Studies of new neural bHLH genes in *Drosophila melanogaster*

Sarah Goulding

A thesis presented for the degree of PhD

University of Edinburgh

1999

Declaration

I declare that this thesis is my own work, unless otherwise stated in the text.

Acknowledgements

First and foremost I would like to thank The Darwin Trust and Dr. Andrew Jarman for providing me with the means and opportunity to come and study here in Edinburgh. Andy has been a five star supervisor over the last few years, and has given me constant support, guidance and more than a glimpse of his infinite wisdom. The lab members Ruth, Petra, Neil, Davey and Emma have made it an unforgettable journey. Thanks are also due to past members If and Julie for technical help. Special thanks go to Andy, Petra, Neil and Ruth for their help and instruction in the lab.

Many friends, too numerous to mention by name, have helped (and hindered!) the production of this thesis. Ruth has been as much fun outside the lab, as in it. I salute the Collective (Jarman) and all the Davis group members – Secret Santa rules! Rachel has provided invaluable Glenogle coffee, tea and courage. Thanks to Damion for weekend entertainment, Nicola for long distance support and Susie for strength and support. Sander has given me much undeserved TLC and distraction, and embellished my vocabulary to boot! Mum and Becs have been instrumental in providing love, support and unfailing belief.

I would also like to acknowledge Yoda, for such inspiring words as 'May the force be with you'.

Finally, I acknowledge myself. So, a toast, Here's to me! Those of you who know me will hopefully realise that this is not conceit, but actually a warped way of stating that I can now recognise when I have achieved something; in my book, a step forward.

Abstract

In the *Drosophila* peripheral nervous system (PNS), neuronal cell fates are specified by a hierarchy of events that are ultimately regulated at the level of transcription. The basic helix-loop helix (bHLH) protein family of transcription regulators are widely implicated throughout development.

The first bHLH proteins shown to be involved in neurogenesis were those encoded by the *Drosophila achaete-scute* complex (AS-C). These are known as proneural genes because they are required during the earliest step of neurogenesis in which naive ectodermal cells are selected to become neural precursors. The AS-C govern the determination of precursors for the external sensory bristles. Another proneural gene, *ato*, is required for the precursors of the chordotonal organs, photoreceptors and some olfactory sensilla. However, these genes alone do not account for the formation of the rest of the nervous system, thus there must be other proneural candidate genes to be discovered. Furthermore, there is much evidence to suggest that other bHLH factors are required throughout neurogenesis, acting in interlinked cascades at each successive level of cell differentiation. These too, have yet to be identified.

In the expectation that more *Drosophila* neural bHLH genes exist, I searched for new genes on the basis of homology to *ato*. Indeed in this way I identified two new *ato*-like bHLH proteins. The molecular characterisation, expression and functional analyses of these genes are presented in this thesis. *amos* (absent MD neurons and olfactory sensilla) is a new proneural gene, required for the selection of neural precursors of larval md neurons and adult olfactory sensilla. On the other hand, *cato* (cousin of *ato*), ensures the proper differentiation of the sensory neurons.

Abbreviations

A	- adenosine
amp	- ampicillin
AP	- alkaline phosphatase
ATP	- adenosine 5'-triphosphatase
ß gal	- ß-galactosidase
bHLH	- basic helix-loop-helix
bp	- base pair
BSA	- bovine serum albumin
C	- cytidine
°C	- degrees celcius
cDNA	- complementary DNA
CNS	- central nervous system
d	- deoxy
DAB	- 3,3'-diaminobenzidine
dd	- dideoxy
DEPC	- diethyl pyrocarbonate
DIG	- digoxygenin
DMSO	- dimethyl sulphoxide
DNA	- deoxyribonucleic acid
DNAse	- deoxyribonuclease
dNTP	- deoxyribonucleoside triphosphate
ds	- double stranded
EDTA	- diaminoethanetetra-acetic acid
g	- gram

G	- guanosine
h	- hour
HEPES	- N-[2-Hydroxyethyl]piperazine-N'-[2-ethane-suplonic acid]
HRP	- horse radish peroxidase
IPTG	- isopropyl-b-D-thiogalactopyranoside
kb(s)	- kilobase(s)
λ	- Lambda
l	- litre
L-broth	- Luria broth
M	- molar
Mb	- megabase(s)
mg	- milligram
min	- minute
ml	- millilitre
mol	- mole
MOPS	- 3-[N-Morpholino]propanesulphonic acid
mRNA	- messenger RNA
MW	- molecular weight
ng	- nanogram
nm	- nanometre
O.D.	- optical density
OrR	- OregonR
p	- plasmid
p	- pico
PBS	- phosphate buffered saline
PBTW	- phosphate buffered saline with Tween20
PBTX	- phosphate buffered saline with Triton X-100

PCR	- polymerase chain reaction
pers. comm.	- personal communication
PEG	- polyethylene glycol
pH	- $\log_{10}$ (hydrogen ion concentration)
PIPES	- piperazine-N,N'-bis[2-ethane-sulphonic acid]; 1,4-piperazine
PNC	- proneural cluster
PNS	- peripheral nervous system
POD	- peroxidase
Rnase	- ribonuclease
rpm	- revolutions per minute
sec	- second
SDS	- sodium dodecyl sulphate
SOC	- sense organ cell
SOP	- sense organ precursor
ssDNA	- single-stranded deoxyribonucleic acid
T	- thymidine
TEMED	- N,N,N'-tetramethylethylenediamine
Tm	- melting temperature
Tris	- tris (hydroxymethyl) aminomethane
U	- uridine
UAS	- upstream activation sequence
μg	- microgram
μl	- microlitre
UV	- ultraviolet
V	- volt
v	- volume
v/v	- volume per volume

w/v	- weight per volume
wt	- wildtpye
X-gal	- 5-bromo-4-chloro-indol-b-D-galatopyranoside
X-phosphate	-5-bromo-4-chloro-indolyl-phosphate, 4-toluidine salt

Table of contents
Title page
Declaration
Abstract
Abbreviations
Table of contents
Chapter 1 Introduction
1.1 Summary 1
1.2 The *Drosophila* PNS as a model system 2
1.3 Development of the PNS 6
1.4 The genetics of sensory development 11
1.5 The basic helix-loop-helix (bHLH) protein family 24
1.6 Aims of this thesis 40
Chapter 2 Isolation and characterisation of new *ato*-related genes
2.1 Introduction 42
2.2 Degenerate PCR screen for *ato* homologues 43
2.3 Isolation and analysis of genomic sequences flanking the *amos* and *cato* bHLH regions 47
2.4 Molecular investigations of *amos* and *cato* expression profiles 49
2.5 Sequence analysis 54
2.6 Protein structure analysis 59
2.7 Mapping of *amos* and *cato* 62
2.8 Discussion 66
Chapter 3 Expression and Functional analysis of *cato*
3.1 Introduction 68
3.2 *cato* mRNA expression 70
3.3 *cato* is involved during sense organ differentiation 76
3.4 Additional deletion of *ase* worsens the *cato* neural defects 80
3.5 Misexpression analysis of *cato* 82
3.6 Making flies transgenic for UAS-*cato* 83
3.7 Mapping and testing of UAS-*cato* lines 84
3.8 Misexpression of UAS-*cato* with $Gal4^{109\text{-}68}$ in the adult 85
3.9 Misexpression of UAS constructs using other Gal4 drivers 91
3.10 An effect of *cato* misexpression is the activation of *ato* 94
3.11 Misexpression in the embryo with *hairyGal4* 96
3.12 *cato* is actively repressed in the CNS 97

Chapter 4 Expression and Functional analysis of *amos*

4.1 Introduction 101

4.2 *amos* mRNA expression in the embryo 102

4.3 Loss of *amos* function results in the loss of a subset of larval sensilla 105

4.4 In imaginal discs *amos* is expressed in the proneural domains of the antennal olfactory sensilla and leg tarsal claw 107

4.5 Summary of experiments performed to analyse the role of *amos* during adult neurogenesis 109

4.6 Discussion 112

Chapter 5 General Discussion and Conclusions

5.1 The potential molecular bases of functional specificity 115

5.2 Possible experimental approaches to explore the bases of bHLH functional specificity 118

5.3 Potential screening strategies for specific mutations 118

Chapter 6 Materials and Methods **121**

Appendices **145**

References **150**

Chapter 1

Introduction

1.1 Summary

This introduction begins with a description of the unique features of the *Drosophila* peripheral nervous system (PNS) that have led to its emergence as a powerful model system for studying the mechanisms of neurogenesis, and development in general. An understanding of the PNS touches on general processes such as cell fate determination, pattern formation, transcriptional activation and repression, equivalence group formation, and lateral inhibition.

In particular, the determination of cell fate is a central question to developmental biology. It seems that many features of this process have been conserved in evolution, allowing the extrapolation of insect studies onto the broader context of vertebrate neural diversity (Arendt and Nübler-Jung, 1999, Chan and Jan, 1999).

In the PNS, neuronal cell fates are specified by a hierarchy of events that are ultimately regulated at the level of transcription. Key players in this process are the basic helix-loop-helix (bHLH) family of transcriptional regulators, proteins which are, in fact, widely implicated throughout development. The structure and various functions of these proteins are discussed, with particular attention to their roles in neurogenesis.

1.2 The *Drosophila* PNS as a model system

Several features of the *Drosophila* PNS make this a useful model to study the mechanisms behind neurogenesis. Insects are provided with a variety of sense organs located in stereotyped pattterns over and in the body, and their presence, location and number can be easily ascertained (reviewed in Jan and Jan, 1993, 1994, Campos-Ortega and Hartenstein, 1998). Furthermore though morphologically and functionally diverse, these organs are relatively simple in structure, being composed of limited cell types. Most sense organs comprise one or more neurons, and several specialised support cells (Jan and Jan, 1993; McIver, 1985; Zacharuk, 1985).

The embryonic sensory nervous system in particular has proved an excellent model for genetic analysis. The anatomy and origins of the larval sense organs has been well documented (Campos-Ortega and Hartenstein, 1998) and each of the sense organ cells can be identified and their development monitored using a wide range of cell specific markers and enhancer trap lines (listed in Jan and Jan, 1993). Thus mutations in genes that affect sense organ formation, number, identity, location, type and axonal pathways can be readily analysed. The advent of sophisticated mosaic analysis and powerful microscopy techniques mean that similar studies of gene function in the adult are now much easier. The adult compound eye is proving an excellent model system for studying the genetics of development. Similarly, the mechanosenseory lineage is a favourite model for studying cell fate processes. In many cases, the cells of the sense organ are clonally related, a common precursor dividing by a strict cell lineage to give rise to the different cells of the organ (Brewster and Bodmer, 1995, Gho et al., 1999) providing an ideal model to investigate the basis of cell fate decisions.

1.2.1 The elements of the PNS

The *Drosophila* PNS comprises a stereotyped pattern of sensory neurons in and under the cuticle. In the larva and adult, the size and form of sensory organs varies widely however, their individual structures and architecture can be accounted for by a small number of basic designs. There are four major classes of sensory element: external sense organs (such as the prominent mechanosensory bristles) (McIver, 1985), chorodotonal organs (stretch and vibration receptors) (Moulins, 1976, McIver, 1985, Sugawara, 1996), multiple dendritic neurons and photoreceptors (Ready,

1989). Modifications of basic structures give rise to functional derivatives, such as the olfactory sensilla (Zacharuk, 1985, Riesgo-Escovar et al., 1997). Arrays of individual units form more complex sensory structures, such as the compound eye (Wolff and Ready, 1993) and chordotonal Johnston's organ (McIver, 1985).

External Sensory Organs

Drosophila are covered with sense organs that detect mechanical or chemical signals from the environment via external sensory structures such as bristles (hairs). Typically, each organ is comprised of four or more cellular components. The trichogen cell forms the sensory structure (hair) and the tormogen cell forms the supportive socket cell (Campos-Ortega and Hartenstein, 1998). Bristles have many morphologies, open tipped to sense chemical stimuli, or pointed to act as mechanosensory organs, or even dome shaped campaniform sensilla, thought to sense cuticular strain (McIver, 1985). Internally, one or more bipolar neuron cells with single dendrites (type I) innervate the external structures, their axons projecting into the CNS, and the thecogen support, or glial, cell ensheathes part of the neuron(s) (Fig. 1.2.1).

Chordotonal Organs

Chordotonal sensilla are completely internal structures thought to function as proprioceptors and, in some insects, as hearing organs (van Staaden and Römer, 1998, Eberl, 1999). These organs are found in the body wall, or joints, as single units (scolopidium), or as arrays of aligned scolopidia. The organ is an elegant elongated tube structure comprising four main cellular components, a single neuron and three support cells namely, the ligament, sheath and cap cells (Fig. 1.2.1). Chordotonal neurons have highly specialised dendritic strucutures that have been well characterised by electron micrograph (EM) studies (Moulins, 1976 Merritt, 1997). The dendrite is closely associated with a modified structure of the sheath cell, a highly refractile structure termed the scolopale. The cap cell attaches to the body wall whereas the ligament cell joins the structure to muscles. Similar to the external sensory organs, the chordotonal neurons are type 1 monodendritic.

Multiple dendritic neurons

As yet only described for the larva, the multiple dendritic (md) neurons are also internal sensory organs of unknown function (Campos-Ortega and Hartenstein, 1998)

(Fig. 1.2.1). They may be touch receptors or proprioceptors (Bodmer and Jan, 1987). Classed as type II on the basis of their multiple dendritic structure, they are further subdivided into three groups: those with arborisations underneath the epidermis (da), those innervating the tracheal branches (td) and those with two opposing dendrites (bd) (Bodmer and Jan, 1987). md neurons have extensive dendritic arborisations that form from 16h to beyond hatching (Gao et al., 1999) and occur as clusters attached to the basal surface of the epidermis or internal organs, such as tracheae, peripheral nerves, and muscles.

Photoreceptors

In adult flies, the compound eye is composed of hundreds of individual units, the ommatidia (reviewed in Wolff and Ready, 1993) which are formed by a recruitment process. Each ommatidium is a precise 19-cell assembly of 8 photoreceptors (R1-R8) and 11 accessory cells including pigment and cone cells (Fig. 1.2.1).

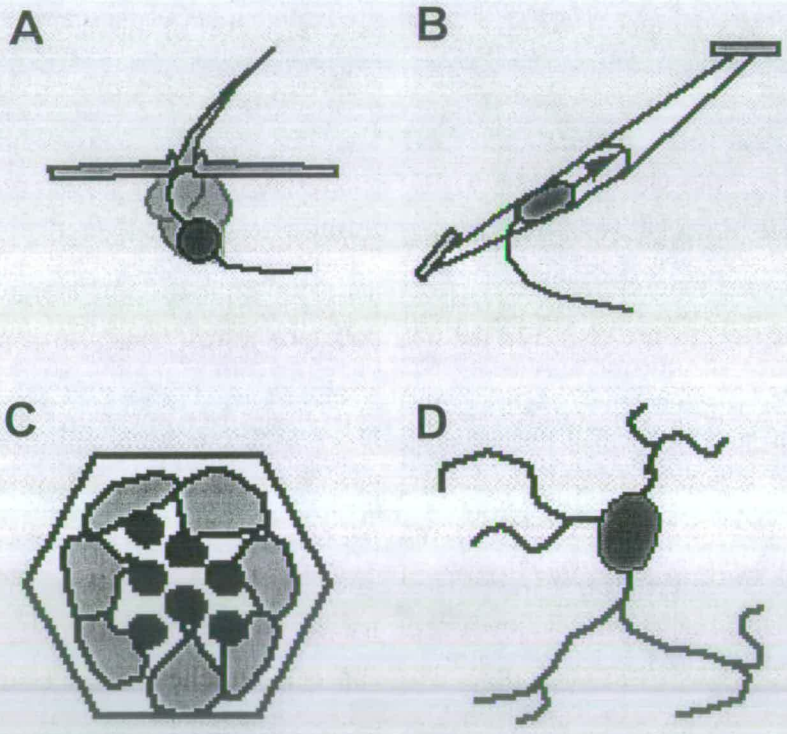


Figure 1.2.1 Elements of the PNS

Schematic diagram of the sense organs. The PNS comprises external sense organs (A), chordotonal organs (B), ommatidia (C) and multiple dendritic neurons (D)

1.2.2 Sense organ distribution and arrangement

The number and position of larval sense organs is essentially invariant (Ghysen et al., 1986; Bodmer and Jan, 1987, Hartenstein, 1988, Campos-Ortega and Hartenstein, 1998). Each body segment contains a fixed number of sensory neurons in a reproducible pattern and the identity of each individual organ has been investigated (Fig. 1.2.2) (Campos-Ortega and Hartenstein, 1998 and refs. therein). The adult sensory system comprises a massive array of different sensory organs, distributed in a recognisable pattern over the body. The arrangement and origins have not been as comprehensively characterised as for the larva, nevertheless, many of the sense organs are readily identifiable under the light miscroscope (Fig.1.2.2).

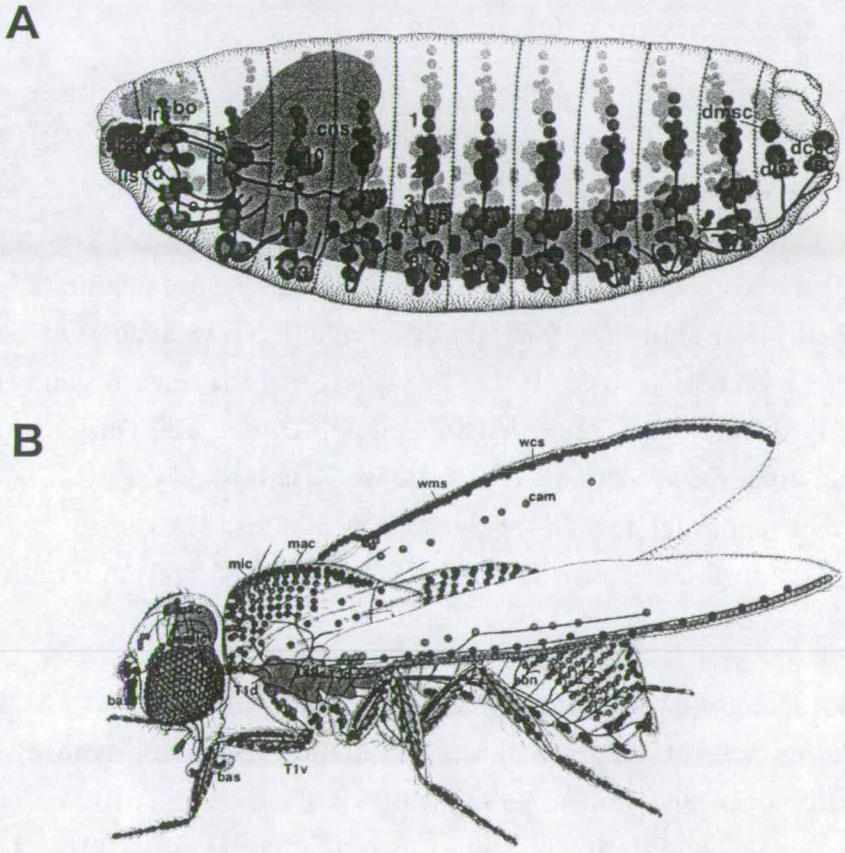


Figure 1.2.2 Sense organ distribution in *Drosophila*

Schematic digram showing sense organ distribution in the larva (A) and adult fly (B). Each circle represents a sense organ. Taken from the Atlas of *Drosophila* development, 1993.

1.3 Development of the PNS

The formation of the *Drosophila* PNS comprises several steps: first, the neurogenic ectoderm is specified followed by the formation of proneural clusters, and the determination of a single sense organ precursor (SOP) from each cluster. Then, additional cells are generated either by stereotyped cell divisions of the SOP or recruitment of nearby cells by signalling from the precursor. The final stage is the differentiation and maturation of individual sense organ cells, referred to here as SOCs. Superimposed on these steps are the patterning events that specify the locations of the proneural clusters (and thus the sensory structures) as well as contributing to their neuronal subtype and circuitry.

1.3.1 The genesis of sensory organs

Larval sense organs are formed during embryogenesis. Moreover, each sense organ is derived from precursors that do not migrate. Thus the organisation of the PNS can be traced from very early embryogenesis, with the aid of specific cell markers, and enhancer trap lines that mark precursors and their derivatives (e.g. Ghysen and O'Kane, 1989). For instance, the enhancer trap line A37 marks neural precursors in the embryo, from their birth, until their progeny differentiate (Ghysen and O'Kane, 1989). The first SOPs arise during stage 10, and by stage 13 all the sense organ cells are born and ready to differentiate (reviewed in Campos-Ortega and Hartenstein, 1998). By cuticle secretion, the sensory neurons have well established the process of axonogenesis, and their dendrites invade the sensory structures. Thus in the short space of less than 24 hours, the entire larval PNS is specified and undergoing differentiation.

In the adult, however, sense organ formation proceeds less synchronously. The SOPs are selected from the monolayer epithelial sheets of the imaginal discs, themselves laid down during embryogenesis (Cohen, 1993). How SOPs arise and differentiate has been most extensively studied in the wing discs (Huang et al., 1991). The different external sense organ SOPs appear in a strict sequence over a period of more than 35 hrs, during most of the third larval instar and the first 10 hours of puparium formation (Hartenstein and Posakony, 1989, Huang et al., 1991). In the wing the earliest formed SOPs are delayed in their mitoses for up to 30 hours whilst their late forming counterparts have a 2hr delay (Huang et al., 1991).

Once the sense organ cells have been specified, they maintain a fixed spatial relationship throughout development. The morphogenesis of larval sensory organs at the ultrastructural level was recently reviewed by Campos-Ortega and Hartenstein (1998). In the embryo, SOCs are spatially restricted throughout development and little migration takes place. External sense organs have external structures lying within the epidermal layer whilst those of chordotonal organs form a linear asubepidermal array of cells. Intricate cell-cell communications are maintained among the cells of the sense organs (Hartenstein and Posakony, 1989). Adult sense organs develop similarly (Hartenstein and Posakony, 1989), except for the timing of terminal differentiation which can be far more protracted in the adult (Huang et al., 1991).

All sensory neurons initiate axonal outgrowth shortly after their terminal division, and the course and distribution of both sensory and motor fibres can be distinguished reasonably well using neural markers in fixed wholemount embryos (reviewed in Goodman and Doe, 1993, Campos-Ortega and Hartenstein, 1998). Recently developed techniques utilising the green fluorescent protein (GFP) are now being applied to examine the dynamics of axon pathfinding *in vivo* (e.g Murray et al., 1998).

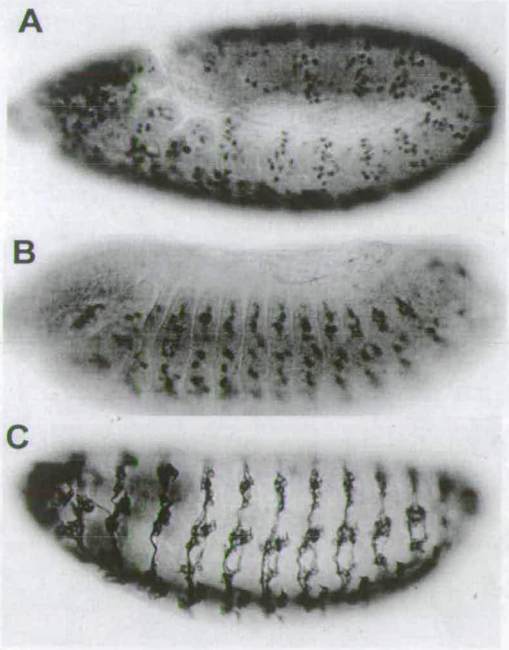


Figure 1.3.1 Sense organ development in the larva.

(A,B) Embryos stained for markers to detect elements of the PNS. The SOPs form during stage 11 (A). By stage 12 the arrangement of the SOCs (B) mirrors their final pattern (C). (C) Stage 17 embryo stained with neural marker Mab22C10, showing neuronal components.

1.3.2 Developmental strategies for sense organ formation

Two main strategies have been described for the development of the sense organs. Individual external sense organs and chordotonal organs utilise a mechanism that is largely lineage restricted (reviewed in Ghysen et al., 1993, Jan and Jan, 1993, Fig. 1.3.2) in which an SOP undergoes a series of stereotyped divisions to give rise to all the cells within the sensillum. This strategy is ideally suited for the formation of spatially isolated sense organs. However, it is now evident that different rules apply in domains where sensilla are densely packed, (such as the wing margin and antenna of the adult), or form arrays of individual units (for instance, photoreceptors of the compound eye and scolopidial array of the leg).

(i) Cell lineage mechanisms for sense organ formation

Several studies have demonstrated that, in the larva, the various cells that make up a single sense organ are related by cell lineage (Hartenstein, 1988, Bodmer et al., 1989, Brewster and Bodmer, 1995, Vervoort et al., 1997b). By inducing the formation of lacZ-expressing mitotic clones during early embryogenesis, Brewster and Bodmer (1995) followed the fate of individual SOPs. In this way, they found that there were a fixed number of possible cell lineages to account for the formation of all the different classes of sense organ in the embryo. In the adult, the best described sense organ formation, where the cells are all derived from a single cell lineage is that of the mechanosensory bristle (Hartenstein and Posakony, 1989, Manning and Doe, 1999, Gho and Schweisguth, 1998, Gho et al., 1999).

In the simplest case of a singly innervated external sensillum, the SOP divides to give two daughter cells which then each divide again, one giving rise to the external structural cells, the hair and socket cells, the other the glial and neuron cells Hartenstein and Posakony, 1989). The sense organ cell (SOC) which gives rise to the neuron is thought to divide further, if a multiply innervated sense organ is to form (Bodmer et al., 1989). In the case of notal bristles, it was recently demonstrated that a fifth (soma sheath) cell associated with the sense organ, appears to derive from a further division of one of the primary divisions products (pIIb) (Gho et al., 1999).

Although the cell division patterns are slightly different, the cells comprising individual chordotonal organs are similarly related by lineage (Bodmer et al., 1989, Brewster and Bodmer, 1995). These studies further demonstrated that a large subset

of md neurons are clonally derived from external sense organ and chordotonal lineages (Brewster and Bodmer, 1995, Vervoort et al., 1997b).

(ii) Cell-cell interactions and recruitment mechanisms for precursor determination

Although morphologically distinct, chordotonal organs and photoreceptors are united by the fact that they are composed of arrays of repeated units. It has now been established that similar mechanisms are employed by founder precursors to recruit further cells in order to establish the correct number of units to form an organ.

At one extreme, the photoreceptors (R1-R8) are not related to each other by cell lineage, but depend on cellular interactions for their development (reviewed by Dickson and Hafen, 1993). The mechanism involves the specification of founder photoreceptors, similar to the development of external sense organ SOPs, but then, these precursors recruit adjacent cells by cell-cell signalling to become photoreceptors (Fig. 1.3.2). Variations of this two-step process are employed during the development of arrays of chordotonal organs in the embryo (zur Lage et al., 1997) and femur (zur Lage and Jarman, 1999). In these cases however, and by contrast to photoreceptor cells, once all the precursors have been recruited, SOPs are then free to divide by cell lineage, to form the individual components of each sensillum.

A variation on this theme appears to be used during olfactory sensilla formation on the adult antenna (Ray and Rodrigues, 1995, Reddy et al., 1997): founder precursors recruit extra cells to form a small presensillum cluster (PSC) of a few cells (Ray and Rodrigues, 1995). Each cell of the PSC divides once, and then differentiate, or die by programmed cell death, resulting in the final sensory structure (Reddy et al., 1997). Interestingly, cells of individual olfactory sensilla are not clonally derived (Reddy et al., 1997), similar to observations for some sense organs along the wing margin which are similarly clustered and often of mixed lineage (Hartenstein and Posakony, 1989). It may be that the strategy for wing margin sense organ formation is more similar to that proposed for olfactory sensilla.

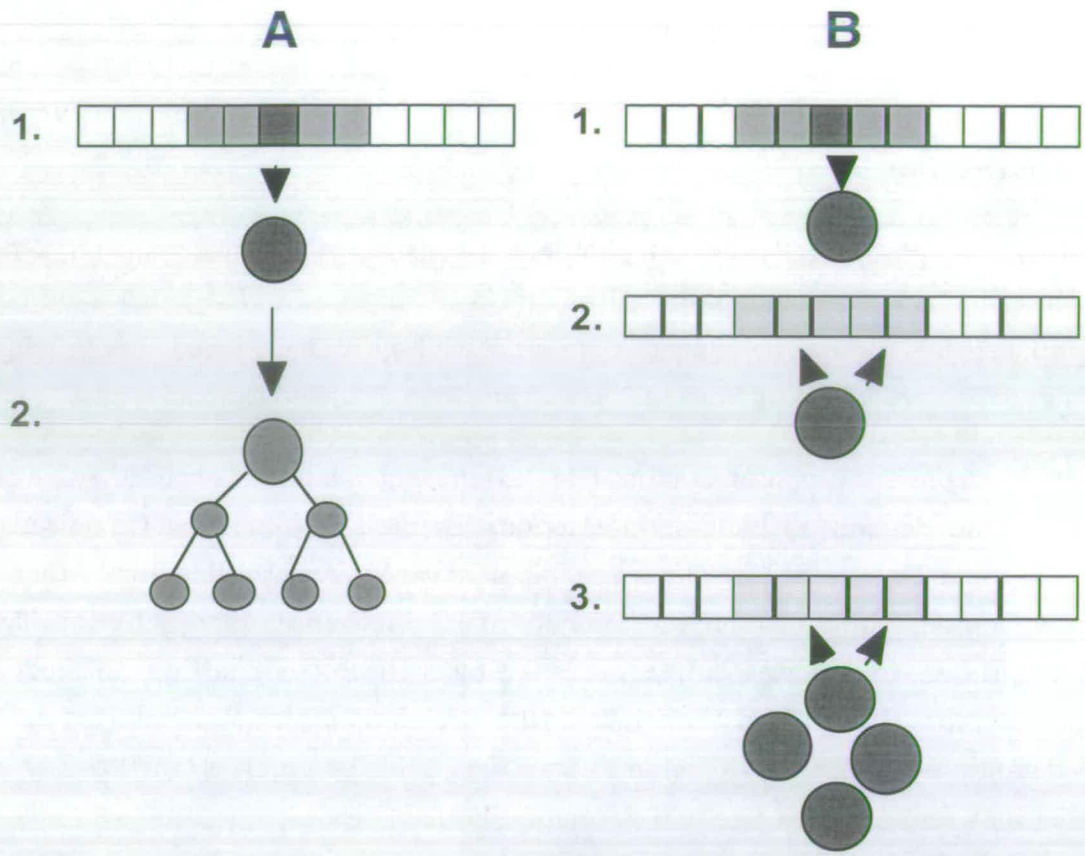


Figure 1.3.2 Developmental strategies for sense organ formation

(A) Sense organ formation by SOP determination (1) followed by cell lineage (2). (B) Sense organ formation by recruitment from founding precursors. (1) SOPs are first determined as in A, emerging from the proneural cluster. (2) This cell then initiates recruitment of other SOPs from the ectoderm. (3) In some cases this process is reiterative, allowing the recruitment of large numbers of SOPs Squares depict ectoderm, gray squares, proneural clusters, dark gray circles, SOPs.

1.4 The genetics of sensory development

1.4.1 The Progressive Determination Model

Regardless of the strategy used to specify individual organs, the early steps of precursor specification, and many aspects of cellular differentiation, appear to be governed by similar genetic mechanisms. Historically, the progressive determination model for the formation of a sensory bristle (Ghysen and Dambly-Chaudière, 1989) is used as a framework to describe the genetics of PNS development.

In essence, the model proposes that the developmental potential of sensory precursor cells becomes progressively restricted as development proceeds. This process is the outcome of a number of distinct steps, each the result of the action of a small number of genes (Ghysen and Dambly-Chaudière, 1989, Jan and Jan, 1993, Jan and Jan, 1994, Fig. 1.4.1). The model is briefly described here, and then information regarding the genes necessary to promote each step is set out in more detail afterwards.

The first step is resolving where and when neural precursors may be specified. In the early embryo a tartan-like expression of pre-pattern genes ensures the unequal developmental potential of cells at different locations in the ectodermal layer (1.4.3). Similarly, during adult development pre-pattern genes have been identified that refine the position and potential of ectodermal cells within the imaginal discs (see 1.4.3). The expression of proneural genes in cell clusters signifies the first step of neural commitment, and the subsequent refinement to one or two precursors is governed by the action of neurogenic genes. The establishment and subsequent development of these neural precursors may involve the actions of regulatory genes such as the pan-neuronal precursor genes which may also attribute general neuronal properties to these cells. The identity, or subtype of individual precursors is influenced by the neuron-type selector genes, as well as earlier information to which the cell has been exposed. The fate of individual cells that make up the sense organ are dictated by a combination of cell intrinsic, and extrinsic determinants that are again at least partially dependent on the genetic history of the precursors.

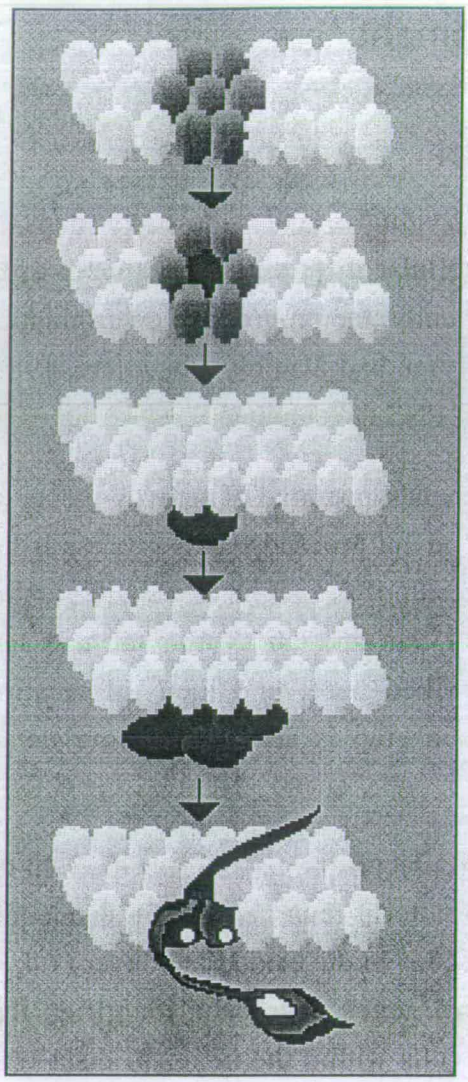


Figure 1.4.1 The Progressive Determination Model

After Ghysen et al., 1989 and Jan and Jan, 1994. This model describes the steps for the formation of an external sense organ. With modifications, the model can be applied to formation of clustered sense organs, such as chordotonal organs and photoreceptors. Shaded circles indicate proneural clusters, and dark circles the SOP and division products.

1.4.2 Proneural genes: defining neural competence

The primary function of the proneural genes is to endow clusters of uncommitted ectodermal cells with neural potential (Ghysen and Dambly-Chaudière, 1989). Their expression thus marks the first asymmetry among cells of the ectoderm, setting apart neural precursors from their epidermal neighbours (Cubas et al., 1991; Romani et al., 1989; Skeath and Carroll, 1991). Given the limited migration of neurons in *Drosophila*, proneural expression foreshadows the distribution of sense organs in both the larva and adult. Proneural gene expression is typified by transient expression in ectodermal cell clusters that becomes restricted to a single neural precursor (Cubas et al., 1991; Skeath and Carroll, 1991).

Furthermore, different proneural genes determine the formation of different subsets of the nervous system. The genes of the AS-C complex and *ato* govern much of the PNS (Jarman and Jan, 1995). In the CNS, the AS-C, *single-minded* (*sim*) and *ventral nervous system condensation defective* (*vnd*) are required for different subsets of neuroblasts (reviewed in Campos-Ortega, 1993). Each of these genes is expressed in a pattern that is for the most part, specific to the subset of structures for which they are required.

Proneural genes for the PNS encode transcription factors of the bHLH family (see 1.5). The individual gene products require a binding partner to effect their proneural functions and this role is supplied by the ubiquitously expressed *daughterless* (*da*,Caudy et al., 1988a; Caudy et al., 1988b) which also encodes a bHLH factor. The universal role of *da* is exemplified by the observation that *da* mutations affect the development of the entire nervous system (Caudy et al., 1988a). *In vitro* experiments and genetic interaction analysis indicate that the bHLH proteins encoded by the AS-C genes function as heterodimers with the ubiquitously distributed bHLH factor Da (Caudy et al., 1988a).

The achaete-scute complex (AS-C)

Initial studies of proneural gene function centred on the AS-C, the phenotpyes of various mutations for these genes hinting heavily at their fundamental role during adult sensory bristle development, (reviewed in Ghysen and Dambly-Chaudière, 1988). The complex comprises four highly related bHLH-encoding genes, namely

achaete (*ac*), *scute* (*sc*), *lethal of scute* (*l'sc*) and *asense* (*ase*) (Alonso and Cabrera, 1988). Three of these genes provide proneural functions during larval and adult neurogenesis (reviewed by Ghysen and Dambly-Chaudière, 1989, Jan and Jan, 1993) and are expressed in a 'classic' proneural pattern. In loss of function mutations for these genes, fewer neural precursors, and hence sense organs, are formed (Dambly-Chaudière and Ghysen, 1987). Genetically, the absense of one gene may be compensated for by the presence of the other, reflecting their structural identity (reviewed in Ghysen and Dambly-Chaudière, 1988). However each gene is required for a specific subset of sense organs, highlighted by their complementary expression patterns and evidenced by genetic analysis and dissection of their differential regulation by cis-control elements and enhancers (Modolell and Campuzano, 1998).

In the adult, *ac* and *sc* are required for the formation of external sensory organs. In the embryo the same two genes govern the formation of the external sense organs and many of the multidendritic neurons of the PNS (Dambly-Chaudière and Ghysen, 1987), as well as several subsets of neuroblast lineages, (Skeath and Doe, 1996). *l'sc* is required for some neuroblast lineages of the embryonic CNS (reviewed in Campos-Ortega, 1993; Skeath and Doe, 1996). In the larva, deletion of the entire complex results in the absense of all external sense organs, many md neurons, and subsets of the CNS.

By contrast, misexpression of any of the genes of the complex results in ectopic sense organ formation (Brand et al., 1993; Dominguez and Campuzano, 1993; Hinz et al., 1994; Rodríguez et al., 1990). These results reinforce the functional redundancy among these structurally similar proteins.

atonal (ato)

ato was identified on the basis of sequence similarity with the AS-C genes (Jarman et al., 1993b) and was shown by expression and mutant analysis to function as a proneural gene (Jarman et al., 1993b; Jarman et al., 1994; Jarman et al., 1995). *ato* expression is restricted mainly to the PNS and does not overlap with that of the AS-C genes. This correlates with its requirement for a complementary subset of sense organs, namely the photoreceptors (Jarman et al., 1994), chordotonal organs (Jarman et al., 1993b, Jarman et al., 1995) and a subset of olfactory sensilla (Gupta and Rodrigues, 1997). *ato* mutant larvae lack chordotonal organs and a subset of multiple dendritic neurons (Jarman et al., 1993b Jarman et al., 1995). *ato* mutant flies are

viable, though lack photoreceptors (Jarman et al., 1994), have fewer olfactory sensilla (Gupta and Rodrigues, 1997), and have no chordotonal organs thus are uncoordinated (Jarman et al., 1995), and deaf (missing the Johnston's organ, Eberl, 1999).

The role of *ato* in the specification of neural precursors for chrodotonal organs is a direct parallel to that of the AS-C genes for external sense organs, both types of sense organ arising by strict cell lineages. Furthermore *ato* is similarly affected by the neurogenic genes and the process of lateral inhibition (Goriely et al., 1991, A.P.J. unpublished observations). At this point however, the functions of these genes diverge (Fig. 1.4.2). *ato* turns out to be instrumental in recruiting extra precursors for the formation of both photoreceptors and scolopidial arrays (Jarman et al., 1995, zur Lage et al., 1997; zur Lage and Jarman, 1999, see below). Furthermore, the AS-C and *ato* encode different neuronal subtype specificities.

The specificity of the proneural genes

A second feature of proneural gene function is to influence the subtype of sense organ formed. This role for proneural genes is shown by misexpression experiments. When any proneural gene is misexpressed in imaginal discs, ectopic SOP determination occurs. AS-C misexpression only induces ectopic external sense organ formation (Rodríguez et al., 1990, Brand et al., 1993, Dominguez and Campuzano, 1993, Hinz et al., 1994). However, *ato* misexpression in these areas gives rise (mostly) to ectopic chordotonal sensilla (Jarman et al., 1993b). The molecular basis for this functional specificity is in the sequence of the transcription factor (bHLH) domains (Chien et al., 1996). Genetically, this specificity was pinpointed as the differential regulation of a single gene, the subtype selector gene *cut* (Jarman and Ahmed, 1998). *ato* inhibits *cut* thereby promoting chordotonal SOP identity whereas the AS-C activate *cut*, initiating the external sensory SOP pathway (Jarman and Ahmed, 1998, Chien et al., 1996). In *cut* mutants, larval external sense organs are completely converted to chordotonal identity (Bodmer et al., 1987, Merritt, 1997) demonstrating the fundamental role of this gene in determining sense organ subtype. Interestingly, *cut* is not expressed in any of the domains where *ato* functions as a proneural gene, suggesting its inhibition is a general requirement for *ato*-dependent SOP development.

Notably, misexpression of *ato* in different domains gives rise to different phenotypes – in domains where *ato* is not normally expressed, misexpression yields chordotonal organ formation. However, in the eye and antenna, *ato* misexpression induces the formation of ectopic photoreceptors and olfactory sensilla, respectively, suggesting that in these regions *ato* function is modulated in some way.

Although any member of the AS-C can substitute for the other during external sensory bristle formation, this is not the case for the other functions of the AS-C genes. For example during sex determination, *ac* cannot substitute for *sc* (*sis-b*) (Parkhurst and Meneely, 1994). Furthermore, studies of the roles of the AS-C during neuroblast formation suggest that there is individual specificity encoded by each protein (Skeath and Doe, 1996).

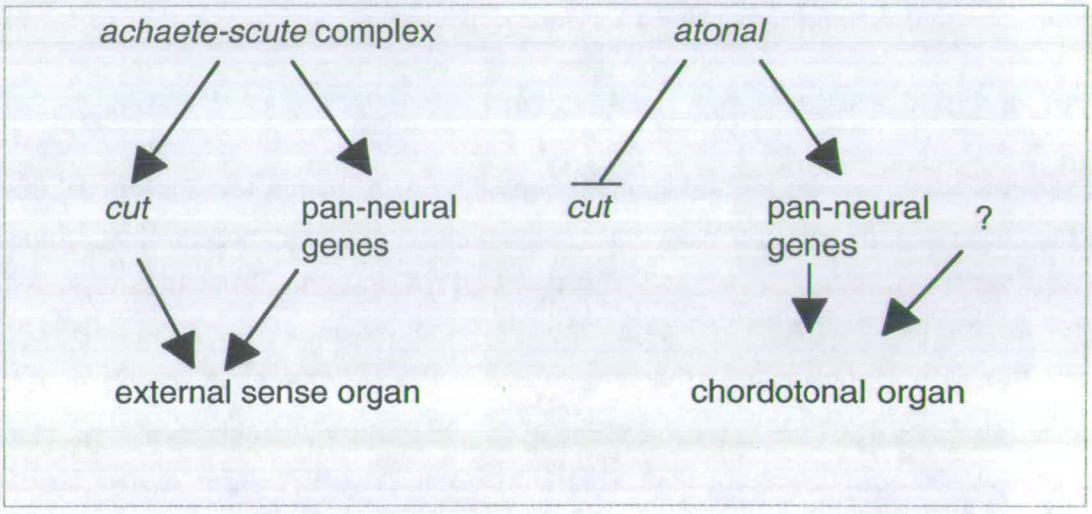


Figure 1.4.2 The different modes of action of the *AS-C* and *ato*

See text for details. In the absense of *ato* residual chordotonal organs form in the larva, presumably the action of another gene (s) can promote chordotonal identity in its absence.

1.4.3 Defining the pattern of neurogenic ectoderm

A complex network of prepattern genes and the regulation of proneural activity ensures the exact positioning of sensory precursors in both the embryo and adult. These processes involve the regulation of proneural expression, and interactions at the protein level. As well as the thoroughly characterised embryonic PNS, the easily identifiable adult mechanosensory bristles have provided excellent models in assessing the molecular basis of prepattern information.

In the embryo the early action of pair-rule genes and segment polarity genes set the co-ordinates for the expression of the proneural genes *ac* and *sc* (Skeath and Carroll, 1992), perhaps even directly affecting their expression (Jan and Jan, 1993).

Complex regulatory elements have been identified for the AS-C genes and for *ato* (Sun et al., 1998). The dynamic expression of the AS-C genes is controlled through the action of cis-regulatory sequences that are distributed throughout the complex (Ruizgomez and Modolell, 1987, Gomez-Skarmeta et al., 1995). It is thought that in the imaginal epithelium, a prepattern of asymmetrically distributed factors regulate the expression of *ac* and *sc* via these enhancer sequences.

Many regulators of proneural function have now been identified. As their names suggest, the loss of function of the genes *hairy* (*h*) and *extramacrochaetae* (*emc*) results in the production of ectopic sensory bristles. Indeed, the basis for these phentoypes is overexpression of the proneural genes *ac* and *sc*. H and Emc are structurally related to the proneural proteins (Rushlow et al., 1989, Garrell and Modolell, 1990). These factors refine *ac* and *sc* expression precisely to proneural clusters by either sequestering functional Achaete or Scute protein (Van Doren et al., 1991) or repressing their expression in cells outside the proneural clusters (in the case of *h*) (Skeath and Carroll, 1991; Van Doren et al., 1994).

The GATA-type transcription factor Pannier (*pnr*) and its zinc-finger regulator U-shaped, (*ush*) are required for the spatial regulation of *ac* and *sc* in a local manner (Ramain et al., 1993, Cubadda et al., 1997). They regulate *ac* and *sc* expression through the same dorsocentral-specific enhancer element (Haenlin et al., 1997). In addition the homeobox genes of the *Bar* (Sato et al., 1999) and *Iroquois* complexes

(Grillenzoni et al., 1998, Gomez-Skarmeta et al., 1996) have been implicated in the positioning of the adult sensory organs.

1.4.4 Refining neural competence to a single precursor

Once the proneural cluster has been established, proneural activity normally becomes restricted to a single precursor, the SOP. This process is the result of combined actions of the proneural genes and the neurogenic genes, which encode an inhibitory cell communication pathway (reviewed in Jan and Jan, 1994, Vervoort et al., 1997a). The proneural genes themselves orchestrate first mutual inhibition, then lateral inhibition, to resolve the expressing proneural cluster to a single SOP. In essence, proneural products in each cell of the cluster activate this pathway, which in turn acts to inhibit proneural expression in neighbouring cells. From this state of mutual inhibition, a single cell eventually predominates and completely represses proneural expression in the surrounding cells; expression is thus refined to a single SOP, which inhibits neural fate in neighbouring cells, ensuring its sole emergence from the proneural cluster.

The neurogenic genes include the products of *Notch*, *Delta*, the *Enhancer of split* genes (*E(spl)*), and *numb*, among others (reviewed in Campos-Ortega, 1993, Jan and Jan, 1993). Mutations for neurogenic genes results in higher levels of proneural expression, and therefore, supernumary sense organs (reviewed in Campos-Ortega, 1993). Among the best characterised neurogenic genes are those of the Notch pathway (Simpson, 1997).

The Notch pathway is unilaterally required for processes involving lateral signalling (Artavanis-Tsakonas et al., 1995 reviewed in Simpson, 1997, Bray, 1998, Chan and Jan, 1999) but the first roles for this pathway during neurogenesis are in mediating the selection of the neural precursor from the proneural cluster. Expression of activated forms of Notch causes all cells to differentiate as epidermis (Rebay et al., 1993 Sep;). The pathway consists of the ligands Delta and Serrate, the receptor Notch, and the downstream transcription factor Suppressor of Hairless (Su(H)), (reviewed in Simpson, 1997). Studies of Notch signal transduction have suggested a model whereby the Notch receptor, upon binding Delta, is cleaved at a site in or near its transmembrane domain to release the Notch intracellular domain from the membrane which then translocates to the nucleus and, with Su(H), activates the transcription of downstream target genes (Struhl and Adachi, 1998, Kidd et al., 1998,

reviewed in Bray, 1998). Such target genes include the genes of the *E(spl)* complex, the activation of which repress proneural expression in the proneural cluster (Jennings et al., 1995, Bailey and Posakony, 1995).

What determines which cell is chosen is not entirely clear. However, the decision ranges from an essentially random one, to a bias via the involvement of extrinsic signals, to predetermination of the precursor, depending on the epidermal versus neural fate circumstances (reviewed in Simpson, 1997, Baker and Yu, 1997).

1.4.5 SOP recruitment - neural clustering

The process of lateral inhibition ensures that only isolated SOPs are formed. However, as described earlier, many sense organ arrays arise from dense clusters of neural precursors (1.3). In particular, the mechanisms underlying the neural clustering of chordotonal arrays in the embryo and adult has recently been addressed (Okabe and Okano, 1997, zur Lage et al., 1997; zur Lage and Jarman, 1999). In these situations, lateral inhibition still limits chordotonal organ SOP formation by *ato*, as shown by the phenotypes of mutants defective for lateral inhibition. Such mutants have increased numbers of chordotonal organs (Hartenstein and Campos-Ortega, 1986) and neural precursors (Goriely et al., 1991), and an increase in *ato* expression (A.P.J. unpublished observations).

Although different organs operate slightly differing modes of recruitment, *ato* is at the top of the signalling cascade, mediating the process at the level of transcriptional regulation, activating the DER (EGFR) pathway, and quenching Notch signalling (zur Lage et al., 1997; zur Lage and Jarman, 1999, White and Jarman, in press) (Fig. 1.4.3).

In the embryo lch5 there are two distinct steps of chordotonal SOP formation (Okabe and Okano, 1997; zur Lage et al., 1997; zur Lage and Jarman, 1999). In the proneural stage, *ato* is refined to separate SOPs by lateral inhibition. After this, these founding precursors recruit further SOPs from surrounding uncommitted cells by EGFR signalling (Fig. 1.4.3). Adult chordotonal organs contain larger clusters, up to 80 scolopidia in the case of the femoral organ (zur Lage and Jarman, 1999). zur Lage and Jarman (1999) recently described a novel process in the femoral chordotonal organ formation (Fig. 1.4.3). The dynamics of femoral chordotonal organ SOP formation were analysed in relation to proneural nad neural markers. Furthermore,

the effects of diminishing or promoting the N lateral inhibition and EGFR inductive signalling pathways were studied with respect to SOP commitment and clustering. It was found that the large chordotonal array arises by progressive accumulation of a large number of SOPs from a persistent proneural cluster (PNC). This is achieved by an antagonistic relationship between inductive EGFR and competitive N signals. It was demonstrated that EGFR acts at the level of commitment as opposed to the proliferation of SOPs. EGFR acts in opposition to N in two ways: first it promotes continuous SOP recruitment despite lateral inhibition, and secondly it attenuates the effect of lateral inhibition on the proneural cluster equivalence group, thus maintaining the persistent PNC. SOP recruitment is iterative because the inductive signal comes from previously recruited SOPs.

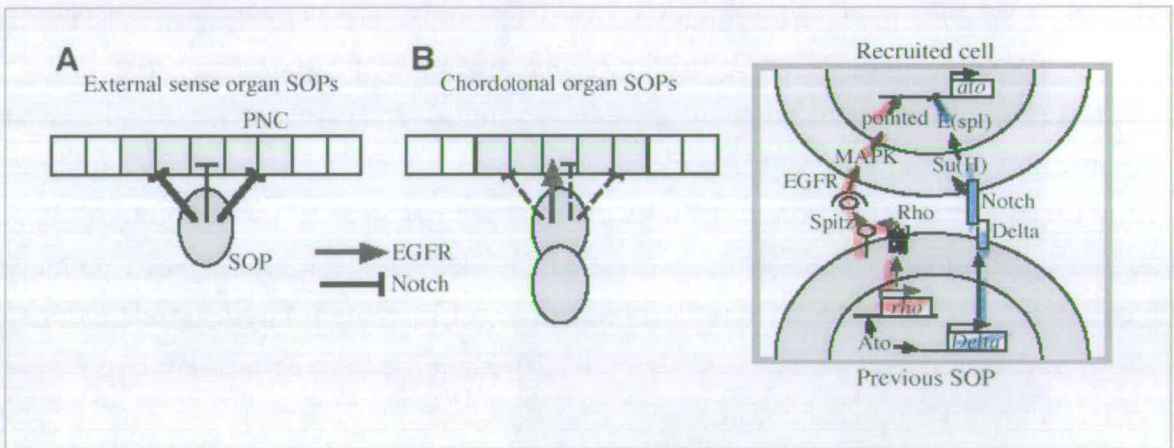


Figure 1.4.3 Model for the selection of SOPs for the femoral chordotonal organ.

(A) Solitary SOPs are formed by lateral inhibition. (B) By contrast, to generate the cluster in the femoral chordotonal organ, a trade-off between Notch and EGFR signalling acts to limit SOP commitment, but at the same time allows reiterative recruitment of SOPs and maintenance of the proneural cluster. In the panel on the right, some of the components of the EGFR (in red) and N (in blue) pathways are highlighted. This serves to demonstrate how the previously SOPS gives opposing signals to the next recruited cell. Taken from zur Lage and Jarman, 1999.

1.4.6 Differentiation of the SOPs

Once the SOPs have been determined, and delaminate from the ectoderm, these cells must divide, and/or differentiate as the individual sense organ cells. The SOPs themselves are characterised by their distinctive size, sub-epidermal location, and the expression of a group of genes, the pan-neural, or neural precursor genes. The action of pan-neural genes are hypothesised to establish and maintain the neural precursor fate and provide further neural information for their development (Bier et al., 1989; Cabrera and Alonso, 1991; Vaessin et al., 1991). Admittedly, the evidence for such roles is circumstantial in many cases, and as yet not well understood.

The pan neural genes identified to date are a diverse family of regulatory proteins thought to be expressed in most, if not all, neural precursors, and many products are present in the later division products of these cells. One such gene is *ase*, the only non-proneural member of the AS-C (Brand et al., 1993, Dominguez and Campuzano, 1993). Mutations for *ase* are viable, however, some sense organs are missing, presumably because SOPs fail to become established properly and do not differentiate (Brand et al., 1993). Binding sites for the proneural proteins *ac* and *sc* have been identified in the regulatory region of *ase* (Jarman et al., 1993a) indicating that expression is downstream of the proneural genes.

However, other pan-neural expression is essentially independent of proneural activity. For instance, *snail* (*sna*) pan neural expression is largely unaffected in proneural and *da* mutant backgrounds, suggesting other modes of regulation (Ip et al., 1994). Furthermore, although the gene appears to be globally expressed in neural precursors, the gene has separable CNS and PNS regulatory elements (Ip et al., 1994). Indeed, such separable enhancer elements have also been identified for the pan-neural genes, *deadpan* (*dpn*), and *scratched* (*scrt*) (Emery and Bier, 1995).

While the individual roles of these pan-neural genes are often difficult to assess, given the remarkable subtleties of their loss of function phenotypes, concomitant deletion of more than one of these genes often has dramatic effects on neural development (Roark et al., 1995). For instance, the deletion of both *scrt* and *dpn* causes a significant loss of neurons (Roark et al., 1995). Interestingly, *scrt* ectopic expression leads to supernumary neurons, and the reduced expression of some

neurogenic genes indicating a role for scrt in the promotion of neuronal fate (Roark et al., 1995).

Other known pan-neural genes may collaborate to establish neuronal fates by different mechanisms. For example, *prospero* (*pros*) turns off expression of pan-neural genes (Doe et al., 1991), such as *dpn* and *ase*, in ganglion mother cells (GMCs). *scabrous* (*sca*) (Mlodzik et al., 1990) encodes a secreted factor that inhibits neighbouring non-neural cells from adopting a neuronal fate, and cyclinA (Lehner and O'Farrell, 1989), regulates cell cycle progression. These examples serve to indicate that distinct parallel functions rather than a single orchestrating master gene promote neuronal fates.

1.4.7 What kind of sense organ? - subtype selector genes

Different subtypes of sense organs are specified by the activity of neuronal-type selector genes. The best characterised of these encodes the homeodomain protein Cut (Blochlinger et al., 1990; Blochlinger et al., 1991), the expression of which is differentially regulated by the different proneural genes to function as a binary switch for the cell fate choice of external sensory versus chordotonal organ formation (Jarman and Ahmed, 1998). In *cut* mutant embryos, the wildtpye external sense organs are completely transformed to chordotonal identity, thus indicating the complete control this gene displays over subtype fate (Bodmer et al., 1987; Merritt, 1997).

Other genes that govern sense organ identity include *paired-box neural* (*poxn*),*target of poxn* (*tap*) and the *Bar* genes. *poxn* encodes a pair-box containing protein and functions to discriminate between polyinnervated chemosensory versus singly innervated (mechanosensory) external sense organs in the embryo (Dambly-Chaudiere et al., 1992). One of its target genes, *tap* (Gautier et al., 1997), encodes a bHLH transcription factor expressed at the onset of differentiation in the embryo in a subset of chemosensory neurons. The *BarH1* and *BarH2* homeobox genes promote the formation of campaniform-like sensilla as opposed to trichoid sensilla (Higashijima et al., 1992).

1.4.8 Cell fate determination genes

Sense organ formation provides a model to address how a single cell can give rise to two cells that have distinct fates. In the context of cell fate determination, cell divisions that gives rise to two daughters which adopt different identities are termed asymmetric divisions (reviewed in Jan and Jan, 1998). In principle, asymmetric divisions may involve intrinsic or extrinsic factors. With extrinsic factors, daughter cells are born equivalent, but adopt different fates as a consequence of interactions with sibling cells, or their environment. Alternatively, daughter cells are provided with unequal amounts of cell fate determinants, intrinsic factors, that therefore bias the cell fate.

A detailed description of these mechanisms is beyond the scope of this introduction However, much has been learned in recent years regarding the genetics and molecular controls underlying these processes and have been reviewed recently by Vervoort et al., (1997a) and Jan and Jan, (1998). Indeed, many of the genes already described here are reiteratively used throughout development, and have specific roles during the specification of individual sense organ cell fates. Taking *Notch* and *numb* as examples, these represent key genes for cell-cell interaction and the intrinsic mechanisms used in generating asymmetry of the SOP cell divisions. In *Notch* and *numb* mutants, the four progeny cells of an external sensory SOP develop into four neurons or four socket cells, respectively (Hartenstein and Posakony, 1989; Rhyu et al., 1994). Indeed, inactivation of these genes at different points in the cell lineage causes different cell fate decisions (Vervoort et al., 1997a).

Many of the genes required for PNS development encode bHLH transcription factors. These include the proneural products of AS-C and *ato*, the ubiquitous *da*, the neurogenic *E(spl)*, pan-neural *dpn* and *ase* and the differentiation factor *tap*. The structure and function of these proteins and how they are utilised in other processes and organisms is addressed in the next section.

1.5 The basic helix-loop-helix (bHLH) protein family

bHLH transcription factors are now either known, or suspected, to be required during a wide variety of developmental processes in many different organisms (Carmeliet, 1999; Lee, 1997a; Littlewood and Evan, 1998). Indeed, it is now a broadly held view that these genes function in interlinked cascades, a related family of proteins acting within individual lineages to promote their development (Kageyama and Nakanishi, 1997; Kanekar et al., 1997; Lee, 1997a).

The large bHLH protein family of both invertebrates and vertebrates are united only by the conservation of a single functional domain, the bHLH motif (Murre et al., 1989a). When aligned by sequence identity, bHLH proteins fall into sub-categories that largely reflect their principal functions during development (Fig. 1.5.1). The role of this domain in the regulation of transcription has been well established by *in vitro* and *in vivo* studies (reviewed in Littlewood and Evan, 1998). The scaffold sequence of the motif is highly conserved across the entire family, so much so that crystal structure predictions of the well characterised MyoD and E47 (Ellenberger et al., 1994; Ma et al., 1994) can be used to computer-simulate the bHLH regions of other related, but functionally distinct, proteins (Chien et al., 1996).

1.5.1 Structural and functional relationships of bHLH motifs

Classically, the bHLH motif comprises a short stretch of hydrophilic residues (often basic) followed by a set of mostly hydrophobic residues located in two short segments (helix 1 and helix 2) and separated by a non-conserved sequence of variable length termed the loop region (Fig. 1.5.1) (reviewed in Littlewood and Evan, 1998). X-ray crystallography structural studies of bHLH proteins including MyoD (Ma et al., 1994, Fig. 1.5.2) and E47 (Ellenberger et al., 1994) have confirmed predictions based on *in vitro* binding studies and seqence analysis (e.g. Murre et al., 1989a; Murre et al., 1989b) as to how bHLH proteins function as regulators of transcription. bHLH proteins bind DNA as dimers by recognition of short palindromic sequences (E-boxes, CANNTG, Blackwell and Weintraub, 1990) such that each monomer contacts one half-site. The motif has two distinct functional compenents: a protein-protein interacting HLH region which promotes dimerisation with similar HLH containing proteins, and importantly constrains the basic (b) domain which in turn makes the principal contacts with DNA.

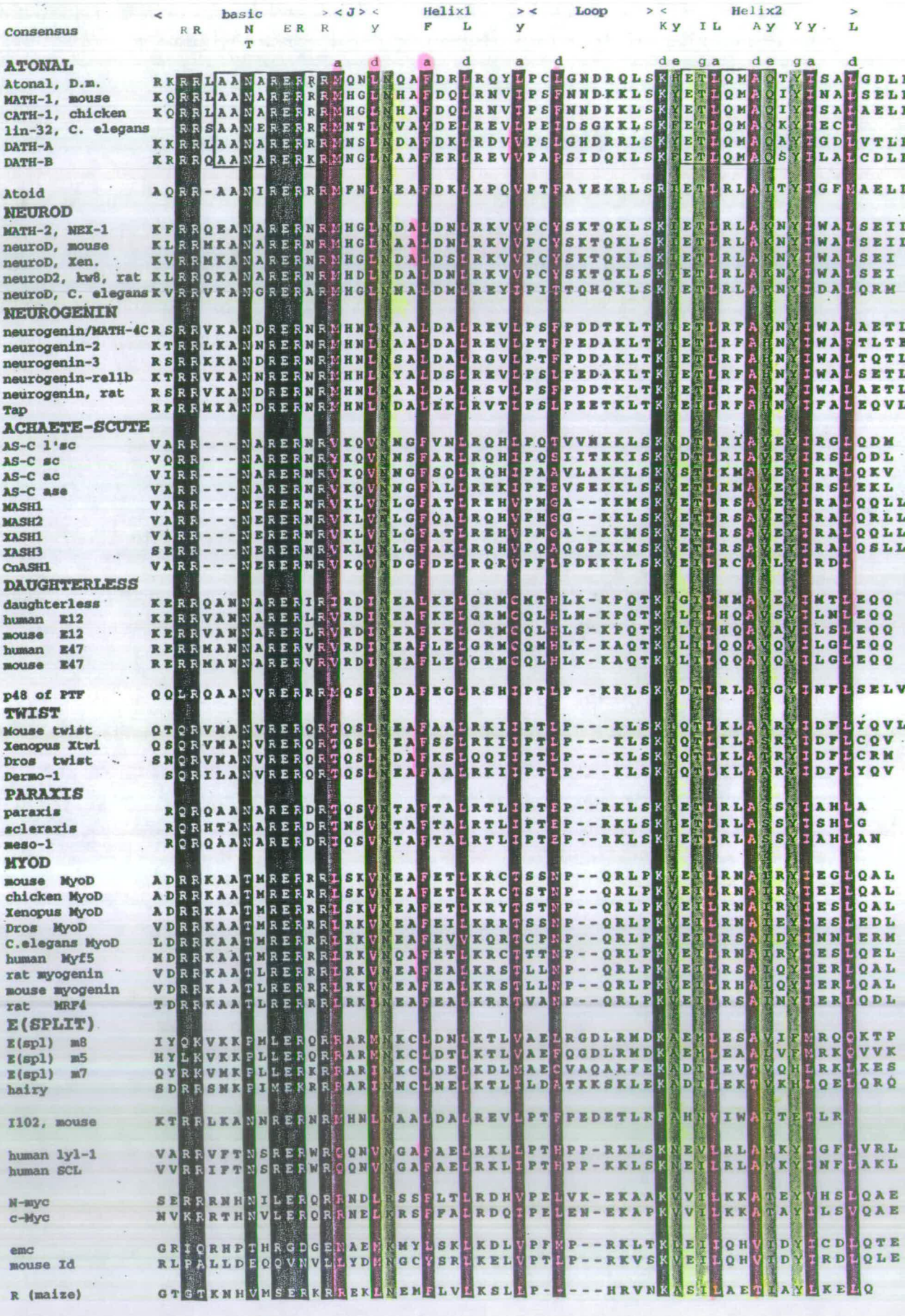
The crystal structures of E47, MyoD and Max bHLH/DNA complexes provide a basis for understanding the diverse dimerisation and DNA-binding properties of these and other bHLH domains (Ellenberger et al., 1994; Ferre-D'Amare et al., 1993; Ma et al., 1994) (Fig. 1.5.2).

Dimerisation of HLH proteins

In the HLH region, hydrophobic amino acids are positioned every third or fourth position throughout the helices so that the helix presents a surface of hydrophobic residues to the environment (Fig. 1.5.1 and Fig. 1.5.2). The distribution of these conserved residues allow pairing of monomers by hydrophobic interaction with similar surfaces. Disruption of these residues prevents dimerisation of bHLH proteins (Davis et al., 1990; Murre et al., 1989a).

Upon dimerisation, the HLH region forms a compact parallel four-helical bundle, which constrains the basic region (as an extension of helix1) to allow interaction with the major groove of the DNA binding site (Ellenberger et al., 1994; Ma et al., 1994) (Fig. 1.5.2). These hydrophobic residues are similarly conserved for function.

The least constrained region of the HLH motif is the loop region which is of variable length in different proteins (Fig. 1.5.1). The region appears not to be necessary for bHLH region conformation but does contribute to DNA contacts, a minimal loop size (4-5 amino acids) being necessary for DNA-binding (Littlewood and Evan, 1998). The surfaces of both helices are decorated with charged or polar residues. These are not conserved between different bHLH proteins suggesting they are not essential for function, but may confer specificity.


basic
Helix1
Loop
Helix2
Consensus
ATONAL
Atonal, D.m.
MATH-1, mouse
CATH-1, chicken
lin-32, C. elegans
DATH-A
DATH-B
Atoid
NEUROD
MATH-2, NEX-1
neuroD, mouse
neuroD, Xen.
neuroD2, kw8, rat
neuroD, C. elegans
NEUROGENIN
neurogenin/MATH-4C
neurogenin-2
neurogenin-3
neurogenin-rel1b
neurogenin, rat
Tap
ACHAETE-SCUTE
AS-C l'sc
AS-C sc
AS-C ac
AS-C ase
MASH1
MASH2
XASH1
XASH3
CnASH1
DAUGHTERLESS
daughterless
human E12
mouse E12
human E47
mouse E47
p48 of PTF
TWIST
Mouse twist
Xenopus Xtwi
Dros twist
Dermo-1
PARAXIS
paraxis
scleraxis
meso-1
MYOD
mouse MyoD
chicken MyoD
Xenopus MyoD
Dros MyoD
C.elegans MyoD
human Myf5
rat myogenin
mouse myogenin
rat MRF4
E(SPLIT)
E(spl) m8
E(spl) m5
E(spl) m7
hairy
I102, mouse
human lyl-1
human SCL
N-myc
c-Myc
emc
mouse Id
R (maize)

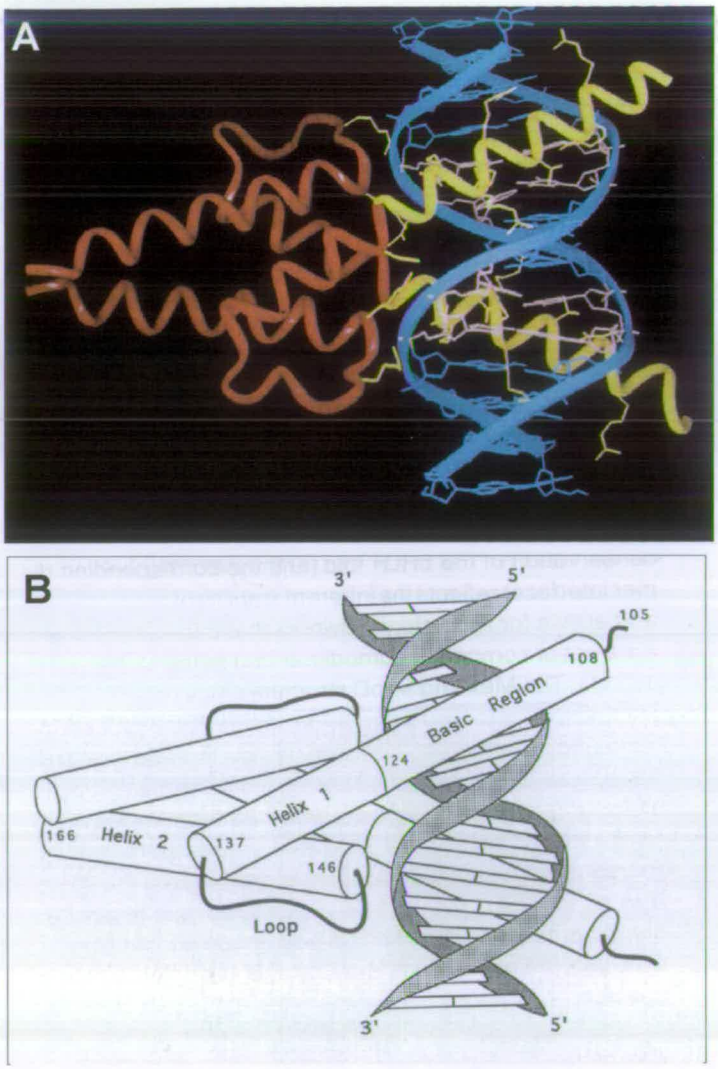


Figure 1.5.2 Overview of bHLH-DNA complex

(A) Overview of the MyoD-DNA complex. The blue region depicts 14bp of DNA. The violet residues denote the E-box. The basic region is highlighted in yellow, and the HLH region is red. Side chains that make contacts to the DNA are also highlighted. (B) Sketch of (A). Adapted from Ma et al., 1994.

Figure 1.5.1 Line-up of bHLH regions from a range of factors

(facing page) In this lineup the loop regions from longer proteins (e.g. AS-C) are not represented. The residues conserved for function are shown in the top line (consensus). Different sub-families have more closely related sequences. Black denotes residues required for DNA binding, in the basic and H2 regions. Residues marked in pink (contacts between the two H2 helices) and yellow (contacts between H2 with the H1 helices) are necessary for dimerisation of the HLH region

---

DNA binding of bHLH proteins

The basic domain is a helical extension of helix 1 (H1) comprising about 15 residues of the domain, and a junction region to H1 of 4 or so residues. The basic domain has at least two separable functions. In the main, this region is necessary for recognition of the consensus binding sites, making the principal contacts with the DNA, through the phosphate backbone, and the bases themselves (Ellenberger et al., 1994; Littlewood and Evan, 1998; Ma et al., 1994). It has been established that over half of the basic domain residues are required in mediating these contacts with the sequence of the E-box. The ability of individual bHLH proteins to recognise distinct half-sites within the dyad symmetry of the E-box consensus sequence therefore allows different combinations of HLH proteins to discriminate between different E boxes (Blackwell and Weintraub, 1990).

The binding preferences of different types of bHLH proteins are often similar (e.g. MyoD and Mash1) and the specificity of DNA binding is relatively low (Kunne et al., 1996). Thus the individual specificities of these proteins is most likely achieved through the co-operative interaction with other proteins of the transcriptional machinery. Indeed, the myocyte enhancer factors-2 (MEF2) can act as co-regulators to potentiate on the one hand, the myogenic effects of MyoD and neurogenic action of Mash1 on the other, through interactions between the DNA binding domains (Mao and Nadalginard, 1996; Molkentin et al., 1995).

Tissue-specific DNA transcriptional activation is a separable function of the basic domain (Davis et al., 1990; Davis and Weintraub, 1992). Experiments which disrupt the basic region of bHLH proteins have demonstrated that such proteins still bind DNA specifically, but fail to initiate developmental programmes such as myogenesis (Davis et al., 1990). Indeed, this myogenic specificity has been pinpointed to three residues of the MyoD basic domain, not conserved with other bHLH factors (Davis and Weintraub, 1992). Similarly, functional differences between *ato* and *sc* are attributable to several residues within the basic domain (Chien et al., 1996). Interestingly, in these cases, the residues that are implicated for specificity are not conserved or required for DNA contact suggesting that these positions of the bHLH domain may be important for interactions with other factors. It is noteworthy that bHLH dimers are inherently unstable, and presumed to require additional interactions to stabilise them as active transcriptional complexes (Littlewood and Evan, 1998 and refs. therein), as mentioned above.

1.5.2 Classes of bHLH factors

The widely accepted view of tissue-specific regulation by the proneural and myogenic bHLH activators is that dimerisation of a tissue-restricted bHLH factor with a ubiquitously expressed member is essential for functionality. Thus the bHLH proteins can be grouped broadly into three distinct classes; those that are ubiquitously expressed (Class A), those that are tissue specific and promote transcription when complexed with a class A factor (Class B), and finally those which negatively regulate the activity of Class A and Class B proteins (Class C).

Ubiquitous bHLH proteins (Class A factors)

Among the first bHLH proteins to be characterised were the gene products of the vertebrate *E2A* gene, E12 and E47 which were identified in a screen for proteins that bound labelled E-box sequences (Murre et al., 1989a). The single *Drosophila* homologue of the E-proteins is Da (Caudy et al., 1988b). The evidence that these proteins act as the dimerisation partners for tissue specific factors to effect DNA binding and thus transcriptional regulation has several sources. *In vitro* studies have demonstrated the preferential binding affinities of bHLH proteins as heterodimers with Class A factors, in the presence of specific DNA sequences (Cabrera and Alonso, 1991; Murre et al., 1989b). Genetic analysis revealed *da* as an essential gene for nervous system development: all precursors of the PNS fail to become established in *da* null mutants (Bodmer et al., 1989; Caudy et al., 1988a), whereas only specific subsets are deleted in proneural mutants (Dambly-Chaudière and Ghysen, 1987; Ghysen and O'Kane, 1989). Furthermore, dominant genetic interactions have been observed between *da* and AS-C (Dambly-Chaudière et al., 1988) as well as for *da* and *ato* (Gupta and Rodrigues, 1997), emphasising their functional cooperativity.

Negative bHLH and HLH regulators (Class C factors)

Just as there are bHLH transcription activators, so there are also proteins that counter-regulate the activities of these factors to restrict and control the timing of development. There are two classes of such negative regulators.

1. Hairy-related bHLH proteins

One type of negative regulator encodes bHLH DNA-binding transcriptional repressors. In *Drosophila* this repressor family comprises proteins such as Hairy (Rushlow et al., 1989; Van Doren et al., 1994), the E(spl) complex (Nakao and Campos-Ortega, 1996) and Dpn (Bier et al., 1992, Younger-Shepherd, 1992). The vertebrate counterparts of these genes are the HES proteins (e.g. Sasai et al., 1992). Together these genes are referred to as *hairy*-related genes because of their structural and functional similarities to Hairy (Fisher and Caudy, 1998).

All *hairy*-related genes share two features that distinguish them from the rest of the bHLH superfamily (Fisher and Caudy, 1998). In the first instance, the basic domain of these proteins has a different arrangement of basic residues, and a conserved proline at a specific location within the domain. These changes alter the binding specificities of hairy-related proteins that preferentially bind an alternative hexamer, the so-called N-box (CACGCG, Ohsako et al., 1994; Van Doren et al., 1994). The second feature of these proteins essential for their function as repressors is the presence of a short 4-amino acid C-terminal motif (WRPW) which mediates an interaction with co-repressor molecules (Fisher, 1996). In *Drosophila* one such co-repressor was identified as Groucho, itself encoded by the E(spl)-C (Paroush et al., 1994). Thus Hairy-related proteins do not interfere with DNA-binding of bHLH transcriptional activators but instead function as active repressors of transcription, including that of proneural bHLH genes themselves (Fisher and Caudy, 1998).

2. The HLH negative regulators

The second class of negative regulators are exemplified by *Drosophila* Emc (Garrell and Modolell, 1990) and the family of vertebrate Id (inhibitor of differentiation) homologues (Benezra et al., 1990; Carmeliet, 1999; Lyden et al., 1999). These HLH proteins lack a basic domain and thus dimers of the proteins cannot bind DNA. Proteins of this kind directly antagonise tissue specific bHLH function by sequestering ubiquitous bHLH factors as inactive heterodimers (Lee, 1997a; Littlewood and Evan, 1998 and refs. therein).

1.5.3 Tissue specific bHLH proteins in development

Many different tissue specific bHLH factors have been identified for a range of developmental processes that include haematopoiesis, angiogenesis, myogenesis, dermal cell differentiation and neurogenesis (Lee, 1997). This section focusses on some of the best characterised roles for bHLH proteins in development.

Myogenesis

Both vertebrates and invertebrates require the activity of the MyoD bHLH family for myogenesis (reviewed in Baylies et al., 1998; Littlewood and Evan, 1998). Indeed, vertebrate skeletal muscle is one of the most thoroughly characterised developmental systems utilising bHLH-mediated transcriptional networks Molkentin and Olson, 1996). At least four key regulatory genes of the bHLH family have been identified: MyoD, myogenin, Myf5 and MRF4 act at multiple steps in the myogenic lineage to control muscle specific gene expression (reviewed in Littlewood and Evan, 1998; Molkentin and Olson, 1996). Strikingly, forced expression of any of these genes in non-muscle cells results in myogenic conversion, measured by expression of skeletal muscle myosin (Molkentin and Olson, 1996). Indeed sequence conservation is high among these genes, and even extends beyond the bHLH region itself (Atchley et al., 1994), a feature unusual among members of this family. However, though they share common properties, each gene is differentially expressed during embryonic development, suggestive of non-overlapping functions in the process of muscle formation (Buckingham, 1992). Furthermore, close analysis of single and double knockout mutations in the mouse has confirmed this idea, demonstrating that MyoD

family members play critical roles in both the determination and differentiation of striated muscle (reviewed in Littlewood and Evan, 1998; Molkentin and Olson, 1996).

At present, there is only one known fly MyoD homologue – *nautilus* (*nau*) (Michelson et al., 1990). Expression of the gene is limited to a subset of muscle precursors and differentiated fibres (Michelson et al., 1990) however, like its vertebrate counterparts, *nau* is capable of inducing muscle specific gene expression when ectopically expressed in other cells of mesodermal origin (Keller et al., 1997). Deficiency phenotype data suggest that *nau* functions more like myogenin, i.e. during the differentiation of muscle cells (Keller et al., 1998). However, a conflicting report based on RNAi data among other analyses seems to support a more fundamental role for *nau* in the formation of muscle precursors (Misquitta and Paterson, 1999). Isolation and analysis of genetic null mutations of this gene will be necessary to be able to tackle this issue with certainty.

Indeed, many other bHLH proteins are required for myogenesis. In *Drosophila*, Twist is both essential for the establishment of mesodermal fate before gastrulation and is involved in subdividing the mesoderm later in development (reviewed in Baylies, 1998). Several vertebrate homologues of *twi* have been identified, also with early roles during mesoderm specification (Spicer et al., 1996). Other regulators of myogenesis include the closely related *capsulin* (or Pod-1, Lu et al., 1998) and MyoR (Lu et al., 1999) which are thought to act as lineage-restricted transcriptional repressors and inhibitors of myogenesis. In *Drosophila*, the proneural gene *l'sc* functions early in the specification of groups of cells to acquire a myogenic fate (Carmena et al., 19955). Late-acting factors such as bHLH54F distinguish longitudinal visceral muscles in the embryo (Georgias et al., 1997).

Sex Determination

Interestingly there is striking overlap between the genetic control of neurogenesis and sex determination in *Drosophila* (Younger-Shepherd, 1992, reviewed in Parkhurst and Meneely, 1994). Many of the HLH proteins utilised in the early steps of neurogenesis are also used earlier on during embryonic development to regulate the initial steps of sex determination. Shared regulators include *sc* (called *sis-b* for sex-determination), *da*, *emc*, *dpn*, and *h* (reviewed in Parkhurst and Meneely, 1994). Although functionally equivalent for neurogenesis, the other members of the AS-C cannot substitute for *sc* during sex determination (Parkhurst et al., 1993). This finding highlights the existence of functionally relevant domains outside the bHLH region.

1.5.4 bHLH proteins in neurogenesis

The roles of bHLH factors during *Drosophila* neurogenesis have already been described (see 1.4). So far, bHLH proneural homologues have been identified in every model organism examined. Vertebrates express scores of neural bHLH genes during their development (Anderson, 1999; Kageyama and Nakanishi, 1997; Lee, 1997a). However, their expression patterns are spatially and temporally complex and by comparison with the proneural functions of *Drosophila* the study of these proteins in vertebrates is in its infancy. Nevertheless, it is clear that much can be learned from vertebrate studies in terms of neurogenesis in both organisms, in particular, the roles of bHLH cascades and networks (Kanekar et al., 1997; Roztocil et al., 1997).

The rapid identification of new factors makes for somewhat confusing nomenclature and indeed, any attempt to comprehensively list these proteins is likely to be quickly out of date (for recent review see Littlewood and Evan, 1998). Instead, I present in this section, a description of the shared features of particular factors that give further insights into the mechanisms of bHLH function during neurogenesis.

Vertebrate proneural homologues

Structurally, vertebrate neural bHLH genes can be divided into two broad classes, those that are AS-C homologues, such as *Mash1-2*, *Xash1*, *Xash3*, and *Cash4* and alternatively a large *ato*-like class which include the *Math* genes, *neurogenins* (*Ngn*) and *NeuroD* (reviewed in Lee, 1997a). Targeted disruptions in many of these genes causes the loss of specific subsets of neurons, demonstrating their essential roles in neurogenesis (reviewed Anderson, 1999; Kageyama and Nakanishi, 1997; Lee, 1997a). However, the highly complex expression patterns of these genes make it difficult to discern whether they act as proneural genes in these systems.

For example, mutations in *Mash-1* results in the loss of subsets of olfactory, retinal and autonomic neurons (Guillemot, 1995; Guillemot et al., 1993). However, *Mash-1* is not expressed during the determination of neural precursors, nor is it required for their specification; analysis of *Mash-1*-null embryos indicates that it is the transition from precursor to post-mitotic cell which seems to be affected suggesting a later role for *Mash-1* (Guillemot, 1995; Guillemot et al., 1993). Thus this gene must promote the switch from growth phase to the neuronal differentiation phase. Similarly, of the *ato*-like *Math* genes, few appear to be expressed early enough to be classed as proneural. Phenotypically, the best characterised of these is *Math1* which appears to be essential for a substantial population of cerebellar neurons (Ben-Arie et al., 1997). *Math1* expression in these cells led the authors suggest a role for the gene either in granule cell fate specification or proliferation.

Given that the expression and functional requirements are in cells that are already fated to become neural, genes such as *Mash1* and *Math1* cannot be considered as proneural in the classic sense of conferring neural competence upon uncommitted cells.

However, several homologues have now been identified which are more likely candidates as vertebrate proneural genes. For instance *Cash4*, a chick AS-C homologue, is expressed in response to neural inducing signals suggestive of an early role in precursor specification (Henrique et al., 1997). *Math5* expression is significantly upstream of other neural bHLH factors during retinogenesis, correlating with the appearance of the first-born retinal neurons, and appears to be regulated by *Pax6* (Brown, 1998). Intriguingly, *Math1*-null mice also lack cochlear and vestibular

hair cells of the inner ear sensory epithelia, leading the authors to propose that in this instance, *Math1* is acting as a pro-hair cell gene, a function analagous to that of *ato* during chordotonal and ommatidia development (Bermingham et al., 1999).

Proneural characteristics of the vertebrate homologues

Regardless of their respective roles, it is quite clear that many of these vertebrate bHLH factors have much in common with their fly counterparts in terms of the functional potential to act as proneural factors.

Many of these genes have been characterised as 'proneural' by virtue of their behaviour in gain of function experiments in *Xenopus* embryos. In such assays, the genes *Xash3* (Turner and Weintraub, 1994), *NeuroD* (Lee et al., 1995), *Ngn1* (Ma et al., 1996), *Xath5* (Kanekar et al., 1997), and *Cash4* (Henrique et al., 1997), promote neuronal differentiation in ectodermal tissue. Furthermore, some of these genes are sensitive to the process of lateral inhibition, a hallmark of fly proneural gene function. For instance, *X-ngnr1* and *Xash3* are sensitive to lateral inhibition whereas *NeuroD* is not (Lee, 1997b).

However, given the complexity of expression patterns for these genes, it is difficult to draw many conclusions as to whether the functions of these genes are intrinsic to their different structures, or merely the result of differential regulation and consequences of the developmental context of their expression. For instance, *Mash1* and *Ngn1/2* are expressed in complementary patterns in several tissues but these genes are also expressed much later during neurogenesis (Lee, 1997a). It seems likely that a combination of intrinsic differences, regulation and positional information must act together in some way to modulate their activities.

In *Drosophila*, misexpression experiments have indicated that proneural genes also contribute to the specification of neuronal identity (Chien et al., 1996; Jarman and Ahmed, 1998; Jarman et al., 1993b). It is now evident that different vertebrate neural factors similarly encode neuronal specificity. For example, the forced expression of *Ngns* in chick embryos causes the activation of several different sensory-specific markers, in addition to pan-neural genes, reflecting their requirement for sensory ganglia in the neural crest (Anderson, 1999 and refs. therein). Indeed, the role of *Mash1* in promoting adrenergic fate in subsets of neurons in which it is expressed

has been assessed by reference to specific markers such as the *Phox* genes (Hirsch et al., 1998; Lo et al., 1998).

Differentiation factors

Many of the bHLH factors identified are expressed later in development, during the differentiation stage, once precursors have proliferated (Lee, 1997a). In particular, the roles of *NeuroD* have been well characterised in terms of its neural differentiation expression and misexpression phenotypes (Lee, 1997b; Lee et al., 1995). In several different tissues, its expression is downstream of genes such as *Mash1* (Cau et al., 1997) and the *Ngn*s (Ma et al., 1996). In both mice and frogs, *NeuroD* is highly expressed in structures that contain differentiating neurons (Lee et al., 1995). Thus, *NeuroD* is suggested to function in the terminal differentiation of postmitotic cells. This gene has different functional capabilities than other early genes; in a Xenopus assay, while *Ngnr1* and *NeuroD* can both induce neurogenesis, in the latter case this is limited to ectodermal derivatives. Furthermore, *Ngnr-1* induces *NeuroD* activation, but the reverse is not so (Ma et al., 1996, Lee et al., 1995, reviewed in Kageyama and Nakanishi, 1997).

Other late-expressing factors that have been identified include *NeuroD2*, *Math2* and *NSCL1-2* and *NeuroM* (reviewed in Lee, 1997, Kageyama and Nakanishi, 1997, Littlewood and Evan, 1998, Anderson, 1999).

Negative regulators of neurogenesis

The *hairy*-related genes and *emc* related *Id* genes are required to regulate the timing of neuronal differentiation (Carmeliet, 1999; Fisher and Caudy, 1998).

hairy-related genes appear to be required throughout development (Lee, 1997). Taking *Drosophila* neurogenesis as an example, there are the early pre-pattern roles of *hairy* itself, through the involvement of the *E(spl)* genes during lateral inhibition to the poorly understood functions of *dpn* during differentiation. Similarly, in vertebrates, the HES genes are widely expressed during neurogenesis, and are

required to ensure the correct timing of neuronal differentiation (Lee, 1997a). The absence of *Hes1* promotes acceleration of precursor differentiation in the neural tube and retinal tissues (Ishibashi et al., 1995, Tomita et al., 1996) causing disproportionate neuronal ratios and small tissue sizes. Furthermore, *Mash1* expression is upregulated in *Hes1*-null embryos, suggesting that one role of this gene is to repress neuronal determination factors, maintaining the cells as precursors (Ishibashi et al., 1995). As in *Drosophila*, genes such as *Hes1* and *Hes5* can be induced by the activation of the Notch lateral inhibition pathway (Kageyama and Naganishi, 1997).

Similarly, though distinct structurally and functionally, the *Id* genes are expressed in a variety of tissues and act as are temporal regulators of neuronal differentiation (Carmeliet, 1999; Lyden et al., 1999). Targeted disruption of Id factors causes cell-cycle withdrawal, premature differentiation and reduced proliferation (Lyden et al., 1999).

Networks and cascades of bHLH genes

It is now a generally held view that bHLH proteins function in cascades in a given tissue. The characterisation of the vertebrate bHLH factors at the level of expression indicates that many networks and cascades of bHLH genes exist, with different genes having different roles at distinct developmental stages. For several lineages, cascades of bHLH gene expression have been described that appear to account for all of the separable stages of neural development (Cau et al., 1997, Kanekar et al., 1997, Rottgen et al., 1998; Roztocil et al., 1997).

In contrast to its late role in autonomic neuron development, *Mash1* heads a cascade of expression of bHLH factors including *Ngn1* and *NeuroD* in the mammalian olfactory neuron lineage (Cau et al., 1997) (Fig. 1.5.3). In the absence of *Mash1* these genes are not expressed, suggesting that in some way *Mash1* is involved in initiating, or continuing the cycle of expression. A second cascade was identified in a separate *Mash1* independent population, headed in this instance by *Ngn1* (Cau et al., 1997).

In the chick, the identification of *NeuroM*, which is expressed throughout the developing nervous system, defined a new stage in the cell-differentiation process (Roztocil et al., 1997, Fig.1.5.3). *NeuroM* is defined as a postmitotic, premigratory neuron factor. Its expression follows that of *ASH1* and *Ngn*, but precedes *NeuroD*, thus defining another expression cascade. The location of *NeuroM* expressing cells was found to be at the interface between proliferating and non-proliferating zones of the neural tube. Thus for this particular lineage, a different bHLH protein was expressed for every defined stage of the cell's development (Roztocil, 1997).

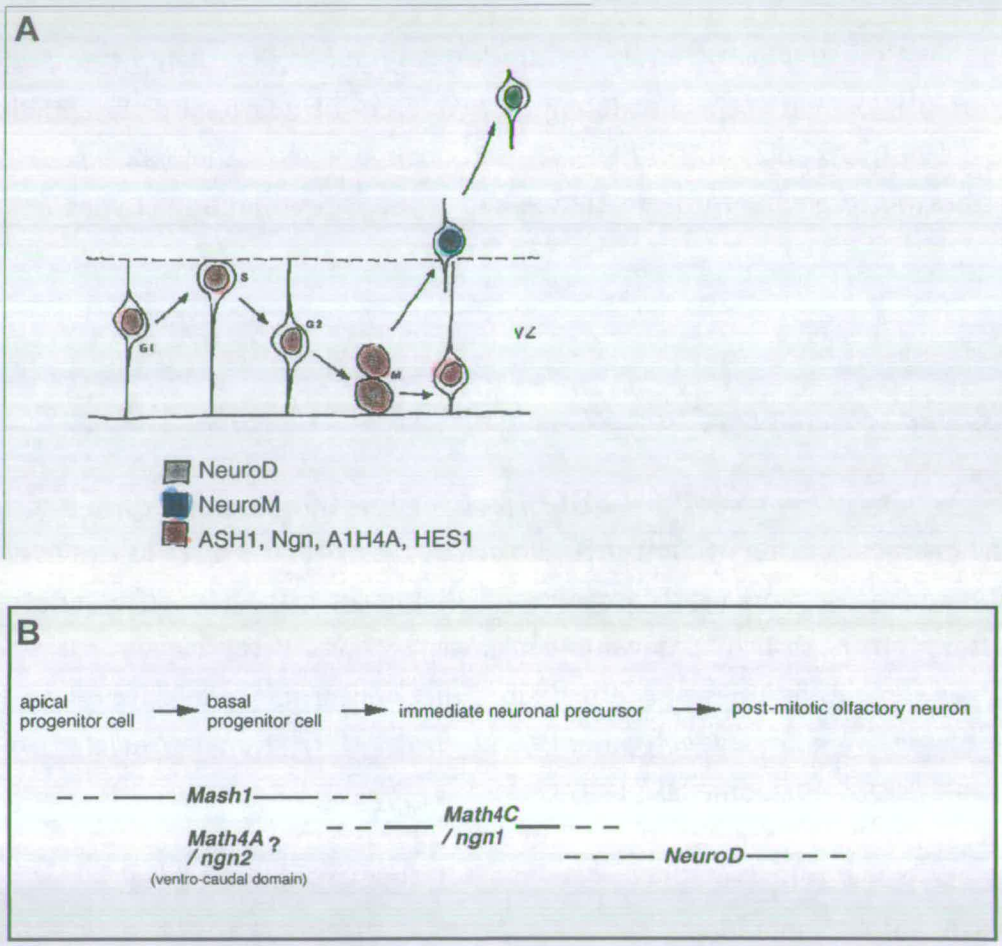


Figure 1.5.3 Vertebrate bHLH expression cascades

(A) Taken from Roztocil et al., 1997. (B) Taken from Cau et al., 1997. See text for details

1.6 Aims of this thesis

1.6.1 The likelihood of other proneural genes in *Drosophila*

At present, the proneural gene family of *Drosophila* comprises the AS-C and *ato* but these genes alone do not account for the formation of the entire nervous system (Jarman and Jan, 1995). Indeed, deletion of both AS-C and *ato* still leaves portions of the PNS and most of the CNS intact (Jarman et al., 1993b). Although it is conceivable that other mechanisms are responsible for these additional proneural functions, there is some evidence that suggests it is worthwhile searching for other bHLH candidates in *Drosophila*. For example, in embryos, deletion of the ubiquitously required *da* disrupts the entire PNS and CNS (Caudy et al., 1988a). Thus there must be other proneural candidates that interact with *da* to account for the rest of the larval nervous system, likely of the bHLH family. Moreover, negative HLH regulators such as *emc*, are expressed in the anlage of nervous system components for which there are no known proneural functions (Gupta and Rodrigues, 1997), again auspicious of undiscovered bHLH factors.

1.6.2 The likelihood of downstream bHLH factors in *Drosophila*

Interestingly, searches for vertebrate proneural homologues and bHLH proteins in general, have uncovered several genes that are expressed not during the determination of cell fate, but at later stages as cells differentiate (1.5.4). Although most evidence so far is based on expression patterns, it seems likely that for vertebrate neurogenesis and myogenesis at least, related bHLH proteins operate in functional cascades. In support of this, in *Drosophila, da* and some of the negative HLH regulators (e.g. *dpn*) are expressed throughout development. Indeed, there are some known tissue-specific bHLH genes that are expressed during, and presumably act at, later stages of neurogenesis, once precursors have been selected. Examples such as *ase*, the pan-neural gene thought to be required for the proper establishment and maturation of SOPs (Brand et al., 1993; Jarman et al., 1993a), and *tap,* expressed in a subset of neurons during their final differentiation (Gautier et al., 1997), hint at

the possibility of similar cascades in *Drosophila*. These cases aside however, most information to date concerning bHLH function during *Drosophila* neurogenesis centres on the proneural determination factors, and little is known regarding the existence and function of the positive bHLH regulators later during development.

A secondary aspect of this thesis concerns the functional specificity of the bHLH proteins themselves, in particular that of *ato*. Several studies have shown that it is the bHLH sequence alone that encodes the particular proneural functions of these proteins (Chien et al., 1996; Jarman and Ahmed, 1998).However, sequence identity among this family rarely extends beyond the bHLH region itself thus making it difficult to determine whether other domains of structural or functional relevance are conserved. Indeed, recent evidence suggests that, in one instance, DNA-binding properties of the bHLH domain are not necessary for function (Porcher et al., 1999). These findings make it even more pertinent to identify functional domains outside the bHLH region, and to pinpoint within it specific residues important for different activities.

Inter-species comparisons have been used to identify and characterise bHLH homologues in related species with a view to uncovering extra conserved domains (e.g. Botella et al., 1996). Identification and analysis of *ato*-related genes or *ato* homologues from other species that retain similar functions, may aid future investigations of the mechanics of bHLH protein function, using *ato* as a model.

Clearly, to address these issues, new neural bHLH factors must first be identified. This was the specific aim of this thesis, i.e. to identify and characterise new bHLH candidate genes in *Drosophila melanogaster*.

Chapter 2

Isolation and characterisation of new *ato*-related genes

2.1 Introduction

Despite extensive genetic screens for neural genes few, if any, new bHLH genes have been identified in this way. Indeed, there are no P-element mutations for any of the known proneural genes and EMS mutagenesis proves difficult because these proteins are generally encoded by small, single exon genes (Jarman et al., 1993b). Reverse genetics seems to be a more tractable approach to finding these genes, as has been borne out by the number of vertebrate homologues that have been identified by such methods. Indeed, both *ato* and *tap* were identified in the first instance by reverse genetics, using degenerate PCR and low stringency hybridisation respectively (Gautier et al., 1997; Jarman et al., 1993b). Several other bHLH genes with homology to either vertebrate or fly genes that function in other developmental processes have been similarly identified.

This chapter details the isolation and molecular characterisation of two *ato*-related genes, identified by degenerate PCR. The genes have been termed *amos*, for absent MD neurons and olfactory sensilla, and *cato*, for cousin of *ato*.

2.2 Degenerate PCR screen for *ato* homologues

A PCR approach was used to identify new *ato*-like genes in *D. melanogaster* and to isolate *ato* homologues from other dipteran species. Degenerate primers were designed to the parts of the *ato* bHLH domain that are most conserved between Ato and its closest vertebrate homologues (Math1 and Math5) (Fig. 2.2.1). These residues in particular distinguish the Ato subfamily from all other bHLH proteins. Genomic DNA from several dipteran species (Table 2.2.1) was used in PCR with these primers. PCR conditions used were 30 cycles of 55°C (1 min), 72°C (1 min) and 94°C (1 min). Products of approximately 130 bp (as estimated by agarose gel electrophoresis) were amplified in each case and then cloned into pGEM-T and DNA from a number of clones sequenced (see Table 2.2.2).

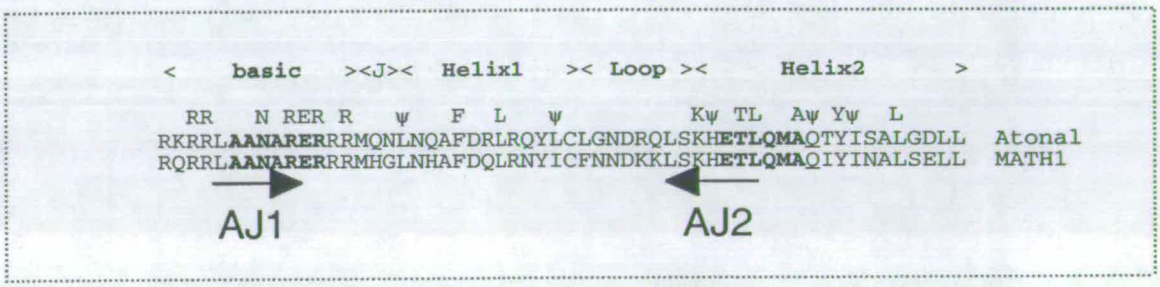


Figure 2.2.1. The bHLH region showing degenerate primer locations

Degenerate primers (AJ1: 5'-GCYGCYAAYGCHCGYGARMG-3' and AJ2: 5'-TGRGCCATYTGBARDGTYTC-3') were designed to the conserved ends of the *ato* and MATH1 bHLH regions. Most conserved residues are indicated above sequences. Regions of basic, helix and loop regions are marked above. An arrow indicates the direction of each primer. The predicted PCR product size for *ato* 126bp.

Table 2.2.1 Species used in degenerate PCR

diptera
Drosophila melanogaster
Drosophila virilis
Musca domestica
Calliphora erythrocephala

The PCR products contained a mixture of different sequences. *ato* bHLH region equivalents were amplified from all species and were identical to *ato* at the amino acid level (Fig. 2.2.2). In addition, two novel sequences were identified that, at that time, had no counterparts in any sequence database (Fig. 2.2.3). One was named ***a**bsent **m**d neurons and **o**lfactory **s**ensilla* (*amos*) and was initially identified in *D. melanogaster*. The second sequence was originally identified only in products from *Musca domestica* and *D. virilis*. Therefore, to isolate the *D. melanogaster* orthologue of this new gene, I screened over 300 cloned *D. melanogaster* PCR products by filter colony hybridisation with *ato* and *amos* probes. 78 colonies did not hybridise to either probe and, of these, four were chosen at random and sequenced. All four sequences closely resembled the *M. domestica* sequence at the amino acid level and the novel gene was termed ***c**ousin of **ato*** (*cato*) (Fig. 2.2.4). Table 2.2.2 summarises the number of clones sequenced, and their assigned identities based on sequence translations and comparisons.

Table 2.2.2 Summary of sequences isolated by degenerate PCR

Species	No. of clones sequenced	*Ato*-homologue	Other sequences
D. melanogaster	9	*ato*	new gene: *amos*
			new gene: *cato*
D. virilis	10	+	*cato-like*
M. domestica	8	+	*cato-like*
C. erythrocephala	4	+	-

```
                A  A  N  A  R  E  R  R  R  M  Q  N  L  N  Q  A  F  D  R  L
D. melanogaster GCCGCCAATGCACGTGAGCGTCGTCGGATGCAGAACCTCAACCAGGCCTTCGATCGTCTC
D.virilis                     CGTGAACGTCGCCGCATGCAAAACCTCAACCAGGCCTTCGATCGACTG
Calliphora      GCTGCTAATGCTCGCGAGCGTCGTAGAATGCAAAATCTCAATCAAGCTTTCGATCGTCTA
Musca           GCCGCTAATGCTCGTGAACGTCGCCGCATGCAAAATCTCAATCAAGCCTTCGATCGTCTG
                         ** ** *****  * ***** ** ***** ** ** ******** **

                R  Q  Y  L  P  C  L  G  N  D  R  Q  L  S  K  H  E  T  L  Q
D. melanogaster CGCCAGTACCTTCCCTGTCTGGGAAACGATCGCCAGCTGTCCAAACACGAGACCCTCCAA
D.virilis       CGTCAGTATCTGCCCTGTCTGGGCAACGATCGGCAACTGTCCAAGCATGAAACACTACAA
Calliphora      CGCCAATATCTTCCCTGCCTGGGCAATGATCGTCAATTGTCCAAACACGAAACCCTGCAG
Musca           CGTCAATATCTCCCCTGTCTGGGTAATGATCGCCAATTGTCCAAACACGAAACCCTGCAA
                ** ** ** ** ***** ***** ** ***** **  ******* ** ** ** ** **

                M  A  Q
D. melanogaster ATGGCCC---
D.virilis       ATGGCCCAAA
Calliphora      ATGGCCCAA-
Musca           ATGGCCCA--
                *******
```

Figure 2.2.2 *ato* bHLH regions from Dipteran species

Top line represents the most conserved portion of the Ato bHLH domain. This amino acid sequence is identical for all species examined. Below, the corresponding DNA sequences. Nucleotide differences are silent substitutions. Asterisks indicate conserved bases.

```
ato             AANARERRRMQNLNQAFDRLRQYLPCLGNDRQLSKHETLQMA
cato            AANARERKRMNGLNAAFERLREVVPAPSIDQKLSKFETLQMA
amos            AANARERRRMNSLNDAFDKLRDVVPSLGHDRRLSKYETLQMA
                *******:**:.** **::**: :*. . *::***.******
```

Figure 2.2.3 bHLH regions of *ato*-related genes in *D. melanogaster*

Alignment created using Clustalx. Asterisks mark identities, colons indicate conservative similarities, periods weaker similarities.

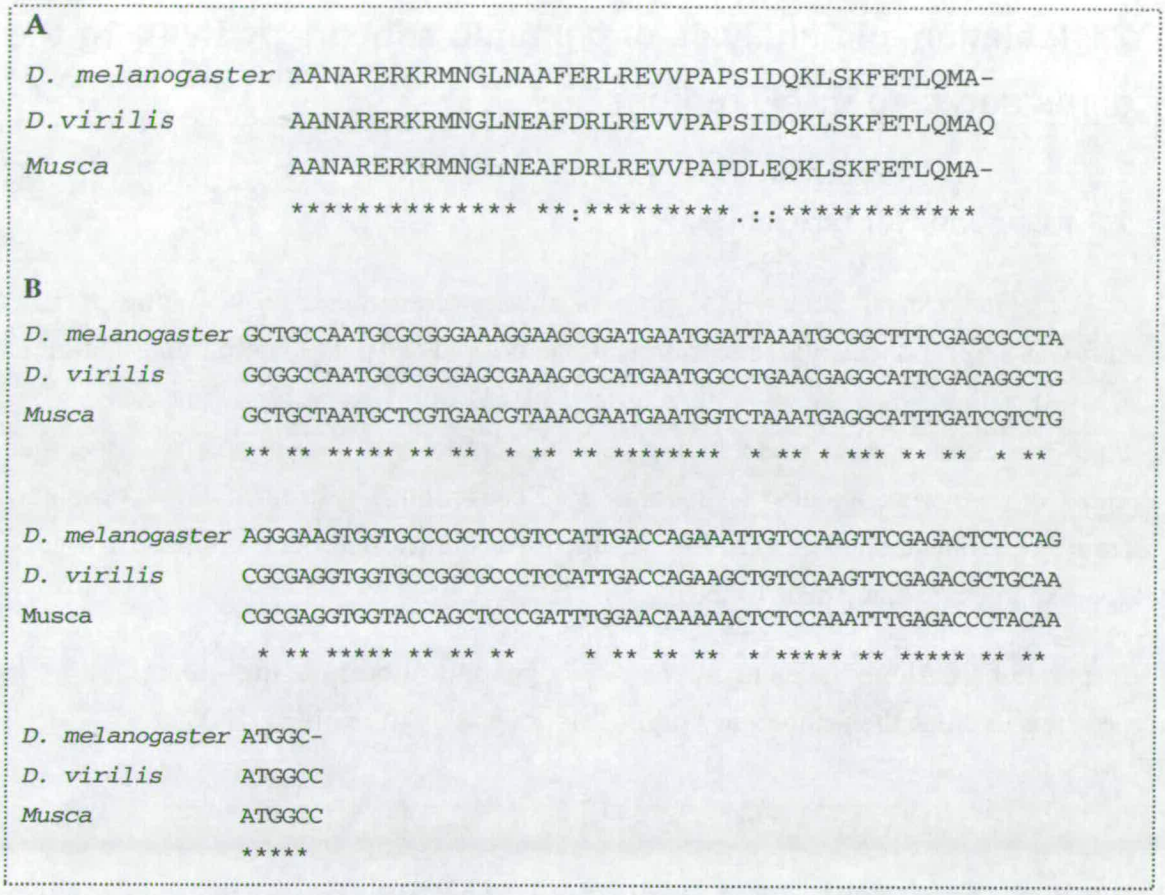


Figure 2.2.4 bHLH regions of *cato* homologues

(A) amino acid sequences of Cato bHLH fragments isolated from different species. **(B)** DNA sequences of the same domains.

Initial database searching using the cloned PCR bHLH fragments of *amos* and *cato* indicated that these represented new sequences, not yet reported by any genome or EST sequencing project. Therefore I used standard molecular methods to further characterise these genes at the genomic and transcript levels in *D. melanogaster*. Comparisons using the *ato* and *ato*-related bHLH fragments from other Dipteran species were not useful in that these regions are fully conserved across species for the different genes. However these fragments provided a starting point to isolate the complete coding sequences from *D. virilis* and thus allow further analysis of individual protein structure (I. Ahmed and A.P.J, unpublished).

2.3 Isolation and analysis of genomic sequences flanking the *amos* and *cato* bHLH regions

2.3.1 Genomic Library screen

A *D. melanogaster* lambda-FIXII genomic library (constructed by N. White, Jarman lab) was screened to isolate recombinant bacteriophage (phage) containing genomic DNA that flanked the newly characterised bHLH domains of *amos* and *cato*. After three successive rounds of screening several recombinant phages positive for either *amos* or *cato* were isolated (Appendix 2). Restriction digestion of DNA prepared from these phage showed that five independent clones had been obtained for each gene, of average insert size 15kb.

In parallel, If Ahmed screened a *D. virilis* genomic library using the cloned PCR products to clone the orthologues of *ato*, *amos* and *cato* (unpublished data).

2.3.2 Construction of a physical map for the genomic region surrounding each gene

Using southern blotting techniques, the recombinant phage clones were employed to establish restriction maps for the genomic region around each (Fig. 2.3.1) gene. These physical maps demonstrated that overlapping genomic clones had been isolated in both cases. Some phages were further analysed to produce a finer physical map of each region and to pinpoint useful genomic fragments for subcloning. Figure 2.3.1 depicts the restriction maps compiled for *amos* and *cato* respectively.

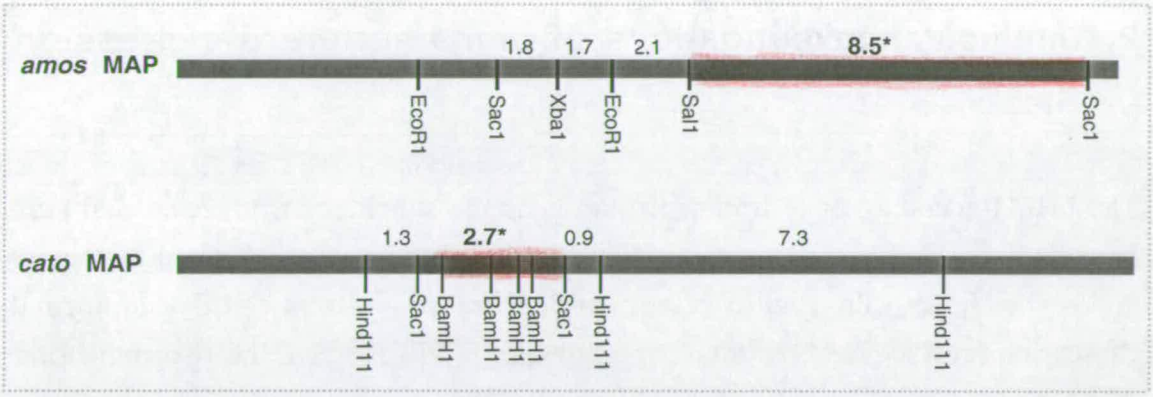


Fig 2.3.1 Physical map of the *amos* and *cato*-containing regions

Physical maps were constructed using restriction mapping data from lambda genomic clones. Red boxes and asterisks indicate the regions from which subclones were isolated. Black arrows mark the location of the open reading frames. Some subclones contained sites from the lambda vector arms. The vector arms are not indicated on this diagram.

2.3.4 Subcloning from the recombinant phage

Drosophila bHLH proteins are often encoded by small non-intron containing ORFs. Therefore I selected and isolated genomic fragments of sizes most amenable to molecular cloning and thought to encompass a significant portion (if not the entire coding region) of each gene. For *amos*, a 3.7kb SalI fragment of the phage AA1 was chosen for further analysis. In the case of *cato*, I isolated a 2.7kb and a 2.2kb SacI fragment from phage BE and BC respectively (Fig. 2.3.1). The fragments were subcloned into pBlueScript (pBS) vector. The resulting plasmids were named pBSSalI-AA, pBSSacI-BE and pBSSacI-BC and used as template for automated sequencing.

2.4 Molecular investigations of *amos* and *cato* expression profiles

The bHLH-encoding gene fragments and genomic subclones from *amos* and *cato* have been used to analyse the expression profiles of these genes. Several molecular approaches have been used to determine their transcript sizes and developmental expression profiles. Reverse transcriptase-PCR (RT-PCR) gave the first indication that *amos* and *cato* are normally transcribed during specific developmental stages. Northern blot analysis showed that each gene encodes a single embryonic transcript. 3'RACE (rapid amplification of cDNA ends) experiments yielded products of a size that indicated that, similar to *ato,* the bHLH domain lies in the carboxy terminal region of each gene. Based on these results, a suitable cDNA library was screened, first by PCR and then by hybridisation, to isolate cDNA clones representative of the *amos* and *cato* mRNAs.

2.4.1 Northern blot analysis

cato encodes a major embryonic transcript of ~750 nucleotides

Northern blots were prepared using stocks of polyA$^+$ and total RNA from staged embryonic extracts. To confirm that these stocks contained full length transcripts, one blot was hybridised with an *ato* probe. A single transcript was observed for *ato* of approximately correct size (Fig. 2.4.1). When other blots were hybridised with a *cato* probe a single band was observed in the 4-8hr polyA$^+$ RNA lane. A longer exposure revealed the same size band was also present in both polyA$^+$ and total RNA stocks of 0-4h embryos. Selection for polyA$^+$ RNA from total RNA can promote 5' truncation of RNA messages. However, in this instance both the polyA$^+$ and total RNA contain the same size *cato*-specific band indicating the presence of full length transcripts in the polyA$^+$ RNA. In addition, two sources of RNA were used, providing further evidence that the full length transcript was detected. Rf values were estimated for the ladder fragments and plotted against their sizes and from this graph (data not shown) the *cato* specific band was calculated to be between 750 to 800 nucleotides.

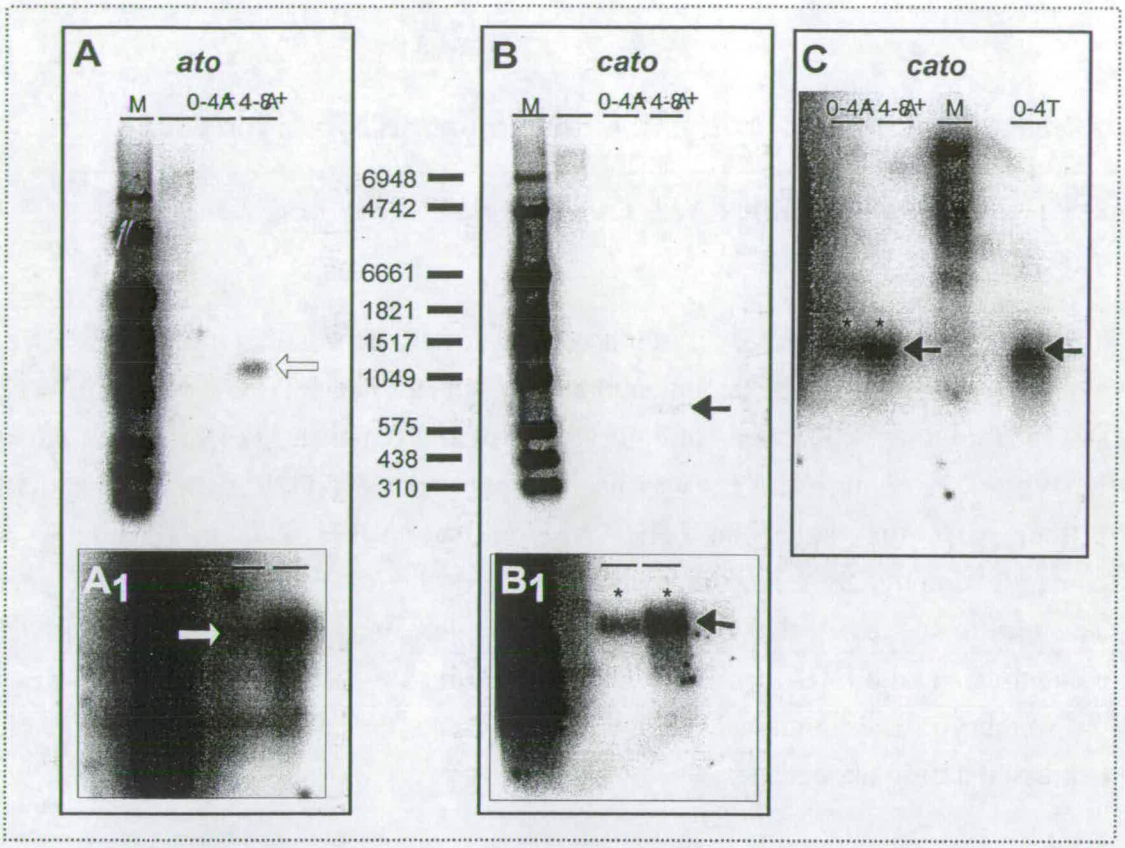


Figure 2.4.1 Northern analysis of *cato* transcripts

Northern blots (N-blot) were made containing 1µg each of polyadenylated (polyA$^+$) RNA from 0-4h and 4-8h embryo stocks and 20µg of a 0-4h embryo total RNA stock from an independent source (L. Thompson). A Digoxygenin-labelled RNA ladder (M) was included to provide a size standard. (A) N-blot hybridised with a probe to ato (white arrow). A1, longer exposure shows *ato* mRNA is present in 0-4h as well as 4-8h polyA$^+$ RNA stocks. A single transcript was seen of 1200 nt. (B) N-blot hybridised for the presence of *cato*. A single transcript was observed in the 4-8h RNA stock (black arrow). (B1), the same sized band was seen in 0-4h RNA after a longer exposure (*). (C) A third N-blot was run that included 0-4h total RNA. Although the marker was degraded, nevertheless a single band that corresponded to the same sized band from polyA+ RNA was seen for *cato*.

amos encodes a major embryonic transcript of ~800 nt

Similar northern blots were made for analysis of *amos* transcripts (Julie Ferguson, pers. comm.). From these results it was estimated that the *amos* mRNA was approximately 800 nt long, and was present in 0-4h RNA stocks.

2.4.2 RT-PCR and 3' RACE using developmentally-staged RNA

amos and cato are expressed during both larval and adult development

To confirm the timing of expression and to determine whether these genes were expressed during other developmental stages I performed RT-PCR with total RNA stocks from timed collections of embryo, larval and pupal stages (S.E.G., R. Kirby, N. White). A commercial kit was used to carry out RT-PCR with gene specific primer pairs that span the bHLH region (see Table 2.4.1). To check for contamination by DNA, negative controls without the RTase enzyme were included. The results suggested that both the 12-16h embryo and larval RNA stocks were contaminated with DNA. However, unequivocal results were obtained in the case of 0-4h embryo, 4-8h embryo and pupal RNA stocks that suggested both genes are expressed during these stages.

Table 2.4.1 RT-PCR results

Primer pairs A3 and A4 for *amos* and B3 and B4 for *cato* gave products of 260bp and 217bp, respectively. Products were analysed by agarose gel electrophoresis.

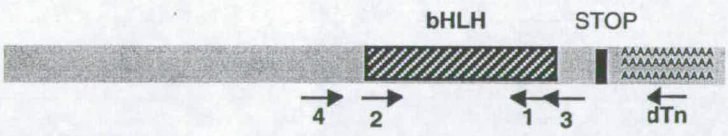
RNA stock	*amos*		*cato*	
	+RTase: +ve	-RTase: -ve	+RTase: +ve	-RTase: -ve
0-4h embryonic	+	-	+	-
4-8h embryonic	+	-	+	-
12-16h embryonic	+	+	+	+
Larval	+	+	+	+
Pupal	+	-	+	-

The amos and cato bHLH domains lie in their carboxy termini

3'RACE was subsequently carried out to amplify the 3' ends of the mRNA transcripts of *amos* and *cato* (Table 2.4.2). Using nested primers that extend 3' from the bHLH domain together with a poly dT primer (from polyadenyl tail), a PCR product of ~320bp was amplified from 0-4h embryonic total RNA for *amos*. In the case of *cato* a similarly sized product (340bp) was isolated from 4-8h embryonic RNA (both total and polyA$^+$). Given the bHLH domain itself is predicted to be at least 170 bp, these results indicate that the bHLH domains are quite 3'. The RACE products were cloned and sequenced.

Table 2.4.2 Summary of primers used and their predicted product sizes

Based on the RT-PCR results, RNA stocks prepared from early embryonic stages (N. White, S.E.G.) were used for first strand cDNA synthesis with an oligo dT (dT_n) primer (Promega 3'RACE kit). The cDNA stocks were used as template in nested PCR reactions using the primer pairs summarised below. The primers marked A are *amos* primers, and those marked 'C' are *cato* primers. For each procedure primer pairs are indicated in each row by ticks (√), and product sizes listed in the righthand column. First round PCR was performed on the synthesised cDNA with dT_n and the gene specific primer A4 or C4 (3'RACE1°). To increase the specificity of the reaction, a second round of PCR was carried out using the internal primers A2 or C2 in conjunction with the dT_n (3'RACE2°).

Procedure					Product Size (bp) *amos, cato*
	A4 or C4	A2 or C2	A3 or C3	dT_n	
RT-PCR	√	-	√	-	260, 217
3'RACE1°	√	-	-	√	
3'RACE2°		√	-	√	320, 340

2.4.3 cDNA library screen and clone isolation

cDNA libraries were screened to determine the structure of each transcript at the sequence level. In the first instance, PCR was used to check for the presence of *amos* and *cato* in 0-4h and 4-8h embryonic Lambda-ZAP cDNA libraries (constructed by N. White). Appropriate libraries were then screened for each gene to isolate cDNA clones. Three independent positives were isolated for *cato*, two of which (B1 and B2) were shown by PCR to contain the bHLH region of *cato*. After standard screening the number of *amos* cDNA clones was narrowed down to five (A1,A5,A6,A9 and A10). PCR on these isolates suggested that only A1,A6 and A9 contained the *amos* bHLH sequence (Julie Ferguson, pers. comm.).

The cDNA inserts were isolated by phagemid rescue as plasmids of the vector pBK-CMV. The *cato* clones, termed pBK-CMV-B1 and pBK-CMV-B2 contained inserts of approximately 750bp and 700bp respectively. The longest *amos* cDNA clone was 782bp long. These sizes agreed with those predicted by the N-blot analysis thus indicating that these genes do not contain introns.

2.5 Sequence analysis

2.5.1 Compilation of a sequence contig for *cato*

Two independently isolated subclones pBSSal1-BE and pBSSacI-BC were used for sequence analysis and a genomic consensus sequence of 1,476 bp was produced from these sequences. The directionally cloned cDNAs were also sequenced: pBKCMV-B1 and pBKCMV-B2 contained 681 bp and 749bp of *cato* sequence, respectively, excluding their polyadenyl tails. The cDNA and genomic contigs were amalgamated and comparison of these sequence data showed that no introns were present, thus *cato* is a single exon gene. The final consensus sequence of 1,476bp depicted in Fig 2.5.1 was deposited into the public databases (AC No. AF134869).

Open Reading Frame (ORF) prediction

When translated in the bHLH reading frame, the consensus sequence was found to contain two possible methionine start codons, termed M336 and M426. These give rise to ORFs of 656bp and 569 bp to the stop codon (Fig. 2.5.1), By comparison of the genomic and cDNA data the 3'UTR sequence downstream of the stop codon was calculated to be 101 bp and a polyadenylation was identified at position 1096 in the consensus sequence (see Fig. 2.5.1).

Both cDNA clones encode the second (M426), but not the upstream, start codon. Given that the longest cDNA clone is of similar size to that predicted from N-blot analysis, it seems that this likely represents the full length transcript. Furthermore, when compared with the *cato* genomic sequence from *D. virilis*, there was no significant homology upstream of the M426 site, thus the M336 start site is not conserved (If Ahmed and APJ, unpublished). Potential consensus translation initiation sites (Kozak sequences) were identified using this data for comparison, although the sites are not well conserved. Together these observations led to the conclusion that the M426 site was the correct initiation codon.

```
   1                                                        CTCCACCCTCGTGTGCCGTTGAG
  24 CGTGAAAGTCACAAATTATGGGCGACAGAGGGAGAGGGAGAGAGAGAGAGGGGAAGATCCCATGGCA
  91 TGATAGTACCAATTTGGCAAGATATGACGCGTGCCTGGCGCGAGAGAGATGACAACAGGCGAAGTGT
 158 AGGCGTTTCACCACCGAGCGAAAGGGAGGGAAACATATCTAACCCNTAAGTCACCCCAGNTCCTCGG
 225 GGAGATCCTTTGGCCGGGCCATGGTAATAGCTGCGAGGATCCTGGGGCCAGGGGTCATTTGCGGGCC
 292 ATTTACACAAACTTGGTCAGAACGGCGGTGAAAATATTTGTAGATGCCATCTCGCGGGGGTTGACCG
 359 TAGTCAAGTCTGGATCCGATCGGATCGTTTCAGTTGCAACGAAACTTTCAAGCCGCGCGGATCTGCA
     M   S   Y   Y   Y   S   S   A   S   E   E   D   G   S   S   Q   Y   17
 426 ATG AGC TAC TAC TAC TCG TCT GCC TCC GAG GAG GAT GGC AGT TCC CAG TAT
     L   G   S   P   N   Y   N   L   T   Q   L   P   P   V   S   G   Q   34
 477 CTG GGC AGT CCC AAC TAC AAC TTG ACC CAG TTG CCG CCA GTT TCT GGC CAG
     D   Y   G   Q   G   A   F   L   S   P   E   W   Q   F   L   D   A   51
 528 GAT TAC GGA CAG GGG GCT TTC TTA TCG CCG GAA TGG CAA TTC TTG GAT GCC
     A   G   G   T   Q   T   E   L   G   P   I   M   E   V   Q   G   Q   68
 579 GCA GGC GGA ACT CAA ACG GAA CTA GGA CCC ATA ATG GAG GTG CAA GGA CAG
     H   T   Q   P   Q   T   K   R   R   S   N   S   S   T   G   S   D   85
 630 CAC ACC CAG CCG CAG ACC AAA CGG CGG AGT AAC AGC TCC ACA GGA TCG GAC
     G   R   K   S   S   P   E   Q   T   N   L   S   P   T   V   Q   K  102
 681 GGT AGG AAG AGC AGT CCA GAG CAG ACC AAT CTC AGT CCC ACG GTC CAG AAG
     R   R   R   Q   A   A   N   A   R   E   R   K   R   M   N   G   L  119
 732 AGA AGA CGA CAG GCT GCC AAT GCG CGG GAA AGG AAG CGG ATG AAT GGA TTA
     N   A   A   F   E   R   L   R   E   V   V   P   A   P   S   I   D  136
 783 AAT GCG GCT TTC GAG CGC CTA AGG GAA GTG GTG CCC GCT CCG TCC ATT GAC
     Q   K   L   S   K   F   E   T   L   Q   M   A   Q   S   Y   I   L  153
 834 CAG AAA TTG TCC AAG TTC GAG ACT CTC CAG ATG GCC CAG TCC TAC ATC CTG
     A   L   C   D   L   L   N   N   G   D   V   E   V   D   A   A   A  170
 885 GCG CTG TGC GAT CTC CTC AAC AAC GGG GAC GTG GAA GTG GAT GCC GCT GCA
     Y   T   I   F   G   D   S   D   S   G   F   G   L   S   G   G   S  187
 936 TAC ACC ATC TTC GGC GAC AGC GAT AGT GGA TTT GGA TTG AGC GGA GGA TCC
     L   S   *                                                          189
 987 TTG TCA TAG ATGGATGATACTAGACTAAAGTTATGTGATTTTCTTCTTACTGTAGATTAAGTTA
1051 AATATGTAATGAAATAAATTGAAATGTTTAATTGAATGCTAAAAATAATTTCAATTTCAAAGGCATT
1118 ATTCATGGAACGCATCGCTTGATTGTGAACCAAGTGTTCAATCCAAAATGAGCTACTCATCTAACCC
1185 TTTCTACCAATTATAACCAGCCCACTACTATACACAGATCCCCAGATTCAATTGGCCATCAGATCGT
1252 TTGGCTGCGCCTCAATGATTTCGAAAACGATCATAAATTAATTACTCAACAGTTCATTGATAGGCGT
1319 ACTTCAAGTTTTGRTCTAACCCCGAACGATTCCGAAACTGCGAAGCCACCACAGTGGGCGGTGTAGC
1386 GTGACACCCCTGAACTCCTGGCCAGAACTCCTCTCATTGAATAAAAAAGGCAGCTGAGTCCTTCGGC
1453 TTCGGTTGGGTGCTGTTTTTTTTT
```

Figure 2.5.1. Sequence of the genomic region of *cato*

The DNA sequence contig of *cato* from genomic and cDNA data. Favoured ORF starts at position 426 in the DNA sequence, and ends at 996. The other possible start site is highlighted in green at position 336. The 189 amino acid translation is positioned above the DNA sequence. **M**, start methionine. (*) Stop codon. The bHLH domain is highlighted in blue. Putative consensus translation initiation sites are indicated in red. A polyadenylation signal sequence is underlined.

Comparison of consensus with genome project sequence data

The region around *cato* has now been sequenced by BDGP from both P1 and BAC (Bacterial Artificial Chromosomes) clones (see Appendix 3). The *cato* genomic sequence consensus and that of the 80,000 bp of P1 #DS06464 (AC004318) were compared and found to be 98% identical over the 1,476bp contig. Analysis with the P1 data revealed several mistakes in the genome project sequence. However there were also differences that could not be resolved as errors. These correlated with ambiguities in my own sequence data which gave rise to silent substitutions within the coding region. Thus it appears that polymorphism exists even at the level of individual lab strains.

Searches of the P1 and BAC sequences indicated that *cato* is the only bHLH protein in this region, thus does not form a complex with related genes. Access to the sequence around the *cato* gene will allow future analysis of the promoter and enhancer elements necessary for the regulation of *cato* transcription.

2.5.2 Compilation of a sequence contig for *amos*

A 1157 bp consensus sequence for the *amos* genomic region was generated from sequence of the genomic sub-clone pBS-AASal1, as well as cDNA clones of the region. cDNA clones were sequenced by J. Ferguson. The *amos* sequence was submitted to the public databases (AC. No. AF166113). When translated, the sequence consensus contained a single bHLH-containing ORF of 198 amino acids. Comparison with *D.virilis* data pinpointed consensus translation initiation sites and polyadenylation signals.

Comparison with genome sequence data

amos has been sequenced from several BAC clones by BDGP, and again is over 98% identical to the genomic consensus sequence (see Appendix 4). In addition, *amos* has been independently cloned by another lab, (Huang, M-L., Hsu, C-H., Chien, C-T., in press). Their sequence is very similar with two or three exceptions, perhaps due to differing lab strains.

```
   1                                                    GTCGACCTGTAGCTGATACAGTAAAT
  27 ACCTGGGAGAAAGAGGGAGAGAGTTGCAGCCAGAGACAATTAAACGTCTTGACCGCAAGCCAATTTA
  94 GGGCACGGGCTGTGCTGCTGATCCTCAAGAGGTTGGCAATCGGGTACCTGAGCGGATCGGATCAGCT
 161 TGAGGCAGCGAATCAGGTAACRGRACATATGTAGTGAACCAATCACGTGAGAGCTGAATCCTCGGCA
 228 GGCAGSTATATAAAGACCCATCCTCAAGCAGCCACTTCAGTTGTCCCTTGAGCCTGCAAACGTAAAC
      M   L   T   N   N   E   L   M   E   Q   F   Y   F   P   D   E   A      17
 295 ATG TTG ACC AAC AAC GAG CTA ATG GAG CAG TTC TAC TTC CCC GAC GAA GCC
      P   A   I   P   E   F   L   G   N   D   T   F   Q   Q   L   E   Q      34
 346 CCA GCG ATT CCC GAG TTC CTG GGC AAC GAC ACC TTC CAG CAG TTG GAG CAG
      L   M   Y   Q   Q   E   F   S   T   S   D   S   Q   S   D   G   A      51
 397 CTC ATG TAC CAG CAG GAG TTC AGC ACC AGC GAC AGC CAG TCG GAT GGC GCC
      N   S   C   S   L   E   M   Y   Y   D   T   P   S   V   L   E   L      68
 448 AAC AGT TGC TCC TTG GAG ATG TAT TAC GAT ACG CCG TCT GTC CTG GAA TTG
      E   H   M   L   N   A   Q   E   Q   Q   Q   H   H   L   Q   A   N      85
 499 GAG CAC ATG CTG AAT GCC CAG GAG CAG CAG CAG CAC CAC CTT CAA GCG AAT
      P   L   G   K   N   Q   G   R   S   P   R   Y   W   N   K   Q   Q     102
 550 CCC TTG GGC AAG AAT CAG GGC AGA AGT CCA AGG TAC TGG AAC AAG CAG CAG
      R   S   K   P   Y   D   K   L   S   T   S   M   S   S   S   T   S     119
 601 AGG AGC AAG CCA TAC GAC AAG CTG TCC ACT TCC ATG TCA TCA TCT ACA TCC
      S   A   S   S   S   S   S   S   S   A   G   F   G   G   E   V   L     136
 652 TCC GCC TCT TCG AGC AGT TCA TCG TCC GCG GGA TTC GGT GGC GAA GTC CTC
      K   K   R   R   L   A   A   N   A   R   E   R   R   R   M   N   S     153
 703 AAA AAA CGG CGA CTG GCC GCC AAT GCT CGG GAA CGG AGG CGG ATG AAC AGC
      L   N   D   A   F   D   K   L   R   D   V   V   P   S   L   G   H     170
 754 CTG AAC GAT GCC TTC GAC AAG TTG AGA GAT GTG GTT CCA TCA CTC GGC CAC
      D   R   R   L   S   K   Y   E   T   L   Q   M   A   Q   A   Y   I     187
 805 GAT CGG CGA CTC TCC AAA TAC GAA ACT CTG CAA ATG GCG CAA GCA TAC ATC
      G   D   L   V   T   L   L   S   R   D   Y   *                         198
 856 GGG GAT CTG GTC ACG TTG CTG TCC AGA GAC TAC TAG CCAGTGTGGGCGATCCTTT
 911 ATCCTTTCTTCCTCAAATGGAAGTTCCTTTTGCGGGCTGTGTTGCAGCAACACCTTCCATATCCTAG
 978 TGGAAATCTTATAAAGGCTGTTAGTTTTACGTTTATTATCATANTTGTACNCAATTCAAGCAATAGT
1045 TTTATAATAAAAATGAATACAAAATATCAATTATATTGTTTTAAATTCATATCGTATGAAATGGTGG
1112 AGTNGGAATGAAAATNATGTTATGCGAACTTGGGAAATTTATATTT
```

Figure 2.5.2. Sequence of the genomic region of *amos*

The DNA sequence contig of *amos* from genomic and cDNA data. ORF starts at position 295 in the DNA sequence, and ends at 891. The 198 amino acid translation is positioned above the DNA sequence. M, start methionine. (*) Stop codon. The bHLH domain is highlighted in blue. Putative consensus translation initiation sites are indicated in red. A putative polyadenylation signal is underlined in the DNA sequence.

2.6 Protein structure analysis

2.6.1 *amos* and *cato* encode *ato*-like bHLH proteins

The predicted protein sequences of Amos and Ato are 74% identical over the entire bHLH region and share an identical basic domain except for an R to K conservative change (Fig. 2.6.1B). This compares with ~70% identity between bHLH domains of the AS-C, and ~40% between Amos and Scute. The bHLH domain of the conceptual Cato protein is 64% identical to Ato, 66% identical to Amos and 41% identical to Sc. This compares with 42% identity between Ato and Sc, and ~70% identity between members of the AS-C. Among Ato-like bHLH proteins, both Amos and Cato form part of the Ato subfamily along with its closest vertebrate homologues, and are closer to Ato than to the neuroD or neurogenin subfamilies (Fig. 1B). In fact, Amos and Ato appear slightly more closely related to the identified vertebrate Ato-like homologues. Interestingly, this suggests that an *ato–amos* gene duplication occurred after the invertebrate-vertebrate split, so that vertebrate *ato*-like genes may exhibit functional similarity with either *amos* or *ato*.

2.6.2 Amos and Cato retain residues necessary for function as positive transcription regulators of neurogenesis

As might be expected from sequence identity with Ato, both Amos and Cato have the conserved residues important for the three-dimensional structure of the bHLH domain and interaction with DNA (Ellenberger et al., 1994; Littlewood and Evan, 1998; Ma et al., 1994) (Fig. 2.6.1A). Furthermore, all residues required for E-box binding, both in basic, helix-2 N-terminus, and loop regions, are found in Cato and Amos, suggesting that these proteins function as positive regulators of transcription. Such conservation leads to the prediction that the proteins will function as a heterodimers with Da protein. Indeed, Amos has been shown to bind to DNA as a heterodimer with Da *in vitro* (Huang, M-L., Hsu, C-H., Chien, C-T., in press).

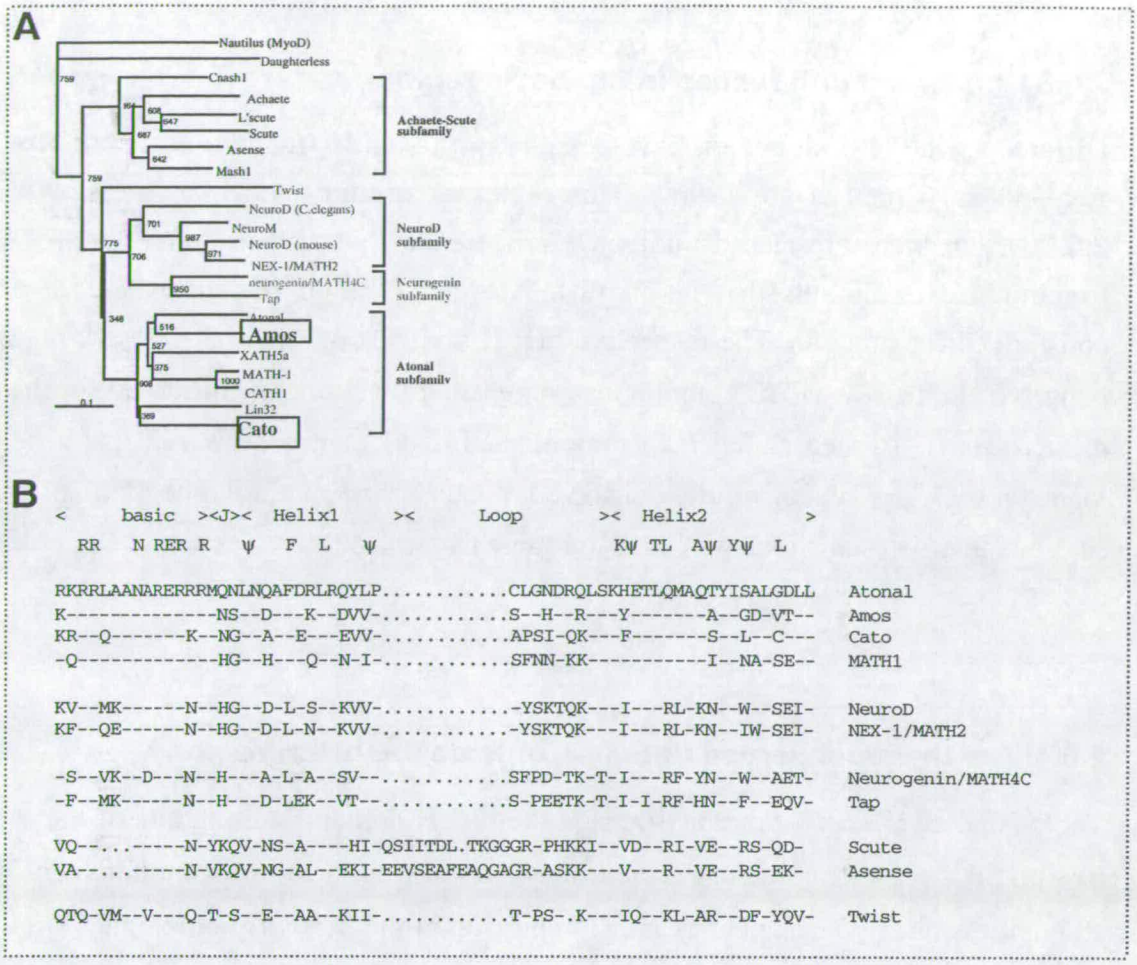


Figure 2.6.1 Relationships among the bHLH protein family

(A) Phylogenetic analysis of bHLH relationships for Amos and selected proteins. Bootstraps values for 1000 runs are indicated (A.P.J., pers. comm.). (B) Sequence comparisons of bHLH domains from Atonal and selected proteins. Dashes represent identities and dots represent gaps. The regions recognised by the degenerate primers are underlined. Consensus residues important for bHLH structure are shown above the sequences

2.6.3 Conserved differences in the basic regions

Differences in the basic region have been shown to confer functional specificity to Ato and Sc (Chien et al., 1996). This region is strongly Ato-like in Cato (Fig. 2.6.1B), but there are also a number of differences, which may either represent functional determinants that distinguish Ato and Cato or residues that are not constrained for function. The respective bHLH domains of Ato and Cato are highly conserved in their *D. virilis* homologues, suggesting functional significance for these differences (I. Ahmed and A.P.J., unpublished data). Amos however, has a near identical basic domain, also fully conserved in the *D. virilis* homologue. Perhaps this suggests that Ato and Amos will have the same target genes.

2.6.4 Are there conserved domains outside the bHLH region?

There is no significant homology outside the bHLH domains among any of the *ato*-like bHLH genes. This indicates then, that the gene-duplication event between *ato* and *amos* is clearly an ancient event. Comparison of the *D.virilis* homologues shows that there are regions of conservation in the orthologues, but these do not represent characterised functional domains (A.P.J., pers. comm.)

2.7 Mapping of *amos* and *cato*

Several methods were used to determine the precise cytological location of both genes. Preliminary mapping was carried out by classical *in situ* hybridisation to polytene chromosomes. These positions were used to search the stock centres and databases (such as Flybase and BDGP) for deficiency (Df) stocks that spanned the region. In addition a P1 gridded array was screened to determine which genomic clones contained either gene. Most recently map locations have been confirmed by reference to the published BAC and sequence scaffold maps (BDGP). In summary, both genes have been mapped to the second chromosome, *amos* to the left arm at position 36F2-6 and *cato* to the right arm at 53A3-5.

2.7.1 *cato* maps to 53A3-A5

A 15 kb DNA probe of a *cato* genomic phage (BF) was used in *in situ* hybridisation to prepared polytene chromosome squashes. By this method *cato* maps to the right arm of chromosome 2R, at 53A (Figure 2.7.1). I examined the relevant databases for all stocks containing aberrations or mutations in or near this region (Figure 2.7.2). There did not appear to be any known specific mutations of this region for which *cato* could be a candidate.

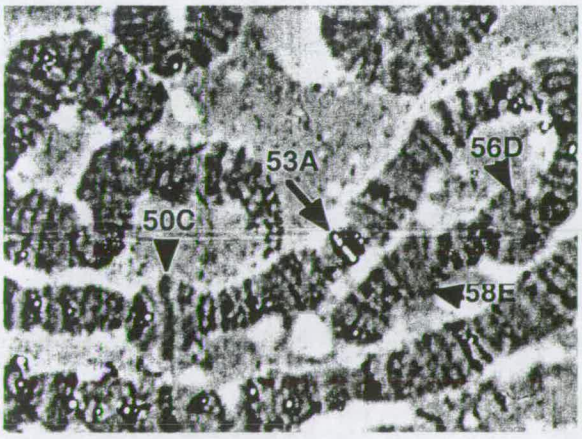


Figure 2.7.1 *cato* maps to 53A

Salivary gland polytene chromosome *in situ* hybridisation using *cato* DNA probe. Landmark puffs are indicated by arrowheads. A band of hybridisation was observed at the region 53A, indicated by the black arrow.

Though the cytological mapping gives a reasonable estimate of the gene location, a finer map, which is also useful genetically, can often be produced by screening deficiency (Df) stocks that span the region. Several Df stocks spanned the position of 53A and these were obtained to investigate which ones uncovered *cato*. In particular the *Df(2R)'Jp'/CyO* series were most useful in refining the position of *cato*. These stocks were generated in a screen for alleles of the *kinesin heavy chain* (*khc*) locus (Saxton et al., 1991) which maps close to *cato*. By this time, an embryonic expression pattern for *cato* had been established. Thus, where appropriate, embryos of Df stocks were examined by *in situ* hybridisation to mRNA, the assumption being that embryos homozygously deficient for any gene will not show expression of it (see Chapter 3). If the proportion of unstained embryos exceeded 25%, then that Df stock was described as 'uncovering' *cato*. Crosses between Df stocks were also performed to create smaller deletions. The smallest region thought to uncover *cato* was estimated to be between 52F5-9 to 53A3-A5, by crossing *Df(2R)Jp7* to *Df(2R)Jp8*. Figure 2.7.2 summarises the results of these experiments for *cato*.

A P1 gridded array was screened using a *cato* specific probe to determine which P1 phage contained genomic DNA from the *cato* locus. Three clones were identified as containing *cato* sequence and all have been mapped cytologically to the region 53A1-A5 (DS00517, DS03927, DS06464). The full sequence of DS06464 is known and analysis using this data indicated that *cato* was 7kb away from the *khc* locus. The direction of the clone is unknown.

Thus *cato* was successfully mapped to 53A3-A5 by three independent methods.

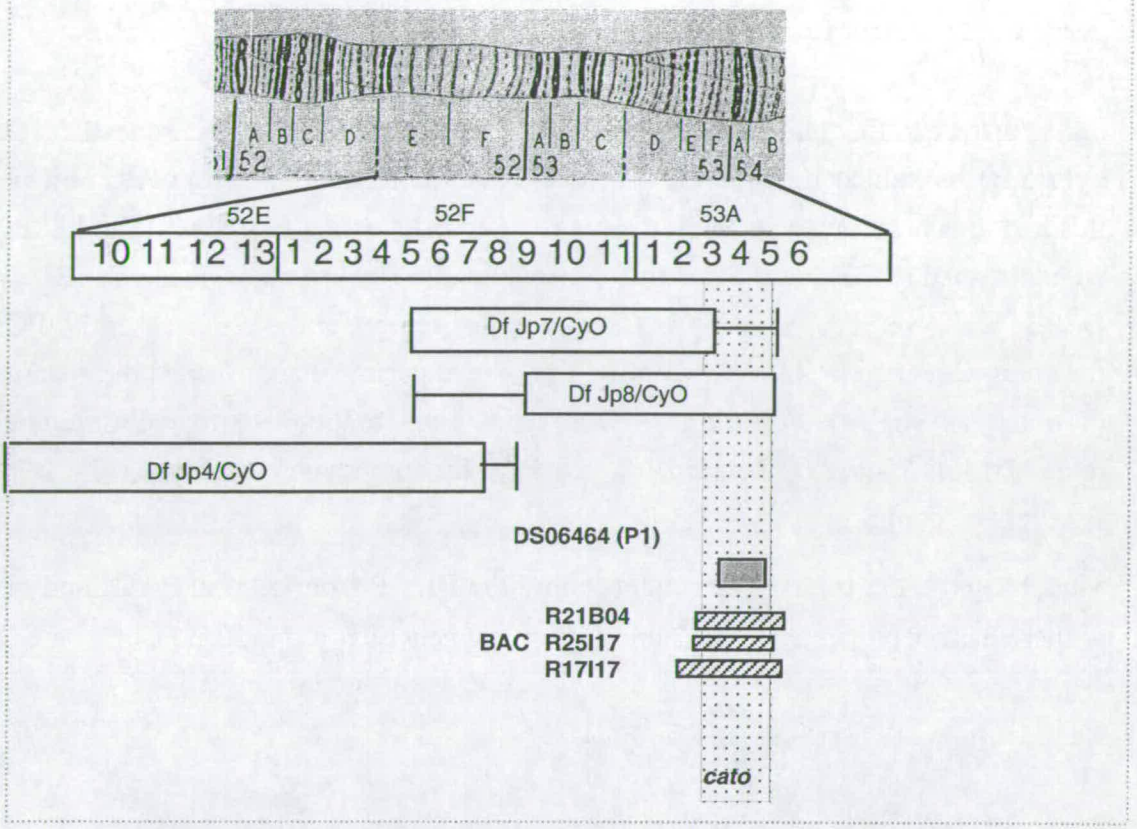


Figure 2.7.2 Mapping of *cato*

Polytene segment and schematic showing location of *cato* by deficiency mapping, and reference to P1 and BAC sequences (the sequence boxes are not to scale)

2.7.2 *amos* maps to 36F2-F6

amos proved rather more difficult to map. Polytene chromosome squash *in situ* hybridisation indicated that *amos* was located on the centromeric end of 2L however this region is notoriously difficult to analyse. Separate estimates of 36C-D and 38E-F were obtained (SEG) but the position was eventually finalised by this method as 36F (A.P.J., pers. comm.). Several YACs map to this region, and were checked by PCR for *amos* sequence (P.z.L., pers. comm.). These experiments confirmed the location of *amos* as 36F2-6. Furthermore, *amos* has been mapped to two deficiencies, (*Df(2L)M36F-S5 and Df(2L)M36F-S6*) by PCR on homozygous embryos (Fig. 2.7.3, P.z.L., pers. comm.)

Most recently, *amos* has been sequenced by the BDGP from several BACs and one of these has been mapped to the same region cytogenetically (Fig. 2.7.3).

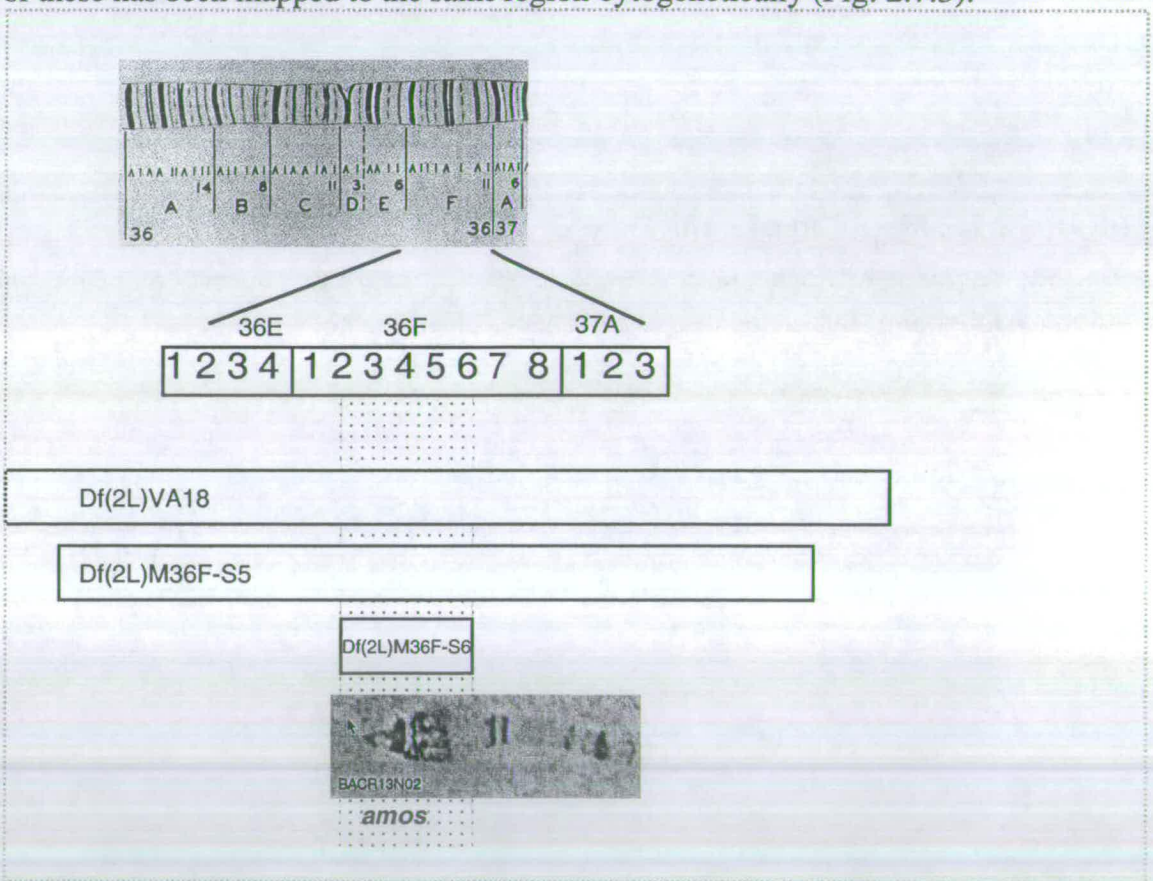


Figure 2.7.2 Mapping of *amos*

Polytene segment and schematic showing location of *cato* by deficiency mapping, and reference to BAC in situ, known to contain *amos* (the sequence boxes are not to scale)

2.8 Discussion

2.8.1 *amos* and *cato* encode new bHLH genes with closest similarity to *ato*

amos and *cato* are new genes with similarity to *ato* that map to independent chromosomal locations. Thus, unlike the AS-C, these closely related genes do not form a gene complex with *ato*. Much like other bHLH family members, their sequence homology is restricted to the bHLH domain and diverges completely outside this region. However both *amos* and *cato* have been conserved in evolution, and are present in the related species *D. virilis*. Both genes likely encode positive regulators of transcription.

2.8.2 Are *amos* and *cato* functionally related to *ato*?

Given that both genes fall within the neural sub-family of bHLH proteins, and their transcripts are expressed at times consistent with the onset of neurogenesis it is not unreasonable to propose that both *amos* and *cato* function during this process. Perhaps, given their relationship to *ato*, these genes form part of a functionally interlinked bHLH cascade, reminiscent of the expression cascades observed for similar vertebrate genes.

It is difficult to tell at this stage whether the complete conservation of the Amos and Ato basic domains has functional relevance, for example both proteins are thus likely to bind similar E-box sequences. The differences present in the bHLH domains for both *amos* and *cato* are conserved in their *D. virilis* orthologues, suggesting that these residues have functional significance. It may be that while each gene has unique functions, like other bHLH proteins, they share some functional features by virtue of their sequence homology. Such information may prove useful in future experiments designed to identify specific co-factors for these proteins. These elusive factors are hypothesised to interact with specific bHLH proteins to modulate their sub-type specificity (Jarman and Ahmed, 1998; Jarman et al., 1993b), potentially via the basic domain.

2.8.3 Are there other bHLH regulators in *Drosophila*?

These genes were isolated on the basis of their similarity with the Ato bHLH domain, under reasonably stringent conditions, but with relative ease. Given then, that this was not an exhaustive search for *ato*-related genes, it could be predicted that other, less related, bHLH genes are present in *Drosophila*. In fact, database searching with the *ato* bHLH domain identified an novel gene called *atoid* (APJ, unpublished). *atoid* would not have been picked up by this PCR procedure because of weak conservation in the primer regions. Not much is known about the gene except that it is expressed during neurogenesis (unpublished data). With the imminent completion of *Drosophila* genome sequencing it should soon be possible to address this question comprehensively and easily.

Chapter 3

Expression and Functional analysis of *cato*

3.1 Introduction

Taken together, the molecular and sequence data presented in the last chapter strongly suggest that *amos* and *cato* encode neural bHLH positive regulators of transcription. To address whether these genes did indeed have roles during neurogenesis, I carried out expression and functional studies for each gene. In this chapter I present an analysis of *cato* expression and function, and in the next, a similar description of results for *amos*.

Based on the assumption that these genes would be expressed during neurogenesis, embryos and imaginal disc tissues were analysed for mRNA expression by *in situ* hybridisation. Individual mRNA expression patterns were determined for each gene, as well as their localisation with respect to other neural markers in double labelling experiments (see Chapter 1 for details). Armed with such knowledge it is possible to predict, more precisely, potential functions for the genes, and aid interpretation of any functional data. Such information is invaluable to designing any appropriate screening strategy to generate specific mutations.

However, since specific point mutations are not yet available for either *amos* or *cato*, functional analysis is limited. Nevertheless, several approaches have been used to address if, and how, the *amos* and *cato* gene products act during development. Traditionally, phenotype analysis of appropriate deficiency stocks (lethal chromosomal deletions of specific genomic regions) gives a first impression of possible gene function, the caveat being that many genes are deleted in such stocks. More recently, the technique of RNAi has been applied to *Drosophila* (Kennerdell and Carthew, 1998; Misquitta and Paterson, 1999), although this appears to be restricted to larval development. Even so, phenotypic analysis requires some prediction in terms of gene function, usually based on gene expression patterns.

Misexpression studies using the Gal4/UAS system (Brand and Perrimon, 1993) have been employed as another approach used to assess the functional specificities of these new genes. Such assays are especially useful in the characterisation of functional properties encoded by bHLH genes (see Chapter 1).

A battery of markers are available with which to examine neural phenotpyes. For instance, Cpo, Elav and Pros are cell indentity markers, each specific for different

subsets of sense organs cells (see text, Chapter 1). There are also sense organ-subtype-specific markers such as Cut for external sense organs (Blochlinger et al., 1990) and Mab1188, an antibody that marks the scolopale of the chordotonal organs (Jarman et al., 1993). Perhaps the most widely used marker is Mab22C10, which is a marker of neuronal morphology, staining the neuron cell bodies and their processes (Zipursky et al., 1984).

Analysis of the loss of function phenotypes for *cato* (and *amos*) have been carried out with reference to such markers, and is described here for *cato*, and in Chapter 4 for *amos*.

3.2 *cato* mRNA expression

Digoxygenin-labelled RNA probes were made from the cloned PCR products and, once isolated, the cDNA and genomic subclones. In general cDNA-derived probes gave the strongest and cleanest results.

Despite its close sequence relationship with *ato* and *amos*, *cato* is not expressed as a proneural gene. Instead, *cato* is expressed specifically and widely during both larval and adult PNS development in a manner that is reminiscent of the pan-neural genes. *cato* mRNA is confined to the established SOPs themselves, and their resulting daughter cells, the SOCs. *cato* is further distinguished from other classes of neural bHLH factors in that its expression is restricted to the PNS, or more specifically still to sensory organ development. Thus, unlike the global neural expression of bHLH proteins such as *ase* (Brand et al., 1993; Dominguez and Campuzano, 1993) and *dpn* (Bier et al., 1992), *cato* can be termed a 'pan-PNS' gene.

3.2.1 *cato* mRNA expression in the embryo

cato is expressed in a dynamic pattern, initiating in SOPs after their formation, and remaining for a time in their division products (SOCs). In *in situ* hybridisation to wholemount embryos, mRNA was first detectable during stage 10 when the first PNS SOPs are formed (Fig. 3.2.1A,E-G), namely the A, P and the presumed dbd precursor (referred to here as the M cell) (Campos-Ortega and Hartenstein, 1998; Ghysen and O'Kane, 1989). The identity of these *cato*-expressing cells as SOPs was confirmed by double labelling experiments using antibodies to Ase or Ato in conjunction with a *cato* RNA probe (Fig 3.2.1F-H,K).

cato expression is activated earliest in the chordotonal precursors

Early *cato* expression in stages 10/11 strongly marks the appearance of the chordotonal SOPs and this pattern is strikingly similar to the later expression of *ato* as it becomes refined from the proneural clusters to SOPs (Fig. 3.2.1B). Indeed, in embryos stained for both *cato* mRNA and Ato protein it is clear that the expression of these genes overlaps in the chordotonal SOPs as they are formed (Fig. 3.2.1H). *ato* is then switched off whereas *cato* remains expressed in the SOP and resulting SOCs.

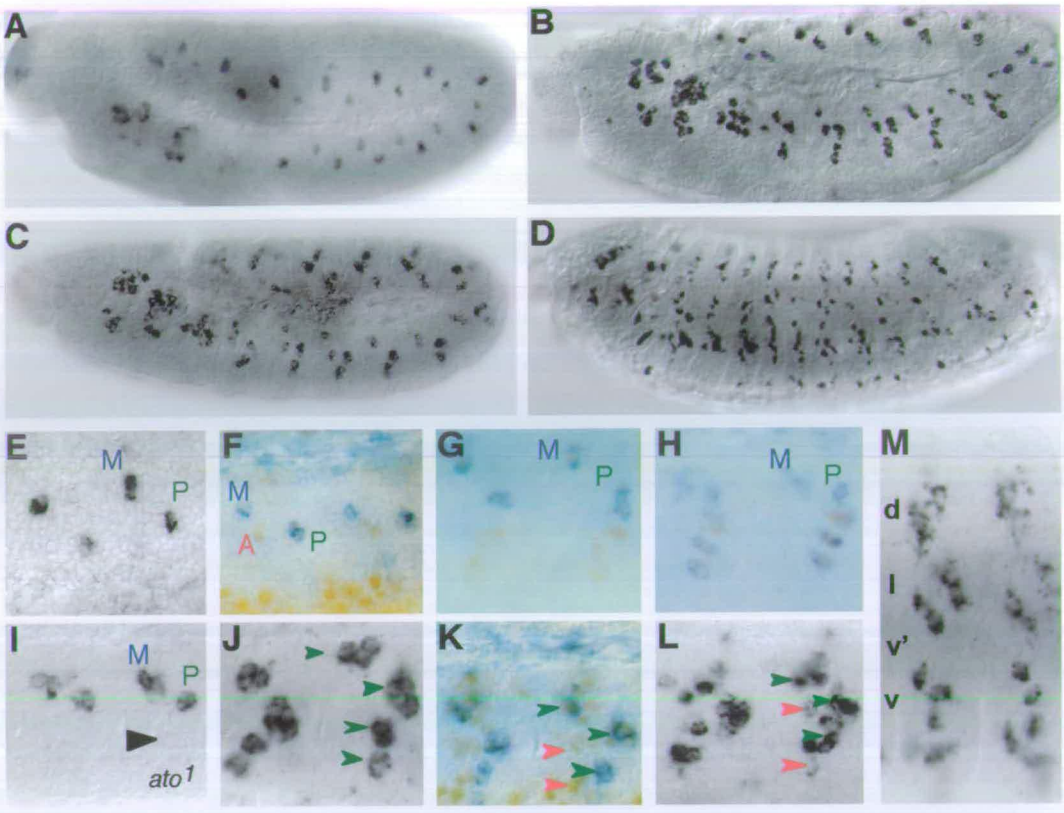


Figure 3.2.1 mRNA expression of *cato* in the embryo

(A-H, J-M) Wildtpye, (I) *ato*^{1}. (E-M) Close-up of two abdominal segments. RNA is black or purple, protein is brown. (A) Stage 10, showing expression in isolated sub-epidermal cells in the head regions and body segments. (B) Early stage 11, each body segment expresses *cato* in 5-8 cells and further expression is present in the head regions. (C) Later stage 11, *cato* is expressed in the division products (SOCs). (D) By stage 12 the complex pattern resembles that of *ase* in the PNS. (E,F,G) Stage 10 embryo. *cato* stains only two of the early precursors, the M and P cells; no expression is detected in the other early SOP (the A cell), marked by the presence of Ase protein (F). (G) *cato* mRNA overlaps Ato protein in the M and P cells. Ato does not stain the A cell. (H) By later stage eleven *cato* overlaps with Ato in all the chordotonal SOPs. (I) *ato*^{1} embryo, this stage is comparable with H and some *cato* expression is missing, corresponding to the missing chordotonal SOPs. mRNA is still present in the M and P cells. The A cell would be visible in this case, if stained with an appropriate marker. (J) Late stage 11. *cato* expression continues in chordotonal precursor division products. (K) *cato* is delayed in external sensory SOPs, marked by Ase, relative to chordotonal SOP division. Green arrows indicate chordotonal SOPS, red arrows show external SOPs (L) Slightly later, these external SOPS begin to express *cato*. (M) Late stage 12, the expression of *cato* foreshadows the final PNS arrangement into four clusters, dorsal (d), lateral (l), ventral prime (v') and ventral (v).

cato expression is delayed in other PNS cells relative to SOP formation

In contrast to observations for the *ato*-expressing SOPs, other types of SOPs have a temporally different profile of *cato* expression. Co-labelling of *cato* mRNA with antibodies either to Cut (Blochlinger et al., 1990), which is restricted to external sense organ precursors and their division products (not shown), or Ase (Brand et al., 1993), indicated that *cato* expression is delayed in such SOPs, relative to their formation (Fig. 3.2.1F,K). It seems that the difference is a temporal consequence, rather than a quantitative difference in *cato* expression levels. This is most clearly seen when the first external SOP (the A cell) and chordotonal SOP (the P cell) are formed in the body segments during stage 10. Although these cells are formed at the similar times, *cato* is initially only expressed in the P cell, appearing later in the A cell, during stage 11. Indeed this delay of *cato* expression in non-chordotonal SOPs is observed in other late-forming SOPs. The temporal difference in *cato* expression between different subtypes of sense organ precursors suggests that there may be separable functions for *cato*.

cato expression in ato mutant embryos

The early *cato* mRNA pattern suggests that there may be an early expression of *cato* that is *ato* dependent. With the exception of one precursor, the P cell, no chordotonal SOPs form in *ato* mutant embryos (Jarman et al. 1995). In addition, the M cell is also formed. I observed that both SOPs retain *cato* expression in ato^{1} embryos (Fig.3.2.1I). Thus *ato* is not necessary for *cato* activation within these cells.

Later cato expression is pan-neural

cato expression continues in the SOP progeny, the SOC cells, such that by late stage 11 and early 12 the pattern closely resembles the pan-neural expression of *ase* (Fig. 3.2.1D). During these stages *cato* expression is too complex to discern the identity of individual cells. However as germ band retraction proceeds, *cato* becomes restricted to a subset of precursors and the pattern foreshadows the final arrangement of the PNS organs into four clusters (Fig. 3.2.1M). Expression is extinguished by stage 13 when only a few cells per segment, which correlate with the position of lch5, still stain for *cato*.

cato becomes restricted to a single SOC, possibly the neuron

By stage 12 *cato* mRNA is restricted to one of the SOCs before being finally switched off. To discern which cell this corresponds to, double-labelling experiments were performed using a *cato* RNA probe and antibodies to either Cut, or Pros. The cells that co-express Cut and *cato* seem to correspond to the inner cells of the sense organ, one of which is the neuron (Fig. 3.2.2A). The inner cells are distiguishable by a weaker staining for cut than the corresponding outer support cells. Similarly, in the lch5 SOCs, Pros protein and *cato* mRNA do not overlap during later stages, and the *cato* RNA appears ventral to the pros-positive scolopale cell, again suggesting that *cato* is restricted to the neuron (Fig. 3.2.2B).

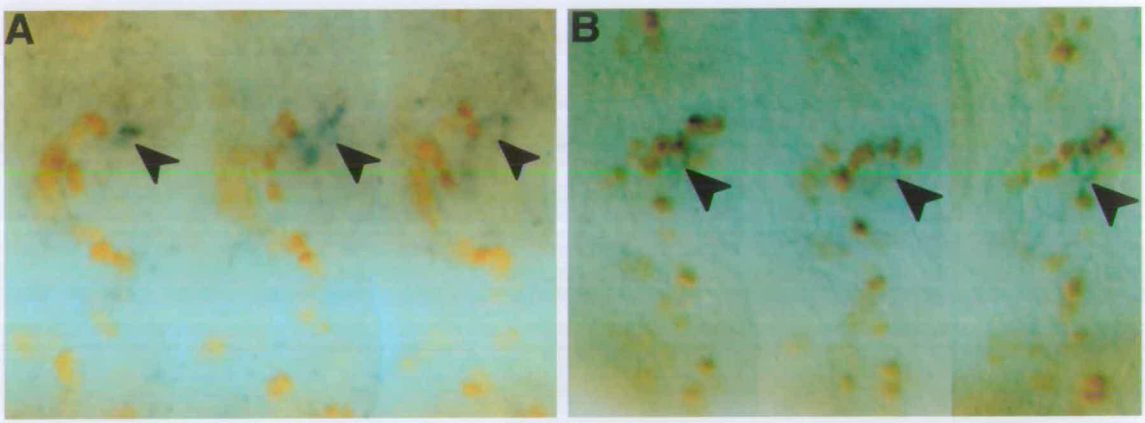


Figure 3.2.2 *cato* appears to localise to the sensory neuron precursor

Abdominal segments of wildtype stage 13 embryos. Purple marks mRNA, orange denotes protein (A) Ventral external sense organs, labelled for Cut and *cato* expression. *cato* mRNA is assciated with the weaker expressing cut-positive cells (arrowhead), one of which corresponds to the neuron. (B) the lch5, labelled for Pros and *cato* expression. Pros localises to the scolopale (sheath) cell of each sense organ. Residual *cato* expression is seen in the cells ventral to the scolopale, likely the neurons (arrowhead).

3.2.3 *cato* mRNA expression during adult development

There are parallels between the expression of *cato* during larval and adult development. In imaginal disc tissue *cato* expression is again PNS specific, and confined to the SOPs and their progeny. Overall, the *cato* mRNA pattern resembles that of the SOP-specific enhancer trap line A101 (Huang et al., 1991), but mRNA is detected in some subsets of SOPs before being switched on in others, reminiscent of the two-phase embryonic pattern. In particular, *cato* expression is markedly delayed in the precursors of the wing margin sensory bristles which are marked by A101 and express *ase* in third instar larval imaginal discs (Brand et al., 1993). In contrast, *cato* is not expressed in these precursors until much later in their development, after puparium formation (Fig. 3.2.3A-C).

Like in the embryo, *cato* mRNA appears in the chordotonal SOPs soon after their formation. In the antenna this appears to be true for the *ato*-dependent olfactory SOPs as well: *cato* mRNA is detected in these SOPs prior to external SOPs of the same age (Fig. 3.2.3D). Thus, it seems that for all *ato*-expressing SOPs, *cato* expression follows soon after the SOP formation, likely overlapping with *ato* itself.

Significantly, *cato* is not expressed in developing photoreceptors in the eye disc, a feature which sets this gene apart from its pan-neural cousins, and is in contrast to the overlap of *cato* and *ato* expression in most other domains (Fig. 3.2.3D).

Much like *ase*, *cato* mRNA persists through at least one round of cell division in the case of bristle precursors although this does not seem to be the case for all SOPs. For instance, in the precursors that give rise to the large femoral chordotonal array, *cato* expression is far more transient than elsewhere (Fig. 3.2.3E,F). In these SOPs *cato* expression is tightly linked with the status of *ato*: it appears that *cato* is quickly switched off as Ato is abruptly down regulated in these precursors specifically (zur Lage and Jarman, 1999). By contrast, *ase* is expressed continually, throughout the development and division of these SOPs (Brand et al., 1993).

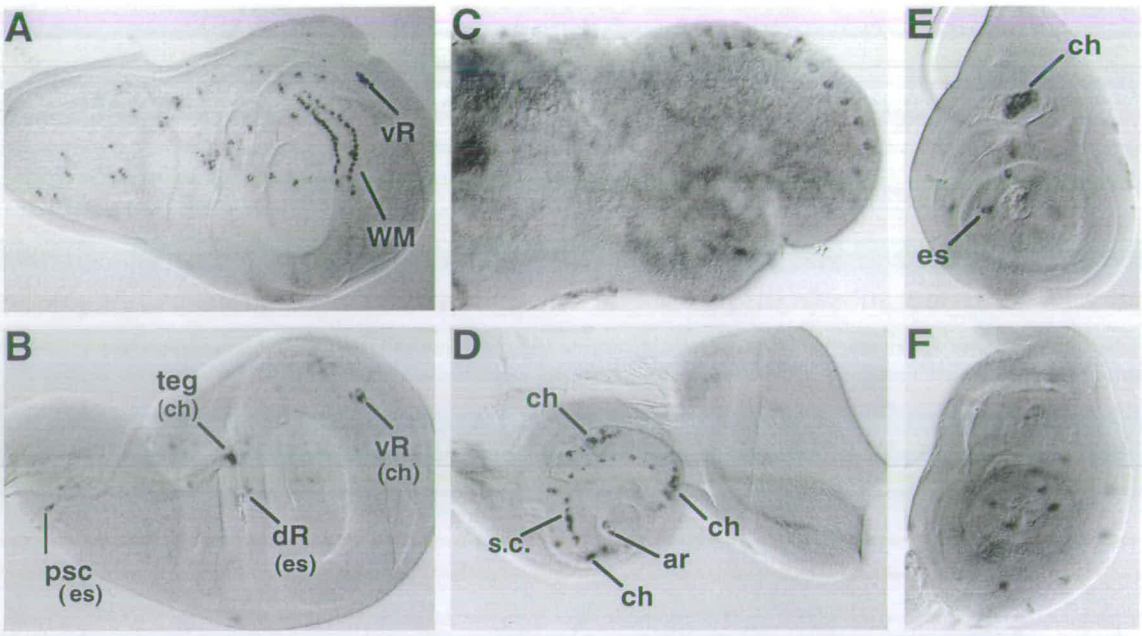


Figure 3.2.3 *cato* mRNA in wildtype 3rd instar larval imaginal discs

(A) A101 enhancer trap line, wing disc stained with antibody to ß-galactosidase. All SOPs are stained. (B-F) Wildtype imaginal discs stained for cato mRNA. (B) Wing disc, showing *cato* mRNA in both external sensory and chordotonal SOPs, and some of their division products. Notably *cato* is absent from the wing margin (WM) external SOPs at this stage. (C) A pupal wing disc shows *cato* staining in the late-forming wing margin precursors. (D) eye-antennal disc, *cato* is expressed in the SOPs of the Johnston's organ (ch), and some olfactory sensilla (s.c.) (all of which depend on *ato* function). No expression is detectable in photoreceptors of the eye disc. (E,F) Leg discs. (E) *cato* is transiently expressed in the cluster of SOPs of the femoral chordotonal organ, as well as isolated external sensory SOPs in other segments. (F) Slightly later, *cato* is switched off in the chordotonal cells and is only present in the external sensory precursors. s.c. sensilla coeloconica, ar arista, teg tegulla, psc posterior scutellar, dR dorsal radius, vR ventral radius.

3.3 *cato* is involved during sense organ differentiation

3.3.1 Embryos deficient for *cato* have neural defects

Embryos of two lethal chromosomal deficiency stocks that delete *cato*, *Df(2R)Jp7/CyO* and *Df(2R)Jp8/CyO*, were examined for PNS defects using neural cell identity and morphology markers. For the most part, transheterozygous embryos (*Df(2R)Jp7/ Df(2R)Jp8*) were analysed, to minimise the chances of extraneous mutation interfering with the phenotype. Indeed, such embryos were less morphologically aberrant than either stock alone. The chordotonal organs were the main focus of my analysis, the precise array of the lch5 and distinctive pattern of their neurons and dendrites making these organs easier to observe than other sense organs.

sense organ type and cell identity are not affected

DfJp7/DfJp8 embryos have relatively subtle PNS defects. The numbers of sense organs and their individual cell constituents appeared grossly normal, when stained with cell markers such as *couch potato* (*cpo*) (Bellen et al., 1992) (Fig. 3.3.1). This suggests that *cato* does not play a role during the establishment (and maturation) of the SOPs themselves. Furthermore, correct cell fate is assigned to each sense organ cell, as indicated by the correct staining for *elav* (neuron specific) and *pros*, which localises to scolopale (sheath) cell nuclei. Indeed, the highly refractile scolopale cells were observable using light microscopy alone. Thus *cato* does not influence cell fate during, or after, asymmetric cell division.

The v'ch1 neuron is often duplicated in cato deficient embryos

Although the numbers and overall pattern of sense organs is mostly unperturbed by the deletion of *cato* the single obvious exception to the alteration of sense organ number is the chordotonal v'ch1 organ (Fig 3.3.1). In the abdominal segments of DfJp7/DfJp8 embryos this neuron was often twinned or duplicated (Figure 3.3.1). Markers such as anti-Pros and anti-Cpo confirmed that only the neuron was affected, the other sense organ cells were present in normal numbers (Fig. 3.3.1). Thus it

appears that the defect is associated with a further division of the neuron cell as opposed to a cell fate transformation.

However, embryos injected with *cato* dsRNA do not show this aberration (see 3.3.2), thus it seems likely that this particular aspect of the phenotype is a genetic artefact rather than a reflection of *cato* function. Interestingly, it has been noted that deficiency stocks often have defects in the v'ch1 organ (Alain Ghysen, pers. comm.).

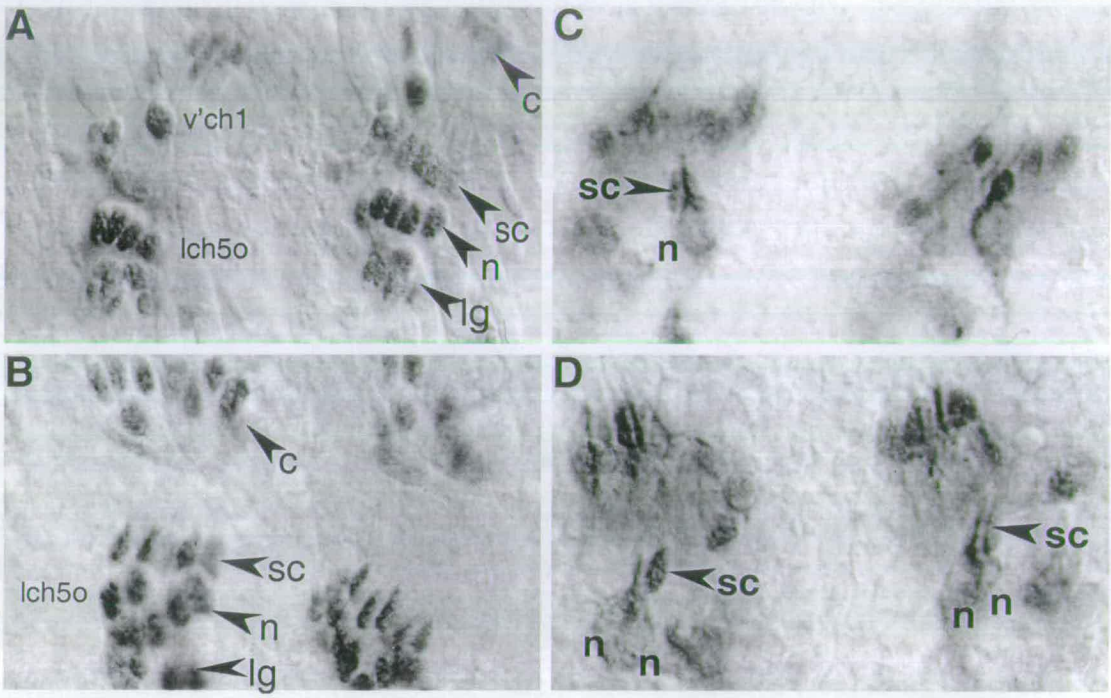


Figure 3.3.1 sense organ type and cell identity are not grossly affected in *cato* deficiency embryos

(A-D) Two abdominal segments showing the lch5 and v'ch1 of stage 17 embryos (A, B) Stained with anti-Cpo antibody. (A) Wildtype. The four cells of each sense organ are clearly visible. (B) *Df(2)Jp7/ Df(2)Jp8*. Although quite disorganised, each cell of the sense organ is present and stained with Cpo. The v'ch1 is out of focus in this view. (C,D) co-labelled for anti Pros and Mab22C10 antibodies. (C) Wildtype, the v'ch1 organ has one scolopale (sc) cell (Pros +ve) associated with its neuron (n) (Mab22C10). (D) *Df(2)Jp7/ Df(2)Jp8*. The v'ch1 organ has a duplicated neuron, but only one associated scolopale. c, cap, lg, ligament.

Neuron morphology and differentiation are grossly affected in embryos deficient for cato

When stained with the marker Mab22C10, specific for the sensory neuronal cell bodies and their processes, widespread defects in neuronal morphology and arrangement were observed in *Df(2R)Jp7/Df(2R)Jp8* embryos. In particular the dendrites of chordotonal neurons are malformed and appear to be longer and often thicker or thinner than their wildtype equivalents (Fig. 3.3.2E-H). The tips of these dendrites are often enlarged, appearing to surround a non-staining hole at the point where the dendrite normally enters the scolopale cell. Furthermore, the highly specialised scolopale cells, normally closely associated with the dendritic apparatus, are often quite disorganised, as revealed by the scolopale-specific marker Ab1188 (Jarman et al., 1993, not shown). Indeed it is highly unlikely that such badly differentiated neurons could function properly, if at all.

In addition to their poor differentiation, the arrangement of the neurons within each particular cluster was also severely affected. Again the most obvious defects were the disruption of the linear array of the lch5 and the ventral vchA/B neurons whose characteristic dendritic orientation was often quite awry (Fig. 3.3.2B).

3.3.2 The effects of *cato* dsRNAi in the embryo

The overall PNS phenotype of these deficiencies suggest a role for *cato* during neural differentiation, consistent with it's later mRNA expression pattern. However, given the size of the deletions in these stocks, it may be that many other genes also have effects on neurogenesis. Therefore, to obtain an independent analysis of *cato* function, in the absence of a specific mutation for the gene, I used the technique of RNAi to gain supporting evidence for these results. Similar experiments were also performed using *ato* dsRNA, as a control for the specificity of the RNAi technique (see Chapter 4.3).

cato dsRNA was injected into w^{1118} embryos which were later stained with Mab22C10. This procedure was attempted exhaustively and the results seem to suggest that similar defects in neuronal morphology are caused by interference with *cato* function (Fig. 3.3.2C,D) but unequivocal results were not obtained using this method. It appeared that neural phenotypes were observed in at least 25% of injected embryos that developed well, however similar phenotypes were also obtained when

embryos were injected with buffer (5-10%). However, it is more difficult to observe cell morphology as opposed to cell markers, thus this may be a limitation of the RNAi method for analysis of late phenotypes. It is clear however, that the v'ch1 is not affected in RNAi embryos. Indeed, no (0%) embryos injected with *cato* dsRNA, or injection buffer had altered numbers of sense organs.

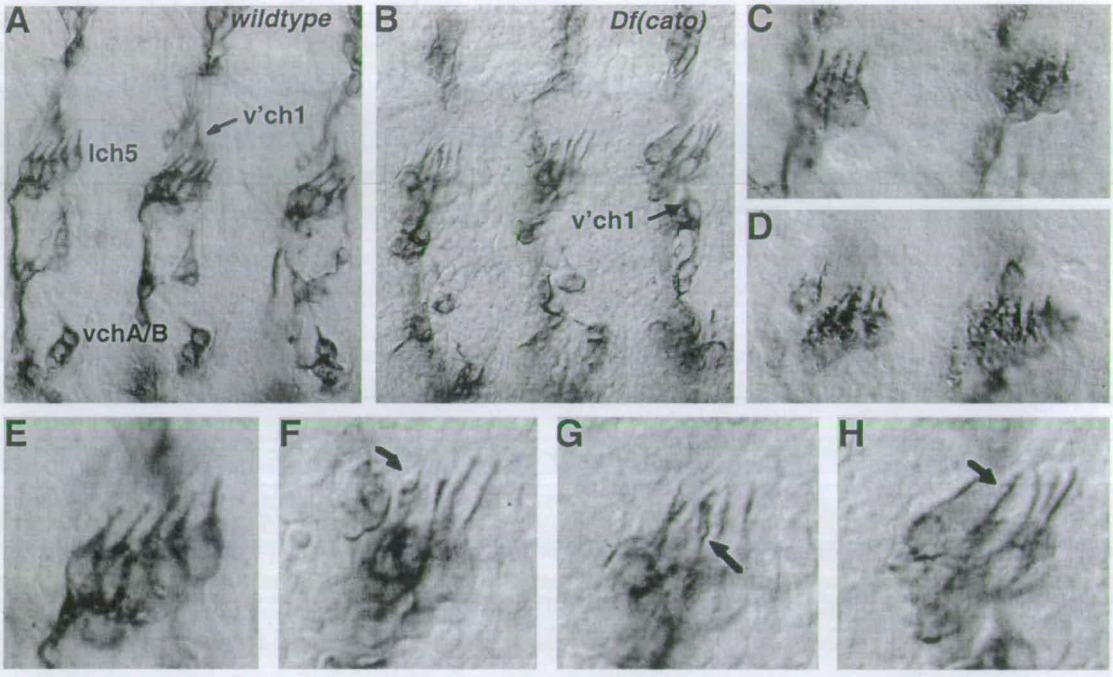


Figure 3.3.2 Loss of *cato* gives rise to neuronal defects

(A-H) Stage 17 embryos stained with Mab22C10 to reveal sensory neurons, (A-B) Three abdominal segments, focussing on the lch5 and vchA/B. (C,D) Close-up of lch5 in two segments from *cato* dsRNA embryos. (E-H) Close-up of lch5. (A,E) Wildtype. Each segment has a similar precise pattern and arrangement of neurons. The lch5 neurons and their dendrites have a characteristic alignment. (B,F-H) *Df(2R)Jp7/Df(2R)Jp8*. In embryos deficient for *cato* the arrangement and morphology of these neurons are perturbed. Dendrites are often misaligned and malformed, with frequent non-staining 'holes' (arrows in F,G,H), or point in the wrong direction (B). RNAi of *cato* appears to disrupt dendrite morphology (C,D).

3.4 Additional deletion of *ase* worsens the *cato* neural defects

The defects observed in *cato* mutant sensory neurons, though probably of functional significance, are relatively subtle. Such subtle phenotypes are not uncommon among the pan-neural genes (Bier et al., 1992; Roark et al., 1995). Indeed, it has often been suggested that the overlapping expression, and relatedness of some pan-neural genes, and their individual weak phenotypes, may be indicative of some functional redundancy. In this case, *ase* is particularly interesting since it also encodes a bHLH protein, and greatly overlaps *cato* expression in the PNS. ase^1 mutant embryos are missing a few external sense organs, however they do not show any significant morphological or differentiation defects when examined with the Mab22C10 marker (Brand et al., 1993). The ase^1 mutation was crossed into the genetic background of *Df(2R)Jp7/CyO* and embryos were collected and stained as before, using Mab22C10.

Indeed, the additional deletion of *ase* increases the severity of neuronal defects of *cato*-deficient embryos (Fig. 3.4.1). In some cases it appears as if the scolopale cell itself is stained, or alternatively that the dendrite extends around it rather than entering into it. Certain cells, although stained for the neural marker, appeared to have either lost, or not developed, any shape or distinctive morphology particular to neuronal cells (Fig. 3.4.1F).

In addition to the dendritic defects, many of the neurons of the dorsal and lateral clusters had poorly formed or misrouted axons. In extreme cases axons were entirely missing (Fig 3.4.1A,B). This suggests that *cato* and *ase* may be required to ensure correct axon pathfinding to the CNS, or indeed for the more general process of axon outgrowth itself. Axon misrouting was most common among the dorsal clusters (Fig. 3.4.1B), which reflects observations made in $pros^-$ embryos which also show axon outgrowth defects (Doe et al., 1991; Vaessin et al., 1991). This may be because the more distal neurons are more sensitive in their response to signals provided by the ventral CNS to aid pathfinding.

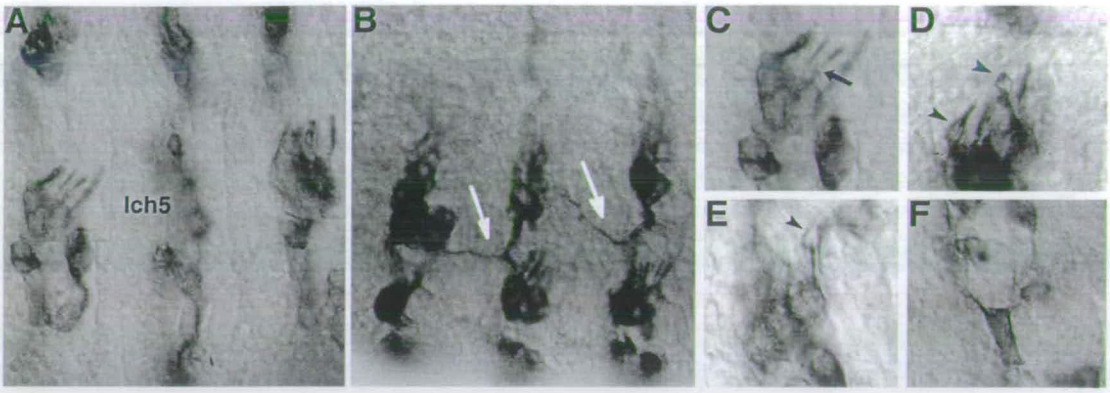


Figure 3.4.1 *ase*^{1};*Df(2)Jp7* embryos have severe neuronal defects

(A-F) Embryos deficient for both *cato* and *ase* (*ase*^{1};*Df(2R)Jp7*) stained with Mab22C10. (A,B) Three abdominal segments. (C-E) Close-up of lch5. Neuronal morphology and alignment are more severely defective than in the *cato* deficiencies alone (A). A significant proportion show axon outgrowth defects from the dorsal and lateral clusters (white arrows, B). Many 'holes' are visible in the dendrites (arrows in C,D,) and there is ectopic staining in and around the scolopales (arrow in (E). Some cells express the neuronal antigen but have no clear neuronal morphology (F).

In summary, the loss of *cato* function does not affect sense organ cell identity or number. However it clearly results in the disruption of sense organ morphology, in particular that of the neuron. Furthermore, concomitant deletion of similarly expressed genes exacerbates these effects

3.5 Misexpression analysis of *cato*

In parallel with loss of function analysis, the UAS-GAL4 misexpression system provides a powerful tool to address how particular genes can behave in an *in vivo* gain of function assay (Brand and Perrimon, 1993). This system takes advantage of the yeast protein Gal4 which binds to a specific site called an upstream activation sequence (UAS) thereby inducing the transcription of a downstream target gene. The *Gal4* gene itself is engineered in a construct downstream of a promoter to generate specific and controllable expression of the protein. Transgenic fly lines are then generated using these constructs. Indeed, many such specific Gal4 lines are available that are particularly useful for misexpression studies of neural genes. To misexpress the gene of interest, such as *cato* or *amos*, the gene ORF is engineered into a misexpression vector, downstream of a UAS. The UAS construct is only expressed in the presence of Gal4 protein, i.e. by crossing two strains of flies.

Several aspects of proneural gene function have been extensively studied in this manner (Jarman and Ahmed, 1998; Jarman et al., 1993b; Jarman and Jan, 1995). In general, ectopic proneural activity within the proneural cluster leads to an elevated number of SOPs that arise from that group, thus giving rise to ectopic sense organs (Brand et al., 1993; Hinz et al., 1994; Jarman et al., 1993b; Rodríguez et al., 1990). In the case of UAS-*sc* and UAS-*ato* their misexpression phenotypes reflect their different activities in terms of function and specificity (Jarman and Ahmed, 1998 see below). For instance, UAS-*sc* gives rise to ectopic external sense organs whereas UAS-*ato* primarily promotes the production of ectopic scolopidia. What underlies the subtype specificity of these genes? Jarman and Ahmed confirmed that the basis of sense organ subtype specificity is the status of *cut* activity. Sc activates *cut*, thereby promoting external sense organ identity, whereas Ato antagonises this, by inhibition of *cut*, to direct scolopidia formation (Chapter 1, Jarman and Ahmed, 1998). Furthermore, misexpression studies revealed that this functional specificity is encoded by the basic domain of the Ato bHLH region (Chien et al., 1996). Thus, not surprisingly, the vertebrate *ato* homologue, MATH1, which has a near identical basic domain, also gives rise to chordotonal organs when misexpressed in *Drosophila* (A.P.J., B.A.Hassan, unpublished data).

Interestingly, despite their endogenous functions, all neural bHLH proteins encode the property of directing external sense organ formation when misexpressed in *Drosophila*, including *ato* itself, under certain conditions (Jarman et al., 1993).

How will the *ato*-like *cato* behave when misexpressed? It is likely that, like other neural bHLH genes, *cato* encodes the property to act as a proneural gene, inducing the formation of external sense organs when misexpressed. But does *cato* also have Ato-like properties? Furthermore, does *cato* misexpression have phenotypic consequences that reflects its endogenous functions? These questions are addressed in this section, with respect to *cato* function.

3.6 Making flies transgenic for UAS-*cato*

At the time these experiments were designed, it was not clear which of the two start codons predicted by the *cato* genomic sequence was correct. Therefore both possible *cato* ORFs were amplified from genomic DNA by PCR and cloned into the pUAST vector to generate UAS-*cato* constructs (Fig. 3.6.1). One encodes the more likely (short) version of *cato*, the other contains both possible start methionines. An antisense construct (UAS-AS*cato*) was prepared to serve as a control but also to investigate the possibility of inducing a loss of function phenotype for *cato*. Although it has been reported that misexpression of antisense RNA can yield specific loss of function phenotypes for bHLH proteins (this study, Misquitta and Paterson, 1999, M. Bownes, pers. comm.) the possible mechanism for this is completely unknown. Once sequenced the UAS-*cato* constructs were injected into w^{1118} embryos to produce transgenic flies.

3.7 Mapping and testing of UAS-*cato* lines

The transgenic UAS-*cato* and UAS-AS*cato* lines were genetically mapped and then tested by crossing them to several *Gal4* driver lines over a range of temperatures. Each line was classed as weak, intermediate or strong, depending on the overall strength of the phenotypes observed. A subset of lines representing a range of UAS-*cato* expression from weak to strong was selected for further analysis.

The *Gal4* lines used in this study have been used previously in the lab and therefore have been relatively well characterised in terms of their expression domains and strength. The lines chosen express Gal4 at times consistent with neurogenesis to conserve the developmental context for the genes of interest.

The results for UAS-*cato* misexpression are described in detail in the following sections. The test experiments demonstrated that there was no difference between the different ORFs of UAS-*cato* constructs. Furthermore, no phenotypes were obtained when the antisense UAS-AS*cato* lines were misexpressed. This may be because all the lines obtained were weakly expressing.

3.8 Misexpression of UAS-*cato* with *Gal4*$^{109-68}$ in the adult

The enhancer trap line *Gal4*$^{109-68}$, thought to be an insertion in *scabrous* (Mlodzik et al., 1990, White and Jarman, in press), expresses Gal4 in proneural clusters and SOPs. This line is therefore suited to address two questions regarding *cato* function. First, how does *cato* behave when misexpressed in the proneural cluster? Does it have *ato*-like properties? Secondly, are there any phenotypic consequences of overexpressing *cato* in the SOPs themselves, the gene's normal domain of expression, that might reflect the loss of function data?

The expression and phenotypes of *Gal4*$^{109-68}$ have been well characterised during experiments to analyse the effects of proneural gene misexpression (Jarman and Ahmed, 1998). The published data describing phenotypes of misexpressing the proneural genes with *Gal4*$^{109-68}$ (Jarman and Ahmed, 1998) provide a good set of data for comparison of the effects of misexpressing UAS-*cato*. Therefore I included UAS-*sc* and UAS-*ato* as controls in these experiments that set out to examine the phenotypic effects of UAS-*cato* misexpression under similar conditions. I first describe the effects of proneural gene misexpression using this line by reference to my own observations in addition to the data published by Jarman and Ahmed (1998). The phenotypes and observations in *Gal4*$^{109-68}$/UAS-*cato* flies are then presented in the following section.

3.8.1 Consequences of misexpressing the proneural genes *sc* and *ato*

UAS-*sc* and UAS-*ato* give rise to distinct proneural phenotypes when misexpressed under the influence of *Gal4*$^{109-68}$. *Gal4*$^{109-68}$/UAS-*sc* flies develop extra bristles in defined regions of the cuticle that correlate with *Gal4* expression; the strongest phenotypes were observed in the scutellum, where both ectopic macrochaetae and microchaetae were formed (Fig. 3.8.1), along the third wing vein (L3) and to a lesser extent in the dorsocentral area of the notum. In general these areas appear particularly susceptible to ectopic gene activity. The phenotype increased in strength at higher temperatures, consistent with higher Gal4 activity: the more UAS-*sc* produced, the ' hairier' the resulting fly (Table 3.8.1). UAS-*sc* does not have the ability to induce chordotonal organ formation.

The phenotype profile of *Gal4*$^{109\text{-}68}$/UAS-*ato* flies is qualitatively dependent on the level of *ato* misexpression. Weak levels of UAS-*ato* expression resulted in ectopic bristle formation as well as a limited number of new scolopidia under the cuticle of the scutellum and along the L3 wing vein (Jarman and Ahmed, 1998) (Table 3.8.1). Wildtype chordotonal organs were hyperplasic, containing a higher number of aligned scolopidial units. These results were similar to those obtained in an earlier study of UAS-*ato* misexpression with *hsGal4* (Jarman et al., 1993). Why does Ato produce ectopic external sense organs in these situations? It has been argued that the low levels of ectopic Ato produced is sufficient to induce ectopic external sense organ formation, but barely enough to produce new scolopidia (Jarman and Ahmed, 1998).

However, at higher expression levels few or no ectopic bristles were observed and specific wildtype external sense organs were missing, most noticeably on the scutellum (Fig. 3.8.1C), L3 and the abdominal sternites (data not shown). Instead in these areas large arrays of scolopidia were visible under the cuticle and the wildtype chordotonal organs were hyperplasic. In these instances the inappropriate *ato* expression is sufficiently high to promote the chordotonal organ fate in the ectopic SOPs, at the expense of sensory bristle formation. Furthermore, in the scutellum mixed organs were also evident in which some eternal sensory features appeared to be associated with chordotonal-specific scolopale structures. Jarman and Ahmed (1998) argued that the presence of such organs, together with the correlation between ectopic chordotonal organ formation and loss of wildtype bristles, suggested a transformation of sense organ subtype under the influence of atonal misexpression.

Finally, *Gal4*$^{109\text{-}68}$/UAS-*ato* flies have rough eyes, there are obvious defects in ommatidia pattern, spacing and photoreceptor identity reflecting *ato*'s function as the proneural gene for R8 photoreceptors (see Chapter 1.x). The basis of this phenotype has been well characterised (White and Jarman, in press), and appears to be the result of ectopic photoreceptor formation, specifically of the founder R8 precursors.

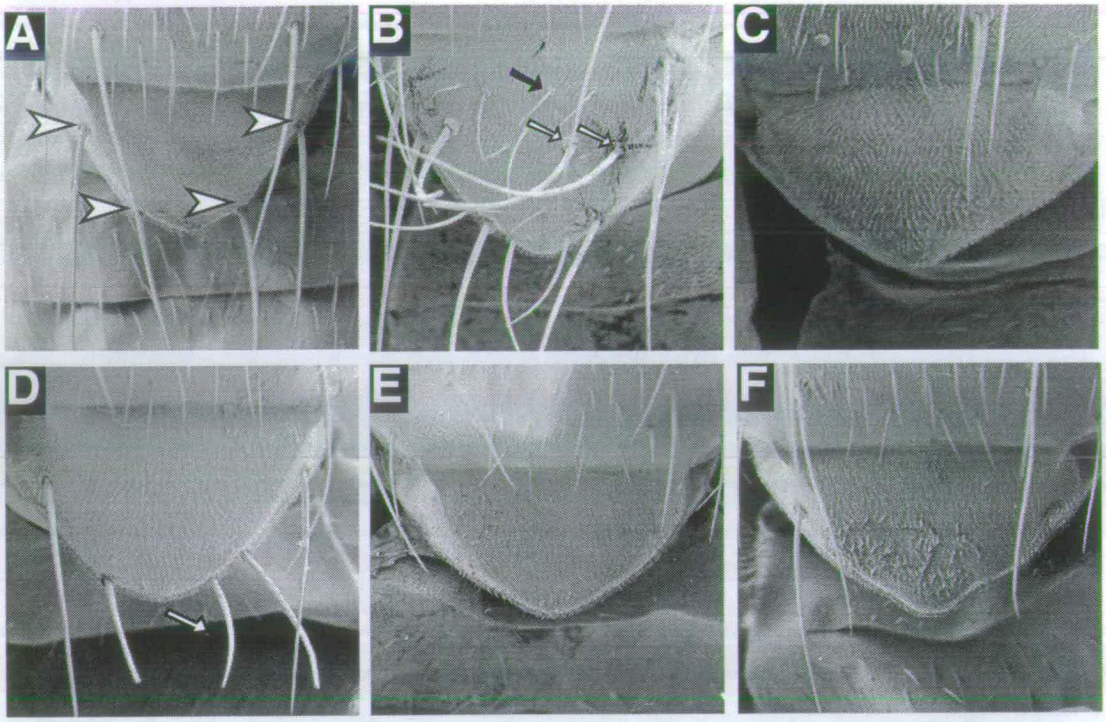


Figure 3.8.1 Misexpression of neural genes gives rise to ectopic sense organs

Scanning electron micrographs of adult scutellum. (A) Wild-type, with the four distinctive macrochaetae (arrowheads). (B) *Gal4*$^{109\text{-}68}$/UAS-*sc*. Many ectopic macrochaetae (white arrow) and microchaetae are formed. (C) *Gal4*$^{109\text{-}68}$/UAS-*ato*. Under conditions of strong UAS-*ato* misexpression the wildtype organs are lost and scolopidia form under the cuticle (Fig. 3.8.2). (D-F) *Gal4*$^{109\text{-}68}$/UAS-*cato* (C) *Gal4*$^{109\text{-}68}$/UAS-*cato*. When weakly expressed, UAS-*cato* induces the formation of ectopic macrochaetae. (D) Under conditions of strong misexpression, the wildtype bristles disappear (E, F).

3.8.2 UAS-*cato* can direct the formation of ectopic sense organs

Like other neural bHLH proteins, *cato* can direct ectopic sense organ formation when misexpressed. Initial observations of *Gal4*$^{109\text{-}68}$/UAS-*cato* flies indicated that *cato* can behave more like *ato* than the AS-C when misexpressed (Fig. 3.8.1D-F and Fig. 3.8.2). For instance, low levels of UAS-*cato* produce ectopic bristles but wildtype organs were often lost at elevated levels of *cato* misexpression (Fig. 3.8.1), and, similar to observations for UAS-*ato*, ectopic scolopidia were observed under the

3.8.3 *cato* misexpression partially mimics *ato*

A more detailed analysis of *Gal4*$^{109\text{-}68}$/UAS-*cato* and *Gal4*$^{109\text{-}68}$/UAS-*ato* flies showed several qualitative and quantitative differences between their respective phenotypes. In contrast to UAS-*ato*, ectopic chordotonal organs were only generated when high levels of UAS-*cato* were present. In fact, even the strongest UAS-*cato* lines are not as effective as UAS-*ato* in forming chordotonal organs - *Gal4*$^{109\text{-}68}$/UAS-*cato* flies have few new chordotonal organs which tend to be isolated in their position (Fig. 3.8.2 C, D). Some clusters of chordotonal organs were observed under the cuticle of the scutellum, but these form long thin disorganised arrays that generally associated with axon structures. Such clusters were quite different in appearance to the densely packed arrays produced in *Gal4*$^{109\text{-}68}$/UAS-*ato* flies (Fig. 3.8.2A). These results indicate that although *cato* can partially mimic the function of *ato* as a proneural gene for chordotonal organs, the proteins are functionally distinct.

This idea is further reinforced by the observation that *Gal4*$^{109\text{-}68}$/UAS-*cato* flies show no phenotype in the eye. This phenotypic difference also correlates with endogenous mRNA expression of each gene: *ato* is expressed in the morphogenetic furrow and emerging R8 photoreceptors in the eye imaginal disc (Jarman et al., 1994, White and Jarman, in press). However *cato* mRNA is not detectable in this domain at any stage during development. *cato* cannot mimic *ato* to produce ectopic photoreceptors.

3.8.4 *Gal4*$^{109\text{-}68}$/UAS-*cato* flies contain poorly differentiated sense organs

A specific feature of the UAS-*cato* phenotype was the large number of mixed and poorly formed sensilla produced; malformed organs consisting of a mixture of bristle-specific and chordotonal-specific structures were observed (Fig. 3.8.2, e.g. F). Indeed, it was often difficult to identify ectopic chordotonal organs on morphology alone, their scolopale structures being poorly differentiated. To aid identification of ectopic organs, scutellae were fixed and stained with the neuron-specific antibody Mab22C10 (Fig 3.8.2). In some cases, scolopale cells associated with individual neurons appeared to stain with the antibody suggesting these cells may have acquired neuronal characteristics. The differentiation of sensory bristles was also affected, in particular the socket and support cells. Together, these data demonstrate that, at least

with one *Gal4* line, *cato* affects the differentiation of sense organs when misexpressed. These results are summarised in Table 3.8.1.

Table 3.8.1 Summary of misexpression with *Gal4*$^{109\text{-}68}$ in the scutellum
es, external sense organs; ch, chordotonal organs, malf, malformed.
↑ increase, ↓ decrease, – absence.

UAS Line	Phenotype at differing expression levels with *Gal4*$^{109\text{-}68}$		
	LOW	MEDIUM	HIGH
ato	es ↑↑	es ↓	es -
	ch ↑↑	ch ↑↑↑	ch ↑↑↑↑↑
		mixed ↑	mixed/malf ↑↑
sc	es ↑↑	es ↑↑↑	es ↑↑↑↑↑↑
cato	es ↑↑	es ↓↓	lethal
	ch ↑	ch ↑↑	
		mixed/malf ↑↑↑	

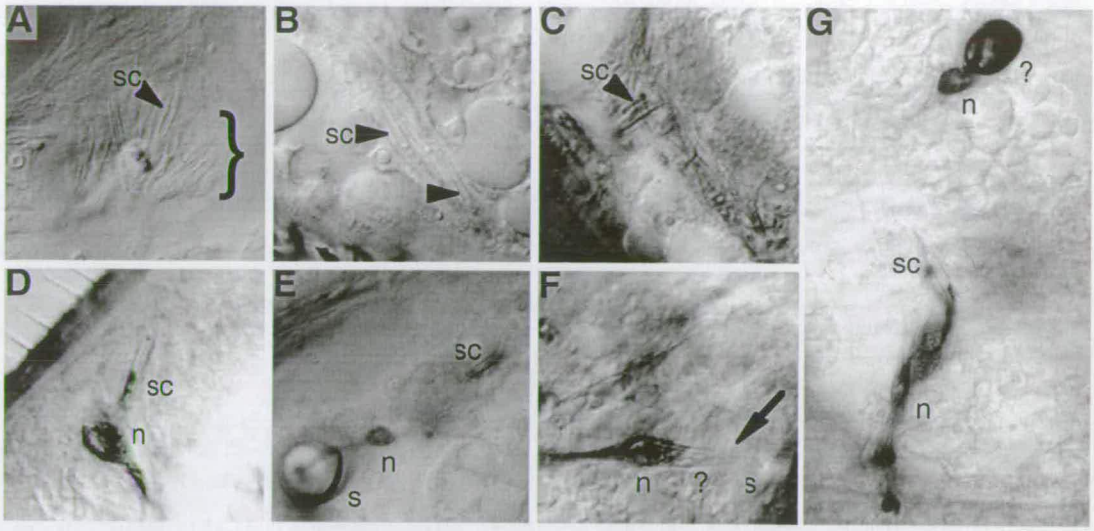


Figure 3.8.2 Ectopic expression of *cato* yields ectopic sense organs and causes differentiation defects.

(A-G) Differential interference contrast (DIC) microscopy of tissue beneath the scutellar cuticle regions, unstained (A) or stained with Mab22C10 (B-G). (A) *Gal4*$^{109\text{-}68}$/UAS-*ato*, a large cluster of ectopic scolopidia (sc). (B-G) *Gal4*$^{109\text{-}68}$/UAS-*cato*. Small arrays of scolopidia are visible under the cuticle (B). More often, ectopic chordotonal organs are isolated (C, D, E, G). Scolopale cells are often malformed, some aberrantly stain for the neural antibody Mab22C10 (C, D, E) Organs of mixed identity are visible (scolopales are arrowed, F). Malformed external sense organs also form (G). n, neuron, ? unknown identity, s, socket cell.

3.9 Misexpression of UAS constructs using other Gal4 drivers

To establish that the *Gal4*$^{109\text{-}68}$/UAS-*cato* phenotypes were not unique to the *Gal4*$^{109\text{-}68}$ stock, several other Gal4 lines were used to investigate whether, under different conditions, UAS-*cato* could mimic proneural gene function, or induce sense organ differentiation defects.

Dpp is an axis-patterning gene and *ptc*, a segment polarity gene, thus the corresponding *Gal4* lines have broader expression domains than the neurogenic *Gal4*$^{109\text{-}68}$ line. This expression includes ectodermal tissue and so misexpression of proneural genes using these lines still induces ectopic sense organ formation. For the most part however, these lines caused a far higher incidence of lethality when used to misexpress UAS-*cato* constructs and therefore direct comparisons between the effects caused by UAS-*cato* and misexpression of the proneural genes are limited. However the phenotypes observed with these lines confirmed that *cato* can behave as an *ato*-like proneural gene in its ability to generate both ectopic external and chordotonal organs.

dppGal4/UAS-*cato* flies showed some distinctive differentiation defects (data not shown) however such phenotypes were inconsistent and mostly not reproduced with other *Gal4* lines, limiting the interpretation of the results.

3.10 An effect of *cato* misexpression is the activation of *ato*

Several possibilities may underlie the ability of UAS-*cato* to produce chordotonal organs. Although *cato* is not a proneural gene, it appears sufficiently similar in its protein structure to partially mimic *ato* and carry out some of its functions, which may include activity in the proneural cluster, inhibition of *cut* and autoregulation of *ato* itself. But is UAS-*cato* doing this directly, or it is acting through *ato?* Indeed activation of *ato* by *cato* may be a true function of the gene, their expression being known to overlap in certain domains. To investigate whether *ato* expression is affected, imaginal discs of *Gal4*$^{109\text{-}68}$/UAS-*cato* were stained for Ato protein.

3.10.1 *cato* can activate *ato* expression under certain conditions

Strong and weak UAS-*cato* lines were examined for their potential to activate *ato* expression (Fig. 3.10.1). 3rd instar larval imaginal discs of $Gal4^{109\text{-}68}$/UAS-*cato* and $Gal4^{109\text{-}68}$/UAS-*ato* were stained for the presence of Ato protein (Fig. 3.10.1). As expected, extra Ato was detected in $Gal4^{109\text{-}68}$/UAS-*ato* discs in proneural clusters and new and wildtype SOPs, including *sc*-dependent precursors. *cato* mRNA was detected ectopically in proneural clusters and SOPs in the discs of $Gal4^{109\text{-}68}$/UAS-*cato* larvae.

Under conditions of strong *cato* misexpression *ato* was ectopically activated throughout the imaginal discs in SOPs, but not in the preceding proneural clusters. Notably no extra Ato was seen across the eye portion of the eye-antennal disc, despite *cato* mRNA expression in the region. In addition, it was evident that weak UAS-*cato* misexpression showed weaker activation of *ato*, and this was restricted further to domains fated to produce chordotonal sensilla. Therefore *cato* can act by itself in proneural clusters to produce ectopic SOPs, but their fate might depend on its ability to activate *ato* rather than an intrinsic ability of *cato* itself.

Anti-Ato antibody does not cross-react with Cato protein

It was important to confirm that the antibody to Ato did not cross-react with Cato protein by misexpressing *cato* in regions where neither *cato* nor *ato* are normally expressed. The *aseGal4* line drives expression in neural precursors of the CNS and PNS as well as all developing photoreceptors. *aseGal4*/UAS-*ato* have extra Ato protein in a pattern that reflects the driver line (Fig. 3.10.1). Strong UAS-*cato* misexpression with *aseGal4* causes lethality, thus it was assumed that Cato protein is produced ectopically. *aseGal4*/UAS-*cato* imaginal discs have a normal Ato protein distribution in the eye portion of the disc (Fig. 3.10.1) indicating that the antibody does not cross-react with Cato protein. Interestingly, in the same larvae, extra Ato was detected in SOPs in regions of the wing and leg discs which suggests that *cato* activation of *ato* is a general feature of *cato* misexpression.

These results correlate with the finding that UAS-*cato* gives rise to fewer and isolated chordotonal organs than *ato*. Indeed, the presence of extra Ato in isolated SOPs seems to correlate with the frequency and pattern of ectopic chordotonal organ observed in the misexpression assay.

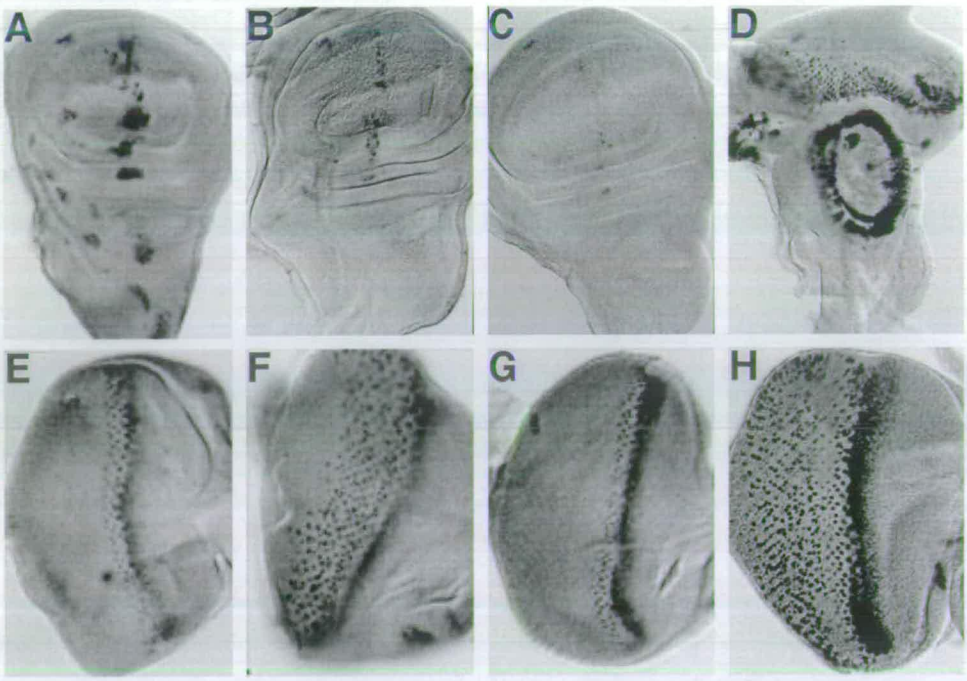


Figure 3.10.1 *cato* activates *ato* when misexpressed

(A-C,E-H) Third larval instar imaginal discs stained for Ato protein. (D) stain for *cato* mRNA. (A-C) Wing. (D-H) Eye. (A) $Gal4^{109\text{-}68}$/UAS-*ato*. Ectopic Ato is present in many proneural clusters and SOPS. (B-D) $Gal4^{109\text{-}68}$/UAS-*cato*. Under conditions of strong misexpression, ectopic Ato is present in external sensory and chordotonal SOPs (B). When weakly misexpressed, Ato expression is more wildtype, and ectopic expression is confined to domains normally fated to give rise to chordotonal organs (C). (D) *cato* mRNA is present in proneural clusters and photoreceptors when misexpressed with this line. (E) Expression of Ato in $Gal4^{109\text{-}68}$/UAS-*cato* eye discs is wildtype. (F) $Gal4^{109\text{-}68}$/UAS-*ato* have ectopic Ato across the eye disc. (G) *aseGal4*/UAS-*cato* flies have a normal distribution of Ato protein. (F) *aseGal4*/UAS-*ato* have Ato across the entire eye disc. For wildtype mRNA expression of *cato* please refer to Fig. 3.2.3.

3.11 Misexpression in the embryo with *hairyGal4*

The adult misexpression studies have demonstrated that, under certain conditions, *cato* can inappropriately behave as a proneural gene, and also give rise to differentiation defects in the ectopic organs that form. I attempted to reproduce these phenotypes in the embryo, to more directly correlate the loss of function data with misexpression studies. The *hairyGal4* (*hGal4*) line has been used previously to characterise the effects of UAS-*ato* misexpression (Jarman et al., 1993), since its expression as a pair rule gene precedes and overlaps that of normal proneural gene function. This line further provides an internal control, since only alternate segments are affected.

Alternate segments of *hGal4*/UAS-*ato* embryos have elevated numbers of chordotonal organs (Jarman et al., 1993). The lch5 is often expanded to include up to ten scolopidia, and the v'ch1 duplicated, or even triplicated. Interestingly, by contrast with the phenotypes of $Gal4^{109\text{-}68}$/UAS-*ato* flies, no transformation of other sensilla to chordotonal identity, or formation of mixed organs were observed in *hGal4*/UAS-*ato* embryos. Furthermore, the ectopic organs produced appeared to be well formed and indistinguishable from their wildtype counterparts.

hGal4/UAS-*cato* embryos also contain ectopic scolopidia in alternate segments. The chordotonal v'ch1 and lch5 organs were most clearly affected, their numbers being increased by one or two sensilla (Fig. 3.11.1B,C), the neurons often being difficult to count because of their aberrant alignment and morphology (Fig. 3.11.1B,C). Again, such observations suggest that UAS-*cato* is less efficient than *ato* as a proneural gene.

The morphology and arrangement of larval sensilla is clearly disrupted in *hGal4*/UAS-*cato* segments (Fig 3.11.1). Indeed, not only are the sense organs misaligned but their dendrites appear to be poorly formed. These defects are reminiscent of those observed in *cato* deficient embryos, and reflect the phenotypes of $Gal4^{109\text{-}68}$/UAS-*cato* flies.

Figure 3.11.1 UAS-*cato* causes differentiation defects in the larval PNS

Stage 17 embryos stained with Mab22C10, two abdominal segments. (A) Wild-type lateral chordotonal neurons (lch5). (B,C) *hairyGal4*/UAS-*cato* lch5. Neuronal morphology is severely affected (B). In addition the lch5 is often expanded by one or two neurons (B). The whole cluster is often shifted dorsally (bracket, C).

In summary, *cato* can behave as an *ato*-like proneural gene when misexpressed. In addition to such properties, the misexpression phenotypes appear to reflect at least one of the *in vivo* roles of *cato*, a function during sense organ differentiation. Thus *cato* demonstrates some functional specificity when misexpressed.

3.12 *cato* is actively repressed in the CNS

3.12.1 *cato* is ectopically expressed in the developing CNS in *prospero* mutant embryos

It seems that loss or gain of *cato* function causes neuronal defects. Like *cato*, embryos mutant for *pros* show neuronal defects, although the basis of these defects is not known. *pros* was recently shown to be important for dendrite morphology (Gao et al., 1999). Furthermore, *pros* mutant embryos exhibit similar axon misrouting defects to those observed in *ase*^{1};*cato Df* embryos. With this in mind, I determined whether *cato* expression depended on *pros* function. Strikingly, loss of *pros* function results in the ectopic appearance of *cato* transcription within the ganglion mother cells (GMCs) and neurons of the developing CNS (Fig. 4.13.1). This correlates with the wild-type expression of Pros in GMCs (Vaessin et al., 1991), indicating that Pros is normally a transcriptional repressor of *cato* in the CNS.

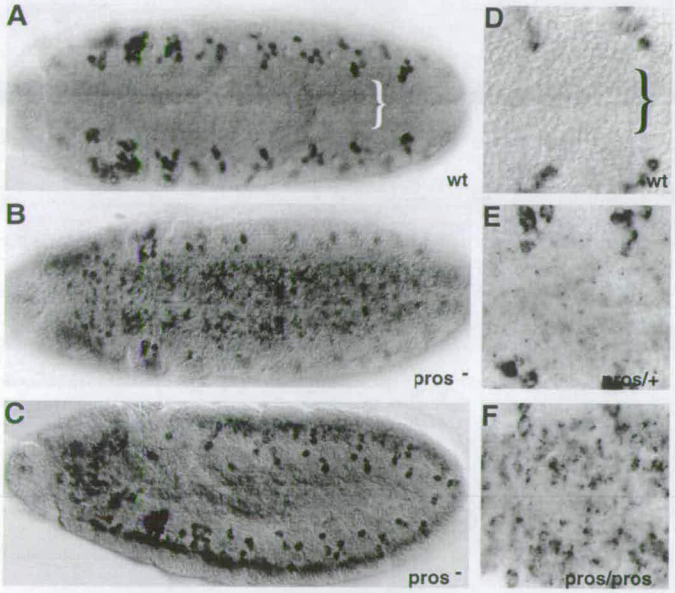


Figure 3.12.1 *cato* is ectopically expressed in *pros* mutant embryos

cato mRNA expression in stage 11 embryos. All views ventral, except (C) (lateral). (A,D) Wild type. *cato* is not expressed in the developing CNS (bracket). (B,C,F) *pros* mutant embryos. *cato* is expressed strongly in the CNS. (E) Heterozygous *pros*/+ embryo. Weak derepression of *cato* is visible as nascent transcripts.

3.13 Discussion

3.13.1 *cato* may be the next step in a hierarchy of bHLH gene function

The findings that in many vertebrate lineages, closely related bHLH proteins are expressed in an overlapping and sequential manner (e.g. Brown et al., 1998; Cau et al., 1997), have led to the theory that such proteins act in interlinked functional cascades (Kageyama and Nakanishi, 1997; Kanekar et al., 1997). That the same may be true for *Drosophila* is supported by these studies of *cato* expression in relation to other bHLH genes.

cato has specific stages to its expression, perhaps allowing the separation of the functions it may have during neurogenesis. During early neurogenesis, *cato* was found to be confined to the *ato*-dependent SOPs, but delayed in external sensory SOPs, relative to their formation, and indeed, *ase* expression. Thus, following the fashion of vertebrate expression cascades, a general order of bHLH gene expression is *ato>cato>ase* for chordotonal organ development, and *sc/ac>ase>cato* in the case of external sense organs.

However, whether such cascades have functional relevance remains to be determined. In this case, it is not clear whether *cato* is actually regulated by *ato*. Indeed, sequence analysis revealed few strong candidates for Ato/Da heterodimer E-box binding sites (A.P.J.), which is in contrast to the high numbers of proneural-binding sites in the genomic sequence of *ase*. Furthermore, *cato* expression is retained in the remaining chordotonal SOP (the P cell) of *ato* mutant embryos. Thus, other factors in addition to, or aside from Ato, are capable of activating *cato* expression in these cells.

3.13.2 Does *cato* regulate *ato*?

It remains an interesting conundrum that a chordotonal organ can still derive from the P cell devoid of *ato* function (Jarman et al., 1993b; zur Lage et al., 1997). Paradoxically, it is redundancy with *sc* expression that allows the formation and differentiation of the precursor (Jarman et al, 1993b). But *cut* expression is not altered in *ato* mutant embryos, i.e. even in the absence of the inhibitory *ato* and the presence of *sc*, *cut* is not expressed in the P cell thus the precursor differentiates as a scolopidium. Possibly there is other positional information that contributes at a very

early stage to proneural cluster and SOP identity. Alternatively, other factors must be able to inhibit *cut* in the absence of *ato* itself. *cato* may be this factor. To address this more directly, future experiments would aim to remove both *ato* and *cato* function simultaneously, potentially using RNAi.

Perhaps there is a more general requirement for the tight link observed between *ato* and *cato* expression. In order to effect *cut* inhibition *ato* expression may have to be prolonged (Jarman and Ahmed, 1998) and indeed, *ato* is different from other proneural genes in that the protein is present relatively long after SOP formation. Is *cato* involved in maintaining or reinforcing that expression? Such a function would also explain why, in the femoral chordotonal precursors, *cato* expression is shut down as *ato* itself is abruptly downregulated, whereas *ase* expression continues (zur Lage and Jarman, 1999). Support for such a role comes from the observation that *cato* can ectopically activate *ato* when misexpressed, but only in the context of an SOP, its normal domain of expression. There maybe cofactors present in the SOP but not the proneural cluster with which *cato* specifically interacts to promote *ato* expression. It is equally possible that the observation is an artefact of misexpression, in this case *cato* mimicing *ato* in an autoregulatory capacity.

3.13.3 *cato* is required for sensory neuron differentiation

cato appears to be the only known transcription factor that is expressed widely but specifically in the developing sensory organs of the PNS, excluded from both the CNS and developing photoreceptors. Furthermore, *cato* localises to the neuron cell. The major defects observed in *cato* deficient embryos are in dendrite and axon morphology making it unlikely that such mutant neurons could function properly, if at all. Sensory neurons display unique characteristics (McIver, 1985; Moulins, 1976), not shared by either CNS neurons, or those of photoreceptors. In particular, the sense organ dendrites are highly differentiated cytoskeletal structures formed by specialised ciliary derivatives (Moulins, 1976). Thus *cato* may supply specific information, crucial for the correct formation of sensory neurons, that is unnecessary for these other nervous system components. Perhaps *cato* regulates genes that are specifically required to modify, or further regulate the modification of, the form of the dendritic cytoskeleton.

Misexpression of *cato* induces defects in sense organ morphology, reminiscent of the loss of function data. This may be because ectopic *cato* expression in a cell other

than the neuron, or at an inappropriate time, interferes with the specialised morphologies of these cells. Thus it is significant that *pros* actively represses *cato* expression in the CNS. If *cato* is required specifically for sensory neuron characteristics, then its derepression in the CNS in *pros* mutants could have deleterious effects on neuronal morphology and function.

Indeed, the fact that *cato* is not normally expressed in the CNS, but can be derepressed, suggests there are specific enhancer elements for the PNS and CNS. Such separable sequence features have been identified in the case of *dpn* (Emery and Bier, 1995). Perhaps *cato* was initially a true pan-neural gene, but functional divergence or specialisation has led to the requirement for active repression in the CNS.

3.13.4 Functional redundancy among neural regulators?

Many neural transcription factors though expressed widely in SOPs, have little detectable effect on PNS development in terms of structure and morphology (Bellen et al., 1992; Bier et al., 1992). In fact, the *cato Df* phenotype is one of the more obvious so far described. Perhaps this reflects the later functions of these genes in neural function, physiology or morphology rather than early fate determination. Concomitant deletion of some of these genes have more dramatic effects on PNS development, suggestive of some functional redundancy among this class of proteins. For instance the loss of *dpn* and *scrt* individually have no visible effects, but loss of both genes causes significant changes in sense organ number and morphology (Roark et al., 1995) Similarly, when *ase* is simultaneously deleted with *cato*, more severe differentiation defects were observed. Do these findings point to a functional redundancy among these proteins, or do they merely reflect the fact that a combination of subtle but distinct phenotypes is functionally disastrous for the developing embryo? The latter possibility seems more likely for several reasons. In the first instance, *dpn* and *scrt* encode very different classes of transcription regulators (Bier et al., 1992; Roark et al., 1995). Moreover, *cato* expression is more specific than the pan-neural *ase*, and sequence differences between these proteins make it likely that they are quite distinct functionally. Lastly, the analyses here are based almost solely on the expression of one neural marker. At this time, it is completely unknown whether each gene may regulate different neural components.

3.13.5 Functional differences among bHLH proteins

cato is clearly not a proneural gene. However, taken out of its developmental context it was demonstrated that *cato* can function as an *ato*-like proneural gene, partially mimicing *ato* function. But unlike the case of *ase* and *sc*, where misexpression phenotypes are indistinguishable (Brand et al., 1993), *cato* is not entirely *ato*-like, and indeed is not very efficient as a chordotonal proneural gene: ectopic SOPs are fewer than when *ato* is misexpressed, and fail to develop properly. Thus *cato* and *ato* must have distinct functional properties.

Perhaps the differences between the two genes in the bHLH domain that are conserved in the respective *D. virilis* orthologues, are sufficient to account for their different properties. Indeed, it may be region outside the bHLH domain, not conserved at all.

These misexpression studies highlight one function of *ato* that *cato* does not encode, which is the recruitment of extra SOPs to generate large clusters of sense organs. This is a key feature of *ato* function during chordotonal organ and ommatidial development (zur Lage and Jarman, 1999). It will be interesting to determine if these differences are attributable to differences in the bHLH domains of these proteins. A complementary experiment might be to examine whether *cato* misexpression can rescue an *ato* mutant phenotype. The different misexpression profiles of *ato* and *cato* may help to pinpoint how this feature of *ato* function is encoded in its protein structure.

Chapter 4

Expression and Functional analysis of *amos*

4.1 Introduction

Similar methodologies to those described in 3.1 have been employed in the analysis of *amos* expression and function.

During the course of this thesis, much of my efforts were diverted to analysing the roles of *cato* in neurogenesis. I have carried out a preliminary expression analysis for *amos* and examined its possible functions during larval development. Interestingly, *amos* has been independently isolated by another group (Huang et al., 2000), and their functional characterisation of *amos* confirms and complements the results presented here for *amos* function in the embryo.

In the meantime, Dr. Petra zur Lage took over the *amos* project, to address how this gene functions during adult neurogenesis. This includes expression, misexpression and genetic analysis using deficiency stocks that map to the *amos* region (Goulding et al., 2000).

4.2 *amos* mRNA expression in the embryo

The *amos* mRNA expression pattern suggests it is a new proneural gene. Typically, such genes are transiently expressed in clusters of ectodermal cells at the onset of neurogenesis. I found that *amos* is expressed in this fashion during both larval and adult development: just like proneural genes, *amos* mRNA quickly becomes restricted from clusters to a single cell, the SOP, vanishing soon after.

4.2.1 *amos* has restricted proneural-like mRNA expression

amos is expressed in a small subset of cells in each body segment

in situ hybridisation to mRNA in wholemount embryos showed that *amos* is expressed very transiently during embryogenesis. *amos* mRNA is present in a small cluster of ectodermal cells in each body segment (thoracic T1-T3, and abdominal A1-A7) during stages nine and 10 (Fig.4.2.1). By early stage 11 expression is restricted to a single cell per segment, and is shut down shortly after this. The timing and ectodermal location of the expression suggests that these cells correspond to proneural clusters that resolve to a SOP by lateral inhibition.

In addition, I examined *amos* mRNA expression with respect to the distribution of the proneural Ato protein (Fig. 4.2.1). These double-labelling experiments demonstrated that the timing of expression for both genes is similar, and even overlaps to some extent in the proneural cluster. Ato expression is the first to become restricted to a single cell but *amos* mRNA then becomes limited to the same cell. Both genes co-label this cell for a short while before expression is shut down. The identity of this *ato*-expressing precursor has not yet been established (Jarman et al., 1995), however it is not a chordotonal precursor (Jarman et al., 1995). Instead it may be the precursor of the dorsal bipolar dendritic (dbd) (Campos-Ortega and Hartenstein, 1998) (Fig. 4.2.1) and one of the dmd5 neurons. This is rather significant because these are the only two remaining neurons per segment in larvae deficient for both the AS-C and *ato* (Jarman et al., 1993). *amos* is therefore a candidate as the proneural gene required for their formation.

In addition to this highly temporal body expression, *amos* has complex expression in the head regions of the embryo, again in a proneural-like fashion (Fig. 4.2.1). This expression corresponds to the location of the antenno-maxillary complex (AMC) anlage (Fig. 4.2.1), whence sensory organs such as taste and olfactory receptors of the larva are formed (Campos-Ortega and Hartenstein, 1998).

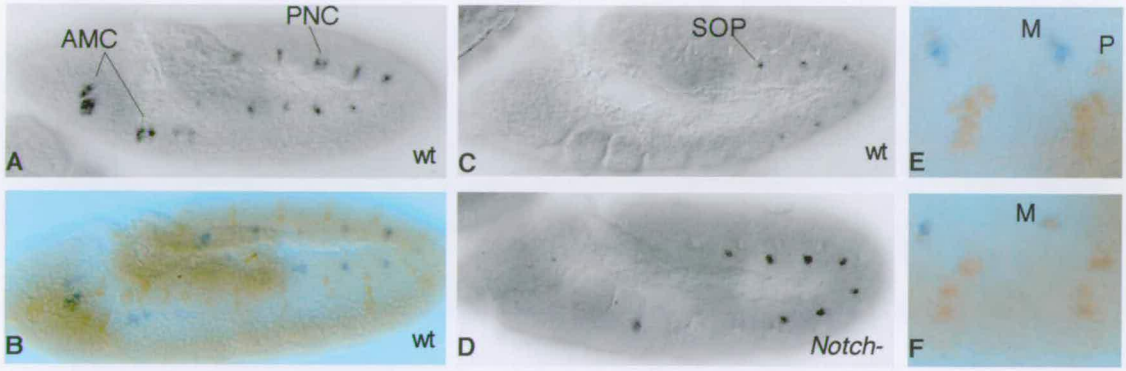


Figure 4.2.1 *amos* expression in the embryo

(A-C, E,F) Wild type (D) Notch mutant. (A) Stage 10 embryo, showing expression in antennomaxillary complex domains (AMC) and a single cluster of ~5 cells in each thoracic and abdominal segment. (B) Wild type stage 10, stained for Ato protein and *amos* RNA. There appears to be some overlap in their expression. (C) Stage 11, abdominal expression is now refined to a single SOP, thoracic and head expression have already ceased. (D) Notch mutant, stage 11, showing continued expression in all cells of the abdominal clusters. (E) Two abdominal segments showing proneural expression of Ato and *amos* from stage 10 and (F) stage 11. Their expression overlaps in the presumed dbd precursor (M). PNC, proneural cluster, SOP, sense organ precursor.

amos expression is influenced by lateral inhibition

When the lateral inhibition signalling process is abolished proneural gene expression is extended in the proneural cluster (reviewed in Campos-Ortega, 1993). Consistent with the notion that *amos* is indeed a proneural gene, I found that, in Notch mutant embryos, *amos* expression persists in groups of cells instead of becoming restricted to a single SOP, as in wildtype (Fig. 4.2.1).

amos is expressed in non-neural tissues

Curiously, in addition to the expression detected in neural tissues, *amos* mRNA is present in cellular blastoderm embryos. A single dorso-ventral band of expression extends two-thirds of the way around the posterior of the embryo (data not shown). Furthermore, *amos* is also expressed during specific stages of oogenesis; transcripts have been detected in nurse cells, the centripetal follicle cells and the oocyte itself (R. Kirby, data not shown).

Thus *amos* is not solely a neural gene, a feature shared with the AS-C genes which also have roles in the development of other tissues. In the case of *amos* it is not clear whether such expression has functional basis, however it's presumed binding partner, *da,* is also expressed ubiquitously in these domains.

4.3 Loss of *amos* function results in the loss of a subset of larval sensilla

Some difficulty was encountered in analysing loss of function phenotypes for *amos* deficiency stocks. The cytogenetics of the region are particularly notorious and known deficiencies of the region are large and difficult to work with; homozygous deficiency embryos did not appear to develop far enough to distinguish the PNS with clarity (APJ). Instead, RNAi was used to examine the effects of disrupting *amos* function in the early embryo. dsRNA was injected into 0-1h w^{1118} embryos which were allowed to develop and were then prepared for analysis with the neural marker Mab22C10. Similar experiments were also performed using *ato* dsRNA, as a control for the specificity of the RNAi technique. In this case an *enGal4tauGFP* line was used, which expresses GFP in a subset of the larval chordotonal neurons (B. Sanson, pers. comm). In over 85% of injected embryos *ato* RNAi caused the loss of chordotonal neurons identical to that of *ato* mutant embryos (Fig. 4.3.1). However, in embryos injected with either *amos* antisense RNA or dsRNA the characteristic dorsal bipolar dendritic (dbd) neuron, which sends lateral axons from each body segment, appeared to be missing in many segments. In these cases the phenotype was reproducible in 30-50% of injected embryos, and was more local to the site of injection. Notably, the dbd is one of the neurons which still form in AS-C;*ato* deficient embryos (Jarman et al., 1993b). These experiments have been independently carried out by others (Huang et al., in press) who further show that *amos* is also required for another solo md neuron of the dmd5. Thus together, *amos*, *ato* and AS-C provide the proneural functions for the entire PNS of the larval body segments.

The expression of *amos* indicated that it may also play some role in the development of the AMC, the major site of larval olfaction. Preliminary RNAi data seems to support this possibility, but the effect appeared subtle and was difficult to analyse. Future experiments would aim to remove *amos* and *ato* simultaneously to see if there is a stronger effect.

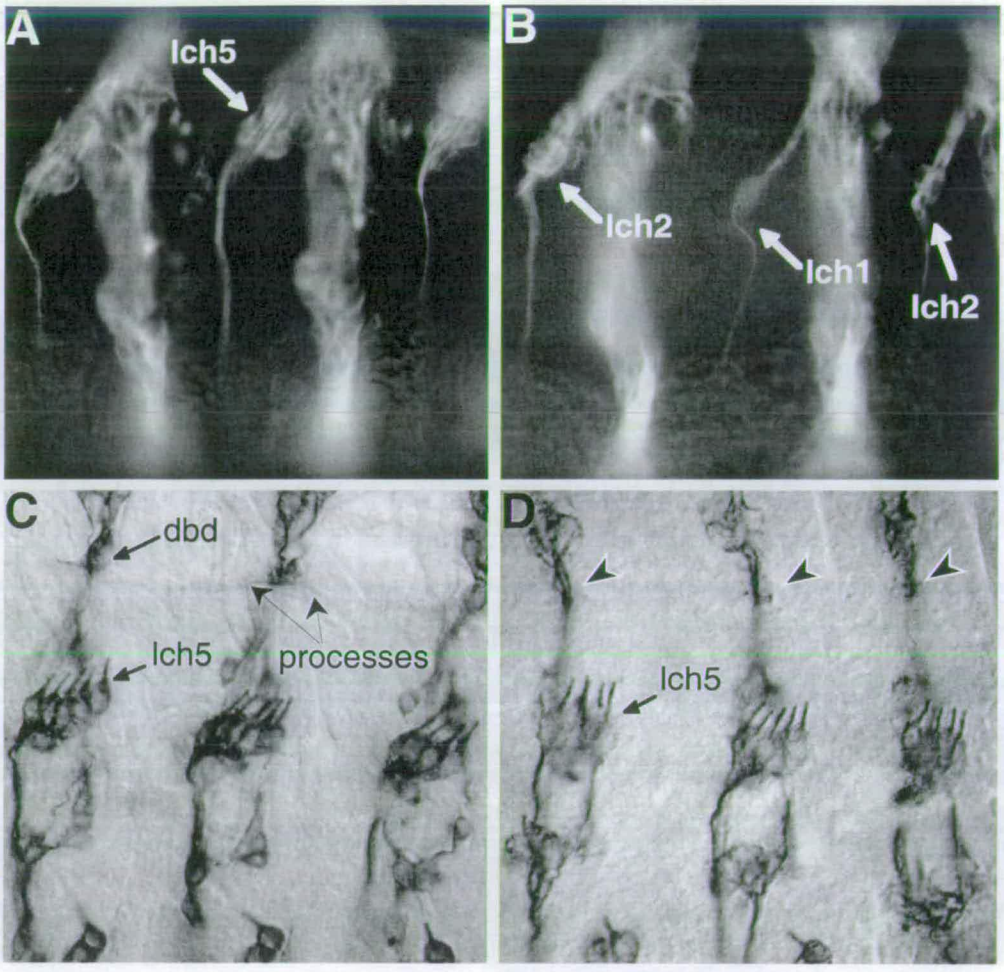


Figure 4.3.1 *amos* affects the formation of the dbd neuron

(A-D) Three abdominal segments from stage 17 embryos. (A,B) *enGal4,UAS-GFP*, view of the lateral chordotonal organs (lch5, white arrow). (A) Control injected with buffer. The lch5 is present in each segment. (B) An embryo injected with *ato* dsRNA. Fewer chordotonal organs are visible. (C,D) w^{1118}. (C) control injected with buffer. The dbd neurons (black arrow) with their distinctive lateral processes (thin arrows), are clearly visible in each segment. (D) Embryo injected with *amos* asRNA. The dbd neurons are missing in each segment (black arrowhead). lch5, lateral chordotonal organ 5.

4.4 In imaginal discs *amos* is expressed in the proneural domains of the antennal olfactory sensilla and leg tarsal claw

amos expression is extremely restricted during adult development and was only observed in pupal imaginal discs. The main expression initiates approximately 4h after puparium formation (APF) in three broad semicircular domains on the medial side of the third antennal segment (Fig. 4.4.1A). In the adult, this region houses over 450 olfactory sensilla of three morphological classes: basiconic, coeloconic and trichoid sensilla. The three *amos*-expressing domains correspond to the sites fated to give rise to olfactory sensilla precursors (Reddy et al., 1997). The outer domain (3) has the strongest *amos* expression and this region widens by ~8h APF. *amos* expression continues until at least 21h APF during which time the two inner domains (1 and 2) appear to fuse to give a single inner region of expression (Fig. 4.4.1C) (P.z.L. pers. comm.). It was difficult to analyse the expression beyond 21h APF because of probe penetration problems. Notably this *amos* expression resembles the three crescents reported for *lozenge* (*lz*) (Gupta et al., 1998). *lz* encodes a transcription factor (Daga et al., 1996) that affects the formation of at least two types of olfactory sensilla but does not fulfill the criteria required for 'proneural' function (Gupta et al., 1998). Significantly, *amos* expression begins rather later than *lz*, suggesting that *lz* is upstream of *amos*.

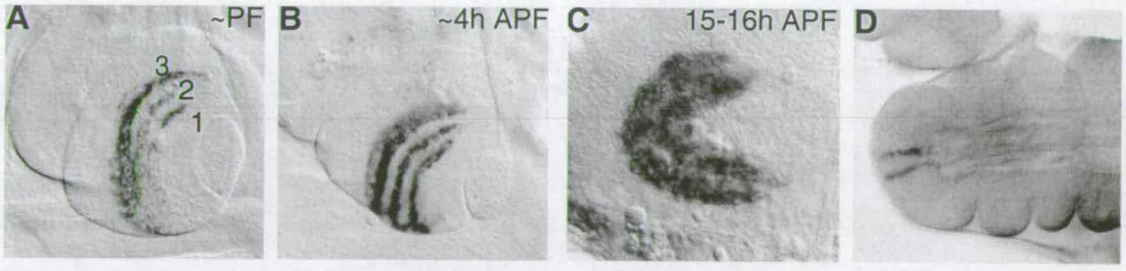


Figure 4.4.1 *amos* expression during adult development.

(A-C) Wild-type pupal antennal discs. (A) Puparium formation (PF). Expression initiates in 3 semicircular bands on the developing third segment. (B) ~4 h APF. (C) 15–16 h APF. Expression broadens, and the inner two bands appear to merge. (D) Wildtype distal leg disc, 0–4 h APF. Expression in two cells in the region fated to give rise to the tarsal claw.

Thus the expression of *amos* strongly suggests this gene as a candidate proneural gene for olfactory sensilla. The provenance of the three sensilla types is not well known. However by analysis of *lz* mutants, it is possible to infer that the basiconic sensilla probably arise from domain 3. In strong *Lz* mutants this domain of A101 expression is missing, which would correspond to the missing basiconic sensilla in adult *lz* flies (Gupta et al., 1998). Furthermore, the likely origins of coeloconic sensilla are the inner two domains, given *ato* is expressed there and is known to be the proneural gene for these sensilla (Gupta and Rodrigues, 1997). This leaves the trichoid sensilla, of which the least is known. Based on its expression in all three domains, *amos* may be required for any or all three types of sensilla.

In addition, limited *amos* expression was detected in the pupal distal leg discs in two cells (Fig. 4.4.1D). This expression was extremely transient and correlates with the position of the innervated tarsal claw for which there is no known proneural function. Interestingly, *lz* is also expressed in a similar pattern in this region (P.z.L., pers. comm.).

4.5 Summary of experiments performed to analyse the role of *amos* during adult neurogenesis

amos mRNA expression in the antennal imaginal discs was suggestive of a role during the specification of olfactory SOPs. Many experiments have been carried out that confirm this hypothesis (Goulding, zur Lage and Jarman, 2000) and these results are primarily the work of Dr. Petra zur Lage. I have summarised these experiments in the following section to allow a more general discussion regarding the functions of *amos* during neurogenesis.

4.5.1 *amos* expression is confined to the proneural clusters of the olfactory precursors

Double labelling experiments with the A101 enhancer-trap line, which stains all SOPs in the adult (Huang et al., 1991), confirmed that SOPs indeed arise from *amos*-expressing ectodermal bands. However, *amos* appeared to be restricted to the proneural clusters, there being little overlap in the A101 cells themselves. Thus *amos* expression is shut down as the SOPs form. This contrasts with observations for *ato*, where protein is detectable in the SOPs (Gupta and Rodrigues, 1997). However, it is possible that Amos protein may perdure in the SOPs, thus not detected in these experiments.

Despite overlapping expression in the inner domains of the antenna, analysis of *amos* mRNA in *ato* mutants and *ato* in *lz* mutants demonstrated that *amos* and *ato* expression are largely independent of each other.

4.5.2 Genetic evidence for the requirement of Amos/Da heterodimers in olfactory SOP formation

SOP formation is sensitive to the simultaneous reduction in copy number of proneural genes and *da*, which encodes their common heterodimer partner (Dambly-Chaudiere et al., 1988,). For example, reducing the dosage of *ato* and *da* genes results in fewer sensilla coeloconica (Gupta and Rodrigues, 1997). Similar gene dosage experiments were carried out using small lethal chromosomal deficiency stocks (*Df(2L)M36F-S5* and *Df(2L)M36F-S6*) that delete *amos*. In summary, flies

lacking one copy each of *amos* and *da* had significantly lower numbers (by 30%) of sensilla basiconica, when compared to individual stocks. To a lesser extent, the numbers of sensilla trichodea were affected.

4.5.3 Misexpression of *amos* gives rise to ectopic olfactory sensilla

When proneural genes are misexpressed, ectopic sense organs are formed. Moreover, the type of ectopic sense organ produced often reflects the neuronal subtype specificity of individual genes (Jarman and Ahmed, 1998; Jarman et al., 1993; Jarman et al., 1995). For example, misexpression of *ato* results mostly in the formation of ectopic chordotonal organs (Jarman et al., 1993) and sensilla coeloconica (Gupta and Rodrigues, 1997).

Gal4/UAS misexpression studies were carried out with *amos* to determine if it can function as a proneural gene, and to assess its subtype-determining properties. Indeed, misexpression of *amos* in the antenna yields ectopic sense organs, but by contrast with *ato*, UAS-*amos* induces the formation of supernumerary olfactory sensilla of all three subtypes. Interestingly, a proportion of ectopic sensilla were of intermediate basiconica/trichodea morphology. This had also been reported as an effect of *lz* misexpression (Gupta et al., 1998).

The subtypes of ectopic sense organ produced by *ato* misexpression are dependent on the tissue. Thus, on the antennal third segment, sensilla coeloconica form whereas chordotonal organs form at most other sites. In regions outside the antenna, *amos* appeared to mimic *ato* function, producing a mixture of ectopic chordotonal organs, and bristles, though in much fewer numbers. Strikingly, *amos* misexpression generated unusual organs that appeared morphologically similar to olfactory sensilla (particularly sensilla trichodea). In fact, on the thorax, many of the wildtype bristles were apparently transformed to olfactory-like sensilla. Indeed, such converted organs were sometimes innervated by two or more sensory neurons (A.P.J., pers. comm.), a characteristic of olfactory sensilla.

4.5.4 A relationship between *amos* and *lz*

As mentioned earlier, the pattern of *amos* expression resembles that of *lz* in the antenna and leg discs (Gupta et al., 1998), but *amos* expression initiates later than *lz*.

Does *lz* regulate *amos* expression during the process leading to the formation of basiconic and trichoid SOPs and the tarsal claw? Several lines of evidence suggest this is so. Firstly, a range of *lz* alleles were examined for alterations in *amos* expression. In strong *lz* alleles that almost completely lack sensilla basiconica and have only 50% of sensilla trichodea, *amos* expression was missing from the middle of all three antennal bands. In weaker *lz* alleles, *amos* expression was less affected, but patchy. In a complementary experiment it was determined that ubiquitous *lz* expression could induce ectopic *amos* expression in many regions of the imaginal discs.

Furthermore, genetic interactions between *lz* alleles and *amos* deficiencies were determined; in flies lacking full *lz* function and one copy of *amos*, the number of olfactory sensilla were further reduced. Finally, when *amos* expression was induced in a *lz* mutant background, the formation of sensilla basiconica was partially rescued. Significantly, *ato* was unable to direct this rescue.

Together these results indicate that *amos* transcription is partly downstream of *lz* and that its loss of expression may explain the loss of sensilla basiconica and trichodea in *lz* mutants.

lz independent roles for amos

amos expression was retained at the ends of the ectodermal antennal bands in *lz* alleles. This expression may account for the remaining sensilla trichodea unaffected by the deletion of *lz*.

4.6 Discussion

4.6.1 *amos* is a new PNS specific proneural gene

amos displays many characteristics of proneural genes. Not only does the transient, ectodermal specific expression pattern reflect those of the known proneural genes, but its expression partly overlaps with at least one of these (*ato*) during larval and adult development. Furthermore, *amos* expression is sensitive to the process of lateral inhibition. Like other proneural genes, loss of *amos* function results in the reduction of sense organs whereas its misexpression induces the formation of ectopic sensilla. Finally, as has been characterised in the case of *ato* (Chien et al., 1996; Jarman and Ahmed, 1998), *amos* encodes neuronal subtype specificity.

With the discovery of *amos*, the origin of almost the entire PNS can be accounted for by the three proneural functions of *amos*, *ato* and the AS-C.

4.6.2 The proneural genes for olfactory sensilla

The formation of olfactory sensilla require two genes of the *ato*-like subfamily, *amos* and *ato* itself, and these genes appear to have complementary functions. The requirement for *ato* is restricted to the adult sensilla coeloconica Gupta and Rodrigues, 1997) and data presented here indicates that *amos* is necessary for the specification of sensilla basiconica and trichodea. The evidence is clearest in the case of basiconica, in that genetic interaction was observed between *amos* and *da*. Clearly, definitive evidence will require the isolation of specific *amos* mutations.

In the larva both *amos* and *ato* are expressed in the domains that give rise to the AMC, the major site of olfaction (and gustation). It will be interesting to determine conclusively, whether, as in the adult, *amos* and *ato* have complementary functions in the specification of larval olfactory sensilla, or whether one or other gene is the major proneural determinant in this domain.

Once specific mutations for *amos* have been isolated it may be possible to address whether olfactory specificity can be discerned among different subsets of sensilla.

Simple behavioural assays have been described for analysis of larval olfaction and gustation (Heimbeck et al., 1999).

4.6.3 *amos* is required for the *ato*- and AS-C- independent larval md neurons

In the larval body segments the only organs for which there are no known proneural function are two md neurons and their location correlates with *amos* expression, making this gene a likely candidate for their formation. Indeed, these two neurons are absent when *amos* function is disrupted (this study, Huang et al, in press). In addition, these results demonstrate that there is no functional redundancy between *amos* and *ato* for the formation of these neurons, contrary to suspicions from expression data. Perhaps the co-expression of *ato* and *amos* in these precursors is an evolutionary hangover, or indicate that similar regulatory pathways still exist for these genes in this context.

4.6.4 Neuronal subtype specificity of bHLH proneural genes

Proneural genes have dual roles. In the first instance, they provide cells with neural competence. In addition, they specify neuronal subtype specificity (Chien, 1996, Jarman and Ahmed, 1998, S.E.G., P.z.L., A.P.J., in press). Misexpression assays provide evidence that this specificity is not just a feature of the developmental context in which such genes are normally expressed (Jarman and Ahmed, 1998). For instance, only *ato*-like bHLH genes can promote chordotonal organ formation, not the AS-C genes. Shared features of the *amos* and *ato* bHLH domains, or possibly other structural regions of the proteins provide important information for olfactory sensillum determination. Perhaps one such function is the inhibition of *cut*. Like photoreceptors and chordotonal organs, olfactory sensilla do not require *cut*. Interestingly, all three types of sensory structure require some form of local cell recruitment in their development. EGFR signalling is required for this in ommatidia and chordotonal organ clustering, perhaps directly controlled by *ato* itself (Freeman, 1996; zur Lage et al., 1997; zur Lage and Jarman, 1999, White and Jarman, in press). It is not known whether the same signalling pathways are invoked for olfactory SOPs, but it is possible that both *amos* and *ato* control this process.

However, *amos* and *ato* are not functionally interchangeable: *amos* cannot mimic *ato* fully (nor vice versa) and *amos* affects all types of olfactory sensilla. In fact, the subtype specificity of *amos* is less constrained than that of *ato*, which is actually constrained by developmental context. It will be interesting to determine whether the basis for these differences in subtype specificity reside in the conserved differences of the bHLH domains or other parts of the protein not conserved with *ato*.

4.6.5 Is *lz* is a prepattern gene and specific cofactor?

Gupta et al. (1998) recently proposed that, analogous to its function during eye development (Daga et al., 1996; Flores et al., 1998), *lz* functions to set up the correct expression patterns of fate determining factors in the antenna. Data presented here indicates that *amos* is one of these factors, though whether *amos* is directly regulated by *lz* is not clear. It is notable that this *lz-amos* regulatory link may also function during tarsal claw development. Interestingly, this link is restricted to *amos* but not *ato*.

In addition to its role in *amos* activation, *lz* influences the choice between basiconic and trichoid fate (Stocker et al., 1993). Gupta et al, 1998 proposed that high *lz* results in basiconic fate whereas lower levels resulted in trichodea fate formation (Gupta et al 1998). Perhaps *amos*-dependent SOPS are sensitive to *lz* dosage and as a consequence have alternative fate potentials. In the absence of *lz*, the remaining *amos*-dependent SOPs assume a 'basal' trichoid fate. So, without *lz* (or low levels) trichoid SOPs develop, but in the presence of high *lz*, basiconic SOPs form. Thus *lz* may act at the level of *amos* activation for trichodea, but directly modulate *amos* function to promote basiconic fate. This possibility is supported by observations that misexpression of either gene can yield ectopic organs of mixed morphology (Gupta et al., 1998).

Thus in this case, *lz* could be a cofactor for *amos* in the decision of olfactory SOP subtype, or it could be activating an unknown co-factor to interact with *amos* to the same end. Experiments are underway to address whether *lz* and *amos* proteins can interact directly (P.z.L).

Chapter 5

General Discussion and Conclusions

General Discussion and Conclusions

This thesis presents the identification and analysis of two further bHLH genes required during *Drosophila* neurogenesis. The isolation of *amos* heralds the (near) completion of the proneural story for PNS formation in the larva and adult. *cato* represents a new class of *Drosophila* neural bHLH genes, being specific for the later development of the sensory component of the PNS.

5.1 The potential molecular bases of functional specificity

The requirement for cofactors

Several levels of functional specificity have already been established for bHLH proteins, and ascribed to molecular features of the bHLH domain itself. For instance, based on the possession of particular residues within the domain, a protein can be classed as myogenic, or neural, and indeed have been shown to encode such properties by their conversion of uncommitted cells to specific tissue types (e.g. Davis and Weintraub, 1992; Jarman et al., 1993; Lee et al., 1995). Furthermore, the proneural genes encode neuronal subtype specificity, as evidenced by the different properties of the fly AS-C and *ato* genes when misexpressed (Davis and Weintraub, 1992; Jarman and Ahmed, 1998; Jarman et al., 1993; Lee et al., 1995). Yet again, the basis for this turns out to be a property of the bHLH domain (Chien et al., 1996) in particular, non-conserved residues of the basic region hypothesised to interact with other unknown factors (Jarman and Ahmed, 1998; Jarman et al., 1993). Interestingly, the misexpression phenotypes for *ato* are dependent on the developmental context in which the gene is misexpressed. Are there different cofactors present in each of these domains that specifically modulate *ato* function?

Indeed, what are these cofactors, and how do they modulate the subtype specificity of bHLH proteins? In the case of the chordotonal organ versus external sense organ decision, *ato* misexpression can transform wildtype sensory bristle SOPs to a chordotonal fate, suggesting that any interaction must be with a factor normally present in both types of SOP. This candidate cofactor may be one of the many prepattern genes expressed to define the regions of neurogenic ectoderm.

Several interactions have been described that allow modulation of bHLH function. For instance, the MADS-box cofactor MEF2 promotes the myogenic activity of MyoD in one tissue (Molkentin et al., 1995) and the neural activity of Mash1 in another (Molkentin et al., 1995). Furthermore, in *Drosophila*, the Dorsal morphogen interacts with bHLH factors to initiate the differentiation of the embryonic mesoderm (with Twi, Gonzalez-Crespo and Levine, 1993; Shirokawa and Courey, 1997) and neuroectoderm (with AS-C and Da, Gonzalez-Crespo and Levine, 1993).

Although misexpression of *amos* can direct chordotonal identity in the adult, it is not as effective as *ato* in achieving this, perhaps because in this instance, *amos* preferentially converts ectopic SOPs to an olfactory fate. It seems from the studies presented here that *amos* is less constrained by developmental context than *ato*. It may be that *amos* does not require the same modulation by cofactors as *ato* does. Indeed, subtype specificity for *amos* seems to be active at the level of individual olfactory SOPs, as studies with *lz* suggest (Chapter 4).

Other shared features among ato-related genes

Interestingly, several neural bHLH factors have now been identified that promote chordotonal identity when misexpressed, namely *amos*, *cato* and *Math1*. These genes have highly similar basic domains to *ato*, thus it may be possible to pinpoint more precisely the residues important for this functional specificity. Moreover both *Math1* and *amos* can partially rescue the *ato* phenotype in the embryo, suggesting that these genes are directly promoting chordotonal identity. This has not yet been demonstrated for the more divergent *cato*. In fact it may well turn out that *cato* does not encode this property, requiring the activity of *ato* to promote chordotonal identity when misexpressed. If this turns out to be the case, then only a few residues of the basic region will be implicated for the inhibition of *cut*.

Other shared features of the bHLH *amos* and *ato* are implied by their roles during olfactory SOP formation, which must utilise some form of cell recruitment mechanism (Reddy and Rodrigues, 1997). Whether this mechanism echoes that used for chordotonal and ommatidial clustering has not yet been established. Significantly, *amos* can fully rescue the *ato* phenotype in the embryo (Huang et al., 2000) whereas *Math1* cannot. UAS-*Math1* misexpression in an *ato* background gives rise to a few chordotonal organs per segment (B. Hassan, unpublished data) whereas UAS-*amos*

misexpression generates the full lch5 (Huang et al., 2000) perhaps indicating that this protein can initiate the recruitment mechanisms normally orchestrated by *ato* (zur Lage, 1997). Thus sequence comparisons between these three genes may highlight molecular features important for neural precursor recruitment.

What distinguishes neural determination factors from differentiation factors?

By now, many neural bHLH factors in vertebrates have been identified and characterised. On the one hand, much like their fly counterparts, these genes have been described as having proneural function by virtue of their behaviour in *in vitro* gain of function assays in vertebrate systems (Chapter 1). However, although expression and functional analysis of several of such genes certainly indicates their fundamental requirement during neural precursor development they do not necessarily point to a role in the proneural 'specification' event itself (Chapter 1.5.4). Furthermore, the same gene is often used reiteratively at different stages of neurogenesis, in the same and other neural subtypes. Together these observations might lead one to conclude that it is developmental context which is the key factor in determining the specificity of these proteins during vertebrate development.

But intrinsic functional differences are evident. For example, although *Ngn* can induce the expression of *NeuroD*, the reverse is not so, indicating a unidirectional regulation between these two genes, reflected in their expression, and presumably their different sequences. *cato* appears to promote sensory neuron differentiation, whereas *ase* does not. At the same time, the notion (and evidence) that related bHLH proteins function in interlinked functional cascades and networks imply that these genes must encode different properties and specificities.

Until now, the paucity of known *Drosophila* neural differentiation bHLH factors has made it difficult to address fully the basis of functional specificity, and determine the relevance and roles of developmental context. The identification of genes such as *cato* and *atoid* as members of an *ato*-like subfamily, and the less related *tap,* present a challenge and an opportunity to discover how functional specificity is encoded in their protein structures. Given the specific differentiation misexpression phenotypes for *cato*, it may be interesting to determine whether such properties are similarly conserved in other neural bHLH factors, of vertebrates and invertebrates alike.

Finally, this analysis has presented another instance of an expression cascade of related bHLH proteins, adding weight to the idea that such cascades are generic to both vertebrate and invertebrate development. Indeed with this in mind, it will be interesting to establish whether a similar temporal link exists for the expression of *amos* with respect to *cato* and investigate whether such systems have functional relevance in *Drosophila*.

5.2 Possible experimental approaches to explore the bases of bHLH functional specificity.

There are many possible approached to assessing the different functions of these genes in neurogenesis. The tried and tested method of domain swop experiments and site directed mutagenesis (e.g. Chien et al., 1996) will be useful in pinpointing functional differences attributable to the bHLH region. Biochemical approaches, such as immunoprecpitation (pull-down) assays may be useful to assess the differences between Ato and Amos. Notably, Amos was isolated independently by yeast two hybrid interaction with Numb (Huang et al., in press). It will be interesting to determine the functional significance of this and whether it can be repeated with Ato. Searching for genetic interactions between these bHLH genes and known neural genes may indicate avenues to be explored and back up molecular interaction data.

5.3 Potential screening strategies for specific mutations

Clearly, the identification of specific mutations for *amos* and *cato* is paramount for effective future functional studies of these genes. The analyses presented here in terms of their structure, expression and function hint at possible strategies that may be used to screen for such mutations.

5.2.1 Possible mutagenesis strategies for *cato*

Although the neuron morphology phenotypes of *cato* deficient embryos suggests that *cato* larvae would be inviable, it may be naïve to assume that deletion of the gene causes lethality. Furthermore, screening for such relatively subtle morphological defects would be difficult. Indeed, several EMS alleles that map to the *cato* region have already been analysed on the basis of neuronal morphology. Although PNS defects and morphology were observed in several cases, there was no clear evidence

that any of these mutations disrupted *cato* function. It is also notable that mutations for similarly expressed genes, such as *ase,* are quite viable (Brand et al., 1993; Dominguez and Campuzano, 1993). Therefore, until more functional information is isolated for *cato,* such as the specific regulation of a particular target gene, efforts to screen for mutations may require a molecular approach. For instance, these EMS alleles could be examined using PCR to clone the *cato* ORF, and sequencing to detect any mutations. Another possibility may be to put the alleles in the background of the *ase* mutation to see if embryos deficient for both functions then have similar axon misrouting defects.

The pros and cons of using a P-element hopping strategy to generate specific mutations have been well rehearsed elsewhere (Greenspan, 1997). In fact, a local P-hop screen based on the close proximity of the *khc* locus to *cato* has already been exhaustively attempted in the lab. Sequence and PCR analysis indicated that a P-element disrupting the *khc* gene was ~7kb away from *cato.* Based on this knowledge, local hopping has been carried out using this stock, and the hops generated were analysed using a PCR strategy to see if new insertions were made in the *cato* locus (A.P.J., D. Prentice). Unfortunately this approach was deemed untenable, for although local hops were generated frequently, these always appeared to be in a direction 3' to the *cato* gene itself. Another strategy being used is the method of male recombination from an existing P-element, to generate smaller deletions of the region, in an attempt to create alleles for *cato*.

Given the obstacles already encountered in searching for mutations of *cato*, it might be more fruitful to analyse *cato* function further, in terms of morphological defects, or determining its downstream target genes.

5.1.2 Screening strategies for *amos* mutations

In the case of *amos*, it is unlikely that disrupting *amos* function in the adult would cause lethality. Null alleles for the proneural *ato* are viable, though poorly (Jarman et al., 1995).

At present an EMS mutagenesis screen is underway in the lab (A.P.J., D.P.) that is based on three assumptions. The first two criteria rely on the prediction that *amos* alleles will be viable. The evidence suggests that in the absence of *amos*, at least one,

if not two, subsets of adult olfactory sensilla will be missing. Mutations will be screened on this basis, it being just possible to note the presence or absence of these sensilla under the dissecting microscope. Another possibility is that *amos* may be required for the formation of the leg tarsal claw. The evidence for such a role is somewhat circumstantial however, given the regulatory link proposed for *amos* and *lz*, it is significant that some strong *lz* alleles are missing the tarsal claw (Stocker, P.z.L. pers. comm.). Mutations will also be screened for the more obvious phenotype that, without the claw, flies are unable to crawl up vertical surfaces.

The third criterion is that *amos* function in the embryo, or a non-neural function earlier in development, may be lethal. For example, it remains a possibility that *amos* function in the embryo may give rise to many of the gustatory and olfactory organs of the larva, without which the larva may starve. If *amos* is required for a specific subset of larval olfactory organs that sense specific odours and tastants, a behavioural screen could be designed on this basis. In fact, only recently very simple assays have been described for analysis of larval olfactory behaviour (Heimbeck et al., 1999). Finally, given the requirement of *amos* for the formation of the distinctive dbd neurons, an F2 screen of embryos missing these neurons could be attempted. Recently, GFP lines specific for different subsets of larval neurons have been described (Gao et al., 1999). One such line was used in the identification of genes required for dendritic branching in the larva (Gao et al., 1999).

Conclusions

The identification of these *ato*-like genes, *amos* and *cato* will allow future investigations of the basis of olfactory sensilla formation and function as well as the later roles of bHLH proteins in neurogenesis.

Chapter 6

Materials and Methods

6.1 Media and solutions

Table 6.1.1 Media recipes

Media Type	Components (Per litre)
L-Broth (Luria-Bertani Broth)	10g Difco Bacto-tryptone, 5g Difco Bacto-yeast extract, 5g Sodium Chloride, adjusted to pH 7.2
L- Agar	16g Difco Bacto-tryptone, 10g Difco Bacto-yeast extract, 5g Sodium Chloride, adjusted to pH 7.2
BBL-Bottom	10g Baltimore Biological Laboratories Trypticase, 10g Difco Bacto agar, 5g Sodium Chloride
BBL-Top	10g Baltimore Biological Laboratories Trypticase, 6.5g Difco Bacto agar, 5g Sodium Chloride
SM buffer	5.8g Sodium Chloride, 2g Magnesium Sulphate, 50mM Tris-HCl (pH 7.5), 0.01% Gelatine
2 x YT	16g Bacto-tryptone, 10g Bacto-yeast extract, 5g Sodium Chloride
Top agarose	0.6g agarose in 100ml LB, 1ml 1M $_{MgSO4}$

Table 6.1.2 List of general solutions and their composition

Solution	Composition
1 x TE	10 mM Tris-HCl, 1mM EDTA, adjusted to pH 8.0
10 x TBE	0.89M Tris-Borate, 0.89M Boric acid, 10mM EDTA
20 x SSC	3M Sodium Chloride, 0.3M Tri-Sodium Citrate, adjusted to pH 7.0
0.5M EDTA	0.5M Diaminoethanetetra-acetic acid, adjusted to pH 8.0
10 x MOPS	0.2M Sodium-MOPS, 50mM Sodium Acetate, 10mM EDTA, adjusted to pH 7.0
100 x Denhardts solution	2% (w/v) Bovine Serum Albumin, 2% (w/v) Polyvinylpyrrolidone (ml. wt. 400 000), 2% (w/v) Ficoll (ml. wt. 400 000)
Herring sperm DNA (purchased from Sigma)	10 mg/ml stock solution. Sonicated and heat denatured.
RNA Formaldehyde Sample Buffer (FSB)	50 % (v/v) Formamide, 25% (v/v) Formaldehyde (at 14.8% w/v), 25% (v/v) 10 x MOPS buffer
10 x DNA gel loading buffer	0.25% (w/v) Bromophenol Blue, 0.25% (w/v) Xylene Cyanol FF, 30% (v/v) Glycerol
10 x RNA gel loading buffer	0.25% (w/v) Bromophenol Blue, 0.25% (w/v) Xylene Cyanol FF, 1mM EDTA (pH 8.0), 50% (v/v) Glycerol
RNA denaturing buffer	4M Guanidine Thiocyanate, 42mM Sodium Citrate, 0.83% (w/v) Lauryl Sarcosine, 0.2mM β-mercaptoethanol
Phenol/Chloroform	Phenol was purchased from SIGMA (re-distilled and pre-equilibrated with 100mM Tris-HCl pH 8.0) and mixed with Chloroform and Isoamyl Alcohol in the ratio 25:24:1 respectively. To prevent oxidation, 8-Hydroxyquinoline was added to 0.1% (w/v) and the solution stored at 4°C in the dark
25mM dNTP stock solution	Made by dilution of 100mM stocks of dATP, dCTP, dGTP, and dTTP nucleotide solutions (pH 7.0) purchased from Pharmacia

5 x T4 DNA Ligase buffer	5mM ATP, 20mM Tris pH7.6, 50mM $MgCl_2$, 50mM DTT, 500μg/ml BSA Stored at –20°C
Deoxyribonuclease I (DNAase I)	20 mg/ml in 50 % (v/v) Glycerol, stored at –20°C
Ribonuclease A (RNAase A)	20 mg/ml in 50 % (v/v) Glycerol, boiled for 5 min to inactivate contaminating DNAase's, and stored at –20°C
Proteinase K	20 mg/ml in 50% (v/v) Glycerol, stored at –20°C
Lysozyme	8 mg/ml in sterile double distilled water, used fresh
10 % SDS	10% (w/v) Sodium Dodecyl Sulphate in sterile double distilled water. Filter sterilised
TM buffer (small scale phage DNA extraction)	50mM Tris-HCl (pH 7.5), 10mM Magnesium Sulphate
DEPC-water	Sterile distilled water rendered RNAase free by treating with Di-ethyl Pyrocarbonate
Ethidium Bromide	10mg/ml in sterile distilled water
STET	5% sucrose, 0.5% TritonX100, 50mM EDTA, pH 8.0, 50mM Tris-HCl pH 8.0
20x TAE	Tris base, 96.8g, Glacial acetic acid 22.8 ml, 0.5M EDTA ph8.0, 40ml, made up to 1l with H_2O
20xTBE	Tris base, 216g; Boric acid, 110g; EDTA 18.6g to 1l with H_2O
10x PBS	NaCl 80g, KCl 20g; Na_2PO4 0.2g; KH_2PO_4 20g to 1l with H_2O
PBTX	1xPBS with 0.3% TritonX100
PBTW	1xPBS with 0.1% Tween20

6.1.3 Suppliers

Chemicals were obtained from SIGMA, BDH, Aldrich, Boehringer Mannheim, and Fisons. Enzymes were supplied by Boehringer Mannheim and Promega. Commercial kits were purchased from Pharmacia, Qiagen, Promega and Boehringer Mannheim.

6.2 Manipulations of bacteria and bacteriophage

6.2.1 Bacterial cultures

The medium used for culture of *E. coli* was autoclaved Luria-Bertani (LB) broth or occasionally a richer broth like 'superbroth' which supported higher bacterial populations and so gave higher yields in bulk preparations of plasmids. Liquid cultures were grown by incubation at 37°C in an orbital shaker, or single colonies were streaked out on solid LB agarose medium and incubated at 37°C overnight. Media were supplied with appropriate nutrients and antibiotics after sterilisation. Ampicillin was added to a concentration of 50μg/ml, sometimes alternated with the higher 100μg/ml concentration to reduce likelihood of resistant bacterial populations.

Stocks of *E. coli* strains and of transformants were stored either as colonies on agar plates at 4°C for short periods (up to one month), or as 50% (v/v) glycerol mixtures at –20°C or –80°C.

6.2.2 Transformation of *E. Coli*

Competent cells used for transformation were prepared using a simple, rapid $CaCl_2$ procedure (Sambrook et al., 1989). Such cells gave adequate transformation efficiencies for the purposes of subcloning (~10^5/μg DNA). Cells were used either at least 2 hours after preparation, or were stored as small (200 μl) aliquots in 50% glycerol at –80°C.

For transformation, between 10-100ng of DNA in a small volume (up to 10μl) of TE, H_2O or ligation buffer was mixed into 200μl of competent cells, which were left on ice for 45 min for adsorption. After a brief heatshock (42°C, 2 mins) for DNA uptake, cells were allowed to recover on ice for a further two minutes, then added to a small volume of LB broth (up to 1 ml) and incubated with gentle shaking at 37°C for up to one hour (expression phase). The cells (max. 200μl/plate) were then spread onto LB plates containing appropriate antibiotic and were incubated at 37°C overnight (12-16 hrs). For blue/white selection, 100μl 100mM IPTG and 20μl 50mg/ml X-gal were added to agar plates prior to cells.

Single colonies were picked using a sterile loop and grown up in 10mls of LB+ antibiotic for screening.

If higher transformation efficiencies were required, commercially available strains of cells were used (Appendix 2) and the appropriate transformation protocol followed according to manufacturers instructions (Stratagene, Promega, Invitrogen).

6.2.3 Preparation of plating cells for λ bacteriophage infection

Cells of the appropriate strain were streaked out on LB-agar plates and grown overnight at 37°C. A single colony was used to inoculate 10ml of LB, which was incubated overnight at 37°C. 1ml of this was used to prime a large scale culture of LB-broth supplemented with 0.2% Maltose and 10mM $MgSO_4$ which was grown at 37°C for ~6 hrs and then cells harvested by centrifugation (4,000 rpm, 10 min, 4°C). Cells were resuspended in 10mM $MgSO_4$ and stored for several days at 4°C.

6.2.4 Plating and titration of λ bacteriophage stocks and libraries

Prepared packaged bacteriophage λ (phage) particles in phage buffer were kept at 4°C. Serial dilutions of phage solution from 10^{-2} through to 10^{-8} were prepared in phage buffer. To infect cells, 10 μl of each dilution were added to 100μl of plating cells and incubated at 37°C for 15 min. 3mls of molten BBL-top agar (supplemented with 10mM $Mg2^+$ ions) cooled to 45°C were then added to each dilution and subsequently poured onto L-agar plates. When set, the plates were incubated overnight at 37°C. To calculate the titre of the phage stock, the number of plaques formed on each plate was counted: each plaque represents a single infecting phage particle. Units are therefore calculated as plaque forming units, or PFU, per ml.

6.2.5 Growth of λ bacteriophage on agar plates

Phage stocks were diluted to the desired PFU/ml and added to a quantity of plating cells (typically 100μl per 100pfu) and adsorbed at 37°C for 15 mins. 3 mls of molten BBL-Top agar was added, the mixture poured over agar plates and allowed to set. The plates were then incubated at 37°C until plaques were 1mm in diameter (6-8h).

A similar procedure was used to establish large plates for screening purposes (23x23cm). Approximately 300,000 PFU were mixed with 3mls plating cells and adsorbed as before. 30ml of molten top agar was added (this time the more robust agarose-top agar), supplemented as before, and poured onto a large BBL-agar plate and allowed to set. Plates were incubated until near-confluent lysis was achieved. All

plates to be screened were placed at 4°C for at least 1hr to firm the top agar to prevent sticking to nylon membranes.

6.2.6 Isolating cultures from individual plaques

After screening, individual phage were scored from the agar using the wide end of a 1ml pipette tip. The agar pieces were transferred to 500 ml of SM buffer in an eppendorf, and vortexed with a few drops of chloroform added to prevent bacterial growth. These liquid stocks were kept at 4°C. Stocks were titred as described in 6.2.4.

6.2.7 High-yield DNA preparation from liquid phage lysates

The method of Burmeister and Lehrach (Burmeister and Lehrach, 1996)was used for isolating large scale cultures of individual phage strains (recombinants).

A few μl of the particular phage stock was streaked onto an L-agar plate as if streaking for single colonies, and allowed to dry. Fresh plating cells were added to molten top agar, supplemented with 10 mM Mg2+ and then poured over the phage on the plate which was incubated overnight at 37°C. Complete lysis was usually visible in one area of the plate, and other regions contained only single plaques. A single plaque, and the bacterial growth surrounding it was scored from the agar plate using the thin end of a pasteur pipette and placed directly into 200mls of LB + 10mM $MgCl_2$. The liquid culture was incubated overnight at 37°C in a 1000 ml flask. The next morning the culture had lysed, with bacterial debris floating as stringy aggregates. A small volume of chloroform was added and the culture then centrifuged to remove bacterial debris. The lysate was then decanted into 50ml sterile falcon tubes, a little chloroform added, and stored at 4°C, until required.

6.3 Preparation of DNA

DNA was resuspended in either TE or sterile or DEPC-treated H_20, depending on future usage.

6.3.1. Genomic DNA

Insect genomic DNA was prepared using a commercial kit, (Puregene, Gentra). Typically, the yield from 20-30 flies was between 10-40μg DNA.

6.3.2. Plasmid preparation procedures.

Plasmid DNA was prepared by a variety of methods, depending on the later application of DNA. For cloning purposes, to screen transformants for correct plasmids, quick boiling lysis or alkaline lysis methods were used (see Sambrook et al., 1989). For sequencing purposes, cleaner DNA was prepared using commercial spin column (B.M, Pharmacia). Larger scale plasmid preps (to isolate fragments for subcloning) were purified using Qiagen midiprep procedure.

For DNA injection of embryos to generate transformants, the large scale bulk prep procedure detailed below was used.

Plasmid bulk prep protocol

Bacterial cultures were transfered to 50 ml Falcon tubes and centrifuged at 5000 rpm for 15 min. The pellets were drained thoroughly and resuspended carefully using a pastette in 2 ml of Solution I per 50 ml of culture and left at room temperature for 10 mins. Then, 4 ml of Solution II was added and mixed thoroughly but not vigorously. After 10 min on ice, with mixing occasionally, 3 ml Solution III were added. At this point the solution was mix quickly and thoroughly and placed on on ice for 15 mins. The mixture was centrifuged at 7000 rpm for 15 min. The supernatant was transfered (strained through inverted serological pipette) to a clean tube, avoiding transfer of any precipitate. The pellets were discarded. 0.6 volumes of 100% isopropanol was added and then let stand for 5 min followed by centrifugation at 4000rpm, 10 min. The supernatant was discarded and pellet briefly rinsed with ~2 ml 70% EtOH. The tube was wiped out and the still wet pellet dissolved in 1 ml H_2O or TE. At this point the DNA solution was transfered to eppendorfs and placed on ice for 5-10 min. An

equal volume of cold 5M LiCl. (LiCl stored at –20°C) was added, and placed again on ice for 5 min. The tubes were centrifuged for 5 min (14,000 rpm). The supernatant was transferred to clean eppendorf tubes (on ice) and an equal volume of isopropanol was added. This was placed on ice for 10 min then spun for 5 min. The supernatant was discarded and the pellets air dried. These were then redissolved in a total of 300 µl TE. 1 µl DNase-free RNase (10 mg/ml stock)was added an then the mixture incubated at 37 C for 30 mins and then transferred to ice. An equal volume of PEG/NaCl (15% PEG, 1.6 M NaCl) was added, and the tube left on ice, 5 min and spun for 5 min. All the supernatant was discarded and then pellet resuspended in 300 µl TE. A PhOH/$CHCl_3$ extraction was carried out (add equal vol PhOH/$ChCl_3$, vortex, spin, remove top aqueous phase) followed by a $CHCl_3$ extraction. The aqueous phase was transferred to clean tube(s). The DNA was precipitated by addition of 1/20 vol 3M NaOAc (pH5.6–6) and 2 vol 100% EtOH, This was mixed thoroughly, but vortexing avoided. A large precipitate was usually seen. The tubes were centrifuged for 5min. The pellets were washed with 70% EtOH and air dried, then the pellets were resuspended in ~500 µl TE.

Solutions

Solution I : 50 mM Glucose, 25 mM Tris pH 8, 10 mM EDTA, 5 mg/ml lysozyme, added just before use

Solution II: 0.2 M NaOH, 1% SDS (Made just before use from 10 or 1 M NaOH and 10% SDS stocks)

Solution III: 3 M KOAc/1.3 M HCOOH

6.3.3 Phage DNA preparation

Phage DNA was isolated using Qiagen midiprep procedure for liquid phage lysates.

6.4. Polymerase Chain Reaction (PCR) procedures

6.4.1 Standard and degenerate PCR

Several PCR procedures were used. Degenerate PCR was employed to identify new genes in *D. melanogaster* and other dipteran species. PCR was also used to amplify fragments for subcloning, and to check for the presence of specific sequences of DNA (or RNA) in libraries. In all cases, primers were initially tested using appropriate positive control template DNA. Melting temperatures were estimated using the formula ([4x(G+C) content] + [2x(A+T) content], and this T_m estimate was used as a starting annealing temperature. Boehringer Mannheim and Promega *Taq* polymerases and their supplied buffers were used. Generally a standard magnesium chloride concentration of 1.5mM was used, but titrations were carried out where necessary. Between 30-50 pmol of each primer was used per 50μl reaction, and 10nmols of dNTPs. For most applications, standard PCR cycling parameters of initial denaturing step 96°C 3min, followed by 30 cycles of 96°C,up to 1 min, 50-60°C annealing for up to 1min, extension at 72°C for 1min, followed by final extension cycle of 10 min, then left at 4°C. More complex cycles were used for longer primers.

6.4.2 3'RACE and RT-PCR

All RACE and RT-PCR applications were carried out using the Promega 'Access RT-PCR' kit according to instructions. Total and polyA+ RNA stocks from early embryo collections were used as template.

6.5 Isolation of total cellular and polyadenylated RNA

mRNA was isolated by guanidine thiocyanate methods and purified using the PolyATract system III (Promega).

6.6 DNA/RNA manipulation and detection

6.6.1 Phenol extraction of proteins from DNA/RNA

An equal volume of phenol/chloroform was added and the solution mixed by vortexing for 1 min. The sample was centrifuged for 5 min at 14,000 rpm, and the aqueous phase transferred to a fresh tube. A chloroform:Isoamyl alcohol (24:1) extraction was performed, to remove traces of phenol, and the aqueous phase was transferred to a new eppendorf.

6.6.2 Precipitation of nucleic acids

Nucleic acid was generally precipitated by the addition of 1/10 volume NaOAc (3M, pH 5.2-5.8) and 2.5 v/v 100% EtOH (or 0.6 or equal vol. of isopropanol). The solution was mixed thoroughly and placed on ice (or –20°C) for at least 15 min, then pelleted by centrifugation (14 000 rpm, 10 min). The pellet was washed with either 70% or 80% EtOH and then dried at room temperature (RT). The sample was then dissolved in an appropriate volume of TE or sterile H_2O, depending on subsequent applications.

6.6.3 Separation of DNA fragments by gel electrophoresis

DNA and RNA was analysed using standard agarose gel electrophoresis as described. Typically, 0.8% agarose in 1xTAE or 0.5xTBE gels containing 0.5μg/ml EtBr solutions were prepared. For separation of low molecular weight fragments, 2% agarose gels were used. 1xTAE gels were used preferentially for excision of DNA for further purification. Before loading, DNA was mixed with standard 1/6 vol FDE mix (150 g/l ficoll 400; 2.5g/l bromophenol blue). Gels were run at 80V for 30 min to 1h.

6.6.4 Purification of DNA

DNA fragments or PCR products were purified by gel extraction followed by spin columns. The DNA to be purified was run on an appropriate 1xTAE agarose gel and the desired band excised, and weighed. The DNA from the gel slice was extracted using a commercial kit procedure (Pharmacia, Boehringer Mannheim). DNA was resuspended in H_2O for ligation.

6.6.6 Estimation of nucleic acid concentration

The concentration of nucleic acid solutions was determined, either by gel electrophoresis (EtBr), with reference to molecular weight standards (Boehringer Mannheim, Bioline) or by spectrophotometry. Absorbence of a sample was read at 260nm in a quartz cuvette. For assessment of nucleic acid purity, an A260:A280 ration was determined. A ratio of greater than 1.8 indicated the solution was free from protein contamination.

6.6.7 Enzymatic reactions

Restriction Digestion

For site specific DNA digestion, restriction endonucleases were used according to manufacturers instructions. Typically, to check plasmid DNA, approximately 3 units of enzyme per μg DNA was incubated in the appropriate buffer for up to one hour at 37°C. For larger scale digestions, reactions were scaled up and allowed to digest overnight.

5' dephosphorylation

The restriction of DNA yields blunt or sticky ends which may be religated. To prevent recircularisation in ligation reactions, vectors were generally dephosphorylated at their 5' ends using calf intestine phosphatase (CIAP) according to the following protocol: DNA in either digestion buffer, or a tris buffer, was incubated with CIAP (at concentration of 0.01U per pmol of 5' termini) for 30 minutes at 37°C, and a further 1μl of CIAP was then added for another 30 min incubation period. Enzymes were then removed by phenol/chloroform extraction followed by EtOH/Ch3CooNa precipitation for nucleic acids

Ligation

To promote ligation between vector and insert fragments, established molar rations were used for reactions. Generally a high 3:1 ratio of insert to vector (I:V) was chosen, where the size of the insert was substantially smaller than that of the vectors. For roughly equal fragment sizes, an equimolar 1:1 ratio was attempted. Generally a range of ratios was set up to obtain transformants.

Ligations were carried out in ligase buffer (1x) using 1Weiss unit of T4 ligase in a total volume not exceeding 15 μl. Ligations were performed at between 12-16°C, overnight.

6.7 Nonradioactive filter hybridisation procedures

Unless otherwise stated, standard protocols were followed using guidelines provided in the Boehringer Mannheim's publication 'The DIG System User's Guide for Filter Hybridisation (1995). Southern blotting techniques were used for analysis of DNA fragments or PCR products separated by gel electrophoresis. Colony and plaque hybridisation methods were used for screening genomic and cDNA libraries, and detecting positive plasmid colonies in a heterologous background. A P1 gridded array filter was screened using standard protocols.

In summary, nucleic acids were transferred to filter membranes (Hybond-N, Amersham) and covalently bound by UV crosslinking (Stratalinker) and later hybridised using a DIG-labelled probe of interest. Excess probe was washed off the filters which were then incubated with an anti-digoxygenin antibody alkaline phosphatase conjugate (Boehringer Mannheim). The signal was detected with a chemiluminescent alkaline phosphate substrate (CSPD, Boehringer Mannheim) by autoradiography.

6.7.1 Synthesis of Digoxygenin (DIG)-labelled probes

DIG-labelled probes were prepared using a variety of methods, using Boehringer Mannheim's kits for nucleic acid labelling. The majority of probes used were DIG-RNA probes, produced using T7/T3-mediated transcription for the synthesis of strand specific RNA probes. In addition, digoxygenin-11-UTP was incorporated into DNA by PCR or the random-priming labelling methods. Table 6.7.1 lists the probes prepared and applications for which they were used. Probe concentrations and hybridisation temperatures were selected using guidelines set out in the manual.

Table 6.7.1 DIG-labelled Probes preparation and applications

Tpye	Template	Application
dsDNA bHLH product	*cato* and *amos* PCR bHLH clones	genomic S-Blot
ssRNA bHLH product	as above	genomic S-Blot Library screens
antisense RNA *amos* sense RNA *amos*	pBS-SalIAA pBK-CMVA1,A6 or A9	mRNA ISH P1-Blot N-blot
antisense RNA *cato* sense RNA *cato*	pBS-SacIBE pBK-CMVB1	mRNA ISH P1-Blot N-blot
dsDNA *amos* genomic region	Phage AA	polytene chromosome ISH
dsDNA *cato*genomic region	Phage BF	polytene chromosome ISH

6.7.2 Southern blotting

Low percentage agarose gels (0.6%) were used for genomic and recombinant phage clone southern blots. For genomic DNA, approximately 5-10μg of restriction digested DNA was loaded per well. For subcloned fragments and PCR products far less was loaded, approximately 100ng. DIG-labelled DNA markers were also loaded to provide size references. Before transfer, the agarose gels were stained with EtBr and photographed for later reference. To enhance the transfer of large DNA fragments, depurination was carried out in some instances.

6.7.3 Northern blot

For northern blotting, formaldehyde denaturing gels were prepared. Typically gels with a final percentage agarose concentration of between 1.3 and 1.5 were used. Gels were prepared by dissolving agarose in 1xMOPS (in the microwave), cooling to approximately 50°C and adding formaldehyde to 17.3% (v/v), such that the final volume gave the desired agarose concentration. Gel solutions were then mixed by swirling and poured into prepared gel trays, and allowed to set for a minimum of 30

min. Gels were run in 1xMOPS buffer at similar voltages to those employed for electrophoresis of DNA.

RNA samples were prepared for loading by heat denaturing at 65°C for one minute and adding an equal volume of formaldehyde sample buffer (FSB). EtBR (5ng per sample) and RNA loading buffer (to 1x) were added prior to loading.

Formaldehyde denaturing RNA gels do not require denaturation or fixation, however before transfer, formaldehyde was extracted from the gel by washing gently in sterile H_2O for 30min prior to blotting.

6.7.4 Colony and Plaque library screening protocols

Plates were chilled at 4°C for one hour before transfer to membranes. Duplicate lifts were taken to aid identification of positives. The protocol was followed as described in the Boehringer Mannheim manual.

6.8 Sequencing of double stranded DNA templates

6.8.1 Sequence reactions

All sequence reactions were performed using BigDye Dye terminator kits (Perkin Elmer), and analysed using an ABI automated sequencer (service provided at ICMB). Individual sequence reactions contained 8.0μl of reaction mix, between 250-500ng dsDNA template, 3.2pmol of primer and sterile H_2O to a final volume of 20μl. It was found that these reactions could be halved with no detrimental effects to the quality of sequence data. Reactions were prepared in a 0.5μl eppendorf and cycled in a PCR machine using the cycle recommended for the particular kit. In this case, the programme used comprised a 3 min 96°C step followed by 25 cycles of 96°C for 30sec, 50°C for 15 sec and 60°C for 4 mins. Subsequent to cycling, extension products were purified by EtOH precipitation to removed the unincorporated nucleotides.

6.8.2 Analysis of sequence data

Sequence information (trace and text files) were analysed using the GCG package and Genejockey programmes. Sequence contigs were assembled using Genejockey, and sequence alignments (DNA and amino acid) were generated using the Clustalx

algorithm. (Program manual for the Wisconsin Package, version 9.0, available from the University BioServer).

6.9 *in situ* hybridisation and immunohistochemical procedures

6.9.1 Salivary gland polytene chromosome *in situ* hybridisation (ISH)

Preparation of chromosome squashes.

For larvae, 8-10 adult female flies were placed in a fresh bottle supplemented with fresh yeast paste. Bottles were kept at 18°C to encourage large larvae, and minimise overcrowding. Larvae close to pupal formation were chosen for dissection. Fresh fix was prepared (1:2:3 lactic acid:water:glacial acetic acid) and three to four larva were dissected at one time in this solution in a watchglass. Mouthhooks and rear were grabbed with forceps, and the larva pulled to extricate the glands. The best salivary glands were those which were turning transparent; whitish ones were too young, and the chromosomes therefore too thin. If cell borders were seen, that indicated the chromosomes were too old, and likely to be brittle. The lateral fat was dissected off the glands. The glands were sucked up in 7μl fix with a pipette, and placed on a glass slide. A siliconised 22x22 mm coverslip was placed over the glands and held in place with one finger. The rubber-end of a pencil was used to tap the coverslip in a spiralling motion and then spread the glands by pressing firmly in a side-to-side motion. The slide was examined under 10x phase lens of a light microscope and allowed to sit 2-5 min. To squash the glands, a tissue was placed over the slide and then was pressed hard with the thumb for 1 min. The slide was left overnight. The next day, prepared squashes were plunged into liquid nitrogen and then taken out, breathed upon, and the coverslip immediately popped off with a razor blade. Each slide was then immersed in dry-ice-cold 95% EtOH for 5 min. The slides were then transferred to room temperature 100% EtOH for 5 min and then air-dried. The chromosome squash preparations were then ready for ISH.

Hybridisation and detection protocol

The chromosome squash slides were first denatured by incubation in 2xSSC at 65°C for 30min, followed by 2xSSC at RT for 2-10 min and an incubation in 70mM NaOH for 3 min. The slides were then rinsed in 2xSSC twice for 5 min. Slides were

then dehydrated by soaking in 70% EtOH (2x5min) followed by 95% EtOH (2x5 min). The slides were air-dried and ready for hybridisation.

DIG-labelled DNA probes were prepared from large templates (typically 15kb inserts of lambda recombinant clones). The probe was diluted to 100ng/ml in hybridisation soultion (2xSSC, 50% deionised formamide, 12.3mM Tris pH 7.5, 0.6M NaCl, 5x Denhardt's solution, 1mM EDTA, 30μg/ml salmon sperm DNA, 10% dextran sulphate). The probe solution was heat denatured and 15-40μl was added to a siliconised coverslip. The coverslip was inverted and placed onto the slide. Slides were then laid flat in a humidified chamber (sealed box with wet paper towels) and placed at 45°C overnight.

Following hybridisation, coverslips were removed by moving slides gently in a beaker of 2x SSC. Slides were washed in 2xSSC at 53°C (3x20 min) and then rinsed in PBS at RT (2x5min). The slides were further rinsed for 2 min in PBT, in PBS (2x5 min). Anti-DIG horse radish peroxidase conjugate (HRP) was diluted to 1:50 in PBS with 5% NGS and 100-200μl of this solution was added to the slides, and covered with a coverslip to prevent drying. Slides were incubated with the antibody for one hour at RT and then rinsed in PBS (3x5 min) followed by PBT (2 min) and further PBS rinses (2x5min).

DAB solution was freshly prepared using the vectastain kit (1:2:1 drops Buffer:H_2O_2:DAB substrate). The DAB mix was added to the slides which were watched to prevent drying. A signal was usually detectable in the range 5min-1hr. Slides were rinsed in PBS and then counterstained with Giemsa (5% Giemsa in 50mM NaPi, pH 7.2) for 5 mins before washing with H_2O. Slides were then examined with phase contrast microscopy for signal. Slides were stored at 4°C for some time.

6.9.2 *in situ* hybridisation to mRNA in wholemount tissues

For *in situ* hybridisation to mRNA in wholemount embryos and imaginal discs, the protocol of Tautz and Pfeifle (Tautz and Pfeifle, 1989) was used, with modifications. DIG-labelled RNA probes were prepared using the Boehringer Mannheim digoxygenin RNA-labelling kit. Probes were used at 1:500 or 1:1000 dilution.

Embryos or imaginal discs were transferred from storage EtOH (6.10) and rinsed in PBTW. The samples were postfixed in 4% formaldehyde in PBTW for 15-20 min,

with gentle shaking. To remove fix, PBTW rinses and washes were carried out (5x5 min). The samples were then washed in 1:1 PBTW: hybridisation solution (HYBE) (50% deionised formamide, 5xSSC, 100μg/ml tRNA, 50μg/ml heparin, 0.1% Tween20, pH 6.5) for 10 min, followed by a wash in HYBE for a further 10 min. The samples were then prehybridised for at least 2h at 70°C on a heating block. After prehybridisation, the probe in HYBE (150μl was sufficient) was added to the samples which were incubated at 70°C overnight.

The following day, the probe solution was removed and a series of washes carried out at 70°C: HYBE (1x30 min), 1:1 HYBE:PBTW (1x30 min), PBTW (4x30 min). Following the washes, samples were incubated in 1:2000 anti-DIG antibody alkaline phosphatase conjugate (anti-DIG-AP) in PBTW for 2h at RT on a rotating wheel. For lowest background, the antibody was preabsorbed over embryos overnight at 4°C at 1:50 and then stored at 4°C. The excess antibody was removed by washes in PBTW (3x20 min). Samples were then transferred from eppendorfs to a microtitre well dish and rinsed several times in reaction solution (100mM Tris, pH 9.5, 100mM NaCl, 50mM MgCl2). A colour reaction was performed by adding a solution containing NBT and X-phosphate (4.5μl and 3.5μl respectively, per ml of reaction solution) to the samples and allowed to develop in the dark. The colour reaction typically took between 5 min to 2 hrs to develop.

Once the purple colour reaction was strong enough to visualise, the reaction was stopped by addition of 200mM EDTA in PBTW and rinsing several times. Samples were mounted as described in 6.10.

6.9.3 Immunohistochemistry and immunofluorescence

For immunohistochemistry embryos and imaginal discs were transferred immediately after the fixation procedure (6.10) to PBTX. Samples were washed thorou-ghly (3x20 min) in PBTX to remove traces of methanol and then blocked in 2% bovine serum albumin solution (BSA in PBTX) for at least 2h at RT, with gentle shaking. Primary antibody was then added to the appropriate dilution, supplemented with NGS and 2%BSA and the samples were incubated at 4°C overnight. Table 6.9.1 lists the antibodies used, and their concentration. The following day, samples were rinsed several times and washed in PBTX (4x15 min). The antibody solution was retained for future use (by addition of Sodium Azide to a concentration of 0.01%) and stored at 4°C. The secondary antibody was then added in PBTX, (either an HRP

or flourochrome conjugate) to a concentration of 1:500, for a period of 2h at RT. Excess antibody was washed away by several PBTX rinses and washes (3x15 min). For immunofluorescence, samples were mounted (dissected first in the case of imaginal discs) in vectashield medium and stored in the dark at 4°C. For HRP-conjugated antibodies, samples were transferred to microtitre well dish in PBTX and then rinsed several times in PBS to remove the triton. Samples were then incubated in freshly prepared DAB as described in 6.9.1) for a short time, typically less than 5 min. To stop the DAB reaction, samples were rinsed in PBTX several times, then dissected and mounted as described (6.10).

6.9.4 Double labelling for RNA and protein

For double labellings, the RNA *in situ* hybridisation protocol was carried out first, then the samples were transferred to PBTX and protein *in situ* procedure then continued.

Table 6.9.1 Primary antibodies used and their concentrations

Antibody	Concentration	Staining Pattern	Reference/Source
RbαAto	1:2000	chordotonal PNC, SOP	Jarman et al., 1994
RbαAse, GpαAse	1:500	neural precursors	Brand et al., 1993
MαPros	1:4, 1:10	outer support cell (scolopale, sheath)	C. Doe
αCpo	1:500	sense organ cells	H. Bellen
MAb1188	1:500	scolopale, cap and neuron cells	Jarman et al., 1993
MAb22C10	1:100	neuronal cell body and processes	Zipursky et al., 1984
RαHRP	1:50	neuron and scolopale	
Mαelav	1:100	neuron cell	Developmental Studies Hybridoma Bank, University of Iowa
Mαcut	1:100	external sense organ outer support cells	as above

6.10 Fixation, Preparation and Mounting procedures

6.10.1 Fixation of embryos and imaginal discs

For staining procedures, embryos were collected on grape juice agar plates and pipetted into a sieve. The embryos were then dechorionated in 50% bleach for 3-5 min and transferred to a prepared scintillation vial containing a 1:1 mix of 4% formaldehyde in PBS and heptane. Embryos were fixed for 20 mins and the fix replaced with two volumes of methanol, for devitellinisation. The vial was shaken vigorously for 30 sec and let stand. Embryos which fell to the bottom of the vial were transferred to an eppendorf and rinsed quickly with MeOH, to remove residual heptane. For immunohistochemical techniques, embryos were transferred to PBTX. In the case of *in situ* hybridisation to mRNA, the embryos were stored in EtOH at –20°C, until required.

Larval or pupal imaginal discs were dissected in PBS, then fixed for 1h in 4% formaldehyde. The discs were then rinsed in PBTW (x3) and either used directly for immunohistochemistry, or put through an EtOH series to dehydrate them for storage in 100% EtOH at –20°C.

6.10.2 Hand devitellinisation of fixed embryos

Embryo injection creates a hole in the vitelline membrane thus precluding the methanol method to remove it. Therefore to devitellinise them, the fix was first removed from the vial and the embryos rinsed several times in heptane to remove residual fix solution. Using a long glass pasteur pipette embryos were sucked and transferred onto an upturned fine gauge sieve stuffed with tissue paper. The excess heptane was blotted away from the sieve using tissue. Then a piece of double-sided sellotape was pressed gently over the sieve, to pick up the embryos. Once all the embryos were stuck to the tape, the sellotape was stuck onto a small plastic petri-dish lid, embryos facing up! PBS was then quickly poured over the embryos. The transluscent reflective vitelline membrane was then torn away from each embryo individually using a mounted fine needle and forceps. Once all the embryos has been devitellinised, the PBS was replaced with PBTX and the staining procedure initiated After staining, the embryos were mounted individually as described in 6.10.4.

6.10.3 Preparation of adult tissues

To observe external and internal sense organs, flies were partially dissected and fixed in 4% formaldehyde in PBT overnight. After rinsing, wings, thoraces etc., were dissected off the carcass, and extraneous tissues were removed. When necessary, staining procedures were carried out as described, then body parts were further dissected and mounted in glycerol on glass slides and examined by DIC miscroscopy.

For scanning electron microscopy (SEM), flies were killed by freezing, mounted on stubs, and then air-dried before being sputtered with gold (service by John Findlay, ICMB).

6.10.4 Mounting samples

Embryos were mounted in 80% glycerol in PBTX on microscope slides sealed with a coverslip and nail varnish. Imaginal discs were dissected away from larval tissues and spread out in a small volume of 80% glycerol and mounted similarly.

6.10.5 Mounting individual embryos.

For some stainings, very few embryos were available, and a particular orientation was necessary to analyse the results. To this end, embryos had to be mounted individually. In this procedure, embryos were first dehydrated through an EtOH series of 5 min washes for each of 70%, 95%, 100% and then 2x5 min washes with Ultradry EtOH. These washes were carried out in a watchglass with gentle agitation to keep the embryos swirling in the centre. The embryos were then infiltrated with acetone for 5 min, and then again with a 1:1 mix of acetone:durpecan resin* and left overnight for the acetone to evaporate. Embryos were then transferred in a small droplet of resin to a microscope slide and aligned individually into rows. Each embryo was orientated to the correct view. The slide was then placed at 65°C to allow resin to harden. The orientation of embryos was checked occasionally and after a few hours, when the resin was no longer viscous, a fresh drop of resin was placed over the embryos and a coverslip placed on top. After 30 min, or when the resin had spread to the edges of the coverslip, the slide was placed on a 65°C heatblock with a lead weight on top, overnight.

*Durpecan resin was prepared by making two mixtures one of 56.5g solution A (resin) and 52.5g solution B (hardener) and the other of 1.75 ml solution C (accelerator) and 1ml solution D (softener). The individual mixtures were left for 10 min and then mixed together, aliquoted into 10ml syringes and stored at –20°C for many months. Resin could be thawed and refrozen a maximum of three times.

6.11 Injection of DNA and RNA

6.11.1 Injection procedure

Large cages of young, healthy flies were set up and the grape-juice agar plate with yeast paste changed regularly to promote laying. Plates were collected after 30min to one hour and these used for injection with the next hour. The rest of the procedure was carried out at 18°C. Embryos were dechorionated for 3min in 50% bleach and then rinsed with H2O. Embryos were lined up along the edge of a piece of agar in the same orientation. The line of about 50-70 embros was then transferred to a coverslip coated with a thin film of glue (made by dissolving glue from sellotape with heptane). The coverslip was placed onto a microscope slide using a drop of oil, and placed in silica gel at RT for dehydration, typically, 9-11 min. Once embryos were sufficiently dehydrated, they were covered with series 700 halocarbon oil and mounted on the microscope for injection. Pre-cellular blastoderm embryos were microinjected using an Eppendorf system. The injected embryos were covered in series 900 halocarbon oil, and allowed to develop at 25°C, until the required stage for a particular procedure.

6.11.2 DNA injection for transformation

P-element transformation is mediated by the introduction of two separate constructs into pre-cellular blastoderm embryos. The pUAST constructs contain the gene of interest and attenuated P-element vector, that is unable to transpose without the aid of a second construct, the helper plasmid. The helper plasmid (Δ2-3) supplies the transposase, though it itself cannot insert into the genome. To generate transgenic lines, foreign DNA is introduced into embryos by injection, and taken up by cells before cellularisation and integrated into the genome by random transposition events.

DNA of each pUAST construct and the helper plasmid were prepared using Bulk prep method. For each construct, the pUAST and helper were co-precipitated such

that the final concentration of each was 500ng/μl and 250ng/μl, respectively. DNA was resuspended in 1x injection buffer. The tube was centrifuged and solution transferred to a new tube to remove any dust particles.

w^{1118} embryos were used for injection. When hatched, first instar larvae were transferred, first to an agar plate to allow removal of excess oil, and then to food vials. About 20 larvae were placed in each vial , and these places at 25°C until eclosure. Individual flies were then crossed to w^{1118} males or virgins, and placed at 25°C until eclosure. Each cross was then screened for the presence of flies with coloured eyes, ranging from pale yellow to intense orange. These flies represented the transformants. Insertions were mapped genetically ().

6.11.3 RNA interference

RNA interference (RNAi) was essentially carried out as described by Kennerdell and Carthew Kennerdell and Carthew, 1998. Here I have described the method in brief.

Preparation and injection of dsRNA and asRNA

ds RNA was made by on of two methods. One method entailed using the cloned cDNA as template in individual transcription reactions with T7 and T3 polymerase, and then annealing the two RNA strands isolated from these reactions according to protocol of Kennerdell and Carthew (1998). The prefered method was using PCR. Primers were designed for the gene ORF of interest, but which included the minimal T7 RNA polymerase promoter site on the ends (Appendix 1). PCR with these primers was used to amplify a product for template in transcripion reactions with T7 RNA polymerase. In this case, both RNA strands are present and anneal *in situ*, with the efficiency of dsRNA production increased by heat denaturation followed by cooling. RiboMax transcription kits were used for transcription reactions (Promega).

Dilutions of dsRNA in injection buffer were injected into embryos on a trial and error basis, each gene appearing to have different conditions for interference. Embryo injections were as described in section 6.10.1. However two different approaches were used. In some cases, embryos expressing UAS-GFP were injected. These were left on the slide to develop at 25°C and later examined directly using a Leica immunofluoresence microscope under the 40x lens. For higher power

magnifications, the coverslip was inverted onto another slide, cleaned carefully, and a drop of imersion oil placed onto it for 60x and 100x oil immersion lenses.

In other cases, wildtype embryos were injected, and similarly allowed to develop under halocarbon oil. Once the embryos reached stage 17 however, the embryos were fixed and stained by immunohistochemistry, as described in 6.9.

Transfer of embryos to fix

Each coverslip of embryos was removed from the slide and as much oil as possible was scraped away. Heptane was then squirted over the embryos to dissolve the glue attaching the embryos, and these were collected in a scintillation vial. Once all embryos were transferred, any oil remaining in the heptane solution was removed, and the embryos rinsed several times in heptane. The embryos were then fixed as described in 6.10. After fixation the embryos were devitellinised by hand (6.10).

6.12 Genetic analysis

6.12.1 Maintenance of fly stocks

Flies stocks were kept in vials or bottles of 'Dundee Food'* at 18°C, or RT. Sickly stocks such as deficiencies were maintained at RT or 25°C. For experiments, crosses were carried out at 25°C. For misexpression analysis, a range of temperatures were used, i.e, 18°C, 25°C and 29°C.

*Dundee Food was prepared by Swann Media Kitchen Staff

6.12.2 Misexpression analysis

Many UAS*cato* lines were generated and crossed to different *Gal4* lines (Appendix 4) to assess their strength of expression. Each line was mapped genetically by crossing to different marked chromosome balancers. The table below sets out the different lines and their relative strengths, dependent on the range of phenotypes given with a range of *Gal4* line at different temperatures. Some lines were selected for experimental analysis under controlled conditions, and these represented a range from weak to strong. (R-red, M1M2 longer ORF, B-blue, M1, shorter ORF+HA tag, G-green, M2-shorter ORF). The UAS-AS*cato* did not give any phenotypes with the

Gal4 lines used, and all had very pale eye colours, thus these lines were all deemed to be weak.

Table 6.12.1 Transgenic lines produced for *cato*

Line	Eye Colour	Strength	Chromosome	Homozygous Viable
R5.1	red	strong	2	no
R4.1a	orange	medium	2	yes
R4.1b	dk red	strong	3	yes
R4.1c	red	medium	3	yes
R7.1b	pale	weak	X	yes
R6.1a	orange	medium	2	?
R9.1a	pale	v. weak	3	yes
B8.1a	pale	strong	2	yes
B8.1b	dark		X	?
G2.1a	red med.	weak*	3	yes
G3.1b	red med.	v. strong*	3	yes
G3.1a	pale	medium	3	yes
G6.1a	pale	weak	2	no
G1.1b	red	v. strong	3	yes
G3.2a	med	weak	3	no
G9.1	pale		X	yes

Appendices

Appendix 1

Primers for PCR and sequencing

Name	Sequence	Description
AJ1	GCYGCYAAYGCHCGYGARMG	degenerate primer to the conserved 5' end of bHLH region
AJ2	TGRGCCATYTGBARDGTYTC	degenerate primer to the conserved 3' end of bHLH region
SGB1	AATTAATCCATTCATCCGCTTC	sequence and PCR primer to 3' end of *cato* bHLH region
SGB2	CGTCCATTGACCAGAAATTG	sequence and PCR primer to 5' end of *cato* bHLH region
SGA1	GAACCACATCTCTCAACTTC	sequence and PCR primer to 3' end of *amos* bHLH region
SGA2	CTCCAATACGAAACCCTAC	sequence and PCR primer to 5' end of *amos* bHLH region
SGB3	GAGTGTCTCGAACTTGGACA	RT sequence and PCR primer at just downstream of SGB1
SGB4	CAGACCAACGGCGGAGTAA	sequence and PCR primer upstream of *cato* bHLH region
SGA3	GTCTCGTATTTGGAGAGTCG	sequence and PCR primer at just downstream of SGA1
SGA4	CAGAAGTCCAAGGTAGTGGA	sequence and PCR primer upstream of *amos* bHLH region
SGB8	TCAGAATTCTTGGTCAGAACGGC GGT	5' of M1 of *cato* ORF, *Eco*RI site for cloning
SG9	GCGGAATTCCTTTAGTCTAGTAT CATC	3' of STOP of *cato* ORF, *Eco*RI site for cloning
SG10	GGCGAATTCTCGGATCGTTTCAG TTG	5' of M2 of *cato* ORF, *Eco*RI site for cloning
SG11	TCAGAATTCCCTGACAAGGATC CTC	3' of ORF, to delete STOP codon, for HA tag cloning, *Eco*RI site
SG13	TAATACGACTCACTATAGGGACA GGCGAAGTGTAGGCG	T7 RNA pol site, catoF, for preparation of dsRNA
SG14	TAATACGACTCACTATAGGGCAA GCGATGCGTTCCATG	T7 RNA pol site, catoR, for preparation of dsRNA
SG15	TAATACGACTCACTATAGGGAGA CCGCTGCAACTCACC	T7 RNA pol site, atoF, for preparation of dsRNA
SG16	TAATACGACTCACTATAGGGCAA CATTTGGCAAGCTC	T7 RNA pol site, atoR, for preparation of dsRNA

Appendix 2

Recombinant Lambda phage isolated for *amos* and *cato*.

Over 300,000 plaques were transferred to filters in duplicate and individually hybridised to Digoxigenin-labelled RNA probes prepared from the *amos* and *cato* PCR clones. Preliminary genomic southern blot analysis established that the probes were specific to each gene concerned. Original names of individual clones are in brackets. In the case of *cato*, restriction digestion suggested that BC and BF were the same isolate. The second column indicates which phage fragments were subcloned for further analysis.

amos	SUBCLONE	*cato*	SUBCLONE
AA (A111)	3.7 kb SalI	BA (D221)	
AB (D133)		BB (C222)	
AC (D232)		BD (C111)	
AD (D314)		BE (D132)	2.7 kb SacI
AE (E111)		BC (E221)	2.2 kb SacI
		BF (F132)	

Appendix 3

Sequence data available for *cato*

Source	Clone name	Accession number	Authors
Genomic. wildtpye		AF134869	Goulding and Jarman
P1	DS06464	AC004318	BDGP
BAC	R21B04 R25I17 R17I17	AC007572 AC009356 AC008230	BDGP

Appendix 4

Sequence data available for *amos*

Source	Clone name	Acession number	Authors
Genomic. wildtpye		AF134869	Goulding and Jarman
Genomic, wildtype		-	Huang, et al., in press
BAC	R25C02 R13N02	AC007137 Ac006590	BDGP

Appendix 5: Phage Vectors, Plasmid and Bacterial strains

Bacteriophage vectors

Phage vector	Description	Reference/Source
λFIXII-	Genomic library cloning vector	Stratagene
λZAP Express™-	cDNA library cloning vector - allows rapid excision of pBS based phagemid for characterisation in plasmid system	Stratagene
ExAssit™ helper phage	For excision of pBK-CMV phagemid from ZAP Express vector	Stratagene

E. coli Bacterial strains used

Name	Relevant use	Source
XL1-blue	Cloning, library screening	Stratagene
XL1-Blue MRF'	library screening	Stratagene
JM109	Transformation of ligations	Promega
TOP10F'	Transformation of ligations	Invitrogen
XLOLR	For phagemid rescue	Stratagene

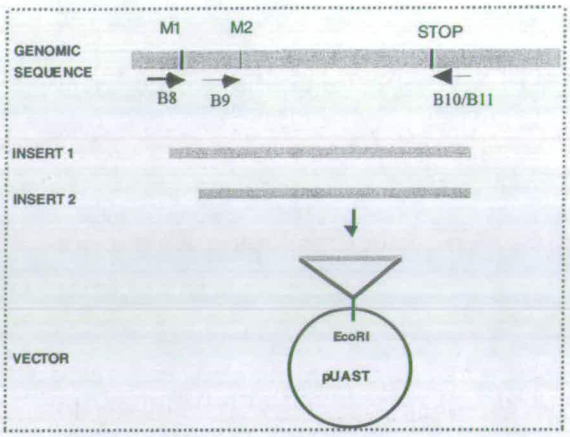
Plasmid vectors used

Name	Source
PBluescript® II	Stratagene
pBK-CMV	Stratagene
pUAST	Lab stock
PGEM®-T	Promega
pCR®-TOPO	Invitrogen

Appendix 6: Plasmids isolated and prepared in the lab

Name	Description
pBS-SalI-AA	3.7kb SalI fragment of λAA subcloned into pBlueScript – *amos* subclone
pBS-SacI-BE	2.7kb SacI fragment of λBE subcloned into pBlueScript– *cato* subclone
pBS-SacI-BC	2.2kb SacI fragment of λBC subcloned into pBlueScript– *cato* subclone
pBK-CMV-B1	*cato* cDNA (681bp) cloned into pBK-CMV
pBK-CMV-B2	*cato* cDNA (749bp) cloned into pBK-CMV
pUAST-M1M2cato	*cato* ORF cloned into pUAST, long version
pUAST-M2cato	*cato* ORF cloned into pUAST,short version
pUAST-HA-M1M2cato	*cato* ORF (long version) cloned into pUAST with HA tag
pUAST-HA-M2cato	*cato* ORF (short version) cloned into pUAST with HA tag
pUAST-M2antisense-cato	

Appendix 7: Strategy to clone the *cato* ORF into pUAST



The coding region of *cato* was cloned by PCR from genomic DNA and the genomic phage subclone pBSBE9. Two inserts were generated, the shorter one containing only the second start methionine. Primer locations are indicated by arrows. M1 and M2 refer to the first and second possible start codons respectively. PCR products were cut with EcoRI, cleaned and ligated into the pUAST vector.

Appendix 7: *D. melanogaster* stocks used and their origins

Stock	Description	Origin/Ref
OregonR (OrR)	wildtype strain	Lab stock
w1118	mutation of the *white* gene	Lab stock
Df(2R)Jp7/CyO	deletion of 52F5-9 to 53A	Bloomington
Df(2R)Jp8/CyO	deletion of 52F5-9 to 53A	Bloomington
yw; Df(2R)Jp7/CyO		this study
yw; Df(2R)Jp8/CyO		this study
ato1	EMS null allele of *atonal*	Jarman, 1993
Df(2L)M36F-S5/CyO	deletion of 36F	Bloomington
Df(2L)M36F-S6/CyO	deletion of 36F	Bloomington, Umea
A101	enhancer trap	C. O'Kane
ase1	asense	Brand, 1993
ase1; Df(2R)Jp7/CyO		this study
prosxx	lethal recessive of pros	Bloomington
Gal4109-69/CyO	Gal4 driver in PNC and SOP	Jan Lab
aseGal4	Gal4 driver	Jarman, 1998
hairyGal4	Gal4 driver	I. Davis
dppGal4	Gal4 driver	Bloomington
ptcGal4	Gal4 driver	G. Jiminez
enGal4UASGFP	Gal4 /UAS-GFP line	B. Sanson
enGal4UAStauGFP/CyO	Gal4 /UAS-tauGFP line	B. Sanson

References

References

Anderson, D. J. (1999). Lineages and transcription factors in the specification of vertebrate primary sensory neurons. *Curr Opin Neurobiol* **9**, 517-524.

Arendt, D. and Nübler-Jung, K. (1999). Comparison of early nerve cord development in insects and vertebrates. *Development* **126**, 2309-2325.

Artavanis-Tsakonas, S., Matsuno, K. and Fortini, M. E. (1995). Notch Signaling. *Science* **268**, 225-232.

Atchley, W. R., Fitch, W. M. and Bronner-Fraser, M. (1994). Molecular evolution of the MyoD family of transcription factors. *Proc Natl Acad Sci U S A* **91**, 11522-11526.

Bailey, A. M. and Posakony, J. W. (1995). Suppressor of hairless directly activates transcription of enhancer of split complex genes in response to Notch receptor activity. *Genes Dev* **9**, 2609-2622.

Baker, N. E. and Yu, S. Y. (1997). Proneural function of neurogenic genes in the developing *Drosophila* eye. *Current Biology* **7**, 122-132.

Baylies, M. K., Bate, M. and Gomez, M. R. (1998). Myogenesis: a view from *Drosophila*. *Cell* **93**, 921-927.

Bellen, H. J., Kooyer, S., D'Evelyn, D. and Pearlman, J. (1992). The *Drosophila* couch potato protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. *Genes and Development* **6**, 2125-2136.

Ben-Arie, N., Bellen, H. J., Armstrong, D. L., McCall, A. E., Gordadze, P. R., Guo, Q., Matzuk, M. M. and Zoghli, H. Y. (1997). Math1 is essential for genesis of cerebellar granule neurons. *Nature* **390**, 169-172.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. and Weintraub, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61**, 49-59.

Bermingham, N. A., Hassan, B. A., Price, S. D., Vollrath, M. A., Ben-Arie, N., Eatock, R. A., Bellen, H. J., Lysakowski, A. and Zoghbi, H. Y. (1999). Math1: an essential gene for the generation of inner ear hair cells. *Science* **284**, 1837-1841.

Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., Ackerman, L., Carretto, R., Uemura, T. and Grell, E. e. a. (1989). Searching for pattern and mutation in the *Drosophila* genome with a P-lacZ vector. *Genes Dev* **3**, 1273-1287.

Bier, E., Vaessin, H., Younger-Shepherd, S., Jan, L. Y. and Jan, Y. N. (1992). *deadpan*, an essential pan-neural gene in *Drosophila* encodes a helix-

loop-helix protein similar to the *hairy* gene product. *Genes and Development* **6**, 2137-2151.

Blackwell, T. K. and Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**, 1104-1110.

Blochlinger, K., Bodmer, R., Jan, L. Y. and Jan, Y. N. (1990). Patterns of expression of Cut, a protein required for external sensory organ development in wild-type and *cut* mutant *Drosophila* embryos. *Genes and Development* **4**, 1322-1331.

Blochlinger, K., Jan, L. Y. and Jan, Y. N. (1991). Transformation of sensory organ identity by ectopic expression of Cut in *Drosophila*. *Genes and Development* **5**, 1124-1135.

Bodmer, R., Carretto, R. and Jan, Y. N. (1989). Neurogenesis of the peripheral nervous system in *Drosophila* embryos: DNA replication patterns and cell lineages. *Neuron* **3**, 21-32.

Bodmer, R. and Jan, Y. N. (1987). Morphological differentiation of the embryonic peripheral neurons in *Drosophila*. *Roux's Archives of Developmental Biology* **196**, 69-77.

Botella, L. M., Donoro, C., Sanchez, L., Segarra, C. and Granadino, B. (1996). Cloning and characterization of the *scute* (sc) gene of *Drosophila* subobscura. *Genetics* **144**, 1043-1051.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415.

Brand, M., Jarman, A. P., Jan, L. Y. and Jan, Y. N. (1993). *asense* is a *Drosophila* neural precursor gene and is capable of initiating sense organ formation. *Development* **119**, 1-17.

Bray, S. (1998). A Notch affair. *Cell* **93**, 499-503.

Brewster, R. and Bodmer, R. (1995). Origin and specification of type II sensory neurons in *Drosophila*. *Development* **121**, 2923-2936.

Brown, N. L., Kanekar, S., Vetter, M. L., Tucker, P. K., Gemza, D. L. and Glaser, T. (1998). Math5 encodes a murine basic helix-loop-helix transcription factor expressed during early stages of retinal neurogenesis. *Development* **125**, 4821-4833.

Buckingham, M. (1992). Making muscle in mammals. *Trends Genet* **8**, 144-148.

Burmeister, M. and Lehrach, H. (1996). High-yield DNA preparation from liquid phage lambda cultures. *Trends Genet* **12**, 389.

Cabrera, C. V. and Alonso, M. C. (1991). Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products in *Drosophila*. *EMBO Journal* **10**, 965-973.

Campos-Ortega, J. (1993). Early neurogenesis in *Drosophila melanogaster*. In *The Development of Drosophila melanogaster*, vol. 2 (ed. M. Bate and A. Martinez-Arias), pp. 1091-1130. New York: Cold Spring Harbor Press.

Campos-Ortega, J. A. and Hartenstein, V. (1998). The Embryonic Development of *Drosophila melanogaster*. : Springer.

Carmeliet, P. (1999). Developmental biology. Controlling the cellular brakes. *Nature* **401**, 657-658.

Carmena, A., Bate, M. and Jiménez, F. (1995). *lethal of scute*, a proneural gene, participates in the specification of muscle progenitors during *Drosophila* embryogenesis. *Genes and Development* **9**, 2373-2383.

Cau, E., Gradwohl, G., Fode, C. and Guillemot, F. (1997). Mash1 activates a cascade of bHLH regulators in olfactory neuron progenitors. *Development* **124**, 1611-1621.

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

Caudy, M., Vassin, H., Brand, M., Tuma, R., Jan, L. Y. and Jan, Y. N. (1988b). daughterless, a *Drosophila* gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the *achaete-scute* complex. *Cell* **55**, 1061-1067.

Chan, Y. M. and Jan, Y. N. (1999). Conservation of neurogenic genes and mechanisms. *Curr Opin Neurobiol* **9**, 582-588.

Chien, C.-T., Hsiao, C.-D., Jan, L. Y. and Jan, Y. N. (1996). Neuronal type information encoded in the basic-helix-loop-helix domain of proneural genes. *Proceedings of the National Academy of Sciences, USA* **93**, 13239-13244.

Cohen, S. M. (1993). Imaginal disc development. In *The Development of Drosophila melanogaster*, vol. 2 (ed. M. Bate and A. Martinez-Arias), pp. 747-842. New York: Cold Spring Harbor Press.

Cubadda, Y., Heitzler, P., Ray, R. P., Bourouis, M., Ramain, P., Gelbart, W., Simpson, P. and Haenlin, M. (1997). u-shaped encodes a zinc finger protein that regulates the proneural genes *achaete* and *scute* during the formation of bristles in *Drosophila*. *Genes Dev* **11**, 3083-3095.

Cubas, P., de Celis, J.-F., Campuzano, S. and Modolell, J. (1991). Proneural clusters of *achaete-scute* expression and the generation of sensory organs in the *Drosophila* wing disc. *Genes and Development* **5**, 996-1008.

Daga, A., Karlovich, C. A., Dumstrei, K. and Banerjee, U. (1996). Patterning of cells in the *Drosophila* eye by Lozenge, which shares homologous domains with AML1. *Genes and Development* **10**, 1194-1205.

Dambly-Chaudière, C. and Ghysen, A. (1987). Independent subpatterns of sense organs require independent genes of the *achaete-scute* complex in *Drosophila* larvae. *Genes and Development* **1**, 297-306.

Dambly-Chaudière, C., Ghysen, A., Jan, L. Y. and Jan, Y. N. (1988). The determination of sense organs in *Drosophila*: interaction of *scute* with *daughterless*. *Roux's Archives of Developmental Biology* **197**, 419-423.

Dambly-Chaudière, C., Jamet, E., Burri, M., Bopp, D., Basler, K., Hafen, E., Dumont, N., Spielmann, P., Ghysen, A. and Noll, M. (1992). The Paired Box Gene Pox-Neuro - a Determinant of Poly-Innervated Sense-Organs in *Drosophila*. *Cell* **69**, 159-172.

Davis, R. L., Cheng, P. F., Lassar, A. B. and Weintraub, H. (1990). The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**, 733-746.

Davis, R. L. and Weintraub, H. (1992). Acquisition of myogenic specificity by replacement of three amino acid residues from MyoD into E12. *Science* **256**, 1027-1030.

Dickson, B. and Hafen, E. (1993). Genetic dissection of eye development in *Drosophila*. In *The Development of Drosophila melanogaster*, vol. 2 (ed. M. Bate and A. Martinez-Arias), pp. 1327-1361. New York: Cold Spring Harbor Press.

Doe, C. Q., Chu-LaGraff, Q., Wright, D. M. and Scott, M. P. (1991). The *prospero* gene specifies cell fates in the *Drosophila* central nervous system. **65**, 451-464.

Dominguez, M. and Campuzano, S. (1993). *asense*, a member of the *Drosophila achaete-scute* complex, is a proneural and neural differentiation gene. *EMBO Journal* **12**, 2049-2060.

Eberl, D. F. (1999). Feeling the vibes: chordotonal mechanisms in insect hearing. *Current Opinion in Neurobiology* **9**, 389-393.

Ellenberger, T., Fass, D., Arnaud, M. and Harrison, S. C. (1994). Crystal structure of transcription factor E47: E-box recognition by a basic region helix-loop-helix dimer. *Genes and Development* **8**, 970-980.

Emery, J. F. and Bier, E. (1995). Specificity of CNS and PNS regulatory subelements comprising pan-neural enhancers of the *deadpan* and *scratch* genes is achieved by repression. *Development* **121**, 3549-3560.

Fisher, A. and Caudy, M. (1998). The function of hairy-related bHLH repressor proteins in cell fate decisions. *Bioessays* **20**, 298-306.

Fisher, A. L., Ohsako, S. and Caudy, M. (1996). The Wrpw Motif Of the Hairy-Related Basic Helix-Loop-Helix Repressor Proteins Acts As a 4-Amino-Acid

Transcription Repression and Protein- Protein Interaction Domain. *Molecular and Cellular Biology* **16**, 2670-2677.

Flores, G. V., Daga, A., Kalhor, H. R. and Banerjee, U. (1998). Lozenge is expressed in pluripotent precursor cells and patterns multiple cell types in the *Drosophila* eye through control of cell-specific transcription factors. *Development* **125**, 3681-3687.

Freeman, M. (1996). Reiterative use of the EGF receptor triggers differentiation of all cell types in the *Drosophila* eye. *Cell* **87**, 651-660.

Gao, F. B., Brenman, J. E., Jan, L. Y. and Jan, Y. N. (1999). Genes regulating dendritic outgrowth, branching, and routing in *Drosophila*. *Genes Dev* **13**, 2549-2561.

Garrell, J. and Modolell, J. (1990). The *Drosophila* Extramacrochaetae Locus, an Antagonist Of Proneural Genes That, Like These Genes, Encodes a Helix-Loop-Helix Protein. *Cell* **61**, 39-48.

Gautier, P., Ledent, V., Massaer, M., Dambly-Chaudiere, C. and Ghysen, A. (1997). tap, a *Drosophila* bHLH gene expressed in chemosensory organs. *Gene* **191**, 15-21.

Georgias, C., Wasser, M. and Hinz, U. (1997). A basic-helix-loop-helix protein expressed in precursors of *Drosophila* longitudinal visceral muscles. *Mechanisms Of Development* **69**, 115-124.

Gho, M. and Schweisguth, F. (1998). Frizzled signalling controls orientation of asymmetric sense organ precursor cell divisions in *Drosophila*. *Nature* **393**, 178-181.

Ghysen, A. and Dambly-Chaudiere, C. (1988). From DNA to form: the *achaete-scute* complex. *Genes Dev* **2**, 495-501.

Ghysen, A. and Dambly-Chaudière, C. (1989). Genesis of the *Drosophila* peripheral nervous system. *Trends in Genetics* **5**, 251-255.

Ghysen, A. and Dambly-Chaudière, C. (1993). The specification of sensory neuron identity in *Drosophila*. *BioEssays* **15**, 202-208.

Ghysen, A., Dambly-Chaudière, C., Aceves, E., Jan, L. Y. and Jan, Y. N. (1986). Sensory neurons and peripheral pathways in *Drosophila* embryos. *Roux's Archives of Developmental Biology* **195**, 281-289.

Ghysen, A., Dambly-Chaudière, C., Jan, L. Y. and Jan, Y. N. (1993). Cell interactions and gene interactions in peripheral neurogenesis. *Genes and Development* **7**, 723-733.

Ghysen, A. and O'Kane, C. (1989). Neural enhancer-like elements as specific cell markers in *Drosophila*. *Development* **105**, 35-52.

Gomez-Skarmeta, J. L., Del Corral, R. D., De la CalleMustienes, E., Ferres Marco, D. and Modolell, J. (1996). Araucan and Caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. *Cell* **85**, 95-105.

Gomez-Skarmeta, J. L., Rodriguez, I., Martinez, C., Culi, J., FerresMarco, D., Beamonte, D. and Modolell, J. (1995). Cis-regulation of *achaete* and *scute*: Shared enhancer-like elements drive their coexpression in proneural clusters of the imaginal discs. *Genes and Development* **9**, 1869-1882.

Gonzalez-Crespo, S. and Levine, M. (1993). Interactions between dorsal and helix-loop-helix proteins initiate the differentiation of the embryonic mesoderm and neuroectoderm in *Drosophila*. *Genes Dev* **7**, 1703-1713.

Goodman, C. and Doe, C. Q. (1993). Embryonic Neurogenesis in *Drosophila melanogaster*. In *The Development of Drosophila melanogaster*, vol. 2 (ed. M. Bate and A. Martinez-Arias), pp. 1131-1206. New York: Cold Spring Harbor Press.

Goriely, A., Dumont, N., Dambly-Chaudière, C. and Ghysen, A. (1991). The determination of sense organs in *Drosophila*: effect of the neurogenic mutations in the embryo. *Development* **113**, 1395-1404.

Goulding, S. E.*, zur Lage, P.* and Jarman, A. P. (2000). *amos*, a proneural gene for *Drosophila* olfactory sense organs that is regulated by *lozenge*. *Neuron* **25**,69-78.

Goulding, S. E., White, N.M. and Jarman, A. P. (2000). *cato* encodes a basic-helix-loop-helix implicated in the correct differentiation of *Drosophila* sense organs. *Dev. Biol.* **in press.**

Greenspan, R. J. (1997). Fly Pushing. New York: Cold Spring Harbor Laboratory Press.

Grillenzoni, N., van Helden, J., Dambly-Chaudiere, C. and Ghysen, A. (1998). The iroquois complex controls the somatotopy of *Drosophila* notum mechanosensory projections. *Development* **125**, 3563-3569.

Guillemot, F. (1995). Analysis of the role of basic-helix-loop-helix transcription factors in the development of neural lineages in the mouse. *Biology of the Cell* **84**, 3-6.

Guillemot, F., Lo, L. C., Johnson, J. E., Auerbach, A., Anderson, D. J. and Joyner, A. L. (1993). Mammalian *achaete-scute* homolog-1 is required for the early development of olfactory and autonomic neurons. *Cell* **75**, 463-476.

Gupta, B. P., Flores, G. V., Banerjee, U. and Rodrigues, V. (1998). Pattterning an epidermal field: *Drosophila* Lozenge, a member of the AML/Runt family of transcription factors, specifies olfactory sense organ type in a dose-dependent manner. *Developmental Biology* **203**, 400-411.

Gupta, B. P. and Rodrigues, V. (1997). *atonal* is a proneural gene for a subset of olfactory sense organs in *Drosophila*. *Genes to Cells* **2**, 225-233.

Haenlin, M., Cubadda, Y., Blondeau, F., Heitzler, P., Lutz, Y., Simpson, P. and Ramain, P. (1997). Transcriptional activity of *pannier* is regulated negatively by heterodimerization of the GATA DNA-binding domain with a cofactor encoded by the *u-shaped* gene of *Drosophila*. *Genes Dev* **11**, 3096-3108.

Hartenstein, V. (1988). Development of *Drosophila* larval sensory organs: spatiotemporal pattern of sensory neurones, peripheral pathways and sensilla differentiation. *Development* **102**, 869-886.

Hartenstein, V. and Campos-Ortega, J. A. (1986). The peripheral nervous system of mutants of early neurogenesis in *Drosophila melanogaster*. *Roux's Archives of Developmental Biology* **195**, 210-221.

Hartenstein, V. and Posakony, J. W. (1989). Development of adult sensilla on the wing and notum of *Drosophila melanogaster*. *Development* **107**, 389-405.

Heimbeck, G., Bugnon, V., Gendre, N., Haberlin, C. and Stocker, R. F. (1999). Smell and taste perception in *Drosophila melanogaster* larva: toxin expression studies in chemosensory neurons. *J Neurosci* **19**, 6599-6609.

Henrique, D., Tyler, D., Kintner, C., Heath, J. K., Lewis, J. H., IshHorowicz, D. and Storey, K. G. (1997). *cash4*, a novel *achaete-scute* homolog induced by Hensen's node during generation of the posterior nervous system. *GENES & DEVELOPMENT* **11**, 603-615.

Higashijima, S., Michiue, T., Emori, Y. and Saigo, K. (1992). Subtype determination of *Drosophila* embryonic external sensory organs by redundant homeo box genes *BarH1* and *BarH2*. *Genes Dev* **6**, 1005-1018.

Hinz, U., Giebel, B. and CamposOrtega, J. A. (1994). The basic-helix-loop-helix domain of *Drosophila* lethal of *scute* protein is sufficient for proneural function and activates neurogenic genes. *Cell* **76**, 77-87.

Hirsch, M. R., Tiveron, M. C., Guillemot, F., Brunet, J. F. and Goridis, C. (1998). Control of noradrenergic differentiation and Phox2a expression by MASH1 in the central and peripheral nervous system. *Development* **125**, 599-608.

Huang, F., Dambly-Chaudière, C. and Ghysen, A. (1991). The emergence of sense organs in the wing disc of *Drosophila*. *Development* **111**, 1087-1095.

Huang, M.-L., Hsu, C.-H. and Chien, C.-H. (1999). The proneural gene *amos* promotes multiple denndritic neuron formation in *Drosophila* peripheral nervous system. *Neuron* **25**, 57-67.

Ip, Y. T., Levine, M. and Bier, E. (1994). Neurogenic expression of snail is controlled by separable CNS and PNS promoter elements. *Development* **120**, 199-207.

Ishibashi, M., Ang, S. L., Shiota, K., Nakanishi, S., Kageyama, R. and Guillemot, F. (1995). Targeted disruption of mammalian hairy and Enhancer of split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. *Genes Dev* **9**, 3136-3148.

Jan, Y. N. and Jan, L. Y. (1993). The peripheral nervous system. In *The Development of Drosophila melanogaster*, vol. 2 (ed. M. Bate and A. Martinez-Arias), pp. 1207-1244. New York: Cold Spring Harbor Press.

Jan, Y. N. and Jan, L. Y. (1994). Genetic control of cell fate specification in *Drosophila* peripheral nervous system. *Annu Rev Genet* **28**, 373-393.

Jan, Y. N. and Jan, L. Y. (1998). Asymmetric cell division. *Nature* **392**, 775-778.

Jarman, A. P. and Ahmed, I. (1998). The specificity of proneural genes in determining *Drosophila* sense organ identity. *Mechanisms of Development* **76**, 117-125.

Jarman, A. P., Brand, M., Jan, L. Y. and Jan, Y. N. (1993a). The regulation and function of the helix-loop-helix gene, *asense*, in *Drosophila* neural precursors. *Development* **119**, 19-29.

Jarman, A. P., Grau, Y., Jan, L. Y. and Jan, Y. N. (1993b). *atonal* is a proneural gene that directs chordotonal organ formation in the *Drosophila* peripheral nervous system. *Cell* **73**, 1307-1321.

Jarman, A. P., Grell, E. H., Ackerman, L., Jan, L. Y. and Jan, Y. N. (1994). *atonal* is the proneural gene for *Drosophila* photoreceptors. *Nature* **369**, 398-400.

Jarman, A. P. and Jan, Y. N. (1995). Multiple roles for proneural genes in *Drosophila* neurogenesis. In *Neural Cell Specification: Molecular Mechanisms and Neurotherapeutic Implications*, vol. 3 (ed. B. H. J. Juurlink, P. H. Krone, W. M. Kulyk, V. M. K. Verge and J. R. Doucette), pp. 97-104. New York: Plenum Press.

Jarman, A. P., Sun, Y., Jan, L. Y. and Jan, Y. N. (1995). Role of the proneural gene, *atonal*, in formation of *Drosophila* chordotonal organs and photoreceptors. *Development* **121**, 2019-2030.

Jennings, B., De Celis, J., Delidakis, C., Preiss, A. and Bray, S. (1995). Role of Notch and *achaete-scute* complex in the expression of Enhancer of split bHLH proteins. *Development* **121**, 3745-3752.

Kageyama, R. and Nakanishi, S. (1997). Helix-loop-helix factors in growth and differentiation of the vertebrate nervous system. *Curr Opin Genet Dev* **7**, 659-665.

Kanekar, S., Perron, M., Dorsky, R., Harris, W. A., Jan, L. Y., Jan, Y. N. and Vetter, M. L. (1997a). *Xath5* participates in a network of bHLH genes in the developing Xenopus retina. *Neuron* **19**, 981-994.

Kanekar, S., Perron, M., Harris, W. A., Jan, L. Y., Jan, Y. N. and Vetter, M. L. (1997b). Xath5, a Xenopus homolog of the *Drosophila* proneural gene *atonal*, functions with Xash3 and NeuroD to regulate retinal neurogenesis. *Developmental Biology* **186**, A239-A239.

Keller, C. A., Erickson, M. S. and Abmayr, S. M. (1997). Misexpression of nautilus induces myogenesis in cardioblasts and alters the pattern of somatic muscle fibers. *Dev Biol* **181**, 197-212.

Keller, C. A., Grill, M. A. and Abmayr, S. M. (1998). A role for nautilus in the differentiation of muscle precursors. *Dev Biol* **202**, 157-171.

Kennerdell, J. R. and Carthew, R. W. (1998). Use of dsRNA-mediated genetic interference to demonstrage that frizzled and frizzled 2 act in the wingless pathway. *Cell* **95**, 1017-1026.

Kidd, S., Lieber, T. and Young, M. W. (1998). Ligand-induced cleavage and regulation of nuclear entry of Notch in *Drosophila melanogaster* embryos. *Genes Dev* **12**, 3728-3740.

Kunne, A. G., Meierhans, D. and Allemann, R. K. (1996). Basic helix-loop-helix protein MyoD displays modest DNA binding specificity. *FEBS Lett* **391**, 79-83.

Lee, J. E. (1997). Basic helix-loop-helix genes in neural development. *Current Opinion in Neurobiology* **7**, 13-20.

Lee, J. E., Hollenberg, S. M., Snider, L., Turner, D. L., Lipnick, N. and Weintraub, H. (1995). Conversion of *Xenopus* ectoderm into neurons by *NeuroD*, a basic helix-loop-helix protein. *Science* **268**, 836-844.

Lehner, C. F. and O'Farrell, P. H. (1989). Expression and function of *Drosophila* cyclin A during embryonic cell cycle progression. *Cell* **56**, 957-968.

Littlewood, T. and Evan, G. I. (1998). Helix-loop-helix transcription factors. Oxford: Oxford University Press.

Lo, L., Tiveron, M. C. and Anderson, D. J. (1998). MASH1 activates expression of the paired homeodomain transcription factor Phox2a, and couples pan-neuronal and subtype-specific components of autonomic neuronal identity. *Development* **125**, 609-620.

Lu, J., Richardson, J. A. and Olson, E. N. (1998). Capsulin: a novel bHLH transcription factor expressed in epicardial progenitors and mesenchyme of visceral organs. *Mech Dev* **73**, 23-32.

Lu, J., Webb, R., Richardson, J. A. and Olson, E. N. (1999). MyoR: a muscle-restricted basic helix-loop-helix transcription factor that antagonizes the actions of MyoD. *Proc Natl Acad Sci U S A* **96**, 552-557.

Lyden, D., Young, A. Z., Zagzag, D., Yan, W., Gerald, W., O'Reilly, R., Bader, B., Hynes, R. O., Zhuang, Y., Manova, K. et

al. (1999). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. *Nature* **401**, 670-677.

Ma, P. C. M., Rould, M. A., Weintraub, H. and Pabo, C. O. (1994). Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. *Cell* **77**, 451-459.

Ma, Q. F., Kintner, C. and Anderson, D. J. (1996). Identification of neurogenin, a vertebrate neuronal determination gene. *Cell* **87**, 43-52.

Manning, L. and Doe, C. Q. (1999). Prospero distinguishes sibling cell fate without asymmetric localization in the *Drosophila* adult external sense organ lineage. *Development* **126**, 2063-2071.

Mao, Z. X. and Nadalginard, B. (1996). Functional and physical interactions between mammalian *achaete-scute* homolog 1 and myocyte enhancer factor 2A. *Journal of Biological Chemistry* **271**, 14371-14375.

McIver, S. B. (1985). Mechanoreception. In *Comprehensive Insect Physiology, Biochemistry and Pharmacology*, vol. 6 (ed. L. I. Gilbert and D. A. Kerkut), pp. 71-132. New York/London: Pergamon Press.

Merritt, D. J. (1997). Transformation of external sensilla to chordotonal sensilla in the cut mutant of *Drosophila* assessed by single-cell marking in the embryo and larva. *Microsc Res Tech* **39**, 492-505.

Michelson, A. M., Abmayr, S. M., Bate, M., Arias, A. M. and Maniatis, T. (1990). Expression of a MyoD family member prefigures muscle pattern in *Drosophila* embryos. *Genes Dev* **4**, 2086-2097.

Misquitta, L. and Paterson, B. M. (1999). Targeted disruption of gene function in *Drosophila* by RNA interference (RNA-i): a role for nautilus in embryonic somatic muscle formation. *Proceedings of the National Academy of Sciences, USA* **96**, 1451-1456.

Mlodzik, M., Baker, N. E. and Rubin, G. M. (1990). Isolation and expression of *scabrous*, a gene regulating neurogenesis in *Drosophila*. *Genes and Development* **4**, 1848-1861.

Modolell, J. and Campuzano, S. (1998). The *achaete-scute* complex as an integrating device. *International Journal Of Developmental Biology* **42**, 275-282.

Molkentin, J. D., Black, B. L., Martin, J. F. and Olson, E. N. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* **83**, 1125-1136.

Molkentin, J. D. and Olson, E. N. (1996). Defining the regulatory networks for muscle development. *Curr Opin Genet Dev* **6**, 445-453.

Moulins, M. (1976). Ultrastructure of chordotonal organs. In *Structure and Function of Proprioceptors in the Invertebrates*, (ed. P. J. Mill), pp. 387-426. London: Chapman and Hall.

Murray, M. J., Merritt, D. J., Brand, A. H. and Whitington, P. M. (1998). In vivo dynamics of axon pathfinding in the Drosophilia CNS: a time-lapse study of an identified motorneuron. *J Neurobiol* **37**, 607-621.

Murre, C., McCaw, P. S. and Baltimore, D. (1989a). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell 1989 Mar 10;56(5):777-83* .

Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D. and Lassar, A. B. e. a. (1989b). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**, 537-544.

Nakao, K. and Campos-Ortega, J. A. (1996). Persistent Expression Of Genes Of the Enhancer Of Split Complex Suppresses Neural Development In *Drosophila*. *Neuron* **16**, 275-286.

Ohsako, S., Hyer, J., Panganiban, G., Oliver, I. and Caudy, M. (1994). Hairy function as a DNA-binding helix-loop-helix repressor of *Drosophila* sensory organ formation. *Genes and Development* **8**, 2743-2755.

Okabe, M. and Okano, H. (1997). Two-step induction of chordotonal organ precursors in *Drosophila* embryogenesis. *Development* **124**, 1045-1053.

Parkhurst, S. M., Lipshitz, H. D. and Ish-Horowicz, D. (1993). *achaete-scute* feminizing activities and *Drosophila* sex determination. *Development* **117**, 737-749.

Parkhurst, S. M. and Meneely, P. M. (1994). Sex determination and dosage compensation: lessons from flies and worms. *Science* **264**, 924-932.

Paroush, Z., Jr., F. R. L., Kidd, T., Wainwright, S. M., Ingham, P. W., Brent, R. and Ish-Horowicz, D. (1994). Groucho is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* **79**, 805-815.

Porcher, C., Liao, E. C., Fujiwara, Y., Zon, L. I. and Orkin, S. H. (1999). Specification of hematopoietic and vascular development by the bHLH transcription factor SCL without direct DNA binding. *Development* **126**, 4603-4615.

Ramain, P., Heitzler, P., Haenlin, M. and Simpson, P. (1993). pannier, a negative regulator of *achaete* and *scute* in *Drosophila*, encodes a zinc finger protein with homology to the vertebrate transcription factor GATA-1. *Development* **119**, 1277-1291.

Ray, K. and Rodrigues, V. (1995). Cellular events during development of the olfactory sense organs in *Drosophila melanogaster*. *Developmental Biology* **167**, 426-438.

Ready, D. F. (1989). A multifaceted approach to neural development. *Trends In Neurosciences* **12**, 102-110.

Rebay, I., Fortini, M. E. and Artavanis-Tsakonas, S. (1993 Sep;). Analysis of phenotypic abnormalities and cell fate changes caused by dominant activated and dominant negative forms of the Notch receptor in *Drosophila* development. *C R Acad Sci III* **316**, 1097-1123.

Reddy, G. V., Gupta, B., Ray, K. and Rodrigues, V. (1997). Development of the *Drosophila* olfactory sense organs utilizes cell- cell interactions as well as lineage. *Development* **124**, 703-712.

Rhyu, M. S., Jan, L. Y. and Jan, Y. N. (1994). Asymmetric distribution of numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells. *Cell* **76**, 477-491.

Riesgo-Escovar, J. R., Peikos, W. B. and Carlson, J. R. (1997). The *Drosophila* antenna: ultrastructural and physiological studies in wild-type and *lozenge* mutants. *J Comp. Physiol. A* **180**, 151-160.

Roark, M., Sturtevant, M. A., Emery, J., Vaessin, H., Grell, E. and Bier, E. (1995). *scratch*, a pan-neural gene encoding a zinc finger protein related to Snail, promotes neuronal development. *Genes and Development* **9**, 2384-2398.

Rodríguez, I., Hernández, R., Modolell, J. and Ruiz-Gómez, M. (1990). Competence to develop sensory organs is temporally and spatially regulated in *Drosophila* epidermal primordia. *EMBO Journal* **9**, 3583-3592.

Romani, S., Campuzano, S., Macagno, E. R. and Modolell, J. (1989). Expression Of *Achaete* and *Scute* Genes In *Drosophila* Imaginal Disks and Their Function In Sensory Organ Development. *GENES & DEVELOPMENT* **3**, 997-1007.

Roztocil, T., Matter-Sadzinski, L., Alliod, C., Ballivet, M. and Matter, J.-M. (1997). NeuroM, a neural helix-loop-helix transcription factor, defines a new transition stage in neurogenesis. *Development* **124**, 3263-3272.

Ruizgomez, M. and Modolell, J. (1987). Deletion Analysis Of the *Achaete-Scute* Locus Of *Drosophila- Melanogaster*. *GENES & DEVELOPMENT* **1**, 1238-1246.

Rushlow, C. A., Hogan, A., Pinchin, S. M., Howe, K. M., Lardelli, M. and Ish-Horowicz, D. (1989). The *Drosophila* hairy protein acts in both segmentation and bristle patterning and shows homology to N-myc. *EMBO J* **8**, 3095-3103.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning - A Laboratory Manual. New York: Cold Spring Harbour LAboratory Press.

Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R. and Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to *Drosophila* hairy and Enhancer of split. *Genes Dev* **6**, 2620-2634.

Sato, M., Kojima, T., Michiue, T. and Saigo, K. (1999). Bar homeobox genes are latitudinal prepattern genes in the developing *Drosophila* notum whose

expression is regulated by the concerted functions of decapentaplegic and wingless. *Development* **126**, 1457-1466.

Saxton, W. M., Hicks, J., Goldstein, L. S. B. and Raff, E. C. (1991). Kinesin heavy chain is essential for viability and neuromuscular functions in *Drosophila*, but mutants show no defects in mitosis. *Cell* **64**, 1093-1102.

Shirokawa, J. M. and Courey, A. J. (1997). A direct contact between the dorsal rel homology domain and Twist may mediate transcriptional synergy. *Mol Cell Biol* **17**, 3345-3355.

Simpson, P. (1997). Notch signalling in development: on equivalence groups and asymmetric developmental potential. *CURRENT OPINION IN GENETICS & DEVELOPMENT* **7**, 537-542.

Skeath, J. B. and Carroll, S. B. (1991). Regulation of *achaete-scute* gene expression and sensory organ formation in the *Drosophila* wing. *Genes and Development* **5**, 984-995.

Skeath, J. B. and Carroll, S. B. (1992). Regulation Of Proneural Gene-Expression and Cell Fate During Neuroblast Segregation In the *Drosophila* Embryo. *Development* **114**, 939-946.

Skeath, J. B. and Doe, C. Q. (1996). The *achaete-scute* complex proneural genes contribute to neural precursor specification in the *Drosophila* CNS. *Curr Biol* **6**, 1146-1152.

Spicer, D. B., Rhee, J., Cheung, W. L. and Lassar, A. B. (1996). Inhibition of myogenic bHLH and MEF2 transcription factors by the bHLH protein Twist. *Science* **272**, 1476-1480.

Stocker, R. F., Gendre, N. and Batterham, P. (1993). Analysis of the antennal phenotype in the *Drosophila* mutant *lozenge*. *Journal of Neurogenetics* **9**, 29-53.

Struhl, G. and Adachi, A. (1998). Nuclear access and action of notch in vivo. *Cell* **93**, 649-660.

Sugawara, T. (1996). Chordotonal sensilla embedded in the epidermis of the soft integument of the cricket, Teleogryllus commodus. *Cell and Tissue Research* **284**, 125-142.

Sun, Y., Jan, L. Y. and Jan, Y. N. (1998). Transcriptional regulation of *atonal* during development of the *Drosophila* peripheral nervous system. *Development* **125**, 3731-3740.

Tautz, D. and Pfeifle, C. (1989). A nonradioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translation control of the segmentation gene *hunchback*. *Chromosoma* **98**, 81-85.

Tomita, K., Ishibashi, M., Nakahara, K., Ang, S. L., Nakanishi, S., Guillemot, F. and Kageyama, R. (1996). Mammalian hairy and Enhancer of

split homolog 1 regulates differentiation of retinal neurons and is essential for eye morphogenesis. *Neuron* **16**, 723-734.

Turner, D. L. and Weintraub, H. (1994). Expression of *achaete-scute* homolog 3 in Xenopus embryos converts ectodermal cells to a neural fate. *Genes and Development* **8**, 1434-1447.

Vaessin, H., Grell, E., Wolff, E., Bier, E., Jan, L. Y. and Jan, Y. N. (1991). *prospero* is expressed in neuronal precursors and encodes a nuclear protein that is involved in the control of axonal outgrowth in *Drosophila*. *Cell* **67**, 941-953.

Van Doren, M., Bailey, A. M., Esnayra, J., Ede, K. and Posakony, J. W. (1994). Negative regulation of proneural gene activity: Hairy is a direct transcriptional repressor of *achaete*. *Genes and Development* **8**, 2729-2742.

Van Doren, M., Ellis, H. M. and Posakony, J. W. (1991). The *Drosophila* Extramacrochaetae protein antagonizes sequence-specific DNA binding by Daughterless/*Achaete-Scute* protein complexes. *Development* **113**, 245-255.

van Staaden, M. and Romer, H. (1998). Evolutionary transition from stretch to hearing organs in ancient grasshoppers. *Nature* **394**, 773-776.

Vervoort, M., Dambly-Chaudière, C. and Ghysen, A. (1997a). Cell fate determination in *Drosophila*. *Current Opinion in Neurobiology* **7**, 21-28.

Vervoort, M., Merritt, D. J., Ghysen, A. and DamblyChaudiere, C. (1997b). Genetic basis of the formation and identity of type I and type II neurons in *Drosophila* embryos. *Development* **124**, 2819-2828.

White, N. M. and Jarman, A. P. (1999). *Atonal* controls photoreceptor R8-specific properties and modulates both RTK and Hedgehog signalling during *Drosophila* eye development. *Development* **in press**.

Wolff, T. and Ready, D. F. (1993). Pattern Formation in the *Drosophila* retina. In *The Development of Drosophila melanogaster*, vol. 2 (ed. M. Bate and A. Martinez-Arias), pp. . New York: Cold Spring Harbor Press.

Younger-Shepherd, S., Vaessin, H., Bier, E., Jan, L. Y. and Jan, Y. N. (1992). deadpan, an essential pan-neural gene encoding an HLH protein, acts as a denominator in *Drosophila* sex determination. *Cell* **70**, 911-922.

Zacharuk, R. Y. (1985). Antennae and Sensilla. In *Comprehensive Insect Physiology, Biochemistry and Pharmacology*, vol. 6 (ed. L. I. Gilbert and D. A. Kerkut), pp. 1-71. New York/London: Pergamon Press.

Zipursky, S. L., Venkatesh, T. R., Teplow, D. B. and Benzer, S. (1984). Neuronal development in the *Drosophila* retina: monoclonal antibodies as molecular probes. *Cell* **36**, 15-26.

zur Lage, P., Jan, Y. N. and Jarman, A. P. (1997). Requirement for EGF receptor signalling in neural recruitment during formation of *Drosophila* chordotonal sense organ clusters. *Current Biology* **7**, 166-175.

zur Lage, P. and Jarman, A. P. (1999). Antagonism of EGFR and Notch signalling in the reiterative recruitment of *Drosophila* adult chordotonal sense organ precursors. *Development* **126**, 3149-3157.

amos, a Proneural Gene for *Drosophila* Olfactory Sense Organs that Is Regulated by *lozenge*

Sarah E. Goulding,† Petra zur Lage,† and Andrew P. Jarman*
Institute of Cell and Molecular Biology
University of Edinburgh
King's Buildings
Edinburgh EH9 3JR
United Kingdom

Summary

In a variety of organisms, early neurogenesis requires the function of basic-helix-loop-helix (bHLH) transcription factors. For the *Drosophila* PNS, such transcription factors are encoded by the proneural genes (*atonal* and the *achaete–scute* complex, AS-C). We have identified a proneural gene, *amos*, that has strong similarity with *atonal* in its bHLH domain. We present evidence that *amos* is required for olfactory sensilla and is regulated by the prepattern gene *lozenge*. Between them, *amos*, *atonal*, and the AS-C can potentially account for the origin of the entire PNS.

Introduction

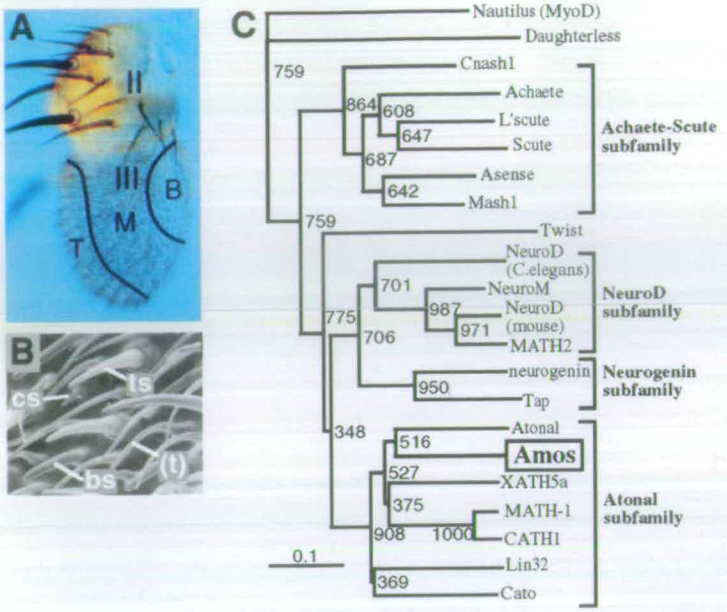
Insect behavior relies heavily on the sensation and interpretation of olfactory stimuli (Siddiqi, 1987). In the *Drosophila melanogaster* larva, olfaction is thought to be mediated by a group of sense organs (sensilla) that form the antenno-maxillary complex (Carlson, 1996). The adult has two main sites of olfaction. Some 450 olfactory sensilla are housed on the third antennal segment of the adult fly (Figure 1A), while an additional 80 sensilla cover the maxillary palp (Carlson, 1996). The function of the adult sensilla as olfactory receptors has been established electrophysiologically for a range of chemical stimuli (Siddiqi, 1987; Carlson, 1996). Moreover, they have recently been shown to express a family of odorant receptor proteins (Clyne et al., 1999). Furthermore, flies carrying mutations that affect the development of subsets of olfactory sensilla have defective olfactory perception (Riesgo-Escovar et al., 1997). Externally, each sensillum comprises a 4–20 μm cuticular protuberance that has microscopic pores and grooves through which it is presumed odours diffuse or are relayed to the internal sensory lymph (Riesgo-Escovar et al., 1997). Each is innervated by up to four bipolar sensory neurons so that the olfactory system comprises about one thousand sensory neurons. The sensilla are classed in three main morphological subtypes, termed sensilla basiconica, trichodea, and coeloconica (Figure 1B).

The development of olfactory sensilla is not as well characterized as it is for sense organs such as sensory bristles (external sense organs) or chordotonal organs (internal stretch receptors) (Jan and Jan, 1993). Nevertheless, the initial steps appear similar to these other sense organs. For each adult olfactory sensillum, a precursor cell is selected from the ectoderm of the early pupal antennal imaginal disc. These cells, generally termed sense organ precursors (SOPs), express well-known neural precursor markers such as the A101 enhancer trap line, and they arise over an extended period of time in the developing pupal antennal imaginal disc (Reddy et al., 1997). Despite their resemblance to sensory bristles, the cells of an olfactory sensillum do not appear to arise solely by division of the SOP (Ray and Rodrigues, 1995). Instead, the SOP (dubbed the founder cell) apparently recruits surrounding ectodermal cells to become the support cells that form the external sensillum structure (Reddy et al., 1997).

Despite their importance, relatively little is known about the developmental genetics of olfactory SOP selection. For other sense organs, SOP selection requires the function of proneural genes (Campuzano and Modolell, 1992; Jarman and Jan, 1995), which encode transcription factors with a basic helix-loop-helix (bHLH) domain for dimerization and DNA binding. These proteins bind to DNA as heterodimers with the ubiquitously expressed bHLH product of the *daughterless* (*da*) gene (Cabrera and Alonso, 1991). Furthermore, this interaction (and hence neurogenesis) is inhibited by negative regulators of the same protein family. Thus, *extramacrochaetae* (*emc*) encodes an HLH protein (lacking the basic DNA-binding domain) that functions by sequestering proneural proteins in inactive heterodimers (Van Doren et al., 1991; Cabrera et al., 1994). Characteristically, the transient expression of proneural genes in clusters of ectodermal cells (proneural clusters) marks these cells as having neural competence (Campuzano and Modolell, 1992). Only one or a few cells realize this potential, and these (the SOPs) inhibit the competence and proneural gene expression of the remaining cells via Notch signaling (lateral inhibition) (Ghysen et al., 1993). There are two known types of proneural gene, and these are required for separate subtypes of SOP. The *achaete-scute* complex genes (AS-C) are required for the SOPs of external sense organs (including mechanosensory and gustatory bristles) (Campuzano and Modolell, 1992). *atonal* (*ato*) governs precursor selection for chordotonal organs and R8 photoreceptors (Jarman et al., 1993, 1994, 1995).

Between them, the AS-C and *ato* account for SOP selection for much of the PNS, with some notable exceptions. In each embryonic abdominal and thoracic hemisegment, two sensory neurons of unknown function are unaffected in AS-C and *ato* double mutants (the dbd and dmd neurons, Jan and Jan, 1993; Jarman et al., 1993). Furthermore, much of the olfactory system develops independently of these genes. Perhaps surprisingly, the adult olfactory SOPs are not specified by the AS-C (Gupta and Rodrigues, 1997), confirming that these sensilla are distinct from external sense organs. Instead, some maxillary palp sensilla and the antennal sensilla coeloconica require *ato* (Gupta and Rodrigues, 1997).

*To whom correspondence should be addressed (e-mail: andrew.jarman@ed.ac.uk).
†These authors contributed equally to this work.



D

```
<    basic    ><J><   Helix1    >< Loop ><      Helix2       >
  RR    N RER R   ψ    F  L   ψ          Kψ TL  Aψ Yψ  L
KKRRLAANARERRRMNSLNDAFDKLRDVVPSLGHDRRLSKYETLQMAQAYIGDLVTLL  Amos
R--------------QN--Q---R--QYL-C--N--Q---H-------T--SA-GD--  Atonal
-Q-------------HG--H---Q--N-I--FNN-KK-----------I--NA-SE--  MATH1
-R--Q-------K---G--A--ER--E---APSI-QK---F-------S--LA-CD--  Cato

-V--MK------N--HG----L-S--K---CYSKTQK---I---RL-KN--WA-SEI-  NeuroD
-F--QE------N--HG----L-N--K---CYSKTQK---I---RL-KN--WA-SEI-  MATH2

RS--VK--D---N--HN--A-L-A--S-L-SFPD-TK-T-I---RF-YN--WA-AET-  Neurogenin
RF--MK--D---N--HN----LE---VTL-S-PEETK-T-I-I-RF-HN--FA-EQV-  Tap

VQ--...-----N-YKQV-NS-AR--QHI-GR-PHKKI--VD--RI-VE--RS-QD-  Scute
VA--...-----N-VKQV-NG-AL--EKI-GR-ASKK---V---R--VE--RS-EK-  Asense

QTQ-VM--V---Q-TQ---E--AA--KII-T-PS-.K---IQ--KL-AR--DF-YQV-  Twist
```

Figure 1. Olfactory Sensilla

(A) Wild-type antenna showing olfactory sensilla on third segment (III) and mechanosensory bristles (external sense organs) on the second (II). Three domains of olfactory sensilla are recognized that contain largely sensilla basiconica (B), sensilla trichodea (T), or a mixture of all three sensillum types (M).

(B) Scanning electron micrograph (SEM) from the "M" domain, showing sensilla basiconica (bs), trichodea (ts), and coeloconica (cs). Also apparent are the noninnervated epidermal hairs known as trichomes (t) or spinules.

(C and D) Amos has an Ato-like bHLH domain.

(C) Phylogenetic analysis of bHLH relationships for Amos and selected proteins, compiled using ClustalX. Bootstraps values for 1000 runs are indicated.

(D) Sequence comparisons of bHLH domains from Amos and selected proteins. Dashes represent identities, and dots represent gaps. The extended loop sequences of Scute and Asense have been removed for clarity. The regions recognized by the degenerate primers are underlined. Consensus residues important for bHLH structure are shown above the sequences.

None of these genes, however, are required for the sensilla basiconica and trichodea. Nevertheless, there is good evidence that these sensilla require the activity of an unidentified bHLH proneural gene(s). Gupta and Rodrigues (1997) showed that misexpression of *emc* in the pupal antennal disc causes a decrease in olfactory sensilla of all types, suggestive of antagonism of a bHLH protein in addition to *ato*. Interestingly, basiconic and trichoid sensillum formation requires the function of a Runt domain transcription factor encoded by *lozenge* (*lz*), with sensilla basiconica being completely absent in strong *lz* mutants (Stocker et al., 1993). A recent analysis suggests that *lz* acts upstream of SOP formation but is not itself a proneural gene (Gupta et al., 1998).

Here, we describe the isolation and characterization of *amos* (*absent md neurons and olfactory sensilla*), a bHLH gene related to *ato*. Its mRNA expression is highly localized, being mostly restricted to regions in which olfactory precursors arise in both embryo and imaginal discs, as well as the putative precursor of the dbd neuron in the embryo. We provide loss-of-function and gain-of-function evidence that *amos* is a proneural gene for olfactory sensilla, most likely the sensilla basiconica and trichodea. Moreover, *amos* is partly regulated by *lz*, and this regulation can account for the effect of *lz* mutations on olfactory SOP formation.

Results

Isolation of a Gene Containing an Ato-like bHLH Coding Region

bHLH-containing DNA fragments were isolated by PCR amplification of *Drosophila* genomic DNA using degenerate primers designed to the ends of the bHLH region that are conserved between Ato and its closest vertebrate homologs (Math1 and Math5) (Figure 1D). These conserved peptides distinguish the Ato subfamily from other bHLH proteins, including more divergent Ato-related ones such as the NeuroD subfamily (Figure 1D). When sequenced, the ~130 bp PCR products proved to be a mixture of three independent sequences. The bHLH domain of *ato* itself was amplified as expected, but, in addition, two novel bHLH sequences were identified that we name *cousin of ato* (*cato*) (S. E. G., N. M. White, and A. P. J., submitted) and *amos*. Northern blot analysis revealed a single *amos* embryonic transcript of about 800 nucleotides. The longest cDNA clone isolated was 782 bp, which contained a 198-amino-acid open reading frame (ORF) with the bHLH domain at the carboxyl terminus. Sequencing of the corresponding genomic DNA showed that *amos* contains no introns (GenBank accession number AF166113). Further analysis of sequence from a 170 kb BAC clone in which *amos* is

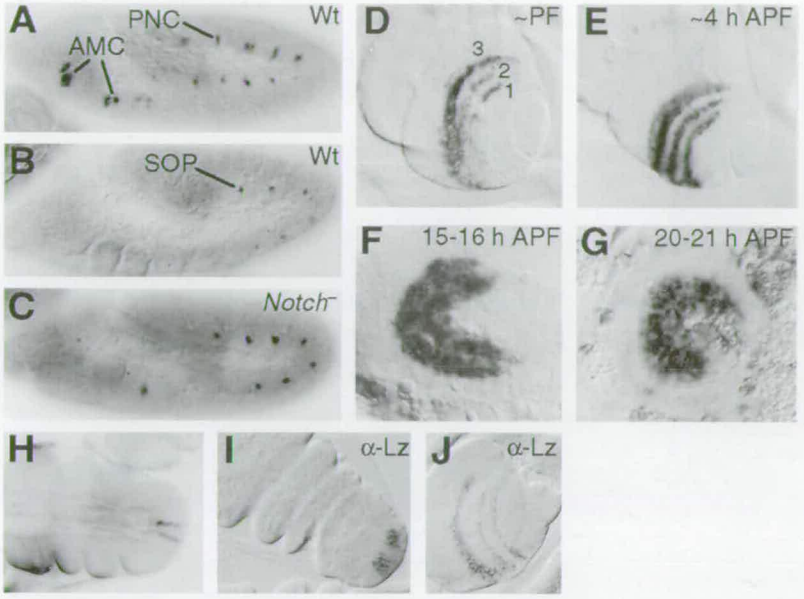


Figure 2. Expression of *amos* RNA

(A–C) Embryo.

(A) Wild-type, stage 10, showing expression in antennomaxillary complex domains (AMC) and a single cluster of ~5 cells in each thoracic and abdominal segment.

(B) Wild-type, stage 11, abdominal expression is now refined to a single SOP; thoracic and head expression have already ceased.

(C) Notch mutant, stage 11, showing continued expression in all cells of the abdominal clusters.

(D–H) Wild-type pupal antennal discs.

(D) Puparium formation (PF). Expression beginning in three semicircular bands on the developing third segment.

(E) ~4 hr APF.

(F) 15–16 hr APF. Expression is broader, and the inner two bands appear to merge.

(G) 20–21 hr APF. Expression is patchy, and bands are less apparent.

(H) Wild-type distal leg disc, 0–4 hr APF. Expression in two cells in the region fated to give the tarsal claw.

(I and J) Lz expression strongly resembles that of *amos* in the distal leg disc, 0–4 hr APF (I), and antennal disc, PF (J).

located (Berkeley *Drosophila* Genome Project, unpublished; accession number AC007137) showed that *amos* is the only bHLH-encoding gene in this region and therefore does not form a complex with other bHLH genes.

The predicted protein sequences of Amos and Ato are 74% identical over the entire bHLH region and share an identical basic domain except for an R to K conservative change (Figure 1D). This compares with ~70% identity between bHLH domains of the AS-C and ~40% between Amos and Scute. Among bHLH proteins, Amos forms part of the Ato subfamily, being most closely related to Ato followed by Ato's closest vertebrate homologs (Figure 1C). Interestingly, this suggests that an *ato–amos* gene duplication occurred after the invertebrate-vertebrate split, so that vertebrate *ato*-like genes may exhibit functional similarity with either *amos* or *ato*. The gene duplication is nevertheless an ancient event, because Ato and Amos share no similarity outside the bHLH domains.

As might be expected from its sequence identity with Ato, Amos has the conserved residues required for correct folding of the bHLH domain and interaction with DNA (Ellenberger et al., 1994; Ma et al., 1994). Such conservation also leads us to predict that Amos functions as a heterodimer with Da protein. Indeed, Amos has been shown to bind to DNA as a heterodimer with Da in vitro (Huang et al., 2000 [this issue of *Neuron*]).

amos Has Proneural-like mRNA Expression in the Embryo

amos is expressed very transiently and dynamically during embryogenesis (Figures 2A and 2B). *amos* mRNA is present in a small cluster of ectodermal cells in each thoracic and abdominal segment during stage 10 (Figure 2A). Later, in stage 11, expression is restricted to a single cell per segment (Figure 2B). This cell quickly ceases to express *amos*, first in the thoracic segments and then in the abdomen. As with SOP formation by *ato* and the AS-C, this restriction of *amos* expression suggests that lateral inhibition functions within a proneural cluster defined by *amos*. Consistent with this, *amos* expression fails to resolve to single cells in *Notch* mutant embryos, which lack lateral inhibition (Figure 2C). Significant *amos* expression was also observed in developing head segments (including antennal, mandibular, and labial segments) in areas that correspond to the anlage of the olfactory sense organs of the larval antennomaxillary complex (Figure 2A) (Campos-Ortega and Hartenstein, 1997). The expression pattern of *amos* thus makes this gene a likely candidate for the AS-C- and *ato*-independent larval sense organs.

In addition to the above, *amos* mRNA was also transiently detected at the cellular blastoderm stage of embryogenesis in a dorsoventral band in the posterior of the embryo (data not shown) and during oogenesis in nurse cells, the centripetal follicle cells, and the oocyte itself (data not shown).

In Imaginal Discs, *amos* Is Expressed in the Proneural Domains of the Antennal Olfactory Sensilla

amos expression is extremely restricted during adult development. Very transient *amos* expression was detected in distal leg discs at approximately 0–4 hr after puparium formation (APF), correlating with the anlage of the innervated tarsal claw (Figure 2H). The main site of expression initiates in the antennal disc at approximately puparium formation (PF) in three semicircular bands on the medial side of the developing third segment (Figures 2D–2F). These three bands correspond to sites from which olfactory sensillum precursors are selected (Reddy et al., 1997). The two inner bands partly coincide with *ato* expression in the antennal disc (data not shown). The outer band (3) widens by 4–8 hr APF (Figure 2E) and persists until at least 21 hr APF, during which time the two inner bands (1 and 2) appear to fuse to give a single inner region of expression (Figure 2F).

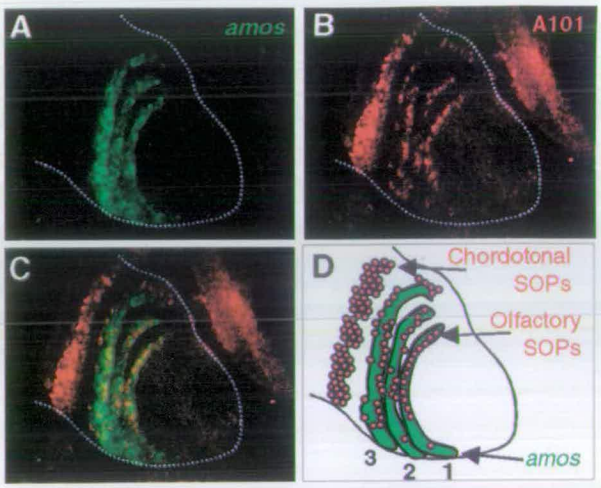


Figure 3. Olfactory SOPs Arise from *amos*-Expressing Bands

(A–C) Antennal disc from A101 enhancer trap line (SOP marker), 4 hr APF. SOPs (marked by β-galactosidase expression [red]) appear from the three bands of *amos* mRNA (green). Chordotonal SOPs are also seen in the second segment.

(D) Schematic representation of the staining patterns. Note that some SOPs do not show *amos* expression even though they are closely associated with the expression domains. Either *amos* is switched off rapidly after SOP formation or some olfactory SOPs never express *amos*.

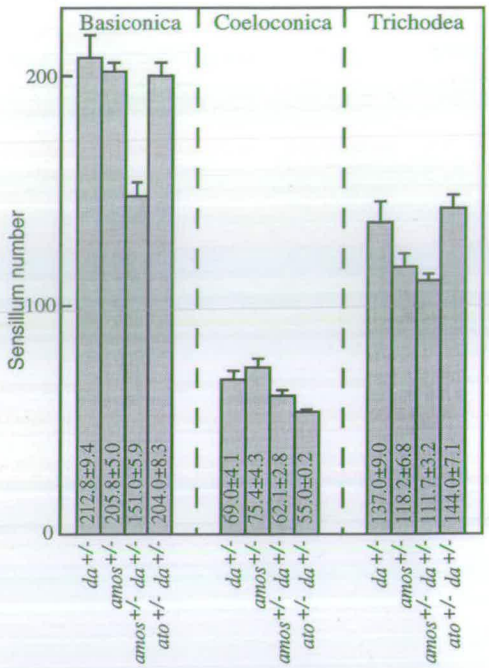


Figure 4. *amos* Interacts with *da* in Sensillum Basiconica Development

Graph of sensillum numbers in flies with reduced gene dosages of *amos*, *ato*, and/or *da*. $da^{+/-}$ represents flies with one copy of the *da* gene (*Df(2R)da*^{KX136}/+); $amos^{+/-}$ represents flies with one copy of a deficiency that deletes *amos* and its surrounding genomic area (*Df(2L)M36-S5*/+). Statistically similar figures were obtained for a second *amos*-containing deficiency (*Df(2L)M36-S6*/+). $ato^{+/-}$ represents flies heterozygous for the *ato*^{1} mutation. The mean ± SEM for six antennae were determined. Note that sensilla coeloconica numbers are rather variable, but the number is not significantly reduced in $amos^{+/-}$:$da^{+/-}$ flies compared with $amos^{+/-}$ or $da^{+/-}$ flies.

Double labeling of antennal discs from the A101 enhancer-trap line, which marks all SOPs, showed that SOPs indeed arise from the *amos*-expressing ectodermal bands (Figure 3). The expression pattern thus identifies *amos* as a candidate proneural gene for olfactory sensilla. Furthermore, we can infer that *amos* may be required for any or all three types of olfactory sensillum.

Genetic Evidence for the Requirement of Amos/Da Heterodimers in Olfactory SOP Formation

SOP formation is very sensitive to simultaneous reduction in copy number of proneural genes and *da*, which encodes their common heterodimer partner. For instance, simultaneously removing one copy of the AS-C and *da* genes (i.e., transheterozygotes) results in adults with a proportion of missing sensory bristles (Dambly-Chaudiere et al., 1988). Likewise, reducing the dosage of *ato* and *da* genes results in reduction in the number of chordotonal organs (A. P. J., unpublished data) and sensilla coeloconica (Gupta and Rodrigues, 1997).

Although there is no specific mutation of *amos*, we investigated lethal chromosomal deficiencies that delete *amos* (*Df(2L)M36F-S5* and *Df(2L)M36F-S6*) for genetic interaction with *da*. Olfactory sensillum numbers were analyzed in flies heterozygous for an *amos* deficiency either alone or in combination with the loss of one copy of *da* (*Df(2R)da*^{KX136}/+). In flies with a single copy of each gene (abbreviated as $amos^{+/-}$:$da^{+/-}$), the number of sensilla basiconica were significantly reduced (by 30%) compared with wild-type, $amos^{+/-}$, $da^{+/-}$, or $ato^{+/-}$:$da^{+/-}$ flies (Figure 4). This genetic interaction suggests functional cooperation between *da* and a bHLH gene in the *amos* genomic region. Given that *amos* is the only bHLH-encoding gene in this region, these data are consistent with the function of Amos/Da heterodimers during the formation of sensilla basiconica. Sensilla trichodea were also significantly reduced in $amos^{+/-}$:$da^{+/-}$ flies, but we also observed a reduction in $amos^{+/-}$ flies compared with wild-type. Therefore, although consistent with a requirement for *amos* in trichodea formation, such a requirement seems to be less sensitive to *da* gene dosage. Although sensillum coeloconica numbers are rather variable (see also Gupta et al., 1998), $amos^{+/-}$:$da^{+/-}$ flies have only slightly fewer sensilla coeloconica than flies with either mutation alone, whereas there are significantly fewer in $ato^{+/-}$:$da^{+/-}$ flies as expected. In summary, these data support a role for the chromosomal region containing *amos* in sensillum basiconica formation and are suggestive of a role during sensillum trichodea development.

Misexpression of *amos* Increases Olfactory Sensillum Numbers on the Antenna

One shared characteristic of proneural genes is that ectopic expression leads to ectopic sense organ formation, which is consistent with their neural competence function. Moreover, specific subtypes of ectopic organ are yielded, demonstrating that different proneural genes regulate distinct neuronal subtype properties (Jarman et al., 1993, 1995; Jarman and Ahmed, 1998).

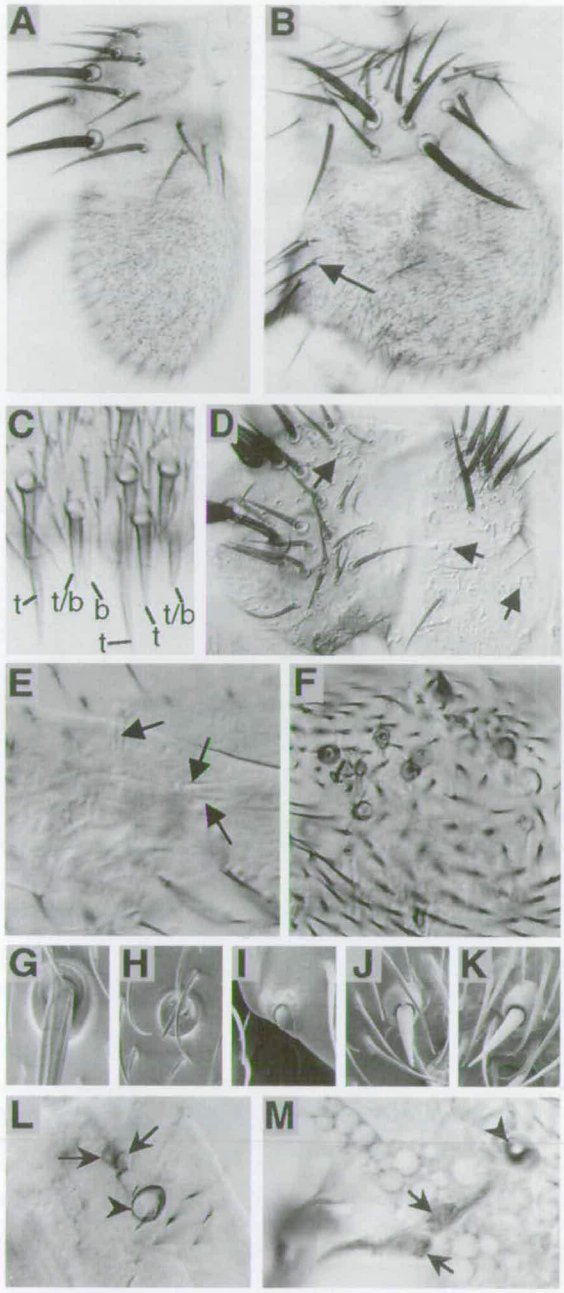


Figure 5. *amos* Misexpression Results in Increased Olfactory Sensillum Formation

(A) Wild-type antenna.

(B) *Gal4*^{OK384}/UAS-*amos*, showing antenna bloated with a large excess of olfactory sensilla. Occasional sensory bristles are also formed (arrow).

(C) Closeup of *Gal4*^{OK384}/UAS-*amos* antenna, showing apparent hybrid trichoid/basiconic sensilla (t/b) that are shaped like basiconica (b) but more pigmented like trichodea (t).

(D) *hsGal4*/UAS-*amos*, first and second antennal segments (third segment is beyond bottom left of picture) showing large numbers of ectopic olfactory-like sensilla compared with (A) (arrows indicate examples) as well as large dark bristles.

(E and F) *Gal4*$^{109\text{-}68}$/UAS-*amos* wing, middle of the third wing vein.

(E) Scolopales of ectopic chordotonal organs are indicated (arrows). This area normally contains no chordotonal organs.

(F) Ectopic sensilla of olfactory-like morphology on the wing.

(G–K) SEM of wild-type and *amos*-induced sensilla on thorax (G and H) and wing (I–K).

When AS-C genes are overexpressed, they each give a similar phenotype of ectopic external sense organs (Rodríguez et al., 1990; Brand et al., 1993; Hinz et al., 1994). Misexpression of *ato* results largely in the formation of ectopic chordotonal organs (Jarman et al., 1993) and sensilla coeloconica (Gupta and Rodrigues, 1997). We carried out *Gal4*/UAS misexpression experiments to determine whether *amos* can function as a proneural gene and assess its subtype-determining properties. Using *Gal4*^{OK384} to drive expression in the larval eye-antennal disc, misexpression of UAS-*ato* causes increased formation of sensilla coeloconica (Gupta and Rodrigues, 1997). Upon misexpression of UAS-*amos*, we observed a dramatic increase in olfactory sensilla (Figures 5A and 5B). The antenna itself was bloated, perhaps a secondary effect of housing more sensilla. Accurate counting of sensilla was difficult because of the antennal morphology and uncertainty in assigning sensilla to specific classes. Nevertheless, sensillum counts suggested that there was an increase of 24%–44% for all sensillum types. Thus, unlike *ato*, *amos* seems capable of directing formation of sensilla basiconica and trichodea. In addition, extra sensillum coeloconica formation may result from an ability of *amos* to substitute functionally for (mimic) *ato* (see below). Interestingly, a proportion of sensilla appeared to be of intermediate basiconica/trichodea morphology (Figure 5C). This has also been reported as an effect of *lz* misexpression (Gupta et al., 1998).

Outside the Antenna, *amos* Misexpression Mimics *ato* and Also Promotes Ectopic Olfactory-like Sensilla

The subtypes of ectopic sense organ produced by misexpression of *ato* are dependent on the tissue. Thus, it directs sensilla coeloconica formation on the third antennal segment but chordotonal formation in most other sites. To see how *amos* function might be modulated outside the antenna, we induced *amos* expression in many proneural clusters in all imaginal discs using *Gal*$^{109\text{-}68}$ (Jarman and Ahmed, 1998). With this driver, UAS-*amos* misexpression resulted in a mixture of ectopic sense organs (Figures 5E–5M). Part of this phenotype resembles that obtained from *ato* misexpression (Jarman and Ahmed, 1998). First, ectopic chordotonal organs are formed in the scutellum and third wing vein (Figure 5E). In the scutellum, these were similar in number to that induced by strong *ato* misexpression, although much more disorganized (UAS-*amos*: 69 ± 13, n = 4, UAS-*ato*: 70.3 ± 10.3 [Jarman and Ahmed, 1998]). Therefore, *amos* can indeed mimic *ato* outside its normal

(G) Wild-type microchaeta (external sense organ).

(H) UAS-*amos* microchaeta transformed to resemble an olfactory sensillum.

(I) Wild-type sensillum campaniformia (external sense organ).

(J and K) UAS-*amos* induced olfactory-like sensilla that closely resemble sensilla trichodea (Figure 1B).

(L and M) UAS-*amos*: polyinnervation of transformed sensilla on the thorax. Dissected adult thoraces stained with neuronal antibodies mAb22C10 (L) or αHRP (M) showing doubled sensory neuron cell bodies (arrows) innervating olfactory-like sensilla (arrowheads). In the wild-type thorax, all bristles are innervated by a single neuron.

developmental context, as suspected above. Second, some extra external sense organs are formed, mostly along the third wing vein (number of ectopic bristles: 4.8 ± 1, n = 5 compared with 5.4 ± 0.7, n = 5 for UAS-*ato*). In the case of *ato*, such bristle formation has been interpreted as an artefact resulting from incomplete subtype determination of ectopic SOPs under the conditions of misexpression, external sense organ apparently being a default fate for such SOPs (Jarman and Ahmed, 1998).

Strikingly, *amos* misexpression also produces a third phenotype: on the thorax, head, and along the third wing vein were unusual organs that appeared morphologically similar to olfactory sensilla, particularly sensilla trichodea (Figures 5D, 5F, and 5I–5K). Although some were difficult to assign to bristle or olfactory-like classes, we estimated that 19.4 ± 2.8 olfactory-like sensilla were observed along the third wing vein (n = 5). Moreover, at high expression levels, some of the sensory bristles on the thorax were transformed toward similar olfactory-like morphology (16 ± 3.3 [n = 4] transformed bristles, Figures 5G, 5H, and 5L). While wild-type thoracic bristles are always innervated by a single sensory neuron, a small proportion of these *amos*-converted sensilla were found to have two or more sensory neurons, a characteristic of olfactory sensilla (Figures 5L and 5M). Thus, *amos* can form olfactory-like sensilla outside its normal developmental context of the antenna. It also appears able to impose olfactory sensillum fate on external sense organs in a manner similar to the chordotonal fate transformation imposed by *ato* (Jarman and Ahmed, 1998). The ability of *amos* to form nonantennal olfactory sensilla is apparently not shared by *ato*, since strong *ato* misexpression resulted in no or very few ectopic olfactory-like sensilla (1.2 ± 0.4 possible olfactory-like sensilla along the third wing vein [n = 5]), and no olfactory-like transformation of thoracic bristles.

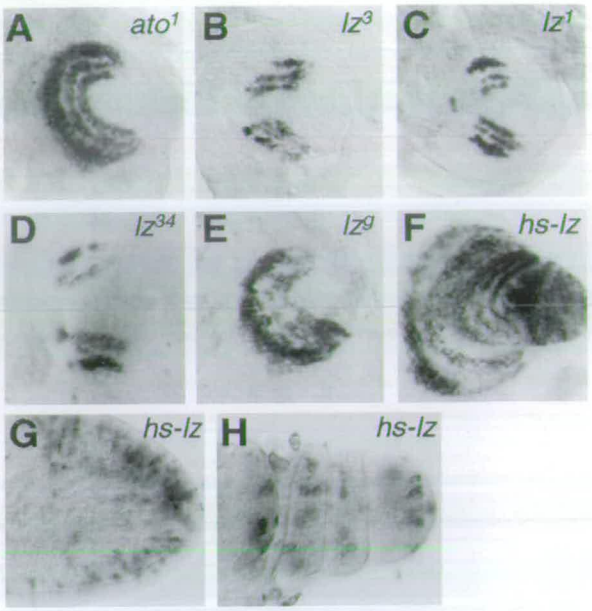


Figure 6. *amos* Expression Is Regulated by *lz*

Antennal discs 0–4 hr APF stained to detect *amos* RNA.

(A) *ato*^{1} mutant, showing no difference from wild-type (Figure 2).

(B–E) *amos* expression is reduced in *lz* mutants.

(B) *lz*^{3}, showing loss of medial *amos* RNA from all three bands. The remaining expression appears patchier.

(C) *lz*^{1}, showing a similar effect, but with a small group of expressing cells remaining in the medial band 3.

(D) *lz*^{34}, at a slightly later stage, showing loss of medial expression.

(E) *lz*^{9}, showing patchy expression.

(F–H) Imaginal discs from hs-*lz* pupae at 0–4 hr APF, showing that *amos* expression is ectopically induced by ubiquitous *lz* expression.

(F) Antenna; *amos* expression is delocalized, but a ring-like pattern can be discerned.

(G) Wing, showing widespread, patchy *amos* induction. No *amos* expression is observed in the wild-type wing.

(H) Leg (c/f Figure 2H).

amos Expression Is Regulated by *lz*

lz functions in eye, antennal, and tarsal claw development (Lindsley and Zimm, 1992). In leg and antennal discs, the pattern of *lz* expression strongly resembles that of *amos* (Figures 2I and 2J; Gupta et al., 1998), although *amos* expression begins later than *lz*. These observations suggest that *lz* might regulate *amos* expression during the process leading to the formation of basiconic and trichoid SOPs in the antenna and perhaps in the tarsal claw. Therefore, we looked for changes in the expression pattern of *amos* in *lz* mutants. Strong *lz* alleles (including *lz*^{1}, *lz*^{3}, and *lz*^{34}) almost completely lack sensilla basiconica and exhibit up to a 50% reduction in sensilla trichodea (Stocker et al., 1993). The number of sensilla coeloconica is reported to be unaffected. In these strong alleles, *amos* mRNA was absent from the middle of all three antennal bands (Figures 6B–6D). For band 3, the affected region corresponds to the area fated to form sensilla basiconica SOPs (Gupta et al., 1998). The correlation between this loss of *amos* expression and the loss of sensilla basiconica is therefore consistent with a requirement for *amos* in sensillum basiconica formation. In addition, we may deduce that the middle regions of the other two bands (2 and 1) give rise to those sensilla trichodea that are missing in strong *lz* mutants. Conversely, *lz*-independent sensilla trichodea may arise from the *lz*-independent tips of the *amos*-expressing bands. Topologically, SOPs from the band tips will end up on the lateral edge of the antenna after metamorphosis, which is where the sensilla trichodea are concentrated (Figure 1A). Interestingly, comparison with the *ato* expression pattern suggests that *amos* is also expressed in regions of sensillum coeloconica formation in bands 1 and 2 (Gupta et al., 1998). Since these SOPs are not lost in *lz* mutants, the loss of *amos* expression from the middle of these regions provides evidence that *amos* is not required at least for many sensilla coeloconica.

In weaker *lz* alleles (such as *lz*^{9}), *amos* expression appears patchy but spatially normal (Figure 6E), suggesting that SOP selection itself is not strongly altered. This would be consistent with observation that the major phenotype of weak *lz* alleles is one of subtype transformation from basiconic to trichoid fate rather than sensillum loss. This is postulated to result from a role of *lz* in subtype specification, such that higher levels are required for SOPs to take on basiconic fate while lower levels are sufficient for trichoid fate (Gupta et al., 1998).

We determined in a complementary experiment whether

ectopic *lz* expression could induce ectopic *amos* expression. When ubiquitous *lz* expression was activated in pupae containing a heatshock-inducible *lz* construct (hs-*lz*) (Gupta et al., 1998), a strong expansion of *amos* expression was observed in the antenna (Figure 6F). *lz* misexpression also resulted in ectopic *amos* expression in pupal wings and legs (Figures 6G and 6H). These experiments show that *lz* is both necessary for much of *amos's* expression pattern and also sufficient to drive ectopic *amos* activation in many other locations.

To investigate further the relationship between *lz* and *amos*, we determined whether *amos* gene dosage reduction would modify the number of sensilla formed in *lz* mutants. In the intermediate allele, lz^g, the number of basiconica is reduced to 28% of wild-type (56.7 ± 2.6 sensilla basiconica). Removing one copy of the chromosomal region containing *amos* results in a further 70% reduction in this number (17.0 ± 3.1). The number of sensilla trichodea was unaltered, probably because these are not affected in this intermediate *lz* allele. To gauge the effect on sensilla trichodea, we therefore examined *amos's* modification of a strong *lz* allele, lz^3. In addition to a total lack of sensilla basiconica, lz^3 exhibits a strong reduction of sensilla trichodea (78.8 ± 3.5, 50% of wild-type). We found that, in the absence of one copy of *amos*, sensilla trichodea were reduced by a further 54% in lz^3 (36.3 ± 1.3, therefore to 24% of wild-type).

amos Misexpression Can Bypass the Requirement for *lz* in Sensillum Basiconica Formation

From the genetic and expression analyses, we conclude that *amos* transcription is partly downstream of *lz* and that its loss of expression may explain the loss of sensilla basiconica and trichodea in *lz* mutants. We therefore tested whether experimentally induced *amos* expression could rescue the loss of sensilla basiconica in strong *lz* mutants. Using *hsGal4* as a driver, UAS-*amos* was misexpressed in lz^3 pupal antennae. Such misexpression resulted in a significant recovery of sensilla basiconica (36.2 ± 8) when compared with lz^3 alone (0.33 ± 0.33 sensilla basiconica) (Figures 7A and 7B). This rescue is still far short of wild-type levels, perhaps because *amos* is not optimally expressed using *hsGal4*. Alternatively, *lz* might need to activate other genes required for basiconic fate in addition to *amos* (i.e., *amos* alone cannot replace all the functions of *lz*). Significantly, *ato* was unable to direct any rescue under the same conditions (1.0 ± 0.6 sensilla), even though the number of sensilla coeloconica was increased (data not shown). Therefore, *amos*, but not *ato*, can partially bypass the requirement for *lz*. Interestingly, many of the rescued basiconica were located in the lateral region of the antenna (T domain; Figure 6A). Such a distribution was also observed upon rescue of *lz* mutants by hs-*lz* (Gupta et al., 1998).

Evidence that *amos* and *ato* Are Independently Regulated

Since the expression of *ato* overlaps with the inner two bands of *amos* expression, it is possible that one gene may be dependent on the other. However, we observed no defect in *amos* expression in *ato* mutant antennal discs (Figure 6A). Furthermore, it was reported that *ato*

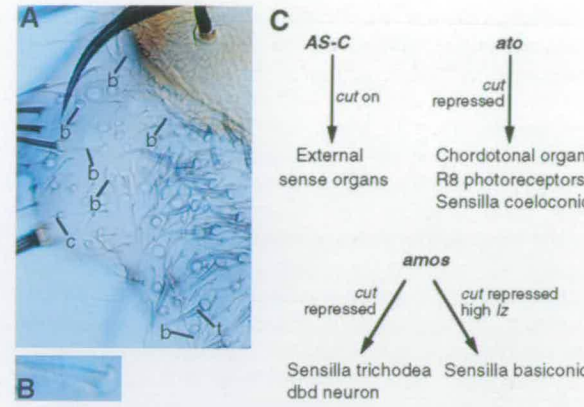


Figure 7. Relationship of *amos* with *lz*

(A and B) *amos* misexpression can rescue a proportion of the sensilla basiconica missing in *lz* mutants.

(A) Lateral region of antenna from lz^3; *hs-Gal4*/UAS-*amos* fly, showing a number of sensilla basiconica (b). Normally, lz^3 antennae have no basiconica. Representative sensilla coeloconica (c) and trichodea (t) are also indicated.

(B) Higher magnification view of one of the replaced sensilla basiconica.

(C) Summary of genetic relationships in SOP formation. The three types of *Drosophila* proneural gene are required for different subtypes of sense organ. This specificity is contingent on developmental context such that a proneural gene can be responsible for different sense organs in different anatomical locations. Subtype specificity may be achieved by the differential regulation of subtype selector genes (such as *cut*) or cooperation with subtype selector genes that are expressed in parallel to proneural genes (perhaps *lz* or a *lz*-regulated gene).

expression is not dependent on *lz* (Gupta et al., 1998), and therefore by inference *ato* does not depend on *amos* at least in the medial antennal region. We conclude that the two olfactory proneural genes, *ato* and *amos*, are largely independent of each other. Furthermore, *ato* shows no interaction with *lz*. Thus, lz^3; ato^1/ato^1 double mutants exhibited a complete absence of sensilla basiconica and coeloconica, as expected from the loss of *lz* and *ato* functions, respectively. However, the number of sensilla trichodea was not reduced below that observed in lz^3 mutant flies (data not shown). This suggests that there is no redundancy between *lz* and *ato* in formation of the remaining sensilla trichodea, which are instead likely to require the *lz*-independent part of *amos's* expression.

Discussion

amos as a Proneural Gene

A combination of sequence, expression pattern, loss-of-function, and gain-of-function data leads us to conclude that *amos* is a bHLH proneural gene for a subset of olfactory sensilla. With the discovery of *amos*, the origin of almost the entire larval and adult PNS can be accounted for by *amos*, *ato*, and the AS-C. Olfactory sensilla require two proneural genes of the *ato*-like subfamily, *amos* and *ato* itself. These genes appear to have complementary functions. While *ato* is required for sensilla coeloconica (Gupta and Rodrigues, 1997), our data strongly suggest that *amos* can account for the formation of the sensilla basiconica and trichodea, although

definitive evidence will require the isolation of specific *amos* mutations. The evidence is clearest for sensilla basiconica in that a genetic interaction was detected between *da* and *amos*. Although the *amos*-containing deficiencies used in these experiments delete other genes in addition to *amos*, the interaction with *da* strongly implicates *amos*, since no other candidate bHLH gene exists in the 170 kb region around *amos*. For the sensilla trichodea, evidence from *da* interaction is more equivocal, but a strong genetic interaction was detected between *amos* and *lz*, a gene required for the formation of many sensilla trichodea. In contrast, sensilla coeloconica do not appear to require *amos*. Notably, *amos* expression is absent from regions of coeloconic SOP formation in *lz* mutants.

During embryogenesis, the expression pattern of *amos* also suggests a function in the development of larval olfactory sensilla of the antennomaxillary complex (Carlson, 1996), and preliminary double-stranded RNA interference (RNAi) experiments are consistent with this (S. E. G. and A. P. J., unpublished data). Additionally, the expression of *amos* in embryonic abdominal and thoracic segments suggests a function in the formation of at least one SOP. Evidence from RNAi experiments suggests that this SOP forms the dbd neuron (data not shown), which is AS-C and *ato* independent (Jarman et al., 1993). Moreover, Chien and colleagues have also isolated *amos*, and provide strong loss-of-function and gain-of-function evidence for its requirement in dbd neuron formation (Huang et al., 2000). The function of these embryonic neurons is unknown, but it appears unlikely from their morphology that they are olfactory. *amos* expression during dbd SOP formation is typically proneural: expression begins in a small proneural cluster of ectodermal cells and then is rapidly refined to the dbd SOP in a Notch-dependent process. In contrast, the pupal antennal expression shows unusual features. Here, *amos* is expressed persistently in ectodermal bands from which multiple olfactory SOPs arise continuously. This superficially resembles the persistent ectodermal expression of *ato* during continuous chordotonal SOP formation in the leg disc (zur Lage and Jarman, 1999). Like the leg chordotonal array, persistent PNC expression may support continuous olfactory SOP determination. The persistent ectodermal expression of *amos* may also reflect another unusual feature of olfactory sensillum development: the recruitment of support cells by the SOP (or founder cell, Reddy et al., 1997). It can be envisaged that these cells must be recruited from the *amos*-expressing bands in order to attain their correct fate. It is notable that *ato* expression in coeloconic SOP formation is more orthodox in that it proceeds from small proneural clusters to single SOPs (Gupta and Rodrigues, 1997).

lz as a Prepattern Gene Regulating *amos* Expression

We have demonstrated that one role of *lz* is to regulate *amos* expression, since *lz* is both necessary for much of *amos's* antennal expression pattern and sufficient to drive ectopic *amos* expression. The partial loss of sense organs in *lz* mutants can thus be accounted for by a commensurate loss of *amos* expression. Thus, consistent with the proposal by Gupta and Rodrigues (1998), *lz* is acting as a prepattern gene. That is, Lz does not function as a cell type-specific transcription factor itself, but to set up the correct expression patterns of fate determining factors. This function is closely analogous to its function in eye development (Daga et al., 1996; Flores et al., 1998), where it is also expressed prior to fate determination and helps pattern the expression of the fate-determining transcription factors BarHI, Sevenup, and Prospero. It is clearly possible that *amos* is a direct target gene of *lz*, but it is also interesting that the induction of *amos* by hs-*lz* is not ubiquitous. In the antenna, for instance, ubiquitous induction of *lz* leads to *amos* activation in a pattern of rings, suggesting that *lz* acts in concert with other prepatterning activities in the antenna (Figure 6F). It is notable that the *lz–amos* regulatory link may also function during tarsal claw development.

Neuronal Subtype Specificity of bHLH Proneural Genes

It is increasingly apparent that proneural bHLH factors not only confer neural competence for SOP formation but also endow SOPs with neural subtype information. While expression and loss-of-function analyses show that *amos*, *ato*, and the AS-C are required for different subtypes of sense organ, a direct test of the functional specificity of proneural proteins comes from comparison of their capabilities in misexpression assays. Misexpression of all proneural genes results in ectopic SOP formation, but the subtype of the sense organ depends on both the gene misexpressed and the developmental context of misexpression (Rodríguez et al., 1990; Brand et al., 1993; Jarman et al., 1993; Jarman and Ahmed, 1998). In the context of antennal development, misexpression of *amos* or *ato*, but not an AS-C gene (Reddy et al., 1997), drives ectopic olfactory sensillum formation. This suggests that the shared features of the *amos* and *ato* bHLH domains provide functional information important for olfactory sensillum determination. One of the specific functions ascribed to *ato* in chordotonal SOP formation is the inhibition of *cut*, a subtype selector gene for external sense organs (Jarman and Ahmed, 1998). Olfactory sensilla also do not require *cut*, and so inhibition of *cut* expression might be an important conserved function of *ato* and *amos* in olfactory SOP formation (Figure 7C). Another shared function may be the activation of cell recruitment, since *amos*- and *ato*-dependent sense organs (olfactory sensilla, chordotonal organs, and ommatidia) all require some form of local cell recruitment in their development. The recruitment of cells to an ommatidium and the recruitment of SOPs into chordotonal clusters both require EGFR signaling, perhaps under the direct control of *ato* (Freeman, 1996; zur Lage et al., 1997; zur Lage and Jarman, 1999; N. M. White and A. P. J., submitted). It is not yet known whether EGFR is involved in the process of cell recruitment by olfactory SOPs (Reddy et al., 1997), but it is clearly possible that *ato* and *amos* directly control the recruitment process.

Outside the antenna, misexpressed *ato* drives ectopic chordotonal organ formation (Jarman et al., 1993). This functional specificity was found to reside in *ato's* bHLH domain and largely in its basic region (Chien et al.,

1996). *ato's* basic region is completely conserved in *amos* apart from an R to K conservative substitution. Not surprisingly, therefore, *amos* can also mimic *ato* in directing chordotonal organ formation when misexpressed outside the antenna. Nevertheless, despite *amos's* ability to mimic *ato* and their common requirement in olfactory sensillum formation, these genes are not functionally interchangeable. Differences in their ability to determine olfactory sensillum subtypes are initially suggested by their differing loss-of-function phenotypes despite their partly overlapping expression domains. Also, Gupta and Rodrigues (1997) reported that *ato* misexpression in the antenna specifically results in increased sensillum coeloconica numbers, while *amos* increases all sensilla types in our experiments and can rescue the loss of sensilla basiconica in *lz* mutants. At the least, this shows that *amos* can supply some functional information that *ato* apparently cannot (i.e., basiconic/trichoid fate). Most significantly, only *amos* misexpression can drive ectopic sensilla of olfactory-like morphology outside the antenna. Thus, unlike *ato*, *amos* function in olfactory sensillum formation appears to be able partially to overcome the constraints of developmental context. Presumably, one or more target genes must be differentially regulated by *amos* and *ato*.

In contrast to these functional differences between *amos* and *ato*, misexpression of the different genes of the AS-C produces identical phenotypes of ectopic external sense organs (Rodríguez et al., 1990; Brand et al., 1993; Hinz et al., 1994), even though the level of sequence identity they share in their bHLH domains is similar to that observed between *amos* and *ato*. It will be important to determine whether the sequence basis for the functional differences between *amos* and *ato* resides in the 26% of divergent residues in their bHLH domains or in other parts of the protein that are not conserved with *ato*. It is notable, however, that most of the amino acid differences between the *ato* and *amos* bHLH domains are conserved in their respective *Drosophila virilis* orthologs, implying functional constraint (I. Ahmed and A. P. J., unpublished data).

lz as a Modulator of *amos* Subtype Specificity

In addition to its role in *amos* activation and SOP commitment, *lz* also influences the choice between basiconic and trichoid fate (Stocker et al., 1993), it being recently proposed that high *lz* levels are required for basiconica fate and low levels for trichodea fate (Gupta et al., 1998). One possibility is that *amos*-dependent SOPs are competent to form basiconica or trichodea, depending on the level of *lz* to which they are exposed. In the absence of *lz*, remaining *amos*-dependent SOPs assume a "basal" trichoid fate. High levels of *lz* modify the specificity of *amos* so that SOPs take on the alternative basiconic fate. Thus, *lz* requirement in sensillum trichodea formation would be at the level of *amos* activation, whereas for sensilla basiconica *lz* is required not only to activate *amos* but also to modulate its subtype-determining properties. This might explain why *amos* misexpression could not completely bypass the need for *lz* in the formation of sensilla basiconica. However, the fact that some rescue was observed suggests *amos* itself also influences this decision directly, perhaps through differences in level or timing of expression. Supporting the possibility that both *lz* and *amos* influence this fate decision, it is significant that misexpression of either gene yields some sensilla of intermediate morphology (Gupta et al., 1998).

To explain the context-dependent subtype specificity of proneural proteins, it has been proposed that they require regionally restricted transcriptional cofactors that directly modulate their specificity (Chien et al., 1996; Jarman and Ahmed, 1998). No such cofactors are yet known. For Amos, however, Lz could be a candidate cofactor for the choice between basiconic and trichoid sensillum fate. *lz* expression, however, ceases at 10 hr APF while *amos* expression continues for longer, raising the possibility that *lz* functions to activate an unknown subtype-modulating cofactor in addition to activating *amos*.

Experimental Procedures

PCR

Degenerate primers were 5′-GCYGCYAAYGCHCGYGARMG-3′ and 5′-TGRGCCATYTGBARDGTYTC-3′, designed to peptide regions AANARER and ETLQMAQ, respectively. PCR conditions were 30 cycles of 55°C (1 min), 72°C (1 min), and 94°C (1 min). PCR products were cloned into pGEM-T and plasmid DNA from nine clones sequenced. Sequence analysis of these showed the presence of *ato*, *cato*, and *amos*. Futher colony hybridization studies suggested that no other genes were represented among the clones of the PCR product.

Molecular Analysis

A *Drosophila* λFIXII genomic library was screened to isolate genomic DNA that flanked the *cato* bHLH domain. Genomic subclones were then used to screen a λZAP-express embryonic cDNA library. Both libraries were constructed by N. White (S. E. G., N. M. White, and A. P. J., submitted). Two independent cDNAs were isolated for *amos*. Sequencing of cDNAs and genomic clones was using Big Dye terminator kit (Perkin Elmer). The gene was localized to 36F2–6 by in situ hybridization to polytene chromosomes and the presence of the gene in YACs from this location. For expression analysis, mRNA was isolated by acidic phenol and purified using the PolyATtract system III (Promega). Northern blotting was by standard protocols.

Immunohistochemistry

Wholemount in situ hybridization was achieved using digoxygenin-labeled RNA probes as described (Tautz and Pfeifle, 1989). Double labellings were carried out by detecting RNA followed by protein (zur Lage and Jarman, 1999).

Counting Olfactory Sensilla

Antennae from adult males were cleared and mounted in Hoyers mountant. Sensilla were scored by video projection of a light microscope image and marking on an overlying acetate sheet. In each experiment, from three to six antennae were scored and the results presented as mean and standard error.

Fly Stocks

All deficiencies and *lz* mutations were from the Bloomington or Umea stock centers. Deficiencies used to remove *amos* function were *Df(2L)M36F-S5* (breakpoints, 36D1-3;36F7-11) and *Df(2L)M36F-S6* (breakpoints, 36F2-6;36F2-6). Deletion of *amos* was confirmed by loss of *amos* DNA and RNA from homozygous embryos. *Gal4*^{OK384} and hs-*lz* stocks were obtained from C. O'Kane and U. Banerjee, respectively. For UAS-*amos*, the *amos* ORF was cloned into pUAST and transformant flies produced by microinjection. Lines of varying expression level were obtained, and most of the experiments described here were using a strongly expressing line.

Misexpression

Crosses were performed at 25°C. For experiments with *hsGal4*, heatshocks were performed on 0–4 hr APF pupae for 20 min at 39°C. For hs-*lz*, 0–4 hr APF pupae were heatshocked for 30 min at 39°C and then allowed to recover at 25°C for 1 hr before being dissected for in situ hybridization.

Acknowledgments

We thank D. Hom, J. Ferguson, N. White, R. Kirby, and S. Khalid for technical help with various aspects of this work. We thank C-T. Chien for discussions and exchange of results prior to publication, and U. Banerjee and C. O'Kane for *lz* reagents and *Gal4* lines, respectively. This work was supported by a Darwin Trust studentship (S. E. G.) and The Wellcome Trust (project grant 055851 and Wellcome Trust Senior Fellowship 042182).

Received September 20, 1999; revised October 19, 1999.

References

Brand, M., Jarman, A.P., Jan, L.Y., and Jan, Y.N. (1993). *asense* is a *Drosophila* neural precursor gene and is capable of initiating sense organ formation. Development *119*, 1–17.

Cabrera, C.V., and Alonso, M.C. (1991). Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products in *Drosophila*. EMBO J. *10*, 965–973.

Cabrera, C.V., Alonso, M.C., and Huikeshoven, H. (1994). Regulation of Scute function by Extramacrochaete in vitro and *in vivo*. Development *120*, 3595–3603.

Campos-Ortega, J.A., and Hartenstein, V. (1997). The embryonic development of *Drosophila melanogaster*, 2nd Edition (Berlin: Springer).

Campuzano, S., and Modolell, J. (1992). Patterning of the *Drosophila* nervous system: the *achaete-scute* gene complex. Trends Genet. *8*, 202–208.

Carlson, J.R. (1996). Olfaction in *Drosophila*: from odor to behavior. Trends Genet. *12*, 175–180.

Chien, C.-T., Hsiao, C.-D., Jan, L.Y., and Jan, Y.N. (1996). Neuronal type information encoded in the basic-helix-loop-helix domain of proneural genes. Proc. Natl. Acad. Sci. USA *93*, 13239–13244.

Clyne, P.J., Warr, C.G., Freeman, M.R., Lassing, D., Kim, J., and Carlson, J.R. (1999). A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. Neuron *22*, 327–338.

Daga, A., Karlovich, C.A., Dumstrei, K., and Banerjee, U. (1996). Patterning of cells in the Drosophila eye by Lozenge, which shares homologous domains with AML1. Genes Dev. *10*, 1194–1205.

Dambly-Chaudiere, C., Ghysen, A., Jan, L.Y., and Jan, Y.N. (1988). The determination of sense organs in *Drosophila*: interaction of *scute* with *daughterless*. Roux's Arch. Dev. Biol. *197*, 419–423.

Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S.C. (1994). Crystal structure of transcription factor E47: E-box recognition by a basic region helix-loop-helix dimer. Genes Dev. *8*, 970–980.

Flores, G.V., Daga, A., Kalhor, H.R., and Banerjee, U. (1998). Lozenge is expressed in pluripotent precursor cells and patterns multiple cell types in the Drosophila eye through control of cell-specific transcription factors. Development *125*, 3681–3687.

Freeman, M. (1996). Reiterative use of the EGF receptor triggers differentiation of all cell types in the *Drosophila* eye. Cell *87*, 651–660.

Ghysen, A., Dambly-Chaudière, C., Jan, L.Y., and Jan, Y.N. (1993). Cell interactions and gene interactions in peripheral neurogenesis. Genes Dev. *7*, 723–733.

Gupta, B.P., and Rodrigues, V. (1997). *atonal* is a proneural gene for a subset of olfactory sense organs in *Drosophila*. Genes Cells *2*, 225–233.

Gupta, B.P., Flores, G.V., Banerjee, U., and Rodrigues, V. (1998). Pattterning an epidermal field: *Drosophila* Lozenge, a member of the AML/Runt family of transcription factors, specifies olfactory sense organ type in a dose-dependent manner. Dev. Biol. *203*, 400–411.

Hinz, U., Giebel, B., and CamposOrtega, J.A. (1994). The basic-helix-loop-helix domain of *Drosophila* lethal of scute protein is sufficient for proneural function and activates neurogenic genes. Cell *76*, 77–87.

Huang, M-L., Hsu, C-H., and Chien, C-T. (2000). The proneural gene *amos* promotes multiple dendritic neuron formation in the *Drosophila* peripheral nervous system. Neuron, this issue, 57–67.

Jan, Y.N., and Jan, L.Y. (1993). The peripheral nervous system. In The Development of *Drosophila melanogaster*, M. Bate and A. Martinez-Arias, eds. (New York: Cold Spring Harbor Press), pp. 1207–1244.

Jarman, A.P., and Jan, Y.N. (1995). Multiple roles for proneural genes in *Drosophila* neurogenesis. In Neural Cell Specification: Molecular Mechanisms and Neurotherapeutic Implications, B.H.J. Juurlink, P.H. Krone, W.M. Kulyk, V.M.K. Verge, and J.R. Doucette, eds. (New York: Plenum Press), pp. 97–104.

Jarman, A.P., and Ahmed, I. (1998). The specificity of proneural genes in determining *Drosophila* sense organ identity. Mech. Dev. *76*, 117–125.

Jarman, A.P., Grau, Y., Jan, L.Y., and Jan, Y.N. (1993). *atonal* is a proneural gene that directs chordotonal organ formation in the *Drosophila* peripheral nervous system. Cell *73*, 1307–1321.

Jarman, A.P., Grell, E.H., Ackerman, L., Jan, L.Y., and Jan, Y.N. (1994). *atonal* is the proneural gene for *Drosophila* photoreceptors. Nature *369*, 398–400.

Jarman, A.P., Sun, Y., Jan, L.Y., and Jan, Y.N. (1995). Role of the proneural gene, *atonal*, in formation of *Drosophila* chordotonal organs and photoreceptors. Development *121*, 2019–2030.

Lindsley, D.L., and Zimm, G.G. (1992). The Genome of *Drosophila melanogaster* (San Diego: Academic Press).

Ma, P.C.M., Rould, M.A., Weintraub, H., and Pabo, C.O. (1994). Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. Cell *77*, 451–459.

Ray, K., and Rodrigues, V. (1995). Cellular events during development of the olfactory sense organs in *Drosophila melanogaster*. Dev. Biol. *167*, 426–438.

Reddy, G.V., Gupta, B., Ray, K., and Rodrigues, V. (1997). Development of the *Drosophila* olfactory sense organs utilizes cell-cell interactions as well as lineage. Development *124*, 703–712.

Riesgo-Escovar, J.R., Peikos, W.B., and Carlson, J.R. (1997). The *Drosophila* antenna: ultrastructural and physiological studies in wild-type and *lozenge* mutants. J. Comp. Physiol. A *180*, 151–160.

Rodríguez, I., Hernández, R., Modolell, J., and Ruiz-Gómez, M. (1990). Competence to develop sensory organs is temporally and spatially regulated in *Drosophila* epidermal primordia. EMBO J. *9*, 3583–3592.

Siddiqi, O. (1987). Neurogenetics of olfaction in *Drosophila melanogaster*. Trends Genet. *3*, 137–142.

Stocker, R.F., Gendre, N., and Batterham, P. (1993). Analysis of the antennal phenotype in the *Drosophila* mutant *lozenge*. J. Neurogenet. *9*, 29–53.

Tautz, D., and Pfeifle, C. (1989). A nonradioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translation control of the segmentation gene *hunchback*. Chromosoma *98*, 81–85.

Van Doren, M., Ellis, H.M., and Posakony, J.W. (1991). The *Drosophila* Extramacrochaetae protein antagonizes sequence-specific DNA binding by Daughterless/Achaete-Scute protein complexes. Development *113*, 245–255.

zur Lage, P., and Jarman, A.P. (1999). Antagonism of EGFR and Notch signaling in the reiterative recruitment of *Drosophila* adult chordotonal sense organ precursors. Development *126*, 3149–3157.

zur Lage, P., Jan, Y.N., and Jarman, A.P. (1997). Requirement for EGF receptor signaling in neural recruitment during formation of *Drosophila* chordotonal sense organ clusters. Curr. Biol. *7*, 166–175.

GenBank Accession Number

The GenBank accession number for the *amos* sequence reported in this paper is AF166113.