UNIVERSITY OF CALIFORNIA, SAN DIEGO

Characterization of two Dscam orthologues in *Hirudo medicinalis*

A thesis submitted in partial satisfaction of the
requirements for the degree Master of Science

in

Biology

by

Eunice Seunga Kym

Committee in charge:

Professor Eduardo Macagno, Chair
Professor Jayant Ghiara
Professor Yimin Zou

2011

The Thesis of Eunice Seunga Kym is approved and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California, San Diego

2011

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Eduardo Macagno for the opportunity to perform this research. I am grateful for his unwavering guidance and support throughout the program. I would also like to thank project scientist, Dr. Michael Baker, for patiently teaching me to perform these techniques and providing me with expert advice. Lastly, I would like to thank bioinformatics graduate student Allan Wu for tirelessly helping me sequence the two genes.

ABSTRACT OF THE THESIS

Characterization of two Dscam orthologues in *Hirudo medicinalis*

by

Eunice Seunga Kym

Master of Science in Biology

University of California, San Diego, 2011

Professor Eduardo Macagno, Chair

Dscam is a transmembrane protein of the Ig superfamily that has been shown to play critical roles in the developing nervous system. Its known functions include isoneuronal dendritic and axonal self-avoidance, heteroneuronal avoidance in a process called "tiling", heteroneuronal synapse formation, and axon guidance. The importance of Dscam is further demonstrated by its high degree of sequence, structural and functional conservation in organisms across several phyla, including chordates, mollusks, platyhelminthes and

arthropods. In arthropods, the Dscam gene is alternatively spliced and can create tens of thousands of isoforms, which are used for neuronal identity and may also serve a novel form of adaptive immune function. This ability to create multiple isoforms is not seen in chordates and mollusks. Here, we report on two Dscam orthologues that we have found in the medicinal leech (*Hirudo medicinalis*), which belongs to the annelid phylum. Both Dscam's are expressed in nervous tissue, indicating a conservation of neural function, but they do not appear to be alternatively spliced, thus placing the annelids with the chordates and mollusks in terms of Dscam structure.

INTRODUCTION

As the nervous system develops, there exist several hurdles that must be resolved in shaping neurons to optimize their central roles in the communication of information. Dendrites of the same neuron must fan out to increase the surface area from which they receive signals, dendrites of neighboring neurons must recognize whether they should overlap or avoid each other, axons of individual neurons must identify where to bifurcate in order to reach different targets, and axons of neighboring neurons must identify where and with whom to fasciculate in order to follow specific paths. Axons need to project into the correct locations. One protein that provides several of these functions across species is Down Syndrome Cell Adhesion Molecule (DSCAM in mammals and Dscam in insects; here we will use Dscam for the leech proteins).

Dscam is a member of the immunoglobulin superfamily and is expressed in developing and adult nervous tissue of several organisms (Shmucker et al. 2000, Agarwala et al. 2001a, Brites et al. 2008, Chou et al. 2011, Li et al. 2009, Fusaoka et al. 2006). As a transmembrane receptor, it consists of an ectodomain, a transmembrane domain, and an endodomain. In general, the ectodomain contains 10 immunoglobulin (Ig) domains and 6 fibronectin type III (FNIII) domains. Dscam in *D.melanogaster* undergoes alternative splicing and is capable of creating 38,016 distinct ectodomains. The N terminal of Ig2, the N terminal of Ig3, and all of Ig7 are encoded by several alternative exons (Schmucker et al. 2000). The genomic encoding of alternative exons and the expression of multiple isoforms has been demonstrated in other arthropods as well, including insects and crustaceans (Dong et al. 2006, Gravely et al. 2004, Brites et al. 2008, Watthanasurorort et al. 2011, Chou et al. 2011). The capacity to create multiple isoforms has not been seen in mammalian DSCAM's or Dscam in the California sea slug, *Aplysia californica* (ApDscam), though vertebrates can encode an

isoform missing it's transmembrane domain through exon skipping (Yamakawa et al. 1998, Agarwala et al. 2000, Agarwala et al. 2001, Li et al. 2009). The ectodomain of Dscam is capable of binding another Dscam ectodomain *in trans*. In systems with alternative splicing, only Dscam's that are the same isoform will bind to each other (Wojtowicz et al. 2004). In mammals, DSCAM or DSCAML1 will form paralogue-specific homophilic but not heterophilic interactions (Agarwala et al. 2000, Agarwala et al. 2001, Yamagata and Sanes 2008).

Dscam plays a critical role in dendritic and axonal self-avoidance in *D.melanogaster*. Dendritic self-avoidance has been demonstrated in the arbors of the da sensory neurons and in olfactory projection neurons. In these systems, dendrites of the same neuron have a planar distribution and display isoneuronal self-avoidance and do not touch or cross over each other. However, dendrites of neurons with Dscam mutations show extensive self-crossing, bundling, and have a tangled appearance (Greuber et al. 2002, Zhu et al. 2006, Matthews et al. 2007, Hughes et al. 2007, Soba et al. 2007). In addition to dendritic self-avoidance, Dscam is implicated in axonal self-avoidance in *Drosophila*. Kenyon cells in the mushroom body, a brain structure important in learning and memory, have axons that bifurcate and send sibling axons dorsally and medially (Heisenberg et al. 1985, Mercer et al. 1983). Axons of Dscam mutant neurons bifurcate but the sibling axons do not segregate and both go medially or dorsally (Wang et al. 2002).

Dscam also plays a role in a process known as "tiling", when the processes of neurons of the same type or similar function overlap minimally, resulting in full coverage by the innervation of a specific area. Dscam2 plays a role in axonal tiling in *D.melanogaster* visual system while DSCAML and DSCAML1 mediate tiling and isoneuronal self-avoidance in the mouse retina (Millard et al. 2007, Fuerst et al. 2008, Fuerst et al. 2009). DSCAM is expressed

in dopaminergic and bNOS positive amacrine cells that normally have evenly spaced cell bodies and neurite arbors that do not overlap the processes of cells of the same type. DSCAM mutations lead to heteroneuronal fasciculation of neurite arbors and disordered spacing of cell bodies (Fuerst et al. 2008). This phenotype is also seen in DSCAML1 positive cells that are found in the rod circuit (Fuerst et al. 2009).

The previous examples of dendritic and axonal self-avoidance and tiling support the putative repulsive functions of Dscam. Dendrites or axons that express the same Dscam isoform or a paralogue repel after binding. Demonstrating the complexity of one molecule, Dscam also plays roles that promote cell adhesion upon homophilic binding. In the chick retina, interneurons must synapse on retinal ganglion cells (RGCs) in the appropriate inner plexiform layer sublaminae. DSCAM and DSCAML1 are expressed in distinct subsets of interneurons and RGCs in the sublaminae that synapse on each other. Introducing DSCAM or DSCAML1 into an interneuron reroutes the neuron to synapse in the sublaminae expressing that paralogue (Yamagata and Sanes 2008). In *Aplysia*, ApDscam is expressed in the monosynaptic sensory neuron and it's cognate L7 motor neuron but not a mismatched L11 motor neuron. ApDscam is thought to stabilize the formation of new synapses for long-term facilitation through these interactions (Li et al. 2009).

In addition to neuronal adhesion and repulsion, Dscam also plays a critical role in axon guidance and axonal targeting. In *Drosophila*, Dscam mutant neurons in Bolwig's nerve fail to meet the correct target in the brain and olfactory receptor neurons fail to innervate the appropriate glomeruli in the antennal lobe (Shmucker et al. 2000, Hummel et al. 2003). In mammals, DSCAM has been found to participate in commissural axon guidance by acting as a netrin-1 receptor (Ly et al. 2008, Liu et al. 2009).

The importance of Dscam in circuit formation and neuronal development make it an exciting molecule to study in the medicinal leech (*H.medicinalis*). This is interesting from an evolutionary perspective because neuronal functions of Dscam are conserved across four phyla: arthropods, platyhelminthes, mollusks, and chordates. While arthropod Dscam's are highly spliced and have been shown to play roles in the immune system on top of neuronal development (Watson et al. 2005, Dong et al. 2006, Chou et al. 2009, Smith et al. 2011, Watthanasurorot et al. 2011), mollusk and chordate Dscam's are not known to be highly spliced and an immune function has not been identified. In this thesis we explored Dscam expression and function in Annelids. Annelids are also the first organisms where neuronal self avoidance was described (Kramer and Stent 1985, Kramer et al. 1985). We have identified two Dscam orthologues in the medicinal leech: HmDscam1 and HmDscam2. Both HmDscam's appear to be expressed in nervous tissue but, like mollusks and chordates, do not appear to be alternatively spliced.

MATERIALS AND METHODS

Molecular Cloning of HmDscam1

Dscam protein sequences from several organisms including *Tribolium castaneum*, *Drosophila melanogaster*, *Aplysia californica*, *Mus musculus*, *Anopheles gambiae*, *Helobdella robusta*, *Apis mellifera*, and *Litopenaeus vannamei* were aligned using the web based program CLUSTAL W. Areas with amino acid conservation for at least five consecutive amino acids were identified. Genomic contigs containing nucleotides that coded for the amino acid sequence was found using a translated genome search. These contigs were put into the blastx program on the National Center for Biotechnology Information (NCBI) website. Contigs that brought up matches for Dscam in any organism were identified. Primers were designed from these contigs using the program Fast PCR. PCR was performed with the forward degenerate primer and a reverse primer specific to the 5' end of the known HmDscam sequence. The primer for the most 5' end of the known HmDscam1 sequence was designed from genomic data.

PCR products were separated by size and visualized via gel electrophoresis using a 2% agarose gel containing ethidium bromide. PCR reactions that generated distinct bands were isolated using the Promega Wizard Gel Clean up Kit. PCR products were cloned into the vector pGEM using the pGEM T- Easy Vector kit (Promega). Vectors were transfected into Max Efficiency DH5α *E. coli* cells (Invitrogen). Colonies containing the pGEM vector and insert were identified by blue/white selection. The plasmid was purified using a Mini Prep Kit (Qiagen). Sequencing was performed by the company Eton Bioscience Inc.

Molecular Cloning of HmDscam2

Exons predicted from the Glimmer software off of a genomic contig showed high homology to Dscam using NCBI blastx were identified. Primers were designed from these predicted exons. Five primer sets (HmDscam2 1-5) were designed to cover regions from the predicted 5' end of the expressed gene to the 3' end. PCR was performed using the five primer sets, bands of expected size were gel purified, and cloned into pGEM T –Easy. Plasmids were purified and sequenced. Primers designed from genomic data amplified regions of HmDscam2 aligned to amino acids 560-828 and 1058-1270 of *A.californica* Dscam. A third set of primers was designed to amplify the region that was predicted to align to amino acids 828-1057 of *Aplysia.* 2454 continuous base pairs of the middle of the HmDscam2 sequence have been cloned and sequenced.

NCBI blast-n alignment software was used to align and stitch together clones that blasted to Dscam. The DNA sequence was translated using expasy translate to ensure there was an appropriate open reading frame.

***In Situ* Hybridization**

HmDscam1 and HmDscam2 specific primers were designed with T7 Promoter sequences on the 5' end of the reverse primer. Target areas of the genes ranging from 230 bp – 1100 bp were identified to synthesize a riboprobe. Target areas were amplified with the gene specific primers through PCR. Appropriate PCR products were isolated via gel purification.

After gel purification, PCR products were further purified by Phenol:Chloroform extraction. PCR products were mixed with an equal amount of phenol (Invitrogen), vortexed thoroughly, and centrifuged at room temperature at maximum speed (>17,000 rpm) for 10 minutes. The aqueous layer was extracted and mixed with an equal amount of chloroform, vortexed thoroughly, and centrifuged at room temperature (RT) at maximum speed for 10

minutes. The aqueous layer was extracted and precipitated with 1/10th the volume Sodium Acetate and 2.5x the volume 200 proof ethanol. 2 ul of glycoblue was added to visualize the pellet. The solution was placed at -80^{o} C for at least 30 minutes and centrifuged at maximum speed at 4^{o}C for 30 minutes. The supernatant was discarded and the pellet was washed with 70% EtOH in DEPC. The solution was vortexed thoroughly and centrifuged at maximum speed for 5-10 minutes. The supernatant was discarded and the pellet was air dried for 6-7 minutes. The pellet was resuspended in DEPC. A spectrophotometer was used to confirm DNA was free of contaminants.

The DNA template was transcribed into single stranded RNA using Fermentas T7 RNA Polymerase and digoxigenin labeled ribonucleotides. The reaction was left for 2 hours at 37^{o}C. The reaction was stopped with 4 ul of LiCl and 75 ul of 200 proof EtOH. 2 ul of glycoblue was added to the reaction to visualize the pellet. The reaction was placed at -80^{o}C for 30 minutes. The solution was then centrifuged for 30 minutes at 4^{o}C. The pellet was washed with 70% EtOH in DEPC and air dried for 6-7 minutes. The pellet was resuspended in 10 ul of DEPC and 90 ul of hybridization buffer. Strict RNAse free techniques were maintained throughout riboprobe synthesis.

Embryos from an in house breeding colony or supplied by the Kristan Lab were anesthetized in 8% EtOH in Helobdella pond water. Embryos were dissected, pinned onto a Sylgard dish, and fixed in 4% Paraformaldehyde in PBS for 30 minutes. Embryos were washed with PBT (50 ml 1x PBS + 50 ul of Tween), permeabilized with Pronase E for 3-5 minutes (older embryos were Pronase E treated for a longer period of time), and then washed with PBT for 25 more minutes. Embryos were washed with SSCHAPS for 30 minutes. Embryos were then unpinned from the Sylgard dish and equilibrated with hybridization buffer (HB) for 5 minutes at RT in an epitube. Embryos were kept in HB at 55^{o}C for 5 – 60 minutes.

The riboprobe was diluted 1:20 or 1:10 with HB and denatured at 95^{o}C for 6 minutes. After denaturing, the riboprobe was immediately kept in ice until use. The HB solution the embryos were in was replaced with 100 ul of riboprobe. Probe hybridization took place for 1 or 2 nights at 55^{o}C. RNAse free techniques were maintained until after the probe was hybridized.

Embryos were equilibrated back into SSCHAPS and washed for a total of 3 hours with PBT. Embryos were pinned back onto Sylgard dishes and blocked with a solution of 5% FBS in PBT that had been denatured at 55^{o}C for >1 hour. Embryos were blocked in the 5% FBS solution for 1 hour. A α-Digoxigenin antibody from Roche Diagnostics was diluted 1:5000 in PBT and applied to the embryos at 4^{o}C overnight.

Embryos were washed for a total of 3.5 hours with PBT. A NET/BCIP tablet was dissolved in 10 ml of water. The NET/BCIP solution was applied to the embryos and embryos were developed in the dark until an appropriate color.

RNA Isolation from Embryos and Cells and cDNA Synthesis

3-4 whole embryos were suspended in 1 ml of Trizol (Invitrogen). The embryos in trizol were vortexed and put on a 55^{o}C heating block to facilitate homogenization of tissues. Phase separation was performed using .2 ml of chloroform (Invitrogen), vortexing for 15 seconds, incubating at room temperature for 3 minutes, and centrifuging at 12,000 x g for 15 minutes at 2^{o}-8^{o}C. The aqueous was removed and RNA was precipitated using .5 ml of isopropanol (Invitrogen), incubating at room temperature for 10 minutes, and centrifuging 12,000 x g for 15 minutes at 2^{o}-8^{o}C. The pellet was washed using .5 ml of 75% EtOh in DEPC, centrifuging at >12,000 x g for 5 minutes at room temperature, withdrawing the supernatant, and repeating once. The pellet was left to air dry for 5 minutes at room temperature and resuspended in 10 ul of DEPC. RNA was extracted from single cells using

Aboslutely RNA (Stratagene). cDNA was synthesized using the cDNA SuperScript® III cDNA Synthesis kit from Invitrogen.

Quantitative PCR

Primers were designed to amplify regions of HmDscam1 and RS60 that were 150-200 nucleotides long. The qPCR reaction was mixed in the following ratios: 10 ul SYBR Green (BioPioneer), 8 ul nuclease free water (BioPioneer), 1 ul cDNA template, 1 ul of 2mM primer. HmDscam1 expression was normalized as a ratio to RS60 expression.

RNA Interference

A target area of 200 -250 bp of the HmDscam1 and HmDscam2 gene was selected to create a dsRNA cocktail. Specific primers with T7 promoter sequences on the 5' end of each primer were designed and PCR was used to amplify the target sequence. PCR products were gel purified, phenol:chloroform purified, and ethanol precipitated. Double stranded RNA (dsRNA) was made using Fermentas T7 RNA Polymerase. The reaction was treated with Invitrogen Amplification Grade Dnase I for 15 min. at 37^{o}C. The reaction was stopped using EDTA. dsRNA was purified using phenol:chloroform extraction and ethanol precipitation. dsRNA was treated with Rnase III from an Ambion siRNA Cocktail Kit, or with Shortcut RNAse III (NEB). The Rnase III reaction was permitted to go for 1 hour at 37^{o}C. siRNA was purified with a phenol:chloroform extraction and ethanol precipitation. RNAse free techniques were maintained.

The siRNA cocktail was mixed with fast green dye. The siRNA solution was loaded into a glass capillary tube that had been pulled to a fine point. Embryos were anesthetized in 8% EtOH in *Helobdella* pond water. The tip of the capillary tube was gently broken with forceps and inserted under the germinal plate of the embryo. The siRNA cocktail was injected under the germinal plate. For HmDscam1 knockdown, embryos were injected when they were

7 days old with.4-.8 ug of siRNA/embryo and dissected when they were 10 days old. For HmDscam2 knockdown, embryos were injected with .04-.05 ug of siRNA when they were 6 days old and dissected when they were 8 days old.

Immunohistochemistry

Embryos were anesthetized in 8% EtOH in *Helobdella* pond water. Embryos were dissected and fixed in 4% PFA for 1 hour. Embryos were washed with a solution of .5% Triton X in PBS with Sodium Azide for 2 hours at room temperature on a low speed shaker. Embryos were blocked in 2% FBS in the wash solution for 1 hour at room temperature. The primary Ab (Lan 3-2) was diluted 1:50 in the blocking solution and applied to the embryos overnight at 4^{o}C. Embryos were washed for a total of 3.5 hours at RT. Embryos were blocked for 5 minutes. The secondary Ab (Donkey α-Mouse) was added to the embryos in a 1:500 dilution overnight at 4^{o}C. Embryos were washed at RT for 3.5 hours, mounted in glycerol, and visualized with fluorescent confocal microscopy.

P cell dye fill

The P_D and P_V cells were injected with a 2% solution of Alexa 488 Dextran (10,000 MW; Invitrogen) dissolved in diH_2O. The dye was passed from the tip of a sharp microelectrode into the cell with pulses of positive current (.5-2 nA) at 500 ms and at 1 Hz. Upon visualization of the neuronal arbor to confirm cellular identity, the preparation was fixed in 4% paraformaldehyde for 30 minutes.

RESULTS

HmDscam1 amd HmDscam2 Sequence and Structure

A schematic diagram of the distribution of immunoglobulin (Ig) and fibronectin type III (FNIII) domains in the extracellular portion of the canonical Dscam protein is shown in Figure1A. Including the transmembrane and intracellular domains, this structure is encoded on average by a full-length sequence of approximately 6000 nucleotides. In comparison, the currently best estimate of the HmDscam1 full-length sequence is that it is 5418 bp long, with a predicted 10 Ig domains but only 5 of the 6 expected FNIII domains. HmDscam1 appears to be missing the highly conserved FNIII #5 domain (Figure 1B). At present, we only have obtained a partial sequence of the HmDscam2 coding region, a region 3362 nucleotides long in the middle of the transcript that encodes only 7 Ig domains and 3 adjacent FNIII domains, (Figure 1C). Both HmDscam1 and HmDscam2 expressed sequences have premature stop codons which are not found in the genome, and further work remains to correctly sequence through this area. The nucleotide and protein sequences, shown in Appendix 1, will be deposited in GenBank in the near future.

Figure 1. Schematic diagrams of Dscam. Canonical model of Dscam (A), and the known domains of (B) HmDscam1 and (C) HmDscam2.

HmDscam1 is expressed in the PNS of the embryo and the CNS of the adult

HmDscam1 expression in the embryo was characterized by using *in situ* hybridization (ISH) on whole mounted E10-E12 embryos (between 10-12 days of age, as judged by morphological criteria; Fernandez and Stent 1982). Expression was observed in clusters of what are judged to be peripheral sensory cells throughout the body wall of the specimens, from the head to the tail, based on peripheral sensory cells described in other papers (Huang et al. 1998, Johansen et al. 1992, DeRosa et al.1981). At these stages, the germinal plate is still expanding but has not closed along the dorsal midline, and much of the cryptolarva is clearly visible on the dorsal areas of the head and extending dorsally past the region of the posterior sucker. As the embryo is opened along the dorsal midline and flattened, the ventral midline in the preparation becomes the antero-posterior axis of symmetry (see schematic, Figure 2A).

In the head region, expression of HmDscam1 was observed to be particularly strong in numerous small clusters of cells distributed throughout the mouth area but also all the way to the margin (Figure 2B). In the 21 segments of the body, expression was observed in every annulus, in rows of 12-24 small clusters of cells located mostly along the middle of the annulus but with some variation (Figure 2C). In the tail, the clusters covered the whole surface in an irregular but evenly spaced distribution (Figure 2D). A close up of a few labeled clusters in the tail sucker of an E10 embryo shows ~2-3 cells with perinuclear staining per cluster (Figure 2E). No expression of HmDscam1 was observed in the embryonic CNS in these preparations, nor was any detected using qPCR on a cDNA template from E16 central nervous system ganglia. However, HmDscam1 was detected in cDNA of excised adult Retzius cell somata using qPCR (Figure 3), so it is possible that expression is more widely present in adult ganglia. Given their locations and distributions, it is highly likely that the cells that express HmDscam1 are a subset of the peripheral sensory neurons in the epidermis.

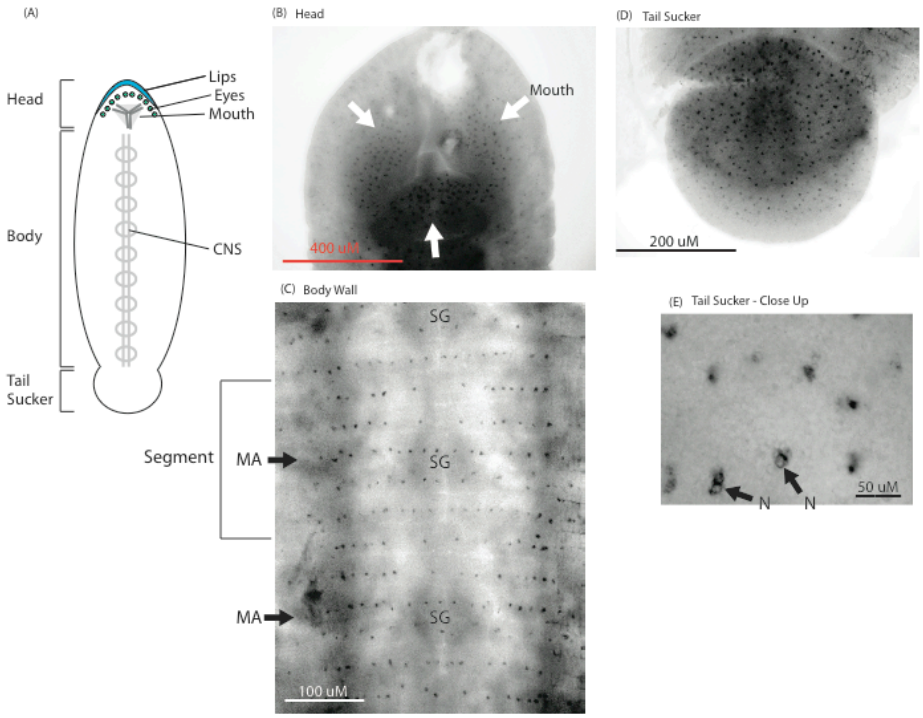

Figure 2. HmDscam1 expression in the E12 leech embryo. (A) Schematic diagram showing leech embryo anatomy. (B) HmDscam1 is expressed in peripheral sensory cells in the mouth region of the head (white arrows), (C) in transverse rows in every annulus of the 21 body segments,, and (D) groups of cells throughout the tail sucker. (E) Close up of expression in the tail sucker of an E10 embryo shows that Dscam1 is expressed by putative sensory neurons in small clusters of 2-3 cells. SG, segmental ganglion, out of focus; MA, middle annulus; bracket denotes a segment (5 annuli); N, nucleus.

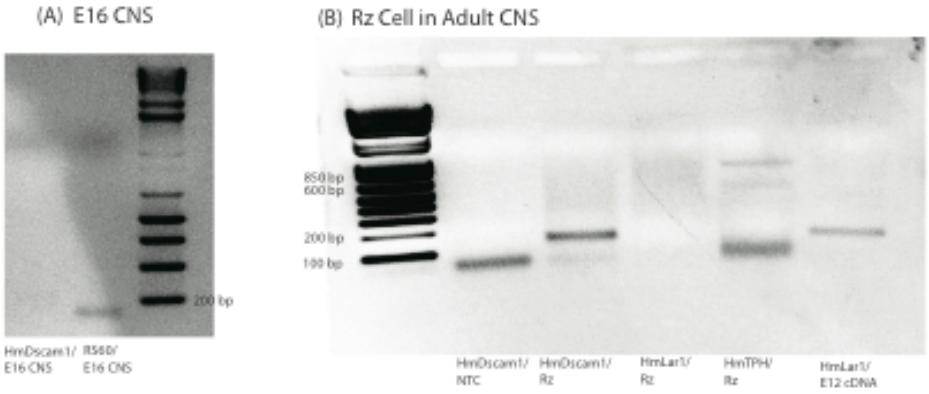

Figure 3 HmDscam1 is not expressed in the embryonic CNS but is expressed in Rz cells of the Adult CNS. Gels from qPCR that show HmDscam1 is not expressed in the (A) E16 CNS (n=3), but is expressed in (B) the Rz (n=5) cell of adult ganglia. HmDscam1/NTC is a negative control using nuclease free water instead of Rz cell cDNA, HmLar1/Rz is a negative control and HmTPH/Rz is a positive control to verify the cell is the Rz cell. HmLar1/E12 cDNA is to show that the HmLar1 primers work.

Whole embryo HmDscam1 RNA interference does not affect the morphology of Lan 3-2 positive axons overtly

Peripheral sensory neurons send afferent projections to the CNS along stereotypic pathways, by receiving guidance cues from the CNS and traveling along nerves established by pioneer P cells in the CNS (Jellies et al. 1994, Jellies et al. 1995 Kuwada 1985, Gan et al. 1995). Peripheral sensory axons fasciculate, travel together to the CNS, and after reaching the neuropil bifurcate into anterior and posterior branches that assort into different, probably modality-specific (e.g., mechanosensory, chemosensory and photosensory) fascicles (Peinado et al. 1987, Hockfield et al. 1983, Johansen et al. 1992, Briggs et al. 1983, Phillips and Friesen 1982). I presume that these afferent bundles include the projections of the HmDscam1-positive peripheral cells, though lacking an antibody, I am currently unable to test this hypothesis directly. Nevertheless, we do have a general marker for *Hirudo* afferent projections, the monoclonal antibody Lan3-2 generated some time ago by Zipser and

colleagues (Peinado et al. 1987, Zipser et al. 1981).

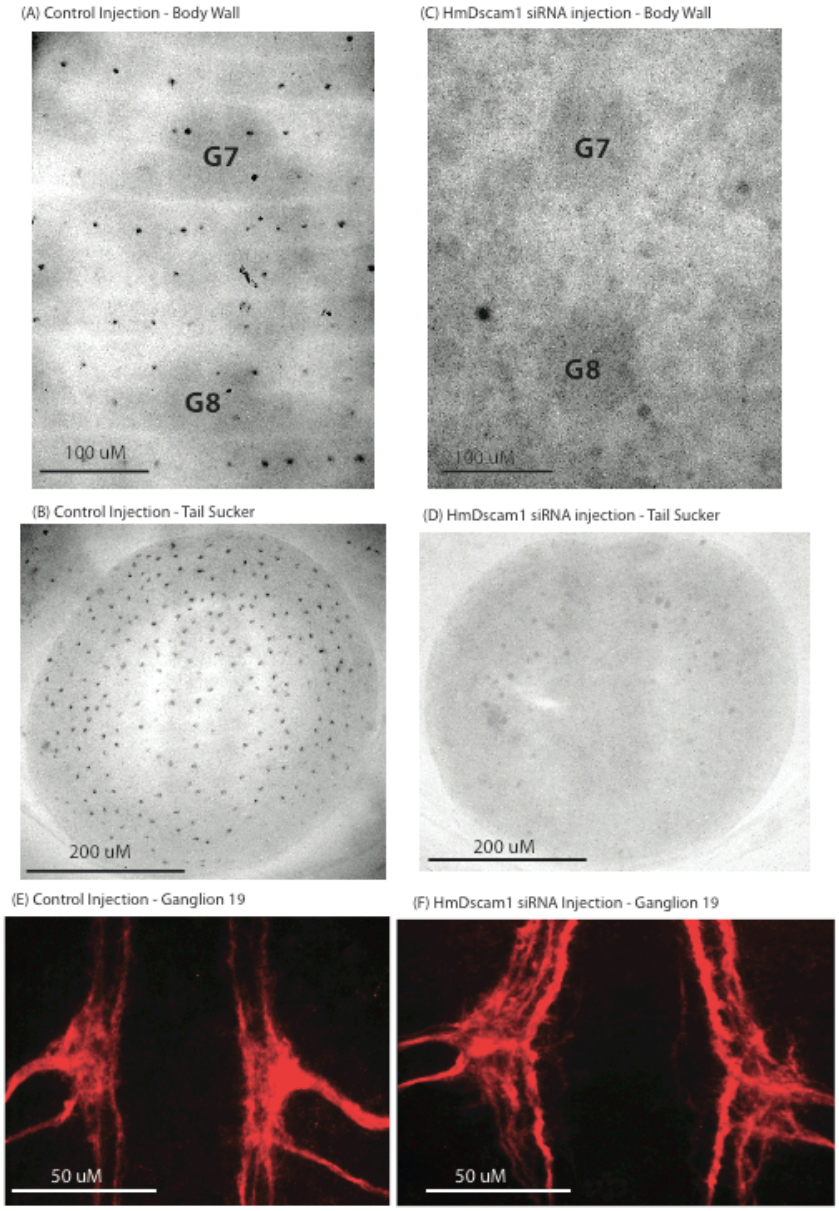

Figure 4. HmDscam1 knockdown does not affect morphology of Lan 3-2 positive axons. (A,B) HmDscam1 *in situ* hybridization from control injected embryos, showing the body wall over Ganglia 7 and 8 (C,D) HmDscam1 *in situ* hybridization from HmDscam1 siRNA injected embryos shows decreased gene expression. (E,F) Antibody staining against Lan3-2 in a control and siRNA embryo.

To begin to test whether HmDscam1 might be involved in the guidance of these projections or their behavior after they entered the CNS, I injected an HmDscam1 siRNA cocktail injected under the germinal plate of the embryo at E7 to reduce HmDscam1 expression. Embryos were dissected and *in situ* hybridization and Lan3-2 antibody staining was performed on them at E10. As shown by *in situ* hybridization, there was decreased HmDscam1 signal from the anterior ganglion 7 and 8 to the posterior tail sucker (Figure 4).

Antibody staining for axons of sensory afferents using the Lan 3-2 antibody did not reveal an obvious phenotype from HmDscam1 knockdown (Figure 4 E,F). However, this result is not totally unexpected, as there are many non-HmDscam1-expressing sensory neurons that are stained by Lan3-2 but are not affected by the RNAi. These other cells might serve as replacement templates for the HmDscam1$^{(-)}$ cells using a redundant signaling pathway, or perhaps the knockdown affects the peripheral arbors rather than the central projections, as similar experiments affect the peripheral arbors of the DA neurons in the fly larva (Matthews et al. 2007, Soba et al. 2007, Hughes et al. 2007). These experiments need to be greatly expanded and markers for these particular sensory neurons need to be developed in order to determine what role HmDscam1 plays in peripheral sensory cell differentiation.

HmDscam2 is expressed in both the CNS and the PNS of the embryo

By contrast with HmDscam1 expression exclusively in the PNS, HmDscam2 is expressed by both peripheral and central neurons. HmDscam2 expression was characterized in embryos between 7-13 days of age using *in situ* hybridization. In the CNS, it is expressed in the mechanosensory cells: the three T cells (T_D, T_V, T_L), the two P cells (P_D, P_V), and one or both of the N cells. There is also faint or strong expression in the macroglial cells of the neuropil, varying by embryo (Figure 5). These cells are identified by their placement, shape,

and size, and expression is consistent between ganglia. Expression in the P cells was confirmed by filling the P cells with dye, observing their arbors to confirm the dye filled cell was a P cell, and performing an *in situ* hybridization to see if the same cell is marked by the riboprobe and fluorescence (Figure 6).

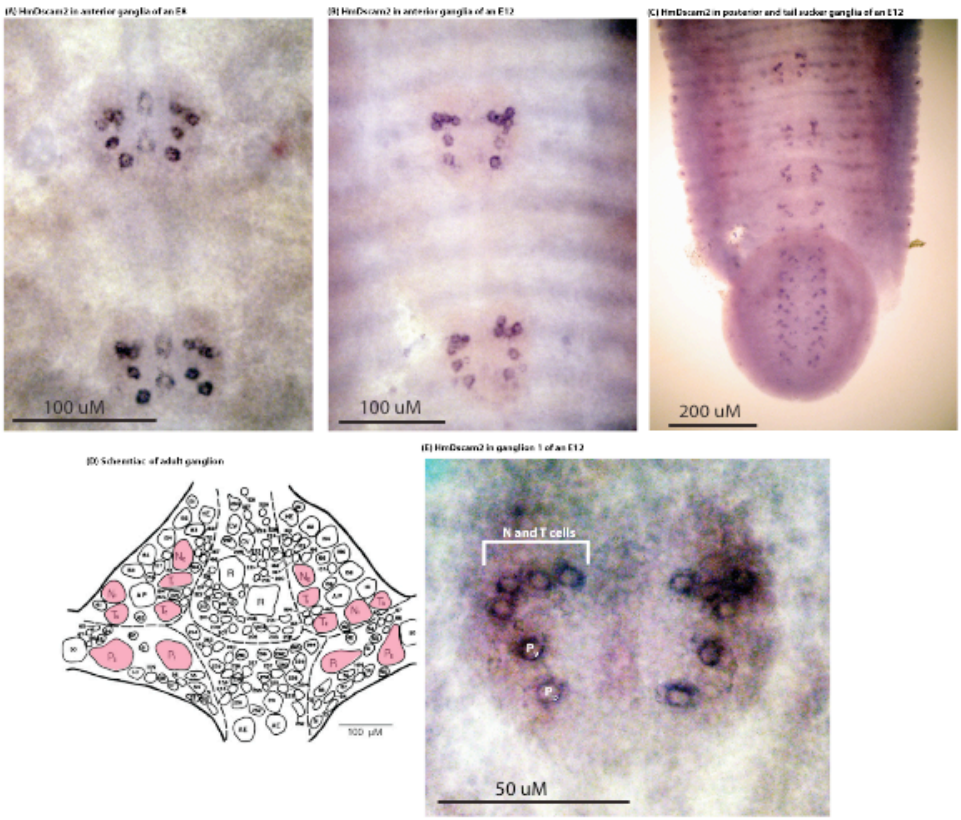

Figure 5. HmDscam2 expression in the CNS of an embryo. (A,B,C) HmDscam2 is expressed in the T, N, and P cells and, at times, the macroglia in the CNS (MG, black arrows); TS, tail sucker; SG, segmental ganglion (D) Schematic diagram of an adult ganglion to show location of identified cells. (E) P_D, P_V, N and T cells labeled on an E12 ganglion, putatively identified by location.

Expression in the neuropil macroglia has not been identified as age dependent, but *in situ* hybridizations or qPCR studies on extracted macroglia in a larger sample of embryos of different ages could elucidate whether expression is developmentally controlled. At E6, posterior ganglia have clear expression of HmDscam2 in only P cells while anterior ganglia

have clear expression of HmDscam2 in T or N cells as well, demonstrating that HmDscam2 is expressed in the P cells before it is expressed in other mechanosensory cells (Figure 7). Posterior ganglia can be up to 2-3 days or 10% behind in development compared to anterior ganglia (Johansen et al. 1992). This is an interesting result because P cells are known to be pioneer cells for other neurons in the CNS sending efferent projections into the periphery (Kuwada 1985, Gan et al. 1995).

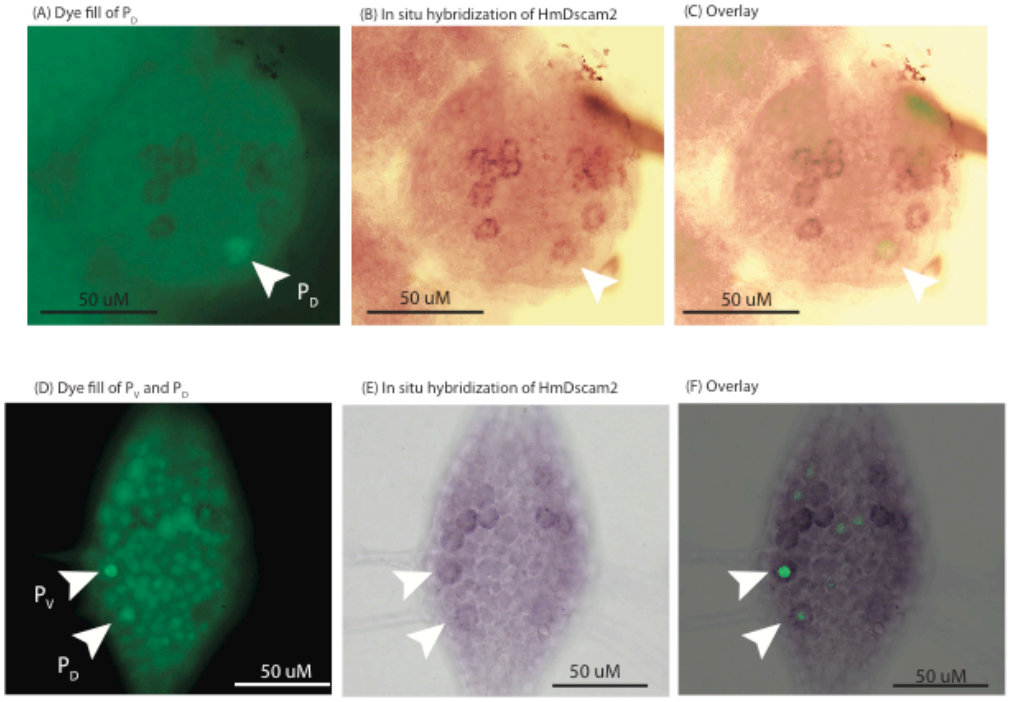

Figure 6. HmDscam2 expression in the P_D and P_V cell is confirmed with double label experiments. Further identification of HmDscam2 positive cells in the embryo by (A,D) Filling a P_D cell (D) or both the P_D and P_V cells with dye, (B,E) *in situ* hybridization of the ganglion, and (C,F) an overlay showing the dye filled cells are HmDscam2 positive.

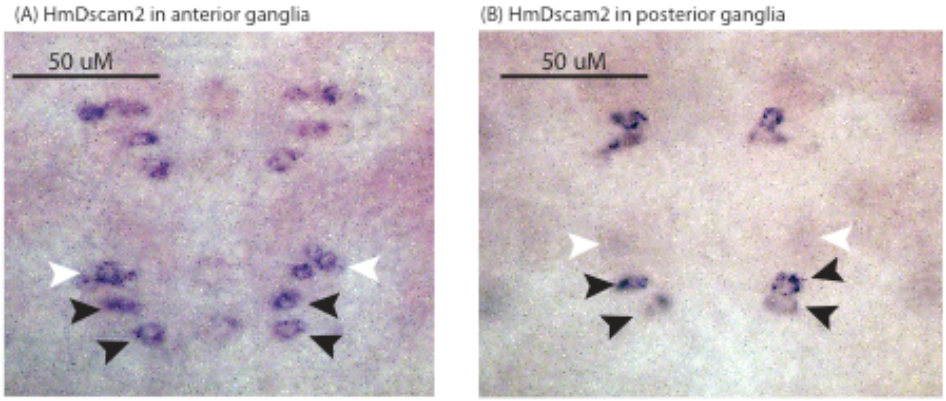

Figure 7. HmDscam2 is first expressed in P cells. HmDscam2 *in situ* hybridization in an E6 showing expression in (A) older anterior ganglia and (B) younger posterior ganglia. HmDscam2 is first expressed by both P cells (black arrows) and then T and N cells (white arrows).

HmDscam2 is also expressed in the peripheral nervous system. Between E6-E8 HmDscam2 expression is observed in large bilateral pairs of clusters along the middle annulus of each segment, which morphologically and by location correspond to the developing Lan3-2 positive sensilla organs previously described at E6 (McGlade-McCulloch et al. 1990) (Figure 8). The criteria used by McGlade-MucCulloch for staging embryos was different than the one we used, however judging by similar morphological features such as a thin tail sucker region and a large, ball shaped, larval mouth, we conclude that the image in Figure 8A of Lan3-2 positive sensilla was taken at an age similar to what refer to as E6. Lan 3-2 positive sensilla in the embryo pictured occur in clusters of 5 while HmDscam2 positive cells appear in clusters of 3 at E6, 5 at E7, and 6 at E8. At E8, extra sensillar peripheral sensory cells express HmDscam2 and occur in clusters of 2-3 cells while the HmDscam2 positive cells in the sensilla begin to appear in clusters up to 5. By E12, HmDscam2 is expressed in many peripheral sensory cells in all 5 annuli, and there are several more HmDscam2 expressing cells in sensillar clusters. In the head region at age E12, HmDscam2 is strongly expressed in the 5

bilateral pairs of eyes and along the anterior lip regions (Figure 9). As with HmDscam1, these peripheral cells are identified as putative sensory neurons based on their placement close to the epidermis, their locations and their clustered appearance.

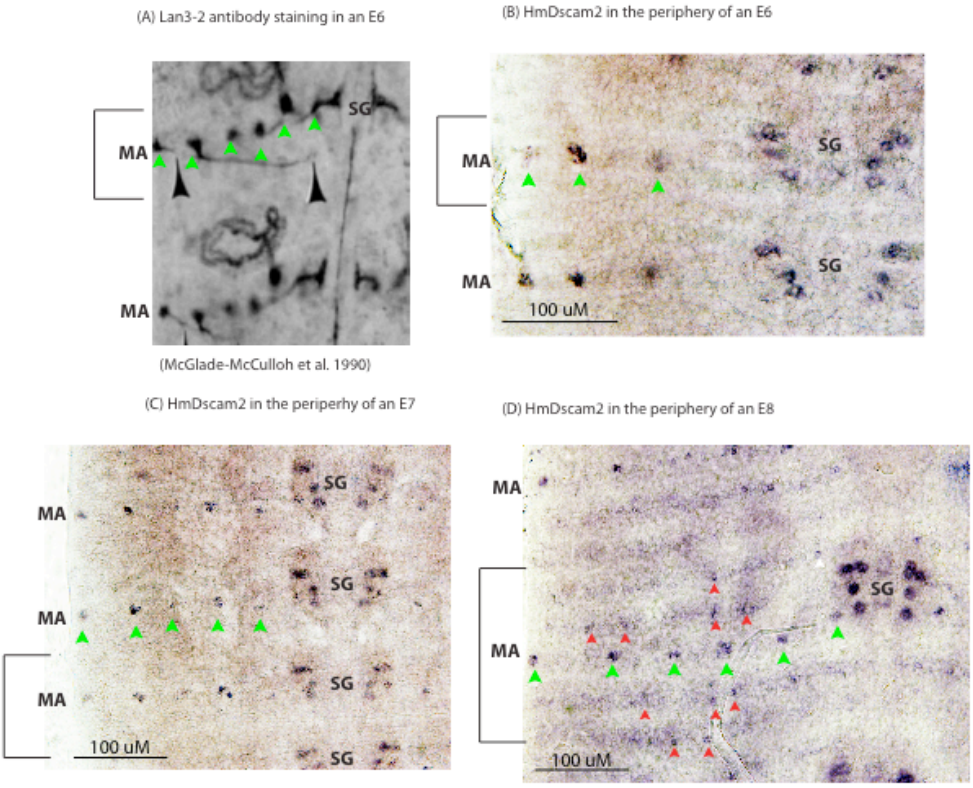

Figure 8. HmDscam2 appears to be expressed in developing sensilla in E6 - E8. (A) Lan 3-2 positive sensilla (green arrows) at a similar embryonic age in morphology and location suggesting HmDscam2 is expressed in the sensillar organ. Projections of Lan3-2 positive sensilla into the CNS (black arrows) (B) At E6 HmDscam2 is expressed in 3 bilateral cell clusters (green arrows) located on the middle annulus (MA) of each segment (SG: segmental ganglion, brackets denote a segment) (C) At E7 HmDscam2 is expressed in 5 bilateral cells clusters and at (D) E8 HmDscam2 is expressed in at least 6 bilateral cell clusters on the middle annulus and several extra sensillar peripheral sensory cells on other annuli (red arrows).

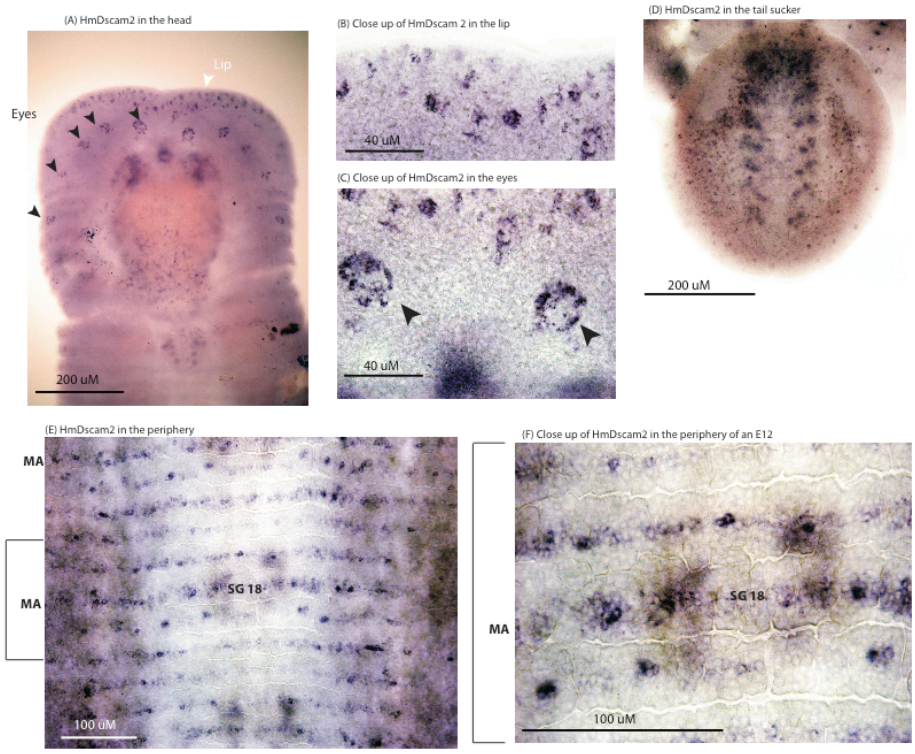

Figure 9. HmDscam2 is expressed in several peripheral sensory cells in an E12. HmDscam2 is expressed in the (A,B) 5 bilateral photoreceptors of the eyes (black arrows) and the (A,C) lip (white arrow) of the head, (D) in some regions of the tail sucker, and (E,F) in peripheral sensory cells located on all five annuli in the body wall MA, middle annulus, SG, segmental ganglion, brackets denote segments.

HmDscam2 RNAi appears to affect clustering of cells in the sensilla

Embryos were injected with an HmDscam2 siRNA cocktail mixed with fast green or nuclease free water mixed with fast green at E6 and dissected at E8. Because HmDscam2 was expressed in peripheral sensory cells including cells in the sensilla, we used the Lan 3-2 antibody to observe a phenotype. (Figure 10). Compared to the control embryo, the HmDscam2 RNAi embryo showed defects in clustering of sensillar cells in the periphery. Instead of appearing as small clusters of 2-4 cells, Lan3-2 positive cells in S1 and S2 appeared separated from each other. There is also increased Lan3-2 staining around the nephridiapore, a phenotype seen in experiments where peripheral sensory afferents can no

longer receive guidance cues from the CNS (Jellies et al. 1995). The effect on cell bodies is reminiscent of a phenotype seen in the retina of Dscam mutants, where cell bodies become irregularly spaced and lose their regular mosaic pattern (Fuerst et al. 2008, Fuerst et al. 2009). This data is very preliminary and needs to be repeated with another dsRNA injection control for a gene not expressed in *H.medicinalis*, or a gene that should have no effects on peripheral sensory cell. Additionally, the Lan 3-2 antibody we used showed much higher background than before, suggesting that this experiment should be repeated with a fresher aliquot. We saw no obvious effects on Lan 3-2 tract fasciculation in peripheral nerves or in the CNS, but this could be due to low visibility due to high background.

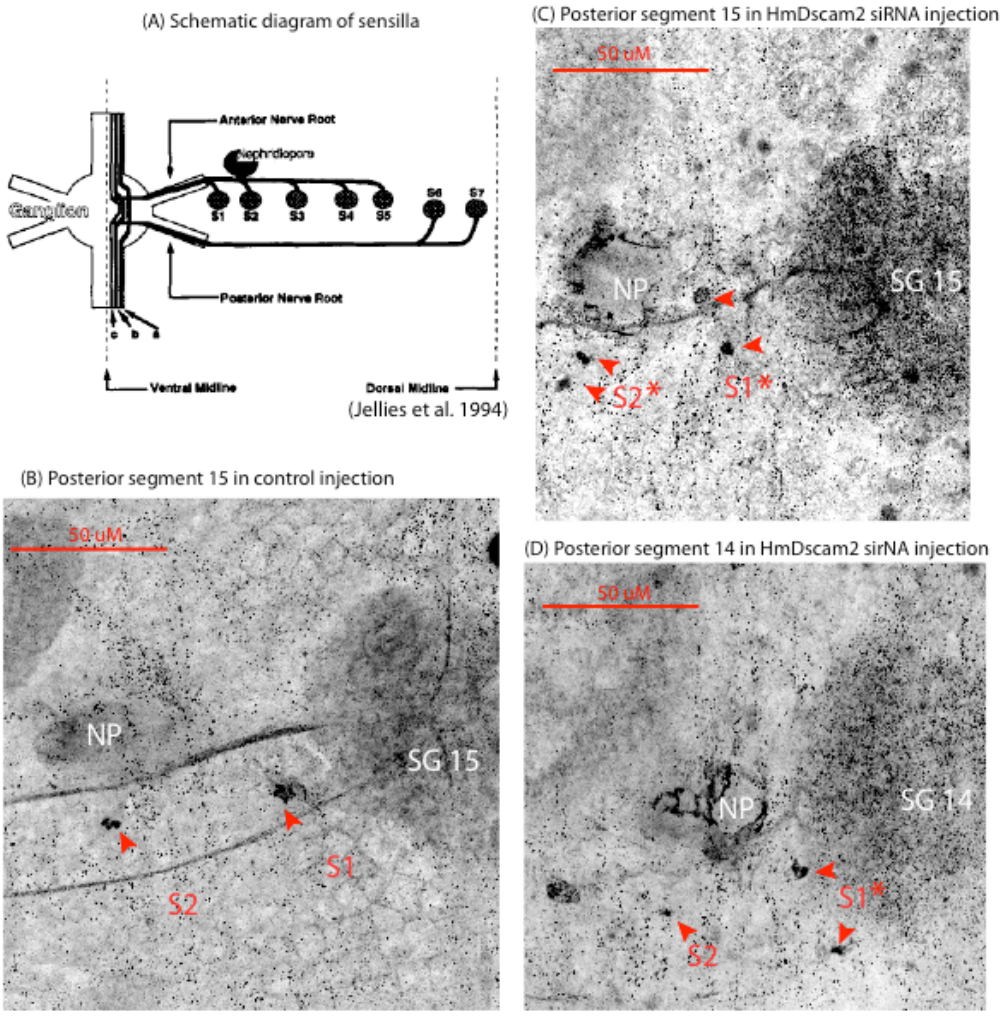

Figure 10. HmDscam2 knockdown affects sensillar cell bodies. (A) Schematic diagram of Lan 3-2 positive sensilla. SG, segmental ganglia; NP, nephridiapore; S(1-7), sensilla 1-7. Lan3-2 antibody staining of S1 and S2 in (B) midbody segment 15 of a control injected embryo shows S1 and S2 as tightly clusters of 2-4 cells (red arrows). Cells in S1 and S2 appear to lose their cohesiveness in (C,D) HmDscam2 RNAi injected embryos (denoted by asterisk) (C) cell bodies of S1 and S2 are separated in midbody segment 15, and (D) cell bodies of S1 are separated in midbody segment 14 , while S2 appears to be smaller than in the control.

Comparison of HmDscam1 and HmDscam2 expression in the periphery

According to a NCBI blast-n alignment, HmDscam1 and HmDscam2 show no significant identity on a nucleotide level. Our *in situ* hybridization results further confirm that HmDscam1 and HmDscam2 are two different genes. HmDscam1 and HmDscam2 are expressed in different areas of the head and tail sucker. HmDscam1 is strongly expressed in the mouth area of the head while HmDscam2 is expressed in the ten eyes and the lips. HmDscam1 is expressed throughout the tail sucker while HmDscam2 in the tail sucker is restricted to certain regions (Figure 9). The peripheral sensory cells that express HmDscam1 and HmDscam2 occur on each of the 5 annuli, but HmDscam2 is expressed in many more sensory cells. HmDscam2 positive peripheral sensory cells may contain the subset of cells that are also HmDscam1 positive. Furthermore, HmDscam1 is not expressed in the embryonic CNS while HmDscam2 is strongly expressed in mechanosensory cells in each ganglia.

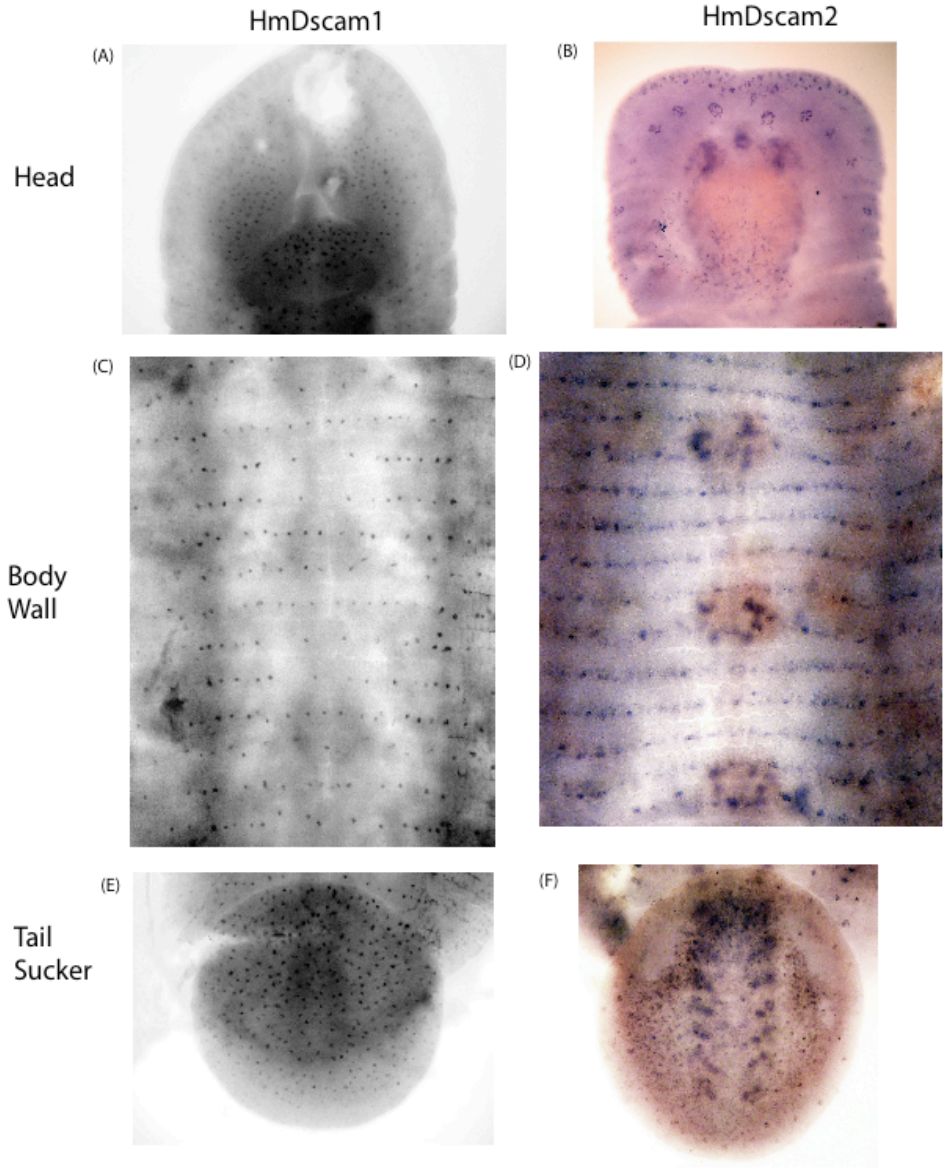

Figure 11. Comparison of HmDscam1 and HmDscam2 expression. (A,C,E) HmDscam1 and (B,D,F) HmDscam2 expression in the peripheral sensory cells of an E12. (A,B) HmDscam1 is expressed in the mouth region of the head while HmDscam2 is expressed in the eyes. (C,D) HmDscam2 is expressed in several peripheral sensory cells while HmDscam1 is expressed in a subset, some of which may or may not be HmDscam2 positive. (E,F) HmDscam1 is expressed throughout the tail sucker while HmDscam2 is expressed in certain parts of the tail sucker.

Evolutionary relationships of HmDscams to other Dscam's

A cladogram generated on CLUSTAL W using the neighbor joining method and excluding gaps shows that the *Hirudo medicinalis* Dscam's are most similar to Dscam from *Helobdella robusta*, another leech, and Dscam in *Dugesia japonica*, a flatworm in the Platyhelminthes phylum. There are three evolutionary groups of DSCAM/Dscam. One encompasses the arthropods, another encompasses chordates and mollusks and the third encompasses the known annelid and Platyhelminthes Dscams (Figure 11).

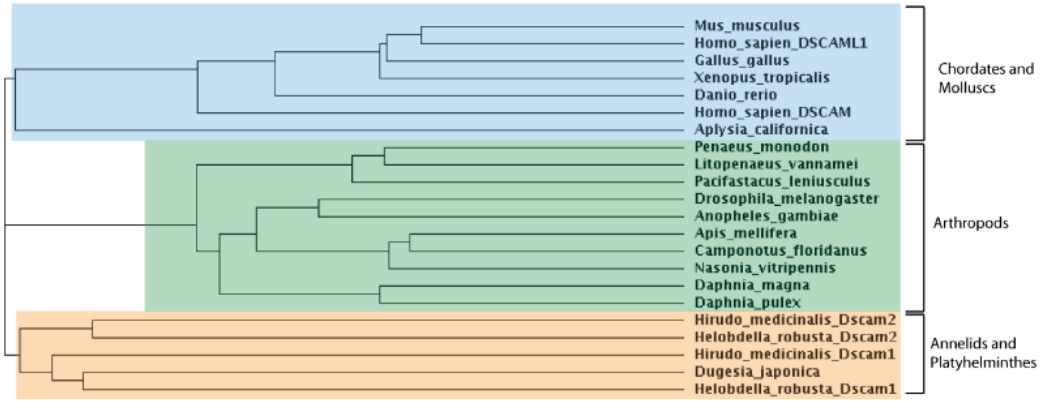

Figure 12. Cladogram of Dscam across species. There are three evolutionary groups of Dscams: one for arthropods, another for chordates and mollusks, and a third for annelids and platyhelminthes.

HmDscams do not appear to be alternatively spliced

Exons and introns were identified by aligning genomic and expressed sequences using NCBI blastn (Figure 12). When the alignment was optimized for "Highly similar sequences (megablast)", the exon encoding the N terminus of Ig8 in HmDscam1 was not identified. Changing optimization to "More dissimilar sequence's (discontiguous megablast)" identified this exon. The problem was due to a 26 nucleotide region of the genomic contig that is located in the middle of the exon but was not found in the expressed sequence. This region is not large enough to be an intron. We modeled our exon-intron figure using the discontinuous megablast

optimization because the N terminus of Ig8 is expressed and this could be a problem with sequencing errors or genomic mis-assembly.

HmDscam1 does not appear to be alternatively spliced in the same way that arthropod Dscam's are alternatively spliced for three reasons. (1) Exons encoding the three spliced Ig domains of HmDscam1 were not present in large tandem arrays of alternative exons and (2) the size of the genomic locus is not large enough to support such tandem arrays: HmDscam1 is 12 kb while *Drosophila* Dscam1 is 61 kb with 12.7 kb devoted to Ig3 tandem repeats (Shmucker 2000). (3) Additionally, Ig3 of HmDscam1 is encoded by one exon and Ig7 is encoded by two exons in HmDscam1. In order to have arthropod like alternative splicing, Ig3 must be encoded by two exons and Ig7 must only be encoded by one exon.

HmDscam2 also does not appear to be alternatively spliced for two reasons. (1) The area of the HmDscam2 genomic locus that codes for Ig1-4 is predicted to be 5 kb while exons coding for Ig2 and Ig3 alone take up 19 kb in the *Drosophila* Dscam1 gene. (2) The sequence of the only known Ig that is alternatively spliced, Ig7, is encoded by two exons instead of one as seen in HmDscam1. However, the genomic and expressed sequences of HmDscam2 are not fully known and there are two identified genomic mis-assemblies in HmDscam2.

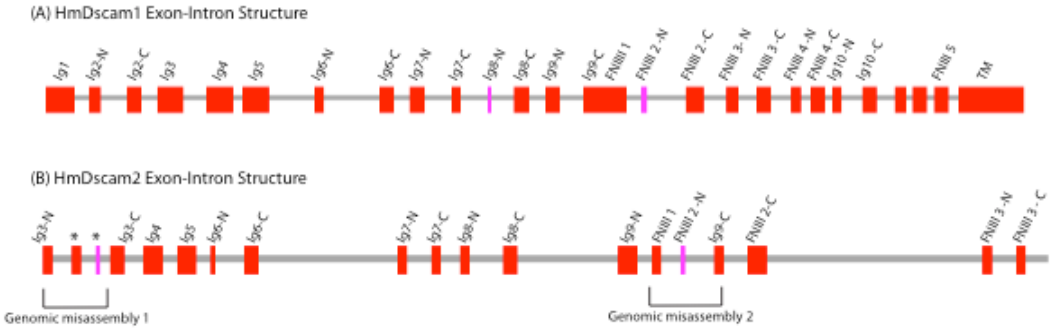

Figure 13. Exon-Intron structure of the known regions of (A) HmDscam1 and (B) HmDscam2. Exon-Intron structure was created by aligning the genomic and expressed sequences of the HmDscam's using NCBI blastn. Exons are shown as red or pink blocks, introns are shown by a grey line. Genomic mis-asemblies in HmDscam2 is denoted by brackets. In genomic mis-assembly 1, the exons with asterisks should ber 5' of Ig3-N. In Genomic mis-assembly 2, the exon for Ig9-C should be 4' of the exon for FNIII-1.

DISCUSSION

We have identified two *H. medicinalis* homologues of Dscam: HmDscam1 and HmDscam2. The HmDscam's show structural similarity to other Dscam's and are expressed in the developing peripheral and central nervous systems, implying a neuronal role for Dscam in annelids in addition to chordates, arthropods, platyhelminthes, and mollusks. The data we have now show that the HmDscam's are unlikely to be alternatively spliced, which is fitting from an evolutionary perspective as annelids are more similar to mollusks and chordates than arthropods.

HmDscam1

HmDscam1 is expressed in peripheral sensory cells in the epidermis of the mouth region, the body wall, and the tail sucker. Leeches have a variety of peripheral sensory receptors including but not limited to chemoreceptors on the upper lip, light sensitive receptors in the eyes, mechanosensory, and stretch and movement receptors in the body wall (Kretz et al. 1976, Elliot et al. 1987, Blackshaw et al. 1982, Blackshaw et al. 1993, DeRosa and Friesen 1981, Phillips and Friesen 1982). Comparing HmDscam1 and HmDscam2 expression in the head, HmDscam1 positive cells could be chemoreceptors on the mouth while HmDscam2 positive cells could be photoreceptors in the eyes. The cells that express HmDscam1 are currently unidentified but are likely to be peripheral sensory cells due to their location in the epidermis, their regular spacing, and the fact that they are made of small clusters of cells. Peripheral sensory cell bodies of different modalities (chemoreceptors, mechanoreceptors, and photoreceptors) cluster together in small ciliated organs called sensilla located on the middle annulus of each segment (Phillips and Friesen 1982, Huang et al. 1998

Johansen et al. 1992, DeRosa et al. 1981). A population of HmDscam1 positive cells could belong to the sensilla while some may belong to extra-sensillar sensory populations. Afferents of several different sensory cells fasciculate and project into the CNS where axons bifurcate and segregate into one of three axonal tracts (Peinado et al. 1987, Hockfield et al. 1983, Johnasen et al. 1992, Jellies et al. 1994). The mAb Lan 3-2 recognizes an epitope found on sensory afferents that enter all three of tracts in the CNS (Peinado et al. 1987, Zipser et al. 1981). HmDscam1 knockdown using a siRNA cocktail appears to have no observable effect on Lan3-2 positive neurons, however. It is possible that HmDscam1 is only expressed in a small subset of peripheral sensory cells and a phenotype may be observable if we used an antibody that only recognizes a subset of sensory afferents, such as Lan 4-2 or mAb's in the Laz series (Peinado et al. 1987, Johansen et al. 1992).

HmDscam1 is structurally similar to Dscam's of other species except that it is missing the fifth FNIII domain. This is unique to HmDscam1 of the leech and has not been seen in other organisms, though LvDscam in the whiteleg shrimp is missing a transmembrane and intracellular domain (Chou 2009). This lack of transmembrane domain is hypothesized to allow LvDscam to be secreted like an antibody and play an adaptive immune function.

Considering that HmDscam1 is not highly spliced, it is mechanistically improbable that HmDscam1 confers a unique identity to each neuron as it does in the fruit fly (Zhan et al. 2004, Neves et al. 2004, Hattori et al. 2007). Comparing Dscam roles across many species, there are many functions HmDscam1 may serve in the sensilla. In the periphery, it is especially important that cell bodies and neurites of sensory cells are spaced evenly and thoroughly across the epidermis to maximize information sampling of the environment. HmDscam1 may maintain dendritic self avoidance within a sensory cell of sensilla, similar to the role of Dscam in *Drosophila* da neurons (Matthews et al. 2007, Hughes et al. 2007, Soba

et al. 2007). Because HmDscam1 is not highly spliced, if it results in repulsion from homophilic binding it may play a role in both isoneuronal avoidance and tiling, as seen in the mouse retina (Fuerst et al. 2008, Fuerst et al. 2009). In the CNS, HmDscam1 may be playing a role in sensory afferent axonal bifurcation or targeting. Similar to Dscam's role in the fly mushroom body, where axons must bifurcate and send sister axons medially and dorsally, leech sensory afferents must bifurcate and send axons in anterior and posterior directions (Wang et al. 2002). Furthermore, sensory afferents must target their axonal neurites to the correct location once in the CNS, a function for Dscam that has been identified in *Drosophila* mechanosensory neurons (Chen et al. 2006).

The lack of a FNIII domain in HmDscam1 is intriguing and may affect HmDscam1 signal transduction or confer an ability to bind novel ligands. The structure of the first 8 Ig domains has been elucidated from crystal structure studies. Homophilic binding of two Dscam monomers causes a structural change in Ig4 - Ig8 that is hypothesized to play a role in signal transduction (Meijers et al. 2007, Sawaya et al. 2008). While little is known about the structure and function of the FNIII domains, their placement between Ig8 and the transmembrane domain suggest they may be important in propagating information about the structural change to the intracellular domain. Dscam in chordates can bind the axonal guidance cue, Netrin -1 (Ly 2008, Liu 2009). Perhaps the lack a FNIII domain in HmDscam1 creates a binding pocket that allows these sensory cells to respond to other axonal guidance cues in the medicinal leech. Finally, HmDscam1 is expressed in the Rz cell of the adult ganglia, suggesting that HmDscam1 function may change from the embryo to the adult.

HmDscam2

HmDscam2 is expressed in sensory cells in both the CNS and the PNS. It is expressed in the mechanosensory neurons in the CNS: the three T, two P, and either one or both sets of N cells, and sometimes the neuropil macroglia. Expression in the glial cells is not seen in all preparations and the strength of expression varies from faint to strong, suggesting that HmDscam2 expression is induced in response to an unidentified cue. HmDscam2 expression in the mechanosensory cells was identified by the relative placement, size, and shape of the cell. Its expression in the P cells has been validated via a dye fill and *in situ* hybridization experiment. HmDscam2 expression is unique and can be used to label mechanosensory neurons for further experiments in the medicinal leech. In the periphery, it is expressed in several peripheral sensory cells, including what appear to be the sensilla, as well as possible photoreceptors in the five bilateral eyes on the dorsal rim of the head and chemoreceptors on the lip (Peterson 1984, Elliot 1987).

Preliminary evidence suggests that HmDscam2 knockdown affects clustering of cells in the sensilla. HmDscam2 RNAi led to Lan3-2 positive cell bodies in the sensilla to fail to cluster together or perhaps to move away from each other. Whether this is the only phenotype of HmDscam2 knockdown on Lan3-2 positive cells needs to be explored further. In the mouse retina, Dscam plays a role in self-avoidance and tiling and Dscam mutants cause defects in mosaic spacing of cell bodies, demonstrating that this phenotype is not completely unexpected in a Dscam knockdown (Fuerst et al. 2008, Fuerst et al. 2009). Additionally, there appears to be more Lan3-2 staining around the nephridiopore as a result of HmDscam2 RNAi. This is an intriguing result considering that experiments where the CNS was excised, preventing sensory afferents from receiving guidance cues from the CNS, Lan 3-2 positive afferents fasciculated around the nephridiopore (Jellies et al. 1995). The phenotype of unclustered sensillar cell

bodies could result from the loss of an adhesion factor, or it could result from the loss of ability for the sensillar cells to migrate to the correct locations. One way to determine if this is a result of disrupted cellular migration could be to explore the HmDscam2 orthologue in *H. robusta* (HrDscam2). The cellular lineages of neurons in *H. robusta* are well defined and if HrDscam2 shows the same expression pattern in HmDscam2 in peripheral sensory cells, it would be possible to inject precursors to cells in the sensilla and observe how it is affected from HrDscam2 knockdown (Weisblat et al. 1980). Further experiments are needed to validate and explore reasons for this phenotype, but this result is helpful in showing us that Lan 3-2 is a good antibody to use for future HmDscam2 knockdown studies.

It is difficult to speculate on HmDscam2 function in the CNS without further experiments, but knowing where it is expressed can help us build some models. T cells respond to light touch, P cells respond to pressure, and N cells respond to noxious stimuli. Each neuron has a clearly identified territory of skin and forms its receptive field by sending an axon into the periphery through an ipsilateral route and forming an arbor with specialized mechanoreceptors (Blackshaw et al. 1981, Nicholls and Baylor 1968).

HmDscam2 is first expressed in P cells which are pioneer neurons (Kuwada 1985, Gan et al. 1995). There are two types of P cells: P_D which innervates the dorsal body wall and P_V which innervates the ventral body wall. P cells exhibit isoneuronal but not heteroneuronal self avoidance. In *H. ghiliani*, the P cell receptive field is formed when a central branch exits the CNS and branches several times in a stereotyped and predetermined manner (Kramer and Kuwada 1983). The outgrowths from the same P neuron will never overlap. This is in comparison to branches of adjacent and contralateral P cells which can touch and overlap extensively. Similarly, the major and minor fields of the same P cell will not overlap though the major field of a P cell will overlap with the minor field of a neighboring P cell (Kramer

and Kuwada 1983). These P cells are homologous to the P cells in *H. medicinalis,* which show isoneuronal self-avoidance unless a branch is severed from the arbor (Wang et al. 1998). The role of P cells as pioneer neurons has been studied in both leech species. The growth cone of the P_D neuron is the first to project into the periphery in *H. ghiliani* and establishes the efferent DP nerve (Kuwada 1985). Ablating the P_D neuron in *H.medicinalis* AP and T_D neurons which follow along its path to target to the correct area but have reduced arbors (Gan et al. 1995).

Like HmDscam1, HmDscam2 does not appear to be alternatively spliced therefore probably does not confer unique cellular identity to each neuron. There is a possibility that HmDscam2 plays an adhesive instead of a repulsive role in this system. Perhaps the phenotype of HmDscam2 RNAi in the sensilla is due to a loss of HmDscam2 mediated homophilic binding and adhesion. Though mechanosensory neurons in *H. ghiliani* were the first model where isoneuronal self avoidance was described (Kramer and Stent 1985), it is unlikely that HmDscam2 is fully responsible for isoneuronal self avoidance in P cells for two reasons: (1) If the same HmDscam2 isoform is expressed in all P cells and it mediates isoneurite repulsion via homophilic binding, it should also mediate neurite repulsion between two heteroneuronal cells that overlap, and (2) isoneuronal self-avoidance is abolished within 24 hours when a first order P cell branch is detached from the cell body. In this time, the severed branch shows sprouting and growth, showing that molecular components in the severed branch were still functioning and surface molecules were likely to still be in the plasma membrane (Wang et al. 1998).

In the vertebrate system, cultured neural cells with Dscam expression extended neurites that sometimes fasciculate (Yamagata and Sanes 2008). Fasciculation has been previously identified to be necessary for growth along pioneer projections (Johansen et al. 1992, Raper review 2010). Dscam has been shown to be important for outgrowth following

pioneer neurons in *Drosophila*. In the *Drosophila* mushroom body α'/β' axons serve as pioneer axons for α/β axons. Mutating Dscam before the α'/β' axons develop causes α/β axons to segregate inappropriately into medial and dorsal lobes while mutating Dscam after α'/β' axons have developed allows α/β axons to segregate appropriately, though unevenly (Wang et al. 2002). Additionally, mutations that reduced Dscam diversity in mechanosensory neurons had similar phenotypes to ablating the pioneer fibers they rely on (Chen et al. 2006). HmDscam2 is also expressed in the three T cells that have secondary processes that appear to grow along one another and make close contact with one another (DeRiemer and Macagno 1981, Macagno et al. 1985). Moreover, T cell axons tend to follow one another within connectives and regenerating T cell axons grow along these connectives (Macagno et al. 1985). At least one T cell, the T_D cell relies on guidance cues from the HmDscam2 positive P cells (Gan et al. 1995) The idea that HmDscam2 plays an adhesive function would fit with it's expression pattern. Though all neurons in segmental ganglion are born at about the same time (Kristan et al. review 2005), the P cells express HmDscam2 before the T cells. Perhaps the P cells express HmDscam2 as they form their arbors, and the T cells express HmDscam2 to follow those arbors. Additionally, HmDscam2 is expressed in all three T cells whose axons maintain contact between ganglia and whose secondary processes maintain contact within the same side of the same ganglia. In this way, HmDscam2 may serve a more canonical IgSF adhesive function upon homophilic binding such as N-CAM, Fasciclin II, and Contactin (Yoshihara et al. review 1990).

An interesting experiment would be to knockdown HmDscam2 expression before and after the P_D cell forms it's arbor, either by injecting a siRNA cocktail under the germinal plate at an early age or injecting the cocktail into the P_D cell at a later age. P_D and T_D arbors could then be filled with dyes and help determine if HmDscam2 function is important for the T_D cell

to use the P_D dendritic arbor to shape it's own arbor. We should also repeat experiments of HmDscam2 knockdown and Lan 3-2 antibody staining to better image if HmDscam2 knockdown has an effect on sensory afferents in the periphery or CNS.

Evolutionary Relationship of HmDscam's to Dscam's of Other Species

Dscam structure is conserved with > 20% amino acid identity across deuterostomes and protostomes with the ancestral Dscam homologue being predicted to be 600 million years old (Shmucker et al. review 2009). The HmDscam's cluster together and share a parental node of with platyhelminthes. The platyhelminthes Dscam is expressed in nervous tissue but it is not known whether it is alternatively spliced or not (Fusaoka et al. 2006, personal communication with Fusaoka). The arthropod Dscam's share a common node as do the chordate/mollusk Dscams. More sophisticated phylogenetic software should be used to study the evolutionary relationship between chordate, mollusk, and annelid Dscam's, especially because we predicted that mollusk Dscam should cluster with platyhelminthes and annelids. It is important to note that annelid Dscam's do not share a common node with arthropod Dscam's, further supporting that they are not heavily spliced and suggesting that their role in neural circuit formation is not to confer a unique identity to each neuron. Chordates and annelids have two Dscam paralogues so it is likely that mollusks and platyhelminthes have a second unidentified Dscam paralogue as well.

In conclusion, HmDscam's are expressed in nervous tissue, showing shared neuronal expression of Dscam across four phyla. However, a review of the literature highlights the multiple neuronal roles of Dscam across the animal kingdom rather than one specific role. It is fascinating that Dscam is of critical importance to neuronal development but each phylum utilizes it in a different manner. Further experiments can show the function of HmDscam1 and

HmDscam2 in the annelid and demonstrate how annelid neuronal development evolved in relation to other phyla.

APPENDIX

HmDscam1 expressed nucleotide sequence - 5418 bp

Premature stop codon is underlined.

```
   1 TACGAGGTTC GATGGACCGA TTTTCACTTC CGAATTCCCG CGAGAGATCA TCCAGGTGGA
  61 AAAAGAGAGC GTAAATGTTT CCTGTTCTGC ACTAGGAGAC CCTCCAGCCA CCATCACCTG
 121 GATGAATTCG GACACAAACC AGGTCCTGGA TAACTTTCCA GGATCGGTGG AGTTGAAGGA
 181 CTCCACAGTG ATTTACCTGC CTTTCCAGGA GCACAAGCTG ACTGCACACG ACGTCACGAG
 241 GAGGATCAGA TGCCTCGCTT CCAATGTGCA CGGCTCCATT CTCAGCCCGG AAATCTACGT
 301 AAGACCAATC ATTCTGAAAC ACCATGCAAA CTTTCAAGCA GAAATTTCTC AGTCCGATTC
 361 TTGTATGGAA GGGACTGCTG TTTTGAGTTG TTCAGTCCTT CCTGACCATG CCAGAGAATT
 421 CACCAATGTA CTTCGCTGGA CTAGAAACGA TGTGCCTCTT ATACAAGACG AGAAACACAG
 481 GATGATTGGA AAATCACTTC ACATCTTCAA CGTTACAACG GAGGACACCG GTAATACCTT
 541 CAGGTGTGTC GTGAGGCACT CGTTGTTTGG AACTACCAGG TTGAGTCCAC CTGTCAAACT
 601 CGTGCTGAAC GGACCGATAA ATACAGCTCC AATCATGGTT TCCTTTGCAA ACGTGGTAAA
 661 AACGAAAATG GGCGACAATG TAGAACTTCC GTGCTCTGCA AACGGTGCTT CTCCAATGAA
 721 ATACACCTGG AAAAGAATGG TTCATATTTC CACCTGGGGT GTAATGTCTT CTCTAGAAGA
 781 GTCTATAAAC AGCTCGAATG GCAGGATGTC TATGATTGGT CAGAACCTCG TGATCTTGTC
 841 TGTCAAGAGG CAAGATGTTG GTCCATACGT TTGCAGGGCT GAAAACCACC TTGGAAGAAT
 901 GGAACGTGTT GTTCAACTCG TCCTCGCGAG TGCATCACTT ATTGCCTTCC TCGAGCCTCA
 961 ATTCCAGCAG TTCAACCTCA GCCAATCAGT GACCTTGAAA TGTCAAGTCT TGACCTCTGA
1021 TTCTGTTGAC CTCATCACGT GGTTCCACAA TGGGAAACCG CTCGACAAGA CTTCTGCTAA
1081 TTCGGAGGGC AAATATCACT TCTTAGAGAA GAGTTCGACG TTGATGATAA CAAACATGGA
1141 TTTGAACGAT GCAGGGATGT ACCAGTGCTT TGTTGGCATA GGTGGAGAGG TTGTGCAGGC
1201 TGTTGCCCAA GTTAAGCTTG GAGCTCTTCC AGCGGAAGTC CTGAGCGGCT TTGCGACTCA
1261 ATTCACTCAA CCAGGACCGA ACGTGTCAAT GCGATGCTTG GTTCGCGGCA AACCGGAACC
1321 GGAAGTAACC TGGATACACG ACGGAGTGAG ACTCTCGTCT GACCTGGAAA GCTCCAAAAG
1381 ATCGATCACC AGCTTCCGAT TGCACAATGG AAGTTCATTG ATCGAGTTGA CGTTGACAAG
1441 GATCACGATC CATGATGGCG GAACTTATGC GCGCGTTGCT CACAATCGAT TTGGTTCGAT
1501 GTTCCACGAG GCCAGGTTGG AAGTCTACGG AATTCCTGAC ATTCGTCCCA TGAATGACGT
1561 CACCGTGGTG ACCGGAAGTA CGGTCGTTCT TCCGTGTTTC GTGATTGGAT ATCCTGTGCA
1621 CAGGGTCGCG TGGCGCAAAA ACGACGGGGT CATCCTACCG ACCAATCACA GACAGATAGT
1681 GGCTCTAAAC CACTCCCTCG TAATCCACTT AAGTGACAGA CTCCAGGACT CGGGTATTTA
1741 CACCTGCGAA GCAACCGATG AACATGGAGC AGGTGTGGAA AGAAAGATGA ACCTGTTTCT
1801 CATTGACGTG CCGAAAATAG ATCCGTTTCA TCTGCCGGAG AAAAAGTTTG AAGGAAGGTT
1861 GTCGGTGACA TGTGCAGTCA GTGAGGGTGA TTTGCCCATG GAGGTTTACT GGAGTAAAGA
1921 TGGAAAGAAA ATATCTGAGG AAGACGCCGT AATTAATATT CAGATGATAA GTTCATTCAC
1981 CAGTCTCCTT TCGATAGAAC ATTTGACCTG GTCGCACCCA GGTAATTACA CCTGTACAGC
2041 ATCAAATCGT GCAGTGATCG TGTAACCAGC CCCACTTCAC TGCTCAATGT GGATGTTCCG
2101 CCATTTTGGG TCGTCGAACC AAACGACGAA GAAACTGTGG TTGGAAATCA ACTCAGGTGT
2161 GCCTCAGCCA TCAATATCTT GGAGTAGAAG TTCAACAGTC ACAGACCAAT TCAATGTAAC
2221 CGCGGAGAAT TTTTCTGCCG AATCCAAAAG GAAGCCGTTG CTTAGCAACG GCAGTTTGTT
2281 TATAAACGGA GTACTCAAGT GTGATGAGGG ATTCTACGTC TGCAGGTCGT GGAACGGAGT
2341 TGGACCTGAC ATCAGCAGGA TCGTCAGCCT GGTTGTGCAC GCACCTGCAA GATTCCTAAC
2401 GAACGGTGAA AATGAATTAA GTGGAAAGAT CGGTTCGAAG GTCACGTTGA GATGTGACGT
2461 CACAGGGGAC CACCCAACAT CCATAGCCTG GAACAAAGAT AAGGAAAACC TCGACTGGAC
2521 GAATGATAAG AGACTGAGGT TGTCTTTCTT CAATGATTCA GACATCATCC GATCCTCAAC
2581 ACTGGAGATT TCATCGCTCC AAAGTAAAGA CACAGGTTGG TATTACTGCA TCGCTCAAAA
2641 CAAATACGGA AGTCACGTCA TCAGGAAGCA CCTGATCGTC CTCGAGCCTC CTCAACCTCC
2701 TCGTCAACTA CAAACGACGG AAGTGACGCC TCGCAGCTTC AAGATCGCAT GGAAAACGGA
2761 GAGGCAGGAC CCACCTCTCC TTGGACACTT TGTCAGCGTC AGAAGCAACC GAGATCTCAC
2821 AAATGGAAGC AAAAATTCTG AGTTTCACGC ACGAGAAATA TTCCTCAACT CATCAATTTC
2881 TGAGGAAGAA TTTGAGAAGG AAATGTTCAG CGTCGACCAT CTAAGGCCCG GAAGCGAATA
2941 CAAAGTCAAA ATAAGAGCAG TAAATGAAGT CGGGAGAGGA TCGTGGAGTG ATGATGTCGT
3001 TGTTGTCACA CCAGAGGAAG TTCCATCGAC TGCACCTGAA GGCGTCCAAG TTGTTGCTGT
3061 TGATTCTCAA TCTCTCAACA TCTCGTGGGG GCCACCGTCC AAAGATAGAA ACGGGGCGAT
3121 CCTTGGATAT CACGTGGGTG TCAGGAGAGA AGGAGGAAGA CATCAGGAGC CCTTCATCTT
3181 TTATTCGTTC AAAGTCATTC CATTTTGTGT CCTGGAAAAT CTTAACAAGT TCACTGAGTA
3241 CTCGGTCGTT GTGCAGGCCT TCAACTCAAA AGGCGTCGGT CCAAGGTCAA ACATTGTCAA
3301 AGCCAAGACT CTGGAGGATG TGCCAGACCA AGCGCCACAA TCCGTGAAGG GTCAAGCAGT
3361 TGATTCAGAG ACAATACACA TTTCTTGGAA GTCACCAGTG TCTGCAAATG GCATCCTTTT
3421 GGGTTACAAC ATGTTCTATC GAAGAACCAA TAGTTTAGAT CCAAATGAAC TTCTGGTGAC
3481 CAACAGAAGT TACATAACAT TGTCGGGACT CTTCAAATTC ACAAATTACA CGTTCACCCT
3541 GGCAGCTTTT ACTGCTGCCG GAACAGGAGC TCTCAGTACT CCGGTGGTTG TTGCTACTTT
```

```
3601 GGAAGATGTT CCAGGCGCTC CAGCAGATGT GATTGCCTTT CCATCAGCTC CGGGAGAAAT
3661 ATCTGTACAC TGGAAACCAC CAACGGAAAA AAATGGAGTT CTTACCGGCT ACGTGCTCAA
3721 TTTGTTTCAA ATGGATGACC AGAATTTCAG TAAAGAGTTC CACTTTCCCC CTTCAGCCGT
3781 GAAAGGCGTC ATTTCAAATC TGACCTTGAA CGTTAATGTC AGCATTAAGG TCAAAGCATT
3841 GACTCAAATT GGAGTTGGGA TATGCTCGAA GACAGTTTTT ATTCAAGTTC AAAACTCAGC
3901 CCCTGCAAGA GTGGTTGAGT CTTCAAAAAG CATAGCGACT GACTTGAAAG AAACAGTCAT
3961 ACTTCCGTGC ACGACAATTG GATATCCACA ACCAAGCAGG ATTTGGAAGA AACAGGGAGA
4021 AAGTGATGAT TGGATAAACA CTTCCTCTTG TTGTACTTTC AACGAAGATG GATCGCTGAC
4081 GATTAAGCAG TTGGAGGAAG ACATGACTGG AATTTATGAG TGTGAAGTGG AAAACATCCA
4141 CGGAAAGGAT AAAACGGAAT TTGAGGTCTC TGCTGCAAAT CTCTCATCAC CAATTAAGTT
4201 GATCGCTAGA TATAACGATT ACCTTCAGTT CCAATGGACT CTGAAATATT CAAAGAAATT
4261 GAAGGAAATT CTTGTACAAT ACAGGGAGTA TAATAATATT GGATTATCTG ATGATCTTAT
4321 TTACAATAAT GCAATACTAA GTGACAATTG GATTGCTGTG AACATTTCGA TGAGTCGAAA
4381 TTTCACGATA AATGATCTCA AATGTGGTCG AAAGTATGAA ATTAGAGTAA AAAAGCAAGA
4441 AATGCTGAGC GACAGCACCA TAGAGCGGTA CTTGTATAGT ACCTCCGGTG AAAGACCTAT
4501 CATGCCCGCC CAGAGCCTAC TAATCCAAAA GATCAACTCC ACTTTTATTT CTCTCAACCT
4561 TTCTACCTGG ATGGATGGCG GATGTCCTAT ATTAAGATAT AAGGTCGAAC TTCAGACGAA
4621 TGCTGCAGAA GAGTGGCATC TTGTAGTTGA CGACTTATTC GGTCATAAGG ACACTTTTCA
4681 TTTAGAGCTT CATCCAGACA CCTGGTACAG GATGCGTGTA ACAGCACAAA ACGATGCCGG
4741 AAAGACGGAA TGTATGCTAA ACTTTGCAAC TCCGCCACTG GAGTTGCTCG GAAGTTCATC
4801 AAAACAAGTA GAGCAGCGCG GATGGGCTGC TCGCATGACA AAATTAAAAA TTCAAAAGCA
4861 AACTCCCTTA TATGATAGGC TGGAATTCAT GCTTCCATTA ACGACTGCCA TTATGTTGAC
4921 TATCGTAGCT ATCACGATAA TCATTTTGCG CAGGAAACGC AAACTCCAGG AAAAACAATC
4981 TCAAAGCAAC CACAATTCTT TGCTTCGAGT ACCTAAACAC GACATCGAAA CAATACACTT
5041 TCGCACTCAG CAGTCATTTT CTTCGGCTCT CAGAAATGAA GCAGAGTTCA ATTACGAAAG
5101 TCAAGTTGGA GATGTCATTT ATAAAAAGTT TCGAAATTGT GGACATAAAG TCAAGGATGG
5161 AATCGTACGC AATCAGGAAC TGCGGAAAAA GCGGATATTC CCTTCGACGC CGATCAGTTC
5221 TGTGTGCGAA GTTGAAAAGC TTCCAAACGA ATGAAAAAGA ACATACTGAG AGATGTACAG
5281 ATGGAAAGGT TTTTTCGACG TTTTGATTCG AGTAGTGGCT ATACAACTCC TCCACTTTCA
5341 AATTTTTCTG GAGATCCGAA CGCTTCGTAG GAACTGAGGA CATTAGTGTG ATCTATAAAT
5401 AAATGACTTT TAATCTTG
```

HmDscam1 expressed protein sequence - 1749 aa

```
   1 TRFDGPIFTS EFPREIIQVE KESVNVSCSA LGDPPATITW MNSDTNQVLD NFPGSVELKD
  61 STVIYLPFQE HKLTAHDVTR RIRCLASNVH GSILSPEIYV RPIILKHHAN FQAEISQSDS
 121 CMEGTAVLSC SVLPDHAREF TNVLRWTRND VPLIQDEKHR MIGKSLHIFN VTTEDTGNTF
 181 RCVVRHSLFG TTRLSPPVKL VLNGPINTAP IMVSFANVVK TKMGDNVELP CSANGASPMK
 241 YTWKRMVHIS TWGVMSSLEE SINSSNGRMS MIGQNLVILS VKRQDVGPYV CRAENHLGRM
 301 ERVVQLVLAS ASLIAFLEPQ FQQFNLSQSV TLKCQVLTSD SVDLITWFHN GKPLDKTSAN
 361 SEGKYHFLEK SSTLMITNMD LNDAGMYQCF VGIGGEVVQA VAQVKLGALP AEVLSGFATQ
 421 FTQPGPNVSM RCLVRGKPEP EVTWIHDGVR LSSDLESSKR SITSFRLHNG SSLIELTLTR
 481 ITIHDGGTYA RVAHNRFGSM FHEARLEVYG IPDIRPMNDV TVVTGSTVVL PCFVIGYPVH
 541 RVAWRKNDGV ILPTNHRQIV ALNHSLVIHL SDRLQDSGIY TCEATDEHGA GVERKMNLFL
 601 IDVPKIDPFH LPEKKFEGRL SVTCAVSEGD LPMEVYWSKD GKKISEEDAV INIQMISSFT
 661 SLLSIEHLTW SHPGNYTCTA SNRAVIVPAP LHCSMWMFRH FGSSNQTTKK LWLEINSGVP
 721 QPSISWSRSS TVTDQFNVTA ENFSAESKRK PLLSNGSLFI NGVLKCDEGF YVCRSWNGVG
 781 PDISRIVSLV VHAPARFLTN GENELSGKIG SKVTLRCDVT GDHPTSIAWN KDKENLDWTN
 841 DKRLRLSFFN DSDIIRSSTL EISSLQSKDT GWYYCIAQNK YGSHVIRKHL IVLEPPQPPR
 901 QLQTTEVTPR SFKIAWKTER QDPPLLGHFV SVRSNRDLTN GSKNSEFHAR EIFLNSSISE
 961 EEFEKEMFSV DHLRPGSEYK VKIRAVNEVG RGSWSDDVVV VTPEEVPSTA PEGVQVVAVD
1021 SQSLNISWGP PSKDRNGAIL GYHVGVRREG GRHQEPFIFY SFKVIPFCVL ENLNKFTEYS
1081 VVVQAFNSKG VGPRSNIVKA KTLEDVPDQA PQSVKGQAVD SETIHISWKS PVSANGILLG
1141 YNMFYRRTNS LDPNELLVTN RSYITLSGLF KFTNYTFTLA AFTAAGTGAL STPVVVATLE
1201 DVPGAPADVI AFPSAPGEIS VHWKPPTEKN GVLTGYVLNL FQMDDQNFSK EFHFPPSAVK
1261 GVISNLTLNV NVSIKVKALT QIGVGICSKT VFIQVQNSAP ARVVESSKSI ATDLKETVIL
1321 PCTTIGYPQP SRIWKKQGES DDWINTSSCC TFNEDGSLTI KQLEEDMTGI YECEVENIHG
1381 KDKTEFEVSA ANLSSPIKLI ARYNDYLQFQ WTLKYSKKLK EILVQYREYN NIGLSDDLIY
1441 NNAILSDNWI AVNISMSRNF TINDLKCGRK YEIRVKKQEM LSDSTIERYL YSTSGERPIM
1501 PAQSLLIQKI NSTFISLNLS TWMDGGCPIL RYKVELQTNA AEEWHLVVDD LFGHKDTFHL
1561 ELHPDTWYRM RVTAQNDAGK TECMLNFATP PLELLGSSSK QVEQRGWAAR MTKLKIQKQT
1621 PLYDRLEFML PLTTAIMLTI VAITIIILRR KRKLQEKQSQ SNHNSLLRVP KHDIETIHFR
1681 TQQSFSSALR NEAEFNYESQ VGDVIYKKFR NCGHKVKDGI VRNQELRKKR IFPSTPISSV
1741 CEVEKLPNE
```

HmDscam2 expressed nucleotide sequence – 3362 bp

Premature stop codon is underlined.

```
   1 GGGNTTGACG CGGGTCTTCA CTGCACCATC AACCCGTACT TCATGCCTGA GTTTGTAGAG
  61 GTGGTCCTGT GGCAAGAGGA GAACCATGAC GTCATTGCCA CAGTCGGAGG CAGGTACAGC
 121 CTGATGGATC ACGGGGACCT AAGGAATGGC AAGCTCCAGG TGAGGAATGT CAGTGTGTCT
 181 GACAGTTACA AGAAGTTCAG ATGTCTGACC AGGAACATAC TGACCGGGAA GGAGAGAAGC
 241 AGTAACCCCG CCCACCTCAT TGTCATTGAG GCCTCAAACT CTCCGCCCAG AATGACTGAT
 301 CACGTGAGTT CAATGAGTGC AACCGGAGGT GCGACCGTGG AGCTTCCGTG CGTTGCCATA
 361 GCAATACCGG TGGCACGCTA CCGGTGGAGC AAGAATGGAA TTCCCATCCG TGGAGGACGC
 421 AGTGATGATG ACGACGATGA TGGCGATCGT ATTTCAGGTG GAAGCATGAG ATTGGTGGGC
 481 GGCAATCTGT TGGTCGAAAA CGTCCAGGTG GATGACTCAG GTGATTACAG GTGTGTGGCT
 541 GAAAATTACA TAGGACACCA GGGGACATCC GTGAGACTGA GCGTCACAGT GCCCCTGTCA
 601 GTCCACATAG AACCACAGTT ACAAATTGTG GATTCGGGAA AGGAAGCACT ACTCCGTTGT
 661 CACGTGTATG GCCATCCGCT GGAAAACATC ACGTGGTTGA AGGACGGAAA GGTGCTAATG
 721 ACACAGGACG AGTCAGGGGT CGTGACCGAA GGCCAGGGGG GTTCGGTGCT GAGGGTCGAA
 781 AAGGTCAAGA GGTCAGACGG CGGGATGTAC CAGTGTGTAG TAGCCCGGTG GGGGTAACGA
 841 GATTGTCCAT GCCTCCGCCC AAGTGTCCCT AGGGGCCTTT CCTCCGTACA TCCTCAAGGG
 901 CATACCCTCT GAATACTTGG ATCCAGGAAC TCGCGTCTCC ATGGCCTGCT CTGCCATCGG
 961 GAATCCAAGG CCCACGATCA CGTGGTCAGT GGACGGAGAA CTACTTCCTT CCGGTCAACA
1021 TGGTGACCGG ATATCAGTCG GAGATTATGT CGGCGGAGTT GGAGAGGTTG TCAGTCACGT
1081 GAACATCACC GGAGCGAAGG TCCAAGATGG CGGGAGGTAC GAATGTCTAG CAGCCAATGA
1141 CATGGGAGAG GCAAGGGAGG TTGCCACGCT CAGCGTTTAT GGAATTCCCC ACATCAAACC
1201 AATGAGGAAC ATCACAGCTC TTGCCGGAAG TGAACTGGAA GTCATCTGTC GTGTGACAGG
1261 ATATCCGGTC AAGGAAATAT TTTGGAAGAA AGGTGCTCAG CTGCTACCGA CCAATCACAG
1321 GCAGGCCACG TCATCCAATG GATCGCTATT GGTCAGTCGA TTGAACCGCC ACCTGGACGA
1381 GGGTTGGTAC TCGTGTCACG TGACTAATCC GAACGAACAG GGCACCGAGA GGAGGTTCTA
1441 CGTTCGTGTT ATAGATCCAC CTAAAATAGA CCCCTTCAAC TTTCCTCCCC GAAAGCAAGG
1501 CGATCGTGTT TCCGTGTCGT GTGTCATGAG CAGTGGTGAC CTTCCGATGA CCTTGGAGTG
1561 GAGGAAGGAT GGTGACGTCA TCGACCCTGA CCTCGGAGTT CAAGTTCACC AATCTGGGTC
1621 GTACTCCAGT TTCCTGTCCA TTGAATATGC AACACCTCAT CACGACGGTA ACTACACCTG
1681 TATTGTCACC AACGAGGGAG CCACCACGAG CCACACCTCC CAACTCCACA TTAACGTTCC
1741 TCCGACCTGG GTCTCAGAAC CCGAGGACAC GAGCATCAGC TTGTCTGGCA CTGCCACCAT
1801 CCACTGCCAG GCCTTTGGGG ATCCTGAGCC CAAAGTCGTT TGGAGGAAGA GCGTTGGCGG
1861 AATCATGAGC TCGTACCGTG ACATCGATCT GGACGGAGAG GAGACAAGGC TGACTCACCT
1921 TGGCAACGGG ACGCTCGTCA TCAATTCGGT TGCTAAGGAA GATCAGGGCT TCTACCTGTG
1981 TCATGCCAAC AACAACATTG CCCCAGACCT CAGCAAAGTG ATTCACCTCA CAGTGCATGT
2041 TCCAGCAACC TTTAAGGTCC ACGATTCCAA CGTGTCGGTG TCCAGAGGTG AGACTGCGGT
2101 CATGGAATGT CACTCGGAAG GTGACCTCCC TACGACCTTG AACTGGAGTT CAAGCAACAA
2161 GCAGCCAGCC GATTGGGGTC ATGACCCAAG GTACCAGGTG ACCTATTTCC CCACCGACAA
2221 AGGTCAGGGT TCAAGGCTGA CAATCAATCA CGTGACGAAA GAAGATTCTG GATTCTACGT
2281 CTGCGCGGCC AAAAATAACT ACGGAAGTGA CGTAACGAGC ATCCACCTCA TGGTCCTCCA
2341 GTTGCCCGAT GCACCCTTTG ACTTTGGGGT CGTGGTGACC GGTGTCGAGT GGGCGGAGTT
2401 GAGGTGGAAA GAGGGTTTCC ATGGAAACTC GGAGATCGTG GAGTATCTGG TCGAGATTTA
2461 CAGGGAGGGA GATTCCTGGA AAAACCCGGT GGCCTCAGTG GAAGTGGAGG CTTCTTTCAG
2521 ATCAGTGAGG GTGGACAAAT TGAGGGCAGC AACAACTTAC AAGGCGAGGG TAAAGGCAAG
2581 AAACAAGGTC GGATTTGGAG AACCCAGTGA GGTCGTGACC TTTGACACGC TGGAACTGGA
2641 ACCCTCCTCT CCGCCCATCA ACGTGTCAGC CATGGCGGTC GGATCTCAGT CCATTCTGGT
2701 TACTTGGCAG CCAGTCCCGA GGTCGCACGA AAACGGTCGG CTACTCGGCT ACTCAGTCGG
2761 CTACAGGAGA GTCAACTCCA CCTCTCCTTA CCTCTACGTC ACCGTCACCA ATGACAACCA
2821 GCGACGCATC GATGACCTGG AAAAGTTTTC CACCTATTCC ATCAACGTGA GGGCCTTCAA
2881 TGGCAAGGGA GCTGGGCCGA TGTCGGAAGA CGTGCTCGTC ATGACGCTTG AGGATGTACC
2941 CAGTCACGCC CCAGTGAACG TGACCGCCAG CAGTCAATCA TCCGAGACGA TTTCAGTGAC
3001 CTGGAGTCCA CCCCCTCTCC ACTCTCTCAA CGGGGTCATC GTGGGGTACA AGGTCATCTA
3061 TAAATACGTT CGCCATGATG AGGATGAATC CGGCTCCATG TCGCAGACGA CCGAGCAACT
3121 GAGCACTACA CTCAGACATC TCAAGAAGTT CACCAACTAC AGTGTTCAAG TCCTGGCTCG
3181 GACCAAGATC GGGGAGGGCA TCGCAAGTTC GCCCGTCTAC GTCACCACTT TGGAAGATGT
3241 TCCAGGTCCA CCTGAGGACA TGAAAGTGCT GGTGATGAAC AGGAGCAGCA TAATGGTGGT
3301 GTGGAAACAA CCTCTCCAGG TCAACGGTGT CCTCACCAGG TATTTCGTCA ACGTCACACA
3361 AC
```

HmDscam2 expressed protein sequence – 1122 aa

```
   1 DAGLHCTINP YFMPEFVEVV LWQEENHDVI ATVGGRYSLM DHGDLRNGKL QVRNVSVSDS
  61 YKKFRCLTRN ILTGKERSSN PAHLIVIEAS NSPPRMTDHV SSMSATGGAT VELPCVAIAI
 121 PVARYRWSKN GIPIRGGRSD DDDDDGDRIS GGSMRLVGGN LLVENVQVDD SGDYRCVAEN
 181 YIGHQGTSVR LSVTVPLSVH IEPQLQIVDS GKEALLRCHV YGHPLENITW LKDGKVLMTQ
 241 DESGVVTEGQ GGSVLRVEKV KRSDGGMYQC VVARWGPGGG NEIVHASAQV SLGAFPPYIL
 301 KGIPSEYLDP GTRVSMACSA IGNPRPTITW SVDGELLPSG QHGDRISVGD YVGGVGEVVS
 361 HVNITGAKVQ DGGRYECLAA NDMGEAREVA TLSVYGIPHI KPMRNITALA GSELEVICRV
 421 TGYPVKEIFW KKGAQLLPTN HRQATSSNGS LLVSRLNRHL DEGWYSCHVT NPNEQGTERR
 481 FYVRVIDPPK IDPFNFPPRK QGDRVSVSCV MSSGDLPMTL EWRKDGDVID PDLGVQVHQS
 541 GSYSSFLSIE YATPHHDGNY TCIVTNEGAT TSHTSQLHIN VPPTWVSEPE DTSISLSGTA
 601 TIHCQAFGDP EPKVVWRKSV GGIMSSYRDI DLDGEETRLT HLGNGTLVIN SVAKEDQGFY
 661 LCHANNNIAP DLSKVIHLTV HVPATFKVHD SNVSVSRGET AVMECHSEGD LPTTLNWSSS
 721 NKQPADWGHD PRYQVTYFPT DKGQGSRLTI NHVTKEDSGF YVCAAKNNYG SDVTSIHLMV
 781 LQLPDAPFDF GVVVTGVEWA ELRWKEGFHG NSEIVEYLVE IYREGDSWKN PVASVEVEAS
 841 FRSVRVDKLR AATTYKARVK ARNKVGFGEP SEVVTFDTLE LEPSSPPINV SAMAVGSQSI
 901 LVTWQPVPRS HENGRLLGYS VGYRRVNSTS PYLYVTVTND NQRRIDDLEK FSTYSINVRA
 961 FNGKGAGPMS EDVLVMTLED VPSHAPVNVT ASSQSSETIS VTWSPPPLHS LNGVIVGYKV
1021 IYKYVRHDED ESGSMSQTTE QLSTTLRHLK KFTNYSVQVL ARTKIGEGIA SSPVYVTTLE
1081 DVPGPPEDMK VLVMNRSSIM VVWKQPLQVN GVLTRYFVNV TQ
```

REFERENCES

Agarwala KL, Nakamura S, Tsutsumi Y, Yamakawa K. Down syndrome cell adhesion molecule DSCAM mediates homophilic intercellular adhesion. Brain Res Mol Brain Res. 2000 Jun 23;79(1-2):118-26. PubMed PMID: 10925149.

Agarwala KL, Ganesh S, Amano K, Suzuki T, Yamakawa K. DSCAM, a highly conserved gene in mammals, expressed in differentiating mouse brain. Biochem Biophys Res Commun. 2001 Mar 2;281(3):697-705. PubMed PMID: 11237714. (A)

Agarwala KL, Ganesh S, Tsutsumi Y, Suzuki T, Amano K, Yamakawa K. Cloning and functional characterization of DSCAML1, a novel DSCAM-like cell adhesion molecule that mediates homophilic intercellular adhesion. Biochem Biophys Res Commun. 2001 Jul 20;285(3):760-72. PubMed PMID: 11453658.

Blackshaw SE. Morphology and distribution of touch cell terminals in the skin of the leech. J Physiol. 1981 Nov;320:219-28. PubMed PMID: 7320936; PubMed Central PMCID: PMC1244043.

Blackshaw SE, Nicholls JG, Parnas I. Expanded receptive fields of cutaneous mechanoreceptor cells after single neurone deletion in leech central nervous system. J Physiol. 1982 May;326:261-8. PubMed PMID: 7108791; PubMed Central PMCID: PMC1251472.

Blackshaw SE. 1993. Stretch receptors and body wall muscle in leeches. Comp Biochem Physiol 105A:643–652.

Briggs KK, Johansen KM, Johansen J. Selective pathway choice of a single central axonal fascicle by a subset of peripheral neurons during leech development. Dev Biol. 1993 Aug;158(2):380-9. PubMed PMID: 8344457.

Brites D, McTaggart S, Morris K, Anderson J, Thomas K, Colson I, Fabbro T, Little TJ, Ebert D, Du Pasquier L. The Dscam homologue of the crustacean Daphnia is diversified by alternative splicing like in insects. Mol Biol Evol. 2008 Jul;25(7):1429-39. Epub 2008 Apr 9. PubMed PMID: 18403399.

Chen BE, Kondo M, Garnier A, Watson FL, Püettmann-Holgado R, Lamar DR, Schmucker D. The molecular diversity of Dscam is functionally required for neuronal wiring specificity in Drosophila. Cell. 2006 May 5;125(3):607-20. PubMed PMID: 16678102.

Chou PH, Chang HS, Chen IT, Lin HY, Chen YM, Yang HL, Wang KC. The putative invertebrate adaptive immune protein Litopenaeus vannamei Dscam (LvDscam) is the first reported Dscam to lack a transmembrane domain and cytoplasmic tail. Dev Comp Immunol. 2009 Dec;33(12):1258-67. Epub 2009 Aug 5. PubMed PMID: 19635499.

Chou PH, Chang HS, Chen IT, Lee CW, Hung HY, Han-Ching Wang KC. Penaeus monodon Dscam (PmDscam) has a highly diverse cytoplasmic tail and is the first

membrane-bound shrimp Dscam to be reported. Fish Shellfish Immunol. 2011 Apr-May;30(4-5):1109-23. Epub 2011 Feb 19. PubMed PMID: 21338689.

DeRiemer SA, Macagno ER. Light microscopic analysis of contacts between pairs of identified leech neurons with combined use of horseradish peroxidase and lucifer yellow. J Neurosci. 1981 Jun;1(6):650-7. PubMed PMID: 7346573.

DeRosa SY, Friesen WO. Morphology of leech sensilla: observations with scanning electron microscope. Biol. Bull. 1981 Jun; 160: 383-393.

Dong Y, Taylor HE, Dimopoulos G. AgDscam, a hypervariable immunoglobulin domain-containing receptor of the Anopheles gambiae innate immune system. PLoS Biol. 2006 Jul;4(7):e229. PubMed PMID: 16774454; PubMed Central PMCID: PMC1479700.

Elliott EJ. Morphology of chemosensory organs required for feeding in the leech Hirudo medicinalis. J Morphol. 1987 May;192(2):181-7. PubMed PMID: 3599081.

Fernández J, Stent GS. Embryonic development of the hirudinid leech Hirudo medicinalis: structure, development and segmentation of the germinal plate. J Embryol Exp Morphol. 1982 Dec;72:71-96. PubMed PMID: 7183746.

Fuerst PG, Koizumi A, Masland RH, Burgess RW. Neurite arborization and mosaic spacing in the mouse retina require DSCAM. Nature. 2008 Jan 24;451(7177):470-4. PubMed PMID: 18216855; PubMed Central PMCID: PMC2259282.

Fuerst PG, Bruce F, Tian M, Wei W, Elstrott J, Feller MB, Erskine L, Singer JH, Burgess RW. DSCAM and DSCAML1 function in self-avoidance in multiple cell types in the developing mouse retina. Neuron. 2009 Nov 25;64(4):484-97. PubMed PMID: 19945391; PubMed Central PMCID: PMC2850049.

Funada M, Hara H, Sasagawa H, Kitagawa Y, Kadowaki T. A honey bee Dscam family member, AbsCAM, is a brain-specific cell adhesion molecule with the neurite outgrowth activity which influences neuronal wiring during development. Eur J Neurosci. 2007 Jan;25(1):168-80. PubMed PMID: 17241278.

Fusaoka E, Inoue T, Mineta K, Agata K, Takeuchi K. Structure and function of primitive immunoglobulin superfamily neural cell adhesion molecules: a lesson from studies on planarian. Genes Cells. 2006 May;11(5):541-55. PubMed PMID: 16629906.

Gan WB, Macagno ER. Developing neurons use a putative pioneer's peripheral arbor to establish their terminal fields. J Neurosci. 1995 May;15(5 Pt 1):3254-62. PubMed PMID: 7751908.

Graveley BR, Kaur A, Gunning D, Zipursky SL, Rowen L, Clemens JC. The organization and evolution of the dipteran and hymenopteran Down syndrome cell adhesion molecule (Dscam) genes. RNA. 2004 Oct;10(10):1499-506. PubMed PMID: 15383675; PubMed Central PMCID: PMC1370636.

Grueber WB, Jan LY, Jan YN. Tiling of the Drosophila epidermis by multidendritic sensory neurons. Development. 2002 Jun;129(12):2867-78. PubMed PMID: 12050135.

Hattori D, Demir E, Kim HW, Viragh E, Zipursky SL, Dickson BJ. Dscam diversity is essential for neuronal wiring and self-recognition. Nature. 2007 Sep 13;449(7159):223-7. PubMed PMID: 17851526; PubMed Central PMCID: PMC2691715.

Heisenberg M, Borst A, Wagner S, Byers D. Drosophila mushroom body mutants are deficient in olfactory learning. J Neurogenet. 1985 Feb;2(1):1-30. PubMed PMID: 4020527.

Huang Y, Jellies J, Johansen KM, Johansen J. Development and pathway formation of peripheral neurons during leech embryogenesis. J Comp Neurol. 1998 Aug 3;397(3):394-402. PubMed PMID: 9674564.

Hughes ME, Bortnick R, Tsubouchi A, Bäumer P, Kondo M, Uemura T, Schmucker D. Homophilic Dscam interactions control complex dendrite morphogenesis. Neuron. 2007 May 3;54(3):417-27. PubMed PMID: 17481395; PubMed Central PMCID: PMC1963440.

Hummel T, Vasconcelos ML, Clemens JC, Fishilevich Y, Vosshall LB, Zipursky SL. Axonal targeting of olfactory receptor neurons in Drosophila is controlled by Dscam. Neuron. 2003 Jan 23;37(2):221-31. PubMed PMID: 12546818.

Jellies J, Johansen K, Johansen J. Specific pathway selection by the early projections of individual peripheral sensory neurons in the embryonic medicinal leech. J Neurobiol. 1994 Oct;25(10):1187-99. PubMed PMID: 7815053.

Jellies J, Johansen KM, Johansen J. Peripheral neurons depend on CNS-derived guidance cues for proper navigation during leech development. Dev Biol. 1995 Oct;171(2):471-82. PubMed PMID: 7556929.

Johansen KM, Kopp DM, Jellies J, Johansen J. Tract formation and axon fasciculation of molecularly distinct peripheral neuron subpopulations during leech embryogenesis. Neuron. 1992 Mar;8(3):559-72. PubMed PMID: 1550678.

Kramer AP, Kuwada JY. Formation of the receptive fields of leech mechanosensory neurons during embryonic development. J Neurosci. 1983 Dec;3(12):2474-86. PubMed PMID: 6317810.

Kramer AP, Stent GS. Developmental arborization of sensory neurons in the leech Haementeria ghilianii. II. Experimentally induced variations in the branching pattern. J Neurosci. 1985 Mar;5(3):768-75. PubMed PMID: 3973696.

Kretz JR, Stent GS, Kristan WB. Photosensory input pathways in the medicinal leech, *J. comp. Physiol. A* **106** (1976), pp. 1–37

Kristan WB Jr, Eisenhart FJ, Johnson LA, French KA. Development of neuronal circuits and behaviors in the medicinal leech. Brain Res Bull. 2000 Nov 15;53(5):561-70. Review. PubMed PMID: 11165792.

Kuwada JY. Pioneering and pathfinding by an identified neuron in the embryonic leech. J Embryol Exp Morphol. 1985 Apr;86:155-67. PubMed PMID: 4031737.

Li HL, Huang BS, Vishwasrao H, Sutedja N, Chen W, Jin I, Hawkins RD, Bailey CH, Kandel ER. Dscam mediates remodeling of glutamate receptors in Aplysia during de novo and learning-related synapse formation. Neuron. 2009 Feb 26;61(4):527-40. PubMed PMID: 19249274.

Liu G, Li W, Wang L, Kar A, Guan KL, Rao Y, Wu JY. DSCAM functions as a netrin receptor in commissural axon pathfinding. Proc Natl Acad Sci U S A. 2009 Feb 24;106(8):2951-6. Epub 2009 Feb 5. PubMed PMID: 19196994; PubMed Central PMCID: PMC2650370.

Ly A, Nikolaev A, Suresh G, Zheng Y, Tessier-Lavigne M, Stein E. DSCAM is a netrin receptor that collaborates with DCC in mediating turning responses to netrin-1. Cell. 2008 Jun 27;133(7):1241-54. PubMed PMID: 18585357; PubMed Central PMCID: PMC2491333.

Macagno ER, Muller KJ, DeRiemer SA. Regeneration of axons and synaptic connections by touch sensory neurons in the leech central nervous system. J Neurosci. 1985 Sep;5(9):2510-21. PubMed PMID: 2993546.

Matthews BJ, Kim ME, Flanagan JJ, Hattori D, Clemens JC, Zipursky SL, Grueber WB. Dendrite self-avoidance is controlled by Dscam. Cell. 2007 May 4;129(3):593-604. PubMed PMID: 17482551.

McGlade-McCulloh E, Muller KJ, Zipser B. Expression of surface glycoproteins early in leech neural development. J Comp Neurol. 1990 Sep 1;299(1):123-31. PubMed PMID: 1698836.

McKay R., Hockfield S. Monoclonal antibodies demonstrate the organization of axons in the leech. J Neurosci. 1983 Feb;3(2):369-75. PubMed PMID: 6337238.

Meijers R, Puettmann-Holgado R, Skiniotis G, Liu JH, Walz T, Wang JH, Schmucker D. Structural basis of Dscam isoform specificity. Nature. 2007 Sep 27;449(7161):487-91. Epub 2007 Aug 26. PubMed PMID: 17721508.

Mercer AR, Mobbs PG, Davenport AP, Evans PD. Biogenic amines in the brain of the honeybee, Apis mellifera. Cell Tissue Res. 1983;234(3):655-77. PubMed PMID: 6420063.

Millard SS, Flanagan JJ, Pappu KS, Wu W, Zipursky SL. Dscam2 mediates axonal tiling in the Drosophila visual system. Nature. 2007 Jun 7;447(7145):720-4.

PubMed PMID: 17554308; PubMed Central PMCID: PMC2691714.

Neves G, Zucker J, Daly M, Chess A. Stochastic yet biased expression of multiple Dscam splice variants by individual cells. Nat Genet. 2004 Mar;36(3):240-6. Epub 2004 Feb 1. PubMed PMID: 14758360.

Nicholls JG, Baylor DA. Specific modalities and receptive fields of sensory neurons in CNS of the leech. J Neurophysiol. 1968 Sep;31(5):740-56. PubMed PMID: 5711143.

Peinado A, Macagno ER, Zipser B. A group of related surface glycoproteins distinguish sets and subsets of sensory afferents in the leech nervous system. Brain Res. 1987 May 5;410(2):335-9. PubMed PMID: 3297253.

Peterson EL. Photoreceptors and visual interneurons in the medicinal leech. J Neurobiol. 1984 Nov;15(6):413-28. PubMed PMID: 6520610.

Phillips CE, Friesen WO. Ultrastructure of the water-movement-sensitive sensilla in the medicinal leech. J Neurobiol. 1982 Nov;13(6):473-86. PubMed PMID: 7175520.

Raper J, Mason C. Cellular strategies of axonal pathfinding. Cold Spring Harb Perspect Biol. 2010 Sep;2(9):a001933. Epub 2010 Jun 30. Review. PubMed PMID: 20591992.

Sawaya MR, Wojtowicz WM, Andre I, Qian B, Wu W, Baker D, Eisenberg D, Zipursky SL. A double S shape provides the structural basis for the extraordinary binding specificity of Dscam isoforms. Cell. 2008 Sep 19;134(6):1007-18. PubMed PMID: 18805093; PubMed Central PMCID: PMC2701508.

Schmucker D, Clemens JC, Shu H, Worby CA, Xiao J, Muda M, Dixon JE, Zipursky SL. Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell. 2000 Jun 9;101(6):671-84. PubMed PMID: 10892653.

Schmucker D, Chen B. Dscam and DSCAM: complex genes in simple animals, complex animals yet simple genes. Genes Dev. 2009 Jan 15;23(2):147-56. Review. PubMed PMID: 19171779.

Smith PH, Mwangi JM, Afrane YA, Yan G, Obbard DJ, Ranford-Cartwright LC, Little TJ. Alternative splicing of the Anopheles gambiae Dscam gene in diverse Plasmodium falciparum infections. Malar J. 2011 Jun 8;10:156. PubMed PMID: 21651790; PubMed Central PMCID: PMC3118162.

Soba P, Zhu S, Emoto K, Younger S, Yang SJ, Yu HH, Lee T, Jan LY, Jan YN. Drosophila sensory neurons require Dscam for dendritic self-avoidance and proper dendritic field organization. Neuron. 2007 May 3;54(3):403-16. PubMed PMID: 17481394; PubMed Central PMCID: PMC1963441.

Wang H, Macagno ER. A detached branch stops being recognized as self by other branches of a neuron. J Neurobiol. 1998 Apr;35(1):53-64. PubMed PMID: 9552166.

Wang J, Zugates CT, Liang IH, Lee CH, Lee T. Drosophila Dscam is required for divergent segregation of sister branches and suppresses ectopic bifurcation of axons. Neuron. 2002 Feb 14;33(4):559-71. PubMed PMID: 11856530.

Watson FL, Püttmann-Holgado R, Thomas F, Lamar DL, Hughes M, Kondo M, Rebel VI, Schmucker D. Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science. 2005 Sep 16;309(5742):1874-8. Epub 2005 Aug 18. PubMed PMID: 16109846.

Watthanasurorot A, Jiravanichpaisal P, Liu H, Söderhäll I, Söderhäll K. Bacteria-Induced Dscam Isoforms of the Crustacean, Pacifastacus leniusculus. PLoS Pathog. 2011 Jun;7(6):e1002062. Epub 2011 Jun 9. PubMed PMID: 21695245; PubMed Central PMCID: PMC3111544.

Weisblat DA, Harper G, Stent GS, Sawyer RT. Embryonic cell lineages in the nervous system of the glossiphoniid leech Helobdella triserialis. Dev Biol. 1980 Apr;76(1):58-78. PubMed PMID: 7380099.

Wojtowicz WM, Flanagan JJ, Millard SS, Zipursky SL, Clemens JC. Alternative splicing of Drosophila Dscam generates axon guidance receptors that exhibit isoform-specific homophilic binding. Cell. 2004 Sep 3;118(5):619-33. PubMed PMID: 15339666; PubMed Central PMCID: PMC2691713.

Yamagata M, Sanes JR. Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. Nature. 2008 Jan 24;451(7177):465-9. PubMed PMID: 18216854.

Yamakawa K, Huot YK, Haendelt MA, Hubert R, Chen XN, Lyons GE, Korenberg JR. DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. Hum Mol Genet. 1998 Feb;7(2):227-37. PubMed PMID: 9426258.

Yoshihara Y, Oka S, Ikeda J, Mori K. Immunoglobulin superfamily molecules in the nervous system. Neuroscience Res. 1991 Mar;2(10): 83-105. Review.

Zhan XL, Clemens JC, Neves G, Hattori D, Flanagan JJ, Hummel T, Vasconcelos ML, Chess A, Zipursky SL. Analysis of Dscam diversity in regulating axon guidance in Drosophila mushroom bodies. Neuron. 2004 Sep 2;43(5):673-86. PubMed PMID: 15339649.

Zhu H, Hummel T, Clemens JC, Berdnik D, Zipursky SL, Luo L. Dendritic patterning by Dscam and synaptic partner matching in the Drosophila antennal lobe. Nat Neurosci. 2006 Mar;9(3):349-55. Epub 2006 Feb 12. PubMed PMID: 16474389.

Zipser B, McKay R. Monoclonal antibodies distinguish identifiable neurones in the leech. Nature. 1981 Feb 12;289(5798):549-54. PubMed PMID: 616