INVESTIGATING THE REGULATION AND FUNCTIONING OF *RNT-1* AND *BRO-1* IN *C. ELEGANS*

THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

CHARLES BRABIN

MAGDALEN COLLEGE

DEPARTMENT OF BIOCHEMISTRY

UNIVERSITY OF OXFORD

TRINITY TERM, 2011

Investigating the regulation and functioning of *rnt-1* and *bro-1* in *C. elegans*

Charles Brabin
Magdalen College

Submitted for the degree of D.Phil.
Trinity Term 2011

Abstract

The stem cell-like seam cells of the nematode, *Caenorhabditis elegans*, represent a tractable and powerful model for studying stem cell biology. *rnt-1,* the worm homologue of the mammalian RUNX family of transcription factors, together with the CBFβ homologue *bro-1*, is essential for the proliferation of the seam cells. RUNX genes and CBFβ are important regulators of stem cell development in mammals, and are associated with a variety of cancers. The worm seam cell model offers an opportunity to examine how these genes function in stem cell biology. The aim of this work was to shed light on the genetic network in which *bro-1* and *rnt-1* function, and to reveal the identity of regulators of these genes as well the downstream targets of the *bro-1/rnt-1* pathway.

Here, a number of genes that interact with *bro-1* and *rnt-1* have been identified. ELT-1, a GATA transcription factor, is shown to be a direct regulator of *bro-1*. Findings which show that the MEIS gene *unc-62* acts upstream of *bro-1/rnt-1* and regulates the symmetry of seam cell divisions are also presented. The seam cell marker, *scm::gfp*, is widely used in studies of the seam cells; here the results of an investigation into its identity and functional links are described. In addition, the mechanism underlying spatial regulation of *rnt-1* was examined; this led to the discovery of distinct tissue-specific enhancer modules within an intron of this gene. Finally, interactions between *pal-1* and *bro-1/rnt-1* are reported and described.

Together, these findings provide a framework for furthering our understanding of the mechanisms and genes associated with the functioning of *bro-1* and *rnt-1* in the worm.

Thesis supervisor: Dr. Alison Woollard

Acknowledgements

The time I have spent in Alison Woollard's laboratory has been rewarding and thoroughly enjoyable. Obviously, I am grateful to Alison for letting me work here in the first place, and for obtaining the MRC funding which supported me throughout, but my appreciation extends far beyond that. I have been incredibly lucky in having Alison as a supervisor – someone who is not just a good scientist but who takes a real interest in the progress and well-being of their students; there can be few supervisors who put aside so much time and invest so much effort for their students' benefit.

In addition, I have been fortunate to work in a lab and in a 'Genetics Unit' with a number of people whom I can call good friends, as well as colleagues. I am extremely grateful to them for making me so welcome in what is almost a family. Their company, friendship, help, advice, tolerance and humour over the course of my D.Phil. has made an enormous difference and meant a great deal to me. In particular, I must thank Pete, who has taught me much about many things and been a truly great companion. I will miss working with you.

TABLE OF CONTENTS

TABLE OF FIGURES

List of abbreviations

3-AT, 3-aminotriazole

3'UTR, 3' Untranslated Region

5'UTR, 5' Untranslated Region

ARF, ADP-ribosylation factor

ARL, Arf-like protein

ATP, adenosine triphosphate

CBFβ, Core Binding Factor

cDNA, complementary DNA

CFP, cyan fluorescent protein

CGC, *Caenorhabditis* Genetics Center

CNE, conserved non-coding element

DE-cadherin, *Drosophila* E-cadherin

DIC, differential interference contrast

DNA-BD, DNA binding domain

DTC, distal tip cell

EMSA, Electrophoretic Mobility Shift Assay

ETO, eight-twenty-one (a gene)

exd, extradenticle

GAP, GTPase Activating Protein

GEF, guanine exchange factor

GFP, green fluorescent protein

GSCs, germ line stem cells

HD, homeodomain

HM domain, Hth/Meis domain

hth, homothorax

IPTG, Isopropyl β-D-1-thiogalactopyranoside

kb, kilobase(s)

Mab, Male ABnormal

mb, megabase(s)

NGM, nematode growth medium

ORF, open reading frame

Pbx, pre-B-cell leukemia homeobox

PCR, polymerase chain reaction

PEBP2, polyomavirus enhancer binding protein 2

PNK, polynucleotide kinase

qRT-PCR, quantitative real time PCR

RNAi, RNA-interference

SAR, Secretion-Associated and Ras-related protein

SCM, seam cell marker

TALE, three amino loop extension

TCF/LEF, T-cell factor/lymphocyte enhancer factor

VPCs, vulval precursor cells

WT, wild type

YFP, yellow fluorescent protein

CHAPTER 1

General Introduction

Developmental Biology

One of the most fascinating questions in biology is how multicellular life forms, with their breath-taking complexity, develop from a single fertilised egg. The field of developmental biology has arisen in response to this question, with the aim of understanding the factors and mechanisms involved in this process. Over the centuries, the specific questions asked have changed with the prevailing scientific interests and fashions, together with the available technology and latest innovations, yet throughout this time the study of developmental biology has involved the study of three broad processes.

Firstly, for organisms to grow, whether in size or just in terms of the number of cells they contain, they must undergo cell proliferation; their constituent cells must divide, giving rise to daughter cells which go on to serve their own roles in the developmental process.

Secondly, multicellular life forms are comprised of different cell types; each cell type has its own specific functions and thus specific features and adaptations for fulfilling these functions. During development, these specialised cell types must arise from the egg through the process of differentiation. Just as differentiation leads to the acquisition of specialised features and functions, it also leads to loss of proliferative ability and exit from the cell division cycle. Thus, proliferation and differentiation are essentially opposite and mutually exclusive processes. The balance between these opposing processes is absolutely critical for development.

Thirdly, the cells which result from the processes of proliferation and differentiation must arrange themselves into tissues and, in all but the simplest animals, into organs through the

process of morphogenesis. This process distinguishes the eumetazoa from the simpler metazoan groups; placozoans and sponges (though the identity, nature and homology of sponge ‘tissues’ remains controversial (Ereskovsky and Dondua 2006)). In eumetazoans, the formation of tissues is not only dependent on proliferation and differentiation but also involves an additional level of complexity, requiring adhesion and communication between the constituent cells, as well as cell death (Grosberg and Strathmann 2007); only through the integration of these processes can tissues and organs be sculpted during development.

These three aspects of developmental biology can be investigated and analysed at different levels. Traditionally, owing to technological limitations, it has been at the anatomical level that studies have been performed. Subsequently though, it became increasingly feasible to answer questions about development at the cellular level and, more recently still, at the level of genes and molecules; it is now possible to trace the basis of development to the genes which control proliferation, differentiation and morphogenesis. Identifying and elucidating the roles of the genes involved, and putting them into the contexts of networks, will prove crucial for further advancing the field of developmental biology.

Overall, developmental biology has become a synthesis of anatomy, genetics, molecular biology and biochemistry. When properly integrated, these components complement each other and open the way for a fuller and ever more comprehensive understanding of the complex and intricate mechanisms underlying the transition from a zygote to a fully developed organism.

RUNX genes

Introduction

The RUNX genes make up a family of evolutionarily conserved transcription factors which play central roles in regulating both the proliferation and differentiation of cells. They are defined by the presence of the 'Runt-box' DNA binding domain, a 128-amino acid motif which facilitates protein-DNA interactions, allowing RUNX genes to act as transcription factors by binding to the promoters of target genes. Specifically, it is the consensus sequence R/TACCRCA which is necessary – but not sufficient – for recognition by and binding of RUNX family transcription factors (Kamachi, Ogawa et al. 1990). The Runt domain also mediates protein-protein interactions, and is necessary and sufficient for RUNX genes to heterodimerise with their binding partner CBFβ, which both increases the DNA binding affinity of RUNX proteins and protects them from proteosomal degradation (Adya, Castilla et al. 2000).

RUNX genes are a metazoan synapomorphy

The first RUNX family member to be described was *runt*, in *Drosophila melanogaster* (Gergen and Butler 1988) followed by numerous and re-iterative discoveries in other systems. In the fly, *runt* has roles in embryonic pattern formation, sex determination and neural development and, since its discovery, three more *Drosophila* RUNX genes have been identified. The involvement of RUNX genes in differentiation became apparent with the identification of polyomavirus enhancer binding protein 2 (PEBP2) in a screen for activators of polyomavirus in differentiated – but not undifferentiated – mouse embryonic cells (Kamachi, Ogawa et al. 1990). Purification and characterisation of PEBP2 revealed the presence of two subunits; PEBP2α and PEBP2β. PEBP2α showed high levels of homology with *Drosophila runt* whilst PEBP2β was identified as Core Binding Factor (CBFβ). Since

then RUNX/CBFβ have been identified in a wide phylogenetic range of organisms, including sponges (Sullivan, Sher et al. 2008) and placozoans (Srivastava, Begovic et al. 2008) both of which reside at the base of the metazoan clade. Indeed, despite extensive searches, RUNX genes have only been found within the metazoa (Rennert, Coffman et al. 2003; Sullivan, Sher et al. 2008) a fact which perhaps reflects their importance in the evolution of complex metazoan multicellular developmental processes (Robertson, Larroux et al. 2009).

RUNX genes have undergone evolutionary expansion and divergence

During evolution, the RUNX family has expanded through gene duplications and, though represented by a single copy in many basal groups, often has several paralogues in higher organisms. In mammals, for example, three RUNX genes are present, owing to previous rounds of duplication (and subsequent loss) within the vertebrate lineage (Strippoli, D'Addabbo et al. 2002): *RUNX1*, *RUNX2* and *RUNX3*. All vertebrate representatives of the RUNX family share, in addition to the Runt domain, the use of two promoters from which they are expressed (Blyth, Cameron et al. 2005), nuclear localisation signals which allow RUNX proteins to attach to the nuclear matrix (Zeng, van Wijnen et al. 1997), and the VWRPY motif (Levanon, Negreanu et al. 1994), which is present in all RUNT proteins and plays a crucial role in RUNX interactions with co-repressors (Aronson, Fisher et al. 1997).

Nevertheless, the structural conservation of RUNX genes should not obscure the fact that the process of family expansion has been accompanied by significant functional divergence, as evidenced by the differing roles of the mammalian RUNX genes; whilst *RUNX1* is required for definitive haematopoiesis (Okuda, van Deursen et al. 1996; Wang, Stacy et al. 1996; Chen, Yokomizo et al. 2009), *RUNX2* plays a critical role in bone formation (Komori, Yagi et al. 1997; Otto, Thornell et al. 1997; Hassan, Gordon et al. 2010) and *RUNX3* is

necessary for neurogenesis (Inoue, Ozaki et al. 2002; Levanon, Bettoun et al. 2002) as well as thymogenesis (Taniuchi, Osato et al. 2002) and gut development (Li, Ito et al. 2002).

RUNX genes and health

The important role of RUNX genes in normal development is highlighted by the number of cancers which are associated with mutations and chromosomal rearrangements involving members of the RUNX family and CBFβ. Acute myeloid leukaemia, for example, is commonly caused by a chromosomal translocation in which the *Runx1* runt domain is fused to *ETO*, resulting in a fusion protein which interferes with the normal functioning of Runx1 and thereby causing constitutive repression of target genes (Meyers, Lenny et al. 1995; Lutterbach, Westendorf et al. 1998; Wildonger and Mann 2005). Whilst *Runx1* appears to act as a tumour suppressor in this context, there are other examples of RUNX involvement in cancer where these genes play oncogenic roles. For example, the oncogenic potential of *Runx2* has been demonstrated by the finding that it is a frequent target site for viral insertions in mouse systems, with the induced insertions causing over-expression of the gene. *Runx1* and *Runx3* have also been found to be targets for viral insertion in lymphomas, with over-expression also resulting from the disruption (Blyth, Cameron et al. 2005). Furthermore, that over-expression of RUNX genes can be the cause of certain human cancers is supported by strong evidence, derived from cases of acute B lymphoblastic leukaemia and myeloid leukaemia in which genomic regions encompassing RUNX1 have undergone amplification (Blyth, Cameron et al. 2005).

Context is critical for RUNX gene function

As well as highlighting the importance of RUNX genes in development – and of understanding their functioning and regulation – the variety of cancers associated with members of this gene family and, in particular, the fact that RUNX genes have the potential to act as both tumour suppressors and oncogenes, emphasises the importance of context in their developmental roles. This context dependency for RUNX function is further highlighted by the fact that these proteins can act as both activators and repressors of transcription, depending on the accessibility of target promoters and the combinations of proteins with which they work in the cell at any given time (Coffman 2003).

RUNX genes are key regulators of proliferation and differentiation

The expansion of RUNX genes, which has been coupled with their adoption of numerous different developmental roles in a wide range of tissues and organisms, should not obscure the unifying theme which runs through the family; RUNX genes are central to the developmental processes of cellular proliferation and differentiation. For example, the sea urchin Runx gene, *SpRUNT,* is, at different stages in the development of the sea urchin embryo, required for both proliferation in actively cycling cells (Robertson, Dickey et al. 2002) and differentiation of aboral ectoderm (Coffman, Kirchhamer et al. 1997). In mammals too, dual roles for RUNX genes in proliferation and differentiation have been characterised; *RUNX2* appears to be associated with cell cycle exit and differentiation (Rossi, MacLean et al. 2002) whilst the role of *RUNX1* in driving the cell cycle into S phase has also been firmly established (Strom, Nip et al. 2000). Given the opposing nature of these two developmental processes, it is at first perhaps surprising that RUNX genes can play central roles in both. This paradox can be rationalised, however, by the role context plays in determining the functioning of RUNX genes in development. Indeed, the importance of context in RUNX

regulation of target genes – and their consequent ability to function in a wide range of different ways – places them in a powerful position as key regulators of development.

C. elegans as model system in developmental biology

The nematode worm, *C. elegans*, is now a well-established model organism for the study of development, genetics and genomics as well as a seemingly ever-expanding field of related areas of research.

Naturally a soil-dwelling organism, *C. elegans* (specifically, the reference strain, N2 (Bristol)) was pioneered as a model organism in order to address the question of how genetic information could specify the development of higher, complex organisms (Brenner 1974); the worm's small size (adults reach a length of approximately 1mm), relatively simple nervous system and anatomy were key advantages. Figure 1.1 shows the basic anatomy of the worm. As a tool for studying genetics, it was also well-suited. Populations are comprised mainly of self-fertile hermaphrodites, with the occasional spontaneous male, capable of mating with the hermaphrodites.

The *C. elegans* lifecycle

The lifecycle of the worm is rapid; after laying, eggs hatch within approximately ten hours, giving rise to larvae which undergo four moults during the transition to adulthood, which takes three-and-a-half days at 20°C (Byerly, Cassada et al. 1976). Descriptions of the developmental stage of larvae are given relative to these larval moults; after hatching, larvae are designated L1, followed by L2, L3 and L4, after which adulthood is reached.

During their two to three week-long life, each hermaphrodite is capable of laying around 300 progeny (Byerly, Cassada et al. 1976). Consequently, genetic crosses can be performed

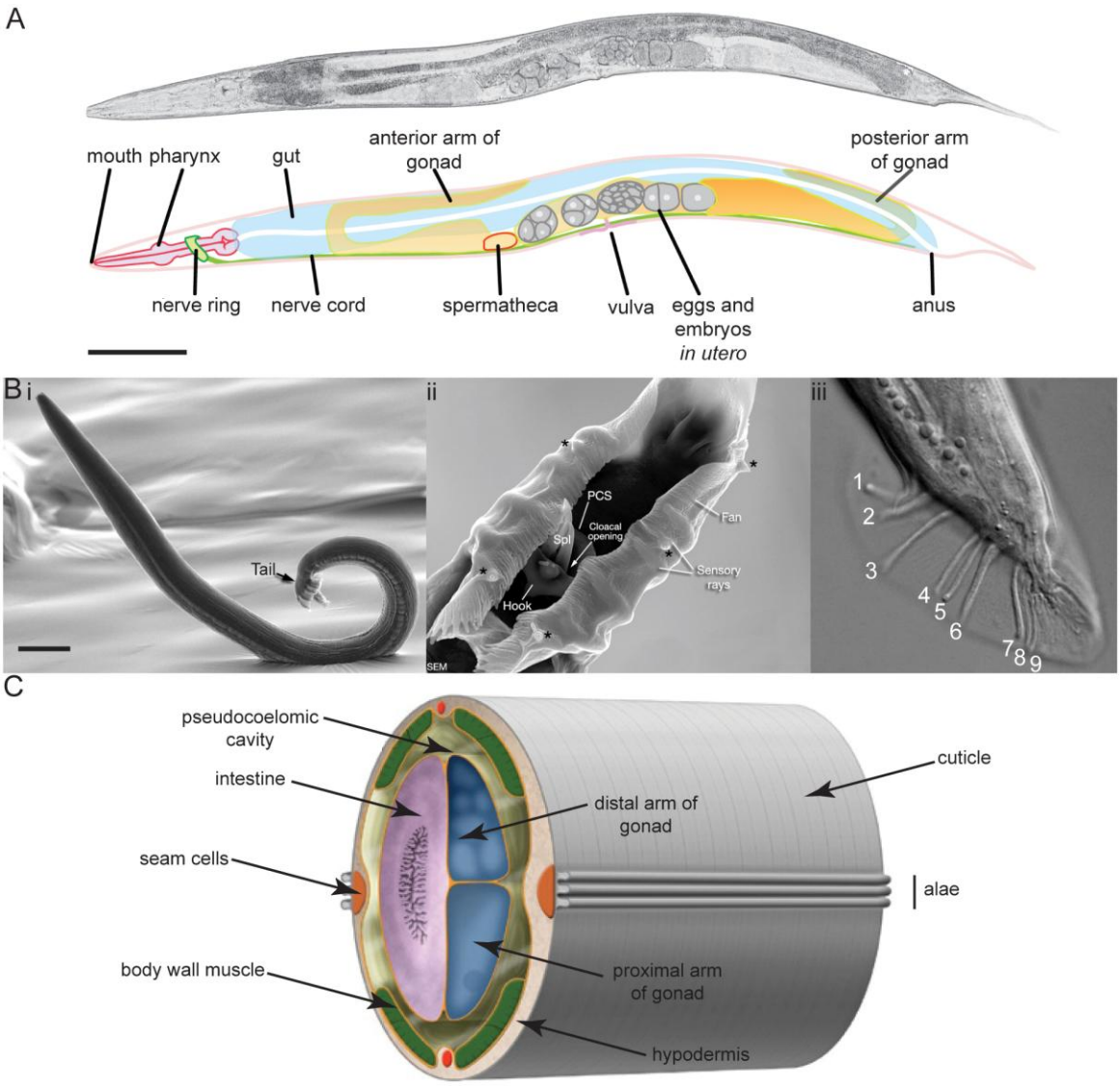


Figure 1.1 General Anatomy of *C. elegans*
A. Nomarski photomicrograph (top) and schematic representation (below) of the worm, *C. elegans*. Scale bar = 0.1mm. B. Males are easily distinguished from hermaphrodites by their elaborate tails, used for mating. Bi. SEM of a male *C. elegans*, with the tail clearly visible at the posterior. Scale bar = 0.1mm. Bii. SEM of ventral male tail detail. Asterisks mark ray openings; PCS, post-cloacal sensilla; Spl, spicules. Biii. Nomarski photomicrograph of one side of a male tail, showing the 9 rays. Scale bar in Bii and Biii = 20µm. C. Cross-section of an adult hermaphrodite, showing the arrangement of the body layers. A. Modified from Blaxter, 2011.Bi, Bii and C adapted from Altun, Z. F. and D. H. Hall, 2009.

relatively quickly and efficiently and, given the small size of the worm, experiments can easily be performed on a large scale, involving thousands or even millions of individuals.

C. elegans is naturally a 'pioneer' organism; such life forms have evolved a strategy of being able to take advantage of favourable conditions by reproducing rapidly, whilst food is abundant and competition and predation relatively low. The 'pioneer' evolutionary life history, therefore, depends on being present in such favourable environments and, in general, two alternative solutions to this problem have evolved. Pioneers either tend to have highly efficient spatial dispersal mechanisms, enabling them to be represented in such environments, or else they have evolved long-term survival strategies, usually involving adaptations to remain viable in the face of adverse – often severely adverse – conditions; when conditions improve, such organisms can simply recover from their state of quiescence and take advantage of the environment. The nematode *C. elegans* falls into the latter category, owing to modifications of its lifecycle which enable it to withstand prolonged periods of starvation and desiccation.

In addition to the basic lifecycle of the worm (Figure 1.2), there are two alternative developmental routes, either of which is taken in response to food deprivation. In situations where eggs hatch in the absence of food, the resulting L1 larvae arrest soon afterwards, moving into a state referred to as L1 diapause in which they can remain viable for up to three weeks (Baugh and Sternberg 2006). Alternatively, starvation or perception of intense competition (the latter based on high population density) during L2 can trigger a separate developmental pathway, in which larvae enter a different 'quiescent' state and become 'dauer-larvae' (Cassada and Russell 1975). 'Dauers' are extremely well-adapted, both physiologically and behaviourally, to survive in unfavourable conditions. Their mouths become plugged, they possess a modified cuticle which enhances water-retention, and they cease pharyngeal pumping and general locomotion almost completely (Golden and Riddle 1984). Indeed, so effective are these modifications that dauers can remain viable in this form for months, retaining the ability

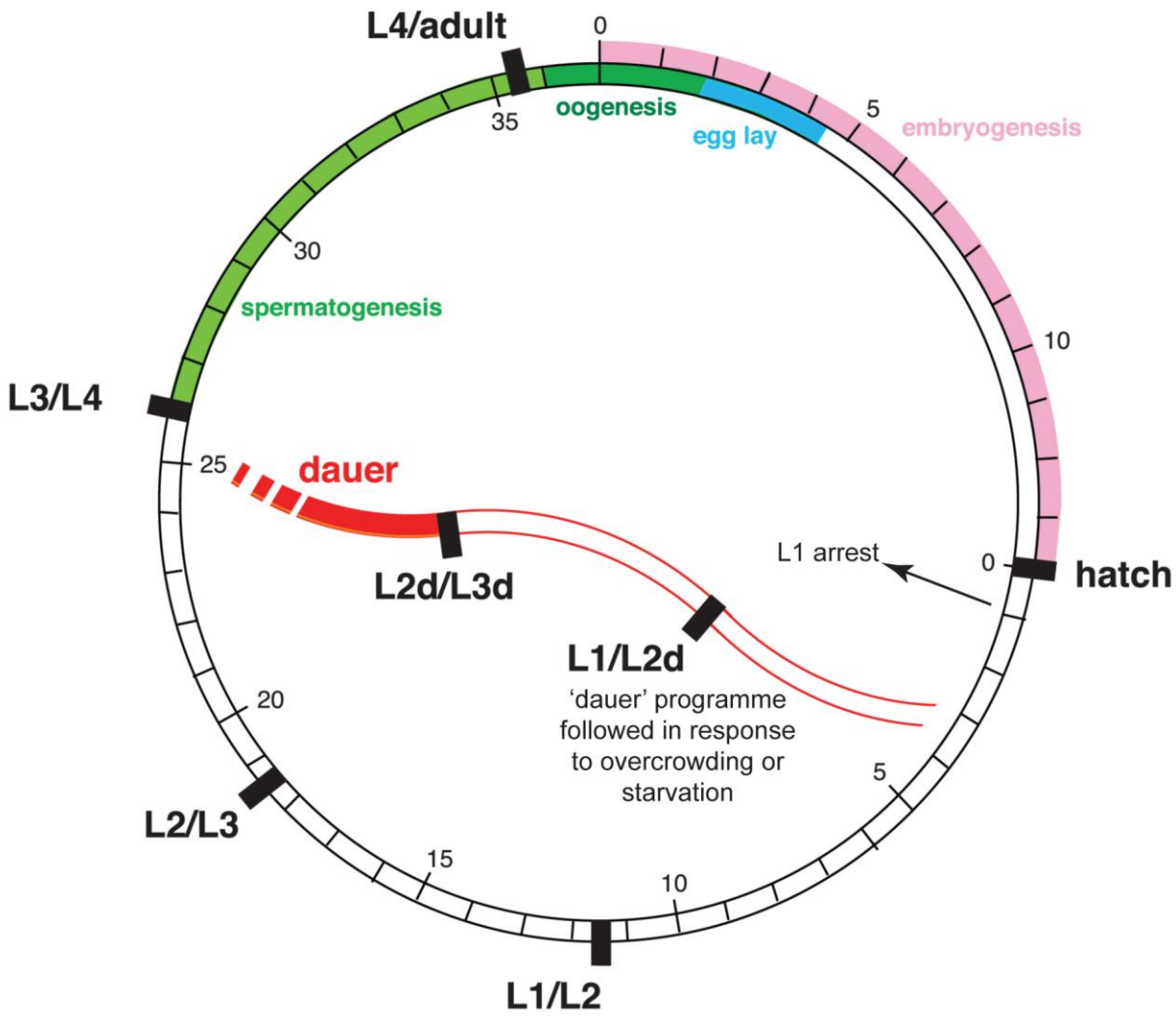


Figure 1.2 The *C. elegans* lifecycle
Schematic showing the lifecycle of the worm, with the time taken for each developmental stage at 25°C shown in hours; numbers on the outside of the circle denote hours from laying, whilst numbers on the inside are hours from hatching. Thick black lines indicate moults. Alternative lifecycle options are also outlined (L1 diapuase and dauer pathways). Modified from Blaxter, 2009.

to subsequently progress through the later larval stages and become fertile adults (Golden and Riddle 1984).

Studying development and genetics in *C. elegans*

The advantages of the worm were augmented by the mapping of the invariant lineage of the transparent *C. elegans*, first for larval development (Sulston and Horvitz 1977) and, later, for embryonic development (Sulston, Schierenberg et al. 1983). The ability to perform lineage analysis, coupled with the availability of a comprehensive reference for normal development, meant that any developmental abnormality could be traced back to its origin at single cell resolution.

The availability of increasing numbers of genetic mutants, together with the ability of worms to survive and remain recoverable after being frozen at -80°C (Brenner 1974), further enhanced the utility of this model organism. Similarly, the finding that genes can be selectively and efficiently switched off through the use of RNA interference (RNAi) (Guo and Kemphues 1995; Fire, Xu et al. 1998), whether administered by injection (Fire, Xu et al. 1998), feeding (Timmons and Fire 1998) or soaking (Tabara, Grishok et al. 1998), further expanded the number of genes which can be readily investigated. The construction and availability of libraries of RNAi clones (Fraser, Kamath et al. 2000; Rual, Ceron et al. 2004), covering the majority of genes in the *C. elegans* genome, augmented the potential for high-throughput screening, facilitating the discovery of gene functions and interactions.

The *C. elegans* genome

The sequencing of the genome (The *C. elegans* Sequencing Consortium, 1998), at the time the first complete genome of a multicellular organism – proved invaluable and further accelerated the rate of growth of understanding of the development and genetics of the

worm, offering the prospect of the discovery of all the genes required for the development and functioning of an animal, and, just as importantly, of the additional genetic material which is also of significance. With over 100Mb of sequence, encoding 20,514 protein-coding genes (Wormbase Release WS229), the sequenced genome represents an invaluable tool for unpicking the complex, multi-layered network of interactions which underlie development.

C. elegans and comparative genomics

The utility of the *C. elegans* genome sequence data has been augmented significantly by the publication of the genomes of other Caenorhabditid (Stein, Bao et al. 2003; Dieterich, Clifton et al. 2008; Blaxter 2011) and nematode species. Comparative genomics is a hugely powerful approach to recognising and analysing functional regions of the genome, whether at the level of operons, genes or transcription factor binding sites. The approach is based on the premise that, over evolutionary time, the genome changes. The speed of this change will depend upon the selection pressures acting on the region of sequence in question. Regions which are of functional significance should be under relatively high selection pressure compared to regions of no functional significance. Thus, by aligning and comparing one or more genomes from different organisms and searching for regions which exhibit high levels of conservation, it is possible to glean clues as to the identity and location of a range of 'functional units' within the genome and thereby probe the developmental genetics of the worm.

In addition, the availability of complete genome sequences for several relatively closely related nematode species affords an opportunity to examine the working of evolution and how evolution acts upon development. For example, this approach has been used to assess the occurrence and rate of changes in cis-regulatory elements which regulate the expression of genes, and to determine the importance of such changes for the evolution of developmental mechanisms (Barrière, Gordon et al. 2011; He, Eichel et al. 2011).

C. elegans as serves as a representation of higher groups

A crucial aspect of any model organism is its relevance to other, more complex systems which it is supposed to represent. That around 38% of human disease genes have an orthologue in *C. elegans* (Shaye and Greenwald 2011) suggests that this is indeed the case. Moreover, there is often conservation at the level of DNA and protein between such homologues, providing valuable functional information on regions of sequence, protein domains and the overall function of the genes in question. Such an approach has led to numerous studies of human disease genes in *C. elegans* (reviewed in Markaki and Tavernarakis 2010).

C. elegans anatomy

Caenorhabditis species, including *C. elegans*, are, like all roundworms, unsegmented pseudocoelomates. Individuals comprise around 1000 somatic nuclei and are approximately 1mm long and 50µm wide and, in the case of hermaphrodites, tapered at both ends (Figure 1.1A). At the most basic level, they comprise two tubes – one within the other – separated by a cavity (Figure 1.1C). On the outer side of the cavity lie the muscles, hypodermis, excretory and neuronal systems and, at the interface with the outside world, a collagenous cuticle. The inner tube consists of the alimentary system, with an anterior pharynx and grinder, for ingesting and lysing bacterial cells, followed by an intestine which runs almost the entire remaining length of the worm. Also present in the inner tube is the gonad. The cavity between the inner and outer tissue layers of the worm is filled with pseudocoelomic fluid which, through hydrostatic pressure, serves to maintain the shape and relative rigidity of the worm (Altun and Hall 2009).

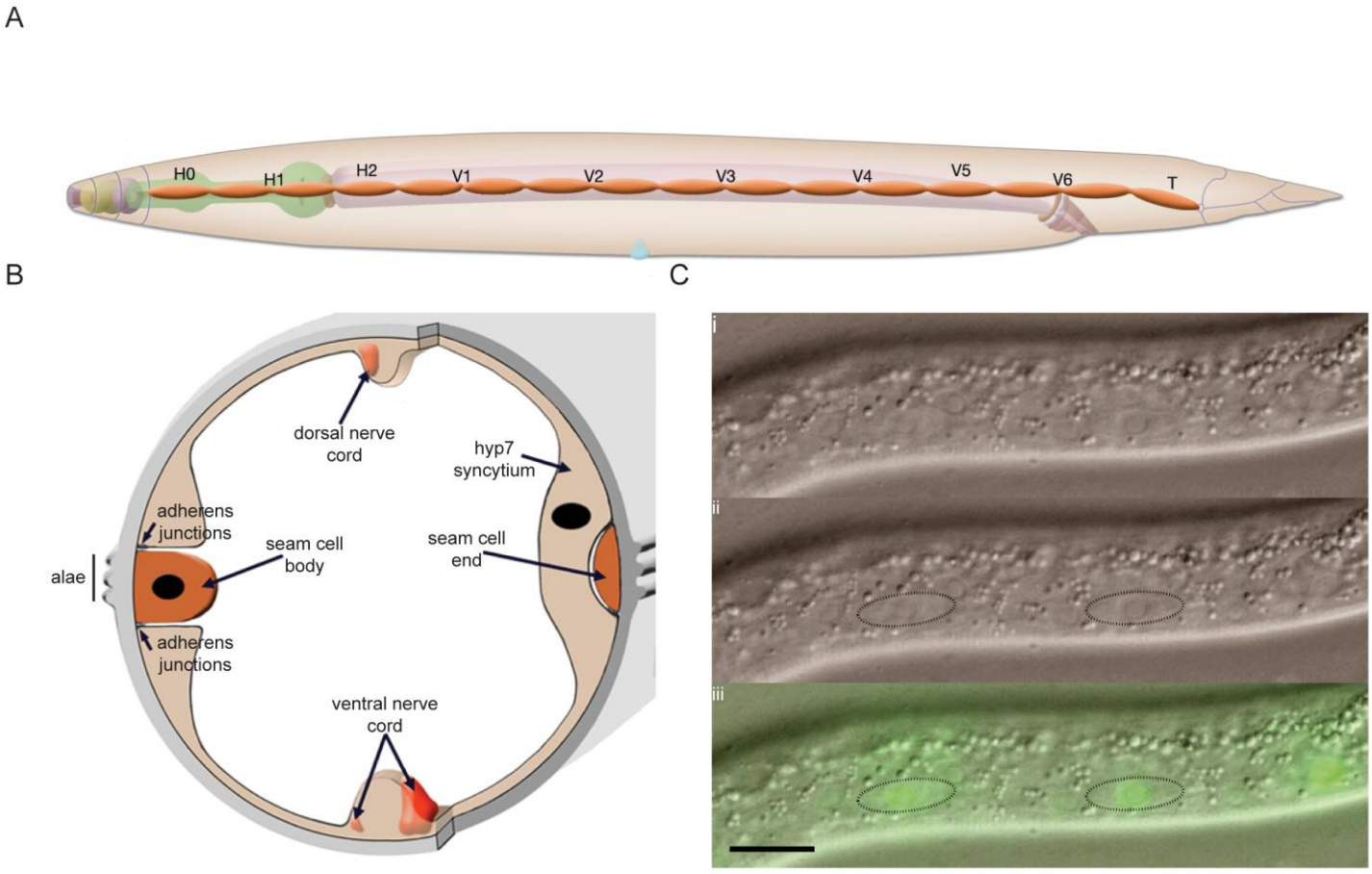


Figure 1.3 Seam cell anatomy
A. Lateral schematic view of an L4 hermaphrodite with 16 seam cells (orange). The letters above the cells denote the lineages from which the respective seam cells are derived. B. Cross section of larva, showing the position of seam cells in the worm relative to other tissues and structures. C. Photomicrographs showing the charateristic 'eye-shape' of seam cells between divisions. The seam cells are clearly visible using Nomarski optics (i); the outlines of two seam cells are marked by the dotted line (ii) and the nuclei by the seam cell marker, SCM::GFP (iii). Scale bar = 25μm. A and B modified from Altun, Z. F. and D. H. Hall, 2009.

Although at the coarsest level of organisation hermaphrodites and males appear similar, almost all tissues show some degree of sex-specific specialisation. At the gross morphological level, males are slightly shorter and thinner than hermaphrodites, but it is the tail which is the most distinctive feature of males; contrasting with the tapered tails of hermaphrodites, L4 and adult males possess an elaborate fan structure (see Figure 1.1B), the initial development of which is clearly visible during L3. The tail comprises: 18 rays which contain sensory dendrites and which, with two exceptions, possess an opening at their tip; the 'hook', an additional sensory structure; post-cloacal sensilla, which also contain neurones, and spicules, which function together with the other sensory tail structures to locate and hold on to the vulva and transfer sperm during mating (Altun and Hall 2009).

The seam lineage in *C. elegans* – a model for stem cell biology

The anatomy of the *C. elegans* seam

Beneath the tough cuticle of the worm lies the epithelium, comprising both hypodermis and specialised epithelial cells. The seam cells, also referred to as lateral hypodermal cells, fall into the second category and are present as ten cells on either side of the worm at hatching; H0 is the anterior-most seam cell, and the only lateral hypodermal cell which does not divide. Posterior to H0 is H1, followed by H2 and then by the V lineage seam cells, V1-6. The T lineage represents the posterior-most end of the seam, located in the tail (Sulston and Horvitz 1977; Sulston, Albertson et al. 1980; Sulston, Schierenberg et al. 1983).

The seam cells are embedded in the hyp7 syncytium which runs along the body of the worm, starting just posterior to the head and ending anterior to the tail (Figure 1.3A and B). Dorsally, ventrally and internally, the seam cells are surrounded by hypodermis. Throughout larval development, the seam cells continue to divide, both increasing the number of seam

cells present and contributing hypodermal daughters to the surrounding syncytium. The seam also gives rise to neurones and glial cells and, in males, the lineages of the posterior seam cells V5, V6 and T produce the sensory rays of the male tail (Sulston and Horvitz 1977). In addition to contributing these cellular lineages, the seam also synthesises the alae, longitudinal ridges in the cuticle along the sides of worms during L1, adulthood and dauer (Cox, Staprans et al. 1981). Indeed, differentiation of the seam lineage is required for the formation of alae (George, Simokat et al. 1998; Petalcorin, Oka et al. 1999; Antebi, Yeh et al. 2000; Koh and Rothman 2001; Frand, Russel et al. 2005).

When the worm reaches adulthood, by which stage 16 seam cells are present on either side of the worm, seam divisions cease; the cells undergo homotypic fusion, joining together to form a syncytium which nevertheless remains distinct from the surrounding hypodermis (Podbilewicz and White 1994).

Stem cells in developmental biology

As organisms develop, they undergo both cell division and proliferation. The regulation of these two processes, and of the balance between them, is of the utmost importance. It is for this reason that stem cells, which have the ability to both self-renew and give rise to specialised, differentiated cell types, play such important roles in development, aging and tissue repair and regeneration. Moreover, it is this property of stem cells which makes them such potent triggers of developmental disorders, such as cancer. This important ability is one of the defining features of stem cells. In addition, stem cells reside within a niche, a microenvironment which maintains them in the proliferative, undifferentiated state, providing signals which both promote mitosis and prevent differentiation (Ohlstein, Kai et al. 2004).

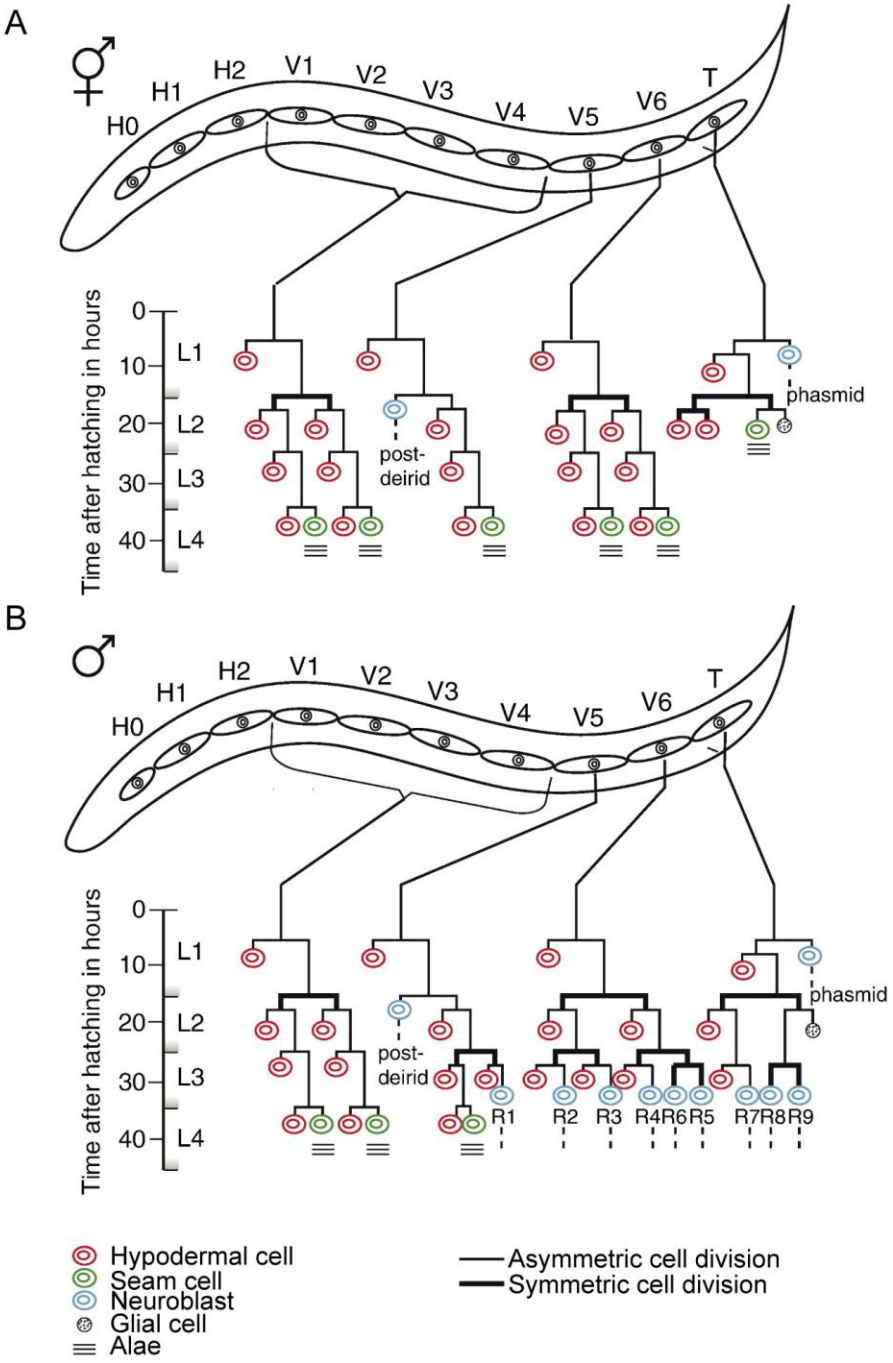


Figure 1.4 Seam cell divisions in *C. elegans*
Postembryonic V and T lineage seam cell divisions in hermaphrodite (A) and male (B) worms. Asymmetric seam cell divisions occur during each larval stage, with an additional symmetric division taking place during L2. Following asymmetric divisions, hypodermal daughters of seam divisions move out of the line of the seam and into the hyp7 syncytium. H1 and H2 divisions not shown. Modified from Nimmo and Woollard, 2008.

The importance of stem cell divisions

When a stem cell divides, the fates of the daughter cells are intimately linked with the mode of division. Traditionally, the asymmetry of stem cell divisions has been seen as key to the ability of stem cells to both self-renew and produce cells that differentiate; one daughter of the division retains the stem fate and the ability to divide again, while the other daughter loses its proliferative potential, instead exiting the cell cycle and adopting any one of a vast range of differentiated cell fates (Doe and Bowerman 2001; Clevers 2005; Yamashita, Fuller et al. 2005; Morrison and Spradling 2008). However, more recently it has been recognised that stem cell divisions are more flexible and adaptable than this model suggests, able to adjust the balance between stem cell proliferation and differentiation to meet the needs of the organism, either in the course of normal development or to redress an imbalance caused by perturbation of the system (Morrison and Kimble 2006). Thus, whilst stem cells can undergo asymmetric divisions which give rise to one daughter cell of each fate, they can also divide symmetrically to produce two stem cell daughters. This mode of division allows for expansion of the stem cell pool, whether at early stages of development, or after, for example, injury which generates the need for growth of the stem cell population (Morrison and Kimble 2006). This flexibility, however, comes with the need for an additional level of regulation; the mode of division is another important factor in the regulation of stem cells, and will have a dramatic impact on the output of stem cell divisions. Too many symmetric divisions will lead to over-expansion of the stem cell pool and a shortage of differentiated cell types, whilst purely asymmetric divisions (or divisions which are symmetrised towards the non-stem fate) will have the opposite outcome, resulting in the ultimately unsustainable depletion of cells with the stem fate.

Stem cells in *C. elegans*

The fact that the worm possesses stem cells, taken together with its long list of other advantages as a model organism, has made it an attractive tool for the study of stem cell biology.

The germ line represents the only true population of stem cells in the worm, retaining cells which continue to proliferate throughout the life of the worm (Hirsh, Oppenheim et al. 1976). Situated in the gonad, the germ nuclei share a common syncytium but are each partially surrounded by cell membranes and are arranged along a proximo-distal axis of development. At the distal end of the gonad, germ ‘cells’ undergo mitosis, proliferating and pushing their neighbours proximally, towards the opposite end of the U-shaped gonad. As cells move away from the proximal region, their mitotic divisions stop and they instead divide meiotically, before undergoing complete cellularisation and becoming oocytes ready for fertilisation.

Clearly, these cells fit the stem cell criteria of being capable of both proliferation and production of differentiated cell types. Furthermore, for the germ line, it has also been possible, owing in part to the simple anatomy of the worm, to demonstrate the existence of a niche which promotes the mitotic (stem) fate and prevents meiosis. The distal tip cell (DTC), so-called because of its position at the distal end of the gonad creates a microenvironment around the distal-most germ line stem cells (GSCs). Though only a single cell, the DTC exerts its influence on a population of cells through the utilisation of fine cytoplasmic processes which pass through the densely packed GSCs, stretching over 10 cell diameters from the cell body (Kimble 1981).

Seam cells as stem cells

The germ line has been an invaluable model for augmenting understanding of the stem cell niche, particularly with respect to the signalling pathways involved. However, *C. elegans* provides another opportunity to examine stem cell biology, and in particular the mechanisms underlying stem cell divisions.

The seam lineage represents an attractive and tractable model of the stem cell mode of division. During larval development, the seam cells undergo a series of invariant divisions, both symmetric and asymmetric, and thus afford an opportunity to investigate not just the factors which control the balance between proliferation versus differentiation but also the regulation of stem cell divisions.

Using differential interference contrast (DIC) microscopy, the seam cells are identifiable by their distinctive 'eye' shape (Figure 1.3C). Just prior to division, the cells become rounder, losing contact with their neighbours in the seam. Subsequent to the division, the cells elongate and return to their 'eye' shape, re-establishing contact with adjacent seam cells (Austin and Kenyon 1994; Podbilewicz and White 1994).

With the exception of H0, which does not divide at all, the seam cells divide in each of the four larval stages (Figure 1.4). The pattern of divisions is different amongst different lineages of the seam. For example, the T lineage undergoes a very different division pattern to the H1, H2 and V lineages. Similarly, V5 which, like the T cell, eventually gives rise to neurones and glial cells, proceeds through a series of divisions which differ markedly from the other V lineage divisions from L2 onwards.

In the context of the choice between symmetrical and asymmetrical divisions, it is perhaps the V cells 1-4 and 6 which are most informative. These cells first divide in L1, where they undergo an asymmetrical division, producing one anterior daughter cell which is destined to join the hyp7 syncytium and one posterior daughter cell which remains in the seam.

This anterior-posterior fate choice is reiterated in all subsequent asymmetric divisions of these cells and their descendants. Soon after the division, the anterior daughter becomes much rounder and, whilst the posterior daughter stretches out to re-join adjacent cells in the seam line, becomes relatively much smaller. Simultaneously, the cell loses expression of the Seam Cell Marker, *scm::gfp* (Koh and Rothman 2001), and the cell membrane breaks down, allowing the cell to move either dorsally or ventrally, depending on its lineage, out of the seam and into the hyp7 syncytium.

After dividing asymmetrically in L1, the seam cells undergo a symmetrical division at the beginning of L2, followed by a further asymmetrical division. The symmetrical division facilitates expansion of the seam, increasing the number of cells with the stem-like fate at the cost of hypodermal daughters during that division (overall though, this will result in the seam being able to produce a greater number of hypodermal daughters at each round of division). The seam daughters of the L2 divisions divide asymmetrically again at the beginning of the L3 and L4 stages (Sulston and Horvitz 1977).

The robust distribution of seam cell divisions throughout the development of the worm demonstrates the precise regulation of the cell cycle with respect to developmental stage. The period between divisions is up to ten hours, yet the seam divisions reliably happen within minutes of their scheduled occurrence. Clearly, there is tight control over when the seam cells enter the cell cycle but the factors which regulate this aspect of development remain unknown in source, origin and mechanism (Kipreos 2005).

As with other cell divisions in *C. elegans*, the pattern of seam divisions is invariant. Thus, there is precise coordination between the regulation of division symmetry and the timing of these divisions relative to the development of the worm. Deviations from the normal pattern could have dramatic developmental consequences. Indeed, such aberrations can be seen in heterochronic mutants, where re-iteration of the L1 asymmetrical division, for example, occurs in mutants for the microRNA gene *lin-4*, resulting in reduced numbers of seam cells

in the worm (Ambros and Horvitz 1984). In contrast, worms in which the target of *lin-4*, *lin-28*, is not repressed reiterate the L2 symmetric division, resulting in not fewer but more seam cells (Moss, Lee et al. 1997; Seggerson, Tang et al. 2002).

Overall then, correct development therefore requires the integration of cell cycle decisions, division symmetry as well as developmental timing. Understanding the factors involved and how they interact has progressed significantly in recent years, but is far from complete.

It should be noted that, despite their clear relevance to stem cell biology, the seam cells are not considered to be true stem cells, like the GSCs, but rather 'stem cell-like'. Significantly, they do eventually terminally differentiate; at adulthood, fusion of the seam cells to form a continuous seam syncytium along each side of the worm is coupled with cessation of division (Sulston and Horvitz 1977; Singh and Sulston 1978) and the production of alae (Singh and Sulston 1978). Furthermore, the definition of stem cells usually includes reference to a niche. As yet though, no niche has been identified in the context of the seam. However, this is not to say that the concept is irrelevant for this tissue; the importance of cell-cell interactions for the normal development and proliferation of the seam has been well established (Austin and Kenyon 1994; Silhankova, Jindra et al. 2005) and recent observations are consistent with the hypothesis that the seam cells reside in a microenvironment, protected from differentiation-inducing signals and provided with proliferation-inducing cues (Brabin, Appleford et al. 2011). The precise nature, origin and identity of these influences remain to be determined.

The *C. elegans* RUNX and CBFβ homologues function in the seam cells

In *C. elegans*, the themes of RUNX genes and stem cell biology are united by *rnt-1*. As the only worm homologue of the RUNX family of transcription factors (Nam, Jin et al. 2002; Nimmo, Antebi et al. 2005; Kagoshima, Nimmo et al. 2007), the study of *rnt-1* does not suffer

from the problems of redundancy experienced in other model organisms, where multiple RUNX homologues are present. As in mammals and other groups, RNT-1 acts as a transcription factor together with the sole *C. elegans* CBFβ homologue, *bro-1*, which serves to increase the affinity and specificity of RNT-1 for DNA (Kagoshima, Nimmo et al. 2007).

Both *rnt-1* and *bro-1* are expressed in the seam cells and body wall muscle (Nimmo, Antebi et al. 2005). As yet, a role of these genes in muscle tissue has not been identified, but the genes' function in the seam is well established. *rnt-1*, together with its binding partner *bro-1*, is required for the correct proliferation of the seam cells during larval development. *rnt-1* mutants were identified based on their male tail phenotype, following a screen for mutants with defects in male tail development in which various Mab (Male ABnormal) mutants were isolated (Hodgkin 1983). *mab-2* was subsequently found to encode RNT-1 (Nimmo, Antebi et al. 2005).

Although initial searches for a CBFβ homologue – based on the known interaction between RUNX and CBFβ – were unfruitful (Adya, Castilla et al. 2000), a *C. elegans* homologue was later proposed (Lee, Ahnn et al. 2004) and subsequently confirmed (Kagoshima, Nimmo et al. 2007).

rnt-1 is required for seam cell proliferation in *C. elegans*

Lineage analysis revealed that the developmental basis of the Mab phenotype in *rnt-1* mutants stems from division failures in the V and T lineage seam cells, which contribute to the sensory rays of the tail (Nimmo, Antebi et al. 2005). As a result, rays are frequently missing in *rnt-1* mutants; whereas wild type males have nine rays on either side of the tail, *rnt-1* mutants on average have between four and five (Nimmo, Antebi et al. 2005). Ray numbers for *bro-1* mutants are similar (Kagoshima, Nimmo et al. 2007). Recognition of a role for *rnt-1* in the seam led to examination of seam cells in the rest of the body and the

observation that division failures occur throughout the V lineage. Compared to adult wild type worms, with 16 seam cells on either side of the body (or 18, in the case of males), *rnt-1* mutants have significantly fewer; hermaphrodites have around 13 seam cells per side, whilst in males the number is around 11 (Nimmo, Antebi et al. 2005). Similarly, *bro-1* mutants also have reduced numbers of seam cells; in hermaphrodites, an average of around 11 are present on each side of the adult worm. In *rnt-1; bro-1* double mutants, the number of seam cells is similar, with adults having an average of 11 on each side (Kagoshima, Nimmo et al. 2007).

Lineage analysis of the seam throughout development revealed that the reduced seam cell number in these mutants is a result of post-embryonic division defects. In some cases, division failures are observed. For example, V lineage seam cells may undergo the first, symmetrical L2 division normally but then not undergo the second scheduled division, instead waiting until the worm has passed through the L3 moult before dividing. On other occasions, the symmetry of divisions is perturbed. This can result in symmetric divisions being transformed into asymmetric divisions, thus resulting in one seam daughter and one hypodermal daughter rather than two seam daughters. Alternatively, asymmetric divisions can be symmetrised in favour of the non-stem, hypodermal fate (Nimmo, Antebi et al. 2005). Perhaps not surprisingly, given the similarity of *rnt-1* and *bro-1* phenotypes and the interdependence of these genes, similar division defects have been observed in *bro-1* mutants (Kagoshima, Nimmo et al. 2007). Thus, *rnt-1* and *bro-1* are necessary for normal proliferation of the seam.

rnt-1 and *bro-1* are rate-limiting for seam proliferation

rnt-1 and *bro-1* are not just necessary for seam proliferation but are rate-limiting factors. Over-expression of either *rnt-1* or *bro-1* results in an increase in the number of seam cells, whilst over-expression of both genes together yields an even more dramatic expansion of

seam numbers owing to a combination of symmetrised divisions (favouring the stem, rather than the hypodermal, fate) and extra divisions (Kagoshima, Nimmo et al. 2007).

As for the mechanism by which *rnt-1* and *bro-1* exert their effect on the seam, *cki*-1, a cyclin-dependent kinase inhibitor, has been identified as a likely target, potentially acting as a bridge between development and cell cycle control (Hong, Roy et al. 1998; Boxem and van den Heuvel 2001; Fukuyama, Gendreau et al. 2003). Based on a number of observations (Nimmo, Antebi et al. 2005; Kagoshima, Nimmo et al. 2007), *rnt*-1, together with *bro*-1, appears to repress *cki-1*, which in turn promotes arrest of the cell cycle (Hong, Roy et al. 1998; Fukuyama, Gendreau et al. 2003). Thus, *rnt-1* is a repressor of a repressor, thereby effectively serving as an activator of cell division. Here again then, RUNX genes are seen to be at the heart of the developmental decision between proliferation and differentiation and play a central role in stem cell biology.

Summary

In *C. elegans*, *rnt-1* and *bro-1* work to promote cell proliferation in the stem-like seam lineage, just as they are able to operate as oncogenes in other systems. This suggests functional conservation between these genes in the worm and in higher organisms and underscores the value of *C. elegans* as a model to further enhance our understanding of the roles and mechanisms of action of RUNX genes.

As yet, relatively little is known about the network in which *C. elegans rnt-1* and *bro-1* operate. Elucidating the pathways in which these genes work – identifying the genes they regulate and revealing the factors by which they themselves are regulated – should reveal important information on the regulation of proliferation, differentiation and stem cell divisions. This study aims to shed light on these issues and provide context for the functioning of RUNX genes in the worm and, ultimately, in developmental biology in its broadest sense.

Aims of this work

The overall aim of this study was to shed more light on the workings of the *rnt-1* pathway. This involves identifying genes which regulate *bro-1* and *rnt-1*, as well as targets which are regulated by the *bro-1/rnt-1* complex. Furthermore, the nature of interactions between *bro-1/rnt-1*, upstream regulators and downstream targets was of interest.

In Chapter 3, *elt-1* is identified as a direct regulator of *bro-1*, driving *elt-1* expression by binding to an intronic enhancer element. Subsequent analysis of *elt-1* revealed that it works both through *bro-1/rnt-1* and, separately, through the fusogen *eff-1*, to maintain the stem-like fate of seam cells and prevent their differentiation.

Chapter 4 details the identification of *unc-62*, a Hox gene co-factor, as an additional upstream regulator of *bro-1/rnt-1*, though in this case the interaction is likely not direct. *unc-62* appears to play a critical role in regulating the symmetry of seam divisions, thus determining whether or not the daughter cells of divisions express *bro-1* and *rnt-1* and retain the stem-like fate. In addition, *unc-62* is found to act redundantly with *bro-1* and *rnt-1* in embryonic seam development, demonstrating previously unknown roles for *bro-1/rnt-1* function before the larval stage.

The widely used *scm::gfp* seam cell marker, utilised throughout this and many other studies, is examined in Chapter 5. The marker is found to reflect expression of a GTPase, *arf-3*, leading to the finding that this gene also plays a role in seam cell development. The links between this GTPase, intracellular trafficking components and known regulators of stem cell divisions are examined.

Chapter 6 focuses on transcriptional regulation of *rnt-1.* Findings concerning the modular structure of seam enhancers within a long intron of *rnt-1* are detailed, together with the utility of these results for finding new upstream regulators of *rnt-1* and the wider implications for transcriptional regulation.

Finally, in Chapter 7 an additional novel *rnt-1* interactor is described. The caudal homologue, *pal-1*, a homeobox gene, appears to function both with and independently of *rnt-1* in determining seam development. Multiple stages of development are affected, apparently in different ways. The implications for *bro-1/rnt-1* function, and for the structure of developmental networks, are considered.

Together, these results sections link a number of genes with the *bro-1/rnt-1* pathway. These interactions, some of which are now well defined, whilst others require further characterisation, represent a framework for significantly broadening our understanding of how *bro-1* and *rnt-1* function in the development of *C. elegans*.

CHAPTER 2

Materials and Methods

Strains and maintenance of worms

All strains used were derived from the wild type N2 Bristol strain. Manipulations and maintenance of strains were performed as previously described (Sulston and Hodgkin 1988). Strains used are detailed in the Appendix.

Genotyping

Where it was not possible to follow mutant alleles phenotypically, PCR was used for genotyping. In the case of deletion alleles, PCR primers were designed to anneal either side of the deletion, resulting in a single PCR product in wild type animals, a single, smaller PCR product in individuals homozygous for the deletion, and two differently-sized products in heterozygotes. Unless otherwise specified, PCR was performed using NEB Taq polymerase, using an annealing temperature of 56°C and all other parameters as standard.

rnt-1(tm388) genotyping

Primers *CB232* and *CB233* were used, with an extension time of 2 minutes, 10 seconds.

bro-1(tm1183) genotyping

Primers *CB164* and *CB216* were used, with an extension time of 2 minutes.

arf-3 (tm1877) genotyping

Primers *CB279* and *CB281* were used, with an extension time of 2minutes, 10 seconds.

mec-17(ok2109) genotyping

Primers *CB282* and *CB298* were used, with an extension time of 2 minutes, 30 seconds.

In the case of mutants carrying point mutations, PCR genotyping was performed using the tetra arms PCR approach (Ye, Dhillon et al. 2001).

unc-62(e644) genotyping

Primers *CB152, CB153, CB154* and *CB155* were used, with an annealing temperature of 61°C and an extension time of 1 minute.

unc-62(ku234) genotyping

Primers CB196, CB197, CB198 and CB199 were used, with an annealing temperature of 61°C and an extension time of 1 minute 30 seconds.

pal-1(e2091) genotyping

Primers CB128, CB129, CB133 and CB134 were used, with an annealing temperature of 61°C and an extension time of 1 minute.

rde-1(ne219) genotyping

Primers CB291, CB292, CB293 and CB294 were used, with an annealing temperature of 58°C and an extension time of 1 minute, 20 seconds.

Lineage analysis and microscopy

Lineage analysis was performed as previously described (Nimmo, Antebi et al. 2005). For lineage analysis of *elt-1(RNAi)* animals, 3µl of a freshly prepared solution of M9 and *HT115 E. coli* cells expressing the *elt-1* dsRNA (scraped from a 2mM IPTG NGM plate, on which they had been growing at 20°C for several days) was placed on the pad. For analysis of wild type and mutant worms, OP50 was used as a food source. For each slide, a single worm was transferred into this drop with an eyelash pick. A coverslip was then slowly lowered on top of the worm, and microscopy performed with Nomarski (DIC) optics and a 100x oil immersion objective (Zeiss). Photomicrographs were taken using a 63x Zeiss oil immersion objective and Axiovision software (Release 4.5).

RNAi

For RNAi knockdown of *pal-1*, *arf-3*, *nhr-72*, *nhr-74*, *nhr-75*, *nhr-77*, *nhr-81*, *nhr-82*, *nhr-89*, *psa*-3, *ceh-20* and *ceh-16*, feeding clones from the Ahringer library were checked and sequenced.

elt-1 knockdown by RNAi was performed as described previously (Smith, McGarr et al. 2005), using an identical feeding construct to *pPM88*, named *pAW565*.

In the case of RNAi knockdown of *cnt-2, arf-3* and *unc-62* feeding constructs were made.

Seam-specific RNAi

This was performed essentially as described (Qadota, Inoue et al. 2007). *pAW559*, containing *rde-1* cDNA driven by the seam-specific *"scm"* promoter (Koh and Rothman 2001), was co-injected into *AW553*, together with *scm::gfp* (*pMF1*), *ajm-1::mCherry* and an *unc-119*$^{+}$ rescuing plasmid to make strain *AW552*. *ajm-1* RNAi was performed by injection of dsRNA into *AW552*.

Single worm lysis

Genomic DNA from *C. briggsae* for construction of the *bro-1 ::gfp* reporter construct was obtained by single worm lysis (Williams, Schrank et al. 1992). Two adult hermaphrodites were placed into 2.5µl of lysis mix (100µl lysis buffer consisting of 50µM KCl, 2.5µM $MgCl_2$, 10µM Tris-HCl pH 8.3, 0.45% NP40, 0.45%Tween 20, 0.01% gelatine plus 1µl proteinase K (10mg/ml)), and the solution frozen at -80°C, in a thin-walled 0.5ml Eppendorf tube, for 1 hour. 30µl mineral oil was added and the tube was incubated for 1 hour at 60°C and subsequently for 15 minutes at 95°C.

Primer phosphorylation

Primers were phosphorylated using NEB T4 polynucleotide kinase (PNK). 2.5µl of 40mM primer were mixed with 9.5µl water, 2µl PNK buffer, 5µl 10mM ATP and 1µl PNK. The reaction was incubated at 37°C for 30 minutes.

Plasmid construction

Plasmids used in transgenic strains are listed and described in the Appendix. Where PCR products were generated from genomic DNA, N2 worms were lysed and the entire 2.5µl resultant lysis mix used as a template for the reaction.

bro-1 CNE minimal promoter

The *bro-1* CNE minimal promoter construct was made by first amplifying the 122bp region from genomic DNA (using primers *CB1* and *CB2*) and cloning the PCR product into the *pCR2.1* vector (TOPO TA cloning kit, Invitrogen), producing *pAW553*. The cloned region was then cut from this construct using *HindIII* and *XbaI*, and inserted into the Fire Lab vector *pPD107.94*, cut with the same restriction endonucleases. This vector contains the *pes-10* minimal promoter sequence and GFP, downstream of the multiple cloning site into which the CNE was inserted. The resulting construct was named *pAW304* (*bro-1* CNE::*GFP*).

bro-1 CNE::bro-1 cDNA::gfp

pAW373 (*bro-1* CNE driving *bro-1* cDNA::*gfp*) was made in two steps. Initially, the Fire Lab vector *pPD49.26* was cut with *HindIII* and *BamHI*. The *bro-1* CNE was then cut from *pAW553* using *HindIII* and *PstI* and, from *pPD107.94*, the *pes-10* promoter (Fire Lab vector,

1997 Kit) was cut with *PstI* and *BamHI*. These three fragments were then ligated together, making *pAW400*. The *bro-1*cDNA::*gfp* PCR product was produced in two stages, by fusion PCR (Hobert 2002). To amplify *bro-1* cDNA with a 3'GFP tag, the primers *CB164* and *CB23* were used. Simultaneously, the GFP ORF was amplified from the Fire Lab vector *pPD95.75* using primers *CB90* and *CB92*. After gel purification of both PCR products, the two were mixed and used as template for a fusion PCR, using *CB164* and *CB91*. The primers for the fusion reaction were phosphorylated prior to the PCR, allowing the product to be ligated into the *EcoRV* site of *pAW400* to produce plasmid *pAW373*. This construct was modified to produce *pAW390*, in which GATA site A was deleted, and *pAW393*, in which GATA site B was deleted. *pAW390* was made by PCR, using phosphorylated primers *CB31* and *CB30*. *pAW393* was derived in a similar fashion, using primers *CB22* and *CB60*.

∆*CNEbro-1::dsRED2*

The mutagenesis of the *bro-1::dsRED2* construct, *pAW303* (Kagoshima, Nimmo et al. 2007), to give *pAW305*, was performed by PCR, using phosphorylated primers *CB47* and *CB46*, flanking the 122bp CNE. After the PCR, the template was digested with *DpnI*, and the PCR product purified and self-ligated.

bro-1 control intron::*gfp*/*bro-1* upstream sequence::*gfp*

To make plasmids *pAW545* (*bro-1* control intron::GFP) and *pAW546* (*bro-1* upstream sequence::GFP), the appropriate regions of *bro-1* were amplified with *CB150* and *CB151*, and *CB32* and *CB33*, respectively. Using the tags in the primers, these PCR products were digested with *HindIII* and *XbaI*. The minimal promoter plasmid *pPD107.94* was cut with the same enzymes and ligated with the inserts.

bro-1 5' intergenic region::*gfp*

Primers *CB150* and *CB151* were used to amplify the 864bp intergenic region between bro-1 and the neighbouring 5' gene, *slx-1*. This fragment was cloned into *pCR®-XL-TOPO,* making *pAW645*. *HindIII* and *XbaI* were then used to clone this fragment into *pPD107.94*, giving *pAW646*.

bro-1 first intron::*gfp*

Primers *CB299* and *CB2* were used to amplify the first *bro-1* intron. The PCR product was cloned into the *pCR®-XL-TOPO* vector, giving *pAW647*. *HindIII* and *XbaI* were then used to clone the inserted fragment into the minimal promoter vector *pPD107.94*, making *pAW648*.

scm::rde-1 cDNA

pAW559 (*scm::rde-1cDNA*) was made in two stages. First, *rde-1* cDNA was amplified from a first strand cDNA preparation, using primers *CB289* and *CB290*, and cloned into the *pCR®-XL-TOPO* vector (Invitrogen) creating *pAW570*, and sequenced. In parallel, the *scm* promoter was amplified from *pMF1* (*scm::gfp*) using primers *CB147* and *CB148* and also cloned into *pCR®-XL-TOPO* vector, creating *pAW505*. Once the sequence of the cloned *rde-1* cDNA had been verified, it was re-amplified with phosphorylated primers and blunt-cloned into the *EcoRV* site of *pAW505*, giving *pAW559*.

eff-1p::gfp

pAW549 (*eff-1* transcriptional reporter) was made in two stages. Firstly, a 4kb region of the *eff-1* promoter was amplified using primers *CB267* and *CB266* and cloned in the reverse

direction into *pCR®-XL-TOPO* vector, creating *pAW571.* *pAW571* was then digested with *BamHI* and *PstI*; the promoter-containing fragment was ligated into *pPD95.75*, which had been cut with the same enzymes.

unc-62 feeding RNAi clone

The *unc-62* RNAi feeding construct was made by first amplifying a 1.4kb fragment of *unc-62* cDNA, using primers *CB111* and *CB94*, and cloning it into the *pCR®-XL-TOPO* vector, giving pAW632. From here, the insert was cut using *PstI* and *SpeI* and inserted into the same restriction enzyme sites in the RNAi feeding vector *L4440*, generating *pAW633*.

∆UNC-62 sites *bro-1* cDNA::*gfp*

Site-directed mutagenesis was performed on *pAW553* using primers *CB375* and *CB376*. The protocol used was the same as that used to make *pAW305* (∆CNE*bro-1::dsRED2)*. The mutated CNE was then cloned from this plasmid, *pAW661*, into *pPD107.94* using *HindIII* and *XbaI*. The resultant plasmid was named *pAW660*.

Proximal 4kb of *unc-62* 1a isoform 5' sequence::*gfp*

pAW603 was made by first amplifying the region of sequence flanked by the primers *CB268* and *CB269* and cloning the fragment in the forward direction into the *pCR®-XL-TOPO* vector, making *pAW606*. From this plasmid, the insert was cut with *SphI* and *SpeI* and ligated into *pPD107.94*, which had been cut with the same enzymes.

Distal 4kb of *unc-62* 1a isoform 5' sequence::*gfp*

pAW604 was made by first amplifying the region of sequence flanked by the primers *CB268* and *CB269* and cloning the fragment in the forward direction into the *pCR®-XL-TOPO* vector, making *pAW605*. From this plasmid, the insert was cut with *SphI* and *SpeI* and ligated into *pPD107.94*, which had been cut with the same enzymes.

8kb 5' of *unc-62* 1a isoform 5' sequence::*gfp*

Primers *CB237* and *CB268* were phosphorylated and used to amplify the 8kb promoter fragment which was subsequently blunted in to the *StuI* site of the *pPD107.94* minimal promoter vector, giving *pAW607*.

unc-62 1b promoter::*gfp*

pAW608 was made by first amplifying the 1b putative promoter element from genomic DNA and cloning the fragment into the *pCR®-XL-TOPO* vector in the forward direction, creating *pAW609*. The *unc-62* fragment was then cut from this construct using *XbaI* and *SpeI*, and ligated into *pPD95.75*, which had been linearised with *XbaI*.

scm::unc-62 1b cDNA and *scm::unc-62* 1b(*ku234*) cDNA

The wild type and *ku234* mutant forms of the *unc-62* cDNA were amplified (using primers *CB247* and *CB176* and *CB377* and *CB176* respectively) and cloned into the *pCR®-XL-TOPO* vector (giving plasmids *pAW616* and *pAW617*). From these plasmids, the cDNAs were cut out using *NsiI*, which cuts either side of the insert, and ligated into a vector containing the *scm::gfp* promoter, *pAW505*, which had been linearised with *PstI*. pAW505

was made by cloning the *scm::gfp* promoter, using primers *CB247* and *CB248*, into the *pCR®-XL-TOPO* vector. The resultant plasmids were named *pAW618* (WT cDNA construct) and *pAW619* (*ku234 cDNA* construct).

unc-62 activation domain yeast assay

To make *pAW655*, a positive control, the GAL4 activation domain was amplified from *pGADT7* (*pAW369*) using phosphorylated primers *CB390* and *CB391*.The PCR produce was then blunted into the *SmaI* site of *pGBKT7*.

To make *pAW653*, a fusion PCR was performed. First, primers *CB390* and *CB392* were used to amplify the nuclear localisation sequence positioned 5' to the AD in *pAW369*. This PCR product possessed a tag complementary to the 5' end of the *unc-62* 1b cDNA. The *unc-62* 1b cDNA was also amplified, using primers *CB247* and *CB176*, and a fusion PCR (using phosphorylated primers *CB390* and *CB176*) then performed on the two PCR products, giving an NLS-*unc-62* cDNA fusion. This PCR product was then blunt-cloned into the *SmaI* site of *pAW370* (*pGBKT7*). Exactly the same procedure was performed to make the truncated *unc-62* 1b cDNA equivalent construct, but *unc-62* cDNA was amplified with *CB247* and *CB315*, giving *pAW654*.

scm::gfp dissection constructs

pAW620, pAW621, pAW622 and pAW623 were made by cutting the digesting the plasmid containing the *scm::gfp* reporter construct (pMF1) with *HindIII* and ligating the 1830, 816, 2614 and 2349bp fragments, respectively, into the *HindIII* site of the minimal promoter vector *pPD107.94*.

pAW624 was constructed by digesting *pAW623* with *HindIII* and *XbaI* and ligating the 221bp fragment into the *HindIII* and *XbaI* sites of *pPD107.94*. Making *pAW625* involved cutting *pAW623* with *XbaI* and ligating the 1330bp fragment into the *XbaI* site of *pPD107.94*. Finally, to construct *pAW626* and *pAW627, pAW623* was digested with *XbaI* and *SpeI*; the resultant 727 and 603bp fragments were inserted into the *XbaI* site of *pPD107.94*.

arf-3 feeding RNAi clone

arf-3 cDNA was amplified using primers *CB277* and *CB278*, and the resulting PCR product was cloned into *pCR®-XL-TOPO,* giving *pAW649*, in which the insert lies in the reverse orientation. The cDNA was then cut from this plasmid with *NotI* and *SpeI*, and transferred into the *L4440* RNAi feeding vector (*pAW522*), which had been cut with *NotI* and *NheI*. This plasmid was named *pAW650*.

arf-3::gfp reporter construct

The full length *arf-3* translational reporter construct was made by amplifying the gene, together with 620bp of 5' intergenic sequence, which stretches to the next gene, *unc-24*, from genomic DNA. Primers *CB279* and *CB276* were used; the latter, reverse primer includes a *gfp* tag at its 3' end. Simultaneously, the *gfp* gene was amplified from *pPD95.75*, using primers *CB90* and *CB92*. A fusion PCR was then performed to join and amplify the two PCR products (Hobert 2002), using primers *CB279* and *CB91*. The fusion PCR product was then cloned into the *pCR®-XL-TOPO* vector, giving *pAW628*.

cnt-2 feeding RNAi construct

First, pAW651 was made, by cloning the PCR fragment produced using primers CB307 and CB308 into *pCR®-XL-TOPO*. This plasmid was then cut with *NotI* and *SpeI*, and the feeding vector L4440 was cut with *NotI* and *NheI*, allowing the *cnt-2* fragment to be cloned into the latter, giving *pAW652*.

rnt-1 promoter dissection constructs

pAW376 was made by amplifying the *rnt-1* intron from genomic DNA using primers *CB5* and *CB6*, the former containing an *NheI* tag. The PCR product was cloned into the *pCR®-XL-TOPO* vector, giving *pAW629*. From this plasmid, the insert was liberated using *NheI* and XbaI and subsequently cloned into the *XbaI* site of the minimal promoter vector *pPD107.94*.

pAW397 was made by first cloning the region flanked by primers CB5 and CB21 into pCR®2.1-TOPO®, giving *pAW367*. The insert was then cut from this plasmid and inserted into *pPD107.94* using *HindIII* and *XbaI*. The same approach was taken for the construction of *pAW364; pAW504* was first made, using primers *CB27* and *CB21*, and the insert was placed into *pPD107.94* with *HindIII* and *XbaI*.

To make *pAW398*, primers *CB17* and *CB11* were used to amplify the relevant promoter region from genomic DNA; this fragment was cloned into *pCR®-XL-TOPO,* giving *pAW396*. Again, *HindIII* and *XbaI* were used to clone this region into *pPD107.94*.

To make *pAW395*, primers *CB26* and *CB7* were used and the amplified fragment was then clone into pCR®2.1-TOPO®, making *pAW630*. The fragment was then excised using *XbaI* and *SpeI* and cloned into the *XbaI* site of *pPD107.94*.

Construction of *pAW399* involved amplification of the genomic region flanked by *CB26* and *CB36* using phosphorylated primers. The fragment was then blunt-cloned into *pPD107.95*, which had been linearised with *StuI*.

pAW394 was made by amplifying the fragment flanked by *CB41* and *CB7*, and cloning this region into pCR®2.1-TOPO®, giving *pAW631*. *HindIII* and *XbaI* were then used to clone a portion of this fragment into the *HindIII* and *XbaI* sites of *pPD107.94*.

pal-1::gfp reporter

The *pal-1::gfp* PCR product was produced in two stages, by fusion PCR (Hobert 2002). First, the *pal-1* ORF and promoter were amplified using primers *CB221* and *CB223*, the latter possessing a *gfp* tag. Simultaneously, the GFP ORF was amplified from *pPD95.75* using primers *CB90* and *CB92*. After gel purification of both PCR products, the two were mixed and used as template for a fusion PCR, using *CB221* and *CB91*. The PCR product was then cloned into *pCR®-XL-TOPO*, giving *pAW634*.

pal-1 promoter transcriptional reporter

The 1.3kb intergenic region between *pal-1* and the neighbouring upstream gene, *C38D4.7*, was amplified using primers *CB221* and *CB317*; the resulting PCR product was cloned into *pCR®-XL-TOPO*, giving *pAW635*. From this plasmid, containing the *pal-1* promoter in the reverse orientation, the promoter was cut, using *SpeI* and *SphI*. The insert was then ligated into the Firelab vector *pPD95.75*, which had been cut using *XbaI* and *SphI*, giving *pAW636*.

pal-1 last intron minimal promoter construct

A region of sequence which included the last intron of *pal-1* was amplified from genomic DNA with primers *CB128* and *CB223*. This fragment was cloned into *pCR®-XL-TOPO* and the resultant plasmid, which contains the insert in the reverse orientation, named *pAW637*. From this, a smaller region consisting only of the intron and a few bases of the surrounding exons (exons 5 and 6) was cut using *SpeI* and *HindIII*, and inserted into the minimal promoter vector *pPD107.94* which had been cut with *HindIII* and *XbaI*. Plasmid *pAW638* resulted.

pal-1 last intron minimal promoter construct with *e2091* mutation

pAW640, which consists of the last intron of *pal-1(e2091)* mutants in the minimal promoter vector, *pPD107.94*, was made in the same way as *pAW638*, described above. The TOPO intermediate construct was designated *pAW639*. The only difference was that the initial genomic amplification step was performed on *pal-1(e2091)* animals, rather than on wild type worms.

All constructs were checked by sequencing before use.

Construction of transgenic worms

Injections were performed as described previously (Mello and Fire 1995) using the *unc-119*$^{+}$ (*pDP#MM016β*) transformation marker (Maduro and Pilgrim 1995). Constructs were injected at 10-20ng/µl.

Band shift experiments

cDNAs of *elt-1*, *elt-2*, *elt-3*, *elt-5* and *elt-6* were amplified from a mixed stage cDNA preparation using Phusion polymerase (Finnzymes). (*elt-4* was not amplified, but no function has yet been detected for this gene (Fukushige, Goszczynski et al. 2003) The PCR products were cloned into the *pCR®-XL-TOPO* vector (Invitrogen) and the *TNT®* Quick Coupled Transcription/Translation kit (Promega) was then used for *in vitro* transcription and translation. To make the labelled probes, oligonucleotides covering 'GATA site B' were synthesised (WT probes*: CB204*, *CB206*; mutant probes: *CB205*, *CB207*), annealed by heating and gradual cooling, and labelled with [γ-^{32}P] dATP (for hot WT probes) or without [γ-^{32}P] dATP (for cold, competitor probes). The DNA binding reaction was carried out on ice for 30 minutes before the reaction mixture was loaded onto a 7% non-denaturing polyacrylamide gel and run at 4°C in 0.5xTBE.

Construction of yeast strains and yeast one-hybrid screen

3 copies of the 122bp *bro-1* CNE were inserted in the forward direction into the Clontech Matchmaker vectors *pHisi-1* and *pLacZi* by blunt-cloning the CNE PCR product (made using phosphorylated primers *CB1* and *CB2*) into the vectors which were cut with *SmaI*. The plasmids were integrated into the yeast strain *YM4271* (*MATa*, *ura3-52*, *his3-Δ200*, *ade2-101*, *lys2-801*, *leu2-3, trp1-901*, *tyr1-501*, *gal4-Δ512*, *gal80-Δ538*, *ade5::hisG* (Liu, Wilson et al. 1993)) creating the double integrant strain *YM4271* [p*bro-1*HIS; p*bro-1*LAC], using the LiAc transformation procedure, as described in the Clontech Matchmaker manual. *YM4271* [p*bro-1*HIS; p*bro-1*LAC] was then transformed with a mixed stage transcription factor cDNA library and plated onto *–his, -trp, -ura* 15mM 3-AT SD plates. All colonies that grew within 3 days were assayed for lacZ expression (Breeden and Nasmyth 1985), and after re-isolating plasmids and re-checking, positive clones were sequenced.

The same protocol was followed for the *arf-3* 5' intergenic region yeast 1-hybrid screen. Three copies of the bait sequence were sequentially cloned into *pHisi-1.* Initially, the region of interest was amplified using primers *CB279 and CB280* and cloned into *pCR®-XL-TOPO*, giving *pAW656*. From this plasmid, the *arf-3* fragment was cut, using *SpeI* and *XbaI* and cloned into the *NheI* site *pHisi-1*. This plasmid, *pAW657*, was used as the basis for the next round of cloning; the *arf-3* region was cut from *pAW656* and cloned into *pAW657*, giving *pAW658*, which contained two copies of the sequence. This process was repeated once more, giving *pAW659* (*pHisi-1* with three copies of the fragment *CB279-CB280*).

Quantitative RT-PCR analysis

qRT-PCR was performed on synchronised worms obtained by bleaching gravid animals and seeding the eggs onto *elt-1* RNAi or L4440 control plates. Larvae were harvested after 1-2 days by washing with M9. RNA was extracted by the hot phenol method (Furger, Monks et al. 2001) and mRNA levels of *elt-1/bro-1* and a normaliser (*nuo-2*, expressed in all seam cells) were assessed using SYBR-Green and a Qiagen Rotor-Gene Q machine. Expression levels were assayed by the $2^{-\Delta\Delta C}$T method (Livak and Schmittgen 2001).

Acknowledgements

The transcription factor library used in the Y 1-H screen was kindly provided by Marian Walhout. I would like to acknowledge Iva Greenwald, for the gift of the strain *arIs99*, Jeff Simske for the *ajm-1::mCherry* construct, Joel Rothman for *pMF1* and the *Caenorhabditis* Genetics Center (CGC) for certain strains.

CHAPTER 3

The transcription factor ELT-1 works through BRO-*1* and the fusogen EFF-1 to maintain the seam fate

Introduction

The regulation of the decision between cell proliferation and differentiation is a key aspect of metazoan development, ensuring that the correct number and type of cells are present. The RUNX family of transcription factors (*RUNX1, 2* and *3*), together with their binding partner CBFβ, are key players in the control of stem cell proliferation in haematopoiesis (Okuda, van Deursen et al. 1996; Wang, Stacy et al. 1996; Chen, Yokomizo et al. 2009), osteogenesis (Komori, Yagi et al. 1997; Otto, Thornell et al. 1997; Hassan, Gordon et al.) and neurogenesis (Inoue, Ozaki et al. 2002; Levanon, Bettoun et al. 2002). Moreover, mutations in these genes are known to cause a variety of diseases, with both CBFβ and RUNX genes having the potential to act as either oncogenes or tumour suppressors, depending on the nature of the mutations and the context in which they act (Blyth, Cameron et al. 2005).

As outlined in Chapter 1, in *Caenorhabditis elegans*, RUNX and CBFβ are each represented by a single gene, *rnt-1* and *bro-1* respectively. Reflecting their role in mammalian tissues, RNT-1 and BRO-1 are involved in the regulation of the division patterns of the stem cell-like seam cells (Nimmo, Antebi et al. 2005; Kagoshima, Nimmo et al. 2007; Xia, Zhang et al. 2007). The seam comprises a specialized epithelial tissue of lateral hypodermal cells located along each side of the worm and provides a tractable model system for studying the balance between proliferative and differentiative developmental decisions. The seam cells are considered stem cell-like because of their ability to both self-renew and to produce a variety of differentiated cell types. Seam cells do not self-renew throughout adult life, and have not

yet been shown to reside in a classic “niche”, hence the qualification stem-like, but they do provide a useful, simplified and well established model for the stem cell mode of division throughout larval development (Eisenmann, 2010; Silhankova, Jindra et al. 2005; Albert Hubbard 2007; Mizumoto and Sawa 2007; Huang, Tian et al. 2009; Nimmo and Slack 2009; Gleason and Eisenmann ; Ren and Zhang 2010).

In *C. elegans*, *rnt-1* and *bro-1* are rate-limiting regulators of seam cell divisions, with both genes being required for proliferation in this tissue. In *rnt-1* and *bro-1* mutants, the number of seam cells present in individuals is reduced due to division failures within the seam lineage (Nimmo, Antebi et al. 2005). Furthermore, over-expression of *rnt-1* and *bro-1* results in seam hyperplasia (Kagoshima, Nimmo et al. 2007). Work is underway to further our understanding of the temporal dynamics of *rnt-1* expression, and of the mechanisms underlying what appears to be a tightly controlled system (A. Woollard, T. Braun and R. Nimmo, pers. comm.). The spatial regulation of *rnt-1* has also been investigated (Chapter 6), and specific, relatively small enhancer regions which direct tissue-specific expression have been identified.

In the case of *bro-1*, however, the situation is somewhat different. In temporal terms, *bro-1* expression appears to be much less dynamic; in the cells in which it is expressed, constant expression is seen throughout the development of the worm, from embryogenesis to adulthood. Compared with *rnt-1,* the activity of which appears to be regulated on at least three levels (by transcription factors, microRNAs and protein degradation, T. Braun, R. Nimmo and A. Woollard, pers. comm.), the relatively simple expression pattern of *bro-1* has perhaps provided fewer points of attack for investigation. Here, the mechanisms underlying the spatial expression pattern of *bro-1* – in the seam cells and in body wall muscle – are investigated. The data presented in this chapter led to the discovery of a seam-specific enhancer within an intron of *bro-1*, and of the identification of ELT-1 as a direct transcriptional regulator acting via this site. Furthermore, subsequent investigation of ELT-1 revealed that it operates in the seam both together with and independently of *bro-1*, having

close links with the cell fusion pathway which mediates seam cell differentiation and loss of stem-like characteristics. As well as allowing dissection of the differentiation process, this work also raised the issue of the stem cell niche, and whether such a concept is appropriate for the seam cells of *C. elegans*. Further supporting evidence is required, as discussed below, but it is at least plausible, if not likely, that the seam cells do indeed reside in some kind of niche, sheltered from differentiative signals emanating from surrounding tissues and thereby maintained in the proliferative, stem-like state.

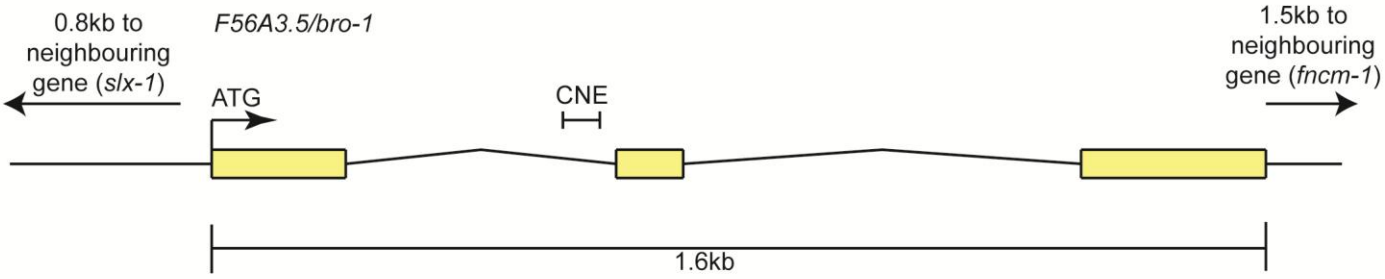
Results

The first intron of *bro-1* contains a highly conserved sequence element that is both necessary and sufficient for *bro-1* expression in seam cells

Comparative sequence analysis revealed high levels of conservation between exonic regions of *bro-1* in *C. elegans* and *C. briggsae* (73, 85 and 75% identity for exons 1, 2 and 3 respectively, compared to an average of 53% for introns 1 and 2). In contrast, no significant blocks of conservation were found in the 0.8kb between *bro-1* and the next upstream gene. However, a 122 bp region within the first intron was found to be highly conserved (3 species alignment shown in Figure 3.1B) (69% identity compared to 48% for the rest of the introns). This is henceforth termed the "*bro-1* Conserved Non-coding Element" (*bro-1* CNE).

Next the ability of the *bro-1* CNE to act as an enhancer for cell-specific expression was tested. When used in conjunction with the *pes-10* minimal promoter (*bro-1* CNE::*gfp*), the CNE was able to drive seam cell-specific GFP expression in worms (Figure 3.2A,B,C). A vector containing a similar-sized fragment of a different, non-conserved region of intron 2 failed to drive any discernible GFP expression (data not shown). Furthermore, when the intergenic region between *bro-1* and the nearest upstream gene was used, no GFP expression was evident (data not shown). To further characterise the importance of the CNE in regulating *bro-1* expression, the 122bp region was deleted from the *bro-1::dsRED2* construct (Kagoshima, Nimmo et al. 2007). The wild type construct (containing the full *bro-1* genomic region) drives expression in seam cells and rescues the *bro-1* mutant male tail phenotype caused by division failures in V and T lineage seam cells (Figure 3.2J). In contrast, the mutagenised *bro-1::dsRED2* construct did not express in seam cells, and was unable to rescue the *bro-1* male tail phenotype (Figure 3.2K). Secondly, the CNE (together with the *pes-10* minimal promoter) was used to drive expression of *bro-1 cDNA::gfp* (see Materials and Methods and (Hobert 2002)). This construct drove seam cell expression, and rescued the

A



B

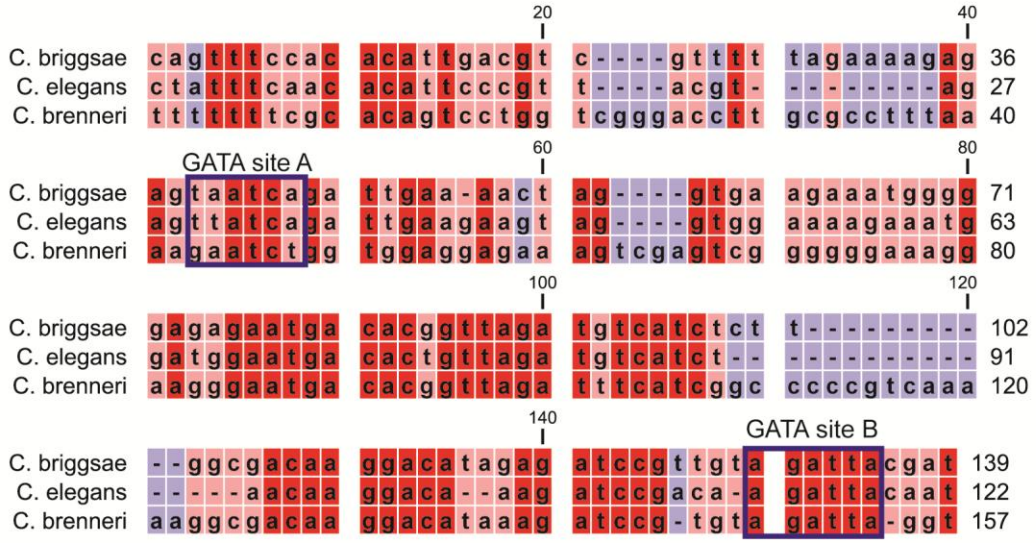


Figure 3.1 A. Genomic structure of *bro-1*, showing the size and position of the CNE. B. Alignment of the *C. elegans bro-1* CNE with orthologues from *C. briggsae* and *C. brenneri* reveals high levels of sequence conservation at the 3' end of intron 1. Conservation is still evident, though less pronounced, when *C. remanei* and *C. japonica* are included in the alignment (no stretches of sequence longer than three base pairs are 10% conserved across the five species and conservation across exonic regions is also significantly lower when all five species are included in the alignment; data not shown). Predicted GATA binding sites are labelled. Sequence alignments were performed using CLC Sequence Viewer v6.4 (CLC bio, Aarhus, Denmark). Red shading denotes that sequences are conserved across all 3 species of nematode; pink denotes conservation in 2 out of 3 species; blue denotes no conservation.

bro-1 mutant phenotype (Figure 3.2L). Thus, the *bro-1* CNE is both necessary and sufficient for *bro-1* expression in seam cells.

The *bro-1* first intron also drives muscle expression

As described above, the mutagenised *bro-1::dsRED2* construct, which lacks the CNE, fails to drive seam expression. Deletion of the CNE does, however, cause *bro-1* to be expressed relatively strongly in body wall muscle (Figure 3.2E-G). Thus, the CNE is clearly not necessary for muscle expression of *bro-1* and, furthermore, likely has a repressive effect on expression in muscle tissue; deletion of this region thus allows *bro-1* to be expressed at higher levels in the muscle. As described above, no expression in either muscle or seam was observed when the *bro-1* 5' intergenic region was cloned into the minimal promoter vector *pPD107.94*. Given the ability of the *bro-1ΔCNE::dsRED2* construct to drive muscle expression, and the fact that 5' intergenic sequence had been ruled out as being responsible for muscle expression, the entire first intron was cloned into the *pPD107.94*. As shown in Figure 3.2H, muscle expression was observed, together with the seam expression driven by the CNE. The precise location of muscle enhancer region within the *bro-1* intron has not been narrowed down further, but it is interesting to note that the intron contains two motifs (1x CACTTCT and 2x WCTTTGM) which have been identified as driving muscle-specific expression (Zhao, Schriefer et al. 2007). Furthermore, without additional dissection of this region of sequence, it is at least apparent that the first intron of *bro-1* is able to drive the complete endogenous expression pattern of this gene.

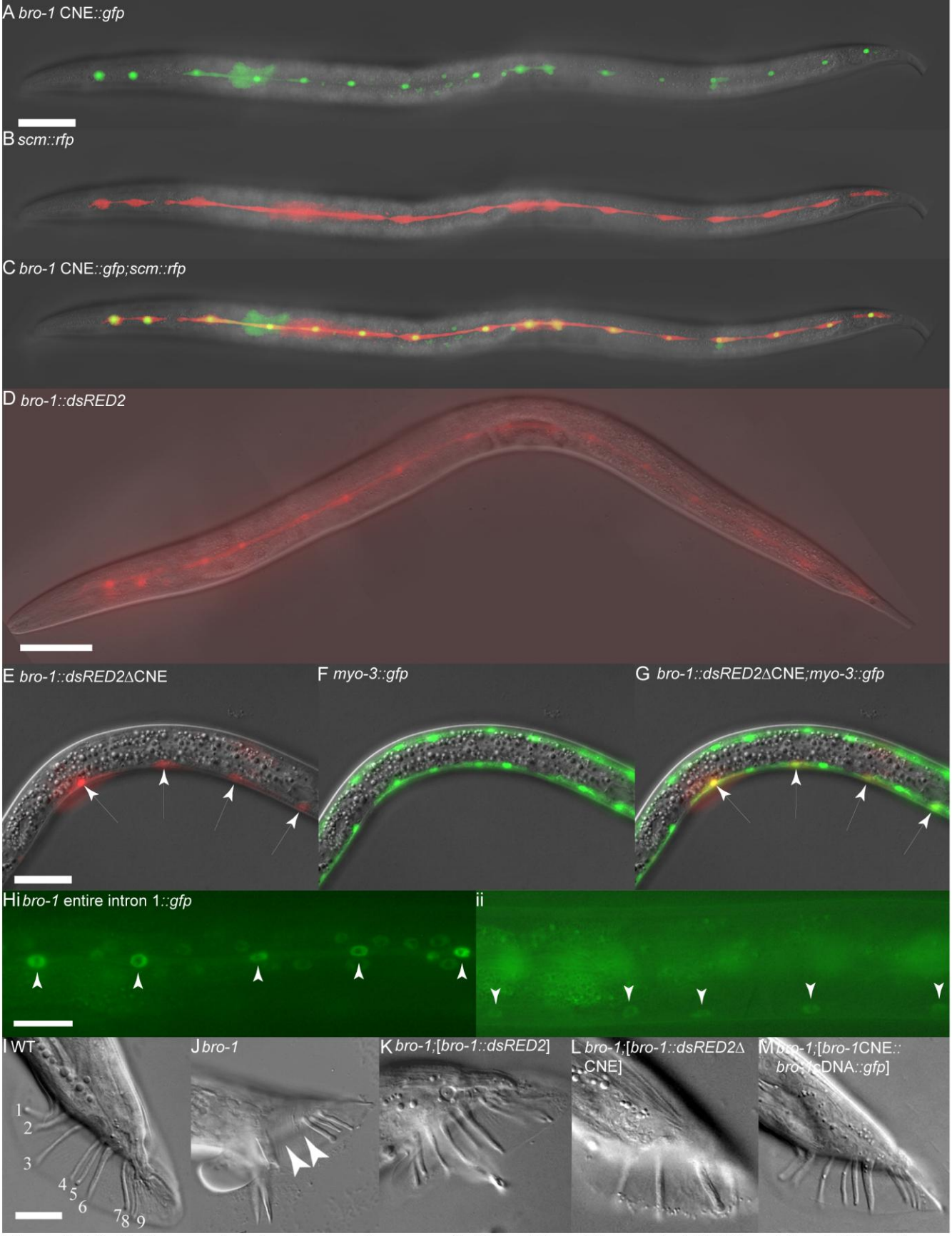


Figure 3.2 A–C. Transgenic animal expressing *scm::rfp* (Kagoshima, Nimmo et al. 2007) and *bro-1 CNE::gfp.* The co-localisation of these two reporters in seam cells demonstrates that the 122 bp CNE is sufficient to drive seam cell expression of GFP when present as a single copy upstream of the *pes-10* minimal promoter in *pPD107.94* (containing an NLS). RFP, driven by *scm::rfp,* localises to both the nucleus and cytoplasm of seam cells. A = *bro-1 CNE::gfp* only, B = *scm::rfp* only, C = merge. D. *bro-1::dsRED2* expression in the seam cells of a transgenic hermaphrodite. E–G. Transgenic animal expressing the *bro-1::dsRED2* construct with the CNE deleted and *myo-3::gfp* to mark body wall muscle cells. A = *bro-1::dsRED2ΔCNE* only, B = *myo-3::gfp* only, C = merge. Deletion of the CNE abolishes seam expression of *bro-1* and results in high levels of expression in body wall muscle (arrows), where *bro-1* is usually expressed only very weakly. *dsRED2* expression is variably absent from some nuclei owing to mosaicism of the extrachromosomal array; the *myo-3::gfp* marker is integrated. Scale bars for A–G, 50 µm. H. Expression of *GFP* driven by the entire *bro-1* first intron in the minimal promoter vector *pPD107.94.* Expression, marked by arrowheads, is apparent in the seam (Hi) and body wall muscle (Hii). Scale bar, 25µm. (I) WT male tail with 18 sensory rays, 9 on each side (numbered 1–9). J. *bro-1* male tail, exhibiting missing rays (arrowheads). K. *bro-1* male tail, rescued by full length *bro-1::dsRED2.* L. Transgenic *bro-1* animal expressing *bro-1::dsRED2*, from which the CNE has been deleted. No rescue of the male tail phenotype is observed. M. *bro-1* male tail, rescued by *bro-1 cDNA::gfp* driven by the CNE. Scale bars for I–L, 20µm.

The GATA transcription factor ELT-1 interacts directly with the *bro-1* CNE

Next, a yeast one-hybrid system was used to determine which transcription factors are able to bind to the *bro-1* CNE, driving seam cell expression. Three tandem copies of the CNE were used as bait and screened with a mixed stage *C. elegans* transcription factor cDNA library. Of 120 positive colonies, 110 contained the clone encoding the GATA factor ELT-1, demonstrating that ELT-1 protein binds to the *bro-1* CNE in this assay.

The bioinformatics software Patch (Matys, Fricke et al. 2003) revealed the presence of possible binding sites for GATA transcription factors in the CNE (Figure 3.1B), two of which are relatively well conserved across *Caenorhabditis* species (Evans, Reitman et al. 1988; Sakai, Nakagawa et al. 1998). Site A matches the GATA consensus sequence WGATAR, while site B is slightly different (AGATTA), but matches the target sequence for the human GATA family member GATA6 (Sakai, Nakagawa et al. 1998). To investigate the significance of these sites, both were deleted separately from the *bro-1::dsRED2* rescuing construct. While deletion of site A had no effect on tail rescuing ability, deletion of site B abolished the ability of the construct to rescue *bro-1* mutant tails (Figure 3.3A).

ELT-1 specifically binds GATA site B within the *bro-1* CNE

To confirm this interaction, an Electrophoretic Mobility Shift Assay (EMSA) was performed using *in vitro* translated ELT-1 protein and a portion of the *bro-1* CNE containing GATA site B as a probe. ELT-1 protein was able to bind the CNE, resulting in a shift in the position of the labelled probe (Figure 3.3B). Cold competitor probe was able to compete with the identical labelled probe, resulting in a diminution in intensity of the shifted band. However, cold probe with a mutation in GATA site B (AGATTA to ATAGTA) was unable to compete with the wild type labelled probe, suggesting that ELT-1 protein interacts directly with GATA site B.

Other *C. elegans* members of the GATA family of transcription factors were tested for their ability to bind the *bro-1* CNE using the same band shift assay; all failed to shift the labelled probe (Figure 3.3C). Taken together, these data suggest that amongst the members of the GATA family in *C. elegans*, ELT-1 alone mediates *bro-1* transcriptional activity in the seam by direct binding to GATA site B within the CNE.

ELT-1 regulates *bro-1* expression *in vivo*

The fact that *bro-1* seam expression disappears when ELT-1 binding site B within the *bro-1* regulatory region is deleted suggests that ELT-1 is an activator of *bro-1* expression. To confirm this, *bro-1::gfp* expression was monitored in animals subjected to *elt-1* RNAi and a decrease in signal intensity was observed (data not shown). It was necessary to reduce the level of *elt-1* expression by RNAi as *elt-1* alleles are embryonic lethal; therefore only the effects of a small reduction in *elt-1* expression could be analysed. This decrease of *elt-1* transcript in animals surviving the RNAi treatment was quantified by qRT-PCR (Pete Appleford, pers. comm.). These animals (most of which exhibited seam defects) were found to have a 1.5 fold reduction in *elt-1* transcript levels and a 1.7 fold decrease in *bro-1* transcript levels, demonstrating that ELT-1 plays a major role in regulating *bro-1* expression *in vivo*.

ELT-1-deficient worms, like *bro-1* mutants, exhibit division failures in the seam lineage

ELT-1 has previously been reported to be essential for seam differentiation and maintenance, with *elt-1(RNAi*) animals having fewer seam cells as assayed by the seam cell marker *scm::gfp* (Smith, McGarr et al. 2005). Significantly, the average number of seam cells in *elt-1(RNAi)* animals (13.6 seam cells per side, ±0.3) is very similar to that of *bro-1* mutants

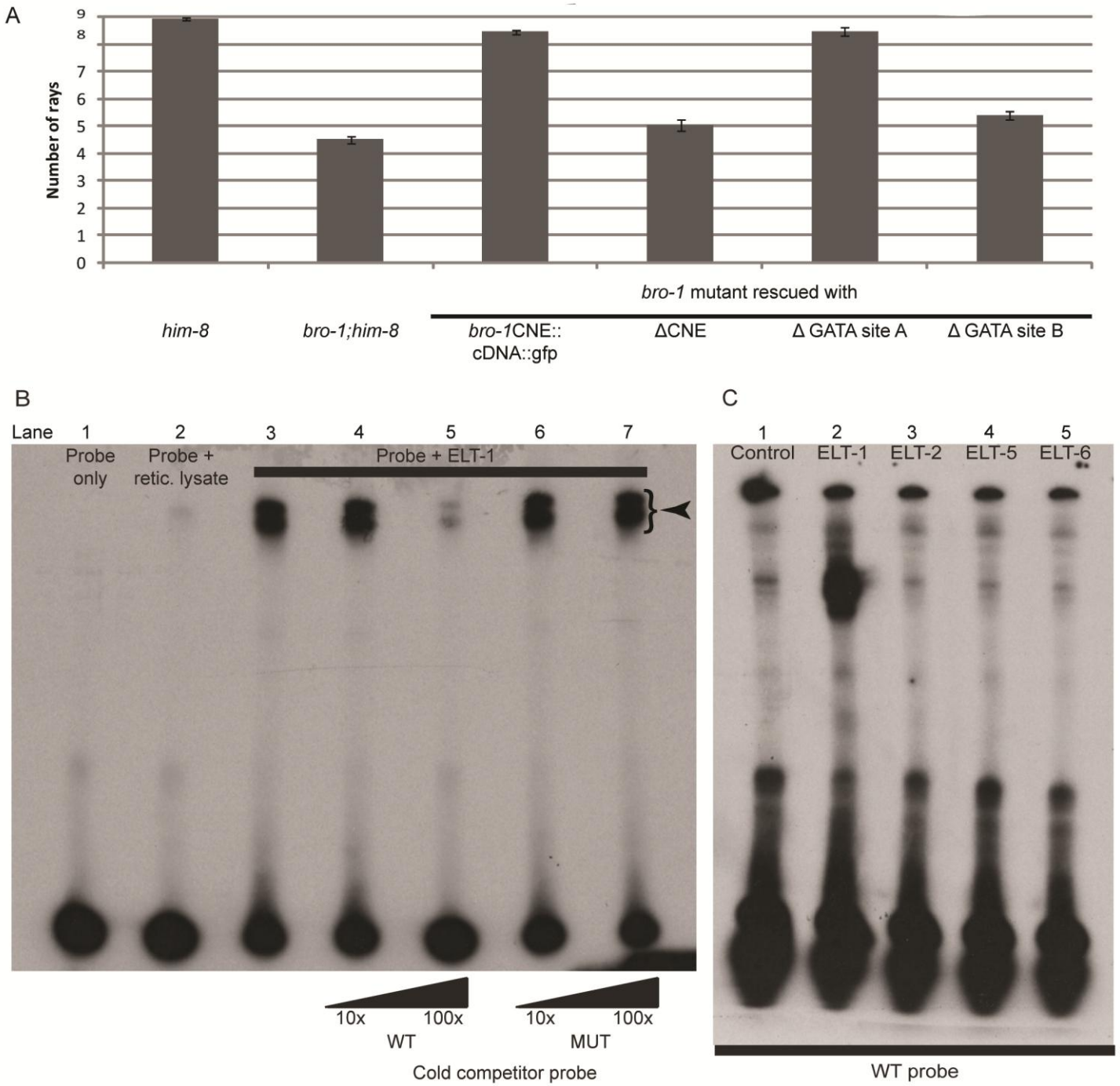


Figure 3.3 A. Bar chart showing ray number in the following strains: *him-8, bro-1;him-8, bro-1;him-8; bro-1CNE::bro-1cDNA::gfp, bro-1;him-8;bro-1ΔCNE::dsRED2, bro-1;him-8;bro-1CNEΔGATA site A::bro-1cDNA::gfp, bro-1;him-8;bro-1CNEΔGATA site B::bro-1cDNA::gfp*. Deletion of GATA site A has no effect on the ability of the *CNE::bro-1 cDNA* construct to rescue the *bro-1* mutant tail phenotype, while deletion of site B abolishes the rescuing ability of the construct. Error bars show SEM. B. EMSA showing that in vitro translated ELT-1 shifts labelled probe (lane 3, DNA-protein complexes labelled with an arrowhead), which is not shifted on its own or in the presence of reticulocyte lysate only (lanes one and two respectively). Non-labelled WT cold competitor probe at 100x concentration (lane 5) significantly reduces the intensity of the shifted bands. Addition of mutated competitor probe has no such effect at 100x concentration (lane 7), demonstrating that the ELT-1-CNE interaction is dependent on GATA site B, which is altered in the mutated probe. C. The interaction between ELT-1 and the *bro-1* CNE is specific. Incubation of ELT-1 protein with labelled probe results in a shift (lane 2) relative to the control incubation (lane 1). Incubation of other GATA transcription factors (ELT-2, ELT-5 and ELT-6) with the labelled probe resulted in no shift. ELT-3 was not tested as in vitro transcription and translation proved problematic but, like ELT-2, ELT-3 is not expressed in the seam (Gilleard, Shafi et al,1999) so would not bind to *bro-1* in seam cells, and was not detected amongst positive clones in the yeast one-hybrid screen, despite being present in the TF library used.

(14.0 seam cells per side, ±0.3) and significantly lower than animals fed control *HT115* bacteria containing the empty feeding vector L4440 (15.8 seam cells per side, ±0.1). The cellular basis of this phenotype has not been previously described, therefore lineage analysis was performed to elucidate the cellular mechanism of seam cell loss.

elt-1(RNAi) worms have variable seam division failures (Figure 3.4A). These defects were observed during both the symmetrical and asymmetrical divisions of L2, suggesting that the role of ELT-1 is not limited to one type of division. However, given the difficulties in pursuing the lineage analysis past L3 (these worms are very sick), it is not possible to exclude the possibility that the role of ELT-1 in regulating seam cell division is limited to the L2 stage. Division failures were not observed during the L1 asymmetric division, although embryonic developmental abnormalities often meant that the number of seam cells present at the start of L1 was lower than that of wild type worms. Nevertheless, the ability of the remaining seam cells to undergo the asymmetric L1 division parallels the situation in *rnt-1* and *bro-1* mutants, where division failures are restricted to the subsequent divisions. These division defects are distinct from the classic retarded heterochronic phenotypes, in which stage-specific stereotypical division "cassettes" occur at the wrong stage. In the case of *lin-4*, the L1 pattern of division is re-iterated at later stages, resulting in fewer seam cells overall (Ambros and Horvitz 1984). In the case of *elt-1(RNAi)* and *rnt-1/bro-1* mutants, the defects seen were not representative of stereotypical division patterns occurring at the wrong stage, but rather involved outright division failures, or occasionally symmetrisation towards the hypodermal fate.

To assess whether the division failures observed in *elt-1(RNAi)* worms were correlated with loss of seam fate, the integrated *scm::gfp* reporter *wIs51* (Koh and Rothman 2001) was used as a seam cell marker. Even worms which do not show gross morphological abnormalities, and which therefore can be followed by lineage analysis throughout development, frequently exhibit loss of *scm::gfp* from usually one or two seam cells from the late L1 stage onwards. While the seam cell remains clearly visible under DIC (retaining its distinctive eye-shaped

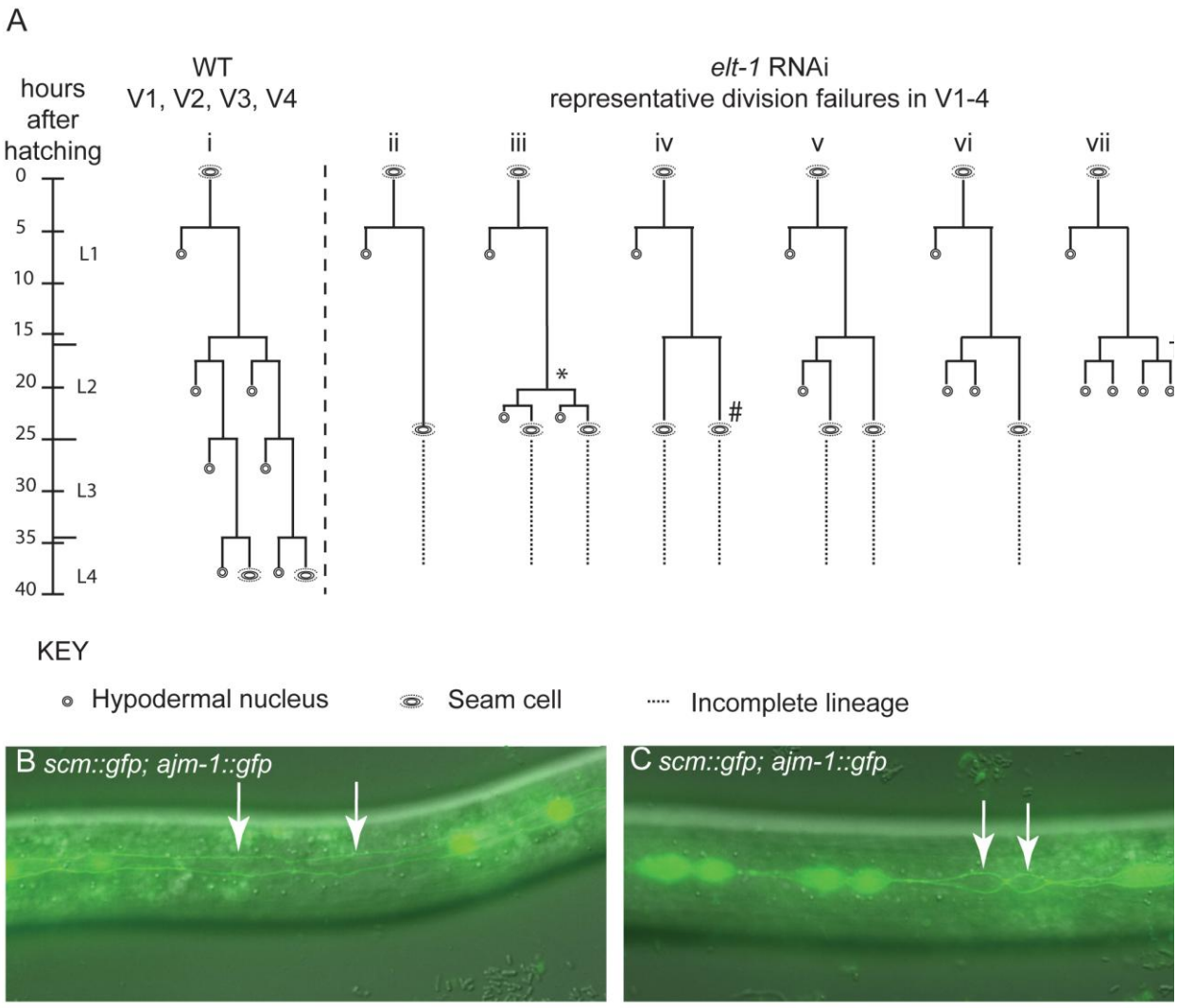


Figure 3.4 A. Lineage traces of WT and *elt-1(RNAi)* hermaphrodites. In *elt-1(RNAi)* worms, seam cells often fail to divide but remain intact and retain their stereotypical seam cell 'eye' shape. The 'extent' of division failure during development is variable; the L2 symmetric division can fail totally (as in trace ii) or can progress normally, only for subsequent division failures to occur in both daughters (trace iv) or one daughter (trace v). One or more divisions be missed (compare, for example, lineage traces ii and iv (#)) and divisions were affected at different stages. From this data, there was no obvious bias for the V cell involved but it is possible that, were lineage analysis to be carri out on more animals, patterns may emerge. A similar principle applies to the question of whether there is an anter posterior bias of division failures; the number of worms followed here allows conclusions to be drawn about the na of division failures but cannot be used as the basis of more detailed inferences about the likelihood of each kind o failure arising - larger n values would be required.

Alternatively, certain V lineage seam cells sometimes undergo the symmetrical and asymmetrical divisions during but significantly later relative to both other V cells and to the time of moulting. For example, whereas Vn.p cells sh divide first symmetrically and then asymmetrically at the start of L2, these divisions have been observed to occur before the L3 moult in *elt-1(RNAi)* animals, around 8 hours after they should have occurred (as in lineage trace iii Similarly, the L2 symmetric and asymmetric divisions in V1-4 should take place slightly ahead of the comparable V divisions in WT worms. *elt-1* RNAi, however, often delays these divisions such that they occur after V6p has divid symmetrically and then asymmetrically in L2.

In addition to these proliferation defects, aberrant fate changes at division were also occasionally observed (approximately 10% of defects observed); instead of the division producing two seam cells or one seam daughter one hypodermal daughter, both cells produced adopted the hypodermal fate, becoming smaller and rounder than normal seam cells, losing expression of the *scm::gfp* marker and failing to divide further (e.g. trace vi and trace vii This defect was also observed in *rnt-1* and *bro-1* animals (Kagoshima, Nimmo et al. 2007). Designation of the sea fate was based on cell shape, position and division potential, where possible, not *scm::gfp* expression. In no worm were L1 division defects observed.

B. and C. *elt-1(RNAi)* animals carrying the *scm::gfp* and *ajm-1::gfp* transgenes. Seam cells often lose *scm::gfp* expression (white arrows). 52% of *elt-1(RNAi)* worms were found to have one or more cells bounded by AJM-1::G but lacking SCM::GFP (n=61) compared to 5% in control worms, fed on HT115 bacteria containing the empty L44 RNAi feeding vector (n= 62). As can be seen, these cells retain their AJM-1::GFP boundary and seam morpholog and were not observed to fuse with hyp7. The loss of SCM::GFP appears to be independent of the cells' stage in cell cycle and can happen before (B) or after (C) division. Scale bar, 25µm.

seam morphology), GFP expression fades over 30-60 minutes until the cell shows no fluorescence at all (Figure 3.4B). To confirm the identity of these 'seam' cells that fail to express *scm::gfp*, cell boundaries were visualised using the *ajm-1::gfp* reporter (which marks the apical junctions between seam cells and the surrounding hyp7 syncytium, in which it is not expressed). These cells were always bounded by AJM-1::GFP, confirming their seam identity (as suggested by their morphology). Thus, when seam cells are counted in *elt-1(RNAi)* animals using *scm::gfp*, not all of the seam cells present will be observed. In addition to division failures, therefore, the loss of expression of the 'seam' marker *scm::gfp* also likely accounts for the apparent progressive loss of seam cells previously reported in *elt-1(RNAi)* worms (Smith, McGarr et al. 2005).

Interestingly, a similar phenotype is observed in *bro-1* and *rnt-1* mutants (data not shown). However, while loss of this marker has been attributed to degeneration of seam cells (Smith, McGarr et al. 2005), this phenomenon has not been observed in either *rnt-1*/*bro-1* mutants or *elt-1(RNAi)* worms; seam cells retain their characteristic "eye" shape and position within the seam line, but may lose *scm::gfp* in these backgrounds. Given the prevalence of using *scm::gfp* as a marker of seam fate, it was surprising to find that *scm::*gfp expression does not in fact seem to be tightly correlated with proliferative potential; lineage analysis revealed that cells that stop expressing *scm::gfp* often undergo the correct number of divisions at the appropriate times. Conversely, division failures were observed in seam cells strongly expressing *scm::gfp*. Taken together, this suggests that while *scm::gfp* may mark seam cells during normal development, it is not an infallible marker for seam cell fate, perhaps because particular elements of seam 'fate' are uncoupled in *elt-1(RNAi)* and *rnt-1/bro-1* mutant animals.

Cells that remained in the seam line but failed to divide, always retained expression of *ajm-1::gfp*, regardless of whether or not they expressed *scm::gfp* (Figure 3.4B and 3.4C). This suggests that their failure to divide results not from a change of fate *per se* (from seam to hypodermis), but rather from a loss of proliferative ability. This is supported by *dpy-7p::yfp*

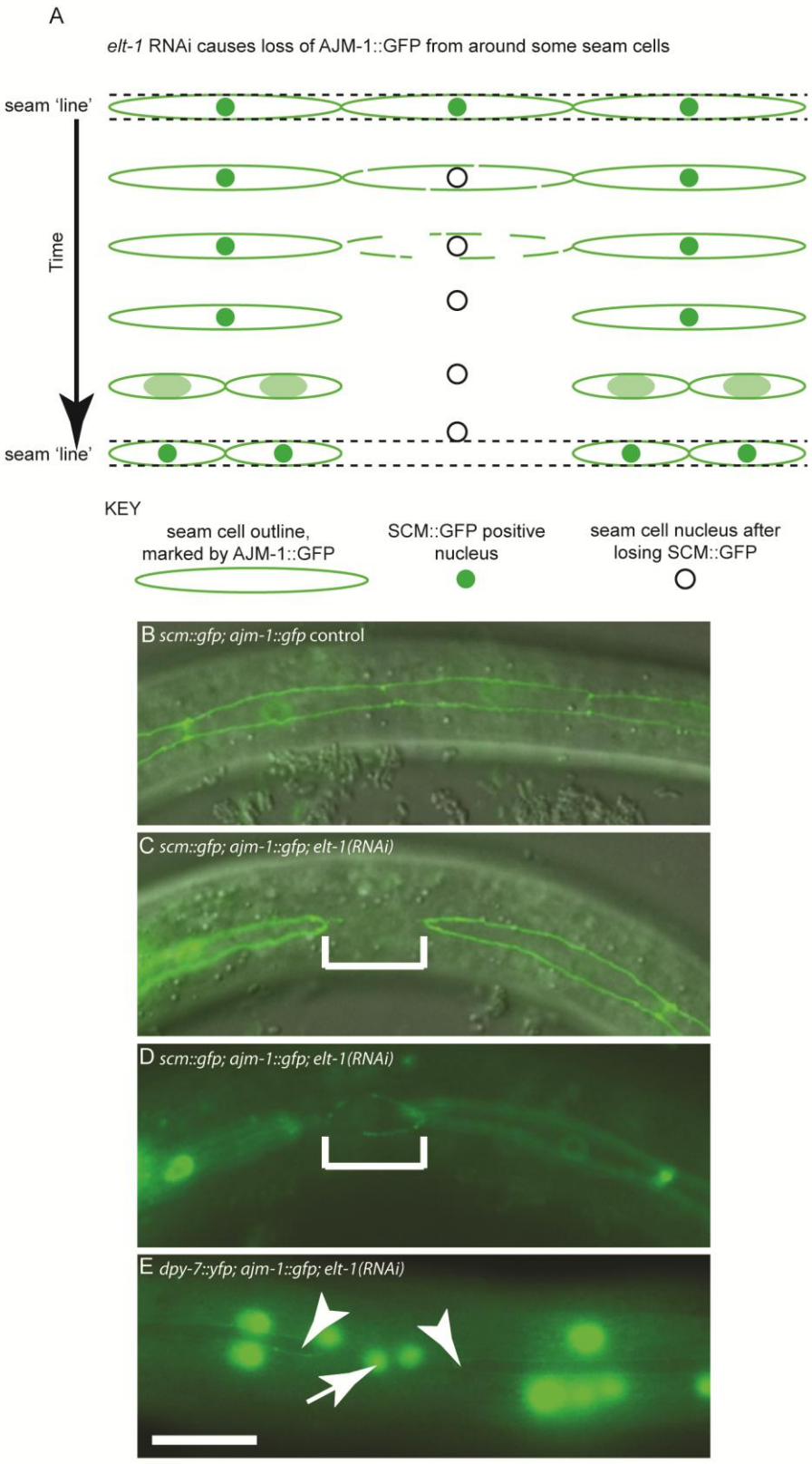


Figure 3.5 In addition to division defects, loss of seam fate was also observed between cell divisions follc *elt-1* knockdown. A. Schematic diagram of AJM-1 loss from one cell in the V lineage, involving the gradua AJM-1::GFP from around the cell until the cell is completely unbounded. SCM::GFP expression is always in such cells. Over a period of several hours, the cell becomes smaller and rounder and moves out of the the seam, into the hyp7 syncytium. Subsequently, these cells fail to divide, in contrast to cells that remain seam lineage and thus retain their stem fate. B. Wild-type seam lineage in animals expressing both *ajm-1* and *scm::gfp*, showing cells in a continuous line. C and D show the same cell in two different planes in an *elt-1(RNAi)* animal in which seam cells lose their AJM-1::GFP boundary (representative cell shown in bra region), changing shape and plane as they move into the hyp7 syncytium (change of shape and plane ob D). In order to quantify inappropriate fusion in these worms, animals with one or more breaks in the seam were scored. In *elt-1(RNAi)* animals 68% of animals had breaks in the seam (n=44). In control animals (e to HT115 bacteria containing the empty RNAi feeding vector *L4440*) 1.8% of animals had breaks (n=56). worms expressing both AJM-1::GFP and *dpy-7p::yfp*, it is possible to show that cells which lose their AJM boundary and move out of the seam lineage as a result of *elt-1 RNAi*, subsequently switch on *dpy-7* expr (indicated by white arrow), a marker of hypodermal fate. Adjacent seam cells which retain AJM-1::GFP e show no *dpy-7* expression (white arrowheads).

expression, which marks cells that have adopted a hypodermal fate but is not expressed in seam cells; cells bounded by AJM-1::GFP never expressed *dpy-7p::yfp*, even when they fail to undergo scheduled divisions (data not shown).

ELT-1 works independently from BRO-1 to prevent seam cells fusing inappropriately with the hypodermis

Unlike the situation in *bro-1* mutants, however, observation of *elt-1(RNAi)* animals carrying the integrated *ajm-1::gfp* construct revealed that some cells, at the same time as losing *scm::gfp* expression, lose AJM-1::GFP, used here as a marker of seam cell boundaries (data not shown and Figure 3.5A,B,C,D). This may be predictive of inappropriate fusion with the hyp7 syncytium. Cells were never observed that lost *ajm-1::gfp* but retained *scm::gfp*. The 'disintegration' of the AJM-1::GFP continues until only vestigial traces are evident between the two seam cells on either side of the affected cell, as shown in Figure 3.5D.

In the hours subsequent to losing *ajm-1::gfp* expression, seam cells were found to become rounder and move out of the line of the seam, with both the morphological and positional changes being indicative of the acquisition of hypodermal fate. To test this, a strain carrying the integrated transgenes *ajm-1::gfp* and *dpy-7::yfp* was used. In several cases, cells were observed during lineage analysis losing *ajm-1::gfp* and then acquiring *dpy-7::yfp* expression. Gaps in *ajm-1::gfp* expression in the seam were correlated with the presence of *dpy-7::yfp*-expressing nuclei which had not yet, or only partially, moved out of the line of the seam (Figure 3.5E). Thus, *elt-1* RNAi causes an additional phenotype not observed in *rnt-1* or *bro-1* mutants, whereby some seam cells differentiate inappropriately by fusing with the hypodermal syncytium.

Overall, the data presented here suggests three distinct phenotypes observed in *elt-1(RNAi)* animals. The first, in common with *rnt-1* and *bro-1* mutants, involves division failure in the absence of a permanent change in cell fate (i.e. not involving fusion with the hypodermis).

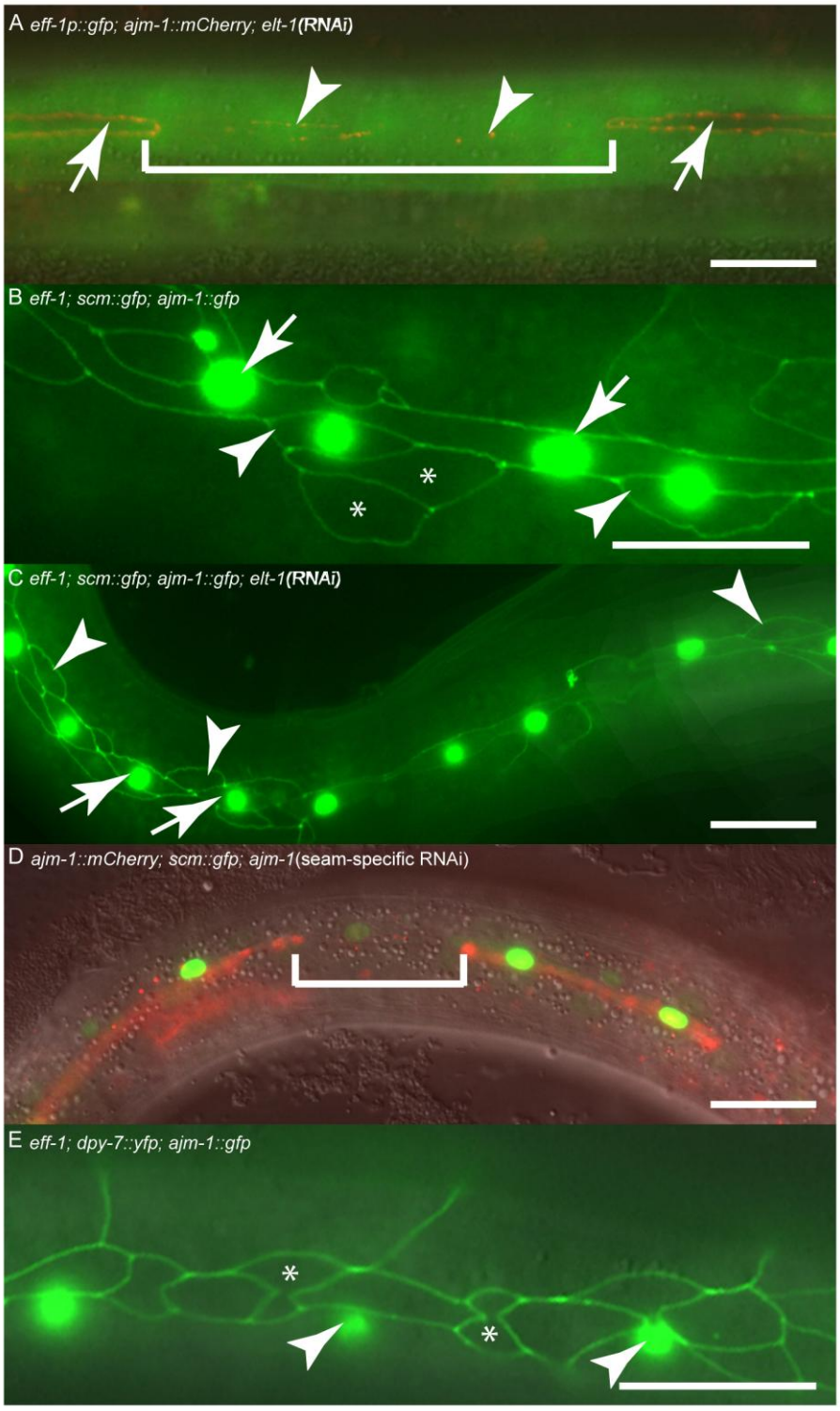


Figure 3.6 A. Hermaphrodite carrying *eff-1p::gfp* and *ajm-1::mCherry* transgenes, subjected to *elt-1 RNAi*. \
bracketed region a seam cell is in the process of losing *ajm-1* expression, and only has vestiges of AJM-1:: remaining (arrowheads). This cell is now expressing *eff-1p::gfp*. In contrast, cells retaining their AJM-1 bord express *eff-1p::gfp* (white arrows). B. *eff-1* mutant hermaphrodite expressing *scm::gfp* and *ajm-1::gfp*. *scm:* expression in the "true" seam is shown with white arrows. Anterior daughters of seam divisions sometimes expression of *scm::gfp* (white arrowheads), particularly where they remain in contact with the true seam (pr these are the most recent daughters). Other anterior daughters do not express *scm::gfp* but retain their AJM ary (asterisks). In order to quantify inappropriate fusion in these worms, animals with one or more breaks ir line were scored. In *eff-1* mutants 3% of animals had breaks in the seam (n=31). C. *eff-1* mutant hermaphr expressing *scm::gfp* and *ajm-1::gfp*, subjected to *elt-1* RNAi. In these animals, no AJM-1 breakdown is obs either around seam cells (white arrows) or around anterior daughters of asymmetric seam divisions (white a heads), demonstrating that the *elt-1* (RNAi) induced fusion of seam cells with hyp7 is suppressed in *eff-1* m *eff-1* mutants subjected to *elt-1* RNAi 0% of animals had breaks in the seam (n=20), compared with 68% in *elt-1(RNAi)* animals alone (n=44, Figure 5A,C,D). D. Seam-specific RNAi strain hermaphrodite subjected to *ajm-1*. In this experiment the RNAi deficient mutant *rde-1* was used, rescued with *rde-1* cDNA expressed u control of the seam-specific SCM promoter, as described in detail in Materials and Methods. The bracketed lost its AJM-1 border and the SCM::GFP has become much fainter. E. *eff-1* hermaphrodite carrying *dpy-7p: ajm-1::gfp* transgenes. Anterior daughters of seam divisions retain their AJM-1 borders but never express *d* (asterisks). *dpy-7* expression is only observed in the syncytial hypodermis surrounding these cells (white ar heads). Scale bar, 25µm.

Secondly, loss of SCM::GFP is also observed in *elt-1(RNAi)* animals (as well as in *rnt-1* and *bro-1* mutants), but this was found to be independent of division failure. Thirdly, in *elt-1(RNAi)* animals but not in *rnt-1/bro-1* mutants, inappropriate adoption of the hypodermal fate was observed, as shown by the acquisition of DPY-7::YFP, preceded by loss of *scm::gfp* expression and of AJM-1::GFP from the apical junctions. While these cells always lose SCM::GFP, fusion with the hypodermis is not always the consequence of SCM::GFP loss. However, AJM-1::GFP is very tightly linked to the seam fate; loss of AJM-1::GFP is always coupled with fusion to the hyp7 syncytium and acquisition of the hypodermal fate.

The *elt-1* (RNAi) fusion defect is dependent on EFF-1

If cells are lost from the seam in *elt-1(RNAi)* animals because of inappropriate fusion with the hyp7 syncytium, one would anticipate that ectopic *eff-1* expression would be evident, as the fusogen EFF-1 is known to be required for this fusion event (Mohler, Shemer et al. 2002). To test this, a transgenic strain carrying both an *eff-1* transcriptional reporter, which normally expresses in dorsal and ventral hypodermis but not in seam cells, and *ajm-1::mCherry* was used. In *elt-1(RNAi)* animals, ectopic *eff-1p::gfp* expression was observed in seam cells that had lost their *AJM-1* boundary (Figure 3.6A). Furthermore, *elt-1* RNAi-induced fusion of seam cells with the hyp7 syncytium was found to be suppressed in *eff-1* mutants (Figures 3.5A,C,D and 3.6C). Taken together, these data suggest that ELT-1 represses *eff-1* expression in seam cells in order to prevent fusion of these cells with the surrounding hypodermal syncytium, thus maintaining their distinct stem cell-like fate.

Seam cell membrane integrity, as marked by apical junctions, is required to maintain stem cell-like identity and sufficient to prevent differentiation

In order to address more directly the role of seam cell boundaries in determining the stem cell-like identity of seam cells, expression of components of apical junctions was knocked

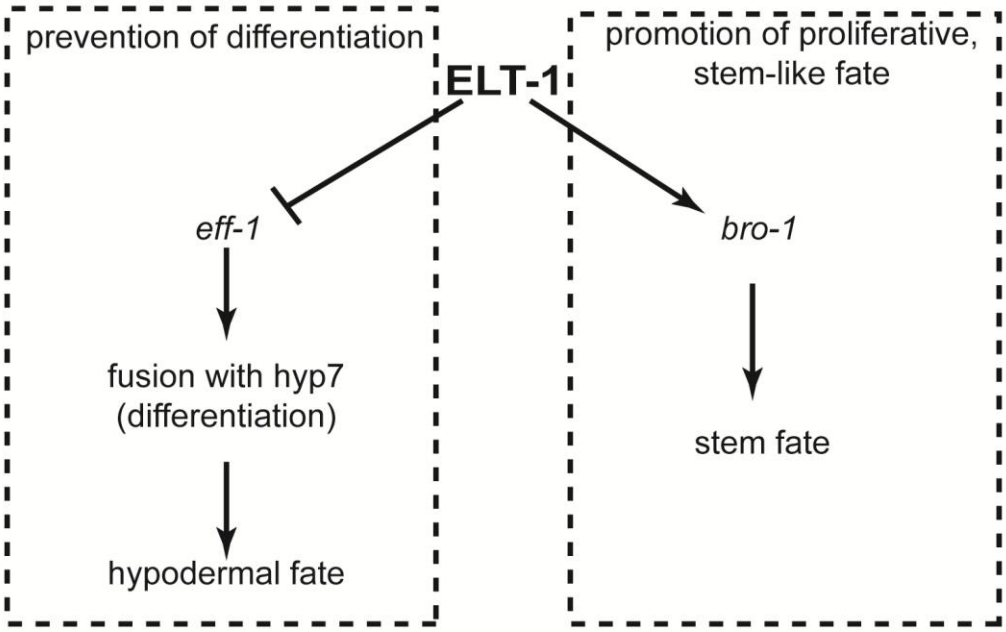


Figure 3.7 Model to illustrate the relationship between ELT-1, BRO-1 and EFF-1 in the maintenance of the seam stem cell-like fate. Direct transcriptional activation of *bro-1* by ELT-1 promotes proliferation, while repression of *eff-1* by ELT-1 (either directly or indirectly, for example through *elt-5/6*) represses fusion mediated by EFF-1, thereby preventing differentiation.

down by RNAi. Given the essential functions of these proteins in a variety of cell types throughout development, it was necessary to use a seam-specific RNAi approach (Qadota, Inoue et al. 2007). Knockdown of either *ajm-1*, *let-413* or *dlg-1* gave a similar phenotype, involving breaks in the AJM-1::GFP boundary, associated loss of *scm::gfp* expression and withdrawal from further division (Figure 3.6D; *let-413* and *dlg-1* data not shown). When *ajm-1* was knocked down by RNAi, 66% of animals displayed breaks in the seam boundary as marked by *ajm-1::mCherry* (n=50), whereas in controls not subjected to RNAi 5% of animals displayed breaks (n=36) (the odd break in the seam would be expected due to mosaicism of the *ajm-1::mCherry scm::gfp* array). Therefore, seam cell membrane integrity, as marked by apical junctions, is essential for the maintenance of seam stem cell-like fate. An important caveat to this experiment is that it is conceivable that, in the case of knock-down of apical junction components, cell integrity is perturbed, rather than 'stemness' *per se*. A useful follow-up experiment would be to follow affected cells and check that, in line with the prediction, they (or rather, their nuclei) continue to reside in the hyp7 syncytium; use of the *dpy-7::yfp* reporter would should also demonstrate hypodermis-specific gene expression.

Next, it was important to examine whether the presence of an intact boundary around the seam cells is sufficient to specify seam fate. This could happen in two ways: the boundary could either promote seam fate, or block differentiation signals from the surrounding environment. To test this, an *eff-1* mutant was used in which anterior daughters retain their AJM-1::GFP marked boundary and remain in contact with the seam line instead of moving into the hypodermis. Whether these anterior daughters continue to express *ajm-1::gfp*, or whether they are marked by existing protein which remains in the membrane and has a low rate of turnover, is unclear. The effect of the *eff-1* mutation is that ectopic AJM-1::GFP boundaries are present in this strain. To assess whether these cells inappropriately retain the seam fate, *scm::gfp* expression was monitored, as well as *dpy-7::yfp* expression. Firstly, it was found that ectopic AJM-1::GFP bordered cells do not express *scm::gfp* (Figure 3.6B), suggesting that these cells do not retain all aspects of seam fate. Perhaps this is not

surprising, given that these cells are the anterior daughters of asymmetric seam divisions and would have already received the instruction to withdraw from further proliferation. However, *dpy-7::yfp* expression was never observed in these cells (Figure 3.6E), indicating that differentiation is repressed. This suggests that fusion of seam cells with hyp7 (mediated by EFF-1) is essential as a trigger for differentiation in this context. These cells are therefore in developmental 'limbo', being neither completely seam nor hypodermis. Thus, although the anterior daughters of asymmetric seam divisions are destined to differentiate at division, the nature of the differentiation event is not specified at this stage and requires further inputs; it is not until cell fusion and membrane breakdown occurs, as marked by the loss of apical junction boundaries, that the hypodermal fate is adopted.

Overall, therefore, seam cell boundaries may not be sufficient to specify all aspects of seam fate, but they are sufficient to prevent differentiation. The data presented here therefore suggest that the seam stem cell-like fate is retained by the blocking of "hypodermalizing" signals in cells that are bounded by apical junctions. ELT-1, therefore, plays dual roles in specifying the stem cell-like properties of the seam cells; on the one hand activating proliferative potential *via bro-1*, and on the other preventing inappropriate differentiation through the repression of *eff-1* and consequent maintenance of seam boundaries (Figure 3.7).

Discussion

The GATA transcription factor *elt-1* directly regulates *bro-1* in the stem cell-like seam cells

C. elegans seam cells are an excellent model for stem cell-like modes of division, sitting as they do at the crossroads between the developmental decision to proliferate (thus renewing the pool of pluripotent precursors), or to pursue a differentiation pathway towards one or more specialised cell types. Comparative genomics was used, coupled with a yeast one-hybrid screen and promoter deletion analysis, to identify upstream regulators of *bro-1* expression; a gene known to be essential for seam cell proliferation. Surprisingly, sequence necessary and sufficient for the expression of *bro-1* was found to lie exclusively within an intron. This evidence adds strength to the concept that highly conserved non-coding sequences correlate with biologically important enhancer elements (Ruvinsky and Ruvkun 2003; Marri and Gupta 2009). However, within the CNE, the putative ELT-1 binding site identified here and implicated in *bro-1* regulation is neither completely conserved across multiple Caenorhabditis species, nor a perfect match to the characterised consensus binding site for GATA transcription factors, WGATAR (Evans, Reitman et al. 1988). Indeed, no sites have been identified within the CNE which are perfect matches to characterised *C. elegans* seam-expressed transcription factor consensus binding sites and completely conserved across five Caenorhabditis species.

The fact remains, however, that multiple lines of evidence suggest that ELT-1 binds to this intronic sequence; in a yeast one-hybrid screen using the CNE as bait, ELT-1 was identified in 92% of positive clones, whilst an EMSA band shift experiment not only confirmed this finding, but demonstrated that the putative GATA site towards the 3' end of the CNE (GATA site B) is essential for ELT-1 binding, an observation confirmed by rescue experiments.

This model is supported by the similarities between *elt-1* (RNAi) and *bro-1* mutant phenotypes in terms of failures of seam cell divisions, and loss of *scm::gfp* expression. It

therefore seems highly likely that *elt-1* and *bro-1* (as well as *rnt-1*) function within the same developmental pathway to promote seam cell proliferation.

The lack of conserved consensus sites for GATA transcription factors, taken together with the strong evidence in support of ELT-1 binding to the *bro-1* CNE, presents, *prima facie*, a contradiction. One way of addressing this apparent paradox is by seeing the interactions between transcription factors and their binding sites not as fixed and immutable but as dynamic and evolvable.

Cross-species experiments have demonstrated that cis-regulatory elements in the promoters of genes can change over relatively short time-scales, to the extent that even closely related species such as *C. elegans* and *C. briggsae* can evolve subtle but significantly different expression profiles of orthologous genes (Ruvinsky I. and Ruvkun G. 2003). However, just as important was the finding that, despite changes at the level of primary sequence within promoter regions, the output of certain genetic pathways remained the same, a fact which can be accounted for by invoking co-evolution of transcription factors with their changing binding sites. Thus, whilst conservation can point to important regions of sequence and provide clues as to the mechanisms by which genes are regulated, by only considering highly conserved regions of putative regulatory sequence as interesting one makes the false assumption that functional regions of the genome are fixed and unchanging.

However, whilst co-evolution of transcription factors and their binding sites may be possible, this mechanism is perhaps unlikely to account for all – or even most of – the evolutionary changes in transcription factor sites within the genome. If the interactions between transcription factors and the sites to which they bind are so precise and specific, it would presumably be necessary for all of the sites bound by a given transcription factor to undergo the same change; in other words, co-evolution would have to occur across thousands of target sites in parallel. Perhaps a more likely explanation, therefore, is that the interactions between transcription factors and their target sites are relatively weak, allowing flexibility

both in terms of binding and in terms of evolutionary change. This makes the evolution of new sites relatively easy – a perfect match is not required, so sites which are recognised by a given transcription factor are much more likely to be arise and provide material for selection – and thus could account for sites effectively moving with respect to the genes they affect.

It is interesting to note that the L1 (asymmetric) division is robust and appears unaffected in both *elt-1* RNAi and *bro-1/rnt-1* mutant worms. This division is most likely regulated by a separate pathway. The involvement of both the CBFβ/RUNX complex and a GATA transcription factor in the regulation of the proliferation of the stem cell-like seam cells is reminiscent of the situation in *Drosophila* and in mammals, where GATA and RUNX factors work together to regulate blood cell formation (Waltzer, Ferjoux et al. 2003). Indeed, mammalian *Runx1* is known to be transcriptionally activated by Gata2 (Nottingham, Jarratt et al. 2007). Thus, the direct regulatory relationship between ELT-1 and *bro-1* suggests yet another mode of interaction between members of these two gene families.

elt-1 has *bro-1*-independent functions

elt-1(RNAi) worms display a striking phenotype which is not observed in *bro-1* mutant animals, however. This involves loss of the apical cell junction marker *ajm-1::gfp* from some seam cells, followed by movement of the cell out of the seam and subsequent differentiation into the hypodermal fate. It has previously been argued that seam cell phenotypes (loss of *scm::gfp* positive cells and gaps in *ajm-1::gfp* expression) in *elt-1(RNAi)* worms do not result from inappropriate fusion of the seam with the hypodermis, on the basis that *scm::gfp*-expressing cells are never observed in the hypodermal syncytium (Smith, McGarr et al. 2005). However, here it is argued that cells do indeed fuse with hyp7 in *elt-1(RNAi)* animals, but first lose their *scm::gfp* and AJM-1 boundary, as well as changing their morphology as they undergo the transition from seam to hypodermal fate and begin to express hypodermal

markers. Thus, the breaks in AJM-1::GFP likely result not from degeneration of the seam cells but are the result of a transition between two cell fates; the proliferative seam fate, which is associated with the AJM-1::GFP boundary, and the differentiated syncytial hyp7 fate, which is not.

elt-1 functions upstream of the fusogen *eff-1*

In terms of its *bro-1*-independent role in the seam, *elt-1* appears to function upstream of EFF-1. The role of the fusogen *eff-1* is critical in the seam (Mohler, Shemer et al. 2002) and is responsible for promoting the fusion of hypodermal seam daughters with the hyp7 syncytium by causing the formation of 'fusion pores' in the adjacent membranes (Mohler, Simske et al. 1998). Indeed, *eff-1* over-expression has been shown to result in inappropriate fusion of seam cells with hyp7 (Shemer, Suissa et al. 2004). This phenotype is strikingly similar to that seen in *elt-1(RNAi)* animals, and suggests that ELT-1 acts to repress *eff-1* in the seam, thereby preventing seam cells from fusing with hyp7. The finding that ectopic *eff-1* expression is observed in *elt-1(RNAi)* animals confirms this.

Inappropriate fusion of the seam cells with the hypodermis has been reported previously; this phenomenon has been observed in embryos and newly hatched animals in which *elt-5* and *elt-6* (which act redundantly) had been knocked down by RNAi (Koh and Rothman 2001). Moreover, similar to the observations described here for *elt-1(RNAi)* animals, fusion in the case of *elt-5/elt-6* RNAi was accompanied by gradual dissolution of the AJM-1::GFP boundary around the seam cells (Koh and Rothman 2001). It therefore seems likely that there is a network of GATA factors acting to prevent inappropriate differentiation of seam cells throughout development. In addition, other transcription factors have been shown to regulate seam cell development, for example NHR-25, BAF-1 and CEH-16 as well as heterochronic regulators like LIN-14 and LIN-29 (Ambros and Horvitz 1984; Chen, Eastburn et al. 2004; Cassata, Shemer et al. 2005; Silhankova, Jindra et al. 2005; Margalit, Neufeld et

al. 2007; Huang, Tian et al. 2009). The interactions between all these genes will be an interesting area for future study.

Seam cells and the stem cell niche

The apparent close relationship between the presence of intact apical junctions and the stem cell-like properties of seam cells is suggestive of similarities with stem cells in *Drosophila*. In the testes and ovaries of *Drosophila,* germline stem cells (GSCs) are retained in what has been termed a niche; the niche concept, introduced over 30 years ago (Schofield 1978), describes how stem cells can be maintained in a proliferative state by signals from a microenvironment, consisting of cells and the extracellular components they produce. The niche is essential for the stem properties of such cells and, as these cells move out of the niche, so they lose these properties in favour of differentiated cell fates. In this way, far from merely creating an inert environment in which the *Drosophila* GSCs reside, the cells around these stem cells provide cues which regulate the maintenance of the stem cell pool, physically anchor the GSCs to the niche, and even control the polarity of the stem cells, determining the positions of the daughters of GSC divisions relative to one another and to the niche. The mechanisms underlying the complex interactions between GSCs and their niche microenvironment involves extracellular signalling (Matunis, Tran et al. 1997; Xie and Spradling 1998; Xie and Spradling 2000; Kiger, Jones et al. 2001; Kai and Spradling 2003) as well as physical adhesion of stem cells to the niche (Song, Zhu et al. 2002), with the DE-cadherin and Armadillo/β-catenin apical junction complex being both important in recruiting GSCs to the niche and required for the maintenance of the stem cell pool. Loss of either of these proteins results in dramatic depletion of GSCs from the niche. Here, *C. elegans* apical junction proteins are shown to be required to maintain the undifferentiated stem cell-like fate of the seam cells.

Perhaps there is something analogous to a seam stem cell “niche” in *C. elegans*, in the sense that cell contacts (marked by, and dependent on, apical junction proteins) are required to maintain a microenvironment in which the seam cells are prevented from differentiation. The importance of cellular contacts for seam development has previously been recognized. For example, the developmental fate of the V5 seam cell has been shown to be dependent on correct seam cell contacts either side (Austin and Kenyon 1994), and proliferation of the seam cells has been shown to be perturbed when contacts between seam cells are not properly re-established following division (Silhankova, Jindra et al. 2005). Thus, both the proliferation and differentiation of the seam cells has been shown to be dependent on signals from the surrounding microenvironment, raising the question of whether they do in fact reside in a niche. In support of the niche concept, this work demonstrates that cells that have withdrawn from the proliferation programme as a result of asymmetric division, but which have failed to fuse with the hyp7 syncytium (as a result of *eff-1* mutation), do not express markers of differentiation like *dpy-7*. In other words, the boundary provides protection from differentiation signals. However, this notion of a niche has to remain speculative in the absence of defining the nature of these signals.

Intriguingly, the mechanism of regulation discussed here, in which ELT-1 represses *eff-1* in order to prevent fusion and differentiation of cells that have the proliferative fate, mirrors the situation in vulval precursor cells (VPCs). In the developing vulva, the 6 VPCs P3.p-P8.p are prevented from fusing with the hyp7 syncytium in early larval stages (in L3 this exclusion from hyp7 is limited to P5.p-P7.p) (Clark, Chisholm et al. 1993). Fusion with hyp7 acts to limit the developmental potential of Pn.p cells (as it does with anterior seam daughters) and is restricted to those that flank the developing vulva (Clark, Chisholm et al. 1993). LIN-39 acts to prevent this fusion by repressing *eff-1* in P3.p-P8.p during L1 and in P5.p-P7.p during L3 (Maloof and Kenyon 1998; Ch'ng and Kenyon 1999; Gleason, Korswagen et al. 2002). In *eff-1* mutants, unfused VPCs fail to differentiate into the hypodermal fate, retaining their AJM-1 boundary and at least some aspects of the vulval fate. These cells fail to proliferate,

however, suggesting that other signals are required for the induction of normal vulval development. This is analogous to the situation with unfused seam cells in *eff-1* mutants, which also fail to divide once they leave the seam line, remaining in developmental “limbo”. In both cases, however, fusion with hyp7 and associated breakdown of cell boundaries is required for cells to take on the differentiated hypodermal fate. In both the seam and the developing vulva, therefore, only those cells that are prevented from fusing with hyp7 (thereby retaining their boundaries) are protected from differentiation and retain further developmental and proliferative potential.

Conclusion

Overall, a model is presented (Figure 3.7) in which the GATA factor ELT-1 plays important dual roles in maintaining the 'stemness' of the seam cells, by both promoting the proliferative fate and preventing differentiation. Firstly, ELT-1 acts directly through *bro-1* to promote proliferation and self-renewal of the seam. Secondly, ELT-1 is essential for maintaining the integrity of the seam cell compartment, as marked by apical junctions. In fulfilling this latter role, ELT-1 works through EFF-1. When *eff-1* is repressed, the boundaries around the seam cells are maintained, and thus differentiation is prevented. Apical junction components themselves are also shown to be important for maintaining seam cell fate, though this does not necessarily mean that ELT-1 acts directly on apical junction components. Indeed, as has been previously suggested, apical junction breakdown could be a relatively late event in the cell fusion process (Mohler, Simske et al. 1998; Gattegno, Mittal et al. 2007). Taken together, these data suggest that the seam cells reside in a microenvironment in which they are protected from differentiation by the boundary that separates them from the hyp7 syncytium. Thus, the seam microenvironment may satisfy the criteria of a niche in certain respects, protecting seam cells from influences that would otherwise trigger differentiation. The GATA factor ELT-1 works through *bro-1* to promote seam cell proliferation and through *eff-1* to maintain seam cells in the undifferentiated state.

CHAPTER 4

The Hox cofactor *unc-62* regulates division symmetry and functions redundantly with *bro-1/rnt-1* in embryonic seam development

Introduction

Identification of *unc-62* as a potential regulator of *bro-1*

Identification of the *bro-1* CNE (see Chapter 3 and Figure 4.1A) provided an opportunity to discover the identity of transcription factors which are responsible for driving *bro-1* expression; transcription factor sites within the CNE, particularly those which are well conserved, provide clues to candidate regulators.

Within the *bro-1* CNE are two adjacent 6bp regions which are particularly well conserved across three Caenorhabditis species (Figure 4.1B). Both regions are close matches to the consensus binding site for the MEIS family of transcription factors (TGACAG) (Chang, Jacobs et al. 1997). Thus, this potential interaction was examined further.

Meis proteins and their partners in development

Meis proteins are of fundamental importance from the early stages of animal development and are intimately involved in early patterning processes. Their importance stems at least in part from interactions with Hox genes. Whilst Hox genes are central to the process of laying-out the animal body plan, they bind DNA with relatively low specificity (Hoey and Levine

1988; Ekker, Jackson et al. 1994; Mann 1995). Specificity is, however, essential for their role as transcriptional regulators. This apparent paradox is resolved by the fact that additional homeodomain proteins, Hox cofactors, form complexes with Hox genes and enhance their specificity for target genes (Kurant, Pai et al. 1998; Mann and Affolter 1998).

A subset of these Hox cofactors is the TALE (Three Amino acid Loop Extension) class of homeodomain proteins. This group is further divided into PBC and Meis classes; the former are represented in vertebrates by Pbx proteins and in *Drosophila* by Extradenticle, grouped together because of their conserved PBC-A and PBC-B domains which lie towards the N-terminal end of the protein relative to the homeodomain (Bürglin 1997; Mukherjee and Bürglin 2007).

The PBC-B domain of PBC class Hox cofactors is thought to be responsible for controlling nuclear import and export, based on the phosphorylation state of residues within the domain (Kilstrup-Nielsen, Alessio et al. 2003). Independently, the PBC-A domain facilitates interactions with another class of Hox cofactors, the MEIS proteins (Chang, Jacobs et al. 1997; Knoepfler, Calvo et al. 1997). Present in *Drosophila* as Homothorax (Hth) and vertebrates as Meis/Prep, this class of TALE proteins is united by the presence of a homeodomain and two ‘Hth/Meis’ (HM) domains, which, like the PBC-A domain of PBC proteins, facilitate interactions with binding partners (in this case, PBC proteins)(Bürglin 1997; Rieckhof, Casares et al. 1997; Bürglin 1998; Ryoo, Marty et al. 1999; Mukherjee and Bürglin 2007).

The modes of action of the members of these two groups of TALE proteins appear diverse. PBC and Meis proteins can form both dimeric and trimeric complexes with Hox proteins; they can operate with or independently of HOX proteins and they can bind to DNA directly or merely serve as stabilising elements in transcriptional complexes, without binding to the DNA being targeted (Jiang, Shi et al. 2009). Furthermore, there is mutual regulation between PBC and Meis members. For example, Meis proteins influence the stability and subcellular

localisation of PBC proteins (Kurant, Pai et al. 1998; Pai, Kuo et al. 1998; Abu-Shaar, Ryoo et al. 1999; Affolter, Marty et al. 1999; Ferretti, Schulz et al. 1999; Stevens and Mann 2007), whilst, in Drosophila, Exd levels have been demonstrated to affect Hth expression (Kurant, Pai et al. 1998; Nagao, Endo et al. 2000).

Meis proteins and their partners in *C. elegans*

In *C. elegans*, both Meis and PBC classes of TALE domain proteins are represented. Meis genes are present in the form of *unc-62* and *psa-3*, the latter lacking a homeodomain (Bürglin 1997; Van Auken, Weaver et al. 2002; Arata, Kouike et al. 2006; Mukherjee and Bürglin 2007), whilst the closest homologues to Drosophila Hth are *ceh-20* and the two more diverged genes *ceh-40* and *ceh-60* (Bürglin 1997; Bürglin 1998; Van Auken, Weaver et al. 2002; Mukherjee and Bürglin 2007).

Overviews of the layouts of *unc-62* and *ceh-20* are given in Figure 4.2. The HM domain of *unc-62* is encoded by exons 2,3,4 and 5, whilst the homeodomain is at the C-terminal end of the protein, encoded by exons 7, 8 and 9 (Figure 4.2A). Both exons 1 and 7 have alternative splice variants, reportedly resulting in four alternative transcripts with combinations of exons 1a, 1b, 7a and 7b (Van Auken, Weaver et al. 2002). Although even the longest *unc-62* cDNA is relatively small (just under 1.7kb), the whole genomic ORF is spread across almost 14kb.

In *ceh-20*, the situation is very different; there are no splice variants and the entire ORF lies within 2kb of sequence. As shown in Figure 4.2B, the PBC-A and PBC-B domains are N-terminal to the homeodomain.

As in mammals, there is evidence for interaction between the *C. elegans* Meis and PBC homologues, but findings concerning interactions with Hox genes have been contradictory. In some cases, the evidence has argued against such genetic interactions (Van Auken,

Weaver et al. 2002) whilst a number of studies have identified situations where Hox genes and Meis/PBC class genes act cooperatively to regulate gene expression (Liu and Fire 2000; Koh, Peyrot et al. 2002; Arata, Kouike et al. 2006)

There are, however, several well-established Hox-independent roles for these genes in development (Van Auken, Weaver et al. 2002; Yang, Sym et al. 2005). *unc*-62, *ceh-20* and *ceh-40* have important – and potentially overlapping – roles in early development. Certain *unc-62* mutants exhibit embryonic and larval lethality, suggestive of functions at multiple developmental stages. Similarly, whilst there is a degree of redundancy between *ceh-20* and *ceh-40*, loss of function of both genes results in high levels of embryonic lethality which can be further increased by the presence of *unc-62* mutations (Van Auken, Weaver et al. 2002). In addition, *unc-62* and *ceh-20* mutants exhibit defects in cell migration (Yang, Sym et al. 2005); single mutants for both genes exhibit migration failures, whereby the Q neuroblasts, which usually migrate anteriorwards in the case of the QR lineage, or posteriorwards in the case of the QL lineage, only partially complete their journey and are found to lie in between their point of origin and their wild-type destination (Yang, Sym et al. 2005).

unc-62 and *ceh-20* are also required for vulva formation, playing several different roles during the development process. Both genes are required for fusion of the cells anterior and posterior to the VPCs with the hyp7 syncytium (in a *lin-39*-independent manner), for vulval versus non-vulval fate specification decisions and for correct morphogenesis of the vulval lineage descendants (Yang, Sym et al. 2005).

The role of *unc-62* in the *C. elegans* seam cells

Little is known about any role for *C. elegans* Meis and PBC homologues in the seam, though there is some evidence of these genes functioning in this lineage. In *unc-62* mutants, Vn.a cells – the anterior daughters of the L1 asymmetric division, which differentiate and adopt

hypodermal fates in wild-type animals – occasionally divide, the incidence of this rising in *ceh-20 (mu290); unc-62 (mu232)* double mutants, suggesting that both genes have a role to play in preventing supernumerary divisions in these cells (Yang, Sym et al. 2005).

A second line of evidence for a seam role for members of these gene classes, discussed in more detail below, comes from studies of the T lineage, where *psa-3* expression is regulated by NOB-1 and CEH-20, determining the symmetry of the T cell division (Arata, Kouike et al. 2006). Significantly, however, apart from these roles, no other major functions in the seam have been identified.

Results

unc-62(RNAi) animals exhibit seam hyperplasia

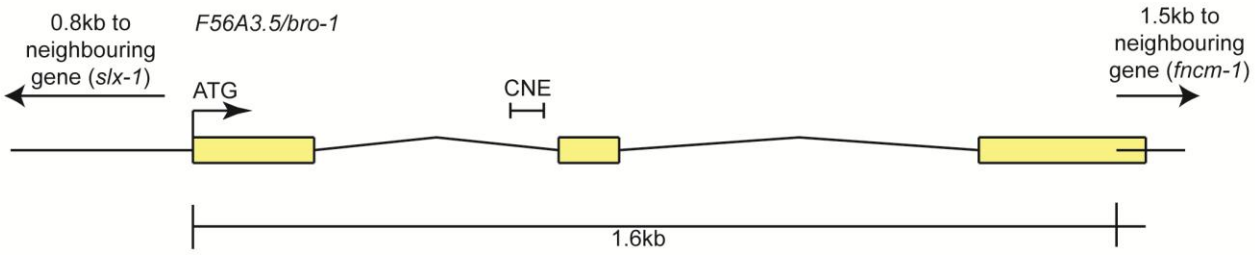
Identification of putative *unc-62* binding sites within the *bro-1* CNE suggested a role for *unc-62* in the seam, specifically in the context of *bro-1* regulation. To determine whether *unc-62* did indeed function in the seam, an *unc-62* RNAi feeding clone was constructed (see Materials and Methods) and worms carrying the *scm::gfp* reporter were subjected to *unc-62* knockdown. Compared to control worms, *unc-62(RNAi)* worms had elevated numbers of seam cells (19.8±0.5 versus 15.8±0.1, Figure 4.3A and B).

unc-62 lesions cause seam hyperplasia

Since RNAi-mediated knockdown of *unc-62* expression yielded increased numbers of seam cells, further investigations were performed using *unc-62* mutants, which were crossed with a strain carrying the integrated *scm::gfp* reporter. Three alleles were obtained: *ku234* (MU1092) kindly provided by W. Wood), which carries a point mutation in the start codon of the 1b transcript, changing the ATG (Met) to ATA (Ile) (Van Auken, Weaver et al. 2002); *mu232* (a kind gift of C. Kenyon), which carries a different point mutation in the same start codon, that of the 1b transcript, changing the ATG to AAG (Lys) (Yang, Sym et al. 2005); and *e644*, a nonsense mutation which introduces a stop codon (TAG, in this case) to the Trp410 residue in exon 7b (Van Auken, Weaver et al. 2002).

In all alleles, seam cell numbers were elevated. *e644* mutants (CB644) have an average of 17.5 (±0.2) seam cells, significantly higher than wild type (Figure 4.3Ai, ii and B). *ku234* and *mu232* mutants showed even more severe hyperplasia, with 36.8(±0.8) seam cells on any given side. Phenotypically, these two alleles appear indistinguishable, likely reflecting the similar nature of the molecular lesions. It is noticeable that, whilst the seam line is

A



B

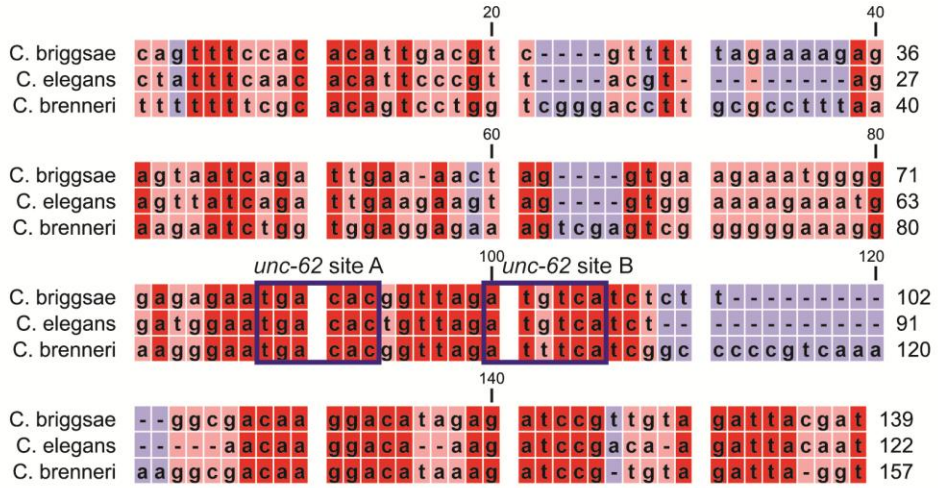


Figure 4.1 A. Genomic structure of *bro-1*, showing the size and position of the CNE. B. Alignment of the *C. elegans bro-1* CNE with orthologues from *C. briggsae* and *C. brenneri* reveals high levels of sequence conservation at the 3' end of intron 1. Conservation is still evident, though less pronounced, when *C. remanei* and *C. japonica* are included in the alignment (data not shown). Predicted *unc-62* binding sites are labelled. Sequence alignments were performed using CLC Sequence Viewer v6.4 (CLC bio, Aarhus, Denmark). Red shading denotes that sequences are conserved across all 3 species of nematode; pink denotes conservation in 2 out of 3 species; blue denotes no conservation.

perturbed along the length of the body, the hyperplasia in *ku234* and *mu292* mutants is concentrated in the anterior. *e644* mutants, in contrast, show no such bias (Figure 4.3Aiii).

The seam phenotype exhibited by *unc-62(ku234)* mutants is very similar to that seen in *ceh-20(mu290)* mutants and worms subjected to *ceh-20* RNAi (a gene also under investigation in the lab following a genome-wide screen for genes affecting seam cell number; RNAi performed and assayed by S. Hughes, pers. comm.). Pronounced seam hyperplasia is also evident, though the average number of seam cells is significantly lower than in *unc-62(ku234)* animals (Figure 4.4C). Following identification of *ceh-20* as a gene involved in regulation of seam cell divisions, several mutant alleles were obtained. *ceh-20(mu290)* carries a single G→A transition which changes a conserved arginine to a histidine (Figure 4.2B, Yang, Sym et al. 2005). Significantly, this arginine lies within the homeodomain and is thought to form part of a nuclear localisation signal (Yang, Sym et al. 2005). A missense mutation is present in *ceh-20(ay9)*, resulting in a methionine within the PBC-A domain being converted into isoleucine (Takacs-Vellai, Vellai et al. 2007). The integrated *scm::gfp* reporter construct was crossed into both mutant backgrounds. Interestingly, *mu290* is the only *ceh-20* allele found to have a seam hyperplasia phenotype; *ay9* shows no such expansion of the seam lineage (S. Hughes, pers. comm.). Given the known interaction between these genes and the similar mutant phenotypes of the *unc-62(ku234)* and *ceh-20(mu290)* alleles with respect to the seam cells, it was inferred that *ceh-20* and *unc-62* likely work together in the seam.

psa-3 is not required for H- and V-lineage seam development

Owing to the apparent involvement of *unc-62* in the seam, the possibility of a role for the only other *C. elegans* Meis homologue, *psa-3*, was investigated. Previously, it has been established that *psa-3* is only expressed in the T cell (Arata, Kouike et al. 2006). Thus, given the hyperplasia phenotype evident throughout the seam lineage in *unc-62* and *ceh-20*

mutant and RNAi worms, particularly in the anterior regions, it seems unlikely that *psa-3* is a major player. To confirm this, *psa-3* RNAi was performed, using a feeding clone from the Ahringer library. No change in seam cell numbers were observed, relative to control worms exposed to RNAi feeding bacteria carrying the empty feeding vector ($p>0.9$) (Figure 4.3B).

The *unc-62* seam phenotype involves perturbation of cell division symmetry

To investigate the developmental basis of the *unc-62* hyperplasia phenotype, lineage analysis was performed on *ku234* mutants carrying the *scm::gfp* reporter. The results are summarised in Figure 4.4A.

The nature of the seam phenotype depends on the region of the worm in question. In the head, where the hyperplasia is most pronounced, additional seam cells deriving from H1 and H2 first appear at the L1 asymmetric division. In *ku234* mutants, this division is symmetrised to the seam fate, resulting in the production of two seam cells as opposed to one seam cell and one daughter which subsequently differentiates and adopts the hypodermal fate. The seam identity of these cells was established by the fact that they continued to express *scm::gfp* up until the next, L2, division (whereas non-seam daughters usually lose their GFP within one or two hours of the L1 division) and by the fact that both daughters were able to divide again in L2; non-seam, hypodermal daughters are unable to divide. Significantly, the next division occurred as normal at the L1-L2 moult. Furthermore, the anterior daughter of the H2 L1 division did not divide several hours before the L2 division, as is normal, but instead waited to divide simultaneously with H2.p. Thus, the timing of divisions *per se* is not perturbed in *unc-62(ku234)* mutants; rather, cells adopt the incorrect fate as a result of perturbation of cell division symmetry and adjust the timing of their divisions accordingly.

The result of this perturbation of the symmetry of seam divisions is that, in affected lineages of the seam, the number of cells with the stem fate doubles at every division. Conversely,

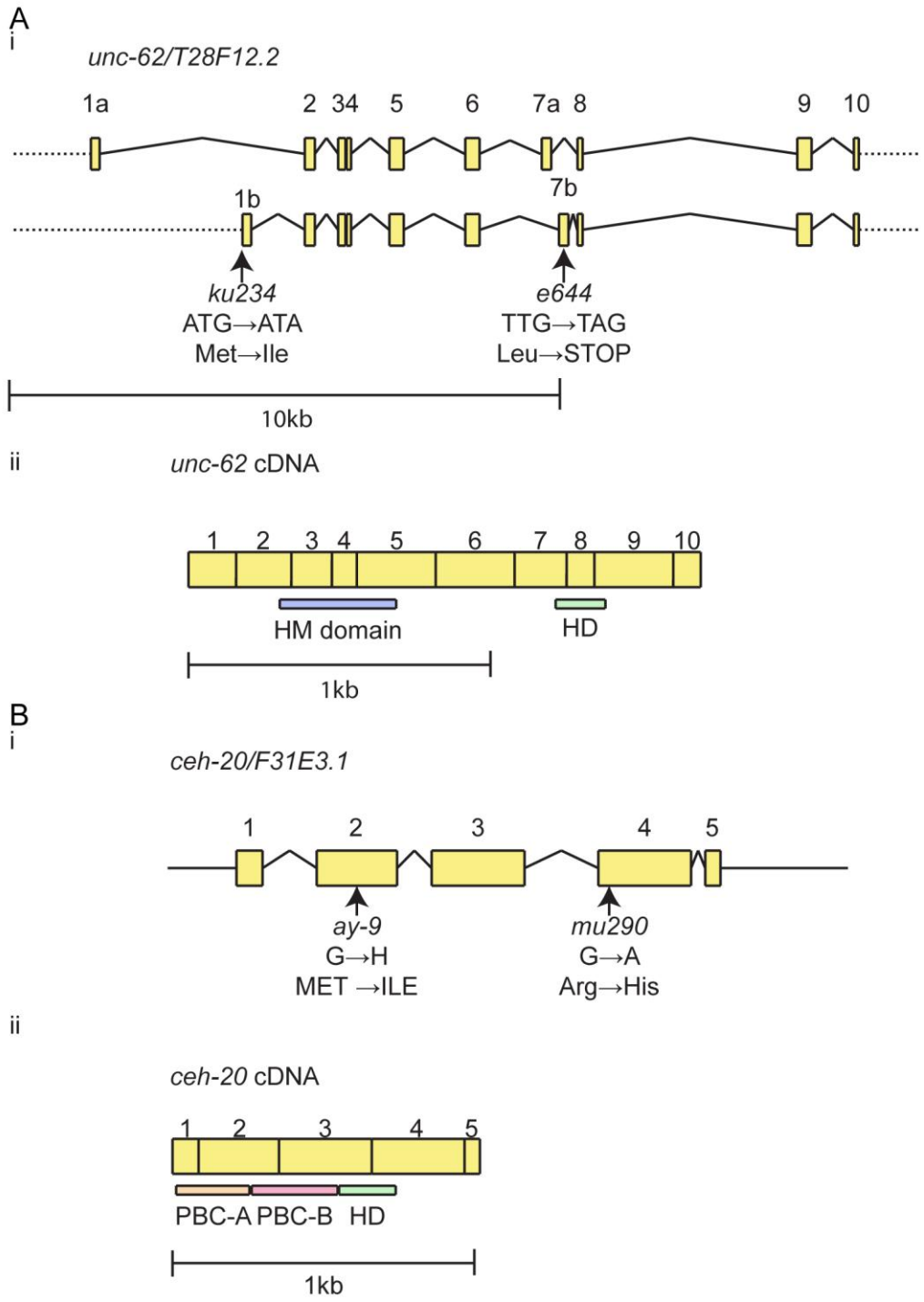


Figure 4.2 A. Overview of *unc-62* ORF (Ai) and protein domains (Aii). The positions of the 1a, 1b, 7a and 7b alternative exons are shown, together with the position and nature of the *ku234* and *e644* mutations. HM domain = Hth/Meis interaction domain. HD = homeodomain.

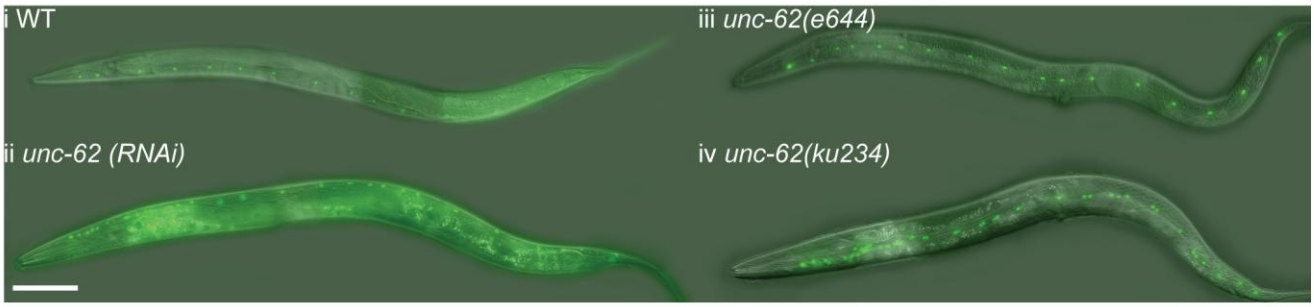
B. Layout of the *ceh-20* ORF (Bi) and protein domains (Bii). The *ceh-20* ORF and cDNA are drawn to the same scale. The location and molecular details of the *ay9* and *mu290* mutations are shown.

this rapid and abnormal expansion of the seam occurs at the cost of the hypodermal fate; lineage analysis revealed that cells which adopt the hyp7 fate during normal development instead become seam cells in *unc-62(ku234)* mutants (Figure 4.4A). This latter point is also illustrated by use of an integrated *dpy-7p::yfp* reporter (a kind gift of Iva Greenwald), which expresses only in differentiated hypodermal cells and not in the seam. In *unc-62(ku234)* mutants, there are significantly fewer nuclei expressing *dpy-7::yfp* compared to wild type worms (Figure 4.4B).

Whilst the hyperplasia of *unc-62* mutants is most apparent in the head, the more posterior seam lineages (V1-6) are also affected, though in a different way. Here again, division symmetry is perturbed. However, whereas in the head the *unc-62(ku234)* mutation symmetrises all seam divisions, in the V lineages the seam fate is not always favoured. In these divisions the situation seen in the head lineages is often reversed; the L2 division, which is symmetric in wild type worms, is often transformed into an asymmetric division, yielding only one, instead of two, daughter cells with the stem fate. These division defects were variable between cells, between divisions at different stages within the same lineage, and between worms. The result is an increase in the range of seam cell numbers and irregularity of seam cell positioning in adult worms: with respect to seam cell numbers, the apparently stochastic symmetry defects in the body region result in some worms having more seam cells than normal, and some having fewer; with respect to the regular positioning of the seam cells, the same stochasticity means that whilst one lineage may produce more cells with the stem fate than normal, others may produce fewer, the result being groups of seam cells produced from ‘proliferative’ lineages, and gaps in the seam corresponding to ‘differentiative’ lineages (Figure 4.3Aiv).

In addition, confirming previous reports (Yang, Sym et al. 2005), L1 V-lineage divisions were often aberrant, with the anterior daughter of the division appearing not to reacquire its characteristic ‘eye’ shape after division, instead remaining relatively small and round, but

A



B

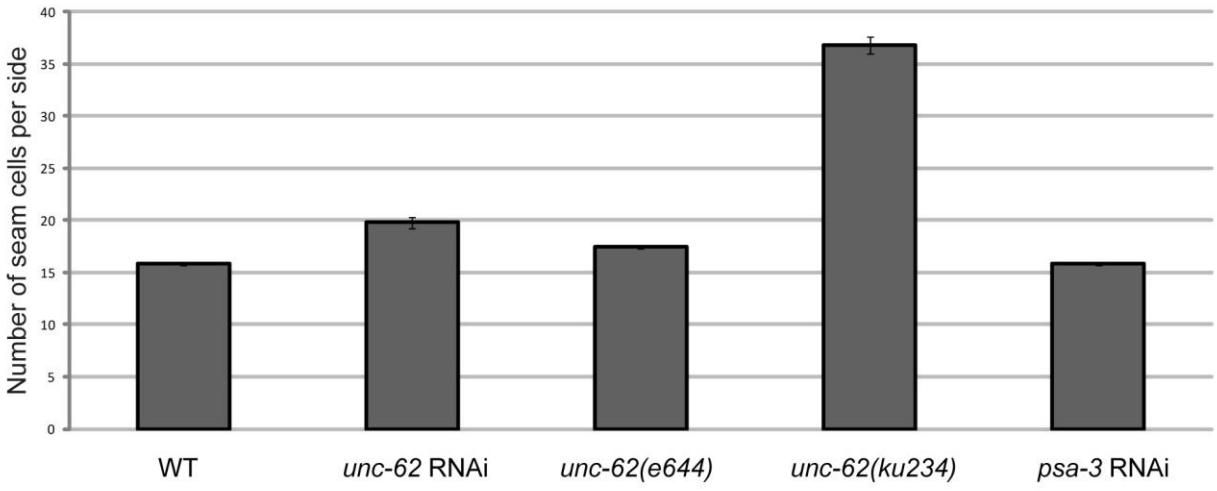


Figure 4.3 A. Seam cell phenotypes in wild-type and *unc-62* mutant and RNAi worms. (Ai) Wild type hermaphrodite with 16 seam cells, hermaphrodite exposed to *unc-62* RNAi, exhibiting seam hyperplasia. (Aii) *unc-62(RNAi)* adults have more seam cells than control worms. (Aiii) *unc-62(e644)* animal, with 18 seam cells. (Aiv) *unc-62(ku234)* animal, displaying pronounced hyperplasia, particularly in the anterior region of the worm. Scale bar, 100µm. B. Graph showing seam cell numbers in wild type (15.8±0.1), *unc-62(RNAi)* (19.8±0.5), *unc-62(e644)* (17.5±0.12), *unc-62(ku234)* (36.8±0.8) and *psa-3*(RNAi) (15.8±0.1) worms, counted using the seam cell marker, SCM::GFP. Error bars show SEM.

nevertheless dividing again and, in contrast to previous reports, often continuing to retain expression of the seam cell specific marker, *scm::gfp*.

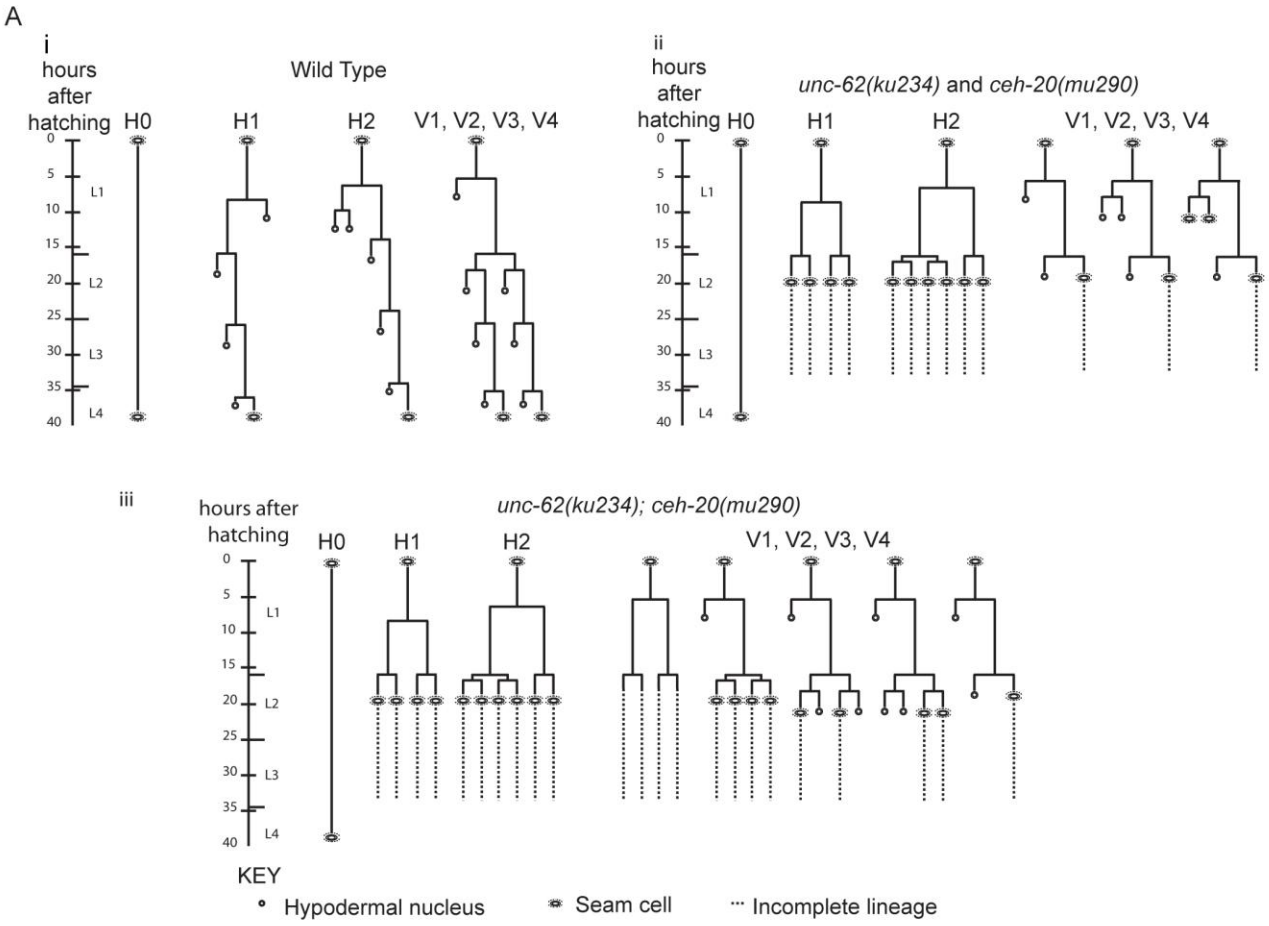
The *ceh-20* seam phenotype also involves perturbation of cell division symmetry

The superficially similar phenotypes of *unc-62* and *ceh-20* mutants (and worms subjected to RNAi knockdown of both genes) are suggestive of these genes performing similar functions in the seam. To determine whether the developmental basis for the phenotypes is the same, lineage analysis was performed on *ceh-20* RNAi animals. Exactly the same phenotypes were observed; hyperplasia in the head occurred due to symmetrisation of all divisions towards the seam fate, whilst, in the V lineages, division symmetry was perturbed but in a way which did not always favour the seam fate, resulting in irregular spacing of the resulting seam cells and variation in seam cell number (Figure 4.4A).

unc-62; ceh-20 double mutants exhibit more severe phenotypes

Given the similar phenotypes of *unc-62* and *ceh-20* animals, the genetic interaction between these genes was investigated by the creation of double mutants, *unc-62(ku234); ceh-20(mu290)*. Worsening of the seam hyperplasia would be suggestive either of redundancy or of the two alleles being hypomorphic, with their combination weakening their function further. Conversely, no worsening of the phenotypes of the single mutants would be indicative of the genes functioning non-redundantly in the same pathway. The double mutants created exhibited manifestations of both of these scenarios (Figure 4.4Aiii and C).

Overall, seam cell numbers of *unc-62(ku234); ceh-20(mu290)* were higher than those of the single mutants (52.1 (±2.0) seam cells per side, Figure 4.4C). However, the increase in seam cell number was specifically due to increased numbers of seam cells in the body region of the worm, posterior to the head.



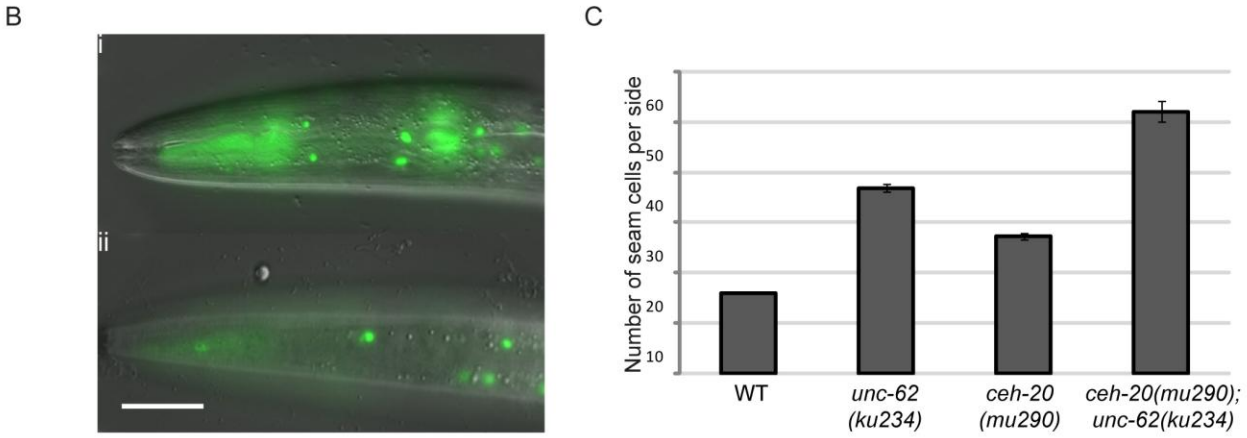


Figure 4.4 A. Symmetrisation defects are prevalent in *unc-62(ku234)* and *ceh-20(mu290)* mutants. Ai. Lineage trace of wild type hermaphrodites, from hatching to L4, in H0, H1, H2 and V1-4 seam lineages. Aii. Lineage trace showing representative division defects in *unc-62(ku234)* and *ceh-20(mu290)* single mutants; the most penetrant aspect of the phenotype is observed in the head, where symmetrisation of divisions to the seam fate results in hyperplasia of the seam lineage. Divisions in V1-4 are often unaffected, but defects are nevertheless observed. In the examples shown, the L2 division has been transformed from a symmetric division to an asymmetric division, resulting in only one cell adopting the seam fate and, in the second trace, the anterior daughter of the first L1 division has divided again. Aiii. In *unc-62(ku234); ceh-20(mu290)* animals, in addition to the hyperplasia in the H1 and H2 lineages the V lineages also give rise to supernumary seam cells, following symmetrisation of divisions towards the seam fate. It is worth noting that considerable variation was evident and it is likely that the traces here do not represent an exhaustive description of the abnormalities in *unc-62(ku234); ceh-20(mu290)* animals. Further lineaging would likely allow any biases (e.g. with respect to the aberrant symmetrisation of the divisions of anterior versus posterior daughters of L2 symmetric divisions) to be revealed. B. Wild type (Bi) and *unc-62(ku234)* (Bii) worms, showing that the latter exhibit reduced numbers of hypodermal nuclei (marked by *dpy-7::yfp*) in the head. Symmetrisation of seam divisions to the seam fate comes at the cost of daughters destined to adopt the differentiated, hypodermal fate. Scale bar = 50µm. C. Graph showing seam cell numbers in wild type, *unc-62(ku234), ceh-20(mu290)* and *unc-62(ku234);ceh-20(mu290)* worms, counted using the seam cell marker, *scm::gfp*. Error bars show SEM.

In the head, the extent of hyperplasia was identical to that observed in either single mutant. Given that both single mutants display symmetrisation of all head divisions (i.e. divisions of the H1 and H2 lineages), it is not surprising that the extent of hyperplasia was not greater in the double mutant; there are simply no more opportunities to increase seam cell number without the invocation of additional divisions. Lineage analysis was performed to confirm that additional divisions do not occur (Figure 4.4Aiii).

In the rest of the body, however, the situation is very different. In the double mutant, divisions are very often symmetrised to the seam fate, hence the increase in seam cell number in this strain. Less frequently, a range of other symmetry defects were observed (Figure 4.4Aiii).

In no cases were extra rounds of divisions observed; thus, neither single allele nor the combination of both alleles can affect entry into the cell cycle. Instead, they function, apparently redundantly in the body but non-redundantly in the head, to regulate the symmetry of cell divisions.

Hyperplasia in *unc-62(ku234)* mutants is dependent on *rnt-1* and *bro-1*

The seam hyperplasia exhibited by *unc-62(ku234)* is much more pronounced than that seen in worms which have been exposed to *unc-62* RNAi feeding bacteria. To determine whether this phenotype could be suppressed by the absence of functional *rnt-1* and *bro-1*, a triple mutant, *rnt-1; bro-1; unc-62* strain – which also carried the *scm::gfp* reporter – was generated. Strikingly, not only was the hyperplasia completely suppressed, but seam cell numbers fall below those seen in *rnt-1; bro-1* double mutants, with these worms having an average of 10.1 (±0.56) seam cells per side (Figure 4.5A and B).

In *rnt-1* and *bro-1* mutants, seam cell numbers are reduced due to division failures and division symmetry defects (Nimmo, Antebi et al. 2005; Kagoshima, Nimmo et al. 2007). In

unc-62(ku234) mutants, the hyperplasia in the seam lineage results from symmetrisation of seam divisions. The triple mutant strain, *rnt-1; bro-1; unc-62*, exhibits both of these defects; division failures and symmetry defects are evident (see Figure 4.5C).

unc-62, together with *rnt-1/bro-1*, is rate-limiting for seam proliferation

The dependence of the *unc-62* seam phenotype on the presence of functional RNT-1 and BRO-1 raised the question of whether *unc-62* is also rate-limiting for seam proliferation. To investigate this, the *unc-62(ku234)* allele was crossed into AW257, which carries multiple integrated copies of BRO-1::GFP and RNT-1::GFP and has elevated seam cell numbers as a result (Kagoshima, Nimmo et al. 2007). The resultant strain, AW665, exhibited even higher levels of hyperplasia, with an average of 65.2 (±2.75) seam cells per side and 87 cells counted in one seam lineage.

unc-62 and *bro-1/rnt-1* have redundant roles in embryonic seam development

In order to understand the relationship between *bro-1/rnt-1* and *unc-62*, it was necessary to establish why the seam cell number is lower in the triple mutant than in either the *unc-62* single mutant, or in the *bro-1; rnt-1* double mutants. Examination of triple mutant L1 larvae just after hatching revealed that seam cell number is significantly less than 10 (8 ±0.17, Figure 4.6A and B). Wild type, *unc-62(ku234)* and *bro-1; rnt-1* L1s are all born with 10 seam cells. Since the L1s examined were not old enough to have undergone their first larval divisions in the seam lineage, this finding points to an embryonic defect. Whether the reduction in seam cell number is due to a failure to specify cells in the seam lineage, or due to inappropriate loss of seam cells after normal specification, remains to be determined. The increase in severity of seam phenotype in the triple mutants, compared to the single mutant

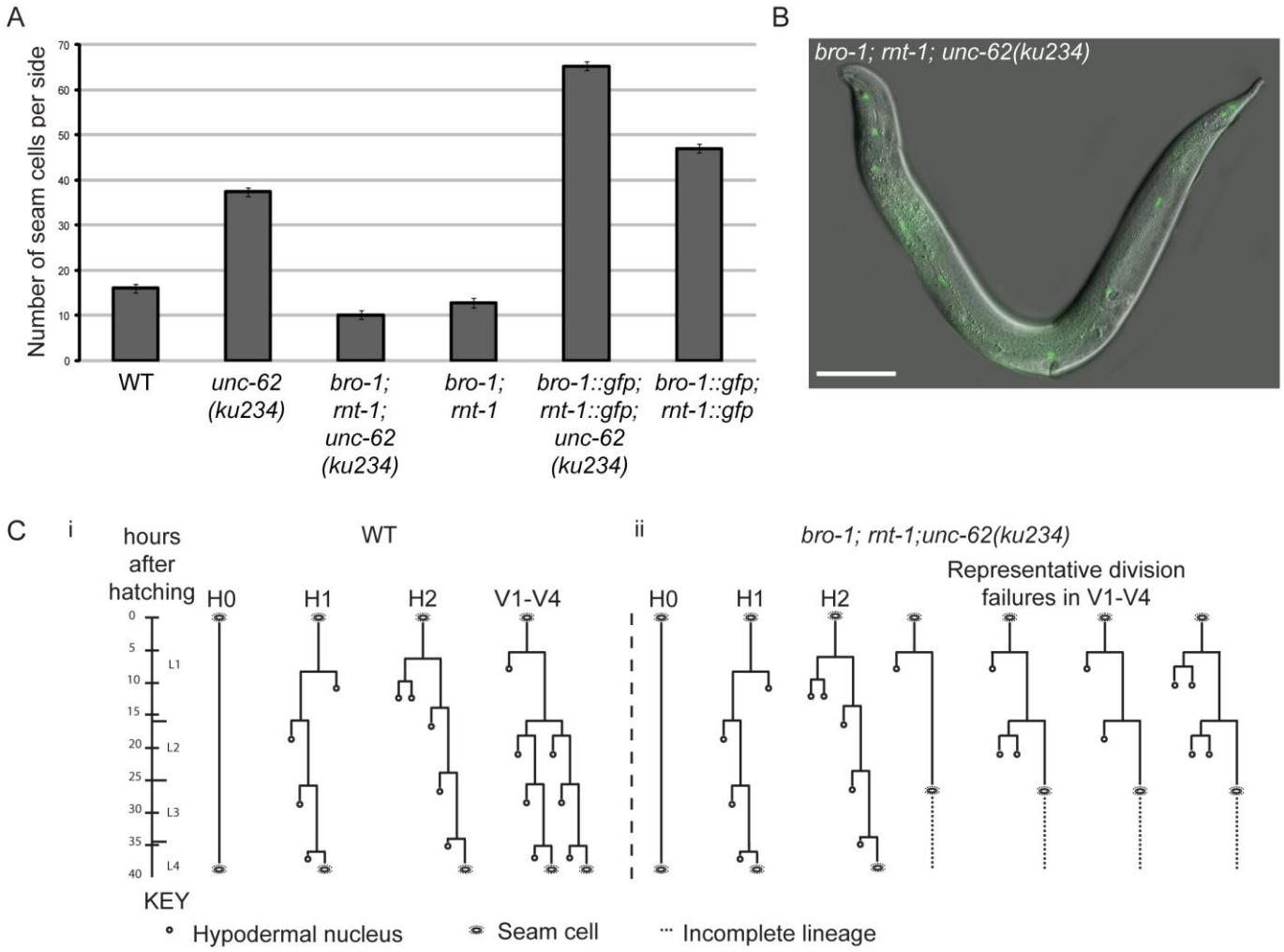


Figure 4.5 A. The seam hyperplasia evident in *unc-62(ku234)* mutants is dependent on *rnt-1/bro-1*. In *bro-1; rnt-1; unc-62(ku234)* triple mutants, no hyperplasia is observed. Furthermore, the average number of seam cells in these animals is lower than in *rnt-1;bro-1* double mutants owing to embryonic development defects. *unc-62*, together with *rnt-1/bro-1*, is also shown to be rate-limiting for seam cell number; the presence of the *unc-62(ku234)* allele in a background in which *rnt-1* and *bro-1* are over-expressed further increases the average number of seam cells per side. B. *bro-1; rnt-1; unc-62(ku234)* adult hermaphrodite with 10 seam cells. Scale bar = 100µm. C. Lineage traces showing divisions in wild-type (Ci) and *bro-1; rnt-1; unc-62(ku234)* mutant animals (Cii).

phenotypes, suggests a degree of redundancy between the functions of *unc-62* and *rnt-1/bro-1*, given that the *rnt-1* and *bro-1* mutants used are presumed null alleles (Kagoshima, Sawa et al. 2005).

In addition to this embryonic seam phenotype, the triple mutant strain is extremely sick, with high levels of embryonic and larval lethality (33.7% of embryos fail to hatch and, of those that do, 81.2% of the resultant larvae die before adulthood, Figure 4.6C). Qualitatively, *ceh-20* RNAi had the same effect as the *unc-62* background on *bro-1/rnt-1* mutants, suggesting that *ceh-20/unc-62* act redundantly with *bro-1/rnt-1* in embryonic and early larval stages.

unc-62 likely does not directly regulate *bro-1* transcription

To examine whether normal *bro-1* expression is dependent on the putative UNC-62 binding sites within the CNE, point mutations were introduced to key bases in both of the predicted binding sites. The resulting sequence was then cloned into the *pes-10* minimal promoter vector, using the same restriction sites as had been used to make the wild type CNE transcriptional reporter. Thus, the resultant plasmid was identical to the wild type version, except for the mutations introduced to the putative UNC-62 binding sites. Examination of transgenic lines carrying this construct revealed normal seam expression (Figure 4.7A), suggesting that UNC-62 does not directly regulate *bro-1* expression.

unc-62 is expressed uniformly in seam cells, throughout development

Given the apparent involvement of *unc-62* in seam development, the expression pattern of *unc-62* was investigated, using a full length genomic reporter construct (gratefully received from Scott Cameron). The reporter includes around 8kb of sequence upstream of the 1a start codon, in which various enhancer elements are known to lie, and should therefore provide an accurate view of the endogenous expression of the gene. As *unc-62* is expressed

in many different tissues and at low levels (Potts, Wang et al. 2009), the integrated *ajm-1::gfp* reporter was crossed into the *unc-62::cfp* reporter strain, allowing the boundaries of the seam cells to be visualised. As shown in Figure 4.7B , *unc-62* expression is clearly visible within the AJM-1::GFP-demarcated seam cells.

However, given that lineage analysis of *ku234* mutants suggest involvement of *unc-62* in division symmetry, it was important to establish whether there is an element of spatial control of expression which could account for the phenotype observed. However, no differences were detected between UNC-62::CFP expression or localisation in anterior and posterior daughters seam daughters at any stage during the division process.

Similarly, no temporal changes in expression level were observed. Expression of the reporter remained low throughout the development of the worm, present from the L1 stage onwards and throughout adulthood.

Consistent with the idea that *unc-62* and *ceh-20* work together, the latter gene is also expressed in the seam (Potts, Wang et al. 2009). Furthermore, mutations in *unc-62* affect CEH-20 nuclear localisation. In wild type worms, CEH-20::GFP is predominantly nuclear, yet introduction of the CEH-20 reporter into an *unc-62(ku234)* background resulted in a loss of this bias – nuclear levels of CEH-20 were found to be dramatically decreased (S. Hughes, pers. comm. and (Potts, Wang et al. 2009)).

Seam expression of *unc-62* is driven by a region 3' of the 1a transcript start site

Understanding the physical manifestations of *unc-62* lesions on seam development is without doubt important for a more general understanding of *unc-62* function, but it is only one facet of this gene's biology. The utility of dissecting a gene's mode and mechanism of action is most fully realised when the information gleaned is put in context – the context of other genes, upstream and downstream, which form the network surrounding *unc-62*.

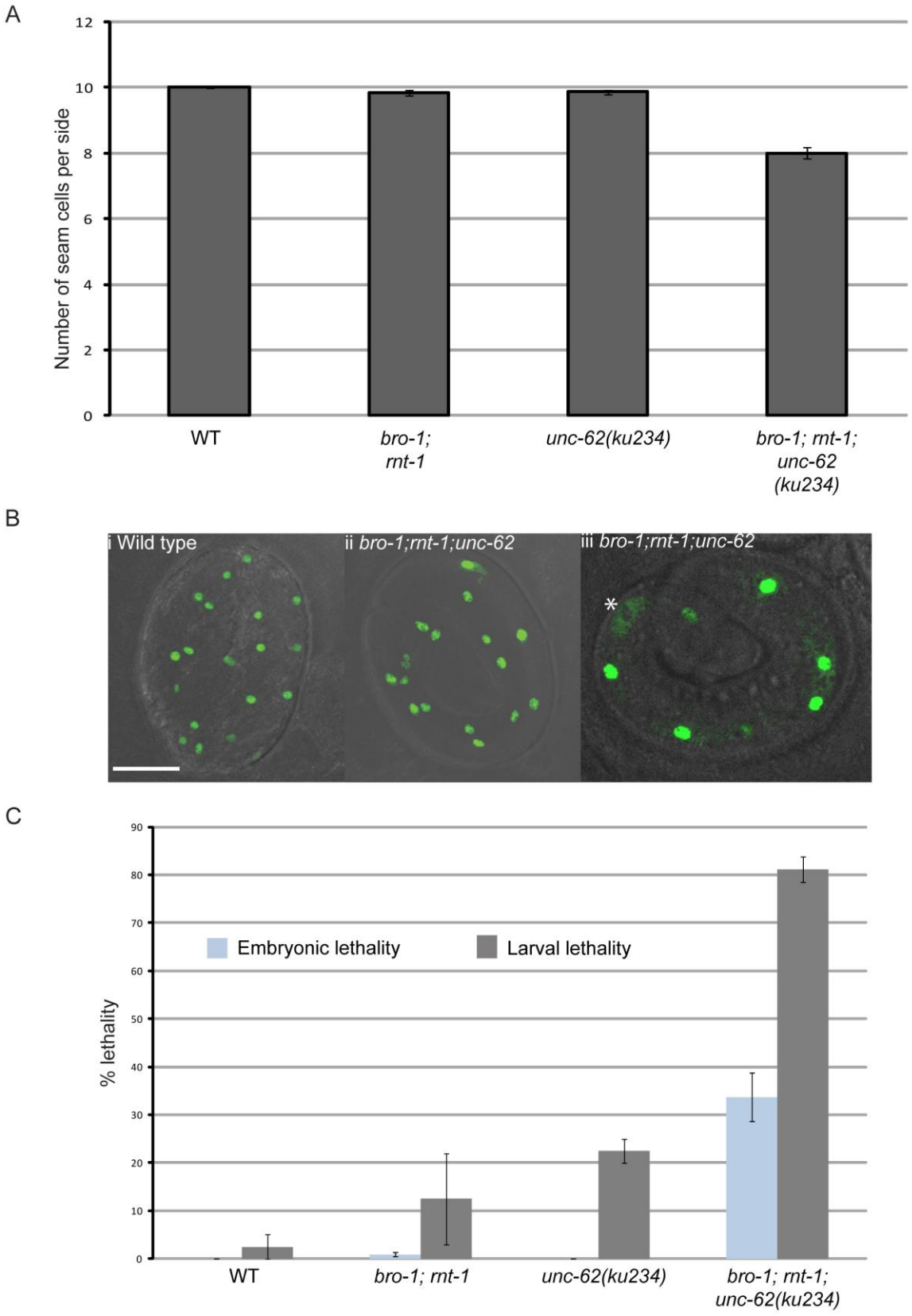


Figure 4.6 A. Seam cell numbers in wild type, *unc-62(ku234)*, *bro-1;rnt-1* and *unc-62(ku234);bro-1;rnt-1* mutants, counted at L1 just after hatching, before any seam divisions have taken place. B. Unlike wild-type embryos (Bi), triple mutant embryos often have fewer than 10 seam cells (In Bii, 8 seam cells are present on one side of the worm, and 9 on the other). The cause of the reduced seam cell number is unclear, but in the triple mutant cells are sometimes observed to be in the process of degradation (marked by asterisk in Biii). Scale bar = 10µm. C. Graph showing embryonic lethality (blue bars) and larval lethality of those larvae which surived embryonic development (grey bars) in wild-type, *bro-1; rnt-1* double mutants, *unc-62(ku234)* mutants and *bro-1; rnt-1; unc-62(ku234)* triple mutants.

In the case of upstream genes, interactions between these regulators and their targets can be discovered by examining the promoters and enhancer elements of these targets. With this in mind, a search for the regions of the *unc-62* sequence which are responsible for seam expression was performed; this gene is expressed in multiple tissues throughout development (Van Auken, Weaver et al. 2002; Yang, Sym et al. 2005; Potts, Wang et al. 2009) but here it is its expression in the seam which is of most interest. The rescuing *unc-62p::unc-62::cfp* construct contains all the promoter and enhancer elements required for endogenous expression, including that required for seam expression. Therefore, the regions of sequence included in this reporter were used as a basis for further analysis.

Initially, two constructs were made, pAW603 and pAW604, consisting of the 4kb directly upstream of the 1a start codon, and the 4kb of sequence 5' to this region, respectively. No seam expression was evident. To ensure that this result was not due to disruption of transcription factor binding sites at the mid-point of the 8kb covered by the two (non-overlapping) constructs, the entire 8kb promoter region was amplified and cloned into the same GFP vector. Here again, whilst expression was observed in other cells, no seam expression was discernable.

The search for seam enhancers of *unc-62* was then focussed on sequence 3' to the conventional promoter of the 1a construct. As an aid, alignment of *unc-62* sequences from *C. elegans, C. briggsae, C. remanei, C. brenneri* and *C. japonica* were performed. In general, exonic sequences showed higher levels of conservation than intronic sequences, but the region between the 1a and 1b transcript start sites could be seen to be particularly well conserved for non-exonic sequence (see Appendix). The region is referred to henceforth as the '1b promoter'. When this region of sequence, just over 2kb in length, was cloned upstream of *gfp*, fluorescence was observed in the seam, but also in muscle and hypodermal nuclei (Figure 4.7C). The expression pattern driven by this region may be complicated by the fact that it is out of its normal, genomic context, but nevertheless suggests that this region plays an important role in directing *unc-62* expression.

unc-62 could act as a repressor or an activator of *rnt-1/bro-1*

One interpretation of the *unc-62* phenotypes and their dependence on *rnt-1/bro-1* function is that UNC-62 acts as a repressor of these genes. Assuming that *unc-62* mutants are loss of function alleles (whether complete or incomplete), the hyperplasia observed could occur due to de-repression of *rnt-1* and/or *bro-1*. This mode of action for UNC-62 will be termed the 'repressor hypothesis'; according to this scenario, the hyperplasia in *unc-62* mutants results from the absence of a UNC-62-mediated repression.

However, there is a possible alternative mechanistic basis for the *unc-62* phenotypes observed. In mammals, Meis1, Meis2 and Prep1 have been found to be capable of acting as activators of transcription; in the cases of Meis1 and Meis2, the proteins possess an activation domain towards the C-terminus of the protein (Huang, Rastegar et al. 2005; Hyman-Walsh, Bjerke et al. 2010). Moreover, this activation domain is subject to autoinhibitory control from the HM domain of the same protein; the interaction with PBX proteins at the HM domain is essential for the regulation of this autoinhibition, the strength of which depends upon the presence or absence of PBX bound to the HM region of Meis2. This paradigm of Meis action, in conjunction with PBX interactions, is the basis for an alternative interpretation of the action of UNC-62 in the worm, the 'activator hypothesis'. According to this explanation, the seam hyperplasia of *unc-62* mutants occurs due to the presence of mutant UNC-62 protein which, as a result of the lesion, acts as a constitutive activator of downstream targets.

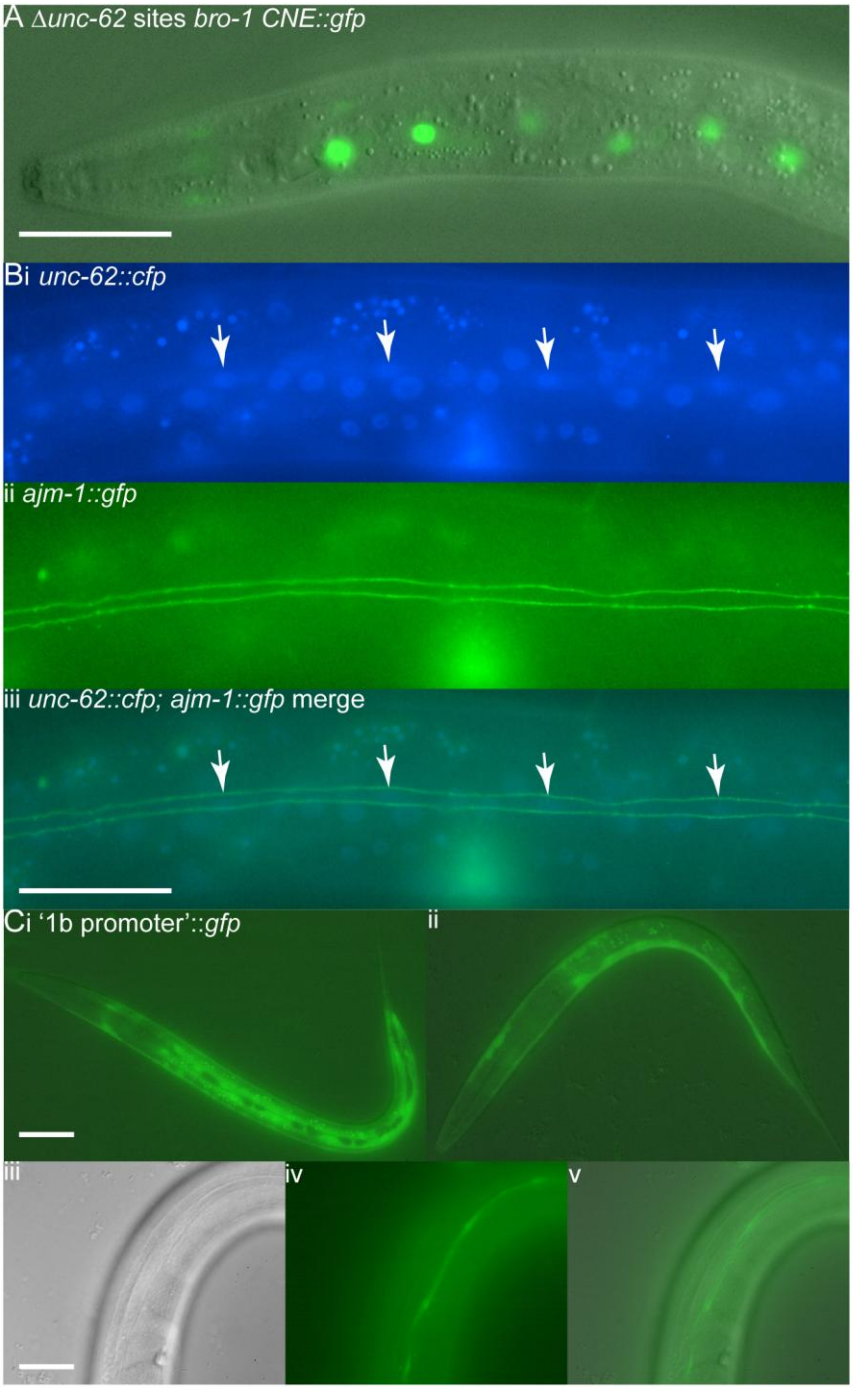


Figure 4.7 A. The introduction of point mutations at key bases of the putative *unc-62* binding sites within the *bro-1* CNE does not abolish seam expression.
B. *unc-62* is expressed in the seam cells (marked with arrows), as can be seen when AJM-1::GFP, which marks the seam cell boundaries, is overlaid with *unc-62::cfp* expression. Bi, CFP only, Bii, AJM-1::GFP only, Biii, merged.
C. A region of sequence lying between the 1a and 1b transcript start sites drives hypodermal (Ci), muscle (Cii) and seam expression (iii, iv and v, showing DIC, GFP and merged images respectively). Scale bars = 100µm.

unc-62 possesses a putative activation domain at its C-terminus

To investigate whether the *C. elegans* Meis homologue UNC-62 contains an activation domain, the protein sequence was aligned with the sequences from other nematode species in addition to those of mammalian Meis proteins. Across mammals and nematodes, both the HM and HD domains show high levels of conservation. Between the two phyla, there was no significant conservation towards the C-terminal domains of the proteins. However, within the worm group, this region of the protein contained a highly conserved region of around 110 amino acids (Figure 4.9A), suggesting that, in addition to the two domain types already identified (HM and HD), there may be an additional functional domain at the C-terminus of the protein, perhaps with an analogous, if not homologous, function to that of the equivalent domain in Meis2 and Prep1.

The *unc-62(ku234)* allele may encode a constitutively active protein

Further insights into the 'activator hypothesis' can be gleaned from the nature of the lesions in the *unc-62* alleles used in this study. In both *ku234* and *mu232* alleles, the start codon of the 1b transcript, which starts with the HM domain, is mutated. As a result, translation has been assumed to begin at a subsequent start codon, C-terminal to the normal start codon and the protein produced lacks the HM domain (Van Auken, Weaver et al. 2002); this was confirmed in this study (see below). The 1a or 'long' transcript is unaffected as the mutant start site is spliced out of the processed transcript.

With respect to the 1b transcript, therefore, the HM domain is absent in *ku234* mutants, whilst the rest of the protein, including the putative activation domain, is present. Thus, any autoinhibitory effect of the N-terminal region of the protein will be relieved, allowing the resultant protein to activate genes. In this case, *rnt-1* and *bro-1*, could be downstream targets of this activation.

The *ceh-20(mu290)* phenotype is consistent with the activator hypothesis

The autoinhibition model (Hyman-Walsh, Bjerke et al. 2010) postulates that interaction with PBX is required for the functioning of the autoinhibition domain within Meis proteins. Thus, mutations in the *C. elegans* PBX homologue, *ceh-20*, which affect the ability of CEH-20 to bind to the HM domain of UNC-62, would be predicted to have the same effect as the *ku234* mutation; prevention of autoinhibition and activation of downstream targets.

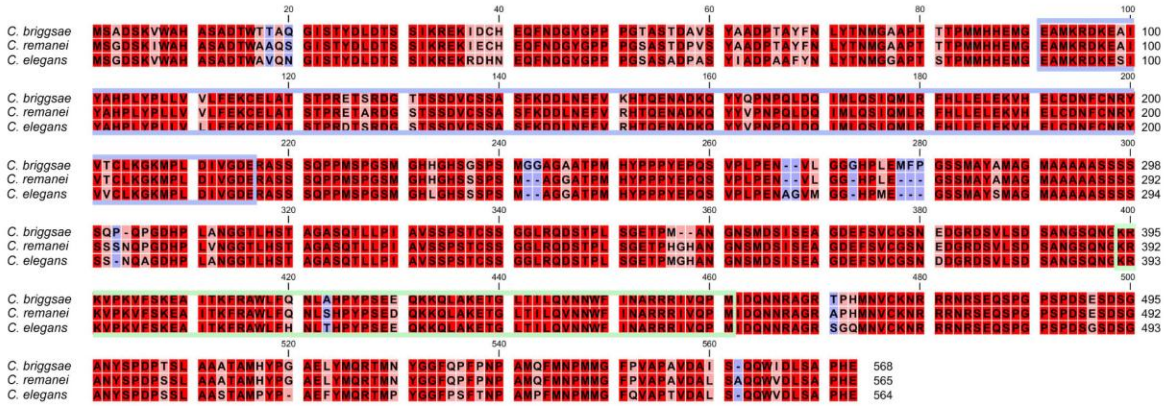
Over-expression of the *ku234* 1b cDNA is able to induce seam hyperplasia

If the activator hypothesis is true, then the *ku234* allele should be able to drive seam hyperplasia even in the presence of functional UNC-62. One approach to testing this possibility was to place the *ku234* 1b cDNA under the control of the scm promoter, thus ensuring that it is expressed strongly and specifically in seam cells. This construct was made and injected, together with *scm::gfp*, into *unc-119* worms which were wild type with respect to *unc-62*. Strikingly, this construct was able to drive hyperplasia identical to that seen in the *ku234* mutants (Figure 4.8C); as in *unc-62(ku234)* mutants, significant numbers of seam cells were present in the head, derived from H1 and H2. Thus, over-expression of the *ku234* cDNA of *unc-62* in the seam is able to induce seam proliferation specifically in the head.

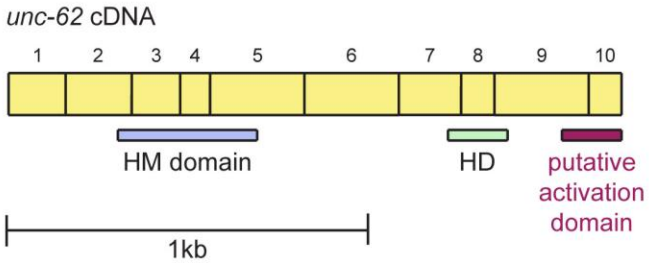
Over-expression of the wild type 1b cDNA is able to induce seam hyperplasia

A control plasmid, containing wild type *unc-62* cDNA under the control of the same promoter, was also made and injected in the same way as for the *ku234* cDNA. Surprisingly, the wild type cDNA, when over-expressed in the seam, is also able to induce hyperplasia (Figure 4.2C). As outlined in the discussion, this can, however, be interpreted in support of the activator hypothesis.

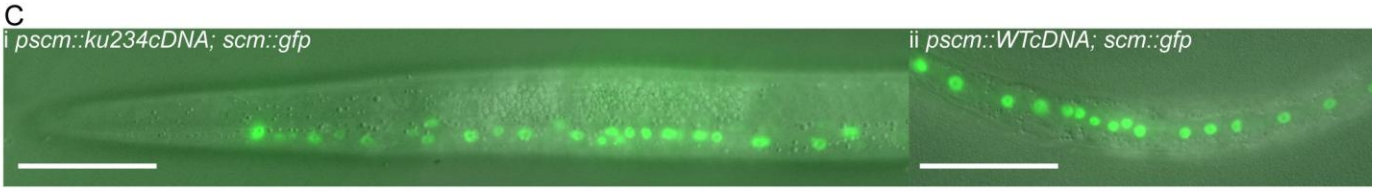
A



B



C



D

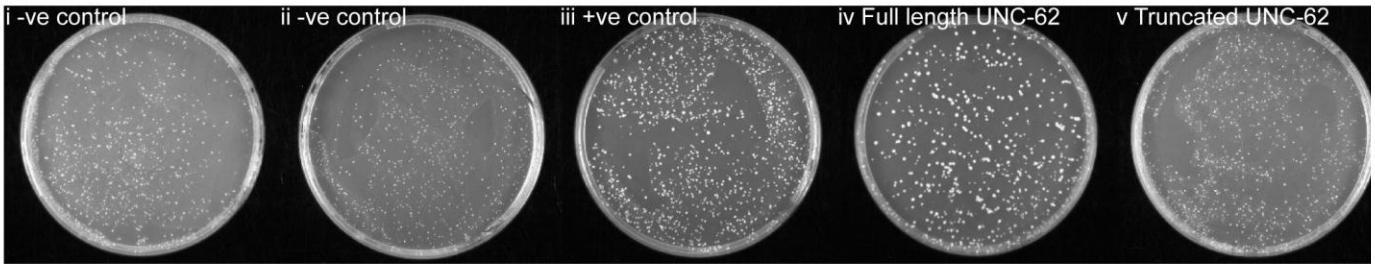


Figure 4.8. A. Alignment of *C. briggsae, C. remanei* and *C. elegans* homologues of *unc-62*, with the HM domain and homeodomain outlined. Following the homeodomain is a stretch of well conserved sequence. B. Schematic of *unc-62* cDNA, showing the relative positions of the HM domain, homeodomain, and putative activation domain. C. Overexpression of the *ku234* mutant (Ci) and wild-ty (Cii) cDNA causes hyperplasia in the head. Scale bars = 100µm.
D. Yeast activation test for *unc-62*. Colony growth after 4 days is shown for yeast integrated with plasmids *pAW370* (negative contro i), *pAW369* (negative control, ii), *pAW655* (positive control, iii), *pAW653* (full length UNC-62, iv), *pAW654* (truncated UNC-62, lackin the putative activation domain at the C-terminus of the protein, v).

As an aside, it is perhaps worth noting that, in this work, only two alternative *unc-62* transcripts were cloned; the 1a-7a transcript and the 1b-7b transcript (n=6). Whilst the total number of transcripts sequenced is small and further work will be required before firm conclusions can be drawn, this result raises the question of whether *unc-62* is represented at the transcriptional level mainly by just two of the possible four (1a-7a; 1a-7b; 1b-7a; 1b-7b; Van Auken, Weaver et al. 2002) splice variants.

Seam hyperplasia is not evident in *unc-62* heterozygotes

To complement the over-expression experiments described above, the effect of the *unc-62(ku234)* and *unc-62(e644)* alleles in heterozygous animals was examined. Based on the 'activator' hypothesis, it might be expected that animals heterozygous for the *ku234* allele would exhibit hyperplasia similar to that seen in homozygous animals. Males carrying the integrated *scm::gfp* array, but wild-type for *unc-62*, were mated with *unc-62(ku234)* and *unc-62(e644)* hermaphrodites. Cross progeny were identifiable by the presence of the *scm::gfp* marker. Interestingly, F1 progeny from neither cross displayed seam hyperplasia.

UNC-62 can act as an activator in yeast in a C-terminus-dependent manner

To test directly whether UNC-62 can operate as an activator of transcription, a yeast system was used. The UNC-62 protein was fused to the GAL4 DNA binding domain (DNA-BD) and transformed into a yeast strain carrying an integrated GAL4-dependent HIS gene. In the absence of activity at the GAL4 promoter, the HIS gene is inactive and, on –his selective plates, minimal growth occurs. If the GAL4 DNA binding domain is able to bring an activation domain into proximity to the GAL4 promoter, the HIS gene is activated and the yeast grows well. As a positive control, the GAL4 DNA-BD was fused to the GAL4 activation domain (AD). Compared with yeast transformed with empty vector controls (see Materials and

Methods), the UNC-62-DNA-BD induced much faster growth, similar to that seen in yeast transformed with the positive control (Figure 4.8D).

The importance of the C-terminal region of UNC-62 was determined by making an identical construct in which the putative activation domain of the protein was absent (the HM domain and TALE homeodomain were kept intact). Strikingly, this plasmid failed to induce growth, suggesting that the activation ability of UNC-62 depends on the C-terminal domain of the protein (Figure 4.8D).

Discussion

UNC-62 acts with CEH-20 to regulate the symmetry of seam divisions

The expression of *unc-62* and *ceh-20* in the seam, taken together with the hyperplasia of the seam lineage observed in these mutants, demonstrates that both genes play important roles in seam development.

Lineage analysis has shown that this role is specific to division symmetry, rather than the decision of whether or not cells enter the cell cycle. In *unc-62(ku234)* and *ceh-20(mu290)* mutants, as well as in *unc-62(ku234); ceh-20(mu290)* double mutants, symmetry defects are observed in the seam throughout the length of the worm. In the head, divisions are almost invariably symmetrised towards the stem fate, in both single mutants, resulting in massive hyperplasia. In the body region, the symmetry defects are more variable.

Notably, extra divisions were not observed; the *unc-62(ku234)* and *ceh-20(mu290)* mutations only affected the symmetry of scheduled divisions. Thus, whilst these two genes do appear to have control over division symmetry, they are not able to drive the resultant stem-fate cells through the cell division cycle. This is supported by three lines of evidence.

Firstly, extra divisions were not observed in the V lineage seam cells in *unc-62(ku234)* or *ceh-20(mu290)* mutants.

Secondly, in the H1 lineage, the anterior daughter of the L1 division adopts the stem fate in *unc-62(ku234)* or *ceh-20(mu290)* mutants. However, even though this anterior daughter it is scheduled to divide several hours before the L2 moult in normal development (Figure 4.3Ci), its aberrantly-adopted seam fate causes it to divide synchronously with the posterior daughter of the L1 division. Thus, for cells which acquire the seam fate, an additional signal or factor must be required to enter the cell cycle; the mutations in *unc-62* and *ceh-20* animals are able to cause cells to incorrectly adopt the seam fate – and thus become receptive to this signal – but are not able to make them divide independently.

Thirdly, none of the inappropriately-specified seam cells in *unc-62* or *ceh-20* mutants derived from H0; this is the only cell in the seam lineage which, during normal development, does not divide. The fact that these mutations are unable to cause extra seam cells to be derived from H0 supports the idea that they are only able to modify the symmetry of divisions which occur, rather than inducing divisions *per se*.

The difference in the seam cell phenotype between *unc-62/ceh-20* single mutants, and the *ceh-20; unc-62* double mutant is striking. In the former, although division symmetry defects are observed throughout the body, from head to tail, seam hyperplasia is largely confined to the head. In the more posterior regions, the defects result in a slight increase in seam cell number but the L2 symmetric division is often observed to be asymmetric, resulting in one, not two, seam cells (Figure 4.3Cii). In the double mutant, however, the situation is very different. Here, as in the head, divisions which are usually asymmetric are frequently symmetrised, during both L1 and L2 stages, resulting in severe hyperplasia throughout the worm (Figure 4.3Ciii). This is clearly reflected in the increase in average seam cell number in this strain (Figure 4.3A).

This finding suggests a clear partitioning of mechanisms between the head and the rest of the body. Both the *unc-62(ku234)* mutation and the *ceh-20(mu290)* mutation/*ceh-20 (RNAi)* are sufficient, separately, to drive hyperplasia, but this hyperplasia is strictly confined to the head. In the body, it appears that there is redundancy between *ceh-20* and *unc-62*; to achieve the same hyperplasia phenotype, mutations in both genes must be present. The reasons underlying this difference are currently unclear, but almost certainly involve interactions with other genes which confer positional information. Both *unc-62* (see above) and *ceh-20* (S. Hughes, pers. comm.) are expressed throughout the length of the anterior-posterior axis, yet their mutant phenotypes have distinct positional biases. Similarly, expression of the wild type and *ku234* mutant *unc-62* cDNAs under the control of the *scm* promoter resulted in hyperplasia only in the head, even though this promoter drives strong seam expression throughout the length of the worm. One possible group of interactors are

the Hox genes, with their distinct domains of expression throughout the worm. *unc-62* and *ceh-20* are well established as Hox-cofactors (Van Auken, Weaver et al. 2002; Yang, Sym et al. 2005; Jiang, Shi et al. 2009) and their precise functions and interactions in the head and body could differ depending on the expression of the Hox genes with which they are co-expressed. These potential interactions are under investigation.

While the exact reason for the difference between head and body regions phenotypes for *unc-62/ceh-20* remains mysterious, the nature of the effect of *unc-62/ceh-20* mutations in the head is clear-cut; even from the first L1 division in both H1 and H2, seam cells almost invariably divide symmetrically, giving rise to two more seam daughters at every division. This almost certainly points to an interaction with the Wnt signalling pathway, which is established as a key regulator of cell division symmetry (Arata, Kouike et al. 2006).

Interestingly, there is already an established link between MEIS, PBX and cell division symmetry in the worm. In the T cell, the most posterior lineage of the seam, division symmetry is also tightly regulated. In the V lineages anterior daughters of asymmetric cell divisions adopt the differentiated, hypodermal fate, whilst posterior daughters remain in the seam line, retaining the stem-like fate. In the T cell, the L1 division is also asymmetric, though the fates of the daughter cells are different; the anterior daughter goes on to produce hypodermal cells, whilst the posterior contributes neural cells, some of which give rise to the tail phasmids. The WNT pathway, the output of which determines levels of nuclear POP-1 in each of the T cell daughters, is central to determining the symmetry of the division(Arata, Kouike et al. 2006) . However, since non-canonical WNT signalling regulates division symmetry throughout *C. elegans* development, this symmetry information must be integrated with tissue-specific, positional information in order to generate cells of the appropriate fates. In the case of the T cell, *psa-3*, *ceh-20* and *nob-1* play central roles (Arata, Kouike et al. 2006).

In the posterior T cell daughter, POP-1 is active and, together with CEH-20 and NOB-1, which bind as a complex to an intronic enhancer element, promotes transcription of psa-3. Furthermore, PSA-3, CEH-20 and NOB-1 then form a complex (CEH-20 binds to PSA-3 and NOB-1, but no physical interaction was detected between PSA-3 and NOB-1) which directs the fate choice of the posterior T daughter (Arata, Kouike et al. 2006).

The combination of MEIS, *ceh-20* and the positional information conferred by *nob-1*, all working together with the WNT symmetry pathway, appears to present a system by which a general symmetry mechanism can be integrated to accommodate local developmental requirements. The possibility that a similar system operates in the rest of the seam lineage – and in particular in the head – is intriguing. *unc-62* and *ceh-20* interact in the seam and clearly have position-specific effects. Thus the anterior Hox genes *lin-39* or *ceh-13* could confer the anterior specificity required. However, the situation in the H and V lineages with *unc-62* does not appear to be a mirror of that seen in the T cell, with *psa-3*. Firstly, the UNC-62::CFP reporter allowed examination of the expression and protein localisation, but this appeared to be uniform throughout the seam, with no apparent bias to either daughter of asymmetric seam divisions. Thus, unlike PSA-3, UNC-62 distribution is not asymmetric. CEH-20, on the other hand, does appear to show a degree of asymmetry in its nuclear localisation (expression appears slightly stronger in anterior daughters), so perhaps this is the asymmetric effector of the system in the non-T lineages of the seam.

An additional difference comes from preliminary evidence which suggests that *unc-62/ceh-20* lie upstream of *pop-1*. POP-1 nuclear asymmetry is set up by the actions of the β-catenin WRM-1, acting in a complex with LIT-1 (Phillips, Kidd et al. 2007). WRM-1 protein is itself asymmetrically distributed. In asymmetric V lineage divisions, the anterior, non-stem-like daughter exhibits a largely cortical distribution of WRM-1, whilst in the posterior daughter WRM-1 protein localises to the nucleus, where it acts to reduce nuclear levels of POP-1, thus ‘activating’ the remaining low levels of this protein. In *ceh-20* mutants, WRM-1::GFP localisation appears to be perturbed (S. Hughes, pers. comm.), with both anterior and

posterior cells possessing high levels of nuclear, rather than cortical, WRM-1. Thus, the situation in the T cell provides a glimpse of the potential interactions which bring together symmetry and positional pathways. However, it is certainly not possible to relate this model directly to the more anterior lineages by merely substituting *unc-62* for *psa-3*. Currently, the anterior Hox genes *ceh-13* and *lin-39* are likely the best candidates for interacting partners with *unc-62*.

unc-62/ceh-20 likely act as indirect upstream regulators of *bro-1/rnt-1*

The severe seam hyperplasia in *unc-62(ku234)* and *ceh-20(mu290)* mutants is totally dependent on functional *bro-1* and *rnt-1*; in a *bro-1; rnt-1* background this phenotype is completely suppressed. This suggests that the *bro-1/rnt-1* complex acts downstream of *unc-62/ceh-20*, but whether the latter act on the former directly is not certain. The *bro-1* CNE appeared to offer the most likely point of interaction, given the presence of two well conserved MEIS sites. However, mutagenesis of these sites in the CNE (see above) resulted in no change in the seam expression driven by this intronic element, suggesting that, despite their conservation and similarity to the MEIS consensus (TGACAG(Chang, Jacobs et al. 1997)), they are not in fact genuine sites.

The size of the *rnt-1* long intron which drives expression of this gene (see Chapter 6) makes the search for genuine *unc-62* binding sites laborious. There are, however, no particularly well-conserved sites within the intron.

Given the finding that WRM-1::GFP localisation is perturbed in *ceh-20/unc-62* mutants (S. Hughes, pers. comm.), it seems more likely that the interaction between *unc-62/ceh-20* and *bro-1/rnt-1* is indirect; the MEIS/PBX complex sets up the symmetry (or asymmetry) of the seam cell divisions, and *bro-1/rnt-1* are downstream targets, facilitating the divisions in those cells designated to have the seam fate.

unc-62/ceh-20 and *bro-1/rnt-1* also act redundantly in embryonic seam development

That *unc-62; bro-1; rnt-1* triple mutants possess even fewer seam cells than *bro-1; rnt-1* double mutants presents a paradoxical situation. In an otherwise wild type background, *unc-62(ku234)* causes a dramatic elevation in seam cell number. This increase is dependent on *bro-1; rnt-1*. However, in a *bro-1; rnt-1* background, this hyperplasia is not just suppressed but reversed; triple mutant worms have even fewer seam cells than the double mutants. It would appear that the *ku234* mutation has a positive effect on seam cell number in wild type worms, and a negative effect in *bro-1; rnt-1* worms, requiring the postulation of a relatively complex mechanism, the workings of which are as yet very much mysterious.

An alternative solution is to separate the embryonic and larval stages of seam development. Crucially, whilst wild type, *unc-62(ku234)* and *bro-1; rnt-1* mutant worms all have 10 seam cells at hatching, only the triple mutant strain exhibits reduced numbers of seam cells at this early stage. Thus, *unc-62, bro-1* and *rnt-1* appear to act redundantly during embryonic development, in an interaction which is distinct from that during larval development. Indeed, this idea matches the data obtained from seam cell counts. On average, *unc-62; bro-1; rnt-1* triple mutant larvae hatch with 8 seam cells on each side, compared to 10 in wild type worms, *unc-62* single mutants and *bro-1;rnt-1* double mutants. This 2-cell deficiency matches the decrease in overall seam cell number, from 12 in *bro-1; rnt-1* double mutants to 10 in *unc-62; bro-1; rnt-1* triple mutants. Thus, *unc-62* appears to act upstream of – and depend on – *bro-1/rnt-1* during larval seam development, but redundantly with these genes during embryonic development.

Exactly how these genes interact and perform their role in the embryonic seam remains unclear. One possibility is that seam cells fail to form, perhaps again due to symmetry defects, perhaps due to division failures. An alternative explanation stems from the recent finding that both *unc-62* and *ceh-20* have novel roles in preventing cell death (Potts, Wang et al. 2009). *unc-62*, by promoting the nuclear localisation of *ceh-20*, represses transcription of

the pro-apoptotic gene *egl-1*, allowing the survival of a subset of motor neurones in the worm. Here again, *unc-62* acts to regulate the subcellular distribution of CEH-20, but this finding demonstrates the breadth of *unc-62* downstream targets and, especially given the observation of cells which appear to be in the process of degrading in *unc-62; bro-1; rnt-1* triple mutant embryos (Figure 4.6Biii), raises the possibility that this function may apply to the seam, in conjunction with the *bro-1/rnt-1* complex.

High levels of synthetic embryonic and lethality suggest wider redundant roles for *unc-62* and *bro-1/rnt-1*

Both *bro-1/rnt-1* and *unc-62(ku234)* mutants show low levels of embryonic lethality. The increase in lethality seen when these mutants are combined is dramatic (Figure 4.6C). In terms of larval lethality, the results are similar; although *bro-1; rnt-1* and *unc-62(ku234)* animals show higher levels of larval lethality than wild type animals (12.5%, 22.5% and 2.5% respectively, Figure 4.6C), the increase to 81.1% in triple mutant animals is nevertheless dramatic.

As for the reason for the extreme sickness of these animals, it is perhaps unlikely that seam defects are solely to blame. On average, *unc-62; bro-1; rnt-1* animals have only two fewer seam cells than *bro-1; rnt-1* double mutants. It seems unlikely that this modest decrease would cause 33% and 69% increases in embryonic and larval lethality respectively, especially given that the decrease from 16 seam cells in wild type worms to 12 in *bro-1;rnt-1* mutants results in only 0.87% embryonic lethality and 12% larval lethality. Furthermore, the larvae which die – and many of those that survive – are severely deformed; such a phenotype is likely due to more than just seam defects. It is however possible that there is a threshold number of seam cells required for survival, below which lethality levels and more general developmental abnormalities dramatically increase.

An alternative explanation is that *unc-62* and *bro-1; rnt-1* perform functions outside the seam. Given that both genes are expressed in muscle, this is perhaps not just possible, but likely. Certainly, the morphological defects observed are consistent with defects in muscle development (Hresko, Williams et al. 1994; Williams and Waterston 1994; Baugh, Wen et al. 2005).

Work is underway to determine whether the embryonic and larval lethality phenotypes can be rescued by the *rnt-1* specifically expressed either in the seam or in muscle (using the tissue-specific promoter modules described in Chapter 6). As a control, the full length *rnt-1* construct, driving *rnt-1* expression in both muscle and seam, has been crossed into the *bro-1;rnt-1;unc-62(ku234);[scm::gfp]* strain, and *rnt-1; unc-62(ku234)* mutants carrying the array have been obtained. As expected, the *unc-62(ku234)* phenotype is de-suppressed, and seam hyperplasia in the head is evident. Qualitatively, the worms also appear considerably less sick. The 'muscle-only *rnt-1*' construct has also been crossed into the triple mutant background. So far, healthy F2s carrying the array have been isolated, but their genotype with respect to *unc-62(ku234)* remains to be determined (though these worms carry the integrated *scm::gfp* array, which is linked to *unc-62*, so they are likely to also carry the *unc-62(ku234)* mutation). Should the *unc-62(ku234)* background be confirmed, the rescue seen in these worms would be strong evidence for a redundant role in muscle underlying the lethality phenotypes observed.

The array of functions of *unc-62* is reflected by the expression of this gene in so many different tissue types, from the seam cells, hyp7 syncytial cells, ventral hypodermis, the P lineage, the ventral nerve cord and the embryonic M lineage (Jiang, Shi et al. 2009). One intriguing possibility is that the different splice variants of *unc-62* reflect the partitioning of some of its differing roles in development. The finding that seam expression is specifically driven by the small region of sequence just 5' to the 1b isoform start codon raises the question of whether the seam functions of *unc-62* are fulfilled by this isoform in particular. It

would be interesting to use different fluorophores fused to the 1a and 1b isoforms to assay for tissue-specific expression patterns of each.

Seam hyperplasia in *unc-62(ku234)* and *ceh-20(mu290)* mutants likely results from constitutive activation of targets

Whilst TALE-family homeodomain proteins are involved in both repressing and activating a range of different transcription factor complexes, Meis-Pbx complexes – those which are most closely related to *C. elegans unc-62* and *ceh-20* – are mainly involved in transcriptional activation (Hyman-Walsh, Bjerke et al. 2010). This finding is supported by the recent discoveries that the mammalian Meis genes Meis1 and Meis2 possess activation domains at their C-termini (Huang, Rastegar et al. 2005; Hyman-Walsh, Bjerke et al. 2010). The activity of these activation domains depends on interactions with the N-terminal end of the protein, which interacts with and inhibits the activation domain. In the mammalian system, this 'autoinhibition' is in turn regulated by interactions with Pbx, which binds to the N-terminus of the Meis protein. In the absence of Pbx, autoinhibition occurs; when Pbx is present, autoinhibition is relieved and the C-terminal end of the protein is free to operate as an activator. It may well be that more detailed studies on other mammalian Meis proteins – and Meis homologues in other animal groups – reveal that this autoinhibition system is even more prevalent than currently thought.

The discovery of activation and autoinhibition domains in mammalian Meis genes raised the question of whether they were also present in *C. elegans* UNC-62 and whether, in the worm, UNC-62 has the potential to work as an activator. Is the seam hyperplasia phenotype manifest in *unc-62(ku234)* mutants the result of a lack of repression of *bro-1/rnt-1* (albeit indirectly), or a result of inappropriate activation? Given the nature of the *unc-62(ku234)* mutation, the latter option is certainly intriguing; although the 1b isoform start codon is lost in *ku234* mutants, the fact that over-expression of the *ku234* mutant cDNA drives the same

hyperplasia phenotype as that seen in *unc-62(ku234)* mutants is strongly suggestive of this cDNA having functionality. One possible counter argument is, whilst no protein is translated from this cDNA, the many copies of the array titrate 'seam cell expression' transcription factors away from the endogenous *unc-62* locus, resulting in less functional protein being produced overall. However, the finding that over-expression of the wild type cDNA has the same effect suggests this is not the case; even if transcription factors were attracted to the array-based copies of *unc-62*, functional protein would be produced. The phenotype, therefore, is likely not due to a decrease in UNC-62 levels.

Instead, the observations can be explained by considering that introduction of the *ku234* cDNA will lead to an abundance of *ku234* protein, which lacks the N-terminal region of wild type UNC-62. As a result, the putative activation domain will be free to activate targets, the end result being an increase in the number of cells with the seam fate.

The phenotype seen in *ceh-20* mutants and RNAi animals could also be consistent with the 'activator' hypothesis. Here, although full length UNC-62 is present, CEH-20 protein will not be available (or available at relatively low levels) to bind to UNC-62. If CEH-20, as in the mammalian context, acts to relieve autoinhibition of the putative activation domain, it is conceivable that the mutant CEH-20 protein over-inhibits this autoinhibition, thereby leading to the seam hyperplasia observed. Accounting for the effects of CEH-20 RNAi using this model is, however, more difficult. An alternative interpretation is that in *C. elegans* CEH-20 acts to promote autoinhibition of the activation domain, which would explain the seam phenotype of *ceh-20* mutants and the effect of *ceh-20* RNAi. This latter hypothesis is supported by more recent data from the yeast system (discussed below).

How then is it possible to rationalise these ideas with the finding that over-expression of wild type UNC-62 can also drive an increase in seam cell number? If CEH-20 protein is present in limiting amounts, it is possible to imagine a situation where introduction of multiple copies of the array lead to a significant increase in UNC-62 protein levels. If the ratio of UNC-62 to

CEH-20 is sufficiently high, a proportion of UNC-62 could remain unbound by CEH-20 and, again, able to serve as an activator. Again, hyperplasia would result. Indeed, this was seen to be the case (Figure 4.8C).

The importance of the relative abundance of UNC-62 and CEH-20 can also account for the absence of a seam hyperplasia phenotype in *unc-62* heterozygotes. It may be that, in *ku234*/+ heterozygotes, the abundance of mutant transcripts is not sufficient to lead to constitutive activation of UNC-62 targets. This situation can be compared to that in animals carrying multiple copies of the *ku234* allele; here, many more copies will be present than the single copy in heterozygotes. Indeed, the effect of the presence of multi-copy extrachromosomal arrays carrying the wild-type *unc-62* gene, usually represented by just two copies, demonstrates the importance of 'dosage' in this context.

The yeast ‘activator’ assay provided a further and perhaps more direct test of the ability of UNC-62 to serve as an activator of transcription. Here again, this mode of action is supported, particularly with the finding that the C-terminus of the protein is required for activator activity. As a confirmatory experiment, it was of interest to determine whether the presence of CEH-20 was able to abolish the ability of the full length UNC-62 protein to activate transcription. Recently, this question has been addressed by the introduction of CEH-20 into the yeast system, in conjunction with the other constructs used (S. Hughes and A. Woollard, pers. comm.). Supporting the view that the presence of CEH-20 facilitates autoinhibition of the activation domain in *C. elegans*, introduction of CEH-20 into yeast carrying the activating, full length UNC-62 construct, led to a decrease in yeast growth.

Conclusion

As was the case with *elt-1* (Chapter 3), *unc-62* was identified as a candidate for involvement in seam cell development based on putative binding sites within the *bro-1* CNE. In the case of this gene, however, this predicted interaction proved false, demonstrating the need for experimental verification of predicted interactions between transcription factors and enhancer sequences.

Fortuitously though, *unc-62* clearly has an important role in the seam lineage, demonstrated most obviously by the extreme hyperplasia phenotype of *unc-62(ku234)* mutants. The independent discovery of a role for *ceh-20* in the seam led to the finding that these genes work together as key regulators of division symmetry in the seam, likely operating upstream of components of the WNT signalling pathway and depending on the actions of *bro-1* and *rnt-1* for their function. How this integration is coupled with positional information, which gives rise to the head-specific phenotypes observed in *unc-62* mutant animals, remains a key question to be addressed.

unc-62 also appears to act redundantly with the *bro-1/rnt-1* complex in embryonic and early larval development, certainly in the seam lineage and likely in muscle as well, defining novel roles and interactions for all three genes. Defining these early-stage functions, particularly with respect to these genes' joint roles in muscle has the potential to shed light not just on embryonic development but on a completely unexplored facet of *bro-1/rnt-1* function in the worm.

CHAPTER 5

Dissection of the *scm::gfp* seam cell marker

Introduction

The *scm::gfp* reporter is a well-established marker of seam cells

In almost all publications which cover the development of the seam lineage in the worm, the 'Seam Cell Marker', *scm::gfp*, is used (for example Smith, McGarr et al. 2005; Zhang, Ding et al. 2007; Kanamori, Inoue et al. 2008; Fritz and Behm 2009). This GFP transcriptional reporter is expressed strongly in seam cells, from the two-fold stage of embryonic development right through to and during adulthood (Koh and Rothman 2001), making them easy to identify and count. Furthermore, it fades relatively rapidly in cells which leave the seam and differentiate. Thus, it is used not just as a marker of the seam lineage, but of the seam stem-like fate. However, the promoter used in the *scm::gfp* reporter is large (7.8kb) and it is not known how much of this sequence is required for seam expression. Moreover, the normal function of this promoter sequence – and of any genes which may be under its control – remains to be identified.

Before the *scm::gfp* transcriptional fusion was made, the *scm* promoter was identified in a promoter trapping screen, whereby genomic fragments were cloned upstream of a *lacZ* gene, allowing β-galactosidase expression to be used as a read-out of the endogenous expression of the promoter elements identified (Hope 1991). One 'promoter' element was found to drive expression in seam cells and used in this construct, as an integrated array, to mark the seam lineage (Gendreau, Moskowitz et al. 1994). Subsequently, the same

promoter element was used to drive *gfp* instead of *lacZ*, making visualisation of expression easier and feasible for live worms (Koh and Rothman 2001).

Discovering the identity and possible function of the scm promoter is of value for several reasons. Firstly, it is a widely used marker of seam cells. Understanding the endogenous role of this promoter and any associated gene is therefore of general relevance. If, for example, the scm promoter has only a tenuous developmental link with the seam fate, there may well be cases where its expression is uncoupled from other aspects of the seam fate (indeed, such cases are outlined and discussed in Chapters 3 and 7). Understanding if – and, if so, then why – this could happen will make *scm::gfp* a much more powerful and useful marker of the seam.

Secondly, knowledge of the workings of promoters is far from complete. The *scm* promoter provides a prime example of strong, tissue-specific expression and offers the opportunity to dissect the mechanistic basis of transcriptional regulation. Revealing the size of such a promoter region, or the number of transcription factors which bind to it in order to drive seam expression, will reveal more about the workings of promoters in general. Furthermore, with respect to promoters which drive seam expression, a compilation of the promoters of different seam-expressed genes will provide a foundation for searching for common motifs or modules, shedding light on whether similar expression patterns utilise similar elements at the sequence level.

Thirdly, there appears to be a link between *scm::gfp* expression and seam development. In *rnt-1*, *bro-1* and *pal-1* mutants, and in *elt-1(RNAi)* worms, loss of *scm::gfp* expression is consistently observed. Although not tightly correlated with the loss of proliferative ability also seen in these worms, it is tightly correlated with loss or reduction in function of these genes, and thus with defects in seam development. Understanding the nature of this marker, why its expression is lost in numerous cells in these backgrounds and what roles the underlying gene is doing will benefit our overall understanding of seam development.

Results

The *scm* promoter consists of an 7.8kb region of chromosome IV, overlapping three genes

Alignment of the *scm* promoter revealed a 100% match with a region of chromosome IV. Three annotated genes are overlapped by the fragment, two partially and one completely. These are *arf-3*, a member of the Arf family of GTPases (Li, Kelly et al. 2004), *mec*-17, which is required for the function of touch receptors in the worm (Chalfie and Au 1989; Zhang, Ma et al. 2002), and the collagen gene *col-115* (Figure 5.1A).

Two conserved regions of sequence associated with *arf-3*, an ADP-ribosylation factor, drive seam expression

To determine whether the whole promoter region, or one or more subsections, are required for seam expression, the 7.8kb region was cut into smaller fragments which were cloned into the minimal promoter vector *pPD107.94*. Figure 5.1B summarises the results from these manipulations. The region of sequence required for seam cell expression was narrowed down to a 603bp fragment which included the last intron of *arf-3*. Primers were designed to precisely amplify this intron (Figure 5.2A, region 2) and the resulting fragment, when placed in pPD107.94 (giving pAW627), was found to drive seam expression.

Given that promoter elements are generally considered to lie in sequence 5' of the ATG of the relevant gene, this region too (spanning the sequence from *unc-24* to the ATG of *arf-3,* Figure 5.2A) was cloned into *pPD107.94* and assayed for ability to drive seam expression. GFP expression was indeed observed in the seam cells (data not shown). Since this construct was made, an additional gene, *F57H12.10*, has been annotated as lying between *unc-24* and *arf-3* (region 1, Figure 5.2A). This gene remains otherwise uncharacterised and

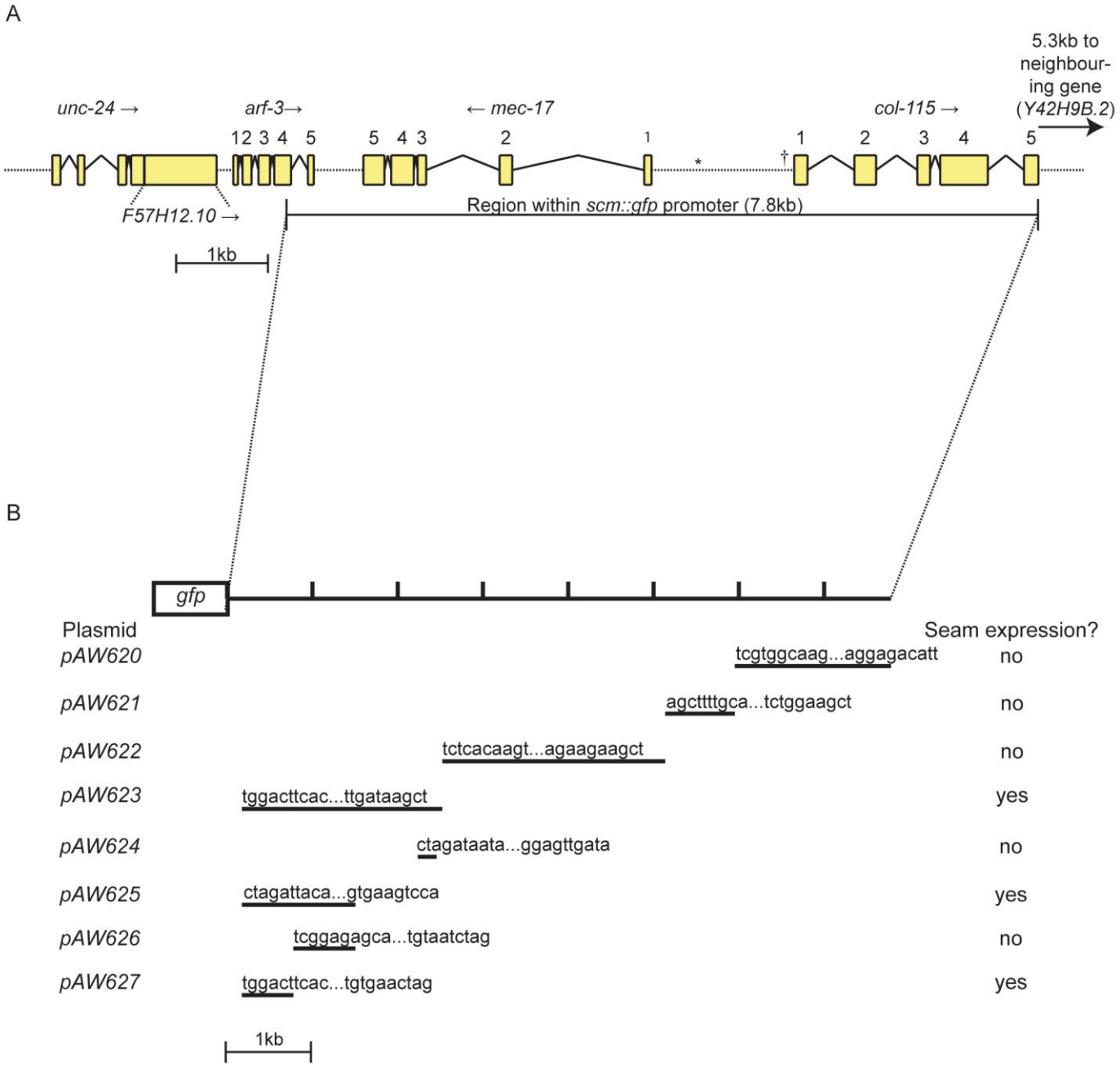


Figure 5.1 A. Alignment of the genomic sequence of the *scm::gfp* promoter showed a match with a region on chromosome IV. The 7.8kb region spans three genes: *arf-3* (left) and *col-115* (right), with which it overlaps partially, and *mec-17*, which it covers completely. The asterisk and dagger show the position of two genes encoding 21 nucleotide small RNAs (*F57H12.8* and *Y42H9B.4* respectively).
B. Dissection of the *scm::gfp* promoter was performed by cloning the fragments shown into the minimal promoter vector *pPD107.94*, giving the plasmids listed on the left-hand side. This approach allowed the region required for seam expression to be narrowed down to the 603bp fragment in *pAW627*.

it is possible that the enhancer identified here described here reflects transcriptional activity of *arf-3*.

RNAi knockdown of *arf-3* leads to variable seam cell numbers

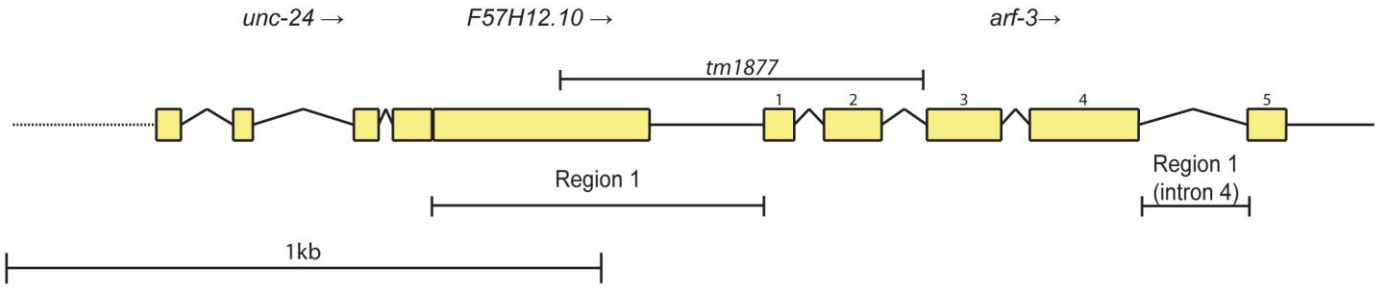
arf-3 is known to encode an ADP-ribosylation factor, a member of the RAS superfamily of GTPases (Li, Kelly et al. 2004). However, relatively little else is known about the function of this gene in *C. elegans*. To determine whether *arf-3* plays a role in seam cell divisions, RNAi by feeding was used to reduce levels of *arf-3* transcript. The result was a significant increase in the variation of seam cell number, relative to controls ($p<0.003$), and a slight but significant overall increase in the average number of seam cells per side(Figure 5.2B); this suggests that *arf-3* does indeed play a role in the seam (though, as discussed below, it is possible that sequence similarity between *arf-3* and *arf-1* could lead to inadvertent knockdown of both of these genes even when only one is targeted for knockdown by RNAi).

The increase in variation was apparent only after embryonic stages; *arf-3(RNAi)* L1s invariably had 10 seam cells at hatching, suggesting that the seam cell phenotype observed is due to larval-stage defects.

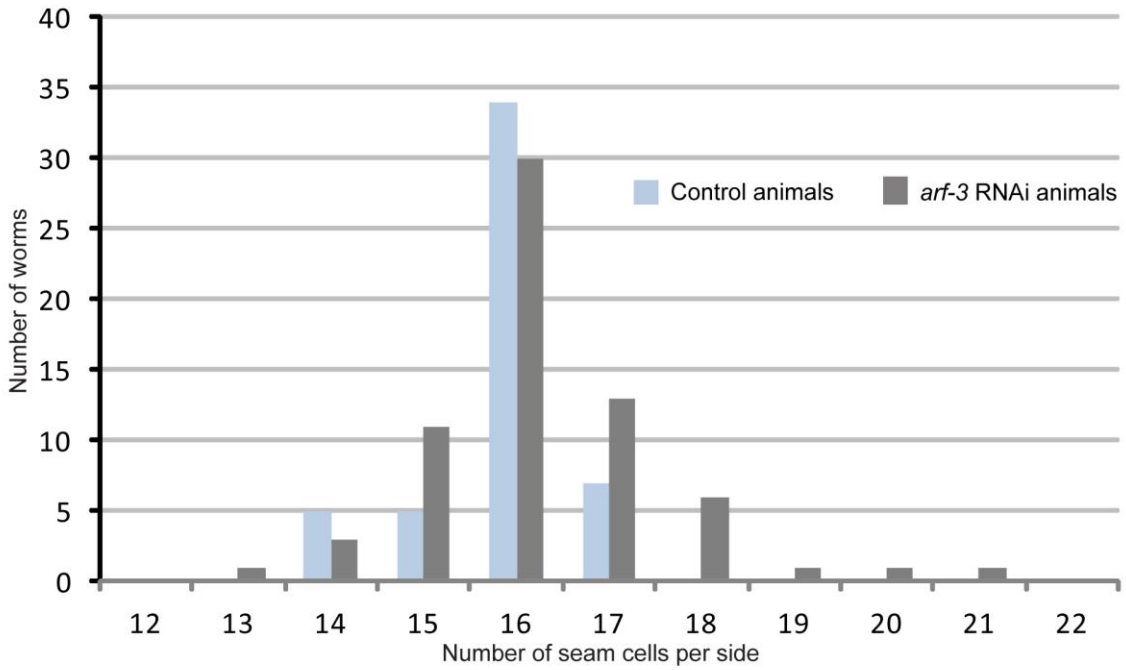
arf-3(tm1877) mutants were obtained from the Mitani lab in the hope that these may allow further dissection of *arf-3* function. This strain, *tm1877*, carries a 631bp deletion which results in embryonic or early larval lethality. The mutation was balanced with *dpy-20(e1282); unc-24(e138)* and the *scm::gfp* marker was crossed into the resultant strain. *tm1877* homozygotes exhibited seam defects (Figure 5.1C). Gaps in the seam were present and there was evidence of aberrant division patterns and orientations (though the developmental basis of these defects has not yet been established by lineage analysis).

Initially, the *tm1877* lesion was annotated as affecting only *arf-3* but, as noted above, an additional, as yet uncharacterised gene, F57H12.10, has been predicted to lie within the

A



B



C

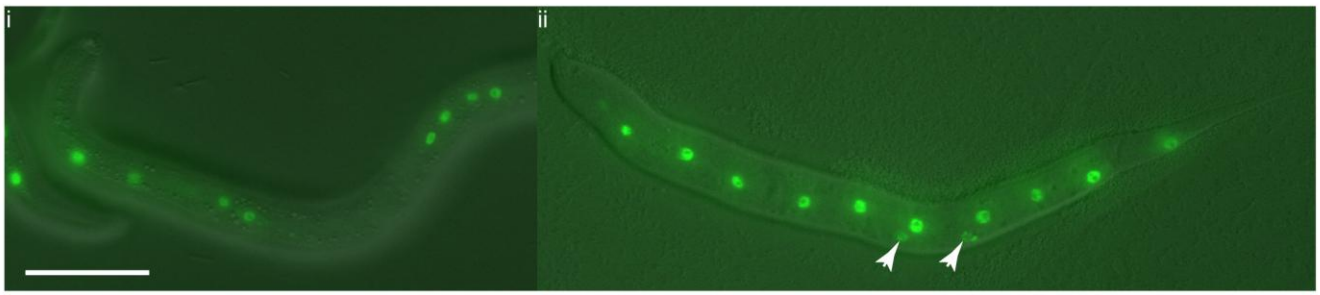


Figure 5.2 A. Detailed schematic overview of the *arf-3* locus, showing the postion of the *tm1877* deletion and highlighting Regions 1 and 2, both of which are sufficient to drive seam expression (see main text).
B. Seam cell numbers in adult hermaphrodites fed on control and *arf-3* RNAi feeding bacteria. *arf-3* RNAi leads to a small but significant increase in average seam cell number compared to controls (16.2±0.2 versus 15.8±0.1 respectively), together with a significant increase in seam cell number variation; up to 21 and as few as 13 seam cells were counted in *arf-3(RNAi)* worms, compared to control worms, where the range was from 14 to 17 seam cells.
C. *arf-3* homozygotes die as larvae, but larvae from heterozygous mothers exhibit seam defects, including gaps in the seam (i) and aberrant divisions (ii, arrowheads). Scale bar = 100µm.

deleted region. Thus, the *tm1877* phenotype, 100% early larval lethality, may be due in part to the effects of *F57H12.10* loss of function, in addition to disruption of *arf-3*.

As discussed later, one potential regulator of *arf-3* is *cnt-2*. Intriguingly, *cnt-2* is also expressed in the seam (Wormbase), but RNAi knockdown of this gene did not result in a significant change in seam cell number (data not shown).

ARF-3::GFP accumulates as granules in the cytosol of seam cells

arf-3 expression in the seam, together with expression in other cells, has been previously reported for embryos (Li, Kelly et al. 2004). However, the *scm::gfp* reporter is clearly expressed in larvae and adults, as well as in embryos, but *arf-3* expression at these stages has not been demonstrated. In addition, the subcellular localisation of ARF-3 and the distribution of the protein with respect to cell division remain unknown. To shed light on these issues, a C-terminal genomic ARF-3::GFP fusion reporter was constructed, using the full length genomic open reading frame and 5' sequence stretching to *unc-24.*

Strong seam expression of the resultant construct was observed at all stages, from the two-fold embryonic stage until adulthood, entirely consistent with expression of *scm::gfp*. As previously reported, some expression was apparent in other cells (head neurones, intestine) though this was extremely faint. Strikingly, the GFP localised to distinct puncta within the cytosol of the seam cells and was clearly excluded from the nucleus (Figure 5.3A). As discussed below, the identity of *arf-3*, a GTPase G protein, strongly suggests that it will localise to membranous structures. It is likely that the granules evident in Figure 5.3 are therefore associated with structures such as endoplasmic reticulum or the Golgi stacks. Expression and protein localisation was examined with respect to the cell division cycle and the asymmetric fates of daughter cells. GFP levels decline sharply soon after cells leave the seam and move into the hyp7 syncytium (following asymmetric seam cell division), but there

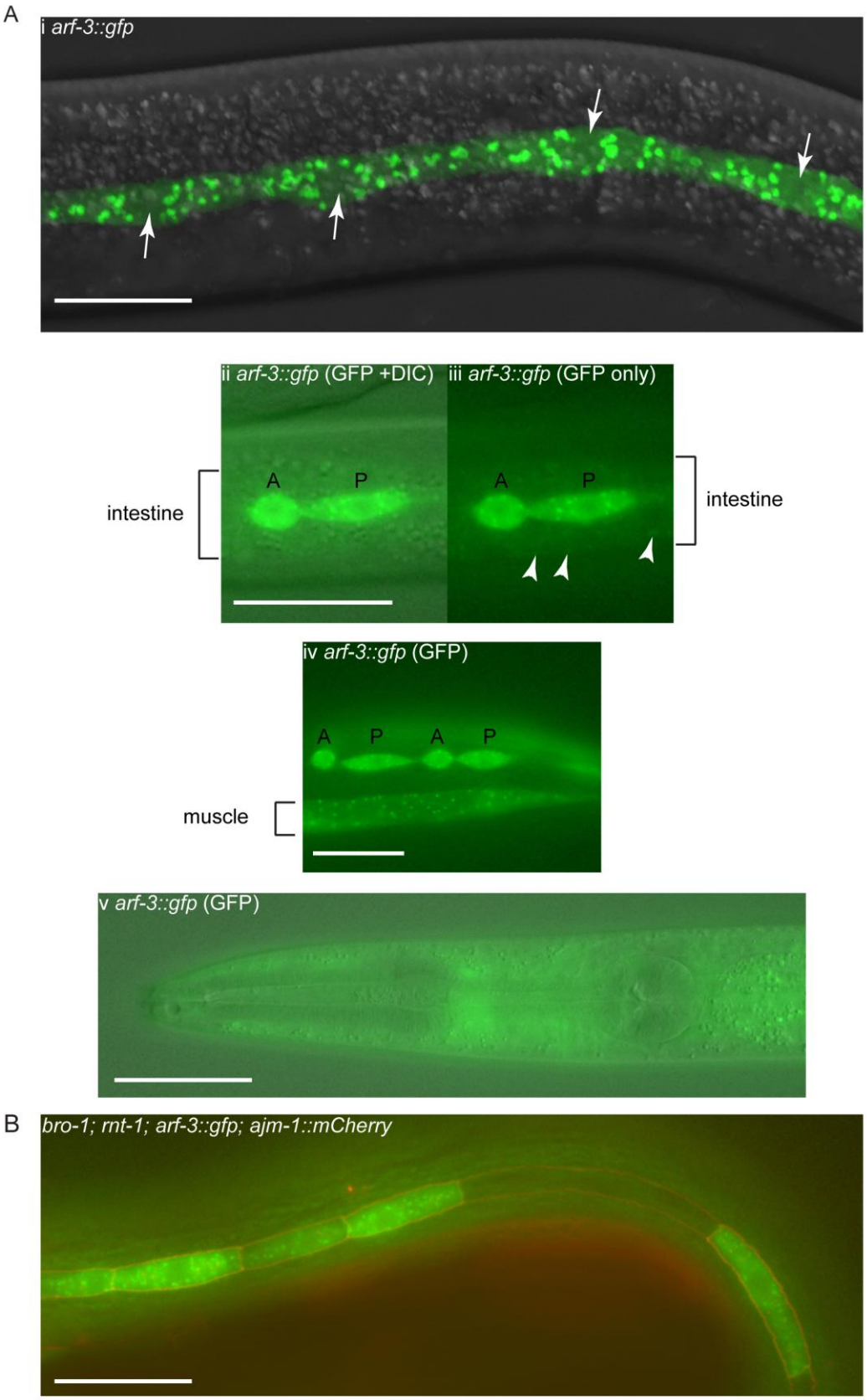


Figure 5.3 A. The ARF-3::GFP translational reporter drives expression in the seam, with clusters of protein aggregating exclusively within the cytosol (Ai). Arrows mark the positions of the seam cell nuclei, which are noticeably lacking GFP. Expression of the ARF-3::GFP reporter is evident in both daughters of asymmetric seam divisions (Aii, iii and iv). Intestinal expression is also evident (Aii and iii); GFP puncta (arrowheads) are visible within the the bracketed region. Muscle expression is also frequently visible (Aiv). A = anterior daughter, P = posterior daughter. As previously reported (Li, Kelly et al, 2004), pharyngeal muscle expression was evident (Av). B. In a *bro-1(tm1183); rnt-1(tm388)* double mutant background, expression of the reporter is lost from a variable subset of the seam cells. This is not simply due to mosaicism, as shown by the AJM-1::Cherry boundary of the seam cells, demonstrating that the extrachromosomal array containing *unc-119*, AJM-1::Cherry and ARF-3::GFP is present in the cell. The presence and distribution of the protein in cells expressing the reporter appears unaffected. Scale bars = 25μm.

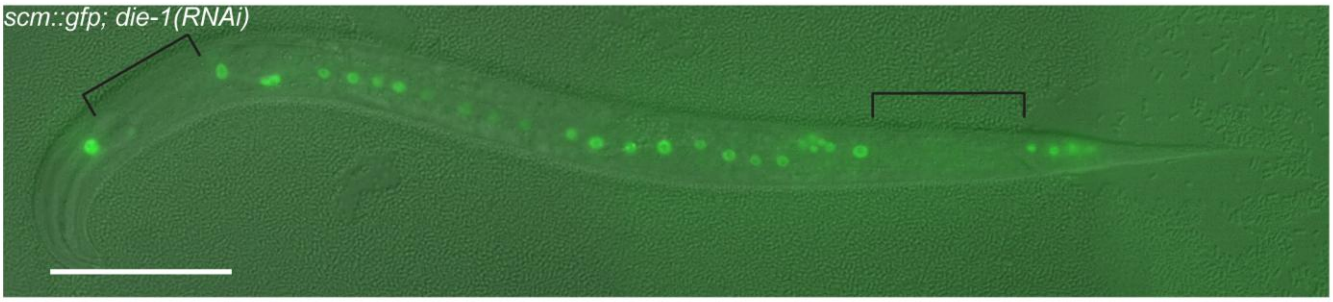
was no apparent difference in the distribution of ARF-3::GFP granules between anterior and posterior daughters of asymmetric seam divisions during or immediately following division.

The full length ARF-3::GFP translational reporter, which includes 5' sequence up to *unc-54* (thereby covering the entire region which is deleted in tm*1877*), was crossed into a balanced *tm1877* background, to rescue the embryonic lethality phenotype, presumably caused by loss of function of *arf-3* and *F57H12.10*. However, it was not possible to rescue the *tm1877* larval lethal phenotype with the full length reporter, suggesting perhaps that the presence of *gfp* at the C terminus of the *arf-3* protein may interfere with its function, or that *tm1877* has a second site mutation.

arf-3 expression – but not ARF-3 subcellular localisation – is partially dependent on *rnt-1/bro-1*

Given the relationship between *rnt-1/bro-1* and *scm::gfp* expression, whereby mutants of the former exhibit consistent loss of expression of the latter, ARF-3::GFP expression and localisation was examined in *rnt-1*; *bro-1* double mutants. Here again, numerous – but not all – seam cells showed loss of ARF-3::GFP (Figure 5.3B). This was not due to mosaicism of the extrachromosomal array as such cells were outlined by AJM-1::mCherry, which was co-injected with the ARF-3::GFP reporter. Rather, loss of GFP in these cells resulted directly from loss of *rnt-1/bro-1*. ARF-3::GFP localisation within GFP-positive cells was not noticeably affected. Taken together with the similar findings from the transcriptional *scm::gfp* reporter, these findings suggest that *rnt-1/bro-1* act as important regulators of *arf-3* at the transcriptional level but not at translational or post-translational stages of processing.

A



B

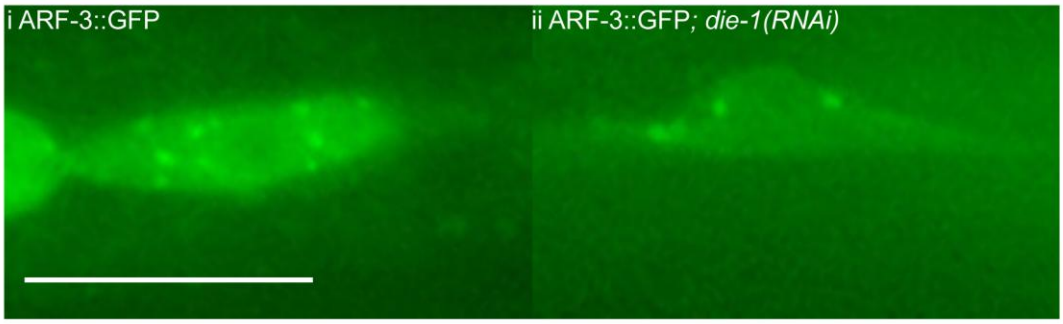


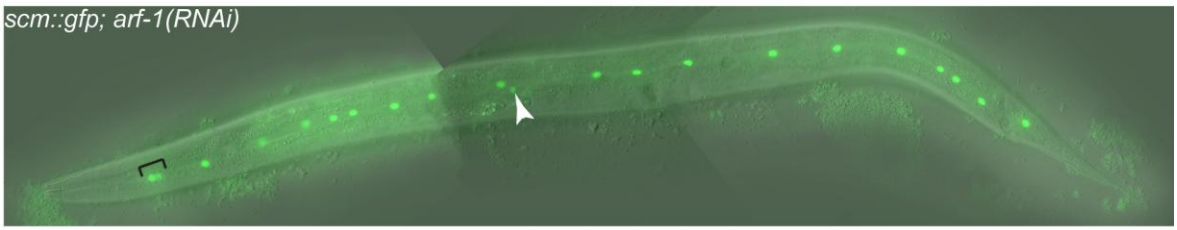
Figure 5.4 A. RNAi knockdown of *die-1* results in high levels of early larval lethality, but worms which survive exhibit gaps in the seam line (bracketed regions). Seam cells marked by *scm::gfp*.
B. In *die-1(RNAi)* worms expressing the ARF-3::GFP translational reporter (Bii), cells are visible which possess fewer GFP granules in the cytoplasm than controls (Bi), suggesting that reduction in function of die-1 leads to lower levels of *arf-3* transcription. Scale bars = 20μm.

Identification of potential upstream regulators of *arf-3*

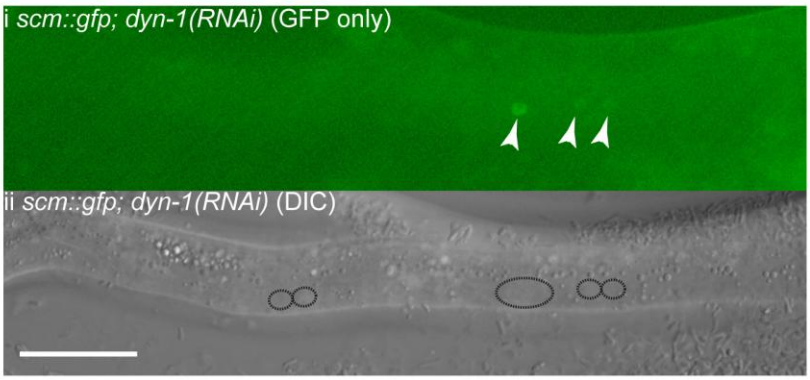
The two small, relatively well conserved regions of 5' and intronic sequence which are apparently associated with *arf-3* and capable of driving seam expression appeared to be ideal candidates for a yeast 1-hybrid screen. Three copies of each region were cloned into the appropriate vectors (*phisi-1* and *placzi*, see Chapter 2, Materials and Methods). Constructs carrying copies of the 5' promoter region were successfully integrated, though it proved impossible to integrate the reporters containing copies of the last intron of *arf-3*. The screen was therefore performed with the presumed *arf-3* 5' promoter. Of the several hundred colonies obtained following transformation with a transcription factor cDNA library, one grew particularly well on the selective plates used, suggesting a strong interaction between the transgenic transcription factor present in the colony and the *arf-3* promoter. Subsequent re-streaking of this and other colonies confirmed this difference. Sequencing of this apparently positive clone revealed the identity of the cDNA to be *die-1*, a transcription factor required for dorsal intercalation during embryogenesis (Heid, Raich et al. 2001), but which is also expressed in the seam where its role remains largely uncharacterised. In embryogenesis, the seam cells of *die-1* mutants are often found to be 'pinched out', but larval phenotypes have not been described.

To investigate the potential role of *die-1* in the seam, RNAi was used to knock down levels of this gene. Although levels of embryonic and early larval lethality were high, a small proportion of larvae survived the L1 stage. In these, expression of *scm::gfp* revealed gaps in the seam line, confirming that this gene does impact upon seam development (Figure 5.4A). It will be interesting to determine whether this effect is a direct consequence of the early seam morphogenetic defects seen in *die-1* RNAi embryos. ARF-3::GFP expression was also examined in worms subjected to *die-1* RNAi. In *die-1(RNAi)* worms, seam cells were often found to contain fewer GFP puncta than cells in control animals (Figure 5.4Bi and ii). This

A



B



C

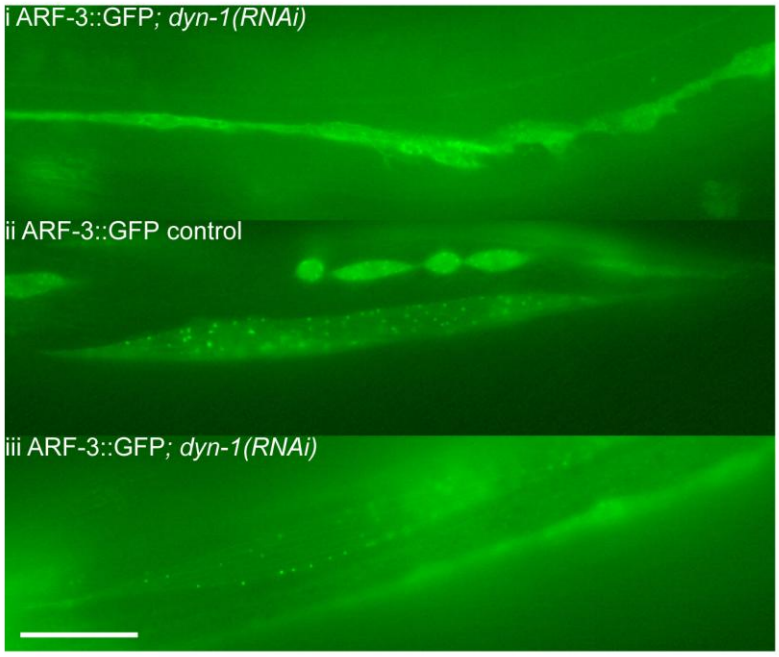


Figure 5.5 A. Hyperplasia in a late L4 hermaphrodite, caused by RNAi knockdown of *arf-1*. Mispositioning c cells is also observed, with two adjacent cells at the anterior of the worm (bracketed) and a cell clearly out of line with the rest of the seam (arrowhead). Given that the seam cells have already fused, GFP signal fro this cell likely represents continued active expression of the *scm::gfp* marker, and not perdurance.
B. RNAi-knockdown of *dyn-1*, resulted in loss seam cells expressing the *scm::gfp* reporter. RNAi was performed on animals carrying *scm::gfp* as an integrated array, ruling out the possibility that the absence of the reporter was due to mosaicism of an extrachromosomal array. In the worm shown, only three cells showed expression of *scm::gfp* (i, arrowheads). 'Seam cells' visible under DIC optics but which are not expressing *scm::gfp* are circled (ii).
C. Performing *dyn-1* RNAi on animals expressing the ARF-3::GFP reporter appears to lead to a reduction i the number of GFP puncta in seam (i) and muscle cells (compare wild-type (ii) with *dyn-1* RNAi (iii)).
Scale bars = 100µm.

supports the idea that *die-1* may directly regulated *arf-3* expression, as indicated by the results of the yeast 1-hybrid screen.

arf-3 may represent a link between sub-cellular trafficking and seam development

As discussed below, *arf-3*, a GTPase likely plays a role in regulating sub-cellular trafficking, together with other GTPases and GTPase regulators. To investigate this possibility, and the link between these roles and seam development, knockdown of two genes was performed; *arf-1* and *dyn-1*. *arf-1* is another GTPase, closely related to *arf-3*. In the case of knockdown of this gene, seam hyperplasia was observed (Figure 5.5A), reflecting the increase in seam cell number seen in *arf-3* RNAi worms. The similarity of members of the Arf family brings the risk of off-target RNAi effects, but whether the phenotype is a result of knockdown of *arf-1, arf-3* or both genes, the involvement of the GTPases in seam development is clear.

The *C. elegans* dynamin homologue, *dyn-1*, another GTPase, is known to function in vesicular trafficking (Yu, Odera et al. 2006). Intriguingly, knockdown of *dyn-1* by RNAi also affected seam cells in the worm; affected worms which survived to larval stages possessed very few seam cells expressing the *scm::gfp* marker (Figure 5.5B), suggesting either that expression of *scm::gfp* is directly down-regulated, or that the fate of the seam cells is perturbed by loss of *dyn-1* function, resulting loss of GFP expression as a consequence. Examination of ARF-3::GFP localisation in *dyn-1(RNAi)* worms revealed fewer puncta in both seam and muscle and a relatively diffuse distribution of the GFP (Figure 5.5C).

Discussion

Sequence within the *scm::gfp* reporter encodes *arf-3*, an ADP-ribosylation factor GTPase

Alignment and dissection of the *scm::gfp* transcriptional reporter have led to the identification of relatively small regions of sequence capable of driving seam-specific expression. Physically, these regions are closest to *arf-3* (lying 5' of the ORF and within the last intron of the gene), an ADP-ribosylation factor (ARF) which resides within the large Ras family of GTPases. It is worth noting that the physical association of these regions with *arf-3* does not necessarily mean that they are functionally associated with this gene; indeed, the *scm::gfp* promoter region also covers the first four exons of *col-115* and, given that enhancers are not always tightly physically associated with the genes they control, it is possible that these regions have a functional relationship with this gene. One approach to determine whether the enhancers identified here play a role in regulating *col-115* might be to determine whether their inclusion or absence in *col-115* translational reporter constructs lead to differing expression patterns.

Certainly, the difference in the expression patterns of *scm::gfp* and the ARF-3::GFP translational reporter could be taken as evidence that the former is not a transcriptional reporter of the latter and that the seam expression of *scm::gfp* represents the transcriptional profile of a different gene. However, an alternative interpretation is that this difference is due to the fact that the region of sequence which drives *scm::gfp* expression extends only up to the last intron of *arf-3* and does not include the rest of the protein coding region or, critically, upstream sequence. It is at least conceivable that *scm::gfp* therefore reflects *arf-3* expression – but only in part, as it includes only the enhancer within the last intron of this gene.

Seam expression of the *scm::gfp* reporter is driven by an enhancer element lying within the last intron of *arf-3*. However, analysis of the sequence 5’ of the *arf-3* start codon revealed

that the intergenic region between *unc-24* and *arf-3* (which is not encompassed by the *scm* promoter) is also capable of driving seam expression. Thus, as with the seam-specific enhancer modules in *rnt-1*, *arf-3* expression appears to be controlled by redundant enhancers (bearing in mind the caveat mentioned above; it would be worthwhile further dissecting this 'intergenic' region to reveal exactly where the seam enhancer lies, relative to *F57H12.10* and *arf-3*), reflecting perhaps the evolutionary mechanisms by which enhancer modules are generated or move around the genome, or perhaps the need for robustness in the expression of this gene. Either way, these findings highlight the importance of examining intronic as well as 5' sequence for function if a comprehensive understanding of upstream regulators is being sought. Furthermore, the revelation of these two additional seam-specific enhancer regions should prove useful data, together with the other seam modules found in this study, for working out whether there is a motif or set of motifs commonly employed to drive seam expression of genes; understanding this better should not only yield information on genome evolution but have predictive power, and provide an alternative, computational route for identifying candidate genes involved in seam cell development and function.

The *scm::gfp* reporter, and possibly *arf-3*, is linked with the stem-like fate of seam cells

That *arf-3* is so strongly expressed in the stem cell-like seam cells is intriguing and in itself raises the question of whether this gene plays a role in an aspect of the stem cell-like fate. Furthermore, the loss of *arf-3::gfp* expression from a subset of seam cells in *bro-1;rnt-1* animals mirrors the loss of *scm::gfp* observed *in bro-1, rnt-1, pal-1* and *elt-1(RNAi)* animals and is strongly suggestive of a link, whether direct or indirect, between these genes and *arf-3* transcriptional regulation. Moreover, these genes have been established to have important roles in determining the seam fate (this study; Nimmo, Antebi et al. 2005; Kagoshima, Nimmo et al. 2007), thus linking *arf-3* with the maintenance of stem cell characteristics. Importantly however, just as *scm::gfp* expression does not appear to be tightly correlated

with proliferative potential (as discussed in Chapter 3), it may be that *arf-3::gfp* expression can be similarly uncoupled; examining expression of the reporter throughout the development of, for example, *elt-1(RNAi)* animals, where division failures can be observed to occur, may shed light on this question. Thus, whilst there is evidence for a relationship between seam cell fate and *arf-3*, the nature of this relationship remains somewhat unclear.

The fact that *arf-3* knockdown in the seam has an effect on seam cell development, together with the seam defects seen in *arf-3* mutants, lends further weight to the link between this gene and stem cell regulation. It is worth considering that the high degree of conservation between Arf family members (see below) means that there is a risk of off-target effects when one of these genes is knocked down by RNAi. There is conflicting data concerning the expression of other members of the Arf family. When reporters were made for the three *C. elegans* Arfs, and the Arl genes *arl-1* and *arl-3*, only *arf-3* was found to be expressed in the seam (Li, Kelly et al. 2004), suggesting that, at least in the context of seam development, the effects of multiple Arfs being knocked down inadvertently are likely to be limited. However, a more recent study found *arf-1* to be expressed in a much wider range of tissues, including seam cells (Hunt-Newbury, Viveiros et al. 2007). Either way, whether the relevant Arf is *arf-1* or *arf-3*, knowledge of involvement of these GTPases in the seam is significant and allows for the postulation of a mechanism (see below) which has not, as yet, been put forward in the context of the seam cells.

The increase in variation in seam cell number caused by *arf-3* RNAi could be explained by this gene playing a role in the choice between symmetric versus asymmetric cell divisions, as described below. Given that most of the divisions during larval development are asymmetric (apart from the first L2 division, which is symmetric and sees each seam cell produce two daughters which both retain the seam fate, divisions in all other larval stages are asymmetric), it is conceivable that perturbing the decision between asymmetric and symmetric division might have the effect of increasing overall seam cell number, as has been shown to be the case following *arf-3* RNAi.

Exactly which genes directly regulate *arf-3* expression in the seam remains to be established. There are no perfect matches to RUNX (RACCRCA) (Kamachi, Ogawa et al. 1990), GATA (WGATAR) (Ko and Engel 1993; Merika and Orkin 1993) or PAL-1 (TTTATG) (Dearolf, Topol et al. 1989; Maurer, Chiorazzi et al. 2007) consensus sites within the two seam-specific enhancers identified. There is, however, within a well-conserved region of the *arf-3* last intron, a match to a recognised permutation of the RUNX binding site (TCAGTGGTAC, or, reverse complemented, GTACCACTGA (Sandelin, Wasserman et al. 2004). Determining the veracity of this site will require further experimentation.

die-1, identified as a candidate regulator from the yeast 1-hybrid screen, is expressed in the seam (Heid, Raich et al. 2001) and although *die-1* has been studied principally in relation to dorsal intercalation, seam defects have been observed, with embryonic seam cells 'pinching out' of the line during late embryonic development. Furthermore, larval developmental defects were not examined in detail, leaving open the possibility of later role for *die-1* in the seam.

GTPases, their regulation and stem cell divisions

It is the unifying GTPase domain which is central to the function of Ras family members. The activity of Ras GTPase proteins is determined by whether they are bound by GTP, in which case they are in the 'active' state, or by GDP, which renders them inactive. The transition between GTP and GDP is facilitated by the GTPase of the protein itself, and by interactions with two antagonistic types of regulator. Guanine exchange factors (GEFs) allow bound GDP to be exchanged for GTP whilst GTPase activating proteins (GAPs) increase the otherwise negligible hydrolysis activity of the GTPase domain, promoting conversion of GTP to GDP (Donaldson and Jackson 2011).

The ARF family, the most divergent group within the Ras superfamily, can itself be subdivided into three groups; Arf proteins, Arf-like (ARL) proteins and the highly diverged Secretion-Associated and Ras-related proteins (SAR) group. *arf-3* lies within the first of these three groups. The importance of this family of genes is reflected by the presence of representatives throughout the eukaryotic kingdom, with a strong correlation between organismal complexity and the number of Arf family members (Li, Kelly et al. 2004). In mammals, the ARF family comprises six members, which are grouped into three classes based on their phylogenetic relationships: Class I, containing ARF1, ARF2 and ARF3, Class II, containing ARF4 and ARF5, and Class III, with ARF6. In *C. elegans*, each of the three classes is represented, but possesses only one homologue. As in the case of RUNX genes, the divergence of the Arf gene family in higher organisms makes *C. elegans* an attractive, relatively simple system in which to study these genes. Somewhat confusingly though, the nomenclature of *C. elegans* ARF genes does not reflect their relationships with their mammalian counterparts. Thus, whilst the *C. elegans arf-1* is most closely related to Class I mammalian Arfs, *arf-3* is nested within the Class II Arfs and *arf-6*, within Class III.

Although existing knowledge of Arf function in *C. elegans* is sparse, the high degree of conservation within the Arf family (Donaldson and Jackson 2011) makes it worthwhile considering what is known about mammalian Arfs. In the context of *arf-3*, the details of how ARF4 and ARF5 operate are likely most relevant; such clues should form the foundations of understanding the relevance of these genes to the seam cells, and to stem cell biology.

One important role for ARF4, recently elucidated, is to facilitate the transport of rhodopsin in rod cells via vesicles which travel from the inner to the outer segment of the rod (Deretic, Williams et al. 2005), though the detailed mechanism is still unclear. ARF4 and ARF5 are also both involved in secretion, recruiting proteins to the *trans*-Golgi network and thereby establishing the complexes required to ferry proteins to the cell membrane (Sadakata, Shinoda et al. 2010). More generally, Arfs have been shown to be involved in the recruitment of coat proteins for vesicle formation, recruitment and activation of membrane-modifying

enzymes in a variety of ways, making them established key regulators of membrane trafficking and organelle structure (Donaldson and Jackson 2011).

In keeping with these roles, *arf-3*, expressed in the seam cells, appears to localise to granules within the cytoplasm, likely associated with membranes such as the endoplasmic reticulum or Golgi apparatus. Currently the dynamics of these granules are unknown, but observation of their mobility (or lack of it) over time may shed more light on their precise function.

Clues as to how *arf-3* may be involved in stem cell regulation can perhaps be gleaned from recent progress in linking the cellular functions of Arf GTPases and their regulators to the developmental context of stem cell divisions. Specifically, *cnt-2*, identified as a GAP, has been found to have an important role in determining the symmetry of divisions of the Q.p neuroblast (Singhvi, Teuliere et al. 2011). In wild type worms, this cell divides asymmetrically, producing an anterior daughter, which has the ability to divide further, yielding three neuronal cells, and a posterior daughter which undergoes programmed cell death. In *cnt-2* mutants, the division of Q.p is symmetrised, resulting in two cells with proliferative ability, and thus double the number of neurons being produced. As well as demonstrating the GAP activity of *cnt-2*, the authors also showed that the division defect in *cnt-2* mutants was coupled with defects in endocytosis, implicating this process in the developmental decision of division symmetry. Furthermore, the involvement of a GAP points to the involvement of a GTPase (presumably regulated by this GAP). As yet, the GTPase which is assumed to be regulated by *cnt-2* has not been identified. *arf-1* and *arf-6* were both dismissed as the relevant Arfs in the Q.p division (Singhvi, Teuliere et al. 2011). As for *arf-3*, the embryonic lethality of *arf-3* RNAi made it impossible to examine the Q.p division, and no defects were observed in the sick *arf-3(tm1877)* early larvae. As the authors point out, this latter observation could be accounted for by the ‘rescuing’ effect of maternally contributed *arf-3* RNA or protein. Alternatively, redundancy between Arfs could mean that two or more must be knocked down before their roles in Q.p division symmetry are fully exposed.

Of course, the work described above is based on division symmetry in a lineage which is distinct from the seam. Nevertheless, developmental mechanisms are frequently reused in different cell types; given the similar need of seam cells for a mechanism for determination of division symmetry, it is at least possible, if not likely, that Arfs and their regulators are operating here too in a similar way, especially given the apparent involvement of *arf-3* in some aspect of seam fate determination.

Asymmetric subcellular localisation of components of the WNT signalling pathway is fundamental to regulation of cell division symmetry (Thorpe, Schlesinger et al. 2000; Michael A 2002; Mizumoto and Sawa 2007). Specifically, the *C. elegans* TCF/LEF homologue, POP-1, is distributed asymmetrically during asymmetric seam cell divisions. In the posterior nucleus, destined to retain the stem-like fate, nuclear POP-1 levels are low, whilst in the anterior nucleus relatively high levels of POP-1 are maintained (Phillips, Kidd et al. 2007). In turn, this is regulated by the asymmetric distribution of the β-catenin WRM-1 which, working together with LIT-1, phosphorylates POP-1 and causes it to be exported from the nucleus (Phillips, Kidd et al. 2007). Thus, posterior daughters have high levels of nuclear WRM-1, and therefore low levels of nuclear POP-1, whilst in anterior daughters WRM-1 is found to be cortically localised, allowing POP-1 levels in the nucleus to rise. Although this mechanism has been described, the control of WRM-1 distribution is less well understood, but presumably involves intracellular trafficking machinery. The recent finding that a phospholipase, *ipla-1*, plays an important role in determining spindle orientation, WRM-1 localisation and division symmetry supports this idea (Kanamori, Inoue et al. 2008). *ipla-1* is thought to operate by regulating the balance between lysosomal degradation of WRM-1 and retrograde trafficking of this protein, from the endosome to the Golgi. This balance impinges on subcellular distribution of WRM-1 and thus on division symmetry.

These findings are highly pertinent to the results presented here, which show that *arf-3*, a GTPase which likely influences intracellular trafficking, is also implicated in the regulation of division symmetry. Further examination of the *C. elegans* Arf family members, *arf-1* and *arf-*

6, together with known regulators of these genes, in particular *cnt-2*, will be important in revealing the link between *arf-3,* seam cell divisions and known components of the symmetry determination pathway.

Conclusion

The *scm::gfp* reporter is an established and widely-used marker of the stem cell-like seam cells. The location and endogenous function of enhancers within it, together with its relationship to the development and identity of the seam cells has, however, not been investigated. Here it is shown to include a seam-specific enhancer which lies within the last intron of the Arf family member *arf-3*.

Despite the fact that *scm::gfp* is commonly used as a seam cell marker, here (and in Chapters 3 and 7) it has been shown that the expression of *scm::gfp* can be uncoupled from the stem, proliferative fate of seam cells. When studying seam development, it is therefore important to bear in mind the fact that, depending on one's definition of 'seam cells', the number of *scm::gfp*-positive nuclei does not necessarily reflect the number of seam cells in the worm.

arf-3 appears to operate at a point downstream of a number of genes which function in the seam, including *rnt-1, bro-1, pal-1* and *elt-1*. Thus it should not be treated as an independent marker of seam fate but as a part of the network required to direct seam development, linked to, but not perfectly correlated with, the seam fate. This knowledge should make the *scm::gfp* marker more useful and seam cell studies more informative.

In addition, dissection of the *scm* promoter has reinforced the importance of examining intronic sequence, not just regions 5' of the ATG, for regulatory elements. Furthermore, these findings provide two more examples of seam-specific enhancers, which should help further our understanding of how promoter and enhancer regions evolve and operate.

Finally, dissection of the scm marker has revealed the involvement of *arf-3* in the seam; this gene has the potential to play a major role in determining the symmetry of seam cell divisions and thus is potentially of significant interest to seam cell – and stem cell – biology.

CHAPTER 6

Spatial regulation of *rnt-1*

Introduction

RNT-1 plays a core role in seam cell development in the worm and conservation of its function as a regulator of stem cell proliferation and differentiation is well established (Nimmo, Antebi et al. 2005; Kagoshima, Nimmo et al. 2007). Understanding how *rnt-1* itself is regulated is therefore essential for revealing both where *rnt-1* lies in the genetic networks which underlie stem cell biology, and how these networks function, thereby shedding light on the topology and components of these networks in higher animals. Only by piecing together the manifold and complex interactions between individual genes can the overall workings, input and output of the system be understood.

In the worm, *rnt-1* is regulated spatially and temporally; expression is seen only in the seam and muscle (Nimmo, Antebi et al. 2005) and RNT-1 levels are tightly correlated with both cell cycle progression and developmental stage (A. Woollard and T. Braun, pers. comm.). Whilst the mechanisms of temporal regulation have been studied in detail, the spatial aspects of *rnt-1* regulation have received less attention. Indeed, even the location of the true 'promoter' of *rnt-1* is incomplete and contentious; how transcription factors might bind to the promoter or to enhancer elements to drive tissue-specific expression is far from clear.

In addition to the general benefits of identifying and analysing upstream regulators of *rnt-1*, outlined above, there are more specific advantages to gaining a greater understanding of the mechanisms of transcriptional regulation of *rnt-1*. Given that *rnt-1* works closely with *bro-1*, and both genes are co-expressed in – and only in – seam and muscle, the evolution of the regulation of these genes' expression patterns presents an intriguing question. Are common

mechanisms, perhaps even modules of sequence, responsible for their regulation, following co-evolution or perhaps even duplication of promoter and enhancer elements? Alternatively, is there little similarity between such elements in *rnt-1* and *bro-1,* either due to independent evolution of promoter elements and networks or flexibility of DNA-transcription factor interactions and subsequent sequence drift? To shed light on these questions, the basis of *rnt-1* transcriptional regulation has been examined.

Results

The *rnt-1* long intron contains distinct muscle and seam enhancer regions

Although the *rnt-1* ORF is less than 1kb in length, the entire gene is considerably larger (12.5 kb), owing in part to being interrupted by introns but, more significantly, due to the presence of a large region of sequence (7.2kb) between the 5' UTR and the start codon, henceforth referred to as the 'long intron' of *rnt-1* (Figure 6.1A). It has been established that the entire sequence between the *rnt-1* start codon and the end of the upstream gene, *hil-5*, is able to drive the normal expression pattern of *rnt-1* (Nimmo, Antebi et al. 2005). To further narrow down the regions required for the endogenous expression of *rnt-1*, stretches of sequence 5' of the ATG were cloned into the minimal promoter construct, pPD107.94, which contains the *pes-10* minimal promoter sequence. Initially, as show in Figure 6.1B, the entire intron was used. This drove expression in three tissues: body wall muscle, intestinal cells and the seam.

Subsequently, smaller regions of the intron – based on regions previously reported to drive tissue-specific expression patterns (Nam, Jin et al. 2002) – were tested for their ability to drive expression. The results are summarised in Figure 6.1B. There are several principal findings.

Firstly, the intron contains elements which drive both muscle and seam expression, but these are distinct and independent 'modules'. Secondly, expression is observed in intestinal nuclei (Figure 6.2Ai and ii). Thirdly, the expression patterns driven by these parts of the *rnt-1* intron are distinctly different from those previously reported (Nam, Jin et al. 2002). Specifically, the region identified here as driving muscle expression has previously been reported to drive seam expression. Conversely, the 3' end of the intron, which has been shown here to drive strong seam expression, has been reported by the same group to

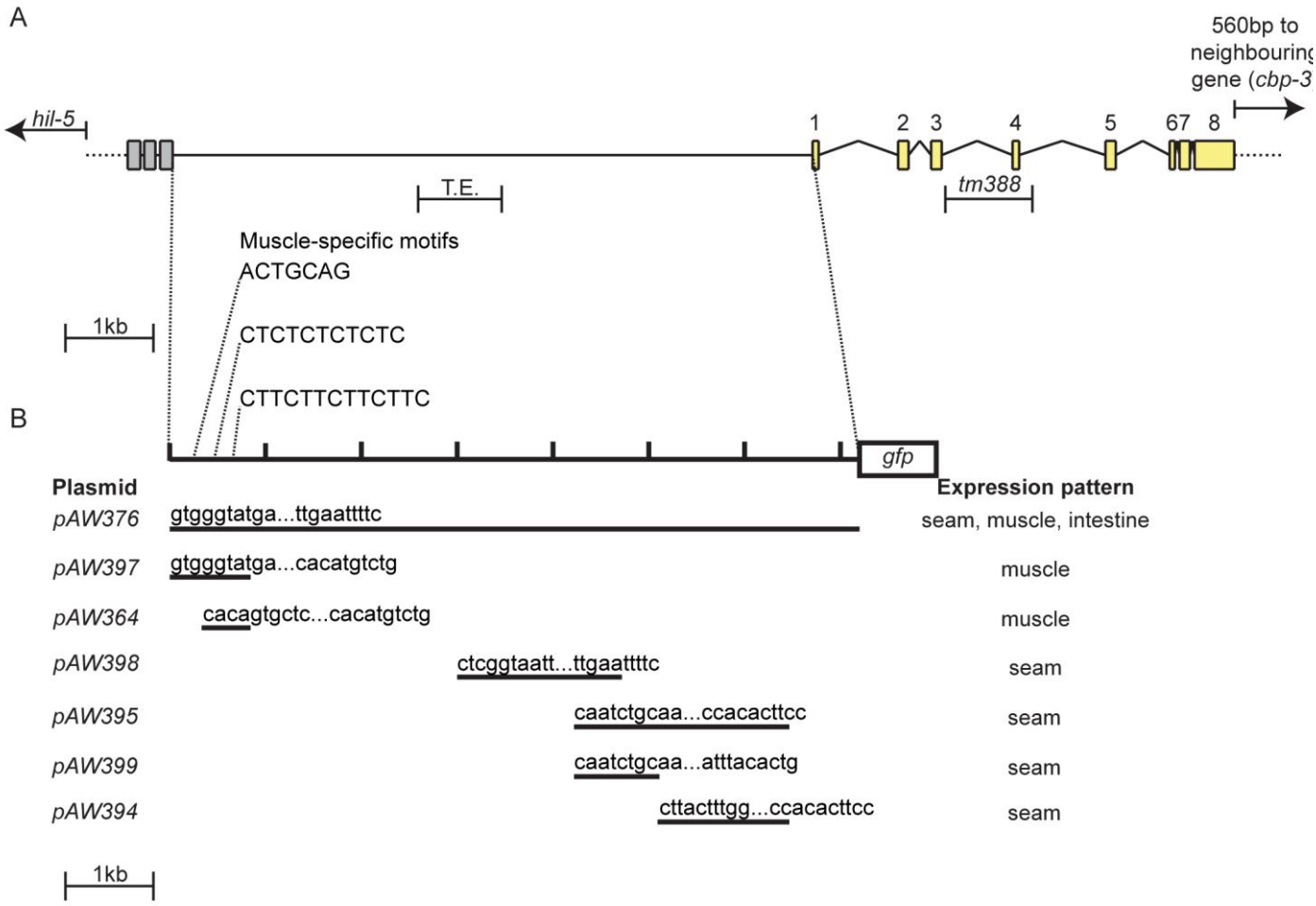


Figure 6.1 A. Schematic of *rnt-1*, with 5' UTR (grey shading). The location of the *tm388* deletion is shown. T.E. = putative transposable element. The positions of recognised muscle motifs (Zhao, Schriefer et al. 2007) within the *rnt-1* long intron are shown. B. Dissection of the *rnt-1* long intron was performed by cloning fragments of the intron into the minimal promoter vector *pPD107.94*. The location and sizes of the fragments are shown, together with the names of the resultant plasmids. The expression pattern driven by each fragment is shown.

promote intestine-specific expression, whereas here this region of the promoter drives seam expression.

Putative tissue-specific 'enhancer modules' are present within the *rnt-1* long intron

The mechanisms by which tissue-specific expression patterns are determined have been studied in numerous different systems. One theme to arise from these efforts is that in higher organisms (Kirchhamer, Yuh et al. 1996; Arnone and Davidson 1997) and in *C. elegans* (Jantsch-Plunger and Fire 1994), transcription factors frequently work together, binding at neighbouring sites arranged in clusters or modules, which have the potential to act in isolation and direct expression in a specific tissue type. Increases in the power and reliability of the predictive tools available have led to more information on the identity of these modules. Recently, a number of *C. elegans* muscle-specific modules were identified based on their occurrence within the promoters or introns of muscle-expressed genes (Zhao, Schriefer et al. 2007). Furthermore, their occurrence was tested and proven to be predictive of muscle expression. Strikingly, several of these modules are found within the *rnt-1* intron, within – and only within – the regions shown here to drive muscle expression: CTCTCTCTCTC (Figure 6.1A and B, pAW397); ACTGCAG and CTTCTTCTTCTTC (Figure 6.1A and B, pAW397 and pAW364). Confirmation of the functionality of some or all of these putative modules would require further work, either involving deletion of these sites or demonstration of their ability to drive muscle expression out of context. However, even the occurrence of these stretches of sequence specifically within the relatively small muscle 'module' of *rnt-1* is strongly suggestive of their significance. The presence of three muscle motifs, each of which has been shown to be capable of driving muscle expression, raises the additional issue of redundancy.

As yet, no comparable stereotypical modules have been identified for seam expressed genes, making it difficult to determine whether *rnt-1* seam expression depends on common

motifs, shared by numerous other seam-expressed genes. What is apparent, though, is that seam expression is driven by at least two separate, non-overlapping regions within the promoter (represented by plasmids pAW398 and pAW394). Based on this finding, it is at least plausible, if not likely, that *rnt-1* expression is driven by additional, as yet unidentified enhancers which lie within – and perhaps even outside – the long intron. From the point of view of trying to define a consensus sequence (or collection of sequences) associated with and capable of driving seam expression, this information should be useful; further investigation on this aspect of *rnt-1* function is likely worthwhile.

The *rnt-1* long intron contains a putative transposable element

The modular arrangement of the *rnt-1* intron, with its distinct 'muscle' and 'seam' promoter regions, raised the question of whether the same modules were employed in other genes in the genome. To address this question, the *rnt-1* intron was aligned with the *C. elegans* genome. No large regions of homology were identified, with the exception of a region of sequence, just under 1kb in length, showing 100% nucleotide sequence conservation with several regions of the genome (see Appendix). Significantly, this region is delimited by two inverted terminal repeats, 29bp in length, indicating that it is, or is derived from, a transposable element.

In order to determine the nature of this element, the sequence was aligned with a database of repeat elements (Kohany, Gentles et al. 2006). No exact match was found, but similarities were identified between the queried *rnt-1* sequence and class II, DNA transposons. This region of the *rnt-1* intron, therefore, is likely an as yet uncharacterised DNA transposon.

Alignment of the element with the *C. elegans* genome revealed multiple almost exact matches on each of the chromosomes. Interestingly, however, no such matches were found in the genomes of *C. briggsae, C. brenneri, C. remanei* or *C. japonica*. Furthermore, a

BLAST search against the NCBI nucleotide database yielded only *C. elegans* hits, suggesting that the element has most likely arisen after *C. elegans* diverged from its other Caenorhabditis relatives.

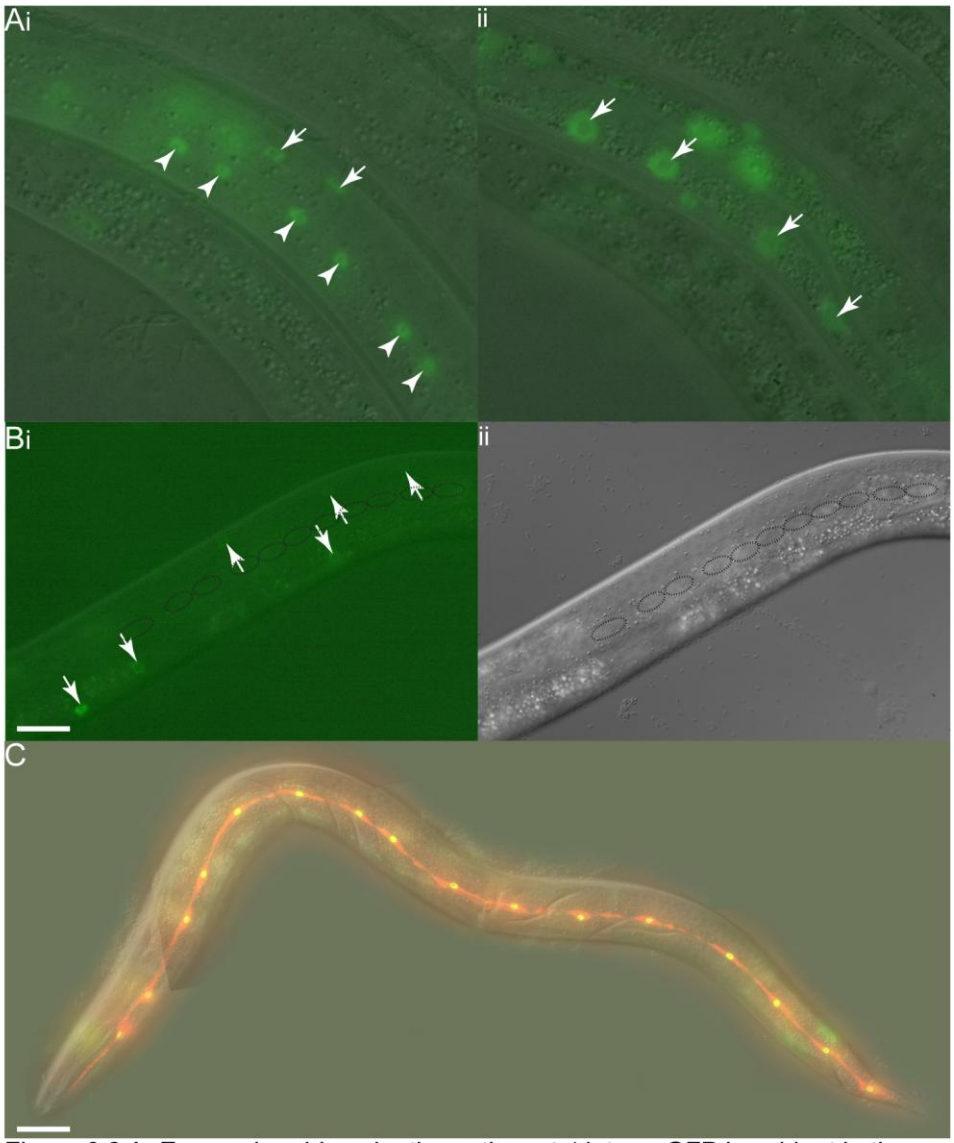


Figure 6.2 A. Expression driven by the entire *rnt-1* intron. GFP is evident in the seam (arrowheads, Ai) and body wall muscle (arrows, Ai), as well as intestinal nuclei (arrows, Aii). B. The 'muscle' region of the *rnt-1* intron is capable of driving GFP expression in body wall muscle cells (arrows, Bi), but not in seam cells (outlined in merged image, Bi, and Nomarski image, Bii). Scale bars for A and B = 50μm.
C. Seam expression, as driven by one of the regions within the 'seam' module of the *rnt-1* promoter (*pAW399*). The seam cells are marked with *scm::rfp*, with co-localisation of the seam cell marker RFP and the *rnt-1* promoter GFP evident as yellow dots in seam nuclei. Scale bar = 25μm.

Discussion

The *rnt-1* long intron contains modules which drive seam, muscle and intestinal expression

Dissection of the *rnt-1* intron which lies between the 5'UTR and the first exon revealed that it is sufficient for driving the endogenous expression pattern of *rnt-1*; expression is observed in the seam lineage and in body wall muscle, matching the expression seen with the full length rescuing construct (Nimmo, Antebi et al. 2005; Kagoshima, Nimmo et al. 2007). Within the intron, two broad regions can be demarcated as driving expression in these tissues; the 5' end is responsible for muscle expression and the 3' end for seam expression. Significantly the region identified here as driving muscle expression was previously reported as driving seam expression (Nam, Jin et al. 2002).

The discrepancy between the expression patterns observed in the two labs is paradoxical given that in both cases the same, distinct 'modules' of the *rnt-1* promoter have been identified, albeit identified as driving expression in different tissues. Perhaps the most likely cause of this difference is the misidentification of the cells in which GFP was observed. There are two lines of evidence which support this possibility. Firstly, the micrographs in (Nam, Jin et al. 2002) which are described as showing seam expression of various *rnt-1* translational reporters appear instead to show muscle expression. This would account for the fact that the 'seam' expressing region of the *rnt-1* promoter (Nam, Jin et al. 2002) is found here to drive muscle expression. Secondly, the determination of cell type in this study was performed either with the use of tissue-specific markers (Figure 6.2C) or when cells were clearly visible using DIC optics, allowing their fate to be unambiguously determined (Figure 6.2Bi and ii).

In addition to driving expression in seam and muscle, the *rnt-1* intron promotes intestinal expression. As yet, it has not been possible to identify a specific region within the intron which drives this intestinal expression. Perhaps the most likely reason for this is that the

region responsible for intestinal expression spans one or more of the clones used to dissect the *rnt-1* promoter and was not left intact when regions of the intron were isolated. Similarly, it is plausible that separate, non-adjacent regions of the intron are required combinatorially for intestinal expression; again, the cloning strategy adopted would likely have disrupted the necessary context if this were the case.

Nevertheless, the fact that intestinal expression is observed in both transcriptional (this study) and translational (Nam, Jin et al. 2002) *rnt-1* reporters raises several important questions. One is why other studies have found no evidence of intestinal expression from independently-constructed *rnt-1* reporters (Nimmo, Antebi et al. 2005; Kagoshima, Nimmo et al. 2007). A possible explanation for the discrepancy between the expression patterns observed in the two labs had been that the reporter construct used differed slightly, with the 'muscle/seam' reporter including a small region of 3'UTR absent from the 'seam/intestine' reporter (Nimmo, Antebi et al. 2005). Perhaps the inclusion of the 3'UTR in the original *rnt-1* construct (which likely reflects the endogenous expression pattern of *rnt-1* most accurately) leads to down-regulation of *rnt-1* transcript in the intestine.

Interestingly, the *rnt-1* long intron has also been postulated to contain sites which allow autoinhibitory regulation of *rnt-1*, mediated by BRO-1 (Shim and Lee 2008). There is compelling evidence to support a *rnt-1*-repressive role for BRO-1 (Kagoshima, Nimmo et al. 2007), but the mechanism is unclear. Coupled with two-species comparative sequence analysis, a series of deletion experiments was used to suggest that one site in particular ("site b") within the long intron is key for the autorepression of *rnt-1* transcription during larval and adult development. *bro-1* RNAi is reported to have led to significant upregulation of *rnt-1* expression during larval stages but this effect was dependent on the presence of site b within the intron (Shim and Lee 2008). What is not clear from the experiments, however, is whether this site is merely important for *rnt-1* expression, rather than specifically *rnt-1* de-repression mediated by *bro-1* RNAi knockdown. Indeed, site b falls within the regions shown here to drive seam expression and, in the experiment where site b was specifically deleted,

upstream sequence of the intron (which, in this study, is shown to redundantly drive seam expression – Figure 6.1) is missing. Clarifying this situation should reveal the importance of these regions of sequence and their relevance for seam expression and, perhaps, *rnt-1* autoinhibition during larval stages.

A role for *rnt-1* in the intestine?

It remains unclear whether endogenous *rnt-1* is down-regulated completely in the intestine, such that it is essentially not expressed at all, or whether it is just regulated so that its levels are below the threshold of detectability with full-length GFP reporter constructs? The former is of course possible and a newly developed *rnt-1* antibody may shed light on this question (P. Appleford, pers. comm.), but given the interaction between *rnt-1*, the seam cells and nutritional status (T. Braun, N. Saad and A. Woollard, pers. comm.), the latter possibility is an intriguing one.

rnt-1 mutants are particularly sensitive to starvation; recovery from starvation-induced L1 arrest and subsequent survival during larval development and adulthood are dramatically reduced in *rnt-1* mutants, compared to wild type worms. The mechanisms underlying this relationship are unclear, but a recent study was performed to determine whether this role of *rnt-1* is dependent on its expression in the seam, or in muscle (B. Shaw, pers. comm.). The *rnt-1* gene was placed under the control of the endogenous seam- and muscle-specific regions of the long intron and the resultant constructs injected separately into *rnt-1* mutant worms, which were then assayed for their ability to survive recovery from L1 arrest. As a positive control, the full length *rnt-1* promoter was used. This was found to effect full length rescue of the 'starvation recovery' phenotype. In contrast, muscle-expressed *rnt-1* did not rescue *rnt-1* mutants. Strikingly, expression of *rnt-1* from the seam-specific promoter region did rescue the starvation phenotype, demonstrating that it is seam expression of *rnt-1* which is required for the transition from arrested L1 to normal larval development (B. Shaw, pers.

comm.). Here, *rnt-1* is likely operating in a manner independent of its role in regulating seam cell division *per se*.

However, the finding that the *rnt-1* intron drives intestinal expression raises an important caveat to these findings. The region used to express *rnt-1* in the seam was the entire promoter region minus the 5' stretch of sequence which has been identified as driving muscle expression. Thus, it is likely that the sequence used included the enhancer regions which are now known to be capable of driving intestinal expression. It is therefore possible that the rescuing of the '*rnt-1* starvation phenotype' was in fact due to low levels of *rnt-1* in the intestine. Dissection of the promoter to reveal the region required for intestinal expression and the use of any such region in this context could provide a valuable insight into potential roles of *rnt-1* outside the seam, in a tissue in which no function for *rnt-1* has been previously described. If endogenous *rnt-1* really is expressed in intestinal nuclei, the link between nutritional status, the gut and a gene hitherto only thought to be expressed in seam and muscle would be much easier to rationalise.

A role for *rnt-1* in body wall muscle?

The reason for *rnt-1* expression in body wall muscle also presents an interesting question. Examination of muscle cells in *rnt-1* and *bro-1* mutants, using the body wall muscle marker *myo-3::gfp*, revealed no abnormalities (Nimmo, Antebi et al. 2005). Of course, it is possible that muscle expression of *rnt-1* and *bro-1* is an evolutionary remnant; perhaps they were once expressed in muscle to fulfil a function in this tissue, only for this function, but not the expression pattern, to be lost or taken up by another gene during the evolution of development. Alternatively, the muscle enhancers in *rnt-1* and *bro-1* could have been incorporated into these genes serendipitously, perhaps in the same way as the transposon appears to have become incorporated into the long intron of *rnt-1*. Such a coincidence, though, seems unlikely, given that *rnt-1* and *bro-1* work closely together to regulate the

transcription of target genes in the seam and are also both expressed in muscle, apparently under the direction of independent enhancer modules. Furthermore, the fact that in both cases the seam and muscle enhancers are physically separate effectively rules out the possibility that degeneration of the muscle-specific enhancers would be constrained by strong selection pressure for maintenance of binding sites for seam expression of *rnt-1* and *bro-1*. The most probable explanation for muscle expression of these two genes, therefore, is that they perform a function in this tissue, but that they act redundantly with one or more additional genes. As discussed in Chapters 4 and 7, *unc-62* and *pal-1* are strong candidates for redundant partners which could work together in the muscle with the RNT-1/BRO-1 complex.

The overall picture which is emerging is that *rnt-1* has multiple interactions and functions, most of which are currently relatively poorly understood. Recently, evidence of an additional interaction – with the TGFβ signalling pathway – has been put forward (Ji, Nam et al. 2004; Ji, Singaravelu et al. 2005; Shim and Lee, 2008). The initial link between the TGFβ pathway and *rnt-1* was made on the basis of gross phenotypical similarities, particularly with respect to the male tail phenotypes of *dbl-1* and *rnt-1* mutants. It is worth noting, however, that the tail phenotypes of the two mutants are by no means identical; whilst ray fusion is the predominant phenotype in TGFβ pathway mutants (reviewed in Ji, Nam et al. 2004), fusion is a low penetrance aspect of the *rnt-1* phenotype, with the complete absence of rays being much more common (Nimmo, Antebi et al. 2005). Nevertheless, this line of investigation has led to the assertion (Shim and Lee, 2008) that DBL-1 acts as an indirect activator of *rnt-1*, though only during larval development, and thus presents yet another potential *rnt-1*-interaction. Exactly how intimate the link between the TGFβ pathway and *rnt-1* is remains to be revealed; it could be that the two genes occupy relatively distant – but nevertheless linked – positions within their developmental pathways. So far, it has been shown that upregulation of *dbl-1* leads to increases in the levels of RNT-1::GFP (Shim and Lee, 2008), but the question of how this is brought about has not been answered. Furthermore, the phenotypes

of double mutants, carrying mutations in *rnt-1* and TGFβ family members, are not suggestive of these genes acting in a linear pathway, but rather of the existence of parallel pathways (Ji, Singaravelu et al. 2005).

Given the apparent links between *rnt-1*, nutritional status and recovery from starvation, an interaction with the TGFβ pathway, which plays a role in the worm's perception of unfavourable conditions and developmental decision as to whether to enter the dauer state, is intriguing and appears worthy of further investigation.

Another potential interaction of interest is that between *rnt-1* and the WNT signalling pathway. Clearly, there must be developmental integration of cell division and division symmetry, requiring some degree of interaction between *rnt-1* and components of the WNT division symmetry pathway such as *pop-1*. However, it has been suggested that *rnt-1* itself plays a role within the WNT symmetry pathway (Ji, Singaravelu et al. 2005), based on the T cell defects observed in *rnt-1* mutants, where the asymmetry of the T cell division appears to be lost (Kagoshima, Sawa et al. 2005). An alternative view of the T cell phenotype though is that it results merely from cell cycle arrest, which blocks the subsequent differentiation of posterior T cell daughters; this interpretation therefore leaves *rnt-1* with a role confined to cell cycle progression, rather than division symmetry *per se* (Kagoshima, Nimmo et al. 2005). Either way, as noted above, there must still be co-ordination between these two pathways, the details of which remain to be unravelled.

The relationship between function and sequence conservation is not infallible

The enhancer which drives expression of *bro-1* in the seam was identified through a comparative approach, as detailed in Chapter 3. However, in the case of the enhancers which direct expression of *rnt-1*, their identification entailed a different approach; the entire region known to be required for *rnt-1* expression was physically dissected, clone-by-clone,

and each fragment tested for the ability to drive expression. The reason for the difference in approaches is that the *rnt-1* intron shows relatively low levels of conservation with other Caenorhabditis species, with the exception of one region, just under 300bp in length, which falls within the seam expression modules described here (Figure 6.1, pAW398, pAW395 and pAW399; see Appendix for alignment). Neither the seam nor muscle enhancer regions of the *rnt-1* intron – including the putative muscle-specific modules shown to be predictive of muscle enhancers (Zhao, Schriefer et al. 2007) – show pronounced conservation with sequence in the same intron in *C. briggsae, C. brenneri, C. remanei* or *C. japonica*. It could simply be that these other species do not exhibit the same expression pattern of *rnt-1* or, alternatively, they have the same pattern of endogenous *rnt-1* expression whilst the sequence within the *rnt-1* intron has diverged. Either way, this demonstrates that lack of sequence conservation even between closely related species is not necessarily indicative of functional insignificance; they may not be conserved, but these modules clearly do drive expression in seam cells and muscle cells, even when taken out of their normal context. Furthermore, the *rnt-1* intron also provides an example of sequence conservation analysis identifying an element which is likely not of significance for the functioning of the promoter – the transposon sequence. These findings highlight that, whilst often informative, sequence conservation is predictive of function, not proof, and is by no means always a reliable indicator.

The identification of relatively small (<2kb) endogenous modules which are sufficient to drive tissue-specific expression of *rnt-1* lays the foundations for further developing our understanding of the factors which lie upstream of *rnt-1* and regulate the spatial expression of this gene. With the aim of identifying the transcription factors which act on the *rnt-1* promoter, the enhancer regions identified have been cloned into suitable vectors to allow a yeast-1-hybrid screen using tiled transcription factor libraries to be performed, in collaboration with John Reece-Hoyes (University of Massachusetts). As well as potentially revealing the identity of proteins which drive *rnt-1* expression in the seam, this may also

shed light on the mystery surrounding *rnt-1* regulation in muscle, perhaps explaining how *rnt-1* operates in the context of a cell specification network other than that involved in seam cell development.

Conclusion

In summary, the long intron of *rnt-1* has been shown to contain all the necessary enhancers for endogenous spatial expression; seam, muscle and intestinal expression are all driven by this intron. Moreover, these enhancers appear to lie in a modular arrangement, perhaps reflecting the evolutionary process by which they arise and move around the genome. This intron, therefore, serves as the regulator of *rnt-1* expression, driving *rnt-1* activity at least two – and possibly three – tissues during development. Furthermore, the use of transcriptional reporters has raised the prospect that *rnt-1* may play a role in intestinal cells, in addition to muscle and seam cells, and broadens the scope of this gene's potential roles in the worm, particularly in the context of a known role in responding to the nutritional environment.

Now that enhancers have been identified within the *rnt-1* long intron, together with the expression patterns they drive, it should be possible to use a yeast one-hybrid approach to reveal the identity of the transcription factors which lie upstream of *rnt-1*, shedding light on the context in which *rnt-1* functions.

CHAPTER 7

The caudal homologue, *pal-1*, works both together with and independently of bro-1/rnt-1 in regulating seam development

Introduction

Whilst *rnt-1* and *bro-1* undoubtedly play important roles in the seam, they do not work alone. Firstly, there are apparently parallel pathways which regulate seam development, as illustrated by the fact that, even in *rnt-1* and *bro-1* single and double mutants, the L1 division of the seam cells is unperturbed. Furthermore, even though some seam divisions in *bro-1* and *rnt-1* mutants fail or occur abnormally, many proceed as in wild-type worms. Secondly, as with any gene or, more generally, any component of a network, *rnt-1* and *bro-1* depend for their function upon the genes which regulate them, and the genes which they regulate. Only by knowing the identity of the upstream and downstream genes, and the genes operating in parallel pathways, can one expect to gain comprehensive understanding of the system.

With this in mind, a small-scale RNAi screen was performed with a number of genes known to be expressed in the seam, with the aim of revealing either direct effects of these genes on seam cell number or interactions with *bro-1* and *rnt-1*. Seam cell numbers were counted in wild-type as well as *bro-1* and *rnt-1* mutant worms. It was, however, the unexpected finding of synthetic lethality between *rnt-1* and the caudal homologue, *pal-1*, which became the focus of this work.

The *C. elegans* homologue of the Drosophila homeobox gene Caudal, *pal-1* was first identified by the abnormal morphology of *pal-1* mutant male tails (Waring and Kenyon 1990;

Waring and Kenyon 1991). Subsequently, the *pal-1(e2091)* phenotype was found to stem from the role of this gene in the specification of cells derived from the C and D blastomeres of the early embryo (Hunter and Kenyon 1996). In addition, *pal-1* is required during the later stages of embryogenesis, both in the C and D lineages and in the AB and MS lineages, where it plays a crucial role in gastrulation, the ventral movement of the seam cells during early embryonic morphogenesis and for the formation of the rectum (Edgar, Carr et al. 2001). During larval development, *pal-1* has been shown to be required in the sensory rays of the male tail; in *pal-1* mutants the seam cells V5 and V6 give rise to ectopic alae, instead of tail rays, and it is from this 'Posterior ALae' phenotype that the gene's name derives (Hunter, Harris et al. 1999).

At the molecular level, *pal-1* is a transcription factor and has recently been shown to be regulated by the Mediator complex. This assortment of proteins is responsible for the recruitment of the DNA polymerase complex to the appropriate promoter, as well as integrating multiple regulatory inputs and communicating them to the polymerase (Zhang and Emmons 2000).

Whilst the function of *pal-1* in the V5 and V6 lineages is relatively well understood, this gene is not known to have roles in the rest of the seam lineage. Here, evidence is presented for such a role. Furthermore, *pal-1* appears to work redundantly with *bro-1* and *rnt-1* during development, adding context to the network in which these genes are known to work.

Results

The caudal homologue *pal-1* interacts with *rnt-1*

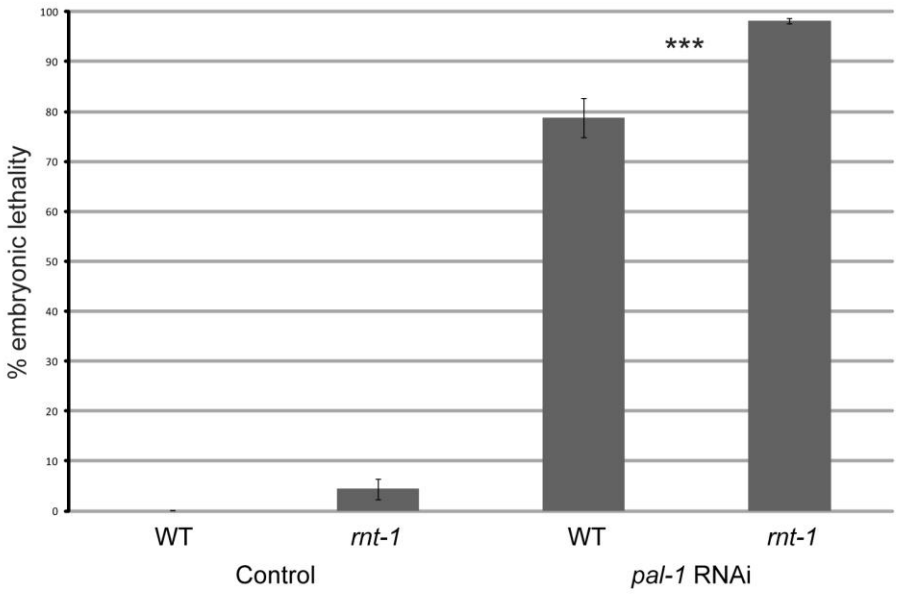
A small number of genes (see Appendix) were knocked down in worms carrying the integrated *scm::gfp* reporter, using RNAi by feeding, and the numbers of seam cells in treated animals were counted. To allow for potential interactions between these genes and *rnt-1* to be revealed, this approach was performed on *rnt-1* mutants, also carrying the *scm::gfp* marker.

In the cases of all genes tested by RNAi (RNAi knockdown was performed on wild-type worms for *nhr-72, nhr-74, nhr-77, ceh-22, nhr-75, nhr-89, ceh-16, nhr-81, nob-1, elt-6* and *pal-1*), the seam cell number was found to be similar to the wild-type number of 16. However, in the case of *pal-1*, particularly high levels of embryonic lethality were observed when knockdown was performed on *rnt-1* animals, suggestive of a synthetic interaction (Figure 7.1A). Knockdown of *pal-1* by RNAi causes high levels of embryonic lethality (79%) in otherwise wild-type worms, but this figure rises to 98% in *rnt-1* mutant worms. This effect is not just additive, as *rnt-1* mutants not subjected to *pal-1* RNAi show only 4% lethality.

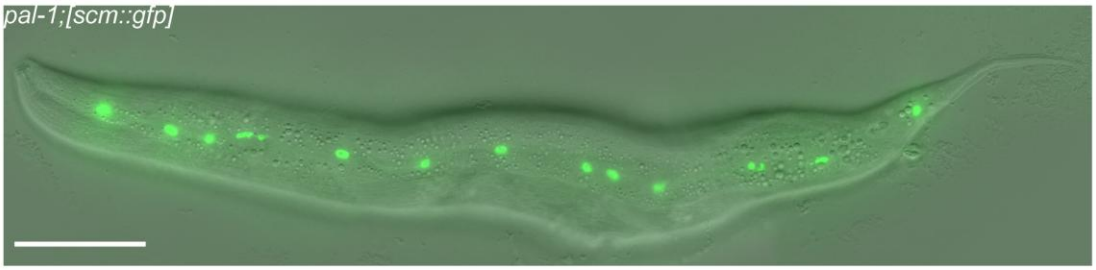
pal-1(e2091) mutants have fewer seam cells

To examine further the apparent relationship between *rnt-1* and *pal-1*, the *pal-1* mutant *e2091* was obtained. As discussed later, this is a hypomorphic rather than a null allele; complete loss of *pal-1* function leads to early embryonic lethality (Edgar, Carr et al. 2001). Given the known role of *rnt-1* in the seam, the *scm::gfp* array was crossed into the *pal-1(e2091)* background, enabling seam cell numbers to be assayed in worms carrying the *pal-1* mutation. In contrast to *pal-1* RNAi animals, which have WT seam cell numbers, *pal-1(e2091)* animals possess reduced numbers of seam cells compared to WT worms (Figure 7.1B and C).

A.



B.



C.

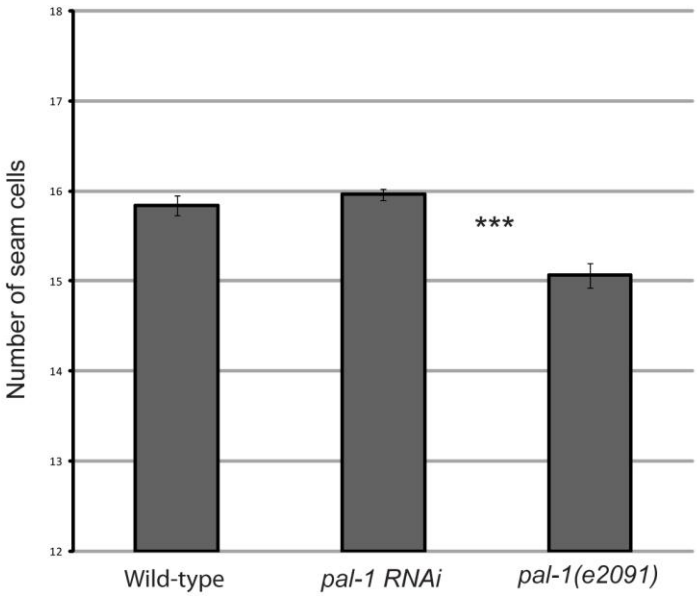


Figure 7.1 A. *pal-1* RNAi causes embryonic synthetic lethality with *rnt-1* mutants. Percentage embryonic lethality is shown for wild-type worms and rnt-1 mutants on control RNAi plates, feeding on HT115 bacteria containing the empty *L4440* RNAi feeding vector, and for wild type worms and *rnt-1* mutants fed on HT115 bacteria expressing dsRNA from the *pal-1* RNAi feeding clone. Embryonic lethality values are 0, 4.4, 78.8 and 98.3 respectively.
B. Seam cell numbers are reduced in *pal-1(e2091)* animals. Here, the *scm::gfp* reporter is used to mark seam cells in an adult hermaphrodite. Scale bar = 100µm
C. Graph showing the significant reduction in seam cell number in *pal-1(e2091)* animals. *pal-1* RNAi has no significant effect on seam cell number. Seam cell averages are 15.8±0.1 (wild-type), 15.96±0.4 (*pal-1* RNAi), 15.1±0.1(*pal-1(e2091)*).

The *pal-1(e2091)* mutation affects embryonic seam development

To determine whether the reduced number of seam cells in *pal-1(e2091)* mutants results from embryonic or larval defects, L1 larvae were examined just after hatching and before the first larval divisions had taken place. Counting seam cells in larvae, in which they are aligned in a single plane on each side of the worm, is considerably easier and quicker than in embryos, where the seam line follows the folds of the developing worm across multiple planes. Wild-type worms possess 10 seam cells at hatching, but in *pal-1(e2091)* mutants the average number of seam cells was slightly but significantly lower (9.7±0.06, Figure 7.2A). Having established this early-stage seam phenotype, embryos were then examined. As shown in Figure 7.2B, embryos too have fewer seam cells than wild-type embryos, demonstrating a role for *pal-1* in embryonic seam determination.

Larval seam abnormalities are also observed in *pal-1 (e2091)* mutants

It is of course possible that the *e2091* allele could also lead to reduced seam cell number during larval development. Lineage analysis was performed on *pal-1(e2091); him-5 [scm::gfp]* animals to determine whether or not this was the case. Of the animals observed in detail (n=31) no outright division failures were observed, but as in *rnt-1(tm388)*, *bro-1(tm1183)* and *elt-1RNAi* (but not in wild-type animals, see Chapters 3 and 5) loss of *scm::gfp* was observed in some seam cells, which were clearly visibly under DIC optics (n seam cells = 6, in 5 worms) and which went on to divide normally (Figure 7.3A). In addition, the relative times of divisions were perturbed, with V6 often dividing slightly before some or all of the other V lineage cells during the L2 symmetric and asymmetric divisions (n=5). Overall, loss of the *scm::gfp* marker from seam cells could contribute to the reduction of ‘green’ seam cells in *pal-1* adult worms, but division failures do not seem to play a role.

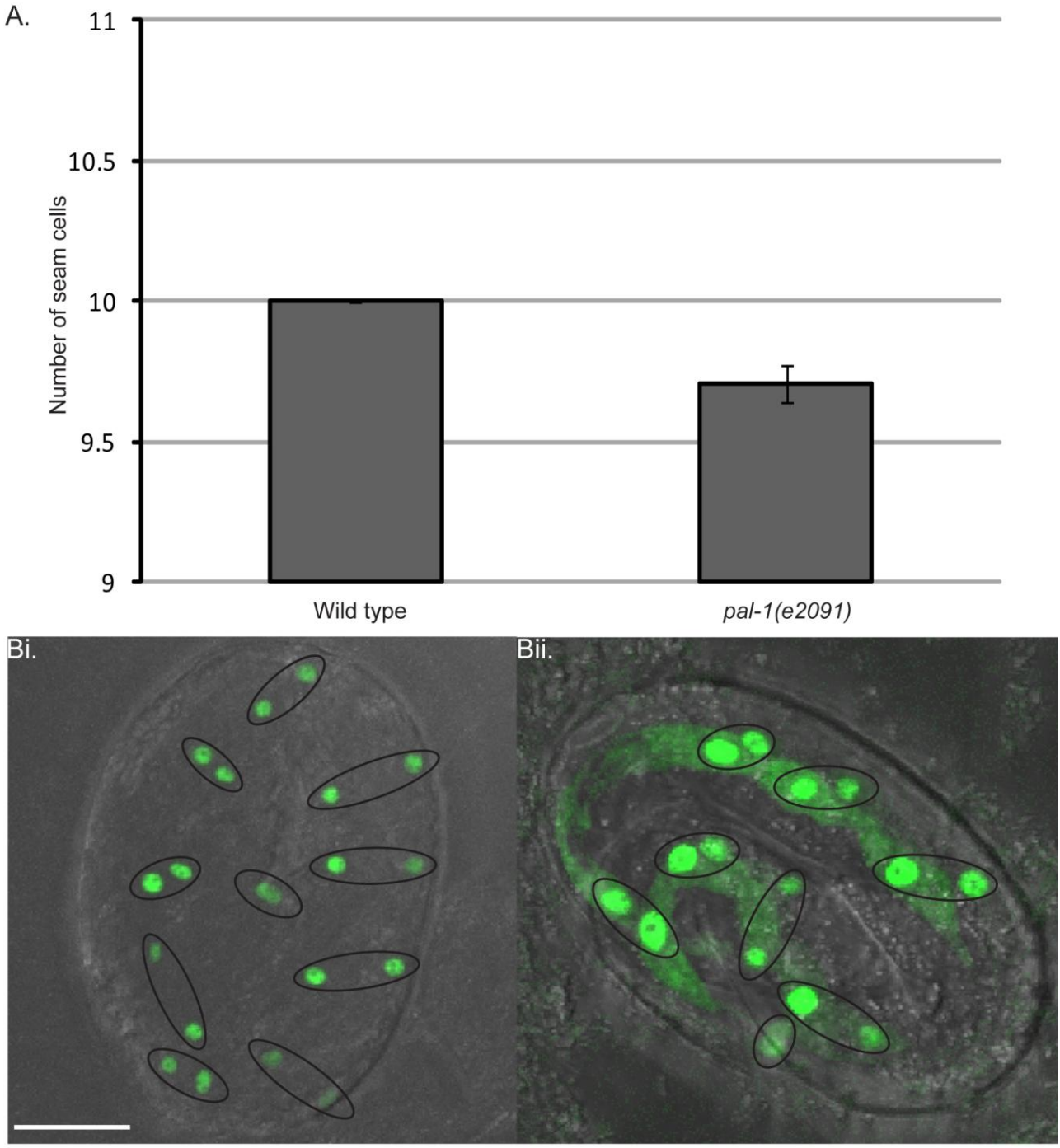


Figure 7.2 A. Graph showing seam cell numbers in wild type and *pal-1(e2091)* L1 larvae, counted before the L1 asymmetric division. B. Wild type (Bi) and *pal-1(e2091)* (Bii) embyros, with seam cells marked by *scm::gfp*. Left-right seam cell pairs are ringed.
Scale bar = 10μm

In addition to these 'fate-themed' aspects of the *pal-1* mutant phenotype, an additional striking abnormality was evident. Frequently, the seam cells were not distributed evenly along the length of the worm (as they are in wild-type animals, or even in, for example, *bro-1* and *rnt-1* mutants, which have several seam cells missing), but were distributed unevenly. Specifically, from the L1 stage onwards, the V2, V3 and V4 seam lineages were often found to be adjacent or even overlapping (Figure 7.3B), the V3 and sometimes also the V2 lineage having become shifted posteriorwards. As described above, the seam cells were observed to divide normally, but owing to the initial positioning of the cells this results in an uneven distribution of seam cells along the worm during larval development.

pal-1(e2091) mutants possess seam-derived 'tumours'

In addition to these phenotypes, *pal-1(e2091)* animals frequently (32% of animals) exhibited a 'bulge' in the middle of the body, often visible at the L1 stage and throughout subsequent development (Figure 7.4Ai). Examination of these bulges in adult *pal-1(e2091)* worms carrying the *scm::gfp* marker revealed no extra seam cells in the affected part of the body. However, examination of expression of the same marker during larval development revealed that additional seam cells often are present in the deformed region of the worm (Figure 7.4Aii). This paradox can be explained if the seam cells present in the bulge undergo a fate transition to another cell type before adulthood. The most obvious scenario was that such a fate change would involve seam cells losing their stem-like characteristics (together with the *scm::gfp* marker) and adopting a hypodermal fate. To determine whether this was indeed happening, the hyp7 marker, *dpy-7::yfp*, was crossed into the *pal-1* mutant background, together with the apical junction marker *ajm-1::gfp*, allowing cell outlines to be visualized. As shown in Figure 7.4B, the bulge region is packed with cells expressing *dpy-7::yfp*, demonstrating that hypodermal nuclei account for the large mass of cells in this part of the worm.

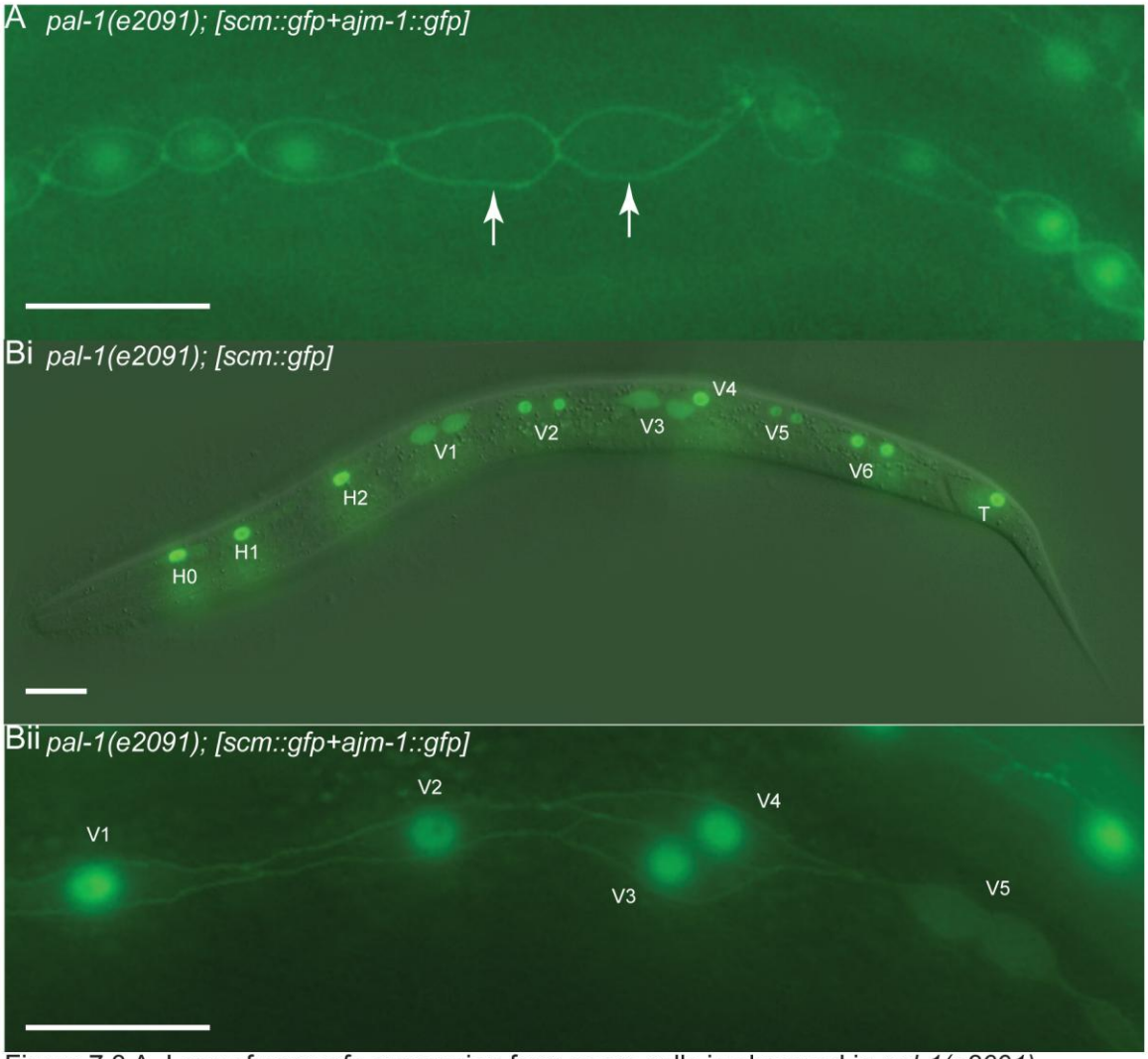


Figure 7.3 A. Loss of *scm::gfp* expression from seam cells is observed in *pal-1(e2091)* mutants. Here, the outlines of the seam cells are marked with AJM-1::GFP, with the two cells marked by arrows clearly lacking *scm::gfp*. B. In *pal-1(e2091)* mutant larvae, seam cells in the middle of the body region of the worm - specifically V3 and V4 - are often found to be abnormally close to one another (Bi, where V1 and V3 are in the process of dividing, and V4 has yet to divide), or even to overlap completely (Bii). Scale bars = 50μm.

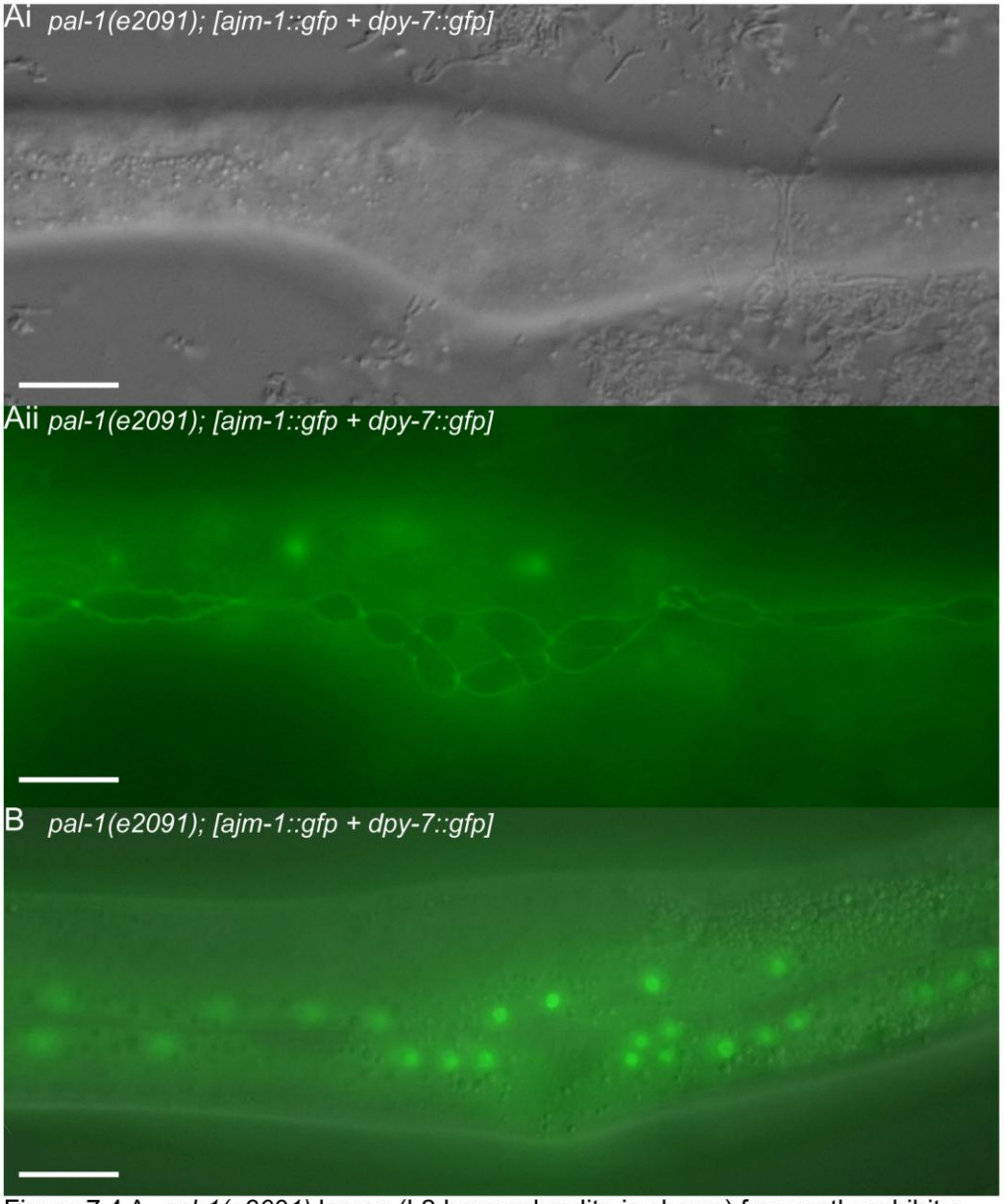


Figure 7.4 A. *pal-1(e2091)* larvae (L2 hermaphrodite is shown) frequently exhibit a distinct bulge in the mid-body region, shown here with Nomarski optics (Ai). During larval development, the bulge comprises in large part numerous seam cells, the outlines of which are marked with AJM-1::GFP (Aii). The same worm is shown in Ai and Aii. The DPY-7::YFP reporter is not in focus, as the nuclei lie in a different plane to the *ajm-1* seam cell boundaries. B. The cells produced by seam divisions in the bulge region of the worm (in this case, an L3 hermaphrodite) adopt a hypodermal fate and, after losing expression of *scm::gfp*, express the hypodermal marker *dpy-7::yfp*. Scale bars = 25µm.

pal-1 is expressed in the seam

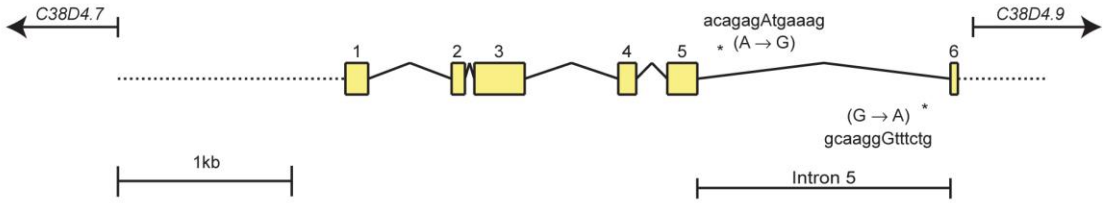
Previous studies addressing the role of *pal-1* have revealed expression from early embryonic stages, in the EMS and P_2 cells and their descendants, and subsequently in the C lineage (Hunter and Kenyon 1996), which contributes to muscle, and, during the larval development of males, in the V5 and V6 lineages, which give rise to cells which form the rays of the male tail. The seam phenotypes observed here, described above, raised the question of whether *pal-1* was also expressed throughout the seam lineage. A *pal-1* reporter construct was made using the entire open reading frame and intergenic region (Figure 7.5A) and found to express not only in these tissues but also in the seam lineage and larval body wall muscle (Figure 7.5B). Interestingly, although expression was evident in the V and T lineages of the seam, no expression was observed in the H0, H1 and H2 lineages. In addition, expression was observed in two cells near the tail, at least one of which appears to be neuronal (Figure 7.5C).

Seam expression of *pal-1* is driven by an intronic enhancer element

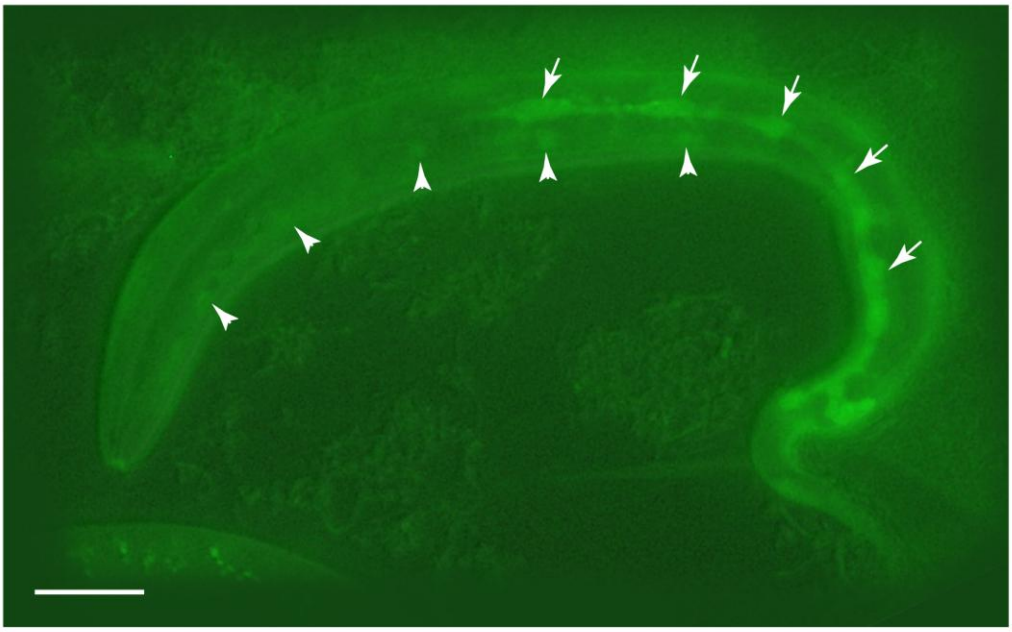
Knowledge of what lies upstream of a gene of interest is essential for understanding how the gene functions in the context of the network in which it performs its role. In an attempt to glean information on which genes might regulate *pal-1* and cause it to be expressed in the seam, the presumed promoter of *pal-1* (Figure 7.5A), comprising the 1.3kb intergenic region between *pal-1* and the neighbouring upstream gene, *C38D4.7*, was cloned into the GFP transcriptional reporter vector *pPD95.75*. No seam expression was apparent.

Given that the 5' intergenic region had failed to drive seam expression, the gene was examined for other candidate enhancer regions. To do this, a comparative approach was taken. Alignment of *C. elegans pal-1* with the homologues from *C. briggsae, C. brenneri, C. remanei* and *C. japonica* revealed particularly high levels of sequence conservation in the

A



B



C

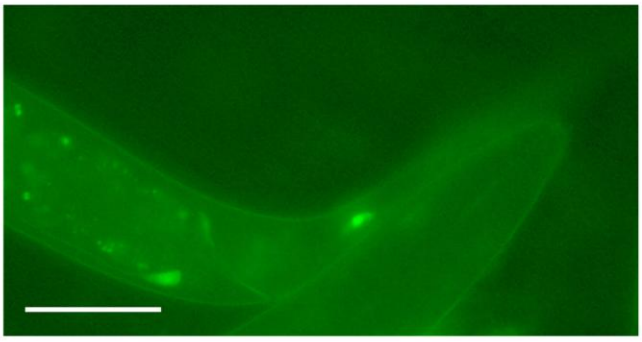


Figure 7.5 A. Schematic overview of *pal-1*. The two base changes in *pal-1(e2091)* mutants are annotated. Both lie within the fifth intron of the gene.
B. Expression pattern of the full length *pal-1::gfp* translational reporter. In addition to expressing in cells where *pal-1* activity has been previously reported (see main text), *pal-1* expression was observed in V-T lineage seam cells seam and in body wall muscle. In this picture, the seam is in focus (arrows) but the GFP is nevertheless visible in body wall muscle cells (arrowheads).
C. Expression of *pal-1::gfp* was also evident in two cells in the posterior of the worm, one of which is clearly neuronal, based on the long axon extending from the cell body. Scale bars = 25µm

last intron of *pal-1* (see Appendix). This intron was of particular interest because, as discussed below, the *pal-1(e2091)* allele contains two mutations within this region. The intron was therefore cloned into the *pPD107.94* and tested for its ability to drive expression in the seam. As shown in Figure 7.6A, seam expression was observed (both in larvae, as shown, and in adults). In addition, the early embryonic expression observed with the full length reporter, as well as strong expression in the two tail cells in larvae, was evident, suggesting that this intron directs most, if not all, of the complex expression pattern of *pal-1*.

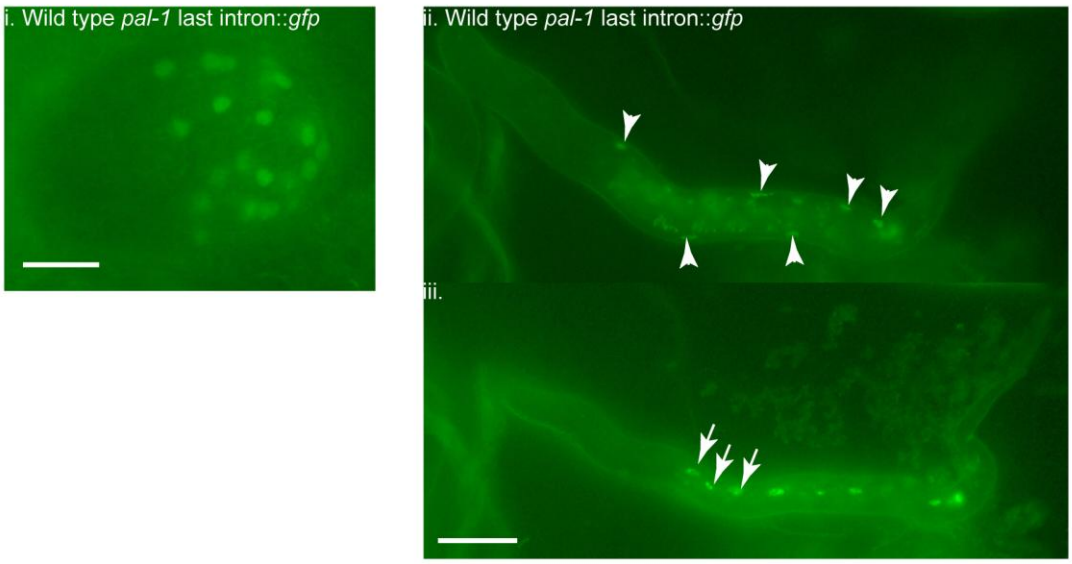
The *pal-1(e2091)* mutation abolishes seam expression from the intronic enhancer

The *e2091* mutation thought to underlie *pal-1* phenotypes – a T→C transition – is believed to affect a cis-acting enhancer element within the last intron (Zhang and Emmons 2000). To examine the effects of this mutation on the ability of the last intron to drive seam expression, the intron was amplified from *e2091* mutant worms and cloned into the minimal promoter vector *pPD107.94*, just as the wild-type version was cloned. In this construct, whilst expression was seen in the two tail cells and body wall muscle cells, no seam expression was observed in L1 larvae (Figure 7.6B). Thus, although the construct was present in the worms and expressing in some cells, the *e2091* mutation appears to interfere with the ability of the intron to drive early larval seam expression. Interestingly, although L1 seam expression was absent in the *e2091* construct, adult seam expression was evident. Analysis of putative binding sites within this region was performed but yielded no matches.

bro-1/rnt-1 work redundantly with *pal-1* in embryonic seam development

Given the apparent interaction between *rnt-1* mutants and *pal-1* RNAi, the mutant strains were crossed to give a *rnt-1; bro-1; pal-1* triple mutant, with the seam cells marked with the

A.



B.

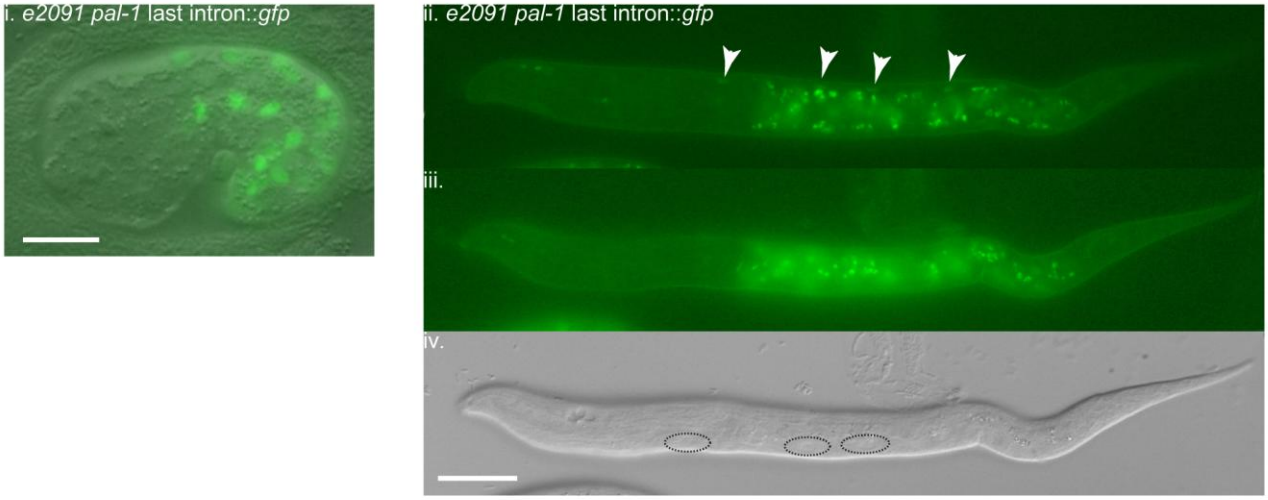


Figure 7.6 The fifth intron of *pal-1* is able to drive embryonic muscle expression, as well as expression in larval muscle and seam. A. Expression of GFP driven by a wild type copy of the intron is evident in muscle lineages at the 2-fold embryonic stage (Ai). In L1 animals, expression is evident in body wall muscle (Aii, arrowheads) and throughout the V seam lineages (Aiii, arrows). B. When the same intronic element is amplified from *pal-1(e2091)* mutants it is able to drive expression in embryonic muscle lineages (Bi) and L1 body wall muscle (Bii, arrowheads), but not L1 seam cells (Biii and iv). Several seam cells show no fluorescence (Biii) despite being clearly visible under DIC optics (Biv, seam cells circled). Scale bars = 10µm.

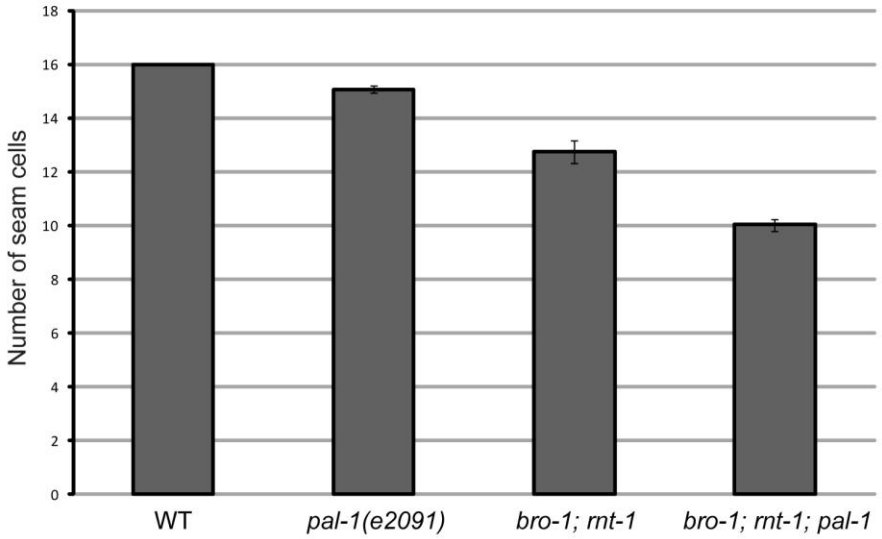
integrated *scm::gfp* array. Although the combination of the *rnt-1* mutation and *pal-1(RNAi)* results in almost complete embryonic lethality, the *pal-1(e2091)* mutant is hypomorphic and the triple mutant strain viable. Adult seam cell numbers were examined in these animals, and found to be significantly lower than in *rnt-1; bro-1* double mutants (10.0±0.2 compared to 12.8±0.4, Figure 7.7A) or in *pal-1* single mutants (15.1±0.1, Figure 7.7A). This phenotype appears to represent a synergistic, not merely an additive effect of the *pal-1(e2091)* and *rnt-1/bro-1* mutant phenotypes.

Again, to determine how the reduction in seam cell number arises, *bro-1; rnt-1; pal-1* early L1 larvae were examined and found to possess significantly fewer seam cells (8.7±0.19) than *pal-1* single mutants (9.7±0.06, see above) and *bro-1; rnt-1* L1s, which, like wild-type worms, invariably have 10 seam cells at hatching (Nimmo, Antebi et al. 2005). *pal-1* and *rnt-1/bro-1*, therefore, appear to have partially redundant roles in the embryonic seam lineage; *pal-1* clearly plays an essential role and is able to work independently of *bro-1/rnt-1*, but the *pal-1* mutant phenotype is enhanced by loss of functional *bro-1/rnt-1*, which would otherwise have no effect on embryonic seam development.

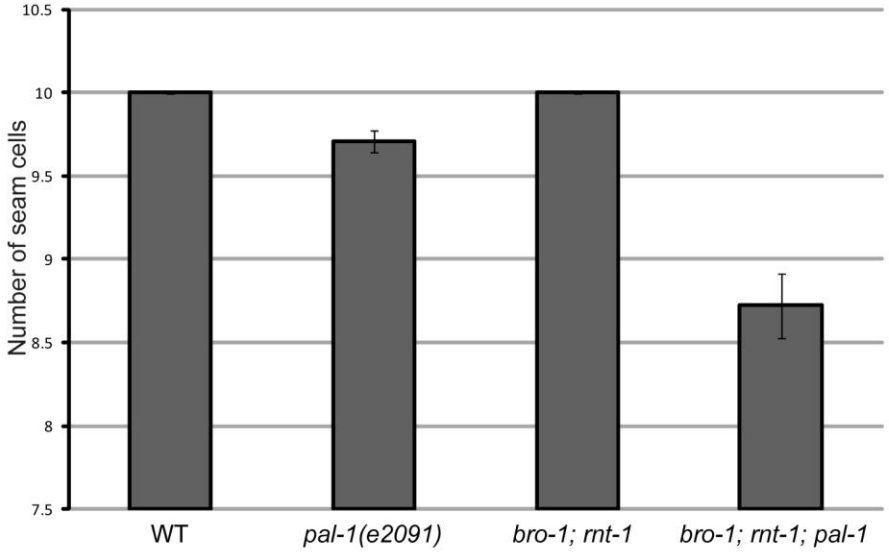
The *bro-1; rnt-1* and *pal-1* mutations cause synthetic increases in larval lethality

The *bro-1; rnt-1; pal-1* triple mutant also provided an additional line of evidence in support of redundancy between *bro-1/rnt-1* and *pal-1*. Whilst there was no significant difference in embryonic lethality between *pal-1* (20%) and *bro-1; rnt-1; pal-1* (26%) animals ($p > 0.4$), larval lethality increased synergistically in the triple mutant strain (from 11.2% in *pal-1(e2091)* mutants and 12.5% *in bro-1;rnt-1* mutants to 68% in *bro-1; rnt-1; pal-1* animals, Figure 7.8A). This may be due to the reduction in seam cell number observed in embryos and L1 stage worms. Furthermore, the finding that early larval seam expression is disrupted as a result of the *e2091* mutation suggests that seam defects may lie behind the high levels

A.



B.



C.

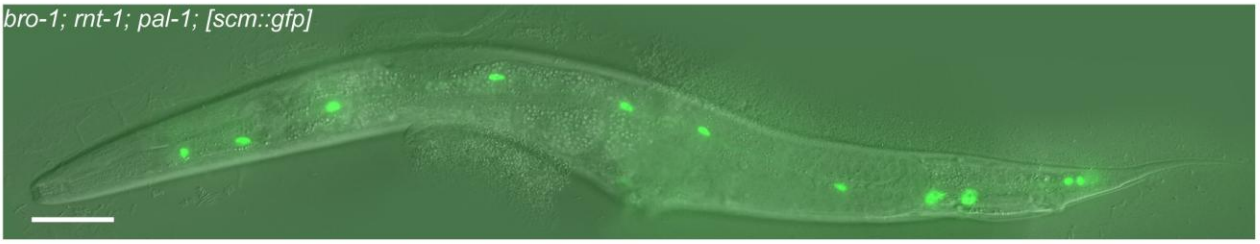


Figure 7.7 A. Seam cell numbers in adult hermaphrodites. *pal-1(e2091)* mutants and *bro-1;rnt-1* double mutants both have significantly fewer seam cells than wild type worms (15 and 12.8 compared to 16, respectively). However, in *bro-1; rnt-1; pal-1* mutants, the number of seam cells is lower still (10.2), with the decrease being more than the sum of that observed in *pal-1* and *bro-1;rnt-1* mutant strains, indicative of a synergistic effect. B. Seam cell counts in newly hatched larvae reveal partial redundancy between *pal-1* and *bro-1/rnt-1*. Although *bro-1; rnt-1* double mutants have an average of 10 seam cells per L1, the triple mutant strain has significantly fewer. Again, comparison with *pal-1* single mutants suggests and additive effect. Error bars show SEM.
C. A *bro-1; rnt-1; pal-1* adult hermaphrodite, possessing only 11 seam cells, as opposed to the wild-type number of 16. Scale bar = 100μm.

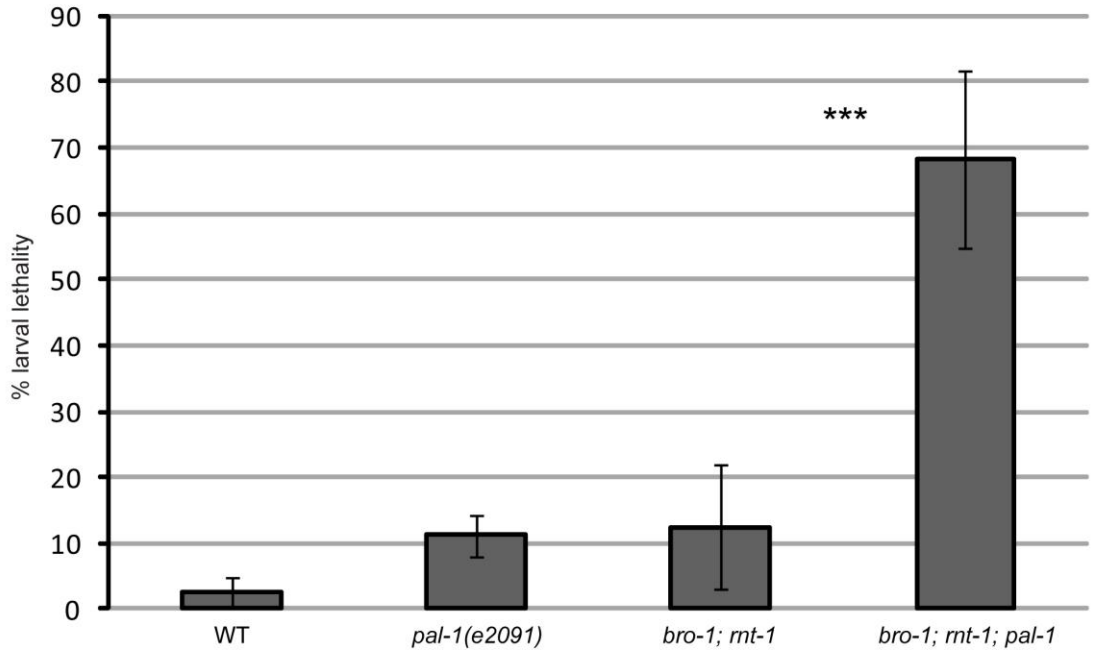
of larval lethality seen in *pal-1* mutants, and may account for the synthetic increase in larval lethality in the triple mutant strain, compared to *pal-1* and *rnt-1; bro-1* mutants. However, it is worth noting that, whilst the number of body wall muscle nuclei in both strains is normal, examination of muscle fibres reveals significant disorganisation (Figure 7.8B). In addition, larvae are severely deformed, consistent with previous reports of muscle defect phenotypes (Hresko, Williams et al. 1994; Williams, Waterston et al. 1994; Baugh, Wen et al. 2005).

pal-1 is known to have a role in muscle development (Hunter and Kenyon 1996; Edgar, Carr et al. 2001; Lei, Liu et al. 2009) and here *pal-1(e2091)* mutant larvae are found to exhibit muscle defects. Significantly, the fact that the *e2091* mutant exhibits such abnormalities demonstrates that the allele does not just affect seam development. In *bro-1; rnt-1; pal-1* triple mutants, larvae are also found to have the muscle fibre phenotype shown in Figure 7.8B. One explanation for this phenotype is therefore that the loss of both functional *bro-1/rnt-1* complex and of *pal-1* results in more worms having muscle defects and therefore in more worms dying as larvae as a result.

The developmental basis of synthetic embryonic lethality *in rnt-1; pal-1(RNAi)* worms

The reason for the high levels of embryonic lethality in rnt-1 mutants subjected to *pal-1* RNAi is not clear. Given that the role of *rnt-1* is best characterised in the seam, it was important to first determine whether *pal-1* RNAi, on its own, had an effect on seam cell number. Seam cells were counted in *pal-1* RNAi worms, but found to be not significantly different from wild-type ($p>0.3$). In the case of *rnt-1* worms subjected to *pal-1* RNAi, seam cells were counted in adult escapers. No significant change in seam cell number was observed (13.2 compared to 13.74, $p = 0.24$). As noted above, *pal-1* is known to be expressed in muscle. *rnt-1* is also expressed in muscle but no muscle defects have been detected in *rnt-1* mutants and, as yet, no function for *rnt-1* in these cells has been determined (Nimmo, Antebi et al. 2005). One possibility, therefore, is that these genes also have redundant functions in this tissue.

A.



B.

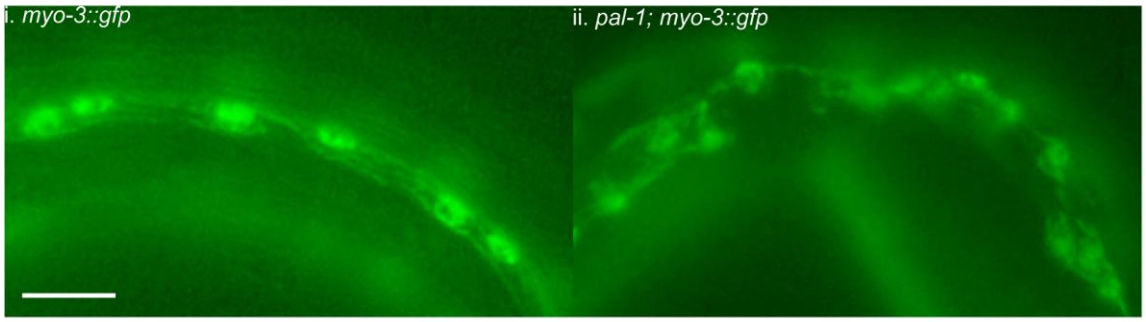


Figure 7.8 A. Graph showing synthetic larval lethality in *bro-1; rnt-1; pal-1* mutants; in the triple mutant, 68% of those larvae which hatch die before adulthood. This compares to 2.5%, 11.2% and 12.5% for wild type, *pal-1(e2091)* and *bro-1; rnt-1* mutants respectively. B. Muscle fibre organisation in wild type (Bi) and *pal-1(e2091)* mutant (Bii) worms. Muscle cells are marked by the *myo-3::gfp* reporter. Significant disorganisation is evident in the muscle fibres of *pal-1* mutant larvae. Scale bar = 25μm

Discussion

The Caudal homologue *pal-1* interacts with *rnt-1*

The findings presented here demonstrate a novel interaction for *rnt-1*, with the Caudal homologue *pal-1*. The nature of this interaction appears complex and is manifest at both embryonic and larval stages. Furthermore, this interaction appears to represent partially overlapping functions of *rnt-1* and *pal-1*, but not complete functional redundancy; both genes have independent roles and distinct phenotypes.

The initial discovery that *pal-1* RNAi synergises with *rnt-1* mutants, leading to almost total embryonic lethality, suggests that there is at least some redundancy between these genes, and that this redundancy involves relatively early developmental processes. As well as revealing the interaction between *pal-1* and the *bro-1/rnt-1* complex, this finding also demonstrates a role for *rnt-1* and *bro-1* in embryos. Although *rnt-1* and *bro-1* expression has been observed from the two-fold embryonic stage (Nimmo, Antebi et al. 2005), no embryonic function has previously been assigned to these genes. This result complements those presented in Chapter 4, where redundant embryonic function was revealed between the *bro-1/rnt-1* complex and *unc-62*, and highlights the utility of screens for synthetic lethality. It appears that *bro-1/rnt-1* function in embryonic development, certainly in the seam and possibly in muscle as well. There has been speculation about a direct interaction between *unc-62* and *pal-1* (Van Auken, Weaver et al. 2002), but little evidence of such a relationship has followed. Either way, the functions of *bro-1* and *rnt-1* are clearly multi-faceted and more complex than previously thought.

Given that *rnt-1* has only so far been found to function in the stem-like seam cells of the worm, the most obvious explanation for this synthetic lethality is that *rnt-1* and *pal-1* function partially redundantly in seam development. It was to investigate this possibility that seam cell numbers were counted in *rnt-1;pal-1(RNAi)* worms which survive embryonic development. The fact that no change in seam cell number was observed is suggestive that a seam defect is not responsible for the synthetic lethality. However, it could be that those embryos which die before the larval stage are more severely affected in terms of seam cell number; the animals in which seam cell numbers are counted likely represent those least affected by the *pal-1* RNAi.

Another line of evidence against a seam cause for the synthetic lethality seen in *rnt-1; pal-1(RNAi)* animals is that, whilst the *pal-1(e2091)* allele affects seam cell number, *pal-1* RNAi does not. Although *pal-1* mutants exhibit severe seam defects and a significant decrease in seam cell number, *pal-1* RNAi animals do not. Thus, it may be that *pal-1* RNAi causes relatively weak knockdown of *pal-1*, but across the wide range of tissues in which this gene acts. Given the severity of lethality in *rnt-1; pal-1(RNAi)* worms, if this was caused by seam defects it might be expected that *pal-1 RNAi* alone would have a seam phenotype at least as pronounced as that seen in *pal-1(e2091)* mutants. This has been shown not to be the case.

An alternative explanation for the synthetic lethality between *rnt-1* mutants and *pal-1* RNAi is that another tissue is affected. *rnt-1*, in addition to being expressed in the seam, is expressed in muscle (Nimmo, Antebi et al. 2005). As discussed above, *pal-1* is also expressed here. Thus, there is a strong possibility that the embryonic defects observed in *rnt-1; pal-1(RNAi)* worms are due to muscle-specific defects. Furthermore, muscle defects could account for the embryonic lethality and developmental abnormalities seen in *rnt-1; pal-1(RNAi)* embryos.

Establishing the precise cause of the synthetic embryonic lethality will likely require more detailed analysis of the affected embryos. Given the availability of *rnt-1* mutant strains rescued with transgenic *rnt-1* expressed specifically in either seam or muscle (see Chapter 6), it should be possible to determine in which tissue functional *rnt-1* is required to alleviate the synthetic lethal phenotype. Nevertheless, the synthetic lethality observed demonstrates two points. Firstly, *rnt-1* functions partially redundantly with *pal-1*. Secondly, *rnt-1* must function earlier in development than previously thought; although *rnt-1* expression has been observed in embryos in both seam and muscle cells, no function has been assigned to this expression.

Just as the embryonic lethality observed in *rnt-1; pal-1(RNAi)* worms demonstrates a novel role for *rnt-1* in embryos, the finding that *pal-1(e2091)* embryos have reduced seam cell numbers and that *pal-1* larvae have numerous seam defects, demonstrates a novel role for *pal-1* in the seam.

With respect to the reduction in embryonic seam cell number, *pal-1* appears to again work both with and independently of *rnt-1/bro-1*. *pal-1(e2091)* mutant embryos possess fewer seam cells, and this number is further reduced in a *rnt-1;bro-1* mutant background (whilst *rnt-1; bro-1* double mutant embryos invariably possess 10 seam cells at hatching). It remains to be determined whether the reduction in seam cell number at this early stage results from division failures, division symmetry defects, or from inappropriate cell death. Based on current knowledge of the functions of *rnt-1* and *bro-1*, the two former options seem most likely.

A novel role for *pal-1* in larval seam development

The *pal-1(e2091)* 'bulge' phenotype reveals yet another role of *pal-1* in the larval seam cells. Examination of the bulge in larvae and adults revealed that, in large part, affected regions are comprised of hypodermal nuclei. However, although this provides an answer to the cellular 'make-up' of the abnormal bulges observed in *pal-1* mutants, it leaves open the question of how and why this feature develops. There are several potential explanations:

First, it could be that hypodermal cells, after leaving the seam following asymmetric cell divisions, continue to divide whilst expressing the *dpy-7::yfp* marker and after losing their AJM-1::GFP boundary. This explanation seems unlikely. As discussed in detail in Chapter 3, loss of the AJM-1::GFP boundary which surrounds seam cells, and concomitant acquisition of *dpy-7::yfp* expression, appear to be definitive markers for the loss of division potential. Furthermore, despite extensive examination of larvae (n=>500), no cell divisions were observed in those cells expressing the *dpy-7::gfp* marker.

Second, it could be that the seam cells undergo supernumerary divisions. For example, extra asymmetric divisions could take place, or perhaps divisions could be symmetrised to the seam fate, followed by some of these resultant cells undergoing divisions symmetrised to the hypodermal fate. The lineage analysis described above revealed no evidence of this.

Instead, it appears that the shifted positions of the mid-body seam cells, in particular V3 (and occasionally V2) results in a concentration of otherwise normal seam divisions in the middle of the worm. As a result, the hypodermal daughters which result from the reiterated asymmetric divisions of the seam cells accumulate within a relatively small region of the

worm. Thus, it seems likely that it is the aberrant position of seam cells, rather than any defects specifically linked to, for example, cell cycle or division orientation, which give rise to the distinctive morphology. *pal-1*, it seems, is therefore important for the positioning of the seam cells during early development, with defects occurring at early larval stages having severe consequences right up to adulthood. This underlines the importance of cell positioning for correct morphogenesis. The precise mechanism underlying the seam position defects in *pal-1(e2091)* mutants remains to be determined, but identification of the exact stage at which abnormalities first arise will be important. Embryonic lineage analysis may help in this respect, potentially revealing whether the mispositioning occurs as a result of aberrant migration or perhaps because of defects during elongation, where the seam cells alter their shape through cytoskeletal rearrangement.

The high levels of larval lethality observed in *bro-1; rnt-1; pal-1* mutants demonstrate that the effects of interactions between *rnt-1* and *pal-1* are not restricted to embryonic development, but it is hard to determine whether seam or muscle defects, or a combination of both, are to blame. Certainly, seam cell numbers are synergistically reduced in triple mutant animals. However, larvae also exhibit severe disruption of muscle fibres, as well as dumpy and uncoordinated phenotypes which are associated with muscle defects (Baugh, Wen et al. 2005). Here again, the *rnt-1* tissue-specific rescuing constructs described in Chapter 6 should shed light on this question.

The last intron of *pal-1* contains key regulatory elements

Dissection of the genomic regions required for *pal-1* expression explained the significance of the molecular lesions present in *pal-1(e2091)* mutants. The *pal-1(e2091)* mutant contains two base changes in the fifth intron of the gene (Zhang and Emmons 2000), as depicted in

Figure 7.5A. Of these two mutations, the 5' most was determined to be behind the *pal-1* phenotype which manifests itself in the V5 and V6 lineages. According to the model proposed, an unknown factor is required to bind to the region encompassing this mutation to allow transcription of *pal-1*. The finding that the intron within which this region lies can effectively act as a key enhancer of *pal-1* supports this model. Furthermore, that the mutation apparently abolishes seam expression of *pal-1* in early larvae provides a mechanism by which the molecular lesion can affect seam development, and affect *pal-1* interactions with *rnt-1*. In the *pal-1(e2091)* mutant, *pal-1* fails to be expressed in the seam and is therefore unavailable to interact with *rnt-1*.

The molecular nature of the interaction between *rnt-1* and *pal-1*, however, remains mysterious but the link between these genes and the mediator complex is intriguing. PAL-1 is known to be regulated by this complex (Zhang and Emmons 2000) and, intriguingly, RNT-1 has been found to show synthetic lethality with SOP-1 (Suppressor Of Pal), a mediator complex component. Indeed, the interactions which have been described for *sop-1* and *bro-1/rnt-1* mirror the *pal-1* interactions described here. In a wild-type background, *sop-1* RNAi causes a decrease in seam cell number, from 16 to 13.9 seam cells per side, a reduction similar to that seen in *bro-1* mutants (Xia, Zhang et al. 2007). However, *sop-1* RNAi in a *bro-1* background synergistically reduced seam cell number to 9.2. Furthermore, just as in the case of *bro-1; rnt-1; pal-1* triple mutants and *rnt-1* mutants subjected to *pal-1* RNAi, *sop-1* RNAi synergised with *bro-1* and *rnt-1* mutations to cause dramatically increased levels of embryonic and larval lethality. That *bro-1* and *rnt-1* mutations synergise with both *sop-1* and *pal-1* defects is suggestive of a function for *bro-1/rnt-1* associated with the mediator complex, known for its role in repressing gene expression. It will be interesting to investigate whether there is synergism with other components. Either way, the apparent link between the *bro-1/rnt*-1 pathway, *pal-1* and *sop-1* provides a basis for further studies which could identify new interacting partners for *bro-1* and *rnt*-1. One interesting avenue of exploration would be the use of other, hypomorphic *pal-1* alleles such as *ns114* and *ns115*; these alleles

are known to give rise to weak tail phenotypes, but no investigations as to whether they have an effect on seam development have so far been reported.

Taking together the enhancer activity of the *pal-1* fifth intron and the effect of the *e2091* mutation on seam expression, analysis of binding sites within this intron – and particularly within the region containing the mutation – may provide important information on the identity of the factors regulating *pal-1* expression in the seam. As noted in the Results section, no putative binding sites have so far been identified. However, a yeast one-hybrid screen is planned with the intron. By using both wild-type and *e2091* copies separately, it should be possible to perform a powerful search and test for *pal-1* regulators.

Of course, the minimal promoter experiments performed here have their limitations. Transcription factors are dependent on context for their function, both in terms of the sequence to which they bind and in terms of the presence of other transcription factors with which they interact. It would be worth introducing the *e2091* mutations into the full length *pal-1* reporter to confirm their importance and effect in a less artificial situation. Furthermore, the assumption that it is the 5' most of the two mutations which affects seam expression (and therefore the ability of *pal-1* to interact with *rnt-1*) could be tested by introducing each of these mutations separately and examining the effect of each on *pal-1* expression.

The effect of the *pal-1* mutation on seam expression makes it somewhat surprising that muscle phenotypes are also apparent in this strain (Figure 7.8). Clearly, the relatively coarse examination of the expression patterns from the minimal promoter experiments described here does not give a complete picture of *pal-1(e2091)* defects. It could be, for example, that subtle defects in the timing of expression result from the mutation. Certainly, this is a critical and often tightly controlled aspect of regulation. In the case of *mab-5*, for example, expression of which happens to be activated by *pal-1* (Costa, Weir et al. 1988; Hunter,

Harris et al. 1999), normal development requires mab-5 to switch on, then switch off, and then switch on again within the same lineage (Salser and Kenyon 1996).

Conclusion

pal-1 represents a new component of the genetic network in which *rnt-1* operates. Unlike *bro-1*, which appears to work so closely with *rnt-1* that the phenotypes of *bro-1* and *rnt-1* mutants are almost identical, *pal-1* appears to have a much wider range of functions than *rnt-1*. Nevertheless, their functions clearly overlap in the seam and, given the severity and nature of defects observed in *rnt-1; pal-1* mutants, and in *rnt-1; pal-1(RNAi)* animals, likely also in muscle.

Furthermore, as well as revealing a novel interacting partner for *rnt-1*, these findings also demonstrate that *rnt-1* has important functions in the developing embryo, at least in the seam and possibly in muscle too.

Finally, it is once again apparent that intronic sequence can have significant roles in gene function. Indeed, in the case of *pal-1*, analysis of mutations within an intron have yielded valuable information on the workings and interactions of this gene and further analysis of this region looks likely to shed more light on the identity of upstream components of this network.

CHAPTER 8

General Discussion

The *C. elegans* RUNX homologue, *rnt-1*, together with its cofactor *bro-1*, is firmly established as being of central importance for the proliferation of the stem cell-like seam cells. However, understanding how these genes regulate seam cell divisions requires knowledge of the genetic context in which they work; this context comprises the genes which regulate *bro-1* and *rnt-1*, and the downstream targets of the *bro-1/rnt-1* pathway. Furthermore, a full understanding of the genetic network which regulates seam development involves knowledge not just of the components but of the nature and mechanisms of the interactions between these components.

The search for components of the *bro-1/rnt-1* pathway described in this thesis has yielded a number of genes. *elt-1*, discussed in Chapter 3, acts as a direct transcriptional regulator of *bro-1*. *unc-62* appears to operate upstream of *bro-1* and *rnt-1*, regulating the symmetry of larval seam cell divisions, but also acting redundantly with these genes in embryonic seam determination and likely in body wall muscle during embryonic and larval development. These findings are covered in Chapter 4. Operating downstream of the *bro-1/rnt-1* complex is *arf-3*, a gene identified because of its inclusion within the widely used *scm::gfp* seam cell marker. This interaction, and the putative function of *arf-3*, are discussed in Chapter 5. Indeed, *arf-3* appears to act in the seam downstream of several of the genes discussed in this thesis; *bro-1*, *rnt-1*, *elt-1* and *pal-1* all appear to regulate *arf-3* expression and uncouple the activity of this gene from the seam cell fate, such that *arf-3* expression is frequently found to be lost from seam cells in worms in which the functionality of these genes is reduced. Chapter 6 describes the modular structure of the *rnt-1* promoter and thus provides a foundation for revealing more about the regulation of *rnt-1* expression in both seam and

body wall muscle lineages. Finally, as detailed in Chapter 7, *pal-1* appears to work together with *bro-1* and with *rnt-1* during embryonic and likely during larval development. In the process of identifying these *bro-1/rnt-1* interactors, a number of themes have become apparent.

The utility (and limitations) of comparative genomics

One of the themes to emerge from this work is the power and utility of comparative genomics, both for identifying relatively large (>100bp) regions of sequence which act as promoters or enhancer elements for genes, and for revealing specific transcription factor binding sites. The case of *elt-1* provides the clearest illustration of both instances; sequence alignment and comparison between several *Caenorhabditis* species revealed the existence of the *bro-1* CNE. Moreover, within this region, a comparative approach made it possible to identify a transcription factor binding site which was subsequently shown to be bound by the GATA family transcription factor ELT-1.

At the same time though, this work also demonstrates the shortfalls of the comparative approach and of the use of transcription factor binding site predictions. In the above case of *elt-1*, it is worth noting that two putative GATA transcription factor sites were identified, but only one was shown to be necessary for ELT-1 binding to the *bro-1* CNE. Furthermore, the work performed on *unc-62* was initiated on the premise that this gene likely bound to two adjacent, well-conserved sites within the same intronic region. Although *unc-62* clearly operates upstream of *bro-1* (and thus of *rnt-1*), mutation of these putative binding sites within the intron suggests that they are not necessary for *bro-1* expression in the seam. In this case, the comparative sequence analysis and binding site prediction was misleading, albeit in a fortuitous way, highlighting the limitations of this approach.

Firstly, the designation of sequence as ‘conserved’ or ‘non-conserved’ is by no means an exact science; 100% conservation between several species obviously falls within the ‘conserved’ category, but the significance of, for example, 70% conservation, is harder to determine.

Secondly, added to this issue is the fact that sequence conservation provides only an indication, not proof, of function. Conversely, as with the *rnt-1* promoter modules revealed in Chapter 6, lack of conservation does not demonstrate lack of function. As discussed in Chapter 3, binding sites and their respective transcription factors can co-evolve, meaning that whilst the same protein-DNA interactions may occur in different species, the DNA binding sites and recognition properties of the transcription factors may differ. It is also worth noting that it is conceivable that networks can change between species; when a new transcription factor binding site evolves in one species, it will by definition not be present even in the most closely related species. In addition, the small size of the motifs bound by transcription factors makes the likelihood of chance occurrence of any given binding site relatively high.

Thirdly, predictions of binding sites are limited by prediction software. If the database of transcription factors used is not comprehensive (and, given that binding site consensus sequences are not available for all transcription factors, this is inevitably the case), the list of candidate regulators will be limited from the outset.

Reflections on the structure and location of ‘promoters’ and ‘enhancers’

Stemming from the comparative approach are findings concerning the nature of promoter and enhancer elements. It is very striking that in the case of every gene whose transcriptional regulation has been investigated in detail here – *bro-1, unc-62, arf-3, rnt-1*

and *pal-1* – intronic sequence has played a major role in directing tissue-specific expression. In many cases, gene reporters include several kilobases of sequence upstream of the open reading frame, but the findings presented here suggest that enhancer elements within the protein coding region may be at least as significant for correct expression and therefore for correct function. Furthermore, when upstream regulators of any given gene are sought, this work highlights the importance of analysing not just upstream sequence but also non-coding sequence within the protein coding region.

Analysis of enhancer elements has also revealed modularity in their structure. In *bro-1* and *rnt-1*, separate regions of sequence drive seam and muscle expression. Furthermore, the mutations present in the *pal-1(e2091)* last intron, shown here to drive expression of the gene, specifically abolish early larval seam expression but appear not to disrupt expression in at least some of the other tissues in which *pal-1* is expressed, and even adult seam expression from this intronic element remains intact.

Redundancy between enhancer modules is also evident in the cases of *rnt-1*, where separate regions of the long intron are capable of driving seam expression, and *arf-3*, where both 5' sequence and the last intron of the gene are sufficient (and therefore not independently necessary) for strong expression in the seam.

Both the modularity and redundancy of promoter elements likely reflect the mechanisms of the evolution of gene function, with distinct regions of sequence being duplicated and co-opted into the function of new or existing genes and thereby modifying their expression patterns. Revealing more about the extent of this modularity, and the degree to which tissue-specific modules are shared between genes expressed in the same cells, should provide further insights into the plasticity of developmental networks and how they have changed over evolutionary time. Moreover, promoter modularity has important predictive potential. If indeed many or even most genes are regulated by common promoter modules, identification of these modules should allow the identification of novel candidate regulators in a range of

tissues of interest. Combined with other predictive approaches and experimental verification, this should be a powerful tool in better annotating and understanding the workings of genes in development.

It is also worth noting that the intronic enhancer elements identified here are much bigger than the average transcription factor binding site consensus sequence – indeed, they are at least an order of magnitude bigger. It is possible that, with further dissection, these elements could be pared down but it also seems likely that, for example, a single binding site is simply not enough to drive precise, tissue-specific expression patterns; not only would this be suggestive of unrealistically simple interactions between transcription factors and the genes they regulate, it would also mean that each occurrence of a given consensus sequence – of which there are hundreds of thousands throughout the genome – would represent an enhancer. Instead, the interactions between transcription factors and binding sites are likely much more subtle and complex. In the case of the *bro-1* CNE, for example, ELT-1 is probably not the only transcription factor involved in sequence recognition and regulation of the gene's expression pattern, as perhaps suggested by the fact that only ELT-1 – and not the other GATA family members – appears to bind to the GATA site within the intron; interactions with other regulatory proteins, which perhaps bind to adjacent sites in the CNE, likely prevent GATA family members other than ELT-1 from binding – the interaction is highly specific. The importance of flanking sequence in this way may therefore account for the relatively large region of conservation in the CNE and, more broadly, for the fact that enhancers are often distinguishable as conserved regions, often several hundred base pairs in length, rather than individual transcription factor binding sites.

Even in the cases here, where putative enhancer regions have been identified and verified, it is unusual to find significant stretches of sequence showing100% conservation across five Caenorhabditis species; *C. remanei* and *C. japonica* often show significant divergence of sites which are much better conserved across *C. elegans*, *C. briggsae* and *C. brenneri*. It appears that the rate of evolutionary change of these sites is relatively high. Whether the

developmental pathways underlying, for example, seam development, are also diverging, remains to be seen.

Dissecting the identity and underlying function of the *scm::gfp* marker

Since all of the genes investigated here play important roles in the seam lineage, the seam cell marker, *scm::gfp* has been relied upon heavily. Indeed, this marker is widely used in the *C. elegans* seam cell field because of its association with the seam fate. However, in the cases of *bro-1, rnt-1, elt-1* and *pal-1*, this association is found to be uncoupled, with the presence or absence of *scm::gfp* expression failing to predict the 'stem fate' of the cell in question; expression is often lost in cells which go on to divide, whilst some cells expressing *scm::gfp* fail to divide. Moreover, all of these genes appear to regulate *scm::gfp* expression; reduced functionality of these genes frequently results in a proportion of seam cells losing expression of this marker. Significantly, these genes are also important for establishment and maintenance of the seam fate. No explicit link between specific genes and regulation of this marker – as opposed to regulation of the seam fate – has previously been made. Taken together, these findings suggest that, whilst not clear-cut, there is nevertheless a link between the *scm::gfp* marker and the seam fate.

Following on from the observations, the genetic identity of the *scm::gfp* marker was investigated and found to depend upon *arf-3*. Unlike the other genes covered in this study, *arf-3* is not a transcription factor but a GTPase, likely involved in regulating vesicle transport (discussed in Chapter 5). Thus, recognition of the role of *arf-3* in determining seam fate represents a broadening of the scope of components known to be involved in this lineage and in the regulation of seam divisions. The apparent involvement of vesicle transport in determination of seam cell division potential and symmetry represents an exciting link

between current knowledge of symmetry determination, mediated by the WNT signalling pathway, and transcription factors such as *bro-1* and *rnt-1*.

Seam cells and the stem cell niche

The focus on *bro-1* and *rnt-1* is intrinsically linked with a focus on the stem cell-like seam cells of the worm. Correct development of seam cell divisions appears to involve regulation at several levels. Thus, whilst *unc-62* regulates division symmetry and *pal-1* regulates seam cell positions, *elt-1* works with *bro-1* and *rnt-1* to maintaining seam proliferative potential and, via *eff-1*, to prevent differentiation. This last aspect of 'seam maintenance' is particularly pertinent to the concept of the seam cells as stem cells. Although the *C. elegans* seam has the potential to both differentiate and self-renew, the absence of a defined niche raises the question of whether seam cells are really stem cells, or merely 'stem cell-like'. The finding that seam cell differentiation is blocked by maintenance of a physical boundary between the seam and the surrounding syncytial hypodermis (Chapter 3) suggests that the seam cells may in fact reside in a niche – a microenvironment which protects them from differentiation signals – and provides an incentive to identify and characterise the signals which are likely involved. Parallels between this system and the situation in the vulval development are worth noting. Here too, differentiation of cells is prevented by maintenance of cell boundaries between vulval precursor cells and the neighbouring hypodermis (hyp7). Thus, the employment of pro-differentiation signals emanating from hyp7 is common to at least two developmental pathways.

The findings here also reveal the importance of apical junction components in maintenance of the seam fate. Parallels with other systems, discussed in Chapter 3, where cell junctions are central to the niche concept, lend additional weight to the concept of the 'seam cell niche'.

Revealing the network in which *bro-1* and *rnt-1* operate

The work outlined above has filled in some of the gaps in our knowledge of the context in which *bro-1* and *rnt-1* function. Many remain, and even in the cases of the genes investigated in this study, the precise nature of their interactions with *bro-1* and *rnt-1*, or with the seam fate (or both), requires further investigation. What is clear is that the pathways in which these genes operate are branched and interwoven; several of the genes (notably *elt-1, pal-1* and *unc-62*, and likely *bro-1* and *rnt-1* too) have roles outside of the seam and within the seam lineage they appear to work together in both linear and parallel pathways. Revelation of the latter has led to identification of previously uncharacterised roles of *bro-1* and *rnt-1* in embryonic development. Mapping out this tangle of interactions will likely be complicated and certainly requires further investigation. However, the findings from this study should act as a starting point for a thorough understanding of the mechanisms and genes associated with the functioning of *rnt-1* in the worm.

Appendix

Table of primers

Primer	Sequence (5'→3')
CB1	AAGCTTCTATTTCAACACATTCCCGTTAC
CB2	AGATCTATTGTAATCTTGTCGGATCTTTG
CB5	GCTAGCGTGGGTATGACTCATGATTCTCAAC
CB6	ACCGGTGAAAATTCAAGAACAATGAAACTATTGC
CB7	GGAAGTGTGGTTTGGAAAGAAAC
CB11	GAAAATTCAAGAACAATGAAACTATTGC
CB17	CTCGGTAATTTGGCGAACTC
CB21	TCTAGACAGACATGTGAAAAGAGGCG
CB22	TCTAGAGTCGGATCTTTGTCCTTGTTAG
CB23	CGACCTGCAGGCATGCCAGCAAATGGGAAGACCATCGC
CB26	CAATCTGCAACCACCCATC
CB27	AAGCTTCACAGTGCTCTTCTAATTCACTATCAG
CB30	CTCTACGTAACGGGAATGTGTTG
CB31	AAGAAGTAGGTGGAAAAGAAATGG
CB32	AAGCTTGAAACCAAATCAAAATTTAGCCC
CB33	TCTAGAGTAGGGCTCAAATCTGCAGTAATAC
CB36	TCTAGACAGTGTAAATGAAAATAGCAGAACTGA
CB41	GCTAGCCCACACCACCCTTACTTTGG
CB46	GCAAAGAGTGTATGTTGTTGTAGG
CB47	ATCGTATGTGTGTTTTCAGACTTTC
CB60	CAATAGATCTAAGGGCGAATTCTG
CB90	GCTTGCATGCCTGCAGGTCG
CB90	GCTTGCATGCCTGCAGGTCG
CB91	AAACAGTTATGTTTGGTATATTGGG
CB91	AAACAGTTATGTTTGGTATATTGGG
CB92	AAGGGCCCGTACGGCCGACTAGTAGG
CB92	AAGGGCCCGTACGGCCGACTAGTAGG
CB94	GAATGATGGAAATCCACCGTAC
CB111	ACGTCAACTCCGATGATGC
CB128	CAGTGATGGGAAAACATTACCAAC
CB147	AATGTCTCCTTGTAGTCCTGGTG
CB148	GTCCACTCTGGAGGCACTATTTC
CB150	CCCAAGCTTCAATTTCGTCGAAAAAATAACG
CB151	CTCTAGAGTTTGAGAATATTCACCAAGTGG
CB152	GGTGACTATGGTACACCTGACG
CB153	GAAGAATTGTGAGACCCGTTTC
CB154	GATTACCAAATTCCGCGCTTA
CB155	CTTACCGTCAAATTGTGAAATACCC

Primer	Sequence (5'→3')
CB164	ATGAAAAGAACGACAACGGAC
CB176	TTATTCGTGTGGTGCTGAGAGATC
CB196	CCACACATCCACGATATAATATGC
CB197	CATAAATAGACTCCTTATCCCGTTTC
CB198	GTGTCGGCAACGACGCTG
CB199	AAGAAGTACCCTCTGCTCT
CB204	GATCCGACAAGATTACAATCCACAT
CB205	GATCCGACAATAGTACAATCCACAT
CB206	ATGTGGATTGTAATCTTGTCGGATC
CB207	ATGTGGATTGTACTATTGTCGGATC
CB216	TCAAATGGGAAGACCATCG
CB221	ACCAGCTATCTAACTATTCCCAACAG
CB223	CGACCTGCAGGCATGCAAGCTTAGCCGAATCTTCTGTTTGTCAC
CB232	CATGCGAAGAAGTTCGAAATG
CB233	CCGATTACTAGGCTTTCTGAAAAC
CB237	GCTCATGTTCCTGCAAGAGAG
CB247	ATGGCGCAGAGGCAGTTC
CB266	GCGGTTCCATGCTTCTG
CB267	CCAGAACCCAATTGACTGAGG
CB268	GTATTACTTCGGCGGGGAG
CB269	CGAGCAAACTTGGACCAGTC
CB276	CGACCTGCAGGCATGCAAGCTGGTCTTGGAAAGCTGGTTG
CB277	ATGGGTTTAACAATCTCCTCCC
CB278	TTAGGTCTTGGAAAGCTGGTTG
CB279	GAGCGAATGTCAGATTGGC
CB279	GAGCGAATGTCAGATTGGC
CB281	GTGATTCGGATGTGGACATTG
CB282	GCACAACTTACATGTCCAAATTC
CB289	ATGTCCTCGAATTTTCCCG
CB290	TTATGCGAACGACATTCCAG
CB291	CTTTCCCAAAAAAGACAGAAATTC
CB292	GCTGATTAGCTACACCGATGG
CB293	GAATACGCTGTACCAAGGG
CB294	ATGAACTTCAAGATGGTT
CB298	GAAGTGTCACTTAAAATGGTCCG
CB299	GGTGAGTGTTCCAAGGAGG
CB307	CGT GTG CCT GTA CAG TGA TG
CB308	CCA AAT AAG CAG CTG AGC C
CB315	TTACATTGGCTGTACAATTCGTCG
CB317	CGACTTGACATCGACCGAC
CB318	CAATCATTCCCCGCTTCTC

Primer	Sequence (5'→3')
CB319	GATCCATCAGGATCTGTCACTTG
CB354	CATCAAATGGGATGTTCTACATG
CB355	GAGATTGGCTGGGTGTGTC
CB375	GGAAAGACACTGTTAGATGTCTTCTAAC
CB376	ATCCATTTCTTTTCCACCTACTTC
CB377	ATAGCGCAGAGGCAGTTC
CB390	AGCGGAATTAATTCCCGAG
CB391	CTCTTTTTTTGGGTTTGGTGG
CB392	GTTGAACTGCCTCTGCGCCATGGCGGTACCCAATTCGAC

List of strains used in this study

Strain	Genotype	Plasmids in extrachromosomal arrays
AW60	*him-5 (e1490);wIs51* [*scm::gfp + unc-119*$^{+}$]V	
AW186	*bro-1 (tm1183) I; him-5 (e1490) wIs51 [scm::gfp + unc-119*$^{+}$*] V*	
AW187	*rnt-1 (tm388) I; him-5 (e1490) wIs51 [scm::gfp + unc-119*$^{+}$*] V*	
AW196	*unc-119(ed3) III; him-8(e1489) IV; ouEx44[pAW304 + pAW302 +unc-119*$^{+}$*]*	*pAW304 = bro-1CNE::gfp* *pAW302 = scm::rfp*
AW197	*ccIs4251[myo-3::gfp] I; unc-119(ed3) III; him-8(e1489) IV; ouEx45[pAW305 + unc-119*$^{+}$*]*	*pAW305 = bro-1*ΔCNE*::dsRED2*
AW198	*bro-1(tm1183) I; unc-119(ed3) III; him-8(e1489) IV; ouEx45[pAW305+ unc-119*$^{+}$*]*	*pAW305 = bro-1*ΔCNE*::dsRED2*
AW204	*bro-1(tm1183) I; unc-119(ed3) III; him-8(e1489) IV; ouEx49[pAW303 + unc-119*$^{+}$*]*	*pAW303 = bro-1::dsRED2*
AW298	*wIs78 [scm::gfp + ajm-1::gfp + unc-119*$^{+}$*] IV; him-5(e1490) V*	
AW302	*rnt-1(tm388); bro-1(tm1183); wIs51 [scm::gfp + unc-119*$^{+}$*] V*	
AW313	*bro-1(tm1183) I; him-8(e1489) IV*	
AW326	*bro-1(tm1183) I; unc-119(ed3) III; him-8(e1489) IV; ouEx115[pAW390 + unc-119*$^{+}$*]*	*pAW390 = bro-1*CNE ΔGATA site A*::bro-1*cDNA*::gfp*
AW392	*unc-62(e644)*V; *wIs51 [scm::gfp + unc-119*$^{+}$*] V*	
AW525	*unc-119(ed3) III; syIs78 [ajm-1::gfp + unc-119*$^{+}$*]; arIs99 [dpy-7p::yfp]*	
AW527	*unc-119(ed3) III; ouEx213* [*ajm-1::mcherry + pAW549 + unc-119*$^{+}$]	*pAW549 = eff-1p::gfp*
AW528	*bro-1(tm1183) I; unc-119(ed3) III; him-8(e1489) IV; ouEx214[pAW373+ unc-119*$^{+}$*]*	*pAW373 = bro-1*CNE*::bro-1*cDNA*::gfp*
AW529	*bro-1(tm1183) I; unc-119(ed3) III; him-8(e1489) IV; ouEx215[pAW393+ unc-119*$^{+}$*]*	*pAW393= bro-*1CNE ΔGATA site B*::bro-1*cDNA*::gfp*
AW534	*bro-1 (tm1183) I; wIs78 [scm::gfp + ajm-1::gfp + unc-119*$^{+}$*]IV*	
AW536	*eff-1(hy21) II; arIs99 [dpy-7p::yfp]; syIs78*	

	[ajm-1::gfp + unc-119^{+}]	
AW551	*eff-1(hy21) II; wIs78 [scm::gfp + ajm-1::gfp + unc-119^{+}]IV*	
AW552	*unc-119(ed3) III; rde-1(ne219) V; ouEx233 [pMF1 + ajm-1::mCherry + pAW559 + unc-119^{+}]*	*pAW559 = scm::rde-1*cDNA *pMF1 = scm::gfp*
AW553	*unc-119(ed4)III; rde-1(ne219)V*	
AW560	*unc-119(ed3) III;ouEx240 [pAW564 + ajm-1::mCherry + unc-119^{+}]*	*pAW564 = eff-1p::gfp*
AW561	*unc-119(ed3) III; ouEx241 [pAW545 + unc-119^{+}]*	*pAW545* = non-conserved region of intron 2*::gfp*
AW562	*unc-119(ed3) III; ouEx242 [pAW546 + unc-119^{+}]*	*pAW546* = intergenic region upstream of *bro-1* ORF::*gfp*
AW669	*unc-119(ed3) III; him-8(e1489) IV; ouEx614[pAW645 +unc-119^{+}]*	*pAW645 = bro-1* 5' intergenic region
AW670	*unc-119(ed3) III; him-8(e1489) IV; ouEx615[pAW647 +unc-119^{+}]*	*pAW647*= entire *bro-1* first intron
AW673	*unc-62(ku234)*V; *wIs51 [scm::gfp + unc-119^{+}] V*	
AW674	*bro-1(tm1183); rnt-1(tm388); unc-62(ku234); wIs51[scm::gfp + unc-119^{+}] V*	
AW675	*unc-62(ku234)*V;msIs114; msIs344	msIs114 = *rnt-1::gfp*; msIs344 = *bro-1::gfp*
AW676	*unc-119(ed3) III; syIs78 [ajm-1::gfp + unc-119^{+}]; [unc-62::cfp+ unc-119^{+}]*	
AW677	*unc-119(ed3) III;* ouEx616[pAW618+ *unc-119^{+}]*	*pAW618 = pscm::WT unc-62* 1b cDNA
AW678	*unc-119(ed3) III;* ouEx617[pAW619+ *unc-119^{+}]*	*pAW619 = pscm::ku234 unc-62* 1b cDNA
AW679	*unc-62(e644)*V; ceh-20(mu290)III; *wIs51 [scm::gfp + unc-119^{+}] V*	
AW680	*unc-119(ed3) III;unc-62(ku234); [unc-62::cfp+ unc-119^{+}]*	
AW681	*unc-119(ed3) III;unc-62(e644); [unc-62::cfp+ unc-119^{+}]*	
AW682	*arIs99 [dpy-7p::yfp]; unc-62(ku234)V*	
AW683	*unc-119(ed3) III; him-8(e1489) IV; ouEx618[pAW661 +unc-119^{+}]*	*pAW661* = ∆UNC-62 sites *bro-1* cDNA*::gfp*
AW684	*unc-119(ed3) III; him-8(e1489) IV; ouEx619[pAW603 +unc-119^{+}]*	*pAW603* = Proximal 4kb of *unc-62* 1a isoform 5' sequence::*gfp*
AW685	*unc-119(ed3) III; him-8(e1489) IV;*	*pAW604* = Distal 4kb of *unc-62* 1a

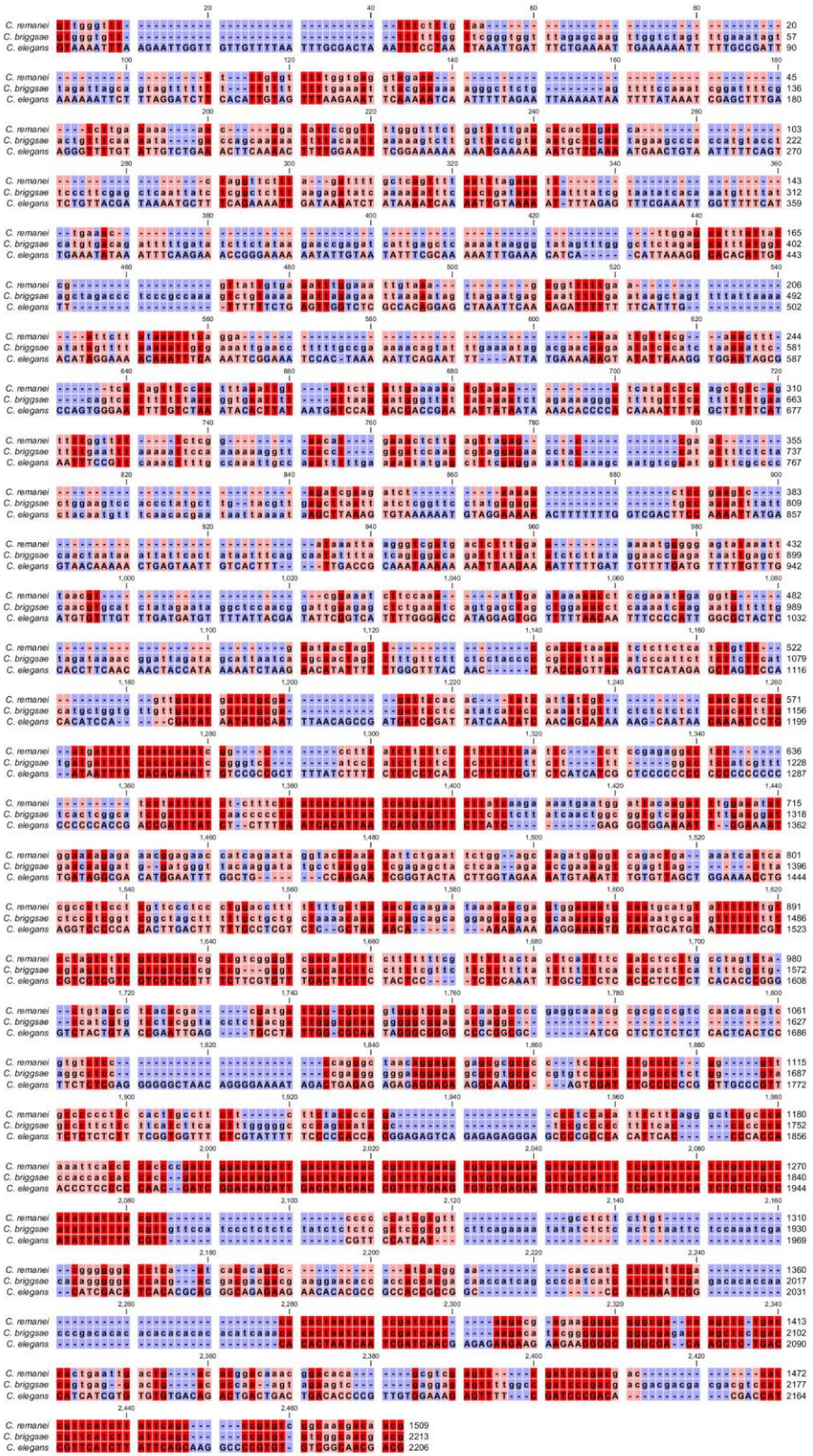
	ouEx620[pAW604 +unc-119^{+}]	isoform 5' sequence::*gfp*
AW686	*unc-119(ed3) III; him-8(e1489) IV; ouEx621[pAW607 +unc-119^{+}]*	*pAW607* = 8kb 5' of *unc-62* 1a isoform 5' sequence::*gfp*
AW687	*unc-119(ed3) III; him-8(e1489) IV; ouEx622[pAW608 +unc-119^{+}]*	*pAW608= unc-62* 1b promoter::*gfp*
AW688	*unc-119(ed3) III; him-8(e1489) IV; ouEx623[pAW620 +unc-119^{+}]*	*pAW620 = scm::gfp* fragment
AW689	*unc-119(ed3) III; him-8(e1489) IV; ouEx624[pAW621 +unc-119^{+}]*	*pAW621 = scm::gfp* fragment
AW690	*unc-119(ed3) III; him-8(e1489) IV; ouEx625[pAW622 +unc-119^{+}]*	*pAW622 = scm::gfp* fragment
AW691	*unc-119(ed3) III; him-8(e1489) IV; ouEx626[pAW623 +unc-119^{+}]*	*pAW623 = scm::gfp* fragment
AW692	*unc-119(ed3) III; him-8(e1489) IV; ouEx627[pAW624 +unc-119^{+}]*	*pAW624 = scm::gfp* fragment
AW693	*unc-119(ed3) III; him-8(e1489) IV; ouEx628[pAW625 +unc-119^{+}]*	*pAW625 = scm::gfp* fragment
AW694	*unc-119(ed3) III; him-8(e1489) IV; ouEx629[pAW626 +unc-119^{+}]*	*pAW626 = scm::gfp* fragment
AW695	*unc-119(ed3) III; him-8(e1489) IV; ouEx630[pAW627+unc-119^{+}]*	*pAW627= scm::gfp* fragment
AW696	*unc-119(ed3) III; him-8(e1489) IV; ouEx631[pAW623 +unc-119^{+}]*	*pAW623 = scm::gfp* fragment
AW697	*unc-119(ed3) III; him-8(e1489) IV; ouEx632[pAW628 +unc-119^{+}]*	*pAW628 = arf-3::gfp*
AW698	*unc-119(ed3) III; him-8(e1489) IV; ouEx633[pAW376+unc-119^{+}]*	*pAW376 = rnt-1* long intron fragment
AW699	*unc-119(ed3) III; him-8(e1489) IV; ouEx634[pAW397 +unc-119^{+}]*	*pAW397 = rnt-1* long intron fragment
AW700	*unc-119(ed3) III; him-8(e1489) IV; ouEx635[pAW364 +unc-119^{+}]*	*pAW364= rnt-1* long intron fragment
AW701	*unc-119(ed3) III; him-8(e1489) IV; ouEx636[pAW398 +unc-119^{+}]*	*pAW398= rnt-1* long intron fragment
AW702	*unc-119(ed3) III; him-8(e1489) IV; ouEx637[pAW395 +unc-119^{+}]*	*pAW395= rnt-1* long intron fragment
AW703	*unc-119(ed3) III; him-8(e1489) IV; ouEx638[pAW399+unc-119^{+}]*	*pAW399 = rnt-1* long intron fragment
AW704	*unc-119(ed3) III; him-8(e1489) IV; ouEx639[pAW394 +unc-119^{+}]*	*pAW394= rnt-1* long intron fragment
AW705	*unc-119(ed3) III; him-8(e1489) IV; ouEx640[pAW364 +unc-119^{+}]*	*pAW634 = pal-1::gfp*

AW706	*unc-119(ed3) III; him-8(e1489) IV; ouEx641[pAW636+unc-119^{+}]*	*pAW636 = pal-1 5'* intergenic region*::gfp*
AW707	*unc-119(ed3) III; him-8(e1489) IV; ouEx642[pAW638+unc-119^{+}]*	*pAW638 = pal-1* last intron in *pPD107.94*
AW708	*unc-119(ed3) III; him-8(e1489) IV; ouEx643[pAW640+unc-119^{+}]*	*pAW640 = pal-1(e2091)* last intron in *pPD107.94*
AW709	*pal-1(e2091)III; him-5(e1490)V*	
AW710	*pal-1(e2091)III; him-5(e1490)V; wIs51 [scm::gfp + unc-119^{+}] V*	
AW711	*pal-1(e2091); rnt-1(tm388); bro-1(tm1183); wIs51 [scm::gfp + unc-119^{+}]V*	
AW712	*pal-1(e2091); him-5 V; wIs78 [scm::gfp + ajm-1::gfp + unc-119^{+}]IV*	
AW713	*pal-1(e2091); him-5 V; arIs99 [dpy-7p::yfp]; syIs78 [ajm-1::gfp + unc-119^{+}]*	
AW714	*arf-3(tm1877)IV/+*	
AW715	*arf-3(tm1877)IV; dpy-20(e1282)IV; unc-24(e138)IV*	
AW716	*arf-3(tm1877)IV; dpy-20(e1282)IV; unc-24(e138)IV; wIs51 [scm::gfp + unc-119^{+}] V*	
AW717	*bro-1(tm1183); rnt-1(tm388);unc-119(ed3) III; him-8(e1489) IV; ouEx644[pAW628 + ajm-1::mCherry +unc-119^{+}]*	*pAW628 = arf-3::gfp*
AW718	*rrf-3(pk1426)II; wIs51 [scm::gfp + unc-119^{+}] V*	
AW719	*pal-1(e2091); rnt-1(tm388); bro-1(tm1183); Is[myo-3::GFP]*	
CB1489	*him-8(e1489) IV*	
CB644	*unc-62(e644)V*	
DL314	*unc-119(ed3) III; [unc-62::cfp+ unc-119+]*	
MU1092	*unc-62(ku234)V*	

Supplementary sequence alignment information

Alignment 1

3-species alignment of *unc-62* intronic sequence between the 1a and 1b cDNA translational start sites. Relatively high levels of conservation are apparent towards the 3' end of the intron.

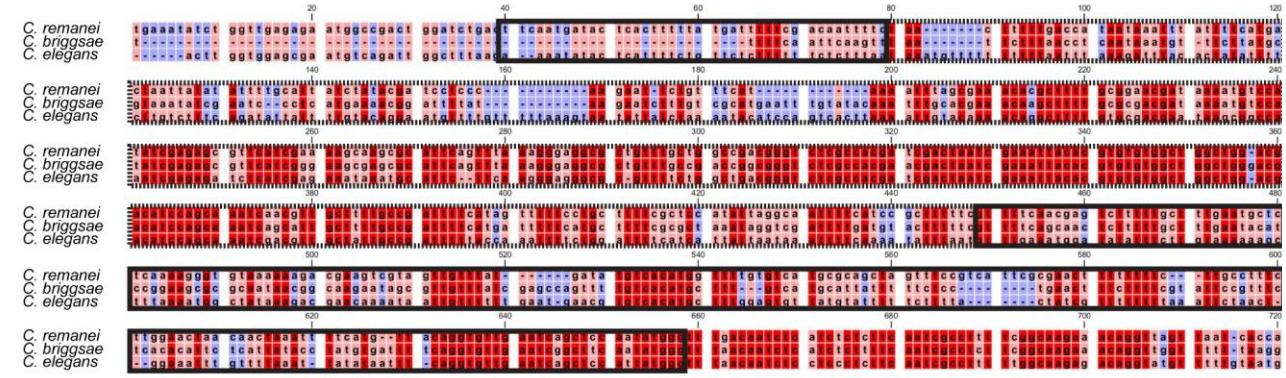


Alignment 2

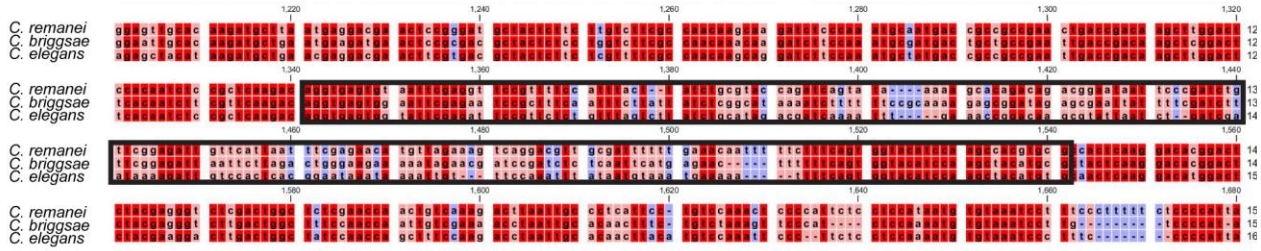
A.The 5' intergenic region between *unc-24* and *arf-3*. The uncharacterised gene *F57H12.10* is bounded by the dashed line. The solid border 5' to this gene extends up to the end of *unc-24* and contains little conservation. The solid box 3' to the dashed box shows the intergenic region between *F57H12.10* and *arf-3*. Here, particularly towards the 3' end, the sequence is relatively well conserved.

B. Conservation is also discernable at the 5' and 3' ends of the last intron of *arf-3* (black box).

A

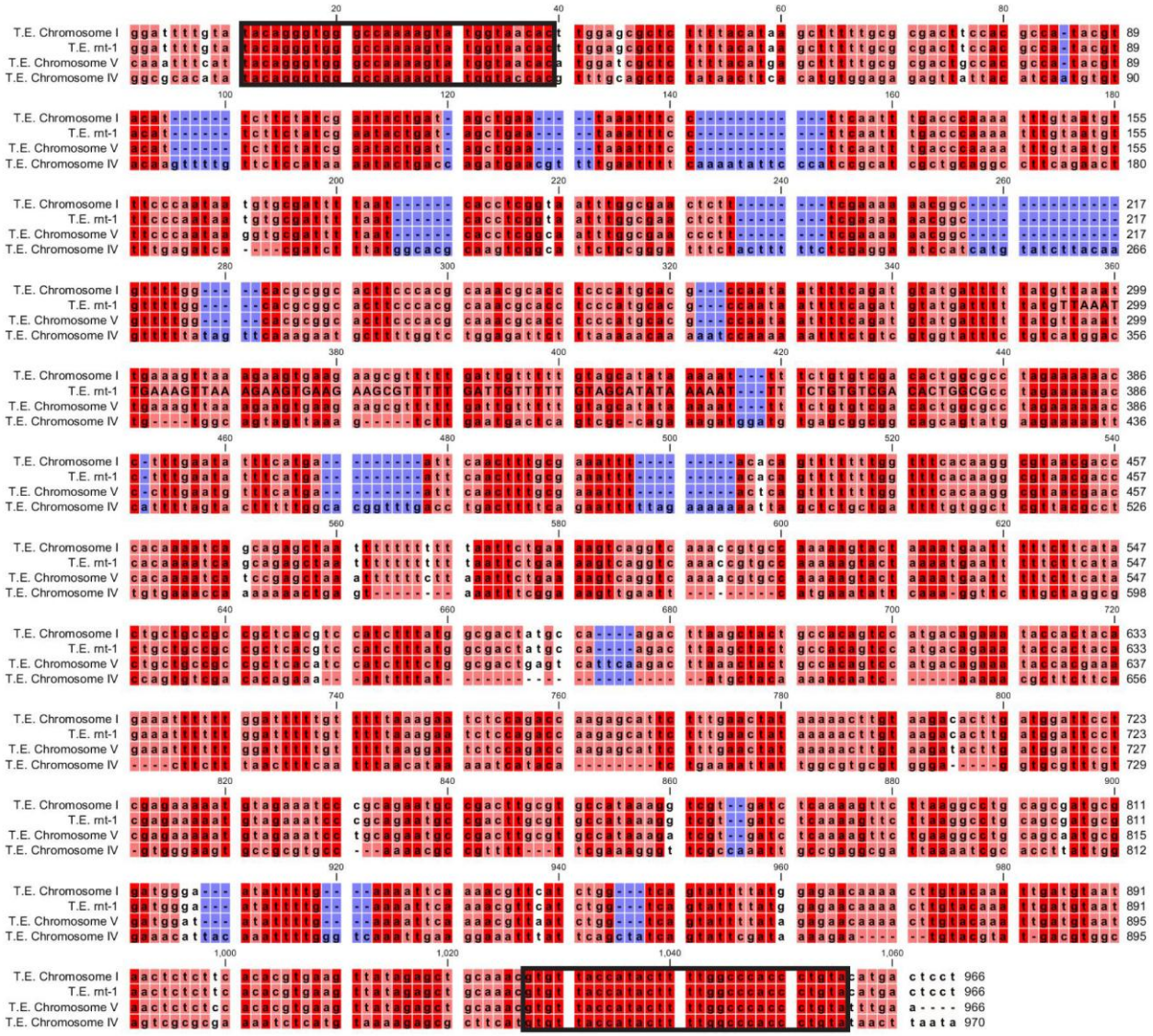


B



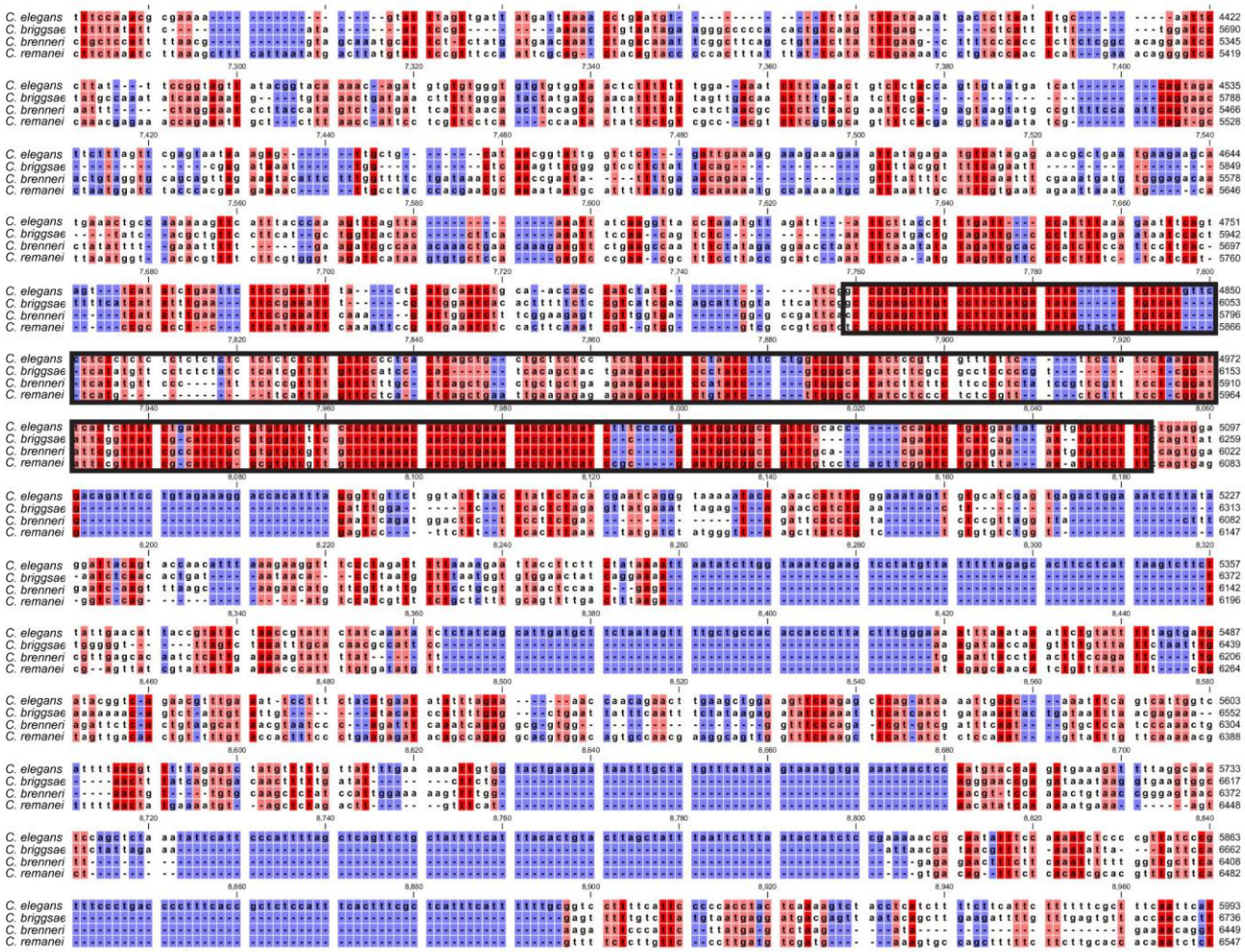
Alignment 3

The *rnt-1* long intron contains a putative transposable element. Alignment of this element with the rest of the *C. elegans* genome yields multiple close matches. Three of these are shown here. The element is bounded by inverted repeats and, as shown below, these are 100% conserved between all four sequences.



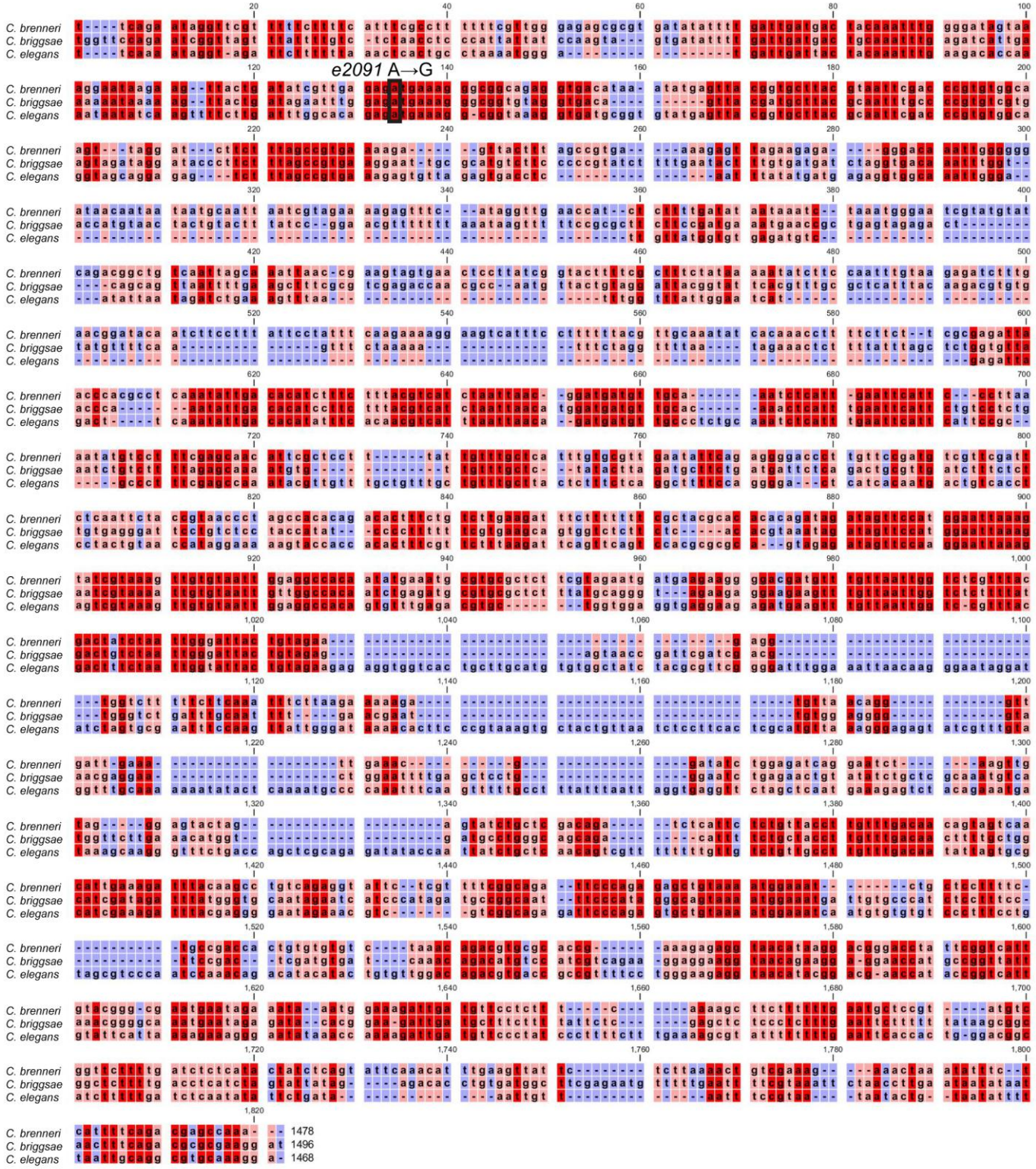
Alignment 4

4-species alignment of the *rnt-1* long intron reveals a region of high conservation across all species (bounded by black box).



Alignment 5

The last intron of *pal-1* contains several blocks of relatively well-conserved sequence. The position of the *e2091* mutation, which lies within one such region, is highlighted.



References

Abu-Shaar, M., Ryoo, H. D. and Mann, R. S. (1999). "Control of the nuclear localization of Extradenticle by competing nuclear import and export signals." Genes & Development **13**(8): 935-945.

Adya, N., Castilla, L. H. and Liu, P. P. (2000). "Function of CBFb/Bro proteins." Seminars in Cell & Developmental Biology **11**(5): 361-368.

Affolter, M., Marty, T. and Vigano, M. A. (1999). "Balancing import and export in development." Genes & Development **13**(8): 913-915.

Albert Hubbard, E. J. (2007). "*Caenorhabditis elegans* germ line: A model for stem cell biology." Developmental Dynamics **236**(12): 3343-3357.

Altun, Z. F. and Hall, D. H. (2009). Introduction. WormAtlas.

Ambros, V. and Horvitz, H. R. (1984). "Heterochronic mutants of the nematode Caenorhabditis elegans." Science **226**(4673): 409-416.

Antebi, A., Yeh, W.-H., Tait, D., Hedgecock, E. M. and Riddle, D. L. (2000). "*daf-12* encodes a nuclear receptor that regulates the dauer diapause and developmental age in *C. elegans*." Genes & Development **14**(12): 1512-1527.

Arata, Y., Kouike, H., Zhang, Y., Herman, M. A., Okano, H. and Sawa, H. (2006). "Wnt Signaling and a Hox Protein Cooperatively Regulate PSA-3/Meis to Determine Daughter Cell Fate after Asymmetric Cell Division in *C. elegans*." Developmental Cell **11**(1): 105-115.

Arnone, M. I. and Davidson, E. H. (1997). "The hardwiring of development: organization and function of genomic regulatory systems." Development **124**(10): 1851-1864.

Aronson, B., Fisher, A., Blechman, K., Caudy, M. and Gergen, J. P. (1997). "Groucho-dependent and -independent repression activities of Runt domain proteins." Mol. Cell. Biol. **17**(9): 5581-5587.

Austin, J. and Kenyon, C. (1994). "Cell contact regulates neuroblast formation in the *Caenorhabditis elegans* lateral epidermis." Development **120**(2): 313-323.

Barrière, A., Gordon, K. L. and Ruvinsky, I. (2011). "Distinct Functional Constraints Partition Sequence Conservation in a *cis*-Regulatory Element." PLoS Genet **7**(6): e1002095.

Baugh, L. R. and Sternberg, P. W. (2006). "DAF-16/FOXO Regulates Transcription of *cki-1*/Cip/Kip and Repression of *lin-4* during *C. elegans* L1 Arrest." Current Biology **16**(8): 780-785.

Baugh, L. R., Wen, J., Hill, A. A., Slonim, D.K., Brown, E.L. and Hunter, C.P. (2005). "Synthetic lethal analysis of *Caenorhabditis elegans* posterior embryonic patterning genes identifies conserved genetic interactions." Genome Biology **6**(5): R45.

Blaxter, M. (2011). "Nematodes: The Worm and Its Relatives." PLoS Biol **9**(4): e1001050.

Blyth, K., Cameron, E.R. and Neil, J.C. (2005). "The RUNX genes: gain or loss of function in cancer." Nat Rev Cancer **5**(5): 376-387.

Boxem, M. and van den Heuvel, S. (2001). "*lin-35* Rb and *cki-1* Cip/Kip cooperate in developmental regulation of G1 progression in *C. elegans*." Development **128**(21): 4349-4359.

Brabin, C., Appleford, P. J. and Woollard, A. (2011). "The *Caenorhabditis elegans* GATA Factor ELT-1 Works through the Cell Proliferation Regulator BRO-1 and the Fusogen EFF-1 to Maintain the Seam Stem-Like Fate." PLoS Genet **7**(8): e1002200.

Breeden, L. and Nasmyth, K. (1985). "Regulation of the yeast HO gene." Cold Spring Harb Symp Quant Biol. **50**: 643-650.

Brenner, S. (1974). "The genetics of *Caenorhabditis elegans*" Genetics **77**(1): 71-94.

Bürglin, T. R. (1997). "Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals." Nucleic Acids Research **25**(21): 4173-4180.

Bürglin, T. R. (1998). "The PBC domain contains a MEINOX domain: Coevolution of Hox and TALE homeobox genes?" Development Genes and Evolution **208**(2): 113-116.

Byerly, L., Cassada, R. C. and Russell, R.L. (1976). "The life cycle of the nematode *Caenorhabditis elegans*: I. Wild-type growth and reproduction." Developmental Biology **51**(1): 23-33.

Cassada, R. C. and Russell, R. L. (1975). "The dauerlarva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*." Developmental Biology **46**(2): 326-342.

Cassata, G., Shemer, G., Morandi, P., Donhauser, R., Podbilewicz, B. and Baumeister, R. (2005). "*ceh-16*/engrailed patterns the embryonic epidermis of *Caenorhabditis elegans*." Development **132**(4): 739-749.

Ch'ng, Q. and Kenyon, C. (1999). "*egl-27* generates anteroposterior patterns of cell fusion in *C. elegans* by regulating Hox gene expression and Hox protein function." Development **126**(15): 3303-3312.

Chalfie, M. and Au, M. (1989). "Genetic control of differentiation of the *Caenorhabditis elegans* touch receptor neurons." Science **243**(4894): 1027-1033.

Chang, C., Jacobs, Y., Nakamura, T., Jenkins, N. A., Copeland, N. G. and Cleary, M. L. (1997). "Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins." Mol. Cell. Biol. **17**(10): 5679-5687.

Chen, M. J., Yokomizo, T., Zeigler, B. M., Dzierzak, E. and Speck, N. A. (2009). "*Runx1* is required for the endothelial to haematopoietic cell transition but not thereafter." Nature **457**(7231): 887-891.

Chen, Z., Eastburn, D. J. and Han, M. (2004). "The *Caenorhabditis elegans* Nuclear Receptor Gene *nhr-25* Regulates Epidermal Cell Development." Mol. Cell. Biol. **24**(17): 7345-7358.

Clark, S. G., Chisholm, A. D. and Horvitz, H. R. (1993). "Control of cell fates in the central body region of *C. elegans* by the homeobox gene *lin-39*." Cell **74**(1): 43-55.

Clevers, H. (2005). "Stem cells, asymmetric division and cancer." Nat Genet **37**(10): 1027-1028.

Coffman, J. (2003). "Runx transcription factors and the developmental balance between cell proliferation and differentiation." Cell Biol Int **27**(4): 315 - 324.

Coffman, J. A., Kirchhamer, C. V., Harrington, M. G. and Davidson, E. H. (1997). "SpMyb functions as an intramodular repressor to regulate spatial expression of CyIIIa in sea urchin embryos." Development **124**(23): 4717-4727.

The *C. elegans* Sequencing Consortium (1998). "Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology." Science **282**(5396): 2012-2018.

Costa, M., Weir, M., Coulson, A., Sulston, J. and Kenyon, C. (1988). "Posterior pattern formation in *C. elegans* involves position-specific expression of a gene containing a homeobox." Cell **55**(5): 747-756.

Cox, G. N., Staprans, S. and Edgar, R. S. (1981). "The cuticle of *Caenorhabditis elegans*: II. Stage-specific changes in ultrastructure and protein composition during postembryonic development." Developmental Biology **86**(2): 456-470.

Dearolf, C. R., Topol, J. and Parker, C.S. (1989). "The caudal gene product is a direct activator of fushi tarazu transcription during *Drosophila* embryogenesis." Nature **341**(6240): 340-343.

Deretic, D., Williams, A.H., Ransom, N., Morel, V., Hargrave, P.A. and Arendt, A. (2005). "Rhodopsin C terminus, the site of mutations causing retinal disease, regulates trafficking by binding to ADP-ribosylation factor 4 (ARF4)." Proceedings of the National Academy of Sciences of the United States of America **102**(9): 3301-3306.

Dieterich, C., Clifton, S.W., Schuster, L.N., Chinwalla, A., Delehaunty, K., Dinkelacker, I., Fulton, L., Fulton, R., Godfrey, J., Minx, P., Mitreva, M., Roeseler, W., Tian, H., Witte, H., Yang, S.P., Wilson, R.K. and Sommer, R.J. (2008). "The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism." Nat Genet **40**(10): 1193-1198.

Doe, C. Q. and Bowerman, B. (2001). "Asymmetric cell division: fly neuroblast meets worm zygote." Current Opinion in Cell Biology **13**(1): 68-75.

Donaldson, J. G. and Jackson, C. L. (2011). "ARF family G proteins and their regulators: roles in membrane transport, development and disease." Nature Reviews Molecular Cell Biology **12**(6): 362-375.

Edgar, L. G., Carr, S., Wang, H. and Wood, W. B. (2001). "Zygotic Expression of the caudal Homolog *pal-1* Is Required for Posterior Patterning in *Caenorhabditis elegans* Embryogenesis." Developmental Biology **229**(1): 71-88.

Eisenmann, D. M. (2010). "*C. elegans* seam cells as stem cells: Wnt signaling and casein kinase I alpha regulate asymmetric cell divisions in an epidermal progenitor cell type." Cell Cycle **10**(1): 20-21.

Ekker, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E. and Beachy, P. A. (1994). "The degree of variation in DNA-sequence recognition among 4 *Drosophila* homeotic proteins." Embo Journal **13**(15): 3551-3560.

Ereskovsky, A. V. and Dondua, A. K. (2006). "The problem of germ layers in sponges (Porifera) and some issues concerning early metazoan evolution." Zoologischer Anzeiger - A Journal of Comparative Zoology **245**(2): 65-76.

Evans, T., Reitman, M. and Felsenfeld, G. (1988). "An erythrocyte-spefic DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes." Proceedings of the National Academy of Sciences of the United States of America **85**(16): 5976-5980.

Ferretti, E., Schulz, H., Talarico, D., Blasi, F. and Berthelsen, J. (1999). "The PBX-Regulating Protein PREP1 is present in different PBX-complexed forms in mouse." Mechanisms of Development **83**(1-2): 53-64.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., Mello, C. C. (1998). "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*." Nature **391**(6669): 806-811.

Frand, A. R., Russel, S. and Ruvkun, G. (2005). "Functional Genomic Analysis of *C. elegans* Molting." PLoS Biol **3**(10): e312.

Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M. and Ahringer, J. (2000). "Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference." Nature **408**(6810): 325-330.

Fritz, J.-A. and Behm, C. A. (2009). "CUTI-1: A Novel Tetraspan Protein Involved in *C. elegan*s CUTicle Formation and Epithelial Integrity." PLoS ONE **4**(4): e5117.

Fukushige, T., Goszczynski, B. Tian, H. and McGhee, J. D. (2003). "The evolutionary duplication and probable demise of an endodermal GATA factor in *Caenorhabditis elegans.*" Genetics **165**(2): 575-588.

Fukuyama, M., Gendreau, S. B., Derry, W. B. and Rothman, J. H. (2003). "Essential embryonic roles of the CKI-1 cyclin-dependent kinase inhibitor in cell-cycle exit and morphogenesis in *C. elegans*." Developmental Biology **260**(1): 273-286.

Furger, A., Monks, J. and Proudfoot, N. J. (2001). "The retroviruses Human Immunodeficiency Virus Type 1 and Moloney Murine Leukemia Virus adopt radically different strategies to regulate promoter-proximal polyadenylation." J. Virol. **75**(23): 11735-11746.

Gattegno, T., Mittal, A., Valansi, C., Nguyen, K. C., Hall, D. H., Chernomordik, L.V. and Podbilewicz, B. (2007). "Genetic Control of Fusion Pore Expansion in the Epidermis of *Caenorhabditis elegans*." Mol Biol Cell **18**(4): 1153-1166.

Gendreau, S. B., Moskowitz, I. P., Terns, R. M. and Rothman, J. H. (1994). "The Potential to Differentiate Epidermis Is Unequally Distributed in the AB Lineage during Early Embryonic Development in *C. elegans*." Developmental Biology **166**(2): 770-781.

George, S. E., Simokat, K., Hardin, J., Chisholm, A. D. (1998). "The VAB-1 Eph Receptor Tyrosine Kinase Functions in Neural and Epithelial Morphogenesis in *C. elegans*." Cell **92**(5): 633-643.

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

Gleason, J. E. and Eisenmann, D. M. (2010). "Wnt signaling controls the stem cell-like asymmetric division of the epithelial seam cells during *C. elegans* larval development." Dev Biol **348**(1): 58-66.

Gleason, J. E., Korswagen, H. C. and Eisenmann, D. M. (2002). "Activation of Wnt signaling bypasses the requirement for RTK/Ras signaling during *C. elegans* vulval induction." Genes & Development **16**(10): 1281-1290.

Golden, J. W. and Riddle, D. L. (1984). "The *Caenorhabditis elegans* dauer larva: Developmental effects of pheromone, food, and temperature." Developmental Biology **102**(2): 368-378.

Grosberg, R. K. and Strathmann, R. R. (2007). "The evolution of multicellularity: A minor major transition?" Annual Review of Ecology Evolution and Systematics **38**: 621-654.

Guo, S. and Kemphues, K. J. (1995). "*par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed." Cell **81**(4): 611-620.

Hassan, M. Q., Gordon, J. A., Beloti, M. M., Croce, C. M., van Wijnen, A. J., Stein, J. L., Stein, G. S. and Lian, J. B. (2010). "A network connecting Runx2, SATB2, and the miR-23a similar to 27a similar to 24-2 cluster regulates the osteoblast differentiation program." Proc. Natl. Acad. Sci. U. S. A. **107**(46): 19879-19884.

He, Z., K. Eichel, K. and Ruvinsky, I. (2011). "Functional Conservation of *Cis*-Regulatory Elements of Heat-Shock Genes over Long Evolutionary Distances." PLoS ONE **6**(7): e22677.

Heid, P. J., Raich, W. B., Smith, R., Mohler, W. A., Simokat, K., Gendreau, S. B., Rothman, J. H., and Hardin, J. (2001). "The Zinc Finger Protein DIE-1 Is Required for Late Events during Epithelial Cell Rearrangement in *C. elegans*." Developmental Biology **236**(1): 165-180.

Hirsh, D., Hirsh, D., Oppenheim, D. and Klass, M. (1976). "Development of the reproductive system of *Caenorhabditis elegans*." Developmental Biology **49**(1): 200-219.

Hobert, O. (2002). "PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*." Biotechniques **32**(4): 728-730.

Hodgkin, J. (1983). "Male phenotypes and mating efficiency in *Caenorhabditis elegans*." Genetics **103**(1): 43-64.

Hoey, T. and Levine, M. (1988). "Divergent homeobox proteins recognize similar DNA sequences in *Drosophila*." Nature **332**(6167): 858-861.

Hong, Y., Roy, R. and Ambros, V. (1998). "Developmental regulation of a cyclin-dependent kinase inhibitor controls postembryonic cell cycle progression in *Caenorhabditis elegans*." Development **125**(18): 3585-3597.

Hope, I. A. (1991). "'Promoter trapping' in *Caenorhabditis elegans*." Development **113**(2): 399-408.

Hresko, M., Williams, B. D. and Waterston, R. H. (1994). "Assembly of body wall muscle and muscle cell attachment structures in *Caenorhabditis elegans*." The Journal of Cell Biology **124**(4): 491-506.

Huang, H., Rastegar, M., Bodner, C., Goh, S. L., Rambaldi, I. and Featherstone, M. (2005). "MEIS C Termini Harbor Transcriptional Activation Domains That Respond to Cell Signaling." Journal of Biological Chemistry **280**(11): 10119-10127.

Huang, X., Tian, E., Xu, Y. and Zhang, H. (2009). "The *C. elegans* engrailed homolog *ceh-16* regulates the self-renewal expansion division of stem cell-like seam cells." Dev Biol **333**(2): 337-347.

Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., Halfnight, E., Lee, D., Lin, J., Lorch, A., McKay, S., Okada, H. M., Pan, J., Schulz, A. K., Tu, D., Wong, K., Zhao, Z., Alexeyenko, A., Burglin, T., Sonnhammer, E., Schnabel, R., Jones, S. J., Marra, M. A., Baillie, D. L. and Moerman, D. G. (2007). "High-Throughput In Vivo Analysis of Gene Expression in *Caenorhabditis elegans*." PLoS Biol **5**(9): e237.

Hunter, C. P., Harris, J. M., Maloof, J. N. and Kenyon, C. (1999). "Hox gene expression in a single *Caenorhabditis elegans* cell is regulated by a caudal homolog and intercellular signals that inhibit wnt signaling." Development **126**(4): 805-814.

Hunter, C. P. and Kenyon, C. (1996). "Spatial and Temporal Controls Target *pal-1* Blastomere-Specification Activity to a Single Blastomere Lineage in *C. elegans* Embryos." Cell **87**(2): 217-226.

Hyman-Walsh, C., Bjerke, G. A., Wotton, D. (2010). "An autoinhibitory effect of the homothorax domain of Meis2." FEBS Journal **277**(12): 2584-2597.

Inoue, K., Ozaki, S., Shiga, T., Ito, K., Masuda, T., Okado, N., Iseda, T., Kawaguchi, S., Ogawa, M., Bae, S. C., Yamashita, N., Itohara, S., Kudo, N. and Ito, Y. (2002). "*Runx3* controls the axonal projection of proprioceptive dorsal root ganglion neurons." Nat Neurosci **5**(10): 946-954.

Jantsch-Plunger, V. and Fire, A. (1994). "Combinatorial structure of a body muscle-specific transcriptional enhancer in *Caenorhabditis elegans*." Journal of Biological Chemistry **269**(43): 27021-27028.

Jiang, Y., Shi, H. and Liu, J. (2009). "Two Hox cofactors, the Meis/Hth homolog UNC-62 and the Pbx/Exd homolog CEH-20, function together during *C. elegans* postembryonic mesodermal development." Developmental Biology **334**(2): 535-546.

Ji, Y. J., Nam, S., Jin, Y. H., Cha, E. J., Lee, K. S., Choi, K. Y., Song, H. O., Lee, J., Bae, S. C. and Ahnn, J. (2004). "RNT-1, the *C. elegans* homologue of mammalian RUNX transcription factors, regulates body size and male tail development" Dev Biol **274**(2): 402-412.

Ji, Y. J., Singaravelu, G. and Ahnn, J. (2005). "RNT-1 Regulation in *C. elegans*." J Cell Biochem **96**(1): 8-15.

Kagoshima, H., Nimmo, R., Saad, N., Tanaka, J., Miwa, Y., Mitani, S., Kohara, Y. and Woollard, A. (2007). "The *C. elegans* CBF-beta homologue BRO-1 interacts with the RUNX factor, RNT-1, to promote stem cell proliferation and self-renewal." Development **134**(21): 3905-3915.

Kagoshima, H., Sawa, H., Mitani, S., Bürglin, T.R., Shigesada, K. and Kohara, Y. (2005). "The *C. elegans* RUNX transcription factor RNT-1/MAB-2 is required for asymmetrical cell division of the T blast cell." Developmental Biology **287**(2): 262-273.

Kai, T. S. and Spradling, A. (2003). "An empty *Drosophila* stem cell niche reactivates the proliferation of ectopic cells." Proc. Natl. Acad. Sci. U. S. A. **100**(8): 4633-4638.

Kamachi, Y., Ogawa, E., Asano, M., Ishida, S., Murakami, Y., Satake, M., Ito, Y. and Shigesada, K. (1990). "Purification of a mouse nuclear factor that binds to both the A and B cores of the polyomavirus enhancer." J. Virol. **64**(10): 4808-4819.

Kanamori, T., Inoue, T., Sakamoto, T., Gengyo-Ando, K., Tsujimoto, M., Mitani, S., Sawa, H., Aoki, J. and Arai H. (2008). "b-Catenin asymmetry is regulated by PLA_1 and retrograde traffic in *C. elegans* stem cell divisions." Embo J **27**(12): 1647-1657.

Kiger, A. A., Jones, D. L., Schulz, C., Rogers, M. B. and Fuller, M. T. (2001). "Stem cell self-renewal specified by JAK-STAT activation in response to a support cell cue." Science **294**(5551): 2542-2545.

Kilstrup-Nielsen, C., Alessio, M. and Zappavigna, V. (2003). "PBX1 nuclear export is regulated independently of PBX-MEINOX interaction by PKA phosphorylation of the PBC-B domain." Embo J **22**(1): 89-99.

Kimble, J. (1981). "Alterations in cell lineage following laser ablation of cells in the somatic gonad of *Caenorhabditis elegans*." Developmental Biology **87**(2): 286-300.

Kipreos, E. T. (2005). "*C. elegans* cell cycles: invariance and stem cell divisions." Nat Rev Mol Cell Biol **6**(10): 766-776.

Kirchhamer, C. V., Yuh, C.H. and Davidson, E.H. (1996). "Modular cis-regulatory organization of developmentally expressed genes: two genes transcribed territorially in the sea urchin embryo, and additional examples." Proceedings of the National Academy of Sciences **93**(18): 9322-9328.

Knoepfler, P. S., Calvo, K. R., Chen, H., Antonarakis, S. E. and Kamps, M. P. (1997). "Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface

disrupted in oncoprotein E2a-Pbx1." Proceedings of the National Academy of Sciences **94**(26): 14553-14558.

Ko, L. J. and Engel, J. D. (1993). "DNA-binding specificities of the GATA transcription factor family." Mol. Cell. Biol. **13**(7): 4011-4022.

Koh, K., Peyrot, S. M., Wood, C. G., Wagmaister, J. A., Maduro, M. F., Eisenmann, D. M. and Rothman, J. H. (2002). "Cell fates and fusion in the *C. elegans* vulval primordium are regulated by the EGL-18 and ELT-6 GATA factors - apparent direct targets of the LIN-39 Hox protein." Development **129**(22): 5171-5180.

Koh, K. and Rothman, J. H. (2001). "ELT-5 and ELT-6 are required continuously to regulate epidermal seam cell differentiation and cell fusion in *C. elegans*." Development **128**(15): 2867-2880.

Kohany, O., Gentles, A.J., Hankus, L. and Jurka, J. (2006). "Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor." BMC Bioinformatics **7**(1): 474.

Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., Shimizu, Y., Bronson, R. T., Gao, Y. H., Inada, M., Sato, M., Okamoto, R., Kitamura, Y., Yoshiki, S. and Kishimoto T. (1997). "Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts." Cell **89**(5): 755-764.

Kurant, E., Pai, C. Y., Sharf, R., Halachmi, N., Sun, Y.H. and Salzberg A. (1998). "*Dorsotonals/homothorax*, the Drosophila homologue of *meis1* interacts with extradenticle in patterning of the embryonic PNS." Development **125**(6): 1037-1048.

Lee, J., Ahnn, J. and Bae, S. C. (2004). "Homologs of RUNX and CBFb/PEBP2b in *C. elegans*." Oncogene **23**(24): 4346-4352.

Lei, H., Liu, J., Fukushige, T., Fire, A. and Krause, M. (2009). "Caudal-like PAL-1 directly activates the bodywall muscle module regulator *hlh-1* in *C. elegans* to initiate the embryonic muscle gene regulatory network." Development **136**(8): 1241-1249.

Levanon, D., Bettoun, D., Harris-Cerruti, C., Woolf, E., Negreanu, V., Eilam, R., Bernstein, Y., Goldenberg, D., Xiao, C., Fliegauf, M., Kremer, E., Otto, F., Brenner, O., Lev-Tov,

A. and Groner Y. (2002). "The *Runx3* transcription factor regulates development and survival of TrkC dorsal root ganglia neurons." Embo Journal **21**(13): 3454-3463.

Levanon, D., Negreanu, V., Bernstein, Y., Bar-Am, I., Avivi, L. and Groner, Y. (1994). "AML1, AML2, and AML3, the Human Members of the *runt* domain Gene-Family: cDNA Structure, Expression, and Chromosomal Localization." Genomics **23**(2): 425-432.

Li, Q. L., Ito, K., Sakakura, C., Fukamachi, H., Inoue, K., Chi, X. Z., Lee, K. Y., Nomura, S., Lee, C. W., Han, S. B., Kim, H. M., Kim, W. J., Yamamoto, H., Yamashita, N., Yano, T., Ikeda, T., Itohara, S., Inazawa, J. , Abe, T., Hagiwara, A., Yamagishi, H., Ooe, A., Kaneda, A., Sugimura, T., Ushijima, T., Bae, S. C. and Ito, Y. (2002). "Causal relationship between the loss of *RUNX3* expression and gastric cancer." Cell **109**(1): 113-124.

Li, Y., Kelly, W. G., Logsdon, J. M. Jr., Schurko, A. M., Harfe, B. D., Hill-Harfe, K. L. and Kahn, R. A. (2004). "Functional genomic analysis of the ADP-ribosylation factor family of GTPases: phylogeny among diverse eukaryotes and function in *C. elegans*." The FASEB Journal **18**(15): 1834-1850.

Liu, J. and Fire, A. (2000). "Overlapping roles of two Hox genes and the exd ortholog *ceh-20* in diversification of the *C. elegans* postembryonic mesoderm." Development **127**(23): 5179-5190.

Liu, J., Wilson, T. E., Milbrandt, J. and Johnston, M. (1993). "Identifying DNA-Binding Sites and Analyzing DNA-Binding Domains Using a Yeast Selection System." Methods **5**(2): 125-137.

Livak, K. J. and Schmittgen, T. D. (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-DDC}T method." Methods **25**(4): 402-408.

Lutterbach, B., Westendorf, J. J., Linggi, B., Patten, A., Moniwa, M., Davie, J. R., Huynh, K. D., Bardwell, V. J., Lavinsky, R. M., Rosenfeld, M. G., Glass, C., Seto, E. and Hiebert, S. W. (1998). "ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors." Molecular and Cellular Biology **18**(12): 7176-7184.

Maduro, M. and Pilgrim, D. (1995). "Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system." Genetics **141**: 977-988.

Maloof, J. N. and Kenyon, C. (1998). "The Hox gene *lin-39* is required during *C. elegans* vulval induction to select the outcome of Ras signaling." Development **125**(2): 181-190.

Mann, R. S. (1995). "THE SPECIFICITY OF HOMEOTIC GENE-FUNCTION." Bioessays **17**(10): 855-863.

Mann, R. S. and Affolter, M. (1998). "Hox proteins meet more partners." Current Opinion in Genetics & Development **8**(4): 423-429.

Margalit, A., Neufeld, E., Feinstein, N., Wilson, K. L., Podbilewicz, B. and Gruenbaum, Y. (2007). "Barrier to autointegration factor blocks premature cell fusion and maintains adult muscle integrity in *C. elegans*." J Cell Biol **178**(4): 661-673.

Markaki, M. and Tavernarakis, N. (2010). "Modeling human diseases in *Caenorhabditis elegans*." Biotechnology Journal **5**(12): 1261-1276.

Marri, S. and Gupta, B. P. (2009). "Dissection of *lin-11* enhancer regions in *Caenorhabditis elegans* and other nematodes." Dev Biol **325**(2): 402-411.

Matunis, E., Tran, J., Gönczy, P., Caldwell, K. and DiNardo, S. (1997). "punt and schnurri regulate a somatically derived signal that restricts proliferation of committed progenitors in the germline." Development **124**(21): 4383-4391.

Matys, V., Fricke, E., Geffers, R., Gössling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A. E., Kel-Margoulis, O. V., Kloos, D. U., Land, S., Lewicki-Potapov, B., Michael, H., Münch, R., Reuter, I., Rotert, S., Saxel, H., Scheer M, Thiele S and Wingender E. (2003). "TRANSFAC: transcriptional regulation, from patterns to profiles." Nucl. Acids Res. **31**(1): 374-378.

Maurer, C. W., Chiorazzi, M. and Shaham, S. (2007). "Timing of the onset of a developmental cell death is controlled by transcriptional induction of the *C. elegans* ced-3 caspase-encoding gene." Development **134**(7): 1357-1368.

Mello, C. and Fire, A. (1995). DNA transformation. *Caenorhabditis elegans*: Modern Biological Analysis of an Organism. Epstein, H. F. and Shakes, D. (eds), H. San Diego, Academic Press. **48:** 451-482.

Merika, M. and Orkin, S. H. (1993). "DNA-binding specificity of GATA family transcription factors." Mol. Cell. Biol. **13**(7): 3999-4010.

Meyers, S., Lenny, N. and Hiebert, S.W. (1995). "The t(8;21) fusion protein interferes with AML-1B-dependent transcriptional activation." Molecular and Cellular Biology **15**(4): 1974-1982.

Herman, M. A. (2002). "Control of cell polarity by noncanonical Wnt signaling in *C. elegans*." Seminars in Cell & Developmental Biology **13**(3): 233-241.

Mizumoto, K. and Sawa, H. (2007). "Two bs or not two bs: regulation of asymmetric division by β-catenin." Trends in Cell Biology **17**(10): 465-473.

Mohler, W. A., G. Shemer, et al. (2002). "The Type I Membrane Protein EFF-1 Is Essential for Developmental Cell Fusion." Dev Cell **2**(3): 355-362.

Mohler, W. A., Simske, J. S., Williams-Masson, E. M., Hardin, J. D. and White, J.G. (1998). "Dynamics and ultrastructure of developmental cell fusions in the *Caenorhabditis elegans* hypodermis." Current Biology **8**(19): 1087-1091.

Morrison, S. J. and Kimble, J. (2006). "Asymmetric and symmetric stem-cell divisions in development and cancer." Nature **441**(7097): 1068-1074.

Morrison, S. J. and Spradling, A. C. (2008). "Stem Cells and Niches: Mechanisms That Promote Stem Cell Maintenance throughout Life." Cell **132**(4): 598-611.

Moss, E. G., Lee, R. C. and Ambros, V. (1997). "The Cold Shock Domain Protein LIN-28 Controls Developmental Timing in *C. elegans* and is Regulated by the lin-4 RNA." Cell **88**(5): 637-646.

Mukherjee, K. and Bürglin, T. (2007). "Comprehensive Analysis of Animal TALE Homeobox Genes: New Conserved Motifs and Cases of Accelerated Evolution." Journal of Molecular Evolution **65**(2): 137-153.

Nagao, T., Endo, K., Kawauchi, H., Walldorf, U. and Furukubo-Tokunaga, K. (2000). "Patterning defects in the primary axonal scaffolds caused by the mutations of the *extradenticle* and *homothorax* genes in the embryonic *Drosophila* brain." Development Genes and Evolution **210**(6): 289-299.

Nam, S., Jin, Y. H., Li, Q. L., Lee, K. Y., Jeong, G. B., Ito, Y., Lee, J. and Bae, S.C. (2002). "Expression Pattern, Regulation, and Biological Role of Runt Domain Transcription Factor, *run*, in *Caenorhabditis elegans*

" Mol. Cell. Biol. **22**(2): 547-554.

Nimmo, R., Antebi, A. and Woollard, A. (2005). "*mab-2* encodes RNT-1, a *C. elegans* RUNX homologue essential for controlling cell proliferation in a stem cell-like developmental lineage." Development **132**(22): 5043-5054.

Nimmo, R. and Slack, F. (2009). "An elegant miRror: microRNAs in stem cells, developmental timing and cancer." Chromosoma **118**(4): 405-418.

Nottingham, W. T., Jarratt, A., Burgess, M., Speck, C. L., Cheng, J. F., Prabhakar, S., Rubin, E. M., Li, P. S., Sloane-Stanley, J., Kong-A-San, J., de Bruijn, M. F. (2007). "*Runx1*-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer." Blood **110**(13): 4188-4197.

Ohlstein, B., Kai, T., Decotto, E. and Spradling, A. (2004). "The stem cell niche: theme and variations." Current Opinion in Cell Biology **16**(6): 693-699.

Okuda, T., van Deursen, J., Hiebert, S. W., Grosveld, G. and Downing, J. R. (1996). "AML1, the Target of Multiple Chromosomal Translocations in Human Leukemia, Is Essential for Normal Fetal Liver Hematopoiesis." Cell **84**(2): 321-330.

Otto, F., Thornell, A. P., Crompton, T., Denzel, A., Gilmour, K. C., Rosewell, I. R., Stamp, G.W., Beddington, R. S., Mundlos, S., Olsen, B. R., Selby, P. B. and Owen, M. J. (1997). "Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development." Cell **89**(5): 765 - 771.

Pai, C.-Y., Kuo, T. S., Jaw, T. J., Kurant, E., Chen, C. T., Bessarab, D. A., Salzberg, A. and Sun, Y. H. (1998). "The Homothorax homeoprotein activates the nuclear localization of another homeoprotein, Extradenticle, and suppresses eye development in Drosophila." Genes & Development **12**(3): 435-446.

Petalcorin, M. I. R., Oka, T., Koga, M., Ogura, K., Wada, Y., Ohshima, Y. and Futai, M. (1999). "Disruption of *clh-1*, a chloride channel gene, results in a wider body of *Caenorhabditis elegans*." Journal of Molecular Biology **294**(2): 347-355.

Phillips, B. T., Kidd, A. R. 3rd, King, R., Hardin, J. and Kimble J. (2007). "Reciprocal asymmetry of SYS-1/beta-catenin and POP-1/TCF controls asymmetric divisions in *Caenorhabditis elegans*." Proceedings of the National Academy of Sciences **104**(9): 3231-3236.

Podbilewicz, B. and White, J. G. (1994). "Cell Fusions in the Developing Epithelia of *C. elegans*." Developmental Biology **161**(2): 408-424.

Potts, M. B., Wang, D. P. and Cameron, S. (2009). "Trithorax, Hox, and TALE-class homeodomain proteins ensure cell survival through repression of the BH3-only gene *egl-1*." Developmental Biology **329**(2): 374-385.

Qadota, H., Inoue, M., Hikita, T., Köppen, M., Hardin, J. D., Amano, M., Moerman, D. G. and Kaibuchi, K. (2007). "Establishment of a tissue-specific RNAi system in *C. elegans*." Gene **400**(1-2): 166-173.

Ren, H. and Zhang, H. (2010). "Wnt signaling controls temporal identities of seam cells in *Caenorhabditis elegans*." Dev Biol **345**(2): 144-155.

Rennert, J., Coffman, J. A., Mushegian, A. R. and Robertson, A. J. (2003). "The evolution of Runx genes I. A comparative study of sequences from phylogenetically diverse model organisms." BMC Evolutionary Biology **3**(1): 4.

Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M. and Mann, R. S. (1997). "Nuclear Translocation of Extradenticle Requires homothorax, which Encodes an Extradenticle-Related Homeodomain Protein." Cell **91**(2): 171-183.

Robertson, A., Dickey, C. E., McCarthy, J. J. and Coffman, J. A. (2002). "The expression of SpRunt during sea urchin embryogenesis." Mech Dev **117**(1-2): 327 - 330.

Robertson, A., Larroux, C., Degnan, B. M. and Coffman, J. A. (2009). "The evolution of Runx genes II. The C-terminal Groucho recruitment motif is present in both eumetazoans and homoscleromorphs but absent in a haplosclerid demosponge." BMC Research Notes **2**(1): 59.

Rossi, F., MacLean, H.E., Yuan, W., Francis, R.O., Semenova, E., Lin, C.S., Kronenberg, H.M. and Cobrinik, D. (2002). "p107 and p130 Coordinately Regulate Proliferation, Cbfa1 Expression, and Hypertrophic Differentiation during Endochondral Bone Development." Developmental Biology **247**(2): 271-285.

Rual, J.-F., Ceron, J., Koreth, J., Hao, T., Nicot, A. S., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., van den Heuvel, S. and Vidal, M. (2004). "Toward Improving *Caenorhabditis elegans* Phenome Mapping With an ORFeome-Based RNAi Library." Genome Research **14**(10b): 2162-2168.

Ruvinsky, I. and Ruvkun, G. (2003). "Functional tests of enhancer conservation between distantly related species." Development **130**(21): 5133-5142.

Ryoo, H. D., Marty, T., Casares, F., Affolter, M. and Mann, R.S. (1999). "Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex." Development **126**(22): 5137-5148.

Sadakata, T., Y., Shinoda, Y., Sekine, Y., Saruta, C., Itakura, M., Takahashi, M. and Furuichi T. (2010). "Interaction of Calcium-dependent Activator Protein for Secretion 1 (CAPS1) with the Class II ADP-ribosylation Factor Small GTPases Is Required for Dense-core Vesicle Trafficking in the trans-Golgi Network." Journal of Biological Chemistry **285**(49): 38710-38719.

Sakai, Y., R., Nakagawa, R., Sato, R. and Maeda, M. (1998). "Selection of DNA Binding Sites for Human Transcriptional Regulator GATA-6." Biochem Biophys Res Comm **250**(3): 682-688.

Salser, S. J. and Kenyon, C. (1996). "A *C. elegans* Hox gene switches on, off, on and off again to regulate proliferation, differentiation and morphogenesis." Development **122**(5): 1651-1661.

Sandelin, A., Wasserman, W. W. and Lenhard, B. (2004). "ConSite: web-based prediction of regulatory elements using cross-species comparison." Nucleic Acids Research **32**(suppl 2): W249-W252.

Schofield, R. (1978). "The relationship between the spleen colony-forming cell and the haemopoietic stem cell." Blood Cells **4**: 7-25.

Seggerson, K., Tang, L. and Moss, E. G. (2002). "Two Genetic Circuits Repress the Caenorhabditis elegans Heterochronic Gene *lin-28* after Translation Initiation." Developmental Biology **243**(2): 215-225.

Shaye, D. D. and Greenwald, I. (2011). "OrthoList: A Compendium of *C. elegans* Genes with Human Orthologs." PLoS ONE **6**(5): e20085.

Shemer, G., Suissa, M., Kolotuev, I., Nguyen, K. C., Hall, D.H. and Podbilewicz, B. (2004). "EFF-1 is sufficient to initiate and execute tissue-specific cell fusion in *C. elegans*." Cur Biol **14**(17): 1587-1591.

Shim, J. and Lee, J. (2008). "Regulation of *rnt-1* expression mediated by the opposing effects of BRO-1 and DBL-1 in the nematode *Caenorhabditis elegans.*" Biochem Biophys Res Commun. **367**(1): 130-136.

Silhankova, M., Jindra, M. and Asahina, M. (2005). "Nuclear receptor NHR-25 is required for cell-shape dynamics during epidermal differentiation in *Caenorhabditis elegans*." J Cell Sci **118**(1): 223-232.

Singh, R. N. and Sulston, J. E. (1978). "Some Observations On Moulting in *Caenorhabditis elegans*." Nematologica **24**: 63-71.

Singhvi, A., Teuliere, J., Talavera, K., Cordes, S., Ou, G., Vale, R. D., Prasad, B. C., Clark, S. G. and Garriga, G. (2011). "The Arf GAP CNT-2 Regulates the Apoptotic Fate in *C. elegans* Asymmetric Neuroblast Divisions." Current biology **21**(11): 948-954.

Smith, J. A., McGarr, P. and Gilleard, J. S. (2005). "The *Caenorhabditis elegans* GATA factor *elt-1* is essential for differentiation and maintenance of hypodermal seam cells and for normal locomotion." J Cell Sci **118**(24): 5709-5719.

Song, X., Zhu, C. H., Doan, C. and Xie, T. (2002). "Germline stem cells anchored by adherens junctions in the *Drosophila* ovary niches." Science **296**(5574): 1855-1857.

Srivastava, M., Begovic, E., Chapman, J., Putnam, N. H., Hellsten, U., Kawashima, T., Kuo, A., Mitros, T., Salamov, A., Carpenter, M. L., Signorovitch, A. Y., Moreno, M. A., Kamm, K., Grimwood, J., Schmutz, J., Shapiro, H., Grigoriev, I. V., Buss, L. W., Schierwater, B., Dellaporta, S. L. and Rokhsar, D. S. (2008). "The Trichoplax genome and the nature of placozoans." Nature **454**(7207): 955-960.

Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D. H., Fulton, L. A., Fulton, R. E., Griffiths-Jones, S., Harris, T. W., Hillier, L. W., Kamath, R., Kuwabara, P. E., Mardis, E. R., Marra, M. A., Miner, T. L., Minx, P., Mullikin, J. C., Plumb, R. W., Rogers, J., Schein, J. E., Sohrmann, M., Spieth, J., Stajich, J. E., Wei, C., Willey, D., Wilson, R. K., Durbin, R. and Waterston, R.H. (2003). "The Genome Sequence of *Caenorhabditis briggsae*: A Platform for Comparative Genomics." PLoS Biol **1**(2): e45.

Stevens, K. E. and Mann, R. S. (2007). "A Balance Between Two Nuclear Localization Sequences and a Nuclear Export Sequence Governs Extradenticle Subcellular Localization." Genetics **175**(4): 1625-1636.

Strippoli, P., D'Addabbo, P., Lenzi, L., Giannone, S., Canaider, S., Casadei, R., Vitale, L., Carinci, P. and Zannotti, M. (2002). "Segmental paralogy in the human genome: a large-scale triplication on 1p, 6p, and 21q." Mammalian Genome **13**(8): 456-462.

Strom, D. K., Nip, J., Westendorf, J. J., Linggi, B., Lutterbach, B., Downing, J. R., Lenny, N. and Hiebert, S. W. (2000). "Expression of the AML-1 Oncogene Shortens the G1 Phase of the Cell Cycle." Journal of Biological Chemistry **275**(5): 3438-3445.

Sullivan, J., Sher, D., Eisenstein, M., Shigesada, K., Reitzel, A. M., Marlow, H., Levanon, D., Groner, Y., Finnerty, J. R. and Gat, U. (2008). "The evolutionary origin of the Runx/CBFb transcription factors - Studies of the most basal metazoans." BMC Evolutionary Biology **8**(1): 228.

Sulston, J. E., Albertson, D. G. and Thomson, J. N. (1980). "The *Caenorhabditis elegans* male: Postembryonic development of nongonadal structures." Developmental Biology **78**(2): 542-576.

Sulston, J. E. and Hodgkin, J. (1988). The nematode *Caenorhabditis elegans*. Methods. W. B. Wood, Cold Spring Harbor Laboratory.

Sulston, J. E. and Horvitz, H. R. (1977). "Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*." Dev. Biol. **56**: 110-156.

Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N. (1983). "The embryonic cell lineage of the nematode *Caenorhabditis elegans*." Developmental Biology **100**(1): 64-119.

Tabara, H., Grishok, A. and Mello, C. C. (1998). "RNAi in *C. elegans*: Soaking in the Genome Sequence." Science **282**(5388): 430-431.

Takacs-Vellai, K., Vellai, T., Chen, E. B., Zhang, Y., Guerry, F., Stern, M. J. and Müller, F. (2007). "Transcriptional control of Notch signaling by a HOX and a PBX/EXD protein during vulval development in *C. elegans*." Developmental Biology **302**(2): 661-669.

Taniuchi, I., Osato, M., Egawa, T., Sunshine, M. J., Bae, S. C., Komori, T., Ito, Y. and Littman, D. R. (2002). "Differential Requirements for Runx Proteins in CD4 Repression and Epigenetic Silencing during T Lymphocyte Development." Cell **111**(5): 621-633.

Thorpe, C. J., Schlesinger, A. and Bowerman, B. (2000). "Wnt signalling in *Caenorhabditis elegans*: regulating repressors and polarizing the cytoskeleton." Trends in Cell Biology **10**(1): 10-17.

Timmons, L. and Fire, A. (1998). "Specific interference by ingested dsRNA." Nature **395**(6705): 854-854.

Van Auken, K., Weaver, D., Robertson, B., Sundaram, M., Saldi, T., Edgar, L., Elling, U., Lee, M., Boese, Q. and Wood, W. B. (2002). "Roles of the Homothorax/Meis/Prep homolog UNC-62 and the Exd/Pbx homologs CEH-20 and CEH-40 in *C. elegans* embryogenesis." Development **129**(22): 5255-5268.

Waltzer, L., Ferjoux, G., Bataillé, L. and Haenlin, M. (2003). "Cooperation between the GATA and RUNX factors Serpent and Lozenge during *Drosophila* hematopoiesis." EMBO J **22**(24): 6516-6525.

Wang, Q., Stacy, T., Binder, M., Marin-Padilla, M., Sharpe, A. H. and Speck, N. A. (1996). "Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis." Proc Natl Acad Sci U S A **93**(8): 3444 - 3449.

Waring, D. A. and Kenyon, C. (1990). "Selective silencing of cell communication influences anteroposterior pattern formation in *C. elegans*." Cell **60**(1): 123-131.

Waring, D. A. and Kenyon, C. (1991). "Regulation of cellular responsiveness to inductive signals in the developing *C. elegans* nervous system." Nature **350**: 712-715.

Wildonger, J. and Mann, R. S. (2005). "The t(8;21) translocation converts AML1 into a constitutive transcriptional repressor." Development **132**(10): 2263-2272.

Williams, B. D. and Waterston, R. H. (1994). "Genes critical for muscle development and function in *Caenorhabditis elegan*s identified through lethal mutations." J. Cell Biol. **124**: 475-490.

Williams, B. D., Schrank, B., Huynh, C., Shownkeen, R. and Waterston, R. H. (1992). "A Genetic Mapping System in *Caenorhabditis elegans* Based on Polymorphic Sequence-Tagged Sites." Genetics **131**(3): 609-624.

Xia, D., Zhang, Y., Huang, X., Sun, Y. and Zhang, H. (2007). "The *C. elegans* CBF-beta homolog, BRO-1, regulates the proliferation, differentiation and specification of the stem cell-like seam cell lineages." Dev Biol **309**(2): 259-272.

Xie, T. and Spradling, A. C. (1998). "decapentaplegic is essential for the maintenance and division of germline stem cells in the *Drosophila* ovary." Cell **94**(2): 251-260.

Xie, T. and Spradling, A. C. (2000). "A niche maintaining germ line stem cells in the *Drosophila* ovary." Science **290**(5490): 328-330.

Yamashita, Y. M., Fuller, M. T. and Jones, D. L. (2005). "Signaling in stem cell niches: lessons from the *Drosophila* germline." Journal of Cell Science **118**(4): 665-672.

Yang, L., Sym, M. and Kenyon, C. (2005). "The roles of two *C. elegans* HOX co-factor orthologs in cell migration and vulva development." Development **132**(6): 1413-1428.

Ye, S., Dhillon, S., Ke, X., Collins, A. R. and Day, I.N. (2001). "An efficient procedure for genotyping single nucleotide polymorphisms." Nucleic Acids Research **29**(17): e88.

Yu, X., Odera, S., Chuang, C. H., Lu, N. and Zhou, Z. (2006). "*C. elegans* Dynamin Mediates the Signaling of Phagocytic Receptor CED-1 for the Engulfment and Degradation of Apoptotic Cells." Developmental Cell **10**(6): 743-757.

Zeng, C., van Wijnen, A. J., Stein, J. L., Meyers, S., Sun, W., Shopland, L., Lawrence, J. B., Penman, S., Lian, J. B., Stein, G. S. and Hiebert, S. W. (1997). "Identification of a nuclear matrix targeting signal in the leukemia and bone-related AML/CBF-α transcription factors." Proceedings of the National Academy of Sciences **94**(13): 6746-6751.

Zhang, H. and Emmons, S. W. (2000). "A *C. elegans* mediator protein confers regulatory selectivity on lineage-specific expression of a transcription factor gene." Genes & Development **14**(17): 2161-2172.

Zhang, L., Ding, L., Cheung, T. H., Dong, M. Q., Chen, J., Sewell, A. K., Liu, X., Yates, J. R. 3rd and Han, M. (2007). "Systematic Identification of *C. elegans* miRISC Proteins, miRNAs, and mRNA Targets by Their Interactions with GW182 Proteins AIN-1 and AIN-2." Molecular Cell **28**(4): 598-613.

Zhang, Y., Ma, C. Delohery, T., Nasipak, B., Foat, B. C., Bounoutas, A., Bussemaker, H. J., Kim, S.K., and Chalfie, M. (2002). "Identification of genes expressed in *C. elegans* touch receptor neurons." Nature **418**(6895): 331-335.

Zhao, G. Y., Schriefer, L. A. and Stormo, G. D. (2007). "Identification of muscle-specific regulatory modules in *Caenorhabditis elegans*." Genome Research **17**(3): 348-357.