). The engrailed locus of Drosophlia melanogaster provides an essential zygotic function in precellular embryos. Cell *43*, 491-501

Karr, T., Weir, M., Au, Z., and Kornberg, T. (1989). Patterns of engrailed protein in early Drosophila embryos. Development *105*, 605-612.

Kassis, J., Poole, S., Wright, D., and O'Farrell, P. (1986). Sequence conservation in the protein coding and intron regions of the engrailed transcription unit. EMBO J *5*, 3583-3589.

King, R., and Buning, J. (1985). The origin and functioning of insect oocytes and nurse cells. *In*: Comprehensive insect physiology, biochemistry, and pharmacology, eds Kerkut, G, A,. and Gilbert, L, I,. (Pergammon Press, Oxford), p38-82.

Kinzler, K., and Vogelstein, B. (1990). The GLI gene encodes a nuclear protein which binds specific sequences in the human genome. Mol Cell Biol *10*, 634-642.

Klein, P., Kanehisa, M., and DeLisi, C. (1985). The detection and classification of membrane-spanning proteins. Biochim Biophys Acta *815*, 468-476.

Klingensmith, J., Noll, E., and Perrimon, N. (1989). The segment polarity phenotype of Drosophila involves differential tendencies toward transformation and cell death. Dev Biol *134*, 130-145.

Klingensmith, J., Nusse, R., and Perrimon, N. (1994). The Drosophila segment polarity gene dishevelled encodes a novel protein required for response to the wingless signal. Genes Dev *18*, 118-130.

Klinger, M., Erdelyi, M., Szabad, J., and Nusslein-Volhard, C. (1988). Function of torso in determining the terminal anlagen of the Drosophila embryo. Nature *335*, 275-277.

Knipple, D., Seiffert, E., Rosenberg, U., Preiss, A., and Jackle, H. (1985). Spatial and temporal patterns of Kruppel gene expression in early Drosophila embryos. Nature *317*, 40-44.

Kornberg, T., Siden, I., O, F., P, and Simon, M. (1985). The engrailed locus of Drosophila; in situ localization reveals compartment-specific expression. Cell *40*, 45-53.

Kraft, R., and Jackle, H. (1994). Drosophila mode of metamerization in the embryogenesis of the lepidopteran insect Manduca sexta. Proc. Natl. Acad. Sci. USA *91*, 6634-6638.

Krause, G. (1939). Die Eitypen der Insekten. Biol Zentralbl *59*, 495-536.

Kraut, R., and Levine, M. (1991). Mutually repressive interactions between the gap genesgiant and Kruppel define middle body regions of the Drosophila embryo. Development *111*, 611-622.

Kyte, J., and Doolittle, R. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol *5*, 105-132.

Lawrence, P. (1966). Gradients in the insect segment: the orientation of hairs in the milkweed bug, Oncopeltus fasciatus. J Exp Biol *44*, 607-620.

Lecuit, T., Brook, W., Ng, M., Calleja, M., Sun, H., and Cohen, S. (1996). Two distinct mechanisms for long range patterning by decapentaplegic in the Drosophila wing. Nature *381*, 387-393.

Lee, J., Ekker, S., von Kessler, D., Porter, J., Sun, B., and Beachy, P. (1994). Autoproteolysis in hedgehog protein biogenesis. Science *266*, 1528-1537.

Lee, J., Von Kessler, D., Parks, S., and Beachy, P. (1992). Secretion and localised transcription suggest a role in positional signaling for products of the segmentation gene hedgehog. Cell *71*, 33-50.

Lehmann, R., and Nusslein-Volhard, C. (1987). Hunchback, a gene required for segmentation of an anterior and posterior region of the Drosophila embryo. Dev Biol *119*, 402-417.

Lepage, T., Cohen, S., Diaz-Benjumea, F., and Parkhurst, S. (1995). Signal transduction by cAMP-dependent protein kinase A in Drosophila limb patterning. Nature *373*, 711-715.

Li, W., Talavera-Ohlmeyer, J., Lane, M., and Kalderon, D. (1995). Function of protein kinase A in hedgehog signal transduction and Drosophila imaginal disc development. Cell *80*, 553-562.

Libert, F., Parmentier, M., Lefort, A., Dinsart, C., Van Sande, J., Maenhau, t. C., Simons, M., Dumont, J., and Vassart, G. (1989). Selective amplification and cloning of four new members of the G protein-coupled receptor family. Science *244*, 569-572.

Locke, M. (1967). The development of patterns in the integument of insects. Adv in Morph *6*, 33-88.

Lohs-Schardin, M., Sander, K., Cremer, C., Cremer, T., and Zorn, C. (1979). Localized ultraviolet laser microbeam irradiation of early Drosophila embryos: Fate maps based on location and frequencies of adult defects. Dev Biol *68*, 533-545.

Lundquist, A. (1981). A quantitative analysis of mitotic gradients during early development in Calliphora erythrocephala Meig (Diptera). J Morphol *168*, 239-245.

Ma, C., and Moses, K. (1995). Wingless and patched are negative regulators of the morphogenetic furrow and can affect tissue polarity in the developing Drosophila compound eye. Development *121*, 2279-2289.

Ma, C., Zhou, Y., Beachy, P., and Moses, K. (1993). The segment polarity gene hedgehog is required for progression of the morphogenetic furrow in the developing Drosophila eye. Cell *75*, 927-938.

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

Macdonald, P., and Struhl, G. (1986). A molecular gradient in early Drosophila embryos and its role in specifying body pattern. Nature *324*, 537-545.

Mahowald, A. (1972). Oogenesis. In Developmental Systems Insects, S. Counce and C. Waddington, eds. (London and New York: Academic Press), pp. 1-47.

Mahr, E. (1960). Struktur und entwicklungsfunktion des dotter-entoplasmasystems im ei des heimchens (Gryllus domesticus). Wilhelm Roux Arch. EntwMech. Org *152*, 263-302.

Manoukian, A., and Krause, H. (1993). Control of segmental asymmetry in Drosophila embryos. Development *118*, 785-796.

Manoukian, A., Yoffe, K., Wilder, E., and Perrimon, N. (1995). The porcupine gene is required for wingless autoregulation in Drosophila. Development *121*, 4037-4044.

Marigo, V., Scott, M., Johnson, R., Goodrich, L., and Tabin, C. (1996). Conservation in hedgehog signaling induction of a chicken patched homolog by sonic hedgehog in the developing limb. Development *122*, 1225-1233.

Martinez-Arias, A., Baker, N., and Ingham, P. (1988). Role of segment polarity genes in the definition and maintenance of cell states in the Drosophila embryo. Development *103*, 157-170.

Martinez-Arias, A., and Lawrence, P. (1985). Parasegments and compartments in the Drosophila embryo. Nature *313*, 639-642.

Martinez-Arias, A., and White, R. (1988). Ultrabithorax and engrailed expression in Drosophila embryos mutant for segmentation genes of the pair-rule class. Development *102*, 325-338.

McGinnis, W., Levine, M., Hafen, E., Kurolwa, A., and Gehring, W. (1984). A conserved DNA sequence in homeotic genes of the Drosophila Antennapedia and bithorax complexes. Nature *308*, 428-433.

Micchelli, C., Rulifson, E., and Blair, S. (1997). The function and regulation of cut expression on the wing margin of Drosophila: Notch, Wingless and a dominant negative role for Delta and Serrate. Development *124*, 1485-1495.

Mlodzik, M., and Gehring, W. (1987). Expression of the caudal gene in the germ line of Drosophila: formation of an RNA and protein gradient during early embryogenesis. Cell *48*, 465-478.

Mohler, J. (1988). Requirements for hedgehog, a segmental polarity gene in patterning larval and adult cuticle in Drosophila. Genetics *120*, 1061-1072.

Mohler, J., ED, E., and Pirrotta, V. (1989). A novel transcription pattern associated with the segmentation gene giant, of Drosophila. EMBO J *8*, 1539-1548.

Mohler, J., and Vani, K. (1992). Molecular organization and embryonic expression of the hedgehog gene involved in cell-cell communication in segmental patterning of Drosophila. Development *115*, 957-971.

Mullen, J., and DiNardo, S. (1995). Establishing parasegments in Drosophila embryos. Roles of the odd skipped and naked genes. Dev Biol *169*, 295-308.

Mullis, K. (1991). The polymerase chain reaction in an anemic mode: how to avoid cold oligodeoxyribonuclear fusion. PCR Methods Appl *1*, 1-4.

Nagy, L., and Carroll, S. (1994). Conservation of wingless pattering functions in the short germ embryos of Tribolium castaneum. Nature *367*, 460-463.

Nakano, Y., Guerrero, I., Hidalgo, A., Taylor, A., Whittle, J., and Ingham, P. (1989). A protein with several possible membrane-spanning domains encoded by the Drosophila segment polarity gene patched. Nature *341*, 508-513.

Naqvi, K., Gonzalez-Rodriguez, J., Cherry, R., and Chapman, D. (1973).
Developmental compartmentalisation of wing disc of Drosophila. Nature New
Biology *245*, 251-252.

Nauber, U., Pankratz, M., Kienlin, A., Siefert, E., Klemm, U., and Jackle, H. (1988).
Abdominal segmentation of the Drosophila embryo requires a hormone receptor like
protein encoded by the gap gene knirps. Nature *336*.

Neumann, C., and Cohen, S. (1996). A hierarchy of cross-regulation involving
Notch, wingless, vestigial and cut organizes the dorsal/ventral axis of the Drosophila
wing. Development *122*, 3477-3485.

Nikaido, H., and Saier, M. (1992). Transport proteins in bacteria common themes in
their design. Science *258*, 936-942.

Noordermer, J., Klingensmith, J., Perrimon, N., and Nusse, R. (1994). Dishevelled
and armadillo act in the wingless signaling pathway in Drosophila. Nature *367*, 80-
82.

Nothinger, R. (1972). The larval development of imaginal discs. In The Biology of
Imaginal Discs, H. Ursprung and R. Nothinger, eds. (Berlin: Springer-Verlag), pp. 1-
34.

Nusslein-Volhard, C. (1991). Determination of the embryonic axes of Drosophila. Development *Supplement*, 1-10.

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

Nusslein-Volhard, C., Wieschaus, E., and Kluding, H. (1984). Mutations affecting the pattern of larval cuticle in Drosophila melanogaster. Zygotic loci on the second chromesome. Wilhelm Roux Arch Dev Biol *193*, 267-282.

Orenic, T., slusarski, D., Kroll, K., and Holmgren, R. (1990). Cloning and characterization of the segment polarity gene Cubitus Interruptus Dominant of Drosophila. Genes Dev *4*, 1053-1067.

Pan, D., and Rubin, G. (1995). cAMP-dependant protein-kinase and hedgehog act antagonistically in regulating decapentaplegic transcription in imaginal discs. Cell *80*, 543-552.

Pan, D., and Rubin, G. (1995). cAMP-dependent protein kinase and hedgehog act antagonistically in regulating decapentaplegic transcription in Drosophila imaginal discs. Cell *80*, 543-552.

Pankratz, M., Busch, M., Hoch, M., Seifert, E., and Jackle, H. (1992). Spatial control of the gap gene knirps in the Drosophila embryo by posterior morphogen system. Science *255*, 986-989.

Pankratz, M., Hoch, M., Seifert, E., and Jackle, H. (1989). Kruppel requirement for enhancement reflects overlapping gene activities in the Drosophila embryo. Nature *341*, 337-340.

Patel, N., Ball, E., and Goodman, C. (1992). Changing role of even-skipped during the evolution of insect pattern formation.. Nature *357*, 339-342.

Patel, N., Condron, B., and Zinn, K. (1994b). Pair-rule expression patterns of even-skipped are found in both short and long germ beetles. Nature *367*, 429-434.

Patel, N., Kornberg, T., and Goodman, C. (1989a). Expression of engrailed during segmentation in grasshopper and crayfish. Development *107*, 201-212.

Patel, N., Martin-Bianco, E., Coleman, K., Poole, S., Ellis, M., Kornberg, T., and Goodman, C. (1989b). Expression of engrailed proteins in arthropods, annelids and chordates. Cell *58*, 955-968.

Perrimon, N., and Mahowald, A. (1987). Multiple functions of segment polarity genes in Drosophila. Dev Biol *119*, 587-600.

Phillips, R., Roberts, I., Ingham, P., and Whittle, J. (1990). The Drosophila segment polarity gene patched is involved in a position signaling mechanism in imaginal discs. Development *110*, 105-114.

Pignoni, F., Baldarelli, R., Steingrimsson, E., Diaz, R., Pataoutian, A., Merriam, J., and Lengyel, J. (1990). The Drosophila gene tailess is expressed at the embryonic termini and is a member of the steroid receptor superfamily. Cell *62*, 151-163.

Pignoni, F., Steingrimsson, E., and Lengyel, J. (1992). Bicoid and the terminal system activate tailess expression in the early Drosophila embryo. Development *115*, 239-251.

Poole, S., Kauvar, L., Drees, B., and Kornberg, T. (1985). The engrailed locus of Drosophila; structural analysis of an embryonic transcript. Cell *40*, 37-43.

Porter, J., von Kessler, D., Ekker, S., Young, K., Lee, J., Moses, K., and Beachy, P. (1995). The product of hedgehog autoproteolytic cleavage active in local and long range signaling. Nature *374*, 363-366.

Preat, T., Therond, P., Lamour-Isnard, C., Limbourg-Bouchon, B., Tricoire, H., Erk, I., Mariol, M., and Busson, D. (1991). A putative serine/threonine protein kinase encoded by the segment-polarity fused gene of Drosophila. Nature *347*, 87-89.

Reagan, J. (1996). Molecular cloning and function expression of a diuretic hormone receptor from the house cricket, Acheta domesticus. Insect Biochem. Mol. Biol 26, 1-6.

Riese, J., Yu, X., Munnerlyn, A., Eresh, S., Hsu, S., Grosschedl, R., and Bienz, M. (1997). LEF-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic. Cell 88, 777-787.

Rijsewijk, F., Schuermann, N., Wagenaar, E., Parren, P., Weigel, D., and Nusse, R. (1987). The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. Cell 50, 649-657.

Rivera-Pomar, R., and Jackle, H. (1996). From gradients to stripes in Drosophila embryogenesis - filling in the gaps. TIGs 12, 478-483.

Rosenberg, U., Schroder, C., Preiss, A., Kienlin, A., Cote, S., I, R., and Jackle, H. (1986). Structural homology of the Drosophila Kruppel gene with Xenopus transcription factor IIIA. Nature 319, 336-339.

Roth, S., Stein, D., and Nusslein-Volhard, C. (1989). A gradient of nuclear localisation of the dorsal protein determined dorsoventral pattern in the Drosophila embryo. Cell 59, 1189-1202.

Ruiz i Altaba, A., and Melton, D. (1989). Involvement of the Xenopus homeobox gene Xhox3 in pattern formation along the anterior-posterior axis. Cell *57*, 317-326.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol *4*, 406-425.

Sambrook, J., Fritsch, E., and Maniatis, T. (1989). Molecular Cloning: A laboratory manual: (Cold Spring Harbor Laboratory Press: New York).

Sampedro, J., and Guerrero, I. (1991). Unrestricted expression of the Drosophila gene patched allows a normal segment polarity. Nature *353*, 187-190.

Sanchez-Herrero, E., Couso, J., Capdevila, J., and Guerrero, I. (1996). The fu gene discriminates between pathways to control dpp expression in Drosophila imaginal discs. Mech Dev *55*, 159-170.

Sander, K. (1985). Fertilization and egg cell activation in insects. In Biology of Fertilization, pp. 409-430.

Sander, K. (1976). Specification of the Basic Body Pattern in Insect Embryogenesis. In Advances in Insect Physiology, pp. 125-238.

Sanger, F., Nicklen, S., and Coulson, A. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA *74*, 5463-5467.

Sanicola, M., Sekelsky, J., Elson, S., and Gelbart, W. (1995). Drawing a stripe in Drosophila imaginal disks: negative regulation of decapentaplegic and patched expression by engrailed. Genetics *1*, 745-756.

Sauer, F., Hansen, S., and Tjian, R. (1995b). DNA template and activator-coactivator requirements for transcriptional synergism by Drosophila bicoid. Science *270*, 1825-1828.

Sauer, F., Hansen, S., and Tjian, R. (1995a). Multiple TAFIIs directing synergistic activation of transcription. Science *270*, 1783-1788.

Sauer, G. (1962). Beziehungen zwischen der bewegungsmoglichkeit und dem regulationsvermogen von blastodermbereichen in Grillenei. Verh. dt. zool. Ges *55*, 323-329.

Sauer, H. (1966). Zeitraffer - mikro - film - analyse embryonaler differenzierungsphasen von Gryllus domesticus. Z. Morph. Okol. Tierre *56*, 143-251.

Scott, M., and Weiner, A. (1984). Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarzu loci of Drosophila. Proc Nat Acad Sci USA *81*, 4115-4119.

Seeger, M., and Kaufman, T. (1990). Molecular analysis of the bicoid gene from Drosophila pseudoobscura: identification of conserved domains within coding and noncoding regions of the bicoid mRNA. EMBO J *9*, 2977-2987.

Seidel, F. (1964). Analyse des differenzierungsverlaufs im insektenei (Gryllus) mittels UV- und rontgenbestrahlungen. Verh. dt. zool. Ges. *57*, 121-143.

Seigfreid, E., Perkins, L., Capaci, T., and Perrimon, N. (1990). Putative protein kinase product of the Drosophila segment polarity gene zeste-white 3. Nature, 825-829.

Siegfreid, E., Wilder, E., and Perrimon, N. (1994). Components of wingless signalling in Drosophila. Nature *367*, 76-80.

Simpson-Brose, M., Treisman, J., and Desplan, C. (1994). Synergy between the hunchback and bicoid morphogens is required for anterior patterning in Drosophila. Cell *78*, 855-865.

Sisson, J., Ho, K., Suyama, K., and Scott, M. (1997). Costal2, a novel kinesin-related protein in the Hedgehog signaling pathway. Cell *90*, 235-245.

Slack, J. (1992). From egg to embryo: regional specification in early development. (Cambridge University Press: Cambridge).

Small, S., Amosti, D., and Levine, M. (1993). Spacing ensures autonomous expression of different stripe enhancers in the even-skipped promoter. Development *119*, 767-772.

Small, S., Kraut, R., Hoey, T., Warrior, R., and Levine, M. (1991). Transcriptional regulation of a pair-ruled stripe in Drosophila. Genes Dev *5*, 827-839.

Smith, A., Owen, L., Strobel, L., Chen, H., and Kanost, M. (1994). Exchangeable apolipoproteins of insects share a common structural motif. J Lipid Res *35*, 1976-1984.

Sneath, P., and Sokal, R. (1973). Numerical taxonomy (San Francisco: Freeman).

Sommer, R., Retzlaff, M., Goerlich, K., Sander, K., and Tautz, D. (1989). Evolutionary conservation pattern of zinc-finger domains of Drosophila segmentation genes. Proc Natl Acad Sci USA *89*, 10782-10786.

Sommer, R., and Tautz, D. (1991a). Asynchronous mitotic domains during blastoderm formation in Musca domestica L(Diptera). Wilhelm Roux Arch Dev Biol *199*, 373-376.

Sommer, R., and Tautz, D. (1993). Involvement of an orthologue of the Drosophila pair-rule gene Hairy in segment formation of the short germ band embryo of Tribolium (coleoptera). Nature *361*, 448-450.

Sommer, R., and Tautz, D. (1989). Minimal homology requirements for PCR primers. Nucleic Acids Research *17*, 6749.

Sommer, R., and Tautz, D. (1991b). Segmentation gene expression in the housefly Musca domestica. Development *113*, 419-430.

Sprengler, F., Stevens, L., and Nusslein-Volhard, C. (1989). The Drosophila gene torso encodes a putative receptor tyrosine kinase. Nature *338*, 478-483.

St.Johston, R., and Gelbart, W. (1987). Decapentaplegic transcripts are localised along the dorsalventral axis of the Drosophila embryo. EMBO J *6*, 2785-2791.

Stanojevic, D., Hoey, T., and Levine, M. (1989). Sequence-specific DNA-binding activities of the gap proteins encoded by hunchback and Kruppel in Drosophila. Nature *28*, 331-335.

Stone, D., Hynes, M., Armanini, M., Swanson, T., Gu, Q., Johnson, R., Scott, M., Pennica, D., Goddard, A., Phillips, H., Noll, M., Hooper, J., de Sauvage, F., and Rosenthal, A. (1996). The tumour-suppressor gene patched encodes a candidate receptor for Sonic hedgehog. Nature *384*, 129-134.

Struhl, G., Struhl, K., and Macdonald, P. (1989). The gradient morphogen bicoid is a concentration-dependent transcriptional activator. Cell *57*, 1259-1273.

Stuart, J., Brown, S., Beeman, P., and Dennell, R. (1991). A deficiency of the homeotic complex of the beetle Tribolium. Nature *350*, 72-74.

Studier, J., and Keppler, K. (1988). A note on the neighbor-joining algorithm of Saitou and Nei. Mol Biol Evol *5*, 729-731.

Stumpf, H. (1966). Mechanism by which cells estimate their location within the body. Nature *212*, 430-431.

Sulston, I., and Anderson, K. (1996). Embryonic patterning mutants in Tribolium castaneum. Development *122*, 805-814.

Tabata, T., Eaton, S., and Kornberg, T. (1992). The Drosophila hedgehog gene is expressed specifically in posterior compartment cells and is a target for engrailed regulation. Genes Dev *6*, 2635-2645.

Tabata, T., and Kornberg, T. (1994). Hedgehog is a signaling protein with a key role in patterning Drosophila imaginal discs. Cell *76*, 89-102.

Tautz, D. (1988a). Regulation of the Drosophila segmentation gene hunchback by two maternal morphogenetic centres. Nature *17*, 281-284.

Tautz, D., Lehmann, R., Schnurch, H., Schuh, R., Seifert, E., Kienlin, K., and Jackle, H. (1987a). Finger protein of novel structure encoded by hunchback, a second member of the gap class of Drosophila segmentation genes. Nature *327*, 383-389.

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

Tautz, D., Tautz, C., Webb, D., and Dover, G. (1987b). Evolutionary divergence of promoters and spacers in the rDNA family of four Drosophila species. Implications for molecular coevolution in multigene families. J Mol Biol *195*, 525-542.

Taylor, A., Nakano, Y., Mohler, J., and Ingham, P. (1993). Contrasting distributions of patched and hedgehog proteins in the Drosophila embryo. Mech Dev *42*, 89-96.

Theisen, H., Purcell, J., M, B., Kansagara, D., Syed, A., and Marsh, J. (1994). Dishevelled is required during wingless signaling to establish both cell polarity and cell identity. Development *120*, 347-360.

Therond, P., Knight, J., Kornberg, T., and Bishop, J. (1996). Phosphorylation of the fused protein kinase in response to signaling from hedgehog. Dev Biol *93*, 4224-4228.

Tillib, S., Sedkov, Y., Mizrokhi, L., and Mazo, A. (1995). Conservation of structure and expression of the trithorax gene between Drosophila virilis and Drosophila melanogaster. Mech Dev *53*, 113-122.

Treier, M., Pfeifle, C., and Tautz, D. (1989). Comparison of the gap segmentation gene hunchback between Drosophila melanogaster and Drosophila virilis reveals novel modes of evolutionary change. EMBO J *8*, 1517-1525.

Treisman, J., and Desplan, C. (1989). The products of the Drosophila gap genes hunchback and Kruppel bind to the hunchback promoters. Nature *341*, 335-337.

Turner, F., and Mahowald, A. (1977). Scanning electron microscopy of Drosophila embryogenesis II Gastrulation and segmentation. Dev Biol *57*, 43-416.

van den Heuvel, M., Harryman-Samos, C., Klingensmith, J., Perrimon, N., and Nusse, R. (1993). Mutations in the segment polarity genes wingless and porcupine impair secretion of the wingless protein. EMBO J *12*, 5293-5302.

van den Heuvel, M., and Ingham, P. (1996). smoothened encodes a receptor-like serpentine protein required for hedgehog signalling. Nature *382*, 547-551.

Vollmar, H. (1971). Musterbildungsprozesse und fruhembryonale gestaltungsbewegungen im ei von Acheta domesticus nach fragmentierung, vitalfarbung und zeitraffer-filmaufnahmen: (Freiburg).

Von Ohlen, T., Lessing, D., Nusse, R., and Hooper, J. (1997). Hedgehog signaling regulates transcription through cubitus interruptus, a sequence-specific DNA binding protein. Proc Natl Acad Sci USA *94*, 2404-2409.

Warrior, R., and Levine, M. (1990). Dose-dependent regulation of pair-rule stripes by gap proteins and the initiation of segment polarity. Development *110*, 759-767.

Weir, M., and Kornberg, T. (1985). Patterns of engrailed and fushi tarazu transcript reveal novel intermediate stages in Drosophila segmentation. Nature *318*, 433-439.

Weischaus, E., and Gehring, W. (1976a). Clonal analysis of primordial disc cells in the early embryo of Drosophila melanogaster. Dev Biol *50*, 249-263.

Weischaus, E., and Gehring, W. (1976b). Gynandromorph analysis of the thoracic disc primordia in Drosophila melanogaster. Wilhelm Roux Arch Dev Biol *180*, 31-46.

Whittle, J. (1990). Pattern formation in imaginal discs. In seminars in Cell Biology, pp. 241-252.

Wilson, R., and al, e. (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of C elegans. Nature *368*, 32-38.

Wolff, C., Sommer, R., Schroder, R., Glaser, G., and Tautz, D. (1995). Conserved and divergent expression aspects of the Drosophila segmentation gene hunchback in the short germ band embryo of the flour beetle Tribolium. Development *121*, 4227-4236.

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

Yu, Y., and Pick, L. (1995). Nonperiodic cues generate seven ftz stripes in the Drosophila embryo. Mech Dev *50*, 163-175

Zarkower, D., and Hodgkin, J. (1992). Molecular analysis of the C. elegans sex determining gene tra-1: A gene encoding two zinc finger proteins. Cell *70*, 237-249.

Zarkower, D., and Hodgkin, J. (1993). Zinc fingers in sex determination only one of the two C elegans TRA-I proteins binds DNA in vitro. Nucleic Acid Research *21*, 3691-3698.

Zecca, M., Basler, K., and Struhl, G. (1995). Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the Drosophila wing. Development *121*, 2265-2278.