Graduate Programme in Molecular Biology and Biomedicine

UNIVERSITY OF CRETE

SCHOOL OF SCIENCES & ENGINEERING

DEPARTMENTS OF BIOLOGY AND MEDICINE

FOUNDATION FOR RESEARCH AND TECHNOLOGY HELLAS

INSTITUTE OF MOLECULAR BIOLOGY AND BIOTECHNOLOGY

MASTER THESIS

Role of leucine-rich repeats and teneurin transmembrane proteins in pancrustacean germband organization

Gentian Kapai

Supervisor:

Dr. Anastasios Pavlopoulos, Developmental Morphogenesis Laboratory, IMBB-FORTH

Thesis Committee:

Dr. Anastasios Pavlopoulos, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas

Prof. Christos Delidakis, Department of Biology, University of Crete

Dr. Christos Zervas, Biomedical Research Foundation, Academy of Athens

October 2021

Acknowledgements

This project would not have been possible without the support of many people. First and foremost, I would like to thank my supervisor, Dr. Anastasios (Tassos) Pavlopoulos, for constantly encouraging me and always being available for me. Many thanks to Dr. Evangelia (Valia) Stamataki and Dr. Marina Ioannou for their valuable comments and guidance. I would also like to thank my lab mates, John Rallis, Eirini Daskalaki, Themistoklis (Themis) Archontidis, Myrto Ziogas, Georgia Chalkia and John Liaskas for their company in the lab.

Also thanks to my committee members, Christos Delidakis and Christos Zervas, who offered guidance and advice.

Finally, I would like to express my deepest gratitude to my family and friends for their unfailing support and continuous encouragement.

This thesis was partly funded by Fondation Santé, who provided me with a Graduate Student Scholarship.

Abstract

The development of multicellular organisms relies on the specification of different cell types and their sorting and assembly into distinct tissues and organs. Cell sorting refers to the segregation of a mixed population of non-identical cells into distinct domains as well as the active maintenance of the segregated compartments. The process of cell sorting appears to be conserved between vertebrates and invertebrates, and was postulated to occur in embryonic, postnatal, and adult tissues. Cell sorting and assembly is controlled by various molecular interactions among the constituent cells, but is ultimately governed by differences between the physical properties of the specific cell types.

In metazoans, the elongation of the anteroposterior axis is a conserved developmental process and an exemplary developmental event relying on tissue convergence along the dorsal-ventral (DV) axis and extension along the anterior-posterior axis. In *Drosophila*, a common model organism in developmental biology, genetic studies have revealed patterns of transcription factor expression that provide spatial and temporal cues that induce oriented cell movements and promote germband extension. Nonetheless, the cellular mechanisms by which patterned transcriptional inputs orchestrate cell polarity and behaviour events had long been elusive. Zallen and colleagues have recently identified leucine-rich repeat and other transmembrane receptors that are expressed in transverse stripes along the head-to-tail axis and direct planar polarity and polarized cell rearrangements during germband convergent extension in *Drosophila*. Thus, tissue-level patterns of expression of transmembrane receptors provide spatial signals that link positional information from the AP patterning system to the essential cell behaviors that drive convergent extension

To determine how leucine-rich repeat and other transmembrane orthologs pattern the germband during the AP axis formation in other arthropods, I have chosen to study the amphipod crustacean *Parhyale hawaiensis*. *Parhyale* appears to be an exemplary model system that allows for single-cell analysis of the morphological process of germband formation. Similarly to other amphipod crustaceans, the ectoderm of *Parhyale* initially condenses from an unorganized population of cells into an organized grid of transverse rows of cells (perpendicular to AP axis) and longitudinal columns of cells (parallel to AP axis). Each transverse row of cells corresponds to one parasegment (ParaSegment Precursor Row - PSPR) termed ‘abcd’, will undergo two rounds of division along the AP axis. The first division generatesC two rows termed 'ab' and 'cd', which in turn become rows 'a','b','c', and 'd' following the second division. 4-row parasegments are easily discerned by the expression of the segment polarity gene *engrailed* in the anterior ‘a’ row. The progressive addition of new PSPRs at the posterior end of the germband and their stereotypic longitudinal divisions contribute to *Parhyale* germband elongation.

In this dissertation, I give a brief overview of the principles underlying cell sorting in development, and describe exemplary developmental processes in which cell sorting is observed. Additionally, I identify the orthologs of 4 *Drosophila melanogaster* genes, *tartan, capricious, Tanascin-major* and *Tenascin-accessory* in *Parhyale* and investigate their phylogeny across the phylum of Arthropoda. Subsequently, I use labeled

RNA probes to examine their expression during the embryonic stages of germband formation in *Parhyale*.

Our results indicate that *Parhyale* has four *tartan/capricious* orthologs and three *ten-m/-a* orthologs in total. All of them are expressed at the forming germand (embryonic stages 14 - 17) and exhibit almost identical expression patterns. The phylogenetic analysis and survey of the *Parhyale* genome support the theory that these genes are not the result of recent gene duplication events in *Parhyale,* raising questions about the specific roles in development for each gene. Functional analysis of these genes using CRISPR to knock out single genes and also groups of similar genes will reveal further information about their possible roles in the organization of the *Parhyale* germband.

Περίληψη

Η ανάπτυξη των πολυκύτταρων οργανισμών βασίζεται στον καθορισμό διαφορετικών τύπων κυττάρων και στην οργάνωση τους σε διακριτούς ιστούς και όργανα. Η οργάνωση κυττάρων αναφέρεται στον διαχωρισμό ενός μικτού πληθυσμού μη πανομοιότυπων κυττάρων σε διακριτές περιοχές καθώς και στην ενεργή διατήρηση των διαχωρισμένων διαμερισμάτων. Αυτή η διαδικασία φαίνεται να διατηρείται μεταξύ σπονδυλωτών και ασπόνδυλων και λαμβάνει χώρα σε εμβρυϊκούς, μεταγεννητικούς και ενήλικους ιστούς. Η οργάνωση των κυττάρων ελέγχεται από διάφορες μοριακές αλληλεπιδράσεις μεταξύ των συστατικών κυττάρων και διέπεται από διαφορές μεταξύ των φυσικών ιδιοτήτων των συγκεκριμένων τύπων κυττάρων.

Στα μετάζωα, η επιμήκυνση του πρόσθιου-οπίσθιου άξονα είναι μια διατηρημένη αναπτυξιακή διαδικασία και ένα υποδειγματικό αναπτυξιακό γεγονός που βασίζεται στη σύγκλιση ιστών των παρααξονικών κυττάρων προς τον AP άξονα του σώματος. Στη Δροσόφιλα, ένα πρότυπο οργανισμό στην αναπτυξιακή βιολογία, γενετικές μελέτες έχουν αποκαλύψει μοτίβα έκφρασης μεταγραφικών παραγόντων που παρέχουν χωρικές και χρονικές πληροφορίες οι οποίες καθορίζουν προσανατολισμένες κυτταρικές κινήσεις και επιμήκυνση ιστού. Παρ' όλα αυτά, οι κυτταρικοί μηχανισμοί με τους οποίους τα μοτίβα έκφρασης μεταγραφικών παραγόντων ενορχηστρώνουν την πολικότητα και τη συμπεριφορά των κυττάρων παρέμεναν άγνωστοι. Η Zallen και οι συνεργάτες της (2014) εντόπισαν τρεις υποδοχείς της οικογένειας Toll, τους Toll-2, Toll-6 και Toll-8, των οποίων η επικαλυπτόμενη έκφραση σε εγκάρσιες λωρίδες κατά μήκος του πρόσθιου-οπίσθιου άξονα κατευθύνει την επίπεδη πολικότητα και τις πολωμένες κυτταρικές αναδιατάξεις κατά τη συγκλίνουσα επέκταση στη Δροσόφιλα. Με αυτο το μηχανισμό, οι πληροφορίες θέσης κατά μήκος του πρόσθιου-οπίσθιου άξονα δημιουργούν μοτίβα έκφρασης του υποδοχέα Toll σε επίπεδο ιστού που παρέχουν χωρικά σήματα απαραίτητα για τις κυτταρικές συμπεριφορές που οδηγούν σε συγκλίνουσα επέκταση.

Για να προσδιορίσω πώς ορθολόγοι διαμεμβρανικοί υποδοχείς διαμορφώνουν το germband κατά τον σχηματισμό του άξονα AP σε άλλα αρθρόποδα, επέλεξα να μελετήσω το αμφίποδο *Parhyale hawaiensis*. Το *Parhyale* φαίνεται να είναι ένα υποδειγματικό οργανισμός μοντέλο που επιτρέπει την ανάλυση της μορφολογικής διαδικασίας σχηματισμού του germband σε επίπεδο κυττάρου. Ομοίως με άλλα αμφίποδα μαλακόστρακα, το εξώδερμα του Parhyale συμπυκνώνεται αρχικά από έναν μη οργανωμένο πληθυσμό κυττάρων σε ένα οργανωμένο πλέγμα στηλών και σειρών. Κάθε σειρά της ζώνης αναπαραγωγής (Parasegment Precursor Row - PSPR) είναι πρόδρομος ενός παρα-μεταμερούς. Τα κύτταρα διαιρούνται δύο φορές κατά μήκος του πρόσθιου-οπίσθιου άξονα δημιουργώντας δύο σειρές που ονομάζονται «A/B» και «C/D» στην αρχή, οι οποίες με τη σειρά τους γίνονται σειρές «a», «b», «c» και «d» ακολουθώντας τη δεύτερη διαίρεση. Τα παρα-μεταμερή 4 σειρών διακρίνονται εύκολα από την έκφραση του γονιδίου engrailed στην πρόσθια σειρά «a».

Σε αυτή τη διατριβή, δίνω μια σύντομη επισκόπηση των αρχών που διέπουν την ταξινόμηση κυττάρων στην ανάπτυξη και περιγράφω υποδειγματικές αναπτυξιακές διαδικασίες στις οποίες παρατηρείται ταξινόμηση κυττάρων. Επιπρόσθετα, εντοπίζω τα ορθόλογα 4 γονιδίων, *tartan, capricious, Tanascin-major* and *Tenascin-accessory* στο *Parhyale* και διερευνώ τη φυλογένειά τους σε όλο το γένος των Arthropoda. Στη συνέχεια, χρησιμοποιώ σημασμένους ανιχνευτές RNA για να εξετάσω την έκφρασή τους κατά τη διάρκεια των εμβρυϊκών σταδίων σχηματισμού του germband στο *Parhyale*.

Τα αποτελέσματά μας υποδεικνύουν ότι το Parhyale έχει τέσσερα ορθολόγα tartan/capricious και τρεια ορθολόγα ten-m/-a γονίδια συνολικά. Όλα εκφράζονται στο σχηματισμό germband (εμβρυϊκά στάδια 14 - 17) και παρουσιάζουν σχεδόν πανομοιότυπα μοτίβα έκφρασης. Η φυλογενετική ανάλυση και η έρευνα του γονιδιώματος Parhyale υποστηρίζουν τη θεωρία ότι αυτά τα γονίδια δεν είναι αποτέλεσμα πρόσφατων γεγονότων διπλασιασμού γονιδίων στο *Parhyale*, εγείροντας ερωτήματα σχετικά με τους συγκεκριμένους ρόλους του κάθε γονιδίου στην ανάπτυξη. Η λειτουργική ανάλυση αυτών των γονιδίων χρησιμοποιώντας το CRISPR σύστημα για την εξάλειψη μεμονωμένων γονιδίων και επίσης ομάδων παρόμοιων γονιδίων θα αποκαλύψει περαιτέρω πληροφορίες σχετικά με τους πιθανούς ρόλους τους στην οργάνωση του germband στο *Parhyale*.

Table of Contents

Acknowledgements 2

Abstract 3

Περίληψη 5

Chapter I 9

- **Principles of cell sorting** 9
- **Sorting by specific cell adhesion** 9
- **Forces driving cell sorting: Specific adhesion, differential adhesion, and adhesion strength** 10

Chapter II 12

- **Molecular origins of cell adhesion** 12
  - **Selectins** 12
  - **Cadherins** 12
  - **Integrins** 13
  - Immunoglobulin Superfamily 13
  - **Leucine Rich Repeat Proteins** 14
  - **Teneurins** 15
- **Molecular origins of cortical forces** 16
  - **Myosin II assembly** 16
  - **Actin assembly at the cell cortex** 16
  - **F-actin cross-linkers and the mechanical properties of actomyosin networks** 17
  - **Cortical forces and adhesion structures in epithelia** 17

Chapter III 19

- **Cell sorting in development** 19
- **Conserved mechanisms of cell sorting in different organisms** 19
  - **Cell segregation during zebrafish and Xenopus gastrulation** 20
  - **Border formation during vertebrate somitogenesis and hindbrain segmentation** 20

Chapter IV 22

- **The crustacean Parhyale hawaiensis as a model system for arthropod development** 22
  - **The biology of Parhyale hawaiensis** 22
  - **Uses of the P. hawaiensis model system** 25
  - **P. hawaiensis ectodermal germband formation** 25

Chapter V 28

Summary of this study 28

Chapter VI 30
Genome/ transcriptome search for leucine-rich repeat and teneurin proteins in P.hawaiensis 30
Materials and Methods 30
Results 30

Chapter VII 40
Phylogenetics analysis of tartan, capricious, tenascin major and accessory in arthropods 40
Materials and Methods 40
Results 43

Chapter VIII 45
Cloning of Parhyale tartan, capricious, tenascin major and accessory orthologs 45
Materials and Methods 45

Chapter IX 47
Expression analysis of Parhyale tartan, capricious, tenascin major and accessory during germband formation and maturation using in situ hybridization 47
Materials and Methods 47
Results 48

Chapter X 51
Discussion 51
Arthropod segmentation: The Drosophila paradigm and Parhyale as a model system to study morphological movements and temporal patterning in sequentially segmenting arthropods 52
Expression of trn/ caps and ten-m/ -a orthologs in relationship to the expression of segmentation genes in Parhyale 53
Expression of genes during PSPR cell migration 53
Expression of genes during PSPR division 54
Expression of genes after PSPR division in A/B or C/D 54
Expression of genes after A/B or C/D division in ‘a’,’b’ or ‘c’,’d’ 54
Ph_trn/caps_2 expression not being apparent in the germband 55

Chapter XI 56
Future directions 56

References 57

Chapter I

Principles of cell sorting

Multicellular, eukaryotic organisms consist of hundreds of distinct cells, with different attributes, dimensions and roles. Yet, all of them originate from a single zygote, whose progeny cells divide and specialize in order to give form to the organism. The segregation of heterotypic cells is a process known as cell sorting. Multiple cell-sorting events occurring *in vivo* have been successfully replicated *ex vivo* using co-cultures of primary cells, despite the fundamental differences of these micro-environments (Armstrong, 1989). In order to link cell sorting phenomena to the mechanical and biochemical properties of the constituent cells, different theories were proposed. Holtfreter's work in the amphibian embryo put forward basic concepts that advanced our understanding of cell sorting and tissue segregation (Armstrong, 1989; Townes & Holtfreter, 1955). Early *ex vivo* experiments revealed that dissociated cells from distinct tissues invariably reaggregated into a single cell mass when mixed, suggesting the existence of a generic adhesion system universal to all cell types in the early embryo. Nonetheless, eventually cells would sort out according to their origins, indicating the existence of a second, cell type-specific adhesion system. Even when entire sections of tissue were mixed, different types of cells displayed varying degrees of attraction or avoidance, which was encapsulated under the notion of tissue affinity (Townes & Holtfreter, 1955).

Sorting by specific cell adhesion

Sorting experiments provided evidence for the existence of two adhesion principles, which acted in conjunction with each other to control cell adhesion. Non-selective adhesion was thought to promote overall cell segregation, while selective adhesion caused sorting and compartmentalization. Both modes of adhesion could be present simultaneously but act at different time scales during development (Townes & Holtfreter, 1955). At the same time, experimental evidence for the presence of specific adhesion components and their biochemical properties was growing. In the 1960s, glycoproteins with the ability to cause species- or tissue-specific sorting were characterized as soluble, aggregation promoting factors in sponges and chicken embryos (Moscona, 1968). Neural cell adhesion molecule, NCAM (Brackenbury et al., 1977; Thiery et al., 1977) and E-cadherin (Takeichi, 1977), were also identified as transmembrane receptors that showed homophilic binding, but were insufficient to delineate the sorting phenomena observed in organisms (Prakasam et al., 2006; Shimoyama et al., 2000). Eventually, factors that could modulate adhesion were characterized such as the adhesion-antagonizing ephrins and their Eph receptors (Cheng & Flanagan, 1994; Drescher et al., 1995). Contrary to someone's expectations, Eph-ephrin signaling, which inherently requires cell-cell contact, leads to cell repulsion and separation. This leads to the interaction being disrupted, signaling ceases, and cells can re-attach, thus creating a potentially dynamically control cycle of repeating adhesion and de-adhesion events (Mellitzer

et al., 1999; Tanaka et al., 2003, Batlle & Wilkinson, 2012; Fagotto, 2014, 2015, Rohani et al., 2011). In addition, adhesion to the extracellular matrix (ECM), despite being fundamentally different from cell-cell adhesion, can promote tissue cohesion by embedding cells in intervening ECM, leading to their "attachment". For example the ECM protein fibronectin interacts with the transmembrane receptor α5β1 integrin and causes cell-cell adhesion when co-expressed in Chinese Hamster Ovary cells (Robinson et al., 2003, 2004). Therefore, the overall adhesive properties of a specific cell type is the summary of distinct adhesive factors and modulators, each with varying degrees of expression levels and binding affinities for its different interaction partners that are expressed in adjacent cells of the surrounding ECM. Overall, the qualitative concept of specific adhesion can account for many, though not all of the observed sorting phenomena.

Forces driving cell sorting: Specific adhesion, differential adhesion, and adhesion strength

The mechanics of cell-cell adhesion are arguably the most significant aspect of understanding morphogenesis (Shawky & Davidson, 2015). This includes the mechanical strength of cell-cell adhesion, a central variable in Steinberg's Differential Adhesion Hypothesis that defined these parameter as the work of adhesion, usually expressed as the energy released upon binding of adhesion molecules, analogously to the thermodynamic explanation of liquid molecules held together by attractive forces (Steinberg, 1963, 1970). Although, this is an elegant quantitative theory of cell sorting, the binding energies per unit area of cell-cell contact are much too low to overcome measured tissue surface tension values, meaning that cells would spontaneously round up due to the much stronger cortical force of the cytoskeleton, instead of attaching to each other (David et al., 2014; Maître et al., 2012; Stirbat et al., 2013; Youssef et al., 2011). This implied the existence of mechanism responsible for the regulated reduction of cortical tension at sites of cell-cell contact that would allow the adhesion strength to overcome the cortical tension in that area (Amack & Manning, 2012; David et al., 2014; Maître et al., 2012; Stirbat et al., 2013; Winklbauer, 2015). The antagonistic relation of cortical tension and adhesion strength creates an additional layer of complexity to the mechanics of cell adhesion (Manning et al., 2010; Wayne Brodland & Chen, 2000). The strength of the adhesion is thus not only affected quantitatively by the specific cell adhesion molecules that contribute, negatively or positively, to the binding energy, but also by any factor that alters the cortical tension in the contact area (Amack & Manning, 2012; Winklbauer, 2015). This explains why multiple known cytoskeletal regulators can also affect cell adhesion (A. K. Harris, 1976). This novel modification of the Differential Adhesion theory implies that cortex tension simply imitates the surface tension of liquids, since it requires active maintenance at the expense of metabolic energy (Clark et al., 2014, A. K. Harris, 1976; Winklbauer, 2015). Another significant application of this updated theory arises as a result of the increased molecular complexity of the mechanical mechanism controlling cell adhesion events. Differential adhesion can be generated by varying the cell surface density of a single adhesion factor, as well as by combining factors

of different specificities, affinities and expression levels, which was later confirmed experimentally (Duguay et al., 2003; Friedlander et al., 1989).

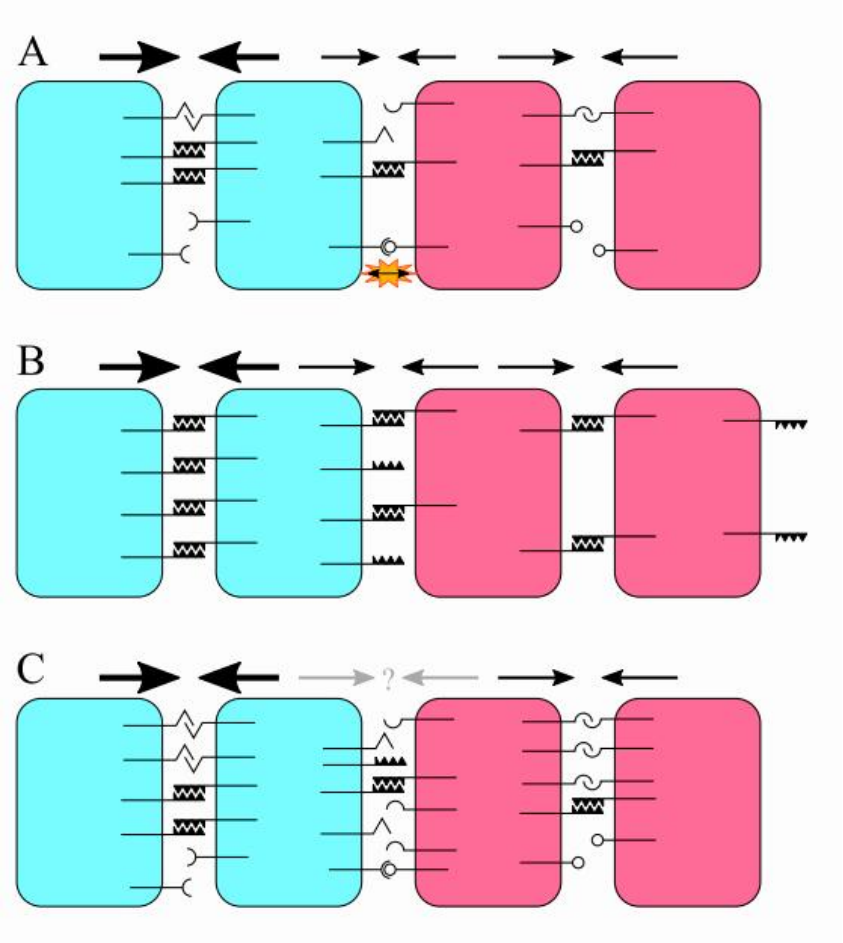


Fig. I.1. Differential affinities and expression levels of adhesion molecules can result in degrees of adhesion within and between tissues. (A) Separation due to expression of selective adhesion molecules. Ectodermal cell interfaces (blue-blue) show greater adhesion (large arrows) than heterotypic (blue-red) interfaces or mesodermal cell interfaces (red-red) (small arrows). This may be due to either a lack of interacting adhesion molecules at the heterophilic interfaces (curved and angular receptors) and intra-mesodermal interfaces (sawtooth receptors) or due to repulsive signaling from heterophilic interactions (cup and circle receptors). A combination of adhesion and repulsion may lead to labile adhesions between cells. (B) Separation due to differential expression of identical adhesion molecules. Cells expressing more molecules are more cohesive and bind to one another more strongly (blue; sawtooth receptors), preventing mixing with less cohesive cells that express fewer adhesion molecules (red; sawtooth receptors). (C) To predict adhesion between different cell types, every interaction between homophilic and heterophilic adhesion molecules or repulsion factors must be quantified. Arrows represent adhesiveness between cells. *Figure and legend copied from Winklbauer & Parent, 2017.*

Chapter II

Molecular origins of cell adhesion

The arrangement of particular cell types in space and time is a prerequisite for animal development, and spatial patterns need to arise in a robust manner. The interaction of cells with their surrounding micro-environment via adhesion molecules, as well as the release and adsorption of soluble substances, largely remodels biological tissues (Singh et al., 2015). Cell surfaces display an immense range of cellular adhesion molecules (CAMs), which mediate and regulate cell-cell and cell-ECM interactions and hold the trillions of cells in the organism tied together. CAMs include major classes of transmembrane proteins, namely selectins, cadherins, integrins, immunoglobulin superfamily (Ig-SF) proteins, leucine rich repeat (LRRs) proteins and teneurins.

Selectins

Selectins are a family of cell adhesion molecules consisting of three Ca^{2+}-dependent lectins: E-selectin (expressed on the surfaces of endothelial cells), L-selectin (expressed on the surfaces of leukocytes, monocytes, neutrophils and eosinophils), and P-selectin (expressed platelets and leukocytes). All selectins have similar structure, which includes an epidermal growth factor-like domain, a transmembrane domain, a repeat sequence, and a cytoplasmic tail (McEver, 2015). Selectins mediate tumor cell interactions by activating integrins and secreting pro-metastatic chemokines in the tumor microenvironment (Läubli & Borsig, 2010). Furthermore, signaling from selectin-selectin ligand interactions between endothelial and tumor cells can result in tumor cell extravasation (Witz, 2008) and is also involved in constitutive lymphocyte homing (Angiari, 2015). Selectins play a major role in physiological and pathological processes such as inflammation, immunological response, and cancer metastasis.

Cadherins

The cadherin superfamily is a multigene family of proteins with diverse structures and functions. The four main cadherins subfamilies include classical cadherins, desmosomal cadherins, protocadherins, and atypical cadherins,. Among them, classical cadherins are the best studied, while the others are less extensively characterized. The prominent classical cadherins are epithelial (E)-, neuronal (N)-, and vascular endothelial (VE)-cadherins. Cadherins localize at adherens junctions at cell-cell adhesion interfaces. (Biswas & Zaidel-Bar, 2017).

Although cadherins are main mediators of cell adhesion, a direct relationship between cadherin expression and differential adhesion in a morphogenetic process has yet to be shown *in vivo*. For

example, *Xenopus laevis* embryos are extremely tolerant of experimentally induced C-cadherin-based adhesion differences during gastrulation, in spite of the fact that these changes are sufficient to promote cell sorting *in vitro* (Ninomiya et al., 2012). Nonetheless, Megason and colleagues have recently provided cadherin-involving evidence for the differential adhesion model. They demonstrate that cell type-specific combinatorial expression of different classes of cadherins (N-cadherin, cadherin 11, and protocadherin 19) results in homotypic preference *ex vivo* and patterning robustness *in vivo* in the zebrafish spinal cord. They directly measured adhesion forces and preferences for three types of endogenous neural progenitor cells and revealed that the sonic hedgehog morphogen gradient regulates the differential adhesion pattern (Tsai et al., 2020).

Integrins

Integrins are a broad family of type-I transmembrane glycoproteins found in all animals. They consist of non-covalently-linked αβ-heterodimeric receptors with a short cytoplasmic domain and a large extracellular domain. Mammalian cells possess integrins with 18 α-subunits (α1–11, αV, αIIb, αL, αM, αX, αD and αE) and eight β-subunits (β1–8), which are assembled into 24 known αβ-heterodimers (Campbell & Humphries, 2011). Integrins can interact either with components of the extracellular matrix, such as fibronectin or collagen, or with ligands on the surfaces of juxtaposed cells, thus being important in a plethora of cell behaviors that rely on cell adhesion. Integrins alter their conformational state due to inside-out activating signal, which in turn leads to high affinity for extracellular ligands. Ligand binding then creates an outside-in signaling cascade that forms a focal adhesion complex at the integrin cytoplasmic tail. This bi-directional signaling enables cells to rapidly respond to surrounding environmental changes (Fu et al., 2012). Additionally, integrins facilitate a connection between the ECM and the actin cytoskeleton (Tamkun et al., 1986). Integrin-mediated adhesion is controlled by a number of factors, including integrin receptors' specific affinity for their extracellular ligands, mechanical forces acting on the point of adhesion, intracellular partners that bind integrins into discrete adhesive structures, and intracellular trafficking of integrins (Iwamoto & Calderwood, 2015; Tamkun et al., 1986).

Immunoglobulin Superfamily

Ig-SF is the most numerous and diverse class of cell adhesion molecules. Ig-SF proteins are characterized by the presence of at least one Ig-like extracellular domain encompassing 70–110 amino acids that mediate calcium-independent cell adhesion (Matthäus et al., 2017). The adhesion events mediated by Ig-SF molecules such as intercellular adhesion molecule (ICAM), vascular cell adhesion molecule (VCAM), and platelet endothelial cell adhesion molecule (PECAM) are critical for important normal physiological processes like immunological recognition and inflammation, as well as pathological processes such as tumor metastasis.

Leucine Rich Repeat Proteins

Leucine rich repeats (LRRs) are widely common proteins, present in over 60,000 proteins, identified in all phyla. All LRR units can be divided into a highly conserved segment (HCS) and a variable segment (VS). The HCS part consists of eleven, LxxLxLxxNxL, or twelve, LxxLxLxxCxxL, residues, in which "L" is Leu, Ile, Val, or Phe, "N" is Asn, Thr, Ser, or Cys, and "C" is Cys, Ser or Asn. Three residues at positions 3 to 5 in the highly conserved segments form a short b-strand, and multiple b-strands stack in parallel to form an arc. Canonical LRRs have been classified into eight groups: RI-like, Cysteine Containing (CC), SDS22-like, IRREKO, Bacterial, Plant specific, Typical, and TpLRR (Kobe & Kajava, 2001; Matsushima et al., 2010).

Several LRR receptors have been shown to contribute to boundary formation between distinctly specified cell populations or direct planar polarity and polarized cell rearrangements. For example, the LRR proteins *capricious* (caps) and *tartan* (trn) contribute to Dorsal-Ventral (DV) boundary formation in the *Drosophila* wing disc. The selector gene, apterous, regulates the development of the DV compartment border in the wing disc and is required to facilitate inter-compartment signaling, as well as prevent inter-mingling of dorsal and ventral cells in the wing disc. In addition, Apterous also regulates the expression of Fringe and the Notch ligands Serrate and Delta in dorsal cells (Marco Milán & Cohen, 2003). The LRR protein Caps, on the other hand, promotes border development but does not activate Notch signaling. Caps and Trn, when produced in ventral cells, induce cells to extend processes toward dorsal cells and seek to sort-out into their compartment. Caps- and Trn-expressing cells, on the other hand, maintain ventral signaling characteristics and are unable to cross the barrier (M. Milán et al., 2001).

In a different example from *Drosophila*, three Toll family receptors, Toll-2, Toll-6 and Toll-8, expressed in overlapping stripes along the AP axis, direct planar polarity and polarized cell rearrangements during convergent extension of the germband. Patterned transcriptional inputs along the AP axis provide spatial cues that regulate the expression of these three Toll receptors. Toll-2, Toll-6, and Toll-8 disruptions severely affect planar polarity, cell intercalation, and convergent extension, and deleting one or two receptors disturbs planar polarity in different groups of cells, showing that these proteins have nonredundant and highly localized roles. These findings lend credence to the theory that planar polarity is produced by interactions between neighboring cells with varying degrees of Toll receptor activation (Paré et al., 2014). Furthermore, Toll-8 forms a molecular complex with the adhesion GPCR Cirl/Latrophilin, which is critical for the activation of junctional Myosin-II. Quantitative variations in Toll-8 transcription levels of neighboring cells can cause Cirl asymmetric localisation, which in turn causes Myosin-II asymmetric activation and results in cell polarization (Lavalou J, Mao Q, Harmansa S, Kerridge S, Lellouch AC, Philippe J, Audebert S, Camoin L, Lecuit T, 2020).

A recently identified sub-class of Toll genes, known as Long Toll Genes (Loto genes), appear to have crucial roles in the embryogenesis of arthropods. Loto genes are distinguished by their relatively high number of leucine-rich repeat elements (LRRs) in comparison to other Toll genes, as well as the fact that they are expressed in transverse stripes in all or a subset of segments, patterns similar to classical pair-rule genes (Benton et al., 2016; Janssen & Lionel, 2018).

Teneurins

Teneurins (TENs) are type II single transmembrane proteins found in all multicellular organisms. TENs are expressed in a variety of organs throughout embryonic development, including the heart and brain. (Levine A, Bashan-Ahrend A, Budai-Hadrian O, Gartenberg D, Menasherow S, Wides R, 1994; Lossie et al., 2005; Nakamura et al., 2013). They possess a large extracellular domain that is involved in several homophilic and heterophilic interactions (Leamey & Sawatari, 2014; Südhof, 2017; Woelfle et al., 2016).

TENs interact with *trans*-cellular latrophilins to regulate synapse formation and organization (Boucard et al., 2014). Moreover, TEN2 interacts with latrophilins and mediates *trans*-cellular signaling to modulate cAMP levels in neighboring cells, in mammalian cell cultures (Li et al., 2018). Studies in *Drosophila, were teneurin* was firstly identified as a pair rule gene, namely *odd Oz (odz),* revealed that homophilic cell-adhesion of TENs plays a key function in selecting the right targets between individual presynaptic and postsynaptic neurons. (Mosca et al., 2012). Recently, it was also shown that during convergent extension in Drosophila, the leucine-rich-repeat receptor Tartan and the teneurin Ten-m convey polarity signals at epithelial compartment borders and promote myosin accumulation at stripe boundaries. Although the mechanisms by which these different receptors interface with the actomyosin contractile machinery remain unknown, Tartan and Ten-m extracellular domains interact *in vitro*, and Tartan enhances Ten-m compartment boundary localization *in vivo*. Tartan and Ten-m play a vital role in the planar polarity of compartment boundary cells (Paré et al., 2019).

Molecular origins of cortical forces

Unraveling the forces that form and reshape multicellular structures is integral to our understanding of development. During morphogenesis, cells alter their morphology and give rise to a multitude of forms. The ability of cells to change their shape relies mainly on forces that are produced at cell surfaces, and are transmitted through cell interfaces. These forces, called cortical forces, are generated in the cell cortex which is a 50-nm to 2-mm thick layer of cytoskeleton underneath the cell membrane, rich in Myosin II, actin filaments, and actin-binding proteins. Cortical forces build up from a range of molecular mechanisms including Myosin II and actin filaments assembly, which are spatially and temporally controlled in the cell.

Myosin II assembly

Non-muscle Myosin II is a hexamer composed of two heavy chains, two essential light chains, and two regulatory light chains (RLC). Each of the two heavy chains includes a globular head domain that binds F-actin and ATP in the presence of actin filaments, and undergoes a mechanochemical cycle of binding, hydrolysis, and release of ATP. These steps are tightly coupled to filament binding, conformational change, and force production. Each heavy chain contains a tail domain where heptad repeat sequences interact to promote dimerization and form a rod-like a-helical coiled coil. Additional interactions of antiparallel coiled–coil domains mediate the self-assembly of Myosin II into bipolar minifilaments, comprising a few dozen of discrete motor heads. Myosin II is primarily regulated by phosphorylation of its regulatory light chains. When RLC is unphosphorylated, Myosin II adopts a folded conformation in which (i) binding of the two heads to one another mutually prevents ATPase activity and actin binding and (ii) head–tail interactions inhibit minifilament assembly. Upon RLC phosphorylation, Myosin II unfolds into a conformation, which relieves states (i) and (ii) (Rauzi & Lenne, 2011).

Actin assembly at the cell cortex

The actin cytoskeleton at the cell's periphery consists of complex structured networks of F-actin that are coupled with the plasma membrane. Actin filament growth relies first on the pool of actin monomers, which add to the barbed ends of existing filaments, allowing their fast growth. Because cytoplasm contains a high concentration of actin monomers, regulation by proteins distributed at the cortex is essential. The actin-related protein Arp2/3 complex and formins are primary examples of proteins that regulate actin assembly and initiate new actin filaments The Arp2/3 complex generates branching filaments that drive the leading edge of motile cells forward and assist in endocytosis. The Arp2/3 complex connects a new filament to an existing actin network. The new filament grows until a capping protein binds to the barbed end and stops expansion. Formins nucleate unbranched filaments and remain attached to their barbed ends as they lengthen, preventing capping proteins from attaching.

Formins help to generate actin bundles, which are present in filopodia and cytokinetic contractile rings (Pollard, 2007).

F-actin cross-linkers and the mechanical properties of actomyosin networks

Additional levels of F-actin assembly rely on F-actin cross-linkers, which offer mechanical anchoring points and are essential for the emergence of contractility in actomyosin networks. While Arp2/3 complex and formins nucleate and assemble actin in branched and unbranched filament, cross-linkers can organize filaments into different higher-order structures: loose/tight networks, orthogonal networks, parallel/antiparallel bundles. In vitro studies have demonstrated the wide range of mechanical properties that cross-linkers confer to actomyosin networks. Whatever the origins of forces (internally generated or externally applied), actomyosin networks have complex mechanical properties, which are dependent on timescale. To some extent, actomyosin networks behave like viscoelastic materials. On fast time scales relative to actin turnover and cross-links, they resist deformation like springs and restore their shape after the force is released. At slow strain rates, they can flow like fluids, given that cross-links in the networks have enough time to bind/unbind and to allow actin networks reshaping. Branched actin networks in vitro are intrinsically very stiff with elastic moduli of 1000–10,000 Pa (Chaudhuri et al., 2007). In vivo measurements indicate comparable properties. Stiff networks are important to produce and resist force at the cell periphery of crawling cells (Mullins et al., 1998; Pollard & Borisy, 2003; Svitkina & Borisy, 1999). In contrast, unbranched actin networks are intrinsically soft, unless organized in higher-order structures by cross-linkers. Cross-linkers affect the organization of actin networks depending on their kinetics and geometry (Fletcher & Mullins, 2010). Due to its strong coupling to actin-binding sites, the cross-linker fascin preferentially organizes formin-mediated actin filaments into rigid bundles to generate protrusive forces in filopodia. In contrast, actin cross-linkers such as a-catenin can stabilize either orthogonal networks or parallel bundles, depending on the kinetics of interactions (Wachsstock et al., 1994).

Cortical forces and adhesion structures in epithelia

Building up forces in actomyosin cortical networks requires anchoring points in the networks but also at cell surfaces in order to produce cell shape changes. In epithelial cells, cell–cell adhesion structures are essential to transmit internally generated forces to other cells through cell surfaces or to extracellular matrix, and to reshape cell contours. Between the different molecular components that mediate cell adhesion, cadherin-based structures which promote cell–cell adhesion are presented here (Albiges-Rizo et al., 2009; Geiger et al., 2009). Observations in both cultured epithelial mammalian cells (Angres et al., 1996; Kametani & Takeichi, 2007) and in early epithelia of nonvertebrates (Cavey et al., 2008; T. J. C. Harris & Peifer, 2004; Müller & Wieschaus, 1996; U. Tepass & Hartenstein, 1994) indicated that E-cadherin forms dense protein clusters, which are thought to represent clusters of homophilic dimers in transassociation (Kametani & Takeichi, 2007). At cell junctions, E-cadherin binds to b-catenin, a-catenin binds b-catenin, and mediates interactions with the actomyosin cytoskeleton (Abe & Takeichi, 2008;

Cavey et al., 2008). Despite the fact that the connections between E-cadherin/b-catenin and actin via a-catenin are dynamic (Drees et al., 2005; Yamada et al., 2005), several studies have shown that a-catenin is responsible for mechanical interactions between E-cadherin clusters and actin at adherens junctions. Another study shows that a-catenin recruits vinculin, an important actin-binding protein of adherens junctions, via force-dependent changes in a-catenin conformation (Yonemura et al., 2010). Unfolding of a-catenin would expose cryptic sites for vinculin binding, as it has been shown for talin in integrin-mediated adhesion (del Rio et al., 2009). This could allow biochemical amplification upon application of force.

Adhesion, on the other hand, has the potential to adjust forces by modulating actomyosin assembly. Actin bundles stabilize cadherin clusters at the terminal ends of cell contacting filopodia (Vasioukhin et al., 2000). Cadherin clusters, in turn, can regulate actin assembly via the Arp2/3 complex and formins (Helwani et al., 2004; Kovacs et al., 2002; Verma et al., 2004). Formin-1 binds a-catenin directly, whereas the Arp2/3 complex binds b-catenin in competition with a-catenin (Drees et al., 2005).

Chapter III

Cell sorting in development

Bona fide cell sorting is a rare event in development as tissues containing random mixtures of different cell types are difficult to find. Instead, tissues usually contain partially segregated populations of cells with different induced fates, which are separated into distinct domains by changing their relative positions to each other during the course of development. These processes contain some aspects of cell sorting, but also involve distinct processes such as tissue border formation, border maintenance, and border maturation. Initial evidence for cell sorting in development came from classical experiments on cell sorting *in vitro*, in which whole embryos were dissociated into single cells, and their aggregation and segregation into distinct tissues were documented (Davis et al., 1997; Townes & Holtfreter, 1955). Prominent examples of this include the reaggregation and segregation of germ layer progenitor cells from gastrulating Rana pipiens, Xenopus, and zebrafish embryos (Davis et al., 1997; Klopper et al., 2010; Ninomiya & Winklbauer, 2008; Townes & Holtfreter, 1955). These experiments revealed important information about the sorting potential of different cell mixtures. However, they do not provide evidence for a critical involvement of cell sorting in vivo.

Conserved mechanisms of cell sorting in different organisms

During development, progenitor cells of different tissues and organs, which are initially defined in partially overlapping domains, need to segregate into distinct compartments. The early "sorting out" of various progenitor cell populations into segregated compartments is thought to be directed by differences in adhesion, mechanical, and motile cell properties. The advent of various genetic tools has also allowed for the analysis of induced cell sorting *in vivo*. To date, this has primarily been achieved by generating genetic mosaic embryos through the transplantation of genetically modified cells or the stochastic/targeted induction of gene expression in a subset of cells within the embryos. The emergence of such molecular toolkits has facilitated the elucidation of the roles of cell sorting and segregation in various developmental processes, such as germ layer formation during zebrafish and *Xenopus* gastrulation, blastocyst development in mouse, limb bud formation in chicken, and boundary formation in *Drosophila* wing imaginal disk. Studying such exemplary developmental processes has shown that analogous developmental events in different organisms may follow distinct cell segregation and boundary formation mechanisms. Nonetheless, the molecular architecture and the key inducers of these developmental processes are widely conserved among animals.

Cell segregation during zebrafish and Xenopus gastrulation

Gastrulation, a key morphogenetic process that entails a set of evolutionarily conserved morphogenetic movements which ultimately result in the specification and formation of distinct cell layers, namely ectoderm, mesoderm and endoderm, differs in regard to the internalization movements of different cells types observed among different organisms. For example, in mouse and chick gastrulation single mesoderm and endoderm progenitor cells delaminate from the epiblast through a process termed single cell ingression (Chuai & Weijer, 2008; Vasiev et al., 2010). On the other hand, in zebrafish, progenitor cells internalize as a stream of coordinated cells, a process called coordinated cell ingression, while in Xenopus, they travel as a sheet of coherent cells, a kind of morphogenetic movement named involution (Shook & Keller, 2008; Winklbauer, 2009). Although the mode of cell internalization differs among organisms, in all cases mesoderm and endoderm progenitor cells adopt more mesenchymal and migratory traits and cannot intermix with the epiblast cells above once segregated. TGFb-type Nodal signals, which are considered to be crucial inducers of mesoderm and endoderm cell fate in vertebrates, are responsible for this shift in adhesion properties of internalizing cells (Schier, 2003, 2009).

Border formation during vertebrate somitogenesis and hindbrain segmentation

Wing development (Garcia-Bellido et al., 1973) and abdomen formation (Lawrence et al., 1978) in *Drosophila*, were the first developmental events used to study the emergence of boundary structures aimed at preventing the intermixing of cells. Naturally, this morphogenetic process has been extensively studied in various other developmental processes in both vertebrates and invertebrates (J. E. Cooke & Moens, 2002; McGrew & Pourquié, 1998; Ulrich Tepass et al., 2002). D/V boundary formation in the *Drosophila* wing disk has already been mentioned when discussing the roles of LRRs proteins Capricious and Tartan in cell adhesion. Other examples of analogous events are border formation during somitogenesis and hindbrain segmentation in vertebrates.

Somites are groups of mesoderm cells located on either side of the neural tube and the notochord in the developing vertebrate embryo. These cells give rise to all skeletal muscles of the body, the axial skeleton, and part of the dermis (McGrew & Pourquié, 1998). Somites are derived from the sequential segmentation of the presomitic mesoderm (PSM) by a clock and wave front mechanism (J. Cooke & Zeeman, 1976). The segmentation of each somite is followed by the establishment of a cell boundary between the newly formed somite from the PSM. Eph receptors and Ephrin ligands are important mediators of somite border formation, as they mediate repulsion of opposing border cells (Durbin et al., 1998; Watanabe et al., 2009). N-cadherin (Rachel M. Warga & Kane, 2007) and PAPC (Kim et al., 1998) are also expressed at the forming somite border in zebrafish and Xenopus, yet their function in somitogenesis remains unclear.

In another example from the vertebrates, the vertebrate hindbrain is transiently segmented into seven rhombomeres in a stereotypical manner with each rhombomere acquiring an unambiguous identity as

the neuroepithelium divides itself. Nonetheless, cells from odd-numbered rhombomeres preferably mix with cells from other odd-numbered segments compared to cells from even-numbered segments and when opposed experimentally, they form borders with even-numbered segments and tend to fuse with other odd-numbered segments. This sorting behaviour indicates that although cells within an individual rhombomere possess unique traits, there are also alternating morphological properties common to all rhombomeres (Fraser et al., 1990; Guthrie et al., 1993; Guthrie & Lumsden, 1991; Wizenmann & Lumsden, 1997). The unique identity of each rhombomeres also creates discrete boundaries between each pair of contacting rhombomeres. Several members of various ephrins and Eph receptor protein families are simultaneously expressed in overlapping, dynamically shifting patterns during rhombomere boundary formation in the mouse, Xenopus and zebrafish. Overall, there are commonalities in the expression of Eph receptors and their ligands in the hindbrain, yet areas of expression seem to vary throughout evolution.

Chapter IV

The crustacean *Parhyale hawaiensis* as a model system for arthropod development

Arthropods are an ideal taxon for studying the evolutionary diversification of developmental patterns due to their diverse body plans. Our understanding of developmental processes and their underlying genetic architecture is significantly impacted by the vast knowledge gained from studies in species such as Drosophila melanogaster and Caenorhabditis elegans, thanks to forward genetics methods in these model species. Only a few other animal species (such as Xenopus laevis, Gallus domesticus, Mus musculus, and Danio rerio) have received the same level of scrutiny as D. melanogaster and C. elegans.

As the closest outgroup to the insects, crustaceans may shed light on the ancestral mode of insect development, yet crustacean development is not as well understood as that of insects. The amphipod *Parhyale hawaiensis* (Dana 1853) is a crustacean species, whose morphological traits render it especially suitable for developmental, genetic, and evolutionary analyses (Rehm et al., 2009d).

The biology of *Parhyale hawaiensis*

P. hawaiensis has a worldwide, circumtropical, intertidal, and shallow-water marine distribution (Barnard, 1969). As a detritivore *Parhyale* feeds on decaying organic material and can aggregate in large population densities (>3000/m^2). Due to the nature of its habitat, it is also extremely tolerant to environmental changes in salinity and temperature. The founding population of laboratory cultured *P. hawaiensis* was established from an initial population isolated in 1997 from the marine filtration system of the John G. Shedd Aquarium in Chicago. *Parhyale* is a robust species maintained in plastic trays filled with artificially prepared saltwater to a specific gravity of 1.018-1.022, adequate air circulation and at temperatures of 25-27°C. *P. hawaiensis* colonies can be maintained on a diet of carrots, but are preferably bred on a combination of kelp and spirulina granules and a liquid mixture of fatty acids, plankton, and vitamins. *P. hawaiensis* has a life cycle of 7–8 weeks at 26°C. Embryogenesis spans approximately 10 days, and the emerging juvenile resembles a tiny adult. Females are distinguishable from males by their prominent gonads and smaller gripping appendages (limbs) in the thorax. Mating occurs by sexually matured *Parhyale* males capturing and holding the females until copulation takes place. Following sperm transfer, the female separates from the male, molts and oviposits her fertilized eggs in a ventral brood pouch (Stamataki & Pavlopoulos, 2016). *Parhyale's* body is divided along the AP axis into the head, thorax (pereon) and abdomen (pleon) regions. The orientation of the T4 and T5 (pointed anterior) walking appendages relative to those of T6-T8 (pointed posterior) is responsible for the name of the group: “amphipod” (Rehm et al., 2009d).

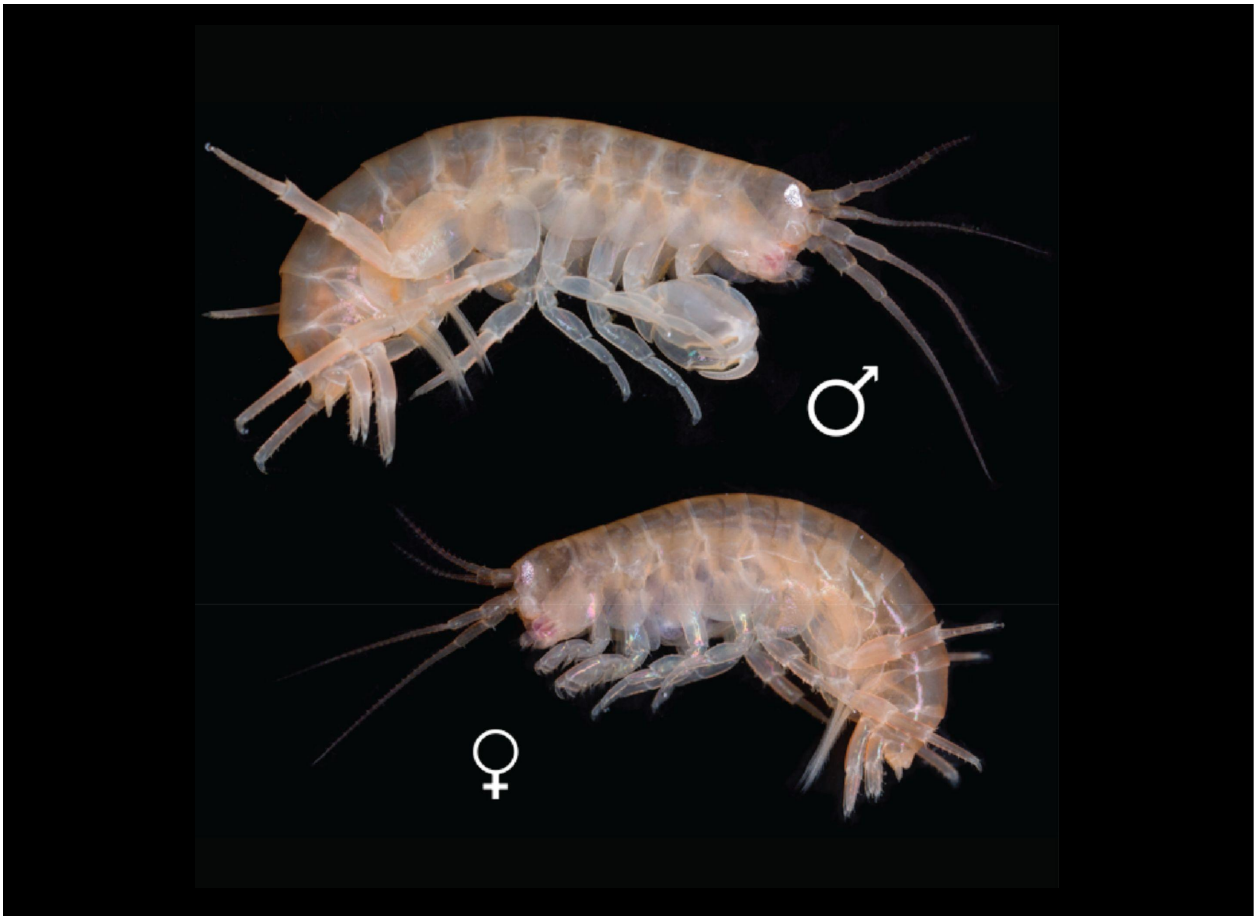


Fig. IV.1. Adult male (top) and female (bottom) forms of *Parhyale hawaiensis*. Sexually mature adults display sexual dimorphism: females are smaller and have visible ovaries, while males are larger and have larger chelipeds on T3. *Figure copied Dennis A. Sun, Nipam H. Patel 2019.*

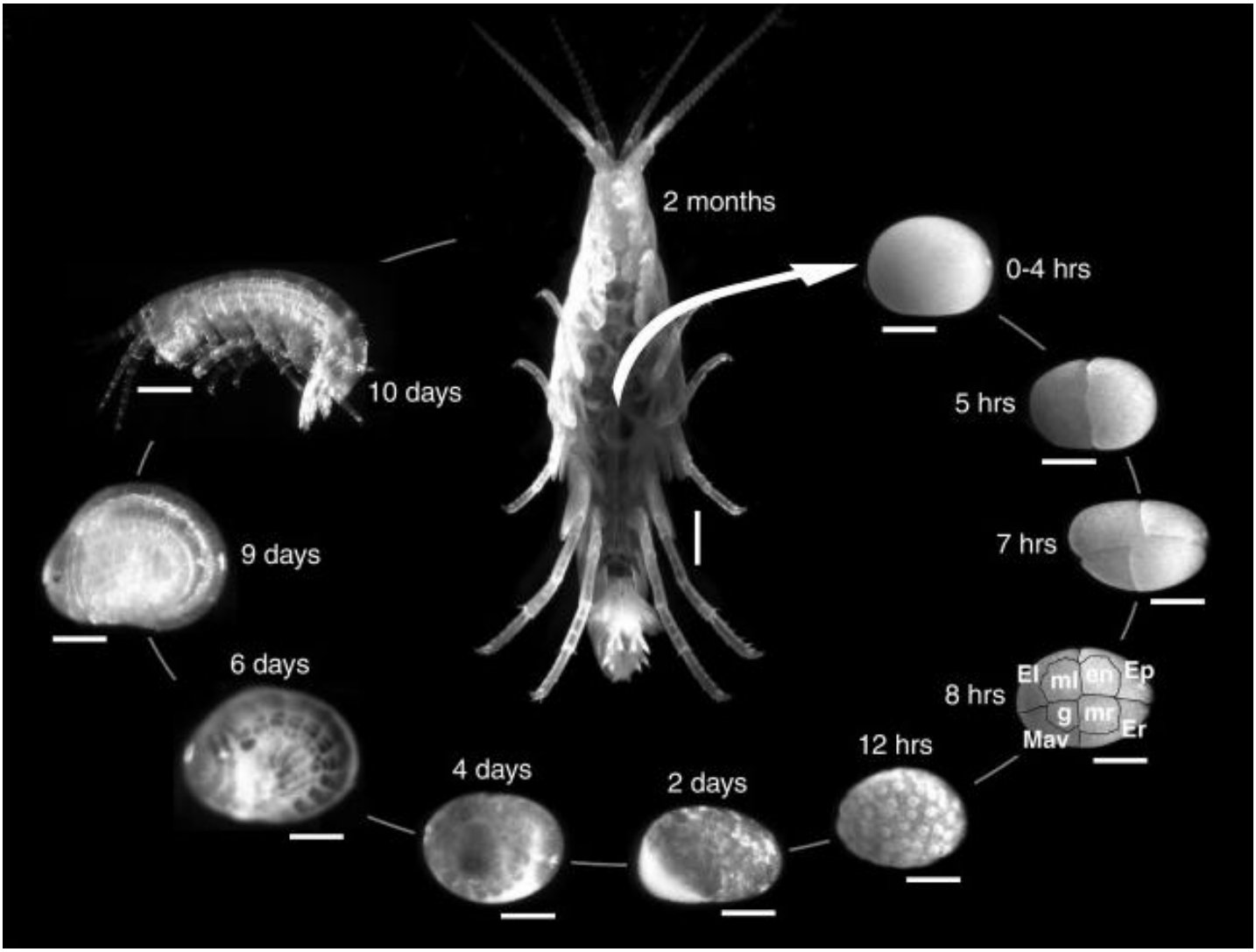


Fig. IV.2. *Parhyale* life cycle. *Parhyale* eggs can be dissected from the female's ventral brood pouch at any stage of development and can be cultured in seawater. During the first 8 hours after egg lay, each egg undergoes three total cleavages producing a stereotyped arrangement of four macromeres and four micromeres with restricted cell fates: the three El, Er and Ep macromeres give rise to the ectoderm, the fourth Mav macromere gives rise to the visceral and anterior mesoderm, the ml and mr micromeres form the rest somatic mesoderm, the en and g micromeres give rise to the endoderm and germline, respectively. Later divisions produce yolk-free cells (12 h) that aggregate ventrally and anteriorly to form the embryo rudiment (2 days). During subsequent segmentation stages, the embryo elongates posteriorly and the appendage buds develop in an anterior to posterior progression (4 days). Appendages continue to grow as the yolk gets sequestered in the developing midgut and the head region separates from the trunk (6 days). Organogenesis appears complete during the last days of embryogenesis when the pigmented compound eyes form (9 days). The hatchling that emerges from the egg looks like a miniature adult (day 10). It increases in size through successive molts and reaches sexual maturation about 2 months after egg lay. All scale bars are 200 mm except in the adult female that is 1000 mm. *Figure copied from* ***Stamataki & Pavlopoulos, 2016.***

Uses of the P. hawaiensis model system

Parhyale has been historically used in toxicological and ecological studies on a sporadic basis. However, a coordinated effort by multiple laboratories has expanded the experimental toolkit (which includes fixation, dissection, and antibody staining of embryos, *in situ* hybridization of labeled RNA probes to fixed *P. hawaiensis* embryos, and injection of *Parhyale* embryos in order to perform transposon- and/or CRISPR -based functional studies) and has established the amphipod P. hawaiensis as a new crustacean model organism (Browne et al., 2005, Extavour, 2005; Gerberding et al., 2002; Rehm et al., 2009a, 2009b, 2009c). Furthermore, the huge *Parhyale* genome has also been sequenced, assembled *de novo* and annotated *(http://www.ncbi.nlm.nih.gov/genome/15533)*. These advancements have enabled the study of appendage development and diversification in Parhyale (Hughes & Kaufman, 2002; Pavlopoulos & Averof, 2002). Moreover, because *Parhyale* eggs undergo a series of stereotyped, holoblastic cleavages, the resulting early fate-restricted blastomeres are easily identifiable by their size and relative position (Gerberding et al., 2002). These properties make the *Parhyale* embryo an ideal model to study the relative contribution of cell history versus cell communication in various developmental processes and investigate the biochemical nature of cell fate determinants. Furthermore, *Parhyale* are able to regenerate their amputated appendages (Kontarakis et al., 2011), which combined with the availability of transgenic lines labeling specific cell types has paved the way for ongoing studies in *Parhyale* that offer innovative approaches at the molecular and cellular basis of regeneration (Konstantinides & Averof, 2014). In addition, the annotation of the *P.hawaiensis* genome revealed that it contains the glycosyl hydrolase enzymes necessary to perform lignocellulose digestion. The investigation of 'wood-eating' in *Parhyale* may provide fresh insights on lignocellulose digestion in a sterile gut, devoid of microorganisms.

P. hawaiensis ectodermal germband formation

Parhyale follows a standard mode of germband development in malacostraca. The germband of the amphipod *P. hawaiensis* consists of an anterior head (naupliar) region and a posterior (post-naupliar) region that sequentially produces all trunk segments in an anterior-to-posterior progression system. The germband begins at the anteriorly at the level of the mandibular segment. It is composed of a grid of ectodermal cells arranged in rows and columns. Each row of ectodermal cells corresponds to a developing parasegment, numbered in an anterior to posterior fashion. With the exception of parasegments 0 and 1, each cell row undergoes two mediolateral, mitotic divisions, strictly oriented along the AP axis that transiently create a parasegment unit composed of four cell rows. Although highly stereotypical, there are some asymmetries visible between the right and left half of the developing germband. Moreover, division of a midline cell lags slightly relative to its neighbours (Browne et al., 2005).

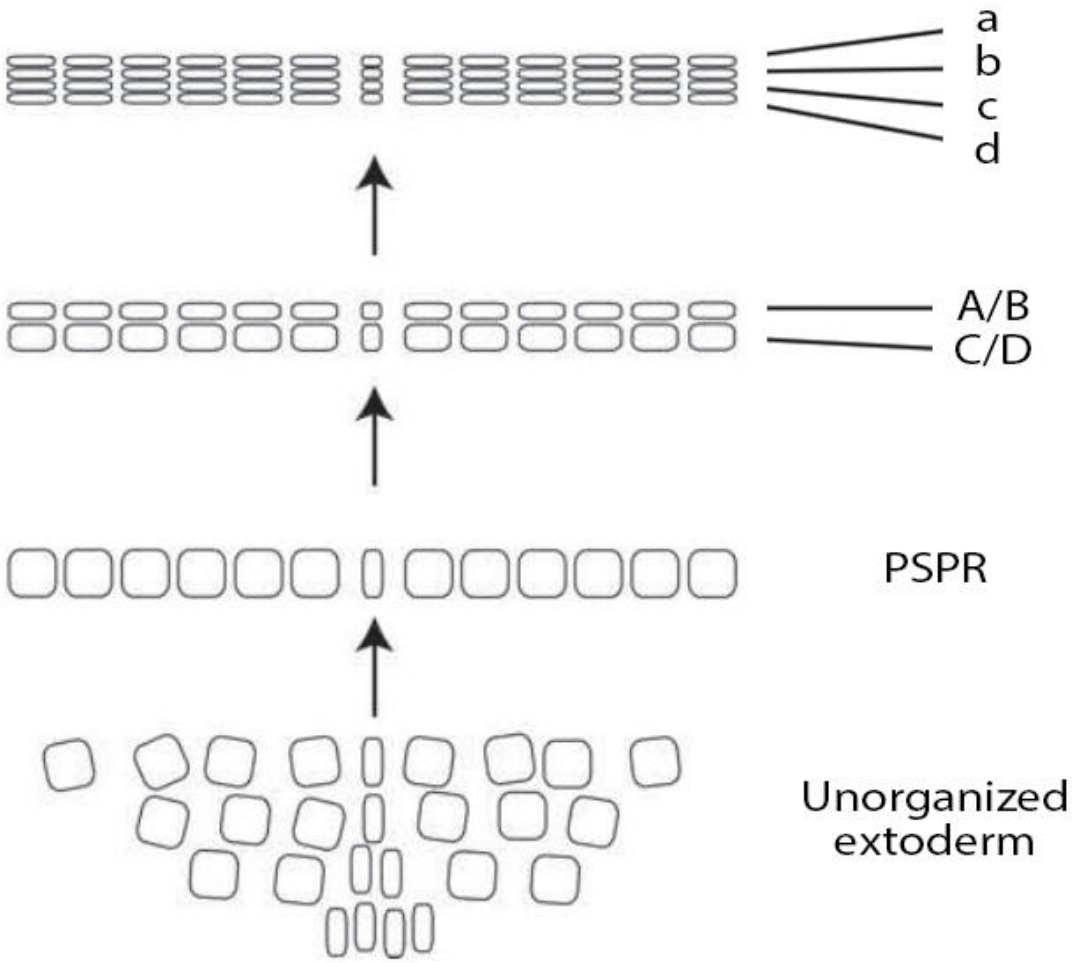


Fig. IV.3. Grid formation in *Parhyale.* The ectodermal grid of *Parhyale* forms as cells begin condensing on the anterior-ventral side of the egg beginning at approximately 24 hours of development. As these cells aggregate, they form ordered rows and columns. Each organized row is termed a parasegment precursor row (PSPR). Each PSPR undergoes a division to yield two daughter rows termed anterior-posterior (A/B and C/D). Each of these rows will then undergo an additional division to form a four cell row unit. These rows are named 'a', 'b', 'c', 'd'.

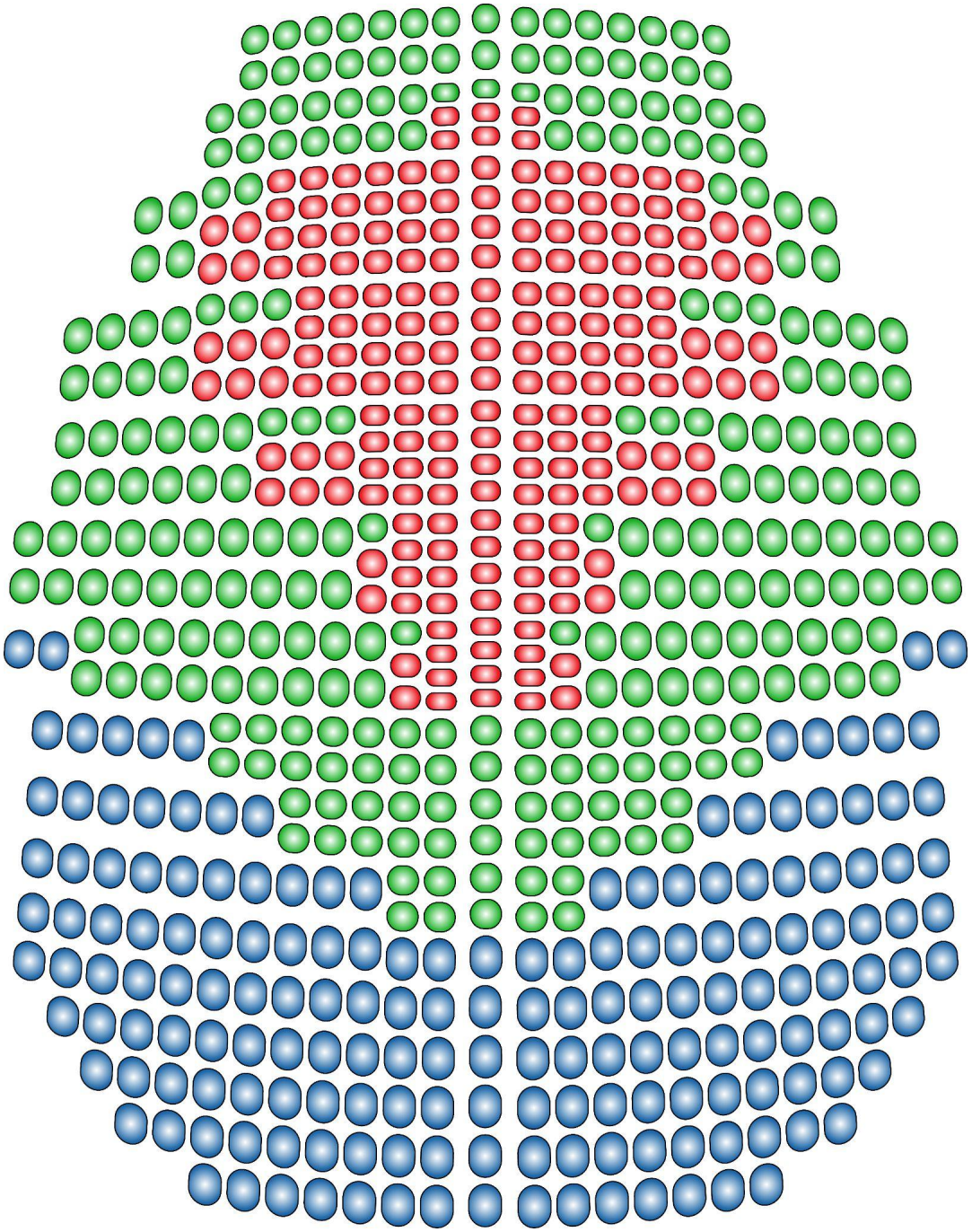


Fig. IV.4. Diagram of germband grid The ectodermal grid of Parhyale develops in an anterior-posterior progression. Outlined here is a typical germband at embryonic stage 16 showing the relative distribution of cells in each step of parasegment formation. Shown in red are progeny of PSPR cells that have divided twice ('a', 'b', 'c', or 'd'). Those in green have divided once (A/B or C/D) and those in blue are PSPRs. This diagram shows the complex pattern of stereotyped divisions that occur during germband elongation.

Chapter V

Summary of this study

Polarized cell rearrangements can change tissue proportions during embryogenesis contributing to tissue and organ morphogenesis. The genetic and physical basis of this universal process across metazoans has been best studied in embryos of the insect model *Drosophila melanogaster*. Elongation of the *Drosophila* anterior-posterior body axis, termed germband extension, is the result of convergent extension, which in turn is driven by dorsoventral cell intercalation. It has been known for many years that this process requires the patterned activity of upstream segmentation genes and downstream cytoskeletal and junctional networks. More recently, a subfamily of Toll receptors containing a high number of leucine-rich repeat elements (LRRs) has been identified as the long-sought intermediate link between regulatory and effector molecules controlling differential cell affinities and promoting oriented cell intercalations. Further reports have provided evidence for a conserved role of these Toll genes in embryonic morphogenesis across holometabolous and hemimetabolous insects, crustaceans, centipedes, spiders and even onychophorans; they are downstream targets of segmentation pair-rule genes, they are expressed in metameric stripes across the AP axis of the cell body, and are required for normal germband convergent extension movements.

During my Master's dissertation, I bioinformatically scanned the genome and transcriptome of the amphipod *P. hawaiensis* for the orthologs of the leucine-rich repeat transmembrane proteins tartan and capricious. It has been shown that one LRR receptor, named *Ph-LongTollA* (abbreviated to *Ph-LotoA*) exhibits stripy expression in embryos during segment formation like in other arthropods. However, unlike in other arthropods, these stripes are dynamically expressed only in newly formed segments. Furthermore, there is no evidence about cell intercalation events during *Parhyale* segmentation, raising questions about the role of LRR transmembrane proteins in this species and the actual cell behaviors that contribute to *Parhyale* germband morphogenesis.

In addition, I also identified the *Parhyale* teneurins transmembrane proteins. In *D. melanogaster* the teneurin Ten-m receptor interacts with tartan to direct planar polarity at compartment boundaries during convergent extension of the germband. Tartan was shown to recruit ten-m *in trans* and also inhibit its membrane localization *in cis.* Differences in Tartan and Ten-m levels provide cells at compartment boundaries with unique behaviors that prohibit crossing into neighboring domains and are essential for the maintenance of the border structure. The differences in the mode of germband extension between *Parhyale* and *Drosophila* combined with the fact that teneurins are possible ligands for LRR transmembrane proteins, call into question their possible role in *Parhyale* germband morphogenesis as well.

The characterization of these genes was accomplished using *D. melanogaster* proteins as query sequences. The putative protein motifs/ domains of the selected homologs were also examined using

the Expasy ScanProsite tool to scan the protein sequences against the PROSITE collection of motifs. In order to assess whether the multiple *tartan, capricious* and *teneurin* genes identified in the *Parhyale* genome are the result of gene duplication events that took place recently in the speciation process of Parhyale or date back in the arthropods' origin, a phylogenetics analysis of both teneurins and LRR genes from 18 species across the arthropods' phylum was completed using RAxML (Randomized Axelerated Maximum Likelihood), a tool for phylogenetic analysis using the maximum likelihood method. The expression analysis of these *Parhyale* receptors was performed using colorimetric *in situ* hybridization to decipher the expression pattern in the various stages of Parhyale development that germband extension spans. Last but not least, genome sequences of around 500-1000 base pairs (bp) from each gene, containing sites for CRISPR/Cas-based mutagenesis, were cloned and partially sequenced. Knock-out of these genes individually and in combinations will enable the comparison of the germband architecture between wildtype and mutant embryos, thus shedding light on their functional role during germband formation in Parhyale.

Chapter VI

Genome/ transcriptome search for leucine-rich repeat and teneurin proteins in *P.hawaiensis*

Materials and Methods

The gene, transcript and protein sequences of 18 leucine-rich repeat or teneurin transmembrane proteins were collected from the FlyBase database to be used as query sequences for the search of their orthologs in *Parhyale.* Each of the selected proteins were scanned using the Expasy ScanProsite tool against the PROSITE collection of motifs. Motifs with a high probability of occurrence were excluded from the scan and the scans were run in both high and low sensitivity settings. In addition, the protein sequences were locally pblasted against a blast database of *Drosophila* CDS sequences in order to identify their isoforms and possibly other proteins that show a high degree of similarity.

Local Blast databases were created using the available genome (version 5.0), transcriptome (version 3.0) and proteome (version 3.0) of *P.hawaiensis.* 14 out of the 18 selected *Drosophila* protein query sequences were blasted using blastp against the Parhyale proteome database. For every blast hit with high Bit score and significant e-value (< 1e-5) the respective transcript was mined from the *Parhyale* transcriptome and used in a nblast against the *Parhyale* genome blast database. These transcripts were also nblasted against the entire phylum of Arthropoda. In addition, the *Parhyale* protein sequences were scanned using the Expasy ScanProsite tool to identify predicted motifs.

Results

The genes acquired from FlyBase were the Toll family of receptors (*Toll-1,...,-9)*, *Tartan*, *Capricious, Tenascin major, Tenascin accessory, Chaoptin, Connectin, Fish-lips, Slit and Kekkon-1.* The predominant isoform of each protein was used to predict the motifs contained on each protein **(Table VI.1)**. All members of the Toll family receptors were predicted to have a TIR motif, yet only 7 out of 9 had LRR motifs, ranging in number from 7 to 25, with the predictions for Toll-3/MstProx and Toll-9 showing no LRR motifs. The Tenascin major and Tenascin accessory proteins were shown to contain 4 EGF-like motifs, while Fish-lips had 7 such motifs, a Laminin-G motif, a C-terminal cysteine knot motif and 10 LRR motifs. Chaoptin was shown to have 29 LRRs and 1 ARM motif. Connectin had 8 LRR motifs. Slit had 20 LRRs, 3 Fibronectin type-III motifs and 5 Ig-like motifs. Lastly, Kekkon-1 had 6 LRR motifs and 1 Ig-like motif being predicted.

Gene Symbol	Gene Name	FlyBase CG#	LRR motif	TIR motif	Other motifs
Tollo	*Toll-8 / Tollo*	CG6890	23	1	-
18w	*Toll-2 / 18-wheeler*	CG8896	25	1	-
Toll-6	*Toll-6*	CG7250	24	1	-
Toll-7	*Toll-7*	CG8595	23	1	-
Tl	*Toll-1 / Toll*	CG5490	13	1	-
MstProx	*Toll-3 / MstProx*	CG1149	0	1	-
Toll-4	*Toll-4*	CG18241	7	1	-
Tehao	*Toll-5 / Tehao*	CG7121	8	1	-
Toll-9	*Toll-9*	CG5528	0	1	-
Ten-m	*Tenascin major*	CG5723	0	0	EGF-like: 4
Ten-a	*Tenascin accessory*	CG42338	0	0	EGF-like: 4
trn	*Tartan*	CG11280	12	0	-
caps	*Capricious*	CG11282	12	0	-
chp	*Chaoptin*	CG1744	29	0	ARM: 1
Con	*Connectin*	CG7503	8	0	-
Fili	*Fish-lips*	CG34368	10	0	EGF-like: 7 Laminin G: 1 C-terminal cystine knot: 1
sli	*Slit*	CG43758	20	0	Fibronectin type-III: 3 Ig-like: 5
kek1	*Kekkon 1*	CG12283	6	0	Ig-like: 1

Table VI.1. Gene symbol, name and FlyBase CG ascension number of the leucine-rich repeat or teneurin transmembrane proteins used as query sequences for the search of their orthologs in *Parhyale.* The number of Leucine-rich repeats (LRR), Toll/interleukin-1 receptor (TIR) homology domain and various other motifs, as predicted by the Expasy ScanProsite tool are also shown.

Toll receptors (Toll-1,...,-9), Tartan, Capricious, Tenascin major, Tenascin accessory and Chaoptin protein sequences were pblasted against the *Parhyale* proteome (**Table VI.2**). Each query sequence had about 4 hits with high Bit score values and e-value < 1e-5. For each of the top 10 blast hits for every query sequence, the protein motifs were also predicted using the ScanProsite tool (data not shown). However, since the expression of Toll receptor had already been studied in *Parhyale,* only the top hits of the Tartan, Capricious, Tenascin major and Tenascin accessory query sequences were selected to study their expression during germband formation. This included, 4 transcripts that appeared as top hits for both Dm.Tartan and Dm.Capricious query sequences, although in a different order (phaw_30_tra_m.004692, phaw_30_tra_m.019648, phaw_30_tra_m.020608 and phaw_30_tra_m.023635), and 3 additional transcripts that were the top hits for both Dm.tenascin-major and Dm.tenascin-accessory (phaw_30_tra_m.015575, phaw_30_tra_m.006773 and phaw_30_tra_m.00161). The predicted motifs for the protein sequences coded by these 7 transcripts are listed in (**Table VI.3**). Furthermore, using the Integrative Genomics Viewer (IGV) to visualize the *Parhyale* genome, it was confirmed that all 7 transcripts are transcribed by different loci in the genome, neither of which is adjacent to any of the others.

Last but not least, blasting these *Parhyale* transcripts against the transcriptome of other arthropods, reveals that in other species, the top results returned by the nblast have been identified as homologous to the same *Drosophila* transcripts that were initially used as query sequences in the case of *Parhyale* teneurin transcripts and as leucine-rich repeat-containing proteins in the case of *Parhyale* tartan and capricious transcripts (**Table VI.4 and VI.5**).

Gene Name (*Dm*)	Total Blast Hits in *Ph* Proteome*	Top Transcripts in *Ph*	Bit score	E-value
Toll-8 / Tollo	51	phaw_30_tra_m.014482	1096	0
		phaw_30_tra_m.020724	884	0
		phaw_30_tra_m.000707	746	0
		phaw_30_tra_m.022043	661	0
Toll-2 / 18-wheeler	49	phaw_30_tra_m.014482	826	0
		phaw_30_tra_m.000707	770	0
		phaw_30_tra_m.020724	747	0
		phaw_30_tra_m.022043	642	0
Toll-6	52	phaw_30_tra_m.000707	990	0
		phaw_30_tra_m.014482	976	0
		phaw_30_tra_m.020724	857	0
		phaw_30_tra_m.022043	692	0
Toll-7	52	phaw_30_tra_m.014482	887	0
		phaw_30_tra_m.000707	813	0
		phaw_30_tra_m.020724	778	0
		phaw_30_tra_m.022043	668	0
Toll-1 / Toll	41	phaw_30_tra_m.004118	454	1.00E-142
		phaw_30_tra_m.014482	295	4.00E-83
Toll-3 / MstProx	5	phaw_30_tra_m.004118	185	6.00E-48
		phaw_30_tra_m.014482	137	9.00E-33
Toll-4	10	phaw_30_tra_m.004118	216	2.00E-57
		phaw_30_tra_m.014482	150	1.00E-36
Toll-5 / Tehao	25	phaw_30_tra_m.004118	239	8.00E-67
		phaw_30_tra_m.014482	186	1.00E-48
		phaw_30_tra_m.000707	150	3.00E-37
Toll-9	26	phaw_30_tra_m.012432	108	7.00E-24
		phaw_30_tra_m.022670	84.7	2.00E-16

		phaw_30_tra_m.014482	82	8.00E-16
		phaw_30_tra_m.000707	80.5	3.00E-15
Tenascin major	31	phaw_30_tra_m.015575	2524	0
		phaw_30_tra_m.006773	1501	0
		phaw_30_tra_m.001611	694	0
Tenascin accessory	32	phaw_30_tra_m.006773	1257	0
		phaw_30_tra_m.015575	1071	0
		phaw_30_tra_m.001611	1014	0
Tartan	49	phaw_30_tra_m.004692	233	8.00E-67
		phaw_30_tra_m.019648	232	4.00E-66
		phaw_30_tra_m.020608	206	4.00E-58
		phaw_30_tra_m.023635	201	1.00E-54
Capricious	50	phaw_30_tra_m.004692	234	8.00E-69
		phaw_30_tra_m.019648	225	2.00E-65
		phaw_30_tra_m.023635	207	1.00E-58
		phaw_30_tra_m.020608	188	3.00E-53
Chaoptin	45	phaw_30_tra_m.022521	375	2.00E-107
		phaw_30_tra_m.023009	211	2.00E-55
		phaw_30_tra_m.009982	210	4.00E-55
		phaw_30_tra_m.022670	198	2.00E-51

Table VI.2. Total number of blast hits for each query protein. Column N_o 3 shows the names of *Parhyale* transcripts that were considered to be significant (high Bit-score, low E-value) blast hits for each gene. Inside the red border are the *Parhyale* transcripts with the query sequence used to identify them whose expression during germband formation was analyzed via *in situ* hybridization.

Gene name (Dm)	*Parhyale* transcript	Size (bp)	LRR motif	TIR motif	Other motifs
Tenascin major / accessory	phaw_30_tra_m.015575	9554	0	0	EGF-like: 3
	phaw_30_tra_m.006773	10198	0	0	EGF-like: 4
	phaw_30_tra_m.001611	5265	0	0	-
Tartan / Capricious	phaw_30_tra_m.004692	4606	8	0	-
	phaw_30_tra_m.019648	8386	12	0	-
	phaw_30_tra_m.020608	6585	13	0	-
	phaw_30_tra_m.023635	4702	12	0	-

Table VI.3. The predicted motifs for the protein sequences coded by the 7 *Parhyale* transcripts that were selected for further analysis, according to the Expasy ScanProsite tool.

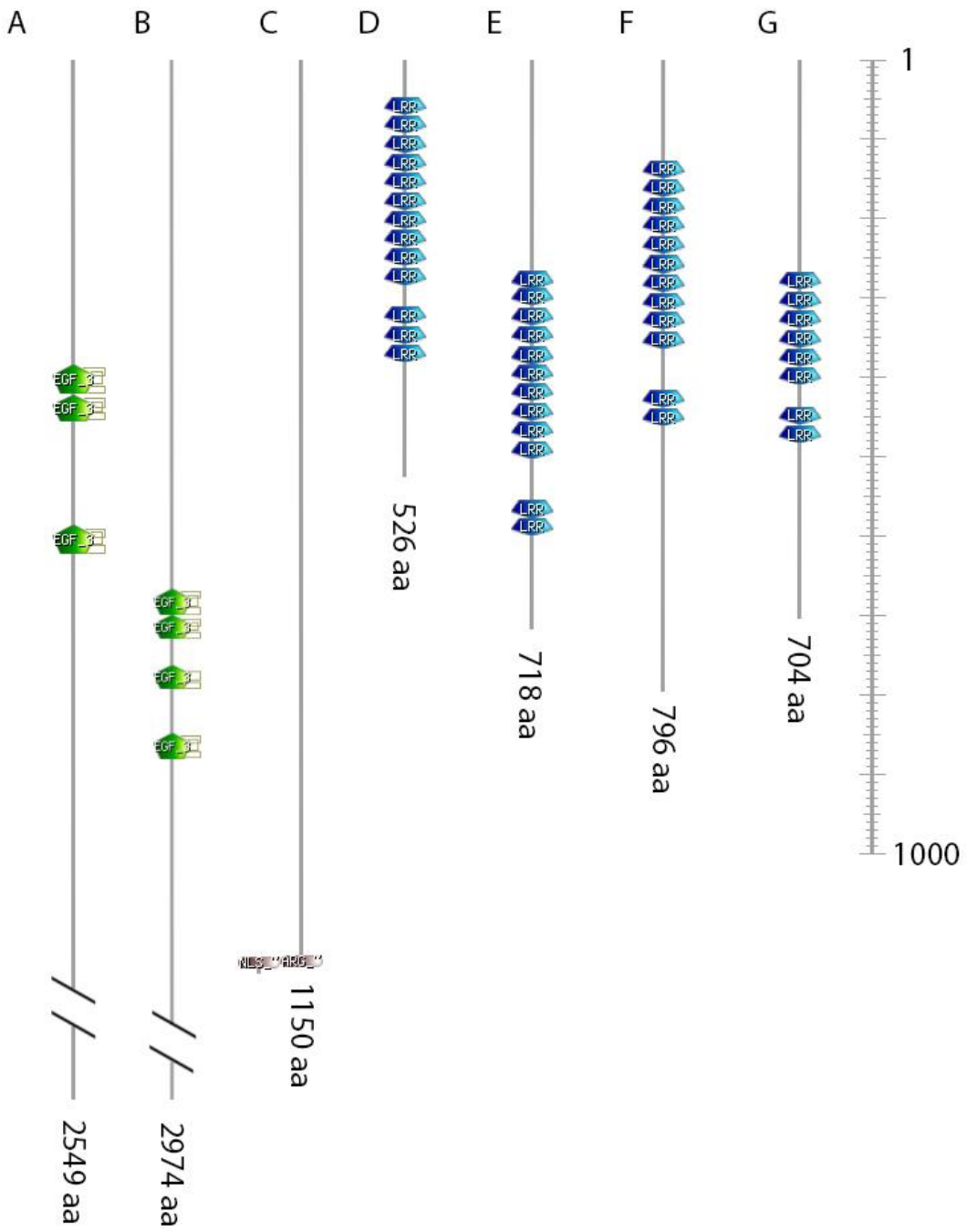


Fig. VI.1. Graphical representations of domains and length of the 7 *Parhyale* proteins selected for further analysis, according to the Expasy ScanProsite tool. (A) phaw_30_tra_m.015575, (B) phaw_30_tra_m.006773, (C) phaw_30_tra_m.001611, (D) phaw_30_tra_m.020608, (E) phaw_30_tra_m.019648, (F) phaw_30_tra_m.023635, (G) phaw_30_tra_m.004692.

Teneurins Orthologs						
***Parhyale* Transcript**	**Description (Top 5 Hits)**	**Organism**	**Total score**	**Query cover**	**E-value**	**Per. Iden.**
phaw_30_tra_m.015575	PREDICTED: teneurin-m-like	Hyalella azteca	3593	82%	0	86.69
	LOW QUALITY PROTEIN: teneurin-m-like	Penaeus vannamei	2672	82%	0	64.96
	teneurin-m isoform X4	Pseudomyrmex gracilis	2657	96%	0	49.79
	teneurin-m isoform X3	Pseudomyrmex gracilis	2656	96%	0	49.79
	teneurin-m isoform X1	Pseudomyrmex gracilis	2655	96%	0	49.79
phaw_30_tra_m.006773	PREDICTED: teneurin-a-like	Hyalella azteca	3870	99%	0	78.39
	PREDICTED: teneurin-a-like	Hyalella azteca	2848	90%	0	53.85
	PREDICTED:teneurin-a	Neodiprion lecontei	2249	86%	0	44.24
	teneurin-a isoform X4	Solenopsis invicta	2248	86%	0	43.79
	PREDICTED: teneurin-a-like isoform X3	Apis dorsata	2246	91%	0	42.85
phaw_30_tra_m.001611	hypothetical protein HAZT_HAZT008617	Hyalella azteca	2101	0.95	0	0.8932
	PREDICTED: teneurin-a-like	Hyalella azteca	1448	99%	0.00E+00	0.5338
	PREDICTED: teneurin-a isoform X6	Polistes dominula	1083	99%	0.00E+00	0.4159
	PREDICTED: teneurin-a isoform X5	Polistes dominula	1083	99%	0.00E+00	0.4159
	PREDICTED: teneurin-a isoform X3	Polistes dominula	1082	99%	0.00E+00	0.4154

Table VI.4. *Parhyale tenascin-major* and *tenascin-accessory* orthologs in Arthropoda. Per. Iden. stands for Percent Identity.

Tartan/ Capricious Orthologs						
***Parhyale* Transcript**	**Description (Top 5 Hits)**	**Organism**	**Total score**	**Query cover**	**E-value**	**Per. Iden.**
phaw_30_tra_m.004692	PREDICTED: leucine-rich repeat and immunoglobulin-like domain containing-NOGO receptor-interacting protein 4	Hyalella azteca	873	0.88	0	0.7098
	leucine-rich repeat neuronal protein 2-like	Penaeus vannamei	381	0.69	0	0.4271
	Platelet glycoprotein V	Armadillidium vulgare	362	0.72	0	0.4288
	Insulin-like growth factor-binding protein complex acid labile chain	Penaeus vannamei	339	0.69	0	0.3984
	Leucine-rich repeat neuronal protein 2	Portunus trituberculatus	300	0.59	0	0.4043
phaw_30_tra_m.019648	PREDICTED: leucine-rich repeat-containing protein 15-like	Hyalella azteca	803	0.99	0	0.5766
	insulin-like growth factor-binding protein complex acid labile subunit	Penaeus vannamei	506	0.69	0	0.497
	Insulin-like growth factor-binding protein complex acid labile chain	Penaeus vannamei	503	0.68	0	0.4949
	Leucine-rich repeat neuronal protein 2	Portunus trituberculatus	508	0.67	0	0.5143
	insulin-like growth factor-binding protein complex acid labile subunit isoform X2	Trichogramma pretiosum	287	0.71	0	0.3503
phaw_30_tra_m.020608	PREDICTED: leucine-rich repeats and immunoglobulin-like domains protein sma-10	Hyalella azteca	612	0.77	0	0.7794
	chondroadherin-like protein	Penaeus vannamei	462	0.94	0	0.5337
	putative insulin-like growth factor-binding protein complex acid labile subunit isoform X1	Penaeus vannamei	409	0.77	0	0.5217
	chondroadherin-like protein	Penaeus vannamei	406	0.77	0	0.5319
	Insulin-like growth factor-binding protein complex acid labile chain	Penaeus vannamei	366	0.94	0	0.444
phaw_30_tra_m.023635	PREDICTED: leucine-rich repeat neuronal protein 4-like	Hyalella azteca	989	0.91	0	0.7517

	leucine-rich repeat neuronal protein 3-like	Penaeus vannamei	375	0.57	0	0.4968
	Insulin-like growth factor-binding protein complex acid labile subunit	Portunus trituberculatus	334	0.52	0	0.4764
	putative insulin-like growth factor-binding protein complex acid labile subunit isoform X1	Penaeus vannamei	332	0.52	0	0.4941
	insulin-like growth factor-binding protein complex acid l abile subunit	Penaeus vannamei	252	0.58	0	0.3424

Table VI.5. *Parhyale tartan* and *capricious* orthologs in Arthropoda. Per. Iden. stands for Percent Identity.

Chapter VII

Phylogenetics analysis of tartan, capricious, tenascin major and accessory in arthropods

Materials and Methods

A phylogenetic analysis containing the tartan/ capricious and tenascin-major/ -accessory proteins of 21 representative species in the Arthropoda group was carried out in order to examine whether or not the multiple Tartan/ Capricious and Tenascin genes identified in the Parhyale genome are the result of gene duplication events that took place recently in the speciation process of *Parhyale*. The genomes, transcriptomes and proteomes of 21 species (**Table VII.1**) were downloaded from the NCBI depository and used for the assembly of local blast databases. The CDS of the 7 *Parhyale* transcripts (4 tartan/ capricious and 3 tenascin-major/ -accessory homologs) were used as query sequences against the local blast database of every species. The same process was also repeated using the CDS sequences of *Drosophila* tartan, capricious, tenascin-major and tenascin-accessory proteins. The top hits for each species with both high Bit score and significant e-value (< 1e-5) were selected and the entire protein sequence was mined from the species proteome. In addition, IGV was used to visualise the genome of each species. For every transcript selected, the genomic loci was located to ensure that none of the selected protein sequences, in any of the species, are isoforms originating from the same gene, but instead are distinct proteins. The acquired sequences were separated into 2 groups, those derived when using tartan/ capricious as query sequence and those derived by using tenascin-major/ -accessory as queries.

As an outgroup to Arthropoda, 3 species were selected (**Table VII.2**); *Ramazzottius varieornatus* belongs in the phylum of Tardigrada, which together with the Onychophora phylum and the Arthropoda phylum make up the Panarthropoda clade. *Priapulus caudatus,* which belongs in the Priapulida phylum, and *Caenorhabditis elegans,* which belongs in the Nematoda phylum, were selected as members of the Ecdysozoa clade, which engulfs the Panarthropoda clade.

Order	Species	trn/ caps Hits	ten-m/ -a Hits
Homoptera	*Acyrthosiphon pisum*	1	2
Sessilia	*Amphibalanus amphitrite*	5	3
Diptera	*Anopheles gambiae*	3	2
Isopoda	*Armadillidium nasatum*	2	2
Isopoda	*Armadillidium vulgare*	1	2
Calanoida	*Calanus finmarchicus*	-	-
Cladocera	*Daphnia magna*	3	2
Cladocera	*Daphnia pulex*	-	-
Diptera	*Drosophila melanogaster*	2	2
Calanoida	*Eurytemora affinis*	5	3
Amphipoda	*Hyalella azteca*	-	-
Ixodida	*Ixodes scapularis*	3	3
Araneae	*Parasteatoda tepidariorum*	-	-
Amphipoda	*Parhyale hawaiensis*	4	3
Decapoda	*Portunus trituberculatus*	6	2
Decapoda	*Penaeus vannamei*	6	4
Harpacticoida	*Tigriopus californicus*	3	3
Coleoptera	*Tribolium castaneum*	2	2
Amphipoda	*Trinorchestia longiramus*	4	3
Geophilomorpha	*Strigamia maritima*	-	-

Table VII.1. Arthropoda species and their Order selected for the inference of *tartan/ capricious* and *tenascin-major/ -accessory* phylogenies. Statistically important number of blast hits for each group of genes is shown in the last two columns.

Phylum	Species	trn/ caps Hits	ten-m/ -a Hits
Nematoda	*Caenorhabditis elegans*	3	1
Priapulida	*Priapulus caudatus*	1	1
Tardigrada	*Ramazzottius varieornatus*	3	1

Table VII.1. Species and their phylum selected as the outgroup for the inference of *tartan/ capricious* and *tenascin-major/ -accessory* phylogenies. Statistically important number of blast hits for each group of genes is shown in the last two columns. These phyla together with the Panarthropoda make up the Ecdysozoa clade.

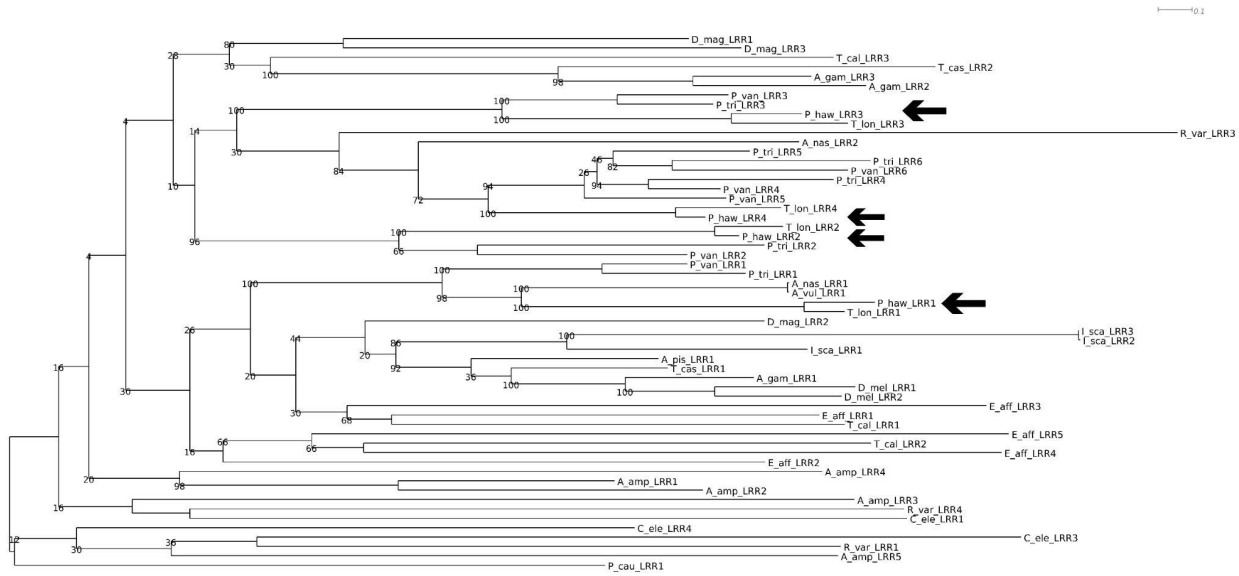
MAFFT (multiple alignment using fast Fourier transform) was used for the multiple alignment of the protein sequences within each group. BLOSUM62 was used as the scoring matrix for amino acid sequences and the program was set to align gappy regions. The sequence alignments were also curated using Gblocks, which eliminates poorly aligned positions and divergent regions, removing positions with gaps and subsequently alignment noise. However, this curation method resulted in the alignment of only a small section (~100 aa or ~15% of the total sequence) from each protein sequence.

RaxML (Randomized Axelerated Maximum Likelihood) was used for the inference of phylogenies with maximum likelihood. Given a randomly created tree, maximum likelihood provides probabilities of the sequences according to an evolution model. This "probability" is presented as a log likelihood, thus the less negative (or larger) the number, the greater the probability. The program was compiled to run in parallel on several cores that share the same memory, but also use vector instructions within each core/processor to increase the arithmetic instructions that can be executed per processor clock cycle, thus increasing the speed of the analysis. The GAMMA model of rate heterogeneity was used and the GAMMA Model parameters were estimated up to an accuracy of 0.1000000000 Log Likelihood units. Instead of choosing among a set of standard models with fixed transition rates, e.g. *DAYHOFF, JTT, BLOSUM62, MTMAM, LG* etc, the optimal protein substitution model with respect to the likelihood on a fixed, reasonable tree was automatically determined by RAxML according to the base frequencies drawn from the alignment. RAxML uses randomized stepwise addition parsimony trees as distinct starting trees and selects the tree with the best likelihood. A bootstrap analysis with 500 replicates was conducted to get support values for this tree. The resulting trees were visualized using Dendroscope, an interactive program for viewing and editing phylogenetic trees.

Results

Maximum likelihood phylogenetic analyses performed with RAxML identified both ancestral and more recent lineage-specific duplication events for these genes **(Fig. VII.1 and VII.2)**. Both trees were unrooted. *Parhyale* orthologs formed clusters with homologs from other crustacea. Dipteran homologs also formed separate clusters. The support values for the nodes within these clusters were high, which is strong evidence that the sequences to the right of the node cluster together to the exclusion of any other. However, there were also nodes with low support values. One possibility is that omitting certain orders within the Arthropods from the analysis, due to lack of reliable protein sequences, has inevitably removed some common ancestors and made it more difficult to connect some of the nodes. An alternative possibility is that the selected sequences needed to be further curated in order to reduce noise.

A



B

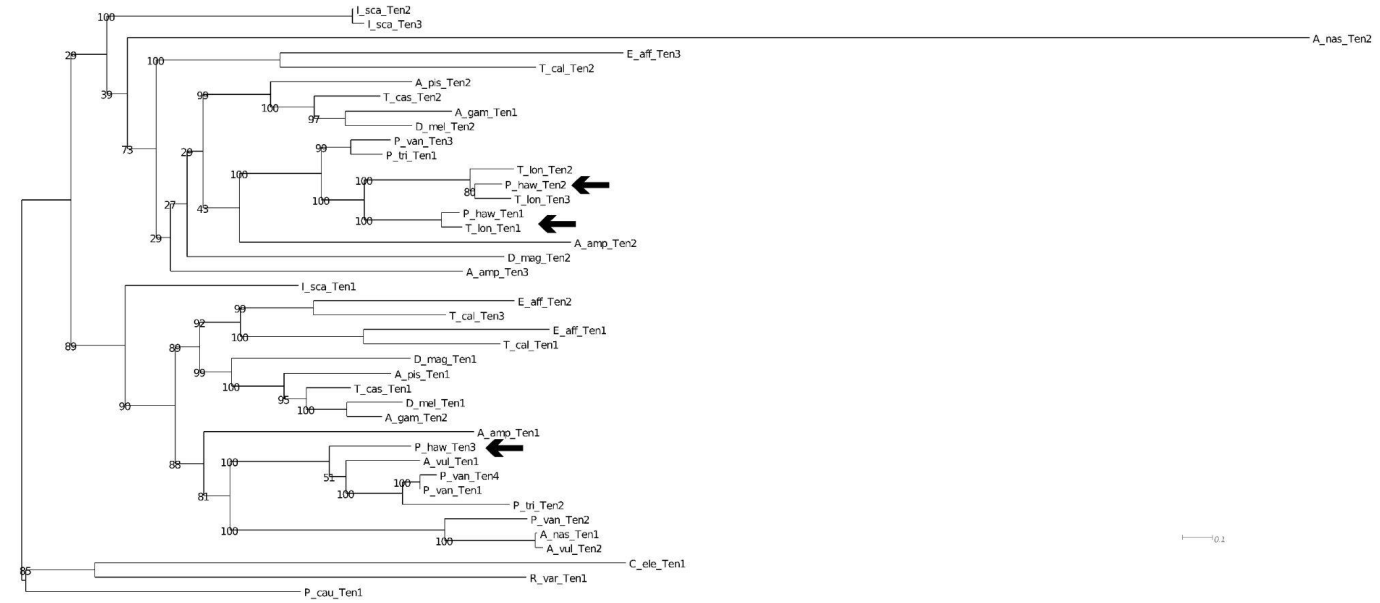


Fig. VII.1 and 2. Maximum likelihood based phylogenetic trees of arthropod (A) tartan and (B) teneurin receptors. Arrows indicate the position of Parhyale proteins.

Chapter VIII

Cloning of Parhyale tartan, capricious, tenascin major and accessory orthologs

Materials and Methods

tartan, capricious, tenascin major and accessory orthologs were partially cloned using PCR (*Phusion DNA polymerase*). For each of the transcripts, all possible ORFs were inferred bioinformatically. The largest ORF was selected as the main one, i.e. the ORF containing the information for the protein product translated from each transcript.

Forward and Reverse primers were designed to amplify sequences of 1500-2000 bp in size that contained the start of the ORF and extended as far as practically possible. cDNA derived from *Parhyale* embryos stage 13-19 was used as template for the PCR reactions, which were carried out in a gradient of annealing temperatures (55-70 ^{o}C). The PCR products were analyzed by gel electrophoresis, using the 1Kb^{+} DNA ladder as reference and were purified using Zymo's "DNA Clean & Concentrator" Kit.

Results

For every transcript, a single ORF was identified **(Table VIII.1 and VIII.2)**. Additional, smaller, possible ORFs (~300bp), that may be translated, producing small peptides, were also identified in some of the transcripts (1-4 small "ORFs" per transcript)

The PCR amplification of the desired fragments was successful for all 7 of the transcripts, albeit in different annealing temperatures.

tartan/ capricious Orthologs				
Transcript	**Name**	**Primers**	**ORF (bp)**	**Amplicon (bp)**
phaw_30_tra_m.004692	Ph_trn/caps_1	FWD: GCGAAGTTCTTCGTCACCAGTG	2115	1488
		REV: TCCTTTCTAACTCAGGGCAGCC		
phaw_30_tra_m.019648	Ph_trn/caps_2	FWD: GGAAATTTGGTCCGGCGAAGTT	2157	1968
		REV: GTGTCTCTCCAAGTGGTGACGA		
phaw_30_tra_m.020608	Ph_trn/caps_3	FWD: CTGGGATGGCTCCTCGTGAG	1581	1534
		REV: TTTCCTTAGGGCCGTGCATGTA		
phaw_30_tra_m.023635	Ph_trn/caps_4	FWD: CCAAGTGTGGTCGCCATTACAC	2391	1873
		REV: CCTTGTCTTCCTCCCAGTAGCC		

Table VIII.1. *tartan/ capricious* orthologs in *Parhyale.* For each transcript, the ORF was bioinformatically inferred and partially cloned.

tenascin -major/ -accessory Orthologs				
Transcript	**Name**	**Primers**	**ORF (bp)**	**Amplicon (bp)**
phaw_30_tra_m.015575	Ph_ten-m/-a_1	FWD: ACAGAAAGCGTCACTCAAAGCG	7650	2204
		REV: GTCGTTTGGTCGGAACATGGAC		
phaw_30_tra_m.006773	Ph_ten-m/-a_2	FWD: TGGAAGTACAGACGGGTGTGTG	8925	1924
		REV: TGAGGCGTTGGTAGCTCTTCAA		
phaw_30_tra_m.001611	Ph_ten-m/-a_3	FWD: GAACGTTCTGGGCGACATTACG	3453	1824
		REV: TTGTCCAAAGTTGACGATGCGG		

Table VIII.2. *tenascin-major/ -accessory* orthologs in *Parhyale.* For each transcript, the ORF was bioinformatically inferred and partially cloned.

Chapter IX

Expression analysis of *Parhyale tartan, capricious, tenascin major and accessory* during germband formation and maturation using *in situ* hybridization

Materials and Methods

In order to analyze the expression of *tartan, capricious, tenascin major and accessory in Parhyale,* the MEGAscript *in vitro* transcription kit (Ambion) was used to generate labeled RNA probes from the cloned DNA fragments. For each transcript, 2 antisense RNA probes, ~550 bp in length, were generated. Digoxigenin-11-uridine triphosphate (DIG-11-UTP) was used for labelling the RNA probes, by replacing UTP in the *in vitro* transcription reaction in a ratio of 35:65%. Probe solutions were stored at -20 °C at a concentration of 100ng/µl in Hyb.

Parhyale embryos were removed from the ventral brood pouch of gravid females that were allowed to develop in petri dishes with artificial seawater for specified time periods after their separation from males. Embryos at the required stages, S14 to S17, were collected in 2 ml eppendorf tubes, rinsed with cold deionized water, and then fixed in ice cold formaldehyde, while gently mixing on a nutator for specified times depending on the temperature, formaldehyde concentration and embryonic stage (Table &). Fixed embryos were rinsed 2-3 times in 1xPBS and washed by gently mixing on a nutator for at least 3 hrs. Washed embryos were transferred onto a Sylgard plate in a big drop of 1xPBS and dissected using fine tungsten needles. Finally, dissected embryos were dehydrate through a series of Methanol/ PBS washes (5 min each in 30% MeOH in 1xPBS, 50% MeOH in 1xPBS, 70% MeOH in 1xPBS, 90% MeOH in 1xPBS, 100% MeOH) and stored in 100% MeOH at -20°C.

Whole-mount in situ hybridization was performed as follows: Embryos were gradually rehydrated in PTw (1x PBS, pH 7.4 with 0.1% Tween-20) and fixed for 30 min in 4% formaldehyde in PTw. After 4 x 5 min washes in PTw, embryos were washed in Detergent solution (1.0% SDS, 0.5% Tween-20, 50mM Tris-HCl, 1mM EDTA pH8.0, 150mM NaCl) for 30 min and washed again 6 x 5 min in PTw. At this point, embryos were transferred from eppendorf tubes into baskets in PTw. Subsequent incubations and washes were carried out by moving the baskets from one well to the next in multi-well plates. Embryos were washed for 10 min in 1.5ml 50% Hyb/ 50% PTw, followed by another wash for 10 min in 1.5ml Hyb (50% Formamide, 5xSSC, 50µg/ml heparin, 0.25%Tween-20, 1% SDS, 100µg/ml Sonicated salmon sperm DNA). Then, the embryos were transferred to fresh Hyb solution and placed for 5 hrs inside a 65 °C incubator. The hybridization reaction was done by incubating stock probe solutions at 37-42 °C for 10 min to bring precipitated SDS in suspension and then diluting each probe to a final concentration 10 ng/µl in 1.5 ml

Hyb. The probes were denatured for 20 min at 80-90 °C, before being pipetted in a multi-well plate kept at 65 °C. Baskets with embryos were transferred into the Hyb/ probes solution and were incubated at 65 °C for 35 hrs. After the hybridization, embryos were washed by carefully transferring the baskets containing the embryos into multi-well plates with fresh Hyb, pre-heated at 65 °C. After 4 hrs of total washes, they were gradually brought to room temperature and washed with increasing concentrations of TBST/ Hyb to gradually replace the Hyb solution. After an hour of washes with TBST (125ml 1M Tris pH 7.5, 40g NaCl, 1g KCl, 5ml 10% Tween-20, ddH_20 up to 500ml) at RT, the embryos were washed for 1 hr in TBST+BSA, at 4 °C. The embryos were then incubated at 4 °C overnight with 1.6 ml of anti-DIG-AP (1:3000) antibody in TBST+BSA. The following day, the embryos were washed several times in TBST at room temperature, before washing in freshly prepared alkaline phosphatase (AP) buffer (5mM MgCl2, 100mM NaCl, 100mM Tris pH9.5, 0.1% Tween-20) for 3 x 5 min at room temperature. 1.6 ml of fresh 5-Bromo-4-chloro-3-indolyl phosphate (BCIP)/Nitro-Blue Tetrazolium (NBT) solution was added to a multi-well plate and the embryos were submerged in the solution. The incubation was done at room temp in the dark, because the substrates are photosensitive. The development of the colour reaction was periodically examined under the stereoscope (first check after 15-30 min, then depending on the signal after every hour). When the signal had developed to the desired intensity, the reaction was stopped with a few TBST washes. Following the reaction, embryos were washed with TBST, stained with DAPI and stored in 70% glycerol in 1xPBS, ready to be mounted for microscopy.

Results

The geometric precision and stereotypy of the grid pattern enables to study gene expression patterns with single-cell resolution and identify the exact stage of parasegment formation and maturation in which the genes of interest are expressed. My gene expression studies gave reproducible patterns with a high signal-to-noise ratio for the 3 out of the 4 Parhyale tartan genes (Ph_trn/caprs_1, -3 and - 4) and all 3 teneurin genes(Ph_ten-m/-a_1, -2 and -3). Ph_trn/caprs_2 colorimetric *in situ* experiments had low signal to noise ratio and did not allow for the identification of its expression pattern. Increasing both the RNA probes and the incubation time did alter these results. Surprisingly, all 6 genes exhibited essentially identical expression patterns in the posterior unsegmented region and in metameric stripes in the more anterior formed parasegments **(Fig. IX.1)**. Expression appeared to start in PSPR cells prior to their first mitosis. Post mitosis, it persists in the daughter A/B and C/D cells, as RNA levels were higher in the 2-row parasegments after they had divided for the first time. After the second mitotic division high expression levels persisted in the posterior two rows of cells in the 4-row parasegments. In this area, cells had higher expression levels as we got closer to the midline.The stripy and localized expression of these genes provides some first evidence that the tartan-teneurin system might have a conserved role, at least within the pancrustaceans, given the similarities in germband formation and number of homologs for these genes in members of this clade, in controlling cell behaviors at compartment boundaries. However, unlike in the Drosophila germband where asymmetric actomyosin distribution drives polarized cell intercalation Paré et al. 2019), no such cell movements are evident during Parhyale grid formation and maturation.

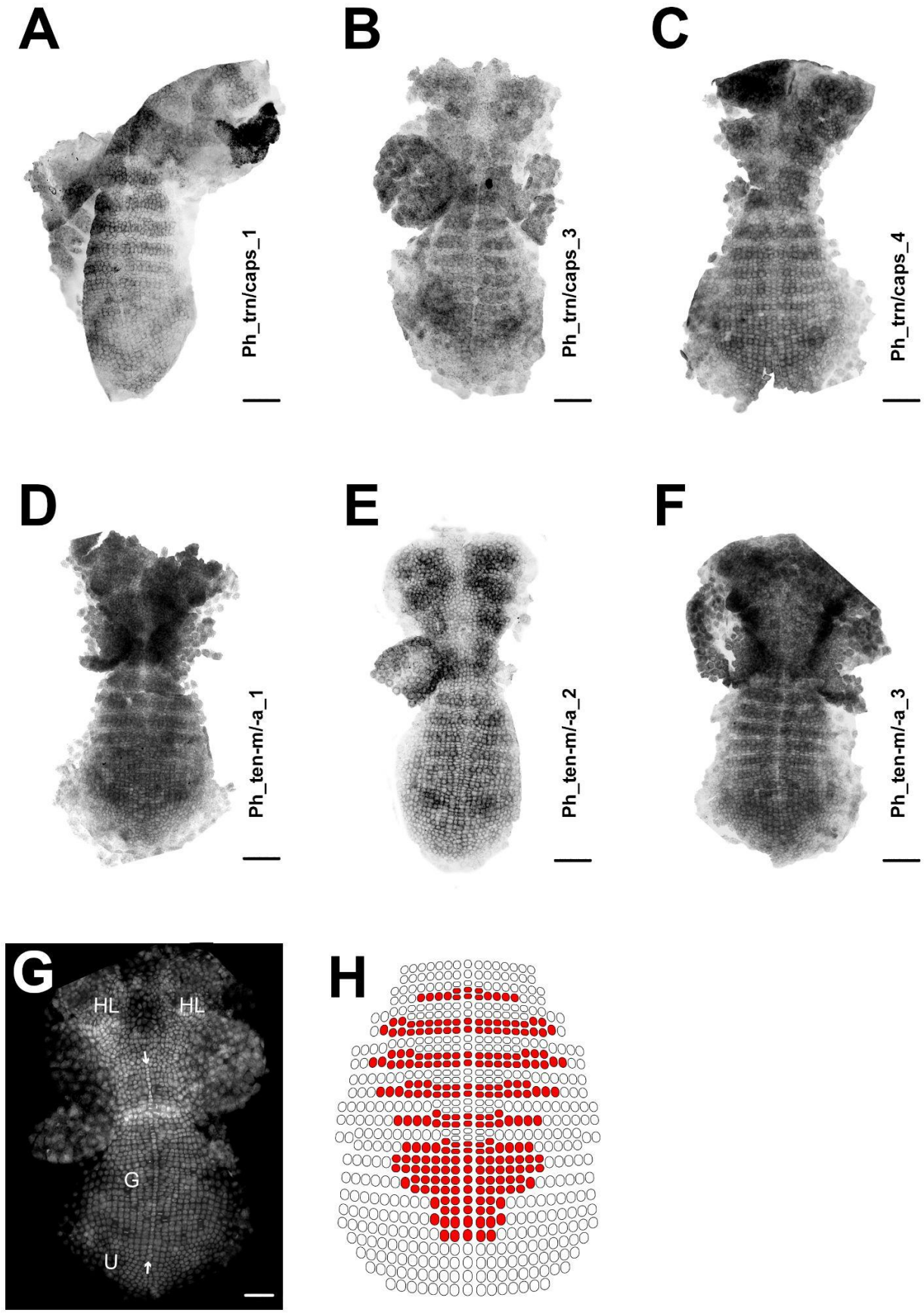

A
Ph_trn/caps_1
B
Ph_trn/caps_3
C
Ph_trn/caps_4
D
Ph_ten-m/-a_1
E
Ph_ten-m/-a_2
F
Ph_ten-m/-a_3
G
HL
HL
G
U
H

Fig. IX.1. Expression analysis of the Parhyale tartan and teneurin genes. (A-F) Brightfield images of Parhyale embryos at the mid germband stage S16 (ventral views like in G) stained by colorimetric *in situ* hybridization for **(A)** Ph_trn/caps_1, **(B)** Ph_trn/caps_3, **(C)** Ph_trn/caps_4, **(D)** Ph_ten-m/-a_1, **(E)** Ph_ten-m/-a_2 and **(F)** Ph_ten-m/-a_3 RNA expression. **(G)** Ventral view of a Parhyale embryo with DAPI-stained nuclei at the mid germband stage. From anterior (top) to posterior (bottom) the embryo is organized into the bilateral pair of head lobes (HL), the conspicuous column of ectodermal cells marking the ventral midline (flanked by arrows), the ectodermal grid (G) with the constellation of transverse parasegmental rows, and the unorganized posterior cells (U) before they become aligned in rows. **(H)** Schematic representation of Ph_trn/caps and Ph_ten-m/-a genes expression pattern in grid cells. Cells that express these genes are shown in red. All scales are 100 nm.

Chapter X

Discussion

Understanding the emergence of biological form in developing multicellular organisms is one of the biggest challenges in modern developmental biology. One productive way to approach this problem is to ask how the activities of transcription factors and morphogen gradients, which are well-known for specifying cell fates and patterning developing tissues, control cells in space and time. In this context, recent studies in model organisms have identified a number of transmembrane proteins, including cadherins, Toll receptors and others leucine-rich-repeat proteins, as transcriptional targets of these regulatory pathways (Paré et al., 2014, 2019; Tsai et al., 2020). Previous work by our lab and collaborators revealed that the same subfamily of Toll genes that controls germband morphogenesis in the insect model *Drosophila melanogaster* (Paré et al., 2014), is also expressed during germband formation in a more basal holometabolous insect, the beetle *Tribolium castaneum*, in a crustacean, the amphipod *Parhyale hawaiensis*, as well as in other pancrustacean, centipede and spider embryos (Benton et al., 2016). In all species tested, these Toll receptors are expressed under the control of pair-rule genes in metameric stripes along the anterior-posterior axis and are required for normal cell movements during axial elongation of the embryos. These results suggest that these segmentally expressed Toll receptors represent a deeply conserved pan-arthropod system for epithelial organization and remodeling in response to body patterning gene regulatory networks.

During my Master's thesis, I explored another family of leucine-rich-repeat receptors, the Tartan/ Capricious receptors, and their interactors from the Teneurin family that have been also shown to direct cell polarity and epithelial organization at compartment boundaries during Drosophila germband extension (Paré et al., 2019). The results indicate that *P. hawaiensis* has four *tartan/ capricious* homologs and three teneurin homologs. The orthology of these genes is supported by bioinformatic analysis of the *Parhyale* genome, by protein motif and domain analyses using the Expasy ScanProsite tool (Castro et al., 2006) and by maximum likelihood phylogenetic analyses using RAxML (Stamatakis, 2015). The expression analysis carried out by colorimetric in situ hybridizations and three different probe combinations for each gene to verify the specificity of the signal, suggests that genes exhibited essentially identical expression patterns in the posterior unsegmented region and in metameric stripes in the more anterior formed parasegments **(Fig IX.1)**.

Arthropod segmentation: The *Drosophila* paradigm and *Parhyale* as a model system to study morphological movements and temporal patterning in sequentially segmenting arthropods

The arthropod body plan consists of repeated units called segments along the anterior-posterior axis. Segments can be seen not only in the morphology of the adult, but also during embryogenesis. This is particularly evident at the phylotypic germband stage where the expression of segment polarity genes demarcates the molecular boundaries of the parasegment, the developmental unit underlying the morphological segment. Due to the seminal work in Drosophila on the genetics of early patterning, we now understand in tremendous detail the genetic hierarchy underlying Drosophila segmentation. Beginning with gap genes and continuing though the pair rule and segment polarity genes, the segmentation gene cascade acts to successively refine the blastoderm into finer and finer regions, ultimately resulting in the segmented fly embryo. This mode of embryogenesis, where the entire body axis is simultaneously subdivided into parasegments is termed 'long germ' embryogenesis. However, it turns out that long germ embryogenesis is a derived form of segmentation, seen only in a subset of insects. Other arthropods, and indeed many insects, undergo what is termed ‘short germ’ or sequential segmentation. In these animals, only the anterior segments are specified when gastrulation begins. Additional segments are added during a later stage of embryogenesis and are specified sequentially in an anterior to posterior progression. This fundamental embryological difference implies differences in the molecular genetics regulating segmentation in these two modes of development. Studies in several arthropods have shown that many of the orthologs of Drosophila segmentation genes are expressed during embryogenesis, though some differently than in Drosophila, implying different roles in segmentation (reviewed in (Peel et al., 2005; Liu and Kaufman, 2005; Damen, 2007)).

In the study of sequentially segmenting arthropods, one must account for cellular activities such as migration, division, or death that may occur during the process of segmentation. Expression studies and functional analyses of arthropod segmentation outside of Drosophila have only just started to characterize results with respect to these cellular activities. For example, functional studies in sequentially segmenting spiders and insects reveal that knockdown of genes expressed in the posterior germband causes a "lack of growth" (Stollewerk et al., 2003; Liu and Kaufman, 2005; Choe et al., 2006). It is not known if the lack of growth is due to cell death, loss of cell divisions, lack of cell movements, or some combination thereof. Importantly, these results indicate that patterning within sequ(Stamatakis, 2015)ential segmentation is closely linked with aspects of development that rely upon mediators of growth such as cell division or rearrangement. Additionally, it also remains to be determined what role, if any, orthologs of Drosophila segmentation genes have in the clock-like rate at which segments are added.

As an outgroup to the insects, and part of the sister clade to chelicerates and myriapods, crustaceans occupy a pivotal position in the arthropod phylogeny and may provide valuable insight into arthropod evolution and the unique changes that have occurred among insects. Unfortunately, there has been little

data addressing the role of pair rule orthologs in crustacean segmentation. The amphipod Parhyale hawaiensis is an ideal crustacean species in which to study segmentation. In addition to being easy to culture in the lab, several aspects of early development have been well-studied, including the morphological and cellular basis of segmentation. Parchem et al. analyzed the expression of segmentation gene orthologs in the amphipod crustacean *Parhyale hawaiensis*. His results indicate that in *Parhyale*, the regulation of these genes may be closely linked to cell biological activities such as migration and division. In contrast to the transcriptional network that functions in the well understood system of Drosophila where divisions and migration are not part of the segmentation process leading up to parasegment formation, the expression of Parhyale segmentation orthologs supports the hypothesis that each of the stereotyped divisions in the germband of *Parhyale* is intimately tied to a patterning event, as unique molecular identity is present prior, during, and after each set of stereotyped divisions. These results could provide a new insight into the relationship between the expression of *Parhyale* segmentation genes and the expression of its *tartan/ capricious* and *teneurins* orthologs, during the different phases of *Parhyale* grid formation.

Expression of *trn/ caps* and *ten-m/ -a* orthologs in relationship to the expression of segmentation genes in *Parhyale*

Expression of genes during PSPR cell migration

Parhyale grid formation begins early in development. Just after gastrulation, scattered ectodermal cells migrate to their proper position and begin condensing into the future germband. Parchem et al. shows that at least three genes are expressed in these cells as they begin to migrate into the presumptive grid. Once there, it appears that the expression of these genes is upregulated in cells as they form PSPRs. The expression during migration and within PSPRs may represent two different roles for these genes during Parhyale segmentation. In other arthropods, expression of pair rule orthologs is seen in the most posterior region of the developing germband before resolving into stripes anteriorly (reviewed in (Peel et al., 2005)). The expression of pair rule orthologs during migration and within PSPRs could be analogous to these patterns of expression. Nonetheless, the *trn/ caps* and *ten-m/ -a* orthologs show no expression during PSPR cell migration. Given that we identified all possible homologs of these genes in *Parhyale,* it is hard to imagine that any of them is needed for the migration or condensing of PSPR cells. Further study of cell movements of scattered ectodermal cells migrating to their proper position in the forming germband in *trn/ caps* and/or *ten-m/ -a* null *Parhyale* embryos may provide insight into this issue.

Expression of genes during PSPR division

Ph_trn/caps_1 to 4 and *Ph_ten-m/ -a_1 to 3* are all expressed for the first time just before PSPR division. Their RNA expression levels seem to persist in 2-row parasegments. Interestingly, *Ph-hh, Ph-odd-1, Ph-odd-2, Ph-odd-5, and Ph-hes-1* are also expressed just before PSPR division implying that a PSPR cell adopts a new molecular identity just before it divides. Furthermore, the expression of these genes is reactivated in the anterior daughter (A/B) immediately following the division. It is tempting to hypothesize that the expression of these transmembrane receptors is a direct result of the combinatorial expression of some or all of these segmentation genes. Additionally, RNAi phenotype of *Ph-hh* (and *Ph-eve-2)* does not show defects in the initial condensation of the grid. Instead, it appears that *Ph-hh* (and *Ph-eve-2)* are necessary for the proper formation of columns and rows within the grid as well as the initiation of PSPR divisions. The similarity in phenotype between *Ph-hh* and *Ph-eve-2* may reflect a shared role in regulation of each other in coordinating PSPR mitosis. In insects, divisions within the growing germband have not been characterized. Although orthologs of eve do result in similar phenotypes in insects, whether or not this is because of a block in cell division or morphological movements has not been determined.

Expression of genes after PSPR division in A/B or C/D

Following the division of PSPRs, the future parasegment now consists of two rows. Expression of *Ph_trn/caps* and *Ph_ten-m/-a* genes persists in this region after the first mitotic division. However, in addition to the segmentation genes that are expressed in A/B following expression just prior to mitosis, another set of genes is only expressed after the division. For example, following PSPR division *Ph-hes-2* is expressed in A/B and *Ph-opa-1* is expressed in C/D. That these rows of cells have unique anterior and posterior molecular identities implies that asymmetric patterning of the parasegment occurs very early. Similar in some respects to the complementary expression patterns of primary pair rule genes in Drosophila, it appears that an early pattern exists within the germband that is subsequently elaborated into the Parasegment. How this transcriptional pattern guides the expression of transmembrane genes like *Ph_trn/caps* and *Ph_ten-m/-a* remains to be seen.

Expression of genes after A/B or C/D division in ‘a’,’b’ or ‘c’,’d’

After the second mitotic division the expression of *Ph_trn/caps* and *Ph_ten-m/-a* genes is restricted in the posterior two rows of cells in the now 4-row parasegments. The parasegment is considered as having been formed upon expression of *Ph-en-1* and *Ph-en-2* in the 'a' row. Additional genes such as *Ph-pax3/7-1* (‘a’ and ‘d’ row) and *Ph-slp-1* (‘d’ row) are expressed in specific rows of the four-cell row parasegment and are likely to be involved in the maintenance of segment polarity. *Ph-pax3/7-1* may act in a homologous manner to the Drosophila segment polarity gene gooseberry. Interestingly, orthologs of the pair rule gene *slp* do not appear to have a role prior to parasegment formation in *Parhyale*. Similarly,

other studies have found orthologs of *slp* to be expressed much later than other pair rule genes in several arthropods (Choe et al., 2006; Damen et al., 2005). This may indicate that ancestrally, *slp* performed a segment polarity function before being co-opted into the pair rule level of patterning in some insects. Given the roles of Tartan and Ten-m receptors in directing planar polarity at compartment boundaries in *Drosophila,* where Tartan sets the position of compartment boundaries by recruiting Ten-m *in trans* and inhibiting its membrane localization *in cis*, together with the identical stripy expression pattern their orthologs exhibit in *Parhyale,* it is likely that segment polarity genes like *Ph-pax3/7-1* and *Ph-slp-1* orchestrate the expression of one or both groups of receptors in these cell rows.

Ph_trn/caps_2 expression not being apparent in the germband

Bioinformatic analysis of the *Ph_trn/caps_2* transcript showed that it had one of the highest similarities to *Drosophila* tartan, it contained 12 LRR motifs and was expressed by a distinct genomic locus. Nonetheless, despite being a true ortholog of *tartan,* this was the only gene whose expression was not successfully analysed by colorimetric *in situ* hybridization. Fine tuning parameters such as probe concentration, development time or antibody concentration did not aid in detecting its expression. A few possible reasons could explain this result. The first one that should be considered is experimental practicalities. Designing better RNA probes or using fluorescently marked ones might have allowed for the detection of Ph_trn/caps_2 expression. Since colorimetric *in situ* relies on the contrast in signal between the expressing and non-expressing cells it is intrinsically harder to detect with high confidence the expression of genes with minute expression levels. Another possible reason is that Ph_trn/caps_2 may not be expressed in the ectodermal grid. It may play a similar (suspected) role as the other Ph_trn/caps proteins, but in different developmental stages. Or perhaps it may have adopted a unique function compared to the others. For example, Capricious has been shown, in other species, to be expressed in neural tissues and to guide elongating neurons as they transverse their environment. Further analysis of Ph_trn/caps_2 expression is needed in order to answer some of these questions.

Chapter XI

Future directions

Using high-resolution RNA fluorescence in situ hybridization (HCR RNA-FISH) could enable a more detailed expression analysis of these genes. This would allow for the quantification of expression levels in early stage embryos (S14) or even prior to the formation of the germband, i.e in migrating ectodermal cells. In addition, such precise and high-resolution methods could be used to detect the expression of Ph_trn/caps_2 and deliver an answer on whether or not it is expressed in the germaband.

Furthermore, future experiments combining genetic perturbation of the Parhyale *tartan* and *teneurin* genes with imaging of actomyosin dynamics and grid architecture during *Parhyale* germband morphogenesis will help resolve how their patterned expression is converted into tissue structure in this species. Functional analysis of these receptors using CRISPR to knock out all or specific combinations of them was set in motion during this study. gRNAs targeting all 7 of the *Parhyale* genes (two per gene, in total 14 gRNAs) have already been designed. Assuming that deletion of these genes is not lethal in the early embryonic stages prior to germband formation, injecting parhyale embryos with Cas9-gRNA mixture targeting individual genes will shed light to their role in *Parhyale* germband morphogenesis. These experiments could also be performed in conjecture with the genetic perturbation of segmentation genes in Parhyale.

In addition, developing *in vitro* cultures of *Parhyale* cells expressing tagged forms of these receptors, would help determine the specific homophilic or heterophilic interactions among them and their effect on the cells adhesion properties, as well as their planar polarity and actomyosin contractility.

References

Abe, K., & Takeichi, M. (2008). EPLIN mediates linkage of the cadherin catenin complex to F-actin and stabilizes the circumferential actin belt. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(1), 13–19.

Albiges-Rizo, C., Destaing, O., Fourcade, B., Planus, E., & Block, M. R. (2009). Actin machinery and mechanosensitivity in invadopodia, podosomes and focal adhesions. *Journal of Cell Science*, *122*(Pt 17), 3037–3049.

Amack, J. D., & Manning, M. L. (2012). Knowing the boundaries: extending the differential adhesion hypothesis in embryonic cell sorting. *Science*, *338*(6104), 212–215.

Angiari, S. (2015). Selectin-mediated leukocyte trafficking during the development of autoimmune disease. *Autoimmunity Reviews*, *14*(11), 984–995.

Angres, B., Barth, A., & Nelson, W. J. (1996). Mechanism for transition from initial to stable cell-cell adhesion: kinetic analysis of E-cadherin-mediated adhesion using a quantitative adhesion assay. *The Journal of Cell Biology*, *134*(2), 549–557.

Arboleda-Estudillo, Y., Krieg, M., Stühmer, J., Licata, N. A., Muller, D. J., & Heisenberg, C.-P. (2010). Movement directionality in collective migration of germ layer progenitors. *Current Biology: CB*, *20*(2), 161–169.

Armstrong, P. B. (1989). Cell sorting out: the self-assembly of tissues in vitro. *Critical Reviews in Biochemistry and Molecular Biology*, *24*(2), 119–149.

Barnard, J. L. (1969). *The Families and Genera of Marine Gammaridean Amphipoda*.

Batlle, E., & Wilkinson, D. G. (2012). Molecular mechanisms of cell segregation and boundary formation in development and tumorigenesis. *Cold Spring Harbor Perspectives in Biology*, *4*(1), a008227.

Benton, M. A., Pechmann, M., Frey, N., Stappert, D., Conrads, K. H., Chen, Y.-T., Stamataki, E., Pavlopoulos, A., & Roth, S. (2016). Toll Genes Have an Ancestral Role in Axis Elongation. *Current Biology: CB*, *26*(12), 1609–1615.

Biswas, K. H., & Zaidel-Bar, R. (2017). Early events in the assembly of E-cadherin adhesions. *Experimental Cell Research*, *358*(1), 14–19.

Boucard, A. A., Maxeiner, S., & Südhof, T. C. (2014). Latrophilins function as heterophilic cell-adhesion molecules by binding to teneurins: regulation by alternative splicing. *The Journal of Biological Chemistry*, *289*(1), 387–402.

Brackenbury, R., Thiery, J. P., Rutishauser, U., & Edelman, G. M. (1977). Adhesion among neural cells of the chick embryo. I. An immunological assay for molecules involved in cell-cell binding. *The Journal of Biological Chemistry*, *252*(19), 6835–6840.

Browne, W. E., Price, A. L., Gerberding, M., & Patel, N. H. (2005). Stages of embryonic development in the amphipod crustacean, Parhyale hawaiensis. *Genesis* , *42*(3), 124–149.

Campbell, I. D., & Humphries, M. J. (2011). Integrin structure, activation, and interactions. *Cold Spring Harbor Perspectives in Biology*, *3*(3). https://doi.org/10.1101/cshperspect.a004994

Castro, E. de, de Castro, E., Sigrist, C. J. A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., Bairoch, A., & Hulo, N. (2006). ScanProsite: detection of PROSITE signature matches and

ProRule-associated functional and structural residues in proteins. In *Nucleic Acids Research* (Vol. 34, Issue Web Server, pp. W362–W365). https://doi.org/10.1093/nar/gkl124

Cavey, M., Rauzi, M., Lenne, P.-F., & Lecuit, T. (2008). A two-tiered mechanism for stabilization and immobilization of E-cadherin. *Nature*, *453*(7196), 751–756.

Chaudhuri, O., Parekh, S. H., & Fletcher, D. A. (2007). Reversible stress softening of actin networks. *Nature*, *445*(7125), 295–298.

Cheng, H. J., & Flanagan, J. G. (1994). Identification and cloning of ELF-1, a developmentally expressed ligand for the Mek4 and Sek receptor tyrosine kinases. *Cell*, *79*(1), 157–168.

Chuai, M., & Weijer, C. J. (2008). The mechanisms underlying primitive streak formation in the chick embryo. *Current Topics in Developmental Biology*, *81*, 135–156.

Clark, A. G., Wartlick, O., Salbreux, G., & Paluch, E. K. (2014). Stresses at the cell surface during animal cell morphogenesis. *Current Biology: CB*, *24*(10), R484–R494.

Cooke, J. E., & Moens, C. B. (2002). Boundary formation in the hindbrain: Eph only it were simple. *Trends in Neurosciences*, *25*(5), 260–267.

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.

David, R., Luu, O., Damm, E. W., Wen, J. W. H., Nagel, M., & Winklbauer, R. (2014). Tissue cohesion and the mechanics of cell rearrangement. *Development* , *141*(19), 3672–3682.

Davis, G. S., Phillips, H. M., & Steinberg, M. S. (1997). Germ-layer surface tensions and “tissue affinities” in Rana pipiens gastrulae: quantitative measurements. *Developmental Biology*, *192*(2), 630–644.

del Rio, A., Perez-Jimenez, R., Liu, R., Roca-Cusachs, P., Fernandez, J. M., & Sheetz, M. P. (2009). Stretching single talin rod molecules activates vinculin binding. *Science*, *323*(5914), 638–641.

Drees, F., Pokutta, S., Yamada, S., Nelson, W. J., & Weis, W. I. (2005). Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. *Cell*, *123*(5), 903–915.

Drescher, U., Kremoser, C., Handwerker, C., Löschinger, J., Noda, M., & Bonhoeffer, F. (1995). In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal protein related to ligands for Eph receptor tyrosine kinases. *Cell*, *82*(3), 359–370.

Duguay, D., Foty, R. A., & Steinberg, M. S. (2003). Cadherin-mediated cell adhesion and tissue segregation: qualitative and quantitative determinants. *Developmental Biology*, *253*(2), 309–323.

Durbin, L., Brennan, C., Shiomi, K., Cooke, J., Barrios, A., Shanmugalingam, S., Guthrie, B., Lindberg, R., & Holder, N. (1998). Eph signaling is required for segmentation and differentiation of the somites. *Genes & Development*, *12*(19), 3096–3109.

Extavour, C. G. (2005). The fate of isolated blastomeres with respect to germ cell formation in the amphipod crustacean Parhyale hawaiensis. *Developmental Biology*, *277*(2), 387–402.

Fagotto, F. (2014). The cellular basis of tissue separation. *Development* , *141*(17), 3303–3318.

Fagotto, F. (2015). Regulation of cell adhesion and cell sorting at embryonic boundaries. *Current Topics in Developmental Biology*, *112*, 19–64.

Fletcher, D. A., & Mullins, R. D. (2010). Cell mechanics and the cytoskeleton. *Nature*, *463*(7280), 485–492.

Foty, R. A., Pfleger, C. M., Forgacs, G., & Steinberg, M. S. (1996). Surface tensions of embryonic tissues predict their mutual envelopment behavior. *Development* , *122*(5), 1611–1620.

Fraser, S., Keynes, R., & Lumsden, A. (1990). Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature*, *344*(6265), 431–435.

Friedlander, D. R., Mège, R. M., Cunningham, B. A., & Edelman, G. M. (1989). Cell sorting-out is modulated by both the specificity and amount of different cell adhesion molecules (CAMs) expressed on cell surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(18), 7043–7047.

Fu, G., Wang, W., & Luo, B.-H. (2012). Overview: structural biology of integrins. *Methods in Molecular Biology* , *757*, 81–99.

García-Bellido, A. (1975). Genetic control of wing disc development in Drosophila. *Ciba Foundation Symposium*, *0*(29), 161–182.

Garcia-Bellido, A., Ripoll, P., & Morata, G. (1973). Developmental compartmentalisation of the wing disk of Drosophila. *Nature: New Biology*, *245*(147), 251–253.

Geiger, B., Spatz, J. P., & Bershadsky, A. D. (2009). Environmental sensing through focal adhesions. *Nature Reviews. Molecular Cell Biology*, *10*(1), 21–33.

Gerberding, M., Browne, W. E., & Patel, N. H. (2002). Cell lineage analysis of the amphipod crustacean Parhyale hawaiensis reveals an early restriction of cell fates. *Development* , *129*(24), 5789–5801.

Guthrie, S., & Lumsden, A. (1991). Formation and regeneration of rhombomere boundaries in the developing chick hindbrain. *Development* , *112*(1), 221–229.

Guthrie, S., Prince, V., & Lumsden, A. (1993). Selective dispersal of avian rhombomere cells in orthotopic and heterotopic grafts. *Development* , *118*(2), 527–538.

Harris, A. K. (1976). Is Cell sorting caused by differences in the work of intercellular adhesion? A critique of the Steinberg hypothesis. *Journal of Theoretical Biology*, *61*(2), 267–285.

Harris, T. J. C., & Peifer, M. (2004). Adherens junction-dependent and -independent steps in the establishment of epithelial cell polarity in Drosophila. *The Journal of Cell Biology*, *167*(1), 135–147.

Helwani, F. M., Kovacs, E. M., Paterson, A. D., Verma, S., Ali, R. G., Fanning, A. S., Weed, S. A., & Yap, A. S. (2004). Cortactin is necessary for E-cadherin-mediated contact formation and actin reorganization. *The Journal of Cell Biology*, *164*(6), 899–910.

Hughes, C. L., & Kaufman, T. C. (2002). Hox genes and the evolution of the arthropod body plan. *Evolution & Development*, *4*(6), 459–499.

Iwamoto, D. V., & Calderwood, D. A. (2015). Regulation of integrin-mediated adhesions. *Current Opinion in Cell Biology*, *36*, 41–47.

Janssen, R., & Lionel, L. (2018). Embryonic expression of a Long Toll (Loto) gene in the onychophorans Euperipatoides kanangrensis and Cephalofovea clandestina. *Development Genes and Evolution*, *228*(3-4), 171–178.

Kametani, Y., & Takeichi, M. (2007). Basal-to-apical cadherin flow at cell junctions. *Nature Cell Biology*, *9*(1), 92–98.

Kim, S. H., Yamamoto, A., Bouwmeester, T., Agius, E., & Robertis, E. M. (1998). The role of paraxial protocadherin in selective adhesion and cell movements of the mesoderm during Xenopus gastrulation. *Development* , *125*(23), 4681–4690.

Klopper, A. V., Krens, G., Grill, S. W., & Heisenberg, C.-P. (2010). Finite-size corrections to scaling behavior in sorted cell aggregates. *The European Physical Journal. E, Soft Matter*, *33*(2), 99–103.

Kobe, B., & Kajava, A. V. (2001). The leucine-rich repeat as a protein recognition motif. *Current Opinion in Structural Biology*, *11*(6), 725–732.

Konstantinides, N., & Averof, M. (2014). A common cellular basis for muscle regeneration in arthropods

and vertebrates. *Science*, *343*(6172), 788–791.

Kontarakis, Z., Pavlopoulos, A., Kiupakis, A., Konstantinides, N., Douris, V., & Averof, M. (2011). A versatile strategy for gene trapping and trap conversion in emerging model organisms. *Development* , *138*(12), 2625–2630.

Kovacs, E. M., Goodwin, M., Ali, R. G., Paterson, A. D., & Yap, A. S. (2002). Cadherin-directed actin assembly: E-cadherin physically associates with the Arp2/3 complex to direct actin assembly in nascent adhesive contacts. *Current Biology: CB*, *12*(5), 379–382.

Läubli, H., & Borsig, L. (2010). Selectins promote tumor metastasis. *Seminars in Cancer Biology*, *20*(3), 169–177.

Lavalou J, Mao Q, Harmansa S, Kerridge S, Lellouch AC, Philippe J, Audebert S, Camoin L, Lecuit T. (2020). Formation of polarized contractile interfaces by self-organized Toll-8/Cirl GPCR asymmetry. *bioRxiv*.

Lawrence, P. A., Green, S. M., & Johnston, P. (1978). Compartmentalization and growth of the Drosophila abdomen. *Journal of Embryology and Experimental Morphology*, *43*, 233–245.

Leamey, C. A., & Sawatari, A. (2014). The teneurins: new players in the generation of visual topography. *Seminars in Cell & Developmental Biology*, *35*, 173–179.

Levine A, Bashan-Ahrend A, Budai-Hadrian O, Gartenberg D, Menasherow S, Wides R. (1994). Odd Oz: a novel Drosophila pair rule gene. *Cell*, *77*, 587–598.

Li, J., Shalev-Benami, M., Sando, R., Jiang, X., Kibrom, A., Wang, J., Leon, K., Katanski, C., Nazarko, O., Lu, Y. C., Südhof, T. C., Skiniotis, G., & Araç, D. (2018). Structural Basis for Teneurin Function in Circuit-Wiring: A Toxin Motif at the Synapse. *Cell*, *173*(3), 735–748.e15.

Lossie, A. C., Nakamura, H., Thomas, S. E., & Justice, M. J. (2005). Mutation of l7Rn3 shows that Odz4 is required for mouse gastrulation. *Genetics*, *169*(1), 285–299.

Maître, J.-L., Berthoumieux, H., Krens, S. F. G., Salbreux, G., Jülicher, F., Paluch, E., & Heisenberg, C.-P. (2012). Adhesion functions in cell sorting by mechanically coupling the cortices of adhering cells. *Science*, *338*(6104), 253–256.

Manning, M. L., Foty, R. A., Steinberg, M. S., & Schoetz, E.-M. (2010). Coaction of intercellular adhesion and cortical tension specifies tissue surface tension. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(28), 12517–12522.

Matsushima, N., Miyashita, H., Mikami, T., & Kuroki, Y. (2010). A nested leucine rich repeat (LRR) domain: the precursor of LRRs is a ten or eleven residue motif. *BMC Microbiology*, *10*, 235.

Matthäus, C., Langhorst, H., Schütz, L., Jüttner, R., & Rathjen, F. G. (2017). Cell-cell communication mediated by the CAR subgroup of immunoglobulin cell adhesion molecules in health and disease. *Molecular and Cellular Neurosciences*, *81*, 32–40.

McEver, R. P. (2015). Selectins: initiators of leucocyte adhesion and signalling at the vascular wall. *Cardiovascular Research*, *107*(3), 331–339.

McGrew, M. J., & Pourquié, O. (1998). Somitogenesis: segmenting a vertebrate. *Current Opinion in Genetics & Development*, *8*(4), 487–493.

Mellitzer, G., Xu, Q., & Wilkinson, D. G. (1999). Eph receptors and ephrins restrict cell intermingling and communication. *Nature*, *400*(6739), 77–81.

Milán, M., & Cohen, S. M. (2003). A re-evaluation of the contributions of Apterous and Notch to the dorsoventral lineage restriction boundary in the Drosophila wing. *Development* , *130*(3), 553–562.

Milán, M., Weihe, U., Pérez, L., & Cohen, S. M. (2001). The LRR proteins capricious and Tartan mediate

cell interactions during DV boundary formation in the Drosophila wing. *Cell*, *106*(6), 785–794.
Montero, J.-A., Carvalho, L., Wilsch-Bräuninger, M., Kilian, B., Mustafa, C., & Heisenberg, C.-P. (2005). Shield formation at the onset of zebrafish gastrulation. *Development* , *132*(6), 1187–1198.
Mosca, T. J., Hong, W., Dani, V. S., Favaloro, V., & Luo, L. (2012). Trans-synaptic Teneurin signalling in neuromuscular synapse organization and target choice. *Nature*, *484*(7393), 237–241.
Moscona, A. A. (1968). Cell aggregation: properties of specific cell-ligands and their role in the formation of multicellular systems. *Developmental Biology*, *18*(3), 250–277.
Müller, H. A., & Wieschaus, E. (1996). armadillo, bazooka, and stardust are critical for early stages in formation of the zonula adherens and maintenance of the polarized blastoderm epithelium in Drosophila. *The Journal of Cell Biology*, *134*(1), 149–163.
Mullins, R. D., Heuser, J. A., & Pollard, T. D. (1998). The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(11), 6181–6186.
Nakamura, H., Cook, R. N., & Justice, M. J. (2013). Mouse Tenm4 is required for mesoderm induction. *BMC Developmental Biology*, *13*, 9.
Ninomiya, H., David, R., Damm, E. W., Fagotto, F., Niessen, C. M., & Winklbauer, R. (2012). Cadherin-dependent differential cell adhesion in Xenopus causes cell sorting in vitro but not in the embryo. *Journal of Cell Science*, *125*(Pt 8), 1877–1883.
Ninomiya, H., & Winklbauer, R. (2008). Epithelial coating controls mesenchymal shape change through tissue-positioning effects and reduction of surface-minimizing tension. *Nature Cell Biology*, *10*(1), 61–69.
Paré, A. C., Naik, P., Shi, J., Mirman, Z., Palmquist, K. H., & Zallen, J. A. (2019). An LRR Receptor-Teneurin System Directs Planar Polarity at Compartment Boundaries. *Developmental Cell*, *51*(2), 208–221.e6.
Paré, A. C., Vichas, A., Fincher, C. T., Mirman, Z., Farrell, D. L., Mainieri, A., & Zallen, J. A. (2014). A positional Toll receptor code directs convergent extension in Drosophila. *Nature*, *515*(7528), 523–527.
Pavlopoulos, A., & Averof, M. (2002). Developmental evolution: Hox proteins ring the changes. *Current Biology: CB*, *12*(8), R291–R293.
Pollard, T. D. (2007). Regulation of actin filament assembly by Arp2/3 complex and formins. *Annual Review of Biophysics and Biomolecular Structure*, *36*, 451–477.
Pollard, T. D., & Borisy, G. G. (2003). Cellular motility driven by assembly and disassembly of actin filaments. *Cell*, *112*(4), 453–465.
Prakasam, A. K., Maruthamuthu, V., & Leckband, D. E. (2006). Similarities between heterophilic and homophilic cadherin adhesion. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(42), 15434–15439.
Rauzi, M., & Lenne, P.-F. (2011). Cortical forces in cell shape changes and tissue morphogenesis. *Current Topics in Developmental Biology*, *95*, 93–144.
Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., & Patel, N. H. (2009a). Antibody staining of Parhyale hawaiensis embryos. *Cold Spring Harbor Protocols*, *2009*(1), db.prot5129.
Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., & Patel, N. H. (2009b). Injection of Parhyale hawaiensis blastomeres with fluorescently labeled tracers. *Cold Spring Harbor Protocols*, *2009*(1),

db.prot5128.
Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., & Patel, N. H. (2009c). In situ hybridization of labeled RNA probes to fixed Parhyale hawaiensis embryos. *Cold Spring Harbor Protocols*, *2009*(1), db.prot5130.
Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A., & Patel, N. H. (2009d). The crustacean Parhyale hawaiensis: a new model for arthropod development. *Cold Spring Harbor Protocols*, *2009*(1), db.emo114.
Robinson, E. E., Foty, R. A., & Corbett, S. A. (2004). Fibronectin matrix assembly regulates alpha5beta1-mediated cell cohesion. *Molecular Biology of the Cell*, *15*(3), 973–981.
Robinson, E. E., Zazzali, K. M., Corbett, S. A., & Foty, R. A. (2003). Alpha5beta1 integrin mediates strong tissue cohesion. *Journal of Cell Science*, *116*(Pt 2), 377–386.
Rohani, N., Canty, L., Luu, O., Fagotto, F., & Winklbauer, R. (2011). EphrinB/EphB signaling controls embryonic germ layer separation by contact-induced cell detachment. *PLoS Biology*, *9*(3), e1000597.
Schier, A. F. (2003). Nodal signaling in vertebrate development. *Annual Review of Cell and Developmental Biology*, *19*, 589–621.
Schier, A. F. (2009). Nodal morphogens. *Cold Spring Harbor Perspectives in Biology*, *1*(5), a003459.
Shawky, J. H., & Davidson, L. A. (2015). Tissue mechanics and adhesion during embryo development. *Developmental Biology*, *401*(1), 152–164.
Shimoyama, Y., Tsujimoto, G., Kitajima, M., & Natori, M. (2000). Identification of three human type-II classic cadherins and frequent heterophilic interactions between different subclasses of type-II classic cadherins. *Biochemical Journal*, *349*(Pt 1), 159–167.
Shook, D. R., & Keller, R. (2008). Morphogenic machines evolve more rapidly than the signals that pattern them: lessons from amphibians. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, *310*(1), 111–135.
Singh, J., Hussain, F., & Decuzzi, P. (2015). Role of differential adhesion in cell cluster evolution: from vasculogenesis to cancer metastasis. *Computer Methods in Biomechanics and Biomedical Engineering*, *18*(3), 282–292.
Stamataki, E., & Pavlopoulos, A. (2016). Non-insect crustacean models in developmental genetics including an encomium to Parhyale hawaiensis. *Current Opinion in Genetics & Development*, *39*, 149–156.
Stamatakis, A. (2015). Using RAxML to Infer Phylogenies. In *Current Protocols in Bioinformatics* (Vol. 51, Issue 1). https://doi.org/10.1002/0471250953.bi0614s51
Steinberg, M. S. (1963). Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. *Science*, *141*(3579), 401–408.
Steinberg, M. S. (1970). Does differential adhesion govern self-assembly processes in histogenesis? Equilibrium configurations and the emergence of a hierarchy among populations of embryonic cells. *The Journal of Experimental Zoology*, *173*(4), 395–433.
Stirbat, T. V., Mgharbel, A., Bodennec, S., Ferri, K., Mertani, H. C., Rieu, J.-P., & Delanoë-Ayari, H. (2013). Fine tuning of tissues' viscosity and surface tension through contractility suggests a new role for α-catenin. *PloS One*, *8*(2), e52554.
Südhof, T. C. (2017). Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits.

Cell, *171*(4), 745–769.

Svitkina, T. M., & Borisy, G. G. (1999). Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *The Journal of Cell Biology*, *145*(5), 1009–1026.

Takeichi, M. (1977). Functional correlation between cell adhesive properties and some cell surface proteins. *The Journal of Cell Biology*, *75*(2 Pt 1), 464–474.

Tamkun, J. W., DeSimone, D. W., Fonda, D., Patel, R. S., Buck, C., Horwitz, A. F., & Hynes, R. O. (1986). Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell*, *46*(2), 271–282.

Tanaka, M., Kamo, T., Ota, S., & Sugimura, H. (2003). Association of Dishevelled with Eph tyrosine kinase receptor and ephrin mediates cell repulsion. *The EMBO Journal*, *22*(4), 847–858.

Tepass, U., Godt, D., & Winklbauer, R. (2002). Cell sorting in animal development: signalling and adhesive mechanisms in the formation of tissue boundaries. *Current Opinion in Genetics & Development*, *12*(5), 572–582.

Tepass, U., & Hartenstein, V. (1994). The development of cellular junctions in the Drosophila embryo. *Developmental Biology*, *161*(2), 563–596.

Thiery, J. P., Brackenbury, R., Rutishauser, U., & Edelman, G. M. (1977). Adhesion among neural cells of the chick embryo. II. Purification and characterization of a cell adhesion molecule from neural retina. *The Journal of Biological Chemistry*, *252*(19), 6841–6845.

Townes, P. L., & Holtfreter, J. (1955). *Directed Movements and Selective Adhesion of Embryonic Amphibian Cells*.

Tsai, T. Y.-C., Sikora, M., Xia, P., Colak-Champollion, T., Knaut, H., Heisenberg, C.-P., & Megason, S. G. (2020). An adhesion code ensures robust pattern formation during tissue morphogenesis. *Science*, *370*(6512), 113–116.

Vasiev, B., Balter, A., Chaplain, M., Glazier, J. A., & Weijer, C. J. (2010). Modeling gastrulation in the chick embryo: formation of the primitive streak. *PloS One*, *5*(5), e10571.

Vasioukhin, V., Bauer, C., Yin, M., & Fuchs, E. (2000). Directed actin polymerization is the driving force for epithelial cell-cell adhesion. *Cell*, *100*(2), 209–219.

Verma, S., Shewan, A. M., Scott, J. A., Helwani, F. M., den Elzen, N. R., Miki, H., Takenawa, T., & Yap, A. S. (2004). Arp2/3 activity is necessary for efficient formation of E-cadherin adhesive contacts. *The Journal of Biological Chemistry*, *279*(32), 34062–34070.

Wachsstock, D. H., Schwarz, W. H., & Pollard, T. D. (1994). Cross-linker dynamics determine the mechanical properties of actin gels. *Biophysical Journal*, *66*(3 Pt 1), 801–809.

Warga, R. M., & Kane, D. A. (2007). A role for N-cadherin in mesodermal morphogenesis during gastrulation. *Developmental Biology*, *310*(2), 211–225.

Warga, R. M., & Kimmel, C. B. (1990). Cell movements during epiboly and gastrulation in zebrafish. *Development* , *108*(4), 569–580.

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.

Wayne Brodland, G., & Chen, H. H. (2000). The mechanics of cell sorting and envelopment. *Journal of Biomechanics*, *33*(7), 845–851.

Winklbauer, R. (2009). Cell adhesion in amphibian gastrulation. *International Review of Cell and Molecular Biology*, *278*, 215–275.

Winklbauer, R. (2015). Cell adhesion strength from cortical tension - an integration of concepts. *Journal of Cell Science*, *128*(20), 3687–3693.

Witz, I. P. (2008). The selectin-selectin ligand axis in tumor progression. *Cancer Metastasis Reviews*, *27*(1), 19–30.

Wizenmann, A., & Lumsden, A. (1997). Segregation of rhombomeres by differential chemoaffinity. *Molecular and Cellular Neurosciences*, *9*(5-6), 448–459.

Woelfle, R., D'Aquila, A. L., & Lovejoy, D. A. (2016). Teneurins, TCAP, and latrophilins: roles in the etiology of mood disorders. *Translational Neuroscience*, *7*(1), 17–23.

Yamada, S., Pokutta, S., Drees, F., Weis, W. I., & Nelson, W. J. (2005). Deconstructing the cadherin-catenin-actin complex. *Cell*, *123*(5), 889–901.

Yonemura, S., Wada, Y., Watanabe, T., Nagafuchi, A., & Shibata, M. (2010). alpha-Catenin as a tension transducer that induces adherens junction development. *Nature Cell Biology*, *12*(6), 533–542.

Youssef, J., Nurse, A. K., Freund, L. B., & Morgan, J. R. (2011). Quantification of the forces driving self-assembly of three-dimensional microtissues. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(17), 6993–6998.