Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht-Karls-University

Heidelberg

Presented by:

M.Sc. Bruno Gideon Bergheim

Born in: Köln

Oral examination: 24th January 2024

Molecular Atlas

and

Developmental Dynamics

of the

Nematostella vectensis

Mesoglea

Referees:

apl. Prof. Dr. Suat Özbek

Prof. Dr. Thomas Holstein

In the intricate dance of cells,

the extracellular matrix is the choreographer.

Abstract

Cnidarians (corals, jellyfish and sea anemones) constitute ideal model systems to investigate the dynamics and interactions of extracellular matrix (ECM) components. Many cnidarians undergo a complex life cycle characterized by transitions through different life stages with distinct morphologies. These morphological changes recapitulate developmental processes observed in vertebrates and involve the synthesis, degradation and remodeling of the ECM.

The simple body plan of cnidarians provides exceptional accessibility for ECM imaging and manipulation, and previous studies have shown that the cnidarian ECM shares striking compositional and structural similarities with the vertebrate ECM. However, a major hurdle in the investigation of the cnidarian ECM is the lack of fully annotated ECM datasets allowing for a deeper understanding of its interactions and functions.

In this thesis, I used the starlet sea anemone *Nematostella vectensis* as a model system to fully characterize the components of the cnidarian ECM by a combination of electron microscopy, single cell transcriptomics, *in silico* matrisome prediction and mass spectroscopic analysis of isolated ECM from three life stages (larvae, primary polyp and adult).

I present the fully annotated *in silico* matrisome comprising 843 proteins, categorized into 246 core matrisome, 309 matrisome-associated and 289 other ECM-like proteins. 182 of these components are specific to the cnidocysts. My research revealed that the primary source of ECM is the gastroderm while the contribution of the ectoderm to ECM production is minimal.

The integration of single cell transcriptome analysis and stage-specific mass spec-

trometry unraveled dramatic changes of the ECM during development, showing that the basal lamina gets established well in advance of the fibrillar interstitial matrix.

Among the various ECM components upregulated during the larva-to-polyp transition, I detected an expanded polydom/SVEP1 family in *Nematostella*. This family encompasses several novel proteins, one of which emerges as a promising candidate for genetic perturbation using shRNAs. The resulting knockdown phenotype displayed defects in epithelial organization, body elongation, and development of internal mesenteries. This underscores that individual ECM components can heavily affect crucial developmental processes.

In summary, my work establishes a foundational framework for future studies of the ECM by providing the first annotated molecular atlas for a cnidarian life cycle. In addition, it provides evidence for an ancient specification of ECM-producing cells within the endomesodermal lineage, along with a hierarchical function of basement membrane and interstitial matrix components in the life history of *Nematostella*.

Zusammenfassung

Nesseltiere (Korallen, Quallen, Seefedern und Seeanemonen) stellen ideale Modellsysteme dar, um die Funktion und Interaktion von extrazellulären Matrix (ECM)-Komponenten zu untersuchen. Viele Nesseltiere besitzen komplexe Lebenszyklen, in denen sie verschiedene Lebensstadien mit unterschiedlicher Morphologie durchlaufen. Diese Veränderungen in der Morphologie spiegeln bekannte Entwicklungsprozesse bei Wirbeltieren wider und beinhalten die Synthese, den Abbau und die Umgestaltung der ECM. Aufgrund des einfachen Körperbaus von Nesseltieren ist ihre ECM leicht für Mikroskopie und Manipulation zugänglich. Frühere Studien haben gezeigt, dass die Struktur und Komposition der Nesseltier-ECM der von Wirbeltieren ähnelt.

Ein wesentliches Hindernis bei der Untersuchung der Nesseltier-ECM ist jedoch das Fehlen vollständig annotierter ECM-Datensätze, die ein tieferes Verständnis ihrer Interaktionen und Funktionen ermöglichen.

In der vorliegenden Arbeit habe ich die zu den Anthozoen gehörende Seeanemone *Nematostella vectensis* als Modellsystem verwendet, um die Komponenten ihrer ECM vollständig zu charakterisieren. Dies erfolgte durch eine Kombination aus Elektronenmikroskopie, Single Cell-Transkriptomik, *in silico* Matrisom-Vorhersage und massenspektrometrischer Analyse der isolierten ECM von drei Lebensstadien (Larve, primärer Polyp und adulte Tiere).

Ich präsentiere hier das vollständig annotierte *in silico* Matrisom von *Nematostella vectensis*, das aus 843 Proteinen besteht und in 246 Core Matrisome, 309 Matrisom-assoziierte und 289 sonstige ECM-ähnliche Proteine unterteilt werden kann. Ich zeige, dass die Mehrzahl der ECM-Komponenten vom Gastroderm se-

kretiert wird und das Ektoderm nur eine geringe Rolle in der ECM-Produktion spielt. Durch den Vergleich von Orthogruppen zwischen verschiedenen basalen Tieren bestätige ich, dass die ECM von Anthozoen der von Bilateriern ähnelt, jedoch seine eigenen spezialisierten Orthogruppen enthält, wie etwa die 182 Matrisomkomponenten der Cnidocysten.

Die Einbeziehung des Single Cell-Transkriptoms und der stadienspezifischen Massenspektrometrie zeigten, dass sich die ECM während der Entwicklung dramatisch verändert und die Basallamina vor der fibrösen interstitiellen Matrix etabliert wird.

Unter den verschiedenen ECM-Komponenten, die während des Übergangs von der Larve zum Polypen hochreguliert sind, habe ich eine ungewöhnlich expandierte Polydom/SVEP1-Familie in *Nematostella* identifiziert. Diese Familie enthält eine Reihe von neuartigen Faktoren, von denen einer für genetische Manipulationen mit shRNAs ausgewählt wurde. Der erhaltene Knockdown-Phänotyp zeigt Defekte in der epithelialen Organisation, der embryonalen Körperstreckung und der Entwicklung innerer Strukturen wie Mesenterien. Dies zeigt eindrucksvoll, dass selbst einzelne ECM-Komponenten erheblichen Einfluss auf wichtige Entwicklungsprozesse haben können.

Insgesamt bildet meine Arbeit eine wichtige Grundlage für zukünftige Studien zur ECM, indem sie den ersten annotierten molekularen Atlas für einen Vertreter der Nesseltiere mit komplexem Lebenszyklus bereitstellt. Darüber hinaus liefere ich Hinweise auf eine phylogenetisch frühe Spezifikation von ECM-produzierenden Zellen aus der endomesodermalen Linie und eine hierarchische Abfolge der Basallamia und Interstitalmatrix-Komponenten im Lebenszyklus von *Nematostella*.

Acknowledgements

First of all, I want to thank my girlfriend Ina for supporting me throughout my whole studies, and especially the hard weeks of writing this thesis, for lovingly preparing food when I was hungry and reminding me that there is more to life than just science. This thesis would not be possible without you.

I want to thank my parents for setting me on the path of biology and art and for enabeling me to study for all these years and achieve this incredible achievement.

I want to thank Suat for being a great supervisor throughout my PhD studies and seeing the story in my project even during the times when I could not see it.

Further thanks go to Thomas and Aissam for being part of my thesis advisory committee. For giving advice and support and telling me to drop the additional projects that I so desperately wanted to add to my already large workload.

A big thanks goes to my colleagues, Niharika, Bérénice, Ibrahim and Maike for making my time in the Özbek lab so much fun and for spending way to much time drinking tea and talking during the work hours.

I want to thank everybody on the 6th floor for spending the last 4 years together, for all the fun social evenings, game nights and of course also for all the scientific discussions and help throughout the years.

Last but not least I want to thank all the colleagues, friends and former students that spend a lot of time reading through and spellchecking my thesis.

Publication Statement

The PhD candidate prepared the following publications during the time of the thesis:

Bergheim BG, Cole A, Stein F, Rettel M, Redl S, Hess M, Ikmi A, Adams JC, Özbek S. Mesoglea dynamics across the Nematostella life cycle reveals novel factors of epithelial organization during larva-to-polyp transition. Manuscript in preparation, this publication contains the main findings of this thesis therefore figures and methods are partially reused here.

Veschgini M, Petersen HO, Suzuki R, Kling S, Bergheim BG, Abuillan W, Linke P, Kaufmann S, Burghammer M, Özbek S, Holstein TW, Tanaka M. Wnt/β-Catenin Signaling Controls Spatio-Temporal Elasticity Patterns in Extracellular Matrix during Hydra Morphogenesis. *iScience*, 26(4), 106416, 2023 10.1016/j.isci.2023.106416

Garg N, Štibler UK, Eismann B, Mercker M, Bergheim BG, Linn A, Tuchscherer P, Engel U, Redl S, Marciniak-Czochra A, Holstein TW, Hess MW, Özbek S. Non-muscle myosin II drives critical steps of nematocyst morphogenesis. *iScience*, 106291202) 10.1016/j.isci.2023.106291

Hofmann, D, Garg, N, Grässle, S, Vanderheiden S, Bergheim BG, Bräse S, Jung N, Özbek S. A small molecule screen identifies novel inhibitors of mechanosensory nematocyst discharge in Hydra. *Sci Rep* 11, 20627, 2021. 10.1038/s41598-021-99974-7

Gornik S[1], Bergheim BG[1], Morel B, Stamatakis A, Foulkes NS, Guse A. Photoreceptor Diversification Accompanies the Evolution of Anthozoa. *Mol. Biol. Evol.* 2020, Nov, doi: 10.1093/molbev/msaa304

Bergheim BG, Özbek S. Extracellular matrix and morphogenesis in cnidarians: a tightly knit relationship. *Essays Biochem.* 2019 Sep 13;63(3):407-416. PMID:31462530 doi: 10.1042/EBC20190021.

[1]Shared first authorship.

Collaboration Statement

Within this thesis, I present data, which was generated by collaborators and interpreted by me.

All data has been included with the permission and knowledge of the collaborator and I made all effort to acknowledge their work.

Mandy Rettel and **Frank Stein** (European Molecular Biology Laboratory (EMBL) Protein Core Facility): Performed the mass spectrometry analysis of ECM material and the statistical analysis of the results.

Alison Cole (University of Vienna, Department for Neuroscience and Developmental Biology): Performed the single cell transcriptome analysis.

Michael Hess and **Stefan Redl** (Medizinische Universität Innsbruck, Institute of Histology and Embryology): Performed parts of the the electron microscopy experiments presented.

Chalotta Funaya and **Stefan Hillmer** (Electron Microscopy Core Facility, Uni Heidelberg): Performed the other initial electron microscopy.

Contents

List of Figures

List of Tables

Abbreviations

aa	amino acids
ADAM	a disintegrin and metalloproteinase
ADAMTS	ADAM with thrombospondin motifs
BL	basal lamina
BLAST	basic local alignment search tool
BMP	bone morphogenetic protein
CAM	cell adhesion, molecule
CCP	complement control protein
COLIG	collagen and immunoglobulin
CPP	cnidarian poly proline protein
CRD	cysteine-rich-domain (CxxxCxxxCxxxCxxxCC)
CUB	C1r/C1s, Uegf, Bmp
DAPI	4′,6-diamidino-2-phenylindole
dda	Data dependent acquisition
dpf	days post fertilization
DTT	dithiothreitol
ECM	extracellular matrix
EGF	epidermal growth factor
EM	electron microscopy
EMBL	European Molecular Biology Laboratory
ER	endoplasmatic reticulum
GO	gene onthology
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HYR	hyalin repeat
IM	interstitial matrix
LDL	low-density-lipoproteins
LDLR	LDL receptor
LRP	LDL receptor-related protein
MAM	meprin, A-5 protein, and receptor protein-tyrosine phosphatase mu
MASP	mannose-associated serine protease
MMP	matrix metalloprotease
MS	mass spectrometry
nep8	nematocyte expressed protein 8
NOWA	novel outer wall antigen
OG	Orthogroup
PAN	plasminogen-apple-nematode
PAS	periodic acid–Schiff
PBS	phosphate-buffered saline
PFA	performaldehyde
PGC	primordial germ cells
ppm	parts per million
pSC	putative stem cells
rpm	rotations per minute
rt	room temperature
SCR	short consensus repeat
SDS	sodiumdodecyl sulfate
SEA	sea urchin sperm-coating protein, enterokinase, agrin
shRNA	short hairpin RNAs
SVEP	Sushi, von Willebrandt factor type A, EGF and pentraxin containing protein
TCA	trichloroacetic acid
TKE	thyroxine kinase ephrin
TMT	tandem mass tag
TSP	thromospondin
UMAP	uniform manifold approximation and projection
vWF	von Willebrandt factor
WAP	whey acidic protein
Wnt	wingless and Int-1

Introduction

1.1 The Extracellular Matrix

Colloquially, it is stated that the fundamental building blocks of any organism are its cells. However, this assertion, while partially accurate, does not encapsulate the complete truth. The ECM, a proteinaceous hydrogel between the cells, constitutes a substantial portion of most animals. It is a diverse compartment that forms the structural framework of the body and serves as the immediate environment and niche of cells and tissues.

In all multicellular organisms, the ECM is essential as it mediates cell adhesion and enables communication between epithelia and tissues [Adams and Engel, 2007, Ellis and Tanentzapf, 2010, Hynes, 2009]. As such, it is a very dynamic compartment that not only differs between different organs but also changes depending on signaling cues and developmental state [Bonnans et al., 2014, Daley et al., 2008, Epp et al., 1986, Ingber, 2006]. While the extracellular compartment is visibly conspicuous in plants and fungi due to their rigid cell walls, the connective tissue of animals is less visible and has not been researched extensively.

1.1.1 The Structure and Components of the Extracellular Matrix

In animals, the ECM is mostly made up of a proteinaceous hydrogel containing large fibrillar proteins such as collagen, elastin and fibronectin, highly modular glycoproteins such as thromospondin (TSP), as well as remodeling factors such as different proteases [Hynes and Naba, 2012, Naba et al., 2012, Petrov et al., 2023, Wang et al., 2021].

The ECM can be categorized into two primary layers. The predominant layer is the fibrous interstitial matrix (IM), composed mainly of fibrillar collagens, including collagen types I, II, III, V, XI, XXIV and XXVII [Bella and Hulmes, 2017]. The IM is the main contributor to the structural properties of the ECM, filling the interstitial space between tissues and providing a stable yet elastic substrate for the development of organs and body structures. Owing to its thickness, the IM can act as a cushion that mitigates tissue damage and enables elastic deformation, which, for example, is especially important in vertebrate lung tissue [Dolega et al., 2018, Stauber et al., 1990, Tschumperlin, 2015]. In addition, the sponge-like composition of the IM enables it to serve as a reservoir for many secreted molecules and proteins, including growth factors [Flaumenhaft and Rifkin, 1991, Hynes, 2009, Taipale and Keski-Oja, 1997].

Conversely, the second main layer of the ECM, the basal lamina (BL), lines the adjacent epithelial tissues. It forms a thin layer, which not only facilitates direct cell-matrix interactions but also creates a tight meshwork of fibers, which functions as a stable substrate for cells. At the molecular level, the BL consists of integrins and laminins that mediate cell-matrix interactions (also called hemidesmosomes)

and several network-forming collagens (Collagen Types IV, VI, VIII, X) [Knupp and Squire, 2005], as well as proteoglycans such as nidogen [Zhou et al., 2022]. Due to its close association with the adjacent cells, the BL has a major impact on the differentiation and formation of epithelia and specific cellular niches [Lu et al., 2012, Noh et al., 2023].

With few exceptions, the ECM of vertebrates can be considered a product of the mesoderm, as most components are secreted by fibroblasts, osteocytes and other mesodermal cell lines, whereas the endo- and ectodermal cells play only a supporting role in the secretion of ECM factors [Brown and Sanders, 1991, Huss et al., 2019, Lin et al., 2020, Ruggieri et al., 2021]. Surprisingly, it was shown that, in early vertebrate embryos, primordial germ cells (PGC) are involved in the secretion of the initial ECM [Huss et al., 2019].

Both the BL and the IM undergo extensive post-translational modifications and remodeling by the action of many regulatory factors [Adams, 2023, Leeming et al., 2011]. The most prominent of these are the members of the matrix metalloprotease (MMP) family, which not only process the structural elements of the ECM, but also many of the secreted enzymes and growth factors [Ziegler et al., 2021].

1.1.2 The Matrisome: Diverse Proteins with common domains

While the proteins constituting the ECM are very diverse in form and function, they share a conserved set of protein domains. These domains are arranged in various domain architectures across different ECM proteins and are encoded on different exons with extensive shuffling between the different ECM genes during the course of evolution [Adams and Engel, 2007, Engel, 1996, Hohenester and Engel, 2002, Patthy, 1999, Whittaker et al., 2006]. Many of these domains facilitate

fiber- or network-formation either by β-sheet layering as in the case of laminins and fibronectins or by cross-linking as in the case of collagens and elastins (briefly reviewed by Ahn et al. [2018]). Other common ECM domains (e.g. lectin domains) specifically bind to and connect these fibers or bind to polycarbohydrates, which make up 5-20% of the ECM dry weight depending on the tissue [Detamore et al., 2005, Lundell et al., 2004, Pesciaroli et al., 2013]. The presence of these fiber-forming domains across diverse protein families leads to the self-assembly of composite fibers and the creation of heterogeneous fibers. The order in which the different components are incorporated determines the size and unique physical properties of the resulting fibers [Ahn et al., 2018].

In addition to these structural domains, ECM proteins commonly feature other domains with various functions:

I Cell-matrix interaction and signaling domains, such as TSP- or epidermal growth factor (EGF)-like domains.

II Secreted/soluble factor binding domains, for instance, Sushi/short consensus repeat (SCR)/complement control protein (CCP)- or von Willebrandt factor (vWF) domains.

III Domains whose functions remain unknown, like sea urchin sperm-coating protein, enterokinase, agrin (SEA)- or C1r/C1s, Uegf, Bmp (CUB)-like domains.

In total, 127 ECM domains that commonly occur in ECM proteins have been described from mouse and human samples [Naba et al., 2012]. Many structural components of the ECM consist of a combination of these domains and can reach lengths of thousands of amino acids (aa). One prominent example of this is the

aptly named human Polydom/SVEP1 proteins, which is 3571 aa long and consists of 49 domains.

Conversely to these diagnostic ECM domains, 49 non-ECM domains, which are never found within ECM proteins of vertebrates have been described that can help to exclude false positives [Naba et al., 2012]. Taken together, these domains enable the identification of ECM proteins based on their sequences and domain annotations through a process of positive and negative selection [Hynes, 2012, Naba et al., 2012].

1.1.3 The Diverse Biological Functions of the ECM

The best-studied function of the ECM is its role as a filler material that separates epithelia and occupies empty spaces and body cavities. As such, it provides a foundational structure upon which the intricate shapes of different organs are built while also serving as a substrate for cells. Cells closely interact with their niche, receiving vital signals for homeostasis and differentiation [Ellis and Tanentzapf, 2010, Pennings et al., 2018]. Thus, the physical properties and composition of the substrate can significantly influence the direction of cell differentiation [Noh et al., 2023]. Furthermore, cell-matrix connections contribute to establishing the basal-apical organization of epithelial tissues [Lee et al., 2021, Varshney et al., 2015]. It is important to note that these cell-matrix connections are not permanent; they can be loosened to facilitate cell migration on top of (or within) the ECM. Alternatively, cell “migration” can be achieved by the movement of the ECM itself [Aufschnaiter et al., 2011, Zamir et al., 2008, 2006].

For the correct migration and differentiation of cells, it is important that morphogen gradients are developed and maintained, and the ECM plays multiple important

roles in the maintenance of these gradients. It can act passively as a reservoir for secreted morphogens, or it can limit the diffusion of factors due to variations in interactions and binding [Flaumenhaft and Rifkin, 1991, Hynes, 2009, Taipale and Keski-Oja, 1997]. Some ECM factors can actively modulate or degrade signaling factors, thus influencing the extent of gradient formation [Chang and Sun, 2014, Kakugawa et al., 2015, Lommel et al., 2018, Norman et al., 2016, Piccolo et al., 1997, Ziegler et al., 2021]

ECM proteins can also act as signaling molecules themselves. As mentioned above many ECM proteins possess EGF-like and other growth factor-like domains, which are thought to be cleaved off to function as growth factors or growth factor receptor inhibitors [Davis, 1990, Engel, 1989]. Taken together, the high diversity of ECM proteins and their domains suggests that the ECM can influence a multitude of cellular functions, many of which have yet to be investigated, especially since new ECM components are identified regularly.

1.1.4 The Identification of Matrisomes

There are three commonly used methods to identify matrisome components *in silico*.

I Sequence similarity

II gene onthology (GO)-term based search

III Domain-based filtering.

All of these methods in the end rely on the conservation and shuffling of domains between ECM proteins. Sequence similarity and GO-term analysis require some previous knowledge; either in the form of a known ECM factor or in the form of

GO-term annotation of a novel protein.

The only bioinformatic method currently able to identify novel and unknown ECM factors is *de novo* domain annotation followed by domain-based filtering based on the diagnostic and exclusionary ECM domains and exclusion of the known non-ECM domains. This method is sufficient to find nearly all ECM proteins in vertebrates [Naba et al., 2012]. In invertebrates this method has been used successfully to identify many novel and species-specific proteins [Davis et al., 2019, Teuscher et al., 2019]. However, it must be noted that this approach only identifies proteins which do contain at least one of the vertebrate ECM domains. This means that specialized proteins that only contain invertebrate-specific domains must be added manually.

For cnidarians a few specific domains have been identified in the past, especially those that are part of the cnidocyte-specific matrisome. One example is cnidocyst-specific cysteine-rich-domain (CxxxCxxxCxxxCxxxCC) (CRD) domain [Adamczyk et al., 2008, Bentele et al., 2019, Tursch et al., 2016]. While these domains often co-occur with classical ECM domains, the example of the cnidarian poly proline protein (CPP) protein family, known from *Hydra* cnidocysts includes members that do not (Bentele [2018] and unpublished results).

Thus far, the field of ECM research has therefore focused on well-established model systems with well-annotated resources. Comprehensive matrisome annotations are available for model organisms such as mouse [Naba et al., 2012]), human [Naba et al., 2012], cow [Listrat et al., 2023], zebrafish [Nauroy et al., 2018], fruit fly [Davis et al., 2019], quail [Huss et al., 2019], and the nematode worm *C. elegans* [Teuscher et al., 2019] (Table 1.1). However, while annotating proteins in these species is relatively straightforward, characterizing the functions of these proteins

is exceptionally challenging due to the intricacies of their body plans and the ECM's specialization in various organs.

The lack of fully annotated datasets for "basal" animals currently blocks the efforts to use them as model systems for functional ECM research. However, these animals and particularly cnidarians, offer a significant opportunity for molecular and functional ECM research. With their comparatively simple body plans and more accessible ECM material, they hold great promise as models for advancing our understanding of the ECM.

Table 1.1: **Statistics of Published Matrisomes.**
In silico matrisomes have been published for a number of model organisms that possess well annotated sequence datasets. While *in silico* matrisomes are easy to create for these species, complete experimental matrisomes are more complicated to obtain for animals with complex body plans. In basal animals, the experimental definition is easier but the annotations of the components is not as straight-forward due to a lack of annotated datasets.

	Core Matrisome	ECM Associated	Species specific	Total Matrisome	Reference
H. sapiens (human)	278	778		1057	Naba et al. [2012]
M. musculus (mouse)	274	824		1098	Naba et al. [2012]
D. rerio (zebrafish)	333	669		1002	Nauroy et al. [2018]
B. taurus (cow)	287	736		1032	Listrat et al. [2023]
C. japonica (quail)	238	458		706	Hu et al. [2020]
D. melanogaster (fruit fly)	34	279	328	641	Davis et al. [2019]
C. elegans (nematode)	238	481	12	731	Teuscher et al. [2019]
S. mediterranea (planarian)	141	115		256	Cote et al. [2019]
	?	?		1038	Sonpho et al. [2020]

1.2 Cnidarians as Extracellular Matrix Models

Cnidarians might be considered ideal model systems for ECM research. Their ECM, often referred to as the mesoglea, is so prominent that it is name-giving for the group of jellyfish. The colloquial gel-like "jelly" in "jellyfish" refers to the ECM, which is the primary structural component of the bell in most medusa forms. In fact, the mesoglea constitutes the majority of the weight and body shape of many

cnidarians.

The cnidarian body possesses only two cell layers (the ecto- and gastroderm), separated by the mesoglea. Since only one cell layer separates it from the surrounding medium, the mesoglea is exceptionally accessible for ECM research. In addition, the mesoglea can be mostly or entirely cell-free, depending on the species, which makes the extraction of pure ECM material in large quantities feasible [Day and Lenhoff, 1981, Tucker et al., 2011]

The most researched cnidarian model systems are the hydrozoan freshwater polyp *Hydra* and the anthozoan starlet sea anemone *Nematostella vectensis* Stephenson [1935] (hereafter called *Nematostella*). In both of these species, comprehensive morphological and molecular analyses of individual aspects of ECM biology have been conducted.

1.2.1 What is known about the Cnidarian ECM

The mesoglea of *Hydra* has been investigated molecularly for many years with individual components being identified by biochemistry, cloning [Barzansky et al., 1975, Fowler et al., 2000, Sarras et al., 1994, Zhang et al., 2002] and bioinformatic analysis [Bentley and Adams, 2010, Özbek et al., 2010]. Similarly, various ECM components have been isolated from different cnidarians as reviewed by Tucker and Adams [2014]. Several researchers have managed to isolate cell-free samples of *Hydra* mesoglea [Barzansky et al., 1975, Day and Lenhoff, 1981].

A study by Lommel et al. [2018] investigated the molecular composition of isolated, cell-free *Hydra* mesoglea samples using mass spectrometry and identified only 37 unique proteins. Veschgini et al. [2023] -with whom I collaborated to investigate

the *Hydra* mesoglea- independently identified 183 unique proteins from isolated *Hydra* mesoglea; however again only 36 of these proteins could be unequivocally assigned ECM identity based on thier domain architecture. Together, this suggests that *Hydra* has an unexpectedly low number of ECM proteins, compared to other animals (Table tab:published_matrisome).

Due to the accessibility of the mesoglea, it has been used in the past to investigate functional questions which would be impossible to answer in more complex animals. For example, to show the connection between cell migration and mesoglea dynamics [Aufschnaiter et al., 2011], the interdependence of wingless and Int-1 (Wnt) signaling with mesoglea composition and its physical properties [Veschgini et al., 2023]. Another well-researched aspect of cnidarian ECM biology is the role of matrix metalloproteases. In *Hydra*, these enzymes play a crucial role in various developmental processes, including foot epithelia differentiation, foot regeneration [Kumpfmüller et al., 1999, Leontovich et al., 2000], ECM turnover [Yan et al., 1995], and Wnt cleavage [Ziegler et al., 2021].

It was also possible to show that in *Hydra* both the ectoderm and the gastroderm secrete ECM factors, whereby e.g. laminins are secreted by the gastroderm and collagens are secreted by the ectoderm [Zhang et al., 2007].

In recent years, the starlet sea anemone *Nematostella* has been established as an additional cnidarian model for developmental and evolutionary biology [Layden et al., 2016]. In contrast to *Hydra*, which hatches as fully formed polyps, *Nematostella* possesses a complex life cycle (Fig 1.1). Their eggs initially develop into spherical planula larvae with a size of 200-300 µm, followed by a settlement and metamorphosis phase during which the larvae elongate and develop four tentacles [Babonis et al., 2016, Hand and Uhlinger, 1992]. Interestingly, the elongation and

tentacle formation processes resemble the gut formation in vertebrates [Technau, 2020, Technau and Steele, 2011, Ikmi et al., 2020].

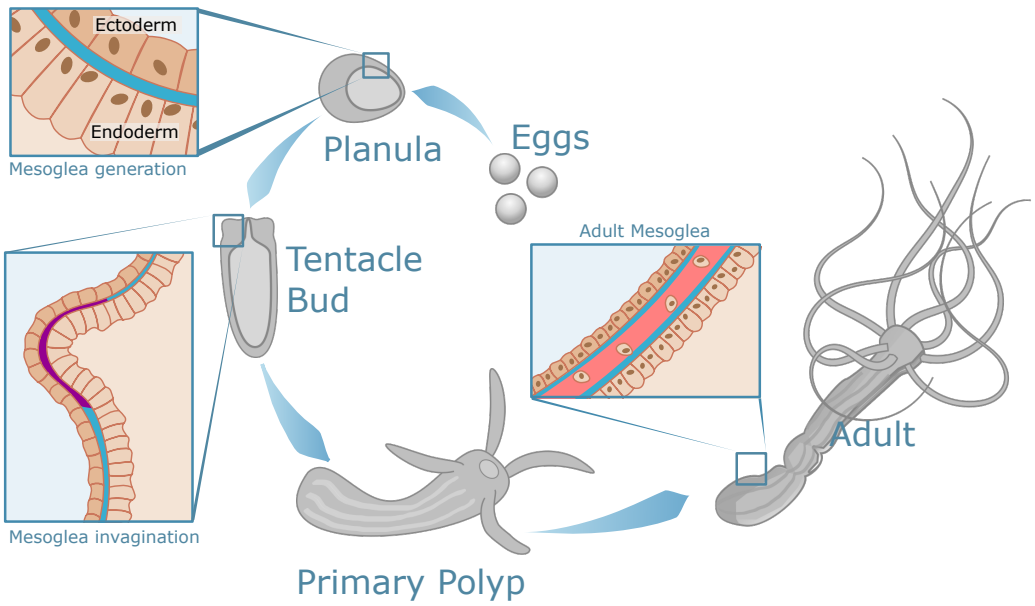


Figure 1.1. The Complex Life Cycle of *Nematostella*. Like many cnidarians, *Nematostella* goes through a complex life cycle. Its eggs first develop into free-swimming planula larvae (approximately 200 µm in length). At this stage, the mesoglea is initially visible. The larvae elongates and develops tentacle buds. During tentacle formation, the mesoglea needs to be reorganized to enable tissue invagination. After approx. 10 days, the 0.5 cm long primary polyps are fully formed. These possess four tentacles and can remain dormant until they catch prey. The up to 15 cm long adult stage is reached in approximately 3 months if the primary polyps are fed.

Before and during metamorphosis the anemones develop mesenteries, which are internal structures and folds within the body wall that increase the surface area of the gastroderm. Additionally, these mesenteries harbor large retractor muscles, which facilitate body contraction and the oocytes, which develop within the mesoglea and are released from it during spawning [Eckelbarger et al., 2008, Moiseeva et al., 2017, Tucker et al., 2011].

Upon completion of metamorphosis, the embryos reach the primary polyp stage in which they can remain for several months [Ikmi et al., 2020]. The development from egg to primary polyp typically spans about 10-14 days at room temperature [Stokkermans et al., 2022]. If primary polyps are able to feed, they continue development into mature polyps, characterized by numerous tentacles and complex mesenteries. These adult *Nematostella* can grow to a length of 15 cm under laboratory conditions.

The mesoglea plays an integral role in various developmental processes of *Nematostella*. Each morphological transformation necessitates the synthesis and remodeling of this critical structure.

The structure of the *Nematostella* mesoglea during development has been investigated by electron microscopy by Tucker et al. [2011]. However, the earliest life stage investigated were 2-week old primary polyps. They noted that the mesoglea morphology and thickness differs depending on the body region. In the tentacles of early primary polyps, the mesoglea thickness was measured to be only 0.2-3 µm, with fused BL and nearly no discernable IM. Conversely, in other body regions, the mesoglea was notably thicker, ranging from 1.5-2.5 µm with an 80-100 nm thick basal lamina. Additionally, the ultrastructure of the mesoglea was noted to be asymmetric, featuring thicker, bundled fibers closer to the gastroderm and thinner fibers towards the epidermis.

The oocytes are not the only cell type found in the mesoglea of *Nematostella*, electron micrographs of the adult mesoglea sporadically reveal the presence of cells within the ECM [Frank and Bleakney, 1976, Tucker et al., 2011]. These cells have been termed ’amoebocytes,’ drawing parallels to the amoeboid cells found in the connective tissue of animals, in fact, Tucker et al. [2011] described these amoebo-

cytes as ”remarkable fibroblast-like”. Nevertheless, the true nature of these cells remains to be investigated. They are observed in close proximity to the IM but do not make contact with the BL [Tucker et al., 2011].

All studies in cnidarians confirmed that while their basic body is comparatively simple, the mesoglea exhibits striking parallels to the complex ECM of vertebrates. The morphological resemblances between the mesoglea to the connective tissues of vertebrates were noted many years ago and the molecular composition suggested that the components of the mesoglea and the vertebrate ECM are similar [Barzansky and Lenhoff, 1974, Burnett, 1973, Lenhoff and Loomis, 1961, Sarras et al., 1994, 1991, Takeuchi et al., 2018, Zhang and Sarras, 1994]. Recent comparative genetic analyses have revealed that Cnidarians represent the first animals to possess all fundamental components that constitute an animal ECM, thereby establishing them as the first evolutionary lineage with a genuine ECM and showing the evolutionary conservation of the ECM components[Özbek et al., 2010].

The publication of the *Nematostella* genome by Putnam et al. [2007] was particularly relevant for this finding. It revealed that even among anthozoans, the earliest branch of cnidarians, all core ECM genes are present. Subsequent investigations have confirmed both the functionality of these genes and their integral role within the mesoglea.

The conserved set of ECM proteins that are found in all animal ECMs are called the core adhesome. They encompass not only structural proteins such as collagens, but also cell-matrix adhesion factors, such as laminins, as well as ECM remodeling factors like MMPs [Hynes, 2012]. Together, these elements constitute the essential building blocks for organizing an ECM and enabling adhesion of epithelia. Some of these factors already evolved in unicellular organisms (Fig 1.2)

and sponges already possess rudimentary epithelia demonstrating that their extracellular compartment is nearly completely present [Adams et al., 2010, Fahey and Degnan, 2010, Leys et al., 2009, Leys and Riesgo, 2012].

Throughout evolution, the core adhesome has undergone expansions, modifications, and in some cases, losses of components to adapt to the diverse body plans and lifestyles of different animals. However, even in the seemingly simple cnidarians, the ECM plays a pivotal role in numerous crucial processes. This underscores the far-reaching implications for understanding the broader functions of ECM proteins across the entire animal kingdom.

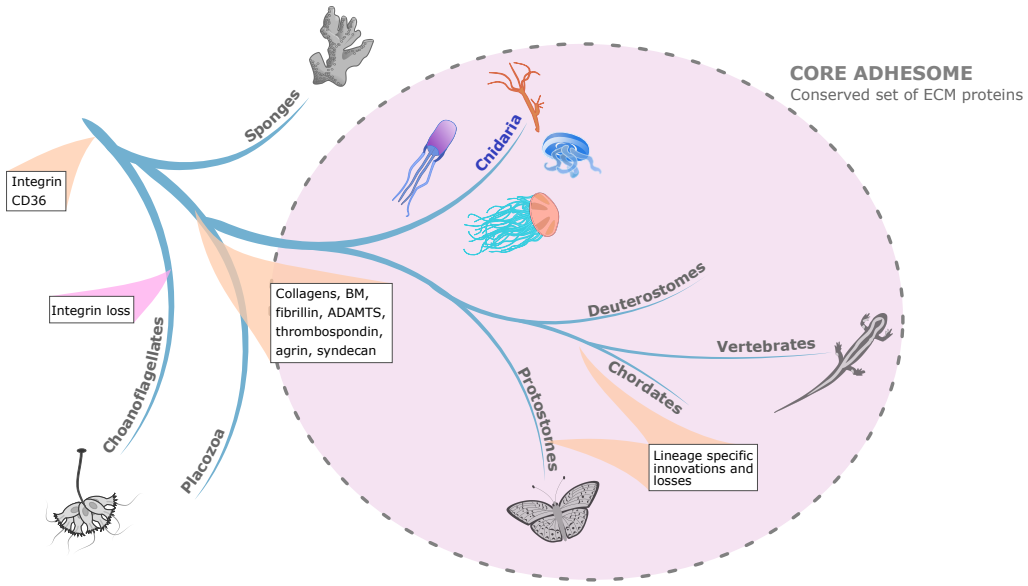


Figure 1.2. The Evolution of the Core Matrisome. A number of ECM proteins can be found in coloninal choanoflagellates and sponges, which possess rudimentary ECMs however cnidarians are the earliest branch of the metazoans that possess the full core adhesome which consists of a set of conserved ECM proteins (Image modified from Özbek et al. [2010].

1.2.2 The Specialized Extracellular Matrix of Cnidarians

In addition to the internal mesoglea, cnidarians evolved two kinds of specialized ECM. Firstly, they evolved bioadhesive matrices that allow the polyp stage to adhere to inorganic surfaces. In these, the structural ECM proteins were adapted to connect the cells to the minerals of the substrate while ECM remodeling factors enable a reversal of attachment [Davey et al., 2021, 2019, Rodrigues et al., 2016, Young et al., 1988]. These adhesive matrices are further specialized in the skeletal organic matrix of corals where the ECM components not only attach to a substrate but also act as nucleation points for the further deposition of minerals [Drake et al., 2020, Levy et al., 2021]. This leads to the formation of coral reefs which host the vast majority of marine life.

The second type of specialized ECM, produced by the cnidarian stinging cells, are the stinging capsules (cnidocysts). These are highly complex, harpoon-like organelles made up of a venom-filled capsule, a long flexible tubule, spines and complex lid structures. ECM-like proteins feature heavily in the building blocks of these capsules [Balasubramanian et al., 2012]. Prominently, the capsule wall is made up in large parts of minicollagens [Adamczyk et al., 2008, Balasubramanian et al., 2012, Holstein et al., 1994, Kurz et al., 1991]. These possess a short collagen repeat flanked by CRDs [Engel et al., 2001, Meier et al., 2004, Milbradt et al., 2005, Pokidysheva et al., 2004, Tursch et al., 2016]. In combination with other likely ECM derived proteins such as cnidoin [Beckmann et al., 2015] and CPP proteins [Bentele et al., 2019, Bentele, 2018], these structural proteins enable the capsule wall to withstand immense internal pressures and at the same time store kinetic energy through its elasticity. The spines and presumably the lid structures (although there are no detailed molecular studies of this structure) are

likewise formed by ECM derived proteins [Hellstern et al., 2006]. While the tubule is membranous, recent studies have shown that lectin-type proteins and a chondroitin matrix are involved in its formation and structure [Garg et al., 2023]. Finally, the venom components include many metalloproteases and protease inhibitors [Jouiaei et al., 2015].

Most cnidarians possess multiple capsule types, which fulfil different functions in prey killing and/or attachment. For the evolution of which the innovation of ECM-like components played an important role [David et al., 2008]. Adult *Nematostella*s possess three capsule types: Spirocysts, basitrichous haplonemas, and microbasic mastigophores [Adamczyk et al., 2010]. However, single cell analysis and gene expression analysis have revealed an exclusively embryonal subtype of spirocysts and an additional capsule type as well as a gland-like cnidocyte subtype characterized by a specific nematocyte expressed protein 8 (nep8) [Columbus-Shenkar et al., 2018, Steger et al., 2022].

Aim of this Thesis

In summary, *Nematostella* combines a several favorable aspects required for the functional investigation of the ECM in various biological processes as well as for the investigation of ECM evolution. The major hurdle for this is the lack of reliably annotated component databases.

Within this thesis, I therefore aim to achieve the following:

I Identify and annotate the *in silico* matrisome of *Nematostella vectensis*.

II Investigate the molecular dynamics of the identified ECM proteins throughout

the larvae, primary polyp, and adult life stages.

III Identify the cell lineages and cell types that secrete the ECM proteins.

IV Put the *Nematostella* matrisome into an evolutionary context.

Together, these aims provide a much needed overview and annotation basis for the cnidarian matrisome and shed further light on the multitude of questions that can be investigated using cnidarian ECM models.

Materials and Methods

2.1 Animals

2.1.1 *Nematostella* Culture

Nematostellas were cultured and spawning was induced as described in Stefanik et al. [2013].

2.2 Molecular Biology/Biochemistry

2.2.1 Mesoglea Decellularization

About 500 000-1 000 000 larvae (3 days post fertilization (dpf)) and primary polyps (10 dpf), and four adult animals were collected for each biological replicate. The adult polyps were cut open along the oral-aboral axis using a scalpel to open the gastroderm for decellularization. All samples were incubated in 0.5 % *N*-lauryl sarcosinate for 5 min and then frozen in liquid N_2. After thawing at room temperature (rt), the samples were transferred to ddH_2O using a 70 µm sieve for larvae and pri-

mary polyps. The mesogleas were decellularized in ddH_2O by repeated pipetting with a flamed glass pipette and frequent water changes. The progress of decellularization was checked repeatedly by phase contrast microscopy at 60x magnification using a Nikon 80i microscope. As a final quality control, a few sample mesogleas were stained with 4′,6-diamidino-2-phenylindole (DAPI) for 10 min in phosphate-buffered saline (PBS) to visualize residual cells or nematocysts. Decellularized mesogleas were then picked individually for further analysis.

2.2.2 Immunocytochemistry

Larvae at 3 dpf, primary polyps at 10 dpf, and small adult polyps were collected and left to relax at 27 °C in direct light for 30 min. A solution of 7 % $MgCl_2$ in seawater was slowly added to anesthetized the animals for 20 min. Fixation was performed with Lavdovsky's fixative (50 % ethanol, 36 % H_2O, 10 % formaldehyde, 4 % acetic acid) for 30 min at rt. After fixation the samples were incubated in 150 mM Tris pH 9.0, 0.05 % Tween 20 for 10 min at rt and then at 70 °C for 10 min with subsequent cooling to rt as suggested by Inoue and Wittbrodt [2011]. This was followed by three washing steps of 10 min each in PBS, 0.1 % Tween-20, PBS, 0.1 % Tween-20, 0.1 % Triton-X100, and PBS, 0.1 % Tween-20. Primary antibody incubation (rabbit anti-NvLaminin, rat anti-NvCollagen2, rat anti-NvPanCollagen) was performed at 1:100 in 0.5 % milk powder overnight at 4 °C. The samples were washed three times for 10 min in PBS, 0.1 % Tween-20 and incubated with secondary antibodies at 1:400 for 2.5 h at rt. Prior to mounting on object slides with Mowiol, DAPI was added at 1:1 000 for 30 min. Decellularized mesogleas were transferred onto a microscopy slide lined with liquid blocker (PAP pen, Newcomer Supply). The mesoglea of larvae and primary polyps were carefully stuck onto the

slide using an eye lash. The staining protocol followed the whole mount immunocytochemistry protocol with 10 min fixation and only one washing step per wash to avoid washing the mesogleas off the slide. The antibodies were incubated on the slides in a petri dish with a wet paper towel to prevent evaporation. Images were acquired with an A1R microscope at the Nikon Imaging Facility Heidelberg. Further image processing was performed with FijiImageJ v1.53t.

2.3 Mass Spectrometry

2.3.1 Protein Extraction for Mass Spectrometry

Sample preparation for mass spectrometry (MS) was carried out as described in Veschgini et al. [2023] with the following changes: For larvae and primary polyps, at least 150 decellularized mesogleas were collected per sample. For adults, 4 mesogleas were collected. The mesogleas were solubilized in 1 M dithiothreitol (DTT) for 30 min at 90 °C and immediately extracted using standard trichloroacetic acid (TCA) precipitation. The resulting pellets were dissolved in 10 mM DTT 1 % sodiumdodecyl sulfate (SDS) TCA pellets were resuspended in 50 µL 1 % SDS, 50 mM
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 8.5. DTT (56 °C, 30 min, 10 mM in 50 mM HEPES, pH 8.5) was used to reduce the disulfide bridges in cysteine containing proteins. After completing the SP3 protocol [Hughes et al., 2014], trypsin (sequencing grade, Promega) was added in an enzyme to protein ratio 1:50 for overnight digestion at 37 °C. Peptides were labelled with tandem mass tag (TMT)-10plex Isobaric Label Reagent (ThermoFisher) according to the manufacturer's instructions [Werner et al., 2014]. Each sample was treated separately.

A control run was performed to ensure equally mixed peptide amounts based on the MS signal in each run. Following this, samples were combined for TMT9plex.

2.3.2 LC–MS/MS

MS was performed as described in Veschgini et al. [2023] with the following adaptations: a Nanospray Flex™ ion source was used at a spray voltage of 2.3 kV, a capillary temperature of 320 °C, and a fly time of 10 ms. Data dependent acquisition (dda) was performed with the resolution of the Orbitrap set to 35 000 and a fill time of 120 ms. The dynamic exclusion time was set to 30 s. IsobarQuant [Franken et al., 2015] and Mascot (v2.2.07) were used to process the acquired data, which was searched against the *Nematostella vectensis* NV2 (wein_nvec200_tcsv2) protein models (SIMRbase) in addition to a sequence set containing common contaminants and reversed sequences. The following modifications were included into the search parameters:
Carbamidomethyl (C) and TMT10 (K) (fixed modification), Acetyl (Protein N-term), Oxidation (M) and TMT10 (N-term) (variable modifications). For the full scan (MS1) we set a mass error tolerance of 10 parts per million (ppm) and for MS/MS (MS2) spectra it was set to 0.02 Da. Further parameters were set: Trypsin as protease with an allowance of maximally two missed cleavages; a minimum peptide length of seven amino acids; at least two unique peptides were required for protein identification. The false discovery rate on the peptide and protein level was set to 0.01. In total, 5056 proteins passed the quality control filters.

2.4 Electron Microscopy

2.4.1 Sample Preparation at EMCF Heidelberg

Nematostella larvae and primary polyps were immobilized in $MgCl_2$ (7 %) for 10 min before fixation in 2.5 % glutaraldehyde and 0.1 % tannic acid in 0.1 M cacodylate buffer. Fixation was done at 4 °C overnight. The samples were centrifuged to a pellet in an Eppendorf tube, washed 4 times in cacodylate buffer on ice, and fixed with 1 % osmium in 0.1 M cacodylate buffer for 1 h on ice. After washing twice in buffer and twice in water for 15 min each, the samples were further stained with 1 % uranylacetate in water overnight at 4 °C, followed by four rinses in water. Dehydration was achieved in 10 min steps with increasing acetone concentrations of 30 %, 50 %, 70 %, 90 % and 2 times 100 %, respectively. Spurr resin was infiltrated in steps of 45 min with 25 %, 50 %, 75 % Spurr resin in acetone, followed by 100 % Spurr resin overnight. After a final change of the resin, the samples were polymerized inside Eppendorf tubes at 60 °C for 24 h. The blocks were cut in 70 nm thin sections using a Leica UC6 ultramicrotome (Leica Microsystems Vienna) and collected on pioloform coated mesh grids. The sections were post stained using uranylacetate and lead citrate and imaged on a JEOL Jem-1400 electron microscope (JEOL, Tokyo) operating at 80 kV and equipped with a 4K TemCam F416 (Tietz Video and Image Processing Systems GmBH, Gautig).

2.4.2 Sample Preparation at Innsbruck Medical University

Cryobased electron microscopy (EM) was performed essentially as described in [Garg et al., 2023, Holstein et al., 2010]. Intact larvae and primary polyps were sub-

jected to rapid cryofixation (high-pressure freezing, freeze-substitution) followed by epoxy resin embedding and optional section post staining with heavy metals (UranyLESS EM stainTM (a mixture of lanthanides); lead). periodic acid–Schiff (PAS)-cytochemistry at EM-level was performed at rt according to standard protocols [Thiery, 1967] as previously described [Böttger et al., 2012]. Unstained sections were treated with periodic acid (1 %, 30 min.), thiocarbohydrazide (0.2 2 h), and silver-proteinate (SP: 1 %, 30 min).

2.5 Short Hairpin RNA Knockdown

The short hairpin RNAs (shRNA) was transcribed according to the MEGAshortscript^{TM} T7 Transcription Kit (Thermo Fisher Scientific) and then purified according to the Direct-zol Kit (Zymogen). Each shRNA (1 µg/mL) sample was mixed with FITC fluorescent dye 1:5 and injected into fertilized *Nematostella* eggs at the EMBL. The sequences of all used shRNAs can be found in the Supplementary Data 6.3.2. The eggs were allowed to develop for 3 days and then transported to the lab where they were imaged once per hour using a SMZ 1500 with a DS Fi3 Camera for 7 days in constant light at 22-23 °C in a 12-well plate with 2 mL *Nematostella* medium. To prevent condensation, the lids were coated with 50 % ethanol, 1 % Tween-20, and then air-dried 1 day in advance of imaging. The embryos were transferred to a new dish in intervals of 2 days. Subsequently, they were fixed and imaged on 10 dpf.

2.6 Bioinformatic MEthods

2.6.1 Accessing Published Datasets

Excel tables of the published matrisomes of *Homo sapiens* , *Mus musculus*, *Bos taurus*, *Coturnix japonica*, *Danio rerio*, *Caenorhabditis elegans*, *Schmidtea mediterranea*, and *Drosophia melanogaster* were downloaded from the MatrisomeProject webpage. Where RefSeq- or Uniprot, WormBase or FlyBase IDs were provided, these were downloaded. In the cases where multiple IDs were indicated for one gene, the longest version was chosen. In the cases where the provided IDs were either not included or deprecated, the gene name was searched in combination with the organism using an Entrez query (see Source Code 6.3.1).

2.6.2 *In Silico* Matrisome Prediction

Domains for the protein models of the SIMRbase *Nematostella vectensis* NV2 (wein_nvec200_tcsv2) transcriptome (SIMRbase/Nematostella) were annotated de novo using InterProScan [Jones et al., 2014]. The proteins were then filtered and sorted using a custom python [van Rossum, 2010] script (Source Code 6.3.1) by filtering ECM inclusion and exclusion domains [Naba et al., 2012] using the pandas package. In addition to *Nematostella*, I predicted *in silico* matrisomes for the following species.

CHOANOFLAGELLATA: *Monosiga brevicollis* UP000001357, *Salpingoeca rosetta* UP000007799; PORIFERA: *Amphimedon queenslandica* UP000007879, *Ephydatia muelleri* [Kenny et al., 2020]; CTENOPHORA: *Mnemiopsis leidyi* (Mnemiopsis Genome Project), *Beroe ovata* (unpublished, RyanLab Webpage), *Pleurobrachia*

bachei [Moroz et al., 2014]; PLACOZOA: *Trichoplax adhaerens* (UP000009022); CNIDARIA: *Aurelia aurita* [Gold et al., 2019], *Clytia hemispherica* (Marimba), *Exaiptasia diaphana* [Oakley et al., 2016], *Hydra vulgaris* (*Hydra* Genome Portal), *Stylophora pistillata* (NCBI *Stylophora pistillata* Annotation Release 100), *Acropora digitifera* [Shinzato et al., 2021], *Renilla reniformis* [Kayal et al., 2018], *Calvadosia cruxmelitensis* [Ohdera et al., 2019], *Cassiopea xamachana* [Ohdera et al., 2018], *Morbakka virulenta* [Khalturin et al., 2019], *Acropora cervicornis* [Libro et al., 2013], *Acropora millepora* [Moya et al., 2012], *Acropora muricata* [Tingyu Han and Chunpeng He, 2022], *Hydra viridissima* [Hamada et al., 2020], *Hydractinia echinata* (Hydractinia Genome Portal), *Hydractinia symbiolongicarpus* (Hydractinia Genome Portal), *Montipora aequituberculata* [van de Water et al., 2018], *Montipora foliosa* [Tingyu Han and Chunpeng He, 2022], *Myxobolus honghuensis* [Guo, 2021], *Pocillopora damicornis* [Tingyu Han and Chunpeng He, 2022], *Pocillopora verrucosa* [Tingyu Han and Chunpeng He, 2022], *Porites asteroids* [Kenkel et al., 2013], *Xenia spec.* , *Acropora hyacinthu*s [Wright et al., 2015].

2.6.3 *Nematostella* Matrisome Annotation

The domains predicted using InterProScan were transferred to the raw protein sequences using the pandas package and biopython [Cock et al., 2009]. Signal peptides were predicted using (SignalP-6.0) [Teufel et al., 2022] (prediction-cutoff = 50 %) and DeepLoc-2 [Thumuluri et al., 2022]. The latter was also used to infer the cellular localization of proteins. OrthoFinder-2.5.4 [Emms and Kelly, 2019] was employed to predict Orthogroup (OG)s using the *in silico* matrisomes for all species and the published matrisomes. The combined information was used to manually compare the proteins, to published protein families, and domain architectures.

2.7 Single Cell Data Analysis

The single cell transcriptome analysis of matrisomal genes was performed using a published dataset. Further details can be found in the respective publication by Steger et al. [2022].

Results

In order to develop a complete picture of *Nematostella*'s matrisome throughout development, I combined *in silico* and experimental techniques into a workflow (Figure 3.1). I developed a pipeline to automatically predict in *in silico* matrisomes from protein sequence databases and created a manually curated *in silico* matrisome of *Nematostella*. I supplemented this curated matrisome with evolutionary comparisons and OG annotation using published matrisomes and newly predicted data for many basal animals.

I chose larvae (3 dpf), primary polyps (10 dpf), and adults as representative life stages encompassing the major morphological changes throughout development and investigated these changes using fluorescence- and electron microscopy. I developed an ECM isolation protocol for the chosen life stages and assayed the composition of the mesoglea of these stages using MS. In combination with an analysis of published single cell data, this allowed me to determine the expression patterns and overall composition of the mesoglea at different life stages. Finally, I knocked down some differentially expressed genes to identify developmental phenotypes caused by ECM factor depletion.

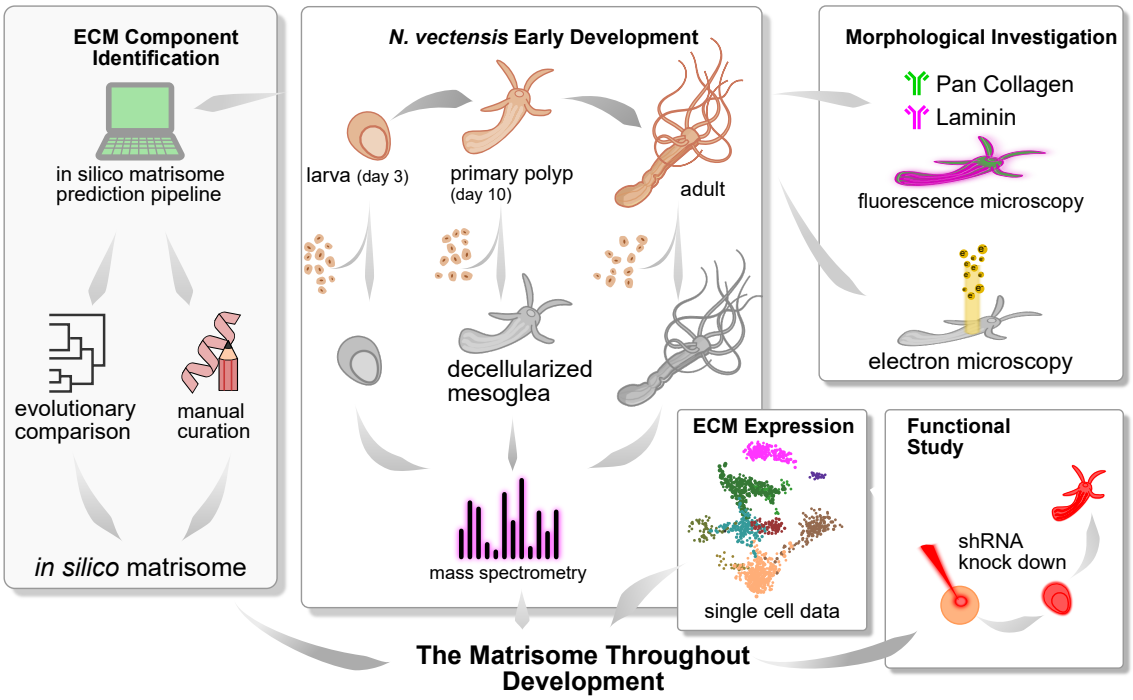


Figure 3.1. The Results of this Thesis. For this study, *in silico* methods and experimental studies were combined to identify the compositional changes of the mesoglea throughout development. A prediction pipeline for *in silico* matrisomes was written and used to predict an evolutionary dataset of matrisomes. The *Nematostella in silico* matrisome was manually curated and annotated. Mesoglea from different life stages was extracted and analyzed by MS. Through combination of the *in silico* and MS datasets with published single cell expression data the complete description of the developing matrisome of *Nematostella* was achieved. These studies were extended by morphological investigation and functional knock downs of candidate ECM factors.

3.1 The Morphological Changes of the Developing Mesoglea

The mesoglea has not been described in literature at earlier life stages than primary polyps at 10 dpf. Which means that the most interesting morphological changes during the larvae-primary polyp transition have not been investigated thoroughly.

Hence, I first investigated the changes in morphology during the larvae-primary polyp transition using immunofluorescence and EM. While EM protocols were already established for *Nematostella* [Tucker et al., 2011], no antibodies against ECM factors were available at the time. Therefore, I designed antibodies against epitopes in previously known matrisome factors of the BL and IM with the aim to enable costaining of both ECM subcompartments.

I chose NvLaminin (NV2t025737001.1) and NvCollagen-4 (NV2t021411001.1) as both of these factors should be highly abundant in the BL. As a marker for the IM, I chose to target fibrillar collagens. Initially, I designed antibodies against an individual collagen (Collagen2c | NV2t010959003.1), however, stainings performed with this yielded very weak staining with strong background signal (Figure 3.2 B). Therefore, I designed an additional PanCollagen antibody against a conserved region of the fibrillar collagens known at the time (Suppl. Figure 6.1). This antibody would stain the entire mesoglea as it targets the collagens of the IM, as well as, the network forming collagens of the BL. To confirm the specificity of the antibodies to the different sublayers of the mesoglea, I first stained adult *Nematostella*s to take advantage of their thicker mesoglea (Figure 3.2).

As predicted, the NvLaminin signal localizes to thin layers lining the basal sides of the gastroderm and ectoderm but is not present in the central region of the mesoglea (Figure 3.2 C). The staining was very specific to the BL and did not yield background staining. However, the NvLaminin antibody was sensitive to the fixative. It only stained the BL when the animals were treated with Lavdowsky's fixative and blocked with milk powder, whereas fixation with 5 % performaldehyde (PFA) led to a weak staining of neurons in early life stages. NvCollagen-4 staining was likewise detected in the BL, yet in comparison to NvLaminin the signal was

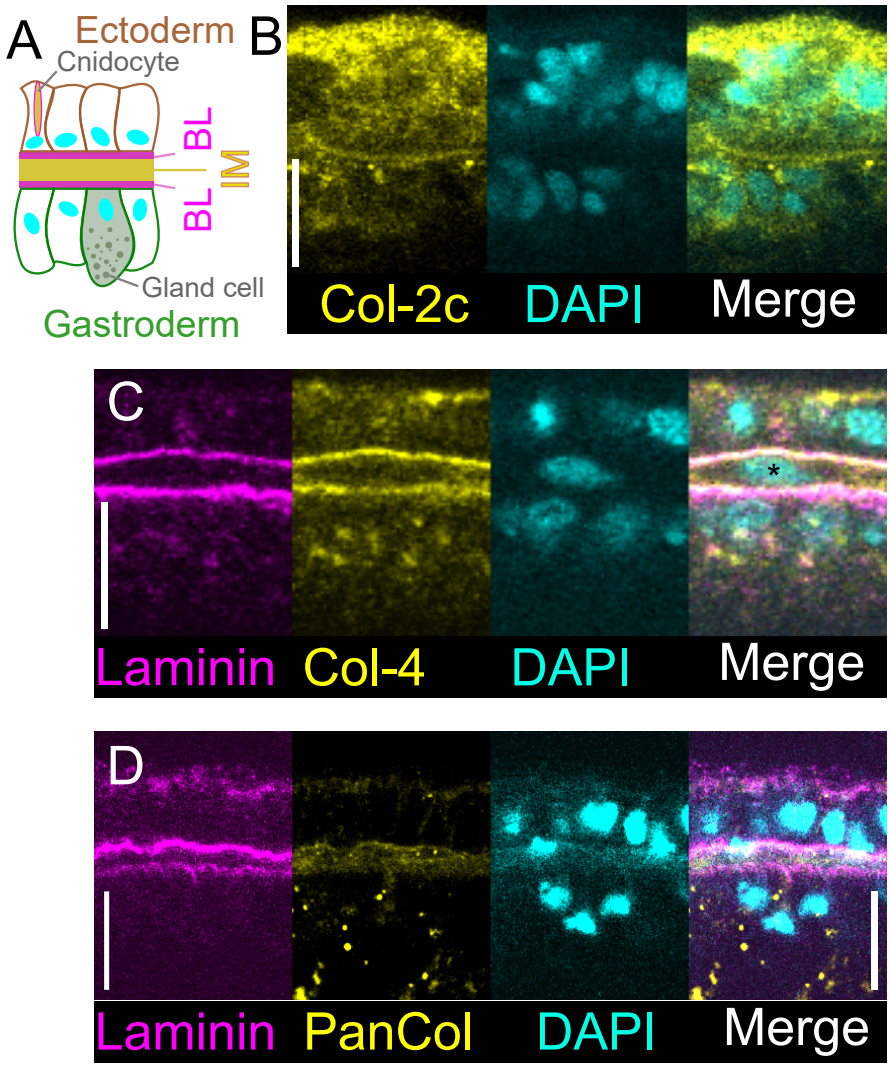


Figure 3.2. Immunofluorescence Staining Using Different ECM-Targeted Antibodies. Schematic visualization of the body wall of *Nematostella*. The ectoderm harbors cnidocytes while the digestive glands are found in the gastroderm. The mesoglea separates these epithelia and consists of two BLs lining the cells and a central IM. (B) An antibody against Collagen-2c (`NV2t010959003.1`) shows weak staining of mesoglea material (likely BL) but is incompatible with BL costaining. (C) Laminin and Collagen-4 antibodies selectively stain the BL. An amoebocyte (*) is visible in the IM. (D) The PanCollagen antibody targeting a conserved region of all fibrillar collagens shows selective albeit weak staining of the whole mesoglea and is compatible with Laminin costaining. (Scale bars = 10 µm)

generally weaker and yielded additional unspecific background signal (Figure 3.2 C).

In contrast to the BL antibodies, the NvCollagen2c antibody did not yield any IM-specific signal and was abandoned after multiple trials (Figure 3.2 B). The NvPanCollagen antibody on the other hand yielded signals localized throughout the mesoglea spanning both the BL and IM (Figure 3.2 D). In combination with the Laminin antibody this allowed me to distinguish between the different mesoglea sublayers. In addition, NvPanCollagen localized to vesicles/puncta in the cytoplasm of the gastrodermal cells. Notably, the staining of NvPanCollagen was not consistent in all body parts. While the Laminin staining is detected throughout the whole body, the PanCollagen signal was decreased in the tentacles (Figure 3.3).

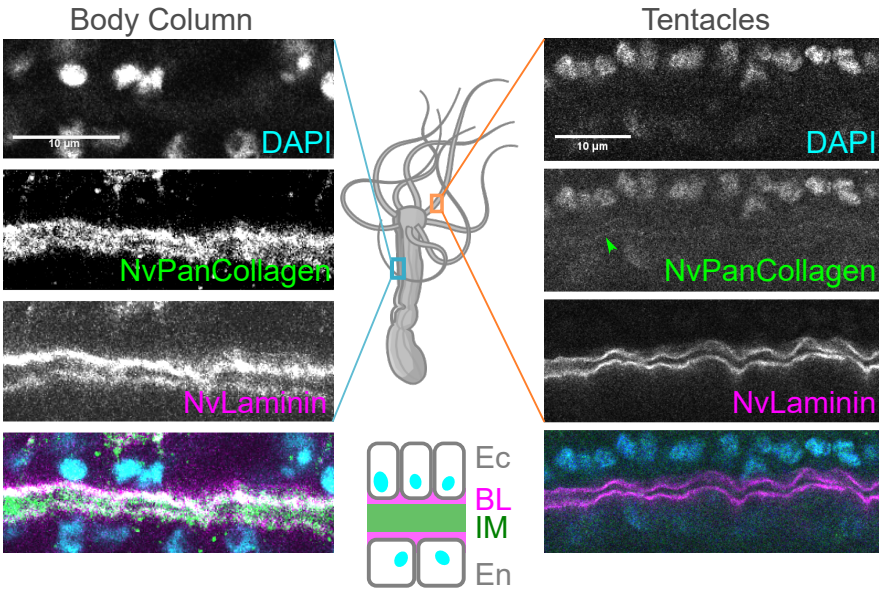


Figure 3.3. Body Region-Dependent Staining of the Interstitial Matrix An adult *Nematostella* polyp was stained for NvLaminin and NvPanCollagen to visualize the different sublayers of the mesoglea. NvLaminin selectively stains the BL in the body column and the tentacle region. NvPanCollagen stains the whole mesoglea in the central body column. However, it only yields faint signal in the BL in the tentacles

After establishing that the antibodies specifically stain the different layers of the mesoglea, I was interested in the morphological features of the mesoglea throughout development. For this purpose, I stained larvae and primary polyps using the antibodies and collaborated with *Stefan Hillmer* and *Chalotta Funyama* from the EMCF Heidelberg and *Michael Hess* and *Stefan Reibl* from the University Innsbruck to investigate the ultrastructure of the mesoglea by EM.

The mesoglea of larvae is 0.4-1 µm thick and mostly consists of BL (Figure 3.4). In the sporadically occurring thicker regions, a central fibrous layer is visible. In the regions where septs are forming and the gastroderm is invaginated the mesoglea is expanded into a triangular shape (Figure 3.4 B-E). This triangle is lined with a Laminin-containing BL and filled with fibrous IM-like material. The central region is not stained by NvPanCollagen (data not shown) suggesting that its compounds are not detected by this antibody.

In contrast to this seemingly BL-based mesoglea of larvae, whole-mount staining of primary polyps revealed similar patterns to those found in adults (Figure 3.2 and Figure 3.3). The NvLaminin stained BL of primary polyps is separated by a 4-6 µm thick IM. Despite this, the two sheets of the BL are not separated completely and seem to conjoin in some regions. This IM can partially be stained using Immunofluoresence with NvPanCollagen which results in the staining of individual fibers. Additionally, I was able to confirm the presence of cells in the IM by DAPI staining.

The NvLaminin staining of primary polyps further revealed an aboral structure where the BL feathers out into the ectoderm (Figure 3.5 C). Close examination seemed to suggest that a pore is formed in the mesoglea at this point (the existence of this structure was first noted by *Soham Basu*, EMBL).

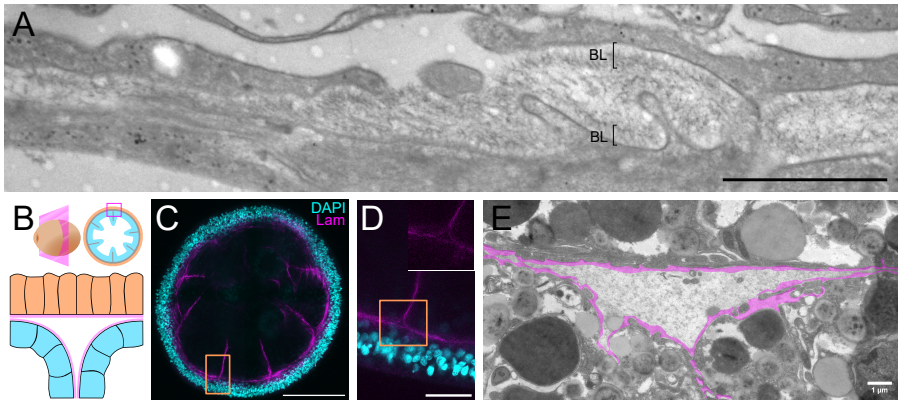


Figure 3.4. The Basal Lamina-based Larvae Mesoglea. (A) Electron micrograph of the larvae mesoglea (scale bar = 1 µm). The mesoglea of larvae has a thickness of 0.41 µm. It is mostly made up of BL material with a thin, sporadic fibrillar central region. (B) The mesoglea thickens into a triangular shape at the septa branching points. (C) Larvae stained with DAPI (cyan) and NvLaminin (magenta). DAPI hardly reaches the gastroderm. The mesoglea and developing septs are shown by Laminin staining. (D) the triangular expansion of the mesoglea at the septa branch is lined with Laminin. The central region does not show any signal. (E) Electron micrograph of the septa branch. The BL is overlaid in magenta. The central region of the septa triangle is filled with IM-like fibrous material. (Scale bars: A + E 1 µm, C+D 10 µm)

EM confirmed the separation between a thick fibrous IM and two thin electron dense BLs lining the epithelia (Figure 3.5 B). Interestingly, at the primary polyp stage, the BL can be subdivided into an electron-passable lamina lucida towards the epithelia and an electron denser lamina densa (Figure 3.5 B).

While the fibers of the IM are in general loosely organized there is a general trend for them to orientate along the oral-aboral axis parallel to the direction of the mesoglea of the primary polyp (Figure 3.5 D-F'). However, in one electron micrograph showing the region close to the aboral pore it was revealed that the IM fibers in the aboral tip are very densely arranged perpendicularly to the direction of the mesoglea.

After establishing that the mesoglea of *Nematostella* changes morphologically dur-

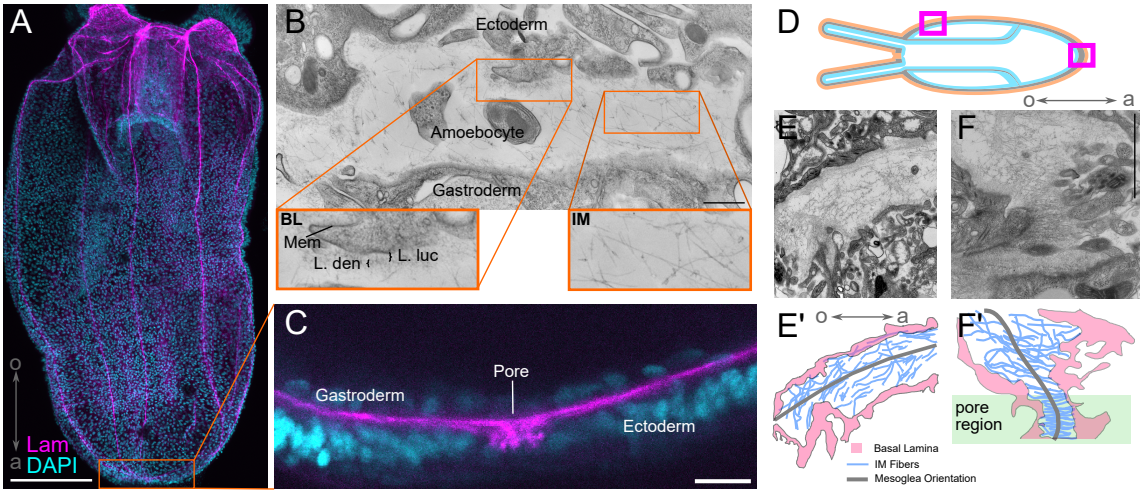


Figure 3.5. Mesoglea Morphology of Primary Polyps (A) Whole mount staining of a tentacle bud stage stained with DAPI and NvLaminin-γ. (B) The mesoglea is 4-10 µm thick and contains a thick IM made up of loose fibrous material. The BL lining the cell membrane (Mem) is separated into an electron passable Lamina lucida (L.luc) and an electron dense Lamina densa (L.den). At the aboral end (C) the mesoglea flares out into the ectoderm layer and forms a small pore. (D-F') The IM fibers in the body column are loosely orientated along the oral-aboral (o-a) axis following the overall orientation of the mesoglea. In the aboral region adjacent to the pore the IM fibers are dense and perpendicular to the mesoglea orientated along the o-a axis Scale bars = A 100 µm, B, E + F 1 µm, C 10 µm.

ing development, I wanted to identify its molecular changes during this process. To be able to follow the expression and protein abundance of these ECM factors I first needed to identify and classify the entire matrisome of *Nematostella*.

3.2 The *Nematostella In Silico* Matrisome Consists of 843 Proteins

Due to the huge evolutionary distance between *Nematostella* and the (mostly vertebrate) reference matrisomes available at the time of writing, and the low percentage of well annotated *Nematostella* proteins, I decided to use domain-based

filtering to identify the complete *in silico* matrisome of *Nematostella* as a basis for all further investigations.

To do this, I established a bioinformatic pipeline designed to automatically predict the *in silico* matrisome from a given set of protein models (Source Code 6.3.1). In this pipeline, InterProScan is used to predict and annotate protein domains. These predicted domains are then used to identify ECM candidates based on their domain architecture according to a list of signature domains proposed by Naba et al. [2012]. These candidates were then filtered for the occurrence of ECM exclusion domains to negatively select cytoplasmatic contaminants.

The resulting raw *in silico* matrisome of *Nematostella* consists of 1810 proteins (See Supplementary Data). However, closer examination revealed many of these sequences to be isoforms and severely fragmented partial sequences. I reviewed each protein model, which resulted in a curated *in silico* matrisome of 843 sequences (See Supplementary Data), equaling to approximately 3 % of the total proteome of *Nematostella* [Artamonova and Mushegian, 2013].

3.2.1 Annotation of the *Nematostella* Matrisome

To be able to use this *in silico* matrisome as a reference list for future cnidarian matrisomes, I attempted to annotate the curated matrisome automatically by BLAST (basic local alignment search tool) against the NCBI non-redundant and Swissprot protein databases. However, these algorithmic annotations returned a high ratio of unknown or undefined protein definitions and often resulted in obvious misannotations, e.g. sequences containing a single vWF type A domain were classified as “Collagen alpha chain”. Due to these problems of automated annotation, I decided to manually annotate and classify the proteins of the curated matrisome.

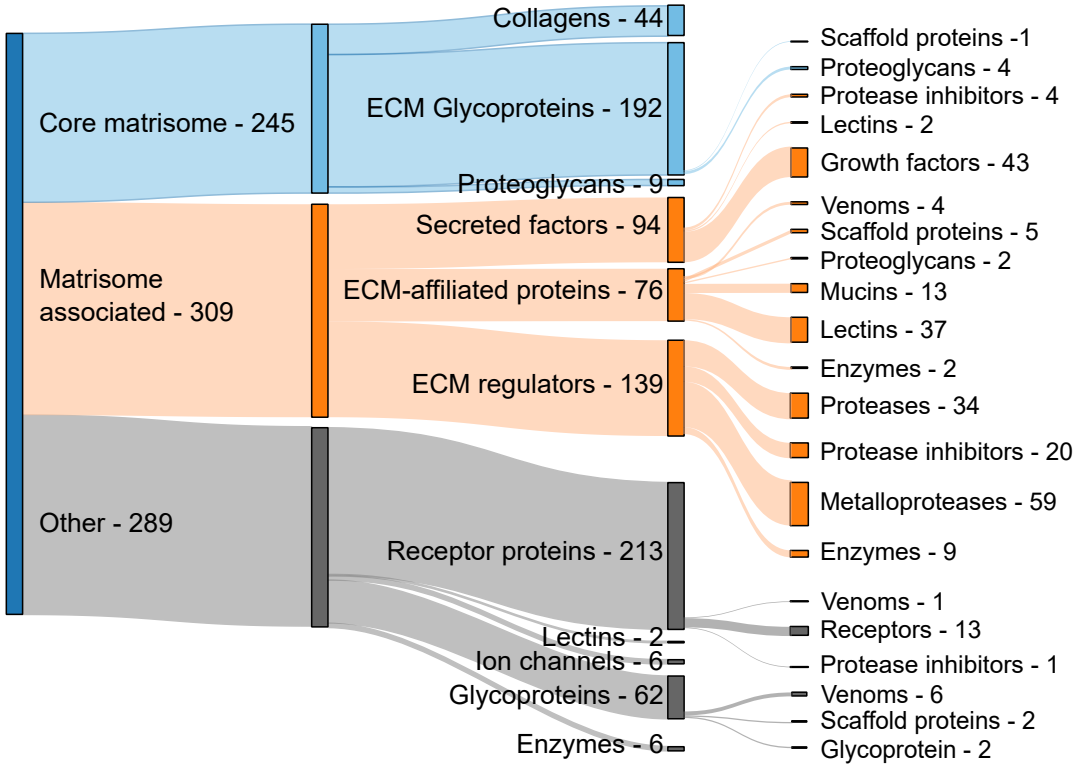


Figure 3.6. Manual Classification of Matrisome Proteins. The predicted *in silico* matrisome sequences were curated and all isoforms and fragmentary sequences were excluded. The resulting sequences were grouped into 245 components of the core matrisome and 309 Matrisome-associated proteins. Within these ECM divisions the proteins were classified into different protein families. In addition to these traditional ECM divisions, we categorized 289 proteins that are potentially interacting with the ECM.

Each of the identified potential ECM protein sequences were compared to published sequences and protein families by full and domain-specific sequence alignment and manual comparison. In addition to this, I predicted the presence of signal peptides and the likely subcellular localization of the proteins by applying SignalP and DeepLoc. This enabled us to categorize the matrisome proteins into protein families (Figure 3.6) and classify them into 245 core matrisome proteins consisting of collagens, glycoproteins and proteoglycan, 309 matrisome-associated proteins comprising secreted/soluble factors (such as growth factors), ECM regula-

tors (such as MMPs) and protease inhibitors) and ECM associated proteins (e.g. lectins and mucins). Finally, we identified a further 289 proteins that possess ECM domains but have been excluded from the classical ECM divisions. This group mostly consists of membrane bound receptors that have a role in cell signaling and adhesion such as cadherins and cell adhesion, molecule (CAM)s.

3.2.2 The Matrisome is Secreted by the Gastroderm

The ECM-secreting cells in vertebrates are of mesodermal origin. Therefore, I wanted to know which cell types contribute to the secretion of the mesoglea in *Nematostella*. To gain insights into the expression of the matrisome, I collaborated with *Alison Cole* from the University of Vienna, who mapped the genes transcribing the identified matrisomal proteins onto the recently published *Nematostella* single cell transcriptome comprising different life stages and all cell types [Steger et al., 2022].

> The protein IDs can be converted into their corresponding gene IDs as follows:
>
> `NV2t000245001.1` → `NV2.245`

To assess the general expression pattern of the matrisome, we grouped the matrisomal genes according to our annotation as core matrisome, ECM associated and Other. For each of these divisions we calculated the average expression of ECM proteins in each cell type and visualized this using uniform manifold approximation and projection (UMAP)s (Figure 3.7).

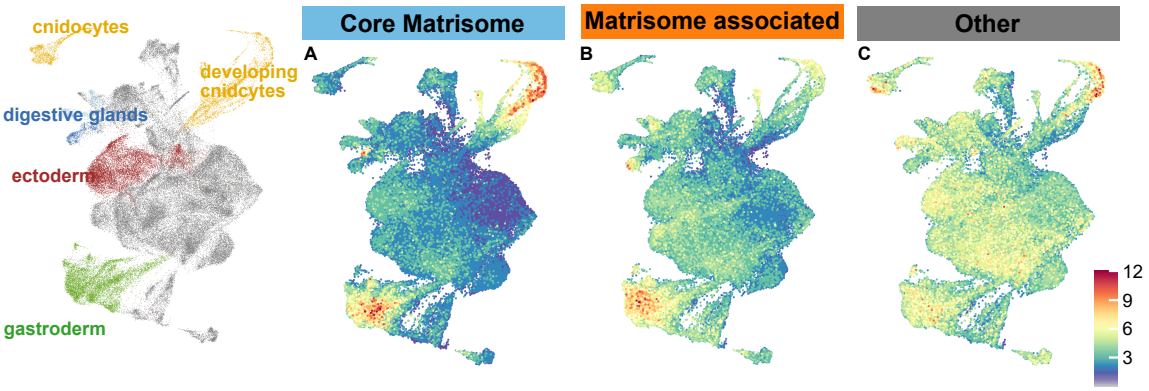


Figure 3.7. The Gastroderm and the Cnidocytes are Hot Spots of ECM Gene Expression. The single cell expression of the ECM proteins classified by core matrisome, matrisome-associated and other was averaged and visualized in UMAP single cell expression plots. (A) The majority of core matrisome genes is expressed in the gastroderm, the cnidocytes and a small cluster of glandular cells. (B)The matrisome-associated genes follow a similar trend but have a broader expression distribution. Some members of this division are also expressed in glandular cells, however, in a different subpopulation than the core matrisome. (C) The members of the "other" division are widely expressed in all cell types with two prominent expression clusters in the developing and mature cnidocytes

The average expression of the core matrisome is distributed across three distinct clusters (Figure 3.7 A). Predominantly, it is expressed by cells within the gastroderm. The second major cluster of core matrisome gene expression is detected in cnidocytes, which can be attributed to minicollagens and other ECM-type structural proteins included in the core matrisome list. The third, relatively small cluster of core matrisome expression is detected in a specific subtype of neuroglandular cells which were predicted to have digestive gland identity.

Matrisome-associated genes exhibit a similar expression pattern, primarily being expressed in the gastroderm (Figure 3.7 B). Similar to the core matrisome, a subtype of neuroglandular cells contributes to the expression of matrisome-associated genes. This cluster is distinct from the neuroglandular cells expressing core matrisomal genes. The expression of matrisome-associated genes in the cnidocytes is

less pronounced.

The category termed “Other” exhibits widespread expression across all cell types, possibly owing to the ubiquitous expression of cell surface receptors (Figure 3.7 C). Nevertheless, two elevated expression clusters in the developing and mature cnidocytes suggest that that a significant portion of the genes classified as ”other” are cnidocyte-specific. Interestingly, these genes are cnidarian-specific indicating a specialized function in nematocyst formation.

Considering that the ECM factors expressed in cnidocytes may not play a significant role in the mesoglea, it appears that the ECM in *Nematostella* is primarily produced by the gastroderm, with a minor contribution from neuroglandular cell types. In contrast, the ectoderm does not seem to be involved in ECM deposition.

3.2.3 The Cell Lineage-specific ECM

After identifying these three broad cell clusters of ECM component expression I was interested to identify the subset of ECM proteins secreted by the different cell lines. I especially wanted to annotate the cnidocyte-specific matrisome as it is a hallmark of cnidarians and I wanted to detect if it is completely distinct from the overall mesoglea matrisome or if certain factors are shared.

In cooperation with *Alison Cole* a filter algorithm was developed to identify all differentially expressed genes in all cell clusters (Figure 3.8 and Suppl. Table 6.3.2).

This more detailed view of differential gene expression confirms that the gastroderm is the primary contributor to the secretion of ECM proteins while the ectoderm seems to play a negligible role in matrisome synthesis. Neither the developing nor the adult ectoderm secrete more than a few factors and only one core

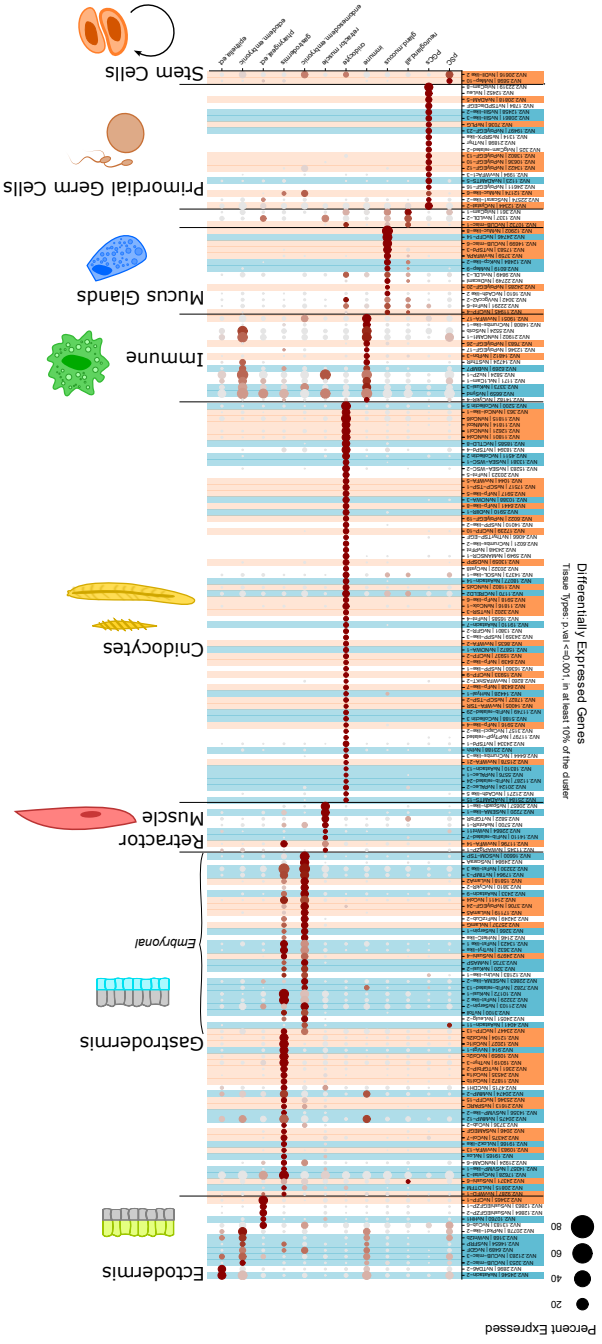


Figure 3.8. Differentially Expressed Matrisome Genes All Matrisomal genes which are differentially expressed in at least one cell type color coded by ECM division (Core matrisome = orange, matrisome associated = blue). The matrisome is expressed by the gastroderm and the cnidocytes. Smaller expression clusters can be found in the primordial germ cells (PGCs), mucus gland cells and immune-like cells. The adult and developing gastroderm have diverging expression patterns. The embryonal gastroderm expresses less core matrisome genes and more matrisome-associated genes. The three clusters of the ectoderm (pharyngeal, embryonic and adult ectoderm) do not express many ECM factors.

matrisome factor is among the top five differentially upregulated in the pharyngeal ectoderm (NvProperidin-1 | NV2.23465). Interestingly, there is a marked difference between the expression profiles of the embryonic and adult gastroderm. The developing gastroderm expresses many matrisome-associated proteins, such as MMPs and other ECM remodeling factors. However, it does not express many structural components. In fact, of the core matrisome factors only BL factors such as Laminins (NvLamA5 | NV2.17119 and NvLamG | NV2.25737) and Collagen-4 (NvCol4 | NV2.21411) are upregulated in the developing gastroderm. The structural components of the IM are upregulated in the adult gastroderm.

Additional cell types that contribute to the secretion of matrisome factors are mucus gland cells, PGC and an unspecified ”immune” cell type.

As expected, numerous ECM-like proteins are epressed in the developing cnidocytes, due to their essential role in assembling cnidocysts. The precise combination and temporal sequence of these structural factors are crucial for the correct assembly of the capsule. Variations in the developmental trajectories of different capsule types may offer insights into the molecular disparities leading to diverse capsule morphologies. Therefore, the cnidocyte-specific matrisome warrants closer examination.

3.2.4 The Cnidocyte-Specific Matrisome

To comprehensively characterize the cnidocyte matrisome, we extracted all ECM genes expressed in at least 10% of the cnidocyte cell clusters, from putative stem cells to mature cnidocytes. This resulted in 182 genes with differential expression in the cnidocyte subclusters as shown in (Figure 3.9 and Suppl. Table 6.3.2). These included well-known cnidocyte-specific factors, such as minicollagens, novel

outer wall antigen (NOWA) proteins and Collectins. It is worth mentioning that the cnidocyte matrisome is highly likely to contain cnidarian-specific proteins not detected by our generalized *in silico* matrisome identification pipeline. This is due to cnidocyte-specific domains, such as the minicollagen CRD, not yet being included in the domain prediction algorithms and databases.

Given the distinct requirements of different capsule types, we anticipated variations in the expression of these cnidocyte matrisome factors. To address this, we examined the differential expression of ECM genes within the cnidocyte subclusters identified by Steger et al. [2022]. These subclusters not only differentiate between capsule types but also between early cnidoblasts and later maturing and mature cnidocytes. This approach allowed us to track the expression of the cnidocyte matrisome from putative stem cells to fully mature cnidocytes.

The putative stem cells (pSC) of the cnidocyte lineage showed limited expression of matrisomal genes, with early expression of a minicollagen-related protein (NvNCol-like-1| NV2.363). In early cnidocytes, many genes are expressed at lower levels, with a prominent expression of cnidocyte-specific Collectin (NvCnCollectin-2 | NV2.4511)

Comparing different capsule developmental trajectories revealed type-specific gene expression patterns. Spirocysts, for example, were characterized by the near exclusive expression of three Minicollagens (NvNCol-5 | NV2.11802, NvCol-like-3 | NV2.11816, and NvCol-like-4 | NV2.11818) and two exclusive isoforms of NOWA (NvNOWA−2 | NV2.10392 and NvNOWA−4 | NV2.10390). Conversely, nep8-cnidocytes which do not form capsules [Columbus-Shenkar et al., 2018, Moran et al., 2013] do not express particular structural proteins but exhibited high expression of specific genes, such as nep8 (NV2.20322), Properidin (NV2.23465), an

MMP (NV2.20475), and a Kunitz-type serine protease inihibitor (NV2.10172). The mastigophore and haplonema capsules (named cnidocyte 1 and cnidocyte 2 in the single cell data) express a similar array of structural and venom components (Figure 3.9).

Additionally, 83 genes were expressed in maturing and fully matured cnidocytes. Spirocysts, in particular, showed increased expression of specific factors before the developmental trajectories of all cnidocytes merge, and the cnidocytes mature. The mature cnidocytes expressed various ECM factors, including proteins from the ”Other” division of the matrisome, MMPs, and protease inhibitors.

3.3 Domain Architecture and Expression of Selected Protein Families

In this section, I take a closer look at some of the most important proteins families of the matrisome and present their domain architecture based on InterProScan annotation and the respective single cell-expression of each member of the protein family. For these, the genes encoding for the identified proteins were extracted from single cell data [Steger et al., 2022] by *Alison Cole*.

3.3.1 The Collagens of *Nematostella*

Collagens are the main component of the animal connective tissue and occur in diverse forms [Kadler et al., 2007]. Fibrillar Collagens have long uninterrupted collagen triple helix domains consisting of Gly-X-Y repeats. Network forming collagens possess long stretches of interrupted collagen triple helix domains [Czarny-Ratajczak and Latos-Bieleńska, 2000, Kadler et al., 2007]. Finally, short chain

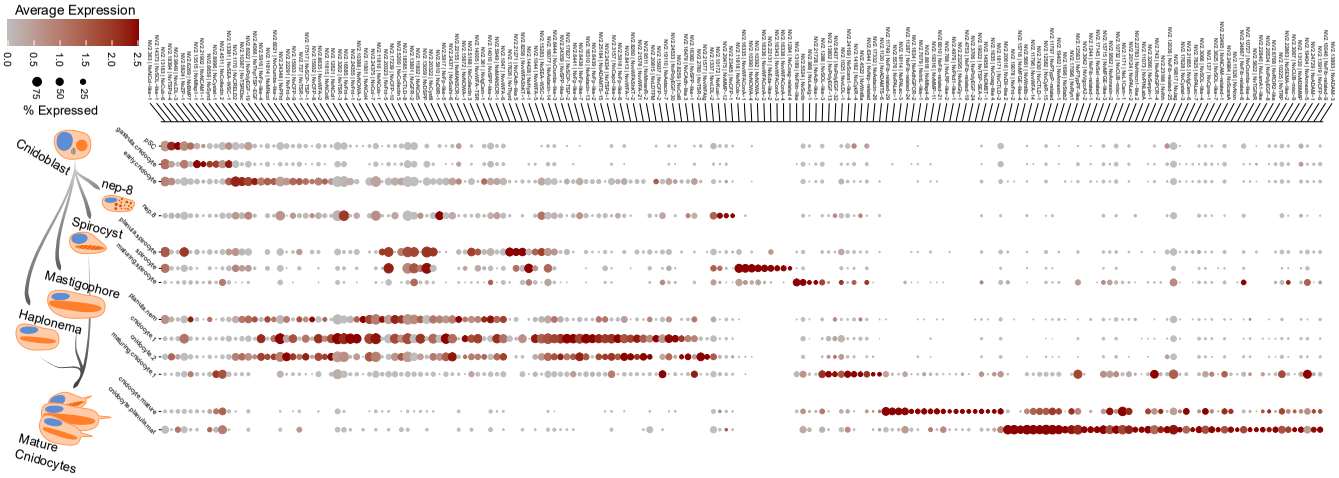


Figure 3.9. The Cnidocytes Express Differential ECM Components. Differential expression analysis of the overall *Nematostella* matrisome revealed 182 genes which are differentially regulated during the cnidocyte developmental trajectory. These can be subdivided into genes expressed during cnidocyte initiation, genes that are expressed for different capsule types and genes that are expressed during the maturation and maintenance of the cnidocysts.

collagens which are often (but not exclusively) found in invertebrates possess only one or very few short collagen triple helix domains [Exposito et al., 2002, 1991, Schmid et al., 1984, Sutmuller et al., 1997, Teuscher et al., 2019].

Within the *Nematostella* matrisome, I identified 44 collagen-related proteins (Figure 3.10). These can be divided into multiple subclasses: I identified 10 fibrillar collagens and classified them according to the *Hydra* nomenclature [Zhang et al., 2007](3 NvCollagen-1 isoforms, 4 NvCollagen-2 isoforms, a NvCollagen-5, NvCollagen-7 and a NvCollagen-18-like). Additionally, I found 6 putative collagen-6-related proteins, 2 with collagen triple helix domains, and 4 characterized by collagen-6-like, repetitive whey acidic protein (WAP)-vWFA domain architecture (IPR002035, IPR008197). The expression analysis in Figure 3.11 shows that most fibrillar collagens are upregulated in the gastroderm of primary polyps and adults, except for BM collagen type IV (NvCollagen-4), which is already upregulated in the endomesoderm of early larval stages. Fibrillar Collagens are also detected at low expression levels in various cell types.

In addition to these vertebrate-type Collagens, I identified several short chain Collagens, including 8 miniCollagen-related sequences, which were identified by the occurrence of a propeptide, miniCollagen CRDs and proline-rich domains. Notably, a miniCollagen protein with an unusually extended Collagen domain comprising extended repetitive Collagen blocks, tentatively designated as mega-miniCollagen (NV2t011814001.1), was found among those. Its genetic locus is in a repetitive region of the genome (data not shown) which makes exon-intron boundary determination difficult. As expected, all miniCollagens, including the mega-miniCollagen, are expressed in cnidocytes, with low expression detected in other cell types. This is likely caused by small cnidoblasts sticking to other cells during the processing

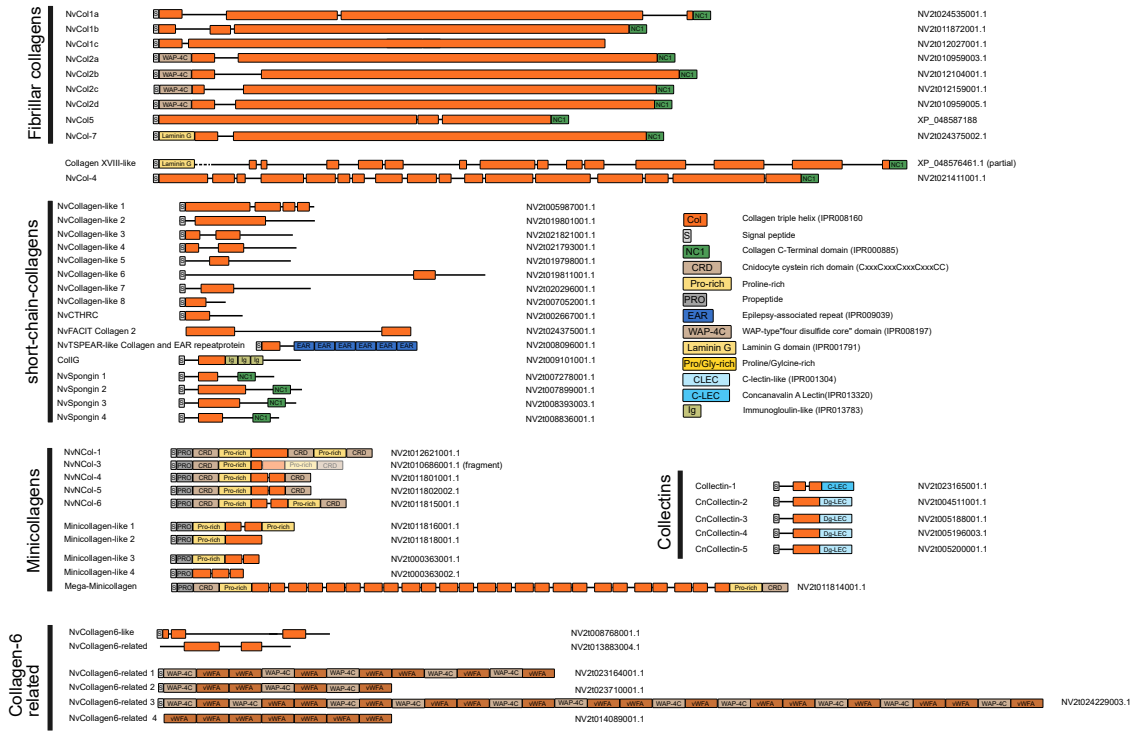


Figure 3.10. Domain Architecture of *Nematostella* Collagens. The Collagens of *Nematostella* were identified by the presence of a 'Collagen triple helix' domain (G-X-Y)>=5.11 fibrillar Collagens were detected. The known cnidarian-specific Collagens are 9 miniCollagens, 1 collagen and immunoglobulin (COLIG) 1 Collectin, and 4 CnCollectins. Six sequences resembling the beaded-filament forming Collagen-6 were identified, however, four of these lack the Collagen domains and were identified based in 'WAP-vWFA' domain architecture. In addition to these, 4 Spongin-like Collagens and 11 Collagen-domain containing proteins with unknown designation were identified.

of the single cells.

Furthermore, a member of the potentially anthozoan-specific COLIG [Knack, 2011] family was found. These proteins possess a short Collagen triple helix and three Immunoglobulin domains. The COLIG gene (NV2.9101) is expressed in the gastroderm and in putative stem cells of larvae and primary polyps. While the function of COLIGs has not been investigated they are speculated to play a role in immune

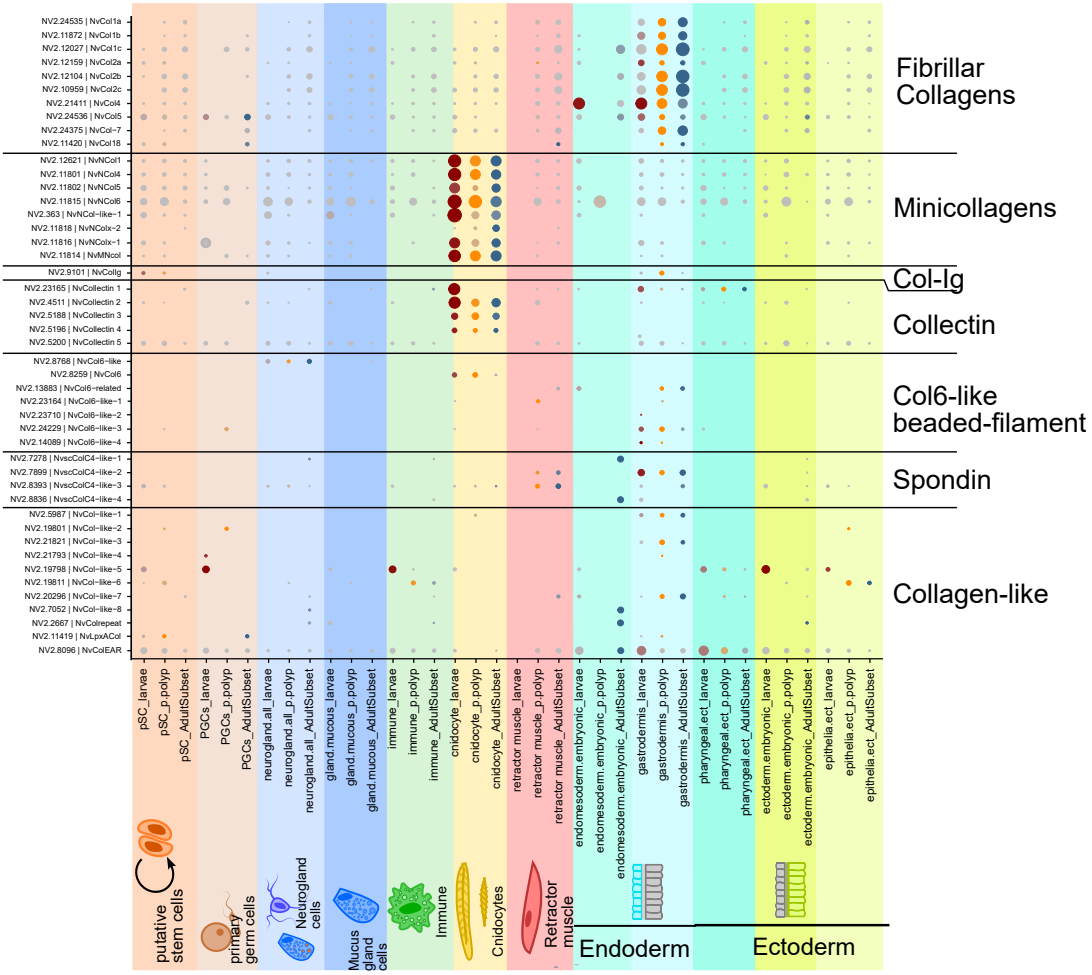


Figure 3.11. Cellular Expression of *Nematostella* Collagens Single cell transcriptome analysis [Steger et al., 2022] of Collagen expression throughout development (analysis by *Alison Cole*). The size of dots represents the cell proportion that expresses the respective gene. The color intensity of the dots shows differential regulation between life stages. While fibrillar Collagen expression was detected at low levels in nearly all cell types they are distinctly upregulated in the gastroderm. All fibrillar Collagens except Collagen-4 were upregulated in the primary polyp stage (yellow) and the adult (blue). Collagen-4 is already upregulated in larvae (red). All miniCollagens and 3 CnCollectins showed cnidocyte-specific expression patterns.

defense by opsonisation of microorganisms [Knack, 2011].

I also identified 5 members of the Collectin family which are characterized by a Collagen triple helix domain followed by a lectin domain. Four of these possess a cnidarian-specific G-lectin and were therefore named cnidarian Collectins (CnCollectins). The majority of these CnCollectins are expressed in cnidocytes. However, NvCnCollectin-5 is expressed at low levels in all cell types except cnidocytes, indicating a more generalized function. The vertebrate-type Collectin (NV2.23165), which possesses a C-type lectin domain, is expressed initially in larval cnidocytes and subsequently gets upregulated in the pharyngeal ectoderm.

Finally, several short-chain Collagens with broadly distributed expression patterns were identified, including four members of the Spongin-like family with a short Collagen triple helix followed by a Collagen C-terminal domain (IPR000885), and 11 Collagen-like proteins, characterized by Collagen triple helices not combined with any other defined domains. Interestingly Collagen-like protein 5 (NV2.19798) is differentially upregulated in larvae. The function of most of these short chain Collagens is currently elusive. However, short chain Collagens like CTHRC (NV2t002667001.1) have been shown to increase cell migration ability and inihibt the expression of fibrillar Collagens [Pyagay et al., 2005].

3.3.2 The Laminins of *Nematostella*

Laminins constitute a significantly to the structure and function of the BL [Beck et al., 1990, Simon-Assmann, 2013] and play a crucial role in the formation of hemidesmosomes, establishing vital cell-matrix connections. The BM-specific antibody shown in Figure 3.2 targets NvLaminin-γ (Suppl. Figure 6.1). We identified 11 proteins characterized by 'Laminin-type EGF' domains (IPR002049), collectively

categorized as Laminin-like proteins (Figure 3.12). These were further classified into 5 Laminin-α-like, 5 Laminin-β-like and 1 Laminin-γ-like, with the latter serving as the antigen for the NvLaminin antibody.

The single cell transcriptome expression analysis of the respective Laminin genes (Figure 3.13) shows that most of them are upregulated in the endomesoderm of larvae (red). Interestingly, NvLaminin-γ (`NV2.25737`) is upregulated in a neuroglandular subtype with a likely neuronal identity (neurogland.all.N1.e3 - Steger et al. [2022]) in larvae, which might explain the neuronal staining of PFA fixed larvae.

3.3.3 The Metalloporteases of *Nematostella*

MMPs play a crucial role in the dynamic restructuring of the ECM, contributing to diverse functions such as ECM reorganization during development and regeneration, as well as the processing of growth factors [Babonis et al., 2019, Leontovich et al., 2000, Sarras, 2017, Shimizu et al., 2002].

Overall, we identified 91 MMPs in the matrisome of *Nematostella*. We classified these based on their domain architecture into the following subclasses: 6 a disintegrin and metalloproteinase (ADAM)s, 15 ADAM with thrombospondin motifs (ADAMTS) + 5 ADAMTS-like, 47 astacins (among them 2 bone morphogenetic protein (BMP)/Tolloid and 18 mephrin-like), 1 mannose-associated serine protease (MASP)-like, 1 M28 like, 1 pappalysin and, finally, 15 miscellaneous MMPs. Each of these subclasses is defined by a specific domain architecture (Figure 3.14).

The single cell expression of the MMPs showed that different MMP subtypes and variations exhibited specific expression profiles across various cell types (Figure 3.15). Notably, theADAM proteases are broadly expressed across different cell

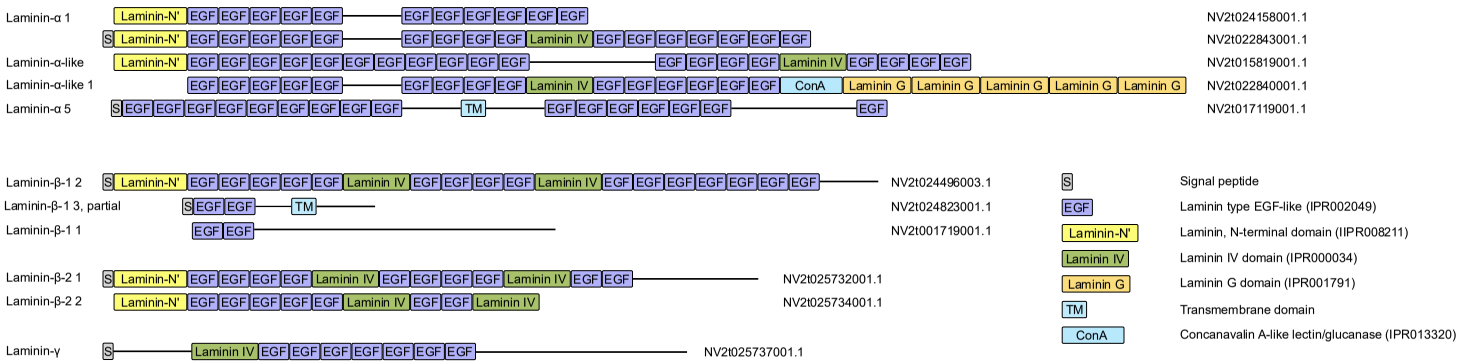


Figure 3.12. Domain Architecture of *Nematostella* Laminins. We identified 11 Laminin-like proteins that share Laminin-type EGF domains. 5 of them were classified as Laminin-α-like, 5 as Laminin-β-like and 1 Laminin-γ-like.

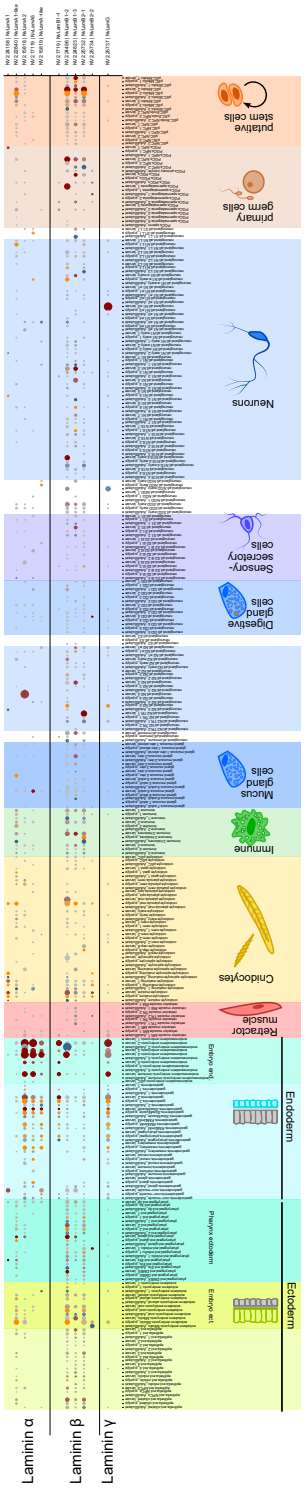


Figure 3.13. Single Cell Expression of *Nematostella* Laminins Single cell transcriptome analysis [Steger et al., 2022] of the Laminin expression throughout development (analysis by *Alison Cole*). The size of dots represents the cell proportion that expresses the gene. Color Intensity of the dots shows differential regulation between life stages (red = larvae, yellow= primary polyp, blue = adult). Most Laminins are upregulated in the larval stage (red dots). Laminin αand Laminin γare primarily expressed in the endomesoderm, while Laminin βhas a broader expression pattern and is upregulated in putative stem cells. Interestingly, Laminin γis highly upregulated in a larval neuron-type cell line (`larva.neurogland.N1.e3`).

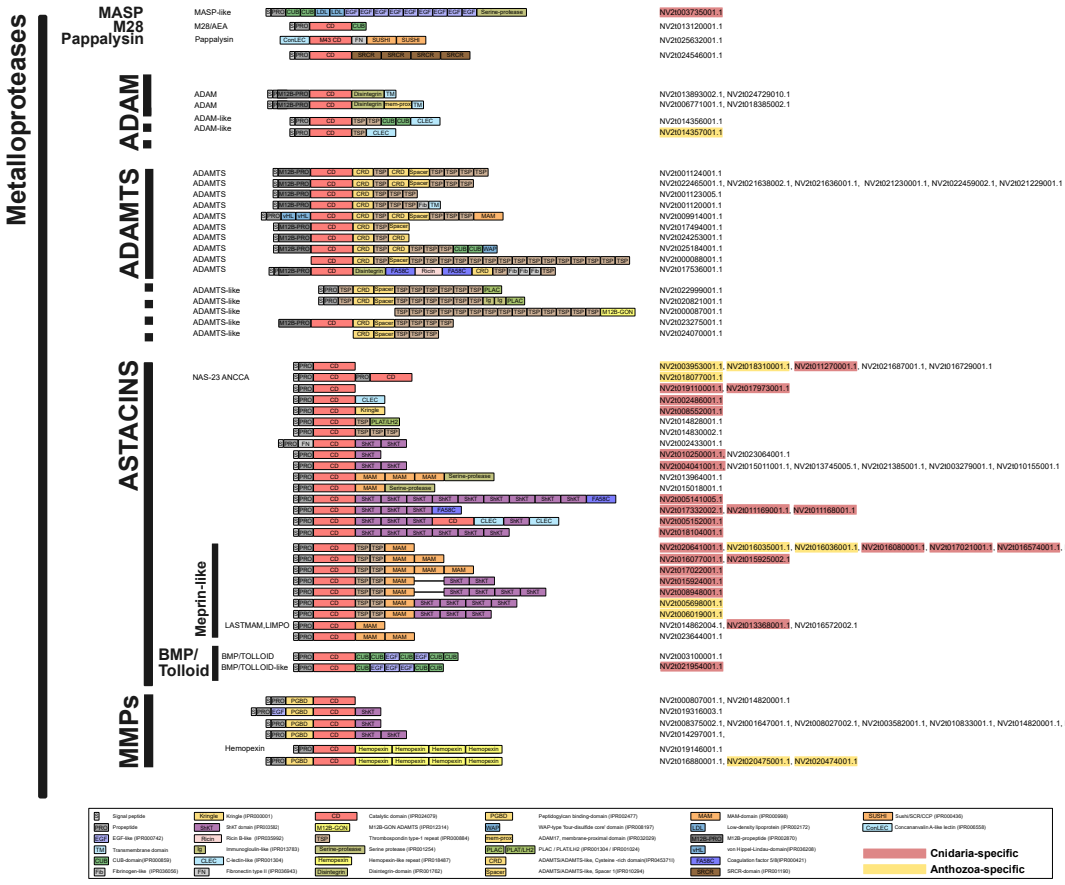


Figure 3.14. Domain Architecture of *Nematostella* Metalloproteases. *Nematostella* has a large number of MMPs. They can be classified into ADAM, ADAMTS, astacins and other MMPs. Their common feature is the catalytic domain but the different subtypes possess a large variety of different domains.

types. Interestingly, ADAM-4 (NV2.18385) and ADAM 1 (NV2.24729) showed up-regulation in larval datasets, suggesting activity in putative stem and germ cell progenitor cells. ADAMTS-16 (NV2.18541) as well as Astacins 7, 14 and 15 (NV2.19110, NV2.18077 and NV2.3279) are expressed in cnidocytes suggesting a potential role as venom components. Additionally, many MMPs were observed to be expressed in the endomesoderm/gastroderm and in neurogland cells.

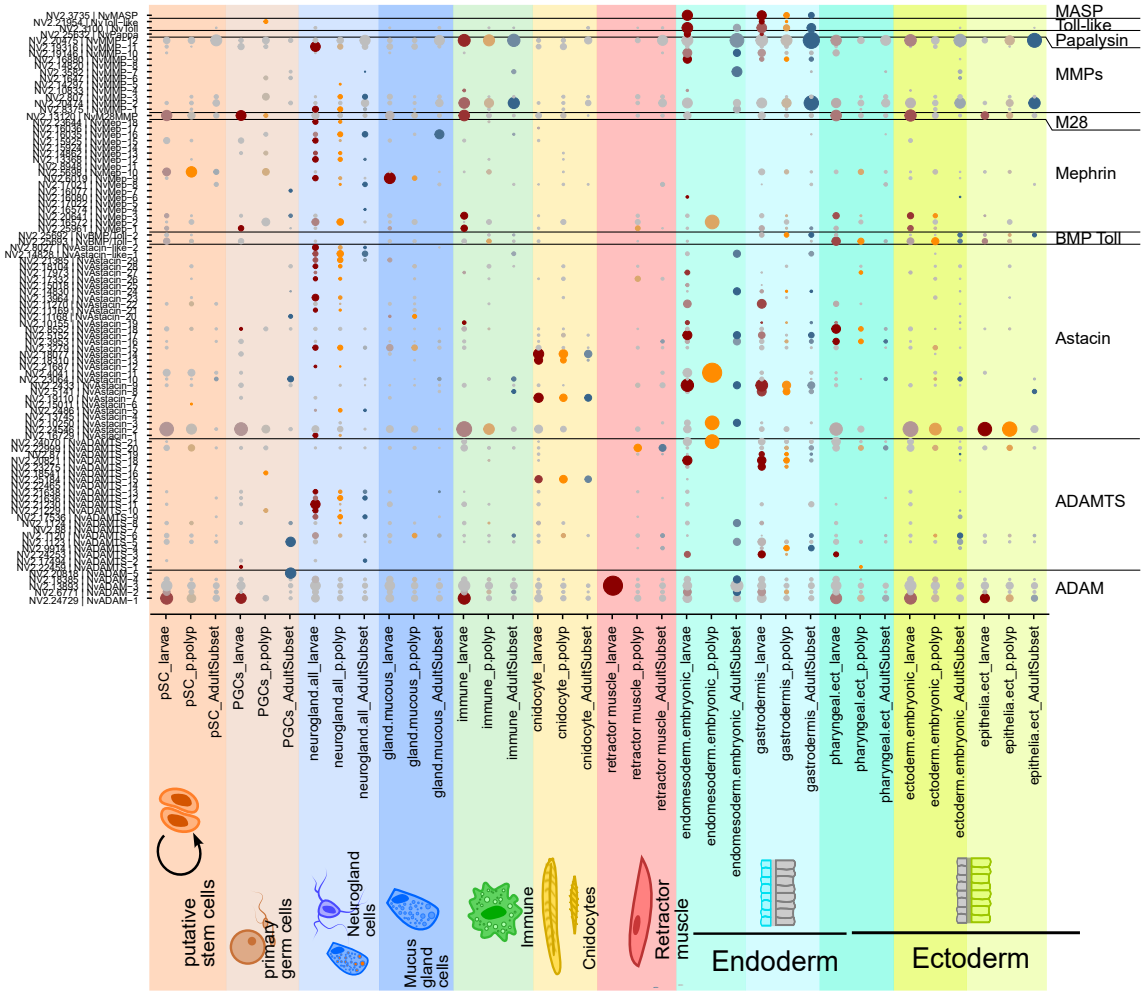


Figure 3.15. Cellular Expression of MMPs Single cell transcriptome analysis [Steger et al., 2022] of MMP expression throughout development (analysis by *Alison Cole*). The size of dots represents the cell proportion that expresses the gene. Color Intensity of the dots shows differential regulation between life stages. The MMPs of *Nematostella* are widely expressed through many cell types. The majority of MMPs are expressed by gastrodermal cells. The ADAMs are differentially regulated in larval stages. Cnidocyte expressed MMPs are possibly venom components..

3.3.4 The Polydoms of *Nematostella*

Polydom/Sushi, von Willebrandt factor type A, EGF and pentraxin containing protein (SVEP)1 is a large, multi-domain ECM protein, originally identified in murine bone marrow stromal cells as a Notch-like protein [Gilgés et al., 2000]. Subsequently, it was described as a ligand for Integrin α9β1 [Sato-Nishiuchi et al., 2012]. Polydom proteins are characterized by repetitive CCP domains, along with a vWFA, a pentraxin and multiple EGF-like domains, Due to their apparent role in cell-matrix contacts they are very likely part of the BL [Samuelov et al., 2017, Sato-Nishiuchi et al., 2012].

The Polydom protein family was of special interest for me as several of its members in *Nematostella* were detected to be differentially regulated in in the larvae-primary polyps comparison hinting at a possible role during metamorphosis. I was able to identify 3 full length Polydoms in *Nematostella* as well as 6 putative Polydom-related/-like proteins. To classify these proteins, I compared them to the domain architecture of described vertebrate and cnidarian Polydoms (CnPolydoms) (Figure 3.16).

CnPolydoms differ from vertebrate Polydoms by a CnPolydom core motif consisting of repeated Sushi (IPR000436), hyalin repeat (HYR) (IPR003410) and thyroxine kinase ephrin (TKE) (IPR011641) domains. In addition, their pentraxin (IPR001759) domain is followed by a plasminogen-apple-nematode (PAN)/Apple disIntegrin- (IPR003609) and a 5/8 coagulation factor (IPR024715) domain. They are also generally shorter than vertebrate Polydoms (vPolydom), which possess a long tail consisting of 35 Sushi domains.

The three Polydoms found in the matrisome of *Nematostella* constitue variations of

a CnPolydom. However, only NvCnPolydom 3 (NV2t003792001.1) possesses the disIntegrin and coagulation factor domains known from hydrozoan CnPolydoms. NvCnPolydom 2 (NV2t003300001.1) possesses a coagulation factor domain while both domains are missing in NvCnPolydom 1.

In addition to these three full-length Polydoms, I identified four Polydom-like proteins. These exhibit variations of the CnPolydom core motif following then N-terminal Coagulation factor 5/8 and poly-EGF motif. In full-length CnPolydoms a poly-EGF motif is located N-terminal to the CnPolydom core. Furthermore, the majority Polydom-like proteins share a C-terminal Immunoglobulin domain. NvPoldom-like-4 is distinguished by having a C-terminal poly TSP-1 (IPR000884) motif and a Ricin (IPR000772) domain. The shortest member of the Polydom family in *Nematostella*, designated as Polydom-related, comprises only the CnPolydom core motif (NV2t011255001.1).

Interestingly, I found an LRP-like protein, which possesses a C-terminal CnPolydom core motif in addition to repetitive low density lipoprotein receptor (LDL receptor (LDLR)) domains at the N-terminus. This protein possesses a C-terminal transmembrane domain suggesting that it is membrane-bound like other LRP receptors.

The single cell transcriptome data revealed that all proteins of this group appeared to be expressed in neural gland cells (Figure 3.17). These consists of multiple cellular subtypes including neurons, sensory secretory cells and digestive gland cells [Steger et al., 2022].

When these subtypes are further resolved it appears that all Polydoms are expressed in digestive gland cells. However, NvPolydom-like-2 (NV2.2755) and

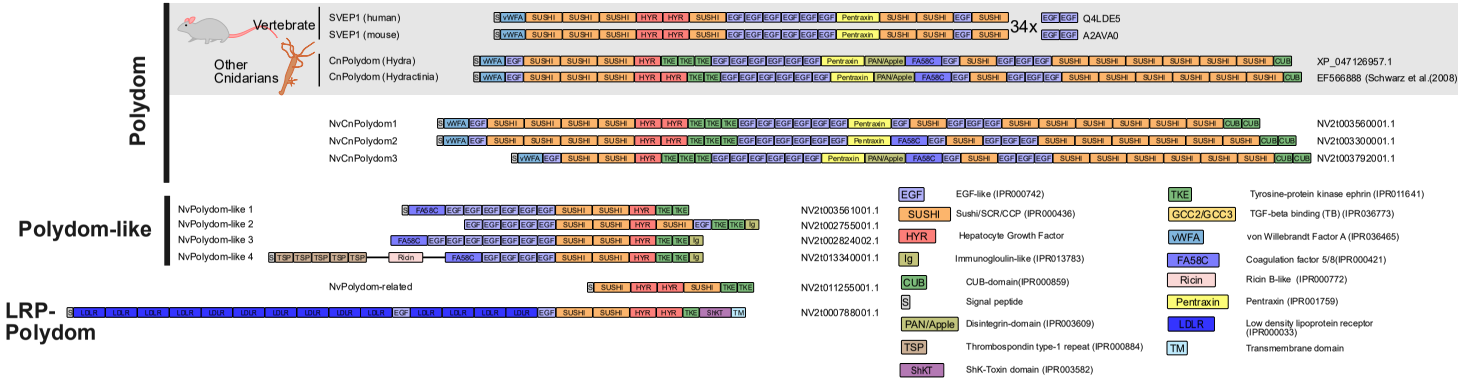


Figure 3.16. Domain Architecture of the Polydom Proteins of *Nematostella*. A comparison of the *Nematostella* Polydom protein family with the vertebrate SVEP1 (mouse/human) and the cnidarian-specific CnPolydom (*Hydra*/*Hydractinia*). Three Polydom proteins were detected in *Nematostella*. They resemble the CnPolydom, however, they have slightly different domain architectures. In addition to this 5 short proteins containing a partial Polydom core consisting of 2 Sushi-, an HYR- and 2 TKE-domains were found. Four of these also possess the EGF repeats found in full length Polydoms and were therefore designated Polydom-like. The shortest member of the family which we named Polydom-related only consists of the Polydom core motif. In addition to these we found a proteins with multiple low-density-lipoproteins (LDL) receptor repeats followed by the Polydom core motif which resembles a fusion of am LDL receptor-related protein (LRP) and a Polydom-like protein.

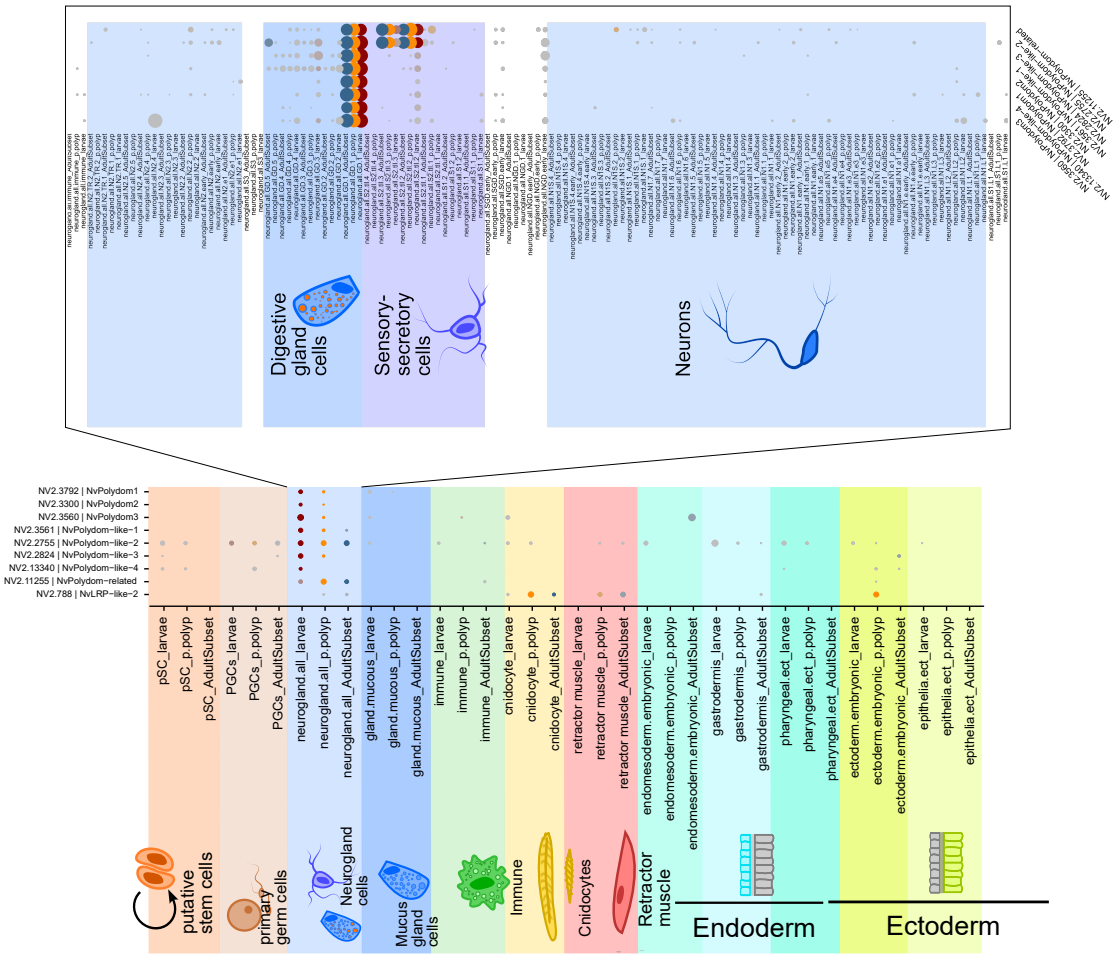


Figure 3.17. Single Cell Expression of the Polydom Proteins of *Nematostella*. Single cell transcriptome analysis [Steger et al., 2022] of the Polydom expression throughout development (analysis by *Alison Cole*). The size of dots represents the cell proportion that expresses the gene. Color Intensity of the dots shows differential regulation between life stages (red = larvae, yellow= primary polyp, blue = adult). All Polydom-related genes seem to be expressed in the neuroglandular cells of *Nematostella*. A closer look at the different subtypes of neuroglandular cells shows that all Polydoms are expressed in the digestive gland cells of all live stages. In addition to this Polydom-related and Polydom-like 2 are expresses in sensory/secretory cells..

NvPolydom-related (NV2.11255) are also differentially regulated in sensory secretory cells. Distinct from the other Polydom proteins, the LRP-like Polydom is expressed in cnidocytes.

3.4 The Developmental Dynamics of the Matrisome

After identifying the multitude of mesoglea components and their expression profiles *in silico* I was interested to investigate the composition of the mesoglea at the protein level in the different life stages. However, several challenges had to be addressed to facilitate the quantitative measurement of ECM proteins.

3.4.1 Mass Spectrometry Reveals Molecular Plasticity

The main challenge was to extract equally pure and intact mesoglea in sufficient amounts from 150-200 µm long larvae as well as 10 cm long adults. I had to adjust the pre-existing protocol for *Hydra* mesoglea extraction [Sarras et al., 1991] to accommodate for the different sizes and fragility of the individual life stages.

Generally, the extraction protocol consists of incubation in a detergent (0.5% N-laurylsarcosinate) followed by a freeze-thaw procedure in liquid N_2 and mechanical decellularization by repeated pipetting in distilled water. I opted for mechanical decellularization over chemical or enzymatic methods to minimize the risk of digesting or damaging mesoglea components.

To facilitate the decellularization of the adult gastroderm it was necessary to open the body wall along the oral-aboral axis using microscissors. In addition, caution was taken to preserve the tentacles during the pipetting steps. This involved carefully lifting the animals from the aboral end and by manually peeling off ectoderm

from the tentacles using a fine forceps.

Larvae and primary polyps were processed in large batches due to their small size, which led to significant material loss. They were strained through a 70 µm sieve to separate them from the detergent and vigorously pipetted using fine glass pipettes with flame-blunted tips. The main challenge was to visualize the decellularized mesoglea in the washing wells, as the water was saturated with cell debris and residual detergent over time, rendering the mesoglea transparent. To address this, a strong lateral light source illuminated the water, making floating mesogleas visible. The decellularized mesogleas stick to plastic and, to a lesser extend to glass. To mitigate this, I recommend keeping the water in motion during the washing steps. Mesogleas that adhere to the vials can be dislodged with a strong water current or by carefully peeling them off with a forceps, though there was a risk of ripping or damaging them. When the water becomes too turbid, transferring mesogleas to fresh water was an option, either through individual pipetting or centrifugation in a table centrifuge (1 min at 13,000 rotations per minute (rpm)). However, cleaner mesogleas resulted in more material loss during these transferal steps, as they tended to stick to the pipette or centrifuge vials.

To confirm that cells were removed from the mesogleas, they were checked using phase contrast microscopy during the wash steps and a fraction of the extracted mesogleas was stained using DAPI and NvLaminin antibody to visualize remaining nuclei, aproviding evidence for the integrity of the extracted mesogleas (Figure 3.18 D-F).

Both quality controls showed that, while complete decellularization of the mesogleas proved challenging, a notable reduction in cell count was achieved, resulting in an enrichment of mesoglea material across all life stages. Most of the remaining cells

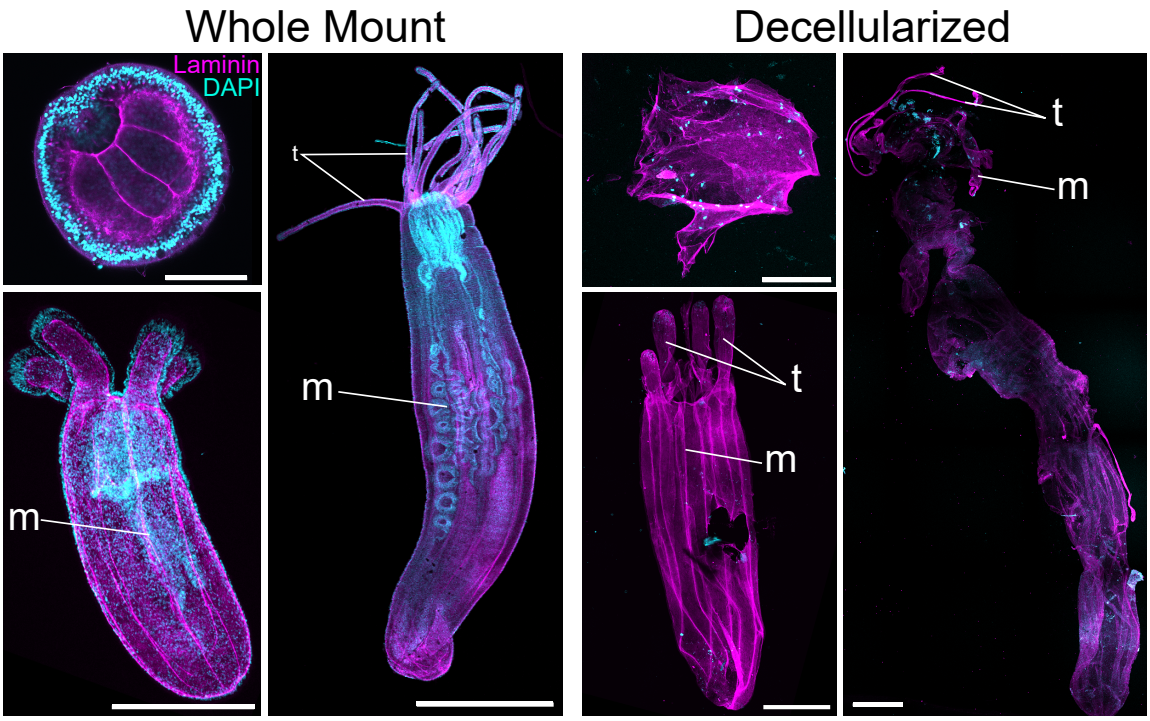


Figure 3.18. Decellularization of *Nematostella* Mesoglea at Three Life Stages. Whole mount stainings of (A) a *Nematostella* larva, (B) primary polyp and (C) adult using NvLaminin (magenta) to visualize the mesoglea and DAPI (cyan) to stain nuclei and cnidocysts. (D-F) Decellularized mesoglea from the respective life stages. The decellularized mesogleas retain some residual nuclei and cnidocysts but the overall number of cells is significantly reduced. Morphological structures such as tentacles (t) and mesenteries (m) are preserved during decellularization and the whole mesoglea is stained with NvLaminin suggesting that the BL remains intact as well.

were amoebocytes that reside within the mesoglea. Notably, some gastrodermal cells remained in the distal tentacles and the aboral pole (especially in primary polyps), where exposure to mechanical decellularization and detergents was minimal. In addition to this, discharged cnidocysts sporadically stuck to the extracted mesogleas.

Removing these remaining contaminants would have necessitated compromising the morphological and molecular integrity of the mesoglea. Instead, I decided

to exclude cytosolic components during post-processing by mapping the *in silico* matrisome on the experimentally identified proteins.

The sample preparation for MS was performed by *Mandy Rettel* (EMBL Proteomics Core Facility); the post processing and initial statistical analysis of the resulting data was performed by *Frank Stein* (EMBL Proteomics core facility)

For MS analysis, I collected larvae, primary polyps, and adults and decellularized their mesoglea in three replicates. Approximately 100-200 decellularized larvae and primary polyp mesogleas and four adult mesogleas were collected per MS run, which amounted to 1 µg proteins for each measurement. To exclude any contaminant sequences from bacteria or human proteins we mapped the results to the *Nematostella* NV2 protein model dataset.

In total, 5081 proteins were detected across the three life stages. Principal component analysis (Figure 3.19) effectively discriminated between the different samples, illustrating the distinct molecular compositions of the mesoglea across various life stages.

To exclude proteins originating from residual cells, we used the previously determined raw and curated *in silico* matrisome to extract ECM proteins (Table 3.1). Initially, we identified 491 proteins within the raw *in silico* matrisome, which subsequently diminished to 287 proteins following exclusion of isoforms and fragments. This amounts to 34%̇ of the curated *in silico* matrisome.

Table 3.1: Proteins Detected Through Mass Spectrometry

Total unique proteins found in MS	5081
Raw *in silico* matrisome	491/1810 (27%)
Curated *in silico* matrisome	287/843 (34%)

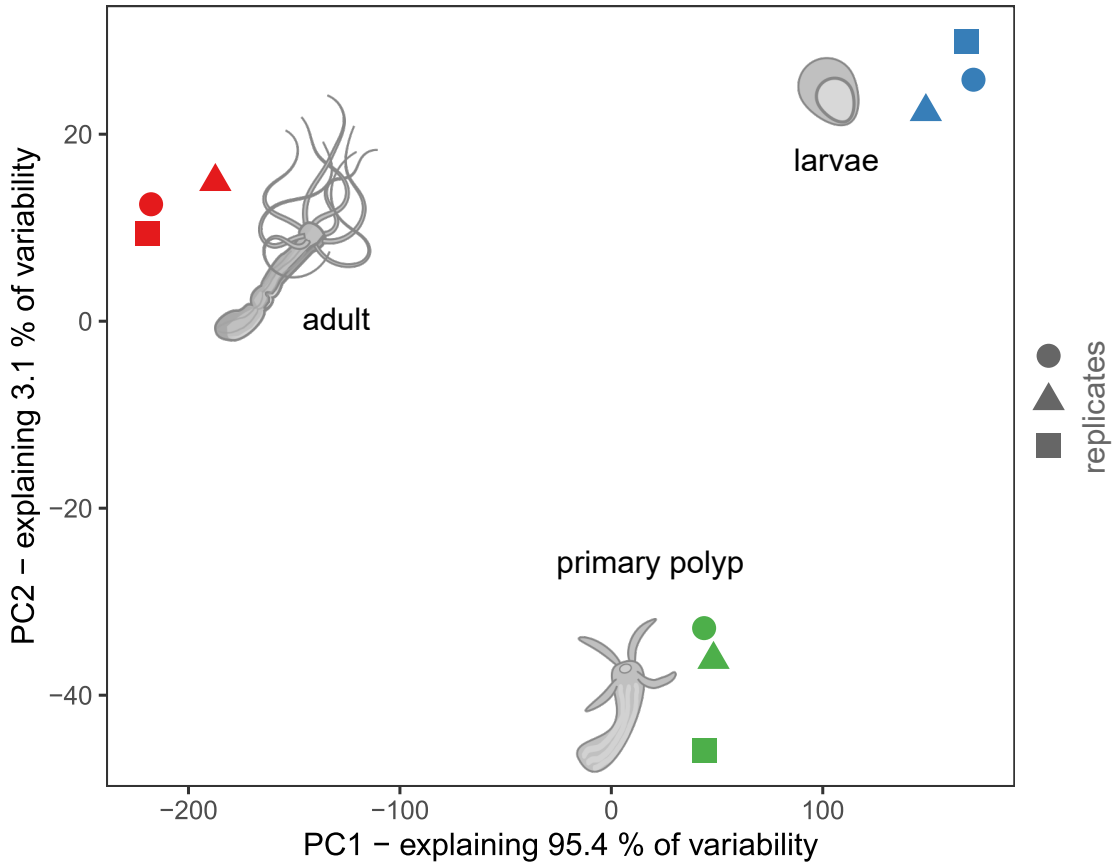


Figure 3.19. The Different Life Stages have Different Molecular Signatures. Principal component analysis of the molecular signatures resulting from the MS of the mesoglea of the three life stages shows a clear separation of the three samples showing their distinct protein composition.

Given the low percentage of ECM proteins in the MS dataset, I wanted to ensure that distinctions among life stages extended to the mesoglea composition. Therefore, we statistically analyzed the peptide hits to identify the number of differentially abundant mesoglea proteins. For this, we defined proteins as statistical hits (upregulation > 2 fold and p-value <= 0.05) and candidates (upregulation >= 1.5 fold and p-value < 0.2 Table 3.2)

Table 3.2: Statistical Analysis of the Differential Abundance of Matrisomal Proteins

Life Stage Comparison	Hits (>2-fold, p<0.05)	Candidates (>1.5-fold, p<0.2)	No Hit
Larvae - Primary polyp	98	59	130
Primary polyp - Adult	157	41	89
Larvae - Adult	143	56	88

Surprisingly, the quantity of differentially abundant proteins is nearly equal across life stage comparisons. This implies that the mesoglea of adults is about as different from that of primary polyps as from that of larvae. Intriguingly, during the in larvae-primary polyps transition were most large-scale morphological changes take place, the distinctions in mesoglea composition are less pronounced than in any other comparison.

3.4.2 Compositional Changes of the Matrisome

After establishing that there are indeed molecular differences at the protein level between the different life stages, I wanted to investigate the exact compositional changes during development. The objective was to understand if these difference stem from a simple increase in mesoglea complexity over time or if each life stage has a distinct protein composition.

The comparison of protein abundance across different life stages (Figure 3.20) suggests a stage-specific composition of the mesoglea. The mesoglea of primary polyps is rich in BL factors, particularly those associated with epithelial organization, such as Laminins and Polydoms, along with cell adhesion receptors categorized as “Other”. In addition, I found many ECM regulators to be abundant during this life stage (Figure 3.21 and Suppl. Table 3). As expected, only 10 proteins

are upregulated in larvae compared to 115 proteins upregulated in primary polyps (Table 3.3 and Suppl. Table 3). In summary, the life stage comparison implies that a general trend towards increased complexity in the mesoglea of older stages. However, the mesoglea composition remains distinct for each stage, underscoring that each life stage depends on a unique set of ECM factors.

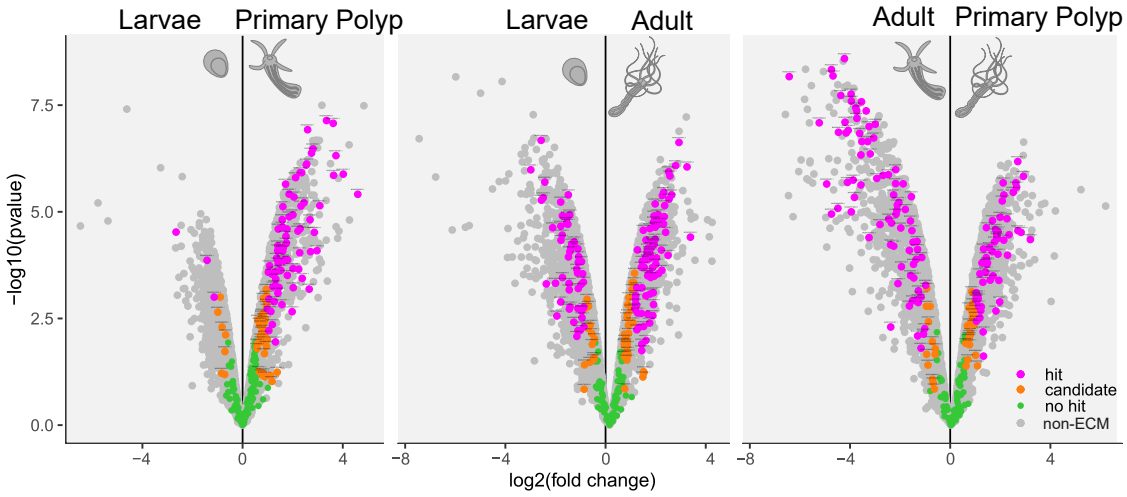


Figure 3.20. Changes in Protein Composition of the Mesoglea. These Vulcano plots depict the relative change of abundance of matrisomal proteins between larvae, primary polyp and adult. The colors depict statistical significance (green = insignificant; orange = fold change >1.5 + p value <0.2; magenta = fold change > 2 + p value < 0.05). Gray dots show the background of non-matrisomal proteins. In the larvae- primary polyp transition a large number of matrisomal proteins get upregulated and there are only few proteins that are more abundant in larvae. The primary polyp- adult comparison shows a nearly equal distribution of proteins that are upregulated in the two life stages showing that the composition is stage specific rather than constantly increasing in protein diversity.

3.4.3 The Interstitial Matrix develops after Metamorphosis

The comprehensive examination of experimental matrisomes suggest that each life stage possesses a distinct ECM composition. I was interested in understanding the individual dynamics of different matrisomal proteins and protein families which

Table 3.3: Upregulated Proteins in Larvae and Primary Polyps

Upregulated in larvae	ID	Description	Upregulated in primary polyps
NV2t003648001.1	Coagulation factor-like	NV2t000650001.1	Secreted Integrin-N-related protein
NV2t002433001.1	Astacin Metalloprotease	NV2t014921001.1	TLDc domain-containing protein
NV2t000914001.1	Vitellogenin	NV2t011903001.1	Crumbs-like
NV2t019166001.1	Lysyl oxidase subfamily 2-like	NV2t014862004.1	Meprin-like Astacin metalloprotease
NV2t010745001.1	Leucine-rich repeat-containing protein-like	NV2t017536001.1	ADAMTS metalloprotease
NV2t013515001.1	Microfibrillar-associated protein 1	NV2t014399001.1	meprin, A-5 protein, and receptor protein-tyrosine phosphatase mu (MAM) and Fibronectin type II domain-containing protein
NV2t000890002.1	SHOC2 leucine rich repeat scaffold protein-like	NV2t016479001.1	Coadhesin-like
NV2t008618001.1	vWFA domain-containing protein	NV2t021578001.1	vWFA domain-containing protein
NV2t020560001.1	vWFA domain-containing protein	NV2t008768001.1	Collagen6-like
NV2t001994001.1	vWFA and cache domain-containing protein 1	NV2t023165001.1	Collectin-like
		NV2t002780001.1	TSR domain-containing protein
		NV2t008769001.1	Anthrax toxin receptor 1-like
		NV2t012484002.1	Kielin/Chordin-like
		NV2t011255001.1	Polydom-related
		NV2t023593001.1	Protocadherin Fat 4-like

Table 3.4: Upregulated Proteins in Primary Polyps and Adults

Upregulated in primary polyps	ID	Description	Upregulated in adults
NV2t011903001.1	Crumbs-like	NV2t016574001.1	Meprin-like astacin metalloprotease
NV2t014862004.1	Meprin-like Astacin metalloprotease	NV2t006705001.1	TLDc domain-containing protein
NV2t000650001.1	Secreted Integrin-N-related protein	NV2t008618001.1	vWFA domain-containing protein
NV2t014399001.1	MAM and Fibronectin type II domain-containing protein	NV2t003648001.1	Coagulation factor-like
NV2t011255001.1	Polydom-related	NV2t013423001.1	Follistatin-like
NV2t002780001.1	TSR domain-containing protein	NV2t022930001.1	Fibrillin-like
NV2t008406003.1	Anthrax toxin receptor-like	NV2t011069001.1	Zona pellucida sperm-binding protein-like
NV2t006021001.1	Crumbs-like	NV2t022931001.1	Fibrillin
NV2t017536001.1	ADAMTS metalloprotease	NV2t022932001.1	Fibrillin-like
NV2t023593001.1	Protocadherin Fat 4-like	NV2t006476001.1	Insulin-like growth factor-binding protein and Kazal domain-containing protein
NV2t002755001.1	Polydom-like	NV2t000320001.1	Kunitz-type serine protease inhibitor
NV2t013340001.1	Polydom-like	NV2t023229001.1	Follistatin-like
NV2t008768001.1	Collagen6-like	NV2t012181001.1	EGF and Trimeric LpxA-like domain-containing protein
NV2t015872001.1	NOWA-1	NV2t022057001.1	Tumor necrosis factor
NV2t002824002.1	Polydom-like	NV2t017096001.1	Ryncolin-1

contribute to this diversity within the developing matrisome. Therefore, I analyzed the average abundance of the identified proteins (Figure 3.21).

This analysis showed that numerous proteins are specifically detected in certain life stages, underscoring distinct dynamics across various protein families within the ECM. Only a minority of the detected proteins appeared to be increased in relative abundance throughout development.

An example of this dynamic behavior is observed in the Collagen family, which increases in abundance towards older life stages. This aligns with the previously shown immunofluorescence analysis, revealing Collagen as integral part of the IM in primary and adult polyps, while being absent in lava. By contrast, Collagen-6-like (NV2t008768001.1), Collagen-8-like (NV2t007052001.1) and Collagen triple helix repeat- containing (CTHRC) protein (NV2t002667001.1), showed the highest

abundance in the primary polyp. In addition, Collagen-4 (NV2t021411001.1) and Collagen-18 (NV2t011420001.1) are upregulated in primary polyps and remain on a constant level in the adult sample. Notably, all of these exceptions are likely components of the BL.

All Polydom-like proteins are likewise abundant in primary polyps, yet they are scarcely detectable in both larvae and adult samples. The Polydoms are likely BL components involved in cell-matrix interactions providing further evidence that the BL and its interaction with the epithelial cells play and important role during metamorphosis. These dynamics underscore the significance of exploring BL components, particularly in the context of functional investigations into stage-specific developmental processes governing the transition from larvae to primary polyps.

Several other protein families, such as the MMPs and various ECM-affiliated factors show abundances that suggest specific variations at each life stage. Given that these families include most of the ECM- regulatory factors, the stage-specific variations likely contribute significantly to the mesoglea's ability to adapt according to the developmental requirements of different life stages.

3.5 Matrisome Components Affect the Morphogenesis of *Nematostella*

Two hypotheses are possible to explain the compositional shifts during development. Either the developmental processes cause a change in the deposition of ECM factors or the ECM itself influences the development of the surrounding tissues. While the truth is most likely a reciprocal interaction between both processes, I attempted to demonstrate whether individual mesoglea components can affect

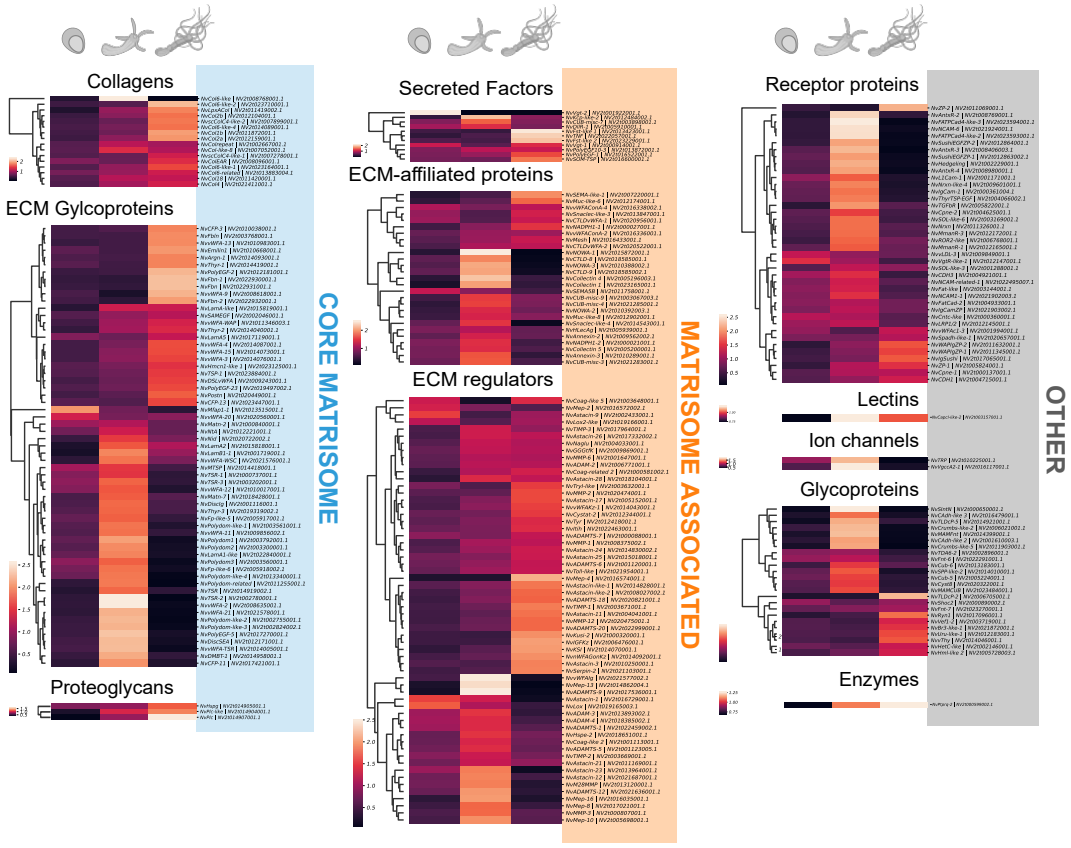


Figure 3.21. Developmental Dynamic of Different Matrisome Components. Comparison of the average abundance of different mesoglea components across the early development sorted by ECM divisions and protein families. Most core matrisome factors have clear differential abundance in either the primary polyp stage (most of the components) or the adult (most IM factors). The matrisome-associated factors are variably abundant across the different life stages. In most of these groups, life stage-specific variants can be identified as well as proteins that continuously increase in abundance. The proteins of the "Other" category are most abundant in the primary polyp.

development. For this, I performed a shRNA knockdown screen against a set of components which were detected as differentially abundant in the larvae-primary polyp comparison and therefore possibly involved in metamorphosis.

An effective shRNA knockdown protocol has been established for *Nematostella*, which allows suppressing of gene activity over multiple weeks covering the entire

Table 3.5: **shRNA Knockdown Candidates.** Differentially abundant proteins from the larvae-primary polyp comparison were chosen and screened for maternal deposition as well as gene duplication which would inhibit a knock down effect. Fifteen candidate genes were injected into fertilized eggs. Three of these resulted in knock-down phenotypes.

Gene ID	Name	Description	Upregulated in	KD Effect
Nv2.1124	NvADAMTS-8	ADAMTS Metalloprotease	Primary Polyps	-
Nv2.22999	NvADAMTS-20	ADAMTS Metalloprotease	Primary Polyps	-
Nv2.11904	NvCRUMBS-like	CRUMBS-like	Primary Polyps	-
Nv2.22931	NvFbn	Fibrillin	Primary Polyps	-
Nv2.650	NvSIntN	Secreted Integrin-N-related protein	Primary Polyps	- Early development stop
Nv2.17119	NvLamA5	Laminin subunit alpha 5	Primary Polyps	-
Nv2.3735	NvMASP	MASP-like Metalloprotease	Primary Polyps	-
Nv2.20956	NvCTLDvWFA-1	CTLD and vWFA domain-containing protein	Primary Polyps	-
Nv2.8618	NvvWFA-9	vWFA domain-containing protein	Primary Polyps	-
Nv2.3300	NvPolydom-2	CnPolydom	Primary Polyps	-
Nv2.2824	NvPolydom-related	Polydom-related (CnPdom-core)	Primary Polyps	- Defects in mesentery formation, Small primary polyps
Nv2.6768	NvROR2-like	ROR2-like	Primary Polyps	- Small primary polyps
Nv2.914	NvVgt-1	Vitellogenin	Larvae	-
Nv2t1994	NvvWFD-1	vWFD domain-containing protein	Larvae	-

development from larvae to primary polyp [Hill et al., 2022]. However, not all genes are suitable for shRNA knockdown; genes that are duplicated and maternally deposited genes present challenges for efficient knockdown (Aissam Ikmi, personal communication). Therefore, I used published transcriptome data of unfertilized *Nematostella* eggs, and developing embryos [Fischer et al., 2014] to exclude all candidates that are maternally deposited and not upregulated during the first 15 days of development. I used basic local alignment search tool (BLAST) to identify duplicated genes and excluded these as well. This resulted a list of candidates, which I used for shRNA generation and injection into freshly fertilized eggs (Table 3.5). The injected eggs were then imaged from day 2 to day 10, which is sufficient for the full transition to a primary polyp. Finally, the resulting embryos were fixed and imaged.

The water-injected control animals generally exhibited normal development, with most embryos progressing to the primary polyp stage. A small fraction of the embryos showed delayed development and were still in the larvae or tentacle bud

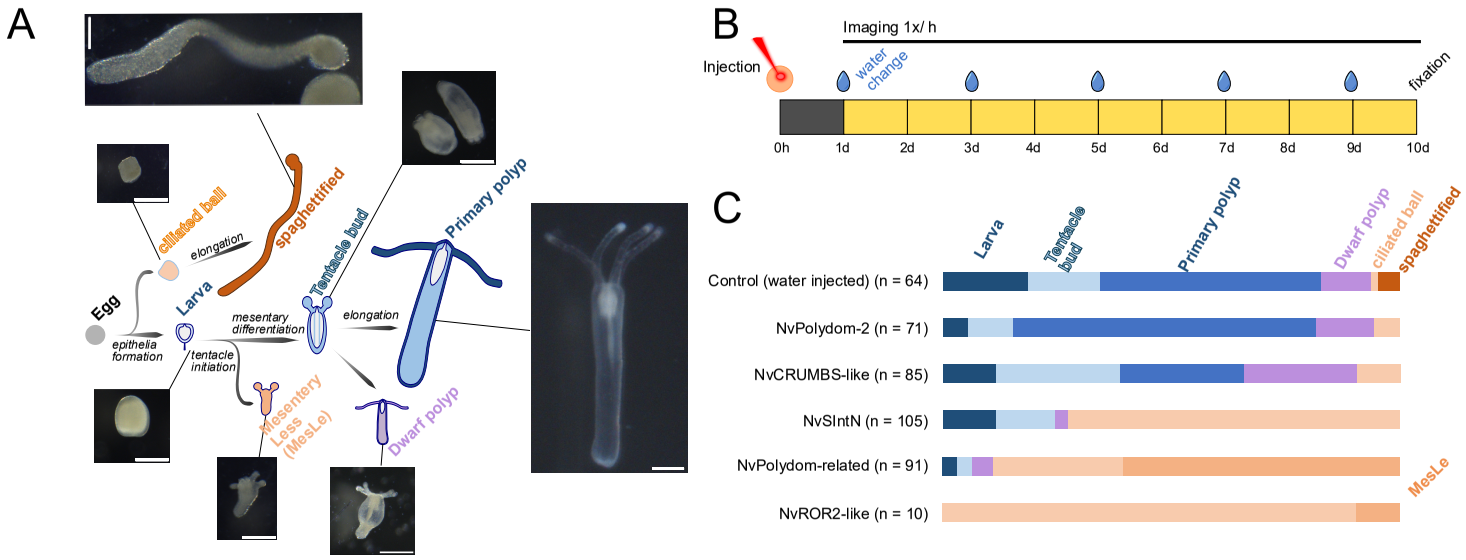


Figure 3.22. The Effect of shRNA Knockdown of ECM Candidate Genes (A) Different phenotypes were observed following knock down of ECM proteins. Apart from the normal developmental stages (larvae, tentacle bud and primary polyp) some embryos failed to establish correct epithelia. These developed into ciliated balls that randomly moved around. In some cases these undefined balls elongated into thin, elongated 'spaghettified' embryos that lack internal structures. Another phenotype observed is the dwarf polyp that morphologically seems to be completely developed but do not develop the full length of average polyps. Finally, I observed embryos that did not develop mesenteries or elongated but were able to establish tentacle buds and mouth structures. These were termed mesentery-less (*MesLe*). (B) Setup of the knockdown. Injected animals developed for 24 in the dark and were then imaged every hour while the medium was changed periodically. On day 10 the animals were fixed. (C) Quantification of phenotypes in five knock down conditions. NvPolydom did not significantly alter the development, NvCRUMBs KD slightly increased the number of dwarf polyps. NvSIntN led to the formation of ciliated ball phenotypes. *MesLe* phenotypes were observed in NvPolydom-related KDs and in one case in an NvROR2-like KD.

stage. A few embryos displayed more severe developmental phenotypes were observed characterized by ciliated cell clumps or spaghetti-like elongated but likewise undefined stages.

Embryos injected with shRNA against NvCnPolydom-2 (NV2.2824) and CRUMBS-like (NV2.11904) developed mostly normal. While the CRUMBS-like knockdown had a minor impact on the number of fully grown primary polyps, the knockdown of Integrin-like (NV2.650), which was anticipated to affect cell-matrix connections, resulted in a substantial fraction of ciliated ball phenotypes. These structures lacked discernable embryonic features but exhibited cilia and showed random rotational movement.

This phenotype was also observed in the Polydom-related (NV2.2824) and ROR2-like (NV2.6768) knockdowns. In addition to the ciliated ball phenotype a second phenotype, characterized by a failure to elongate and form mesenteries, was observed here. This phenotype was named mesentery-less (*MesLe*). During the fixation process, these phenotypes as well as the previously mentioned cilia ball phenotypes proved fragile,partially dissolving into cell clumps when the fixative was washed out. This might indicate severe impairment in cell connections under these knockdown conditions.

The fragility of the embryos made a closer investigation of the *MesLe* phenotype challenging. Initially, I observed that while the *MesLe* phenotypes developed tentacle structures these tentacles did not fully develop but instead they remained in a transition state between tentacle buds and fully developed tentacles as evident from the roundish tentacle tips. The mouth structures were visible as small elevations at the oral end, however, I was not able to test if the mouths are functional and able to take up food particles yet. The mesoglea as well as the ecto-

and gastroderm were discernable in the body wall, however, I was not able to detect internal structures, suggesting a potential role of Polydom/Laminin-mediated cell-matrix contact in the correct invagination of the body wall gastroderm.

These knock down experiments underscore that a pertubation of individual ECM components can severely impact the metamorphosis of *Nematostella*.

3.6 The Matrisome of *Nematostella* in the Context of Evolution

At the time of writing, the taxonomic sampling for full matrisomes is highly biased towards vertebrates and conventional model organisms. In contrast, the investigation of the matrisome of other species has thus far been done mostly based on the identification of individual factors in different taxa. Therefore, I decided to create a larger taxonomical dataset of basal animals to create a *Nematostella* matrisome and put it into an evolutionary context.

The pipeline I used for the identification of the raw *Nematostella* matrisome was successful in identifying a large number of ECM factors while still resulting in a significant amount of false positive hits, which had to be removed manually. It is not feasible to do use the same approach for a larger taxonomical dataset. However, even the determination of raw *in silico* matrisomes can provide valuable insights in the evolution and taxonomical distribution of ECM factors. Further, the addition of the evolutionary dimension facilitates better preliminary classification of ECM orthologues, enabling the classification of many proteins at the same time.

I used the *in silico* matrisome prediction pipeline (Source Code 1) to predict the

matrisomes of 44 additional basal animal species: 26 cnidarians (1 cubozoan, 5 hydrozoans, 3 scyphozoans, 2 myxozoans, 15 anthozoans), 3 ctenophores, 2 sponges, 2 choanoflagellates, 1 flat worm , 2 placozoans and 3 lancet fish and compared these to published matrisome sizes (Table 7).

The average predicted matrisome size of multicellular animals is relatively variable and consists of 1037 ± 479 proteins. As expected, the average matrisome size of unicellular organisms is significantly lower (200 ± 60 proteins). Similarly, the ratio of matrisome genes within all protein-coding genes varies between 2.2 – 7.7 % (ø 4.4 ± 1.3 %.) in multicellular animals and 1.1 – 2.5 % (ø 1.7 ± 0.7 %) in unicellular organisms. Notably, the lancet fish (Branchiostoma) possesses extremely large overall proteomes and thus, also extremely sizable matrisomes. It is unclear if these large numbers of genes are due to increased evolutionary speed in this lineage or due to gene annotation errors [Bányai and Patthy, 2016].

3.6.1 Taxon-Specific and Conserved Matrisomal Orthologues

To understand the evolutionary conservation of different matrisome factors I classified the predicted *in silico* matrisomes onto OGs using OrthoFinder [Emms and Kelly, 2019]. This resulted in 4240 different OGs across all species. As part of the quality control, the OrthoFinder algorithm creates a phylogenetic tree based on the most likely gain and loss events of OGs (Figure 3.23). The generated phylogeny largely recapitulates the generally accepted evolutionary tree of metazoans with some notable exceptions. Firstly, the anthozoans are placed as a sister group of bilaterians outside of the other cnidarians, suggesting that the matrisome of medusozoan cnidarians might have gone through additional diversifications. Indeed, when looking at the abundance of different orthogroups in cnidarians (Figure 3.24

B) a distinct set of OGs is detectable, which is most abundant in medusozoans. Interestingly, the myxozoans and the ctenophores are placed basally to all other species. As myxozoans have become single-celled organisms secondarily, they have lost many ECM genes or possess highly derived variations so their unusual placement at the base of the tree was expected. The ctenophores have an average matrisome size (ø = 772.3) so their placement at the base of the phylogeny suggests that they possess a unique or highly modified matrisome.

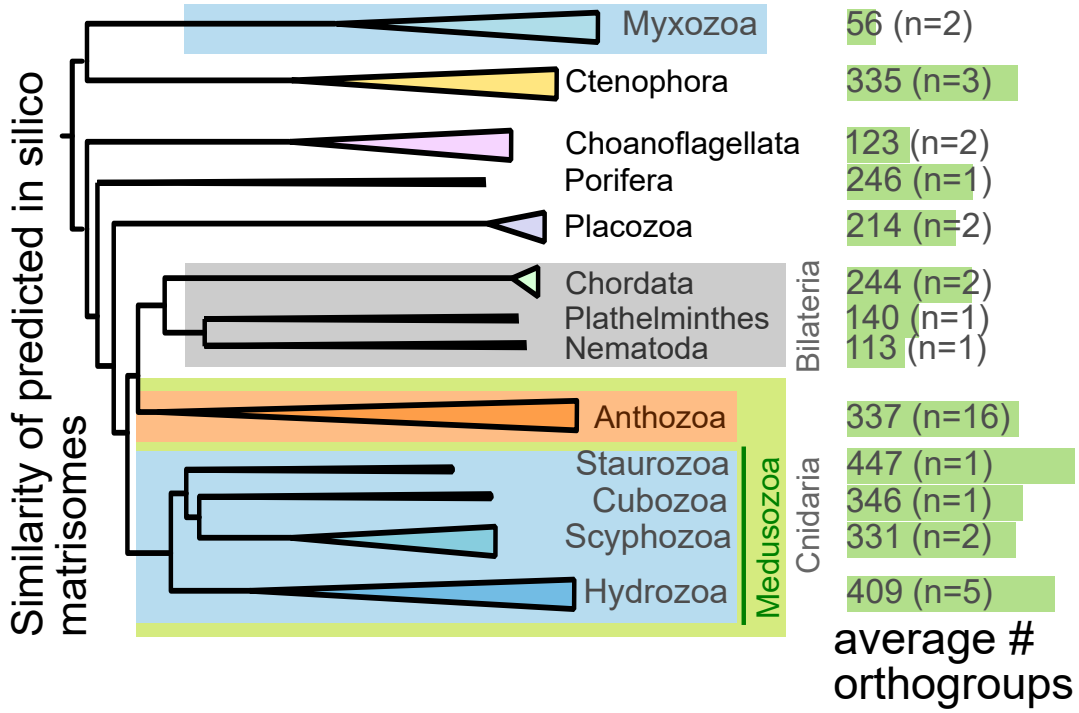


Figure 3.23. Species Phylogeny based on Matrisome orthogroups. The species tree generated by OrthoFinder based on the predicted OGs of the predicted matrisomes. This matrisome tree recapitulates the generally accepted evolutionary tree of animals. The only exceptions are the myxozoans and ctenophores which are places basally to all other animals.

The detected OGs can be used as a proxy for the diversity of matrisomal factors. As the *in silico* matrisomes are not curated this analysis likely overestimates the

true diversity of the different matrisomes. However, lacking curated datasets, this estimation can help to identify trends and future research directions. I was interested in identifying species that have an unusually low or high matrisome diversity so I plotted the overall size and number of OGs of all (raw *in silico* and published curated) matrisomes, with the exception of the lancet fish 3.24. The matrisome size is proportional to the number of OGs (matrisome diversity) for most species. The only exception are the two species of sponges (*A. queenslandica* and *E. muelleri*) which have a reduced matrisome diversity compared to their overall size of the matrisome. When looking closer at these species I found that 37 % of their matrisome is made up of expanded fibronectin-like proteins (sorted into 4 OGs) which explains the lower diversity of their matrisome.

I aimed to reduce the number of false positives by excluding all OGs that were neither found in *Nematostella* nor in any of the previously published matrisomes.

By removing all OGs that were not found in any of the previously published matrisomes (Figure 3.24 C + C') it is possible to assay the conservation of the vertebrate matrisome orthologues through evolution. Surprisingly, although the lancet fish have much larger matrisome sizes than the vertebrates they have a comparatively low conservation of the vertebrate factors.

To further investigate the taxonomical distribution of OGs, I extracted taxonomically restricted OGs from the dataset (Figure 3.25 A). As expected, the chordates contribute a large amount (829) of the identified OGs, however, most of these OGs stem from the Amphioxiformes (lancet fish). Vertebrates contribute 289 OGs. As I included a large amount of cnidarians in my analysis, this taxon contributes nearly a quarter of all identified OGs. The other clades each contribute a number of taxon-specific OGs showing that each taxon possess a few innovations of the

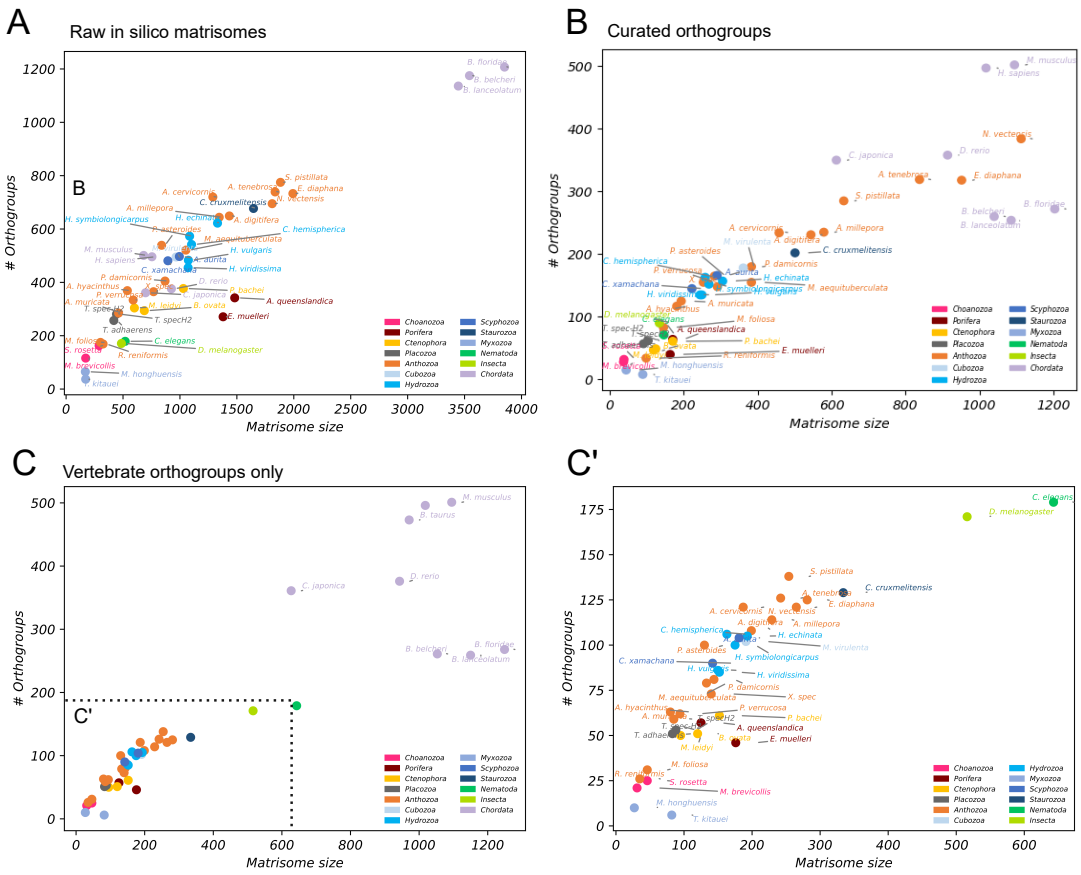


Figure 3.24. Evolutionary Comparison of Matrisome Size and Diversity. (A) The sequences of the *in silico* matrisomes were sorted into OGs using OrthoFinder as an approximation for matrisome diversity. While the diversity and the number of sequences is linearly correlated in most species the sponges (*A. queenslandica* and *E. muelleri*) have a comparatively low matrisome diversity. (B) Reducing the results to only include curated OGs drops some of the incorrectly assigned groups, however, due to the limited taxonomical sampling this analysis is biased towards species that are closely related to *Nematostella*. (C) To visualize the conservation of known matrisome OGs all OGs that do not contain any previously published matrisome sequence were excluded. The invertebrates (C') generally possess fewer of these conserved OGs. The large number of OGs in the raw matrisomes is likely due to invertebrate-specific innovations.

matrisome genes. Within the different OGs, different numbers of orthologues are found (Figure 3.25 B).

As the matrisome of Ctenophores was identified to be very divergent from the other animal matrisomes I wanted to take a closer look at the ctenophore OGs and identify the taxon that shares the most OGs with the ctenophores. I therefore extracted the exclusive and shared OGs between ctenophores, sponges, cnidarians and vertebrates (Figure 3.26).

This comparison revealed that ctenophores and cnidarians share 292 OGs (180 shared only between cnidaria and ctenophora + 63 shared between cnidaria, porifera and ctenophora + 31 shared between vertebrates, cnidaria and ctenophora + 18 shared between all groups) while the shared OGs with sponges and chordates is only 90 and 62 OGs respectively. These are more shared OGs than the OGs shared between any of the other taxon comparisons. In other words, my result provide evidence that the matrisome of ctenophores might be most closely related to that of cnidarians.

At the same time, ctenophores also possess 201 ctenophore-specific OGs, possibly explaining the divergent position of ctenophores in the matrisome phylogeny. The cnidarians contribute the largest amount of taxon-restricted OGs and share roughly the same amount of OGs with the other taxa (292 with ctenophores, 289 with sponges and 234 with vertebrates) placing them in the “center” of these evolutionarily interesting taxa.

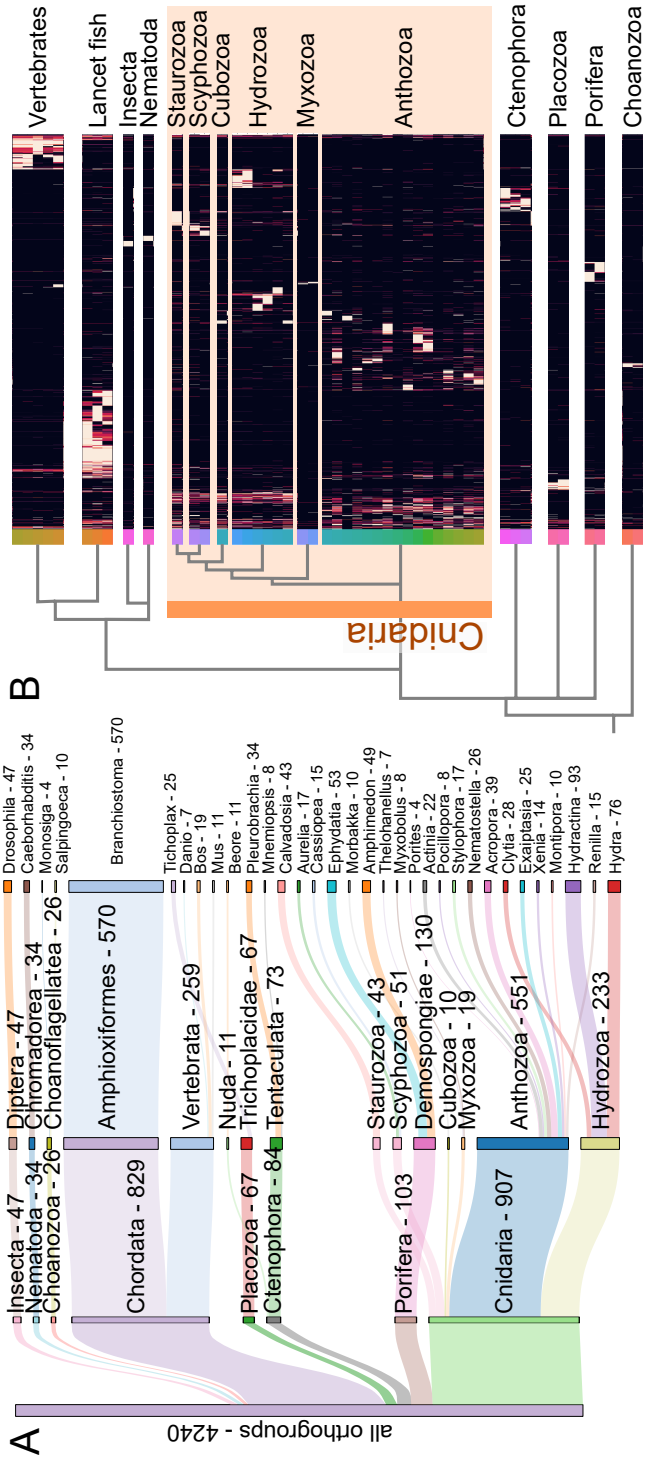


Figure 3.25. Taxonomic Distribution of Orthogroups. (A) The identified OGs of the matrisomal dataset were analyzed for shared and taxon exclusive OGs. Most identified OGs (2097) can be found across at least two phyla. All other OGs are restricted to individual phyla. For example, 907 OGs are only found in cnidarians, most of these are restricted to specific genera. (B) Number of orthologues in the different OGs across taxa normalized to each OG.

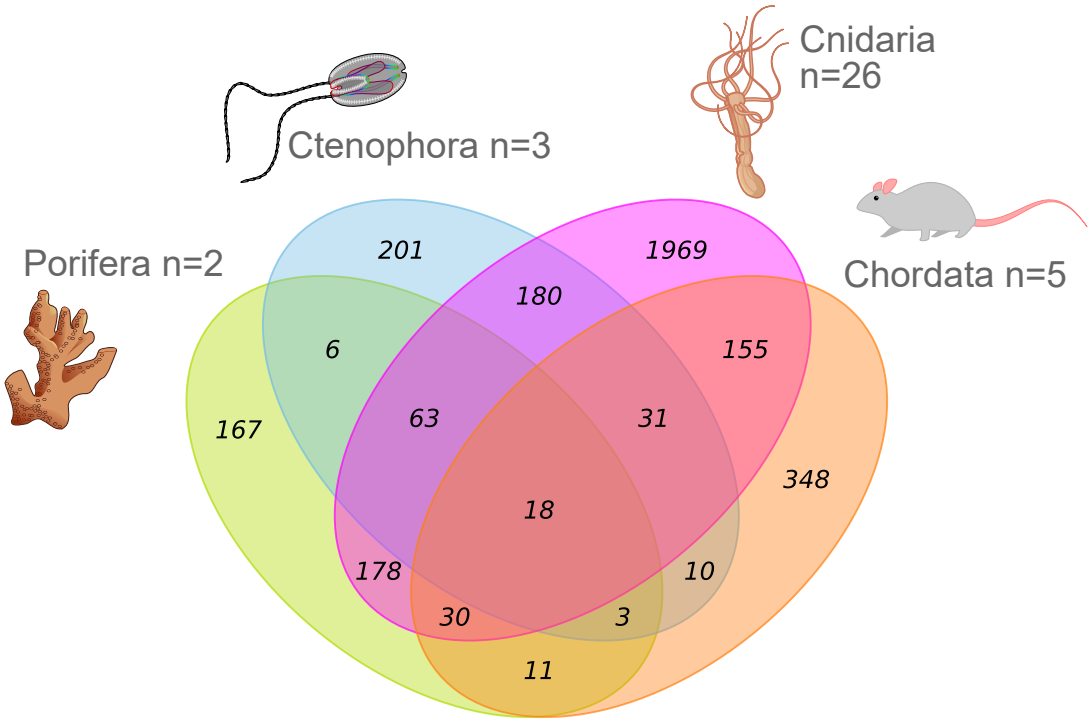


Figure 3.26. Orthogroup Comparison Between Basal Animals. The number of orthologues in each of the OGs shows the evolutionary distribution of the different OGs and the relative number of orthologues in each of the species (normalized for each OGs). (Yellow) OGs found across taxons. (Green) OGs that are enriched in cnidarians. (Orange) OGs shared between cnidarians and vertebrates.(Cyan) Taxonomically-restricted OGs in other taxa.

Table 3.6: **Comparison of *In Silico* Matrisome Sizes Across Different Taxa** *In silico* matrisome sizes predicted based on domain filtering from complete protein model predictions. The ratio of matrisomal proteins was calculated from published estimations of the amounts of coding genes (see Suppl. Table 3 for more information). The predicted matrisomes were sorted into OGs using OrthoFinder

Taxon	Species	*in silico* Matrisome Size	% of Coding Genes	# of OGs
Choanoflagellatea	*Monosiga brevicollis*	172	1.9%	103
	Salpingoeca rosetta	290	2.5%	142
Cnidaria Anthozoa	*Nematostella vectensis*	1810	3.4%	443
	Actina tenebrosa	1836	7.7%	447
	Acropora cervicornis	1289	4.3%	443
	Acropora digitifera	1434	5.5%	424
	Acropora hyacinthus	538	2.6%	251
	Acropora millepora	1345	5.0%	408
	Acropora muricata	461	3.4%	218
	Exaiptasia diaphana	1992	7.0%	468
	Montipora aequituberculata	1052	3.7%	314
	Montipora foliosa	302	2.9%	139
	Pocillopora damicornis	869	6.2%	308
	Pocillopora verrucosa	589	4.6%	257
	Porites asteroides	839	4.5%	319
	Renilla reniformis	325	2.6%	135
	Stylophora pistillata	1883	7.6%	496
	Xenia spec.	768	3.7%	266
Cnidaria Cubozoa	*Morbakka virulenta*	966	4.0%	346
Cnidaria Hydrozoa	*Clytia hemispherica*	1102	4.9%	392
	Hydra viridissima	1072	5.0%	356
	Hydra vulgaris	1074	5.4%	367
	Hydractinia echinata	1330	4.6%	484
	Hydractinia symbiolongicarpus	1085	4.9%	446
Cnidaria Scyphozoa	*Aurelia aurita*	995	3.5%	345
	Cassiopea xamachana	894	2.8%	317
Cnidaria Staurozoa	*Calvadosia cruxmelitensis*	1646	2.2%	447
Ctenophora	*Beroe ovata*	687	4.2%	301
	Mnemiopsis leidyi	600	3.6%	313
	Pleurobrachia bachei	1030	5.3%	391
Cnidaria Myxozoa	*Thelohanellus kitauei*	172	1.1%	72
	Myxobolus honghuensis	169	1.1%	40
Placozoa	*Tricoplax spec.* H2	443	3.6%	213
	Tricoplax adhaerens	419	3.5%	214
Porifera	*Amphimedon queenslandica*	1481	3.7%	246
	Ephydatia mulleri	1379	3.50%	271
Chordata	*Branchiostoma belcheri*	3543	11.20%	1175
	Branchiostoma floridae	3851	10.00%	1207
	Branchiostoma lanceolatum	3445	8.90%	1136
Published Matrisomes	*Mus musculus* (Naba et al 2012)	1096	5.0%	501
	Homo sapiens (Naba et al 2012)	1056	5.1%	496
	Danio rerio (Nauroy 2018)	1002	3.3%	376
	Bos Taurus (Listrat 2023)	1178	4.9%	473
	C. elegans (Teuscher et al 2019)	719	3.6%	179
	Drosophila melanogaster (Davis et al 2019)	641	3.6%	171
	Schmidtea mediterranea (Sonpho et al 2020)	1712	5.5%	154

Discussion

4.1 The Sequential Development of the Mesoglea

My results demonstrate that the mesoglea of *Nematostella* is a highly dynamic compartment. I was able to show that the molecular composition of the mesoglea is changing from one life stage to the next and that matrix components have different temporal dynamics. Overall, my morphological observations, the MS, and single cell expression data suggest that the mesoglea develops in a sequential manner, starting with the establishment of the BL and then subsequently adding the IM.

4.1.1 The Sequential Development of the Mesoglea

It is currently unknown which ECM factors are deposited first during development, however, it is likely that the earliest ECM factors are already deposited earlier than the life stages i investigated in my study. According to Wolpert et al. [2015], the baso-apical (ba) axis of the embryo stages is established during the blastula stage, followed immediately by deposition of ECM factors.

The initial mesoglea of larvae, consisting mostly of BL, is present in early larvae stages soon after gastrulation. It is possible, that in *Nematostella*, the PGCs play a role in the establishment of the primary mesoglea as it is the case in avian embryos [Brown and Sanders, 1991, Huss et al., 2019]. Immunofluorescence studies revealed signals for NvCollagen4 and NvLaminin signals but not for fibrillar collagens in the mesoglea of 3 dpf larvae (Figure 3.4). The primary mesoglea provides a basic structure and cell-attachment sites. These two functions are the basis for the formation of the epithelia. Furthermore, this limited ECM might facilitate the remodeling of the mesoglea during the following morphological changes.

During the primary polyp stage, the primary mesoglea is expanded as further components are added to it. Proteins such as Polydoms and BL-associated collagens 4, 6, 8 and 18 [Bernardi and Bonaldo, 2008, Halfter et al., 1998, Muragaki et al., 1995, Shuttleworth, 1997] were detected to be highly abundant in primary polyps showing that their BL is already rich in structural components.

Starting from the primary polyp stage, the fibrillar components of the IM are deposited in the center of the BL. After the metamorphosis into a primary polyp, the trilaminar structure of the mesoglea is fully formed. However, dynamic mesoglea protein changes suggest ongoing incorporation of components into both the BL and the IM during later development and until maturity in the adult stages (Figure 3.21).

A similar sequential pattern of ECM establishment has been observed during head regeneration in *Hydra* [Bergheim and Özbek, 2019, Petersen et al., 2015, Shimizu et al., 2002]. During, this process the ECM first retracts [Shimizu et al., 2002] or is digested [Bergheim and Özbek, 2019, Epp et al., 1986] and the b-a orientation of the cells is lost. Similar to embryonic development, the BL is established before

the IM during regeneration, reflecting a common principle in cnidarians [Croce and Röttinger, 2022]. The establishment of the b-a re-orientation of the epithelia (indicated by the expression of *snail* 3-12 h after wounding) aligns with the formation of the initial BL 6 h-12 h after wounding. Upregulation of IM components occurs 12-48 h after wounding, coinciding with the formation of new tentacles buds and mouth structures form.

4.1.2 Cell Type-Specific Expression of ECM Components

I have found that the mesoglea components are nearly exclusively expressed by the gastroderm (Figure 3.7 + Figure 3.8). This observation is intriguing, given that studies indicate the gastroderm of *Nematostella* bears closer resemblance to the mesoderm of triploblastic animals rather than the endoderm [Martindale et al., 2004, Seipel and Schmid, 2006, Technau, 2020, Technau and Steele, 2011]. The mesoderm is the origin of fibroblasts and other ECM secreting cell types such as osteoblasts in vertebrates [Fritzenwanker et al., 2004, Martindale et al., 2004]. Cnidarian epithelial cells are known to combine cellular functions distributed among multiple specialized cell types in vertebrates [Arendt et al., 2016, Arendt, 2008]. This sheds a new light on the evolution of fibroblasts, which might have evolved as a specialization of the gastroderm epithelia.

When the amoebocytes within the *Nematostella* mesoglea were initially examined, it was noted that they are fibroblast-like and it was assumed that these cells function as a combination of fibroblast and immune cells [Palmer et al., 2011, Parisi et al., 2021, Tucker et al., 2011]. Other studies have shown that amoebocytes produce pigments and are involved in regeneration, or have interpreted them simply as migrating epithelial cells traversing the mesoglea for relocation e.g. for the pur-

pose of regeneration [Gold and Jacobs, 2013, Meszaros and Bigger, 1999, Mydlarz et al., 2008, Palmer et al., 2011, Parisi et al., 2020]. While an amoebocyte is not annotated in the single cell transcriptomic data, the wide-spread expression of the ECM in the entire gastroderm makes it unlikely that a specialized ECM-secreting, fibroblast-like cell type exists in *Nematostella* (Figure 3.7). However, If the amoebocytes are indeed function as migratory epithelial cells, as suggested by Meszaros and Bigger [1999], they might deposit ECM components during this migration.

Whole mount stainings with Collagen 4 and PanCollagen antibodies showed punctate signals in the gastroderm (Figure 3.2). Similar vesicles were seen by Aufschnaiter et al. [2011] in the ectoderm of *Hydra* after ECM grafting. They interpreted these vesicles as lysosomes in the process of mesoglea digestion. In our case, it is more likely that the puncta represent post-Golgi vesicles containing pro-Collagen [McCaughey et al., 2019]. The synthesis of fibrillar Collagens starts with the post-translational modification of pro-Collagens in the endoplasmatic reticulum (ER) and Golgi apparatus, which are then secreted into the extracellular space, cleaved by Collagen peptidases, and forms fibers [Wu et al., 2023].

While the Collagen synthesis in vertebrates is facilitated by fibroblasts which are roaming the IM, it is currently unknown how the Collagen fiber formation is restricted to the central IM in cnidarians. This is particularly interesting as all investigated cnidarians secrete mesoglea components asymetrically. Either, as I showed for Nematostella (Figure 3.7), only the gastroderm expresses the main ECM components or as in the case of *Hydra*, different ECM components are exclusively secreted by one of the two epithelia [Epp et al., 1986, Sarras et al., 1994, Shimizu et al., 2002, Zhang et al., 2002]. This presents an interesting example of

the self-assembly of a complex biological structure. Ahn et al. [2018] demonstrated self-assembly and limited complexity *in vitro* for ECM components, showing that layered fibers can be formed *in vitro de novo* by sequential incubation of ECM components.

It is possible that the differences in expression of ECM factors by the epithelia in *Nematostella* and *Hydra* stems from the variations in body plans of anthozoans and medusozoans. Anthozoans possess mesenteries formed by gastrodermal (and pharyngeal) epithelia, separated by a mesoglea Tucker et al. [2011]. The formation of these internal structures, therefore requires a gastroderm that is capable of secreting all necessary ECM components. In the medusozoans, however the mesoglea is strongly thickened in the bells of the medusa stage. This expansion might necessitate that both epithelia contribute to ECM deposition of ECM factors. To date, *Hydra* is the only medusozoan in which the expression of multiple ECM genes has been investigated and it has been shown that ECM genes here are expressed by both epithelia [Epp et al., 1986, Sarras et al., 1994, Shimizu et al., 2002, Zhang et al., 2002]. However, *Hydra* has a reduced life cycle without a medusa stage, so this separation of labor between the epithelia might have evolved secondarily.

Wound healing of the outer bell epithelia in *Clytia* , which still has a medusa stage, involves the mesoglea [Lee et al., 2023], which the ectoderm of *Clytia* is apparently capable of ECM synthesis without the supply of additional factors from the gastroderm. A single-cell transcriptome is available for *Clytia* and delving into the *in silico* matrisome genes I predicted for *Clytia* may help investigate the cellular origin of ECM genes in the medusa.

Furthermore, the single cell transcriptome of *Nematostella* has revealed that diges-

tive gland cells also contribute to the secretion of ECM factors (Figure 3.7). This is in accordance with the observation that astacin metalloproteases and some important developmental genes such as *Dickkopf-1/2/4* are secreted from gland cells in *Hydra* [Guder et al., 2006, Ziegler et al., 2021]. It remains unclear, however, if and how the factors secreted from these cells are distributed throughout the mesoglea.

4.1.3 The Dynamics of the Developing Mesoglea

Even following the full development of the mesoglea, it appears to remain highly dynamic both morphologically and molecularly. In *Hydra*, the newly synthesized mesoglea "migrates" towards the tentacles and the foot regions where it is digested and resorbed [Aufschnaiter et al., 2011]. It remains uncertain if the mesoglea of *Nematostella* is similarly dynamic. The movement of the *Hydra* mesoglea was investigated by the injection of fluorescent Collagen mesoglea antibodies between the mesoglea [Aufschnaiter et al., 2011]. The mesoglea antibodies established in this study offer the opportunity to perform a comparative study and to visualize possible mesoglea dynamics in *Nematostella*.

During early development, the matrix is not fully matured and therefore easier to remodel based on the specific requirements of developmental processes. We have evidence of mesoglea remodeling through the abundance of various MMPs at different life stages (Figure 3.21). MMPs play a crucial role in ECM reorganization and modification [Roycik et al., 2009, Theocharis et al., 2019, 2016]. The MS analysis shows that throughout the larvae and primary polyp stages, when the mesoglea needs to be most malleable, distinct MMP are active. By combining the injection of mesoglea-specific antibodies into life animals pioneered by [Aufschnaiter et al., 2011] with MMP inhibitors such as Batimastat it would be possible

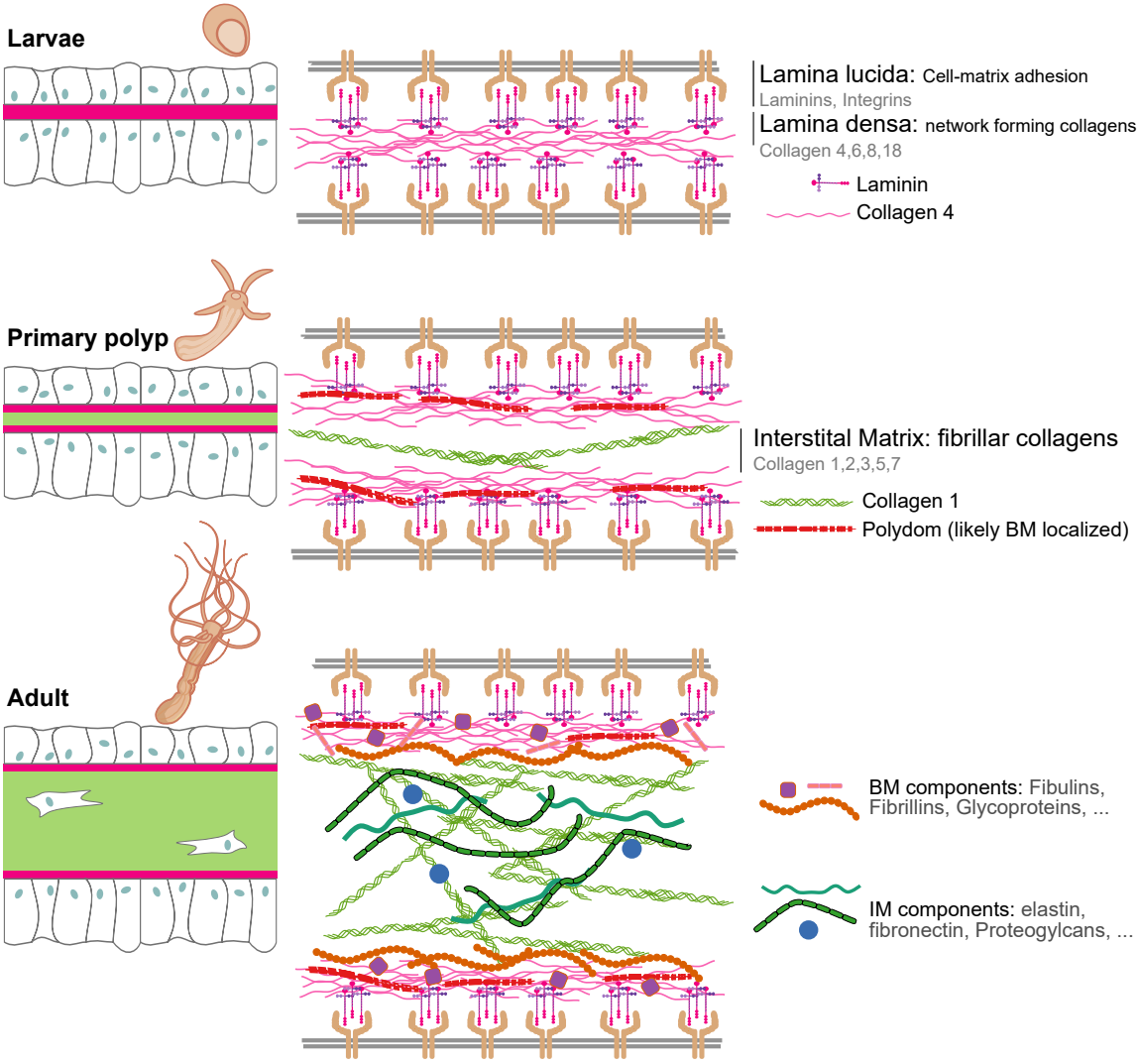


Figure 4.1. The Maturation of the Mesoglea. The mesoglea is initialized in the larvae as a thin “primary mesoglea” consisting mostly of BL, which provides an initial structure and attachment point for the epithelial cells. During metamorphosis, this basement membrane is enriched by the incorporation of additional factors such as the Polydoms and an initial IM is formed in the center of the protomatrix consisting of fibrillar Collagens. During further development, both subdivisions of the mesoglea are further maturing, and additional components are incorporated.

to observe changes in mesoglea remodeling and dynamics under conditions of reduced MMP activity [Giavazzi et al., 1998].

In the mature mesoglea, the MMPs are probably being inhibited by the abundant protease inhibitors. However, MMPs are also important in the adult stages for maintaining homeostasis. Ziegler et al. [2021] showed that an astacin is involved in Wnt degradation and for the shaping of morphogen gradients in steady-state animals. Some processes in adult *Nematostellas* are likely to involve MMP activity as well. For instance, as *Nematostella* oocytes form within the mesoglea , the release during spawning requires protease activity [Eckelbarger et al., 2008, Moiseeva et al., 2017].

Controlling protease activity through inhibitors, rather than enzyme production, likely enables finely tuned, localized regulation of MMP activity. Precise control of MMPs is also important during wound closure and regeneration, especially because aged mesoglea needs to be digested relatively quickly in a confined body region [Tucker et al., 2011]. The different roles of MMPs in these life stage explain the observed abundance of different MMPs in the mesoglea of primary polyps and adults (Figure 3.21).

4.1.4 The Physical Properties of Developing Mesoglea

Veschgini et al. [2023] demonstrated that the elasticity of *Hydra*'s mesoglea differs depending on the body region. The physical properties of hydrogels are influenced by water content, fiber thickness, fiber length, interconnectivity between fibers, and fiber density [Rigogliuso et al., 2020, Zhu et al., 2012]). Overall, the mesoglea of cnidarians has been shown to possess interesting physical properties for material science, some of which surpass the capabilities of synthetic hydrogels [Zhu et al.,

2012].

To fully understand the dynamics and the adaptive range of its physical properties it is necessary to investigate the mesoglea across different body regions. My study provides the following, initial evidence for these localized differences in the physical properties of the mesoglea.

The immunofluorescence analysis using the PanCollagen antibody revealed differential staining of tentacle and body column IM suggesting either molecular differences or impaired antibody penetration due to the tentacle mesoglea's physical density (Figure 3.3). In either case, it is very likely that the physical properties differ between these two regions.

Further support for this hypothesis comes from the observation that the aboral pore (described in Basu et al., manuscript in preparation) is surrounded by dense IM fibers ordered perpendicular to the direction of the mesoglea (Figure 3.5). This dense and ordered formation of IM fibers suggests that the physical properties of the mesoglea at the aboral pore vastly differ from that of the surrounding mesoglea. Stokkermans et al. [2022] showed that attachment and physical stretching is essential for the elongation of *Nematostella* larvae. During this process the embryo pulls itself into an elongated shape. The dense IM fibers at the aboral pole might be essential to counteract the pulling forces generated during body elongation. Consistently, the IM appears to be the main factor for the differences in morphology/physical properties, while changes of the BL are not visible.

It is likely that the ECM evolved in the same sequtial manner as its appearance during development [Huxley-Jones et al., 2009, Hynes, 2012, Özbek et al., 2010]. Initially, the BL evolved to facilitate cell attachment to inorganic substrates or the

formation of small colonies in unicellular organisms. Basic attachment of cells to surfaces or to each other is known from many unicellular organisms such as choanoflagellates, other protists, and from bacterial mats [Abedin and King, 2010, Kang et al., 2021, King et al., 2003]. Subsequently the factors facilitating attachment radiated and the ECM repositiry became expanded until the evolution of IM factors allowed for the development of even thicker and diverse matrices, which are more easily adapted.

4.1.5 Conclusion

Both morphological and molecular evidence confirms that the mesoglea of *Nematostella* is highly dynamic during development. Initially constituted by BL, the mesoglea undergoes a significant expansion in both morphological thickness and molecular diversity through the incorporation of IM. The IM is first deposited in the branching points of the septs and then forms the intermittent layer between the BL. I provide evidence that the morphological thickening of the mesoglea is mainly driven by structural components such as fibrillar Collagens and that each life stage possesses a distinct composition of ECM remodeling factors and ECM-associated proteins.

4.2 The Definition of ECM Components Remains a Challenge

A core aim of my thesis was the definition of *in silico* matrisomes. However, the definition of ‘ECM factors’ is relatively challenging both conceptually and technically.

The definition of the matrisome used in this thesis largely follows the concept established by Hynes and Naba [2012] who categorize the matrisome into two subdivisions: The core matrisome and the ECM-associated proteins [Naba et al., 2012, Hynes and Naba, 2012]. These divisions are based on the occurrence or absence of indicator domains, with the core matrisome predominantly comprising structural components, while ECM-associated proteins contain ECM regulators and growth factors interacting with the core ECM. According to this definition, membrane-bound receptors like CAMs, which only partially extend into the ECM or transmit signals into the cells, are excluded from the matrisome classification and are instead categorized under the broader 'adhesome'. As cell adhesion can be counted as one of the major functions of the ECM. I decided to include these proteins along with several other ECM-related proteins not fitting into either the core matrisome or the ECM-associated division into a separate third category named "Other".

This "Other" division also contains various potential cnidarian-specific factors of the cnidocytes. As these factors do not possess equivalent vertebrate orthologues, they were not considered in the exsisting matrisome definition. In fact, the analysis of the *Drosophila* [Davis et al., 2019] and *C. elegans* [Teuscher et al., 2019] *in silico* matrisomes likewise resulted in 328/12 unclassified genes, termed "organism-specific proteins". This nomenclature, however, is somewhat misleading, as certain protein families, such as the tenascins and some other factors, which evolved in the vertebrate lineage [Hynes, 2012, Hynes and Naba, 2012] are not designated as termed organism-specific but are instead integrated into the core matrisome or ECM-
associated divisions.

This underscores a vertebrate-centric bias in current matrisome definition protocols, probably resulting in the 3-fold difference in size between the *in silico* matrisome of *S. mediterranea* presented by Cote et al. [2019] which likely excluded all species-specific proteins and the experimental matrisome of the same species by Sonpho et al. [2020]. This is an important fact to keep in mind when defining the matrisomes of invertebrates.

In fact, one could argue that a core matrisome signifying 'the essential, conserved factors' can only be accurately defined under the consideration of vertebrate *and* invertebrate species and across an evolutionary comparison that includes basal animal groups. The annotation of a cnidarian matrisome is an important step towards a better understanding of the matrisome and its definition.

4.2.1 Challenges of the In Silico Matrisome Definition

Sequence homology comparisons and GO-term filtering can be very accurate, but they rely on evolutionary conservation, well-assembled protein models, and of course the prior identification of ECM factors. Even those cnidarian sequence datasets with the highest assembly quality are fragmentary and annotated algorithmically by predicted protein models, which makes homology comparisons and GO-term filtering unreliable. In general, my analysis has proven how challenging the annotation of basal animal protein databases can be. These databases are annotated automatically by propagating descriptions based on sequence similarities [Haft et al., 2018, Li et al., 2021, Tatusova et al., 2016]. While this approach is generally useful, the evolution of the ECM by domain shuffling leads to similar sequences appearing in functionally diverse proteins [Engel, 1996, Hohenester and Engel, 2002, Patthy, 1999]. This results in a significant number of miss-annotated

ECM genes.

One of the goals of my thesis was to define the matrisomal proteins in *Nematostella* to enable sequence homology comparisons in cnidarian species.

It has to be noted, though, that the domain-based identification of ECM proteins at the time of writing is not sufficient to identify all ECM proteins. Even in humans and mice some factors such as dermatopontin and dentin sialophosphoprotein can only be identified manually [Naba et al., 2012]. Notably, some cnidarian-specific domain/motif markers are not registered in any of the domain databases. For instance, the Minicollagen CRD found in many structural ECM-like components of cnidocysts can therefore be used as a diagnostic domain for cnidocyst-specific protein components [Beckmann et al., 2015, Beckmann, 2013, Engel et al., 2002, Meier et al., 2007, Tursch et al., 2016]. It is highly likely that there are more ECM-defining domains, particularly in invertebrates, which are yet to be discovered. It is possible to include such novel ECM domains in a filter system e.g. by employing custom regular expression (RegEx) search (See Appendix 6.2 for possible RegEx patterns). However, this approach requires detailed preliminary knowledge of the sequence patterns and manual alterations, which is why I did not include it in the annotation pipeline presented in this study.

Moreover, the set of defined ECM domains requires constant review and updates. In this study, I relied on the domain list provided by Naba et al. [2012], but since this list was published, more ECM proteins have been found and additional domain types have been deposited in the domain identification databases. A revised list of domains is likely to reveal more ECM factors.

The fragmentary nature of the sequence datasets of non-standard model organ-

isms is additionally complicating the search for matrisome components. In many cases, my analysis resulted in shortened protein models or in some cases in proteins that were assembled fragmentary sequence models or previously published compllete sequences. An example for this is NvNCol-3 for which only a partial sequence was identified in the available proteins sequence dataset and which was completed based on published sequences [David et al., 2008].

Future approaches should consider reviewing and expanding the ECM domain definitions and using multiple variations for the common matrisome domains to ensure that these variations are also captured.

4.2.2 Challenges of Experimental Matrisome Extraction

Similar to the difficulties of the bioinformatic approach, the experimental definition of the matrisome faces several challenges. Firstly, ECM material needs to be extracted without cellular contamination. This step proves to be exceptionally challenging, as most organisms possess cells that reside within the ECM, making their removal a complex task. Different extraction techniques have been used depending on the organism and specific techniques. These can be categorized as mechanical, chemical, and enzymatic decellularization techniques. All of these techniques introduce some sort of bias, because they rely on enzymatic digestion of structural components, long washing steps which remove soluble fractors or are unable to completely remove cell debris.

Cnidarians present an interesting case in this respect, with many species possessing nearly cell-free mesogleas and a simple body structure, facilitating decellularization [Davis and Haynes, 1968, Sarras et al., 1991, Sarras and Deutzmann, 2001]. In addition, medusozoans possess thick mesogleas within their bells which

can be extracted easily in large amounts. This availability of ECM material en mass was utilized by material scientist to investigate the properties of biological hydrogels [Yang et al., 2015].

Specialized ECMs as in cnidocysts or the substrate attachment matrices requires specialized extraction methods [Davey et al., 2021, 2019, Liu, 2022, Marchini et al., 2004, Rodrigues et al., 2016, Weber, 1989, Young et al., 1988, Zenkert et al., 2011]. Moreover, different life stages exhibit distinct ECM compositions, requiring adapted extraction protocols based on size and fragility. Similarly, the ECM of different tissues and organs in vertebrates differ in composition and physical properties [Goh et al., 2021, 2020, McCabe et al., 2022]. Therefore, a diverse array of ECM extraction methods is needed to investigate the full range of ECM across life phases and body regions.

The extraction method used in this study required me to cut open the body wall of the adult *Nematostella*s to facilitate more complete decellularization, introducing an injury signal just before the extraction of the mesoglea. Studies in *Hydra* and *Clytia* have shown that the mesoglea is heavily involved in regenerative processes [Lee et al., 2023, Sarras, 2012]. Structural mesoglea factors have only been shown to be upregulated 3 h after an injury and are therefore likely not affected by this protocol [Bergheim and Özbek, 2019]. However, other injury signals react much faster to wounding [Tursch et al., 2022], therefore this protocol might affect the abundance of soluble factors.

Even after successful extraction of pure material, solubilizing the ECM remains challenging, due to the high grade of cross-linkage between ECM components, their high molecular weights and fibrillar nature. Usually, strong detergents, high temperatures and enzymatic digestion are used to decellularize and solubilize

ECM samples [Fernández-Pérez and Ahearne, 2019, McCabe et al., 2021, McInnes et al., 2022, Sonpho et al., 2020]. Sonpho et al. [2020] have shown that different solubilization protocols yield different subsets of the matrisome with only partial overlap. Therefore, the complete experimental investigation of the ECM by MS likely requires the investigation of different life stages and body regions with multiple extraction and solubilization protocols. Such elaborate experiments are more feasable in animals with limited body complexity and short development such cnidarians.

My MS study focused on only three life stages and did not consider the differences between different body regions and can therefore only be seen as an initial survey of the cnidarian matrisome. Future, more detailed studies might therefore uncover additional interesting interactions between the matrisome components. This might also close the gap between the size of the *in silico* matrisome and the much smaller number of the experimentally confirmed matrisome factors.

4.2.3 Functional Investigation of ECM components

Ultimately, the definitive confirmation of a specific protein's inclusion in the matrisome relies on demonstrating its localization within the ECM. To achieve this, antibodies for the different predicted ECM components were raised and then investigated in a tractable model organism. In this thesis, I presented antibodies for the entire ECM and, more specifically, the BL. It is likely that not all matrisome components are solely confined to the mesoglea or that they are uniformly localized within it, as I demonstrated for the targets of the PanCollagen antibody (Figure 3.3). Reliable mesoglea antibodies, such as the NvLaminin and NvCollagen-4 antibody are important references for future investigations into ECM components.

The prospective functional analysis of ECM components poses several challenges for the future. Many of the matrisome proteins identified during my study exhibit multiple isoforms, and the high similarity among different proteins suggests potential redundancies in their functions. In addition, the interactions between the different ECM factors within the ECM are complex and not well understood. Consequently, phenotypes resulting from ECM protein knockdowns/-oust might be subtle and challenging to interpret without detailed ultrastructural investigations unless a factor crucial of cell-matrix connection is affected. At the same time, the functional investigation of these components has far-reaching perspectives. It is estimated that a large number of human diseases are caused by, or enhanced by yet pooly understood ECM deficiencies or misregulations [Bonnans et al., 2014, Lu et al., 2012, 2011, Theocharis et al., 2019].

4.2.4 Conclusion

My study establishes *Nematostella* as a tractable model organism for multifaceted investigations into the functions of relevant ECM components. Overall, the *in silico* matrisome presented in this thesis can be considered a reliable approximation of *Nematostella*'s complete matrisome, and a significant step forward in the identification of cnidarian ECM factors. However, the experimental MS-based matrisome is unlikely incomplete as I only investigated three life stages and applied only a single decellularization and solubilization method. To achieve a comprehensive molecular map of the *Nematostella* matrisome, further enhancements to sequence and domain annotation databases are necessary for refining bioinformatic annotations. Additionally, more diverse approaches of decellularization and protein solubilization should be explored to achieve a more holistic MS investigation.

4.3 Cnidarian Specializations of the ECM

4.3.1 The Building Blocks of the Capsule Types

Cnidocysts are complex organelles that self-assemble from many different components. To facilitate this assembly, the components need to be secreted in a distinct temporal sequence and then interact at the correct time and space. The importance of temporal sequencing in ECM 2D fiber formation has been shown *in vitro* by Ahn et al. [2018]. The decision-making process directing the trajectory towards different capsule types has been investigated at the level of transcriptions factors by Steger et al. [2022]. However, the capsule-specificity and temporal sequence of the ECM components that shape the capsule has not been investigated in *Nematostella*.

The development of cnidocytes occurs in multiple steps (Figure 4.2). Initially, protein components are secreted into a post-Golgi vesicle. Some of these components get densely packed into pseudo-crystals [Adamczyk et al., 2010, Garg et al., 2023]. This is followed by the differentiation of the different capsule types, which requires the deposition of capsule-specific components. The evolution of novel components is one of the driving forces of capsule type innovations [David et al., 2008]. Similar to the initial components these capsule type-specific factors are deposited in the vesicle and partially form clusters or pseudo-crystals. Finally, the capsules mature and become active. Upon maturation, the cnidocytes switch their expression profile and become a distinct mature cnidocyte cell type. This cell type expresses the factors required for maintenance and the receptors essential for the chemically and mechanically controlled discharge of the capsule. This trajectory

pattern of divergence into capsule-specific development and convergence to the mature cnidocytes during capsule development has also been observed in the single cell transcriptomes of *Hydra* [Siebert et al., 2019], *Xenia* [Hu et al., 2020], *Hydractinia* (*Hydractinia* Genome Project), *Clytia* [Chari et al., 2021], and *Exaiptasia* (Sebastian Gornik, personal communication).

Nematostella's capsule types exhibit differences in overall morphology, size, and tubule and spine architecture [Zenkert et al., 2011]. Analysis of cnidocyte-specific matrisome expression reveals type-specific variants of Minicollagens and NOWA proteins. Two Minicollagens emerge as the most abundant components of the capsule wall, while NOWA plays an integral role in tubule assembly [David et al., 2008, Garg et al., 2023, Holstein et al., 1994]. Spirocysts, characterized by a spiral tuble lacking spines, express two specific NOWA proteins (NvNOWA−2 | NV2.10392 and NvNOWA−4| NV2.10390) as well as two specific Minicollagen variants (NvNCol-5 | NV2.11802 and NvNCol-like 3| NV2.11816). These might be responsible for the unique morphology of this capsule type. Columbus-Shenkar et al. [2018] describe a pharyngeal cnidocyte type which is characterized by the expression of nep8. This cell type expresses only Minicollagen-1 (NV2.12621) as structural component supporting the idea that it does not form a capsule and instead acts as a gland cell secreting venom factors into the pharyngeal space. I was able to show that the nep8 cells express a number of protease inhibitors, such as (Cytostatin-8 | NV2.20322 and Kunitz-type protease inhibitor | NV2.10172), which might function as venom factors as well.

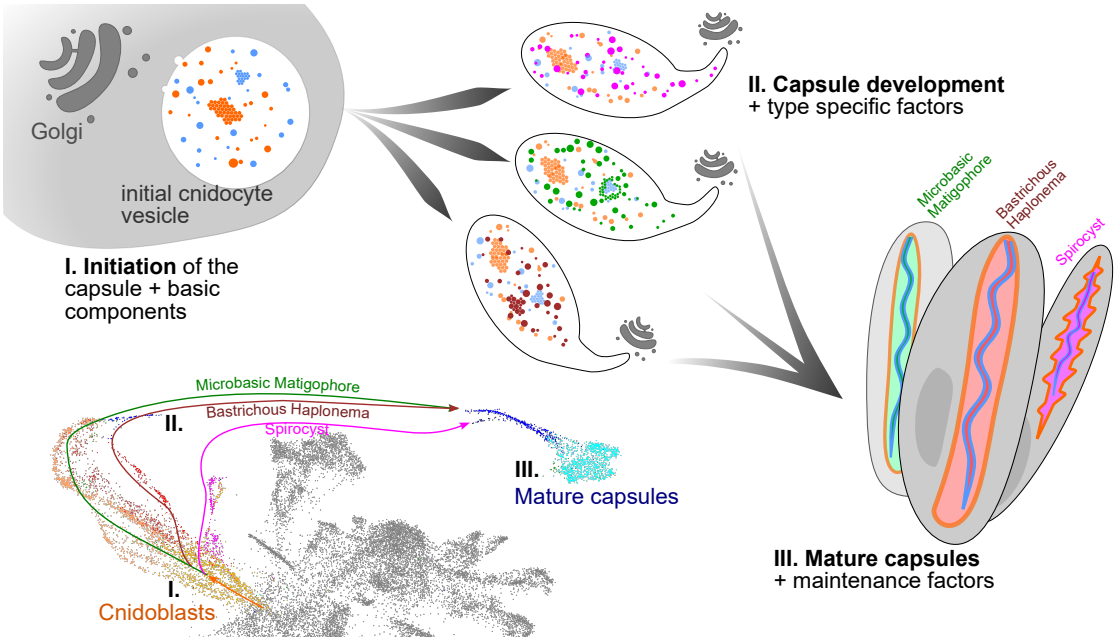


Figure 4.2. The Sequential Secretion of Cnidocyte-Specific Proteins. From the perspective of gene expression, cnidocyte development can be split into three phases. (I.) During the initiation phase all basic cnidocyte factors are expressed and secreted into a post-Golgi vesicle, where some of them obtain a crystalline organization [Garg et al., 2023]. (II.) During development, capsule type-specific genes are expressed and the developmental trajectories (bottom left, from Steger et al. [2022]) split into separated paths. Finally, during the maturation phase (III.) these paths converge as the capsules become mature. The fully mature cnidocytes express genes necessary for the maintenance and discharge of the capsules.

4.3.2 A New Type of Long Minicollagens

Minicollagens have been known to be the structural components of the cnidarian cnidocytes for many years [Kurz et al., 1991]. The identification of Spongins, short chain Collagen family originally found in sponges, shows that the basic building block of Collagens evolved before the eumetazoan split [Aouacheria et al., 2006, Garrone, 1999, Garrone and Trelstad, 1984]. It is therefore likely that these short chain Collagens are the origin of the Minicollagens. My analysis recovered a mega-

Minicollagen, characterized by an unexpectedly long Collagen sequence, bridging the gap between short-chain Collagens and the longer fibrillar Collagen types. This protein is expressed in haplonemas and mastigophores and previous immunofluorescence staining done in our lab showed that the protein is indeed localized in the capsules of these cnidocytes (data unpublished). While the protein model is well supported by read mapping attempts to clone the full gene or to confirm the size of the protein product by PCR amplification were not successful (data not published). Therefore, we cannot completely rule out the possibility that this gene represents either an assembly mistake or an expression cassette that enables the assembly of diverse short Minicollagens (Figure 4.3). Minicollagen genes in cnidarians have been shown to be in close proximity to each [Kyslík et al., 2021] other or even arranged in expression cassettes sharing the same N'-terminal signal- and propeptide [David et al., 2008]. Future, experiments will need to test which of these hypotheses are true. Should the *in silico* predictions hold true, rendering the mega-Minicollagen as a full-length protein, it would offer a compelling model to unravel the evolutionary transition from fibrillar Collagens of the ECM to cnidocyte-specific Minicollagens,

4.3.3 Cnidarian Collectins Distinct from Classical Collectins

One of the interesting factors of the cnidocyst matrisome with unknown function are the CnCollectins consisting of a Collagen and galactose lectin domain that were first described by Grasso et al. [2008] as anthozoan-specific proteins. They showed that Anthozoans possess CnCollectins in addition to vertebrate collectins. The domain architecture of vCollectin consist of a Collagen and C-Lectin domain. While the vCollectin was shown to be expressed in aboral ectodermal cells of

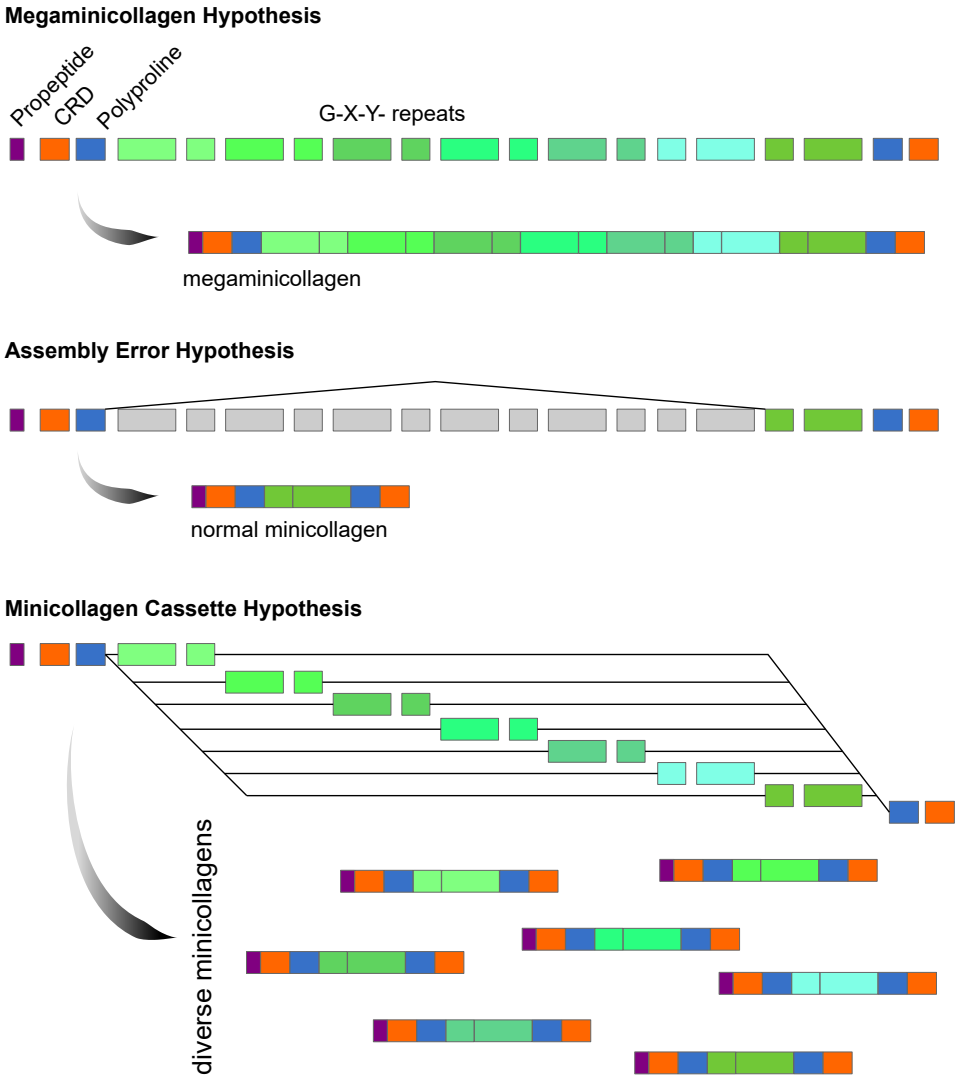


Figure 4.3. Different Hypotheses of the Protein Product of the MegaMinicollagen Gene (A) The full-length mega Minicollagen is well supported by read mapping and exon prediction in *in silico* analysis, however, the full-length product has not yet been confirmed experimentally therefore alternative hypotheses are possible. (B) The mega Minicollagen might be the result of an assembly error due to the highly repetitive nature of the Collagen helix region and might actually only result in a normal sized Minicollagen. (C) Another alternative is that the gene is an expression cassette which is able to assemble multiple Minicollagen variations that share the same N- and C-terminus.

Acropora, the CnCollectins were either expressed in putative cnidoblasts or in one case in the gastroderm [Grasso et al., 2008]. Based on these expression patterns, Grasso et al. [2008] proposed that CnCollectins play different roles, depending on their localization. They proposed a role in symbiont uptake in the gastroderm and as structural components of the cnidocytes due to their Minicollagen-like short Collagen domain.

Knack [2011] showed that CnCollectins are found in hydrozoans as well and reclassified them as cnidaria-specific. He noted their similarity to vertebrate collectins (vCollectins). These, have been shown to be involved in adhesive processes of the immune response [Gupta and Surolia, 2007, van de Wetering et al., 2004]. Consequently, Knack [2011] classified the CnCollectins as cnidarian-specific cell adhesion factors of the immune system in line with the previously proposed function in symbiont recognition. As the vCollectin (NV2.23165) and one of the CnCollectins (NV2.5200) are expressed in epithelial cells in *Nematostella* according to the single cell data (Figure 3.11) these might be involved in the immune response.

The remaining three CnCollectins (NV2.4511, NV2.5188, NV2.5196) are exclusively expressed in cnidocytes in accordance with the *in situ* hybridization data from Grasso et al. [2008] in *Acropora*. The function of the CnCollectins in cnidocytes is unclear. Their domain architecture suggests that they can connect Collagenous matrices to carbohydrates suggesting structural role; however, it is unclear if they are components of the capsule wall, tubule apparatus or even a venom factor.

4.3.4 Conclusion

My results provide deeper insights into the cnidarian-specific cnidocyte matrisome. A clear definition of the components of this matrisome subtype is needed for the in-

vestigation of the evolutionary origins of the cnidocyst structural components and to differentiate between the cnidocyte-specific factors and the proteins, which make up the mesoglea. This is essential for establishing cnidarians as model systems for ECM research. The inclusion of cnidocysts-specific proteins could potentially confound future investigations aimed at unraveling general ECM factors, and vice versa. 4.4. Polydom Knockdown .

4.4 Polydom Knockdown Affects the Epithelia and the Formation of Mesenteries

My results revealed the Polydom/SVEP1 family as an interesting new component of the ECM. Polydoms were first identified as new pentraxin domain-containing proteins, characterized by their large size and rich variety of domains [Gilgés et al., 2000]. They have been shown to be involved in various processes, such as cell adhesion via integrin interaction [Sato-Nishiuchi et al., 2012, Shur et al., 2006, Winn et al., 2007], blood and lymphatic vessel formation [Karpanen et al., 2017, Michelini et al., 2021, Morooka et al., 2017, Yoshida et al., 2021] and epidermal development [Samuelov et al., 2017].

In humans and mice, Polydoms are widely expressed [Samuelov et al., 2017] and have been identified in the transcriptome of osteoblasts and the placenta where they are involved in the connection of mother and child [Shur et al., 2006]. In contrast, the CnPolydoms of *Nematostella* which were first found by Knack [2011], are expressed in mucus gland cells (and in the case of Polydom-related and Polydom-like 2 in sensory secretory cells). CnPolydoms are large proteins that are not likely to traverse the dense fiber mesh of the ECM. In vertebrates, Polydoms have been

shown to be a component of the BL where they seem to directly interact with integrins [Sato-Nishiuchi et al., 2012]. Notably, their abundance in *Nematostella* correlates with the primary polyp life stage, aligning with the peak presence of other BL components such as Collagen 4.

The CnPolydoms of *Nematostella* offer intriguing insights into ECM evolution. Firstly, *Nematostella* possesses multiple orthologues while for *Hydra* and *Hydractinina* only one CnPolydom has been described [Schwarz et al., 2008]. In addition to this, the three different *Nematostella* orthologues differ in the amount and identity of their domains, bridging the gap between the vertebrate Polydoms and the CnPolydoms of *Hydra* and *Hydractinia* [Schwarz et al., 2008]. NvPolydom-1 is the most vertebrate-like CnPolydom, possessing only EGF and Sushi domains following the Pentraxin domain (Figure 3.16). In the other orthologues, a coagulation-factor and a Pan/Apple domain are added to the structure and NvCnPolydom-3 resembles the previously described CnPolydoms. As the Polydoms are apparently conserved from cnidarians to vertebrates, they could be an essential ECM factor worthy to be included in the basic basal membrane toolkit [Hynes, 2009].

We were surprised to find an LRP-like Polydom (`NV2t000788001.1`) which appears to be a fusion product of the LRP N'-terminus and a C-terminus of Polydom, expressed in cnidocytes (Figure 3.17). Barring a possible assembly error, this is a striking example of how entire domain complexes can be shuffled between ECM proteins through domain shuffling [Adams and Engel, 2007, Engel, 1996, Hohenester and Engel, 2002, Patthy, 1999, Whittaker et al., 2006]. Given that neither Polydoms nor LRPs are typically found in cnidocytes, the function of this protein remains unknown.

The significance of Polydoms lies in their size and the multitude of domains, sug-

gesting a role as ”Swiss army knife”-like proteins capable of adapting to diverse functions through specific folding or domain cleavage. Activation of ECM proteins by cleavage is common, and it has been suggested that some ECM proteins that possess e.g. poly-EGF motifs (as do Polydoms) might serve as reservoirs for growth factors which can be shed to perform various function [Bonnans et al., 2014, Giannelli et al., 1997].

The presented preliminary siRNA knockdown results suggests that Polydom-related (NV2t011255001.1) has an essential role for development and epithelia organization. I observed that the embryos failed to elongate and form internal structures but were still able to form head structures, such as tentacles and mouths. In addition to this, the epithelia organization seemed to be disturbed and the cells were only loosely connected to the mesoglea. This fragility of the epithelia aligns well to the findings of Samuelov et al. [2017] who showed that a knockdown of Polydoms and treatment with antibodies against Polydom in zebrafish and mice leads to abnormal epidermis morphology likely due to affected cell-matrix adhesion.

The observed lack of mesenteries and elongation in the knockdown may indicate a missing feedback mechanism between cells and the mesoglea, crucial for these morphological processes. Thus, an individual ECM protein can have significant impact on the overall development of *Nematostella*. In line with the defects in elongation in these phenotypes, the human genome-wide association score catalogue shows that four single nucleotide polymorphisms in the human SVEP-1 gene are associated with body height (GeneCard:SVEP1). The role of ECM proteins in development and especially mesentery formation has also been demonstrated before forBMP2/4 kockdown [Leclère and Rentzsch, 2014, Saina et al., 2009, Stokkermans et al., 2022]. A knock-down of this protein leads to severely elongated (which

I called 'spaghettified') embryos that likewise did not develop internal structures. A BMP gradient is important for the symmetry break, which induces a bilaterian body plan in animals. Interestingly the mesentery organization of Anthozoans such as *Nematostella* is pseudo-bilaterian and controlled by Hox-like genes [Berking, 2007, DuBuc et al., 2018, Steinworth et al., 2023, Technau and Genikhovich, 2018]

While the knockdown of the short Polydom-related protein resulted in this interesting phenotype, the preliminary knock down of the full length NvCnPolydoms did not result in any significant phenotype, which might be explained by redundancy between the three CnPolydoms of *Nematostella*. Here a simultaneous knockdown of all three genes might elucidate the function of CnPolydoms during development.

In future studies further knockdowns or knockouts of the other members of the Polydom family in combination with immunofluorescence and electron microscopy, might give further insights into the general function of the Polydom core motif and the specific interactions of the different Polydom-like variants.

4.5 The Matrisome of *Nematostella* in the Context of Evolution

The base of the animal kingdom is very interesting for the ECM evolution, as the development of the core matrisome is one of the prerequisites for multicellular animals. While it was not a focus of my thesis, I partially elucidated the evolutionary history of the ECM at the base of the animal evolution by expanding the scope of the *in silico* matrisome prediction to multiple early diverging animal species.

Unfortunately, proteomic datasets of organisms at the base of animal evolution are

relatively sparse and, in many cases, fragmentary and not curated. Therefore, all purely bioinformatic analyses as the one in this study have to be considered with care. Fragmentary or misannotated sequences may lead to misleading interpretations. This is especially true for the protein families of the ECM where domain shuffling is common and redundancies are regular occurrences.

The examples of the two sponge species (*A. queenslandica* and *E. muelleri*) show that *in silico* matrisome size alone can be misleading. In these two sponges, a few families of fibronectin-like proteins are expanded, makeing up nearly 30 % of the overall matrisome size. As both species belong to the order Haplosclerida, it is unclear if this is a general feature of sponges or just of the members of this order. This shows the importance of defining a measure for the diversity of the matrisome in addition to its size. The analysis of OGs proved to be invaluable for the large matrisome datasets. However, it is affected by the fragmentary state of the sequence datasets, and in many cases, overestimates the actual diversity of these datasets. In addition, the software used to infer OGs (OrthoFinder), relies on iterative gene- and species tree predictions for the prediction of OGs [Emms and Kelly, 2019]. It is therefore subject to the limitations of phylogenetic analysis, particularly the evolutionary sampling and the quality of the input datasets can affect the results of such phylogenetic analyses [Nabhan and Sarkar, 2012, Smith et al., 2021].

The *in silico* matrisomes which I predicted for these taxa seem to support a gradual evolution of additional ECM components. The size and complexity of the *in silico* matrisomes (Figure 3.24) roughly reflects the generally accepted evolutionary history of early animals. Notably, single-celled chanoflagellates, possessing ECM-like proteins but lacking complete ECMs, exhibit the smallest and least com-

plex *in silico* matrisomes, followed by placozoans, ctenophores, and cnidarians.

Within the cnidarians there was a marked difference between the anthozoans and medusozoans, with anthozoans being phylogenetically placed closer to the bilaterians than to the other cnidarians (Figure 3.23).

This might be explained by adaptations of the mesoglea to the free-swimming lifestyle of medusozoans. In most medusozoan jellyfish, the mesoglea is much thicker and can act as a counterbalance for muscle activity [Gladfelter, 1973, 1972]. Interestingly, anthozoans also show additional morphological similarities to bilaterians in that they possess internal structures such as mesenteries and septas and that in many species these features are arranged in pseudo-bilaterian anatomy [McFadden et al., 2021]. My preliminary knock-down experiments have shown that a knockdown of matrisomal proteins might influence the development of mesenteries. It might therefore be possible that some of the OGs shared between anthozoans and vertebrates are required for the development of internal bilateralism.

4.5.1 The Exceptional Matrisome of the Ctenophores

The evolutionary position of ctenophores has been a matter of debate among different researchers. Based on their anatomical features they have been placed in within the eumetazoans, however, several molecular features have placed them as a sister group to eumetazoans [Jékely et al., 2015, Daley and Antcliffe, 2019] suggesting that they are molecularly isolated from other extant species. Multiple theories have been suggested to bridge this apparent divergence [Jager and Manuel, 2016]. A recent analysis based on synteny patterns, defines ctenophores as the sister to all animals, while sponges, bilaterians and cnidarians show derived chromosome rearrangements [Schultz et al., 2023]. Concerning the ECM of

ctenophores in particular, Draper et al. [2019] noted that many ECM genes are not conserved in ctenophores.

In my analysis, which only considers similarities of the matrisomes, the ctenophores were placed as a sister group to all other species and clustered with the myxozoans (Figure 3.23). This can be explained by the high number of taxon-specific OGs that might represent unique ECM adaptations. Interestingly, ctenophores share a large number of OGs with cnidarians, which suggests that a phylogeny that excludes all taxon-specific OGs might result in ctenophores being placed close to cnidarians. This supports the hypothesis proposed by Daley and Antcliffe [2019] that ctenophores likely evolved from cnidarian-like organisms.

At the same time morphological analysis of the ctenophore 'mesoglea' has revealed complex chamber and tube structures that are not found in cnidarians [Bargmann, 1972]. Ctenophores and cnidarians possess very similar lifestyles as planktonic predators, however, there are also marked differences. Cnidarian medusa forms swim by utilizing full body contractions while ctenophores use cilia organs and smaller scale deformations of the mesoglea to adjust the swimming directions. Both methods involve compression of the mesoglea, however, the physical stress exhibited differs [Bargmann, 1972, Gladfelter, 1973, 1972]. In this context an investigation of the differences and similarities of the matrisome of cnidarians and ctenophores might not only help to validate the phylogenetic position of the ctenophores but also shine light on different ways of how the ECM adapts to differences in life-style.

General Conclusion

Overall, my study provides a significant step forward in the investigation of the early animal ECM. I was able to establish a generalized pipeline for the *de novo* prediction of ECM components from raw protein models. This pipeline can be used to quickly generate evolutionary diverse matrisome datasets that enable taxonomic comparisons and the identification of taxon-specific ECM factors.

With the newly curated matrisome of *Nematostella* presented here, I contribute the first completely annotated cnidarian matrisome. I was able to show that its matrisome has a size of 554 proteins (+ 289 non-classical ECM components). This is comparable to previously published invertebrate matrisome sizes. This dataset will be a valuable reference for future investigation of the ECM in early branching animals. As cnidarians are ideal model organisms for the investigation of ECM factors at the level of whole organisms the presented dataset will enable future detailed analyses of individual ECM factors and their interactions.

I was able to supplement the definition of the matrisome of *Nematostella* with data from single cell transcriptomes investigating the cellular origin and secretion of the matrisome. I showed that the ECM of *Nematostella* is secreted by mesoderm-

like gastroderm without contribution of the ectoderm. In addition, I defined 182 cnidocyte-specific matrisome components and showed their capsule type-specific expression,adding insights into molecular differences that lead to different capsule type morphologies.

Through the study of three life stages by MS, I showed that the ECM is compositionally dynamic during development. Together with morphological studies the data showed that the BL develops early in embryogenesis and that the IM is subsequently added during and after metamorphosis. According to my analysis, every life stage possesses a distinct ECM composition suggesting that the developmental dimension is important for the functional investigation of ECM components.

Through my preliminary knock down of Polydom-related NV2.2824 I presented new evidence that individual ECM factors can vastly impact morphogenesis during early development. This knockdown affects the cell-matrix connection but has the additional, intriguing phenotype to inhibit the elongation of the embryo and the formation of the mesenteries, indicating an acient role of Polydoms for epithelial organisation.

Bibliography

M. Abedin and N. King. Diverse evolutionary paths to cell adhesion. *Trends in cell biology*, 20(12):734–742, 2010. doi: 10.1016/j.tcb.2010.08.002.

P. Adamczyk, S. Meier, T. Gross, B. Hobmayer, S. Grzesiek, H. P. Bächinger, T. W. Holstein, and S. Özbek. Minicollagen-15, a novel minicollagen isolated from hydra, forms tubule structures in nematocysts. *Journal of molecular biology*, 376(4):1008–1020, 2008. doi: 10.1016/j.jmb.2007.10.090.

P. Adamczyk, C. Zenkert, G. P. Balasubramanian, S. Yamada, S. Murakoshi, K. Sugahara, J. S. Hwang, T. Gojobori, T. W. Holstein, and S. Özbek. A non-sulfated chondroitin stabilizes membrane tubulation in cnidarian organelles. *The Journal of biological chemistry*, 285(33):25613–25623, 2010. ISSN 0021-9258. doi: 10.1074/jbc.M110.107904.

E. D. M. Adams, G. G. Goss, and S. P. Leys. Freshwater sponges have functional, sealing epithelia with high transepithelial resistance and negative transepithelial potential. *PloS one*, 5(11):e15040, 2010. doi: 10.1371/journal.pone.0015040.

J. C. Adams. Passing the post: roles of posttranslational modifications in the

form and function of extracellular matrix. *American journal of physiology. Cell physiology*, 324(5):C1179–C1197, 2023. doi: 10.1152/ajpcell.00054.2023.

J. C. Adams and J. Engel. Bioinformatic analysis of adhesion proteins. *Methods in molecular biology (Clifton, N.J.)*, 370:147–172, 2007. doi: 10.1007/978-1-59745-353-0{\textunderscore}12.

S. Ahn, K. Y. Lee, K. K. Parker, and K. Shin. Formation of multi-component extracellular matrix protein fibers. *Scientific reports*, 8(1):1913, 2018. doi: 10.1038/s41598-018-20371-8.

A. Aouacheria, C. Geourjon, N. Aghajari, V. Navratil, G. Deléage, C. Lethias, and J.-Y. Exposito. Insights into early extracellular matrix evolution: spongin short chain collagen-related proteins are homologous to basement membrane type iv collagens and form a novel family widely distributed in invertebrates. *Molecular biology and evolution*, 23(12):2288–2302, 2006. doi: 10.1093/molbev/msl100.

D. Arendt. The evolution of cell types in animals: emerging principles from molecular studies. *Nature reviews. Genetics*, 9(11):868–882, 2008. doi: 10.1038/nrg2416.

D. Arendt, J. M. Musser, C. V. H. Baker, A. Bergman, C. Cepko, D. H. Erwin, M. Pavlicev, G. Schlosser, S. Widder, M. D. Laubichler, and G. P. Wagner. The origin and evolution of cell types. *Nature reviews. Genetics*, 17(12):744–757, 2016. doi: 10.1038/nrg.2016.127.

I. I. Artamonova and A. R. Mushegian. Genome sequence analysis indicates that the model eukaryote nematostella vectensis harbors bacterial consorts. *Applied*

and environmental microbiology, 79(22):6868–6873, 2013. doi: 10.1128/AEM.01635-13.

R. Aufschnaiter, E. A. Zamir, C. D. Little, S. Özbek, S. Münder, C. N. David, L. Li, M. P. Sarras, and X. Zhang. In vivo imaging of basement membrane movement: Ecm patterning shapes hydra polyps. *Journal of cell science*, 124(Pt 23):4027–4038, 2011. doi: 10.1242/jcs.087239.

L. S. Babonis, M. Q. Martindale, and J. F. Ryan. Do novel genes drive morphological novelty? an investigation of the nematosomes in the sea anemone nematostella vectensis. *BMC Evolutionary Biology*, 16(1):114, 2016. doi: 10.1186/s12862-016-0683-3.

L. S. Babonis, J. F. Ryan, C. Enjolras, and M. Q. Martindale. Genomic analysis of the tryptome reveals molecular mechanisms of gland cell evolution. *EvoDevo*, 10:23, 2019. ISSN 2041-9139. doi: 10.1186/s13227-019-0138-1.

G. P. Balasubramanian, A. Beckmann, U. Warnken, M. Schnölzer, A. Schüler, E. Bornberg-Bauer, T. W. Holstein, and S. Özbek. Proteome of hydra nematocyst. *The Journal of biological chemistry*, 287(13):9672–9681, 2012. ISSN 0021-9258. doi: 10.1074/jbc.M111.328203.

L. Bányai and L. Patthy. Putative extremely high rate of proteome innovation in lancelets might be explained by high rate of gene prediction errors. *Scientific reports*, 6:30700, 2016. doi: 10.1038/srep30700.

W. Bargmann. Zur architektur der mesogloea: Untersuchungen an der rippenqualle pleurobrachia pileus. *Zeitschrift fur Zellforschung und Mikroskopische*

Anatomie, 123:66–81, 1972. ISSN 0340-0336. URL https://link.springer.com/article/10.1007/BF00337674.

B. Barzansky and H. M. Lenhoff. On the chemical composition and developmental role of the mesoglea of hydra. *American zoologist*, 14(2):575–581, 1974. ISSN 0003-1569. doi: 10.1093/icb/14.2.575.

B. Barzansky, H. M. Lenhoff, and H. R. Bode. Hydra mesoglea: similarity of its amino acid and neutral sugar composition to that of vertebrate basal lamina. *Comparative biochemistry and physiology. B, Comparative biochemistry*, 50(3): 419–424, 1975. ISSN 1096-4959.

K. Beck, I. Hunter, and J. Engel. Structure and function of laminin: anatomy of a multidomain glycoprotein. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 4(2):148–160, 1990. doi: 10.1096/fasebj.4.2.2404817.

A. Beckmann. *Molecular Factors of Nematocyst Morphogenesis and Discharge in the Freshwater Polyp Hydra*. Phd thesis, Ruperto-Carola University, Heidelberg, Germany, 2013.

A. Beckmann, S. Xiao, J. P. Müller, D. Mercadante, T. Nüchter, N. Kröger, F. Langhojer, W. Petrich, T. W. Holstein, M. Benoit, F. Gräter, and S. Özbek. A fast recoiling silk-like elastomer facilitates nanosecond nematocyst discharge. *BMC biology*, 13:3, 2015. ISSN 101190720. doi: 10.1186/s12915-014-0113-1.

J. Bella and D. J. S. Hulmes. Fibrillar collagens. *Sub-cellular biochemistry*, 82:457–490, 2017. ISSN 0306-0225. doi: 10.1007/978-3-319-49674-0{\textunderscore}14.

T. Bentele. *Der molekulare Aufbau der Nematocystenkapsel bei Hydra als Modell für synthetische Polymere*. PhD thesis, Ruprecht–Karls–Universität, Heidelberg, 2018.

T. Bentele, F. Amadei, E. Kimmle, M. Veschgini, P. Linke, M. Sontag-González, J. Tennigkeit, A. D. Ho, S. Özbek, and M. Tanaka. New class of crosslinker-free nanofiber biomaterials from hydra nematocyst proteins. *Scientific reports*, 9(1): 19116, 2019. doi: 10.1038/s41598-019-55655-0.

A. A. Bentley and J. C. Adams. The evolution of thrombospondins and their ligand-binding activities. *Molecular biology and evolution*, 27(9):2187–2197, 2010. doi: 10.1093/molbev/msq107.

B. G. Bergheim and S. Özbek. Extracellular matrix and morphogenesis in cnidarians: a tightly knit relationship. *Essays In Biochemistry*, 127:1–10, 2019. ISSN 0071-1365. doi: 10.1042/EBC20190021.

S. Berking. Generation of bilateral symmetry in anthozoa: a model. *Journal of theoretical Biology*, 246(3):477–490, 2007. doi: 10.1016/j.jtbi.2007.01.008.

P. Bernardi and P. Bonaldo. Dysfunction of mitochondria and sarcoplasmic reticulum in the pathogenesis of collagen vi muscular dystrophies. *Annals of the New York Academy of Sciences*, 1147:303–311, 2008. doi: 10.1196/annals.1427.009.

C. Bonnans, J. Chou, and Z. Werb. Remodelling the extracellular matrix in development and disease. *Nature reviews. Molecular cell biology*, 15(12):786–801, 2014. doi: 10.1038/nrm3904.

A. Böttger, A. C. Doxey, M. W. Hess, K. Pfaller, W. Salvenmoser, R. Deutzmann, A. Geissner, B. Pauly, J. Altstätter, S. Münder, A. Heim, H.-J. Gabius, B. J. McConkey, and C. N. David. Horizontal gene transfer contributed to the evolution of extracellular surface structures: the freshwater polyp hydra is covered by a complex fibrous cuticle containing glycosaminoglycans and proteins of the ppod and swt (sweet tooth) families. *PloS one*, 7(12):e52278, 2012. doi: 10.1371/journal.pone.0052278.

A. J. Brown and E. J. Sanders. Interactions between mesoderm cells and the extracellular matrix following gastrulation in the chick embryo. *Journal of cell science*, 99(2):431–441, 1991. doi: 10.1242/jcs.99.2.431.

A. Burnett, editor. *Biology of hydra*. Academic Press, New York, 1973. ISBN 9780323148696. URL https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=1167220.

Y.-H. Chang and Y. H. Sun. Carrier of wingless (cow), a secreted heparan sulfate proteoglycan, promotes extracellular transport of wingless. *PloS one*, 9(10): e111573, 2014. doi: 10.1371/journal.pone.0111573.

T. Chari, B. Weissbourd, J. Gehring, A. Ferraioli, L. Leclère, M. Herl, F. Gao, S. Chevalier, R. R. Copley, E. Houliston, D. J. Anderson, and L. Pachter. Whole-animal multiplexed single-cell rna-seq reveals transcriptional shifts across clytia medusa cell types. *Science advances*, 7(48):eabh1683, 2021. doi: 10.1126/sciadv.abh1683.

P. J. A. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, and M. J. L. de Hoon. Biopython: freely available python tools for computational molecular biology and

bioinformatics. *Bioinformatics (Oxford, England)*, 25(11):1422–1423, 2009. doi: 10.1093/bioinformatics/btp163.

Y. Y. Columbus-Shenkar, M. Y. Sachkova, J. Macrander, A. Fridrich, V. Modepalli, A. M. Reitzel, K. Sunagar, and Y. Moran. Dynamics of venom composition across a complex life cycle. *eLife*, 7, 2018. ISSN 2050-084X. doi: 10.7554/eLife.35014.

L. E. Cote, E. Simental, and P. W. Reddien. Muscle functions as a connective tissue and source of extracellular matrix in planarians. *Nature communications*, 10(1):1592, 2019. doi: 10.1038/s41467-019-09539-6.

O. Croce and E. Röttinger. Creating a user-friendly and open-access gene expression database for comparing embryonic development and regeneration in nematostella vectensis. *Methods in molecular biology (Clifton, N.J.)*, 2450:649–662, 2022. doi: 10.1007/978-1-0716-2172-1{\textunderscore}35.

M. Czarny-Ratajczak and A. Latos-Bieleńska. Collagens_the_basic_proteins_of_-the. *Journal of Applied Genetics*, 41(4):317–330, 2000.

A. C. Daley and J. B. Antcliffe. Evolution: The battle of the first animals. *Current biology : CB*, 29(7):R257–R259, 2019. ISSN 0960-9822. doi: 10.1016/j.cub.2019.02.031.

W. P. Daley, S. B. Peters, and M. Larsen. Extracellular matrix dynamics in development and regenerative medicine. *Journal of cell science*, 121(Pt 3):255–264, 2008. doi: 10.1242/jcs.006064.

P. A. Davey, J. L. Clarke, M. Rodrigues, and N. Aldred. Transcriptional characterisation of the exaiptasia pallida pedal disc. *BMC genomics*, 2019. doi: 10.21203/rs.2.9670/v2.

P. A. Davey, A. M. Power, R. Santos, P. Bertemes, P. Ladurner, P. Palmowski, J. Clarke, P. Flammang, B. Lengerer, E. Hennebert, U. Rothbächer, R. Pjeta, J. Wunderer, M. Zurovec, and N. Aldred. Omics-based molecular analyses of adhesion by aquatic invertebrates. *Biological reviews of the Cambridge Philosophical Society*, 96(3):1051–1075, 2021. doi: 10.1111/brv.12691.

C. N. David, S. Özbek, P. Adamczyk, S. Meier, B. Pauly, J. Chapman, J. S. Hwang, T. Gojobori, and T. W. Holstein. Evolution of complex structures: minicollagens shape the cnidarian nematocyst. *Trends in genetics*, 24(9):431–438, 2008. ISSN 0168-9525. doi: 10.1016/j.tig.2008.07.001.

C. G. Davis. The many faces of epidermal growth factor repeats. *The New biologist*, 2(5):410–419, 1990. ISSN 1043-4674.

L. E. Davis and J. F. Haynes. An ultrastructural examination of the mesoglea of hydra. *Zeitschrift fur Zellforschung und Mikroskopische Anatomie*, 92(2):149–158, 1968. ISSN 0340-0336.

M. N. Davis, S. Horne-Badovinac, and A. Naba. In-silico definition of the drosophila melanogaster matrisome. *Matrix Biology Plus*, 4:100015, 2019. doi: 10.1016/j.mbplus.2019.100015.

R. M. Day and H. M. Lenhoff. Hydra mesoglea: a model for investigating epithelial cell–basement membrane interactions. *Science (New York, N.Y.)*, 211(4479): 291–294, 1981. ISSN 0036-8075.

M. S. Detamore, J. G. Orfanos, A. J. Almarza, M. M. French, M. E. Wong, and K. A. Athanasiou. Quantitative analysis and comparative regional investigation of the extracellular matrix of the porcine temporomandibular joint disc. *Matrix biology*

: journal of the International Society for Matrix Biology, 24(1):45–57, 2005. doi: 10.1016/J.MATBIO.2004.11.006.

M. E. Dolega, B. Brunel, M. Le Goff, M. Greda, C. Verdier, J.-F. Joanny, P. Recho, and G. Cappello. Extracellular matrix acts as pressure detector in biological tissues. *BioRxiv*, 2018. doi: 10.1101/488635.

J. L. Drake, J. P. Whitelegge, and D. K. Jacobs. First sequencing of ancient coral skeletal proteins. *Scientific reports*, 10(1):19407, 2020. doi: 10.1038/s41598-020-75846-4.

G. W. Draper, D. K. Shoemark, and J. C. Adams. Modelling the early evolution of extracellular matrix from modern ctenophores and sponges. *Essays in biochemistry*, 63(3):389–405, 2019. doi: 10.1042/EBC20180048.

T. Q. DuBuc, T. B. Stephenson, A. Q. Rock, and M. Q. Martindale. Hox and wnt pattern the primary body axis of an anthozoan cnidarian before gastrulation. *Nature communications*, 9(1):2007, 2018. doi: 10.1038/s41467-018-04184-x.

K. J. Eckelbarger, C. Hand, and K. R. Uhlinger. Ultrastructural features of the trophonema and oogenesis in the starlet sea anemone, nematostella vectensis (edwardsiidae). *Invertebrate Biology*, 127(4):381–395, 2008. ISSN 10778306. doi: 10.1111/j.1744-7410.2008.00146.x.

S. J. Ellis and G. Tanentzapf. Integrin-mediated adhesion and stem-cell-niche interactions. *Cell and Tissue Research*, 339(1):121–130, 2010. ISSN 0302-766X. doi: 10.1007/s00441-009-0828-4.

D. M. Emms and S. Kelly. Orthofinder: phylogenetic orthology inference for

comparative genomics. *Genome biology*, 20(1):238, 2019. doi: 10.1186/s13059-019-1832-y.

J. Engel. Egf-like domains in extracellular matrix proteins: localized signals for growth and differentiation? *FEBS Letters*, 251(1-2):1–7, 1989. ISSN 00145793. doi: 10.1016/0014-5793(89)81417-6.

J. Engel. Domain organizations of modular extracellular matrix proteins and their evolution. *Matrix biology : journal of the International Society for Matrix Biology*, 15(5):295–299, 1996. doi: 10.1016/s0945-053x(96)90130-4.

U. Engel, O. Pertz, C. Fauser, J. Engel, C. N. David, and T. W. Holstein. A switch in disulfide linkage during minicollagen assembly in hydra nematocysts. *The EMBO Journal*, 20(12):3063–3073, 2001. ISSN 0261-4189. doi: 10.1093/emboj/20.12.3063.

U. Engel, S. Özbek, R. Streitwolf-Engel, B. M. Petri, F. Lottspeich, T. W. Holstein, and R. Engel. Nowa, a novel protein with minicollagen cys-rich domains, is involved in nematocyst formation in hydra. *Journal of cell science*, 115(Pt 20): 3923–3934, 2002. doi: 10.1242/jcs.00084.

L. Epp, I. Smid, and P. Tardent. Synthesis of the mesoglea by ectoderm and endoderm in reassembled hydra. *Journal of morphology*, 189(3):271–279, 1986. doi: 10.1002/jmor.1051890306.

J. Y. Exposito, D. Le Guellec, Q. Lu, and R. Garrone. Short chain collagens in sponges are encoded by a family of closely related genes. *The Journal of biological chemistry*, 266(32):21923–21928, 1991. ISSN 0021-9258. doi: 10.1016/S0021-9258(18)54725-6.

J.-Y. Exposito, C. Cluzel, R. Garrone, and C. Lethias. Evolution of collagens. *The Anatomical Record*, 268:302–316, 2002. doi: 10.1002/ar.10162.

B. Fahey and B. M. Degnan. Origin of animal epithelia: insights from the sponge genome. *Evolution & development*, 12(6):601–617, 2010. doi: 10.1111/j.1525-142X.2010.00445.x.

J. Fernández-Pérez and M. Ahearne. The impact of decellularization methods on extracellular matrix derived hydrogels. *Scientific reports*, 9(1):14933, 2019. doi: 10.1038/s41598-019-49575-2.

A. H. L. Fischer, D. Mozzherin, A. M. Eren, K. D. Lans, N. Wilson, C. Cosentino, and J. Smith. Seabase: a multispecies transcriptomic resource and platform for gene network inference. *Integrative and comparative biology*, 54(2):250–263, 2014. doi: 10.1093/icb/icu065.

R. Flaumenhaft and D. B. Rifkin. Extracellular matrix regulation of growth factor and protease activity. *Current opinion in cell biology*, 3(5):817–823, 1991. ISSN 0955-0674. doi: 10.1016/0955-0674{\%}2891{\%}2990055-4.

S. J. Fowler, S. Jose, X. Zhang, R. Deutzmann, M. P. Sarras, and R. P. Boot-Handford. Characterization of hydra type iv collagen. type iv collagen is essential for head regeneration and its expression is up-regulated upon exposure to glucose. *The Journal of biological chemistry*, 275(50):39589–39599, 2000. ISSN 0021-9258. doi: 10.1074/jbc.M005871200.

P. Frank and J. S. Bleakney. Histology and sexual reproduction of the anemone nematostella vectensis stephenson 1935. *Journal of Natural History*, 10(4):441–449, 1976. ISSN 0022-2933. doi: 10.1080/00222937600770331.

H. Franken, T. Mathieson, D. Childs, G. M. A. Sweetman, T. Werner, I. Tögel, C. Doce, S. Gade, M. Bantscheff, G. Drewes, F. B. M. Reinhard, W. Huber, and M. M. Savitski. Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nature protocols*, 10(10):1567–1593, 2015. doi: 10.1038/nprot.2015.101.

J. H. Fritzenwanker, M. Saina, and U. Technau. Analysis of forkhead and snail expression reveals epithelial-mesenchymal transitions during embryonic and larval development of nematostella vectensis. *Developmental biology*, 275(2):389–402, 2004. ISSN 0012-1606. doi: 10.1016/j.ydbio.2004.08.014.

N. Garg, U. K. Štibler, B. Eismann, M. Mercker, B. G. Bergheim, A. Linn, P. Tuchscherer, U. Engel, S. Redl, A. Marciniak-Czochra, T. W. Holstein, M. W. Hess, and S. Özbek. Non-muscle myosin ii drives critical steps of nematocyst morphogenesis. *iScience*, 26(3):106291, 2023. doi: 10.1016/j.isci.2023.106291.

R. Garrone. Collagen, a common thread in extracellular matrix evolution. *Proceedings of the Indian Academy of Science*, 1:51–66, 1999.

R. Garrone and R. Trelstad. *Formation and involvement of extracellular matrix in the development of sponges, a primitive multicellular system: The role of extracellular matrix in development*. Liss, New York, 1984.

G. Giannelli, J. Falk-Marzillier, O. Schiraldi, W. G. Stetler-Stevenson, and V. Quaranta. Induction of cell migration by matrix metalloprotease-2 cleavage of laminin-5. *Science (New York, N.Y.)*, 277(5323):225–228, 1997. ISSN 0036-8075. doi: 10.1126/science.277.5323.225.

R. Giavazzi, A. Garofalo, C. Ferri, V. Lucchini, E. A. Bone, S. Chiari, P. D. Brown,

M. I. Nicoletti, and G. Taraboletti. Batimastat, a synthetic inhibitor of matrix metalloproteinases, potentiates the antitumor activity of cisplatin in ovarian carcinoma xenografts. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 4(4):985–992, 1998. ISSN 1078-0432.

D. Gilgés, M.-A. VINIT, I. CALLEBAUT, L. COULOMBEL, V. Cacgeux, P.-H. ROMEO, and I. VIGON. Polydom: a secreted protein with pentraxin, complement control protein, epidermal growth factor and von willebrand factor a domains. *Biochemical Journal*, 352(1):49–59, 2000. ISSN 0264-6021. doi: 10.1042/bj3520049.

W. B. Gladfelter. Structure and function of the locomotory system ofpolyorchis montereyensis (cnidaria, hydrozoa). *Helgoländer Wissenschaftliche Meeresuntersuchungen*, 23(1):38–79, 1972. ISSN 0017-9957. doi: 10.1007/BF01616310.

W. G. Gladfelter. A comparative analysis of the locomotory systems of medusoid cnidaria. *Helgoländer Wissenschaftliche Meeresuntersuchungen*, 25(2-3):228–272, 1973. ISSN 0017-9957. doi: 10.1007/BF01611199.

S.-K. Goh, W. Halfter, T. Richardson, S. Bertera, V. Vaidya, J. Candiello, M. Bradford, and I. Banerjee. Organ-specific ecm arrays for investigating cell-ecm interactions during stem cell differentiation. *Biofabrication*, 13(1), 2020. doi: 10.1088/1758-5090/abc05f.

S.-K. Goh, W. Halfter, T. Richardson, S. Bertera, V. Vaidya, J. Candiello, M. Bradford, and I. Banerjee. Organ-specific ecm arrays for investigating cell-ecm interactions during stem cell differentiation. *Biofabrication*, 13(015015), 2021. doi: 10.1088/1758-5090/abc05f.

D. A. Gold and D. K. Jacobs. Stem cell dynamics in cnidaria: are there unifying principles? *Development genes and evolution*, 223(1-2):53–66, 2013. ISSN 0949-944X. doi: 10.1007/s00427-012-0429-1.

D. A. Gold, T. Katsuki, Y. Li, X. Yan, M. Regulski, D. Ibberson, T. Holstein, R. E. Steele, D. K. Jacobs, and R. J. Greenspan. The genome of the jellyfish aurelia and the evolution of animal complexity. *Nature ecology & evolution*, 3(1):96–104, 2019. doi: 10.1038/s41559-018-0719-8.

L. C. Grasso, J. Maindonald, S. Rudd, D. C. Hayward, R. Saint, D. J. Miller, and E. E. Ball. Microarray analysis identifies candidate genes for key roles in coral development. *BMC genomics*, 9:1–18, 2008. doi: 10.1186/1471-2164-9-540.

C. Guder, S. Pinho, T. G. Nacak, H. A. Schmidt, B. Hobmayer, C. Niehrs, and T. W. Holstein. An ancient wnt-dickkopf antagonism in hydra. *Development*, 133(5): 901–911, 2006. ISSN 0950-1991. doi: 10.1242/dev.02265.

q. Guo. A myxozoan genome reveals mosaic evolution in a parasitic cnidarian. *BMC Biology*, 20(51), 2021. doi: 10.7910/DVN/INLEPM.

G. Gupta and A. Surolia. Collectins: sentinels of innate immunity. *BioEssays*, 29 (5):452–464, 2007. ISSN 0265-9247. doi: 10.1002/bies.20573. URL https://pubmed.ncbi.nlm.nih.gov/17450595/.

D. H. Haft, M. DiCuccio, A. Badretdin, V. Brover, V. Chetvernin, K. O'Neill, W. Li, F. Chitsaz, M. K. Derbyshire, N. R. Gonzales, M. Gwadz, F. Lu, G. H. Marchler, J. S. Song, N. Thanki, R. A. Yamashita, C. Zheng, F. Thibaud-Nissen, L. Y. Geer, A. Marchler-Bauer, and K. D. Pruitt. Refseq: an update on prokaryotic genome

annotation and curation. *Nucleic Acids Research*, 46(D1):D851–D860, 2018. ISSN 0305-1048. doi: 10.1093/nar/gkx1068.

W. Halfter, S. Dong, B. Schurer, and G. J. Cole. Collagen xviii is a basement membrane heparan sulfate proteoglycan. *The Journal of biological chemistry*, 273(39):25404–25412, 1998. ISSN 0021-9258. doi: 10.1074/jbc.273.39.25404.

M. Hamada, N. Satoh, and K. Khalturin. A reference genome from the symbiotic hydrozoan, hydra viridissima. *G3 (Bethesda, Md.)*, 10(11):3883–3895, 2020. doi: 10.1534/g3.120.401411.

C. Hand and K. R. Uhlinger. The culture, sexual and asexual reproduction, and growth of the sea anemone nematostella vectensis. *The Biological bulletin*, 182 (2):169–176, 1992. ISSN 0006-3185. doi: 10.2307/1542110.

S. Hellstern, J. Stetefeld, C. Fauser, A. Lustig, J. Engel, T. W. Holstein, and S. Ozbek. Structure/function analysis of spinalin, a spine protein of hydra nematocysts. *The FEBS journal*, 273(14):3230–3237, 2006. doi: 10.1111/j. 1742-4658.2006.05331.x.

E. M. Hill, C.-Y. Chen, F. Del Viso, L. R. Ellington, S. He, A. Karabulut, A. Paulson, and M. C. Gibson. Manipulation of gene activity in the regenerative model sea anemone, nematostella vectensis. *Methods in molecular biology (Clifton, N.J.)*, 2450:437–465, 2022. doi: 10.1007/978-1-0716-2172-1{\textunderscore}23.

E. Hohenester and J. Engel. Domain structure and organisation in extracellular matrix proteins. *Matrix biology : journal of the International Society for Matrix Biology*, 21(2):115–128, 2002. doi: 10.1016/s0945-053x(01)00191-3.

T. W. Holstein, M. Benoit, G. V. Herder, C. N. David, G. Wanner, and H. E. Gaub. Fibrous mini-collagens in hydra nematocysts. *Science (New York, N.Y.)*, 265 (5170):402–404, 1994. ISSN 0036-8075. doi: 10.1126/science.265.5170.402.

T. W. Holstein, M. W. Hess, and W. Salvenmoser. Preparation techniques for transmission electron microscopy of hydra. *Methods in cell biology*, 96:285–306, 2010. ISSN 0091-679X. doi: 10.1016/S0091-679X(10)96013-5.

M. Hu, X. Zheng, C.-M. Fan, and Y. Zheng. Lineage dynamics of the endosymbiotic cell type in the soft coral xenia. *Nature*, 582(7813):534–538, 2020. doi: 10.1038/s41586-020-2385-7.

C. S. Hughes, S. Foehr, D. A. Garfield, E. E. Furlong, L. M. Steinmetz, and J. Krijgsveld. Ultrasensitive proteome analysis using paramagnetic bead technology. *Molecular systems biology*, 10(10):757, 2014. doi: 10.15252/msb.20145625.

D. J. Huss, S. Saias, S. Hamamah, J. M. Singh, J. Wang, M. Dave, J. Kim, J. Eberwine, and R. Lansford. Avian primordial germ cells contribute to and interact with the extracellular matrix during early migration. *Frontiers in cell and developmental biology*, 7:35, 2019. ISSN 2296-634X. doi: 10.3389/fcell.2019.00035.

J. Huxley-Jones, J. W. Pinney, J. Archer, D. L. Robertson, and R. P. Boot-Handford. Back to basics–how the evolution of the extracellular matrix underpinned vertebrate evolution. *International journal of experimental pathology*, 90 (2):95–100, 2009. doi: 10.1111/j.1365-2613.2008.00637.x.

R. O. Hynes. The extracellular matrix: not just pretty fibrils. *Science (New York, N.Y.)*, 326(5957):1216–1219, 2009. ISSN 0036-8075. doi: 10.1126/science.1176009.

R. O. Hynes. The evolution of metazoan extracellular matrix. *The Journal of Cell Biology*, 196(6):671–679, 2012. ISSN 0021-9525. doi: 10.1083/jcb.201109041.

R. O. Hynes and A. Naba. Overview of the matrisome–an inventory of extracellular matrix constituents and functions. *Cold Spring Harbor Perspectives in Biology*, 4(1):a004903, 2012. doi: 10.1101/cshperspect.a004903.

A. Ikmi, P. J. Steenbergen, M. Anzo, M. R. McMullen, A. Stokkermans, L. R. Ellington, and M. C. Gibson. Feeding-dependent tentacle development in the sea anemone nematostella vectensis. *Nature communications*, 11(1):4399, 2020. doi: 10.1038/s41467-020-18133-0.

D. E. Ingber. Mechanical control of tissue morphogenesis during embryological development. *The International journal of developmental biology*, 50(2-3):255–266, 2006. doi: 10.1387/ijdb.052044di.

D. Inoue and J. Wittbrodt. One for all–a highly efficient and versatile method for fluorescent immunostaining in fish embryos. *PloS one*, 6(5):e19713, 2011. doi: 10.1371/journal.pone.0019713.

M. Jager and M. Manuel. Ctenophores: an evolutionary-developmental perspective. *Current opinion in genetics & development*, 39:85–92, 2016. doi: 10.1016/j.gde.2016.05.020.

G. Jékely, J. Paps, and C. Nielsen. The phylogenetic position of ctenophores and the origin(s) of nervous systems. *EvoDevo*, 6:1, 2015. ISSN 2041-9139. doi: 10.1186/2041-9139-6-1.

P. Jones, D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas,

M. Scheremetjew, S.-Y. Yong, R. Lopez, and S. Hunter. Interproscan 5: genome-scale protein function classification. *Bioinformatics (Oxford, England)*, 30(9): 1236–1240, 2014. doi: 10.1093/bioinformatics/btu031.

M. Jouiaei, A. A. Yanagihara, B. Madio, T. J. Nevalainen, P. F. Alewood, and B. G. Fry. Ancient venom systems: A review on cnidaria toxins. *Toxins*, 7(6):2251–2271, 2015. doi: 10.3390/toxins7062251.

K. E. Kadler, C. Baldock, J. Bella, and R. P. Boot-Handford. Collagens at a glance. *Journal of cell science*, 120(Pt 12):1955–1958, 2007. doi: 10.1242/jcs.03453.

S. Kakugawa, P. F. Langton, M. Zebisch, S. Howell, T.-H. Chang, Y. Liu, Feizi, G. Bineva, N. O'Reilly, A. P. Snijders, E. Y. Jones, and J.-P. Vincent. Notum deacylates wnt proteins to suppress signalling activity. *Nature*, 519(7542):187–192, 2015. doi: 10.1038/nature14259.

S. Kang, A. K. Tice, C. W. Stairs, R. E. Jones, D. J. G. Lahr, and M. W. Brown. The integrin-mediated adhesive complex in the ancestor of animals, fungi, and amoebae. *Current biology : CB*, 31(14):3073–3085.e3, 2021. ISSN 0960-9822. doi: 10.1016/j.cub.2021.04.076.

T. Karpanen, Y. Padberg, S. A. van de Pavert, C. Dierkes, N. Morooka, J. Peterson-Maduro, G. van de Hoek, M. Adrian, N. Mochizuki, K. Sekiguchi, F. Kiefer, D. Schulte, and S. Schulte-Merker. An evolutionarily conserved role for polydom/svep1 during lymphatic vessel formation. *Circulation research*, 120(8): 1263–1275, 2017. doi: 10.1161/CIRCRESAHA.116.308813.

E. Kayal, B. Bentlage, M. Sabrina Pankey, A. H. Ohdera, M. Medina, D. C. Plachetzki, A. G. Collins, and J. F. Ryan. Phylogenomics provides a robust topology of

the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evolutionary Biology*, 18(1):139, 2018. doi: 10.1186/s12862-018-1142-0.

C. D. Kenkel, E. Meyer, and M. V. Matz. Gene expression under chronic heat stress in populations of the mustard hill coral (porites astreoides) from different thermal environments. *Molecular ecology*, 22(16):4322–4334, 2013. doi: 10.1111/mec.12390.

N. J. Kenny, W. R. Francis, R. E. Rivera-Vicéns, K. Juravel, A. de Mendoza, C. Díez-Vives, R. Lister, L. A. Bezares-Calderón, L. Grombacher, M. Roller, L. D. Barlow, S. Camilli, J. F. Ryan, G. Wörheide, A. L. Hill, A. Riesgo, and S. P. Leys. Tracing animal genomic evolution with the chromosomal-level assembly of the freshwater sponge ephydatia muelleri. *Nature communications*, 11(1):3676, 2020. doi: 10.1038/s41467-020-17397-w.

K. Khalturin, C. Shinzato, M. Khalturina, M. Hamada, M. Fujie, R. Koyanagi, M. Kanda, H. Goto, F. Anton-Erxleben, M. Toyokawa, S. Toshino, and N. Satoh. Medusozoan genomes inform the evolution of the jellyfish body plan. *Nature ecology & evolution*, 3(5):811–822, 2019. doi: 10.1038/s41559-019-0853-y.

N. King, C. T. Hittinger, and S. B. Carroll. Evolution of key cell signaling and adhesion protein families predates animal origins. *Science (New York, N.Y.)*, 301 (5631):361–363, 2003. ISSN 0036-8075. doi: 10.1126/SCIENCE.1083853.

B. A. Knack. *Cell adhesion factors in cnidarians*. Phd thesis, James Cook University, 2011. URL https://researchonline.jcu.edu.au/36991/.

C. Knupp and J. M. Squire. Molecular packing in network-forming collagens.

Advances in protein chemistry, 70:375–403, 2005. ISSN 0065-3233. doi: 10.1016/S0065-3233(05)70011-5.

G. Kumpfmüller, V. Rybakine, T. Takahashi, T. Fujisawa, and T. C. Bosch. Identification of an astacin matrix metalloprotease as target gene for hydra foot activator peptides. *Development genes and evolution*, 209(10):601–607, 1999. ISSN 0949-944X. doi: 10.1007/s004270050294.

E. M. Kurz, T. W. Holstein, B. M. Petri, J. Engel, and C. N. David. Mini-collagens in hydra nematocytes. *The Journal of Cell Biology*, 115(4):1159–1169, 1991. ISSN 0021-9525.

J. Kyslík, A. Kosakyan, S. Nenarokov, A. S. Holzer, and I. Fiala. The myxozoan minicollagen gene repertoire was not simplified by the parasitic lifestyle: computational identification of a novel myxozoan minicollagen gene. *BMC Genomics*, 22(1):198, 2021. doi: 10.1186/s12864-021-07515-3.

M. J. Layden, F. Rentzsch, and E. Röttinger. 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, 2016. ISSN 1759-7684. doi: 10.1002/wdev.222.

L. Leclère and F. Rentzsch. Rgm regulates bmp-mediated secondary axis formation in the sea anemone nematostella vectensis. *Cell reports*, 9(5):1921–1930, 2014. doi: 10.1016/j.celrep.2014.11.009.

E. E. L. Lee, E. Watto, and J. Malamy. Characterizing epithelial wound healing in vivo using the cnidarian model organism clytia hemisphaerica. *Journal of visualized experiments : JoVE*, 192, 2023. doi: 10.3791/65081.

G. Lee, S.-B. Han, and D.-H. Kim. Cell-ecm contact-guided intracellular polarization is mediated via lamin a/c dependent nucleus-cytoskeletal connection. *Biomaterials*, 268:120548, 2021. doi: 10.1016/j.biomaterials.2020.120548.

D. J. Leeming, A. C. Bay-Jensen, E. Vassiliadis, M. R. Larsen, K. Henriksen, and M. A. Karsdal. Post-translational modifications of the extracellular matrix are key events in cancer progression: opportunities for biochemical marker development. *Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals*, 16(3):193–205, 2011. doi: 10.3109/1354750X.2011.557440.

H. M. Lenhoff and W. F. Loomis. *The biology of hydra and of some other coelenterates*. University of Miami Press, 1961.

A. A. Leontovich, J. Zhang, K. Shimokawa, H. Nagase, and M. P. Sarras. A novel hydra matrix metalloproteinase (hmmp) functions in extracellular matrix degradation, morphogenesis and the maintenance of differentiated cells in the foot process. *Development*, 127(4):907–920, 2000. ISSN 0950-1991.

S. Levy, A. Elek, X. Grau-Bové, S. Menéndez-Bravo, M. Iglesias, A. Tanay, T. Mass, and A. Sebé-Pedrós. A stony coral cell atlas illuminates the molecular and cellular basis of coral symbiosis, calcification, and immunity. *Cell*, 2021. ISSN 0092-8674. doi: 10.1016/j.cell.2021.04.005.

S. P. Leys and A. Riesgo. Epithelia, an evolutionary novelty of metazoans. *Journal of experimental zoology. Part B, Molecular and developmental evolution*, 318(6): 438–447, 2012. doi: 10.1002/jez.b.21442.

S. P. Leys, S. A. Nichols, and E. D. M. Adams. Epithelia and integration in sponges.

Integrative and comparative biology, 49(2):167–177, 2009. doi: 10.1093/icb/icp038.

W. Li, K. R. O'Neill, D. H. Haft, M. DiCuccio, V. Chetvernin, A. Badretdin, G. Coulouris, F. Chitsaz, M. K. Derbyshire, A. S. Durkin, N. R. Gonzales, M. Gwadz, C. J. Lanczycki, J. S. Song, N. Thanki, J. Wang, R. A. Yamashita, M. Yang, C. Zheng, A. Marchler-Bauer, and F. Thibaud-Nissen. Refseq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Research*, 49(D1):D1020–D1028, 2021. ISSN 0305-1048. doi: 10.1093/nar/gkaa1105.

S. Libro, S. T. Kaluziak, and S. V. Vollmer. Rna-seq profiles of immune related genes in the staghorn coral acropora cervicornis infected with white band disease. *PloS one*, 8(11):e81821, 2013. doi: 10.1371/journal.pone.0081821.

X. Lin, S. Patil, Y.-G. Gao, and A. Qian. The bone extracellular matrix in bone formation and regeneration. *Frontiers in pharmacology*, 11:757, 2020. ISSN 1663-9812. doi: 10.3389/fphar.2020.00757.

A. Listrat, C. Boby, J. Tournayre, and C. Jousse. Bovine extracellular matrix proteins and potential role in meat quality: First in silico bos taurus compendium. *Journal of proteomics*, 279:104891, 2023. doi: 10.1016/j.jprot.2023.104891.

C. Liu. Comparative proteomics for an in-depth understanding of bioadhesion mechanisms and evolution across metazoans. *Journal of proteomics*, 256: 104506, 2022. doi: 10.1016/j.jprot.2022.104506.

M. Lommel, J. Strompen, A. L. Hellewell, G. P. Balasubramanian, E. D. Christofidou, A. R. Thomson, A. L. Boyle, D. N. Woolfson, K. Puglisi, M. Hartl, T. W.

Holstein, J. C. Adams, and S. Özbek. Hydra mesoglea proteome identifies thrombospondin as a conserved component active in head organizer restriction. *Scientific reports*, 8(1):11753, 2018. doi: 10.1038/s41598-018-30035-2.

P. Lu, K. Takai, V. M. Weaver, and Z. Werb. Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harbor Perspectives in Biology*, 3(12), 2011. doi: 10.1101/cshperspect.a005058.

P. Lu, V. M. Weaver, and Z. Werb. The extracellular matrix: a dynamic niche in cancer progression. *The Journal of Cell Biology*, 196(4):395–406, 2012. ISSN 0021-9525. doi: 10.1083/jcb.201102147.

A. Lundell, A. I. Olin, M. Mörgelin, S. al Karadaghi, A. Aspberg, and D. T. Logan. Structural basis for interactions between tenascins and lectican c-type lectin domains: evidence for a crosslinking role for tenascins. *Structure (London, England : 1993)*, 12(8):1495–1506, 2004. ISSN 0969-2126. doi: 10.1016/j.str.2004.05.021.

B. Marchini, L. de Nuccio, M. Mazzei, and G. L. Mariottini. A fast centrifuge method for nematocyst isolation from pelagia noctiluca forskal (cnidaria: Scyphozoa). *Rivista di biologia*, 97(3):505–515, 2004. ISSN 0035-6050.

M. Q. Martindale, K. Pang, and J. R. Finnerty. 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, 2004. ISSN 0950-1991. doi: 10.1242/dev.01119.

M. C. McCabe, L. R. Schmitt, R. C. Hill, M. Dzieciatkowska, M. Maslanka, W. F. Daamen, T. H. van Kuppevelt, D. J. Hof, and K. C. Hansen. Evaluation and

refinement of sample preparation methods for extracellular matrix proteome coverage. *Molecular & cellular proteomics : MCP*, 20:100079, 2021. doi: 10.1016/j.mcpro.2021.100079.

M. C. McCabe, A. J. Saviola, and K. C. Hansen. A mass spectrometry-based atlas of extracellular matrix proteins across 25 mouse organs. *BioRxive*, 2022. doi: 10.1101/2022.03.04.482898.

J. McCaughey, N. L. Stevenson, S. Cross, and D. J. Stephens. Er-to-golgi trafficking of procollagen in the absence of large carriers. *The Journal of Cell Biology*, 218(3):929–948, 2019. ISSN 0021-9525. doi: 10.1083/jcb.201806035.

C. S. McFadden, A. M. Quattrini, M. R. Brugler, P. F. Cowman, L. F. Dueñas, M. V. Kitahara, D. A. Paz-García, J. D. Reimer, and E. Rodríguez. Phylogenomics, origin and diversification of anthozoans (phylum cnidaria). *Systematic biology*, 2021. doi: 10.1093/sysbio/syaa103.

A. D. McInnes, M. A. J. Moser, and X. Chen. Preparation and use of decellularized extracellular matrix for tissue engineering. *Journal of functional biomaterials*, 13 (4), 2022. ISSN 2079-4983. doi: 10.3390/jfb13040240.

S. Meier, D. Häussinger, E. Pokidysheva, H. P. Bächinger, and S. Grzesiek. Determination of a high-precision nmr structure of the minicollagen cysteine rich domain from hydra and characterization of its disulfide bond formation. *FEBS Letters*, 569(1-3):112–116, 2004. ISSN 00145793. doi: 10.1016/j.febslet.2004.05.034.

S. Meier, P. R. Jensen, C. N. David, J. Chapman, T. W. Holstein, S. Grzesiek, and S. Özbek. Continuous molecular evolution of protein-domain structures by

single amino acid changes. *Current biology : CB*, 17(2):173–178, 2007. ISSN 0960-9822. doi: 10.1016/j.cub.2006.10.063.

A. Meszaros and C. Bigger. Qualitative and quantitative study of wound healing processes in the coelenterate, plexaurella fusifera: spatial, temporal, and environmental (light attenuation) influences. *Journal of invertebrate pathology*, 73 (3):321–331, 1999. ISSN 0022-2011. doi: 10.1006/jipa.1999.4851.

S. Michelini, B. Amato, M. Ricci, R. Serrani, D. Veselenyiova, S. Kenanoglu, D. Kurti, A. Dautaj, M. Baglivo, R. Compagna, J. Krajcovic, M. Dundar, S. H. Basha, S. Priya, J. P. Belgrado, and M. Bertelli. Svep1 is important for morphogenesis of lymphatic system: Possible implications in lymphedema. *Lymphology*, 54(1), 2021. doi: 10.2458/lymph.4678.

A. G. Milbradt, C. Boulegue, L. Moroder, and C. Renner. The two cysteine-rich head domains of minicollagen from hydra nematocysts differ in their cystine framework and overall fold despite an identical cysteine sequence pattern. *Journal of molecular biology*, 354(3):591–600, 2005. doi: 10.1016/j.jmb.2005.09.080.

E. Moiseeva, C. Rabinowitz, G. Paz, and B. Rinkevich. Histological study on maturation, fertilization and the state of gonadal region following spawning in the model sea anemone, nematostella vectensis. *PloS one*, 12(8):e0182677, 2017. doi: 10.1371/journal.pone.0182677.

Y. Moran, D. Praher, A. Schlesinger, A. Ayalon, Y. Tal, and U. Technau. Analysis of soluble protein contents from the nematocysts of a model sea anemone sheds light on venom evolution. *Marine biotechnology (New York, N.Y.)*, 15(3):329–339, 2013. doi: 10.1007/s10126-012-9491-y.

N. Morooka, S. Futaki, R. Sato-Nishiuchi, M. Nishino, Y. Totani, C. Shimono, I. Nakano, H. Nakajima, N. Mochizuki, and K. Sekiguchi. Polydom is an extracellular matrix protein involved in lymphatic vessel remodeling. *Circulation research*, 120(8):1276–1288, 2017. doi: 10.1161/CIRCRESAHA.116.308825.

L. L. Moroz, K. M. Kocot, M. R. Citarella, S. Dosung, T. P. Norekian, I. S. Povolotskaya, A. P. Grigorenko, C. Dailey, E. Berezikov, K. M. Buckley, A. Ptitsyn, D. Reshetov, K. Mukherjee, T. P. Moroz, Y. Bobkova, F. Yu, V. V. Kapitonov, J. Jurka, Y. V. Bobkov, J. J. Swore, D. O. Girardo, A. Fodor, F. Gusev, R. Sanford, R. Bruders, E. Kittler, C. E. Mills, J. P. Rast, R. Derelle, V. V. Solovyev, F. A. Kondrashov, B. J. Swalla, J. V. Sweedler, E. I. Rogaev, K. M. Halanych, and A. B. Kohn. The ctenophore genome and the evolutionary origins of neural systems. *Nature*, 510(7503):109–114, 2014. doi: 10.1038/nature13400.

A. Moya, L. Huisman, E. E. Ball, D. C. Hayward, L. C. Grasso, C. M. Chua, H. N. Woo, J.-P. Gattuso, S. Forêt, and D. J. Miller. Whole transcriptome analysis of the coral acropora millepora reveals complex responses to co_2-driven acidification during the initiation of calcification. *Molecular ecology*, 21(10):2440–2454, 2012. doi: 10.1111/j.1365-294X.2012.05554.x.

Y. Muragaki, S. Timmons, C. M. Griffith, S. P. Oh, B. Fadel, T. Quertermous, and B. R. Olsen. Mouse col18a1 is expressed in a tissue-specific manner as three alternative variants and is localized in basement membrane zones. *Proceedings of the National Academy of Sciences of the United States of America*, 92(19): 8763–8767, 1995. doi: 10.1073/pnas.92.19.8763.

L. D. Mydlarz, S. F. Holthouse, E. C. Peters, and C. D. Harvell. Cellular responses

in sea fan corals: granular amoebocytes react to pathogen and climate stressors. *PloS one*, 3(3):e1811, 2008. doi: 10.1371/journal.pone.0001811.

A. Naba, K. R. Clauser, S. Hoersch, H. Liu, S. A. Carr, and R. O. Hynes. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Molecular & cellular proteomics : MCP*, 11(4): M111.014647, 2012. doi: 10.1074/mcp.M111.014647.

A. R. Nabhan and I. N. Sarkar. The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. *Briefings in bioinformatics*, 13(1):122–134, 2012. doi: 10.1093/bib/bbr014.

P. Nauroy, S. Hughes, A. Naba, and F. Ruggiero. The in-silico zebrafish matrisome: A new tool to study extracellular matrix gene and protein functions. *Matrix biology : journal of the International Society for Matrix Biology*, 65:5–13, 2018. doi: 10.1016/j.matbio.2017.07.001.

K. M. Noh, S.-J. Park, S.-H. Moon, and S. Y. Jung. Extracellular matrix cues regulate the differentiation of pluripotent stem cell-derived endothelial cells. *Frontiers in cardiovascular medicine*, 10:1169331, 2023. ISSN 2297-055X. doi: 10.3389/fcvm.2023.1169331.

M. Norman, R. Vuilleumier, A. Springhorn, J. Gawlik, and G. Pyrowolakis. Pentagone internalises glypicans to fine-tune multiple signalling pathways. *eLife*, 5, 2016. ISSN 2050-084X. doi: 10.7554/eLife.13301.

C. A. Oakley, M. F. Ameismeier, L. Peng, V. M. Weis, A. R. Grossman, and S. K. Davy. Symbiosis induces widespread changes in the proteome of the model

cnidarian aiptasia. *Cellular microbiology*, 18(7):1009–1023, 2016. doi: 10.1111/cmi.12564.

A. Ohdera, C. L. Ames, R. B. Dikow, E. Kayal, M. Chiodin, B. Busby, S. La, S. Pirro, A. G. Collins, M. Medina, and J. F. Ryan. Box, stalked, and upside-down? draft genomes from diverse jellyfish (cnidaria, acraspeda) lineages: Alatina alata (cubozoa), calvadosia cruxmelitensis (staurozoa), and cassiopea xamachana (scyphozoa). *GigaScience*, 8(7), 2019. doi: 10.1093/gigascience/giz069.

A. H. Ohdera, M. J. Abrams, C. L. Ames, D. M. Baker, L. P. Suescún-Bolívar, A. G. Collins, C. J. Freeman, E. Gamero-Mora, T. L. Goulet, D. K. Hofmann, A. Jaimes-Becerra, P. F. Long, A. C. Marques, L. A. Miller, L. D. Mydlarz, A. C. Morandini, C. R. Newkirk, S. P. Putri, J. E. Samson, S. N. Stampar, B. Steinworth, M. Templeman, P. E. Thomé, M. Vlok, C. M. Woodley, J. C. Wong, M. Q. Martindale, W. K. Fitt, and M. Medina. Upside-down but headed in the right direction: Review of the highly versatile cassiopea xamachana system. *Frontiers in Ecology and Evolution*, 6, 2018. doi: 10.3389/fevo.2018.00035.

S. Özbek, G. P. Balasubramanian, R. Chiquet-Ehrismann, R. P. Tucker, and J. C. Adams. The evolution of extracellular matrix. *Molecular biology of the cell*, 21 (24):4300–4305, 2010. doi: 10.1091/mbc.E10-03-0251.

C. V. Palmer, N. G. Traylor-Knowles, B. L. Willis, and J. C. Bythell. Corals use similar immune cells and wound-healing processes as those of higher organisms. *PloS one*, 6(8):e23992, 2011. doi: 10.1371/journal.pone.0023992.

M. G. Parisi, D. Parrinello, L. Stabili, and M. Cammarata. Cnidarian immunity and the repertoire of defense mechanisms in anthozoans. *Biology*, 9(9), 2020. ISSN 2079-7737. doi: 10.3390/biology9090283.

M. G. Parisi, A. Grimaldi, N. Baranzini, C. La Corte, M. Dara, D. Parrinello, and M. Cammarata. Mesoglea extracellular matrix reorganization during regenerative process in anemonia viridis (forskål, 1775). *International journal of molecular sciences*, 22(11), 2021. doi: 10.3390/ijms22115971.

L. Patthy. Genome evolution and the evolution of exon-shuffling–a review. *Gene*, 238(1):103–114, 1999. ISSN 0378-1119. doi: 10.1016/s0378-1119(99)00228-0.

S. Pennings, K. J. Liu, and H. Qian. The stem cell niche: Interactions between stem cells and their environment. *Stem cells international*, 2018:4879379, 2018. ISSN 1687-966X. doi: 10.1155/2018/4879379.

L. Pesciaroli, M. Petruccioli, S. Fedi, A. Firrincieli, F. Federici, and A. D'Annibale. Characterization of pleurotus ostreatus biofilms by using the calgary biofilm device. *Applied and environmental microbiology*, 79(19):6083–6092, 2013. doi: 10.1128/AEM.02099-13.

H. O. Petersen, S. K. Höger, M. Looso, T. Lengfeld, A. Kuhn, U. Warnken, C. Nishimiya-Fujisawa, M. Schnölzer, M. Krüger, S. Özbek, O. Simakov, and T. W. Holstein. A comprehensive transcriptomic and proteomic analysis of hydra head regeneration. *Molecular biology and evolution*, 32(8):1928–1947, 2015. doi: 10.1093/molbev/msv079.

P. B. Petrov, J. M. Considine, V. Izzi, and A. Naba. Matrisome analyzer: A suite of tools to annotate and quantify ecm molecules in big datasets across organisms. *bioRxiv : the preprint server for biology*, 2023. doi: 10.1101/2023.04.18.537378.

S. Piccolo, E. Agius, B. Lu, S. Goodman, L. Dale, and E. M. de Robertis. Cleavage of chordin by xolloid metalloprotease suggests a role for proteolytic processing

in the regulation of spemann organizer activity. *Cell*, 91(3):407–416, 1997. ISSN 0092-8674. doi: 10.1016/s0092-8674(00)80424-9.

E. Pokidysheva, A. G. Milbradt, S. Meier, C. Renner, D. Häussinger, H. P. Bächinger, L. Moroder, S. Grzesiek, T. W. Holstein, S. Özbek, and J. Engel. The structure of the cys-rich terminal domain of hydra minicollagen, which is involved in disulfide networks of the nematocyst wall. *The Journal of biological chemistry*, 279(29):30395–30401, 2004. ISSN 0021-9258. doi: 10.1074/jbc.M403734200.

N. H. Putnam, M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, A. Salamov, A. Terry, H. Shapiro, E. Lindquist, V. V. Kapitonov, J. Jurka, G. Genikhovich, I. V. Grigoriev, S. M. Lucas, R. E. Steele, J. R. Finnerty, U. Technau, M. Q. Martindale, and D. S. Rokhsar. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science (New York, N.Y.)*, 317(5834):86–94, 2007. ISSN 0036-8075. doi: 10.1126/science.1139158.

P. Pyagay, M. Heroult, Q. Wang, W. Lehnert, J. Belden, L. Liaw, R. E. Friesel, and V. Lindner. Collagen triple helix repeat containing 1, a novel secreted protein in injured and diseased arteries, inhibits collagen expression and promotes cell migration. *Circulation research*, 96(2):261–268, 2005. doi: 10.1161/01.RES.0000154262.07264.12.

S. Rigogliuso, M. Salamone, E. Barbarino, M. Barbarino, A. Nicosia, and G. Ghersi. Production of injectable marine collagen-based hydrogel for the maintenance of differentiated chondrocytes in tissue engineering applications. *International journal of molecular sciences*, 21(16), 2020. doi: 10.3390/ijms21165798.

M. Rodrigues, T. Ostermann, L. Kremeser, H. Lindner, C. Beisel, E. Berezikov, B. Hobmayer, and P. Ladurner. Profiling of adhesive-related genes in the fresh-

water cnidarian hydra magnipapillata by transcriptomics and proteomics. *Biofouling*, 32(9):1115–1129, 2016. doi: 10.1080/08927014.2016.1233325.

M. D. Roycik, X. Fang, and Q.-X. Sang. A fresh prospect of extracellular matrix hydrolytic enzymes and their substrates. *Current pharmaceutical design*, 15(12): 1295–1308, 2009. doi: 10.2174/138161209787846676.

I. N. C. Ruggieri, A. M. Cícero, J. P. M. Issa, and S. Feldman. Bone fracture healing: perspectives according to molecular basis. *Journal of bone and mineral metabolism*, 39(3):311–331, 2021. doi: 10.1007/s00774-020-01168-0.

M. Saina, G. Genikhovich, E. Renfer, and U. Technau. Bmps and chordin regulate patterning of the directive axis in a sea anemone. *Proceedings of the National Academy of Sciences of the United States of America*, 106(44):18592–18597, 2009. doi: 10.1073/pnas.0900151106.

L. Samuelov, Q. Li, R. Bochner, N. A. Najor, L. Albrecht, N. Malchin, T. Goldsmith, M. Grafi-Cohen, D. Vodo, G. Fainberg, B. Meilik, I. Goldberg, E. Warshauer, T. Rogers, S. Edie, A. Ishida-Yamamoto, L. Burzenski, N. Erez, S. A. Murray, A. D. Irvine, L. Shultz, K. J. Green, J. Uitto, E. Sprecher, and O. Sarig. Svep1 plays a crucial role in epidermal differentiation. *Experimental dermatology*, 26 (5):423–430, 2017. doi: 10.1111/exd.13256.

M. P. Sarras. Components, structure, biogenesis and function of the hydra extracellular matrix in regeneration, pattern formation and cell differentiation. *The International journal of developmental biology*, 56(6-8):567–576, 2012. doi: 10.1387/ijdb.113445ms.

M. P. Sarras. Hydra's ecm/integrin/matrix metalloproteinase network: Compara-

tive analysis of hydra and vertebrate ecm signaling systems based on current genomic and est databases. *MOJ Anatomy & Physiology*, 3(3), 2017. doi: 10.15406/mojap.2017.03.00094.

M. P. Sarras and R. Deutzmann. Hydra and niccolo paganini (1782-1840)–two peas in a pod? the molecular basis of extracellular matrix structure in the invertebrate, hydra. *BioEssays*, 23(8):716–724, 2001. ISSN 0265-9247. doi: 10.1002/bies.1101.

M. P. Sarras, M. E. Madden, X. Zhang, S. Gunwar, J. K. Huff, and B. G. Hudson. Extracellular matrix (mesoglea) of hydra vulgaris: Isolation and characterization. *Developmental biology*, 148:481–495, 1991. ISSN 0012-1606.

M. P. Sarras, L. Yan, A. Grens, X. Zhang, A. Ağbaş, J. K. Huff, P. L. St John, and D. R. Abrahamson. Cloning and biological function of laminin in hydra vulgaris. *Developmental biology*, 164(1):312–324, 1994. ISSN 0012-1606. doi: 10.1006/dbio.1994.1201.

R. Sato-Nishiuchi, I. Nakano, A. Ozawa, Y. Sato, M. Takeichi, D. Kiyozumi, K. Yamazaki, T. Yasunaga, S. Futaki, and K. Sekiguchi. Polydom/svep1 is a ligand for integrin $\alpha 9 \beta 1$. *The Journal of biological chemistry*, 287(30):25615–25630, 2012. ISSN 0021-9258. doi: 10.1074/jbc.M112.355016.

T. M. Schmid, R. Mayne, R. R. Bruns, and T. F. Linsenmayer. Molecular structure of short-chain (sc) cartilage collagen by electron microscopy. *Journal of ultrastructure research*, 86(2):186–191, 1984. ISSN 0022-5320. doi: 10.1016/s0022-5320(84)80057-x.

D. T. Schultz, S. H. D. Haddock, J. V. Bredeson, R. E. Green, O. Simakov, and D. S.

Rokhsar. Ancient gene linkages support ctenophores as sister to other animals. *Nature*, 618(7963):110–117, 2023. doi: 10.1038/s41586-023-05936-6.

R. S. Schwarz, T. C. G. Bosch, and L. F. Cadavid. Evolution of polydom-like molecules: identification and characterization of cnidarian polydom (cnpolydom) in the basal metazoan hydractinia. *Developmental and comparative immunology*, 32(10):1192–1210, 2008. ISSN 0145-305X. doi: 10.1016/j.dci.2008.03.007.

K. Seipel and V. Schmid. Mesodermal anatomies in cnidarian polyps and medusae. *The International journal of developmental biology*, 50(7):589–599, 2006. doi: 10.1387/ijdb.062150ks.

H. Shimizu, X. Zhang, J. Zhang, A. A. Leontovich, K. Fei, L. Yan, and M. P. Sarras. Epithelial morphogenesis in hydra requires de novo expression of extracellular matrix components and matrix metalloproteinases. *Development*, 129(6):1521–1532, 2002. ISSN 0950-1991.

C. Shinzato, K. Khalturin, J. Inoue, Y. Zayasu, M. Kanda, M. Kawamitsu, Y. Yoshioka, H. Yamashita, G. Suzuki, and N. Satoh. Eighteen coral genomes reveal the evolutionary origin of acropora strategies to accommodate environmental changes. *Molecular biology and evolution*, 38(1):16–30, 2021. doi: 10.1093/molbev/msaa216.

I. Shur, R. Socher, M. Hameiri, A. Fried, and D. Benayahu. Molecular and cellular characterization of sel-ob/svep1 in osteogenic cells in vivo and in vitro. *Journal of cellular physiology*, 206(2):420–427, 2006. doi: 10.1002/jcp.20497.

A. C. Shuttleworth. Type viii collagen. *The international journal of biochem-*

istry & cell biology, 29(10):1145–1148, 1997. ISSN 1357-2725. doi: 10.1016/S1357-2725(97)00033-2.

S. Siebert, J. A. Farrell, J. F. Cazet, Y. Abeykoon, A. S. Primack, C. E. Schnitzler, and C. E. Juliano. Stem cell differentiation trajectories in hydra resolved at single-cell resolution. *Science*, 365(6451):eaav9314, 2019. doi: 10.1126/science.aav9314.

P. Simon-Assmann. The laminin family: founding members of the basement membrane. *Cell adhesion & migration*, 7(1):44–47, 2013. doi: 10.4161/cam.23276.

T. J. Smith, M. N. Puttick, J. E. O'Reilly, D. Pisani, and P. C. J. Donoghue. Phylogenetic sampling affects evolutionary patterns of morphological disparity. *Palaeontology*, 64(6):765–787, 2021. ISSN 0031-0239. doi: 10.1111/pala.12569.

E. Sonpho, F. G. Mann, M. Levy, E. J. Ross, C. Guerrero-Hernández, L. Florens, A. Saraf, V. Doddihal, P. Ounjai, and A. S. Alvarado. Decellularization enables functional analysis of ecm remodeling in planarian regeneration. *Molecular Cell Proteomics*, 20(100137), 2020. doi: 10.1101/2020.09.11.293936.

W. T. Stauber, V. K. Fritz, and B. Dahlmann. Extracellular matrix changes following blunt trauma to rat skeletal muscles. *Experimental and molecular pathology*, 52(1):69–86, 1990. ISSN 0014-4800. doi: 10.1016/0014-4800{\%}2890{\%}2990060-Q.

D. J. Stefanik, L. E. Friedman, and J. R. Finnerty. Collecting, rearing, spawning and inducing regeneration of the starlet sea anemone, nematostella vectensis. *Nature protocols*, 8(5):916–923, 2013. doi: 10.1038/nprot.2013.044.

J. Steger, A. G. Cole, A. Denner, T. Lebedeva, G. Genikhovich, A. Ries, R. Reischl, E. Taudes, M. Lassnig, and U. Technau. Single-cell transcriptomics identifies conserved regulators of neuroglandular lineages. *Cell reports*, 40(12):111370, 2022. doi: 10.1016/j.celrep.2022.111370.

B. M. Steinworth, M. Q. Martindale, and J. F. Ryan. Gene loss may have shaped the cnidarian and bilaterian hox and parahox complement. *Genome biology and evolution*, 15(1), 2023. doi: 10.1093/gbe/evac172.

T. Stephenson. *The British Sea Anemones volume II*. The Ray Society, London, 1935. ISBN 978-1-904690-10-8.

A. Stokkermans, A. Chakrabarti, K. Subramanian, L. Wang, S. Yin, P. Moghe, P. Steenbergen, G. Mönke, T. Hiiragi, R. Prevedel, L. Mahadevan, and A. Ikmi. Muscular hydraulics drive larva-polyp morphogenesis. *Current biology : CB*, 2022. ISSN 0960-9822. doi: 10.1016/j.cub.2022.08.065.

M. Sutmuller, J. A. Bruijn, and E. de Heer. Collagen types viii and x, two non-fibrillar, short-chain collagens. structure homologies, functions and involvement in pathology. *Histology and histopathology*, 12(2):557–566, 1997. ISSN 0213-3911.

J. Taipale and J. Keski-Oja. Growth factors in the extracellular matrix. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 11(1):51–59, 1997. doi: 10.1096/fasebj.11.1.9034166.

T. Takeuchi, L. Plasseraud, I. Ziegler-Devin, N. Brosse, C. Shinzato, N. Satoh, and F. Marin. Biochemical characterization of the skeletal matrix of the massive coral,

porites australiensis - the saccharide moieties and their localization. *Journal of structural biology*, 203(3):219–229, 2018. doi: 10.1016/j.jsb.2018.05.011.

T. Tatusova, M. DiCuccio, A. Badretdin, V. Chetvernin, E. P. Nawrocki, L. Zaslavsky, A. Lomsadze, K. D. Pruitt, M. Borodovsky, and J. Ostell. Ncbi prokaryotic genome annotation pipeline. *Nucleic Acids Research*, 44(14):6614–6624, 2016. ISSN 0305-1048. doi: 10.1093/nar/gkw569.

U. Technau. Gastrulation and germ layer formation in the sea anemone nematostella vectensis and other cnidarians. *Mechanisms of development*, page 103628, 2020. doi: 10.1016/j.mod.2020.103628.

U. Technau and G. Genikhovich. Evolution: Directives from sea anemone hox genes. *Current biology : CB*, 28(22):R1303–R1305, 2018. ISSN 0960-9822. doi: 10.1016/j.cub.2018.09.040.

U. Technau and R. E. Steele. Evolutionary crossroads in developmental biology: Cnidaria. *Development*, 138(8):1447–1458, 2011. ISSN 0950-1991. doi: 10.1242/dev.048959.

F. Teufel, J. J. Almagro Armenteros, A. R. Johansen, M. H. Gíslason, S. I. Pihl, K. D. Tsirigos, O. Winther, S. Brunak, G. von Heijne, and H. Nielsen. Signalp 6.0 predicts all five types of signal peptides using protein language models. *Nature biotechnology*, 40(7):1023–1025, 2022. doi: 10.1038/s41587-021-01156-3.

A. C. Teuscher, E. Jongsma, M. N. Davis, C. Statzer, J. M. Gebauer, A. Naba, and C. Y. Ewald. The in-silico characterization of the caenorhabditis elegans matrisome and proposal of a novel collagen classification. *Matrix Biology Plus*, 1, 2019. doi: 10.1016/j.mbplus.2018.11.001.

A. D. Theocharis, S. S. Skandalis, C. Gialeli, and N. K. Karamanos. Extracellular matrix structure. *Advanced drug delivery reviews*, 97:4–27, 2016. doi: 10.1016/j.addr.2015.11.001.

A. D. Theocharis, D. Manou, and N. K. Karamanos. The extracellular matrix as a multitasking player in disease. *The FEBS journal*, 286(15):2830–2869, 2019. doi: 10.1111/febs.14818.

J. Thiery. Mise en evidence des polysaccharides sur coupes fines en microscopie electronique. *Journal of Microscopy*, 6:987–1018, 1967.

V. Thumuluri, J. J. Almagro Armenteros, A. R. Johansen, H. Nielsen, and O. Winther. Deeploc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Research*, 50(W1):W228–W234, 2022. ISSN 0305-1048. doi: 10.1093/nar/gkac278.

Tingyu Han and Chunpeng He. Full-length transcriptome maps of reef-building coral illuminate the molecular basis of calcification, symbiosis, and circadian genes. *International journal of molecular sciences*, 23(11135), 2022. doi: 10.3390/ijms231911135.

D. J. Tschumperlin. Matrix, mesenchyme, and mechanotransduction. *Annals of the American Thoracic Society*, 12 Suppl 1(Suppl 1):S24–9, 2015. doi: 10.1513/AnnalsATS.201407-320MG.

R. P. Tucker and J. C. Adams. Adhesion networks of cnidarians: a postgenomic view. *International review of cell and molecular biology*, 308:323–377, 2014. ISSN 1937-6448. doi: 10.1016/B978-0-12-800097-7.00008-7.

R. P. Tucker, B. Shibata, and T. N. Blankenship. Ultrastructure of the mesoglea of the sea anemone nematostella vectensis (edwardsiidae). *Invertebrate Biology*, 130(1):11–24, 2011. ISSN 10778306. doi: 10.1111/j.1744-7410.2010.00219.x.

A. Tursch, D. Mercadante, J. Tennigkeit, F. Gräter, and S. Özbek. Minicollagen cysteine-rich domains encode distinct modes of polymerization to form stable nematocyst capsules. *Scientific reports*, 6:25709, 2016. doi: 10.1038/srep25709.

A. Tursch, N. Bartsch, M. Mercker, J. Schlüter, M. Lommel, A. Marciniak-Czochra, S. Özbek, and T. W. Holstein. Injury-induced mapk activation triggers body axis formation in hydra by default wnt signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 119(35):e2204122119, 2022. doi: 10.1073/pnas.2204122119.

J. A. J. M. van de Water, M. Chaib De Mares, G. B. Dixon, J.-B. Raina, B. L. Willis, D. G. Bourne, and M. J. H. van Oppen. Antimicrobial and stress responses to increased temperature and bacterial pathogen challenge in the holobiont of a reef-building coral. *Molecular ecology*, 27(4):1065–1080, 2018. doi: 10.1111/mec.14489.

J. K. van de Wetering, L. M. G. van Golde, and J. J. Batenburg. Collectins: players of the innate immune system. *European Journal of Biochemistry*, 271(7):1229–1249, 2004. ISSN 0014-2956. doi: 10.1111/j.1432-1033.2004.04040.x.

G. van Rossum. *The Python language reference*, volume Pt. 2 of *Documentation for Python*. Python Software Foundation and SoHo Books, Hampton, NH and Redwood City, Calif., release 3.0.1 [repr.] edition, 2010. ISBN 1441412697.

S. Varshney, D. D. Hunter, and W. J. Brunken. Extracellular matrix components

regulate cellular polarity and tissue structure in the developing and mature retina. *Journal of ophthalmic & vision research*, 10(3):329–339, 2015. ISSN 2008-2010. doi: 10.4103/2008-322X.170354.

M. Veschgini, R. Suzuki, S. Kling, H. O. Petersen, B. G. Bergheim, W. Abuillan, P. Linke, S. Kaufmann, M. Burghammer, U. Engel, F. Stein, S. Özbek, T. W. Holstein, and M. Tanaka. Wnt/β-catenin signaling induces axial elasticity patterns of hydra extracellular matrix. *iScience*, 26(4):106416, 2023. doi: 10.1016/j.isci.2023.106416.

K. Wang, X. Meng, and Z. Guo. Elastin structure, synthesis, regulatory mechanism and relationship with cardiovascular diseases. *Frontiers in cell and developmental biology*, 9:596702, 2021. ISSN 2296-634X. doi: 10.3389/fcell.2021.596702.

J. Weber. Nematocysts (stinging capsules of cnidaria) as donnan-potential-dominated osmotic systems. *European Journal of Biochemistry*, 184(2):465–476, 1989. ISSN 0014-2956. doi: 10.1111/j.1432-1033.1989.tb15039.x.

T. Werner, G. Sweetman, M. F. Savitski, T. Mathieson, M. Bantscheff, and M. M. Savitski. Ion coalescence of neutron encoded tmt 10-plex reporter ions. *Analytical chemistry*, 86(7):3594–3601, 2014. doi: 10.1021/ac500140s.

C. A. Whittaker, K.-F. Bergeron, J. Whittle, B. P. Brandhorst, R. D. Burke, and R. O. Hynes. The echinoderm adhesome. *Developmental biology*, 300(1):252–266, 2006. ISSN 0012-1606. doi: 10.1016/j.ydbio.2006.07.044.

V. D. Winn, R. Haimov-Kochman, A. C. Paquet, Y. J. Yang, M. S. Madhusudhan, M. Gormley, K.-T. V. Feng, D. A. Bernlohr, S. McDonagh, L. Pereira, A. Sali, and S. J. Fisher. Gene expression profiling of the human maternal-fetal interface

reveals dramatic changes between midgestation and term. *Endocrinology*, 148 (3):1059–1079, 2007. ISSN 0013-7227. doi: 10.1210/en.2006-0683.

L. Wolpert, C. Tickle, and A. Martines Arias. *Principles of development*. Oxford Univ. Press, Oxford, 5. ed. edition, 2015. ISBN 9780199678143.

R. M. Wright, G. V. Aglyamova, E. Meyer, and M. V. Matz. Gene expression associated with white syndromes in a reef building coral, acropora hyacinthus. *BMC Genomics*, 16(1):371, 2015. doi: 10.1186/s12864-015-1540-2.

M. Wu, K. Cronin, and J. S. Crane. *StatPearls: Biochemistry, Collagen Synthesis*. NIH, NCBI, Treasure Island (FL), 2023.

L. Yan, G. H. Pollock, H. Nagase, and M. P. Sarras. A 25.7 x 10(3) m(r) hydra metalloproteinase (hmp1), a member of the astacin family, localizes to the extracellular matrix of hydra vulgaris in a head-specific manner and has a developmental function. *Development*, 121(6):1591–1602, 1995. ISSN 0950-1991. doi: 10.1242/dev.121.6.1591.

Y. J. Yang, C. S. Kim, B.-H. Choi, and H. J. Cha. Mechanically durable and biologically favorable protein hydrogel based on elastic silklike protein derived from sea anemone. *Biomacromolecules*, 16(12):3819–3826, 2015. ISSN 1525-7797. doi: 10.1021/acs.biomac.5b01130.

N. Yoshida, S. Yamamoto, T. Hamashima, N. Okuno, N. Okita, S. Horikawa, M. Hayashi, T. C. Dang, Q. L. Nguyen, K. Nishiyama, T. Makino, Y. Ishii, K. Tomihara, T. Shimizu, M. Shibuya, M. Noguchi, and M. Sasahara. Dysregulation of amphiregulin stimulates the pathogenesis of cystic lymphangioma. *Proceedings*

of the National Academy of Sciences of the United States of America, 118(19), 2021. doi: 10.1073/pnas.2019580118.

G. A. Young, A. B. Yule, and G. Walker. Adhesion in the sea anemones actinia equina l. and metridium senile (l.). *Biofouling*, 1(2):137–146, 1988. doi: 10.1080/08927018809378102.

E. A. Zamir, A. Czirók, C. Cui, C. D. Little, and B. J. Rongish. Mesodermal cell displacements during avian gastrulation are due to both individual cell-autonomous and convective tissue movements. *Proceedings of the National Academy of Sciences of the United States of America*, 103(52):19806–19811, 2006. doi: 10.1073/pnas.0606100103.

E. A. Zamir, B. J. Rongish, and C. D. Little. The ecm moves during primitive streak formation–computation of ecm versus cellular motion. *PLoS biology*, 6(10):e247, 2008. doi: 10.1371/journal.pbio.0060247.

C. Zenkert, T. Takahashi, M.-O. Diesner, and S. Özbek. Morphological and molecular analysis of the nematostella vectensis cnidom. *PloS one*, 6(7):e22725, 2011. doi: 10.1371/journal.pone.0022725.

X. Zhang and M. P. Sarras. Cell-extracellular matrix interactions under in vivo conditions during interstitial cell migration in hydra vulgaris. *Development*, 120 (2):425–432, 1994. ISSN 0950-1991.

X. Zhang, K. Fei, A. Ağbaş, L. Yan, J. Zhang, B. O'Reilly, R. Deutzmann, and M. P. Sarras. Structure and function of an early divergent form of laminin in hydra: a structurally conserved ecm component that is essential for epithelial mor-

phogenesis. *Development genes and evolution*, 212(4):159–172, 2002. ISSN 0949-944X. doi: 10.1007/s00427-002-0225-4.

X. Zhang, R. P. Boot-Handford, J. Huxley-Jones, L. N. Forse, A. P. Mould, D. L. Robertson, Lili, M. Athiyal, and M. P. Sarras. The collagens of hydra provide insight into the evolution of metazoan extracellular matrices. *The Journal of biological chemistry*, 282(9):6792–6802, 2007. ISSN 0021-9258. doi: 10.1074/jbc.M607528200.

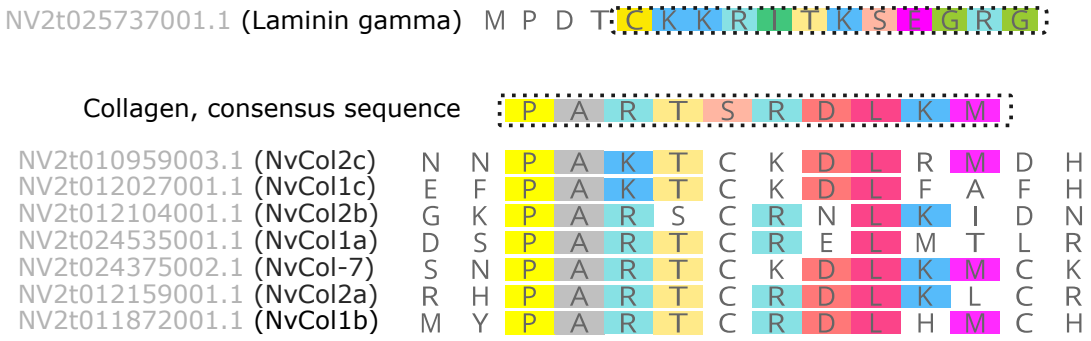
S. Zhou, S. Chen, Y. A. Pei, and M. Pei. Nidogen: A matrix protein with potential roles in musculoskeletal tissue regeneration. *Genes & diseases*, 9(3):598–609, 2022. doi: 10.1016/j.gendis.2021.03.004.

J. Zhu, X. Wang, C. He, and H. Wang. Mechanical properties, anisotropic swelling behaviours and structures of jellyfish mesogloea. *Journal of the mechanical behavior of biomedical materials*, 6:63–73, 2012. doi: 10.1016/j.jmbbm.2011.10.005.

B. Ziegler, I. Yiallouros, B. Trageser, S. Kumar, M. Mercker, S. Kling, M. Fath, U. Warnken, M. Schnölzer, T. W. Holstein, M. Hartl, A. Marciniak-Czochra, J. Stetefeld, W. Stöcker, and S. Özbek. The wnt-specific astacin proteinase has-7 restricts head organizer formation in hydra. *BMC biology*, 19(1):120, 2021. ISSN 101190720. doi: 10.1186/s12915-021-01046-9.

Appendix

6.1 Supplementray Images



Supplementary Figure 6.1. Epitopes of the Mesoglea Antibodies The NvLaminin antibody was raised against the C-terminal end of NvLaminin-γ. The Collagen-2c and NvPan Collagen antibodies were raised against a conserved region of fibrillar collagens

6.2 Supplementary Notes

The following RegEx patterns can be used to identify different known domains of the cnidocysts:

CRD

```
(?:C(?:[^C]{3})){2,4}CC
```

C-CRD

```
C[^C]{2}V(?:C[^C]{3}){2}CP[^C]{2}CC
```

N-CRD

```
(C[^C]{2}[^V](?:C[^C]{3}){2}C[^P][^C]{2}CC|
C[^C]{2}[^V](?:C[^C]{3}){2}C[^C]{3}CC|
C[^C]{2}V(?:C[^C]{3}){2}C[^P][^C]{2}CC)
```

Polyproline

```
P{9,}
```

6.3 Supplementary Data and Data Availability

All Supplementary Files and Datasets are available at the following repostories:

Github: https://heibox.uni-heidelberg.de/d/fc1348298c8045df9ca5/

Laboratory Server: /Gideon/Thesis_Organisation/Supplementary_Data

6.3.1 Source Code

Source Code 1: *In Silico* Matrisome Definition

Source Code 2: Orthogroup analysis

Source Code 3: Domain Analysis

Source Code 4: Statistical Analysis of MS Data (by Frank Stein)

Source Code 5: Analysis of MS Data

Source Code 6: Published Matrisome Databases Query

Source Code 7: Published matrisome only Analysis

Source Code 8: Supplementary Table Generation

6.3.2 Supplementary Tables

MS_curated_matrisome

This table contains the curated in silico matrisome and the respective MS data for each of the sequences found experimentally.

Full_MS_data.csv

This table contains all Proteins found by MS.

Nematostella_Matrisome.xlsx

This table contains the annotations of all proteins of the curated matrisome of Nematostella.

shRNA.xlsx

This table contains the annotations teh sequences of all ShRNAs used in this thesis.

Suppl_mass_spec_by_stage.xlsx

This table contains the up and downregulated genes for each life stage comparison.

ECMGeneExpressionAllCellTypes.xlsx

Differential expression of matrisome genes across all cell types

CnidocyteDEG.xlsx

Differentially Expressed genes in the Cnidocyte clusters