IDENTIFICATION OF TWO GENETIC ENHANCERS OF CELL MIGRATION DEFECTS AND THEIR ROLES IN DISTAL TIP CELL GUIDANCE IN *C.ELEGANS*.

BY

MEGAN SCHWABIUK

A Thesis submitted to

The Faculty of Graduate Studies

in Partial Fulfillment of the Requirements for the Degree of:

MASTER OF SCIENCE

Department of Biochemistry and Medical Genetics

University of Manitoba

Winnipeg, Manitoba, Canada

October 2006

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

IDENTIFICATION OF TWO GENETIC ENHANCERS OF CELL MIGRATION DEFECTS AND THEIR ROLES IN DISTAL TIP CELL GUIDANCE IN *C.ELEGANS.*

BY

Megan Schwabiuk

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

OF

MASTER OF SCIENCE

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Ant.	Anterior
ARP2/3	Actin related protein
β	beta
BLAST	Basic local alignment search tool
bp	base pair
Ca^{2+}	Calcium
cAMP	cyclic adenosine monophosphate
$CHCl_3$	Chloroform
C.elegans	*Caenorhabditis Elegans*
CGC	*Caenorhabditis Elegans* Genetics Centre.
cGMP	Cyclic guanosine monophosphate
cm	centimetre
^0C	degrees Centigrade
DCC	Deleted in Colorectal Cancer
ddH_20	Double distilled water
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytosine triphosphate
dGTP	Deoxyguanosine triphosphate
DIC	Differential interference contrast
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
dsRNA	Double stranded ribonucleic acid
dTTP	Deoxythymidine triphosphate
Dpy	Dumpy phenotype, reduced body length
DTC	Distal tip cell
EBI	European Bioinformatics Institute
E.Coli	*Escherichia coli*
ECM	Extracellular matrix
EDTA	Ethylene diaminetetraacetic acid
EGF	Epidermal growth factor
Egl	Egg-laying defect phenotype.
EMS	Ethyl methane sulfonate
EtBr	Ethidium bromide
EtOH	Ethanol
FAK	Focal adhesion kinase
FGF	Fibroblast growth factor
FGF-R	Fibroblast growth factor receptor
F1	First generation progeny from one mating.
g	gram
g	Earth's gravitational constant
GFP	Green fluorescent protein
HA	Hemagglutin
HSPGs	Heparan sulphate proteoglycans

IDT	Intergrated DNA Technologies
Ig	Immunoglobulin
IPTG	isopropylthiogalactoside
Kb	Kilobase
LB	Luria Bertani
L1	Larval stage one.
L2	Larval stage two.
L3	Larval stage three.
L4	Larval stage four.
LGX	Linkage group X
M	Molar
µg	microgram
mins	minutes
µl	microlitre
ml	millilitre
mM	millimolar
mm	millimetre
NCBI	National Centre for Biotechnology Information
NDPase	nucleoside diphosphatase
ng	nanogram
NGM	Nematode growth medium
nM	nanomolar
N2	Wild-type *C.elegans* strain.
PCR	Polymerase chain reaction
PKA	Cyclic AMP-dependent protein kinase
pM	picomolar
PM	Plasma membrane
Post.	Posterior
RNA	Ribonucleic acid
ROBO	Roundabout
Rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SH2	Src homology 2
SH3	Src homology 3
SNP	Single nucleotide polymorphism
TAE	Tris-Acetic Acid EDTA buffer
Taq	*Thermus aquaticus*
TGFβ	Transforming growth factor beta
Tyr	Tyrosine
Unc	uncoordinated, locomotion defect phenotype.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Dr. David Merz for providing me with the opportunity to learn and practice science in his laboratory. Your guidance, advice and beliefs in the value of academic freedom have made this an incredibly rewarding learning experience for me. My committee members Dr. Dana Schroeder and Dr. Jeff Wigle, thank-you for your time, guidance and interest in this project.

Thank-you Merz lab members, Allison for sharing our graduate student experiences together, Lu for the refreshing conversations and advice and Cheryl; your help, counselling and companionship throughout this entire project was invaluable and will always be greatly appreciated.

For Rick, Dr. Pind, Dr. Allison Elliot. The guidance, scientific knowledge and expertise you have offered me is greatly appreciated and thank-you for always taking a moment from your busy schedules to listen and share your opinions. Thank-you Dr. Barbara Triggs-Raine for the counselling and critique you took the time to offer. Our talks were essential to leading me in the right direction for my future endeavours.

Dianna, Scarlet and Shauna. I am very grateful to have become acquainted with three remarkable women in science and I value the friendships we have made throughout the past years.

Thank you to all of my friends, whom I treasure for your continual encouragement. To my families, your never ending patience, support and faith in me have meant so much. Mom and Dad, you taught me that the world is an open door and I can achieve anything if I put my mind to it, so I did and Dev, your compassion, kindness and most of all your laughter brightens up my days and my life.

And Rob, merci pour ton amour, ton sourire et un rapel continuel de la joie de vivre.

Megan.

ABSTRACT

Cell motility is critical for many biological processes such as cellular re-organizations during organogenesis and disease states such as tumour metastasis. The secreted netrin guidance molecule and its receptors play roles in directing cell and growth cone migrations during axon patterning, lung and vascular development. Despite the significance of this cell guidance pathway, there is still much to learn about the signalling and regulatory mechanisms of netrin receptors. In *C.elegans*, ventrally expressed UNC-6/Netrin and netrin receptors UNC-5 and UNC-40 guide the migration of growth cones and distal tip cells (DTCs) along the ventral/dorsal axis. Migrating DTCs cap each tip of the gonad arm and direct the formation of the C-shaped bi-lobed gonad. Mutations in *unc-6, unc-5* or *unc-40* genes disrupt the ventral to dorsal DTC migration phase resulting in visibly misshapen gonads providing an ideal model for studying UNC-6/UNC-40/UNC-5 cell guidance mechanisms *in vivo*.

A genetic screen for enhancers of DTC migration defects generated 28 mutants. Two of these mutants, defined by enhancer alleles *mig(ev648)* and *enh(ev697)* mapped to chromosomal regions containing no known DTC migration genes. We have identified the *mig(ev648)* enhancer as an allele of the gene *mig-23* that encodes a nucleoside diphosphatase and have shown the role of *mig-23(ev648)* in DTC guidance is not limited to UNC-6/UNC-5/UNC-40 mediated guidance. *enh(ev697)* has been identified as an allele of *sdn-1* (syndecan), a trans-membrane heparan sulfate proteoglycan. The *ev697* allele encodes a truncated form of SDN-1 and *sdn-1(ev697)* guidance mechanisms appear to only affect *unc-5* mediated DTC guidance. *sdn-1* is functions cell non-autonomously and appears to be

involved in limiting growth factor molecules EGL-17/FGF, UNC-129/TGF-β, DBL-1/TGF-β EGL-20/WNT and LIN-3/EGF within the extra-cellular environment for DTC guidance.

1. INTRODUCTION

1.1 Cell guidance and biological processes.

Directional guidance of motile cells is required in diverse biological and developmental processes. For example, chemokines produced during an inflammatory response attract monocytes/macrophages from the blood stream towards the site of tissue damage (Ridley 2001). During oogenesis in *Drosophila*, migration of border cells along a specified path from the most anterior region of the oocyte to the midline is required for oocyte fertility and embryonic patterning (Montell 2003). Wiring the human nervous system involves guidance of growth cone-tipped axonal projections from neurons to specified distant targets for the formation of neurological synapses. Cells can also acquire aberrant migratory abilities resulting in disease states such as tumour metastasis and invasion. By modulating cellular interactions with the extra-cellular micro-environment to favour aberrant migration/invasion and initiating intra-cellular pathways for cellular motility, tumour cells gain the ability to become motile and invade through tissues ultimately attaining the blood stream or lymphatic system (Hanahan and Weinberg 2000). Despite the significance of guided cellular migrations there is still much to learn about the temporal and spatial regulation of cell motility and guidance.

1.2 Mechanisms of cellular migration.

In order to become motile cells execute several key processes. These include breaking initial contacts with the ECM(extra-cellular matrix)/neighbouring cells and extending portions of plasma membrane, remodelling the ECM in its pathway, initiating new contacts with the ECM/neighbouring cells and retracting the trailing edge of the cell. A

moving cell (broadly speaking) can be segregated into two parts, a leading edge and a retracting edge (Lauffenburger and Horwitz 1996). At the leading edge, actin filament polymerization (Pollard and Borisy 2003) via activation of ARP2/3s(Actin Related Proteins) forces the plasma membrane (PM) to protrude in the form of lamellipodia and filopodia. Activation of ARP2/3 is indirectly mediated by RhoGTPases Rac and Cdc42. RhoGTPases Rac and Cdc42(for protrusions) and Rho(for retraction) play a major role in regulating the intra-cellular signalling pathways involved in the regulation of actin dynamics (Raftopoulou and Hall 2004). The leading edge of a migrating cell favours the formation of focal adhesion integrin clusters, arbitrated by RhoGTPases, that mediate adhesion of cell membrane protrusions by linking actin filaments within the cell to the ECM (Lauffenburger and Horwitz 1996). Integrin adhesion also triggers signalling pathways inducing actin polymerization regulated by cAMP/PKA (cyclic AMP-dependent protein kinase) signalling pathways (Howe 2004). Within the retracting edge, actin polymerization and the formation of focal adhesions is attenuated and cross-linking of myosin light chains to actin filaments for contraction is mediated by Rho, ultimately causing retraction of the PM and breakage of focal adhesion contacts from the ECM (Lauffenburger and Horwitz 1996). Increased levels of Ca^{2+} within the trailing edge have been associated with the activation of contractile pathways and detachment of adhesion structures (Lee *et al.* 1999). In addition to the intracellular dynamics, proteases for ECM remodelling and degradation, such as the matrix metalloproteases (Vu and Werb 2000), are required on the cell surface in order to create a pathway within the ECM through which the cell can move. Conferring migratory direction to a motile cell requires spatial and temporal activation and co-ordination of each

aforementioned process mediated via ligand/receptor interactions for activation of intra-cellular pathways.

1.3 Cell guidance molecules

Guidance cues work through their respective receptors to trigger and co-ordinate intra-cellular events promoting cellular motility and guidance of a cell towards or away from their source. Guidance molecules are either bound to extra-cellular substrates (contact repulsion or attraction) or disseminated usually as a gradient throughout the extra-cellular environment of the migrating cell (chemoattraction/chemorepulsion) (Tessier-Lavigne and Goodman 1996). At any given moment, a cell's migration pathway can be influenced by a number of guidance cues acting on the cell at once. The numerous cues/receptor interactions regulating cell motility processes, coupled with the finding that most cell guidance pathways are redundant, renders the study of cellular guidance challenging. Conserved families of guidance molecules and their respective receptors Slit/ROBO (Roundabout), semaphorins/Plexin and Neuropilin, ephrins/Ephs and netrin/Unc5/DCC (Deleted in Colorectal Cancer) (Dickson 2002) have been identified. However, mechanisms linking guidance receptors to the activation and co-ordination of intracellular cell motility pathways are not well defined. In addition, regulation of the spatial distribution of guidance cues used by these receptors within the extra-cellular environment and the integration and co-ordination of various receptor responses within the cell required to confer direction to a migrating cell have yet to be fully elucidated.

1.4 Netrin/UNC-6 and receptors UNC-40 and UNC-5

Netrins and their receptors make up one of the conserved cell guidance systems. Orthologues of netrins and netrin receptors have been identified in *Drosophila, C.elegans*, and vertebrates (outlined in Table 1) exhibit strong sequence similarities and their roles in mediating cell and axon guidance along the ventral-dorsal axis are highly conserved.

Table 1: Orthologues of *C.elegans* UNC-6, UNC-5 and UNC-40s.

C.elegans	*Drosophila*	Vertebrate
UNC-6	NetA, NetB	Netrin1-4, Netrin-G
UNC-5	Unc-5	UNC5A-UNC5D
UNC-40	Frazzled	DCC/Neogenin

Adapted from (Barallobre *et al.* 2005)

Netrin ligands are secreted into the extra-cellular space and act as both a chemoattractant (for cells expressing trans-membrane receptor DCC/UNC-40 on their surface) and as a chemorepellant (for cells expressing trans-membrane receptors UNC5 and DCC/UNC-40 on their surface). Both netrin receptors are members of the immunoglobulin superfamily (Leung-Hagesteijn *et al.* 1992, Chan *et al.* 1996). Roles of proteins in the immunoglobulin superfamily include mediating adhesive interactions between other immunoglobulin family proteins, integrin recognition and binding to the ECM (Brümmendorf, Rathjen 1996). Although each receptor has putative intracellular protein binding domains, it is still a mystery how these receptors work to mediate a response to netrins for cellular guidance. Adding to the complexity of this guidance system, netrins play the dual role of mediating cellular repulsion or attraction depending on the netrin receptor

expressed on the surface of the migrating cell. For the context of this project a review of the structural and functional properties of netrins and netrin receptors will focus on UNC-6, UNC-40 and UNC-5 in *C.elegans*.

The secreted, laminin-related netrin ligand was initially identified in *C.elegans* as an axon guidance molecule and termed *unc-6* (Brenner 1974; Hedgecock, Culotti and Hall 1990). *C.elegans unc-6* mutants are Unc (**unc**oordinated, locomotion defects) due to axon guidance defects, exhibit slight egg-laying defects and mesodermal DTC migration defects (Hedgecock, Culotti and Hall 1990). *unc-6* expression in motorneurons within the ventral nerve cord during the first larval molt into adult stages is required for the migration of pioneer axons and DTCs along the ventral/dorsal axis (Wadsworth, Bhatt and Hedgecock 1996). The regulation of the UNC-6 diffusion pattern from these cells, which results in a ventralized restriction of UNC-6, is still uncharacterized. The UNC-6 ligand consists of conserved domains VI, V-1,V-2, V-3 and C (Figure 1) (Ishii *et al.* 1992). Each V domain exhibits similarities to laminin subunits while the C domain is not similar to laminins. *unc-6* plays a dual guidance role by interacting with *unc-40* for ventral guidance of axons (Chan *et al.* 1996) and with both *unc-40* and *unc-5* for ventral to dorsal guidance of axons (Colavita *et al.* 1998) and mesodermal cells (Merz *et al.* 2001). Functional analysis of UNC-6 domains has demonstrated both V-2 and V-3 domains are involved in dorsal axon and cell guidance while V-3 is required for ventral cell guidance, suggesting the V-2 domain possibly interacts with UNC-5 to mediate a repulsive response to UNC-6 (Lim and Wadsworth 2002). Supporting this observation, the *rh202* allele of *unc-6* encoding a V-2 domain deletion behaves like an *unc-5* loss of function mutation during DTC guidance (Merz *et al.* 2001).

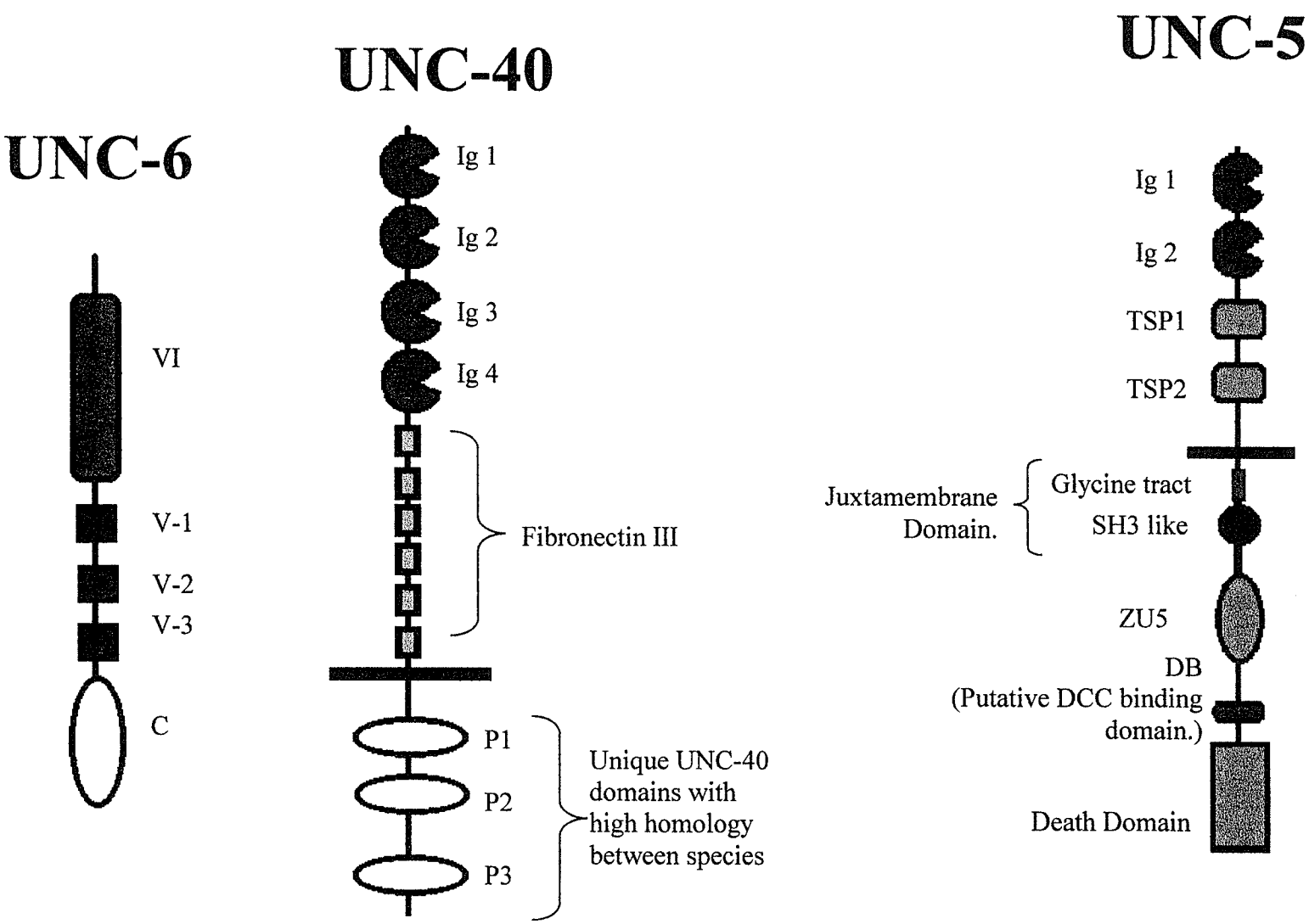


Figure 1: Schematic of UNC-6, UNC-40 and UNC-5 protein domains.

The C domain is not required for guidance but rather has a role in inhibiting axon branching and deletion of V-3 and C domains abolishes UNC-6 guidance functions (Wadsworth, Bhatt and Hedgecock 1996). The β subunit motif within the VI domain is required for all UNC-6 guidance functions, possibly via interactions with the ECM (Lim and Wadsworth 2002). Although the roles of each UNC-6 domain have been characterized by genetic interaction analysis, direct *in vivo* interactions of UNC-40, UNC-5 and possibly additional molecules with the UNC-6 ligand have yet to be elucidated.

Deleted in Colorectal Cancer (DCC), the UNC-40 vertebrate homologue, was identified as a tumour suppressor gene inactivated in colorectal cancer (Hedrick *et al.* 1994) prior to being recognized as a netrin receptor in *C.elegans* (Hedgecock, Culotti and Hall 1990; Chan *et al.* 1996) and rodents (Keino-Masu *et al.* 1996). *C.elegans unc-40* mutants are uncoordinated and exhibit a slight reduction in body length in addition to defects in DTC migration patterns. UNC-40 functions cell-autonomously in migrating DTCs and axons to guide cells along a ventral source of UNC-6 (Chan *et al.* 1996). In addition, UNC-40 can mediate repulsion from UNC-6 independently or in combination with the UNC-5 receptor in *C.elegans* DTCs (Merz *et al.* 2001). The ectodomain of *C.elegans* UNC-40 trans-membrane receptor consists of four V-like immunoglobulin domains and four fibronectin type 3 domains (Figure 1). *In vitro* pull-down assays have demonstrated vertebrate DCC binds directly to a recombinant form of Netrin-1 via its 5th fibronectin type III domain (Geisbrecht *et al.* 2003), the same DCC domain that is found to interact with heparan sulfate (Bennett *et al.* 1997). However, binding of the recombinant Netrin-1 appears to abrogate DCC and heparan sulfate interactions (Geisbrecht *et al.* 2003). UNC-40 interactions with heparan sulfate in *C.elegans* have not been characterized to date. The UNC-40 cytodomain consists of P1, P2 and P3 domains with little resemblance to other protein domains but high sequence

similarities between UNC-40 homologues. *In vitro* experiments suggest that the DCC P1 domain is required for DCC interactions with the UNC5 receptor in response to Netrin-1 (Hong *et al.* 1999). Recently, FAK(focal adhesion kinase), a tyrosine kinase with a role in mediating downstream signalling of integrins, was found to act downstream of DCC in a Netrin-1 dependent matter (Li *et al.* 2004). Additional *in vitro* assays have shown DCC mediates neurite outgrowth in response to Netrin-1 by associating directly with the SH3 domains of the RhoGTPase activator Nck adaptor protein (Li *et al.* 2002a) and via indirect activation of Rac1 and Cdc42 RhoGTPases (Li *et al.* 2002b), suggesting a direct link between the DCC receptor and control of actin polymerization. *In vivo* assays in *C.elegans* have confirmed that *unc-40* genetically interacts with *ced-10*/Rac in one pathway (possibly through the P2 domain) and *unc-34*/Enabled (possibly through the P1 domain) in a parallel pathway to mediate axon attraction to UNC-6 (Gitai *et al.* 2003). An additional, UNC-6 independent guidance role of the UNC-40 receptor in *C.elegans* involves direct interactions with the ROBO receptor in mediating cellular repulsion from dorsal *slt-1* expression for ventral axon and neuron guidance (Yu *et al.* 2002). However, UNC-40/ROBO interactions have not been defined for DTC guidance.

The UNC-5 trans membrane receptor was also initially identified in *C.elegans* (Brenner 1974; Hedgecock, Culotti and Hall 1990a). *unc-5* mutants are uncoordinated and exhibit DTC migration defects similar to *unc-40* and *unc-6* mutants. In *C.elegans*, ectopic expression of *unc-5* in migrating touch neurons turns them away from their ventral migration pathway along an UNC-6 source (Hamelin *et al.* 1993) indicating that UNC-5 mediates axons guidance cell-autonomously away from an UNC-6 . This guidance role of UNC-5 has also been demonstrated in the DTCs (Su *et al.* 2000). UNC-5 mediates DTC migration away from UNC-6 independently or in concert with UNC-40, although repulsion from UNC-6 is

most efficient when both receptors are present on the surface of the migrating cell (Merz *et al.* 2001). The UNC-5 ectodomain domain consists of two immunoglobulin domains followed by two thrombospondin type 1 domains (Leung-Hagesteijn *et al.* 1992) (Figure 1). *In vitro* assays using a recombinant form of Netrin-1 have shown Netrin-1 interacts with both Ig domains of UNC5 (Geisbrecht *et al.* 2003), however the direct interactions between *C.elegans* UNC-6 and UNC-5 have yet to be characterized. The UNC-5 trans-membrane domain is followed by a short glycine tract (possibly acting as a hinge adjusting UNC-5 conformation upon ligand or receptor binding or possibly creating conditions for homo-dimerization) and a sequence weakly resembling an SH3 domain that is characteristic of intra-cellular signalling molecules involved in regulating actin, suggesting UNC-5 may have a role in directly regulating cell motility (Leung-Hagesteijn *et al.* 1992). The cytodomain of the UNC-5 receptor between the trans-membrane domain and the ZU5 domain, encompassing the previously mentioned regions has been named the juxtamembrane domain (Figure 1). The juxtamembrane domain is required for UNC-40 dependent UNC-5 mediated repulsion and harbours a tyrosine phosphorylation site required for UNC-5 mediated guidance in DTCs and motorneurons *in vivo* (Killeen *et al.* 2002). Genetic interaction analysis in *C.elegans* has confirmed these results as the *e152* allele of *unc-5*, encoding an UNC-5 protein with a deletion after the sixth amino acid of the ZU5 domain, retains the UNC-40 dependent functions of UNC-5 (Merz *et al.* 2001). In the ZU5 domain, *in vivo* assays have shown that an HA(hemagglutin) tagged UNC-5 in *C.elegans* is phosphorylated on Tyr^{568} upon UNC-6 stimulation and co-immunoprecipitates with the tyrosine phosphatase Shp2 (Tong *et al.* 2001), which has been associated with regulating cell motility processes including cell spreading and focal adhesion turnover (Yu *et al.* 1998). Interactions between UNC-5 and SRC-1 kinases (which associate with FAKs) have been identified although the

phosphorylated tyrosine residues within the UNC-5 cytodomain that mediate SRC-1/UNC-5 interactions are not defined (Lee, Li and Guan 2005). In addition, SRC-1 was found to be essential for UNC-5 mediated DTC guidance, as animals able to escape a lethal phenotype caused by *src-1* RNAi knockdown exhibit DTC migration defects (Lee, Li and Guan 2005). The UNC-5 C terminal death domain, a domain characteristic of proteins involved in cell death signalling and innate immune responses, is dispensable for UNC-5 mediated guidance. Interestingly, the UNC-5 protein sequence does not code for a signal sequence (Leung-Hagesteijn *et al.* 1992).

Several genetic screens conducted in *C.elegans* have identified genes that interact genetically with *unc-5*. A screen for suppressors of aberrant growth cone steering caused by ectopic expression of *unc-5* identified axon guidance genes *unc-44*/ankyrin and *unc-34*/Enabled, novel genes *seu-1*(suppressors of ectopic unc-5), *seu-2, seu-3* and *unc-129* (TGFβ (transforming growth factor)) as suppressors of aberrant growth cone migration, suggesting these genes interact in an *unc-5* mediated guidance mechanism to steer axons dorsally (Colavita and Culotti 1998). *unc-129*(TGFβ) has been implicated in *unc-5* mediated DTC migrations (Merz *et al.* 2003) and ectopic expression of *unc-129* in ventral muscle band in addition to the dorsal muscle band causes the aberrant DTC migration patterns observed in *unc-40*, *unc-5* and *unc-6* mutants (Colavita *et al.* 1998). Thus a dorsal/ventral distribution pattern of *unc-129* appears to be required for dorsal/ventral guidance functions of the DTC. The aforementioned genetic assays provide evidence for interactions between these genes and the UNC-5 receptor both in axon and DTC guidance, however their roles in the guidance mechanisms of the UNC-5 receptor have yet to be defined.

The study of cellular guidance mediated by netrins and their receptors is complex due to the receptors abilities to either work together or separately to instruct a cell to migrate

away from or towards an UNC-6 source. *In vitro* analysis of UNC5 and DCC receptor interactions in *Xenopus* cultured spinal neurons demonstrated that UNC5 expressing cells were repelled from a Netrin-1 source and that the intracellular domains of both DCC and UNC5 interact directly to mediate Netrin-1 repulsion, as receptors co-immunoprecipitate together via their cytodomains (Hong *et al.* 1999). *In vivo* assays in *C.elegans* support the direct interactions between each UNC-6 receptor (Merz *et al.* 2001). In addition, UNC-40 and UNC-5 both appear to signal via *unc-34*/Enabled in *C.elegans* axons (Colavita and Culotti 1998, Gitai *et al.* 2003). Another shared signalling mechanism of UNC5 and DCC is the ability of DCC to associate with FAK and both DCC and UNC5 to bind Src-1 (Li *et al.* 2006). These findings suggest FAK/Src may be the key players in initiating the switch in cell motility towards or away from a Netrin-1 source. However, UNC-40 and FAK associations have yet to be identified in *C.elegans*. Another event associated with the directional guidance switch of netrin receptors is an apparent variation in intra-cellular Ca^{2+} concentrations. UNC5 binding to DCC has been shown to modulate cyclic nucleotide signalling pathways. *Xenopus* spinal neuron growth cones that are attracted to a Netrin-1 source exhibit increased intracellular levels of Ca^{2+} the via activation of cAMP signalling pathways. and cells expressing UNC5, in response to a Netrin-1 source, exhibit reduced Ca^{2+} gradients by the activation of cGMP pathways (Nishiyama *et al.* 2003). It is still unclear how netrin receptors are directly involved in regulating intracellular levels of Ca^{2+}, what other proteins are involved and whether the same modulation of intra-cellular calcium levels occurs *in vivo*.

1.5 The UNC5 receptor and human disease.

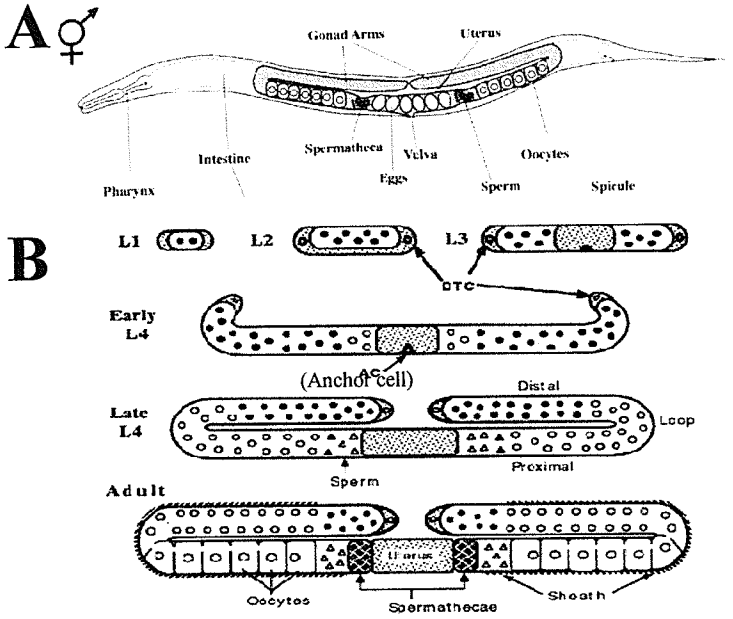
Netrins and the UNC5 receptor were initially identified and studied for their role in axon guidance in the model organism *C.elegans.* Mice homozygous for mutations in *Unc5c* were ataxic and exhibited small cerebella and abnormal migration of granule cells and Purkinje cell precursors (Ackerman *et al.* 1997; Przyborski, Knowles and Ackerman 1998), indicating that the UNC5C receptor is also a key player in neuronal development of higher organisms. Expression of UNC5A-D was detected in various tissues other than those of the nervous system and vascular system, such as testis, ovaries, thymus, spleen, bladder and kidney tissues (Ackerman *et al.* 1997), suggesting that the UNC5 receptor may have additional roles in biological processes other than axon guidance during development. Expression of *Unc5b* was observed in the endothelial tip cells located at the leading edge of developing vessels in the vascular system (Lu *et al.* 2004) and *Unc5b* mutant mice exhibited abnormal extension of filopodia and navigation defects in endothelial tip cells. *Unc5b* mutants died at day 12.5 due to heart failure demonstrating an important role for UNC5B in the morphogenesis of the vascular system. In addition to various roles in development, recent evidence suggests UNC5 is functioning as a tumour suppressor since UNC5A-C expression was down regulated in a number of cancers (Thiebault *et al.* 2003). UNC5 has been termed a dependence receptor with the ability to initiate apoptosis and prevent tumour cell extension, invasion and/or metastasis upon exiting an environment containing the Netrin-1 ligand. *Unc5b* is a target of the tumour suppressor p53 transcription factor (Tanikawa *et al.* 2003) and is implicated in mediating the ability of p53 to suppress tumour cell transformation processes including growth and invasion. The recurring role of the UNC-5 receptor in the guidance and regulation of cell migrations is evidently conserved among various biological processes. Using *C.elegans* as a model for studying the mechanisms of UNC-5 in cell

guidance enables a simplified compilation of data and results that can later be examined in higher organisms.

1.6 *C.elegans* distal tip cells and their role in patterning the gonad morphology.

The *C.elegans* hermaphrodite gonad is a tubular structure in the form of two mirror image, C shaped arms (Figure 2). The gonad ultra-structure has been reviewed by Hall *et al.* (1999). Five gonadal sheath cell pairs shape each anterior and posterior tubular arm. In the most distal region of each gonad arm, germ line nuclei proliferation occurs within a syncytium and is regulated by the DTC (Kimble and White 1981). The germ cells mature into oocytes as they migrate along the gonad arm bend and then pass through the spermatheca were fertilization occurs. Fertilized oocytes move into the common uterus shared by each gonad arm. Here, embryos begin to mature and are then laid through the vulva. The structure, function and formation of the gonad in *C.elegans* males differs from that of the hermaphrodite gonad and will not be addressed here.

The gonad morphology of a hermaphrodite is patterned by the mesodermal DTCs, one of the largest cell types in *C.elegans*. DTC have an asymmetric shape and cap the distal end of the forming gonad. Two DTCs are born at L1 (larval stage 1) from Z1 and Z4 precursors cells located centrally along the anterior/posterior axis and each begins their migration pattern towards opposite ends of the animal initiating the first of three migration phases (Figure 2). Interestingly evidence suggests that the posterior migration pattern of the DTCs is regulated differently than the anterior migration pattern (Nishiwaki 1999). Defects in DTC migration patterns result in aberrant gonad morphology phenotypes visible at low magnification, simplifying the identification of defective cell migration patterns throughout large numbers of animals.



C.elegans II, 1997.

Figure 2: The *C.elegans* hermaphrodite gonad morphology and the DTC migration pattern. A) A schematic depicting hermaphrodite gonad morphology and associated structures. B) Outline of the DTC migration pattern. In L1, DTCs arise from Z1 and Z4 precursors and migrate along the ventral muscle band until early L4, where the DTCs turn dorsally and migrate towards the dorsal muscle band. During L4, DTCs return towards the central anterior/posterior axis along the dorsal muscle band forming two mirror image, C shaped gonad arms.

The DTC migration pattern consists of three distinct migration phases each occurring along a different substrate. Migration phase I, initiated in early L2 larval stage occurs, longitudinally and centrifugally on the ventral side of the animal along the basement membrane of ventral muscle band. In *gon-1* (a secreted metalloprotease) mutants the DTCs fail to extend away the gonadal primordium resulting in a failure to initiate the first DTC migration phase (Blelloch *et al.* 1999). Migration phase II is initiated during the late L3 stage, when the DTC, now located at the furthest anterior/posterior end, at a precise time makes a 90 degree turn and begins migrating towards the dorsal side of the animal along the basement membrane of the hypodermis. Initiation of this migration phase is largely dependent on the up-regulation of *unc-5* expression by *daf-12* (nuclear hormone receptor) (Su *et al.* 2000a). Other genes involved in the DTC migration phase II include TGFβ growth factors, *dbl-1* and *unc-129* (apparently in an *unc-5* dependent mechanism) (Merz *et al.* 2003), *src-1* as *src-1* RNAi knockdown causes a failure in the DTC to initiate the second migration phase and the third migration phase resulting in a straight gonad arm that fails to reflux back toward the midline (Lee, Li and Guan 2005) and *clr-1*. CLR-1 (CLeaR-1) is a receptor tyrosine phosphatase implicated in limiting UNC-40 mediated attraction in AMV axons to an UNC-6 source thus negatively regulating netrin attraction in axons (Chang *et al.* 2004). Supporting this theory, *clr-1* positively regulates DTCs migration away from an UNC-6 source as *clr-1* enhances the DTC migration defects in *unc-5(e152)* and *unc-5(e53)* mutants indicating that *clr-1* functions in an UNC-5 parallel guidance pathway for DTC guidance during the second migration phase (Merz *et al.* 2003). However it is still unclear exactly how *clr-1* is working within the UNC-40/UNC-6 guidance pathway in DTCs.

Once the DTC reaches the dorsal muscle band it makes another 90 degree turn back towards the centre of the anterior/posterior axis and migrates centripetally and longitudinally

along the basement membrane of dorsal muscle band to complete the DTC migration phase III. The *ced-5* (Wu and Horvitz 1998) *ced-2* and *ced-10* (Reddien and Horvitz 2000) genes encode *C.elegans* homologues of CrkII, DOCK180(associates with the Crk adaptor involved in integrin signalling to the actin cytoskeleton) and RacGTPase respectively. Gonad morphology defects in *ced-2*, *ced-5* and *ced-10* mutants illustrate that the DTCs in these mutants stop prematurely along the ventral muscle band during migration phase I and make extra turns, suggesting a role for cytoskeletal regulatory elements in DTC pathfinding. MIG-17, a disintegrin and matrix metalloprotease is also involved in DTC pathfinding as DTCs in *mig-17* mutants either do not execute the second ventral to dorsal migration phase or do so and are unable to migrate in a straight line along the dorsal muscle band (Nishiwaki, Hisamoto and Matsumoto 2000).

Mechanisms regulating the spatial and temporal guidance of DTC migration are still largely unknown and appear to be somewhat distinctive due to a lack of hallmark cell migration structures such as filopodia, lamellipodia and pseudopodia at the leading edge of the DTC. It has been suggested that DTCs are not propelled by an inner source but rather are pushed by developing gonadal sheath cells and guide the extension of the gonad tubular arms by changing the substrate over which the DTC migrates, from the ventral muscle band to the hypodermis to finally the dorsal muscle band. However these speculations have not been proven.

1.7 UNC-6/UNC-5/UNC-40 and their roles in DTC guidance.

Although *unc-6*, *unc-40* and *unc-5* are involved in guiding several cell types in *C.elegans*, their role in initiating and maintaining the second ventral to dorsal DTC migration phase was used as the basis for our genetic screen to identify novel DTC guidance genes.

Mutations in *unc-6, unc-5* or *unc-40* genes disrupt the ventral to dorsal migration phase causing the DTC to reflux and complete the third migratory phase along the ventral muscle band. As the DTCs pattern the gonad morphology, mis-positioning of the gonad arm pushes the intestine dorsally resulting in visibly misshapen gonads (Figure 3), thus providing an ideal model for studying UNC-6/UNC-40/UNC-5 cell guidance.

In *C. elegans*, UNC-6 is secreted along the ventral midline. During the first ventral migration phase of DTC migration along the UNC-6 source, *unc-40* is expressed at a constant level in the DTC (Chan *et al.* 1996). Interestingly, *unc-40* mutants exhibit no defects in DTC migration along the ventral muscle band (Chan *et al.* 1996), suggesting this migration phase is regulated by another UNC-40 independent DTC guidance/adhesion mechanism. Reporter constructs have not been able to demonstrate that *unc-5* expression in the DTC during migration phase I, however genetic interaction assays suggest low levels of *unc-5* expression occur during this migration phase (Su *et al.* 2000). At a precise time during the late L3 larval stage, as the DTC is migrating along the ventral muscle band, expression of *unc-5* in the DTC is up-regulated by DAF-12 and DAF-9 resulting in DTC migration away from UNC-6 towards the dorsal muscle band (Su *et al.* 2000). Taken together, these data suggest *unc-5* expression is somehow down regulated or inhibited prior to the DTC turning time and at the turning time, the inhibitor is removed and expression is up-regulated by DAF-12. Up-regulation of *unc-5* expression plays a key role in initiating the second DTC migratory phase. The initiation of DTC turning is determined intrinsically by the DTC rather than the surrounding environment. For example, DTCs migrate the same distance along the ventral muscle band in *lon-2* mutants with elongated body length as they do in wild-type animals as it is only the environment at DTC turning onset has been changed in *lon-2* mutants and not the instrinsic DTC turning program (Su *et al.* 2000). As previously

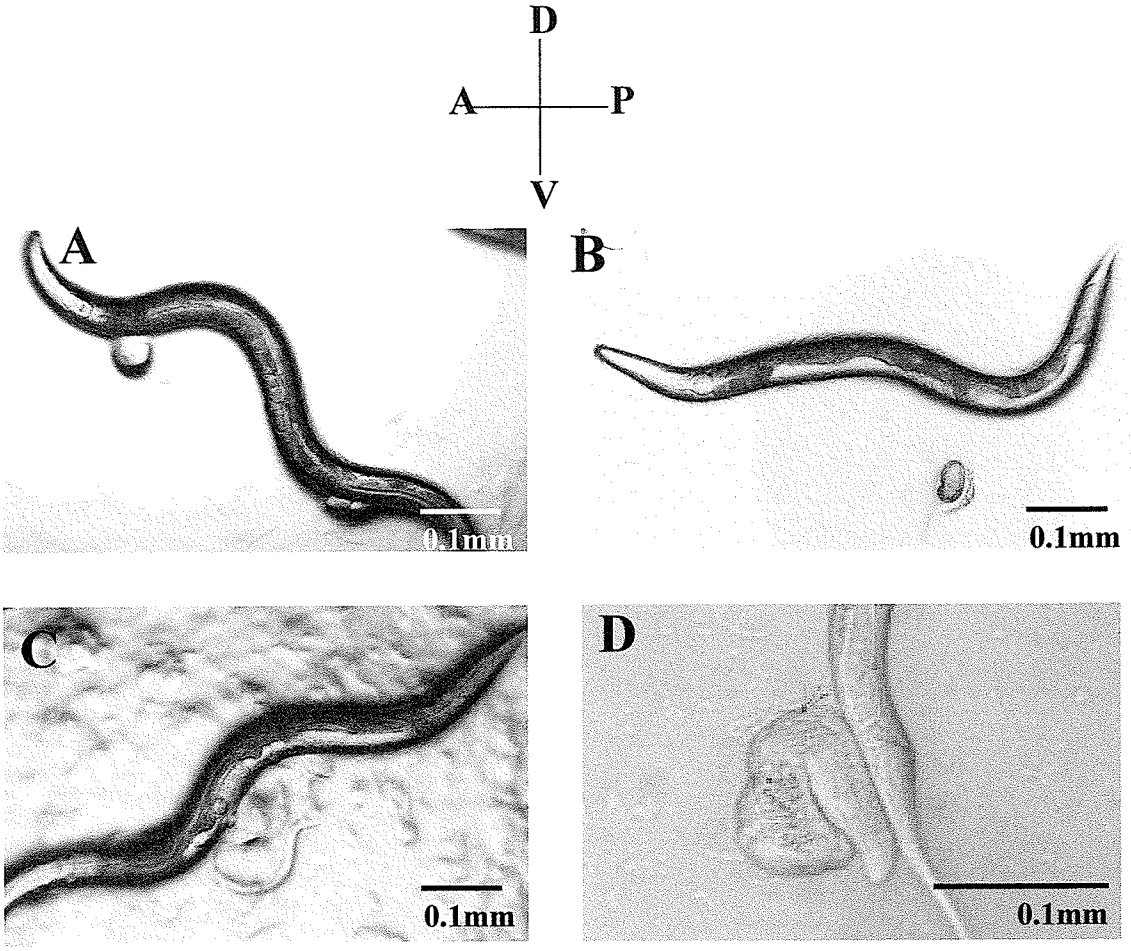


Figure 3 : Photographs of N2 wild-type, *unc-5*, *mig(ev648)* and *enh(ev697)* mutants. Bright field live images taken at 10X magnification for A) wild-type, B) *unc-5(e152),* C) *mig(ev648)* and 20X magnification for D) *enh(ev697).*

mentioned, this migration phase is most efficiently executed when both UNC-5 and UNC-40 are present in the DTC. However each receptor can mediate this migration phase independently (Merz *et al* 2003). Genetic interactions analysis suggests *unc-129*(TGFβ) and *dbl-1*TGFβ are limited to an *unc-5* mediated guidance mechanism during the ventral to dorsal migration phase. However, other than the role of UNC-52/perlecan in limiting *unc-129* and *dbl-1* distribution with the extracellular environment for the *unc-5* receptor, the exact role of *unc-129* and *dbl-1* in guiding the DTCs along the dorsal/ventral axis is not known (Merz *et al.* 2003).

Defects in the second DTC migration pattern of *unc-5* null mutants are not fully penetrant suggesting a parallel *unc-5* independent DTC guidance pathway is involved in guiding the DTC dorsally, compensating for the residual DTC guidance observed in mutants without a functional UNC-5 (a null mutant). One such pathway may involve the *egl*-17/FGF (fibroblast growth factor), as it appears to be functioning in an *unc-5* parallel guidance pathway for DTC guidance (Merz *et al.* 2003). Interestingly, the activity of *C.elegans* FGF-R(fibroblast growth factor receptor) EGL-15 is attenuated by CLR-1 (Kokel *et al.* 1998) and CLR-1 has been shown to positively regulate DTC migration away from UNC-6 (Merz *et al.* 2003). Thus CLR-1, FGF and the FGF-R possibly represent an additional, *unc-5* independent dorsal/ventral DTC guidance pathway.

It should be noted that the UNC-6/UNC-40/UNC-5 guidance system is not the sole guidance system used by the DTC (although they do appear to be the dominant guidance mechanism) as DTC migration defects in the second DTC migration phase are not fully penetrant in *unc-6* null mutants representing a complete loss of *unc-5* and *unc-40* function in the DTC. Thus netrin independent guidance mechanisms must be functioning in the DTC to account for the residual DTC guidance along the ventral/dorsal axis observed in these *unc-6*

null mutants. In addition, slight differences are observed in the guidance mechanisms of the UNC-6/UNC-40/UNC-5 pathway in DTCs and in axon guidance. Studying the role of UNC-6, UNC-40 and UNC-5 proteins in the guidance of *C.elegans* DTCs provides a model for understanding the mechanisms, the regulation and the interactions of these guidance molecules within an *in vivo* system.

1.8 The genetic screen for enhancers of DTC migration defects.

C.elegans is an invaluable model system that can be employed for facilitating mass genetic screens in order to identify and characterize genetic interactions between genes involved in a signalling pathway for a particular biological process (Jorgensen and Mango 2002). This model was used for our genetic screen designed to identify candidate *C. elegans* genes involved in DTC guidance during the second phase of DTC migration.

unc-5(e152) (an *unc-5* allele causing significantly lower DTC migration defect frequencies compared to those caused by an *unc-5* null allele) were mutagenized with EMS (ethyl methane sulphonate) to induce small deletions and point mutations, particularily GC-AT transitions in the DNA (Meuth and Arrand 1982). F1 progeny were cloned out and allowed to replicate. Plates with F2 progeny exhibiting higher DTC migration defect frequencies were examined for phenotypes other than those observed in *unc-5* mutants. These phenotypes define the alleles of candidate DTC guidance genes. For example, an *unc-40* null mutation *(e1430)* enhances the frequency of DTC migration defects in *unc-5(e152)*, as *unc-40* is itself a DTC migration gene and is directly involved in *unc-5* mediated DTC guidance. *unc-40* mutants are uncoordinated and slightly Dpy (dumpy, reduced body length), thus these phenotypes define the DTC migration defect enhancer allele *e1430* of the gene *unc-40*.

From this screen, 28 enhancers of ventral to dorsal DTC migration defects were isolated. The majority of the enhancer alleles cloned to date have been identified as alleles of previously characterized DTC migration genes (Table 2).

Table 2 : Summary of the alleles identified in the screen for enhancers of DTC defects in *unc-5(e152)* mutants.

Mutant	LG	# of Alleles	Identity
unc-53	II	5	actin-binding
unc-52	II	3	Perlecan
unc-40	II	2	netrin-receptor
unc-5	IV	3	netrin-receptor
lon-2	X	2	novel, secreted
ced-5	IV	1	DOCK 180 homologue.
lin-7	II	1	PDZ domain
sma-9	X+2.5	1	Schnurri homologue.
enu(IVA)	IV+3.5	?	?
ev675	V+6.5	?	?
ev676	III +4.1	?	?
ev648	X -2.9	?	?
ev697	X+2.0	?	?

Two enhancer mutants, *mig(ev648)* and *enh(ev697)* were genetically mapped to chromosomal regions lacking any known DTC migration genes at the time of mapping. Thus *mig(ev648)* and *enh(ev697)* appear to be alleles of novel cell migration genes involved in DTC guidance either via UNC-6/UNC-5 or parallel DTC guidance pathways.

1.9 Project summary.

Gonad morphology patterning, regulated by the migration and guidance of DTCs in *C.elegans* provides an ideal model for studying UNC-6, UNC-5, UNC-40 and additional cell guidance molecules functioning in their endogenous environment. The capability to organize mass genetic screens in *C.elegans* is a powerful resource for taking the preliminary steps to

identify genes involved in regulating and guiding cell migrations. Once these genes are identified further analysis can be continued in *C.elegans* and later their function can be assayed in more complex model organisms. Identifying genes involved in regulating and guiding cellular migrations is a crucial first step towards elucidating the fundamentals of cell motility and cell guidance.

For this project, I propose to **1)Physically map and clone enhancers *mig(ev648)* and *enh(ev697)* and 2) Define their roles in DTC guidance.**

Cloning each enhancer can result in two possible outcomes. The enhancer may be an allele of a previously cloned gene and thus this screen identifies the genes involvement in DTC migrations and possibly in UNC-6/UNC-5 mediated DTC guidance. However, the enhancer allele may be an allele of a novel DTC migration gene and the identification and characterization of this novel gene requires analysis of its roles in DTC guidance and its interactions in UNC-6/UNC-40/UNC-5 DTC mediated guidance.

2. MATERIALS AND METHODS

2.1 Solutions and media preparation.

Chemicals used for this project were purchased from Sigma, Fisher, Invitrogen New England Biolabs, USB, Qiagen, Roche and Promega and are all of molecular biology grade. All solutions and media used in this project are listed in Section 6.1.

2.2 Maintenance and handling of *C.elegans.*

Bristol *C.elegans* strains were cultured on NGM (nematode growth medium) agar in sterile 9cm or 5cm Petri Dishes (Fisher) with a lawn of *E.coli(Escherichia coli)*OP50 (*Caenorhabditis Elegans* Genetic Centre) on which the nematode feeds. The *E.coli*OP50 strain is a uracil auxotroph with limited growth preventing bacterial over-growth on plates. NGM agar was prepared, autoclaved, poured into petri dishes to 7mm thickness and left to solidify for one day. Once the medium solidified, plates were seeded with 1ml of liquid *E.coli*OP50 cultured in LB(Luria Bertani) broth overnight in a shaker incubator at 37^0C. For each plate, 1ml of liquid *E.coli*OP50 was spread using a flame sterilized glass "hockey stick" ensuring no contact with the edge of the plate creating a centralized lawn of bacteria on each plate. The *E.coli*OP50 was left to dry on the plates overnight and the following day plates were ready for use.

Animals were transferred between plates with a sterilized worm pick; a pasteur pipette with a 1.5cm long platinum wire melted into the pipette tip. Between each animal transfer the wire was flame sterilized and cooled on an area of NGM agar devoid of *E.coli*OP50. Animals were collected under a LeicaMZ6 dissecting microscope by gently amassing *E.coli*OP50 onto the worm pick tip and lightly tapping the sticky *E.coli*OP50

covered pick tip on top of the animal. Once adhered, animals were transferred to new plates by lightly pressing the pick onto the NMG agar of the new plate and allowing animals to crawl off. Strains were grown in a 20^0C incubator (VWR).

2.3 *C.elegans* strains.

The *C.elegans* strains utilized for this project are summarized in Table 3.

2.4 *C.elegans* mutant strain generation.

For this project, the following strains were generated as outlined in Section 6.2:

6.2.1 *tnIs5;mig(ev648)*

6.2.2 *unc-40(e1430);unc-5(e53);mig(ev648)*

6.2.3 *dpy-6(e14)enh(ev697)egl-15(n484)*

6.2.4 *unc-5(e152);sdn-1(zh20)* and *unc-5(e53);sdn-1(zh20)*

6.2.5 *unc-40(e1430);sdn-1(ev697)* and *unc-40(e1430);sdn-1(ev697)*

6.2.6 *unc-6(ev400)sdn-1(ev697)*

6.2.7 *unc-5(e152);sdn-1(zh20);opEx1159,*
unc-5(e152);sdn-1(zh20);opEx1206 and
unc-5(e152);sdn-1(zh20);opEx1198.

6.2.8 *unc-129(ev554) unc-5(e152);sdn-1(ev697)*

6.2.9 *unc-5(e152)egl-20(mu39);sdn-1(ev697)*

6.2.10 *unc-5(e152)lin-3(e1413);sdn-1(ev697)*

6.2.11 *unc-5(e152);dbl-1(ev580);sdn-1(ev697)*

6.2.12 *unc-5(e152);egl-17(e1313)sdn-1(ev697)*

For each mating, a ratio of five males to one hermaphrodite was utilized.

Table 3 : Summary of the *C.elegans* strains used in this project.

Strain name	Source	Phenotype	Description
N2	*Caenorhabditis elegans* Genetics Centre (CGC).	Wild-type.	The *C.elegans* Bristol wild-type reference strain.
mig(ev648)	EMS, screen for enhancers of DTC migration defects.	DTC migration defects in the 2nd and 3rd DTC migration phases resulting in ventral clear patches.	?
enh(ev697)	EMS, screen for enhancers of DTC migration defects.	Embryonic elongation defects (low penetrance).	?
unc-5(e152)	EMS, (Hedgecock, Culotti and Hall 1990) (CGC).	Uncoordinated, (defective backwards locomotion), moderate penetrance of DTC migration defects in the second DTC migration phase resulting in ventral clear patches.	UNC-5 receptor cytoplasmic truncation in the ZU5 domain resulting in a partially functional UNC-5. (Merz et al. 2001)
unc-5(e53)	EMS, (Brenner 1974), (CGC).	Severly uncoordinated, defective backwards locomotion, DTC migration defects in the second DTC migration phase resulting in ventral clear patches.	UNC-5 receptor truncation prior to the first Ig domain in the extra-cellular region of the protein. (NULL) (Killeen et al. 2002)
unc-5(dm11)	Merz, (unpublished).	Slightly uncoordinated, very, very low penetrance of DTC migration defects in the 2nd DTC migration phase resulting in ventral clear patches.	UNC-5 receptor cytoplasmic truncation after the ZU5 domain, retains the majority of UNC-5 functions.

unc-6(ev400)	EMS (Hedgecock, Culotti and Hall 1990),CGC	Uncoordinated, moderately egg-laying defective (animals are bloated), moderate penetrance of DTC migration defects in the 2nd DTC migration phase resulting in ventral clear patches.	Stop codon before the VI domain. (NULL) (Wadsworth, Bhatt and Hedgecock 1996).
unc-40(e1430)	EMS (Hedgecock, Culotti and Hall 1990) CGC.	Uncoordinated, slightly dumpy(reduced body length), low penetrance of DTC migration defects in the 2nd DTC migration phase resulting in ventral clear patches.	Stop codon after the first Ig domain. (NULL) (Chan et al. 1996)
unc-5(e53);mig(ev648)	Merz	Uncoordinated with *unc-5(e53)* and *mig(ev648)* DTC migration defects.	
unc-5(e152);mig(ev648)	Merz	Uncoordinated with *unc-5(e152)* and *mig(ev648)* DTC migration defects.	
unc-40(e1430);unc-5(e53)	Merz.	Severely uncoordinated, DTC migration defects, slightly dumpy and egg-laying defective.	
unc-5(e53);enh(ev697)	Merz	Uncoordinated with *unc-5(e53)* DTC migration defects and *enh(ev697)* embryonic elongation defects.	
unc-5(e152);enh(ev697)	Merz	Uncoordinated with *unc-5(e152)* DTC migration defects and *enh(ev697)* embryonic elongation defects.	

*evIs99**	(Su et al. 2000)	Dorsal clear patches due to precocious turning of the DTC dorsally during the first DTC migration phase.	Transgenic strain with *unc-5* expression regulated by the *emb-9* promoter, resulting in early ectopic *unc-5* expression in DTCs during the first, ventral migration phase. Array integrated on LGI.
evIs99; mig(ev648)	Merz	Dorsal clear patches due to precocious turning of the DTC dorsally and *mig(ev648)* DTC migration defects.	
evIs99; enh(ev697)	Merz	Dorsal clear patches due to precocious turning of the DTC dorsally and *enh(ev697)* embryonic elongation defects.	
dpy-6(e14)unc-115(e2225)	Merz D, (unpublished).	Reduced body length (dumpy, dpy) and uncoordinated.	
dpy-6(e14)egl-15(n484)	Merz D, (unpublished).	Reduced body length (dumpy, dpy) and egg-laying (egl) defective resulting in the formation of live worm sacs.	
tnIs5	(Hall et al. 1999), CGC	Gonadal sheath cells 1-4 express GFP, outling gonad morphology.	A transgenic strain with 2.23Kb of *lim-7* upstream regulatory sequence fused to the genetic coding sequence for the first 61 amino acids of LIM-7 protein in frame with a GFP sequence, resulting in expression of the LIM-7::GFP fusion protein in gonadal sheath cells.

mig-23(k180)	(Nishiwaki et al. 2003) from a genetic screen for defects in gonad morphogenesis. (CGC).	DTC migration defects in the second DTC migration phase, not fully penetrant.	NULL (Nishiwaki et al. 2003)
sdn-1(zh20)	EMS, (Rhiner et al 2005) obtained from C. Rhiner via personal communication.	Variable egg-laying defects resulting in bloated animals, defects in backward locomotion and a low penetrance of embryonic elongation defects.	Deletion of 1260bp within the coding region. NULL (Rhiner et al 2005)
sdn-1(ev697)	EMS, screen for enhancers of DTC migration defects.	Embryonic elongation defects.	A 610 G>T mutation in the *sdn-1* coding sequence resulting in a E203X creating a truncation in SDN-1 before the trans-membrane domain.
mig-23(ev648)	EMS, screen for enhancers of DTC migration defects.	DTC migration defects in the second and third DTC migration phases resulting in ventral clear patches.	A 335C>T mutation in the *mig-23* coding sequence resulting in a A112V substitution.
dbl-1(ev580)	CGC.	Reduced body length (fully penetrant).	
egl-17(e1313)	CGC.	Severe bloating due to egg laying defects (fully penetrant).	
unc-129(ev554)unc-5(e152)	(Merz et al 2003).	Uncoordinated, moderate penetrance of DTC migration defects in the 2nd DTC migration phase resulting in ventral clear patches.	*unc-129(ev554)* NULL.

unc-5(e152)egl-20(mu39)	(Merz et al. 2003) .	Uncoordinated, DTC migration defects in the 2nd DTC migration phase resulting in ventral clear patches, bloated due to defects in egg-laying.	
unc-5(e152)lin-3(e1413)	(Merz et al. 2003)	Uncoordinated, moderate incidence of DTC migration defects in the 2nd DTC migration phase resulting in ventral clear patches, vulvaless and bloated due to defects in egg-laying.	
*opEx1159***	(Rhiner et al. 2005) obtained from C.Rhiner via personal communication.	Strong GFP expression in the head and tail, weaker GFP expression in all 6 touch cell neurons, variable egg laying defects (bloated animals) and embryonic elongation defects.	A transgenic strain, *sdn-1(zh20)* with an extra-chromosomal array comprising 216bp of *dpy-7* promoter fused to *sdn-1* cDNA (hypodermal expression of *sdn-1)* and a *lin-48* promoter fused to GFP as a marker. (Rhiner et al. 2005)
opEx1206	(Rhiner et al. 2005) obtained from C.Rhiner via personal communication.	Strong GFP expression in the head and tail, weaker GFP expression in all 6 touch cell neurons, variable egg laying defects (bloated animals) and embryonic elongation defects.	A transgenic *sdn-1(zh20)* strain with an extra-chromosomal array comprising 2189bp of *unc-119* promoter fused to *sdn-1* cDNA (hypodermal expression of *sdn-1)* and *lin-48* promoter fused to GFP as a marker. (Rhiner et al. 2005)

opEx1198	(Rhiner 2005) obtained from C.Rhiner via personal communication.	Strong GFP expression in the head and tail, weaker GFP expression in all 6 touch cell neurons, variable egg laying defects (bloated animals) and embryonic elongation defects.	A transgenic strain with an extra-chromosomal comprising 2.8Kb of *sdn-1* promoter fused to *sdn-1* cDNA (hypodermal expression of *sdn-1)* and *lin-48* promoter fused to GFP as a marker. In addition, a *mec-4* ::gfp transgene in integrated on LGI for visualization.(Rhiner et al. 2005)
opIs170	(Rhiner et al. 2005) obtained from C. Rhiner via personal communication.	SDN-1::GFP expression in hypodermis, ventral nerve cord, nerve ring and commissures.	A transgenic sdn-1(zh20) strain with an integrated array comprised of 2.8Kb of *sdn-1* promoter sequence followed by the entire *sdn-1* coding sequence in frame with a GFP sequence.(Rhiner et al. 2005)

Note : Phenotype descriptions for each strain only describes the phenotypes relevant to this project. Additional phenotype information for each strain can be obtained from www.wormbase.org. * *Is:* strain has an integrated transgenic DNA array, ***Ex:* strain

2.5 *C.elegans* phenotype analysis.

2.5.1 General phenotype analysis.

In *C.elegans*, phenotypes are used to deduce the genotype of a particular strain. For example, animals of a mutant *unc-5* strain exhibit ventral clear patches and uncoordinated movement. A LeicaMZ6 dissecting microscope was utilized for the majority of the *C.elegans* general manipulations such as transferring animals between plates as well as for general phenotype analysis for deducing strain genotype. For phenotype analysis requiring higher magnification a Zeiss Stemi M2BIO QUAD stereomicroscope with 10X and 20X magnification was utilized. This microscope is equipped with a Zeiss AxioCam HRm and AxioCam 4.0 software and was also used for imaging.

2.5.2 Identifying and scoring DTC migration defects.

DTCs that fail to reach the dorsal muscle band during their second migration phase cause a gonad morphology defect that is identified by a ventral clear patch in the animal. The clear patch is easily visualized under the LeicaMZ6 dissecting microscope at low magnification facilitating DTC migration defect scoring.

DTC migration defects were scored in animals at the L4 larval stage of F1 progeny from a single clone. DTC migration defects in the ventral to dorsal migration phase were counted based on their occurrence (anterior or posterior/ventral or dorsal clear patches) for each F1 progeny. Proportions of anterior and posterior DTC migration defects for each strain were calculated assuming a binomial distribution. Proportions were expressed as percentages and standard errors calculated as described in (Hedgecock, Culotti and Hall 1990).

2.5.3 Live *C.elegans* imaging.

Live photographs of animals (bright field or with a 470 GFP filter) on NGM agar plates were taken using a Zeiss Stemi M2BIO QUAD stereomicroscope equipped with a Zeiss AxioCam HRm and AxioCam 4.0 software. Magnifications (10X or 20X) are specified in each photograph. For all live photographs, the anterior posterior axis runs from left to right and the dorsal ventral axis runs top to bottom.

2.5.4 Still *C.elegans* imaging.

For phenotype analysis requiring higher magnifications, animals were mounted on a 2.5% agarose pad dried to a thin cover slip and immobilized with 1mM levamisol (Brenner 1974). A Zeiss Axio Imager equipped with an Axiocam MRc camera and Axiovision Rel 4.4 software was used and photos were taken with a 63X lens with DIC (differential interference contrast). For all still photographs, the anterior posterior axis runs from left to right and the dorsal ventral axis runs top to bottom.

2.6 Preparation of agarose gels

Agarose gels were prepared in either a 100x115mm small casting tray (Fisherbrand horizontal unit mini-plus) or 130x150mm medium casting tray (Fisherbrand horizontal unit midi). Appropriate amounts of agarose (Promega) were weighed out and mixed with the appropriate amount of 1X TAE (depending on the desired size and width of the gel) in a glass 250ml flask (Pyrex). The mixture was gently shaken and placed to boil in a microwave until the agarose dissolved. A magnetic stir bar was added to the flask and the mixture was placed on a P-353 stirrer on a low setting. Ethidium Bromide (10mg/ml) was added to the mixture (10% of total TAE buffer volume) and stirred. Casting trays were positioned

vertically in the trays, combs (with various well numbers) inserted and the gel was poured into the tray and left to set. Once the gel solidified the comb was remove and the tray was rearranged horizontally and submerged in 1 X TAE buffer with 10% ethidium bromide. Samples were loaded with an OrangeG running dye (1:6 of total volume being loaded) and run against a 1Kb DNA ladder (Invitrogen). The gels were run at 80-90 volts (mini-plus) or 100-120 volts (midi) (Fisher FB300 power pack) until desired band resolution was achieved. Gels were visualized on an AlphaImager2200 trans-illuminator and photographed with the AlphaEaseFC software.

2.7 Cosmid microinjections

2.7.1 Cosmid preparation

Cosmids C03B1, T22E5, K10C2, K04E7, C15B12, F22A3, T14E8 and T28B4 (spanning the region that *mig(ev648)* was mapped to) and cosmids F41E7, R07E3, F46F6, ZC504, C39B10, C33D3, F14F3, F59F5, F57C7, M79, F11A1, F13E6, C46B5, T01C1, R07A4 (spanning the area that *enh(ev697)* was mapped to) were obtained from Alan Coulsan, Cambridge University UK (Described in Section 6.3). Cosmids arrived previously transformed in *E.coli* and were immediately streaked out onto LB agar (in 9mm diameter plates) with kanamycin (75µg/ml) or LB agar plates with ampicillin (75µg/ml). Plates were incubated at 37^0C overnight. Ampicillin/Kanamycin resistant colonies were picked with a 1-10µl pipette tip and aseptically placed in 10ml sterile culture tubes (Simport) containing 4ml of liquid LB broth with 50µg/ml of kanamycin or ampicillin. Liquid cultures were incubated at 37^0C overnight in a shaking incubator. A Qiagen mini-prep kit was utilized to isolate cosmids (as per manufacturers instructions). Cosmid identity was verified by resolving restriction enzyme digest patterns (New England Biolabs) on a 1% agarose gel.

2.7.2 Injection mixture preparation.

The cosmid injection mixture consisted of 10ng/µl of the appropriate cosmid in elution buffer, 50ng/µl of a pTG96GFP plasmid used as a co-transformation marker, 40ng/µl pKS plasmid and ddH_20 for a total volume of 20µl. The pTG96GFP construct (Gu, Orita and Han 1998) contains a GFP coding sequence regulated by the nuclear specific promoter of *sur-5* inducing GFP expression in all nuclei of transgenic animals.

2.7.3 Microinjection.

Each cosmid mixture was microinjected using a fine capillary needle and compressed air (nitrogen) into the syncytial distal region of adult *mig(ev648)* and *enh(ev697)* animal gonads. Using a dissecting microscope, animals were placed in mineral oil covering a thin pad of 2.0% agarose that was dried onto a 24X50 microscope cover glass (Fisher). Once the animal adhered to the pad, the cover glass was transferred to an inverted LeicaDMIL DIC microscope equipped with a manoeuvrable stage and injection needle micromanipulator. Injection needles were made from glass capillaries using a needle puller. The injection mixture was added into a capillary needle and the needle was attached to the micromanipulator. The needle micromanipulator was connected to tubing and joined to a Nitrogen pressure tank forcing air through the needle and ejecting the mixture in a controllable manner. Using the movable stage the animal was positioned onto the needle ensuring that the needle pierced through the cuticle and into the gonad syncytium. The air was activated and the cosmid mixture was forced into the distal gonad region. Microinjection of transgenic DNA in *C.elegans* is reviewed in (Mello and Fire 1995). Briefly, DNA fragments contained within the mixture microinjected into the gonad syncytium of *C.elegans* undergo recombination forming multi-copy, extra-chomosomal (*Ex*)

arrays of repeating GFP, pKS and cosmid DNA fragments in no particular order or number. Arrays within the syncytium are taken up by germ cells as their plasma membrane is formed and expressed using the animals endogenous transcription factors. The extra-chromosomal arrays remain present in the nucleus and have a 5%-95% transmission frequency.

Injected animals were cloned onto NGM plates and F1 progeny were analysed for GFP expression. F1 progeny expressing GFP were cloned out and the stability of the transgene was examined in the F2 progeny. At least two independent stable transgenic lines for each cosmid microinjected were generated and used for analysis. Rescue was deduced by phenotype analysis using a Zeiss Stemi M2BIO QUAD microscope equipped with a 470 GFP lense.

2.8 Sequencing.

2.8.1 *C.elegans* genomic DNA isolation.

C.elegans genomic DNA was prepared from N2 (wild-type), *mig(ev648)* and *enh(ev697)* strains. Each strain was grown up on three, 150mm thick Rich Agarose Plates until just before starvation for a total of nine plates. Each strain was washed from plates with M9 buffer and collected into a 15ml plastic conical tube (Corning). To remove bacterial residue, each tube was centrifuged for 30 seconds at 300-500x*g*, the supernatant was removed and the pellet was re-suspended in ddH_20 and centrifuged again. After three washings, the animal pellet was re-suspended in 1ml H_20, transferred to a 1.5ml Eppendorf tube and centrifuged for 1 min at 13 200rpm. The supernatant was removed and 500ul of genomic worm lysis solution was added. Tubes were incubated in a -80^0 C dry ice bath for 30 mins, thawed and then re-incubate for 30-60 mins in a water bath at 55-65^0C with occasional agitation. Lysate was then centrifuged at 13 200 rpm and the aqueous phase was

transferred to a new tube, leaving behind the eggshells and worm carcasses. To clean the DNA, an equal volume (~ 500µl) of phenol/$CHCl_3$ was added to the tube and centrifuged at 13 200 rpm for 1 min. The aqueous phase was transferred to a new sterile Eppendorf tube, an equal volume of $CHCl_3$ was added and the mixture was re-centrifuged at 13 200 rpm for 1 min. Addition of $CHCl_3$ followed by centrifugation and removal of the aqueous phase was repeated for a total of three times. After the third centrifugation, the aqueous phase was transferred into another tube and 2.5 volumes of 100% EtOH was added to precipitate the DNA from solution. The mixture was centrifuged for 15mins at 16,000 x*g*. The aqueous phase was then removed and 1ml of low TE was added. The tube was sealed with parafilm and the DNA was re-suspended into solution on a slow moving rotator at -4^0C.

2.8.2 Primer design and PCR amplification.

Primers for *mig-23, sdn-1* and *nas-33* sequencing were designed based in the most recent gene sequences submitted to WormBase (www.wormbase.org). Primers for *nas-33* were designed to amplify exonic regions and primers for *mig-23* and *sdn-1* were designed to amplify the entire gene coding sequence including introns. Primer sets were chosen based on their melting temperatures, tendencies to form hair-pins and their abilities to homodimerize and heterodimerize as deduced by IDT Oligoanalyzed3.0. (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx). Primers were blasted against the entire *C.elegans* genome (NCBI BLAST) to ensure specificity. Primers for each gene are described in Tables 4, 5 and 6 and were ordered through Qiagen(*nas-33*) and Invitrogen (*mig-23* and *sdn-1*).

Table 4 : Primers used for sequencing *nas-33*

Primer Name	Primer Sequence (5' to 3')
MSK04-F1	CAACACCACCCATTGAGACG
MSK04-R1	GCAAAGCCGTCCTTGGTGAAC
MSK04- F2	GGTCTTTGAGTGGGAGCTGT
MSK04-R2	CACTCTCAAAGCAGCAACGA
MSK04 - F3	TCGTTGCTGCTTTGAGAGTG
MSK04 - R3	CACCGTTGAACCCAGTAGACCT
MSK04 - F4	GGCAGTTTATCCGATTGGTGC
MSK04 - R4	GGTGATGGATGTGAAACGGT
MSK04 - F5	GGTCTCGCTTGTTCATGCCA
MSK04 - R5	GCTCAAAACGGCTTTCGTGT
MSK04 - F6	CGAGACCACGTTGTTCCGTT
MSK04 - R6	CTGAGCCGGAACCACAACAA
MSK04 - F7	CCACTGGACGGGATTACAGT
MSK04 - R7	CCTTGGCATCGGTGAGATTT

Table 5 : Primers used for sequencing *mig-23*

Primer Name	Primer sequence	#bp	Tm
Mig-1A	TCG GAA GTG CGC TTT GAA TG	20	56.1
Mig-1B	GACACGCATAGGATCACCGCA	21	59.6
Mig-2A	TCC GAA TTG CAG CGT CCG A	19	59.7
Mig-2B	CTGAACGACCGATTTCCACCA	21	57.4
Mig-3A	TCC CTG AAA AAC CGC CGA AA	20	57.5
Mig-3B	CTTGGGCAGGTTTTGTTCCA	20	56
Mig-4A	GTG ATG CAG GGT CAA CTG GA	20	57
Mig-4B	CTT GCA TGT GCT GGC GAG GTT	21	60.9
Mig-5A	TGGAACAAAACCTGCCCAAG	20	56
Mig-5B	GTCTTCCCTGCATCCGAGGT	20	59
Mig-6A	GATATGGGTGGAGCAAGTGCT	21	57
Mig-6B	TCCGCATCATACTGTCCTCCA	21	57
Mig-7A	GTCTGTAAAGCTGAAGCGGCA	21	57
Mig-7B	AGCTGAAACCGTGCTGCAA	20	59
Mig-8A	GTACCCGAGAGCTGACGAGGA	21	60
Mig-8B	GAAACCAAGGCCCAATCCCA	20	59
Mig-9A	CAATGGGCTCTCGGAGCAATG	21	58
Mig-9B	GAGGACCGACGTTTGTCATC	20	55.6
Mig-10A	TGGGATTGGGCCTTGGTTTC	21	58.6
Mig-10B	CATCTGGAGGTTCCTGCTTG	20	55.6

Table 6 : Primers used for sequencing *sdn-1*

Primer Name	Primer Sequence	#bp	Tm
sdn-1A	TCCTCCTCCACCACAACACCA	21	60.5
sdn-1B	TTCGTCGTCGGTTGGGTAG	19	56.9
sdn-2A	TTGCAGCAGGTCGAAGGAAG	20	57.8
sdn-2B	CTCCTTGTCGTTTGCCGCTG	20	59.2
sdn-3A	CTACAGCGGTTTGTGTCGGC	20	58.8
sdn-3B	GAAGCCATTTGCCAGTGTCT	20	55
sdn-4A	CAGAACGCCAAGGTCAGCAG	20	58
sdn-4B	AAGAGGCCACGCCATCTGTC	20	60
sdn-5A	CAAGCCTATCCGTTCCGTCTG	21	57.7
sdn-5B	CTCTCATCGTCTTCCCACCA	20	56.8
sdn-6A	CAGAATGGGGACCCCTTCGT	20	59.4
sdn-6B	CTTCCGTCCCACCATCCCGA	20	61
sdn-7A	CTGAGCAGCATCCCACATC	19	56
sdn-7B	CACCGCAACGAGAACACCT	19	58
sdn-8A	GTGTCTGTGAGGAAAAGGGGA	21	56.7
sdn-8B	GTCTTGCTTGCTTGGGTTCATC	21	57
sdn-9A	AGGTGTTCTCGTTGCGGTG	19	58
sdn-9B	TCCCTACCCCTAAGTGGGTCT	21	59

PCR fragments for sequencing were amplified using an Invitrogen Platinum *Taq* DNA polymerase kit. For each primer set, a temperature gradient setting on the Eppendorf Mastercycler was used to determine primer set melting temperatures for optimal and specific DNA amplification. Each gene was sequenced in the *C.elegans* N2 wild-type strain in addition to the mutant strain.

2.8.3 Gel extraction and quantification of DNA fragments.

PCR amplified DNA fragments to be sequenced were resolved on a 1.5% agarose gel and excised from the gel using a clean scalpel. A Qiagen Gel extraction kit was utilized to extract product from the gel (as per manufacturers instructions). To ensure that the final DNA fragment concentration was at least 50ng/µl (as recommended by the The Centre for Applied Genetics Sequencing Facility in Toronto), 1µl of the fragment solution was diluted

in 9ul of low TE and resolved on a 1% agarose gel against a low mass ladder (Invitrogen). PCR fragments were sent for sequencing (along with their respective primers at a concentration of 5 pmol/μl and a minimum volume of 10μl) to The Centre for Applied Genetics Sequencing Facility in Toronto. Each fragment was sequenced using both forward and reverse primers in order to validate the accuracy of all sequencing results.

2.8.4 DNA sequence alignments.

Automated sequencing results were sent from Toronto to Winnipeg via email in text format. The ClustalW alignment program provided by EBI (www.ebi.ac.uk/clustalw/) was used for all alignment (sequences and protein). The most recent sequences submitted to Wormbase (www.wormbase.org) (*nas-33*: GenBank Assession# U39666; *mig-23*:GenBank Assession#NM_076593; *sdn-1*: GenBank Assession#NM_171972) were used to compare the sequences obtained from N2 and mutant strains for each fragment. Mutations were deduced based on observed variations in the sequences compared by alignment.

2.9 DTC migration defect enhancement/suppression assays.

Previous studies have established a methodology for analyzing genetic interactions of *unc-5/unc-40/unc-6* with various other cell migration genes in order to deduce their roles in cell guidance and place them within the hierarchies of cell guidance signalling pathways. For example, the frequency of DTC migration defects in *unc-5(e53);unc-6(ev400)* mutants and *unc-40(e1430);unc-6(ev400)* mutants is not enhanced compared when to the frequency of DTC migration defects in *unc-6(ev400)* (Hedgecock, Culotti and Hall 1990; Merz *et al.* 2001). Eliminating *unc-6* and its receptors *unc-5* and *unc-40* results in the same outcome as simply eliminating *unc-6* as *unc-6* signals through its downstream receptors for ventral-

dorsal DTC guidance. Placing each enhancer allele in *unc-5/unc-40/unc-6* genetic backgrounds can assess whether the roles of *mig(ev648)* and *enh(ev697)* alleles in DTC guidance are limited to *unc-5/unc-40/unc-6* or whether each allele functions in a parallel DTC guidance pathway.

2.10 SNP mapping.

2.10.1 Selection of genetic recombination markers.

Three factor genetic mapping mapped *enh(ev697)* to LGX:+2.0. The genes *dpy-6*(LGX:0.0) and *egl-15*(LGX:2.86) were chosen as genetic markers for identifying recombinant animals. Phenotypes of both *dpy-6(e14)* and *egl-15(n484)* mutant strains, reduced body length and egg-laying deficiencies respectively, are easily distinguishable at low magnification facilitating phenotypic analysis of high numbers of animals.

2.10.2 SNP selection and amplification.

An alternate *C.elegans* strain, CB4856 Hawaiian, bears SNPs (single nucleotide polymorphisms) in its DNA sequence relative to the commonly used Bristol strain. These SNPs have been identified, mapped and recorded in Wormbase (www.wormbase.org). SNPs in the Bristol strain within the genetic region between LGX:1.8 and LGX:2.86 were selected. Three verified SNPs were identified, pkP6128(LGX:1.932), pkP6040(LGX:2.31) and pkP6160(LGX:2.54) (Table 7) and primers for amplifying the area surrounding and including each SNP were ordered (Qiagen).

2.10.3 Recombinant identification.

A *dpy-6(e14)enh(ev697)egl-15(n484)* mutant strain was generated as outlined in Section 6.2.3. Hawaiian males were crossed to a Bristol *dpy-6(e14)enh(ev697)egl-15(n484)* hermaphrodite and wild-type F1 progeny (*Bristol-dpy-6(e14)enh(ev697)egl-15(n484)/Hawaiian-+;+;+*) were selected and cloned out. Recombinations occurring between the Bristol *dpy-6(e14)enh(ev697)egl-15(n484)* chromatid and Hawaiian +;+;+ chromatid were recovered by isolating F2 progeny that were Egl and non Dpy or Dpy and non Egl. From these progeny, seven *egl-15(n484),* non-Dpy lines were isolated and labelled C1 to C9.

2.10.4 DNA isolation from recombinant strains.

DNA from the recombinant strains was isolated using the single worm lysis method. Five animals from each recombinant strain were placed in 0.2ml thin walled PCR tubes (Biocan) filled with single worm lysis buffer. Samples were placed in an ethanol and dry ice bath for 15mins. 1µl of mineral oil was added to each sample and the tubes were placed in the Eppendorf Mastercycler where they were incubated at 60^0C for 1 hour, 95^0C for 15 mins and immediately stored at -20^oC.

2.10.5 PCR amplification.

SNPs in each recombinant were PCR amplified with an Invitrogen Platinum *Taq* DNA polymerase kit utilizing the appropriate primers. The temperature gradient setting on the Eppendorf Mastercycler was used for each primer set to determine melting temperatures for optimal and specific DNA amplification. PCR amplification products were resolved on a 1% agarose gel to ensure that the correct fragment was amplified.

Table7 : SNPs used for mapping *enh(ev697)*

SNP Name	Genetic Locus	Forward Primer	Reverse Primer	Fragment Length	Restriction Enzyme	Bristol Fragments	Hawaiian Fragments
pkP6128	LGX 1.932	ACTTGGTGAGCATTCCGCAC	ACCATATCAAGTGGTGTCGG	649bp	NsiI	378bp/271bp	649bp
pkP6040	LGX 2.38	CCGTTTGAACTTCTAGGTCG	CGTCCGTATCGTTTTCCTC	725bp	EcoRV	725bp	374bp/351bp
pkP6160	LGX 2.54	ATTATGAACGTGGTCCTTTCCG	ATACATATTTCTCGCACCGTTC	418bp	HinfI	418bp	296bp/122bp

2.10.6 Restriction digest of PCR amplified DNA fragments.

PCR amplification fragments spanning each SNP were treated with the appropriate restriction enzyme to verify their restriction pattern and determine whether fragments were of Bristol or Hawaiian origin. All restriction enzyme reactions were performed in Eppendorf tubes and consisted of 1µl of the PCR amplified DNA fragment, 1µl of enzyme (NEB), 2µl of the appropriate restriction enzyme buffer (NEB) and 16µl of ddH_20. All restriction digests were performed in a 37^0C water bath (enzyme manufacturers instructions) for at least 1 hour. Digested products were resolved on a 1% agarose gel and visualize as described in Section 2.6.

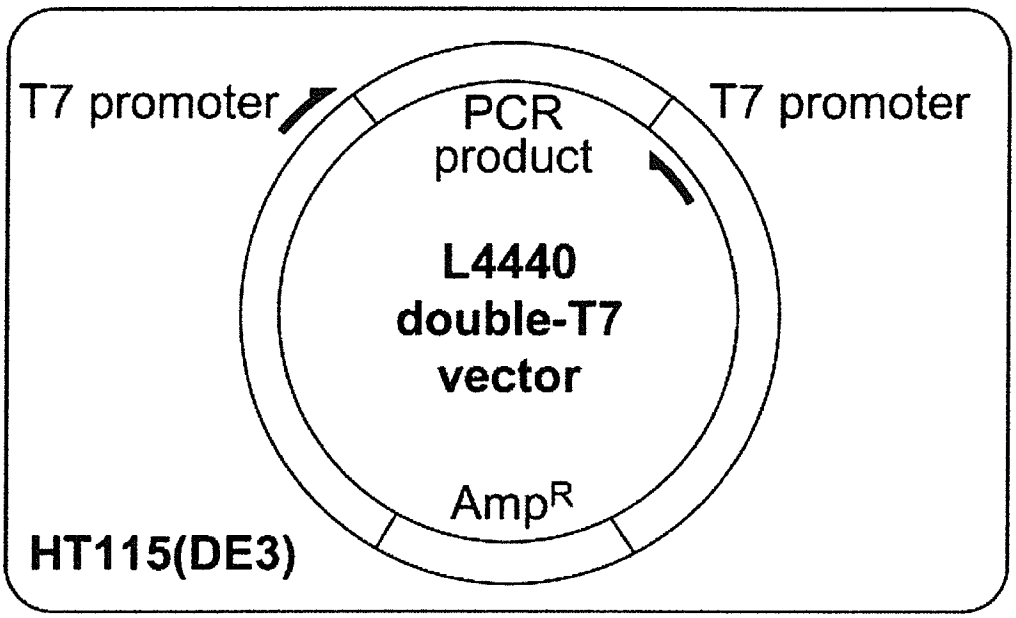
2.10.7 Mapping the recombination.

The restriction enzyme digest pattern of each SNP DNA fragment was used to deduce the origin (Bristol or Hawaiian) of each SNP fragment from the recombinant strains. Once the origin of each SNP fragment from each recombinant was deduced, the genetic recombination position for each recombinant strain was inferred. Recombinants were scored for the presence of the *enh(ev697)* embryonic elongation phenotype and *enh(ev697)* was mapped relative to the location of the recombination between the two SNPs.

2.11 RNAi media preparation.

RNA interference by feeding method is described in (Kammath 2001). The *E.coli* HT115 strain X-5K23 was obtained from P.Roy (University of Toronto). This bacterial strain is deficient in RNAse III, can induce T7 bacteriophage polymerase activity in the presence of isopropylthiogalactoside (IPTG) and contains the L4440 vector described in Figure 4 (Timmons and Fire 1998). The vector contains an ampicillin resistance gene for

selection and the *abl-1* fragment inserted into an EcoRV site flanked with T7-polymerase promoters allowing bi-directional transcription of the fragment to produce *abl-1* dsRNA. NGM agar was prepared and carbenicillin (Fisher) [25ug/ml] and IPTG [1mM] (Fisher) (1.5mg/ml) were added prior to pouring. The media was poured into 9cm petri dishes (Fisher) and allowed to solidify for 4-7 days. A 10ml liquid culture of the X-5K23 *E.Coli* was prepared in liquid LB broth with 50ug/ml ampicillin and incubated in a shaking incubator overnight at 37^0C. The following day the plates were seeded with X-5K23 *E.Coli* and allowed to dry overnight. Plates were stored at $+4^0$C until used to maintain the stability of the IPTG. *enh(ev697)* mutants at the L3 larval stage were cloned out onto the prepared RNAi plates. Plates were left at room temperature and F1 progeny were analysed within three days.



(Kamath 2001)

Figure 4 : The L4440 vector for RNAi assays in *C.elegans.*
Schematic diagram depicting the L4440 vector and its components in an HT115(DE3) *E.coli* cell.

3 RESULTS AND DISCUSSION

3.1 Cloning and characterizing the DTC guidance roles of *mig(ev648).*

The *mig(ev648)* allele was identified in a genetic screen for enhancers of DTC migration defects in *unc-5(e152)* mutants. *mig(ev648)* mutants in an *unc-5* wild-type or *unc-5* mutant background exhibit DTC migration defects in the second and third DTC migratory phases. Specifically, the DTCs fail to migrate dorsally causing the third migration phase to occur along the ventral muscle band (as seen in *unc-5*, *unc-6* or *unc-40* mutants) or the DTC executes the ventral to dorsal migration phase and during the third migration phase wanders on and off of the dorsal muscle band. These aberrant DTC migratory pathways in *mig(ev648)* mutants were observed upon examining differences in clear patches of *unc-5(e152)* and *mig(ev648)* (Figure 3).

3.1.1 The gonad morphology of the *mig(ev648)* mutants.

Phenotypic differences in the ventral clear patches of *unc-5* and *mig(ev648)* mutants resulting from the gonad morphology defects caused by aberrant DTC migration patterns were observed at low magnification (Figure 3). Clear ventral patches in *mig(ev648)* mutants appear blurry and the edges not well defined when compared to *unc-5(e152)* clear patches, suggesting a structural gonad morphology defect (possibly due to defects in the position or structure of the gonadal sheath cells) in *mig(ev648)* mutants rather than a positional gonad morphology defect due to aberrant DTC migration patterns such as those observed in *unc-5/unc-40/unc-6* mutants. The adult hermaphrodite gonad structure has been reviewed (Hall *et al.* 1999). Five gonadal sheath cell pairs join together to form each tubular gonad arm. The most distal sheath cell pair enclosing the syncytium containing the germline nuclei, are

located in the area of the gonad arm along the dorsal muscle band patterned by the third phase of DTC migration. The transgenic strain *tnIs5* expresses GFP under the regulation of the gonadal sheath cell specific *lim-7* promoter allowing for the visualization of the distal gonad morphology outline. The *tnIs5;mig(ev648)* strain was generated as outlined in Section 6.2.1 and the gonad morphology of *tnIs5;mig(ev648)* mutants was compared to the gonad morphology in wild-type *tnIs5* (Section 2.5.3). An irregular gonad morphology was observed in *tnIs5;mig(ev648)* (Figure 5). Sheath cells appeared to be normal and properly localized in *tnIs5;mig(ev648)* indicating gonad morphology defects observed in *mig(ev648)* appear to be a result of aberrant DTC migration patterns rather than structural defects caused by the aberrant structure of the sheath cells. In *tnIs5*, DTC migrate along the dorsal muscle band resulting in a straight, linear gonad arm. In *tnIs5;mig(ev648)* the DTC wanders off the dorsal muscle band during the third DTC migration phase resulting in a meandering distal gonad arm. Thus the *mig(ev648)* allele appears to play a role in DTC guidance during the second and third DTC migration phases and possibly is involved in either mediating the DTCs or the gonadal sheath cells ability to adhere to the dorsal muscle band or is required for selecting the substrate over which the DTC migrates.

3.1.2 Genetic interactions of *mig(ev648)*, *unc-5* and *unc-6*.

Our genetic screen identified *mig (ev648)* as an enhancer of DTC migration defects in *unc-5(e152)* mutants indicating *mig(ev648)* has a role in guiding DTC migrations. However it remains to be shown whether or not *mig(ev648)* is involved in *unc-6/unc-5/unc-40* DTC guidance mechanisms during the second migration phase or whether it is limited to parallel, *unc-5* independent guidance pathways. This uncertainty can be addressed by employing

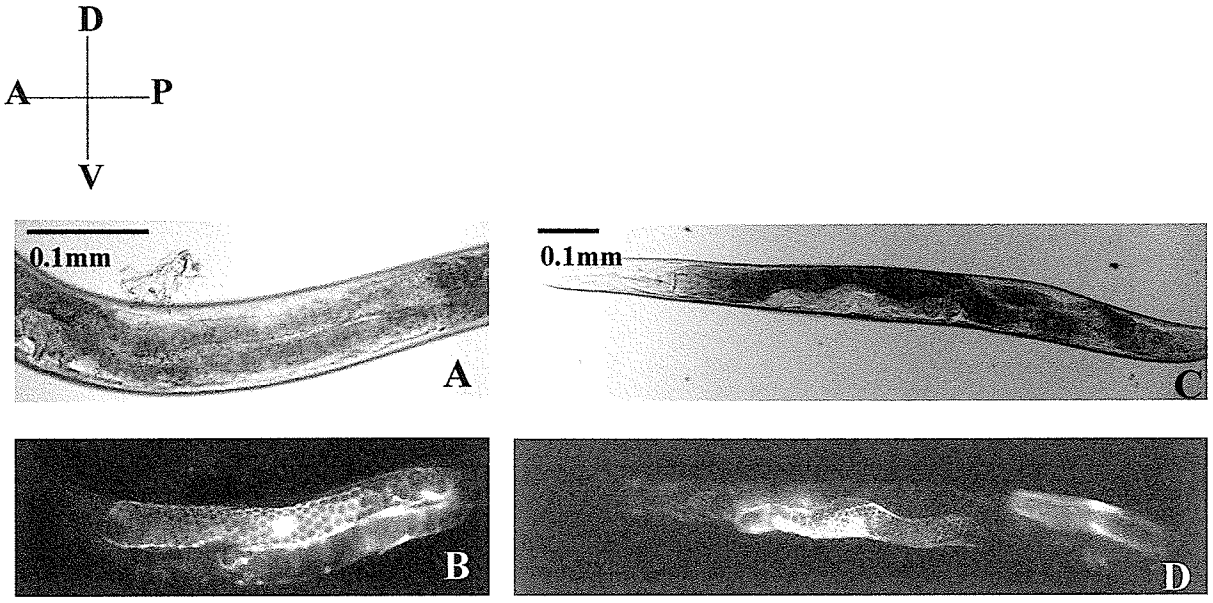


Figure 5 : Images of *tnIs5* and *tnIs5;mig(ev648)* gonad morphologies. Live, bright-field photographs. Gonad morphology of *tnIs5*(A,B) and *tnIs5;mig(ev648)* (C,D) mutants visualized with a GFP regulated by the *lim-7* promoter for gonadal sheath cell specific expression. The gonad morphology of *tnIs5* (B) is representative of a straight DTC migration pattern whereas the gonad morphology of *tnIs5;mig(ev648)* (D) is representative of an aberrant, non-linear DTC migration pathway along the dorsal muscle band. Brightfield images depict the clear patches (representing DTC migration defects) present in *mig(ev648)* mutants (C) and not observed in N2 (A). Photos were taken on the M2BIO QUAD stereoscope.

classical genetic methods to elucidate genetic interactions of *mig(ev648)* with *unc-5/unc-6/unc-40* for DTC guidance (Section 2.9).

The *unc-5(e152)* allele encodes a truncated form of the UNC-5 protein retaining partial DTC guidance function. The *e152* allele was used in the genetic screen in order to recover DTC guidance genes otherwise undetectable with a complete loss of *unc-5* function as all components directly involved in UNC-5 guidance functions and signalling pathways are silenced. Therefore, if a gene does not enhance the frequency of DTC migration defects in *unc-5(e53)*(a null), this gene must be dependent on *unc-5* function for DTC guidance. Alternatively, if a gene enhances the frequency of DTC migration defects in *unc-5(e53)*, this gene must be functioning, at least in part, in some parallel guidance pathway that has also been disrupted in addition to the complete loss of *unc-5* guidance.

To determine the role of *mig(ev648)* in *unc-5* mediated DTC guidance, DTC migration defects in *unc-5(e53);mig(ev648)* were scored and compared to the frequency of DTC migration defects in *unc-5(e53).* The frequency of posterior DTC migration defects in the *unc-5(e53);mig(ev648)* mutants was significantly increased in the posterior compared to the frequency observed in *unc-5(e53)* (Table 8).

Table 8 : DTC migration defects in *unc-5* and *mig-23(ev648)* mutants.

unc-5	*mig-23*	Anterior DTC	Posterior DTC	*n*
e152	WT	8 ± 1	40 ± 1	1464
WT	*mig(ev648)*	36 ± 2	54 ± 2	400
e152	*mig(ev648)*	32 ± 2**[a]	71 ± 2**	201
e53	WT	28 ± 2	53 ± 2	951
e53	*mig(ev648)*	31 ± 2*	76 ± 2**	407
e53;unc-40(e1430)	WT			
e53;unc-40(e1430)	*mig(ev648)*	66 ± 2	88 ± 1	708

[a] Each statistical comparison is against the frequency of DTC migration defects in *unc-5* or *unc-5;unc-40* strain alone. *$P<0.05$; **$P<0.001$.

The enhancement of DTC migration defects in *unc-5(e53);mig(ev648)* double mutants demonstrates *mig(ev648)* is functioning, at least in part in a parallel DTC guidance pathway and is not limited to *unc-5* mediated guidance of the DTCs. To assess the possibility that *mig(ev648)* is limited to *unc-40* or *unc-6* for DTC guidance, an *unc-40(e1430);unc-5(e53);mig(ev648)* triple mutant was generated (Section 6.2.2). *unc-40(e1430);unc-5(e53)* mutants (both null alleles) represent a complete loss of the UNC-6 guidance pathway as DTC migration defect frequencies in *unc-40(e1430);unc-5(e53)* are identical to the frequencies in *unc-6(ev400)* (Hedgecock, Culotti and Hall 1990; Merz *et al.* 2001). An *unc-6(ev400)mig (ev648)* mutant is the ideal strain for this assay. However, *unc-6* and *mig(ev648)* are closely linked on LGX and isolating recombinants would be extremely difficult. The frequency of DTC migration defects in *unc-40(e1430);unc-5(e53);mig(ev648)* mutants was calculated and compared to *unc-40(e1430);unc-5(e53)*. The frequency of DTC migration defects in triple mutants was significantly greater compared to the frequency in *unc-40(e1430);unc-5(e53)* (Table 8). This increase is conclusive evidence confirming the role of *mig-23* in DTC migrations is not limited to the *unc-6/unc-40/unc-5* guidance pathway and that it is acting, at least in part, in a parallel DTC guidance pathway.

3.1.3 Suppression/enhancement of *evIs99* DTC migration defects in a *mig(ev648)* background.

The previous data has indicated *mig(ev648)* is not limited to the *unc-6/unc-5/unc-40* guidance pathway but has not ruled out the possibility that *mig(ev648)* is directly involved in *unc-5/unc-40/unc-6* mediated DTC guidance. To assess whether *mig(ev648)* has a role in *unc-5* mediated DTC guidance, an *evIs99;mig(ev648)* strain (Section 2.3) was generated. The *evIs99* transgenic strain contains an integrated DNA array with the entire *unc-5* coding

sequence regulated by the *emb-9* promoter, resulting in the early expression of *unc-5* during the first ventral DTC migration phase thus causing a precocious dorsalward turn of the DTC (Su *et al* 2000). This aberrant DTC migration pattern is identified at low magnification by a dorsal clear patch in the animal (Figure 6). If a gene directly involved in *unc-5* mediated DTC guidance is placed in an *evIs99* background, the frequency of precocious DTC migration defects is suppressed due the disruption of a component required for *unc-5* mediated guidance. For example, the frequency of precocious DTC turns in *evIs99;unc-6(ev400)* is 0% for both anterior and posterior DTC compared to 66% (anterior) and 75% (posterior) precocious DTC turns in *evIs99* animals (Su *et al* 2000). A gene dispensable for *unc-5* mediated guidance will not suppress DTC migration defects in an *evIs99* background as *unc-5* retains its ability to turn the DTC dorsally. *evIs99;mig(ev648)* was scored for the frequency of precocious DTC turns and compared to *evIs99* alone (Table 9).

Table 9 : DTC migration defects of *evIs99* and *evIs99;mig(ev648)*

	Anterior DTC	Posterior DTC	*n*
evIs99	24 ± 2	40 ± 2	392
evIs99;mig(ev648)	31 ± 4	42 ± 4	108

The statistical comparison is against the frequency of DTC migration defects in the *evIs99* strain alone. **P*<0.05; ***P*<0.001.

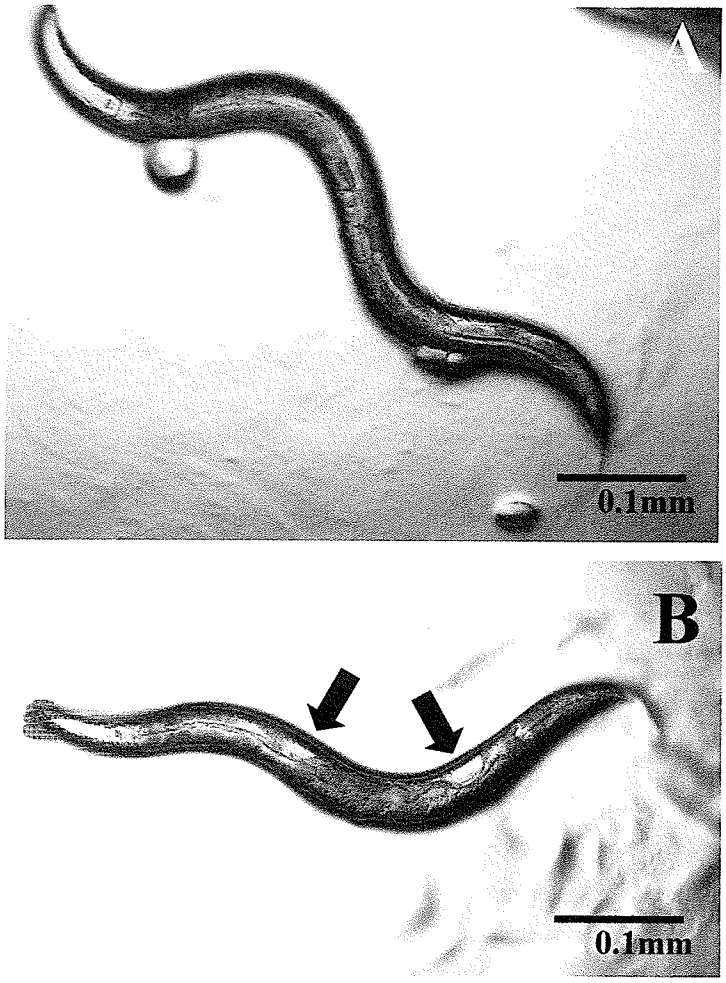


Figure 6: Images of N2 wild-type and *evIs99*. Two clear dorsal patches (denoted by the two arrows) are visible in *evIs99* (B) and are a result of the precocious DTC turn caused by early *unc-5* expression during the first DTC migration phase. These patches are not observed in N2 (A).

The frequency of precocious DTC turns was not significantly changed in *evIs99;mig(ev648)*, suggesting that *mig(ev648)* does not play a role in *unc-5* mediated DTC guidance.

3.1.4 Cosmid phenotype rescue in *mig(ev648)*.

The *mig(ev648)* enhancer allele was mapped to LGX-2.9 between genes *lon-2*(-6.75) and *unc-6*(-2.0), a genetic region with (at the time) no known genes involved in DTC migrations or guidance, (Figure 7) suggesting that *mig(ev648)* was possibly an allele of a novel DTC migration gene. Cloning *mig(ev648)* began by microinjection of wild-type cosmids into *mig(ev648)* mutants and testing whether the cosmids rescued the DTC phenotype (Section 2.7). Overlapping cosmids spanning the region LGX–3.0 to LGX–2.74 were obtained from the CGC and prepared as described in Section 2.7.1. Eight cosmid mixtures (Section 2.7.2) were microinjected separately into individual *mig(ev648)* mutants and eight transgenic lines stably expressing each cosmid (denoted by GFP expression) were generated. Each transgenic line expressed GFP and exhibited the *mig(ev648)* DTC migration defect phenotype indicating that cosmids spanning LGX–3.05 to LGX–2.74 did not rescue the *mig(ev648)* DTC migration defect phenotype. This finding suggests that the *mig(ev648)* enhancer allele maps to an alternate genetic region not covered by these cosmids. The lack of rescue observed could be explained by limitations associated with microinjecting DNA arrays for creation of transgenic *C.elegans* lines. Possible complications with DNA microinjections include gene over-expression due to increased gene copies contained within the transgenic array, variations in gene expression patterns from one animal to the next due to varying amounts of the array being transmitted to progeny and the presence of a gene silencing mechanism in *C.elegans* triggered by tandem sequence repeats

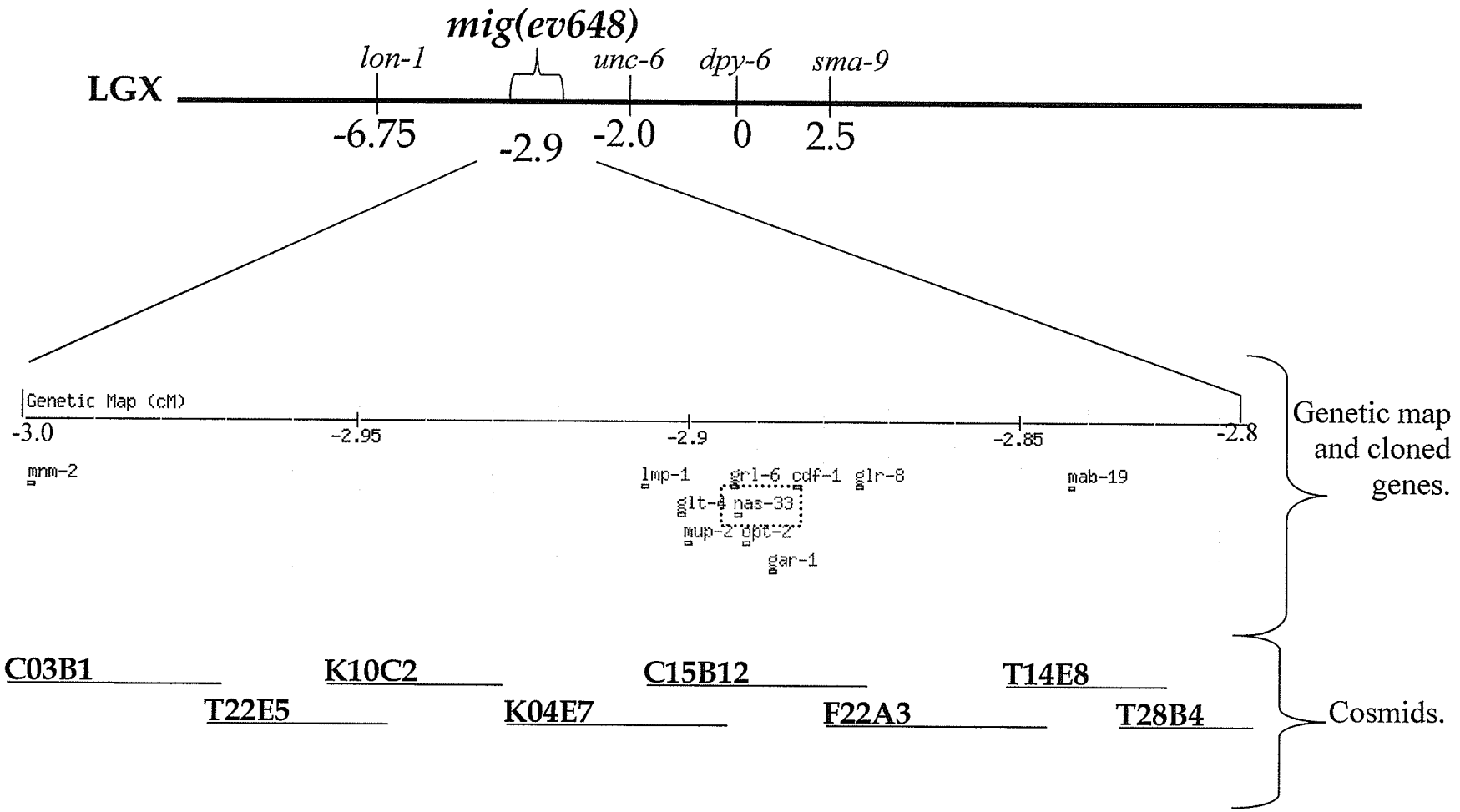


Figure 7 : A summary of the *mig(ev648)* genetic region. A schematic of the cloned genes (www.wormbase.org, 2003) and overlapping cosmids within the genetic region surrounding *mig(ev648)*. Actual cosmid sizes summarized in Section 6.3.1.

that occur within the injected array (Praitis *et al.* 2001). To refine *mig(ev648)* mapping and to confirm the cosmid rescue data obtained was not a result of transformation complications, higher resolution gene mapping using SNPs (single nucleotide polymorphisms) was undertaken.

3.1.5 Sequencing *nas-33*

Genes within the mapped *mig(ev648)* region were evaluated for a potential role in DTC guidance that has not been identified to date. The gene *nas-33* located at LGX-2.9 (Figure 7) encodes an astacin-like protein (Mohrlen, Hutter and Zwilling 2003). Astacins are a family of metalloproteases with various defined roles in pattern formation, morphogenesis and cell migrations (Basbaum and Werb 1996). As roles for metalloproteases *mig-17* (Nishiwaki, Hisamoto and Matsumoto 2000) and *gon-1* (Blelloch *et al.* 1999) in DTC guidance have been previously described, *nas-33* was a putative candidate gene for the *mig(ev648)* enhancer allele.

The *nas-33* gene in was sequenced in N2 wild-type and *mig(ev648)* strains (Section 2.8) and sequencing results were aligned and analysed for any differences (Section 6.4). The N2 wild-type *nas-33* sequence was identical to the *nas-33* gene sequence in the *mig(ev648)* mutant strain confirming that the *mig(ev648)* enhancer is not an allele of the *nas-33* gene.

3.1.6 *mig(ev648)* and *mig-23* complementation test.

A paper was published in January 2004 identifying and describing a novel gene involved in DTC migrations, *mig-23* (Nishiwaki *et al.* 2004). *mig-23* encodes a membrane bound nucleotide diphosphatase (NDPase) involved in glycosylating and localizing MIG-17, a disintegrin and matrix metalloprotease with a role in DTC guidance. Gonad morphology

defects in *mig-23* mutants closely resembled the wandering gonad arm observed in our *tnIs5;mig(ev648)* strain. Although *mig-23* maps to LGX:-4.0 (Nishiwaki *et al.* 2004), the similarities in the gonad morphology defects observed in *mig-23* and *mig(ev648)* strongly suggested that *mig-23* was a good candidate gene.

The *mig-23(k180)* strain (a *mig-23* null allele) was obtained from the CGC (Caenorhabditis Elegans Genetics Centre). At low magnification, gonad morphology defects in the *mig-23(k180)* strain closely resembled those observed in *mig(ev648)* (Figure 8). To assess whether *mig(ev648)* was an allele of *mig-23* a complementation test was performed by crossing *mig(ev648)* males to *mig-23(k180)* hermaphrodites. If F1 progeny from this cross were phenotypically wild-type, *mig(ev648)* and *mig-23(k180)* complement each other and thus are not the same gene. On the contrary, if F1 progeny from the cross exhibited identical gonad morphology defects observed in *mig(ev648)*, *mig(ev648)* and *mig-23(k180)* do not complement each other and thus are the same gene.

Five *mig(ev648)* males were mated with one *mig-23(k180)* hermaphrodite. To ensure mating occurred and F1 progeny were cross progeny and not *mig-23(k180)* hermaphrodite self-progeny, a high presence of males in the F1 progeny was used as an indication that *mig(ev648)* males mated with the *mig-23(k180)* hermaphrodite. In each of the four arranged crosses, a strong presence of F1 males was observed and F1 hermaphrodites exhibited a high frequency of the identical gonad morphology defects observed in both *mig(ev648)* and *mig-23(k180)* mutants. Thus *mig-23(k180)* and *mig(ev648)* do not complement each other indicating that *mig(ev648)* is indeed an allele of the gene *mig-23*.

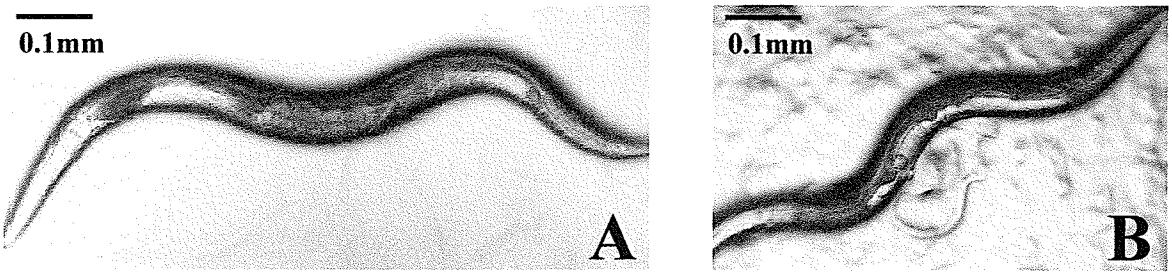


Figure 8: Images of *mig-23(k180)* and *mig(ev648)*. Gonad morphology defects in *mig-23(k180)* (A) and *mig(ev648)* (B) mutants bear significant similarities.

3.1.7 Sequencing *mig-23.*

To confirm that enhancer *mig(ev648)* is an allele of *mig-23*, *mig-23* was sequenced in both N2 wild-type and *mig(ev648)* strains (Section 2.8) and sequencing results were aligned and analysed for differences (Section 6.5). The *mig-23* sequence in *mig(ev648)* mutants differed from the wild-type *mig-23* sequence at base pair 335 in exon 3 of the *mig-23* coding sequence resulting in a GC-AT transition. The mutation causes an amino acid substitution at position 112 in the protein of an alanine for a valine (Figure 9). MIG-23, a member of the apyrase protein family has 5 apyrase conserved regions and trans-membrane domains at both the carboxy and amino termini (Nishiwaki *et al.* 2004). In *mig-23(ev648)* mutants the alanine to valine substitution occurred in a region of the protein that is relatively conserved but not within an apyrase conserved region described by (Nishiwaki *et al.* 2004) (Figure 10).

Assays used to genetically map *mig(ev648)* suggested that the enhancer allele was temperature sensitive. For example, the *unc-130(ev505)* allele is a temperature sensitive null allele, as DTC migration defect frequencies in *unc-130(ev505)* grown at 25°C exhibit more than double the percentage of DTC migration defects than *unc-130(ev505)* mutants grown at 16 °C in the posterior and more than triple the percentage in the anterior (Nash *et al.* 2000). *mig(ev648)* mutants were grown at 16°C, 20°C and 25°C and the frequency of DTC migration defects was determined for each temperature (Section 2.5.2). *mig(ev648)* mutants grown at 25°C exhibited a significant increase in the frequency of posterior and anterior DTC migration defects when compared to *mig(ev648)* mutants grown at 20°C and 16°C (Table 10), suggesting the *mig(ev648)* allele is temperature sensitive.

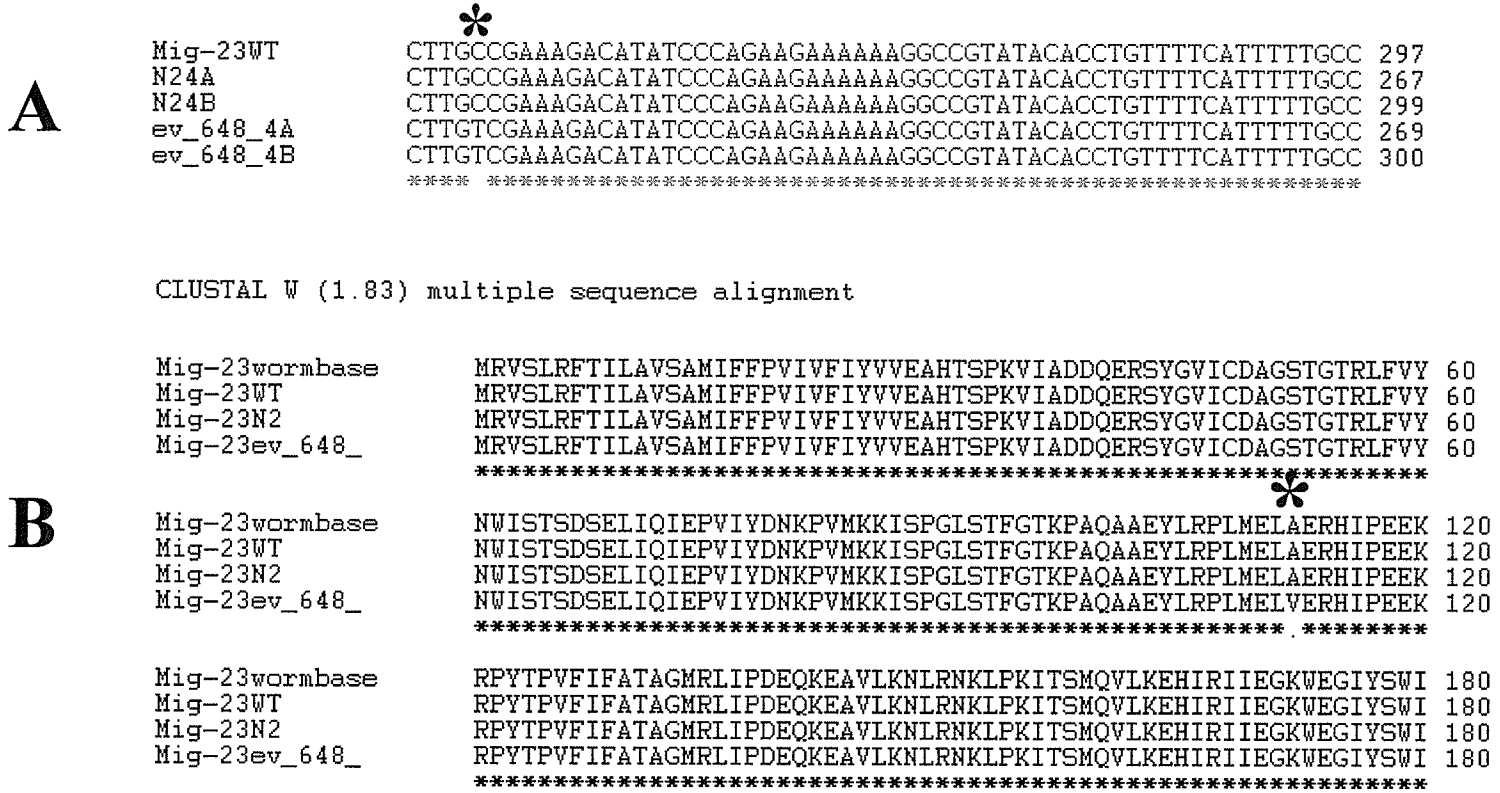


Figure 9: A summary of *mig-23* gene and MIG-23 protein sequence alignments. A)The region taken from the fourth *mig-23* DNA fragment sequenced demonstrating the CG-TA transition. (For entire sequence alignment, see Section 6.5). *mig-23* wt is the genetic *mig-23* sequence obtained from www.wormbase.org. DNA sequenced from wild-type animals is denoted as N2, DNA sequenced from *mig(ev648)* mutants is denoted as ev_648 and for each, A represents sequence results from forward primers and B represents sequence results from reverse primers. B) A region taken from the MIG-23 protein alignments demonstrating the alanine to valine amino acid substitution at position 112 in *mig-23(ev648)* mutants. (For entire sequence alignment, see Section 6.6). Mig-23wormbase: MIG-23 protein sequence from www.wormbase.org, Mig-23WT: translated *mig-23* wormbase gene sequence, Mig-23N2: translated *mig-23* gene sequence obtained from the N2 strain, Mig-23ev_648_: translated *mig-23* gene sequence obtained from the *mig(ev648)* strain.

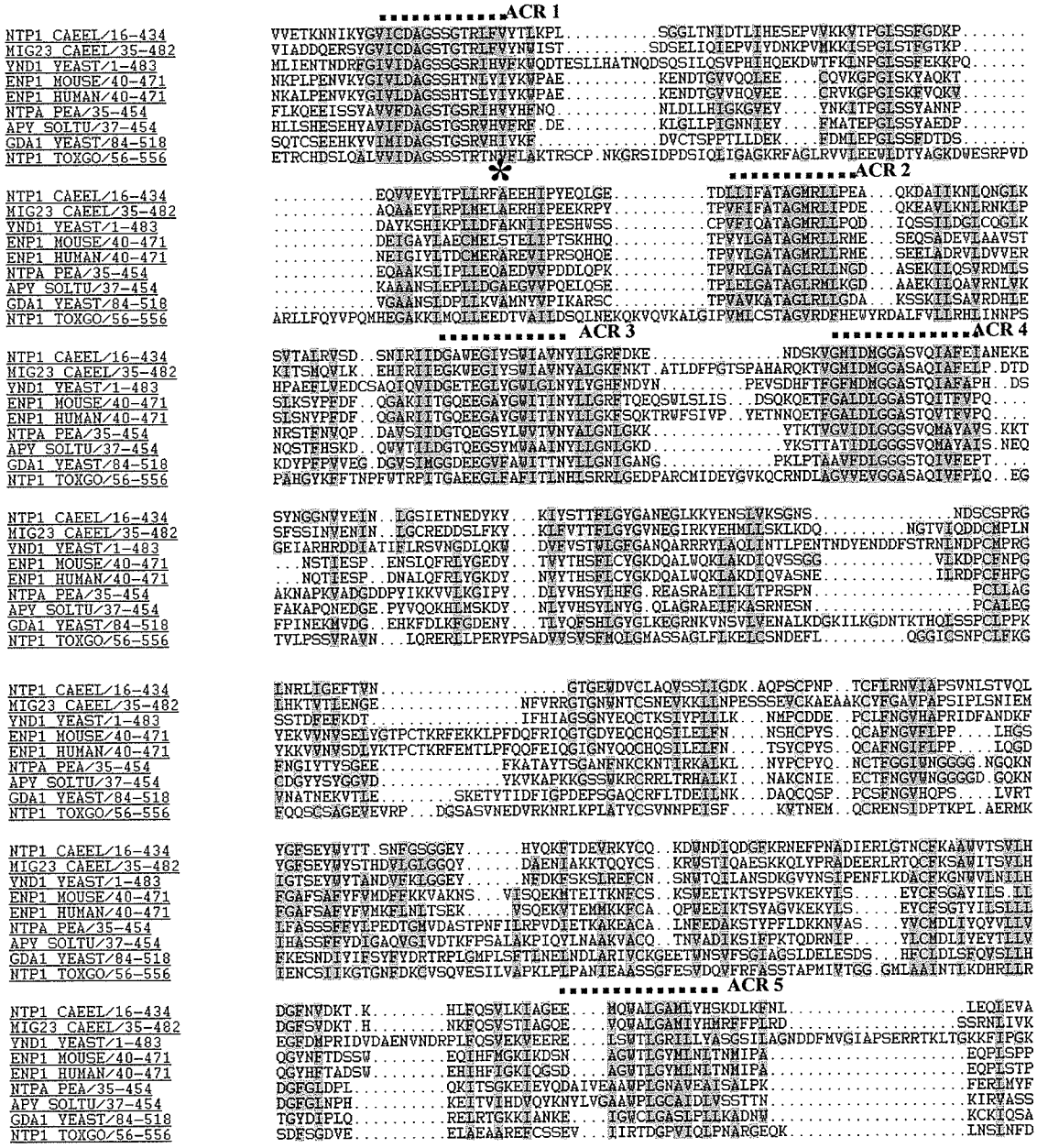


Figure 10: Cross species MIG-23/NDPase protein sequence alignments. A multiple alignment of NDPase/MIG-23 related proteins obtained from the Pfam database. The A112V mutation encoded by *mig-23(ev648)* is within a relatively conserved domain of the protein (*) although not with one of the five **A**pyrase **C**onserved **R**egions (ACR) defined by Nishiwaki et al. 2004 (-----).

Table 10 : DTC migration defects of *mig(ev648)* at 16°C, 20°C and 25°C.

	Anterior DTC	Posterior DTC	*n*
mig(ev648) @ 16°C	30 ± 2*[a]	58 ± 2	475
mig(ev648) @ 20°C	36 ± 2	54 ± 2	400
mig(ev648) @ 25°C	53 ± 3**	63 ± 3*	279

[a] Each statistical comparison is against the frequency of DTC migration defects in *mig(ev648)* grown at 20°C. *$P<0.05$; **$P<0.001$.

MIG-23 functions in the muscle cells of the body wall and is essential for the glycosylation and localization of MIG-17 for DTC guidance during the second and third DTC migration phase. Genetic interactions between *mig-17* and *mig-23* demonstrated *mig-23* has *mig-17* independent DTC guidance roles (Nishiwaki *et al.* 2004). Nishiwaki *et al.* (2004) have shown *mig-23(k180)* is a DTC migration gene directly interacting with *mig-17* for DTC guidance. However *mig-17* enhances DTC migration defects in an *unc-6(ev400)* background (Nishiwaki, Hisamoto and Matsumoto 2000), suggesting *mig-23/mig-17* guidance mechanisms are *unc-6* independent. We have confirmed the role of *mig-23(ev648)* in DTC guidance is *unc-5* independent and have demonstrated that *mig(ev648)* is a temperature sensitive allele.

3.2 Cloning and characterizing the DTC guidance roles of *enh(ev697)*.

The *enh(ev697)* allele was identified in the genetic screen for enhancers of DTC migration defects indicating that the allele has a role in DTC guidance. However, *enh(ev697)* mutants do not exhibit DTC migration defects unless they are in an *unc-5(e152)* background suggesting the role of *enh(ev697)* in DTC guidance can be compensated by additional guidance mechanisms. The *enh(ev697)* mutant strain exhibits a low frequency of embryonic elongation defects (Figure 3) and in a heterozygous state (*enh(ev697)/+*), *enh(ev697)*

enhances DTC migration defects in *unc-5(e152)* mutants, suggesting this allele may cause a dominant effect or a gain of function which in turn results in the enhancement of DTC migration defects in *unc-5(e152).*

3.2.1 High resolution SNP mapping

enh(ev697) was mapped by three factor mapping to LGX between LGX 1.88 and LGX 2.86, a genetic region covered by over 40 cosmids and encompassing over 20 genes. In order to refine the genetic mapping of *enh(ev697)* within this region and narrow down the number of candidate cosmids required for phenotype rescue, *enh(ev697)* mapping using SNPs(single nucleotide polymorphisms) was outlined. An alternate *C.elegans* strain Hawaiian, with almost complete genome sequence similarities to the commonly used *C.elegans* Bristol strain bears SNPs that are mapped throughout the entire Bristol genome. The SNPs of each strain can be distinguished by PCR amplification and restriction enzyme digest patterns and utilized to map recombination events between a Bristol chromatid and a Hawaiian chromatid that span a genomic region flanked by two phenotype markers for recombinant animal identification. If a gene of interest is situated between these markers the gene can be mapped relative to the position of the recombination (identified using SNPs) between a Bristol and Hawaiian chromatid (Figure 11).

Genes *dpy-6(e14)*(LGX:0.0) and *egl-15(n484)*(LGX:2.86) flanking the region to which *enh(ev697)* maps were chosen as markers for recombinant animal identification as their phenotypes (Dpy; reduced body length and Egl; egg-laying defective and bloated) are easily distinguishable. Within the genetic region flanked by these markers, three verified SNPs were identified (Figure 12) and are described in Section 2.10.2. A *dpy-6(e14)enh(ev697)egl-15(n484)* strain was generated as outlined in Section 6.2.3

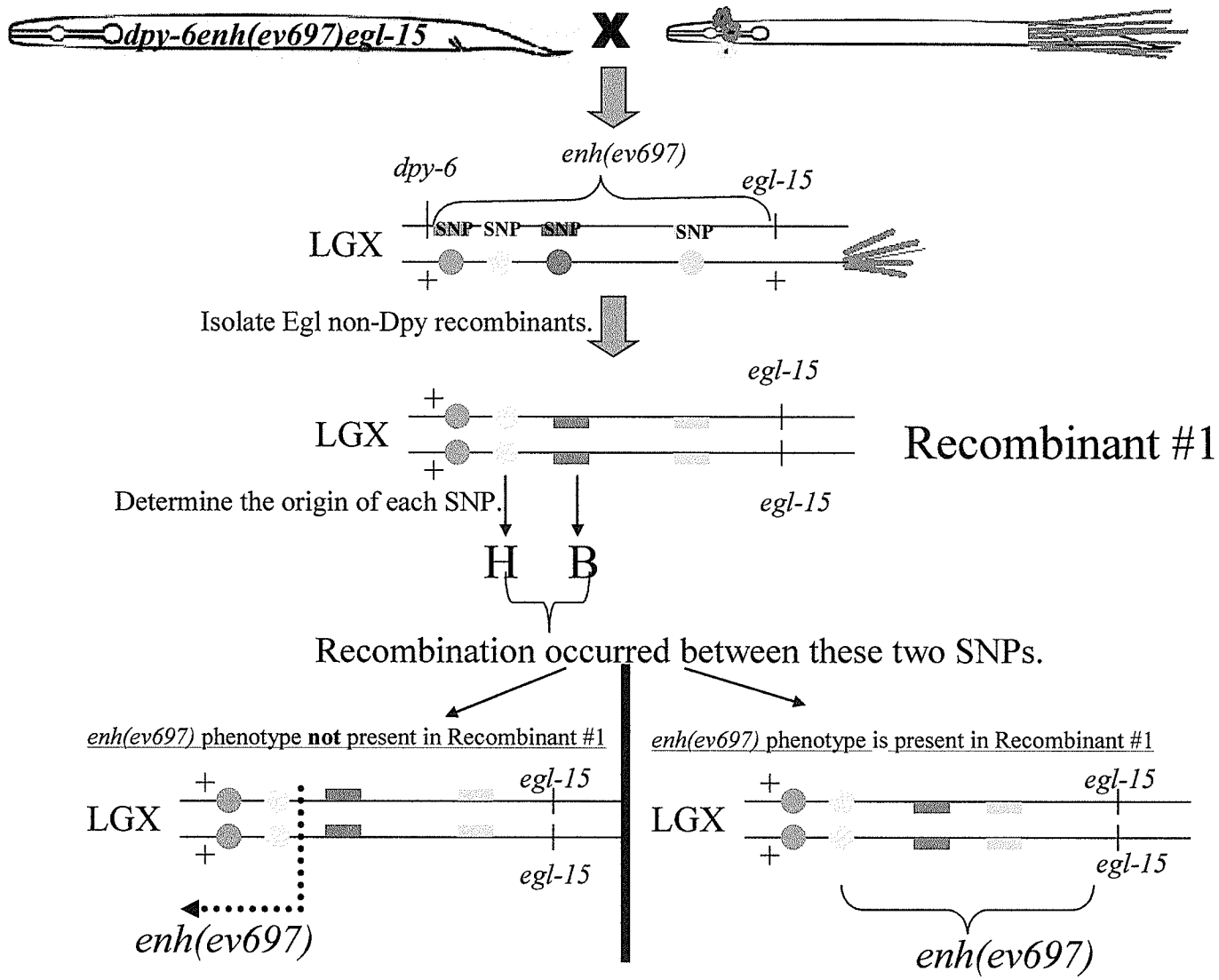


Figure 11: SNP mapping outline.

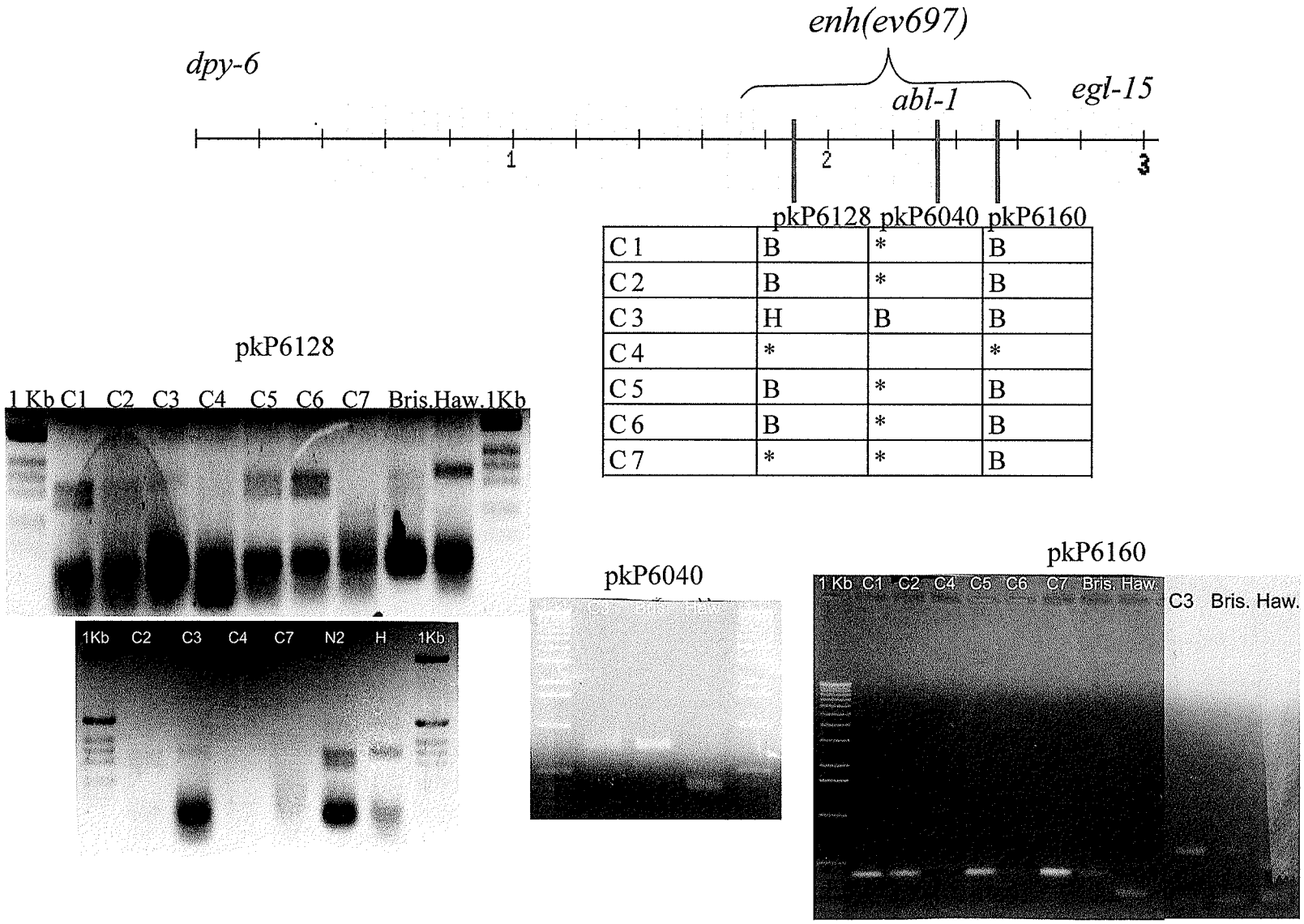


Figure 12: SNP mapping results.

and *dpy-6(e14)enh(ev697)egl-15(n484)* hermaphrodites were crossed with Hawaiian males. Seven Egl non-Dpy recombinants were identified and labelled C1-C7 (Recombinant Clones). DNA was prepared from each recombinant strain (Section 2.10.4) and all three SNP fragments were amplified in each recombinant (Section 2.10.5). In order to identify the origin of each SNP, fragments were digested with their respective restriction enzymes (Table 7) and resolved on an agarose gel (Section 2.10.6).

The first SNP analysed in the recombinants was pkP6160. From the restriction digest patterns resolved on the pKP6160 gel in Figure 12, the pkP6160 SNP in all recombinants was of Bristol origin indicating that the recombination in each strain occurred to the left of this SNP. Analysis of the pkP6128 SNP followed as it was the furthest to the left and would determine whether or not the recombination took place before or within the group of SNPs. The restriction enzyme digest patterns on the pkP6128 gel in Figure 12 indicated the only recombinant line with a Hawaiian SNP at this loci was C3. As the recombination in C1, C2, C5 and C6 occurred to the left of pkP6128 and amplifying DNA from the C4 and C7 lines was problematic, C1, C2, C4, C5, C6 and C7 were put aside for the time being and the C3 recombinant was used for the remaining SNP mapping analysis.

Examining egg-laying defective and bloated recombinants for the presence of the *enh(ev697)* embryonic elongation defect was challenging. The *enh(ev697)* phenotype is present at a low penetrance and Egl animals not only produce lower brood sizes but also form sacs of eggs and young larvae that mask the *enh(ev697)* phenotype. As the recombination in C3 was clearly mapped and I had confidently confirmed the presence of *enh(ev697)* in C3, SNP mapping continued in this recombinant strain. The final SNP, pkP6040 was analysed in the C3 recombinant to further delineate the region between the SNPs pkP6128 and pkP6160 in which the recombination occurred. The final pkP6040 agarose gel in Figure 12

demonstrated the pkP6040 SNP in C3 was of Bristol origin. Thus the recombination happened between 1.932 and 2.38 on LGX in C3. Together these data confirm the *enh(ev697)* allele is situated between pkP6128 at 1.932 and *egl-15* at 2.86. The genetic area flanked by these markers contains no additional verified SNPs and thus the *dpy-6(e14);enh(ev697);egl-15(e484)* strain was once again mated to Hawaiian males in order to isolate additional recombinants and continue mapping. As this was underway, the genetic region between 1.932 and 2.86 was analysed for candidate genes with potential roles in DTC guidance.

3.2.2 *abl-1* RNAi in *enh(ev697)* mutants.

The gene *abl-1* maps to LGX+2.35 and was chosen as a possible candidate gene for the *enh(ev697)* allele. ABL-1 is a non-receptor tyrosine kinase with SH2 and SH3 domains and has been shown to play a role in axon guidance (Wills *et al.* 1999). In Drosophila, Abl (Abelson) mediates downstream signalling of the ROBO axon guidance receptor in combination with Ena,(Enabled, Abl substrate) an actin cytoskeletal regulator for cell migrations (Bashaw *et al.* 2000). In a screen for suppressors of ectopically expressed *unc-5, unc-34* (a *C.elegans* Ena homologue) suppressed a dorsal reorientation of axon growth cones resulting from ectopically expressed *unc-5* suggesting *unc-34* plays a role in UNC-5 mediated axon guidance (Colavita and Culotti 1998b). Thus it is plausible that *abl-1* is working with *unc-34* to mediate UNC-5 responses during cell guidance.

An RNAi vector containing the *C.elegans abl-1* double stranded gene fragment transformed into the appropriate *E.coli* strain was readily available and was kindly provided by P.Roy at the University of Toronto. Cultures and media for the RNAi experiments were prepared as described in Section 2.11. Wild-type N2 animals and *enh(ev697)* animals were

picked at the L3 stage and cloned separately onto plates seeded with *E.Coli* expressing *abl-1* dsRNA. Animals were left to lay their eggs and F1 progeny were examined at the L4 stage for the presence of the *enh(ev697)* embryonic elongation defect. If *enh(ev697)* is a loss of function allele of *abl-1*, it would be expected that RNAi knockdown of *abl-1* in N2 animals would induce the embryonic elongation defect observed in *enh(ev697).* However, N2 progeny cloned onto the RNAi plates did not exhibit this phenotype. As our previous data suggests *enh(ev697)* causes some type of dominant effect, *enh(ev697)* animals were cloned onto the RNAi plates to determine whether *abl-1* RNAi knockdown could rescue the wild-type phenotype in these mutants. If *enh(ev697)* is an allele of *abl-1* and is causing a gain of function mutation, silencing *abl-1* expression thus reducing the dominant effect of the allele should result in the elimination of the embryonic elongation defect phenotype. This was not the case as *enh(ev697)* animals grown on *abl-1* dsRNA still exhibited the embryonic elongation defect phenotype.

As *enh(ev697)* is suspected to have a role in DTC guidance, we assessed the outcome of *abl-1* knockdown on the role of *enh(ev697)* in enhancing DTC migration defects. *unc-5(e152);enh(ev697)* animals were grown on *E.coli* expressing *abl-1* dsRNA and their DTC migration defects scored. If *enh(ev697)* is an allele of *abl-1*, then *abl-1* knockdown should suppress the enhancement of DTC migration defects observed in *unc-5(e152);enh(ev697)* mutants. Data summarized in Table 11 confirms *abl-1* RNAi knockdown does not suppress the frequency of DTC migration defects in *unc-5(e152);enh(ev697).*

Table 11 : DTC migration defects of *unc-5(e152);enh(ev697)* fed with *abl-1* RNAi.

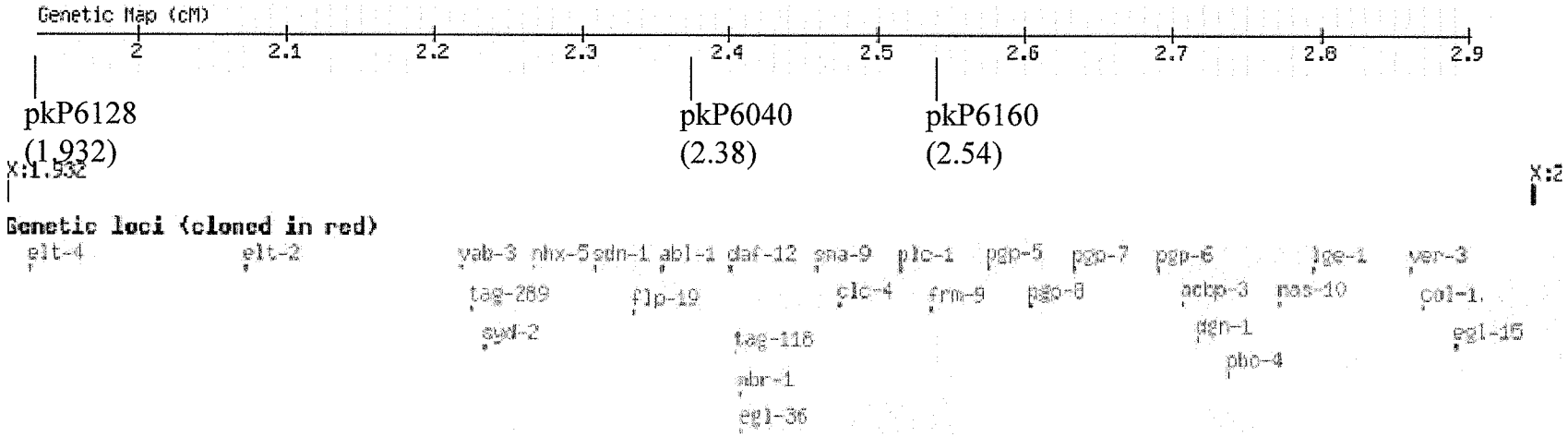
	Anterior DTC	Posterior DTC	*n*
unc-5(e152);enh(ev697)	22 ± 2	64 ± 2	713
unc-5(e152);enh(ev697). abl-1 RNAi	23 ± 2	65 ± 3	522

Taken together the *abl-1* RNAi assays indicate *enh(ev697)* is not an allele of the *abl-1* gene.

3.2.3 Cosmid phenotype rescue in *enh(ev697).*

SNP mapping refined *enh(ev697)* to the area between 1.932 and 2.86. Cosmids described in Section 6.3.2 spanning this region (Figure 13) were obtained and prepared as described (Section 2.7.1). As previous data suggests *enh(ev697)* could be a gain of function allele the cosmid rescue strategy used for cosmid phenotype rescue in the *mig-23(ev648)* strain was altered accordingly for cosmid rescue in *enh(ev697).* If the *enh(ev697)* allele is causing a gain of function, microinjecting the wild-type version of the allele into *enh(ev697)* mutants will likely not inhibit the dominant effect resulting in continued occurrence of the embryonic elongation phenotype in *enh(ev697)* mutants. If *enh(ev697)* were a loss of function allele, microinjecting the wild-type version of the allele should rescue the wild-type phenotype in *enh(ev697)* mutants. Thus each cosmid mixture in addition to being microinjected into *enh(ev697)* mutants, was microinjected into wild-type N2 mutants to assay whether increasing the gene copy and expression of the *enh(ev697)* allele could induce the *enh(ev697)* phenotype in a wild-type background. As embryonic elongation defects in *enh(ev697)* are not fully penetrant, cosmid phenotype rescue was deduced by a lack of GFP expressing *enh(ev697)* "commas" on the plate, as those "commas" appearing in a strain with cosmid rescue would be those that have lost the array and are no longer being rescued. On the contrary, stable injected GFP lines exhibiting *enh(ev697)* "commas" with GFP indicate that the cosmid does not rescue.

Several cloned genes are located centrally in the genomic region delineated by SNP mapping (Figure 13). As *abl-1* was a candidate gene for *enh(ev697),* cosmid phenotype



F41E7 **F46F6** **C39B10** **F14F3** **F57C7** **F11A1** **C46B5** **R07A4**

R07E3 **ZC504** **C33D3** **F59F5** **M79** **F13E6** **T01C1**

Figure 13: A summary of the *enh(ev697)* genetic region. A schematic representation of the genetic region between pK6128 and *egl-15* markers to which *enh(ev697)* maps. Cloned genes are shown in addition to the cosmids spanning the region selected for phenotype rescue assays.

rescue assays began with the F57C7 cosmid. Four stable N2 and four stable *enh(ev697)* transgenic lines each expressing the GFP and cosmids F57C7, M79, F11A1 and F59F5 were generated and analysed. N2 and *enh(ev697)* lines expressing M79, F11A1 and F59F5 appeared phenotypically normal. The presence of the *enh(ev697)* embryonic elongation defect was not detected in each of the GFP expressing N2 wild-type lines and GFP expression was detected in the "commas" of the *enh(ev697)* lines, indicating that M79, F11A1 and F59F5 do not rescue *enh(ev697)* and thus do not carry the gene of the *enh(ev697)* allele. However, this was not the case for the F57C7 cosmid. Cosmid phenotype rescue of *enh(ev697)* began by microinjection of each M79 and F57C7 cosmid mixtures with identical cosmid concentrations into *enh(ev697)* mutants taken from the same plate. When the F1 progeny from the hermaphrodites microinjected with the M79 cosmid were verified for GFP expression, more than ten GFP expressing progeny were cloned out to isolate stable transgenic lines. The hermaphrodites microinjected with the F57C7 cosmid mixture only produced one single transgenic GFP F1 progeny which itself, produced a very small amount of progeny, although one of those was an *enh(ev697)* GFP "comma". As microinjections of the F11A1 and F59F5 cosmids were generating many transgenic animals, a new cosmid mixture with the F57C7 cosmid concentration reduced from 10ng/μl to 5ng/μl was prepared. The mixture with a reduced F57C7 cosmid concentration was microinjected into *enh(ev697)* animals and F1 GFP eggs were observed the following day. However, GFP expressing larvae hatched from the eggs did not survive to produce any progeny. A new cosmid mixture was prepared with the F57C7 cosmid concentration reduced to 0.5ng/μl and microinjected into *enh(ev697)* animals. Reducing the F57C7 cosmid concentration resulted in the generation of transgenic GFP F1 progeny that were cloned out to isolate stable transgenic lines expressing F57C7. Interestingly, progeny from the transgenic, GFP expressing F1s

exhibited an observable increase in the frequency of the embryonic elongation defects expressing the GFP. Thus for each transgenic line, the F57C7 cosmid was being expressed (denoted by the GFP expression), was toxic in higher concentrations and in lower doses affected the appearance of *enh(ev697)* embryonic elongation defects. Together these results strongly suggest F57C7 contains the gene of the *enh(ev697)* allele but could not be complemented by phenotype rescue. The genetic area covered by F57C7 was analysed for candidate genes while strategies for further characterizing the effects of the F57C7 cosmid, including quantification of the elongation defect in transgenic animals and microinjection into *unc-5(e152)* mutants were outlined.

3.2.4 Sequencing *sdn-1*.

Results from the microinjections of the F57C7 cosmid in 3.2.3 suggested that F57C7 contains the gene of the *enh(ev697)* allele. The F57C7 cosmid spans a genetic region that includes genes *nhx-5* and *sdn-1*. *nhx-5* encodes a sodium/proton exchanger and mutants exhibit no obvious phenotype (www.wormbase.org) and *sdn-1* encodes a heparan sulfate proteoglycan and mutants are slightly egg-laying defect (Minniti *et al.* 2004). Although the reported phenotype of *sdn-1* mutants does not resemble the phenotype of *enh(ev697)*, a role for proteoglycans in DTC guidance has previously been characterized (Merz *et al.* 2003) and thus *sdn-1* was chosen as a candidate gene for sequencing.

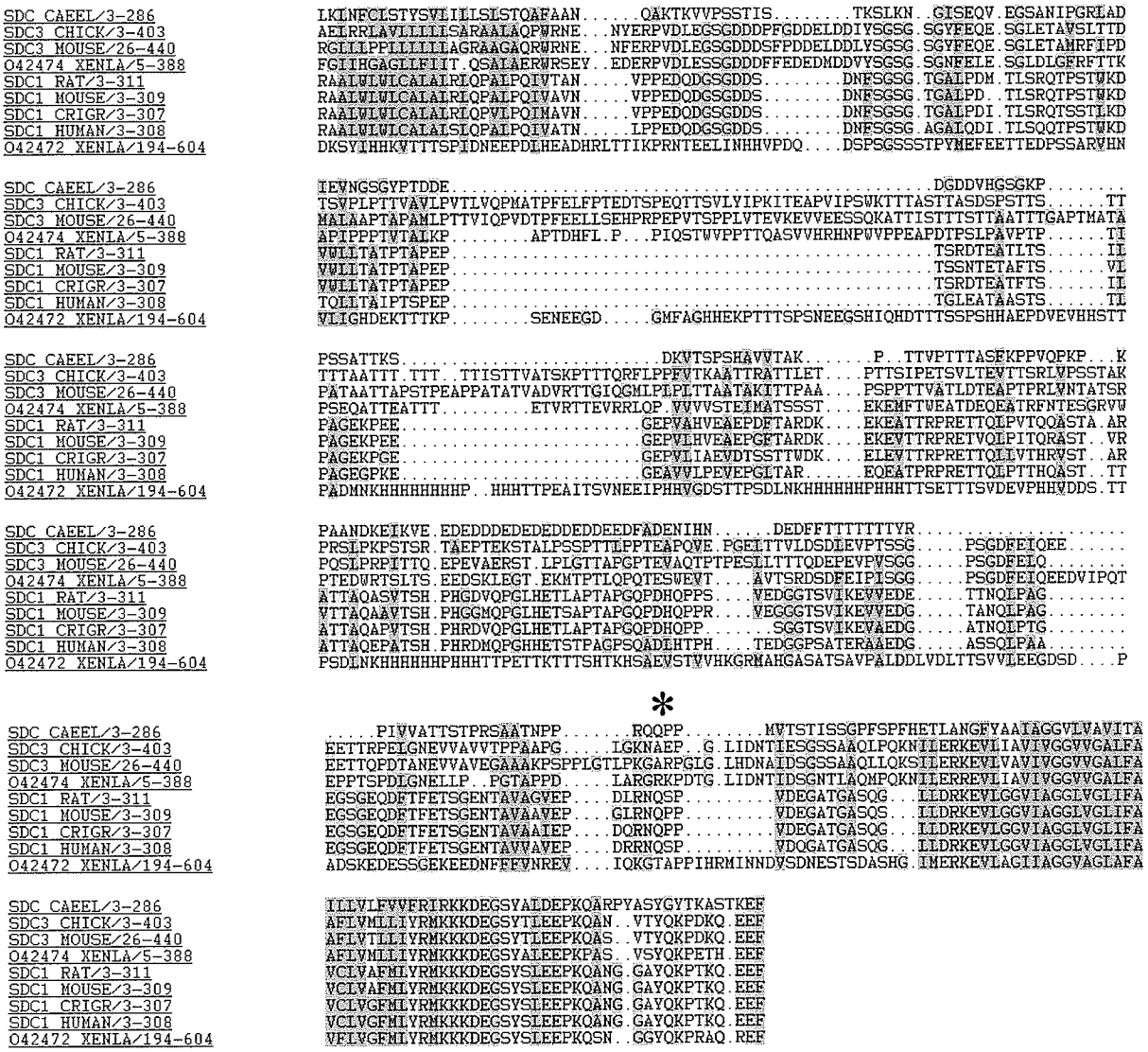
The *sdn-1* gene was sequenced in both N2 wild-type and *enh(ev697)* strains (Section 2.8) and sequencing results were aligned and analysed for differences (Section 6.7). The *sdn-1* sequence in *enh(ev697)* mutants differed from the wild-type *sdn-1* sequence at base pair 610 in exon 5 of the *sdn-1* coding sequence resulting in a GC-AT transition and causing a premature stop codon that truncates the protein at amino acid 203 (Figure 14). The *sdn-1*

gene sequencing results in Figure 14 also demonstrated that our N2 wild-type strain had an apparent polymorphism at the same base pair position of the *sdn-1* sequence in which the *enh(ev697)* mutation occurs. Although the gene was not sequenced, it was presumed the *sdn-1* sequence in the *unc-5(e152)* strain used for the genetic screen, in which the initial *enh(ev697)* mutation would have occurred, matched the wild-type sequence in Wormbase. This result, together with the mapping data confirms *enh(ev697)* is an allele of the *sdn-1* gene, *sdn-1(ev697)*.

SDN-1 is a trans-membrane proteoglycan whose ectodomain consists of conserved serine residues that connect modifiable heparan sulfate side-chains (Minniti *et al.* 2004) and a highly conserved short cytodomain containing a PDZ-binding motif (Rhiner *et al.* 2005). An alignment between SDN-1 and syndecans proteins in other species identifies a highly conserved endodomain of SDN-1 and a highly conserved trans-membrane domain (Figure 15). As the *ev697* allele encodes a stop codon that precedes the trans-membrane domain, *sdn-1(ev697)* mutants may be expressing an unbound form of SDN-1.

Two *sdn-1* alleles have been identified to date, *ok449* (Minniti *et al.* 2004) and *zh20* (Rhiner *et al.* 2005). The *ok449* allele encodes an SDN-1 protein with a deletion from amino acid 52 to 120 abolishing two glycosaminoglycan attachment sites. The *zh20* allele encodes a deletion of exons 1-5 resulting in a presumed null. Both *zh20* and *ok449* mutants exhibit egg-laying deficiencies in addition to neural cell and axon migration defects. Specifically DD/VD commissures fail to reach the dorsal nerve cord or inappropriately branch out and PQV axons exhibit aberrant midline crossings. The role of *sdn-1* in axon guidance is cell-autonomous and different modifications on the SDN-1 GAG chains are associated with SDN-1 guidance mechanisms of the different axon types (Rhiner *et al.* 2005).

Figure 14: A summary of the *sdn-1* sequence and SDN-1 protein alignments. A)The regions taken from the third and fourth *sdn-1* DNA fragments sequenced demonstrating the GC-TA transition*. (For entire sequence alignments, see Appendix 6.7). Sdn3/Sdn4WT is the *sdn-1* coding sequence obtained from www.wormbase.org. DNA sequenced from wild-type animals is denoted as N2, DNA sequenced from *enh(ev697)* mutants is denoted as 5E68 and for each, A represents the sequence results from forward primers and B represents sequence results from reverse primers. B) A region taken from the SDN-1 protein alignment demonstrating the premature truncation (-) at position 203 in *sdn-1(ev697)* mutants. (For entire sequence alignment, see Appendix 6.8).Sdnwormbase: SDN-1 protein sequence from Wormbase, SdnWt: translated *sdn-1* wormbase gene sequence, SdnN: translated *sdn-1* gene sequence obtained from the N2 strain, SdnEv: translated *sdn-1* gene sequence obtained from the *sdn-1(ev697)* strain.



The coloured markup was created by Jalview (Michele Clamp)

Alignments are coloured using the ClustalX scheme in Jalview (orange:glycine (G); yellow: Proline (P); blue: small and hydrophobic amino-acids (A, V, L, W); green: hydroxyl and amine amino-acids (S, T, N, Q); red: charged amino-acids (D, E, R, K); cyan: histidine (H) and tyrosine(Y)).

Figure 15 : Cross-species syndecan protein alignments. Obtained from the Pfam database, alignment compares the syndecan protein sequences across various species (C.elegans SDN-1is the first sequence). The N terminal at the beginning of the sequence (ectodomain) and the C terminus is at the end of the sequences (endodomain). * denotes where the truncation occurs in *sdn-1(ev697)*. The trans-membrane domain is located at the end of the fifth sequence row, denoted by the blue, hydrophobic amino acids. Note the highly conserved cytoplasmic domain at the terminal end of the sequence.

A GFP tagged SDN-1 localizes to the ventral nerve cord motor neurons, the nerve ring, the vulva and the hypodermis.

The *sdn-1(zh20)* strain was a kind gift from C. Rhiner. Upon phenotype analysis at low magnification the embryonic elongation defects observed in *enh(ev697)* were apparent in *sdn-1(zh20)* although at a much lower frequency (Figure 16). A complementation test between the alleles was attempted but not achieved due to low mating efficiencies of males in both *ev697* and *zh20* strains.

3.2.5 Genetic interactions of *enh(ev697)*, *unc-5*,*unc-40* and *unc-6*.

In our genetic screen for enhancers of DTC migration defects, *sdn-1(ev697)* was identified for its ability to enhance the frequency of DTC migration defects in *unc-5(e152)* mutants indicating that *sdn-1* has a role in guiding DTC migrations and suggesting the possibility that *sdn-1* may be involved in *unc-5* mediated DTC guidance. Classical genetic methods were employed to elucidate the genetic interactions of *sdn-1* with *unc-5/unc-6/unc-40* for DTC guidance (Section 2.9).

An *unc-5(e152);sdn-1(zh20)* mutant was generated as outlined in Section 6.2.4 and the DTC migration defects scored. As summarized in Table 12, *sdn-1(zh20)* enhanced the frequency of DTC migration defects in *unc-5(e152)* mutants.

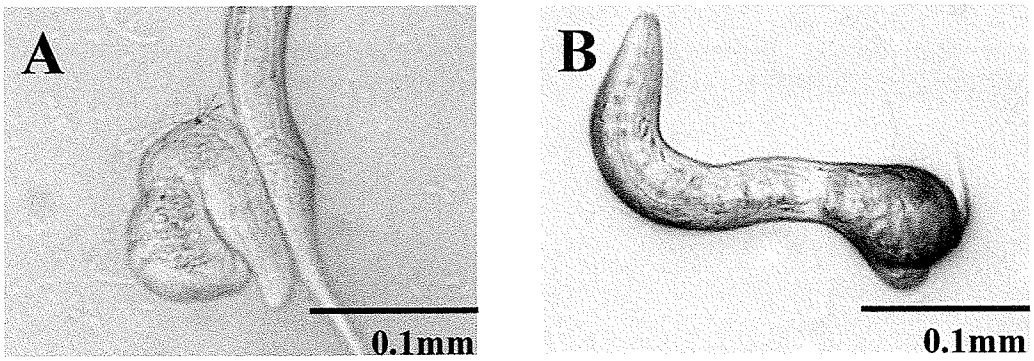


Figure 16: Images of embryonic elongation defects in *sdn-1(ev697)* (A) and *sdn-1(zh20)* (B) mutants.

Table 12 : DTC migration defects of *unc-5, unc-40, unc-6* and *sdn-1*.

unc-5	*sdn-1*	Ant. DTC	Post. DTC	*n*
WT	*ev697*	0	0	200
WT	*zh20*	0	0	150
e152	WT	8 ± 1	40 ± 1	1464
e152	*ev697*	20 ± 1**[a]	63 ± 2**	476
e152	*zh20*	36 ± 2**	77 ± 1**	249
dm11	WT	0.2 ± 0.1	3 ± 0.5	948
dm11	*ev697*	2 ± 0.5**	15 ± 1**	708
dm11	*zh20*	0.5 ± 0.2	8 ± 1**	847
e53	WT	28 ± 1	53 ± 2	951
e53	*ev697*	25 ± 3	56 ± 3	186
e53	*zh20*	23 ± 2*	60 ± 3*	321
unc-40(e1430)	WT	5 ± 1	24 ± 1	833
unc-40(e1430)	*zh20*	3 ± 1	40 ± 3**	339
unc-6(ev400)	WT	34 ± 2	68 ± 2	434
unc-6(ev400)	*ev697*	41 ± 3*	71 ± 3	299

[a] Each statistical comparison is against the frequency of DTC migration defects in *unc-5, unc-40* or *unc-6* strain alone. *P<0.05; **P<0.001.

In addition, both *ev697* and *zh20* enhanced the frequency of DTC migration defects in *unc-5(dm11)* mutants (a very weak *unc-5* allele) (Table 12) clearly indicating that *sdn-1* has a role in DTC guidance.

To determine whether the role of *sdn-1* in DTC guidance is limited to *unc-5* mediated guidance, *unc-5(e53);sdn-1(ev697)* and *unc-5(e53);sdn-1(zh20)* strains were generated (Section 6.2.4) and DTC migration defects scored. As previously described, the *e53* allele represents a complete loss of *unc-5* function in the DTC and thus an enhancement in the frequency of DTC migration defects in *unc-5(e53);sdn-1(ev697)* or *unc-5(e53);sdn-1(zh20)*

suggests that the role *sdn-1* in DTC guidance is not limited to *unc-5,* and the enhancement accounts for the additional role of *sdn-1.* In the *unc-5(e53);sdn-1(ev697)* mutants, the frequency of anterior and posterior DTC migration defects did not increase when compared to the frequency observed in *unc-5(e53)* (Table 12) and only a small increase was observed in *unc-5(e53);sdn-1(ev697).* These results suggest the role of *sdn-1* for DTC guidance along the ventral/dorsal axis is limited mainly to the UNC-5 mediated DTC guidance pathway.

As a principle goal of the genetic screen was to identify genes involved with *unc-5* for DTC guidance, these results prompted further analysis of the role of *sdn-1* not only in DTC guidance but in the UNC-6/UNC-40/UNC-5 guidance pathway. Thus genetic interactions between *sdn-1* the DTC guidance receptor gene *unc-40* were analysed to identify whether the role of *sdn-1* in DTC guidance involves *unc-40.* As with *unc-5* this entailed putting each *sdn-1* allele into an *unc-40* null (*e1430*) background and comparing the DTC migration defect frequencies of both double mutants to *unc-40(e1430)* mutant. Strains *unc-40(e1430);sdn-1(ev697)* and *unc-40(e1430);sdn-1(zh20)* were generated as outlined in Section 6.2.5 and the frequencies of DTC migration defects scored (Section 2.5.2). The *unc-40(e1430);sdn-1(zh20)* demonstrated significantly higher frequencies of DTC migration defects when compare to *unc-40(e1430)* mutants (Table 12) indicating the role of *sdn-1* in DTC guidance is not limited to *unc-40,* consistent with the *unc-5* results. However, *sdn-1(ev697)* in an *unc-40(e1430)* background causes lethality, thus scoring DTC migration defects in *unc-40(e1430);sdn-1(ev697)* was not possible. Taken together, these results strongly suggest that the *ev697* allele behaves differently than the *zh20 sdn-1* allele. To further characterize the role of *sdn-1* in *unc-5* mediated DTC guidance, an *unc-6(ev400)sdn-1(ev697)* strain was generated as outlined in section 6.2.6. The null *unc-6* allele *ev400* causes a complete loss of UNC-6 function (Wadsworth, Bhatt and Hedgecock 1996) and thus a loss

of UNC-40 and UNC-5 mediated DTC guidance as *unc-5(e53);unc-6(ev400)* and *unc-40(e1430);unc-6(ev400)* (Hedgecock, Culotti and Hall 1990) do not enhance DTC migration defect frequencies observed in *unc-6(ev400)*. An enhancement of DTC migration defects in *unc-6(ev400)sdn-1(ev697)* was observed only in the anterior, suggesting that the role of *sdn-1* in DTC guidance is limited to *unc-6* for the posterior DTC (Table 12). Thus the genetic interaction assays have demonstrated the role of *sdn-1* in DTC guidance is mostly limited to *unc-5* for anterior and posterior DTC migrations and limited to *unc-6* for posterior DTC guidance.

3.2.6 Suppression/enhancement of *evIs99* DTC migration defects in an *enh(ev697)* background.

To confirm the role of *sdn-1* in DTC guidance is limited to *unc-5* (Section 3.3.1) an *evIs99;sdn-1(ev697)* strain (Section 2.3) was generated. The *evIs99* transgenic strain contains an integrated DNA array with the entire *unc-5* gene coding sequence regulated by the *emb-9* promoter causing premature expression of *unc-5* during the first ventral DTC migration phase resulting in a precocious dorsalward turn of the DTC. This precocious turn is identified at low magnification by a dorsal clear patch in the animal. If a gene directly involved with *unc-5* mediated DTC guidance is placed in an *evIs99* background the frequency of precocious DTC migration defects is suppressed due to a disruption in a component required for *unc-5* mediated guidance preventing the precocious DTC turn. However if the gene is dispensable for *unc-5* function placing it in an *evIs99* background will not suppress DTC migration defects as *unc-5* still retains its ability to turn the DTC dorsally. The *evIs99; sdn-1(ev697)* strain was scored for the frequency of DTC migration defects and compared to *evIs99* alone. As expected, *sdn-1(ev697)* suppressed the precocious turn of the

DTC in *evIs99* (Table 13), confirming *sdn-1(ev697)* has a direct role in *unc-5* mediated DTC guidance.

Table 13 : DTC migration defects of *evIs99* and *evIs99;sdn-1(ev697)*

	Anterior DTC	Posterior DTC	*n*
evIs99	24 ± 2	40 ± 2	392
evIs99;sdn-1(ev697)	9 ± 2**[a]	28 ± 2**	250

[a] The statistical comparison is against the frequency of DTC migration defects in the *evIs99* strain alone. *$P<0.05$; **$P<0.001$.

However, before we can begin to understand the link between *sdn-1* and *unc-5,* the role of *sdn-1* in DTC guidance requires further characterization.

3.2.7 Determining the presence of SDN-1 within the DTCs.

The initial step taken to deduce the role of *sdn-1* in DTC guidance was to determine whether *sdn-1* acts cell-autonomously or cell non-autonomously. Transgenic lines *sdn-1(zh20);opEx1206* (*sdn-1* cDNA under the regulation of the *unc-119* pan-neural promoter), *sdn-1(zh20);opEx1159* (*sdn-1* cDNA under the regulation of the *dpy-7* hypodermal promoter) and *sdn-1(zh20);opEx1198* (*sdn-1* cDNA under the regulation of the *sdn-1* promoter) as the control (Section 2.3) were generated by Rhiner *et al.* (2005) to demonstrate that *sdn-1* acts cell-autonomously in axons for guidance. Our results have confirmed the involvement of *sdn-1* in *unc-5* mediated DTC guidance and interestingly, SDN-1 localization studies in both Minniti *et al.* (2004) (using an anti-syndecan-4 phospho-specific polyclonal antibody) and Rhiner *et al.* (2005) (using a GFP tagged SDN-1) experiments did not detect the presence of

SDN-1 in the DTCs, suggesting that the role of *sdn-1* in DTC guidance is not cell-autonomous and thus represents an uncharacterized *sdn-1* cell guidance role in *C.elegans*

The transgenic line *opIs170* (Section 2.3) with a GFP tagged SDN-1 protein was a kind gift from C. Rhiner. *opIs170* animals were prepared and analysed as described in Section 2.5.4. In agreement with the previous studies, SDN-1 was not detected in the DTCs (Figure 17) during the second DTC migration phase where it would be required for *unc-5* mediated turning of the DTC towards the dorsal muscle band. However, this result requires further analysis as Rhiner *et al.* (2005) have not indicated whether the GFP tagged SDN-1 can rescue the wild-type SDN-1 functions. As our cosmid phenotype rescue assay demonstrates the complications that occur by introducing an ectopic source of SDN-1, it is possible the SDN-1::GFP does not accurately represent endogenous SDN-1 localization within the animal. Phenotypic analysis suggests the SDN-1::GFP cannot rescue the egg-laying defect in *sdn-1* mutants as *opIs170* mutants are still egg-laying defective. In addition, the GFP tag is attached to the C terminus of the conserved endodomain of SDN-1, possibly interfering with endogenous SDN-1 functions, expression or localization. In order to confirm the GFP tag is not interfering with endogenous SDN-1 functions, an *opIs170;unc-5(e152);sdn-1(zh20)* strain would have to be generated and the frequency of DTC migration defects in this strain determined and compared to those of *unc-5(e152);sdn-1(zh20)*. If *opIs170;unc-5(e152);sdn-1(zh20)* can suppress the enhancement of DTC migration defects in *unc-5(e152);sdn-1(zh20),* the SDN-1::GFP tagged protein retains wild-type SDN-1 mechanisms for DTC guidance. As independent SDN-1 localization studies demonstrate that SDN-1 is not present in the DTC, it is highly probable *sdn-1* is functioning cell non-autonomously for DTC guidance.

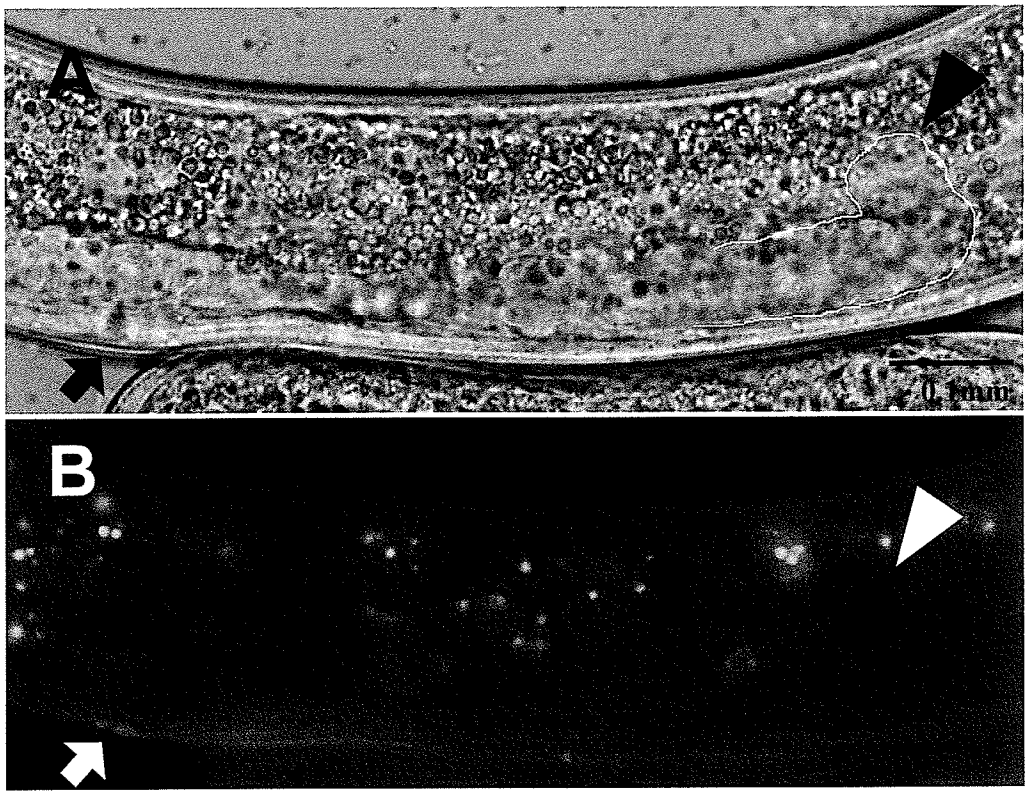


Figure 17 : Images of SDN-1::GFP localization in the DTC of *opIs170* animals. A) DIC bright-field image of an L1 animal. (Arrowhead denotes the DTC leading the gonad arm along the second dorsal-ward migratory phase and arrow denotes the position of the vulva.) The DTC nucleus is visible in this photograph. B) SDN-1::GFP expression in the same animal. SDN-1 expression is detected in the vulva and the ventral nerve cord (arrows) but is not present in the DTCs (arrowhead).

3.2.8 Hypodermal vs. axonal expression of *sdn-1* for DTC guidance.

As previously described, transgenic lines *sdn-1(zh20);opEx1206* (*sdn-1* cDNA under the regulation of the *unc-119* pan-neural promoter), *sdn-1(zh20);opEx1159* (*sdn-1* cDNA under the regulation of the *dpy-7* hypodermal promoter) and *sdn-1(zh20);opEx1198* (*sdn-1* cDNA under the regulation of the *sdn-1* promoter) as the control (Section 2.3) were generated to demonstrated *sdn-1* is required cell-autonomously in axons for guidance (Rhiner *et al.* 2005). Results from the previous section indicate that SDN-1 is not present in the DTCs during the second migration phase, arguing that SDN-1 is functioning cell non-autonomously for DTC guidance. Each transgenic array was separately crossed into an *unc-5(e152)* background to determine whether hypodermal or axonal expression of *sdn-1* is critical for DTC guidance. The *unc-5(e152);sdn-1(zh20);opEx1206[*$P_{unc\text{-}119}$*::sdn-1], unc-5(e152);sdn-1(zh20);opEx1159[*$P_{dpy\text{-}7}$*::sdn-1]* and *unc-5(e152);sdn-1(zh20);opEx1198[*$P_{sdn\text{-}1}$*::sdn-1]* strains were generated as outlined in Section 6.2.7. Each transgenic line was generated in a *sdn-1(zh20)* background resulting in exclusive SDN-1 expression from either the hypodermis (*opEx1159[*$P_{dpy\text{-}7}$*::sdn-1]*) or axons (*opEx1206[*$P_{unc\text{-}119}$*::sdn-1]*). As the control, *unc-5(e152);opEx1198[*$P_{sdn\text{-}1}$*::sdn-1]* should suppress the enhancement of DTC migration defect frequencies in *unc-5(e152);sdn-1(zh20).* Summarized in Table 14, *sdn-1* expression regulated by the *sdn-1* promoter (*opEx1198*) did not rescue the enhancement of DTC migration defects in *unc-5(e152);sdn-1(zh20)* mutants and of the three constructs the most significant rescue was observed with *opEx1206[*$P_{unc\text{-}119}$*::sdn-1]* indicating axonal expression and to a lesser extent hypodermal expression of *sdn-1* is required for DTC guidance.

Table 14 : Hypodermal/axonal restricted *sdn-1* expression and DTC migration defects.

unc-5	*Sdn-1*	Other	Ant. DTC	Post. DTC	*N*
e152	*zh20*	WT	36 ± 4	78 ± 3	831
e152	*zh20*	*Psdn-1::sdn-1*	26 ± 3*	67 ± 4*	156
e152	*zh20*	*Punc-119::sdn-1*	18 ± 2**	52 ± 3**	339
e152	*zh20*	*Pdpy-7::sdn-1*	26 ± 4*	69 ± 4*	117

[a] Each statistical comparison is against the frequency of DTC migration defects in *unc-5, sdn-1* strain alone. *$P<0.05$; **$P<0.001$.

The data in this table are in agreement with the data in Rhiner *et al.* (2005) as *opEx1198[P_{sdn-1}::sdn-1]* did not fully rescue axon guidance defects in *sdn-1(zh20)* mutants and axonal expression of *sdn-1* rescued DTC migration defects to a significant extent while hypodermal *sdn-1* expression rescued to a lesser extent. The inability of the *[P_{sdn-1}::sdn-1]* construct to fully rescue DTC migration defects in *unc-5(e152);sdn-1(zh20)* is not surprising as cosmid rescue assays with F57C7 (with the *sdn-1* gene) caused lethality when microinjected at high concentrations. The amount of SDN-1 produced from each construct has not been shown but could be verified using anti-syndecan-4 immunofluorescence in each transgenic strain. In combination with the SDN-1 GFP localization results, this data confirms *sdn-1* is not required cell-autonomously for DTC guidance as axonal and to a lesser extent hypodermal expression of *sdn-1* significantly suppressed DTC migration defects in *unc-5(e152);sdn-1(zh20)* mutants.

Heparan sulfate proteoglycans (HSPGs) are the core proteins, either secreted (perlecan) or membrane bound (syndecan/glypican) for the attachment of heparan sulfate

(HS) polysaccharides side chains at key serine residues. These side chains are extensively modified and vary in length accounting for the diversity of ligand/receptor interactions with HSPGs. HSPGs have numerous diverse roles [reviewed in (Bernfield *et al.* 1999)] including interactions within the TGFβ, WNT, EGF and FGF signalling pathways during development (Baeg and Perrimon 2000) and during axon guidance (Charron and Tessier-Lavigne 2005). In *C.elegans,* a role for HSPG UNC-52/perlecan in localizing growth factors for DTC guidance has been characterized (Merz *et al.* 2003). The heparan sulfate proteoglycan *sdn-1* has been shown to have a role in axon guidance by acting cell-autonomously in neurons in *C.elegans* and additionally, a role in DTC guidance that is cell non-autonomous. As migrating DTCs appear to require an axonal source of SDN-1 for guidance, *sdn-1* is most probably acting in a manner similar to UNC-52 by limiting or localizing growth factors for DTC guidance.

3.2.9 Genetic interactions of *sdn-1*, growth factors and *unc-5.*

unc-52 was identified in the genetic screen for enhancers of DTC migration defects in *unc-5(e152)* mutants. However unlike *sdn-1*, *unc-52* is not limited to *unc-5* mediated DTC guidance as *unc-52(e1421);unc-5(e53)* mutants exhibit significant increases in DTC migration defects when compared to frequencies in *unc-5(e53)* alone (Merz *et al.* 2003). The role of growth factors UNC-129(TGFβ), DBL-1(TGFβ), EGL-20(WNT) and EGL-17(FGF) in DTC guidance is an apparent gain of function mechanism that disrupts DTC guidance in the absence of UNC-52, suggesting UNC-52 localizes or limits these growth factors in a specific manner required for DTC migration. For example, *unc-52* enhances DTC migration defects in *unc-5(e152)*, but this enhancement is partially suppressed when *unc-52* and *unc-5* are in an *unc-129*, *dbl-1*, *egl-20* or *egl-17* background (Merz *et al.*

2003), suggesting UNC-52 is involved in suppressing a mechanism involving each growth factor that causes disruptions in the wild-type DTC migration pattern and removal of this suppression (*unc-52;unc-5(e152)*) induces DTC guidance disruptions. However, removal of each growth factor in addition to their suppressor reverts the DTC migration closer to wild-type. Guidance mechanisms of growth factors, regulated by *unc-52* for DTC guidance are redundant with other DTC guidance pathways as DTC migrations defects are not observed in *unc-52* mutants nor any of the aforementioned growth factor mutants. Their guidance roles are not apparent unless DTC migration defects are sensitized with an *unc-5* mutation.

Growth factors themselves are directly involved with DTC guidance as *dbl-1, unc-129* and *egl-20* (posterior only) enhanced DTC migration defect frequencies in *unc-5(e152)* mutants (Table 15) in a wild-type *unc-52* background.

Table 15 : DTC migration defects of *unc-5* and growth factor like mutants.

unc-5	Other	Ant. DTC	Post. DTC	*n*
e152	WT	7 ± 1	40 ± 1	1464
e152	*lin-3(e1413)*	10 ± 2	46 ± 3*[a]	252
e152	*egl-17(e1313)*	9 ± 2	43 ± 3	213
e152	*egl-20(mu39)*	21 ± 1**	49 ± 2**	738
e152	*dbl-1(ev580)*	24 ± 2**	59 ± 3**	322
e152	*unc-129(ev554)*	27 ± 4**	69 ± 4**	146

[a] Each statistical comparison is against the frequency of DTC migration defects in the *unc-5* strain alone. *P<0.05; **P<0.001.

C.elegans growth factors UNC-129/TGFβ (Colavita and Culotti 1998a)), DBL-1/TGFβ, (Suzuki *et al.* 1999), EGL-20/WNT (Maloof *et al.* 1999), EGL-17/FGF (Burdine *et al.* 1997) and LIN-3/EGF (Hill and Sternberg 1992) were selected as candidates for characterizing the interactions of growth factors with *sdn-1* for DTC guidance. Each growth

factor gene was placed in an *unc-5(e152);sdn-1(zh20)* and *unc-5(e152);sdn-1(ev697)* background (outlined in Section 6.2.8). As described for *unc-52*, if removing *sdn-1* function results in the delocalisation/gain of function of growth factors resulting in the enhancement of DTC migration defects in *unc-5(e152);sdn-1(zh20)* mutants, removing the growth factors should suppress the enhancement. Summarized in Table 16, suppression in *unc-5(e152);sdn-1(ev697)* by *egl-17* and *egl-20* and to a lesser extent *unc-129* and *lin-3* was limited to posterior DTC migrations and *dbl-1* did not suppress the enhancement of *unc-5(e152);sdn-1(ev697)*. However, each growth factor partially suppressed DTC migration defects mutants with a complete loss of *sdn-1* function (*zh20*) in both anterior and posterior DTC, indicating a role for SDN-1 in limiting growth factors UNC-129, DBL-1, EGL-20, LIN-3 and EGL-17. In addition, variation of the results between each allele was observed, further confirming *zh20* and *ev697 sdn-1* alleles behave differently.

Table 16 : DTC migration defects of *unc-5*, *sdn-1* and growth factor mutants.

unc-5	*sdn-1*	Other	Ant. DTC	Post. DTC	*n*
e152	*ev697*	WT	20 ± 1	63 ± 2	476
e152	*ev697*	*lin-3(e1413)*	22 ± 2	55 ± 5*	400
e152	*ev697*	*egl-17(e1313)*	18 ± 2	43 ± 3**	249
e152	*ev697*	*egl-20(mu39)*	22 ± 2	52 ± 2**	512
e152	*ev697*	*dbl-1(ev580)*	21 ± 2	68 ± 2	369
e152	*ev697*	*unc-129(ev554)*	23 ± 2	56 ± 2*	407
e152	*zh20*	*WT*	36 ± 2	77 ± 1	831
e152	*zh20*	*lin-3(e1413)*	17 ± 2**	55 ± 3**	330
e152	*zh20*	*egl-17(e1313)*	18 ± 2**	49 ± 3**	314
e152	*zh20*	*egl-20(mu39)*	16 ± 3**	46 ± 2**	502
e152	*zh20*	*dbl-1(ev580)*	17 ± 2**	51 ± 3**	296
e152	*zh20*	*unc-129(ev554)*	16 ± 2**	45 ± 2**	453

[a] Each statistical comparison is against the frequency of DTC migration defects in *unc-5;sdn-1* strain alone. *$P<0.05$; **$P<0.001$.

A recent review paper (Lee and Chien 2004) proposed four models for the roles of HSPGs in axon guidance:

1) HSPGs act as co-receptors and mediate ternary receptor complex formation between a HSPG, a ligand and its receptor.

2) HSPGs locally increase ligand concentrations at the cell surface and recruit membrane receptors forming lipid rafts.

3) HSPGs themselves can act as axon guidance ligands and receptors.

4) HSPGs regulate and limit the distribution of axon guidance ligands.

In the context of the DTCs and HSPG *sdn-1*, the first three of the four proposed models likely do not depict the role of *sdn-1* in DTC guidance, as HSPGs mediate axon guidance in these models in a cell-autonomous fashion and our results have demonstrated HSPG *sdn-1* is functioning cell non-autonomously for DTC guidance. If *sdn-1* were required specifically for the mechanisms of UNC-5/UNC-40 receptor mediated guidance within the DTC by formation of ternary complexes between UNC-40/UNC-5 and UNC-6 or as a ligand itself, one would expect to see DTC migration defects in the *sdn-1* mutant which is not the case. However, a role for heparan sulfate in stabilizing FGF and FGFR interactions has been described (Ornitz 2000) and interestingly in our study, *egl-17*(FGF) significantly suppressed DTC migration defects in *unc-5(e152);sdn-1(zh20)* mutants, suggesting a possible cell non-autonomous role for HSPGs in mediating FGF/FGFR interactions.

Cell guidance roles have been described for *egl-17* (Sex myoblast migration, (Burdine *et al.* 1997)) and *egl-20* (QL/QR neuroblasts, (Whangbo and Kenyon 1999)) and a role for *unc-129*, *dbl-1* and *egl-20* in DTC guidance has been shown (Merz *et al.* 2003), Table 15. However, *dbl-1, unc-129, lin-3, egl-17* and *egl-20* mutants, like HSPG/*sdn-1* mutants and HSPG/*unc-52* mutants do not exhibit DTC migration defects in an otherwise wild-type background, suggesting the limitation of growth factors by HSPGs in addition to the cell guidance roles of growth factors for DTC guidance are redundant. In addition, the limitation/regulation of one growth factor could directly affect another growth factor. For example, *lin-3*(EGF) is involved in inducing vulva cell fate in the vulva precursor cells and is expressed by the anchor cell located centrally amongst the vulva precursor cells (Hill and Sternberg 1992). *egl-17*(FGF) is expressed in vulva precursor cell P.6.p. and expression is required for sex myoblast (sex muscle precursors) migration to the developing gonad

(Burdine, Branda and Stern 1998). Taken together, one could speculate that *lin-3* is possibly involved in regulating *egl-17* function or expression for the formation of a functioning vulva.

A model for the role of SDN-1 limited growth factor distribution in DTC guidance can be postulated. As we have shown the role of *sdn-1* in DTC guidance is limited to *unc-5*, one possible model pertains to the either the regulation of *unc-5* expression or the activation of UNC-5 in response to UNC-6, in that *sdn-1* in the ventral nerve cord or hypodermis limits the localization of growth factors in a manner that sequesters growth factors away from *unc-5*/UNC-5 (if the growth factors are acting as activators) or makes them available to *unc-5*/UNC-5 (if the growth factor are acting as repressors) during the first migration phase and both require a reversal at the time of DTC turning towards the dorsal muscle band. The role of HSPGs in either restricting or facilitating diffusion of morphogens is supported by recent findings in *Drosophila*. The diffusion patterns of Hh (hedgehog) and Wgl (wingless, WNT) are altered in *sotv (sister of tout velu)* and *ttv (tout velu)* mutants and in addition, Dpp (Decapentaplegic, BMP/TGFβ) signalling is disrupted as these mutants exhibit defects in glycosaminoglycan (GAG) chain synthesis and thus have reduced levels of heparan sulfate (Bornemann *et al.* 2004). In addition, the HSPG Dally (glypican) was found to modulate the Dpp gradient in *Drosophila* (Fujise *et al.* 2003). Interestingly, results in Table 16 indicate DTC migration defect frequencies in triple mutants *unc-5(e152);dbl-1(ev580);sdn-1(zh20)* and *unc-129(554)unc-5(e152);sdn-1(zh20)* are much lower then the DTC migration defect frequencies in the double mutants *unc-5(e152);dbl-1(ev580)* and *unc-129(554)unc-5(e152)* (Table 15). The enhancement of DTC migration defect frequencies caused by each TGFβ in *unc-5(e152)* mutants indicates their involvement in the ventral to dorsal guidance of the DTC. In addition, Merz *et al.* (2003) have proposed the *unc-129* and *dbl-1* DTC guidance mechanisms are limited to *unc-5* mediated guidance and that both *unc-129* and *dbl-1* interact

within the same guidance pathway. Our *sdn-1* data suggests two possible theories, 1) *sdn-1* limits *dbl-1* and *unc-129* (removal of *sdn-1* disrupts the *dbl-1* and *unc-129* role in DTC guidance) and 2) *dbl-1* and *unc-129* can compensate for each other's guidance functions. For example, when you take one TGFβ away (*unc-5(e152);dbl-1(ev580)* mutant) and delimit UNC-129 (*unc-5(e152); dbl-1(ev580);sdn-1(zh20)* mutant), delimited UNC-129 can compensate for the loss of DBL-1 resulting in a lower frequency of DTC migration defects observed in *unc-5(e152); dbl-1(ev580);sdn-1(zh20)*). As *unc-129* is expressed dorsally and *dbl-1* is expressed ventrally, it is quite possible a specific gradient of each TGFβ growth factor, limited by *sdn-1*, is required for DTC guidance along the ventral-dorsal axis. However further characterization of the role for each growth factor and their respective receptors in DTC guidance is required before we can begin to understand the complete picture.

We have proposed that HSPG *sdn-1* has a role in limiting EGL-17(FGF), UNC-129(TGFβ), DBL-1(TGFβ), EGL-20(WNT) and LIN-3(EGF) growth factor molecules for DTC guidance. As we have shown the role of *sdn-1* in DTC guidance is limited to *unc-5* guidance mechanisms, our finding suggests these growth factors are functioning within an *unc-5* mediated guidance mechanism. However further analysis is require to confirm that the role of these growth factors, regulated by *sdn-1,* in DTC guidance are limited to *unc-5.*

We have identified an allele of *sdn-1* encoding the truncated, SDN-1 ectodomain, which appears to be present and possibly functional in the animal as suggested by the repeated differences between *ev697* and *zh20* alleles in our results. Shedding of the syndecan-1 ectodomain, induced by FGF2 and MMP7 (matrix metalloprotease 7) was shown to be associated with tumour progression in pancreatic cancer cells (Ding *et al.* 2005). Thus the *ev697* allele may serve as a potential model in *C.elegans* for further study of the role of

sdn-1 in the progression of this disease. Further more, *sdn-1* and DTC guidance provides a system for modelling the diverse roles of *sdn-1* in cell migrations.

4.0 CONCLUSION

4.1 The *mig(ev648)* enhancer allele.

The enhancer *mig(ev648)* identified in the genetic screen for enhancers of DTC migration defects is an allele of the nucleoside diphosphatase *mig-23* gene. *mig-23* is required for the glycosylation of the matrix metalloprotease *mig-17* for proper localization during DTC migrations (Nishiwaki *et al.* 2004). We have identified a previously uncharacterized allele of *mig-23, ev648* consisting of a 335C-T mutation in the *mig-23* coding sequence resulting in an A112V mutation in the MIG-23 protein. This mutation is not within an apyrase conserved domain of the protein described by (Nishiwaki *et al.* 2004) but is within a relatively conserved region of NDPases and the allele appears to be partially temperature sensitive.

We have demonstrated the role of *mig-23* in DTC guidance is not limited to *unc-5/unc-40/unc-6* guidance mechanisms and does not have a role in *unc-5* mediated DTC guidance. Genetic interactions have suggested *mig-17* guidance mechanisms are not limited to *unc-5/unc-6/unc-40* DTC guidance mechanisms (Nishiwaki *et al.* 2000) and genetic interactions of *mig-23* and *mig-17* suggests *mig-23* has additional DTC guidance roles independent of *mig-17* mediated guidance (Nishiwaki *et al.* 2004). As *mig-17* did not arise in our genetic screen for enhancers of DTC migrations in *unc-5(e152)* mutants, it is reasonable to hypothesize that the *mig-17* independent DTC guidance mechanisms of *mig-23* are perhaps associated with *unc-40* or *unc-6* DTC guidance mechanisms. However, further analysis is required.

4.2 The *enh(ev697)* enhancer allele.

The enhancer *enh(ev697)* allele identified in the genetic screen for enhancers of DTC migration defects is an allele of the heparan sulfate proteoglycan *sdn-1*. *sdn-1* is involved in guiding axons and neurons along specified migration patterns (Rhiner *et al.* 2005). We have identified a previously uncharacterized allele of *sdn-1, ev697,* consisting of a 610C-T mutation in the *sdn-1* coding sequence resulting in a Q203X mutation in the SDN-1 protein, possibly resulting in a truncated, un-tethered form of *sdn-1* whose presence is suggested by the differences in behaviour observed between the *ev697* and null *zh20* alleles. The role of *sdn-1* in DTC guidance appears to be limited to *unc-5* mediated guidance. In contrast to the role of *sdn-1* in axon guidance, *sdn-1* functions cell non-autonomously for DTC guidance as an axonal source of SDN-1 rescues DTC migration defects, as does a hypodermal source but to a lesser extent. In accordance with these data, our genetic interaction analysis suggests *sdn-1* has a role in limiting growth factors *unc-129*(TGFβ), *dbl*-1(TGFβ), *egl-17*(FGF), *lin*-3(EGF), and (*egl-20*)WNT within the extra-cellular environment for DTC guidance, possibly in an *unc-5* dependent manner.

Recently roles for semaphorin guidance cues and their receptors plexins/neutropilins have been identified in biological processes other then axon guidance including immune function regulation, angiogenesis and cancer [reviewed in (Tamagnone and Comoglio 2000)]. As Netrins and their receptors appear to be following the same pathway, a concrete understanding of the mechanisms and regulation of Netrins and their receptors is essential. Identifying genes involved in regulating cell motility and assembling genes into pathways and hierarchies contributes to our understanding of the mechanisms of cell guidance that are crucial for diverse biological and developmental processes.

5 FUTURE DIRECTIONS

The roles of *mig-23* in DTC cell guidance have been characterized and are relatively well defined (Nishiwaki *et al.* 2004). However, the possibility that the MIG-23 NDPase is functioning with UNC-6 or UNC-40 for DTC guidance has yet to be investigated. Further characterization of the genetic interactions between *mig-23* and *unc-6* or *unc-40* is required to determine whether *mig-23/mig-17* play a role in UNC-6 or UNC-40 DTC guidance mechanisms.

This is the first study to identify a cell non-autonomous role for SDN-1 in DTC guidance involving growth factors. The *sdn-1* allele we have generated encodes a truncation before the trans-membrane domain of SDN-1, potentially resulting in an unbound form of the syndecan ectodomain. Our data suggests the *ev697* form of SDN-1 is present, however Western blot analysis would be required to confirm the presence of this truncated form of SDN-1.

To support our finding that *sdn-1* is functioning cell non-autonomously for DTC guidance, a construct with the *sdn-1* coding sequence regulated by the DTC specific promoter *lag-2* in an *unc-5(e152);sdn-1(zh20)* would conclusively determine whether *sdn-1* within the DTC is enough to rescue the enhancement of DTC migration defects in *unc-5(e152)sdn-1(zh20).* In addition, (Rhiner *et al.* 2005) have generated additional transgenic lines with the constructs used in this study to determine whether expression of *sdn-1* in axon or the hypodermis is required for DTC migration.

Embryonic elongation defects present in *sdn-1(ev697)* suggests a role for *sdn-1* during embryogenesis. An increase in penetrance of this phenotype in *unc-40(e1430);sdn-1(ev697)* suggests *unc-40* is involved in this uncharacterized role of *sdn-1*. Further analysis of the HS binding abilities of UNC-40 may elucidate these potential interactions.

As our data suggests *egl-17*(FGF) is possibly directly involved in *unc-5* mediated DTC guidance, further analysis of the genetic interaction between *clr-1* and *unc-5* and *unc-5* and *egl-15* using the *evIs99* strain would determine whether these genes mediate DTC guidance in an *unc-5* dependent manner. As CLR-1 is a tyrosine phosphatase, analysis of UNC-5 phosphorylation in *egl-15*, *egl-17* and *clr-1* mutants would determine whether CLR-1 has direct interactions with UNC-5. In addition, as *egl-17*, *egl-20*, *lin-3*, *dbl-1* and *unc-129* suppressed DTC migration defects in *unc-5(e152);sdn-1(zh20),* analysing *unc-5* expression in a mutant background of each growth factor would determine whether growth factors are involved in regulating *unc-5* expression. *sdn-1* was originally identified in *C.elegans* for its role in axon guidance and analysis of its interactions with growth factors in mediating axon guidance defects would determine whether *sdn-1* guidance mechanisms are conserved among different cell types.

SDN-1 is a heparan sulfate proteoglycan with modifiable heparan sulfate side chains. Genetic interactions in *C.elegans* between *sdn-1* and heparan sulfate modifying enzymes *hse-5* (C5-epimerase) and *hst-2*(*20*-sulfotransferase) in axons have indicated that different heparan sulfate modifications of *sdn-1* are required for the guidance of different neurons (Rhiner *et al.* 2005). The nature of *sdn-1* heparan sulfate side-chain modifications required for DTC guidance has yet to be shown.

Thus by employing a combination of genetic interaction analysis and biochemical assays, the model organism *C.elegans* can be used to further our understanding of the roles of growth factors and heparan sulfate proteoglycans in cell guidance and axon guidance.

6 APPENDIX

6.1 : Solutions

Nematode Growth Medium Agar	
NaCl	3g
Agar	17g
Peptone	2.5g
Cholesterol (5mg/ml in EtOH)	1ml
dH_20	975ml
Autoclave, then add the following using sterile technique	
$CaCl_2$ 1M	1ml
$MgSO_4$ 1M	1ml
potassium phosphate 1M pH6	25ml

LB broth	
Tryptone	10g
Yeast Extract	5g
NaCl	10g
H_20	1L

0.5M EDTA (pH8.0)	
EDTA	14.61g
NaOH	2g
H_20	to 80 µl

50 X TAE	
Tris base	242g
Glacial acetic acid	57.1ml
0.5 EDTA (pH8.0)	100ml
ddH20	1L
For 1X TAE (Agarose gel electrophoresis running buffer), add 20ml 50X TAE to 1000ml dH20	

1% agarose gel	5mm		10mm	
	Agarose	TAE	Agarose	TAE
Small	0.2g	20ml	0.45g	45ml
Medium	0.6g	60ml	1.25g	125ml

1.5% agarose gel	5mm		10mm	
	Agarose	TAE	Agarose	TAE
Small	0.3g	20ml	0.675g	45ml
Medium	0.9g	60ml	1.875g	125ml

Ethidium Bromide	
EtBr	1g
dH_20	100ml
Stir for several hours in container covered with foil.	

Orange G (6X loading dye)	
Glycerol	30ml
Orange G	0.25g
0.5M EDTA	400 µl
dH_20	100ml

Low TE (10mM Tris, 1mM EDTA pH8)	
1M Tris, pH8	1ml
0.5M EDTA, pH8	0.2ml
dH_2O	98.8ml

DNA Ladder (Invitrogen)	
Orange G dye	170µl
DNA ladder	50 µl
Low TE	780 µl

LB agar	
Tryptone	10g
Yeast Extract	5g
NaCl	10g
Agar	15g
H_20	1L

Rich Agarose Plates (500ml=20 plates)	
50mM NaCl	1.56g
5ug/ml cholesterol (autoclaved)	500µl
1.5% agarose	7.5g
Mix in 500ml water and autoclave.	
1mM $CaCl_2$ (autoclaved)	500µl
1mM $MgSO_4$ (autoclaved)	500µl
25 mM K-PO_4 (pH 6.0) (autoclaved)	12.5ml

M9 Buffer	
KH_2PO_4	3g
Na_2HPO_4	6g
NaCl	5g
$MgSO_4$ 1M	1ml
H_20	1L

Genomic Worm Lysis Solution (store at -20^{0}C)	
1M Tris (pH 8.5)	1ml
100mM NaCl	0.058g
0.5M EDTA	1ml
10% SDS	1ml
1% beta-mercaptoethanol	100µl
100 ug/ml proteinase K	65.36µl
ddH_20	6.2ml

Phenol/$CHCl_3$ (Phenol alcohol:$CHCl_3$:isoamyl alcohol, 25:24:1)	
Phenol alcohol	250ml
$CHCl_3$	240ml
Isoamyl alcohol	10ml

$CHCl_3$ (24:1)	
isoamyl alcohol	10ml
$CHCl_3$	240ml

10mM dNTP (Invitrogen kit)	
100mM dTTP	10µl
100mM dGTP	10µl
100mM dATP	10µl
100mM dCTP	10µl
ddH_20	60µl

General mixture for 50ul PCR**.	
Worm lysis	2µl
10 X PCR Buffer (Invitrogen)	5µl
25 mM $MgCl_2$ (Invitrogen)	3µl
10 mM dNTP	1µl
10 pM primer 1	1µl
10 pM primer 2	1ul
Taq (DNA polymerase or Platinum, Invitrogen)	0.125µl
ddH_20	36.875 µl

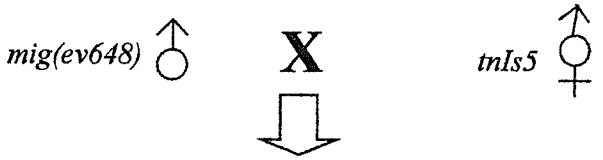
**PCR reactions were optimized by adjusting worm lysis and primer concentrations.

Single worm lysis buffer (1ml aliquots in -20^{0}C)	
1M KCl	0.5ml
1M Tris	0.1ml
1M $MgCl_2$	0.025ml
10% Triton	0.45ml
10% Tween-20	0.45ml
10% gelatin	0.1ml
20mg/ml proteinase K	30µl
ddH_20	8.345ml

Agarose 2%	
Agarose	1g
dH_20	5ml

6.2 *C.elegans* mutant strain generation outlines

6.2.1 *tnIs5;mig(ev648)*



F1

Clone out wild-type, GFP heterozygotes.

From GFP clones, pick progeny with *mig(ev648)* DTC migration defects and strong GFP

Ensure all clones carry the GFP and exhibit *mig(ev648)* DTC migration defects.

tnIs5; mig(ev648)

6.2.2 *unc-40(e1430);unc-5(e53);mig(ev648)*

F1

Clone out wild-type heterozygotes

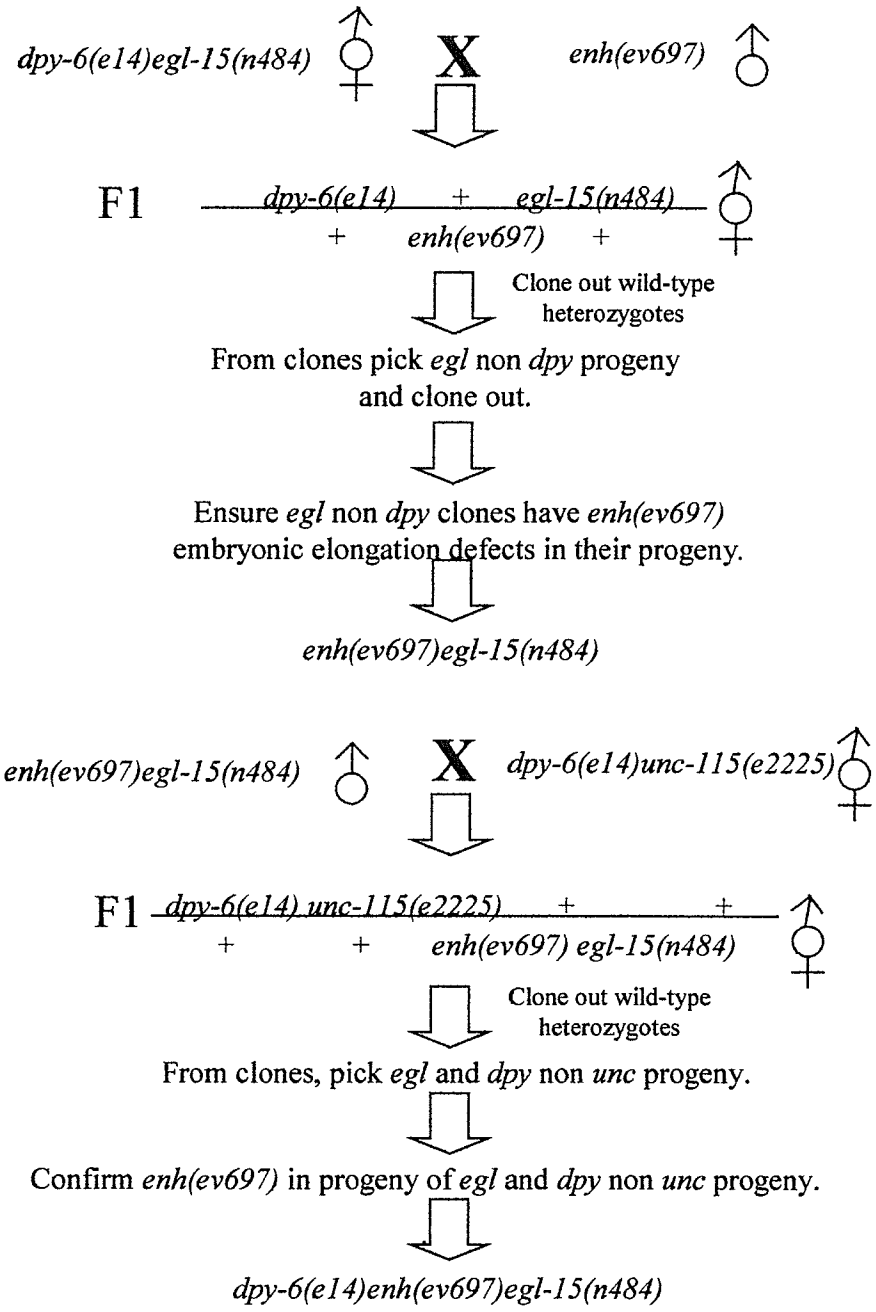
From clones, pick progeny with *mig(ev648)* DTC migration defects and clone out.

From *mig(ev648)* clones, pick uncoordinated progeny with *mig(ev648)* DTC migration defects and clone out. [*unc-5(e53);mig(ev648)*]

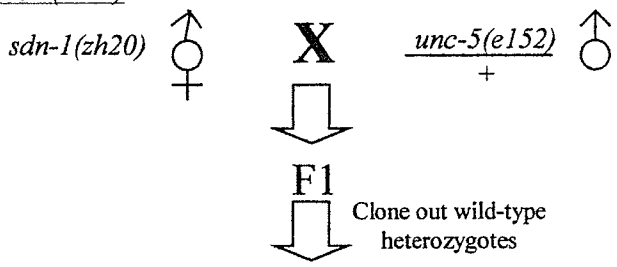
From *unc-5(e53);mig(ev648)* clones, pick egl,, dpy and uncoordinated progeny and clone out.

unc-40(e1430);unc-5(e53);mig(ev648)

6.2.3 *dpy-6(e14)enh(ev697)egl-15(n484)*



6.2.4 *unc-5(e152);sdn-1(zh20)*



From clones, pick unc animals with DTC migration defects and clone out. [*unc-5(e152)*].

From *unc-5(e152)* clones, pick uncoordinated progeny that are bloated and clone out.

From the suspected *unc-5(e152);sdn-1(zh20)* clones, pick at least 10 uncoordinated progeny that are bloated from plates with *sdn-1* embryonic elongation defects.

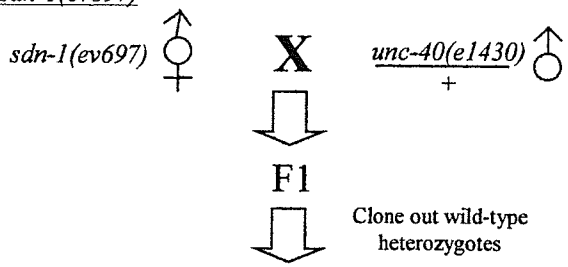
Ensure all 10 uncoordinated and bloated animals chosen have *sdn-1* embryonic elongation defects in their progeny.

unc-5(e152);sdn-1(zh20)

unc-5(e53);sdn-1(zh20)

Repeat the same instructions for *unc-5(e152);sdn-1(zh20)* mutant generation but use *unc-5(e53)*/+ ♂ s

6.2.5 *unc-40(ev1430);sdn-1(ev697)*



From clones, pick uncoordinated, bloated progeny with DTC migration defects and clone out.[(*unc-40(e1430)*)]

In *unc-40(e1430)* clones, look for *ev697* embryonic elongation defects in the progeny and from these plates, clone out at least 10 uncoordinated animals.

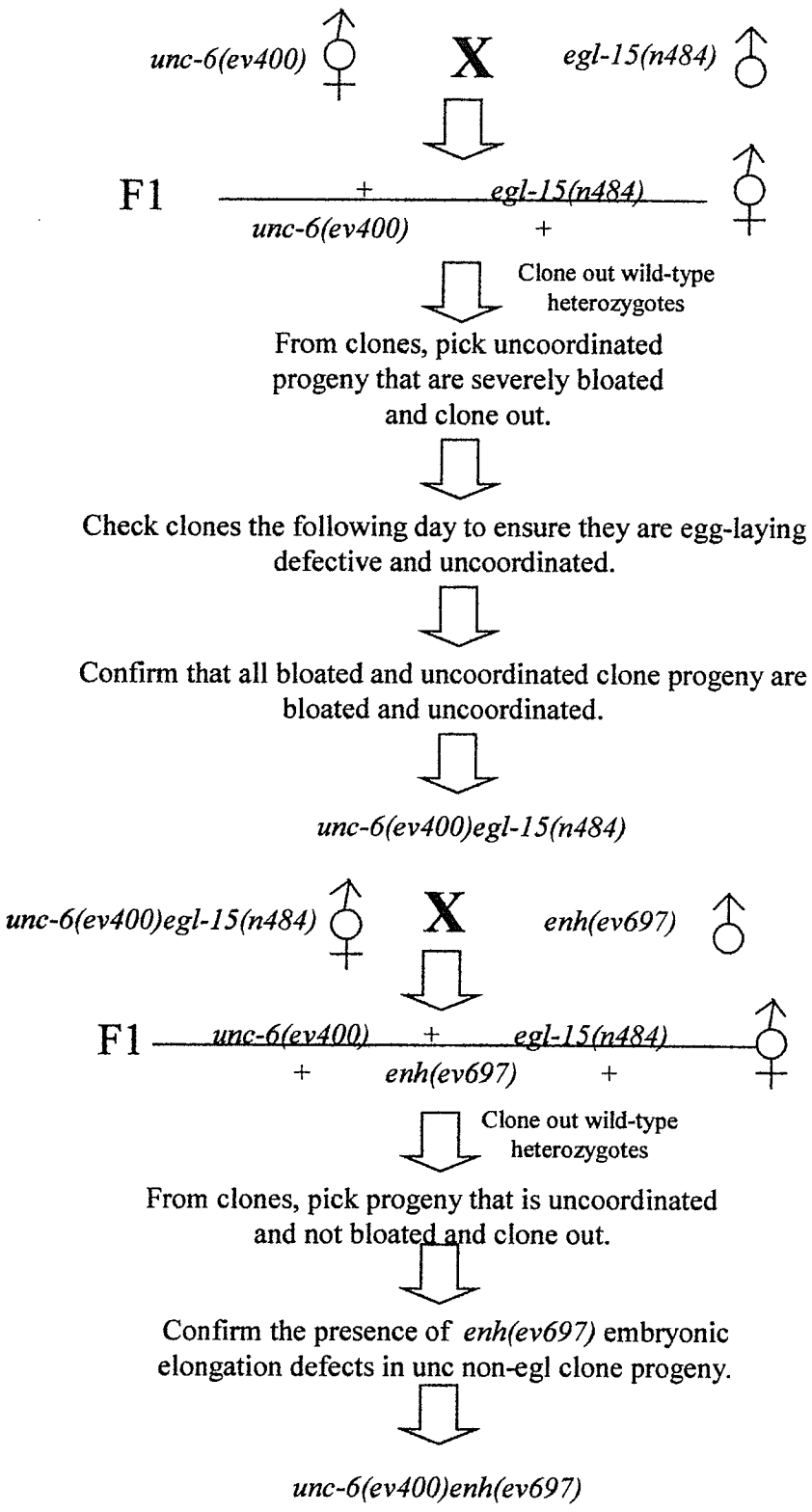
Ensure all 10 uncoordinated and bloated clones have *ev697* embryonic elongation defects in their progeny.

unc-40(ev1430);sdn-1(ev697)

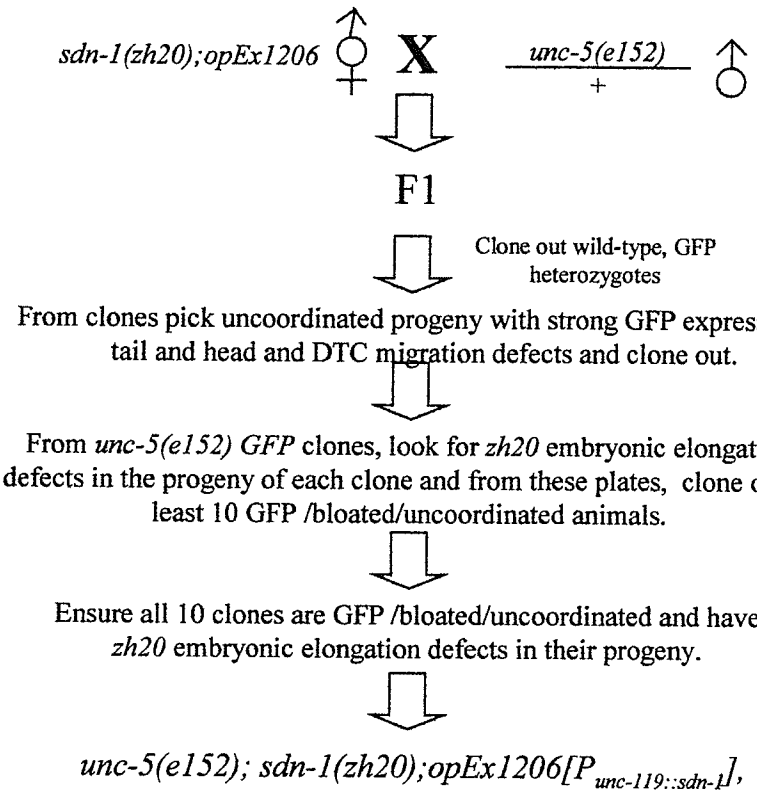
unc-40(ev1430);sdn-1(zh20)

Repeat the same outline for *unc-40(ev1430);sdn-1(ev697)* mutant generation but use *sdn-1(zh20)* / + ♀ s

6.2.6 *unc-6(e400)enh(ev697)*



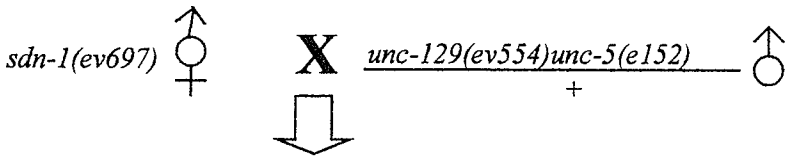
6.2.7 *unc-5(e152);sdn-1(zh20);opEx1206[$P_{unc-119}$::sdn-1]*



The same outline was utilized for generating:

- *unc-5(e152);sdn-1(zh20);opEx1159[P_{dpy7}::sdn-1]*
- *unc-5(e152);sdn-1(zh20);opEx1198[P_{sdn-1}::sdn-1]*

6.2.8 *unc-129(ev554)unc-5(e152);sdn-1(ev697)*



F1

Clone out wild-type heterozygotes

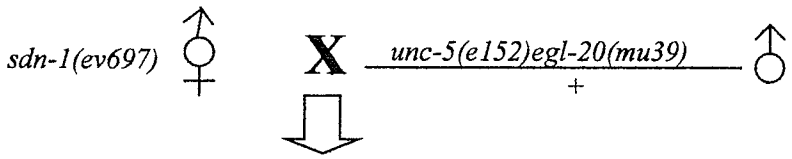
From clones, pick uncoordinated progeny with DTC migration defects and clone out.

Look for *ev697* elongation defects in the progeny of each clone and from these plates, clone out at least 10 uncoordinated animals.

Ensure all 10 clones are uncoordinated and have *ev697* embryonic elongation defects in their progeny.

unc-5(e152);unc-129(ev554);sdn-1(ev697)

6.2.9 *unc-5(e152)egl-20(mu39);sdn-1(ev697)*



F1

Clone out wild-type heterozygotes

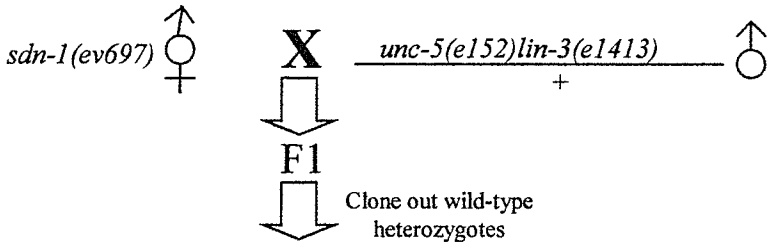
From clone,s pick uncoordinated bloated progeny with DTC cell migration defects and clone out.

Look for *ev697* elongation defects in the progeny of each clone and from these plates, clone out at least 10 uncoordinated and bloated animals.

Ensure all 10 clones are uncoordinated and bloated and have *ev697* elongation defects in their progeny.

unc-5(e152)egl-20(mu39);sdn-1(ev697)

6.2.10 *unc-5(e152)lin-3(e1413);sdn-1(ev697)*



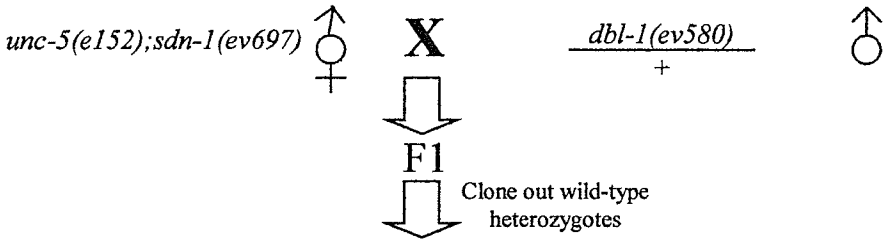
From clones, pick uncoordinated, bloated and valvaless progeny with DTC migration defects and clone out.

Look for *ev697* embryonic elongation defects in the progeny of each clone and from these plates, clone out at least 10 uncoordinated and bloated animals.

Ensure all 10 clones are uncoordinated and bloated and have *ev697* embryonic elongation defects in their progeny.

unc-5(e152)lin-3(e1413);sdn-1(ev697)

6.2.11 *unc-5(e152);dbl-1(ev580);sdn-1(ev697)*



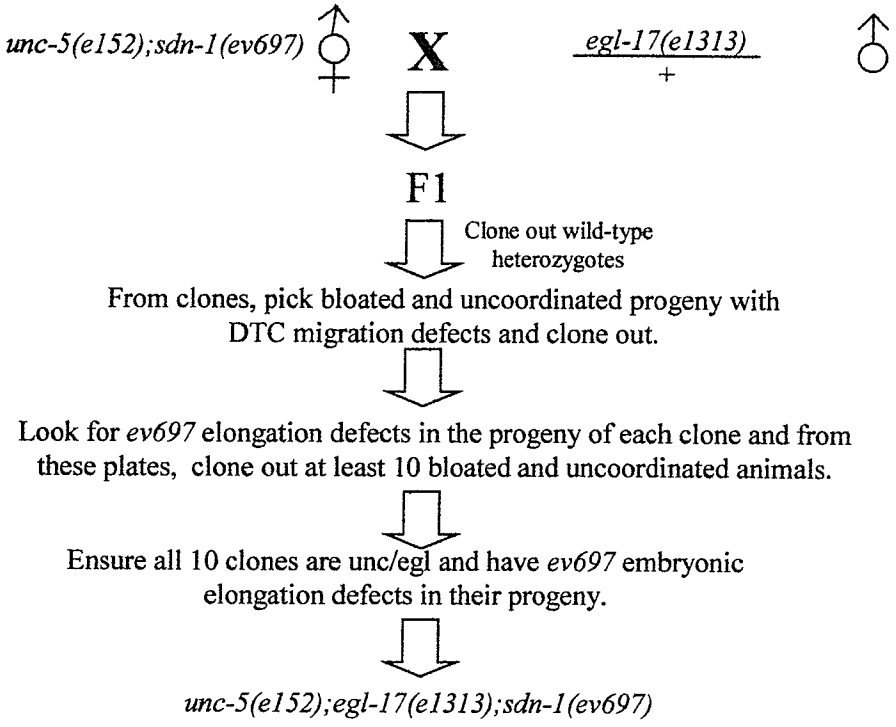
From clones, pick dumpy, uncoordinated progeny with DTC migration defects and clone out.

Look for *ev697* elongation defects in the progeny of each clone and from these plates, clone out at least 10 dumpy and uncoordinated animals.

Ensure all 10 clones are unc/egl and have *ev697* embryonic elongation defects in their progeny.

unc-5(e152);dbl-1(ev580);sdn-1(ev697)

6.2.12 *unc-5(e152);egl-17(e1313);sdn-1(ev697).*



6.3 Cosmid descriptions.

6.3.1 Summary of the cosmids used for *mig(ev648)* rescue.

Cosmid	Approximate size (Kb)	Kan/Amp?
C03B1	43	AMP
T22E5	34.5	KAN
K10	34	KAN
K07	11	KAN
C15B12	44	AMP
F22A3	30	KAN
T14E8	37	KAN
T28B4	27	KAN

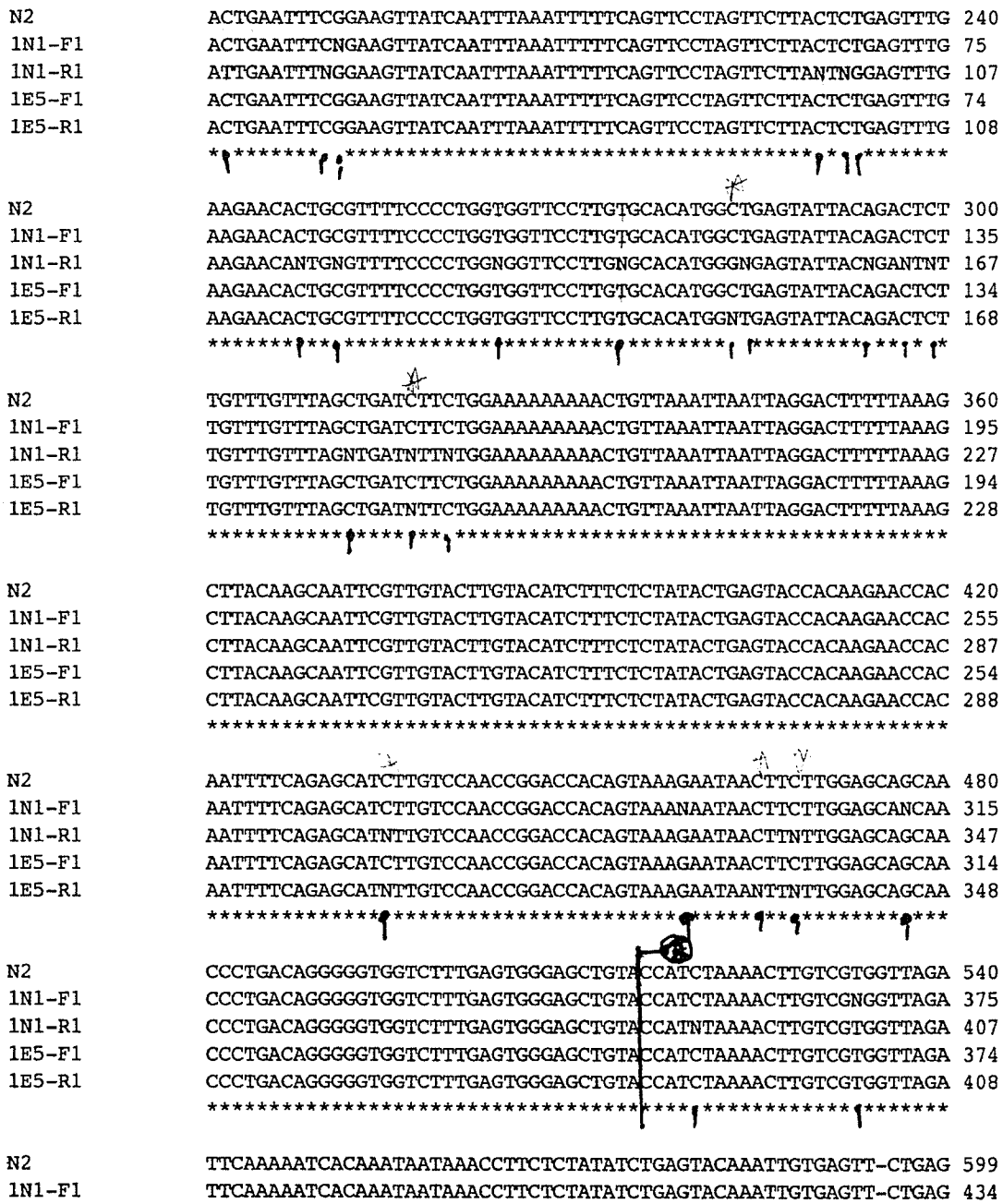
6.3.2 Summary of the cosmids used for *enh(ev697)* rescue.

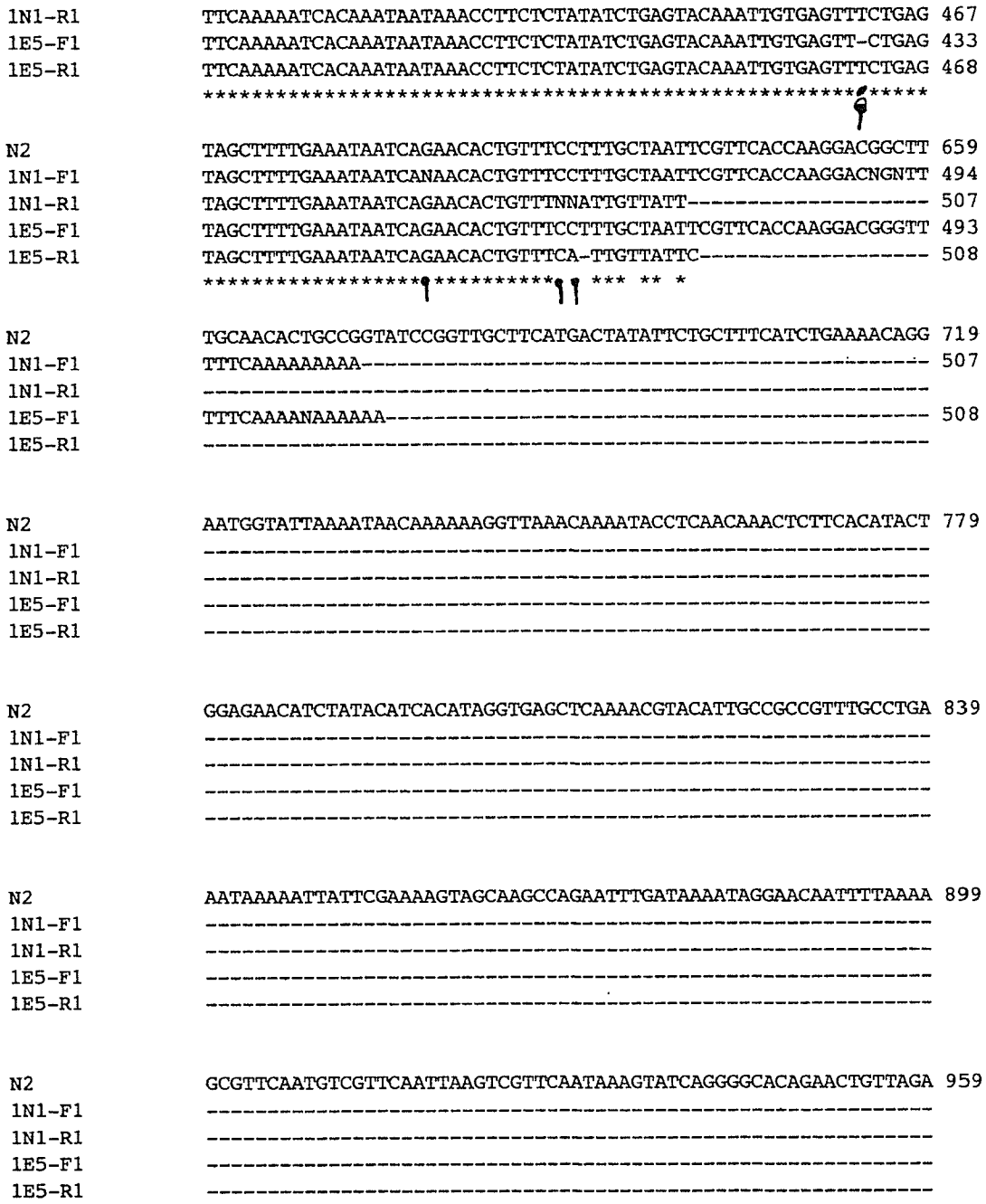
Cosmid	Approximate size (Kb)	Kan/Amp?
F41E7	36	KAN
R07E3	40	AMP
F46F6	25	KAN
ZC504	38	AMP
C39B10	38	AMP
C33D3	18	AMP
F14F3	35	KAN
F59F5	35	KAN
F57C7	30	KAN
M79	34	AMP
F11A1	40	KAN
F13E6	40	KAN
C46B5	1.6	AMP
T01C1	28	AMP
R07A4	37	AMP

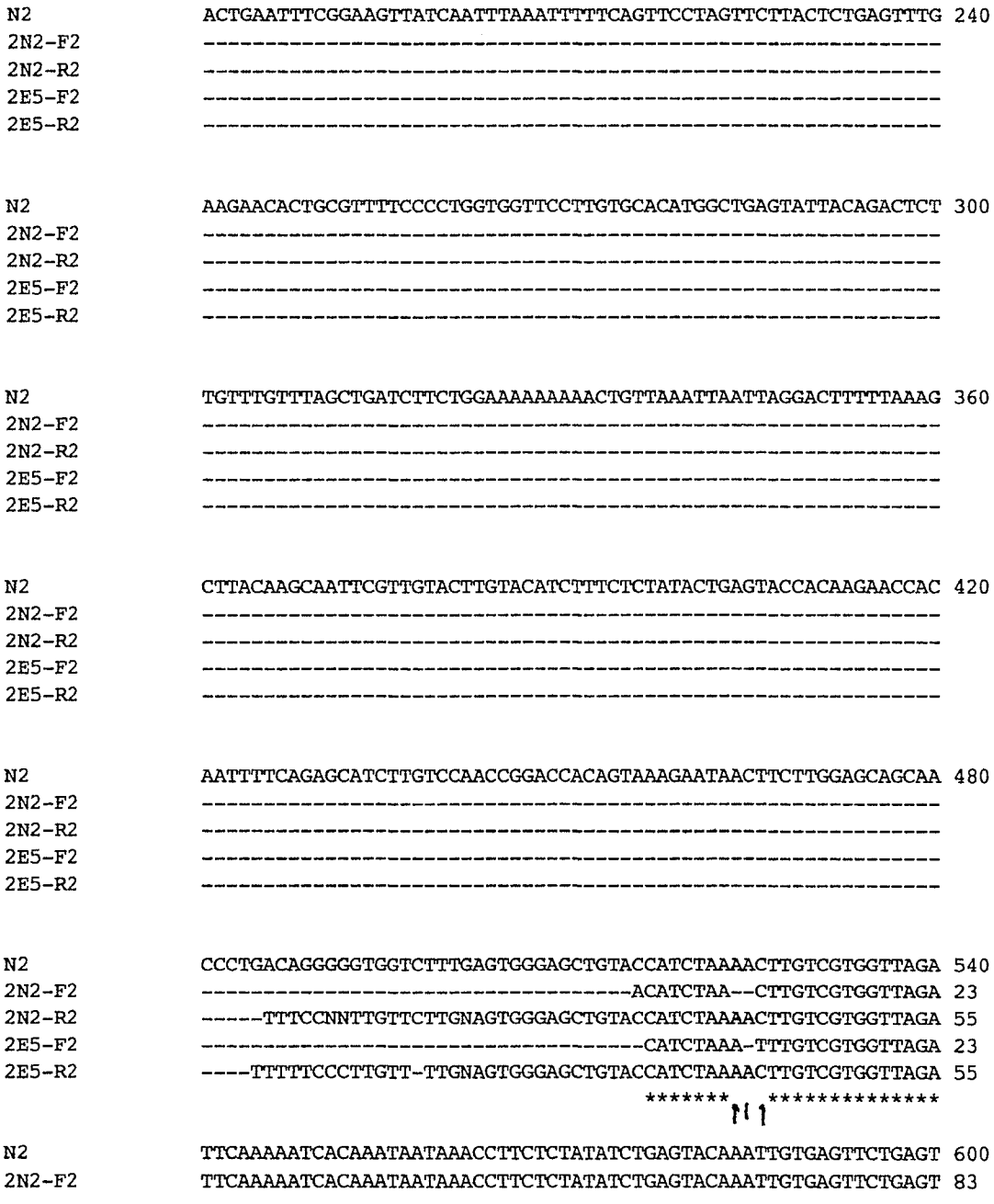
6.4 Alignments of *nas-33* sequencing results.

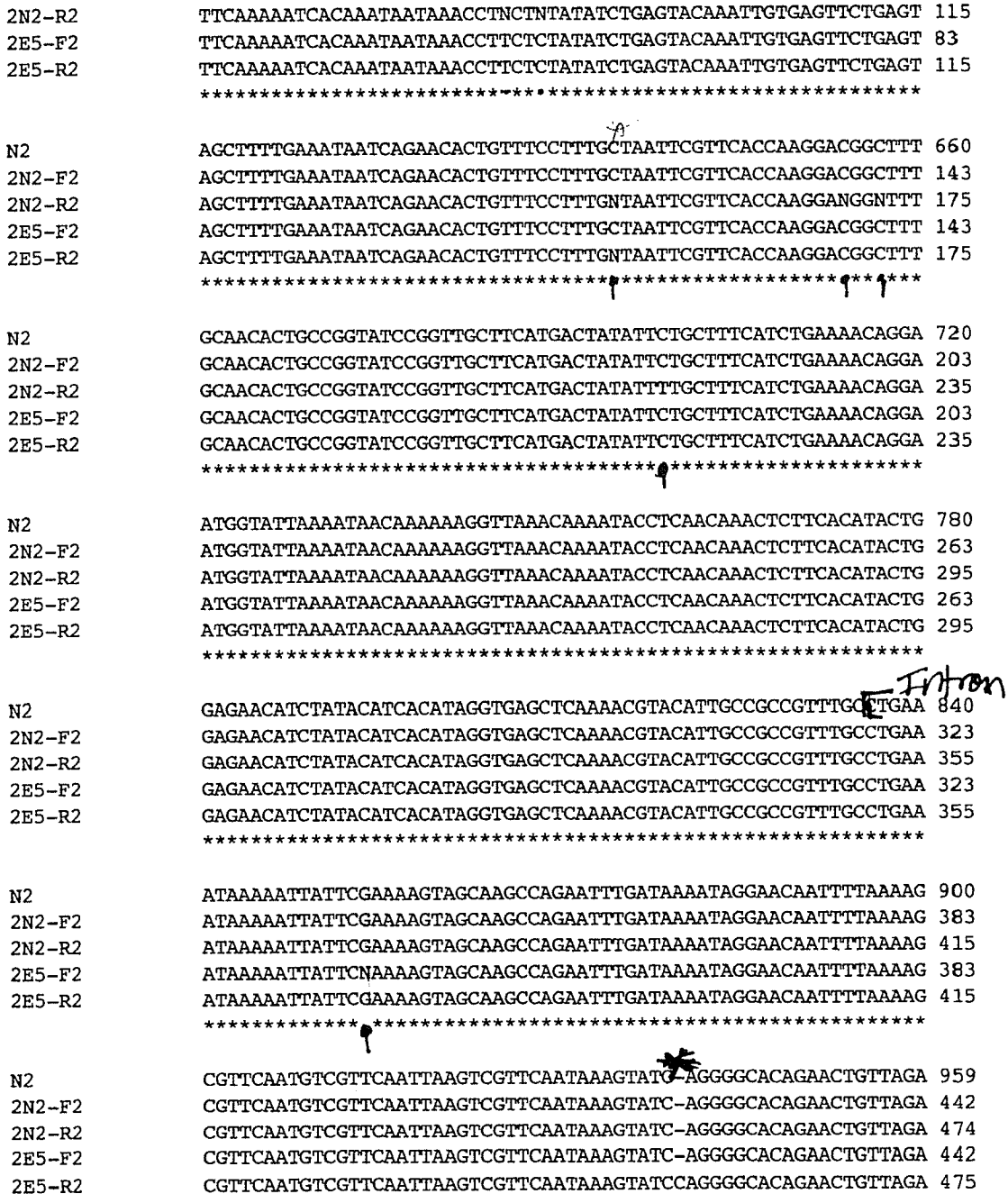
clustalw-20031208-18091412.aln K04E7 cosmid seq alignment need GC-AT

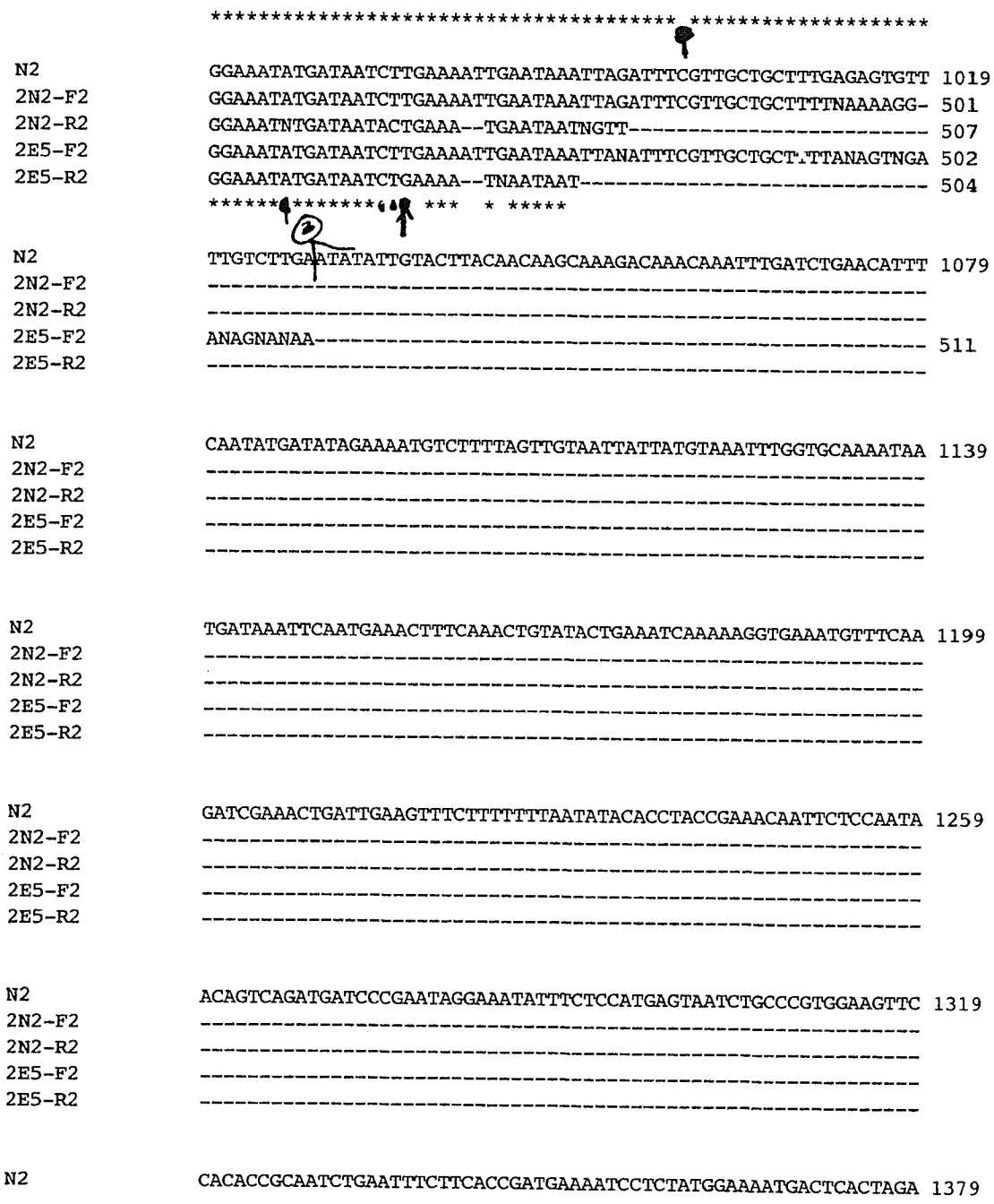
```
CLUSTAL W (1.82) multiple sequence alignment


N2              GATCCTTAGACAGTTATCTGAAGGTCATAGATGTTATGATTCAAACTAGATTCCGCTTCT 60
1N1-F1          ------------------------------------------------------------
1N1-R1          ------------------------------------------------------------
1E5-F1          ------------------------------------------------------------
1E5-R1          ------------------------------------------------------------


N2              CCTTCTATTTTTCTTATTCAGTTCGGATGCGTCACCCCTGTTCTCACTGGAACACAGACA 120
1N1-F1          ------------------------------------------------------------
1N1-R1          ------------------------------------------------------------
1E5-F1          ------------------------------------------------------------
1E5-R1          ------------------------------------------------------------


N2              CCTCTTGAAAGAACAAATCTGGCACAACACCACCCATTGAGACGTTTTACACGATGTGAG 180
1N1-F1          ----------------------------------------------TTTTAC-NGANTTGAN  15
1N1-R1          ----------TTTTNTAAANCCAACNCCACCCATTGAGANGTTTTACANGATGTGAG  47
1E5-F1          ----------------------------------------------TTTTAA-CGANTN-AN  14
1E5-R1          ----------TTTTTNNNNCCCNAACACNCCCCATTGAGACGTTTTACACGATGTGAG  48
                                                             *****111**111*1
```

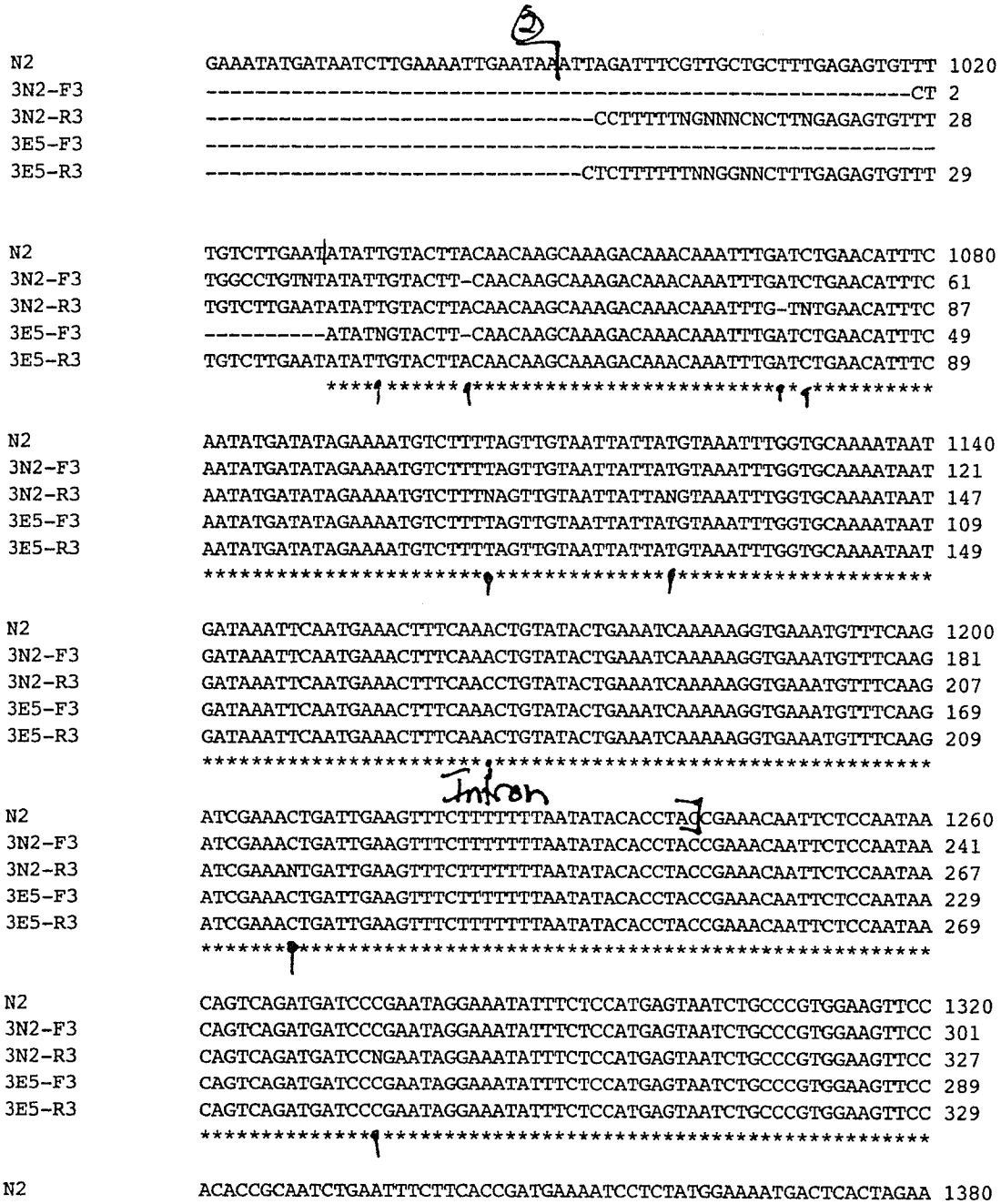

N2 ACTGAATTTCGGAAGTTATCAATTTAAATTTTTCAGTTCCTAGTTCTTACTCTGAGTTTG 240
1N1-F1 ACTGAATTTCNGAAGTTATCAATTTAAATTTTTCAGTTCCTAGTTCTTACTCTGAGTTTG 75
1N1-R1 ATTGAATTTNGGAAGTTATCAATTTAAATTTTTCAGTTCCTAGTTCTTANTNGGAGTTTG 107
1E5-F1 ACTGAATTTCGGAAGTTATCAATTTAAATTTTTCAGTTCCTAGTTCTTACTCTGAGTTTG 74
1E5-R1 ACTGAATTTCGGAAGTTATCAATTTAAATTTTTCAGTTCCTAGTTCTTACTCTGAGTTTG 108
N2 AAGAACACTGCGTTTTCCCCTGGTGGTTCCTTGTGCACATGGCTGAGTATTACAGACTCT 300
1N1-F1 AAGAACACTGCGTTTTCCCCTGGTGGTTCCTTGTGCACATGGCTGAGTATTACAGACTCT 135
1N1-R1 AAGAACANTGNGTTTTCCCCTGGNGGTTCCTTGNGCACATGGGNGAGTATTACNGANTNT 167
1E5-F1 AAGAACACTGCGTTTTCCCCTGGTGGTTCCTTGTGCACATGGCTGAGTATTACAGACTCT 134
1E5-R1 AAGAACACTGCGTTTTCCCCTGGTGGTTCCTTGTGCACATGGNTGAGTATTACAGACTCT 168
N2 TGTTTGTTTAGCTGATCTTCTGGAAAAAAAAACTGTTAAATTAATTAGGACTTTTTAAAG 360
1N1-F1 TGTTTGTTTAGCTGATCTTCTGGAAAAAAAAACTGTTAAATTAATTAGGACTTTTTAAAG 195
1N1-R1 TGTTTGTTTAGNTGATNTTNTGGAAAAAAAAACTGTTAAATTAATTAGGACTTTTTAAAG 227
1E5-F1 TGTTTGTTTAGCTGATCTTCTGGAAAAAAAAACTGTTAAATTAATTAGGACTTTTTAAAG 194
1E5-R1 TGTTTGTTTAGCTGATNTTCTGGAAAAAAAAACTGTTAAATTAATTAGGACTTTTTAAAG 228
N2 CTTACAAGCAATTCGTTGTACTTGTACATCTTTCTCTATACTGAGTACCACAAGAACCAC 420
1N1-F1 CTTACAAGCAATTCGTTGTACTTGTACATCTTTCTCTATACTGAGTACCACAAGAACCAC 255
1N1-R1 CTTACAAGCAATTCGTTGTACTTGTACATCTTTCTCTATACTGAGTACCACAAGAACCAC 287
1E5-F1 CTTACAAGCAATTCGTTGTACTTGTACATCTTTCTCTATACTGAGTACCACAAGAACCAC 254
1E5-R1 CTTACAAGCAATTCGTTGTACTTGTACATCTTTCTCTATACTGAGTACCACAAGAACCAC 288
N2 AATTTTCAGAGCATCTTGTCCAACCGGACCACAGTAAAGAATAACTTCTTGGAGCAGCAA 480
1N1-F1 AATTTTCAGAGCATCTTGTCCAACCGGACCACAGTAAANAATAACTTCTTGGAGCANCAA 315
1N1-R1 AATTTTCAGAGCATNTTGTCCAACCGGACCACAGTAAAGAATAACTTNTTGGAGCAGCAA 347
1E5-F1 AATTTTCAGAGCATCTTGTCCAACCGGACCACAGTAAAGAATAACTTCTTGGAGCAGCAA 314
1E5-R1 AATTTTCAGAGCATNTTGTCCAACCGGACCACAGTAAAGAATAANTTNTTGGAGCAGCAA 348
N2 CCCTGACAGGGGGTGGTCTTTGAGTGGGAGCTGTACCATCTAAAACTTGTCGTGGTTAGA 540
1N1-F1 CCCTGACAGGGGGTGGTCTTTGAGTGGGAGCTGTACCATCTAAAACTTGTCGNGGTTAGA 375
1N1-R1 CCCTGACAGGGGGTGGTCTTTGAGTGGGAGCTGTACCATNTAAAACTTGTCGTGGTTAGA 407
1E5-F1 CCCTGACAGGGGGTGGTCTTTGAGTGGGAGCTGTACCATCTAAAACTTGTCGTGGTTAGA 374
1E5-R1 CCCTGACAGGGGGTGGTCTTTGAGTGGGAGCTGTACCATCTAAAACTTGTCGTGGTTAGA 408
N2 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTT-CTGAG 599
1N1-F1 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTT-CTGAG 434

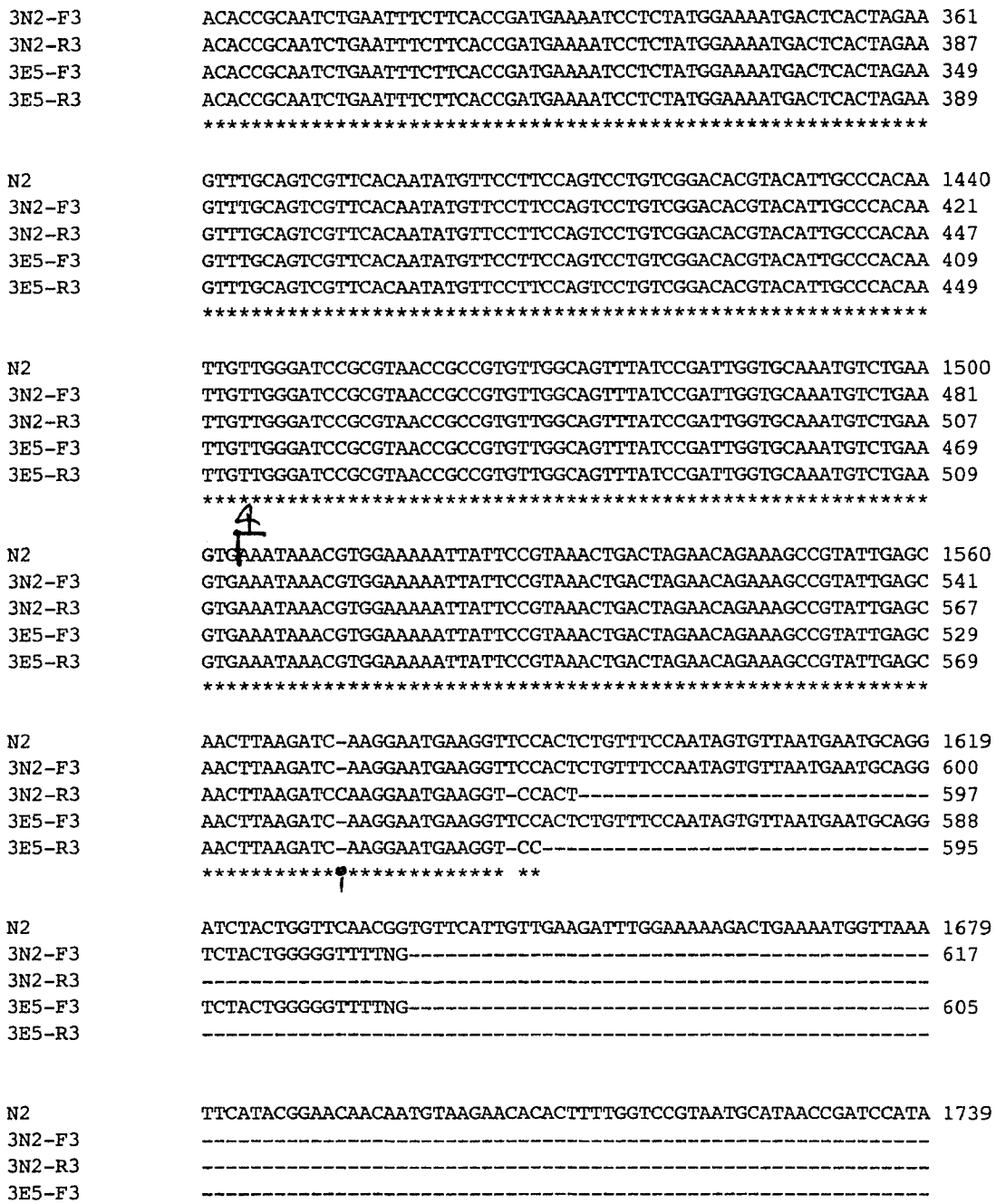

1N1-R1 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTTTCTGAG 467
1E5-F1 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTT-CTGAG 433
1E5-R1 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTTTCTGAG 468
N2 TAGCTTTTGAAATAATCAGAACACTGTTTCCTTTGCTAATTCGTTCACCAAGGACGGCTT 659
1N1-F1 TAGCTTTTGAAATAATCANAACACTGTTTCCTTTGCTAATTCGTTCACCAAGGACNGNTT 494
1N1-R1 TAGCTTTTGAAATAATCAGAACACTGTTTNNATTGTTATT---------- 507
1E5-F1 TAGCTTTTGAAATAATCAGAACACTGTTTCCTTTGCTAATTCGTTCACCAAGGACGGGTT 493
1E5-R1 TAGCTTTTGAAATAATCAGAACACTGTTTCA-TTGTTATTC---------- 508
N2 TGCAACACTGCCGGTATCCGGTTGCTTCATGACTATATTCTGCTTTCATCTGAAAACAGG 719
1N1-F1 TTTCAAAAAAAAA---------- 507
1E5-F1 TTTCAAAANAAAAAA---------- 508
N2 AATGGTATTAAAATAACAAAAAAGGTTAAACAAAATACCTCAACAAACTCTTCACATACT 779
N2 GGAGAACATCTATACATCACATAGGTGAGCTCAAAACGTACATTGCCGCCGTTTGCCTGA 839
N2 AATAAAAATTATTCGAAAAGTAGCAAGCCAGAATTTGATAAAATAGGAACAATTTTAAAA 899
N2 GCGTTCAATGTCGTTCAATTAAGTCGTTCAATAAAGTATCAGGGGCACAGAACTGTTAGA 959

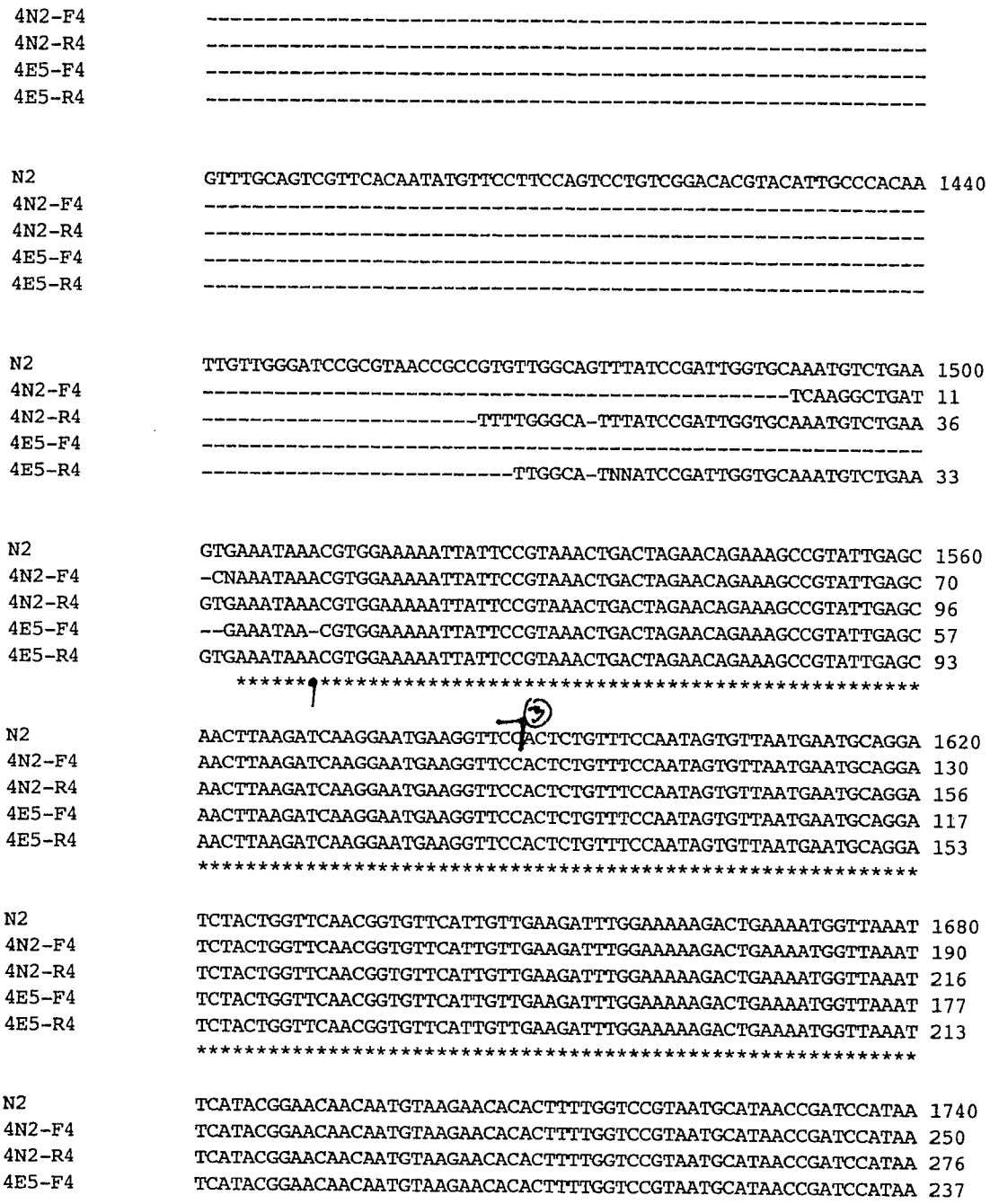

N2 ACTGAATTTCGGAAGTTATCAATTTAAATTTTTCAGTTCCTAGTTCTTACTCTGAGTTTG 240
2N2-F2 ----------
2N2-R2 ----------
2E5-F2 ----------
2E5-R2 ----------
N2 AAGAACACTGCGTTTTCCCCTGGTGGTTCCTTGTGCACATGGCTGAGTATTACAGACTCT 300
2N2-F2 ----------
2N2-R2 ----------
2E5-F2 ----------
2E5-R2 ----------
N2 TGTTTGTTTAGCTGATCTTCTGGAAAAAAAAACTGTTAAATTAATTAGGACTTTTTAAAG 360
2N2-F2 ----------
2N2-R2 ----------
2E5-F2 ----------
2E5-R2 ----------
N2 CTTACAAGCAATTCGTTGTACTTGTACATCTTTCTCTATACTGAGTACCACAAGAACCAC 420
2N2-F2 ----------
2N2-R2 ----------
2E5-F2 ----------
2E5-R2 ----------
N2 AATTTTCAGAGCATCTTGTCCAACCGGACCACAGTAAAGAATAACTTCTTGGAGCAGCAA 480
2N2-F2 ----------
2N2-R2 ----------
2E5-F2 ----------
2E5-R2 ----------
N2 CCCTGACAGGGGGTGGTCTTTGAGTGGGAGCTGTACCATCTAAAACTTGTCGTGGTTAGA 540
2N2-F2 ----------ACATCTAA--CTTGTCGTGGTTAGA 23
2N2-R2 -----TTTCCNNTTGTTCTTGNAGTGGGAGCTGTACCATCTAAAACTTGTCGTGGTTAGA 55
2E5-F2 ----------CATCTAAA-TTTGTCGTGGTTAGA 23
2E5-R2 ----TTTTTCCCTTGTT-TTGNAGTGGGAGCTGTACCATCTAAAACTTGTCGTGGTTAGA 55
******* **********
N2 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTTCTGAGT 600
2N2-F2 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTTCTGAGT 83


2N2-R2 TTCAAAAATCACAAATAATAAACCTNCTNTATATCTGAGTACAAATTGTGAGTTCTGAGT 115
2E5-F2 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTTCTGAGT 83
2E5-R2 TTCAAAAATCACAAATAATAAACCTTCTCTATATCTGAGTACAAATTGTGAGTTCTGAGT 115
N2 AGCTTTTGAAATAATCAGAACACTGTTTCCTTTGCTAATTCGTTCACCAAGGACGGCTTT 660
2N2-F2 AGCTTTTGAAATAATCAGAACACTGTTTCCTTTGCTAATTCGTTCACCAAGGACGGCTTT 143
2N2-R2 AGCTTTTGAAATAATCAGAACACTGTTTCCTTTGNTAATTCGTTCACCAAGGANGGNTTT 175
2E5-F2 AGCTTTTGAAATAATCAGAACACTGTTTCCTTTGCTAATTCGTTCACCAAGGACGGCTTT 143
2E5-R2 AGCTTTTGAAATAATCAGAACACTGTTTCCTTTGNTAATTCGTTCACCAAGGACGGCTTT 175
N2 GCAACACTGCCGGTATCCGGTTGCTTCATGACTATATTCTGCTTTCATCTGAAAACAGGA 720
2N2-F2 GCAACACTGCCGGTATCCGGTTGCTTCATGACTATATTCTGCTTTCATCTGAAAACAGGA 203
2N2-R2 GCAACACTGCCGGTATCCGGTTGCTTCATGACTATATTTTGCTTTCATCTGAAAACAGGA 235
2E5-F2 GCAACACTGCCGGTATCCGGTTGCTTCATGACTATATTCTGCTTTCATCTGAAAACAGGA 203
2E5-R2 GCAACACTGCCGGTATCCGGTTGCTTCATGACTATATTCTGCTTTCATCTGAAAACAGGA 235
N2 ATGGTATTAAAATAACAAAAAAGGTTAAACAAAATACCTCAACAAACTCTTCACATACTG 780
2N2-F2 ATGGTATTAAAATAACAAAAAAGGTTAAACAAAATACCTCAACAAACTCTTCACATACTG 263
2N2-R2 ATGGTATTAAAATAACAAAAAAGGTTAAACAAAATACCTCAACAAACTCTTCACATACTG 295
2E5-F2 ATGGTATTAAAATAACAAAAAAGGTTAAACAAAATACCTCAACAAACTCTTCACATACTG 263
2E5-R2 ATGGTATTAAAATAACAAAAAAGGTTAAACAAAATACCTCAACAAACTCTTCACATACTG 295
Intron
N2 GAGAACATCTATACATCACATAGGTGAGCTCAAAACGTACATTGCCGCCGTTTGCCTGAA 840
2N2-F2 GAGAACATCTATACATCACATAGGTGAGCTCAAAACGTACATTGCCGCCGTTTGCCTGAA 323
2N2-R2 GAGAACATCTATACATCACATAGGTGAGCTCAAAACGTACATTGCCGCCGTTTGCCTGAA 355
2E5-F2 GAGAACATCTATACATCACATAGGTGAGCTCAAAACGTACATTGCCGCCGTTTGCCTGAA 323
2E5-R2 GAGAACATCTATACATCACATAGGTGAGCTCAAAACGTACATTGCCGCCGTTTGCCTGAA 355
N2 ATAAAAATTATTCGAAAAGTAGCAAGCCAGAATTTGATAAAATAGGAACAATTTTAAAAG 900
2N2-F2 ATAAAAATTATTCGAAAAGTAGCAAGCCAGAATTTGATAAAATAGGAACAATTTTAAAAG 383
2N2-R2 ATAAAAATTATTCGAAAAGTAGCAAGCCAGAATTTGATAAAATAGGAACAATTTTAAAAG 415
2E5-F2 ATAAAAATTATTCNAAAAGTAGCAAGCCAGAATTTGATAAAATAGGAACAATTTTAAAAG 383
2E5-R2 ATAAAAATTATTCGAAAAGTAGCAAGCCAGAATTTGATAAAATAGGAACAATTTTAAAAG 415
N2 CGTTCAATGTCGTTCAATTAAGTCGTTCAATAAAGTATC-AGGGGCACAGAACTGTTAGA 959
2N2-F2 CGTTCAATGTCGTTCAATTAAGTCGTTCAATAAAGTATC-AGGGGCACAGAACTGTTAGA 442
2N2-R2 CGTTCAATGTCGTTCAATTAAGTCGTTCAATAAAGTATC-AGGGGCACAGAACTGTTAGA 474
2E5-F2 CGTTCAATGTCGTTCAATTAAGTCGTTCAATAAAGTATC-AGGGGCACAGAACTGTTAGA 442
2E5-R2 CGTTCAATGTCGTTCAATTAAGTCGTTCAATAAAGTATCCAGGGGCACAGAACTGTTAGA 475

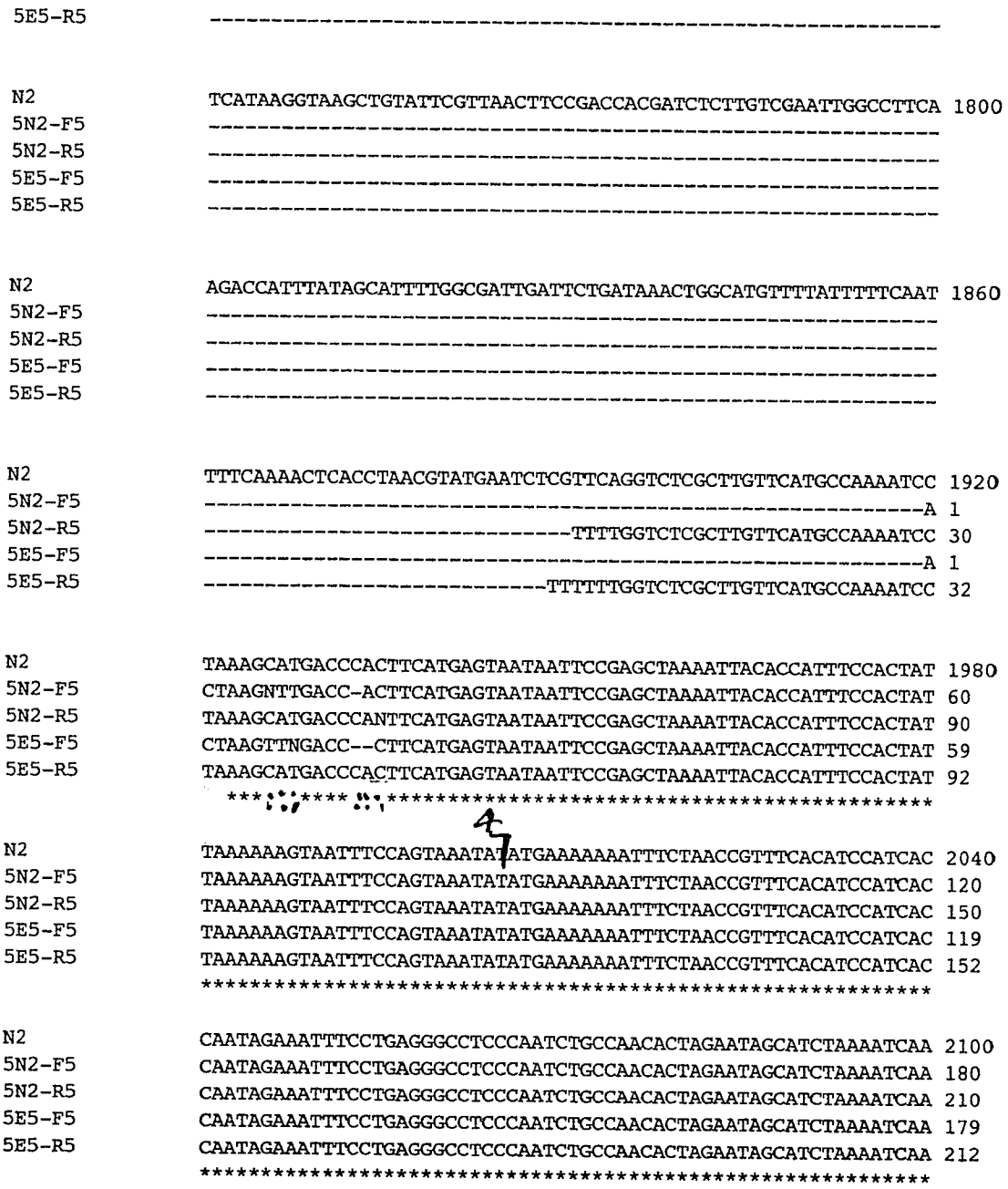

N2 GGAAATATGATAATCTTGAAAATTGAATAAATTAGATTTCGTTGCTGCTTTGAGAGTGTT 1019
2N2-F2 GGAAATATGATAATCTTGAAAATTGAATAAATTAGATTTCGTTGCTGCTTTTNAAAAGG- 501
2N2-R2 GGAAATNTGATAATACTGAAA--TGAATAATNGTT---------- 507
2E5-F2 GGAAATATGATAATCTTGAAAATTGAATAAATTANATTTCGTTGCTGCT.TTANAGTNGA 502
2E5-R2 GGAAATATGATAATCTGAAAA--TNAATAAT---------- 504
N2 TTGTCTTGAATATATTGTACTTACAACAAGCAAAGACAAACAAATTTGATCTGAACATTT 1079
2N2-F2
2N2-R2
2E5-F2 ANAGNANAA---------- 511
2E5-R2
N2 CAATATGATATAGAAAATGTCTTTTAGTTGTAATTATTATGTAAATTTGGTGCAAAATAA 1139
2N2-F2
2N2-R2
2E5-F2
2E5-R2
N2 TGATAAATTCAATGAAACTTTCAAACTGTATACTGAAATCAAAAAGGTGAAATGTTTCAA 1199
2N2-F2
2N2-R2
2E5-F2
2E5-R2
N2 GATCGAAACTGATTGAAGTTTCTTTTTTTTAATATACACCTACCGAAACAATTCTCCAATA 1259
2N2-F2
2N2-R2
2E5-F2
2E5-R2
N2 ACAGTCAGATGATCCCGAATAGGAAATATTTCTCCATGAGTAATCTGCCCGTGGAAGTTC 1319
2N2-F2
2N2-R2
2E5-F2
2E5-R2
N2 CACACCGCAATCTGAATTTCTTCACCGATGAAAATCCTCTATGGAAAATGACTCACTAGA 1379


2
N2 GAAATATGATAATCTTGAAAATTGAATAAATTAGATTTCGTTGCTGCTTTGAGAGTGTTT 1020
3N2-F3 ----------CT 2
3N2-R3 ----------CCTTTTTNGNNNCNCTTNGAGAGTGTTT 28
3E5-F3 ----------
3E5-R3 ----------CTCTTTTTTNNGGNNCTTTGAGAGTGTTT 29
N2 TGTCTTGAATATATTGTACTTACAACAAGCAAAGACAAACAAATTTGATCTGAACATTTC 1080
3N2-F3 TGGCCTGTNTATATTGTACTT-CAACAAGCAAAGACAAACAAATTTGATCTGAACATTTC 61
3N2-R3 TGTCTTGAATATATTGTACTTACAACAAGCAAAGACAAACAAATTTG-TNTGAACATTTC 87
3E5-F3 ----------ATATNGTACTT-CAACAAGCAAAGACAAACAAATTTGATCTGAACATTTC 49
3E5-R3 TGTCTTGAATATATTGTACTTACAACAAGCAAAGACAAACAAATTTGATCTGAACATTTC 89
N2 AATATGATATAGAAAATGTCTTTTAGTTGTAATTATTATGTAAATTTGGTGCAAAATAAT 1140
3N2-F3 AATATGATATAGAAAATGTCTTTTAGTTGTAATTATTATGTAAATTTGGTGCAAAATAAT 121
3N2-R3 AATATGATATAGAAAATGTCTTTTNAGTTGTAATTATTANGTAAATTTGGTGCAAAATAAT 147
3E5-F3 AATATGATATAGAAAATGTCTTTTAGTTGTAATTATTATGTAAATTTGGTGCAAAATAAT 109
3E5-R3 AATATGATATAGAAAATGTCTTTTAGTTGTAATTATTATGTAAATTTGGTGCAAAATAAT 149
N2 GATAAATTCAATGAAACTTTCAAACTGTATACTGAAATCAAAAAGGTGAAATGTTTCAAG 1200
3N2-F3 GATAAATTCAATGAAACTTTCAAACTGTATACTGAAATCAAAAAGGTGAAATGTTTCAAG 181
3N2-R3 GATAAATTCAATGAAACTTTCAACCTGTATACTGAAATCAAAAAGGTGAAATGTTTCAAG 207
3E5-F3 GATAAATTCAATGAAACTTTCAAACTGTATACTGAAATCAAAAAGGTGAAATGTTTCAAG 169
3E5-R3 GATAAATTCAATGAAACTTTCAAACTGTATACTGAAATCAAAAAGGTGAAATGTTTCAAG 209
Intron
N2 ATCGAAACTGATTGAAGTTTCTTTTTTTTAATATACACCTACCGAAACAATTCTCCAATAA 1260
3N2-F3 ATCGAAACTGATTGAAGTTTCTTTTTTTTAATATACACCTACCGAAACAATTCTCCAATAA 241
3N2-R3 ATCGAAANTGATTGAAGTTTCTTTTTTTTAATATACACCTACCGAAACAATTCTCCAATAA 267
3E5-F3 ATCGAAACTGATTGAAGTTTCTTTTTTTTAATATACACCTACCGAAACAATTCTCCAATAA 229
3E5-R3 ATCGAAACTGATTGAAGTTTCTTTTTTTTAATATACACCTACCGAAACAATTCTCCAATAA 269
N2 CAGTCAGATGATCCCGAATAGGAAATATTTCTCCATGAGTAATCTGCCCGTGGAAGTTCC 1320
3N2-F3 CAGTCAGATGATCCCGAATAGGAAATATTTCTCCATGAGTAATCTGCCCGTGGAAGTTCC 301
3N2-R3 CAGTCAGATGATCCNGAATAGGAAATATTTCTCCATGAGTAATCTGCCCGTGGAAGTTCC 327
3E5-F3 CAGTCAGATGATCCCGAATAGGAAATATTTCTCCATGAGTAATCTGCCCGTGGAAGTTCC 289
3E5-R3 CAGTCAGATGATCCCGAATAGGAAATATTTCTCCATGAGTAATCTGCCCGTGGAAGTTCC 329
N2 ACACCGCAATCTGAATTTCTTCACCGATGAAAATCCTCTATGGAAAATGACTCACTAGAA 1380

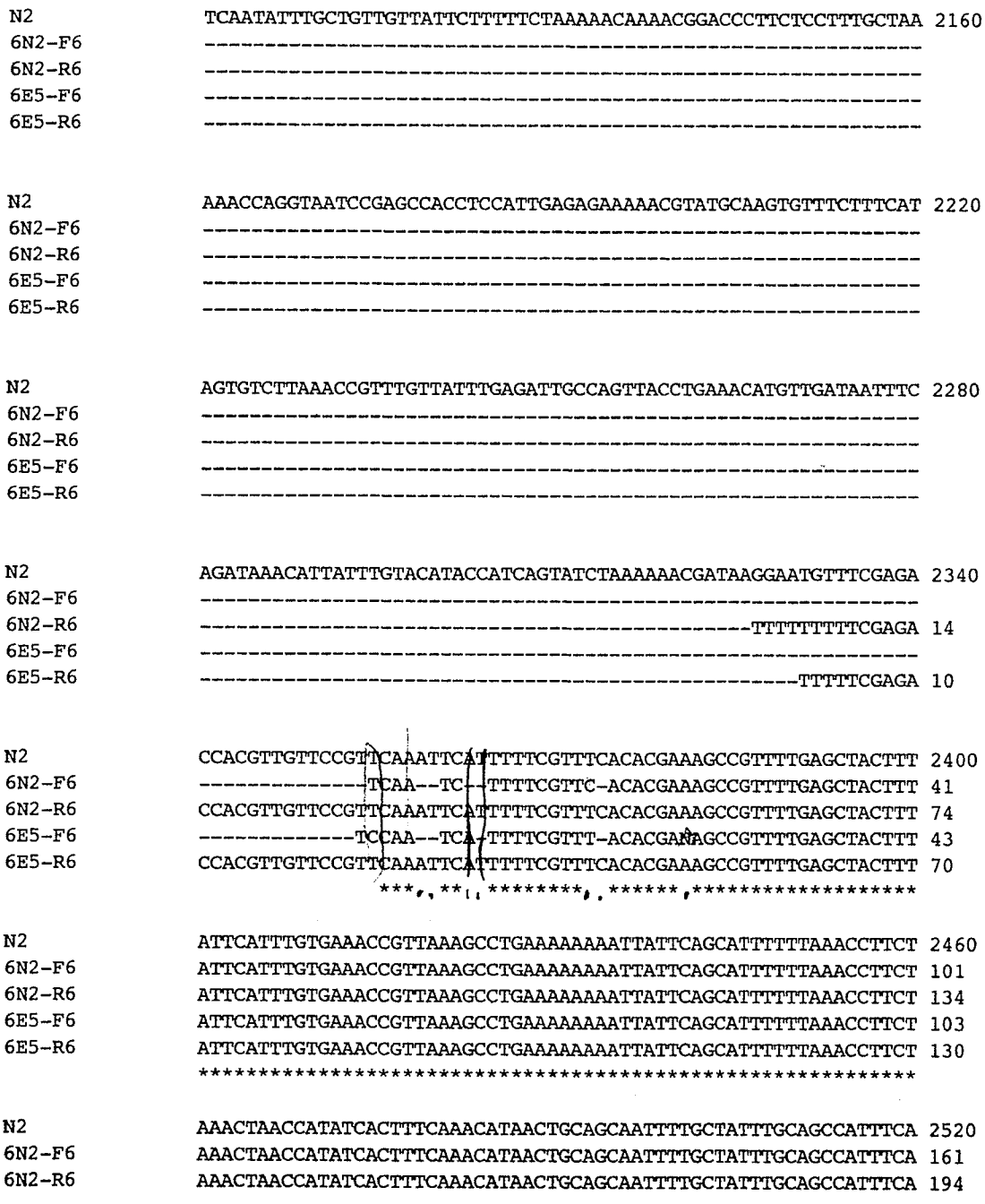

3N2-F3 ACACCGCAATCTGAATTTCTTCACCGATGAAAATCCTCTATGGAAAATGACTCACTAGAA 361
3N2-R3 ACACCGCAATCTGAATTTCTTCACCGATGAAAATCCTCTATGGAAAATGACTCACTAGAA 387
3E5-F3 ACACCGCAATCTGAATTTCTTCACCGATGAAAATCCTCTATGGAAAATGACTCACTAGAA 349
3E5-R3 ACACCGCAATCTGAATTTCTTCACCGATGAAAATCCTCTATGGAAAATGACTCACTAGAA 389
N2 GTTTGCAGTCGTTCACAATATGTTCCTTCCAGTCCTGTCGGACACGTACATTGCCCACAA 1440
3N2-F3 GTTTGCAGTCGTTCACAATATGTTCCTTCCAGTCCTGTCGGACACGTACATTGCCCACAA 421
3N2-R3 GTTTGCAGTCGTTCACAATATGTTCCTTCCAGTCCTGTCGGACACGTACATTGCCCACAA 447
3E5-F3 GTTTGCAGTCGTTCACAATATGTTCCTTCCAGTCCTGTCGGACACGTACATTGCCCACAA 409
3E5-R3 GTTTGCAGTCGTTCACAATATGTTCCTTCCAGTCCTGTCGGACACGTACATTGCCCACAA 449
N2 TTGTTGGGATCCGCGTAACCGCCGTGTTGGCAGTTTATCCGATTGGTGCAAATGTCTGAA 1500
3N2-F3 TTGTTGGGATCCGCGTAACCGCCGTGTTGGCAGTTTATCCGATTGGTGCAAATGTCTGAA 481
3N2-R3 TTGTTGGGATCCGCGTAACCGCCGTGTTGGCAGTTTATCCGATTGGTGCAAATGTCTGAA 507
3E5-F3 TTGTTGGGATCCGCGTAACCGCCGTGTTGGCAGTTTATCCGATTGGTGCAAATGTCTGAA 469
3E5-R3 TTGTTGGGATCCGCGTAACCGCCGTGTTGGCAGTTTATCCGATTGGTGCAAATGTCTGAA 509
4
N2 GTGAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 1560
3N2-F3 GTGAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 541
3N2-R3 GTGAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 567
3E5-F3 GTGAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 529
3E5-R3 GTGAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 569
N2 AACTTAAGATC-AAGGAATGAAGGTTCCACTCTGTTTCCAATAGTGTTAATGAATGCAGG 1619
3N2-F3 AACTTAAGATC-AAGGAATGAAGGTTCCACTCTGTTTCCAATAGTGTTAATGAATGCAGG 600
3N2-R3 AACTTAAGATCCAAGGAATGAAGGT-CCACT---------- 597
3E5-F3 AACTTAAGATC-AAGGAATGAAGGTTCCACTCTGTTTCCAATAGTGTTAATGAATGCAGG 588
3E5-R3 AACTTAAGATC-AAGGAATGAAGGT-CC---------- 595
1
N2 ATCTACTGGTTCAACGGTGTTCATTGTTGAAGATTTGGAAAAAGACTGAAAATGGTTAAA 1679
3N2-F3 TCTACTGGGGGTTTTNG---------- 617
3N2-R3 ----------
3E5-F3 TCTACTGGGGGTTTTNG---------- 605
3E5-R3 ----------
N2 TTCATACGGAACAACAATGTAAGAACACACTTTTGGTCCGTAATGCATAACCGATCCATA 1739
3N2-F3 ----------
3N2-R3 ----------
3E5-F3 ----------

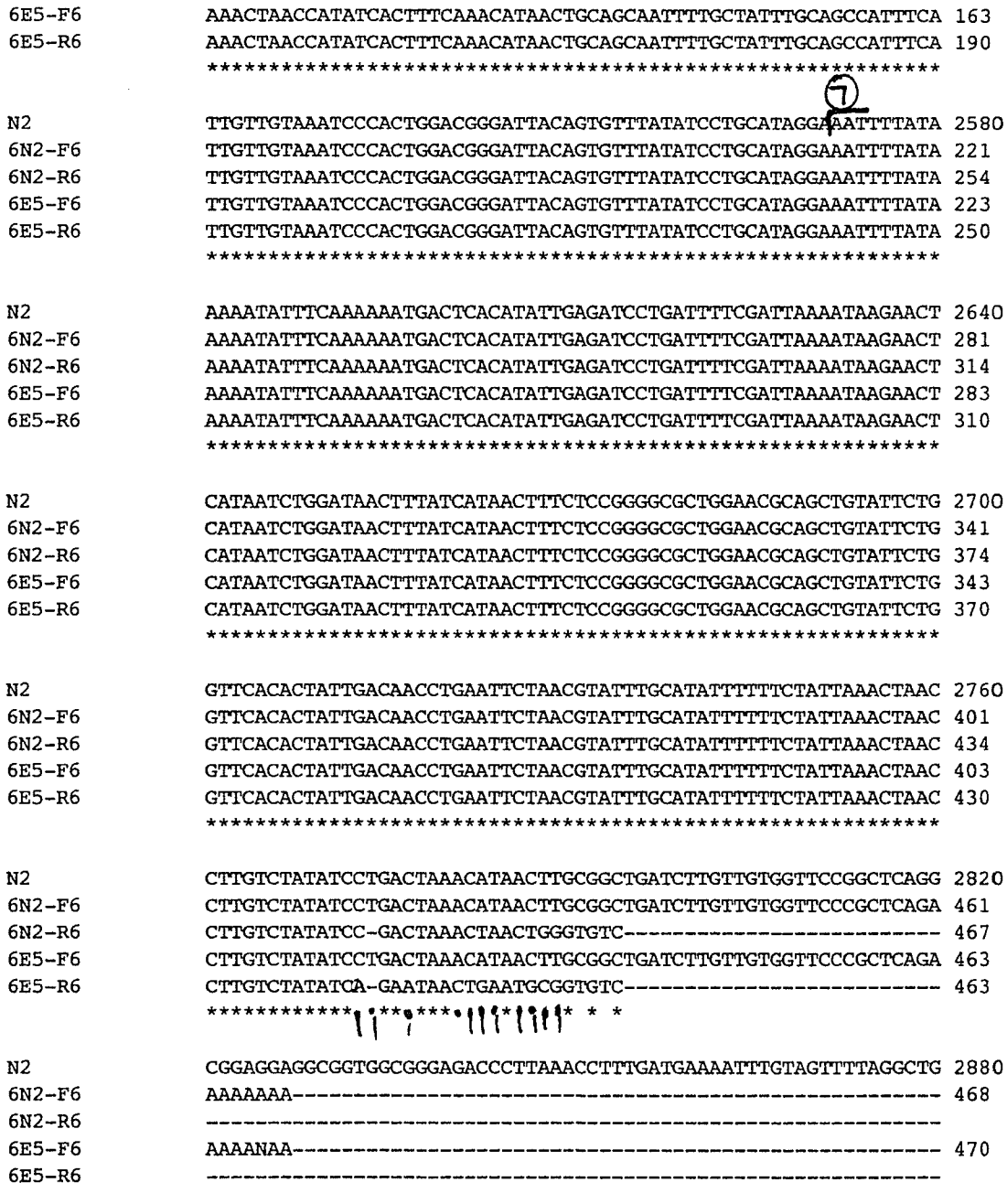

4N2-F4
4N2-R4
4E5-F4
4E5-R4
N2
GTTTGCAGTCGTTCACAATATGTTCCTTCCAGTCCTGTCGGACACGTACATTGCCCACAA 1440
TTGTTGGGATCCGCGTAACCGCCGTGTTGGCAGTTTATCCGATTGGTGCAAATGTCTGAA 1500
-TCAAGGCTGAT 11
-TTTTGGGCA-TTTATCCGATTGGTGCAAATGTCTGAA 36
-TTGGCA-TNNATCCGATTGGTGCAAATGTCTGAA 33
GTGAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 1560
-CNAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 70
GTGAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 96
--GAAATAA-CGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 57
GTGAAATAAACGTGGAAAAATTATTCCGTAAACTGACTAGAACAGAAAGCCGTATTGAGC 93
AACTTAAGATCAAGGAATGAAGGTTCCACTCTGTTTCCAATAGTGTTAATGAATGCAGGA 1620
AACTTAAGATCAAGGAATGAAGGTTCCACTCTGTTTCCAATAGTGTTAATGAATGCAGGA 130
AACTTAAGATCAAGGAATGAAGGTTCCACTCTGTTTCCAATAGTGTTAATGAATGCAGGA 156
AACTTAAGATCAAGGAATGAAGGTTCCACTCTGTTTCCAATAGTGTTAATGAATGCAGGA 117
AACTTAAGATCAAGGAATGAAGGTTCCACTCTGTTTCCAATAGTGTTAATGAATGCAGGA 153
TCTACTGGTTCAACGGTGTTCATTGTTGAAGATTTGGAAAAAGACTGAAAATGGTTAAAT 1680
TCTACTGGTTCAACGGTGTTCATTGTTGAAGATTTGGAAAAAGACTGAAAATGGTTAAAT 190
TCTACTGGTTCAACGGTGTTCATTGTTGAAGATTTGGAAAAAGACTGAAAATGGTTAAAT 216
TCTACTGGTTCAACGGTGTTCATTGTTGAAGATTTGGAAAAAGACTGAAAATGGTTAAAT 177
TCTACTGGTTCAACGGTGTTCATTGTTGAAGATTTGGAAAAAGACTGAAAATGGTTAAAT 213
TCATACGGAACAACAATGTAAGAACACACTTTTGGTCCGTAATGCATAACCGATCCATAA 1740
TCATACGGAACAACAATGTAAGAACACACTTTTGGTCCGTAATGCATAACCGATCCATAA 250
TCATACGGAACAACAATGTAAGAACACACTTTTGGTCCGTAATGCATAACCGATCCATAA 276
TCATACGGAACAACAATGTAAGAACACACTTTTGGTCCGTAATGCATAACCGATCCATAA 237

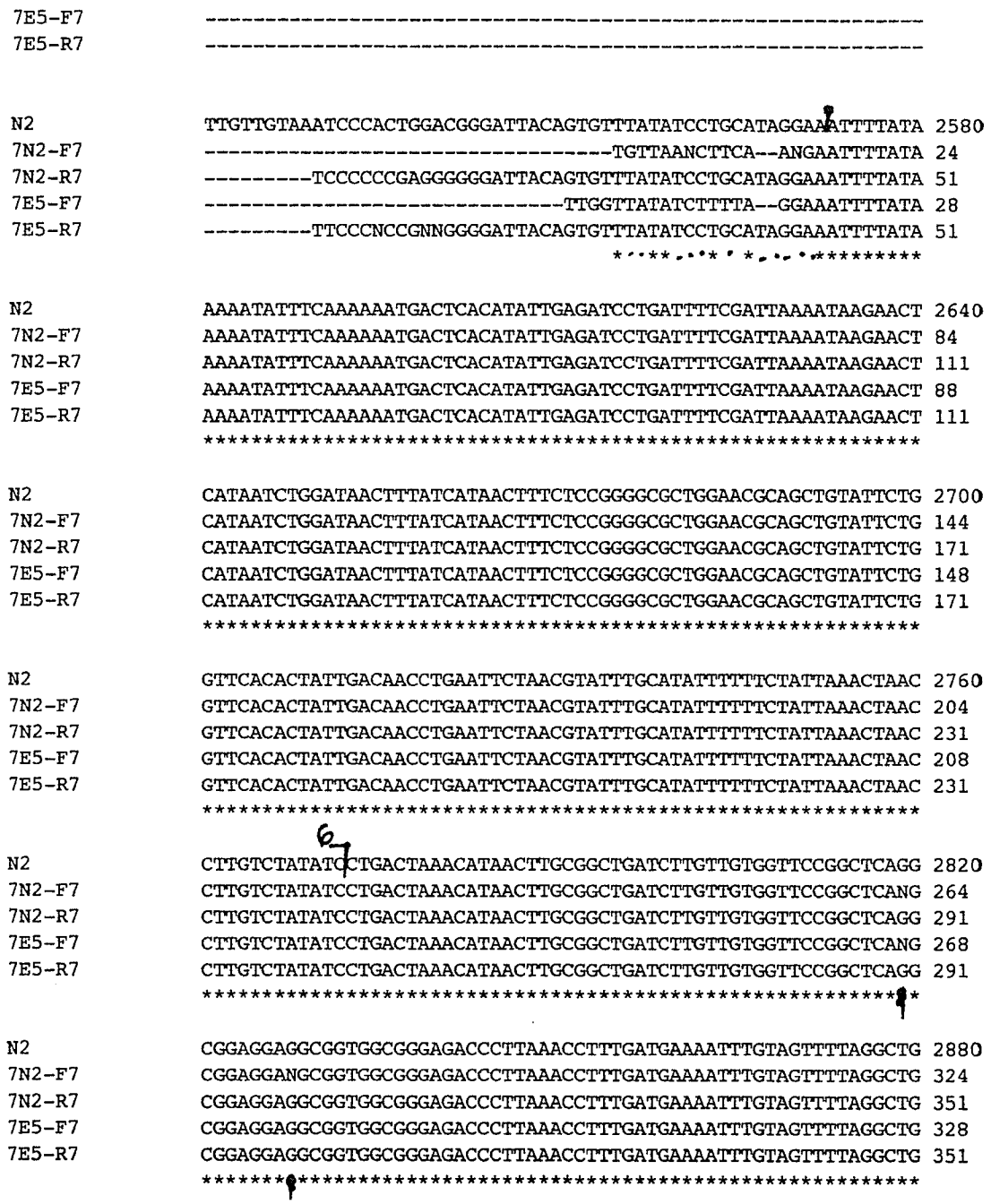
```
4E5-R4      TCATACGGAACAACAATGTAAGAACACACTTTTGGTCCGTAATGCATAACCGATCCATAA 273
            ************************************************************

N2          TCATAAGGTAAGCTGTATTCGTTAACTTCCGACCACGATCTCTTGTCGAATTGGCCTTCA 1800
4N2-F4      TCATAAGGTAAGCTGTATTCGTTAACTTCCGACCACGATCTCTTGTCGAATTGGCCTTCA 310
4N2-R4      TCATAAGGTAAGCTGTATTCGTTAACTTCCGACCACGATCTCTTGTCGAATTGGCCTTCA 336
4E5-F4      TCATAAGGTAAGCTGTATTCGTTAACTTCCGACCACGATCTCTTGTCGAATTGGCCTTCA 297
4E5-R4      TCATAAGGTAAGCTGTATTCGTTAACTTCCGACCACGATCTCTTGTCGAATTGGCCTTCA 333
            ************************************************************

N2          AGACCATTTATAGCATTTTGGCGATTGATTCTGATAAACTGGCATGTTTTATTTTTCAAT 1860
4N2-F4      AGACCATTTATAGCATTTTGGCGATTGATTCTGATAAACTGGCATGTTTTATTTTTCAAT 370
4N2-R4      AGACCATTTATAGCATTTTGGCGATTGATTCTGATAAACTGGCATGTTTTATTTTTCAAT 396
4E5-F4      AGACCATTTATAGCATTTTGGCGATTGATTCTGATAAACTGGCATGTTTTATTTTTCAAT 357
4E5-R4      AGACCATTTATAGCATTTTGGCGATTGATTCTGATAAACTGGCATGTTTTATTTTTCAAT 393
            ************************************************************

N2          TTTCAAAACTCACCTAACGTATGAATCTCGTTCAGGTCTCGCTTGTTCATGCCAAAATCC 1920
4N2-F4      TTTCAAAACTCACCTAACGTATGAATCTCGTTCAGGTCTCGCTTGTTCATGCCAAAATCC 430
4N2-R4      TTTCAAAACTCACCTAACGTATGAATCTCGTTCAGGTCTCGCTTGTTCATGCCAAAATCC 456
4E5-F4      TTTCAAAACTCACCTAACGTATGAATCTCGTTCAGGTCTCGCTTGTTCATGCCAAAATCC 417
4E5-R4      TTTCAAAACTCACCTAACGTATGAATCTCGTTCAGGTCTCGCTTGTTCATGCCAAAATCC 453
            ************************************************************

N2          TAAAGCATGACCCACTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 1980
4N2-F4      TAAAGCATGACCCACTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 490
4N2-R4      TAAAGCATGACCCACTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 516
4E5-F4      TAAAGCATGACCCACTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 477
4E5-R4      TAAAGCATGACCCACTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 513
            ************************************************************

N2          TAAAAAAGTAATTTCCAGTAAATATATGAAAAAAATTTCTAACCGTTTCACATCCATCAC 2040
4N2-F4      TAAAAAAGTAATTTCCAGTAAATATATGAAAAAAATTTCTAACCGTTTCACATTTCACCC 550
4N2-R4      TAAAAAAGTAATTTCNCAAAAATNTTGAAAAATTTCTA---------------------- 554
4E5-F4      TAAAAAAGTAATTTCCAGTAAATATATGAAAAAAATTTCTAACCGTTTCACATTTCANCC 537
4E5-R4      TAAAAAAGTAATTTCCNGNAAATCTATG------------------------------- 541
            ***************** ***** *

N2          CAATAGAAATTTCCTGAGGGCCTCCCAATCTGCCAACACTAGAATAGCATCTAAAATCAA 2100
4N2-F4      CNAAA------------------------------------------------------- 555
4N2-R4      ------------------------------------------------------------
4E5-F4      CAAAA------------------------------------------------------- 542
4E5-R4      ------------------------------------------------------------
```


5E5-R5
N2
5N2-F5
5N2-R5
5E5-F5
5E5-R5
TCATAAGGTAAGCTGTATTCGTTAACTTCCGACCACGATCTCTTGTCGAATTGGCCTTCA 1800
AGACCATTTATAGCATTTTGGCGATTGATTCTGATAAACTGGCATGTTTTATTTTTCAAT 1860
TTTCAAAACTCACCTAACGTATGAATCTCGTTCAGGTCTCGCTTGTTCATGCCAAAATCC 1920
-A 1
TTTTGGTCTCGCTTGTTCATGCCAAAATCC 30
-A 1
TTTTTTGGTCTCGCTTGTTCATGCCAAAATCC 32
TAAAGCATGACCCACTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 1980
CTAAGNTTGACC-ACTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 60
TAAAGCATGACCCANTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 90
CTAAGTTNGACC--CTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 59
TAAAGCATGACCCACTTCATGAGTAATAATTCCGAGCTAAAATTACACCATTTCCACTAT 92
TAAAAAAGTAATTTCCAGTAAATATATGAAAAAAATTTCTAACCGTTTCACATCCATCAC 2040
TAAAAAAGTAATTTCCAGTAAATATATGAAAAAAATTTCTAACCGTTTCACATCCATCAC 120
TAAAAAAGTAATTTCCAGTAAATATATGAAAAAAATTTCTAACCGTTTCACATCCATCAC 150
TAAAAAAGTAATTTCCAGTAAATATATGAAAAAAATTTCTAACCGTTTCACATCCATCAC 119
TAAAAAAGTAATTTCCAGTAAATATATGAAAAAAATTTCTAACCGTTTCACATCCATCAC 152
CAATAGAAATTTCCTGAGGGCCTCCCAATCTGCCAACACTAGAATAGCATCTAAAATCAA 2100
CAATAGAAATTTCCTGAGGGCCTCCCAATCTGCCAACACTAGAATAGCATCTAAAATCAA 180
CAATAGAAATTTCCTGAGGGCCTCCCAATCTGCCAACACTAGAATAGCATCTAAAATCAA 210
CAATAGAAATTTCCTGAGGGCCTCCCAATCTGCCAACACTAGAATAGCATCTAAAATCAA 179
CAATAGAAATTTCCTGAGGGCCTCCCAATCTGCCAACACTAGAATAGCATCTAAAATCAA 212

```
N2          TCAATATTTGCTGTTGTTATTCTTTTTCTAAAAACAAAACGGACCCTTCTCCTTTGCTAA 2160
5N2-F5      TCAATATTTGCTGTTGTTATTCTTTTTCTAAAAACAAAACGGACCCTTCTCCTTTGCTAA 240
5N2-R5      TCAATATTTGCTGTTGTTATTCTTTTTCTAAAAACAAAACGGACCCTTNTCCTTTGCTAA 270
5E5-F5      TCAATATTTGCTGTTGTTATTCTTTTTCTAAAAACAAAACGGACCCTTCTCCTTTGCTAA 239
5E5-R5      TCAATATTTGCTGTTGTTATTCTTTTTCTAAAAACAAAACGGACCCTTCTCCTTTGCTAA 272
            **************************************************.**********

N2          AAACCAGGTAATCCGAGCCACCTCCATTGAGAGAAAAACGTATGCAAGTGTTTCTTTCAT 2220
5N2-F5      AAACCAGGTAATCCGAGCCACCTCCATTGAGAGAAAAACGTATGCAAGTGTTTCTTTCAT 300
5N2-R5      AAACCAGGTAATCCGAGCCACCTCCATTGAGAGAAAAACGTATGCAAGTGTTTCTTTCAT 330
5E5-F5      AAACCAGGTAATCCGAGCCACCTCCATTGAGAGAAAAACGTATGCAAGTGTTTCTTTCAT 299
5E5-R5      AAACCAGGTAATCCGAGCCACCTCCATTGAGAGAAAAACGTATGCAAGTGTTTCTTTCAT 332
            ************************************************************

N2          AGTGTCTTAAACCGTTTGTTATTTGAGATTGCCAGTTACCTGAAACATGTTGATAATTTC 2280
5N2-F5      AGTGTCTTAAACCGTTTGTTATTTGAGATTGCCAGTTACCTGAAACATGTTGATAATTTC 360
5N2-R5      AGTGTCTTAAACCGTTTGTTATTTGAGATTGCCAGTTACCTGAAACATGTTGATAATTTC 390
5E5-F5      AGTGTCTTAAACCGTTTGTTATTTGAGATTGCCAGTTACCTGAAACATGTTGATAATTTC 359
5E5-R5      AGTGTCTTAAACCGTTTGTTATTTGAGATTGCCAGTTACCTGAAACATGTTGATAATTTC 392
            ************************************************************

N2          AGATAAACATTATTTGTACATACCATCAGTATCTAAAAAACGATAAGGAATGTTTCGAGA 2340
5N2-F5      AGATAAACATTATTTGTACATACCATCAGTATCTAAAAAACGATAAGGAATGTTTCGAGA 420
5N2-R5      AGATAAACATTATTTGTACATACCATCAGTATCTAAAAAACGATAAGGAATGTTTCGAGA 450
5E5-F5      AGATAAACATTATTTGTACATACCATCAGTATCTAAAAAACGATAAGGAATGTTTCGAGA 419
5E5-R5      AGATAAACATTATTTGTACATACCATCAGTATCTAAAAAACGATAAGGAATGTTTCGAGA 452
            **********.*************************************************

N2          CCACGTTGTTCCGTTCA--AATTCATTTTTCGTTTCACACGAAAGCCGTTTTGAGCTACT 2398
5N2-F5      CCACGTTGTTCCGTTCA--AATTCATTTTTCGTTTCACACGAAAGCCGTTTTGAGCTACT 478
5N2-R5      CCACGTTGTGAAGTTCA--AATT---------------------------------------- 471
5E5-F5      CCACGTTGTTCCGTTCA--AATTCATTTTTCGTTTCACACGAAAGCCGNTTTGAGCAAAA 477
5E5-R5      CCACGT-GTGAAGTCAAATAATTTTCNTC------------------------------- 480
            ******.*****.*****.*****

N2          TTATTCATTTGTGAAACCGTTAAAGCCTGAAAAAAAATTATTCAGCATTTTTTAAACCTT 2458
5N2-F5      A---------------------------------------------------------- 479
5N2-R5      ----------------------------------------------------------
5E5-F5      ----------------------------------------------------------
5E5-R5      ----------------------------------------------------------

N2          CTAAACTAACCATATCACTTTCAAACATAACTGCAGCAATTTTGCTATTTGCAGCCATTT 2518
5N2-F5      ----------------------------------------------------------
5N2-R5      ----------------------------------------------------------
```


N2 TCAATATTTGCTGTTGTTATTCTTTTTCTAAAAACAAAACGGACCCTTCTCCTTTGCTAA 2160
6N2-F6 ----------
6N2-R6 ----------
6E5-F6 ----------
6E5-R6 ----------
N2 AAACCAGGTAATCCGAGCCACCTCCATTGAGAGAAAAACGTATGCAAGTGTTTCTTTCAT 2220
6N2-F6 ----------
6N2-R6 ----------
6E5-F6 ----------
6E5-R6 ----------
N2 AGTGTCTTAAACCGTTTGTTATTTGAGATTGCCAGTTACCTGAAACATGTTGATAATTTC 2280
6N2-F6 ----------
6N2-R6 ----------
6E5-F6 ----------
6E5-R6 ----------
N2 AGATAAACATTATTTGTACATACCATCAGTATCTAAAAAACGATAAGGAATGTTTCGAGA 2340
6N2-F6 ----------
6N2-R6 ----------TTTTTTTTTCGAGA 14
6E5-F6 ----------
6E5-R6 ----------TTTTTCGAGA 10
N2 CCACGTTGTTCCGTTCAAATTCATTTTTCGTTTCACACGAAAGCCGTTTTGAGCTACTTT 2400
6N2-F6 ----------TCAA--TC--TTTTCGTTC-ACACGAAAGCCGTTTTGAGCTACTTT 41
6N2-R6 CCACGTTGTTCCGTTCAAATTCATTTTTCGTTTCACACGAAAGCCGTTTTGAGCTACTTT 74
6E5-F6 ----------TCCAA--TCA-TTTTCGTTT-ACACGAAAGCCGTTTTGAGCTACTTT 43
6E5-R6 CCACGTTGTTCCGTTCAAATTCATTTTTCGTTTCACACGAAAGCCGTTTTGAGCTACTTT 70
********.**.********.******.******************
N2 ATTCATTTGTGAAACCGTTAAAGCCTGAAAAAAAATTATTCAGCATTTTTTAAACCTTCT 2460
6N2-F6 ATTCATTTGTGAAACCGTTAAAGCCTGAAAAAAAATTATTCAGCATTTTTTAAACCTTCT 101
6N2-R6 ATTCATTTGTGAAACCGTTAAAGCCTGAAAAAAAATTATTCAGCATTTTTTAAACCTTCT 134
6E5-F6 ATTCATTTGTGAAACCGTTAAAGCCTGAAAAAAAATTATTCAGCATTTTTTAAACCTTCT 103
6E5-R6 ATTCATTTGTGAAACCGTTAAAGCCTGAAAAAAAATTATTCAGCATTTTTTAAACCTTCT 130
****
N2 AAACTAACCATATCACTTTCAAACATAACTGCAGCAATTTTGCTATTTGCAGCCATTTCA 2520
6N2-F6 AAACTAACCATATCACTTTCAAACATAACTGCAGCAATTTTGCTATTTGCAGCCATTTCA 161
6N2-R6 AAACTAACCATATCACTTTCAAACATAACTGCAGCAATTTTGCTATTTGCAGCCATTTCA 194

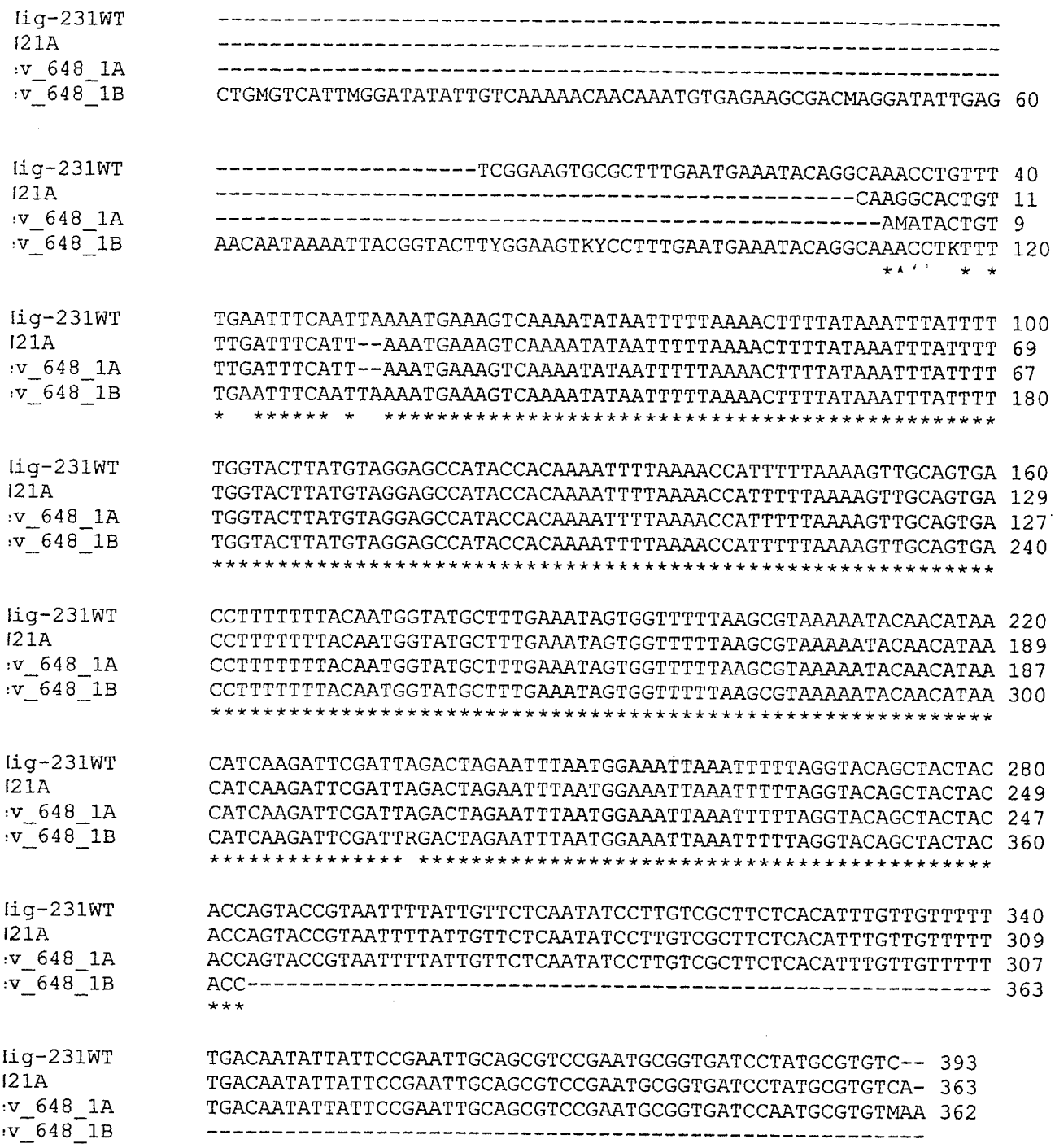

6E5-F6 AAACTAACCATATCACTTTCAAACATAACTGCAGCAATTTTGCTATTTGCAGCCATTTCA 163
6E5-R6 AAACTAACCATATCACTTTCAAACATAACTGCAGCAATTTTGCTATTTGCAGCCATTTCA 190
N2 TTGTTGTAAATCCCACTGGACGGGATTACAGTGTTTATATCCTGCATAGGAAATTTTATA 2580
6N2-F6 TTGTTGTAAATCCCACTGGACGGGATTACAGTGTTTATATCCTGCATAGGAAATTTTATA 221
6N2-R6 TTGTTGTAAATCCCACTGGACGGGATTACAGTGTTTATATCCTGCATAGGAAATTTTATA 254
6E5-F6 TTGTTGTAAATCCCACTGGACGGGATTACAGTGTTTATATCCTGCATAGGAAATTTTATA 223
6E5-R6 TTGTTGTAAATCCCACTGGACGGGATTACAGTGTTTATATCCTGCATAGGAAATTTTATA 250
N2 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 2640
6N2-F6 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 281
6N2-R6 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 314
6E5-F6 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 283
6E5-R6 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 310
N2 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 2700
6N2-F6 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 341
6N2-R6 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 374
6E5-F6 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 343
6E5-R6 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 370
N2 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 2760
6N2-F6 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 401
6N2-R6 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 434
6E5-F6 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 403
6E5-R6 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 430
N2 CTTGTCTATATCCTGACTAAACATAACTTGCGGCTGATCTTGTTGTGGTTCCGGCTCAGG 2820
6N2-F6 CTTGTCTATATCCTGACTAAACATAACTTGCGGCTGATCTTGTTGTGGTTCCCGCTCAGA 461
6N2-R6 CTTGTCTATATCC-GACTAAACTAACTGGGTGTC---------- 467
6E5-F6 CTTGTCTATATCCTGACTAAACATAACTTGCGGCTGATCTTGTTGTGGTTCCCGCTCAGA 463
6E5-R6 CTTGTCTATATCA-GAATAACTGAATGCGGTGTC---------- 463
N2 CGGAGGAGGCGGTGGCGGGAGACCCTTAAACCTTTGATGAAAATTTGTAGTTTTAGGCTG 2880
6N2-F6 AAAAAAA---------- 468
6N2-R6 ----------
6E5-F6 AAAANAA---------- 470
6E5-R6 ----------


7E5-F7
7E5-R7
N2 TTGTTGTAAATCCCACTGGACGGGATTACAGTGTTTATATCCTGCATAGGAAATTTTATA 2580
7N2-F7 -----TGTTAANCTTCA--ANGAATTTTATA 24
7N2-R7 ---------TCCCCCCGAGGGGGGATTACAGTGTTTATATCCTGCATAGGAAATTTTATA 51
7E5-F7 ----------TTGGTTATATCTTTTA--GGAAATTTTATA 28
7E5-R7 ---------TTCCCNCCGNNGGGGATTACAGTGTTTATATCCTGCATAGGAAATTTTATA 51
N2 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 2640
7N2-F7 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 84
7N2-R7 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 111
7E5-F7 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 88
7E5-R7 AAAATATTTCAAAAAATGACTCACATATTGAGATCCTGATTTTCGATTAAAATAAGAACT 111
N2 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 2700
7N2-F7 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 144
7N2-R7 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 171
7E5-F7 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 148
7E5-R7 CATAATCTGGATAACTTTATCATAACTTTCTCCGGGGCGCTGGAACGCAGCTGTATTCTG 171
N2 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 2760
7N2-F7 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 204
7N2-R7 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 231
7E5-F7 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 208
7E5-R7 GTTCACACTATTGACAACCTGAATTCTAACGTATTTGCATATTTTTTCTATTAAACTAAC 231
N2 CTTGTCTATATCCTGACTAAACATAACTTGCGGCTGATCTTGTTGTGGTTCCGGCTCAGG 2820
7N2-F7 CTTGTCTATATCCTGACTAAACATAACTTGCGGCTGATCTTGTTGTGGTTCCGGCTCANG 264
7N2-R7 CTTGTCTATATCCTGACTAAACATAACTTGCGGCTGATCTTGTTGTGGTTCCGGCTCAGG 291
7E5-F7 CTTGTCTATATCCTGACTAAACATAACTTGCGGCTGATCTTGTTGTGGTTCCGGCTCANG 268
7E5-R7 CTTGTCTATATCCTGACTAAACATAACTTGCGGCTGATCTTGTTGTGGTTCCGGCTCAGG 291
N2 CGGAGGAGGCGGTGGCGGGAGACCCTTAAACCTTTGATGAAAATTTGTAGTTTTAGGCTG 2880
7N2-F7 CGGAGGANGCGGTGGCGGGAGACCCTTAAACCTTTGATGAAAATTTGTAGTTTTAGGCTG 324
7N2-R7 CGGAGGAGGCGGTGGCGGGAGACCCTTAAACCTTTGATGAAAATTTGTAGTTTTAGGCTG 351
7E5-F7 CGGAGGAGGCGGTGGCGGGAGACCCTTAAACCTTTGATGAAAATTTGTAGTTTTAGGCTG 328
7E5-R7 CGGAGGAGGCGGTGGCGGGAGACCCTTAAACCTTTGATGAAAATTTGTAGTTTTAGGCTG 351

```
N2          GTTACCCATGGAGGTCTTCTATGCCATGGAGGTCTTCGCCAAGGTGGCGGTGGTCGATCC 2940
7N2-F7      GTTACCCATGGAGGTCTTCTATGCCATGGAGGTCTTCGCCAAGGTGGCGGTGGTCGATCC 384
7N2-R7      GTTACCCATGGAGGTCTTCTATGCCATGGAGGTNTTCGCCAAGGTGGCGGTGGTCGATCC 411
7E5-F7      GTTACCCATGGAGGTCTTCTATGCCATGGAGGTCTTCGCCAAGGTGGCGGTGGTCGATCC 388
7E5-R7      GTTACCCATGGAGGTCTTCTATGCCATGGAGGTNTTCGCCAAGGTGGCGGTGGTCGATCC 411
            ********************************************************** 

N2          CATGGAGGTGGGGGTCCGAATGGCGGAGGACCTCGGTTTCTAAAACTTTGTTTTAATATG 3000
7N2-F7      CATGGAGGTGGGGGTCCGAATGGCGGANGACCTCGGTTTCTAAAACTTTGTTTTAATATG 444
7N2-R7      CATGGAGGTGGGGGTCCGAATGGCGGAGGACCTCGGTTTCTAAAACTTTGTTTTAATATG 471
7E5-F7      CATGGAGGTGGGGGTCCGAATGGCGGAGGACCTCGGTTTCTAAAACTTTGTTTTAATATG 448
7E5-R7      CATGGAGGTGGGGGTCCGAATGGCGGAGGACCTCGGTTTCTAAAACTTTGTTTTAATATG 471
            ************************************************************ 

N2          AAAACCTTGGCATTCTACACGTGCGAAAACGTGCTTTTACAAAGTTCTATTTATATATTC 3060
7N2-F7      AAAACCTTGGCATTCTACACGTGCGAAAACGTGCTTTTACAAAGTTCTATTTATATATTC 504
7N2-R7      AAAACCTTGGCATTCTACACGTGCGAAAACGTGCTTTTACAAAGTTCTATTTATATATTC 531
7E5-F7      AAAACCTTGGCATTCTACACGTGCGAAAACGTGCTTTTACAAAGTTCTATTTATATATTC 508
7E5-R7      AAAACCTTGGCATTCTACACGTGCGAAAACGTGCTTTTACAAAGTTCTATTTATATATTC 531
            ************************************************************ 

N2          TTGGATTAGCATGCTGATTTTTTCAAATTCAGTTATGTTTGCAATTCTTCTGGAATCAAA 3120
7N2-F7      TTGGATTAGCATGCTGATTTTTTCAAATTCAGTTATGTTTGCAATTCTTCTGGAATCAAA 564
7N2-R7      TTGGATTAGCATGCTG-TTTTTTCAAATTCAGTTATNG-----AACATT--TTNANTCAAG 584
7E5-F7      TTGGATTAGCATGCTGATTTTTTCAAATTCAGTTATGTTTGCAATTCTTCTGGAATCAAA 568
7E5-R7      TTGGATTAGCATGCTGATTTTTTCAAATTCAGTTATNT----AACATTCNTNAANTNAAG 587
            ****************************************      **   *  *  * * ** 

N2          GTGAAATCTCACCGATGCCAAGGTGGAGGCGGTCCCCAAGGCCGAAGAGGTCCTGGAAAA 3180
7N2-F7      GTGAAATCTCACC--CNCNAGCGNGGAANA---------------------------- 592
7N2-R7      G------------------------------------------------------------ 585
7E5-F7      GTGAAATCTCACC--CCCNAGCGGGGAANA---------------------------- 596
7E5-R7      G------------------------------------------------------------ 588
            *

N2          AACCAATCTGTAAACCGGAACGTATTTCATTATTTTTATACTTCCGTTTTTATCAAATTC 3240
7N2-F7      ------------------------------------------------------------
7N2-R7      ------------------------------------------------------------
7E5-F7      ------------------------------------------------------------
7E5-R7      ------------------------------------------------------------

N2          CAGGCAAAATTTTTCAATTTTCAGATAAAAAAATAAGTAAGTGTCAGCTGATGGCGATAA 3300
7N2-F7      ------------------------------------------------------------
```

6.5 Alignments of *mig-23* sequencing results.

```
lig-231WT       ------------------------------------------------------------
l21A            ------------------------------------------------------------
:v_648_1A       ------------------------------------------------------------
:v_648_1B       CTGMGTCATTMGGATATATTGTCAAAAACAACAAATGTGAGAAGCGACMAGGATATTGAG 60


lig-231WT       --------------------TCGGAAGTGCGCTTTGAATGAAATACAGGCAAACCTGTTT 40
l21A            ------------------------------------------------CAAGGCACTGT 11
:v_648_1A       ------------------------------------------------AMATACTGT 9
:v_648_1B       AACAATAAAATTACGGTACTTYGGAAGTKYCCTTTGAATGAAATACAGGCAAACCTKTTT 120
                                                                 *A'  * *


lig-231WT       TGAATTTCAATTAAAATGAAAGTCAAAATATAATTTTTAAAACTTTTATAAATTTATTTT 100
l21A            TTGATTTCATT--AAATGAAAGTCAAAATATAATTTTTAAAACTTTTATAAATTTATTTT 69
:v_648_1A       TTGATTTCATT--AAATGAAAGTCAAAATATAATTTTTAAAACTTTTATAAATTTATTTT 67
:v_648_1B       TGAATTTCAATTAAAATGAAAGTCAAAATATAATTTTTAAAACTTTTATAAATTTATTTT 180
                *  ****** *  **************************************************


lig-231WT       TGGTACTTATGTAGGAGCCATACCACAAAATTTTAAAACCATTTTTAAAAGTTGCAGTGA 160
l21A            TGGTACTTATGTAGGAGCCATACCACAAAATTTTAAAACCATTTTTAAAAGTTGCAGTGA 129
:v_648_1A       TGGTACTTATGTAGGAGCCATACCACAAAATTTTAAAACCATTTTTAAAAGTTGCAGTGA 127
:v_648_1B       TGGTACTTATGTAGGAGCCATACCACAAAATTTTAAAACCATTTTTAAAAGTTGCAGTGA 240
                ************************************************************


lig-231WT       CCTTTTTTTACAATGGTATGCTTTGAAATAGTGGTTTTTAAGCGTAAAAATACAACATAA 220
l21A            CCTTTTTTTACAATGGTATGCTTTGAAATAGTGGTTTTTAAGCGTAAAAATACAACATAA 189
:v_648_1A       CCTTTTTTTACAATGGTATGCTTTGAAATAGTGGTTTTTAAGCGTAAAAATACAACATAA 187
:v_648_1B       CCTTTTTTTACAATGGTATGCTTTGAAATAGTGGTTTTTAAGCGTAAAAATACAACATAA 300
                ************************************************************


lig-231WT       CATCAAGATTCGATTAGACTAGAATTTAATGGAAATTAAATTTTTAGGTACAGCTACTAC 280
l21A            CATCAAGATTCGATTAGACTAGAATTTAATGGAAATTAAATTTTTAGGTACAGCTACTAC 249
:v_648_1A       CATCAAGATTCGATTAGACTAGAATTTAATGGAAATTAAATTTTTAGGTACAGCTACTAC 247
:v_648_1B       CATCAAGATTCGATTRGACTAGAATTTAATGGAAATTAAATTTTTAGGTACAGCTACTAC 360
                *************** ********************************************


lig-231WT       ACCAGTACCGTAATTTTATTGTTCTCAATATCCTTGTCGCTTCTCACATTTGTTGTTTTT 340
l21A            ACCAGTACCGTAATTTTATTGTTCTCAATATCCTTGTCGCTTCTCACATTTGTTGTTTTT 309
:v_648_1A       ACCAGTACCGTAATTTTATTGTTCTCAATATCCTTGTCGCTTCTCACATTTGTTGTTTTT 307
:v_648_1B       ACC---------------------------------------------------------- 363
                ***


lig-231WT       TGACAATATTATTCCGAATTGCAGCGTCCGAATGCGGTGATCCTATGCGTGTC-- 393
l21A            TGACAATATTATTCCGAATTGCAGCGTCCGAATGCGGTGATCCTATGCGTGTCA- 363
:v_648_1A       TGACAATATTATTCCGAATTGCAGCGTCCGAATGCGGTGATCCAATGCGTGTMAA 362
:v_648_1B       ----------------------------------------------------
```

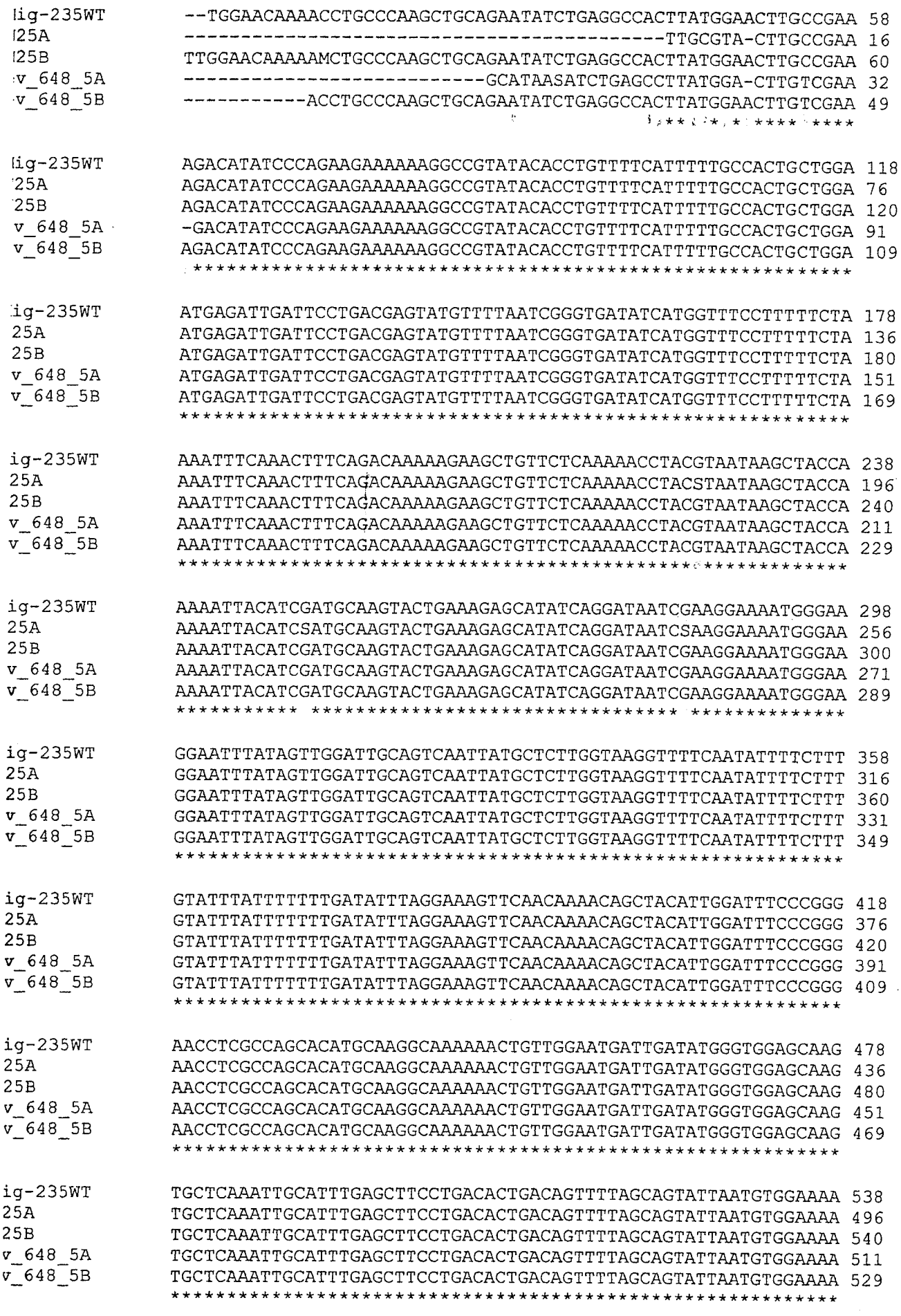

lig-231WT
l21A
:v_648_1A
:v_648_1B
CTGMGTCATTMGGATATATTGTCAAAAACAACAAATGTGAGAAGCGACMAGGATATTGAG 60
TCGGAAGTGCGCTTTGAATGAAATACAGGCAAACCTGTTT 40
CAAGGCACTGT 11
AMATACTGT 9
AACAATAAAATTACGGTACTTYGGAAGTKYCCTTTGAATGAAATACAGGCAAACCTKTTT 120
TGAATTTCAATTAAAATGAAAGTCAAAATATAATTTTTAAAACTTTTATAAATTTATTTT 100
TTGATTTCATT--AAATGAAAGTCAAAATATAATTTTTAAAACTTTTATAAATTTATTTT 69
TTGATTTCATT--AAATGAAAGTCAAAATATAATTTTTAAAACTTTTATAAATTTATTTT 67
TGAATTTCAATTAAAATGAAAGTCAAAATATAATTTTTAAAACTTTTATAAATTTATTTT 180
TGGTACTTATGTAGGAGCCATACCACAAAATTTTAAAACCATTTTTAAAAGTTGCAGTGA 160
TGGTACTTATGTAGGAGCCATACCACAAAATTTTAAAACCATTTTTAAAAGTTGCAGTGA 129
TGGTACTTATGTAGGAGCCATACCACAAAATTTTAAAACCATTTTTAAAAGTTGCAGTGA 127
TGGTACTTATGTAGGAGCCATACCACAAAATTTTAAAACCATTTTTAAAAGTTGCAGTGA 240
CCTTTTTTTACAATGGTATGCTTTGAAATAGTGGTTTTTAAGCGTAAAAATACAACATAA 220
CCTTTTTTTACAATGGTATGCTTTGAAATAGTGGTTTTTAAGCGTAAAAATACAACATAA 189
CCTTTTTTTACAATGGTATGCTTTGAAATAGTGGTTTTTAAGCGTAAAAATACAACATAA 187
CCTTTTTTTACAATGGTATGCTTTGAAATAGTGGTTTTTAAGCGTAAAAATACAACATAA 300
CATCAAGATTCGATTAGACTAGAATTTAATGGAAATTAAATTTTTAGGTACAGCTACTAC 280
CATCAAGATTCGATTAGACTAGAATTTAATGGAAATTAAATTTTTAGGTACAGCTACTAC 249
CATCAAGATTCGATTAGACTAGAATTTAATGGAAATTAAATTTTTAGGTACAGCTACTAC 247
CATCAAGATTCGATTRGACTAGAATTTAATGGAAATTAAATTTTTAGGTACAGCTACTAC 360
ACCAGTACCGTAATTTTATTGTTCTCAATATCCTTGTCGCTTCTCACATTTGTTGTTTTT 340
ACCAGTACCGTAATTTTATTGTTCTCAATATCCTTGTCGCTTCTCACATTTGTTGTTTTT 309
ACCAGTACCGTAATTTTATTGTTCTCAATATCCTTGTCGCTTCTCACATTTGTTGTTTTT 307
ACC 363
TGACAATATTATTCCGAATTGCAGCGTCCGAATGCGGTGATCCTATGCGTGTC-- 393
TGACAATATTATTCCGAATTGCAGCGTCCGAATGCGGTGATCCTATGCGTGTCA- 363
TGACAATATTATTCCGAATTGCAGCGTCCGAATGCGGTGATCCAATGCGTGTMAA 362

```
lig-23WT     -TCCGAATTGCAGCGTCCGAATGCGGTGATCCTATGCGTGTCAGTCTTAGATTCACCATT 59
l2-2A        --------------------------------GGCAKCGTGTM-GTCTTAGATTC-CCAYT 27
l2-2B        TTYCGAATTGCAGCGTCCGAATGCGGTGATCCTATGCGTGTCAGTYTTAGATTCACCATT 60
:v6482A      ---------------------------GACGTAATGCGTGTM-GTCTTAGATTM-CCATT 31
:v6482B      -------------SGTCCGAATGCGGKGATCCTATGCGKGTCAGTYTTAGATTCACCATT 47
                                                 **, **  **  *******   *** *

lig-23WT     CTTGCCGTTTCGGCAATGATATTCTTTCCAGTTATTGTATTTATTTATGTGGTAGAAGCA 119
l2-2A        CTTGCCGTTTCGGCAATGATATTCYTTCCMGTTATTGTATTTATTTATGTGGTAGAAGCM 87
l2-2B        CTTGCCGTTTYGGCAATGATATTCTTTCCAGTTATTGTATTTATTTATGTGGTAGAAGCA 120
:v6482A      CTTGCCGTTTCGGCAATGATATTCTTTCCAGTTATTGTATTTATTTATGTGGTAGAAGCA 91
:v6482B      CTTGCCGTTTCGGCAATGATATTCTTTCCAGTTATTGTATTTATTTATGTGGTAGAAGCA 107
             ********** ************* **** ******************************

lig-23WT     CACACATCTCCAAAAGTGATAGCAGATGGTGAGTTGTTTGGATTTTTGAAGTTTTTATTA 179
l2-2A        CMCMCMTCYCCMAAAGTGATAGCMGATGGTGAGTTGTTTGGATTTTTGAAGTTTTTATTA 147
l2-2B        CACACATCTCCAAAAGTGATAGCAGATGGTGAGTTGTTTGGATTTTTGAAGTTTTTWTTA 180
:v6482A      CACACATCTCCAAAAGTGATAGCAGATGGTGAGTTGTTTGGATTTTTGAAGTTTTTATTA 151
:v6482B      CACACATCTCCAAAAGTGATAGCAGATGGTGAGTTGTTTGGATTTTTGAAGTTTTTWTTA 167
             * * * ** ** *********** ************************************

lig-23WT     TACAAAAGTATAAAAAAGATTATAAATAATTTTTGGAAAAAAAACTTTGATACATAAATT 239
l2-2A        TACMAAAGTATAAAAAAGATTATAAATAATTTTTGGAAAAAAAACTTTGATACMTAAATT 207
l2-2B        TACAAAAGTATAAAAAAGATTATAAATAATTTTTGGAAAAAAAACTTTGATACATAAATT 240
:v6482A      TACAAAAGTATAAAAAAGATTATAAATAATTTTTGGAAAAAAAACTTTGATACATAAATT 211
:v6482B      TACAAAAGTATAAAAAAGATTATAAATAATTTTTGGAAAAAAAACTTTGATACATAAATT 227
             *** ************************************************* ******

lig-23WT     TGAAACAATAGAATCCGTATTATCAGAACCTATTGAAAAACATCATTAATAACGTGTAAT 299
l2-2A        TGAAACAATAGAATCCGTATTATCARAACCYATTGAAAAACMTCATTAATAACGTGTAAT 267
l2-2B        TGAAACAATAGAATCCGTATTATCAGAACCTATTGAAAAACATCATTAATAACGTGTAAT 300
:v6482A      TGAAACAATAGAATCCGTATTATCARAACCTATTGAAAAACATCATTAATAACGTGTAAT 271
:v6482B      TGAAACAATAGAATCCGTATTATCAGAACCTATTGAAAAACATCATTAATAACGKGTAAT 287
             ************************* **** *********** ************ *****

lig-23WT     TGGCTAATAGAAACGTTAACTAATTAAGTAATTTCGATAACTAACCGAGCAGGACCGTGT 359
l2-2A        TGGCTAATARAAACGTTAACTAATTAAGTAATTTCGATAACTAACCGAGCAGGACCGTGT 327
l2-2B        TGGCTAATAGAAACGTTAACTAATTAAGTAATTTCGATAACTAACCGAGCAGGACCGTGT 360
:v6482A      TGGCTAAWARAAACGTTAACTAATTAAGTAATTTCRATAACTAACCGAGCAGGACCGTGT 331
:v6482B      KGGCTAATAGAAACGTTAACTAATTAAGTAATTTSGATAACTAACCGAGCAGGACCGKGK 347
              ******* * ************************ ********************** * 

lig-23WT     GTAACTTCCCTGAAAAACCGCCGAAAGTATTGTTAGTTTTTGAAATTCTTTGCAAACTTT 419
l2-2A        GTAACTTCCCTGAAAAACCGCCRAAAGTATTGTTAGTTTTTGAAATTCTTTGCAAACTTT 387
l2-2B        GTAACTTCCCTGAAAAACCGCCGAAAGTATTGTTAGTTTTTGAAATTCTTTGCAAACTTT 420
:v6482A      GTAACTTCCCTGAAAAACCGCCRAAAGTATTGTTAKTTTTTGAAATTCTTTGCAAACTTT 391
:v6482B      GTAACTTCCCTGAAAAACCGCCGAAAGTATTGTTAGTTTTTGAAATTCTTTGCAAACTTT 407
             ********************** ************ ************************

lig-23WT     ACTTTAAATGATTTCAAAAATAAAAAAACAATTAAAGCTTCACAGACTTTGTTAGTTTTG 479
l2-2A        ACTTTAAATGATTTCAAAAATAAAAAAACAATTAAAGCTTCMCAGACTTTGTTAGTTTTG 447
l2-2B        ACTTTAAATGATT-CAAAAATAAAAAAACAAT-AAAGCT-CACAGACTTT---------- 467
:v6482A      ACTTTAAATGATTTCAAAAATAAAAAAACAATTAAAGCTTCACARACTTTGTTAGTTTTG 451
:v6482B      ACTTTAAATGATTTCAAAAATAAAAAAACAAT-AAAGCT-CACAGACTT----------- 454
             ************* ****************** ****** * ** ****

lig-23WT     GTGGAAATCGGTCGTTCAG- 498
l2-2A        GTGGAATG------------ 455
l2-2B        --------------------
:v6482A      GTGGAAATCGGTCGTTCAGA 471
:v6482B      --------------------
```

```
lig-233WT       TCCCTGAAAAACCGCCGAAAGTATTGTTAGTTTTTGAAAT---TCTTTGCAAACTTTACT 57
l23A            ----------------------------GGGTKGCMT-GA---TCTTTGCAA-CTTTACT 27
l23B            --TTCCCTGAAAAACCGCCGAAAGTATTGTTAGTTTTTGAAATTCTTTGCAAACTTTACT 58
:v6483A         ----------------------------GCGCARCTTTGA---TCTTTGCAA-CTTTACT 28
:v6483B         -------------CCGCCGAAAGTATTGTTAGTTTTTGAAATTCTTTGCAAACTTTACT 46
                                                          ********:*******

lig-233WT       TTAAATGATTTCAAAAATAAAAAAACAATTAAAGCTTCACAGACTTTGTTAGTTTTGGTG 117
l23A            TTAA-TGATTTCAAAAATAAAAAAACAATTAAAGCTTCACAGACTTTGTTAGTTTTGGTG 86
l23B            TTAAATGATTTCAAAAATAAAAAAACAATTAAAGCTTCACAGACTTTGTTAGTTTTGGTG 118
:v6483A         TTAAATGATTTCAAAAATAAAAAAACAATTAAAGCTTCACAGACTTTGTTAGTTTTGGTG 88
:v6483B         TTAAATGATTTCAAAAATAAAAAAACAATTAAAGCTTCACAGACTTTGTTAGTTTTGGTG 106
                ****:**********************************************************

lig-233WT       GAAATCGGTCGTTCAGTATTTGCTTTATTATGTATTTTATTATTCACAACGGAAATTTGA 177
l23A            GAAATCGGTCGTTCAGTATTTGCTTTATTATGTATTTTATTATTCACAACGGAAATTTGA 146
l23B            GAAATCGGTCGTTCAGTATTTGCTTTATTATGTATTTTATTATTCACAACGGAAATTTGA 178
:v6483A         GAAATCGGTCGTTCAGTATTTGCTTTATTATGTATTTTATTATTCACAACGGAAATTTGA 148
:v6483B         GAAATCGGTCGTTCAGTATTTGCTTTATTATGTATTTTATTATTCACAACGGAAATTTGA 166
                ************************************************************

lig-233WT       GGTTTGACAGACTTGAGAAGTTTAAATTTTGCAGTAAAATTAATGTTCGGTAATACTCAT 237
'23A            GGTTTGACAGACTTGAGAAGTTTAAATTTTGCAGTAAAATTAATGTTCGGTAATACTCAT 206
'23B            GGTTTGACAGACTTGAGAAGTTTAAATTTTGCAGTAAAATTAATGTTCGGTAATACTCAT 238
v6483A          GGTTTGACAGACTTGAGAAGTTTAAATTTTGCAGTAAAATTAATGTTCGGTAATACTCAT 208
v6483B          GGTTTGACAGACTTGAGAAGTTTAAATTTTGCAGTAAAATTAATGTTCGGTAATACTCAT 226
                ************************************************************

ig-233WT        TTCAGTACTATTTGCATTATTGAGCTTAACTTTTTCACAATTAGTACACATATTTAGAAA 297
23A             TTCAGTACTATTTGCATTATTGAGCTTAACTTTTTCACAATTAGTACACATATTTAGAAA 266
23B             TTCAGTACTATTTGCATTATTGAGCTTAACTTTTTCACAATTAGTACACATATTTAGAAA 298
v6483A          TTCAGTACTATTTGCATTATTGAGCTTAACTTTTTCACAATTAGTACACATATTTAGAAA 268
v6483B          TTCAGTACTATTTGCATTATTGAGCTTAACTTTTTCACAATTAGTACACATATTTAGAAA 286
                ************************************************************

ig-233WT        CAAAATCTAATACGCCAACTCAATCCTTATTTTTACAGATCAAGAAAGGTCGTATGGGGT 357
23A             CAAAATCTAATACGCCAACTCAATCCTTATTTTTACAGATCAAGAAAGGTCGTATGGGGT 326
23B             CAAAATCTAATACGCCAACTCAATCCTTATTTTTACAGATCAAGAAAGGTCGTATGGGGT 358
v6483A          CAAAATCTAATACGCCAACTCAATCCTTATTTTTACAGATCAAGAAAGGTCGTATGGGGT 328
v6483B          CAAAATCTAATACGCCAACTCAATCCTTATTTTTACAGATCAAGAAAGGTCGTATGGGGT 346
                ************************************************************

ig-233WT        GATTTGTGATGCAGGGTCAACTGGAACACGGTTATTCGTTTACAACTGGATTAGTACTTC 417
23A             GATTTGTGATGCAGGGTCAACTGGAACACGGTTATTCGTTTACAACTGGATTAGTACTTC 386
23B             GATTTGTGATGCAGGGTCAACTGGAACACGGTTATTCGTTTACAACTGGATTAGTACTTC 418
v6483A          GATTTGTGATGCAGGGTCAACTGGAACACGGTTATTCGTTTACAACTGGATTAGTACTTC 388
v6483B          GATTTGTGATGCAGGGTCAACTGGAACACGGTTATTCGTTTACAACTGGATTAGTACTTC 406
                ************************************************************

ig-233WT        AGGTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAAT 477
23A             AGGTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAAT 446
23B             AGGTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAAT 478
v6483A          AGGTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAAT 448
v6483B          AGGTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAAT 466
                ************************************************************

ig-233WT        TGATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCATGAAGAAGATCAGTCCCG 537
23A             TGATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCATGAAGAAGATCAGTCCCG 506
23B             TGATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCAK-AAGAAGATCATGTAWY 537
v6483A          TGATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCATGAAGAAGATCAGTCCCG 508
v6483B          TGATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCAK-AAGAAGATCATATGAA 525
                ****************************************:.**********
```

```
ig-23WT        --GTGATGCAGGG-TCAACTGGAACACGGTTATTCGTTTACAACTGGATTAGTACTTCAG 57
24A            ------------------------------GCGACACG--TCCTGGATTAGTACTTCAG 27
24B            TTKTGATGCMRGG-TCAACTGGAACACGGTTATTYGTTTACAACTGGATTAGTACTTCAG 59
v_648_4A       ------------------------------GCGMCATCGCTCCTGGATTAGTACTTCAG 29
v_648_4B       TTKTGATGCMARGGTCAACTGGAACACGGTTATTYGTTTACAACTGGATTAGTACTTCAG 60
                                                **********

ig-23WT        GTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAATTG 117
24A            GTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAATTG 87
24B            GTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAATTG 119
v_648_4A       GTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAATTG 89
v_648_4B       GTTTTACTTTAAATTATAAAACTAATACATTTTTCACAATTGTTTCTAGACTCCGAATTG 120
               ************************************************************

ig-23WT        ATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCATGAAGAAGATCAGTCCCGGA 177
24A            ATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCATGAAGAAGATCAGTCCCGGA 147
24B            ATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCATGAAGAAGATCAGTCCCGGA 179
v_648_4A       ATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCATGAAGAAGATCAGTCCCGGA 149
v_648_4B       ATCCAAATTGAACCAGTGATATACGATAACAAGCCGGTCATGAAGAAGATCAGTCCCGGA 180
               ************************************************************

ig-23WT        CTGAGCACATTTGGAACAAAACCTGCCCAAGCTGCAGAATATCTGAGGCCACTTATGGAA 237
24A            CTGAGCACATTTGGAACAAAACCTGCCCAAGCTGCAGAATATCTGAGGCCACTTATGGAA 207
24B            CTGAGCACATTTGGAACAAAACCTGCCCAAGCTGCAGAATATYTGAGGCCACTTATGGAA 239
v_648_4A       CTGAGCACATTTGGAACAAAACCTGCCCAAGCTGCAGAATATCTGAGGCCACTTATGGAA 209
v_648_4B       CTGAGCACATTTGGAACAAAACCTGCCCAAGCTGCAGAATATYTGAGGCCACTTATGGAA 240
               ****************************************** ****************

ig-23WT        CTTGCCGAAAGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCC 297
24A            CTTGCCGAAAGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCC 267
24B            CTTGCCGAAAGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCC 299
v_648_4A       CTTGTCGAAAGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCC 269
v_648_4B       CTTGTCGAAAGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCC 300
               **** ********************************************************

ig-23WT        ACTGCTGGAATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTC 357
24A            ACTGCTGGAATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTC 327
24B            ACTGCTGGAATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTC 359
v_648_4A       ACTGCTGGAATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTC 329
v_648_4B       ACTGCTGGAATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTC 360
               ************************************************************

ig-23WT        CTTTTTCTAAAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAAT 417
24A            CTTTTTCTAAAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAAT 387
24B            CTTTTTYTAAAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAAT 419
v_648_4A       CTTTTTCTAAAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAAT 389
v_648_4B       CTTTTTYTAAAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAAT 420
               ****** ******************************************************

ig-23WT        AAGCTACCAAAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGA 477
24A            AAGCTACCAAAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGA 447
24B            AAGCTACCAAAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGA 479
v_648_4A       AAGCTACCAAAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGA 449
v_648_4B       AAGCTACCAAAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGA 480
               ************************************************************

ig-23WT        AAATGGGAAGGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAAT 537
24A            AAATGGGAAGGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAAT 507
24B            AAATGGGAAGGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAAT 539
v_648_4A       AAATGGGAAGGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAAT 509
v_648_4B       AAATGGGAAGGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAAT 540
               ************************************************************
```

```
lig-23WT        ATTTTCTTTGTATTTATTTTTTTGATATTTAGGAAAGTTCAACAAAACAGCTACATTGGA 597
l24A            ATTTTCTTTGTATTTATTTTTTTGATATTTAGGAAAGTTCAACAAAACAGCTACATTGGA 567
l24B            ATTTTCTTTGTATTTATTTTTTTGATATTTAGGAAAGT-CAACAAAACAGCTACAGAACC 598
:v_648_4A       ATTTTCTTTGTATTTATTTTTTTGATATTTAGGAAAGTTCAACAAAACAGCTACATTGGA 569
:v_648_4B       ATTTTCTTTGTATTTATTTTTTTGATATTTAGGAAAGT-CAACAAAACAKCWCGAKT-CC 598
                ************************************ ********** *   *

lig-23WT        TTTCCCGGGAACCTCGCCAGCACATGCA--- 625
l24A            TTTCCCGGGAACCTCGCCASA---------- 588
l24B            GGTA---------------------------- 602
:v_648_4A       TTTCCCGGGAACCTCGCCAGCACATGCAAGA 600
:v_648_4B       GC------------------------------ 600
```


lig-235WT --TGGAACAAAACCTGCCCAAGCTGCAGAATATCTGAGGCCACTTATGGAACTTGCCGAA 58
l25A ----------TTGCGTA-CTTGCCGAA 16
l25B TTGGAACAAAAAMCTGCCCAAGCTGCAGAATATCTGAGGCCACTTATGGAACTTGCCGAA 60
v_648_5A ----------GCATAASATCTGAGCCTTATGGA-CTTGTCGAA 32
v_648_5B ----------ACCTGCCCAAGCTGCAGAATATCTGAGGCCACTTATGGAACTTGTCGAA 49
lig-235WT AGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCCACTGCTGGA 118
l25A AGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCCACTGCTGGA 76
l25B AGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCCACTGCTGGA 120
v_648_5A -GACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCCACTGCTGGA 91
v_648_5B AGACATATCCCAGAAGAAAAAAGGCCGTATACACCTGTTTTCATTTTTGCCACTGCTGGA 109
ig-235WT ATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTCCTTTTTCTA 178
25A ATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTCCTTTTTCTA 136
25B ATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTCCTTTTTCTA 180
v_648_5A ATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTCCTTTTTCTA 151
v_648_5B ATGAGATTGATTCCTGACGAGTATGTTTTAATCGGGTGATATCATGGTTTCCTTTTTCTA 169
ig-235WT AAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAATAAGCTACCA 238
25A AAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACSTAATAAGCTACCA 196
25B AAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAATAAGCTACCA 240
v_648_5A AAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAATAAGCTACCA 211
v_648_5B AAATTTCAAACTTTCAGACAAAAAGAAGCTGTTCTCAAAAACCTACGTAATAAGCTACCA 229
ig-235WT AAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGAAAATGGGAA 298
25A AAAATTACATCSATGCAAGTACTGAAAGAGCATATCAGGATAATCSAAGGAAAATGGGAA 256
25B AAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGAAAATGGGAA 300
v_648_5A AAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGAAAATGGGAA 271
v_648_5B AAAATTACATCGATGCAAGTACTGAAAGAGCATATCAGGATAATCGAAGGAAAATGGGAA 289
ig-235WT GGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAATATTTTCTTT 358
25A GGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAATATTTTCTTT 316
25B GGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAATATTTTCTTT 360
v_648_5A GGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAATATTTTCTTT 331
v_648_5B GGAATTTATAGTTGGATTGCAGTCAATTATGCTCTTGGTAAGGTTTTCAATATTTTCTTT 349
ig-235WT GTATTTATTTTTTTGATATTTAGGAAAGTTCAACAAAACAGCTACATTGGATTTCCCGGG 418
25A GTATTTATTTTTTTGATATTTAGGAAAGTTCAACAAAACAGCTACATTGGATTTCCCGGG 376
25B GTATTTATTTTTTTGATATTTAGGAAAGTTCAACAAAACAGCTACATTGGATTTCCCGGG 420
v_648_5A GTATTTATTTTTTTGATATTTAGGAAAGTTCAACAAAACAGCTACATTGGATTTCCCGGG 391
v_648_5B GTATTTATTTTTTTGATATTTAGGAAAGTTCAACAAAACAGCTACATTGGATTTCCCGGG 409
ig-235WT AACCTCGCCAGCACATGCAAGGCAAAAAACTGTTGGAATGATTGATATGGGTGGAGCAAG 478
25A AACCTCGCCAGCACATGCAAGGCAAAAAACTGTTGGAATGATTGATATGGGTGGAGCAAG 436
25B AACCTCGCCAGCACATGCAAGGCAAAAAACTGTTGGAATGATTGATATGGGTGGAGCAAG 480
v_648_5A AACCTCGCCAGCACATGCAAGGCAAAAAACTGTTGGAATGATTGATATGGGTGGAGCAAG 451
v_648_5B AACCTCGCCAGCACATGCAAGGCAAAAAACTGTTGGAATGATTGATATGGGTGGAGCAAG 469
ig-235WT TGCTCAAATTGCATTTGAGCTTCCTGACACTGACAGTTTTAGCAGTATTAATGTGGAAAA 538
25A TGCTCAAATTGCATTTGAGCTTCCTGACACTGACAGTTTTAGCAGTATTAATGTGGAAAA 496
25B TGCTCAAATTGCATTTGAGCTTCCTGACACTGACAGTTTTAGCAGTATTAATGTGGAAAA 540
v_648_5A TGCTCAAATTGCATTTGAGCTTCCTGACACTGACAGTTTTAGCAGTATTAATGTGGAAAA 511
v_648_5B TGCTCAAATTGCATTTGAGCTTCCTGACACTGACAGTTTTAGCAGTATTAATGTGGAAAA 529

```
ig-235WT        TGTGAGTTTTGCTTTAAGTAACTACTTTTTAAATATGTGCAATTTTAAGATTAACCTCGG 598
25A             TGTGAGTTTTGCTTTAAGTAACTACTTTTTAAATATGTGCAATTTTAAGATTAAMMSCRR 556
25B             TGTGAGTTTTGCTTTAAGTAACTACT---TAAATATGCATTCT----------------- 580
v_648_5A        TGTGAGTTTTGCTTTAAGTAACTACTTTTTAAATATGTGCAATTTTAARATTAACCTCGG 571
v_648_5B        TGTGAGTTTTGCTTTAAGTAACTACT---TAAATATGCATGG------------------ 568
                ************************** ********

ig-235WT        ATGCAGGG-AAGAC---- 611
25A             SWAMAGGGGAARACCAMA 574
25B             ------------------
v_648_5A        ATGARGGG-AARACCA-- 586
v_648_5B        ------------------
```

```
ig-236WT        -GATATGGGTGG-AGCAAGTGCTCAAATTGCATTTGAGCTTCCTGACACTGACAGTTTTA 58
26A             -------------------------------TTKCR---CTCTGACC-TGACAGTTTTA 24
26B             TKATATGGGGKKGAGCAAGTGCTCAAATTGCATTTGAGCTTCCTGACACTGACAGTTTTA 60
v_648_6A        ---------------CGRRKTRK-RCRTSGCAKMTGR---CTCTGMC--TGAC-GTTTTA 38
v_648_6B        TKATATGGGKKG-AGCAAGTGCTCAAATTGCATTTGAGCTTCCTGACACTGACAGTTTTA 59
                                                          ******.**********

ig-236WT        GCAGTATTAATGTGGAAAATGTGAGTTTTGCTTTAAGTAACTACTTTTTAAATATGTGCA 118
26A             GCAGTATTAATGTGGAAAATGTGAGTTTTGCTTTAAGTAACTACTTTTTAAATATGTGCA 84
26B             GCAGTATTAATGTGGAAAATGTGAGTTTTGCTTTAAGTAACTACTTTTTAAATATGTGCA 120
v_648_6A        GCAGTATTAATGTGGAAAATGTGAGTTTTGCTTTAAGTAACTACTTTTTAAATATGTGCA 98
v_648_6B        GCAGTATTAATGTGGAAAATGTGAGTTTTGCTTTAAGTAACTACTTTTTAAATATGTGCA 119
                ************************************************************

ig-236WT        ATTTTAAGATTAACCTCGGATGCAGGGAAGACGACTCGTTATTCAAGTATAAACTGTTTG 178
26A             ATTTTAAGATTAACCTCGGATGCAGGGAAGACGACTCGTTATTCAAGTATAAACTGTTTG 144
26B             ATTTTAAGATTAACCTCGGATGCAGGGAAGACGACTCGTTATTCAAGTATAAACTGTTTG 180
v_648_6A        ATTTTAAGATTAACCTCGGATGCAGGGAAGACGACTCGTTATTCAAGTATAAACTGTTTG 158
v_648_6B        ATTTTAAGATTAACCTCGGATGCAGGGAAGACGACTCGTTATTCAAGTATAAACTGTTTG 179
                ************************************************************

ig-236WT        TGACAACATTTCTTGGCTACGGAGTCAACGAAGGAATCCGGAAGTACGAGCACATGTTGC 238
26A             TGACAACATTTCTTGGCTACGGAGTCAACGAAGGAATCCGGAAGTACGAGCACATGTTGC 204
26B             TGACAACATTTCTTGGCTACGGAGTCAACGAAGGAATCCGGAAGTACGAGCACATGTTGC 240
v_648_6A        TGACAACATTTCTTGGCTACGGAGTCAACGAAGGAATCCGGAAGTACGAGCACATGTTGC 218
v_648_6B        TGACAACATTTCTTGGCTACGGAGTCAACGAAGGAATCCGGAAGTACGAGCACATGTTGC 239
                ************************************************************

ig-236WT        TGTCAAAATTGAAAGATCAAAATGGAACAGTCATTCAAGATGATTGCATGCCACTGAACT 298
26A             TGTCAAAATTGAAAGATCAAAATGGAACAGTCATTCAAGATGATTGCATGCCACTGAACT 264
26B             TGTCAAAATTGAAAGATCAAAATGGAACAGTCATTCAAGATGATTGCATGCCACTGAACT 300
v_648_6A        TGTCAAAATTGAAAGATCAAAATGGAACAGTCATTCAAGATGATTGCATGCCACTGAACT 278
v_648_6B        TGTCAAAATTGAAAGATCAAAATGGAACAGTCATTCAAGATGATTGCATGCCACTGAACT 299
                ************************************************************

ig-236WT        TACATAAAACCGTCACACTGGAAAACGGAGAAAATTTTGTGCGAAGAGTATGTTGTTTTT 358
26A             TACATAAAACCGTCACACTGGAAAACGGARAAAATTTTGTGCGAAGAGTATGTTGTTTTT 324
26B             TACATAAAACCGTCACACTGGAAAACGGAGAAAATTTTGTGCGAAGAGTATGTTGTTTTT 360
v_648_6A        TACATAAAACCGTCACACTGGAAAACGGARAAWATTTTGTGCGAAGAGTATGTTGTTTTT 338
v_648_6B        TACATAAAACCGTCACACTGGAAAACGGAGAAAATTTTGTGCGAAGAGTATGTTGTTTTT 359
                ********************************.*****************************

ig-236WT        TGGGCTTATTTGAAAATTCGAATAAAATTATTTTTAGGGTACCGGAAACTGGAATACTTG 418
26A             TGGGCTTATTTGAAAATTCRAATAAAATTATTTTTAGGGTACCGGAAACTGGAATACTTG 384
26B             TGGGCTTATTTGAAAATTCGAATAAAATTATTTTTAGGGTACCGGAAACTGGAATACTTG 420
v_648_6A        TGGGCTTATTTGAAAATTCGAATAAAATTATTTTTAGGGTACCGGAAACTGGAATACTTG 398
v_648_6B        TGGGCTTATTTGAAAATTCGAATAAAATTATTTTTAGGGTACCGGAAACTGGAATACTTG 419
                ******************.*********************************************

ig-236WT        TTCAAATGAAGTGAAAAAGCTTCTCAATCCGGAATCAAGTTCGGAAGTCTGTAAAGCTGA 478
26A             TTCAAATGAAGTGAAAAAGCTTCTCAATCCGGAATCAAGTTCGGAAGTCTGTAAAGCTGA 444
26B             TTCAAATGAAGTGAAAAAGCTTCTCAATCCGGAATCAAGTTCGGAAGTCTGTAAAGCTGA 480
v_648_6A        TTCAAATGAAGTGAAAAAGCTTCTCAATCCGGAATCAAGTTCGGAAGTCTGTAAAGCTGA 458
v_648_6B        TTCAAATGAAGTGAAAAAGCTTCTCAATCCGGAATCAAGTTCGGAAGTCTGTAAAGCTGA 479
                ************************************************************

ig-236WT        AGCGGCAAAATGCTATTTTGGAGCTGTTCCTGCTCCGAGTATACCACTGTCGAACATTGA 538
26A             AGCGGCAAAATGCTATTTTGGAGCTGTTCCTGCTCCGAGTATACCACTGTCGAACATTGA 504
26B             AGCGGCAAAATGCTATTTTGGAGCTGTTCCTGCTCCGAGTATACCACTGTCGAACATTGA 540
v_648_6A        AGCGGCAAAATGCTATTTTGGAGCTGTTCCTGCTCCGAGTATACCACTGTCGAACATTGA 518
v_648_6B        AGCGGCAAAATGCTATTTTGGAGCTGTTCCTGCTCCGAGTATACCACTGTCGAACATTGA 539
                ************************************************************
```

```
lig-236WT       GATGTATGGGTTCTCAGAATACTGGTACTCAACCCATGATGTATTGGGTCTTGGAGGACA 598
l26A            GATGTATGGGTTCTCAGAATACTGGTACTCAACCCATGATGTATTGGGTCTTGGAGGACA 564
l26B            GATGTATGGGTTCTCAGAATACTGGTACTCAACGCAKAKTGG------------------ 582
:v_648_6A       GATGTATGGGTTCTCARAATACTGGTACTCAACCCATGATGTATTGGGTCTTGRAGGACA 578
:v_648_6B       GATGTATGGGTTCTCAGAATACTGGTACTCAACTCAKAAGCCC----------------- 582
                ********************,*********************,**

lig-236WT       GTATGATGCGGA- 610
l26A            G-AWG-------- 568
l26B            ------------
:v_648_6A       GTATGATGCGGAA 591
:v_648_6B       ------------
```

```
lig-237WT       --GTCTGTAAAGCTGAAGCGGCAAAATGCTATTTTGGAGCTGTTCCTGCTCCGAGTATAC  58
l27A            ----------------------------GTRMAWTTTGGRCTGTTC-TGCTC-GAGTATAC  31
l27B            TGGTCTGTAAAACTGAAGCGGCAAAATGCTATTTTGGAGCTGTTCCTGCTCCGAGTATAC  60
:v_648_7A       --------------------CMKKRWWKGCGACGAAGTGACTGTTC-TGCTC-GAGTATAC  39
:v_648_7B       --GTCTGTAAAACTGAAGCGGCAAAATGCTATTTTGGAGCTGTTCCTGCTCCGAGTATAC  58
                                     *  .  .     .  ****** ***** ********

lig-237WT       CACTGTCGAACATTGAGATGTATGGGTTCTCAGAATACTGGTACTCAACCCATGATGTAT 118
:27A            CACTGTCSAACATTGAGATGTATGGGTTCTCAGAATACTGGTACTCAACCCATGATGTAT  91
:27B            CACTGTCGAACATTGAGATGTATGGGTTCTCAGAATACTGGTACTCAACCCATGATGTAT 120
v_648_7A        CACTGTCGAACATTGAGATGTATGGGTTCTCAGAATACTGGTACTCAACCCATGATGTAT  99
v_648_7B        CACTGTCGAACATTGAGATGTATGGGTTCTCAGAATACTGGTACTCAACCCATGATGTAT 118
                ******* **************************************************

ig-237WT        TGGGTCTTGGAGGACAGTATGATGCGGAAAATATTGCGAAAAAGACGCAGCAGTATTGTA 178
27A             TGGGTCTTGGAGGACAGTATGATGCGGAAAATATTGCGAAAAAGACGCAGCAGTATTGTA 151
27B             TGGGTCTTGGAGGACAGTATGATGCGGAAAATATTGCGAAAAAGACGCAGCAGTATTGTA 180
v_648_7A        TGGGTCTTGGAGGACAGTATGATGCGGAAAATATTGCGAAAAAGACGCAGCAGTATTGTA 159
v_648_7B        TGGGTCTTGGAGGACAGTATGATGCGGAAAATATTGCGAAAAAGACGCAGCAGTATTGTA 178
                ************************************************************

ig-237WT        GTAAAAGATGGTCCACGATCCAAGCAGAGTCGAAGAAACAGTTGTACCCGAGAGCTGACG 238
27A             GTAAAAGATGGTCCACGATCCAAGCAGAGTCGAAGAAACAGTTGTACCCGAGAGCTGACG 211
27B             GTAAAAGATGGTCCACGATCCAAGCAGAGTCGAAGAAACAGTTGTACCCGAGAGCTGACG 240
v_648_7A        GTAAAAGATGGTCCACGATCCAAGCAGAGTCGAAGAAACAGTTGTACCCGAGAGCTGACG 219
v_648_7B        GTAAAAGATGGTCCACGATCCAAGCAGAGTCGAAGAAACAGTTGTACCCGAGAGCTGACG 238
                ************************************************************

ig-237WT        AGGAAAGATTAAGAACTCAGTGCTTTAAGTCGGCATGGATAACATCAGTGTTGCATGATG 298
27A             AGGAAAGATTAAGAACTCAGTGCTTTAAGTCGGCATGGATAACATCAGTGTTGCATGATG 271
27B             AGGAAAGATTAAGAACTCAGTGCTTTAAGTCGGCATGGATAACATCAGTGTTGCATGATG 300
v_648_7A        AGGAAAGATTAAGAACTCAGTGCTTTAAGTCGGCATGGATAACATCAGTGTTGCATGATG 279
v_648_7B        AGGAAAGATTAAGAACTCAGTGCTTTAAGTCGGCATGGATAACATCAGTGTTGCATGATG 298
                ************************************************************

ig-237WT        GGTTCTCAGTAGATAAGACTCACAACAAATTCCAGGTAATACAATTTGAACAAATCTTGA 358
27A             GGTTCTCAGTAGATAAGACTCACAACAAATTCCAGGTAATACAATTTGAACAAATCTTGA 331
27B             GGTTCTCAGTAGATAAGACTCACAACAAATTCCAGGTAATACAATTTGAACAAATCTTGA 360
v_648_7A        GGTTCTCAGTAGATAAGACTCACAACAAATTCCAGGTAATACAATTTGAACAAATCTTGA 339
v_648_7B        GGTTCTCAGTAGATAAGACTCACAACAAATTCCAGGTAATACAATTTGAACAAATCTTGA 358
                ************************************************************

ig-237WT        ATAAGTTGAGCATTGATCGAATGAAAAAGTACCTTAATTAGCATAATTATTTTGAATTAA 418
27A             ATAAGTTGAGCATTGATCGAATGAAAAAGTACCTTAATTAGCATAATTATTTTGAATTAA 391
27B             ATAAGTTGAGCATTGATCGAATGAAAAAGTACCTTAATTAGCATAATTATTTTGAATTAA 420
v_648_7A        ATAAGTTMAGCATTGATCGAATGAAAAAGTACCTTAATTAGCATAATTATTTTGAATTAA 399
v_648_7B        ATAAGTTGAGCATTGATCGAATGAAAAAGTACCTTAATTAGCATAATTATTTTGAATTAA 418
                ************************************************************

ig-237WT        TTTTTCGCCACCGCTAAACTATTGTGGTCAAGGTTTGAAATTGCAGCACGGTTTCAGCTT 478
27A             TTTTTCGCCACCGCTAAACTATTGTGGTCAAGGTTTGAAATTGCAGCACGGTTTCAGCT- 450
27B             TGT--CGCCACCGCTAAACTATGTCATKT------------------------------- 447
v_648_7A        TTTTTCGCCACCGCTAAACTATTGTGGTCAAGGTTTGAAATTGCAGCACGGTTTCAGCTA 459
v_648_7B        TAT--CGCCACCGCTAAACTATTGAACT-------------------------------- 444
                ***  *****************. . .
```

```
lig-238WT       --GTACCCGAGAGCTGACGAGGAAAGATTAAGAACTCAGTGCTTTAAGTCGGCATGGATA 58
l28A            ------------------------------GCKACTM-GTGCTTTA-GTCGGCATGGATA 28
l28B            TTKTWCCCGAGGGCTGACGRGGAAAGATTAAGAAMTCAGTGCTTTAAGTYGGCATGGATA 60
:v_648_8A       ------------------------------TCGGACTM-GTGCTTTA-GTCGGCATGGATA 29
:v_648_8B       TTKTACCCGAGGGCTGRCGRGGAAAGATTAAGAAMTCAGTGCTTTAAGTYGGCATGGATA 60
                                               *:*: ******** ** **********

lig-238WT       ACATCAGTGTTGCATGATGGGTTCTCAGTAGATAAGACTCACAACAAATTCCAGGTAATA 118
l28A            ACATCAGTGTTGCATGATGGGTTCTCAGTAGATAAGACTCACAACAAATTCCAGGTAATA 88
l28B            AMATCAGTGTTGCMTGATGGGTTYTCMGTAGATAAGACTCMCAACAAATTCCMGGTAATA 120
:v_648_8A       ACATCAGTGTTGCATGATGGGTTCTCAGTAGATAAGACTCACAACAAATTCCAGGTAATA 89
:v_648_8B       ACATCAGTGTTGCMTGATGGGTTYTCMGTAGATAAGACTCMCAACAAATTCCMGGTAATA 120
                ************.********:**:************.**********.*******

lig-238WT       CAATTTGAACAAATCTTGAATAAGTTGAGCATTGATCGAATGAAAAAGTACCTTAATTAG 178
l28A            CAATTTGAACAAATCTTGAATAAGTTGAGCATTGATCGAATGAAAAAGTACCTTAATTAG 148
l28B            CAATTTGAACAAATYTTGAATAARTTGAGCMTTGATYGAAWGAAAAARTACCTTAATTAG 180
v_648_8A        CAATTTGAACAAATCTTGAATAAGTTGAGCATTGATCGAATGAAAAAGTACCTTAATTAG 149
v_648_8B        CAATTTGAACAAATYTTGAATAARTTGAGCATTGATYGAAWGAAAAARTACCTTAATTAG 180
                **************.********'******.*****.***.******.************

ig-238WT        CATAATTATTTTGAATTAATTTTTCGCCACCGCTAAACTATTGTGGTCAAGGTTTGAAAT 238
28A             CATAATTATTTTGAATTAATTTTTCGCCACCGCTAAACTATTGTGGTCAAGGTTTGAAAT 208
28B             CATAATTATTTTGAATTAATTTTTYGCCMCCSCTAAAMTATTGTGGTCAAGGTTTGAAAW 240
v_648_8A        CATAATTATTTTGAATTAATTTTTCGCCACCGCTAAACTATTGTGGTCAAGGTTTGAAAT 209
v_648_8B        CATAATTATTTTGAATTAATTTTTYGCCMCCSSTAAAMTATTGTGGTCAAGGTTTGAAAW 240
                ************************.***.**  ****.**********************

ig-238WT        TGCAGCACGGTTTCAGCTTTTTTCGATTTTTTTTTGAAACGCTATATTTTCTTTTTAAAA 298
28A             TGCAGCACGGTTTCAGCTTTTTTCGATTTTTTTTTGAAACGCTATATTTTCTTTTTAAAA 268
28B             TGCAGCMCGGTTTYAGCTTTTTTYGATTTTTTTTTGAAACGCTATATTTTCTTTTTAAAA 300
v_648_8A        TGCAGCACGGTTTCAGCTTTTTTCGATTTTTTTTTGAAACGCTATATTTTCTTTTTAAAA 269
v_648_8B        TGCARCMCGGTTTYARCTTTTTTYGATTTTTTTTTGAAACGCTATATTTTCTTTTTAAAA 300
                *****.******.*.*******.************************************

ig-238WT        TTTTTAAAATGTCTCCTTTAAATTCTCGATCCGTCTAGACAATTTTTAGGATTAAAAAAA 358
28A             TTTTTAAAATGTCYCCTTTAAATTCYCGATCCGTCTARACAATTTTTAGGATTAAAAAAA 328
28B             TTTTTAAAATGTCTCCTTTAAATTCTCGATCCGTCTAGACAATTTTTAGGATTAAAAAAA 360
v_648_8A        TTTTTAAAATGTCTCCTTTAAATTCTCRATCCGTCTARACAATTTTTAGGATTAAAAAAA 329
v_648_8B        TTTTTAAAATGTCTCCTTTAAATTCTCGATCCGTCTAGACAATTTTTAGGATTAAAAAAA 360
                ************.**********.*.*********.************************

ig-238WT        ATTCGCCACCACATGTATAAATAATTTAGAGTGTTTCCACAATAGCAGGACAAGAAGTTC 418
28A             ATTCSCCMCCACATGTATAAATAATTTARAGKGTTTCCMCAAWAGCAGGACAARAAKTTC 388
28B             ATTCGCCACCACATGTATAAATAATTTAGAGTGTTTCCACAATAGCAGGACAAGAAGTTC 420
v_648_8A        ATTCSCCMCCACAKGTATAAATAATTTARAGKGTTTCCMCAAWAGCAGGACAARAAKTTC 389
v_648_8B        ATTCGCCACCACATGTATAAATAATTTAGAGTGTTTCCACAATAGCAGGACAAGAAGTTC 420
                ****:**:*****;************** **.******.*** ********** **:***

ig-238WT        AATGGGCTCTCGGAGCAATGATCTATCATATGAGATTCTTTCCCCTTCGGGATTCTTCTA 478
28A             AAKGGGCTCYCGRAGCAATGATCTATCAWATGARATTCTTTCCCCTTCGGGATTCTTCTA 448
28B             AATGGGCTCTCGGAGCAATGATCTATCATATGAGATTCTTTCCCCTTCGGGATTCTTCTA 480
v_648_8A        AAKGGGCTCYCGRAGCAATGATCTATCAWATGARATTCTTTCCCCTTCGGGATTCTTCTA 449
v_648_8B        AATGGGCTCTCGGAGCAATGATCTATCATATGAGATTCTTTCCCCTTCGGGATTCTTCTA 480
                ** ****** ** ***************.****.**************************

ig-238WT        GAAATCTCATCGTTAAAGAGTATGTTCCTAAAAATGGGATTGGGCCTTGGTTTC- 532
28A             RAAATCTCATCGTTAAARAGTATGTTCCTAAAAAKGGGATTGGCCTTKGGTTYCA 503
28B             GAAATCTCATCGT-AAAGAGWACTC---------------------------- 504
v_648_8A        RAAATCTCATCGTTAAARAGTATGTTCCTAAAAAKGGGATKGGGCCTTGGTTTMA 504
v_648_8B        GAAATCTCATCGT-AC-GAGWTCC---------------------------- 502
                ************ *   **
```

```
lig-239WT       -CAATGGGCTCTCGGAGCAATGATCTATCATATGAGATTCTTTCCCCTTCGGGATTCTTC 59
l29A            ----------------------------AAWRCRGATCTTT--CCCTTCGGGAT-CTTC 28
l29A2           ----------------------------AAWRCRGATCTTT--CCCTTCGGGAT-CTTC 28
l29B            TTAATGGGCTCTCGGAGCAATGATCTATCATATGAGATTCTTTCCCCTTCGGGATTCTTC 60
:v_648_9A       --------------------AATCACCTCAGCCAGRWYTTT--CCCTTCGGGATTCTTC 37
:v_648_9B       -WAAWGGGCTYTSGGAGCAATGATMTATCATATGAGATTCTTTCCCCTTCGGGATTCTTC 59
                                     **  **********  ****

lig-239WT       TAGAAATCTCATCGTTAAAGAGTATGTTCCTAAAAATGGGATTGGGCCTTGGTTTCATAT 119
l29A            TAGAA-TCTCATCGWTAAAGAGTATGTTCCTAAAAATGGGATTGGGCCTTGGTTTCATAT 87
l29A2           TAGAA-TCTCATCGWTAAAGAGTATGTTCCTAAAAATGGGATTGGGCCTTGGTTTCATAT 87
l29B            TAGAAATCTCATCGTTAAAGAGTATGTTCCTAAAAAWGGGATTGGGCCTTGGTTTCATAT 120
:v_648_9A       TAGAAATCTCATCGTTAAAGAGTATGTTCCTAAAAATGGGATTGGGCCTTGGTTTCATAT 97
:v_648_9B       TAGAAATCTCATCGTTAAAGAGTATGTTCCTAAAAATGGGATTGGGCCTTGGTTTCATAT 119
                **************************************************************

lig-239WT       TTTGGGTTATTCCAGGACTCATTCTTCCTCCGAAAGCTTATGGGCTCCGCTATTCTTCCT 179
l29A            TTTGGGTTATTCCAGGACTCATTCTTCCTCCGAAAGCTTATGGGCTCCGCTATTCTTCCT 147
l29A2           TTTGGGTTATTCCAGGACTCATTCTTCCTCCGAAAGCTTATGGGCTCCGCTATTCTTCCT 147
l29B            TTTGGGTTATTCCAGGACTCATTCTTCCTCCGAAAGCTTATGGGCTCCGCTATTCTTCCT 180
:v_648_9A       TTTGGGTTATTCCAGGACTCATTCTTCCTCCRAAAGMTTATGGGCTCCGCTATTCTTCCT 157
:v_648_9B       TTTGGGTTATTCCAGGACTCATTCTTCCTCCGAAAGMWTATGGGMTCCGCTATTMTTCCT 179
                ******************************************************

lig-239WT       TTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCGCTAAGGAGCAGTCTGTACTATGCTT 239
29A             TTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCGCTAAGGAGCAGTCTGTACTATGCTT 207
29A2            TTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCGCTAAGGAGCAGTCTGTACTATGCTT 207
29B             TTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCGCTAAG-AGCAGTCRWCAC------- 232
v_648_9A        TTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCGCTAAGGAGCAGTCTGTACTATGCTT 217
v_648_9B        TTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCGCTAAG-AGCAKTCWWTCC------- 231
                ******************************************************

ig-239WT        TGATGACAAACGTCGGTCCTC---- 260
29A             TGATGACAAACGTMRGTCCTC---- 228
29A2            TGATGACAAACGTMRGTCCTC---- 228
29B             ------------------------
v_648_9A        TGATGACAAACGSCSRTCCTCCAAA 242
v_648_9B        ------------------------
```

```
ig-2310WT       -TGGGATTGGGCCTTGGTTTCATATTTTGGGTTATTCCAGGACTCATTCTTCCTCCGAAA 59
210A            ---------------------------------TATGTTMTGACTMTTCTTCCTCCGAA- 26
210B            TTKGGATTGGGGCTTGGTTTCATATTTTGGGTTATTCCAGGACTCATTCTTCCTCCGAAA 60
v_648_10A       ----------------------------AAATGGCTTWAATGACTMTTCTTCCTCCGAA- 31
v_648_10B       TTKGGATTGGGCCTTGGTTTCATATTTTGGGTTATTCCAGGACTCATTYTTCCTYCGAAA 60
                                                *       ' ,**,*****,****,

ig-2310WT       GCTTATGGGCTCCGCTATTCTTCCTTTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCG 119
210A            GCTTATGGGCTCCGCTATTCTTCCTTTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCG 86
210B            GCTTATGGGCTCCGCTATTYTTCCTTTCCGCCGTTTTCTGTCTTTTYGTCTTGGTATGCG 120
v_648_10A       GCTTATGGGCTCCGCTATTCTTCCTTTCCGCCGTTTTCTGTCTTTTCGTCTTGGTATGCG 91
v_648_10B       GCTTATGGGCTCCGCTATTYTTCCTTTCCGCCGTTTTYTGTCTTTTYGTCTTGGTATGCG 120
                ********************,**************** ******** ************

ig-2310WT       CTAAGGAGCAGTCTGTACTATGCTTTGATGACAAACGTCGGTCCTCTTTTGGGATGTCAC 179
210A            CTAAGGAGCAGTCTGTACTATGCTTTGATGACAAACGTCGGTCCTCTTTTGGGATGTCAC 146
210B            CTAAGGAGCAGTYTGTACTATGCTTTGATGACAAACGTCGGTCCTYTTTTGGGATGTCAC 180
v_648_10A       CTAAGGAGCAGTCTGTACTATGCTTTGATGACAAACGTCGGTCCTCTTTTGGGATGTCAC 151
v_648_10B       STAAGGAGCAGTYTGTACTATGCTTTGATGACAAACGTCGGTCCTYTTTTGGGATGTCAC 180
                ,**********,****************************************,********** 

ig-2310WT       GCAGTCAATACTCCTACAAAATGCTTAAAGAAAACCGTACTTCTTCTTCATTTCTTGAGA 239
210A            GCAGTCAATACTCCTACAAAATGCTTAAAGAAAACCGTACTTCTTCTTCATTTCTTGAGA 206
210B            GCAGTCAATACTCCTACAAAATGCTTAAAGAAAACCGTACTTYTTYTTCATTTYTTGAGA 240
v_648_10A       GCAGTCAATACTCCTACAAAATGCTTAAAGAAAACCGTACTTCTTCTTCATTTCTTGAGA 211
v_648_10B       GCMGTCAATACTCCTACAAAATGCTTAAAGAAAACCGTACTTYTTYTTCATTTYTTGAGA 240
                **********************************************,******* ****** 

ig-2310WT       ACTTTGCCTAGTCAATCTTTCTAATCGTGTGTGTTCCAATACGTGTTTTATTGTCAAATC 299
210A            ACTTTGCCTAGTCAATCTTTCTAATCGTGTGTGTTCCAATACGTGTTTTATTGTCAAATC 266
210B            ACTTTGCCTAGTCAATCTTTYTAATCGTGTGTGTTCCAATACGTGTTTTATTGTCAAATC 300
v_648_10A       ACTTTGCCTAGTCAATCTTTCTAATCGTGTGTGTTCCAATACGTGTTTTATTGTCAAATC 271
v_648_10B       ACTTTGCCTAGTCAATCTTTYTAATYGTGTGTGTTYCAATACGTGTTTTATTGTCAAATC 300
                ********************,*****,*********,************************ 

ig-2310WT       ACATCGCACTTCAATTGCCTTCCAAAGTTTTATTGTCCTGTCTTTTTTGTTAGATCTTAC 359
210A            ACATCGCACTTCAATTGCCTTCCAAAGTTTTATTGTCCTGTCTTTTTTGTTAGATCTTAC 326
210B            ACATCGCACTTCAATTGCCTTCCAAARTTTTATTGTCCTGTCTTTTTTGTTAGATYTTAC 360
v_648_10A       ACATCGCACTTCAATTGCCTTCCAAAGTTTTATTGTCCTGTCTTTTTTGTTAGATCTTAC 331
v_648_10B       ACATCGCACTTCAATTGCCTTCCAAARTTTTATTGTCCTGTCTTTTTTGTTAGATYTTAC 360
                ************************* ,*****************************,**** 

ig-2310WT       GGGTTTAGTAATTTATATGTGATACTTTTTAATGTATTGAAAATAATTTATCTTGACATG 419
210A            GGGTTTAGTAATTTATATGTGATACTTTTTAATGTATTGAAAATAATTTATCTTGACATG 386
210B            GGGTTTAGTAATTTATATGTGATACTTTTTAATGTATTGAAAATAATTTATCTTGACATG 420
v_648_10A       GGGTTTAGTAATTTATATGTGATACTTTTTAATGTATTGAAAATAATTTATCTTGACATG 391
v_648_10B       GGGTTTAGTAATTTATATGTGATACTTTTTAATGTATTGAAAATAATTTATCTTGACATG 420
                ************************************************************ 

ig-2310WT       TAATCAGTTACGATTTATATTTTTACCGGGAATTTTGATAATTTTTCAATTCTAAATAAA 479
210A            TAATCAGTTACGATTTATATTTTTACCGGGAATTTTGATAATTTTTCAATTCTAAATAAA 446
210B            TAATCAGTTACGATTTATATTTTTACCGGGAATTTTGATAATTTTTCAATTCTAAATAAA 480
v_648_10A       TAATCAGTTACGATTTATATTTTTACCGGGAATTTTGATAATTTTTCAATTCTAAATAAA 451
v_648_10B       TAATCAGTTACGATTTATATTTTTACCGGGAATTTTGATAATTTTTCAATTYTAAATAAA 480
                ************************************************,******** 

ig-2310WT       TTTTTATTTATTTATTTTTTAATGGCAACAATACAAGTTCGAGATACAAGTGTCTAGTAT 539
210A            TTTTTATTTATTTATTTTTTAATGGCAACAATACAAGTTCGAGATACAAGTGTCTAGTAT 506
210B            TTTTTATTTATTTATTTTTTAATGGCAACAATACAAGTTCGAGATACAAGTGTYTAGTAT 540
v_648_10A       TTTTTATTTATTTATTTTTTAATGGCAACAATACAAGTTCGAGATACAAGTGTCTAGTAT 511
v_648_10B       TTTTTWTTTATTTWTTTTTTAATGGCAACAATACAAGTTCGAGATACAAGTGTYTWGTAT 540
                *****,******* ****************************************,* **** 
```

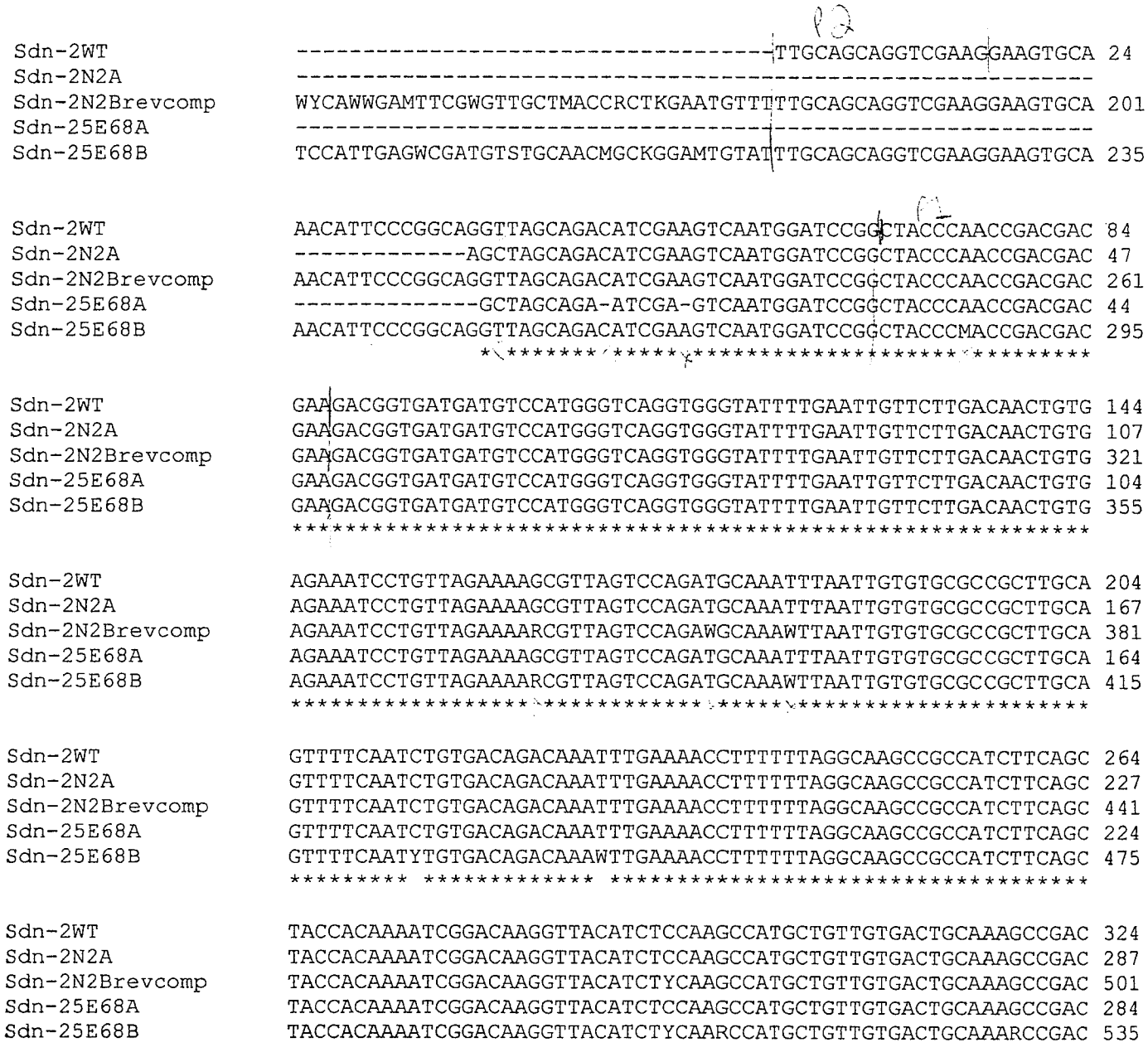
```
lig-2310WT      ACAATTATGACCCGCAATCGGAAAAATTAACTTTCAAATGCCCACCAAAAAACGAATGTA 599
l210A           ACAATTATGACCCGCAATCGGAAAAATTAACTTTCAAATGCCCACCAAAAAACGAATGTA 566
l210B           ACAATTATGACCCGCAATCGGAAAAATTAACTTTCAAATGCCCACCAAAAAACGAATGTA 600
:v_648_10A      ACAATTATGACCCGCAATCGGAAAAATTAACTTTCAAATGCCCACCAAAAAACGAATGTA 571
:v_648_10B      ACAATTATGACCCGCMATCGGAAAAATTAACTTTCAAATGCCCACCAAAAAACGAATGTA 600
                ************************************************************

lig-2310WT      TGTTTTCTGTTTATTGCTACCAAAAAAAATATCCAAAGACAATCTCCAAAAAATATCAAG 659
l210A           TGTTTTCTGTTTATTGCTACCAAAAAAAATATCCAAAGACAATCTCCAAAAAATATCAAG 626
l210B           TGTTTTCTGTTTATTGCTACCAAAAAAAAW-TCCAAAGACA---TCWTCAGGTTAGARAT 656
:v_648_10A      TGTTTTCTGTTTATTGCTACCAAAAAAAATATCCAAAGACAATCTCCAAAAAATATCARG 631
:v_648_10B      TGTTTTYTGTTTATTGCTACCAAAAAAAAT-WCCAA-GACA---TCCAGMACAATGCGRT 655
                ****** ************************ ****** ****   **

lig-2310WT      CAGGAACCTCCAGATG- 675
l210A           CAGGCC---------- 632
l210B           C---------------- 657
:v_648_10A      CAGGAACCTCCARAKGA 648
:v_648_10B      T---------------- 656
```

6.6 MIG-23 protein alignments from translated *mig-23* sequences.

```
CLUSTAL W (1.83) multiple sequence alignment


mig-23wormbase      MRVSLRFTILAVSAMIFFPVIVFIYVVEAHTSPKVIADDQERSYGVICDAGSTGTRLFVY 60
mig-23WT            MRVSLRFTILAVSAMIFFPVIVFIYVVEAHTSPKVIADDQERSYGVICDAGSTGTRLFVY 60
mig-23N2            MRVSLRFTILAVSAMIFFPVIVFIYVVEAHTSPKVIADDQERSYGVICDAGSTGTRLFVY 60
mig-23_ev648        MRVSLRFTILAVSAMIFFPVIVFIYVVEAHTSPKVIADDQERSYGVICDAGSTGTRLFVY 60
                    ************************************************************

mig-23wormbase      NWISTSDSELIQIEPVIYDNKPVMKKISPGLSTFGTKPAQAAEYLRPLMELAERHIPEEK 120
mig-23WT            NWISTSDSELIQIEPVIYDNKPVMKKISPGLSTFGTKPAQAAEYLRPLMELAERHIPEEK 120
mig-23N2            NWISTSDSELIQIEPVIYDNKPVMKKISPGLSTFGTKPAQAAEYLRPLMELAERHIPEEK 120
mig-23_ev648        NWISTSDSELIQIEPVIYDNKPVMKKISPGLSTFGTKPAQAAEYLRPLMELVERHIPEEK 120
                    ************************************************.********** 

mig-23wormbase      RPYTPVFIFATAGMRLIPDEQKEAVLKNLRNKLPKITSMQVLKEHIRIIEGKWEGIYSWI 180
mig-23WT            RPYTPVFIFATAGMRLIPDEQKEAVLKNLRNKLPKITSMQVLKEHIRIIEGKWEGIYSWI 180
mig-23N2            RPYTPVFIFATAGMRLIPDEQKEAVLKNLRNKLPKITSMQVLKEHIRIIEGKWEGIYSWI 180
mig-23_ev648        RPYTPVFIFATAGMRLIPDEQKEAVLKNLRNKLPKITSMQVLKEHIRIIEGKWEGIYSWI 180
                    ************************************************************

mig-23wormbase      AVNYALGKFNKTATLDFPGTSPAHARQKTVGMIDMGGASAQIAFELPDTDSFSSINVENI 240
mig-23WT            AVNYALGKFNKTATLDFPGTSPAHARQKTVGMIDMGGASAQIAFELPDTDSFSSINVENI 240
mig-23N2            AVNYALGKFNKTATLDFPGTSPAHARQKTVGMIDMGGASAQIAFELPDTDSFSSINVENI 240
mig-23_ev648        AVNYALGKFNKTATLDFPGTSPAHARQKTVGMIDMGGASAQIAFELPDTDSFSSINVENI 240
                    ************************************************************

mig-23wormbase      NLGCREDDSLFKYKLFVTTFLGYGVNEGIRKYEHMLLSKLKDQNGTVIQDDCMPLNLHKT 300
mig-23WT            NLGCREDDSLFKYKLFVTTFLGYGVNEGIRKYEHMLLSKLKDQNGTVIQDDCMPLNLHKT 300
mig-23N2            NLGCREDDSLFKYKLFVTTFLGYGVNEGIRKYEHMLLSKLKDQNGTVIQDDCMPLNLHKT 300
mig-23_ev648        NLGCREDDSLFKYKLFVTTFLGYGVNEGIRKYEHMLLSKLKDQNGTVIQDDCMPLNLHKT 300
                    ************************************************************

mig-23wormbase      VTLENGENFVRRGTGNWNTCSNEVKKLLNPESSSEVCKAEAAKCYFGAVPAPSIPLSNIE 360
mig-23WT            VTLENGENFVRRGTGNWNTCSNEVKKLLNPESSSEVCKAEAAKCYFGAVPAPSIPLSNIE 360
mig-23N2            VTLENGENFVRRGTGNWNTCSNEVKKLLNPESSSEVCKAEAAKCYFGAVPAPSIPLSNIE 360
mig-23_ev648        VTLENGENFVRRGTGNWNTCSNEVKKLLNPESSSEVCKAEAAKCYFGAVPAPSIPLSNIE 360
                    ************************************************************

mig-23wormbase      MYGFSEYWYSTHDVLGLGGQYDAENIAKKTQQYCSKRWSTIQAESKKQLYPRADEERLRT 420
mig-23WT            MYGFSEYWYSTHDVLGLGGQYDAENIAKKTQQYCSKRWSTIQAESKKQLYPRADEERLRT 420
mig-23N2            MYGFSEYWYSTHDVLGLGGQYDAENIAKKTQQYCSKRWSTIQAESKKQLYPRADEERLRT 420
mig-23_ev648        MYGFSEYWYSTHDVLGLGGQYDAENIAKKTQQYCSKRWSTIQAESKKQLYPRADEERLRT 420
                    ************************************************************

mig-23wormbase      QCFKSAWITSVLHDGFSVDKTHNKFQSVSTIAGQEVQWALGAMIYHMRFFPLRDSSRNLI 480
mig-23WT            QCFKSAWITSVLHDGFSVDKTHNKFQSVSTIAGQEVQWALGAMIYHMRFFPLRDSSRNLI 480
mig-23N2            QCFKSAWITSVLHDGFSVDKTHNKFQSVSTIAGQEVQWALGAMIYHMRFFPLRDSSRNLI 480
mig-23_ev648        QCFKSAWITSVLHDGFSVDKTHNKFQSVSTIAGQEVQWALGAMIYHMRFFPLRDSSRNLI 480
                    ************************************************************

mig-23wormbase      VKETHSSSESLWAPLFFLSAVFCLFVLVCAKEQSVLCFDDKRRSSFGMSRSQYSYKMLKE 540
mig-23WT            VKETHSSSESLWAPLFFLSAVFCLFVLVCAKEQSVLCFDDKRRSSFGMSRSQYSYKMLKE 540
mig-23N2            VKETHSSSESLWAPLFFLSAVFCLFVLVCAKEQSVLCFDDKRRSSFGMSRSQYSYKMLKE 540
mig-23_ev648        VKETHSSSESLWAPLFFLSAVFCLFVLVCAKEQSVLCFDDKRRSSFGMSRSQYSYKMLKE 540
                    ************************************************************

mig-23wormbase      NRTSSSFLENFA 552
mig-23WT            NRTSSSFLENFA 552
mig-23N2            NRTSSSFLENFA 552
```

6.7 Alignments of *sdn-1* sequencing results.

```
1Sdn-1WT          -TCCTCCTCCACCACAACACCAATTGCTTTTTCAGGGGTAAAGAGAGTTTCCAAAGAGAT 59
2Sdn-1N2B         -TCCTCCTCCCMCACAACACCAATTGCTTTTTCAGGGGTAAAGAGAGTTTCCAAAGAGAT 59
3Sdn-15E68Aasis   ------------------------------------TAGAGAGTTTCCAA-GAGAT 19
4Sdn-15E68B       TTCCTCCTCCACCACAACACCAATTGCTTTTTCAGGGGTAAAGAGAGTTTCCAAAGAGAT 60
                                                      ************ *****

1Sdn-1WT          GCCGGTCAGGTGATTACACCAACAAGACAATATGATTCTGAAACTCAATTTCTGCCTGTC 119
2Sdn-1N2B         GCCGGTCAGGTGATTACACCAACAAGACAATATGATTCTGAAACTCAATTTCTGCCTGTC 119
3Sdn-15E68Aasis   GCCGGTCAGGTGATTACACCAACAAGACAATATGATTCTGAAACTCAATTTCTGCCTGTC 79
4Sdn-15E68B       GCCGGTCAGGTGATTACACCAACAAGACAATATGATTCTGAAACTCAATTTCTGCCTGTC 120
                  ************************************************************

1Sdn-1WT          GACTTACTCGGTCCTCATTCTCCTATCCTTATCTACACAAGCTTTCGCCGCAAATCAAGC 179
2Sdn-1N2B         GACTTACTCGGTCCTCATTCTCCTATCCTTATCTACACAAGCTTTCGCCGCAAATCAAGC 179
3Sdn-15E68Aasis   GACTTACTCGGTCCTCATTCTCCTATCCTTATCTACACAAGCTTTCGCCGCAAATCAAGC 139
4Sdn-15E68B       GACTTACTCGGTCCTCATTCTCCTATCCTTATCTACACAAGCTTTCGCCGCAAATCAAGC 180
                  ************************************************************

1Sdn-1WT          AAAAACGAAAGTGGTCCCTTCGTCAACTATTTCTACGAAAAGCTTGAAAAATGGAATATC 239
2Sdn-1N2B         AAAAACGAAAGTGGTCCCTTCGTCAACTATTTCTACGAAAAGCTTGAAAAATGGAATATC 239
3Sdn-15E68Aasis   AAAAACGAAAGTGGTCCCTTCGTCAACTATTTCTACGAAAAGCTTGAAAAATGGAATATC 199
4Sdn-15E68B       AAAAACGAAAGTGGTCCCTTCGTCAACTATTTCTACGAAAAGCTTGAAAAATGGAATATC 240
                  ************************************************************

1Sdn-1WT          TGAAGTAAGTTATTATATAAATTTATCATAATGTTACAATACGATACTACGAAATTCATT 299
2Sdn-1N2B         TGAAGTAAGTTATTATATAAATTTATCATAATGTTACAATACGATACTACGAAATTCATT 299
3Sdn-15E68Aasis   TGAAGTAAGTTATTATATAAATTTATCATAATGTTACAATACGATACTACGAAATTCATT 259
4Sdn-15E68B       TGAAGTAAGTTATTATATAAATTTATCATAATGTTACAATACGATACTACGAAATTCATT 300
                  ************************************************************

1Sdn-1WT          CACCAGCGCACTTTTTGTTTGCGCAAAGCAGGCTCGAATCTGAAGCGCGGCAAATGCGCA 359
2Sdn-1N2B         CACCAGCGCACTTTTTGTTTGCGCAAAGCAGGCTCGAATCTGAAGCGCGGCAAATGCGCA 359
3Sdn-15E68Aasis   CACCAGCGCACTTTTTGTTTGCGCAAAGCAGGCTCGAATCTGAAGCGCGGCAAATGCGCA 319
4Sdn-15E68B       CACCAGCGCACTTTTTGTTTGCGCAAAGCAGGCTCGAATCTGAAGCGCGGCAAATGCGCA 360
                  ************************************************************

1Sdn-1WT          CCAAAATATTTGCATTATTCAAACAGATATAATGGGCGATGGAGCGCGCGTAACATAAAT 419
2Sdn-1N2B         CCAAAATATTTGCATTATTCAAACAGATATAATGGGCGATGGAGCGCGCGTAACATAAAT 419
3Sdn-15E68Aasis   CCAAAATATTTGCATTATTCAAACAGATATAATGGGCGATGGAGCGCGCGTAACATAAAT 379
4Sdn-15E68B       CCAAAATATTTGCATTATTCAAACAGATATAATGGGCGATGGAGCGCGCGTAACATAAAT 420
                  ************************************************************

1Sdn-1WT          AATAATATTAATAATAATAACGAATAATATTTATTTAAATAATATAATAATAATATTGGC 479
2Sdn-1N2B         AATAATATTAATAATAATAACGAATAATATTTATTTAAATAATATAATAATAATATTGGC 479
3Sdn-15E68Aasis   AATAATATTAATAATAATAACGAATAATATTTATTTAAATAATATAATAATAATATTGGC 439
4Sdn-15E68B       AATAATATTAATAATAATAACGAATAATATTTATTTAAATAATATAATAATAATATTGGC 480
                  ************************************************************

1Sdn-1WT          AATATGAATAATTTGAAACTTTTGCAGCAGGTCGAAGGAAGTGCAAACATTCCCGGCAGG 539
2Sdn-1N2B         AATATGAATAATTTGAAACTTTTGCAGCAGGTCGAAGGAAGTGCAAACAT-CCCGGCAGG 538
3Sdn-15E68Aasis   AATATGAATAATTTGAAACTTTTGCAGCAGGTCGAAGGAAGTGCAAACATTCCCGGCAGG 499
4Sdn-15E68B       AATATGAATAATTTGAAACTTTTGCAGCAGGTCGAAGGAAGTGCAAACAT-CCCGGCAGG 539
                  ************************************************************

1Sdn-1WT          TTAGCAGACATCGAAGTCAATGGATCCGGCTACCCAACCGACGACGAA-- 587
2Sdn-1N2B         T-AGCAGACAT---------------------------------------- 548
3Sdn-15E68Aasis   TTAGCAGACATCGAAGTCAATGGATCCGGCTACCCAACCGAACGACRAAA 549
4Sdn-15E68B       T-AGCAGACAT---------------------------------------- 549
                  * *********
```


Sdn-2WT ----------TTGCAGCAGGTCGAAGGAAGTGCA 24
Sdn-2N2A ----------
Sdn-2N2Brevcomp WYCAWWGAMTTCGWGTTGCTMACCRCTKGAATGTTTTTGCAGCAGGTCGAAGGAAGTGCA 201
Sdn-25E68A ----------
Sdn-25E68B TCCATTGAGWCGATGTSTGCAACMGCKGGAMTGTATTTGCAGCAGGTCGAAGGAAGTGCA 235
Sdn-2WT AACATTCCCGGCAGGTTAGCAGACATCGAAGTCAATGGATCCGGCTACCCAACCGACGAC 84
Sdn-2N2A ----------AGCTAGCAGACATCGAAGTCAATGGATCCGGCTACCCAACCGACGAC 47
Sdn-2N2Brevcomp AACATTCCCGGCAGGTTAGCAGACATCGAAGTCAATGGATCCGGCTACCCAACCGACGAC 261
Sdn-25E68A ----------GCTAGCAGA-ATCGA-GTCAATGGATCCGGCTACCCAACCGACGAC 44
Sdn-25E68B AACATTCCCGGCAGGTTAGCAGACATCGAAGTCAATGGATCCGGCTACCCMACCGACGAC 295
Sdn-2WT GAAGACGGTGATGATGTCCATGGGTCAGGTGGGTATTTTGAATTGTTCTTGACAACTGTG 144
Sdn-2N2A GAAGACGGTGATGATGTCCATGGGTCAGGTGGGTATTTTGAATTGTTCTTGACAACTGTG 107
Sdn-2N2Brevcomp GAAGACGGTGATGATGTCCATGGGTCAGGTGGGTATTTTGAATTGTTCTTGACAACTGTG 321
Sdn-25E68A GAAGACGGTGATGATGTCCATGGGTCAGGTGGGTATTTTGAATTGTTCTTGACAACTGTG 104
Sdn-25E68B GAAGACGGTGATGATGTCCATGGGTCAGGTGGGTATTTTGAATTGTTCTTGACAACTGTG 355
Sdn-2WT AGAAATCCTGTTAGAAAAGCGTTAGTCCAGATGCAAATTTAATTGTGTGCGCCGCTTGCA 204
Sdn-2N2A AGAAATCCTGTTAGAAAAGCGTTAGTCCAGATGCAAATTTAATTGTGTGCGCCGCTTGCA 167
Sdn-2N2Brevcomp AGAAATCCTGTTAGAAAARCGTTAGTCCAGAWGCAAAWTTAATTGTGTGCGCCGCTTGCA 381
Sdn-25E68A AGAAATCCTGTTAGAAAAGCGTTAGTCCAGATGCAAATTTAATTGTGTGCGCCGCTTGCA 164
Sdn-25E68B AGAAATCCTGTTAGAAAARCGTTAGTCCAGATGCAAAWTTAATTGTGTGCGCCGCTTGCA 415
Sdn-2WT GTTTTCAATCTGTGACAGACAAATTTGAAAACCTTTTTTAGGCAAGCCGCCATCTTCAGC 264
Sdn-2N2A GTTTTCAATCTGTGACAGACAAATTTGAAAACCTTTTTTAGGCAAGCCGCCATCTTCAGC 227
Sdn-2N2Brevcomp GTTTTCAATCTGTGACAGACAAATTTGAAAACCTTTTTTAGGCAAGCCGCCATCTTCAGC 441
Sdn-25E68A GTTTTCAATCTGTGACAGACAAATTTGAAAACCTTTTTTAGGCAAGCCGCCATCTTCAGC 224
Sdn-25E68B GTTTTCAATYTGTGACAGACAAAWTTGAAAACCTTTTTTAGGCAAGCCGCCATCTTCAGC 475
Sdn-2WT TACCACAAAATCGGACAAGGTTACATCTCCAAGCCATGCTGTTGTGACTGCAAAGCCGAC 324
Sdn-2N2A TACCACAAAATCGGACAAGGTTACATCTCCAAGCCATGCTGTTGTGACTGCAAAGCCGAC 287
Sdn-2N2Brevcomp TACCACAAAATCGGACAAGGTTACATCTYCAAGCCATGCTGTTGTGACTGCAAAGCCGAC 501
Sdn-25E68A TACCACAAAATCGGACAAGGTTACATCTCCAAGCCATGCTGTTGTGACTGCAAAGCCGAC 284
Sdn-25E68B TACCACAAAATCGGACAAGGTTACATCTYCAARCCATGCTGTTGTGACTGCAAARCCGAC 535

```
Sdn-2WT              AACGGTACCTACTACTACAGCGGTTTGTGTCGGCTTTTTCTTTTGATGTGTTTTACCGCA 384
Sdn-2N2A             AACGGTACCTACTACTACAGCGGTTTGTGTCGGCTTTTTCTTTTGATGTGTTTTACCGCA 347
Sdn-2N2Brevcomp      AACGGTACCTACTACTACAGCGGTTTGTGTCGGCTTTTTCTTTTGATGTGTTTTACCGCA 561
Sdn-25E68A           AACGGTACCTACTACTACAGCGGTTTGTGTCGGCTTTTTCTTTTGATGTGTTTTACCGCA 344
Sdn-25E68B           AACGGTACCTACTWCTACAGCGGTTTGTGTCGGCTTTTTCTTTTGATGTGTTTTACCGCA 595
                     ************ ********************************************

Sdn-2WT              ATCTTTCATTCTTCAGAGCTTCAAGCCTCCTGTTCAGCCCAAGCCTAAGCCAGCGGCAAA 444
Sdn-2N2A             ATCTTTCATTCTTCAGAGCTTCAAGCCTCCTGTTCAGCCCAAGCCTAAGCCAGCGGCAAA 407
Sdn-2N2Brevcomp      ATCTTTCAT--CTCAGAGCT-CAAGCTCC------------------------------ 587
Sdn-25E68A           ATCTTTCATTCTTCARAGCTTCAAGCCTCCTGTTCAGCCCAAGCCTAAGCCAGCGGCAAA 404
Sdn-25E68B           ATCTT-CAT--CTCAGAGCT-CAAGCTCAT------------------------------ 621
                     *****.****    *** **** *****

Sdn-2WT              CGA-CAAGGAG- 454
Sdn-2N2A             CGA-CAAGGAGA 418
Sdn-2N2Brevcomp      ------------
Sdn-25E68A           CGAACAAGGAGA 416
Sdn-25E68B           ------------
```

```
sdn3          CTACAGCGG-TTTGTGTCGGCTTTTTCTTTTGATGTGTTTTACCGCAATCTTTCATTCTT 59
sdn3N2A       ----------------------------------------TACCGCA--TCTTTMTTCTT 18
sdn3N2B       CTACARCSGGTTTGTGTCGGCTTTTTCTTTTGATGTGTTTTACCGCAATCTTTCATTYTT 60
sdn35E68A     -----------------------------CATRCTGGGTTACGCA--TCTTTMTTCTT 27
Sdn35E68B     -TACARCGG-TTTGTGTCGGCTTTTTCTTTTGATGTGTTTTACCGCAATCTTTCATTYTT 58
                                             *  ****    **  ** **

sdn3          CAGAGCTTCAAGCCTCCTGTTCAGCCCAAGCCTAAGCCAGCGGCAAACGACAAGGAGATC 119
sdn3N2A       CAGAGCTTCARGCCTCCKGTTCAGCCCARGCCWARGCCASCGGYAAACGACAAGGAGATC 78
sdn3N2B       CAGAGCTTCAAGCCTCCTGTTCAGCCCAAGCCTAAGCCAGCGGCAAACGACAAGGAGATC 120
sdn35E68A     C-GAGCTTCARGCCTCCKGTTCAGCCCAAGCCTAAGCCMGCGGCAAACGACAAGGAGATC 86
Sdn35E68B     YAGAGCTTCAARCCTCCTGTTCAGCCCAAGCCTAAGCCAGCGGCAAACGACAAGGAGATC 118
              *********  ***** ********** *** * ***  *** ****************

sdn3          AAGGTCGAGGAGGACGAGGACGATGATGAAGATGAGGATGAAGATGATGAGGATGATGAA 179
sdn3N2A       ARGGTCGAGRAGGACGAGGACRATGATGAARATGAGGATGAAGATGATGAGGATGATGAA 138
sdn3N2B       AAGGTCGAGGAGGACGAGGACGATGATGAAGATGAGGATGAAGATGATGAGGATGATGAA 180
sdn35E68A     AAGGTCGAGGAGGACGAGGACGATGATGAAGATGAGGATGAAGATGATGAGGATGATGAA 146
Sdn35E68B     AAGGTCGAGGAGGACGAGGACGATGATGAAGATGAGGATGAAGATGATGAGGATGATGAA 178
              ********** ********** ******** ****************************

sdn3          GAAGATTTTGCTGATGAGAATATTCATAATGATGAAGATTTCTTCACAACTACTACTACA 239
sdn3N2A       RAAGATTTTGCTGATGARAATATTCATAAKGATGAARATTTCTTCACAACTACWACTACA 198
sdn3N2B       GAAGATTTTGCTGATGAGAATATTCATAATGATGAAGATTTYTTCACAACTACTACTACA 240
sdn35E68A     GAAGATTTTGCTGATGAGAATATTCATAATGATGAAGATTTCTTCACAACTACTACTACA 206
Sdn35E68B     GAAGATTTTGCTGATGAGAATATTCATAATGATGAAGATTTYTTCACAACTACTACTACA 238
              ***************** ********** ****** **** ********** ******

sdn3          ACAACGTATCGACCTATTGTTGTAGCTACCACCTCGTATGTTTTTCATTTGAACAAAAAA 299
sdn3N2A       ACAACGTATCGACCTATTGTTGTAGCTACCACCTCGTATGTTTTTCATTTGAACAAAAAA 258
sdn3N2B       ACAACGTATCGACCTATTGTTGTAGCTACCACCTCGTATGTTTTTCATTTGAACAAAAAA 300
sdn35E68A     ACAACGTATCGACCTATTGTTGTAGCTACCACCTCGTATGTTTTTCATTTGAACAAAAAA 266
Sdn35E68B     ACAACGTATCGACCTATTGTTGTAGCTACCMCCTYGTATGTTTTTCATTTGAACAAAAAA 298
              ******************************* **** *************************

sdn3          CAAAACTATTCTAATGCCCTCCAGGGCCCAGGATTTATGCATCTAGAAATGTATTAATTA 359
sdn3N2A       CAAAAYTATTCTAATGCCCTCCAGGGCCCAGGATTTATGCATCTAGAAAKGTATTAATTA 318
sdn3N2B       CAAAACTATTYTAATGCCCTCCAGGGCCCAGGATTTATGCATCTAGAAATGTATTAATTA 360
sdn35E68A     CAAAACTATTCTAAKGCCCTCCAGGGCCCAGGATTTATGCATCTARAAATGTATTAATTA 326
Sdn35E68B     CAAAACTATTYTAATGCCCTCCAGGGCCCAGGATTTATGCATCTAGAAATGTATTAATTA 358
              ***** **** *** ******************************** *** **********

sdn3          TCTTGTCAATCAACAAATCCCCGAAAACATCTAGCAGCCAATTTAATTTTCAATTTTCCA 419
sdn3N2A       TCTTGTCAATCAACAAATCCCCRAAAACWTCTAGCAGCCAWTTTAWTTTTCAWTTTTCCA 378
sdn3N2B       TCTTGTCAATCAACAAATCCCCGAAAACATYTAGCAGCCAATTTAATTTTCAATTTTCCA 420
sdn35E68A     TCTTGTCAATCAACAAATCCCCRAAAACATCTAGCAGCCAATTTAATTTTCAATTTTCCA 386
Sdn35E68B     TCTTGTCAATCAACAAATCCCCGAAAACATCTAGCAGCCAATTTAATTTTCAATTTTCCA 418
              ********************** ***** * ********* **** ****** *******

sdn3          GAACGCCAAGGTCAGCAGCCACAAATCCACCTCGACAGCAGCCACCAATGGTCACATCAA 479
sdn3N2A       RAACYCCAAGGTCAGCASCCACAAATCCACCTCGACRGGAGCCACCAAKGGTCACATCAA 438
sdn3N2B       GAACGCCAAGGTCAGCAGCCACAAATCCACCTYGACAGGAGCCACCAATGGTCACATCAA 480
sdn35E68A     RAACGCCAAGGTCAGCAGCCACAAATCCACCTCGACAGTAGCCACCAATGGTCACATCAA 446
Sdn35E68B     GAACGCCAAGGTCAGCRGCCACAAATCCACCTCGACAGTAGCCACCAATGGTCACATCAA 478
               *** ***********  ************** *** * ********* ***********

sdn3          CCATCTCATCTGGACCATTCTCGCCATTCCATGAGACACTGGCAAATGGCTTC 532
sdn3N2A       CCAYCTCATCKGGACCATTCTCGCCWTTCCATGARACACKGGCAAAKGG---- 487
sdn3N2B       CCATCTCATCTGATCATYCSACGCCA---------------------------- 506
sdn35E68A     CCATCTCATCKGGACCATTCTCGCCATTCCATGARACACTGGCAAAKGGCTYM 499
Sdn35E68B     CCATCTCATCTGAGCAT----------------------------------- 495
```

```
sdn4WT          --CAGAACGCCAAGGTCAGCAGCCACAAATCCACCTCGACAGCAGCCACCAATGGTCACA 58
Sdn4N2A         ------------------------------------CTCGAC-GGAGCCACCAATGGTCACA 25
Sdn4N2B         -TAGAACGCCCCAGGTCAGCAGCCACAAATCCACCTCGACAGGAGCCACCAATGGTCACA 59
Sdn45E68A       ---------------------------------CTCGTM-GTAGCCACCAATGGTCACA 25
Sdn45E68B       TTAGAACGCCA-AGGTCAGCAGCCACAAATCCACCTCGACAGTAGCCACCAATGGTCACA 59
                                                 ****   *.******************

sdn4WT          TCAACCATCTCATCTGGACCATTCTCGCCATTCCATGAGACACTGGCAAATGGCTTCTAT 118
Sdn4N2A         TCAACCATCTCATCTGGACCATTCTCGCCATTCCATGARACACTGGCAAATGGCTTCTAT 85
Sdn4N2B         TCAACCATCTCATYTGGACCATTCTCGCCATTCCATGAGACACTGGCAAATGGCTTCTAT 119
Sdn45E68A       TCAACCATCTCATCTGGACCATTCTCGCCATTCCATGARACACTGGCAAATGGCTTCTAT 85
Sdn45E68B       TCAACCATCTCATYTGGACCATTCTCGCCATTCCATGAGACACTGGCAAATGGCTTCTAT 119
                ************* ************************ **********************

sdn4WT          GCAGGTATGTGTGAAGCAAAACTCATTTCAATTAATTTTGCATTGGGACCTTTGATCGTC 178
Sdn4N2A         GCAGGTATGTGTGAAGCAAAACTCATTTCAATTAATTTTGCATTGGGACCTTTGATCGTC 145
Sdn4N2B         GCAGGTATGTGTGAAGCAAAACTCATTTCAATTAATTTTGCATTGGGACCTTTGATCGTC 179
Sdn45E68A       GCAGGTATGTGTGAAGCAAAACTCATTTCAATTAATTTTGCATTGGGACCTTTGATCGTC 145
Sdn45E68B       GCAGGTATGTGTGAAGCAAAACTCATTTCAATTAATTTTGCATTGGGACCTTTGATCGTC 179
                ************************************************************

sdn4WT          CGGAGTGTAACTATCATTAGTGGGAGCAGATGCAAATGTACATAAATCGCCCGGTCAGTG 238
Sdn4N2A         CGGAGTGTAACTATCATTAGTGGGAGCAGATGCAAATGTACATAAATCGCCCGGTCAGTG 205
Sdn4N2B         CGGAGTGTAACTATCATTAGTGGGAGCAGATGCAAATGTACATAAATCGCCCGGTCAGTG 239
Sdn45E68A       CGGAGTGTAACTATCATTAGTGGGAGCAGATGCAAATGTACATAAATCGCCCGGTCAGTG 205
Sdn45E68B       CGGAGTGTAACTATCATTAGTGGGAGCAGATGCAAATGTACATAAATCGCCCGGTCAGTG 239
                ************************************************************

sdn4WT          TGATGGCGTCGGCGAATGATTTGTTGCGAGATGATAAATTGATGTCTTCTCGTCAATCGA 298
Sdn4N2A         TGATGGCGTCGGCGAATGATTTGTTGCGAGATGATAAATTGATGTCTTCTCGTCAATCGA 265
Sdn4N2B         TGATGGCGTCGGCGAATGATTTGTTGCGAGATGATAAATTGATGTYTTYTCGTCAATCGA 299
Sdn45E68A       TGATGGCGTCGGCGAATGATTTGTTGCRAGATGATAAATTGATGTCTTCTCGTCAATCGA 265
Sdn45E68B       TGATGGCGTCGGCGAATGATTTGTTGCGAGATGATAAATTGATGTCTTCTCGTCAATCGA 299
                *************************** ***************** ** ***********

sdn4WT          AGGCAACACGGTTGCATTCAATATATTCGGCCACTGTATTCCGGGGGGAATATGCTGCCTC 358
Sdn4N2A         AGGCAACACGGTTGCATTCAATATATTCGGCCACTGTATTCCGGGGGGAATATGCTGCCTC 325
Sdn4N2B         AGGCAACACGGTTGCATTCAATATATTCGGCCACTGTATTCCGGGGGGAATATGCTGCCTC 359
Sdn45E68A       AGGCAACACGGTTGCATTCAATATATTCGGCCACTGTATTCCGGGGGGAATATGCTGCCTC 325
Sdn45E68B       AGGCAACACGGTTGCATTCAATATATTCGGCCACTGTATTCCGGGGGGAATATGCTGCCTC 359
                ************************************************************

sdn4WT          AAGCCTATCCGTTCCGTCTGTTCCGGATGGGCAAATATACACAAAAATATGAGACTTGCA 418
Sdn4N2A         AAGCCTATCCGTTCCGTCTGTTCCGGATGGGCAAATATACACAAAAATATGAGACTTGCA 385
Sdn4N2B         AAGCCTATCCGTTCCGTCTGTTCCGGATGGGCAAATATACACAAAAATATGAGACTTGCA 419
Sdn45E68A       AAGCCTATCCGTTCCGTCTGTTCCGGATGGGCAAATATACACAAAAATATGAGACTTGCA 385
Sdn45E68B       AAGCCTATCCGTTCCGTCTGTTCCGGATGGGCAAATATACACAAAAATATGAGACTTGCA 419
                ************************************************************

sdn4WT          TAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCAGT 478
Sdn4N2A         TAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCAGT 445
Sdn4N2B         TAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCAGT 479
Sdn45E68A       TAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCAGT 445
Sdn45E68B       TAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCSGTCCATTTTTCGCGACCAGT 479
                ************************************** ********************
```

```
sdn4WT          CTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTTTGTTGTCGAGCTGTTTCTTAAT 538
Sdn4N2A         CTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTTTGTTGTCGAGCTGTTTCTTAAT 505
Sdn4N2B         CTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTT-GTG--CGAGCTC--------- 527
Sdn45E68A       CTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTTTGTTGTCGAGCTGTTTCTTAAT 505
Sdn45E68B       CTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTT-GTG--CGAGGC---------- 526
                ************************************ **   ****


sdn4WT          GACAGATGGCGTGGCCTCTT--- 558
Sdn4N2A         GACAGATGGCKKGGGCCTCTATG 528
Sdn4N2B         ----------------------
Sdn45E68A       GACAGATGGCKKGGGCCTWTA-- 526
Sdn45E68B       ----------------------
```

```
Sdn5WT          -CAAGCCTATCCGTTCCGTCTGTTCCGGATGGGCAAATATACACAAAAATATGAGACTTG 59
Sdn5N2A         ----------------------------------------CAAAATATGAGACTTG 16
Sdn5N2B         TCAAGCCTATCCGTTCCGTCTGTTCCGGATGGGCAAATATACACAAAAATATGAGACTTG 60
Sdn55E68A       -----------------------------------GCATTAMCAGAATATGAGACTTG 23
Sdn55E68B       TCAAGCCTATYCGTTCCGTCTGTTCCGGATGGGCAAATATACACAAAAATATGAGACTTG 60
                                                    * ************

Sdn5WT          CATAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCA 119
Sdn5N2A         CATAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCA 76
Sdn5N2B         CATAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCA 120
Sdn55E68A       CATAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCA 83
Sdn55E68B       CATAAAAGTATGCATACTTTTTGCATTCCGGACACTCATCCCGTCCATTTTTCGCGACCA 120
                ************************************************************

Sdn5WT          GTCTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTTTGTTGTCGAGCTGTTTCTTA 179
Sdn5N2A         GTCTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTTTGTTGTCGAGCTGTTTCTTA 136
Sdn5N2B         GTCTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTTTGTTGTCGAGCTGTTTCTTA 180
Sdn55E68A       GTCTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTTTGTTGTCGAGCTGTTTCTTA 143
Sdn55E68B       GTCTATCAATCAATGTGTACAGCTCGAGGATGTGCTAGTTTGTTGTCGAGCTGTTTCTTA 180
                ************************************************************

Sdn5WT          ATGACAGATGGCGTGGCCTCTTCTTCGTCTTCTTGTGCTTGTGTCACAGAATGGGGACCC 239
Sdn5N2A         ATGACAGATGGCGTGGCCTCTTCTTCGTCTTCTTGTGCTTGTGTCACAGAATGGGGACCC 196
Sdn5N2B         ATGACAGATGGCGTGGCCTCTTCTTCGTCTTCTTGTGCTTGTGTCACAGAATGGGGACCC 240
Sdn55E68A       ATGACAGATGGCGTGGCCTCTTCTTCGTCTTCTTGTGCTTGTGTCACAGAATGGGGACCC 203
Sdn55E68B       ATGACAGATGGCGTGGCCTCTTCTTCGTCTTCTTGTGCTTGTGTCACAGAATGGGGACCC 240
                ************************************************************

Sdn5WT          CTTCGTATGAATGATGCACCACCACACATACTCTTCTTTCCCCCATGTGTCTAATTATTC 299
Sdn5N2A         CTTCGTATGAATGATGCACCACCACACATACTCTTCTTTCCCCCATGTGTCTAATTATTC 256
Sdn5N2B         CTTCGTATGAATGATGCACCACCACACATACTCTCTTACCCCC----------------- 283
Sdn55E68A       CTTCGTATGAATGATGCACCACCACACATACTCTTCTTTCCCCCATGTGTCTAATTATTC 263
Sdn55E68B       CTTCGTATGAATGATGCACCACCACACATACTCTCTCCCCC------------------- 281
                ********************************     **

Sdn5WT          CTGGTGGGAAGACGATGAGAG- 320
Sdn5N2A         CTGGTGGGAAGACRATGAGAGA 278
Sdn5N2B         ----------------------
Sdn55E68A       CTGGTGGGAAGACRATGAGAG- 284
Sdn55E68B       ----------------------
```

```
Sdn6WT          -CAGAATGGGGACCCCTTCGTATGAATGATGCACCACCACACATACTCTTCTTTCCCCCA 59
Sdn6N2A         ------------------------------------ACACACATACTCTTCTTTCCCC-A 23
Sdn6N2B         TCAGAATGGGGACCCCTTYGTATGAATGATGCACCACCACACATACTCTTCTTTCCCCCA 60
Sdn65E68A       ------------------------------------ACATACTCTTCTTTCCCC-A 19
Sdn65E68B       TCAGAAWRRKGACCCMTTYGTATKAWWRKKGCACCACCSCWCATKCTMTWSTTTCCCCCA 60
                                                     *** ** *  ******* *

Sdn6WT          TGTGTCTAATTATTCCTGGTGGGAAGACGATGAGAGCTCTTCTTTCTTATTGAAAGAAAC 119
Sdn6N2A         TGTGTCTAATTATTCCTGGTGGGAAGACGATGAGAGCTCTTCTTTCTTATTGAAAGAAAC 83
Sdn6N2B         TGTGTCTAATTATTCCTGGTGGGAAGACGATGAGAGCTYTTCTTTYTTATTGAAAGAAAC 120
Sdn65E68A       TGTGTCTAATTATTCCTGGTGGGAAGACGATGAGAGCTCTTCTTTCTTATTGAAAGAAAC 79
Sdn65E68B       TSTGTSTAATTAWTCCTGGTGGGAAGMCGATGASAGCTYTTCTTTCTTATTGAAAGAAAC 120
                * *** ****** ************* ****** **** ****** **************

Sdn6WT          CAACAAGTATTGGTGAATGACACGCCAGAAATTATATTATTGAAAACGGAGAGATGATGA 179
Sdn6N2A         CAACAAGTATTGGTGAATGACACGCCAGAAATTATATTATTGAAAACGGAGAGATGATGA 143
Sdn6N2B         CAACAAGTATTGGTGAATGACACGCCAGAAATTATATTATTGAAAACGGAGAGATGATGA 180
Sdn65E68A       CAACAAGTATTGGTGAATGACACGCCAGAAATTATATTATTGAAAACGGAGAGATGATGA 139
Sdn65E68B       CAACAAGTATTGGKGAATGACAMGCCAGAAATTATATTAWTGAAAACGGMRAGATGATGA 180
                ************* ******** ***************** *********  *********

Sdn6WT          TGATGATGAGGATCAAGATGGGTCCAAATTATATATGAGTGCCACTGAACGAGACACTTC 239
Sdn6N2A         TGATGATGAGGATCAAGATGGGTCCAAATTATATATGAGTGCCACTGAACGAGACACTTC 203
Sdn6N2B         TGATGATGAGGATCAAGATGGGTCCAAATTATATATGAGTGCCACTGAACGAGACACTTY 240
Sdn65E68A       TGATGATGAGGATCAAGATGGGTCCAAATTATATATGAGTGCCACTGAACGAGACACTTC 199
Sdn65E68B       TGATGATGAGGATCAAGWTGGGTCCAAATTATATAWGAGTGCCYCYGAACGAGACACTTY 240
                ****************** ****************** ******* * *************

Sdn6WT          TCAAGTTTTCTATTTTGGCGCAAAAATGTAGGCAGAGATGTTTAGTTTTTTTCTTTCCAC 299
Sdn6N2A         TCAAGTTTTCTATTTTGGCGCAAAAATGTAGGAAGAGATGTTTAGTTTTTTTCTTTCCAC 263
Sdn6N2B         TCAAGTTTTYTATTTTGGCGCAAAAATGTAGGAAGAGATGTTTAGTTTTTTTCTTTCCAC 300
Sdn65E68A       TCAAGTTTTCTATTTTGGCGCAAAAATGTAGGAAGAGATGTTTAGTTTTTTTCTTTCCAC 259
Sdn65E68B       TCAAGTWTTYTATTTTGGCGCAAAAATGTAGGAAGAGATGTTTAGTTTTTTTCTTTCCAC 300
                ****** ** ********************** ****************************

Sdn6WT          AACACTTGCTTTGAACAAACTTTTTTGTGTTTACTGCCTTTTGTGAGGAAGTACAGATGC 359
Sdn6N2A         AACACTTGCTTTGAACAAACTTTTTTGTGTTTACTGCYTTTTGTGAGGAAGTACAGATGC 323
Sdn6N2B         AACACTTGCTTTGAACAAACTTTTTTGTGTTTACTGCCTTTTGTGAGGAAGTACAGATGC 360
Sdn65E68A       AACACTTGCTTTGAACAAAYTTTTTTGTGTTTACTGCCTTTTGTGAGGAAGTACAGATGC 319
Sdn65E68B       AACACTTGCTTTGAACAAACTTTTTTGTGTTTACTGCCTTTTGTGAGGAAGTACAGATGC 360
                ****************** ***************** **********************

Sdn6WT          CTTGAAATGGGTTTTTAAAGAAAACATAGACCCCTTATGACGCGTGTCTCAAAATGTGGG 419
Sdn6N2A         CTTGAAATGGGTTTTTAAARAAAACATAGACCCCTTATGACGCGTGTCTCAAAATGTGGG 383
Sdn6N2B         CTTGAAATGGGTTTTTAAAGAAAACATAGACCCCTTATGACGCGTGTCTCAAAATGTGGG 420
Sdn65E68A       CTTGAAATGGGTTTTTAAAGAAAACATAGACCCCTTATGACGCGTGTCTCAAAATGTGGG 379
Sdn65E68B       CTTGAAATGSGTTTTTAAAGAAAACATAGACCCCTTATGACGCGTGTCTCAAAATGTGGG 420
                ********* ********* ************************************

Sdn6WT          GAAAGACGAATTGTTCAGTAAAGTTGACATCTCTTCTCAAAACTTGCTATGCTTTCTTCT 479
Sdn6N2A         GAAAGACRAATTGTTCAGTAAAGTTGACATCTCTTCTCAAAACTTGCTATGCTTTCTTCT 443
Sdn6N2B         GAAAGACGAATTGTTCAGTAAAGTTGACATCTCTTCTCAAAACTTGCTATGCTTTCTTCT 480
Sdn65E68A       GAAAGACRAATTGTTCAGTAAAGTTGACATCTCTTCTCAAAACTTGCTATGCTTTCTTCT 439
Sdn65E68B       GAAAGACGAAWTGTTCAGTAAAGTKGACATCTCTTCTCAAAACTTGCTATGCTTTCTTCT 480
                ******* ** ************* ***********************************
```

```
Sdn6WT          TTTGTGGTTATCAAAATCTGAGCAGCATCCCACATCTAACAAAGGAGCCCCAGCTCTATT 539
Sdn6N2A         TTTGTGGTTATCAAAATCTGAGCAGCATCCCACATCTAACAAAGGAGCCCCAGYTCTATT 503
Sdn6N2B         TTTGTGGTTATCAAAATCTGAGCAGCATCCCACATCTAACAAAGGAGCCCCAGCTCTAT- 539
Sdn65E68A       TTTGTGGTTATCAAAATCTGAGCAGCATCCCACATCTAACAAAGGASCCCCAKYTCTATT 499
Sdn65E68B       TTTGTGGTTATCAAAATCTGAGCAGCATCCCACATCTAACAAAGGAGCCCCAGCTCTATC 540
                ********************************************** *****  *****
```

```
Sdn6WT          CTCTTGAGAAAAAGGGCCTGAGGTTCGGGATGGTGGGACGGAAG- 583
Sdn6N2A         CTCTTGAGAAAAARGGCCTGAGGTTCGGGATGGTGGGACGGAAGA 548
Sdn6N2B         CTCKAGAAATTCTT------------------------------- 553
Sdn65E68A       CTCTTGAGAAWAAGGGCYTKARGTTCGGRATGGTGGGACGGAAG- 543
Sdn65E68B       TCTKWKAAAAACGT------------------------------- 554
                     * *
```

```
Sdn7WT          -CTGAGCAGCATCCCACATCTAACAAAGGAGCCCCAGCTCTATTCTCTTGAGAAAAAGGG 59
Sdn7N2A         ----------------------------------------CTTGAGAAAAGRKC 14
Sdn7N2B         TCTGAGCAGCCACCCACATCTAACAAAGGAGCCCCAGCTCTATTCTCTTGAGAAAAAGGG 60
Sdn75E68A       --------------------------------GCAGCTCTATTCTCTTGAGAAA--GGG 25
Sdn75E68B       ------------------------------------------------------------


Sdn7WT          CCTGAGGTTCGGGATGGTGGGACGGAAGCCGCTGAAGAAAGAAGGAAGGCAAGGTCAAGG 119
Sdn7N2A         CKCRARGTTCRGGATGGTGGGACGSAAGCCGCWGAW-MAASMAGGAMRGCAAGGTCAAGG 73
Sdn7N2B         CCTGAGGTTCGGGATGGTGGGACGGAAGCCGCTGAAGAAAGAAGGAAGGCAAGGTCAAGG 120
Sdn75E68A       CCTGAGGTTCGGGATGGTGGGACGGAAGCCGCTGAAGAAAGAAGGAAGGCAAGGTCAAGG 85
Sdn75E68B       ------------------------------------------------------------


Sdn7WT          GTCGCATTTTGTGCATATTTGATTTTATGACACTGAGTGGAGGAAGTGTCTGTGAGGAAA 179
Sdn7N2A         GTCSCATTTTGYGCATATTTGATTTTMTRACACTGASTGGAGGAAGKGYCWGYGAGGAMA 133
Sdn7N2B         GTCGCATTTTGTGCATATTTGATTTTATGACACTGAGTGGAGGAAGTGTYTGTGAGGAAA 180
Sdn75E68A       GTCGCATTTTGTGCATATTTGATTTTATGACACTGAGTGGAGGAAGTGTCTGTGAGGAAA 145
Sdn75E68B       ------------------------------------------------------------


Sdn7WT          AGGGGAACCTTTTGAATGAAAGGTTTTGAGAGAAGATGACGGTTTTGAAAGGATTTAGTG 239
Sdn7N2A         AGGGGAMCCTTTTGAAKGAAAGGTTYTGMGASAASATGACSGTTTTGAAMGGATTKAGTG 193
Sdn7N2B         AGGGGAACCTTTTGAATGAAAGGTTTTGAGAGAAGATGACGGTTTTGAAAGGATTTAGTG 240
Sdn75E68A       AGGGGAACCTTTTGAATGAAAGGTTTTGAGAGAAGATGACGGTTTTGAAAGGATTTAGTG 205
Sdn75E68B       ------------------------------------------------------------


Sdn7WT          AGTTGGAAAGGAAAACTAAAATTTTTTTTCACTTAGTTACAAAATTAAACTCTTATTGAA 299
Sdn7N2A         ASTTGGAAAGGAAAACTAAAATTTTTTTTCMCTTAKYTACAAAATTAAACSCTTATTGAA 253
Sdn7N2B         AGTTGGAAAGGAAAACTAAAATTTTTTTTCACTTAGTTACAAAATTAAACTCTTATTGAA 300
Sdn75E68A       AGTTGGAAAGGAAAACTAAAATTTTTTTTCACTTAGTTACAAAATTAAACTCTTATTGAA 265
Sdn75E68B       ------------------------------------------------------------


Sdn7WT          ATTATTCACTACCAAGAAAGTTCAATTTAGGACAATGCAATATTTCTTACATGCATCGCA 359
Sdn7N2A         ATTATTCACTACCAARAAAGTTCAATTTAGGACAATGCAATATTTCTTACATGCATCGCA 313
Sdn7N2B         ATTATTCACTACCAAGAAAGTTCAATTTAGGACAATGCAATATTTCTTACATGCATCGCA 360
Sdn75E68A       ATTATTCACTACCAAGAAAGTTCAATTTAGGACAATGCAATATTTCTTACATGCATCGCA 325
Sdn75E68B       ------------------------------------------------------------


Sdn7WT          TTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAAAATATGTTTGGAATAAAG 419
Sdn7N2A         TTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAAAATATGTTTGGAATAAAG 373
Sdn7N2B         TTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAAAATATGTTTGGAATAAAG 420
Sdn75E68A       TTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAAAATATGTTTGGAATAAAG 385
Sdn75E68B       ------------------------------------------------------------


Sdn7WT          CACAAAGCTTATTTATTTTTTAGTAACGAAAAGATTTAAAGTTGTAGAATTATCTATTTT 479
Sdn7N2A         CACAAAGCTTATTTATTTTTTAKTAACGAAAAGATTTAAAGTTGTARAATTATCTATTTT 433
Sdn7N2B         CACAAAGCTTATTTATTTTTTAGTAACGAAAAGATTTAAAGTTGTAGAATTATCTATTTT 480
Sdn75E68A       CACAAAGCTTATTTATTTTTTAGTAACGAAAAGATTTAAAGTTGTAGAATTATCTATTTT 445
Sdn75E68B       ------------------------------------------------------------


Sdn7WT          AATTTGATTTAATTTGATATTTTCAGCTATTGCCGGAGGTGTTCTCGTTGCGGTG- 534
Sdn7N2A         AATTTGATTTAATTTGATATTTTCAGCTATTGCCGGAGGTGTTCTCGTTSCSGKG- 488
Sdn7N2B         AATTGA--TTAATTGAATG---------------------------------- 497
Sdn75E68A       AATTTGATTTAATTTGATATTTTCAGCTATTGCCGGAGGTGTTCTCGTTGCGGTGA 501
Sdn75E68B       ----------------A---------------------------------- 1
```

```
Sdn8WT          --GTGTCTGTGAGGAAAAGGGGAACCTTTTGAATGAAAGGTTTTGAGAGAAGATGACGGT 58
Sdn8N2A         ----------------------------------------------------------CGGT 4
Sdn8N2B         TTKTGTYTGTTGGGAAAAGGGGAACCTTTTGAATGAAAGGTTTTGAGAGAAGATGACGGT 60
Sdn85E68A       ------------------------------------------TGAGAGAGATGAC 13
Sdn85E68B       --GTGTYTGTGAGGAAAAGGGGAACCTTTTGAATGAAAGGTTTTGAGAGAAGATGACGGT 58
                                                                        *

Sdn8WT          T-TTGAAAGGATTTAGTGAGTTGGAAAGGAAAACTAAAATTTTTTTTCACTTAGTTACAA 117
Sdn8N2A         TTTYGAAAGGWTTYMGTGAGTTGGAAAGGAAAACTAAAATTYTTTYYYACTTAGTTACAA 64
Sdn8N2B         T-TTGAAAGGATTTAGTGAGTTGGAAAGGAAAACTAAAATTTTTTTTCACTTAGTTACAA 119
Sdn85E68A       GGTTTTGAAGGWTTAGTGAGTTGGAAAGGAAAACTAAAATTTTTTTTCACTTAGTTACAA 73
Sdn85E68B       T-TTGAAAGGATTTAGTGAGTTGGAAAGGAAAACTAAAATTTTTTTTCACTTAGTTACAA 117
                  *    * *  *  *************************** ***   ************

Sdn8WT          AATTAAACTCTTATTGAAATTATTCACTACCAAGAAAGTTCAATTTAGGACAATGCAATA 177
Sdn8N2A         AATTAAACTCTTATTGAAATTATTCACTACCAAGAAAGTTCAATTTAGGACAATGCAATA 124
Sdn8N2B         AATTAAACTCTTATTGAAATTATTCACTACCAAGAAAGTTCAATTTAGGACAATGCAATA 179
Sdn85E68A       AATTAAACTCTTATTGAAATTATTCACTACCAAGAAAGTTCAATTTAGGACAATGCAATA 133
Sdn85E68B       AATTAAACTCTTATTGAAATTATTCACTACCAAGAAAGTTCAATTTAGGACAATGCAATA 177
                ************************************************************

Sdn8WT          TTTCTTACATGCATCGCATTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAA 237
Sdn8N2A         TTTCTTACATGCATCGCATTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAA 184
Sdn8N2B         TTTCTTACATGCATCGCATTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAA 239
Sdn85E68A       TTTCTTACATGCATCGCATTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAA 193
Sdn85E68B       TTTYTTACATGCATCGCATTACTGATATTATATGTTGTCCTTACGGGTTGAGATCATAAA 237
                *** ********************************************************

Sdn8WT          AATATGTTTGGAATAAAGCACAAAGCTTATTTATTTTTTAGTAACGAAAAGATTTAAAGT 297
Sdn8N2A         AATATGTTTGGAATAAAGCACAAAGCTTATTTATTTTTTAGTAACGAAAAGATTTAAAGT 244
Sdn8N2B         AATATGTTTGGAATAAAGCACAAAGCTTATTTATTTTTTAGTAACGAAAAGATTTAAAGT 299
Sdn85E68A       AATATGTTTGGAATAAAGCACAAAGCTTATTTATTTTTTAGTAACGAAAAGATTTAAAGT 253
Sdn85E68B       AATATGTTTGGAATAAAGCACAAAGCTTATTTATTTTTTAGTAACGAAAAGATTTAAAGT 297
                ************************************************************

Sdn8WT          TGTAGAATTATCTATTTTAATTTGATTTAATTTGATATTTTCAGCTATTGCCGGAGGTGT 357
Sdn8N2A         TGTARAATTATCTATTTTAATTTGATTTAATTTGATATTTTCAGCTATTGCCGGAGGTGT 304
Sdn8N2B         TGTAGAATTATCTATTTTAATTTGATTTAATTTGATATTTTCAGCTATTGCCGGAGGTGT 359
Sdn85E68A       TGTAGAATTATCTATTTTAATTTGATTTAATTTGATATTTTCAGCTATTGCCGGAGGTGT 313
Sdn85E68B       TGTAGAATTATCTATTTTAATTTGATTTAATTTGATATTTTCAGCTATKGCCGGMKGW-T 356
                **** ******************************************* *****  *  *

Sdn8WT          TCTCGTTGCGGTGATTACAGCCATTCTACTCGTGCTCTTTGTAGTCTTTCGAATCAGGAA 417
Sdn8N2A         TCTCGTTGCGGTGATTACAGCCATTCTACTCGTGCTCTTTGTAGTCTTTCRAATCAGGAA 364
Sdn8N2B         TCTCGTTGCGGTGATTACAGCCATTCTACTCGTGCTCTTTGTAGTCTTTCGAATCAGG-- 417
Sdn85E68A       TCTCGTTGCGGTGATTACAGCCATTCTACTCGTGCTCTTTGTAGTCTTTCRAATCAGGAA 373
Sdn85E68B       TCTSGTTGMGWTGRYTMYWMCCMTYWTACTSGTGCTCKTTGTWTCSAKGMWWCWGAGSMW 416
                *** **** * **  *    ** *  **** ****** ****             **

Sdn8WT          AAA-AGATGAAGGGTCATACGCATTGGATGAACCCAAGCAAGCAAGAC- 464
Sdn8N2A         AAA-ARATGAAGGGTCATACGCATTGGATGAACCCAAGAAAGCAARAMY 412
Sdn8N2B         ------------------------------------------------
Sdn85E68A       AAA-AGATGAAGGGTCATACGCATTGGATGAACCCAAGAAAGCAARAMY 421
Sdn85E68B       AGACAAGTRWWRWKTCCMRYKYGC----------------------- 440
```

```
Sdn9WT          --AGGTGTTCTCGTTGCGGTGATTACAGCCATTCTACTCGTGCTCTTTGTAGTCTTTCGA 58
Sdn9N2A         ------------------------CGGMMGTC--WCTCGTGCTCTTTGTAGTCTTTCGA 33
Sdn9N2B         TRAWGKGKTYYYSKKGSGGGGRWTWMMSSCMWTYTWMTYKKGKTTTTTKKWRKYTTTYGR 60
Sdn95E68A       ------------------------AGCMATC--WCTCGTGCTCTTTGTAGTCTTTCGA 32
Sdn95E68B       TSWRWGKKTYYYGKKGSGGKGRWTWMMRSCMWTYTWMTYGKKSTYTTTKKWRTYTTTYGR 60
                                                *     * ***      *** *

Sdn9WT          ATCAGGAAAAAAGATGAAGGGTCATACGCATTGGATGAACCCAAGCAAGCAAGACCATAT 118
Sdn9N2A         -TCAGGAAAAAAGATGAAGGGTCATACGCATTGGATGAACCCAAGCAAGCAAGACCATAT 92
Sdn9N2B         AWYMRGRAAAAARRWGRARGGKYWWWMSCMWTGGRWKRAMCCMARCMARRMARRMCMWWW 120
Sdn95E68A       -TCAGGAAAAAAGATGAAGGGTCATACGCATTGGATGAACCCAAGCAAGCAAGACCATAT 91
Sdn95E68B       AWYMRGRAAAAARRWGRARGGKYWWWMSCMWTGGRWGRAMCCMARSMARRMARRCCMWWW 120
                     * *****   * * **       *  ***    * ** *   *   *   *

Sdn9WT          GCCTCGTATGGTTATACCAAAGCATCGACAAAAGAATTTTACGCGTAATCTCTACTGTCA 178
Sdn9N2A         GCCTCGTATGGTTATACCAAAGCATCGACAAAAGAATTTTACGCGTAATCTCTACTGTCA 152
Sdn9N2B         KCCTYGKWWGGKTWTWMCMAARSMWYSRMMAAAGRAWTTTWMGSGKWAWYTYTWWYKKYM 180
Sdn95E68A       GCCTCGTATGGTTATACCAAAGCATCGACAAAAGAATTTTACGCGTAATCTCTACTGTCA 151
Sdn95E68B       KCCTYGKWWKGKTWWWMCMAARSKWYGRMMAAARRAWTTTWRGSGKWAWYTYTWMYKKYM 180
                 *** *    * *    * **         ***  * ***  * *  *  * *

Sdn9WT          TTTGTTCAAAATCTTCTATCACTCAATCACCTTTCAAATCATTTTTATGATTCTGTTCTC 238
Sdn9N2A         TTTGTTCAAAATCTTCTATCACTCAATCACCTTTCAAATCATTTTTATGATTCTGTTCTC 212
Sdn9N2B         WTTKKTYMAAATTTTYTWWCMMYYMAWYMCCYTTYMAATYMWTTTTWWKRWTYYKKTTTY 240
Sdn95E68A       TTTGTTCAAAATCTTCTATCACTCAATCACCTTTCAAATCATTTTTATGATTCTGTTCTC 211
Sdn95E68B       WTTKKTYMAAAWYTTYTWWCMMYCMATYMCCYTTYAAAWYMWTTTTWWGRWTYSGKTYTY 240
                 **  *  ***  ** *  *     *   ** **  **    ****     *    * *

Sdn9WT          AGACTTCATTCCAATCTGCCATACCTTTCAATTTGTTTTTTTCCCACTCCCATTTTTTTC 298
Sdn9N2A         AGACTTCATTCCAATCTGCCATACCTTTCAATTTGTTTTTTTCCCACTCCCATTTTTTTC 272
Sdn9N2B         MRRMYTYMWTYCMATTTKSCMWWCCYTTYMAWTTKKTTTTTTYCCMMYYCCMWTTTTTTY 300
Sdn95E68A       AGACTTCATTCCAATCTGCCATACCTTTCAATTTGTTTTTTTCCCACTCCCATTTTTTTC 271
Sdn95E68B       SRRMTTYMWTYCMATTTKCCMWWCCYTTYMAWTTKKTTTTTTYCCMMYYCCMWTTTTTTY 300
                     *   * * ** *  *   ** **  * **  ****** **    **  ******

Sdn9WT          AAACCCTCCCCCCCCCCGCCTTCCTTTCGTAAAGGTCATTACTCTCTGTTCTACTCGTGA 358
Sdn9N2A         AAACCCTCCCCCCCCCSSCYTYCYTTYCKWAARGGYMWTWAYYYYYKKTYYYYYYSKWRA 332
Sdn9N2B         MAAMCCYYCCCCCCCCCGCCTTCCTTTCGTAAAGGTCATTACTMTSTGTTCTACTCGTGA 360
Sdn95E68A       AAACCCTCCCCCCCCCCSCYTYCYTTYCSWAARGKYMWTWYYYYYYKKSYCWYYYCKKRW 331
Sdn95E68B       MAAMCCYYCCCCCCCCCGCCTTCCTTTCGTAAAGGTCATTACTCTCTGTTCTACTCGTGA 360
                 ** **  ********  * * * ** *  ** *    *

Sdn9WT          TAATTTGATAATATAAACTGATCTGACTCCATGGTGCCAAATATTTGAATATAATCTTTT 418
Sdn9N2A         WAWTTKRAWAAWAWAAMYKRYYYKRAYYCMRKGGKGCCAAAWWTTKRAWWWWAWYYTTTK 392
Sdn9N2B         TAATTTGATAATATAAACTGATCTGACTCCATGGTGCCAAATATTG--AAWWWWYCTT-- 416
Sdn95E68A       WAWTTTRAWAAWAWAAAYKRAYYKRMYYCCWKGGKGCCAAAWWTTTRAAWWWAWYYTTTK 391
Sdn95E68B       TAATTTGATAATATAAACTGATCTGACTCCATGGTGCCAAATATK----AATWWTCT--- 413
                 * **  * ** * **           *   ** ******  *           *

Sdn9WT          GTAGACCCACTTAGGGGTAGGGA-- 441
Sdn9N2A         GWAAMCCCMYTWAGGGGKWRGGRAA 417
Sdn9N2B         -------------------------
Sdn95E68A       GRARMCCCMYTWAGGGKWRGGRAA- 415
Sdn95E68B       -------------------------
```

6.8 SDN-1 protein alignments from translated *sdn-1* sequences.

```
SdnWormbase     MILKLNFCLSTYSVLILLSLSTQAFAANQAKTKVVPSSTISTKSLKNGISEQVEGSANIP 60
SdnWt           MILKLNFCLSTYSVLILLSLSTQAFAANQAKTKVVPSSTISTKSLKNGISEQVEGSANIP 60
SdnN            MILKLNFCLSTYSVLILLSLSTQAFAANQAKTKVVPSSTISTKSLKNGISEQVEGSANIP 60
SdnEv           MILKLNFCLSTYSVLILLSLSTQAFAANQAKTKVVPSSTISTKSLKNGISEQVEGSANIP 60
                ************************************************************

SdnWormbase     GRLADIEVNGSGYPTDDEDGDDVHGSGKPPSSATTKSDKVTSPSHAVVTAKPTTVPTTTA 120
SdnWt           GRLADIEVNGSGYPTDDEDGDDVHGSGKPPSSATTKSDKVTSPSHAVVTAKPTTVPTTTA 120
SdnN            GRLADIEVNGSGYPTDDEDGDDVHGSGKPPSSATTKSDKVTSPSHAVVTAKPTTVPTTTA 120
SdnEv           GRLADIEVNGSGYPTDDEDGDDVHGSGKPPSSATTKSDKVTSPSHAVVTAKPTTVPTTTA 120
                ************************************************************

SdnWormbase     SFKPPVQPKPKPAANDKEIKVEEDEDDDEDEDEDDEDDEEDFADENIHNDEDFFTTTTTT 180
SdnWt           SFKPPVQPKPKPAANDKEIKVEEDEDDDEDEDEDDEDDEEDFADENIHNDEDFFTTTTTT 180
SdnN            SFKPPVQPKPKPAANDKEIKVEEDEDDDEDEDEDDEDDEEDFADENIHNDEDFFTTTTTT 180
SdnEv           SFKPPVQPKPKPAANDKEIKVEEDEDDDEDEDEDDEDDEEDFADENIHNDEDFFTTTTTT 180
                ************************************************************

SdnWormbase     TYRPIVVATTSTPRSAATNPPRQQPPMVTSTISSGPFSPFHETLANGFYAAIAGGVLVAV 240
SdnWt           TYRPIVVATTSTPRSAATNPPRQQPPMVTSTISSGPFSPFHETLANGFYAAIAGGVLVAV 240
SdnN            TYRPIVVATTSTPRSAATNPPRQEPPMVTSTISSGPFSPFHETLANGFYAAIAGGVLVAV 240
SdnEv           TYRPIVVATTSTPRSAATNPPRQ-PPMVTSTISSGPFSPFHETLANGFYAAIAGGVLVAV 239
                ********************** ************************''''''''''''

SdnWormbase     ITAILLVLFVVFRIRKKDEGSYALDEPKQARPYASYGYTKASTKEFYA- 2
SdnWt           ITAILLVLFVVFRIRKKDEGSYALDEPKQARPYASYGYTKASTKEFYA- 2
SdnN            ITAILLVLFVVFRIRKKDEGSYALDEPKQARPYASYGYTKASTKEFYA- 2
SdnEv           ITAILLVLFVVFRIRKKDEGSYALDEPKQARPYASYGYTKASTKEFYA- 2
                ************************************************
```

7 REFERENCES.

Ackerman SL, Kozak LP, Przyborski SA, Rund LA, Boyer BB, Knowles BB (1997) The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein. Nature 386:838-842

Baeg GH, Perrimon N (2000) Functional binding of secreted molecules to heparan sulfate proteoglycans in drosophila. Curr Opin Cell Biol 12:575-580

Barallobre MJ, Pascual M, Del Rio JA, Soriano E (2005) The netrin family of guidance factors: Emphasis on netrin-1 signalling. Brain Res Brain Res Rev 49:22-47

Basbaum CB, Werb Z (1996) Focalized proteolysis: Spatial and temporal regulation of extracellular matrix degradation at the cell surface. Curr Opin Cell Biol 8:731-738

Bashaw GJ, Kidd T, Murray D, Pawson T, Goodman CS (2000) Repulsive axon guidance: Abelson and enabled play opposing roles downstream of the roundabout receptor. Cell 101:703-715

Bennett KL, Bradshaw J, Youngman T, Rodgers J, Greenfield B, Aruffo A, Linsley PS (1997) Deleted in colorectal carcinoma (DCC) binds heparin via its fifth fibronectin type III domain. J Biol Chem 272:26940-26946

Bernfield M, Gotte M, Park PW, Reizes O, Fitzgerald ML, Lincecum J, Zako M (1999) Functions of cell surface heparan sulfate proteoglycans. Annu Rev Biochem 68:729-777

Blelloch R, Anna-Arriola SS, Gao D, Li Y, Hodgkin J, Kimble J (1999) The gon-1 gene is required for gonadal morphogenesis in caenorhabditis elegans. Dev Biol 216:382-393

Bornemann DJ, Duncan JE, Staatz W, Selleck S, Warrior R (2004) Abrogation of heparan sulfate synthesis in drosophila disrupts the wingless, hedgehog and decapentaplegic signaling pathways. Development 131:1927-1938

Brenner S (1974) The genetics of caenorhabditis elegans. Genetics 77:71-94

Brümmendorf T, Rathjen F.G (1996) Structure/function relationships of axon-associated adhesion receptors of the immunoglobulin superfamily. Curr Opin Neurobio 6:584-593.

Burdine RD, Branda CS, Stern MJ (1998) EGL-17(FGF) expression coordinates the attraction of the migrating sex myoblasts with vulval induction in C. elegans. Development 125:1083-1093

Burdine RD, Chen EB, Kwok SF, Stern MJ (1997) Egl-17 encodes an invertebrate fibroblast growth factor family member required specifically for sex myoblast migration in caenorhabditis elegans. Proc Natl Acad Sci U S A 94:2433-2437

Chan SS, Zheng H, Su MW, Wilk R, Killeen MT, Hedgecock EM, Culotti JG (1996) UNC-40, a C. elegans homolog of DCC (deleted in colorectal cancer), is required in motile cells responding to UNC-6 netrin cues. Cell 87:187-195

Chang C, Yu TW, Bargmann CI, Tessier-Lavigne M (2004) Inhibition of netrin-mediated axon attraction by a receptor protein tyrosine phosphatase. Science 305:103-106

Charron F, Tessier-Lavigne M (2005) Novel brain wiring functions for classical morphogens: A role as graded positional cues in axon guidance. Development 132:2251-2262

Colavita A, Culotti JG (1998) Suppressors of ectopic UNC-5 growth cone steering identify eight genes involved in axon guidance in caenorhabditis elegans. Dev Biol 194:72-85

Colavita A, Krishna S, Zheng H, Padgett RW, Culotti JG (1998) Pioneer axon guidance by UNC-129, a C. elegans TGF-beta. Science 281:706-709

Dickson BJ (2002) Molecular mechanisms of axon guidance. Science 298:1959-1964

Ding K, Lopez-Burks M, Sanchez-Duran JA, Korc M, Lander AD (2005) Growth factor-induced shedding of syndecan-1 confers glypican-1 dependence on mitogenic responses of cancer cells. J Cell Biol 171:729-738

Fujise M, Takeo S, Kamimura K, Matsuo T, Aigaki T, Izumi S, Nakato H (2003) Dally regulates dpp morphogen gradient formation in the drosophila wing. Development 130:1515-1522

Geisbrecht BV, Dowd KA, Barfield RW, Longo PA, Leahy DJ (2003) Netrin binds discrete subdomains of DCC and UNC5 and mediates interactions between DCC and heparin. J Biol Chem 278:32561-32568

Gitai Z, Yu TW, Lundquist EA, Tessier-Lavigne M, Bargmann CI (2003) The netrin receptor UNC-40/DCC stimulates axon attraction and outgrowth through enabled and, in parallel, rac and UNC-115/AbLIM. Neuron 37:53-65

Gu T, Orita S, Han M (1998) Caenorhabditis elegans SUR-5, a novel but conserved protein, negatively regulates LET-60 ras activity during vulval induction. Mol Cell Biol 18:4556-4564

Hall DH, Winfrey VP, Blaeuer G, Hoffman LH, Furuta T, Rose KL, Hobert O, Greenstein D (1999) Ultrastructural features of the adult hermaphrodite gonad of caenorhabditis elegans: Relations between the germ line and soma. Dev Biol 212:101-123

Hamelin M, Zhou Y, Su MW, Scott IM, Culotti JG (1993) Expression of the UNC-5 guidance receptor in the touch neurons of C. elegans steers their axons dorsally. Nature 364:327-330

Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57-70

Hedgecock EM, Culotti JG, Hall DH (1990) The unc-5, unc-6, and unc-40 genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in C. elegans. Neuron 4:61-85

Hedrick L, Cho KR, Fearon ER, Wu TC, Kinzler KW, Vogelstein B (1994) The DCC gene product in cellular differentiation and colorectal tumorigenesis. Genes Dev 8:1174-1183

Hill RJ, Sternberg PW (1992) The gene lin-3 encodes an inductive signal for vulval development in C. elegans. Nature 358:470-476

Hong K, Hinck L, Nishiyama M, Poo MM, Tessier-Lavigne M, Stein E (1999) A ligand-gated association between cytoplasmic domains of UNC5 and DCC family receptors converts netrin-induced growth cone attraction to repulsion. Cell 97:927-941

Howe AK (2004) Regulation of actin-based cell migration by cAMP/PKA. Biochim Biophys Acta 1692:159-174

Ishii N, Wadsworth WG, Stern BD, Culotti JG, Hedgecock EM (1992) UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in C. elegans. Neuron 9:873-881

Jorgensen EM, Mango SE (2002) The art and design of genetic screens: Caenorhabditis elegans. Nat Rev Genet 3:356-369

Kamath RS, Martinez-Campos M, Zipperlen P, Fraser AG, Ahringer J (2001) Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in caenorhabditis elegans. Genome Biol 2:RESEARCH0002

Keino-Masu K, Masu M, Hinck L, Leonardo ED, Chan SS, Culotti JG, Tessier-Lavigne M (1996) Deleted in colorectal cancer (DCC) encodes a netrin receptor. Cell 87:175-185

Killeen M, Tong J, Krizus A, Steven R, Scott I, Pawson T, Culotti J (2002) UNC-5 function requires phosphorylation of cytoplasmic tyrosine 482, but its UNC-40-independent functions also require a region between the ZU-5 and death domains. Dev Biol 251:348-366

Kimble JE, White JG (1981) On the control of germ cell development in caenorhabditis elegans. Dev Biol 81:208-219

Kokel M, Borland CZ, DeLong L, Horvitz HR, Stern MJ (1998) Clr-1 encodes a receptor tyrosine phosphatase that negatively regulates an FGF receptor signaling pathway in caenorhabditis elegans. Genes Dev 12:1425-1437

Lauffenburger DA, Horwitz AF (1996) Cell migration: A physically integrated molecular process. Cell 84:359-369

Lee J, Ishihara A, Oxford G, Johnson B, Jacobson K (1999) Regulation of cell movement is mediated by stretch-activated calcium channels. Nature 400:382-386

Lee J, Li W, Guan KL (2005) SRC-1 mediates UNC-5 signaling in caenorhabditis elegans. Mol Cell Biol 25:6485-6495

Lee JS, Chien CB (2004) When sugars guide axons: Insights from heparan sulphate proteoglycan mutants. Nat Rev Genet 5:923-935

Leung-Hagesteijn C, Spence AM, Stern BD, Zhou Y, Su MW, Hedgecock EM, Culotti JG (1992) UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in C. elegans. Cell 71:289-299

Li W, Aurandt J, Jurgense C, Rao Y, Guan KL (2006) FAK and src kinases are required for netrin-induced tyrosine phosphorylation of UNC5. J Cell Sci 119:47-55

Li W, Lee J, Vikis HG, Lee SH, Liu G, Aurandt J, Shen TL, Fearon ER, Guan JL, Han M, Rao Y, Hong K, Guan KL (2004) Activation of FAK and src are receptor-proximal events required for netrin signaling. Nat Neurosci 7:1213-1221

Li X, Meriane M, Triki I, Shekarabi M, Kennedy TE, Larose L, Lamarche-Vane N (2002a) The adaptor protein nck-1 couples the netrin-1 receptor DCC (deleted in colorectal cancer) to the activation of the small GTPase Rac1 through an atypical mechanism. J Biol Chem 277:37788-37797

Li X, Saint-Cyr-Proulx E, Aktories K, Lamarche-Vane N (2002b) Rac1 and Cdc42 but not RhoA or rho kinase activities are required for neurite outgrowth induced by the netrin-1 receptor DCC (deleted in colorectal cancer) in N1E-115 neuroblastoma cells. J Biol Chem 277:15207-15214

Lim YS, Wadsworth WG (2002) Identification of domains of netrin UNC-6 that mediate attractive and repulsive guidance and responses from cells and growth cones. J Neurosci 22:7080-7087

Lu X, Le Noble F, Yuan L, Jiang Q, De Lafarge B, Sugiyama D, Breant C, Claes F, De Smet F, Thomas JL, Autiero M, Carmeliet P, Tessier-Lavigne M, Eichmann A (2004) The netrin receptor UNC5B mediates guidance events controlling morphogenesis of the vascular system. Nature 432:179-186

Maloof JN, Whangbo J, Harris JM, Jongeward GD, Kenyon C (1999) A wnt signaling pathway controls hox gene expression and neuroblast migration in C. elegans. Development 126:37-49

Mello C, Fire A (1995) DNA transformation. Methods Cell Biol 48:451-482

Merz DC, Alves G, Kawano T, Zheng H, Culotti JG (2003) UNC-52/perlecan affects gonadal leader cell migrations in C. elegans hermaphrodites through alterations in growth factor signaling. Dev Biol 256:173-186

Merz DC, Zheng H, Killeen MT, Krizus A, Culotti JG (2001) Multiple signaling mechanisms of the UNC-6/netrin receptors UNC-5 and UNC-40/DCC in vivo. Genetics 158:1071-1080

Meuth M, Arrand JE (1982) Alterations of gene structure in ethyl methane sulfonate-induced mutants of mammalian cells. Mol Cell Biol 2:1459-1462

Minniti AN, Labarca M, Hurtado C, Brandan E (2004) Caenorhabditis elegans syndecan (SDN-1) is required for normal egg laying and associates with the nervous system and the vulva. J Cell Sci 117:5179-5190

Montell DJ (2003) Border-cell migration: The race is on. Nat Rev Mol Cell Biol 4:13-24

Mohrlen F, Hutter H, Zwilling R (2003) The astacin protein family in caenorhabditis elegans. Eur J Biochem 270:4909-4920

Nash B, Colavita A, Zheng H, Roy PJ, Culotti JG (2000) The forkhead transcription factor UNC-130 is required for the graded spatial expression of the UNC-129 TGF-beta guidance factor in C. elegans. Genes Dev 14:2486-2500

Nishiwaki K (1999) Mutations affecting symmetrical migration of distal tip cells in caenorhabditis elegans. Genetics 152:985-997

Nishiwaki K, Hisamoto N, Matsumoto K (2000) A metalloprotease disintegrin that controls cell migration in caenorhabditis elegans. Science 288:2205-2208

Nishiwaki K, Kubota Y, Chigira Y, Roy SK, Suzuki M, Schvarzstein M, Jigami Y, Hisamoto N, Matsumoto K (2004) An NDPase links ADAM protease glycosylation with organ morphogenesis in C. elegans. Nat Cell Biol 6:31-37

Nishiyama M, Hoshino A, Tsai L, Henley JR, Goshima Y, Tessier-Lavigne M, Poo MM, Hong K (2003) Cyclic AMP/GMP-dependent modulation of Ca2+ channels sets the polarity of nerve growth-cone turning. Nature 423:990-995

Ornitz, DM (2000) FGFs, heparan sulfate and FGFRs: complex interactions essential for development. BioEssays 22:108-112

Pollard TD, Borisy GG (2003) Cellular motility driven by assembly and disassembly of actin filaments. Cell 112:453-465

Praitis V, Casey E, Collar D, Austin J (2001) Creation of low-copy integrated transgenic lines in *caenorhabditis elegans*. Genetics 157:1217-1226

Przyborski SA, Knowles BB, Ackerman SL (1998) Embryonic phenotype of Unc5h3 mutant mice suggests chemorepulsion during the formation of the rostral cerebellar boundary. Development 125:41-50

Raftopoulou M, Hall A (2004) Cell migration: Rho GTPases lead the way. Dev Biol 265:23-32

Reddien PW, Horvitz HR (2000) CED-2/CrkII and CED-10/Rac control phagocytosis and cell migration in caenorhabditis elegans. Nat Cell Biol 2:131-136

Ridley AJ (2001) Rho proteins, PI 3-kinases, and monocyte/macrophage motility. FEBS Lett 498:168-171

Rhiner C, Gysi S, Frohli E, Hengartner MO, Hajnal A (2005) Syndecan regulates cell migration and axon guidance in C. elegans. Development 132:4621-4633

Su M, Merz DC, Killeen MT, Zhou Y, Zheng H, Kramer JM, Hedgecock EM, Culotti JG (2000) Regulation of the UNC-5 netrin receptor initiates the first reorientation of migrating distal tip cells in caenorhabditis elegans. Development 127:585-594

Suzuki Y, Yandell MD, Roy PJ, Krishna S, Savage-Dunn C, Ross RM, Padgett RW, Wood WB (1999) A BMP homolog acts as a dose-dependent regulator of body size and male tail patterning in caenorhabditis elegans. Development 126:241-250

Tamagnone L, Comoglio PM (2000) Signalling by semaphorin receptors: Cell guidance and beyond. Trends Cell Biol 10:377-383

Tanikawa C, Matsuda K, Fukuda S, Nakamura Y, Arakawa H (2003) p53RDL1 regulates p53-dependent apoptosis. Nat Cell Biol 5:216-223

Tessier-Lavigne M, Goodman CS (1996) The molecular biology of axon guidance. Science 274:1123-1133

Thiebault K, Mazelin L, Pays L, Llambi F, Joly MO, Scoazec JY, Saurin JC, Romeo G, Mehlen P (2003) The netrin-1 receptors UNC5H are putative tumor suppressors controlling cell death commitment. Proc Natl Acad Sci U S A 100:4173-4178

Timmons L, Fire A (1998) Specific interference by ingested dsRNA. Nature 395:854

Tong J, Killeen M, Steven R, Binns KL, Culotti J, Pawson T (2001) Netrin stimulates tyrosine phosphorylation of the UNC-5 family of netrin receptors and induces Shp2 binding to the RCM cytodomain. J Biol Chem 276:40917-40925

Vu TH, Werb Z (2000) Matrix metalloproteinases: Effectors of development and normal physiology. Genes Dev 14:2123-2133

Wadsworth WG, Bhatt H, Hedgecock EM (1996) Neuroglia and pioneer neurons express UNC-6 to provide global and local netrin cues for guiding migrations in C. elegans. Neuron 16:35-46

Whangbo J, Kenyon C (1999) A wnt signaling system that specifies two patterns of cell migration in C. elegans. Mol Cell 4:851-858

Wills Z, Marr L, Zinn K, Goodman CS, Van Vactor D (1999) Profilin and the abl tyrosine kinase are required for motor axon outgrowth in the drosophila embryo. Neuron 22:291-299

Wu YC, Horvitz HR (1998) C. elegans phagocytosis and cell-migration protein CED-5 is similar to human DOCK180. Nature 392:501-504

Yu DH, Qu CK, Henegariu O, Lu X, Feng GS (1998) Protein-tyrosine phosphatase shp-2 regulates cell spreading, migration, and focal adhesion. J Biol Chem 273:21125-21131

Yu TW, Hao JC, Lim W, Tessier-Lavigne M, Bargmann CI (2002) Shared receptors in axon guidance: SAX-3/Robo signals via UNC-34/Enabled and a netrin-independent UNC-40/DCC function. Nat Neurosci 5:1147-1154