

# IGSF9 Family Proteins

Maria Hansen · Peter Schledermann Walmod

Received: 2 January 2013 / Revised: 6 February 2013 / Accepted: 8 February 2013 / Published online: 16 February 2013  
© Springer Science+Business Media New York 2013

**Abstract** The *Drosophila* protein Turtle and the vertebrate proteins immunoglobulin superfamily (IgSF), member 9 (IGSF9/Dasm1) and IGSF9B are members of an evolutionarily ancient protein family. A bioinformatics analysis of the protein family revealed that invertebrates contain only a single IGSF9 family gene, whereas vertebrates contain two to four genes. In cnidarians, the gene appears to encode a secreted protein, but transmembrane isoforms of the protein have also evolved, and in many species, alternative splicing facilitates the expression of both transmembrane and secreted isoforms. In most species, the longest isoforms of the proteins have the same general organization as the neural cell adhesion molecule family of cell adhesion molecule proteins, and like this family of proteins, IGSF9 family members are expressed in the nervous system. A review of the literature revealed that *Drosophila* Turtle facilitates homophilic cell adhesion. Moreover, IGSF9 family proteins have been implicated in the outgrowth and branching of neurites, axon guidance, synapse maturation, self-avoidance, and tiling. However, despite the few published studies on IGSF9 family proteins, reports on the functions of both Turtle and mammalian IGSF9 proteins are contradictory.

**Keywords** Cell adhesion molecule · Dasm1 · igcm-2 · IGSF9 · IGSF9B · SSSD1 · Turtle

## Introduction

Cell adhesion molecules (CAMs) constitute a large group of proteins with a diverse range of functions. As the name implies, they mediate intercellular interactions and interactions between cells and the extracellular matrix (ECM). Additionally, CAMs can modulate intracellular signal transduction through interactions with other cell surface receptors or intracellular signaling molecules via their extra- and intracellular domains [1]. Consequently, CAMs are involved in various biological and pathological processes, including the survival, proliferation, differentiation, and migration of cells, organogenesis, angiogenesis, and inflammation [2–5]. Moreover, CAMs expressed in the nervous system also regulate neuronal regeneration, axon guidance, tiling, self-avoidance, and the formation and maturation of synapses [6–8].

Traditionally, CAMs have been divided into four families: cadherins, selectins, integrins, and members of the immunoglobulin (Ig) superfamily (IgSF). However, additional protein families also contain CAMs, including neu-rexins/neuroligins [9] and several subfamilies of the leucine-rich repeat family [10]. Of the traditional CAMs, cadherins, integrins, and IgSF-CAMs are evolutionarily ancient families with members that are believed to be expressed throughout *Animalia* [11, 12].

One of the first CAMs identified was the vertebrate IgSF member neural cell adhesion molecule 1 (NCAM1) that was described already in the mid-1970s [13, 14]. In 2002, a new gene, *IGSF9*, encoding a member of the vertebrate IgSF family was cloned [15] (the gene product had been identified already 2 years earlier [16]). The protein encoded by *IGSF9* was found to be a homolog of the *Drosophila* protein Turtle identified the previous year [17], and in mammals, the gene family has been found to contain one additional member, *IGSF9B* [18–20].

---

Special Issue: In Honor of Elisabeth Bock.

---

M. Hansen · P. S. Walmod (✉)  
Protein Laboratory, Department of Neuroscience and  
Pharmacology, Faculty of Health Sciences, Panum Institute,  
University of Copenhagen, Building 24.2, Blegdamsvej 3, 2200  
Copenhagen N, Denmark  
e-mail: psw@sund.ku.dk

Interestingly, members of the NCAM family and what in the following sections will be referred to as the IGSF9 family demonstrate the same overall organization of their extracellular domains [15]. However, whereas several members of the NCAM family, particularly NCAM1, have been studied in detail, the knowledge is still sparse about the members of. Nevertheless, members of the IGSF9 family have been identified throughout *Animalia*, with the exception of *Porifera*, suggesting that the proteins have important, conserved functions.

We provide a bioinformatics summary and review of the current knowledge about members of the IGSF9 family, with an emphasis on similarities and differences between IGSF9 family proteins, NCAM family proteins, and other related CAMs. We also provide a comparison of the reported vertebrate and invertebrate functions of IGSF9 proteins.

## Materials and Methods

### Amino Acid Sequences

The list below contains species names and the corresponding database IDs for the respective amino acid sequences used in the study. The gene names are indicated in brackets after the IDs. If no gene name is indicated, then the sequence corresponds to the transcript from an invertebrate *tutl* gene.

- *Acromyrmex echinator* (Panamanian leafcutter ant): GenBank ID EGI70599.1
- *Acyrtosiphon pisum* (pea aphid): NCBI RefSeq XP\_003240846.1
- *Ascaris suum* (pig roundworm): GenBank ID ADY42573.1
- *Bos Taurus* (cow): NCBI RefSeq NP\_001192461.1; GenBank ID DAA31922.1 (*IGSF9*); UniProtKB/TrEMBL ID F1ML81 (*IGSF9B*)
- *Caenorhabditis elegans* (roundworm): NCBI RefSeq NP\_509073.3
- *Camponotus floridanus* (Florida carpenter ant): GenBank ID EFN60857.1
- *Canis lupus familiaris* (dog): NCBI RefSeqs XP\_545751.4 and XP\_003434369.2 (*IGSF9*); UniProtKB/TrEMBL ID F1P9I8 (*IGSF9B*)
- *Cavia porcellus* (guinea pig): NCBI RefSeqs XP\_003466605.1 and XP\_003466606.1 (*Igsf9*)
- *Ciona intestinalis* (vase tunicate): A combination of Ensembl transcript ID ENSCINT00000010210 (amino acids 1-80) and NCBI RefSeq XP\_002122656.1 (amino acids 81-322 of ENSCINT00000010210 correspond to amino acids 1-242 of XP\_002122656.1)
- *Cricetulus griseus* (Chinese hamster): NCBI RefSeq XP\_003500305.1 (*Igsf9*); UniProtKB/TrEMBL ID H0UV25 (*Igsf9B*)
- *Danio rerio* (zebrafish): NCBI RefSeq XP\_685629.2 (*Igsf9a*); GenBank ID CAQ13296.1 (*Igsf9b*); NCBI RefSeq XP\_695776.3 (*Igsf9Ba*); NCBI RefSeq XP\_698041.4 (*Igsf9Bb*)
- *Drosophila melanogaster* (common fruit fly): Flybase IDs FBpp0291078, FBpp0305668, FBpp0088327, FBpp0088326, FBpp0291077, FBpp0088328, FBpp0305669, and FBpp0088325 (*tutl*); UniProtKB/TrEMBL ID Q9U4G1 (*BcDNA.GH11322*)
- *Gallus gallus* (chicken): NCBI RefSeq XP\_423214.3 (*IGSF9*); UniProtKB/TrEMBL ID F1NSC3 (*IGSF9B*)
- *Homo sapiens* (human): NCBI RefSeqs NP\_001128522.1 and NP\_065840.2; GenBank IDs AAQ88495.1 and EAW52756.1 (*IGSF9*); UniProtKB/Swiss-Prot ID Q9UPX0 (*IGSF9B*); UniProtKB/Swiss-Prot ID P13591.3 (*NCAM1*); NCBI RefSeq NP\_004531.2 (*NCAM2*)
- *Hydra magnipapillata* (freshwater polyp): NCBI RefSeq XP\_002160325.1
- *Ixodes scapularis* (black-legged tick): NCBI RefSeq XP\_002401962.1
- *Mus musculus* (mouse): NCBI RefSeqs NP\_001139272.1; GenBank IDs AAL37164.1 and EDL39012.1 (*Igsf9*); UniProtKB/TrEMBL IDs E9PZ19, D3Z169, and D3Z1P3 (*Igsf9b*)
- *Nematostella vectensis* (starlet sea anemone): UniProtKB/TrEMBL ID A7TDK3 (*v1g225659*)
- *Oryzias latipes* (Japanese medaka): UniProtKB/TrEMBL IDs H2M9R6 (*Igsf9*, 1 of 2), H2M9X3 (*Igsf9*, 2 of 2), H2L3M5 (*Igsf9B*, 1 of 2), and H2LKH1 (*Igsf9B*, 2 of 2)
- *Rattus norvegicus* (brown rat): NCBI RefSeq NP\_001100667.1 (*Igsf9*); UniProtKB/TrEMBL ID D3ZB51 (*Igsf9b*)
- *Strongylocentrotus purpuratus* (purple sea urchin): NCBI RefSeq XP\_788219.3
- *Taeniopygia guttata* (zebra finch): UniProtKB/TrEMBL IDs H1A0X2 (*Igsf9*) and H0YQX1 (*Igsf9B*)
- *Xenopus tropicalis* (Western clawed frog): NCBI RefSeq NP\_001090640.1 (*Igsf9*); UniProtKB/TrEMBL ID F7CTP9 (*Igsf9B*).

### Programs Used for Bioinformatics Analysis of Amino Acid Sequences

Alignments and determinations of amino acid identities were performed using ClustalW2 [21]. Phylogenetic relationships were calculated using ClustalW2—Phylogeny [22–24] and visualized using Treeview 1.6.6 [25, 26] in combination with CoreIDRAW 11.0 (Corel Corporation).

Domains and motifs were located using the Eukaryotic Linear Motifs (ELM) database [27, 28]). Additionally, the identification of signal peptides, Ig modules, fibronectin 3 (Fn3) modules, and transmembrane helices was made using

the Simple Modular Architecture Research tool (SMART; [29, 30]). Estimates of signal peptides and transmembrane helices were performed using SignalP 4.0 [31] and TMHMM Server v. 2.0 [32].

Estimates of *N*-linked glycosylation sites were performed using NetNGlyc 1.0 [33]. Estimates of *O*-linked glycosylation sites were performed using NetOGlyc 3.1 [34]. Potential palmitoylation and SUMOylation sites were identified using CSS-Palm 3.0 and SUMOsp 3.0, respectively [35, 36]. Potential prenylation sites were identified using Prenylation Prediction Suite (PrePS; [37, 38]). Illustrations of domain organizations were obtained using Prosite MyDomains Image Creator [39]. The structural modeling of three-dimensional protein structures was performed using GENO3D [40, 41].

## Results

### Evolutionary Relationships Between IGSF9 Family Proteins

In mammals, the IGSF9 family consists of two members, IGSF9 and IGSF9B [43], sometimes referred to as Protein Turtle homolog A and B, respectively (see e.g. UniProt IDs Q9P2J2 and Q9UPX0). Sequences for these proteins exist for several mammalian species, some of which were utilized to prepare Fig. 1. Additionally, orthologs of the mammalian members of the IGSF9 family have been identified in several non-mammalian vertebrates and invertebrates. Birds [e.g., chicken (*Gallus gallus*) and zebra finch (*Taeniopygia guttata*)] and amphibians [e.g., frog (*Xenopus tropicalis*)] also have two genes, whereas bony fish have four genes that encode IGSF9 family proteins. In zebrafish (*Danio rerio*), the four genes are *Igsf9a* (located on chromosome 10), *Igsf9b* (located on chromosome 8), *Igsf9Ba* (located on chromosome 15), and *Igsf9Bb* (located on chromosome 21).

The *Drosophila* ortholog of the vertebrate IGSF9 family genes is *tutl*, encoding the protein Turtle. Another *Drosophila* gene, *BcDNA.GH11322*, encodes a protein with a high degree of sequence homology and structural organization similarities to Turtle (see below). *BcDNA.GH11322* is located only 1.2 kb downstream of *tutl*, suggesting that the genes originated by tandem duplication [17].

Interestingly, vertebrate IGSF9 family genes have an invertebrate ortholog not only in insects but also in several other invertebrates, including cnidarians (e.g., *Hydra magnipapillata* and *Nematostella vectensis*). In *Hydra*, the IGSF9 ortholog is one of only 12 proteins that contain both Ig and Fn3 homology modules; other members including orthologs to DSCAM, Roundabout 2, and TAG-1/axonin 1 (NCBI RefSeqs XP\_002169672.1, XP\_002166740.1, and

XP\_002154651.1, respectively) [44, 45]. This makes the IGSF9 family not only an evolutionarily ancient family but also one of the most ancient families of IgSF CAMs (although the protein may in fact not be a CAM in cnidarians; see below).

From the description above it is clear that the IGSF9 family contains one member in invertebrates, two members in mammals, and four members in bony fish (e.g., *Danio rerio*). This is paralleled by the evolution of NCAM family genes, which also consists of one member in invertebrates (Fasciclin II encoded by *Fas2* [46]), two members in mammals (NCAM1 and NCAM2/OCAM/RNCAM encoded by *NCAM1* and *NCAM2*, respectively), and four members in bony fish (in *Danio rerio* encoded by *ncam1a*, *ncam1b*, *ncam2*, and *zgc:152904* [47]). These differences in gene numbers are generally explained by repeated whole-genome duplications. Thus, the increased number of genes in bony fish is believed to be the result of a fish-specific genome duplication, the so-called 3R [48]. Consequently, vertebrate NCAM and IGSF9 family genes are ohnologs, i.e. paralogs generated by whole-genome duplication [49], which is in accordance with the fact that the genes are located on separate chromosomes (see [46, 47] and the *Danio rerio* genes mentioned above).

Figure 2 shows an unrooted phylogenetic tree of selected members of the IGSF9 family. Invertebrate Turtle-expressing genes are shown on the right side of the figure. The division of vertebrate genes that encode IGSF9 and IGSF9B proteins, respectively, can be seen on the left side of the figure. Moreover, *Danio rerio Igsf9a* and *Igsf9b* are evidently orthologs of human *IGSF9*, whereas *Danio rerio Igsf9Ba* and *Igsf9Bb* are orthologs of human *IGSF9B*.

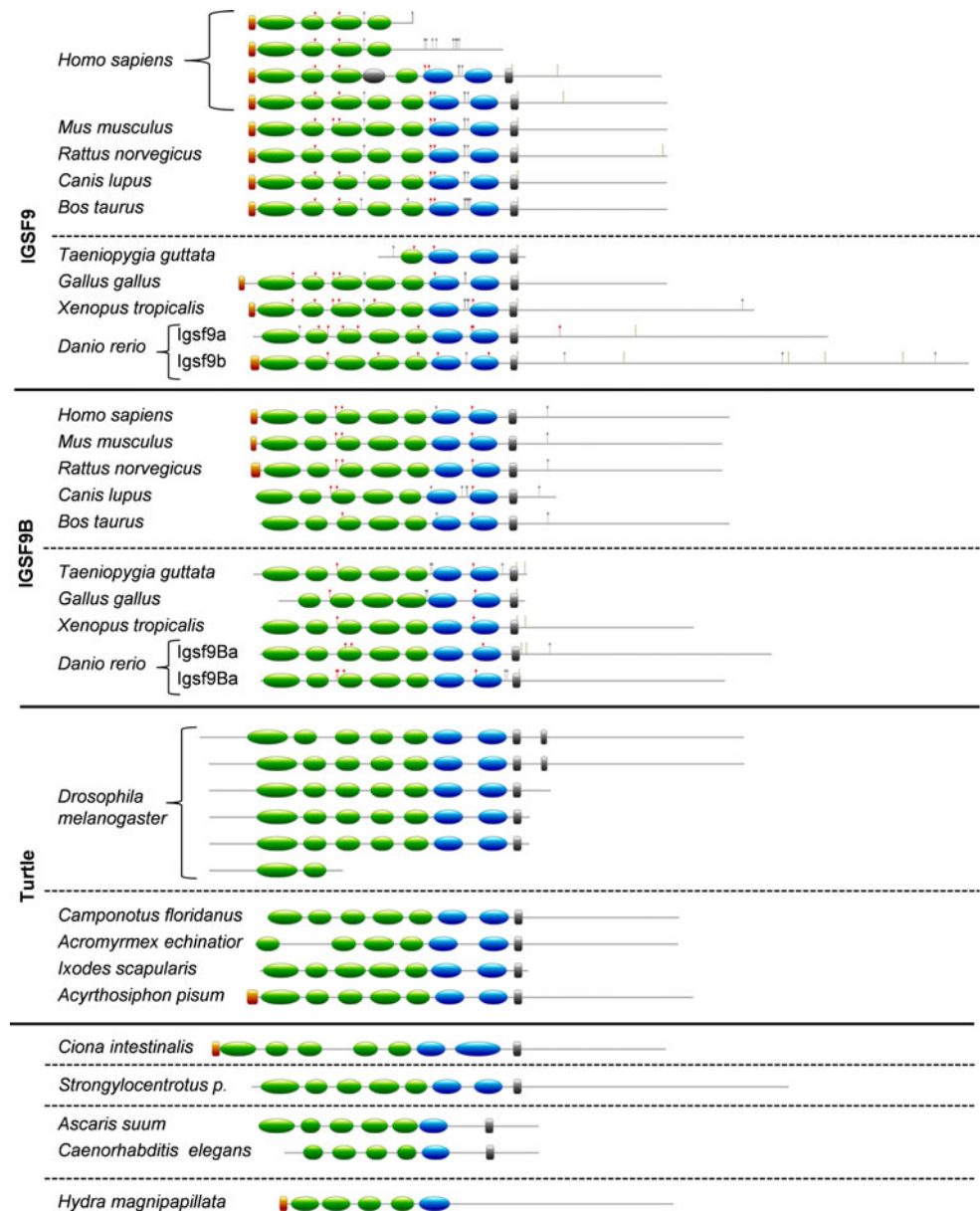
The IGSF9 family ortholog that is present in the nematode *Caenorhabditis elegans* is called SSSD1 and encoded by the gene *igcm-2* [50, 51], whereas the product of the corresponding gene in another nematode, *Ascaris suum*, is called Turtle. In *C. elegans*, the gene constitutes one of only four members of the immunoglobulin-like cell adhesion molecule (IGCM) gene family [52], further demonstrating the evolutionarily ancient origin of the IGSF9 family.

### Structural Organization of IGSF9 Family Proteins

IgSF proteins, by definition, contain one or several Ig modules. Additionally, IgSF CAMs often contain numerous Fn3 modules. The two types of modules are both composed of a  $\beta$ -sandwich formed by two layers of anti-parallel  $\beta$ -sheets. In contrast to the typical Ig modules, however, the  $\beta$ -sheets of Fn3 modules are not connected by a disulfide bond [53, 54].

Most members of the IGSF9 family, similar to members of the NCAM family, contain an extracellular domain that is

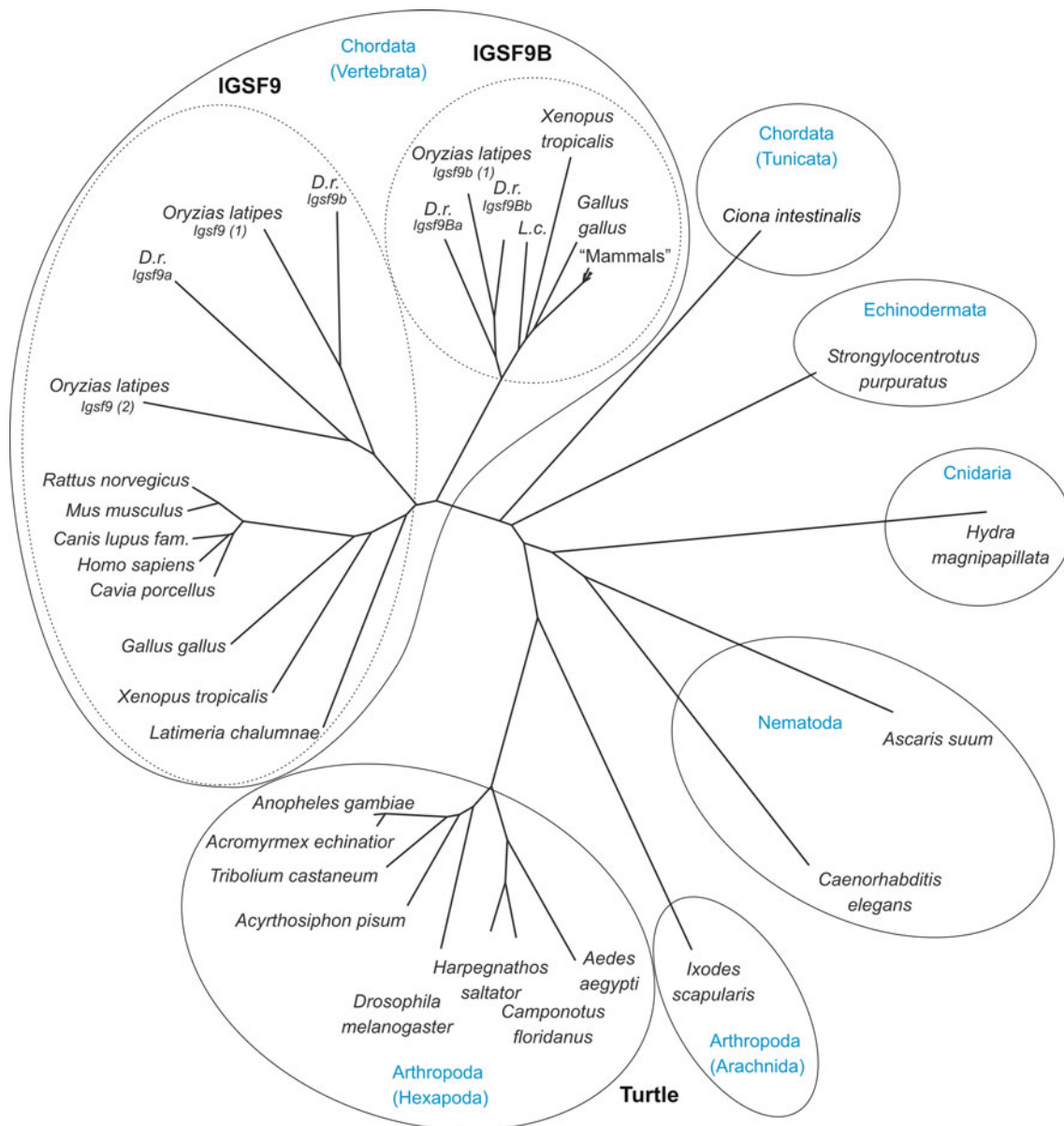
**Fig. 1** Structural organization and potential posttranslational modifications of selected IGSF9 family proteins. Domains and posttranslational modifications were identified as described in the “Materials and Methods”. Orange boxes indicate signal peptides. Green and blue spheres indicate Ig and Fn3 modules, respectively. The gray sphere corresponds to the IgIV module of human IGSF9 that lacks  $\beta$ -strand A and therefore is not recognized by SMART or ELM as an Ig module. Gray boxes indicate transmembrane domains. In the extracellular domains, red and gray markers indicate potential N- and O-linked glycosylations, respectively. In the intracellular domains, gray markers indicate potential SUMOylation sites, whereas gray lines indicate potential palmitoylation sites. Potential posttranslational modifications are only indicated for vertebrate proteins. Some of the presented structures may be made from sequences that are only fragments of a transcript



composed of up to five N-terminal Ig modules (denoted IgI-V below), followed by up to two Fn3 modules (denoted Fn3I and Fn3II below). The extracellular domains are typically followed by a single transmembrane segment and cytoplasmic domain [15]. Human IGSF9 and IGSF9B have 737 and 726 amino-acid-long extracellular regions, respectively, followed by a transmembrane region of 23 amino acids and cytoplasmic domains of 419 and 600 amino acids, respectively. However, as described below, some IGSF9 family members may diverge from this general organization.

As shown in Fig. 1, *Hydra* IGSF9 contains only four Ig modules and a single Fn3 module. The protein contains a signal peptide but does not appear to contain a transmembrane domain, suggesting that it is secreted. IGSF9 proteins from all other investigated species contain a

transmembrane domain, although alternative splicing may generate secreted isoforms (see below). Nematode IGSF9 proteins contain five Ig modules but apparently still only a single Fn3 module, whereas IGSF9 proteins from other species investigated contain five Ig modules and two Fn3 modules. In *C. elegans*, the protein only appears to contain four Ig modules (Fig. 1), but the published sequence is likely to be incomplete, since the signal peptide also appears to be absent. Likewise, the sequences of some of the other genes investigated may be incomplete, and it is possible that additional exons have been overlooked in the original cloning of the respective sequences. Sequence alignments reveal that the single Fn3 module of cnidarian and nematode IGSF9 proteins is homologous to the Fn3I module of other IGSF9 proteins, whereas the four Ig



**Fig. 2** Unrooted phylogenetic *tree* of IGSF9 family proteins. The unrooted *tree* shows the relationships between selected IGSF9 family proteins. Different phyla and sub-phyla are indicated by circles, with the phyla shown in *blue*. The groups that correspond to the genes that encode IGSF9, IGSF9B, and Turtle proteins are shown in *bold*. The amino acid sequences utilized for the initial alignments included a region that stretches from the beginning of IgII to the end of Fn3I (corresponding to amino acids 156–593 of human IGSF9 [NP\_001128522.1]). The names of individual genes that encode the

respective *Danio rerio* and *Oryzias latipes* proteins are indicated. Similar to *Danio rerio*, *Oryzias latipes* also has four separate genes that express IGSF9 proteins. However, for one of these proteins, the available sequence fragment did not include the region utilized for the alignments; consequently, only three *Oryzias latipes* proteins appear in the figure. “Mammals” indicates refers to the same mammalian species presented for the IGSF9 proteins relative to the IGSF9B proteins. Dr, *Danio rerio*; Lc, *Latimeria chalumnae*

modules of cnidarian IGSF9 are homologous to IgII–IgV of IGSF9 proteins from other species.

The *tutl*-related *Drosophila* gene *BcDNA.GH11322* mentioned above encodes a 719-amino acid-long protein that contains four Ig modules [homologous to IgI, IgIII, IgIV, and IgV of Turtle (amino acid sequence identities of 40 and 47 %, respectively)], two Fn3 modules (amino acid

sequence identity of 26 % compared with Turtle), a transmembrane domain, and a short cytoplasmic tail.

As shown in Table 1, IGSF9 and IGSF9B have high module-to-module similarity in their extracellular domains, supporting the interpretation that the proteins are paralogs. However, despite their structural similarities, such module-to-module identity does not exist between NCAM and

IGSF9 family proteins. For example, the IgI module of human NCAM1 only exhibits amino acid sequence identities of 6 and 7 % compared with the IgI modules of human IGSF9 and IGSF9B, respectively, whereas it exhibits sequence identities of 18 and 16 %, respectively, compared with the IgIV modules of human IGSF9 and IGSF9B, respectively. Thus, despite the structural similarities, NCAM family proteins are not paralogs of IGSF9 family proteins. Instead, the structural similarities of the two protein families are believed to be generated by convergent evolution [17].

A comparison of the extracellular domains of human IGSF9 and IGSF9B with other proteins revealed that the IgI–IgV modules resemble regions of the large proteins titin (NCBI RefSeq NP\_001243779), hemicentin-1 (NCBI RefSeq NP\_114141.2), and hemicentin-2 (NCBI RefSeq XP\_001715206.1), which all contain numerous Ig modules. Of more interest, the modules resemble regions of DSCAM and DSCAML1 (NCBI RefSeqs NP\_001380.2 and NP\_065744.2, respectively), the SLIT receptors ROBO1, -2, and -3 (NCBI RefSeqs NP\_002932.1, NP\_001122401.1, and NP\_071765.2, respectively), which also contain five Ig modules (overall amino acid sequence identities of 18–21 %), and NCAM1/2 (NCBI RefSeqs P13591 and NP\_004531.2, respectively; overall amino acid sequence identities of 17–20 %).

The Fn3I modules of human IGSF9 and IGSF9B resemble the single Fn3 module of the kidney-specific CAM nephrin (NCBI RefSeq NP\_004637.1; amino acid sequence identities of 28 and 24 %, respectively), the fifth of the six Fn3 modules of DSCAM (NCBI RefSeq NP\_001380.2, 22 and 24 %, respectively), and the second Fn3II module of the synaptic ECM protein pikachurin (NCBI RefSeqNP\_689616.2; 21 and 26 %, respectively).

The nuclear magnetic resonance structure of the Fn3II module of human IGSF9 has been released (PDB ID:

1V5J), but no accompanying biochemical or biological studies have yet been published. This module and the corresponding module from human IGSF9B resemble the third Fn3 module of the netrin receptor neogenin (NCBI RefSeqNP\_002490.2; amino acid sequence identities of 21 and 17 %, respectively) and the second Fn3 module of the ubiquitously expressed CAM brother of CDO (BOC; NCBI RefSeqNP\_150279.1; 15 and 21 %, respectively).

The aforementioned similarities between the extracellular modules of IGSF9 proteins and other proteins suggest that IGSF9 proteins might be engaged in protein–protein interactions in the nervous system that are involved in cell adhesion and differentiation, axon guidance, tiling, self-avoidance, and synapse formation [55–59]. As described below, IGSF9 family proteins may indeed be involved in such processes.

### Motifs and Posttranslational Modifications

A comparison of the mammalian forms of IGSF9 and IGSF9B revealed that IGSF9 is potentially more glycosylated than IGSF9B. IGSF9 contains conserved sites for potential *N*-linked glycosylation in IgII and Fn3II and conserved sites for *O*-linked glycosylation in the IgIII–IgIV and Fn3I–Fn3II linker regions, whereas the IgI, IgIV, IgV, and Fn3II modules are devoid of any glycosylations. Mammalian forms of IGSF9B have fewer sites for potential glycosylation than IGSF9. They have conserved sites for *N*-linked glycosylation in IgIII and Fn3II but no conserved sites for *O*-linked glycosylation. IgI, IgII, IgIV, IgV, and FnI are devoid of glycosylation sites (Fig. 1).

Glycosylation on CAMs is known to be involved in a range of biological and pathological processes, including the adhesion, growth, migration, and differentiation of cells, cancer invasion and metastasis, synaptogenesis, and neurodegeneration [60–62]. The functions of IGSF9 family

**Table 1** Comparison of primary sequence identities between *H. sapiens* Igsf9, *H. sapiens* Igsf9B, and *D. melanogaster* Turtle

Regions compared	Entire protein	Extracellular domain							TM domain	Cytoplasmic domain
		Ig1	Ig2	Ig3	Ig4	Ig5	Fn3I	Fn3II		
Amino acids <sup>b</sup>	20–1,179	20–133	134–224	225–320	321–415	416–504	505–604	605–737	738–760	761–1,179
Compared protein sequences <sup>a</sup>										
Igsf9 versus Igsf9B	38 %	50 %							47 %	22 %
		54 %	45 %	59 %	50 %	48 %	58 %	37 %		
Igsf9 versus Turtle	25 %	31 %							43 %	12 %
		35 %	28 %	42 %	41 %	26 %	30 %	17 %		
Igsf9B versus Turtle	22 %	33 %							34 %	10 %
		41 %	32 %	42 %	33 %	30 %	37 %	19 %		

<sup>a</sup> Primary sequences were from *H. sapiens* IGSF9 (UniProt: Q9P2J2), *H. sapiens* Igsf9B (UniProt: Q9UPX0), and *D. melanogaster* Turtle (UniProt: Q967D7)

<sup>b</sup> Amino acid numbers refer to the sequence of Igsf9, including the signal peptide. The sequences were aligned pairwise using ClustalW2 [21]

glycosylation are unknown, but specifically the region that contains the conserved site for *O*-linked glycosylation between IgIII and IgIV in mammals is absent in one alternatively spliced isoform of IGSF9 (see below).

IGSF9 from mouse in the IgV module was previously shown to contain an arginine-glycine-aspartate (RGD) motif, which is known from integrins to facilitate adhesion to ECM components [63]. However, the RGD motif is not conserved between IGSF9 proteins from different species and, therefore, may not play any functional role.

The intracellular domain of IGSF9 from all investigated vertebrate species contains a conserved site for potential palmitoylation proximal to the transmembrane region. This potential palmitoylation site is missing from vertebrate forms of IGSF9B, which instead contains a conserved site for potential SUMOylation. Neither of the proteins contains any sites for potential prenylation.

Palmitoylation is the post-translational attachment of a palmitic acid group onto cysteine residues. The modification facilitates the attachment of the palmitoylated proteins to membranes. In contrast to prenylation and myristoylation, the process is reversible and regulated by both palmitoylating and depalmitoylating enzymes [64].

Many proteins that are involved in neuritogenesis and synaptic spine formation have been shown to be palmitoylated. In the case of transmembrane proteins like CAMs, palmitoylation has several potential effects. For example, it may affect the subcellular localization of proteins that target them to lipid rafts [64]. This has been demonstrated for NCAM1, which is palmitoylated in hippocampal neurons in response to fibroblast growth factor 2 (FGF2)-mediated signaling and subsequently translocated to lipid rafts. Moreover, palmitoylation may alter the conformation of a protein and therefore its activity and ability to bind other proteins [64]. For example, NCAM1-mediated neurite outgrowth facilitated by FGF receptor (FGFR) activation has been shown to be inhibited when the palmitoylation of NCAM1 is prevented [65]. Moreover, the palmitoylation of NCAM1 has been shown to inhibit the *cis*-homodimerization of the protein (interactions between NCAM1 molecules located on the same cell surface [66]). Many CAMs, including NCAM1, are believed to form zipper-like structures generated by a combination of homophilic *cis*- and *trans*-interactions (i.e. interactions between molecules on opposing cell surfaces) [67, 68]. Therefore, the inhibition of CAM *cis*-dimerization may subsequently affect cell–cell interactions. The fact that IGSF9 but not ISF9B is potentially palmitoylated suggests that the two proteins, at least under certain conditions, may have different subcellular localizations, and their activities may be differentially regulated.

The effects of the SUMOylation (i.e., the covalent attachment of ubiquitin-like proteins to lysine residues) of

cell surface receptors have not been well-studied, but the modification is known to be able to induce endocytosis (e.g., of the GluR6 receptor; [69]).

As previously reported, human IGSF9 contains a C-terminal PDZ-binding motif [70]. This motif is conserved between all mammalian species investigated and is also found in *Xenopus*. Human IGSF9B also contains a C-terminal PDZ-binding motif. In IGSF9B, however, the motif is not conserved between species. The lack of a conserved PDZ-binding motif in IGSF9B implies that IGSF9 and IGSF9B have partially different cytoplasmic binding partners, suggesting that IGSF9B does not have the same potential effects on synapse maturation as IGSF9 (see below).

IGSF9 family proteins contain a large number of potential sites for various types of phosphorylation, including the so-called immunoreceptor tyrosine-based inhibitory motif (ITIM; [ILV]-x-x-Y-x-[LV] [24]). Tyrosine phosphorylation at ITIMs in the cytoplasmic tail of immune receptors contributes to the regulation of immune cell activation. However, ITIM-regulated signaling has been hypothesized to not be restricted to immune cells. Using a new database search strategy, the number of human transmembrane receptors with potential ITIMs was reduced to 109 proteins, including DSCAM and IGSF9 [71]. The potential ITIM in IGSF9 (amino acids 908–913 of NP\_065840.1) is present in all mammalian species investigated in this study, as well as in chicken and frog. However, whether the potential ITIM in IGSF9 has functional significance remains to be determined.

### Isoforms

Only the full-length mammalian isoforms of IGSF9 have been investigated *in vivo* [15, 70, 72]. However, the alternative splicing of the transcripts appears to generate several mammalian isoforms of IGSF9. As shown in Fig. 1, four different human transcripts have been reported: two membrane-attached isoforms and two secreted isoforms. The two secreted isoforms both consist of IgI–IgIV, followed by a C-terminal tail of variable length and degree of potential *O*-linked glycosylation. When comparing the two membrane-attached isoforms, a 17-amino-acid-long sequence in one isoform that corresponds to the linker-region between IgIII and IgIV and  $\beta$ -strand A of IgIV is replaced by a single cysteine residue. Consequently, the conserved site for potential *O*-linked glycosylation between IgIII and IgIV is missing. Moreover, the difference must cause pronounced differences in the arrangement of the two Ig modules (if at all, the IgIV module is folded when the “A”  $\beta$ -strand is missing). It is tempting to speculate that the short isoform may not exist, and that the published sequence may be the result of an artifact. However, similar isoforms have been reported for several other mammals, including cow, dog, and

guinea pig (GenBank ID DAA31922.1, and NCBI RefSeqs XP\_003434369.2 and XP\_003466606.1, respectively), arguing against this explanation.

No alternatively spliced isoforms have been published for the vertebrate isoforms of IGSF9B [18, 73].

Insect Turtle proteins have also been reported to exist in multiple isoforms [17, 51]. According to the Flybase [74, 75], the *Drosophila* Turtle protein exists in six different isoforms (see Fig. 1), including five transmembrane isoforms and one secreted isoform that consists of only IgI and IgII. The secreted isoform and longest transmembrane isoform have a 25-amino-acid-long insert after IgII or between IgII and IgIII, respectively. Apart from this difference, all five transmembrane isoforms have identical N-terminal extracellular domains. Three of the transmembrane isoforms have a single transmembrane domain followed by 23- to 84-amino-acid-long cytoplasmic domains. The remaining two transmembrane isoforms each appear to have two transmembrane regions separated by a 56-amino-acid-long region, followed by a 555-amino-acid-long C-terminal domain. The C-terminal domains do not appear to possess any globular domains, suggesting that they are intracellular domains, despite the apparent presence of two transmembrane regions.

In addition to the *Drosophila* isoforms mentioned in the Flybase, other isoforms have been reported. For example, Al-Anzi et al. [51] reported the expression of two additional secreted isoforms that consist of IgI–Fn3I and three additional transmembrane isoforms with ectodomains that consist of IgII–Fn3II, IgIII–Fn3I, and IgIV–Fn3II, respectively.

As a result of alternative splicing, members of the NCAM family and several other IgSF CAMs exist not only as transmembrane proteins but also as isoforms in which the extracellular domain is attached to the plasma membrane with a glycosylphosphatidylinositol anchor [76]. However, such isoforms have not been reported for IGSF9 family proteins.

#### Expression of IGSF9 Family Proteins

##### *Expression of C. elegans SSSD1*

According to the Wormbase, *igcm-2* promoter::green fluorescence protein (GFP) experiments, in which GFP is expressed from the promoter of the gene that encodes the *C. elegans* homolog of Turtle, SSSD1, suggest that the protein is expressed in the pharynx and intestine, some head neurons, and paraventricular thalamic interneurons of the preanal ganglion [52].

##### *Expression of Drosophila Turtle*

The expression of Turtle in *Drosophila* has been investigated in detail in several studies [17, 51, 77–79]. Briefly,

the expression of Turtle is detectable from embryonic stage 9 and persists throughout the embryogenesis, larval, pupal, and adult stages. Transcripts from *tutl* have predominantly been detected in the nervous system and primarily in the central nervous system (CNS). At stages 12–13, *tutl* is expressed close to the midline, but the expression spreads throughout the CNS later in development [51]. However, it is also expressed in a limited group of sensory neurons in the peripheral nervous system (PNS) [17]. For example, in larvae, Turtle is expressed in the cell bodies and along the dendrites of dendritic arborization neurons of the PNS. The effects of the protein on the morphology and organization of these neurons have been studied in detail [78, 79] (see below). During embryogenesis, Turtle is expressed throughout the CNS, but the expression becomes restricted to defined regions and subsets of neurons in the adult brain. For example, strong expression is seen in the superior lateral protocerebrum [17], a region that receives inputs from the antennal lobes directly via projection neurons or indirectly via mushroom bodies. This region is believed to be involved in experience-independent olfactory responses and male courtship behavior [80]. In the visual system, the protein is expressed at axon terminals of R7 photoreceptor neurons in the medulla neuropil (see below) and in non-R7 neurons in the medulla neuropil and inner optic lobe [77].

##### *Expression of Mammalian IGSF9 Proteins*

In mammals, the expression of IGSF9 has been investigated in mouse, rat, and human tissues [15, 43, 70, 72]. Using real-time polymerase chain reaction (RT-PCR), *Igsf9* transcripts were originally detected during mouse embryogenesis from embryonic stage 7.5 to 16.5, whereas no expression was detected in adult animals [15]. However, the gene was later shown to be expressed in adult animals, but the level of expression decreases postnatally following high expression during mid- and late gestation [43]. In situ hybridization studies revealed that the gene is expressed in many regions of the CNS during embryogenesis, but with particularly high expression in the dorsal root ganglia, olfactory epithelium, trigeminal ganglia, and hindbrain neuroepithelium. Less pronounced expression is found in the cranial ganglia, retina, and ventricular zones of the forebrain neuroepithelium. *Igsf9* is also expressed in the spinal cord; In the upper spine, it is predominantly expressed in the dorsal-most neuroepithelium; In the caudal spine, the expression shifts to a ventral location. Moreover, *Igsf9* expression is detectable in limb buds and the hindgut [15].

Northern blot results have demonstrated the expression of three different *Igsf9* transcripts in mouse brain, in which the longest transcript (4.2 kb) is the predominant form [70]. In the brains of newborn mice, *Igsf9* exhibits broad



expression. In day 18 mouse brain, it is highly expressed in the forebrain and intermediate and ventricular zone. In the adult brain, the expression becomes more restricted. In particular, strong expression has been detected in the CA1 area and dentate gyrus of the hippocampus, regions known to be involved in learning and memory formation [81] and where the protein is located in cell bodies and dendrites and exhibits high expression in the postsynaptic densities (PSDs) of excitatory synapses [43, 70, 72]. Moreover, the protein is expressed in cell bodies and dendrites but not axons of adult cerebellar Purkinje cells [70].

In humans, the expression of *IGSF9* has been examined by RT-PCR in selected tissues from 8- to 14-week-old embryos. The gene appears to have broad expression and is detectable in the heart, placenta, eye, brain, umbilical cord, intestine, and kidney. The expression is most apparent in the eye, brain, and intestinal tissues [15].

*IGSF9B* has by in situ hybridization been shown to be expressed in the hippocampus [43], but a thorough description of *IGSF9B* expression has not been published.

Interestingly, the expression of *Drosophila* Turtle is stimulated by the transcription factor Cut [79], a protein known to be involved in the regulation of dendritic branching [82]. The mammalian homologs to *Drosophila* Cut are *Cux1* and *Cux2*, which are also important for the regulation of neuronal morphology and function. For example, they are reported to control dendritic branching and the formation of dendritic spines and synapses [83, 84]. However, whether mammalian *IGSF9* family genes are the targets of Cux proteins has not been determined.

#### Abundance of *IGSF9* Proteins

*IGSF9* family proteins are not abundantly expressed. Available data are limited, but according to PaxDb, both *Drosophila* and mammalian members of the *IGSF9* family belong to the bottom 5 %, in which proteins are ranked according to the level of expression. *Drosophila* Turtle has an expression level (total organism) of approximately 0.48 ppm. There are no available data on mammalian *IGSF9*, but the expression level for *IGSF9B* is 1.03 ppm in mouse brain and 0.06 ppm in human (total organism). For comparison, the reported whole-organism expression levels of *Drosophila* FasII and human NCAM1 and NCAM2 are 19.1, 20.7, and 3.05 ppm, respectively [85].

#### Protein Interactions Mediated by *IGSF9* Family Proteins

As mentioned above, the sequence homologies between *IGSF9* family proteins and NCAM and ROBO family proteins, DSCAM, and neogenin suggest that *IGSF9* family proteins are involved in extracellular protein interactions

that lead to cell adhesion or the modulation of intracellular signal transduction. Neogenin is a receptor for netrin and repulsive guidance molecules (RGMs) [86]. ROBO proteins are receptors for SLIT proteins [87], and DSCAM and NCAM family proteins form *trans*-homophilic interactions and several other extracellular protein interactions [47, 88, 89]. However, extracellular protein interactions mediated by *IGSF9* family proteins have not been investigated in detail. The potential receptors for *IGSF9* family proteins are unknown, and the intracellular signal transduction pathways regulated by *IGSF9* family proteins remain to be characterized. Nevertheless, the transient expression of transmembrane *Drosophila* Turtle in *Drosophila* Schneider 2 cells has been shown to lead to the aggregation of Turtle-expressing cells, whereas Turtle-expressing cells do not form aggregates with untransfected cells. This demonstrates that Turtle indeed is a CAM that is able to mediate homophilic *trans*-interactions [77]. Whether vertebrate *IGSF9* and *IGSF9B* are also engaged in homophilic *trans*-interactions remains to be determined.

Intracellular interactions mediated by *IGSF9* proteins have only been investigated for mammalian forms of *IGSF9*, in which yeast two-hybrid screening and subsequent immunoprecipitation experiments demonstrated that the conserved PDZ-binding domain in the C-terminal tail of *IGSF9* in mammals interacts with the proteins Shank and synaptic scaffolding molecule (S-SCAM)/membrane-associated guanylate kinase inverted 2 (MAGI-2) but not PSD-95, GRIP, or SAP102 [72]. Only the 60 most C-terminal amino acids were used as bait in the screen. Therefore, the two identified proteins may only reflect a fraction of the intracellular binding partners of *IGSF9*. Both S-SCAM and Shank are multidomain scaffolding proteins that are enriched at synapses in the nervous system, where they facilitate the formation of multimeric protein complexes. The expression of S-SCAM has been shown to correlate with the level of  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid (AMPA) receptors at the cell surface, AMPA receptor-mediated synaptic transmission, and the size of dendritic spines [90]. Shank constitutes a family of proteins that is highly expressed at the PSD, and increased Shank levels have been shown to facilitate postsynaptic maturation and spine formation (reviewed in [91]).

#### Functions of *IGSF9* Family Proteins

Of the many *IGSF9* family proteins described above, only *Drosophila* Turtle proteins and human and mouse *IGSF9* have been investigated in detail. Virtually nothing is known about *IGSF9B*. Surprisingly, despite the scarcity of studies, the published data on the function of Turtle and *IGSF9* are somewhat contradictory.

### Effects of IGSF9 Family Proteins on Behavior

*Drosophila tutl* is an essential gene. Bodily et al. generated a number of *tutl* mutants and demonstrated that the complete absence of Turtle expression results in postembryonic lethality during late pupation. At the larval stage, the *tutl* mutants are unable to perform normal escape responses and exhibit a reduced ability to return to an upright position when placed in an inverted position (hence, the name “Turtle”). Moreover, whereas mutant larvae move by peristaltic crawling, moving almost as fast as wildtype animals, they demonstrate a marked reduction of directional persistence. Some *Drosophila tutl* mutants are able to survive to adulthood. These flies are able to walk but unable to fly, although they still possess wings and the muscles and ability to move them. Altogether, these observations demonstrate that the absence of or mutations in *tutl* lead to defects in coordinated motor control [17].

Consistent with its importance for motor control, the expression of Turtle has been found to be upregulated by 15 % in migrating, juvenile hormone-deficient North American monarch butterflies (*Danaus plexippus*) compared with non-migrating summer butterflies of the same species [92].

In contrast to *tutl*, mouse *Igsf9* is not an essential gene. *Igsf9* knockout mice are viable and fertile. They develop without any apparent abnormalities, and their coordinated motor control does not appear to be affected. Moreover, they do not demonstrate any disabilities in spatial learning in Morris water maze. The absence of behavioral (and other) defects in *Igsf9* knockout mice might be explained by functional redundancy caused by the expression of *Igsf9B* [43]. However, as mentioned above, the function of IGSF9B is unknown, and the ability of IGSF9 and IGSF9B to functionally replace each other remains to be determined.

### Effects of Mammalian IGSF9 Family Proteins on Axon Guidance and Neurite Growth and Branching

The outgrowth and guidance/pathfinding of axons is an important process during neural development and regeneration, ensuring that axons grow in the proper direction to form connections with the proper targets [93]. During the outgrowth of axons and dendrites, branching leads to the formation of axonal and dendritic arbors [94].

In contrast to original reports [17], the abrogation of Turtle expression in *Drosophila* has recently been reported to lead to defects in axon growth and guidance [51]. Thus, a lack of Turtle expression during embryonic development leads to defects in the growth and pathfinding of axons from motor and retinal neurons. The defects can be rescued by the expression of individual Turtle isoforms, in which the

secreted isoforms that contain IgI–Fn3I are the most potent. The overexpression of the same Turtle isoforms in a wild-type background leads to the increased branching of axons from motor and retinal neurons, and the diffusible isoforms are more potent in this regard than the transmembrane isoforms. These studies suggest that Turtle is involved in the outgrowth, guidance, and branching of neurites.

Mammalian forms of IGSF9 have been reported to regulate the branching of dendrites. In primary cultures of mouse hippocampal neurons, the knockdown of *Igsf9* transcripts by RNA interference (RNAi) or expression of a chimeric protein, in which the ~382 most C-terminal amino acids of IGSF9 are replaced with GFP (Dasm1[delC]-EGFP), was initially found to lead to a reduction of dendritic but not axonal arborization. These and other observations led the authors to name the product of *Igsf9* Dendrite Arborization and Synapse Maturation 1 (Dasm1) [70].

Later, the effects of Dasm1(delC)-EGFP expression was reproduced, whereas the expression of another C-terminally truncated version of IGSF9 that lacks the 200 most C-terminal amino acids (FLAG-C200) was found to have no effects on dendritic arborization [43]. Moreover, the observed effect of the aforementioned RNAi experiments was reported to be attributable to off-target effects. The morphology of hippocampal neurons from *Igsf9* knockout mice is similar to neurons from wildtype animals both in vivo and in vitro. The absence of IGSF9 does not affect brain-derived neurotrophic factor (BDNF)-stimulated dendritic growth and branching from cultured mouse hippocampal neurons [43].

Again, unknown is the extent to which the expression of IGSF9B can compensate for a lack of IGSF9 expression or an abrogation of IGSF9 function. However, judging from the aforementioned data, IGSF9 does not appear to play a pivotal role in dendritic arborization. Nevertheless, if the two tested C-terminally truncated forms of IGSF9 indeed have different effects on dendritic arborization, it is noteworthy that FLAG-C200 contains a large cytoplasmic domain, including the region with the aforementioned ITIM, whereas Dasm1(delC)-EGFP does not.

### Effects of Turtle on Neurite Growth and Branching, Tiling, and Self-avoidance

Tiling refers to a phenomenon in which the axonal or dendritic arbors of individual neurons become organized in a manner in which a given space is covered in an optimal, nonredundant manner (i.e., with minimal or no overlap between receptive fields). Self-avoidance refers to the ability of neurites within a single arbor to avoid interactions with neurites from the same neuron (i.e., isoneuronal interactions). This facilitates optimal, uniform coverage of the individual receptive fields (see [7, 94, 95] for recent reviews).

Neurite outgrowth, dendritic branching, and self-avoidance shape the receptive fields of individual neurons, whereas tiling optimizes the coverage and separation of regions of signal transmission for groups of neurons. Together, these processes are important for the proper processing of sensory information. Whereas tiling and self-avoidance imply a lack of isoneuronal interactions, heteroneuronal adhesive interactions can bundle neuronal connections in columns for specific neuronal connections. However, as described by Ferguson et al. [77], the restrictions of such columns are conceptually similar to tiling. One molecule that potentially contributes to such bundling of neuronal connections is NCAM2, which has been implicated in the formation of dendritic bundles in the granular retrosplenial cortex, where the protein is expressed at interbundle zones and appears to have a repellent effect on the dendritic branching of neighboring dendritic bundles [96, see [97] for review].

Both *Drosophila* and mammalian members of the family of DSCAMs have been implicated in tiling and self-avoidance (reviewed in [7, 56, 88, 94]). Moreover, Turtle has also been recently suggested to be involved in these processes [78]. The peripheral nervous system of *Drosophila* contains four classes of dendritic arborization neurons (I–IV) that all express Turtle. The morphology of the four classes of neurons differs greatly. Class I neurons have the simplest dendritic arbors, and class IV neurons have the most complex dendritic arbors. In one study, class IV neurons of larvae were found to exhibit an increased degree of isoneuronal dendritic crossing points, with no change in dendritic length, the number of termini, or the size of the receptive field, in response to the abrogation of Turtle expression, suggesting that the protein is important for self-avoidance in these neurons. DSCAM is also involved in self-avoidance in class IV neurons, but the two proteins appeared to mediate self-avoidance independently. Moreover, the authors found no evidence of an effect of Turtle on tiling. In class II and III neurons, a lack of Turtle expression was found to have no effect on branch numbers. In class I neurons, however, the absence of Turtle led to an increase in dendritic branching in a manner that was independent from the effects on dendritic growth, suggesting that the protein prevents excessive dendritic branching in neurons with simple arbors [78].

Surprisingly, in another study, the absence of Turtle expression in class II–IV neurons was found to have no effect on self-avoidance but instead led to significant decreases in dendritic length, the size of receptive fields, and the number of branches [79]. Generally, the protein appeared to promote branching in neurons with simple dendritic arbors and promote dendritic outgrowth in neurons with more complex arbors. Moreover, the absence of Turtle expression had no effects on the morphology of class I neurons [79].

No single explanation exists for these apparently contradictory data. The studies may have had allelic variations. One study investigated a subtype of class II and III neurons, whereas the other study did not [79]. However, additional factors might also contribute to the differences. Turtle has been reported to control tiling in the visual system of *Drosophila* [77]. The *Drosophila* compound eye consists of ommatidia that each contain eight different photoreceptor neurons, R1–R8. The expression of Turtle has been shown to be critical for R7 tiling by preventing interactions between adjacent R7 axons [77].

#### *Effects of IGSF9 Family Proteins on Synapse Maturation and Synaptic Transmission*

AMPA receptors are glutamate receptors that are crucial for rapid excitatory synaptic transmission. As mentioned above, mammalian forms of IGSF9 partially localize to the PSDs of excitatory synapses where they colocalize with the AMPA glutamate receptor subunits GluR2 and GluR4, suggesting that IGSF9 might affect the function of AMPA receptors. Indeed, experiments with RNAi knockdown of *Igsf9* have demonstrated that a lack of IGSF9 impairs the expression of AMPA receptors but not *N*-methyl-D-aspartate receptors [72]. Whether these observations, like the results from the other *Igsf9* knockdown experiments mentioned above, are the result of the off-target effects of the RNAi construct is unknown. However, *Igsf9* knockout mice are not impaired in basal synaptic transmission in the hippocampus, including input–output relationships of field excitatory postsynaptic currents and paired-pulse facilitation at CA1 synapses [43]. As mentioned above, the expression of the IGSF9-binding protein S-SCAM affects the level and function of AMPA receptors [90], so a potential effect of IGSF9 expression on AMPA-mediated processes should not be dismissed.

#### *IGSF9 Family Proteins in Diseases*

The expression of IGSF9 has been reported to be modulated in several types of cancer [98–101]. In a whole-genome expression profile of melanoma progression, *IGSF9* was found to be among the 50 most downregulated genes in advanced-stage melanomas (by approximately sevenfold) [98]. In another whole-genome expression study, *IGSF9* was among the genes downregulated twofold or more in colorectal familial adenomatous polyposis compared with normal colon tissue [99]. In apparent contrast to these studies, *IGSF9* was found to be one of the most highly overexpressed genes in gallbladder cancer (by approximately 18-fold) compared with normal gallbladder tissue [100]. In mouse and human fallopian tube carcinomas, from which high-grade serous ovarian cancer originates, *Igsf9/IGSF9* was found to be upregulated by

approximately 5- to 9-times fold compared with normal fallopian tube tissue [101]. Another observation with potential relevance to cancer is that *IGSF9* was shown to be a target of the AP-2 family member TFAP2C, a transcription factor involved in mammary development and potentially mammary cancer progression [102].

The expression of IGSF9B also appears to be modulated in some cancers [103, 104]. Sézary syndrome is a cutaneous lymphoma that originates from malignant T-cells. In a whole-genome expression analysis, *IGSFB* was among the 10 most upregulated genes (by approximately 11-fold) in malignant, circulating CD4+ T-cells compared with non-malignant T-cells from the same patients [103]. Moreover, using gene co-expression analysis, *IGSF9B* was among the few clusters of genes identified as potential prognostic markers of glioblastoma [104].

Cell adhesion molecules can be either up- or downregulated in different types of cancer, and the effects of changes in CAM expression on cancer progression are not always consistent. For example, the increased expression of NCAM1 correlates with the increased progression of neuroblastoma, myeloma, acute myeloid leukemia, and small-cell lung cancer, whereas the decreased expression of NCAM1 correlates with the increased progression of gliomas and thyroid and colon cancers [105]. Alterations in CAM expression can modulate cell–cell and cell–matrix interactions, thereby facilitating or reducing the migration and invasion of cancer cells. Moreover, CAMs can modulate cancer progression by modulating intracellular signaling in cancer cells. For example, NCAM1 promotes the metastatic process of epithelial ovarian carcinomas by interacting with and activating FGFR [106]. Whether the up- and downregulation of IGSF9-family proteins in various types of cancer have negative or positive effects on cancer progression remains to be determined.

IGSF9 may also play a role in seizure susceptibility. Thus, differences in seizure susceptibility between C57BL/6 (B6) and DBA/2 (D2) mice have been mapped to a region of chromosome 1 that includes *Igsf9*. Based on the expression profile of *Igsf9* and existence of an exon single-nucleotide polymorphism that leads to amino acid substitutions, *Igsf9* is one of 12 candidate genes responsible for the changes in seizure susceptibility. The difference between IGSF9 in relatively seizure-resistant B6 mice and relatively seizure-susceptible D2 mice is a Gln-to-Arg change at position 411 in IgIV and Pro-to-Ser change at position 871 in the cytoplasmic tail [107].

## Future Directions

The research reviewed above highlights the fact that the first decade of studies of the function of IGSF9-family

proteins has raised many questions but produced few answers, and the function of vertebrate members of the family is still elusive. As described earlier (see Fig. 1), *Hydra* IGSF9/Turtle lacks a transmembrane region. This indicates that the protein is secreted and suggests that IGSF9/Turtle was originally a secreted protein, but some IGSF9/Turtle-expressing genes have been modulated to also encode membrane-attached isoforms later during evolution. In *Drosophila*, the secreted isoforms still appear to be the most important because they have been reported to be able to cause stronger phenotypic changes than the membrane-attached isoforms [51]. In vertebrates, only the functions of transmembrane isoforms have been investigated so far. Seeing results from studies of secreted vertebrate isoforms of IGSF9 family proteins would be interesting.

The lack of an identified phenotype for the *Igsf9* knockout mouse may be attributable to redundancy caused, at least partially, by the expression of IGSF9B. Indeed, the sequences of the ectodomains of the two proteins are 50 % identical (see Table 1). However, as described above several differences between IGSF9 and IGSF9B exist, which suggest that the two proteins have somewhat different functions and therefore can replace each other only to some extent. For example, IGSF9 but not IGSF9B is potentially palmitoylated, suggesting that the two proteins have different subcellular localizations, and IGSF9B but not IGSF9 is potentially SUMOylated, further suggesting that the subcellular localizations of the two proteins are differentially regulated. Moreover, IGSF9 is potentially more glycosylated than IGSF9B. Finally, in contrast to IGSF9B, IGSF9 contains a C-terminal PDZ-binding domain. Together, these differences imply variations in the extracellular and intracellular binding partners for the two proteins.

Turtle has been shown to mediate homophilic *trans*-interactions [77], but whether vertebrate members of the protein family also serve as CAMs remains to be determined. In a recent study that pharmacologically inhibited Rho kinase (ROCK) to reduce glial scarring, the treatment of astrocytes with the ROCK inhibitor Fasudil in vitro led to morphological changes accompanied by the prolonged downregulation of *Igsf9* expression, suggesting that IGSF9 is involved in astrocyte-ECM interactions [108]. However, any heterophilic ligands of IGSF9-family proteins remain to be identified.

The importance of Turtle for *Drosophila* development is well documented [17], but the potential function of vertebrate IGSF9 proteins during embryonic development is unclear. *Igsf9* knockout mice do not exhibit apparent developmental defects [43]. However, a recent study of global gene expression alterations in *Danio rerio* in response to environmental lead poisoning [109] demonstrated that *Igsf9* was among the genes whose expression

was downregulated in response to lead exposure, although the study did not specify whether the data were derived from *Igsf9a* or *Igsf9b*. The expression of the gene that encodes the IGSF9-interacting synaptic protein SHANK1 in mammals was also downregulated. These observations suggest that a perturbation of IGSF9 expression during embryonic development could contribute to the neurological defects caused by environmental lead exposure [109]. In this case, clarifying the function of the protein during normal embryonic development would be relevant.

The function of vertebrate IGSF9 family proteins in adult organisms also remains to be determined. In a study of the transcriptome of the rodent pineal gland (i.e., a gland known to produce the circadian clock-related compound melatonin), *Igsf9* was one of the only CAM-expressing genes that was differentially expressed day and night (reviewed in [110]). This does not necessarily suggest that the protein is involved in controlling circadian rhythms but implies that the protein has a function in the adult organism that requires changes in expression levels. Moreover, the data suggest that the protein has functions that are different from the learning- and memory-related functions previously investigated [43] and that alternative assays should be utilized to unravel the functions of the protein.

Whether the up- and downregulation of IGSF9-family proteins in various types of cancer have negative or positive effects on cancer progression (if any) and the extent to which the proteins can be utilized as markers for specific types of cancer remain to be determined.

One study of the potential effects of epigenetic drift on ageing found that the DNA methylation of CpG island promoters was increased in *IGSF9B* in centenarians compared with newborns. Subsequently, quantitative PCR experiments revealed that the expression of *IGSF9B* was reduced by ~90 % in centenarians compared with newborns [111]. This is the first report of which we are aware that has revealed information about the potential function of IGSF9B. However, to elucidate the effects of reduced IGSF9B expression, basic knowledge about the function of IGSF9B is required.

In conclusion, the IGSF9 protein family constitutes an evolutionarily ancient group of proteins that to a large extent are uncharacterized. In *Drosophila*, Turtle seems to be one of the proteins, which in addition to DSCAM is involved in the regulation of tiling and self-avoidance. In mammals, DSCAM do not play the same pivotal role in the processes, and additional proteins must be involved in the regulation of the phenomena [88]. A recent study suggests that protocadherins in part serve this function [112], but IGSF9 and IGSF9B may still also be among the proteins regulating mammalian tiling and self-avoidance. Moreover, the expression of both IGSF9 and IGSF9B in e.g. the hippocampus [43] suggests that they may also be involved in processes related to learning and

memory formation, and their modulated expression in various cancers [98–101, 103, 104] may also have functional significance. It will be interesting to see future studies clarifying the functions of the IGSF9 family proteins in relation all these processes.

**Acknowledgments** This paper was written in honor of Prof. Elisabeth Bock (Protein Laboratory, Department of Neuroscience and Pharmacology, University of Copenhagen, Denmark) who for many years has dedicated her time to unraveling the function of proteins in the nervous system in general and NCAM1 in particular.

## References

- Walmod PS, Pedersen MV, Berezin V, Bock E (2007) Cell adhesion molecules of the immunoglobulin superfamily in the nervous system. In: Lajtha A, Banik N (eds) Handbook of neurochemistry and molecular neurobiology. Neural protein metabolism and function, 3rd edn. Springer, Berlin, pp 35–151. ISBN: 978-0-387-30379-6
- Li DM, Feng YM (2011) Signaling mechanism of cell adhesion molecules in breast cancer metastasis: potential therapeutic targets. *Breast Cancer Res Treat* 128(1):7–21. doi:10.1007/s10549-011-1499-x
- Redies C, Neudert F, Lin J (2011) Cadherins in cerebellar development: translation of embryonic patterning into mature functional compartmentalization. *Cerebellum* 10(3):393–408. doi:10.1007/s12311-010-0207-4
- Weber GF, Bjerke MA, DeSimone DW (2011) Integrins and cadherins join forces to form adhesive networks. *J Cell Sci* 124(8):1183–1193. doi:10.1242/jcs.064618
- Zhong X, Rescorla FJ (2012) Cell surface adhesion molecules and adhesion-initiated signaling: understanding of anoikis resistance mechanisms and therapeutic opportunities. *Cell Signal* 24(2):393–401. doi:10.1016/j.cellsig.2011.10.005
- Zhang Y, Yeh J, Richardson PM, Bo X (2008) Cell adhesion molecules of the immunoglobulin superfamily in axonal regeneration and neural repair. *Restor Neurol Neurosci* 26(2–3): 81–96
- Cameron S, Rao Y (2010) Molecular mechanisms of tiling and self-avoidance in neural development. *Mol Brain* 3(1):28
- Tallafuss A, Constable JR, Washbourne P (2010) Organization of central synapses by adhesion molecules. *Eur J Neurosci* 32(2):198–206. doi:10.1111/j.1460-9568.2010.07340.x
- Bottos A, Rissone A, Bussolino F, Arese M (2011) Neurexins and neuroligins: synapses look out of the nervous system. *Cell Mol Life Sci* 68(16):2655–2666. doi:10.1007/s00018-011-0664-z
- Chen Y, Aulia S, Li L, Tang BL (2006) AMIGO and friends: an emerging family of brain-enriched, neuronal growth modulating, type I transmembrane proteins with leucine-rich repeats (LRR) and cell adhesion molecule motifs. *Brain Res Rev* 51(2): 265–274
- Harwood A, Coates JC (2004) A prehistory of cell adhesion. *Curr Opin Cell Biol* 16(5):470–476. doi:10.1016/j.ceb.2004.07.011
- Hynes RO (2002) Integrins: bidirectional, allosteric signaling machines. *Cell* 110(6):673–687
- Jorgensen OS, Bock E (1974) Brain specific synaptosomal membrane proteins demonstrated by crossed immunoelectrophoresis. *J Neurochem* 23(4):879–880
- Rutishauser U, Thiery JP, Brackenbury R, Sela BA, Edelman GM (1976) Mechanisms of adhesion among cells from neural

- tissues of the chick embryo. *Proc Natl Acad Sci U S A* 73(2):577–581
15. Doudney K, Murdoch JN, Braybrook C, Paternotte C, Bentley L, Copp AJ, Stanier P (2002) Cloning and characterization of Iggsf9 in mouse and human: a new member of the immunoglobulin superfamily expressed in the developing nervous system. *Genomics* 79(5):663–670
  16. Nagase T, Kikuno R, Ishikawa KI, Hirotsawa M, Ohara O (2000) Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* 7(1):65–73
  17. Bodily KD, Morrison CM, Renden RB, Broadie K (2001) A novel member of the Ig superfamily, turtle, is a CNS-specific protein required for coordinated motor control. *J Neurosci* 21(9):3113–3125
  18. Kikuno R, Nagase T, Ishikawa K, Hirotsawa M, Miyajima N, Tanaka A, Kotani H, Nomura N, Ohara O (1999) Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* 6(3):197–205
  19. Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smailus DE, Schnerch A, Schein JE, Jones SJ, Marra MA (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci U S A* 99(26):16899–16903. doi:10.1073/pnas.242603899
  20. Clark HF, Gurney AL, Abaya E, Baker K, Baldwin D, Brush J, Chen J, Chow B, Chui C, Crowley C, Currell B, Deuel B, Dowd P, Eaton D, Foster J, Grimaldi C, Gu Q, Hass PE, Heldens S, Huang A, Kim HS, Klimowski L, Jin Y, Johnson S, Lee J, Lewis L, Liao D, Mark M, Robbie E, Sanchez C, Schoenfeld J, Seshagiri S, Simmons L, Singh J, Smith V, Stinson J, Vagts A, Vandlen R, Watanabe C, Wieand D, Woods K, Xie MH, Yansura D, Yi S, Yu G, Yuan J, Zhang M, Zhang Z, Goddard A, Wood WI, Godowski P, Gray A (2003) The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. *Genome Res* 13(10):2265–2270. doi:10.1101/gr.1293003
  21. ClustalW2. <http://www.ebi.ac.uk/Tools/msa/clustalw2/>. Accessed Aug 2012
  22. ClustalW2—phylogeny. [http://www.ebi.ac.uk/Tools/phylogeny/clustalw2\\_phylogeny/](http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/). Accessed Aug 2012
  23. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948. doi:10.1093/bioinformatics/btm404
  24. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38(suppl 2):W695–W699. doi:10.1093/nar/gkq313
  25. Treeview. <http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>. Accessed Aug 2012
  26. Page RDM (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12(4):357–358
  27. Eukaryotic linear motifs, ELM. <http://elm.eu.org/>. Accessed Aug 2012
  28. Dinkel H, Michael S, Weatheritt RJ, Davey NE, Van Roey K, Altenberg B, Toedt G, Uyar B, Seiler M, Budd A, Jödicke L, Dammert MA, Schroeter C, Hammer M, Schmidt T, Jehl P, McGuigan C, Dymecka M, Chica C, Luck K, Via A, Chatr-aryamontri A, Haslam N, Grebnev G, Edwards RJ, Steinmetz MO, Meiselbach H, Diella F, Gibson TJ (2011) ELM—the database of eukaryotic linear motifs. *Nucleic Acids Res*. doi:10.1093/nar/gkr1064
  29. Simple Modular Architecture Research tool, SMART. <http://smart.embl-heidelberg.de/>. Accessed Aug 2012
  30. Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 40(D1):D302–D305. doi:10.1093/nar/gkr931
  31. SignalP 4.0. <http://www.cbs.dtu.dk/services/SignalP/>. Accessed Aug 2012
  32. TMHMM Server v. 2.0. <http://www.cbs.dtu.dk/services/TMHMM/>. Accessed Aug 2012
  33. NetNGlyc 1.0. <http://www.cbs.dtu.dk/services/NetNGlyc/>. Accessed Aug 2012
  34. NetOGlyc 3.1. <http://www.cbs.dtu.dk/services/NetOGlyc/>. Accessed Aug 2012
  35. Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X (2008) CSS-Palm 2.0: an updated software for palmitoylation sites prediction. *Protein Eng Des Sel* 21(11):639–644. doi:10.1093/protein/gzn039
  36. Ren J, Gao X, Jin C, Zhu M, Wang X, Shaw A, Wen L, Yao X, Xue Y (2009) Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0. *Proteomics* 9(12):3409–3412. doi:10.1002/pmic.200800646
  37. Prenylation Prediction Suite, PrePS. <http://mendel.imp.ac.at/sat/PrePS/index.html>. Accessed Aug 2012
  38. Maurer-Stroh S, Eisenhaber F (2005) Refinement and prediction of protein prenylation motifs. *Genome Biol* 6(6):R55
  39. Prosite MyDomains Image Creator. <http://prosite.expasy.org/cgi-bin/prosite/mydomains/>. Accessed Aug 2012
  40. GENO3D. <http://pbil.ibcp.fr/htm/index.php>. Accessed Aug 2012
  41. Combet C, Jambon M, Deléage G, Geourjon C (2002) Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics* 18(1):213–214. doi:10.1093/bioinformatics/18.1.213
  42. PyMOL Molecular Graphics System. <http://www.pymol.org/pymol>. Accessed June 2012
  43. Mishra A, Knerr B, Paixao S, Kramer ER, Klein R (2008) The protein dendrite arborization and synapse maturation 1 (Dasm-1) is dispensable for dendrite arborization. *Mol Cell Biol* 28(8):2782–2791
  44. Superfamily: HMM library and genome assignments server. Accessed Aug 2012
  45. Gough J, Karplus K, Hughey R, Chothia C (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J Mol Biol* 313(4):903–919. doi:10.1006/jmbi.2001.5080
  46. Kristiansen LV, Hortsch M (2010) Fasciclin II: the NCAM Ortholog in *Drosophila melanogaster*. In: Berezin V (ed) Structure and function of the neural cell adhesion molecule NCAM, vol 663. Springer, Berlin, pp 387–401
  47. Owczarek SE, Kristiansen LV, Hortsch M, Walmod PS (2009) Cell adhesion molecules of the NCAM family and their role at synapses. In: Hortsch M, Umemori H (eds) The sticky synapse, 1st edn. Springer, Berlin, pp 265–299. ISBN: 978-0-387-92708-4

48. Meyer A, Van de Peer Y (2005) From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *BioEssays* 27(9):937–945. doi:10.1002/bies.20293
49. Makino T, McLysaght A (2010) Ohnologs in the human genome are dosage balanced and frequently associated with disease. *Proc Natl Acad Sci* 107(20):9270–9274. doi:10.1073/pnas.0914697107
50. Vogel C, Teichmann SA, Chothia C (2003) The immunoglobulin superfamily in *Drosophila melanogaster* and *Caenorhabditis elegans* and the evolution of complexity. *Development* 130(25):6317–6328. doi:10.1242/dev.00848
51. Al-Anzi B, Wyman RJ (2009) The *Drosophila* immunoglobulin gene turtle encodes guidance molecules involved in axon path-finding. *Neural Dev* 4:31
52. Wormbase. Accessed Aug 2012
53. Bork P, Holm L, Sander C (1994) The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol* 242(4):309–320. doi:10.1006/jmbi.1994.1582
54. Bloom L, Calabro V (2009) FN3: a new protein scaffold reaches the clinic. *Drug Discov Today* 14(19–20):949–955. doi:10.1016/j.drudis.2009.06.007
55. Connor RM, Allen CL, Devine CA, Claxton C, Key B (2005) BOC, brother of CDO, is a dorsoventral axon-guidance molecule in the embryonic vertebrate brain. *J Comp Neurol* 485(1):32–42. doi:10.1002/cne.20503
56. Millard SS, Zipursky SL (2008) Dscam-mediated repulsion controls tiling and self-avoidance. *Curr Opin Neurobiol* 18(1):84–89. doi:10.1016/j.conb.2008.05.005
57. Muller D, Mendez P, Deroo M, Klausner P, Steen S, Pogliola L (2010) Role of NCAM in spine dynamics and synaptogenesis. *Adv Exp Med Biol* 663:245–256. doi:10.1007/978-1-4419-1170-4\_16
58. Nguyen-Ba-Charvet KT, Chedotal A (2002) Role of slit proteins in the vertebrate brain. *J Physiol Paris* 96(1–2):91–98
59. Sato S, Omori Y, Katoh K, Kondo M, Kanagawa M, Miyata K, Funabiki K, Koyasu T, Kajimura N, Miyoshi T, Sawai H, Kobayashi K, Tani A, Toda T, Usukura J, Tano Y, Fujikado T, Furukawa T (2008) Pikachurin, a dystroglycan ligand, is essential for photoreceptor ribbon synapse formation. *Nat Neurosci* 11(8):923–931. doi:10.1038/nn.2160
60. Zhao YY, Takahashi M, Gu JG, Miyoshi E, Matsumoto A, Kitazume S, Taniguchi N (2008) Functional roles of N-glycans in cell signaling and cell adhesion in cancer. *Cancer Sci* 99(7):1304–1310. doi:10.1111/j.1349-7006.2008.00839.x
61. Vicente Miranda H, Outeiro TF (2010) The sour side of neurodegenerative disorders: the effects of protein glycation. *J Pathol* 221(1):13–25. doi:10.1002/path.2682
62. Dani N, Broadie K (2012) Glycosylated synaptomatrix regulation of trans-synaptic signaling. *Dev Neurobiol* 72(1):2–21. doi:10.1002/dneu.20891
63. Barczyk M, Carracedo S, Gullberg D (2010) Integrins. *Cell Tissue Res* 339(1):269–280. doi:10.1007/s00441-009-0834-6
64. Fukata Y, Fukata M (2010) Protein palmitoylation in neuronal development and synaptic plasticity. *Nat Rev Neurosci* 11(3):161–175. doi:10.1038/nrn2788
65. Niethammer P, Dellling M, Sytnyk V, Dityatev A, Fukami K, Schachner M (2002) Cosignaling of NCAM via lipid rafts and the FGF receptor is required for neurogenesis. *J Cell Biol* 157(3):521–532
66. Kulahin N, Grunnet LG, Lundh M, Christensen DP, Jorgensen R, Heding A, Billestrup N, Berezin V, Bock E, Mandrup-Poulsen T (2011) Direct demonstration of NCAM cis-dimerization and inhibitory effect of palmitoylation using the BRET(2) technique. *FEBS Lett* 585(1):58–64. doi:10.1016/j.febslet.2010.11.043
67. Walmod PS, Kolkova K, Berezin V, Bock E (2004) Zippers make signals: NCAM-mediated molecular interactions and signal transduction. *Neurochem Res* 29(11):2015–2035
68. Aricescu AR, Jones EY (2007) Immunoglobulin superfamily cell adhesion molecules: zippers and signals. *Curr Opin Cell Biol* 19(5):543–550. doi:10.1016/j.ceb.2007.09.010
69. Wilkinson KA, Henley JM (2010) Mechanisms, regulation and consequences of protein SUMOylation. *Biochem J* 428(2):133–145. doi:10.1042/bj20100158
70. Shi SH, Cox DN, Wang D, Jan LY, Jan YN (2004) Control of dendrite arborization by an Ig family member, dendrite arborization and synapse maturation 1 (Dasm1). *Proc Natl Acad Sci U S A* 101(36):13341–13345
71. Staub E, Rosenthal A, Hinzmann B (2004) Systematic identification of immunoreceptor tyrosine-based inhibitory motifs in the human proteome. *Cell Signal* 16(4):435–456
72. Shi SH, Cheng T, Jan LY, Jan YN (2004) The immunoglobulin family member dendrite arborization and synapse maturation 1 (Dasm1) controls excitatory synapse maturation. *Proc Natl Acad Sci U S A* 101(36):13346–13351
73. Taylor TD, Noguchi H, Totoki Y, Toyoda A, Kuroki Y, Dewar K, Lloyd C, Itoh T, Takeda T, Kim DW, She X, Barlow KF, Bloom T, Bruford E, Chang JL, Cuomo CA, Eichler E, Fitz-Gerald MG, Jaffe DB, LaButti K, Nicol R, Park HS, Seaman C, Sougnez C, Yang X, Zimmer AR, Zody MC, Birren BW, Nusbaum C, Fujiyama A, Hattori M, Rogers J, Lander ES, Sakaki Y (2006) Human chromosome 11 DNA sequence and analysis including novel gene identification. *Nature* 440(7083):497–500. doi:10.1038/nature04632
74. Flybase. Accessed Aug 2012
75. McQuilton P, St. Pierre SE, Thurmond J, FlyBase Consortium (2012) FlyBase 101—the basics of navigating FlyBase. *Nucleic Acids Res* 40(D1):D706–D714. doi:10.1093/nar/gkr1030
76. Karagozeos D (2003) Neural GPI-anchored cell adhesion molecules. *Front Biosci* 8:s1304–s1320
77. Ferguson K, Long H, Cameron S, Chang WT, Rao Y (2009) The conserved Ig superfamily member turtle mediates axonal tiling in *Drosophila*. *J Neurosci* 29(45):14151–14159
78. Long H, Ou Y, Rao Y, van Meyel DJ (2009) Dendrite branching and self-avoidance are controlled by turtle, a conserved IgSF protein in *Drosophila*. *Development* 136(20):3475–3484
79. Sulkowski MJ, Iyer SC, Kurosawa MS, Iyer EP, Cox DN (2011) Turtle functions downstream of cut in differentially regulating class specific dendrite morphogenesis in *Drosophila*. *PLoS ONE* 6(7):e22611. doi:10.1371/journal.pone.0022611
80. Heimbeck G, Bugnon V, Gendre N, Keller A, Stocker RF (2001) A central neural circuit for experience-independent olfactory and courtship behavior in *Drosophila melanogaster*. *Proc Natl Acad Sci* 98(26):15336–15341. doi:10.1073/pnas.011314898
81. Okada T, Yamada N, Tsuzuki K, Horikawa HPM, Tanaka K, Ozawa S (2003) Long-term potentiation in the hippocampal CA1 area and dentate gyrus plays different roles in spatial learning. *Eur J Neurosci* 17(2):341–349. doi:10.1046/j.1460-9568.2003.02458.x
82. Grueber WB, Jan LY, Jan YN (2003) Different levels of the homeodomain protein cut regulate distinct dendrite branching patterns of *Drosophila* multidendritic neurons. *Cell* 112(6):805–818. doi:10.1016/s0092-8674(03)00160-0
83. Cubelos B, Sebastián-Serrano A, Beccari L, Calcagnotto ME, Cisneros E, Kim S, Dopazo A, Alvarez-Dolado M, Redondo JM, Bovolenta P, Walsh CA, Nieto M (2010) Cux1 and Cux2 regulate dendritic branching, spine morphology, and synapses of the upper layer neurons of the cortex. *Neuron* 66(4):523–535. doi:10.1016/j.neuron.2010.04.038
84. Li N, Zhao C-T, Wang Y, Yuan X-B (2010) The transcription factor Cux1 regulates dendritic morphology of cortical pyramidal neurons. *PLoS ONE* 5(5):e10596. doi:10.1371/journal.pone.0010596
85. PaxDb: protein abundance across organisms. Accessed Aug 2012

86. De Vries M, Cooper HM (2008) Emerging roles for neogenin and its ligands in CNS development. *J Neurochem* 106(4):1483–1492. doi:[10.1111/j.1471-4159.2008.05485.x](https://doi.org/10.1111/j.1471-4159.2008.05485.x)
87. Dickinson RE, Duncan WC (2010) The SLIT–ROBO pathway: a regulator of cell function with implications for the reproductive system. *Reproduction* 139(4):697–704. doi:[10.1530/rep-10-0017](https://doi.org/10.1530/rep-10-0017)
88. Shi L, Lee T (2012) Molecular diversity of Dscam and self-recognition. *Adv Exp Med Biol* 739:262–275. doi:[10.1007/978-1-4614-1704-0\\_17](https://doi.org/10.1007/978-1-4614-1704-0_17)
89. Nielsen J, Kulahin N, Walmod PS (2010) Extracellular protein interactions mediated by the neural cell adhesion molecule, NCAM: Heterophilic interactions between NCAM and cell adhesion molecules, extracellular matrix proteins, and viruses. In: Berezin V (ed) Structure and function of the neural cell adhesion molecule NCAM, vol 663, pp 23–53
90. Danielson E, Zhang N, Metallo J, Kaleka K, Shin SM, Gerges N, Lee SH (2012) S-SCAM/MAGI-2 is an essential synaptic scaffolding molecule for the GluA2-containing maintenance pool of AMPA receptors. *J Neurosci* 32(20):6967–6980. doi:[10.1523/jneurosci.0025-12.2012](https://doi.org/10.1523/jneurosci.0025-12.2012)
91. Kreienkamp HJ (2008) Scaffolding proteins at the postsynaptic density: shank as the architectural framework. *Handb Exp Pharmacol* 186:365–380. doi:[10.1007/978-3-540-72843-6\\_15](https://doi.org/10.1007/978-3-540-72843-6_15)
92. Zhu H, Casselman A, Reppert SM (2008) Chasing migration genes: a brain expressed sequence tag resource for summer and migratory monarch butterflies (*Danaus plexippus*). *PLoS ONE* 3(1):e1345
93. McCormick A, Leipzig N (2012) Neural regenerative strategies incorporating biomolecular axon guidance signals. *Ann Biomed Eng* 40(3):578–597. doi:[10.1007/s10439-011-0505-0](https://doi.org/10.1007/s10439-011-0505-0)
94. Grueber WB, Sagasti A (2010) Self-avoidance and tiling: mechanisms of dendrite and axon spacing. *Cold Spring Harb Perspect Biol* 2(9):a001750. doi:[10.1101/cshperspect.a001750](https://doi.org/10.1101/cshperspect.a001750)
95. Gao FB (2007) Molecular and cellular mechanisms of dendritic morphogenesis. *Curr Opin Neurobiol* 17(5):525–532. doi:[10.1016/j.conb.2007.08.004](https://doi.org/10.1016/j.conb.2007.08.004)
96. Ichinohe N, Yoshihara Y, Hashikawa T, Rockland KS (2003) Developmental study of dendritic bundles in layer 1 of the rat granular retrosplenial cortex with special reference to a cell adhesion molecule, OCAM. *Eur J Neurosci* 18(7):1764–1774
97. Winther M, Berezin V, Walmod PS (2012) NCAM2/OCAM/RNCAM: cell adhesion molecule with a role in neuronal compartmentalization. *Int J Biochem Cell Biol* 44(3):441–446. doi:[10.1016/j.biocel.2011.11.020](https://doi.org/10.1016/j.biocel.2011.11.020)
98. Smith AP, Hoek K, Becker D (2005) Whole-genome expression profiling of the melanoma progression pathway reveals marked molecular differences between nevi/melanoma in situ and advanced-stage melanomas. *Cancer Biol Ther* 4(9):1018–1029
99. Bowden NA, Croft A, Scott RJ (2007) Gene expression profiling in familial adenomatous polyposis adenomas and desmoid disease. *Hered Cancer Clin Pract* 5(2):79–96. doi:[10.1186/1897-4287-5-2-79](https://doi.org/10.1186/1897-4287-5-2-79)
100. Kim JH, Kim HN, Lee KT, Lee JK, Choi SH, Paik SW, Rhee JC, Lowe AW (2008) Gene expression profiles in gallbladder cancer: the close genetic similarity seen for early and advanced gallbladder cancers may explain the poor prognosis. *Tumor Biol* 29:41–49
101. Kim J, Coffey DM, Creighton CJ, Yu Z, Hawkins SM, Matzuk MM (2012) High-grade serous ovarian cancer arises from fallopian tube in a mouse model. *Proc Natl Acad Sci*. doi:[10.1073/pnas.1117135109](https://doi.org/10.1073/pnas.1117135109)
102. Woodfield GW, Chen Y, Bair TB, Domann FE, Weigel RJ (2010) Identification of primary gene targets of TFAP2C in hormone responsive breast carcinoma cells. *Genes Chromosom Cancer* 49(10):948–962. doi:[10.1002/gcc.20807](https://doi.org/10.1002/gcc.20807)
103. Pomerantz RG, Mirvish ED, Erdos G, Falo LD, Geskin LJ (2010) Novel approach to gene expression profiling in Sézary syndrome. *Br J Dermatol* 163(5):1090–1094. doi:[10.1111/j.1365-2133.2010.09973.x](https://doi.org/10.1111/j.1365-2133.2010.09973.x)
104. Uppalapati P, Xiang Y, Huang K (2010) Predicting prognostic markers for glioma using gene co-expression network analysis. Paper presented at the proceedings of the first ACM international conference on bioinformatics and computational biology, Niagara Falls, New York
105. Zecchini S, Cavallaro U (2010) Neural cell adhesion molecule in cancer: expression and mechanisms. *Adv Exp Med Biol* 663:319–333. doi:[10.1007/978-1-4419-1170-4\\_20](https://doi.org/10.1007/978-1-4419-1170-4_20)
106. Zecchini S, Bombardelli L, Decio A, Bianchi M, Mazzarol G, Sanguineti F, Aletti G, Maddaluno L, Berezin V, Bock E, Casadio C, Viale G, Colombo N, Giavazzi R, Cavallaro U (2011) The adhesion molecule NCAM promotes ovarian cancer progression via FGFR signalling. *EMBO Mol Med* 3(8):480–494. doi:[10.1002/emmm.201100152](https://doi.org/10.1002/emmm.201100152)
107. Ferraro T, Golden G, Smith G, Martin J, Lohoff F, Gieringer T, Zamboni D, Schwebel C, Press D, KratzerZhao SH, Berrettini W, Buono R (2004) Fine mapping of a seizure susceptibility locus on mouse Chromosome 1: nomination of Kcnj10 as a causative gene. *Mamm Genome* 15(4):239–251
108. Lau CL, Perreau VM, Chen MJ, Cate HS, Merlo D, Cheung NS, O’Shea RD, Beart PM (2012) Transcriptomic profiling of astrocytes treated with the Rho kinase inhibitor Fasudil reveals cytoskeletal and pro-survival responses. *J Cell Physiol* 227(3):1199–1211. doi:[10.1002/jcp.22838](https://doi.org/10.1002/jcp.22838)
109. Peterson SM, Zhang J, Weber G, Freeman JL (2011) Global gene expression analysis reveals dynamic and developmental stage-dependent enrichment of lead-induced neurological gene alterations. *Environ Health Perspect* 119(5):615–621. doi:[10.1289/ehp.1002590](https://doi.org/10.1289/ehp.1002590)
110. Klein DC, Bailey MJ, Carter DA, Kim J-s, Shi Q, Ho AK, Chik CL, Gaildrat P, Morin F, Ganguly S, Rath MF, Møller M, Sugden D, Rangel ZG, Munson PJ, Weller JL, Coon SL (2010) Pineal function: impact of microarray analysis. *Mol Cell Endocrinol* 314(2):170–183
111. Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, Diez J, Sanchez-Mut JV, Setien F, Carmona FJ, Puca AA, Sayols S, Pujana MA, Serra-Musach J, Iglesias-Platas I, Formiga F, Fernandez AF, Fraga MF, Heath SC, Valencia A, Gut IG, Wang J, Esteller M (2012) Distinct DNA methylomes of newborns and centenarians. *Proc Natl Acad Sci*. doi:[10.1073/pnas.1120658109](https://doi.org/10.1073/pnas.1120658109)
112. Lefebvre JL, Kostadinov D, Chen WV, Maniatis T, Sanes JR (2012) Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* 488(7412):517–521. doi:[10.1038/nature11305](https://doi.org/10.1038/nature11305)