The Role of Paraxis and Integrins in the Development of *Nematostella vectensis*

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

by

Atharva Abhijit Valanju

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

Date: March 2025

Under the guidance of

Supervisor: **Dr. Matthew C. Gibson**,

Stowers Institute for Medical Research

From May 2024 to March 2025

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Certificate

This is to certify that this dissertation entitled “The Role of Paraxis and Integrins in the Development of *Nematostella vectensis”* submitted towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Atharva Abhijit Valanju at the Stowers Institute for Medical Research under the supervision of Dr. Matthew C. Gibson during the academic year 2024-2025.

Dr. Matthew C. Gibson

Committee:

Dr. Matthew C. Gibson

Dr. Girish Ratnaparkhi

This thesis is dedicated to my parents and grandparents,

who allowed my interest in biology to be nourished

DECLARATION

I hereby declare that the matter embodied in the report entitled “The Role of Paraxis and Integrins in the Development of *Nematostella vectensis*” are the results of the work carried out by me at the Stowers Institute for Medical Research, under the supervision of Dr. Matthew C. Gibson, and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.

Atharva Abhijit Valanju

20201173

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Metazoan body plans are built from complex repertoires of specialized cells assembled into functional structures. To achieve this complexity many bilaterian embryos undergo some form of metameric segmentation, which subdivides groups of cells into discrete blocks along the major body axis. Vertebrate somitogenesis is a classic example of metameric development, wherein paraxial mesoderm is sequentially partitioned into epithelial blocks whose axial identities are assigned by differential *Hox* gene expression. Cnidaria, as the sister phyla to Bilateria, until recently was not thought to exhibit segmentation. Remarkably, recent work demonstrated that an ancient segmentation program uses *Hox* genes to subdivide endo-mesodermal cells along the directive body axis in the cnidarian subphylum Anthozoa (corals and sea anemones). This thesis investigates potentially key factors in cnidarian segment morphogenesis that also regulate vertebrate somitogenesis: the transcription factor Paraxis and the integrin family of proteins. To characterize *paraxis* function in the cnidarian model *Nematostella*, short hairpin-mediated knockdowns and CRISPR-induced mutations in the *paraxis* ortholog were analyzed. In vertebrates, Paraxis drives somite epithelialization; however, examination of knockdowns and CRISPR-mediated mutations in *Nematostella* embryos did not reveal a similar role. Since integrin signaling is essential in cell-to-cell interactions and may be transcriptionally regulated by Paraxis, the integrin repertoire of *Nematostella* was characterized, and it was found that specific subunits were restricted to endo-mesodermal tissue, suggesting a potential role in the segmentation program. Together, these observations suggest that while a core set of conserved regulatory genes were present in the common ancestor of cnidarians and bilaterians, some sub-functionalization may have occurred in the 600 million years since their divergence.

ACKNOWLEDGEMENT

I would first like to thank Dr. Matt Gibson for providing me with the opportunity and support throughout my Master's project, as well as for allowing me the freedom to explore my ideas. I would like to thank my mentor Christof Nolte for his help and guidance across all aspects of the project – for supporting me when things did not work, patiently helping me troubleshoot experiments and navigate through a scientific project. I thank Emma Rangel-Huerta, Ruohan Zhong and Anna Klompen for their insightful discussions, technical help, and encouragement throughout my project. I thank Jenny Duong and Lacey Ellington for their help in getting me acquainted with *Nematostella* and teaching me the associated techniques. I am grateful to all the core facilities, especially aquatics and microscopy for their assistance. I thank all members of the Gibson Lab - Abby, Erica, Carlos, Helen, Steffi, Dilan, and everyone else for supporting me throughout my time in the lab and for being so helpful and welcoming.

I am thankful for all the friendships I built at the Stowers Institute and the people I was able to learn from. I am also grateful to my friends from IISER Pune, who despite the time difference, kept in touch and provided a familiar touch to life. I would also like to acknowledge the KVPY-Inspire Fellowship and the IISER Pune IDeaS Scholarship.

My parents and grandparents have been my pillars of support and I am grateful for their unwavering love and encouragement. I hope and I know from within, that my Aaji would be proud of me from above.

CONTRIBUTIONS

Contributor name	Contributor role
Dr. Matt Gibson, Dr. Christof Nolte, Atharva Valanju	Conceptualization Ideas
Dr. Matt Gibson, Dr. Christof Nolte, Jenny Duong, Atharva Valanju	Methodology
Dr. Christof Nolte, Lacey Ellington	Paraxis mutant generation
Dr. Anna Klompen, Jenny Duong	Adult male scRNA seq data
-	Software
Dr. Christof Nolte, Atharva Valanju	Validation
Atharva Valanju	Formal analysis
Atharva Valanju	Investigation
Dr. Matt Gibson	Resources
-	Data Curation
Atharva Valanju	Writing - original draft preparation
Dr. Matt Gibson, Dr. Christof Nolte	Writing - review and editing
Atharva Valanju	Visualization
Dr. Matt Gibson, Dr. Christof Nolte	Supervision
Dr. Matt Gibson	Project administration
Dr. Matt Gibson	Funding acquisition

CHAPTER 1 – INTRODUCTION

"An embryo is not a simple thing, nor a single thing, but a vast community of cells, communicating, cooperating, and competing to shape the body's final form."

— Lewis Wolpert

How did the variety of life and phyla that populate the world evolve from simple, multicellular organisms? One possible mechanism is the organization of cells into discrete blocks of tissue or segments, from which complexity is built (Chipman, 2010). During embryonic development, many metazoans exhibit some form of segmentation, which appears as repetitive units along the primary body axis. While the process of segmentation might appear homologous amongst the different phyla of bilaterians, implying that it is derived from a common ancestor, its evolutionary history is still actively debated. As a paradigm in biology, segmentation is an extensively used term without a clear definition, as it is not simply a bimodal characteristic but rather encompasses a complex suite of traits (Hannibal and Patel, 2013). This loose definition allows for a greater number of animals to be considered as pseudo or partially segmented.

True segmentation is thought to be a characteristic of three distinct phyla: Arthropoda, Annelida, and Chordata (Fig. 1.1). In *Drosophila melanogaster*, the process of segmentation proceeds in hierarchical stages and simultaneously divides the entire length of the embryo into segments (Nüsslein-Volhard and Wieschaus, 1980). However, *Drosophila* embryos exhibit a highly derived mode of segmentation that is not a representation of arthropods in general (Peel et al. 2005). Many other arthropods, some annelids, and even vertebrate embryos exhibit a pattern of segmentation that involves sequential partitioning with an anterior-to-posterior segment boundary formation (Chipman, 2010). Representative species from these three phyla have revealed striking similarities in the molecular mechanisms governing the formation of a segmented body plan during embryonic development (Müller et al., 1996; Stollewerk et al., 2003; Rivera and Weisblat, 2009). This has led to the hypothesis that the common "Urbilaterian"

ancestor might have been segmented, and that segmentation is a derived ancestral trait (Peel and Akam, 2003). Yet this hypothesis cannot account for the majority of phyla that display no evidence of segmentation, implying that this mechanism may have been lost in over twenty phyla.

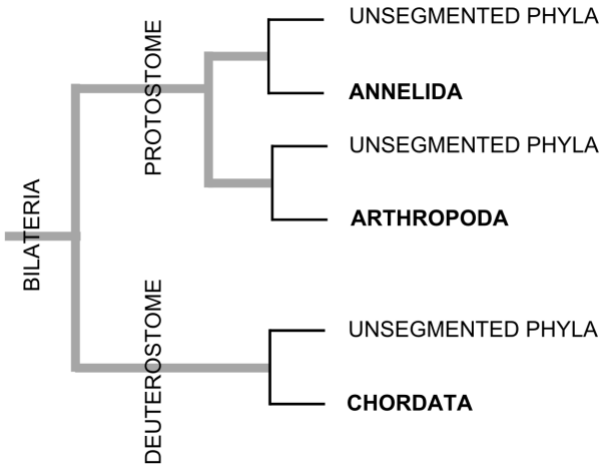


Fig. 1.1 A truncated phylogenetic tree that highlights the three classically segmented phyla (Annelida, Arthropoda and Chordata).

1.1 Vertebrate segmentation

Vertebrate segmentation along the primary body axis occurs through the process of somitogenesis, which involves the periodic budding of blocks of mesenchymal tissues from the presomitic mesoderm (PSM) (Fig. 1.2). These blocks of mesenchymal cells that undergo epithelization are called somites and eventually give rise to important components of the vertebrate body plan - the dermomyotome, which differentiates into the dermis and muscles; and the sclerotome, which forms the axial skeleton (Chal and Pourquie, 2009).

The process begins with an epithelial-to-mesenchymal transition and the specification of stem-like progenitors that ingress and contribute to the PSM during gastrulation (Dequéant and Pourquié, 2008). Early observations revealed that somites are produced

with a cyclic periodicity that is characteristic of the vertebrate species being studied (Romanoff, 1960; Aulehla et al., 2008; Palmeirim et al., 1997; Schröter et al., 2008). A mechanistic model - the clock and wavefront model soon emerged to explain this phenomenon, and it remains a widely accepted and favored framework (Cooke and Zeeman, 1976) (Fig 1.2).

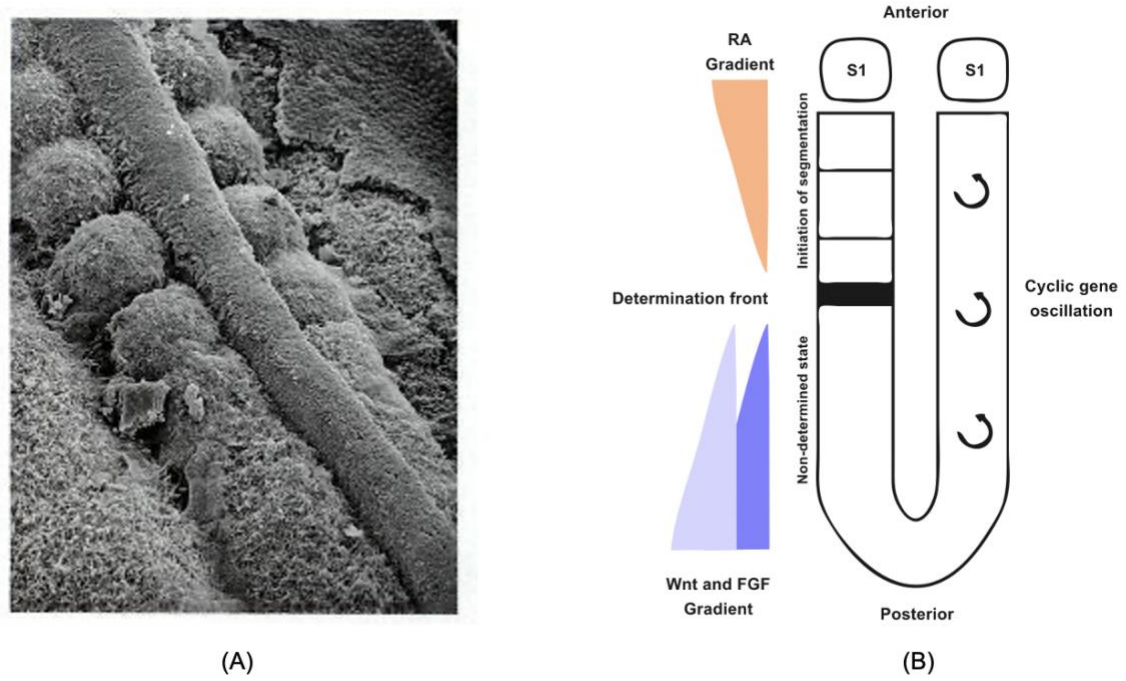


Fig. 1.2 (A) Scanning electron microscopy image of the paraxial mesoderm and the newly formed somites. (B) The clock and wavefront model of somitogenesis. S1 represents the mostly recently formed somite that has separated from the PSM; and as the determination front (black stripe) moves posteriorly, future somite boundaries are specified in concordance with the cyclic oscillation of the clock genes. (A) Image from Gilbert SF. Developmental Biology. 6th edition. (B) Image modified from Gibb et al., 2010.

The segmentation clock drives the cyclical expression of several genes belonging to the *Hairy/Hes/Her* family (in different species) in a posterior to anterior fashion through the PSM (Palmeirim et al., 1997; Bessho et al., 2003; Henry et al. 2002). This expression pattern is intrinsic to the PSM and involves the periodic transcriptional switching of these genes (Maroto and Bone et al., 2012). Notch-Delta signaling has also been found to play a critical role in maintaining this wave of expression and synchronizing it in a

cell-to-cell coordinated fashion (Gibb et al., 2010; Ozbudak and Lewis, 2008). This wave travels through the PSM and collapses at the anterior determination front, where the future somitic boundaries are specified (Maroto and Bone et al., 2012). The determination front (also called the wavefront) is positioned by two opposing gradients - an anterior gradient of retinoic acid (RA) signaling and a posterior gradient of FGF/Wnt signaling (Sawada et al., 2001; Moreno and Kintner, 2004; Diez del Corral and Storey, 2004). The convergence of these two gradients at the wavefront establishes a boundary: posterior cells remain in an immature, non-determined state under the influence of FGF signaling, while anterior cells, upon activation by the cycling genes, initiate the segmentation program (Dubrulle et al., 2001). The precise somite boundary formation is brought about by a complex cascade that involves the restriction of *Mesoderm posterior 2 (Mesp2)* expression to the anterior half of the prospective somite, triggering Eph-ephrin signaling and ultimately leading to the creation of a new furrow at the anterior end of the PSM (Takahashi et al., 2010; Watanabe et al., 2009).

Superimposed on the developing somites is a positional code generated by the combinatorial expression of *Hox* genes along the body axis. Depending on its position along the AP axis, each somite expresses a unique combination of *Hox* genes that results in the characteristic patterning of the somite derivatives (Kessel and Gruss, 1991). Pharmacological-mediated or genetic alterations of the *Hox* code result in homeotic transformations of the skeletal muscles and their associated nerves (Mallo et al., 2011). The expression of the *Hox* genes follows a spatio-temporal colinear order whereby genes located more 3’ within the cluster are expressed earlier and set their border more anteriorly to that of their neighboring *Hox* gene in the cluster (Krumlauf, 1994). Part of the establishment of this border in PSM and lateral plate mesoderm is regulated by the opposing gradients of retinoic acid and FGF signaling, thus coordinating the process of somitogenesis with the acquisition of AP identity (Aulehla et al., 2010).

1.2 Role of *paraxis* in vertebrate segmentation

Many of the factors involved in somitogenesis have been identified and studied. As described above, there are genes whose proteins specify segmental identity and help coordinate cellular associations into metameric units. At the transcriptional level, this involves the differential expression of transcription factors to implement these changes. For example - *paraxis* is a basic helix-loop-helix (bHLH) transcription factor whose expression in multiple vertebrate species is restricted to the anterior presomitic mesoderm and the newly formed somites (Burgess et al., 1995; Barnes et al., 1997; Shanmugalingam et al., 1998). As the somites mature, the expression of *paraxis* starts to gradually decline from some compartments while being maintained in the epithelial dermatome (Burgess et al., 1995; Šošić et al., 1997). Understanding the induction of *paraxis* has led to confounding results from multiple studies, but one hypothesis is that *paraxis* is induced in the PSM by signals from the surface ectoderm and then maintained in the newly formed somites by signals from the neural tube and surface ectoderm (Šošić et al., 1997). The signal from the surface ectoderm is thought to be Wnt6, which, through the activation of its corresponding Frizzled7 receptor, activates beta-catenin signaling, resulting in *paraxis* expression (Linker et al., 2005). However, there have also been other reports that indicate that the surface ectoderm is not essential for initiating *paraxis* expression but instead plays a role in maintaining its expression pattern in maturing somites (Rifes et al., 2007; Linker et al., 2005).

The importance of *paraxis* expression during somitogenesis is attributed to the fact that it is a transcription factor (TF) belonging to the Twist subfamily, potentially controlling a repertoire of genes involved in epithelization. Biochemical studies have shown that Paraxis heterodimerizes with specific E-box proteins (another type of bHLH proteins), leading to the formation of a bipartite DNA binding domain composed of basic residues (Rawls et al., 2004). Studies with homozygous mutant mice have revealed that the loss of *paraxis* function affects the mesenchymal-to-epithelial transition (MET) during somitogenesis (Burgess et al., 1996; Rowton et al., 2013). In these mutants, the PSM undergoes crude segmentation, but epithelial somites fail to form properly (Burgess et al., 1996). Consequently, the skeletal elements and their associated muscles are poorly

patterned (Burgess et al., 1996) and the neonatal mice die shortly after birth. Part of its role in MET is in the demonstration that *paraxis* regulates genes involved in fibronectin assembly and polarization of somites along the A/P axis (Rowton et al., 2013; Johnson et al., 2001). These observations suggest that Paraxis plays a central role in segmentation through the regulation of genes involved in the MET process.

1.3 Interplay of integrins and the extracellular matrix in vertebrate segmentation

How the presomitic and somitic cells interact to coordinate their clustering and movement is facilitated in part by physical interactions through integrins. Integrins are cell surface receptors that primarily mediate contact of the cell with components of the extracellular matrix (ECM). They are composed of two subunits - an alpha and a beta subunit that come together to form a heterodimer via non-covalent interactions (Fig 1.3). Integrins modulate multiple and diverse signaling pathways, including NF-κB, AKT, JNK, and ERK, among others (Hynes, 2002). They are involved in outside-in signaling as well as non-canonical inside-out signaling through changes in their conformation states and modifications to the cytoplasmic tail (Ginsberg et al., 1992; Hynes, 2002). Through adapter proteins interacting with their cytoplasmic domains, integrins also mediate contact with the cytoskeleton, becoming involved in cell motility and adhesion (DeMali et al., 2003).

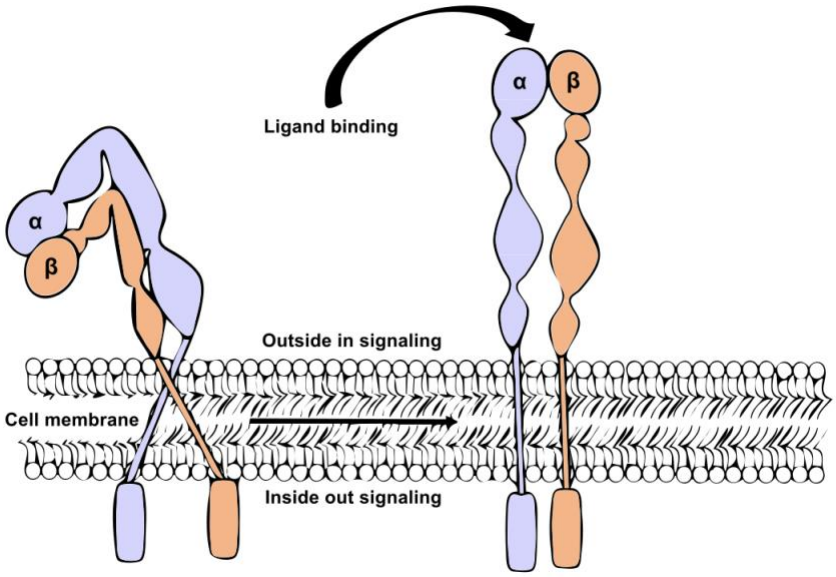


Fig. 1.3 Graphical representation of integrin receptors undergoing a conformational change. The alpha and beta subunits heterodimerize and form a receptor that can change conformation from a closed to an open state depending on the signal received. Image modified from Hynes et al., 2002.

Vertebrates possess a large repertoire of integrin subunits, with at least 18 alpha and 8 beta subunits that are well characterized in humans (Hynes, 2002). The alpha subunits can be further categorized based on their ECM ligand specificity, including RGD binding, LDV binding, collagen binding, and laminin-binding groups (Humphries and Byron et al., 2006). Loss-of-function studies have revealed that integrins are involved in diverse but distinct physiological and developmental processes (Hynes, 2002). For example, Integrin β1 is a subunit that can associate with at least 12 different alpha subunits and was found essential at the earliest stage of gastrulation, with null embryos often failing to develop shortly after implantation (Fässler et al., 1995).

In the context of somitogenesis, Integrin α5 (*Itga5*) has been implicated to play an important role, as homozygous mutant embryos die presenting defects in the posterior somites and the PSM (Yang et al., 1993). When present as an α5β1 complex, this subunit binds to fibronectin, and subsequently, it has also been seen that these defects in somites and mesoderm patterning are more extreme (and extend anteriorly) in

mutants for fibronectin (Yang et al., 1993; George et al., 1993). Fibronectin is a major component of the ECM and plays an important role in cell adhesion and migration at various stages. Subsequent studies have shown that during somitogenesis, fibronectin is localized to the surface ectoderm while its cognate receptor, *Itga5*, is expressed by the PSM (Rifes et al., 2007).

Interestingly, Paraxis, through the regulation of the *fibroblast activating protein alpha* gene, stimulates the production and organization of fibronectin in the ECM (Rowton et al., 2013; Lee et al., 2011). Furthermore, the loss of Paraxis function leads to the downregulation of *Itgav,* which is required for the formation of focal adhesions in somites and is normally expressed throughout the PSM and, along with *Itga5,* binds to fibronectin (Rowton et al., 2013; Yang et al., 1999; Girós et al., 2011). These observations demonstrate a remarkable interplay between the roles of Paraxis, integrin signaling, and ECM components during the process of vertebrate somitogenesis (Fig 1.4).

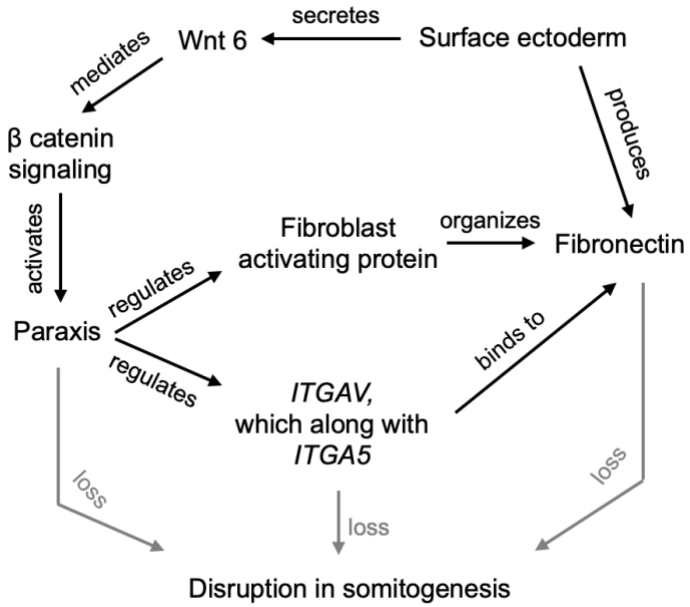


Fig. 1.4 A proposed scheme of interactions involving the regulation of Paraxis, integrins and ECM components. A dysregulation in any of these components leads to defects in somitogenesis. Proposed based on Yang et al., 1993; Linker et al., 2005; Rowton et al., 2013.

1.4 An emerging cnidarian model system

To gain an understanding in the development of complex bilaterian traits, extant species that predate the 'Urbilaterian' split can be a useful resource. As a sister group to Bilateria, the Cnidarian phylum offers the potential to uncover the origins of bilaterian complexity (Layden et al., 2016). Cnidarians are an early branching metazoan lineage consisting of diverse species living in an aquatic environment and are primarily composed of sea anemones, jellyfish, coral and hydroids. The phylum is divided into two major clades: the Medusozoans, which generally undergo a polyp-to-medusae life cycle transition, and the Anthozoans, members of which do not undergo this transition and typically remain as a polyp throughout the reproductive stage (Technau and Steele, 2011). The defining feature of cnidarians is the presence of a special type of stinging cell called the nematocyte, used for predation, defense, and adhesion (Technau and Steele, 2011). Cnidarians are traditionally considered to have two germ layers (the endoderm and the ectoderm), and most possess an axis of radial symmetry. However, anthozoan polyps are an exception to this with an internal asymmetric organization (Technau and Steele, 2011). Lacking a central nervous system, cnidarians are thought to possess a more simplistically organized diffuse 'nerve net' (Röttinger, 2021). With these interesting set of primitive characteristics, cnidarians have gained considerable interest to study key processes involved in the origin of the third germ layer (the mesoderm), the evolution of a central nervous system, and the development of bilateral symmetry.

While Hydra has historically been widely used as a model, particularly for regeneration-based studies, the sea anemone *Nematostella vectensis* has rapidly gained acceptance as a cnidarian model species. It belongs to the Anthozoan clade and is typically found in estuarine environments. The morphology of the adult *Nematostella* animal is simple, consisting of two germ layers - an outer ectoderm and an inner endoderm (Layden et al., 2016), separated by mesoglea, which is primarily an acellular matrix composed of

amoeboid cells and ECM components (Tucker et al., 2011). A single oral opening serves as both the mouth and the anus and can be surrounded by as many as 16 tentacles in a mature adult (Fig. 1.5). The opening then leads to a distinct pharynx and eight radially arranged mesenteries. The mesenteries, running along the oral aboral axis, comprise gonadal cells, myoepithelial cells, cnidocytes, and cells involved in nutrient absorption and storage (Layden et al., 2016).

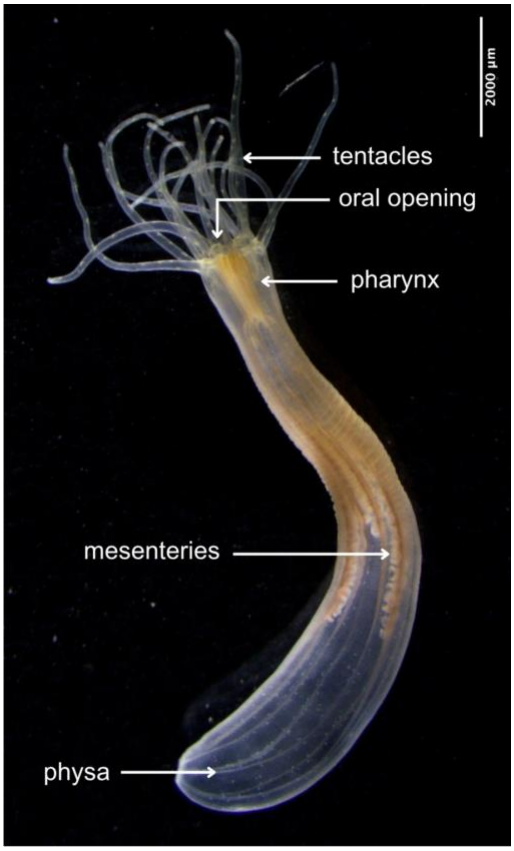


Fig. 1.5 Morphology of an adult *Nematostella vectensis* polyp. The mesenteries are located on the inner side of the translucent body wall ectoderm. The aboral end is also referred to as the physal end. Image acquired on Leica MN165 BW.

Over the years, a wide variety of resources and tools have been developed for this model, ranging from extensive genomic and transcriptomic databases (Putnam et al., 2007; Warner et al., 2018), CRISPR Cas9-based genome editing (Ikmi et al., 2014), shRNA based knockdown (He et al., 2018), and multiple imaging methods (Röttinger, 2021). Its popularity also stems from the fact that methods for its propagation and

spawning in the laboratory are well documented (Hand and Uhlinger, 1992). Being a broadcast spawner induced by light and temperature allows for the regular procurement of fresh embryos for manipulation.

The life cycle of *Nematostella* has also been well studied, and fertilization to gamete formation takes around 10-12 weeks, depending on the temperature and feeding schedule (Hand and Uhlinger, 1992). Males and females are present as distinct sexes; however, they are not overtly distinguishable by any morphological difference (Röttinger, 2021). Soon after fertilization, the zygote starts dividing and goes through classical stages of embryonic development - early cleavage, blastula, gastrula, planula (which is a free-swimming stage), juvenile polyp, and reproductive polyp (Layden et al., 2016).

Studies have shown that an intrinsic polarity exists in *Nematostella* embryos, with later cleavages leading to selective beta-catenin degradation at the future aboral pole and nuclear stabilization at the site of gastrulation, marking the future endoderm (Wikramanayake, 2003). The process of gastrulation is known to initiate at the animal pole and occurs via invagination (Magie et al., 2007; Layden et al., 2016). Gastrulation, in bilaterians as well as in cnidarians, is linked to the formation of an oral-aboral axis. In *Nematostella,* this primary axis runs from the singular oral opening to the opposite aboral end. The role of Wnt signaling in the initial formation of the oral axis is still debated; however, beginning from the mid-gastrula stage, staggered expression domains of individual *wnt* genes are seen in the oral half of the embryo and may be involved in determining separate territories and eventually patterning the oral-aboral axis (Layden et al., 2016).

Nematostella also possesses another axis that runs orthogonal to the oral aboral axis - the directive axis. Functional studies have revealed that this axis is established and patterned by asymmetric BMP signaling (Genikhovich, 2015). Surprisingly, *Nematostella NvDpp* (*Bmp2/4* in vertebrates) and *NvChordin*, unlike in vertebrates, are not expressed in antagonistic gradients but are instead found on the same asymmetric side of the embryo after being initially expressed in the radial blastopore ring (Technau

and Steele, 2011). It has been found that BMP signaling is mediated by its own negative feedback loop and a possible shuttling mechanism, with its peak occurring at the end opposite to the source of *NvDpp* and forming a signaling gradient along the directive axis (Saina et al., 2009).

1.5 Homology of the axes and a segmentation program

Since the discovery that anthozoans possess two orthogonal body axes, there has been a lot of interest in understanding if the two axes are homologous to their bilaterian AP/DV counterparts and if the directive axis in anthozoans could have eventually led to the dorsoventral axis in bilaterians. It has also been proposed that the oral-aboral axis of cnidarians might correspond to the animal-vegetal axis of bilaterians, with both involving canonical Wnt signaling specification and blastopore organizing activity shown by transplantation experiments in *Nematostella* (Technau and Steele, 2011).

In *Nematostella* embryos, during larval development, the invaginating endoderm undergoes a morphological segregation of tissues that becomes apparent as eight bi-radial segments along the directive axis at the planula stage (Fig. 1.6). These anatomical subdivisions correlate with the positioning of the first four tentacles in the polyp and eventually give rise to the eight mesenteries in the adult body.

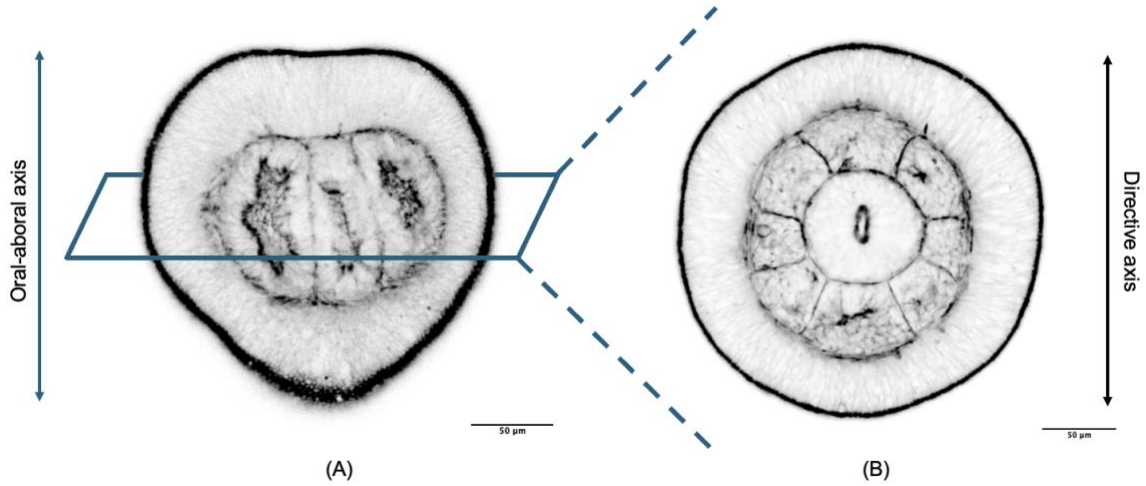


Fig. 1.6 Representative image of planula larvae (wild type) stained to label F-actin architecture. (A) Visualization of the lateral/side view that demonstrates the oral aboral axis with the oral pole facing upward. (B) Visualization of the oral view demonstrating the directive axis. Image acquired on the Andor DragonFly 200 spinning disc confocal microscope.

A recent study has shown that a larval Hox-Gbx code in *Nematostella* specifically defines these endodermal segment boundaries with a nested pattern of *hox* gene expression arising in a stepwise manner as the endodermal tissue develops (He et al., 2018) (Fig. 1.7 A). Functional analysis using shRNA knockdowns and CRISPR Cas9 genome editing also revealed a loss of segmental identity, resulting in endodermal fusion and tentacle patterning defects in the absence of individual *hox* gene expression (He et al., 2018). BMP signaling was found to be acting upstream to the expression of these *hox* genes and possibly regulating their expression in a signaling gradient-dependent manner (Genikhovich et al., 2015).

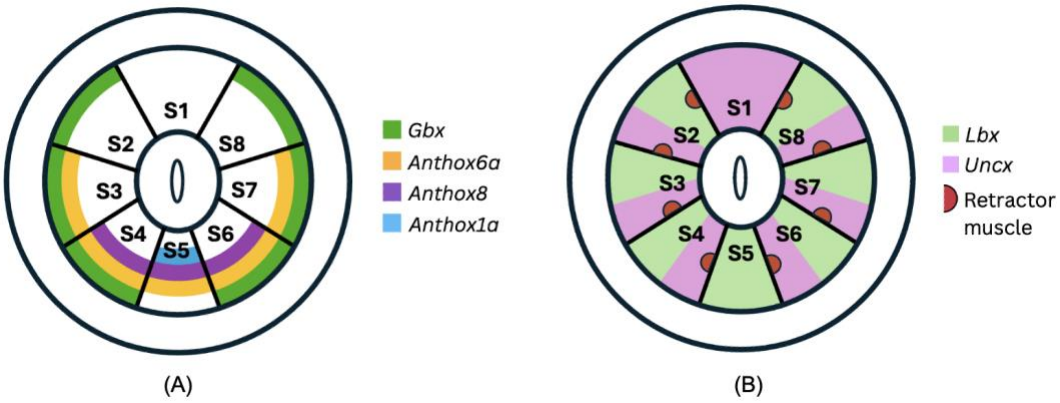


Fig. 1.7 (A) The Hox-Gbx code (B) Segment polarity program in *Nematostella vectensis* planula. The Hox-Gbx genes exhibit a nested pattern of expression and regulate the formation of segmental boundaries. The segments themselves possess an internal polarity marked by domains of *Lbx* and *Uncx* expression. Figures taken from He et al., 2018; He et al., 2023.

Spatial transcriptomics has also revealed the existence of a segmental polarity program in these endodermal segments. Controlled by BMP signaling and the Hox Gbx pathway, two conserved homeobox domain-containing genes, *Lbx* (the *ladybird* homolog) and *Uncx* (the *Unc-4* homolog) establish opposite subsegmental domains concurrently with

the segmentation process at the planula stage (He et al., 2023) (Fig. 1.7 B). Both these genes are considered as polarity genes in a diverse range of segmented bilaterians; particularly, *Uncx4.1* is expressed and patterns the posterior half of vertebrate somites (He et al., 2023; Mansouri et al., 2000). These observations of orthologous genes whose products function in a similar manner have added substantial support to the idea that segment polarization is a distinct characteristic found in a vast majority of segmented animals and allows for the formation of complex metameric structures while maintaining regionalization of function (He et al., 2023).

Recent studies in *Nematostell*a have challenged the two-germ layer homology and have instead proposed that the cnidarian endoderm is homologous to the bilaterian mesoderm, and the cnidarian pharyngeal ectoderm corresponds to the bilaterian endoderm (Steinmetz et al., 2017). This is supported by the enriched expression of orthologs of several genes associated with bilaterian mesoderm specification and differentiation, in *Nematostella* endoderm (which has now been designated as the endomesoderm) (Steinmetz et al., 2017). These potentially conserved expression patterns include multiple genes whose bilaterian orthologs are involved in the formation of skeletal muscle (*eya*, *six1/2*, *dachshund*, *lbx*), cardiac muscle (*nk4*/*nkx2.5*/*tinman*, *hand, gata4*), visceral mesoderm (*foxC*, *nkx3*/*bagpipe*, *six 4/5*) as well as vertebrate somites (*Twist*, *Tbx15*, *Paraxis/Scleraxis*) (He et al., 2023; Steinmetz et al., 2017). These studies establish that the *Nematostella* endomesoderm is a complex molecular tissue, and further experiments can provide insights into the nature and evolutionary history of mesoderm specification and axis formation.

1.6 Paraxis in *Nematostella*

The bHLH superfamily of transcription factors is an evolutionary ancient family with roots going back to the pre-metazoan split. This class of genes appear to have undergone a massive tandem gene duplication event, particularly in *Nematostella vectensis*, giving rise to around 68 bHLH genes that were found in the then-unannotated genome (Simionato et al., 2007). This is significantly more than most invertebrate

bilaterians and added support to the idea that *Nematostella* is not as primitive as its phylogeny may indicate.

Paraxis is a member of the bHLH family and is of particular interest because of its involvement in the vertebrate somitogenesis cascade. *Nematostella* was surprisingly found to possess an ortholog of this vertebrate gene, with its expression appearing restricted to the pharyngeal endoderm at the polyp stage (Steinmetz et al. 2017). Spatial transcriptomic analysis, supported by *in situ* hybridization, further revealed *paraxis* expression in the endomesodermal tissue that gives rise to *Nematostella* segments during the developing planula stage (He et al., 2023).

A previous unpublished study from our lab had identified the *paraxis* ortholog in the annotated *N. vectensis* genome and used a shRNA based knockdown approach to ascertain its function during early developmental stages (RR Talele - MS thesis, IISER Pune 2024). Preliminary data suggested that when *paraxis* was knocked down, the developing endomesodermal segments appeared to display morphological deformities with incomplete segment boundaries and a fused segmental architecture.

1.7 Integrins across evolution and in *Nematostella*

It was originally believed that integrins were a metazoan novelty; however, core components of the integrin-adhesion complex have now been found in ancestral eukaryotic species going back to unicellular protists (Sebé-Pedrós et al., 2010). Early studies in metazoan lineages seemed to suggest that as animals increased in complexity, they acquired a larger repertoire of integrin subunits, with *C. elegans* possessing 2 α and 1 β subunit, *D. melanogaster* possessing 5 α and 2 β subunits, and vertebrates with 18 α and 8 β subunits. This was soon demonstrated not to be true, with a large number of integrin subunits being discovered in multiple basal metazoan species, including cnidarians.

The anthozoan coral *Acropora millepora* is found to possess three integrin subunits whose expression during early development is restricted to the presumptive endoderm

encompassed within the blastopore lip (Knack et al., 2008). In *Nematostella*, a study from 2014 identified 2 α and 4 β integrin subunits. Expression profiling at the adult polyp stage revealed widespread expression across multiple tissues in a subunit-dependent manner, including the gonads, gastrodermis, mesenteric filaments, and oral epidermis (Gong et al., 2014). Changes in the expression of these subunits were examined in a regenerative context through horizontal transection of the body column, revealing a significant upregulation during the regeneration process (Gong et al., 2014). However, no study so far has looked at the expression and role of integrin subunits during early embryonic development in *Nematostella*.

Hypothesis

Nematostella endomesodermal segments are homologous to vertebrate somites. Their formation is governed by an ancestral molecular toolkit utilizing a genetic cascade homologous to vertebrate somitogenesis.

Cnidaria, as a sister group to Bilateria, is perfectly positioned to understand the evolution of complex bilaterian traits, including segmentation. To uncover homology between structures, it is necessary to use integrative approaches as well as comprehensive molecular, morphological and phylogenetic analysis. Driving the argument that vertebrate somites are homologous to endomesodermal segments in *Nematostella* are three main lines of evidence:

1. The endomesoderm in *Nematostella* transcriptionally resembles bilaterian mesoderm with expression of major mesoderm specification and differentiation genes.

2. Existence of a cnidarian Hox code that provides segmental identity along the directive axis. Vertebrates also show a staggered pattern of *Hox* gene expression that patterns the somites along the AP axis.

3. A segment polarity program in the endomesodermal segments that is reminiscent of the polarized patterning in somites.

Morphologically, as well, it might be tempting to draw parallels between the epithelial architecture of the endomesodermal clusters and the epithelialization of the somites.

To further explore this hypothesis, it would be valuable to elucidate additional molecular pathways and their interplay with other components that may contribute to the establishment and patterning of endomesodermal segments in *Nematostella*.

Goal of this thesis

This thesis aims to investigate the function of the bHLH transcriptional factor Paraxis in the formation of *Nematostella* segments, drawing from its vertebrate ortholog's essential contribution to somite epithelialization.

An additional focus is to elucidate the role of integrin signaling during early development in *Nematostella* and its potential involvement in the segmentation program. To achieve this, all potential integrin subunit homologs in *Nematostella* will be identified and analyzed for their evolutionary relationship. Furthermore, the expression pattern of a set of selected subunits will be examined during early development to provide insights into their functional relevance.

CHAPTER TWO - MATERIALS AND METHODS

Materials

2.1 Gene names and IDs for *Nematostella vectensis*

The nomenclature for genes is based on the gene name and description provided in the SIMRBASE annotation of the Nvec200 genome.

Gene name used	Gene ID (SIMRBASE NV2)	Protein ID (UniProt)
NvParaxis	NV2g010865001.1	A0A1T4JGW9
NvITGB1	NV2g012164000.1	A7RGD5
NvITGB	NV2g021343000.1	A7SM23
NvITGB2	NV2g011319000.1	A7RRW0
NvITGBn2	NV2g010694000.1	A7T2U5
NvITGA5	NV2g011383000.1	-
NvITGA8	NV2g012108000.1	A7RG71
NvITGA3	NV2g011909000.1	A7RHH6
NvITGA9	NV2g011572000.1	-

2.2 Protein sequences for phylogenetic analysis

The *Nematostella* integrin protein sequences have been acquired from SIMRBASE using the gene IDs provided in table 2.1, except the sequence for *NvITGBn2*, which is truncated in SIMRBASE and so has been acquired from UniProt (A7T2U5). Sequences

for all other proteins have been acquired from UniProt.

Species name	Integrin α protein subunit name	UniProt ID	Integrin β protein subunit name	UniProt ID
Geodia cydonium	GEOCY ITA	O18428	GEOCY ITGB	O97189
Podocoryna carnea	PODCA ITA	Q9GSF4	PODCA ITGB	Q9GSF3
Acropora millepora	ACRMI ITGA1	B2XW55	ACRMI ITGB_O1	O17494
			ACRMI ITGB_B2	B2XW56
Caenorhabditis elegans	CAEEL INA1	Q03600	CAEEL_ITGB	Q27874
	CAEEL PAT2	P34446		
Drosophila melanogaster	DROME ITA1	Q24247	DROME ITGB_PS	P11584
	DROME ITA2	P12080	DROME ITGB_NU	Q27591
	DROME ITA3	O44386		
Homo sapiens	HUMAN ITGA1	P56199	HUMAN ITGB1	P05556
	HUMAN ITGA2	P17301	HUMAN ITGB2	P05107
	HUMAN ITGA3	P26006	HUMAN ITGB3	P05106
	HUMAN ITGA4	P13612	HUMAN ITGB4	P16144
	HUMAN ITGA5	P08648	HUMAN ITGB5	P18084
	HUMAN ITGA6	P23229	HUMAN ITGB6	P18564
	HUMAN ITGA7	Q13683	HUMAN ITGB7	P26010
	HUMAN ITGA8	P53708	HUMAN ITGB8	P26012
	HUMAN ITGA9	Q13797		
	HUMAN ITGA10	O75578		

2.3 Oligonucleotides

shRNA primers	
Name	Sequence (5’-3’)
paraxis shRNA	AAG ACT CGG ATA GCT CTC TAA TCT CTT GAA TTA GAG AGC TAT CCG AGT CTA TAG TGA GT
ITGB1 - 1 shRNA	AAG GTC TAT GCT TGA GGA CAT TCT CTT GAA ATG TCC TCA AGC ATA GAC CTA TAG TGA GT
ITGB1 - 2 shRNA	AAG GCC CTG AGC ATG GTA TAT TCT CTT GAA ATA TAC CAT GCT CAG GGC CTA TAG TGA GT
ITGB - 1 shRNA	AAG GAC TAC TTC GTC AGA CAG TCT CTT GAA CTG TCT GAC GAA GTA GTC CTA TAG TGA GT
ITGB - 2 shRNA	AAG TAC GAG GCA TCC ACA CTA TCT CTT GAA TAG TGT GGA TGC CTC GTA CTA TAG TGA GT
ITGB2 - 1 shRNA	AAG GCC AGG TAT TCC ATC TAT TCT CTT GAA ATA GAT GGA ATA CCT GGC CTA TAG TGA GT
ITGB2 - 2 shRNA	AAG ACG GAC AGT GCC ATT TAA TCT CTT GAA TTA AAT GGC ACT GTC CGT CTA TAG TGA GT
eGFP shRNA	AAG ACG TAA ACG GCC ACA AGT TCT CTT GAA ACT TGT GGC CGT TTA CGT CTA TAG TGA GT
Universal T7 primer	TAA TAC GAC TCA CTA TA

RT-qPCR primers	
Name	Sequence (5’-3’)
paraxis forward primer 1	CGC ATC TGA GCA CTA TCC TC
paraxis forward primer 2	AAC TAT TTC GCC ATC CTC AGC
paraxis forward primer 3	CCT CAG CTA GAA GGG TTT GC
paraxis reverse primer 1	CTT TTG TTA CCG CCT TCA GC

paraxis reverse primer 2	ATA AAG TTT CTG CTC AGC GG
paraxis reverse primer 3	CTC TTC GAG TCT CGG TTT CG
ITGB1 forward primer	TAC ACT TGA CTA CCC CTC TG
ITGB1 reverse primer	ACT TGA GTC TTT GTG ATG GC
ITGB forward primer	AGG ACA AGG ACT ACT TCG TC
ITGB reverse primer	ACT TTG GGT TCG TCA GTT TC
ITGB2 forward primer	CAC ATA CTC CCC CTT GTT AG
ITGB2 reverse primer	TAA GAG CAT CAA AAC CTC CC
GAPDH forward primer	GGA CCA AGT GCC AAG AAC TG
GAPDH reverse primer	GGA ATG CCA TAC CCG TCA G
18s rRNA forward primer	CTC AGG CTC CTA AAG GTT TCA T
18s rRNA reverse primer	CAT AAG TCC CAG CCC AAG ATA G

Colorimetric in-situ hybridization probe primers	
Name	Sequence (5’-3’)
ITGB1 forward primer	CGT CTC AGC ACA ATC AAG TC
ITGB1 reverse primer	ATA GGT GGG GTT CTG AAA CG
ITGB forward primer	TGT CGC TTC TCA AGA TGA AC
ITGB reverse primer	AAG GTA AAA GTG CAG TCG TC
ITGB2 forward primer	AAA AGG ATA TGG ATG TCG CG
ITGB2 reverse primer	GGC CAG CAT AAG TAG GAT TC

2.4 Antibodies

Name	Source	Type	Catalog
NvParaxis	Rabbit	Polyclonal	Custom Antibody
Integrin α4	Rabbit	Monoclonal	Cell Signaling Technology #8440
Integrin α5	Rabbit	Polyclonal	Cell Signaling Technology #4705
Integrin αV	Rabbit	Polyclonal	Cell Signaling Technology #4711
Integrin β1	Rabbit	Monoclonal	Cell Signaling Technology #9699
Integrin β3	Rabbit	Monoclonal	Cell Signaling Technology #13166
Integrin β4	Rabbit	Monoclonal	Cell Signaling Technology #14803
Integrin β5	Rabbit	Monoclonal	Cell Signaling Technology #3629
Anti Rabbit HRP	Goat	-	Cell Signaling Technology #13166
Anti Rabbit 568	Goat	Polyclonal	Invitrogen #A-11011
Phalloidin 488	-	-	Invitrogen #A12379
Anti-Digoxigenin-AP	Sheep	Polyclonal	Sigma, #11093274910

2.5 Enzymes and reagents

Name	Catalog
Direct-zol™ RNA Miniprep Plus Kit	Zymo Research #R2072
Zymoclean™ Gel DNA Recovery Kit	Zymo Research #D4002
AmpliScribe™ T7-Flash Transcription Kit	Lucigen ASF3507
iScript™ cDNA Synthesis Kit	Bio-Rad 1708890
Minute™ Detergent-Free Protein Extraction Kit	Invent Biotechnologies, Inc. SN-006
Zero Blunt™ TOPO™ PCR Cloning Kit	Thermo Scientific 450245

DirectPlate™ DH5-Alpha Chemically Competent Cells	Intact Genomics # 1013-12
PureYield™ Plasmid Miniprep System	Promega A1223
DIG RNA Labeling Mix	Roche #11277073910
MEGAscript™ T7 Transcription Kit	Invitrogen AM1334
MEGAscript™ SP6 Transcription Kit	Invitrogen AM1330
QIAshredder	Qiagen 79654
Klenow fragment	NEB M0212S
BCIP/NBT Color developing reagent	Promega S3771
SuperSignal™ West Dura Extended Duration Substrate	Thermo Scientific 34075
Blocking Reagent	Roche #46925300
Dextran, Texas Red; 3000 MW	Life Tech #D3328
FITC (5/6-fluorescein isothiocyanate)	Thermo Scientific 46425
Hoechst 33342	Invitrogen H1399
PerfeCTa SYBR™ Green FastMix Low ROX	Quanta bio 95074-012

2.6 Algorithms and software

Name	Source	Link
Nematostella Genome	SIMRBASE	https://simrbase.stowers.org/jb_pub/?data=data/starlet_pub
InvivoGen siRNA Wizard™	Web interface	https://www.invivogen.com/sirnawizard/design.php

EndoAtlas	He et al., 2023	http://simrbase.stowers.org:8888/
MAFFT version 7	Kato et al., 2019	https://mafft.cbrc.jp/alignment/server/
IQ-TREE	Trifinopoulos et al., 2016	http://iqtree.cibiv.univie.ac.at/
Design and Analysis (DA2) software	Thermo Fischer Scientific	Download

Methods

2.7 *Nematostella* husbandry - culture and spawning

Adult *Nematostella* polyps were housed in the aquatics facility at the Stowers Institute for Medical Research. Throughout development, the animals were reared in 12 parts per thousand (ppt) artificial saltwater (ASW) and regularly fed freshly hatched artemia once they had reached the primary polyp stage. Spawning groups were delivered daily to a light box, which was scheduled on a light and heat cycle. To induce spawning, this light and heat treatment was followed by a cold shock in the morning (~9 am), replacing half of the ASW with ASW at 17°C. Once the female bowls had released egg sacs, they were collected and dejellied by gentle oscillation in a solution of 4% L-cysteine in 12 ppt ASW (with 5 drops of 5 M NaOH in 25 mL solution) for 10 mins. Once the eggs had been dejellied and separated from the sacs, they were washed three times with ASW and held at 17°C till fertilized. Fertilization was performed within 2-3 hours of spawning by adding sperm water from the male bowls to the eggs.

Once fertilized, the embryos were cultured in petri dishes in an incubator setup. For a few days post-fertilization, the ASW media in the petri dish was cleaned and refreshed daily to remove unfertilized eggs or dying embryos. Depending on the temperature

chosen, the rate of development could be slowed down or sped up - incubation at 17°C allowed the planula stage to be attained five days post fertilization (dpf) while incubation at 24°C allowed the same stage to be attained at three days post fertilization.

2.8 Short-hairpin RNA (shRNA) - design and synthesis

An shRNA approach can be used to produce targeted gene-specific knockdowns in *Nematostella vectensis* embryos (He et al., 2018). shRNA's, once endogenously processed, mimic the precursor miRNA pathway that regulates target mRNAs by cleavage, involving the Dicer protein and the RNA-induced silencing complex.

The shRNA candidate sequences were selected using a web-based tool - Invivogen siRNA Wizard (Table 2.6). The entire coding sequence for the target gene was obtained from SIMRbase and provided as an input to the software with the motif length set to 19 bp. Two to four candidates for each gene were selected, ensuring that the overall GC content is between 40% and 55% and that BLAST searches against the *Nematostella* transcriptome return complete complementarity only for the targeted gene.

Production of the shRNA was based on in-vitro transcription (IVT) from a template DNA. The template DNA was produced by Klenow extension using a gene-specific reverse strand oligonucleotide and a universal forward oligonucleotide (that contains a T7 promoter sequence to enable IVT). The reverse strand oligonucleotide was designed following the template: 5′-AA-[19 bp shRNA candidate sequence]-TCTCTTGAA-[reverse complement candidate sequence]-TATAGTGAGT-3′. The 3' TATAGTGAGT sequence reverse complements the T7 sequence and enables the Klenow extension to generate a double-stranded DNA template. The nine-nucleotide TCTCTTGAA loop sequence is crucial for shRNA's to be recognized and cleaved by the dicer complex in vivo (Hill et al., 2022).

To prepare the shRNA *in vitro*, the DNA template was synthesized using a Klenow reaction mixture and 100uM forward and reverse oligonucleotides. The oligonucleotides were first denatured at 70°C for 2 minutes and then allowed to anneal at room

temperature (RT) for 5 minutes. The template extension reaction was carried out at 37°C for 30 minutes using NEB 2/2.1 buffer, dNTPs, and the annealed oligos. Once complete, the Klenow polymerase was deactivated by heating to 70°C, and the DNA template was ready for *in vitro* transcription. IVT was performed with the Ampliscribe T7-Flash Transcription kit, and the reaction was allowed to run for 2 hours at 37°C. The transcribed shRNA was precipitated with ice-cold 100% EtOH and purified using Direct-zol RNA MiniPrep Plus kit. Once eluted, it was quantified using a Nanodrop spectrophotometer and stored at -20°C for immediate use or - 70°C for long-term storage.

2.9 Microinjection

Unfertilized dejellied *Nematostella* eggs were plated in a Falcon plastic petri dish in 12 ppt ASW and microinjected using an Eppendorf FemtoJet 4x system. Glass capillary needles that were used for microinjections were pulled using the Sutter Micropipette puller P-87. The injection mix was composed of shRNA diluted to the appropriate concentration (with DEPC/nuclease free water) and FITC and Dextran Texas Red as short-term and long-term tracer dyes, respectively (1uL each in 5uL total injection mix). Eggs were fertilized post-injection and incubated at temperatures specified in the culture conditions. Fluorescence from the tracer dyes was used as a marker for injection, and non-fluorescent embryos were sorted and discarded soon after fertilization.

2.10 Electroporation

Electroporation is an alternative method of delivering shRNAs into the embryo that allows for a much larger amount of embryos to be simultaneously manipulated (Karabulut et al., 2019).

Unfertilized dejellied eggs and ASW amounting to 100uL were collected in a 1.5 mL microcentrifuge tube (MCT). The ASW was replaced with a solution of 15% Ficoll PM

400 (in ASW), and the eggs were resuspended in this solution. The suspended eggs were then transferred to clean 4 mm electroporation cuvettes, where the purified shRNA was added directly to a final concentration of 300 ng/uL. The solution was mixed gently by agitating the cuvette and then placed in the electroporation chamber. Conditions were set to 50V, 25 msec, 1 pulse, and electroporation was initiated. Immediately after electroporation, sperm water was added to the cuvette, and the eggs were transferred to a petri dish for incubation at the appropriate temperature. After successful electroporation, the eggs displayed an altered morphology that recovered within 30-60 minutes when observed under a stereomicroscope. Embryos that have undergone electroporation were seen to be developmentally more fragile than wild-type embryos, and it was optimum to rear them at lower temperatures with regular water changes to remove debris.

2.11 RNA extraction and cDNA synthesis

Embryos at the appropriate developmental stage were collected, and excess ASW was carefully aspirated. Trizol reagent was then added for lysis, and if the samples were to be processed at a later stage, they were flash-frozen in liquid nitrogen and stored at -70°C. For immediate processing, an electric pestle was used to homogenize the sample for 60 seconds. The resulting lysate was transferred to a QIAshredder column and centrifuged. An equal volume of 100% EtOH was added to the flowthrough. This mixture was then purified with the Direct-zol RNA MiniPrep Plus kit. An additional DNAase I digestion step was performed directly on the spin column to remove any genomic DNA contamination. The purified RNA was eluted in nuclease free water, quantified using a Nanodrop spectrophotometer, and immediately used as a template for cDNA synthesis.

cDNA synthesis was carried out using the iScript cDNA synthesis kit, which employs an RNAase H+ reverse transcriptase for first-strand cDNA synthesis and degradation of the residual RNA template. This kit utilizes a combination of oligo(dT) and random hexamer primers to ensure comprehensive transcript coverage and features a broad linear dynamic range to minimize bias in downstream gene expression analysis.

2.12 Reverse transcription - quantitative polymerase chain reaction (RT-qPCR) assay

An RT-qPCR assay was used as a measure of quantifying gene expression based on mRNA levels.
Primers were designed to generate amplicons of 100–200 bp while minimizing the potential for genomic DNA contamination. This was achieved by selecting primer pairs that flanked large introns or spanned exon-intron junctions. A standard dilution RT-qPCR was performed to obtain an estimate of the linear range of amplification and optimize primer selection when multiple candidates were being tested. Wild-type 3 days post-fertilization cDNA was used as a template, with serial five-fold dilutions prepared in technical triplicates. Reactions were set up using a Tecan WorkStation automated liquid handler in a 384-well plate format with Low ROX Perfecta SYBR Green mix, and the plate was run on QuantStudio 7 Pro System. The data was analyzed using the Design and Analysis software, and a standard dilution curve was constructed. Only primers that produce a single melt curve profile were selected for further analysis.

The same setup was followed for experimental runs, with cDNA templates selected based on the target assay. Negative controls included a no-template water control and a no-reverse transcriptase control to ensure minimum genomic DNA contamination. Gene expression changes were quantified using the $\Delta\Delta C_t$ (delta-delta C_t) method (Fig 2.1) (Schmittgen and Livak 2008). *NvGAPDH* was used as an internal housekeeping control for the relative fold change as it had been previously validated for RT-qPCR studies in *Nematostella* (He et al., 2018). *Nv18s rRNA* was also used as another internal housekeeping gene to ensure that the levels of *GAPDH* were not fluctuating in the assay.

$$\Delta C_t = \{C_t(\text{Target gene}) - C_t(\text{Internal control})\}_{\text{sample}}$$

$$\Delta\Delta C_t = \Delta C_t(\text{treated sample}) - \Delta C_t(\text{control sample})$$

$$\text{Fold change} = 2^{-\Delta\Delta C_t}$$

$$\text{Fold change reduction} = -1/2^{-\Delta\Delta C_t}$$

Fig 2.1. Calculation for the fold change in transcript levels (Schmittgen and Livak 2008). C_t is the cycle threshold for the amplification step. The internal control is the housekeeping gene used for calculating the relative abundance (ΔC_t)

Analysis was performed using error propagation and generation of 95% confidence intervals. For the reduction in transcript levels of *paraxis* (Section 3.3), the experiment was performed with samples composed of ~200-300 embryos pooled into three technical replicates for each setup (un-electroporated WT, eGFP shRNA and *paraxis* shRNA). 95% confidence intervals were generated for the $\Delta\Delta C_t$ value using standard error propagation (from the technical replicates) and then these limits were exponentiated to the fold change to generate the error bars. Only one biological replicate is shown for this assay as it was a verification of knockdown assay and was performed once with multiple different shRNAs combinations to identify the one which produced the greatest reduction of transcript levels.

For the relative levels of integrin transcripts during the dissociation reaggregation assay (section 3.11), the experiment was performed with five independent biological replicates, each with three technical replicates. The technical replicates were used to obtain a mean C_t which was used to calculate the ΔC_t and $\Delta\Delta C_t$ for each biological replicate. Using the standard deviation of the five biological replicates a 95% confidence interval was generated for the $\Delta\Delta C_t$ values and then these limits were exponentiated to the fold change to generate the error bars.

2.13 Phalloidin staining and immunofluorescence

Phalloidin staining was used to visualize the actin filament architecture and observe the segmental organization at the planula stage of embryonic development.
Planula-stage embryos were collected and fixed in 4% paraformaldehyde (12 ppt ASW) with gentle shaking on the nutator for 1 hour at room temperature. After fixation, the embryos were washed five times with 0.2% PBSTw (0.2% Tween 20 in 1x Phosphate buffered saline solution) for 5-10 minutes each. They were then incubated with

Phalloidin Alexa Fluor 488 (1:400) and Hoescht 33342 (1:1000) in 0.2% PBSTw and kept shaking overnight at 4°C. Post staining, the embryos were washed again five times with 0.2% PBSTw and then immersed in SCALE A2 (composed of 75% glycerol and 25% 8M urea).

For immunostaining with the custom Paraxis antibody and the commercial integrin antibodies, a modified version of the previous protocol was used with additional steps to reduce non-specific binding. Embryos were collected as usual; however, for polyp stage embryos, a relaxation step was performed by gradual addition of 7% $MgCl_2$ into the ASW media to ensure that the polyps do not contract and alter their morphology upon addition of the fixative. Multiple different fixation protocols were tested for each set of antibodies; however, the one that consistently produced better staining (with less background signal) was NAFA fixation. The NAFA fixation solution comprised 10% pH 7.5 HEPES buffer, 25% of 16% PFA, 5% 0.5M EGTA, and 5% of 95% formic acid in 12 ppt ASW. The embryos were incubated in the fixative solution for 1 hour at room temperature with gentle shaking, followed by five washes with 0.2% PBSTw, each for 5 minutes. If needed, for long-term storage, embryos were gradually dehydrated through a methanol gradient until reaching 100% MeOH and stored at -20°C. Prior to use, embryos were rehydrated stepwise by sequential washes back into 0.2% PBSTw. This was followed by a blocking step wherein the embryos were incubated in the blocking solution (5% goat serum, 1% BSA, 10% Roche whole blocking reagent, 10% DMSO in 0.2% PBSTw) for 1 hour at room temperature. The primary antibody (1:1000 dilution for Paraxis; 1:200 dilution for α4, αV, α5, β1, β3, β4; 1:1200 dilution for β5) was then added to the blocking solution (composition same as before except 0.1% DMSO instead of 10%) and the embryos were incubated in this overnight at 4°C. The next day, the samples were washed five times for 20 minutes each with 0.2% PBSTw. Then, the secondary antibody (Goat anti-rabbit 568) was diluted 1:500 in 0.2% PBSTw along with 1:1000 of Hoescht (for nuclear counterstaining), and the embryos were incubated with it overnight at 4°C. The next day, the stained embryos were washed five times with 0.2% PBSTw for 20 minutes each and then immersed in SCALEA2 for imaging. The embryos in SCALEA2 could be stored at 4°C for a few days, till imaged.

2.14 Western Blot

Whole protein lysates were obtained using the Minute Detergent-Free Protein Extraction Kit. Embryos at the appropriate planula stage were collected, and the ASW media was replaced with 50uL Buffer A (adjusted as needed based on embryo quantity). The samples were transferred to the extraction column and homogenized using a handheld pestle. An equal volume of buffer B was then added, and the homogenization step was repeated for 30-60 seconds. This lysate was centrifuged, and the resulting flowthrough, containing the whole protein extract, was quantified using a Nanodrop or a Qubit assay, and either used immediately or flash frozen and stored at -70°C for later use.

To prepare the samples for loading onto the protein gel, 10 ug of total protein extract solution was mixed with NuPage sample loading buffer and beta-mercaptoethanol. This solution was heated to 70°C for 10 minutes and then loaded on the NuPage 4-12% Bis-Tris gel and run with 1x MES SDS running buffer. Following electrophoresis, protein transfer was performed using a PVDF membrane, pre-wetted with methanol, and assembled in a transfer chamber with blotting pads and filter papers. The transfer was carried out at 30V for 1 hour. After the transfer was complete, the membrane was blocked overnight at 4°C with 1% blotting grade blocker in 0.1% PBSTw. Commercial integrin primary antibodies (#4749, Cell Signalling Technology) were diluted in the blocking buffer (1:1000 dilution) and incubated with the PVDF membrane for 1 hour at room temperature. The membrane was then washed five times with 0.1% PBSTw for 5 minutes each. Subsequently, the membrane was incubated with a 1:2000 dilution of HRP-linked anti-rabbit IgG secondary antibody (#7074, Cell Signaling Technology) in blocking buffer for 1 hour at room temperature, followed by another five washes with 0.1% PBSTw for 5 minutes each. Protein detection was performed using SuperSignal West Dura chemiluminescent substrate, and the signal was imaged using a G:Box imaging system.

2.15 Colorimetric in-situ hybridization

Colorimetric in-situ hybridization enables visualization of gene expression in a spatial manner within fixed developing embryos by utilizing labeled probes that specifically bind to the target mRNA transcript.

Cloning and probe synthesis:

To generate the probes, an appropriate set of forward and reverse primers were designed for each transcript (Table 2.3) such that they amplified ~2kB fragments from the coding sequence of genes obtained from SIMRbase. The fragments of interest were amplified using PCR with Q5 polymerase, utilizing 3 dpf cDNA as the template. The PCR product was band purified and cloned into a pCR -Blunt II-TOPO vector using the Zero Blunt TOPO PCR cloning kit. The ligation mixture was then chemically transformed into *E. coli* One Shot cells, which were plated on LB agar containing kanamycin and X-gal for blue-white screening. After overnight incubation at 37°C, white colonies were obtained, indicative of successful insertion into the lacZ site. Five white colonies were picked and cultured in LB liquid media with kanamycin for 24 hours at 32°C. Plasmids were extracted using the PureYield Plasmid Miniprep System and sent for sequencing with M13 forward and reverse primers flanking the insertion site. Once the inserts were confirmed via sequencing, M13 forward and reverse primers were again used to amplify the entire fragment, which contains the blunt PCR product flanked on either side by a T7 promoter and SP6 promoter (present in the backbone vector). Sequencing confirmed that all the PCR products were cloned in an orientation allowing for antisense strand transcription if the SP6 promoter was used and sense strand transcription (for negative control probe) if the T7 promoter was used. Therefore, in-vitro transcription was performed using the MEGAscript SP6 Transcription Kit/ MEGAscript T7 Transcription Kit (depending on which strand was being synthesized) with the amplified fragment as the template and a DIG-RNA labeling mix. The IVT reaction was allowed to run overnight at 37°C. The following day, TURBO DNase I was added to the reaction, and the mixture was incubated at 37°C for 30 minutes to degrade any residual template. The RNA product was then precipitated with 100% ethanol and purified using

the Direct-zol RNA MiniPrep Plus Kit. The purified RNA probes were eluted in DEPC water, quantified using a Nanodrop spectrophotometer, and stored at -20°C until future use. The same probe could be reused for up to three in-situ hybridizations, with signal improvement and non-specific background gradually decreasing with each use.

In-situ hybridization (ISH):

For the colorimetric ISH, embryos were collected as usual, with an additional relaxation step involving gradual addition of 7% MgCl2 for polyp stage embryos. All the samples were fixed in 4% PFA (12 ppt ASW) for 1 hour at room temperature under gentle shaking. They were then washed five times with 0.1% PBSTw for 5 minutes each at room temperature and then gradually dehydrated through a methanol gradient until reaching 100% MeOH and stored at -20°C (at least overnight, if needed could be kept in 100% MeOH for a few weeks). Once ready to be processed, samples were rehydrated back into 0.1% PBSTw and washed five times. Post rehydration and washing, a critical digestion step was performed by incubation with 60ug/mL of Proteinase K for exactly 2 minutes without shaking. The samples were then immediately post-fixed in 4% PFA (0.1% PBSTw) for 1 hour at room temperature. This was followed by five washes with 0.1% PBSTw for 5 minutes each. The samples were then incubated with a 1:1 mixture of pre-Hybe and 0.1% PBSTw for 10 minutes at room temperature and then with 100% pre-Hybe for 10 minutes at room temperature (pre-Hybe composition – 50% deionized formamide, 25% of 20X SCC pH 7, 100ug/mL Heparin, 1% SDS, 0.1% Tween 20; in DEPC). As a blocking step, the samples were incubated overnight in 100% Hybe (hybridization) solution at 60°C (Hybe composition – same as pre-Hybe with addition of 5% dextran sulphate and 5mg/mL torula yeast RNA).

The next day, the DIG-labeled probes were prepared to a concentration of 1ng/uL in Hybe solution and allowed to denature at 85°C for 10 minutes, followed by cooling on ice. The samples were then incubated in the Hybe solution containing the labeled probes for 72 hours at 60°C. After this step, the probes can be stored for reuse in the Hybe solution itself, at -20°C. Following hybridization, the samples were washed for 30 mins each at 60°C with a gradient of 2X SSC (going from 25%, 50%, 75% - in pre-

Hybe) till they were in 100% 2X SSC. This was followed by more washes of 20 minutes each at 60°C with 0.2X SSC, 0.1X SSC, and a 1:1 mixture of 0.1X SSC and 0.1% PBSTw before they were brought back into complete 0.1% PBSTw. After two more washes with 0.1% PBSTw for 5 minutes each at room temperature, the samples were incubated with MAB buffer for 5 minutes at room temperature. Finally, the samples were blocked with a blocking solution (5% sheep serum, 1% Roche whole blocking reagent, 10% DMSO; in MAB) for 2 hours at room temperature. At the same time, the anti-DIG AP antibody (dilution 1:2000) was also incubated in the blocking solution (same as previous with 0.1% DMSO instead of 10%) for pre-absorption. Once blocking was complete, the samples were incubated in the pre-absorbed anti-DIG antibody solution overnight at 4°C.

After incubation, the samples were washed five times with 0.1% PBSTw over a total duration of 6-7 hours. As the first step to developing the signal, samples were washed once in AP buffer without $MgCl_2$ and twice in AP buffer with $MgCl_2$ for 10 minutes each, at room temperature (AP buffer composition - 100mM NaCl, 50mM $MgCl_2$, 100 mM Tris pH 9.5, 0.1% Tween-20; in MilliQ). The samples were then added to a development solution (3.3uL/mL NBT and 3.3uL/mL BCIP in AP buffer) and kept covered from light for signal development. The signal was checked every 15-30 minutes, and the reaction was stopped by transferring the samples to 0.1% PBSTw. The samples were then washed twice in 0.1% PBSTw and fixed in 4% PFA for 30 minutes at room temperature. Post fixation, they were transferred to 100% EtOH and then to 50% EtOH, and then they were washed twice with 0.1% PBSTw. Finally, they were immersed in SCALEA2 and stored at 4°C till imaged.

2.16 Dissociation reaggregation assay

For this assay, mid gastrula stage embryos (28 HPF at 17°C) were used. Embryos were collected in a 1.5 mL Eppendorf tube in ~100 uL 12 ppt ASW. They were then triturated with a 200uL pipette for 1 minute and 45 seconds to achieve uniform dissociation, after which 1mL of 12ppt ASW was added to the resulting cell suspension. The suspension

was passed through a 40u filter to remove large cellular debris and further diluted with an additional 1mL ASW before being transferred to a 24-well plate. The plate containing the suspension was incubated at 17°C, and the cellular aggregates were either harvested for RT-qPCR at 18 hours post-dissociation or allowed to continue aggregating for imaging, as needed.

2.17 Image acquisition

Images for the immunofluorescence assays and for the samples stained with phalloidin were acquired on the Andor DragonFly 200 Spinning Disc confocal microscope under 250X magnification. The Fusion HD software was used for image acquisition.

Images for the colorimetric in-situ hybridization were obtained on the Axiovert MOT 462 microscope with the Micro-Manager software. Dissociation and reaggregation assay images were obtained on the Leica MN165 BW.

All acquired images were processed in Fiji to adjust the brightness/contrast, create maximum intensity projections from multiple z slices, and for scale bar addition.

2.18 Phylogenetic analyses

Amino acid sequences for the various integrin subunits from different species were obtained from SIMRbase (for *Nematostella* sequences) and Uniprot (for all other species) (Table 2.2).

These sequences were used to generate a FASTA file in AliView and were then aligned on MAFFT version 7 using a BLOSUM 62 substitution scoring matrix. Extremely long gappy sequences were manually trimmed to allow for better alignment and tree construction. The aligned sequences were used to construct a maximum likelihood phylogenetic tree using the IQ-TREE webserver with the substitution model being auto selected and a perturbation strength of 0.5 (Trifinopoulos et al., 2016). Statistical

support was estimated by performing 1000 bootstrap alignment replicates. The tree was visualized using the FigTree software and the sequences from the sponge species *Geodia cydonium* were chosen as an outgroup to root the tree. The bootstrap strength as a percentage (out of 100) is represented by the node labels.

2.19 Text Refinements

ChatGPT, a large language model, was used for the refinement of certain sections of this text. However, its use was restricted to editing text that had already been written, and it was not used to generate new text or content.

CHAPTER THREE - RESULTS

3.1 Paraxis ortholog in *Nematostella vectensis*

The *Nematostella* ortholog of the *paraxis* gene (renamed TCF-15 in mice and humans) was previously mapped and annotated independently in publicly available genomes by the Wellcome Sanger Institute and the Stowers Institute for Medical Research (SIMR). Its preliminary identification was based on computational predictions. To further characterize it, a comparison of its amino acid sequence and 3D structure was conducted.

Multiple sequence alignment of the amino acid residues of the bHLH domain of the putative Paraxis protein in *Nematostella* and TCF15 revealed a high degree of conservation (Fig. 3.1). The basic region, responsible for the binding of the protein to specific sites in the genome, was nearly identical. Outside of the bHLH domain, the flanking N-terminus region displayed no level of sequence conservation, while the C-terminus had a small block of amino acid similarity (not displayed).

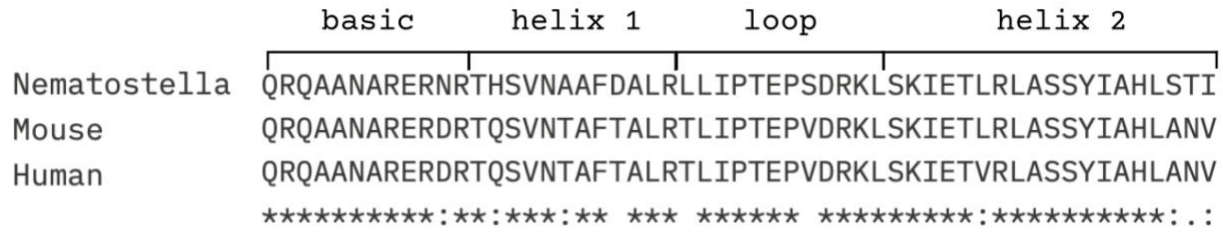


Fig. 3.1 Clustal multiple sequence alignment of the bHLH domain between *Nematostella*, mouse, and human. Identical residues are represented by an asterisk (*) below the alignment, residues with strongly similar properties are represented by a colon (:), and conservation between groups with weakly similar properties is indicated by a period (.). The bHLH domain has been annotated based on the mouse TCF15. UniProt Accession IDs: Nematostella - A0A1T4JGW9, Mouse - Q60756, Human - Q12870

The predicted 3-dimensional structure of the *Nematostella* Paraxis protein also displayed high structural similarity within the bHLH functional domain (the region

colored in blue - also indicative of very high model confidence) to its vertebrate ortholog (Fig. 3.2), while the N- and C-terminus remained relatively unstructured.

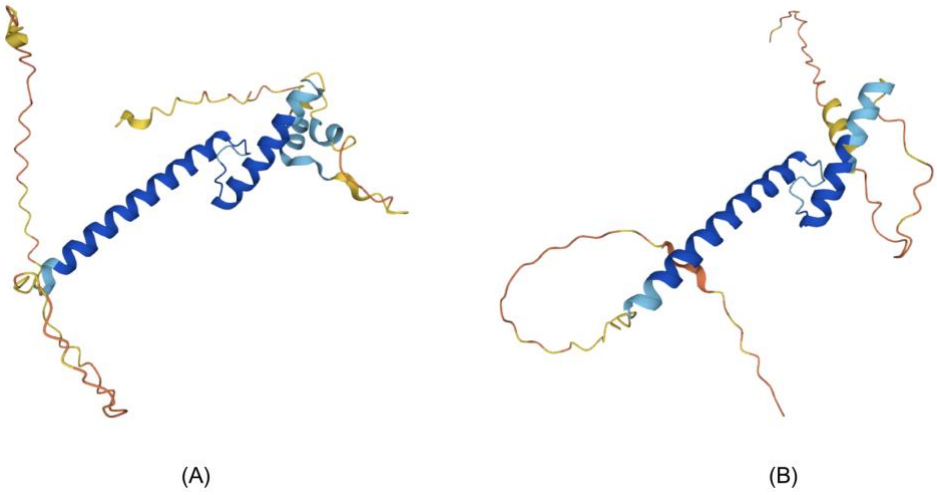


Fig. 3.2 Alphafold2 modeling of the predicted Paraxis protein structure between (A) Mouse and (B) *Nematostella*. The bHLH domain from both organisms was highly structured and similar in arrangement in 3D space.

An examination of previously generated bulk RNA-sequencing data found that the *paraxis* transcript was not maternally deposited, and its expression was absent at the blastula stage. Early *paraxis* transcription began at the post gastrula stage and peaked as the embryo progressed into the planula stage (Fig. 3.3).

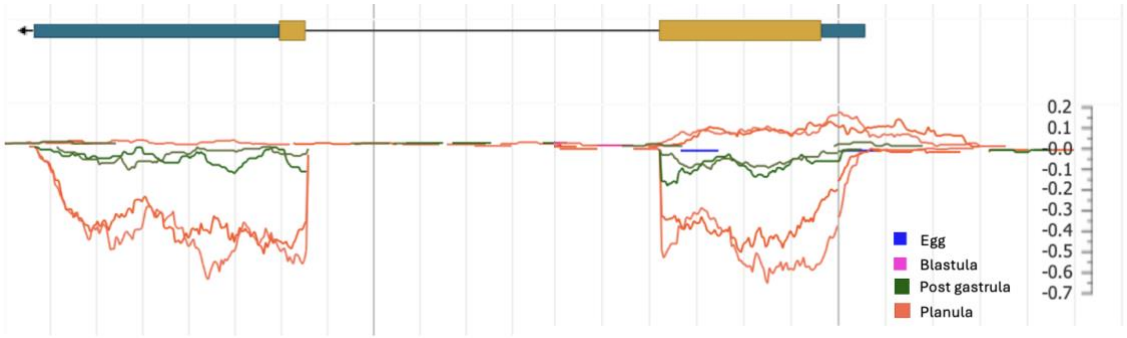


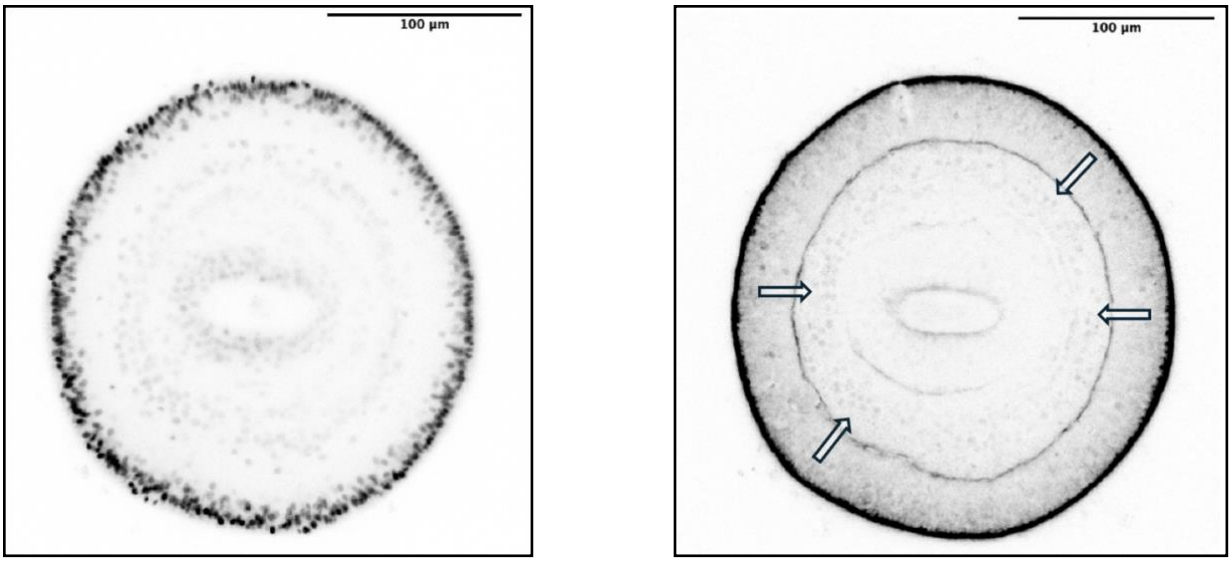
Fig. 3.3 RNA sequencing data profiled beneath a genomic map representing the *Nematostella* locus. The top graphic represents the *Nematostella* paraxis gene, where the ochre-colored boxes represent exons, the blue boxes correspond to untranslated regions

(UTRs), and the black line represents the intron. The organization of the *Paraxis/Tcf15* gene was simple and, for most species, was composed of only two exons separated by a single intron. The RNA sequencing data is presented in the bottom half of the figure, with a color-coded legend detailing the developmental time from which the RNA was extracted.

3.2 Paraxis localization in the developing embryo

Immunohistochemistry (IHC) was performed with a custom antibody generated against *Nematostella* Paraxis. It was expected that Paraxis, being a transcription factor, would be localized primarily in the nucleus. A nuclear-localized signal that co-localized with Hoechst staining in the cells of endomesoderm was observed; however, there was significant staining on the outside of the embryo and between the surface-located ectodermal cells (Fig. 3.4).

To address this background staining issue, multiple different staining protocols were tested including heat induced antigen retrieval, paraformaldehyde fixation, formic acid/glutaraldehyde fixation. However, in each case, there was a significant amount of background or non specific staining being detected in the outer epithelial cells and at cell boundaries. Different concentrations of the primary antibody and different types of secondary antibodies were also tested but produced no significant difference in the staining pattern.



(A) (B)

Fig. 3.4 Optical Z sections through 3 dpf embryos used in immunostaining experiments. (A) - The results of Hoechst staining, which marks the nuclei throughout the embryo. (B) - The results of immunostaining with the Paraxis antibody. The arrows mark examples of punctate nuclear staining in the endomesoderm layer.

3.3 Verification of *paraxis* shRNA knockdown

Electroporation and microinjection are both well-studied and robust methods of introducing shRNAs in the embryo to knockdown translation of specific target genes (Karabulut et al., 2019). Electroporation of shRNAs against the *paraxis* transcript was chosen as a method as it would allow for a larger number of affected embryos to be obtained for cDNA preparation.

RT-qPCR assays were designed and performed to test whether the shRNA targeted against the *paraxis* transcript were indeed reducing its expression in the embryo.

Multiple sets of primers were designed, flanking the intronic sequence of *paraxis,* and tested for their specificity in a standard PCR reaction to verify that single amplicons were being produced. Sequencing of these PCR products confirmed that the primers

were amplifying regions corresponding to the *paraxis* transcript. A standard dilution qPCR was performed, and the optimal primer pair was chosen for further qPCR experiments (Fig 3.5).

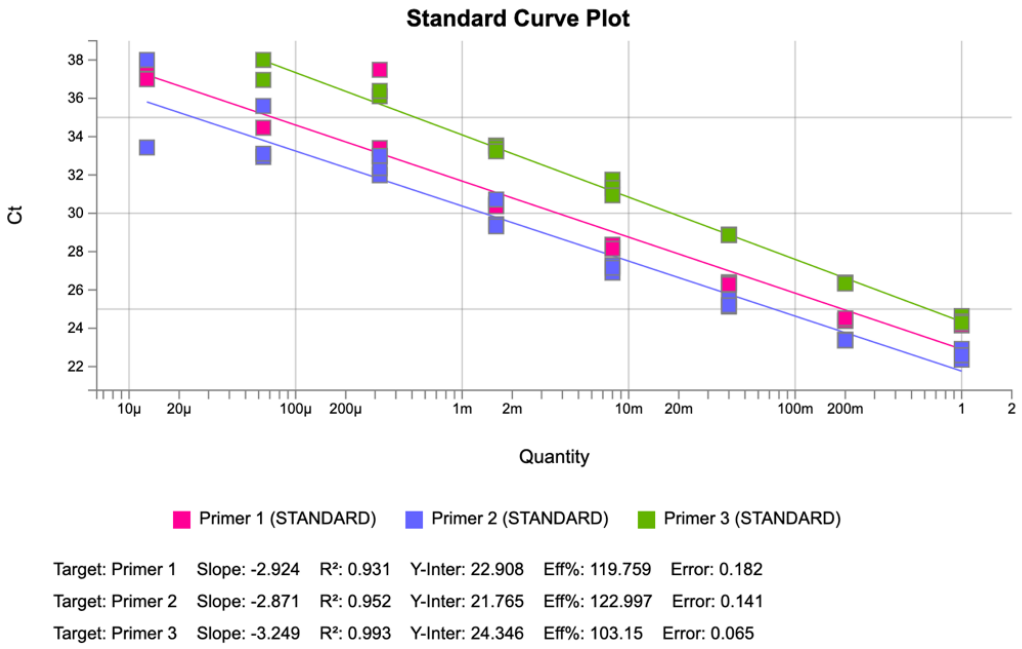


Fig. 3.5 Quantification of PCR primers designed against the *Nematostella paraxis* transcript. A standard curve generated by qPCR, using serial dilutions of wild-type cDNA, was used to determine which primer set was optimal for RT-PCR analysis. Three primer pairs were tested, and primer set number 3 was chosen for further experiments since it had the least error and its efficiency was the closest to 100%.

Previously published and validated housekeeping genes, *GAPDH* and *18s rRNA* were used as internal controls to provide relative transcript abundance (He et al., 2023).

RT-qPCR analysis (Fig. 3.6) using the double delta C_t method exhibited a 1.8-fold reduction in the *paraxis* transcript level in embryos that were electroporated with the appropriate shRNA compared to WT embryos, indicating that the shRNA was successful in reducing levels of the transcript. In the eGFP electroporated embryos, a modest increase in *paraxis* transcript abundance was observed, but this small change was not significant and could be considered near wild-type levels.

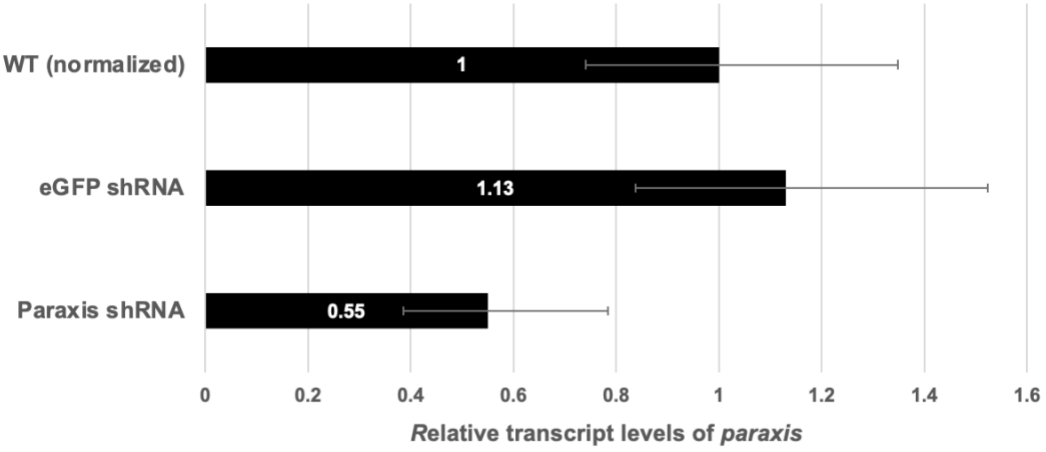


Fig. 3.6 The Results of an RT-qPCR assay performed using RNA extracted from 3 dpf embryos electroporated with shRNAs against *Nematostella paraxis* or eGFP (negative control), and un-electroporated WT embryos. The RT-qPCR data was analyzed using the double delta C_t method and the error bars were generated based on 95% confidence intervals (section 2.12 – three technical replicates consisting of 200-300 pooled embryos for each group). The level of transcripts for *paraxis* were reduced compared to wild-type expression when electroporated with shRNA against *paraxis*.

It is important to note that while the transcript level is being reduced, it has not been completely eliminated. Despite the reduced level, Paraxis, as a transcriptional factor, may still retain its ability to activate downstream signaling cascades. At the time of this experiment, this was proposed as a potential explanation for the low penetrance of the phenotype observed in the previous unpublished study.

3.4 Troubleshooting the phenotype

The previous unpublished study reported that, on average, 42% of embryos injected with *paraxis* shRNA exhibited a loss of segment boundaries and disorganized cellular morphology. This was statistically significant from the negative controls (eGFP shRNA injected embryos) and wild type embryos for which 10%(N=1/10) and 4%(N=1/25) respectively exhibited segmental deformities. However, a major concern was the

variability of the phenotype in the *paraxis* shRNA injected embryos, with some exhibiting fused segments, others displaying partial segment boundaries, and some showing severe developmental defects.

In the current setup, while attempting to replicate the phenotype of the *paraxis* shRNA knockdown, it was observed that there was a significant amount of variation in the negative control as well as in the wild type embryos (Fig. 3.7). In the current trials, 37.5% (N=9/24) of eGFP shRNA injected embryos showed segmental deformities that included incomplete segment boundaries and partial fusions. The *paraxis* shRNA injected animals also displayed a similar percentage (42%, N=6/14) and phenotype of segment boundary deformities. There are multiple variables that can contribute to this kind of variation and to understand them, each was systematically studied.

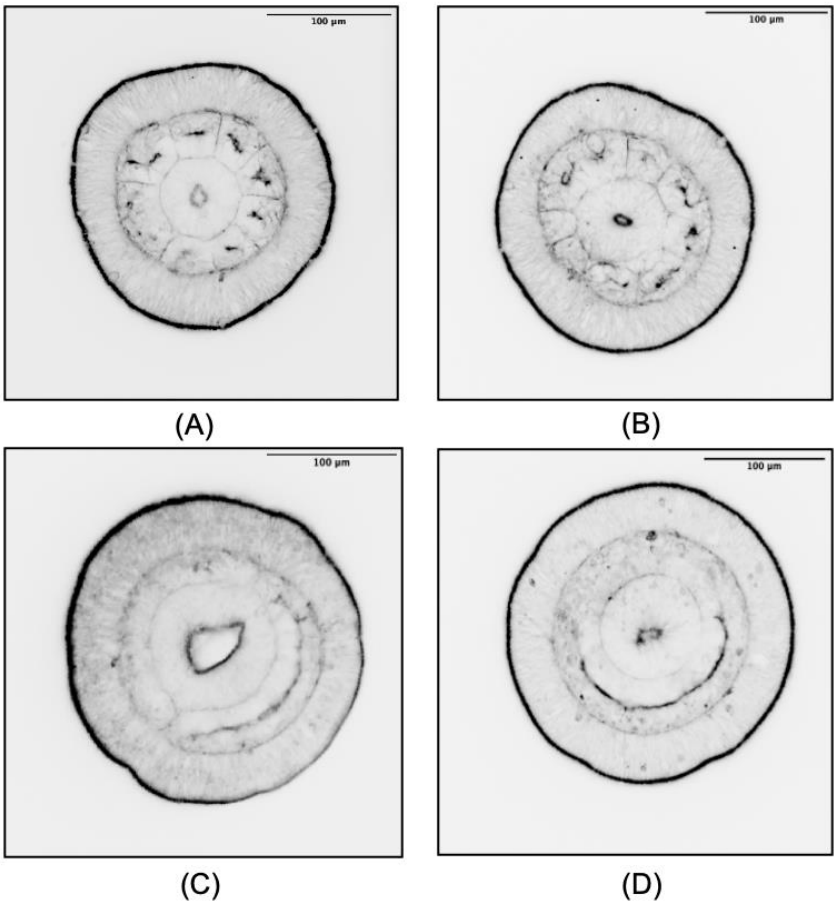


Fig. 3.7 Optical cross-sectional images on a selection of wild type embryos (3 dpf) stained with phalloidin, marking the actin filaments. From this analysis, three types of

segmental defects were observed amongst the wild embryos. (A) representative of a ‘normal’ wild type embryos with eight distinct segments. Segmental defects ranged from (B) incomplete and partial fusion of some segments, (C) complete fusion and (D) total absence of segmental boundaries.

To generate embryos for experiments, adults are induced to spawn and later in the day, egg sacs are collected, de-jellied and then combined with sperm to produce fertilized embryos. However, both eggs and sperm have limited viability which could result in substandard embryos when they are too ‘old’. A comparison between early and late fertilized embryos when morphologically examined at 3 dpf revealed that late fertilization resulted in a significantly higher deformity rate (78%, N=14/18); however, early fertilized embryos still exhibited a substantial proportion of deformities (41%, N=10/17) but at a lower prevalence.

Another possible variable was the process of removing the jelly that surrounds the eggs. This chemical process involves some degree of agitation which could potentially alter maternal deposited transcripts and shift proteins involved in the embryo’s first cleavage when activated. However, 33% (N=4/12) of non de-jellied embryos exhibited similar segmental deformities suggesting that the process of de-jellying was not an underlying cause of deformities.

A third possible source of segmental defect in wild type embryos was temperature. In all the previous trials, embryos were reared at room temperature (22°C), which is identical to the rearing temperature of the reproducing adults. To eliminate temperature as a factor and provide the optimum condition, de-jellied, wild type embryos were fertilized early, reared in a 24°C incubator and examined at 3 dpf. Even with reduced handling and an early fertilization time, these embryos displayed 25% (N=8/31) segmental deformities.

These observations indicated that there is currently a significantly higher basal frequency of segment deformities in wild type embryos which confounds the analysis of the *paraxis* shRNA experiment and needs to be kept in consideration for all future experiments.

3.5 Paraxis mutant analysis

Given that the *paraxis* shRNA knockdown reduced but did not eliminate the expression of *paraxis* and that there was segmental variation in untreated, wild-type embryos, it became necessary to examine embryos in which *paraxis* had been mutated through CRISPR-Cas9 mediated deletion. Multiple heterozygous animals were identified in which a single base pair insertion or deletion was introduced within the first coding exon of *paraxis* resulting in a frameshift that produced a premature stop codon prior to the bHLH region.

F2 offsprings of heterozygous mutant crosses were obtained and phenotyped at 3 dpf and sent for sequencing. The *paraxis* gene appeared to roughly segregate in a Mendelian fashion, with the F2 heterozygous cross (N=54) producing 18% wild-type embryos, 59% heterozygous embryos, and 22% homozygous mutant embryos. Of the 12 embryos genotyped as homozygous mutants for *paraxis*, 50% (6/12) displayed a normal segmentation pattern with eight distinct endomesodermal segments (Fig. 3.8). Embryos with deformed segment boundaries were present in all the genotypes, including the wild type, heterozygous, and homozygous mutant background (Fig. 3.9). This was consistent with the previous findings of naturally occurring segmentation defects.

Additional F2 offsprings, genotyped as homozygous mutants for *paraxis,* when allowed to mature to an adult stage, did not display any overt morphological phenotype. Moreover, crosses between these homozygous mutants produced viable normal-looking embryos, ruling out any requirement for maternally deposited transcripts.

These observations conclusively demonstrated that Paraxis, at least in a non-redundant capacity, does not appear to play a direct role in the formation of segment boundaries in *Nematostella*, as homozygous *paraxis* mutants developed eight fully formed segments with intact boundaries.

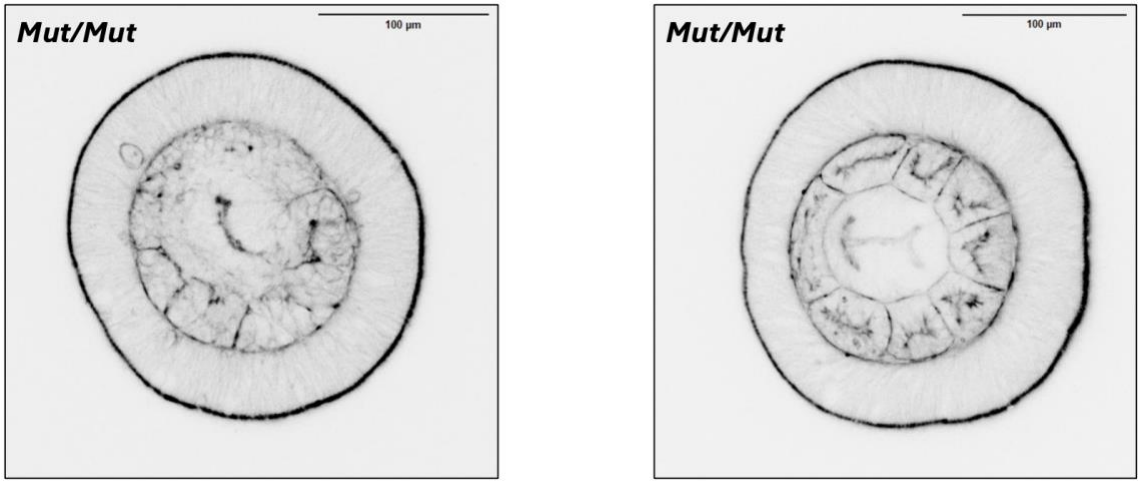


Fig. 3.8 Comparison of *paraxis* homozygous mutant embryos at 3 dpf stained with phalloidin to visualize actin architecture. (A) Incomplete segment boundaries and fused segments (N=6/12) (B) Eight complete segments and fully formed segment boundaries (N=6/12). The incomplete segmentation phenotype is not due to the loss of Paraxis function as half of the homozygous mutant embryos display normal segment formation while the other half display incomplete segment boundaries and fused segments.

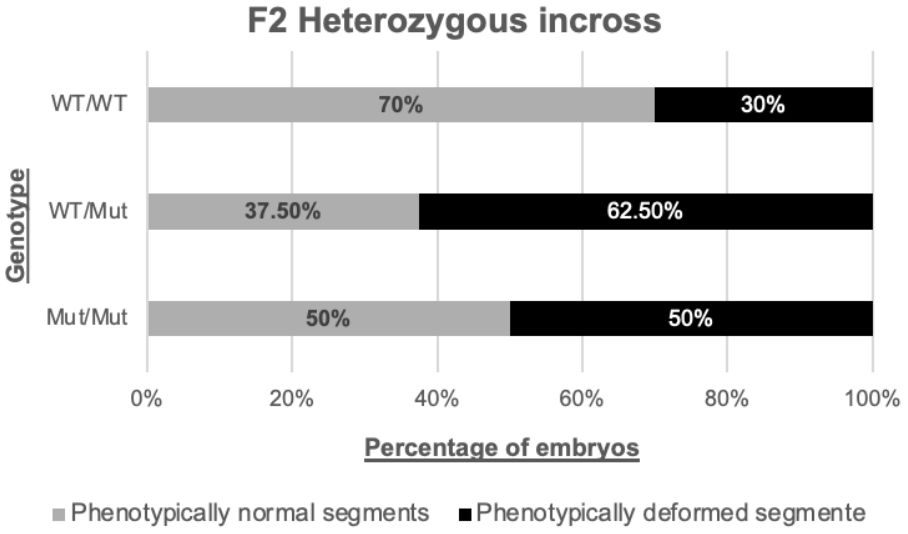


Fig. 3.9 Graph summarizing the outcome of an F2 heterozygous incross for the *paraxis* mutation. A portion of all embryos at 3 dpf, regardless of genotype, present segmentation defects. The embryos examined are wild-type (WT/WT N=10), heterozygous (WT/Mut N=32), and homozygous (Mut/Mut N=12).

3.6 The integrin subunit repertoire in *Nematostella vectensis*

The lack of an obvious and reproducible segmental defect in *paraxis* mutants was surprising. Still, it could be explained by unknown compensation from other genes whose products are involved in segment formation during development.
There is evidence that the vertebrate Paraxis protein regulates the expression of integrins and the deposition of fibronectin around the somites (Rowton et al., 2013). To date, an ortholog of fibronectin has not been identified in the *Nematostella* genome; however, previous studies have reported the existence of integrin subunits in *Nematostella*, with two alpha and four beta subunits being documented (Gong et al., 2014). Certain aspects of integrin signaling are dependent on interactions with fibronectin, but not all integrins are dependent on this one component of the ECM for their function. To understand if there were similar roles for integrins in *Nematostella* and uncover a correlation with the process of segmentation, integrin genes in *Nematostella* were identified and characterized.

A protein BLAST search of the *Nematostella* genome (SIMRbase) revealed a greater diversity of subunits than previously reported, with the identification of genes encoding for four alpha and four beta subunits. When reverse blasted, this set generated hits only within themselves (E value cutoff - 0.1, explored till E = 10) and, therefore appears to encompass the entire repertoire of subunits, at least for the current genome. In addition to this set, there was another subunit that was annotated as a truncated alpha subunit, however, this did not show up in any of the BLAST analyses (UniProt - A7RPV4). When a cross-species BLAST using these putative *Nematostella* candidates was performed, with species ranging from human, mouse, zebrafish, xenopus, medaka, lamprey, lancelet, *Drosophila* and *C. elegans*, the top hits were consistently integrin subunits.

Protein domain analysis performed by InterPro revealed that these proposed subunits consisted of canonical domains found in vertebrate integrins (Sebé-Pedrós et al., 2010). For all the alpha subunits, this included the signal peptide, FG-GAP repeats, non-cytoplasmic regions, a transmembrane domain and a cytoplasmic tail. Three of the FG-GAP repeats also contained intact DXD/NXD/NXXXD cation binding motifs (Fig. 3.10).

The diagnostic feature of integrin alpha subunit - a KXGFFXR motif, present in the cytoplasmic tail and interacting with the beta integrin subunit is also found in all the proposed *Nematostella* subunits (Fig. 3.10) (Sebé-Pedrós et al., 2010).

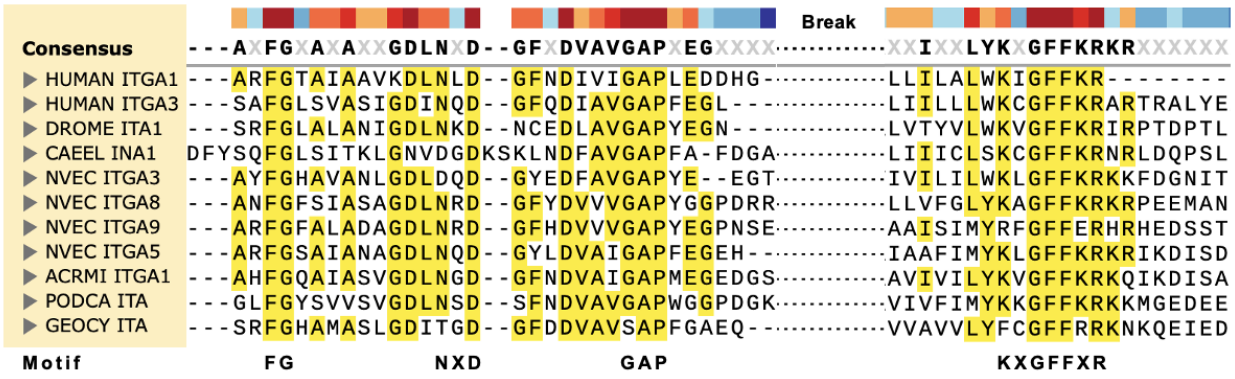


Fig. 3.10 A multiple sequence alignment of the alpha integrin subunit from different species. The alignment has been displayed with a break in the middle to showcase the conservation of the FG-GAP motifs and the KXGFFXR motifs. (Sequences - Table 2.2). Visualized using SnapGene. Species abbreviation: GEOCY - *Geodia cydonium,* PODCA - *Podocoryna carnea,* ACRMI - *Acropora millepora, NVEC – Nematostella vectensis,* CAEEL - *Caenorhabditis elegans,* DROME - *Drosophila melanogaster,* HUMAN - *Homo sapiens.*

The proposed integrin beta subunits when analyzed on InterPro contained the signal peptide (except *NvITGBn2*), vWFA domain (von Willebrand factor A), cysteine-rich stalk, non-cytoplasmic domain, transmembrane and cytoplasmic domains that are characteristic of vertebrate integrin beta subunits (Sebé-Pedrós et al., 2010). The consensus sequence for the Metal-Ion-Dependent-Adhesion-Site (MIDAS) (DXSXS) is completely conserved (Fig. 3.11) among all the subunits (Valdramidou et al., 2008). The ligand-associated metal binding site (NXDXPE), which houses another cation binding motif, also appears to be well conserved in the proposed subunits (Fig. 3.11). A key motif in metazoan beta integrin subunit, the NPXY motif, which facilitates interaction with intracellular scaffolding and signaling proteins, is also fully retained (Fig 3.11) (Sebé-Pedrós et al., 2010).

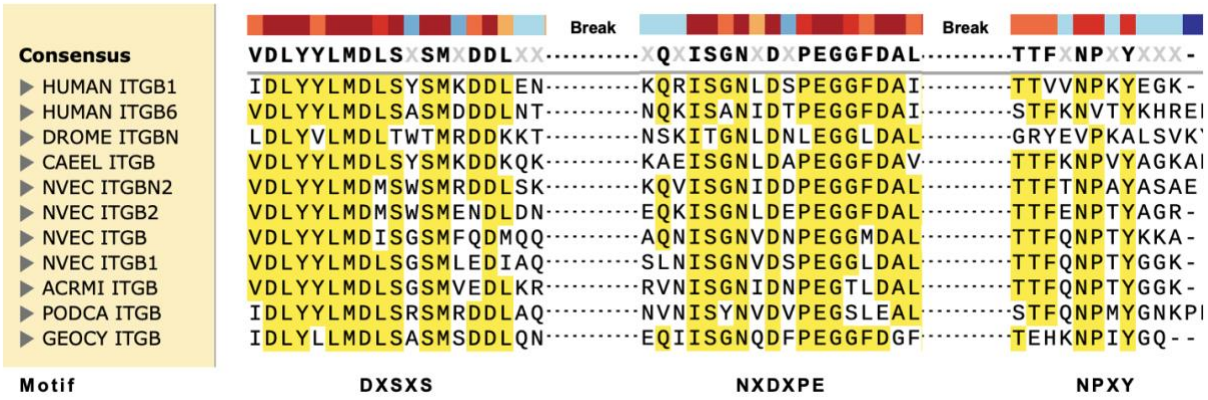


Fig. 3.11 A multiple sequence alignment of the beta integrin subunit from different species. The alignment has been displayed with breaks in the middle to showcase the conservation of the DXSXS, NXDXPE, and NPXY motifs. (Sequences - Table 2.2). Visualized using SnapGene. Species abbreviation: GEOCY - *Geodia cydonium,* PODCA - *Podocoryna carnea,* ACRMI - *Acropora millepora, NVEC – Nematostella vectensis,* CAEEL - *Caenorhabditis elegans,* DROME - *Drosophila melanogaster,* HUMAN - *Homo sapiens.*

While sequence homology and motif conservation are strong indicators of evolutionary ancestry, another parameter that must be considered is the overall protein structure. Therefore, predicted protein structures of the proposed *Nematostella* subunits were generated by AlphaFold and compared to their vertebrate homologs (Fig. 3.12). This computational analysis demonstrated a high degree of structural similarity and conserved domain architecture between the proposed subunits and their known vertebrate counterparts.

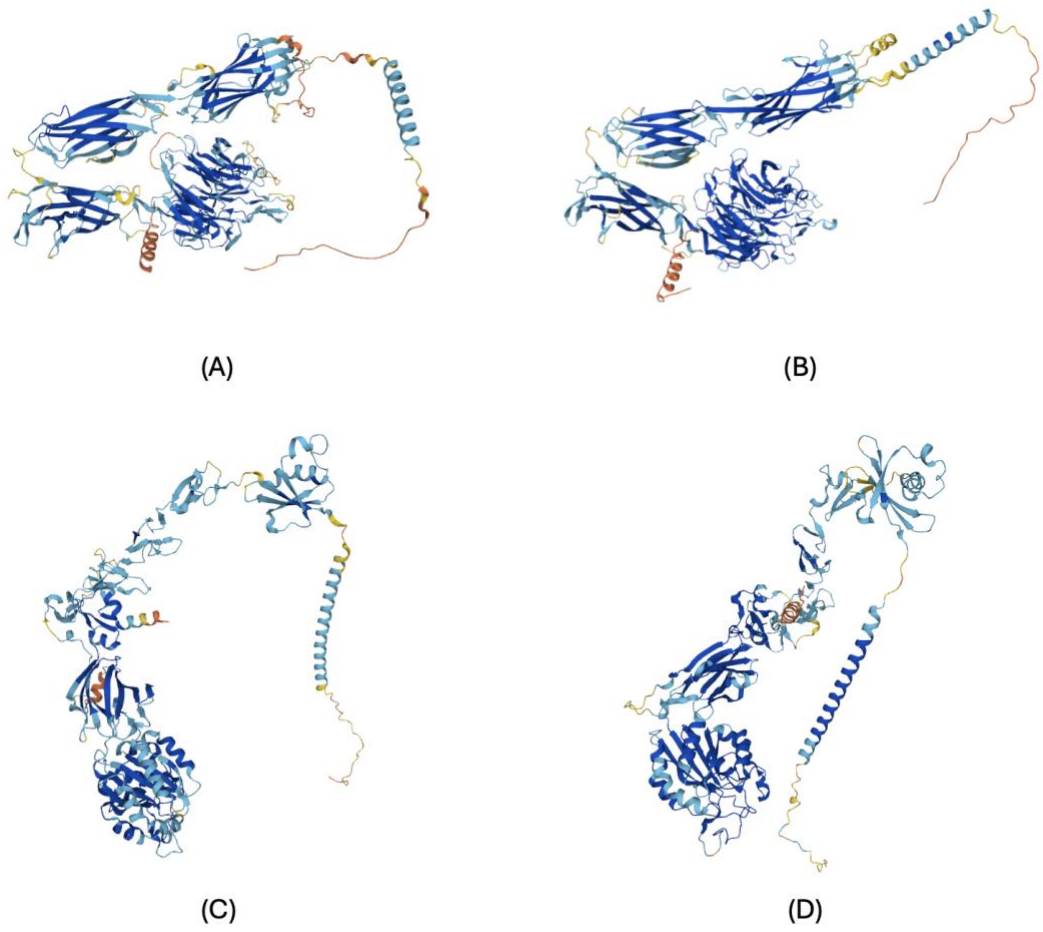


Fig. 3.12 Comparison of the predicted protein structure for *Nematostella* integrins against their human homologs using AlphaFold. The overall structure and subdomains are clearly conserved between (A) *Nematostella* integrin alpha subunit (*NvITGA8*) and (B) its human homolog (*HsITGA4*). Similarly, the organization and structure of (C) *Nematostella* integrin beta subunit (*NvITGB1*) shows comparable 3D organization to (D) its human homolog (*HsITGB1*).

Analyzing previously generated bulk RNA-sequencing data (Supplementary Fig. S1.1 and S1.2) indicated that many of these subunits are maternally deposited, but that the individual expression profiles for each of them varied with most subunits increasing expression as the embryo develops to the planula stage (*NvITGB1, NvITGB2, Nv ITGA5, NvITGA8*) while some others displayed high levels of expression in the egg(*NvITGB*) or in the post gastrula stage(*NvITGA3, NvITGA9*).

A surprising but serendipitous observation was the co-occurrence of cell clusters that express *NvITGB2* and *NvITGA3* in our in-house adult male *Nematostella* single-cell RNA seq dataset (Fig. 3.13). While preliminary and occluded by multiple factors, this could potentially indicate that these two subunits are interacting partners of the integrin receptor heterodimer given their nearly overlapping expression pattern. This type of overlapping expression pattern was not clear seen for all other combination of alpha-beta subunits examined, which might suggest that the other subunits interact with multiple partners (Supplementary Fig. S2).

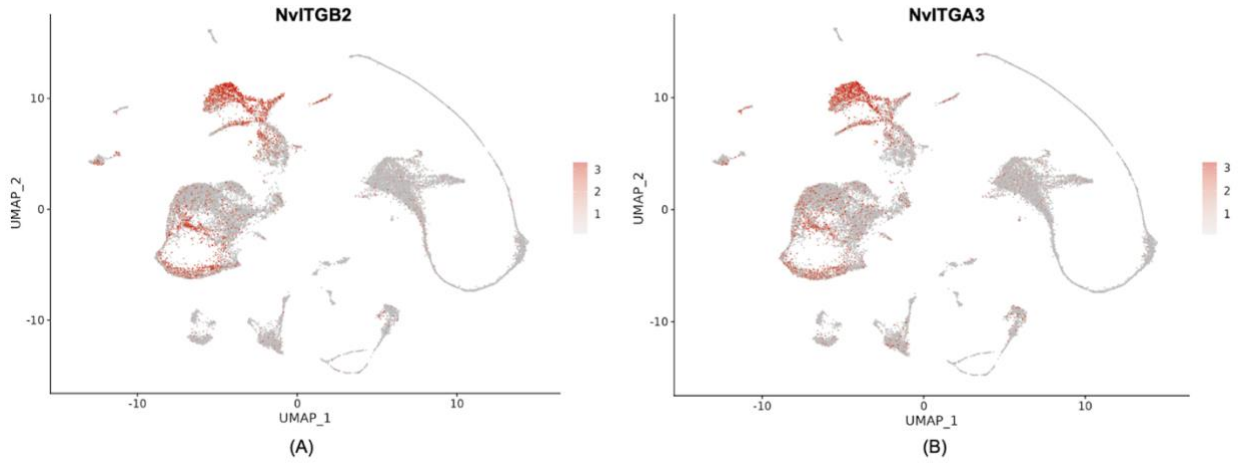


Fig. 3.13 UMAP projection of single-cell RNA-Seq data from a male adult *Nematostella*
Expression profiles for (A) *NvITGB2* and (B) *NvITGA3* are overlapping within the same clusters of cells.

The combined analysis of sequence homology, motif conservation, structural similarity, and expression profiles provides compelling evidence for the existence and interaction of these diverse integrin subunits in *Nematostella*.

3.7 Phylogenetic analysis of integrin subunits

A phylogenetic tree can provide insights into the evolutionary history of a homologous gene family. The maximum likelihood (ML) method is a statistical method that estimates a tree topology based on maximizing the probability of observing the given sequence alignment using an evolutionary substitution model. To understand the relationship between the *Nematostella* integrin sequences and integrins from other known phyla, a maximum likelihood phylogenetic tree was constructed. This included the proposed *Nematostella* integrin alpha and beta subunits along with representative members from other evolutionary-relevant phyla (Fig. 3.14 and Fig. 3.15).

The phylogenetic analysis of alpha subunits was consistent with previous literature, with the grouping of many bilaterian sequences into either an RGD binding or a laminin-binding clade (Fig. 3.14) (Knack et al., 2008). These two clades contain members from protostomia (Nematoda - *C. elegans* and Arthropoda - *D. melanogaster*) as well as deuterostomia (Vertebrata - *H. sapiens*), indicating that this functional divergence of the integrin alpha subunits likely occurred in the common ancestor of bilaterians. One of the *Nematostella* alpha subunits (*NvITGA3*) also clusters in the laminin-binding clade (with a high bootstrap value of 89), suggesting an early functional specialization of subunits. However, all the other cnidarian subunits (including *NvITGA5/8/9*) cluster together into a clade that within itself does not have a bilaterian member (Fig. 3.14).

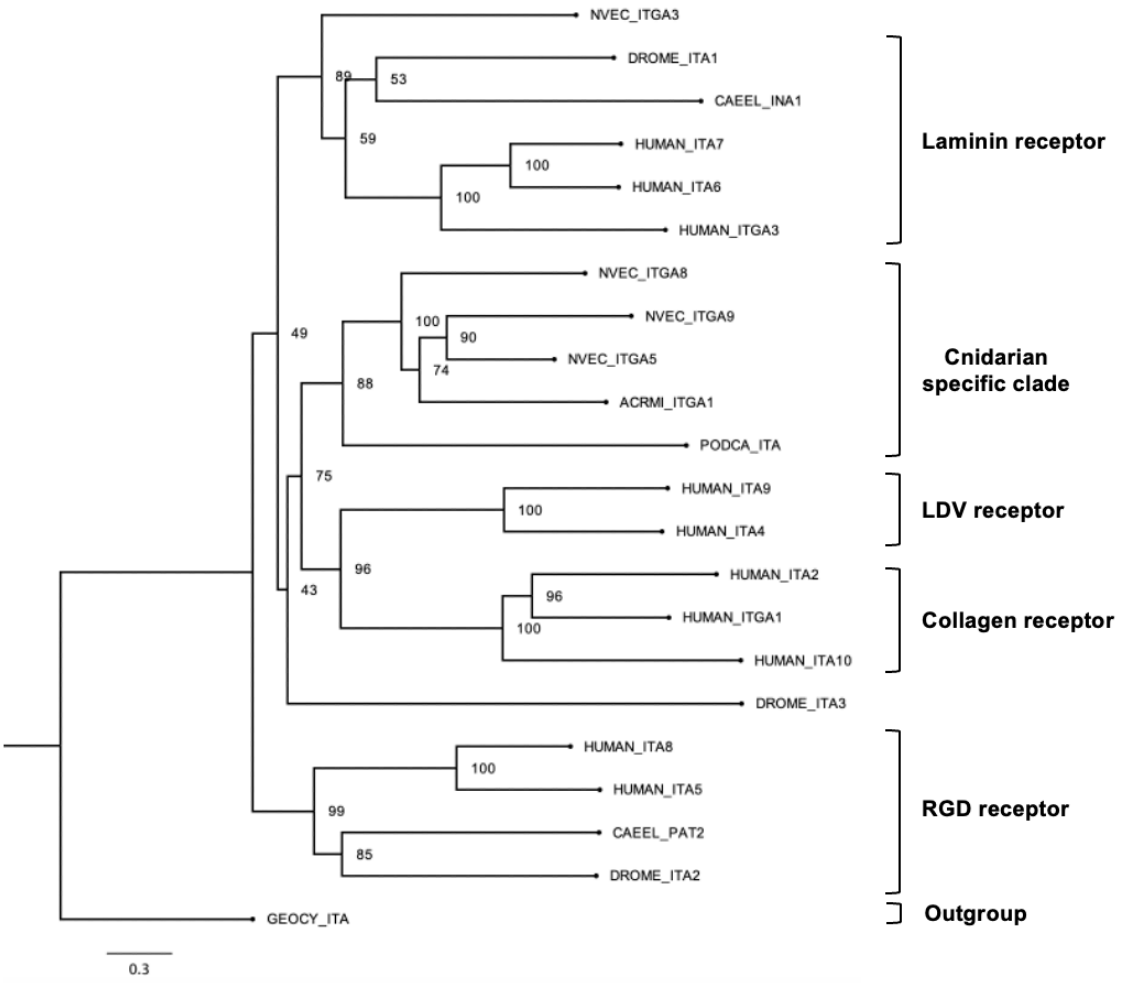


Fig. 3.14 Maximum likelihood phylogenetic tree constructed for integrin alpha subunits. Protein sequences from representative animals were used to generate the tree (Table 2.2). The node labels are the percentage bootstrap strength (out of 100). A sponge integrin alpha subunit (GEOCY_ITA) was chosen as an outgroup to root the tree. Alpha integrin subunits were clustered based on their canonical ECM binding specificity (Knack et al., 2008). Species abbreviation: GEOCY - *Geodia cydonium,* PODCA - *Podocoryna carnea,* ACRMI - *Acropora millepora, NVEC – Nematostella vectensis*, CAEEL - *Caenorhabditis elegans,* DROME - *Drosophila melanogaster,* HUMAN - *Homo sapiens.*

The phylogenetic analysis of the beta subunits also agreed with previous literature, which describes the subdivision of human subunits into three separate clades defined by *β1*, *β3,* and *β4* (Hughes, 2001). The invertebrate integrins cluster separately from their vertebrate homologs, and as such, this indicates the absence of functional

divergence of beta subunits in the bilaterian ancestor (Fig. 3.15). The cnidarian beta subunits cluster independently from all their bilaterian homologs, which further adds to the idea of gene duplication and independent divergence of beta subunits. The *Nematostella* subunits are split between two clades, with *NvITGB* and *NvITGB1* clustering with one of the *Acropora* (ACRMI_ITGB_B2) and *Podocoryne* subunit, while *NvITGB2* and *NvITGBN2* cluster with the other *Acropora* subunit (ACRMI_ITGB_O1) (Fig. 3.15).

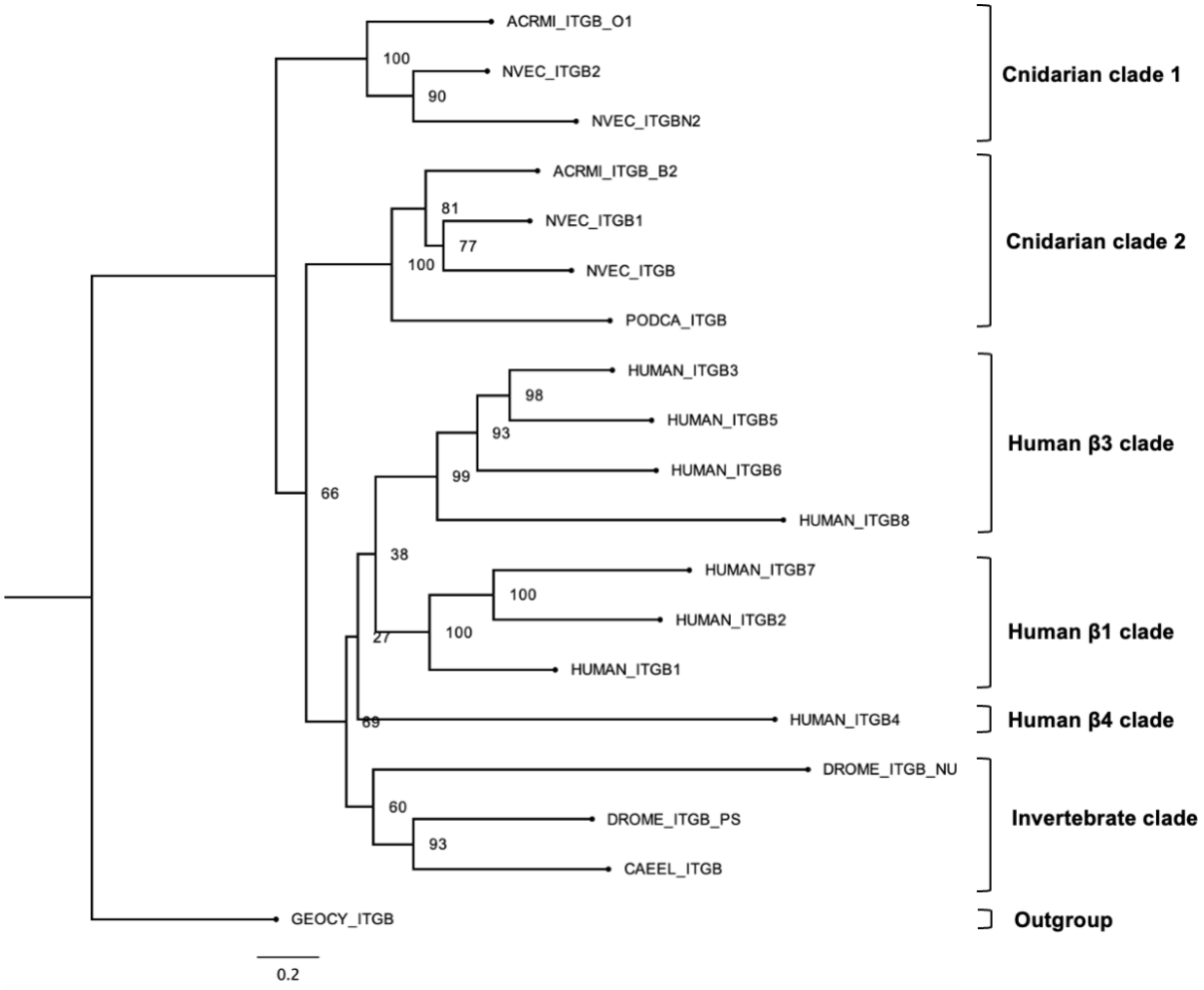


Fig. 3.15 Maximum likelihood phylogenetic tree constructed for integrin beta subunits. Protein sequences from representative animals were used to generate the tree (Table 2.2). The node labels are the percentage bootstrap strength (out of 100). A sponge integrin beta subunit (GEOCY_ITGB) was chosen as an outgroup to root the tree. Species abbreviation: GEOCY -

Geodia cydonium, PODCA - *Podocoryna carnea,* ACRMI - *Acropora millepora, NVEC – Nematostella vectensis,* CAEEL - *Caenorhabditis elegans,* DROME - *Drosophila melanogaster,* HUMAN - *Homo sapiens.*

While the generated phylogenetic trees are insightful in providing homology-based relationship between the integrin subunits among different species, most of the *Nematostella* subunits cluster together into cnidarian-specific clades that have the bilaterian subunits as a sister group making it difficult to derive individual orthology relationships.

3.8 In-situ expression profile for *Nematostella* integrin beta subunits

To understand the spatiotemporal expression pattern of the *N. vectensis* integrin beta subunits, colorimetric RNA in-situ hybridizations (CISH) were performed on embryos at different stages of development. DIG-labeled antisense probes were generated using *in vitro* transcription (IVT) of cloned cDNA fragments that had been previously sequenced and verified. To rule out nonspecific signal, negative controls for each of the subunits consisted of DIG-labeled sense strand probes that produced faint to null background staining (Fig. 3.16).

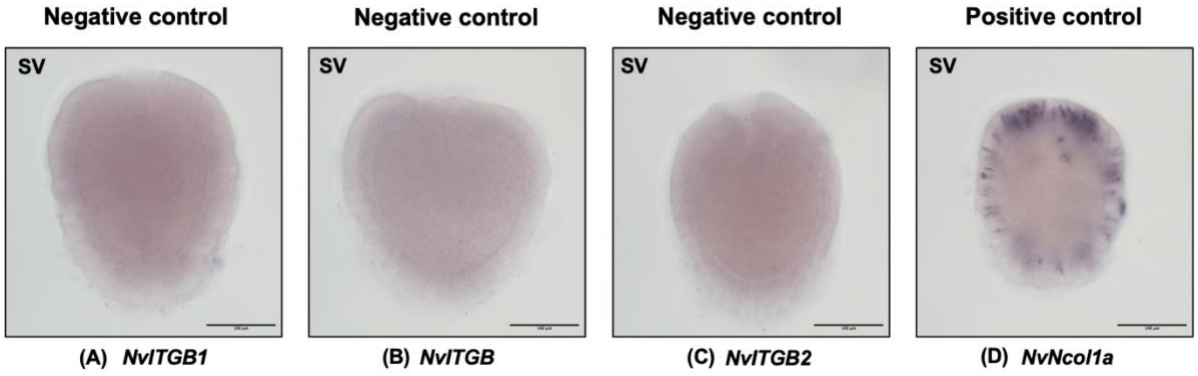


Fig. 3.16 Panel of CISH images for negative controls against (A) *NvITGB1*, (B) *NvITGB,* and (C) *NvITGB2*. The positive control is an antisense strand probe against (D) *NvNcol1a* which marks the cnidocytes. All images are side views (SV) of planula-stage embryos.

NvITGB1 appears to be expressed in the putative endomesoderm at the mid-gastrula stage (Fig. 3.17 A). As the embryo progresses through gastrulation its domain of expression expands but still appears to be restricted by the endomesoderm boundary (Fig. 3.17). At the planula stage it appears to be expressed in all eight segments and laterally throughout the length of each of the segments. As the planula elongates to a polyp, the expression of *NvITGB1* appears to be uniform and widespread throughout the gastrodermis and inner body wall but missing from the ectoderm.

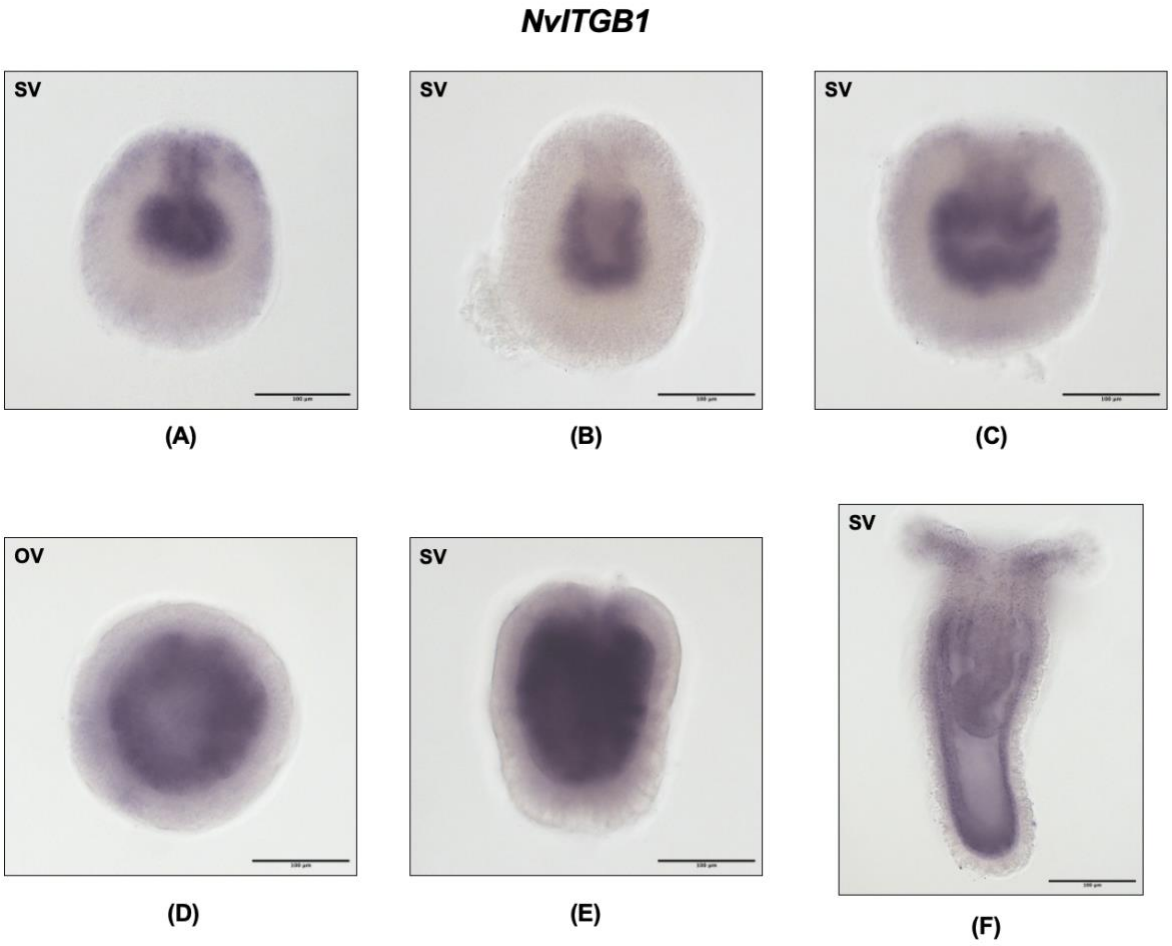


Fig. 3.17 Detection of *NvITGB1* expression using CISH. Expression of *NvITGB1* in (A, B) early and mid gastrula staged embryos (C) early planula, (D) planula, (E) late planula, and (F) primary polyp (views are side view (SV), oral view (OV)) of *Nematostella*.

Surprisingly both *NvITGB* and *NvITGB2* followed a similar pattern of expression with localization of the signal in the endomesoderm as the embryo progressed through the gastrula and planula stage, and the gastrodermis in the primary polyp stage (Fig. 3.18). However, unlike *NvITGB1* and *NvITGB2*, *NvITG*B appeared to be expressed at lower levels with fainter staining being observed at most stages.

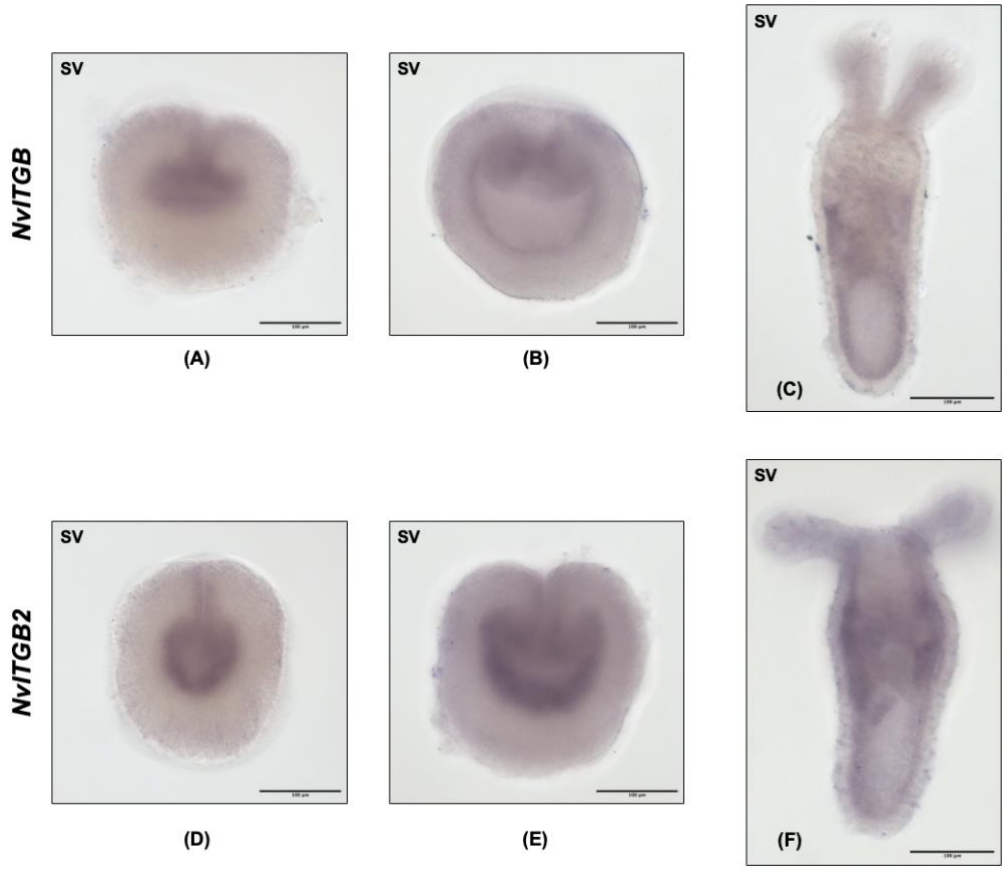


Fig. 3.18 Expression of *NvITGB* and *NvITGB2* using CISH. Side view (SV) images of *Nematostella* embryos stained with antisense probes against either *NvITGB* or *NvITGB2.* Expression in (A) early gastrula, (B) late gastrula, (D) mid gastrula, (E) late planula, and (C, F) primary polyp.

These staining patterns are consistent with the in-house scRNA seq data generated from an adult male *Nematostella* which shows that *NvITGB1* and *NvITGB2* are expressed in many of the same cell clusters (with *NvITGB2* in marginally fewer cells) while *NvITGB* is expressed in a fewer number of cells (Supplementary Fig. S2). Spatial

transcriptomics predictions using EndoAtlas indicate an asymmetric pattern of expression of the subunits at the planula stage (Supplementary Fig. S3). However, it was not possible to observe this asymmetry due to the strength of the CISH probe. To decipher this asymmetry, a fluorescent in-situ hybridization would need to be performed to obtain higher signal resolution. In the future, it would also be insightful to perform in-situ's for the integrin alpha subunits and understand their expression patterns.

NvITGBN2 was not included in the CISH analysis as it was annotated as a truncated protein and showed no significant expression profile in the bulk RNA seq dataset over various developmental time points (Supplementary Fig. S1.2). Nevertheless, these *in-situ* profiles indicate that the *NvITGB/1/2* subunits are expressed in the endomesodermal tissue beginning from early embryonic development in *Nematostella* and that their expression remains restricted to this region as development proceeds. This tissue is where the segmentation program takes place and from which the mesenteries and other tissues are derived in highly morphogenetic processes.

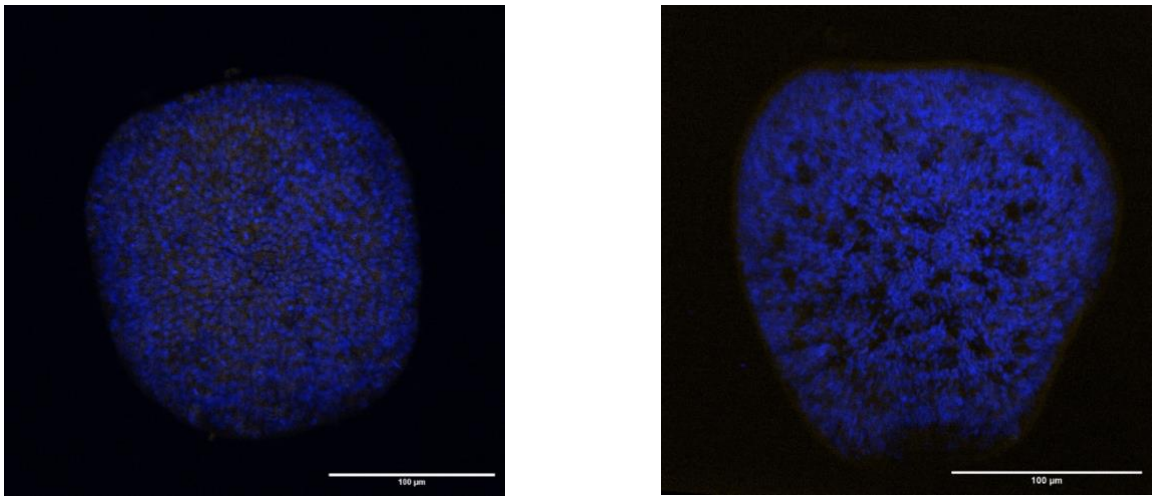
3.9 Cross-reactive integrin antibody trials

Antibodies are a versatile tool that once shown to be specific to a particular protein, can be used for a variety of informative studies ranging from spatio-temporal and conditional protein localization to understanding the interacting partners of the concerned protein.

Given the sequence and structural similarity of the putative *Nematostella* integrins to their vertebrate homologs, it seemed feasible to test commercially available antibodies, created against vertebrate integrins, for their binding in *Nematostella*. Therefore, whole mount immunohistochemical studies were performed with a panel of available human integrin antibodies (α4, αV, α5, β1, β3, β4, β5) at different developmental stages in *Nematostella*.

Most of the antibodies against human integrin subunits (α4, αV, β1, β3, β4) showed no specific staining at the planula and primary polyp stage. As a negative control, a set of embryos was processed without the primary antibody incubation step, using only the

secondary antibody. The non-specific integrin antibodies showed the same background staining that was seen in the negative control, confirming the absence of specific binding (Fig. 3.19). However, there were two antibodies that displayed distinct staining patterns, suggesting specific recognition of target.



(A) **(B)**

Fig. 3.19 Comparison of immunostaining of (A) negative control versus (B) human integrin antibodies that displayed no immunoreactivity in *Nematostella*. Both images are side views of planula stage embryos. The nuclei were stained blue with Hoescht. In the negative control no yellow fluorescence signal was observed which would have indicated binding of the secondary antibody to the primary antibody. Similarly, an antibody specific for (B) human integrin α4 produced no visible signal. Lack of signal was also observed when using antibodies specific for human integrins αV, β1, β3, β4.

The antibody against human integrin α5, when tested, exhibited a pattern of staining that marked the cell boundaries (Fig. 3.20). This was seen for both the planula stage as well as the primary polyp stage (Fig. 3.20 and 3.21). In addition, it also seemed to be staining cnidocytes in the ectoderm which appear as protruding cylindrical capsules. This staining profile appears promising as integrins, being cell surface adhesion receptors, are localized to the cell surface boundary.

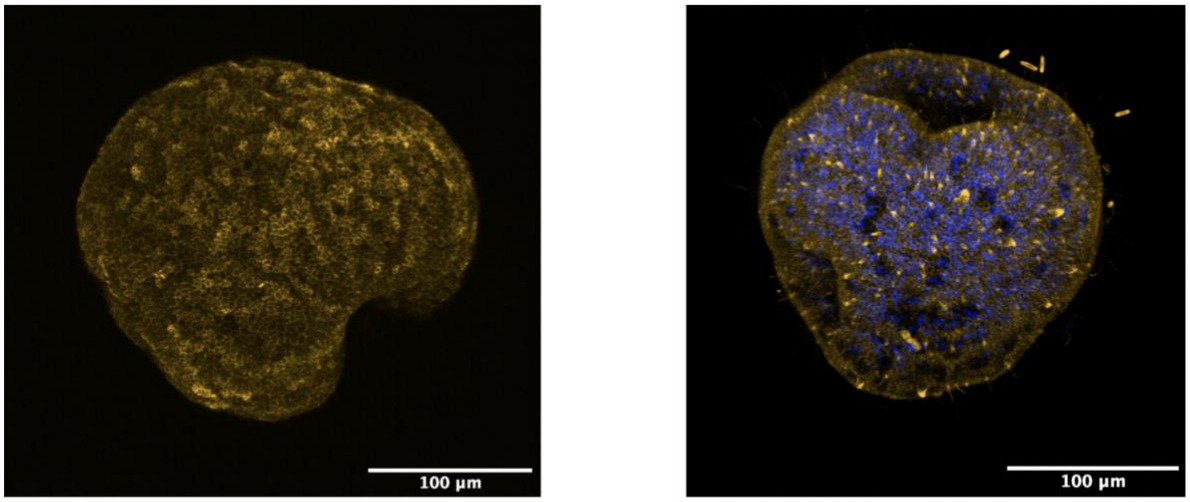


Fig. 3.20 Images of Immunohistochemistry results using an antibody designed to recognize human integrin α5, in *Nematostella* embryos. Immunoreactivity (yellow) was observed in (A, B) planula staged embryos. The nuclei are stained with Hoescht (in blue).

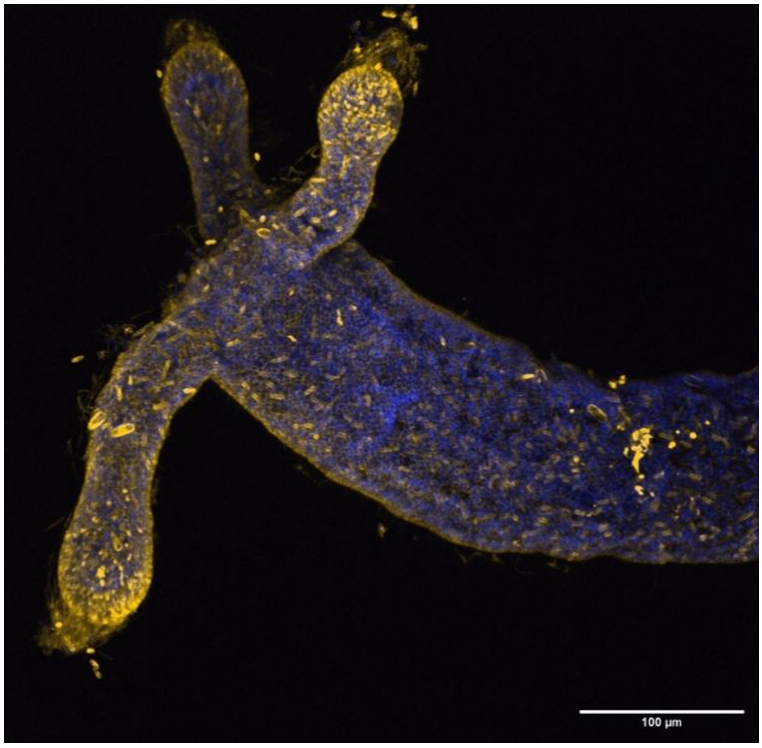


Fig. 3.21 Image of Immunohistochemistry result using an antibody designed to recognize human integrin α5 in *Nematostella* polyp. Immunoreactivity (yellow) was observed with the human integrin α5 antibody used on a primary polyp. Nuclei are stained with Hoechst (in blue).

The other antibody that produced a distinct staining pattern was directed against human integrin β5. At the planula stage, staining was observed as sparse punctae, which later expanded to a reticulate network surrounding brightly labeled punctae in the primary polyp (Fig. 3.22). This network appeared interconnected and ran parallel throughout the length of the body column. Although unexpected and preliminary, this staining pattern appears to specifically label what morphologically looks like the neural cell bodies and their processes. However, this would have to be confirmed by co-staining with a characterized neural marker.

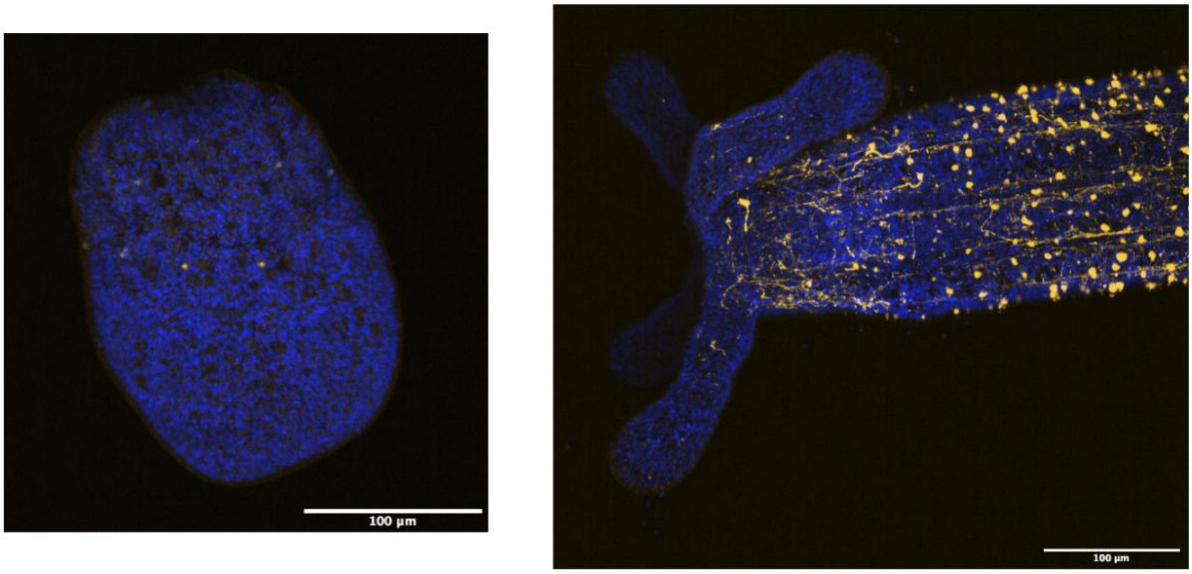


Fig. 3.22 Results of immunohistochemistry using an antibody directed against the human integrin β5 protein. Positive signals (yellow) were detected in (A) planula and (B) primary polyp embryos of *Nematostella*, staining what appears to be neuronal cell bodies based on the reticulate network in (B). The nuclei in both images were stained blue with Hoechst stain.

To assess the binding specificity of these antibodies, a western blot analysis was performed (Fig. 3.23). As expected, antibodies that did not stain specifically in the whole mount IHC (α4, αV, β1, β3, β4) also did not show any binding in the whole protein extracted from planula stage embryos. The integrin α5 antibody detected multiple bands with the most prominent band appearing around 100 kDa. Given that *Nematostella* integrin alpha subunits are approximately 115 kDa (calculated based on amino acid

sequence, Table 2.1) this band closely aligns with the predicted molecular weight. In contrast, the β5 antibody produced two sharp bands, both significantly lower than the predicted 80 kDa size (calculated based on amino acid sequence, Table 2.1) of *Nematostella* beta subunits.

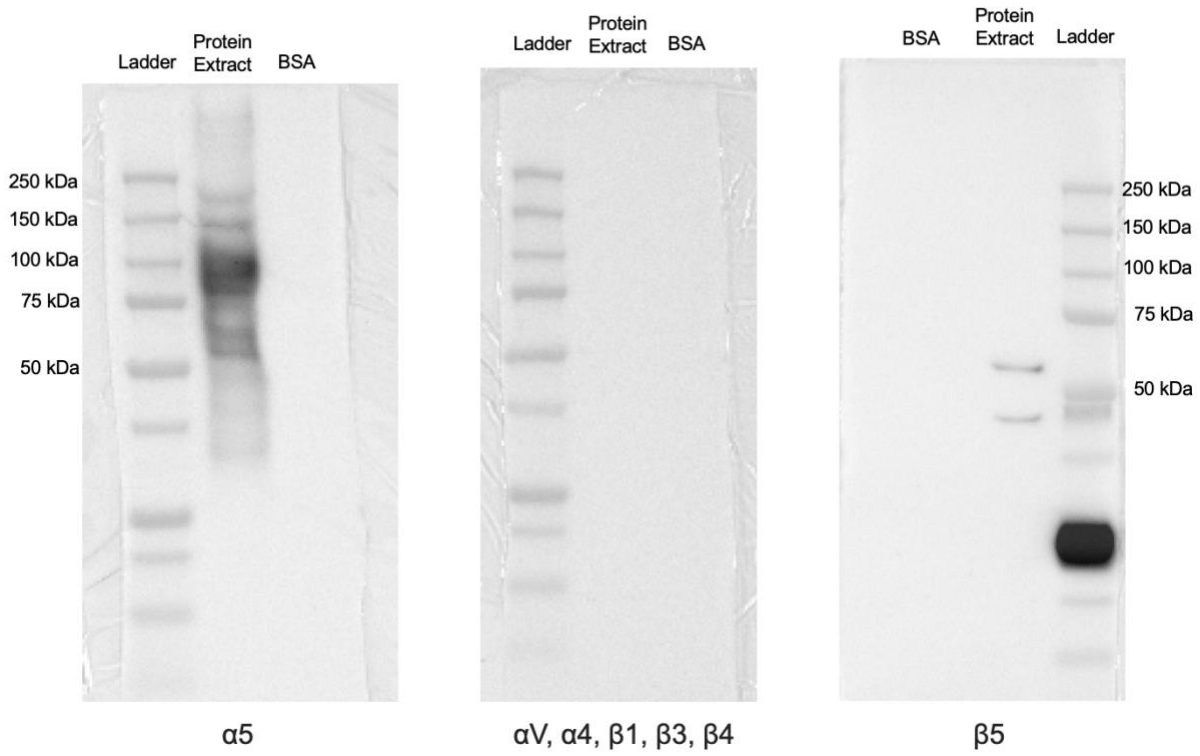


Fig. 3.23 Western blot analysis using whole protein extract from planula stage embryos. Lanes indicate protein reference ladder, whole protein extract (planula stage), and a BSA control. Only antibodies against human integrins α5 (left panel) and β5 (right panel) showed any immunoreactivity with proteins present in *Nematostella* embryos. All other antibodies produced blots with no signal (the center panel represents the result from one negative antibody shown as representation for antibodies that exhibit no binding).

To definitively determine whether these antibodies bind endogenous integrin subunits in *Nematostella*, a co-immunoprecipitation (Co-IP) pulldown followed by mass spectrometry analysis is to be performed in the future.

3.10 Functional analysis trials via integrin shRNA knockdowns

To functionally interrogate the role of these proposed integrin subunits, short hairpins (shRNA) were used to interfere with the translation of their transcripts. This involved designing a shRNA against the desired gene of interest, transcribing it *in vitro* from a template, and then microinjecting it into *Nematostella* embryos to assay for a phenotype during development.

Two shRNAs each were designed for *NvITGB1*, *NvITGB* and *NvITGB2,* targeting an exonic region in the coding sequence. One set of trials was performed for each of the subunits where the two shRNAs were coinjected together (600 ng/uL each), and another set was performed wherein a single shRNA for all the subunits was co-injected together (at a concentration of 400 ng/uL each). The results were compared against eGFP shRNA injected embryos as a negative control to account for injection trauma and overall variability in the wild type. As a positive control, shRNA targeting Nvβ-catenin was used, and as described in previous studies this blocked early gastrulation and exhibited more than 90% penetrance of phenotype (Karabulut et al., 2019). It was expected that with the knock down of integrin beta subunits, the embryos would fail in early development; however, embryos from both sets of microinjection experiments appeared to proceed normally through the blastula and into gastrula stages.

At the planula stage (3 dpf at 24°C or 5 dpf at 17°C), the embryos were fixed, stained with phalloidin and Hoechst and mounted to assay for a segmental phenotype. While the embryos injected with shRNAs against integrin beta subunits displayed segmental defects ranging from partial segment boundaries and a null segmental architecture, similar phenotypic deformities were observed at consistent levels in the control eGFP shRNA injected animals as well. The frequency of segmentation defects appeared to be higher in embryos injected with shRNAs against the integrin beta subunits; however, it was not substantially different from the eGFP shRNA negative control (on average across multiple trials, the fraction of deformed animals when shRNAs against integrin subunits were used - 62.5% (60/96), and for eGFP shRNA the average fraction of deformed embryos was - 46% (39/85)). Using a similar methodology, shRNA’s were

also designed and tested against the alpha integrin subunits, however, these observations were also confounded by the prevalence of deformities in the negative controls and the absence of a distinct early developmental disruption phenotype.

Therefore, the significant rate of segmentation defects in uninjected wild-type embryos (as observed in the experiments for *paraxis*) and the moderate efficiency of shRNA-mediated knockdowns, has currently made it difficult to assay for developmental phenotypes that involve segmental morphology solely based on the shRNA targeted against the integrin subunits.

3.11 Dissociation reaggregation studies – a new phenotypic avenue

Cnidarians are well known for their highly plastic body plan. *Nematostella*, in particular, has been shown to repattern its body plan from a cellular suspension reaggregated by centrifugation (Kirillova et al., 2018). While centrifugation-induced reaggregation provides a tool to study repatterning of the body axis, allowing dissociated cells to reaggregate in suspension might serve as a more suitable experimental setup to specifically assess the role of ECM adhesion molecules, such as integrins. This would also serve as an alternative paradigm to decipher the function of the proposed integrin subunits without the confounding effects of developmental deformities.

The existing protocol for centrifugation-mediated reaggregation was modified to allow for reaggregation from a suspension state. This involved dissociating *Nematostella* at the mid-gastrulae stage, filtering the cellular debris to achieve a uniform suspension of small clusters of cells, and then allowing reaggregation ex-vivo in 24 well plates. Immediately following cellular dissociation, the clusters of cells appear to start clumping together and begin re-aggregation (Fig. 3.24 A). In the beginning, most of these cellular aggregates are motile. At later time points, some form ciliated spheres that remain motile, while others form larger masses that settle to the bottom of the well (Fig. 3.24 C). Over time, these aggregates grow in size, and some form complete body plans, resulting in polyps.

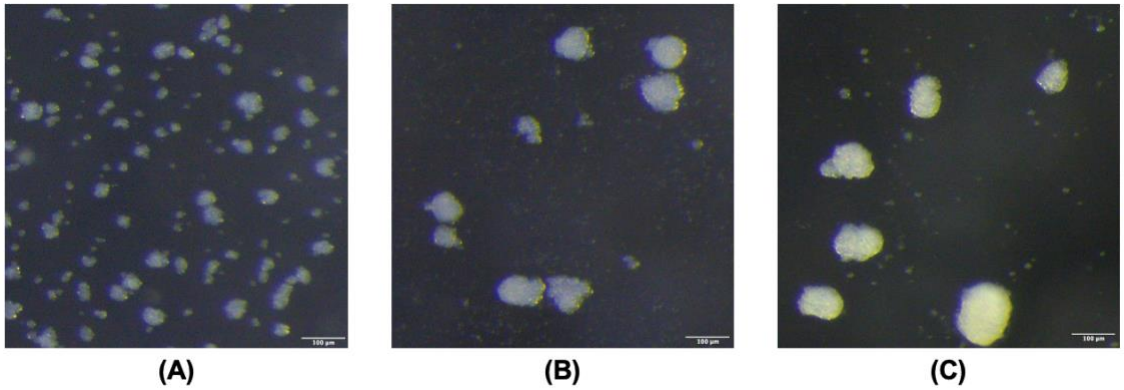


Fig. 3.24 Time course of a dissociation reaggregation experiment on mid-gastrula *Nematostella* embryos. Reaggregations of dissociated embryos at (A) 2-5 minutes post dissociation, (B) 18 hours, and (C) 47 hours post dissociation.

To determine whether integrins are involved in this process of reaggregation, an RT-qPCR assay was performed. Primer pairs were first validated with a serial dilution curve and a linear range of amplification was established (Fig. 3.25). When the experiment was performed with dissociation-reaggregation setups and compared to non-dissociated control sets of embryos it was observed that *NvITGB1* and *NvITGB2* undergo substantial upregulation (around 3.5 fold and 4.5 fold, respectively) while *NvITGB* remains nearly unperturbed during the process of reaggregation (Fig. 3.26). Another internal housekeeping control, *Nv18s* was used to ensure that the assay conditions were not affecting the levels of the internal housekeeping reference control (*NvGAPDH*).

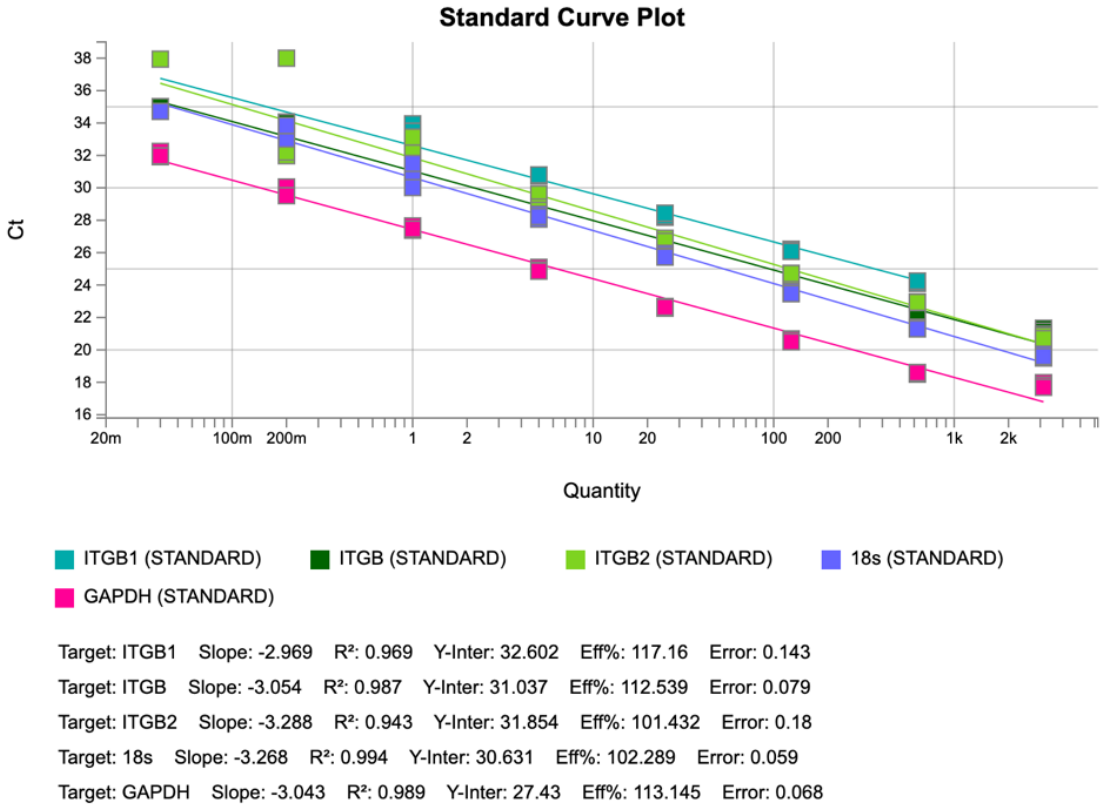


Fig. 3.25 A serial dilution curve for RT-qPCR primers designed against the *Nematostella* integrin beta and housekeeping control genes. The standard curve generated by RT-qPCR, using serial dilutions of wild-type cDNA, was used to validate the primer efficiencies and establish a linear range of amplification.

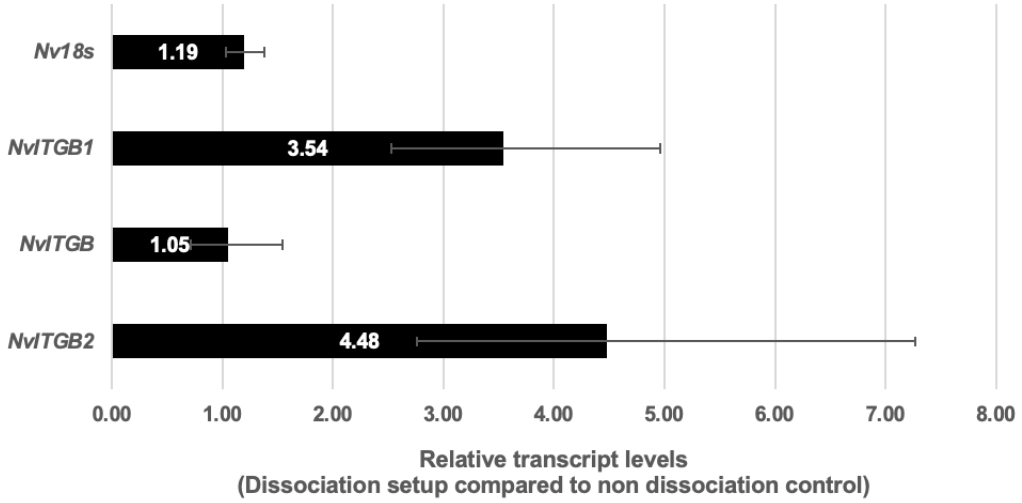


Fig. 3.26 RT-qPCR results examining integrin expression in RNA extracted from dissociated reaggregating *Nematostella* embryos versus wild type, non-dissociated embryos (control). The RT-qPCR data was analyzed using the double delta C_t method and the error bars were generated based on 95% confidence intervals. The changes in gene expression (transcript levels) are represented as a mean of the fold change seen in five biological replicates (section 2.12). Two of the integrin subunits, *NvITGB1* and *NvITGB2* appear to undergo substantial upregulation during the process of re-aggregation.

Although additional setups are needed to validate these findings, this preliminary data suggests that certain integrin subunits in *Nematostella* are involved in the reaggregation process as they might be required for cellular interactions and responding to ECM cues that allow cells to re-associate. To functionally assess the necessity of an upregulation in *NvITGB1/2*, in the future, it would be insightful to perform these reaggregation studies on embryos where these genes have been knocked down or mutated via CRISPR-mediated approaches.

CHAPTER FOUR – DISCUSSION

4.1 The role of *paraxis* in *Nematostella vectensis*

The hypothesis that *paraxis* contributes to the endomesodermal segmentation program in *Nematostella* was proposed based on the established role of its vertebrate counterpart, observation of its upregulation in scRNA data derived from the endomesoderm (He et al., 2023), and prior shRNA mediated knockdown experiments that resulted in malformed segments. This hypothesis was further supported by the high degree of sequence and structural conservation of the bHLH domain in the proposed *Nematostella* Paraxis, reinforcing its orthologous relationship with the vertebrate gene. However, when trying to replicate the earlier shRNA-mediated knockdown results, a new observation came to light. The phenotype of segmental deformity that had originally been ascribed to a knockdown of *paraxis*, was surprisingly being exhibited by the negative control (eGFP shRNA injected) and wild-type animals. And while the prior study had reported this phenotype in the control at negligible levels, in the current study, a significantly higher level of this developmental deformity was observed in the wild-type developing embryos.

There are multiple hypotheses that can explain this high rate of developmental ‘noise’ in the embryos. It was observed that the timing of fertilization, as well as the temperature of incubation, plays an important role in ensuring robust development of the embryo, however even accounting for these factors, there was something else not influenced by these apparent external factors that was going wrong, as upon the optimization of these factors – the appearance of deformed embryos still persisted. One possible hypothesis is that the animals used for the production of eggs and sperm are growing old which is leading to an age dependent decline in gamete quality. While easy to formulate, this topic is sparsely referenced in literature and the idea of aging in cnidarians is still being worked upon. The problem was further complicated by the logistics of the spawning event, with multiple groups of *Nematostella* being housed and spawned together, making it difficult to track and identify the exact individuals that contributed to the

deformed embryos. With the *paraxis* knockdown, this was even more concerning because of the lack of a specific characteristic phenotype, as this meant that the developmental phenotype could have just been an over-representation of the pre-existing background level of deformity. Confounding this analysis was the penetrance of the shRNA knockdowns. While RT-qPCR analysis confirmed that the designed shRNA was indeed reducing the levels of the *paraxis* transcript, the level of reduction was not complete. Given the role of *paraxis* as a transcription factor, even a partial knockdown may have allowed residual expression to activate downstream signaling cascades thereby mitigating the severity of the phenotype.

To address these issues, CRISPR-generated Paraxis mutants were examined. Contrary to what was originally expected, several of the homozygous mutant embryos exhibited the phenotypically normal eight-segmented body plan. These observations suggest that Paraxis may be acting in a redundant fashion with respect to segment boundary formation in *Nematostella*. Given the duplication and gene expansion of the bHLH family of TFs in *Nematostella*, it is entirely possible that there may be redundancy in the function of Paraxis and its effect may be compensated by a yet unidentified factor. For example, both *mesp* and *ripply* orthologs exist in the *Nematostella* genome and are upregulated in the endomesoderm at a similar time as *paraxis* (He et al., 2023). In vertebrates, both of these gene products play roles in establishing the polarity of somites during segmentation, thus contributing to boundary formation. In mouse mutants, all three produce segmentation defects, with the Paraxis mutation producing the mildest phenotype. A similar hierarchy could be present in *Nematostella,* with *NvMesp* and *NvRipply* making larger contributions to the segmentation program and thereby masking the loss or knockdown of Paraxis function. To address these issues, it would be necessary to generate either shRNA knockdowns or CRISPR mutants for *NvMesp* and *NvRipply*. Then, they could be examined singly or in combination to elucidate a potential mechanism. In parallel, the use of bulk RNA sequencing analysis on the Paraxis mutants can provide details into the downstream pathways that are affected in the absence of this transcription factor and whether they show any conservation to the vertebrate somitogenesis program.

Another possibility is that *NvParaxis* functions in a dissimilar fashion from its vertebrate counterpart. To support this, there is more than 600 million years of evolutionary divergence between these species. Paraxis's function may have diverged even though its core element, the bHLH region, is highly conserved, and it continues to be expressed in tissues that undergo segmentation. Nonetheless, additional work is required to determine its role during *Nematostella* development and whether or not it contributes to some aspect of segmentation.

4.2 Integrins in the cnidarian model system – *Nematostella vectensis*

Given the ancient evolutionary history and functional interconnectedness of the integrin receptors, it was surprising to see very sparse literature on integrins in *Nematostella* - and more broadly for the entire cnidarian phylum. In this study, four alpha and four beta integrin subunits were identified that exhibited all the motif signatures of canonical vertebrate integrins and shared considerable structural similarities. With additional support from phylogenetic analysis, it became apparent that these subunits are homologs of their bilaterian counterparts; however, due to the high levels of primary sequence divergence and gene duplication events, it was difficult to derive individual orthology relationships between the cnidarian and bilaterian subunits. Despite this divergence, one of the *Nematostella* alpha subunits clustered with high confidence in the laminin-binding group of bilaterian alpha subunits, and further interaction and functional studies could elaborate on whether this classification underlies an early functional divergence.

The choice to study integrins was based on the functional role of their vertebrate homologs in the process of somitogenesis and early embryonic development. For integrins in *Nematostella* to be involved in the endomesodermal segmentation process, they would need to be expressed in the endomesodermal tissue. Colorimetric in-situ hybridization studies with integrin beta subunits revealed a pattern of expression that appeared to be restricted to the developing endomesoderm as the embryo underwent gastrulation, and persisted throughout development in this tissue. In the future, a

fluorescent in situ hybridization could provide better resolution and help understand whether all the subunits are expressed with the same developmental profile (which might indicate redundancy) or if there are subtle differences in their patterns that could indicate that they have specialized roles.

Since integrins play a fundamental role in cellular adhesion to the extracellular matrix, it was predicted that disruption in their expression would produce morphological defects. However, functional studies, when performed with shRNA-based knockdowns targeting the integrin beta subunits, were confounded by the level of basal deformities in the wild type and the absence of a characteristic phenotype. To some degree, this variability of phenotype was expected, as these transcripts are maternally deposited, have some degree of overlapping expression, and play a fundamental role in cellular adhesion to the extracellular matrix. Consequently, perturbations in integrin levels could manifest as a gross developmental deformity at different stages depending on the extent of knockdown in individual embryos and potential compensating mechanisms. Future work should be focused on generating CRISPR Cas9-mediated mutant lines for these genes, so that their expression can be effectively deleted and they can be studied singly or in combination.

One possible, novel way to address the current shortcomings of this approach was to leverage the animal's ability to reaggregate when completely dissociated into a cellular suspension. Given the established role of integrins in cell-ECM adhesion, their potential involvement in the reaggregation process was of particular interest. Current knowledge of how a reaggregated body plan is built and how ECM components are incorporated during this process is limited. When set up, this assay proved to be both reproducible and robust, with a rapid and consistent phenotypic reaggregation response. Gene expression analysis revealed that *NvITGB1* and *NvITGB2* are substantially upregulated during this process, while *NvITGB* does not undergo a substantial fold change. This preliminary data suggests that some, but not all integrin subunits are involved in the reaggregation process, and in the future, it would be insightful to use chemical inhibitors or knockdown studies in this reaggregation setup to understand the necessity of integrin upregulation.

So far, based on these observations, while it was difficult to directly uncover the role that these proposed integrin subunits might play in the endomesoderm (due to their maternal RNA deposition and variability of the shRNA knockdown), it has been shown that they exhibit interesting homology relationships and spatiotemporal expression profiles that might underlie a functional relationship to the segmentation program.

Conclusion

Nematostella vectensis, the starlet sea anemone, while appearing to be a morphologically simple cnidarian species composed of two germ layers, possesses a surprisingly complex repertoire of genes, many of which have direct orthologs or homologous families in bilaterians. Crucial to its importance as a model for an evolutionary understanding of mesodermal processes such as gastrulation and segmentation, is the expression of a number of genes involved in bilaterian mesoderm specification and differentiation within its ‘endomesodermal’ tissue layer.

This study demonstrates that the *Nematostella* ortholog of Paraxis, a bHLH transcription factor essential for vertebrate somitogenesis, is by itself not crucial to the formation of segment boundaries in *Nematostella*. Elucidating its function would require a more in-depth study of the components involved in setting up the endomesodermal segment architecture. Meanwhile, integrins - cell surface receptors that mediate adhesion to components of the extracellular matrix, emerge as promising candidates for their involvement in segment formation, as suggested by their expression pattern and canonical role.

Studying segmentation as a paradigm in *Nematostella* is particularly compelling due to its striking parallels with the vertebrate segmentation program. This evolutionary comparison provides an opportunity to identify conserved regulatory factors and structural mechanisms underlying segmentation. Insights from this work could either reinforce the hypothesis of deep homology in segmental patterning or highlight an early functional divergence within gene families that have evolved over 600 million years.

REFERENCES

Aulehla, A., Wiegraebe, W., Baubet, V., Wahl, M. B., Deng, C., Taketo, M., Lewandoski, M., & Pourquié, O. (2008). A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. Nature cell biology, 10(2), 186–193.

Aulehla, A., & Pourquié, O. (2010). Signaling gradients during paraxial mesoderm development. *Cold Spring Harbor perspectives in biology*, *2*(2), a000869.

Barnes, G. L., Alexander, P. G., Hsu, C. W., Mariani, B. D., & Tuan, R. S. (1997). Cloning and characterization of chicken Paraxis: a regulator of paraxial mesoderm development and somite formation. Developmental biology, 189(1), 95–111.

Bessho, Y., Hirata, H., Masamizu, Y., & Kageyama, R. (2003). Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. Genes & development, 17(12), 1451–1456.

Burgess, R., Cserjesi, P., Ligon, K. L., & Olson, E. N. (1995). Paraxis: a basic helix-loop-helix protein expressed in paraxial mesoderm and developing somites. Developmental biology, 168(2), 296–306.

Burgess, R., Rawls, A., Brown, D., Bradley, A., & Olson, E. N. (1996). Requirement of the paraxis gene for somite formation and musculoskeletal patterning. Nature, 384(6609), 570–573.

Chal, J.; Pourquie, O. Patterning and differentiation of the vertebrate spine. In: Pourquie, O., editor. The Skeletal System. Cold Spring Harbor Laboratory Press; 2009. p. 41-116.

Chipman A. D. (2010). Parallel evolution of segmentation by co-option of ancestral gene regulatory networks. BioEssays : news and reviews in molecular, cellular and developmental biology, 32(1), 60–70.

Cooke, J., & Zeeman, E. C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. Journal of theoretical biology, 58(2), 455–476.

DeMali, K. A., Wennerberg, K., & Burridge, K. (2003). Integrin signaling to the actin cytoskeleton. Current opinion in cell biology, 15(5), 572–582.

Dequéant, M. L., & Pourquié, O. (2008). Segmental patterning of the vertebrate embryonic axis. Nature reviews. Genetics, 9(5), 370–382.

Diez del Corral, R., & Storey, K. G. (2004). Opposing FGF and retinoid pathways: a signalling switch that controls differentiation and patterning onset in the extending vertebrate body axis. BioEssays : news and reviews in molecular, cellular and developmental biology, 26(8), 857–869.

Dubrulle, J., McGrew, M. J., & Pourquié, O. (2001). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. Cell, 106(2), 219–232.

Fässler, R., & Meyer, M. (1995). Consequences of lack of beta 1 integrin gene expression in mice. Genes & development, 9(15), 1896–1908.

Genikhovich, G., Fried, P., Prünster, M. M., Schinko, J. B., Gilles, A. F., Fredman, D., Meier, K., Iber, D., & Technau, U. (2015). Axis Patterning by BMPs: Cnidarian Network Reveals Evolutionary Constraints. Cell reports, 10(10), 1646–1654.

George, E. L., Georges-Labouesse, E. N., Patel-King, R. S., Rayburn, H., & Hynes, R. O. (1993). Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. Development (Cambridge, England), 119(4), 1079–1091.

Gibb, S., Maroto, M., & Dale, J. K. (2010). The segmentation clock mechanism moves up a notch. Trends in cell biology, 20(10), 593–600.

Ginsberg, M. H., Du, X., & Plow, E. F. (1992). Inside-out integrin signalling. Current opinion in cell biology, 4(5), 766–771.

Girós, A., Grgur, K., Gossler, A., & Costell, M. (2011). α5β1 integrin-mediated adhesion to fibronectin is required for axis elongation and somitogenesis in mice. PloS one, 6(7), e22002.

Gong, Q., Garvey, K., Qian, C., Yin, I., Wong, G., & Tucker, R. P. (2014). Integrins of the starlet sea anemone Nematostella vectensis. The Biological bulletin, 227(3), 211–220.

Hand, C., & Uhlinger, K. R. (1992). The Culture, Sexual and Asexual Reproduction, and Growth of the Sea Anemone Nematostella vectensis. The Biological bulletin, 182(2), 169–176.

Hannibal, R. L., & Patel, N. H. (2013). What is a segment?. EvoDevo, 4(1), 35.

He, S., Del Viso, F., Chen, C. Y., Ikmi, A., Kroesen, A. E., & Gibson, M. C. (2018). An axial Hox code controls tissue segmentation and body patterning in Nematostella vectensis. Science (New York, N.Y.), 361(6409), 1377–1380.

He, S., Shao, W., Chen, S. C., Wang, T., & Gibson, M. C. (2023). Spatial transcriptomics reveals a cnidarian segment polarity program in Nematostella vectensis. Current biology : CB, 33(13), 2678–2689.e5.

Henry, C. A., Urban, M. K., Dill, K. K., Merlie, J. P., Page, M. F., Kimmel, C. B., & Amacher, S. L. (2002). Two linked hairy/Enhancer of split-related zebrafish genes, her1 and her7, function together to refine alternating somite boundaries. Development (Cambridge, England), 129(15), 3693–3704.

Hill, E. M., Chen, C. Y., Del Viso, F., Ellington, L. R., He, S., Karabulut, A., Paulson, A., & Gibson, M. C. (2022). Manipulation of Gene Activity in the Regenerative Model Sea Anemone, Nematostella vectensis. Methods in molecular biology (Clifton, N.J.), 2450, 437–465.

Hughes A. L. (2001). Evolution of the integrin alpha and beta protein families. Journal of molecular evolution, 52(1), 63–72.

Humphries, J. D., Byron, A., & Humphries, M. J. (2006). Integrin ligands at a glance. Journal of cell science, 119(Pt 19), 3901–3903.

Hynes R. O. (2002). Integrins: bidirectional, allosteric signaling machines. Cell, 110(6), 673–687.

Iimura, T., Denans, N., & Pourquié, O. (2009). Establishment of Hox vertebral identities in the embryonic spine precursors. Current topics in developmental biology, 88, 201–234.

Ikmi, A., McKinney, S. A., Delventhal, K. M., & Gibson, M. C. (2014). TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan Nematostella vectensis. Nature communications, 5, 5486.

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in bioinformatics, 20(4), 1160–1166.

Karabulut, A., He, S., Chen, C. Y., McKinney, S. A., & Gibson, M. C. (2019). Electroporation of short hairpin RNAs for rapid and efficient gene knockdown in the starlet sea anemone, Nematostella vectensis. *Developmental biology*, *448*(1), 7–15.

Kessel, M., & Gruss, P. (1991). Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. Cell, 67(1), 89–104.

Kirillova, A., Genikhovich, G., Pukhlyakova, E., Demilly, A., Kraus, Y., & Technau, U. (2018). Germ-layer commitment and axis formation in sea anemone embryonic cell aggregates. Proceedings of the National Academy of Sciences of the United States of America, 115(8), 1813–1818.

Knack, B. A., Iguchi, A., Shinzato, C., Hayward, D. C., Ball, E. E., & Miller, D. J. (2008). Unexpected diversity of cnidarian integrins: expression during coral gastrulation. BMC evolutionary biology, 8, 136.

Krumlauf R. (1994). Hox genes in vertebrate development. Cell, 78(2), 191–201.

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

Lee, H. O., Mullins, S. R., Franco-Barraza, J., Valianou, M., Cukierman, E., & Cheng, J. D. (2011). FAP-overexpressing fibroblasts produce an extracellular matrix that

enhances invasive velocity and directionality of pancreatic cancer cells. BMC cancer, 11, 245.

Linker, C., Lesbros, C., Gros, J., Burrus, L. W., Rawls, A., & Marcelle, C. (2005). beta-Catenin-dependent Wnt signalling controls the epithelial organisation of somites through the activation of paraxis. Development (Cambridge, England), 132(17), 3895–3905.

Magie, C. R., Daly, M., & Martindale, M. Q. (2007). Gastrulation in the cnidarian Nematostella vectensis occurs via invagination not ingression. Developmental biology, 305(2), 483–497.

Mallo, M., Wellik, D. M., & Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. *Developmental biology*, *344*(1), 7–15.

Mansouri, A., Voss, A. K., Thomas, T., Yokota, Y., & Gruss, P. (2000). Uncx4.1 is required for the formation of the pedicles and proximal ribs and acts upstream of Pax9. Development (Cambridge, England), 127(11), 2251–2258.

Maroto, M., Bone, R. A., & Dale, J. K. (2012). Somitogenesis. Development (Cambridge, England), 139(14), 2453–2456.

Moreno, T. A., & Kintner, C. (2004). Regulation of segmental patterning by retinoic acid signaling during Xenopus somitogenesis. Developmental cell, 6(2), 205–218.

Müller, M., v Weizsäcker, E., & Campos-Ortega, J. A. (1996). Expression domains of a zebrafish homologue of the Drosophila pair-rule gene hairy correspond to primordia of alternating somites. Development (Cambridge, England), 122(7), 2071–2078.

Nüsslein-Volhard, C., & Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. Nature, 287(5785), 795–801.

Ozbudak, E. M., & Lewis, J. (2008). Notch signalling synchronizes the zebrafish segmentation clock but is not needed to create somite boundaries. PLoS genetics, 4(2), e15.

Palmeirim, I., Henrique, D., Ish-Horowicz, D., & Pourquié, O. (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell, 91(5), 639–648.

Peel, A. D., Chipman, A. D., & Akam, M. (2005). Arthropod segmentation: beyond the Drosophila paradigm. Nature reviews. Genetics, 6(12), 905–916.

Peel, A., & Akam, M. (2003). Evolution of segmentation: rolling back the clock. Current biology : CB, 13(18), R708–R710.

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

Rifes, P., Carvalho, L., Lopes, C., Andrade, R. P., Rodrigues, G., Palmeirim, I., & Thorsteinsdóttir, S. (2007). Redefining the role of ectoderm in somitogenesis: a player in the formation of the fibronectin matrix of presomitic mesoderm. Development (Cambridge, England), 134(17), 3155–3165.

Rivera, A. S., & Weisblat, D. A. (2009). And Lophotrochozoa makes three: Notch/Hes signaling in annelid segmentation. Development genes and evolution, 219(1), 37–43.

Romanoff, A. L. (1960). The avian embryo. Structural and functional development.

Rowton, M., Ramos, P., Anderson, D. M., Rhee, J. M., Cunliffe, H. E., & Rawls, A. (2013). Regulation of mesenchymal-to-epithelial transition by PARAXIS during somitogenesis. Developmental dynamics : an official publication of the American Association of Anatomists, 242(11), 1332–1344.

Röttinger E. (2021). Nematostella vectensis, an Emerging Model for Deciphering the Molecular and Cellular Mechanisms Underlying Whole-Body Regeneration. Cells, 10(10), 2692.

Saina, M., Genikhovich, G., Renfer, E., & Technau, U. (2009). 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.

Sawada, A., Shinya, M., Jiang, Y. J., Kawakami, A., Kuroiwa, A., & Takeda, H. (2001). Fgf/MAPK signalling is a crucial positional cue in somite boundary formation. Development (Cambridge, England), 128(23), 4873–4880.

Schröter, C., Herrgen, L., Cardona, A., Brouhard, G. J., Feldman, B., & Oates, A. C. (2008). Dynamics of zebrafish somitogenesis. Developmental dynamics : an official publication of the American Association of Anatomists, 237(3), 545–553.

Sebé-Pedrós, A., Roger, A. J., Lang, F. B., King, N., & Ruiz-Trillo, I. (2010). Ancient origin of the integrin-mediated adhesion and signaling machinery. Proceedings of the National Academy of Sciences of the United States of America, 107(22), 10142–10147.

Shanmugalingam, S., & Wilson, S. W. (1998). Isolation, expression and regulation of a zebrafish paraxis homologue. Mechanisms of development, 78(1-2), 85–89.

Simionato, E., Ledent, V., Richards, G., Thomas-Chollier, M., Kerner, P., Coornaert, D., Degnan, B. M., & Vervoort, M. (2007). Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics. BMC evolutionary biology, 7, 33.

Šošić, D., Brand-Saberi, B., Schmidt, C., Christ, B., & Olson, E. N. (1997). Regulation of paraxis expression and somite formation by ectoderm- and neural tube-derived signals. Developmental biology, 185(2), 229–243.

Steinmetz, P. R. H., Aman, A., Kraus, J. E. M., & Technau, U. (2017). Gut-like ectodermal tissue in a sea anemone challenges germ layer homology. Nature ecology & evolution, 1(10), 1535–1542.

Stollewerk, A., Schoppmeier, M., & Damen, W. G. (2003). Involvement of Notch and Delta genes in spider segmentation. Nature, 423(6942), 863–865.

Takahashi, J., Ohbayashi, A., Oginuma, M., Saito, D., Mochizuki, A., Saga, Y., & Takada, S. (2010). Analysis of Ripply1/2-deficient mouse embryos reveals a

mechanism underlying the rostro-caudal patterning within a somite. Developmental biology, 342(2), 134–145.

Technau, U., & Steele, R. E. (2011). Evolutionary crossroads in developmental biology: Cnidaria. Development (Cambridge, England), 138(8), 1447–1458.

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols*, *3*(6), 1101–1108.

Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic acids research, 44(W1), W232–W235.

Tucker, R. P., Shibata, B., & Blankenship, T. N. (2011). Ultrastructure of the mesoglea of the sea anemone Nematostella vectensis (Edwardsiidae). Invertebrate Biology, 130(1), 11-24.

Turner R. L. (2024). The Metameric Echinoderm. Integrative organismal biology (Oxford, England), 6(1), obae005.

Valdramidou, D., Humphries, M. J., & Mould, A. P. (2008). Distinct roles of beta1 metal ion-dependent adhesion site (MIDAS), adjacent to MIDAS (ADMIDAS), and ligand-associated metal-binding site (LIMBS) cation-binding sites in ligand recognition by integrin alpha2beta1. The Journal of biological chemistry, 283(47), 32704–32714.

Watanabe, T., Sato, Y., Saito, D., Tadokoro, R., & Takahashi, Y. (2009). EphrinB2 coordinates the formation of a morphological boundary and cell epithelialization during somite segmentation. Proceedings of the National Academy of Sciences of the United States of America, 106(18), 7467–7472.

Wikramanayake, A. H., Hong, M., Lee, P. N., Pang, K., Byrum, C. A., Bince, J. M., ... & Martindale, M. Q. (2003). An ancient role for nuclear β-catenin in the evolution of axial polarity and germ layer segregation. Nature, 426(6965), 446-450.

Wilson-Rawls, J., Rhee, J. M., & Rawls, A. (2004). Paraxis is a basic helix-loop-helix protein that positively regulates transcription through binding to specific E-box elements. The Journal of biological chemistry, 279(36), 37685–37692.

Yang, J. T., Bader, B. L., Kreidberg, J. A., Ullman-Culleré, M., Trevithick, J. E., & Hynes, R. O. (1999). Overlapping and independent functions of fibronectin receptor integrins in early mesodermal development. Developmental biology, 215(2), 264–277.

Yang, J. T., Rayburn, H., & Hynes, R. O. (1993). Embryonic mesodermal defects in alpha 5 integrin-deficient mice. Development (Cambridge, England), 119(4), 1093–1105.

SUPPLEMENTARY FIGURES

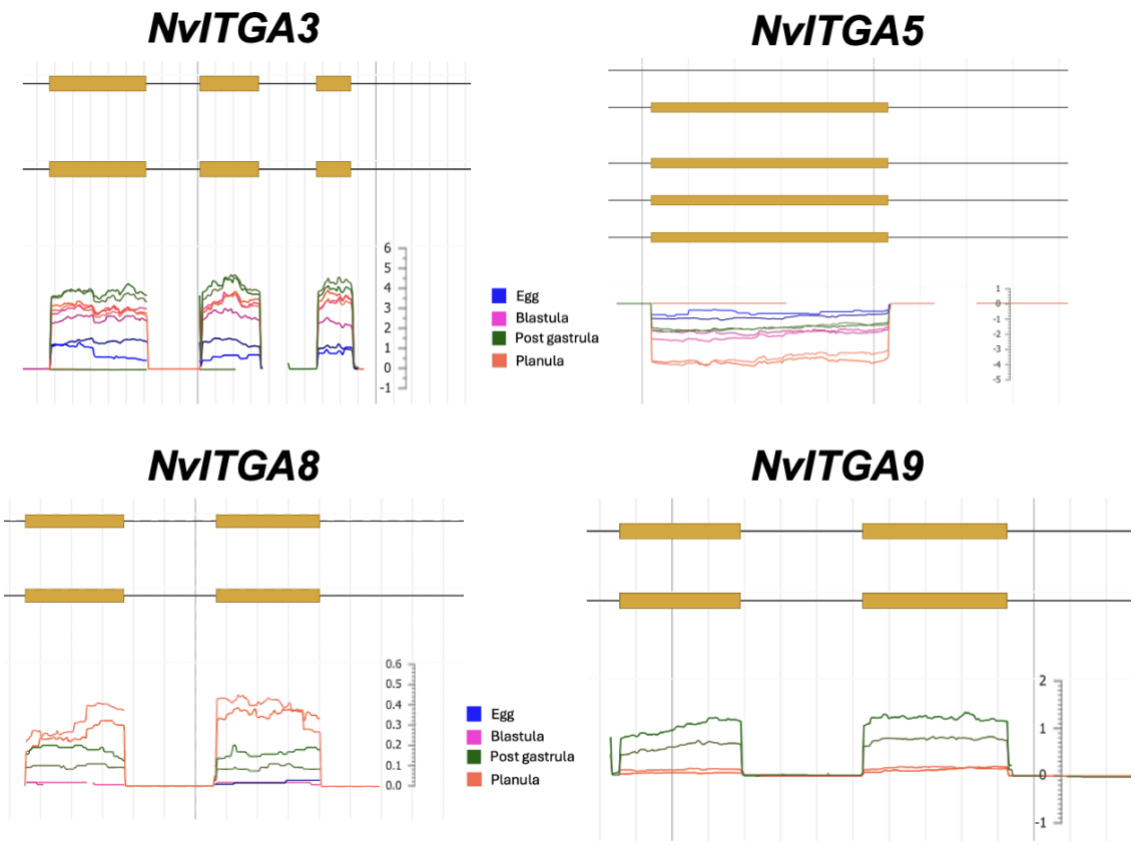


Fig. S1.1 RNA sequencing data profiled beneath a genomic map representing the *Nematostella* integrin alpha subunit loci. The graphic for each figure represents the *Nematostella* genes where the ochre-colored boxes represent exons and the black line represents introns. The organization of the integrin alpha subunits is composed of multiple introns and exons. The RNA sequencing data is presented in the bottom half of each figure, with a colored coded legend detailing the developmental time from which the RNA was extracted. The RNA sequencing data is displayed inverted for *NvITGA5* as it is transcribed in an orientation opposite to all others when displayed on a chromosomal map. Each of the subunits displays a unique expression profile that changes over embryonic development.

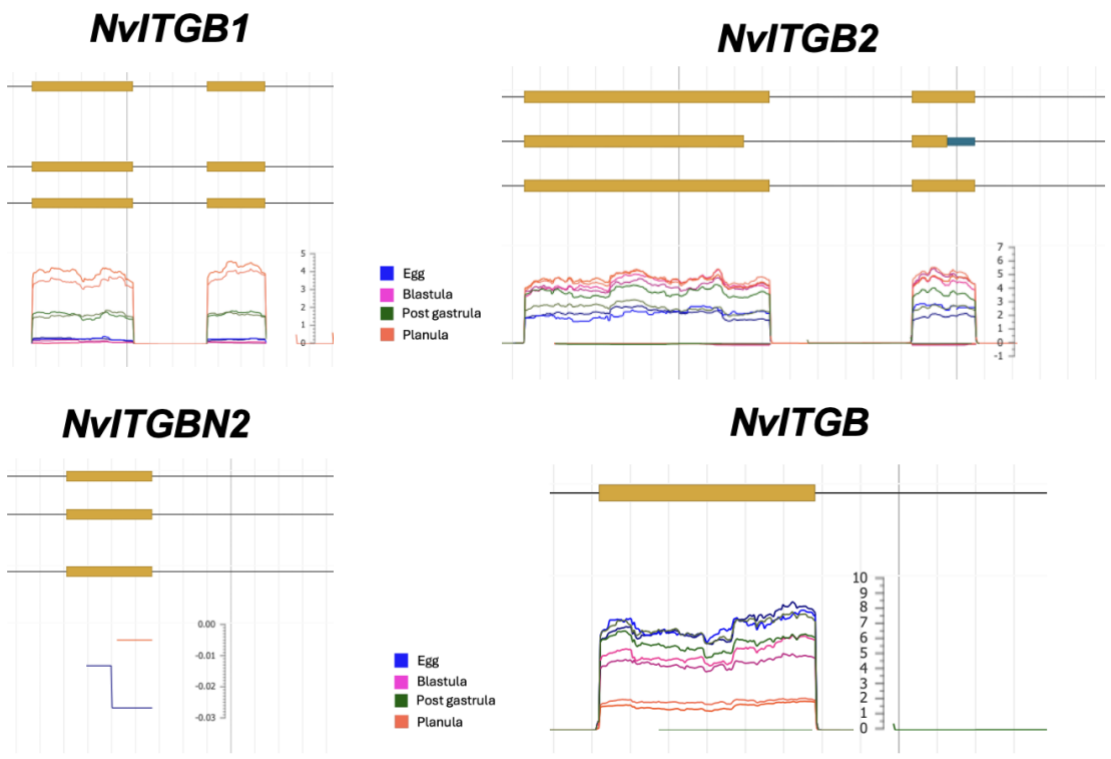


Fig. S1.2 RNA sequencing data profiled beneath a genomic map representing the *Nematostella* integrin beta subunit loci. The graphic for each figure represents the *Nematostella* genes where the ochre-colored boxes represent exons and the black line represents the intron. The organization of the integrin alpha subunits is composed of multiple introns and exons. The RNA sequencing data is presented in the bottom half of each figure, with a colored coded legend detailing the developmental time from which the RNA was extracted. *NvITGBN2* was not found to be expressed in this study and so displays poor RNA sequencing reads. All the other subunits (*NvITGB1*, *NvITGB,* and *NvITGB2*) display a unique expression profile that changes over embryonic development.

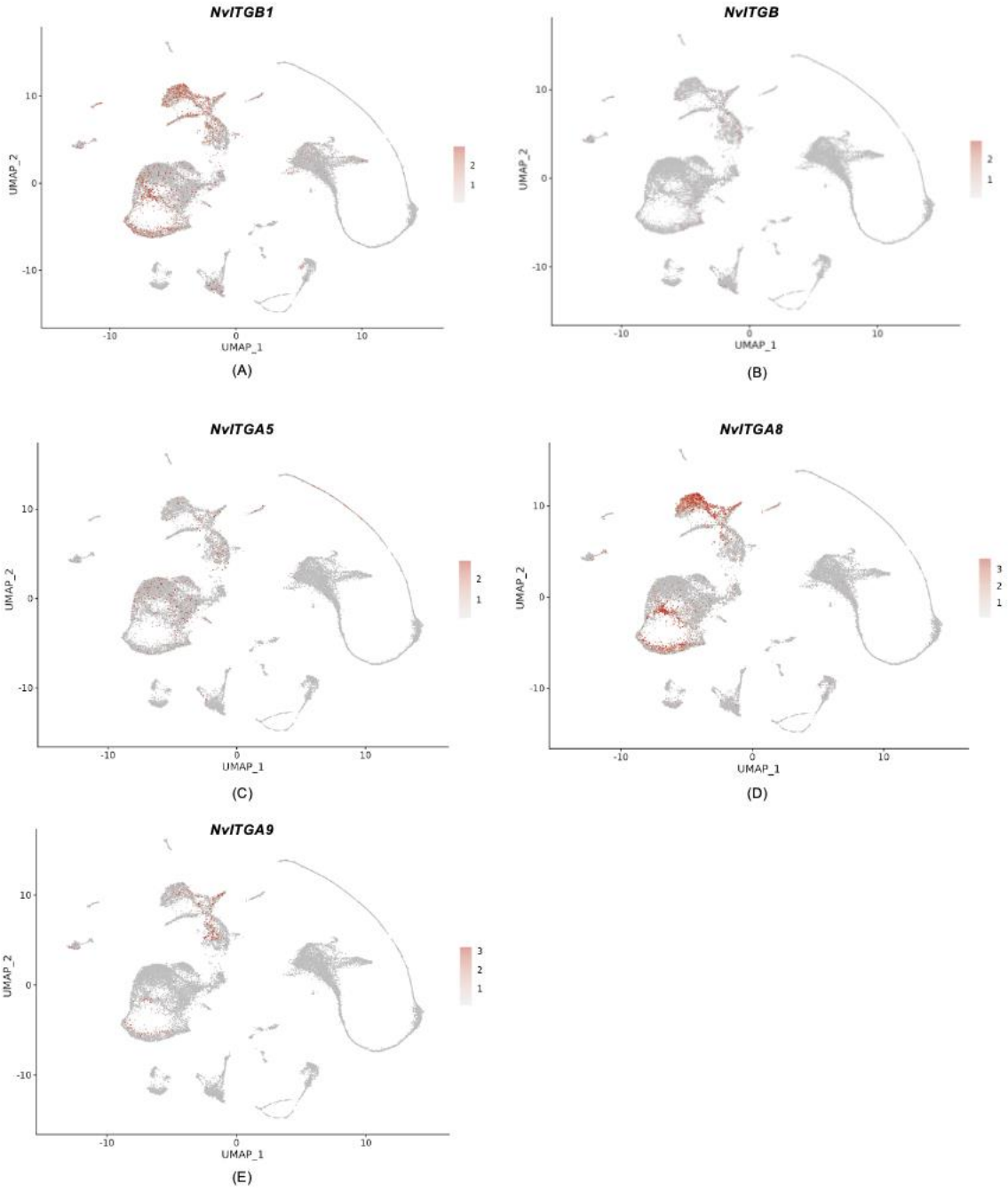


Fig. S2 UMAP projection of single-cell RNA-Seq (scRNA) data from a male adult *Nematostella* highlighting expression of integrin subunits. Expression profile for (A) *NvITGB1* resembles *NvITGB2* to an extent (Fig 6.4). While (B) *NvITGB* was also expressed in some of the same clusters, its expression was limited to fewer cells.

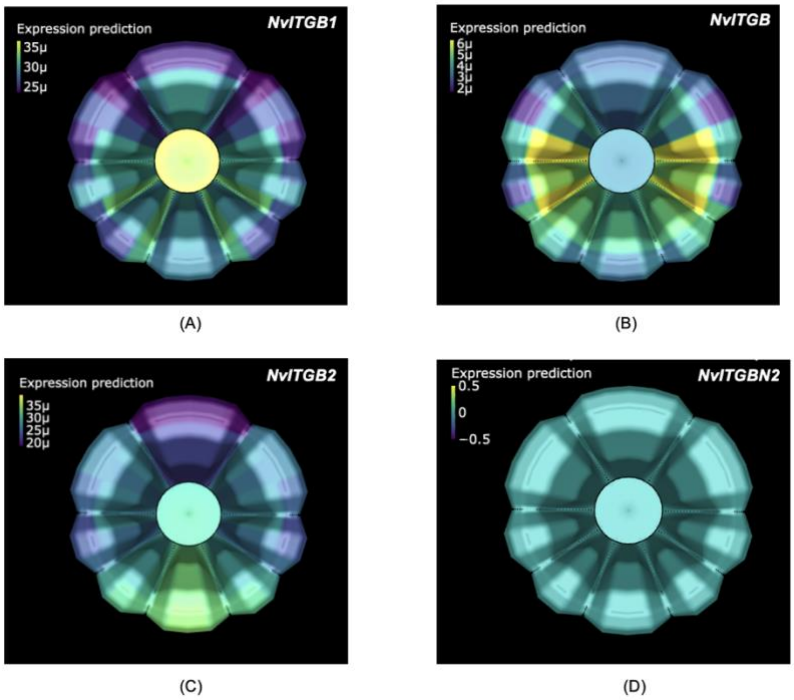


Fig. S3 Predicted spatial expression profile of *Nematostella* integrin beta subunits at the planula stage, based on EndoAtlas. (A) *NvITGB1*, (B) *NvITGB*, and (C) *NvITGB2* exhibited an asymmetric expression pattern in the endomesodermal tissue. (D) *NvITGBN2* did not show any expression in this study.