Okinawa Institute of Science and Technology

Graduate University

Thesis submitted for the degree

Doctor of Philosophy

Molecular dissection of ancestral glia

by

Larisa Sheloukhova

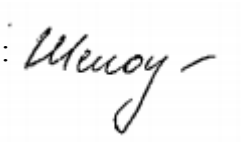
Supervisor: Hiroshi Watanabe

24 March 2022

Declaration of Original and Sole Authorship

I, Larisa Sheloukhova, declare that this thesis entitled **"Molecular dissection of ancestral glia"** and the data presented in it are original and my own work.
I confirm that:

- No part of this work has previously been submitted for a degree at this or any other university.
- References to the work of others have been clearly acknowledged. Quotations from the work of others have been clearly indicated, and attributed to them.
- In cases where others have contributed to part of this work, such contribution has been clearly acknowledged and distinguished from my own work.
- None of this work has been previously published elsewhere.

Signature: 

Date: 24 March 2022

Abstract

Nervous systems of bilaterian animals generally consist of two cell types: neurons and glial cells. Glia participate in almost every process taking place in the nervous system of bilaterians, indicating a crucial role of glial cells in both neurophysiological functions and for the nervous system evolution. Therefore, tracing back the first glia and elucidating its ancestral function is important for understanding the evolutionary sophistications of the nervous system. Histological examinations have not so far revealed any morphological sign of glial cells in *Cnidaria*, the closest outgroup to *Bilateria*. This led to the hypothesis that glial cells appeared after the common bilaterian ancestor had branched off from *Cnidaria*. However, this view has not been well examined at the genetic level. In this work I sought to investigate gliogenic program conservation in non-bilaterian animals, i.e. basal metazoans (*Cnidaria, Placozoa, Ctenophora,* and *Porifera*). First, I performed a phylum-wide and genome-wide survey on representative species of bilaterians and basal metazoans to clarify the conserved genetic repertoire required for bilaterian glial development. I found that the *glial cells missing (Gcm)* is one of the evolutionary conserved glial transcription factors (TF). Second, in *Nematostella vectensis,* a cnidarian model with highly conserved genetic repertoire required for bilaterian glial development, a homolog of *Gcm* is expressed in specific neuronal cells. I analyzed the function of *Gcm* by knocking it down and performing RNA-seq and RT-qPCR analyses. I found that the *Gcm* knockdown in *Nematostella* embryos resulted in expression alterations of cell adhesion proteins, GABA and glutamate transporters, ion channels, metabolic and protein modifying enzymes, as well as zinc finger and *Ets*-related TFs. In addition, *Gcm* seems to control *Notch-Delta* signaling, which is one of the crucial neuro-gliogenic pathways in bilaterians. Immunostaining of a *Gcm* target protein visualized a novel class of cells with flat cell body and no clear neurite process, which were previously classified as neurons as they express neuronal markers (neuropeptides). The major finding of my thesis is that *Nematostella Gcm*-expressing cells demonstrate characteristics of both neurons and glia, suggesting a dual nature of ancestral cells. This may indicate that the ancestral gliogenic program was intertwined with the neurogenic program and separated later in the animal evolution.

Acknowledgments

First I would like to thank my supervisor, Dr. Hiroshi Watanabe, who gave me an opportunity to work on my project in his lab and supported me during the PhD. He not only guided me throughout my research, but also gave helpful scientific and otherwise advice. I would like to thank professor Watanabe for fruitful discussions and continuous encouragement.

I would also like to thank my lab members for their mentorship and support. Many thanks go to Ivan Mbogo and Ryotaro Nakamura, who were always there when I needed help. My gratitude also goes to our lab technicians, Amol Dahal, Akiko Tanimoto, and Junko Higuchi, who took such good care of the animal culture. I would also like to thank Christine Guzman and Mei Fang Lin for their guidance in bioinformatics analysis. Chihiro Arasaki, Chihiro Kawano, Ryo Nakamura, Osamu Horiguchi, thank you for your moral support, informal discussions, and sweets.

I would like to thank my Thesis Committee Members, Dr. Tomoyuki Takahashi and Dr. Bernd Kuhn, for their continuous support along with the skills and knowledge I acquired when I did lab rotations in their units. My gratitude also goes to Konstantin Khalturin for helpful discussions and tips.

To my close friends and family. I would like to thank my mother, Galina Sheloukhova, and sister, Olesya Pechenova, for their love and support. Many thanks to my partner, Saahil Acharya, for his continuous encouragement, thoughtful conversations, fun and caring. I also greatly appreciate the support I got from my friends at OIST and beyond.

Last but not least, I would like to thank OIST graduate school and Health Center for all the support and making my life in Okinawa easy and enjoyable. I would like to express my gratitude to the Okinawan people I was lucky to get to know for their friendliness and welcoming attitude. I truly enjoyed my PhD life at OIST on this beautiful island full of sunshine.

List of Abbreviations

18S - 18S ribosomal RNA
BBB - blood-brain barrier
bHLH - basic helix-loop-helix
BMP - bone morphogenetic protein
DE - differential expression
Eaat - excitatory amino acid transporter
ECM - extracellular matrix
Ef1a - elongation factor 1-alpha
GABA - gamma-Aminobutyric acid
Gapdh - glyceraldehyde 3-phosphate dehydrogenase
Gat - GABA transporter
Gcm - glial cells missing
Gfap - glial fibrillary acidic protein
Glut - glucose transporter
GO - gene ontology
GPCR - G protein-coupled receptor
GS - glutamine synthetase
IF - immunofluorescence
Jak-Stat - The Janus kinase-signal transducer and activator of transcription
KD - knockdown
NfI - nuclear factor I
Nv - *Nematostella vectensis*
OPC - oligodendrocyte precursor cell
Repo - reversed polarity
RGC - radial glial cell
RT-qPCR - real-time quantitative polymerase chain reaction
SCT - single cell transcriptome
TF - transcription factor
TPM - transcripts per million
WB - western blotting

Table of contents

List of figures

List of tables

Chapter I. Introduction

1.1 Problem statement and aims

Metazoans (animals) are broadly subdivided into bilaterians (*Protostomia* and *Deuterostomia*) and basal metazoans (*Porifera* (sponges), *Placozoa*, *Ctenophora* (comb jellies), *Cnidaria*). The nervous system (NS) of bilaterian animals are generally composed of two cell types: neurons and glial cells. Among basal metazoans only *Ctenophora* and *Cnidaria* possess neurons forming a diffuse NS, and are believed to lack glia (Hartline, 2011; Verkhratsky et al., 2019) **(Figure 1.1)**.

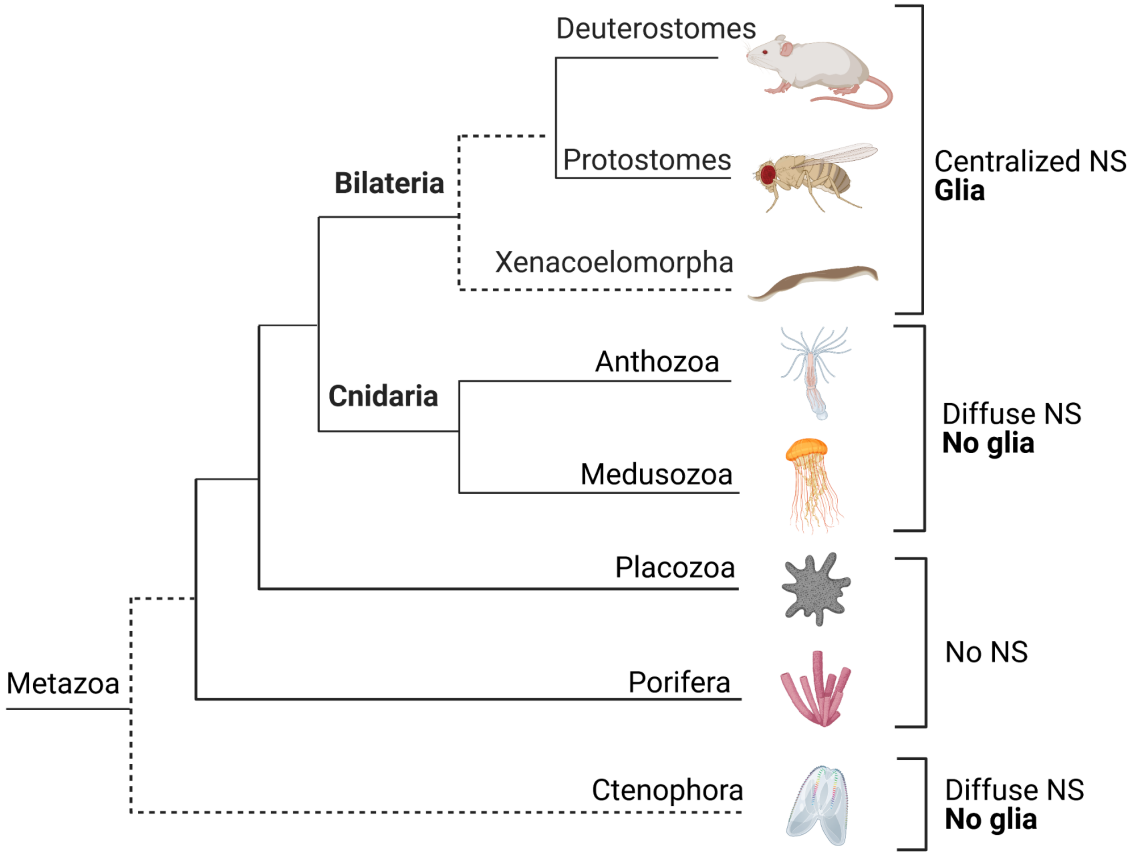


Figure 1.1. Current view of glial evolution. The metazoan phylogenetic tree of the main animal groups includes the unresolved positions of *Ctenophora* and *Xenacoelomorpha* indicated by the dotted branches (Kapli & Telford, 2020). Created with BioRender.com.

Neurons are electrically active which makes them the primary functional component of the NS. Glia are considered to be supportive cells. Nevertheless they take part in almost every process in the nervous system of bilaterians**,** including neurotransmission, homeostasis and an important function of producing neurons (Araque & Navarrete, 2010; Dimou & Götz, 2014). In the course of metazoan evolution glial cells have acquired greater morphological and functional diversity (Verkhratsky & Butt, 2013). The remarkable diversity is obvious especially in higher bilaterians such as vertebrates, as they include distinct types even within

one glial cell class (Zhang & Barres, 2010). The emergence of glial cells is one of the key novelties in the evolution of the nervous system.

The important unanswered questions in glial biology are where glial cells come from, and what their first functions were. Was the first nervous system exclusively composed of multifunctional self-sustaining neurons, which later in evolution became more complex and needed a helpful counterpart - glia? Was the first nervous system composed of glia-neurons that had features of both cell types instead? Or rather did glia predate neurons, and is an indispensable part of the nervous system? Tracing back the first glia and elucidating their original function is therefore crucial for understanding the evolution of the nervous system.

Extensive glia studies have been mostly confined to a few bilaterian model organisms, i.e. *Caenorhabditis elegans* and *Drosophila melanogaster* (protostomes) among invertebrates, *Danio rerio* (zebra fish), *Mus musculus* (mouse), *Rattus norvegicus* (rat) and *Homo sapiens* (human) (deuterostomes) among vertebrates. Some animals from other lineages including basal bilaterians have been screened for the presence of cells morphologically similar to glia. It is currently believed that *Xenacoelamorpha* (interstitial flatworms), supposedly the most basal bilaterians, are the most primitive animals possessing glia-like cells (Verkhratsky et al., 2019). Overall glia are hypothesized to have evolved with the appearance of the centralized nervous system, i.e. after the common bilaterian ancestor had branched off from *Cnidaria* (Freeman & Rowitch, 2013; Verkhratsky & Butt, 2013) **(Figure 1.1)**. However, there have not been any molecular investigations of conserved glial markers in cnidarians, or any other basal metazoans.

Therefore, I set the following **aims for my PhD research**:

1) investigate to what extent the bilaterian gliogenic genetic program is conserved in basal metazoans

2) explore the core evolutionary conserved glial gene(s) in a cnidarian, *Nematostella vectensis*

3) investigate the functions of the core gliogenic genes conserved in *Nematostella*

4) explore the physiological and morphological features of the cells expressing the conserved glial genes in *Nematostella*

1.2 Defining glia

To unravel the evolutionary origins of glial cells, it is important to first define them.

Based on what is known about glial cells and their functions in bilaterians, several definitions of glia emerged. According to Shai Shaham (Shaham, 2005), there are three aspects characteristic of glial cells:

1. Morphology: glial cells are associated with neurons
2. Physiology: glial cells do not conduct fast currents, do not possess neurotransmitter-laden vesicles and do not form presynaptic structures
3. Origin (development): glia, together with neurons, arise from neuroectoderm during embryogenesis

1.2.1 Morphology

Morphological examination has been the most widely used and often the only method of identifying glia in various animals. Despite several morphological variations observed in

glial cells of different species (reviewed in (Hartline, 2011)), close association with neurons is generally regarded as a common glia-specific feature. In addition, glial processes usually ensheath axons and creep in between nerve bundles (Radojcic & Pentreath, 1979). *Cnidaria*, basal metazoans with the most primitive NS along with *Ctenophora*, are believed to lack glial cells based on morphological assessment of their representatives performed in the 1960s (Horridge & Mackay, 1962). The authors examined the ectodermal tissue of two cnidarians (jellyfish) using an electron microscope. They did not observe any cells ensheathing the axons or being associated with neurons otherwise. However, a more exhaustive search is necessary in this phylum to conclude that they do not possess glia. First, their glia population might be composed of just a few cells. Second, their glia-like cells might be associated with a specific group of neurons not examined in the study. Finally, *Cnidaria* consists of two clades, *Anthozoa* (sea anemone, coral) and *Medusozoa* (jellyfish) **(Figure 1.1)**, which differ dramatically not only in their body shape and life cycle, but also genetic composition (Khalturin et al., 2019). While absent in *Medusozoa* glia-like cells might be present in *Anthozoa* and vice versa. In general, glial cells are more or less associated with neurons depending on the phylum/species in question and glia cell type (reviewed in Hartline, 2011). In addition, even *C. elegans*, considered the most simple organism with developed glia, has morphologically diverse glial cells (Oikonomou & Shaham, 2011). Thus, morphology, if used as the only criteria, is not reliable when it comes to identifying glial cell type.

1.2.2 Physiology

It is well known that structure defines function. The astonishing morphological diversity of bilaterian glia is determined by the mirriad of functions these cells perform: providing energy to neurons, maintaining the extracellular environment of neurons, immune response, serving as stem cells for generating glia and neurons in the adult brain, formation of blood-brain barrier (BBB) (Gebara et al., 2016; Michinaga & Koyama, 2019; Sancho et al., 2020) On the other hand, the diversity of functions and glial cell types increases with the complexity of the NS. The increasing number of glia (from 10% of total brain volume in model invertebrates to over 50% in mammals), glia-to-neuron ratio, and glial cells' complexity drives the theory that first neurons did not need supporting glial cells (Verkhratsky & Butt, 2013). Therefore, at some point during evolution, neurons could not energetically support themselves any more, thus glia emerged. This led to a general definition of glia cells as "homeostatic cells of the nervous system" (Verkhratsky & Butt, 2013). According to this definition glia present as housekeeping cells while neurons hold a unique title of information processing units. At the same time, studies on *C.elegans* suggest that the supposedly first glia emerged at the sensory receptive endings to control/support neuronal processing of incoming information about the environment (Oikonomou & Shaham, 2011). In mammals glial cells also play an important role in information processing by modulating synaptic activity and connectivity (Perea et al., 2014; Verkhratsky & Nedergaard, 2014). To assign one function to glial cells in higher animals and use it as a universal feature for all glial cells is therefore problematic.

Certain physiological features are helpful to consider when defining glial cells.
Unlike neurons, glia have not been documented to generate action potentials, even though a fraction of mouse NG-2-expressing oligodendrocyte precursor cells (OPCs) were shown to produce spikes when depolarized (Káradóttir et al., 2008). Nor do we know about a complete

synaptic machinery present in glial cells. Nevertheless, bidirectional communication between neurons and astroglia is known. This involves neurotransmitters being released from neurons and acting on astrocytes, while gliotransmitters are in turn released from glia to modulate neuronal activity (Bezzi & Volterra, 2001). Moreover, some gliotransmitters are packed and released from glial vesicles that share many characteristics of synaptic vesicles including some membrane proteins (Bergersen et al., 2012; Martineau et al., 2013).

Unfortunately, physiological studies are generally limited to bilaterian model organisms. While the development of physiological techniques in basal lineages is in its infancy, genetic manipulations have been established and are widely used in evo-devo studies including the evolution of the NS. Functional analyses of various genes driving bilaterian neurogenesis in basal metazoans have already shed light on the evolution of neurons (Rentzsch et al., 2020; Richards & Rentzsch, 2015; Tournière et al., 2020, 2021). Similar approaches might prove useful to understand the evolution of glial cells.

1.2.3 Molecular markers

Molecular markers have long been used to identify cell types and reflect functional characteristics of the cells in question. With the emergence of omics tools it has become even more prevalent. Surprisingly though, even among bilaterians a handful of glial markers have been identified and searched for in non-model organisms, which makes it difficult to make a conclusion about glial presence or absence in many phyla (Hartline, 2011). Therefore, our knowledge of glial presence even in bilaterian lineages is fragmented. It is thus not clear if glia evolved once or several times.

The most widely used glial marker is an intermediate filament protein *Gfap* (glial fibrillary acidic protein), which was first identified in vertebrates. *Gfap* immunostaining was then used to identify potential glia in various invertebrates including a basal bilaterian *Symsagittifera roscoffensis* (*Xenacoela*) (Bailly et al., 2014). NS-specific staining probably revealing glial cell distribution is reported in these studies. However, it is not clear if the antibody binds a *Gfap*-homologous protein. Moreover, in vertebrates *Gfap* labels only a subset of astroglia (Cahoy et al., 2008), and is also expressed in non-glial cells (Hainfellner et al., 2001). It is a late astrocytic marker, expressed in radial glia, which gives rise to both neurons and astrocytes, and is not expressed in other glial types like oligodendrocytes (Cahoy et al., 2008). Thus, additional molecular markers should be explored to conclude if the observed cells are indeed glia in this impossibly basal bilaterian lineage.

Intermediate filaments are prominent features of glial cells in several bilaterians examined including annelids (Beckers et al., 2019). Apart from *Gfap* another intermediate filament protein, *vimentin,* is used as a glial marker, e.g. in the snail (Dos Santos et al., 2005). In mammals *vimentin* is an early marker of some glial cells (radial glia and astrocytes) which is later replaced by the expression of *Gfap* (Bramanti et al., 2010). Overall it is less glia-specific than *Gfap* even in invertebrates (Cardone & Roots, 1990). Interestingly, *If-1* (intermediate filament-1) protein was shown to be expressed in planarian glia (Wang et al., 2016).

Other vertebrate glial markers searched for and used to identify glia in invertebrates are *Gs* (glutamine synthetase), e.g. in the lobster (*Panulirus argus*) (Linser et al., 1997) and *Aplysia* (Levenson et al., 2000); *S100b* (S100 calcium-binding protein B), e.g. in the giant prawn (*Macrobrachium rosenbergi)* (Allodi et al., 2006) and flatworm (*Christensenella*

minuta) (Biserova et al., 2010); transporters for glutamate (*Eaat)*, GABA (*Gat*), and glucose (*Glut*), e.g. in planaria (*Schmidtea mediterranea*) (Wang et al., 2016). In a recent study the expression of *Eaat*, *Gs*, *GFAP/vimentin/If* genes was explored in the lancelet shining light on glia in *Cephalochordata* (Bozzo et al., 2021). *Sco-spondin* (*sspo*), an extracellular matrix (ECM) glycoprotein, which is the main component of Reissner's fiber (Driever, 2018), is another glial marker. It is secreted by radial glia and was used to identify glial presence in both deuterostomes and protostomes (Conrad et al., 2017; Viehweg et al., 1998). Some glial markers seem phylum/class-specific and have been used to identify glia in certain animals. For example *Gnrh,* gonadotropin-releasing hormone (GnRH), is expressed in glia of a urochordate *Ciona ntestinalis* (Okawa et al., 2020), while *Aplysia* glia secrete a protein called *Ag* (Lockhart et al., 1996).

Some transcription factors (TFs) driving gliogenesis are used as glial markers: *Gcm* (*glial cells missing*) and *repo* in *Drosophila* (*Arthropoda*), and *Olig* in vertebrates. These are discussed in detail in section 1.2.4. **Figure 1.2** summarizes glial markers identified in various bilaterian lineages.

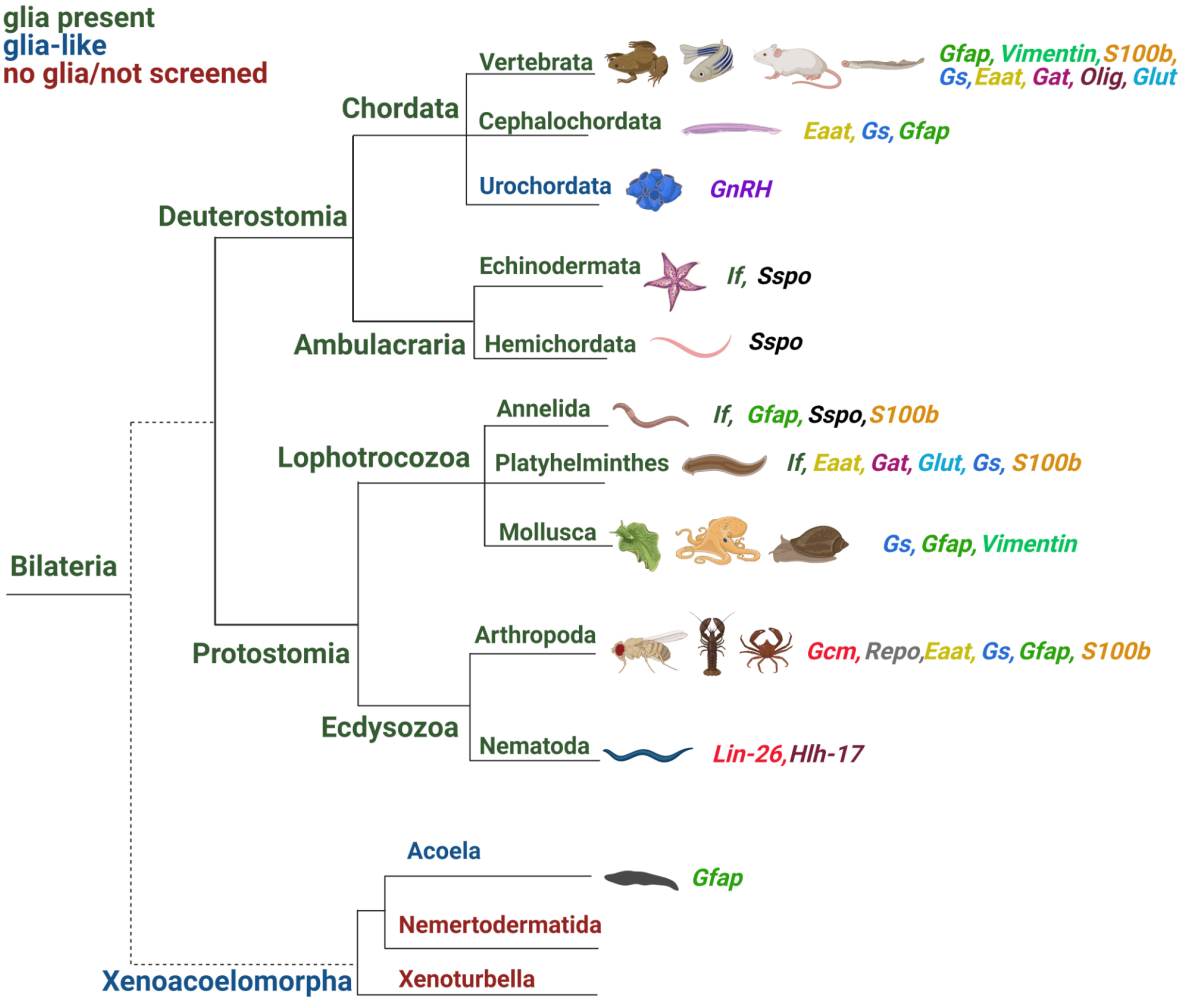


Figure 1.2. Bilaterian glial phylogeny. Presence of glia is inferred from reports on cells morphologically and molecularly similar to well-developed glia of vertebrates. Glial markers identified in

Caption continuation for Figure 1.2
each phylum are specified and color-coded. Note that in Nematoda *Lin-26* and *Hlh-17* are not homologs of *Gcm* and *Olig* respectively, but are colored the same because of their functional similarity. Representative species screened for glial markers are enumerated in Appendix 1 and accompanied by the literature references. Created with BioRender.com.

1.2.4 Gliogenic developmental program

The expression of glial molecular markers, i.e. functional genes, is driven by TFs and extracellular cues. Therefore, by unraveling the genetic developmental program driving gliogenesis in animals with well-developed and thoroughly studied glia it will be possible to not only define glia more precisely but also trace back its origin to the organisms with the most primitive NS.

1.2.4.1. Gliogenic TFs

In *Drosophila* it has been possible to separate out the TFs that drive the expression of glial-specific markers, the so-called 'binary switches', such as *Gcm* and *Repo* (Lee & Jones, 2005; Yuasa et al., 2003). As a result, it is considered that more primitive organisms have a more robust intrinsic system for neuro/gliogenesis. The situation is more complex with vertebrates, particularly mammals. No specific TFs responsible for glial cell fate acquisition have been identified per se, even though some, such as *bHLH*s, *Sox* group E, *Ets* and *NFI* family, are obviously involved (Glasgow et al., 2014; Molofsky et al., 2013; Stolt et al., 2003). These either promote glial marker expression or suppress proneural activators. It is therefore believed that the gliogenic program in protostomes (insects) and deuterostomes (vertebrates) are not homologous.

Gcm

Gcm is a master regulator of gliogenesis in *Drosophila* (Freeman et al., 2003; Hosoya et al., 1995; Soustelle & Giangrande, 2007). *Gcm* was shown to be indispensable for gliogenesis, since a mutation in this gene turns presumptive glia to develop into neurons (Hosoya et al., 1995). All lateral glia of *Drosophila* rely on *Gcm* for their development, but the specific subtypes of glial cells are established via different mechanisms (Yildirim et al., 2019). *Drosophila* has two *Gcm* genes with partially redundant functions (Alfonso & Jones, 2002; Kammerer & Giangrande, 2001). *Gcm2* seems less important for glia differentiation, but is involved in macrophage development (Alfonso & Jones, 2002; Kammerer & Giangrande, 2001). Both *Gcm* genes promote postembryonic neurogenesis in addition to glial development in *Drosophila* CNS (Chotard et al., 2005).

The neuro/gliogenic functions of *Gcm* homologs were shown beyond insects. *Gcm* neural expression and function was also demonstrated in crayfish (crustaceans) (Junkunlo et al., 2020). It is not clear if crustacean *Gcm*-expressing cells are glial. Nevertheless, similarly to the fly crustacean *Gcm* seems to be coexpressed with *Repo,* which is an important gliogenic TF under *Gcm* regulation in *Drosophila*.

Sea urchin (*Echinodermata*) is another invertebrate possessing *Gcm*, but it is not expressed in the nervous tissue (Ransick & Davidson, 2006). It plays a role in pigment cell specification instead (Wessel et al., 2020). Importantly though in both *Drosophila* and sea urchin *Gcm* expression is driven by *Notch* (Udolph et al., 2001). *Notch-Delta* signaling plays a major role in neuro/gliogenesis in bilaterians (Fischer & Gessler, 2007; Zhou et al., 2010). It promotes gliogenesis by activating glial genes and sustains neuronal precursors both in vertebrates and invertebrates (Gaiano & Fishell, 2002; Ge et al., 2002; Wheeler et al., 2008). *Hes/hairy* are primary downstream targets of *Notch*, and play a crucial role in gliogenesis by inhibiting such proneuronal bHLH activators as *Atonal* and *Achaete–Scute complex* (Fischer & Gessler, 2007); (Monastirioti et al., 2010; Wheeler et al., 2008). In *Drosophila Notch* drives *Gcm* expression to give rise to subperineurial glia (SPG) (Udolph et al., 2001). However,

Notch has an opposite effect on *Gcm* in sensory organ precursor (SOP) lineage of the peripheral nervous system (PNS) of *Drosophila* (Bor & Giangrande, 2001). Somewhat similarly there is no unified effect of *Notch* on glial differentiation in vertebrates, where *Notch* promotes certain glial types such as astrocytes and radial glia, but not oligodendrocytes (reviewed in Louvi & Artavanis-Tsakonas, 2006).

The function of *Gcm* homologs in vertebrates is still debated. *Gcma* (or *Gcm1*) and *Gcmb* (or *Gcm2*) were isolated in mammals (Kim et al., 1998). These are expressed predominantly not in the nervous tissue: *Ggma* is expressed in placenta (Schreiber et al., 2000; Yamada et al., 1999), *Gcmb* in parathyroid glands (Yu, et al., 2007). However, both genes induce the generation of neural stem cells (Hitoshi et al., 2011). The neurogenic role of *Gcma* is conserved in the chicken (Hitoshi et al., 2011; Soustelle et al., 2007). Moreover, rodent *Gcma* has a capacity to induce gliogenesis and drive astrocyte differentiation (Iwasaki et al., 2003; Saito et al., 2012). In zebrafish one *Gcm* gene was isolated and coined *Gcmb* due to its similarity to the mammalian *Gcmb.* It is expressed in macrophages and plays an important role in pharyngeal cartilage formation (Hanaoka et al., 2004).

In *C.elegans* all glial cells express Zn-finger transcription factor *Lin-26 (Labouesse et al., 1996)*. Epithelial cells also express *Lin-26* and are transformed into neurons in *Lin-26* mutants, highlighting the conserved linear relationships between glia, neurons and epithelia. The function of *lin-26* is reminiscent of the function of *Gcm*. However, *Lin-26* and *Gcm* proteins are not obviously homologous. Based on *C.elegans* single-cell transcriptome data *Lin-26* is indeed expressed mostly in glia, but also hypodermis, gonad, and some other cells in small amounts (Cao et al., 2017).

Therefore, the following conclusions can be made about *Gcm* genes in protostomes and deuterostomes:

1) In those lineages that have one *Gcm* gene the TF function is not clear, but might be neural in some;
2) In those lineages that have two *Gcm* genes, homologs of Drosophila *Gcm1* show gliogenic functions;
3) In both Protostomes and Deuterostomes *Gcm* is tightly regulated by *Notch* (Hitoshi et al., 2011; Udolph et al., 2001; Umesono et al., 2002).

Thus, the *Notch*-regulated *Gcm*-driven gliogenic program seems conserved in vertebrates and invertebrates. This opens up a possibility of exploring its conservation in their common ancestors.

Repo

Repo is a homeobox gene downstream from *Gcm* in *Drosophila*, which ensures terminal differentiation of glial cells (Trébuchet et al., 2019). *Repo* in turn drives the expression of glial-specific markers, including *Pointed*, which is also under the regulation of *hairy* (Bianchi-Frias et al., 2004; Yuasa et al., 2003; Bianchi-Frias et al., 2004). In addition to *Pointed, Tramtrack69*, *Loco* and *M84* are effectors of *Repo* that play a role in glial differentiation and morphogenesis in *Drosophila* (Yuasa et al., 2003). *Repo* mutants demonstrate reduced cell number and poorly differentiated glia (Halter et al., 1995). *Repo* was not reported to have a gliogenic role in animals other than insects. Vertebrates lack the gene for *Repo* altogether.

Olig

Olig belongs to group A bHLH genes. *Olig1* and *Olig2* were shown to promote oligodendrocyte differentiation in mammals (Dai et al., 2015; Hu et al., 2007). These are used

as mature oligodendrocyte specific markers, even though a subpopulation of astrocytes was shown to express *Olig2* (Ohayon et al., 2020). Importantly, *Olig* are among several TFs that couple neuronal and oligodendrocyte type specification (Lu et al., 2002; Zhou & Anderson, 2002), which means it is not a strictly glial TF. The primary function of oligodendrocytes is to myelinate neurons to ensure fast action potential propagation. Myelin is a new invention from the evolutionary point of view as it is associated with vertebrates (Salzer, 2015; Verkhratsky et al., 2019). On the other hand axonal ensheathment with glial membranes is observed in animals with primitive glia. This probably ensured the first functions of glia such as metabolic support for neurons. Interestingly, *C.elegans* possesses a bHLH gene *Hlh-17* and its expression regulation is similar to the mammalian *Olig2* (Yoshimura et al., 2008). *C.elegans* single-cell data confirms its mostly glial expression, but it is also expressed in neurons (Cao et al., 2017). Therefore, *Olig-like* genes could have emerged as TFs driving glio/neurogenesis and were specified to induce oligodendrocyte differentiation in vertebrates.

Hes

Hes/Hey genes are metazoan specific.They belong to group E of bHLH genes. *Hes* were shown to be involved in many developmental processes, including suppression of proneural bHLH genes, and promotion of gliogenesis (Kageyama et al., 2007; Gazave et al., 2014). Mammalian *Hes* genes are homologs of *Drosophila Hairy* and *Enhancer of Split*. Vertebrate *Hes1*, *Hes3*, and *Hes5* particularly promote gliogenesis at a later stage of the developing brain, but control the production of neural stem cells at an earlier neurogenic stage (Kageyama et al., 2008). Interestingly, *Hes5* is specifically expressed in mammalian Muller glial cells (Hojo et al., 2000). Its expression was shown to be regulated by *Gcm* genes at an early stage to induce neural stem cell generation, which is later replaced by the activation by *Notch* (Hitoshi et al., 2011). *Hes* genes are known effectors of *Notch* in mammals as well as *Drosophila* (Pompa et al., 1997; Kageyama et al., 2008; Kageyama & Ohtsuka, 1999). The *Gcm-Hes-Notch* could therefore be important for gliogenesis in various lineages, even though *Gcm-Notch* synergy driving glia generation in *Drosophila* seems independent from *Hes* (Magadi et al., 2020). Therefore, *Hes* genes have a dual role of promoting neural stem cell and glia generation in mammals, but do not seem to play an important role in invertebrate gliogenesis.

NfI

Nuclear factor I (*NfI*) genes are CCAAT box element-binding TFs (Gronostajski, 2000). In vertebrates, the *NfI* family is composed of four genes: *NfIa, NfIb, NfIc* and *NfIx*. They have been shown to play an important role in the development of various tissues, including the nervous system (Campbell et al., 2008; Gronostajski, 2000). *NfIa, NfIb,* and *NfIc* promote differentiation of radial glial cells into both glia and neurons (Harris et al., 2015). *NfIa* directly induces the expression of glial specific genes (Deneen et al., 2006), and initiates gliogenesis under the regulation of *Sox9* (Kang et al., 2012). *Notch* was shown to induce *NfIa* to drive gliogenesis via *Hes* genes (Deneen et al., 2006). *NfI* are metazoan specific TFs, and in organisms other than vertebrates seem to be represented by a single gene. This is the case in the lancelet *Amphioxus*, *C. elegans*, and *Drosophila* (Fletcher et al., 1999; Gronostajski, 2000; Fletcher et al., 1999; Lazakovitch et al., 2005). *NfI* does not play a role in invertebrate gliogenesis, thus *NfI* family genes could have acquired a gliogenic function after the gene duplication.

SoxE group

SoxE is a group of proteins belonging to a HMG (high mobility group)-box *Sox* family, which plays diverse roles including NS development. In mammals, *SoxE* include *Sox8, Sox9,* and *Sox10*. These are essential for glia generation (Reiprich & Wegner, 2015). *Sox9* is a major neuron-glia switch. It directly regulates the expression of *NfIa,* is indispensable for astrogenesis, and prevents neurogenesis (Glasgow et al., 2014; Kang et al., 2012; Stolt et al., 2003). *Sox8* and *Sox10* play an important role in oligodendrogenesis (Reiprich & Wegner, 2015). *Sox10* interacts with *Olig1* and is induced by *Olig2,* thus driving the expression of myelin genes and suppressing the expression of astrocyte-related genes (Glasgow et al., 2014). In addition to rodents, the gliogenic functions of *SoxE* genes were shown in primitive jawless vertebrates (lampreys) (Yuan et al., 2018). Even though these animals lack oligodendrocytes, they possess the genetic regulatory network required for oligodendrogenesis including *SoxE* and *Olig* orthologs. Gliogenic functions of *SoxE* genes were not reported in protostomes. In crustaceans *SoxE* orthologs regulate gonad and embryo development (Hu et al., 2020; Wan et al., 2021). In *Drosophila* an ortholog of vertebrate *Sox8,9,10* is expressed in the gut and gonads, and required for intestinal epithelium functioning (Jin et al., 2020; Phochanukul & Russell, 2010).The gliogenic function of the class E *Sox* genes seems specific only in mammals or chordates.

Ets family

Members of *Ets* proteins, belonging to group A bHLH TF family, are metazoan-specific and regulate various developmental processes including glial cell differentiation. In *Drosophila* an *Ets* TF, Pointed (*Pnt*), is activated by *Gcm* via *Repo* to induce the expression of glial markers in several glial cell types (Klämbt, 1993; Yuasa et al., 2003). Moreover, *Pnt* drives astrogenesis along with *Notch* signaling, which is also the case in mammals (Peco et al., 2016). Vertebrate homologs of *Pnt* are *Ets-1* and *Ets-2*. They were shown to drive radial glia formation in *Xenopus* (Kiyota et al., 2007). In addition, I extracted two *Ets* family TFs (*Ets1* and *Fli1*) driving glial cell identity (astrocytes+oligodendrocytes) from the mouse single-cell transcriptome data (Tabula Muris Consortium et al., 2018). Other *Ets* family members were reported as driving gliogenesis in PNS and promoting oligodendrocyte proliferation in rodents (Ahmad et al., 2019; Hagedorn et al., 2000). Therefore, *Ets* gliogenic nature is present in both Protostomes and Deuterostomes, and dictates further investigation regarding its conserved function in basal metazoans.

1.2.4.2. Extracellular signaling pathways regulating gliogenesis

In addition to intrinsic factors, extracellular cues are capable of influencing cell fate acquisition. Various signaling pathways have been implied in bilterian gliogenesis (Sauvageot & Stiles, 2002). As stated above *Notch-Delta* signaling is a versatile pathway that plays a major role in gliogenesis in bilaterians (Fischer & Gessler, 2007; Zhou et al., 2010). *Notch* function of maintaining a pool of stem cells in the nervous system in the form of either glio-neuro-precursors or mature glial cells seems conserved among bilaterians. Other signaling pathways driving gliogenesis in vertebrates are *JAK-STAT* (Bonni et al., 1997; He et al., 2005; Lee et al., 2010), which crosstalks with *Notch-Delta* to drive glial differentiation (Kamakura et al., 2004)*, BMP* signaling (Gomes et al., 2003; Gross et al., 1996), and *Hedgehog* (Sauvageot & Stiles, 2002; Wang et al., 2016). Thus, intrinsic as well as extrinsic factors are important drivers of gliogenesis particularly in vertebrates.

1.3 Bilaterian glia

In several model organisms of both deuterostomes and protostomes it was possible to distinguish several glial cell types. The more data are collected the more complex the classification of glia becomes. In vertebrates the term "glia" usually entails macroglia, i.e. radial glia, astrocytes, myelin-producing oligodendrocytes and Schwann cells, and microglia. There are several less numerous populations of glia particularly in mammals such as NG2 glia, known for their expression of a neurexin cell adhesion molecule NG2 and glial progenitor potential (Trotter et al., 2010), pituicytes, tanycytes and others residing in specific brain areas and resembling astrocytes transcriptionally (Chen et al., 2020; Rodríguez et al., 2019; Suess & Pliska, 1981). I will not consider these glial subgroups separately.

In *Drosophila* seven types of glia have been identified (Yildirim et al., 2019), whereas *C.elegans* glia was divided into three groups (sheath glia, socket glia, and mesodermally derived glial cells) (Oikonomou & Shaham, 2011). In other primitive and non-model organisms radial glia is often the only glial type to be reported (Mashanov & Zueva, 2019). On the other hand, more thorough examinations of several non-model bilaterians revealed glial cells actively involved in neurotransmitter metabolism, which is an astrocytic feature (Wang et al., 2016). Therefore, in order to trace back glial origins it is important to consider the characteristics of different bilaterian glial types and see which of them are conserved in more primitive organisms.

1.3.1 Radial glia

Radial glia are neural stem cells and as such are sometimes not considered strictly glial. They express stem cell markers such as *Sox* (Gebara et al., 2016). These cells give rise to neurons and astrocytes in addition to forming a scaffold used by newly generated neurons to travel to their final destination. Radial glia secrete Reissner's fiber components (the major one being *Sco-spondin*) and several markers generally considered specific to astrocytes such as glutamate transporters and intermediate filament proteins (*Gfap* and *Nestin*) (Campbell & Götz, 2002). Radial glia are elongated in shape, and extend long processes through the neuropil **(Figure 1.3)**.

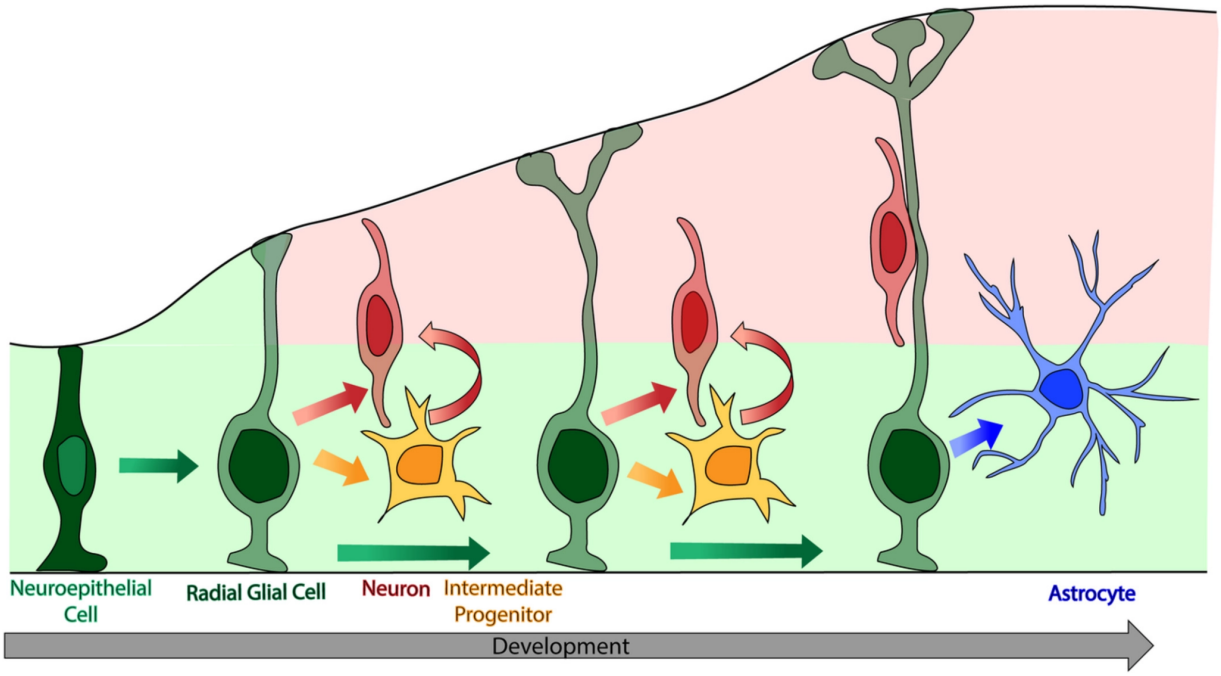


Figure 1.3. Simplified cartoon of mammalian radial glial cells (RGCs) giving rise to neurons and glia. First RGCs produce neurons either directly or via intermediate progenitors. Later in the development RGCs produce glial cells including astrocytes. Image taken from (D'Arcy & Silver, 2020).

They are the earliest type of glia to develop in vertebrates, but are sparse in the adult mammalian brain. This type of glia was reported in both deuterostomes (*Vertebrata*, *Echinodermata*, *Hemichordata*) and protostomes (*Annelida*, *Arthropoda*) (Conrad et al., 2017; Jurisch-Yaksi et al., 2020). It is the only type of glia identified so far in *Echinodermata* (Mashanov & Zueva, 2019), and was shown to fulfill a phagocytic function in addition to its neurogenic and scaffolding functions (Mashanov et al., 2008). Given that radial glial cells are present throughout *Bilateria,* have the potential to generate both neurons and glial cells, and are the first neural cell type to develop, they could have been the first glia to emerge in animals with the nervous system.

1.3.2 Astrocytes/astrocyte-like

Astrocytes are glial cells in the most classical sense. They are closely associated with neurons, not neurons themselves, and are the main cells maintaining the homeostasis of the vertebrate CNS. Astroglia fulfill many functions in the NS such as regulating ionic and neurotransmission composition of the neuronal environment (Boddum et al., 2016; Kimelberg & Nedergaard, 2010); maintaining water homeostasis (Jin et al., 2013); energetic support of neurons (Brown & Ransom, 2007); partitioning of the brain from the rest of the body (Limmer et al., 2014; Michinaga & Koyama, 2019); synaptogenesis (Baldwin & Eroglu, 2017), axon guidance (Joosten & Gribnau, 1989); and phagocytosis (al-Ali & al-Hussain, 1996; Morizawa et al., 2017) **(Figure 1.4)**. Each of these functions is evidenced by the astrocyte-expressing

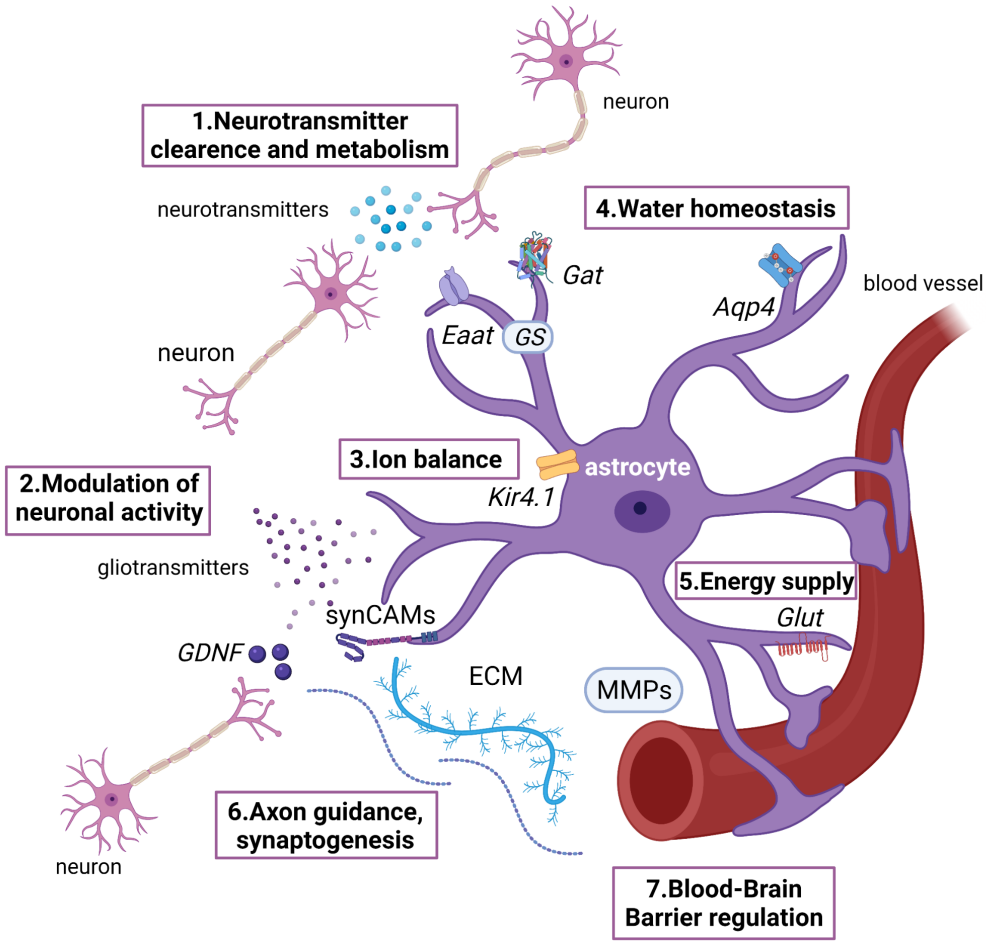


Figure 1.4. Astrocyte functional diversity. Each function is fulfilled by astrocyte-expressed and secreted proteins. Created with BioRender.com.

markers: various ion channels including potassium channels (*Kir4.1*) (Olsen et al., 2015; Shigetomi et al., 2011), GABA and glutamate transporters, glutamine synthetase (*Gat*, *Eaat, Gs*) (Ghirardini et al., 2018; Liang et al., 2008); aquaporin channels (*Aqp4*) (Hubbard et al., 2018); glucose transporters *(Glut)* (Morgello et al., 1995); BBB-regulating factors such as matrix metalloproteinases (*MMP*s) and glial-derived neurotrophic factor (GDNF) (Michinaga & Koyama, 2019); ECM proteins which serve as synaptogenic (e.g. thrombospondins) and axon guiding (e.g. SynCAMs) factors (Frei & Stoeckli, 2014; Christopherson et al., 2005; Frei & Stoeckli, 2014); *ced* pathway components involved in engulfment and phagocytosis (Cahoy et al., 2008) **(Figure 1.4)**. Some of these markers, especially those involved in glutamate metabolism, have been successfully used to identify glial presence in invertebrate bilaterians, e.g. planarias (Wang et al., 2016) **(Figure 1.2)**

No astrocyte-specific TFs were identified per se. *Sox9* and *Nfia/b* are among the major regulators of astroglia development as discussed in 1.2.4.1. The following TFs are also involved in driving astroglial program: *Nfe2l1*(bZIP TF) regulated by *Sox9* (Molofsky et al., 2013), *Klf15* (Kruppel-like family), and *Scl* (bHLH TF) (Fu et al., 2009). A metabolic enzyme *Aldh1l1* (aldehyde dehydrogenase 1 family member L1) and an intracellular signaling

Ras-GCPase *Rab6* are pan-astrocytic markers in rodents (Cahoy et al., 2008; Melzer et al., 2021).

Astrocytes also demonstrate a "classical" glial morphology: multiple processes extending into the neuropil, surrounding synapses and neuronal processes. Based on this criteria in addition to the expression of *Gfap,* astrocyte-like glia was identified in the most basal bilaterians - acoels (Bailly et al., 2014; Bery et al., 2010). Glial cells morphologically and functionally resembling astrocytes are also found in zebrafish (deuterostomes), *C.elegans* and *Drosophila* (protostomes). Despite increasing complexity and heterogeneity of astroglia culminating in the existence of primate-specific subpopulations of these cells (Padmashri et al., 2021; Vasile et al., 2017), astrocyte-like glial cells are present throughout bilaterian lineages. Thus, astroglia could be the first true glia type to emerge in evolution.

1.3.3 Oligodendrocytes, Schwann cells, wrapping glia

Myelin production is the major function of oligodendrocytes in the CNS and Schwann cells in the PNS of vertebrates. Myelin appearance correlates with the jaw development, thus making it a new invention from the evolutionary point of view (Verkhratsky et al., 2019). Myelin-associated proteins are thus used as oligodendrocyte and Schwann cell markers (Hai et al., 2002; Scolding et al., 1989). On the other hand, oligodendrocyte lineage markers are present in those species that do not possess myelin. These include TFs that drive oligodendrogenesis such as *Sox10* and *Olig2* among others, as well as a well-known marker of oligodendrocyte progenitors - platelet derived growth factor receptor (*PDGFRα*) (Yuan et al., 2018) **(Figure 1.5)**. This might indicate a common gliogenic program going back to more primitive bilaterians.

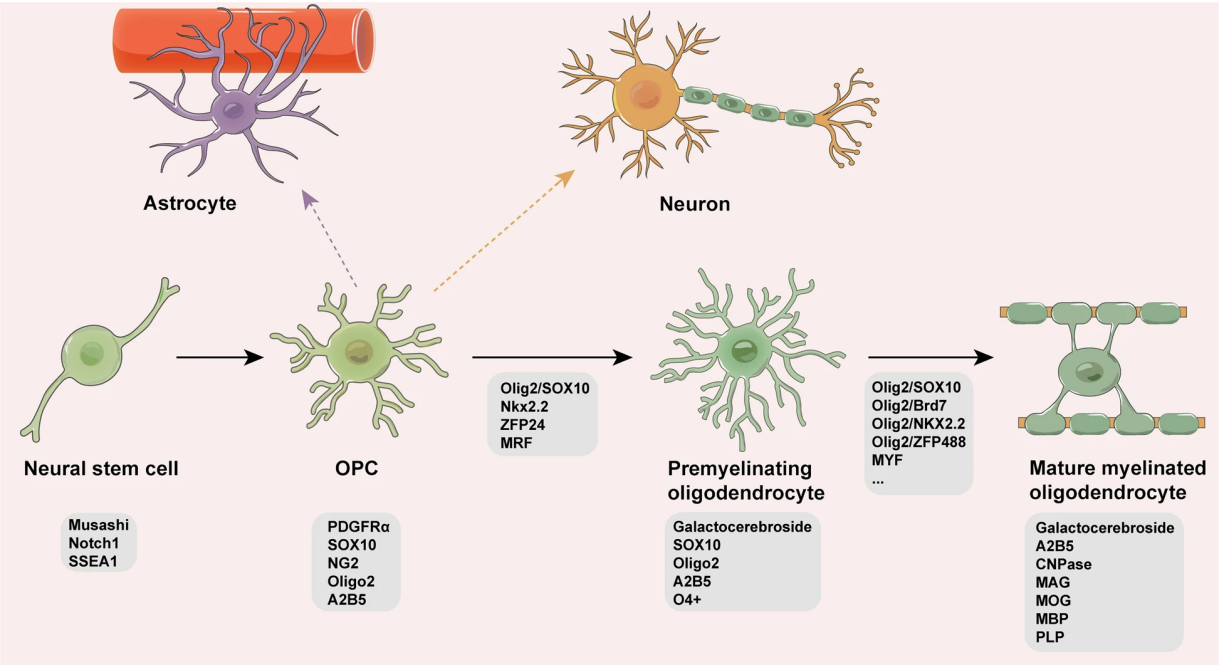


Figure 1.5. Developmental program of mammalian oligodendrocytes. Neural stem cells give rise to oligodendrocyte progenitor cells (OPCs), which begin to express oligodendrocyte markers (*Sox10*, *PDGFRa*). OPCs have a potential to differentiate into astrocytes, neurons, and oligodendrocytes, although the former two signaling pathways are under debate. *Olig2* and *Sox10* genes among others

Caption continuation for Figure 1.5
drive oligodendrocyte maturation. Myelin constituents such as MBP (myelin basic protein) and PLP (proteolipid protein) are expressed by mature myelinated oligodendrocytes. Image taken from (Zhou et al., 2021).

Even in vertebrates not all oligodendrocytes produce myelin. The common feature that oligodendrocytes do share, however, is axonal ensheathment. Glial cells covering axons with their membranes were found in various animals including protostomes (Verkhratsky et al., 2019). It is possible that these cells are primitive oligodendrocytes, even though unless specific markers are detected it's hard to define them. Functionally, similar to astrocytes, oligodendrocytes were shown to express glutamate and GABA transporters and provide metabolic support to neurons (Seifert & Steinhäuser, 2018). They form networks with each other and astrocytes via gap junctions contributing to ionic buffering during neuronal activity (Wasseff & Scherer, 2011). Therefore, distinguishing between oligodendrocytes and astrocytes might be easy only in more advanced animals.

When considering the evolutionary roots of glia, it is helpful to look at the gliogenic vs. neurogenic program. It is generally believed that oligodendrocytes arise from the same progenitors as neurons (Lu et al., 2002). If this was the case only the astrogenic program should be searched for in more primitive organisms to uncover glial evolution. However, some progenitors produce both oligodendrocytes and astrocytes (Chari, 2002). Radial glia was shown to produce all three cell types depending on the developmental stage (Bergles & Richardson, 2015). OPCs might also generate all three cell types **(Figure 1.5)**. Therefore, it's rather difficult to separate out the strictly gliogenic program, particularly in vertebrates.

1.3.4 Microglia

Unlike macroglia which originate in the ectoderm, vertebrate microglia have a mesodermal origin. Similarly *C.elegans* possesses 6 mesodermally derived glial-like cells in the nerve ring (GLR) serving as connections between neurons and muscles (Oikonomou & Shaham, 2011) **(Figure 1.6A)**. The function of GLR is not clear, but it was observed to engulf dead cells (Stout, 2014). In mammals microglia perform a distinct function of being resident macrophages of the CNS (Ginhoux & Prinz, 2015). Because one of the main functions of microglia is debris clearance and degradation, markers of phagocytic pathways such as *P2ry12* and lysosomal enzymes such as *Hesb* are abundantly expressed by these cells (Cahoy et al., 2008; Mazzolini et al., 2020) **(Figure 1.6B)**.

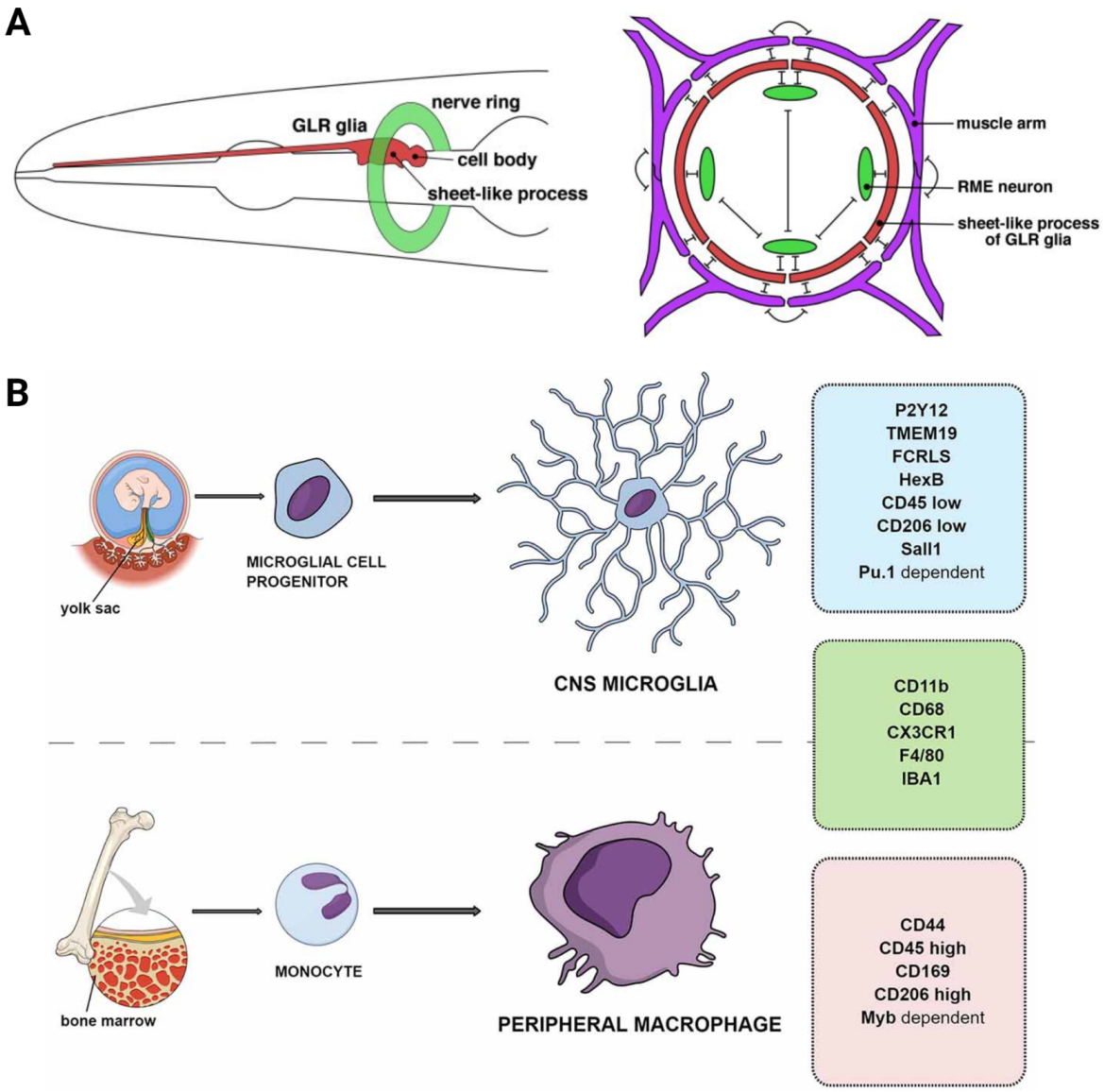


Figure 1.6. Mesodermally derived glial cells in bilaterians. (A) Left: one out of six GLR glia cells of *C.elegans* is depicted. A sheet-like process of the GLR cell lines the inside of the nerve ring. Right: GLR glial sheet-like processes make gap junctions to muscle cells and neurons. Image taken from (Oikonomou & Shaham, 2011). (B) Mammalian microglia and peripheral macrophage ontogeny and markers. The two cell types share a majority of markers, although their origins differ. Image taken from (Jurga et al., 2020).

To this end microglia and astrocytes have the same function of synapse pruning and engulfment (Vainchtein & Molofsky, 2020). Similarly, both cell types become reactive in response to injury and inflammation. Even more so astrocytes assume the role of phagocytes in case of microglial dysfunction (Konishi et al., 2020). Similar to other neural cells, vertebrate microglia express a number of ion channels, neurotransmitter receptors and transporters (Kettenmann et al., 2011). In general, however, the transcriptome of microglia differs significantly from that of macroglia: it is enriched with markers related to the immune

system processes and macrophages (e.g. *CD45*, *CD68*), but also microglia-specific markers (e.g. *TMEM119*) (Bennett et al., 2016; Mazzolini et al., 2020) **(Figure 1.6B)**. The transcriptional program driving microglia identity is drastically different from macroglia (Holtman et al., 2017). *Spi1* (or *Pu.1* - *Ets*-domain TF) and *Irf8* (interferon regulatory factor family) are the main TFs driving microgliogenesis in various vertebrate species (Geirsdottir et al., 2019).

Morphologically resting microglia resembles astroglia: cells extend processes from the central soma. Upon activation in response to injury microglia change their morphology and aggregate at the lesion site. The complexity of the ramified structure of microglia varies with the overall trend of increasing in evolution. However, unlike astrocytes human microglia do not display the most complex morphology compared to other vertebrates (Geirsdottir et al., 2019). Apart from vertebrates, insects (*Arthropoda*), leech (*Annelida*), and molluscs were reported to have microglial cells (Verkhratsky et al., 2019). Among these leech microglia have been studied most thoroughly albeit using few molecular markers (Sharma et al., 2021). Surprisingly, no other glial cell types were identified in the leech. It is not clear to what extent the vertebrate microglial molecular program and functions are conserved in invertebrates due to the limited studies available. Comparative analysis of microglia markers across species should clarify the evolutionary origins of these cells. Meanwhile, given its origin and a distinct immune function, it is unlikely that glia ancestral cells were microglia-like.

Figure 1.7 summarizes glial cell type conservation among bilaterians.

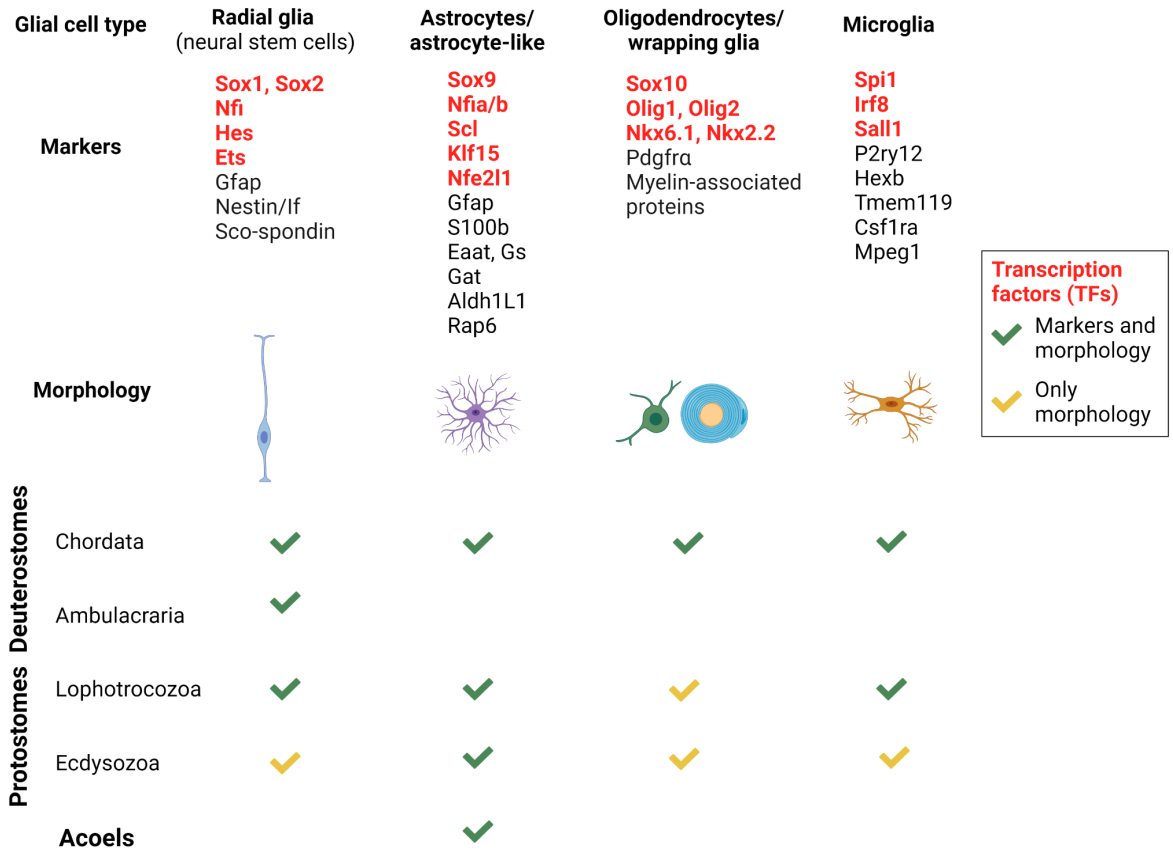


Figure 1.7. Glial cell type conservation in bilaterian phyla. Molecular markers and morphological features of major glial cell types are shown. Created with BioRender.com.

With the emergence of molecular identifiers for each glial cell type it has become possible to identify various glial cells in bilaterian lineages not limited to vertebrates and insects. Nevertheless, few glial markers have been searched for in non-model organisms. Based on the data obtained mostly from the histological studies it is likely that radial glia was the first glial type to evolve. These cells combine the features of both neurons and glia and have the potential to give rise to both. The first true glia to emerge could have been astrocyte-like which assumed several key functions including the regulation of ionic and neurotransmitter composition of the neuronal environment and debris clearance. Glial cells reminiscent of microglia and oligodendrocytes (wrapping glia) can be found in primitive bilaterians, but as such were likely to have emerged later in evolution **(Figure 1.8)**.

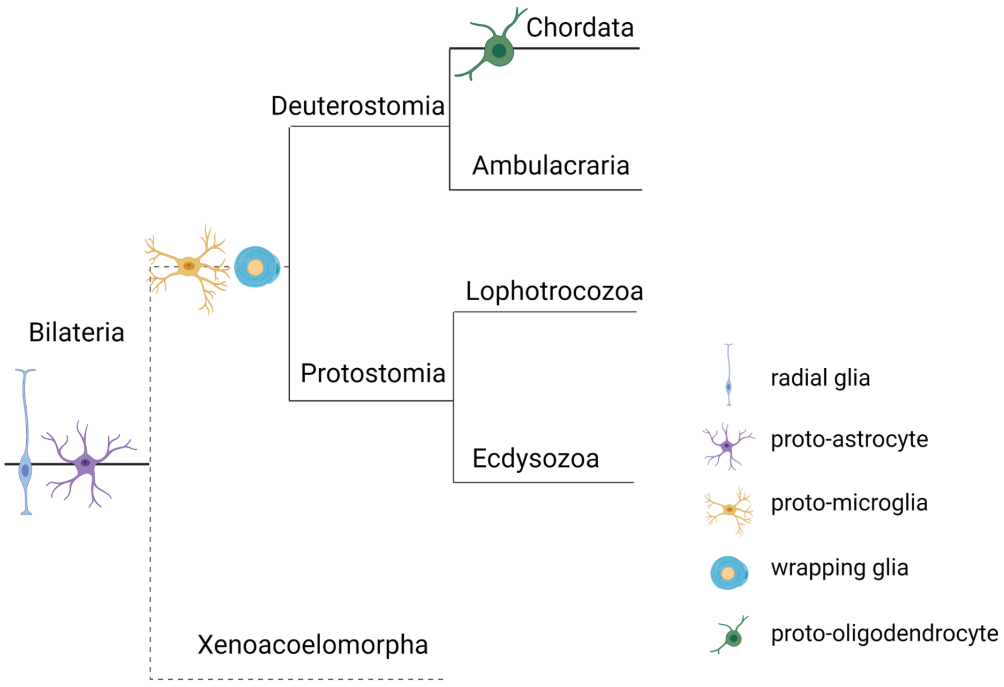


Figure 1.8. Evolutionary tree of glia cell types based on data from bilaterians. Created with BioRender.com.

1.4. Summary and approach

On the basis of the available data it is hypothesized that 1) glia emergence was dictated by the centralization of the nervous system and 2) glial evolved multiple times independently (Hartenstein & Giangrande, 2018; Hartline, 2011; Verkhratsky & Butt, 2013). As discussed above these conclusions are drawn from the data collected mostly from a few model organisms. The glia of some basal bilaterian lineages have been investigated, but no comprehensive molecular or functional analysis was done. It is acknowledged that in order to recreate the evolution of glia, more data should be collected, especially from basal bilaterians (Hartline, 2011). I propose that to recreate the full story of glial evolutionary processes it is necessary to investigate animals other than bilaterians.

Instead of screening various bilaterian lineages, exploring glial feature conservation in animals with the most primitive nervous system, *Ctenophora* and *Cnidaria*, is indispensable to reconstruct the evolution of glia. Surprising molecular and functional complexity of their nervous systems has already been shown (Sebé-Pedrós, Chomsky, et al., 2018; Sebé-Pedrós, Saudemont, et al., 2018). Although no cells morphologically or transcriptionally similar to bilaterian glia have been identified in these animals, no exploration of the bilaterian gliogenic program conservation was carried out in the basal metazoans. Similar to neural modules predating neurons (Liebeskind et al., 2017), certain glial features could have been present in animals in the absence of glia per se. The next chapter explores to what extent the bilaterian glial genetic program is conserved in basal metazoans. Cnidarians are of particular interest, since they occupy a privileged position of a sister group to all bilaterians.

1.4.1 *Nematostella vectensis* as a model cnidarian

Nematostella vectensis is one of the most widely used cnidarian model organisms to address various questions about the evolution of the nervous system (Layden et al., 2016; Watanabe et al., 2014). *Nematostella* is easy to culture as it spawns with increased light and temperature. *Nematostella* is amenable to genetic manipulations thanks to the external fertilization. Males and females are kept in separate boxes for precise timing control of the fertilization process. Females release eggs encapsulated in a gelatinous mass, which can be later fertilized using sperm-rich brackish water obtained from males. Fertilized eggs cleave, organize into a one-dimensional layer of cells called blastula, then transform into a gastrula by invagination **(Figure 1.9)**. After that a planula larva develops into a sessile tentacle bud, and, finally, by 8-9 dpf, a juvenile polyp with tentacles and mesenteries **(Figure 1.9)**. *Nematostella* is diploblastic: all the cells develop from either endoderm or ectoderm (Martindale et al., 2004).

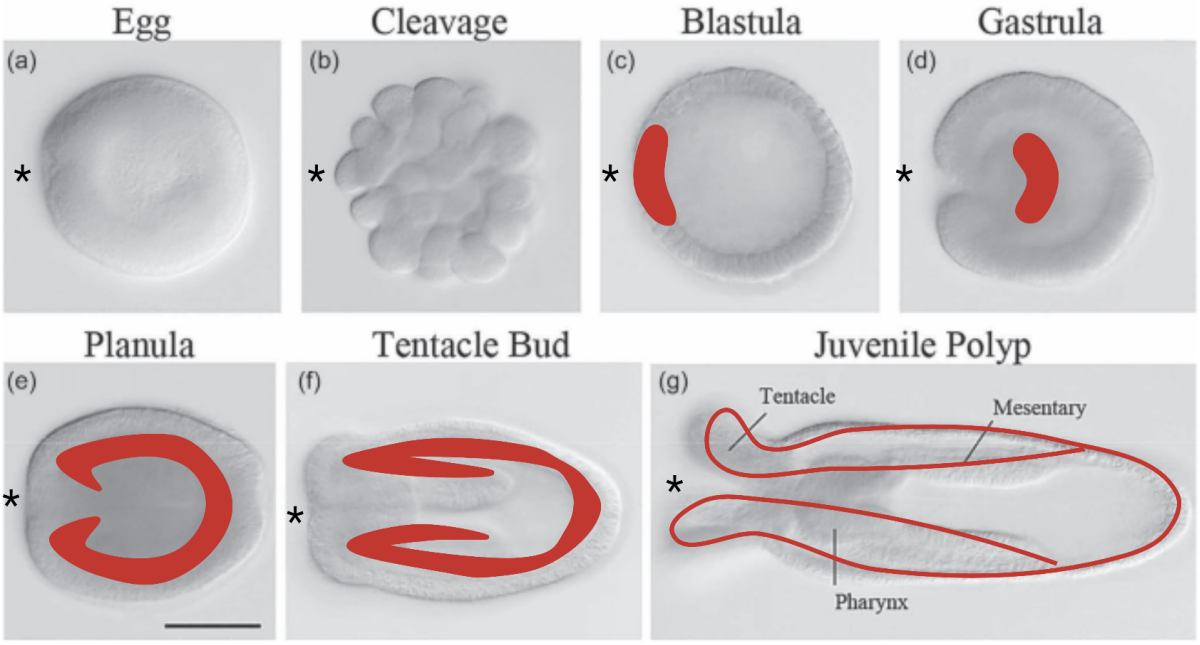


Figure 1.9. Developmental stages of *Nematostella vectensis*. Endoderm is labeled with red. Oral side is marked with an asterisk. Scale bar =100 µm. Modified from Al-Shaer et al., 2021.

Various molecular techniques and tools can be applied to *Nematostella*. Its genome and single-cell transcriptome are now available (Putnam et al., 2007; Sebé-Pedrós et al., 2018; Zimmermann et al., 2020). Gene expression analysis using *in situ* hybridization and immunostaining have been successfully done (Genikhovich & Technau, 2009; Marlow et al., 2009; Wolenski et al., 2013). In addition, gene manipulation techniques for gain and loss of function experiments, as well as transgenic lines have been established (Layden & Martindale, 2014; Renfer et al., 2010). Therefore, it is possible to design experiments combining several techniques using *Nematostella* **(Figure 1.10)**. Despite a plethora of methods available, the establishment of some techniques is still in its infancy such as cell culture development (Nowotny et al., 2021). Nevertheless, I suggest that *Nematostella* could be a useful model system to get an insight into the deep evolutionary roots of glia as argued in the next chapter.

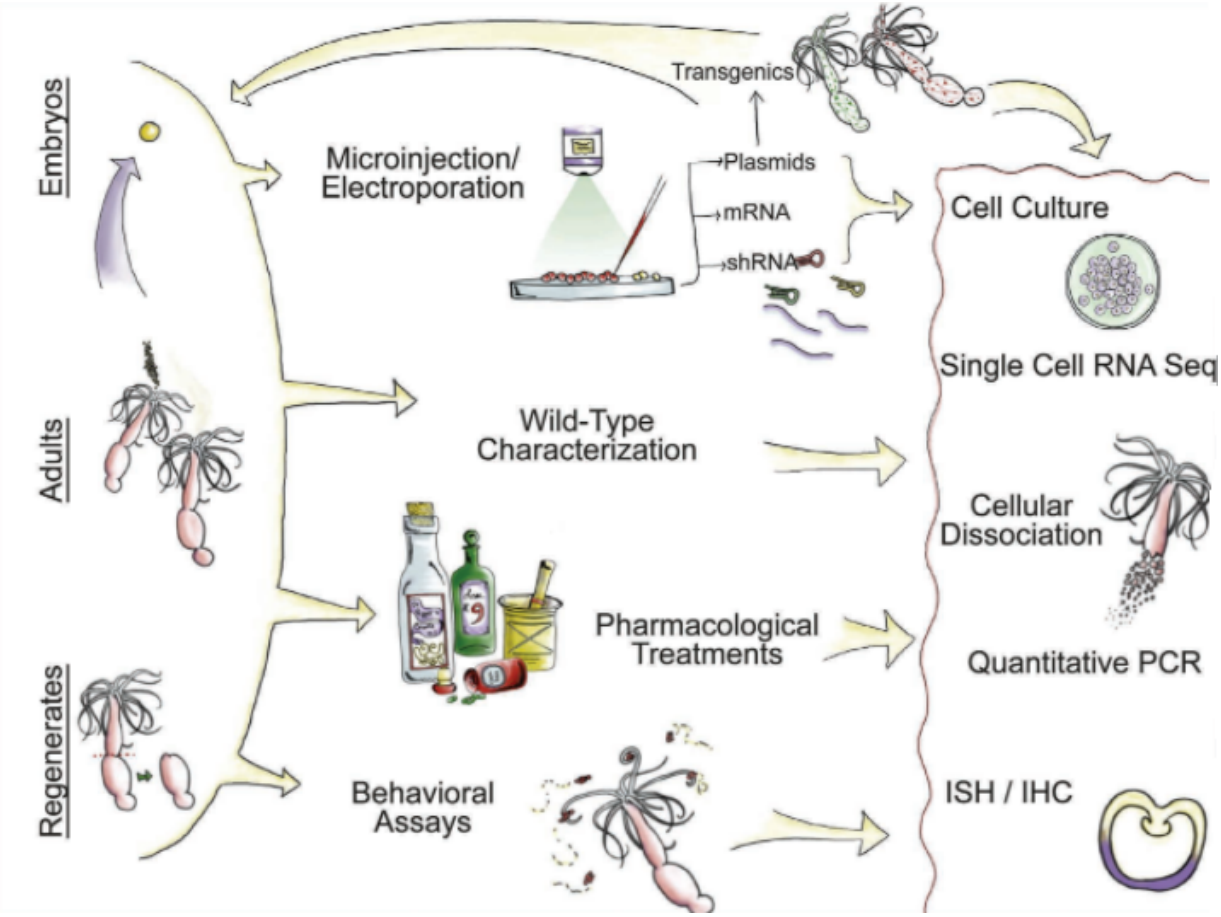


Figure 1.10. Various techniques and their potential integration for experiments in *Nematostella*. Taken from Al-Shaer et al., 2021.

Chapter II. Phylum-wide and genome-wide survey of gliogenic factors in basal metazoans

2.1 Background

Basal metazoans are non-bilaterian animals which diverged early in the evolution. These include *Placozoa*, *Porifera* (sponges), *Ctenophora* (comb jellies), and *Cnidaria* **(Figure 1.1)**. While *Cnidaria* firmly occupies the position of the closest sister group to *Bilateria*, the phylogenetic relationships among the rest of basal metazoans are still debated. Currently the *Ctenophora*-first hypothesis is prevailing (Kapli & Telford, 2020; Li et al., 2021; Shen et al., 2017). Unlike bilaterians, basal metazoans have two embryonic germ layers: endoderm and ectoderm with a non-cellular region called mesoglea in between (Schmidt-Rhaesa et al., 2015). *Ctenophora* seems an exception since their mesoglea gives rise to tissues including neurons suggesting a triploblastic nature (Jager et al., 2011; Schmidt-Rhaesa et al., 2015). Recently these phyla have been actively studied particularly in the context of the nervous system evolution. Thanks to the advancements in sequencing and molecular techniques in these animals it has been possible to address questions regarding the NS of ancestral metazoans. Prior to investigating gliogenic program conservation in basal metazoans, it is paramount to characterize the nervous system of these animals as discussed below.

Placozoans and poriferans do not possess neurons but their genomes contain pro-neural TFs and encode proteins required for synapse formation and neurotransmitter synthesis (Srivastava et al., 2008, 2010). Therefore, studies focused on unraveling the functions of neuro-associated genes in these animals are expected to shed light on the evolution of the NS. In addition, nerveless *Placozoa* and *Porifera* show coordinated behavior usually associated with the nervous system. They display muscle contractile response to fast neurotransmitters such as glutamate, GABA, and glycine (Romanova et al., 2020; Ryan & Chiodin, 2015). Nevertheless their movements are slow compared to other animals and in the absence of the NS the neural function of these chemical transmitters is not possible. Neuropeptides have not been found in sponges. Meanwhile *Placozoa* behavior was shown to be regulated by neuropeptides (Senatore et al., 2017; Varoqueaux et al., 2018). This data suggest that synaptic signaling is pre-dated by peptidergic signaling **(Table 2.1)**.

NS feature	***Porifera***	***Placozoa***	***Ctenophora***	***Cnidaria***	***Bilateria***
Neurons	-	-	+	+	+
NS structure	-	-	Nerve net with regional condensation sites	Nerve net with regional condensation sites	Centralized
Neurogenesis	-	-	Ectoderm, mesoglea	Ectoderm, endoderm	Ectoderm
Neuropeptides	-	+	+	+	+
Neuronal function of small chemical neurotransmitters - Glutamate - GABA - Glycine	 - - -	 - - -	 (+) - (+-)	 + + +	 + + +

Table 2.1. Neuronal features in metazoans. Nerveless phyla are highlighted in yellow, phyla possessing neurons are highlighted in blue.

Indeed, the first nervous systems probably heavily relied on peptidergic signaling as is evidenced by an extensive repertoire of neuropeptides in the basal metazoans with the nervous systems, *Cnidaria* and *Ctenophora* (Hayakawa et al., 2019; Sachkova et al., 2021).

Cnidaria and *Ctenophora* both possess neurons organized in nerve nets with regional condensations **(Table 2.1)**. These two phyla used to be considered closely related and known under a common name *Coelenterata*. However, phylogenomic analyses placed *Cnidaria* as a sister group to all *Bilateria*, whereas *Ctenophora* is one of the early-branching lineages of basal metazoans. It is hypothesized that *Ctenophora* evolved its nervous system independently from all other animals due to its genetic, molecular and structural uniqueness (Norekian & Moroz, 2020). The authors found that many homologs of the bilaterian neuronal genes that exist in *Cnidaria* seem absent in *Ctenophora*. In addition, no conserved neuropeptides were identified in *Ctenophora* (Sachkova et al., 2021). Structural organization of the ctenophoran NS also differs: it is composed of epidermal and mesogleal nerve nets with numerous condensations. Some neurons demonstrate a unique feature of loops formed by neurites (anastomoses) (Sachkova et al., 2021). Nevertheless, certain similar features of the NS in *Bilateria, Cnidaria* and *Ctenophora* can be observed: conserved neurogenic TFs (SoxB, bHLH) and synaptic genes, glutamate as a fast neurotransmitter. Whether chemical neurotransmitters are recruited in *Ctenophora* is not known, but current evidence suggests that glutamate and glycine, but not GABA, might be involved in neuronal transmission (Moroz et al., 2021).

Cnidaria is a highly diverse phylum composed of Class *Anthozoa* (sea anemones, corals, sea pens), and *Medusozoa* (including Classes *Hydrozoa, Cubozoa,* and *Scyphozoa*). The former are considered the basal group within *Cnidaria* (Collins, 2002; Kayal et al., 2013). They have fewer derived features than *Medusozoa* including the absence of a free swimming medusa stage and less elaborate nematocysts, an organelle encapsulated within *Cnidaria*-specific stinging cells (Oren et al., 2014; Sebé-Pedrós et al., 2018; Watanabe, 2017). *Cnidaria* are thought to be descendants of the animals with the first nervous systems. Unlike *Ctenophora*, *Cnidaria* share all key features of the bilaterian NS including a diverse repertoire of neurons and neurotransmitters (Oren et al., 2014; Sebé-Pedrós et al., 2018; Watanabe, 2017). In addition, cnidarians possess almost all the homologs of bilaterian genes driving neurogenesis and neuron specification (Oren et al., 2014; Sebé-Pedrós et al., 2018; Watanabe, 2017).

Cnidarian NS generally consists of a nerve net and regional condensations in the oral ("nerve ring") and aboral regions. The development of neurons takes place in both ectoderm and endoderm **(Table 2.1)**. Morphologically two neuronal cell types are distinguished: sensory neurosecretory cells characterized by an elongated cell body extending a thin process and ganglionic neurons (Watanabe et al., 2009). Cnidarian NS is rich in neuropeptides. In addition, classical chemical neurotransmitters such as nitric oxide (NO), glutamate, GABA, glycine are involved in neural functions in cnidarians **(Table 2.1)**. Small transmitters including glutamate can perform both non-neuronal and neuronal functions in the same organism, which might be the case in *Cnidaria*. Regardless, glutamate and glycine might have gotten recruited as neurotransmitters by neurons at some point in the evolution (Moroz, Romanova, et al., 2021). More functional studies are required to understand whether it took place with the emergence of the first neurons, however. It is also debated if acetylcholine and monoamines function as neurotransmitters in *Cnidaria* since a complete gene set of the canonical pathways for the synthesis of these molecules is absent (Moroz, Romanova, et al., 2021).

To sum up, neuronal genes and modules can be found even in basal metazoans without nervous systems as described above. This suggests their recruitment in neuronal functions later on during evolution. Similarly, glial genes and modules might be present in non-glial and/or nerveless organisms. At the same time, of all the basal metazoans the nervous system of *Cnidaria* is characterized by a high degree of similarity to the bilaterian NS. The critical question we should ask here is to what extent the bilaterian gliogenic program is conserved in basal metazoans. Is it also the case that the program driving gliogenesis in *Bilateria* is highly conserved specifically in *Cnidaria?* To answer these questions, I screened available genomes and transcriptomes of basal metazoans for the homologs of bilaterian glial TFs and members of signaling pathways known to drive gliogenesis in bilaterian animals. Glial and neuronal developmental programs are tightly intertwined, which complicates the identification of strictly glial genes. Nevertheless, as discussed in section 1.2.4.1 and 1.3 several specific glial TFs and effector genes are present throughout *Bilateria*.

In this chapter, using comparative sequence analysis complemented by available literature, I identified the orthologs of bilaterian glial TF in basal metazoans. It is common not to find specific orthologs for bilaterian genes in early branching animals. Orthologous family members should be analyzed instead. Functions of these basal metazoan TFs can be assumed from their sequence similarities to the bilaterian orthologs, but expression patterns and effector genes must be considered for accurate functional assessment. Therefore, in addition to sequence similarities, I explored the expression patterns of some of these glial TF orthologs.

I show that all the constituents of the bilaterian gliogenic program are indeed conserved in *Cnidaria*. In addition, I identified *Gcm* as the "core" bilaterian glial TF conserved only in *Anthozoa* (*Cnidaria)*. *Gcm* is expressed in a subset of neuronal cells in a cnidarian model organism, *Nematostella vectensis*. In the next chapter I focused on the functional analysis of *Nematostella Gcm* by identifying its targets, and describing the morphology of *Gcm*-expressing cells.

2.2 Materials and methods

2.2.1 Phylogenetic analysis of *Gcm*

Bidirectional BLAST (Altschul et al., 1990; States & Gish, 1994) searches using *Drosophila melanogaster Gcm1* protein sequence were done against databases of representative bilaterian and basal metazoan organisms (see appendix 2A). Fungal protein sequences with the highest similarity to animal *Gcm* were used as an outgroup. For the *Cnidaria*-specific phylogenetic tree several databases were used to retrieve the sequences (see appendix 3A).

The sequences were aligned using MUSCLE (Mega7) and trimmed by eye to include only the domain. The trees were constructed with maximum likelihood (ML) method using PhyML (SeaView). Bootstrap support is based on 2000 replicates. See appendix 2B and 3B for complete alignments.

2.2.2 *Nematostella vectensis* single cell data analysis

Raw unique molecule identifier (UMI) counts for each gene were taken from single-cell transcriptome data published by (Sebé-Pedrós, Saudemont, et al., 2018). For gene per cell type expression analysis, total UMI for each gene were summarized across cell clusters of the same cell type, and calculated as a fraction of the total expression (UMI). For gene per cell cluster (metacell) expression analysis, molecules/1,000 UMIs transformation was carried out, i.e. raw UMI of gene/ total UMI of cell cluster x 1,000. Genes with total UMI in all cell clusters < 10 were not included in the analysis. Expression analysis was done for the genes of interest for larval (pooled 2dpf, 4dpf, 7dpf animals) and adult (5 months polyps) stages separately.

2.2.3 *Nematostella vectensis* culture

Nematostella vectensis were cultured in brackish water (108.32 g l−1 Tropic Marine sea salt, pH 7.6) at 18 °C in the dark and fed two-three times a week. Females and males were cultured in separate boxes, which were washed on the day following the feeding day. To stimulate spawning female and male boxes were placed in incubation chambers at 26°C under the light overnight. Eggs were collected and fertilized using the sperm-containing water and cultured at RT.

2.2.4 Whole-mount *in situ* hybridization (WISH)

Conventional *in situ* hybridization was performed as previously described (Martindale et al., 2004) using Intavis Automated In-Situ Hybridization System. NBT/BCIP (Figure 2.3E

left) or BM purple (Figure 2.3E right, 2.3F) were used as color reaction substrates with similar results. DIG labeled RNA gene-specific antisense and sense (control) probes were constructed. The following primers were used to generate the probes designed with Geneious:

	Forward (5'-3')	Reverse (5'-3')
Gcm	GCTGTGGACGAGTTTGACGA	CAGGTGACGTGTGCGATTAG
Repo	GCCTGAAGAGTCGGTCAACA	AACTCCACCTTCTCGCTTCG

To generate the probes, total RNA was extracted from 4dpf planulae using RNeasy Mini Kit (per manufacturer's guidelines; QIAGEN). NanoDropTM 1000 Spectrophotometer (Thermo Fisher Scientific) for the RNA concentration and purity check. Agarose gel was also run to assess RNA integrity. Single-stranded cDNA was synthesized from 300-500 ng of total RNA in a final volume of 20 µL by using oligo(dT)20 and SuperScript IV Reverse Transcriptase (Invitrogen) according to manufacturers' instructions. cDNA was stored at −20 °C for future use. PCR was carried out using GoTaq® DNA Polymerase (Promega). The DNA product was subcloned using pGEM-T Easy Vector (Promega). BigDye™ Terminator v3.1 Cycle Sequencing Kit (ThermoFisher Scientific) was used to prepare the samples for sequencing. The samples were sequenced using 3730 DNA Analyzer (Applied Biosystems, ThermoFisher Scientific).

2.3 Results

2.3.1 Orthologs of glial transcription factors in basal metazoans

2.3.1.1 Nuclear Factor I (*NfI*) genes

TFs belonging to the *NfI* family were shown to have a gliogenic function in vertebrates. As discussed in section 1.2.4.1 a single member of the *NfI* family is present in invertebrates and does not seem to play a role in gliogenesis. Similarly, in *Nematostella* a single *NfIx*-like gene was identified **(Table 2.2)**. It is expressed in the central region at blastula stage, i.e. before the onset of neurogenesis, and is implied in endomesoderm specification (Röttinger et al., 2012). It is unclear if *NfIx*-like gene in *Nematostella* is expressed beyond this developmental stage and what other functions it might have. A single *NfI-like* gene could also be identified in *Porifera* and *Placozoa* **(Table 2.2)**. To date no expression pattern of this gene was reported in these animals.

Glial features	*Porifera*	*Placozoa*	*Ctenophora*	*Cnidaria*	*Bilateria*
Gliogenesis	-	-	?	?	Ectoderm
Glial TFs					
- *Sox (group E)*	+	+	+	+	+
- *Hes* (bHLH E)	+	+	+	+	+
- *Olig/Beta3* (bHLH A)	(+)	(+)	(+)	(+)	+
- *Ets*	+	+	+	+	+
- *Nfl*	+	+	-	+	+
- *Gcm*	-	-	-	+	+
- *Repo*	-	-	-	+	+*
Pathways involved in gliogenesis					
- *Delta-Notch*	+	+	-	+	+
- *BMP*	-	+	+	+	+
- *JAK-STAT*	-	+	-	+	+
- *Hedgehog*	-	-	-	+	+

Table 2.2. Glial genes in metazoans. Nerveless phyla are highlighted in yellow, phyla possessing neurons are highlighted in blue. The transcription factors that seem to be related to neural sophistication are shown in red. Asterisk: *Repo* gene is known to be present in some insects, but not in other bilaterians belonging to *Lophotrochozoa* and *Deuterostomia*.

2.3.1.2 *SoxE* genes

Certain members of the *SoxE* group drive gliogenesis in bilaterians. *SoxE* genes are metazoan-specific and can be found in the genomes of all the basal metazoans (Jager et al., 2006; Miller & Ball, 2008; Schnitzler et al., 2014)) **(Table 2.2)**. In sponges *SoxE* is expressed in choanocytes - flagellum-containing cells filtering particles out of the water (Fortunato et al., 2012). Endodermal expression of two *SoxE* genes was shown in both *Ctenophora* (Jager et al., 2006; Miller & Ball, 2008; Schnitzler et al., 2014) and *Cnidaria* (DuBuc et al., 2014; Magie et al., 2005; Shinzato et al., 2008). The broad expression pattern in the endoderm does not seem to support its involvement in neural function.

2.3.1.3 Gliogenic *bHLH* genes: group A and E

Well-known bilaterian gliogenic TFs *Olig* and *Hes* belong to group A and E of bHLH family respectively. Both bHLH groups are metazoan-specific. Several members of group A (including *Atonal*) and a single member of group E (*Hes/Hey*) of bHLH genes were identified in *Porifera* (Fortunato et al., 2016; Zhou et al., 2012) **(Table 2.2)**. Poriferan *Atonal* is expressed in sensory cells and has a strong proneural activity as evidenced by functional studies in bilaterians (Richards et al., 2008). O*ligo/Beta3-like* and *Hes* ortholog were also identified in *Placozoa* (Gyoja, 2014). Three *Hes* genes are present in the genomes of

Ctenophora **(Table 2.2)**. It is therefore assumed that bHLH group A and E emerged in basal metazoans. However, a significant expansion of these genes occurred in *Cnidaria*. Thirty group A genes and eleven *Hes* copies are present in the genome of a sea anemone *Nematostella.* It should therefore be easier to identify homologs of specific gliogenic bilaterian bHLH genes in *Cnidaria*. Unfortunately, it is not always the case (Gyoja, 2017). Two *Olig-like* genes could be identified in *Nematostella.* One of them is expressed in the oral region of endoderm (Marlow et al., 2009; Watanabe et al., 2014). The expression pattern of another *Olig-like* gene is not known. The expression of *Hes* genes varies (Magie et al., 2005; Marlow et al., 2012). Interestingly, one of *Hes-like* genes (*Nvhl3*) is strongly expressed in a subset of cells, which is reminiscent of the *Gcm* expression pattern (see below).

2.3.1.4 *Ets* family

Ets family genes share a common C-terminal DNA-binding domain and are metazoan-specific (Degnan et al., 1993) **(Table 2.2)**. A subset of *Ets* proteins contains an N-terminal domain known as *Pointed* (Yordy & Muise-Helmericks, 2000). Both vertebrate (*Ets1, Ets2*) and invertebrate (*Pointed*) genes belonging to the *Pointed Ets* group have a gliogenic function (Albagli et al., 1996; Kiyota et al., 2007). In a cnidarian *Nematostella* the expression and a gene regulatory networks (GRN) of a *Pointed*-containing *Ets* gene (*NvERG*) was analyzed (Amiel et al., 2017; Larroux et al., 2006). *NvErg* GRN includes but is not limited to NS components such as *SoxB*, neuropeptides and other members of *Ets* family. Among the members of an apical pole GRN of *NvErg* are the following orthologs of bilaterian glial TFs: *Hes*, *SoxE genes,* and *Gcm*. The authors identified 12 *Ets* genes in *Nematostella*, but the expression and functions of most of them are unknown. Importantly, not only *Pointed* domain-containing bilaterian *Ets* members were shown to have a gliogenic role. As the neuron-less animals like *Porifera* and *Placozoa* possess *Ets* genes, the function of ancestral *Ets* genes should be independent from the nervous (glial) system.

2.3.1.5 *Gcm* and *Repo*

Gcm is a master regulator of gliogenesis in *Drosophila*, which also shows a gliogenic potential in vertebrates (see 1.2.4.1). The phylogenetic analysis of *Gcm* I performed revealed that among basal metazoans *Gcm* domain is highly conserved in *Cnidaria* **(Figure 2.1)**.

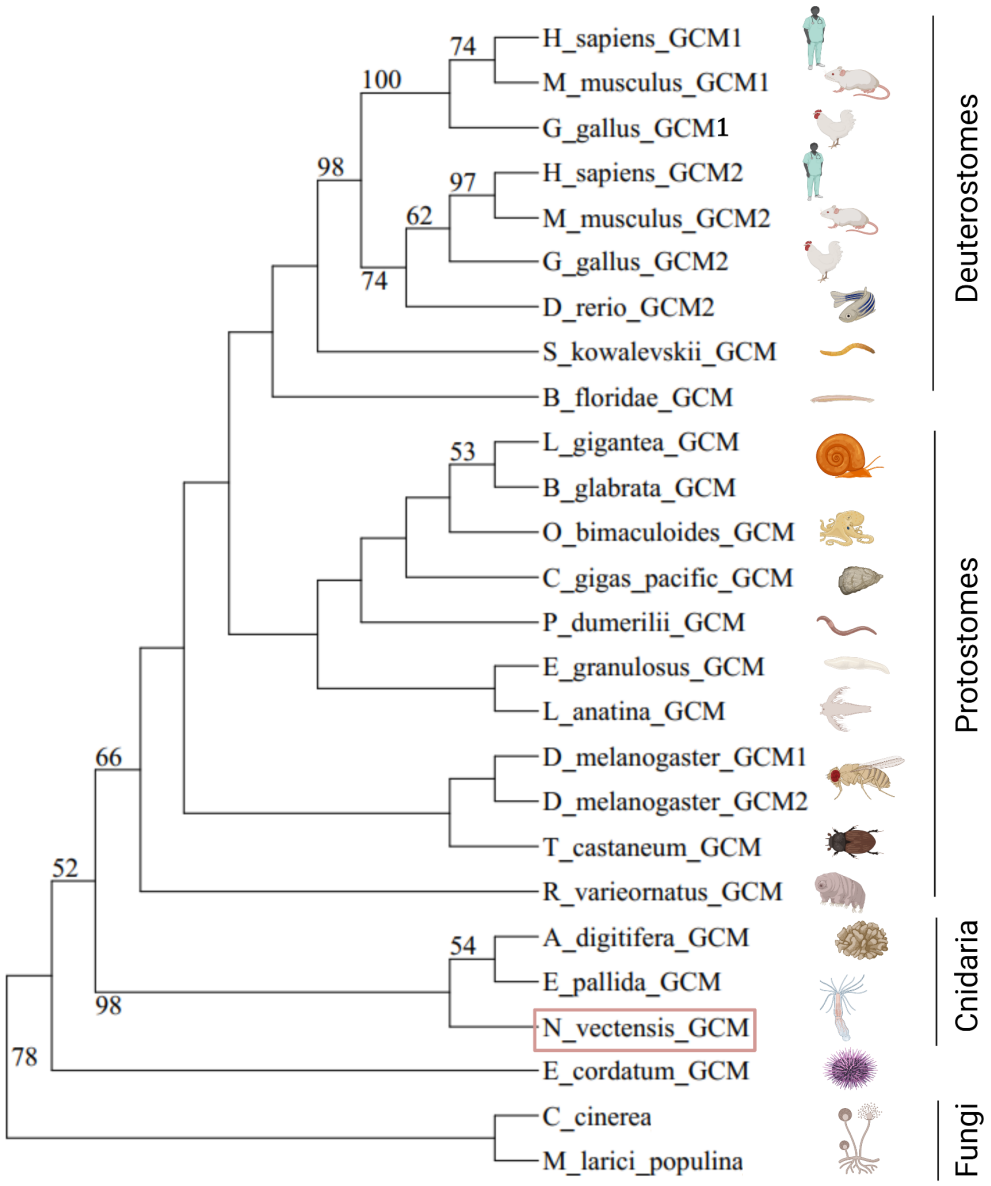


Figure 2.1. Phylogeny of *Gcm* in metazoans. Among basal metazoans *Gcm* is conserved only in *Cnidaria*. Support values >50 are shown at the basal nodes. *Nematostella Gcm* is framed in red.

Repo, another important TF driving gliogenesis under the regulation of *Gcm* in *Drosophila*, is also conserved in *Cnidaria*. In *Nematostella Gcm* is expressed in a subset of cells in the ectoderm at early gastrula stage, and then expands to both endoderm and ectoderm (Marlow et al., 2009; Watanabe et al., 2014). *Repo* is expressed in the oral region of the ectoderm and might participate in the specification of the oral nerve ring (Marlow et al., 2009; Watanabe et al., 2014). Therefore, the two TFs seem to be expressed in different cells, which is unlike *Drosophila Gcm*.

In summary, orthologs of most families containing gliogenic TFs in bilaterians are present in basal metazoans **(Table 2.2)**. However, among basal metazoans all the bilaterian gliogenic TFs are present only in *Cnidaria*. These include *Gcm* and *Repo*. The former is known to have a gliogenic role in *Drosophila* and in vertebrates. I therefore hypothesize that the emergence of these TFs might be linked to glia emergence.

2.3.2 Bilaterian glial pathway conservation in basal metazoans

Several signal transduction pathways play an important role in bilateral gliogenesis. All of them are multifunctional and drive various developmental events. Nonetheless the extent of their conservation in basal metazoans could give us a clue regarding gliogenic program evolution.

2.3.2.1 *Notch-Delta*

This pathway is composed of the *Notch* receptor and its ligands (*Delta/Jagged*). It is metazoa-specific (Gazave et al., 2009). Most *Notch-Delta* components are present in *Porifera*, *Placozoa*, and *Cnidaria* **(Table 2.2)**. The pathway seems functional in all the basal metazoans except *Ctenophora*, since the latter does not have *Delta/Jegged* in their genomes (He et al., 2021). In sponges this pathway has a proneural-like function (Richards et al., 2008). In *Nematostella Notch* signaling regulates neural progenitors and restricts neurogenesis (Richards & Rentzsch, 2015; Layden & Martindale, 2014; Marlow et al., 2012). The function of cnidarian *Notch* is thus reminiscent of bilaterian *Notch,* which acts via repressing proneural genes. At the same time in both *Nematostella* and *Hydra Notch-Delta* drives nematocyte, *Cnidaria*-specific neural cell, generation and differentiation respectively. *Notch* signaling might have different roles in different cnidarian neural cells. Regardless, an important unanswered question is whether cnidarian *Notch* and bilaterian gliogenic orthologs interact with each other.

2.3.2.2 *Hedgehog*

True *Hedgehog (hh)* genes containing both ‘hedge’ and ‘hog’ domains are absent in basal metazoans except *Cnidaria* (Ingham et al., 2011) **(Table 2.2)**. Hog-domain proteins have been identified throughout metazoa. A hedge-domain protein named *‘Hedgling’* was identified in sponges, and is present in *Cnidaria* but not *Bilateria* (Adamska et al., 2007). In *Nematostella Hedgehog* gene expression analysis showed that true *hh* genes play a role in gut formation, and *hh*-related genes are involved in neuronal development (Matus et al., 2008). A more recent study demonstrated that germ cell development is dependent on *Hedgehog* in *Nematostella* (Chen et al., 2020).

2.3.2.3 *Jak-Stat*

Jak-Stat pathway is composed of several proteins made up of specific domains which together assemble into a functional system driving transcriptional response to specific extracellular signals (Liongue & Ward, 2013). Basal metazoans possess most of the *Jak-Stat* components **(Table 2.2)**. *Ctenophora* has the least number of conserved proteins, while *Cnidaria* and *Placozoa* only lack one functional unit exclusively present in bilaterians (Liongue et al., 2016). There are no functional studies of *Jak-Stat* pathway in basal metazoans.

2.3.2.4 Bone morphogenetic proteins

Bone morphogenetic proteins *(BMPs)* belong to the *TGF-b* superfamily and are involved in several aspects of the NS development including glial cell differentiation (Li & Grumet, 2007; See et al., 2007). All basal metazoans except *Porifera* have *BMP-like* genes (Pang et al., 2011) **(Table 2.2)**. In Cnidaria *BMP* signaling is involved in oral NS formation

(Watanabe et al., 2014). Given that *BMP* genes are expressed in a neuron-rich aboral region of ctenophores, they might play a role in the NS development in these animals as well (Pang et al., 2011).

In summary, in addition to gliogenic TF conservation, *Cnidaria* is the only basal metazoa phylum characterized by the conservation of all the signaling pathways required for glial development in bilaterians **(Table 2.2)**. Moreover, many of these show a conserved function of driving the development of the nervous system. Most components of these pathways are present in the other three basal metazoan phyla, but their functional description is limited. *Cnidaria* also possess all the functional glial genes including GABA and glutamate transporters and enzymes required for their synthesis, glucose transporters, *TRPM* ion channels, aquaporins, etc. (Putnam et al., 2007; Sebé-Pedrós et al., 2018). Thus, the existence of a complete set of neuronal and glial genes in *Cnidaria* substantiates the possibility of simultaneous evolution of both cell types as argued by (Rey et al., 2020).

2.3.3 *Nematostella vectensis* as a model organism to study the evolution of glia

2.3.3.1 *Gcm* is conserved in *Anthozoa*

Having carried out the phylum-wide genome-wide analysis of bilaterian glial TF conservation in basal metazoans, I focused on *Gcm* as the "core" bilaterian glial TF. To get further insight I performed *Cnidaria*-specific phylogenetic analysis of *Gcm*, which revealed that it is highly conserved only in anthozoans **(Figure 2.2)**.

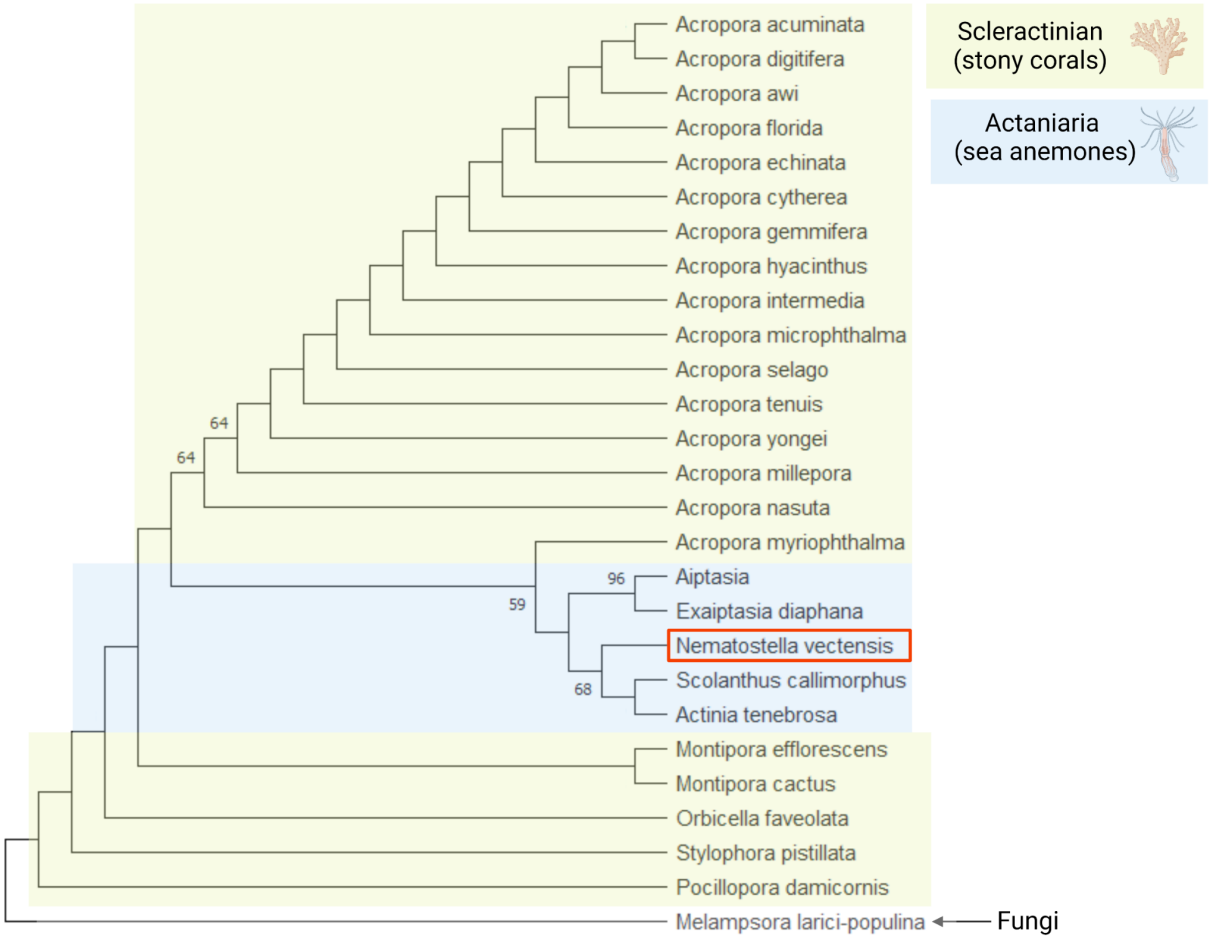


Figure 2.2. Phylogeny of *Gcm* in *Anthozoa* (*Cnidaria*). The domain is highly conserved in stony corals and sea anemones. Support values >50 are shown at the basal nodes. *Nematostella Gcm* is framed in red.

I could not identify *Gcm* orthologs in *Aurelia aurita* (scyphozoan), *Hydra viridissima* (hydrozoan), or *Morbakka virulenta* (cubozoan). Apparently, *Gcm* was lost in *Medusozoa*, which is consistent with the data showing that *Hydra* has lost more transcription factor families than *Nematostella* (reviewed in Steele et al., 2010).

2.3.3.2 Core bilaterian glial TFs in *Nematostella*

The starlet sea anemone *Nematostella vectensis* is an anthozoan that has become an intensely used cnidarian model organism. This animal is easy to culture in the lab; its genome, transcriptome, and single-cell transcriptome are available; numerous gene function manipulation techniques have been developed and transgenic lines of *Nematostella* have been established (Karabulut et al., 2019; Layden et al., 2016; Putnam et al., 2007; Renfer et al., 2010; Sebé-Pedrós et al., 2018). Mounting data regarding *Nematostella* neurogenesis, nervous system development and functioning, as well as neuron type diversity made it possible to look for the conservation of gliogenic program conservation in this animal. Despite the absence of clear glial cell clusters, i.e. glial transcriptome signature in *Nematostella* (Sebé-Pedrós et al., 2018), the extensive literature search and phylogenetic analysis of glial genes I performed confirmed glial program conservation in the sea anemone. In addition, *Gcm* ortholog is not

only present in *Nematostella*, but is expressed in a subset of cells during its development (Marlow et al., 2009). Therefore, *Nematostella* is a perfect cnidarian organism to explore the functions of this bilaterian glial TF.

First, I confirmed the expression pattern of *Nematostella Gcm* (Marlow et al., 2009). Importantly, *Gcm* expression coincides with the formation of the nervous system in *Nematostella* (Rentzsch et al., 2017). Using single-cell transcriptome data, I confirmed the prevalent expression of *Gcm* in neuronal cells in both larval and adult *Nematostella* (**Figure 2.3A,C**).

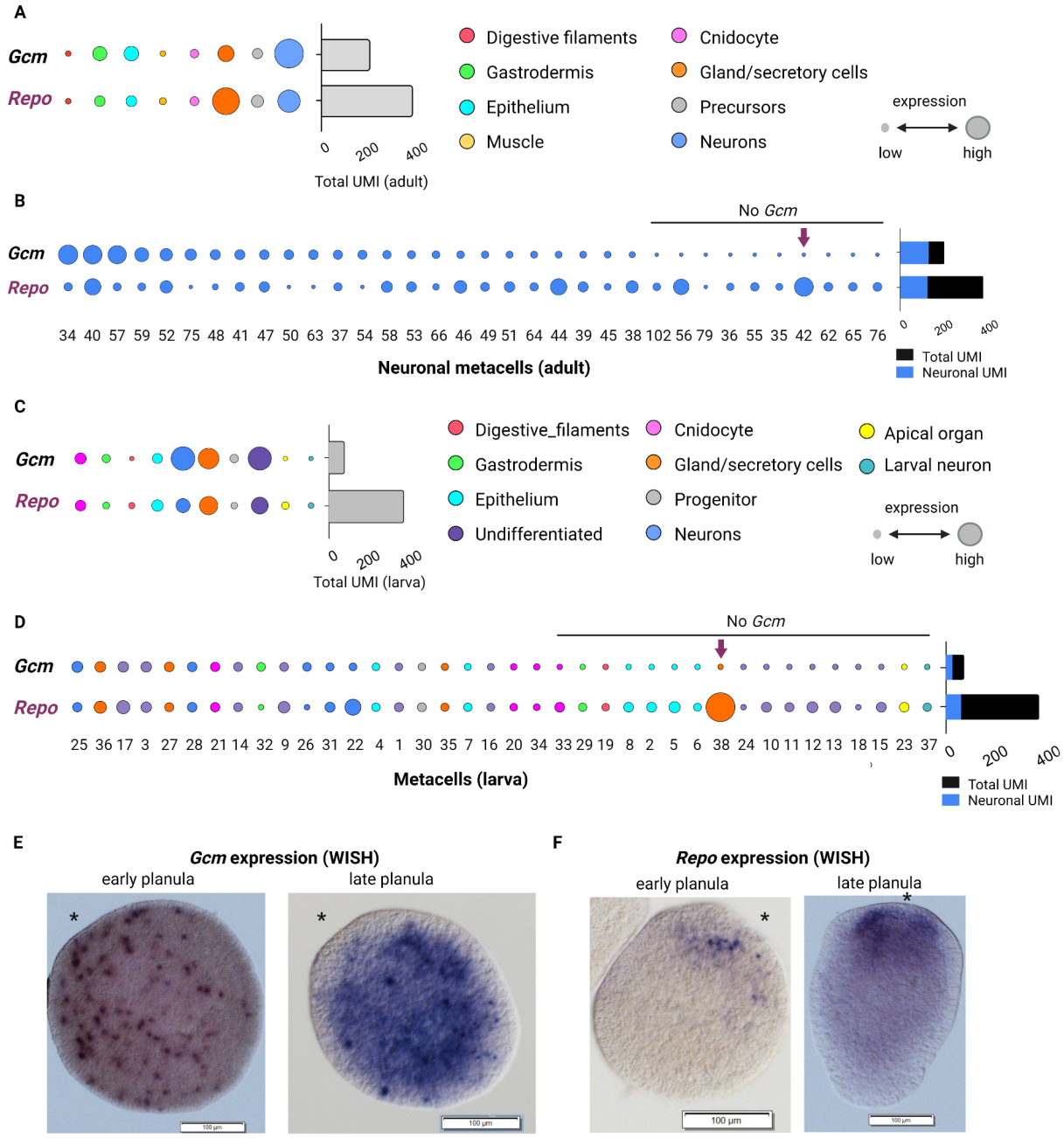


Figure 2.3. Expression of *Nematostella Repo* and *Gcm* does not coincide. Expression of both genes across different cell types in adult animals (A) and at planula stage (C). Expression (molecules/1,000 UMIs) of both genes across neuronal cell clusters (metacells) at adult stage (B) and across all cell clusters at larval stage (D). Neuronal cell clusters are organized left to right in the descending order of *Gcm* expression. Purple arrow indicates the metacell with the highest expression of *Repo*. (E) *Gcm* is expressed in a salt and pepper pattern throughout endoderm and ectoderm. (F) *Repo* expression is restricted to the oral region. WISH technique. Asterisk denotes the oral side of the larvae. Scale bars: 100 µm. The highest expression is defined as the most UMIs/cell type for a given gene relative to its overall expression. Assembled with BioRender.com.

In *Drosophila, repo* is an important downstream target of *Gcm* driving gliogenesis (see 1.2.4.1). Therefore next I confirmed the expression of *Repo* in *Nematostella* (Marlow et al., 2009; Watanabe et al., 2014) **(Figure 2.3F**). The two TFs are not co-expressed, which is also evidenced by the single cell data analysis I performed **(Figure 2.3)**. In adult *Nematostella* the neuronal cell with the highest expression of *Repo* is devoid of *Gcm* **(Figure 2.3B)**. Similarly, at the larval stage *Repo* is most abundant in *Gcm*-negative cell clusters **(Figure 2.3D)**. Moreover, at both adult and larval stages gland/secretory cells demonstrate the highest expression of *Repo* unlike *Gcm*, indicating a non-neuronal function of *Repo* **(Figure 2.3A,C,D)**. Hence, unlike *Drosophila, Nematostella Gcm* does not seem to control *Repo.*

2.4 Discussion

Up until now no comprehensive analysis of bilaterian gliogenic program conservation in basal metazoans has been done. My analysis showed that cnidarians are equipped with all the glial TFs and pathway components required for gliogenesis. Moreover, the core bilaterian glial TF, *Gcm*, is conserved in anthozoans. In *Nematostella Gcm* belongs to a neurogenic gene regulatory network (apical domain) together with *NvashA*, *Hes*, and *Sox* genes (Layden et al., 2016; Amiel et al., 2017; Layden et al., 2016). This, combined with its expression in a subset of neuronal cells, indicates *Gcm* involvement in neural development and its interaction with TFs known to play a gliogenic role in bilaterians. Therefore, I decided to unravel the *Gcm*-controlled program in *Nematostella* in order to provide insights into its function as a conserved neuronal TF and shed light on glial evolution. I knocked down *Nematostella Gcm* and explored its targets as described in the next chapter. To my knowledge this is the first attempt to perform a functional analysis of a conserved glial TF in basal metazoans.

Chapter III. *Nematostella Gcm*-controlled program is glio-neuronal in nature

3.1 Background

Gcm is one of the major gliogenic TFs in bilaterians: it demonstrates gliogenic potential not only in *Drosophila*, but also vertebrates. At the same time, driving gliogenesis is not the primary function of *Gcm* in mammals (see Chapter I). Its non-neuronal expression in vertebrates casts doubt regarding the conserved function of *Gcm* as a neuronal TF. Importantly, *Cnidaria,* the sister group of *Bilateria*, possesses a *Gcm* homolog (see Chapter II). Therefore, the analysis of the cnidarian *Gcm* should unravel if its primary function was indeed neuronal. Gene knockdown (KD) is a widely-used method of exploring the gene regulatory network of a transcription factor (TF) and drawing conclusions regarding its function. In this study, I performed a knockdown of *Gcm* in *Nematostella.* Next I analyzed the genes affected by *Gcm* KD. This approach allowed to clarify the conserved function of *Gcm* as a neuronal TF.

I demonstrate that the *Gcm*-controlled program is enriched with cell adhesion proteins, neurotransmitter transporters, ion channels, metabolic and protein modifying enzymes. Therefore, *Nematostella Gcm* targets are genes involved in fulfilling classical glial functions such as regulating ionic and neurotransmission composition of the neuronal environment, glucose and lipid metabolism. Moreover, similarly to *Drosophila*, *Nematostella Gcm* regulates the expression of a glutamate transporter homolog, *Eaat1*. Immunostaining for *Eaat1* revealed a new cell type in *Nematostella*, which differs from classical neurons but nevertheless does not conform to bilaterian glial morphology. The results of this study indicate that *Gcm* gliogenic potential is conserved in *Nematostella* in the absence of glia per se. It is thus possible in *Nematostella* to observe a subset of neuronal cells already equipped with a glial program. This suggests that glial and neuronal programs co-evolved, and separated later on in the evolution to give rise to cells with either neuronal or glial functions.

3.2 Materials and methods

3.2.1 *Nematostella vectensis* culture

Nematostella vectensis were cultured as described previously (section 2.2.3).

3.2.2 RNA interference

For *Gcm* RNA interference, by transfection, *Nematostella vectensis* embryos were electroporated as described in (Karabulut et al., 2019) with the following modifications: no Ficoll was added to the brackish water medium, since it decreased embryo survival rate; siRNA was used instead of shRNA because of the simplicity of siRNA manufacturing; egg masses were collected, degelled and fertilized prior to transfection **(Figure 3.1)**. Electroporation was carried out using Gene Pulser Xcell electroporation system (BIO-RAD).

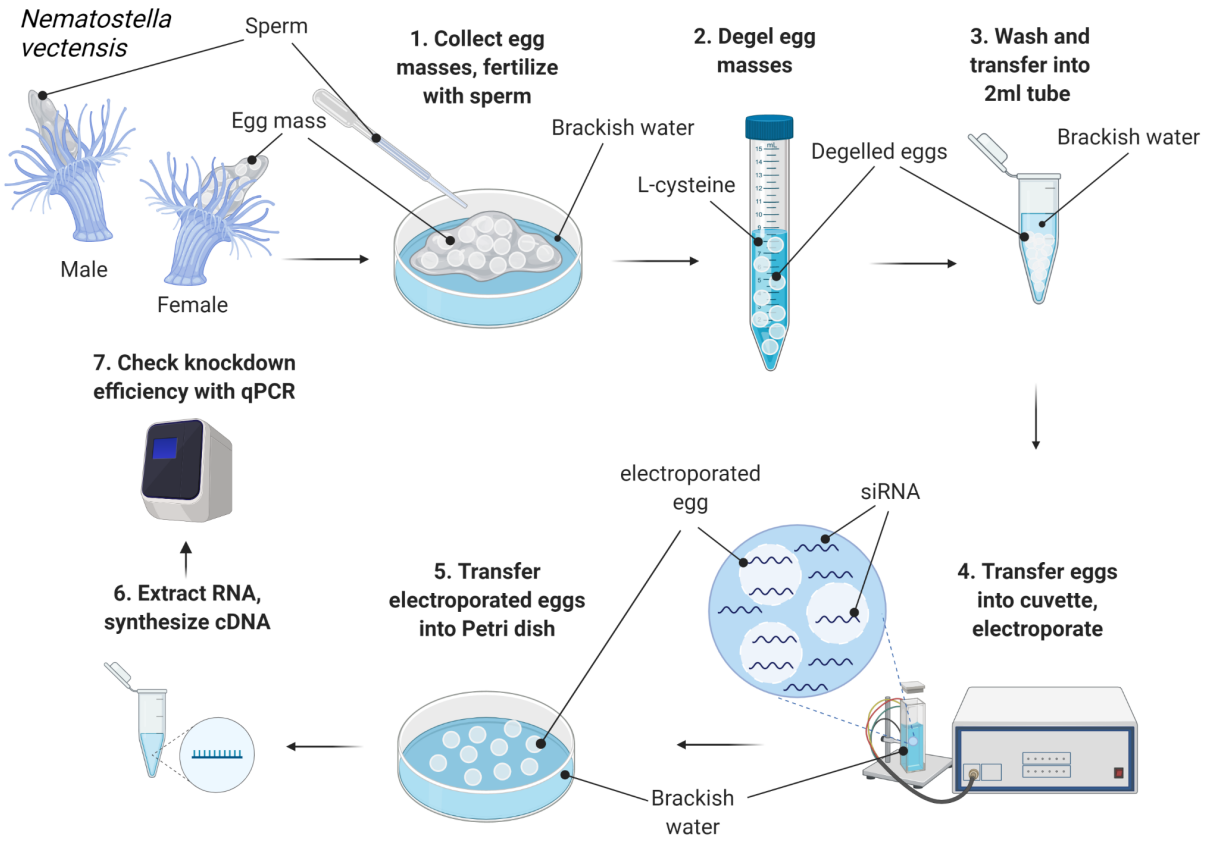


Figure 3.1. Schematic of the experimental procedure for *Nematostella Gcm* knockdown. Created with BioRender.com.

Three pairs of *Gcm*-specific siRNAs sequences were designed to target *Gcm* mRNA: Invivogen sirnawizard was used to design siRNA1 (TTACCTGTGCATTTCCTGTdT); siRNA2 (CUACCCAAUGGACUCAAGAdTdT), siRNA3 (CGUCCGACCCGCGAUAUCAdTdT), and negative control siRNA (GCAACACGCAGAGTCGTAAdT) were designed by Sigma-Aldrich Japan. siRNA concentration of 500 ng/uL was used. 4dpf *Nematostella vectensis* planulae were collected for RNA extraction and knock-down efficiency assessment. Although siRNA KD efficiency varied among experimental batches (Appendix 4A), siRNA1 and siRNA2 were shown to be more efficient than siRNA3, and were thus chosen for the experiments (Appendix 4B).

3.2.3 Sample preparation for RNA sequencing

Total RNA was extracted from 4dpf transfected and control planulae *Nematostella vectensis* using RNeasy Mini Kit (per manufacturer's guidelines; QIAGEN). NanoDropTM 1000 Spectrophotometer (Thermo Fisher Scientific) was used for the RNA concentration and purity check. Samples with absorbance ratios OD260 / OD280 and OD260 / OD230 higher than 1.6 each were used for sequencing. Total RNA was extracted from five biological replicates (15 samples overall - 5 experiments of 3 groups: siRNA1, siRNA2, and control siRNA), dissolved in 30 µL of nuclease-free water and kept at −80 °C until total RNA from all the samples was collected. Once all the fifteen samples were acquired, they were placed on dry

ice, and sent to a third-party company for sequencing. Quality checks were performed by Agilent TapeStation (Takara bio facility). RIN (RNA integrity number) value of all the samples was confirmed to be 10, i.e. sufficient for high-quality sequencing. Library preparation (TruSeq RNA Library Prep Kit) and sequencing (NovaSeq6000) were also done by Takara bio (paired-end: 150bp x 2; depth: 50-80ml). *Gcm* KD efficiency in RNA samples used for sequencing was assessed with RT-qPCR analysis. RNA from different knock-down experiments was used to confirm down/upregulation of genes identified by RNA-seq data analysis.

3.2.4 RNA-seq data quality control and read mapping

Read quality was assessed using FastQC (Andrews, 2010) and visualized with the MultiQC python package (Ewels et al., 2016). The mean sequencing depth of the libraries was 61.76 million reads (Appendix 5). The data were of good quality, and no trimming was required. Paired reads were aligned to the most updated at the time *Nematostella vectensis* gene models available on Figshare using Bowtie 2 aligner (Langmead & Salzberg, 2012) with prior filtering using CD-HIT-EST (similarity threshold set to 94%) to remove isoforms (Li et al., 2012). Overall alignment rate was 76% on average (Appendix 5). SAMtools was used to convert the files from SAM to BAM format. Salmon (Patro et al., 2017) was used to quantify transcript abundances. Salmon-generated pseudocounts represented as normalized TPM (transcripts per million) were converted into non-normalized count estimates for differential gene expression analysis as per the workshop by the teaching team at the Harvard Chan Bioinformatics Core (HBC) - https://github.com/hbctraining/DGE_workshop_salmon. The RNA-seq profiles of 23,913 transcripts were obtained.

3.2.5 Differential gene expression analysis

To identify genes differentially expressed (DE) in *Gcm* KD and control groups DESeq2 was used, which utilizes the median of ratios method to normalize for sequencing depth and RNA composition of the samples (Love et al., 2014). Sample-level quality control (QC) was performed using Principal Component Analysis (PCA) and hierarchical clustering methods. Two out of five experimental batches were selected to ensure the experimental condition represented the major source of variation. *Gcm* KD for both siRNA1 and siRNA2 was most efficient in these batches (Appendix 4C). Experiment dates represented another source of variation, which was controlled for during DEseq analysis. *Gcm* KD siRNA1 and siRNA2 groups were compared to the control group separately. P-adj (an adjusted p value cutoff, i.e. the FDR which is the expected proportion of false positives) was set to 0.05, 01, and finally 0.2 to extract a longer list of DE genes. This list included *Eaat1* that we previously confirmed to be downregulated in *Gcm* morphants using RT-qPCR. Since siRNA2 KD was shown to be consistently more efficient and yielded more DE genes, the results of only siRNA2 vs control are presented. Three common DE genes between siRNA1 vs. control and siRNA2 vs. control groups were identified: *PKM2* (NVE38) - downregulated in KD; *Ets*-related (NVE9883) and uncharacterized (NVE23912) - upregulated in KD (Appendix 8). *PKM2* and *Ets*-related were confirmed to be affected by *Gcm* KD using RT-qPCR (see section 3.3).

3.2.6 Functional analysis of DE genes

Available databases used for gene ontology (GO)/pathway analysis such as KEGG, PANTHER classification system, DAVID have incomplete *Nematostella vectensis* annotations. Therefore, in order to categorize DE genes into functional groups and explore their molecular pathways, each protein function was inferred from its similarity to the bilaterian proteins using CBLAST (protein data bank)/BLAST(Altschul et al., 1990; States & Gish, 1994). The genes were then assigned to the pathway groups as defined in KEGG PATHWAY Database (Kanehisa & Goto, 2000). The workflow is summarized in **figure 3.2**.

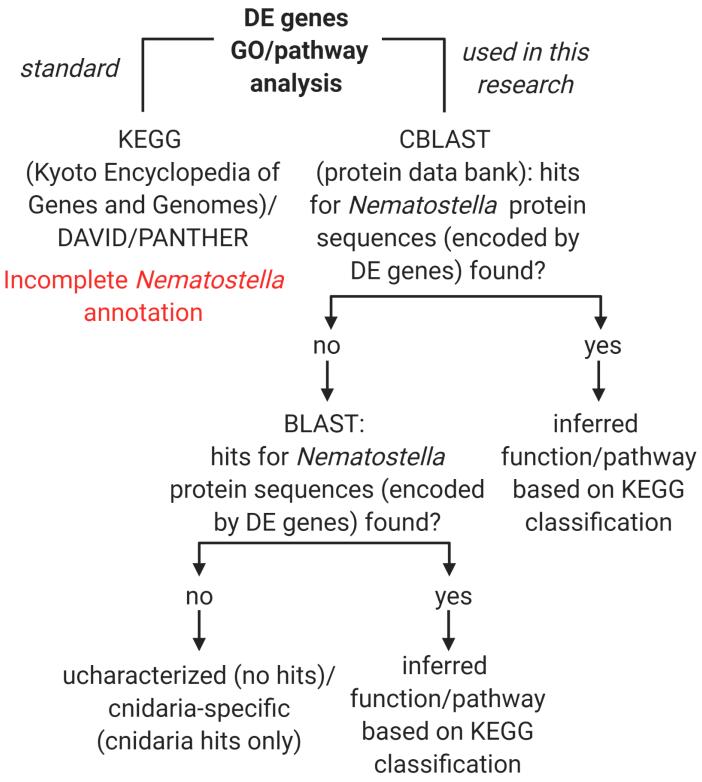


Figure 3.2. DE gene functional analysis approaches: standard vs used in this study.

3.2.7 DE gene verification with RT-qPCR

3.2.7.1 RNA extraction and cDNA synthesis

Total RNA was extracted from 4dpf control and electroporated *Nematostella* using RNeasy Mini Kit (per manufacturer's guidelines; QIAGEN). Total RNA was dissolved in 30 µL RNase-free water and stored at -80°C freezer prior to cDNA synthesis. NanoDropTM 1000 Spectrophotometer (Thermo Fisher Scientific) for the RNA concentration and purity check. Agarose gel was also run to assess RNA integrity. Single-stranded cDNA was synthesized from 300-500 ng of total RNA in a final volume of 20 µL by using oligo(dT)20 and SuperScript IV Reverse Transcriptase (Invitrogen) according to manufacturers' instructions. cDNA was stored at −20 °C for future use. Agarose gel of PCR products obtained using primers for a gene not affected by *Gcm* KD (*Nfix-like,* Gene ID: 5517439) as confirmed by RT-qPCR was run to check for genomic DNA contamination of cDNA. The primers were

designed so as to yield PCR products of different length for mRNA (850 bp) and gDNA (1463 bp). Forward: GACCCTCAGTCACCATCCATC; Reverse:ACTGAGCTACCCTTGACCCA.

3.2.7.2 RT-qPCR settings and primer design

RT-qPCR was performed using StepOne Plus™ Real-Time PCR System (Thermo Fisher Scientific). The amplification program was set as follows (fast mode ~40 min): holding stage at 95 °C for 20 sec followed by 40 cycles of 3 sec at 95 °C and 30 sec at 60 °C. After amplification, a denaturing cycle (15 sec at 95° followed by 1 min at 60 °C and 15 sec at 95°) was performed to obtain the melting curves and verify amplification specificity. Samples for RT-qPCR were prepared using PowerUp SYBR® Green Master Mix as per manufacturer's instructions. Primers were designed using Geneious 10.2.4 (https://www.geneious.com) and Primer-BLAST (Ye et al., 2012). Primer sequences and product sizes are listed in Appendix 6. Primer efficiencies were assessed using cDNA 5-fold serial dilutions through the generation of a standard curve. Primers with 80-110% efficiency were chosen for qPCR analysis. Four samples were run for each gene; triplicates were used for analysis.

3.2.7.3 Candidate reference genes

Three candidate housekeeping genes, *Gapdh*, *Ef1a*, and *18S,* were considered as normalization (reference) genes. Norm-Finder (Andersen et al., 2004) was used to assess the variability of the genes and how much they are affected by the experimental conditions, which are important factors to consider when choosing normalization genes (Huggett et al., 2005). *Ef1a* was consistently shown to be the best candidate among the three with the lowest intra- and intergroup variation scores (3 experiments, 4-8 samples per group, 3 groups: siRNA1, siRNA2, control siRNA). As a result, the expression of the genes of interest was normalized to *Ef1a.* Nevertheless, reproducible results were obtained regardless of using one (*Ef1a*) or several (*Ef1a+Gapdh)* genes for normalization.The comparative Ct (△△Ct) value method was used for relative quantification (Livak & Schmittgen, 2001). Both siRNA1 vs control siRNA and siRNA2 vs control siRNA comparisons were performed and yielded similar results. Since siRNA2 was shown to be more efficient in some batches, the results of only siRNA2 vs control are presented.

3.2.7.4 Statistical analysis

RT-qPCR data are presented as expression fold change (±SD) normalized to one endogenous control (*Ef1a*) and relative to the control siRNA. Means of different groups were compared (siRNA2 and siRNA control) and analyzed using Student's unpaired t-test. Results from three independent experiments were used for the analysis unless stated otherwise. Differences were considered statistically significant when p-value was < 0.05.

3.2.8 Immunostaining techniques

3.2.8.1 Antibodies

The following primary antibodies were used: to detect NvEAAT1 anti-Eaat1 (rabbit, German Research Products, 1:400 for IF (immunofluorescence) and 1:1000 for WB (western blotting)), anti-Human Eaat1/Glast-1 (sheep, Bio-Techne, 10 µg for IF and 5 µg for WB); to

detect NvADGRV1 anti-ADGRV1 (rabbit, Atlas Antibodies, 1:200 for IF and 1:1000 for WB); to detect Nva-tubulin (positive control) anti-a-tubulin (mouse, Sigma, 1:300 for IF and 1:1000 for WB); to detect PRGamide an antibody against the amidated Cys-PRGamide was generated (Scrum, 1:300 for IF). Anti-EAAT1 and anti-ADGRV1 antibodies were selected based on the sequence similarity of the *Nematostella* protein sequences and the epitopes declared to be recognized by the antibodies. See appendix 9A,B for sequence alignments.

The following secondary antibodies were used: Alexa Fluor 488-conjugated anti-rabbit, anti-sheep, and anti-mouse antibodies (goat, Jackson ImmunoResearch, 1:500) for IF; Alexa Fluor 647-conjugated anti-rabbit antibodies (goat, Jackson ImmunoResearch, 1:500) for IF; anti-rabbit, anti-sheep, and anti-mouse peroxidase-conjugated antibodies (goat, Jackson ImmunoResearch, 1:10000) for WB.

3.2.8.2 Western blotting

Approximately 1000 (25 µl) of 4dpf and 5dpf *Nematostella vectensis* planulae were collected for total protein extraction in a 1.5 mL eppendorf tube. 5 µl of Laemmli sample buffer (6x, Bio-Rad) was added, and samples were boiled at 80°C for 5 mins. Protein concentration was measured using EZQ protein quantification kit (Thermo Scientific). 5 µg of total protein were loaded per lane. SDS-PAGE was carried out using 10% polyacrylamide gels (Bio-Rad). The gel was initially run at 10mA/gel (stacking phase), then 20 mA/gel (resolving phase) until the protein ladder was ~1 cm from the base. The gel was then transferred onto a methanol activated polyvinylidene difluoride (PVDF) membrane sandwiched between two filter papers soaked in the 1x transfer buffer. Trans-Blot Turbo System (Bio-Rad) was used to transfer the proteins at 25V for 7 minutes. The PVDF membranes were then blocked with 5% skimmed milk in 0.1% Tween 20 in PBS (PBST) for 1h at RT on a shaker. Subsequently, the membranes were incubated with primary antibodies diluted in 5% skimmed milk in PBST at 4°C o/n. The membranes were then washed 3 times in PBST for 5 mins each. This was followed by an incubation with secondary antibodies diluted in 5% skimmed milk in PBST for 1h at RT. The membranes were then washed 3 times in PBST 10 mins each. Chemiluminescence was developed by treating the membranes with ImmunoStar Zeta (Fujifilm Wako Pure Chemical Industries).

Anti-EAAT1 antibody (German Research Products) specificity was confirmed by knocking down *Nematostella Eaat1* using gene-specific siRNAs and performing western blotting (WB) with a loading control. Three pairs of *NvEaat1*-specific siRNAs were designed by Sigma-Aldrich Japan: siRNA1 (GUCAGUUACGUUAACAGAAdTdT), siRNA2 (CCUAUCAAACGGCACACAUdTdT), siRNA3 (GUAUUACUUAGCUACGACUdTdT). Data for siRNA2 and siRNA3 are presented. The same negative control siRNA was used as described in section 3.2.2. siRNA concentration of 500 ng/uL was used. SiRNA KD efficiency was assessed using RT-qPCR as described in section 3.2. WB with a-tubulin as a loading control was performed to compare NvEAAT1 protein expression in KD and control samples at 4dpf. The optimal total protein load of 6 µg was determined beforehand by creating a standard curve using serial dilutions of a control sample for both NvEAAT1 and Nva-tubulin. Following the gel run of KD and control samples, the membrane was cut in half at ~55 kDa and incubated in anti-EAAT1 and anti-a-tubulin antibodies. WB band intensities were then determined using Image Lab 6.0.1 (Bio-Rad) software, target protein expression was normalized using loading control, and fold difference was calculated for KD samples.

3.2.8.3 Fluorescent immunohistochemistry

Bud/polyp stage *Nematostella* were anesthetized using 2.43% MgCl2 and incubated for 10 mins. Planulae and anesthetized bud/polyps were fixed in ice-cold 4% PFA/PBS+0.1% Triton X-100 (PBSTx) for 1h at RT on a rotator. Samples were washed for 3 times (10 mins each) with PBS+0.1%Tween (PBT). Samples were then incubated in blocking solution (PBT/1%BSA/5% Normal goat serum/0.01%NaN3) for 1h at RT, followed by o/n incubation with primary antibodies at 4°C on the rotator. Samples were washed for 2 times (10 mins each) with PBSTx and incubated in the blocking solution for 1h at RT on the rotator. Samples were incubated in secondary antibodies for 1h at RT on the rotator. Samples were then washed for 3 times with PBSTx (15 mins each) at RT. Samples were then incubated for 30 mins in DAPI (1:1000) and mounted with SlowFade Gold Antifade Mountant (Invitrogen).

3.2.9 RNAscope® whole-mount fluorescent *in situ* hybridization (FISH)

RNAscope® Multiplex Fluorescent Reagent Kit v2 Assay was used to detect *Nematostella Gcm* mRNA in whole-mount specimens. A custom target probe for *Gcm* was designed by Cosmobio, Japan. RNAscope® Negative Control Probe- DapB was used as a negative control. Conventional one-color *in situ* hybridization with *Gcm*-specific probe was used as a positive control (see section 2.2.4) for RNAscope FISH.

RNAscope® FISH fine-tuned protocol developed for zebrafish embryos (Gross-Thebing et al., 2014) was optimized for *Nematostella* as follows. Bud/polyp stage *Nematostella* were anesthetized using 2.43% MgCl2 and incubated for 10 mins. Planulae and anesthetized bud/polyps were fixed in ice-cold 4% PFA/PBS for 1h at RT on a rotator. For all the procedures ~20 embryos were processed in one 2 ml Eppendorf tube coated with 1% BSA/brackish water. Embryos were washed with PBT at RT for 10 mins 3 times. A series of increasing MeOH concentrations (25%, 50%, 75%, 2× 100%) in PBT was used to dehydrate the embryos stepwise in 10-min washes. Embryos were then stored in MeOH at -20°C for at least one night. They were then incubated in fresh MeOH at RT for 30 mins. A series of decreasing MeOH concentrations (100%, 75%, 50%, 25%,) in PBT was used to rehydrate the embryos stepwise in 10-min washes. Embryos were then incubated in PBT+1%BSA at RT for 10 mins. Protease Plus digestion of embryos was performed for 60 mins at 40ºC. Target probe hybridization was performed at 40°C O/N. 50µl of probe mix was used per tube. The embryos were then washed three times for 10 min in PBT at RT. Postfixation step
was performed using 4% PFA in PBS for 10 min at RT. The embryos were then washed three times for 10 min in the wash buffer at RT. Incubation in amplifier solutions was done at 40°C for 30 mins followed by 3 washes in the wash buffer 10 mins each at RT. Label probe hybridization was performed as stated in the original protocol. Opal 520 (1:1000 dilution) and Opal 570 (1:1500 dilution) were used to detect the probe. Embryos were incubated at 4°C in ready-to-use DAPI solution O/N. Embryos were then rinsed in PBS and mounted with SlowFade Gold Antifade Mountant (Invitrogen). Slides were stored in the dark at 4°C before imaging.

3.2.10 Microscopy

Images were captured either on a Zeiss LSM780 confocal microscope system using x40 and x63 objectives or Nikon epifluorescence Eclipse Ni-U microscope using x40 objective. Image manipulation was performed with ZEISS ZEN microscope software.

3.2.11 *Nematostella vectensis* single cell data analysis

Single cell data analysis was carried out as described previously (section 2.2.2) except for figure 3.19, where the expression is shown as a fraction of total expression following molecules/1,000 UMIs transformation.

3.3. Results

3.3.1 siRNA-based knockdown of *Nematostella Gcm*: efficiency and RNA-seq

To determine *Gcm* function in *Nematostella* I knocked it down using siRNA specific for *Gcm*. I did not observe any phenotypic changes in the knockdown (KD) animals. KD efficiency of *Gcm*-targeted siRNA was ~75% for the batches selected for RNA-seq analysis **(Figure 3.3, A)**.

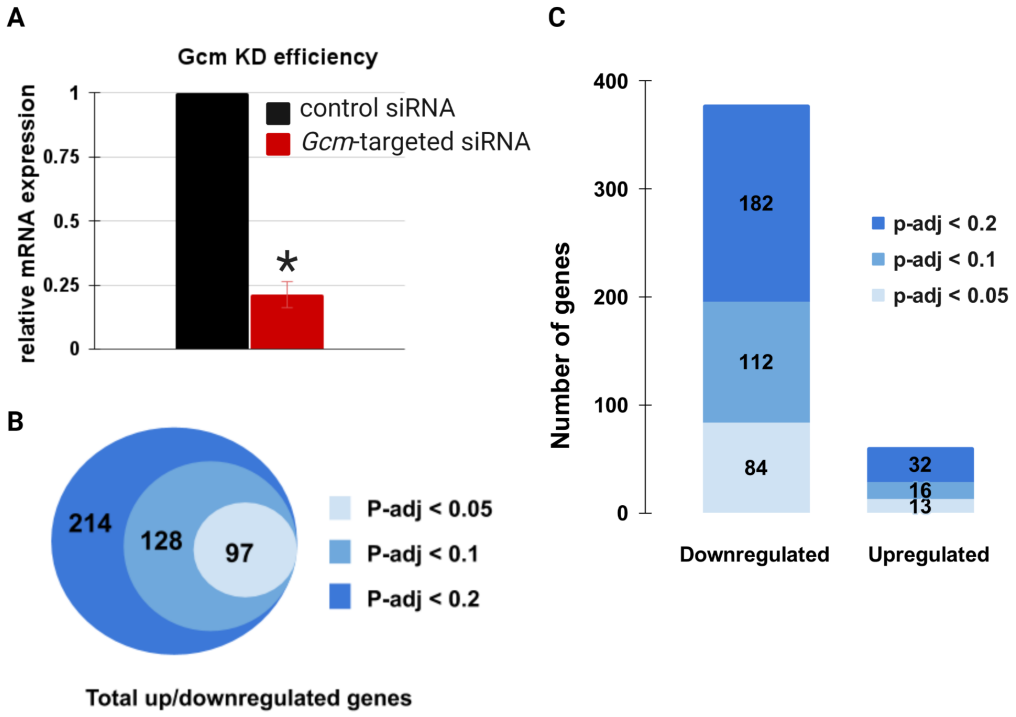


Figure 3.3. *Gcm* KD efficiency and RNA-seq DE gene distribution. (A) *Gcm* KD efficiency of siRNA for the experimental batches selected for RNA-seq analysis. Data are given mean ±SD, N=2, *p < 0.05, unpaired t-test. (B) Total number of affected genes due to *Gcm* KD depending on p-adj. (C) Total number of genes up and downregulated depending on p-adj. Assembled with BioRender.com.

To identify genes affected by *Gcm* KD, I performed differential gene expression (DE) analysis. Depending on the set p-adj, the number of significantly DE genes varied **(Figure 3.3, B)**. More genes were downregulated than upregulated in the *Gcm* morphants **(Figure 3.3, C)**. All the DE genes and their functional classification are listed in Appendix 7.

According to GO (gene ontology)/pathway analysis of the affected genes, uncharacterized/cnidaria-specific genes with unknown functions constitute the majority of both up- and downregulated genes (~25%) **(Figure 3.4)**.

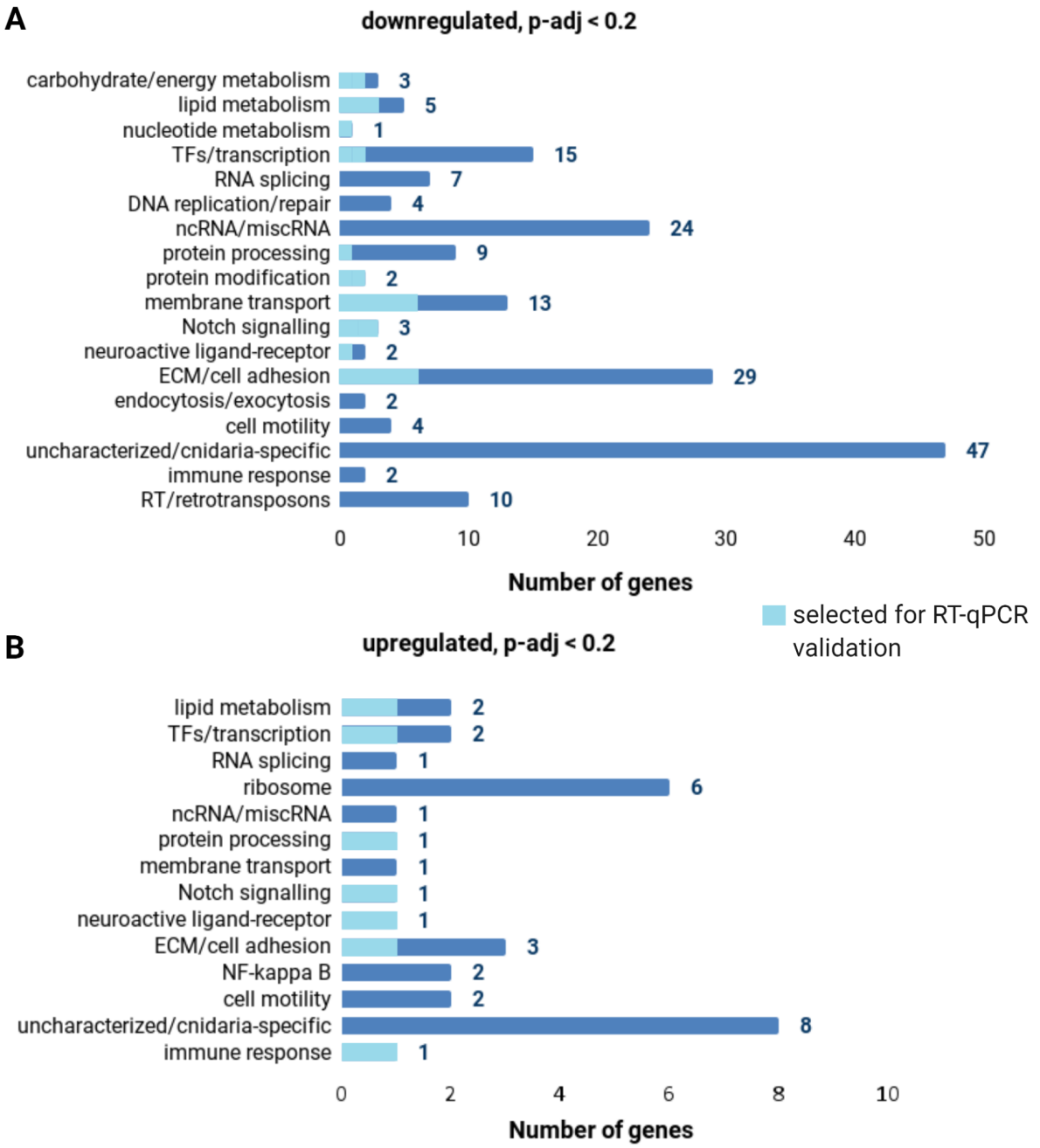


Figure 3.4. GO terms for DE genes in *Gcm* morphants. (A) Total number of upregulated genes according to their functional groups. (B) Total number of downregulated genes according to their functional groups.

Among upregulated genes, ribosomal and ECM (extracellular matrix)/cell adhesion constituents are second (~19%) and third (~9%) largest groups respectively **(Figure 3.4A)**. Interestingly, among the downregulated genes, ECM/cell adhesion is the second largest group (~16%) followed by ncRNA/miscRNA (~13%) **(Figure 3.4B)**. Therefore, ECM/cell adhesion genes are highly abundant in both up- and downregulated groups. Among other major GO groups are TF/transcription (~8%) and membrane transport (~7%) in downregulated

genes. Genes belonging to *Notch* signaling pathway, neuroactive ligand-receptor, and lipid metabolism groups are also present among both up- and downregulated genes, suggesting a neural nature of the *Gcm*-controlled program.

3.3.2 RT-qPCR validated *Gcm* targets

To validate *Gcm* targets I selected several genes from each of the major GO groups for RT-qPCR analysis **(Figure 3.5)**. The exception was those functional groups that would not contribute to the understanding of the *Gcm*-controlled program (ncRNA/miscRNA, uncharacterized/cnidaria-specific, RNA splicing, ribosome, RT/retrotransposons, etc.). I selected 33 genes for RT-qPCR validation in total, and confirmed 24 of them to be up/downregulated in the *Gcm* morphants **(Figure 3.5A)**. *Gcm* KD efficiency was also assessed with RT-qPCR for each experimental batch **(Figure 3.5B)**.

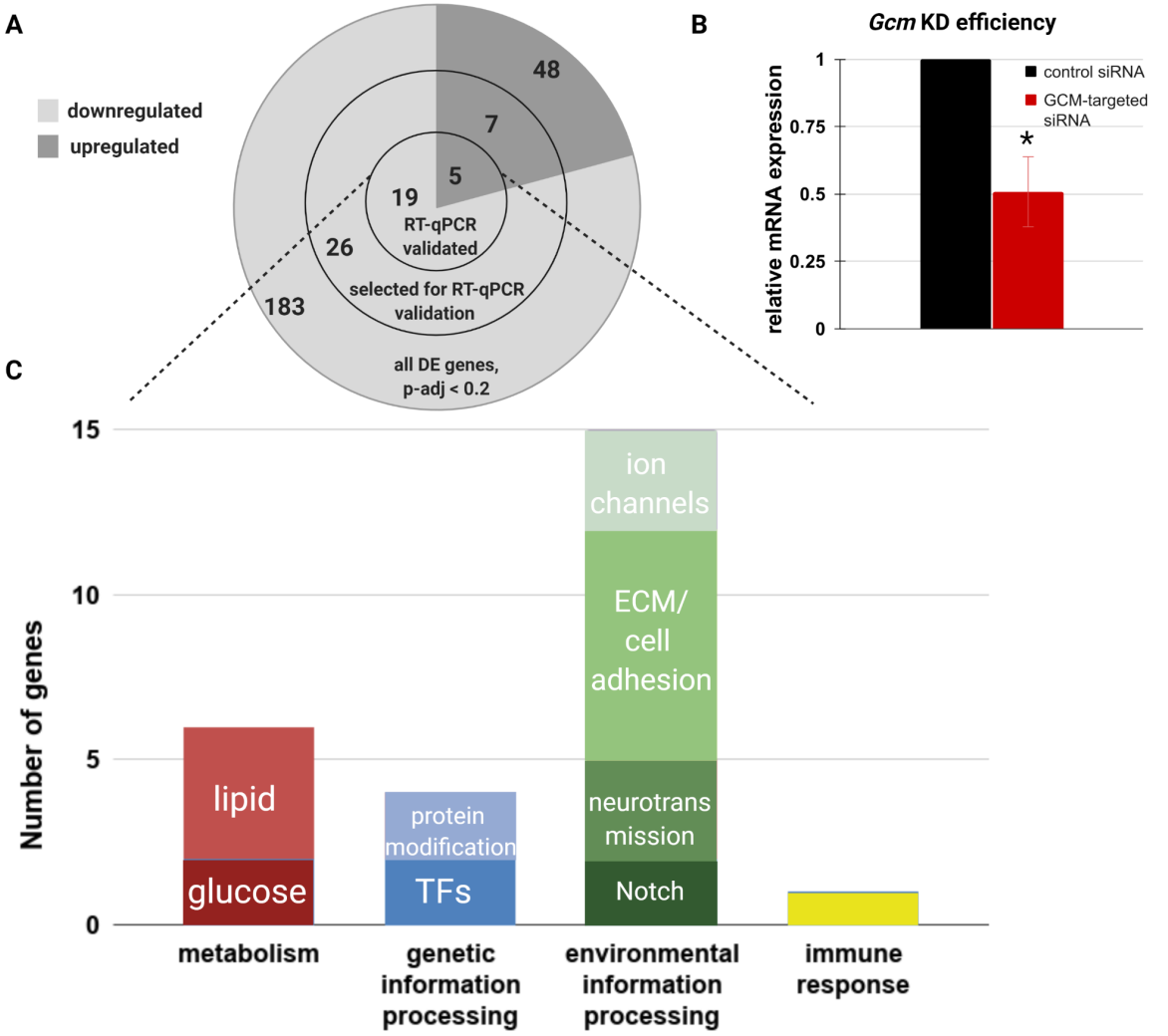


Figure 3.5. RT-qPCR DE gene validation. (A) Total number of DE genes identified by RNA-seq at p-adj < 0.2, number of DE genes selected for RT-qPCR validation, number of genes validated by RT-qPCR. (B) *Gcm* KD efficiency of siRNA2 for the experimental batches used for RT-qPCR validation. Data are given mean ±SD, N=8, *p < 0.05, unpaired t-test. (C) GO term distribution of RT-qPCR validated *Gcm* targets. Assembled with BioRender.com.

All the RT-qPCR validated *Gcm* targets, their functional classification and RNA-seq attributes are presented in Appendix 8. In addition to the 24 genes identified from the RNA-seq data, I confirmed two genes coding for GABA transporters (*GAT*s) to be affected by *Gcm* KD. Instead of a general GO term "membrane transport", I separated the genes belonging to this group into either "ion channel" or "neurotransmission" to describe their functions more precisely. Similarly, instead of a broad "carbohydrate/energy metabolism" GO term I use "glucose metabolism" function. I therefore grouped the validated *Gcm* targets into nine functional categories **(Figure 3.5C)** described below.

3.3.3 *Nematostella Gcm*-controlled program is glio-neuronal in nature

The homologs of *Nematostella Gcm* targets I identified are known to play an important role in neuronal and glial functions in bilaterians. Below I describe these in detail.

3.3.3.1 Metabolism

One of the classical functions of glial cells is to provide energy support to neurons. *Nematostella Gcm* controls the following genes involved in maintaining glucose level: *SGLT* (sodium-dependent glucose cotransporter) and *PKM2 (*pyruvate kinase); as well as lipid metabolism: *SORL1* (neuronal apolipoprotein E receptor), *P2Y6-like* (G protein-coupled P2Y receptor), and *PLA*s (phospholipases) **(Figure 3.6)**.

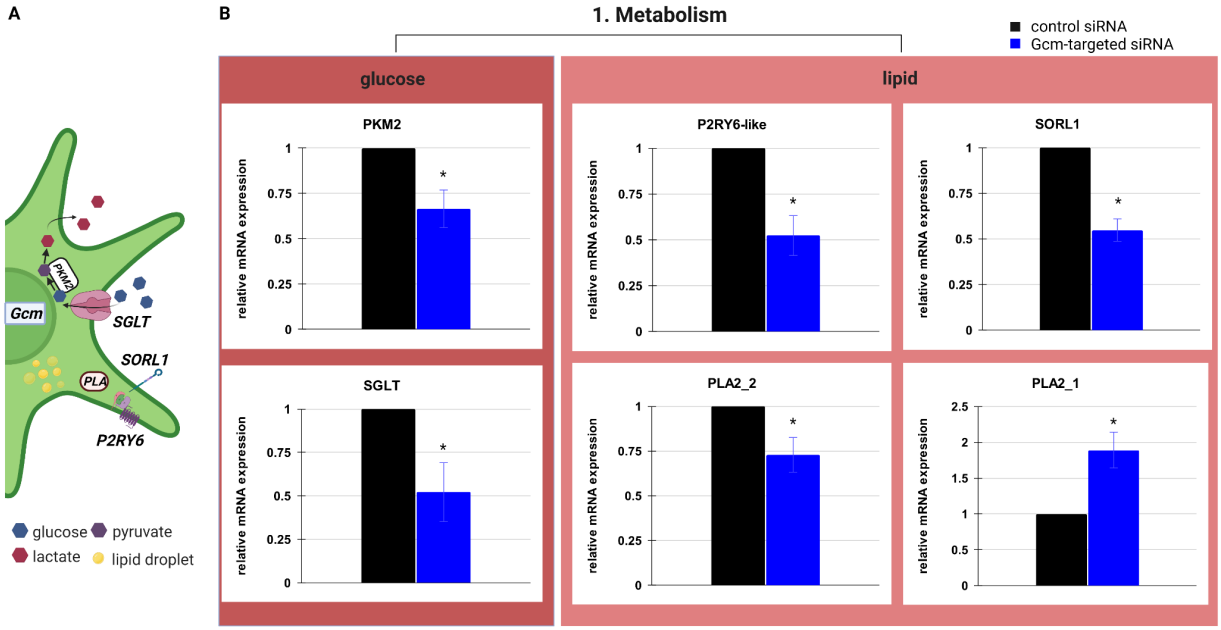


Figure 3.6. *Gcm* KD affects the expression of genes involved in glucose and lipid metabolism. (A) Cartoon of a *Gcm*-expressing cell equipped with receptors and enzymes required for glucose and lipid metabolism and storage. Created with BioRender.com (B) Expression change of identified genes in *Gcm* KD samples relative to control measured by RT-qPCR. Data are given mean ±SD, N=3, *p < 0.05, unpaired t-test.

3.3.3.1.1 Glucose metabolism

Astrocytes are known to provide nutrients to neurons. This includes storing glycogen and degrading it into lactate to provide energy to neurons upon increased neuronal activity

(Brown & Ransom, 2007; Freeman & Doherty, 2006). This process involves glucose entry into the glial cells via *GLUT* transporters and conversion into glycogen by several enzymes, followed by glycogen breakdown by a number of other enzymes (Brown & Ransom, 2007). In general, cellular glucose transport is mediated via either *GLUT* (glucose transporter) or *SGLT* (sodium/glucose cotransporter) protein families. The former allows for passive glucose transport downward concentration gradient. In the case of *SGLT* glucose is transported against its electrochemical gradient which is assisted by Na+ concentration gradient (Zhao & Keating, 2007). Certain members of the *GLUT* family are known to be expressed in the nervous system including vertebrate glial cells (Maher, 1995) as well as planarian glia (Wang et al., 2016). These transporters are considered important for glia to fulfill their function of providing energy to neurons. *SGLT* transporters are expressed mostly in kidneys and small intestine (Zhao & Keating, 2007), even though several studies showed SGLT expression in the brain specifically in the blood-brain barrier (Elfeber et al., 2004; Kepe et al., 2018; Oerter et al., 2019). Therefore, *SGLTs* play a role in the blood-to-brain movement of glucose especially under pathological conditions (Vemula et al., 2009).

I observed that in *Nematostella* a member of the *SGLT* family is downregulated in *Gcm* KD **(Figure 3.6B)**. The expression of another gene, *PKM2,* involved in glucose metabolism is also affected by *Gcm* KD **(Figure 3.6B)**. *PKM2* codes for a metabolic enzyme pyruvate kinase which mediates the last step of glucose to pyruvate conversion **(Figure 3.6A).** In mammals, similarly to other glycolytic enzymes, *PKM2* is expressed in both neurons and astrocytes. However, the latter are known as the primary glycolytic cells of the brain, providing energy to neurons in the form of lactate converted from glucose (astrocyte-neuron lactate shuttle) (Zhang et al., 2014).

3.3.3.1.2 Lipid metabolism

In addition to being a glucose reservoir in the brain, glia play an important role in lipid metabolism of the mammalian nervous systems (Ioannou et al., 2019; Thal, 2012). In *Nematostella Gcm* controls the expression of several genes involved in lipid processing **(Figure 3.6)**.

SORL1 (sortilin-related receptor) is a neuronal apolipoprotein E (*APOE*, protein involved in fat metabolism) receptor. It is a transmembrane receptor which in mammals is known to control intracellular processing of amyloid precursor protein (*APP*): *SORL1* binds *APP* and reduces production of amyloid-β (Aβ) peptide (Grear et al., 2009). *SORL1* was shown to be expressed in neurons and glial cells and secreted by astrocytes (Dowell et al., 2009; Grear et al., 2009). Moreover, *APOE* is mainly produced by astrocytes as well as neuronal stem cells (Zollo et al., 2017). Indeed, astrocytes are indispensable for *APP* processing and Aβ clearance (Thal, 2012). Overall *SORL1* might play an important role in lipid metabolism of the nervous system (T Cuenco et al., 2008). Astrocytes not only synthesize lipids more efficiently than neurons, which is important for synapse development (van Deijk et al., 2017), but also unlike neurons form lipid droplets (LDs) to store energy-rich lipids (Ioannou et al., 2019).

In addition to *SORL1,* I show that two phospholipases (*PLA2s)*, enzymes that cleave fatty acids of phospholipids, are affected by *Gcm* KD **(Figure 3.6B)**. Surprisingly, one *PLA* is downregulated, while the other is upregulated in *Gcm* morphants. This could be explained by genetic compensation, where the expression of *PLA2_1* is increased in response to *PLA2_2* expression decrease. Otherwise *PLA2_1* might be an indirect target of *Gcm* regulated by other

TFs. Indeed, unlike *Gcm*, *PLA2_1* is most abundant in gastrodermal cells at the larval stage of *Nematostella* (see section 3.4).

Another *Gcm* target is a G protein-coupled receptor, *P2RY6-like,* which upon activation in astrocytes was shown to activate *PLA2s* and stimulate amino acid release (Sun et al., 2004). This suggests an important role of *Nematostella Gcm*-expressing cells in lipid metabolism regulation **(Figure 3.6A)**.

3.3.3.2 Genetic information processing

3.3.3.2.1 Transcription factors

I identified and validated two TFs affected by *Gcm* KD. *Ets*-related TF is upregulated and a member of the zinc finger protein family, *ZICA*, is downregulated in *Gcm* morphants **(Figure 3.7)**.

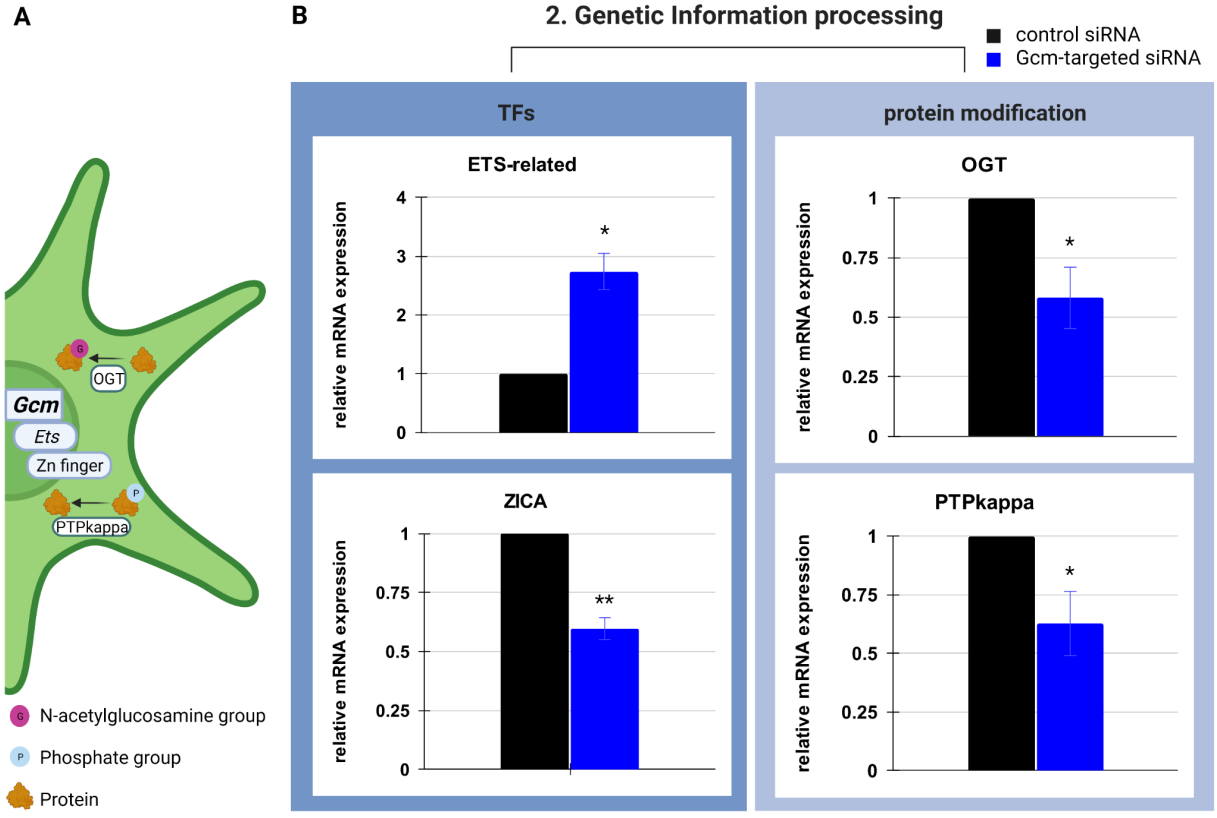


Figure 3.7. *Gcm* KD affects the expression of genes involved in genetic information processing. (A) Cartoon of a *Gcm*-expressing cell. *Gcm* regulated the expression of two TFs, *Ets* and Zn finger, as well as a protein glycosylation enzyme *OGT* and a dephosphorylation enzyme *PTPkappa*. Created with BioRender.com (B) Expression change of identified genes in *Gcm* KD samples relative to control measured by RT-qPCR. Data are given mean ±SD, N=3, *p < 0.05, **p<0.01, unpaired t-test.

Conversely, in *Drosophila PNT* (pointed, *Ets* protein family member) and *TTK69* (tramtrack 69, zinc finger TF), a promoter of glial fate and a suppressor of neuronal fate respectively, are downstream targets of *Gcm.* Various zinc finger and *Ets* family TFs are involved in neuronal and glial specification in bilaterians (see section 2.3.1.4). The *Ets*-related

gene affected by *Gcm* KD in *Nematostella* might drive the neurogenic, rather than gliogenic program. This might explain its upregulation in response to *Gcm* KD.

As expected from the expression analysis (see section 2.3.3.2), I did not observe any expression change of *Repo* (reversed polarity) in *Nematostella Gcm* morphants. In *Drosophila Repo* activates the glial program downstream from *Gcm* (Yuasa et al., 2003), It is possible that *Repo* got recruited as a *Gcm* cofactor once the bilaterians brunched off, and/or *Gcm*-regulated *Repo*-dependant glial developmental program is a protostome-specific feature.

3.3.3.2.2 Post-translational protein modification

Drosophila Gcm controlled program includes enzymes involved in post-translational protein modification (Altenhein et al., 2006). I show that *Nematostella Gcm* targets also include such enzymes, namely *OGT* (UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase) and *PTPkappa* (receptor-type tyrosine-protein phosphatase kappa) **(Figure 3.7)**. *PTPkappa* was shown to be indispensable for neural development in bilaterians because of its effect on adhesion molecules (Crossin & Krushel, 2000). Meanwhile *OGT* is an important metabolic sensor which is involved in glycolysis regulation and mediates glia-neuron interaction by promoting glial axonal support (Kim et al., 2016). This further suggests a high degree of *Gcm*-controlled program conservation in *Nematostella*.

3.3.3.3 Environmental information processing

3.3.3.3.1 Neurotransmitter transporters and ion channels

The regulation of ionic and neurotransmitter composition of the neuronal environment is a well-known glial function. Based on my results, *Nematostella Gcm* controls the following genes involved in fulfilling this function: *EAAT1* (glutamate transporter), *LRRC8A* (volume-regulated anion channel), *GAT*s (GABA transporters), *PKD1L2* (transient receptor potential polycystic (*TRPP*) channel), and *TRPA1* (transient receptor potential cation channel) **(Figure 3.8)**.

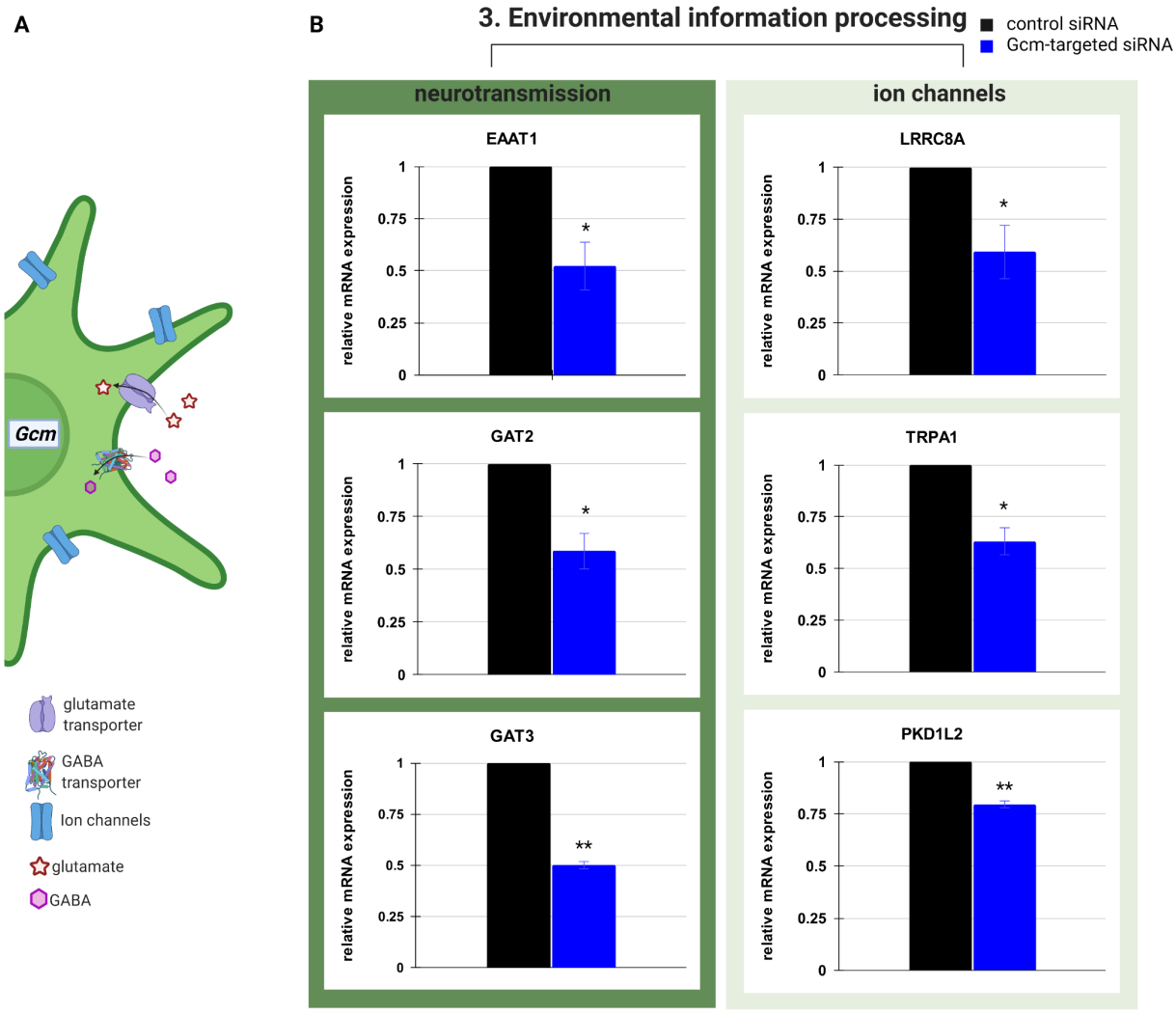


Figure 3.8. *Gcm* KD affects the expression of neurotransmitter transporters and ion channels. (A) Cartoon of a *Gcm*-expressing cell equipped with GABA and glutamate transporters and ion channels. Created with BioRender.com (B) Expression change of identified genes in *Gcm* KD samples relative to control measured by RT-qPCR. Data are given mean ±SD, N=3, *p < 0.05, **p<0.01, unpaired t-test.

Glutamate

In bilaterians *glia* are the key players in the metabolism of glutamate with the highest expression level of *EAAT2*, a transporter accountable for 95% of total glutamate uptake in the brain (Danbolt et al., 2016). *Nematostella* possesses three *EAAT-like* transporters (Anctil, 2009). I observed that one of these transporters, *EAAT1*, is downregulated in the *Gcm* morphants **(Figure 3.8B)**. In addition, a volume-regulated anion channel (*VRAC*), *LRRC8A,* is affected by *Gcm* KD. *VRAC*s are known to regulate cell volume by adjusting the intracellular concentration of negatively charged inorganic anions such as Cl^-, but also organic molecules such as glutamate, aspartate and taurine (Hyzinski-García et al., 2014). A member of the leucine-rich repeat containing protein 8 group, *LRRC8A*, is an essential *VRAC* of the rodent CNS that is indispensable for EAA release (including glutamate) from rodent glial cells (Hyzinski-García et al., 2014). Therefore, the fact that the key bilaterian glial TF, *Gcm,*

controls the expression of genes involved in glutamate metabolism in *Nematostella* points to the deep evolutionary roots of this glial function.

GABA

GABA is another neurotransmitter heavily relying on glia for reuptake from the extracellular environment. Moreover, in mammals GABA was shown to be synthesized and released by astrocytes thus prompting the researchers to consider glial cells GABAergic (Zorec et al., 2016). Indeed, glial cells contain various enzymes for GABA synthesis including glutamic acid decarboxylase (GAD), an enzyme that synthesizes GABA from glutamate - one of the major pathways for GABA synthesis in the nervous system (Ishibashi et al., 2019). Vesicular release of GABA from vertebrate glia is yet to be shown (Zorec et al., 2016). Glia is known to express both ionotropic and metabotropic GABA receptors localized close to the synaptic cleft to sense GABAergic signaling (Ishibashi et al., 2019). Importantly, several GABA transporters (*GAT1, GAT2, GAT3*) are expressed in neurons and astrocytes in vertebrates (Gadea & López-Colomé, 2001). *Drosophila* has one *GABA* transporter (*GAT*), which is expressed exclusively in astrocytes (Stork et al., 2014). Planarian glia was also shown to express *GAT* (Wang et al., 2016). I therefore sought to determine if *GAT*s are under the control of *Nematostella Gcm.*

Although in my RNA-seq data no *GAT*s were affected by *Gcm* KD, I identified *GAT*s in the *Nematostella* genome*,* and checked if *Gcm* KD affects their expression using RT-qPCR. I observed that two *GAT*s are downregulated in *Gcm* morphants **(Figure 3.8B)**.

Ion channels

Within the last few decades glial cells emerged as plastic cells sensing and reacting to neuronal activity by expressing various ion channels (Olsen et al., 2015). *Gcm* KD in *Nematostella* affected the expression of ion channels, *PKD1L2* and *TRPA1* **(Figure 3.8B)**.

PKD1L2 is a member of Polycystin 1 (PC1)-like proteins. PC1 is a 462 kDa integral membrane protein whose domain architecture suggests that it is involved in receptor signaling and cell adhesion. *PKD1L2* belongs to the *TRPP* subfamily of *TRP* (transient receptor potential) channels, which are cationic channels vital for sensory information processing. They transduce information about environmental stimuli by changing intracellular calcium concentration or membrane potential (Samanta et al., 2018). In the mammalian nervous system *PKD1L2* is known to be a marker of a specific subset of neurons (Ghazale et al., 2019). Other members of *TRPs* (*TRPMs*) are detected in various mammalian glial cells as well as planarian glia (Wang et al., 2016). *TRPA1* is another member of the TRP channel protein family. It is the only member of the mammalian *TRPA* subfamily, a chemo-nociceptor (Samanta et al., 2018), which is an important mediator of astrocytic Ca2+ signaling (Bosson et al., 2017).

3.3.3.3.2 The *Notch/delta* signaling pathway

I identified that among the genes affected by *Gcm* KD in *Nematostella* are those involved in the *Notch/delta* signaling pathway **(Figure 3.9)**.

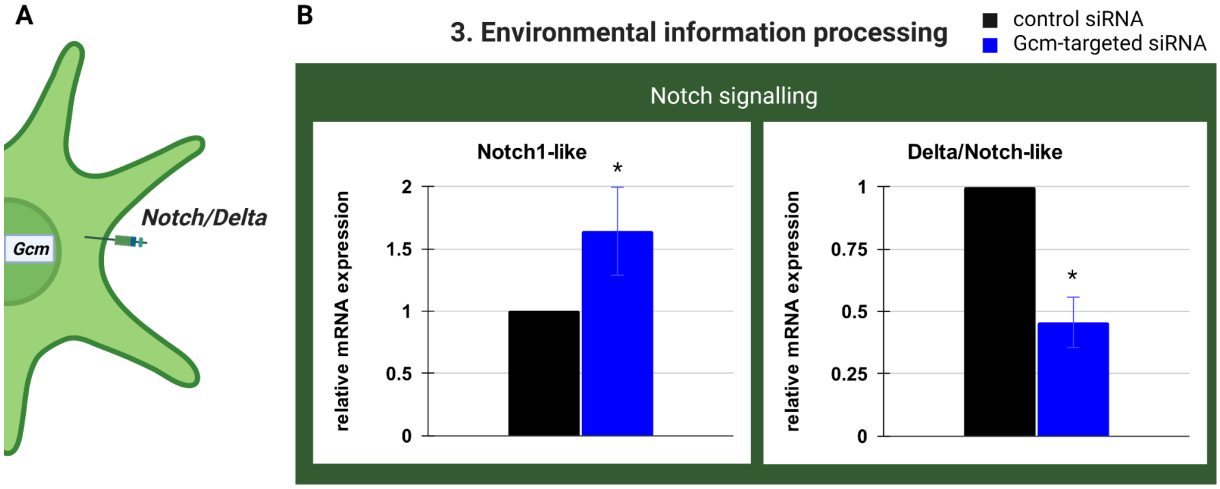


Figure 3.9. *Gcm* KD affects the expression of *Notch/Delta* signaling genes. (A) Cartoon of a *Gcm*-expressing cell equipped with members of *Notch/Delta* signaling pathway. Created with BioRender.com (B) Expression change of identified genes in *Gcm* KD samples relative to control measured by RT-qPCR. Data are given mean ±SD, N=3, *p < 0.05, unpaired t-test.

Notch signaling is a versatile pathway that plays a major role in neuro/gliogenesis in bilaterians (see section 1.2.4.2). Although *Nematostella* genome has several *Notch/Delta* homologs, the expression change of *Delta/Notch-like* genes in *Gcm* morphants implies an important role of the glial TF in regulating neuro/gliogenesis in *Nematostella* via this conserved pathway.

3.3.3.3.3 Cell-cell adhesion proteins and ECM constituents

The prevailing number of genes with known functions among *Gcm* targets I identified in *Nematostella* RNA-seq data and validated with RT-qPCR belong to the ECM/cell adhesion functional group **(Figure 3.4, 3.5)**. These include *CASPR2* (neurexin cell adhesion protein), *Contactin/DSCAM* (neuronal cell adhesion molecule), *SDK1* (cell adhesion protein), *ADGRB1* (thrombospondin 1-like), *ADGRV1* (adhesion G protein-coupled receptor), *EPR1* (aggrecan-related secreted proteoglycan), and an integrin ligand *EDIL3* (EGF-like repeat and discoidin I-like domain-containing protein 3) **(Figure 3.10)**.

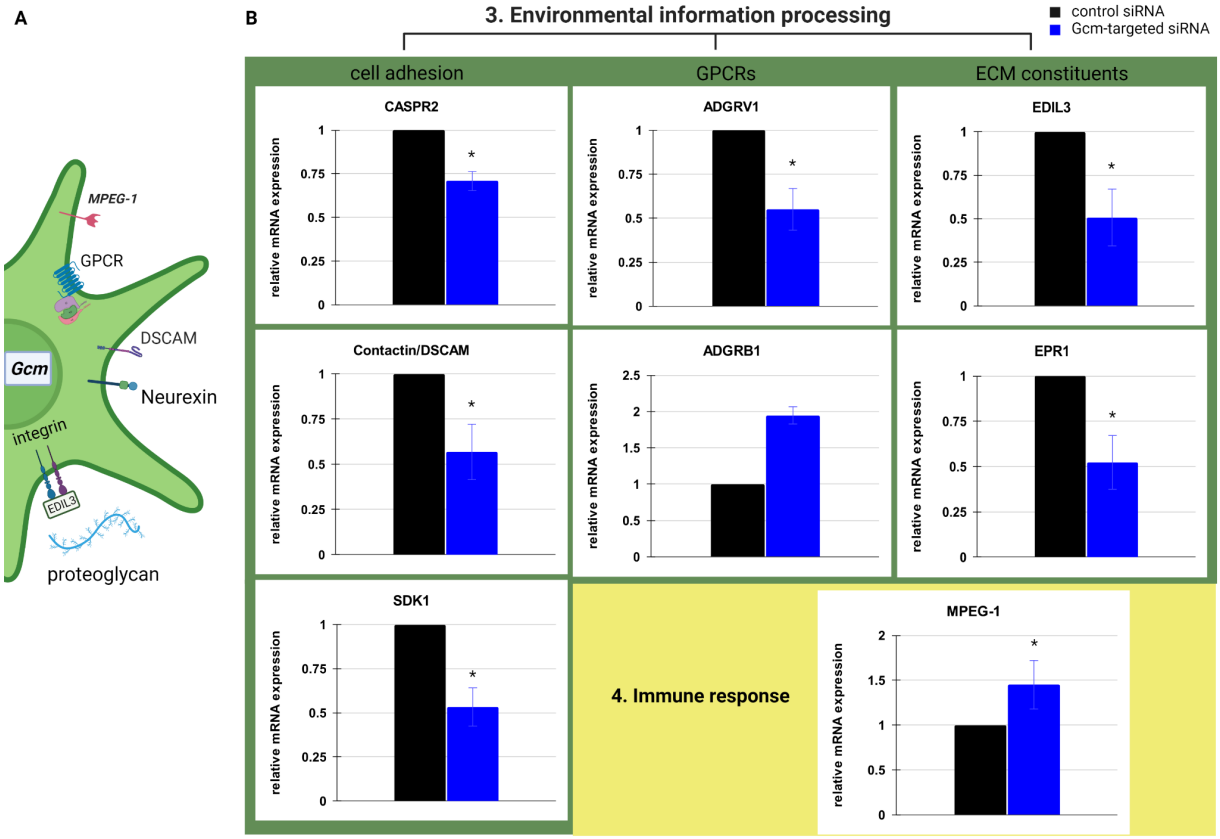


Figure 3.10. *Gcm* KD affects the expression of cell adhesion and ECM genes. (A) Cartoon of a *Gcm*-expressing cell equipped with several cell adhesion and ECM proteins as well as a protein involved in defense response *MPEG-1*. Created with BioRender.com (B) Expression change of identified genes in *Gcm* KD samples relative to control measured by RT-qPCR. Data are given mean ±SD, N=3, *p < 0.05, unpaired t-test. For *ADGRB1*: N=4, p=0.051.

CASPR2 (contactin-associated protein-like 2, encoded by *CNTNAP2* gene) is a cell adhesion protein belonging to the neurexin family (Saint-Martin et al., 2018). These are neuronal presynaptic cell adhesion proteins involved in the nervous system development, mainly in organizing synapses (Bang & Owczarek, 2013). Neurexins of both vertebrates and invertebrates are structural components of axo-glial septate junctions (SJs) (Baumgartner et al., 1996)**.** These genes are primarily expressed by neurons as well as glial cells, and crucial for regulating the interactions between these cell types (Proctor et al., 2015; Stork et al., 2009). Cnidarians possess several members of the neurexin protein family involved in the formation of SJs (Ganot et al., 2015). In *Drosophila* a neurexin protein, axotactin, is secreted by *Gcm*-expressing glial cells to be localized to axonal tracts (Yuan & Ganetzky, 1999). *Nematostella CASPR2* is closely related to *Drosophila* neurexin IV and axotactin, suggesting a conserved role of *Gcm*-expressing cells in establishing contacts with neurons and regulating their membrane excitability.

Another cell adhesion protein I identified as a *Gcm* target is one of *Nematostella* contactins/*DSCAMs* **(Figure 3.10A,B)**. These proteins are known to play an important role in the nervous system including synapse formation and axo-glial interactions (Chatterjee et al., 2019). Similar to neurexins, contactins/*DSCAMs* are involved in SJ organization and axonal guidance (Andrews et al., 2008; Banerjee et al., 2006; Mohebiany et al., 2014). *SDK1*

(sidekick 1) is a *DSCAM*-related adhesion molecule which also participates in nervous system development, particularly in synapse formation. Both *SDK* and *DSCAM* proteins belong to the immunoglobulin superfamily of cell adhesion molecules and have been extensively studied in the context of retina development in both vertebrates and invertebrates (Astigarraga et al., 2018; Yamagata & Sanes, 2008). Overall these proteins are indispensable for establishing neural connectivity in the bilateral nervous system.

Although for *ADGRB1* (adhesion G protein-coupled receptor B1) the expression change in *Gcm* morphants did not reach statistical significance, I found a clear tendency of this gene to be upregulated **(Figure 3.10B)**. *ADGRB1* is a protein containing thrombospondin (TSP) type 1 repeats and CD36 domain. TSP superfamily members are transmembrane and ECM proteins that play a role in nervous system development (Adams & Tucker, 2000). *Nematostella ADGRB1* contains a cell-binding motif (CD36) suggesting its role in cell adhesion and junction organization, and possible ECM interaction since CD36 is known to interact with ECM proteins (Pepino et al., 2014). In bilaterians thrombospondins are expressed by neurons and glial cells. Astrocyte-secreted thrombospondins, particularly TSP1, modulate synapse formation and induce neurogenesis (Dowell et al., 2009; Lu & Kipnis, 2010).

ADGRV1 is a transmembrane adhesion G protein-coupled receptor (GPCR) containing Calx-beta and CaCA (ca(2+):cation antiporter) domains involved in Ca+ transport. In mammals *ADGRV1* is primarily expressed in the nervous system with enriched expression in oligodendrocytes, thus playing an important role in myelination (Mehta & Piao, 2017). *ADGRV1* endogenous ligand is unknown (Mogha et al., 2016). In general adhesion GPCRs are a large group of proteins which play a crucial role in various processes including the nervous system development. It's been shown that adhesion GRCRs fulfill their functions primarily via interacting with ECM proteins such as collagens and laminins (Mehta & Piao, 2017).

I identified two ECM proteins affected by *Gcm* KD in *Nematostella: EPR1* and *EDIL3* **(Figure 3.10)**. Structurally *EPR1* matches aggrecan core protein and belongs to the aggrecan/versican family of proteoglycans that play a major functional role in cell-cell and cell-matrix interactions in the brain (Yamaguchi, 1996). Aggrecan is the primary component of perineuronal nets (PNNs) and an indispensable part required for PNN formation (Rowlands et al., 2018). PNNs are condensed structures of neural ECM that are placed around some neurons in between glial processes surrounding their synapses, which modulate brain development and neuronal plasticity (Giamanco & Matthews, 2012). PMM consists of various ECM proteins such as proteoglycans (heavily glycosylated proteins), glycoproteins, and hyaluronic acids. PNNs seem mammalian-specific, and were not found in invertebrates. In addition, I show that *EDIL3,* a glycoprotein which plays a role in ECM organization, cell adhesion and migration (Gasca et al., 2020), is a validated *Gcm* target**.**

Importantly, in *Drosophila* among common genes expressed in all glial cells are cell adhesion proteins required for axonogenesis and axon ensheathment like *Bdl* (borderless), *Mfas* (midline fasciclin), *Tsp5D* (tetraspanin) (Croset et al., 2018). Two innexins (*Inx2* and *ogre*), which belong to intercellular channels constituting gap junctions, are also shared among all *Drosophila* glia cell types, and are essential for normal development of the insect nervous system (Holcroft et al., 2013). Moreover, *Drosophila Gcm* directly regulates ECM and cell adhesion proteins which are indispensable for glial development and neuronal ensheathment. In addition, apart from several TFs regulated by *Gc,* a glycoprotein (*ana)* secreted by glia is the only other gene required for glial function and neuronal fate suppression in *Drosophila*

(Freeman et al., 2003). Therefore, the *Gcm*-controlled program of *Nematostella* seems to resemble that of *Drosophila*.

3.3.3.4 Immune response

In *Drosophila Gcm* regulates some genes required for fulfilling immune functions. A well-known example is a cell-surface receptor *Draper* (extracellular EGF-repeats and a novel intracellular domain), involved in apoptotic cell corpse engulfment and dendritic pruning (Freeman et al., 2003; Fullard & Baker, 2015). My analysis revealed the expression changes of several immune function-related genes in *Gcm* KD. This suggests that the regulation of microglial (macrophage) functions such as engulfment of cellular debris might be a conserved feature in *Nematostella Gcm.* One of these genes is *ADGRB1* discussed above, which was previously shown to have apoptotic functions in vertebrates (Sokolowski et al., 2011). Another example is *MPEG-1*, a marker of zebrafish microglia known to play a protective role against pathogens (Preston et al., 2018). However, my analysis shows that both *ADGRB1* and *MPEG-1* are upregulated in *Gcm* morphants **(Figure 3.10)**. This suggests that *Nematostella Gcm* could repress microglial fate instead of promoting it.

3.3.4 *Gcm*-expressing cells have a distinct morphology

Analysis of the *Gcm*-controlled program revealed its dual neuro-glial nature. Next I sought to characterize the morphology of *Gcm*-expressing cells to see if they demonstrate features of neurons or glial cells.

First, I analyzed the mRNA expression of *Gcm*. For that purpose, I used RNAscope®, a recently developed technology for mRNA detection, which had not been tried in *Nematostella* before. Although detecting *Gcm* expression is useful to visualize the cells, TFs are mostly confined to the nucleus and the cytoplasm. This means TF detection does not reveal all the morphological features of the cells, e.g. its processes. Therefore, I used commercially available antibodies for two transmembrane proteins, *Eaat1* and *Adgrv1,* affected by *Gcm* KD, to get more insights into the *Gcm*-expressing cell structure.

EAAT1 is a glutamate transporter, which is primarily expressed in astrocyte processes in mammals (Hayashi & Yasui, 2015). In *Nematostella* it is abundantly expressed in neuronal

cells with the highest expression of *Gcm* at both adult and larval stages **(Figure 3.11)**.

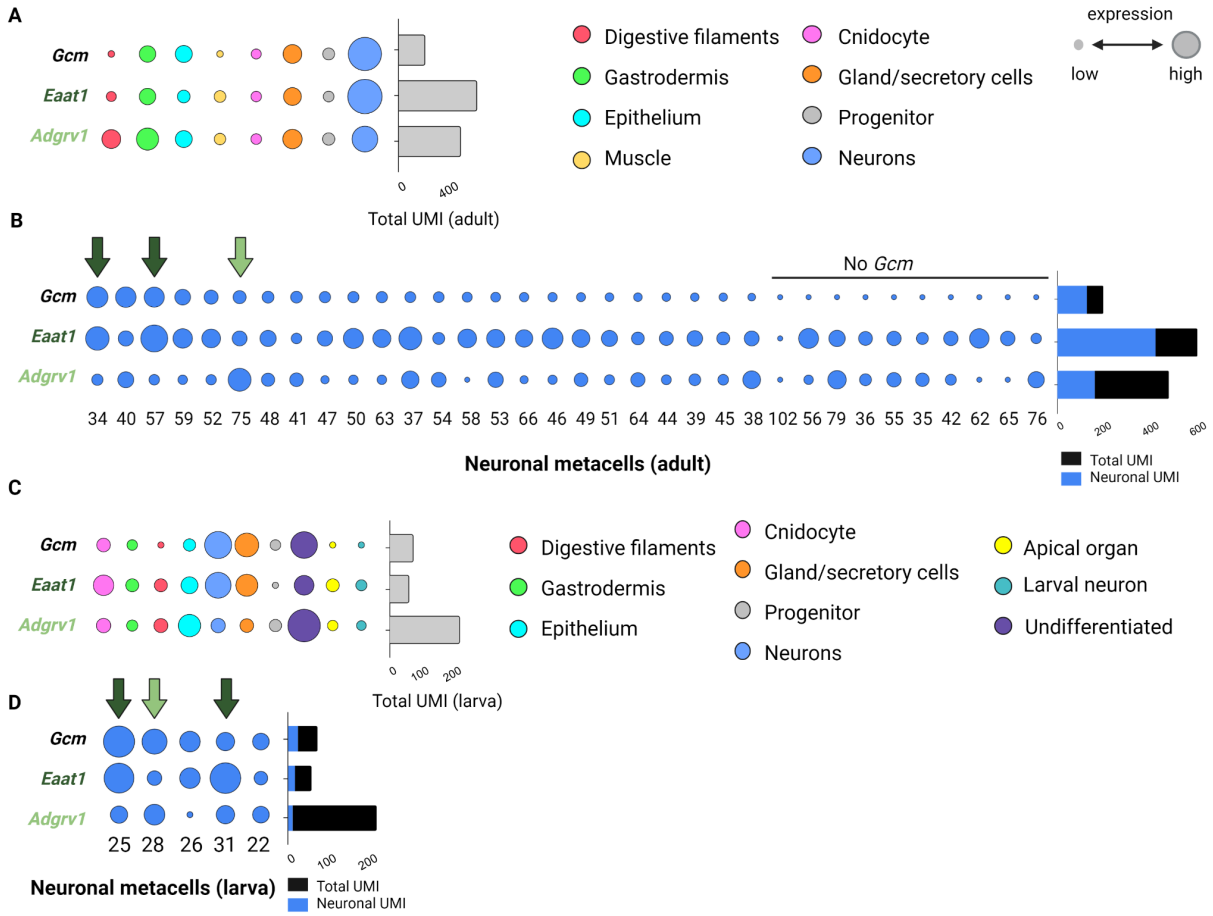


Figure 3.11. Expression of *Gcm* and its targets *Eaat1* and *Adgrv1* in adult and larval *Nematostella*. (A) Expression of the genes across different cell types in adult animals is shown. All three genes demonstrate highest expression in neurons. (B) Expression (molecules/1,000 UMIs) of the genes across neuronal cell clusters (metacells) in adult *Nematostella*. Neuronal cell clusters are organized left to right in the descending order of *Gcm* expression. Dark green arrow indicates the metacell with the highest expression of *Eaat1*. Note that the most abundant expression of *Gcm* and *Eaat1* coincide. Light green arrow indicates the metacell with the highest expression of *Adgrv1*. (C) Expression of the genes across different cell types in larva is shown. *Gcm* and *Eaat1* show predominantly neuronal expression unlike *Adgrv1*. (D) Expression (molecules/1,000 UMIs) of the genes across neuronal cell clusters (metacells) in larval *Nematostella*. Neuronal cell clusters are organized left to right in the descending order of *Gcm* expression. Dark and light green arrows indicate the metacells with the highest expression of *Eaat1* and *Adgrv1* respectively. Assembled with BioRender.com.

This indicates the co-expression of *Gcm* and *Eaat1*. EAAT1 immunostaining should therefore reveal *Gcm*-expressing cell morphology.

ADGRV1 is a GPCR, which is broadly expressed in various cell types, and is not highly abundant although present in *Gcm*-enriched neuronal cells **(Figure 3.11)**. Thus, in addition to revealing some cells expressing *Gcm,* ADGRV1 immunostaining could identify the cells interacting with *Gcm*-expressing cells.

3.3.4.1. Morphology of *Gcm*-positive cells revealed by RNAscope®

RNAscope® is a novel *in situ* hybridization assay, which allows high-resolution detection of transcripts. The original protocol was optimized and successfully used for zebrafish embryos (Gross-Thebing et al., 2014). I adopted the protocol for *Nematostella* to detect *Gcm*-positive cells in the planula stage. These cells are scattered throughout ectoderm and endoderm mostly in the aboral region, and demonstrate an elongated shape **(Figure 3.12)**.

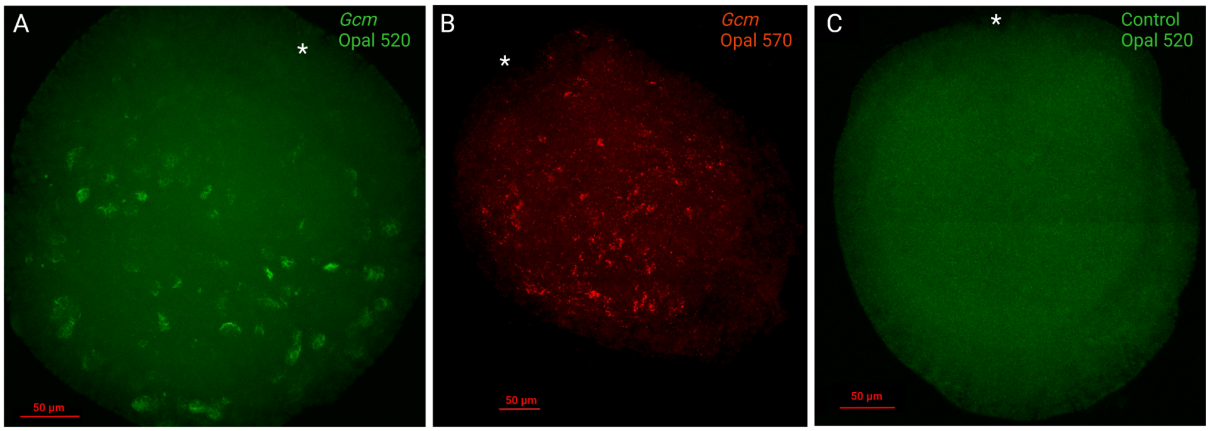


Figure 3.12. *Gcm*-positive cells in *Nematostella* planulae. Two Opal dyes were used to detect the same probe for comparison. (A) and (B) *Gcm*-specific probe was detected using Opal 520 and Opal 570 dyes respectively. (C) Negative control (probe for bacterial DapB) was detected using Opal 520 dye. Scale bar: 50 µm. Oral end is marked by an asterisk. Standard WISH *Gcm* probe was used as a positive control (Figure 2.3).

3.3.4.2. *Gcm*-expressing cells are not classical neurons

3.3.4.2.1 Antibody specificity

To confirm the specificity of commercially available antibodies for *Nematostella* EAAT1 and ADGRV1 I performed Western blot analysis. Both antibodies against EAAT1 detect a protein of the expected size of ~65 kDa **(Figure 3.13A,B)**. EAAT1 expression can be detected at both 4 and 5dpf. An antibody against ADGRV1 detects a protein of the expected size of ~95 kDa with high abundance at 5dpf **(Figure 3.13C)**.

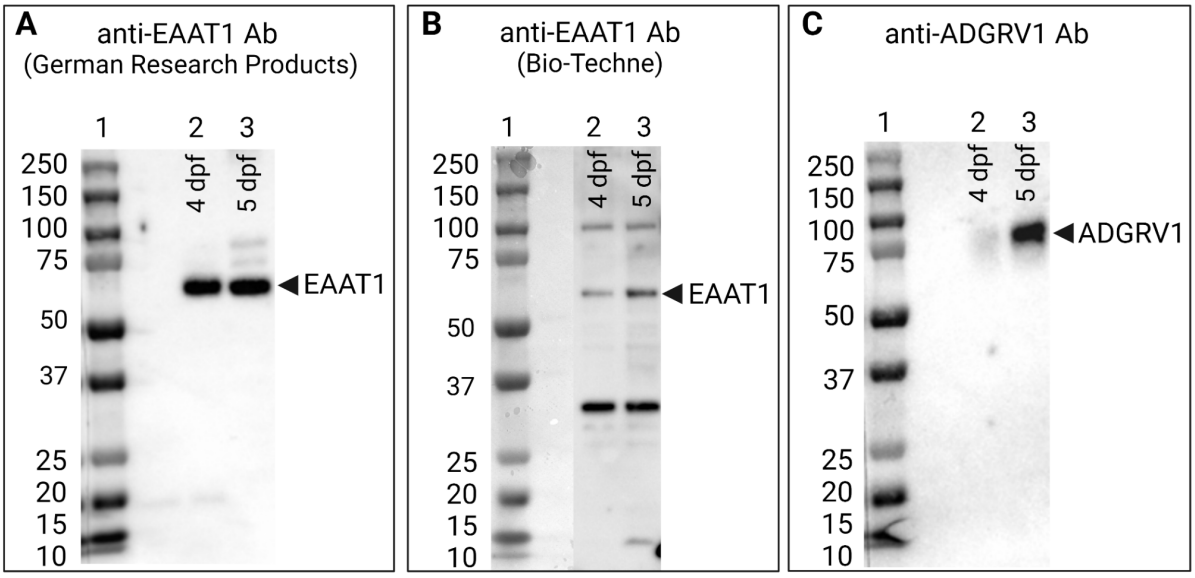


Figure 3.13. Western blots for *Nematostella* EAAT1 and ADGRV1 proteins. Lane 1 shows Molecular weight markers (indicated in kDa, kilodaltons). Lanes 2 and 3 show 4dpf and 5dpf expressions of probed proteins. Blot A was probed with anti-EAAT1 antibody (German Research Products). A single specific band of ~65 KDa is detected as expected based on the *Nematostella* EAAT1 protein sequence. Blot B was probed with anti-EAAT1 antibody (Bio-Techne). Several bands are detected including the band of ~65 KDa (*Eaat1*). Blot C was probed with ADGRV1 antibody. A single band of ~95 KDa is present as expected based on the *Nematostella* ADGRV1 protein sequence. Dpf - days post fertilization.

Next, to confirm that the antibody against EAAT1 (German Research Products) recognizes the protein of interest, I performed *Eaat1* KD using gene-specific siRNAs followed by qPCR and WB with a loading control **(Figure 3.14)**. Following *Eaat1* knockdown, the reduced protein expression was detected with WB, although not significant, indicating that the antibodies should recognize *Nematostella* EAAT1 **(Figure 3.14)**.

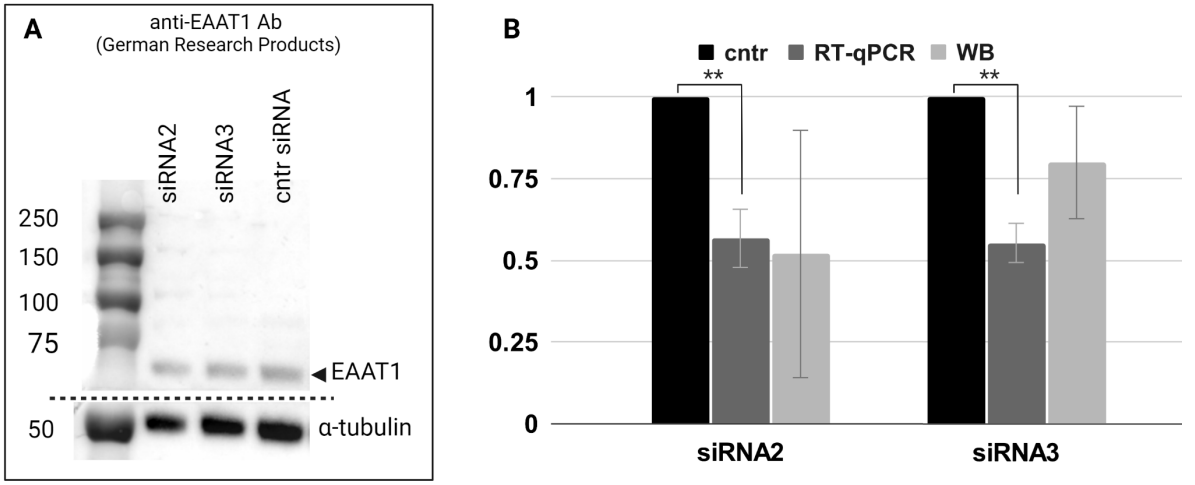


Figure 3.14. Anti-EAAT1 antibody (German Research Products) recognizes *Nematostella* EAAT1. (A) Western blot with a loading control (*a-tubulin*) to compare relative expression of EAAT1 in KD vs. control samples probed with anti-EAAT1 antibody. (B) Expression reduction (fold change) of *Eaat1* following KD as confirmed by RT-qPCR and WB analysis. Data are given mean ±SD, N=3, **p < 0.01, unpaired t-test.

3.3.4.2.2 EAAT1-positive cells are flat and processless

Immunofluorescent staining (IF) using both antibodies against EAAT1 revealed the same expression pattern in primary polyps **(Figure 3.15E,G)**. Thanks to the low background it was possible to detect EAAT1-positive cells using anti-EAAT1 antibody (German Research Products) at different developmental stages **(Figure 3.15A-E)**.

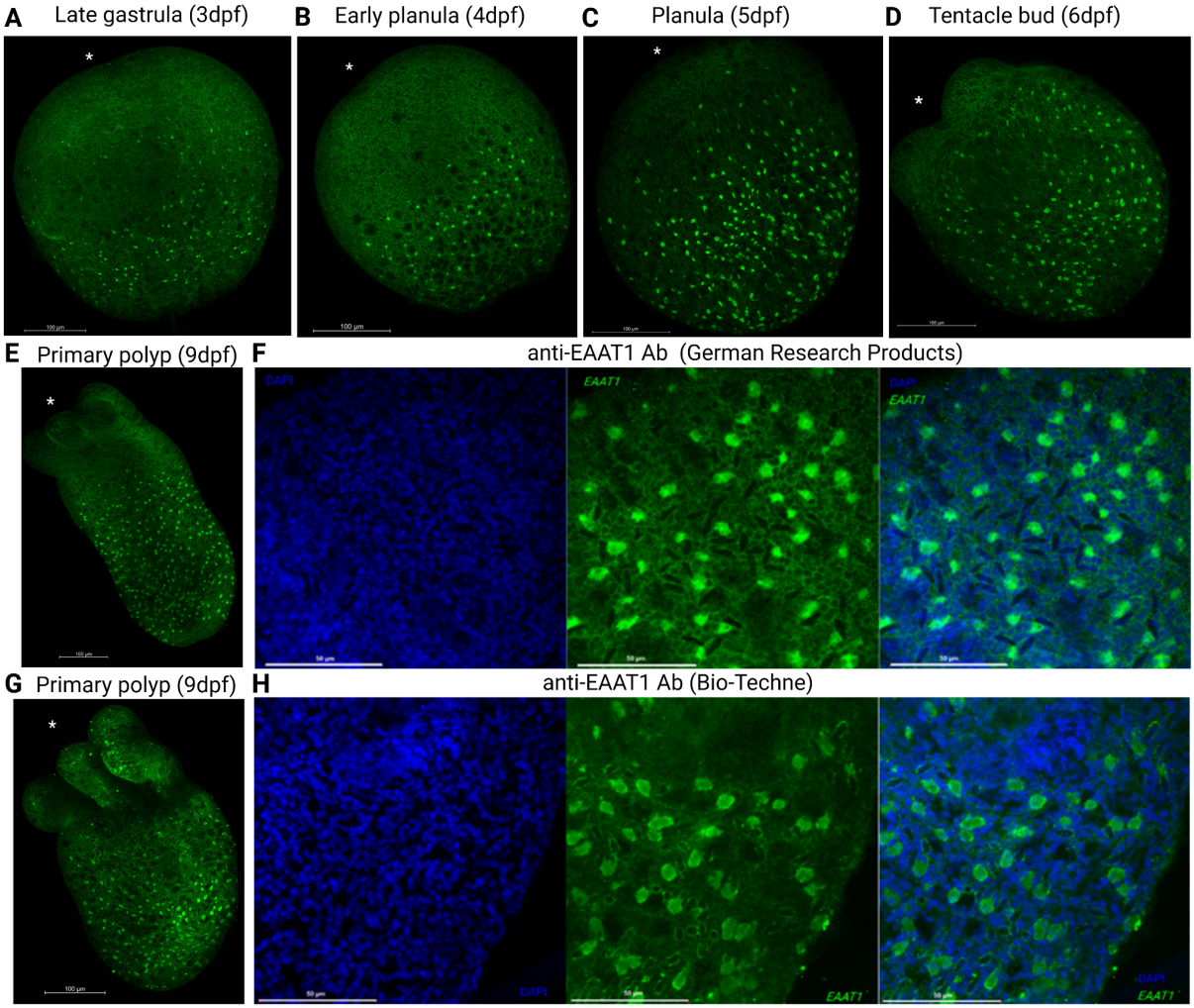


Figure 3.15. EAAT1-positive cells in *Nematostella* at different developmental stages (A) Late gastrula stage (3dpf) (B) Early planula stage (4dpf) (C) Planula stage (5dpf) (D) Tentacle bud stage (6dpf). (E-H) Primary polyp stage (9dpf) (A-F) Immunofluorescence staining using anti-EAAT1 antibody (German Research Products). (G-H) Immunofluorescence staining using anti-EAAT1 antibody (Bio-Techne). (A-E, G) Scale bar: 100 μm. (F,H) EAAT1 and DAPI staining. Scale bar: 50 μm. Oral end is marked by an asterisk.

EAAT1-positive cells start being detected at late gastrula stage in the outer ectoderm of the aboral region **(Figure 3.15A)**. At late gastrula and early planula stages only a few cells are detected **(Figure 3.15A-B)**. Numerous EAAT1-positive cells are found in a broad aboral region from planula through primary polyp stages **(Figure 3.15C-E,G)**. They demonstrate a flat processless morphology and do not extend neurites **(Figure 3.15 F,H)**. The morphological features of EAAT1-positive cells suggest that they are not classical *Nematostella* neurons.

The nervous system of *Nematostella* is composed of several morphologically distinct neurons broadly subdivided into ganglion cells (interneuron-like cells) sensory/sensory-motor cells, and cnidaria-specific nematocytes (Layden et al., 2016; Tournière et al., 2020). ELAV1, SoxB, and various neuropeptides are generally used to visualize *Nematostella* neurons (Galliot et al., 2009; Rentzsch et al., 2020). Some of them are distributed throughout the body forming prominent tracts of neurites along the mesenteries. Others form oral and pharyngeal nerve rings. A typical neuropeptide-positive neuron is spindle-shaped at the planula stage, and has neurites extending from a bulbous soma at polyp stage as revealed by PRGamide immunostaining in **Figure 3.16A,D**. At primary polyp stage PRGamide-positive neurons are detected along the mesenteries **(Figure 3.16B)**.

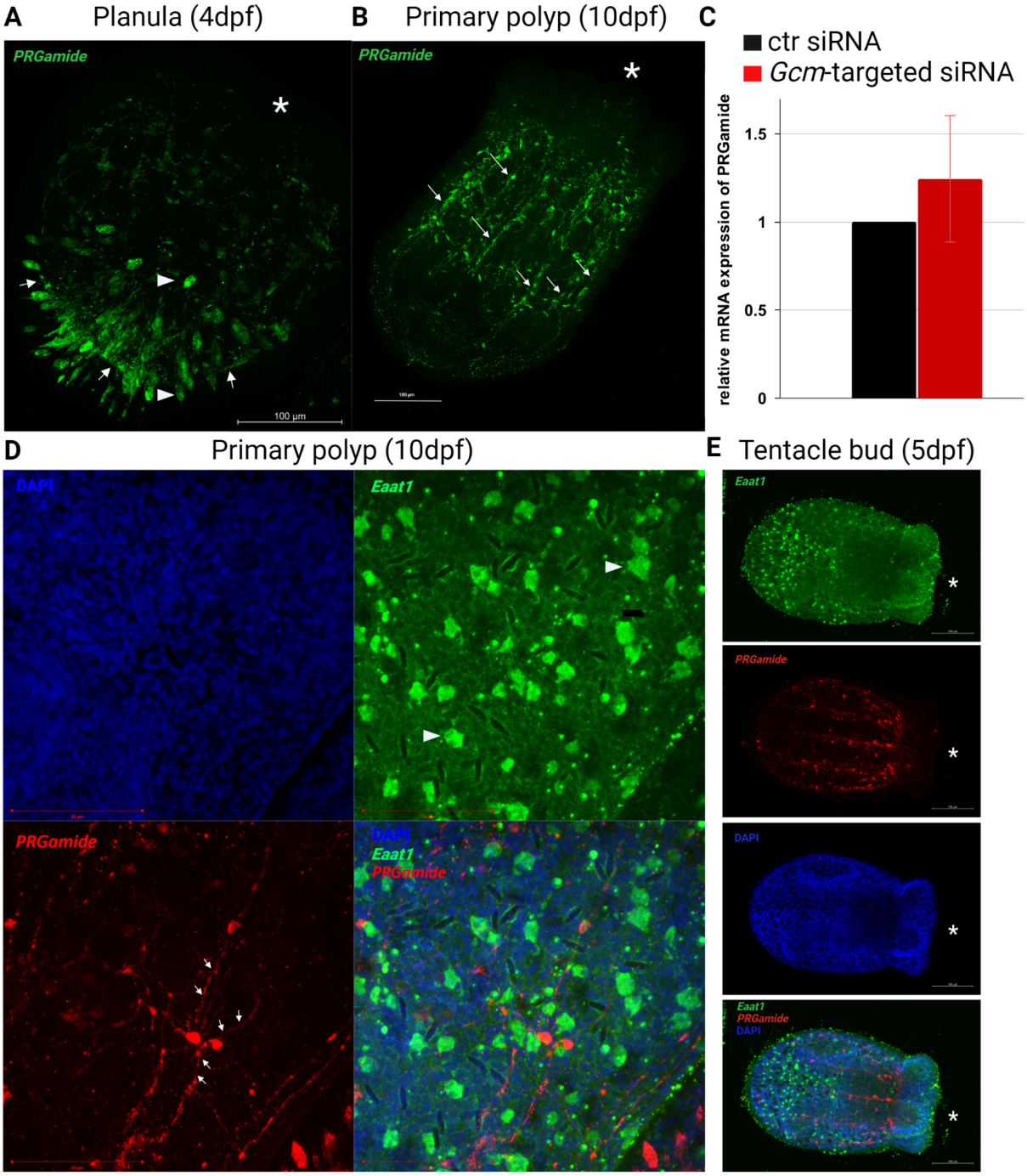


Figure 3.16. Morphology of peptidergic neurons and EAAT1-positive cells in *Nematostella*. (A) Spindle-shaped sensory neurons (white arrowheads) forming an aboral nerve net (white arrows) labeled by PRGamide immunostaining at planula (4dpf) stage (B) PRGamide neuron labeling at primary polyp (10 dpf) stage. White arrows indicate longitudinal tracks of neurites running along the mesenteries. (C) No expression change of PRGamide in Gcm KD samples relative to control was observed as measured by RT-qPCR. Data are given mean ±SD, N=3. (D) Immunofluorescence co-staining using anti-EAAT1 antibody (Bio-Techne) and anti-PRGamide antibody at primary polyp stage (10dpf). EAAT1 and PRGamide stainings do not overlap. EAAT1-positive cells are processless (white arrowheads) unlike PRGamide-positive neurons, which extend neurites (white arrows). (E) At

Caption continuation for Figure 3.16:
tentacle bud stage (5dpf) EAAT1 is expressed in a broad aboral region, whereas PRGamide immunostaining labels neurons along the mesenteries. Oral side is marked by an asterisk. Scale bars: 100 μm (A,B,E), 50 μm (D).

PRGamide is a neuropeptide which at the larval stage of *Nematostella* has the highest expression in the same neuronal cell cluster as *Gcm* (Appendix 10A). However, it is neither affected by *Gcm* KD **(Figure 3.16C)**, nor is it abundant in *Gcm*-rich cell clusters in adult *Nematostella* (Appendix 10B). Moreover, *PRGamide* and *Eaat1* are expressed in different cells in *Nematostella* at tentacle bud and primary polyp stages **(Figure 3.16D-E)**.

RNA-seq analysis of *Gcm* morphants revealed that none of the neuropeptides is affected by *Gcm* KD. In addition, the flat neuriteless morphology of EAAT1-positive cells is strikingly different from typical *Nematostella* neuropeptide-positive neurons. Hence, *Gcm*-expressing cells are a new type of neuronal cells that have not been described in *Nematostella* before.

3.3.4.2.3 ADGRV1-positive cells resemble sensory cells

ADGRV1-positive cells are found in the aboral ectoderm at all developmental stages of *Nematostella* **(Figure 3.17)**.

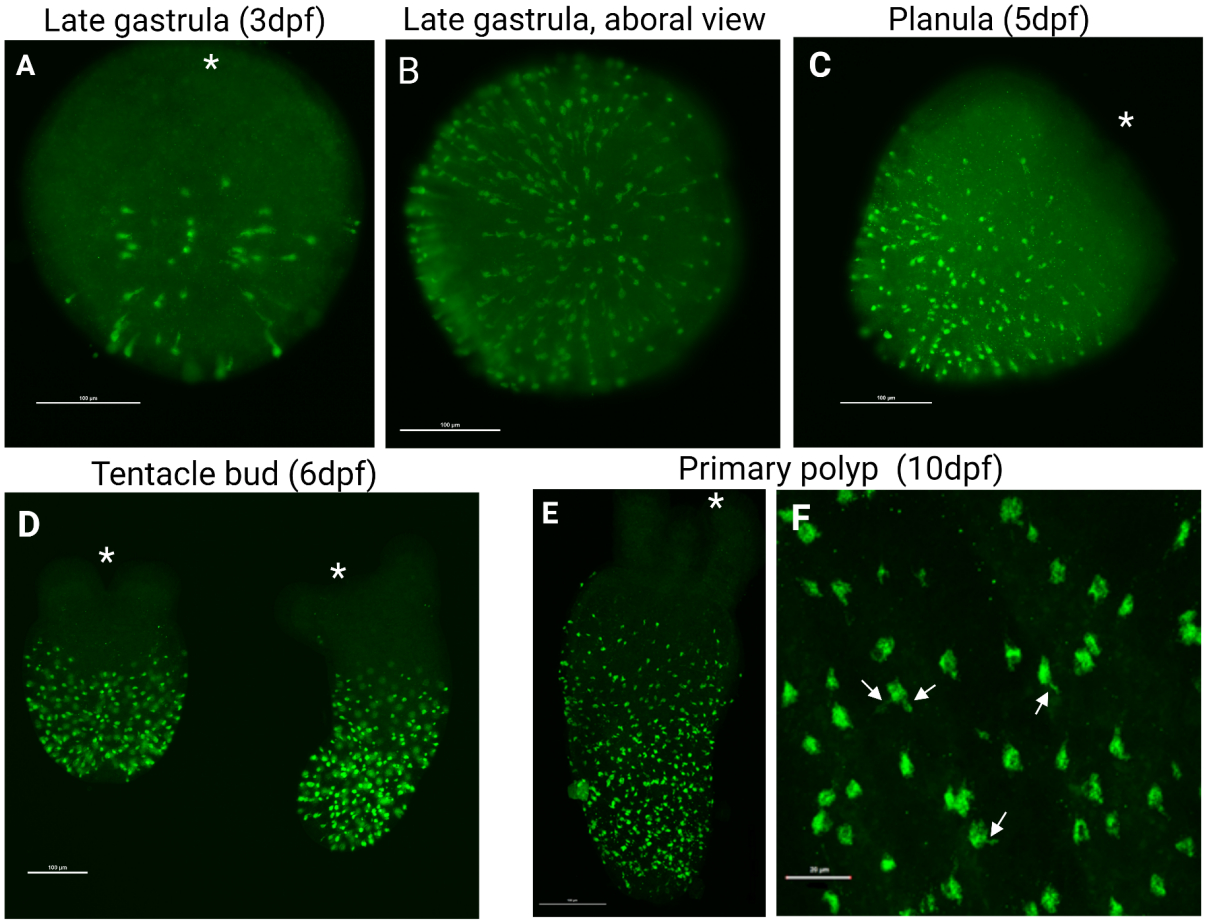


Figure 3.17. ADGRV1-positive cells in *Nematostella* at different developmental stages. In all the stages analyzed the expression of ADGRV1 is confined mostly to the aboral region. (A) Late gastrula stage (3dpf) (B) Aboral view of late gastrula stage (3dpf) (C) Planula stage (5dpf) (D) Tentacle bud stage (6dpf). (E,F) Primary polyp stage (10dpf). Short processes extending from the soma are detected (white arrows). Scale bars: 100 μm (A-E), 20 μm (F). Epifluorescence (A-D) and confocal images (E-F). Oral end is marked by an asterisk.

At a gastrula/early planula stage the expression of ADGRV1 is detected in a few scattered cells in the aboral region **(Figure 3.17A,B)**. The expression then expands towards the oral side, and more cells can be detected **(Figure 3.17C)**. The oral part of *Nematostella* as well as tentacles remain free of ADGRV1 even at tentacle bud and primary polyp stages **(Figure 3.17D,E)**. At planula stage these cells have slender bodies with a single cilium spanning the aboral ectoderm **(Figure 3.17A-C)**. At the primary polyp stage ADGRV1-positive cells extend 1-2 short processes **(Figure 3.17F)**. These morphological features conform to previously described sensory cells in *Nematostella,* e.g. labeled in the NvFoxQ2d::mOrange transgenic line (Busengdal & Rentzsch, 2017). The ADGRV1 Ab staining seems to support the existence of neurite-less cells, but further analysis remains required as I was not able to examine the antibody specificity by WB.

3.4 Discussion

Nematostella vectensis is an excellent model organism to study the evolution of the nervous system thanks to the privileged position of cnidarians as a sister group to *Bilateria*. In this work for the first time *Nematostella* was used as a model to explore the conservation of the bilaterian gliogenic program in cnidarians. I focused on *Gcm* as the "core" glial TF in bilaterian animals. The phylogenetic analysis I performed (see Chapter II) revealed that among basal metazoans *Gcm* is conserved only in cnidarians. I therefore sought to explore if *Gcm* emergence set the stage for glia emergence. In this chapter I employed a well-established way of revealing the functions of a gene by knocking it down and identifying its targets. Using the available single cell transcriptome of *Nematostella* I've explored the expression of *Gcm* and its targets. Immunostaining for a *Gcm* target, glutamate transporter *Eaat1,* revealed a new type of neuronal cells lacking neurites scattered throughout the aboral ectoderm.

This study has certain limitations. First, it is still unclear if *Gcm, Eaat1,* and *PRGamide* are expressed in the same cells at different developmental stages. Double *in situ* combined with immunostaining if technically possible is necessary to prove the co-expression. Second, *Gcm* KD followed by immunostaining for EAAT1 and ADGRV1 should be done to reveal if the two are indeed regulated by *Gcm*. Third, co-labeling of EAAT1 and markers known to be widely expressed in *Nematostella* neurons such as ELAV or SoxB2 would be useful to confirm if EAAT1-expressing cells can be characterized as neuronal. Forth, only partial KD of *Gcm* was achieved using targeted siRNA. Knockout experiments, e.g. using CRISPR/Cas9 generated mutants given the manipulation is not fatal, could be useful to further clarify *Gcm* function. Finally, generating a transgenic line to trace *Gcm*-expressing cells and their progeny would shed light on the development, morphological and molecular characteristics of these cells. Meanwhile SCT is a useful tool to get an insight into *Gcm*-expressing cells.

3.4.1. *Nematostella Gcm* is a neuronal TF with gliogenic potential

My results support a conserved function of *Gcm* as a neuronal TF in *Nematostella*. First of all, it is expressed mostly in neuronal cells at both larval and adult stages **(Figure 3.18)**.

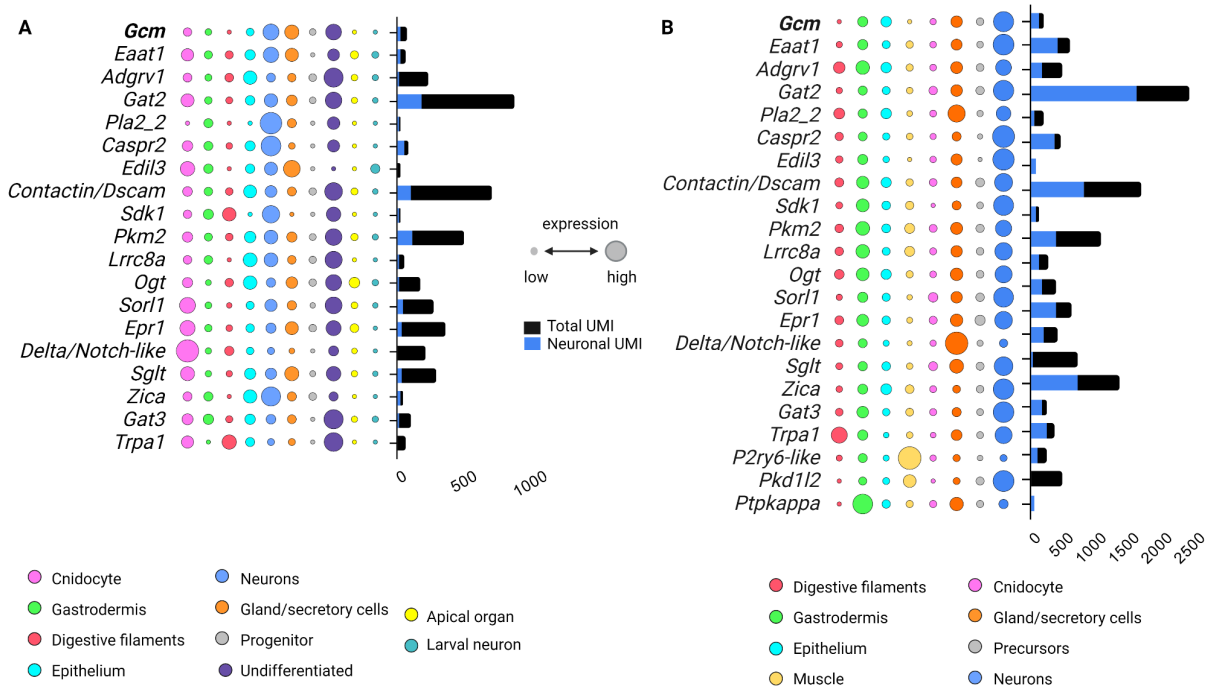


Figure 3.18. Expression of *Gcm* and genes affected by *Gcm* KD in *Nematostella*. (A) Expression (molecules/1,000 UMIs) of the genes across larval cell types. (B) Expression (molecules/1,000 UMIs) of the genes across cell types in the adult animal.

Although the expression level of *Gcm* is low and broad particularly at larval stage **(Figure 3.18A)**, it can be detected and is functional at early planula stage **(Figure 2.3E, 3.12)**, which is when the neurogenesis takes place.

In addition, *Nematostella Gcm* KD uncovered that the *Gcm*-controlled program is neuro-glial and resembles that of *Drosophila*. First, similarly to *Drosophila, Nematostella Gcm* regulates ECM and cell adhesion proteins, post-translational protein modification enzymes, *Ets* and Zn finger TFs, as well as genes involved in immune response. Some of these genes are indeed abundant specifically in neuronal cells of *Nematostella* especially at adult stage, e.g. *Caspr2*, *Sdk1, Zica, Mpeg1* **(Figure 3.18)**. Second, *Nematostella Gcm* regulates genes involved in fulfilling classical glial functions in bilaterians such as lipid and glucose metabolism, and the regulation of ionic and neurotransmitter composition of the neuronal environment. Among these some are abundant in neuronal cells at both larval and adult stages, e.g. *Eaat1, Gat2, Pla2_2* **(Figure 3.18)**. Others, such as *Ogt, Sorl, Gat3*, are predominantly expressed in neuronal cells at adult stage **(Figure 3.18B)**. Although it is difficult to assign a strict glial role to these genes, some of them are known to be expressed specifically in glial cells in bilaterians. Finally, *Nematostella Gcm* regulates the expression of *Notch/Delta* signaling genes. This pathway is vital for both neuro- and gliogenesis in bilaterian animals including *Drosophila*. In *Nematostella Notch* regulates the development of neurons and cnidocytes (Marlow et al., 2012; Rentzsch et al., 2017). Surprisingly, *Gcm* KD increased the expression of one *Notch* homolog while decreasing the expression of another homolog. The former is expressed mostly in cnidocytes, and the latter is most abundant in gastrodermis at planula stage **(Figure 3.18A)**. Likewise, in adult *Nematostella* both genes are not highly expressed in neurons **(Figure 3.1B)**. This might indicate that *Notch* genes are indirect targets

of *Gcm* or a non-neuronal effect of *Gcm* KD, since *Gcm* is expressed in other cell types as well.

Importantly, many of the genes affected by *Gcm* KD are broadly expressed in various cell clusters at both larval and adult stages of *Nematostella*. Some of them are most abundant in non-*Gcm*-rich cell clusters (not shown). This might mean that they are indirect targets of *Gcm*. Additional studies are required to clarify this. Computational analysis to identify *Gcm* functional binding sites combined with chromatin immunoprecipitation (ChIP) technique could be useful to distinguish direct target genes of *Nematostella Gcm*. Meanwhile exploring *Gcm*-rich neuronal cell clusters could provide an in-depth view on the characteristics of these cells.

3.4.2. Glial and peptidergic programs intertwine

Given that *Nematostella* nervous system is highly peptidergic, I explored if *Gcm*-expressing cells are also rich in neuropeptides. At planula stage *Gcm* is most abundant in a peptidergic neuronal cell cluster 25 **(Figure 3.19A**, Appendix 10A).

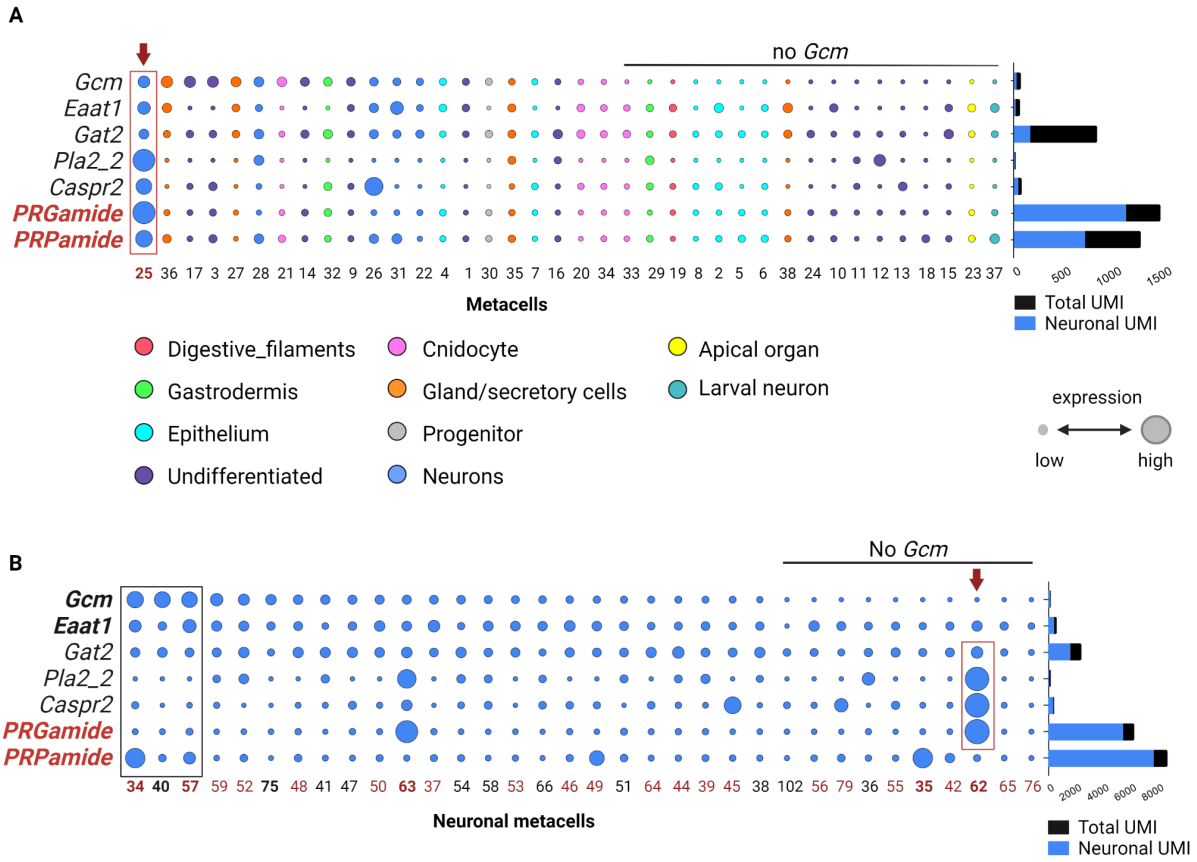


Figure 3.19. Expression of *Gcm*, some genes affected by *Gcm* KD, and two neuropeptides in *Nematostella*. Cell clusters are organized left to right in the descending order of *Gcm* expression. (A) Gene expression across larval cell clusters (metacells). All the genes are abundant in peptidergic cell cluster 25 (red frame and arrow) (B) Gene expression across adult neuronal cell clusters (metacells). *Gcm* is most abundant in the peptidergic neuronal cell cluster 34, followed by the non-peptidergic neuronal cell cluster 40, and the peptidergic neuronal metacell 57, characterized by the highest abundance of *Eaat1*. High expression of *Eaat1* and *PRPamide* is found in these clusters (black frame). *PRGamide* and several genes affected by *Gcm* KD (red frame) are highly co-expressed in the

Caption continuation for Figure 3.19:
peptidergic neuronal cell cluster 62 (red arrow). A cluster is defined as peptidergic (red) if the expression of any neuropeptide was detected in it (data not shown).

This cluster is enriched with neuropeptides, *PRGamide* and *PRPamide.* It is also characterized by the highest expression of some genes affected by *Gcm* KD. These include a glutamate transporter, a GABA transporter, a phospholipase, and a neurexin (**Figure 3.19A**, Appendix 10A). Therefore, early in the development *Nematostella Gcm*-rich cells are peptidergic neurons expressing neural and glial genes, with *Gcm* regulating the glial program of neurotransmitter and lipid metabolism.

At adult stage *Gcm* is most abundant in three neuronal cell clusters 34, 40, 57 (**Figure 3.19B**, Appendix 10B). Cluster 34 can be characterized as peptidergic because of its enrichment with *PRPamide,* although the most abundant expression of *PRPamide* is in neuronal cluster 35. Cluster 40 is not peptidergic. Cluster 57 has some neuropeptide abundance, but *PRPamide* is four times less abundant compared to clusters 34 and 35. Interestingly, only *Gcm* and *Eaat1* stay abundantly co-expressed, whereas the other genes affected by *Gcm* KD are now highly expressed in a separate cell cluster 62 (**Figure 3.19B**, Appendix 10B). This cluster also has the highest expression of *PRGamide*, but is completely devoid of *Gcm*. Thus, as *Nematostella* develops, *Gcm*-expressing cells lose their neuropeptidergic nature.

In accordance with single cell transcriptome (SCT), *Eaat1* and *PRGamide* immunostaining should overlap at the larval stage **(Figure 3.19)**. However, the two antibodies seem to label morphologically different cells at planula through primary polyp stages **(Figure 3.15, 3.16)**. The co-staining using both antibodies confirmed the non-overlapping labeling at the bud stage **(Figure 3.16)**. Unfortunately, it was technically difficult to achieve clear labeling with both antibodies in the same animals at the planula stage. Given that both *PRGamide* and *Eaat1* are expressed in a subset of cells in the aboral region **(Figure 3.15B, 3.16A)**, there might be cells co-expressing both proteins at this stage. Further analysis is required to confirm this. The discrepancy between the single cell transcriptome and immunostaining data might be explained by: 1) larval SCT is the expression profile pooled from several developmental stages (2 dpf, 4dpf, 7dpf), and does not reflect the differences between these stages 2) mRNA and protein expression level do not perfectly correlate. It was previously shown that mRNA expression levels are not always predictive of protein abundances (Kuchta et al., 2018). Studies investigating the correlation between mRNA and protein expression levels focused on either human cells or yeast. Overall positive correlation was found. However, one study found that the correlation varied depending on the gene ontology: genes of regulation showed lowest, while extracellular region and signal transducer activity genes showed the highest correlation (Guo et al., 2008). Therefore, conclusions regarding protein expression based on mRNA abundance should be made with caution, especially in less studied animals such as *Cnidaria*.

If early in the development *Gcm* and neuropeptides are indeed co-expressed, the developmental division of *Nematostella* peptidergic *Gcm*-expressing cells into peptidergic neurons and glia-neuron intermediates might recapitulate the evolutionary path of glial cells. Early in *Nematostella* development *Gcm*-expressing cells are peptidergic progenitors, which by means of consumptive asymmetric division (Busengdal & Rentzsch, 2017) give rise to peptidergic neurons and protoglia in adult animals each with partially retained developmental potential of the parental *Gcm*-expressing cell **(Figure 3.20)**.

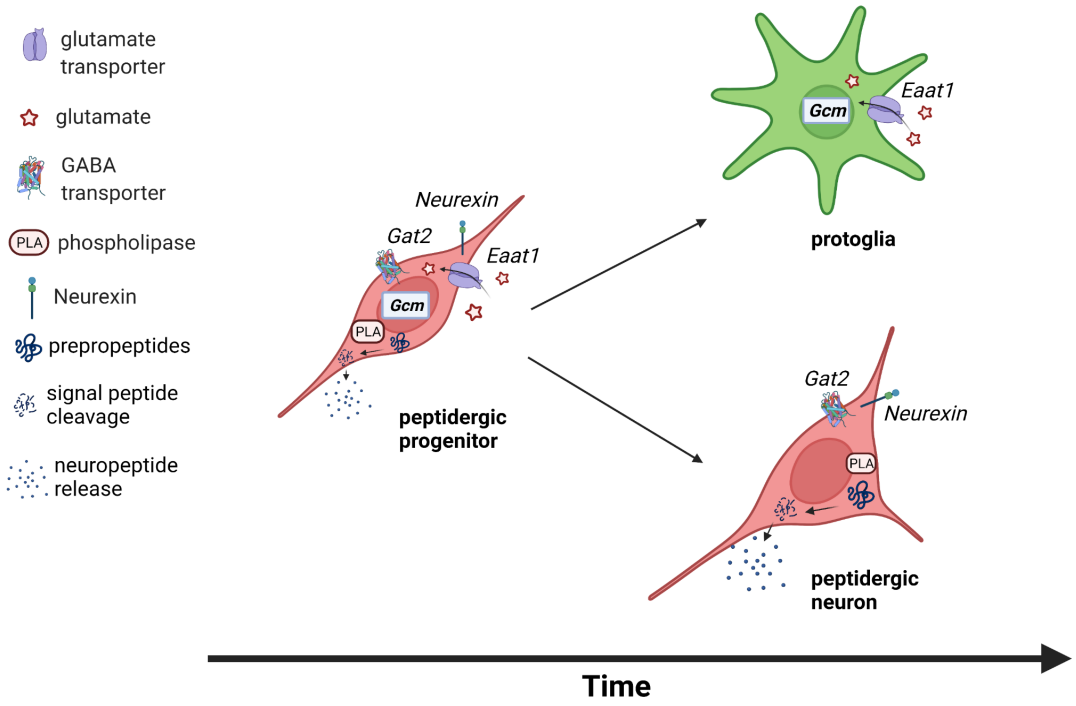


Figure 3.20. Hypothetical representation of how the development of *Nematostella Gcm*-expressing cells might recapitulate glial evolution. Peptidergic pro-neurons might already be equipped with the *Gcm*-controlled glial program which includes a glutamate transporter. These cells could give rise to two neuronal cell types: peptidergic neurons and "protoglia". The latter would inherit the *Gcm*-regulated program and provide glutamate metabolic support to neurons. Created with BioRender.com.

The protoglia is characterized by a high expression of *Eaat1* and demonstrates non-neuronal morphology as revealed by *Eaat1* immunostaining **(Figure 3.15)**. Unlike peptidergic neurons, *Eaat1*-positive cells do not extend neurites and do not form networks. At the same time, processless *Eaat1*-positive cells do not resemble classical glial cells either. Similarly, the evolutionary origins of glia might lie in peptidergic neurons **(Figure 3.20)**. Neurosecretion is thought to predate fast chemical neurotransmission. It's been long debated if cnidarian nervous system is strictly peptidergic. As discussed in chapter 2.1.1 more evidence gets accumulated showing that this is not the case.

3.4.3. Chemical neurotransmission in *Cnidaria* in glial context

The function of small chemical neurotransmitters, GABA and glutamate, is assumed to have evolved from being chemicals to sense the environment (observed in *Porifera* and *Placozoa*, 'chemosensory synapse') to becoming neurotransmitters to relay internal information within a synapse (*Cnidaria*, *Bilateria*, and possibly *Ctenophora*) (Moroz et al., 2021). Therefore, GABA and glutamate transmission predates the emergence of the NS in animal evolution. It is possible that in some animals, like *Cnidaria*, small transmitters perform both neuronal and non-neuronal functions. The detection of glutamate and GABAergic related gene expression in non-neuronal cells in a sea anemone *Nematostella* cast doubt on whether these small molecules act as neurotransmitters in *Cnidaria* (Oren et al., 2014). Nevertheless,

all the receptors, transporters, enzymes required for uptake, vesicular packaging, and production of these neurotransmitters are present in the genomes of several cnidarian models including *Nematostella* (Pierobon, 2021). Moreover, single-cell transcriptome data demonstrates the prevalent neuronal expression, i.e. function, of the genes involved in both glutamate and GABA signaling **(Figure 3.21, 3.22)**.

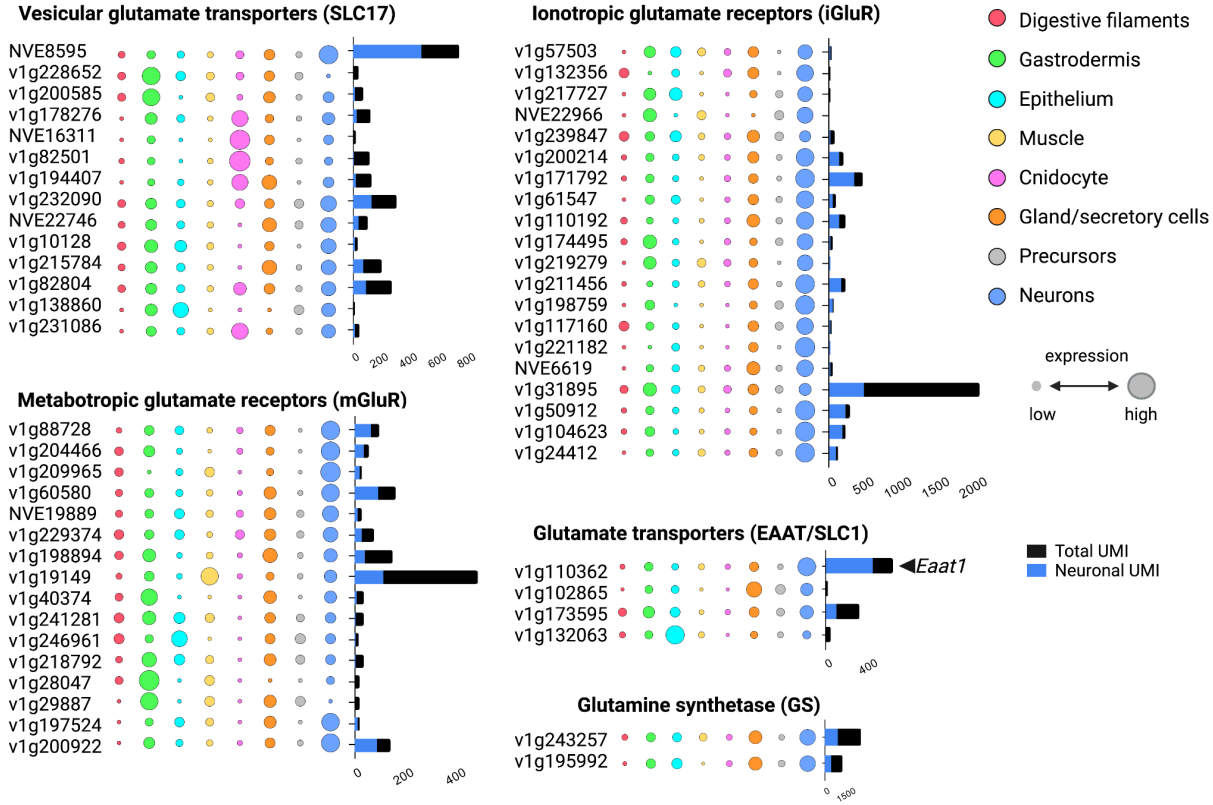


Figure 3.21. Expression of genes involved in glutamate synthesis and signaling in different cell types of adult *Nematostella*. *Eaat1*, a *Gcm* target, identified in this study, is marked with an arrowhead. Assembled with BioRender.com

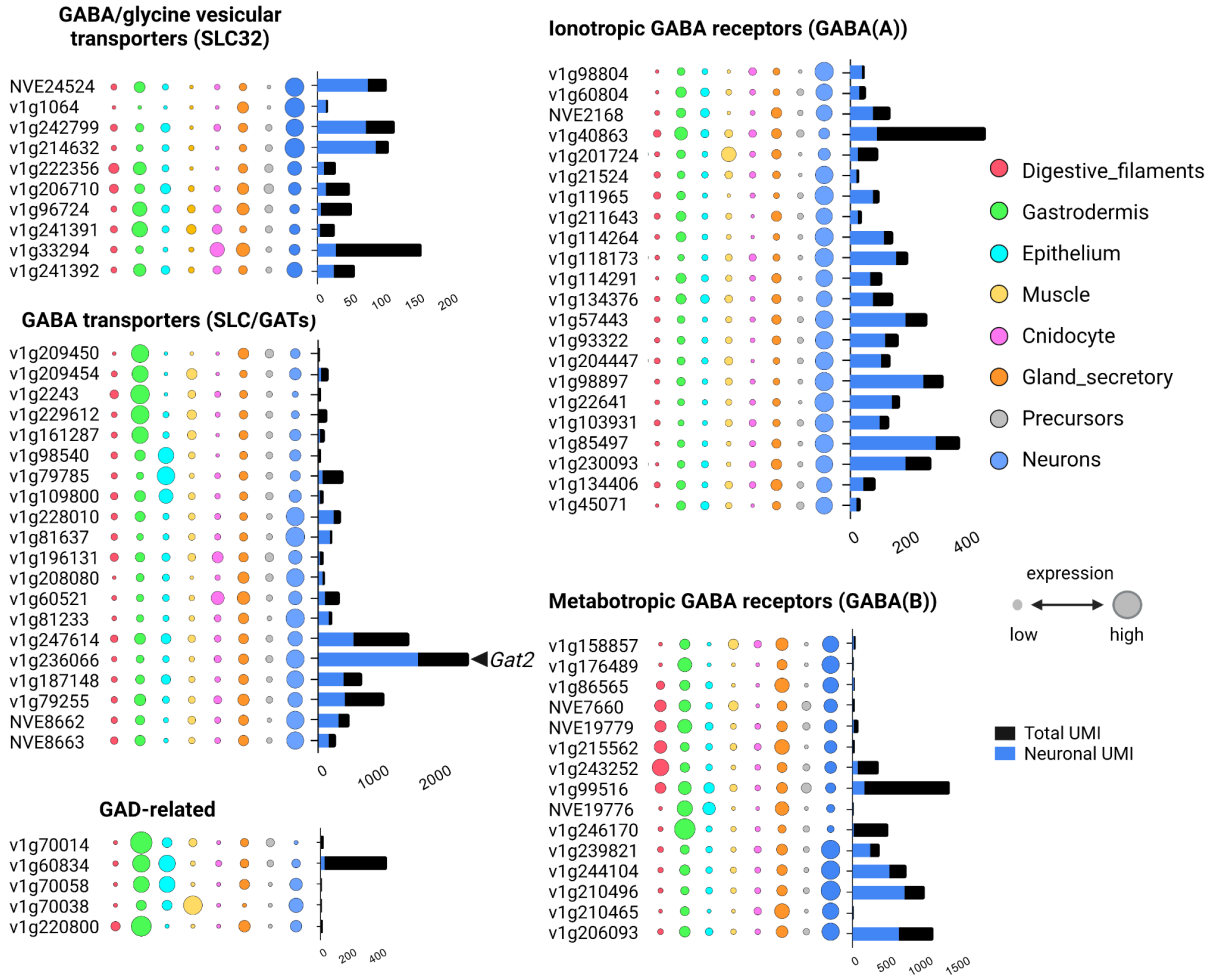


Figure 3.22. Expression of genes involved in GABA synthesis and signaling in different cell types of adult *Nematostella*. *Gat2*, a *Gcm* target, identified in this study, is marked with an arrowhead. Assembled with BioRender.com

Apart from being enriched in neuronal cells, some of these genes are enriched in other cell types. Some of the vesicular glutamate transporters are most abundant in cnidocytes, which are considered a type of cnidarian neuronal cells (**Figure 3.21).** Meanwhile some of the metabotropic glutamate and GABA receptors, as well as plasmalemmal GABA transporters, are highly expressed in *Nematostella* gastrodermis (**Figure 3.21, 3.22**). Moreover, the expression of GAD, an enzyme that catalyzes the conversion of glutamate to GABA, is mostly confined to gastrodermis **(Figure 3.22).** This raises a question regarding neuronal synthesis of GABA to be used as a neurotransmitter in *Nematostella*. While it is still debated if glutamate and GABA are used as neurotransmitters in *Cnidaria* due to the lack of direct evidence, *Nematostella* SCT analysis suggests that at least glutamate is likely to be synthesized and released via vesicular packaging from neurons.

Nevertheless, my analysis showed that *Gcm* regulates a glutamate transporter, *Eaat1*, and a GABA transporter, *Gat2*. Both are most abundant in neuronal cells (**Figure 3.21, 3.22).** Given that *Eaat1* is also co-expressed with *Gcm* (i.e. its putative target) at both larval and adult stages **(Figure 3.19)**, one of the major functions of protoglial cells might be glutamate quenching. This explains a conserved glial function of glutamate metabolism observed

throughout bilaterian glia. Thus, it's possible that both neuropeptide and glutamate neurotransmission programs co-existed and subsequently split up to give rise to not only glutamatergic and peptidergic neurons, but also glutamate metabolizing glia.

3.4.4. Novel view on glial evolution

Although adult *Gcm*-rich cell clusters are classified as neuronal because of their enrichment with classical neuronal markers, *Gcm* KD and the analysis of its targets allowed me to tease out the gliogenic nature of the *Gcm*-regulated program. Indeed, the molecular composition of *Gcm*-rich cell clusters is not obviously homologous to the bilaterian glial transcription signature. Among *Nematostella Gcm* co-expressed genes are neural stem cell markers (*Otx, Gata, meteorin)* and neuronal TFs *(Pou, Elav*), proteins driving synaptogenesis and axon guidance, potassium voltage-gated channels, neuropeptide receptors, as well as receptors for other neurotransmitters (histamine, dopamine, etc.) (Appendix 11). Interestingly, *Gcm* is co-enriched with *meteorin* in cell cluster C40, a secreted protein which in addition to being a neural progenitor marker, drives the generation of glial cells with meteor-like tails, hence the name (Nishino et al., 2004). In addition, these cell clusters contain several glutamate and GABA receptors and transporters. Therefore, *Nematostella Gcm*-expressing cells seem to combine the features of neurons, astrocytes, and radial glia (neural stem cells).

These cells probably represent "protoglia", which diversified into neurons and glia later in evolution. It is assumed that a new cell type emerges by a modification of a pre-existing cell type (Schmidt-Rhaesa et al., 2015). It is possible that the last common ancestor of *Bilateria* and *Cnidaria* possessed cells combining the characteristics of both neurons and glia, which got inherited by both animal groups and subsequently diverged. Later through the "division of labor" various glial cell types emerged in *Bilateria* **(Figure 3.23).**

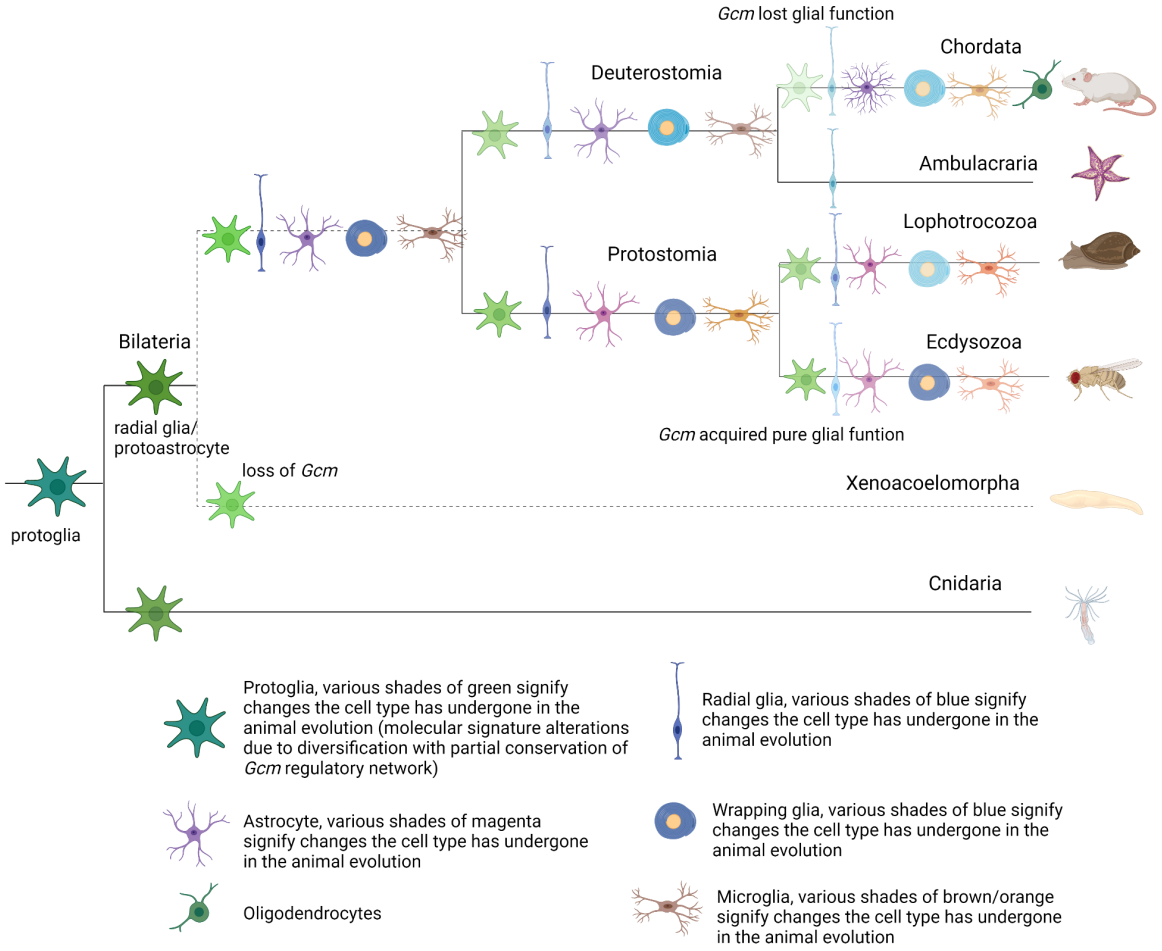


Figure 3.23. Evolutionary tree of glia based on the analysis of *Gcm* program conservation. The diversification of protoglia into various glial cell types with specialized functions is shown. Created with BioRender.com.

The protoglial ancestral cell type had a specific molecular signature that can be traced in *Nematostella* and is conserved to a various degree in different bilaterian lineages. With functional segregation the *Gcm*-controlled program either maintained its glial regulation, which is the case in *Drosophila*, or was significantly modified to have only a potential to induce glial markers as is the case in vertebrates. The program kept being modified so that it's almost impossible to recognize it in the most advanced species **(Figure 3.23)**.

This explains the absence of obvious homologous glial cell types in *Nematostella*. Instead *Gcm*-expressing cells are a subset of neurons possessing glial features. In accordance with the observation that neural modules, such as pre- and post synapses, appeared earlier than the neurons themselves (Achim & Arendt, 2014), glial features could have emerged earlier than glia as a neuronal feature. Metabolic support of neurons could have been the primary glial function that gave rise to a distinct cell type and separated glia from neurons.

An important question is: do all glial cells derive from *Gcm*-expressing protoglia? There might be another program, especially given that there are animals lacking *Gcm* but possibly possessing glia, such as acoels. The BLAST search for *Gcm* I did using the transcriptomes of *Praesagittifera naikaiensis* (OIST marine genome browser) and *Irura*

pulchra (Duruz et al., 2021) did not yield any result. Although, similarly to *Nematostella,* no clearly glial cluster homologous to known bilaterian glia was identified in single-cell transcriptome data of *I.pulchra*, the authors assume that an uncharacterized cell cluster 15 could be glial because of its enrichment with lipoprotein receptors. Future studies could unravel a *Gcm*-independent mechanism driving gliogenesis in these animals, potentially present even in cnidarians.

Conclusions

In recent years basal metazoans have been extensively studied particularly in the context of neuronal development thanks to the rapid development of the 'omics' tools. However, the glial program conservation has not been investigated in these animals. I discovered that *Cnidaria,* the reference taxon in all studies on the origins of the nervous system in *Bilateria,* possesses the highest degree of bilaterian glial program conservation. In this study the first attempt at using a cnidarian model organism to unravel the origins of glial cells was made. I show that *Gcm,* the core bilaterian glial TF, has a conserved function of driving a neuro-glial program in *Nematostella*. This suggests that glial and neuronal programs were tightly intertwined, and separated later during evolution to give rise to separate cell types. The first functions of protoglial cells might have been neurotransmitter quenching and lipid metabolism, which is consistent with the notion of glial cells as the homeostatic cells of the nervous system. Future studies should clarify the functions of other conserved bilaterian glial TFs and functional genes in *Cnidaria* and other basal metazoans.

Bibliography

Achim, K., & Arendt, D. (2014). Structural evolution of cell types by step-wise assembly of cellular modules. *Current Opinion in Genetics & Development*, *27*, 102–108.

Adams, J. C., & Tucker, R. P. (2000). The thrombospondin type 1 repeat (TSR) superfamily: diverse proteins with related roles in neuronal development. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *218*(2), 280–299.

Adamska, M., Matus, D. Q., Adamski, M., Green, K., Rokhsar, D. S., Martindale, M. Q., & Degnan, B. M. (2007). The evolutionary origin of hedgehog proteins. *Current Biology: CB*, *17*(19), R836–R837.

Ahmad, S. T., Rogers, A. D., Chen, M. J., Dixit, R., Adnani, L., Frankiw, L. S., Lawn, S. O., Blough, M. D., Alshehri, M., Wu, W., Marra, M. A., Robbins, S. M., Cairncross, J. G., Schuurmans, C., & Chan, J. A. (2019). Capicua regulates neural stem cell proliferation and lineage specification through control of Ets factors. *Nature Communications*, *10*(1), 2000.

al-Ali, S. Y., & al-Hussain, S. M. (1996). An ultrastructural study of the phagocytic activity of astrocytes in adult rat brain. *Journal of Anatomy*, *188 (Pt 2)*, 257–262.

Albagli, O., Klaes, A., Ferreira, E., Leprince, D., & Klämbt, C. (1996). Function of ets genes is conserved between vertebrates and Drosophila. *Mechanisms of Development*, *59*(1), 29–40.

Alfonso, T. B., & Jones, B. W. (2002). gcm2 promotes glial cell differentiation and is required with glial cells missing for macrophage development in Drosophila. *Developmental Biology*, *248*(2), 369–383.

Allodi, S., Bressan, C. M., Carvalho, S. L., & Cavalcante, L. A. (2006). Regionally specific distribution of the binding of anti‐glutamine synthetase and anti‐S100 antibodies and of Datura stramonium lectin in glial domains of the optic lobe of the giant prawn. *Glia*, *53*(6), 612–620.

Al-Shaer, L., Havrilak, J. A., & Layden, M. J. (2021). Nematostella vectensis as a Model System. In *Handbook of Marine Model Organisms in Experimental Biology* (pp. 107–128). unknown.

Altenhein, B., Becker, A., Busold, C., Beckmann, B., Hoheisel, J. D., & Technau, G. M. (2006). Expression profiling of glial genes during Drosophila embryogenesis. *Developmental Biology*, *296*(2), 545–560.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410.

Amiel, A. R., Johnston, H., Chock, T., Dahlin, P., Iglesias, M., Layden, M., Röttinger, E., & Martindale, M. Q. (2017). A bipolar role of the transcription factor ERG for cnidarian germ layer formation and apical domain patterning. *Developmental Biology*, *430*(2), 346–361.

Anctil, M. (2009). Chemical transmission in the sea anemone Nematostella vectensis: A genomic perspective. *Comparative Biochemistry and Physiology. Part D, Genomics & Proteomics*, *4*(4), 268–289.

Andersen, C. L., Jensen, J. L., & Ørntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research*, *64*(15), 5245–5250.

Andrews, G. L., Tanglao, S., Farmer, W. T., Morin, S., Brotman, S., Berberoglu, M. A., Price, H., Fernandez, G. C., Mastick, G. S., Charron, F., & Kidd, T. (2008). Dscam guides embryonic axons by Netrin-dependent and -independent functions. *Development* , *135*(23), 3839–3848.

Araque, A., & Navarrete, M. (2010). Glial cells in neuronal network function. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1551), 2375–2381.

Astigarraga, S., Douthit, J., Tarnogorska, D., Creamer, M. S., Mano, O., Clark, D. A., Meinertzhagen, I. A., & Treisman, J. E. (2018). Drosophila Sidekick is required in developing photoreceptors to enable visual motion detection. *Development* , *145*(3). https://doi.org/10.1242/dev.158246

Bailly, X., Laguerre, L., Correc, G., Dupont, S., Kurth, T., Pfannkuchen, A., Entzeroth, R., Probert, I., Vinogradov, S., Lechauve, C., Garet-Delmas, M.-J., Reichert, H., & Hartenstein, V. (2014). The chimerical and multifaceted marine acoel Symsagittifera roscoffensis: from photosymbiosis to brain regeneration. *Frontiers in Microbiology*, *5*, 498.

Baldwin, K. T., & Eroglu, C. (2017). Molecular mechanisms of astrocyte-induced synaptogenesis. *Current Opinion in Neurobiology*, *45*, 113–120.

Banerjee, S., Pillai, A. M., Paik, R., Li, J., & Bhat, M. A. (2006). Axonal ensheathment and septate junction formation in the peripheral nervous system of Drosophila. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(12), 3319–3329.

Bang, M. L., & Owczarek, S. (2013). A matter of balance: role of neurexin and neuroligin at the synapse. *Neurochemical Research*, *38*(6), 1174–1189.

Baumgartner, S., Littleton, J. T., Broadie, K., Bhat, M. A., Harbecke, R., Lengyel, J. A., Chiquet-Ehrismann, R., Prokop, A., & Bellen, H. J. (1996). A Drosophila neurexin is required for septate junction and blood-nerve barrier formation and function. *Cell*, *87*(6), 1059–1068.

Beckers, P., Helm, C., & Bartolomaeus, T. (2019). The anatomy and development of the nervous system in Magelonidae (Annelida) – insights into the evolution of the annelid brain. *BMC Evolutionary Biology*, *19*. https://bmcecolevol.biomedcentral.com/articles/10.1186/s12862-019-1498-9

Bennett, M. L., Bennett, F. C., Liddelow, S. A., Ajami, B., Zamanian, J. L., Fernhoff, N. B., Mulinyawe, S. B., Bohlen, C. J., Adil, A., Tucker, A., Weissman, I. L., Chang, E. F., Li, G., Grant, G. A., Hayden Gephart, M. G., & Barres, B. A. (2016). New tools for studying microglia in the mouse and human CNS. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(12), E1738–E1746.

Bergersen, L. H., Morland, C., Ormel, L., Rinholm, J. E., Larsson, M., Wold, J. F. H., Røe, A. T., Stranna, A., Santello, M., Bouvier, D., Ottersen, O. P., Volterra, A., & Gundersen, V. (2012). Immunogold detection of L-glutamate and D-serine in small synaptic-like microvesicles in adult hippocampal astrocytes. *Cerebral Cortex* , *22*(7), 1690–1697.

Bergles, D. E., & Richardson, W. D. (2015). Oligodendrocyte Development and Plasticity. *Cold Spring Harbor Perspectives in Biology*, *8*(2), a020453.

Bery, A., Cardona, A., Martinez, P., & Hartenstein, V. (2010). Structure of the central nervous system of a juvenile acoel, Symsagittifera roscoffensis. *Development Genes and Evolution*, *220*(3-4), 61–76.

Bezzi, P., & Volterra, A. (2001). A neuron–glia signaling network in the active brain. *Current*

Opinion in Neurobiology, *11*(3), 387–394.
Bianchi-Frias, D., Orian, A., Delrow, J. J., Vazquez, J., Rosales-Nieves, A. E., & Parkhurst, S. M. (2004). Hairy transcriptional repression targets and cofactor recruitment in Drosophila. *PLoS Biology*, *2*(7), E178.
Biserova, N. M., Gordeev, I. I., Korneva, J. V., & Salnikova, M. M. (2010). Structure of the glial cells in the nervous system of parasitic and free-living flatworms. *Biology Bulletin*, *37*(3), 277–287.
Boddum, K., Jensen, T. P., Magloire, V., Kristiansen, U., Rusakov, D. A., Pavlov, I., & Walker, M. C. (2016). Astrocytic GABA transporter activity modulates excitatory neurotransmission. *Nature Communications*, *7*, 13572.
Bonni, A., Sun, Y., Nadal-Vicens, M., Bhatt, A., Frank, D. A., Rozovsky, I., Stahl, N., Yancopoulos, G. D., & Greenberg, M. E. (1997). Regulation of gliogenesis in the central nervous system by the JAK-STAT signaling pathway. *Science*, *278*(5337), 477–483.
Bosson, A., Paumier, A., Boisseau, S., Jacquier-Sarlin, M., Buisson, A., & Albrieux, M. (2017). TRPA1 channels promote astrocytic Ca2+ hyperactivity and synaptic dysfunction mediated by oligomeric forms of amyloid-β peptide. *Molecular Neurodegeneration*, *12*(1), 53.
Bozzo, M., Lacalli, T. C., Obino, V., Caicci, F., Marcenaro, E., Bachetti, T., Manni, L., Pestarino, M., Schubert, M., & Candiani, S. (2021). Amphioxus neuroglia: Molecular characterization and evidence for early compartmentalization of the developing nerve cord. *Glia*. https://doi.org/10.1002/glia.23982
Bramanti, V., Tomassoni, D., Avitabile, M., Amenta, F., & Avola, R. (2010). Biomarkers of glial cell proliferation and differentiation in culture. *Frontiers in Bioscience* , *2*, 558–570.
Brown, A. M., & Ransom, B. R. (2007). Astrocyte glycogen and brain energy metabolism. *Glia*, *55*(12), 1263–1271.
Busengdal, H., & Rentzsch, F. (2017). Unipotent progenitors contribute to the generation of sensory cell types in the nervous system of the cnidarian Nematostella vectensis. *Developmental Biology*, *431*(1), 59–68.
Cahoy, J. D., Emery, B., Kaushal, A., Foo, L. C., Zamanian, J. L., Christopherson, K. S., Xing, Y., Lubischer, J. L., Krieg, P. A., Krupenko, S. A., Thompson, W. J., & Barres, B. A. (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(1), 264–278.
Campbell, C. E., Piper, M., Plachez, C., Yeh, Y.-T., Baizer, J. S., Osinski, J. M., Litwack, E. D., Richards, L. J., & Gronostajski, R. M. (2008). The transcription factor Nfix is essential for normal brain development. *BMC Developmental Biology*, *8*, 52.
Campbell, K., & Götz, M. (2002). Radial glia: multi-purpose cells for vertebrate brain development. *Trends in Neurosciences*, *25*(5), 235–238.
Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S. N., Steemers, F. J., Adey, A., Waterston, R. H., Trapnell, C., & Shendure, J. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science*, *357*(6352), 661–667.
Cardone, B., & Roots, B. I. (1990). Comparative immunohistochemical study of glial filament proteins (glial fibrillary acidic protein and vimentin) in goldfish, octopus, and snail. *Glia*, *3*(3), 180–192.
Chari, D. (2002). Glial Cell Development, 2nd Edition. *Neuropathology and Applied*

Neurobiology, *28*(1), 85–86.
Chatterjee, M., Schild, D., & Teunissen, C. E. (2019). Contactins in the central nervous system: role in health and disease. *Neural Regeneration Research*, *14*(2), 206–216.
Chen, C.-Y., McKinney, S. A., Ellington, L. R., & Gibson, M. C. (2020). Hedgehog signaling is required for endomesodermal patterning and germ cell development in the sea anemone Nematostella vectensis. *eLife*, *9*. https://doi.org/10.7554/eLife.54573
Chen, Q., Leshkowitz, D., Blechman, J., & Levkowitz, G. (2020). Single-Cell Molecular and Cellular Architecture of the Mouse Neurohypophysis. *eNeuro*, *7*(1). https://doi.org/10.1523/ENEURO.0345-19.2019
Chotard, C., Leung, W., & Salecker, I. (2005). glial cells missing and gcm2 cell autonomously regulate both glial and neuronal development in the visual system of Drosophila. *Neuron*, *48*(2), 237–251.
Christopherson, K. S., Ullian, E. M., Stokes, C. C. A., Mullowney, C. E., Hell, J. W., Agah, A., Lawler, J., Mosher, D. F., Bornstein, P., & Barres, B. A. (2005). Thrombospondins are astrocyte-secreted proteins that promote CNS synaptogenesis. *Cell*, *120*(3), 421–433.
Collins, A. G. (2002). Phylogeny of Medusozoa and the evolution of cnidarian life cycles. *Journal of Evolutionary Biology*, *15*(3), 418–432.
Croset, V., Treiber, C. D., & Waddell, S. (2018). Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics. *eLife*, *7*. https://doi.org/10.7554/eLife.34550
Crossin, K. L., & Krushel, L. A. (2000). Cellular signaling by neural cell adhesion molecules of the immunoglobulin superfamily. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *218*(2), 260–279.
Dai, J., Bercury, K. K., Ahrendsen, J. T., & Macklin, W. B. (2015). Olig1 function is required for oligodendrocyte differentiation in the mouse brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(10), 4386–4402.
Danbolt, N. C., Furness, D. N., & Zhou, Y. (2016). Neuronal vs glial glutamate uptake: Resolving the conundrum. *Neurochemistry International*, *98*, 29–45.
D'Arcy, B. R., & Silver, D. L. (2020). Local gene regulation in radial glia: Lessons from across the nervous system. *Traffic* . https://doi.org/10.1111/tra.12769
Degnan, B. M., Degnan, S. M., Naganuma, T., & Morse, D. E. (1993). The ets multigene family is conserved throughout the Metazoa. *Nucleic Acids Research*, *21*(15), 3479–3484.
de la Pompa, J. L., Wakeham, A., Correia, K. M., Samper, E., Brown, S., Aguilera, R. J., Nakano, T., Honjo, T., Mak, T. W., Rossant, J., & Conlon, R. A. (1997). Conservation of the Notch signaling pathway in mammalian neurogenesis. *Development* , *124*(6), 1139–1148.
Deneen, B., Ho, R., Lukaszewicz, A., Hochstim, C. J., Gronostajski, R. M., & Anderson, D. J. (2006). The transcription factor NFIA controls the onset of gliogenesis in the developing spinal cord. *Neuron*, *52*(6), 953–968.
Dimou, L., & Götz, M. (2014). Glial cells as progenitors and stem cells: new roles in the healthy and diseased brain. *Physiological Reviews*, *94*(3), 709–737.
Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J.-M., Taly, J.-F., & Notredame, C. (2011). T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Research*, *39*(Web Server issue), W13–W17.
Dos Santos, P. C., Gehlen, G., Faccioni-Heuser, M. C., & Achaval, M. (2005). Detection of glial fibrillary acidic protein (GFAP) and vimentin (Vim) by immunoelectron microscopy

of the glial cells in the central nervous system of the snail Megalobulimus abbreviatus: GFAP and Vim in glial cells of Megalobulimus. *Acta Zoologica* , *86*(2), 135–144.

Dowell, J. A., Johnson, J. A., & Li, L. (2009). Identification of astrocyte secreted proteins with a combination of shotgun proteomics and bioinformatics. *Journal of Proteome Research*, *8*(8), 4135–4143.

Driever, W. (2018). Developmental Biology: Reissner's Fiber and Straightening of the Body Axis. *Current Biology: CB*, *28*(15), R833–R835.

DuBuc, T. Q., Traylor-Knowles, N., & Martindale, M. Q. (2014). Initiating a regenerative response; cellular and molecular features of wound healing in the cnidarian Nematostella vectensis. *BMC Biology*, *12*, 24.

Duruz, J., Kaltenrieder, C., Ladurner, P., Bruggmann, R., Martìnez, P., & Sprecher, S. G. (2021). Acoel Single-Cell Transcriptomics: Cell Type Analysis of a Deep Branching Bilaterian. *Molecular Biology and Evolution*, *38*(5), 1888–1904.

Elfeber, K., Köhler, A., Lutzenburg, M., Osswald, C., Galla, H.-J., Witte, O. W., & Koepsell, H. (2004). Localization of the Na+-D-glucose cotransporter SGLT1 in the blood-brain barrier. *Histochemistry and Cell Biology*, *121*(3), 201–207.

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* , *32*(19), 3047–3048.

Fischer, A., & Gessler, M. (2007). Delta-Notch--and then? Protein interactions and proposed modes of repression by Hes and Hey bHLH factors. *Nucleic Acids Research*, *35*(14), 4583–4596.

Fletcher, C. F., Jenkins, N. A., Copeland, N. G., Chaudhry, A. Z., & Gronostajski, R. M. (1999). Exon structure of the nuclear factor I DNA-binding domain from C. elegans to mammals. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, *10*(4), 390–396.

Fortunato, S., Adamski, M., Bergum, B., Guder, C., Jordal, S., Leininger, S., Zwafink, C., Rapp, H. T., & Adamska, M. (2012). Genome-wide analysis of the sox family in the calcareous sponge Sycon ciliatum: multiple genes with unique expression patterns. *EvoDevo*, *3*(1), 14.

Fortunato, S. A. V., Vervoort, M., Adamski, M., & Adamska, M. (2016). Conservation and divergence of bHLH genes in the calcisponge Sycon ciliatum. *EvoDevo*, *7*, 23.

Freeman, M. R., Delrow, J., Kim, J., Johnson, E., & Doe, C. Q. (2003). Unwrapping glial biology: Gcm target genes regulating glial development, diversification, and function. *Neuron*, *38*(4), 567–580.

Freeman, M. R., & Doherty, J. (2006). Glial cell biology in Drosophila and vertebrates. *Trends in Neurosciences*, *29*(2), 82–90.

Freeman, M. R., & Rowitch, D. H. (2013). Evolving concepts of gliogenesis: a look way back and ahead to the next 25 years. *Neuron*, *80*(3), 613–623.

Frei, J. A., & Stoeckli, E. T. (2014). SynCAMs extend their functions beyond the synapse. *The European Journal of Neuroscience*, *39*(11), 1752–1760.

Fu, H., Cai, J., Clevers, H., Fast, E., Gray, S., Greenberg, R., Jain, M. K., Ma, Q., Qiu, M., Rowitch, D. H., Taylor, C. M., & Stiles, C. D. (2009). A genome-wide screen for spatially restricted expression patterns identifies transcription factors that regulate glial development. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(36), 11399–11408.

Fullard, J. F., & Baker, N. E. (2015). Signaling by the engulfment receptor draper: a screen in Drosophila melanogaster implicates cytoskeletal regulators, Jun N-terminal Kinase, and Yorkie. *Genetics*, *199*(1), 117–134.

Gadea, A., & López-Colomé, A. M. (2001). Glial transporters for glutamate, glycine, and GABA: II. GABA transporters. *Journal of Neuroscience Research*, *63*(6), 461–468.

Gaiano, N., & Fishell, G. (2002). The role of notch in promoting glial and neural stem cell fates. *Annual Review of Neuroscience*, *25*, 471–490.

Galliot, B., Quiquand, M., Ghila, L., de Rosa, R., Miljkovic-Licina, M., & Chera, S. (2009). Origins of neurogenesis, a cnidarian view. *Developmental Biology*, *332*(1), 2–24.

Ganot, P., Zoccola, D., Tambutté, E., Voolstra, C. R., Aranda, M., Allemand, D., & Tambutté, S. (2015). Structural molecular components of septate junctions in cnidarians point to the origin of epithelial junctions in eukaryotes. *Molecular Biology and Evolution*, *32*(1), 44–62.

Gasca, J., Flores, M. L., Jiménez-Guerrero, R., Sáez, M. E., Barragán, I., Ruíz-Borrego, M., Tortolero, M., Romero, F., Sáez, C., & Japón, M. A. (2020). EDIL3 promotes epithelial-mesenchymal transition and paclitaxel resistance through its interaction with integrin αVβ3 in cancer cells. *Cell Death Discovery*, *6*, 86.

Gazave, E., Guillou, A., & Balavoine, G. (2014). History of a prolific family: the Hes/Hey-related genes of the annelid Platynereis. *EvoDevo*, *5*, 29.

Gazave, E., Lapébie, P., Richards, G. S., Brunet, F., Ereskovsky, A. V., Degnan, B. M., Borchiellini, C., Vervoort, M., & Renard, E. (2009). Origin and evolution of the Notch signaling pathway: an overview from eukaryotic genomes. *BMC Evolutionary Biology*, *9*, 249.

Gebara, E., Bonaguidi, M. A., Beckervordersandforth, R., Sultan, S., Udry, F., Gijs, P.-J., Lie, D. C., Ming, G.-L., Song, H., & Toni, N. (2016). Heterogeneity of Radial Glia-Like Cells in the Adult Hippocampus. *Stem Cells*, *34*(4), 997–1010.

Geirsdottir, L., David, E., Keren-Shaul, H., Weiner, A., Bohlen, S. C., Neuber, J., Balic, A., Giladi, A., Sheban, F., Dutertre, C.-A., Pfeifle, C., Peri, F., Raffo-Romero, A., Vizioli, J., Matiasek, K., Scheiwe, C., Meckel, S., Mätz-Rensing, K., van der Meer, F., … Prinz, M. (2019). Cross-Species Single-Cell Analysis Reveals Divergence of the Primate Microglia Program. *Cell*, *179*(7), 1609–1622.e16.

Genikhovich, G., & Technau, U. (2009). The starlet sea anemone Nematostella vectensis: an anthozoan model organism for studies in comparative genomics and functional evolutionary developmental biology. *Cold Spring Harbor Protocols*, *2009*(9), db.emo129.

Ge, W., Martinowich, K., Wu, X., He, F., Miyamoto, A., Fan, G., Weinmaster, G., & Sun, Y. E. (2002). Notch signaling promotes astrogliogenesis via direct CSL-mediated glial gene activation. *Journal of Neuroscience Research*, *69*(6), 848–860.

Ghazale, H., Ripoll, C., Leventoux, N., Jacob, L., Azar, S., Mamaeva, D., Glasson, Y., Calvo, C.-F., Thomas, J.-L., Meneceur, S., Lallemand, Y., Rigau, V., Perrin, F. E., Noristani, H. N., Rocamonde, B., Huillard, E., Bauchet, L., & Hugnot, J.-P. (2019). RNA Profiling of the Human and Mouse Spinal Cord Stem Cell Niches Reveals an Embryonic-like Regionalization with MSX1+ Roof-Plate-Derived Cells. *Stem Cell Reports*, *12*(5), 1159–1177.

Ghirardini, E., Wadle, S. L., Augustin, V., Becker, J., Brill, S., Hammerich, J., Seifert, G., & Stephan, J. (2018). Expression of functional inhibitory neurotransmitter transporters GlyT1, GAT-1, and GAT-3 by astrocytes of inferior colliculus and hippocampus.

Molecular Brain, *11*(1), 4.
Giamanco, K. A., & Matthews, R. T. (2012). Deconstructing the perineuronal net: cellular contributions and molecular composition of the neuronal extracellular matrix. *Neuroscience*, *218*, 367–384.
Ginhoux, F., & Prinz, M. (2015). Origin of microglia: current concepts and past controversies. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a020537.
Glasgow, S. M., Zhu, W., Stolt, C. C., Huang, T.-W., Chen, F., LoTurco, J. J., Neul, J. L., Wegner, M., Mohila, C., & Deneen, B. (2014). Mutual antagonism between Sox10 and NFIA regulates diversification of glial lineages and glioma subtypes. *Nature Neuroscience*, *17*(10), 1322–1329.
Gomes, W. A., Mehler, M. F., & Kessler, J. A. (2003). Transgenic overexpression of BMP4 increases astroglial and decreases oligodendroglial lineage commitment. *Developmental Biology*, *255*(1), 164–177.
Grear, K. E., Ling, I.-F., Simpson, J. F., Furman, J. L., Simmons, C. R., Peterson, S. L., Schmitt, F. A., Markesbery, W. R., Liu, Q., Crook, J. E., Younkin, S. G., Bu, G., & Estus, S. (2009). Expression of SORL1 and a novel SORL1 splice variant in normal and Alzheimers disease brain. *Molecular Neurodegeneration*, *4*, 46.
Gronostajski, R. M. (2000). Roles of the NFI/CTF gene family in transcription and development. *Gene*, *249*(1-2), 31–45.
Gross, R. E., Mehler, M. F., Mabie, P. C., Zang, Z., Santschi, L., & Kessler, J. A. (1996). Bone morphogenetic proteins promote astroglial lineage commitment by mammalian subventricular zone progenitor cells. *Neuron*, *17*(4), 595–606.
Gross-Thebing, T., Paksa, A., & Raz, E. (2014). Simultaneous high-resolution detection of multiple transcripts combined with localization of proteins in whole-mount embryos. *BMC Biology*, *12*, 55.
Guo, Y., Xiao, P., Lei, S., Deng, F., Xiao, G. G., Liu, Y., Chen, X., Li, L., Wu, S., Chen, Y., Jiang, H., Tan, L., Xie, J., Zhu, X., Liang, S., & Deng, H. (2008). How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta Biochimica et Biophysica Sinica*, *40*(5), 426–436.
Gyoja, F. (2014). A genome-wide survey of bHLH transcription factors in the Placozoan Trichoplax adhaerens reveals the ancient repertoire of this gene family in metazoan. *Gene*, *542*(1), 29–37.
Gyoja, F. (2017). Basic helix-loop-helix transcription factors in evolution: Roles in development of mesoderm and neural tissues. *Genesis* , *55*(9). https://doi.org/10.1002/dvg.23051
Hagedorn, L., Paratore, C., Brugnoli, G., Baert, J. L., Mercader, N., Suter, U., & Sommer, L. (2000). The Ets domain transcription factor Erm distinguishes rat satellite glia from Schwann cells and is regulated in satellite cells by neuregulin signaling. *Developmental Biology*, *219*(1), 44–58.
Hai, M., Muja, N., DeVries, G. H., Quarles, R. H., & Patel, P. I. (2002). Comparative analysis of Schwann cell lines as model systems for myelin gene transcription studies. *Journal of Neuroscience Research*, *69*(4), 497–508.
Hainfellner, J. A., Voigtländer, T., Ströbel, T., Mazal, P. R., Maddalena, A. S., Aguzzi, A., & Budka, H. (2001). Fibroblasts can express glial fibrillary acidic protein (GFAP) in vivo. *Journal of Neuropathology and Experimental Neurology*, *60*(5), 449–461.
Halter, D. A., Urban, J., Rickert, C., Ner, S. S., Ito, K., Travers, A. A., & Technau, G. M.

(1995). The homeobox gene repo is required for the differentiation and maintenance of glia function in the embryonic nervous system of Drosophila melanogaster. *Development* , *121*(2), 317–332.

Hanaoka, R., Ohmori, Y., Uyemura, K., Hosoya, T., Hotta, Y., Shirao, T., & Okamoto, H. (2004). Zebrafish gcmb is required for pharyngeal cartilage formation. *Mechanisms of Development*, *121*(10), 1235–1247.

Harris, L., Genovesi, L. A., Gronostajski, R. M., Wainwright, B. J., & Piper, M. (2015). Nuclear factor one transcription factors: Divergent functions in developmental versus adult stem cell populations. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *244*(3), 227–238.

Hartenstein, V., & Giangrande, A. (2018). Connecting the nervous and the immune systems in evolution. *Communications Biology*, *1*(1), 64.

Hartline, D. K. (2011). The evolutionary origins of glia. *Glia*, *59*(9), 1215–1236.

Hayakawa, E., Watanabe, H., Menschaert, G., Holstein, T. W., Baggerman, G., & Schoofs, L. (2019). A combined strategy of neuropeptide prediction and tandem mass spectrometry identifies evolutionarily conserved ancient neuropeptides in the sea anemone Nematostella vectensis. *PloS One*, *14*(9), e0215185.

Hayashi, M. K., & Yasui, M. (2015). The transmembrane transporter domain of glutamate transporters is a process tip localizer. *Scientific Reports*, *5*, 9032.

He, F., Ge, W., Martinowich, K., Becker-Catania, S., Coskun, V., Zhu, W., Wu, H., Castro, D., Guillemot, F., Fan, G., de Vellis, J., & Sun, Y. E. (2005). A positive autoregulatory loop of Jak-STAT signaling controls the onset of astrogliogenesis. *Nature Neuroscience*, *8*(5), 616–625.

Helm Conrad, Karl Anett, Beckers Patrick, Kaul-Strehlow Sabrina, Ulbricht Elke, Kourtesis Ioannis, Kuhrt Heidrun, Hausen Harald, Bartolomaeus Thomas, Reichenbach Andreas, & Bleidorn Christoph. (2017). Early evolution of radial glial cells in Bilateria. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1859), 20170743.

He, X., Wu, F., Zhang, L., Li, L., & Zhang, G. (2021). Comparative and evolutionary analyses reveal conservation and divergence of the notch pathway in lophotrochozoa. *Scientific Reports*, *11*(1), 11378.

Hitoshi, S., Ishino, Y., Kumar, A., Jasmine, S., Tanaka, K. F., Kondo, T., Kato, S., Hosoya, T., Hotta, Y., & Ikenaka, K. (2011). Mammalian Gcm genes induce Hes5 expression by active DNA demethylation and induce neural stem cells. *Nature Neuroscience*, *14*(8), 957–964.

Hojo, M., Ohtsuka, T., Hashimoto, N., Gradwohl, G., Guillemot, F., & Kageyama, R. (2000). Glial cell fate specification modulated by the bHLH gene Hes5 in mouse retina. *Development* , *127*(12), 2515–2522.

Holcroft, C. E., Jackson, W. D., Lin, W.-H., Bassiri, K., Baines, R. A., & Phelan, P. (2013). Innexins Ogre and Inx2 are required in glial cells for normal postembryonic development of the Drosophila central nervous system. *Journal of Cell Science*, *126*(Pt 17), 3823–3834.

Holtman, I. R., Skola, D., & Glass, C. K. (2017). Transcriptional control of microglia phenotypes in health and disease. *The Journal of Clinical Investigation*, *127*(9), 3220–3229.

Horridge, G. A., & Mackay, B. (1962). Naked Axons and Symmetrical Synapses in Coelenterates. *Journal of Cell Science*, *s3-103*(64), 531–541.

Hosoya, T., Takizawa, K., Nitta, K., & Hotta, Y. (1995). glial cells missing: a binary switch between neuronal and glial determination in Drosophila. *Cell*, *82*(6), 1025–1036.

Hubbard, J. A., Szu, J. I., & Binder, D. K. (2018). The role of aquaporin-4 in synaptic plasticity, memory and disease. *Brain Research Bulletin*, *136*, 118–129.

Huggett, J., Dheda, K., Bustin, S., & Zumla, A. (2005). Real-time RT-PCR normalisation; strategies and considerations. *Genes and Immunity*, *6*(4), 279–284.

Hu, Y., Jin, S., Fu, H., Qiao, H., Zhang, W., Jiang, S., Gong, Y., Xiong, Y., & Wu, Y. (2020). Functional analysis of a SoxE gene in the oriental freshwater prawn, Macrobrachium nipponense by molecular cloning, expression pattern analysis, and in situ hybridization (de Haan, 1849). *3 Biotech*, *10*(1), 10.

Hyzinski-García, M. C., Rudkouskaya, A., & Mongin, A. A. (2014). LRRC8A protein is indispensable for swelling-activated and ATP-induced release of excitatory amino acids in rat astrocytes. *The Journal of Physiology*, *592*(22), 4855–4862.

Ingham, P. W., Nakano, Y., & Seger, C. (2011). Mechanisms and functions of Hedgehog signaling across the metazoa. *Nature Reviews. Genetics*, *12*(6), 393–406.

Ioannou, M. S., Jackson, J., Sheu, S.-H., Chang, C.-L., Weigel, A. V., Liu, H., Pasolli, H. A., Xu, C. S., Pang, S., Matthies, D., Hess, H. F., Lippincott-Schwartz, J., & Liu, Z. (2019). Neuron-Astrocyte Metabolic Coupling Protects against Activity-Induced Fatty Acid Toxicity. *Cell*, *177*(6), 1522–1535.e14.

Ishibashi, M., Egawa, K., & Fukuda, A. (2019). Diverse Actions of Astrocytes in GABAergic Signaling. *International Journal of Molecular Sciences*, *20*(12). https://doi.org/10.3390/ijms20122964

Iwasaki, Y., Hosoya, T., Takebayashi, H., Ogawa, Y., Hotta, Y., & Ikenaka, K. (2003). The potential to induce glial differentiation is conserved between Drosophila and mammalian glial cells missing genes. *Development* , *130*(24), 6027–6035.

Jager, M., Chiori, R., Alié, A., Dayraud, C., Quéinnec, E., & Manuel, M. (2011). New insights on ctenophore neural anatomy: immunofluorescence study in Pleurobrachia pileus (Müller, 1776). *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, *316B*(3), 171–187.

Jager, M., Quéinnec, E., Houliston, E., & Manuel, M. (2006). Expansion of the SOX gene family predated the emergence of the Bilateria. *Molecular Phylogenetics and Evolution*, *39*(2), 468–477.

Jin, B.-J., Zhang, H., Binder, D. K., & Verkman, A. S. (2013). Aquaporin-4-dependent K(+) and water transport modeled in brain extracellular space following neuroexcitation. *The Journal of General Physiology*, *141*(1), 119–132.

Jin, Z., Chen, J., Huang, H., Wang, J., Lv, J., Yu, M., Guo, X., Zhang, Y., Cai, T., & Xi, R. (2020). The Drosophila Ortholog of Mammalian Transcription Factor Sox9 Regulates Intestinal Homeostasis and Regeneration at an Appropriate Level. *Cell Reports*, *31*(8), 107683.

Joosten, E. A., & Gribnau, A. A. (1989). Astrocytes and guidance of outgrowing corticospinal tract axons in the rat. An immunocytochemical study using anti-vimentin and anti-glial fibrillary acidic protein. *Neuroscience*, *31*(2), 439–452.

Junkunlo, K., Söderhäll, K., & Söderhäll, I. (2020). A transcription factor glial cell missing (Gcm) in the freshwater crayfish Pacifastacus leniusculus. *Developmental and Comparative Immunology*, *113*, 103782.

Jurga, A. M., Paleczna, M., & Kuter, K. Z. (2020). Overview of General and Discriminating

Markers of Differential Microglia Phenotypes. *Frontiers in Cellular Neuroscience*, *14*, 198.

Jurisch-Yaksi, N., Yaksi, E., & Kizil, C. (2020). Radial glia in the zebrafish brain: Functional, structural, and physiological comparison with the mammalian glia. *Glia*, *68*(12), 2451–2470.

Kageyama, R., & Ohtsuka, T. (1999). The Notch-Hes pathway in mammalian neural development. *Cell Research*, *9*(3), 179–188.

Kageyama, R., Ohtsuka, T., & Kobayashi, T. (2007). The Hes gene family: repressors and oscillators that orchestrate embryogenesis. *Development* , *134*(7), 1243–1251.

Kageyama, R., Ohtsuka, T., & Kobayashi, T. (2008). Roles of Hes genes in neural development. *Development, Growth & Differentiation*, *50 Suppl 1*, S97–S103.

Kamakura, S., Oishi, K., Yoshimatsu, T., Nakafuku, M., Masuyama, N., & Gotoh, Y. (2004). Hes binding to STAT3 mediates crosstalk between Notch and JAK–STAT signaling. *Nature Cell Biology*, *6*(6), 547–554.

Kammerer, M., & Giangrande, A. (2001). Glide2, a second glial promoting factor in Drosophila melanogaster. *The EMBO Journal*, *20*(17), 4664–4673.

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Kang, P., Lee, H. K., Glasgow, S. M., Finley, M., Donti, T., Gaber, Z. B., Graham, B. H., Foster, A. E., Novitch, B. G., Gronostajski, R. M., & Deneen, B. (2012). Sox9 and NFIA coordinate a transcriptional regulatory cascade during the initiation of gliogenesis. *Neuron*, *74*(1), 79–94.

Kapli, P., & Telford, M. J. (2020). Topology-dependent asymmetry in systematic errors affects phylogenetic placement of Ctenophora and Xenacoelomorpha. *Science Advances*, *6*(50). https://doi.org/10.1126/sciadv.abc5162

Karabulut, A., He, S., Chen, C.-Y., McKinney, S. A., & Gibson, M. C. (2019). Electroporation of short hairpin RNAs for rapid and efficient gene knockdown in the starlet sea anemone, Nematostella vectensis. *Developmental Biology*. https://doi.org/10.1016/j.ydbio.2019.01.005

Káradóttir, R., Hamilton, N. B., Bakiri, Y., & Attwell, D. (2008). Spiking and nonspiking classes of oligodendrocyte precursor glia in CNS white matter. *Nature Neuroscience*, *11*(4), 450–456.

Kayal, E., Roure, B., Philippe, H., Collins, A. G., & Lavrov, D. V. (2013). Cnidarian phylogenetic relationships as revealed by mitogenomics. *BMC Evolutionary Biology*, *13*, 5.

Kepe, V., Scafoglio, C., Liu, J., Yong, W. H., Bergsneider, M., Huang, S.-C., Barrio, J. R., & Wright, E. M. (2018). Positron emission tomography of sodium glucose cotransport activity in high grade astrocytomas. *Journal of Neuro-Oncology*, *138*(3), 557–569.

Kettenmann, H., Hanisch, U.-K., Noda, M., & Verkhratsky, A. (2011). Physiology of microglia. *Physiological Reviews*, *91*(2), 461–553.

Khalturin, K., Shinzato, C., Khalturina, M., Hamada, M., Fujie, M., Koyanagi, R., Kanda, M., Goto, H., Anton-Erxleben, F., Toyokawa, M., Toshino, S., & Satoh, N. (2019). Medusozoan genomes inform the evolution of the jellyfish body plan. *Nature Ecology & Evolution*, *3*(5), 811–822.

Kimelberg, H. K., & Nedergaard, M. (2010). Functions of astrocytes and their potential as therapeutic targets. *Neurotherapeutics: The Journal of the American Society for*

Experimental NeuroTherapeutics, *7*(4), 338–353.

Kim, J., Jones, B. W., Zock, C., Chen, Z., Wang, H., Goodman, C. S., & Anderson, D. J. (1998). Isolation and characterization of mammalian homologs of the Drosophila gene glial cells missing. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(21), 12364–12369.

Kim, S., Maynard, J. C., Sasaki, Y., Strickland, A., Sherman, D. L., Brophy, P. J., Burlingame, A. L., & Milbrandt, J. (2016). Schwann Cell O-GlcNAc Glycosylation Is Required for Myelin Maintenance and Axon Integrity. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(37), 9633–9646.

Kiyota, T., Kato, A., & Kato, Y. (2007). Ets-1 regulates radial glia formation during vertebrate embryogenesis. *Organogenesis*, *3*(2), 93–101.

Klämbt, C. (1993). The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* , *117*(1), 163–176.

Konishi, H., Okamoto, T., Hara, Y., Komine, O., Tamada, H., Maeda, M., Osako, F., Kobayashi, M., Nishiyama, A., Kataoka, Y., Takai, T., Udagawa, N., Jung, S., Ozato, K., Tamura, T., Tsuda, M., Yamanaka, K., Ogi, T., Sato, K., & Kiyama, H. (2020). Astrocytic phagocytosis is a compensatory mechanism for microglial dysfunction. *The EMBO Journal*, *39*(22), e104464.

Kuchta, K., Towpik, J., Biernacka, A., Kutner, J., Kudlicki, A., Ginalski, K., & Rowicka, M. (2018). Predicting proteome dynamics using gene expression data. *Scientific Reports*, *8*(1), 13866.

Labouesse, M., Hartwieg, E., & Horvitz, H. R. (1996). The Caenorhabditis elegans LIN-26 protein is required to specify and/or maintain all non-neuronal ectodermal cell fates. *Development* , *122*(9), 2579–2588.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359.

Larroux, C., Fahey, B., Liubicich, D., Hinman, V. F., Gauthier, M., Gongora, M., Green, K., Wörheide, G., Leys, S. P., & Degnan, B. M. (2006). Developmental expression of transcription factor genes in a demosponge: insights into the origin of metazoan multicellularity. *Evolution & Development*, *8*(2), 150–173.

Layden, M. J., Johnston, H., Amiel, A. R., Havrilak, J., Steinworth, B., Chock, T., Röttinger, E., & Martindale, M. Q. (2016). MAPK signaling is necessary for neurogenesis in Nematostella vectensis. *BMC Biology*, *14*, 61.

Layden, M. J., & Martindale, M. Q. (2014). Non-canonical Notch signaling represents an ancestral mechanism to regulate neural differentiation. *EvoDevo*, *5*, 30.

Layden, M. J., Rentzsch, F., & Röttinger, E. (2016). The rise of the starlet sea anemone Nematostella vectensis as a model system to investigate development and regeneration. *Wiley Interdisciplinary Reviews. Developmental Biology*, *5*(4), 408–428.

Lazakovitch, E., Kalb, J. M., Matsumoto, R., Hirono, K., Kohara, Y., & Gronostajski, R. M. (2005). nfi-I affects behavior and life-span in C. elegans but is not essential for DNA replication or survival. *BMC Developmental Biology*, *5*, 24.

Lee, B. P., & Jones, B. W. (2005). Transcriptional regulation of the Drosophila glial gene repo. *Mechanisms of Development*, *122*(6), 849–862.

Lee, H. S., Han, J., Lee, S.-H., Park, J. A., & Kim, K.-W. (2010). Meteorin promotes the formation of GFAP-positive glia via activation of the Jak-STAT3 pathway. *Journal of Cell Science*, *123*(Pt 11), 1959–1968.

Levenson, J., Endo, S., Kategaya, L. S., Fernandez, R. I., Brabham, D. G., Chin, J., Byrne, J. H., & Eskin, A. (2000). Long-term regulation of neuronal high-affinity glutamate and glutamine uptake in Aplysia. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(23), 12858–12863.

Liang, J., Takeuchi, H., Doi, Y., Kawanokuchi, J., Sonobe, Y., Jin, S., Yawata, I., Li, H., Yasuoka, S., Mizuno, T., & Suzumura, A. (2008). Excitatory amino acid transporter expression by astrocytes is neuroprotective against microglial excitotoxicity. *Brain Research*, *1210*, 11–19.

Liebeskind, B. J., Hofmann, H. A., Hillis, D. M., & Zakon, H. H. (2017). Evolution of Animal Neural Systems. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 377–398.

Li, H., & Grumet, M. (2007). BMP and LIF signaling coordinately regulate lineage restriction of radial glia in the developing forebrain. *Glia*, *55*(1), 24–35.

Limmer, S., Weiler, A., Volkenhoff, A., Babatz, F., & Klämbt, C. (2014). The Drosophila blood-brain barrier: development and function of a glial endothelium. *Frontiers in Neuroscience*, *8*, 365.

Linser, P. J., Trapido-Rosenthal, H. G., & Orona, E. (1997). Glutamine synthetase is a glial-specific marker in the olfactory regions of the lobster (Panulirus argus) nervous system. *Glia*, *20*(4), 275–283.

Liongue, C., Taznin, T., & Ward, A. C. (2016). Signaling via the CytoR/JAK/STAT/SOCS pathway: Emergence during evolution. *Molecular Immunology*, *71*, 166–175.

Liongue, C., & Ward, A. C. (2013). Evolution of the JAK-STAT pathway. *JAK-STAT*, *2*(1), e22756.

Liu, Z., Hu, X., Cai, J., Liu, B., Peng, X., Wegner, M., & Qiu, M. (2007). Induction of oligodendrocyte differentiation by Olig2 and Sox10: evidence for reciprocal interactions and dosage-dependent mechanisms. *Developmental Biology*, *302*(2), 683–693.

Liu, Z., Yu, S., & Manley, N. R. (2007). Gcm2 is required for the differentiation and survival of parathyroid precursor cells in the parathyroid/thymus primordia. *Developmental Biology*, *305*(1), 333–346.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* , *25*(4), 402–408.

Li, W., Fu, L., Niu, B., Wu, S., & Wooley, J. (2012). Ultrafast clustering algorithms for metagenomic sequence analysis. *Briefings in Bioinformatics*, *13*(6), 656–668.

Li, Y., Shen, X.-X., Evans, B., Dunn, C. W., & Rokas, A. (2021). Rooting the animal tree of life. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msab170

Lockhart, S. T., Levitan, I. B., & Pikielny, C. W. (1996). Ag, a novel protein secreted from Aplysia glia. *Journal of Neurobiology*, *29*(1), 35–48.

Louvi, A., & Artavanis-Tsakonas, S. (2006). Notch signaling in vertebrate neural development. *Nature Reviews. Neuroscience*, *7*(2), 93–102.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550.

Lu, Q. R., Sun, T., Zhu, Z., Ma, N., Garcia, M., Stiles, C. D., & Rowitch, D. H. (2002). Common developmental requirement for Olig function indicates a motor neuron/oligodendrocyte connection. *Cell*, *109*(1), 75–86.

Lu, Z., & Kipnis, J. (2010). Thrombospondin 1--a key astrocyte-derived neurogenic factor. *FASEB Journal: Official Publication of the Federation of American Societies for*

Experimental Biology, *24*(6), 1925–1934.
Magadi, S. S., Voutyraki, C., Anagnostopoulos, G., Zacharioudaki, E., Poutakidou, I. K., Efraimoglou, C., Stapountzi, M., Theodorou, V., Nikolaou, C., Koumbanakis, K. A., Fullard, J. F., & Delidakis, C. (2020). Dissecting Hes-centred transcriptional networks in neural stem cell maintenance and tumorigenesis in Drosophila. *Development* , *147*(22). https://doi.org/10.1242/dev.191544
Magie, C. R., Pang, K., & Martindale, M. Q. (2005). Genomic inventory and expression of Sox and Fox genes in the cnidarian Nematostella vectensis. *Development Genes and Evolution*, *215*(12), 618–630.
Maher, F. (1995). Immunolocalization of GLUT1 and GLUT3 glucose transporters in primary cultured neurons and glia. *Journal of Neuroscience Research*, *42*(4), 459–469.
Marlow, H. Q., Srivastava, M., Matus, D. Q., Rokhsar, D., & Martindale, M. Q. (2009). Anatomy and development of the nervous system of Nematostella vectensis, an anthozoan cnidarian. *Developmental Neurobiology*, *69*(4), 235–254.
Marlow, H., Roettinger, E., Boekhout, M., & Martindale, M. Q. (2012). Functional roles of Notch signaling in the cnidarian Nematostella vectensis. *Developmental Biology*, *362*(2), 295–308.
Martindale, M. Q., Pang, K., & Finnerty, J. R. (2004). Investigating the origins of triploblasty: "mesodermal" gene expression in a diploblastic animal, the sea anemone Nematostella vectensis (phylum, Cnidaria; class, Anthozoa). *Development* , *131*(10), 2463–2474.
Martineau, M., Shi, T., Puyal, J., Knolhoff, A. M., Dulong, J., Gasnier, B., Klingauf, J., Sweedler, J. V., Jahn, R., & Mothet, J.-P. (2013). Storage and uptake of D-serine into astrocytic synaptic-like vesicles specify gliotransmission. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(8), 3413–3423.
Mashanov, V. S., Zueva, O. R., & Heinzeller, T. (2008). Regeneration of the radial nerve cord in a holothurian: a promising new model system for studying post-traumatic recovery in the adult nervous system. *Tissue & Cell*, *40*(5), 351–372.
Mashanov, V., & Zueva, O. (2019). Radial Glia in Echinoderms. *Developmental Neurobiology*, *79*(5), 396–405.
Matus, D. Q., Magie, C. R., Pang, K., Martindale, M. Q., & Thomsen, G. H. (2008). The Hedgehog gene family of the cnidarian, Nematostella vectensis, and implications for understanding metazoan Hedgehog pathway evolution. *Developmental Biology*, *313*(2), 501–518.
Mazzolini, J., Le Clerc, S., Morisse, G., Coulonges, C., Kuil, L. E., van Ham, T. J., Zagury, J.-F., & Sieger, D. (2020). Gene expression profiling reveals a conserved microglia signature in larval zebrafish. *Glia*, *68*(2), 298–315.
Mehta, P., & Piao, X. (2017). Adhesion G-protein coupled receptors and extracellular matrix proteins: Roles in myelination and glial cell development. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *246*(4), 275–284.
Melzer, L., Freiman, T. M., & Derouiche, A. (2021). Rab6A as a Pan-Astrocytic Marker in Mouse and Human Brain, and Comparison with Other Glial Markers (GFAP, GS, Aldh1L1, SOX9). *Cells* , *10*(1). https://doi.org/10.3390/cells10010072
Michinaga, S., & Koyama, Y. (2019). Dual Roles of Astrocyte-Derived Factors in Regulation of Blood-Brain Barrier Function after Brain Damage. *International Journal of Molecular Sciences*, *20*(3). https://doi.org/10.3390/ijms20030571
Miller, D. J., & Ball, E. E. (2008). Animal evolution: Trichoplax, trees, and taxonomic

turmoil. *Current Biology: CB*, *18*(21), R1003–R1005.
Mogha, A., D'Rozario, M., & Monk, K. R. (2016). G Protein-Coupled Receptors in Myelinating Glia. *Trends in Pharmacological Sciences*, *37*(11), 977–987.
Mohebiany, A. N., Harroch, S., & Bouyain, S. (2014). New Insights into the Roles of the Contactin Cell Adhesion Molecules in Neural Development. In V. Berezin & P. S. Walmod (Eds.), *Cell Adhesion Molecules: Implications in Neurological Diseases* (pp. 165–194). Springer New York.
Molofsky, A. V., Glasgow, S. M., Chaboub, L. S., Tsai, H.-H., Murnen, A. T., Kelley, K. W., Fancy, S. P. J., Yuen, T. J., Madireddy, L., Baranzini, S., Deneen, B., Rowitch, D. H., & Oldham, M. C. (2013). Expression profiling of Aldh1l1-precursors in the developing spinal cord reveals glial lineage-specific genes and direct Sox9-Nfe2l1 interactions. *Glia*, *61*(9), 1518–1532.
Monastirioti, M., Giagtzoglou, N., Koumbanakis, K. A., Zacharioudaki, E., Deligiannaki, M., Wech, I., Almeida, M., Preiss, A., Bray, S., & Delidakis, C. (2010). Drosophila Hey is a target of Notch in asymmetric divisions during embryonic and larval neurogenesis. *Development* , *137*(2), 191–201.
Morgello, S., Uson, R. R., Schwartz, E. J., & Haber, R. S. (1995). The human blood-brain barrier glucose transporter (GLUT1) is a glucose transporter of gray matter astrocytes. *Glia*, *14*(1), 43–54.
Morizawa, Y. M., Hirayama, Y., Ohno, N., Shibata, S., Shigetomi, E., Sui, Y., Nabekura, J., Sato, K., Okajima, F., Takebayashi, H., Okano, H., & Koizumi, S. (2017). Reactive astrocytes function as phagocytes after brain ischemia via ABCA1-mediated pathway. *Nature Communications*, *8*(1), 28.
Moroz, L. L., Nikitin, M. A., Poličar, P. G., Kohn, A. B., & Romanova, D. Y. (2021). Evolution of glutamatergic signaling and synapses. *Neuropharmacology*, *199*, 108740.
Moroz, L. L., Romanova, D. Y., & Kohn, A. B. (2021). Neural versus alternative integrative systems: molecular insights into origins of neurotransmitters. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *376*(1821), 20190762.
Nishino, J., Yamashita, K., Hashiguchi, H., Fujii, H., Shimazaki, T., & Hamada, H. (2004). Meteorin: a secreted protein that regulates glial cell differentiation and promotes axonal extension. *The EMBO Journal*, *23*(9), 1998–2008.
Norekian, T. P., & Moroz, L. L. (2020). Comparative neuroanatomy of ctenophores: Neural and muscular systems in Euplokamis dunlapae and related species. *The Journal of Comparative Neurology*, *528*(3), 481–501.
Nowotny, J. D., Connelly, M. T., & Traylor-Knowles, N. (2021). Novel methods to establish whole-body primary cell cultures for the cnidarians Nematostella vectensis and Pocillopora damicornis. *Scientific Reports*, *11*(1), 4086.
Oerter, S., Förster, C., & Bohnert, M. (2019). Validation of sodium/glucose cotransporter proteins in human brain as a potential marker for temporal narrowing of the trauma formation. *International Journal of Legal Medicine*, *133*(4), 1107–1114.
Ohayon, D., Aguirrebengoa, M., Escalas, N., & Soula, C. (2020). Transcriptome profiling of the Olig2-expressing astrocyte subtype reveals their unique molecular signature. In *bioRxiv* (p. 2020.10.15.340505). https://doi.org/10.1101/2020.10.15.340505
Oikonomou, G., & Shaham, S. (2011). The glia of Caenorhabditis elegans. *Glia*, *59*(9), 1253–1263.
Okawa, N., Shimai, K., Ohnishi, K., Ohkura, M., Nakai, J., Horie, T., Kuhara, A., &

Kusakabe, T. G. (2020). Cellular identity and Ca2+ signaling activity of the non-reproductive GnRH system in the Ciona intestinalis type A (Ciona robusta) larva. *Scientific Reports*, *10*(1), 18590.

Olsen, M. L., Khakh, B. S., Skatchkov, S. N., Zhou, M., Lee, C. J., & Rouach, N. (2015). New Insights on Astrocyte Ion Channels: Critical for Homeostasis and Neuron-Glia Signaling. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(41), 13827–13835.

Oren, M., Brickner, I., Brikner, I., Appelbaum, L., & Levy, O. (2014). Fast neurotransmission related genes are expressed in non nervous endoderm in the sea anemone Nematostella vectensis. *PloS One*, *9*(4), e93832.

Padmashri, R., Ren, B., Oldham, B., Jung, Y., Gough, R., & Dunaevsky, A. (2021). Modeling human-specific interlaminar astrocytes in the mouse cerebral cortex. *The Journal of Comparative Neurology*, *529*(4), 802–810.

Pang, K., Ryan, J. F., Baxevanis, A. D., & Martindale, M. Q. (2011). Evolution of the TGF-β signaling pathway and its potential role in the ctenophore, Mnemiopsis leidyi. *PloS One*, *6*(9), e24152.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419.

Peco, E., Davla, S., Camp, D., Stacey, S. M., Landgraf, M., & van Meyel, D. J. (2016). Drosophila astrocytes cover specific territories of the CNS neuropil and are instructed to differentiate by Prospero, a key effector of Notch. *Development* , *143*(7), 1170–1181.

Pepino, M. Y., Kuda, O., Samovski, D., & Abumrad, N. A. (2014). Structure-function of CD36 and importance of fatty acid signal transduction in fat metabolism. *Annual Review of Nutrition*, *34*, 281–303.

Perea, G., Sur, M., & Araque, A. (2014). Neuron-glia networks: integral gear of brain function. *Frontiers in Cellular Neuroscience*, *8*, 378.

Phochanukul, N., & Russell, S. (2010). No backbone but lots of Sox: Invertebrate Sox genes. *The International Journal of Biochemistry & Cell Biology*, *42*(3), 453–464.

Pierobon, P. (2021). An Interesting Molecule: γ-Aminobutyric Acid. What Can We Learn from Hydra Polyps? *Brain Sciences*, *11*(4). https://doi.org/10.3390/brainsci11040437

Preston, A. N., Farr, J. D., O'Neill, B. K., Thompson, K. K., Tsirka, S. E., & Laughlin, S. T. (2018). Visualizing the Brain's Astrocytes with Diverse Chemical Scaffolds. *ACS Chemical Biology*, *13*(6), 1493–1498.

Proctor, D. T., Stotz, S. C., Scott, L. O. M., de la Hoz, C. L. R., Poon, K. W. C., Stys, P. K., & Colicos, M. A. (2015). Axo-glial communication through neurexin-neuroligin signaling regulates myelination and oligodendrocyte differentiation. *Glia*, *63*(11), 2023–2039.

Putnam, N. H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V. V., Jurka, J., Genikhovich, G., Grigoriev, I. V., Lucas, S. M., Steele, R. E., Finnerty, J. R., Technau, U., Martindale, M. Q., & Rokhsar, D. S. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science*, *317*(5834), 86–94.

Radojcic, T., & Pentreath, V. W. (1979). Invertebrate glia. *Progress in Neurobiology*, *12*(2), 115–179.

Randy F. Stout, A. V. A. V. P., Jr. (2014). Caenorhabditis elegans glia modulate neuronal activity and behavior. *Front. Cell. Neurosci.* https://doi.org/10.3389/fncel.2014.00067

Ransick, A., & Davidson, E. H. (2006). cis-regulatory processing of Notch signaling input to the sea urchin glial cells missing gene during mesoderm specification. *Developmental Biology*, *297*(2), 587–602.

Reiprich, S., & Wegner, M. (2015). From CNS stem cells to neurons and glia: Sox for everyone. *Cell and Tissue Research*, *359*(1), 111–124.

Renfer, E., Amon-Hassenzahl, A., Steinmetz, P. R. H., & Technau, U. (2010). A muscle-specific transgenic reporter line of the sea anemone, Nematostella vectensis. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(1), 104–108.

Rentzsch, F., Layden, M., & Manuel, M. (2017). The cellular and molecular basis of cnidarian neurogenesis. *Wiley Interdisciplinary Reviews. Developmental Biology*, *6*(1). https://doi.org/10.1002/wdev.257

Rentzsch, F., Renfer, E., & Technau, U. (2020). Generating Transgenic Reporter Lines for Studying Nervous System Development in the Cnidarian Nematostella vectensis. In S. G. Sprecher (Ed.), *Brain Development: Methods and Protocols* (pp. 45–57). Springer New York.

Rey, S., Zalc, B., & Klämbt, C. (2020). Evolution of glial wrapping: A new hypothesis. *Developmental Neurobiology*. https://doi.org/10.1002/dneu.22739

Richards, G. S., & Rentzsch, F. (2015). Regulation of Nematostella neural progenitors by SoxB, Notch and bHLH genes. *Development* , *142*(19), 3332–3342.

Richards, G. S., Simionato, E., Perron, M., Adamska, M., Vervoort, M., & Degnan, B. M. (2008). Sponge genes provide new insight into the evolutionary origin of the neurogenic circuit. *Current Biology: CB*, *18*(15), 1156–1161.

Rodríguez, E., Guerra, M., Peruzzo, B., & Blázquez, J. L. (2019). Tanycytes: A rich morphological history to underpin future molecular and physiological investigations. *Journal of Neuroendocrinology*, *31*(3), e12690.

Romanova, D. Y., Heyland, A., Sohn, D., Kohn, A. B., Fasshauer, D., Varoqueaux, F., & Moroz, L. L. (2020). Glycine as a signaling molecule and chemoattractant in Trichoplax (Placozoa): insights into the early evolution of neurotransmitters. *Neuroreport*, *31*(6), 490–497.

Röttinger, E., Dahlin, P., & Martindale, M. Q. (2012). A framework for the establishment of a cnidarian gene regulatory network for “endomesoderm” specification: the inputs of ß-catenin/TCF signaling. *PLoS Genetics*, *8*(12), e1003164.

Rowlands, D., Lensjø, K. K., Dinh, T., Yang, S., Andrews, M. R., Hafting, T., Fyhn, M., Fawcett, J. W., & Dick, G. (2018). Aggrecan Directs Extracellular Matrix-Mediated Neuronal Plasticity. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(47), 10102–10113.

Ryan, J. F., & Chiodin, M. (2015). Where is my mind? How sponges and placozoans may have lost neural cell types. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1684). https://doi.org/10.1098/rstb.2015.0059

Sachkova, M. Y., Nordmann, E.-L., Soto-Àngel, J. J., Meeda, Y., Górski, B., Naumann, B., Dondorp, D., Chatzigeorgiou, M., Kittelmann, M., & Burkhardt, P. (2021). The unique neuronal structure and neuropeptide repertoire in the ctenophore Mnemiopsis leidyi shed light on the evolution of animal nervous systems. In *bioRxiv* (p. 2021.03.31.437758). https://doi.org/10.1101/2021.03.31.437758

Saint-Martin, M., Joubert, B., Pellier-Monnin, V., Pascual, O., Noraz, N., & Honnorat, J.

(2018). Contactin-associated protein-like 2, a protein of the neurexin family involved in several human diseases. *The European Journal of Neuroscience*, *48*(3), 1906–1923.

Saito, A., Kanemoto, S., Kawasaki, N., Asada, R., Iwamoto, H., Oki, M., Miyagi, H., Izumi, S., Sanosaka, T., Nakashima, K., & Imaizumi, K. (2012). Unfolded protein response, activated by OASIS family transcription factors, promotes astrocyte differentiation. *Nature Communications*, *3*, 967.

Salzer, J. L. (2015). Schwann cell myelination. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a020529.

Samanta, A., Hughes, T. E. T., & Moiseenkova-Bell, V. Y. (2018). Transient Receptor Potential (TRP) Channels. *Sub-Cellular Biochemistry*, *87*, 141–165.

Sancho, L., Contreras, M., & Allen, N. J. (2020). Glia as sculptors of synaptic plasticity. *Neuroscience Research*. https://doi.org/10.1016/j.neures.2020.11.005

Sauvageot, C. M., & Stiles, C. D. (2002). Molecular mechanisms controlling cortical gliogenesis. *Current Opinion in Neurobiology*, *12*(3), 244–249.

Schmidt-Rhaesa, A., Harzsch, S., & Purschke, G. (2015). *Structure and Evolution of Invertebrate Nervous Systems*. Oxford University Press.

Schnitzler, C. E., Simmons, D. K., Pang, K., Martindale, M. Q., & Baxevanis, A. D. (2014). Expression of multiple Sox genes through embryonic development in the ctenophore Mnemiopsis leidyi is spatially restricted to zones of cell proliferation. *EvoDevo*, *5*, 15.

Schreiber, J., Riethmacher-Sonnenberg, E., Riethmacher, D., Tuerk, E. E., Enderich, J., Bösl, M. R., & Wegner, M. (2000). Placental failure in mice lacking the mammalian homolog of glial cells missing, GCMa. *Molecular and Cellular Biology*, *20*(7), 2466–2474.

Scolding, N. J., Frith, S., Linington, C., Morgan, B. P., Campbell, A. K., & Compston, D. A. (1989). Myelin-oligodendrocyte glycoprotein (MOG) is a surface marker of oligodendrocyte maturation. *Journal of Neuroimmunology*, *22*(3), 169–176.

Sebé-Pedrós, A., Chomsky, E., Pang, K., Lara-Astiaso, D., Gaiti, F., Mukamel, Z., Amit, I., Hejnol, A., Degnan, B. M., & Tanay, A. (2018). Early metazoan cell type diversity and the evolution of multicellular gene regulation. *Nature Ecology & Evolution*, *2*(7), 1176–1188.

Sebé-Pedrós, A., Saudemont, B., Chomsky, E., Plessier, F., Mailhé, M.-P., Renno, J., Loe-Mie, Y., Lifshitz, A., Mukamel, Z., Schmutz, S., Novault, S., Steinmetz, P. R. H., Spitz, F., Tanay, A., & Marlow, H. (2018). Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. *Cell*, *173*(6), 1520–1534.e20.

See, J., Mamontov, P., Ahn, K., Wine-Lee, L., Crenshaw, E. B., 3rd, & Grinspan, J. B. (2007). BMP signaling mutant mice exhibit glial cell maturation defects. *Molecular and Cellular Neurosciences*, *35*(1), 171–182.

Seifert, G., & Steinhäuser, C. (2018). Heterogeneity and function of hippocampal macroglia. *Cell and Tissue Research*, *373*(3), 653–670.

Senatore, A., Reese, T. S., & Smith, C. L. (2017). Neuropeptidergic integration of behavior in Trichoplax adhaerens, an animal without synapses. *The Journal of Experimental Biology*, *220*(Pt 18), 3381–3390.

Shaham, S. (2005). Glia–neuron interactions in nervous system function and development. *Current Topics in Developmental Biology*. http://www.sciencedirect.com/science/article/pii/S0070215305690035

Sharma, K., Bisht, K., & Eyo, U. B. (2021). A comparative biology of microglia across species. *Frontiers in Cell and Developmental Biology*, *9*, 652748.

Shen, X.-X., Hittinger, C. T., & Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Ecology & Evolution*, *1*(5), s41559–017 – 0126.

Shigetomi, E., Tong, X., Kwan, K. Y., Corey, D. P., & Khakh, B. S. (2011). TRPA1 channels regulate astrocyte resting calcium and inhibitory synapse efficacy through GAT-3. *Nature Neuroscience*, *15*(1), 70–80.

Shinzato, C., Iguchi, A., Hayward, D. C., Technau, U., Ball, E. E., & Miller, D. J. (2008). Sox genes in the coral Acropora millepora: divergent expression patterns reflect differences in developmental mechanisms within the Anthozoa. *BMC Evolutionary Biology*, *8*, 311.

Sokolowski, J. D., Nobles, S. L., Heffron, D. S., Park, D., Ravichandran, K. S., & Mandell, J. W. (2011). Brain-specific angiogenesis inhibitor-1 expression in astrocytes and neurons: implications for its dual function as an apoptotic engulfment receptor. *Brain, Behavior, and Immunity*, *25*(5), 915–921.

Soustelle, L., & Giangrande, A. (2007). Glial differentiation and the Gcm pathway. *Neuron Glia Biology*, *3*(1), 5–16.

Soustelle, L., Trousse, F., Jacques, C., Ceron, J., Cochard, P., Soula, C., & Giangrande, A. (2007). Neurogenic role of Gcm transcription factors is conserved in chicken spinal cord. *Development* , *134*(3), 625–634.

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

Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. A., Mitros, T., Richards, G. S., Conaco, C., Dacre, M., Hellsten, U., Larroux, C., Putnam, N. H., Stanke, M., Adamska, M., Darling, A., Degnan, S. M., Oakley, T. H., Plachetzki, D. C., Zhai, Y., … Rokhsar, D. S. (2010). The Amphimedon queenslandica genome and the evolution of animal complexity. *Nature*, *466*(7307), 720–726.

States, D. J., & Gish, W. (1994). Combined use of sequence similarity and codon bias for coding region identification. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, *1*(1), 39–50.

Stolt, C. C., Lommes, P., Sock, E., Chaboissier, M.-C., Schedl, A., & Wegner, M. (2003). The Sox9 transcription factor determines glial fate choice in the developing spinal cord. *Genes & Development*, *17*(13), 1677–1689.

Stork, T., Sheehan, A., Tasdemir-Yilmaz, O. E., & Freeman, M. R. (2014). Neuron-glia interactions through the Heartless FGF receptor signaling pathway mediate morphogenesis of Drosophila astrocytes. *Neuron*, *83*(2), 388–403.

Stork, T., Thomas, S., Rodrigues, F., Silies, M., Naffin, E., Wenderdel, S., & Klämbt, C. (2009). Drosophila Neurexin IV stabilizes neuron-glia interactions at the CNS midline by binding to Wrapper. *Development* , *136*(8), 1251–1261.

Suess, U., & Pliska, V. (1981). Identification of the pituicytes as astroglial cells by indirect immunofluorescence-staining for the glial fibrillary acidic protein. *Brain Research*, *221*(1), 27–33.

Sun, G. Y., Xu, J., Jensen, M. D., & Simonyi, A. (2004). Phospholipase A2 in the central nervous system: implications for neurodegenerative diseases. *Journal of Lipid Research*, *45*(2), 205–213.

Tabula Muris Consortium, Overall coordination, Logistical coordination, Organ collection and processing, Library preparation and sequencing, Computational data analysis, Cell type annotation, Writing group, Supplemental text writing group, & Principal investigators. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*, *562*(7727), 367–372.

T Cuenco, K., Lunetta, K. L., Baldwin, C. T., McKee, A. C., Guo, J., Cupples, L. A., Green, R. C., St George-Hyslop, P. H., Chui, H., DeCarli, C., Farrer, L. A., & MIRAGE Study Group. (2008). Association of distinct variants in SORL1 with cerebrovascular and neurodegenerative changes related to Alzheimer disease. *Archives of Neurology*, *65*(12), 1640–1648.

Thal, D. R. (2012). The role of astrocytes in amyloid β-protein toxicity and clearance [Review of *The role of astrocytes in amyloid β-protein toxicity and clearance*]. *Experimental Neurology*, *236*(1), 1–5.

Tournière, O., Busengdal, H., Gahan, J. M., & Rentzsch, F. (2021). Insm1-expressing neurons and secretory cells develop from a common pool of progenitors in the sea anemone Nematostella vectensis. In *bioRxiv* (p. 2021.04.09.439178). https://doi.org/10.1101/2021.04.09.439178

Tournière, O., Dolan, D., Richards, G. S., Sunagar, K., Columbus-Shenkar, Y. Y., Moran, Y., & Rentzsch, F. (2020). NvPOU4/Brain3 Functions as a Terminal Selector Gene in the Nervous System of the Cnidarian Nematostella vectensis. *Cell Reports*, *30*(13), 4473–4489.e5.

Trébuchet, G., Cattenoz, P. B., Zsámboki, J., Mazaud, D., Siekhaus, D. E., Fanto, M., & Giangrande, A. (2019). The Repo Homeodomain Transcription Factor Suppresses Hematopoiesis in Drosophila and Preserves the Glial Fate. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(2), 238–255.

Trotter, J., Karram, K., & Nishiyama, A. (2010). NG2 cells: Properties, progeny and origin. *Brain Research Reviews*, *63*(1-2), 72–82.

Udolph, G., Rath, P., & Chia, W. (2001). A requirement for Notch in the genesis of a subset of glial cells in the Drosophila embryonic central nervous system which arise through asymmetric divisions. *Development* , *128*(8), 1457–1466.

Umesono, Y., Hiromi, Y., & Hotta, Y. (2002). Context-dependent utilization of Notch activity in Drosophila glial determination. *Development* , *129*(10), 2391–2399.

Vainchtein, I. D., & Molofsky, A. V. (2020). Astrocytes and Microglia: In Sickness and in Health. *Trends in Neurosciences*, *43*(3), 144–154.

Van De Bor, V., & Giangrande, A. (2001). Notch signaling represses the glial fate in fly PNS. *Development* , *128*(8), 1381–1390.

van Deijk, A.-L. F., Camargo, N., Timmerman, J., Heistek, T., Brouwers, J. F., Mogavero, F., Mansvelder, H. D., Smit, A. B., & Verheijen, M. H. G. (2017). Astrocyte lipid metabolism is critical for synapse development and function in vivo. *Glia*, *65*(4), 670–682.

Varoqueaux, F., Williams, E. A., Grandemange, S., Truscello, L., Kamm, K., Schierwater, B., Jékely, G., & Fasshauer, D. (2018). High Cell Diversity and Complex Peptidergic Signaling Underlie Placozoan Behavior. *Current Biology: CB*, *28*(21), 3495–3501.e2.

Vasile, F., Dossi, E., & Rouach, N. (2017). Human astrocytes: structure and functions in the healthy brain. *Brain Structure & Function*, *222*(5), 2017–2029.

Vemula, S., Roder, K. E., Yang, T., Bhat, G. J., Thekkumkara, T. J., & Abbruscato, T. J.

(2009). A functional role for sodium-dependent glucose transport across the blood-brain barrier during oxygen glucose deprivation. *The Journal of Pharmacology and Experimental Therapeutics*, *328*(2), 487–495.

Verkhratsky, A., & Butt, A. (2013). Neuroglia: Definition, Classification, Evolution, Numbers, Development. In *Glial Physiology and Pathophysiology* (pp. 73–104). John Wiley & Sons, Ltd, UK.

Verkhratsky, A., Ho, M. S., & Parpura, V. (2019). Evolution of Neuroglia. In A. Verkhratsky, M. S. Ho, R. Zorec, & V. Parpura (Eds.), *Neuroglia in Neurodegenerative Diseases* (pp. 15–44). Springer Singapore.

Verkhratsky, A., & Nedergaard, M. (2014). Astroglial cradle in the life of the synapse. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1654), 20130595.

Viehweg, J., Naumann, W. W., & Olsson, R. (1998). Secretory radial Glia in the ectoneural system of the sea StarAsterias rubens(echinodermata). *Acta Zoologica* , *79*(2), 119–131.

Wang, I. E., Lapan, S. W., Scimone, M. L., Clandinin, T. R., & Reddien, P. W. (2016). Hedgehog signaling regulates gene expression in planarian glia. *eLife*, *5*. https://doi.org/10.7554/eLife.16996

Wan, H., Liao, J., Zhang, Z., Zeng, X., Liang, K., & Wang, Y. (2021). Molecular cloning, characterization, and expression analysis of a sex-biased transcriptional factor sox9 gene of mud crab Scylla paramamosain. *Gene*, *774*, 145423.

Wasseff, S. K., & Scherer, S. S. (2011). Cx32 and Cx47 mediate oligodendrocyte:astrocyte and oligodendrocyte:oligodendrocyte gap junction coupling. *Neurobiology of Disease*, *42*(3), 506–513.

Watanabe, H. (2017). Back Through Time: How Cnidarians and Basal Metazoans Shed Light on Ancient Nervous Systems. In *Brain Evolution by Design* (pp. 45–75). Springer, Tokyo.

Watanabe, H., Fujisawa, T., & Holstein, T. W. (2009). Cnidarians and the evolutionary origin of the nervous system. *Development, Growth & Differentiation*, *51*(3), 167–183.

Watanabe, H., Kuhn, A., Fushiki, M., Agata, K., Özbek, S., Fujisawa, T., & Holstein, T. W. (2014). Sequential actions of β-catenin and Bmp pattern the oral nerve net in Nematostella vectensis. *Nature Communications*, *5*, 5536.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* , *25*(9), 1189–1191.

Wessel, G. M., Kiyomoto, M., Shen, T.-L., & Yajima, M. (2020). Genetic manipulation of the pigment pathway in a sea urchin reveals distinct lineage commitment prior to metamorphosis in the bilateral to radial body plan transition. *Scientific Reports*, *10*(1), 1973.

Wheeler, S. R., Stagg, S. B., & Crews, S. T. (2008). Multiple Notch signaling events control Drosophila CNS midline neurogenesis, gliogenesis and neuronal identity. *Development* , *135*(18), 3071–3079.

Wolenski, F. S., Layden, M. J., Martindale, M. Q., Gilmore, T. D., & Finnerty, J. R. (2013). Characterizing the spatiotemporal expression of RNAs and proteins in the starlet sea anemone, Nematostella vectensis. *Nature Protocols*, *8*(5), 900–915.

Yamada, K., Ogawa, H., Honda, S.-I., Harada, N., & Okazaki, T. (1999). A GCM Motif Protein Is Involved in Placenta-specific Expression of Human Aromatase Gene*. *The Journal of Biological Chemistry*, *274*(45), 32279–32286.

Yamagata, M., & Sanes, J. R. (2008). Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. *Nature*, *451*(7177), 465–469.

Yamaguchi, Y. (1996). Brevican: a major proteoglycan in adult brain. *Perspectives on Developmental Neurobiology*, *3*(4), 307–317.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*, 134.

Yildirim, K., Petri, J., Kottmeier, R., & Klämbt, C. (2019). Drosophila glia: Few cell types and many conserved functions. *Glia*, *67*(1), 5–26.

Yordy, J. S., & Muise-Helmericks, R. C. (2000). Signal transduction and the Ets family of transcription factors. *Oncogene*, *19*(55), 6503–6513.

Yoshimura, S., Murray, J. I., Lu, Y., Waterston, R. H., & Shaham, S. (2008). mls-2 and vab-3 Control glia development, hlh-17/Olig expression and glia-dependent neurite extension in C. elegans. *Development* , *135*(13), 2263–2275.

Yuan, L. L., & Ganetzky, B. (1999). A glial-neuronal signaling pathway revealed by mutations in a neurexin-related protein. *Science*, *283*(5406), 1343–1345.

Yuan, T., York, J. R., & McCauley, D. W. (2018). Gliogenesis in lampreys shares gene regulatory interactions with oligodendrocyte development in jawed vertebrates. *Developmental Biology*, *441*(1), 176–190.

Yuasa, Y., Okabe, M., Yoshikawa, S., Tabuchi, K., Xiong, W.-C., Hiromi, Y., & Okano, H. (2003). Drosophila homeodomain protein REPO controls glial differentiation by cooperating with ETS and BTB transcription factors. *Development* , *130*(11), 2419–2428.

Zhang, Y., & Barres, B. A. (2010). Astrocyte heterogeneity: an underappreciated topic in neurobiology. *Current Opinion in Neurobiology*, *20*(5), 588–594.

Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., Phatnani, H. P., Guarnieri, P., Caneda, C., Ruderisch, N., Deng, S., Liddelow, S. A., Zhang, C., Daneman, R., Maniatis, T., Barres, B. A., & Wu, J. Q. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(36), 11929–11947.

Zhao, F.-Q., & Keating, A. F. (2007). Functional properties and genomics of glucose transporters. *Current Genomics*, *8*(2), 113–128.

Zhou, B., Zhu, Z., Ransom, B. R., & Tong, X. (2021). Oligodendrocyte lineage cells and depression. *Molecular Psychiatry*, *26*(1), 103–117.

Zhou, M., Yan, J., Ma, Z., Zhou, Y., Abbood, N. N., Liu, J., Su, L., Jia, H., & Guo, A.-Y. (2012). Comparative and evolutionary analysis of the HES/HEY gene family reveal exon/intron loss and teleost specific duplication events. *PloS One*, *7*(7), e40649.

Zhou, Q., & Anderson, D. J. (2002). The bHLH transcription factors OLIG2 and OLIG1 couple neuronal and glial subtype specification. *Cell*, *109*(1), 61–73.

Zhou, Z.-D., Kumari, U., Xiao, Z.-C., & Tan, E.-K. (2010). Notch as a molecular switch in neural stem cells. *IUBMB Life*, *62*(8), 618–623.

Zimmermann, B., Robb, S. M. C., Genikhovich, G., Fropf, W. J., Weilguny, L., He, S., Chen, S., Lovegrove-Walsh, J., Hill, E. M., Ragkousi, K., Praher, D., Fredman, D., Moran, Y., Gibson, M. C., & Technau, U. (2020). Sea anemone genomes reveal ancestral metazoan chromosomal macrosynteny. In *bioRxiv* (p. 2020.10.30.359448). https://doi.org/10.1101/2020.10.30.359448

Zollo, A., Allen, Z., Rasmussen, H. F., Iannuzzi, F., Shi, Y., Larsen, A., Maier, T. J., & Matrone, C. (2017). Sortilin-Related Receptor Expression in Human Neural Stem Cells Derived from Alzheimer's Disease Patients Carrying the APOE Epsilon 4 Allele. *Neural Plasticity*, *2017*, 1892612.

Zorec, R., Verkhratsky, A., Rodríguez, J. J., & Parpura, V. (2016). Astrocytic vesicles and gliotransmitters: Slowness of vesicular release and synaptobrevin2-laden vesicle nanoarchitecture. *Neuroscience*, *323*, 67–75.

Appendices

Appendix 1. List of species screened for glial markers

Phylum	Species	Glial markers	References
Vertebrata	*Mus musculus*	*Gfap, vimentin, S100b, Gs, Eaat, Gat, Olig*	Cahoy, J. D. et al. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. J. Neurosci. 28, 264–278 (2008)
	Danio rerio	*Gfap, Gs, Eaat*	Grupp, L., Wolburg, H. & Mack, A. F. Astroglial structures in the zebrafish brain. J. Comp. Neurol. 518, 4277–4287 (2010); Jurisch-Yaksi, N., Yaksi, E. & Kizil, C. Radial glia in the zebrafish brain: Functional, structural, and physiological comparison with the mammalian glia. Glia 68, 2451–2470 (2020)
	Xenopus laevis	*Vimentin, Gfap*	Yoshida, M. Glial-defined boundaries in Xenopus CNS. Dev. Neurosci. 23, 299–306 (2001); D'Amico, L. A., Boujard, D. & Coumailleau, P. Proliferation, migration and differentiation in juvenile and adult Xenopus laevis brains. Brain Res. 1405, 31–48 (2011)
	Petromyzon marinus	*Olig1/2, Aldh1l1, Gfap*	Weil, M.-T. et al. Axonal ensheathment in the nervous system of lamprey: Implications for the evolution of myelinating glia. J. Neurosci. (2018) doi:10.1523/JNEUROSCI.1034-18.2018; Yuan, T., York, J. R. & McCauley, D. W. Gliogenesis in lampreys shares gene regulatory interactions with oligodendrocyte development in jawed vertebrates. Dev. Biol. 441, 176–190 (2018);
Cephalochordata	*European amphioxus*	*Eaat, Gs, Gfap*	Bozzo, M. et al. Amphioxus neuroglia: Molecular characterization and evidence for early compartmentalization of the developing nerve cord. Glia (2021) doi:10.1002/glia.23982
Urochordata	*Ciona intestinalis*	*Gonadotropin-releasing hormone (GnRH)*	Okawa, N. et al. Cellular identity and Ca2+ signaling activity of the non-reproductive GnRH system in the Ciona intestinalis type A (Ciona robusta) larva. Sci. Rep. 10, 18590 (2020)
Echinodermata	*Asterias rubens*	*If, sspo*	Helm Conrad et al. Early evolution of radial glial cells in Bilateria. Proceedings of the Royal Society B: Biological Sciences 284, 20170743 (2017)

	Ophioderma brevispinum	*sspo*	Mashanov, V. & Zueva, O. Radial Glia in Echinoderms. Dev. Neurobiol. 79, 396–405 (2019)
Hemichordaa	*Balanoglossus misakiensis*	*sspo*	Helm Conrad et al. Early evolution of radial glial cells in Bilateria. Proceedings of the Royal Society B: Biological Sciences 284, 20170743 (2017)
Annelida	*Magelonidae (mirabilis and alleni)*	*lf*	Beckers, P., Helm, C. & Bartolomaeus, T. The anatomy and development of the nervous system in Magelonidae (Annelida) – insights into the evolution of the annelid brain. BMC Evolutionary Biology 19, (2019)
	Owenia fusiformis	*sspo*	Helm Conrad et al. Early evolution of radial glial cells in Bilateria. Proceedings of the Royal Society B: Biological Sciences 284, 20170743 (2017)
	Eisenia fetida	*Gfap, S100*	Csoknya, M., Dénes, V. & Wilhelm, M. Glial cells in the central nervous system of earthworm, Eisenia fetida. Acta Biol. Hung. 63 Suppl 1, 114–128 (2012)
	Hirudo medicinalis	*Gfap*	Riehl, B. & Schlue, W. R. Morphological organization of neuropile glial cells in the central nervous system of the medicinal leech (Hirudo medicinalis). Tissue Cell 30, 177–186 (1998)
Platyhelminthes	*Schmidtea mediterranea*	*IF, Eaat, Gat, Glut, Gs*	Wang, I. E., Lapan, S. W., Scimone, M. L., Clandinin, T. R. & Reddien, P. W. Hedgehog signaling regulates gene expression in planarian glia. Elife 5, (2016)
	Christianella minuta	*S100b*	Biserova, N. M., Gordeev, I. I., Korneva, J. V. & Salnikova, M. M. Structure of the glial cells in the nervous system of parasitic and free-living flatworms. Biology Bulletin 37, 277–287 (2010)
Mollusca	*Aplysia californica*	*Gs, Ag*	Levenson, J. et al. Long-term regulation of neuronal high-affinity glutamate and glutamine uptake in Aplysia. Proc. Natl. Acad. Sci. U. S. A. 97, 12858–12863 (2000)
	Octopus vulgaris	*Gfap, vimentin*	Cardone, B. & Roots, B. I. Comparative immunohistochemical study of glial filament proteins (glial fibrillary acidic protein and vimentin) in goldfish, octopus, and snail. Glia 3, 180–192 (1990)
	Megalobulimus abbreviatus	*Gfap, vimentin*	Dos Santos, P. C., Gehlen, G., Faccioni-Heuser, M. C. & Achaval, M. Detection of glial fibrillary acidic protein (GFAP) and vimentin (Vim) by immunoelectron microscopy of the glial cells in the central nervous system of the snail Megalobulimus abbreviatus: GFAP and Vim in glial cells of Megalobulimus. Acta Zool. 86, 135–144 (2005)

Arthropoda	*Drosophila melanogaster*	*Gcm, repo, Eaat, Gs*	Yildirim, K., Petri, J., Kottmeier, R. & Klämbt, C. Drosophila glia: Few cell types and many conserved functions. Glia 67, 5–26 (2019); Stacey, S. M. et al. Drosophila glial glutamate transporter Eaat1 is regulated by fringe-mediated notch signaling and is essential for larval locomotion. J. Neurosci. 30, 14446–14457 (2010); Soustelle, L., Besson, M.-T., Rival, T. & Birman, S. Terminal glial differentiation involves regulated expression of the excitatory amino acid transporters in the Drosophila embryonic CNS. Dev. Biol. 248, 294–306 (2002)
	Panulirus argus	*Gs*	Linser, P. J., Trapido-Rosenthal, H. G. & Orona, E. Glutamine synthetase is a glial-specific marker in the olfactory regions of the lobster (Panulirus argus) nervous system. Glia 20, 275–283 (1997)
	Ucides cordatus	*Gfap*	Florim da Silva, S. et al. Glial fibrillary acidic protein (GFAP)-like immunoreactivity in the visual system of the crab Ucides cordatus (Crustacea, Decapoda). Biol. Cell 96, 727–734 (2004)
	Pacifastacus leniusculus	*Gcm*	Junkunlo, K., Söderhäll, K. & Söderhäll, I. A transcription factor glial cell missing (Gcm) in the freshwater crayfish Pacifastacus leniusculus. Dev. Comp. Immunol. 113, 103782 (2020)
	Macrobrachium rosenbergii	*S100b*	Allodi, S., Bressan, C. M., Carvalho, S. L. & Cavalcante, L. A. Regionally specific distribution of the binding of anti-glutamine synthetase and anti-S100 antibodies and of Datura stramonium lectin in glial domains of the optic lobe of the giant prawn. Glia 53, 612–620 (2006)
Nematoda	*Caenorhabditis elegans*	*Lin-26, Hlh-17*	Labouesse, M., Hartwieg, E. & Horvitz, H. R. The Caenorhabditis elegans LIN-26 protein is required to specify and/or maintain all non-neuronal ectodermal cell fates. Development 122, 2579–2588 (1996); Oikonomou, G. & Shaham, S. The glia of Caenorhabditis elegans. Glia 59, 1253–1263 (2011)
Acoela	*Symsagittifera roscoffensis*	*Gfap*	Bailly, X. et al. The chimerical and multifaceted marine acoel Symsagittifera roscoffensis: from photosymbiosis to brain regeneration. Front. Microbiol. 5, 498 (2014)

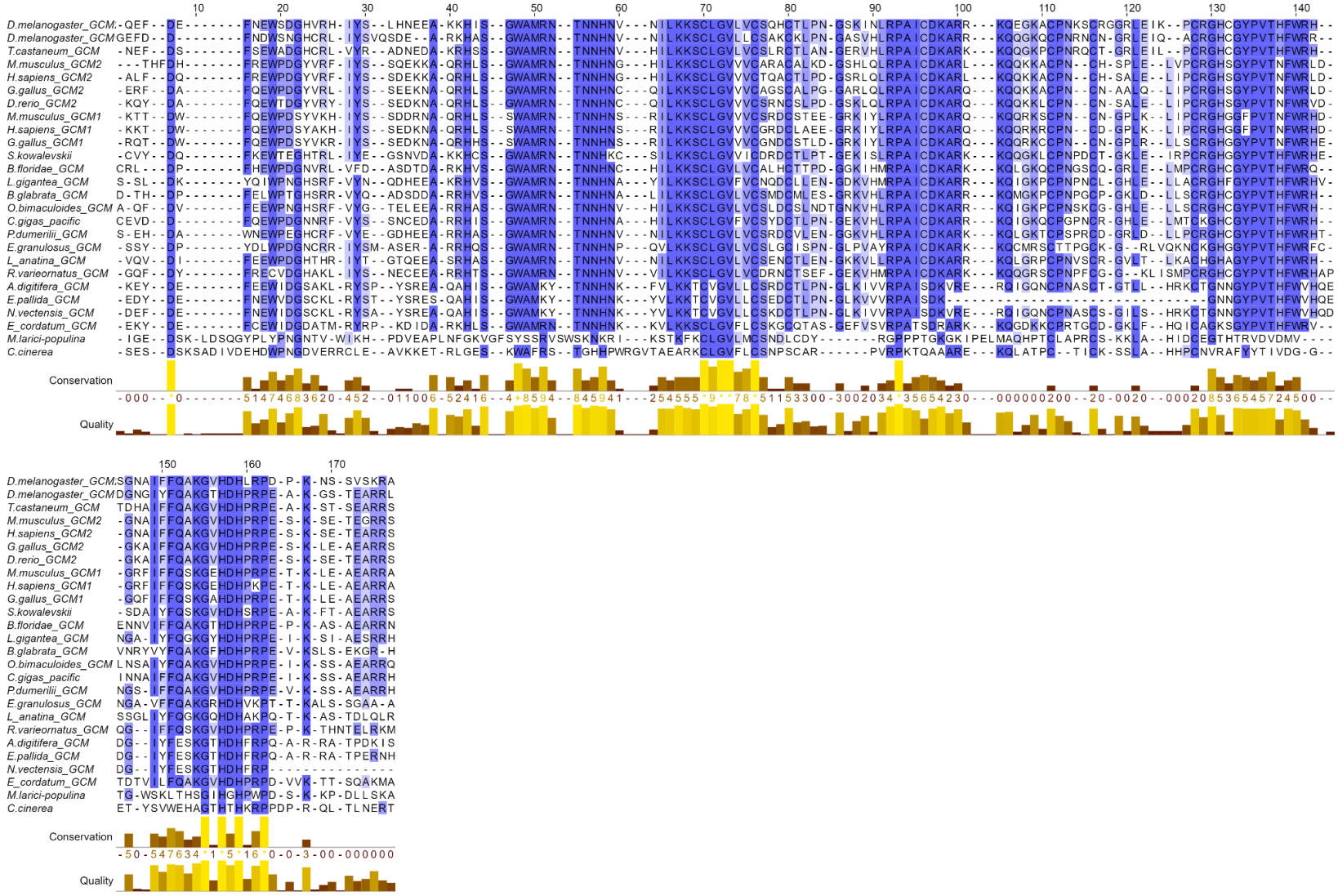
Appendix 2A. Species used for *Gcm* phylogenetic analysis

Phylum	Species	Gene name, Accession ID	Source
Chordata	*Homo sapiens*	GCM1, EAX04412.1	GenBank
		GCM2, NP004743	GenBank
	Mus musculus	GCM1, EDL26359.1	GenBank
		GCM2, NP_032130	GenBank
	Gallus gallus	GCM1, AAR24261	GenBank
		GCM2, BAD72825	GenBank
	Danio rerio	GCM2, AAI62304	GenBank
Hemichordata	*Saccoglossus kowalevskii*	GCM, NP_001161553	GenBank
Cephalochordata	*Branchiostoma floridae*	GCM, XP_002z591781	GenBank
Arthropoda	*Drosophila melanogaster*	GCM1, AAC46912	GenBank
	Drosophila melanogaster	GCM2, NP_609302	GenBank
	Tribolium castaneum	GCM, EFA04430	GenBank
Mollusca	*Lottia gigantea*	GCM, XP_009048823	GenBank
	Biomphalaria glabrata	GCM, XP_013078709	GenBank
	Octopus bimaculoides	GCM, XP_014777859	GenBank
	Crassostrea gigas	GCM, EKC24558	GenBank

Appendix 2A. Species used for *Gcm* phylogenetic analysis

Annelida	*Platynereis dumerilii*	GCM, CCK33024	GenBank
Platyhelminthes	*Echinococcus granulosus*	GCM, CDS20483	GenBank
Brachiopoda	*Lingula anatina*	GCM, XP_013399310	GenBank
Echinodermata	*Echinocardium cordatum*	GCM, BAS66823	GenBank
Tardigrada	*Ramazzottius varieornatus*	GCM, GAU98972	GenBank
Cnidaria	*Acropora digitifera*	GCM, XP_015774704	GenBank
	Exaiptasia pallida	GCM, KXJ08668	GenBank
	Nematostella vectensis	GCM, EDO25565	GenBank
Fungi	*Coprinopsis cinerea okayama*	Hypothetical protein, EAU87004	GenBank
	Melampsora larici-populina	Hypothetical protein, EGG12299	GenBank

Appendix 2B. Alignment of *Gcm* domain. Created with JalView Version 2 (Waterhouse et al., 2009)



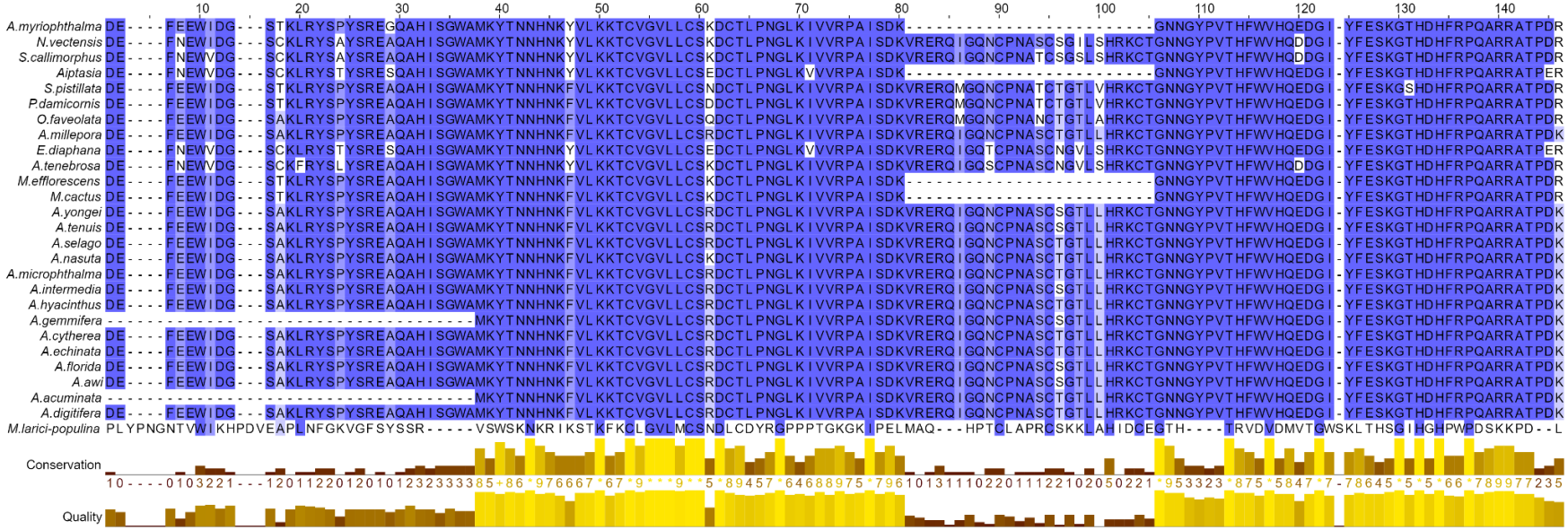
Appendix 3A. Cnidarian species used for *Gcm* phylogenetic analysis

Species	Accession ID	Source
Acropora acuminata	aacu_s0062.g65.t1	OIST marine genome browser
Acropora awi	aawi_s0172.g23.t1	OIST marine genome browser
Acropora cytherea	acyt_s0052.g133.t1	OIST marine genome browser
Acropora digitifera	adig_s0165.g12.t1	OIST marine genome browser
Acropora echinata	aech_s0121.g25.t1	OIST marine genome browser
Acropora florida	aflo_s0478.g2.t1	OIST marine genome browser
Acropora gemmifera	agem_s0269.g18.t1	OIST marine genome browser
Acropora hyacinthus	ahya_s0011.g87.t1	OIST marine genome browser
Acropora intermedia	aint_s0197.g10.t1	OIST marine genome browser
Acropora millepora	XP_029180379	GenBank
Acropora muricata	amur_s0056.g15.t1	OIST marine genome browser
Acropora myriophthalma	amic_s0106.g13.t1	OIST marine genome browser
Acropora nasuta	anas_s0200.g8.t1	OIST marine genome browser
Acropora selago	asel_s0087.g22.t1	OIST marine genome browser
Acropora tenuis	aten_s0286.g9.t1	OIST marine genome browser
Acropora yongei	ayon_s0143.g10.t1	OIST marine genome browser

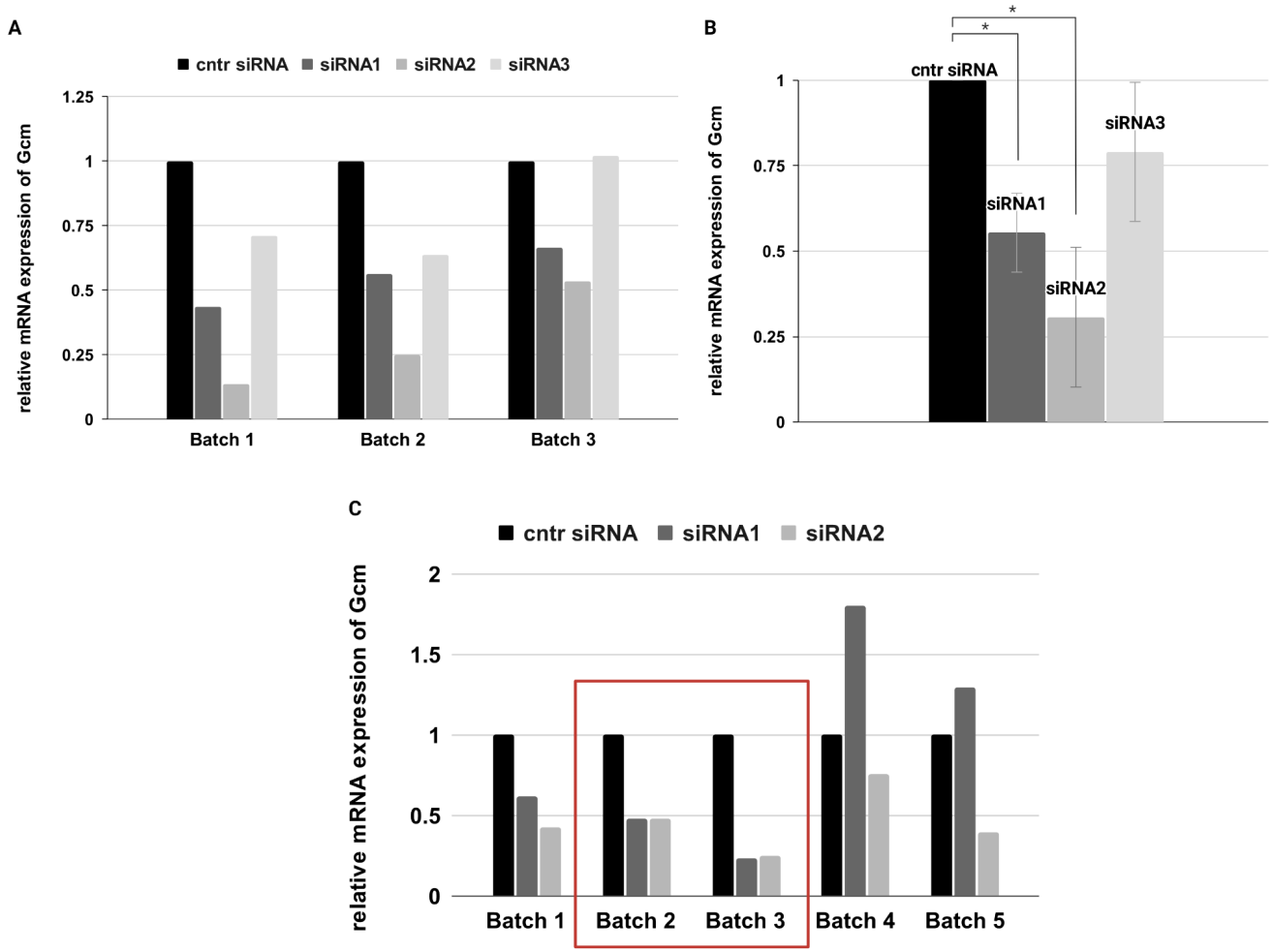
Appendix 3A. Cnidarian species used for *Gcm* phylogenetic analysis

Actinia tenebrosa	XP_031557567	GenBank
Aiptasia	AIPGENE4948 sp|Q9VLA2|GCM2	Reef Genomics database
Nematostella vectensis	EDO25565	GenBank
Exaiptasia diaphana	KXJ08668	GenBank
Scolanthus callimorphus	N/A	SIMRbase
Montipora cactus	mcac_s0146.g32.t1	OIST marine genome browser
Montipora efflorescens	meff_s0379.g17.t1	OIST marine genome browser
Orbicella faveolata	XP_020628548	GenBank
Pocillopora damicornis	pdam_00014897-RA	Reef Genomics database
Stylophora pistillata	SpisGene9252	Reef Genomics database

Appendix 3B. Alignment of *Gcm* domain in cnidarians. Created with JalView Version 2 (Waterhouse et al., 2009).



Appendix 4. KD efficiency of *Gcm*-specific siRNAs. (A) *Gcm* KD efficiency of siRNA1, siRNA2, siRNA3 in three experimental batches (B) Average siRNA1 and siRNA2 KD efficiency is ~50% and ~75% respectively. siRNA3 is not efficient at knocking down *Gcm*. Data are given mean ±SD, N=3, *p < 0.05, unpaired t-test. (C) KD efficiency of *Gcm*-specific siRNA1 and siRNA2 in batches used for RNA-seq. Batch 2 and Batch 3 (framed in red) were selected for DE gene analysis. Assembled with BioRender.com



Appendix 5. Sequencing depths and alignment rates of the RNA-seq libraries. Libraries used for obtaining the final results are highlighted in yellow.

Sample Name	Sample ID	Total reads (millions)	Alignment rate
HTKK5DSXX_PR1396_01A8080_H1_L002_R1	siRNA1_1batch_1	79.3	76%
HTKK5DSXX_PR1396_01A8080_H1_L002_R2	siRNA1_1batch_2	79.3	
HTKK5DSXX_PR1396_02A8181_H1_L002_R1	siRNA2_1batch_1	69.7	77%
HTKK5DSXX_PR1396_02A8181_H1_L002_R2	siRNA2_1batch_2	69.7	
HTKK5DSXX_PR1396_03A8282_H1_L002_R1	control_1batch_1	56.1	77%
HTKK5DSXX_PR1396_03A8282_H1_L002_R2	control_1batch_2	56.1	
HTKK5DSXX_PR1396_04A8383_H1_L002_R1	siRNA1_2batch_1	51.8	76%
HTKK5DSXX_PR1396_04A8383_H1_L002_R2	siRNA1_2batch_2	51.8	
HTKK5DSXX_PR1396_05A8484_H1_L002_R1	siRNA2_2batch_1	76.7	78%
HTKK5DSXX_PR1396_05A8484_H1_L002_R2	siRNA2_2batch_2	76.7	
HTKK5DSXX_PR1396_06A8585_H1_L002_R1	control_2batch_1	58.8	76%
HTKK5DSXX_PR1396_06A8585_H1_L002_R2	control_2batch_2	58.8	
HTKK5DSXX_PR1396_07A8686_H1_L002_R1	siRNA1_3batch_1	54.5	76%
HTKK5DSXX_PR1396_07A8686_H1_L002_R2	siRNA1_3batch_2	54.5	
HTKK5DSXX_PR1396_08A8787_H1_L002_R1	siRNA2_3batch_1	54.3	76%

Appendix 5. Sequencing depths and alignment rates of the RNA-seq libraries.

HTKK5DSXX_PR1396_08A8787_H1_L002_R2	siRNA2_3batch_2	54.3	
HTKK5DSXX_PR1396_09A8989_H1_L002_R1	control_3batch_1	70.1	75%
HTKK5DSXX_PR1396_09A8989_H1_L002_R2	control_3batch_2	70.1	
HTKK5DSXX_PR1396_10A9191_H1_L002_R1	siRNA1_4batch_1	68.1	76%
HTKK5DSXX_PR1396_10A9191_H1_L002_R2	siRNA1_4batch_2	68.1	
HTKK5DSXX_PR1396_11A9292_H1_L002_R1	siRNA2_4batch_1	64.6	76%
HTKK5DSXX_PR1396_11A9292_H1_L002_R2	siRNA2_4batch_2	64.6	
HTKK5DSXX_PR1396_12A9393_H1_L002_R1	control_4batch_1	57.6	77%
HTKK5DSXX_PR1396_12A9393_H1_L002_R2	control_4batch_2	57.6	
HTKK5DSXX_PR1396_13A9494_H1_L002_R1	siRNA1_5batch_1	50.3	77%
HTKK5DSXX_PR1396_13A9494_H1_L002_R2	siRNA1_5batch_2	50.3	
HTKK5DSXX_PR1396_14A9595_H1_L002_R1	siRNA2_5batch_1	59.1	78%
HTKK5DSXX_PR1396_14A9595_H1_L002_R2	siRNA2_5batch_2	59.1	
HTKK5DSXX_PR1396_15A9696_H1_L002_R1	control_5batch_1	55.4	77%
HTKK5DSXX_PR1396_15A9696_H1_L002_R2	control_5batch_2	55.4	
	mean	61.76	76%
	median	58.8	76%

Appendix 6. Primer sequences and expected product sizes for genes used for RT-qPCR validation

Gene symbol	Gene ID	Primer sequence 5'-3'	Expected product size (bp)
Gcm	NVE12024	Forward:CGACCCGCGATATCAGATAAA Reverse:CTGTGCATTTCCTGTGAGAGA	97
EAAT1	NVE443	Forward: ACTCGAAGGAACCCGTTGAC Reverse: GAATGGTGCATGGAAGGGTG	100
GAT2	NVE14046	Forward: TTCTCGCCCTCTCCTCCATT Reverse:GTGGTTACGCTCCTCAGACG	100
GAT3	NVE16531	Forward: ACTCGGCAAAGAAGTCTCGG Reverse: CCAGAAGGGTGAAATCGGCA	100
18S	NVE20421	Forward: CTTAGATGTTCTGGGCCGCA Reverse: CCAGCACGACGATGTTTCAC	100
Gapdh	NVE23813	Forward: GGATGGACCAAGTGCCAAGAAC Reverse: GCTTGCCGTTTACCTCAGGAATGA	119
Ef1a	NVE12051	Forward: GGTTGCCTCTTCGCTTACCACT Reverse: CGTTCCTGGCTTTAGGACAC	101
SORL1	NVE13564	Forward: TGTGACCGACCAACCATCTG Reverse: GATGTGCCAGTCGTCGATCA	100
PTPkappa	NVE25055	Forward: AACTCAGCAAACACCACCCG Reverse: ACCACCCGGGAGTGATCATA	100
CASPR2	NVE830	Forward: CCGCCTGGATCAGTAAAGGG Reverse: GTCTTGCCCTTGGTCGTGAT	100

Appendix 6. Primer sequences and expected product sizes for genes selected for qPCR validation

ZICA	NVE8653	Forward: TGCCGCGGATTGATCGTTAT Reverse: AACATCGCGTTGGATGATGG	100
PKD1L2	NVE22455	Forward: CGGTTGATTCTCCAGAGGCT Reverse: TCGTACTCGAACGGCTTTCC	100
ADGRB1	NVE10004	Forward: TCGTTGTCGTGGCAACTTTG Reverse: CTGACCGCACAATTGACAGC	100
TRPA1	NVE1653	Forward: GTAGAGGCCTTGACTCGAGC Reverse: CCACCTTCCTGAAGCCACTG	100
Notch1-like	NVE8334	Forward: CCCTAGCGGAATCTCTTGGC Reverse: GGTCCCTGTGTGATGGTACC	100
SGLT	NVE18073	Forward: TGGCGGTATGTGGAAGTGAG Reverse: ACAGCCATCATCAGACCACG	100
SDK1	NVE20069	Forward: CGCCTTTCTTGTCAAGTCGC Reverse: CATCAGCACCGTCACTGCTA	100
ADGRV1	NVE12193	Forward: CAAGCAAGTCAACTGTTCGGA Reverse: GGCAGTGTACGCAGAAAATTG	119
EPR1	NVE17236	Forward: GCAAAGCTGTTCCGGACAAG Reverse: TACCACAGACCGACGCAATC	100
P2RY6-like	NVE18731	Forward: CAACCCTGATGTCCATGCCA Reverse: GAGGGCATCGAGTGAAAGGT	100
Contactin/DSN CAM	NVE12757	Forward: GTCAGCTGCGTTCATTGAGC Reverse: ACCACGTCTAGAAACCCTGC	100

Appendix 6. Primer sequences and expected product sizes for genes selected for qPCR validation

LRRC8A	NVE23087	Forward:ACTGCGCTTGGACATATAAGA Reverse: TAGAGACGCGTACCGGGTTA	107
Delta/Notch-like	NVE18894	Forward: CTGGAGTGGTAAACCCTGCC Reverse: AAGATGACAAACAGGCGGTCT	100
Ets-related	NVE9883	Forward: GTGCATTGCTGAAATTGCTTCG Reverse: GTGATTGGTTGATCGGTTGCC	91
PLA2_1	NVE23561	Forward: CCAAGAGGACGCTATGAGGC Reverse: TGATCGCTTCGCCACTTCTT	96
EDIL3	NVE20550	Forward: ATAACGAAACGGGCTGGTGC Reverse: ATGCGTGAGGCTGAACAGTAA	107
OGT	NVE5399	Forward: CCATTGAGACCCAGCCTAGC Reverse: AAGTGGTGTATTGCCAGCCA	93
PKM2	NVE38	Forward: GCACCTTCGTTTTCGTGTCC Reverse: GCGCGATGTGGATCTTCTCT	95
MPEG-1	NVE3218	Forward: GGTTCGGTAGTGTGAGTGGG Reverse: TGTCGTAATTGCACACGGGT	100
PLA2_2	NVE2264	Forward: GCGGGGTTAGGTTAATTGTCG Reverse: TTGCAACCATTTGGCTGTCC	94
PRGamide	NVE226	Forward: GCAGGTCCCTTATTGAGCTTC Reverse: GTCCGACTTCTCAGCAGACC	149

Appendix 7. GO terms and function (UniProt) of DE genes grouped by p-adj.

P-adj< 0.05, downregulated			
Nv_vienna_ID	**NCBI_accession**	**Function /BLAST hit**	**GO/KEGG pathway**
NVE38	XM_001641976	pyruvate kinase PKM. PKM2 inhibition increases substrate flux through the pentose phosphate pathway to generate reducing equivalents (NADPH and GSH) and protect against oxidative stress.	carbohydrate/energy metabolism
NVE22599	XM_032367192	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific
NVE21209	XM_032376191	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific
NVE11899	XM_001641371	uncharactarized, blast hit - [Acropora] TNF receptor-associated factor 5-like	uncharacterized/cnidaria-specific
NVE6277	XM_032376361	collagen alpha-4 chain, integrin-like	ECM/cell adhesion
NVE13045	XM_032366926	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific
NVE14657	XM_032376191	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific
NVE12839	XM_032366921	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific
NVE17933	XM_032370086	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific
NVE7172	XM_032371668	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific
NVE12840	XM_032366924	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific
NVE13565	XR_004291976	ncRNA	ncRNA
NVE24747	XR_004296070	misc_RNA	misc_RNA
NVE4321	XM_032377551	supwaprin-a (antimicrobial)	immune response
NVE13564	XM_032387386	sortilin-related receptor, neuronal apolipoprotein E receptor, the gene for which is predominantly expressed in the central nervous system	lipid metabolism
NVE22531	XM_032375707	uncharactarized, blastp - no hit	uncharacterized/cnidaria-specific

NVE25055	XM_032368708	receptor-type tyrosine-protein phosphatase kappa, blast hit - PTPRD (Receptor-type tyrosine-protein phosphatase delta), can bidirectionally induce pre- and post-synaptic differentiation of neurons by mediating interaction with IL1RAP and IL1RAPL1 trans-synaptically. Involved in pre-synaptic differentiation through interaction with SLITRK2.	protein modification
NVE831	XM_032380069	blast hit - mucin-3A-like isoform X2 [Stylophora pistillata], major glycoprotein component of a variety of mucus gels. Thought to provide a protective, lubricating barrier against particles and infectious agents at mucosal surfaces. May be involved in ligand binding and intracellular signaling.	TF/transcription
NVE24010	none	uncharacterized, blastp - no hit	uncharacterized/cnidaria-specific
NVE24376	XM_032381802	UBX protein domain is found in ubiquitin-regulatory proteins, which are members of the ubiquitination pathway, as well as a number of other ubiquitin-like proteins including FAF-1 (FAS-associated factor 1), required for the progression of DNA replication forks by targeting DNA replication licensing factor CDT1 for degradation	protein modification
NVE16319	XM_032366645	E3 ubiquitin-protein ligase TM129, transmembrane protein, involved in ER-associated protein degradation	protein degradation
NVE3633	XM_032366195	uncharactarized, blast hit - Intraflagellar transport protein 172 (coral)	uncharacterized/cnidaria-specific
NVE22208	XM_032369057	uncharactarized, blast - no hit	uncharacterized/cnidaria-specific
NVE3634	none	uncharactarized, blast hits - hypothetical proteins in corals	uncharacterized/cnidaria-specific
NVE3381	XR_004294124	ncRNA	ncRNA
NVE14629	XM_032371431	uncharactarized, blast - hypothetical proteins in corals	uncharacterized/cnidaria-specific
NVE5822	XM_032367471	uncharactarized, blast hit - E3 ubiquitin-protein ligase DZIP3 [Exaiptasia pallida], mediates ubiquitination and subsequent proteasomal degradation of target proteins	protein degradation
NVE2774	XM_032375707	uncharactarized, blast - no hit	uncharacterized/cnidaria-specific

NVE830	XM_032380063	contactin-associated protein-like 2 (CNTNAP2)	ECM/cell adhesion
NVE18730	XM_032384910	blast hit - Polycystic kidney disease protein 1-like 3 [Exaiptasia pallida], Component of a calcium channel. May act as a sour taste receptor by forming a calcium channel with PKD1L3 in gustatory cells	Membrane transport
NVE15740	XM_032372064	uncharactarized, blast hits - hypothetical proteins from corals and molluscs	uncharacterized/cnidaria-specific
NVE2552	XM_032368045	Integrase zinc core domain, blast hit - TNF receptor-associated factor 3 [Stylophora pistillata] an essential constituent of several E3 ubiquitin-protein ligase complexes. May have E3 ubiquitin-protein ligase activity and promote 'Lys-63'-linked ubiquitination of target proteins.	Protein synthesis/processing/ degradation
NVE8261	XM_032366377	Neurogenic locus notch homolog protein 1 and Notch 2	Notch signaling
NVE10736	XM_032372378	blast hit - Transposon Ty3-G Gag-Pol polyprotein [Exaiptasia pallida]	uncharacterized/cnidaria-specific
NVE141	XM_032367644	ADGRL4, adhesion G protein-coupled receptor L4, Endothelial orphan receptor that acts as a key regulator of angiogenesis	ECM/Cell adhesion
NVE13473	XM_032367885	Replication factor C subunit 1. The elongation of primed DNA templates by DNA polymerase delta and epsilon requires the action of the accessory proteins PCNA and activator 1. This subunit binds to the primer-template junction. Binds the PO-B transcription element as well as other GA rich DNA sequences. Could play a role in DNA transcription regulation as well as DNA replication and/or repair.	DNA replication/repair
NVE8653	AB231867	TF, zinc finger protein ZIC4 (Zinc finger protein of the cerebellum 4) is involved in neurogenesis	TFs/transcription

NVE5670	XM_032376697	sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1 (SVEP1), May play a role in the cell attachment process, multi-domain cell adhesion protein, homologous to the mouse polydom protein, which has been shown to mediate cell-cell adhesion in an integrin dependent-manner in osteogenic cells	ECM/Cell adhesion
NVE16480	XM_032382916	blast hit - E3 ubiquitin-protein ligase DZIP3 [Exaiptasia pallida]	protein degradation
NVE26003	XM_032380640	uncharacterized	uncharacterized/cnidaria-specific
NVE12641	XM_032372363	vicilin-like seed storage protein At2g18540, Seed storage protein.	TFs/transcription
NVE13485	XM_032372265	uncharacterized, blast hit - ATP-binding cassette sub-family D member 3 [Labeo rohita]	uncharacterized/cnidaria-specific
NVE18507	XM_032366926	uncharactarized, blastp hit - none	uncharacterized/cnidaria-specific
NVE13487	XM_032375749	uncharacterized, blast hit - hypothetical protein [Nostoc sp. T09]	uncharacterized/cnidaria-specific
NVE7065	XM_001627778	uncharacterized, TSP1 superfamily domain, blast hit - cell surface glycoprotein 1 [Exaiptasia pallida]	ECM/cell adhesion
NVE7280	XM_032375568	arginine and glutamate-rich protein 1-A, is required for the oestrogen-dependent expression of ESR1 target genes. It functions in cooperation with MED1	TFs/transcription
NVE18121	XR_004295795	ncRNA	ncRNA
NVE9285	none	uncharacterized, blast - no hit	uncharacterized/cnidaria-specific
NVE20550	XM_032369412	EGF-like repeat and discoidin I-like domain-containing protein 3. EDIL3 is a glycoprotein that is secreted by endothelial cells and can associate with the endothelial cell surface and the extracellular matrix.	ECM/cell adhesion
NVE10760	XM_032373473	cell wall protein DAN4-like, DAN4 in yeast is a component of the cell wall	membrane transport
NVE13391	XM_032372047	uncharacterized, Corresponds to Merops family A17. These proteins are homologous to aspartic proteinases encoded by retroposons and retroviruses.	RT/retrotransposons

NVE8561	XM_032374855	organic cation transporter protein (OCTN1). An organic cation transport protein mediates the transport of organic cations across the cell membrane. These proteins are members of the solute carrier family, subfamily 22.	membrane transport
NVE18120	XR_004296913, XM_032385195	ncRNA	ncRNA
NVE444	XM_001631508	YTH domain-containing protein 1, Regulator of alternative splicing that specifically recognizes and binds N6-methyladenosine (m6A)-containing RNAs, directs different complexes to regulate RNA signaling pathways, such as RNA metabolism, RNA splicing, RNA folding, and protein translation.	RNA splicing
NVE9694	XM_032374129	peroxisomal N(1)-acetyl-spermine/spermidine oxidase, Flavoenzyme which catalyzes the oxidation of N1-acetylspermine to spermidine and is thus involved in the polyamine back-conversion	lipid metabolism
NVE18446	XM_032375707	uncharactarized, blast - no hit	uncharacterized/cnidaria-specific
NVE3596	XM_032377969	Sorting nexin-13, RGS-PX1 is a member of both the regulator of G protein signaling (RGS) and sorting nexin (SNX) protein families.	Protein synthesis/processing/ degradation
NVE22605	XM_032376127	blast hit - extensin-like mRNA, extensins are a family of flexuous, rodlike, hydroxyproline-rich glycoproteins (HRGPs) of the plant cell wall or zonadhesin (salmon match). This gene encodes a protein that functions in the species specificity of sperm adhesion to the egg zona pellucida. The encoded protein is located in the acrosome and may be involved in signaling or gamete recognition.	TFs/transcription
NVE10366	XM_032362710	uncharacterized, blast hits - uncharacterized proteins in cnidaria and molluscs	uncharacterized/cnidaria-specific
NVE12411	XM_032372478	bifunctional polynucleotide phosphatase/kinase, Plays a key role in the repair of DNA damage, functioning as part of both the non-homologous end-joining (NHEJ) and base excision repair (BER) pathways.	DNA replication/repair

NVE24235	XM_001636600	DNA replication ATP-dependent helicase/nuclease DNA2, Key enzyme involved in DNA replication and DNA repair in nucleus and mitochondria.	DNA replication/repair
NVE1600	XR_004293153	ncRNA	ncRNA
NVE7392	XM_001620713	cilia- and flagella-associated protein 74, as part of the central apparatus of the cilium axoneme may play a role in cilium movement.	cell motility
NVE18180	XM_032371133	adhesin BpaC-like, cell adhesion protein identified in bacteria	ECM/cell adhesion
NVE12293	XM_032372590	blast hit - cilia- and flagella-associated protein 99 [Delphinapterus leucas]	cell motility
NVE9769	XM_032367534	uncharacterized, blast hit - Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A [Exaiptasia pallida], Catalyzes the addition of N-acetylglucosamine (GlcNAc) in beta 1-6 linkage to the alpha-linked mannose of biantennary N-linked oligosaccharides, Via its role in the biosynthesis of complex N-glycans, plays an important role in the activation of cellular signaling pathways, reorganization of the actin cytoskeleton, cell-cell adhesion and cell migration.	ECM/cell adhesion
NVE22455	XM_032365909	polycystic kidney disease protein 1-like 2, may function as an ion-channel regulator, involved in cell-cell/matrix interactions. May function as a G-protein-coupled receptor.	membrane transport
NVE24871	XM_032381419	Cartilage matrix protein-like, is a major component of the extracellular matrix of non-articular cartilage. It binds to collagen.	ECM/cell adhesion
NVE12227	XM_032386874	uncharacterized	uncharacterized/cnidaria-specific
NVE7553	XM_032387267	reverse transcriptase	RT/retrotransposons
NVE7577	XM_032375405	heterogeneous nuclear ribonucleoprotein L-like (HNRNPLL), RNA-binding protein that functions as regulator of alternative splicing for multiple target mRNAs	RNA splicing
NVE21596	XM_032383320	probable ATP-dependent RNA helicase DDX10 (RNA helicase).	RNA splicing
NVE19901	XR_004292618	misc_RNA	misc_RNA

NVE13277	XM_032372150	potassium channel subfamily T member 1 (voltage-gated ion channels), outwardly rectifying potassium channel subunit that may coassemble with other Slo-type channel subunits. Activated by high intracellular sodium or chloride levels. Activated upon stimulation of G-protein coupled receptors, such as CHRM1 and GRIA1.	membrane transport
NVE24988	XR_004295998	misc_RNA	misc_RNA
NVE17173	XM_032369781	von Willebrand factor D and EGF domain-containing protein (calcium ion binding)	ECM/cell adhesion
NVE16596	XR_004297107	ncRNA, blast hit - hypothetical proteins (Cnidaria)	ncRNA
NVE7066	XM_001627778	same as NVE7065	ECM/cell adhesion
NVE4276	XM_032362635	arginine and glutamate-rich protein 1-A (ARGLU1)	TFs/transcription
NVE20392	XM_032369454	zinc finger C3H1 domain-containing protein	TFs/transcription
NVE14680	XR_004295795	ncRNA, miscRNA	ncRNA
NVE16225	XM_001623467	bZIP transcription factor-like	TFs/transcription
NVE8051	XM_032367102	sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1 (SVEP1), May play a role in the cell attachment process.	ECM/cell adhesion
NVE7655	XM_032375303	predicted, blast hits - uncharacterized proteins in Chidaria, REJ domain - domain is found in PKD1 and the sperm receptor for egg jelly. The function of this domain is unknown. IPT/TIG domain. This family consists of a domain that has an immunoglobulin-like fold. These domains are found in cell surface receptors such as Met and Ron as well as in intracellular transcription factors where it is involved in DNA binding.	membrane transport
p-adj < 0.05, upregulated			
NVE17843	XR_004296956	miscRNA	miscRNA
NVE14448	XM_032371523	uncharacterized	uncharacterized/cnidaria-specific
NVE5252	XM_032368149	uncharacterized, blast hit- galactose-specific lectin [Anthopleura japonica]	uncharacterized/cnidaria-specific
NVE14983	XM_032371282	60S ribosomal protein L15	ribosomal

NVE4717	XM_032364190	uncharacterized, blast hit - Pancreatic secretory granule membrane major glycoprotein GP2 [homo sapiens], LZP variant II [Mus musculus], Zona pellucida (ZP) domain has been recognized in a number of receptor-like eukaryotic glycoproteins, which is involved in many important biological processes, such as signal transduction, development, differentiation.	ECM/cell adhesion
NVE10004	XM_001626357	adhesion G protein-coupled receptor B1 (ADGRB1 or BAI1), Phosphatidylserine receptor which enhances the engulfment of apoptotic cells.	ECM/cell adhesion
NVE3218	XM_032378343	macrophage-expressed gene 1 protein, Macrophage Expressed Gene-1 (MPEG-1; also termed Perforin-2) is an endosomal / phagolysosomal perforin-like protein that is conserved across the metazoan kingdom and that functions within the phagolysosome to damage engulfed microbes. Plays a key role in the innate immune response following bacterial infection by inserting into the bacterial surface to form pores	immune response
NVE23561	XM_032382138	PLA2_1 phospholipase A2 A2-actitoxin-Cgg2a, involved in Lipid metabolism (cleaves fatty acids), phospholipase A2 (PLA2), membrane associated precursor [Homo sapiens], Ca2+-requiring, secretory enzymes that have been implicated in a number of biological processes, such as modification of eicosanoid generation, inflammation, host defense, and atherosclerosis.	lipid metabolism
NVE1644	XM_032379490	uncharacterized	uncharacterized/cnidaria-specific
NVE12742	XM_032385880	uncharacterized	uncharacterized/cnidaria-specific
NVE2195	XM_032364690	blast - galaxin-like, Galaxin, which encodes for a matrix protein suspected to be involved in calcification, is originally identified from the coral. It is hypothesized that the coral galaxin homologs are recruited as biomineralization proteins when Scleractinia diverged from non-biomineralizing taxa during the Triassic56 because galaxin-like proteins are also found in non-calcifying taxa outside Cnidaria	uncharacterized/cnidaria-specific
NVE9883	XM_032362934.1	uncharactarized, blast hit - ETS-related transcription factor Elf-1	TFs/transcription
NVE23912	XM_032365681	uncharactarized	uncharacterized/cnidaria-specific

p-adj < 0.1, downregulated			
NVE2938	XM_001639423	calcineurin-binding protein cabin-1. May be required for replication-independent chromatin assembly. May serve as a negative regulator of T-cell receptor (TCR) signaling via inhibition of calcineurin. Inhibition of activated calcineurin is dependent on both PKC and calcium signals. Acts as a negative regulator of p53/TP53 by keeping p53 in an inactive state on chromatin at promoters of a subset of it's target genes.	TFs/transcription
NVE18895	XM_032366377	neurogenic locus notch homolog protein 1	Notch signaling
NVE23097	XR_004295937	misc_RNA	misc_RNA
NVE21821	XM_032362584	bromodomain-containing protein 7 - is a transcription regulator that is normally highly expressed in the testis, Acts both as coactivator and as corepressor. May play a role in chromatin remodeling. Activator of the Wnt signaling pathway in a DVL1-dependent manner by negatively regulating the GSK3B phosphotransferase activity.	TFs/transcription
NVE2281	XM_032379039	uncharacterized, blast hit - transient receptor potential protein 5 [Mus musculus, homo sapiens], receptor-activated non-selective calcium permeant cation channel, transient receptor potential 5 TRPC5 (also TrpC5, trp-5 or trp5) is one of the seven mammalian TRPC proteins. Its known functional property is that of a mixed cationic plasma membrane channel with calcium permeability.	membrane transport
NVE5335	DQ518750, BR000683 , XM_032362962	Mox gene cluster, gene for opsin (match - melanopsin [mus musculus]), Photoreceptor required for regulation of circadian rhythm. Contributes to pupillar reflex and other non-image forming responses to light. May be able to isomerize covalently bound all-trans retinal back to 11-cis retinal.	TFs/transcription
NVE25263	XM_001640126	proton channel OtopLc. Proton-selective channel that specifically transports protons into cells. Proton-selective channel activity is probably required in cell types that use changes in intracellular pH for cell signaling or to regulate biochemical or developmental processes.	membrane transport
NVE1719	XM_032379414	piwi-like protein 1, Endoribonuclease that plays a central role in postnatal germ cells by repressing transposable elements and preventing their mobilization, which is essential for the germline integrity.	RT/retrotransposons
NVE3923	XR_004293389	ncRNA	ncRNA
NVE13370	XM_032372074	RT-like family, DUF5641 (protein domain family of unknown function).	RT/retrotransposons

NVE13426	XM_001624701	formin-binding protein 4, protein containing two tryptophan-rich WW domains that binds the proline-rich formin homology 1 domains of formin family proteins, suggesting a role in the regulation of cytoskeletal dynamics during cell division and migration.	cell motility
NVE1653	XM_032379482	transient receptor potential cation channel subfamily A member 1, Receptor-activated non-selective cation channel involved in pain detection and possibly also in cold perception, oxygen concentration perception, cough, itch, and inner ear function.	membrane transport
NVE24012	XM_032365673	leukocyte receptor cluster member 8 (LENG8), signal transduction, genes involved in blood-brain barrier (BBB) formation and integrity (Schiavone et al, 2017), ubiquitous expression	membrane transport
NVE25595	XM_032380977	tetratricopeptide repeat protein 12, The TTC12 gene encodes a protein containing several domains known to mediate protein-protein interactions and the assembly of multiprotein complexes. TTC12 is believed to play a role in the assembly of dynein arms in motile cilia in both human respiratory cells and sperm flagella	Protein synthesis/processing/degradation
NVE5337	XM_001628627	uncharacterized, blast hits - hypothetical proteins	uncharacterized/cnidaria-specific
NVE14627	XM_032371433	uncharacterized, blast hits - hypothetical proteins	uncharacterized/cnidaria-specific
NVE4725	XM_032364135	uncharacterized, RT, retrotransposon peptidase, zinc binding domain, retrotransposon	RT/retrotransposons
NVE24441	XM_032385326	uncharacterized, blast hits - hypothetical proteins	uncharacterized/cnidaria-specific
NVE9889	XM_032362858	DNA excision repair protein ERCC-6-like 2. This gene encodes a member of the Snf2 family of helicase-like proteins. The encoded protein may play a role in DNA repair and mitochondrial function.	DNA replication/repair
NVE1830	XM_032366645	E3 ubiquitin-protein ligase TM129, same as NVE16319	Protein synthesis/processing/degradation
NVE21065	XM_032366926	uncharacterized	uncharacterized/cnidaria-specific
NVE25845	XM_032375707	uncharacterized, same as NVE22531	uncharacterized/cnidaria-specific
NVE25715	XM_032366922	uncharacterized, blast hit - LysM peptidoglycan-binding domain-containing protein [Moraxella sp.]	uncharacterized/cnidaria-specific

NVE8970	XM_032386874	uncharacterized, blast hit - zinc-finger DNA-binding protein [Operophtera brumata]	TFs/transcription
NVE10940	XM_032373417	adhesion G-protein coupled receptor D1-like (ADGRD1). Orphan receptor. Signals via G(s)-alpha family of G-proteins (PubMed:22025619, PubMed:22575658). Has protumorigenic function especially in glioblastoma (PubMed:27775701). Plays a role in cell-cell adhesion and neuron guidance via its interactions with FLRT2 and FLRT3 that are expressed at the surface of adjacent cells (PubMed:26235030).	ECM/cell adhesion
NVE20529	XM_032373468	uncharacterized, protein unc-13 homolog C-like [Pocillopora damicornis]. May play a role in vesicle maturation during exocytosis as a target of the diacylglycerol second messenger pathway.	endocytosis/exocytosis
NVE7559	XM_032370283	uncharacterized, blast - no hits	uncharacterized/cnidaria-specific
NVE23107	XM_032381141	uncharacterized, blast - no hits	uncharacterized/cnidaria-specific
p-adj < 0.1, upregulated			
NVE8459	XM_032374949	glycine-rich RNA-binding protein 7, Plays a role in RNA transcription or processing during stress. Binds RNAs and DNAs sequence with a preference to single-stranded nucleic acids. Displays strong affinity to poly(U) and poly(G) sequence. Involved in mRNA alternative splicing of numerous targets by modulating splice site selection.	RNA splicing
NVE13796	XM_032371808	polycystic kidney disease protein 1-like 2, May function as an ion-channel regulator. May function as a G-protein-coupled receptor.	membrane transport
NVE3046	XM_001639565	glutamic acid-rich protein	TFs/transcription
p-adj < 0.2, upregulated			
NVE8526	XM_001627177	uncharacterized, blast - no hits	uncharacterized/cnidaria-specific
NVE17842	XM_032385402	uncharacterized, blast - no hits	uncharacterized/cnidaria-specific
NVE14450	XM_032371522	uncharacterized, galactose-specific lectin [Anthopleura japonica]	ECM/cell adhesion
NVE8043	XM_001641744	40S ribosomal protein S23	ribosome

NVE11578	XM_001620758	40S ribosomal protein S9	ribosome
NVE9450	XM_001638285	60S ribosomal protein L30	ribosome
NVE12548	XM_032361981	tetratricopeptide repeat protein 28-like, TTC28, During mitosis, may be involved in the condensation of spindle midzone microtubules, leading to the formation of midbody	Protein synthesis/processing/degradation
NVE2256	XM_001630433	60S ribosomal protein L34	ribosome
NVE2295	XM_001630427	60S ribosomal protein L42-A	ribosome
NVE7522	XM_032363570	calponin-1, Calponin 1 is a basic smooth muscle protein, The expression of CNN1 is specific to differentiated mature smooth muscle cells, suggesting a role in contractile functions.	cell motility
NVE8334	XM_032363335	neurogenic locus notch homolog protein 1-like	Notch signaling
NVE21167	XM_001634346	LWamide neuropeptides	Neuroactive ligand-receptor
NVE6169	XM_032363930	uncharacterized, blast - Fert2 protein, partial [Mus musculus], Tyrosine-protein kinase that acts downstream of cell surface receptors for growth factors and plays a role in the regulation of the actin cytoskeleton, microtubule assembly, lamellipodia formation, cell adhesion, cell migration and chemotaxis. Plays a role in synapse organization, trafficking of synaptic vesicles, the generation of excitatory postsynaptic currents and neuron-neuron synaptic transmission.	cell motility
NVE9019	XM_032374594	B-cell lymphoma 3 protein (BCL3) Contributes to the regulation of transcriptional activation of NF-kappa-B target genes. In the cytoplasm, inhibits the nuclear translocation of the NF-kappa-B p50 subunit. In the nucleus, acts as a transcriptional activator that promotes transcription of NF-kappa-B target genes. Contributes to the regulation of cell proliferation	NF-kappa B
NVE2142	XM_032364650	beta-ketoacyl-ACP synthase 3 [Nocardioides sp. SLBN-35], Beta-ketoacyl-ACP synthase is found as a component of a number of enzymatic systems, including fatty acid synthetase (FAS)	lipid metabolism

NVE7499	XM_032363508	TNFAIP3-interacting protein 3-like, Binds to zinc finger protein TNFAIP3 and inhibits NF-kappa-B activation induced by tumor necrosis factor, Toll-like receptor 4 (TLR4), interleukin-1 and 12-O-tetradecanoylphorbol-13-acetate. Overexpression inhibits NF-kappa-B-dependent gene expression in response to lipopolysaccharide at a level downstream of TRAF6 and upstream of IKBKB. NF-kappa-B inhibition is independent of TNFAIP3 binding.	NF-kappa B
p-adj < 0.2, downregulated			
NVE18073	XM_032385202	sodium/glucose cotransporter 5, High capacity transporter for mannose and fructose and, to a lesser extent, glucose, AMG, and galactose. found in the intestinal mucosa (enterocytes) of the small intestine (SGLT1)	carbohydrate/energy metabolism
NVE10425	XM_032362651	KH homology domain-containing protein 4 - RNA-binding protein involved in pre-mRNA splicing.	RNA splicing
NVE10681	XM_032370200	uncharacterized, blast hits - hypothetical proteins (cnidaria)	uncharacterized/cnidaria-specific
NVE16938	XM_032362591	uncharacterized	uncharacterized/cnidaria-specific
NVE5342	XM_032376858	uncharacterized, RT-like superfamily, peptidase superfamily, DUF1759 unknown function	RT/retrotransposons
NVE17112	XR_004297054	ncRNA	ncRNA
NVE13369	XR_004293601	miscRNA	miscRNA
NVE6128	XM_032363949	uncharacterized, blast hits - other uncharacterized proteins (incl zebrafish)	uncharacterized/cnidaria-specific
NVE5726	XM_032376662	uncharacterized, blast hits - other uncharacterized proteins (cnidaria)	uncharacterized/cnidaria-specific
NVE5672	XM_032376697	sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1 (SVEP1) calcium ion binding and chromatin binding.	ECM/cell adhesion
NVE24253	XM_032369598	uncharacterized, blast hits - other uncharacterized proteins	uncharacterized/cnidaria-specific
NVE18659	XM_032369784	uncharacterized, blast hits - other uncharacterized proteins	uncharacterized/cnidaria-specific
NVE23504	XR_004292166	miscRNA	miscRNA
NVE2560	XR_004293834	ncRNA	ncRNA

NVE10977	XM_032373395	uncharacterized, Reverse transcriptase (RTs) in retrotransposons. This subfamily represents the RT domain of a multifunctional enzyme.	RT/retrotransposons
NVE19504	XM_032369412	EGF-like repeat and discoidin I-like domain-containing protein 3 (same as NVE20550)	ECM/cell adhesion
NVE14525	XM_032386819	flocculation protein FLO11-like, Cell wall protein that participates in adhesive cell-cell interactions during yeast flocculation, a reversible, asexual and Ca2+-dependent process in which cells adhere to form aggregates (flocs) consisting of thousands of cells	ECM/cell adhesion
NVE10424	XM_032362651	KH homology domain-containing protein 4	RNA splicing
NVE9099	XM_032374520	multidrug resistance-associated protein 4, belongs to the ATP-binding cassette (ABC) superfamily of transporters, is localized to the basolateral membrane of hepatocytes and the apical membrane of renal proximal tubule cells.	Membrane transport
NVE22907	XM_032371133	autotransporter adhesin BpaC-like, Autotransporters form a large family of outer membrane proteins specifying diverse biological traits of Gram-negative bacteria	ECM/cell adhesion
NVE20069	XM_032369520	uncharacterized, SDK1, protein sidekick-1-like [Pocillopora damicornis]. Adhesion molecule that promotes lamina-specific synaptic connections in the retina.	ECM/cell adhesion
NVE1287	XM_032365086	thyroglobulin-like, Thyroglobulin (Tg) is a vertebrate secretory protein synthesized in the thyrocyte endoplasmic reticulum (ER)	ECM/cell adhesion
NVE9264	XM_032374374	myosin-11-like - Muscle contraction - It is a major contractile protein, converting chemical energy into mechanical energy through the hydrolysis of ATP.	cell motility
NVE15277	XR_004293291	ncRNA	ncRNA
NVE11894	XR_004295885	ncRNA	ncRNA
NVE6633	XM_032376014	uncharacterized, blast hits - other uncharacterized proteins, incl zn finger protein	uncharacterized/cnidaria-specific
NVE17538	XM_032369784	uncharacterized, blast hits - other uncharacterized proteins, same as NVE18659	uncharacterized/cnidaria-specific
NVE4885	XR_004292165.1	miscRNA	miscRNA
NVE22620	XM_032382832	tetratricopeptide repeat protein 14-like, TTC14, nucleic acid binding, function unknown	uncharacterized/cnidaria-specific

NVE9699	XM_032374134	uncharacterized, blast hits - other hypothetical proteins	uncharacterized/cnidaria-specific
NVE16190	XR_004293155	ncRNA	ncRNA
NVE11097	XM_032369784	same as NVE18659, uncharacterized	uncharacterized/cnidaria-specific
NVE15479	XM_032386450	multidrug resistance-associated protein 4, cystic fibrosis transmembrane conductance regulator [Mus musculus] - The CFTR gene codes for an ABC transporter-class ion channel protein that conducts chloride ions across epithelial cell membranes.	Membrane transport
NVE2545	XR_004295885	ncRNA	ncRNA
NVE20278	XM_032386874	uncharacterized, blast hits - hypothetical proteins	uncharacterized/cnidaria-specific
NVE18894	XM_032366394	neurogenic locus notch homolog protein 1	Notch signaling
NVE13408	XM_032369784	uncharacterized, blast hits -other uncharacterized proteins, same as NVE18659	uncharacterized/cnidaria-specific
NVE17860	XM_032385326	uncharacterized, blast hits - hypothetical proteins of cnidaria and mollusca, same as NVE24441	uncharacterized/cnidaria-specific
NVE253	XM_032367740	uncharacterized, Collagen alpha-1(VII) chain [Exaiptasia pallida], collagen alpha-4(VI) chain isoform X1 [Mus musculus], integrin-like	ECM/cell adhesion
NVE443	XM_032380333	EAAT1 (excitatory amino acid transporter 1)	Membrane transport, neurotransmission
NVE2551	XR_004295429	ncRNA	ncRNA
NVE12193	XM_032362078	adhesion G-protein coupled receptor V1-like, G-protein coupled receptor which has an essential role in the development of hearing and vision. In response to extracellular calcium, activates kinases PKA and PKC to regulate myelination by inhibiting the ubiquitination of MAG, thus enhancing the stability of this protein in myelin-forming cells of the auditory pathway. In retina photoreceptors, the USH2 complex is required for the maintenance of the periciliary membrane complex that seems to play a role in regulating intracellular protein transport.	ECM/cell adhesion
NVE17236	XM_032370293	proline-rich extensin-like protein EPR1, matches brevican isoform 1 variant, partial [Homo sapiens], In zebrafish matches neurocan core protein [Danio rerio]	ECM/cell adhesion

NVE10653	XM_001625963	RAB11-binding protein RELCH homolog, Regulates intracellular cholesterol distribution from recycling endosomes to the trans-Golgi network through interactions with RAB11 and OSBP (PubMed:29514919). Functions in membrane tethering and promotes OSBP-mediated cholesterol transfer between RAB11-bound recycling endosomes and OSBP-bound Golgi-like membranes.	lipid metabolism
NVE15290	XM_032371075	uncharacterized, blast hits - stabilin-1 [Homo sapiens], Acts as a scavenger receptor for acetylated low density lipoprotein. Binds to both Gram-positive and Gram-negative bacteria and may play a role in defense against bacterial infection. When inhibited in endothelial tube formation assays, there is a marked decrease in cell-cell interactions, suggesting a role in angiogenesis.	ECM/cell adhesion
NVE18731	XM_032384894	P2Y purinoceptor 6-like, belongs to the family of G-protein coupled receptors.	lipid metabolism
NVE22894	XM_032382668	Trace amine-associated receptor 6-like, belongs to a class of G protein-coupled receptors. Trace amines are endogenous amine compounds that are chemically similar to classic biogenic amines like dopamine, norepinephrine, serotonin, and histamine. Trace amines were thought to be 'false transmitters' that displace classic biogenic amines from their storage and act on transporters in a fashion similar to the amphetamines, but the identification of brain receptors specific to trace amines indicates that they also have effects of their own. TAAR6 expressed in kidney and brain. In mouse TAAR6 is pecifically expressed in neurons of the olfactory epithelium, to discrete glomeruli predominantly localized to a confined bulb region.	neuroactive ligand-receptor
NVE16188	XR_004293154	ncRNA	ncRNA
NVE12757	XM_032372270	blast hits - contactin-6 isoform X3 [Mus musculus], neuronal cell adhesion molecule isoform hh [Homo sapiens], the protein encoded by this gene is a member of the immunoglobulin superfamily. It is a glycosylphosphatidylinositol (GPI)-anchored neuronal membrane protein that functions as a cell adhesion molecule. It may play a role in the formation of axon connections in the developing nervous system.	ECM/cell adhesion
NVE5666	XM_032376690	laminin subunit beta-2, Binding to cells via a high affinity receptor, laminin is thought to mediate the attachment, migration and organization of cells into tissues during embryonic development by interacting with other extracellular matrix components. Tissue specificity -Neuromuscular synapse and kidney glomerulus.	ECM/cell adhesion

NVE5768	XM_001628429	protein HEXIM, Probable transcriptional regulator which functions as a general RNA polymerase II transcription inhibitor, Transcriptional regulator which functions as a general RNA polymerase II transcription inhibitor. In cooperation with 7SK snRNA sequesters P-TEFb in a large inactive 7SK snRNP complex preventing RNA polymerase II phosphorylation and subsequent transcriptional elongation.	TFs/transcription
NVE1877	XM_001630624	uncharacterized, blast hits - Zinc finger protein 474 (human, mouse) - ZNF494, Predicted to have metal ion binding activity and nucleic acid binding activity.	TFs/transcription
NVE13233	XM_032382318	ankyrin repeat and KH domain-containing protein 1-like, Could play pivotal roles in cell cycle and DNA regulation (PubMed:19150984). Involved in innate immune defense against virus by positively regulating the viral dsRNA receptors DDX58 and IFIH1 signaling pathways (PubMed:22328336).	immune response
NVE24958	XM_001632239	histone-lysine N-methyltransferase 2D, Histone methyltransferase. Methylates 'Lys-4' of histone H3 (H3K4me). H3K4me represents a specific tag for epigenetic transcriptional activation. Acts as a coactivator for estrogen receptor by being recruited by ESR1, thereby activating transcription.	TFs/transcription
NVE2264	XM_001630435	PLA2_2, blast hits - basic phospholipase A2 nigexine, phospholipase A2 isoform 2 precursor [Mus musculus], Phospholipase A(2) (PLA(2)) enzymes are involved in lipid metabolism and, as such, are central to several cellular processes.	lipid metabolism
NVE23108	LOC116607023	ncRNA	ncRNA
NVE12755	XM_001624903	adhesion G-protein coupled receptor D1, Orphan receptor. Signals via G(s)-alpha family of G-proteins (PubMed:22025619, PubMed:22575658). Has protumorigenic function especially in glioblastoma (PubMed:27775701), mus musculus - adhesion G protein-coupled receptor L3. Plays a role in cell-cell adhesion and neuron guidance via its interactions with FLRT2 and FLRT3 that are expressed at the surface of adjacent cells (PubMed:26235030).	ECM/cell adhesion

NVE16286	XM_032366658	AP-5 complex subunit beta-1, As part of AP-5, a probable fifth adaptor protein complex it may be involved in endosomal transport, Adaptor protein complex 1 is found at the cytoplasmic face of coated vesicles located at the Golgi complex, where it mediates both the recruitment of clathrin to the membrane and the recognition of sorting signals within the cytosolic tails of transmembrane receptors.	endocytosis/exocytosis
NVE17123	XM_001636071	pinin, Transcriptional activator binding to the E-box 1 core sequence of the E-cadherin promoter gene; the core-binding sequence is 5'CAGGTG-3'.	TF/transcription
NVE11362	XM_032362383	uncharacterized, blast hits - integrin and matrilin 1, cartilage matrix protein, isoform CRA_b [Homo sapiens], This gene encodes a member of von Willebrand factor A domain containing protein family.	ECM/cell adhesion
NVE23087	XM_032371074	volume-regulated anion channel subunit LRRC8E-like	membrane transport
NVE8748	XM_032374695	uncharacterized, collagen alpha-6(VI) chain isoform X2 [Xenopus tropicalis], von Willebrand factor type A domains, fibronectin type 2 domains, TSP type 1 repeats	ECM/cell adhesion
NVE15008	XM_032371244	early endosome antigen 1-like, EEA1 localizes exclusively to early endosomes and has an important role in endosomal trafficking.	endocytosis/exocytosis
NVE9385	XM_032365376	uncharacterized, blast hits - other uncharacterized proteins (cnidaria, moth, molluscs)	uncharacterized/cnidaria-specific
NVE5399	XM_001639086	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit, Catalyzes the transfer of a single N-acetylglucosamine from UDP-GlcNAc to a serine or threonine residue in cytoplasmic and nuclear proteins resulting in their modification with a beta-linked N-acetylglucosamine (O-GlcNAc)	protein modification
NVE6210	XM_032363854	neurogenic locus notch homolog protein 1	Notch signaling
NVE5945	XM_032376466	uncharacterized, thrombospondin 1, hemicentin-1 isoform X2 [Mus musculus], This gene encodes a large extracellular member of the immunoglobulin superfamily.	ECM/cell adhesion
NVE23760	XM_001621826	zinc metalloproteinase nas-6, Expressed in pharyngeal and body wall muscles, intestine, hypodermis and pharyngeal mc2 cells.	ECM/cell adhesion

NVE19786	XM_032384482	splicing factor 3B subunit 1, Involved in pre-mRNA splicing as a component of the splicing factor SF3B complex	RNA splicing
NVE9070	XM_032374535	uncharacterized, blast hits - chymotrypsinogen B2 precursor [Homo sapiens], Chymotrypsinis a digestive enzyme component of pancreatic juice acting in the duodenum, where it performs proteolysis, the breakdown of proteins and polypeptides.	protein degradation

Appendix 8. DE genes validated with RT-qPCR, their functional classification and RNA-seq attributes

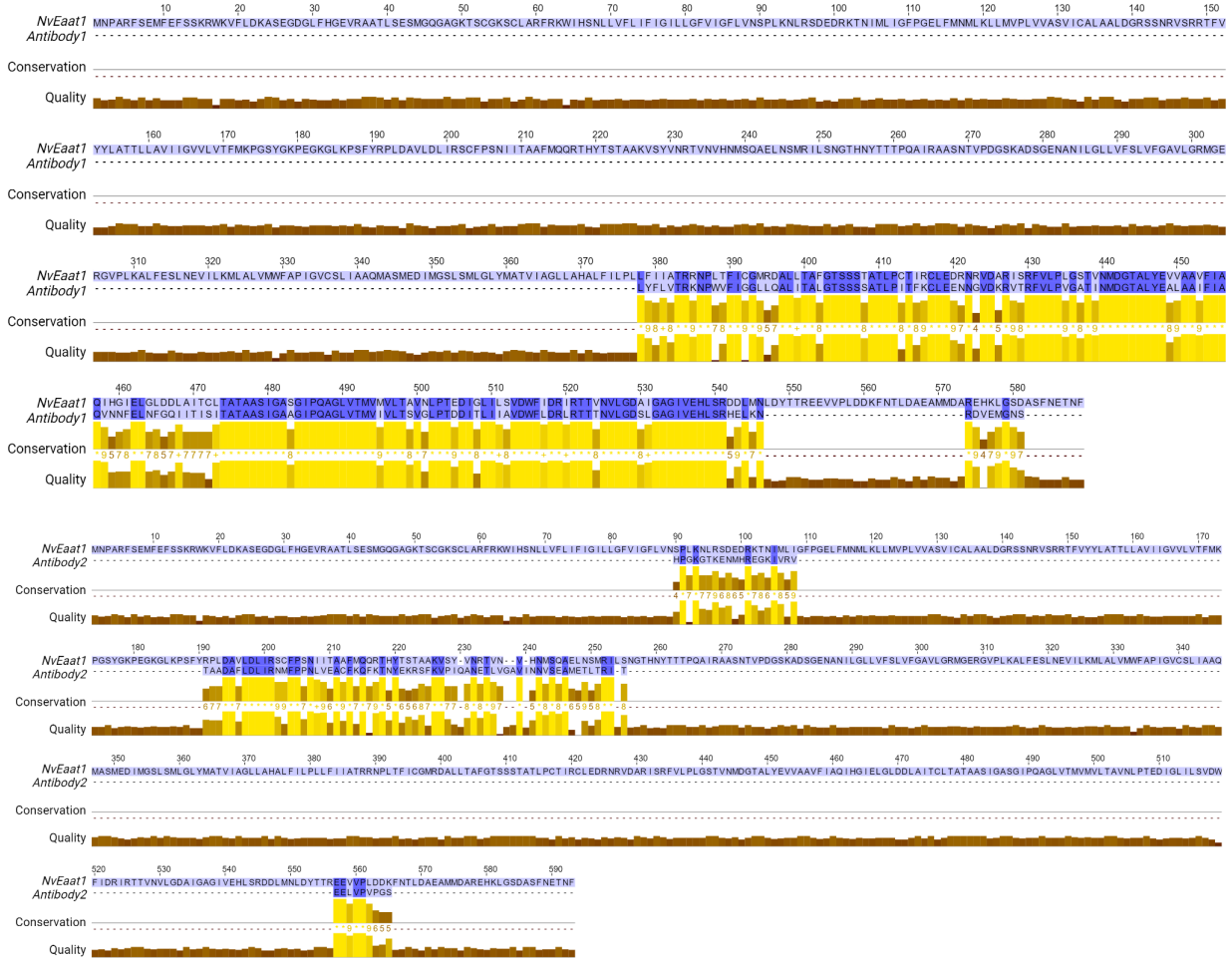
The genes were sorted by log2FoldChange, gene symbols are taken from the names of the bilaterian homologs.

Gene symbol	Nv_vienna_ID	Nv_Gene_ID	Padj	log2FoldChange	Functional classification
SORL1	NVE13564	5517308	4.85E-05	-0.43007	lipid metabolism
PTPkappa	NVE25055	5500246	6.46E-04	-0.42214	post-translational modification of proteins
CASPR2	NVE830	5510871	6.39E-03	-0.38865	ECM/cell adhesion
ZICA	NVE8653	5511658	7.01E-03	-0.37163	TF
EDIL3	NVE20550	116608110	2.39E-02	-0.35463	ECM/cell adhesion
PKD1L2	NVE22455	5520736	4.91E-02	-0.32723	ion channel
SGLT	NVE18073	5515535	1.01E-01	-0.30671	glucose metabolism
TRPA1	NVE1653	5510332	7.09E-02	-0.29384	ion channel
SDK1	NVE20069	116608285	1.77E-01	-0.28481	ECM/cell adhesion
Delta/ Notch-like	NVE18894	5521105	1.69E-01	-0.27414	Notch signaling
EAAT1	NVE443	5511109	1.86E-01	-0.27145	neurotransmission
ADGRV1	NVE12193	116601365	1.25E-01	-0.26955	ECM/cell adhesion
EPR1	NVE17236	116609465	1.23E-01	-0.26700	ECM/cell adhesion
P2RY6 -like	NVE18731	116619764	1.87E-01	-0.26629	lipid metabolism
Contactin/ DSCAM	NVE12757	116611752	1.16E-01	-0.26555	ECM/cell adhesion
PLA2_2	NVE2264	5509975	1.69E-01	-0.25764	lipid metabolism

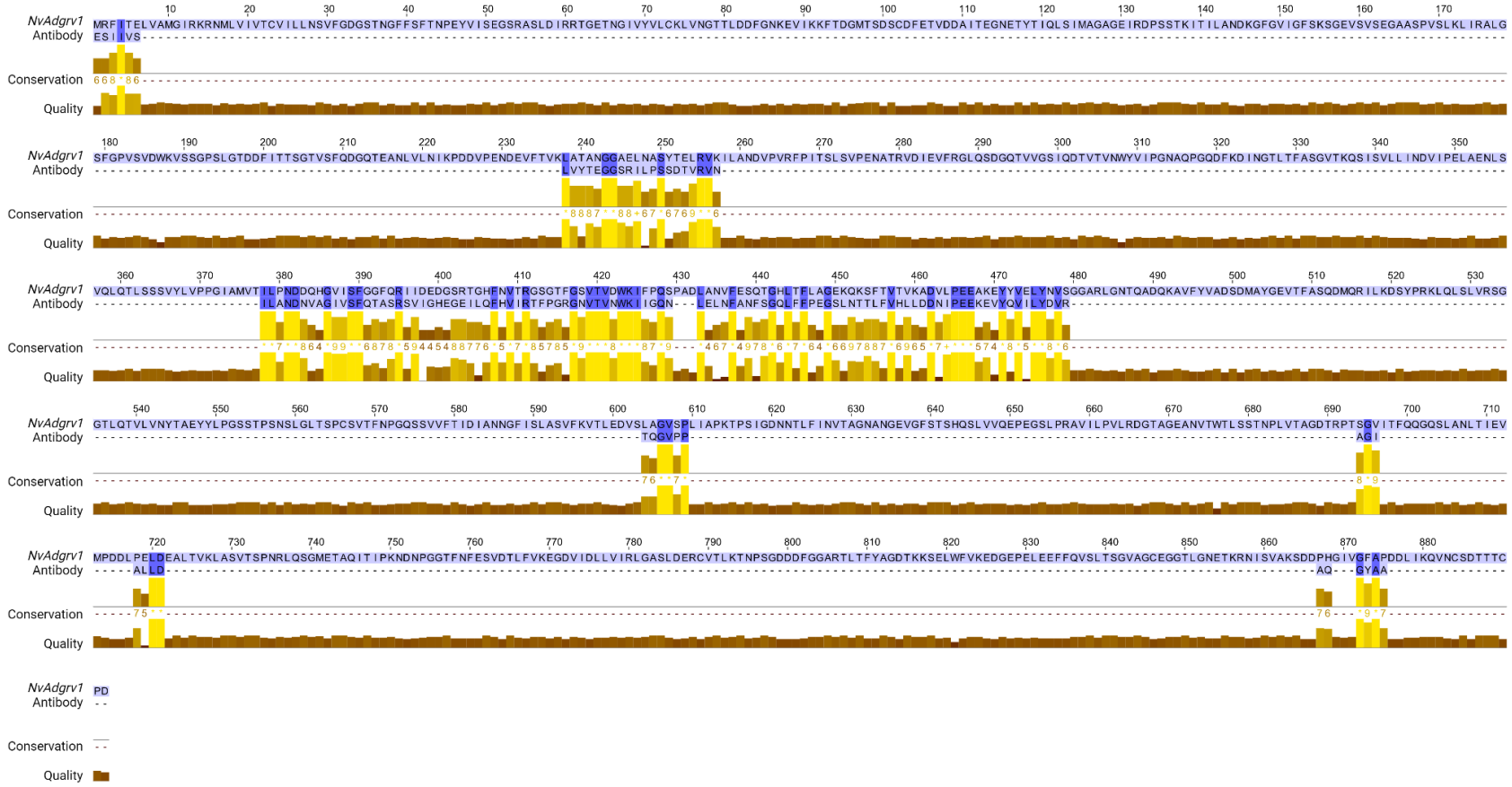
Appendix 8. DE genes validated with RT-qPCR, their functional classification and RNA-seq attributes

LRRC8A	NVE23087	116610263	1.87E-01	-0.24172	ion channel
OGT	NVE5399	5519150	1.55E-01	-0.23520	post-translational modification of proteins
PKM2	NVE38	5522266	2.04E-07	-0.54460	glucose metabolism
Notch1 -like	NVE8334	116602164	1.17E-01	0.26129	Notch signaling
ADGRB1	NVE10004	5505625	4.91E-02	0.31528	ECM/cell adhesion
MPEG-1	NVE3218	5509442	1.34E-02	0.35356	immune response
PLA2_1	NVE23561	5512771	6.06E-03	0.36609	lipid metabolism
Ets-related	NVE9883	5518230	4.05E-04	0.43941	TF
GAT2	NVE14046	5503381	NA	NA	neurotransmission
GAT3	NVE16531	5521615	NA	NA	neurotransmission

Appendix 9A. Alignment of *Nematostella Eaat1* and epitope sequences of two antibodies.
The sequences were aligned in the T-coffee program with the structural alignment (Espresso) method (Di Tommaso et al., 2011). Image created with JalView Version 2 (Waterhouse et al., 2009).



Appendix 9B. Alignment of *Nematostella Adgrv1* and antibody epitope sequence.
The sequences were aligned in the T-coffee program with the structural alignment (Espresso) method (Di Tommaso et al., 2011). Image created with JalView Version 2 (Waterhouse et al., 2009).



Appendix 10A. Expression levels of *Gcm*, some of its targets, and *PRGamide* in larval metacells

The expression (molecules/1,000 UMIs) of each gene is organized in the descending order accompanied by the respected metacell number. All the genes are abundantly expressed in neuronal metacell 25 (white font). Neuronal metacells are highlighted in blue. Raw unique molecule identifier (UMI) counts were taken from single-cell transcriptome ((Sebé-Pedrós, Saudemont, et al., 2018)).

Gcm	Meta cell	*Eaat1*	Meta cell	*PRPamide*	Meta cell	*PRGamide*	Meta cell	*Gat2*	Meta cell	*PLA2_2*	Meta cell	*CASPR2*	Meta cell
0.118708452	25	0.1140998945	31	5.650522317	25	12.93922127	25	0.5579297246	25	0.1305792972	25	0.5183206411	26
0.1095860387	36	0.1068376068	25	1.169523918	31	0.5205847902	32	0.5028984621	28	0.02336994625	12	0.3561253561	25
0.1056127299	17	0.04366335553	38	1.066960251	28	0.3514232642	3	0.4838255977	15	0.01359185033	28	0.0569476082	13
0.1004066469	3	0.04109476453	36	1.037075447	37	0.2940455771	14	0.4500196884	16	0.008011729172	29	0.05020332346	3
0.07124391882	27	0.03987081855	26	0.9568996451	26	0.2779682375	4	0.4338206585	32	0.005920067252	35	0.04338206585	32
0.06795925163	28	0.03889839739	2	0.6849127421	36	0.2199978	10	0.3987081855	26	0.004018032932	16	0.02593226492	2
0.06195403011	21	0.03240860773	37	0.6024398815	3	0.2140518761	5	0.3045366132	20	0.003238048363	11	0.02403518751	29
0.04900759618	14	0.027499725	10	0.5189620758	18	0.2131224211	35	0.2940455771	14	0	21	0.01733522865	7
0.04338206585	32	0.02368026901	35	0.3907244386	35	0.1944516464	37	0.2852497361	31	0	33	0.01473564257	20
0.04311937908	9	0.02282740201	23	0.3717241807	21	0.1944484955	30	0.2838118109	38	0	34	0.01295219345	11
0.03987081855	26	0.02035540538	27	0.3525560312	5	0.1746534221	38	0.2776917408	22	0	20	0.01231982259	9
0.02852497361	31	0.01964752343	20	0.3509849515	6	0.1643790581	36	0.2427227573	35	0	32	0.01173474776	17
0.02136090314	22	0.01853121583	4	0.2396877211	23	0.149526322	22	0.2112146532	33	0	19	0.01096827974	6
0.01853121583	4	0.01806635773	1	0.232340146	29	0.1369644121	23	0.2037905034	24	0	4	0.01070568293	34
0.01806635773	1	0.01770961544	19	0.2169103293	32	0.1358603356	24	0.2035540538	27	0	8	0.01002094377	8

Appendix 10A. Expression levels of *Gcm*, some of its targets, and *PRGamide* in larval metacells

0.01620404129	30	0.01687763713	15	0.2025316456	15	0.1239080602	21	0.2008132938	3	0	2	0.008854807718	19
0.0118401345	35	0.01602345834	29	0.1964850999	38	0.1196124556	26	0.1994907119	17	0	5	0.006795925163	28
0.008667614326	7	0.01359185033	28	0.1944919869	2	0.1173474776	17	0.1944484955	30	0	6	0.005920067252	35
0.008036065864	16	0.01231982259	9	0.1944484955	30	0.1170383146	9	0.1643790581	36	0	7	0.005625879044	15
0.004911880857	20	0.01096827974	6	0.16499835	10	0.1140998945	31	0.1539977824	9	0	22	0.004224293065	33
0.003568560978	34	0.01070568293	34	0.1607213173	16	0.09534375758	7	0.153448122	34	0	31	0.004018032932	16
0	33	0.01068045157	22	0.1482497267	4	0.093479785	12	0.1385041551	5	0	26	0	21
0	29	0.008448586129	33	0.1478378711	9	0.09018849395	8	0.1328221158	19	0	38	0	4
0	19	0	21	0.1470227885	14	0.08774623788	6	0.1296613246	2	0	36	0	5
0	8	0	32	0.1445308619	1	0.07984031936	18	0.1296344309	37	0	27	0	22
0	2	0	8	0.1358603356	24	0.07969326946	19	0.1213466006	7	0	30	0	31
0	5	0	5	0.1141939513	34	0.07779679477	2	0.1083981464	1	0	24	0	38
0	6	0	7	0.1138952164	13	0.07447511236	11	0.106855596	11	0	1	0	36
0	38	0	30	0.09387798209	17	0.07232459277	16	0.09871451762	6	0	3	0	27
0	24	0	24	0.08871015436	33	0.0642340976	34	0.09265607917	4	0	9	0	30
0	10	0	3	0.07010983875	12	0.06336439597	33	0.09130960805	23	0	10	0	24
0	11	0	11	0.06408270942	22	0.06106621613	27	0.08812902089	29	0	13	0	1
0	12	0	12	0.06067330028	7	0.0569476082	13	0.07014660641	8	0	14	0	10
0	13	0	13	0.04427403859	19	0.05436740131	28	0.07010983875	12	0	18	0	12
0	18	0	14	0.04008377509	8	0.04807037503	29	0.06195403011	21	0	17	0	14

Appendix 10A. Expression levels of *Gcm,* some of its targets, and *PRGamide* in larval metacells

0	15	0	18	0.02266633854	11	0.04420692772	20	0.0569476082	13	0	15	0	18
0	23	0	17	0.02035540538	27	0.03613271546	1	0.05499945001	10	0	23	0	23
0	37	0	16	0.01964752343	20	0.01687763713	15	0	18	0	37	0	37

Appendix 10B. Expression levels of *Gcm*, some of its targets, and *PRGamide* in adult neuronal metacells

The expression (molecules/1,000 UMIs) of each gene is organized in the descending order accompanied by the respected metacell number. *Gcm*-rich metacells are highlighted in yellow. Metacell 62 with the highest abundance of *PRGamide* and the *Gcm* targets is highlighted in green. Raw unique molecule identifier (UMI) counts for each gene were taken from single-cell transcriptome data ((Sebé-Pedrós, Saudemont, et al., 2018)).

Gcm	Meta cell	*Eaat1*	Meta cell	*PRPamide*	Meta cell	*PRGamide*	Meta cell	*GAT2*	Meta cell	*PLA2_2*	Meta cell	*CASPR2*	Meta cell
2.253491254	34	4.170390229	57	20.7130661	35	367.919278	62	10.76777893	62	3.024657004	62	19.84174995	62
2.067549806	40	2.982561954	34	20.39409585	34	284.9049498	63	10.68703801	44	1.637384769	63	8.028143605	45
1.957530108	57	2.784418777	37	9.652730908	49	5.927682276	51	8.023185368	63	0.533560986	36	4.20540062	79
0.982929786	59	2.197931014	46	5.038512277	57	5.787480318	57	7.652528813	52	0.2318948125	52	1.964861723	63
0.6956844375	52	1.974171259	50	2.277090682	42	3.636033088	47	7.504946442	54	0.1932049808	39	0.7508796018	55
0.5552162567	75	1.930377711	56	0.9738822488	79	3.503715634	48	7.403012065	48	0.1364535717	54	0.6610969251	47
0.3955807974	48	1.814794202	62	0.8812002875	52	3.463391947	44	7.397942027	38	0.1277122898	41	0.608469901	58
0.3831368693	41	1.802037941	59	0.8785671945	51	3.381087164	39	7.251261311	64	0.1058514692	51	0.5386811546	66
0.3305484625	47	1.623263688	52	0.8476731372	48	3.36050542	37	7.106791945	47	0.1014116502	58	0.51285057	46
0.3290285432	50	1.622586403	58	0.8122517094	40	3.296896521	46	7.024175936	41	0.09895405563	44	0.510652205	64
0.3274769538	63	1.484918794	53	0.6886850511	46	2.934585423	36	5.922513778	50	0.0819108155	59	0.3976749273	34
0.2880433217	37	1.473646292	63	0.6240938637	37	2.467714074	50	5.907285159	40	0.07695445066	66	0.3820342048	49
0.2729071433	54	1.400792084	49	0.6003957154	54	2.451130584	64	5.1034949	34	0.05604501536	38	0.3014227152	102

Appendix 10B. Expression levels of *Gcm*, some of its targets, and *PRGamide* in adult neuronal metacells

0.2028233003	58	1.308225661	66	0.5928327728	62	2.054653791	34	4.981976966	46	0	34	0.2729071433	54
0.1856148492	53	1.106684374	79	0.5922513778	50	2.032685584	35	4.934351669	55	0	40	0.1979081113	44
0.1539089013	66	1.073719307	48	0.5796149425	39	1.919867677	40	4.825986079	53	0	57	0.1645142716	50
0.1465287343	46	1.058514692	51	0.402162023	56	1.681350461	38	4.505094853	59	0	75	0.163821631	59
0.1273447349	49	0.9660249041	39	0.3608154429	45	1.63712402	42	4.316054233	42	0	48	0.1608648092	56
0.1058514692	51	0.909008272	47	0.3463391947	44	1.623263688	52	4.255500234	57	0	47	0.1488294563	42
0.102130441	64	0.8860927739	40	0.3432592465	55	1.533465632	45	4.250509578	39	0	50	0.1277122898	41
0.09895405563	44	0.8406752303	38	0.3334756163	36	1.500989288	54	4.239581454	45	0	37	0.113023085	48
0.09660249041	39	0.8328243851	75	0.3155452436	53	1.484918794	53	4.02235583	51	0	53	0.1101806963	65
0.09020386073	45	0.7712648744	65	0.2947292584	63	1.388040642	75	3.947686783	49	0	46	0.1016342792	35
0.05604501536	38	0.7508796018	55	0.2858295783	38	1.318351452	58	3.840577623	37	0	49	0.09660249041	39
0	102	0.7441472816	42	0.2451130584	64	1.286918474	56	3.650819406	58	0	64	0.08511000468	57
0	56	0.6926783894	44	0.2148565006	47	1.273447349	49	3.334756163	36	0	45	0.05604501536	38
0	79	0.6669512325	36	0.2043396636	41	1.228662233	59	2.744125539	35	0	102	0	40
0	36	0.5412231644	45	0.1723998053	58	1.205690861	102	2.573836947	56	0	56	0	52
0	55	0.510652205	64	0.1432349052	65	0.8853474989	79	2.109959007	102	0	79	0	75
0	35	0.409360715	54	0.1310573048	59	0.8581481164	55	1.652710445	65	0	55	0	37
0	42	0.4065371168	35	0.06028454304	102	0.6925900559	66	1.388040642	75	0	35	0	53
0	62	0.2554245795	41	0.05386811546	66	0.6610841781	65	1.385180112	66	0	42	0	51
0	65	0.2429808896	76	0.03644713343	76	0.6385614488	41	1.328021248	79	0	65	0	36
0	76	0	102	0.02776081284	75	0.3644713343	76	0.4859617791	76	0	76	0	76

Appendix 11. Enriched genes in larval and adult metacells with the highest expression of *Gcm*

Color coding: *Gcm*, glutamate and GABA metabolism, synaptogenesis, neuronal TFs, calcium-binding/sensing, potassium voltage-gated channel, ECM/cell adhesion/axon guidance, neuropeptide receptor, other neurotransmitter receptors, neuronal enzymes, meteorin. Data from (Sebé-Pedrós, Saudemont, et al., 2018).

Larva	Adult		
Neuronal metacell C25	**Neuronal metacell C34**	**Neuronal metacell C40**	**Neuronal metacell C57**
Secretagogin	Tetraspanin (Tetraspanins are often thought to act as scaffolding proteins, anchoring multiple proteins to one area of the cell membrane)	N-acetylated-alpha-linked acidic dipeptidase 2	"ST6 (Alpha-N-acetyl-neuraminyl-2, 3-beta-galactosyl-1, 3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4"
Tubulin alpha-1C chain	Galanin receptor type 2 (Galanin (neuropeptide) is an important neuromodulator present in the brain, gastrointestinal system, and hypothalamopituitary axis)	Homeobox protein EMX2	Tetraspanin
Endophilin-A3	Uncharacterized protein KIAA1958	Solute carrier family 35 member G1	Orexin type-2 receptor isoform 1
HCG15971	Homeobox protein OTX1	UPF0462 protein C4orf33	Histamine H2 receptor
Trans-acting T-cell-specific transcription factor GATA-3	"Adrenergic, alpha-1D-, receptor " (They activate mitogenic responses and regulate growth and proliferation of many cells.)	Probable N-acetyltransferase 16	Fibroblast growth factor 7
Tubulin alpha-1C chain	"ST6 (Alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1, 3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4"	Uncharacterized protein KIAA1467	APBB2 protein
Calmodulin	Orexin type-2 receptor isoform 2	Tetraspanin	"Adrenergic, alpha-1D-, receptor "
"Glutamate receptor, ionotrophic variant "	APBB2 protein (The protein encoded by this gene interacts with the cytoplasmic domains of amyloid beta (A4) precursor protein and amyloid beta (A4) precursor-like protein 2)	Histamine H2 receptor	Peptidyl-glycine alpha-amidating monooxygenase

Appendix 11. Enriched genes in larval and adult metacells with the highest expression of *Gcm*

PR domain zinc finger protein 14	Potassium voltage-gated channel subfamily A member 10	Thyrotropin receptor	Neuroendocrine convertase 1
Tubulin alpha-1A chain	Substance-K receptor	"Glycine amidinotransferase, mitochondrial "	Protein APCDD1
Protein FAM26E	von Willebrand factor A domain-containing protein 5B1	FOLH1 protein	Synaptotagmin I
3'-phosphoadenosine 5'-phosphosulfate synthase 1	Neuroendocrine convertase 1	Hypocretin receptor-2	"cDNA FLJ45846 fis, clone OCBBF2007039, highly similar to ADAMTS-16 (EC 3.4.24.-) "
Frequenin homolog (Drosophila)	Cationic amino acid transporter 4 (Involved in the transport of the cationic amino acids (arginine, lysine and ornithine).)	"b(0,+)-type amino acid transporter 1 "	Metalloendopeptidase
Cubilin variant	"cDNA FLJ45846 fis, clone OCBBF2007039, highly similar to ADAMTS-16 (EC 3.4.24.-) "	Retinoid X nuclear receptor alpha	Neogenin
Major facilitator superfamily domain-containing protein 4	Diencephalon/mesencephalon homeobox protein 1	Neuropeptide FF receptor 2	Homeobox protein HMX3
Reticulon	VEGFC protein (Growth factor active in angiogenesis, and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels.)	UNC homeobox protein transcript variant 1	Amino acid transporter
Vesicle amine transport protein 1 homolog (T californica)	Histamine H2 receptor	Fibrinogen C domain-containing protein 1	Anoctamin-7
Class A basic helix-loop-helix protein 15	Epidermal growth factor-like protein 8	Alpha-1A adrenergic receptor	Cationic amino acid transporter 4
"cDNA FLJ59828, highly similar to Regulator of G-protein signaling 19 "	Reticulon (evolutionary conservative proteins residing predominantly in endoplasmic reticulum, primarily playing a role in promoting membrane curvature)	TRAF-type zinc finger domain containing 1	Trans-acting T-cell-specific transcription factor GATA-3
Vesicular inhibitory amino acid transporter	CTRL protein (is expressed specifically in the pancreas and likely functions as a digestive enzyme.)	"cDNA, FLJ94678, highly similar to Homo sapiens BCL2-associated athanogene 4 (BAG4), mRNA "	Fibrinogen C domain-containing protein 1

Appendix 11. Enriched genes in larval and adult metacells with the highest expression of *Gcm*

Transmembrane protein 77	Metalloendopeptidase	Transmembrane protein 163	Gamma-aminobutyric acid (GABA) B receptor 1 isoform c variant
"Protein kinase, cAMP-dependent, regulatory, type II, alpha"	Neogenin (member of the immunoglobulin superfamily of transmembrane receptors (involved in axon guidence). Neogenin is also a receptor for repulsive guidance molecule, a glycosylphosphatidylinositol-linked protein involved in neuronal differentiation, apoptosis and repulsive axon guidance. Expressed in muscle and cnidocyte clusters)	Transmembrane and TPR repeat-containing protein 2	Synaptotagmin II
Major facilitator superfamily domain-containing protein 6	"cDNA FLJ50459, highly similar to ETS translocation variant 1 " (Transcriptional activator, involved in many pocesses incl. axon guidance)	Synaptotagmin I	Synaptotagmin-7
SEC14-like 4 (S. cerevisiae)	Trans-acting T-cell-specific transcription factor GATA-3 (Transcriptional activator which binds to the enhancer of the T-cell receptor alpha and delta genes. Binds to the consensus sequence 5'-AGATAG-3'. Required for the T-helper 2 (Th2) differentiation process following immune and inflammatory responses.)	Enterokinase catalytic subunit	Reticulon
DBH-like monooxygenase protein 1	Proprotein convertase subtilisin/kexin type 1	Synaptotagmin II	Pancreas transcription factor 1 subunit alpha
Protein-tyrosine-phosphatase	Tachykinin receptor 2 variant (substance K receptor)	HCG27868	Alpha-1A adrenergic receptor
Multiple C2 and transmembrane domain-containing protein 1	Protein APCDD1 (Negative regulator of the Wnt signaling pathway. Inhibits Wnt signaling in a cell-autonomous manner and functions upstream of beta-catenin.)	KIAA1274	Galanin receptor type 2
"Core-binding factor, beta subunit"	Protein phosphatase 1 regulatory subunit 27 (Inhibits phosphatase activity of protein phosphatase 1 (PP1) complexes.)	Protocadherin Fat 4	Homeobox protein OTX1

Appendix 11. Enriched genes in larval and adult metacells with the highest expression of *Gcm*

Glutamate receptor 4	Otoferlin (large multi–C2 domain protein proposed to act as a calcium sensor that regulates synaptic vesicle exocytosis in cochlear hair cells.	Transcription factor AP-2-alpha	Orthodenticle homolog 1 (Drosophila)
Anoctamin-7	Alpha-1A adrenergic receptor (G protein-coupled receptors that are targets of many catecholamines like norepinephrine (noradrenaline) and epinephrine (adrenaline))	Talin-2	KIAA1274
VEGFC protein	Latrophilin-1 (G-protein coupled receptor, Latrophilins may function in both cell adhesion and signal transduction)	Long-chain-fatty-acid--CoA ligase 1	Protocadherin Fat 4
ELAV-like protein	Endothelin-converting enzyme 1	Meteorin	VEGFC protein
CDW9/WDR51B	KIAA1274 (negative regulator of insulin signaling)	Cryptochrome 1 (Photolyase-like)	ELAV-like protein
"POU domain, class 4, transcription factor 3 "	Transmembrane prolyl 4-hydroxylase (Catalyzes the post-translational formation of 4-hydroxyproline in hypoxia-inducible factor (HIF) alpha proteins)	Orthodenticle homolog 1 (Drosophila)	"RAS-like, estrogen-regulated, growth inhibitor"
Synaptotagmin II	Orexin type-2 receptor isoform 1	Endothelin-converting enzyme 1	Adenosine receptor A2
"Cyclin-dependent kinase inhibitor 2D (P19, inhibits CDK4)"	Amino acid transporter	Pim-1 oncogene	Otoferlin
Septin 9	Pancreas transcription factor 1 subunit alpha	Cytochrome P450	von Willebrand factor A domain-containing protein 5B1
Visual system homeobox 1	Adenosine receptor A2	Frequenin homolog (Drosophila)	NK1 transcription factor-related protein 1
Plasma membrane calcium-transporting ATPase 3	Tubulin alpha-1C chain	Epididymis luminal protein 33	"cDNA FLJ50459, highly similar to ETS translocation variant 1 "
Serine/threonine-protein kinase 17A	Neuropeptide FF receptor 2	Transporter	Tubulin alpha-1C chain

Appendix 11. Enriched genes in larval and adult metacells with the highest expression of *Gcm*

Feline leukemia virus subgroup C receptor-related protein 1	Enterokinase catalytic subunit (enzyme involved in digestion)	"cDNA FLJ61245, highly similar to Homo sapiens kinase D-interacting substance of 220 kDa (KIDINS220), mRNA "	
Tetraspanin	Fibrinogen C domain-containing protein 1 (Acetyl group-binding receptor which shows a high-affinity and calcium-dependent binding to acetylated structures such as chitin, some N-acetylated carbohydrates, and amino acids, but not to their non-acetylated counterparts. Can facilitate the endocytosis of acetylated components.)	Uncharacterized protein KIAA1958	
Potassium channel subfamily T member 1	Orthodenticle homolog 1 (Drosophila) (may play a role in brain and sensory organ development.)	Potassium voltage-gated channel subfamily C member 1	
Potassium voltage-gated channel subfamily A member 1	Chorion-specific transcription factor GCMb	Epididymis luminal secretory protein 52	
Protein canopy homolog 4	"cDNA FLJ46798 fis, clone TRACH3031660, highly similar to cAMP-dependent protein kinase type II-beta regulatory subunit "	"Protein kinase, cAMP-dependent, catalytic, beta"	
Rho guanine nucleotide exchange factor 7	Neuropeptide FF receptor 2	"cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A "	
Endothelin-converting enzyme 1	Protein Tob2 (involved in the regulation of cell cycle progression)	"Adrenergic, alpha-1D-, receptor "	
Peptidyl-glycine alpha-amidating monooxygenase		Homeobox protein OTX1	
Contactin associated protein-like 2 isoform B		Protein phosphatase 1 regulatory subunit 27	
Kalirin		Lactadherin	
Proprotein convertase subtilisin/kexin type 1		"Fibroblast growth factor receptor 1 (Fms-related tyrosine kinase 2, Pfeiffer syndrome)"	

Appendix 11. Enriched genes in larval and adult metacells with the highest expression of *Gcm*

"Protein kinase, cAMP-dependent, catalytic, beta"		Fibroblast growth factor 7	
Aquaporin-2		"cDNA, FLJ95945 "	
Orexin type-2 receptor isoform 2		"ST6 (Alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1, 3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4"	
Calponin-3		Heat shock factor protein 1	
Runt-related transcription factor 1/NDC1 transmembrane nucleoporin fusion protein			
Synaptotagmin-14			
RUN and FYVE domain-containing protein 4			
7-dehydrocholesterol reductase isoform A			
Tumor protein p53-inducible protein 11			
Tubulin alpha-1A chain			
Synaptotagmin-7			
Melatonin receptor type 1A			
Melatonin-related receptor			