Characterisation of subdivision 19E3 of chromosome X housing the *shaking-B* gene of *Drosophila melanogaster*

by

Shuqing Ji

This thesis is submitted for the degree of
Master of Science in
the University of Glasgow

November, 1995

ProQuest Number: 11007825

ProQuest 11007825

To:

Xiaoshi and Kelvin

Summary

Subdivision 19E3 is within the proximal region of the X chromosome of *Drosophila melanogaster*. This region has long been known for being transcriptionally active and housing many neurologically active genes, among which the *shaking-B* gene is of special interest.

The *shaking-B* gene encodes two distinct genetic functions: one specifically neural and the other required for viability. Flies carrying neural mutations show a range of defects, of which the best characterised is the disruption of synapse formation in the giant fibre system of *Drosophila melanogaster*. Mutations of another functional domain cause flies to die as first instar larvae.

Extending the chromosomal walking initiated and carried out earlier, 27 kb more genomic DNA in subdivision 19E3 has been cloned. This will benefit the detailed study of this genomic region and the characterisation of the *shaking-B* locus.

In order to isolate transcripts from the *shaking-B* locus, large scale cDNA library screening (approximately one million clones) has been carried out. Meanwhile, a PCR based method has been used to isolated different splicing variants at 5' end of a cDNA known to represent a transcript from the *shaking-B* locus. After characterisation, it appears that three of the clones isolated represent transcripts from the 19E3 genomic region. Among the three, S3A belongs to a cDNA family whose function is unknown. N52 and W2 represent transcripts from the shaking-B locus. W2 is truncated and contains only two 5' exons. It suggests a novel splicing variant from the *shaking-B* locus. N52 encodes a putitive protein which has an identical C-terminus to the two proteins which have been proposed to account for the two distinct genetic functions of the *shaking-B* locus. N52 may serve a specific developmental function of the gene. Otherwise, it may produce a non-functional protein and is simply a reflection of alternative splicing as a mechanism of gene regulation.

Acknowledgments

Acknowledgments are due to the University of Glasgow, especially, to Department of Genetics in which I used its resources during my study from 1991 to 1993.

It is my pleasure to express my gratitude to all the members of staff and postgraduate students in Department of Genetics for their support and help. To mention a few, thanks to Mary on the sixth floor and the ladies in the preparation room for always being there when I needed help. Thanks also to Colin Milligan, Audery Duncanson, Stephen Goodwin and Simon Tomlinson in Dr. Kim Kaiser's group for allowing me to share things from expensive reagents to personal secret methods.

The friendly, helpful, co-operative, enthusiastic, academically active atmosphere in Dr. Jane A. Davies' group was very important for me to finish my study. Special thanks to Marian Wilkin, Douglas Crompton, Alan Griffin, Martin Todman, Anthony Dornan, Margaret Adams, Sergei Korneev and Mary Gardiner for precious discussion and genuine friendship. I will always treasure the experience of working with these wonderful people.

I feel especially indebted to Dr. Jane A. Davies for all her patience, encouragment, deliberate guidance and thorough reviews of this thesis, without whom the completion of this study would be hard to imagine.

Finally, thanks to my husband, Xiaoshi, for his love, support and understanding. Thanks also to my son, Kelvin, for making sure mom would never overwork during the preparation of this thesis.

CONTENTS

Chapter I. Introduction

1.1. GENERAL INTRODUCTION

1.1.1. Molecular Mechanisms of Nervous System Development

How cells construct the specific synaptic connections that form the basis of neuronal circuits has been one of the major problems of developmental neurobiology. Because of the improvement in techniques over the last a few decades, scientists have taken closer observations on the structure of neuronal circuits and made a more detailed study on how neuronal network are set up in a variety of organisms. On the base of those observations and studies, the construction of specific synaptic connections has been described in three major steps.

The first step of the construction of the specific connection is the growing tips of neurons, the growth cones, travelling long distances to find their correct target region. En route, they are confronted by a series of choice points and yet correctly navigate these pathways in a remarkably unerring way till they reach the correct neighborhood (Thomas et al., 1984; Jacobs & Goodman, 1989a; 1989b; Sink & Whitington, 1991; Goodman *et al.*, 1984; Bentley & O'connor, 1992; Broadie *et al.*, 1993). The next step is contacting and recognizing their correct target, a set of muscles, for example. In this way, the overall scaffold of projections and synapses is initially established (Johansen *et al.*, 1989a, 1989b; Halpern *et al.*, 1991; Cash *et al.*, 1992; Keshishian *et al.*, 1994). The last step is remodelling. These initial patterns of connections are refined, as axonal terminals retract and expand to select a specific subset of cells from within the overall target. This process relies on the context of and competition with surrounding inputs and is capable of transforming a coarse-gained and overlapping projection into a stereotyped pattern of connections (Purves & Lichtman, 1980; Thompson, 1986; Lo & Poo, 1991; Dan & Poo, 1992; Keshishian *et al.*, 1994; Lo *et al.*, 1994). Two broad mechanisms have been suggested to account for the formation of precise patterns of neural connections during development: those that require neural activity (activity dependent), and those do not (activity independent). The initial steps of growth cone guidance typically occur before neurons

become functionally active and rely on molecular mechanisms of pathway and target recognition that are largely activity independent. These mechanisms bring together multiple inputs with appropriate targets to form initial patterns of connections. From this point on, the patterns of neuronal activity within these emerging patterns of connections take over as the predominant mechanism that drives the refinement and remodeling of these initial projections into highly tuned and functioning circuits. Different parts of the nervous system use the same mechanisms but in different proportions to achieve the final specificity. In some of the cases, precise patterns of neuronal connectivity can form without neural activity (for review see Goodman and Shatz, 1993). However, we have very little understanding about the molecular details on how precise patterns of neural connections are formed during development.

Over the last decade, an impressive amount has been learned about the cell surface and extracellular matrix molecules and the possible roles they play during neurite outgrowth, axon guidance and target recognition. Results from both *in vitro* and *in vivo* experiments have shown that general cell surface adhesion molecules, such as NCAM and N-cadherin, and components of the extra-cellular matrix have been shown to function in axon outgrowth as well as cellular adhesion. The *in vivo* spatio-temporal distributions of most of them undergo remarkable developmental changes. Such results reveal the roles played by these molecules in controlling neural development. Furthermore, the cloning and sequencing of cDNAs for many of these molecules has been accomplished, making detailed genetic study possible. Thus it was found that the genes encoding some molecules showed differential expression patterns during development (for reviews see Jessell, 1988; Lander,1989; Hynes & Lander, 1992).

Although vertebrates provide very good systems for cell culture through which one can receive much information about cell behaviour *in vitro* , invertebrates are better for investigating the whole organism *in vivo*. As one of the best understood genetic models with a well characterised nervous system, *Drosophila* is an ideal organism for the investigation of molecular mechanisms involved in the development of the nerves system. In the past, some extracellular proteins have been isolated in *Drosophila*, and they are

very similar to their corresponding vertebrate ones (for review see Semeriva *et al.* 1989).

Goodman and colleagues have been using Grasshoppers and *Drosophila* in a complementary way to approach the underlying genetic and molecular mechanisms of cell recognition and neuronal specificity (Thomas *et at.*, 1984). A variety of molecules, Fascilin I, Fascilin II, Fascilin III, neuroglian and semaphorin I (originally named Fascilin IV), have been isolated from *Drosophila* and grasshopper. Genes encoding these proteins were cloned and the gene expression patterns were examined (Bastiani *et al.*, 1987; Elkins et al., 1990a & b; Grenningloh 1991; Snow *et at.*, 1991; Kolodkin *et al.*, 1992 and 1994).

Fascilin I, Fascilin II, Fascilin III, neuroglian and semaphorin I are expressed on a subset of axon pathway. Among these molecules, Fascilin II, Fascilin III and neuroglian are members of the immunoglobulin superfamily and are highly related to a series of vertebrate cell adhesion molecules (Grenningloh *et al.*, 1990). Semaphorin I is a member of the Semaphorins, a family of highly related transmembrane and secreted growth cone guidance molecules (Kolodkin *et al.*, 1994). The *Drosophila ableson* (*abl*) gene encodes a cytoplasmic tyrosine kinase that is expressed almost exclusively in developing CNS axons during embryogenesis. As with *fasI* mutations, mutitions in *abl* typically lead to no gross defects in CNS morphogenesis. However, in homozygous double mutants for *fasI* and *abl*, motor neurons fail to cross over the midline whereas homozygous mutants for either of the two genes are normal. This suggests that in addition to its function acting as a homophilic adhesion molecule, fascilin I may also interact with signal transduction pathways during axon guidance (Elkins *et al.*, 1990). It is generally accepted that differential adhesion guides growth cones along specific pathways. Meanwhile, diffusible gradients can either attract growth cones toward their targets or act as repulsive forces driving growth cones toward their targets. In some cases, guidance appears to involve both attraction and repulsion which may be controlled by different genes (Goodman and Shatz, 1993).

Compared to what is known about growth cone guidance and pathway recognition, little is known about the molecular mechanisms involved in target recognition. The ability of motorneurons to find and recognize their correct muscles has long been a model system for studies on the mechanisms of target recognition in both vertebrates and invertebrates. By screening for defects in neuromuscular connectivity, a few genes affecting neuromuscular connectivity were found in *Drosophila*. The different classes of mutant phenotypes suggest that neural specificity including a motorneuron recognizing its special target muscle is controlled by molecular mechanisms (Vactor *et al.*, 1993; Keshishian *et al.*, 1994; Chiba *et al*, 1995). A homophilic cell adhesion molecule, connectin, expressed on a subset of muscles and the motorneurons that innervate them has been isolated in *Drosophila*. The temporal and spatial expression pattern of *connectin* suggests that it has a role in target recognition (Nose *et al.*, 1992). In addition, fascilin III, a homophilic adhesion molecule, previously known to be expressed by several of the efferent growth cones is also expressed in a site-specific fashion by the target muscle fibre's membrane (Snow *et al*, 1989; Halpern *et al.*, 1991). The fascilin III expression is transient, corresponding to the period in embryogenesis when the first neuromuscular contacts are made. It is a strongly held belief that there must be some sort of molecules acting as signal for neuron-target recognition and synaptic connection formation.

The work presented here is part of the characterization of the *shaking-B* locus. The gene product of this locus is required for the correct formation of certain synapses in a well characterised neuronal pathway, the giant fibre system, in *Drosophila melanogaster*. This neuronal pathway governs the fly's visually induced escape response. *shaking-B* may also be involved in the formation of other synapses in the nervous system of *Drosophila*. Although much has been learnt in the last decade about the molecular cues that guide axons towards their targets, no genes have yet been characterised which are required for selective synaptogenesis. This makes *shaking-B* potentially very interesting.

1.1.2. The Proximal Region of the X Chromosome of *Drosophila melanogaster*

The *shaking-B* locus maps to the euchromatic 19E region of the polytene X chromosome (Miklos *et al.*, 1987; Baird *et al.*, 1990). In order to explain how the characterisation of this gene has been carried out, it seems necessary to give a brief introduction about the polytene chromosomes of *Drosophila melanogaster* and the euchromatic 19E region.

In most adult tissues of *Drosophila,* mitotic chromosomes undergo several rounds of replication without cell division, becoming polytene chromosomes. This observation has been a subject of considerable interest for many years. The euchromatic portion of a mitotic chromosome undergoes about ten cycles of replication to form the familiar banded polytene structure seen, for example, in the *Drosophila* salivary gland (Lefevre, 1976). On the other hand, the heterochromatic region of a mitotic chromosome replicates little, or not at all. The heterochromatic regions of different chromosomes tend to stick together, forming the chromocentre of the polytene chromosomes in each cell. Except for the right arm of chromosome 3, each arm of the polytene chromosome has a poorly banded region which is not visible in the metaphase plate. This poorly banded area is the so-called β-heterochromatin, whereas the heterochromatic portion of a metaphase chromosome is called α-heterochromatin in a polytenized chromosome.

The β-heterochromatin has remained controversial in the past. It has hardly any similarity with α-heterochromatin, whereas it has a few aspects in common with euchromatin. The β-heterochromatin is transcriptionally active, it has almost the same density of complementation group as euchromatin does, and it undergoes replication during polytenisation [Yamamoto *et al.* 1990]. Apart from the cytological appearance, differences between euchromatin and β-heterochromatin have also been found at the molecular level. There are certain repetitive sequences which are confined to either the euchromatic or β–heterochromatic regions. There are certain arrays of short tandem repeats and mobile element families which are highly enriched in β-heterochromatin and not in euchromatin. Likewise there are certain repetitive sequences which are present in euchromatin and not β-heterochromatin. For example, the

repetitive sequence, $(dC-dA)_n \cdot (dG-dT)_n$ is enriched in most of the euchromatin, but not in β-heterochromatin (Pardue *et al.*, 1987; Miklos *et al.*, 1988). It is worth mentioning that the repetitive sequences in the β-heterochromatic region spread into the neighbouring banded euchromatic portion.

Intriguingly, a monoclonal antibody has been generated which binds only to β-heterochromatin and not euchromatin. One explanation for the different cytological appearances of β-heterochromatin and euchromatin is that the differences are caused by the binding of a specific protein to β-heterochromatin. Perhaps the binding of this protein may cause abnormal replication of β-heterochromatin This difference in replicative properties of the β-heterochromatin compared to euchromatin may be the underlying cause of their cytological differences (Miklos and Costell, 1990).

The proximal region of the X chromosome of *Drosophila melanogaster* from conventional euchromatin in polytene division 19 to β-heterochromatin in division 20 has been extensively studied genetically. There are many deficiencies and duplications in this region and their breakpoints have been mapped, together with some complementation groups in this region, which makes genetically analyzing any locus in this region much easier (Schalet & Levfevre. 1976; Miklos *et al.*, 1986; Miklos *et al.*, 1987; Yamamoto & Miklos, 1987).

There are some genes (or mutations) associated with putative neurological phenotypes such as *uncoordinated, uncoordinated-like, shaking-B, flightless, small optic lodes, sluggish, stoned, stress-sensitive* mapped to the β-heterochromatin-euchromatin transition region from polytene division 19 to division 20 (Miklos *et al* 1987; Perrimon *et al.*, 1989). Through the course of trying to characterise these genes, some of the molecular nature of the transition region, the frequent appearances of some repetitive sequences, for example, has been revealed (Miklos *et al.*, 1984; 1988).

1.2. BACKGROUND TO THIS STUDY

1.2.1. The Giant Fibre System in *Drosophila*

In order to understand how the nervous system develops, it is best to choose a simple neural circuit containing few neurons which are relatively easy to study. One such neural circuit in *Drosophila melanogaster* is the giant fibre system (GFS) which is responsible for the visually induced escape response. Most of the neurons in the fly are too small to access, however all the neurons in GFS are large enough to be identified individually and to be studied electrophysiologically.

The GFS is a neuronal circuit consisting of a group of neurons that relay multimodal sensory information from the brain to some of the muscles of the thorax. It is bilaterally symmetrical, and each side consists of 8 neurons: the giant fibre (GF), the peripherally synapsing interneurone (PSI), the 5 dorsal longitudinal motor neurons (DLMn), and the tergotrochanteral motor neuron (TTMn). The TTMn innervates the tergotrochanteral muscle (TTM), the "jump muscle", so it is also called "jump muscle motor neuron". The DLMn innervates the dorsal longitudinal muscle, the "flight muscle", so it is the "flight muscle motor neuron". Normal driving of the flight muscle and the jump muscle is essential for the fly to jump into the air and fly away. Similar systems to the GFS have been found in some other insects (King and Valentino 1983).

The GF cell body is located in the lower protocerebrum of the brain (Koto *et al.*, 1981). It has been found to respond to various stimuli including small field visual motion, flickers, displacement of the antennae, puffs of air onto the head and body , and wide field visual movement of gratings (Strausfeld *et al.*, 1984). These findings correlate with the areas of brain where the dendrites of the GF have been visualized to synapse (Strausfeld and Bassemir, 1983).

The axon of GF descends through the cervical connective to the thoracic ganglion. In the thorax, the GF is unbranched except for a small tuft of processes at the midline of the mesothoracic region of the ganglion. Below

the tuft, the GF bends laterally, forming an electrical synapse directly onto the TTMn. Just anterior to the region where the GF synapses onto the TTMn, it also forms an electrical synapse with the PSI (King & Wyman, 1980; Tanouye & Wyman, 1980).
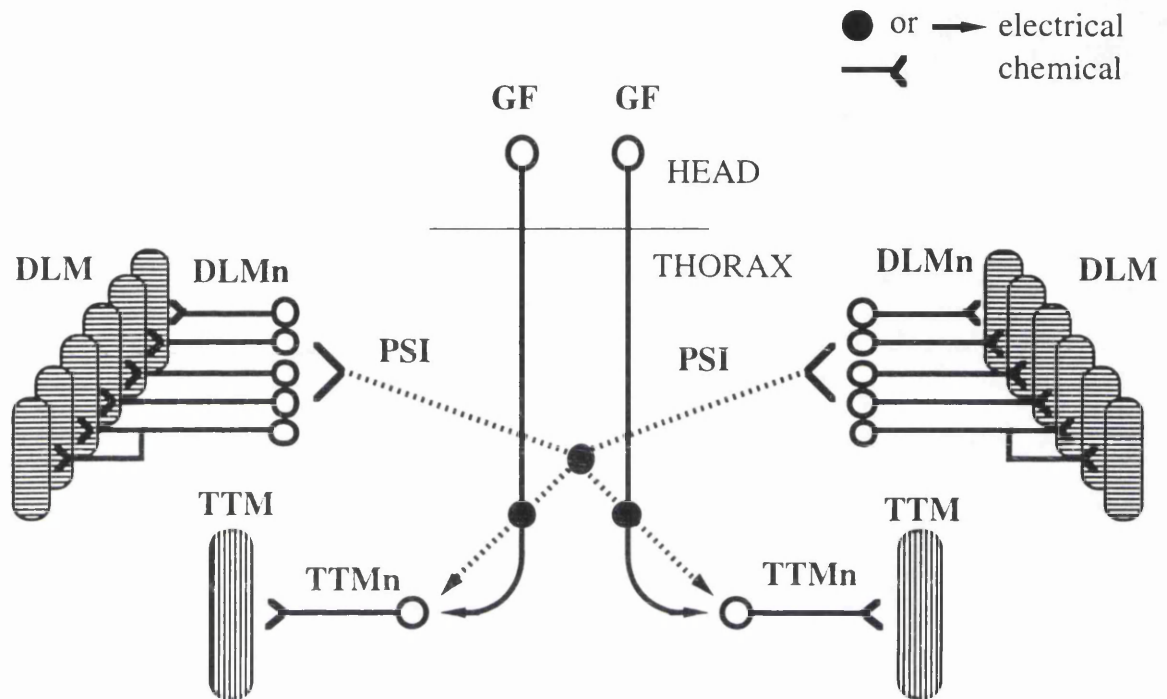
Figure 1.1. Schematic Representation of Giant Fibre System (Taken from Krishnan *et al.*, 1993). The eight neurons on each side (GF, TTMn, and five DLMns) are depicted with their synaptic interconnections. The GF cell bodies are in the brain; all other neurons are thoracic. GF = giant fibre; DLM = dorsal longitudinal muscle; DLMn = dorsal longitudinal neuron; TTM = tergotrochanteral muscle; TTMn = tergotrochanteral motorneuron; PSI = peripherally synapsing interneuron.

The PSIs have cell bodies at the lateral edge of the ganglion. After synapsing with the GF, each PSI then crosses the ganglion and leaves the ganglion in the contralateral posterior dorsal mesothoracic nerve. Along the way, it forms chemical synapses with the DLMns (Ikeda *et al.*, 1980). The PSI also contacts the TTMn (King & Wyman, 1980). Fig. 1.1. which is taken from Krishnan *et al.* (1993), summarizes the proposed connectivity among the interneurons and the motor neurons comprising the GFS in *Drosophila*.

Electrical stimulation of the GF in the brain evokes stereotyped responses in the TTM and DLMs. A single shock activation of the GF elicits a single spike in the TTM with a latency of ~1.0 msec. and spikes of DLMs with a latency of ~1.4 msec. (Baird *et al.*, 1990).

1.2.2. The *shaking-B* (*shak-B*)Locus

In order to understand how the connectivity of the GFS is established, it was necessary to isolate mutants that disrupt the pathway. Mutants were screened by examining the behaviours of the animals. The rational behind this is that small defects in the nervous system can cause large abnormalities in the behaviour. If the GFS is affected, the mutant flies will not jump into the air and fly away when startled by a light-off stimulus.

The most common methods for creating mutations *in vivo* in *Drosophila* include feeding them mutagens such as EMS (ethyl methanesulfonate), or bombarding them with neutrons or X-rays. Generally, EMS causes single base pair substitution or small deletions whereas Neutron and X-ray bombardment tend to cause large DNA deletions.

Passover (*Pas*), one of the mutant alleles of the *shak-B* locus was found originally by feeding male flies EMS and screening their progeny for the inability to jump to a light off stimulus. The purpose was to obtain mutants which would have defects in the GFS. Wild type flies will jump and fly away in response to a rapid light off stimulus. Flies with defects in the giant fibre system fail to do so.

Initially, two mutations were found to disrupt the pathway of the GFS. One is called *bendless* (*ben*), isolated in the screen for jumpless flies. This mutant causes a defect in the giant fibre axon. The other one was *Pas* which had an abnormal connection of the GF with the TTM and the GF-PSI-DLMN circuit was abnormal. In one occasion, it was found that the medial branch of the TTM grew across the midline. Normally it stops and makes the synapse with the GF (the reason why this mutant was originally named "*Passover*"). *Pas* was initially mapped to region between band 19D1 and 20A2 on the X chromosome (Wyman & Thomas, 1983).

shak-B² was isolated in a different behavioural screen. It has been found that *shak-B²* flies have leg tremors under ether anaesthesia (Homyk *et al.* 1980). Subsequently, this mutant was placed in a lethal complementation group, *R-9-29*, which was previously mapped to subdivision 19E3 on the X chromosome (Miklos *et al.*, 1987). Because *shak-B²* is the first allele which was found at this locus, this locus is defined as *shaking-B* (Lindsley & Zimm, 1992).

Further study revealed that *Pas* was a semidominant allele. *Pas/+* flies respond to a light-off stimulus by jumping only 1% of the time among the flies tested, but *+/+* flies have a 89% jump response. *shak-B²* is a recessive allele, and *shak-B² / +* has a normal escape response. Electrophysiological tests have shown that in *Pas* homozygotes, the DLMns do not respond to the stimuli delivered from the GF, and the TTMn responds about 50% of the time or had a delayed response (variable between individuals, within the range of 0.8-2.4 msec). *shak-B²* causes similar defects in the GFS according to electrophysiological studies (Baird *et al.*, 1990). Interestingly, the defects in the GFS have neuroanatomical correlates. The medial branch of the TTMn is often reduced in diameter and in its anterior-posterior extent and the PSI is smaller than in wild-type flies (Baird *et al.*, 1993).

The mutations of *shak-B* not only cause abnormal escape responses, but also confer some other behavioural abnormalities. This suggests that *shaking-B* is required to form neural synapses in other neural circuits as well. It is expected that other defects would be found, as there are not enough genes in the fly to code individually for interconnections between so many neurons. The pleiotropic effects of *shak-B* mutants include:

a) Abnormal gustatory response. *shak-B* flies exhibit increased thresholds of detection in response to sucrose and fructose. Their attractive response to 0.1M NaCl is completely abolished (Balakrishnan and Rodrigues,1991).
b) Abnormal visual transduction (Homyk *et al.*, 1980)
c) Abnormal grooming behaviour (Phillis *et al.*, 1993)
d) Abnormal courtship behaviour (O'Dell, personal communication)

Genetics of the *ShakB* locus

Cytogenetic analysis has mapped the *shak-B* locus to subdivision 19E3 on the X chromosome. Figure 1.2. shows the proximal region of the X chromosome (from subdivision 19D3 to 19E7), the complementation groups in this region and the relevant deficiencies and duplications which were the most useful tools for localising *shak-B*. *Pas* and *shak-B²* have complex complementation relations with the *R-9-29* complementation group which was previously mapped to this region. The *R-9-29* complementation group contains 7 lethal alleles: *17-360*, *E81*, *EC201*, *HM437*, *R-9-29*, *EF535* and *L41*. They cause inviability of the organism when tested *inter se*.

When *shak-B^Pas* and *shaking-B²* are combined with most lethal alleles of the *R-9-29* complementation group, flies are viable but show the neural phenotype. However, two of the lethal alleles, *shak-B^L41* and *shak-B^EF535*, fully complement the neural phenotype of *Pas* and *shak-B²*. This implies that the *shak-B* locus has two independent functional domains: one is essential for the viability of the fly, another one is necessary for the formation of normal synapses. For clarity, the terms *shak-B(lethal)* and *shak-B(neural)* are used to represent the two distinct functional domains, *shak-B(lethal)* being equivalent to the complementation group previously known as *R-9-29*, *shak-B (neural)* being the genetic function formerly known as *Passover* (Crompton *et al.*, 1995)

Trans Effects of 19E5-6 Deficiencies

shak-B^Pas and *shak-B²* cause abnormal driving of the jump muscle when they are combined with *Df(1)17-489*, *Df(1)A53* and *Df(1)T2-14A* in *trans*. All the three deletions uncover 19E5 and 19E6, a more proximal region of the X-chromosome. This region is split from the 19E3 region by at least one complementation group. This naturally raises the question: is there another gene in 19E5-6, which interacts with *shak-B* ? However this was shown not to be the case. When *shak-B^Pas* is combined with these three deficiencies in *cis*, the flies show no neural phenotype. If there were another gene in 19E5-6, interacting with *shak-B^Pas*, it would not matter that *shak-B^Pas* is present with

them in *trans* or in *cis*. It is not known what lies in the 19E5-6 region. It is possible that it houses an exon needed for the neural function or else a *cis*-activator. In addition, one of the alleles of lethal complementation group *R-9-28, R-9-28,* in subdivision 19E4 interacts with *shak-B^{Pas} and shak-B^2* in the gustatory response (Balakrishnan and Rodrigues, 1991). These facts lead to the possibility that regions required for *shak-B* expression span a huge region of DNA, from subdivision 19E3 to 19E6. *shak-B* would appear to be a rather complicated gene.

1.2.3. Molecular Characterisation of the *shaking-B* Locus

The entry point to subdivision 19E3 of chromosome X, where *shak-B* has been localised, was found by combining microdissection technique and the methods of deficiency and duplication mapping. Over 200 kilobases (kb) of genomic DNA has been subsequently cloned by the usage of chromosome walking in and around subdivision 19E3 (Miklos *et al.*, 1988; Crompton *et al.*, 1992). Meanwhile, characterisation of this cloned genomic DNA has generated useful information about the molecular features of this region. The breakpoints of most of the deficiencies and duplications that impinge upon the 19E3 region, including the distal breakpoint of *Df(1)LB6* and the proximal breakpoint of *Df(1)16-3-35* , have been crossed (for details, see Section 3.1.). *Df(1)LB6* and *Df(1)16-3-35* are overlapping deficiencies, deleting a 20 kb stretch of DNA (Crompton *et al.*, 1992). The only known gene removed by these two deficiencies is *shak-B* (Baird *et al.*, 1990).

Several cDNAs have been isolated from this genomic region. They represent transcripts which are produced by differential splicing and alternative promoter usage from the *shak-B* locus. The first cDNA reported is KE2 which was isolated by screening cDNA libraries. It hybridises discontinuously across 18 kb of the genomic DNA between the distal breakpoint of *Df(1)LB6* and the proximal breakpoint of *Df(1)16-3-35* . KE2 is 1817 base pairs (bp) in length and encodes a putative protein sequence of 120 amino acids (aa). It contains a non-coding 5' exon, a second exon containing most of the 120 aa open reading frame, followed by two, 3' end non-coding exons (Crompton *et al.*, 1992).

Later on, Krishnan *et al.* (1993) isolated a P element-induced allele of *shak-B, njP181* and subsequently cloned a new cDNA, designated as P2.4, from the *shak-B* locus. P2.4 is 2726 bp in length, coding a protein of 361 aa.

19D3	19E1	19E2	19E3	19E4	19E5	19E6	19E7
mal	*mell*	*runt*	*shakB*	*R-9-28*	*EC235*	*lf*	*vao*

mal [10]

16-3-35

16-3-32

B57

Ymal [171]

LB6

LB7

17-351

A118

HC279

17-489

A53

T2-14A

Q359

LB23

◼ = deficiency   ▨ = duplication

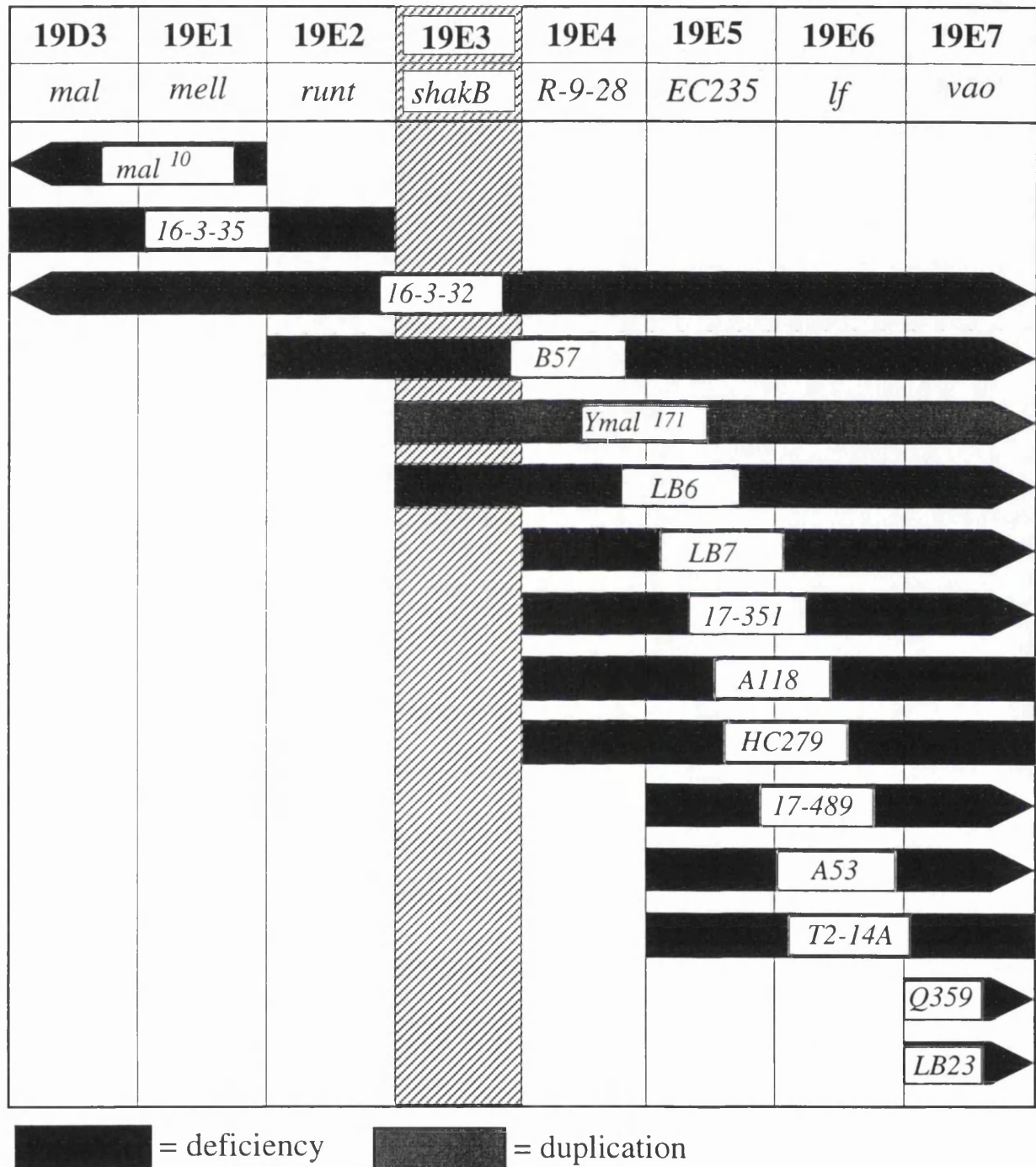Figure 1.2. -- Cytogenetic map of the 19D3 to 19E7 X chromosome interval illustrating the genetic complementation groups and chromosomal aberrations (deficiencies and duplication) used in the cytogenetic and molecular chracterisation of the *shak-B* locus (Baird *et al.*, 1990; Griffin *et al.*, personal communication). Aberrations which extend beyond the boundaries of this map are drawn with an arrowed end.

The cDNA, S2.2, was isolated from a 12-24 hours embryonic cDNA library, by using KE2 as a probe in a PCR-based approach. S2.2 shares common 5' end exons with KE2, whereas, it shares common 3' exons with P2.4. Therefore, the 120 aa at N-terminal region of the S2.2 protein is identical to the KE2 protein. The S2.2 protein and the P2.4 protein share common C-terminal region of 252 aa and have distinct N-termini regions of 120 aa and 109 aa for S2.2 and P2.4, respectively (Crompton *et al.*, 1995).

Molecular Basis of the Complex Genetics of *Shaking-B*

Molecular lesions of some of the *shak-B* mutant alleles have been found in these transcripts. Mutation *shak-B^{L41}* has a 17 bp deletion which removes the common translation start codon of S2.2 and KE2. In *shak-B^{EC201}*, a G to A transition leads to a change of TGG (tryptophan) codon of W$_{273}$ to a TGA (stop) codon, generating truncated S2.2 and P2.4 reading frames (Crompton *et al.*, 1995). The molecular lesions of neural mutations, *shak-B^{Pas}* and *shak-B^2*, have been localised to the sequence of P2.4 on its unique 5' end exon. In *shak-B^2*, a T to A transversion introduce a stop codon, resulting in a premature ending of translation, whereas, in *shak-B^{Pas}*, a C to T transition replaces an arginine with a tryptophan (Krishnan *et al.*, 1993).

The proteins encoded by P2.4 and S2.2 share common C-termini and have distinct N-termini. Based on the positions of different classes of *shak-B* mutant lesions within the proposed coding regions of these transcripts, a simple molecular model has been suggested to account for the complex complementation relationships of *shak-B* alleles (Crompton *et al.*, 1995).

In this model, S2.2 is an essential transcript, thus lesions within its coding sequence can cause lethality, whereas, the P2.4 cDNA represents a *shak-B(neural)* transcript, which has been originally suggested by Krishnan *et al.*(1993). The *shak-B^{L41}* allele, which fully complements the nervous system defects of *shak-B(neural)* alleles, maps to a translation start required for essential but not for neural Shaking-B proteins. Similarly, the neural alleles, *shak-B^2* and *shak-B^{Pas}*, complement the lethal mutations as these mutations map to coding sequence unique to a *shak-B(neural)* protein. The *shak-B^{EC201}*

allele complement neither neural nor lethal *shak-B* alleles, consistent with the fact that it lies within a sequence common to both neural and essential proteins.

Shaking-B(neural), Shaking-B(Lethal), Ogre and Unc-7 Form New Family of Proteins

Shaking-B(neural) (P2.4) and Shaking-B(lethal) (S2.2) share common C-terminus of 252 residues. Their N-termini [109 residues and 120 residues for Shaking-B(neural) and Shaking-B(lethal), respectively] are highly homologous, 75 out of 109 are identical. Thus it has been suggested that the N-terminal coding regions have arisen by partial gene duplication from a common ancestral sequence, and have subsequently been moulded by evolutionary processes to fulfill distinct roles (Crompton *et al.*, 1995).

Moreover, Shaking-B(neural) and Shaking-B(lethal) proteins have been found to be homologous to that encoded by *lethal(1)optic-ganglion-reduced* [*l(1)ogre*]. Shaking-B(neural) and Shaking-B(lethal) are also similar to the product of the C. elegans gene *unc-7*. (Krishnan *et al.*, 1993; Crompton *et al.*, 1995).

The *l(1)ogre* gene plays essential roles for the development and maintenance of the post-embryonic neuroblasts in *Drosophila* including those from which the optic lobe is derived. Mutants of *l(1)ogre* exhibit abnormalities in the CNS (Watanabe & Kankel 1990; 1992). Based on the relative levels of homology among Shaking-B(neural) and Shaking-B(lethal) and Ogre, It has been suggested that *shak-B* and *l(1)ogre* emerged by duplication of a common ancestral gene prior to the partial duplication event that engendered neural and lethal *shak-B* functions. Both the *shak-B* gene and the *l(1)ogre* gene have roles in the development and maintenance of neural pathways in the adult CNS. The expression of both genes were detected in mesodermal derivatives during embryonic development (Watanabe and Kankel, 1992; Crompton *et al.*, 1995). These suggest that the ancestral *shak-B* sequence had a rather universal role in the development of the embryonic mesoderm and adult nervous system, and the partial duplication event enabled

one protein species to a subset of neurons with more specialised properties (Crompton *et al.*, 1995).

The *unc-7* mutation induces certain gap junction formation which does not connect in wild type worms, causes miswiring (Starich *et al.*, 1993). The close similarity among Shaking-B(neural), Shaking-B(lethal), Ogre and Unc-7 suggests that these proteins may form a new family of neural molecules which play roles during nervous system development.

Structure and Function of the Shaking-B Proteins

The currently available evidence suggests a role for *shak-B* in the formation and continued functioning of electrical synapses (Phelan *et al.*, 1995). Crompton *et al.* (1995) have predicted that Shak-B(neural), along with Shak-B(lethal), Ogre and Unc-7, are transmembrane proteins with N- and C-termini cytoplasmic, of which each features 4 transmembrane domains. It has been suggested that the Shaking-B protein are involved in the organisation of cellular membranes. Preliminary antibody staining results has suggested a cell surface location for at least some of the Shaking-B proteins (Crompton *et al.*, 1995; Phelan *et al.*, 1995).

Chapter II. Materials and Methods

2.1. MATERIALS

2.1.1. Chemicals and Biochemicals

All chemicals were of analytical reagent grade and were obtained from one of the following suppliers: Amersham International plc., Amersham, U.K.; BDH Ltd., Poole, U.K.; Boehringer Mannheim, Lewes, U.K.; Formachem Ltd., Strathaven, U.K.; FSA Laboratory Supplies, Loughborough, U.K.; Gibco, BRL Ltd., Paisley, U.K.; Koch-Light Ltd., Haverhill, U.K.; May and Baker, Dagenham, U.K.; Sigma (London) Chemical Co., Poole, Dorset, U.K.; IBI Ltd., Cambridge, U.K.

2.1.2. Enzymes

All restriction enzymes, T4 DNA ligase, Taq DNA Polymerase and T4 Polynucleotide kinase were obtained either from BRL, Gibco Ltd., Paisley, UK., or from Promega Ltd, Southampton, U.K.. Ribonuclease A (RNase A) and lysozyme was obtained from Sigma (London) Chemical Co., Poole, Dorset, U.K..

2.2. GENERAL METHODS

2.2.1. Dialysis Membranes

Dialysis membranes (Scientific Instruments Centre Ltd., London) were cut into appropriate lengths, boiled for 5 minutes in 1% (w/v) EDTA (pH 7.0), stored in 70% ethanol and rinsed in distilled water prior to using.

2.2.2. Spectrophotometric Determination of Nucleic Acid Concentrations

Nucleic acid concentrations were determined spectrophotometrically at 260 nm (Sambrook *et al.*, 1989). In a 1 cm path length quartz cuvette an absorbance of 1.0 corresponds to 50 μg/ml. for double stranded DNA.

2.2.3. Buffer Solutions

10 X TBE Buffer(per litre):	Tris	108 g
	H_3BO_3	55 g
	EDTA (0.5 M, pH 8.0)	40 ml

10 X Loading Buffer:	Bromophenol blue	0.42%
	Xylene cyanol FF	0.42%
	Glycerol	50% in distilled water

TE Buffer: 10 mM Tris, 1mM EDTA, pH to 7.0.

Birnboim Doly Buffer I: 50 mM Glucose, 25 mM Tris, 10 mM EDTA (pH 8.0).

Birnboim Doly Buffer II: 0.2 M NaOH, 1% SDS (made up fresh each time from stoke solution).

Birnboim Doly Buffer III: 5 M KAc, (pH 4.8), made by mixing equal volumes of 3M CH_3COOK and CH_3COOH, pH should be 4.8.

STET: 50 mM Tris-HCl, 50 mM EDTA, 4% Sucrose, 5% Triton-X 100 (pH 8.0).

Oligo-Labelling Buffer (OLB):

Solution A: Tris.HCl (1.25 M pH 8.0),·MgCl$_2$ (0.125 1 ml
M)

Mercaptoethanol	18 µl
dATP (100 mM)	5 µl
dGTP (100 mM)	5 µl
dTTP (100 mM)	5 µl

Solution B: HEPES (2 M pH 6.6)

Solution C: Pharmacia Hexanucleotide primers inTE
(90 OD units/ml)

Prepare OLB by mixing solution A, B and C in the ratio 100:250:150

20 X SSC: 3M NaCl, 0.3 M NaCitrate pH to 7.0.

20 X SSPE: 3.6 M NaCl, 200 mM NaH$_2$PO$_4$, 20 mM EDTA, pH to 7.4.

Denaturing Solution: 1.5 M NaCl, 0.5 M NaOH.

Neutralising Solution: 1.0 M Tris, 1.5 M NaCl, pH to 8.0.

Denhardts Solution: 0.2 mg/ml BSA, 0.2 mg/ml Ficoll-400, 0.2 mg/ml Polyvinylpyrolidne.

Sonicated Salmon Sperm DNA: A 10 mg/ml solution was prepared by dissolving dried salmon sperm DNA (Sigma (London) Chemical Co., Poole, Dorset, U.K.), by vigorous mixing overnight. This was then treated for 7 minutes with a Branson Ultrasonic Processor (Model B-15) in a DAWE Acoustic Booth. Typically the DNA was sheared to lengths between 200 to 1000 bp.

2.3. MICRO BIOLOGICAL TECHNIQUES

2.3.1. Bacterial Strains, Plasmids And Phage

2.3.1a. *E. coli* Strains

Strain	Genotype	Source/Reference
LE392	*supE*44 *supF*58 *hsdR*514 *galT*22 *metB*1 *trR*55 *lacY*1	
XL1-Blue	*supE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*46 *thirelA*1lac- F'[*proAB*+ *lacI*q *lacZ* D M 1 5 Tn*10(tet*r*)*	STRATAGENE

2.3.1b. Plasmids and Bacteriophages

Strain	Genotype	Source/Reference
pBluescript KS(-)		STRATAGENE
R408 (helper phage)		STRATAGENE

2.3.2. *Escherichia coli* Growth Media:

2.3.2a. Liquid Media

L-Broth (per litre):   Bactotryptone   10 g
                           Yeast extract   5 g
                           NaCl   10 g
                           (+5 ml 20% (w/v) glucose)

2 X TY-Medium (per litre): Bacto-tryptone   16 g
                            Yeast extract   10 g
                            NaCl   5 g

SOB Medium (per litre):	bacto-tryptone	20 g
	bacto-yeast extract	5 g
	Nacl	0.5 g
	KCl (250 mM)	10 ml
	Just before use, add 5 ml of 2 M MgCl$_2$	

| SOC Medium (per litre) | SOC medium is identical to SOB medium, except that it contains 20 mM glucose. |

NZCYM Mediam (per litre):	NZ amine	10 g
	NaCl	5 g
	bacto-yeast extract	5 g
	casamino acids	1 g
	MgSO$_4$ · 7H$_2$O	2 g

Phage Dilution Buffer(PDF)	NaCl	5.8 g
(Per litre):	MgSO$_4$	2 g
(Used for the storage and dilution	gelatin,	2%
of phage)	Tris (pH 7.5)	1 mM

All growth media were adjusted to pH 7.0 by adding 5 N NaOH and were sterilised by autoclaving at 120°C for 15 (25) minutes at 15 lb/sq.in on a liquid cycle.

2.3.2b. Media Containing Agar or Agarose

Prepare liquid media according to the recipes given above. Just before autoclaving, add 15 g/litre of bacto-agar or agarose for plates, 7 g/litre of agarose for top agar.

2.3.3. Selection Supplements

The following supplements were added to rich media to select and identify recombinant organisms.

(a) Antibiotics When required either amplicillin, at a final concentration of 100 μg/ml , or tetracycline, at a final concentration of 12.5 μg/ml, were added to broth and agar.

(b) Chromogenic substrates X-gal and IPTG were used at a final concentration of 20 μg/ml. A stock solution of IPTG of 20 mg/ml was filter sterilised and stored at -20°C. A stock solution of X-gal (20 mg/ml) was made up in DMF and stored at -20°C.

2.3.4. Growth Of Bacteria and Bacteriophage

Liquid cultures for transformation, and plasmid DNA preparations were routinely grown in L-broth (NZCYM media for phage DNA preparation) at 37°C with vigorous shaking.

Plate cultures were grown at on L-agar, with antibiotics added as required, and incubated for approximately 16 hours at 37°C.

When plating bacteriophage, phage particles were mixed with plating cells and incubated at 37°C for 20 minutes to allow the phage to absorb to the bacteria. This suspension was then added to 4 (or 8) ml of cooled NZCYM agarose overlay then poured onto hardened NZCYM agar plates (for plate diameter 9 cm, 4 ml top agarose was needed, for plate 10 cm X 10 cm, 8 ml top agarose was needed). After the top agarose had set, the plates were inverted and incubated at 37°C for approximately 16 hours.

2.3.5. Measurement Of Growth

Bacterial cell density was measured as an apparent absorption at 600 nm. The measurements were carried out in 1 cm light path cuvettes using a Beckman DU-50 series spectrophotometer (Beckman Instruments Ltd., U.K.) equipped with a digital read out.

2.3.6. Harvesting of Bacteria

Cells were harvested by centrifugation as soon as the speed reached 8,000 rpm at 4°C in a Beckman JA 14 rotor (holding 250 ml tubes).

2.3.7. Storage of Bacterial Strains

Bacterial strains were stored on L-agar slopes at 4°C, or in 50% LB broth, 40% glycerol at -20°C.

2.4. NUCLEIC ACID ISOLATION AND PURIFICATION

2.4.1. Plasmid DNA

Large scale plasmid DNA isolation was carried out by the alkaline-lysis method of Birnboim and Doly (1979). Further purification by CsCl/EtBr equilibrium centrifugation was described by Sambrook *et al* (1989). Qiagen-tip 100 (QIAGEN Inc.) was also used to isolate up to 100mg DNA according to manufacturer's instruction.

Small scale plasmid DNA preparation was carried out either by the boiling method of Holmes and Quigley (1981), or the alkaline-lysis method of Birnboim and Doly (1979).

2.4.2. Large Scale Isolation of Bacteriophage DNA

0.2 ml of fresh host cells (typically LE392) were mixed with 5 X10^6 to 5 X 10^7 phage particles in phage buffer (plaque was usually picked one day in advance) and incubated for 20 minutes at 37°C. This cell/phage particle mixture was then used to inoculate three conical flasks containing 100 ml of NZCYM broth which were grown at 37°C with vigorous shaking. Good lysis normally appeared after 16-18 hours growth.

The lysates were combined from three flasks and spun at 5000 rpm for 20 minutes in the Beckman JA10 to remove cell debris. The phage particles were

then pelleted by centrifugation at 9000 rpm for 4 hours at 18°C. The phage pellet was air-dried and resuspended in 4 ml of phage buffer.

The suspension was then poured into a Beckman 0.5 X 2" ultra clear tube which contained 3.4 g CsCl. The tube was shaken gently until the CsCl was dissolved. The tube was then centrifuged to equilibrium (16-18 hours at 35,000 rpm in a SW50.1 rotor at 20°C). The phage particle band was formed 1/3 up the tube and was collected by side puncture using a 21G1/2 guage needle inserted just below the band.

CsCl was removed from phage particles by dialysis against a large volume (1-5 litre) of TE for at least 2 hours. After transferring the phage suspension into a centrifuge tube, 0.2 volume of 0.5 M EDTA (pH 8) were added. The phage suspension was incubated at 65°C for 10 minutes. Pronase and SDS were added to a final concentration of 1 mg/ml and 0.5%. The mixture was incubated at 37°C for 1 hour. After phenol-chloroform extraction, the DNA was dialysed against 3 changes of 5 litres of TE for 36 hours at 4°C.

2.4.3. Phenol/Chloroform Extraction of Nucleic Acids

To decontaminate DNA from protein, agarose, etc., a phenol/chloroform extraction was performed. The volume of sample was measured and an equal volume of TE-saturated phenol : chloroform (1:1) was added. The mixture was vortexed for 1 minutes (only gentle shaking was allowed for extraction of lambda phage DNA). The mixture was separated into two layers after centrifugation for 5 minutes. The top aqueous layer was carefully removed to a fresh tube and extracted in an identical manner with an equal volume of chloroform, which remove any residual trace of phenol. The DNA was concentrated (if necessary) by isopropanol or ethanol precipitation.

2.4.4. Concentration of Nucleic Acids by Ethanol or Isopropanol Precipitition

Sodium acetate was added to the DNA solution to a final concentration of 0.3 M by adding a 0.1 volume of a 3.0M solution (pH 5.2), and mixed. Two

volumes of ethanol or one volume isopropanol was added to the mixture, and mixed thoroughly. The mixture was then incubated at -20ºC for a minimum of 2 hours. The DNA was recovered by centrifugation in a microfuge at 4ºC for 15 minutes, or at 10,000 rpm for 30 minutes at 4ºC (Beckman JA20 or JA17 rotor). The DNA pellet was washed in 70% (v/v) ethanol and resuspended in a suitable volume of TE buffer.

2.5. *In Vitro* MANIPULATION OF DNA

2.5.1. Restriction-endonuclease Cleavage of DNA

The methods used were as described by Sambrook *et al.* (1989) in conditions recommended by the manufacturers.

2.5.2. Electrophoresis of DNA

DNA was separated at room temperature on horizontal submerged agarose gels as described (Sambrook *et at., 1989*). A range of agarose concentrations (0.6-1.2%) was used in order to resolve molecules across a range of lengths. For visualization of DNA, EtBr was added to melted agarose at a concentration of 0.1 mg/ml when the agarose was ready to pour into a gel former. Gel buffer is 1 X TBE. Markers were usually 1kb ladder or 123bp ladder (BRL Gibco Ltd.,Paisley, U.K.). Both DNA samples and markers were mixed with 0.1-0.2 volume of loading buffer prior to loading. DNA was visualised by UV fluorescence of the EtBr stained gel.

2.5.3. Photography of Electrophoresis Gels

Gels were photographed with UV translumination (wavelength: 240 nm and 320 nm) using a Polaroid camera loaded with Polaroid 4X5' Land Film (No57) fitted with a Kodak Wratten Filter No23A (red).

2.5.4. Recovery of DNA from Agarose Gels by Electroelution into Dialysis Bags

A gel slice contained the desired DNA band was excised from the EtBr stained gel under a UV lamp (wavelength: 320). The gel slice was then sealed into a piece of dialysis bag along with approximately 500 μl of 1 X TBE to keep the gel slice in constant contact with the buffer. The dialysis bag was immersed in 1 X TBE in a electrophoresis tank. Electric current was passed through the bag (usually 4-5 V/cm for 2-3 hours). After the DNA is electroeluted out of the gel slice, the polarity of the current was reversed for 1 minute to release the DNA from the wall of the bag. The buffer containing the DNA was removed out of the bag. Following isopropanol precipitition, the DNA pellet was dissolved in a suitable volume of TE buffer.

2.5.5. Amplification of DNA by the Polymerase Chain Reaction (PCR)

A 12-24 hours embryonic *Drosophila* cDNA library (kind gift of Nick Brown) was used as the template.

Primer 1 (NB1): 5'-GGAATTCCGG TGACACTATA
GAATACAAGC TTGC-3'
Primer 2 (P11): 5'-CTTCGGGTCC GTTCCTC-3'

The reaction was set up as below:

Template	1 μl
Primer1	15 pmoles (in 2.5 μl of H_2O)
Primer2	15 pmoles (in 2.5 μl of H_2O)
dNTPs (2 mM)	5 μl
10 X amplification buffer	5 μl
Taq DNA polymerase	0.5 μl

H_2O to a total volume of 50 μl

The reaction mixture was in a 0.5 ml tube, overlain with 20 μl of mineral oil. The sample was placed in a THERMAL REACTOR (HYBAID), heated to 94°C for 2 minutes, taken through 30 cycles of 1 minute at 60°C, 2 minutes at 72°C, and 50 seconds at 94°C, and finally incubated for 5 minutes at 60°C and

for 10 minutes at 72°C. After amplification, the PCR mixture was separated by eletrophoresis in 1% agarose gel. The target DNA was recovered from the agarose gel as described in 2.5.4.

The amplified DNA fragments were either directly cloned into T-vectors (Kovalic *et al.*, 1991), or treated with kinase to generate blunt ends and then ligated into EcoRV cleaved pBKS(-).

After ligation, the ligated DNAs were transformed into competent cells made of XL-1-Blue strain for blue-white selection (white colonies were picked up over blue colonies on X-gal and IPTG treated L-agar plates. See Section 2.3.3.)

2.6. RECOMBINANT DNA TECHNIQUES

2.6.1. Ligation

Both insert DNA and vector DNA (typically pBluescript KS-) were excised with designed restriction-endonucleases. DNAs were recovered from agarose gel. A typical ligation reaction was set up as follow:

Foreign DNA:	100 ng
Vector DNA:	100 ng
Bacteriophage T4 DNA ligase:	0.5 unit (weiss unit)
10 X ligation buffer:	1 µl
(supplied by manufacturer)	
dH$_2$O :	to a final volume of 10µl

The ligation mixture was incubated at 16°C for a minimum of 4 hours.

2.6.2. Transformation of *E.coli* with Plasmid DNA

2.6.2a. Preparation of Competent Cells

A single colony of *E coli* XL-1 blue was used to inoculate 10 ml LB which was incubated overnight at 37°C. 4 ml of the culture was used to inoculate

100 ml LB which was grown until an OD_{600} of 0.3-0.4 (this took about 2 hours). The cells were harvested by centrifugation. The cells were then resuspended in 50 ml ice-cold 50 mM $CaCl_2$ and incubated on ice for 30 minutes. Following centrifugation, the cells were resuspended in 10 ml ice-cold 50 mM $CaCl_2$. The cells were ready for transformation after 1-24 hours of incubation on ice.

2.6.2b. Transformation of Competent Cells

An aliquot of ligation mixture (usually 5 µl) or 50 ng plasmid DNA was added to 200 µl of competent cells. The transformation mixture was incubated on ice for 30 minutes. The mixture was then heat-shocked at 42ºC for 2 minutes, then immediately placed on ice. After adding 800 µl of SOC, The transformation mixture was shaken at 37ºC for 30 minutes. The transformed cells were then plated onto antibiotic and chromogenic substrate-containing L-agar plates and incubated overnight at 37ºC.

2.7. HYBRIDISATION OF NUCLEIC ACID

2.7.1. DNA Transfer to Nitrocellulose Membranes

2.7.1a. Southern Blotting

This method is based on that of Southern (1975). After electrophoresis the agarose gel was placed in denaturing solution for 30 minutes and then placed into neutralising solution for 45 to 60 minutes. The DNA was then blotted onto a nitrocellulose (Hybond-C EXTRA, Amersham) membrane by capillary action (20 X SSC as transfer buffer) as described in detail in Sambrook *et al.*,(1989). DNA was bound to the membrane by baking for 2 hours at 80ºC.

2.7.1b. Dot Blotting

This method was modified from that described by David *et al*, (1986). Approximate 10 µg DNA in 5 µl of TE was denatured by heating up at 95ºC for 5 minute, then snap cooled. 10 µl of 20 X SSC was added to each DNA sample. The samples were then immediately spoted on 20 X SSC wetted

nitrocellulose filter in a defined pattern of rows and columns. The filter was then baked at 80°C for 2 hours.

2.7.1c. Lift Filters from Plates with Bacteriophage (Plaque Lifts)

A piece of labeled nitrocellulose filter was laid carefully on the top of a plate with bacteriophage growing on it. The filter and the plate were oriented at three or more asymmetric locations by stubbing through with an 19G2 guage needle. For a single filter lift, the filter was left on the top of the plate for at least 1 minute. For duplicate filter lifts, the first filter was left for approximately 0.5 minute and the second filter should be left for no less than 5 minutes. The second filter was oriented according to the marks through the underlying medium left by the first orientation. The filter was then lifted from the top of the plate. Subsequently, the filter was put in denaturing solution for 30 seconds, in neutralizing solution for 2 minutes and washed in 2 X SSC for 10 minutes. The filter was blotted dry with Whatman 3 MM paper, baked at 80°C for 2 hours.

2.7.1d. Colony Lifts

To obtain duplicated filter lifts from the master-plates for primary screening, labeled dry nitrocellulose filters were placed on the master-plates with bacteria grown for 4-5 hours at 37°C. The filters and the master-plates were oriented also by stabbing through asymmetrically. The filters were then transferred colony side up to a set of fresh plates. Two set of plates were incubated for 16-18 hours at room temperature. The master-plates were put away for picking up positive colonies later on. The filters with colonies were peeled off the plates. Each of the filters was placed colony side down onto a labeled filter supported by a glass plate and a piece of Whatman 3 MM paper. The second filter was oriented by stubbing a needle through the orientation markers on the first filter. After another piece of Whatman paper and a glass plate were placed on the top of the two filters, the colonies were transferred to the second filter by pressing the top glass plate firmly. Lysis of cells was carried out efficiently by autoclaving filters (it also denatured the DNA) with transferred colonies for 5 minutes. The filters should be kept damp and flat during autoclaving. The filters were then immediately baked for 2 hours at

80°C to bind DNA to the filters. Before hybridization, the filters were washed in a solution of 0.1 M NaH_2PO_4, 50 mM $Na_4P_2O_7$, 1% SDS and 1 mM EDTA for 2 hours at 65°C. The two set of filters were separated during the washing

For single set colony lift, the filters were lifted in a similar manner as for single set plaque lift. The filters were then autoclaved, baked and washed as described above.

2.7.2. Radio labelling of DNA

[32]P labelling of DNA was carried out by Random Priming technique of Feinberg and Vogelstein (1983), with modifications as described by Hodgson and Fisk (1987). Random sequence hexanucleotides were used to prime DNA synthesis on a denatured double stranded DNA template by a large fragment of DNA polymerase I (Klenow Fragment). A typical Random Priming reaction was set up as below:

DNA (double stranded, denatured by boiling 5 minutes, then snap-cooled on ice)	10-100 ng in 36μl TE
Oligo-labelling Buffer	10 μl
[32]P-dCTP (800 ci/m mol)	3 μl
Klenow	1 μl

The reaction was left either for 2-4 hours at 37°C or overnight at room temperature.

The unincorporated nucleotides were removed by selective precipitition with Spermine as described by Hoopes and McClure,(1981). To 50 μl labeling mixture, 1.2 μl of 0.1M spermine was added (i.e. the final concentration of spermine was 2.4 mM). After incubation on ice for a minimum of 15 minutes, the labeling mixture was then centrifuged at 14,000 rpm for 15 minutes at 4°C. The supernatant containing unincorporated nucleotides was removed and the pellet was resuspended in 500 μl of a solution of 10 mM EDTA and 0.1% SDS.

The percentage of the incorporation of ^{32}P-dCTP was measured by comparing the cps (counts per second) of the labeled DNA with the cps of the unincorporated nucleotides.

2.7.3. Hybridisation

Nitrocellulose membranes bound with denatured DNA (Southern blots, dot blots and colony/plaque lifts) were pre-hybridised in 5 X SSPE, 5 X Denhardt's solution, 0.5% SDS and 100 μg/ml sonicated and denatured salmon sperm DNA, for at least 1 hour at 65°C. The probe was denatured by heating for 5 minutes at 90 -100°C for 5 minutes and snap-cooled on ice. The hybridisation was carried out at the same condition as that for pre-hybridisation for at least 6 hours.

Blots were pre-washed briefly in 2 X SSC at room temperature and then washed at 65°C for three times :

1°	2 X SSC, 0.1% SDS	20 minutes
2°	2 X SSC, 0.1% SDS	20 minutes
3°	0.5 X SSC, 0.1% SDS	15 minutes

When high stringency was required, an additional wash was carried out in 0.1 X SSC, 0.1% SDS for 15 minutes at 65°C.

The nitro-cellulose membranes were then rinsed in 2 X SSC, blotted dry, and autoradiographed.

2.7.4. Auto-radiography

^{32}P in nitro-celloluse membranes was detected by exposure to either Amersham Hyperfilm or Fuji RX X-ray film using intensifying screens at -70°C. ^{35}S in polyacrylamide gels was detected by exposure to film in the absence of intensifying screens at room temperature. Films were developed by a X-OMAT Automatic Film Processor Model Compact X2, (X-Ograph Ltd.).

2.8. SCREENING cDNA / GENOMIC DNA LIBRARIES

2.8.1. Titering Library

For a cDNA library or a genomic DNA library constructed in bacteriophage vectors, a serial dilutions of the phage particles in PDB were made. After absorbing into the host cells, the phage infected cells were grown on NZCYM plates (see section 2.3.4.). After incubation at 37°C for 16 hours, the number of pfu on each plate was counted. The titre of the DNA library was then calculated. Similar method was used for titering a DNA library constructed into plasmid vectors. Instead of pfu, the number of colonies was counted.

2.8.2. Primary Screening

After titering, the DNA library was plated out in a designed density on 10 X 10 cm² square plates (normally 10,000 pfu/plate, 50,000 colony/plate). Filter lifts (usually duplicated filter lifts for primary screening) were taken according to section 2.7.1. The library was then screened by hybridisation of radio-labeled DNA to duplicated filter lifts simultaneously. After autoradiography, the putative positive plaques or colonies were identified by alignment of plates, filters and autoradiograph of hybridised filters. The putative positive plaques were picked up by an end of a pasteur pipette and transferrd into a tube containing 1 ml PDB and 20 μl chloroform. For the plasmid library, the putative positive colonies were picked and transferred into 1ml L-broth (containing appropriate antibiotics).

2.8.3. Re-screening DNA libraries

In the primary screening, usually more than one plaque/colony was picked. The negative plaques/colonies were eliminated in subsequent rounds of re-screening (i.e. the secondary and the tertiary screenings). The secondary and the tertiary screenings were carried out the same as the primary screening, except that the density of pfu/colonies was lower (usually, the secondary screening: 300-400 pfu/plate, 500-1,000 colonies/plate; the tertiary screening: less than 100 pfu/plate, 100-200 colonies/plate, plate diameter: 9 cm) to ensure that a single positive plaque or colony was picked up finally.

2.8.4. cDNA Library and Genomic Library Screened

cDNA library:

Name	Vector	Source/Reference
pNB40 Drosophila cDNA Library (12-24 hours embryonic)	pNB40	Nicholas Brown
Kauvar cDNA library	λ gt10	
Ron's Drosophila Head cDNA library	Uni-ZAP™ XR	Kim Kaiser's lab
Stratagene Adult Drosophila cDNA library	Uni-ZAP™ XR	Stratagene

Genomic DNA Library:

Name	Vector	Source/Reference
HE	λ EMBL3	
LK-E3	λ EMBL4	

2.9. DNA SEQUENCING

2.9.1. Preparation of DNA for Sequencing

Sequencing of double stranded DNA was carried out by the dideoxy chain termination method first described by Sanger *et al* (1977). Sequenase Version 2.0 (United States Biochemical Corporation) was used. The sequencing reactions were carried out using the conditions recommended in the Sequenase Version 2.0 "step-by step protocols" supplied by the manufacturer.

2.9.2. Polyacrylamide Gel Electrophoresis

Products of DNA sequencing reactions were separated on denaturing polyacrylamide gels. Gels were composed of the following constituents:

40% (w/v) Acrylamide	18 ml
(Acrylamide : Bis-acrylamide, 19:1)	
10 X TBE	12 ml
Urea	60 g
H_2O	36 ml
10% (w/v) ammonium persulphate	900 ml
TEMED	15 ml

After pouring, a gel was left for at least 2 hours to polymerise. Prior to loading sequencing samples, the gel was pre-run at 60 W for 30-40 minutes to heat up to 55°C. Samples were denatured by heating to 75-80°C for 2 minutes, and then loaded onto a gel with the aid of a sharks-tooth comb (IBI). Electrophoresis was carried out at 45-60 W (the gel was kept at 55°C) for 2 to 6 hours. The gel was fixed in a solution of 10% acetic acid and 12% methanol for 20 minutes and then dried onto Whatmann 3 MM paper using an ATTO RapidDry Gel Drier connected to a HOWE Refrigerated Solvent Trap and a Brook Crompton Parkinson Multi-Purpose Vacuum Unit.

2.9.3. Computer Programs Used in the Analysis of DNA and Amino Acid Sequences:

Seqed and **Map** (from the IBI MacVector package)

Fasta (from the Wisconsin GCG package)

Chapter III. Extension of a Chromosomal Walk in Subdivision 19E3

3.1. INTRODUCTION

shaking-B had been mapped cytologically to subdivision 19E3 of the X chromosome (Miklos *et al.*, 1987; Baird *et al.*, 1990). The region from 19D3 to 20F2 has previously been known for its highly repetitive feature (Miklos *et al.*, 1984). In order to clone the DNA in subdivision 19E3 and the region nearby, the microdissection technique developed by Scalenghe *et al.* in 1981 was used. Instead of dealing with total genomic DNA, this method was designed to directly excise the DNA from certain regions of the polytene chromosomes, then to clone the limited amount of genomic DNA into a selected vector, which can save much experimental time. The entry point for the cloning of the *white* locus and 3B1-3C1 region was found by using this method (Pirrota *et al.*, 1983).

The region from 19EF-20A on one polytene X-chromosome was sliced into 4 pieces using a glass needle. The DNA was extracted and cleaved with EcoRI. The DNA fragments were then cloned into λNM1149. Thus 4 mini-libraries: 19E, 19F, 20A and 20B-F were generated (Miklos *et al.*, 1988). The microclones in mini-library 19E were then mapped using deficiency mapping to deficiencies and duplications covering or impinging upon the complementation group *R-9-29*. Chromosomal walking was then initiated with those microclones selected by deficiency and duplication mapping (Miklos & Davies, unpublished).

There were four walks initiated around subdivision 19E3. Three of them were started and named after the number assigned to the microclone i.e. 798, 896 and 952. One walk was started from the *runt* locus, which has been cloned (Gergen & Butler, 1988), so it was named the runt walk. The 952 walk fell into 19E3 and was extended both distally and proximally. The 896 walk was mapped to 19E4 and extended distally. The walks were oriented by localising the end points of various deficiencies (Crompton *et al.*, 1992),So far, the 952 walk has been extended over 170 kb and the 896 walk has been extended some 80 kb. However the two walks have not been joined up yet.

Meanwhile, some characterisation of these cloned genomic regions has been in progress. By using the reverse Northern technique, several transcribed regions have been identified on the walks (Fig. 3.1.1.). The repetitive areas of these walks have also been investigated. Thus we know which part of the genomic DNAs is highly repetitive, which part is moderately repetitive and which part is single copy (Griffin, unpublished).

The breakpoints of most of the deficiencies and duplications that impinge upon the 19E3 region have been mapped. The distal breakpoints of $Dp(1)Ymal^{171}$, $Df(1)LB6$ and $Df(1)17-351$ and proximal breakpoint of $Df(1)16-3-35$ have been localised in the 952 walk, whereas the distal breakpoint of $Df(1)T2-14A$ has been found in the 896 walk. In addition, it is known that the distal breakpoints of $Df(1)A118$ and $Df(1)HC279$ are in the gap between the 952 walk and the 896 walk (see Fig. 3.1.1.) (Crompton et al., 1992). Therefore, the 19E3 subdivision is almost bracketed by the two walks with a gap in between.

One putative cDNA, designated KE2, has been isolated by screening cDNA libraries with probes from the 952 walk. It hybridises discontinuously across 18 kb of the genomic DNA between the distal breakpoint of $Df(1)LB6$ and the proximal breakpoint of $Df(1)16-3-35$ (Fig. 3.1.1.) (Crompton et al., 1992). The only known gene removed by these two deficiencies is shak-B (Baird et al, 1990). It can thus be concluded that the walks definitely contain the genomic region constituting at least part of the shak-B locus.

$Df(1)A118$ and $Df(1)HC279$ which appear to break in the gap between the 952 walk and the 896 walk (Fig. 3.1.1.) cause a severe neuronal phenotype when heterozygous with $shak-B^{Pas}$ (Baird et al., 1990). This fact makes the genetic properties of the(the region between the 2 walks potentially interesting. Therefore, carrying out further chromosomal walking appeared to be necessary in order to characterise the shak-B locus. This is the purpose of the work presented in this chapter.

952 walk

896 walk

Region1

Region2

Region3

λ995

λ683

λ486

?

Dp(1)Ymal¹⁷¹

Df(1)LB6   Df(1)16-3-35

Df(1)7-3-51   Df(1)A118/
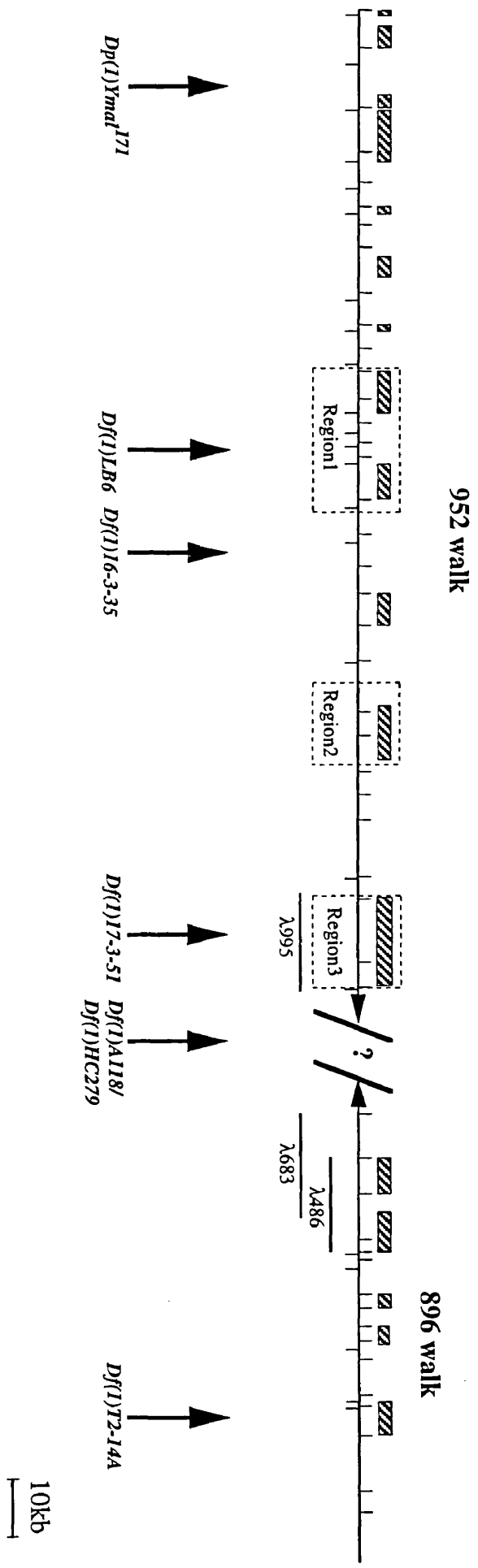Df(1)HC279

Df(1)T2-14A

10kb

Figure 3.1.1. A diagram of the two chromosomal walks (restriction sites shown are only EcoRI). All the breakpoints of duplication and defeciencies mapped are indicated by up-pointed arrows. Hatched bars indicate the transcribed regions at various developmental stages, which were defined by Reverse-Northern technique and *in situ* hybridisation experiemnts (Data from Griffin and Davies).

3.2. RESULTS

3.2.1. First Approach: Bridging the Gap Using Cosmid Clones

Because of the repetitive feature of this genomic region, the 896 walk has terminated in a repetitive area. The 952 walk, probably due to underrepresentation of more proximal sequences in the library, has been difficult to extend proximally.

As it was difficult to extend the walks, we adopted a strategy that involved utilising cosmid clones. This method has worked well in some other cases, for example, the chromosome walk in 3B1-3C1 region (Pirrotta *et al.*, 1983).

We were kindly given cosmid clones which were isolated with probes from this region. These clones were characterised by using the Southern blotting technique (Alan Griffin, personal communication). It was found that two clones, c9A1 and c8C1, extend the 952 walk and the 896 walk in the proximal and distal directions, respectively. Fragments which were shown by Southern blotting not to overlap with the walks were excised and used to screen a genomic phage library, as there were problems with rearrangement of cosmid sequences during growth.

The 952 walk

A 3.8 kb single-copy DNA fragment was excised from c9A1 by a EcoRI-SalI double restriction digest and this DNA fragment was then used to screen the HE genomic library which was cloned in λEMBL3. Three phage clones, λS2, λS4 and λS5, were isolated by using the c9A1-3.8 kb fragment. DNAs were made from these clones for further analysis. When these clones were mapped with restriction enzymes, it was found that λS2, λS4 and λS5 overlapped (Fig. 3.2.1.). However, the restriction maps of the three clones: λS2, λS4 and λS5, had no similarity to that of λ995 which is the most proximal clone of the 952 walk (Fig. 3.1.1. and Fig.3.2.1.). Southern blotting experiments has shown that λS2, λS4 and λS5 do not hybridise to the 952 walk (Griffin, personal communication) Thus, λS2, λS4 and λS5 do not overlap with the 952 walk. Southern blotting has also shown that c9A1 hybridises only to the

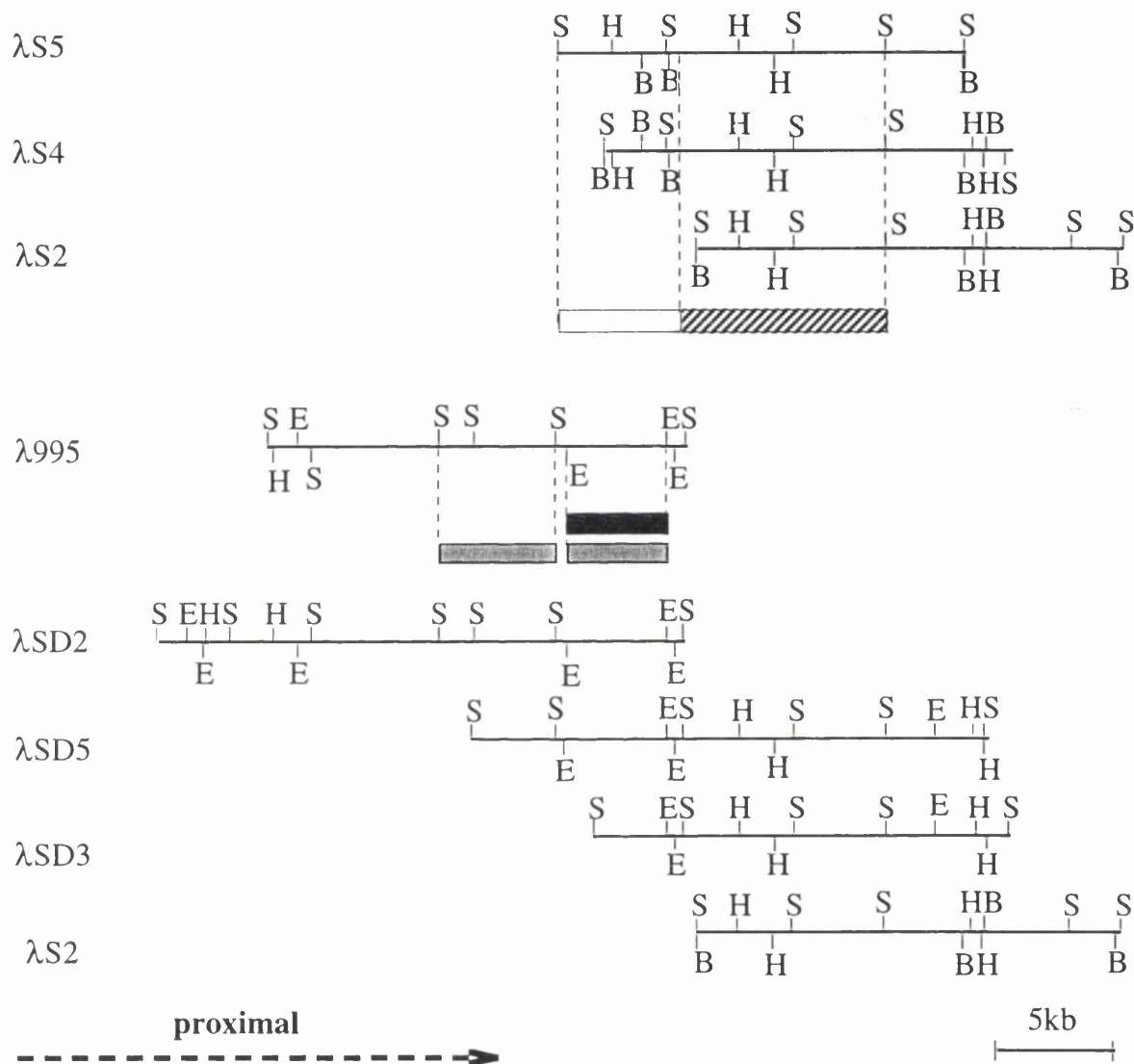middle region of λS2 etc. (Fig. 3.2.1.), which is rather unexpected. Some discrepancies occurred.

Figure 3.2.1. Restriction maps of λS5, λS4, λS2, λ995 (from Davies), λSD2,λSD5 and λSD3.

λS5, λS4 and λS2 overlap with each other . Plain bar indicates the region which has shown repetitive feature and does not hybridise to the rest of the walk (from Davies). Hatched Bar indicates the region which hybridise with c9A1.

λSD2, λSD5 and λSD3 overlap with each other. They overlap with λ995 and with λS5, λS4 and λS2, too.

Black bar under λ995 indicates the DNA fregment used to isolate λSD2, λSD5 and λSD3. Dotted bars indicate the DNA fragments used to isolate 2B and 5A (see Section 4.2.1.).

(E = EcoRL, S = SalI, H = HindIII, B = BamHI.)

N.B. λS5, λS4 and λS2 have no homology to λ995 (see text). All restriction maps were drawn with the help of J. A. Davies.

The 896 walk

To extend the 896 walk, single copy fragments of 2.2 kb and 1.2 kb were excised from c8C1 by a EcoRI-HindIII double restriction digest. Two clones, λSC2 and λSU2, were isolated by using c8C1-1.2 kb and c8C1-2.2 kb EcoRI-HindIII fragments, respectively. λSC2 was used to probe a Southern blot made of all the phage clones composing the 896 walk. The result showed λSC2 did not hybridise to any of these clones. It is likely that there is a small gap between λSC2 and λ683 (the most distal clone of the 896 walk). λSC2 was restriction-mapped (Fig. 3.2.2.). A Southern blot made of λSC2 was probed by c8C1 and the result showed that c8C1 hybridised to the middle region of λSC2 (Fig. 3.2.2.). Discrepancies occurred once again. λSU2 was used to probe the same blot as λSC2, it hybridised to two of the phage clones: to λ683 partially; to λ382 completely. λ683 and λ382 were then used to probe a Southern blot made of λSU2, the results were the same. Since λ683 is the most distal clone of the 896 walk and λ382 is more than 10 kb away from the distal end of the 896 walk, λSU2 appears to overlap completely with the 896 walk. However, the overlap does not extend to the end of the 896 walk. Thus λSU2 can not extend the 896 walk. It was rather unexpected to find out that λSU2 overlapped with the 896 walk without extending it distally as the c8C1-2.2 kb fragment which was used to isolate λSU2 did not hybridise to the 896 walk. One possible explanation is that the 2.2 kb EcoRI-HindIII fregment from c8C1 contains a sequence, for example, an insertion, which is absent from the 896 walk.

3.2.2. Second Approach: Using a Different Genomic DNA Library

Since the first attempt at extending the chromosomal walking in 19E3 subdivision did not succeed, an alternative strategy was adopted, that is to use a genomic DNA library made from a different fly strain.

A 4.3 kb EcoRI-EcoRI fragment was excised from λ995 which was the most proximal fregment of the 952 walk (Figure 3.1.1. and Figure 3.2.1.). A 2.7 kb SalI-HindIII fragment was excised from λ486 which was at the distal end of 896 walk (Figure 3.1.1. and Figure 3.2.2.). These DNA fragments

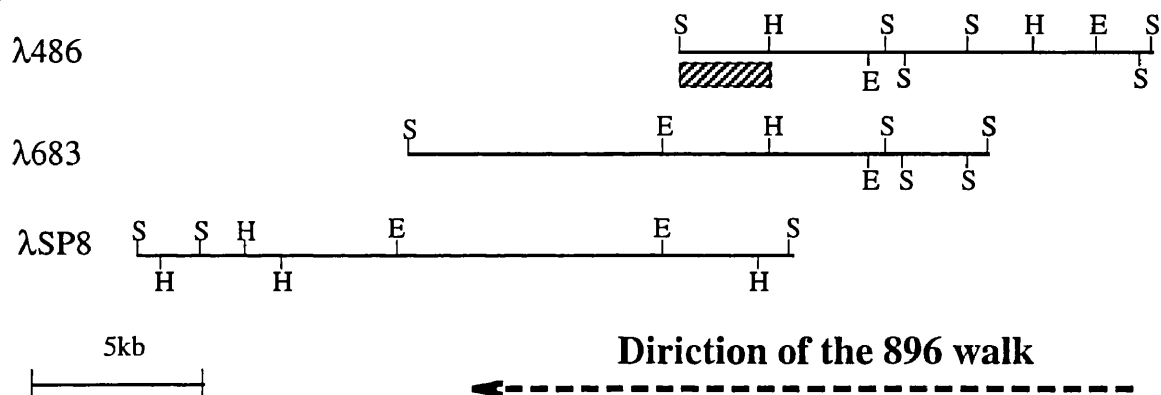were labeled with ^{32}P-dCTP and used to screen genomic DNA library LK-E3.

A.
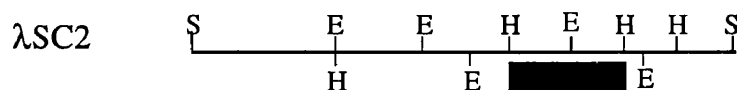
B.

Figure 3.2.2. A. Restriction maps of λSC2. Black bar indicates the region of λSC2 with homology to c8C1.

    B. Restriction maps of λ486, λ683 (from Davies) and λSP8. Hatched bar indicates the DNA fragment used to isolate λSP8.

(E = EcoRI, S = SalI, H = HindIII)

N. B. Restriction maps were drawn with the help of J. A. Davies.

Three clones, λSD2, λSD3 and λSD5, were isolated with the λ995-4.3 kb EcoRI-EcoRI fragment. The results from restriction mapping has shown that λSD2, λSD3 and λSD5 overlap with each other. By comparing the restriction maps, it was found that they overlap with λ995 while λSD3 extends the 952 walk the most. It has also been found that λSD2, λSD3 and λSD5 overlap with λS2, λS4 and λS5 (see section 3.2.1.) although the sizes of some of the fragments varies (Fig. 3.2.1.). It has been known that the distal ends of S2, S4 and S5 are highly repetitive and do not hybridise to the rest of the 952 walk (Fig. 3.2.1.) (Davies, unpublished). It is possible that the fly strain from which the HE genomic DNA library was constructed has a transposable element inserted into this genomic region. This would also explain why c9A1 hybridises only to the middle region of λS2 etc. (see section 3.2.1.) Therefore, the 952 walk has been extended forward 18 kb in total.

For the 896 walk, only one positive clone, λSP8, was isolated when screening LK-E3 with λ486-2.7 kb SalI-HindIII fragment. Results from both restriction mapping and southern blotting have shown that λSP8 overlaps with λ486 and λ683 and extends the 896 walk about 9 kb forward into the gap between the 896 walk and the 952 walk (Fig. 3.2.2.). However, λSP8 does not overlap with previously isolated λSC2. This conclusion has been drawn from the comparison of their restriction maps.

3.3. DISCUSSION

By the end of this study, the chromosomal walk had been extended by 27 kb. In total, about 280 kb DNA had been cloned around subdivision 19E3. The gap between the 952 walk and the 896 walk is still to be filled. However, carring out further chromosomal walking appears unnecessary, at present, because several cDNAs representing transcripts from the shak-B locus have been cloned and characterised, of which, P2.4, almost certainly represents the transcript of shak-B(neural) (Krishnan et al., 1993), whereas, S2.2 appears to represent the transcript of shak-B(lethal) (Crompton et al., 1995). Proteins encoded by these transcripts have been analysed and thus the possible function of these proteins can be postulated. As further research, such as in situ hybridization, antibody staining, and phenotypic rescue, are carried out, the function of these transcripts will be revealed. It will be known by then if it is necessary to carry out further chromosomal walking. Nevertheless, lessons have been learnt through the course of extending the chromosomal walk.

The fact that utilising cosmid clones failed to extend the chromosomal walk any further is very likely due to the characteristics of the genomic region around 19E3 which is a heterochromatin-euchromotin transition region. This region spans19D3 to 20F2 and contains a raletively high density of repetitive sequences (Miklos et al., 1984; Yamamoto et al., 1990). In addition, cosmid clones are known not to be stable and sometimes recombination, resulting in inversions, deletions and so on, occurs when a cosmid clone is grown up.

Most of the repetitive sequences are due to mobile elements, and they are variable from strain to strain. The fact that the DNA at distal end of λS2 etc does not overlap with λ995 is likely to be a consequence of this. Changing to a genomic library made from a different strain might be an alternative to continue the chromosomal walking (it certainly is in this case). This would also overcome the problem of underrepresentation of appropriate regions in other genomic libraries (in this case, the HE library).

By changing the genomic library combined with utilising the cosmid clones, the chromosomal walk has been extended approximately 27 kb towards the gap between the 952 walk and the 896 walk.

Chapter IV. Isolation and Characterisation of cDNAs from the *Shaking-B* Locus

4.1. PREVIOUS WORK

By using chromosomal walking, about 250 kb DNA in subdivision 19E3 on the X chromosome has been cloned. A considerable amount of molecular characterisation in this region has been done. In the 952 walk and the 896 walk, a few regions are known to be transcribed (figure 3.1.1.). However, the corresponding cDNAs derived from these transcribed regions, with two exceptions, have not been found. The two regions from which transcripts have been found (indicated as region 1 and region 2 in figure 3.1.1.) correspond to two separate groups of cDNAs (Davies *et al.*, personal communication).

In region 1, 2 cDNAs including KE2 have been identified at the time this study was begun. KE2 was originally a phage clone isolated from an embryonic cDNA library by Davies. The cDNA was then released from its original vector and re-cloned into pBluescript. It has been found that KE2 is 1817 bp in length and has 4 exons which hybridise discontinuously over 18 kb from the 952 walk. The genomic region to which λKE2 hybridises is bracketed by the distal breakpoint of *Df(1)LB6* and the proximal breakpoint of *Df(1)16-3-35* (Fig. 4.2.11.) (Crompton *et al.*, 1992).

There was evidence suggesting the existence of transcripts produced from *shak-B* by differential splicing and alternative promoter usage. It has been found that *shak-B^{L41}*, one of the alleles of *shak-B(lethal)* has a 17 bp deletion which removes the translation start codon of KE2, but *shak-B^{Pas}/shak-B^{L41}* heterozygotes have no neural phenotype (Baird *et al*, 1990). This implies that KE2 actually represents the transcription species from *shak-B(lethal)*, which does not affect the neural function of this locus. Therefore, there must be transcript(s) different from the transcript represented by KE2 from this locus. Such a transcript could be a splice variant produced from the same genomic region as KE2 and it could share common exons with KE2. It could

also come from one of the other transcribed regions defined by the reverse Northern technique on the cloned genomic DNA from the 952 walk.

There have been 6 putative cDNAs derived from region 2. Five of them, M11B, M12A, M13, M14 and M23, constituting the M21 cDNA family. It appears that these cDNAs are homologous. An additional cDNA, termed pMGB, independent of all members of M21 family was found from this region. However, there has been no evidence which suggests that these cDNAs represent the transcripts from the *shak-B* locus.

In order to find out which transcribed region or cDNA could come from the *shak-B* locus, *in situ* hybridization to fly sections has been carried out. This would also give the expression pattern of the RNA transcribed from these regions. Region 3 covered mostly by λ995 (one of the chromosomal walking phages) (see Fig. 3.1.1 and Fig. 3.2.1.) is transcribed and expressed in a number of cells in the fly head (Griffin, personal communication). This is consistent with the known phenotypes of the *shak-B* mutations. *Df(1)16-3-35* and *Df(1)LB6* are known to affect lethal and neural functions of shak-B (Baird *et al.*, 1990) and so it is likely that shak-B cDNAs span this region .

One of the purposes in this study is to isolate more transcript variants of KE2 and M21 cDNA family, and to isolate cDNAs derived from region 3 (see figure 3.1.1.).

4.2. RESULTS

4.2.1. Isolation of cDNA

Screening cDNA Library

The cDNA libraries screened were listed in Section 2.8.4. Approximately 1,000,000 cDNA clones were screened.

Most of the DNA fragments used for screening cDNA libraries were in three categories: 1) cDNA fragments from KE2, 2) cDNA fragments from

the M21 family, 3) Genomic DNA fragments from λ995. All DNA fragments were labeled with ^{32}P-dCTP in random primer reactions.

When KE2 was used as a probe, no positive plaque was seen. At the same time, cDNA S2.2 (originally known as B10) was isolated by using a PCR based method (Crompton *et al.*, 1995). Extensive analysis has shown that S2.2 is a splice variant of KE2 (Fig. 1.3. shows the organisation of KE2 and S2.2) (Crompton *et al.*, 1995). S2.2 was, therefore, used to probe cDNA libraries. One clone, designated as N52, was isolated from the NB library

The longest cDNA in M21 family, M23PS, was used as a probe. One clone, termed S3A was isolated from the SDA library. At the 3'end of genomic region 3, pMGB was used to screen the SDA library, two clones, S2 and S5, were isolated.

When screening the SDA library with genomic fragments from λ995 (see Fig 3.2.1.). two clones, 5A and 2B, were isolated by homology to a 4.3 kb EcoRI-EcoRI fragment and two adjacent SalI-SalI fragments of 1.4 kb and 3.3 kb , respectively.

Isolation of cDNA with PCR-based method

It would be nice if we could obtain all the cDNA splice variants by screening cDNA libraries. However, it has been difficult to isolate the splice variants of KE2 with this approach. Therefore, apart from screening cDNA library, at the same time we also adopted a PCR based strategy (illustrated in Fig. 4.2.1.).

This polymerase chain reaction (PCR) was used to look for clones related to KE2 (see Fig. 4.2.1.). One μl of the plasmid suspension of NB library was used as template. The primers were derived from both the sequence of the NB library vector at its 3'end (as NB1) and the sequence of KE2 from the second exon (as P11) (see Fig. 4.2.1.):

Primer 1--NB1: 5'-GGAATTCCGG TGACACTATA GAATAAATTG C-3' (After the first 8 bp which has a built-in EcoRI restriction site, it comprises bases 2405-2430 of NB vector which is 2490 bp in length)
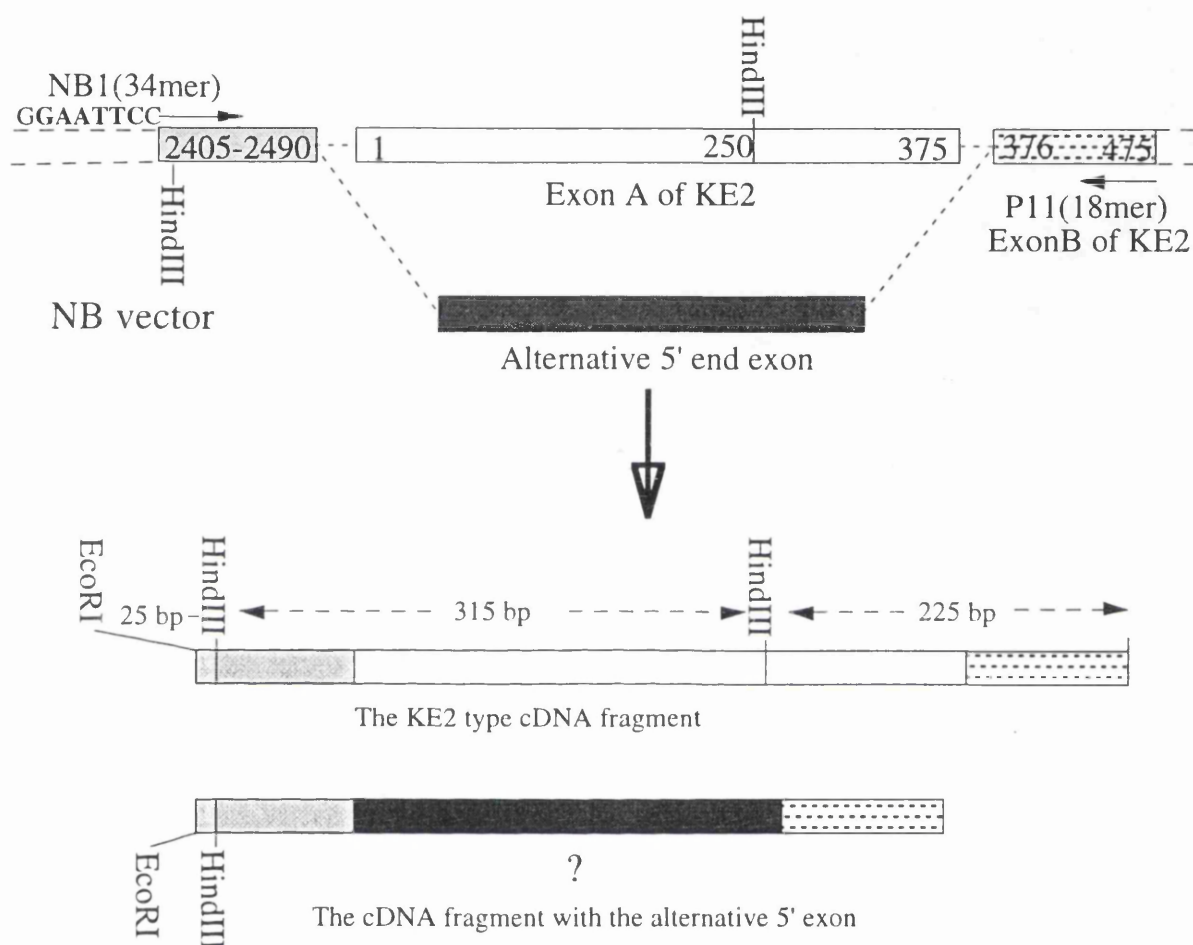

Fig. 4.2.1. A diagram showing the PCR-based strategy for amplifing the cDNA with its own distinct 5' end exon from that of KE2 and distinquishing the cDNA fragment with alternative 5' end exon from the KE2 type cDNA fregments. NB1 and p11 are the two primers for PCR.

Primer 2 --P11: 5'-CTTCGGGTCC GTTCCTC-3' (base 458-base 475 of the non-coding strand of KE2. Primers were designed and synthesized by Crompton.

This experiment was carried out as described in Section 2.5.5. It was performed three times. A total of 109 colonies with PCR-generated cDNA fragments as inserts were recovered (39 colonies at the first time; 0 and 70 at the second and third time, respectively).

4.2.2. Characterisation of Potential cDNAs of the *shaking-B* locus

Several clones which are homologous to either putative cDNAs or to DNA fragments from putative transcribed genomic region have been isolated by screening cDNA libraries. Meanwhile, a number of clones with PCR-amplified cDNA fragments as inserts have been generated. Physical characterisation of these clones are needed in order to identify novel cDNAs from the *shak-B* locus.

Restriction Mapping and Southern Analysis

To characterise the putative cDNA clones isolated from cDNA libraries (see Section 4.2.1.), their restriction maps have to be deduced first. The restriction maps of all the potential cDNAs should then be compared to those of the cDNAs and the genomic DNA fragments which were used to isolate them. This along with the hybridisation patterns of potential cDNA clones to their corresponding genomic regions would provide important evidence to determine if the potential cDNAs are real or not. Nevertheless, the last and most decisive step of cDNA characterisation is, of course, sequence analysis.
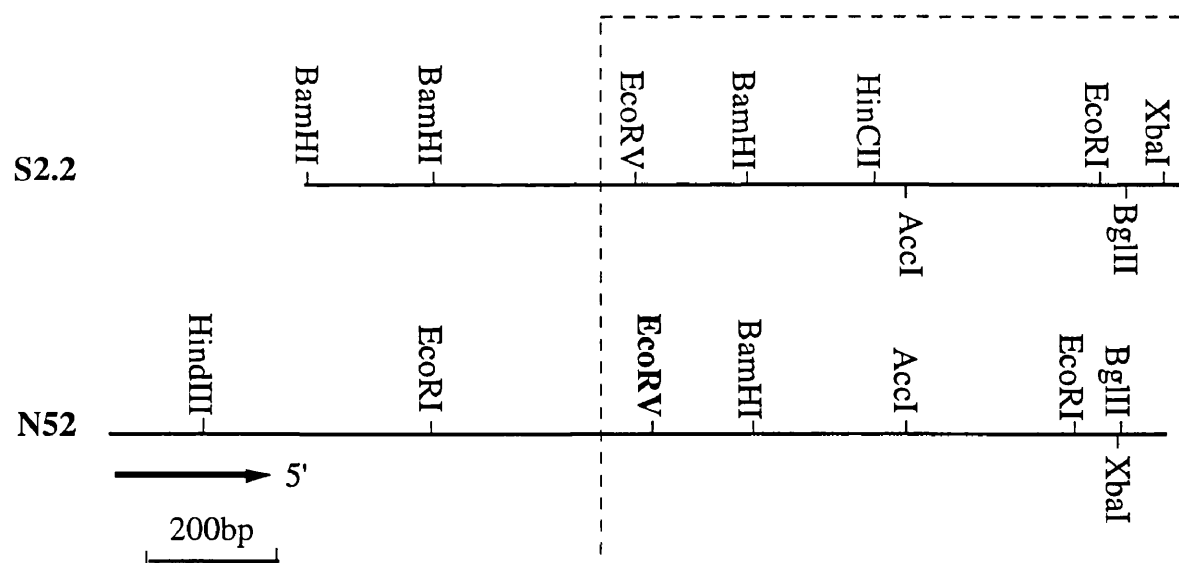

Figure 4.2.2. Comparison of the restriction maps of N52 (botttom) and S2.2 (top redrawn from Wilkin and Crompton). The dashed rectangular box indicates the portions of the two cDNAs with similar restriction patterns.

N52 The restriction map of N52 is shown in Fig. 4.2.2. By comparing N52 with part of S2.2, it has been found that the restriction maps of the two DNA fragments are very similar, although these are some variations in fragment sizes between the two, which could be due to inaccurate measurement of the DNA fragments during restriction mapping.
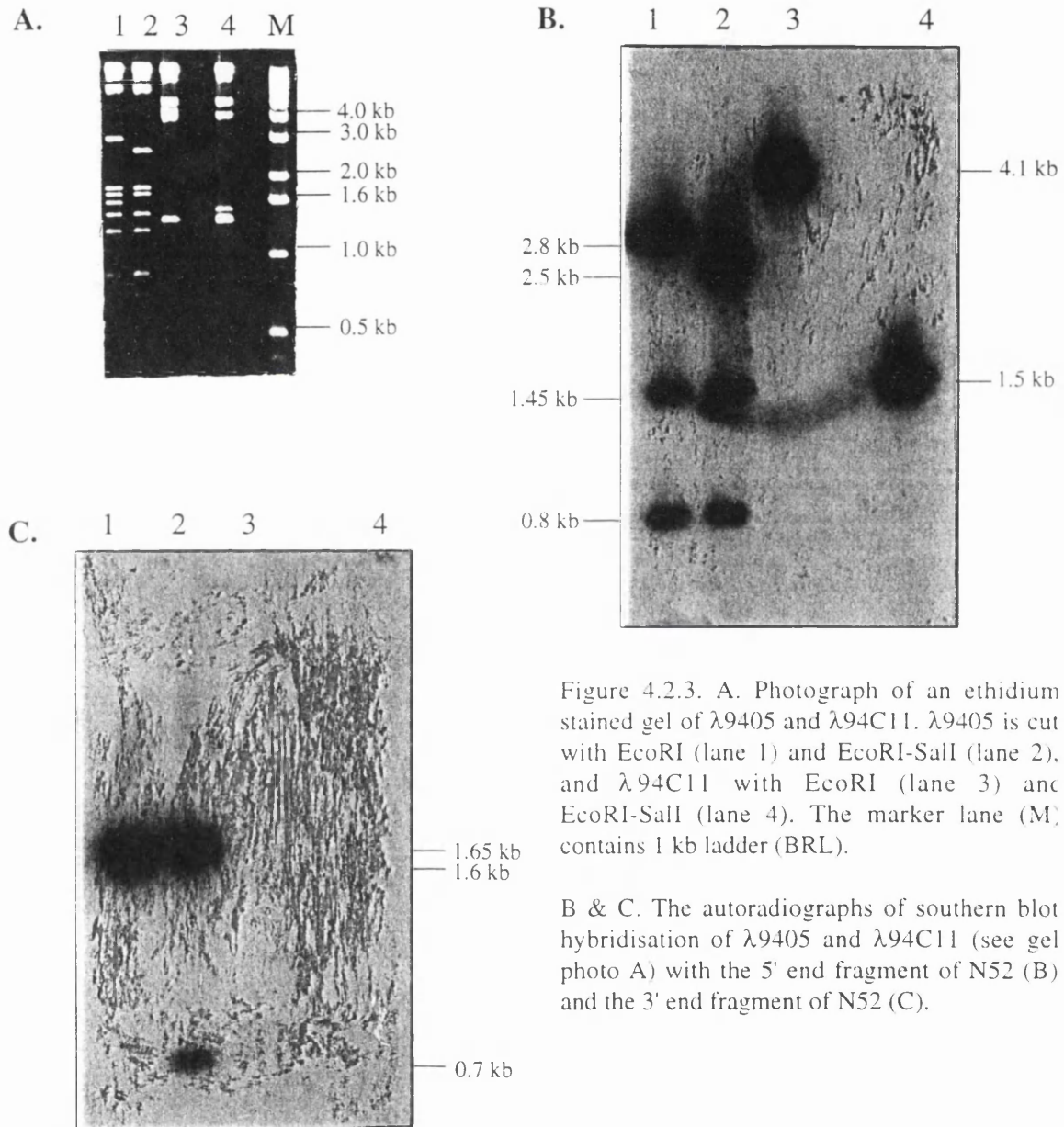

Figure 4.2.3. A. Photograph of an ethidium stained gel of λ9405 and λ94C11. λ9405 is cut with EcoRI (lane 1) and EcoRI-SalI (lane 2), and λ94C11 with EcoRI (lane 3) and EcoRI-SalI (lane 4). The marker lane (M) contains 1 kb ladder (BRL).

B & C. The autoradiographs of southern blot hybridisation of λ9405 and λ94C11 (see gel photo A) with the 5' end fragment of N52 (B) and the 3' end fragment of N52 (C).

To examine the hybridisation pattern of N52 to the corresponding genomic region, N52 was initially used to probe a dot blot made of DNA of all phages of the walk 952. Two phages, λ9405 and λ94C11, hybridised to N52. The insert of N52 was then cut into two fragments: a BglII-EcoRV fragment at the 5' end and a EcoRV-PvuII fragment at the 3' end (BglII and

PvuII are restriction sites at 3' and 5' end of the NB vector) (see Fig. 4.2.2.). These two fragments were used to probe two duplicate Southern blots of walk phages λ9405 and λ94C11 (Fig. 4.2.3a.) to which both KE2 and S2.2 have shown discontinuous hybridization patterns. The BglII-EcoRV fragment hybridises to the 0.8 kb, the 1.45 kb and the 2.8 kb EcoRI-EcoRI fragments of λ9405 and the 4.1 kb EcoRI-EcoRI fragment of λ94C11 (Fig. 4.2.3b.), whereas, the EcoRV-PvuII fragment hybridises to the 1.65 kb and the 1.6 kb EcoRI-EcoRI fragments of λ9405 (Fig. 4.2.3c. and Fig. 4.2.4.). Thus, N52 hybridises discontinuously to genomic DNA of 15 kb, approximately. These results strongly suggest that N52 represents a transcript from the *shak-B* locus

S3A S3A appears to have the same restriction pattern as M23ps with one exception. There is a BamHI site instead of an EcoRI site at the 3' end and the KpnI-BamHI fragment is shorter than the KpnI-EcoRI fragment (Fig. 4.2.5a.). This was found out first by mapping S3A, and later on confirmed by electrophoresis of both S3A and M23ps excised with the same enzymes on one agarose gel (Fig. 4.2.5b.). This suggested that S3A was possibly identical with M23ps with slight variation at 3' end which might be due to a polymorphism in the genome.
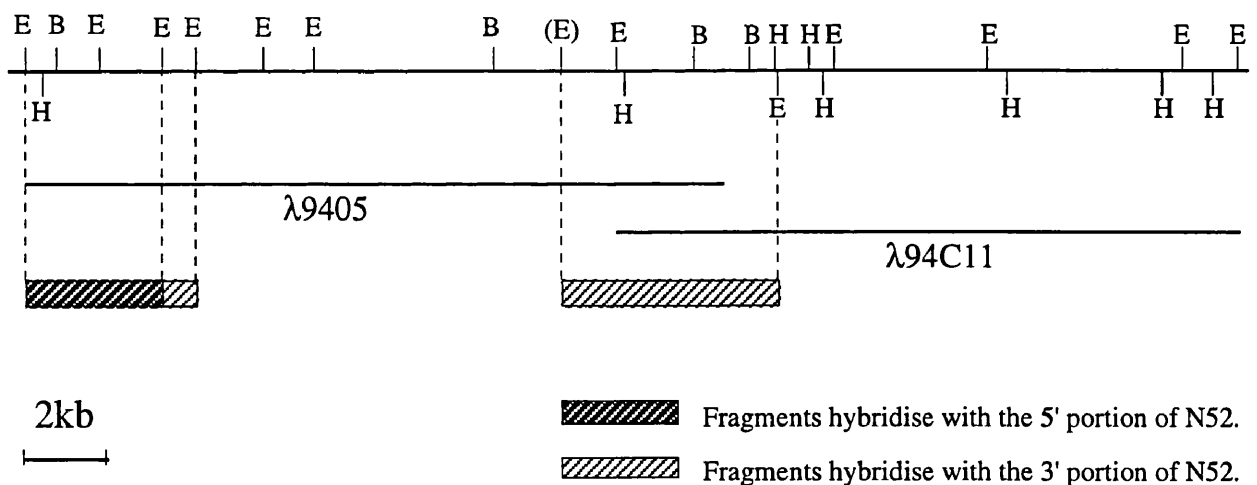

2kb

▓▓▓▓▓▓ Fragments hybridise with the 5' portion of N52.

▓▓▓▓▓▓ Fragments hybridise with the 3' portion of N52.

Figure 4.2.4. The restriction map of the 952 walk covered by λ9405 and λ94C11 (Data from Davies). Hatched bars below indicate the hybridization pattern of N52 to this region.

(E = EcoRI, B = BamHI, H = HindIII)

S2 and S5 S2 and S5 appear to have similar restriction patterns with difference of the fragment sizes (see Fig. 4.2.6a.). They do not have similarity with pMGB with respect to restriction maps. However, there are comparable restriction sites in the corresponding genomic region covered by genomic DNA fragment M2 (Fig. 4.2.6a.).

S2 and S5 were used to probe two duplicate blots of M2 (Fig. 4.2.6b.) The result has shown that both S2 and S5 hybridise to M2-1.1 kb HindIII-XhoI fragment and 0.8 kb HindIII-HindIII fragment (Fig. 4.2.6c). Therefore, it is likely that S2 and S5 are genomic DNA fragments which were cloned into the cDNA library due to genomic DNA contamination. The difference of the fragment sizes might be due to a polymorphism in the genome.

A.   **B.**
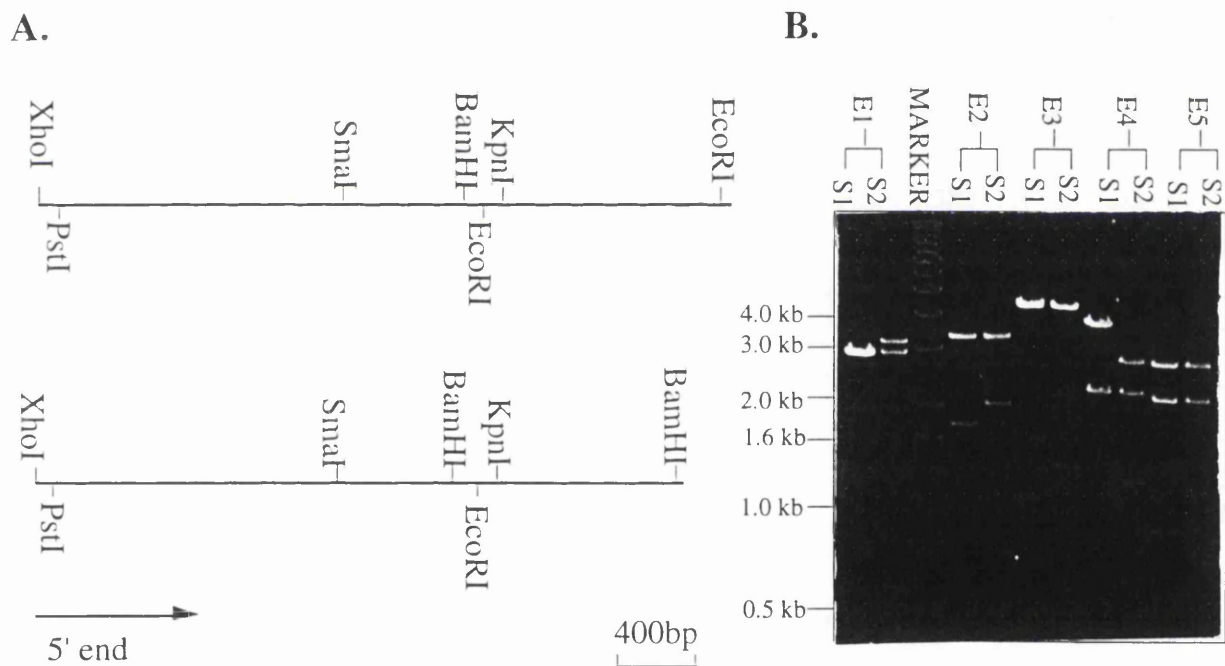

Figure 4.2.5. A. Comparison of the restriction maps of M23 (top, redrawn from Davies *et al.*) and S3A (Bottom).

B. Photograph of electrophoresis gel of S3A(S1) and M23ps (S2) cleaved by E1. PstI-XhoI; E2. PstI-SmaI; E3. SmaI-BamHI; E4. EcoRI-XhoI; E5. KpnI-BamHI. The marker lane contains 1 kb ladder (BRL).

2B and 5A Initially, to check the cross-hybridization pattern of 2B and 5A to the cloned genomic region, 2B and 5A were used to probe Southern blots of all the 952 walk phages. 5A hybridises to the 4.3 kb EcoRI-EcoRI

fragment and the 1.4 kb SalI-SalI fragment of λ995 though the signal was rather weak with the 1.4 kb fragment. 2B hybridises to both the 1.4 kb SalI-SalI fragment and the 3.3 kb SalI-SalI fragment (Fig. 4.2.7.) (Davies and Gardiner, personal communication).

After mapping both 2B and 5A, and the corresponding genomic DNA fragments (i.e. the three DNA fragments of 1.4 kb, 3.3 kb and 4.3 kb from λ995), the restriction maps of 2B and 5A were compared with those of the genomic DNA fragments. The results have shown that the restriction patterns of 2B and 5A are similar to their corresponding genomic regions (Fig 4.2.7.).

To give a more detailed cross-hybridization pattern, the inserts of 2B and 5A were excised with XhoI and PstI. This cut 2B into a 1.7 kb fragment (the 5' end part) and a 1.15 kb fragment (the 3' end part) and cut 5A into a 1.45 kb fragment (the 5' end part) and a 0.95 kb fragment (the 3' end part). When the two fragments of 2B were used to probe two duplicate southern blots of the 1.4 kb and the 3.3 kb SalI-SalI fragments of λ995 and 5A, the 3' end hybridised to the 1.1 kb PstI-PvuII fragment and the 600 bp PstI-SalI fragment and the 5' end hybridised to the 1.85 kb PstI-XbaI fragment. The 200 bp PvuII-SalI fragment did not cross-hybridise with ths 3' end. This might be due to the inefficiency of transfer of small DNA fragment to nitrocellulose membranes (Fig. 4.2.7.). When the two fragments of 5A were used to probe two duplicate Southern blots of the 1.4 kb SalI-SalI fragment, the 4.3 kb EcoRI-EcoRI fragment and 2B, the 5' end hybridised to the 1.7 kb HinCII fragment and the 880 bp HinCII-PstI fragment and the 3' end hybridised to the 200 bp PstI-HincII fragment (Fig 4.2.7. and Fig. 4.2.8.). However, 5A does not hybridise to the 1.4 kb SalI-SalI fragment (Fig. 4.2.8.).

The results from both comparison of restriction maps of 2B and 5A with that of λ995 and the hybridisation patterns of 2B and 5A to λ995 suggests that it is likely that 2B and 5A are genomic DNA fragments. Results from partial sequencing analysis also support this conclusion (data not shown).
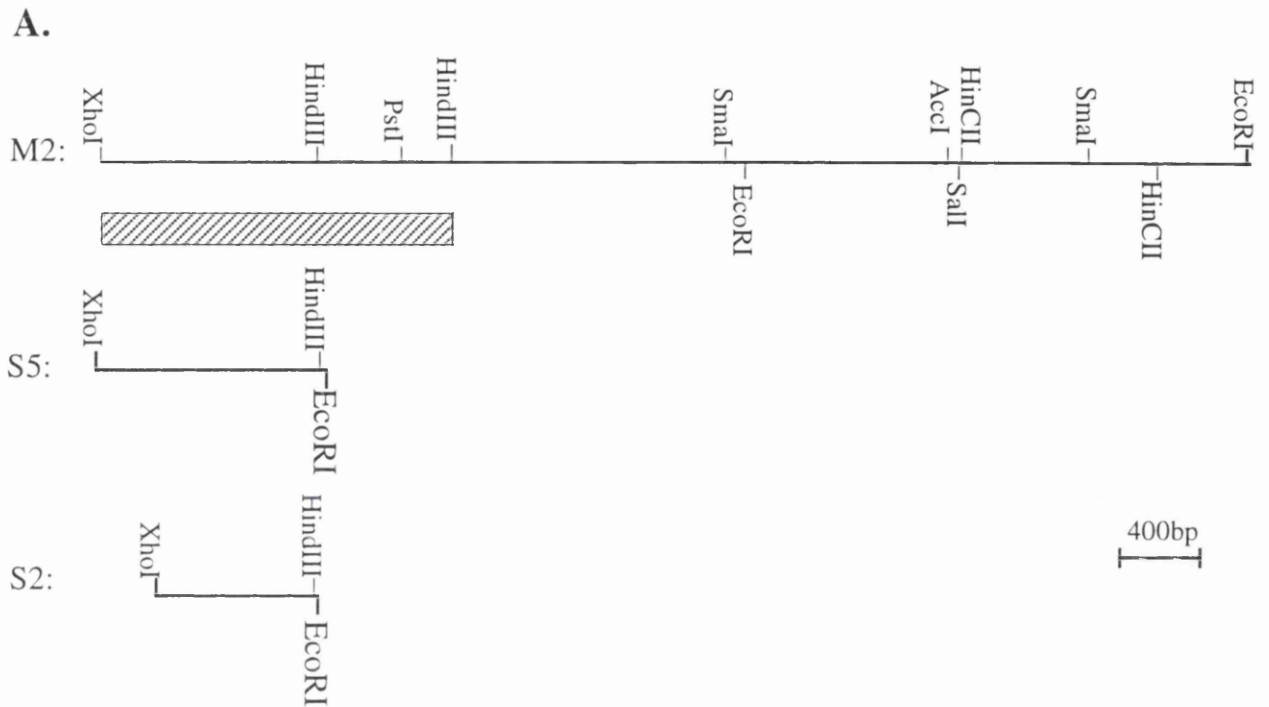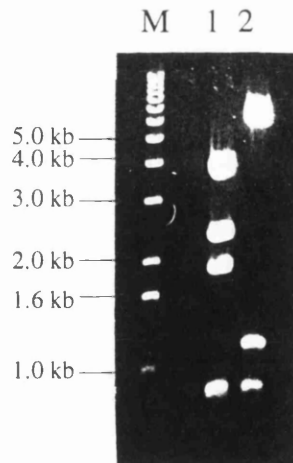
A.

B.

C.

Figure 4.2.6. A. Restriction maps of M2, S5 & S2. Hatched bar under M2 shows the portion of M2 to which S5 and S2 hybridise.
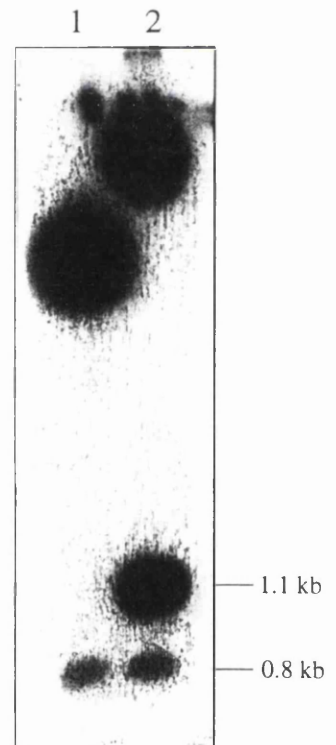
B. The photograph of the electrophoresis gel of M2 cut by EcoRI-HindIII (lane 1) and XhoI-HindIII (lane 2). The marker lane (M) contains a 1 kb ladder (BRL).

C. Autoradiograph of Southern blot hybridisation of M2 with S5 (the result of S2 is identical to S5)
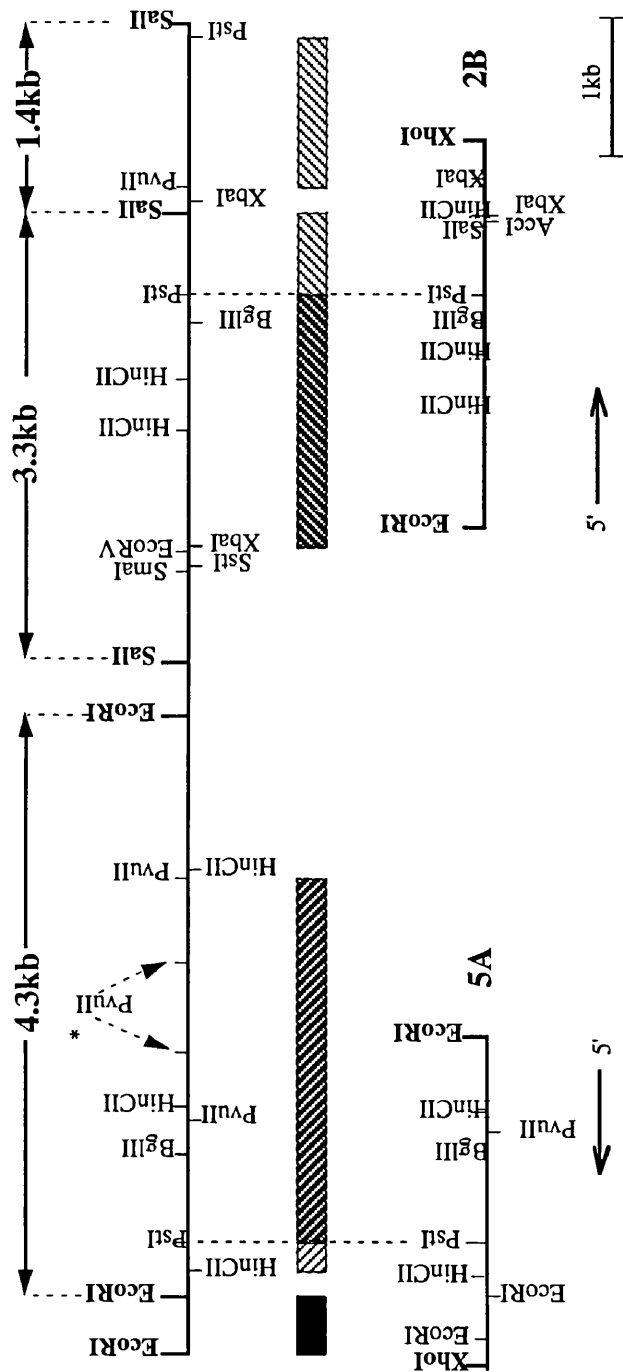
54

Figure 4.2.7. A Comparison of the restriction maps of 2B and 5A with the joined map of the genomic DNA fragments (the 1.4 kb SalI-SalI fragment, the 3.3 kb SalI-SalI fragment and the 4.3 kb EcoRI-EcoRI fragment of λ995).

The hatched bars under the genomic DNA fragments show the hybridisation pattern of 2B and 5A to the genomic region. The light left hatched bar indicates the genomic region to which the 3' portion of 2B hybridises, and the dark left hatched bar vs the 5' portion; the dark right hatched bar vs the 5' portion of 5A, the light right hatched bar vs the 3' portion (see text). The solid bar indicates the DNA fragment which does not hybridise to 5A (Gardiner and Davies), which was not examined in this study.

N. B. A previous Southern blotting experiment performed by Gardiner had shown that 5A hybridised to the 1.4 kb SalI-SalI fragment. The result from this study is opposite to the previous result.

* The PvuII site could be at either of the positions.

A.
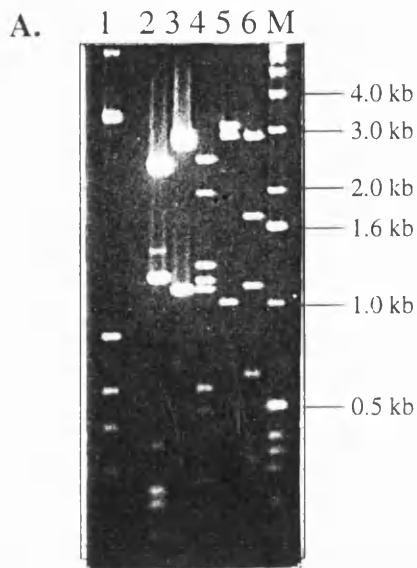
Figure 4.2.8. A. Photograph of electrophoresis gel of 2B, λ995-1.4 kb and λ995-4.3 kb. 2B is cut with PstI-HincII (lane 1), λ995-1.4 kb with SalI-PvuII (lane 2) and XbaI-PstI (lane 3), and λ995-4.3 kb with EcoRI-PvuII (lane 4), EcoRI-BglII (lane 5) and HincII-PstI (lane 6). The marker lane contains 1 kb ladder (BRL).

B &C. The autoradiographs of Southern blot hybridisation of 2B, λ995-1.4 kb and λ995-4.3 kb (see gel photo A) with the 5' end and the 3' end fragments of 5A (B and C, respectively).

B.

C.
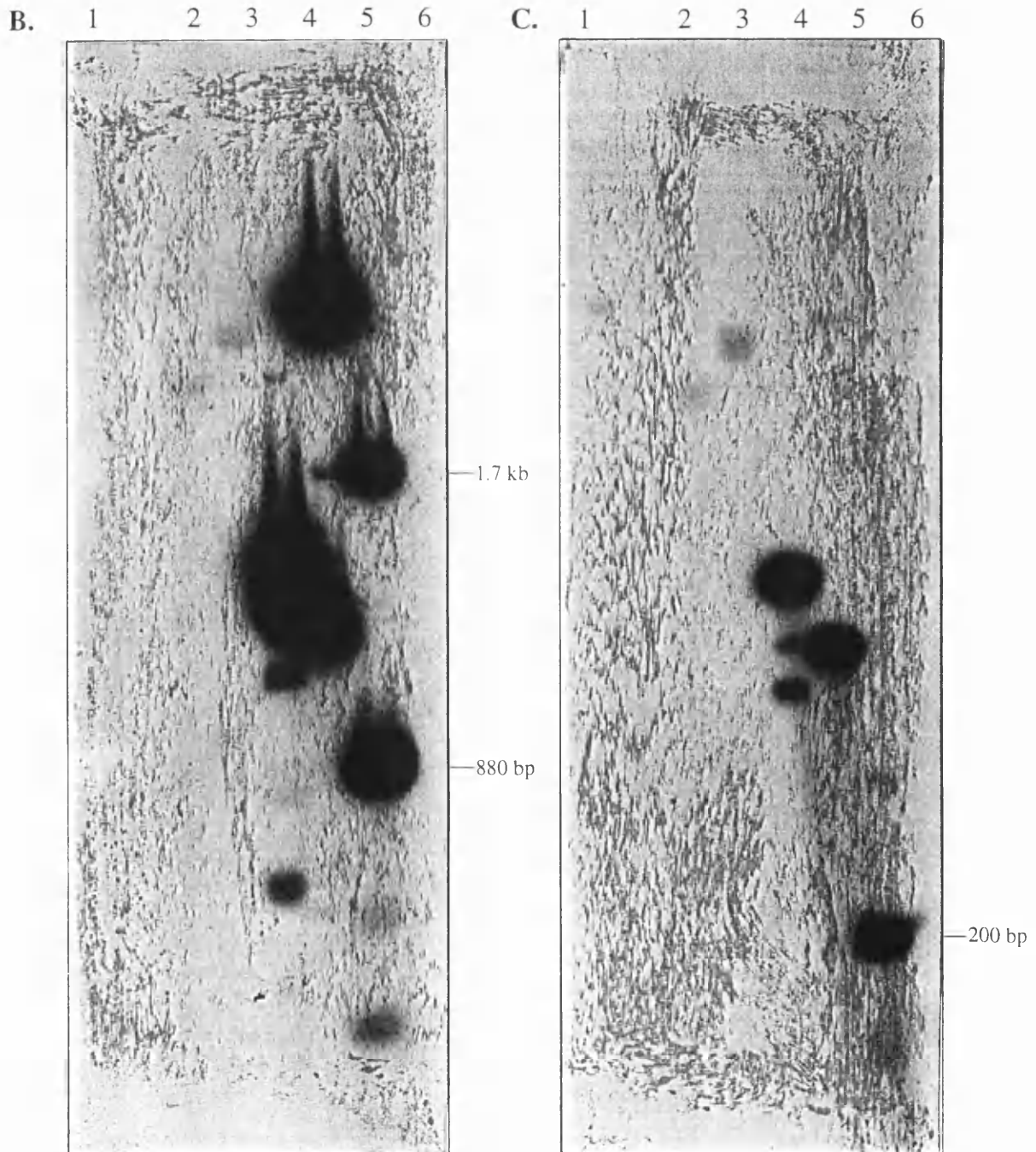
Screening PCR Generated cDNA Clones for Alternative 5' End of KE2

109 clones were generated from cloning PCR products. Thirty-nine of them were cloned into a T-vector and the other 70 were cloned into pBKS(-).

Since one of the primers was from exon 2 of KE2, it was expected that some of these clones would be actually copies of a truncated version of KE2 (from base 1 to base 475). This type of clones would have inserts of 569 bp as there is 94 bp from primer NB1 to the 3' end of the NB vector. They should have a HindIII site at base 344 because KE2 has a HindIII site at base 250 (Fig. 4.2.1.). However, some of the clones would be expected to contain alternative 5' exons and part of exon B of KE2 (from base 376 to base 475) (Fig. 4.2.1.).

To identify the clones with alternative 5' end, all clones were mapped first. They were cleaved with restriction enzymes. The DNA was then separated on agarose gel by electrophoresis and transferred onto a nitro-cellulose membrane and probed with two DNA fragments generated by excising KE2 with HindIII: a 250 bp fragment at 5' the end of KE2 and the rest of it (1550 bp). A clone containing a different 5' exon from KE2 would not hybridise to the 250 bp fragment at the 5' end of KE2.

No. of clones	Features
5	No insert (artifact generated during cloning)
25	No cross-hybridization with KE2 (artifact generated during PCR)
9	KE2 type

Table 4.2.1. -- The result of screening the 39 clones of which the inserts were generated in the first PCR experiment.

Among the 39 clones which were generated in the first PCR experiment, 9 of them were the KE2 type. Two fragments (340 bp and 220 bp in length) were obtained when cleaved by HindIII. The 340 bp long fragment hybridises to the 250 bp long fragment at 5' end of KE2, and the 220 bp long fragment

hybridises to the rest of KE2. None of the 39 clones appeared to be the one with a alternative 5' end (see Table 4.2.1.).

No. of clones	Features
22	Some endonuclease recognition sites, such as EcoRI, missing. Artefact during PCR.
36	No insert (artefact during cloning).
4	KE2 type, e.g. W1.
2	With inserts of approx. 520 bp, which hybridises to the 250 bp fragment at the 5' end of KE2 as well as the rest of KE2, e.g. W4.
6	With inserts of approx. 260 bp, which hybridises to the 1550 bp fragment at the 3' end of KE2, e.g. W2.

Table 4.2.2. -- The result of screening the 39 clones whose inserts were generated in the third PCR experiment.

When the 70 clones which were generated in the third PCR experiment were screened, 22 of them did not have the EcoRI recognition sites which should have been present (see Fig. 4.2.1.). They were thought to be an artefacts of the polymerase chain reaction. Further characterisation of these clones were not carried out. 36 of the 70 clones appeared to have no insert. When the other 12 clones were restriction mapped, blotted and probed with the two KE2 fragments (i.e. the 250 bp fragment at 5' end and the 1550 bp fragment at 3' end), they fell into three groups. The first group contains 4 clones which are standard KE2 type. Two clones, W4 and B4, composing the second group, have similar restriction maps to that of the KE2 type clone, but the inserts are shorter (Fig. 4.2.9.). The other 6 clones in the third group, typical of W2, have inserts of about 280 bp and they hybridise only to the 1550 bp fragment (Fig. 4.2.9.). The results are summarised in table 4.2.2.

Two cDNA fragments representing cDNAs with 5' ends different from KE2 have possibly been obtained from screening these PCR-generated clones. They are represented by strain W2 and W4 and thus W2 and W4 will be further analysed.

Sequence Analysis of Potential cDNAs of the *shaking-B* locus

N52 N52 was cut into 7 fragments and subcloned into pBKS(-). Each fragment was sequenced from both directions with T3 and T7 as primers. Since all the fragments were small enough to complete sequence with these primers, the sequence of N52 was a sum of those of the 7 fragments.

N52 is 1661 bp in length (Fig. 4.2.10.). The putitive protein encoded by N52 is 229 aa. Termination codons exist upstream of the putitive translation start codon in all three frames.
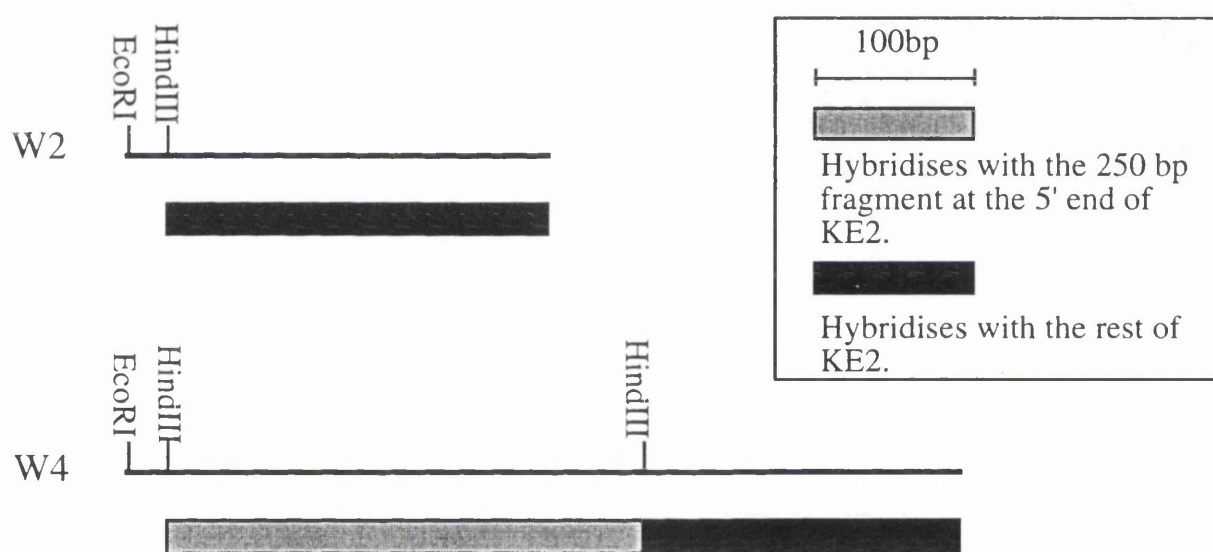

Figure 4.2.9. Restriction maps of W2 and W4. Shadowed bars below each restriction maps indicates the hybridisation patterns of KE2 to W2 and W4.

From the genomic region from which N52 is derived, a few other transcripts, represented by KE2, S2.2, P2.4 and etc., have also been isolated. Sequence analysis of this genomic region has been done, which provides the necessary data for analysis of N52 (Crompton, personal communication). Comparison of N52 and its corresponding genomic DNA reveals that N52 contains 6 exons of which the first one (A1) is 699 bp. The next five exons are found in S2.2 and P2.4 with lengths of 155 bp (D), 65 bp (E), 165 bp (F), 181 bp (G) and 396 bp (H) (the last exon of P2.4 is longer than that of N52 and S2.2) (Fig. 4.2.11) and thus the sequence of N52 from base 700 (exon D) to the end is identical to S2.2 and P2.4. The sequence of exonA1 (the first 291 bp) of N52 is partially identical to that of exon A (from base 87 to base 375)

of KE2 and S2.2 except that there are two additional thymidines (T) after base 105 (Fig. 4.2.10.). The open reading frame of N52 starts at its second exon (D) and ends to its last exon (Fig. 4.2.11.).

Unlike KE2 and S2.2, N52 does not have an extra G residue at its 5' end. The extra G residues at the 5' ends of KE2 and S2.2 are not encoded by the genome. They are likely to be derived from 7-methyl guanosine caps, suggesting that KE2 and S2.2 include the full 5' extents of the transcripts which they represent. The transcript represented by N52 appears to have the same transcription start site as KE2 and S2.2 (Crompton, personal communication), thus N52 is truncated (a 86 bp deletion) at its 5' end. N52 has the same 3' end as S2.2. The poly-A sequence at the 3' ends of the two cDNAs corresponding to an A-rich sequence present in the genome. There is no recognisable polyadenylation signal upstream of the A-rich sequence,suggesting that S2.2 and N52 are likely to be internally primed (Crompton et al., 1995). However, their reading frames are not disrupted by these deletions at the 5' end of N52 and the 3' end.

N52, S2.2, KE2, P2.4 and a few other cDNAs (not mentioned here) (Crompton, personal communication) represent transcripts from the shakB locus, produced by differential splicing and alternative promoter usage (Fig. 4.2.11.). The proteins encoded by these cDNAs shows interesting relationship. Among them, S2.2 and P2.4 share common C-termini of 252 residues and have distinct N-termini of 120 residues and 109 residues, for S2.2 and P2.4, respectively. The 120 aa long KE2 protein is identical to the 120 residues at N-terminus of S2.2 protein, whereas, the 229 aa long N52 protein is identical to S2.2 and P2.4 proteins at their common C-terminus (Fig. 4.2.12.).

Since N52 shares exon D, E, F, G and H with S2.2 and P2.4, the shak-B^{EC201} mutation, which disrupts the open reading frame of S2.2 and P2.4, must also disrupt the open reading frame of N52.

```
  1 CTGATCTCGAAGAAAAGCACGGACAAAAAAAAAAACCTGAAACCGGGGAAAAGCCGTATT 60
                                                  extra bases
                                                     * *
 61 CCGATTTCTTTTCCGTAAACGCAAACCCATCGAAGTTTTTTTTTTTTACCAGTGAGCGAT 120
121 CGAAAAATGTGAGGTTGAGGTGCACAGAGTCACCCAGAGTCATATAAGCTTCGACCATGT 180
181 CGGACACGTGACCCGCGCTTATGCCGCGCGATGAAACCGATAACACGAGCTGACTAAGCC 240
241 GATTAGGCCGATAGCAACGATAGCGTCGATAGCCCAAATCAAACGACGAGTGTGAGAATT 300
301 AAATTCGCGCAACTAGTGTGCAAACAAACACCGAGCGTTTTCTATATTTTCACAACACCC 360
361 ACGAAAGATCATTGACTGCAATGCGGGATCATAACATATAGTAGGAAGCCACATGCCCAC 420
421 AGAATCAGAAATTATATTTATATATACACGTATATATAGTATTGTAGGGGCTGTGTGCAG 480
481 GCCACCGCTTGATCATAATTAACTCTTACAAAGTAAATAAACCAACTGCGAATTCGGGGG 540
541 CCATGTGTTGCCCCCGAAACCGAAAACCAAAAACCGAAAACAAACAGGCGATCCACAAAT 600
601 TTTTACAATTAACATTCGACACGATTGTCCAAAAACCAAAATGGAAACCGAAAAAACCAA 660
661 AATACAAATTTTATTTTTCATTTTATTTTTATATTTTATGCAATCTTATTTTATACACCA 720
721 AGATGGCTGTGGAAATCTTGGGAGGGTGGCAAGATTCATGCGCTCATCATGGACTTAGAC 780
  1                                                    M  D  L  D  4
781 ATAGGCATTTGTTCCGAAGCCGAGAAAAAACAAAAAAAGAAATTACTTTTAGATTATTTG 840
  5  I  G  I  C  S  E  A  E  K  K  Q  K  K  K  L  L  L  D  Y  L  24
841 TGGGAAAATTTAAGATATCACAATTGGTGGGCGTACAGATATTACGTGTGTGAGCTGCTC 900
 25 W  E  N  L  R  Y  H  N  W  W  A  Y  R  Y  Y  V  C  E  L  L  44
901 GCCCTTATAAATGTGATAGGTCAAATGTTTCTTATGAATCGATTTTTCGATGGCGAATTT 960
 45 A  L  I  N  V  I  G  Q  M  F  L  M  N  R  F  F  D  G  E  F  64
961 ATAACATTTGGCCTGAAAGTGATAGATTATATGGAGACCGATCAGGAAGATCGCATGGAT 1020
 65 I  T  F  G  L  K  V  I  D  Y  M  E  T  D  Q  E  D  R  M  D  84
1021 CCGATGATTTACATGTTTCCCAGAATGACCAAATGTACATTTTTTAAATATGGTTCCAGT 1080
 85 P  M  I  Y  M  F  P  R  M  T  K  C  T  F  F  K  Y  G  S  S  104
1080 GGGGAGGTGGAGAAACACGACGCCATTTGCATTTTACCATTAAACGTTGTTAATGAGAAG 1140
105 G  E  V  E  K  H  D  A  I  C  I  L  P  L  N  V  V  N  E  K  124
1141 ATTTACATTTTCCTTTGGTTTTGGTTTATATTATTAACGTTTCTCACATTATTAACGCTA 1200
125 I  Y  I  F  L  W  F  W  F  I  L  L  T ·F  L  T  L  L  T  L  144
1201 ATATACAGGGTGGTTATTATATTCTCTCCTCGAATGAGGGTCTACTTATTTCGTATGCGA 1260
145 I  Y  R  V  V  I  I  F  S  P  R  M  R  V  Y  L  F  R  M  R  164
1261 TTTAGGTTAGTGCGTCGTGACGCTATTGAAATAATCGTTCGTCGTTCAAAGATGGGCGAT 1320
165 F  R  L  V  R  R  D  A  I  E  I  I  V  R  R  S  K  M  G  D  184
1321 TGGTTTTTGTTGTATTTACTAGGTGAAAACATAGATACAGTTATATTTCGTGATGTTGTA 1380
185 W  F  L  L  Y  L  L  G  E  N  I  D  T  V  I  F  R  D  V  V  204
1381 CAGGACTTAGCGAATCGTTTAGGACATAACCAACACCACAGGGTGCCTGGCTTAAAAGGT 1440
205 Q  D  L  A  N  R  L  G  H  N  Q  H  H  R  V  P  G  L  K  G  224
1441 GAAATACAGGATGCATGATATTGGGAGTATTAGAAACAATACAAAATGCAATTTGTCGTC 1500
225 E  I  Q  D  A                                          229
1501 TCCATTTAAAAACCATCGAATTCAATAACAAAAATGTCGAAAGCAAGAAAAAGATTAAGA 1560
1561 AAGGACAATTACAACCACAAAGGAATCTAGAGATCTTCGCAGCAGCCGCTTCATTAAAAC 1620
1621 TTACAACTCAACACACCGCTAAAAAATCTAAAAAAAAAAAA                   1660
```
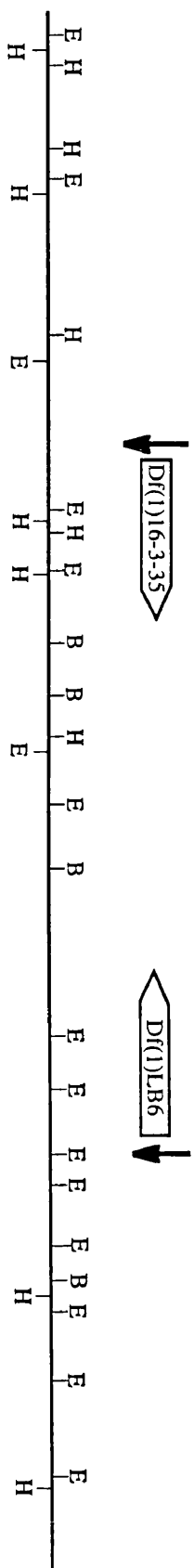
Fig. 4.2.10. The complete nucleotide sequence of N52 with the putative protein sequence which it encodes. The start codon is underlined. The two extra T residues (compare to the sequences of KE2 and S2.2) are indicated.
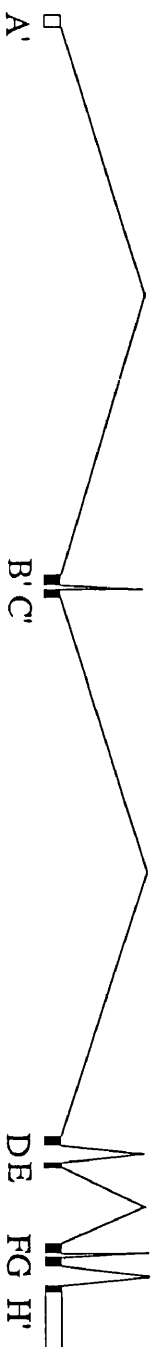
61

Figure 4.2.11. The physical organisation of the *shakB* locus (with major transcripts).

A. Map of genomic DNA with restriction sites shown. Black arrows above the restriction map indicate the breakpoints of *Df(1)16-3-35* and *Df(1)LB6*. The centromere on the DNA map is to the left, the telomere to the right (data are from Davies *et al.*).

B. Organization of the *shakB* major transcripts. (Data are from Crompton *et al.*, 1995 except P2.4* which is from Krishnan *et al.*, 1993). The exons A', B', C' H' are those in P2.4 which are different from those in the other transcripts. The locations of these exons are not accurate because the genomic DNA was mapped with different enzymes between the two groups.

W2 and W4 W2 and W4 are different from the expected KE2 type clones generated by PCR-based method with respect to their sizes and restriction maps. W2 does not hybridise to the 250 bp fragment at the 5'-end of KE2. W2 and W4 might represent the transcripts with 5' exons differnt from that of KE2, thus sequence analysis of them is necessary. W2 and W4 were sequenced from both directions with T3 and T7 primers. Since both of them hybridised to KE2, the sequences of W2 and W4 were compared to that of KE2, first.

The sequence of W2 from base 2 to the end is identical to that of KE2 from base 307 to the base 475. The first residue (G) at the 5' end of W2 is different from the 306th residue (A) of KE2 and it is not encoded by the genome. It is likely that the G residue is derived from 7-methyl guanosine caps, which suggests a new transcription start site of this locus. W2, therefore, might represent a novel transcript which had a distinct 5' non-coding region (see Fig. 4.2.11.). This was later proved to be the case. A cDNA, termed DC224, was isolated by using an adjusted PCR-based method. Preliminary analysis has shown that DC224 has the same 5' exons as W2 (Crompton, personal communication).

W4 contains an insert of 429 bp excluding the 94 bp which is the sequence of NB library vector at it's 3' end and the artificially added 8 bp in primer NB1. It is identical across its whole length with the sequence of the first 429 residues of KE2. Thus, W4 is a KE2 type clone (or S2.2) with a 46 bp deletion (from base 430 to base 475) at its 3' end. The deletion possibly occurred during cloning the PCR products because the sequence of P11 (from base 458 to base 475 of KE2) is missing.
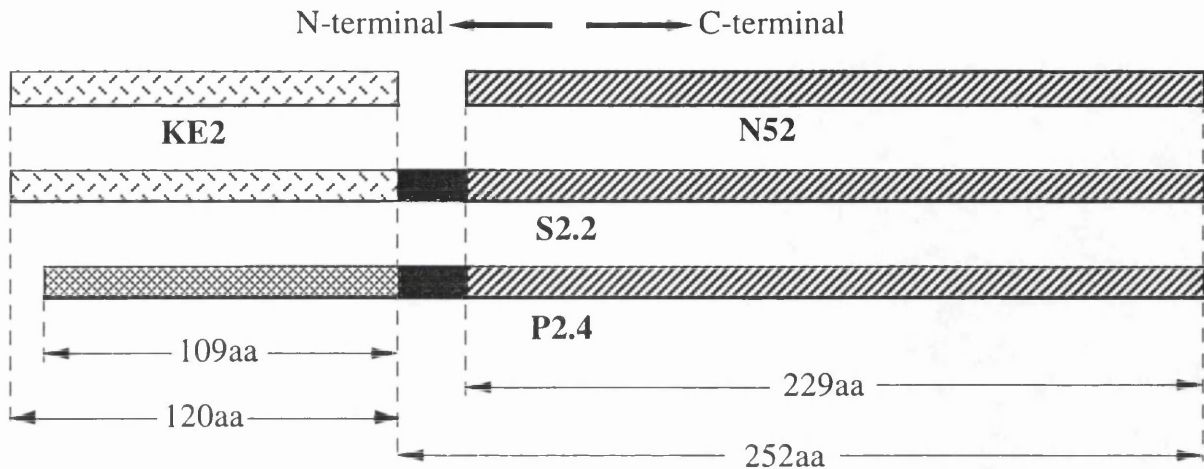
Figure 4.2.12. A diagram comparing the putitive proteins encoded by KE2, N52, S2.2 and P2.4. Each bar represents a protein. Identical amino acid sequences between different proteins are shown with identical patterns.

4.3. DISCUSSION

4.3.1. Isolation of cDNA

An essential step to molecularly analyse a gene is to obtain cDNAs representing the transcripts from this gene. To obtain more cDNA derived from the *shak-B* locus, two strategies were adopted, of which one was the conventional method, cDNA library screening and the other was a PCR based method. The later one appears to be more promising in this case, conversely, the former one give poor results except that N52 was obtained by the former method. The reasons might be:

(1) The level of gene expression is relatively low. It had been known that KE2 represented one of the transcripts of *shak-B* and other transcripts produced by differential splicing and alternative promoter usage must exist (Crompton *et al.*, 1992). However, the attempt to obtain KE2 homologue(s) by screening a cDNA library has not been very successful. This leads to the possibility that the expression level of the transcripts represented by KE2 and related transcripts is low. Although the reading frame of KE2 and the first 120 codons of S2.2 have codon usage in reasonal agreement with that bias in *Drosophila* (Crompton *et al.*, 1992; 1995), P2.4 has unbiased codon usage (Krishnan *et al.*, 1993). Since proteins encoded by S2.2 and N52 share

common C-termini with P2.4, presumably, the C-terminus of S2.2 and N52 would have unbiased codon usage. In less expressed genes, alternative codon usage is usually found instead of standard codon usage (Ikemura, 1985). This suggests that at least some of the transcripts represented by members of KE2 family are of low abundance. This conclusion is consistent with the fact that all attempts to detect them on Northern blots failed with only one exception when KE2 was used as a probe.

(2) The half-life of the mRNA is short. This could lead to rapid degradation of mRNA and thus the cDNA library would be underrepresentitive of some transcripts. However, there is no evidence, so far, to prove that the half life of *shak-B* transcripts are short

(3) The poor quality of cDNA libraries, in particular, the SDA library which appears to be contaminated by genomic DNAs. Several clones, S2, S5, 2B and 5A (and a few more clones which were not mentioned), turned out to be genomic DNA contaminants of the SDA library

On the other hand, the PCR based method is a more sensitive way to detect low abundance transcripts. A PCR-based method has been used to amplify full-length cDNAs of rarely transcribed RNAs by directly using RNAs as templete (Frohman *et al.*, 1988). The method used in this study was more productive in spite of its shortcomings, such as generating truncated cDNAs (e.g. W2). This problem can be overcome by adjusting the PCR based method. Several other transcripts have been obtained with the adjusted method (Crompton, personal communication).

The attempt to obtain cDNAs by using cDNAs M23ps and pMGB and genomic DNA fragments from λ995 was not successful. The reasons might be the same as mentioned above.

It has been suggested that S2.2 and P2.4 represent transcripts of *shak-B*, serving the two dintinct functions of the locus (Crompton *et al.*, 1995). In contrast to those cDNAs of the KE2 family, the roles played by the other transcripts represented by the M23 family and detected by reverse Northerns are completely unknown. Nevertheless, the *shak-B* locus is very complex

genetically (Baird *et al.*, 1990; Chapter I). Its genetic complexity must be underpined by its molecular complexity. It is difficult to come to any kind of conclusion about the roles played by the transcripts other than the KE2 family from subdivision 19E3 at the current stage.

4.3.2. The Roles of the KE2 cDNA Family

The KE2 cDNA family contains a number of cDNAs representing transcripts produced by differential splicing and alternative promoter usage from the *shak-B* locus. Among them, KE2, P2.4, S2.2 and N52 are best characterised (Crompton *et al.*, 1992; Krishnan *et al.*, 1993; Crompton *et at.*, 1995).

The fact that the translation start codon for the putitive proteins encoded by KE2 and S2.2 is removed by lethal mutant *shak-B^{L41}* eliminate the possibility that either of them represents a transcript of *shak-B(neural)* because *shak-B^{L41}* compliments the neural function of *shak-B^{Pas}* and *shak-B^2*, the two major neural mutants of *shak-B* (Baird *et al*, 1990). N52 does not contain exon B which has the translation start codon for the proteins encoded by KE2 and S2.2, which means that the lethal mutant *shak-B^{L41}* can not affect the N52 gene product. We initially thought that N52 protein might contribute to the function of *shak-B(neural)*. However, the possibility that none of these three cDNAs (KE2, S2.2 and N52) was derived from *shak-B(neural)* could not be ruled out, since it was known that some fonctions of *shak-B* lay at least partly upstream of the proximal breakpoint of *16-3-35* (Baird *et al.*, 1990). There might be other gene product(s) serving the neural function. Later on, a cDNA, P2.4, representing such a transcript was isolated by Krishnan *et al.* The first exon of P2.4, which contains the first 2 bp of the initiating methionine, is located beyond the proximal breakpoint of *Df(1)16-3-35*. The neural mutations, *shak-B^2* and *shak-B^{Pas}* map to the unique 5' exons of P2.4 (Krishnan *et al.*, 1993).

Based on the positions of different classes of *shak-B* mutant lesions within the proposed coding regions of these transcripts, a simple molecular model has been suggested to account for the complex complementation relationships of *shak-B* alleles. In this model, S2.2 and P2.4 are essential and neural transcripts, respectively (Crompton *et al.*, 1995).

In this model, however, only two of the *shak-B* transcripts are taken into account for the complex genetics of the locus. Apart from S2.2, P2.4, several other transcripts including those represented by KE2 and N52 have been identified to date. Each of these encodes a different protein. Translation of the products of some of these cDNAs show that one of the putitive proteins, KE2, is initiated from the translation start codon disrupted by the *shak-B*^{L41} lesion, while others, for example, N52, would be disrupted by the *shak-B*^{EC201} mutation. The fact that *shak-B*^{L41} and *shak-B*^{EC201} fail to complement eacn other implies that a protein disrupted by both of these mutitions must be necessary for viability. As yet only the S2.2 reading frame fits these criteria. Likewise, the fact that *shak-B*^{EC201} can not compliment *shak-B*^{Pas} and *shak-B*^{2} indicates that the P2.4 transcript form is required for the formation of synaptic connections between the GFS and its target cells.

If it is, as seems to be the case, that the transcripts represented by P2.4 and S2.2 serve the two distinct genetic functions of the *shak-B* locus. In the adult central nervous system, and in the pupa during GF synapse formation, the P2.4 transcript is expressed in the GF and in cells in the location of its postsynaptic targets (Krishnan *et al.*, 1993). The expression pattern of this *shak-B* transcript suggests it has a role in the formation and maintenance of synaptic connection between GF and its target neurons. The S2.2 transcript is also expressed in the developing pupal nervous system, although not in the GFs (Crompton *et al.*, 1995), suggesting that it also have a role in the pupal nervous system as well as its function in the embryo. Apart from S2.2 and P2.4, do all the other transcripts represented by cDNAs in the KE2 family serve any function at all?

It has been known that other behavioural abnormalities have been observed in *shak-B* flies, which have hardly been characterised physiologically (Balakrishnan and Rodrigues, 1991; Homyk *et al.*, 1980; Phillis *et al.*, 1993; O'Dell, personal communication). The expression of P2.4, according to Krishnan and colleagues, is limited on a small number of cells including GF cell bodies, both or either of TTMn and PSI, starting from 75 hrs after pupation (Krishnan *et al.*, 1993). It is not known if Krishnan and colleagues used the neural specific part of P2.4 (i.e. the 5' end exons) for

in situ hybridisation. If the function of P2.4 is only limited to the development and maintenance of the GFS, at least one more transcript is required for the proper development of the other neural pathways which appear to be afffected by mutations in *shak-B*.

Crompton *et al.* have demonstrated that the *shak-B* transcripts are expressed in a wider range of cells in the nervous system, starting from as early as third instar larvae. The probe (pL) which could not detect P2.4, the neural transcript, was seen to hybridise to cells in various regions of the central nervous system though the hybridisation patterns are not the same as that seen with the probe (pNL) which could detect all known transcripts (Crompton *et al.*, 1995). The differences of the *in situ* hybridisation patterns resulting from the usage of different probes (pL and pNL) imply that the transcripts, such as KE2 and N52, might have their own distinct expression patterns.

At the current stage, it can not yet be argued that any identified *shak-B* splice form or combination thereof is sufficient to account for either identified function of *shaking-B*. Wether all *shak-B* transcript forms isolated to date serve a specific developmental function, or wether some may produce non-functional proteins and are simply a reflection of alternative splicing as a mechanism of gene regulation is not yet clear. We must await results from further experiments, especially, the phenotypic rescue experiments using cDNA constructs alone and in combination, in order to resolve this issue.

REFERENCES

Baird, D. H., Schalet, A. P. & Wyman, R. W. (1990) The *passover* locus in *Drosophila melanogaster*: complex complementation and different effects on the giant fiber neural pathway. *Genetics* **126**: 1045-1059.

Baird, D. H., Koto, M. & Wyman, R. W. (1993) Dentritic reduction in *passover,* a *Drosophila* mutant with a defective giant fiber neural pathway. *J. Neurobiology* **24**: 971-084.

Balakrishnan, R. & Rodrigues, V. (1991) The shaker and shaking-B genes specify element in the processing of gustatory information in *Drosophila melanogaster. J. Exp. Biol.* **157**: 161-181.

Bastiani, M. J., Harrelson, A. L., Snow, P. M. & Goodman, C. S. (1987) Expression of fasciclin I and II glycoproteins on subsets of axon pathways during development in the grasshopper. *Cell* **48**: 745-755.

Bentley, D. & O'Connor, T. P. (1992) Guidance and steering of peripheral pioneer growth cones in grasshopper embryos. In The Nerve Growth Cone, P.C. Letorneau, S. B. Kater, and E. R. Macagno. eds. (New York; Raven Press, Ltd.) pp. 265-282.

Birnboim, H. C. & Doly, J. (1979) A rapid alkeline extraction procedure for screening recombinant DNA. *Nucleic Acids Res.* **7**: 1513-1523.

Broadie, K., Sink, H., Vactor, D. V., Fambrough, D., Whitington, P. M., Bate, M. & Goodman, C. S. (1993) From growth cone to synapse: the life history of the RP3 motor neuron. *Development (Suppl.)* : 227-238.

Cash, S., Chiba, A. & Keshishian, H. (1992) Alternative neuromuscular target selection following the loss of single muscle fibre in *Drosophila. J. Neuro sci.* **10**: 1134-1153

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

Chiba, A., Snow, P., Keshishian, H. & Hotta, Y. (1995) Fasciclin III as a synaptic target recognition molecule in *Drosophila. Nature* **374**: 166-168.

Crompton, D. E., Griffin, A., Davies, J. A. & Miklos, G. L. G. (1992) Analysis of a cDNA from the neuroloinally active locus *shaking-B* (*Passover*) Of *Drosophila melanogaster. Gene* **122**: 385-386.

Crompton, D., Todman, M., Wilkin, M., Ji, S. & Davies, J. A. (1995) Essential and neural transcripts from the *Drosophila shaking-B* locus are

differentially expressed in the embryonic mesoderm and pupal nervous system. *Developmental Biology* **170**: 142-158.

Dan, Y. & Poo, M. M. (1992) Hebbian depression of isolated neuromuscular synapses in vitro. *Science* **256**: 1570-1573.

Elkins, T., Hortsch, M., Bieber, A. J., Snow, P. M. & Goodman, C. S. (1990) *Drosophila* Fasciclin I is a novel homophilic adhesion molecule that along with fasciclin III can mediate cell sorting. *J. Cell Biol.* **110**: 1825-1832.

Elkins, T., Zinn, K., McAllister, L. Hoffmann, F. M. & Goodman, C. S. (1990) Genetics analysis of a *Drosophila* neural cell adhesion molecule: interaction of fasciclin I and Abelson tyrosine kinase mutations. *Cell* **60**: 565-575.

Feinberg, C. P. & Vogelstein, B. (1983) A technique for radiolabling DNA restriction endoclease fragments to high specific activity. *Anal. Biochem.* **132(1)**: 6-13.

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

Gergen, J. P. & Butler, B. A. (1988) Isolation of the *Drosophila* segmentation gene *runt* and analysis of its expression during embryogenesis. Gene *& Development* **2(9)**: 1179-1193.

Goodman, C. S., Bastiani, M. J., Doe, C. Q., du Lac, S., Helfand, S. L., Kuwada, J. Y., & Thomas, J. B. (1984) Cell recognition during neuronal development. *Science* **225**: 1271-1279.

Goodman, C. S. & Shatz, C. J. (1993) Developmental mechanisms that generate precise patterns of neuronal connectivity. *Cell* **72** / *Neuron* **10** (*Suppl.*): 77-98.

Grenningloh, G., Bieber, A. J., Rehm, E. J., Snow, P. M., Traquina, Z. R., Hortsch, M., Patel, N. H. & Goodman, C. S. (1990) Molecular genetics of neuronal recognition in *Drosophila*: evolution and function of immunoglobulin superfamily cell adhesion molecules. *Cold Spring Harbor Symp. Quant. Biol.* **55**: 327-340.

Grenningloh, G., Rehm, E. J. & Goodman, C. S. (1991). Genetic analysis of growth cone guidance in *Drosophila*: Fasciclin II functions as a neuronal recognition molecule. *Cell* **67**: 45-57.

Halpern, M. E., Chiba, A., Johansen, J. & Keshishian, H. (1991). Growth cone behavior Underlying the development of stereotypic Synaptic Connections in Drosophila embryos. *J. Neurosci.* **11(10):** 3227-3238.

Hodgson, C. P. & Fisk, R. Z. (1987) Hybridisation probe size control: optimized oligolabelling. *Nucleic Acids Res.* **15:** 6295.

Holmes D. S. and Quigley, M. (1981) A rapid boiling methods for the preparation of bacterial plasmids. *Anal. Biochem.* **114:** 193-197.

Homyk, T., Szidonya, J. and Suzuki, D. T. (1980) Behavioural mutants of *Drosophila melanogaster* III. Isolation and mapping of mutations by direct visual observation of behavioural phenotype. *Mol. Gen. Genetics* **177:** 553-565.

Hynes, R. O. & Lander, A. D. (1992) contact and adhesive specificities in the associations, migrations, and targeting of cells and axons. *Cell* **68:** 303-322.

Ikeda, K., Koenig, J. H. & Tsuruhara, T. (1980) Organization of identified axons innervating the dorsal longitudinal flight muscle of *Drosophila melanogaster*. *J. Neurocytology* **9:** 799-823.

Ikemura, T. (1985) Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* **2:** 13-34.

Jacobs, J. R. & Goodman, C. S. (1989a) . Embryonic development of axon pathways in the *Drosophila* CNS. I. A glial scaffold appears before the first growth cones. *J. Neurosci.* **9(7):** 2402-2411.

Jacobs, J. R. & Goodman, C. S. (1989b). Embryonic development of axon pathways in the *Drosophila* CNS. II. Behavior of pioneer growth cones. *J. Neurosci.* **9(7):** 2412-2422.

Jessell, T. M. (1988) Adhesion molecules and the hierarchy of neural development. *Neuron* **4:** 3-13.

Johansen, J., Halpern, M. E., Johansen, K. M. & Keshishian, H. (1989a) Stereotypic morphology of glutamatergic synapses on identified muscle cells in *Drosophila* larvae. *J. Neurosci.* **9:** 710-725.

Johansen, J., Halpern, M. E. & Keshishian, H. (1989b) Axonal guidance and the development of muscle fibre-specific innervation in *Drosophila* embryos. *J. Neurosci.* **9:** 4318-4332.

Keshishian, H., Chang, T. N. & Jarecki, J. (1994) Precision and plasticity during *Drosophila* neuromuscular development. *FASEB Journal* **8(10):** 731-737.

King, D. G. & Valentino, K. L. (1983)   On neuronal homology: a comparison of similar axons in *Musca, Sarcophaga,* and *Drosophila* (Diptera: Schizophora) *J. Comp. Neur.* **219:** 1-9.

King, D. G. & Wyman, R. J. (1980) Anatomy of the giant fibre pathway in Drosophila. I. Three thoracic components of the pathway *J. Neurocytology* **9:** 753-770.

Kolodkin, A. L., Matthes, D. J., O'Connor, T. P., Patel, N. H., Admon, A., Bentley, B. & Goodman, C. S. (1992) Fasciclin IV: sequence, Expression, and function during growth cone guidance in the grasshopper embryo. *Neuron* **9**: 831-845.

Koto, M., Tanouye, M. A., Ferrus, A., Thomas, J. B. & Wyman, R. J. (1981) The morphology of the cervical giant fabre neuron of *Drosophila. Brain Research* **221:** 213-217.

Kovalic, D., Kwak, J-H. & Weisblum, B. (1991) General method for direct cloning of DNA fragments generated by the polymerase chain reaction. *Nucleic Acids Research* **19(16):** pp4560.

Krishnan, S. N., Frei, E., Swain, G. P. & Wyman, R. J. (1993) *Passover*: a gene required for synaptic connectivity in the giant fibre system of *Drosophila. Cell* **73**: 967-977.

Lander, A. D.   (1989) Understanding the molecules of neural cell contacts: emerging patterns of structure and function. *TINS*   **12(5):** 189-195.

Lee, K. L., Freeman, M. & Stellar, H. (1991) Expression of the *disconnected* gene during development of *Drosophila EMBO J.* **10:** 817-826.

Lindsley, D. L. & Zimm, G. G. (1992) The genome of *Drosophila melanogaster. Academic Press,* San Diogo, CA.

Lo, Y. J., & Poo, M. M. (1991) Activity-dependent synaptic competition in vitro: heterosynaptic suppression of developing synapses. *Science* **254:** 1019-1022.

Lo. Y. J., Lin, Y. C., Senas, D. M. & Poo, M. M. (1994) Depression of developing neuromuscular synapses induced by repetitive postsynaptic depolarisation. *J. Neurosci.* **14(8):** 4694-4704.

Miklos, G. L. G., Healy, M. J. , Pain, P. , Howells, A. J. & Russell, R. J. (1984) Molecular and genetic studies on the euchromatin-heterochromatin transition region of the X- chromosome of *Drosophila. Chromosoma* **89:** 218-227.

Miklos, G. L. G., Kramers, P. G. N. & Schalet, A. P. (1986) The proximal distal orientation of two lethal complementation groups A112 and LB20 in region 19F at the base of the X- chromosome. *Drosophila Inform Serv.* **63:** 96-97.

Miklos, G. L. G. , Kelley, L. E., Coombe, P. E., lLeeds, C. & Lefevre, G. (1987) Localization of the genes *shaking-B, small optic lobes, sluggish-A, stoned and stress-sensitive-C* to a well defined region on the X-chromosome of *Drosophila. J. Neurogenet.* **4:** 1-19.

Miklos, G. L. G. & Cotsell, J. N. (1990) Chromosome structure at interfaces between major chromatin types: *alpha-* and *beta-*heterochromatin. *BioEssays* **12(1):** 1-6.

Miklos, G. L. G., Yamamoto, M-T, Davies, J. A. & Pirotta, V. (1988) Microcloning reveals a high frequency of repetitive sequences characteristic of chromosome 4 and the B-heterochromatin of *Drosophila. Proc. Natl. Acad. Sci.* USA **85:** 2051-2055.

O'Connor, T. P., Duerr, J. S. & Bentley, D. (1990) Pioneer growth cone steering decisions mediated by single filopodial contacts *in situ .J. Neurosci.* **10(12):** 3935-3946.

Pardue, M. L., Lowenhaupt, K., Rich, A. & Nordheim, A. (1987) (dC-dA)$_n$·(dG-dT)$_n$ sequences have evolutionarily conserved chromosomal locations in *Drosophila* with implications for roles in chromosome structure and function *EMBO J.* **6(6):** 1781-1789.

Patel, N. H. & Goodman, C. S. (1990) Molecular genetics of neuronal recognition in *Drosophila*: evolution and function of immunoglobulin superfamily cell adhesion molecules. *Cold Spring Harbor Symp. Quant. Biol.* **55:** 327-340.

Patel, N. H., Snow, P. M. & Goodman, C. S. (1987) Characterisation and cloning of fasciclin III: A glycoprotein expressed on a subset of neurons and axon pathways in Drosophila. *Cell* **48:** 975-988.

Perrimon, N. , Smouse, D. & Miklos G. L. G. (1989) Developmental genetics of loci at the base of the X- chromosome of *Drosophila Genetics* **121:** 313-331.

Phelan, P., Nakagawa, M., Wilkin, M. B., Moffat, K. G., O'Kane, C. J., Davies, J. A. & Bacon, J. P. (1995) Mutations in *shaking-B* prevent electrical synapse formation in the *Drosophila* giant fibre system (submitted)

Phillis, R. W., Bramlage, A. T., Wotus, C., Whittaker, A., Grametes, L. S., Seppala, D., Farahanchi, F., Caruccio, P. & Murphey, R. K. (1993) Isolation of mutants affecting neural circuitry required for grooming behaviour in *Drosophila melanogaster. Genetics* **133**: 581-592.

Pirrotta, V. (1986) *Drosophila*: a practical approach. (Ed. D. B. Roberts) IRL press.

Pirrotta, V., Hadfield, C. & Pretorius, G. H. (1983) Microdissection and cloning of the *white* locus and the 3B1-3C1 region of the *Drosophila* X chromosome.

Purves, D. & Lichtman, J. W. (1980) Elimination of synapses in the developing nervous system. *Science* **210**: 153-157.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular cloning: a laboratory manual. *Cold Spring Harbor Laboratory Press*, Cold Spring Harbor.

Sanger, F., Nicklen, S. & Coulson, A. P. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.

Schalet, A. and Lefevre, G. (1976) The proximal region of the X chromosome The genetics and biology of *Drosophila*. Vol. 1b Academic Press, Now York 846-902.

Seeger, M., Tear, G., Ferres-Marco, D & Goodmen, C. S. (1993) Mutations affecting growth cone guidance in Drosophila: genes necessary for guidance toward or away from the midline. *Neuron* **10**: 409-426.

Sink, H. & Whitington, P. M. (1991) Pathfinding in the central nervous system and periphery by identified embryonic *Drosophila* motor axons. *Development* **112**: 307-316.

Snow, P. M. , Bieber, A. J. & Goodman C. S. (1989) Fasciclin III: A novel homophilic adhesion molecule in *Drosophila. Cell* **59**: 313-323.

Starich, T. A., Herman, R. K. & Shaw, J. E. (1993) Molecular and genetic analysis of *Unc-7*, a *C. elegans* gene required for coordinated locomotion. *Genetics* **133**: 527-541.

Strausfeld, N. J. & Bassmir, U. K. (1983) Cobalt-coupled neurons of a giant fibre system in diptera. *J. of Neurocytology* **12**: 971-991.

Strausfeld, N. J., Bassmir, U. K., Singh, R. N. & Bacon, J. (1984) Organizational principles of outputs from dipteran brains. *J. Insect Physiol.* **30(1)**: 73-93.

Tanouye, M. A. & Wyman, R. J. (1980) Motor outputs of giant nerve fibre in *Drosophila. J. Neurophysiol.* **44(2)** 405-421.

Thomas, J. B., Bastiani, M. J., Bate, M. & Goodman, C. S. (1984) From grasshopper to *Drosophila* : a common plan for neuronal development. *Nature* **310**: 203-207.

Thompson, W. J. (1986) Changes in the innervation of mammalian skeletal muscle fibres during postnatal development. *Trens Neurosci.* **9**: 25-28

Tian, S-S., Tsoulfas, P. & Zinn, K. (1991) Three receptor-linked protein-tyrosine phosphatases are selectively expressed on central nervous system axons in the *Drosophila* embryo. *Cell* **67**: 675-685.

Vactor, D. V., Sink, H., Fambrough, D., Tsoo, R. & Goodman, C. S. (1993) Genes that control neuromuscular specificity in *Drosophila. Cell* **73**: 1137-1153.

Wyman, R. J. & Thomas, J. B. (1983) What genes are necessary to make an identical synapse? *Cold Spring Harb. Symp. Quant. Biol.* **18**: 611-652.

Watanabe, T. 7 Kankel, D. R. (1990) Molecular cloning and analysis of *l(1) ogre*, a locus of *Drosophila* with prominent effects on the postembryonic development of the central nervous system. *Genetics* **126**: 1033-1044.

Yamamoto, M-T., Mitchelson, A., Tudor, M., Hare, K. O., Davies, J. & Miklos, G. L. G. (1990) Molecular and Cytogenetic analysis of the heterochromatin-euchromatin junction region of the *Drosophila melanogaster* X chromosome using cloned cDNA sequences. *Genetics* **125**: 821-832.