Mechanisms of segmentation in the American cockroach, *Periplaneta americana*

John Chesebro

Doctor of Philosophy
School of Life Sciences
University of Sussex
Submitted December 2012

DECLARATION

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signed……..

UNIVERSITY OF SUSSEX

JOHN CHESEBRO

Thesis submitted for the Degree of Doctor of Philosophy

MECHANISMS OF SEGMENTATION IN THE AMERICAN COCKROACH, *PERIPLANETA AMERICANA*

SUMMARY

A fully segmented body and jointed legs are defining characteristics of the Arthropoda (Insecta, Crustacea, Myriapoda, and Chelicerata). The underlying mechanisms involved in achieving these features are not well understood outside of the insect *Drosophila melanogaster* (fruit fly) – a long germ band organism where segmentation occurs all at once in a syncytial blastoderm. In the more common, ancestral mode of development, short germ band, new segments are added sequentially from the cellular environment of a posteriorly extending growth zone. Segmentation in these organisms may not always be comparable to the "*Drosophila* paradigm" and, therefore, require further analysis. My thesis will explore the conservation and divergence of the molecular mechanisms of segmentation in a phylogenetically basal, short germ band insect, *Periplaneta americana* (American cockroach). Presented over three results chapters, I will discuss aspects of cockroach segmentation processes, from the establishment of a posterior organiser and growth zone, to subsequent posterior growth and the formation of new segments. In particular, Chapter III describes how interactions between the Cad/Wnt-dependent posterior organiser and the Notch-segmentation clock control posterior growth and segmentation. Chapter IV encompasses the expression patterns and potential roles for *Periplaneta* homologues of the pair-rule genes: *even-skipped, runt, pairberry,* and *sloppy-paired* throughout embryogenesis, identifying deviations in function between anterior and posterior segmentation processes. New functions for the non-canonical, polycistronic small Open Reading Frame (smORF) gene *tarsal-less* in body patterning are discussed in Chapter V, along with the conserved roles for *tarsal-less, nubbin, Notch,* and *Delta* in leg and development. Elucidation of the networks involved in these processes will help establish putative ancestral gene functions allowing us to gain further insights into the evolution of insect (and arthropod) body segmentation and leg joint formation.

DEDICATION

For Dave (and Petey)

ACKNOWLEDGMENTS

I would like to thank my PhD supervisor, Prof. Juan Pablo Couso, for his encouragement and guidance (and money) during my time in his laboratory. I would also like to thank my co-supervisors Dr. Robert Ray and Dr. Claudio Alonso for their thoughtful insights throughout my PhD work and the use of their specimens and lab equipment when in need. Special thanks to my *viva voce* examiners Prof. Jonathan Bacon (University of Sussex) and Dr. Alistair McGregor (Oxford Brookes University) for allowing me to receive a doctorate. I am also thankful to the members of the Couso lab: Inyaki Pueyo for his tremendous support and knowledge, Rose Phillips for doing some of my dirty work and taking such tender loving care of the cockroaches, Sarah Bishop for showing me the embryo culture technique, Emile Magny for dancing on tables, Julie Aspen for her laughter, Chris Sampson, Ali Mumtaz for being awesome, and finally Unum Amin for constantly distracting me and being a truly amazing friend, especially this past year. I would also like to thank former lab members: Miguel Céspedes for helping me to adjust and feel at home in a new city, mi novia Amparo Garcia for her love, and former students Rob Tetley, Claire Davison, and Alex Hurst – some of whose work is presented here. Special thanks to Richard Kaschula and Aidan Maartens for critical reading of portions of this thesis and to Gemma Hopley for careful proofreading. Thank you also to the School of Life Sciences and the University of Sussex for the Graduate Teaching Assistantship, for which I otherwise may not have had to opportunity to study here. Lastly, I would like to thank my friends and family for all their unending love and support during my years and years and years of graduate studies.

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

Genes

bcd – bicoid
btd – buttonhead
cad/Cdx – caudal
dac – dachshund
Dl – Delta
Dll – Distal-less
Dll1– Delta-like1
dpp – decapentaplegic
ems – empty spiracles
en – engrailed
eve – even-skipped
E(spl) – Enhancer of split
ftz – fushi tarazu
gby – gooseberry
gt – giant
h – hairy
hes – hairy and E(spl)
her – hairy and E(spl)-related
hb – hunchback
hth – homothorax
kni – knirps
mlpt – mille-pattes
N – Notch
nub – nubbin
nos – nanos
odd – odd-skipped
opa – odd-paired
otd – orthodenticle
prd/pby – paired/pairberry
rn – rotund
run – runt
slp – sloppy-paired
ss – spineless
svb – shaven baby
tal – tarsal-less
twi – twist
wg/Wnt1 – wingless

Anatomy

A1, A2, etc. – 1st, 2nd, etc. abdominal segments
a – abdomen
ant – antennae
cl – clypeus
cx – coxa
fe – femur
fr – frons
ga – galea
ge – gena
GZ – growth zone
h – head
ic – intercalary segment
lb – labium
lr – labrum
mn – mandibles
mx – maxilla
PSM – presomitic mesoderm
T1-T3 – 1st-3rd thoracic segments
t – thorax
ta – tarsus
ta1, ta2, etc. – tarsomere 1, 2, etc.
ti – tibia
tr – trochanter
ve – vertex

LIST OF ABBREVIATIONS

Organisms:
Abbreviation – *Latin name* (common name)

Ad –	*Acheta domesticus*	(house cricket)
Af –	*Artemia franciscana*	(brine shrimp)
Ap –	*Acyrthosiphon pisum*	(pea aphid)
At –	*Achaearanea* (*Parasteatoda*) *tepidariorum*	(common house spider)
Am –	*Apis mellifera*	(western or European honeybee)
Bm –	*Bombyx mori*	(silk moth)
Ca –	*Clogmia albipunctata*	(moth fly, drain fly)
Cap –	*Capitella capitata*	(gallery worm)
Cf –	*Copidosoma floridanum*	(endoparasitic wasp)
Cs –	*Cryptotermes secundus*	(drywood termite)
Cs –	*Cupiennius salei*	(Central American wandering spider)
Dm –	*Drosophila melanogaster*	(fruit fly)
Dr –	*Danio rerio*	(zebrafish)
Gb –	*Gryllus bimaculatus*	(field cricket)
Gm –	*Glomeris marginata*	(pill millipede)
Hro –	*Helobdella robusta*	(Californian leech)
Lf –	*Lithobius forficatus*	(brown/stone centipede)
La –	*Lithobius atkinsoni*	(centipede)
Md –	*Musca domestica*	(house fly)
Ms –	*Manduca sexta*	(tobacco hornworm)
Nv –	*Nasonia vitripennis*	(jewel wasp)
Of –	*Oncopeltus fasciatus*	(milkweed bug)
Pa –	*Periplaneta americana*	(American cockroach)
Pd –	*Platynereis dumerilii*	(Dumeril's clam worm)
Ph –	*Parhyale hawaiensis*	(amphipod crustacean)
Sa –	*Schistocerca americana*	(American bird grasshopper)
Sg –	*Schistocerca gregaria*	(desert locust)
Sm –	*Strigamia maritima*	(geophilomorph/coastal centipede)
Tc –	*Tribolium castaneum*	(red flour beetle)
Tu –	*Tetranychus urticae*	(two-spotted spider mite)
Tl –	*Triops longicaudatus*	(longtail tadpole shrimp)

CHAPTER I

GENERAL INTRODUCTION

Arthropods, annelids, and vertebrates represent some of the most successful and diverse organisms; they are found in all climates and environments. One of the keys to their evolutionary success is the development of a segmented body. The definition of segmentation has been argued by many (Budd, 2001; Couso, 2009; Davis and Patel, 1999; Minelli and Fusco, 2004; Scholtz, 2002), but in essence it is the subdivision of the body and/or organ systems into smaller serially repeated units that may be evident externally, internally, or both. Each segment, or group of segments (tagma), may function as a semi-independent unit (i.e. insect head, thorax, and abdomen) (French, 1983) and over the course of evolution, changes in the early genetic programs during embryogenesis led to modification of these units, resulting in the great morphological diversity we see today. The flexibility and increased mobility provided by this segmented/modular body plan, combined with phenotypic variation, allowed organisms to adapt to new and changing environmental conditions, leading to their rapid spread during the Cambrian radiation.

That segmentation exists among three of the largest and most successful animal phyla begs the question of relationship and the origin of segmentation. But why is segmentation so common among these three seeming divergent clades? There are three main theories on the evolutionary origin of segmentation (Fig. 1.1): 1) it developed independently in each of the three phyla (Chipman, 2010); 2) it arose separately in the protostomes and deuterostomes; or 3) it was already present in the last common bilaterian ancestor, *Urbilateria* (Davis and Patel, 1999; De Robertis, 1997). While data can be found to argue each of the three cases, more information is becoming available suggesting that the origin of segmentation was founded in the Urbilateria (Balavoine and Adoutte, 2003; Couso, 2009; De Robertis, 1997; De Robertis, 2008b; de Rosa et al., 2005; Erwin and

Davidson, 2002; Holland et al., 1997; Kimmel, 1996; Pueyo et al., 2008). To begin to understand the evolutionary relationship between these phyla, we need to examine the mechanisms of segmentation used in each case, comparing the similarities and differences between them. Exploring the available data and investigating more diverse and basal representative organisms from each phylum will expand our understanding of the evolution of segmentation and developmental mechanisms.

POSTERIOR GROWTH VIA A WNT-BASED POSTERIOR ORGANISER

A shared feature among most segmented bilaterians is embryonic growth along the anterior-posterior (AP) axis combined with the sequential addition of segments/somites from a posterior growth zone (GZ) or presomitic mesoderm (PSM – vertebrates) (Balavoine and Adoutte, 2003; Couso, 2009; De Robertis, 1997; Jacobs et al., 2005). An early step in embryonic growth is setting up the posterior axis and a 'posterior organiser' that expresses the Caudal and Wnt proteins; a mechanism that is considered ancestral to all segmented (and non-segmented) bilaterians (Holland, 2002; Kimelman and Martin, 2012; Martin and Kimelman, 2009; McGregor et al., 2009; Niehrs, 2010; Wei et al., 2012). In vertebrates, one of the three segmented phyla, this involves the interplay of several developmental factors including members of the Wnt-signalling pathway (i.e. *Wnt3a*) that regulate expression of the homeobox gene *caudal/Cdx* (Ikeya and Takada, 2001; Lohnes, 2003; Prinos et al., 2001; Shimizu et al., 2005; van de Ven et al., 2011). Wnt and Cdx function in the posterior PSM are required throughout embryonic development for proper growth and somitogenesis, as a loss of function of either leads to posterior segment abnormalities and truncations (Aulehla et al., 2003; Aulehla et al., 2008; Martin and Kimelman, 2009; Takada et al., 1994; Young et al., 2009).

Annelids and arthropods represent the other two segmented phyla, after vertebrates, and similarly grow and segment from the posterior. Although once considered more closely related, the annelids and arthropods are now placed within separate superphyla - the Lophotrochozoa and Ecdysozoa, respectively (Aguinaldo et al., 1997; Dunn et al., 2008; Philippe et al., 2005). Together these sister clades form the Protostomia that split from the Deuterostomia, of which the chordates belong, approximately 570mya. In the annelids, there is conservation in the expression of *cad* and *Wnt* family homologues in the small posterior growth zone of the polychaete worms *Platynereis dumerilii* and *Capitella telata* and the teloblastic 'growth zone' of the leech *Helobdella robusta* (Cho et al., 2010; de Rosa et al., 2005; Dray et al., 2010; Hui et al., 2009; Janssen et al., 2010). While functional analysis has not been carried out in these organisms, the similarity of expression patterns may indicate a conserved function for annelid Cad and Wnt1 in establishing a posterior organiser that regulates growth and segmentation.

Arthropods in which segments are added from a posteriorly extending GZ undergo what is called a short or intermediate germ band mode of development. During the earliest stages of embryogenesis, as the germ primordium condenses, posterior expression of *cad* and *Wnt1/wingless* (*wg*) establish the posterior axis and set up an organiser required for future growth and segmentation. Posterior expression of *Wnt1/wg* and *cad* is conserved in the short germ arthropods studied, including: the insects *Tribolium castaneum* (Bolognesi et al., 2008; Schulz et al., 1998), *Gryllus bimaculatus* (Miyawaki et al., 2004; Shinmyo et al., 2005), and *Oncopeltus fasciatus* (Angelini and Kaufman, 2005a); the crustaceans *Artemia franciscana* (Copf et al., 2003; Copf et al., 2004) and *Triops longicaudatus* (Nulsen and Nagy, 1999); the spider *Achaearanea tepidariorum* (Akiyama-Oda and Oda, 2003; Janssen et al., 2010; McGregor et al., 2008); and the myriapods *Glomeris marginata* (Janssen et al., 2010) and *Strigamia maritima* (Chipman et al., 2004). The function of a

Wnt-Cad organiser, where Wnt-signalling regulates *cad* expression, has been experimentally established in *Gryllus* (Shinmyo et al., 2005) and *Achaearanea* (McGregor et al., 2008). Posterior interaction between Cad and Wnt-signalling is an important feature and, similar to vertebrates, loss of function of either gene leads to posterior truncation (Bolognesi et al., 2008; Copf et al., 2004; Martin and Kimelman, 2009; McGregor et al., 2009; McGregor et al., 2008; Shinmyo et al., 2005).

Caudal and Wnt-signalling in the posterior end of the developing embryo work together to establish the posterior axis, form a posterior organiser, and establish a functional growth zone; processes conserved in most bilaterian animals. Not only are these genes important for continued posterior growth, they are also essential in regulating the processes of segmentation/somitogenesis. The mechanisms of segmentation can vary between phyla, but again, there is incredible similarity in the mechanisms involved in these processes, largely based around the N-signalling pathway. I will first review what is known about vertebrate somitogenesis ('clock and wavefront'), before briefly touching on annelid segmentation, and concluding with the various methods involved in arthropod body and appendage segmentation.

MECHANISMS OF SEGMENTATION

Vertebrate somitogenesis

Vertebrates make up the majority of the phylum Chordata, and exhibit metameric patterning of the muscles, nervous system, and skeleton compared to the more obvious, overt segmentation of the annelid worms and arthropods. Nevertheless, in each case new segments arise sequentially from the posteriorly extending PSM/GZ. The coordination of somite formation in vertebrates occurs through a series of synchronised oscillations of

gene expression that begins in the posterior PSM and move as a wave of transcription towards the anterior, inducing the expression of segmentation genes leading to somite formation in the anterior. This process involves complex interactions between three major signalling pathways: Notch (N), Wnt, and Fibroblast Growth Factor (FGF). Interactions between these pathways form a "clock and wavefront" mechanism of segmentation utilized by all vertebrates studied to date (Fig. 1.2) (Cooke and Zeeman, 1976; Dequeant and Pourquie, 2008; Giudicelli and Lewis, 2004; Jiang et al., 2000; Kageyama et al., 2007; Lewis, 2003; Naiche et al., 2011; Ozbudak and Pourquie, 2008; Palmeirim et al., 1997; Pourquie, 2011).

FGF and N-signalling both regulate the expression of the transcriptional regulator *hairy and Enhancer of split* (*hes/her*). In turn, hes/her negatively regulates its own expression as well as the ability of the N-pathway to signal to neighbouring cells, by inhibiting the expression of the N ligands *Delta* (*Dl*; in zebrafish) or *lunatic fringe* (*lfng*; chick and mouse) (Cinquin, 2007; Evrard et al., 1998; Kageyama et al., 2012; Kawamura et al., 2005; Lewis, 2003; Palmeirim et al., 1997; Pourquie, 2011). As the temporal waves of N-activation travel through the PSM, they are translated into a spatial pattern leading to somite formation in the anterior. The spatial readout is in the form of stripes from which segment boundaries and somite polarity will be regulated by segmentation genes, including *lfng* and *mesoderm posterior protein 2* (*mesp2*) (Aulehla et al., 2003; Dunty et al., 2008; Kawamura et al., 2005; Morimoto et al., 2005; Pourquie, 1999). Defects in any part of this highly complex patterning network may result in abnormal somite formation (Dubrulle et al., 2001; Lewis, 2003; Shimizu et al., 2005; Wahl et al., 2007); where these effects occur because of loss of N-signalling, they are attributed to desychronisation and decoupling of oscillations in the posterior PSM (Horikawa et al., 2006; Jiang et al., 2000; Lewis, 2003; Mara et al., 2007).

A diagrammatic summary of the vertebrate 'clock and wavefront' method of somitogenesis is presented in Figure 1.2. In brief, periodic waves of *Notch* activation travel from the posterior PSM to the anterior activating downstream effector genes, like *hes/her*, forming the synchronised "clock" mechanism (Forsberg et al., 1998; Jiang et al., 2000; Palmeirim et al., 1997). The posterior-to-anterior gradients of Wnt and FGF ("wavefront") signalling are counteracted by an opposite and antagonistic anterior-posterior gradient of retinoic acid. It is here, the determination zone, that the periodic wave of the clock is translated into a patterned stripe of expression and the production of one somite (Aulehla et al., 2003; Dubrulle et al., 2001; Morimoto et al., 2005; Oginuma et al., 2008).

Adding further complexity to the segmentation process are the varied rates of transcription/translation and mRNA/protein degradation for each of the components discussed above, which can have profound effects on segment formation. The intracellular, autoinhibitory oscillations of *hes/her* are correlated with the rates of production and decay, creating a delayed negative feedback loop (Bessho et al., 2003; Lewis, 2003; Oates et al., 2012). This cell-autonomous, autoinhibitory delay is very important as part of the clock mechanism added to coupling of the intracellular oscillations and synchronisation with neighbouring cells (via N-signalling), as loss of expression often leads to segmentation defects (Bessho et al., 2001; Herrgen et al., 2010; Lewis, 2003; Oates et al., 2012; Ozbudak and Lewis, 2008). The regulatory actions of Hes/Her are suggestive of a role as the clock pacemaker, but this responsibility is still unclear, as each of the signalling pathways have also been ruled out in this role, this elusive pace-keeper remains to be identified. Each pathway may have its own independent oscillation mechanism that overlaps at some interval, revealing a coordination between them, synchronized through N-signalling (Goldbeter and Pourquie, 2008; Kageyama et al.,

2012; Krol et al., 2011; Pourquie, 2011) or, alternatively, there may not be an autonomous pacemaker, leaving the oscillations to occur on a multicellular level.

Degradation rates of mRNA or protein also have an important role in the formation of the posterior-to-anterior Wnt and FGF gradients, which are linked to the rate of posterior growth. Both genes are expressed in the posterior, yet the rate of protein decay is quite slow; as the tail bud grows posteriorly this leads to a gradient of Wnt and FGF along the PSM with high levels in the posterior and lower levels in the anterior (Aulehla and Pourquie, 2010; Aulehla et al., 2003; Dubrulle and Pourquie, 2004b). The relative rates of production and decay affect segmentation in a way that can determine how large or small a somite may be. A slower clock, relative to posterior growth, leads to larger and/or fewer segments, whereas a faster clock may lead to smaller and/or more numerous segments (Gomez et al., 2008; Oates et al., 2012; Sawada et al., 2001; Schroter and Oates, 2010). The timing of the clock and the onset of somitogenesis varies amongst the organisms studied, ranging from one somite formed every 30 minutes in zebrafish to one every 120 minutes in mouse (Forsberg et al., 1998; Gomez et al., 2008). These are just a few examples of the flexibility and plasticity of the mechanisms of somitogenesis based around the integration of the conserved core players: FGF, Notch, Wnt, and hes/her. The interplay between these pathways has been extensively studied, but is not completely understood.

Interactions between the early establishment of a Wnt-cad posterior organiser and later segmentation via N-signalling are beginning to emerge. A positive interaction has recently been shown between Cdx/Cad and the Notch ligand *Delta-like1* (*Dll1*) during mouse somitogenesis (Grainger et al., 2012). Here, Cdx directly binds to the *Dll1* promoter, regulating its expression, illustrating the intimate link between posterior growth and segmentation (Aulehla et al., 2003; Aulehla et al., 2008; Dubrulle and Pourquie,

2004a; Grainger et al., 2012). Understanding the connections between and among the different segmented Bilateria will help us elucidate the evolution of the mechanisms involved.

Annelid segmentation

Annelid segmentation is externally visible along the annulated body and is carried through internally as each segment is divided by a septum. Unfortunately, research on segment patterning in this group is lacking; however, exciting results have shown similarities between annelid and vertebrate segmentation in the oscillatory expression of Notch-signalling. In the polychaete *Capitella capitata*, homologues of *N, Dl,* and *hes/her* are expressed in the juvenile sub-terminal growth zone, although functional analysis is still required to determine a possible function in segmentation (Thamm and Seaver, 2008). In the leech *Helobdella robusta*, homologues of *N* and *hes/her* oscillate in a manner that is tightly linked to the cell cycle, relying on the rate of teloblast cell divisions before activation of segmentation mechanisms (Rivera et al., 2005; Song et al., 2004). The presumed segmentation function for *N* and *hes* was confirmed in the leech by inhibition of N-signalling via DAPT treatment in cultured embryos and/or morpholino inhibition of *Hro-hes* translation (Rivera and Weisblat, 2009). Loss of N-signalling led to the reduction of *Hro-hes* expression, indicating *hes* as a downstream target, and loss of both *Hro-N* and *Hro-hes* lead to segmentation defects. The relationship between oscillatory *N* and *hes* expression with segment formation is very similar to the clock-and-wavefront mechanism of vertebrates and indicates a conserved, ancient mechanism between annelids and other segmented phyla. However, interactions between N-mediated segmentation and the presumed mechanisms of posterior growth, via *caudal* and Wnt-signalling, remain to be determined in annelids.

Arthropod segmentation

Highly segmented bodies and appendages are defining features of the Arthropoda. Members of this phylum include insects (flies, cockroaches), crustaceans (crabs, wood lice), chelicerates (spiders, mites), and myriapods (centipedes, millipedes). Overt segmentation of the body and appendages allowed for rapid evolution of the arthropods making them the most speciose and morphologically diverse phylum seen today. There are three main types of segmentation in arthropods: short, intermediate, and long germ band. In short and intermediate germ band organisms, the anterior head and thoracic segments form in the blastoderm while the remaining segments are added sequentially from a posteriorly extending growth zone. The short and intermediate modes differ in the number of segments specified in the blastoderm, but otherwise development is the same and, for simplicity, both will be referred to as short germ band throughout the text. In the long germ band mode of development, all body segments form almost simultaneously in the syncytial blastoderm with no directional posterior growth or sequential segment formation. The long germ band mode is highly derived compared to short germ band development and is the method utilized by the fruit fly, *Drosophila melanogaster*, of which much is known about the processes of segmentation.

Long germ band segmentation

During *Drosophila* embryogenesis, segments are specified simultaneously in the cell-free environment of a syncytial blastoderm. Much of what we know about arthropod segmentation today began with the intensive studies of *Drosophila* mutants displaying segmentation defects by the Nüsslein-Volhard/Wieschaus group (Nüsslein-Volhard and Wieschaus, 1980; Nusslein-Volhard et al., 1984). Examination of these phenotypes and the genes involved led to the establishment of the "*Drosophila* paradigm" – a hierarchic

cascade of transcription factors working together to pattern and segment the fly embryo into successively smaller units (Fig. 1.3A) (Ingham, 1988; Lawrence, 1992; Peel et al., 2005). In this paradigm, the first genes to be expressed are the maternal coordinate genes, such as *bicoid* (*bcd*) and *caudal*, in broad anterior and posterior gradients that define the axes of the egg. The maternal effect genes regulate the gap genes, such as *Krüppel* (*Kr*) and *giant* (*gt*), which are expressed in broad overlapping domains covering several contiguous segments of the developing embryo. The gap genes, in turn, regulate expression of the pair-rule genes (i.e. *even-skipped*, *runt*) in a double segmental manner which ultimately leads to the expression and regulation of the segment polarity genes, *wingless* (*wg*) and *engrailed* (*en*), that establish anteroposterior identity within each segment ('polarity') and delineate the borders between the developing parasegments.

In *Drosophila*, the maternally loaded *bicoid* gene is expressed as a gradient that establishes the anterior axis. *bcd* is a relatively new gene found only in Diptera, where in other organisms the anterior end is specified by *orthodenticle* (*otd*) and *hunchback* (*hb*) (Lynch et al., 2006; Pultz et al., 2005; Rosenberg et al., 2009). Bcd inhibits the anterior spread of *cad* expression leading to the formation of a posterior *cad* gradient that, along with *nanos*, establishes the posterior (Rosenberg et al., 2009; Wang and Lehmann, 1991). The maternal component of *Dm-cad* has a stronger effect on overall body patterning (Macdonald and Struhl, 1986), but the zygotic component has relatively minimal functions, in the posterior-most segments only (Moreno and Morata, 1999). *cad* plays a larger role in other long germ organisms, such as the parasitic wasp *Nasonia vitripennis* (Hymenoptera), where loss of function results in a "head only" phenotype (Olesnicky et al., 2006). The Wnt-Cad posterior organiser found in vertebrates, annelids, and basal arthropods has been lost in long germ band organisms like *Drosophila*, suggesting a

decoupling of posterior growth and segmentation mechanisms, yet remnants of this relationship remain (Vorwald-Denholtz and De Robertis, 2011; Wu and Lengyel, 1998).

The maternal effect genes regulate the broad domains of gap gene expressions along the length of the syncytial blastoderm. The gap gene expression domains overlap and interact with each other to establish expression of the genes in the next step of the segmentation hierarchy: the pair-rule genes. The pair-rule genes are expressed in a 'classic' pattern of seven stripes in alternate parasegments and are, themselves, split into a small hierarchy based on the timing of their expressions (Jaynes and Fujioka, 2004). In *Drosophila*, the first to be expressed are the primary pair-rule genes *even-skipped* (*eve*), *runt* (*run*), and *hairy* (*h*), which are directly regulated by the gap genes. In turn, the primary pair-rule genes, along with the gap genes, regulate the expression of the secondary pair-rule genes *paired* (*prd*), *sloppy-paired* (*slp*), *odd-skipped* (*odd*), and *fushi tarazu* (*ftz*) (Gutjahr et al., 1993; Jaynes and Fujioka, 2004; Peel et al., 2005). The mostly inhibitory interactions between the pair-rule genes are highly complex and differentially regulate the expression of the segment polarity genes *wg* and *en*, defining the borders between adjacent parasegments and future embryonic segments, thus, finalizing the segmentation cascade (Jaynes and Fujioka, 2004).

First described in hypomorphic mutants of *Drosophila*, the 'classic' pair-rule phenotype is embryos lacking every other body segment (Nüsslein-Volhard and Wieschaus, 1980). Subsequent studies showed that the expression patterns of the pair-rule genes correspond to their mutant phenotypes (Gergen and Butler, 1988; Grossniklaus et al., 1992; Holmgren, 1984; Kilchherr et al., 1986; Macdonald and Struhl, 1986). Regulation of the pair-rule genes is the *key* step in long germ band segmentation as it is here that the first sign of a periodic pattern is observed. Initially expressed as seven stripes in alternating segments, each pair-rule gene is eventually expressed in all fourteen future

segments. However, there is some variation in pair-rule expression patterns between the long germ band insects, such as the rapid sequential formation of *eve* stripes in the endoparasitic wasp *Copidosoma floridanum* (Grbic et al., 1996) and stripes of *prd* the honey bee *Apis mellifera* (Osborne and Dearden, 2005).

The underlying mechanisms involved in achieving a segmented body are fairly well understood in *Drosophila*. However, as *Drosophila* undergoes the highly derived and specialised long germ band mode of segmentation, gene expression and consequent functions may not be generally applicable to all arthropods. By studying elements of the *Drosophila* paradigm (maternal → gap → pair-rule → segment polarity) in other organisms, we can begin to understand how the evolution of body patterning mechanisms changed over time and to draw a clearer picture of the putative common ancestor of insects and other arthropods.

Short germ band segmentation

The more common and presumed ancestral mode of segmentation in arthropods is short germ band (Davis and Patel, 2002; Liu and Kaufman, 2005b). In this mode of development, the head and anterior thoracic segments are specified 'all-at-once' at the syncytial blastoderm stage, similar to that described for long germ band insects, while the remaining segments are added sequentially from the extending posterior GZ (Fig. 1.3B) (Davis and Patel, 2002; Liu and Kaufman, 2005b). In addition to the Wnt-Cad posterior organiser, elements of Notch-signalling are becoming recognised as important factors in short germ band arthropod segmentation, similar to their counterparts in vertebrates and annelids. Homologues of *N*, *Dl*, and *h* are expressed in the posterior growth zone of the spiders *Cupiennius salei* (Schoppmeier and Damen, 2005b; Stollewerk, 2002; Stollewerk et al., 2003) and *Achaearanea* (Oda et al., 2007); the myriapods *Strigamia* (Chipman and

Akam, 2008), *Glomeris* (Dove and Stollewerk, 2003), and *Lithobius forficatus* (Kadner and Stollewerk, 2004); the crustacean *Parhyale hawaiensis* (O'Day, 2006); and the basal insects *Periplaneta americana* (Pueyo et al., 2008) and *Gryllus bimaculatus* (Kainz et al., 2011; Mito et al., 2011).

Functional analyses, where possible, have shown the importance for Dl/N-signalling and Hairy in the segmentation process of short germ band organisms. Previous work in our lab has shown that segment formation is disrupted in *Periplaneta* upon the loss of *Pa-N* or *Pa-h* expression resulting in segmentation defects and posterior truncation (Pueyo et al., 2008). Similar results were obtained in the spiders *Cupiennius* and *Achaearanea* (Oda et al., 2007; Stollewerk et al., 2003) and the crustaceans *Artemia* (Williams et al., 2012) and *Parhyale* (O'Day, 2006). This research adds credence to the notion of an evolutionarily conserved N-mediated segmentation mechanism shared between most arthropods, vertebrates, and annelids, which was subsequently lost in highly derived organisms, such as *Drosophila* and other holometabolous insects.

The loss of N-mediated segmentation in Holometabola (Aranda et al., 2008; Tautz, 2004; Wilson et al., 2010b) may have occurred around the time of the split between holometabolous and hemimetabolous insects over 350 million years ago. Even though it has a short germ band mode of segmentation, the holometabolous insect *Tribolium castaneum* does not utilize N-signalling to pattern its segments. This may seem counterintuitive compared to the other short germ arthropods in which N-signalling has a major role; however, *Tribolium* has developed a novel mechanism in the pair-rule gene circuit (Choe and Brown, 2009; Choe et al., 2006). This circuit involves the cyclic activation and repression of the primary pair-rule genes *eve*, *run*, and *odd* in the posterior growth zone during germ band elongation. In this circuit, the initial expression of *Tc-eve* activates the expression of *Tc-run*, which then activates *Tc-odd* expression. *Tc-odd*, in

turn, inhibits the expression of *Tc-eve*, thus completing the circuit (*eve* → *run* → *odd* | *eve*). Recently, Sarrazin et al. (2012) and El-Sherif et al. (2012) showed that *Tc-odd* and *Tc-eve*, respectively, oscillate spatiotemporally and the moving waves of expression through the posterior GZ are part of a cyclic segmentation clock mechanism, thus signifying the importance of an oscillator in regulating sequential addition of segments during short germ band development.

In short germ band organisms, the anterior-most segments form rapidly in a syncytial blastoderm, while the remaining segments are added sequentially from the posterior growth zone. While the posterior addition of segments is regulated in a Notch-dependent manner, the anterior segments form independent of N-signalling. This Notch-independent segmentation mechanism is likely the precursor to what exists in higher insects today. As new methods were gained to speed up embryonic development (i.e. meroistic ovaries and nurse cells), addition of segments from a posterior GZ became less important as more segments formed in the syncytial blastoderm. Along with this, the requirement for Notch became unnecessary as signalling now occurred in a cell-free environment allowing the gap genes to take over control of pair-rule gene expression and, consequently, segmentation (Damen, 2007; Peel, 2004). The expression patterns of the pair-rule genes are highly variable, yet there seems to be some conservation in the pair-rule gene hierarchy and the regulation of the segment polarity genes (Choe and Brown, 2009; Damen et al., 2005; Janssen et al., 2011; Jaynes and Fujioka, 2004), demonstrating the plastic and adaptable nature of these genes during segment formation.

Still, it appears as though the function of the pair-rule genes in segmentation may be restricted to the arthropods. Homologues of the pair-rule genes have been found in annelids and vertebrates, but may not function during segmentation/somitogenesis (Cruz et al., 2010; Jostes et al., 1990; Seaver et al., 2012). On the other hand, neural expression

of many pair-rule genes is conserved among vertebrates, annelids, and arthropods and may represent the true ancestral nature of these genes (Inoue et al., 2002a; Moran-Rivard et al., 2001; Patel et al., 1992; Song et al., 2002). Through the course of evolution, the pair-rule genes may have gained a new function in arthropod body segmentation, initially regulated by N-signalling before coming under control of the gap genes (Damen, 2007; Peel, 2004); however, additional functional analysis is needed in phylogenetically basal organisms to determine the ancestral state.

tarsal-less: *a new segmentation gene?*

Initially identified as a putative non-coding RNA (Tupy et al., 2005), further investigation has shown the non-canonical, polycistronic gene *tarsal-less* (also called *polished rice*) is translated into several short peptides through small Open Reading Frames (smORF) of less than 100 amino acids (Galindo et al., 2007; Kondo et al., 2007; Savard et al., 2006). In *Drosophila*, *tal* encodes three different peptides: Type-A, AA, and B. The Type-A and Type-AA peptides contain a conserved motif of LDPTGXY and have been shown to function, non-autonomously, during embryonic and post-embryonic development (Galindo et al., 2007; Kondo et al., 2007; Pueyo and Couso, 2008; Pueyo and Couso, 2011), while the Type-B peptide is non-functional/non-translated (Galindo et al., 2007).

Embryonic expression of *Dm-tal* is dynamic and is required for the proper formation of the trachea, cephalopharyngeal skeleton, and posterior spiracles (Galindo et al., 2007; Kondo et al., 2007). Although expressed in several stripes in the developing embryo, these segmental stripes do not function in segmentation, instead *Dm-tal* in this location correlates with and is necessary for development of the future denticle belts (Galindo et al., 2007; Kondo et al., 2007). *Dm-tal* function in denticle belt formation is

carried out through regulation of an active form of the transcription factor *shavenbaby* (*svb*); loss of *Dm-tal* expression leads to loss of denticle belt formation as a result of the defective organisation of filamentous actin (F-actin) bundles (Chanut-Delalande et al., 2006; Delon et al., 2003; Galindo et al., 2007; Kondo et al., 2007; Kondo et al., 2010). A similar loss of F-actin bundle organisation leads to a breakdown in taenidial folding and truncated trachea in *Dm-tal* mutants (Galindo et al., 2007; Kondo et al., 2007). However, in this instance, the function of *Dm-tal* is not mediated through *svb*, indicating that *tal* function in ectoderm morphogenesis may operate through different modes in varying developmental contexts (Galindo et al., 2007; Kondo et al., 2007; Kondo et al., 2010; Pueyo and Couso, 2008; Pueyo and Couso, 2011). The post-embryonic functions of *Dm-tal* can be separated into three main functions: 1) tarsal joint formation (Pueyo and Couso, 2011); 2) formation of trachea (Galindo et al., 2007; Pueyo and Couso, 2008); and 3) wing development (Galindo et al., 2007; Pi et al., 2011), of which tarsal leg development will be the focus of the following section.

A *tal* homologue, called *mille-pattes* (*mlpt*), has been found in the short germ band insect *Tribolium*. *Tc-mlpt* encodes for several small peptides that contain the conserved heptapeptide sequence, LDPTGXY (Savard et al., 2006). Similar to *Dm-tal*, *Tc-mlpt* expression is dynamic; however, *Tc-mlpt*^{RNAi} revealed that this gene has differing functions than *Dm-tal*, being required for proper body patterning. *Tc-mlpt* functions in *Tribolium* as a gap gene and, indeed, it has been shown to regulate and be regulated by the other gap genes *Tc-hunchback* (*hb*), *Tc-Krüppel* (*Kr*), and *Tc-giant* (*gt*) (Savard et al., 2006). Recently work by Schnellhammer (2012) indicates a possible interaction between *Tc-mlpt* and a *Tribolium* homologue of *shavenbaby* during segment patterning. Loss-of-function of either gene leads to stumpy legs with malformed joints and truncation of the posterior-most abdomen with the remaining segments taking on a thoracic identity,

including ectopic legs (Savard et al., 2006; Schnellhammer, 2012). With the exception of *Tribolium-mlpt*, the role(s) of this gene have not been studied in other organisms.

Bioinformatic searches of the short peptide motif LDPTGXY revealed *tal* homologues in many insect species and at least one crustacean (Galindo et al., 2007; Savard et al., 2006). These *tal* homologues are variable in length and amino acid content outside of the conserved motif and there appears to be a tendency of increasing copy numbers of *tal* smORFs, from one in crustaceans to an average of three or four in holometabolous insects (Galindo et al., 2007). The variations in sequence, along with the deviations in function between *Tribolium* and *Drosophila*, illustrate the evolvability of this small gene and the need to analyse this, and other smORF genes, in more basal species.

Arthropod leg patterning

Segmented appendages are the key defining characteristic of the Arthropoda and gave the phylum its name (arthro – joint, podos – leg). Arthropod legs, like their bodies, display innumerable morphologic diversity, having been extensively modified and adapted for various functions (running, jumping, swimming, etc.). Though outwardly different in appearance, all arthropod legs are subdivided into several smaller units (podomeres), separated by joints that allow flexibility and ease of motion. The mechanisms of leg development are remarkably conserved, involving the 'leg gap' genes for proximodistal regionalization (Angelini and Kaufman, 2005b; Jockusch et al., 2000; Niwa et al., 2000; Panganiban et al., 1994) and N-signalling in border and joint formation (Bishop et al., 1999; de Celis et al., 1998; Prpic and Damen, 2009; Rauskolb and Irvine, 1999).

Early expression of the leg gap genes *homothorax* (*hth*), *dachshund* (*dac*), and *Distal-less* (*Dll*) respectively pattern the proximal, medial, and distal regions of the developing legs (Abzhanov and Kaufman, 2000; Angelini and Kaufman, 2005b; Blagburn,

2007; Couso and Bishop, 1998; Inoue et al., 2002b; Pechmann et al., 2010; Prpic et al., 2003; Prpic and Tautz, 2003; Wu and Cohen, 1999). Further patterning of the leg involves N-signalling in defining segment borders, which may represent a phylotypic trait in arthropods (Prpic and Damen, 2009; Rauskolb and Irvine, 1999).

Leg development has been well studied in *Drosophila*, in which the legs are composed of five basic segments; from proximal to distal: coxa (cx), trochanter (tr), femur (fe), tibia (ti), and tarsus (ta – itself subdivided into smaller tarsomeres). In *Drosophila* and other holometabolous insects, the legs are derived from thickened invaginated pockets of undifferentiated cells contained within the larval epidermis called imaginal discs. By the end of larval development and through pupariation, concentric rings of gene expression lead to a series of concentric folding of leg tissue, which later evert as fully formed legs. The genes that pattern fly legs are the same as those used to pattern the direct developing legs (limb buds) of lower insects and arthropods during embryogenesis. These include the leg gap genes and members of the N-pathway, such as *Dl* and *Serrate* (*Ser*), important for appendage growth and segmentation (Angelini et al., 2012; Bishop et al., 1999; de Celis et al., 1998; Dearden and Akam, 2000; Greenberg and Hatini, 2009; Mito et al., 2011; Pechmann et al., 2010; Prpic and Damen, 2009; Rauskolb and Irvine, 1999).

In *Drosophila*, *Notch* regulation of leg segmentation varies between the 'true' joints containing muscle attachments (cx, tr, fe, and ti) and the ball-and-socket joints of the tarsomeres. Downstream targets of N-signalling in *Drosophila* leg development include genes important in leg growth and controlling joint intercalation/formation such as *odd-skipped related* (*odd-r*), *Enhancer of split* (*E(spl)*), and *nubbin* (*nub*) (Bishop et al., 1999; de Celis et al., 1998; Hao et al., 2003; Rauskolb and Irvine, 1999). In the true joints, *Dl/Ser* and *N* are expressed in the distal end of each podomere where the genes *odd-r*, *lines*, and *bowl* form a negative feedback mechanism that maintains a boundary between

Dl/Ser expression and the N-responsive region, consequently promoting growth, defining the leg segment, and regulating joint morphogenesis (Greenberg and Hatini, 2009; Kojima, 2004).

Also of note is the N-regulated expression of *nubbin* (a POU homeodomain transcription factor) at the distal end of each true jointed leg segment (Greenberg and Hatini, 2009; Rauskolb and Irvine, 1999). Loss of *Dm-nub* leads to the loss of joint definition and fusions between adjacent podomeres, except for most of the tarsal sub-segments (Turchyn et al., 2011). Although expressed in a broad domain across the presumptive tarsi at early third instar, *Dm-nub* is no longer expressed at the time of tarsomere folding and joint formation, except tarsomere-5 (Natori et al., 2012; Turchyn et al., 2011). Thus, *Dm-nub* functions in the tarsi as a temporal modulator of tarsal gene expression, but not in the actual process of tarsal joint formation (Natori et al., 2012).

One of the main components of tarsal patterning in the fly is the N-regulation of *tarsal-less*, which then goes on to activate the expression of the zinc-finger transcription factor *rotund* (*rn*) (Natori et al., 2012; Pueyo and Couso, 2008). Through mid-third instar, the progressive clearance of *Dm-nub* from the medial tarsus allows derepression of *rn* expression, which, in turn, inhibits *Bar*, leading to the sequential formation of the tarsal subsegments (Kojima et al., 2000; Natori et al., 2012). During pupation, *Dm-tal* is also involved in the intercalation and formation of tarsal joints by modulating the expression of an active form of *shavenbaby* (Galindo et al., 2007; Kondo et al., 2010; Pueyo and Couso, 2008; Pueyo and Couso, 2011). *Dm*-Svb inhibits the expression of *Delta*, thus forming a negative feedback loop between Tal and N, leading to the formation of a sharp Dl+/Dl- border at the future tarsal joint boundaries (Pueyo and Couso, 2011). Depleted *Dm-tal* expression leads to loss of tarsomeres, while gain-of-function of *Dm-tal* leads to the development of ectopic joints, confirming the function of *Dm-tal* in tarsal development

and joint formation (Galindo et al., 2007; Pueyo and Couso, 2008). In *Drosophila*, *tal* is not expressed in the "true joints" and could be an evolutionary novelty in specifying the non-muscle joints of the tarsi, acting as a link between patterning and morphogenesis in this location. In brief, the true joints require Notch regulation of *nub* and *odd-r* expression for proper formation, where the tarsomeres are formed by the initial activation of *tal* by Notch, which leads to the cleavage of the long form of Svb into its active form required for proper joint development.

Leg patterning in other arthropods has not been as extensively studied as in *Drosophila*. Outside of the conserved expression and function of the leg gap genes mentioned above, only a few studies have shown a requirement for N-signalling in leg patterning in other arthropods (Angelini et al., 2012; Mito et al., 2011; Prpic and Damen, 2009). As a putative downstream target of Notch, *nubbin* is variably expressed in the developing appendages of many arthropods, at the distal end of all, or most, podomeres (Hrycaj et al., 2008; Li and Popadić, 2004; Popadic, 2005; Prpic and Damen, 2005), and is important for proper appendage patterning (Prpic and Damen, 2005; Turchyn et al., 2011). Recently, Turchyn et al. (2011) have further described the divergence in expression and function of *nub* in *Acheta domesticus* (house cricket), *Periplaneta*, and *Drosophila*. Their results illustrate that although *nub* is expressed in all (crickets and roaches) or most (flies) leg segments, the major effect of the gene differs in location in different species. *Acheta-nub* functions mainly in the tibia and first tarsus, while *Periplaneta-nub* is more essential in the coxa, trochanter, and femur (Turchyn et al., 2011). *Drosophila-nub* is important in all leg joints except tarsi. Additionally, it was shown that the expression and function of *nub* in joint formation is reliant on Notch signalling in the cockroach, a mechanism that is conserved in flies and spiders (Pechmann et al., 2010; Rauskolb and Irvine, 1999; Turchyn et al., 2011).

The mechanisms of leg segmentation have been well studied in *Drosophila*, but less so in other organisms. In *Drosophila*, *nub* and *tal* have complementary roles in joint formation in true joints versus tarsi, respectively. Conversely, *Tribolium-tal* is expressed in all podomeres and loss of expression via *Tc-tal*^{RNAi} resulted in larvae with short, stubby legs (Savard et al., 2006). Expression and function of Notch and Nubbin in leg development appear to be conserved and there is a potentially conserved role for the newly discovered gene, *tarsal-less*, in leg patterning. Having mainly been studied in *Drosophila* leg and denticle belt formation and *Tribolium* body patterning, the expression and function(s) of this gene have yet to be examined and compared in more basal arthropod species.

Thesis Aims

In this thesis, I present results of my doctoral research investigating the mechanisms of segmentation in the American cockroach, *Periplaneta americana*. *Periplaneta* is a suitable candidate for these investigations, as elucidation of these mechanisms in a phylogenetically basal, short germ band organism will help to establish putative ancestral gene expression and function, allowing us to gain further insights into the evolution of segmentation. Cockroaches are a good system in which to study these mechanisms, as they are highly susceptible to RNA interference (RNAi), embryos can be cultured *ex ovo*, and within a single ootheca there is an age gradient where younger embryos differ from older ones by approximately one half segment. These advantages were instrumental in showing that the sequential addition of segments in *Periplaneta* is comparable to vertebrate somitogenesis involving cyclic waves of Notch-signalling from the posterior (Pueyo et al., 2008); this feature is important when drawing evolutionary corollaries between closely or distantly related species. This thesis is organised into three main results sections discussing the following aims:

1) Investigate the Wnt-Cad posterior organiser and analyse any potential relationship with N-mediated segmentation.

The function of the Wnt-Cad posterior organiser in axis formation and growth is generally accepted to be true for all bilaterians and the function of N-signalling in segmentation/somitogenesis is conserved among the segmented phyla. There is an intimate relationship between posterior growth and segmentation; however, the unification of the Wnt-cad organiser and N-signalling has rarely been examined. In Chapter III, I discuss the interactions between these networks and show that they are dependent on each

other in order to maintain posterior growth and achieve proper segmentation during *Periplaneta* embryogenesis.

2) Examine the role of the pair-rule genes in anterior and posterior segmentation.

The pair-rule genes are important in segmentation of both long and short germ band arthropods, though it is unclear whether the ancestral state of these genes is single or double segmental. Between species, expression patterns vary widely and even within the same embryo there can be differences between expression and function in the anterior and the posterior. I set out to investigate the patterns of expression and putative functions of four pair-rule genes during *Periplaneta* embryogenesis: *even-skipped*, *runt*, *paired*, and *sloppy-paired*. Chapter IV will discuss the requirement for these pair-rule genes in proper anterior segmentation and their apparent semi-redundant functions in posterior patterning in the cockroach.

3) Isolate and analyse the expression and function of the smORF gene *tarsal-less*.

While there is considerable conservation in *tal* sequences between the different arthropod species, this gene has only been studied in two holometabolous insects, *Drosophila* and *Tribolium*. The expression patterns and functions are divergent between these organisms during body segmentation, but may be conserved in leg patterning. I set out to isolate a *tal* homologue in *Periplaneta* in order to determine any conserved mechanisms shared between these three species. *Periplaneta* is the first hemimetabolous insect in which this gene has been studied (Chapter V), showing conserved expression in the legs, but divergent expression and function in body patterning.

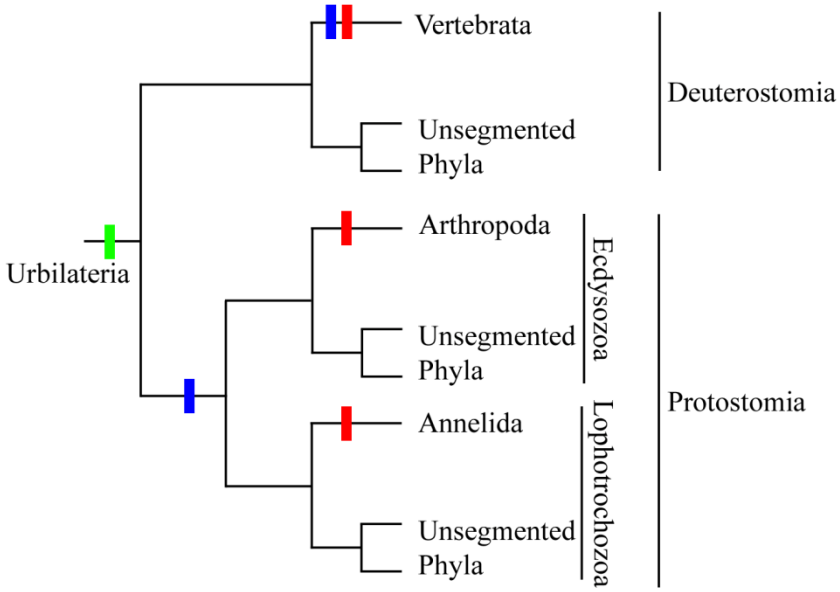


Figure 1.1: Theories on the origin of bilaterian segmentation. A segmented body plan is a common theme amongst the Bilateria, found in the three most successful groups, i.e. vertebrates, annelids, and arthropods. The origin of segmentation is contested and three main views exist: (1) each group evolved segmentation mechanisms independently (red bars), (2) segmentation evolved separately in the deuterostomes (vertebrates) and protostomes (blue bars), or (3) it may have already been present in the last common bilaterian ancestor (green bar). Adapted from Peel and Akam (2003).

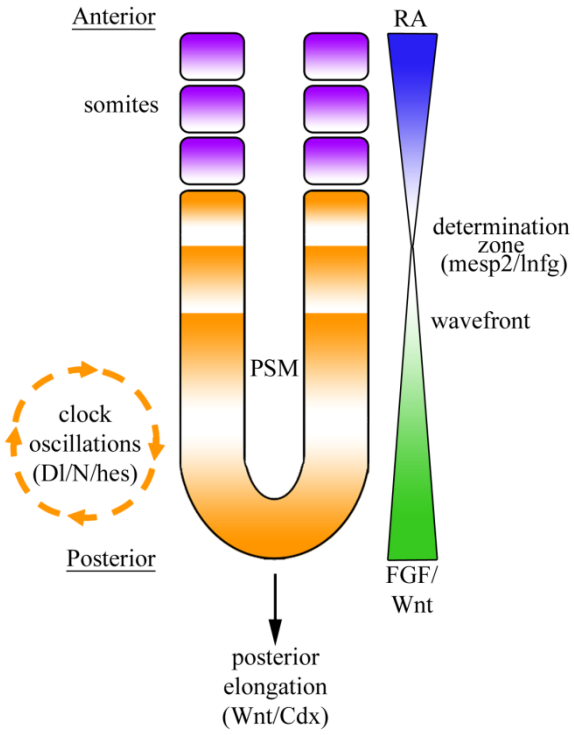


Figure 1.2: Vertebrate 'clock and wavefront' model of segmentation. Vertebrate embryos grow and new segments emerge from the posteriorly extending presomitic mesoderm (PSM). Posterior elongation involves Cdx/cad and Wnt-signalling. Periodic waves of the N-based clock mechanism (orange), involving the oscillatory activation/inhibition of *Delta*, *Notch*, and *hes/her*, move through the PSM resolving into stripes of expression in the anterior leading to boundary and somite formation (purple). Spatiotemporal regulation of somitogenesis occurs through two opposing gradients: 1) a posterior-to-anterior 'wavefront' of FGF and Wnt signalling (green triangle) that is opposed by 2) the inhibitory effects of retinoic acid (RA; blue triangle) expressed in a counter anterior-to-posterior gradient. Where these gradients meet, the determination zone, is where the temporal N-clock is translated into a spatial pattern of gene expressions involved in boundary formation (i.e. *mesp2* and *lfng*).

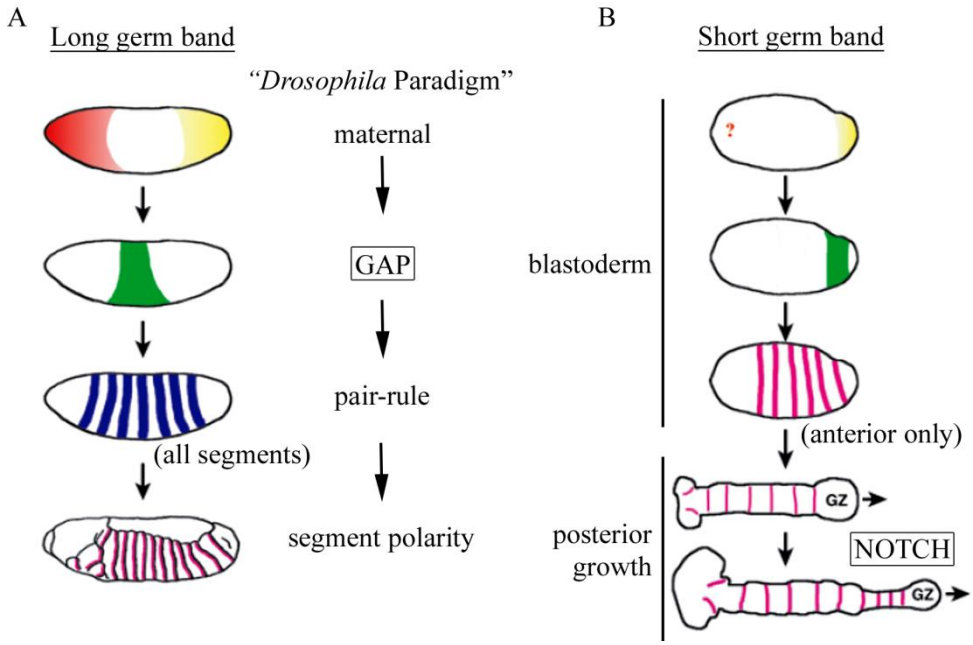


Figure 1.3: Two general modes of arthropod embryonic segmentation, short and long germ band. (A) Long germ band insects are highly derived and all body segments are specified at the same time in the blastoderm. This process is controlled through a highly complex cascade of transcription factors called the "*Drosophila* paradigm". At the top of this cascade are the maternal effect genes that function to establish the AP axis and to regulate the downstream gap genes. Gap genes are expressed in broad domains within the blastoderm and, in turn, regulate the pair-rule genes in the 'classic' double-segment pattern of expression. The pair-rule genes then go on to regulate the segment polarity genes that define parasegment boundaries and give separate anterior and posterior identities within each segmental unit. (B) In contrast, during the short germ band mode of development, only the anterior-most head and some thoracic segments are formed in the early blastoderm, while the remaining posterior segments are added sequentially from a posteriorly extending growth zone, a process likely to be regulated through N-signalling. Modified from Liu and Kaufman (2005b).

CHAPTER II

MATERIALS AND METHODS

A. Animal rearing

A *Periplaneta americana* colony was kept in the laboratory at 29ºC and provided with a regular supply of tap water and Cockroach Diet (Blades Biological Ltd. – Cat#DTS079). Freshly laid oothecae were collected and placed in a humidified incubator at 29º C allowing the embryos to develop to the required stage for dissection (according to Lenoir-Rousseaux and Lender, 1970) or were allowed to hatch to first nymph.

Gryllus bimaculatus eggs were kindly provided by Robert Ray from his laboratory colony. Eggs were collected and placed in a petri dish on wet paper towel for humidity and embryos allowed to develop at 29ºC until the required stage (according to Niwa et al., 1997) before dissecting.

B. Embryo dissection and fixation

Periplaneta and *Gryllus* embryos at the desired stages were dissected in a watch glass containing nuclease free 1X PBS (Phosphate Buffered Saline; Roche – Cat#11666789001). For *Periplaneta*, the top "zipper" part of the ootheca was removed with scissors and the two halves were pried apart using sterile forceps. Each ootheca contains 12-18 eggs, which were individually peeled open to expose the embryo within. After removing the surrounding yolk, embryos were transferred to a 0.5 ml eppendorf tube and fixed with 4% paraformaldehyde in a slow moving rotator at room temperature for one hour (for *in situ* hybridisation) or 20 minutes (for antibody staining). The embryos were then washed several times with 1X PBS, and dehydrated in an increasing ethanol series (30%, 50%, 70%, 90%, and 100%) and stored at -20ºC until ready for use.

Gryllus eggs are laid individually and embryos were removed, in 1X PBS, by cutting off the anterior end of the egg and either peeling the rest open or gently pushing the embryo through the opened end with forceps. Cricket embryos were fixed and stored in the same manner as cockroach embryos.

C. RNA extraction

Total RNA was extracted from *Periplaneta* or *Gryllus* embryos of mixed stages using either of two methods. The first method used the RNAqueous®-4PCR kit (Ambion – Cat#AM1912) following the manufacturer's protocol. Briefly, 40-60 cockroach embryos or cricket eggs were homogenized in a nuclease-free 1.5 mL eppendorf tube using a sterile pestle. The cells were lysed in a solution containing guanidium thiocynate, which also inactivates ribonucleases. After the addition of ethanol, the homogenate was passed through a silica-based filter cartridge to which the RNA is bound. Proteins, DNA, and other contaminants were removed through several washes with reagents provided in the kit and RNA was recovered in a sterile 1.5 mL tube by the addition of the Elution Solution (Ambion). Any residual DNA contamination was removed with DNase1 treatment at 37ºC for 30 minutes.

The second method for RNA isolation used the TRIzol® reagent (Invitrogen – Cat#15596-026), which contains a mixture of phenol, guanidium thiocynate, and ammonium thiocynate for fast, 'one-step' RNA isolation. With this method, the embryos were homogenized and lysed using a sterile pestle in a 1.5 mL eppendorf tube containing a small amount of TRIzol® reagent. After the addition of chloroform:isoamyl (24:1) the tubes were incubated for 5 minutes at room temperature, spun to separate phases, and the resulting supernatant was transferred to a sterile, nuclease-free tube. In subsequent steps, the RNA was precipitated with isopropanol and washed with ethanol before being

resuspended in nuclease-free H_2O. Visualisation of both RNA extraction techniques was done via agarose gel electrophoresis (section G) and RNA was stored at -80ºC.

D. cDNA synthesis

cDNA was synthesized using the RETROscript® kit (Ambion – Cat#AM1710M) following the protocol for the 'Two-step RT-PCR with heat denaturation of RNA' procedure provided by the manufacturer. One microgram total RNA was combined with either random decamers or oligo(dT) primers and nuclease-free water, then denatured at 80ºC before the addition of the remaining RT reagents: 10X RT buffer, dNTPs, RNase inhibitor, and the M-MLV Reverse Transcriptase. Reverse transcription of cDNA was done at 42ºC for 1-2 hours. The reaction was stopped by inactivating the reverse transcriptase at 92ºC for 10 minutes. Newly synthesized cDNA was stored at -80ºC until ready to use in PCR reactions. Visualisation of the newly synthesised cDNA was done via agarose gel electrophoresis, as described in section G.

E. Primer design

Previously published gene sequences from insects and other arthropod species were obtained from GenBank (http://www.ncbi.nlm.nih.gov/) and aligned using Lasergene® MegAlign™ sequencing analysis software (DNASTAR) looking for highly conserved regions. Degenerate primers were designed by hand towards these conserved sequences and prepared by Invitrogen Custom Primers (http://www.invitrogen.com). Upon receipt, all primers were resuspended in sterile water to a stock concentration of 100 µM from which a 10 µM working solution was prepared; both were stored at -20ºC.

Where possible, primer degeneracy was reduced to allow for fewer incidents of non-specific priming by analysing codon usage bias in *P. americana* via the Codon Usage

Database (Appendix 1) (http://www.kazusa.or.jp/codon) (Nakamura et al., 2000). Gene fragments resulting from degenerate PCR were sequenced and used to design gene-specific primers for use in future PCR reactions in order to obtain longer or full-length gene sequences. A complete list of degenerate and specific primers used for each gene can be found in Appendix 2.

F. Polymerase Chain Reaction (PCR)

PCR reactions were conducted using reagents provided in the *Taq* PCR Core Kit (QIAGEN – Cat#201225). PCR reactions were prepared on ice to a total volume of 100 µl as follows: 5 µl 10X PCR buffer; 10 µl 5X Q-Solution; 2-4 µl $MgCl_2$ (25mM); 2 µl dNTP mix (10 mM each); 2 µl each forward/reverse specific primer (5 µl degenerate); 2 µl template cDNA; 0.5 µl *Taq* DNA polymerase (5 Units/µl); H_20 up to 100 µl. Final concentrations of some reagents, such as $MgCl_2$ and dNTP, were adjusted where necessary to improve optimal target amplification. Nested primer pairs were used in order to decrease the number of non-specific target fragments in degenerate PCR reactions – first round PCR using outer primers and second round PCR using inner (nested) primers. PCR was performed using either a Techne TC-3000 or an eppendorf Mastercycler Gradient thermocycler. Cycling conditions were optimized based on use of degenerate or specific primers, melting temperatures (Tm), and target sequence length. Standard PCR conditions:

1 cycle:	extended DNA denaturation at 94ºC – 5 minutes
30 cycles:	denaturation at 94ºC – 30 seconds
	annealing at 3-5ºC below average primer Tm – 30 seconds
	extension at 72ºC – 30 seconds to 2 minutes (depending on target length)
1 cycle:	extended extension at 72ºC – 10 minutes
Hold:	4ºC

Visualisation of PCR products was done via agarose gel electrophoresis as described in section G.

G. Agarose gel electrophoresis

Visualisation of the various RNA, cDNA, and PCR products was done via agarose gel electrophoresis. Gels were prepared by combining 0.5-1.5% agarose, based on expected product size, in 1X TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA) and heated in a microwave to dissolve. For visualisation, 0.5 µg/ml ethidium bromide was added to the liquid agarose before pouring into the gel cast. The DNA or RNA product was combined with MassRuler^{TM} DNA Loading Dye (Fermentas – Cat#SM0403) at a proportion of 1 µl dye for every 5 µl product, and loaded into the wells of the agarose gel alongside the MassRuler^{TM} DNA Ladder Mix (80-10,000 bp fragments; Fermentas – Cat#SM0403) to assess length of product and calculate approximate sample concentration. Gel pictures were taken using an Uvidoc gel documentation system (Uvitec Cambridge) and UviPhotoMW image analysis software.

H. Gene cloning

Upon successful PCR amplification, the desired gene fragment was cloned using the pCR®4-TOPO TA cloning kit (Invitrogen – Cat#K4575-02) and transformed, via heat-shock, into One Shot® TOP10 competent *E. coli* cells (Invitrogen – supplied with kit) following the manufacturer's instructions. The plasmid cloning vector contains antibiotic resistance genes and those bacteria successfully transformed are selected for by first growing overnight at 37ºC on Luria-Bertania (LB) agarose plates containing the antibiotic Kanamycin (50 µg/ml). Individual colonies were picked and transferred with sterile toothpicks into separate bottles of liquid LB media containing Ampicillin (100 µg/ml) in

which they were allowed to grow overnight at 37ºC with constant shaking. On the following day, the plasmids were isolated via miniprep (section I).

I. Minipreparation of plasmid DNA

Plasmid DNA was isolated from bacterial cultures using the QIAprep® Spin Miniprep Kit (QIAGEN – Cat#27106) following the instructions provided by the manufacturer. Two millilitres of overnight culture were spun to pellet the cells and the remaining supernatant was poured off. The cells were lysed in an alkaline solution and cleared via centrifugation. The DNA containing supernatant was transferred to a spin column containing a silica membrane to which the DNA is adsorbed. After several clean-up washes using the kit provided reagents, DNA was removed from the membrane using the low-salt Elution Buffer provided.

The TOPO plasmid vector contains *EcoRI* restriction sites on either side of the insertion site allowing the inserted gene fragment to be excised and isolated by digesting the plasmid with the *EcoRI* restriction enzyme (Invitrogen) for several hours at 37ºC. Confirmation of plasmid insert excision was done via agarose gel electrophoresis (section G). Aliquots of plasmids containing the putative insert were prepared and sent for sequencing with the remaining plasmid stored at -20ºC.

J. Sequencing and phylogenetic analysis

All plasmids were sent for sequencing to Eurofins MWG Operon (http://eurofinsdna.com). Approximately 1-2 µg of plasmid DNA template was sent in nuclease-free water to be sequenced in the forward direction using the universal forward M13uni(-21), that recognise regions on the plasmid flanking the insert, provided by the

company. Larger fragments (>1000 bp) were also sequenced in reverse using the M13rev(-29) primer. Primer sequences can be found on the Eurofins website.

Upon receipt, sequences were analysed using the Lasergene® Software Suite for Sequence Analysis (DNASTAR). Initial trimming of the extraneous plasmid sequence was done using the EditSeq™ program, leaving the insert sequence which was then subjected to BLAST (Basic Local Alignment Search Tool – http://blast.ncbi.nlm. nih.gov/Blast.cgi) (Altschul et al., 1990) analysis to confirm homology of the isolated *Periplaneta* gene. Upon confirmation, orthologous protein sequences from other arthropods were acquired from GenBank (http://www.ncbi.nlm.nih.gov/genbank) and used for alignment and phylogenetic comparison via MegAlign™ software (DNASTAR).

K. 5’ and 3’ RLM-RACE

Periplaneta cDNA sequences resulting from degenerate PCR were extended into the 5’ and 3’ directions using the FirstChoice® RNA Ligase Mediated-Rapid Amplification of cDNA Ends (RLM-RACE; Ambion – Cat#AM1700) technique following the manufacturer’s protocol. In both methods, 1 µg of *Periplaneta* total RNA was used as a template along with gene specific primers (Appendix 2) and the adaptor specific primers provided with the kit. For 5’RACE, the 5’-phosphate from incomplete/uncapped RNAs was first removed using Calf Intestine Alkaline Phosphatase (CIP). Complete, full-length mRNAs were treated with Tobacco Acid Pyrophosphatase (TAP) to remove the cap, leaving a monophosphate at the 5’end to which an adaptor oligonucleotide is ligated; this adaptor cannot be ligated to the previously dephosphorylated, incomplete RNAs. Full-length mRNAs are reverse transcribed and the resultant cDNA used in 5’RACE PCR reactions. 3’RACE uses an adaptor containing a poly-T sequence that specifically binds to the poly-A tail of the mRNA. The adaptor

sequences contain a recognition site for M-MLV reverse transcriptase. Newly synthesised cDNA was used in subsequent 3'RACE PCR reactions. 5' and 3' RLM-RACE PCR thermocycling profile:

1 cycle:	extended DNA denaturation at 94ºC – 5 minutes
30 cycles:	denaturation at 94ºC – 30 seconds
	annealing at 3-5ºC below average primer Tm – 30 seconds
	extension at 72ºC – 2-3 minutes
1 cycle:	extended extension at 72ºC – 10 minutes
Hold:	4ºC

RACE PCR annealing temperature as recommended in the kit ranges from 55-65ºC, therefore is flexible and is optimised according to the Tm of the gene-specific primer. On occasion, the Tm of a gene-specific primer is significantly different from that of the RACE primer and, therefore, a touchdown PCR was performed. In these instances, the thermocycling profile was altered: starting with the higher annealing temperature corresponding to the RACE-specific primer (60ºC for 5 cycles) and subsequently lowering the annealing temperature by 3ºC (5 cycles each) until reaching the optimum gene-specific primer Tm, which was then allowed to cycle 20-25 times. Visualisation of RACE PCR products was done via agarose gel electrophoresis as described in section G.

L. Riboprobe synthesis

Plasmids containing the sequence of a desired gene of interest were linearised with an appropriate restriction enzyme that cut the plasmid one time on either side of the insert, depending on the orientation, and use of either the T7 or T3 polymerase priming site. Plasmids containing a *Pa-nubbin* fragment were kindly provided by A. Popadić (Wayne

State University, Detroit, USA). RE digestions were performed for at least two hours at 37ºC. The linearised plasmid DNA was purified by phenol/chloroform extraction, ethanol precipitated, and resuspended in nuclease-free water. Concentration was checked by running both 0.2 µl and 1µl aliquots on a 1% agarose gel and comparing band intensity with the MassRuler^{TM} DNA Ladder Mix.

RNA riboprobes were synthesised, under RNase free conditions, using the RNA Labelling Kit (Roche – Cat#11175025910) following the manufacturer's protocol. One microgram of DNA was used as a template for transcription and RNA probes were generated by incorporating labelled UTPs [Digoxigenin (DIG; Roche – Cat#11175025910); Biotin (BIO; Roche - 11685597910); Dinitrophenol (DNP; Perkin Elmer – NEL55001EA] during this process. Transcription reactions were conducted for two hours at 37ºC, after which the cDNA template was removed via DNAse I treatment. RNA was precipitated overnight at -20ºC after the addition of 8M LiCl and 100% ethanol, then centrifuged to pellet, washed with 70% ethanol, and resuspended in 40 µl nuclease-free water. Details of probe length and mRNA hybridisation sequences for each gene are described within each chapter.

M. Hydrolysis of riboprobes

Newly synthesised riboprobes were hydrolysed into smaller fragments of 100-200 bp in order to improve signal in subsequent *in situ* hybridisation. Probe hydrolysis, as described by Lanfear (2007), consisted of adding equal volumes of 0.4 M sodium bicarbonate and 0.6 M sodium carbonate to the riboprobe and incubating at 60ºC for a calculated time: $t = \frac{(L_{start} - L_{end})}{0.11 \text{ x } L_{start} \text{ x } L_{end}}$, where L_{start} is starting length of the riboprobe (in kb) and L_{end} is the desired end length (in kb). The reaction was stopped by adding ammonium acetate, 3 M glacial acetic acid, and precipitated with 100% ethanol at -20ºC for 20

minutes. The mixture was then centrifuged and the resulting pellet washed with 70% ethanol, air-dried, and resuspended in 80-100 µl hybridisation solution (Section O).

N. Dot blot

Labelling efficiency and quantification of DIG-labelled riboprobes were done via dot blot. A dilution series (1:100; 1:1000; 1:10000 in nuclease-free water) was prepared for the gene specific riboprobe alongside a DIG-labelled control RNA provided in the Roche DIG-labelling kit. Two microliters of each dilution was blotted onto a positively charged nylon membrane (Roche – Cat#1417240) and fixed using the UV Stratalinker® (Stratagene). The membrane was rinsed twice for 5 minutes with Maleic acid buffer pH 7.5 (0.1 M maleic acid, 0.15 M NaCl) before blocking for 30 minutes in 10% Blocking Solution (0.5 g Blocking Reagent [Roche – Cat# 11096176001] in 50 ml maleic acid buffer). After blocking, the membrane was incubated for 30 minutes with the Alkaline Phosphatase conjugated anti-DIG antibodies, Fab fragments (α-DIG-AP; Roche – Cat#11093274910) at 1:10000 in Blocking Solution. Excess antibody was removed with two 15-minute washes using Wash Buffer (Maleic acid buffer + 0.3% Tween-20). Before detection, the membrane was equilibrated with two 5-minute washes in Detection Buffer (0.1 M Tris-HCl; 0.1 M NaCl; 50 mM $MgCl_2$; pH 9.5). Colour was developed in the dark in Detection Buffer containing NBT (Roche – Cat#11383213001) and BCIP (Roche – Cat#11383221001); when colour developed, the reaction was stopped via several washes with distilled water.

O. Colorimetric *in situ* hybridization

Colorimetric *in situ* hybridization protocols were slightly modified from that described by Marie et al. (2000) and Lanfear (2007). Embryos stored at -20ºC were

removed and allowed to warm up to room temperature, then rehydrated in a decreasing series of ethanol dilutions (90%, 70%, 50%, 30%), and washed several times with 1X PBT (1X PBS + 0.2% Tween-20) to remove excess ethanol. The embryos were then permeabilised using either of two methods. The first method used Proteinase K (Roche – Cat#3115828) digestion at 3 µg/ml in PBT for 1 hour on ice. The reaction was stopped by washing two times with 2 mg/ml glycine in PBT followed by several washes with 1X PBT to remove the glycine. The second treatment used 0.1% sodium borohydride (Acros Organics – Cat#AC41947-1000) in 1X PBT for 10 minutes, uncapped in a fume hood as this produces a large amount of hydrogen gas. The embryos were washed 3 times with PBT before moving on to the next step.

When a DIG-labelled probe was used, the embryos were incubated with 0.2 M HCl for 10 minutes to remove endogenous alkaline phosphatase activity. Embryos were washed several times in 1X PBT then post-fixed in 4% paraformaldehyde for 20 minutes. Embryos were prepared for hybridisation by first washing with a 1:1 solution of PBT:Hybridisation Solution (HS – a mixture of deionised formamide, 20X standard saline citrate, heparin (50 µg/ml), Tween-20, boiled salmon sperm (10 mg/ml), tRNA, and nuclease free H_2O) then washed with neat HS. Pre-hybridisation was performed in HS for two or more hours in a 56ºC water bath. 200-500 ng of hydrolysed probe in HS was heated to 90ºC for one minute to relax any secondary structures before adding to the embryos; the probe was allowed to hybridise at 56ºC overnight, for at least 16 hours.

The probe was removed and stored at -80ºC for later re-use. Embryos were washed several times in HS to remove excess probe and then slowly transferred to PBT in a series of HS:PBT solutions (1:3, 1:1, 3:1), finally being brought down to room temperature in 1X PBT. Embryos were blocked for 2 hours in a solution of 1% bovine serum albumin (Sigma) and 5% normal horse serum (Vector Labs) in 1X PBT. After blocking, embryos

were incubated in a slow rotator at 4ºC overnight in a 1:1000 dilution of α-DIG-AP (Roche) in blocking buffer. The antibody was removed by washing the embryos several times with PBT over 2 hours followed by three rinses with staining solution (5 M NaCl, 1 M Tris-HCl, 1 M $MgCl_2$, 0.2% Tween-20, and H_2O). A developing solution was made by combining staining solution with 4.5 µl/ml NBT and 3.5 µl/ml BCIP. Signal was allowed to develop in the dark and monitored periodically as to not overstain. Embryos were then washed several times in 1X PBS before mounting on glass slides in Aqua-Poly/Mount (Polysciences – Cat#18606-20) prior to imaging.

P. Fluorescence *in situ* hybridisation (FISH)

Embryos undergoing fluorescence *in situ* hybridisation were treated similarly as described above (section N) up to the post-hybridisation PBT washes. Embryos were then subjected to the Tyramide Signal Amplification (TSA) System (PerkinElmer Life Sciences) protocol. Here, the embryos were incubated in 3% hydrogen peroxide for 30 minutes in order to deactivate endogenous peroxidase activity so as not to interfere with the horseradish peroxidase (HRP) driven processes during the signal detection stage. The embryos were washed twice for 10 minutes with 1X PBT, then blocked for 2 hours with TNB Buffer (0.1 M Tris-HCl pH7.5; 15 M NaCl; 0.5% Blocking Reagent [TSA kit], sterile H_2O). For biotin-labelled probes, a 1:100 dilution of the streptavidin-HRP conjugated antibody (TSA kit) in TNB was added to the embryos, which were then incubated overnight at 4ºC. DIG-labelled probes were treated with HRP conjugated α-DIG-POD, Fab fragments (Roche – Cat#11207733910) at 1:100 dilution in TNB and similarly stored overnight at 4ºC. DIG-labelled probes used in double FISH experiments were incubated overnight at 4ºC with sheep α-DIG (Roche – Cat#11333089001) at 1:200

in TNB, then washed three times in TNB and incubated for 2 hours at room temperature with a donkey α-sheep Alexa555 antibody (Molecular Probes – A21436) at 1:200 in TNB.

In either case, embryos were washed 3 times with TNT (0.1 M Tris-HCl pH 7.5; 0.15 M NaCl; 0.05% Tween-20; sterile H_2O) before signal amplification and detection. TSA amplification was done by adding a 1:50 dilution of either Cyanine3 (red; PerkinElmer – Cat#NEL704A001KT) or Fluorescein (green; PerkinElmer – Cat#NEL701A001KT) in 1X Amplification Diluent (all components provided in the TSA kits; PerkinElmer). Embryos were placed in a slow moving rotator for 10-30 minutes and periodically checked for optimal fluorescent signal. Reactions were stopped by washing two times in TNT followed by several washes with PBS. Embryos were mounted on glass slides in Aqua-Poly/Mount (Polysciences) prior to imaging.

Q. Immunocytochemistry

Protein expression was detected using various cross-reactive antibodies following protocols described by Pueyo et al. (2008). Primary monoclonal mouse antibody FP6.87 (α-Ubx/α-AbdA; from Rob White, University of Cambridge, Cambridge, UK) was used at a 1:10 dilution. Mouse α-Eve monoclonal antibody 2B8 from the Developmental Studies Hybridoma Bank at the University of Iowa as described by Patel et al. (1992; 1994) and executed by JP Couso (University of Sussex). Apoptosis was detected using a rabbit α-cleaved caspase 3 antibody (cas3; Cell Signalling Technology) at a 1:50 dilution and cells undergoing mitosis were detected using a polyclonal rabbit α-phosphorylated Histone 3 antibody (H3P; Upstate) at 1:1000 dilution. Nuclei were detected using DAPI (4’,6-diamidino-2-phenylindole; Invitrogen). Secondary antibodies conjugated with biotin or fluorophores were from Jackson Immunochemicals.

Fixed embryos were removed from -20ºC and allowed to warm to room temperature, then washed several times with PBTx (1X PBS, 0.3% Triton X-100, 1% Bovine Serum Albumin). Embryos were incubated with the primary antibody overnight at 4ºC. The following day, embryos were washed two times with PBTx prior to addition of the secondary antibody and incubated at room temperature for 2 hours. A single embryo was removed from the tube periodically in order to examine strength of signal; at optimal staining embryos were washed several times over 1 hour with PBTx before mounting on glass slides in Aqua-Poly/Mount (Polysciences). Images taken as described in Section U.

R. Double stranded RNA synthesis

Gene-specific forward and reverse primers were designed containing the T7 polymerase promoter sequence (TAATACGACTCACTATAGGGA) at the 5' end (Appendix 2); this method allows for dsRNA production directly from the PCR products. A standard PCR reaction was prepared as described in section F, using plasmid clones as the template. PCR products were run on a 0.8% agarose gel and the band of the expected sized was excised on a UV transilluminator (Syngene) with a sterile razor, then extracted and cleaned up using the QIAquick Gel Extraction Kit (QIAGEN – Cat#28706), following the protocol provided by the manufacturer.

The purified DNA fragment was used to synthesise dsRNA with the T7 RiboMAX^{TM} Express Large Scale RNA Production System (Promega – Cat#P1320), according to the manufacturer's instructions. One microgram of cDNA from the PCR reaction above was used per transcription reaction, which was allowed to run at 37ºC for one hour. DNA template was removed via treatment with DNase I for 15 minutes at 37ºC and RNA was extracted with phenol:chloroform (5:1), then chloroform:isoamyl (24:1),

followed by isopropanol precipitation at -20ºC for 2 hours. RNA precipitate was spun to pellet, washed with 70% ethanol, air-dried, and resuspended in nuclease-free H_2O.

S. RNA interference (RNAi)

Periplaneta maternal RNAi was conducted as described by Pueyo et al. (2008). dsRNA, of varying concentration by gene (Table 2.1), was injected into the ventrolateral abdomen of adult virgin females using a 5 ml BD Plastipak syringe (Becton Dickinson) modified to fit pulled glass capillary needles. Prior to injection females were anaesthetized with CO_2. Females were subjected to two injections, first on the left side then on the right side of the abdomen, spaced three hours apart. Post-injection, up to six similarly injected females were placed in isolated containers with 3 males and kept in a 29ºC incubator; oothecae were collected and incubated as described in section A. As a control, several adult females were injected with 10 µl H_2O. Phenotypic analysis of resultant embryos and first nymphs are described within each chapter.

T. RT-PCR

Total RNA isolation and cDNA synthesis from stage 9 wild type, Class 'T' *Pa-Wnt1*^{RNAi}, and Class 'T' *Pa-cad*^{RNAi} embryos were conducted as described in sections C and D above. Equal concentrations of cDNA were used in subsequent RT-PCR reactions using specific primers for *Pa-Wnt* or *Pa-cad* to determine the amount of each in *Pa-cad*^{RNAi} or *Pa-Wnt*^{RNAi} compared to wild type. Specific primers were designed towards the *Periplaneta* 18S ribosomal subunit (Appendix 2I), used as a positive control. PCR reactions were run on a 1% agarose gel as describe in section G.

dsRNA	Concentration	Amount injected	Total injected
Pa-caudal	2.0 µg/µl	2 x 4 µl	16 µg
Pa-Wnt1	0.5 µg/µl	2 x 4 µl	4 µg
Pa-even-skipped	2.0 µg/µl	2 x 5 µl	20 µg
Pa-runt	2.0 µg/µl	2 x 5 µl	20 µg
Pa-pairberry	2.0 µg/µl	2 x 4 µl	16 µg
Pa-sloppy-paired	2.0 µg/µl	2 x 5 µl	20 µg
Pa-tarsal-less	4.0 µg/µl	2 x 4 µl	32 µg
Pa-nubbin	3.0 µg/µl	2 x 5 µl	30 µg
Pa-Notch	0.5 µg/µl	2 x 4 µl	4 µg
Pa-eve/run	2.5 µg/µl each	2 x 2.5 µl each	12.5 µg each
Pa-tal/nub	2.0 µg/µl each	2 x 2.5 µl each	12.5 µg each
H_2O (negative)	---	2 x 5 µl	---

Table 2.1: dsRNA concentration used in RNA interference. Stock solutions of each dsRNA were diluted in water to a final concentration of 0.5-4.0 µg/µl. Two injections of equal volume (2.5-5 µl) were administered at three hour intervals. Sterile water was injected as a negative control.

U. *Periplaneta* embryo culturing and inhibitor treatment

ex ovo cultures of *Periplaneta* embryos were performed according to Wang et al. (1992) and modified by Pueyo et al. (2008). Artificial culture medium was prepared in 100 ml batches as follows: 98 ml Liebovitz L-15 medium (Sigma – Cat#L1518), 6 µl hydroxyecdysterone (2.5 mg/ml; Sigma), 80 µl bovine insulin (10 mg/ml; Sigma), 1 ml penicillin (50000 units/ml; Sigma), 1ml streptomycin (50 mg/ml; Sigma), 400 mg glucose. Embryos were dissected in pre-warmed culture medium and transferred to a sterile low-adhesion 96-well tissue culture plate (Corning Incorporated) containing 200 µl pre-warmed culture medium and one of three treatments: the Notch inhibitor DAPT, the Wnt inhibitor IWP-3, or DMSO as a control.

Inhibition of N-signalling was carried out through the addition of DAPT (N-[N-(3,5-Difluorophenacetyl-L-alanyl)]-S-phenylglycine t-butyl ester; Calbiochem – Cat#565770), which inhibits γ-secretase cleavage of the Notch intracellular domain (Dovey et al., 2001). Embryos were cultured for 24 hours at 29ºC in 100 μM DAPT following the protocol described in Pueyo et al. (2008). Inhibition of Wnt-signalling was conducted by adding the Inhibitor of Wnt Production – compound 3 (IWP-3; 2-(3-(4-fluorophenyl)-3,4,6,7-tetrahydro-4-oxothieno[3,2-d]pyrimidin-2-ylthio)-N-(6-methyl benzo[d] thiazol-2-yl)acetamide; Stemgent – Cat#04-0035), which blocks the secretion of Wnt ligands by inhibiting palmitoylation of Wnt by Porcupine (Chen et al., 2009). Upon receipt, the compound was diluted in DMSO (dimethyl sulfoxide; Sigma – Cat#D8418) to a stock concentration of 10 mM. Various working dilutions were added to the embryo cultures to a final concentration between 20 μM and 40 μM. Exposure response analysis showed various results at different concentrations (see Fig. 3.7): < 20 μM had no effect; 20-30 μM had hypomorph/weak effects; 30-40 μM had strong effects; > 40 μM produced cytotoxic effects. Exposure to IWP-3 for times longer than 18-24 hours also produced cytotoxic effects.

In all cases, embryos from the same ootheca were split into three groups. First, several embryos were fixed immediately to report the developmental stage before culture (0 hours control). The remaining embryos were then split into two groups and cultured for 16-24 hours at 29ºC either with or without inhibitor (culture control). Control cultured embryos were treated with the same amount of DMSO as the respective inhibitor used. Post-culture, embryos were fixed with 4% paraformaldehyde and stored at -20ºC until ready to use for *in situ* hybridisation.

V. Cuticle preparation

Technique modified from the cuticle preparation protocol for flies as described by Alexandre (2008). Hatched first nymphs or late stage embryos (stages 30-32) were dissected and cleaned of debris in 1X PBS, then rinsed twice with sterile H_2O. The embryos or nymphs were then transferred to a 1M NaOH solution and incubated for 7 minutes on a 60ºC hotplate, rinsed 3 times with H_2O, then incubated at 60ºC for 20-40 minutes in a 1:1 solution of Hoyer's:Lactic Acid (Hoyer's mountant/fixative: 30 g gum arabic, 16 ml glycerol, 200 g chloral hydrate, 50 ml distilled H_2O). Dissected first nymph abdomens and embryo cuticles were mounted on a glass slide with 1:1 Hoyer's:Lactic Acid and a coverslip and flattened with a weight on top at 60ºC for several hours, then cooled to room temperature allowing the mountant to solidify.

W. Imaging

Unstained or non-fluorescently stained mounted embryos were analysed and imaged using a Leica DMRB research microscope with a mounted Hamamatsu digital camera and SimplePCI 6 software. Images of whole, unmounted late-stage embryos and first nymphs were taken using a Leica MZ7.5 stereomicroscope with a Leica DFC420 C digital camera and Leica FireCam software (kindly offered for use by the Alonso group; University of Sussex). FISH and fluorescent antibody stained embryos were analyzed and images captured using a Zeiss LSM510 Meta point scanning confocal microscope and the Zeiss LSM software (Sussex Centre for Advanced Microscopy, University of Sussex). Image adjustments and figure preparations were done using Adobe Photoshop CS2 software.

CHAPTER III

Interplay between a Wnt-dependent organiser and the Notch segmentation clock regulates posterior development in *Periplaneta americana*

ABSTRACT

Sequential addition of segments in the posteriorly growing end of the embryo is a developmental mechanism common to many bilaterians. However, posterior growth and patterning in most animals also entails the establishment of a 'posterior organiser' that expresses the Caudal and Wnt proteins and has been proposed to be an ancestral feature of animal development. We have studied the functional relationships between the Wnt-driven organiser and the segmentation mechanisms in a basal insect, the cockroach *Periplaneta americana*. Here, posteriorly-expressed *Wnt1* promotes *caudal* and *Delta* expression early in development to generate a growth zone from which segments will later bud off. *caudal* maintains the undifferentiated growth zone by dampening *Delta* expression and hence Notch-mediated segmentation occurs just outside the *caudal* domain. In turn, *Delta* expression maintains *Wnt1* creating a posterior gene network that functions until all segments have formed. This feedback between *caudal*, Wnt and Notch-signalling in regulating growth and segmentation seems conserved in other arthropods, with some aspects found even in vertebrates. These findings not only support an ancestral Wnt posterior organiser, but also impinge on the proposals for a common origin of segmentation in arthropods, annelids, and vertebrates.

N.B. This chapter has been accepted for publication in Biology Open: Chesebro JE, Pueyo JI, Couso JP. (2012). Interplay between a Wnt-dependent organiser and the Notch segmentation clock regulates posterior development in *Periplaneta americana*. doi: 10.1242/bio.20123699

N.B. Some experiments discussed in this chapter were conducted by Inyaki Pueyo, including: sequencing of *Pa-Wnt1* (with Alex Hurst), cockroach embryo cultures, caspase-3 and Histone-3-P immunocytochemistry.

INTRODUCTION

Animals display a great variety of morphologies and body plans. Despite this diversity, a common form of development entails the early setting up of a small anteriorized embryo which then grows from its posterior end until the final body length is achieved (Jacobs et al., 2005). This is the mode of development for most complex animals with bilateral symmetry, except atypical ones, such as the indirect developing Echinoderms and Tunicates (Holland, 2002; Mooi and David, 2008). In many animals that display posterior growth, this occurs through the sequential addition of metameric, serially repeated, units called segments or somites (Balavoine and Adoutte, 2003; Couso, 2009; De Robertis, 1997). This is the case of vertebrates, annelids, and most arthropods.

To implement this posterior growth mode of development, the developing embryo must establish a 'posterior organiser' that drives development and growth from the posterior end of the embryo. In vertebrates this involves the early expression of genes of the Wnt family (Shimizu et al., 2005; Takada et al., 1994). In fact, Wnts are found to be expressed in the posterior of all bilaterians studied to date and have even been found to pattern the posterior axis in non-bilaterian Porifera larvae (Adamska et al., 2007; Cho et al., 2010; Janssen et al., 2010; Martin and Kimelman, 2009; Niehrs, 2010; Riddiford and Olson, 2011; Ryan and Baxevanis, 2007). Another gene required for proper posterior growth is *Cdx*, the vertebrate homologue of the homeobox transcription factor *caudal* (*cad*). It has been determined that *Cdx* expression is controlled by Wnt and, indeed, that the effects of posterior Wnt-signalling are mediated through Cdx (Lohnes, 2003; Shimizu et al., 2005; van de Ven et al., 2011). Thus, the use of a Wnt-cad posterior organiser may be a common mechanism utilized by animals in which posterior growth occurs through the sequential addition of segments (Martin and Kimelman, 2009).

In arthropods this mode of development is referred to as short germ band and is the most common, and inferred ancestral, developmental mechanism of the group (Anderson, 1972; Davis and Patel, 2002; Liu and Kaufman, 2005b; Peel and Akam, 2003). In short germ arthropods most segments are subsequently added during germ band elongation from an undifferentiated region of proliferating cells at the posterior end of the embryo called the growth zone (GZ – analogous to the presomitic mesoderm of vertebrates) (Davis and Patel, 2002; Liu and Kaufman, 2005b; Peel and Akam, 2003). Comparison of different arthropods has revealed a conserved set of genes involved in early posterior development consisting of *caudal* and the Wnt-signalling pathway. Similar to vertebrates (Chawengsaksophak et al., 2004; van de Ven et al., 2011; van den Akker et al., 2002), knock-down of either *cad* or *Wnt* in arthropods results in posterior truncations (Bolognesi et al., 2008; Copf et al., 2004; Martin and Kimelman, 2009; McGregor et al., 2009; McGregor et al., 2008; Shimizu et al., 2005; Shinmyo et al., 2005; Takada et al., 1994), highlighting their importance in establishing a posterior organiser driving development and growth.

The basal, short germ band insect, *Periplaneta americana* (American cockroach), was used to gain insight about the interaction between this posterior organiser and the segmentation mechanisms. Sequential segmentation in *Periplaneta* involves cyclic waves of *Delta* (*Dl*) and *hairy/hes* (*h*) expression that emanate from the posterior and resolve into segmental stripes in the anterior GZ prior to segment formation (Pueyo et al., 2008). Loss of Notch (N) signalling via RNAi resulted in embryos in which the posterior is truncated and unsegmented (Pueyo et al., 2008). This N-mediated segmentation is reminiscent of the ‘clock and wavefront’ mechanism found in vertebrate somitogenesis (Dequeant and Pourquie, 2008; Jiang et al., 2000). Dynamic expression of *Dl/N* has been found in spiders and determined to be important for proper segmentation and been proposed to be

ancestral, though whether these patterns are also oscillatory in the spider remain to be determined (Oda et al., 2007; Stollewerk et al., 2003). A number of studies have indicated that an interplay between Wnt, Cad and N signalling during somitogenesis in vertebrates may exist (Aulehla and Herrmann, 2004; Grainger et al., 2012; Savory et al., 2011; Young et al., 2009); however, the nature of such a connection has not yet been established definitively.

Discussed in this chapter, *Periplaneta cad* and *Wnt1* are expressed in the posterior GZ and disrupting their functions using RNAi or chemical inhibitors produced embryos with segmentation defects and revealed two distinct phases of *Pa-cad* and *Pa-Wnt1* function. First, early in development, *Pa-Wnt1* is required for *Pa-cad* expression and together they establish a posterior organiser and a functional GZ. Second, during germ band elongation, *Pa-Wnt1* regulates axial growth and posterior segmentation by activating *Pa-Delta* and *Pa-cad* expression in the GZ. Subsequently, *Pa-cad* maintains the GZ in an unsegmented and proliferative state through which the dynamic waves of *Delta* (Pueyo et al., 2008) are allowed to progress in order to form segments outside the *Pa-cad* domain. Reciprocally, Dl-N signalling in the posterior tip is necessary to maintain posterior *Pa-Wnt1* expression. Thus, the interplay between Cad, Wnt, and N signalling pathways regulates posterior growth, elongation, and segmentation in *Periplaneta*. This two-step model can explain results in other arthropods and shows conserved features in vertebrates, suggesting that Wnt and Notch-signalling form an ancestral gene network controlling posterior growth and segmentation.

RESULTS

Isolation and patterns of expression of *Pa-caudal* and *Pa-Wnt1* transcripts

The full-length transcripts of *Pa-cad* and *Pa-Wnt1* were cloned using a degenerate RT-PCR approach (Appendix 2A and 2B, respectively). A single *Pa-cad* transcript of 1420 nucleotides that encodes for a 290aa protein was identified (Appendix 3A). Phylogenetic analysis indicates that *Pa*-Cad is the *Periplaneta* orthologue as it aligns closely to Caudal from related species, such as *Gryllus bimaculatus* (cricket) and *Schistocerca gregaria* (locust) (Appendix 3B). Likewise, a single 2364bp *Pa-Wnt1* transcript was isolated that encodes for a 373aa protein (Appendix 4A-B). According to the alignment with other arthropod Wnt proteins, *Pa-Wnt1* is the *Periplaneta* orthologue closely related to *Cryptotermes* (termite; Blattodea) and the orthopterans *Schistocerca* (locust) and *Gryllus* (cricket), which follows the predicted insect phylogeny (Appendix 4B).

in situ hybridisation using a *Pa-cad* riboprobe shows that at early, post-blastoderm stages of embryogenesis, *Pa-cad* is strongly expressed at the posterior end of the embryo (Fig. 3.1A). This broad domain of expression remains at the onset of germ band elongation (Fig. 3.1B), but is cleared from the posterior tip at late germ band elongation (Fig. 3.1C). This expression pattern of *Pa-cad* is similar to that observed in other short germ band insects (Copf et al., 2004; Dearden and Akam, 2001; Shinmyo et al., 2005). *Pa-Wnt1* transcripts are detected at the post-blastoderm stage in the posterior GZ as two symmetrical clusters of cells (Fig. 3.1D). In addition, anterior *Pa-Wnt1* expression is observed in the head lobes and in the antennal primordia (Fig. 3.1D). During early germ band elongation, anterior stripes of *Pa-Wnt1* expression become apparent in the presumptive gnathal and thoracic segments while the two posterior clusters of expression start to fuse (Fig. 3.1E). By late germ band elongation, the two *Pa-Wnt1* clusters of cells

have joined together forming a wide arc of expression in the posterior end of the GZ, set apart from the posterior tip (Fig. 3.1F). New stripes of *Pa-Wnt1* expression appear one by one in the anterior-most GZ and remain in the developing segments and ventral appendages throughout embryogenesis (Fig. 3.1F). The early, dynamic posterior expression of *Pa-Wnt1* lies close to the expression of *Pa-cad*, and thus would be consistent with a conserved role for these two genes during posterior patterning. The later *Pa-Wnt1* segmental stripe pattern is consistent with *Wnt1* expression in other insects and arthropods in which its function in parasegmental boundary formation, segmentation, and appendage development has been established (Bolognesi et al., 2008; Couso et al., 1993; Grossmann et al., 2009; Martinez-Arias, 1993; Miyawaki et al., 2004).

In order to establish precisely the spatial and temporal details of *Pa-Wnt1* and *Pa-cad* expression, double fluorescence *in situ* hybridization (FISH) was employed, as well as comparison with the pan-segmental marker *engrailed* (*en*) (Patel et al., 1989). *Pa-Wnt1* slightly overlaps and is posterior to *Pa-cad* expression during both post-blastoderm (Fig. 3.2A-A") and germ band elongation (Fig. 3.2B-B"). In comparison, post-blastoderm expression of *Pa-en* (Marie and Bacon, 2000) appears as several stripes from the head segments up to the T2 segment (Fig. 3.1G). During early (Fig. 3.1H) and late (Fig. 3.1I) germ band elongation, additional stripes of *Pa-en* appear sequentially in the anterior GZ, slightly before the corresponding *Pa-Wnt1* segmental stripe expression (compare Fig. 3.1E-F with 3.1H-I). Double FISH experiments confirm the temporal delay of segmental *Pa-Wnt1* expression after *Pa-en* (Fig. 3.2C-C"), as well as their adjacent expression at the parasegmental boundary as *Pa-Wnt1* abuts anterior to *Pa-en* expression (Farzana and Brown, 2008; Martinez Arias and Lawrence, 1985; Prud'homme et al., 2003).

Pa-Cad and *Pa-Wnt1* act as a posterior organiser

Maternal RNAi was used to test the possible conservation of Wnt and *cad* function. In *Pa-Wnt1*^{RNAi} embryos, *Pa-cad* expression is greatly reduced at post-blastoderm (Fig. 3.3B; compare with wild type, Fig. 3.3A). By contrast, posterior *Pa-Wnt1* expression in *Pa-cad*^{RNAi} embryos at this stage was similar to wild type, despite a premature fusion of the two *Pa-Wnt1* expressing clusters (Fig. 3.3D, compare with Fig. 3.3C). *Pa-Wnt1*^{RNAi} and *Pa-cad*^{RNAi} embryos developing up to germ band elongation display a truncated or tapering posterior end with growth and development of new segments arrested at the thoracic region (Fig. 3.3F and H). To ascertain the cellular basis of this reduced and abnormal GZ in *Pa-Wnt1*^{RNAi} embryos, cell death and proliferation were examined (Fig. 3.4). *Pa-Wnt1*^{RNAi} embryos do not show an increase in apoptosis in the GZ (Fig. 3.4A-B, I), but there is a reduction in cell proliferation (Fig. 3.4C-D, I). In these *Pa-Wnt1*^{RNAi} germ band elongation embryos, expression of *Pa-cad* appears completely lost according to both *in situ* hybridisation (compare Figs. 3.3F and 3.3E) and RT-PCR (Fig. 3.5D). However, in the reciprocal experiment, the number of cells expressing *Pa-Wnt1* in *Pa-cad*^{RNAi} appears similar to that of the wild type, albeit the region displays an abnormal shape (compare Figs. 3.3H and 3.3G). Accordingly, RT-PCR shows normal levels of *Pa-Wnt1* transcript in *Pa-cad*^{RNAi} compared to wild type (Fig. 3.5D). These results suggest that *Pa-Wnt1* regulates *Pa-cad* expression and that the apparent alteration of the *Pa-Wnt1* expression domain in *Pa-cad*^{RNAi} is a secondary consequence of abnormal posterior tip development (Fig. 3.3D; 3.3H).

A detailed quantification of the phenotypes observed indicates a graded variability within a conserved requirement for posterior development of the embryo. *Pa-cad*^{RNAi} embryos were classified according to the strength of the phenotype into three different classes (Table 3.1). In Class ‘H’ embryos, only the head segments form properly (Fig.

3.5A) and often a less defined region of undifferentiated tissue forms just posterior to these structures, representing a complete or near complete loss of *Pa-cad* expression. In Class 'H' embryos, no posterior segmental addition is observed (Fig. 3.5B). In Class 'T' embryos, the head and some of the thoracic segments developed normally but with truncations after either T2 or T3 (Fig. 3.3J and Fig. 3.6E), suggesting that some *Pa-cad* expression may have remained in the early post-blastoderm, but later expression was lost, leading to the eventual development of segments only up to T2/T3. This would explain the residual expression of *Pa-cad* detected in RT-PCR using Class 'T' *Pa-cad*^{RNAi} embryos (Fig. 3.5D). Finally, Class 'A' embryos display a reduced GZ and abnormal and thin attempts at abdominal segment development (Fig. 3.5B and Fig. 3.6H).

Pa-Wnt1^{RNAi} embryos displayed a similar range of posterior segmentation phenotypes (Table 3.1). *Pa-Wnt1*^{RNAi} embryos in Class 'T' show severe truncations ranging from the second thoracic segment to the first abdominal segment (Fig. 3.3K). In these embryos, the GZ is narrow and posterior segmentation proceeds abnormally (Fig. 3.6F). Class 'A' *Pa-Wnt1*^{RNAi} embryos displayed moderate truncations involving only a few abdominal segments and only a slight reduction in the GZ (Fig. 3.5C and Fig. 3.6I). The absence of Class 'H' embryos, which correlate with a total loss of *Pa-cad*, in *Pa-Wnt1*^{RNAi} may be due to several reasons. First, *Pa-Wnt1*^{RNAi} may not create a total null condition at stages 4-5 (post-blastoderm) or, second, there could be other Wnt genes acting at these early stages. Third, there could be some Wnt-independent, maternally deposited, *Pa-cad* expression. The absence of *Pa-Wnt1* in *Pa-Wnt1*^{RNAi} RT-PCR (Fig. 3.5D) may discard the first explanation, but in principle either of the three is compatible with both the residual *Pa-cad* expression observed in *Pa-Wnt1*^{RNAi} post-blastoderm embryos (Fig. 3.3B), but not later (Fig. 3.3F), and the similarity in phenotypes of Class 'T' embryos between the two RNAi treatments (Fig. 3.3J and 3.3K).

Specificity of Wnt-Cad posterior organiser phenotypes

Altogether these results are compatible with the model where Wnt-signalling activates zygotic *cad* expression during the early stages of development, establishing a posterior organiser (McGregor et al., 2008; Shinmyo et al., 2005). This *Periplaneta* 'posterior organiser' seems required for proper patterning of the posterior end of the embryo, which includes the GZ, and would be required for subsequent growth and segmentation. However, these RNAi results do not clarify whether the activity of this Wnt-Cad organiser is still required during germ band elongation, or whether the phenotypes observed arise secondarily from earlier defects – perhaps a failure to properly establish the GZ.

In order to clarify this issue, the embryo culture technique developed previously in our laboratory for application of chemical inhibitors of specific protein activity was used (Pueyo et al., 2008). Here, germ band elongating embryos were cultured and exposed to a previously described inhibitor of Wnt protein secretion, IWP-3 (Chen et al., 2009), to decouple the function of Wnt-signalling during germ band elongation from earlier or indirect perturbations. Embryos exposed to IWP-3 showed a marked decrease in *Pa-cad* expression and develop a slightly reduced GZ (Fig. 3.3N, compared to 3.3L-M) and disrupted segmentation (Fig. 3.7), while the expression of other genes, such as *Ubx* and *AbdA*, remain unaffected (Fig. 3.4E-F). Compared to control cultured embryos (Fig. 3.4E,G), a slightly increased rate of apoptosis (Fig. 3.4F,J), but a decrease in cell proliferation (Fig. 3.4H,K), is observed in embryos cultured with IWP-3. Staining with the vital dye DAPI confirms that the overall majority of cells are alive throughout the posterior embryo (Fig. 3.4G-H). Altogether, these effects are in accordance with *Pa-Wnt1*^{RNAi} data where Wnt-signalling seems to be required mainly for the maintenance of cell proliferation in the GZ (Fig. 3.4C-D, I-K). In addition, those cells detected as dying or

proliferating are scattered throughout the embryo and not localised to the posterior GZ such that these changes in growth and death cannot account for the proportional reduction in *Pa-cad* expression in this area. These data confirm that a) Wnt-signalling is required for *Pa-cad* expression, and b) this requirement extends into the germ band elongation stages.

Regulatory interactions between the posterior organiser and the Notch pathway

As dynamic N-signalling is necessary for posterior segmentation in *Periplaneta* (Pueyo et al., 2008), potential interactions between *Pa-cad* and *Pa-Wnt1* with elements of the N-signalling pathway were investigated. First, their patterns of expression in the GZ were compared during post-blastoderm (Fig. 3.8A-B) and germ band elongation (Fig. 3.8C-F). At post-blastoderm, five stripes of *Pa-Dl* expression appear simultaneously in the anterior part of the embryo, yet its expression is noticeably absent from the posterior end (Fig. 3.8A) where *Pa-Wnt1* and *Pa-cad* are expressed (Fig. 3.8B). These early *Pa-Dl* stripes have faded by germ band elongation at stage 6/7 when a new *Pa-Dl* pattern emerges at the posterior to include the new sequentially arising segmental stripes described in Pueyo et al. (2008). This germ band elongation pattern is composed of a small domain in the posterior tip from which cyclic waves emanate periodically (Pueyo et al., 2008) and the resulting 2-3 stripes in the anterior GZ which, interestingly, only form outside of the *Pa-cad* expression domain (Fig. 3.8C-C”). As previously described by Pueyo et al. (2008), these stripes of *Pa-Dl* just outside the mid-GZ regulate the striped expression of *Pa-en* in the anterior GZ (Figs. 3.8E-E”), eventually leading to segment border formation.

When these patterns of *Pa-Dl* and *Pa-en* are overlaid with the patterns of *Pa-cad* and *Pa-Wnt1* described above (Figs. 3.8C’,D), it appears that the posterior expression of these genes during germ band elongation divides the GZ into four distinct gene expression

regions (Fig. 3.8F), from posterior to anterior: 1) the posterior tip expressing *Pa-Dl*; 2) a wide posterior arc of *Pa-Wnt1*; 3) a broad domain of *Pa-cad* in the mid-GZ; and 4) an anterior pre-segmental region expressing first *Pa-Dl*, then *Pa-en*, and finally *Pa-Wnt1*. Based on these spatiotemporal expression patterns it appears that as the early post-blastoderm embryo develops into an elongating germ band where segments are added sequentially at the GZ, the most significant change is the addition of *Pa-Dl*, and hence Notch-signalling, to the posterior gene network.

To understand the regulatory relationships between the genes of this posterior network, their patterns of expression were compared in different maternal RNAi conditions. *Pa-N*^{RNAi} was used instead of *Pa-Dl*^{RNAi}, which has fewer effects on oogenesis, as the latter produces sterile females (Lanfear, 2007). *Pa-N*^{RNAi} leads to a loss of *Pa-Dl* expression in the posterior tip and a failure to form stripes in the anterior GZ (Fig. 3.9A-B), corresponding to the loss of sequential segment addition (Pueyo et al., 2008). A similar effect is observed in embryos cultured in the Notch pathway inhibitor DAPT, compared to controls (Fig. 3.9C-D). In both cases, neurogenic phenotypes are also observed, as expected (Fig. 3.4A'-D').

In post-blastoderm *Pa-Wnt1*^{RNAi} embryos the early segmental stripes of *Pa-Dl* expression remain (Fig. 3.10B, compare with Fig. 3.10A). Conversely, *Pa-N*^{RNAi} had no effect on the activity of the Wnt-Cad organiser at this stage, as indicated by normal *Pa-cad* expression (Fig. 3.10E-F). However, the situation changes during germ band elongation. *Pa-Wnt1*^{RNAi} embryos at this stage show an absence of *Pa-Dl* expression in the posterior tip as well as a loss of the stripes in the anterior GZ even though the neural *Pa-Dl* expression remains (Fig. 3.10C-D). Similarly, embryos cultured with the Wnt-signalling inhibitor, IWP-3, during germ band elongation experienced an almost total loss of *Pa-Dl* in the posterior tip (Fig. 3.10K). Interestingly, the *Pa-Dl* segmental stripes do not

disappear in IWP-3 embryos; however, only one new stripe can form in the anterior GZ (Fig. 3.10K) compared to two stripes that form in the control culture (Fig. 3.10J), indicating a temporal component to the stronger phenotypes revealed previously by *Pa-Wnt1*^{RNAi}.

The posterior tip expression of *Pa-Dl*, which lies adjacent to the posterior arc of *Pa-Wnt1* expression, seems to have an immediate requirement for *Pa*-Wnt1 signalling. The anterior stripes of *Pa-Dl*, which at least initially do not lie near any source of *Pa-Wnt1* (whether the posterior arc or the segmental stripes), seem to have a delayed, most likely indirect, requirement. The simplest explanation is that the loss of *Pa-Dl* segmental stripes in the anterior GZ follows the loss of the posterior tip expression. In wild type, the segmental stripes of *Pa-Dl* are transient, each one arising from a wave in the GZ and disappearing as it moves towards the anterior (Pueyo et al., 2008). Hence, the loss of the posterior tip domain, from which the waves of *Pa-Dl* expression emanate, will lead in time to the absence of segmental stripes. This secondary loss can be seen after 16 hours of development in IWP-3 culture where only one additional stripe forms in the anterior GZ (Fig. 3.10K) compared to the two stripes of *Pa-Dl* that form in control culture (Fig. 3.10J), and is more completely observed in *Pa-Wnt1*^{RNAi} embryos left to develop for several days after blastoderm (Fig. 3.10D).

Interestingly, this function of *Pa*-Wnt1 in activating *Pa-Dl* posterior tip expression does not seem to be mediated by *Pa-cad*. In *Pa-cad*^{RNAi} embryos at germ band elongation, the tip expression of *Pa-Dl* remains unaffected, but the anterior stripes do not form (Fig. 3.10H). Instead, there is a strong, expanded expression of *Pa-Dl* covering the posterior region of the embryo, in what corresponds to the entire GZ region of a wild type embryo. The lack of effect of Pa-Cad on *Pa-Dl* posterior tip expression is not unexpected, as these expression domains do not overlap. However, the revealed repressory function of *Pa-cad*

on *Pa-Dl* expression in the mid-GZ matches the wild type expression patterns, since previous observations showed that the strong segmental stripes of *Pa-Dl* only form outside the *Pa-cad* domain (Fig. 3.8C). These results imply that *Pa-Wnt1* signalling modulates the cyclic expression of *Pa-Dl* during germ band elongation in two opposite ways: it locally promotes *Pa-Dl* in the nearby posterior tip, while the anterior stripes of *Pa-Dl* seem to have a delayed, indirect requirement for *Pa*-Wnt1, regulated through *Pa-cad*. The combined effect of these two inputs is that the posterior tip source from which the cyclic waves of *Pa-Dl* originate is maintained, but the waves are only allowed to coalesce into a stripe at the anterior GZ. This model explains why the loss of *Pa-cad* expression in *Pa-Wnt1*^{RNAi} does not result in the expansion of *Pa-Dl* expression, as in *Pa-Wnt1*^{RNAi} there is also no emanating source of *Pa-Dl* expression from the posterior tip.

Unexpectedly, functional assessment of Notch-signalling indicates that this modulating effect of the posterior Wnt-Cad organiser on *Pa-Dl* expression is reciprocated. *Pa-N*^{RNAi} embryos at germ band elongation have a reduced and tapered GZ that does not express either *Pa-cad* (Fig. 3.11B) or *Pa-Wnt1* (Fig. 3.11D), implying that *Pa-Dl* may be required to maintain the posterior expression of *Pa-Wnt1*, and hence, *Pa-cad*. This inference is corroborated by embryo culture experiments. When embryos were cultured with the Notch inhibitor DAPT there was a noticeable reduction in *Pa-cad* expression (Fig. 3.11G) and the arc of *Pa-Wnt1* is reduced from a 4-cell wide band to a 2-cell wide narrow strip of expression (Fig. 3.11J). These results confirm the role of *Pa-Dl* in maintaining *Pa-Wnt1*, and *Pa-cad*, expression in a feedback mechanism in the posterior GZ. The reduction, but not loss, of both *Pa-Wnt1* and *Pa-cad* expression can be explained by the incomplete block of Notch activity in these experiments. For example, the segmentation process is perturbed but not entirely blocked, as shown by new but incomplete segmental stripes of *Pa-Wnt1* (Fig. 3.11J) and *Pa-en* (Pueyo et al., 2008)

appearing during DAPT culture. Importantly, this hypomorphy allows a normal-sized GZ to remain and consequently shows that the gene expression effects observed here and in *Pa-N*^{RNAi} experiments are not purely a secondary consequence of the lack of expressing cells; see also (Pueyo et al., 2008).

DISCUSSION

A Wnt-dependent posterior organiser is present in *Periplaneta americana*

The results presented above reveal the presence of a posterior organiser regulating posterior development and segmentation in the short germ band insect, *Periplaneta americana*. Shortly after cellularisation and before the sequential addition of segments takes place, *Pa-Wnt1* is observed in two clusters of cells in the posterior-most part of the embryo and *Pa-cad* is strongly detected in nearby cells covering what will become the GZ (Fig. 3.12A). The data presented here suggest that *Pa-Wnt1* expression at this stage activates or maintains *Pa-cad*, which in turn is essential for the proper patterning of the embryo. Since this patterning includes the establishment of a GZ from which most of the segments will form during germ band elongation, perturbation of *Pa-Wnt1* and *Pa-cad* at this early stage precludes the proper development of much of the embryo.

In this sense, this Pa-Wnt1+Pa-Cad module is operationally similar to the 'posterior organiser' observed in other animals. Wnt-signalling knock-down experiments in other arthropods, such as the spider *Achaearanea* and the insects *Gryllus*, *Oncopeltus*, and *Tribolium* all show similar phenotypes to *Periplaneta-Wnt1*^{RNAi} embryos: posterior truncations and a reduced GZ (Angelini and Kaufman, 2005a; Bolognesi et al., 2008; McGregor et al., 2008; Miyawaki et al., 2004). These results suggest a conserved role of Wnt-signalling in establishing and/or maintaining the GZ, including the activation of *caudal* (McGregor et al., 2008; Shinmyo et al., 2005), as in *Periplaneta*. In vertebrates, *Wnt* and *cad/Cdx* genes are also involved in posterior patterning and their regulatory interactions appear to be conserved as well. For instance, Wnt-signalling activates *cad/Cdx* expression in the posterior end of zebrafish and mouse embryos (Ikeya and Takada, 2001; Shimizu et al., 2005) and this activation is essential for the proper development of the animal. In fact, it has been proposed that most metazoans display a

Wnt-dependant posterior organiser and this must be a conserved ancestral feature of animal development (Martin and Kimelman, 2009). Results from *Periplaneta* are consistent with such a model.

Despite early *Pa-Wnt1* expression, no effects on the formation of the most anterior segments in *Pa-Wnt1*^{RNAi} embryos were noted, suggesting that formation of these segments is *Wnt1* independent or that there could be other, functionally redundant, Wnt ligands implicated at this stage (Bolognesi et al., 2008; Janssen et al., 2010; McGregor et al., 2008). Interestingly, formation of these anterior segments does not seem to involve N-signalling either (Pueyo et al., 2008), but a separate head-segmentation mechanism possibly involving the head gap genes *empty spiracles*, *buttonhead*, and *orthodenticle*, which in *Drosophila* and *Tribolium* directly regulate segmental genes such as *engrailed* and *wingless/Wnt1*, and whose expression and function seems widely conserved (Cohen and Jurgens, 1990; Royet and Finkelstein, 1997; Schinko et al., 2008; Schroder, 2003).

Control of segmentation by the posterior organiser

Altogether, the results presented here indicate that there exist interdependent and dual roles of the Wnt and N signalling pathways during posterior development and segmentation in *Periplaneta* (Fig. 3.12B). *Pa-Wnt1* signalling modulates the cyclic expression of *Pa-Dl* during germ band elongation in two opposite ways: it locally promotes *Pa-Dl* expression in the nearby posterior tip, while indirectly and via *Pa-cad*, dampens *Pa-Dl* expression in the mid-GZ. The combined effect of these two inputs is that the posterior tip source from which the cyclic waves of *Pa-Dl* originate is maintained, but the waves are only allowed to coalesce into a stripe at the anterior GZ. Reciprocally, *Pa-Dl* expression at the posterior tip has two distinct roles during posterior segmentation. On the one hand, it gives rise to the cyclical waves that precede and promote the sequential

formation of segments in the anterior GZ (Pueyo et al., 2008). On the other hand, constant N-signalling in the posterior tip of the embryo maintains the non-cyclic expression of *Pa-Wnt1* in nearby cells during germ band elongation, allowing for maintenance of the GZ until all the segments are laid down.

This model links with the earlier findings of Pueyo et al. (2008), and indeed explains some of the features then reported. *Pa-Notch*^{RNAi} embryos rarely display disruptions of segmentation in head or T1 segments, despite the generation of neurogenic phenotypes in these places (caused by the requirement for Notch in arthropod neurogenesis (Jiménez and Campos-Ortega, 1982)). The results reported here also highlight that the posterior tip expression of *Pa-Dl* and other members of the Notch pathway (Pueyo et al., 2008) is a relevant feature during Notch-mediated segmentation. The stripes of N-signalling promote the formation of segments in the anterior GZ but N-signalling also has a role in the maintenance of the GZ itself. This structure is highly reduced in *Pa-N*^{RNAi} embryos, involving a reduction of cell division and a mild increase in cell death (Pueyo et al., 2008). Following these results, this requirement for GZ maintenance can now be attributed to the expression of *Pa-Dl* at the posterior tip, which is essential to maintain *Pa-Wnt1* and *Pa-cad*.

Pa-cad appears as the most direct agent studied here involved in maintaining a GZ of appropriate size. *Pa*-Cad could act through a stimulation of cell division, or through a repression of differentiation as suggested by the lack of *Pa-Dl*, *Pa-en*, and *Pa-Wnt1* segmental stripes in the *Pa-cad* domain. In a unified hypothesis, these cellular and gene regulation effects can be traced to the role of *Pa*-Cad in dampening *Pa-Dl* expression throughout the posterior GZ. This partial repression still allows the progression of Notch-signalling waves through the *Pa-cad* domain, but only allows the formation of segmental Notch-signalling stripes outside it. Since these stripes of Notch-signalling eventually lead

to segment formation and differentiation, the expression of *Pa-cad* might be essential to maintain an undifferentiated, actively dividing cell population at the GZ that can continue to bud off new segments. Loss of *Pa*-Cad eventually allows high *Pa-Dl* levels across the posterior region of the embryo, producing widespread Notch-signalling, and hence, differentiation of all GZ cells and truncation of the embryo. In this model, as in normal development, patterning and growth are inextricably linked. The patterning activity of the Wnt posterior organiser sets up a GZ, which is then needed to provide cells for the formation and patterning of new segments by Notch. Notch itself maintains the posterior organiser, and hence the GZ, thus completing the circle.

The posterior segmentation gene network in other arthropods

Comparing the roles of Cad, Wnt, and Notch-signalling in different arthropods allows us to examine the possible conservation of a 'posterior segmentation gene network' (Wnt → Cad → N). Assessment of Notch-signalling has revealed dynamic expression of N-signalling members in the GZ in two spider species, *Achaearanea tepidariorum* and *Cupiennius salei* (Oda et al., 2007; Schoppmeier and Damen, 2005b; Stollewerk et al., 2003). Knockdown of N-signalling members via RNAi leads to loss of posterior segmentation in both species (Oda et al., 2007; Schoppmeier and Damen, 2005b; Stollewerk et al., 2003), but some differences were reported in the different studies. In *Cupiennius-N*^{RNAi}, only the segmentation of posterior segments was affected, while *Achaearanea-Dl*^{RNAi} and *N*^{RNAi} not only affected posterior segmentation but also had effects earlier in the formation of the caudal lobe (the spider GZ). These differences could be due to the RNAi techniques used. In *Cupiennius* dsRNA was injected into the eggs at blastula stage, whereas in *Achaearanea* RNAi was injected in the mothers, which produced stronger and earlier phenotypes hinting that zygotic RNAi may be less efficient

at interfering with N-signalling. Indeed, the *Achaearanea* results are similar to the maternal *Periplaneta-N*^{RNAi}, whereas the *Cupiennius* results are more similar to the cockroach embryos cultured with the Notch inhibitor DAPT. In *Achaearanea*, N-signalling is revealed as necessary for *caudal* expression in the GZ during later stages (Oda et al., 2007), and is similarly regulated by Wnt-signalling (McGregor et al., 2008). Thus, there seems to exist a connection between N and Wnt signalling in the establishment of the GZ and posterior germ band elongation in spiders, similar to the one reported here in *Periplaneta*, as well as striking similarities in the regulatory interactions between Cad, N, and Wnt signalling pathways during these processes.

This view is supported by a recent report in *Gryllus* where maternal *Gb-Dl*^{RNAi} and *Gb-N*^{RNAi} produced embryos displaying similar truncated phenotypes with reduced GZ similar to those observed in *Achaearanea* and *Periplaneta* (Mito et al., 2011). The authors suggest that posterior segmentation might be controlled by parallel mechanisms, one via stripes of N-signalling (segmentation) and the other through *cad* and Wnt-signalling (posterior growth). These pathways, however, appear to be dynamically linked in *Gryllus*, as in *Periplaneta*. At early stages, *Gb-wg* and *Gb-cad* are not regulated by N-signalling, whereas during germ band elongation *Gb-N* signalling at the posterior tip maintains *Gb-wg* and, therefore, *Gb-cad* which feeds back onto *Gb-Dl* (Mito et al., 2011). As in *Periplaneta*, these regulatory interactions can allow for sustained growth by maintaining a properly sized and functional GZ. Another study in *Gryllus* by Kainz et al. (2011) used zygotic *Gb-Dl*^{RNAi} and observed defects mostly in the development of the nervous system (a universal function of Notch-signalling in arthropods and the function that is most sensitive to loss or partial loss, as shown in *Drosophila* (Mohr, 1924; Van Breugel and Langhout, 1983), but partial loss of segment markers in only a small minority of embryos.

These results could be explained as in spiders where zygotic RNAi produced weaker phenotypes compared to maternal RNAi.

Overall, the different studies of Wnt and Notch signalling in different arthropods can be explained by the two functions of Wnt and Cad in maintaining 1) the GZ and 2) *Delta* expression; and by the two functions of Notch-signalling, 1) maintaining posterior *Wnt1* expression, and 2) in segmental stripes triggering segment formation in the anterior GZ. Different perturbations in different species (zygotic vs. maternal RNAi, embryo culture) reveal different aspects of these functions, but in doing so they offer a temporal window to the regulatory intricacies of the posterior gene segmentation network.

Evolution of the posterior gene network

My findings in *Periplaneta*, along with previous studies in other arthropods, suggest that there exists conservation in both the developmental roles and in the regulatory interactions among the Wnt → Cad → N posterior segmentation network. The most parsimonious explanation is that these different arthropods must have inherited such a network from their last common arthropod ancestor. This hypothesis begs the question whether conservation of the posterior gene network can be pushed back further in time. In vertebrates, Wnt-signalling regulates cyclic expression of N-signalling members in the presomitic mesoderm (PSM) (Dunty et al., 2008; Gibb et al., 2009). Recently, it has been shown that this Wnt-dependant regulation of N-signalling oscillations seems partly mediated through Cdx, as the expression of the mouse *Delta* homologue *Delta-like1* (*Dll1*) is disrupted in the PSM in *Cdx* mutants, and Cdx protein binds to the regulatory regions of *Dll1* (Grainger et al., 2012). Furthermore, patterns of expression of this posterior gene network have also been found in the GZ of several different species of annelids, the third segmented phyla, including segmental stripes of *Delta* and *Notch* and posterior expression

of *Wnt* and *cad* (Cho et al., 2010; de Rosa et al., 2005; Janssen et al., 2010; Rivera and Weisblat, 2009; Thamm and Seaver, 2008).

Curiously, vestiges of a Wnt-Cad organiser remain in insects that seem to have lost the requirement for N-mediated segmentation (Tautz, 2004). In the long germ band insect *Drosophila*, segmentation occurs 'all-at-once' without the need for a posterior organiser. However, a posterior stripe of *Dm-wg* is expressed in the blastoderm, independent of its future segment polarity expression and regulated to *Dm-cad*, though this expression has rarely been examined (Vorwald-Denholtz and De Robertis, 2011; Wu and Lengyel, 1998). While a Wnt-Cad posterior organiser may exist in *Tribolium* (Beermann et al., 2011; Bolognesi et al., 2008; Copf et al., 2004), the sequential addition of posterior segments occurs through a cyclical mechanism which involves pair-rule genes (Choe et al., 2006; El-Sherif et al., 2012; Sarrazin et al., 2012), but is not yet completely understood.

In summary, the available data strongly suggest that at least two components of the regulatory gene network, Wnt → Cad, were involved in posterior organization and growth in the last common bilaterian ancestor, the Urbilateria (De Robertis, 1997; Martin and Kimelman, 2009). In addition, following reports from annelids (Rivera and Weisblat, 2009; Thamm and Seaver, 2008), it seems likely that the third component of this gene network, N-signalling, would also have been involved in the Urbilateria. However, data is as yet unavailable from partially segmented (metameric) phyla and thus one cannot rule out the possibility that each of the segmented clades convergently and independently recruited a N-mediated segmentation mechanism (Chipman, 2010). However, this would require a) an ancestral in-built tendency of Cad to modulate Delta, b) a predisposition for the Notch pathway to form oscillatory clocks, and c) some unknown constraint of selective advantage leading to the repeated recruitment of these pathways in different phyla, in the face of other genes that could fulfil similar roles. The most parsimonious explanation

remains that this posterior gene regulatory network evolved once and is ancestral. Perhaps the Urbilateria developed by posterior elongation (Jacobs et al., 2005) and contained some kind of serially repeated structures, added sequentially from the posterior end of the animal (Balavoine and Adoutte, 2003; Couso, 2009; De Robertis, 2008b), and generated by the interplay between a Notch oscillator and the Wnt-Cad posterior organiser.

	Class H	**Class T**	**Class A**	**Wild Type**
***Pa-cad*$^{\mathbf{RNAi}}$**	202 (20.2%)	346 (34.5%)	167 (16.7%)	287 (28.6%)
***Pa-Wnt1*$^{\mathbf{RNAi}}$**	0 (0%)	309 (36.1 %)	281 (32.9%)	265 (31.0%)
***Pa*-$\mathbf{H_2O}$**	0 (0%)	0 (0%)	0 (0%)	300 (100%)

Table 3.1: Phenotypic series of *Pa-cad*$^{\mathbf{RNAi}}$ and *Pa-Wnt1*$^{\mathbf{RNAi}}$ embryos. RNAi embryos displayed a range of phenotypes categorized into three classes. Class 'H' embryos displayed head only or head with some undifferentiated tissue in the posterior and were found only in *Pa-cad*^{RNAi}. Class 'T' embryos form the head and some thoracic segments properly, but are truncated after either T2 or T3; this was the most common phenotype in both *Pa-cad*^{RNAi} and *Pa-Wnt1*^{RNAi} embryos. Class 'A' embryos were also observed in both *Pa-cad*^{RNAi} and *Pa-Wnt1*^{RNAi}. Embryos in this class developed complete head and thorax with defects in abdominal segmentation, usually tapering at the posterior. Control *Pa*-H_2O embryos showed no abnormalities, all were wild type in appearance.

Figure 3.1: Wild type expression patterns of *Pa-cad* (A-C), *Pa-Wnt1* (D-F), and *Pa-en* (G-I) in *Periplaneta americana*. (A-B) *Pa-cad* is expressed in a broad posterior domain (brackets) in post-blastoderm (stage 5; A) and early germ band elongation (stage 7; B) embryos. (C) During late germ band elongation, stage 9, *Pa-cad* is restricted to the mid-GZ (bracket) and is lost from the posterior tip (arrow). (D) *Pa-Wnt1* stage 4, post-blastoderm expression in head (arrow) and antennae (arrowhead) and in two posterior clusters of cells (*). (E) By early germ band elongation, stage 7, *Pa-Wnt1* is at the posterior (*) and in segmental stripes reaching T3. (F) During late germ band elongation, stage 10, *Pa-Wnt1* is expressed in an arc-like stripe in the posterior GZ (*) – set apart from the posterior tip (arrow), in segmental stripes in the anterior GZ (open arrowheads), in anterior segments and ventral appendages (black arrowheads). (G) *Pa-en* segmental expression up to T2 during post-blastoderm, stage 6. Additional stripes (black arrowheads) are added sequentially from the posterior during early (stage 7; H) and late (stage 10; I) germ band elongation and remain in the posterior of each segment throughout development (open arrowheads). T2, T3: second, third thoracic segment; A1, A4: first, fourth abdominal segment.

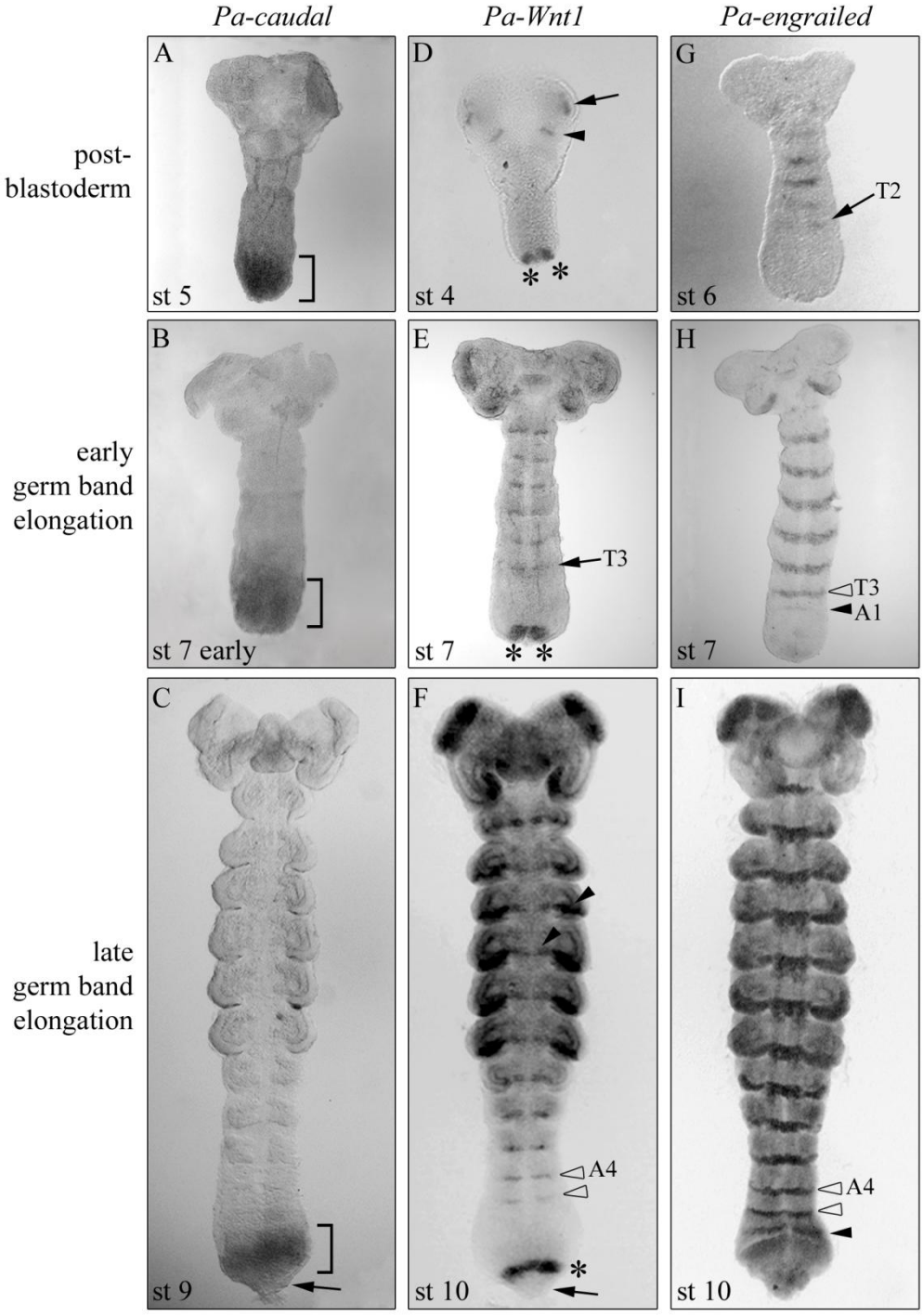

Pa-caudal
Pa-Wnt1
Pa-engrailed
post-
blastoderm
early
germ band
elongation
late
germ band
elongation
A
st 5
D
st 4
G
T2
st 6
B
st 7 early
E
T3
st 7
H
T3
A1
st 7
C
st 9
F
A4
st 10
I
A4
st 10

Figure 3.1

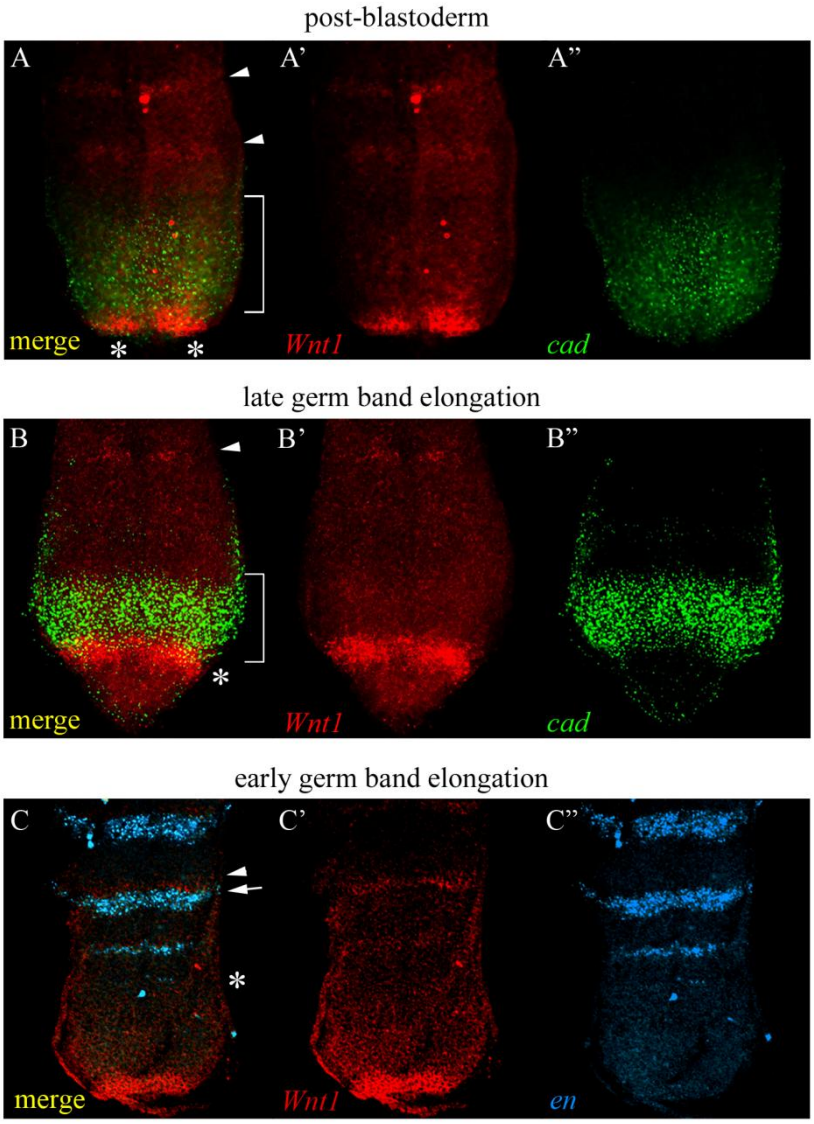


Figure 3.2: FISH reveals spatiotemporal patterns of *Pa-Wnt1*, *Pa-cad,* and *Pa-en*. (A-A") Posterior end of a post-blastoderm embryo (stage 6) showing *Pa-Wnt1* (red, A') and *Pa-cad* (green, A") patterns. *Pa-Wnt1* expression in segmental stripes (arrowheads), and posterior clusters (*) overlapping with *Pa-cad* domain (bracket). (B-B") Posterior end of a germ band elongation embryo (stage 9), showing *Pa-Wnt1* expression (B') in segmental stripes in anterior GZ (arrowhead) and a posterior arc (*) slightly overlapping with *Pa-cad* (B''; bracket). (C-C") *Pa-Wnt1* (C') and *Pa-en* (blue, C'') in the anterior GZ of a stage 7 embryo, with stripes of *Pa-en* (*) developing before *Wnt1* and defining the parasegment boundary. *Pa-Wnt1* (arrowhead) is expressed anterior and adjacent to *Pa-en* (arrow).

Figure 3.3: Disruption of *Wnt1* and *cad* function by maternal RNAi and embryo culture. (A-D) RNAi effects on *Pa-cad* and *Pa-Wnt1* expression at post-blastoderm, stage 6. The broad posterior domain of *Pa-cad* in wild type embryos (bracket, A) is greatly reduced in *Pa-Wnt1*^{RNAi} (bracket, B). Conversely, the wild type expression of *Pa-Wnt1* (arrow, C) is relatively unaltered in *Pa-cad*^{RNAi} (arrow, D; brackets in C-D insets). (E-H) RNAi affects *Pa-cad* and *Pa-Wnt1* during germ band elongation. The *Pa-cad* domain (bracket, E) is absent in *Pa-Wnt1*^{RNAi} embryos (*, F), while the arc of *Pa-Wnt1* (arrow, G) is unaffected in location (arrow, H) and width of expressing cells in *Pa-cad*^{RNAi} embryos (brackets in G-H insets). (I-K) RNAi phenotypes. (I) Stage 22 wild type embryo displaying head, thorax and abdomen. Class 'T' *Pa-cad*^{RNAi} (J) and *Pa-Wnt1*^{RNAi} (K) embryos show a similar body truncation after T3. (L-N) IWP-3 culture inhibition of Wnt-signalling affects *Pa-cad* expression. (L) Prior to culture (0 hour control), *Pa-cad* is in a broad posterior domain (bracket). After 16 hours in DMSO control culture there is no effect on *Pa-cad* (bracket, M), while IWP-3 cultured embryos show a marked decrease in *Pa-cad* (*, N). h: head; t: thorax; a: abdomen.

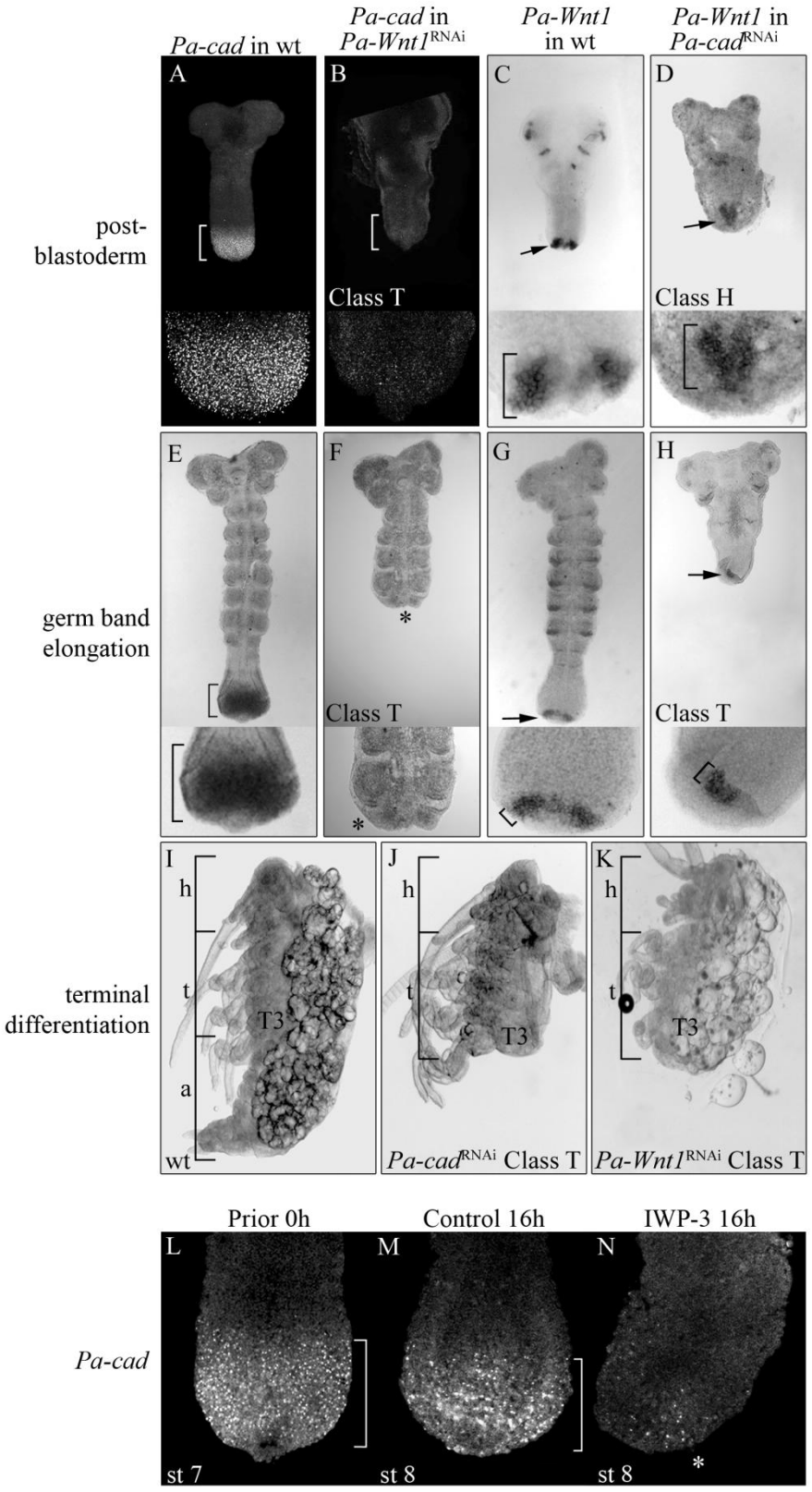
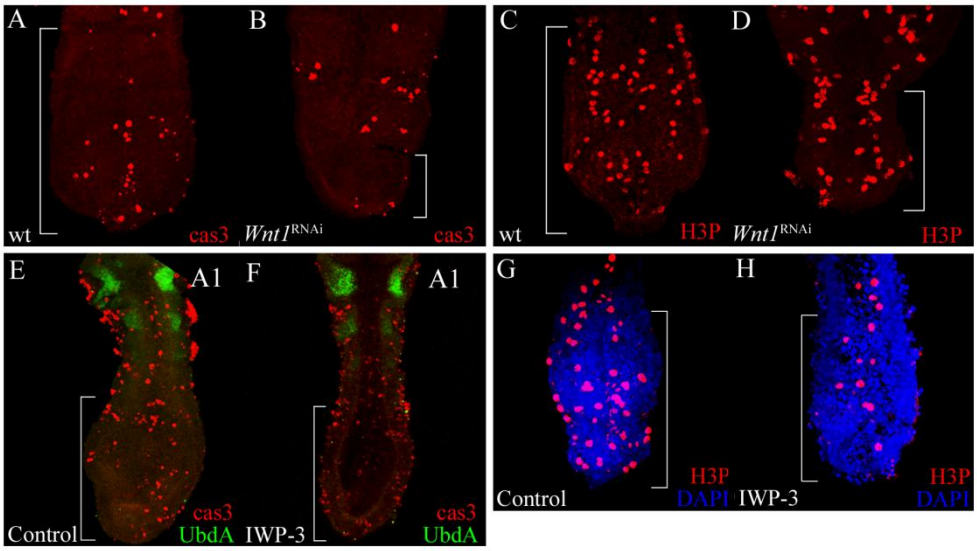

Pa-cad in wt
Pa-cad in Pa-Wnt1RNAi
Pa-Wnt1 in wt
Pa-Wnt1 in Pa-cadRNAi
post-blastoderm
germ band elongation
terminal differentiation
Pa-cad
A
B
C
D
Class T
Class H
E
F
G
H
Class T
Class T
I
J
K
h
t
a
T3
wt
Pa-cadRNAi Class T
Pa-Wnt1RNAi Class T
Prior 0h
Control 16h
IWP-3 16h
L
M
N
st 7
st 8
st 8

Figure 3.3

Figure 3.4: *Wnt1* regulates growth but not cell death in the growth zone. All embryos are stage 9 and the growth zone is marked with a bracket. In both wild type (A) and *Pa-Wnt1*^{RNAi} (B) embryos, apoptosis (as detected by α-cleaved caspase 3; cas3) is minimal and scattered throughout the GZ. (C) In wild type, numerous cells undergoing mitosis are seen (as detected by α-phosphorylated Histone 3; H3P). (D) Cell proliferation is lower in the small GZ of *Pa-Wnt1*^{RNAi} embryos. (E-F) The number of dying cells is increased in DMSO control (E) and IWP-3 (F) cultured embryos compared to A and B, yet similar to each other. This increase in cell death is likely due to the cytotoxic nature of DMSO and not to a role of Wnt signalling in apoptosis. Additionally, expression of the Hox genes Ubx and AbdA (as detected by mouse monoclonal antibody FP6.87; α-UbdA), remain unaffected in DMSO control (E) and IWP-3 (F) treated culture embryos, indicating that IWP-3 only affects Wnt signalling and not other patterning genes. In culture, the number of mitotic cells is reduced in control embryos (G), compared to wt (C), and are much lower in IWP-3 treated embryos (H). Staining with the vital dye DAPI showed similar amounts of living cells in control and IWP-3 treated embryos, indicating a healthy GZ. (I) Table representing the average number of cells counted undergoing apoptosis (cas3) or mitosis (H3P) in the wt and Wnt^{RNAi} embryos. Apoptotic index (J) and mitotic index (K) comparisons between DMSO control and IWP-3 cultured embryos. Indices calculated by dividing the number of dying (cas3) or proliferating (H3P) cells by the total number of cells found in the GZ (DAPI). Statistical analyses determined using a two-tailed T-test. A1 – first abdominal segment.



I

	cas3	H3P
wt	13 (n=4)	62 (n=3)
Pa-Wnt1^{RNAi}	8 (n=2)	44 (n=4)

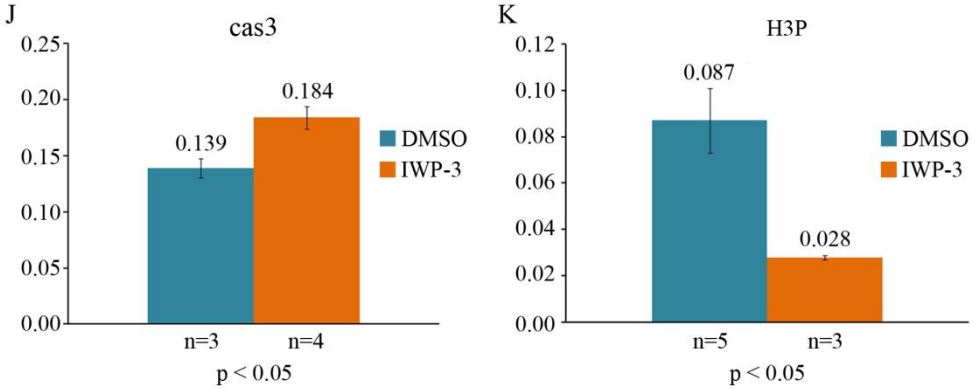


Figure 3.4

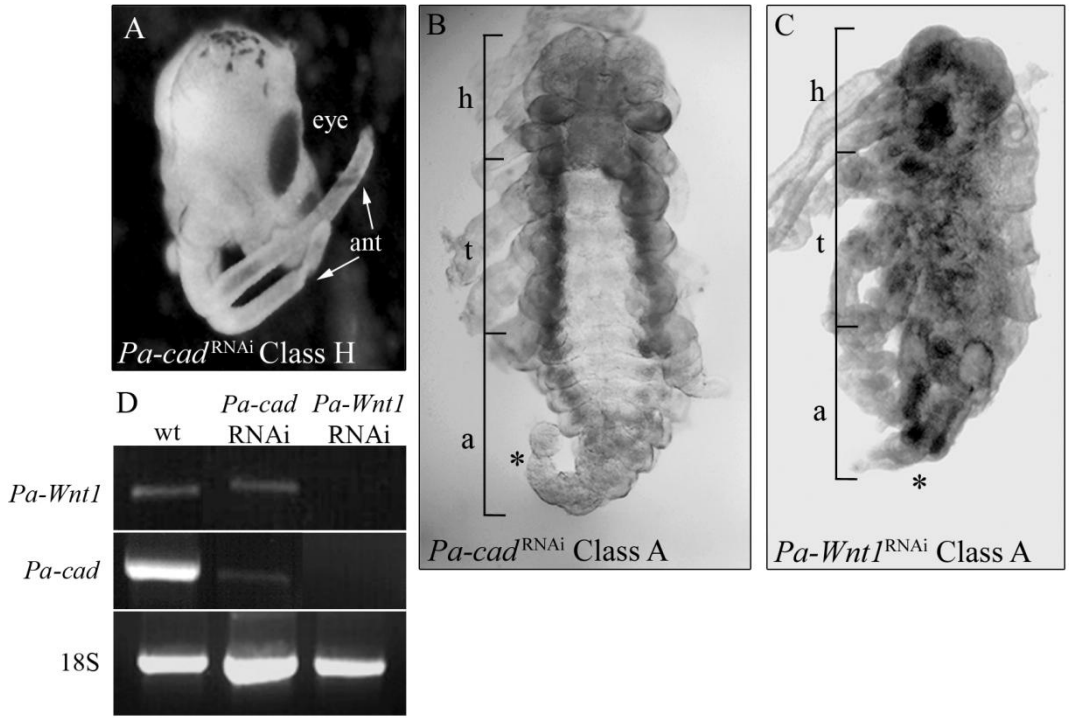


Figure 3.5: *Pa-cad*^{RNAi} and *Pa-Wnt1*^{RNAi} phenotypes and RT-PCR. (A) Class 'H' *Pa-cad*^{RNAi} embryo in which only the anterior-most head segments, eyes and antennae, develop properly. (B) In Class 'A' *Pa-cad*^{RNAi} embryos the head and thorax develop properly, while the abdomen appears normal in the anterior, but tapers into a point at the posterior (*). (C) Similarly, Class 'A' *Pa-Wnt1*^{RNAi} phenotypes have a normally developed head and thorax with a truncated and tapered abdomen (*). (D) RT-PCR analysis of stage 9 Class 'T' *Pa-cad*^{RNAi} and *Pa-Wnt1*^{RNAi} embryos. *Pa-Wnt1* expression remains at wild type levels in *Pa-cad*^{RNAi} but is abolished in *Pa-Wnt1*^{RNAi}. Only trace amounts of *Pa-cad* expression can be detected in *Pa-cad*^{RNAi} and is absent in *Pa-Wnt1*^{RNAi}. *Periplaneta* 18S ribosomal subunit was used as a positive control in each case. These data suggest that the different phenotypic RNAi classes correlate with the degree of hypomorphy (RNA loss) caused by the RNAi treatment on a single, continuous function required for posterior segment development throughout embryogenesis. In the case of *Pa-cad* if enough expression is present during germ band elongation, further segments will develop, if not, posterior segment development will be arrested and the animals will display only the head and some thoracic segments formed at blastoderm. Thus, Class 'H' *Pa-cad*^{RNAi} embryos display a total or very strong loss of *Pa-cad* RNA whereas in Class 'A' *Pa-cad* can remain at near wt levels (not shown). The same holds true for *Pa-cad* expression in *Pa-Wnt1*^{RNAi} embryos. ant – antennae; h – head; t – thorax; a – abdomen.

Figure 3.6: Expression of segmental markers in *Pa-cad*^{RNAi} and *Pa-Wnt1*^{RNAi} phenotypic classes. (A) *Pa-hairy* (*Pa-h*) expression in wild type embryo (stage 7) showing a broad domain at the posterior GZ and in transient segmental stripes (A1 and A2) in the anterior GZ. Expression of *Pa-h* is also observed in the head, CNS midline, and somites. (B) In a class 'H' *Pa-cad*^{RNAi} embryo *Pa-h* expression is in the head and in an expanded domain at the posterior but no segmental stripes are observed. (C) Similarly, in class 'T' *Pa-cad*^{RNAi} embryo, *Pa-h* segmental stripes are absent and *Pa-h* posterior GZ domain expanded. (D) Segmental expression of the *Pa-en* in a wild-type embryo (stage 8) reaching the A3 segment. (E) *Pa-en* expression in a class 'T' *Pa-cad*^{RNAi} embryo shows that segment formation proceeds until the thoracic segments. (F) Likewise, in a class 'T' *Pa-Wnt1*^{RNAi} embryo *Pa-en* segmental stripes are observed up to T3 segment. (G) Expression pattern of *Pa-en* in a stage 9 wild type embryo showing expression up to the A5 segment. (H) Expression of *Pa-en* in a class 'A' *Pa-cad*^{RNAi} embryo with a reduced GZ shows that only two abdominal segments have been formed. (I) Similarly, in a class 'A' *Pa-Wnt1*^{RNAi} embryo the A1 *Pa-en* segmental stripe has been laid down. Note that the GZ is reduced in size.

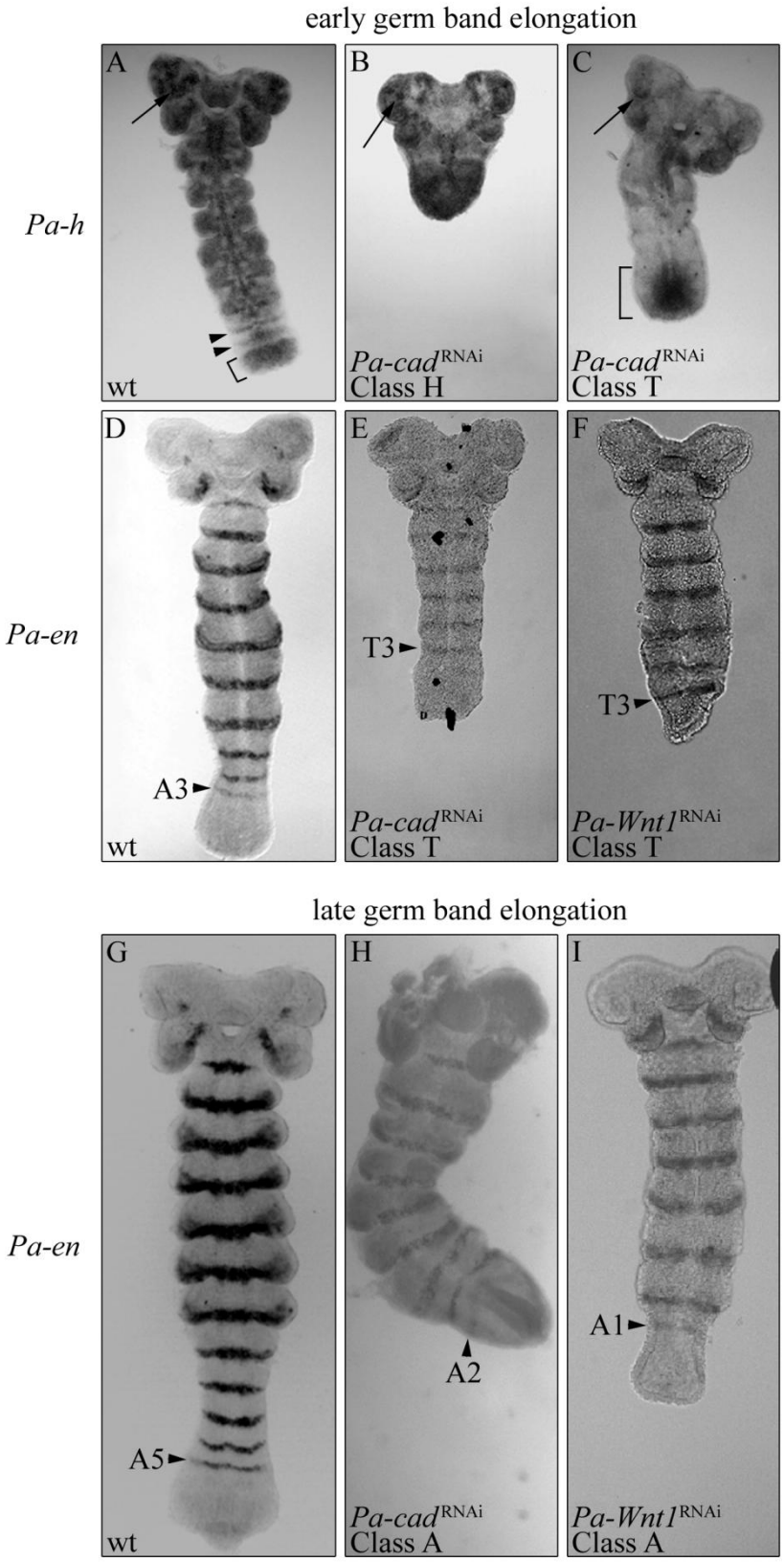

early germ band elongation
Pa-h
Pa-en
late germ band elongation
Pa-en
A
wt
B
Pa-cad RNAi
Class H
C
Pa-cad RNAi
Class T
D
A3
wt
E
T3
Pa-cad RNAi
Class T
F
T3
Pa-Wnt1 RNAi
Class T
G
A5
wt
H
A2
Pa-cad RNAi
Class A
I
A1
Pa-Wnt1 RNAi
Class A

Figure 3.6

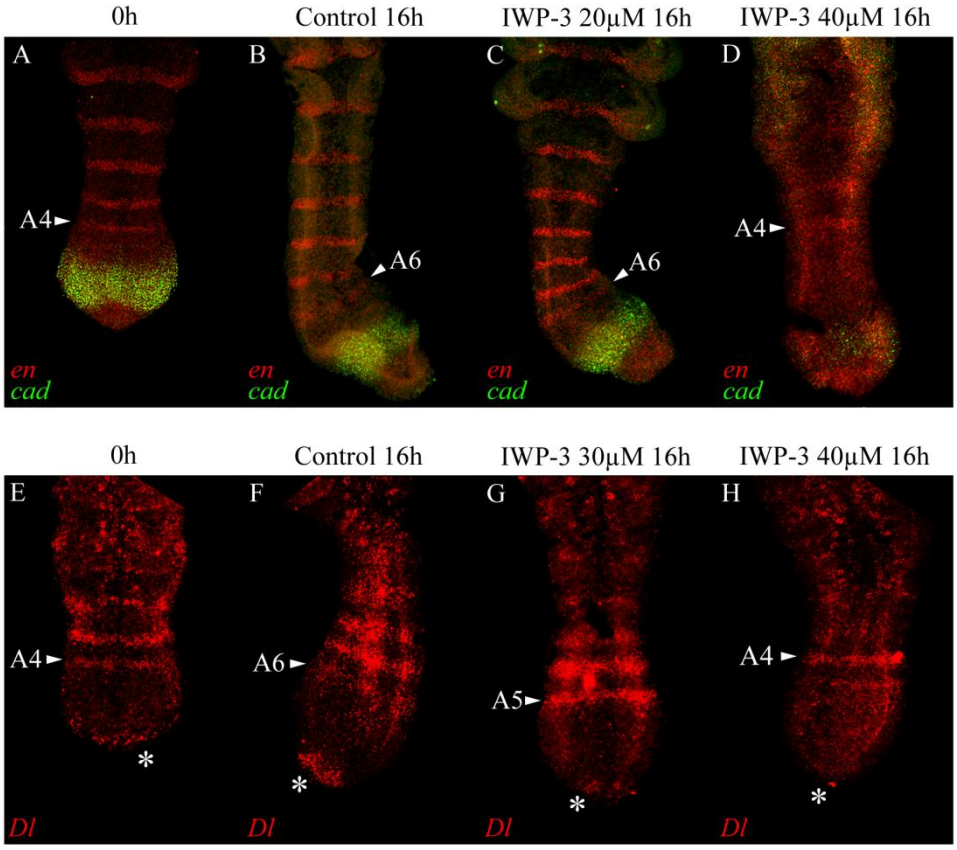


Figure 3.7: IWP 3 dosage effects on *Pa-en*, *Pa-cad* and *Pa-Dl* expression patterns. (A) *Pa-cad* (green) and *Pa-en* (red) patterns of expression in a stage 9 wild type embryo prior culture. (B) A 16h DMSO control culture embryo labelled as in A. *Pa-en* is expressed segmentally up to the A6 segment and *Pa-cad* is expressed in the mid-GZ. (C) *Pa-en* and *Pa-cad* expressions have not been affected in a 16 hour, 20 µM IWP3 culture embryo. (D) In a 16 hour, 40 µM IWP3 culture embryo posterior *Pa-cad* expression is highly reduced and no new *Pa-en* expressing segmental stripes have been added. (E) Pattern of expression of *Pa-Dl* in the posterior tip and segmental stripes reaching the A3 segment in the anterior GZ in a stage 8 wild type embryo before culture. (F) In a 16 hour DMSO control culture embryo *Pa-Dl* is strongly expressed in the posterior tip and in segmental stripes reaching the A6 segment. (G) *Pa-Dl* expression in a 16 hour culture embryo with 30 µM IWP3. *Pa-Dl* posterior tip domain is reduced and A5 *Pa-Dl* segmental stripe has been just laid down. (H) A 16 hour 40 µM IWP3 culture embryo showing complete absence of *Pa-Dl* at the posterior tip and *Pa-Dl* segmental expression reaching the A4 segment.

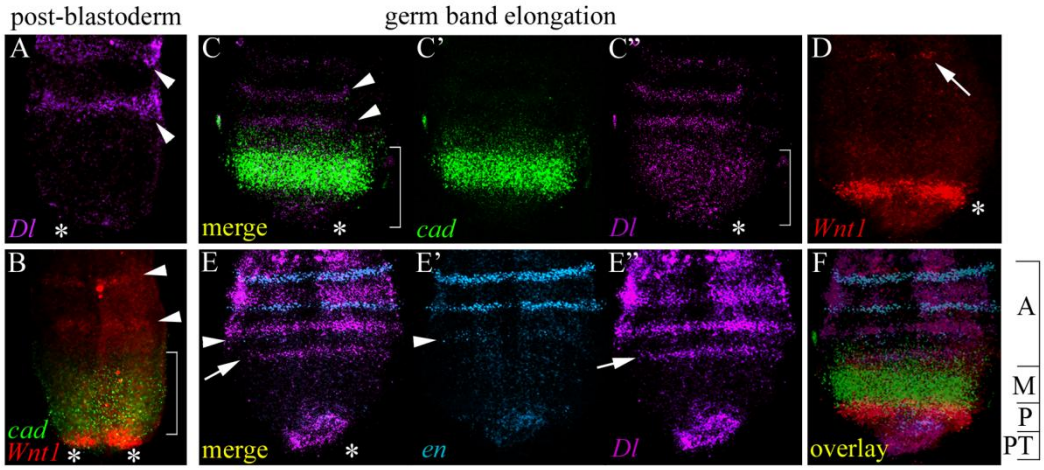


Figure 3.8: Co-expression of 'posterior network' genes at post-blastoderm (stage 6, A-B) and germ band elongation (stage 9, C-F). (A) *Pa-Dl* is expressed in anterior segmental stripes (arrowheads) but not in the posterior (*). (B) Double FISH from Fig. 2, showing *Pa-Wnt1* (red) and *Pa-cad* (green) expression. (C-C") Double FISH showing expression of *Pa-cad* (C') and *Pa-Dl* (purple, C"). The wave of *Pa-Dl* (bracket) emanates from the posterior tip (*), travelling through the *Pa-cad* domain and resolving into stripes in the anterior GZ (arrowheads). (D) *Pa-Wnt1* expression in segmental stripes (arrow) and in a posterior arc (*). (E-E") Double FISH for *Pa-en* (blue, E') and *Pa-Dl* (E"). *Pa-Dl* is expressed in the posterior tip (*) and the wave of expression has coalesced into a stripe in the anterior GZ (arrow) preceding the stripes of *Pa-en* expression (arrowhead). (F) Photomontage of C-E overlaid to show the GZ divided into four distinct posterior gene expression domains; from posterior to anterior - PT: posterior tip expressing *Pa-Dl*; P: posterior GZ arc of *Pa-Wnt1*; M: broad *Pa-cad* domain in mid-GZ; A: anterior GZ region expressing segmental stripes of *Pa-Dl*, *Pa-en*, and *Pa-Wnt1*.

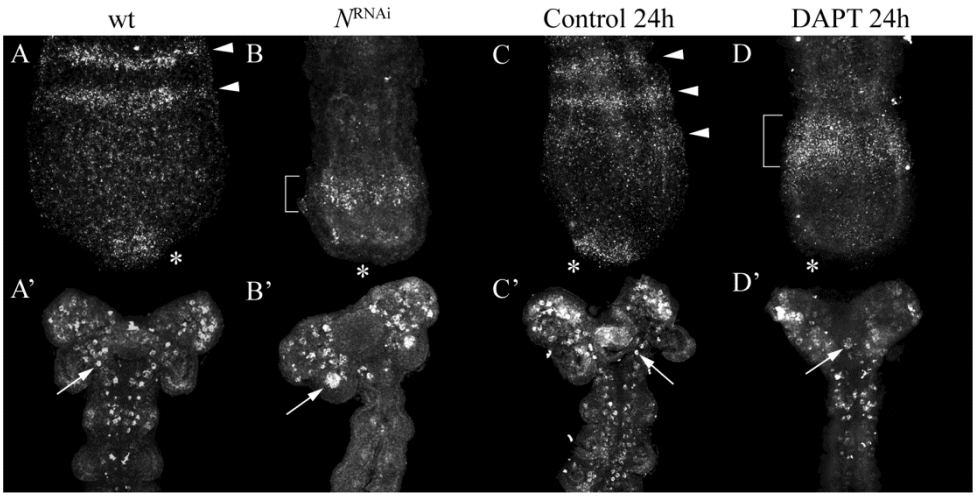


Figure 3.9: Notch signalling regulates *Pa-Dl* in the growth zone. (A-A') Wild type *Pa-Dl* expression in the posterior tip (*, A) and segmental stripes in the anterior GZ (arrowheads, A), as well as in isolated neuroblasts in the developing head (arrow, A') and ventral trunk. (B-B') In N^{RNAi} embryos, *Pa-Dl* expression reveals a loss of expression in the posterior tip (*, B) and a single broad band of expression in the anterior GZ (bracket, B) resulting from a failure to form segmental stripes, along with a neurologic phenotype revealed by clusters of neuroblast cells in the head and trunk (arrow, B'). (C-C') Similar to wild type, DMSO control cultured embryos express *Pa-Dl* in the posterior tip (*, C), in several stripes in the anterior GZ (arrowheads, C), and in the developing neuroblasts (arrow, C'). (D-D') *Pa-Dl* expression in DAPT treated embryos is similar to that in N^{RNAi} embryos; a broad band, but no stripes, of expression in the anterior GZ (bracket, D) and no expression in the posterior tip (*, D) with large clusters of neuroblasts in the head and trunk (arrows, D'). These results confirm that Notch signalling regulates the stripes and cyclic expression of *Delta* in *Periplaneta*.

Figure 3.10: *Pa-Dl* posterior tip expression during germ band elongation requires *Pa-Wnt1* and *Pa-cad*. (A-B) Stage 6, post-blastoderm *Pa-Dl* segmental expression (*) in wild type (A) remains in *Wnt1*^{RNAi} (B), despite the GZ being reduced (GZ outlines in A-B insets). (C) Stage 9, wild type expression of *Pa-Dl* during germ band elongation. (D) Loss of *Pa-Dl* at the posterior tip (arrow) and absence of stripes in the anterior GZ (bracket), but not in neuroblasts (arrowheads) in *Pa-Wnt1*^{RNAi} embryos at stage 9. (E-F) *Pa-cad* expression (brackets) is similar in stage 6 post-blastoderm wild type (E) and *N*^{RNAi} embryos (F). Stage 8, wild type expression of *Pa-Dl* (G) is expanded in *Pa-cad*^{RNAi} embryos (H) covering the posterior region and does not form stripes (bracket); neural expressions remain unaffected (arrowheads). (I-K) Inhibition of Wnt-signalling in culture affects *Pa-Dl* expression. (I) *Pa-Dl* in the posterior tip (arrow) and in segmental stripes (A3) in a 0 hour control embryo. (J) After 16 hours in DMSO control culture *Pa-Dl* is expressed in the posterior tip (arrow) and two additional stripes have been added in the anterior GZ (arrowheads). (K) Embryos cultured for 16 hours in the IWP-3 Wnt-inhibitor show greatly reduced *Pa-Dl* expression in the posterior tip (arrow) and only one new stripe (arrowhead) has formed. T1: first thoracic segment; A3: third abdominal segment.

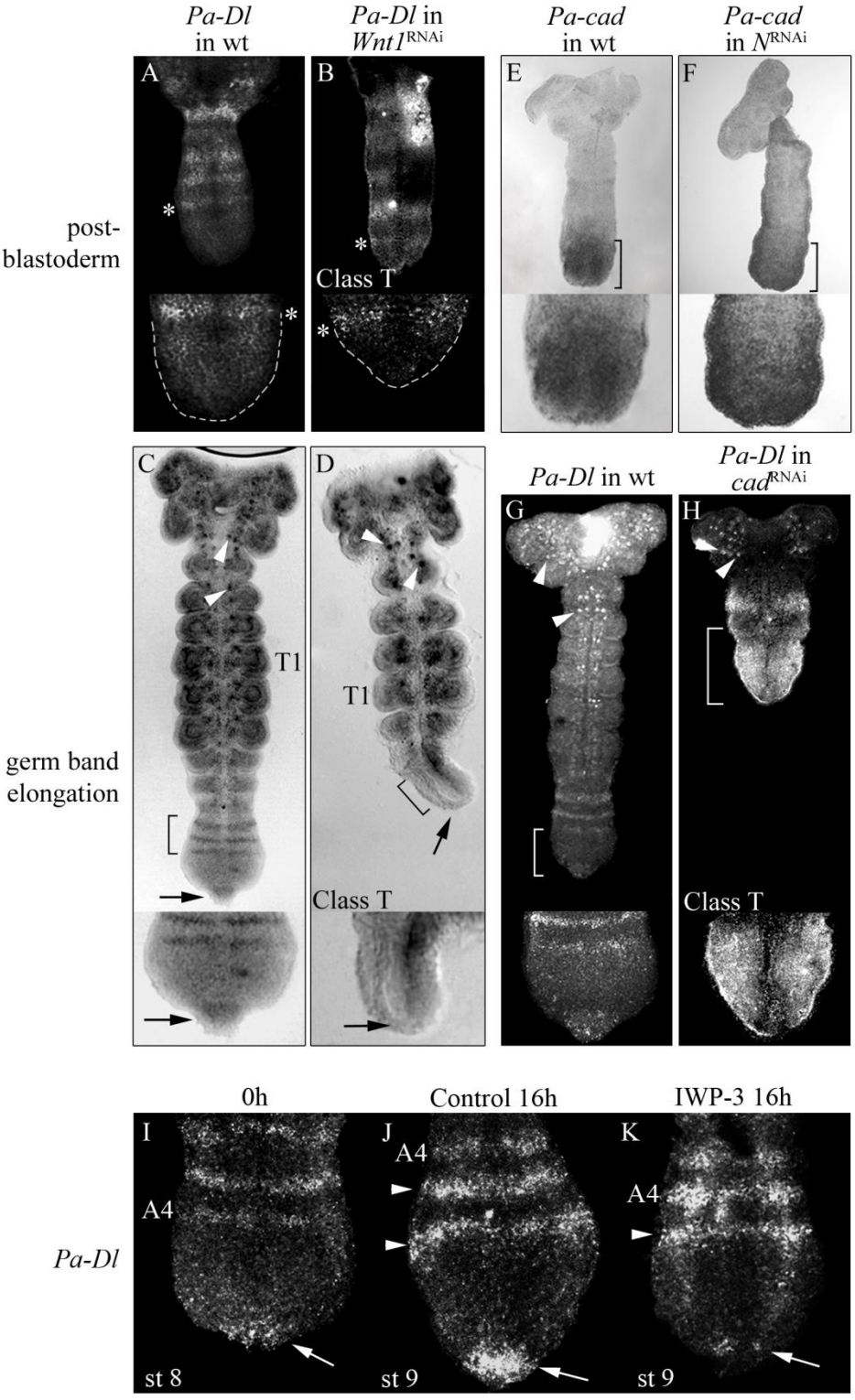

Pa-Dl in wt
Pa-Dl in Wnt1RNAi
Pa-cad in wt
Pa-cad in NRNAi
post-blastoderm
A
B
E
F
Class T
germ band elongation
C
D
T1
Pa-Dl in wt
Pa-Dl in cadRNAi
G
H
Class T
0h
Control 16h
IWP-3 16h
Pa-Dl
I
J
K
A4
st 8
st 9

Figure 3.10

Figure 3.11: N-signalling affects *Pa-cad* and *Pa-Wnt1* during germ band elongation. (A) Wild type, stage 9, *Pa-cad* expression (bracket). (B) Stage 9 N^{RNAi} embryos have a reduced GZ (outline, inset) devoid *Pa-cad* expression (*). (C) Wild type *Pa-Wnt1* expression at stage 8. (D) In N^{RNAi} embryos, stage 8, there is no posterior *Pa-Wnt1* expression (arrowhead), while the anterior segmental stripes remain unaffected (arrow; compare with C). (E-J) Effects of the N-inhibitor DAPT on posterior gene expression. (E-G) *Pa-cad* expression in the 0 hour control embryos (bracket, E) is unaffected in embryos cultured for 24 hours in DMSO control (bracket, F), but is reduced after 24 hours in DAPT culture (bracket, G). (H-J) *Pa-Wnt1* is in segmental stripes (arrowhead) and a posterior arc (arrow) in embryos prior to culture (0h control, H) and after 24 hours in DMSO control culture, in which a new *Pa-Wnt1* stripe forms in the anterior GZ (open arrowhead, I). (J) 24 hour DAPT cultured embryos show a reduction of the *Pa-Wnt1* posterior arc from four to two cells in width and the new A4 stripe is only partially formed (open arrowhead) (add insets of arc cell width). T2: second thoracic segment; A3: third abdominal segment.

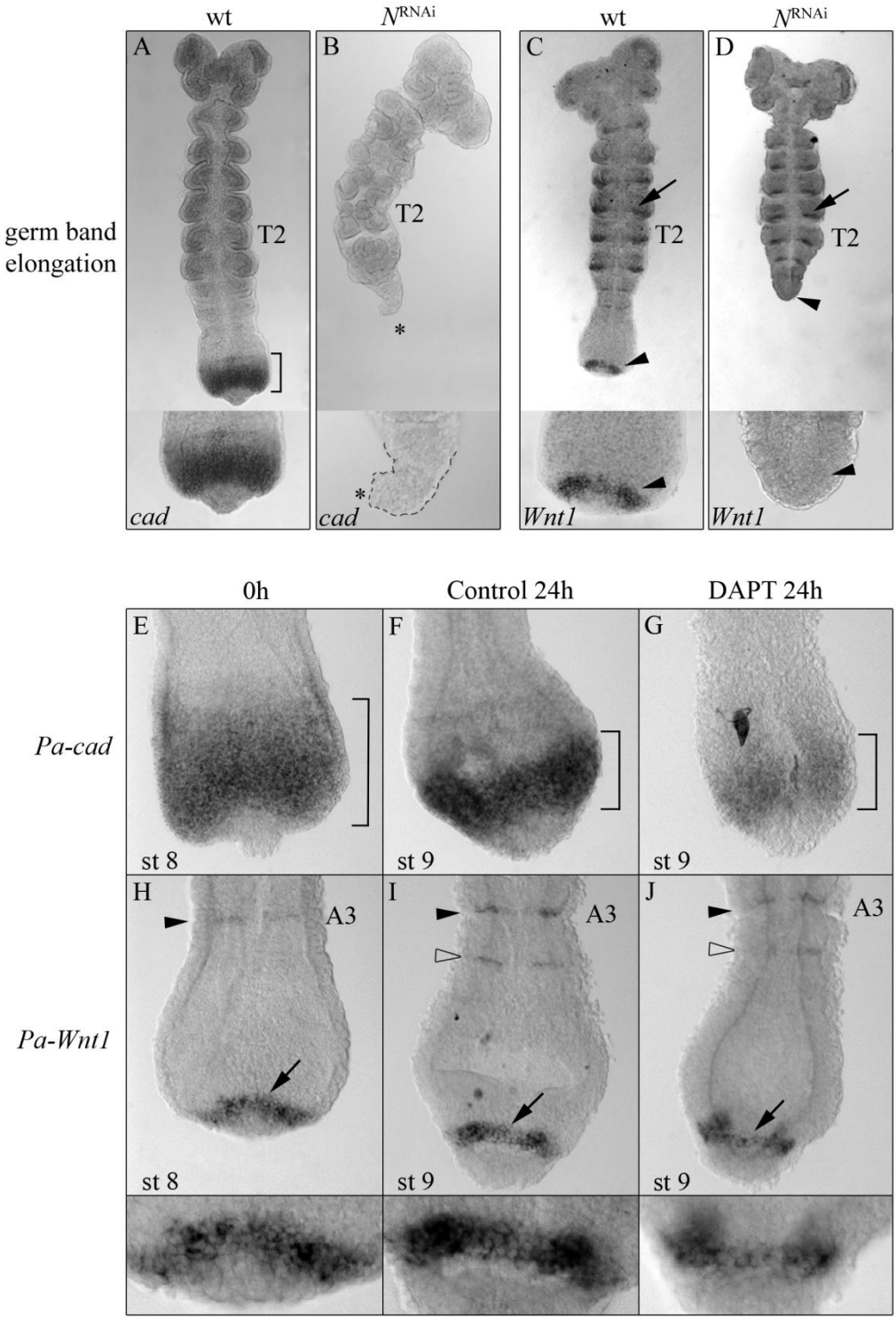

wt
N^RNAi
wt
N^RNAi
A
B
C
D
germ band elongation
T2
T2
T2
T2
*
*
cad
cad
Wnt1
Wnt1
0h
Control 24h
DAPT 24h
E
F
G
Pa-cad
st 8
st 9
st 9
H
I
J
A3
A3
A3
Pa-Wnt1
st 8
st 9
st 9

Figure 3.11

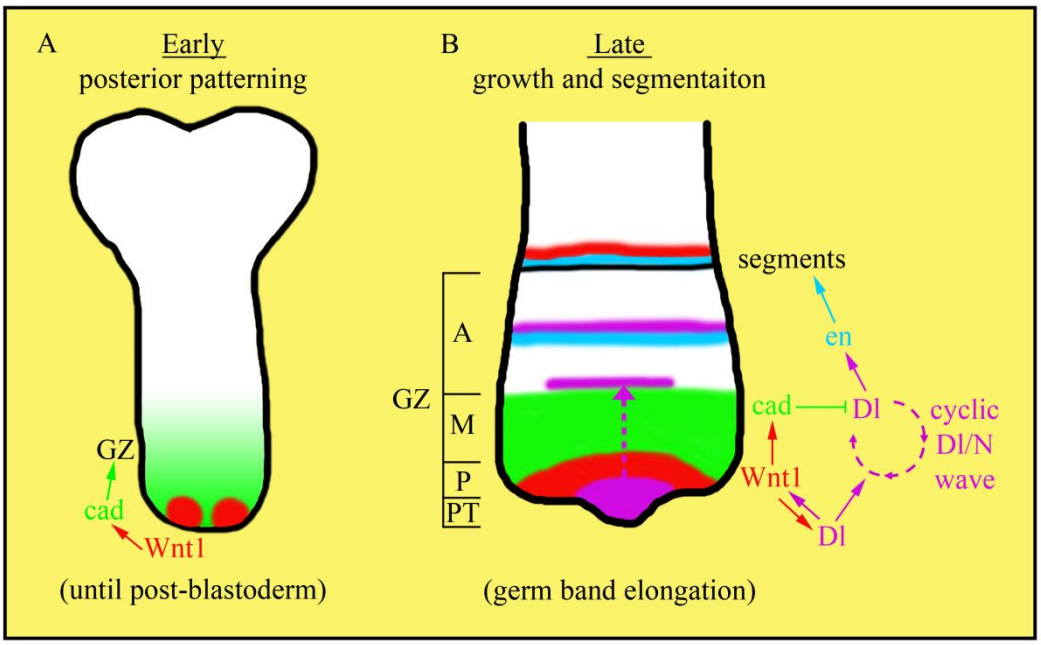


Figure 3.12: Model depicting the regulatory interactions of the posterior gene network in *Periplaneta americana*. (A) In the posterior region of the post-blastoderm embryo there are two clusters of cells expressing *Pa-Wnt1* (red). *Pa-Wnt1* activates the expression of zygotic *Pa-cad* (green) in the neighbouring cells. This activation is required for the posterior organization and establishment of the growth zone (GZ). *Pa-Dl* has no functional role at the posterior end at this stage. (B) During germ band elongation the embryo grows from the posterior and new segments are laid down sequentially from the GZ. The arc of *Pa-Wnt1* at the posterior GZ (P) has two main functions: 1) maintaining *Pa-cad* expression in a broad domain (green) in the mid-GZ (M) and 2) promoting *Pa-Dl* expression (purple) in the posterior tip (PT). Upon activation, *Pa-Dl* forms a positive feedback loop with *Pa-Wnt1*, thereby maintaining the expression of each other. As germ band elongation proceeds, cyclic waves of *Pa-Dl* expression emerge from the posterior tip and pass through the *Pa-cad* domain (dashed arrow). During this process *Pa-cad* dampens the expression of *Pa-Dl*, thus inhibiting the formation of segmental stripes but allowing the wave of expression to travel through. *Pa-Dl* stripes form outside the *Pa-cad* domain in the anterior GZ (A) and activate segmental stripes of *Pa-en* (blue). *Pa-en* then activates a segment polarity gene network involving *Pa-Wnt1* eventually leading to segment border formation. Thus, the interaction between the posterior gene network and the N-dependent oscillator regulates growth and segmentation in *Periplaneta* during germ band elongation.

CHAPTER IV

Pair-rule gene segmentation in *Periplaneta*

ABSTRACT

The *Drosophila* segmentation cascade is fairly well understood, involving highly complex regulatory interactions between groups of transcription factors: maternal coordinate → gap → pair-rule → segment polarity. However, as a derived long germ band organism, segmentation occurs simultaneously in a syncytial environment, while most other arthropods go through a short germband mode of development where segments are added sequentially from the posterior. Activation of the pair-rule genes is a key step in long germ band patterning heralding a change from a non-periodic to a periodic pattern of metameric organisation. Expression patterns of the pair-rule genes are highly variable amongst other arthropods, yet there appears to be conservation in the classification into primary and secondary groups. I have examined the expression and function of four pair-rule gene orthologues in the cockroach, *Periplaneta americana*: two primary pair-rule genes (*even-skipped* and *runt*) and two secondary (*pairberry* and *sloppy-paired*). Each gene is expressed segmentally with no indication of a 'pair-rule' pattern of expression in alternate segments. Subsequent functional analyses indicate a stronger requirement for the pair-rule genes in the anterior, those segments formed in the blastoderm, while their functions may be more redundant during posterior segment patterning, a process that may be dependent on the Notch-signalling pathway. While homologues of the pair-rule genes exist in other segmented bilaterians, such as vertebrates and annelids, they appear to have derived a new function in segmentation in the arthropods. Thorough analyses of the pair-rule and other segmentation genes in 'lower' insects and arthropods will help us to reveal the original nature of their functions and how they have changed over time, perhaps gaining a further understanding of the transition from short to long germ band development.

INTRODUCTION

The mechanisms of segmentation in arthropods are not very well understood outside of the long germ band insect, *Drosophila melanogaster*. Here, the "*Drosophila* paradigm" (maternal → gap → pair-rule → segment polarity) illustrates the progressive patterning of the embryo into smaller units, or segments, through a series of complex regulatory interactions comprising this hierarchic cascade of transcription factors. The pair-rule genes were first identified in a mutagenesis screen by Nüsslein-Volhard and Wieschaus (1980; 1984) in which *Drosophila* mutant larvae were missing every other segment. Subsequent studies of the mutated genes showed that they are expressed early at the blastoderm stage as seven alternating stripes corresponding to the missing larval segments (Gergen and Butler, 1988; Grossniklaus et al., 1992; Holmgren, 1984; Kilchherr et al., 1986; Macdonald et al., 1986).

It was observed that some of these pair-rule genes were expressed before the others: the primary pair-rule genes *even-skipped* (*eve*), *runt* (*run*), and *hairy* (*h*). These genes are actively involved in regulating the downstream secondary pair-rule genes *paired* (*prd*), *sloppy-paired* (*slp*), *odd-skipped* (*odd*), *fushi tarazu* (*ftz*), and *odd-paired* (*opa*) (Akam, 1987b; Howard and Ingham, 1986; Ingham, 1988). Each of these genes is initially expressed in the 'classic' pair-rule pattern of seven stripes in alternating segments before their final expression as fourteen segmental stripes, with the additional stripes arising *de novo* between existing ones or forming by the splitting of the initial stripes into two. The pair-rule genes regulate the expression of the segment polarity genes, *engrailed* (*en*) and *wingless* (*wg*), leading to parasegmental and segmental boundary formation (Jaynes and Fujioka, 2004). This chapter will mainly focus on four of these pair-rule genes: two primary (*eve*, and *run*) and two secondary (*prd* and *slp*), with some attention given to the primary pair-rule gene *h*, previously described by Pueyo et al. (2008).

Pair-rule patterning in *Drosophila*

even-skipped is a homeodomain transcription factor that, in *Drosophila*, is required to pattern the segmental expression of *Dm-en*, through indirect means, by inhibiting the *Dm-en* repressors *Dm-slp* and *Dm-odd* (Jaynes and Fujioka, 2004; Macdonald et al., 1986). *Dm-eve* hypomorphic mutants display the 'classic' pair-rule phenotype, missing the odd numbered segments, while null phenotypes present an asegmental ball of tissue with a lawn of denticle belts (Macdonald et al., 1986; Nüsslein-Volhard and Wieschaus, 1980). The pattern of *eve* expression in *Drosophila* is relatively conserved in other Diptera (Rohr et al., 1999) and Hymenopteran species (Binner and Sander, 1997). However, there are some differences in other long germ band insects, which may show no secondary stripes of *eve* expression (Rohr et al., 1999) or stripes may appear rapidly with no hint of an initial pair-rule pattern of expression (Grbic et al., 1996). These are only a few examples of the wide variation of expression patterns seen for *eve*; results are similar for the other pair-rule genes in both long and short germ band arthropods.

The primary pair-rule gene *runt* belongs to the Runt family of transcription factors that contain a conserved Runt domain capable of dimerisation and DNA binding activity. *Dm-run* interacts with the Groucho co-repressor to inhibit expression of both primary and secondary pair-rule genes, leading to the regulation of *en* and *wg* expression (Prazak et al., 2010). *Dm-run* mutants display a pair-rule phenotype as well as mirror-image duplication of the anterior region of each segment (Gergen and Wieschaus, 1985; Nüsslein-Volhard and Wieschaus, 1980). The function of *Dm-run* may be conserved in other long germ band insects, but has not yet been examined.

Dm-hairy belongs to the *Hairy* and *Enhancer of split* (*HES*) family of basic helix-loop-helix (bHLH) transcription factors. *Dm-h* is as a primary pair-rule gene that negatively regulates the expression of *Dm-ftz* (Howard and Ingham, 1986; Jaynes and

Fujioka, 2004). *Dm-h* mutant larvae are missing portions of the posterior odd and anterior even-numbered segments (Holmgren, 1984). Except for the conserved pair-rule expression of *h* in the housefly (Sommer and Tautz, 1991), the expression and function(s) of this gene have not been well studied in other long germ band insects.

Members of the Pax group III family of transcription factors are unusual in that they contain two DNA binding domains: a conserved Paired domain and a paired-type homeodomain (Bopp et al., 1986; Treisman et al., 1991). Members of this family include several *Drosophila* segmentation genes, such as the secondary pair-rule gene *paired* and the segment polarity genes *gooseberry* and *gooseberry-neuro*. The latter two possibly derived from a duplication event as they are expressed segmentally and function in segment polarity and neurogenesis, functions that are combined in many *paired* homologues in other species, hence the name *pairberry* (*pby*) in these organisms (Davis et al., 2001; Osborne and Dearden, 2005). *Dm-prd* is expressed in stripes that overlap the segmental boundaries (Bertuccioli et al., 1996; DiNardo and O'Farrell, 1987), with the subsequent loss of the posterior portion of the odd and anterior part of the even numbered segments in *Dm-prd* mutants (Nüsslein-Volhard and Wieschaus, 1980). This cross-border expression allows *Dm-prd* to regulate both the expression of *Dm-wg* and *Dm-en* (Baumgartner and Noll, 1990; DiNardo and O'Farrell, 1987).

The secondary pair-rule gene *sloppy-paired* is a member of a class of transcription factors containing the DNA-binding Forkhead domain and functions to regulate the expression of *en* and *wg* (Grossniklaus et al., 1992). Loss of expression of *Dm-slp* results in larvae missing the posterior portions of the even-numbered segments (Grossniklaus et al., 1992; Nusslein-Volhard et al., 1984). *Dm-slp* is unique in that it has an additional expression in the anterior head segments that arises independent of its later pair-rule stripe expression, where it functions as a head gap gene to inhibit the expression of the other

pair-rule genes in this region (Andrioli et al., 2004; Grossniklaus et al., 1992). Loss of *Dm-slp* head expression leads to large deletions in the head, including the antennae and gnathal segments (Andrioli et al., 2004; Grossniklaus et al., 1992), similar to those observed for some of the "head gap" genes that function in patterning the pre-gnathal head, such as *empty spiracles* and *buttonhead* (Cohen and Jurgens, 1990). Unfortunately, the expression and function of *slp* has not been thoroughly examined in other long germ band organisms.

The long germ band mode *Drosophila* development is highly derived and cannot be used as a general representative of other insects or arthropods. The hierarchy of the *Drosophila* segmentation cascade, namely gap regulation of the pair-rule genes within a syncytial environment, is mechanistically sophisticated and seems to have just suddenly appeared in evolution, as there is no real counterpart in other animals outside of insects. Arthropods undergoing the more common and ancestral mode of segmentation, short germ band, only form the anterior-most head, gnathal, and some of the thoracic segments in the early blastoderm while the remaining thoracic and abdominal segments are formed sequentially from the cellularized posterior growth zone (GZ) (Davis and Patel, 2002; Liu and Kaufman, 2005b). By studying elements of the *Drosophila* paradigm in 'lower' arthropods, we can gain a better understanding of how these methods work in different contexts and elucidate their possible evolutionary functions and origin.

Pair-rule genes in *Tribolium*

After *Drosophila*, the pair-rule genes have probably best been studied in the holometabolous, short germ band insect *Tribolium castaneum*. In *Tribolium*, *eve*, *run*, and *odd* are considered primary pair-rule genes and are expressed in broad domains in the posterior GZ from which primary stripes emerge in the mid-anterior before splitting into

thinner secondary stripes (Brown et al., 1997; Choe and Brown, 2009; Patel et al., 1994). RNAi depletion of any of these genes results in severe truncations (Choe et al., 2006). Knock-down of *Tc*-EVE expression via Chromophore Assisted Laser Inactivation (CALI) results in embryos displaying a pair-rule phenotype, missing every other segment (Schroder et al., 1999), confirming a function for *eve* at both the pair-rule and segmental level is conserved between *Tribolium* and *Drosophila*, assuming that truncations in the beetle are equivalent to asegmental mutants in flies. *Tc-eve*, *Tc-run*, and *Tc-odd* form an oscillatory gene circuit that regulates segmentation during *Tribolium* embryogenesis (Choe and Brown, 2009; Choe et al., 2006; El-Sherif et al., 2012; Sarrazin et al., 2012), coordinated/synchronised through a yet undetermined mechanism.

Tc-run plays an important role in inhibiting the expression of *Tc-prd* and *Tc-slp* in alternate segments (Choe and Brown, 2009; Choe et al., 2006). *Tc-prd* and *Tc-slp* activate *Tc-en* in their respective segments and loss of function of either leads to a pair-rule phenotype in the embryos, unlike the truncations observed for the primary pair-rule genes (Choe and Brown, 2007; Choe et al., 2006). Similar to *Drosophila*, *Tc-slp* plays an important role in anterior head development where RNAi loss-of-function phenotypes show deletions of large portions of the head including the antennae and gnathal segments (Choe and Brown, 2007), indicating a conserved function in this region.

Finally, the expression of *Tc-hairy* is in a pair-rule pattern, but functional analysis indicates this gene to have no apparent function during posterior segmentation, being required mainly in the anterior gnathal segments (Aranda et al., 2008). While the functions of the pair-rule genes in *Tribolium* are becoming more known, they are not as well understood in lower insects and other arthropods where functional analysis is still challenging. Expression studies of the pair-rule genes are limited and mainly focus on posterior patterning.

Pair-rule genes in other arthropods

The best studied of the pair-rule genes is *even-skipped*, which shows considerable variation in expression patterns (Fig. 4.1). The 'classic' double-segment pattern of expression seen in *Drosophila* and *Tribolium* appears to be more common in the Holometabola, but less so in hemimetabolous species where single-segmental expression in the posterior is more common. The posterior stripes of *eve* expression in *Gryllus* are a mix of pair-rule and segmental (Mito et al., 2007), while *Schistocerca eve* is only expressed in a broad posterior domain with no stripes forming in the anterior GZ, leading one to question its function in this organism (Patel et al., 1992). Many other species express single-segment stripes of *eve* in the GZ as shown for the insect *Oncopeltus fasciatus* (milkweed bug – Liu and Kaufman, 2005a), and phylogenetically basal arthropods including spiders (Damen et al., 2005), myriapods (Hughes and Kaufman, 2002a; Janssen et al., 2011), and crustaceans (Copf et al., 2003). In short germ band organisms, *eve* expression is often dynamic and transient, where a broad posterior GZ domain is refined into segmental stripes that fade as they move anteriorly. While a double-segment expression of *eve*, and other pair-rule genes, in the centipede *Strigamia maritima* has been observed in the GZ (Chipman and Akam, 2008), it may represent an example of convergent evolution, which possibly evolved as a mechanism to handle the rapid development of so many body segments (Janssen et al., 2011).

Functional data, where available, show better conservation of *eve* in body patterning compared to the deviations in expression. Similar to *Tribolium-eve*^{RNAi}, strong *Oncopeltus-eve*^{RNAi} embryos develop a "head only" phenotype (Liu and Kaufman, 2005a). Weaker *eve*^{RNAi} phenotypes in *Oncopeltus* and *Gryllus* both display a hypomorphic range of segment fusions with the posterior segments affected most, though no pair-rule-like phenotypes were observed (Liu and Kaufman, 2005a; Mito et al., 2007). From these

results, it is evident that the role of *eve* in segmentation is conserved in the Arthropoda, but the way this goal is achieved can vary based on expression, from pair-rule to segmental, or even a mix of the two.

The primary pair-rule gene *runt* has been examined in the holometabolous, short germ band insect *Bombyx mori* (Liu et al., 2008) where it is expressed in a double segment pattern and loss-of-function leads to an asegmental phenotype (Liu et al., 2008). In non-insect arthropods, such as the spider *Cupiennius salei* (Damen et al., 2000; Damen et al., 2005) and the millipede *Glomeris marginata* (Janssen et al., 2011), *run* expression is similar to its *eve* counterpart: a broad posterior domain is refined into a segmental stripe that moves into the anterior GZ where it fades as a new segment forms. While homologues of *runt* have been found in many arthropods and other metazoans (Duncan et al., 2008), functional studies have not yet been conducted to confirm any evolutionary correlation during segmentation processes.

Homologues of *paired/pairberry* have been studied in a range of arthropod species and were the first to show a pair-rule-type expression in a non-holometabolous insect, *Schistocerca americana* (Orthoptera), during both anterior and posterior patterning (Davis et al., 2001). Subsequent studies in several non-insect species showed pair-rule-type expression for *prd/pby* in the spider mite *Tetranychus urticae* (Dearden et al., 2002) and *Glomeris* (Janssen et al., 2012). In both these instances the pair-rule expression is restricted to the anterior segments specified at blastoderm, while *prd/pby* is segmentally expressed during germ band elongation (Dearden et al., 2002; Janssen et al., 2011), an expression pattern also observed in other chelicerates and myriapods (Davis et al., 2005; Schoppmeier and Damen, 2005a; Schwager et al., 2009). Loss-of-function studies have not been carried out in most organisms, so *prd/pby* function in segmentation can only be inferred in lower arthropods by comparison with *Drosophila* and *Tribolium*.

The secondary pair-rule gene *sloppy-paired* is expressed in single-segment stripes in the anterior GZ of *Oncopeltus* (Liu and Patel, 2010), *Cupiennius* (Damen et al., 2005) and *Glomeris* (Janssen et al., 2011). Janssen et al. (2012) recently showed a partial-pair-rule expression in the blastoderm of *Glomeris* for *slp*, as with *pby* and other pair-rule genes, supporting the notion that pair-rule type expression may be ancestral to the arthropods, at least those segments specified in the early blastoderm. However, functional analysis remains to be conducted in most arthropod species.

Of particular note, and special importance, is the pair-rule gene *hairy*, which has been shown to be necessary for proper segmentation in more basal organisms, where it is part of the Notch-segmentation pathway (Pueyo et al., 2008; Stollewerk et al., 2003). This mechanism appears to be conserved in sequentially segmenting animals represented in the three major phyla: arthropods, chordates, and annelids. In the basal arthropods *Periplaneta* and *Cupiennius*, *hairy* expression is dynamic in the GZ – a broad posterior domain is refined into single-segment stripes that move anteriorly until a new segment is formed (Damen et al., 2005; Pueyo et al., 2008). Conversely, *Glomeris-h* is expressed in segmental stripes only in the anterior GZ (Janssen et al., 2011), while *Strigamia-h* is expressed in alternating primary and secondary stripes in a pair-rule-like manner (Chipman and Akam, 2008).

The *Periplaneta* pair-rule gene orthologues *Pa-eve*, *Pa-run*, *Pa-pby*, and *Pa-slp* are expressed segmentally and sequentially during both early (post-blastoderm) and late (germ band elongation) embryonic development. There appears to be a stronger requirement for these genes in the anterior segments, those patterned in the blastoderm, compared to the posterior during germ band elongation where they may have more redundant roles. My data, combined with the literature, also confirm conservation in the primary/secondary pair-rule gene hierarchy based on spatiotemporal expression patterns. The expression

patterns of the pair-rule genes are highly variable (Fig. 4.1), yet the end goal in regulating segment polarity genes may be conserved, thus demonstrating the plastic and adaptable nature of these genes during segment formation.

RESULTS

Cloning of *Periplaneta* pair-rule genes orthologues

The search for pair-rule gene homologues in *Periplaneta* each began with an initial PCR using degenerate primers designed towards areas of high conservation (i.e. homeodomain - HD) of each of the known pair-rule genes. When an initial fragment was isolated, gene specific primers were designed and used in 5' and 3'RLM-RACE in order to obtain a full-length cDNA sequence by extending the sequence in the respective 5' and 3' ends. Once attained, the *Periplaneta* pair-rule gene homology was confirmed by phylogenetic analysis by aligning the conserved region of the predicted protein sequence with known sequences from other arthropods. In this way, I have successfully isolated and cloned four pair-rule gene orthologues in *Periplaneta*: *even-skipped*, *runt*, *paired/pairberry*, and *sloppy-paired*. An orthologue of *Pa-hairy* has previously been cloned and described by Pueyo et al. (2008).

Pa-even-skipped

The initial 113bp of fragment *Pa-eve* was isolated using degenerate primers designed towards the conserved HD (Appendix 2C). *Pa-eve* specific primers (Appendix 2C) used in 5' and 3'RACE resulted in the addition of 274bp in the 5' end and 602bp in the 3'end of the *Pa-eve* sequence giving a total length of 985bp (Appendix 5A). This sequence encodes a 295aa protein that contains the highly conserved 60aa *eve*-like HD as well as a conserved N-terminal domain and a Groucho interaction domain (LFQPYKT) near the C-terminus (Kobayashi et al., 2001; Macdonald et al., 1986). There is also a semi-conserved Alanine/Proline rich region just 3' to the HD, indicative of a Groucho-independent repressive function for *Pa-eve*, and a 5' region rich in PEST residues, typical of proteins with rapid turnover. Protein alignment of the HD with known *eve* protein

sequences from other arthropods revealed that this contig is, indeed, the *Periplaneta* orthologue of *even-skipped* and phylogenetic analysis reveals that it closely aligns to related species (Appendix 5B).

Pa-runt

A sequence for *Pa-run* (Appendix 6A) was assembled after the initial degenerate PCR, using primers designed to a highly conserved region of the Runt domain, yielded a 101bp fragment. The sequence was extended by 842bp and 366bp using *Pa-run* specific primers (Appendix 2D) in 5' and 3'RACE, respectively. This 1305bp sequence codes for a 258aa protein (Appendix 6A) that contains the Runt domain, though is missing the characteristic Groucho interaction motif (VWRPY) at the C-terminus found in most Runt family proteins (Aronson et al., 1997; Duncan et al., 2008). Protein alignment and phylogenetic analysis using a highly conserved portion of the Runt domain shows this sequence to be the *Periplaneta runt* orthologue (Appendix 6B).

Pa-pairberry

The *Periplaneta* orthologue for *paired/pairberry* was assembled as described above. The initial degenerate primers (Appendix 2E) used were designed towards highly conserved regions near the 5' end of Paired domain and the 3'end of homeodomain resulting in an initial 499bp fragment. 5' and 3'RACE extended this sequence by 141bp and 779bp, respectively, using *Pa-pby* specific primers (Appendix 2E). The resulting 1413bp sequence (Appendix 7A) encodes a 398aa protein that contains the conserved 96aa Paired domain, and a 59aa *paired*-like HD, including a 16aa extended region at the N-terminus. Protein alignment of the Paired domain with PaxIII group proteins from other arthropods indicates that this sequence is the *Periplaneta* orthologue of *pairberry*, as it

aligns closely to that of related species (Appendix 7B) and contains the Octapeptide (YSVDGILG) that is only found in *pairberry, gooseberry,* and *gooseberry-neuro* genes that are closely related to *paired* which, itself, does not contain this octapeptide sequence (Bopp et al., 1986; Davis et al., 2001).

Pa-sloppy-paired

A 700bp mRNA fragment of *Pa-slp* (Appendix 8A) was assembled from various PCR fragments starting with an initial 275bp fragment that was isolated using degenerate PCR primers designed towards the highly conserved Forkhead domain (Appendix 2F). 5'RACE using a *Pa-slp* specific primer (Appendix 2F) resulted in the extension of the sequence by 425bp but, unfortunately, the sequence could not be extended in the 3'direction. This *Pa-slp* sequence contains 91 of the 110aa conserved Forkhead domain as well as a Conserved Domain II near the N-terminus that contains a protein motif (FSISSIL) likely to be a Groucho co-repressor interaction domain (Andrioli et al., 2004). Protein alignment of the Forkhead domain reveals this sequence is the *Periplaneta* orthologue of *sloppy-paired* aligning closely with *slp* homologues from related species (Appendix 8B).

Embryonic expression patterns of *Periplaneta* pair-rule genes

Pa-even-skipped

Embryonic *Pa-eve* mRNA expression was examined via *in situ* hybridisation using a 397bp anti-sense riboprobe that hybridised with a sequence from the 5' end through the middle of the HD. Additionally, the 2B8 antibody was used to detect *Pa*-Eve protein expression (Patel et al., 1994). At stage 4 post-blastoderm, *Pa-eve* mRNA (Fig. 4.2A) and protein (Fig. 4.2B) are already expressed in five stripes in the three presumptive gnathal

and the first two thoracic segments, with the weakest expression in the mandibles. By early germ band elongation, stage 7, the anterior-most stripes have already begun to fade, leaving only the lateral edges of T2 mRNA expression remaining (Fig. 4.2C). In the posterior GZ, a faint broad domain is present and a new stripe of expression has formed at the presumptive A1 segment (Fig. 4.2C). A similar pattern of expression for *Pa*-Eve protein is observed at stage 7 (Fig. 4.2D), although perdurance of segmental *Pa*-Eve protein expression in more anterior segments indicates that protein expression/decay lags behind mRNA by several stripes (Fig. 4.2D). As the segment furrows become evident, *Pa-eve* expression is noted in the anterior compartment of each segment (Fig. 4.2D).

During late germ band elongation, the expression of *Pa-eve* (mRNA and protein) has completely diminished from the forming anterior segments and is now only observed in the GZ as a broad posterior domain and one to three stripes of expression in the anterior GZ (Figs. 4.2E-H). New stripes of *Pa-eve* expression emerge from the broad posterior domain in a sequential manner and as the stripes move anteriorly they begin to fade from the middle out to the lateral edges, completely dissipating before the formation of a new segment (Fig. 4.2E-H'). This pattern of expression persists until all of the segments have formed. Overall, *Pa-eve* has a dynamic, transient expression throughout embryogenesis. *Pa-eve* expression can be detected at the earliest post-blastoderm stages, when the germ rudiment is forming, as a series of five stripes corresponding to the future gnathal (mandible, maxilla, labium) and anterior thoracic (T1, T2) segments. These anterior stripes fade as a broad domain is established in the posterior GZ from which new stripes of expression will emerge in a sequential manner during germ band elongation and posterior segmentation.

Pa-runt

A 440bp riboprobe was synthesized that recognized a sequence from the 3' end of the Runt domain up to the poly-A tail and was used to detect *Pa-run* expression via *in situ* hybridisation. Similar to *Pa-eve*, *Pa-run* is expressed very early in development, by stage 4 post-blastoderm, as five gnathal and thoracic stripes (Fig. 4.3A). Similar to *Pa-eve* expression at this stage, the mandibular stripe of expression is the weakest (Fig. 4.3A). In addition, expression can be detected in the developing head lobes, which will remain throughout embryonic development (Fig. 4.3A-G). By stage 5, these initial stripes have widened and a posterior GZ domain becomes evident (Fig. 4.3B). During early germ band elongation, stage 6, the anterior-most stripes of *Pa-run* have started to fade from the midline out to the lateral edges (Fig. 4.3C) and by stage 7 (Fig. 4.3D) most of the original five anterior stripes have gone as new stripes are being added from the posterior domain. The segmental furrows begin to form at early germ band elongation and indicate *Pa-run* expression in the anterior of each developing segment (Fig. 4.3C-D).

As the embryo continues to develop, new stripes emerge sequentially (Figs. 4.3E-G) from the broad domain of expression in the posterior GZ. The broad posterior expression of *Pa-run* is dynamic, but not as strongly detected as *Pa-eve*, and is best seen with fluorescent *in situ* hybridisation (FISH; Figs. 4.3H-J). This dynamic expression resolves into a stripe in the mid-anterior GZ that moves anteriorly so that 2-3 stripes are expressed at any one time. As the stripes move into the more anterior GZ they begin to fade from the midline and are absent by the time the new segment takes shape (Figs. 4.3E-G and H-J). Once a segment has completely formed, expression of *Pa-run* appears *de novo* in the form of two ventral dots in the CNS on either side of the midline, in the anterior-ventral region of each segment, that become clusters as the segment matures (Figs. 4.3G and 4.3J).

Double FISH with *Pa-en* confirms that the early post-blastoderm stripes of *Pa-run* are expressed in the anterior of each segment (Fig. 4.3K) and that the stripes of *Pa-en* expression arise sequentially in an anteroposterior manner, becoming wider and more pronounced as the stripes of *Pa-run* fade in a similar progression. This pattern is observed in the GZ during germ band elongation as well (Fig. 4.3L); stripes of *Pa-run* arise well before those of *Pa-en* in the anterior GZ and their expressions do not overlap, as *Pa-run* expression is already fading when *Pa-en* expression begins to emerge. This suggests that *Pa-run* may negatively regulate the expression of *Pa-en*, most likely indirectly through the regulation of a secondary pair-rule gene.

Pa-pairberry

Pa-pby expression was detected in cockroach embryos using a 511bp riboprobe that hybridised to a sequence spanning the 3' end of the Paired domain to the middle of the homeodomain. *in situ* hybridisation revealed an initial expression of *Pa-pby* at post-blastoderm stages 3 and 4 in a small arc within each of the future antennae and a small cluster of expression in the developing ocular region of the head lobes (Fig. 4.4A-B). Initially, two faint stripes of expression can be detected in the mandible and maxilla at stage 3 (Fig. 4.4A). New stripes are added sequentially from anterior to the posterior until all five post-blastoderm stripes have formed at stage 5 (Fig. 4.4A-C). At late stage 6, a new stripe of expression has developed in the intercalary segment (Fig. 4.4D). These data indicate that the expression of *Pa-pby* lags behind *Pa-eve* and *Pa-run*, which at this stage are already expressed in 6-7 stripes, and shows that aspects of anterior segment specification might occur in a sequential and segmental manner.

The temporal delay observed in anterior *Pa-pby* stripe formation persists during germ band elongation. By stage 7, a weak stripe is forming at T3 (Fig. 4.4E), whereas by

this stage of embryogenesis stripes of *Pa-eve* and *Pa-run* are already expressed in the future A1 segment (Fig. 4.2C and 4.3C). In addition, as the segment furrows become apparent it is easy to see that *Pa-pby* is expressed in the posterior of each developing segment (Fig. 4.4E). The sequentially appearing stripes of *Pa-pby* are formed only in the anterior GZ (Fig. 4.4E-H) and there is no indication of a broad posterior domain like that of *Pa-eve* and *Pa-run*. Initially arising as thin, faint stripes in the mid-anterior GZ the stripes become stronger and more pronounced as they move anteriorly, remaining expressed in the ventral/posterior region of all segments throughout development (Fig. 4.4I). Lastly, the early expression of *Pa-pby* in the antennae and head lobes change during germ band elongation and are observed as two dots in the inner/proximal fold of these developing appendages (Fig. 4.4H).

Double FISH experiments were conducted in order to determine how the expression of *Pa-pby* in the GZ relates to that of the segment polarity gene *Pa-en* during germ band elongation (Fig. 4.5A), previously shown to arise only in the anterior GZ as well (Chapter III – Figure 3.1I) (Marie and Bacon, 2000). *Pa-pby* expression arises prior and anterior to that of *Pa-en*, and their expressions partially overlap (Fig. 4.5A). Double FISH also revealed that the stripes of *Pa-pby* expression develop after and posterior to *Pa-run* and by the time a strong, solid stripe of *Pa-pby* has formed, that of *Pa-run* has already begun to fade from the midline (Fig. 4.5B). This spatiotemporal pattern was also observed for *Pa-run* and *Pa-en* double FISH (Fig. 4.3L). Thin stripes of *Pa-en* appear posterior to, but not overlapping, those of *Pa-run* as they begin to fade from the midline. When the expression patterns of *Pa-run*, *Pa-pby*, and *Pa-en* are overlaid (Fig. 4.5C), it becomes apparent that as the stripe of *Pa-run* emerges from the broad posterior domain and moves towards the anterior GZ, the expression of first *Pa-pby*, then *Pa-en* develop only after that of *Pa-run* has dissipated. In addition, each of these genes is expressed in the developed

segments with *Pa-pby* and *Pa-en* in wide overlapping bands in the posterior of each segment (Fig. 4.5D) and two dots of *Pa-run* in the anterior, away from the other genes (Fig. 4.5E). Altogether, these data indicate that the expression of *Pa-pby* stripes in the anterior GZ may be negatively regulated by *Pa-run*, whereby the clearance of *Pa-run* expression releases the inhibition of *Pa-pby* allowing it to activate *Pa-en* expression.

Compared to *Pa-eve* and *Pa-run*, *Pa-pby* is expressed later during post-blastoderm development, forming sequentially from the anterior to include the five gnathal and thoracic segments. There is also a lag in expression during germ band elongation where the *Pa-pby* stripes only form in the anterior GZ after *Pa-run* (and *Pa-eve*) has already begun to fade. *Pa-pby* stripes arise posterior to and after *Pa-run*, but anterior to and before *Pa-en*. *Pa-pby* and *Pa-en* stripes partially overlap and this expression pattern remains in all segments throughout development.

Pa-sloppy-paired

A 428bp riboprobe was synthesized in order to examine *Pa-slp* expression during cockroach embryogenesis. This probe recognizes a region of the *Pa-slp* sequence spanning from the N-terminus through most of the Forkhead domain. At late stage 3 post-blastoderm, *Pa-slp* expression is detected in the presumptive antennae and mandibles (Fig. 4.6A). By stage 5, *Pa-slp* is expressed in five wide stripes in each segment from mandibles to T2 as well as strong expression in the developing antennae (Fig. 4.6B). These stripes remain strongly expressed at stage 6, when a faint new stripe is forming at T3 (Fig. 4.6C). At this stage, *Pa-slp* expression seems much more broadly expressed in the mandibles, taking up most of the segment (Fig. 4.6C). During germ band elongation, new stripes of *Pa-slp* expression appear in the anterior GZ, becoming stronger as they move further anteriorly (Fig. 4.6D'-E') and remaining in the newly developed segments,

although a bit weak, throughout embryogenesis (Fig. 4.6D-E). Similar to *Pa-pby*, there is no broad domain of expression in the posterior-most end of the GZ (Figs. 4.6D-E') and the appearance of *Pa-slp* segmental stripes of expression are likely to lag behind those of *Pa-eve* and *Pa-run*, although further analysis is required.

Pa-hairy

A *Periplaneta* homologue of the pair-rule gene *hairy* was previously described by Pueyo et al. (2008) in relation to germ band elongation. To examine *Pa-h* expression at earlier stages, a similar riboprobe was used that was available in the lab. The early post-blastoderm expression of *Pa-h*, stage 5, is in five stripes in the gnathal and first two thoracic segments, as well as in the antennae and developing head lobes (Fig. 4.7A). During germ band elongation, new stripes of *Pa-h* expression develop sequentially in the anterior GZ as they emerge from a broad posterior domain (Fig. 4.7B-D; Pueyo et al., 2008). At stage 6, a strong posterior expression develops (Fig. 4.7B) and by stage 7, when the segment furrows are forming, *Pa-h* appears to be expressed in the posterior of each segment (Fig. 4.7C). Also at stage 7, *Pa-h* expression begins to appear as a line in the ventral CNS (Fig. 4.7C), which will remain throughout the rest of embryonic development, as the stripes of expression begin to fade from developed segments (Fig. 4.7D). The broad posterior GZ expression is similar to *Pa-eve* and *Pa-run*, and in each case this expression is dynamic, resolving into segmental stripes of expression in the anterior GZ. These stripes move anteriorly and fade by the time a new segment is established. The posterior expression of *Pa-h* is regulated by the Dl/N pathway, where the stripes form anterior and adjacent to those of *Pa-en*, potentially regulating their expression (Pueyo et al., 2008).

Function of *Periplaneta* pair-rule gene orthologues

Pa-eve

In order to determine the function of *Pa-eve* during embryonic development, a 400bp long dsRNA fragment was injected into virgin female cockroaches from which subsequent embryos and nymphs were analysed for phenotypes. *Pa-eve*^{RNAi} phenotypes provided a hypomorphic series in which affected embryos were placed into two classes (Table 4.1). Class I phenotypes (60/667; 9%) were typically observed in earlier staged embryos and were long and skinny with weakly defined segments (Fig. 4.8B-C), except for the mandibles where *Pa-eve* expression is the weakest (Fig. 4.2A-B), and a normal appearing GZ. *in situ* hybridisation for *Pa-en* in these Class I *Pa-eve*^{RNAi} embryos showed that the expression of *en* was very weak and thin (Fig. 4.8B) or completely absent (Fig. 4.8C) compared to wild type (Fig. 4.8A), indicating a failure to define, or maintain, segment boundaries. *Pa-eve*^{RNAi} Class II phenotypes (179/667; 26.8%) were commonly observed in later stages of embryogenesis when the limb buds become apparent (Fig. 4.8E,G). In these embryos, the mandibles appear unaffected (Fig. 4.8E), while extensive fusion occurs between the more posterior gnathal and thoracic segments (Fig. 4.8E,G). These phenotypes were somewhat erratic, but typically involved fusion between the maxillary and labial segments (Fig. 4.8E) and a varying degree of fusions between the thoracic segments, with the most extreme phenotypes displaying a fusion of all three segments, as shown in Figs. 4.8E and 4.8G.

Pa-eve^{RNAi} first nymphs (Table 4.1) display similar fusions of gnathal and thoracic segments as Class II embryos, although less frequently observed (14/399 – 3.5%; Fig. 4.8I-J'). These nymphs displayed a great deal of ventral and dorsal thoracic fusions (Figs. 4.7I,J), typically between T1/T2 or T2/T3. In one example, all thoracic segments were fused both dorsally and ventrally (Figs. 4.7J-J'). Additional, less common phenotypes

include fusions of lb/T1, T3/A1, and deletion of the maxillary segment. Very weak defects in posterior patterning and segmentation were observed in only two nymphs where two abdominal segments were fused at the midline on both dorsal and ventral sclerites (arrow, Fig. 4.8I-I').

This functional analysis shows that *Pa-eve* is important in anterior segmentation, except in the mandibles. *Pa-eve*^{RNAi} resulted in extensive fusion between the remaining gnathal and/or thoracic segments, especially in the thorax. Stronger phenotypes were observed in the early staged embryos in which segments were only weakly formed, at best, with a marked reduction or absence of *Pa-en* expression. While RNAi may not give a true null phenotype, as residual transcript may be enough to be effective, the Class I phenotypes indicate that complete knockout of *Pa-eve* could lead to asegmental embryos that may continue growing in the posterior from a normal GZ. Perhaps the reason that no Class I nymphs were observed, is that the effects of *Pa-eve*^{RNAi} proved lethal beyond a certain point in development, which could also explain the low penetrance of Class II nymphal phenotypes. Altogether, these results show a conserved function for *Pa-eve* in segment patterning and formation with a strong requirement for proper segmentation of the anterior segments established at post-blastoderm.

Pa-runt

Functional analysis via *Pa-run*^{RNAi} involved injecting a 615bp dsRNA fragment corresponding to part of the 5'UTR through most of the Runt domain. Phenotypic variation of *Pa-run*^{RNAi} embryos allowed them to be placed into either of two classes (Table 4.1). Class I embryos (75/334; 22.5%) showed extensive fusions of gnathal and thoracic segments and typically displayed a pinched abdomen (Fig. 4.9B-C). In addition, partial or total loss of various segments was observed, with the most common being loss

of the maxillae (Fig. 4.9B), whereas the mandibles always remained unaffected (Fig. 4.9B-C). These results may be explained in part as a result of diminished *Pa-en* expression in the anterior segments (Fig. 4.9D) leading to failure of segment boundary formation or maintenance, while the expression of *Pa-en* in the posterior GZ remains relatively unaffected. Class II *Pa-run*^{RNAi} embryos (15/334; 4.5%) showed a unique phenotype of mirror image duplications in some segments (Fig. 4.9E-E'), a phenotype also observed in *Drosophila-run* mutant larvae (Gergen and Wieschaus, 1985; Nüsslein-Volhard and Wieschaus, 1980).

Pa-run^{RNAi} phenotypes observed in nymphs (16/135; 11.9%) resembled the Class I embryonic phenotypes described above, exhibiting segmentation defects in the anterior, but not in the abdominal segments (Fig. 4.9H, J; Table 4.1). The most common first nymph phenotypes include fusion of thoracic segments, typically T1/T2 (Fig. 4.9H), and reduced or missing maxillae, similar to *Pa-eve*^{RNAi} embryos. In addition, some nymphal phenotypes present with partial loss of labium and/or fusion of lb with T1, as well as a loss of one or both T1 legs (Fig. 4.9J).

Although low in number, embryonic *Pa-run*^{RNAi} Class II mirror-image segment phenotypes indicate an important and conserved role for *Pa-run* in the regulation of segment polarity genes possibly through negative regulation of *Pa-pby*. As shown in Fig. 4.5B, the stripes of *Pa-pby* form in the anterior GZ only after the expression of *Pa-run* begins to fade from the midline. *Pa-run*^{RNAi} leads to irregular formation of *Pa-pby* stripes in the GZ (Fig. 4.10B-C), sometimes in misaligned diagonal stripes, which may be able to explain some Class I *Pa-run*^{RNAi} phenotypes, where only one half of some segments are missing. In addition, knockdown of *Pa-run* expression resulted in ectopic expression of *Pa-pby* in the anterior segmental compartments (Fig. 4.10B-B'). This ectopic expression would explain the development of mirror-image segment phenotypes observed in *Pa-*

run^{RNAi}. Overall, these results indicate that *Pa-run* regulates the expression of *Pa-pby* by inhibiting its expression in the anterior compartment of each embryonic segment at both early post-blastoderm and later during germ band elongation; a function that is conserved in *Tribolium* (Choe and Brown, 2009; Choe et al., 2006). Similar to *Pa-eve*, *Pa-run* appears to be more strongly required for proper development of the anterior gnathal and thoracic segments formed at the earliest stages of development, likely the indirect regulation *Pa-en* expression through *Pa-pby*. While some defects were observed in the posterior, in the form of a pinched abdomen and misaligned segments, posterior growth and segmentation otherwise proceeds as normal.

Pa-pairberry

The 405bp dsRNA injected for functional analysis of *Pa-pby* included a region spanning the Paired domain to the 5' end of the homeodomain, including the extended region. Unfortunately, even after several attempts, *Pa-pby*^{RNAi} did not result in a substantial phenotype other than a few embryos (Class II; 16/133; 12%; Table 4.1) which were short and wide with a few exhibiting fusions between lb and T1. Only one embryo (Class I; 1/133; 0.75%) displayed a significant phenotype, having extensive fusions of all segments posterior to the labium (Fig. 4.4J). The abnormal folding of the embryo makes it hard to determine if this region is unsegmented. However, as shown previously, loss of proper segmentation often leads to fusions that cause abnormal bends and twists in the developing embryos, which could be an indication of a function for *Pa-pby* in segment formation. No nymphal phenotypes were observed. While *Pa-pby*^{RNAi} embryonic phenotypes were not totally conclusive, they do suggest a role for *Pa-pby* in posterior patterning in the roach.

One potential reason for such a low penetrance could simply be due to the small number of oothecae collected. *Pa-pby*^{RNAi} was attempted on three separate occasions in small groups of 2-4 females per round. In none of these repeats were more than a dozen eggs collected, which could potentially indicate a previously undescribed role for *Pa-pby* during oogenesis. Alternatively, the dsRNA synthesised for *Pa-pby*^{RNAi} may have recognised off-target sites of different genes required for proper oogenesis, resulting in the low fecundity of these females. However, no such gene was detected in a 1000 gene hits from a BLAST search using the *Pa-pby* dsRNA sequence used for RNAi.

Pa-sloppy-paired

Of all the pair-rule genes studied, it is with *Pa-slp* that the strongest and most consistent phenotypes were displayed. *Pa-slp*^{RNAi} was conducted by injecting a 525bp dsRNA fragment directed towards a region from the start codon to the middle of the Forkhead domain. Class I *Pa-slp*^{RNAi} embryos (Table 4.1; 12/170; 7.1%) were long and skinny with limited observable segmentation, similar to the *Pa-eve* Class I embryos shown in Figs. 4.8B,C. Class II *Pa-slp*^{RNAi} embryos (54/170; 31.8%) displayed malformed mandibles and some segments appeared "lumpy" due to a what may be a mirror-image duplication of some segments (Fig. 4.11B), as seen in *Pa-run*^{RNAi} (Fig. 4.9E').

Pa-slp^{RNAi} first nymphs also displayed a hypomorphic range of phenotypes and were placed into two main classes (Table 4.1). Class I nymphs (209/349; 59.9%) were variably missing one or both mandibles and one or both antennae, though not always on the same side (Figs. 4.11D,G). In some cases the mandibles in Class I nymphs were only partially formed, where the molar was missing (Fig. 4.11D1,D2) and sometimes the incisors were malformed in a way that they looked split or duplicated (Fig. 4.11D2). Class II nymphs invariably died before or just after hatching and displayed large deletions of

head and gnathal segments (20/349; 5.7%; Fig. 4.11H). Not only were the mandibles and antennae always missing, these *Pa-slp*^{RNAi} first nymphs also had deletions of maxillae and either partial or total loss of the labium as well. Two of the Class II nymphs were also missing either one T2 leg or had one severely reduced and malformed T1 leg (Figs 4.11H,I). However, in none of the *Pa-slp* nymphs were dorsal thoracic fusions noted, nor were there any defects in posterior segmentation (Fig. 4.11I).

Closer examination of the head morphology in *Pa-slp*^{RNAi} first nymphs show that the results of such large deletions of the mouthparts led to a misshapen head, as can easily be seen in the *Pa-slp*^{RNAi} nymphs in Figs. 4.11D and 4.11G. Although the contribution of the antennal and gnathal segments to the dorsal head may be minimal, that of the intercalary segment is fairly significant in other species studied (Posnien and Bucher, 2010; Posnien et al., 2010). The intercalary segment gives rise to the sides of the epicranium (the gena or "cheeks"), which is missing in *Pa-slp*^{RNAi} first nymphs (Fig. 4.11D, G-H), indicating an early loss of the intercalary segment along with the antennae and mandibles. Loss of the gena and, most notably, the mandibles causes the anterior portions of the head to bend in along the sides, giving the appearance of a reduced frons, clypeus, and labrum, as well as making the dorsal epicranial plates of the vertex look expanded (Fig. 4.11D). While no posterior phenotypes were noted in *Pa-slp*^{RNAi}, embryos and nymphs show considerable defects in the anterior-most segments – mandibles and antennae, indicative of a strong and conserved requirement for *Pa-slp* for proper pregnathal head and anterior gnathal development (Andrioli et al., 2004; Choe and Brown, 2007; Grossniklaus et al., 1992).

Redundancy among Periplaneta *pair-rule genes*

Segmental stripes of *Periplaneta* pair-rule genes arise sequentially in the anterior GZ, through a yet undetermined mechanism, so each is expressed at one point or another in all developing segments. As posterior segmentation defects were limited in the *Periplaneta* RNAi phenotypes examined, then perhaps the functions of these genes are redundant during posterior growth and segmentation. If the pair-rule genes are redundant during segmentation, then knocking down their expressions in combination should increase the likelihood of creating posterior defects.

As *Pa-eve* and *Pa-run* are primary pair-rule genes in other insects, it is probable they have this role in *Periplaneta* and therefore would have the greatest effects on segmentation if both were knocked-down via double RNAi. Resultant *Pa-eve/run*^{RNAi} embryos displayed two main phenotypes. Class I *Pa-eve/run*^{RNAi} embryos (6/185; 3.2%) had long skinny bodies posterior to the head (Fig. 4.12B). Consequently, staining for *Pa-en* showed this gene to be minimally expressed in narrow stripes in each weakly defined segment (Fig. 4.12B), whereas in wild type embryos *Pa-en* expression is in wide stripes in the posterior of each segment (Fig. 4.12A). Class II embryos (60/185; 32.4%) had short, fat bodies with extensive fusions between segments and a greatly reduced GZ (Fig. 4.12C-D). Expression of *Pa-en* in these embryos shows normal development of the mandibles and irregular formation of the remaining segments and an eventual absence of stripes in the reduced GZ (Fig. 4.12D). No phenotypes were observed in hatched or unhatched first nymphs (98/98; 100% wild type). Combined with the RNAi phenotypes described above, these data indicate an overall requirement on a single gene basis for proper anterior segmentation, while in the posterior these pair-rule genes may act in a redundant manner during germ band elongation.

DISCUSSION

Periplaneta pair-rule genes affect anterior segmentation

In short germ band organisms, the anterior head, gnathal, and thoracic segments are specified in the early blastoderm in a way that is highly similar to *Drosophila*. These segments form quite rapidly compared to the sequential addition of the more posterior segments during germ band elongation and may utilize an alternate mechanism to do so (Dearden et al., 2002; Janssen et al., 2012; Minelli, 2001; Pechmann et al., 2009). In long germ band organisms, regulation of the pair-rule genes is a key step as they signal a change from an unsegmented to a segmented embryo. The 'classic' pair-rule type of expression is in alternating segments. In short germ band organisms, this pattern of expression is less common, except for the pair-rule genes of *Tribolium* and some notable exceptions in the anterior for *Gryllus-eve*, *Schistocera-pby1*, and *Tetranychus-pby* (Choe et al., 2006; Davis et al., 2001; Dearden et al., 2002; Mito et al., 2007). Recently, Janssen et al. (2012) showed that most of the pair-rule gene homologues in *Glomeris* are expressed in a pair-rule manner in the anterior segments during the blastoderm and post-blastoderm stages.

In *Periplaneta*, the pair-rule genes do not show a pair-rule type of expression in the anterior during post-blastoderm development (Fig. 4.13A). While it is still possible that the earliest blastoderm expressions are in a pair-rule-like manner, analysis at this stage is not currently possible in this system, as eggs are laid in an ootheca that cannot be opened at such an early stage without destroying them. However, several observations point to the fact that these genes are, indeed, expressed in a segmental fashion. First, the pattern of expression of *Pa-pby* is seen to emerge in a sequential and segmental order in an anteroposterior direction, which likely reflects the expression of the genes immediately preceding and regulating its expression (i.e. *Pa-run*). Second, in other short germ band

insects, the pair-rule pattern of expression is noted for *Gryllus-eve* (Mito et al., 2007), *Schistocerca-pby* (Davis et al., 2001), and all *Tribolium* pair-rule homologues (Choe et al., 2006) during the post-blastoderm stages, not restricted to the blastoderm (see also Table 4.2). Therefore, it is reasonable to assume one would expect to see a pair-rule pattern of expression for *Periplaneta* pair-rule genes at comparable stages if it existed. Finally, RNAi did not result in a "pair-rule" phenotype in the anterior segments for any of the *Periplaneta* pair-rule genes, where they were observed in the anterior for both *Tribolium* and *Gryllus* embryos subjected to RNAi (Choe et al., 2006; Mito et al., 2007). Support for this argument comes from studies in *Oncopeltus*, where *Of-eve* is also expressed segmentally in the anterior and RNAi did not result in a 'pair-rule' phenotype (Liu and Kaufman, 2005a).

However, there does appear to be some regional differences in the effects of the pair-rule genes in *Periplaneta*. *Pa-eve* and *Pa-run* both show extensive fusion among the gnathal (mx/lb) and thoracic (T1/T2, T2/T3) segments in RNAi embryos and nymphs. While both affected the maxilla, deletions of this segment were more common in *Pa-run*^{RNAi} where fusions between maxilla and labium were increased in *Pa-eve*^{RNAi}. Similarly, while both RNAi treatments caused fusions to occur between T1 and T2, only *Pa-eve*^{RNAi} embryos and nymphs showed fusion between T2/T3 or among all three thoracic segments, sometimes also including the gnathal segments. However, in both *Pa-eve*^{RNAi} and *Pa-run*^{RNAi} the mandibles remained unaltered, even in those embryos showing very strong fusion in the more posterior gnathal and thoracic segments. Effects on the mandibles are mainly observed in *Pa-slp*^{RNAi}.

Pa-slp^{RNAi} embryos and nymphs often show large deletions of the head segments, from antennae to mandibles and up to the labium and thorax. This phenotype is also noted in *Drosophila* and *Tribolium* larvae when *slp* expression is depleted (Choe and Brown,

2007; Grossniklaus et al., 1994). In these insects, *slp* expression arises in the anterior head before, and independent of, its future expression as a secondary pair-rule gene (Andrioli et al., 2004; Choe and Brown, 2007; Fujioka and Jaynes, 2012; Grossniklaus et al., 1992). Formation of the mandibles appears to be a special case in many organisms studied, as it lies in a region of overlap between two defining head segmentation processes. It is the most anterior segment affected by the pair-rule genes and is the posterior-most segment specified by the "head gap" genes, such as *orthodenticle*, *buttonhead*, and *empty spiracles* (Andrioli et al., 2004). Orthologues of these head gap genes remain to be isolated and examined in *Periplaneta*, but it would be likely that they exist based on the high rate of conservation in bilaterians (Ang et al., 1996; Mercier et al., 1995; Pannese et al., 1995).

The overlapping expressions of *Dm-slp* and the other head gap genes work to differentially regulate the formation of pregnathal head segments, while the remaining mouthparts and body are patterned by the pair-rule genes (Grossniklaus et al., 1994). *Dm-slp1* works with the Groucho co-repressor in the pre-gnathal segments in a gap-like manner restricting the expression of the other pair-rule genes posterior to the mandibles; this repression includes the primary pair-rule genes *Dm-eve* and *Dm-run* which regulate the secondary expression of *Dm-slp* in more posterior segments (Andrioli et al., 2004; Cadigan et al., 1994). A putative Groucho binding domain has been identified in the cockroach *slp* homologue indicating a potential interaction with *Pa-slp* and, thus, a conserved repressor function for this gene in head patterning. Perhaps *Pa-slp* plays a role in repressing *Pa-eve* and *Pa-run* function in the mandibles, which could explain why loss of either *Pa-eve* or *Pa-run* expression never result in a mandibular phenotype.

Altogether, the data on *Periplaneta* pair-rule genes studied here show that while they are all expressed segmentally in the anterior, there is some regionalisation to their effects. *Pa-slp* functions primarily in the antennae and mandibles, while *Pa-run* and *Pa-*

eve mainly affect the posterior gnathal and thoracic segments, with *Pa-run* having a stronger effect on the maxilla and *Pa-eve* affecting more of the thoracic segments. Conversely, no overt segmentation phenotypes were observed in the anterior of *Pa-h*^{RNAi} embryos, although there are minor defects in *Pa-en* stripes in this location (Pueyo et al., 2008); however, *Pa-h* plays a much larger role in posterior patterning. Functions for *Pa-pby* in anterior segmentation remain to be fully elucidated and *Periplaneta* homologues for the other pair-rule genes, such as *odd-skipped*, have yet to be isolated and analysed.

Periplaneta pair-rule genes and posterior patterning

After the germ anlage condenses and the anterior segments begin to take form at the post-blastoderm stage, the remaining posterior segments will emerge during germ band elongation from the GZ. Pair-rule gene expression in the GZ can vary depending on the organism in question, but are usually segmental or pair-rule. The *Periplaneta* pair-rule gene homologues are expressed segmentally in the GZ and can be separated into two groups based on when and where their expression initially arises. *Pa-eve*, *Pa-run*, and *Pa-h* (Pueyo et al., 2008) are each expressed in a broad posterior domain from which stripes of expression emerge in the anterior GZ; whereas *Pa-pby* and *Pa-slp* are not expressed in the posterior and only form stripes of expression in the anterior GZ (Fig. 4.13B-C). Thus, it is now possible to define *Pa-eve*, *Pa-run*, and *Pa-h* as primary pair-rule genes, being expressed first and more posterior, and *Pa-pby* and *Pa-slp* secondary, as they arise later and only in the anterior; a similar spatiotemporal hierarchy has been suggested for spider pair-rule genes by Damen et al., (2005). These patterns hold true in the anterior post-blastoderm, as well, where *Pa-eve*, *Pa-run*, and *Pa-h* are expressed earlier than *Pa-pby* and *Pa-slp* (Fig. 4.13A). The genetic interaction between *Pa-run* and *Pa-pby* confirms this relationship, whereby *Pa-pby* is negatively regulated by *Pa-run* in such a way that stripes

can only form in the anterior GZ when the expression of *Pa-run* begins to fade (Fig. 4.5B).

This designation of primary and secondary has been well established in *Drosophila* where the primary pair-rule genes (*eve*, *run*, *h*) are expressed first and, in turn, regulate the downstream secondary pair-rule genes (*prd*, *slp*, *odd*, *ftz*), thus forming a "mini-hierarchy" within the *Drosophila* segmentation cascade. Applying the spatiotemporal relationship described above for *Periplaneta* to the literature reveals a strong conservation of this pair-rule gene mini-hierarchy within the Arthropoda (Table 4.2 and references therein). Based largely on analysis of expression data in the posterior GZ, several patterns begin to emerge. First, the ancestral expression patterns of the pair-rule genes in the posterior GZ was most likely segmental in nature, as this is the most common pattern observed in phylogenetically basal organisms. While a few examples of pair-rule-like expression have been noted, such as several *Strigamia* genes, this pattern is mainly and almost exclusively observed in the holometabolous insects. Second, whether expressed in a segmental or pair-rule manner, nearly all pair-rule genes studied can be classified as either primary or secondary. With this in mind, a third pattern is noted in which *eve* and *run* are always regarded as primary pair-rule genes (blue box; Table 4.2), while *prd/pby* and *slp* are secondary (green box; Table 4.2).

The general classification of pair-rule genes as primary and secondary is conserved among the arthropods; however, the posterior expression patterns can be highly variable (i.e. *eve*; Fig. 4.1). In higher insects, these genes are expressed with double-segment periodicity, while in more basal insects and arthropods the pair-rule genes tend to be expressed segmentally. As most of these inferences rely on expression data only, functional analysis and examination of the interplay between these primary and secondary pair-rule genes remains to be done in order to reveal the true nature of these potential

interactions and the presumed conserved function of these genes in regulating the segment polarity genes and segment formation.

Evolution of pair-rule patterning

RNAi for *Periplaneta* pair-rule genes has a more profound effect on the anterior segments compared to the posterior. Phenotypes commonly involved fusion or deletion of the gnathal and thoracic segments, those formed in the early blastoderm, while effects in the posterior were limited. These results could indicate that 1) RNAi is less effective in the posterior of the cockroach compared to the anterior, or 2) *Periplaneta* pair-rule genes are highly redundant during posterior growth and segmentation. The first option can be ruled out as posterior effects are observed in RNAi for several other genes (see Chapters III and V) (Pueyo et al., 2008). Therefore, the second alternative may be more explanatory to the effects observed in *Periplaneta* pair-rule gene RNAi, which is supported by the increased prevalence of posterior effects upon the loss of both *Pa-eve* and *Pa-run* in double knockdown RNAi. Additionally, these observations may also be partially explained by the differences in the mechanisms involved in segment formation in these two different regions.

Posterior growth and segment formation in the cockroach is under control of Notch-signalling, while the anterior segments form in the blastoderm in a N-free environment (Chapter III). In *Periplaneta*, N-signalling regulates the dynamic expression of the pair-rule gene *Pa-h*, which goes on to regulate *Pa-en* expression and segment formation (Pueyo et al., 2008). N-mediated segmentation is conserved in all segmented phyla, including annelids and vertebrates, and has been suggested to be ancestral to the bilaterians (Pueyo et al., 2008; Stollewerk et al., 2003). However, *hairy* may present a unique case as a pair-rule gene functioning in bilaterian segmentation. While other pair-

rule gene homologues have been found outside of the arthropods they do not usually have a function in segmentation/somitogenesis, instead playing a role in neurogenesis (Bastian and Gruss, 1990; Inoue et al., 2002a; Ruiz i Altaba, 1990; Song et al., 2002). This neural function remains conserved in the arthropods, indicating an ancient origin for these genes and suggesting that the role of these genes in segmentation must have evolved after the protostome-deuterostome split. Even within the protostomes, however, the pair-rule genes have only been found to function in arthropod segmentation, with the possible exception of the polychaete worm *Platynereis dumerilii* (de Rosa et al., 2005; Seaver et al., 2012). Therefore, pair-rule gene expression and function in the arthropods may be a derived state having been co-opted into the segmentation mechanisms.

These results are interesting in light of the fact that in short germ organisms the anterior-most segments form in a syncytial blastoderm, reminiscent of long germ band segmentation. This Notch-independent manner of segmentation is likely the precursor to what exists in higher insects today. As new methods were gained to speed up embryonic development, addition of segments from a posterior growth zone became less important as more segments were being formed in the syncytial blastoderm. Along with this, the requirement for Notch became unnecessary as signalling now occurred in a cell-free environment which allowed for the gap genes to take over control of pair-rule gene expression and segmentation (Damen, 2007; Peel and Akam, 2003).

Expression in the double-segment manner allows for faster development, especially within the syncytial blastoderm of highly derived insects, such as *Drosophila*. This all-at-once approach to segment patterning may have its roots in short germ band arthropods where the anterior segments are pre-patterned in the early blastoderm (Davis and Patel, 2002; Liu and Kaufman, 2005b). Several studies have shown initial pair-rule patterning at the blastoderm/post-blastoderm stages, when the anterior segments are

specified, in basal insects and other arthropods (Davis et al., 2001; Dearden et al., 2002; Janssen et al., 2012; Mito et al., 2007). These data suggest a relationship between the anterior segmentation mechanisms in short germ organisms and the whole body segmentation in long germ band organisms, both occurring within an early blastoderm. Perhaps this early patterning of more basal arthropods was expanded during the transition from short to long germ band as more and more segments became pre-patterned in the blastoderm, until all have been incorporated into this stage (Damen, 2007; Peel and Akam, 2003). This change may have occurred as the pair-rule genes came under the control of the gap genes by gaining cis-regulatory gap response elements (Peel and Akam, 2003). Some studies suggest a conserved function of the gap genes in regulating pair-rule genes (Bucher and Klingler, 2004; Cerny et al., 2005; Cerny et al., 2008; Liu and Patel, 2010; Mito et al., 2006; Mito et al., 2005; Schwager et al., 2009); however, gap gene functions in short germ band organisms remain unclear. Other reports suggest that the pair-rule genes themselves may act as gap genes that even regulate the expression of other, canonical, gap genes (Liu and Kaufman, 2005a; Mito et al., 2007).

It is imperative to study more phylogenetically basal insects and arthropods, to further elucidate and draw stronger conclusions on putative ancestral gene functions. Perhaps the best candidates for studying the evolution of developmental mechanisms are those organisms lying at the border of major lineages. The cockroach is a prime example of an insect with the power of elucidating the functions and evolution of segmentation genes, as *Periplaneta* represents a link between the basal arthropods (Myriapoda and Chelicerata) and the derived insects. This is the first study in which a large sample of the pair-rule genes have been analysed, both expression and function, in a basal insect species.

It appears as though the function of the pair-rule genes in segmentation may be restricted to the arthropods. Through the course of evolution, the pair-rule genes seem to

have gained a new function in arthropod body segmentation; however, without proper functional analysis in more ancestral arthropods, it is hard to determine the true nature of these genes pertaining to segmentation. The results presented here indicate an ancestral segmental expression of the pair-rule genes in the posterior, while leaving the question remaining as to whether the ancestral anterior expression is pair-rule. My data provide additional evidence that a pair-rule hierarchy existed in the arthropod ancestor and further demonstrate the highly flexible nature of the pair-rule genes during arthropod segmentation.

Figure 4.1: Variability in pair-rule gene expression in arthropods. The expression of the pair-rule gene *even-skipped* is one example of the highly variable nature of pair-rule gene expression. Expression may be in the 'classic' double-segment pattern, as in (1) the highly derived, long germ band insect *Drosophila melanogaster* (Diptera) (Macdonald et al., 1986) and in (2) the short germ band, holometabolous insect *Tribolium castaneum* (Coleoptera – Patel et al., 1994). In most short germ band arthropods, such as (3) the hemimetabolous insect *Oncopeltus fasciatus* (Hemiptera – Liu and Kaufman, 2005a) and (4) the crustacean *Artemia franciscana* (Copf et al., 2003), *eve* is expressed broadly in the posterior growth zone (GZ) from which new stripes will emerge with a single-segment periodicity. In the Orthoptera, two different patterns emerge. In the cricket *Gryllus bimaculatus* (5) (Mito et al., 2007), stripes of *eve* emerge in both a single- and a double-segment manner, while in the grasshopper *Schistocerca americana* (6) (Patel et al., 1992) *eve* is expressed broadly in the posterior GZ, but no stripes form in the anterior. Phylogenetic tree reproduced from Trautwein et al. (2012).

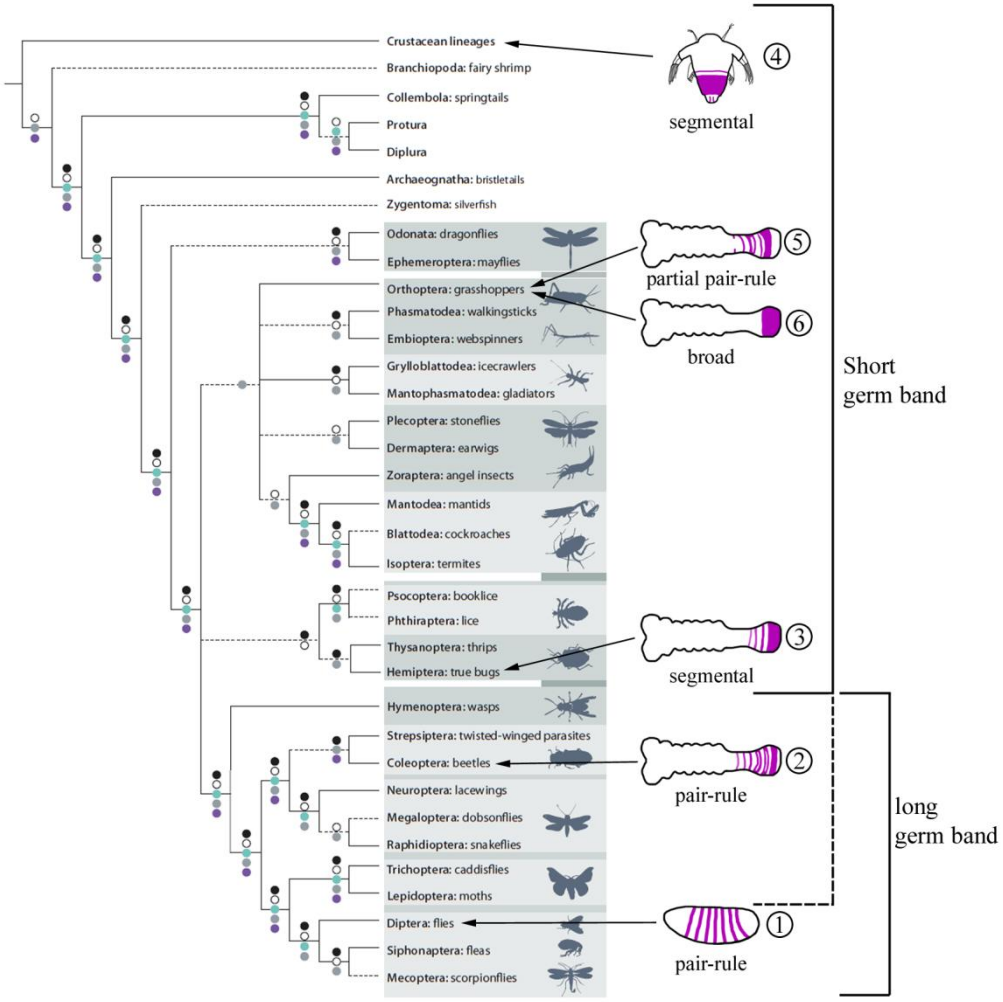

Crustacean lineages
Branchiopoda: fairy shrimp
Collembola: springtails
Protura
Diplura
Archaeognatha: bristletails
Zygentoma: silverfish
Odonata: dragonflies
Ephemeroptera: mayflies
Orthoptera: grasshoppers
Phasmatodea: walkingsticks
Embioptera: webspinners
Grylloblattodea: icecrawlers
Mantophasmatodea: gladiators
Plecoptera: stoneflies
Dermaptera: earwigs
Zoraptera: angel insects
Mantodea: mantids
Blattodea: cockroaches
Isoptera: termites
Psocoptera: booklice
Phthiraptera: lice
Thysanoptera: thrips
Hemiptera: true bugs
Hymenoptera: wasps
Strepsiptera: twisted-winged parasites
Coleoptera: beetles
Neuroptera: lacewings
Megaloptera: dobsonflies
Raphidioptera: snakeflies
Trichoptera: caddisflies
Lepidoptera: moths
Diptera: flies
Siphonaptera: fleas
Mecoptera: scorpionflies
④
segmental
⑤
partial pair-rule
⑥
broad
Short
germ band
③
segmental
②
pair-rule
long
germ band
①
pair-rule

Figure 4.1

Figure 4.2: Wild type expression of *Periplaneta even-skipped*. *Pa-eve* mRNA (greyscale) and protein (green) during post-blastoderm (A-B), early germ band (C-D) and late germ band (E-H) elongation. At stage 4 post-blastoderm, *Pa-eve* mRNA (A) and protein (B) are similarly expressed as five stripes (arrowheads) from the mandibles to the second thoracic segment, with the presumptive mandibular stripe showing the weakest expression. (C) At early germ band elongation, stage 7, the initial anterior segmental stripes of *Pa-eve* have faded down to T2 and expression is now detected weakly in a broad domain in the posterior GZ (bracket) from which new segmental stripes emerge (arrowhead). (D) *Pa*-Eve protein shows a perdurance of expression in the previously established stripes in the anterior portion of each forming segment (outlined T2), and new expression develops in the posterior GZ (bracket) and in a stripe in the anterior (white arrowhead). (E-H') During late germ band elongation, *Pa-eve* expression is dynamic, with a broad domain of expression in the posterior GZ (brackets) from which new stripes emerge (black arrowheads, E-H) that fade as they move anteriorly (open arrowheads, G-H). Up to three stripes are visible at any given time, as in F. a – anterior; A1, A3, A4, A5 – first, third, fourth, and fifth abdominal segment, respectively; mn – mandibles; p – posterior; T2 – second thoracic segment.

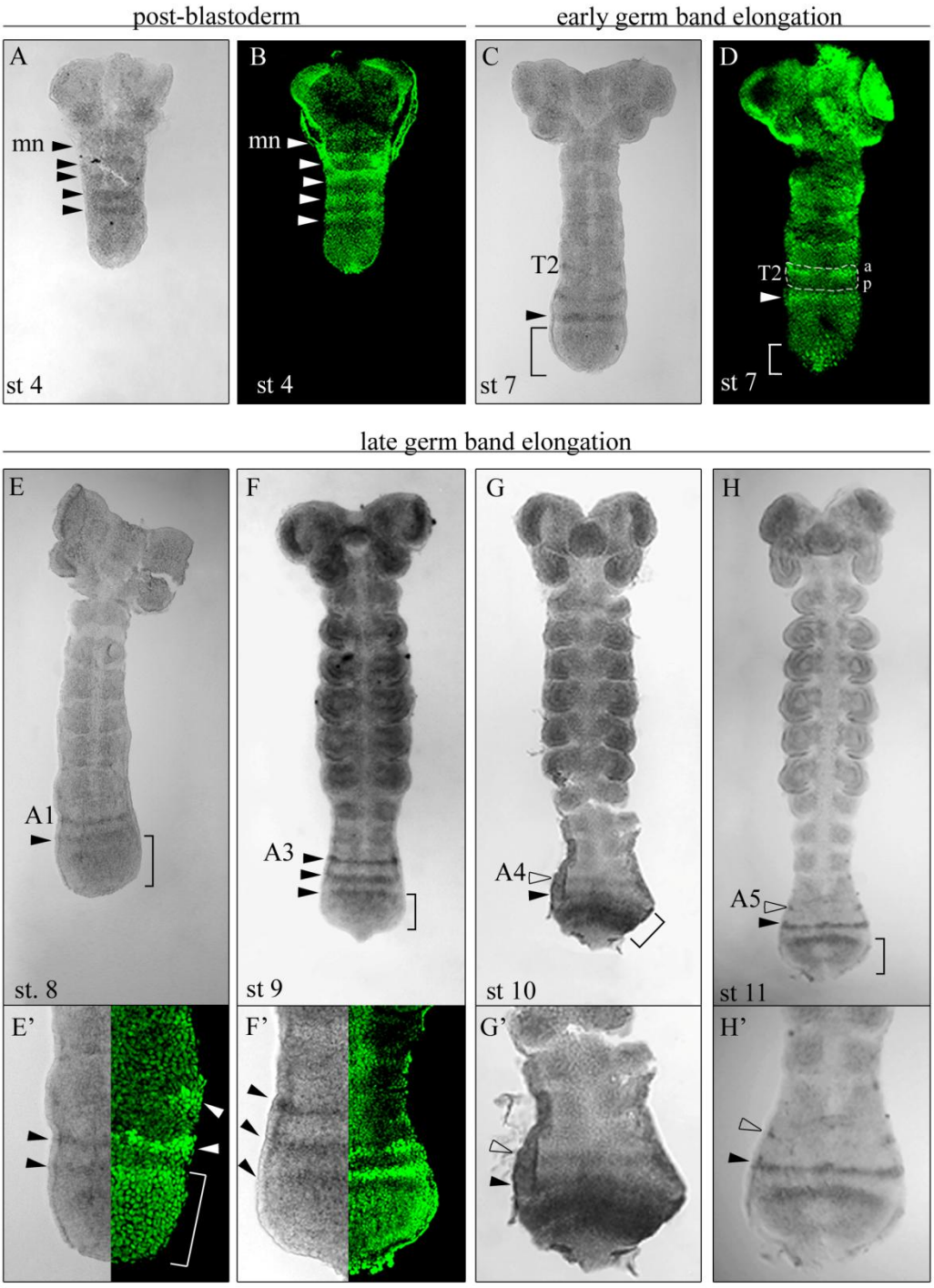

post-blastoderm
early germ band elongation
A
mn
st 4
B
mn
st 4
C
T2
st 7
D
T2
a
p
st 7
late germ band elongation
E
A1
st. 8
F
A3
st 9
G
A4
st 10
H
A5
st 11
E'
F'
G'
H'

Figure 4.2

Figure 4.3: Wild type expression of *Periplaneta runt* and co-expression with *Pa-engrailed*. Expression of *Pa-run* mRNA at post-blastoderm (A-B), early germ band (C-D), and late germ band elongation (E-J). (A) *Pa-run* is already expressed as five stripes at stage 4 post-blastoderm (arrowheads), with the weakest expression in the mandibles, and in the developing head lobes (arrow). (B) At stage 5 post-blastoderm, the stripes of *Pa-run* are wider and a new stripe of expression is beginning to form in the posterior (arrowhead); the anterior head expression remains and will continue to be expressed throughout embryogenesis (black arrow). (C) At early germ ban elongation, stage 6, the anterior stripes of *Pa-run* expression begin to fade from the ventral midline (arrowhead), while a new stripe has formed at A1. At this stage, the segmental furrows begin to form, showing that *Pa-run* expression is in the anterior half of each developing segment (outline). (D) By stage 7, most of the previously formed anterior stripes have faded and a new stripe is resolving (arrowhead) from the broad posterior GZ domain (bracket). (E-J) New stripes of *Pa-run* are added sequentially (arrowheads) from a broad, dynamic posterior domain that is best observed via FISH (H-J). As the stripes move anteriorly they begin to fade from the midline and eventually disappear before the new segment forms (* in H-J). In the forming segments, two ventral dots appear on either side of the midline (open arrowhead, G, J) that become clusters of expression in the developing CNS (open arrow, G). (K-L) Double FISH for *Pa-run* (red) and *Pa-en* (green). At both post-blastoderm (K) and germ band elongation (L), the stripes of *Pa-en* arise only after the expression of *Pa-run* begins to fade and their expressions do not overlap. A1, A2, A3, A5, A6 – first, second, third, fifth, and sixth abdominal segment, respectively; mn – mandible; T2, T3 – second and third thoracic segment.

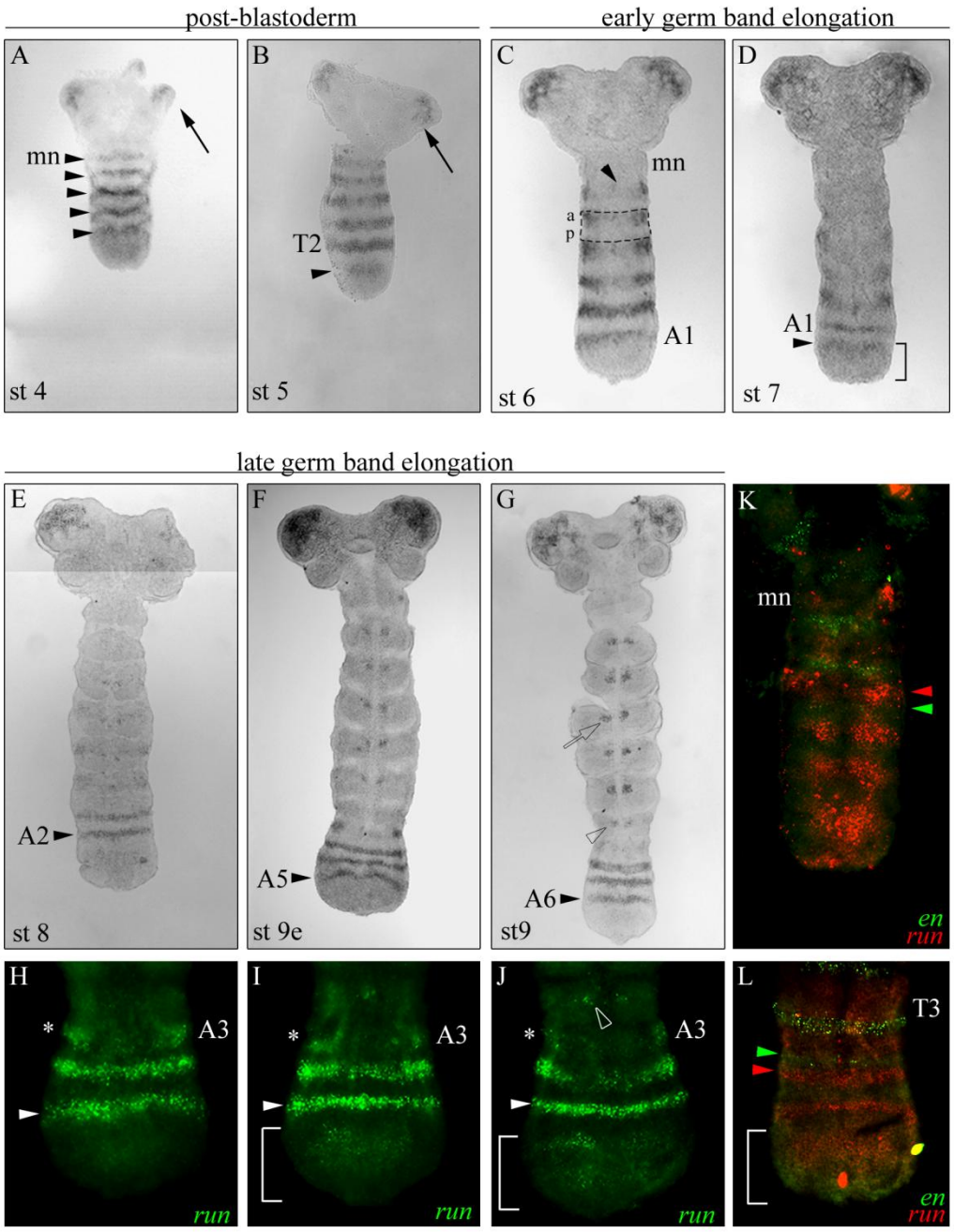

post-blastoderm
early germ band elongation
A
mn
st 4
B
T2
st 5
C
mn
a
p
A1
st 6
D
A1
st 7
late germ band elongation
E
A2
st 8
F
A5
st 9e
G
A6
st9
K
mn
en
run
H
*
A3
run
I
*
A3
run
J
*
A3
run
L
T3
en
run

Figure 4.3

Figure 4.4: Wild type expression of *Periplaneta pair-berry* and effects of *Pa-pby*^{RNAi}. *Pa-pby* expression at post-blastoderm (A-D) and germ band elongation (E-I). At early post-blastoderm, stage 3, *Pa-pby* is only expressed in two gnathal segments (black arrowheads). In addition, expression is detected in the antennal (arrow) and head lobe (open arrowhead) primordial, which remains throughout development. (B-D) Segmental stripes of expression develop sequentially from the anterior until the full complement of five post-blastoderm stripes have formed and, at stage 6, a new stripe forms in the presumptive intercalary segment (arrow; D). (E-I) New stripes of *Pa-pby* expression continue to arise sequentially in the anterior GZ; no expression is detected in the posterior GZ (H). (I) The stripes of *Pa-pby* remain in the posterior region of all developed segments throughout embryonic development. (J) *Pa-pby*^{RNAi} can result in extensive fusion of the posterior thoracic and abdominal segments (bracket). a – anterior; ant – antennae; A1, A2, A3, A5 – first, second, third, fifth abdominal segment, respectively; ic – intercalary segment; lb – labium; mn – mandibles; T1, T2, T3 – first, second, third thoracic segment, respectively; p – posterior.

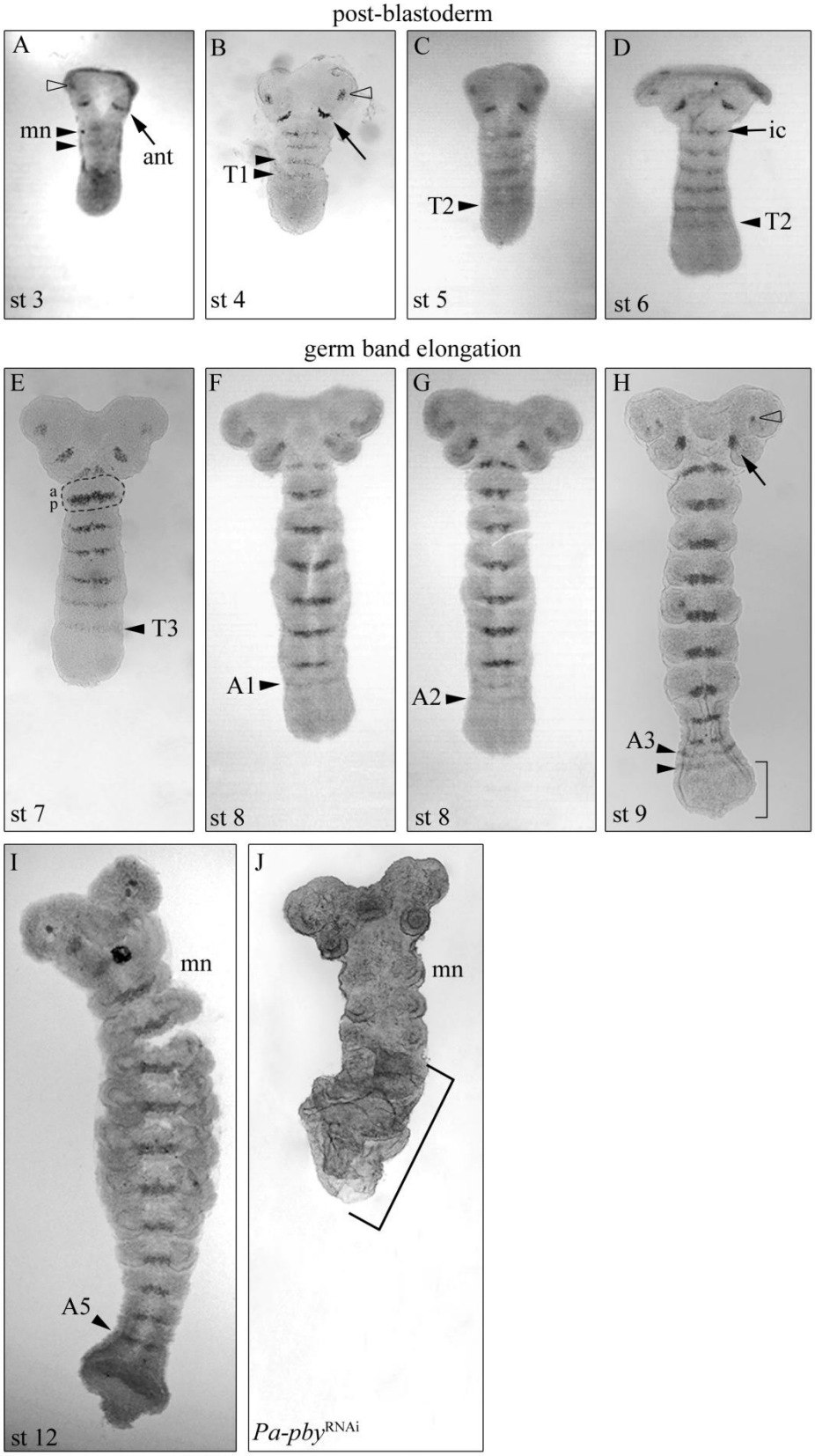

post-blastoderm
A
mn
ant
st 3
B
T1
st 4
C
T2
st 5
D
ic
T2
st 6
germ band elongation
E
a
p
T3
st 7
F
A1
st 8
G
A2
st 8
H
A3
st 9
I
mn
A5
st 12
J
mn
Pa-pby^RNAi

Figure 4.4

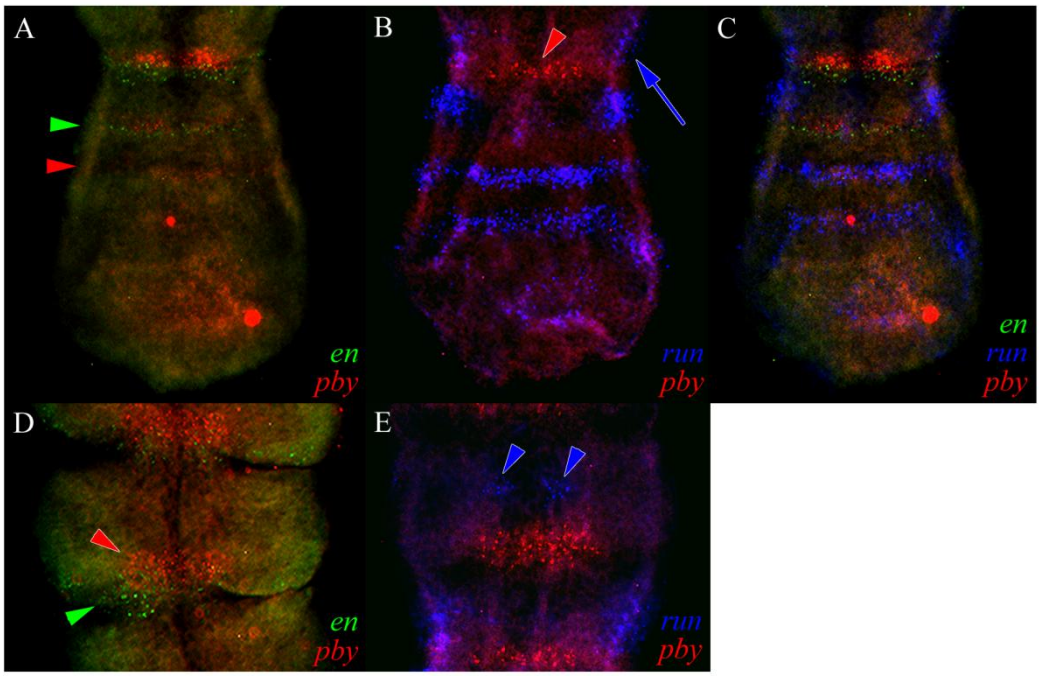


Figure 4.5: Co-expression of *Pa-pby* with *Pa-en* and *Pa-run* during germ band elongation in wild type embryos. (A) Stage 9 embryo showing expression of *Pa-pby* (red) in the anterior GZ developing before that of *Pa-en* (green). Stripes of expression partially overlap with that of *Pa-pby* lying anterior to *Pa-en*. (B) Stage 9 embryo showing co-expression of *Pa-pby* and *Pa-run* (blue). As the stripes of *Pa-run* move anteriorly they begin to fade from the midline (blue arrow) and at this point a stripe of *Pa-pby* expression begins to emerge (red arrowhead), posterior to *Pa-run*, expanding out from the midline to the lateral edges. (C) An overlay of images in A and B show that the expression of *Pa-pby* and *Pa-en* develop only after that of *Pa-run* had faded. (D) Expression of both *Pa-pby* and *Pa-en* remain in the developing segments, partially overlapping with *Pa-pby* anterior to *Pa-en*.

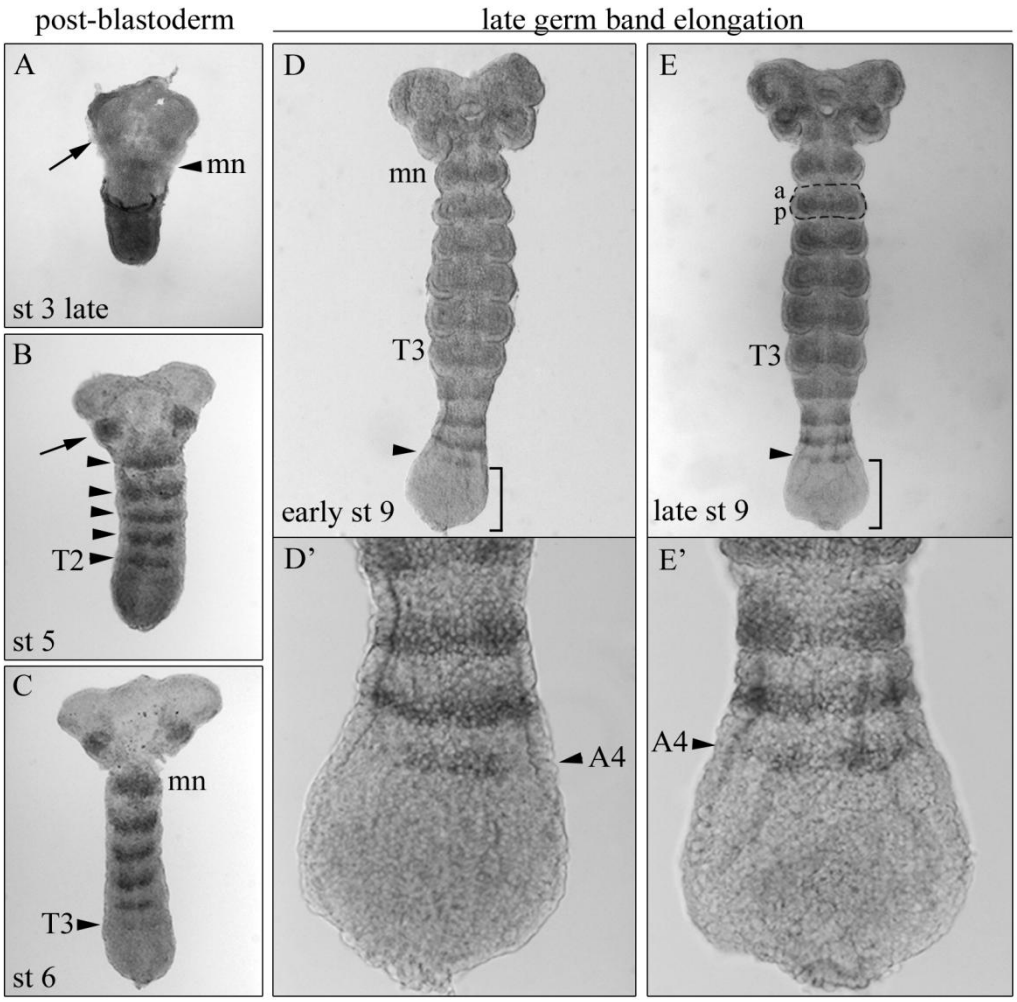


Figure 4.6: Wild type expression of *Pa-sloppy-paired* at post-blastoderm (A-C) and late germ band elongation (D-E). (A) At late stage 3, *Pa-slp* is expressed in the presumptive antennae (arrow) and mandibles (arrowhead). (B) By stage 5, stripes corresponding to the three gnathal and first two thoracic segments can now be detected (arrowheads), as well as broad expression in the antennae (arrow). (C) At late stage 6, a new stripe is beginning to form at T3. (D-E) During late germ band elongation, new stripes of *Pa-slp* form sequentially in the anterior GZ, starting at the midline (arrow, D-D') and becoming stronger and extending laterally as they move further anterior (arrowhead, E-E'). 2-3 stripes can be detected in the anterior GZ at one time and expression remains in the posterior half of all developed segments (arrow, E); no expression is observed in the posterior (brackets, D-E). A4 – fourth abdominal segment; mn – mandibles; T3 – third thoracic segment.

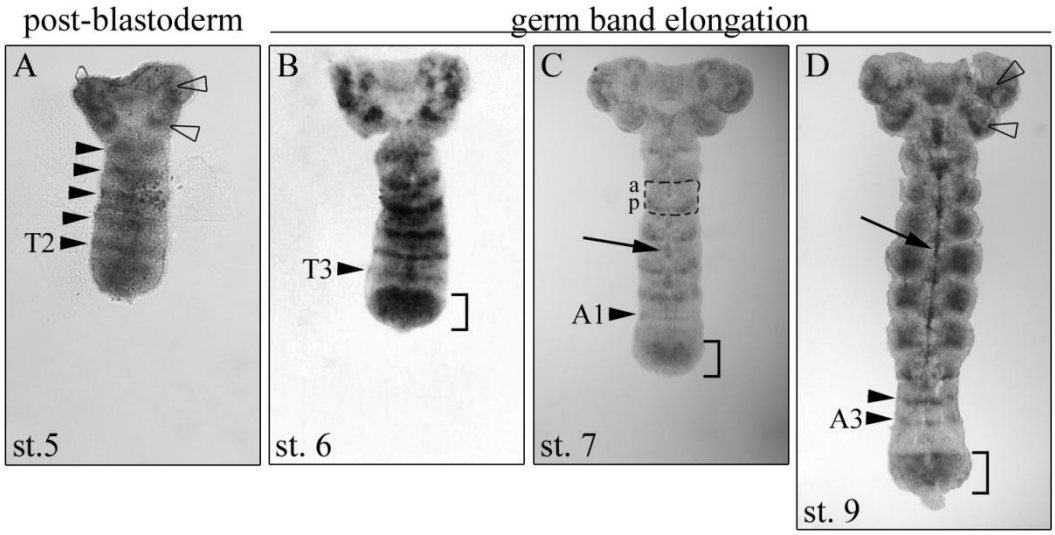


Figure 4.7: Wild type expression of *Pa-hairy*. (A) At post-blastoderm stage 5, *Pa-h* is expressed in five stripes (black arrowheads) corresponding to the three gnathal and first two thoracic segments. Expression is also detected in the head lobes and antennal primordia (open arrowheads). (B-D) *Pa-h* expression is dynamic during germ band elongation. (B) By stage 6 early germ band elongation, *Pa-h* expression is in a broad domain in the posterior GZ (bracket) and a new stripe has formed in the anterior GZ correlating with T3 (black arrowhead). (C) A new stripe has been added at A1 (black arrowhead) from the broad posterior domain (bracket). As segment furrows form, *Pa-h* expression is noted in the posterior half of each segment (outlined) and new expression emerges in the ventral CNS (arrow). (D) At late germ band elongation, the anterior stripes of expression have faded and the CNS expression becomes stronger (arrow). New stripes of expression continue to emerge from the broad posterior domain (black arrowheads and bracket, respectively) and the head and antennal expression remain strong (open arrowheads). A1, A3 – first and third abdominal segment, respectively; T2, T3 – second and third thoracic segment, respectively.

	Embryo			Nymph		
Gene	**Class I**	**Class II**	**wt**	**Class I**	**Class II**	**wt**
Pa-eve	60 (9.0%)	179 (26.8%)	428 (64.2%)	14 (3.5%)	n/a	385 (96.5%)
Pa-run	75 (22.5%)	15 (4.5%)	244 (73.1%)	16 (11.9%)	n/a	119 (88.1%)
Pa-prd	1 (0.75%)	16 (12.0%)	117 (88.0%)	n/a	n/a	182 (99.5%)
Pa-slp	12 (7.1%)	54 (31.8%)	104 (61.2%)	209 (59.9%)	20 (5.7%)	120 (34.4%)
Pa-eve/run	6 (3.2%)	60 (32.4%)	119 (64.3%)	0 (0.0%)	0 (0.0%)	98 (100%)
Control (H2O)	0 (0.0%)	0 (0.0%)	46 (100%)	0 (0.0%)	0 (0.0%)	253 (99.6%)

Table 4.1: Phenotypic series of pair-rule RNAi affected *Periplaneta* embryos and nymphs. Pair-rule RNAi often resulted in a range of hypomorphic embryo phenotypes that could be separated into two classes. In general, Class I phenotypes were stronger than those of the moderate/weak Class II. Most nymphal RNAi phenotypes could only be classified into one Class, which simply separated wild type first nymphs from all others showing any phenotype, usually segment fusions. *Pa-slp*^{RNAi} is the one exception, and here Class I are slightly weaker than the rare Class II first nymphs. For a detailed description of specific RNAi phenotypes for each pair-rule gene, please refer to the appropriate sections of main text.

Figure 4.8: *Pa-eve*^{RNAi} effects anterior segment patterning. (A) *Pa-en* expression in a stage 8 wild type embryo. (B-C) Stage 8, Class I *Pa-eve*^{RNAi} embryos displaying a skinny body phenotype, with a normal GZ. *Pa-en* expression is either weak and thin (arrowheads, B) or may be completely absent (C). *Pa-eve*^{RNAi} does not affect formation of the mandibles (arrow, C). (D) Stage 9 wild type, unstained embryo with mandible (mn) and thoracic segments marked with a bracket. (E) Class II *Pa-eve*^{RNAi} embryos display extensive fusion between adjacent segments, such as the maxilla/labium (*) and up to all three thoracic segments (bracket), while the mandibles remain unaffected (mn). (F) Stage 23 wild type embryo showing separation of the three thoracic segments (brackets). (G) In this stage 23, Class II *Pa-eve*^{RNAi} embryo, all thoracic segments are fused dorsally (bracket and green outline) and the ventral legs are fused at the base (red outline). (H-H') Wild type first nymph shown in dorsal (H) and lateral (H') views; maxillary palp indicated with an arrowhead. (I-I') Class II *Pa-eve*^{RNAi} first nymphs showing partial fusion of dorsal thoracic segments (bracket). In rare cases, fusion between two abdominal segments have occurred (arrow, I') and the maxillae are absent (black arrowhead), while the labium remains unaffected (open arrowhead). (J-J') Another Class II *Pa-eve*^{RNAi} embryo displaying extensive fusion of all thoracic segments both dorsally (bracket, J) and ventrally, as shown by the fused legs (arrow, M). GZ – growth zone; mn – mandibles; T1, T2, T3 – first, second, third thoracic segment, respectively.

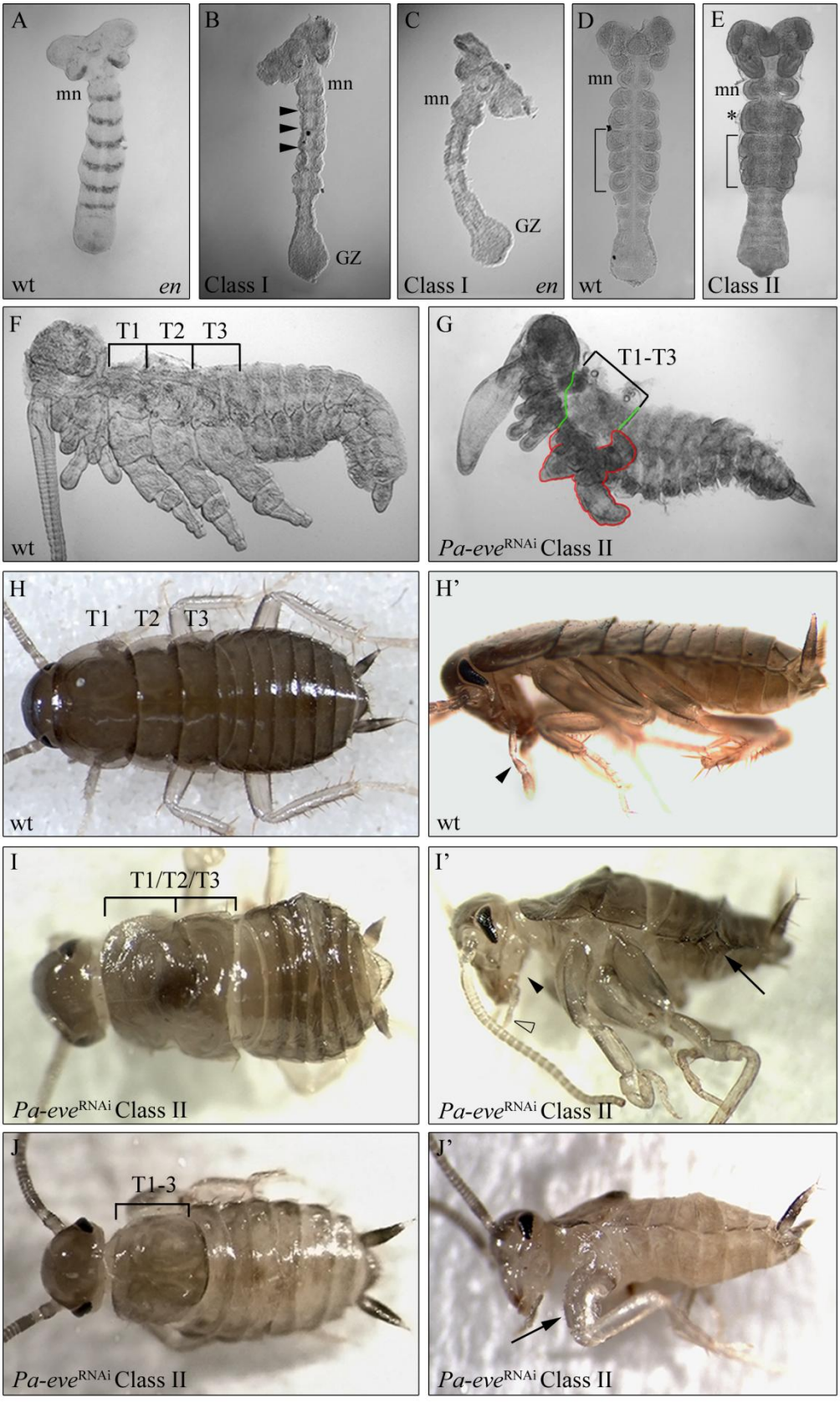

A
mn
wt
en
B
mn
GZ
Class I
C
mn
GZ
Class I
en
D
mn
wt
E
mn
*
Class II
F
T1 T2 T3
wt
G
T1-T3
Pa-eveRNAi Class II
H
T1 T2 T3
wt
H'
wt
I
T1/T2/T3
Pa-eveRNAi Class II
I'
Pa-eveRNAi Class II
J
T1-3
Pa-eveRNAi Class II
J'
Pa-eveRNAi Class II

Figure 4.8

Figure 4.9: *Pa-run*^{RNAi} effects anterior segment patterning. (A) Wild type, stage 11 embryo showing normal expression of *Pa-en*. (B-C) Unstained, stage 11 *Pa-run*^{RNAi} embryos show extensive fusion between gnathal and thoracic segments (bracket, B) and either partial (open arrowheads, C) or complete (arrowhead, C) loss of segments, commonly maxilla. In these embryos, the posterior growth zone is unaffected, but there is a pinch in the abdomen where segments are forming (*, B-C). (D) *Pa-en* expression is unaffected in the pinched posterior abdomen (open arrowheads), but is lost in the ventral area (white arrowhead) of fused segments (bracket). (E-E') In Class II *Pa-run*^{RNAi} embryos, some segments display a 'mirror-image' duplication effect (arrowheads and magnified image E'). In both Class I and Class II the mandibles remain unaffected (B-E). (G) Dorsal view of a wild type first nymph. (H) Dorsal view of a *Pa-run*^{RNAi} first nymph showing fusion of the first and second thoracic segment. (I) Ventral view of a wild type first nymph showing normal formation of the three thoracic segments and the maxillary palp (arrow). (J) Ventral view of a *Pa-run*^{RNAi} embryo in which the T1 legs and maxillary segment (arrow) did not form. GZ – growth zone; mn – mandibles; T1, T2, T3 – first, second, third thoracic segment, respectively.

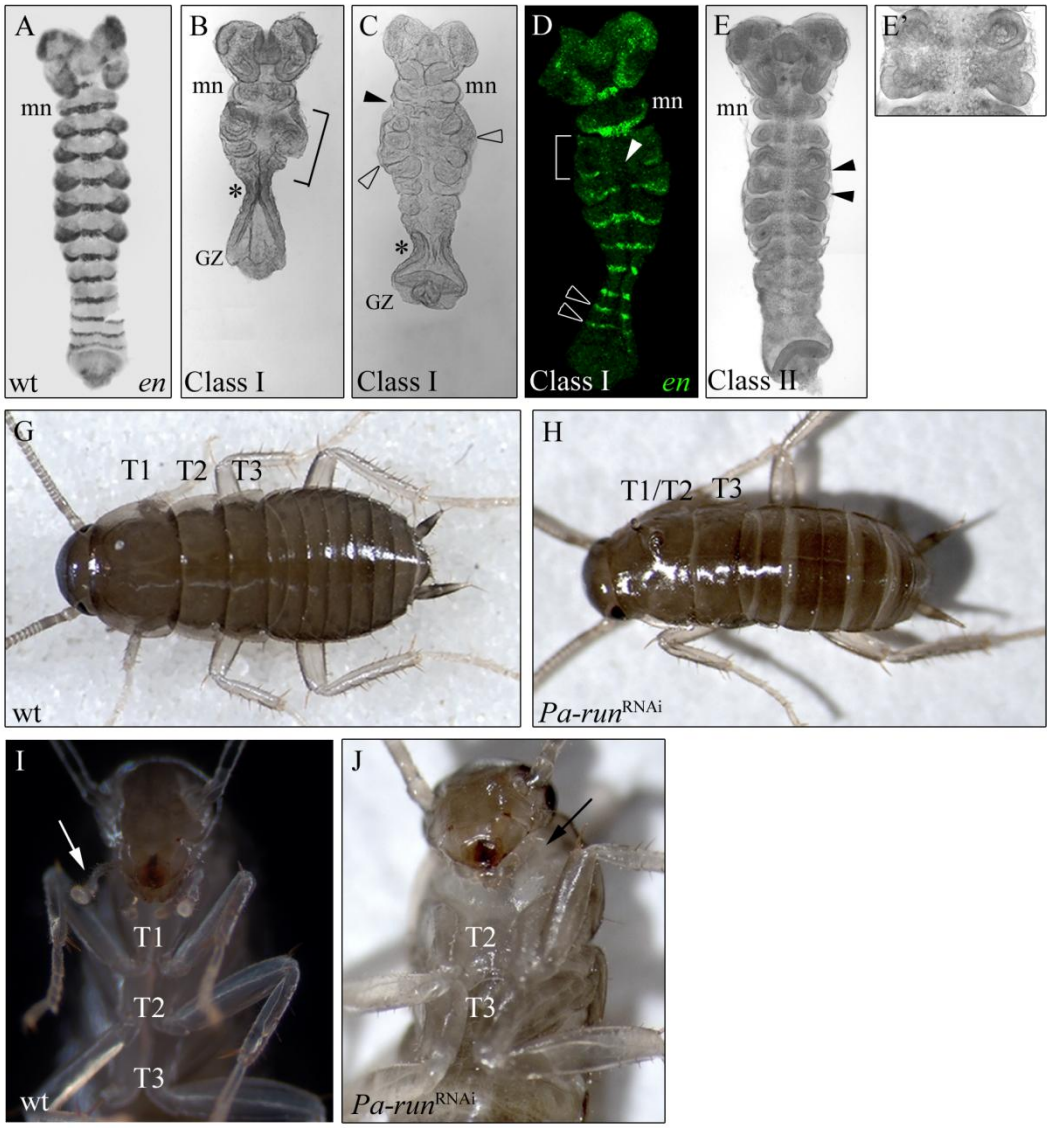

A
mn
wt
en
B
mn
GZ
Class I
C
mn
GZ
Class I
D
mn
Class I
en
E
mn
Class II
E'
G
T1
T2
T3
wt
H
T1/T2
T3
Pa-run RNAi
I
T1
T2
T3
wt
J
T2
T3
Pa-run RNAi

Figure 4.9

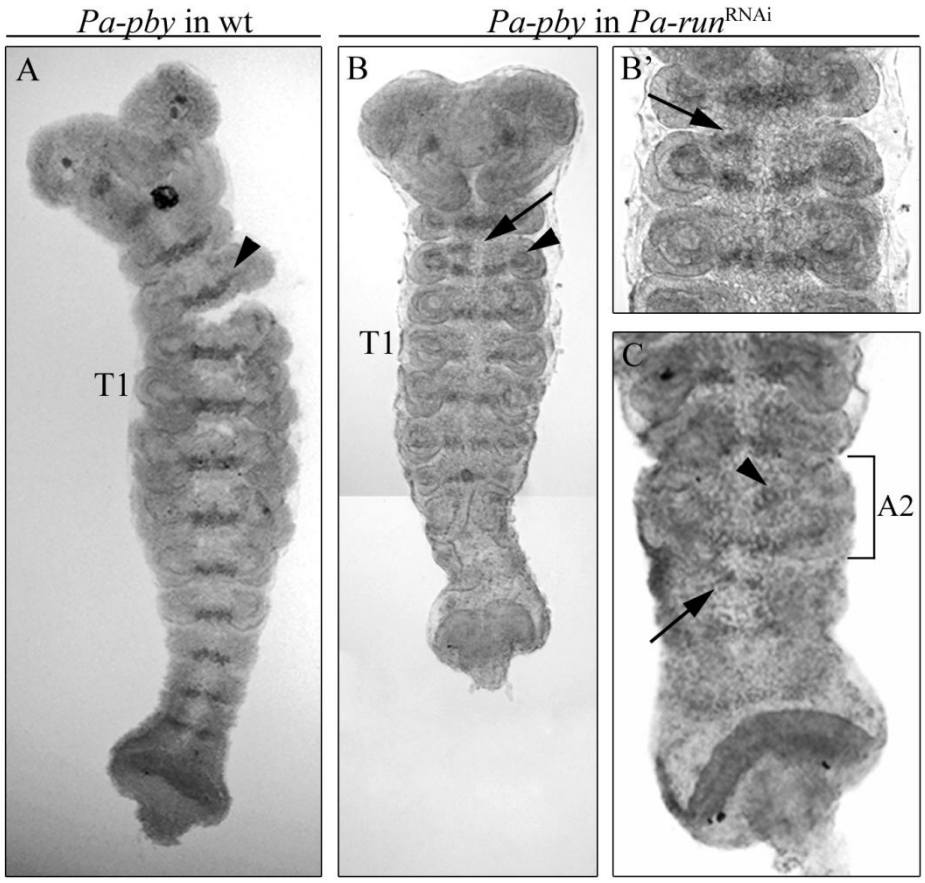


Figure 4.10: *Pa-run* regulates expression of *Pa-pby*. (A) *Pa-pby* expression in a wild type stage 12 embryo is in the posterior half of all segments (arrowhead). (B-C) In *Pa-run*^{RNAi} affected embryos at stage 12, *Pa-pby* stripe expression is in the posterior of most developed segments and is occasionally expressed ectopically in the anterior of some segmental compartments (arrow, B, and magnified image B'). (C) In the posterior abdomen of this similarly staged *Pa-run*^{RNAi} embryo, stripes of *Pa-pby* expression are misaligned (arrow) or only partially formed (arrowhead). A2 – second abdominal segment; T1 – first thoracic segment.

Figure 4.11: Effects of *Pa-slp*^{RNAi} on anterior head segmentation. (A) Unstained, stage 9 wild type embryo. (A') Magnification of box in A showing labium and T1 segments. (B-B') In *Pa-slp*^{RNAi} embryos, the mandibles do not form properly (mn, B) and some segments exhibit a 'mirror-image' duplication in the anterior (arrowheads, B'). (C) Anterior view of a dissected head from a wild type first nymph. The insect head is separated into several cuticular plates. Of particular note here are, from top to bottom: the vertex, followed by the frons, clypeus, and labrum, with the gena forming the sides or cheeks. (C') Dissected mandible from the wild type first nymph represented in C with the molar indicated with an arrow. (D) In Class I *Pa-slp*^{RNAi} affected first nymphs the head appears misshapen due to of a failure to form the mandibles, gena, and antennae. The frons appears reduced in size due to the inward folding of the clypeus and labrum. (D1-D2) In some nymphs the mandibles are only partially formed, often missing the molar (arrows) and occasionally with split or duplicated incisors (arrowheads, D2). (E) Lateral view of wild type first nymph head showing normal formation of antennae and gena. (F) Whole body, side view of a wild type first nymph. (G) In Class I *Pa-slp*^{RNAi} first nymphs, the gena and antenna do not form, though the antennal socket may be present (arrow). (H) Class II *Pa-slp*^{RNAi} first nymphs do not hatch from the egg and are missing all gnathal (*) and some ventral thoracic segments (arrow). (I) Dorsal thoracic and abdominal segments remain relatively unaffected, while the ventral T1 only forms one leg with the other growing as unidentifiable tissue (arrow). ant – antenna; cl – clypeus; fr – frons; ge – gena; lm – labrum; mn – mandibles; T1, T2, T3 – first, second, third thoracic segment, respectively; ve – vertex.

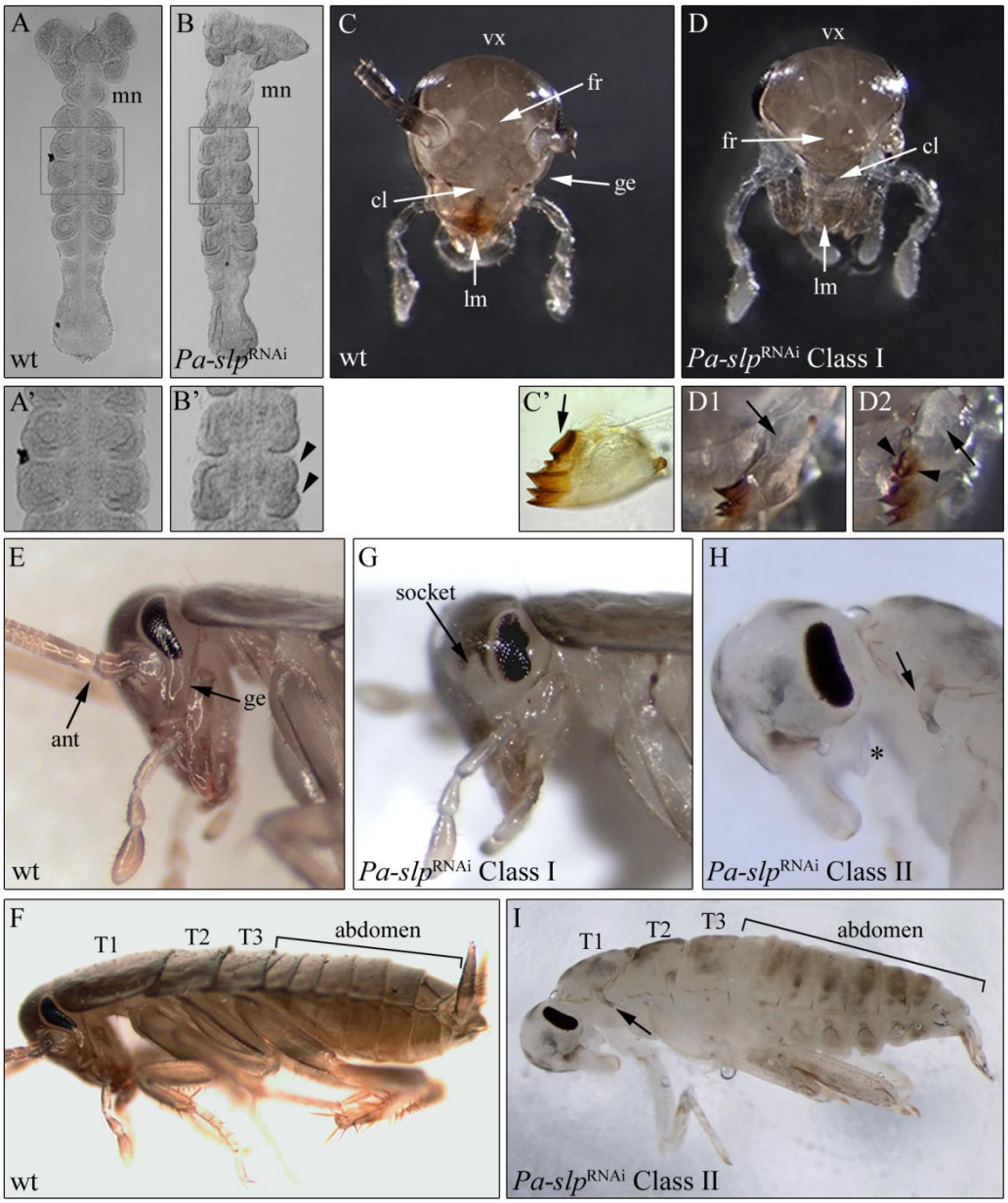

A
mn
wt
B
mn
Pa-slpRNAi
C
vx
fr
ge
cl
lm
wt
D
vx
fr
cl
lm
Pa-slpRNAi Class I
A'
B'
C'
D1
D2
E
ant
ge
wt
G
socket
Pa-slpRNAi Class I
H
*
Pa-slpRNAi Class II
F
T1
T2
T3
abdomen
wt
I
T1
T2
T3
abdomen
Pa-slpRNAi Class II

Figure 4.11

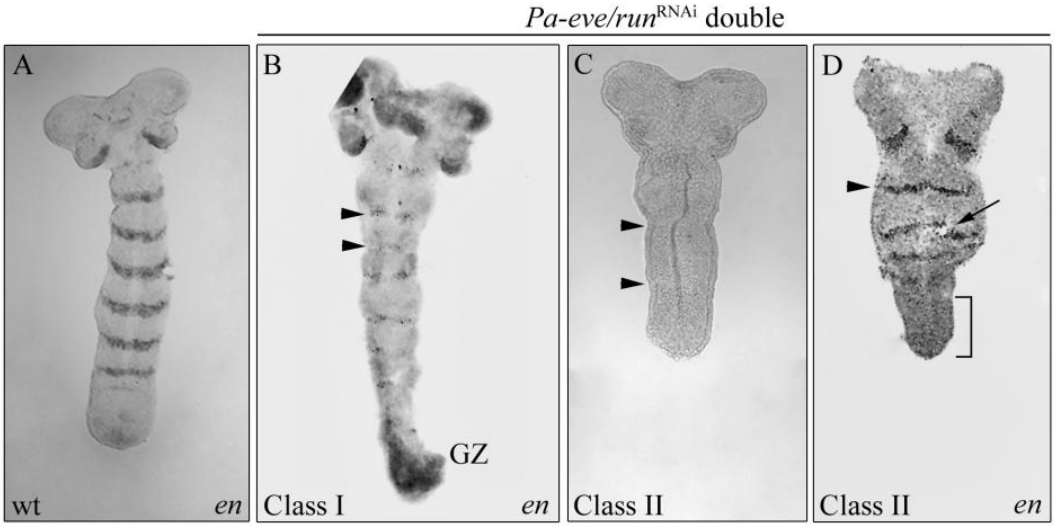


Figure 4.12: *Pa-eve* and *Pa-run* act redundantly to pattern the *Periplaneta* embryo. (A) *Pa-en* expression in a stage 8 wild type embryo. (B-D) Double *Pa-eve/run* knockdown RNAi phenotypes. (B) Stage 8 Class I *Pa-eve/run*^{RNAi} embryo displaying a thin body and a tapering tail ending in a malformed GZ. *Pa-en* expression is reduced to a weak speckled pattern in the posterior of the semi-formed segments (arrowheads). (C) An unstained, Class II *Pa-eve/run*^{RNAi} embryo at stage 8 is truncated in the posterior with only a few weak attempts at segment formation (arrowheads). (D) *Pa-en* expression is mostly normal in the mandibular segment (arrowhead) while other stripes are irregularly formed (arrow). The embryo is truncated in the posterior with a greatly reduced GZ (bracket).

Figure 4.13: Primary and secondary classification of *Periplaneta* pair-rule genes. In *Periplaneta*, the pair-rule genes are expressed in segmental stripes at both post-blastoderm (A) and germ band elongation (B-C). (A) During early post-blastoderm development, stage 4, segmental stripes of *Pa-eve* (green), *Pa-run* (blue), and *Pa-h* (red) are already detected in five stripes, corresponding to the three gnathal and first two thoracic segments. *Pa-pby* (purple) and *Pa-slp* (orange) expression lag behind and the full complement of five stripes is only detected later during post-blastoderm development, stage 6, and are displaced just posterior to the other three gene expressions. *Pa-en* (cyan) is expressed soon afterwards. (B) Wild type patterns of expression of the *Periplaneta* pair-rule genes during germ band elongation, stage 9. (C) Cartoon depiction of presumed co-expression of pair-rule genes shown. *Pa-h*, *Pa-eve*, and *Pa-run* have a dynamic pattern of expression in the posterior and mid-GZ, resolving into stripes of expression that move anteriorly. As the stripes of the primary pair-rule genes travel anteriorly, they begin to fade from the mid-line and only then do the stripes of the secondary pair-rule genes become expressed, followed shortly after by the expression of *Pa-en*. In each case, stripes of expression arise sequentially, with single segment periodicity. Based on these spatiotemporal expression patterns at post-blastoderm and in the GZ during germ band elongation, the *Periplaneta* pair-rule genes can be classified as primary (*eve*, *run*, *h*) or secondary (*pby*, *slp*), with primary expression developing first and anterior to the secondary pair-rule genes. Those segments affected by RNAi are indicated with a * (strong) or a ● (weak or moderate).

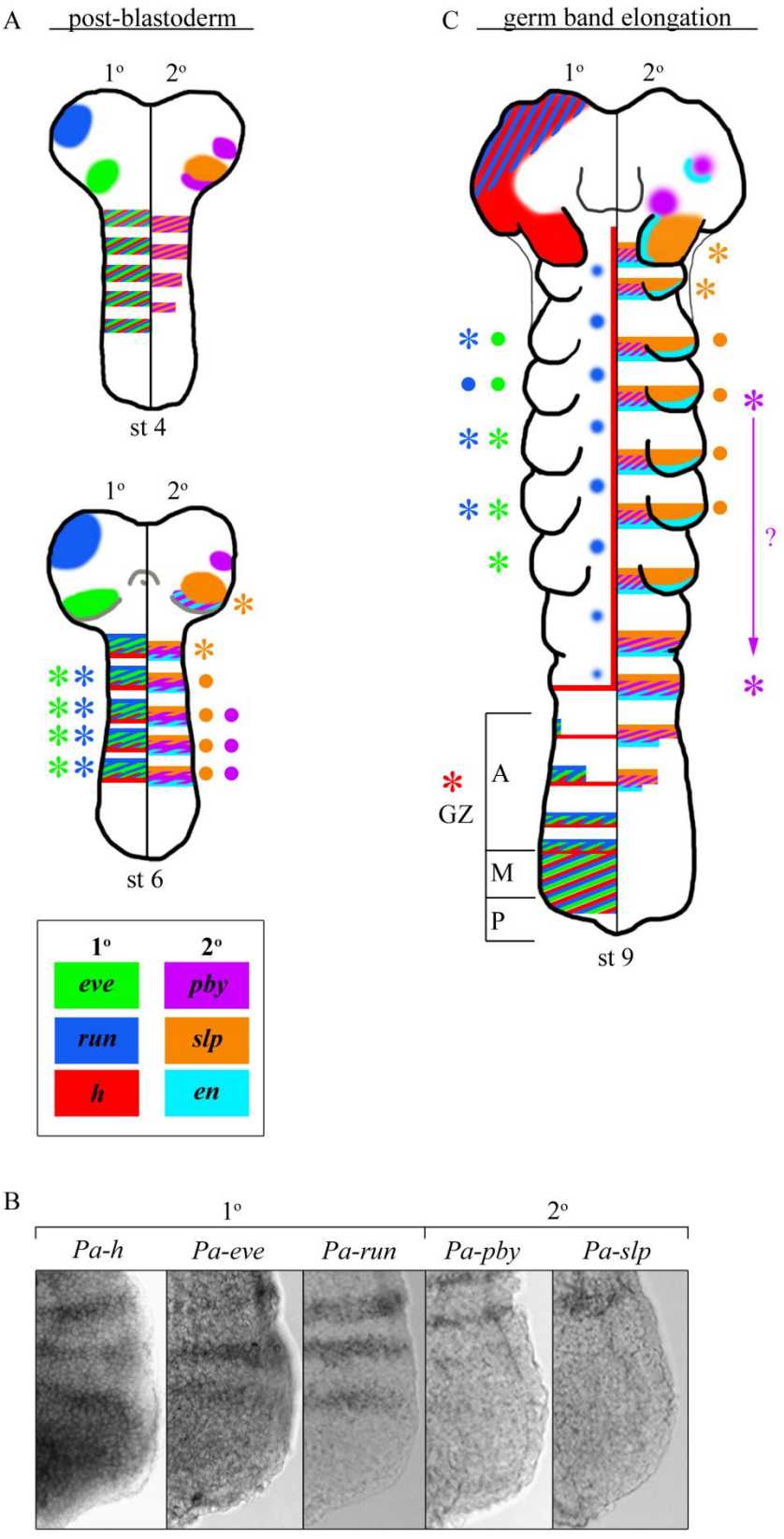

A
post-blastoderm
1°
2°
st 4
1°
2°
st 6
1°
2°
eve
run
h
phy
slp
en
C
germ band elongation
1°
2°
?
GZ
A
M
P
st 9
B
1°
2°
Pa-h
Pa-eve
Pa-run
Pa-pby
Pa-slp

Figure 4.13

Table 4.2: Comparison of arthropod pair-rule gene expression. A review of the literature compiled in this table shows the evolution of eight of the pair-rule genes: *eve*, *run*, *h*, *odd-skipped/odd-skipped-related* (*odd/odd-r*), *prd/pby*, *slp*, *odd-paired* (*opa*), and *fushi tarazu* (*ftz*) among the arthropods including members of the subphyla: Chelicerata, Myriapoda, Crustacea, and many orders of the Insecta. Throughout the course of evolution, the pattern of pair-rule gene expression seems to have changed from the more common segmental pattern (orange) in basal arthropods to that of the 'classic' pair-rule pattern (blue) of double-segment periodicity. However, the patterns of expression are highly variable and may be expressed in only a broad posterior GZ domain (green), a mix of expression patterns (dual-coloured squares) or not expressed at all (red). In some short germ band organisms studied, the patterns of expression can differ between the anterior segments, those patterned at blastoderm/post-blastoderm, and the remaining posterior segments added during germ band elongation. Spatiotemporal analysis in combination with existing functional data, where available, reveals a conserved classification for *eve* and *run* as primary (1º; boxed in blue) while *prd/pby* and *slp* are almost always classified as secondary (2º; green box) pair-rule genes.

Table References: (1-Akam, 1987a); (2-Benedyk et al., 1994); (3-Sommer and Tautz, 1991); (4-Liu et al., 2008); (5-Nakao, 2010); (6-Kraft and Jaeckle, 1994); (7-Davis et al., 2001); (8-Choe and Brown, 2007); (9-Patel et al., 1994); (10-Choe et al., 2006); (11-Binner and Sander, 1997); (12-Osborne and Dearden, 2005); (13-Grbic et al., 1996); (14-Keller et al., 2010); (15-Liu and Kaufman, 2005a); (16-Liu and Patel, 2010); (17-Mito et al., 2007); (18-Davis et al., 2001); (19-Dawes et al., 1994); (20-Patel et al., 1992); (21-this work); (22-Copf et al., 2003); (23-Davis et al., 2005); (24-Janssen et al., 2012); (25-Janssen et al., 2011); (26-Hughes and Kaufman, 2002a); (27-Hughes and Kaufman, 2002b); (28-Chipman et al., 2004); (29-Chipman and Akam, 2008); (Dearden et al., 2002); (31-Pechmann et al., 2009); (32-Schwager et al., 2009); (33-Damen et al., 2005); (34-Schoppmeier and Damen, 2005a).

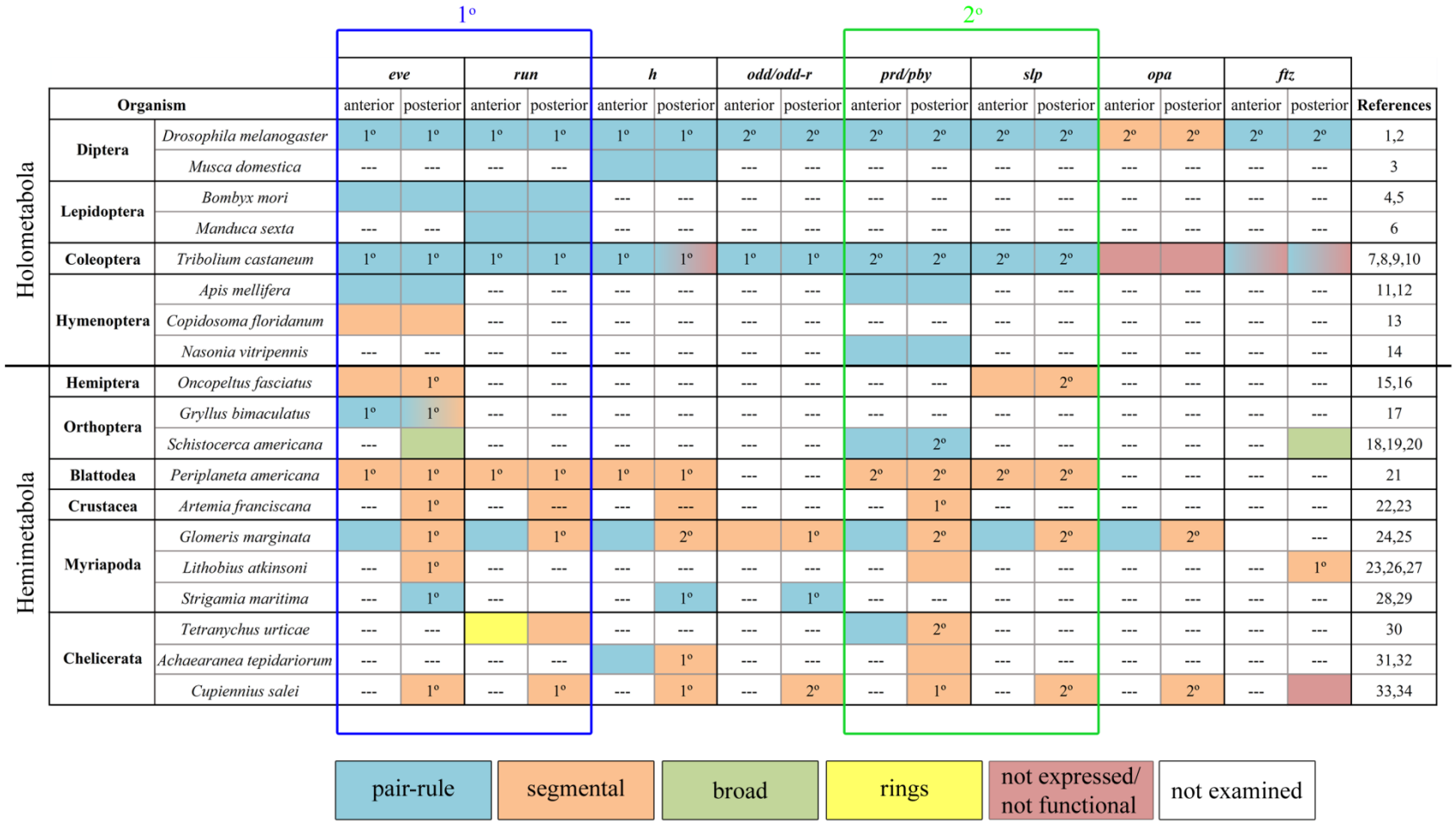

1º
2º
Holometabola
Hemimetabola
eve
run
h
odd/odd-r
prd/pby
slp
opa
ftz
anterior
posterior
Organism
References
Diptera
Drosophila melanogaster
Musca domestica
Lepidoptera
Bombyx mori
Manduca sexta
Coleoptera
Tribolium castaneum
Hymenoptera
Apis mellifera
Copidosoma floridanum
Nasonia vitripennis
Hemiptera
Oncopeltus fasciatus
Orthoptera
Gryllus bimaculatus
Schistocerca americana
Blattodea
Periplaneta americana
Crustacea
Artemia franciscana
Myriapoda
Glomeris marginata
Lithobius atkinsoni
Strigamia maritima
Chelicerata
Tetranychus urticae
Achaearanea tepidariorum
Cupiennius salei
1,2
3
4,5
6
7,8,9,10
11,12
13
14
15,16
17
18,19,20
21
22,23
24,25
23,26,27
28,29
30
31,32
33,34
pair-rule
segmental
broad
rings
not expressed/ not functional
not examined

Table 4.2

CHAPTER V

Expression and functional analysis of the smORF gene *tarsal-less* in *Periplaneta*

ABSTRACT

tarsal-less (*tal*) is a non-canonical, polycistronic gene encoding fully-functional peptide products through small Open Reading Frames (smORFs). These small peptides are essential during embryonic and post-embryonic development of *Drosophila*. RNAi studies on the *Tribolium tal* homologue show this gene to have gap-like properties, resulting in truncation of the posterior abdomen and a transformation of the remaining segments towards a thoracic identity. *tal*-like genes have been discovered in other organisms, indicating an ancient function for this gene; however, the role(s) of *tal* have not been studied in these organisms. I have cloned and analysed *tal* expression during embryonic development in the cockroach, *Periplaneta americana*. *Pa-tal* expression in the growth zone is dynamic and cyclic during germ band elongation, suggesting involvement with the Notch segmentation clock known to function in *Periplaneta*.

The Notch-signalling cascade is a conserved mechanism in leg segmentation (from flies to spiders), suggesting that Notch-mediated joint formation could be the phylotypic trait of arthropods. Our understanding of leg development in *Drosophila* is increasingly detailed, while our knowledge in other arthropods remains comparatively basic. While *Drosophila* Notch-signalling is activated in all developing segments, "true" joints of the proximal leg differ from the joints of the tarsomeres due to changes in gene regulation in the respective leg regions. We have explored the mechanisms of joint formation by analyzing joint patterning genes including *tal*, *Notch*, *Delta*, and *nubbin* in the basal insect, *Periplaneta americana*. *tal* and *nub* have a complementary role in *Drosophila*, but may be redundant in more basal species, underpinning the existence of variable gene functions downstream of a conserved Notch cassette controlling developmental boundaries in a variety of contexts.

N.B. Portions of this chapter, pertaining to *Pa-nubbin* leg expression and function, have been published in Developmental Biology: Turchyn N, **Chesebro J**, Hrycaj S, Couso JP, Popadić A. 2011. Evolution of *nubbin* function in hemimetabolous and holometabolous insect appendages. Dev Biol 357(1): 83-95.

INTRODUCTION

Putative proteins, from predicted coding sequences, of less than 100 amino acids have widely been ignored in bioinformatic searches and gene annotations under the general assumption that they are too small to be relevant. With advancing technology and the scientific wherewithal we are now beginning to understand that entire groups of functional and biologically important peptides may have been overlooked (Tupy et al., 2005). One of the best studied of these small Open Reading Frame (smORF) genes is the newly discovered, polycistronic gene *tarsal-less* (*tal*; also called *polished rice*). In *Drosophila melanogaster*, *tal* encodes four 11- to 32-amino-acid-long, functionally redundant peptides (three Type-A and one Type-AA), each containing a conserved seven amino acid sequence of LDPTGXY (Galindo et al., 2007; Kondo et al., 2007). The Type-A and Type-AA peptides are essential for the proper formation of ectodermal/epithelial structures during both embryonic and post-embryonic development in *Drosophila* (Galindo et al., 2007; Kondo et al., 2007; Pueyo and Couso, 2008). *Dm-tal* also codes for one non-functional/non-translated (Type-B) peptide that does not contain this motif (Galindo et al., 2007).

During embryogenesis, *Dm-tal* expression is highly dynamic. Initially expressed as seven stripes and an anterior cluster of cells in the blastoderm, expression is then restricted to the tracheal precursors before becoming segmentally expressed in appearance (Galindo et al., 2007; Kondo et al., 2007). *Dm-tal* is also expressed in the developing trachea, posterior spiracles, and cephalopharyngeal skeleton (Galindo et al., 2007; Kondo et al., 2007; Kondo et al., 2010; Pueyo and Couso, 2011). *Dm-tal* loss-of-function mutants show malformation or absence of these structures, most notably and best studied are the missing denticle belts (Galindo et al., 2007; Kondo et al., 2007; Kondo et al., 2010). The role of *Dm-tal* in denticle belt formation is carried out through regulation of the zinc-finger

transcription factor *shavenbaby* (*svb*), possibly through cleavage of a long repressor form into a short, activator form that regulates filamentous actin and the protrusion of the ventral denticles (Kondo et al., 2007; Kondo et al., 2010).

During post-embryonic development, *Dm-tal* is expressed in leg imaginal discs at the location of the future tarsal segments where it plays two important roles: 1) intercalation of tarsal segments in third instar larva and 2) tarsal joint formation in the pupa. Loss-of-function of *Dm-tal* at these stages results in adult flies lacking the tarsal sub-segments, hence the name. Several studies have shown that *Dm*-Tal regulates the expression of known leg patterning genes, such as *rotund* (*rn*) and *spineless* (*ss*), leading to the intercalation of tarsal segments ta2-ta4 (Galindo et al., 2007; Pueyo and Couso, 2008; Pueyo and Couso, 2011). This function for *Dm*-Tal is further associated with regulation of the active form of *Dm-svb*, forming a negative feedback loop leading to the inhibition of *Dm-Delta* expression, thus creating the sharp Dl+/Dl- signalling border required for joint formation (Pueyo and Couso, 2011). *Dm-tal* is not expressed and does not function in the other leg segments containing "true" joints (muscle attachments) and could be an evolutionary novelty in specifying non-muscle joints of the tarsomeres. While the functions of *tal* in tarsal joint and denticle belt formation involve regulation of *svb*, this is not always the case, as with taenidial folding in the trachea and tarsal intercalation (Kondo et al., 2010; Pueyo and Couso, 2008; Pueyo and Couso, 2011), indicating that *tal* can function in alternate ways in differing developmental contexts. On a final note, several studies indicate that *tal* is expressed non-cell-autonomously, suggestive of a possible functional role as a signalling factor (Kondo et al., 2007; Pueyo and Couso, 2008; Pueyo and Couso, 2011).

tarsal-less has been studied in one other holometabolous organism, the short germ band insect, *Tribolium castaneum*. The *Tribolium-tal* homologue, *mille-pattes* (*mlpt*),

encodes for three Type-A peptides, each containing the conserved heptapeptide LDPTGXY, one Type B, but no Type-AA peptides (Savard et al., 2006). Like *Drosophila-tal*, *Tc-mlpt* expression is dynamic throughout embryogenesis. Initially expressed as a broad domain in the anterior blastoderm (Savard et al., 2006), at post-blastoderm *Tc-mlpt* is expressed as a single, wide stripe in the future mandibles and a broad posterior domain. During germ band elongation, *Tc-mlpt* is expressed in a gap-like domain covering most of the thorax and the anterior abdomen with an additional domain of expression in the posterior growth zone (GZ). Upon completion of segment formation, *Tc-mlpt* is expressed in the tracheal precursors and as several stripes in the legs corresponding to all future leg joints, not restricted to the tarsus (Savard et al., 2006).

RNA-depletion of *Tc-mlpt* revealed this gene to have differing functions compared to *Drosophila*, being required for proper body segmentation. *Tc-mlpt*^{RNAi} leads to truncation of the posterior-most abdomen with the remaining abdominal segments taking on a thoracic identity, complete with ectopic legs (Savard et al., 2006). This phenotype, along with the interactions between *Tc-mlpt* and several of the gap genes, led the authors to conclude that *Tc-mlpt* functions as a gap gene during *Tribolium* embryogenesis. In addition, the thoracic larval legs are underdeveloped, becoming short and stumpy upon loss of *Tc-mlpt* expression (Savard et al., 2006). Finally, it was also noted that partial fusions occurred between dorsal abdominal segments, while the head and thoracic tagmata were unaffected (Savard et al., 2006). Recently, Schnellhammer (2012) showed a similar expression pattern for a *Tribolium* homologue of *svb*, in the posterior blastoderm and GZ. *Tc-svb*^{RNAi} embryos and larvae display similar phenotypes to those affected by *Tc-tal*^{RNAi} – abdominal truncation, ectopic legs, stumpy thoracic legs with malformed joints (Schnellhammer, 2012), indicating a possible conserved interaction between *tal* and *svb* compared to *Drosophila*. As these expression patterns and functions diverge between

Drosophila and *Tribolium*, especially during segmentation processes, I set out to isolate and analyse a *tal* homologue in the basal, hemimetabolous insect *Periplaneta americana* in hopes of ascertaining a conserved mechanism of *tal* function between these species and gaining some evolutionary insight.

The *Pa-tal* homologue contains two putative Type-A peptides, each containing the conserved LDPTGXY motif. Embryonic expression of *Pa-tal* is dynamic during both body and leg development. During germ band elongation, *Pa-tal* expression in the posterior GZ appears in a cyclic manner similar to that described for N-pathway members in *Periplaneta* (Pueyo and Couso, 2008). Functional analysis via RNA interference shows *Pa-tal* is required for anterior and posterior body segmentation as well as proper leg patterning through an apparent interaction with the POU homeodomain gene *nubbin*. Previous studies on *nub* showed this gene to be important in appendage patterning in many arthropod species, though expression and function are variable (Damen et al., 2002; Hrycaj et al., 2008; Li and Popadić, 2004; Prpic and Damen, 2005; Turchyn et al., 2011).

Recently, Turchyn et al. (2011) described this diversity of *nub* expression and function in *Acheta domesticus* (house cricket), *Drosophila*, and *Periplaneta*. Their results illustrate that, although expressed in all (crickets and cockroaches) or most (flies) leg segments, the major function of *nubbin* differs in location in the different species. *Acheta-nub*^{RNAi} resulted in first nymphs with a reduced trochanter and fusion between the tibia and the first tarsal subsegment, whereas in *Periplaneta* the coxa/trochanter and femur/tibia joints are most affected (Turchyn et al., 2011). *Drosophila nub* mutants also displayed heavy fusion between all leg segments where *Dm-nub* is normally expressed, being the true joints of the coxa, trochanter, femur, and tibia, along with a loss of tarsomere-5 and the distal claws (Turchyn et al., 2011). It appears as though *Pa-nub* also works in conjunction with *Pa-tal* in patterning the cockroach leg and in maintaining the limbless

abdomen. Similar abdominal functions were noted separately for *nub* in *Oncopeltus* (Hrycaj et al., 2008), and *tal* in *Tribolium* (Savard et al., 2006), indicating possibly overlooked functions/interactions for these genes in basal arthropod species.

Bioinformatic searches for *tal* homologues in other organisms revealed the existence of *tal*-like genes, containing the conserved LDPTGXY motif, in most insect lineages and as distant, phylogenetically, as the crustacean *Daphnia pulex* (Cladocera). Although homologues of *tal* have been found in numerous insect and crustacean species, it has not been well studied. With the results presented below, the expression and functional roles for *tarsal-less* have now been examined in three insect species, including two holometabolous insects – the highly derived, long germ band *Drosophila* (Galindo et al., 2007; Kondo et al., 2007; Kondo et al., 2010; Pueyo and Couso, 2008; Pueyo and Couso, 2011) and the short germ band beetle *Tribolium* (Savard et al., 2006) – and now the hemimetabolous, short germ band insect *Periplaneta*.

RESULTS

Isolation and Cloning of *Periplaneta tarsal-less*

Isolation of *Pa-tal* proved difficult as the *Periplaneta* genome has not been sequenced and it is hard to design quality primers for such a small region of conserved sequence (24bp). Initial primers were far too degenerate for practical purposes (~50,000-150,000) and often resulted in abundant off-target priming. In order to lower the degeneracy and to extend the usable region for primer design, homologous *tal* sequences were identified from species closely related to *Periplaneta*, the Orthopterans *Locusta migratoria* (DY229958) and *Gryllus bimaculatus* (AK281313), and used to make a general consensus sequence for each of the two smORFs. Next, the most commonly used codons were examined for *Periplaneta* by analysing the codon usage bias (Appendix 1) (Nakamura et al., 2000), thus allowing the reduction in the number of codon options for some of the more degenerate amino acids, such as Leucine that has six codons attributed to it. Together, these two steps helped to reduce primer degeneracy to 50-600 for each primer.

Taking advantage of the incomplete second smORF of *Gryllus tarsal-less*, which was missing the last two amino acids and the stop codon, made it possible to be used as an experimental control in RACE PCR using the degenerate primers. This step allowed for the optimisation of PCR parameters and increased confidence that an isolated sequence would be a true *tal* homologue, if the completed second smORF ended with a stop codon. This, indeed, was the case and a completed *Gb-tal* sequence of 706bp was assembled that contained two smORFs, each encoding a 12aa-long peptide with the conserved LDPTGXY motif (Appendix 11).

These degenerate *tal* primers were then used with *Periplaneta* cDNA in similar 5' and 3' RACE PCR conditions to construct a 569bp transcript for *Periplaneta-tal*

(Appendix 9). *Pa-tal* contains two smORFs, each encoding a putative 12aa-long peptide with the conserved LDPTGXY sequence at the C-terminal end. Phylogenetic analysis of *tal* sequences proved challenging with conventional methods, as they vary in length and only the heptapeptide sequence is conserved between species. For these and other reasons *tal* homologues were analysed and compared manually, which will be explored further in the Discussion: 'Evolution of tarsal-less peptides'.

Embryonic expression of *Pa-tal*

A 397bp riboprobe covering both smORF sequences was synthesised to detect *Pa-tal* mRNA expression via *in situ* hybridisation at several embryonic stages of development. At post-blastoderm stage 4, there is a wide band of expression in the anterior, just below the protocephalon in the putative mandibles, and a wide stripe near the posterior end of the embryo (Fig. 5.1A). At late stage 5, these expression patterns remain strong (Fig. 5.1B), while a new domain of expression develops in the very posterior of the embryo at early stage 6 (Fig. 5.1C). In late stage 6 embryos, the mandibular stripe of expression has almost completely faded and the initial posterior stripe is beginning to fade from the midline (Figs. 5.1D). At the same time, the broad posterior domain of *Pa-tal* is being refined into a new stripe of expression in the anterior GZ (Fig. 5.1D).

This dynamic posterior pattern continues through germ band elongation, where the broad GZ expression of *Pa-tal* develops into a wide stripe in the mid-anterior GZ (compare black and red arrowheads, Fig. 5.1E-H). As this new stripe forms, the previously established stripe of *Pa-tal* expression moves anteriorly and fades before the formation of a new segment (*, Figs. 5.1E-H). This cyclic pattern of expression continues until all the body segments have formed. As terminal segment addition is completed, appendage development continues and *Pa-tal* is expressed as several stripes in the antennae and

mouthparts, as well as the anal pads and at the distal ends of the pleuropodia (Fig. 5.1I). Several stripes of expression appear throughout leg development, near the future joints (Fig. 5.1I), examined in more detail in a later section. Finally, as the embryo begins dorsal closure, the stripes of *Pa-tal* fade from the appendages and expression is now observed in clusters in the apodemes at the proximal limbs near the body wall (Fig. 5.1J) and in the developing CNS (Fig. 5.1J). Altogether, *Pa-tal* expression is dynamic during cockroach embryogenesis showing some divergent (body) and conserved (leg) patterns compared to *Drosophila* and *Tribolium*.

Pa-tal affects anterior and posterior body patterning

Functional analysis in *Periplaneta* was carried out by injecting virgin females with a 556bp *Pa-tal* dsRNA fragment and analysing the resultant embryonic and first nymph phenotypes. The hypomorphic range of embryonic phenotypes were placed into one of three categories (Table 5.1). The weak Class I *Pa-tal*^{RNAi} embryos (224/1316; 17.0%) develop normally in the anterior head, gnathal, and thoracic segments, while segmentation defects are observed in the abdomen as asymmetric segment fusions, causing a slight bend to one side (Fig. 5.2B). Moderate Class II embryos (121/1316; 9.2%) showed stronger segment fusions in both the posterior and the anterior, often presented as a bulging out (not shown) or 'bunching up' (Fig. 5.2C) of the gnathal and thoracic segments in addition to lateral abdominal segment fusion (Fig. 5.2C). Strong Class III *Pa-tal*^{RNAi} embryos (115/1316; 8.7%) have an 'asegmental' phenotype, as the inter-segmental furrows do not appear to form (Fig. 5.2D). Although overt segmentation did not occur, there appear to be some weak attempts at either segment or limb bud formation, giving the embryo a 'wavy' appearance at the lateral edges (Fig. 5.2D). In addition, some Class III embryos are extremely small in size compared to wild type embryos of a similar stage (Fig. 5.2E

compared to 5.2A). Though small, the embryo in Fig. 5.2E is not simply at an earlier stage and can be roughly placed at around stage 11, as it was found in an ootheca containing larger embryos of a similar age and this is the stage in which the proctodeal pit begins to form in *Periplaneta* embryogenesis (black arrowhead, Fig. 5.2E) (Lenoir-Rousseaux and Lender, 1970). In all classes of *Pa-tal*^{RNAi}, an apparently unaltered posterior GZ and normal pre-gnathal head development occurs (Fig. 5.2A-E).

Pa-tal^{RNAi} nymphal phenotypes were separated into two classes based on the degree of abdominal and leg defects. Class I *Pa-tal*^{RNAi} nymphs (108/534; 20.2%) are likely a reflection of Class I embryos as they showed various degrees of segment fusions only in the abdomen, while the head and thoracic segments remain wild type in appearance (Fig. 5.2G-G'). In these *Pa-tal*^{RNAi} affected first nymphs, some abdominal segments are only partially formed, hemi-segments, missing the left or right half (black dots, Fig. 5.2G), or segments may be fused on one side from the midline to the lateral edge (*, Fig. 5.2G); whole segment deletions nor fusions along the entire length of adjoining segments were not observed. Abdominal segment fusions occur either ventrally or at both dorsal and ventral sides, but never in the dorsal tergites only. The dorsal fusion patterns closely mimic that ventrally, providing a mirror-image effect when the abdomen is dissected and laid flat (Fig. 5.2G'), reflecting the fact that the dorsal tergites develop as lateral outgrowths of the ventral tissue during embryogenesis. There is no indication that fusions occur more often between certain segments over others, nor is there a penchant for left over right side, or vice versa.

Class II nymphs (43/534; 8.1%) also have strong dorsal and ventral abdominal fusions, as described for Class I above, with the addition of variable leg defects – often as a slight bend in the first tarsomere (Fig. 5.2I) or less frequently a bend in the tibia (not shown). Occasionally, neighbouring legs may be fused together proximally (Fig. 5.2J) or

may be severely misshapen (Fig. 5.2K). These leg phenotypes and functions of *Pa-tal* in these appendages will be discussed further in the section on *Periplaneta* leg patterning.

Altogether, the RNAi phenotypes correlate well with the patterns of embryonic expression, affecting anterior and posterior segment formation and leg patterning. Loss of dynamic expression in the posterior GZ leads to asymmetric fusions of the abdominal segments, yet the growth zone remains unaffected. In the legs, *Pa-tal* is expressed in all developing podomeres, just proximal to the developing joints. However, *Pa-tal*^{RNAi} shows only a slight effect on nymph leg patterning which could be due to less efficient RNAi in the developing legs compared to body patterning or *Pa-tal* may play a more redundant role in this region.

Analysis of *Pa-tal*^{RNAi} segmentation phenotypes

Misaligned segment formation

In order to further analyse the asymmetric fusion phenotypes of *Pa-tal*^{RNAi} embryos and nymphs, I examined the expression of the pan-segmental marker *Pa-engrailed*. If *Pa-tal* had a role in segment formation one might expect to see defects in segment patterning genes, such as this. In wild type embryos (Figs. 5.3A,C), *Pa-en* is expressed as stripes that span the width of the embryo in the posterior of all developed and developing segments. In the Class I *Pa-tal*^{RNAi} embryos, stripes of *Pa-en* form normally in the gnathal and thoracic segments, but expression is disrupted in the abdomen, forming only partially on either the left or right hand side of the body (Fig. 5.3B). As a result, two hemi-stripes from either side may become misaligned leading to a cross over between the mismatched pairs, such as the matching up of the A2 hemi-stripe on the right with that of A3 on the left in the embryo shown in Fig. 5.3B. This mismatched alignment leads to the

fusions and partial deletions of body segments observed in Class I and Class II *Pa-tal*^{RNAi} first nymphs.

Class II *Pa-tal*^{RNAi} phenotypes show extensive fusion in both the posterior and the anterior body segments. *Pa-en* staining in these embryos shows a similar misalignment between the left and right hand stripes of expression in the anterior, as evidenced by the left-hand mandible stripe connecting to the right-hand labium stripe in the embryo shown in Fig. 5.3D. These *Pa-en* stripes are weak and thin compared to wild type (Fig. 5.3C) or even Class I *Pa-tal*^{RNAi} embryos (Fig. 5.3B). The asegmental Class III *Pa-tal*^{RNAi} embryos show little to no overt signs of segmentation, yet there are still very weak, speckled stripes of *Pa-en* expression where segments might normally form (Fig. 5.3E and inset). As shown in Figs. 5.3A-E, there is a graded effect on *Pa-en* expression in *Pa-tal*^{RNAi} embryos. It is not yet clear if *Pa-tal* is directly involved in activating *Pa-en* expression or plays a role further upstream in possibly regulating pair-rule gene expression, which may be a more likely situation.

Pa-tal *and the posterior patterning gene network*

In all classes of *Pa-tal*^{RNAi} embryos the posterior GZ appeared more or less normal. To confirm the unaffected state of the GZ, I examined the expression of *Pa-caudal*, a gene whose function is crucial to the establishment and maintenance of the posterior GZ (see Chapter III). In wild type cockroach embryos at late germ band elongation, *Pa-cad* is expressed broadly in the mid-GZ but not in the most posterior end (Fig. 5.3F). In *Pa-tal*^{RNAi} embryos, *Pa-cad* is still expressed in the posterior, complete with a characteristic clearing in the posterior tip, even in the strong Class III *Pa-tal*^{RNAi} embryos (Fig. 5.3G). Conversely, *Pa-tal* expression is absent in Class-A *Pa-cad*^{RNAi} Class 'A' embryos (Fig. 5.3H).

The cyclic expression of *Pa-tal* in the GZ resembles the dynamic pattern of expression of N-pathway genes in *Periplaneta* required for proper posterior patterning (Pueyo et al., 2008). Therefore, I explored the potential interactions between *Pa-tal* and the N-signalling pathway by examining any changes in expression of *Pa-Dl* and *Pa-h*, downstream targets of the N-pathway (Pueyo et al., 2008), in *Pa-tal*^{RNAi} embryos. First considered was *Pa-h*, as similar hemi-stripe formations of *Pa-en* also occurred in *Pa-h*^{RNAi} embryos (Pueyo et al., 2008); although no misaligned stripe connections were noted. In Class I *Pa-tal*^{RNAi} embryos, the broad posterior GZ domain and anterior stripes of *Pa-h* expression remained unaffected (Figs. 5.3I-J). Similarly, the segmental stripes of *Pa-Dl* expression in the anterior GZ were unaffected in Class I and II *Pa-tal*^{RNAi} embryos (Figs. 5.3K-M). In the reciprocal experiment, *Pa-tal* expression was completely absent from the posterior GZ in *Pa-N*^{RNAi} affected embryos (Figs. 5.3N-O). Attempts to analyse *Pa-Dl* and *Pa-h* in Class III *Pa-tal*^{RNAi} embryos were unsuccessful, as these embryos were uncommon and their fragility and small size made them difficult to manipulate during *in situ* hybridisations.

Overall, these results indicate that the early post-blastoderm expression of *Pa-tal* is not involved in formation or maintenance of the posterior GZ and the cyclic posterior expression during germ band elongation may function downstream of Cad and/or the Dl/N-signalling pathway. Alternatively, as RNAi-induced loss of expression of either *Pa-cad* or *Pa-N* results in a reduced and altered GZ (Chapter III) (Pueyo et al., 2008), the loss of *Pa-tal* expression in these backgrounds may be a secondary consequence of growth zone collapse. However, as demonstrated in Chapter III, this explanation may be circumvented via analysis of *Pa-tal* in DAPT cultured embryos in order to fully understand this loss of expression.

Periplaneta leg patterning

Expression of leg patterning genes: tal, nub, Dl, and N

In *Periplaneta*, *Pa-tal* is expressed in the developing limbs from an early stage. At stage 10, as limb buds are emerging, expression is observed at the distal tips of the developing legs and gnathal appendages (Fig. 5.4A). By stage 14, *Pa-tal* is broadly expressed in the middle of the leg, between two constricting points (Fig. 5.4B). As the limbs extend further and the segment constrictions become stronger, stage 16, *Pa-tal* expression is cleared from the central regions so that expression now appears as two medial rings (Fig. 5.4C). By stage 19, when all leg constrictions become evident, *Pa-tal* is expressed at the distal part of all future leg segments, just proximal to the putative joints (Fig. 5.4D). Similar patterns of *tal* expression have been observed in the cricket, *Gryllus bimaculatus* (Appendix 11).

Expression of other known leg patterning genes *nub*, *N*, and *Dl* were also examined in the developing cockroach legs. Previous reports showed that *Pa-nub* has a dynamic expression in the developing cockroach legs (Li and Popadić, 2004; Turchyn et al., 2011). Independent analysis using a 450bp riboprobe that recognized the conserved POU domain and Homeodomain, showed early *Pa-nub* is expressed at the distal tips of the developing buds of the legs and gnathal appendages (Fig. 5.4E). At early stage 13, *Pa-nub* is expressed broadly in the medial leg and a wide stripe is visible in the proximal region of the leg (Fig. 5.4F). At late stage 13, the broad middle stripe splits in two, for a total of three rings of expression and a diffuse pattern at the tip of the limbs (Fig. 5.4G). By stage 17, five stripes of *Pa-nub* are expressed proximal to each of the constricting leg joints (Fig. 5.4H).

As with *Pa-tal* and *Pa-nub*, *Pa-Dl* is expressed at the distal tip of the protruding limb buds, stage 12 (Fig. 5.4I). As the limb buds start to constrict, stage 15, *Pa-Dl* is

expressed broadly in the middle of the leg, between the two constricting points (Figs. 5.4J). As legs continue to grow, stage 16, *Pa-Dl* is expressed in two bands in the middle of the developing leg (Fig. 5.4K). By stage 21, *Pa-Dl* is expressed in all leg segments, proximal to the constrictions of the future leg joints (Figs. 5.4L), including several bands in the developing tarsal sub-segments (Fig. 5.4L). Similarly, expression of *Pa-N* at stage 14 is in a broad domain in the medial leg (Fig. 5.4M), which is then refined into two stripes medially with a new stripe forming in the proximal leg at stage 16 (Fig. 5.4N). In later stages of limb development, stage 19, *Pa-N* is expressed in the distal portion of all podomeres (Fig. 5.4O)

Considering these data, the patterns of gene expression in the legs, including that of *Pa-tal*, are dynamic and closely resemble one another. Each of these genes is expressed at the earliest limb bud formation and the expression changes as the legs develop, eventually becoming restricted to the distal end of each leg segment, just proximal to where the future joints will form. These similar patterns of expression hint at some level of interaction between N-signalling, *tarsal-less*, and *nubbin*; interactions that may be conserved with *Drosophila* and other arthropod species.

Effects of Pa-tal *and* Pa-nub *on leg development*

The effects of *Pa-tal* on nymphal leg patterning, as determined by maternal RNAi, were minimal. *Pa-tal*^{RNAi} Class II nymphs showed only limited effects in leg development, often displayed as a bend in the middle of the first tarsomere (ta1), but not associated with any loss of joint formation (Fig. 5.2I and Fig. 5.5B'). There were a few examples of proximal fusions between several legs (Fig. 5.2J), but this can be attributed to earlier misalignment and fusion of ventral segments and may not be an actual representation of *Pa-tal* function in leg patterning. A few legs displayed gross abnormalities, including

coxa/trochanter fusion (Fig. 5.2K), sometimes accompanied by a reduced femur, malformed tibia, and unformed/missing tarsomeres and claw (Fig. 5.2K). These leg fusion phenotypes are quite similar to those observed in *Pa-nub*^{RNAi} first nymphs (Turchyn et al., 2011).

As shown by Turchyn et al. (2011), late stage *Pa-nub*^{RNAi} embryos and first nymphs displayed a range of leg phenotypes including a reduced trochanter, which may be fused to the coxa, and fusion between femur and tibia. Seeing that *Pa-nub* has a profound effect on leg development and is expressed in all leg segments, similar to *Pa-tal*, analysis of possible interactions between *Pa-tal* and *Pa-nub* was done by performing double RNAi experiments. Diminished expression of both *Pa-tal* and *Pa-nub* resulted in first nymphs (Class I; 50/153; 32.7%) (Table 5.1) displaying a mixture of the phenotypes displayed by single injections, including abdominal segment fusions, crooked antennae, and malformed legs (Fig. 5.2 and Turchyn et al., 2011). The most notable of these double *Pa-tal/nub*^{RNAi} phenotypes is an increased rate of a severely reduced or unformed trochanter (Fig. 5.5D). Where the trochanter is unaltered in single *Pa-tal*^{RNAi} first nymphs (Fig. 5.5C), it is usually reduced in size or fused to the coxa in *Pa-nub*^{RNAi} first nymphs (Fig. 5.5D and Turchyn et al., 2011). Finally, while no effects were observed in tarsal joint formation for either *Pa-tal* or *Pa-nub* single RNAi (Figs. 5.5B' and 5.5C', respectively), loss of expression of both of these genes resulted in small defects in the tarsomeres, apparent as slight malformations between tarsal sub-segments; however, no major effects were observed in tarsal joint formation (Fig. 5.5D'). It must be considered here that RNAi may not be effective in completely eliminating expression of all *Pa-tal* and *Pa-nub* transcripts, whereby, even minimal translation may be enough to allow for complete joint formation.

On a final note, there is some conservation in leg patterning between *Periplaneta* and other arthropods. As N-signalling seems to be required for proper leg segmentation

and joint development in all species studied so far, there is also a conserved requirement for Notch in the partial regulation of *nub* expression (Angelini et al., 2012; Bishop et al., 1999; Mito et al., 2011; Prpic and Damen, 2009; Rauskolb and Irvine, 1999). This requirement is conserved in *Periplaneta*, where *Pa-N*^{RNAi} results in reduced intensity and size of *Pa-nub* stripes in the legs (Fig. 5.6A-B) (Turchyn et al., 2011). It is currently unknown if *Pa-tal* is similarly regulated by N-signalling, as in the fly (Pueyo and Couso, 2011), or how *Pa-tal* and *Pa-nub* may be interacting during cockroach leg development.

Pa-tal and *Pa-nub* help form the limbless abdomen

Determining the effects of *Pa-nub* on *Periplaneta* leg development revealed several unexpected results. While Class I *Pa-nub*^{RNAi} first nymphs (Table 5.1; 44/167; 26.3%) showed appendage defects such as wavy antennae and fused podomeres, in a small fraction of *Pa-nub*^{RNAi} first nymphs (Class II; 6/167; 3.6%), abdominal fusion phenotypes were observed (Fig. 5.7B), similar to those previously described for *Pa-tal*^{RNAi} (Fig. 5.2G-G'). As expression of *Pa-nub* in the GZ has not previously been described, I examined *Pa-nub* expression at early germ band elongation. Preliminary results from *in situ* hybridisation revealed a broad expression of *Pa-nub* in the posterior GZ (Fig. 5.7C); however, further analysis is required to determine the full extent of this expression, especially whether it is cyclic in a similar manner as *Pa-tal* and members of the N-signalling pathway in this location (Pueyo et al., 2008).

A few scattered cases of ectopic appendage-like structures were noted in *Pa-tal*^{RNAi} first nymphs, though these appendages were not restricted to the anterior abdomen (Figs. 5.7D-E). Ectopic outgrowths were found on ventral-lateral A5 (Fig. 5.7D) and as a lateral outgrowth of the ventrally fused segments A2/A3 (Fig. 5.7E). In all cases, the outgrowths were small and usually contain several bristles or hairs. No such outgrowths

were noticed in single *Pa-nub*^{RNAi} nymphs in *Periplaneta* (Turchyn et al., 2011). Surprisingly, several *Pa-tal/nub*^{RNAi} first nymphs, Class II (27/153; 17.6%; Table 5.1), developed ectopic legs on the first and second abdominal segments. In general, the appendage on A1 was short and gnarled (Fig. 5.7F-G), while that on A2 was well-formed (Fig. 5.7F-G). In a similar regard, ectopic legs formed on the anterior abdominal segments in *Oncopeltus-nub*^{RNAi} first nymphs (Hrycaj et al., 2008) and in the remaining abdominal segments in *Tribolium-tal*^{RNAi} larvae (Savard et al., 2006). The emergence of this phenotype in *Periplaneta* might signify important and possibly conserved roles for *Pa-tal* and *Pa-nub* in regulating the appendage free abdomen so characteristic of insects.

DISCUSSION

Evolution of *tarsal-less* peptides

Homologues of *tarsal-less* have been identified in a diverse range of insect orders as well as a crustacean species (Appendix 12). With the exception of the single open reading frame found in a crustacean, most insect *tal* homologues are polycistronic having at least two smORFs. Phylogenetic analysis of the Type-A *tal* peptides proved difficult, as they vary in length, from 9 to 26 aa, and sequence composition with only seven amino acids (LDPTGXY) conserved both between and within the same species (Appendix 12A). Further complications arise in that there is no conservation in flanking sequences and even the spacing between smORFs is inconsistent, ranging from 2 bp between two ORFs in *Bombyx mori* (Lepidoptera) to 314 bp in *Apis mellifera* (Hymenoptera). On average, ORFs are spaced less than 100 nucleotides apart and are often out of frame with each other.

Moving up the evolutionary scale, from basal to derived, there is a trend towards acquiring additional copies of the Type-A Tal peptides (Fig. 5.8; Appendix 12A). While only one Type-A smORF has been identified in the crustacean *Daphnia pulex* (Cladocera) up to 11 have been identified in *Bombyx*, where a gene duplication has also occurred (Appendix 12A). Type-B peptides appear with the arrival of the Holometabola (Fig. 5.8) and share a semi-conserved sequence of 12aa in the middle of the peptide (Appendix 12C), though the reasons for this conservation are elusive, as the Type-B peptides have so far been shown to be non-functional in *Drosophila* (Galindo et al., 2007). The functional Type-AA Tal peptides occur only in the Diptera (Fig. 5.8) and are longer then Type-A and may have arisen through a point deletion or insertion which caused two nearby ORFs to become one, as these sequences have the LDPTGXY motif at both the N- and C-terminal ends (Appendix 12B).

Figure 5.8 shows a generalised phylogenetic tree based largely on Type-A Tal peptide consensus sequences generated manually by comparing the most frequently used amino acids from representative species within each Order. The *tal* gene is conserved at least from before the split of the insects and crustaceans, 440 million years ago. At the base of this pancrustacean tree, the general consensus sequence for the single Type-A Tal peptide is MXPKLDPTGLY. Conserved between the Crustacea and most of the hemimetabolous insects is a Leucine (L) as the penultimate amino acid (green letters in Fig. 5.8) and a two amino acid sequence of Proline (P) and Lysine (K) in the 5' end (PK; blue letters in Fig. 5.8). In most Hemipteran species, the 5' PK is either lost completely – giving rise to some of the shortest Tal peptides (9aa; Hemiptera-1) – or replaced with a variable number of amino acid residues (orange letters, Hemiptera-2; Fig. 5.8), providing some of the longest Tal peptides, up to 26aa in the pink mealybug, *Maconelliococcus hirsutus* (Appendix 12A). Moving up into the Holometabola, there is an increase in number of smORFs to three or more. In these organisms, the 5' PK is completely lost and there is variation in the second to last amino acid, often replaced with a Glutamine (Q) in the Hymenoptera and Coleoptera, but remaining highly variable in Lepidoptera and Diptera (Fig. 5.8).

It is easy to see that numerous changes occurred in the *tal* peptide sequence over a relatively short period of evolutionary time. While there are some weak similarities between closely related species outside of the conserved heptapeptide motif, the sequence is highly variable between orders, even within the same species. Although an amino acid change may only require one nucleotide substitution, these differences would often be deemed unlikely or uncommon based on their negative similarity scores in the BLOSUM and PAM substitution matrices. In addition, there is no apparent conservation of amino acid nature, i.e. hydrophobicity or polarity. This suggests that the highly conserved

LDPTGXY motif is most likely the functional sequence of this small peptide, though how it carries out its role remains uncertain. That it is so small and is expressed non-cell autonomously would make it reasonable to believe that *tal* acts as a signalling molecule where the variable N-terminus may be dispensable for its function.

Pa-tal functions in segment patterning and abdominal identity

Analysis of the *Periplaneta tal* homologue shows a different pattern of expression and functional domains compared to *Drosophila* and *Tribolium*. Strong Class III *Pa-tal*^{RNAi} embryos display an 'asegmental' phenotype and have drastically reduced expression of the segment polarity gene *Pa-en*, indicative of a dysfunctional segmentation mechanism. In the weaker Class I and II *Pa-tal*^{RNAi} embryos, stripes of Pa-*en* expression are misaligned, manifesting in the partially formed and fused segments observed in the hatched first nymphs. As a potential downstream target of N-signalling, *Pa-tal* may act in conjunction, or in parallel with other segmentation genes for proper patterning and segment formation.

The patterning defects seen in the *Pa-tal*^{RNAi} first nymphs are quite similar to the trunk anomalies recently illustrated in natural populations of the centipede *Stigmatogaster subterranean* (Lesniewska et al., 2009). Representatives of all arthropod subphyla and some annelids have shown this spiral cleavage, or helicomerism, but the genetic causes of these 'monsters' have not been identified (Cockayne, 1929; Curcic et al., 1983; Mitic et al., 2011; Morgan, 1892; Ramsay, 1959; Sobels, 1952). In 1952, Sobels began to study helicomerism phenotypes in adult flies and identified a potential genetic constituent, *Abnormal abdomen* (*aa* or *A*), associated with the "spiral" segment phenotype (Sobels, 1952). *Dm-aa* is X-linked but currently unannotated (FlyBaseID: FBgn0000009) and the particulars of its molecular functions are unknown. However, penetrance of the phenotype

is increased in *Dm-aa* heterozygotes when crossed with the transcription factor *Dichaete* (*D*; also known as *fish-hook*) (Sobels, 1952), a Sox-B class gene with many important developmental functions, including segment patterning (Nambu and Nambu, 1996; Russell et al., 1996). In the blastoderm, *Dm-D* regulates the expression of the primary pair-rule genes *eve*, *run*, and *hairy* (Russell et al., 1996) and, surprisingly, *Dm-D* interacts with *Dm-nub* to specifically regulate the expression of *Dm-eve* stripes 3-7 (Ma et al., 1997). *Dichaete* loss-of-function mutants lead to a range of segmentation defects including hemi-segment formation and segment fusions (Russell et al., 1996). These effects were amplified in *Drosophila-D/nub* double mutants (Ma et al., 1998), yet no segmentation defects have been noted when *Dm-nub* was lost on its own (Turchyn et al., 2011).

It is important to recall at this time that no segmentation defects were observed in *Dm-tal* mutants (Galindo et al., 2007; Kondo et al., 2007); while dorsal segment fusions and dorsal closure defects were noted in *Tribolium-tal*^{RNAi} and *Tc-svb*^{RNAi} larvae (Savard et al., 2006; Schnellhammer, 2012). Conversely, abdominal segment fusions were commonly observed in the *Periplaneta-tal*^{RNAi} first nymphs and occasionally in *Pa-nub*^{RNAi} nymphs. It is curious that double *Periplaneta-tal/nub*^{RNAi} and the *Drosophila-nub/D* mutant phenotypes are so similar. Perhaps *Pa-tal* and *Pa-nub* interact with a cockroach *Dichaete* homologue, assuming one exists, during embryonic segment patterning, which may similarly be involved in regulating pair-rule gene expression in the posterior GZ. While a relationship between *tal*, *nub*, and the pair-rule genes in *Periplaneta* is not yet known, such an interaction could explain the RNAi phenotypes where loss of *Pa-tal* and/or *Pa-nub* may result in faulty pair-rule patterning and, thus, misexpression of *Pa-en* stripes leading to segment fusion/deletion. On the other hand, the potential interaction between *tal* and *svb* in *Tribolium* during segment formation (Schnellhammer, 2012) implies a possible conserved role for these genes in *Periplaneta*. This suggestion is

further supported by the similarities in phenotypes between *Tribolium-svb*^{RNAi} and *Periplaneta-tal*^{RNAi} affected embryos, both of which display a bulging of abdominal segments and misalignment of stripes of the segment polarity gene *Tc-gooseberry* (Schnellhammer, 2012) and *Pa-en* expression, respectively. However, attempts at isolating a *Periplaneta* homologue of *svb* have so far been unsuccessful.

Additional phenotypes arose in *Pa-tal*^{RNAi} in the form of small ectopic 'appendages' on the abdomen, a phenotype not observed upon loss of *Pa-nub* expression alone. However, loss of both *Pa-tal* and *Pa-nub* resulted in the formation of nearly fully formed legs in the anterior abdominal segments. This result is interesting when taking into account that loss of *tal* in *Tribolium* (Savard et al., 2006) and loss of *nub* in *Oncopeltus* (Hrycaj et al., 2008) both resulted in the similar formation of ectopic abdominal legs. *Tribolium-tal* was shown to regulate several gap genes, which likely function upstream of the Hox genes *Ubx* and *AbdA*, important in maintaining the identity of this appendage-free tagma (Hughes and Kaufman, 2002c; Lewis et al., 2000; Savard et al., 2006). In *Oncopeltus*, *nub* has an important role in regulating expression of the Hox gene *AbdA* and, indirectly, the inhibition of ectopic limbs on the abdomen (Hrycaj et al., 2008). Ectopic leg formation was not observed in *Drosophila nub* or *tal* mutants (Galindo et al., 2007; Kondo et al., 2007; Turchyn et al., 2011), and expression of *Dm-Ubx* and *Dm-Distal-less* remain unaffected in *Dm-tal* mutant embryos (Galindo et al., 2007). In *Periplaneta*, RNAi knockdown of *Ubx* or *AbdA* results in ectopic abdominal limbs on A1 or on segments A1-A5, respectively (Chesebro, 2008); however an interaction between the Hox genes and *tal* or *nub* is not yet known.

Overall, these data indicate that *Periplaneta-tal* plays multiple roles in posterior segment formation and maintenance of abdominal identity during embryogenesis, functions that may be regulated by the N-signalling cascade. In the anterior-most

segments, *Pa-tal* could work outside of the N-pathway, as segmentation in the anterior gnathal and thoracic segments is N-independent (Chapter III), indicating that *Pa-tal* expression and function in this location may be influenced by some other, yet undetermined mechanism, perhaps maternal and/or anterior gap genes. The synergistic effects of *Pa-tal* and *Pa-nub* in abdominal patterning in the cockroach highlight several divergent mechanisms in the lineage leading to flies. While they may be required together to regulate Hox and pair-rule gene expression on *Periplaneta*, these interactions might have been lost or altered during the course of evolution and may function separately in *Oncopeltus*, *Tribolium*, and *Drosophila*.

Pa-tal and limb patterning

Jointed legs are a defining characteristic of arthropods. Our understanding of leg development in *Drosophila* is increasingly detailed while our knowledge in other arthropods remains comparatively basic. The Notch-signalling cascade is a conserved mechanism in leg segmentation from flies (Bishop et al., 1999; de Celis et al., 1998; Rauskolb and Irvine, 1999) to spiders (Prpic and Damen, 2009), suggesting that Notch-mediated joint formation could be the phylotypic trait of arthropods. I have explored the mechanisms of joint formation by analyzing leg-patterning genes in *Periplaneta*, including *Notch*, *Delta*, *nub*, and *tal*. The data presented here suggest that while *nub* and *tal* have complementary roles in *Drosophila* legs, these genes may be redundant in more basal species, underpinning the existence of variable gene functions downstream of a conserved Notch cassette controlling developmental boundaries in a variety of contexts.

In *Drosophila*, *Notch* is expressed in all leg segments and is required for proper growth and joint formation by defining sharp borders between podomeres in slightly different ways in the true joints and the tarsal sub-segments (Bishop et al., 1999; de Celis

et al., 1998; Pueyo and Couso, 2011; Rauskolb and Irvine, 1999). Those making up the true joints (coxa, trochanter, femur, and tibia) contain muscle attachments and here *N* and *Ser* regulate the expression of *odd-related* (*odd-r*) and *nubbin* at the distal end of each leg segment. In the tarsal sub-segments, *N* regulates the expression of *tal* that, in turn, modulates the expression of the active form of *svb*, leading to the intercalation and formation of tarsal joints. The interaction between *Dm-tal* and *Dm-svb* is similar to that described during denticle belt formation during embryogenesis (Kondo et al., 2010; Pueyo and Couso, 2011). Loss of *Dm-tal* expression leads to loss of tarsomere formation, while gain-of-function of *Dm-tal* leads to the development of ectopic joints, confirming the function of *tal* in joint formation (Galindo et al., 2007; Pueyo and Couso, 2011). In *Tribolium*, *tal* and *svb* are expressed in all larval leg segments and both *Tc-tal*^{RNAi} and *Tc-svb*^{RNAi} result in stubby legs with malformed joints (Savard et al., 2006; Schnellhammer, 2012), perhaps these genes play a role in each leg segment and are not relegated to one type of joint over another (true joint vs. tarsal). In *Periplaneta*, *tal* is similarly expressed in all leg segments, although gene expression knockdown does not result in a strong leg phenotype. *nubbin*, on the other hand, has a more profound effect on leg patterning, where loss of expression results in the fusion of adjacent podomeres in *Periplaneta*, *Gryllus*, and *Drosophila* (Turchyn et al., 2011).

The expression of *Pa-N* and *Pa-Dl* in all developing leg segments is conserved in arthropods (Angelini et al., 2012; Bishop et al., 1999; de Celis et al., 1998; Mito et al., 2011; Prpic and Damen, 2009; Rauskolb and Irvine, 1999); however, downstream targets of the N-pathway differ depending on location. In the true joints, Notch partially regulates *nub* expression and, along with *odd-r*, this leads to proper joint formation (Rauskolb and Irvine, 1999; Turchyn et al., 2011). In the tarsomeres, the initial activation of *tal* by Notch leads to the regulation of an active form of Svb that inhibits expression of *Dl*, thus forming

a negative feedback loop and a sharp Dl+/Dl- border at the location of future joint formation (Pueyo and Couso, 2011). Additionally, it has been shown that Notch signalling is also partially required for the expression of *nub* in *Periplaneta* and *Cupiennius* (Prpic and Damen, 2009; Turchyn et al., 2011). The reduction of the trochanter in *Pa-nub*^{RNAi} first nymphs is enhanced when both *nub* and *tal* expression are knocked down in double RNAi experiments, indicating that these genes work together in proper leg development. A similar relationship may play a role in segment patterning of the cockroach body during embryogenesis.

In flies, *Dm-nub* and *Dm-tal* have a complementary role in joint formation in the true joints or the tarsal segments, respectively (Fig. 5.9A). In phylogenetically basal organisms, *nub* and *tal* appear to be expressed in all developing leg segments (Fig. 5.9B) (Prpic and Damen, 2009; Savard et al., 2006; Turchyn et al., 2011). In *Periplaneta*, *Pa-nub* and *Pa-tal* appear to work synergistically for proper leg segmentation, hinting at a semi-redundant role of these genes in joint formation. These functions vary depending on the segment in question, i.e. trochanter. Overall, my studies in *Periplaneta* indicate that additional factors are at play during the growth and formation of leg segments that may or may not be dependent on the N-signalling pathway. *tal* may be downstream of N-signalling in the legs of both *Periplaneta* and *Drosophila*, but only functions in body segmentation downstream of N-signalling in the cockroach, while *tal* plays a significant role in *Tribolium* posterior segmentation through interactions with the gap genes.

While this non-canonical gene encodes highly conserved peptide sequences, expression and function is variable between species, yet in each case it plays a very important role in proper development and patterning in several developmental contexts. *Periplaneta* is now the third species in which *tal* has been examined, after *Drosophila* and *Tribolium*. *Pa-tal* could act as a link between patterning and morphogenesis (Galindo et

al., 2007), as *Dm-tal* functions in the morphogenic movements of several ectodermal tissues, including tarsomeres, trachea, and denticle belts. Overall, these studies highlight the multi-varied role of *tarsal-less* having very important developmental functions in body and leg patterning, even though it is such a small peptide. There is a clear sign that numerous smORF genes and gene families are waiting to be discovered, many of which may have a limited or subtle role in development, but some, like *tarsal-less*, may prove to be indispensible.

dsRNA	stage	Class I	Class II	Class III	wt
***Pa-tal*^{RNAi}**	embryo	224 (17.0%)	121 (9.2%)	115 (8.7%)	856 (65.1%)
	nymph	108 (20.2%)	43 (8.1%)	---	383 (71.7%)
***Pa-nub*^{RNAi}**	nymph	44 (26.3%)	6 (3.6%)	---	117 (70.1%)
***Pa-tal/nub*^{RNAi}**	nymph	50 (32.7%)	27 (17.6%)	---	76 (49.7%)

Table 5.1: Phenotypic series of *Pa-tal*^{RNAi}, *Pa-nub*^{RNAi}, and *Pa-tal/nub*^{RNAi} embryos and first nymphs. RNAi embryos and nymphs displayed a range of phenotypes, generally placed into two or three separate categories based on severity, with Class I being the weakest, Class II moderate, and Class III giving the strongest phenotypes. For a detailed description of the phenotypes displayed for each RNAi treatment, please refer to the main text.

Figure 5.1: Wild type expression of *tarsal-less* in *Periplaneta* is dynamic. Expression analysis of *Pa-tal* at post-blastoderm (A-D), germ band elongation (E-H), and after terminal differentiation (I-J). (A-B) At stage 4 (A) and stage 5 (B) post-blastoderm, *Pa-tal* is expressed in two wide stripes in the presumptive mandibles (*) and near the posterior end of the embryo (arrowhead). (C) These initial two bands remain at early stage 6 (* and arrowhead), while a new broad domain of expression is detected in the posterior (bracket). (D) By late stage 6, the anterior-most mandible stripe has faded (*) and the initial posterior stripe (arrowhead in A-C) is beginning to fade from the midline (white arrowhead). In addition, the broad posterior domain (bracket) is resolving into a new stripe of expression (arrow). (E-H) During germ band elongation, *Pa-tal* is expressed in a dynamic pattern. Initially expressed in a broad domain in the posterior GZ (bracket, A) at stage 8, this expression is refined into a wide stripe in the anterior GZ by stage 9 (follow progression of red arrowheads compared to black arrowheads). At the same time this new stripe emerges, the previously established stripe (*) dissipates as it moves anteriorly and a new segment begins to form. (I) After all posterior segments have been added and the terminal ends have differentiated, stage 16, *Pa-tal* is expressed is several new regions, including several bands in the developing appendages (black arrowheads), the distal tips of the pleuropodia (arrows), and the anal pads (*). (J) Later in development, stage 19, the stripes of expression in the appendages disappear and *Pa-tal* is now expressed in the apodemes at the base of the limbs (arrowheads) and new expression arises in the developing CNS (arrow).

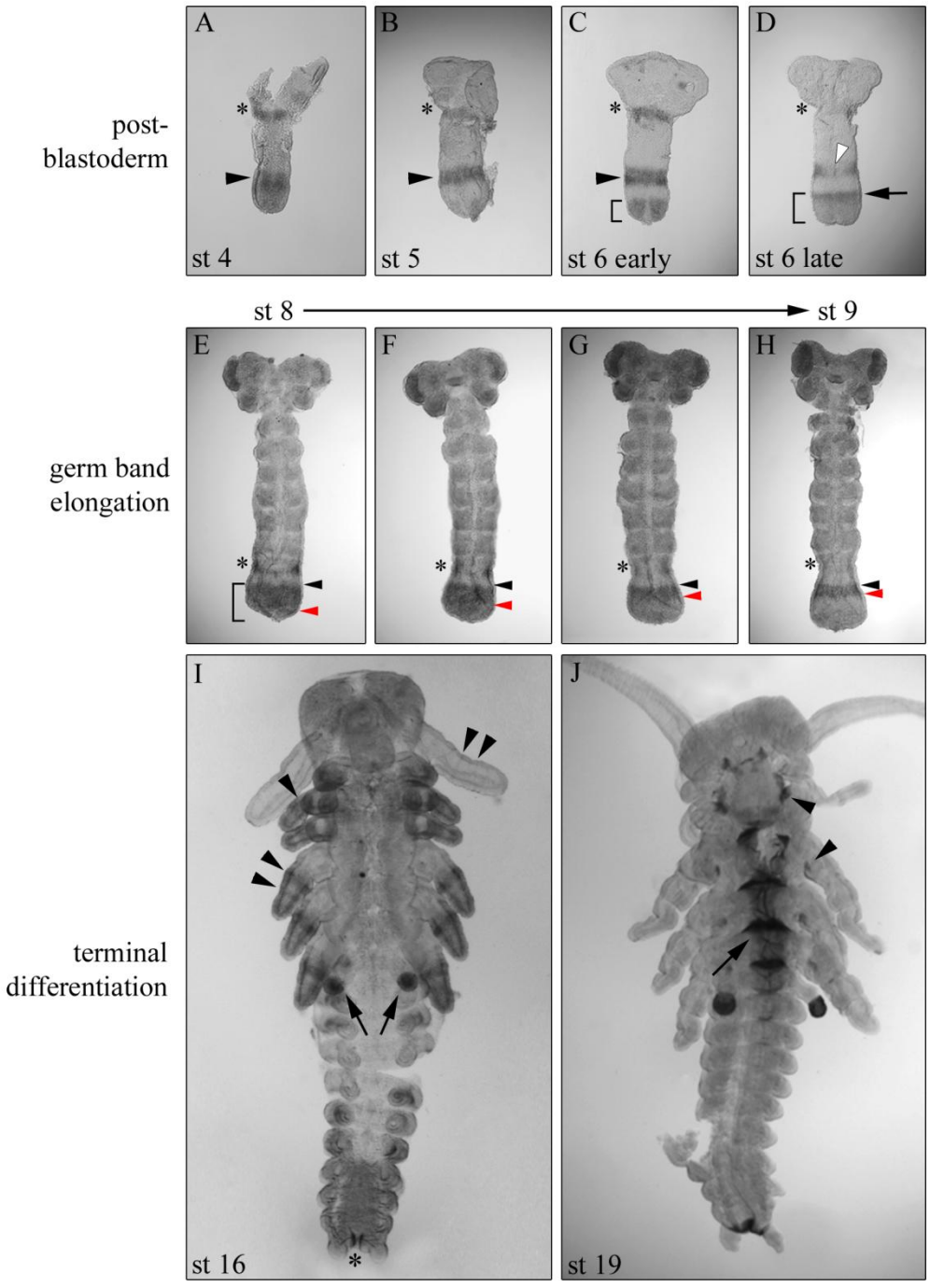

post-
blastoderm
germ band
elongation
terminal
differentiation
A
B
C
D
st 4
st 5
st 6 early
st 6 late
st 8
st 9
E
F
G
H
I
J
st 16
st 19

Figure 5.1

Figure 5.2: Effects of *Pa-tal*^{RNAi} on body patterning. (A) Wild type stage 10 embryo, unstained. (B) Weak Class I *Pa-tal*^{RNAi} showing partial fusion of some abdominal segments giving the embryo a slight bend to one side (arrow). (C) In moderate Class II *Pa-tal*^{RNAi} embryos, extensive segment fusion often leads to a 'bunching up' of adjacent segments (bracket). (D) Class III *Pa-tal*^{RNAi} embryos display an 'asegmental' phenotype. While posterior growth proceeds normally, there are no clearly defined segment borders (*) and weak attempts at limb bud formation give the embryo a wavy appearance (arrowheads). (E) Some asegmental Class III *Pa-tal*^{RNAi} embryos are greatly reduced in size compared to wild type and other RNAi affected embryos of a similar stage. The embryo shown is at stage 11 as indicated by the presence of the proctodeal pit (arrowhead) which forms at this stage. (F) Wild type first nymph showing normal development of head, thorax, and abdomen. (F') Cuticle prep of a dissected wild type first nymph abdomen; ventral to the left and anterior at the top. (G) Class I *Pa-tal*^{RNAi} first nymph with normally developed head and thorax, but with hemi-segments (black dots) and fusion between segments (*) in the abdomen. (G') Cuticle prep of a dissected abdomen of the *Pa-tal*^{RNAi} affected first nymph in G. Here it is easy to see the half-segment formation of abdominal segments A4 and A5, as well as the fusion between segments A4/A6. The fusion pattern is reflected in the dorsal (right) and ventral (left) sclerites; anterior is to the top. (J) Dissected wild type T2 leg with normal development of the five basic podomeres. (K) Dissected T2 leg of a Class II *Pa-tal*^{RNAi} first nymph showing fusion between the coxa/trochanter and a slight bend in the first tarsal subsegment (arrow). (L) In some Class II *Pa-tal*^{RNAi} first nymphs, an affected ventral thorax may result in the fusion at the base of neighbouring legs. (K) In an extreme example, loss of *Pa-tal* may lead to extremely malformed legs, including coxa/trochanter fusion, reduced femur and tibia, and absence of tarsomeres. A4, A5, A6 – fourth, fifth, sixth abdominal segment, respectively; cx – coxa; fe – femur; ta1-ta5 – first through fifth tarsal subsegment; ti – tibia; tr – trochanter.

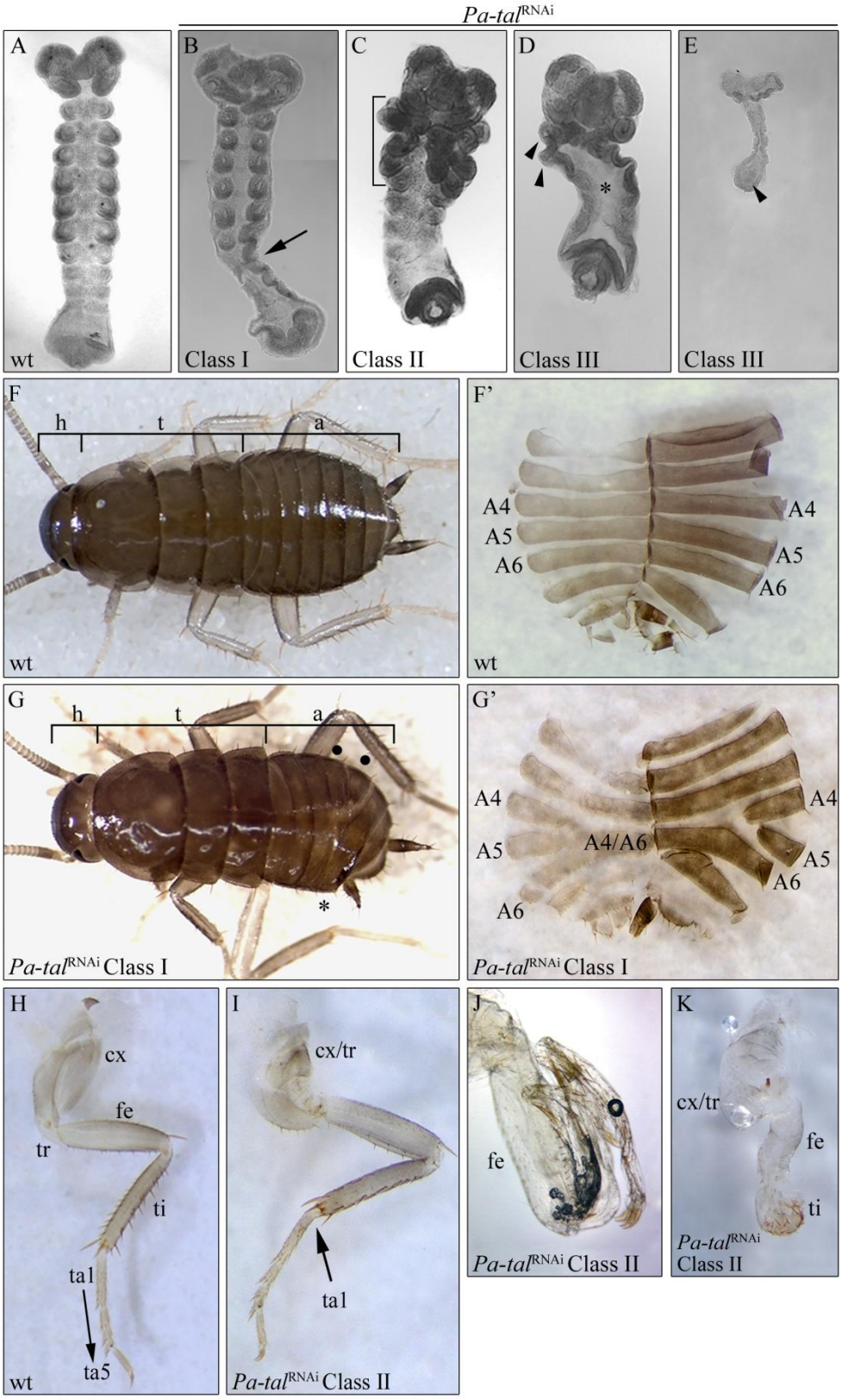

Pa-talRNAi
A
wt
B
Class I
C
Class II
D
Class III
E
Class III
F
h
t
a
wt
F'
A4
A5
A6
A4
A5
A6
wt
G
h
t
a
*
Pa-talRNAi Class I
G'
A4
A5
A6
A4/A6
A4
A5
A6
Pa-talRNAi Class I
H
cx
fe
tr
ti
ta1
ta5
wt
I
cx/tr
ta1
Pa-talRNAi Class II
J
fe
Pa-talRNAi Class II
K
cx/tr
fe
ti
Pa-talRNAi
Class II

Figure 5.2

Figure 5.3: *Pa-tal* affects and is affected by other segmentation genes during germ band elongation. (A) *Pa-en* expression in a wild type stage 10 embryo. (B) In Class I *Pa-tal*^{RNAi} embryos, *Pa-en* expression is normal in the anterior segments, but the stripes of expression in the abdomen are misaligned or only partially formed. (C) *Pa-en* expression in a wild type embryo. (D) In Class II *Pa-tal*^{RNAi} embryos, a similar misalignment and partial formation of *Pa-en* stripes occurs. (E) *Pa-en* expression in the 'asegmental' Class III *Pa-tal*^{RNAi} embryos is largely reduced to only a speckled pattern (arrowheads). (F) In wild type embryos, *Pa-cad* is expressed in a broad domain in the posterior GZ (bracket), except the tip (*). (G) This expression pattern remains unaffected in a Class III *Pa-tal*^{RNAi} affected embryo. (H) Posterior expression of *Pa-tal* is lost (*) in a Class A *Pa-cad*^{RNAi} embryo. (I-J) The wild type expression of *Pa-h* (I) in a broad posterior GZ domain (bracket) and several stripes in the anterior (arrowheads) remains unaffected in *Pa-tal*^{RNAi} (J). (K-M) Wild type expression of *Pa-Dl* (K) is in several stripes in the anterior GZ (arrowheads). This pattern of expression is unaffected in Class I (L) and Class II (M) *Pa-tal*^{RNAi} embryos. (N-O) The wild type expression of *Pa-tal* in the posterior GZ (*) does not form in *Pa-N*^{RNAi} affected embryos (*). A1, A3, A4, A5 – first, third, fourth, fifth abdominal segment, respectively; lb – labium; mn – mandible; mx – maxilla.

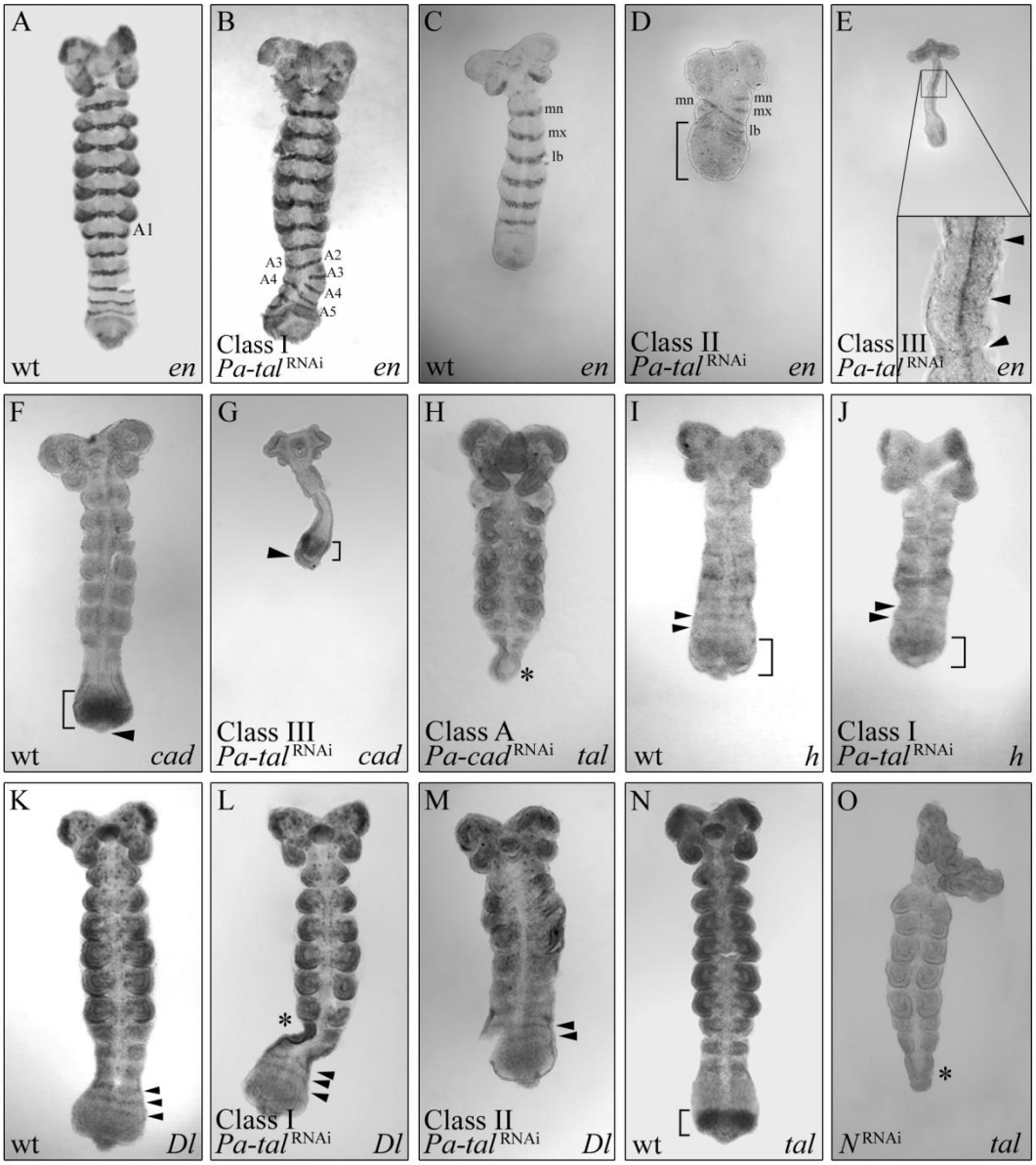

A
A1
wt
en
B
A2
A3
A4
A5
A3
A4
Class I
Pa-tal RNAi
en
C
mn
mx
lb
wt
en
D
mn
mn
mx
lb
Class II
Pa-tal RNAi
en
E
Class III
Pa-tal RNAi
en
F
wt
cad
G
Class III
Pa-tal RNAi
cad
H
Class A
Pa-cad RNAi
tal
I
wt
h
J
Class I
Pa-tal RNAi
h
K
wt
Dl
L
Class I
Pa-tal RNAi
Dl
M
Class II
Pa-tal RNAi
Dl
N
wt
tal
O
N RNAi
tal

Figure 5.3

Figure 5.4: Wild type expression patterns of *Pa-tal* and other leg segmentation genes. (A-D) Expression of *Pa-tal* in the developing embryonic legs. (A) At stage 11, *Pa-tal* is diffusely expressed in the distal half of the protruding limb buds (arrowheads). (B) At stage 14, *Pa-tal* is expressed broadly in the medial leg (bracket). (C) By stage 16, the future leg segments start to constrict and at this time, *Pa-tal* begins to clear from the middle of the leg (*) so that two bands of expression are observed (arrowheads). (D) At stage 19, furrows appear for all future leg segments and *Pa-tal* is expressed at the distal end of each presumptive podomere. (E-H) *Pa-nub* expression during leg development. (E) *Pa-nub* is expressed broadly in the distal half of the nascent limb buds at stage 10. (F) By early stage 13, *Pa-nub* is expressed as a wide band in the medial leg (bracket) and as a ring of expression in the proximal leg. (G) At late stage 13, three stripes of *Pa-nub* are detected (arrowheads) and there is diffuse expression in the distal tip (bracket). (H) At stage 17, podomere constrictions are visible showing that *Pa-nub* is expressed at the distal end of each leg segment. (I-L) *Pa-Dl* leg expression resembles *Pa-tal* during leg development. (I) At stage 12, *Pa-Dl* is expressed at the distal ends of the early limb buds. (J) *Pa-Dl* expression is expressed broadly in the medial leg, between two constriction points (bracket), which is cleared in the middle leaving two rings of expression in the middle of the leg (arrowheads, K). (L) As the separate podomeres are distinguished, *Pa-Dl* is detected at the distal ends of each segment, including several bands in the developing tarsus (open arrowheads). (M-O) *Pa-N* is expressed in a broad medial domain at stage 14 (bracket, M) and as several stripes by stage 16 (arrowheads, N). By stage 20 (O), *Pa-N* is detected at the distal end of all developing podomeres.

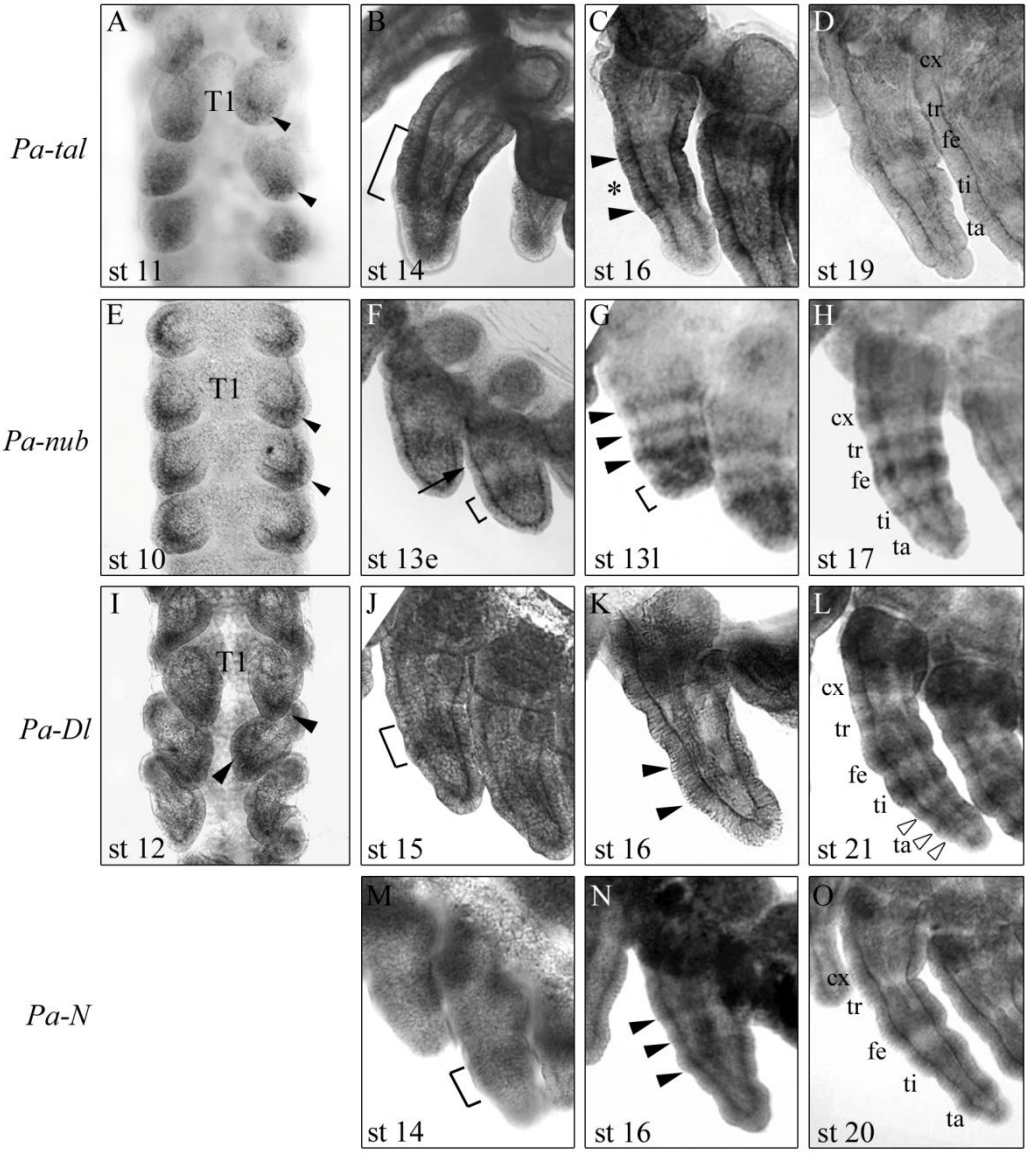

Pa-tal
Pa-nub
Pa-Dl
Pa-N
A
T1
st 11
B
st 14
C
*
st 16
D
cx
tr
fe
ti
ta
st 19
E
T1
st 10
F
st 13e
G
st 13l
H
cx
tr
fe
ti
ta
st 17
I
T1
st 12
J
st 15
K
st 16
L
cx
tr
fe
ti
ta
st 21
M
st 14
N
st 16
O
cx
tr
fe
ti
ta
st 20

Figure 5.4

Figure 5.5: *Pa-tal* and *Pa-nub* affect leg development. All images are of wild type or RNAi affected T3 legs of hatched first nymphs. (A) Wild type trochanter. (A') Wild type tarsus. Each tarsomere is a separate unit connected with a joint (arrowhead). (B-B') *Pa-tal*^{RNAi} does not affect development of the trochanter (B), but there is a noticeable bend in the first tarsal subsegment (arrow, B'). Tarsal joints do not appear to be affected (arrowhead). (C-C') The trochanter is reduced in *Pa-nub*^{RNAi} first nymphs (C), while the tarsomeres and respective joints (arrowhead, C') remain normal. (D-D') Loss of both *Pa-tal* and *Pa-nub* via RNAi result in a greatly reduced or even absent trochanter (D). The first tarsal subsegment is bent (arrow, D'), similar to *Pa-tal*^{RNAi} first nymphs (B'). There are mild effects in the formation of the other tarsomeres, but no noticeable defects in joint formation (arrowhead, D').

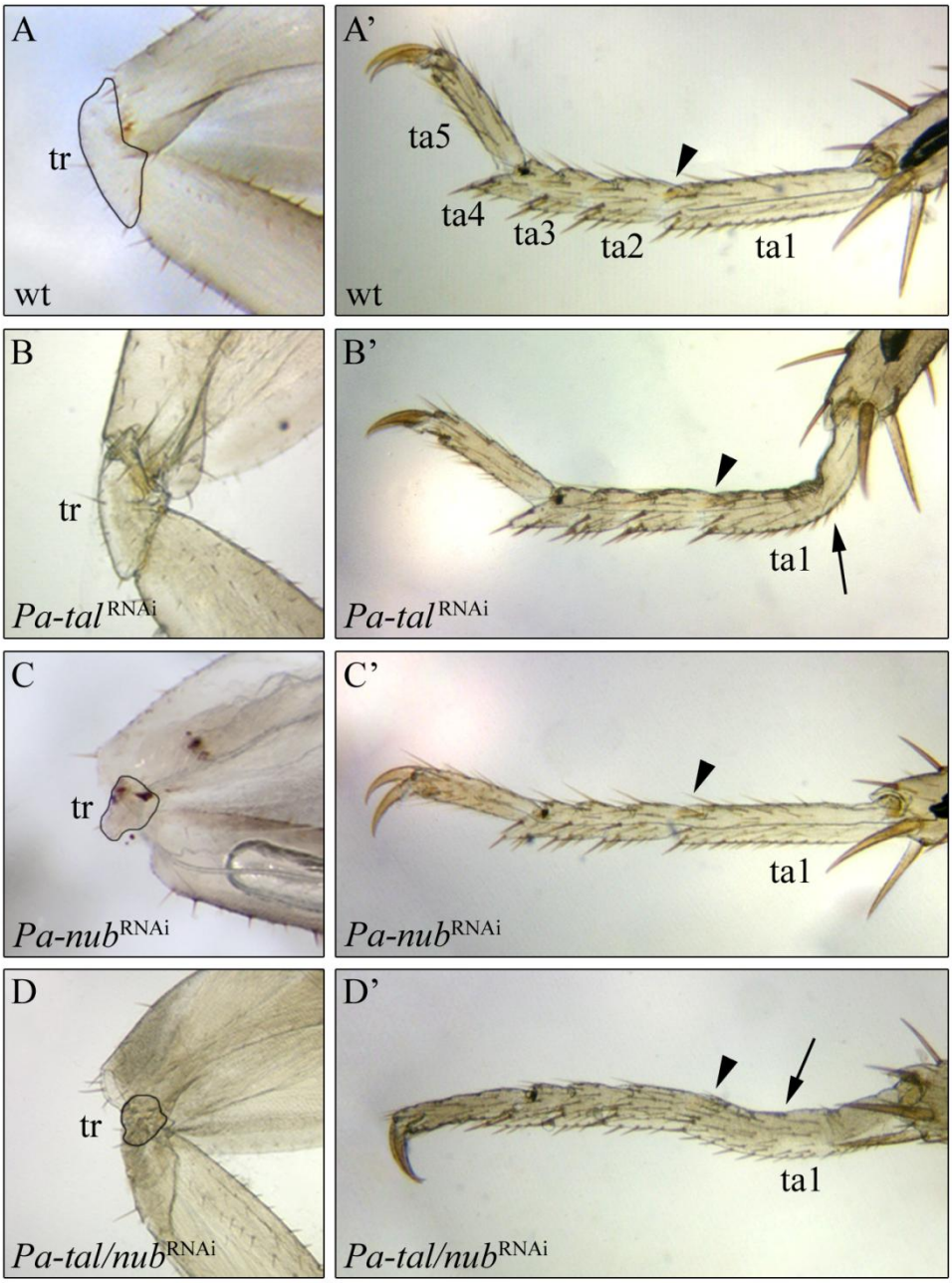

A
tr
wt
A'
ta5
ta4
ta3
ta2
ta1
wt
B
tr
Pa-tal RNAi
B'
ta1
Pa-tal RNAi
C
tr
Pa-nub RNAi
C'
ta1
Pa-nub RNAi
D
tr
Pa-tal/nub RNAi
D'
ta1
Pa-tal/nub RNAi

Figure 5.5

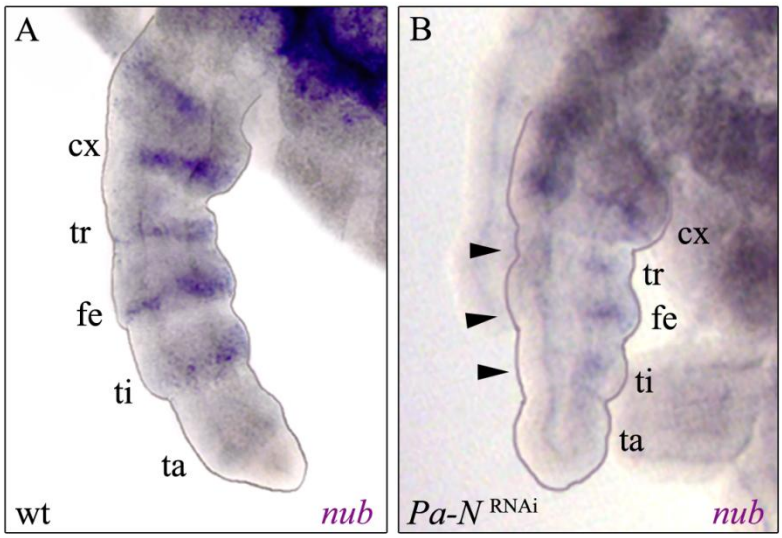


Figure 5.6: *Pa-nub* expression in the legs requires Notch-signalling. The wild type levels of *Pa-nub* expression at the distal portion of each leg segment (A) is significantly reduced, though not completely abolished, upon the depletion of *Notch* expression (B); most notable in the tr, fe, and ti (arrowheads).

Pa-nub expression analysis performed by N. Turchyn and A. Popadić (Wayne State University, Detroit, USA); figure reproduced from (Turchyn et al., 2011).

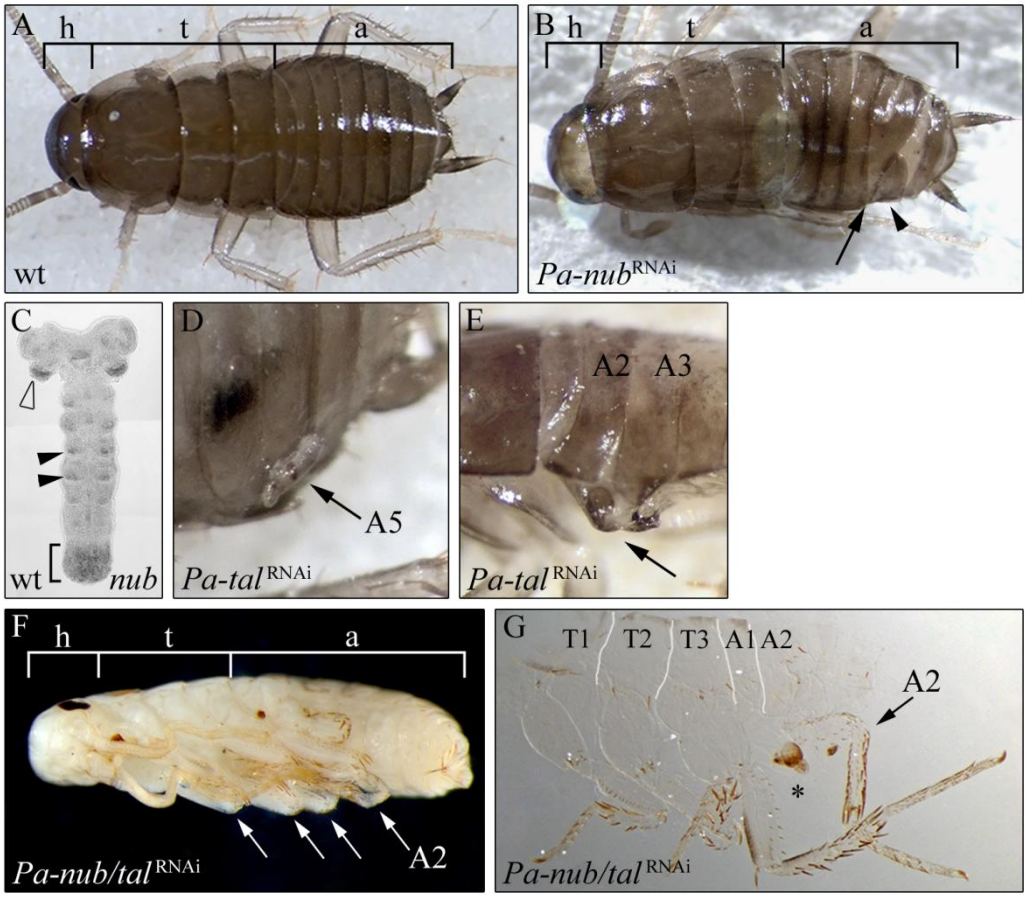


Figure 5.7: *Pa-nub* affects abdominal patterning in conjunction with *Pa-tal*. (A) Wild type first nymph showing normal segment development of the head, thorax, and abdomen. (B) In *Pa-nub*^{RNAi} affected first nymphs, the head and thorax form properly, while some of the abdominal segments are fused (arrow) or only partially formed (arrowhead). (C) Wild type expression of *Pa-nub* at stage 8 is in a broad domain in the posterior GZ (bracket) and the early limb buds (arrowhead) and antennae (open arrowhead). (D-E) In a few *Pa-tal*^{RNAi} first nymphs, ectopic growths (arrows) appear on the ventral (D) and lateral (E) abdomen. (F) Double knockdown phenotype of both *Pa-tal* and *Pa-nub* results in the formation of ectopic legs on the anterior abdomen. (G) Cuticle prep of unhatched nymph shown in F, showing the ectopic tissue on A1 (*) and the almost fully formed leg on A2 (arrow).

Figure 5.8: Evolution of *tarsal-less* in the arthropods. This phylogenetic tree is largely based on the presumptive Type-A Tal peptide sequences (Appendix 12) used to manually generate a consensus for each insect or crustacean Order. At the base of the tree is a single copy of *tal* (Type-A ORF) with the sequence MXPKLDPTGLY; coloured letters representing those amino acids liable to change and featured at the branch nodes of the tree. Changes are most common at the penultimate amino acid (green letters); the N-terminal PK (blue) found in basal hemimetabolous insects; and a variable number of amino acid changes, additions, and deletions in the N-terminal end of the peptide (red and orange letters).

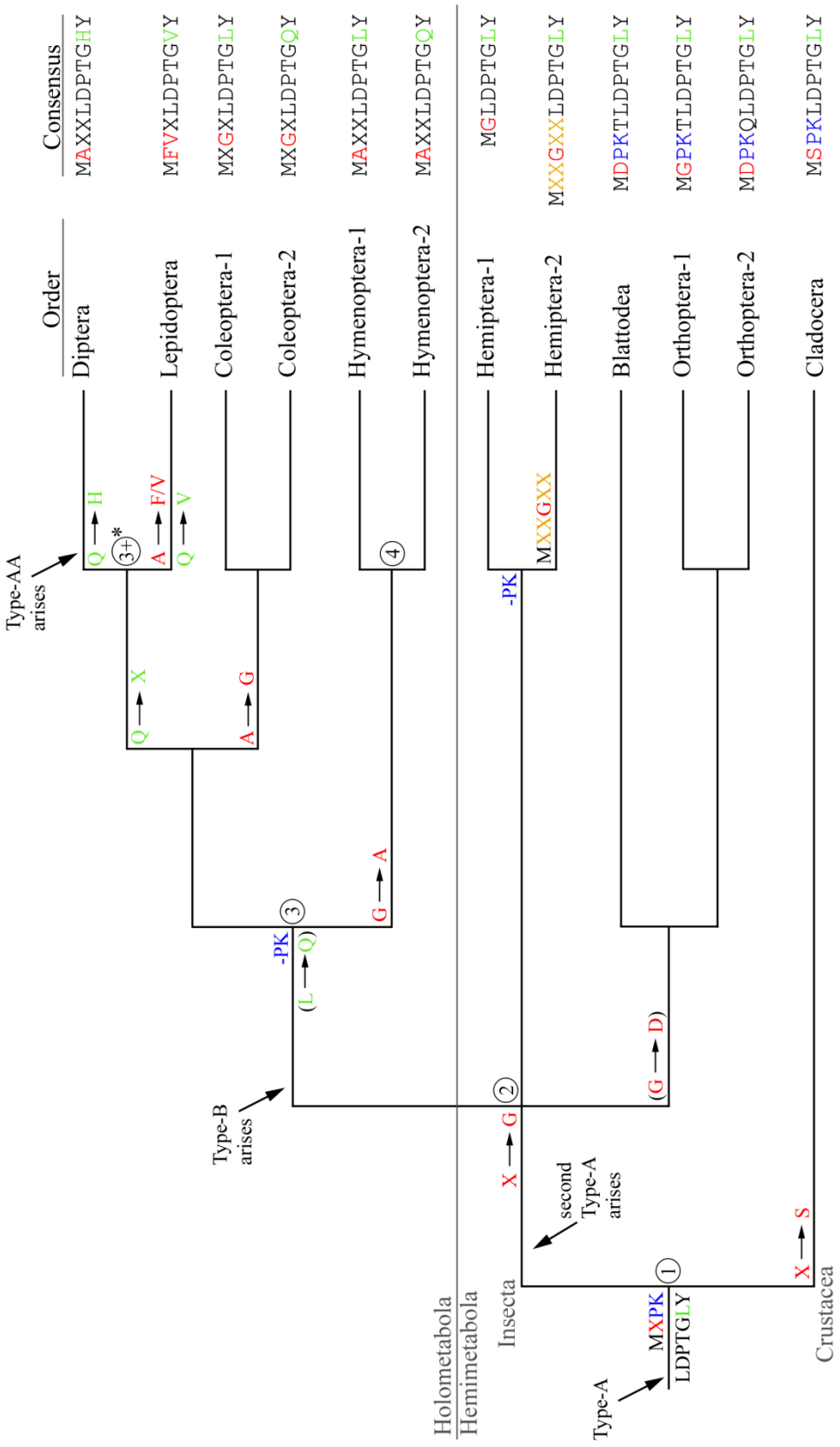

Order
Consensus
Diptera
MAXXLDPTGHY
Lepidoptera
MFVXLDPTGVY
Coleoptera-1
MXGXLDPTGLY
Coleoptera-2
MXGXLDPTGQY
Hymenoptera-1
MAXXLDPTGLY
Hymenoptera-2
MAXXLDPTGQY
Hemiptera-1
MGLDPTGLY
Hemiptera-2
MXXGXXLDPTGLY
Blattodea
MDPKTLDPTGLY
Orthoptera-1
MGPKTLDPTGLY
Orthoptera-2
MDPKQLDPTGLY
Cladocera
MSPKLDPTGLY
Type-AA arises
Q → H
3+*
A → F/V
Q → V
Q → X
A → G
G → A
4
Type-B arises
-PK
(L → Q)
3
-PK
MXXGXX
(G → D)
X → G
2
second Type-A arises
Type-A
MXPK
LDPTGLY
1
X → S
Holometabola
Hemimetabola
Insecta
Crustacea

Figure 5.8

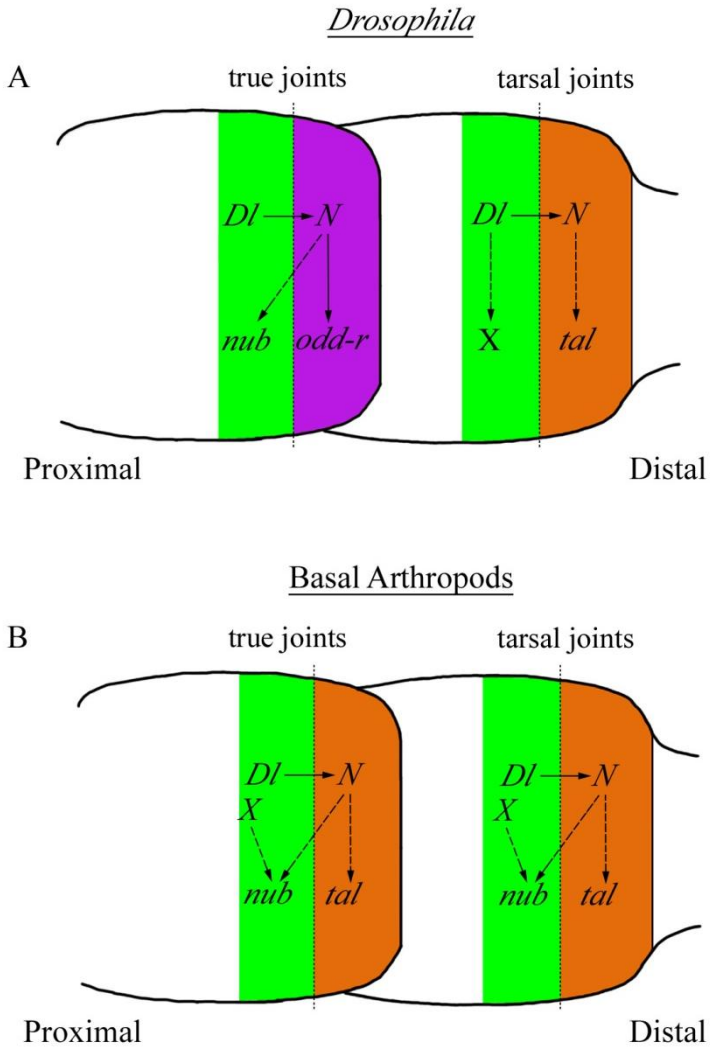


Figure 5.9: Models of leg development in derived versus basal arthropods. N-mediated joint formation could be a phylotypic trait in arthropods as its expression in all putative joints is conserved in organisms studied to date, from flies (Bishop et al., 1999; de Celis et al., 1998; Rauskolb and Irvine, 1999) to spiders (Prpic and Damen, 2009). However, downstream targets of the N-pathway differ depending on location. (A) In flies, *nub* and *tal* have a complementary role in joint formation in the true joint or the tarsal segments, respectively. (B) In phylogenetically basal organisms *nub* and *tal* appear to be expressed in all developing leg segments (Prpic and Damen, 2009; Savard et al., 2006; Turchyn et al., 2011). In *Periplaneta* depletion of *nub* or *tal* results in a minor shortening of leg segments, while in combination *Pa-tal/nub*^{RNAi} has a synergistic effect on joint development, hinting at a partially redundant role of these genes in joint formation. Modified from Turchyn et al. (2011).

CHAPTER VI

DISCUSSION

Results from my work on the cockroach *Periplaneta americana* helps to shed light on the conserved and divergent mechanisms of segmentation between distantly related phyla. In the preceding chapters, various mechanisms used during different phases of embryonic development have been discussed pertaining to *Periplaneta* anterior and posterior body segmentation, as well as segmentation of the legs. There is a conserved mechanism involved in the early establishment of a posterior Wnt-Cad organiser at post-blastoderm that is important in defining the posterior axis and setting up the GZ from which new segments will sequentially emerge during germ band elongation. The genes involved in posterior growth and segmentation form an intimately linked network required for the proper maintenance and continuation of these coupled processes throughout development. At the heart of this network, the core genes (*Wnt*, *caudal*, *Dl*, *N*, and *hes/her*) are conserved among most segmented bilaterians, while downstream targets are variable and divergent, highlighting the flexible nature of this mechanism. There are also lineage-specific mechanisms that have developed in each phylum, such as the pair-rule genes in arthropods. Although variably expressed, the pair-rule genes are required for proper patterning and segmentation in all arthropods examined. Finally, the recently discovered smORF gene *tarsal-less* shows conserved and divergent functions in both body and appendage development. In the sections below, I will briefly summarise and discuss further the evolutionary implications of the segmentation mechanisms shared between the segmented bilaterians.

Conservation of posterior patterning mechanisms

The first step in embryonic patterning is establishing the anterior-posterior (AP) axis, which often involves the posterior determinants *Wnt* and *caudal*, a mechanism common to most metazoans including Porifera and Cnidarians (Adamska et al., 2011; Adamska et al., 2007; Martin and Kimelman, 2009; Ryan and Baxevanis, 2007). Most organisms exhibiting posterior growth generally do so through a Wnt-controlled organiser, involving Wnt regulation of *caudal* expression, required to establish a functional growth zone in the posterior from which sequential segmentation will proceed (Martin and Kimelman, 2009; McGregor et al., 2009). Results presented in Chapter III show that *Periplaneta*-Wnt1 activates zygotic *Pa-cad* expression at the post-blastoderm stage and this relationship is maintained throughout germ band elongation, thus sustaining a functional GZ. Loss of either *Pa-Wnt1* or *Pa-cad* expression results in a failure to properly establish the organiser and GZ, thus leading to severe truncations, similar to what is observed in vertebrates (Martin and Kimelman, 2009; Shimizu et al., 2005; van de Ven et al., 2011). It is still unclear how these genes function in posterior patterning and segmentation in the annelids, though some studies are indicative of a conserved role in these processes (Cho et al., 2010; de Rosa et al., 2005).

Like Wnt, N-signalling also developed very early in animal evolution, coinciding with the development of multicellularity, as a necessary means of cell-to-cell communication (Gazave et al., 2009; Richards and Degnan, 2012; Richards et al., 2008; Technau et al., 2005). Throughout the course of evolution the N-signalling pathway increased in complexity and function, including regulating cell proliferation and differentiation, and by the dawn of the Bilateria acquired a new role in organising borders and segment patterning (Artavanis-Tsakonas et al., 1999; Gazave et al., 2009). N-mediated segmentation is the most common method used by posterior segmenting bilaterians,

including such divergent phyla as vertebrates (reviewed by Dequeant and Pourquie, 2008), annelids (Rivera and Weisblat, 2009; Thamm and Seaver, 2008), and arthropods (Kadner and Stollewerk, 2004; Mito et al., 2011; Pueyo et al., 2008; Stollewerk et al., 2003; Williams et al., 2012). In each of these phyla, dynamic N-signalling involves activation of the transcription factor *hes/her* and waves of expression travel through the posterior GZ/PSM, eventually forming stripes in the anterior, demarcating the future segment/somite boundaries. In *Periplaneta* (Chapter III; Pueyo et al., 2008) and other segmented organisms (Palmeirim et al., 1997; Rivera and Weisblat, 2009), loss of *N* or *hes/her* expression leads to disruptions in proper segment formation and/or truncations of the developing embryo, indicating a conserved requirement for N-signalling in posterior patterning.

Coupling posterior growth and segmentation

Although recently contested (Kainz et al., 2011; Mito et al., 2011; Williams et al., 2012), results presented in Chapter III provide clear evidence that the processes of Wnt1-Cad posterior growth and Dl/N segmentation are intimately linked during *Periplaneta* development. In the cockroach embryo at late post-blastoderm, posterior *Pa*-Wnt1 signalling positively regulates the expression of *Pa-Dl* in the posterior tip, which feeds back to maintain *Pa-Wnt1* expression, and subsequently that of *Pa-cad*, thus sustaining a functional GZ. *Pa-cad* expressing cells in the mid-GZ are kept in an undifferentiated state by repressing their ability to respond to the travelling waves of Dl/N-signalling until the signal is outside of the *Pa-cad* domain. Indeed, loss of *Pa-cad* expression results in a failure of *Pa-Dl* stripe and segment formation, possibly leading to premature determination of the posterior embryo. The regulatory relationships between Wnt/Cad and Dl/N/H, along with the similar phenotypes observed upon loss of expression of any one of

these genes, provides support to the coupling of posterior growth and segmentation within an intimately linked genetic network in this basal insect. In other short germ band arthropods, loss of N-signalling leads to varied phenotypes that generally include loss of segment formation either with (*Gryllus* – Mito et al., 2011; *Parhyale* – O'Day, 2006; *Achaearanea* – Oda et al., 2007) or without (*Cupiennius* – Schoppmeier and Damen, 2005b; Stollewerk et al., 2003; *Artemia* – Williams et al., 2012) concomitant loss of posterior elongation. Furthermore, in *Gryllus* and *Achaearanea* N-signalling is required for *cad* expression and in setting up a proper GZ, thus supporting a conserved coupling of growth and segmentation in arthropods (Mito et al., 2011; Oda et al., 2007). These data illustrate that even between closely related species, the mechanisms of segmentation can be divergent.

In the Arthropoda, posterior expression of *cad* may play a conserved role in regulating the pair-rule genes in both long germ band (Olesnicky et al., 2006; Wilson et al., 2010a) and short germ band (Bolognesi et al., 2008; Copf et al., 2004; Shinmyo et al., 2005) insects. Whether *cad* is involved in regulating the initial activation or later stripe formation, as with *Dl* and *h* in *Periplaneta*, remains to be determined, but either way this represents another level of interaction between the growth and segmentation mechanisms. In addition, further support comes from *Tribolium* in which posterior Wnt-signalling is involved in regulating *Tc-eve* expression in the posterior GZ (Bolognesi et al., 2008). Functional control of the pair-rule genes in *Periplaneta*, and other arthropods, remains to be elucidated and may be regulated through either Wnt-Cad or N-signalling, or both mechanisms.

In vertebrates, the processes of posterior elongation and segment formation are inextricably linked and involve three main signalling pathways: Wnt, N, and FGF (Dubrulle and Pourquie, 2004a; Gibb et al., 2010; Goldbeter and Pourquie, 2008). For

example, the FGF-signalling pathway has two major roles during somitogenesis: maintenance of the posterior PSM and anterior somite positioning. FGF maintains the posterior PSM in an undifferentiated state by regulating cellular responses to N-signalling (Aulehla and Pourquie, 2010; Dubrulle et al., 2001; Dubrulle and Pourquie, 2004a), thus allowing for continued posterior growth. In addition, the posterior-to-anterior gradient of FGF aids in proper positioning of the future somites in the anterior PSM by acting as a wavefront where the temporal waves of N-signalling are translated into a spatially reiterated pattern, leading to somite boundary formation (Dubrulle et al., 2001; Niwa et al., 2011; Sawada et al., 2001). While the function(s) of FGF-signalling have not been extensively studied in arthropod segmentation (Beermann and Schroder, 2008; Huang and Stern, 2005), this result is intriguing and suggestive, though still speculatively, that FGF-signalling could play a similar role in arthropod posterior patterning and segmentation. It is curious that in *Periplaneta*, the *Pa-cad* expression domain functions in a similar manner as vertebrate FGF by repressing the readout of N-signalling in the posterior GZ so that a defined group of cells can only respond to form segments when outside of this *Pa-cad* domain. Perhaps in short germ band arthropods, *Pa-cad* acts as a sort of "wavefront" controlling the spatial readout of temporal N-signalling and acting as an integrator between posterior growth and segmentation.

Pair-rule patterning

Broadly studied in *Drosophila*, the pair-rule genes play a key role in the long germ band segmentation paradigm: maternal → gap → pair-rule → segment polarity. In lower arthropods, pair-rule gene expressions can be quite variable, and may not fit easily into such a hierarchic segmentation paradigm. However, there is conservation of the mini-hierarchy of pair-rule gene expressions, where the primary pair-rule genes (*eve, run*) are

expressed first and regulate the expression of the secondary pair-rule genes (*prd, slp*). The upstream regulation of the pair-rule genes is still a bit unclear and may depend on when and where expression occurs during embryogenesis. In short germ band organisms there are two distinct phases of development in which the anterior segments are patterned first during the early blastoderm/post-blastoderm stages, while the remaining posterior segments are added sequentially from the extending posterior GZ during later stages of development (Davis and Patel, 2002; Liu and Kaufman, 2005b). That the patterning of the anterior (gnathal, thoracic) and posterior (thoracic, abdominal) segments utilise separate mechanisms, highlights the differences between these two environments during embryogenesis (Dearden et al., 2002; Schoppmeier and Damen, 2005a). In *Periplaneta*, N-signalling is an essential mechanism for posterior segment patterning, but is not required for specification of the more anterior head and thoracic segments. However, one method for proper patterning utilised by both anterior and posterior segments in most arthropods is the pair-rule genes.

As discussed in Chapter IV, *Periplaneta* pair-rule genes have stronger effects on the anterior segments and may be regulated upstream by the gap genes present in this area; however, further analysis is required as gap genes have not been studied in the cockroach. In fact, gap genes are not very well studied in short germ band organisms in general and show variable expression and functions that may be equally varied depending on anterior or posterior expression (Bucher and Klingler, 2004; Cerny et al., 2005; Liu and Patel, 2010; Mito et al., 2005; Schwager et al., 2009). Several studies have even revealed a gap-like function for *even-skipped* in both *Gryllus* (Mito et al., 2007) and *Oncopeltus* (Liu and Kaufman, 2005a), including the regulation of other gap genes, stressing the variable and flexible nature of these genes during segmentation processes. These studies suggest the pair-rule genes may have larger and more varied roles in basal organisms, but in the

lineage leading to *Drosophila*, their functions were restricted when they came under the control of the gap genes, as more segments became pre-patterned in the blastoderm (Damen, 2007; Peel and Akam, 2003). However, a recent study has revealed an early gap-like role for *Drosophila-runt* in regulating *bicoid* target genes (Chen et al., 2012) and previous studies have shown a gap function for *sloppy-paired* in the pre-gnathal head segments (Andrioli et al., 2004; Grossniklaus et al., 1992), signifying that any one gene can have multiple functions and may not be so easily relegated into a single category.

During posterior segmentation, *Periplaneta eve* and *run* have apparent redundant roles, whereas *Pa-hairy* plays a major role in proper segment formation. *hairy* is of particular interest as it is the only known pair-rule gene to function during segmentation in arthropods, annelids, and vertebrates. In short germ band arthropods, *hairy* expression is regulated by N-signalling and presumably leads to the indirect regulation of *en* expression (O'Day, 2006; Pueyo et al., 2008; Schoppmeier and Damen, 2005b; Stollewerk et al., 2003). Whether the other pair-rule genes are regulated as part of the N-signalling pathway during germ band elongation remains to be determined, but it could be possible. In *Periplaneta*, and other arthropod species (Choe et al., 2006; Damen et al., 2000; El-Sherif et al., 2012; Janssen et al., 2011; Liu and Kaufman, 2005a; Mito et al., 2007), the expression of *eve* and *run* in the posterior GZ is also dynamic and transient, resulting in stripes of expression at the anterior. The established interactions between N-signalling and *h*, combined with the similar expression patterns of the other pair-rule genes in the posterior GZ, suggests that N-signalling could be involved in regulation of pair-rule gene expression in basal arthropods, a function subsequently lost in the lineage leading to the holometabolous insects in which gap genes have become the major regulator of the pair-rule genes (Damen, 2007; Peel, 2004; Peel and Akam, 2003).

Even in *Tribolium*, a short germ band holometabolous insect, Dl/N-signalling does not appear to have a function in segmentation (Aranda et al., 2008; Tautz, 2004). Instead *Tribolium* utilizes a primary pair-rule gene circuit (*eve* → *run* → *odd* | ~~*eve*~~) that is dynamic and cyclic (Choe et al., 2006; El-Sherif et al., 2012; Sarrazin et al., 2012) and could represent an important change in pair-rule gene regulation upon the loss of N-signalling, yet retaining a still unknown oscillatory/synchronisation mechanism necessary for sequential segment patterning in a cellular environment. With this in mind, the dynamic expression of the pair-rule genes in basal arthropods could, alternatively, be indicative of a conserved pair-rule gene circuit. If this mechanism did exist, perhaps it runs in parallel with the N-segmentation pathway, which could explain why posterior phenotypes were rarely observed in *Periplaneta* pair-rule gene RNAi, except *Pa-h*^{RNAi}. This would also indicate that the N-segmentation mechanism has more relevance than the pair-rule genes, whether in circuit or not, during posterior segment patterning in the cockroach. Further work is required to determine if 1) the pair-rule genes are regulated by N-signalling, 2) a double-segmental, oscillatory pair-rule gene circuit exists outside of *Tribolium*, and 3) these two mechanisms are linked or run parallel to each other during germ band elongation in short germ band arthropods.

Orthologues of the arthropod pair-rule gene *hairy* (*hes/her*) play a significant role in body patterning in most segmented organisms. Orthologues of the other pair-rule genes exist as well, but mainly play a conserved and presumed ancestral role in neurogenesis (Avaron et al., 2003; de Rosa et al., 2005; Patel et al., 1992; Patel et al., 1989; Seaver et al., 2012; Song et al., 2002). Interestingly, the pair-rule gene *even-skipped* is expressed in the posterior GZ/PSM of annelids (de Rosa et al., 2005; Song et al., 2002; Song et al., 2004) and vertebrates (Bastian and Gruss, 1990; Beck and Slack, 1999; Dush and Martin, 1992; Joly et al., 1993; Seebald and Szeto, 2011) where it functions as a transcriptional

repressor in posterior growth, but not in the actual segmentation mechanism. In *Xenopus*, *eve/Xhox3* is downstream of N, Wnt3a, and FGF signalling (Beck and Slack, 1999; Pownall et al., 1996), while *eve1* in zebrafish is regulated by Wnt and FGF signalling where it is a part of the posterior organiser (Cruz et al., 2010). However, in the Arthropoda, *eve* and the other pair-rule genes acquired a new role in the regulation of segment patterning and formation.

Studies on pair-rule genes in other non-model arthropods often focus on posterior expression. However, inferring function from expression is not always straightforward. For example, expression of *Tribolium-hairy* in the posterior GZ would clearly suggest a pair-rule function in posterior segment patterning (Eckert et al., 2004; Sommer and Tautz, 1993). Nevertheless, *Tc-h*^{RNAi} did not reflect this presumed function, as segmentation phenotypes were only observed in the anterior (Aranda et al., 2008). Other than a handful of examples, the functions of the pair-rule genes have not been fully studied in more basal arthropods (tools for which may not yet be available). As technology improves, we will be able to fully examine these genes in other non-model species in hopes of determining an ancestral mechanism for pair-rule genes in segmentation.

A new segmentation gene: *tarsal-less*

To date, putative *tal* homologues have only been identified in one crustacean and numerous hemimetabolous and holometabolous insect species (Appendix 12) (Galindo et al., 2007; Savard et al., 2006), but as more genomes become sequenced it will be easier to determine the origin and possible ancestral functions of this gene. *tal* appears to be highly evolvable, having been duplicated numerous times during insect evolution with a trend towards increased copy number moving up from basal to more derived insect species. While the C-terminal LDPTGXY motif remains conserved, the N-terminus is highly

variable in length and amino acid composition. *tal* function is crucial to several developmental processes in those organisms in which it has been studied (Galindo et al., 2007; Kondo et al., 2007; Savard et al., 2006).

Data from Chapter V shows that *Periplaneta-tal* has a dynamic pattern of expression throughout embryogenesis and functions in both body and appendage development, apparently downstream of the N-pathway. Additionally, *Pa-tal* works in conjunction with *Pa-nubbin* to maintain the limbless abdomen, a function previously shown separately for *Tribolium-tal* (Savard et al., 2006) and *Oncopeltus-nubbin* (Hrycaj et al., 2008). In *Periplaneta*, *Tribolium*, and *Gryllus*, *tal* is expressed in all putative leg joints (Savard et al., 2006); whereas, *Drosophila-tal* is only expressed in the tarsomeres where it is required for proper growth and joint formation (Galindo et al., 2007; Pueyo and Couso, 2008; Pueyo and Couso, 2011). In *Pa-tal*^{RNAi} first nymphs, only minor changes in length and bending of tarsomere-1 were observed, although there were no effects on joint formation. However, the number of leg defects increased upon loss of both *Pa-tal* and *Pa-nub*, suggesting a relationship between *Pa-tal* and *Pa-nub* in leg development, as with body segmentation.

Additionally, single *Pa-tal*^{RNAi} resulted in asymmetric fusions between abdominal segments. A similar phenotype has also been noted in *Drosophila Dichaete* mutants (Russell et al., 1996). *Dm-D* regulates the expression of the pair-rule genes *eve, run*, and *h* (Russell et al., 1996), a function that may be carried out through an interaction with *Dm-nub* (Ma et al., 1998). Perhaps the association between *tal* and *nub* in *Periplaneta* also involves interaction with a *Dichaete* homologue, should it exist, for proper segment patterning. Alternatively, as recently shown in *Tribolium*, an interaction between *tal* and *svb* may be required for segment formation during germ band elongation (Schnellhammer, 2012), an interaction also required for proper denticle belt formation in *Drosophila*

((Galindo et al., 2007; Kondo et al., 2010). As *tal* seems to act downstream of N-signalling in *Periplaneta*, either of these options may add a new level of complexity in pair-rule gene regulation and the segmentation gene network in short germ band organisms. Between the three species in which *tal* function has been studied there are significant similarities and differences in body and leg patterning. Clearly, further examination in a range of insects and other arthropods is required in order to determine the ancestral role of this gene and when it may have arisen during the course of arthropod evolution.

Redundancy, flexibility, and divergence of segmentation mechanisms

During vertebrate somitogenesis there is considerable redundancy between the three major signalling pathways, Wnt/N/FGF, that interact with and regulate the expression of one another as well as similar downstream target genes (Cinquin, 2007; Gibb et al., 2010; Oates et al., 2012). These three pathways can compensate for each other to a certain extent, which may explain why loss of function in one pathway often leads to a breakdown of somitogenesis, but not a complete loss of segmentation (Dequeant et al., 2006; Dequeant and Pourquie, 2008; Ozbudak and Pourquie, 2008; Riedel-Kruse et al., 2007). Two examples of this redundancy include the regulation of *cad/Cdx* expression by Wnt and FGF signalling (Ikeya and Takada, 2001; Lohnes, 2003; Savory et al., 2009) and the regulation of *hes/her* by both N and FGF signalling, which in turn feeds back to regulate each of their expressions as well (Kageyama et al., 2012; Niwa et al., 2011).

The level of complexity within the vertebrate segmentation clock has not been as thoroughly analyzed in other organisms that employ N-mediated segmentation. Thus, redundancy at the core has not yet been determined, but there are additional factors to consider, such as the pair-rule genes that have gained new roles in arthropod

segmentation. For example, in *Periplaneta* two mechanisms at work during posterior patterning, Dl/N and pair-rule, may or may not be intertwined during posterior patterning. The primary pair-rule genes *Pa-eve* and *Pa-run* seem to act redundantly in the posterior, as phenotypes here were only observed upon double *Pa-eve/run*^{RNAi} compared to either gene alone. If Dl/N and pair-rule patterning mechanisms ran concurrently, this would just add another level of complex redundancy.

Along with the various levels of redundancy among the segmentation mechanisms, there is also considerable flexibility, even between closely related species, around the core conserved genes pertaining to their downstream targets. A study presented by Krol et al. (2011) analysed and compared the transcriptomes of three vertebrate species: chicken, zebrafish, and mouse. This study identified 40-100 cyclic genes within the three signalling pathways, but noted that different genes cycle in different species and between them only two *hes/her* orthologues were conserved, signifying the "evolutionary plasticity" of the vertebrate somitogenesis gene network (Krol et al., 2011). Flexibility is also observed in expression patterns of the arthropod pair-rule genes, which may be single or double segmental, a mix of these two, or even just a broad posterior domain; though still carrying out a conserved function in regulating segment polarity gene expression.

Genes involved in the delineation of segment/somite boundaries and the definition of anterior versus posterior (polarity) within each segment also vary between the different segmented bilaterians. In arthropods, this takes place through the segment polarity genes *engrailed* and *wingless*. In the annelids, some studies indicate a conserved mechanism of segment polarity and boundary formation via orthologues of *en* and *wg* or the NK homeobox genes, such as *tinman* and *ladybird* (Prud'homme et al., 2003; Saudemont et al., 2008). However, with the exception of the marine worm *Platynereis* (Dray et al., 2010; Prud'homme et al., 2003; Saudemont et al., 2008), these genes are typically expressed

after segment determination and may not play a role in actual segment formation (Patel et al., 1989; Seaver and Kaneshige, 2006; Seaver and Shankland, 2001), leaving open the question as to how these organisms carry out this function. Vertebrate orthologues of the arthropod segment polarity genes are generally not expressed in the developing somites (Ekker et al., 1992; Marti et al., 1995; Patel et al., 1989); however, *engrailed* is expressed in a segmental pattern in the invertebrate cephalochordate *Amphioxus* (Holland et al., 1997), suggesting some conservation between deuterostomes and protostomes in this process that have diverged as the lineages derived. Nevertheless, a mechanism for somite boundary and polarity does exist in the vertebrates, via segment polarity-like genes that are expressed in the unsegmented anterior region of the PSM before segment formation. These include *Mesp2* (Morimoto et al., 2005; Oginuma et al., 2008; Saga and Takahashi, 2008; Takahashi et al., 2000), *Delta-like1* (Hrabe de Angelis et al., 1997) and *lunatic fringe* (Evrard et al., 1998; Zhang and Gridley, 1998). These few examples highlight the redundant and flexible nature of the conserved segmentation mechanisms, allowing for species-specific divergence, but still coming to the same end – segment/somite formation.

Conclusions

The mechanisms of segmentation involve highly complex networks of positive and negative genetic interactions of which, in some regards, we have a fair amount of understanding (i.e. "*Drosophila* paradigm"; vertebrate "clock and wavefront"). Many elements of AP axis determination, posterior growth, and segmentation are shared between the segmented bilaterians and at the core of the segmentation network lies the conserved expressions of and genetic relationships between Wnt, *caudal*, Dl/N, and *hairy*. Within and between the segmented phyla, there are surprising levels of redundancy and

flexibility around the core members of this intricate gene network, yet the end goal of segmentation/somitogenesis is achieved.

While an ancient origin of the Wnt-Cad posterior organiser for axial growth is widely accepted (Martin and Kimelman, 2009; Petersen and Reddien, 2009), the Urbilaterian origins of a N-based segmentation mechanism have been hotly contested (Figure 1.1). Opponents argue that a single origin would consequently entail numerous losses and suggest, at the very least, two separate origins in the Deuterostomes and the Protostomes (Chipman, 2010). Other authors suggest convergent evolution of N-mediated segmentation that may have developed many times over the course of bilaterian evolution and even on a per species basis (Kainz et al., 2011). N-signalling is utilised in numerous developmental contexts, having arisen very early in animal evolution as a ubiquitous cell-cell communication mechanism (Gazave et al., 2009). The N-pathway may have been recruited into a new role in segmentation numerous times, having been in 'the right place at the right time'. The recruitment of an established network to a new location, or new genes into a pre-existing network, is not unheard of and may be quite common, such as body patterning genes employed in arthropod leg development (Beermann et al., 2004; Estella et al., 2003; Lemons et al., 2010; Panganiban et al., 1997; Schaeper et al., 2009), leg genes used to pattern beetle horns (Wasik and Moczek, 2011), or eye patterning genes used in butterfly wing development (Monteiro, 2012; Reed et al., 2011).

As for posterior segmentation mechanisms, independent recruitment of one signalling cascade for similar functions in posterior segmentation/somitogenesis in such widely divergent organisms as cockroaches and mice is possible, even plausible (i.e. convergent evolution). However, the probability of independently recruiting multiple pathways into a posterior segmentation gene network, arranged with similar regulatory connections and spatiotemporal expression patterns, may be too low to be coincident.

Along with this is an ingrained requirement (some evolutionary constraint) to maintain the 'robustness' of the system (Felix and Wagner, 2008) around the core players, allowing for flexibility in downstream targets leading to segment delineation and morphogenesis. That N-mediated segmentation exists in the three segmented phyla and is comparably linked to the Wnt-Cad posterior organiser, at least in the arthropods and vertebrates, suggests the likely involvement of these mechanisms in the last common ancestor shared between these organisms. Considering that it only takes a single point mutation to lose function of a gene or an entire pathway (Prud'homme et al., 2006; Shashikant et al., 1998; Theron et al., 2001), multiple loses over 570 million years of evolution are not improbable. As the different groups split into their separate lineages, these mechanisms were individually refined to suit their needs for body segmentation, such as pair-rule segmentation in the arthropods. In fact, the pair-rule genes are essential for proper patterning in arthropods, but show great variability in expression between species and even within the same embryo – i.e. anterior versus posterior patterning.

Further evidence of a shared segmentation mechanism with ancestral roots in the Urbilateria could be obtained by analysing the third signalling pathway required for vertebrate somitogenesis (FGF) in arthropods and annelids. A function for FGF-signalling in relation to arthropod segmentation, if found, would likely be intimately tied in with Wnt and N-signalling making it even less probable that these pathways could have become involved in the same ways at different times in the different segmented animals. Finally, investigations into an evolutionarily conserved patterning network could be expanded to include dorsoventral (DV) patterning, as this process is also established in the posterior during gastrulation. Conservation between vertebrate and insect DV patterning mechanisms have already been shown to exist, such as *BMP-4*/*dpp* and *chordin*/*short gastrulation* (De Robertis, 2008a; De Robertis and Sasai, 1996; Nunes da Fonseca et al.,

2010; van der Zee et al., 2006) and connections between DV (BMP/dpp) and AP (Wnt) patterning have already been established (De Robertis, 2008a; Fuentealba et al., 2007). Examining the connections between AP and DV patterning in arthropods will provide a more comprehensible picture of the ancestral model of this complex posterior patterning network. So far, it is apparent that the genes involved in segmentation/somitogenesis in the various segmented organisms function in a redundant and flexible manner around a conserved core network that was likely present in the Urbilaterian ancestor. As the lineages derived and evolved into their present state, downstream targets diverged, but the end goal of a segmented body remained. The flexible nature of this process allowed for the incredible diversity witnessed among the vertebrates, annelids, and especially the arthropods.

REFERENCES

Abzhanov, A. and Kaufman, T. C. (2000). Homologs of Drosophila appendage genes in the patterning of arthropod limbs. *Dev. Biol.* **227**, 673-689.

Adamska, M., Degnan, B. M., Green, K. and Zwafink, C. (2011). What sponges can tell us about the evolution of developmental processes. *Zoology (Jena)* **114**, 1-10.

Adamska, M., Degnan, S. M., Green, K. M., Adamski, M., Craigie, A., Larroux, C. and Degnan, B. M. (2007). Wnt and TGF-beta expression in the sponge Amphimedon queenslandica and the origin of metazoan embryonic patterning. *PLoS One* **2**, e1031.

Aguinaldo, A. M. A., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A. and Lake, J. A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489-493.

Akam, M. (1987a). The molecular basis for metameric pattern in the Drosophila embryo. *Development* **101**, 1-22.

Akam, M. (1987b). Molecules and morphology. *Nature* **327**, 184-185.

Akiyama-Oda, Y. and Oda, H. (2003). Early patterning of the spider embryo: a cluster of mesenchymal cells at the cumulus produces Dpp signals received by germ disc epithelial cells. *Development* **130**, 1735-47.

Alexandre, C. (2008). Cuticle preparation of Drosophila embryos and larvae. *Methods Mol Biol* **420**, 197-205.

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

Anderson, D. T. (1972). The development of hemimetabolous insects. In *Developmental systems: Insects*, vol. 1 (ed. S. J. Counce and C. H. Waddington), pp. 96-162. London: Academic Press Inc. LTD.

Andrioli, L. P., Oberstein, A. L., Corado, M. S., Yu, D. and Small, S. (2004). Groucho-dependent repression by sloppy-paired 1 differentially positions anterior pair-rule stripes in the Drosophila embryo. *Dev Biol* **276**, 541-51.

Ang, S. L., Jin, O., Rhinn, M., Daigle, N., Stevenson, L. and Rossant, J. (1996). A targeted mouse Otx2 mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain. *Development* **122**, 243-52.

Angelini, D. R. and Kaufman, T. C. (2005a). Functional analyses in the milkweed bug Oncopeltus fasciatus (Hemiptera) support a role for Wnt signaling in body segmentation but not appendage development. *Dev Biol* **283**, 409-23.

Angelini, D. R. and Kaufman, T. C. (2005b). Insect appendages and comparative ontogenetics. *Dev Biol* **286**, 57-77.

Angelini, D. R., Smith, F. W. and Jockusch, E. L. (2012). Extent With Modification: Leg Patterning in the Beetle Tribolium castaneum and the Evolution of Serial Homologs. *G3 (Bethesda)* **2**, 235-48.

Aranda, M., Marques-Souza, H., Bayer, T. and Tautz, D. (2008). The role of the segmentation gene hairy in Tribolium. *Dev Genes Evol* **218**, 465-77.

Aronson, B. D., Fisher, A. L., Blechman, K., Caudy, M. and Gergen, J. P. (1997). Groucho-dependent and -independent repression activities of Runt domain proteins. *Mol Cell Biol* **17**, 5581-7.

Artavanis-Tsakonas, S., Rand, M. D. and Lake, R. J. (1999). Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770-6.

Aulehla, A. and Herrmann, B. G. (2004). Segmentation in vertebrates: clock and gradient finally joined. *Genes Dev* **18**, 2060-7.

Aulehla, A. and Pourquie, O. (2010). Signaling gradients during paraxial mesoderm development. *Cold Spring Harb Perspect Biol* **2**, a000869.

Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B. and Herrmann, B. G. (2003). Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev Cell* **4**, 395-406.

Aulehla, A., Wiegraebe, W., Baubet, V., Wahl, M. B., Deng, C., Taketo, M., Lewandoski, M. and Pourquie, O. (2008). A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. *Nat Cell Biol* **10**, 186-93.

Avaron, F., Thaeron-Antono, C., Beck, C. W., Borday-Birraux, V., Geraudie, J., Casane, D. and Laurenti, P. (2003). Comparison of even-skipped related gene expression pattern in vertebrates shows an association between expression domain loss and modification of selective constraints on sequences. *Evol Dev* **5**, 145-56.

Balavoine, G. and Adoutte, A. (2003). The segmented urbilateria: a testable scenario. *Integr Comp Biol* **43**, 137-47.

Bastian, H. and Gruss, P. (1990). A murine even-skipped homologue, Evx 1, is expressed during early embryogenesis and neurogenesis in a biphasic manner. *Embo J* **9**, 1839-52.

Baumgartner, S. and Noll, M. (1990). Network of interactions among pair-rule genes regulating paired expression during primordial segmentation of Drosophila. *Mech Dev* **33**, 1-18.

Beck, C. W. and Slack, J. M. (1999). A developmental pathway controlling outgrowth of the Xenopus tail bud. *Development* **126**, 1611-20.

Beermann, A., Aranda, M. and Schroder, R. (2004). The Sp8 zinc-finger transcription factor is involved in allometric growth of the limbs in the beetle Tribolium castaneum. *Development* **131**, 733-42.

Beermann, A., Pruhs, R., Lutz, R. and Schroder, R. (2011). A context-dependent combination of Wnt receptors controls axis elongation and leg development in a short germ insect. *Development* **138**, 2793-805.

Beermann, A. and Schroder, R. (2008). Sites of Fgf signalling and perception during embryogenesis of the beetle Tribolium castaneum. *Dev Genes Evol* **218**, 153-67.

Benedyk, M. J., Mullen, J. R. and DiNardo, S. (1994). odd-paired: a zinc finger pair-rule protein required for the timely activation of engrailed and wingless in Drosophila embryos. *Genes Dev* **8**, 105-17.

Bertuccioli, C., Fasano, L., Jun, S., Wang, S., Sheng, G. and Desplan, C. (1996). In vivo requirement for the paired domain and homeodomain of the paired segmentation gene product. *Development* **122**, 2673-85.

Bessho, Y., Hirata, H., Masamizu, Y. and Kageyama, R. (2003). Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. *Genes Dev* **17**, 1451-6.

Bessho, Y., Sakata, R., Komatsu, S., Shiota, K., Yamada, S. and Kageyama, R. (2001). Dynamic expression and essential functions of Hes7 in somite segmentation. *Genes Dev* **15**, 2642-7.

Binner, P. and Sander, K. (1997). Pair-rule patterning in the honeybee Apis mellifera: Expression of even-skipped combines traites known from beetles and fruitfly. *Dev Genes Evol* **206**, 447-454.

Bishop, S. A., Klein, T., Arias, A. M. and Couso, J. P. (1999). Composite signalling from Serrate and Delta establishes leg segments in Drosophila through Notch. *Development* **126**, 2993-3003.

Blagburn, J. M. (2007). Co-factors and co-repressors of Engrailed: expression in the central nervous system and cerci of the cockroach, Periplaneta americana. *Cell and Tissue Research* **327**, 177-187.

Bolognesi, R., Farzana, L., Fischer, T. D. and Brown, S. J. (2008). Multiple Wnt genes are required for segmentation in the short-germ embryo of Tribolium castaneum. *Curr Biol* **18**, 1624-9.

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. and Noll, M. (1986). Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. *Cell* **47**, 1033-1040.

Brown, S. J., Parrish, J. K., Beeman, R. W. and Denell, R. E. (1997). Molecular characterization and embryonic expression of the even-skipped ortholog of Tribolium castaneum. *Mech Dev* **61**, 165-73.

Bucher, G. and Klingler, M. (2004). Divergent segmentation mechanism in the short germ insect Tribolium revealed by giant expression and function. *Development* **131**, 1729-40.

Budd, G. E. (2001). Why are arthropods segmented? *Evolution & Development* **3**, 332-342.

Cadigan, K. M., Grossniklaus, U. and Gehring, W. J. (1994). Localized expression of sloppy paired protein maintains the polarity of Drosophila parasegments. *Genes Dev* **8**, 899-913.

Cerny, A. C., Bucher, G., Schroder, R. and Klingler, M. (2005). Breakdown of abdominal patterning in the Tribolium Kruppel mutant jaws. *Development* **132**, 5353-63.

Cerny, A. C., Grossmann, D., Bucher, G. and Klingler, M. (2008). The Tribolium ortholog of knirps and knirps-related is crucial for head segmentation but plays a minor role during abdominal patterning. *Dev Biol* **321**, 284-94.

Chanut-Delalande, H., Fernandes, I., Roch, F., Payre, F. and Plaza, S. (2006). Shavenbaby couples patterning to epidermal cell shape control. *Plos Biology* **4**, 1549-1561.

Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J. and Beck, F. (2004). Cdx2 is essential for axial elongation in mouse development. *Proc Natl Acad Sci U S A* **101**, 7641-5.

Chen, B., Dodge, M. E., Tang, W., Lu, J., Ma, Z., Fan, C. W., Wei, S., Hao, W., Kilgore, J., Williams, N. S. et al. (2009). Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nat Chem Biol* **5**, 100-7.

Chen, H., Xu, Z., Mei, C., Yu, D. and Small, S. (2012). A system of repressor gradients spatially organizes the boundaries of Bicoid-dependent target genes. *Cell* **149**, 618-29.

Chesebro, J. (2008). The role of Scr in two hemimetabolous insect species, Oncopeltus fasciatus and Periplaneta americana. In *Biological Sciences*, vol. MS (ed., pp. 68. Detroit: Wayne State University.

Chipman, A. D. (2010). Parallel evolution of segmentation by co-option of ancestral gene regulatory networks. *Bioessays* **32**, 60-70.

Chipman, A. D. and Akam, M. (2008). The segmentation cascade in the centipede Strigamia maritima: Involvement of the Notch pathway and pair-rule gene homologues. *Developmental Biology* **319**, 160-169.

Chipman, A. D., Arthur, W. and Akam, M. (2004). A double segment periodicity underlies segment generation in centipede development. *Current Biology* **14**, 1250-1255.

Cho, S. J., Valles, Y., Giani, V. C., Jr., Seaver, E. C. and Weisblat, D. A. (2010). Evolutionary dynamics of the wnt gene family: a lophotrochozoan perspective. *Mol Biol Evol* **27**, 1645-58.

Choe, C. P. and Brown, S. J. (2007). Evolutionary flexibility of pair-rule patterning revealed by functional analysis of secondary pair-rule genes, paired and sloppy-paired in the short-germ insect, Tribolium castaneum. *Dev Biol* **302**, 281-94.

Choe, C. P. and Brown, S. J. (2009). Genetic regulation of engrailed and wingless in Tribolium segmentation and the evolution of pair-rule segmentation. *Dev Biol* **325**, 482-91.

Choe, C. P., Miller, S. C. and Brown, S. J. (2006). A pair-rule gene circuit defines segments sequentially in the short-germ insect Tribolium castaneum. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6560-6564.

Cinquin, O. (2007). Understanding the somitogenesis clock: what's missing? *Mech Dev* **124**, 501-17.

Cockayne, E. A. (1929). Spiral and other anomalous forms of segmentation. *Trans Ent Soc Lond* **77**, 177-184.

Cohen, S. M. and Jurgens, G. (1990). Mediation of Drosophila head development by gap-like segmentation genes. *Nature* **346**, 482-5.

Cooke, J. and Zeeman, E. C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *J Theor Biol* **58**, 455-76.

Copf, T., Rabet, N., Celniker, S. E. and Averof, M. (2003). Posterior patterning genes and the identification of a unique body region in the brine shrimp Artemia franciscana. *Development* **130**, 5915-27.

Copf, T., Schroder, R. and Averof, M. (2004). Ancestral role of caudal genes in axis elongation and segmentation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17711-17715.

Couso, J. P. (2009). Segmentation, metamerism and the Cambrian explosion. *Int J Dev Biol* **53**, 1305-16.

Couso, J. P., Bate, M. and Martinez Arias, A. (1993). A wingless-dependent polar coordinate system in Drosophila imaginal discs. *Science* **259**, 484-489.

Couso, J. P. and Bishop, S. A. (1998). Proximo-distal development in the legs of Drosophila. *International Journal of Developmental Biology* **42**, 345-352.

Cruz, C., Maegawa, S., Weinberg, E. S., Wilson, S. W., Dawid, I. B. and Kudoh, T. (2010). Induction and patterning of trunk and tail neural ectoderm by the homeobox gene eve1 in zebrafish embryos. *Proc Natl Acad Sci U S A* **107**, 3564-9.

Curcic, B. P. M., Krunic, M. D. and Brajkovic, M. M. (1983). Tergal and sternal anomalies in Neobisium Chamberlin (Neobisiidae, Pseudoscorpiones, Arachnida). *J Arachnol* **11**, 243-250.

Damen, W. G., Saridaki, T. and Averof, M. (2002). Diverse adaptations of an ancestral gill: a common evolutionary origin for wings, breathing organs, and spinnerets. *Curr Biol* **12**, 1711-6.

Damen, W. G., Weller, M. and Tautz, D. (2000). Expression patterns of hairy, even-skipped, and runt in the spider Cupiennius salei imply that these genes were segmentation genes in a basal arthropod. *Proc Natl Acad Sci U S A* **97**, 4515-4519.

Damen, W. G. M. (2007). Evolutionary conservation and divergence of the segmentation process in arthropods. *Developmental Dynamics* **236**, 1379-1391.

Damen, W. G. M., Janssen, R. and Prpic, N. M. (2005). Pair rule gene orthologs in spider segmentation. *Evolution & Development* **7**, 618-628.

Davis, G. K., D'Alessio, J. A. and Patel, N. H. (2005). Pax3/7 genes reveal conservation and divergence in the arthropod segmentation hierarchy. *Dev Biol* **285**, 169-84.

Davis, G. K., Jaramillo, C. A. and Patel, N. H. (2001). Pax group III genes and the evolution of insect pair-rule patterning. *Development* **128**, 3445-3458.

Davis, G. K. and Patel, N. H. (1999). The origin and evolution of segmentation. *Trends Cell Biol* **9**, M68-72.

Davis, G. K. and Patel, N. H. (2002). Short, long, and beyond: Molecular and embryological approaches to insect segmentation. *Annual Review of Entomology* **47**, 669-699.

Dawes, R., Dawson, I., Falciani, F., Tear, G. and Akam, M. (1994). Dax, a locust Hox gene related to fushi-tarazu but showing no pair-rule expression. *Development* **120**, 1561-72.

de Celis, J. F., Tyler, D. M., de Celis, J. and Bray, S. J. (1998). Notch signalling mediates segmentation of the Drosophila leg. *Development* **125**, 4617-4626.

De Robertis, E. M. (1997). The ancestry of segmentation. *Nature* **387**, 25-26.

De Robertis, E. M. (2008a). Evo-Devo: Variations on Ancestral Themes. *Cell* **132**, 185-195.

De Robertis, E. M. (2008b). The molecular ancestry of segmentation mechanisms. *Proc Natl Acad Sci U S A* **105**, 16411-2.

De Robertis, E. M. and Sasai, Y. (1996). A common plan for dorsoventral patterning in Bilateria. *Nature* **380**, 37-40.

de Rosa, R., Prud'homme, B. and Balavoine, G. (2005). Caudal and even-skipped in the annelid Platynereis dumerilii and the ancestry of posterior growth. *Evol Dev* **7**, 574-87.

Dearden, P. and Akam, M. (2000). A role for Fringe in segment morphogenesis but not segment formation in the grasshopper, Schistocerca gregaria. *Development Genes and Evolution* **210**, 329-336.

Dearden, P. K. and Akam, M. (2001). Early embryo patterning in the grasshopper, Schistocerca gregaria: wingless, decapentaplegic and caudal expression. *Development* **128**, 3435-44.

Dearden, P. K., Donly, C. and Grbic, M. (2002). Expression of pair-rule gene homologues in a chelicerate: early patterning of the two-spotted spider mite Tetranychus urticae. *Development* **129**, 5461-5472.

Delon, I., Chanut-Delalande, H. and Payre, F. (2003). The Ovo/Shavenbaby transcription factor specifies actin remodelling during epidermal differentiation in Drosophila. *Mechanisms of Development* **120**, 747-758.

Dequeant, M. L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A. and Pourquie, O. (2006). A complex oscillating network of signaling genes underlies the mouse segmentation clock. *Science* **314**, 1595-1598.

Dequeant, M. L. and Pourquie, O. (2008). Segmental patterning of the vertebrate embryonic axis. *Nat Rev Genet* **9**, 370-82.

DiNardo, S. and O'Farrell, P. H. (1987). Establishment and refinement of segmental pattern in the Drosophila embryo: spatial control of engrailed expression by pair-rule genes. *Genes Dev* **1**, 1212-1225.

Dove, H. and Stollewerk, A. (2003). Comparative analysis of neurogenesis in the myriapod Glomeris marginata (Diplopoda) suggests more similarities to chelicerates than to insects. *Development* **130**, 2161-71.

Dovey, H. F., John, V., Anderson, J. P., Chen, L. Z., de Saint Andrieu, P., Fang, L. Y., Freedman, S. B., Folmer, B., Goldbach, E., Holsztynska, E. J. et al. (2001). Functional gamma-secretase inhibitors reduce beta-amyloid peptide levels in brain. *J Neurochem* **76**, 173-81.

Dray, N., Tessmar-Raible, K., Le Gouar, M., Vibert, L., Christodoulou, F., Schipany, K., Guillou, A., Zantke, J., Snyman, H., Behague, J. et al. (2010). Hedgehog signaling regulates segment formation in the annelid Platynereis. *Science* **329**, 339-42.

Dubrulle, J., McGrew, M. J. and Pourquie, O. (2001). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* **106**, 219-32.

Dubrulle, J. and Pourquie, O. (2004a). Coupling segmentation to axis formation. *Development* **131**, 5783-93.

Dubrulle, J. and Pourquie, O. (2004b). fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. *Nature* **427**, 419-22.

Duncan, E. J., Wilson, M. J., Smith, J. M. and Dearden, P. K. (2008). Evolutionary origin and genomic organisation of runt-domain containing genes in arthropods. *BMC Genomics* **9**, 558.

Dunn, C. W., Hejnol, A., Matus, D. Q., Pang, K., Browne, W. E., Smith, S. A., Seaver, E., Rouse, G. W., Obst, M., Edgecombe, G. D. et al. (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**, 745-U5.

Dunty, W. C., Jr., Biris, K. K., Chalamalasetty, R. B., Taketo, M. M., Lewandoski, M. and Yamaguchi, T. P. (2008). Wnt3a/beta-catenin signaling controls posterior body development by coordinating mesoderm formation and segmentation. *Development* **135**, 85-94.

Dush, M. K. and Martin, G. R. (1992). Analysis of mouse Evx genes: Evx-1 displays graded expression in the primitive streak. *Dev Biol* **151**, 273-87.

Eckert, C., Aranda, M., Wolff, C. and Tautz, D. (2004). Separable stripe enhancer elements for the pair-rule gene hairy in the beetle Tribolium. *EMBO Rep* **5**, 638-42.

Ekker, M., Wegner, J., Akimenko, M. A. and Westerfield, M. (1992). Coordinate embryonic expression of three zebrafish engrailed genes. *Development* **116**, 1001-10.

El-Sherif, E., Averof, M. and Brown, S. J. (2012). A segmentation clock operating in blastoderm and germband stages of Tribolium development. *Development*.

Erwin, D. H. and Davidson, E. H. (2002). The last common bilaterian ancestor. *Development* **129**, 3021-3032.

Estella, C., Rieckhof, G., Calleja, M. and Morata, G. (2003). The role of buttonhead and Sp1 in the development of the ventral imaginal discs of Drosophila. *Development* **130**, 5929-5941.

Evrard, Y. A., Lun, Y., Aulehla, A., Gan, L. and Johnson, R. L. (1998). lunatic fringe is an essential mediator of somite segmentation and patterning. *Nature* **394**, 377-81.

Farzana, L. and Brown, S. J. (2008). Hedgehog signaling pathway function conserved in Tribolium segmentation. *Dev Genes Evol* **218**, 181-92.

Felix, M. A. and Wagner, A. (2008). Robustness and evolution: concepts, insights and challenges from a developmental model system. *Heredity (Edinb)* **100**, 132-40.

Forsberg, H., Crozet, F. and Brown, N. A. (1998). Waves of mouse Lunatic fringe expression, in four-hour cycles at two-hour intervals, precede somite boundary formation. *Curr Biol* **8**, 1027-30.

French, V. (1983). Development and evolution of the insect segment. In *Dev.and Evol. British Society for Developmental Biology Symposium*, vol. 6 (ed. B. Goodwin), pp. 161-193.

Fuentealba, L. C., Eivers, E., Ikeda, A., Hurtado, C., Kuroda, H., Pera, E. M. and De Robertis, E. M. (2007). Integrating patterning signals: Wnt/GSK3 regulates the duration of the BMP/Smad1 signal. *Cell* **131**, 980-93.

Fujioka, M. and Jaynes, J. B. (2012). Regulation of a duplicated locus: Drosophila sloppy paired is replete with functionally overlapping enhancers. *Dev Biol* **362**, 309-19.

Galindo, M. I., Pueyo, J. I., Fouix, S., Bishop, S. A. and Couso, J. P. (2007). Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *Plos Biology* **5**, 1052-1062.

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

Gergen, J. P. and Butler, B. A. (1988). Isolation of the Drosophila segmentation gene runt and analysis of its expression during embryogenesis. *Genes Dev* **2**, 1179-93.

Gergen, J. P. and Wieschaus, E. F. (1985). The localized requirements for a gene affecting segmentation in Drosophila: analysis of larvae mosaic for runt. *Developmental Biology* **109**, 321-335.

Gibb, S., Maroto, M. and Dale, J. K. (2010). The segmentation clock mechanism moves up a notch. *Trends Cell Biol* **20**, 593-600.

Gibb, S., Zagorska, A., Melton, K., Tenin, G., Vacca, I., Trainor, P., Maroto, M. and Dale, J. K. (2009). Interfering with Wnt signalling alters the periodicity of the segmentation clock. *Dev Biol* **330**, 21-31.

Giudicelli, F. and Lewis, J. (2004). The vertebrate segmentation clock. *Current Opinion in Genetics & Development* **14**, 407-414.

Goldbeter, A. and Pourquie, O. (2008). Modeling the segmentation clock as a network of coupled oscillations in the Notch, Wnt and FGF signaling pathways. *J Theor Biol* **252**, 574-85.

Gomez, C., Ozbudak, E. M., Wunderlich, J., Baumann, D., Lewis, J. and Pourquie, O. (2008). Control of segment number in vertebrate embryos. *Nature* **454**, 335-339.

Grainger, S., Lam, J., Savory, J. G., Mears, A. J., Rijli, F. M. and Lohnes, D. (2012). Cdx regulates Dll1 in multiple lineages. *Dev Biol* **361**, 1-11.

Grbic, M., Nagy, L. M., Carroll, S. B. and Strand, M. (1996). Polyembryonic development: insect pattern formation in a cellularized environment. *Development* **122**, 795-804.

Greenberg, L. and Hatini, V. (2009). Essential roles for lines in mediating leg and antennal proximodistal patterning and generating a stable Notch signaling interface at segment borders. *Developmental Biology* **330**, 93-104.

Grossmann, D., Scholten, J. and Prpic, N. M. (2009). Separable functions of wingless in distal and ventral patterning of the Tribolium leg. *Dev Genes Evol* **219**, 469-79.

Grossniklaus, U., Cadigan, K. M. and Gehring, W. J. (1994). Three maternal coordinate systems cooperate in the patterning of the Drosophila head. *Development* **120**, 3155-71.

Grossniklaus, U., Pearson, R. K. and Gehring, W. J. (1992). The Drosophila sloppy paired locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. *Genes Dev* **6**, 1030-51.

Gutjahr, T., Frei, E. and Noll, M. (1993). Complex regulation of early paired expression: initial activation by gap genes and pattern modulation by pair-rule genes. *Development* **117**, 609-23.

Hao, I., Green, R. B., Dunaevsky, O., Lengyel, J. A. and Rauskolb, C. (2003). The odd-skipped family of zinc finger genes promotes Drosophila leg segmentation. *Dev Biol* **263**, 282-95.

Herrgen, L., Ares, S., Morelli, L. G., Schroter, C., Julicher, F. and Oates, A. C. (2010). Intercellular coupling regulates the period of the segmentation clock. *Curr Biol* **20**, 1244-53.

Holland, L. Z. (2002). Heads or tails? Amphioxus and the evolution of anterior-posterior patterning in deuterostomes. *Dev Biol* **241**, 209-28.

Holland, L. Z., Kene, M., Williams, N. A. and Holland, N. D. (1997). Sequence and embryonic expression of the amphioxus engrailed gene (AmphiEn): the metameric pattern of transcription resembles that of its segment-polarity homolog in Drosophila. *Development* **124**, 1723-32.

Holmgren, R. (1984). Cloning sequences from the hairy gene of Drosophila. *Embo J* **3**, 569-73.

Horikawa, K., Ishimatsu, K., Yoshimoto, E., Kondo, S. and Takeda, H. (2006). Noise-resistant and synchronized oscillation of the segmentation clock. *Nature* **441**, 719-23.

Howard, K. and Ingham, P. (1986). Regulatory interactions between the segmentation genes fushi tarazu, hairy, and engrailed in the Drosophila blastoderm. *Cell* **44**, 949-57.

Hrabe de Angelis, M., McIntyre, J., 2nd and Gossler, A. (1997). Maintenance of somite borders in mice requires the Delta homologue DII1. *Nature* **386**, 717-21.

Hrycaj, S., Mihajlovic, M., Mahfooz, N., Couso, J. P. and Popadic, A. (2008). RNAi analysis of nubbin embryonic functions in a hemimetabolous insect, Oncopeltus fasciatus. *Evol Dev* **10**, 705-16.

Huang, P. and Stern, M. J. (2005). FGF signaling in flies and worms: more and more relevant to vertebrate biology. *Cytokine Growth Factor Rev* **16**, 151-8.

Hughes, C. L. and Kaufman, T. C. (2002a). Exploring myriapod segmentation: The expression patterns of even-skipped, engrailed, and wingless in a centipede. *Developmental Biology* **247**, 47-61.

Hughes, C. L. and Kaufman, T. C. (2002b). Exploring the myriapod body plan: expression patterns of the ten Hox genes in a centipede. *Development* **129**, 1225-38.

Hughes, C. L. and Kaufman, T. C. (2002c). Hox genes and the evolution of the arthropod body plan. *Evol Dev* **4**, 459-99.

Hui, J. H., Raible, F., Korchagina, N., Dray, N., Samain, S., Magdelenat, G., Jubin, C., Segurens, B., Balavoine, G., Arendt, D. et al. (2009). Features of the ancestral bilaterian inferred from Platynereis dumerilii ParaHox genes. *BMC Biol* **7**, 43.

Ikeya, M. and Takada, S. (2001). Wnt-3a is required for somite specification along the anteroposterior axis of the mouse embryo and for regulation of cdx-1 expression. *Mech Dev* **103**, 27-33.

Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in Drosophila [published erratum appears in Nature 1988 Oct 20;335(6192):744]. *Nature* **335**, 25-34.

Inoue, K., Ozaki, S., Shiga, T., Ito, K., Masuda, T., Okado, N., Iseda, T., Kawaguchi, S., Ogawa, M., Bae, S. C. et al. (2002a). Runx3 controls the axonal projection of proprioceptive dorsal root ganglion neurons. *Nat Neurosci* **5**, 946-54.

Inoue, Y., Mito, T., Miyawaki, K., Matsushima, K., Shinmyo, Y., Heanue, T. A., Mardon, G., Ohuchi, H. and Noji, S. (2002b). Correlation of expression patterns of *homothorax*, *dachshund*, and *Distal-less* with the proximodistal segmentation of the cricket leg bud. *Mech. Dev.*

Jacobs, D. K., Hughes, N. C., Fitz-Gibbon, S. T. and Winchell, C. J. (2005). Terminal addition, the Cambrian radiation and the Phanerozoic evolution of bilaterian form. *Evolution & Development* **7**, 498-514.

Janssen, R., Budd, G. E., Prpic, N. M. and Damen, W. G. (2011). Expression of myriapod pair rule gene orthologs. *Evodevo* **2**, 5.

Janssen, R., Damen, W. G. M. and Budd, G. E. (2012). Expression of pair-rule gene orthologs in the blastoderm of a myriapod: evidence of pair rule-like mechanisms? *BMC Dev Biol* **12**, Epub ahead of print.

Janssen, R., Le Gouar, M., Pechmann, M., Poulin, F., Bolognesi, R., Schwager, E. E., Hopfen, C., Colbourne, J. K., Budd, G. E., Brown, S. J. et al. (2010). Conservation, loss, and redeployment of Wnt ligands in protostomes: implications for understanding the evolution of segment formation. *BMC Evol Biol* **10**, 374.

Jaynes, J. B. and Fujioka, M. (2004). Drawing lines in the sand: even skipped et al. and parasegment boundaries. *Developmental Biology* **269**, 609-622.

Jiang, Y. J., Aerne, B. L., Smithers, L., Haddon, C., Ish-Horowicz, D. and Lewis, J. (2000). Notch signalling and the synchronization of the somite segmentation clock. *Nature* **408**, 475-479.

Jiménez, F. and Campos-Ortega, J. A. (1982). Maternal effects of zygotic mutants affecting early neurogenesis in Drosophila. *Development Genes and Evolution* **191**, 191-201.

Jockusch, E. L., Nulsen, C., Newfeld, S. J. and Nagy, L. M. (2000). Leg development in flies versus grasshoppers: differences in dpp expression do not lead to differences in the expression of downstream components of the leg patterning pathway. *Development* **127**, 1617 - 1626.

Joly, J. S., Joly, C., Schulte-Merker, S., Boulekbache, H. and Condamine, H. (1993). The ventral and posterior expression of the zebrafish homeobox gene eve1 is perturbed in dorsalized and mutant embryos. *Development* **119**, 1261-75.

Jostes, B., Walther, C. and Gruss, P. (1990). The murine paired box gene, Pax7, is expressed specifically during the development of the nervous and muscular system. *Mech Dev* **33**, 27-37.

Kadner, D. and Stollewerk, A. (2004). Neurogenesis in the chilopod Lithobius forficatus suggests more similarities to chelicerates than to insects. *Dev Genes Evol* **214**, 367-79.

Kageyama, R., Niwa, Y., Isomura, A., Gonzalez, A. and Harima, Y. (2012). Oscillatory gene expression and somitogenesis. *WIREs Dev Biol* **1**, 629-641.

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

Kainz, F., Ewen-Campen, B., Akam, M. and Extavour, C. G. (2011). Notch/Delta signalling is not required for segment generation in the basally branching insect Gryllus bimaculatus. *Development* **138**, 5015-26.

Kawamura, A., Koshida, S., Hijikata, H., Sakaguchi, T., Kondoh, H. and Takada, S. (2005). Zebrafish hairy/enhancer of split protein links FGF signaling to cyclic gene expression in the periodic segmentation of somites. *Genes Dev* **19**, 1156-61.

Keller, R. G., Desplan, C. and Rosenberg, M. I. (2010). Identification and characterization of Nasonia Pax genes. *Insect Mol Biol* **19 Suppl 1**, 109-20.

Kilchherr, F., Baumgartner, S., Bopp, D., Frei, E. and Noll, M. (1986). Isolation of the paired gene of Drosophila and its spatial expression during early embryogenesis. *Nature* **321**, 493-499.

Kimelman, D. and Martin, B. L. (2012). Anterior-Posterior patterning in early development: three strategies. *WIREs Dev Biol* **1**, 253-266.

Kimmel, C. B. (1996). Was Urbilateria segmented? *Trends Genet* **12**, 329-31.

Kobayashi, M., Goldstein, R. E., Fujioka, M., Paroush, Z. and Jaynes, J. B. (2001). Groucho augments the repression of multiple Even skipped target genes in establishing parasegment boundaries. *Development* **128**, 1805-15.

Kojima, T. (2004). The mechanism of Drosophila leg development along the proximodistal axis. *Dev., Growth Diffn* **46**, 115--129.

Kojima, T., Sato, M. and Saigo, K. (2000). Formation and specification of distal leg segments in Drosophila by dual Bar homeobox genes, BarH1 and BarH2. *Development* **127**, 769-778.

Kondo, T., Hashimoto, Y., Kato, K., Inagaki, S., Hayashi, S. and Kageyama, Y. (2007). Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. *Nature Cell Biology* **9**, 660-U87.

Kondo, T., Plaza, S., Zanet, J., Benrabah, E., Valenti, P., Hashimoto, Y., Kobayashi, S., Payre, F. and Kageyama, Y. (2010). Small peptides switch the transcriptional activity of Shavenbaby during Drosophila embryogenesis. *Science* **329**, 336-9.

Kraft, R. and Jaeckle, H. (1994). Drosophila mode of metamerization in the embryogenesis of the lepidopteran insect Manduca sexta. *Proc. Natl. Acad. Sci. USA* **91**, 6634-6638.

Krol, A. J., Roellig, D., Dequeant, M. L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A. C. and Pourquie, O. (2011). Evolutionary plasticity of segmentation clock networks. *Development* **138**, 2783-92.

Lanfear, R. (2007). The evolution of animal body plans. In *School of Life Sciences*, vol. PhD (ed., pp. 228. Brighton: University of Sussex.

Lawrence, P. A. (1992). The making of a fly. Oxford: Blackwell Scientic Publications.

Lemons, D., Fritzenwanker, J. H., Gerhart, J., Lowe, C. J. and McGinnis, W. (2010). Co-option of an anteroposterior head axis patterning system for proximodistal patterning of appendages in early bilaterian evolution. *Dev Biol* **344**, 358-62.

Lenoir-Rousseaux, J. J. and Lender, T. (1970). Table de developpement embryonnaire de *Periplaneta americana* (L.) insecte dictyptere. *Bulletin de la Societe Zoologique de France.* **95**, 737-751.

Lesniewska, M., Bonato, L., Minelli, A. and Fusco, G. (2009). Trunk anomalies in the centipede Stigmatogaster subterranea provide insight into late-embryonic segmentation. *Arthropod Struct Dev* **38**, 417-26.

Lewis, D. L., DeCamillis, M. and Bennett, R. L. (2000). Distinct roles of the homeotic genes Ubx and abd-A in beetle embryonic abdominal appendage development. *Proc Natl Acad Sci U S A* **97**, 4504-9.

Lewis, J. (2003). Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. *Curr Biol* **13**, 1398-408.

Li, H. and Popadić, A. (2004). Analysis of nubbin expression patterns in insects. *Evolution & Development* **6**, 310-324.

Liu, P. Z. and Kaufman, T. C. (2005a). even-skipped is not a pair-rule gene but has segmental and gap-like functions in Oncopeltus fasciatus, an intermediate germband insect. *Development* **132**, 2081-2092.

Liu, P. Z. and Kaufman, T. C. (2005b). Short and long germ segmentation: unanswered questions in the evolution of a developmental mode. *Evol Dev* **7**, 629-46.

Liu, P. Z. and Patel, N. H. (2010). giant is a bona fide gap gene in the intermediate germband insect, Oncopeltus fasciatus. *Development* **137**, 835-44.

Liu, W., Yang, F., Jia, S., Miao, X. and Huang, Y. (2008). Cloning and characterization of Bmrunt from the silkworm Bombyx mori during embryonic development. *Arch Insect Biochem Physiol* **69**, 47-59.

Lohnes, D. (2003). The Cdx1 homeodomain protein: an integrator of posterior signaling in the mouse. *Bioessays* **25**, 971-80.

Lynch, J. A., Brent, A. E., Leaf, D. S., Pultz, M. A. and Desplan, C. (2006). Localized maternal orthodenticle patterns anterior and posterior in the long germ wasp Nasonia. *Nature* **439**, 728-32.

Ma, Y., Niemitz, E. L., Nambu, P. A., Shan, X., Sackerson, C., Fujioka, M., Goto, T. and Nambu, J. R. (1998). Gene regulatory functions of Drosophila fish-hook, a high mobility group domain Sox protein. *Mech Dev* **73**, 169-82.

Macdonald, P. M., Ingham, P. and Struhl, G. (1986). Isolation, structure, and expression of even-skipped: a second pair-rule gene of Drosophila containing a homeo box. *Cell* **47**, 721-34.

Macdonald, P. M. and Struhl, G. (1986). A molecular gradient in early Drosophila embryos and its role in specifying the body pattern. *Nature* **324**, 537-45.

Mara, A., Schroeder, J., Chalouni, C. and Holley, S. A. (2007). Priming, initiation and synchronization of the segmentation clock by deltaD and deltaC. *Nat Cell Biol* **9**, 523-30.

Marie, B. and Bacon, J. P. (2000). Two engrailed-related genes in the cockroach: cloning, phylogenetic analysis, expression and isolation of splice variants. *Dev. Genes Evol.* **210**, 436-448.

Marie, B., Bacon, J. P. and Blagburn, J. M. (2000). Double-stranded RNA interference shows that Engrailed controls the synaptic specificity of identified sensory neurons. *Curr. Biol.* **10**, 289-292.

Marti, E., Takada, R., Bumcrot, D. A., Sasaki, H. and McMahon, A. P. (1995). Distribution of Sonic hedgehog peptides in the developing chick and mouse embryo. *Development* **121**, 2537-47.

Martin, B. L. and Kimelman, D. (2009). Wnt signaling and the evolution of embryonic posterior development. *Curr Biol* **19**, R215-9.

Martinez-Arias, A. (1993). Development and patterning of the larval epidermis of Drosophila. In *The development of Drosophila melanogaster*, (ed. C. M. Bate and A. Martinez Arias), pp. 517-608: Cold Spring Harbour Laboratory Press.

Martinez Arias, A. and Lawrence, P. A. (1985). Parasegments and compartments in the Drosophila embryo. *Nature* **313**, 639-642.

McGregor, A. P., Pechmann, M., Schwager, E. E. and Damen, W. G. (2009). An ancestral regulatory network for posterior development in arthropods. *Commun Integr Biol* **2**, 174-6.

McGregor, A. P., Pechmann, M., Schwager, E. E., Feitosa, N. M., Kruck, S., Aranda, M. and Damen, W. G. (2008). Wnt8 is required for growth-zone establishment and development of opisthosomal segments in a spider. *Curr Biol* **18**, 1619-23.

Mercier, P., Simeone, A., Cotelli, F. and Boncinelli, E. (1995). Expression pattern of two otx genes suggests a role in specifying anterior body structures in zebrafish. *Int J Dev Biol* **39**, 559-73.

Minelli, A. (2001). A three-phase model of arthropod segmentation. *Dev Genes Evol* **211**, 509-21.

Minelli, A. and Fusco, G. (2004). Evo-devo perspectives on segmentation: model organisms, and beyond. *Trends in Ecology & Evolution* **19**, 423-429.

Mitic, B. M., Makarov, S. E., Ilic, B. S., Stojanovic, D. Z. and Curcic, B. P. M. (2011). Cases of trunk segmenatal anomalies in the geophilomorph centipedes Clinopodes flavidus C. L. Koch and Clinopodes trebevicensis (Verhoff) (Chilopoda: Geophilomorpha). *Arch Biol Sci* **63**, 841-845.

Mito, T., Kobayashi, C., Sarashina, I., Zhang, H. J., Shinahara, W., Miyawaki, K., Shinmyo, Y., Ohuchi, H. and Noji, S. (2007). even-skipped has gap-like, pair-rule-like, and segmental functions in the cricket Gryllus bimaculatus, a basal, intermediate germ insect (Orthoptera). *Developmental Biology* **303**, 202-213.

Mito, T., Okamoto, H., Shinahara, W., Shinmyo, Y., Miyawaki, K., Ohuchi, H. and Noji, S. (2006). Kruppel acts as a gap gene regulating expression of hunchback and even-skipped in the intermediate germ cricket Gryllus bimaculatus. *Dev Biol* **294**, 471-81.

Mito, T., Sarashina, I., Zhang, H., Iwahashi, A., Okamoto, H., Miyawaki, K., Shinmyo, Y., Ohuchi, H. and Noji, S. (2005). Non-canonical functions of hunchback in segment patterning of the intermediate germ cricket Gryllus bimaculatus. *Development* **132**, 2069-79.

Mito, T., Shinmyo, Y., Kurita, K., Nakamura, T., Ohuchi, H. and Noji, S. (2011). Ancestral functions of Delta/Notch signaling in the formation of body and leg segments in the cricket Gryllus bimaculatus. *Development* **138**, 3823-33.

Miyawaki, K., Mito, T., Sarashina, I., Zhang, H. J., Shinmyo, Y., Ohuchi, H. and Noji, S. (2004). Involvement of Wingless/Armadillo signaling in the posterior sequential segmentation in the cricket, Gryllus bimaculatus (Orthoptera), as revealed by RNAi analysis. *Mechanisms of Development* **121**, 119-130.

Mohr, O. L. (1924). A genetic and cytological analysis of a section deficiency involving four units of the X-chromosome in Drosophila melanogaster. *Molecular and General Genetics MGG* **32**, 108-232.

Monteiro, A. (2012). Gene regulatory networks reused to build novel traits: co-option of an eye-related gene regulatory network in eye-like organs and red wing patches on insect wings is suggested by optix expression. *Bioessays* **34**, 181-6.

Mooi, R. and David, B. (2008). Radial Symmetry, the Anterior/Posterior Axis, and Echinoderm Hox Genes. *Annual Review of Ecology, Evolution, and Systematics* **39**, 43-62.

Moran-Rivard, L., Kagawa, T., Saueressig, H., Gross, M. K., Burrill, J. and Goulding, M. (2001). Evx1 is a postmitotic determinant of v0 interneuron identity in the spinal cord. *Neuron* **29**, 385-99.

Moreno, E. and Morata, G. (1999). Caudal is the Hox gene that specifies the most posterior Drosophile segment. *Nature* **400**, 873-7.

Morgan, T. H. (1892). Spiral modification of metamerism. *J Morphol* **7**, 245-251.

Morimoto, M., Takahashi, Y., Endo, M. and Saga, Y. (2005). The Mesp2 transcription factor establishes segmental borders by suppressing Notch activity. *Nature* **435**, 354-9.

Naiche, L. A., Holder, N. and Lewandoski, M. (2011). FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. *Proc Natl Acad Sci U S A* **108**, 4018-23.

Nakamura, Y., Gojobori, T. and Ikemura, T. (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res* **28**, 292.

Nakao, H. (2010). Characterization of Bombyx embryo segmentation process: expression profiles of engrailed, even-skipped, caudal, and wnt1/wingless homologues. *J Exp Zool B Mol Dev Evol* **314**, 224-31.

Nambu, P. A. and Nambu, J. R. (1996). The Drosophila fish-hook gene encodes a HMG domain protein essential for segmentation and CNS development. *Development* **122**, 3467-75.

Natori, K., Tajiri, R., Furukawa, S. and Kojima, T. (2012). Progressive tarsal patterning in the Drosophila by temporally dynamic regulation of transcription factor genes. *Dev Biol* **361**, 450-62.

Niehrs, C. (2010). On growth and form: a Cartesian coordinate system of Wnt and BMP signaling specifies bilaterian body axes. *Development* **137**, 845-57.

Niwa, N., Inoue, Y., Nozawa, A., Saito, M., Misumi, Y., Ohuchi, H., Yoshioka, H. and Noji, S. (2000). Correlation of diversity of leg morphology in Gryllus bimaculatus (cricket) with divergence in dpp expression pattern during leg development. *Development* **127**, 4373-4381.

Niwa, N., Saitoh, M., Ohuchi, H., Yoshioka, H. and Noji, S. (1997). Correlation between Distal-less expression patterns and structures of appendages in development of the two-spotted cricket, Gryllus bimaculatus. *Zoolog Sci* **14**, 115-125.

Niwa, Y., Shimojo, H., Isomura, A., Gonzalez, A., Miyachi, H. and Kageyama, R. (2011). Different types of oscillations in Notch and Fgf signaling regulate the spatiotemporal periodicity of somitogenesis. *Genes Dev* **25**, 1115-20.

Nulsen, C. and Nagy, L. M. (1999). The role of wingless in the development of multibranched crustacean limbs. *Development Genes and Evolution* **209**, 340-348.

Nunes da Fonseca, R., van der Zee, M. and Roth, S. (2010). Evolution of extracellular Dpp modulators in insects: The roles of tolloid and twisted-gastrulation in dorsoventral patterning of the Tribolium embryo. *Dev Biol* **345**, 80-93.

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

Nusslein-Volhard, C., Wieschaus, E. and Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster. *I. Zygotic loci on the second chromosome.* **193**, 267-282.

O'Day, K. E. (2006). Notch signalling and segmentation in Parhyale hawaiensis. In *Molecular and Cell Biology*, vol. PhD (ed., pp. 194. Berkeley: University of California Berkeley.

Oates, A. C., Morelli, L. G. and Ares, S. (2012). Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. *Development* **139**, 625-39.

Oda, H., Nishimura, O., Hirao, Y., Tarui, H., Agata, K. and Akiyama-Oda, Y. (2007). Progressive activation of Delta-Notch signaling from around the blastopore is required to set up a functional caudal lobe in the spider Achaearanea tepidariorum. *Development* **134**, 2195-2205.

Oginuma, M., Niwa, Y., Chapman, D. L. and Saga, Y. (2008). Mesp2 and Tbx6 cooperatively create periodic patterns coupled with the clock machinery during mouse somitogenesis. *Development* **135**, 2555-62.

Olesnicky, E. C., Brent, A. E., Tonnes, L., Walker, M., Pultz, M. A., Leaf, D. and Desplan, C. (2006). A caudal mRNA gradient controls posterior development in the wasp Nasonia. *Development* **133**, 3973-82.

Osborne, P. W. and Dearden, P. K. (2005). Expression of Pax group III genes in the honeybee (Apis mellifera). *Dev Genes Evol* **215**, 499-508.

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

Ozbudak, E. M. and Pourquie, O. (2008). The vertebrate segmentation clock: the tip of the iceberg. *Current Opinion in Genetics & Development* **18**, 317-323.

Palmeirim, I., Henrique, D., IshHorowicz, D. and Pourquie, O. (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91**, 639-648.

Panganiban, G., Irvine, S. M., Lowe, C., Roehl, H., Corley, L. S., Sherbon, B., Grenier, J. K., Fallon, J. F., Kimble, J., Walker, M. et al. (1997). The origin and evolution of animal appendages. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **94**, 5162-5166.

Panganiban, G., Nagy, L. and Carroll, S. B. (1994). The role of the Distal-less gene in the development and evolution of insect limbs. *Curr Biol* **4**, 671-5.

Pannese, M., Polo, C., Andreazzoli, M., Vignali, R., Kablar, B., Barsacchi, G. and Boncinelli, E. (1995). The Xenopus homologue of Otx2 is a maternal homeobox gene that demarcates and specifies anterior body regions. *Development* **121**, 707-20.

Patel, N. H., Ball, E. E. and Goodman, C. S. (1992). Changing role of even-skipped during the evolution of insect pattern formation. *Nature* **357**, 339-42.

Patel, N. H., Condron, B. G. and Zinn, K. (1994). Pair-rule expression patterns of even-skipped are found in both short- and long-germ beetles. *Nature* **367**, 429-434.

Patel, N. H., Martin-Blanco, E., Coleman, K. G., Poole, S. J., Ellis, M. C., Kornberg, T. B. and Goodman, C. S. (1989). Expression of Engrailed Proteins in Arthropods, Annelids, and Chordates. *Cell* **58**, 955-968.

Pechmann, M., Khadjeh, S., Sprenger, F. and Prpic, N. M. (2010). Patterning mechanisms and morphological diversity of spider appendages and their importance for spider evolution. *Arthropod Struct Dev* **39**, 453-67.

Pechmann, M., McGregor, A. P., Schwager, E. E., Feitosa, N. M. and Damen, W. G. (2009). Dynamic gene expression is required for anterior regionalization in a spider. *Proc Natl Acad Sci U S A* **106**, 1468-72.

Peel, A. (2004). The evolution of arthropod segmentation mechanisms. *Bioessays* **26**, 1108-1116.

Peel, A. and Akam, M. (2003). Evolution of segmentation: Rolling back the clock. *Current Biology* **13**, 708-710.

Peel, A. D., Chipman, A. D. and Akam, M. (2005). Arthropod segmentation: Beyond the Drosophila paradigm. *Nature Reviews Genetics* **6**, 905-916.

Petersen, C. P. and Reddien, P. W. (2009). Wnt signaling and the polarity of the primary body axis. *Cell* **139**, 1056-68.

Philippe, H., Lartillot, N. and Brinkmann, H. (2005). Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. *Mol Biol Evol* **22**, 1246-53.

Pi, H., Huang, Y. C., Chen, I. C., Lin, C. D., Yeh, H. F. and Pai, L. M. (2011). Identification of 11-amino acid peptides that disrupt Notch-mediated processes in Drosophila. *J Biomed Sci* **18**, 42.

Popadic, A. (2005). Global evolution of nubbin expression patterns in arthropods: emerging view. *Evol Dev* **7**, 359-61.

Posnien, N. and Bucher, G. (2010). Formation of the insect head involves lateral contribution of the intercalary segment, which depends on Tc-labial function. *Dev Biol* **338**, 107-16.

Posnien, N., Schinko, J. B., Kittelmann, S. and Bucher, G. (2010). Genetics, development and composition of the insect head--a beetle's view. *Arthropod Struct Dev* **39**, 399-410.

Pourquie, O. (1999). Notch around the clock. *Curr Opin Genet Dev* **9**, 559-65.

Pourquie, O. (2011). Vertebrate segmentation: from cyclic gene networks to scoliosis. *Cell* **145**, 650-63.

Pownall, M. E., Tucker, A. S., Slack, J. M. and Isaacs, H. V. (1996). eFGF, Xcad3 and Hox genes form a molecular pathway that establishes the anteroposterior axis in Xenopus. *Development* **122**, 3881-92.

Prazak, L., Fujioka, M. and Gergen, J. P. (2010). Non-additive interactions involving two distinct elements mediate sloppy-paired regulation by pair-rule transcription factors. *Dev Biol* **344**, 1048-59.

Prinos, P., Joseph, S., Oh, K., Meyer, B. I., Gruss, P. and Lohnes, D. (2001). Multiple pathways governing Cdx1 expression during murine development. *Dev Biol* **239**, 257-69.

Prpic, N. M. and Damen, W. G. (2005). Diversification of nubbin expression patterns in arthropods: data from an additional spider species, Cupiennius salei. *Evol Dev* **7**, 276-9.

Prpic, N. M. and Damen, W. G. (2009). Notch-mediated segmentation of the appendages is a molecular phylotypic trait of the arthropods. *Dev Biol* **326**, 262-71.

Prpic, N. M., Janssen, R., Wigand, B., Klingler, M. and Damen, W. G. M. (2003). Gene expression in spider appendages reveals reversal of exd/hth spatial specificity, altered leg gap gene dynamics, and suggests divergent distal morphogen signaling. *Developmental Biology* **264**, 119-140.

Prpic, N. M. and Tautz, D. (2003). The expression of the proximodistal axis patterning genes Distal-less and dachshund in the appendages of Glomeris marginata (Myriapoda: Diplopoda) suggests a special role of these genes in patterning the head appendages. *Dev Biol* **260**, 97-112.

Prud'homme, B., de Rosa, R., Arendt, D., Julien, J. F., Pajaziti, R., Dorresteijn, A. W., Adoutte, A., Wittbrodt, J. and Balavoine, G. (2003). Arthropod-like expression patterns of engrailed and wingless in the annelid Platynereis dumerilii suggest a role in segment formation. *Curr Biol* **13**, 1876-81.

Prud'homme, B., Gompel, N., Rokas, A., Kassner, V. A., Williams, T. M., Yeh, S. D., True, J. R. and Carroll, S. B. (2006). Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* **440**, 1050-3.

Pueyo, J. I. and Couso, J. P. (2008). The 11-aminoacid long Tarsal-less peptides trigger a cell signal in Drosophila leg development. *Developmental Biology* **324**, 192-201.

Pueyo, J. I. and Couso, J. P. (2011). Tarsal-less peptides control Notch signalling through the Shavenbaby transcription factor. *Dev Biol* **355**, 183-93.

Pueyo, J. I., Lanfear, R. and Couso, J. P. (2008). Ancestral Notch-mediated segmentation revealed in the cockroach Periplaneta americana. *Proc Natl Acad Sci U S A* **105**, 16614-9.

Pultz, M. A., Westendorf, L., Gale, S. D., Hawkins, K., Lynch, J., Pitt, J. N., Reeves, N. L., Yao, J. C., Small, S., Desplan, C. et al. (2005). A major role for zygotic hunchback in patterning the Nasonia embryo. *Development* **132**, 3705-15.

Ramsay, G. W. (1959). Spiral segmentation in two species of New Zealand Weta (Orthoptera, Gryllacridoidea, Henicidae). *Trans Roy Soc NZ* **86**, 393-394.

Rauskolb, C. and Irvine, K. D. (1999). Notch-mediated segmentation and growth control of the Drosophila leg. *Developmental Biology* **210**, 339-350.

Reed, R. D., Papa, R., Martin, A., Hines, H. M., Counterman, B. A., Pardo-Diaz, C., Jiggins, C. D., Chamberlain, N. L., Kronforst, M. R., Chen, R. et al. (2011). optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* **333**, 1137-41.

Richards, G. S. and Degnan, B. M. (2012). The expression of Delta ligands in the sponge Amphimedon queenslandica suggests an ancient role for Notch signaling in metazoan development. *Evodevo* **3**, 15.

Richards, G. S., Simionato, E., Perron, M., Adamska, M., Vervoort, M. and Degnan, B. M. (2008). Sponge genes provide new insight into the evolutionary origin of the neurogenic circuit. *Curr Biol* **18**, 1156-61.

Riddiford, N. and Olson, P. D. (2011). Wnt gene loss in flatworms. *Dev Genes Evol* **221**, 187-97.

Riedel-Kruse, I. H., Muller, C. and Oates, A. C. (2007). Synchrony dynamics during initiation, failure, and rescue of the segmentation clock. *Science* **317**, 1911-5.

Rivera, A. S., Gonsalves, F. C., Song, M. H., Norris, B. J. and Weisblat, D. A. (2005). Characterization of Notch-class gene expression in segmentation stem cells and segment founder cells in Helobdella robusta (Lophotrochozoa; Annelida; Clitellata; Hirudinida; Glossiphoniidae). *Evolution & Development* **7**, 588-599.

Rivera, A. S. and Weisblat, D. A. (2009). And Lophotrochozoa makes three: Notch/Hes signaling in annelid segmentation. *Development Genes and Evolution* **219**, 37-43.

Rohr, K. B., Tautz, D. and Sander, K. (1999). Segmentation gene expression in the mothmidge Clogmia albipunctata (Diptera, psychodidae) and other primitive dipterans. *Dev Genes Evol* **209**, 145-54.

Rosenberg, M. I., Lynch, J. A. and Desplan, C. (2009). Heads and tails: evolution of antero-posterior patterning in insects. *Biochim Biophys Acta* **1789**, 333-42.

Royet, J. and Finkelstein, R. (1997). Establishing primordia in the Drosophila eye-antennal imaginal disc: the roles of decapentaplegic, wingless and hedgehog. *Development* **124**, 4793-4800.

Ruiz i Altaba, A. (1990). Neural expression of the Xenopus homeobox gene Xhox3: evidence for a patterning neural signal that spreads through the ectoderm. *Development* **108**, 595-604.

Russell, S. R., Sanchez-Soriano, N., Wright, C. R. and Ashburner, M. (1996). The Dichaete gene of Drosophila melanogaster encodes a SOX-domain protein required for embryonic segmentation. *Development* **122**, 3669-76.

Ryan, J. F. and Baxevanis, A. D. (2007). Hox, Wnt, and the evolution of the primary body axis: insights from the early-divergent phyla. *Biol Direct* **2**, 37.

Saga, Y. and Takahashi, Y. (2008). Mesp-family genes are required for segmental patterning and segmental border formation. *Adv Exp Med Biol* **638**, 113-23.

Sarrazin, A. F., Peel, A. D. and Averof, M. (2012). A segmentation clock with two-segment periodicity in insects. *Science* **336**, 338-41.

Saudemont, A., Dray, N., Hudry, B., Le Gouar, M., Vervoort, M. and Balavoine, G. (2008). Complementary striped expression patterns of NK homeobox genes during segment formation in the annelid Platynereis. *Dev Biol* **317**, 430-43.

Savard, J., Marques-Souza, H., Aranda, M. and Tautz, D. (2006). A segmentation gene in *Tribollium* produces a polycistronic mRNA that codes for multiple conserved peptides. *Cell* **126**, 559-569.

Savory, J. G., Bouchard, N., Pierre, V., Rijli, F. M., De Repentigny, Y., Kothary, R. and Lohnes, D. (2009). Cdx2 regulation of posterior development through non-Hox targets. *Development* **136**, 4099-110.

Savory, J. G., Mansfield, M., Rijli, F. M. and Lohnes, D. (2011). Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7. *Development* **138**, 1361-70.

Sawada, A., Shinya, M., Jiang, Y. J., Kawakami, A., Kuroiwa, A. and Takeda, H. (2001). Fgf/MAPK signalling is a crucial positional cue in somite boundary formation. *Development* **128**, 4873-80.

Schaeper, N. D., Prpic, N. M. and Wimmer, E. A. (2009). A conserved function of the zinc finger transcription factor Sp8/9 in allometric appendage growth in the milkweed bug Oncopeltus fasciatus. *Dev Genes Evol* **219**, 427-35.

Schinko, J. B., Kreuzer, N., Offen, N., Posnien, N., Wimmer, E. A. and Bucher, G. (2008). Divergent functions of orthodenticle, empty spiracles and buttonhead in early head patterning of the beetle Tribolium castaneum (Coleoptera). *Developmental Biology* **317**, 600-613.

Schnellhammer, I. (2012). Evolution früher Faktoren der Segmentierungskaskade: Funktionelle Untersuchungen in Bruchidius, Tribolium und Oncopeltus. In *Der Naturwissenschaftlichen Fakultät*, vol. PhD (ed. Erlangen-Nürnberg: Friedrich-Alexander-Universität Erlangen-Nürnberg.

Scholtz, G. (2002). The articulata hypothesis - or what is a segment? *Org Divers Evol* **2**, 197-215.

Schoppmeier, M. and Damen, W. G. (2005a). Expression of Pax group III genes suggests a single-segmental periodicity for opisthosomal segment patterning in the spider Cupiennius salei. *Evol Dev* **7**, 160-9.

Schoppmeier, M. and Damen, W. G. M. (2005b). Suppressor of Hairless and Presenilin phenotypes imply involvement of canonical Notch-signalling in segmentation of the spider Cupiennius salei. *Developmental Biology* **280**, 211-224.

Schroder, R. (2003). The genes orthodenticle and hunchback substitute for bicoid in the beetle Tribolium. *Nature* **422**, 621-625.

Schroder, R., Jay, D. G. and Tautz, D. (1999). Elimination of EVE protein by CALI in the short germ band insect Tribolium suggests a conserved pair-rule function for even skipped. *Mech Dev* **80**, 191-5.

Schroter, C. and Oates, A. C. (2010). Segment number and axial identity in a segmentation clock period mutant. *Curr Biol* **20**, 1254-1258.

Schulz, C., Schroder, R., Hausdorf, B., Wolff, C. and Tautz, D. (1998). A caudal homologue in the short germ band beetle Tribolium shows similarities to both, the Drosophila and the vertebrate caudal expression patterns. *Dev Genes Evol* **208**, 283-9.

Schwager, E. E., Pechmann, M., Feitosa, N. M., McGregor, A. P. and Damen, W. G. (2009). hunchback functions as a segmentation gene in the spider Achaearanea tepidariorum. *Curr Biol* **19**, 1333-40.

Seaver, E. C. and Kaneshige, L. M. (2006). Expression of 'segmentation' genes during larval and juvenile development in the polychaetes Capitella sp. I and H. elegans. *Dev Biol* **289**, 179-94.

Seaver, E. C. and Shankland, M. (2001). Establishment of segment polarity in the ectoderm of the leech Helobdella. *Development* **128**, 1629-41.

Seaver, E. C., Yamaguchi, E., Richards, G. S. and Meyer, N. P. (2012). Expression of the pair-rule gene homologs runt, Pax3/7, even-skipped-1 and even-skipped-2 during larval and juvenile development of the polychaete annelid Capitella teleta does not support a role in segmentation. *Evodevo* **3**, 8.

Seebald, J. L. and Szeto, D. P. (2011). Zebrafish eve1 regulates the lateral and ventral fates of mesodermal progenitor cells at the onset of gastrulation. *Dev Biol* **349**, 78-89.

Shashikant, C. S., Kim, C. B., Borbely, M. A., Wang, W. C. and Ruddle, F. H. (1998). Comparative studies on mammalian Hoxc8 early enhancer sequence reveal a baleen whale-specific deletion of a cis-acting element. *Proc Natl Acad Sci U S A* **95**, 15446-51.

Shimizu, T., Bae, Y. K., Muraoka, O. and Hibi, M. (2005). Interaction of Wnt and caudal-related genes in zebrafish posterior body formation. *Dev Biol* **279**, 125-41.

Shinmyo, Y., Mito, T., Matsushita, T., Sarashina, I., Miyawaki, K., Ohuchi, H. and Noji, S. (2005). caudal is required for gnathal and thoracic patterning and for posterior elongation in the intermediate-germband cricket Gryllus bimaculatus. *Mech Dev* **122**, 231-9.

Sobels, F. H. (1952). Genetics and morphology of the genotype asymmetric with special reference to its abnormal abdomen character; Drosophila melanogaster. *Genetica* **26**, 117-279.

Sommer, R. and Tautz, D. (1991). Segmentation gene expression in the housefly Musca domestica. *Development* **113**, 419-30.

Sommer, R. J. and Tautz, D. (1993). Involvement of an orthologue of the Drosophila pair-rule gene hairy in segment formation of the short germ-band embryo of Tribolium (Coleoptera). *Nature* **361**, 448-50.

Song, M. H., Huang, F. Z., Chang, G. Y. and Weisblat, D. A. (2002). Expression and function of an even-skipped homolog in the leech Helobdella robusta. *Development* **129**, 3681-92.

Song, M. H., Huang, F. Z., Gonsalves, F. C. and Weisblat, D. A. (2004). Cell cycle-dependent expression of a hairy and Enhancer of split (hes) homolog during cleavage and segmentation in leech embryos. *Dev Biol* **269**, 183-95.

Stollewerk, A. (2002). Recruitment of cell groups through Delta/Notch signalling during spider neurogenesis. *Development* **129**, 5339-5348.

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

Takada, S., Stark, K. L., Shea, M. J., Vassileva, G., McMahon, J. A. and McMahon, A. P. (1994). Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes Dev* **8**, 174-89.

Takahashi, Y., Koizumi, K., Takagi, A., Kitajima, S., Inoue, T., Koseki, H. and Saga, Y. (2000). Mesp2 initiates somite segmentation through the Notch signalling pathway. *Nat Genet* **25**, 390-6.

Tautz, D. (2004). Segmentation. *Developmental Cell* **7**, 301-312.

Technau, U., Rudd, S., Maxwell, P., Gordon, P. M. K., Saina, M., Grasso, L. C., Hayward, D. C., Sensen, C. W., Saint, R., Holstein, T. W. et al. (2005). Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. *Trends in Genetics* **21**, 633-639.

Thamm, K. and Seaver, E. C. (2008). Notch signaling during larval and juvenile development in the polychaete annelid Capitella sp. I. *Developmental Biology* **320**, 304-318.

Theron, E., Hawkins, K., Bermingham, E., Ricklefs, R. E. and Mundy, N. I. (2001). The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, Coereba flaveola. *Curr Biol* **11**, 550-7.

Trautwein, M. D., Wiegmann, B. M., Beutel, R., Kjer, K. M. and Yeates, D. K. (2012). Advances in insect phylogeny at the dawn of the postgenomic era. *Annu Rev Entomol* **57**, 449-68.

Treisman, J., Harris, E. and Desplan, C. (1991). The paired box encodes a second DNA-binding domain in the paired homeo domain protein. *Genes Dev* **5**, 594-604.

Tupy, J. L., Bailey, A. M., Dailey, G., Evans-Holm, M., Siebel, C. W., Misra, S., Celniker, S. E. and Rubin, G. M. (2005). Identification of putative noncoding polyadenylated transcripts in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 5495-5500.

Turchyn, N., Couso, J. P., Hrycaj, S., Chesebro, J. and Popadic, A. (2011). Embryonic functions of nubbin in hemimetabolous insects. *Dev Biol* **357**, 83-95.

Van Breugel, F. M. A. and Langhout, B. V. Z. (1983). The Notch locus of Drosophila hydei: alleles, phenotypes and functional organization. *Genetics* **103**, 197-217.

van de Ven, C., Bialecka, M., Neijts, R., Young, T., Rowland, J. E., Stringer, E. J., Van Rooijen, C., Meijlink, F., NÃ³voa, A., Freund, J.-N. et al. (2011). Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone. *Development* **138**, 3451-3462.

van den Akker, E., Forlani, S., Chawengsaksophak, K., de Graaff, W., Beck, F., Meyer, B. I. and Deschamps, J. (2002). Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* **129**, 2181-2193.

van der Zee, M., Stockhammer, O., von Levetzow, C., Nunes da Fonseca, R. and Roth, S. (2006). Sog/Chordin is required for ventral-to-dorsal Dpp/BMP transport and head formation in a short germ insect. *Proc Natl Acad Sci U S A* **103**, 16307-12.

Vorwald-Denholtz, P. P. and De Robertis, E. M. (2011). Temporal pattern of the posterior expression of Wingless in Drosophila blastoderm. *Gene Expr Patterns* **11**, 456-463.

Wahl, M. B., Deng, C., Lewandoski, M. and Pourquie, O. (2007). FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development* **134**, 4033-41.

Wang, C. and Lehmann, R. (1991). Nanos is the localized posterior determinant in Drosophila. *Cell* **66**, 637-47.

Wang, L. and Denburg, J. L. (1992). A role for proteoglycans in the guidance of a subset of pioneer axons in cultured embryos of the cockroach. *Neuron* **8**, 701-714.

Wasik, B. R. and Moczek, A. P. (2011). Decapentaplegic (dpp) regulates the growth of a morphological novelty, beetle horns. *Dev Genes Evol* **221**, 17-27.

Wei, Z., Range, R., Angerer, R. and Angerer, L. (2012). Axial patterning interactions in the sea urchin embryo: suppression of nodal by Wnt1 signaling. *Development* **139**, 1662-9.

Williams, T., Blachuta, B., Hegna, T. A. and Nagy, L. M. (2012). Decoupling elongation and segmentation: notch involvement in anostracan crustacean segmentation. *Evol Dev* **14**, 372-82.

Wilson, M. J., Havler, M. and Dearden, P. K. (2010a). Giant, Kruppel, and caudal act as gap genes with extensive roles in patterning the honeybee embryo. *Dev Biol* **339**, 200-11.

Wilson, M. J., McKelvey, B. H., van der Heide, S. and Dearden, P. K. (2010b). Notch signaling does not regulate segmentation in the honeybee, Apis mellifera. *Dev Genes Evol* **220**, 179-90.

Wu, J. and Cohen, S. M. (1999). Proximodistal axis formation in the Drosophila leg: subdivision into proximal and distal domains by Homothorax and Distal-less. *Development* **126**, 109-117.

Wu, L. H. and Lengyel, J. A. (1998). Role of caudal in hindgut specification and gastrulation suggests homology between Drosophila amnioproctodeal invagination and vertebrate blastopore. *Development* **125**, 2433-42.

Young, T., Rowland, J. E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J. N. et al. (2009). Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* **17**, 516-26.

Zhang, N. and Gridley, T. (1998). Defects in somite formation in lunatic fringe-deficient mice. *Nature* **394**, 374-7.

APPENDIX 1

***Periplaneta americana* codon usage bias.** The codon usage database (http://kazusa.or.jp/codon) was used in order to reduce the high rates of degeneracy for some of the primers used to isolate *Periplaneta* segmentation genes. The database analysed 46 coding sequences containing 26,676 codons for *Periplaneta americana* found in GenBank. In each box, the codon triplet is in bold, followed by the single letter amino acid code. The remaining numbers represent, from left to right: the fraction of a particular codon used compared to other synonymous codons, the frequency of appearance of that codon per every thousand codons identified, and the total number of times that codon appears in the 21,676 codons analysed. Table reproduced here from Nakamura et al. (2000).

Fields: [**triplet**] [amino acid] [fraction] [frequency: per thousand] (number])

UUU F 0.29 11.4 (248)	**UCU** S 0.17 14.0 (303)	**UAU** Y 0.33 11.2 (243)	**UGU** C 0.43 08.6 (186)
UUC F 0.71 28.4 (616)	**UCC** S 0.18 15.2 (330)	**UAC** Y 0.67 22.6 (489)	**UGC** C 0.57 11.4 (248)
UUA L 0.06 05.4 (116)	**UCA** S 0.13 11.1 (241)	**UAA** * 0.39 00.8 (018)	**UGA** * 0.37 00.8 (017)
UUG L 0.17 14.2 (307)	**UCG** S 0.12 10.1 (220)	**UAG** * 0.24 00.5 (011)	**UGG** W 1.00 10.6 (229)
CUU L 0.13 11.1 (240)	**CCU** P 0.29 14.7 (319)	**CAU** H 0.39 10.1 (220)	**CGU** R 0.15 08.6 (186)
CUC L 0.21 17.7 (383)	**CCC** P 0.26 13.1 (283)	**CAC** H 0.61 15.6 (339)	**CGC** R 0.20 11.5 (249)
CUA L 0.09 07.3 (158)	**CCA** P 0.27 13.9 (301)	**CAA** Q 0.39 14.6 (317)	**CGA** R 0.13 07.6 (165)
CUG L 0.33 27.8 (602)	**CCG** P 0.18 09.1 (198)	**CAG** Q 0.61 22.6 (489)	**CGG** R 0.09 05.4 (118)
AUU I 0.31 15.0 (325)	**ACU** T 0.27 13.3 (289)	**AAU** N 0.38 18.5 (400)	**AGU** S 0.16 13.5 (293)
AUC I 0.50 24.3 (526)	**ACC** T 0.24 12.2 (265)	**AAC** N 0.62 29.7 (643)	**AGC** S 0.23 18.9 (410)
AUA I 0.19 09.3 (201)	**ACA** T 0.27 13.7 (297)	**AAA** K 0.38 22.1 (478)	**AGA** R 0.25 14.3 (310)
AUG M 1.00 24.0 (521)	**ACG** T 0.22 11.0 (239)	**AAG** K 0.62 35.5 (770)	**AGG** R 0.17 09.9 (215)
GUU V 0.22 15.5 (335)	**GCU** A 0.27 18.7 (406)	**GAU** D 0.45 26.8 (581)	**GGU** G 0.21 13.2 (286)
GUC V 0.22 15.2 (329)	**GCC** A 0.31 21.2 (459)	**GAC** D 0.55 32.1 (696)	**GGC** G 0.31 19.3 (419)
GUA V 0.17 11.7 (254)	**GCA** A 0.26 18.1 (392)	**GAA** E 0.47 33.0 (715)	**GGA** G 0.35 21.5 (466)
GUG V 0.38 26.3 (571)	**GCG** A 0.16 10.9 (237)	**GAG** E 0.53 36.8 (798)	**GGG** G 0.12 07.4 (161)

Appendix 1

APPENDIX 2

A) Primers for *Pa-caudal*

Degenerate

Pa-cad forward: 5’ GAGCTSGAGAAGGARTT 3’
Pa-cad reverse : 5’ RCAGCACCARTGRAABGTRCA 3’

Specific

Pa-cad 5’RACE reverse (outer): 5’ TCTTCACCTGTCGTTCTGAGAGTC 3’
Pa-cad 5’RACE reverse (inner): 5’ CTTACGCTTGATGGTGATGTACC 3’
Pa-cad 3’RACE forward (outer): 5’ ACTACAGCCGGTACAT CACCATCA 3’
Pa-cad 3’RACE forward (inner): 5’ CTCTCAGAACGACAGGTGAAGATC 3’

RNAi (T7 polymerase priming sequence underlined)

Pa-cad RNAi forward:
5’ TAATACGACTCACTATAGGGACCGGAGTGGACCAGCGCGG 3’
Pa-cad RNAi reverse:
5’ TAATACGACTCACTATAGGGATCCTTCTCCAGTTCTAGTCGC 3’

B) Primers for *Pa-Wnt1*

Degenerate

Pa-Wnt1 forward 1 (outer): 5’ CGNCCGVTGGAACTGYTCNAC 3’
Pa-Wnt1 reverse 1 (outer): 5’ RCAGCACCARTGRAABGTRCA 3’
Pa-Wnt1 forward 2 (inner): 5’ TGGGGYGGHTGCTSBGAYAAAT 3’
Pa-Wnt1 reverse 2 (inner): 5’ GCCNCKKCCGSAGCACATVAG 3’

Specific

Pa-Wnt1 5’RACE reverse 1 (outer): 5’ TGGAGATTCATCTTCTCGCGGAGGT 3’
Pa-Wnt1 5’RACE reverse 1 (inner): 5’ CTTGTCGAGCAGTTCCACCTCCTG 3’
Pa-Wnt1 5’RACE reverse 2 (outer): 5’ GAATTCGCGGGAGAACTTGAAGCC 3’
Pa-Wnt1 5’RACE reverse 2 (inner): 5’ AACTAGACGTCTCTGCTTCCGCCG 3’
Pa-Wnt1 3’RACE forward 1 (outer): 5’ TTCTGCGAGCGCAACCCGAGA 3’
Pa-Wnt1 3’RACE forward 1 (inner): 5’ GTCAGTGCAACGATACGTCGATAGGCT3’

C) Primers for *Pa-even-skipped*

Degenerate

Pa-eve forward 1 (outer): 5’ ATMMGNMGVTAYMGVACNGC 3’
Pa-eve forward 2 (inner): 5’ CSTTCACBMGSGAVCAGYTG 3’
Pa-eve reverse 1 (outer): 5’ ACCTTGATNGTGSWYTCBGG 3’
Pa-eve reverse 2 (inner): 5’ GTABGGCTKGAASARCTTSAA 3’

Specific

Pa-eve forward 1 (outer): 5’ AGAAGGAGTTCCGCAACTCAACCT 3’
Pa-eve forward 2 (inner): 5’ TGCCCGAGTCCACCATCAAG 3’
Pa-eve reverse 1 (outer): 5’ TTCTGGAACCACACCTTGATGGTG 3’
Pa-eve reverse 2 (inner): 5’ TGGACTCGGGACAGGTTGAGTT 3’

APPENDIX 2

RNAi (T7 polymerase priming sequence underlined)
Pa-eve RNAi forward:
5' TAATACGACTCACTATAGGGACAAGCTCCACCTCAAGACCTCG 3'
Pa-eve RNAi reverse:
5' TAATACGACTCACTATAGGGAGCGGTTCTGGAACCACACCTTG 3'

D) Primers for *Pa-runt*
Degenerate
Pa-run forward 1 (outer): 5' RCNRYNATGAARAAYCARGTNGC 3'
Pa-run forward 2 (inner): 5' MRNTTYAAYGAYYTNMGNTTYGTNGG 3'
Pa-run reverse 1 (outer): 5' CKNGGYTCNCKNGGNCCRTC 3'

Specific
Pa-run 5'RACE reverse 1 (outer): 5' AACACACTGTACGCTGCCGTTAGC 3'
Pa-run 5'RACE reverse 2 (inner): 5' TGGTATTCTTCCGCCTGGAGTC 3'
Pa-run 3'RACE forward 1 (outer): 5' TTCTCGCTGACCATCGTCATCAG 3'
Pa-run 3'RACE forward 2 (inner): 5' TCAGTTCGACGCCCTTTCAGATCG 3'

RNAi (T7 polymerase priming sequence underlined)
Pa-run RNAi forward:
5' TAATACGACTCACTATAGGGAAATGCTAACGGCAGCTACAG 3'
Pa-run RNAi reverse:
5' TAATACGACTCACTATAGGGACTTCCCTCTGCCACTTCTG 3'

E) Primers for *Pa-pairberry*
Degenerate
Pa-pby forward 1 (outer): 5' GGNGGNGTNTTYATHAAYGG 3'
Pa-pby forward 2 (inner): 5' MARATHGTNGARATGGC 3'
Pa-pby reverse 1 (outer): 5' RTTNSWRAACCANACYTG 3'
Pa-pby reverse 2 (inner): 5' RTANACRTCNGGRTAYTG 3'

Specific
Pa-pby 5'RACE reverse 1 (outer): 5' ATGCTGCCGGTCTCCTGGTATCG 3'
Pa-pby 5'RACE reverse 2 (inner): 5' TATCGGTCCAGAATCTTGGACAC 3'

RNAi (T7 polymerase priming sequence underlined)
Pa-pby RNAi forward:
5' TAATACGACTCACTATAGGGAGGAGTTCGACCATGCGTCATC 3'
Pa-pby RNAi reverse:
5' TAATACGACTCACTATAGGGAGCGTTGCTTGCGCTTCAG 3'

APPENDIX 2

F) Primers for *Pa-sloppy-paired*

Degenerate

Pa-slp forward 1 (outer): 5' AARCCNCCNTWYWSNTAYAAYGC 3'
Pa-slp forward 2 (inner): 5' YTNATHATGATGGCNATHMG 3'
Pa-slp reverse 1 (outer): 5' TTNCCNGTNGTNCCNCCDARTAA 3'
Pa-slp reverse2 (inner): 5' TACACCATRTTRCCCNTTYCC 3'

Specific

Pa-slp 5'RACE reverse 1 (outer): 5' ATGCTGAACGAGGACTTGAGG 3'
Pa-slp 5'RACE reverse 2 (inner): 5' ATCTTCACCATGTTGACATGC 3'
Pa-slp 3'RACE forward 1 (outer): 5' ACTACTGGATGCTGGACGCCAG 3'
Pa-slp 3'RACE forward 2 (inner): 5' AGGACGTCTTTATCGGCGGCAC 3'

RNAi (T7 polymerase priming sequence underlined)

Pa-slp RNAi forward:
5' TAATACGACTCACTATAGGGATGTCAACATGGTGAAGATCGAG 3'
Pa-slp RNAi reverse:
5' TAATACGACTCACTATAGGGAAGCATCCAGTAGTTGCCCTTG 3'

G) Primers for *tarsal-less*

Degenerate

tal forward 1.1: 5' ATGGAYCCHAA**G**CARYTSGA 3'
tal forward 1.2: 5' ATGGAYCCHAA**R**CARYTSGA 3'
tal forward 2.1: 5' CAYYTSGAYCCHACNGG 3'
tal reverse 1.1: 5' **Y**YT**G**TASTRKCCNGTDGG 3'
tal reverse 1.2: 5' **A**YT**R**TASTRKCCNGTDGG 3'
tal reverse 2.1: 5' DGGRTCSTRNGT**C**TTDGG 3'
tal reverse 2.2: 5' DGGRTCSTRNGT**Y**TTDGG 3'

Specific (*Periplaneta*)

Pa-tal forward 1 (outer): 5' CACTGCCAAGTACCTCGCAC 3'
Pa-tal forward 2 (inner): 5' CTCAGCCTTCGACATGGATC 3'
Pa-tal reverse 1 (outer): 5' GATCCATGTCGAAGGCTGAG 3'
Pa-tal reverse 2 (inner): 5' GTGCGAGGTACTTGGACAGTG 3'

Periplaneta RNAi (T7 polymerase priming sequence underlined)

Pa-tal RNAi forward:
5' TAATACGACTCACTATAGGGAAGTTGCGGGCTAGCCCTCGA 3'
Pa-tal RNAi reverse:
5' TAATACGACTCACTATAGGGATTCTTCGTATGCCTTGCGTCT 3'

Specific (*Gryllus*)

Gb-tal forward 1 (outer): 5' CCACTTCAACTCGCAGCC 3'
Gb-tal forward 2 (inner): 5' ATGGGCCCCAAGACTCTG 3'
Gb-tal reverse 1 (outer): 5' GCTGCGGTTGTGCTGCCAC 3'
Gb-tal reverse 2 (inner): 5' GTCCACGAAGGCTGCCGTC 3'

APPENDIX 2

H) Primers for *Pa-nub* RNAi (T7 polymerase priming sequence underlined)

Pa-nub RNAi forward:

5’ TAATACGACTCACTATAGGGAGGAATTCGAGCAATTTG 3’

Pa-nub RNAi reverse:

5’ TAATACGACTCACTATAGGGAGCTCTAGAGGGTTGATG 3’

I) Primers for *Pa*-18S rDNA

Pa-18S forward: 5’ GTACCGGCGACGCATCTTTCA 3’

Pa-18S reverse: 5’ CTTTCGGCCAGGCAGGACAC 3’

APPENDIX 3

Nucleotide and protein sequences for *Periplaneta caudal* and phylogenetic analysis. (A) The 1420bp nucleotide sequence of *Pa-cad* codes for a 290 amino acid protein. Double underscore indicates the conserved homeodomain. (B) Phylogenetic analysis of the homeodomain protein sequence with other known *cad* sequences shows that *Pa-cad* is the *Periplaneta* orthologue and aligns with closely related insect species.

A) *Periplaneta caudal* nucleotide and protein sequences.

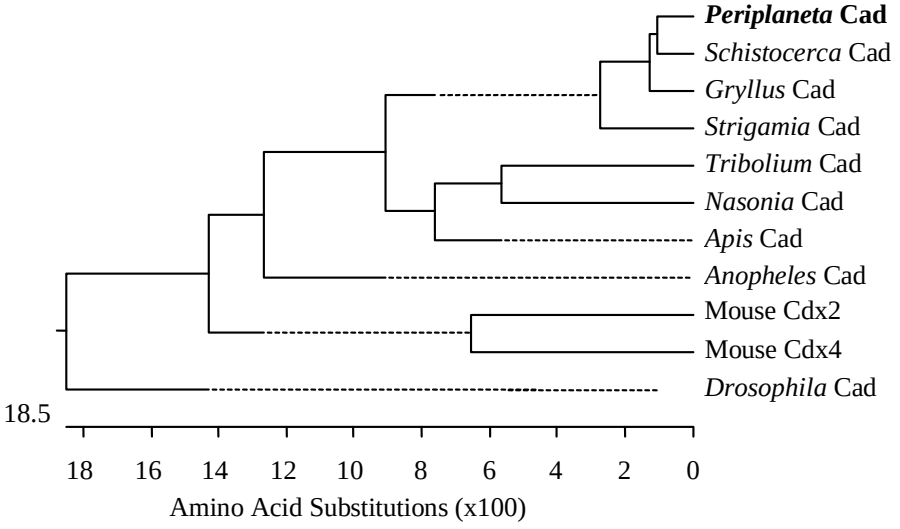
```
1     CGCGGCCGCGTCGACGCGCCGAGTTGTAGCGTGCGGGAGGTGTGCGAGAC   50
51    AGGATACGAACCCGTGTGCGCGAGCATGGAACCCGCGACTCTGTGACTTG  100
101   GGAACATGTGTCCGTGCTGACGAGATGTTGAGGAACAGGATGTGTAGTGT  150
151   CGTCTAGTGTCGGAGTGTGCTGTGCTCGGCTTCCAACAAGGTGGCATTCA  200
201   GACGAAAAAAGTGGAGGAGGAACTGGGTGCAAAGACAATGGTTTCATACT  250
1                                           M  V  S  Y     4
251   ACAACCCCTTAGCGATGTACCGGCACCAGCAGCAGGCCTCGGCTCCGGGG  300
5     Y  N  P  L  A  M  Y  R  H  Q  Q  Q  A  S  A  P  G    21
301   GGACCTCCGCAGTTCCACCACCCAACCAGCCCCGCCGCCACTTGGTACGG  350
22     G  P  P  Q  F  H  H  P  T  S  P  A  A  T  W  Y  G   38
351   GCCCCCCGGGAGTTACCAGCCCTCGCATCACCACCAGGTGCCTCCACCCC  400
39      P  P  G  S  Y  Q  P  S  H  H  H  Q  V  P  P  P    54
401   CGCTGCAACAGTACCCAAATTGCGTGCAGGACGATCAGCAGGGCACCGCG  450
55    P  L  Q  Q  Y  P  N  C  V  Q  D  D  Q  Q  G  T  A   71
451   GGGGCATGGCACCATCACCACCACCACATGTTTCAGCCGGAGTGGACCAG  500
72     G  A  W  H  H  H  H  H  H  M  F  Q  P  E  W  T  S   88
501   CGCGGGGGCCCCGGATTTTGGGAGTGCACCCGGAATGTCGCAGGGCCATC  550
89      A  G  A  P  D  F  G  S  A  P  G  M  S  Q  G  H    104
551   AGCCCTCTTCGGCGGGGCTCGAGGACCCCCAGTTACCGTCCCCGCCTATA  600
105   Q  P  S  S  A  G  L  E  D  P  Q  L  P  S  P  P  I   121
601   ACGGTGTCTAGTAGCGAACTGTCGAGTCCCGGCGCCGTTGGGGGTTCCGT  650
122    T  V  S  S  S  E  L  S  S  P  G  A  V  G  G  S  V  138
651   GACGCCCCCCCAGCACGCTGGGCGCCCCATTCCCGTCCGGAGCCCCTACG  700
139     T  P  P  Q  H  A  G  R  P  I  P  V  R  S  P  Y    154
701   AATGGATGAAGAAGCCGTCCTACCAGAGCCAGCCGAATCCAGTTGGCCCC  750
155   E  W  M  K  K  P  S  Y  Q  S  Q  P  N  P  V  G  P   171
751   AATCCCCCGCTTCTAGACCATACACGTGCCGGAATGCAGGAACTGTTGAG  800
172    N  P  P  L  L  D  H  T  R  A  G  M  Q  E  L  L  S  188
801   CAAAACGCGGACGAAgGACAAGTACCGAGTGGTATATAGCGATCACCAGC  850
189     K  T  R  T  K  D  K  Y  R  V  V  Y  S  D  H  Q    204
851   GACTAGARCTGGAGAAGGAGTTCCACTACAGCCGGTACATCACCATTAGG  900
205   R  L  E  L  E  K  E  F  H  Y  S  R  Y  I  T  I  R   221
901   CGTAAGGCGGAACTCGCTGCCAACCTGGGACTCTCAGAACGACAGGTGAA  950
222    R  K  A  E  L  A  A  N  L  G  L  S  E  R  Q  V  K  238
951   GATCTGGTTCCAGAACCGTCGCGCCAAGGAGCGCAAGCAGGTCAAGAAGC 1000
239     I  W  F  Q  N  R  R  A  K  E  R  K  Q  V  K  K    254
1001  GGGAGGAGCTGCTGCACAAGGGGAAGCTGGAGGCGGTGAGCGCGGCGCAC 1050
255   R  E  E  L  L  H  K  E  K  L  E  A  V  S  A  A  H   271
```

APPENDIX 3

```
1051 CACCAGCTGCAGCACGCGCACCAGCAGCCGCAGCAGGCCGCGCAGGACCT 1100
272   H  Q  L  Q  H  A  H  Q  Q  P  Q  Q  A  A  Q  D  L  288
1101 GCTTCTGTGACGTCGCAGCGCTGCTACTGCTCATCAGCATTCAGACAATA 1150
289    L  L  *                                          290
1151 AACGTGCGAGTCGCGCCATCGTACATGACAACACGGGTTGTTGTCGGTAT 1200
1201 GGATCTGGCACCACGAGTGCTGCAGTGAACTCGGACACGGAACAAACTTC 1250
1251 TACAGGTTGCACAATATCCATGGAATCAGACTTTTTCAATTGACACAAAT 1300
1301 CTGACGACCTCATTCTTTACTGTTATTTATTTTGTCTGTCTTATTTAGCA 1350
1351 TTGAACATACCGCAGTCTGCTAAGTTATCGTAGCCGCCGTCTAAATTTTC 1400
1401 TACTTTGTCCAAAAAAAAAA                               1420
```

B) Alignment and phylogenetic analysis of *Pa*-Cad protein. NCBI Accession numbers: *Schistocerca gregaria* (AAK56940); *Gryllus bimaculatus* (BAD51739); *Strigamia marginata* (AAT35589); *Tribolium castaneum* (AJ005421); *Apis mellifera* (AER27701); *Nasonia vitripennis* (ABI49991); *Anopheles gambiae* (AAD27585); *Drosophila melanogaster* (AAF53923); *Mus musculus*, Mouse Cdx2 (NP_031699); M*us musculus*, Mouse Cdx4 (NP_031700).

```
Majority          GKTRTKDKYRVVYTDHQRLELEKEFHYSRYITIRRKAELAAXLGLSERQVKIWFQNRRAKERKQVKK
                  ---------+---------+---------+---------+---------+---------+-------
                         10        20        30        40        50        60
                  ---------+---------+---------+---------+---------+---------+-------
Periplaneta Cad   SKTRTKDKYRVVYSDHQRLELEKEFHYSRYITIRRKAELAANLGLSERQVKIWFQNRRAKERKQVKK 67
Schistocerca Cad  GKTRTKDKYRVVYSDHQRLELEKEFHYSRYITIRRKAELAANLGLSERQV                  50
Gryllus Cad       GKTRTKDKYRVVYSDHQRLELEKEFHYSRYITIRRKAELAASLGLSERQVKIWFQNRRAKERKQVKK 67
Strigamia Cad     GKTRTKDKYRVVYSDIQRLELEKEFHYSRYITIRRKAELAQLLGLSERQVKIWFQNRRAKERKQVKK 67
Tribolium Cad     GKTRTKDKYRVVYTDHQRVELEKEFYYSRYITIRRKAELANSLGLSERQVKIWFQNRRAKERKQVKK 67
Nasonia Cad       GKTRTKDKYRVVYTEHQRLELEKEFYSSRYITIRRKAELASSLALSERQVKIWFQNRRAKDRKQSKK 67
Apis Cad          GKTRTKDKYRVVYTDHQRLELEKEFHYSRYITIRRKAELALSLSLSERQVKIWFQNRRAKERKQMKK 67
Anopheles Cad     GKTRTKDKYRVVYTDQQRLELEKEFHYTRYITIRRKAELAQNLQLSERQVKIWFQNRRAKDRKQKKK 67
Drosophila Cad    GKTRTKDKYRVVYTDFQRLELEKEYCTSRYITIRRKSELAQTLSLSERQVKIWFQNRRAKERKQNKK 67
Mouse Cdx2        VKTRTKDKYRVVYTDHQRLELEKEFHFSRYITIRRKSELAATLGLSERQVKIWFQNRRAKERKIKKK 67
Mouse Cdx4        GKTRTKEKYRVVYTDHQRLELEKEFHCNRYITIRRKSELAVNLGLSERQVKIWFQNRRAKERKMIKK 67
```



APPENDIX 4

Nucleotide and protein sequences for *Periplaneta Wnt1* and phylogenetic analysis. (A) The 2360bp nucleotide sequence of *Periplaneta Wnt1* codes for a 373 amino acid protein. (B) Phylogenetic analysis of a conserved 108aa region of the *Pa-Wnt1* protein sequence with other known sequences shows that *Pa-Wnt1* is the *Periplaneta* orthologue and aligns with closely related insect species.

A) *Periplaneta americana Wnt1*

```
1     GAAGGCCATTGAAATTGTTTAGAAAATGTACTGTAATGTTAGAATGTCTG    50
51    CGACTAACATACTTCAATATTTTTCCGTGTGAGTGCTTGTTTAAAGTAGA   100
101   TTTTTATGTACAATTAACTACATTACGTTGGGTGATATTTACTTTACCGT   150
151   ATAATTTATGTTGAAGTGAATTTGTACAAGAATGGAAACAAATGTAACGT   200
201   ACAATTATAAGTGATAAAAATATAACCAATGAAACAACGACTTAAGGGAA   250
251   AATATTGTACTGGAATTAAGGTAAAGAGAAAATAAAATATTTTTGGTCAG   300
301   CGATAGCTACAGAATGTGAAATAATACAACATAAAAGTCAGAAATTCCGA   350
351   CGCATCAAAATATATACAAGAAATCTAACAAGAAGAGAATCAAGGGCAAT   400
401   CAAATGAGAGAACAAATTTGCTGACTAGAATAACAAACAATAGGCCTATA   450
451   GAATGTGAAGATGACGATTACAGGAAATGTGAAACAATTTAAGAAAAATT   500
501   GTTCTAAAAATACAAGACGTAAGTTCTTATCACTTACGAATAAAAATAGA   550
551   AATTAAAACAGACATCGTTGTACGTCACATCGACATCGAATATTCGAGGG   600
601   AAGACTTCAGAATACTGGTTGTATAATACTATCCGAAGATCGCTGAGGAT   650
651   TAGCACAAAACTAGACATTGCGGTACAAAACTAAATGCAAGATATTGTGC   700
1                                       M  Q  D  I  V       5
701   GCTCAGGAATACCTGTGTGTTTGTTCCAGGGGATTGCGAAAGCCGGCGAG   750
6     R  S  G  I  P  V  C  L  F  Q  G  I  A  K  A  G  E    22
751   CCCAACAACTTGCTTCCGCAGACGCCGGGCGCGCTCTACATGGACCCGGC   800
23     P  N  N  L  L  P  Q  T  P  G  A  L  Y  M  D  P  A   39
801   CGTGCACGCCATTCTGCGGCGGAAGCAGAGACGTCTAGTTCGGGAGAACC   850
40      V  H  A  I  L  R  R  K  Q  R  R  L  V  R  E  N     55
851   CGGGAGTTCTTGTGGCGGTAGCCAAAGGTGCTAACCAGGCCATCGTGGAA   900
56    P  G  V  L  V  A  V  A  K  G  A  N  Q  A  I  V  E    72
901   TGCCAGTTCCAGTTTCGAAACAGGAGGTGGAACTGCTCGACAAGAAATTT   950
73     C  Q  F  Q  F  R  N  R  R  W  N  C  S  T  R  N  F   89
951   TCTACGAGGCAAAAACCTCTTCGGAAAAATTGTTGACAGAGGTTGTCGGG  1000
90      L  R  G  K  N  L  F  G  K  I  V  D  R  G  C  R    105
1001  AGACGGCGTTCATATACGCGATCACAAGTGCGGGCGTGACACACGCTATC  1050
106   E  T  A  F  I  Y  A  I  T  S  A  G  V  T  H  A  I   122
1051  GCGCGGGCGCGCAGCGAGGGCAGCATCGAGTCGTGCACGTGTGATTACAG  1100
123    A  R  A  R  S  E  G  S  I  E  S  C  T  C  D  Y  S  139
1101  CCACCAGGCGCGGGCGCCGCAGGTGACGTCCGTGCCCGGCCTGCGCGACT  1150
140     H  Q  A  R  A  P  Q  V  T  S  V  P  G  L  R  D    155
1151  GGGAGTGGGGCGGCTGCTCCGACAACATCGGCTACGGCTTCAAGTTCTCC  1200
156   W  E  W  G  G  C  S  D  N  I  G  Y  G  F  K  F  S   172
1201  CGCGAATTCGTCGATACCGGCGAGCGGGGgCGCAACCTCCGCGAGAAGAT  1250
```

APPENDIX 4

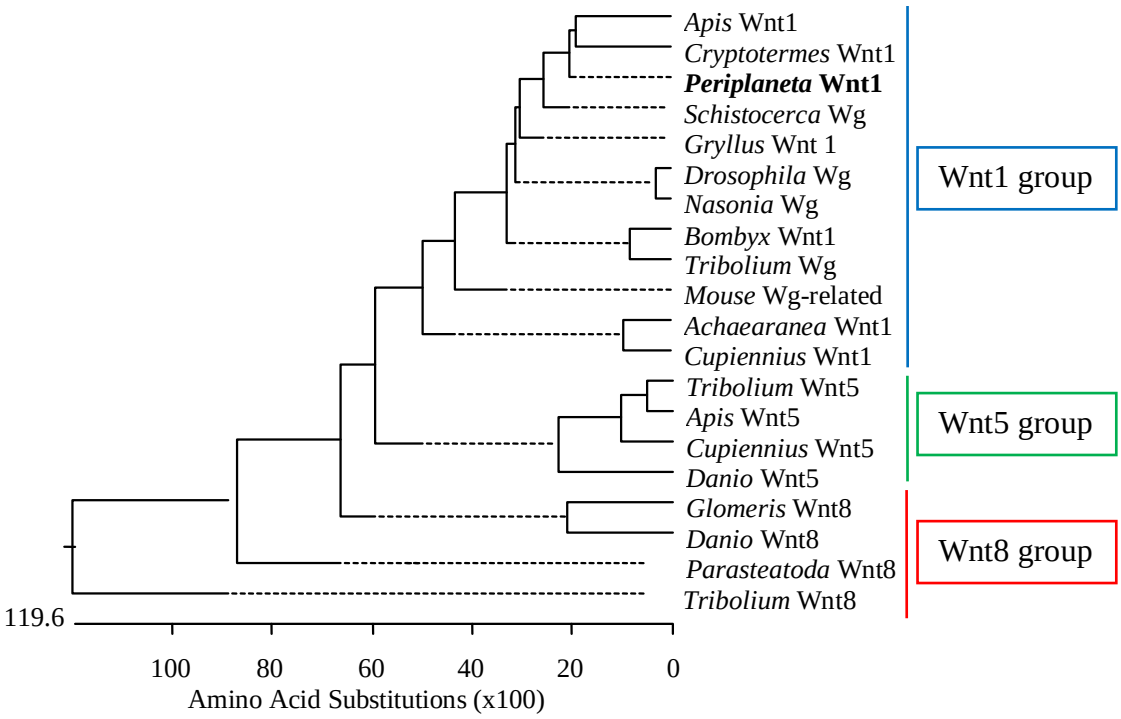
```
173   R  E  F  V  D  T  G  E  R  G  R  N  L  R  E  K  M   189
1251 GAATCTCCACAACAATGAGGCCGGCAGAGCGCACGTTTCCTCGGAGATGC  1300
190    N  L  H  N  N  E  A  G  R  A  H  V  S  S  E  M    205
1301 GTCAAGAATGTAAGTGCCACGGCATGTCTGGCTCCTGCACGGTCAAGACC  1350
206  R  Q  E  C  K  C  H  G  M  S  G  S  C  T  V  K  T   222
1351 TGCTGGATGCGGCTGCCCAGCTTCCGAGTCGTAGGCGACAACCTCAAGGA  1400
223   C  W  M  R  L  P  S  F  R  V  V  G  D  N  L  K  D   239
1401 CCGCTTCGACGGCGCCTCCAGAGTGATGGTGAGTAACGCGGGCAGCCTGC  1450
240    R  F  D  G  A  S  R  V  M  V  S  N  A  G  S  L    255
1451 GCGGCCAGGGTGGTAGCGGCGGCAGCGGCGTCGGTGGTAAGAAGAACAGA  1500
256  R  G  Q  G  G  S  G  G  S  G  V  G  G  K  K  N  R   272
1501 TACAACTTCCAACTGAAACCCTACAACCCGGACCACAAGCCGCCCGGCAC  1550
273   Y  N  F  Q  L  K  P  Y  N  P  D  H  K  P  P  G  T   289
1551 CAAAGACCTGGTCTACTTGGAGCCTTCCCCAGGGTTCTGCGAGCGCAACC  1600
290    K  D  L  V  Y  L  E  P  S  P  G  F  C  E  R  N    305
1601 CGAGACTCGGTATCCAAGGCACGCACGGACGTCAGTGCAACGATACGTCG  1650
306  P  R  L  G  I  Q  G  T  H  G  R  Q  C  N  D  T  S   322
1651 ATAGGCGTGGATGGTTGCGACCTCATGTGTTGTGGGCGAGGATATAGAAC  1700
323   I  G  V  D  G  C  D  L  M  C  C  G  R  G  Y  R  T   339
1701 TCATGAGGTGTCCGTGGTGCAGAGGTGTGCGTGCATGTTCCACTGGTGCT  1750
340    H  E  V  S  V  V  Q  R  C  A  C  M  F  H  W  C    355
1751 GCGAAGTCAAGTGCAACCTCTGTCGGACAAAGAAAACCATTCACACGTGT  1800
356  C  E  V  K  C  N  L  C  R  T  K  K  T  I  H  T  C   372
1801 CTGTGAGTGGTGAAAAAGAAACAATTCACCCATACTTGTGAGTGCTGCAA  1850
373   L  *                                               373
1851 AGAAAACCATCCACACGTGTCTGTGAGTGGTGAAAAAGAAACAATTCACC  1900
1901 CATACTAGTGAGTGCTGCGAAGAAAACCATCCACACGTGTCTGTGAGTGG  1950
1951 TGAAAAAGAAACCATTCACATTTGCTTGTGAGTGCTGCAAAGAAAAACTT  2000
2001 CCACACATATCTGTATGATCAGCATAGTGAAACTCGTTAATTGTCTGCGA  2050
2051 GTGCTTCAAAATCATTCATACATATCGATAAGTATTGGACAAAGAGAAAC  2100
2101 ATCACACTCGTTGTGATCGTTTCACAATGAAAACCATCCAGATGTGTTGT  2150
2151 GGGTGGTAGAAGAGTAACGTCACACTTGTATATGTTGCACATGTATGAGC  2200
2201 GCTGGACAAAGAGAACCACCCAAATGTGTTTACGTTGAGTACGTACTGCA  2250
2251 CAAACAAATACATTCACACGTCTTTATGAGCATACTGCAGAGATACGTAT  2300
2301 CTGTGGGTTACGCAAAAATAAAGTCACAACGCGTTTATGGTGAGTGTTAT  2350
2351 GAAAAAAAAAAAAA                                      2364
```

APPENDIX 4

B) Alignment and phylogenetic analysis of *Pa*-Wnt1 protein. NCBI Accession numbers: *Cryptotermes secundus* Wnt1 (ACN25142); *Gryllus bimaculatus* Wnt1 (BAB19660); *Schistocerca americana* Wg (AAD37798); *Achaearanea* (*Parasteatoda*) *tepidariorum* Wnt1 (BAD12586); *Apis mellifera* Wnt1 (XP_003251712); *Bombyx mori* Wnt1 (BAA03211); *Cupiennius salei* Wnt1 (CAC87040); *Drosophila melanogaster* Wg (AAF52501); *Mus musculus*, Mouse Wnt1/Wingless-related (EDL04166); *Nasonia vitripennis* Wg (XM_001603338); *Tribolium castaneum* Wg (EFA04660); *Cupiennius salei* Wnt5 (CAC87041); *Tribolium castaneum* Wnt5 (XP_974684); *Apis mellifera* Wnt5 (XP_397473); *Danio rerio* Wnt5 (AAA96519); *Glomeris marginata* Wnt8 (CBL52909); *Tribolium castaneum* Wnt8 (XP_971439); *Parasteatoda tepidariorum* Wnt8 (ACH88002); *Danio rerio* Wnt8 (AAI64176).

```
Majority            DNIGXGFKFSREFVDTGERGR--------DLREKMNLHNNEAGRAHVSSEMRQECKCHGMSGSCTVKTCWMRL
                    ---------+---------+---------+---------+---------+---------+---------+---
                             10        20        30        40        50        60        70
                    ---------+---------+---------+---------+---------+---------+---------+---
***Periplaneta* Wnt1**    **DNIGYGFKFSREFVDTGERGR--------NLREKMNLHNNEAGRAHVSSEMRQECKCHGMSGSCTVKTCWMRL  64**
*Cryptotermes* Wnt1 DNIGYGFKFSREFVDRASADQ--------PARED-DVHNNEAGRAHVSS----EC                    42
*Gryllus* Wnt1      DNIEYGFKFSRDFVDTGERGR--------TLREKMNLHNNEAGRLHVREEMRKECKCHGMSGSCTVKTCWMRL  64
*Schistocerca* Wg   -NIDYGFKFSREFVDTGERGR--------SLREKMNLHNNEAGRAHVVSEMRRECKCHGMSGSCTIRTCWMRL  63
*Achaearanea* Wnt1  DNIEFGSKFTKQFVGAAERGK--------DLRFTMNLHNNEAGRTHVAAGMRRQCKCHGMSGSCTVQTCWMQL  64
*Apis* Wnt          DNIGYGFKFSREFVDTGERGR--------NLREKMNLHNNEAGRAHVSSEMRQECKCHGMSGSCTVKTCWMRL  64
*Bombyx* Wnt1       DNIGFGFRFSREFVDTGERGK--------TLREKMNLHNNEAGRRHVQTEMKQECKCHGMSGSCTVKTCWMRL  64
*Cupiennius* Wnt1   DNIDFGAKFSRQFVDASERGK--------DLRYIMNLHNNEAGRAHVIGGMRRQCKCHGMSGSCTVQTCWMQL  64
*Drosophila* Wg     DNIGFGFKFSREFVDTGERGR--------NLREKMNLHNNEAGRAHVQAEMRQECKCHGMSGSCTVKTCWMRL  64
*Mouse* Wg-related  DNIDFGRLFGREFVDSGEKGR--------DLRFLMNLHNNEAGRTTVFSEMRQECKCHGMSGSCTVRTCWMRL  64
*Nasonia* Wg        DNIGYGFRFSREFVDTGERGR--------NLREKMNLHNNEAGRTHVSAEMRQECKCHGMSGSCTVKTCWMRL  64
*Tribolium* Wg      DNIGFGFTVSREFVDAGERGK--------TIREKMNLHNNEAGRWHVKDQMRQECKCHGMSGSCTIKTCWMRL  64
*Cupiennius* Wnt5   ---------------------------------MNLHNNEAGRRAVIRKIKVTCKCHGVSGSCSLVTCWQQL  39
*Tribolium* Wnt5    DNLEYGYKFTQNFVDVREKERKFKRGSKEQGRNLMNLHNNEAGRRAVIKKSKVTCKCHGVSGSCSLITCWQQL  73
*Apis* Wnt5         DNLEYGYKFTQAFVDVKERERSFKRGSREQGRSLMNLHNNEAGRRAVIKRSKVTCKCHGVSGSCSLITCWQQL  73
*Danio* Wnt5        DNVNYGYRFAREFVDAREREKNYPRGSVEHARTLMNLQNNEAGRMAVYNLANVACKCHGVSGSCSLKTCWLQL  73
*Glomeris* Wnt8     ENVHFGIDVTKDFLEARESGK--------DARALVNLHNNEAGRVAVSKTMRRLCKCHGVSGSCSIQTCWMQL  73
*Tribolium* Wnt8    DDSSFGEELVLKLLEDNEESS--------DAQAFINRHNNRIGREIIREKMLKTCKCHGVSGSCSFQTCWMQM  64
*Parasteatoda* Wnt8 DNVKIGNKMAKHYMDSKEHGR--------DIQAMINLHNNRVGRMMVKRNMRRMCKCHGVSGSCEMKTCWMRV  64
*Danio* Wnt8        DNVNFGDRIAKLFVDALENGH--------DSRAAVNLHNNEAGRLAVKATLKRTCKCHGLSGSCSIQTCWMQL  64
```

APPENDIX 4


Apis Wnt1
Cryptotermes Wnt1
Periplaneta Wnt1
Schistocerca Wg
Gryllus Wnt 1
Drosophila Wg
Nasonia Wg
Bombyx Wnt1
Tribolium Wg
Mouse Wg-related
Achaearanea Wnt1
Cupiennius Wnt1
Tribolium Wnt5
Apis Wnt5
Cupiennius Wnt5
Danio Wnt5
Glomeris Wnt8
Danio Wnt8
Parasteatoda Wnt8
Tribolium Wnt8
Wnt1 group
Wnt5 group
Wnt8 group
119.6
100
80
60
40
20
0
Amino Acid Substitutions (x100)

APPENDIX 5

Nucleotide and protein sequences for *Periplaneta even-skipped* and phylogenetic analysis. A) The 985bp nucleotide sequence of *Pa-eve* codes for a 295 amino acid protein containing the conserved regions: 60aa *eve*-like HD in double underscore, conserved N-terminal domain (dashed line), and Groucho interaction domain near the C-terminus (single underscore). B) Phylogenetic analysis of a conserved 77aa region of the *Pa-eve* HD protein sequence with other known sequences shows that *Pa-eve* is the *Periplaneta* orthologue and aligns with closely related insect species.

A) *Periplaneta even-skipped* nucleotide and protein sequences.

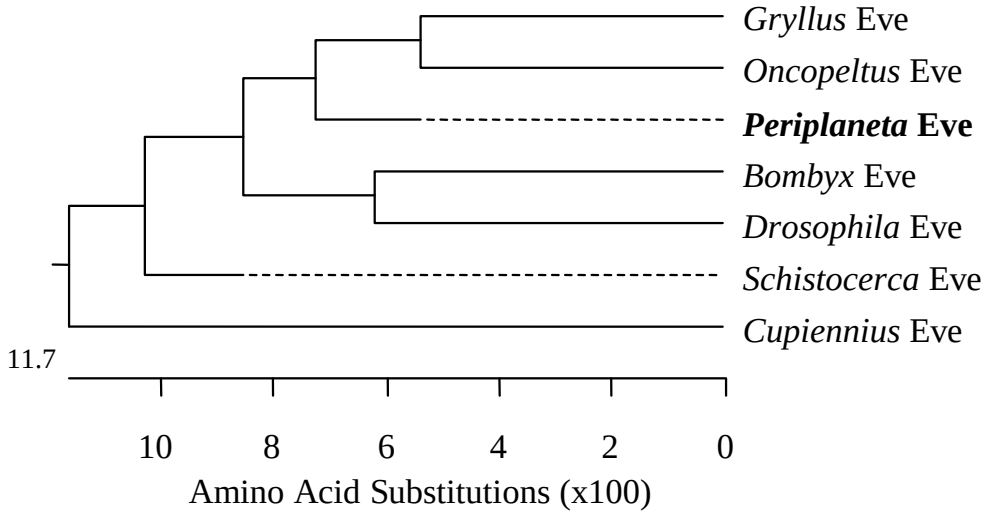
```
1    AATTAGCGCGCCATGCAGCAGTCGTACCACCAGGACAAGCTCCACCTCAA    50
1                M  Q  Q  S  Y  H  Q  D  K  L  H  L  K    13
51   GACCTCGGCCGCCACCGTCGTCGTGGATCTTCTGCCGCCGGCCTATACAC   100
14     T  S  A  A  T  V  V  V  D  L  L  P  P  A  Y  T     29
101  TCGGCACGCACCACCACGCCCCCCCGTCGCCGCCACAGCAGCCACCACAG   150
30   L  G  T  H  H  H  A  P  P  S  P  P  Q  Q  P  P  Q     46
151  CCCCCGTCACAGACGGGCAAGCAGCCCGACTCATCTTTGAAAGATGGCGC   200
47    P  P  S  Q  T  G  K  Q  P  D  S  S  L  K  D  G  A    63
201  CGGGACGTCCGCGGCCGCCGAGCAGAACATCCGGCGCTACCGAACGGCGT   250
64     G  T  S  A  A  A  E  Q  N  I  R  R  Y  R  T  A     79
251  TCACGCGGGAGCAGCTCGCCCGCCTCGAGAAGGAGTTCTACAAGGAGAAC   300
80   F  T  R  E  Q  L  A  R  L  E  K  E  F  Y  K  E  N    96
301  TACGTGTCCCGGCCCCGGCGCTGCGAGTTGGCGGCGCAACTCAACCTGCC   350
97    Y  V  S  R  P  R  R  C  E  L  A  A  Q  L  N  L  P   113
351  CGAGTCCACCATCAAGGTGTGGTTCCAGAACCGGCGCATGAAGGACAAGC   400
114    E  S  T  I  K  V  W  F  Q  N  R  R  M  K  D  K     129
401  GGCAGCGCATGGCGATGGCGTGGCCCTACGCCGTGTACACGGACCCCGCC   450
130  R  Q  R  M  A  M  A  W  P  Y  A  V  Y  T  D  P  A    146
451  TTCGCGGCCAGCATCCTGCAGGCGGCGGCGGCGAGCGCGGGCGGGCTACC   500
147   F  A  A  S  I  L  Q  A  A  A  A  S  A  G  G  L  P   163
501  CGGCATCGCGGCGGCAGCGGCGGCCGCGGCGTACGGCAGCCCTTACTCGT   550
164    G  I  A  A  A  A  A  A  A  A  A  Y  G  S  P  Y  S    179
551  ACTMCCACCCGCACGCACGCTACGGCCCCTACCCGCCGCCACTGCCGCCC   600
180  Y  X  H  P  H  A  R  Y  G  P  Y  P  P  P  L  P  P    196
601  CCGCCACACCGCCCGCAGCCCTACCTGCCGCCGCCCCCGCCCTTCGCCCT   650
197   P  P  H  R  P  Q  P  Y  L  P  P  P  P  P  F  A  L   213
651  CCGCACCGGCAGCCCCGCGGGCGCGGGCCACTCGCCCCGCTCGGAAAACA   700
214    R  T  G  S  P  A  G  A  G  H  S  P  R  S  E  N     229
701  GCACGCCCACGCTCAGCCCGCCCGCCACCAACAACAACGACGACAACTGC   750
230  S  T  P  T  L  S  P  P  A  T  N  N  N  D  D  N  C    246
751  GACGGCTCGCCCAGCTGTCGCTGCGGCATCGTCAACTGCGTCACGGCCTC   800
247   D  G  S  P  S  C  R  C  G  I  V  N  C  V  T  A  S   263
801  CACGCCCTCCCTGCTCATGACCACGGCGCTGAAGAGCCCGCACACGGAGC   850
264    T  P  S  L  L  M  T  T  A  L  K  S  P  H  T  E     279
```

APPENDIX 5

```
851  CGCCCAAAYTGTTCCAGCCCTACAAGACGGATGTGACGGAGCGCGCGTGA   900
280  P  P  K  L  F  Q  P  Y  K  T  D  V  T  E  R  A  *   295
901  GGGACAGAGACGGTCTGCGTTTTGTATTGTGTATAGCCTGTACATAGTTT   950
951  AGAATATAGTGCTGCCTACCACGAAAAAAAAAAAA                  985
```

B) Alignment and phylogenetic analysis of *Pa-eve* protein. NCBI Accession numbers: *Schistocerca americana* Eve (CAA77869); *Gryllus bimaculatus* Eve (BAD12840); *Bombyx mori* Eve (NP_001037327); *Cupiennius salei* Eve (CAB89492); *Drosophila melanogaster* Eve (AAF58865); *Oncopeltus fasciatus* Eve (AAW58076).

```
Majority          XDXSIRRYRTAFTREQLARLEKEFYKENYVSRPRRCELAAQLNLPESTIKVWFQNRRMKDKRQRMAMAWP
                  ---------+---------+---------+---------+---------+---------+---------+
                           10        20        30        40        50        60        70
                  ---------+---------+---------+---------+---------+---------+---------+
Periplaneta Eve   AEQNIRRYRTAFTREQLARLEKEFYKENYVSRPRRCELAAQLNLPESTIKVWFQNRRMKDKRQRMAMAWP  70
Schistocerca Eve  NDQSIRRYRTAFTREQLARLEKEFYKENYVSRPRRCELASQLNLPESTIKVWFQNRRMKDKRQRMAMAWP  70
Bombyx Eve        PDPNIRRYRTTFTREQLARLEKEFMKENYVSRPRRCELAAQLQLPESTIKVWFQNRRMKDKRQRIAVAWP  70
Cupiennius Eve    DLSSIRRYRTAFTREQLARLEKEFMRENYVSRPRRCELATALNLPESTIKVWFQNRRMKDKRQRMSLPWP  70
Drosophila Eve    ADPSVRRYRTAFTRDQLGRLEKEFYKENYVSRPRRCELAAQLNLPESTIKVWFQNRRMKDKRQRIAVAWP  70
Gryllus Eve       GAGSIRRYRTAFTREQLARLEKEFFKENYVSRPRRCELAAQLSLPESTIKVWFQNRRMKDKQQRMAMAWP  70
Oncopeltus Eve    NEVNIRRYRTAFTREQLTRLEKEFFKENYVSRPRRCELAAQLGLPESTIKVWFQNRRMKDKRQRMAMAWP  70
```



APPENDIX 6

Nucleotide and protein sequences for *Periplaneta runt* and phylogenetic analysis. A) The 1305 base pair nucleotide sequence of *Periplaneta run* codes for a 258 amino acid protein containing the highly conserved RUNT protein domain (double underscored). B) Phylogenetic analysis using 78aa of the RUNT domain aligned with other known sequences shows that *Pa-run* is the *Periplaneta* orthologue.

A) *Pa-run* cDNA and protein sequences.

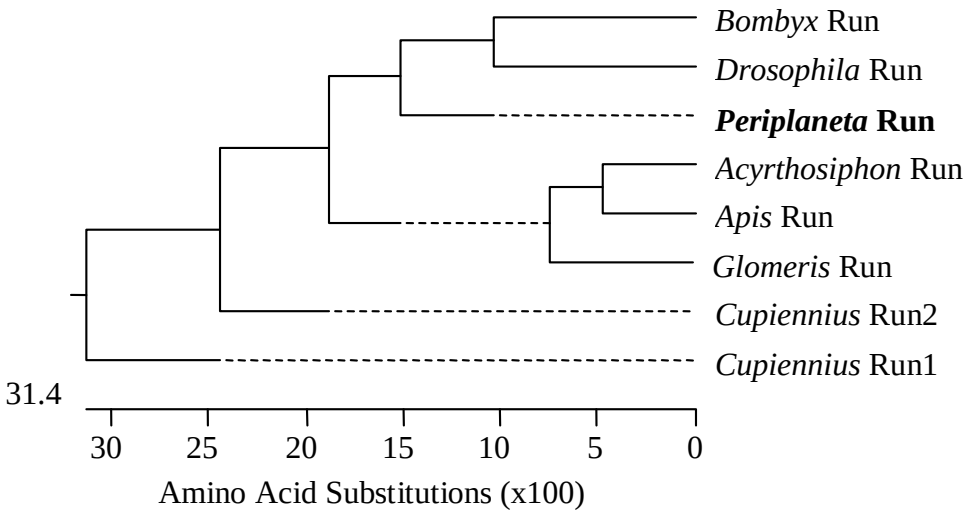
```
1     CAGTGCACGGGCTGCGCGCACGGTAGGGACATGTAGAACTATCAGCATTC    50
51    TTTCCAAAACACCCCCACACACACGACGCTGTGACTTAGCAGAGGAGGGG   100
101   CGAGCGTTTAGCGACAACGTGCCACAATTTCGATTCCGATTTTCAAAAAA   150
151   TAGAAGCGAGGGACTCGCACCCGACAGCCACCTGTTCGTCGCGACTCCAG   200
201   GCGGAAGAATACCACGGCGGCGACTCTTTTCCTGCCGATGAAGAAATGCT   250
251   AACGGCAGCTACAGTGTGTTCCAGAAGAAGAGAAGAGCAACCTCAAGACG   300
301   CTCTACGAAATTAGGAGCTCCTGCCTGTCGACCGGGAACCCTTCTTCAGC   350
351   GGCCACGGAGGAGCCCTCCTTCCTCAATGCATCTACCATCTCCCTACGAC   400
1                                 M  H  L  P  S  P  Y  D     8
401   TTGGACATCCACCACCAGCACCGCTCGAGTCTGTCACTCGTCCTCAACAG   450
9      L  D  I  H  H  Q  H  R  S  S  L  S  L  V  L  N  S    25
451   TAAGAGTAGTGTTAGTGCAAAGTCCGGGGACGGCCACGAATCCTCGCCGG   500
26       K  S  S  V  S  A  K  S  G  D  G  H  E  S  S  P     41
501   CGATGTCAGGCTCGGTGCCCGCGAGTTCGGGCGAGCCTTCCCCGGCAACG   550
42    A  M  S  G  S  V  P  A  S  S  G  E  P  S  P  A  T     58
551   ACGACCACGGTCCCTGGTACGGACTGGCTGCACGAGGCCTTGCAAGAATA   600
59     T  T  T  V  P  G  T  D  W  L  H  E  A  L  Q  E  Y    74
601   CCACGGGGAGCTGGTGCAGACCGGAAGTCCGGCGGTGCTATGCTCGGCGC   650
75       H  G  E  L  V  Q  T  G  S  P  A  V  L  C  S  A     91
651   TGCCTACCCACTGGCGCTCCAACAAGTCCCTGCCCGTCGCCTTCAAAGTG   700
92    L  P  T  H  W  R  S  N  K  S  L  P  V  A  F  K  V    108
701   GTCGCCCTGGACGATATCATGGACGGCACGCTAGTGACTGTGAAGGCCGG   750
109    V  A  L  D  D  I  M  D  G  T  L  V  T  V  K  A  G   124
751   CAACGATGAGAATTTCTGTGGTGAACTAAGGAACTGTACGGCCGTCATGA   800
125      N  D  E  N  F  C  G  E  L  R  N  C  T  A  V  M    141
801   AGAACCAAGTGGCCAAGTTCAACGACCTCAGGTTCGTCGGCAGAAGTGGC   850
142   K  N  Q  V  A  K  F  N  D  L  R  F  V  G  R  S  G   158
851   AGAGGGAAGAGCTTCTCGCTGACCATCGTCATCAGTTCGACGCCCTTTCA   900
159    R  G  K  S  F  S  L  T  I  V  I  S  S  T  P  F  Q   174
901   GATCGCCACGTACAACAAAGCCATCAAAGTGACGGTAGATGGACCCAGAG   950
175      I  A  T  Y  N  K  A  I  K  V  T  V  D  G  P  R    191
951   AACCTCGTACAAAATCTAACTTCCATTACCTCCCAGGCCAGCATCCCGGA  1000
192   E  P  R  T  K  S  N  F  H  Y  L  P  G  Q  H  P  G   208
1001  TTCGGTCCCTTCGCACTCTTTCCAGCGGCCCAGTGGCTGGACACTGCGGC  1050
209    F  G  P  F  A  L  F  P  A  A  Q  W  L  D  T  A  A   224
1051  GTATATGGGCTATCCGTGGCCCGAATACTTCAGGAGGCCTACGACTGCGG  1100
225      Y  M  G  Y  P  W  P  E  Y  F  R  R  P  T  T  A    241
1101  AACTCTGTAAGCTACCGAGCACCTGCGGTTCGTCTATTATAAAAAGTAAG  1150
242   E  L  C  K  L  P  S  T  C  G  S  S  I  I  K  S  K   258
```

APPENDIX 6

```
1151 TAGAGTTCTGACACTTCATCTGTGCGTTCCGAAACTGTTATTGTGTGAGG  1200
      *
1201 GAGAAGGCGTATGAAAGGCAGAAATGGGAGGCTCTGCGCCTGCGTATATC  1250
1251 AGCCCTCCTGCTACTCTACCGCATTTCCAAAAATATCAGTTGTTAAAAAA  1300
1301 AAAAA                                               1305
```

B) Alignment and phylogenetic analysis of *Pa-run* protein. NCBI Accession numbers: *Acyrthosiphon pisum* Run (XP_001950219); *Apis mellifera* Run (XP_394014); *Bombyx mori* Run (ABW84373); *Cupiennius salei* Run1 (CAB89493); *Cupiennius salei* Run2 (CAB89494); *Drosophila melanogaster* Run (AAC27786); *Glomeris marginata* Run (CAK50844)

```
Majority          SNKTLPVAFKVVALGXVPDGTLVTIXAGNDENXCAELRNCTAVMKNQVAKFNDLRFVGRSGRGKSFTLTITISTSPPQ
                  ---------+---------+---------+---------+---------+---------+---------+--------
                           10        20        30        40        50        60        70
                  ---------+---------+---------+---------+---------+---------+---------+--------
Periplaneta Run   SNKSLPVAFKVVALDDIMDGTLVTVKAGNDENFCGELRNCTAVMKNQVAKFNDLRFVGRSGRGKSFSLTIVISSTPFQ  78
Acyrthosiphon Run SNKTLPVAFKVVALGEVPDGTAVTIRAGNDENFCAELRNCTALMKNQVAKFNDLRFVGRSGRGKSFTLTITVSCSPPQ  78
Apis Run          SNKTLPVAFKVVALGEVGDGTLVTVRAGNDENCCAELRNSTALMKNQVAKFNDLRFVGRSGRGKSFTLTITVSTTPPQ  78
Bombyx Run        SNKSLPLAFKVVALDDVQDGTLVTIKAGNDENVMAELRNCTAVMKNQVAKFNDLRFVGRSGRGKSFTLTITISSFPSQ  78
Cupiennius Run1   SNKTLPLPFKVICLGEVADGTMVTIRAGNDENFCGELRNASAVMKNQVAKFNDLRFVGRSGRGKSFSLTISISTSPPH  78
Cupiennius Run2   SNKTLPVAFRVVSLGEVLDGTVVTIKAGNDDNYCAELRNATAVMKNQVAKFNDLRFVGRSGRGKSFSLTITLSTSPPQ  78
Drosophila Run    SNKSLPGAFKVIALDDVPDGTLVSIKCGNDENYCGELRNCTTTMKNQVAKFNDLRFVGRSGRGKSFTLTITIATYPVQ  78
Glomeris Run      SNKTLPVAFKVVALGDINDGTIVTIRAGNDENYCAELRNCTAIMKNQIAKFNDLRFVGRSGRGKSFNLTITVSTNPPQ  78
```



APPENDIX 7

Nucleotide and protein sequences for *Periplaneta pairberry* and phylogenetic analysis. A) The 1413 base pair nucleotide sequence of *Periplaneta pby* codes for a 398 amino acid protein. The conserved Paired domain and *paired*-like HD are indicated by single and double underscore, respectively, and the octapeptide indicated with a dashed line. B) Phylogenetic analysis of a conserved 75aa region of the Paired domain with other known sequences shows that this sequence is the *Periplaneta* orthologue of *pairberry*, as it contains the octapeptide, dashed line, and aligns more closely with those orthologues found in other insects and not found in *paired* proteins in this PaxII group of related genes.

A) *Pa-pby* cDNA and protein sequences.

```
1    GACTCATGGCGGGGCCTGCAGGATTGAACATTCACCCTCTCTTCACCAGA    50
1         M  A  G  P  A  G  L  N  I  H  P  L  F  T  R     15
51   TACTCGTTCCAAGGTCAGGGCAGAGTGAACCAGCTGGGCGGTGTGTTCAT   100
16    Y  S  F  Q  G  Q  G  R  V  N  Q  L  G  G  V  F  I    32
101  AAACGGACGTCCTCTGCCGAACCACATCCGCCTCAARATCGTtGaaATGG   150
33     N  G  R  P  L  P  N  H  I  R  L  K  I  V  E  M     48
151  CAGCCGCCGGAGTTCGACCATGYGTCATCTCGCGGCAGCTCCGGGTGTCG   200
49   A  A  A  G  V  R  P  C  V  I  S  R  Q  L  R  V  S    65
201  CACGGCTGCGTGTCCAAGATTCTGAACCGATACCAGGAGACCGGCAGCAT   250
66    H  G  C  V  S  K  I  L  N  R  Y  Q  E  T  G  S  I    82
251  CCGGCCAGGGGTGATCGGCGGCAGCAAGCCGAGGGTGGCCACCCCCGAAG   300
83     R  P  G  V  I  G  G  S  K  P  R  V  A  T  P  E     98
301  TGGAGCAGCGCATCGAGGACTACAAGAAGGCCAACCCGGGCATCTTCAGC   350
99   V  E  Q  R  I  E  D  Y  K  K  A  N  P  G  I  F  S    115
351  TGGGAGATTCGCGACCGGCTCATCAAGGAGGGCTTATGCGACAGCACCAA   400
116   W  E  I  R  D  R  L  I  K  E  G  L  C  D  S  T  N   132
401  CGCTCCCAGCGTGTCTTCAATCAGCCGGCTCCTCAGGGGCGGAAGAAGAG   450
133    A  P  S  V  S  S  I  S  R  L  L  R  G  G  R  R     148
451  ATGAGACGGACCTCAAAAAGGACTACAGCGTTGACGGGATTCTCGGAGGT   500
149  D  E  T  D  L  K  K  D  Y  S  V  D  G  I  L  G  G    165
501  CGTTGCGGGGACGAGTCGGACACGGAGTCAGAGCCGGGCATCCCGCTGAA   550
166   R  C  G  D  E  S  D  T  E  S  E  P  G  I  P  L  K   182
551  GCGCAAGCAACGCCGCAGCCGCACCACGTTCTCGGGGGACCAGCTGGAGG   600
183    R  K  Q  R  R  S  R  T  T  F  S  G  D  Q  L  E     198
601  AGCTGGAGCGGGCCTTCCAGAGGACGCAGTACCCCGACGTCTACACGCGC   650
199  E  L  E  R  A  F  Q  R  T  Q  Y  P  D  V  Y  T  R    215
651  GAGGAGCTCGCACAACGGACCGGCCTTACGGAGGCCAGGATACAGGTTTG   700
216   E  E  L  A  Q  R  T  G  L  T  E  A  R  I  Q  V  W   232
701  GTTCAGTAATCGACGTGCAAGATGGAGAAAACATACCGGTGGTACGTCGT   750
233    F  S  N  R  R  A  R  W  R  K  H  T  G  G  T  S     248
```

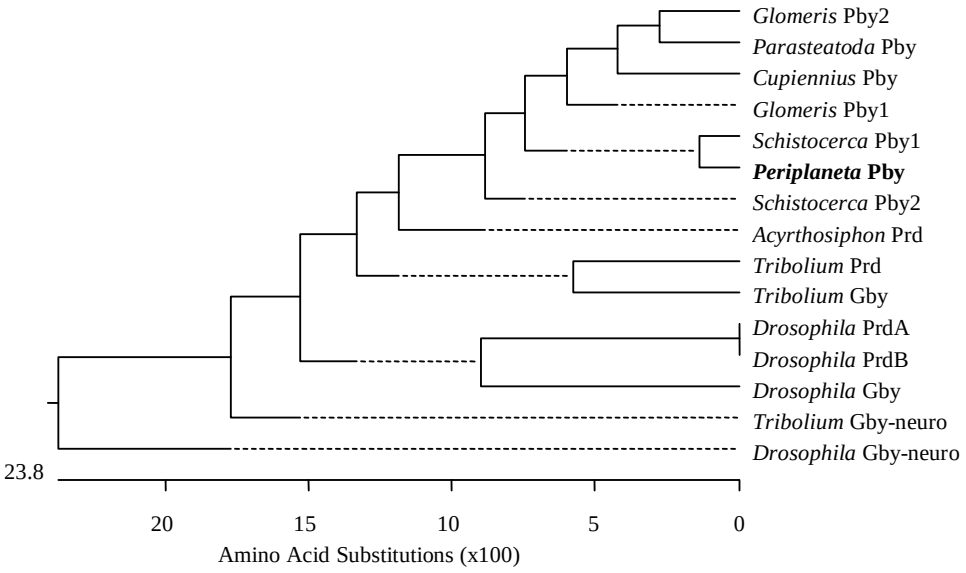
APPENDIX 7

```
751  TCAACCCATTATCTGCAGTTTCTGGTTACCAGTACCCGACAACGAGCTGT   800
249  F  N  P  L  S  A  V  S  G  Y  Q  Y  P  T  T  S  C   265
801  GAAGTCATGGCACTTCATCACAATGCAGGAAACTCCAATTGGCATCGCAC   850
266   E  V  M  A  L  H  H  N  A  G  N  S  N  W  H  R  T  282
851  CGGAAGTCAATTAGCCAACTACTCTGCCTTAATGCAACAATCACACGTCA   900
283    G  S  Q  L  A  N  Y  S  A  L  M  Q  Q  S  H  V   298
901  CTTCTGCTAGTCTGCAGCAATCGAATTTCGCATTATCGGCATCTCAAATG   950
299  T  S  A  S  L  Q  Q  S  N  F  A  L  S  A  S  Q  M   315
951  ATTGAACAAGTGACGACGCCATCTTCCGCAGCGGCTGCCACCACAACGTC  1000
316   I  E  Q  V  T  T  P  S  S  A  A  A  A  T  T  T  S  332
1001 ATCTTGTACAAACTCGATTCAGGGAAATGGCAATGTAATGGGATACACAG  1050
333    S  C  T  N  S  I  Q  G  N  G  N  V  M  G  Y  T   348
1051 TTCCATCTAGTGGAGTCACTGGAGACTACCAGCACTGCGATGCGGGCGCC  1100
349  V  P  S  S  G  V  T  G  D  Y  Q  H  C  D  A  G  A   365
1101 ACTGCAGTCTGGGGGGACAAGACTCAATGCCGACTCTAATTGGAGCCATCA  1150
366   T  A  V  W  G  T  R  L  N  A  D  S  N  W  S  H  H  382
1151 TGGTTTCCCAACAATTGCAGACAATTTCGGACGTGGCAAAAATAGGTCGT  1200
383    G  F  P  T  I  A  D  N  F  G  R  G  K  N  R  S   398
1201 AGAATTTAAAATATAAATTTGAGAAGGAACTCCAGAAGCTCCCTAAGATA  1250
      *
1251 CGAAGCGGTTCCGAAGCTCTTGAATCGCTACATCGTTATGGACGTTTTTC  1300
1301 GACCTGACGATGTTTATGAAGGACGTCGTCCTTTCTGATAGGACTGTAGG  1350
1351 AAATCTTTCCTTCCAGGAGATCCCATCGACTCCAAACAGTCCTGATTTGG  1400
1401 GAAAAAAAAAAAA                                      1413
```

APPENDIX 7

B) Alignment and phylogenetic analysis of *Pa-pby* protein. NCBI Accession numbers: *Acyrthosiphon pisum* Prd (XP_001949018); *Cupiennius salei* Pby (CAG30843); *Glomeris marginata* Pby1 (CBX36141); *Glomeris marginata* Pby2 (CBX36142); *Parasteatoda tepidariorum* Pby (CAX11345); *Schistocerca americana* Pby1 (AAK82936); *Schistocerca americana* Pby2 (AAK82937); *Drosophila melanogaster* PrdA (AAF53160); *Drosophila melanogaster* PrdB (AAN10801); *Drosophila melanogaster* Gby (AAF47315); *Drosophila melanogaster* Gby-neuro (AAF47314); *Tribolium castaneum* Prd (EFA05752); *Tribolium castaneum* Gby (EFA09269); *Tribolium castaneum* Gby-neuro (EFA09140).

```
Majority                AAGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEVEKRIEEYKKENPGIFSWEIRDRLIK
                        ---------+---------+---------+---------+---------+---------+---------+-----
                                 10        20        30        40        50        60        70
                        ---------+---------+---------+---------+---------+---------+---------+-----
Periplaneta Pby         AAGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEVEQRIEDYKKANPGIFSWEIRDRLIK 75
Acyrthosiphon Prd       AAGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPDVERRIEEYKTENPGIFSWEIRDRLIR 75
Cupiennius Pby          AAGVRPCVISRKLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEVEKKIEEYKQDSPGIFSWEIRDRLIN 75
Glomeris Pby1           AAGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEVEKKIDDYKKDNPGIFSWEIRDRLIK 75
Glomeris Pby2           AAVIRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEVEKKIEEYKRDNPGIFSWEIRDRLIK 75
Parasteatoda Pby        AAGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEVEKKIEEYKRDNPGIFSWEIKDRLVK 75
Schistocerca Pby1       AAGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEVEARIDDYKKANPGIFSWEIRDRLIK 75
Schistocerca Pby2       AAGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEIEARIEEYKKANPGIFSWEIRDRLIK 75
Drosophila PrdA         ADGIRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRIATPEIENRIEEYKRSSPGMFSWEIREKLIR 75
Drosophila PrdB         ADGIRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRIATPEIENRIEEYKRSSPGMFSWEIREKLIR 75
Drosophila Gby          AAGVRPCVISRQLRVSHGCVSKILNRFQETGSIRPGVIGGSKPRVATPDIESRIEELKQSQPGIFSWEIRAKLIE 75
Drosophila Gby-neuro    ASGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPKVTSPEIETRIDELRKENPSIFSWEIREKLIK 75
Tribolium Prd           ANGIRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEVENRIEQYKRENPSIFSWEIRDRLVK 75
Tribolium Gby           AAGIRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATLEVEARIEQLKKEEPQIFSWEIRDRLIK 75
Tribolium Gby-neuro     AAGVRPCVISRQLRVSHGCVSKILNRYQETGSIRPGVIGGSKPRVATPEIETRIEQMKKENPTIFSWEIRERLIK 75
```



APPENDIX 8

Nucleotide and protein sequences for *Periplaneta sloppy-paired* and phylogenetic analysis. A) The 700 base pair 5' fragment sequence of *Periplaneta slp* codes for a 278 amino acid protein fragment containing 91aa of the highly conserved Forkhead domain (double underscore) and the 15aa Conserved Domain II near the N-terminal end (dashed line). B) Phylogenetic analysis of a 69aa region of the forkhead domain other known sequences shows that *Pa-slp* is the *Periplaneta* orthologue and aligns with closely related insect species.

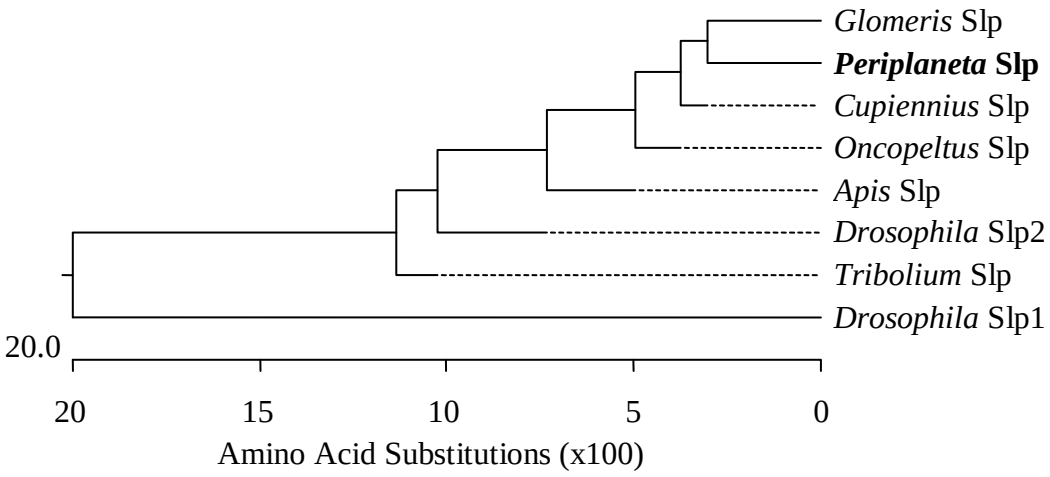
A) *Pa-slp* nucleotide and protein sequences.

```
1    AGTGAGACTCGGCCAGCGGGCAGTGACGCGTCAGTCTGTGTCGGAGCAGT    50
51   GCGGACAGTGCGAGATACCCGACCCGACGTTTCCTCAGGTGAAATTCTCG   100
101  TCTGCGAGAGGACTGCACCCCGCCCGCGCATGTCAACATGGTGAAGATCG   150
1                                        M  V  K  I      4
151  AGGGCCCGCTGCACCCCCTAGTGATGGCCCGAACGCCCCTCAAGTCCTCG   200
5    E  G  P  L  H  P  L  V  M  A  R  T  P  L  K  S  S    24
201  TTCAGCATCAGCTCCATCTTGCCGGAAACGGCGGCCGCGTCGCGCGCGCC   250
25    F  S  I  S  S  I  L  P  E  T  A  A  A  S  R  A  P    39
251  GAGCCCCCCCGACTTGGGGGCCCCGTGCTCCAGCGGCTCGGACAGCGACA   300
40     S  P  P  D  L  G  A  P  C  S  S  G  S  D  S  D     55
301  GCGACCTGGACGTGACGGGGGGCGCCACCCCCCCGCCGCTGGATTGCAGC   350
56   S  D  L  D  V  T  G  G  A  T  P  P  P  L  D  C  S    72
351  ACCAACAAGGACGGCAAGCAGGACGGCCCCGCCGCCGCCGCGGACAAGGC   400
73    T  N  K  D  G  K  Q  D  G  P  A  A  A  A  D  K  A   87
401  CGAGGGCGAGAAGAAGAAGAACGAGAAGCCGCCGTACAGCTACAACGCGC   450
88     E  G  E  K  K  K  N  E  K  P  P  Y  S  Y  N  A     203
451  TCATCATGATGGCCATCCGCCAGAGCCCCGAGAAGCGCCTCACGCTCAAC   500
204  L  I  M  M  A  I  R  Q  S  P  E  K  R  L  T  L  N   219
501  GGCATCTACGAGTTCATCATGAAGAACTTCCCCTACTACCGCGAGAACAA   550
220   G  I  Y  E  F  I  M  K  N  F  P  Y  Y  R  E  N  K   234
551  GCAGGGCTGGCAGAACTCCATCCGCCACAACCTCTCGCTCAACAAGTGCT   600
235    Q  G  W  Q  N  S  I  R  H  N  L  S  L  N  K  C     250
601  TCGTGAAGGTGCCGCGCCACTACGACGACCCGGGCAAGGGCAACTACTGG   650
251  F  V  K  V  P  R  H  Y  D  D  P  G  K  G  N  Y  W   266
651  ATGCTGGACGCCAGCAGCGAGGACGTCTTTATCGGCGGCACGACGGGAAA   700
267   M  L  D  A  S  S  E  D  V  F  I  G  G  T  T  G  K   278
```

APPENDIX 8

B) Alignment and phylogenetic analysis of *Pa-slp* protein. NCBI Accession numbers: *Glomeris marginata* Slp (CBX36143); *Cupiennius salei* Slp (CAI91293); *Oncopeltus fasciatus* Slp (ACZ94039); *Apis mellifera* Slp (XP_003251005); *Tribolium castaneum* Slp (ABD63010); *Drosophila melanogaster* Slp2 (AAF51057); *Drosophila melanogaster* Slp1 (AAF51058).

```
Majority          RQSPEKRLTLNGIYEYIMXNFPYYRENKQGWQNSIRHNLSLNKCFVKVPRHYDDPGKGNYWMLDPSSED
                  ---------+---------+---------+---------+---------+---------+---------
                           10        20        30        40        50        60
                  ---------+---------+---------+---------+---------+---------+---------
Periplaneta Slp   RQSPEKRLTLNGIYEFIMKNFPYYRENKQGWQNSIRHNLSLNKCFVKVPRHYDDPGKGNYWMLDASSED   69
Glomeris Slp      RQSPEKRLTLSGIYEFIMRNFPYYRENKQGWQNSIRHNLSLNKCFVKVPRHYDDPGKGNYWMLDPSSDD   69
Cupiennius Slp    RQSPEKRLTLNGIYEYIMKNFPYYRENKQGWQNSIRHNLSLNKCFVKVPRHYDDPGKGNYWMLDPSSDD   69
Oncopeltus Slp    RQSPEKRLTLNGIYEYIMKNFPYYRDNKQGWQNSIRHNLSLNKCFVKVPRHYDDPGKGNYWMLDPSSED   69
Apis Slp          RQSPEKRLTLNGIYEYIMRHFPYYENNKQGWQNSIRHNLSLNKCFVKVPRHYDDPGKGNYWMLDPSSED   69
Tribolium Slp     RNSPEKRLTLNGIYEYIMRNFPYYRENKQGWQNSIRHNLSLNKCFVKVPRHYDDPGKGNYWMLDPSAED   69
Drosophila Slp1   QDSPEQRLTLNGIYQYLINRFPYFKANKRGWQNSIRHNLSLNKCFTKIPRSYDDPGKGNYWILDPSAEE   69
Drosophila Slp2   RQSSEKRLTLNGIYEYIMTNHPYYRDNKQGWQNSIRHNLSLNKCFVKVPRHYDDPGKGNYWMLDPSAED   69
```



APPENDIX 9

Nucleotide and protein sequences for *Periplaneta tarsal-less*. The 569bp sequence of *Pa-tal* contains two small Open Reading Frames that each encode a 12 amino-acid-long peptide containing the highly conserved heptapeptide sequence, LDPTGXY (underscored).

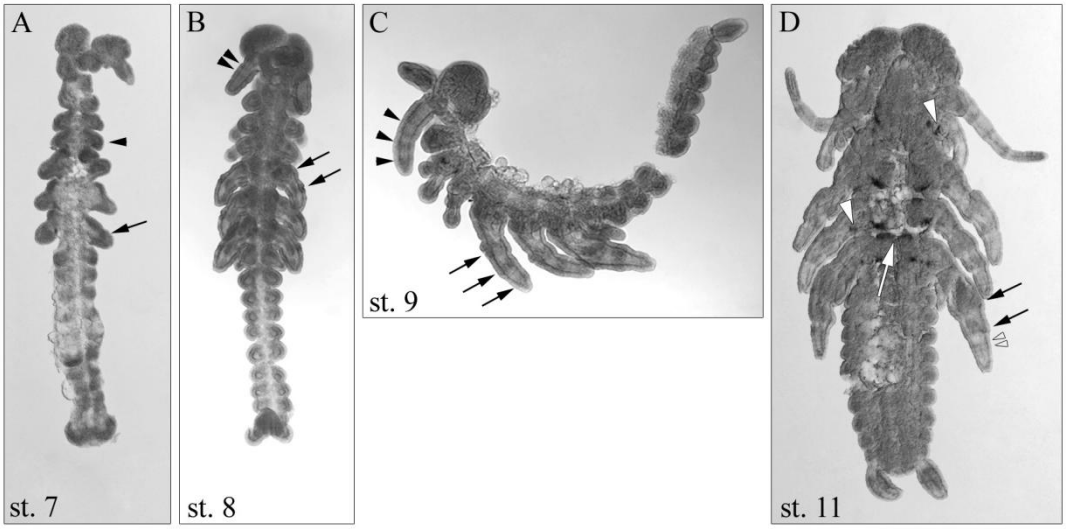
```
1     AAGTTGCGGGCTAGCCCTCGAACGGAGAGGTTCTCTGCTTGCCTGGTGCT    50
51    CGGTTCAATACCGTTGTGACTTGCGTGTTAAGTTCATTGCAGTTCACAAG   100
101   ACAGGAGGTGACGTTCAACCCAGTCTGCAACGACKGTTACTCAACTCAGG   150
151   TATACCATCACTGCCAAGTACCTCGCACGTTGCTGCCCGTTTTTCACTTG   200
201   TTACCTCAGCCTTCGACATGGATCCCAAGACTTTGGATCCCACCGGTCTG   250
1                       M  D  P  K  T  L  D  P  T  G  L     11
251   TACTAGTCAGACGTCTTCCTGTTCCTGTTGCGGGGAAAGAAAATCGGTGA   300
12     Y  *                                                12
301   CATCGACCTTGTGATTCCGTACTTAGTCACGAGAAGAAAAACAAGAAAGA   350
351   TGGATCCTACACATCTAGACCCGACCGGCCTGTACTGAGAGCGAGACTRC   400
1     M  D  P  T  H  L  D  P  T  G  L  Y  *                12
401   CGTCRAGACRTCGGCGATGAAGGTATGTTCGCTTTCTTCTGCACGTGACG   450
451   TTTGAAGTTGAGAGTTCCCTCACACCAGGAAAGGGAACCCTTCAGTAATA   500
501   TATTCCTGGGAGATGATTCAAACCGGCAAGAAAAAGACGCAAGGCATACG   550
551   AAGAAAGAAAAAAAAAAAA                                 569
```

APPENDIX 10

Nucleotide and protein sequences for *Gryllus tarsal-less*. The 706bp sequence of *Gb-tal* contains two small Open Reading Frames that each encode a 12 amino-acid-long peptide containing the highly conserved heptapeptide sequence, LDPTGXY (underscored).

```
1     AGTCGAGTCCCAGGAGCGAGCGAGGCCGTGCACCAGACGCGCAGCAGCTC    50
51    TCGCCACGCGCGGTCGCTCGCCCTTTGCTTCGCCAGCCGTTCACCGCCCG   100
101   CTCGCCCGCCCGTGCGCGCGCGTATAAACCGACGCAGCCGCTCCGTGCGC   150
151   GTTCAGTTGCCCGCTGGCTGCCTTCCGCTTCGGGTCGGTGCTCTCTCTCG   200
201   CTTCTACGCTCGCTCGTGTTGGACGTGCGACGTCGACCGGACGGGTGTGA   250
251   CGCCTGTGATCGTGAACTCTACTGTTTTCTACGTTCCGCCACTTCAACTC   300
301   GCAGCCATGGGCCCCAAGACTCTGGACCCCACCGGCCTGTACTGAAGGAG   350
1           M  G  P  K  T  L  D  P  T  G  L  Y  *          12
351   AGAGAGAGAGGTCGCGGTGACAGTCGCCAGCAGTCGAGCCTTGCAACGGG   400
401   CCCCGCCTGCCTCAGCAGCAGCCACGCCGCTGTTACTTCCTGGTCGAGAG   450
451   ACGGCAGCCTTCGTGGACAGTGCCAGCACAACCGCAGCCATGGACCCCAA   500
1                                              M  D  P  K     4
501   GCAGCTGGACCCCACCGGTCTGTACTAAGGCCCCCGGGCGTGCCTTACCG   550
5       Q  L  D  P  T  G  L  Y  *                          12
551   GCGGCAGCAGCAGCAACAGGCTTCACCAGAGCAGCCACCCTGCTATGCAC   600
601   CCGTGAGCCATGGGACGAGCAGCATCATCATCAGCCCGCGCTTGGTAAGC   650
651   GGCACATATTTATTTAAAAAACGAAGATATTTTTTATATAGAACAAAAAA   700
701   AAAAAA                                              706
```

Appendix 11



Expression of the *Gryllus bimaculatus* homologue of *tarsal-less*. The expression of the *Gryllus-tal* homologue was analysed using a 295bp riboprobe that included both smORF sequences. (A) In the stage 7 germ band elongation embryo (according to Niwa et al., 1997), *Gb-tal* is diffusely expressed at the tips of the extending gnathal and thoracic limb buds (arrows). (B) By stage 8, *Gb-tal* expression is evident as stripes in the antennae (arrowheads) and as two broad bands in the proximal and medial regions of the extending leg bud (arrows). (C) As legs continue to grow and constrictions become evident, stage 9, *Gb-tal* expression is refined into stripes in the distal end of all developing leg segments (arrows) and, by stage 11 (D), can be detected in the developing tarsomeres (open arrowheads). (D) In addition, *Gb-tal* is expressed in the apodemes of the mouthparts and legs (white arrowheads), as well as in the developing CNS (white arrow). Altogether, these preliminary studies in *Gryllus* show a conserved expression in all developing legs segments compared to *Periplaneta* and *Tribolium* (Savard et al., 2006).

Appendix 12

List of *tarsal-less* homologues in arthropod species. (A-C) An expanded list of that by Galindo et al. (2007) showing the putative peptide translations of *tarsal-less* homologues in numerous insect species and in one representative of the crustaceans (*Daphnia pulex*). This list was used to create the phylogenetic tree shown in Figure 8. (A) Numerous copies of Type-A *tal* are found in most insect orders and only a single copy in the crustacean. The presumed peptides range from 9 to 26 amino acids in length. With few exceptions, all peptides contain the conserved LDPTGXY motif (highlighted in red). (B) Type-AA *tal* peptides are only found in the Diptera, are longer then Type-A, and contain the conserved LDPTGXY motif at the 5' and 3' ends of the peptide (in red). (C) The Type-B *tal* peptides are only found in the holometabolous insects and do not contain the conserved motif of the Type-A and Type-AA *tal* peptides, but have a different motif, loosely ETSSCRRRR (highlighted in blue), conserved between them.

Appendix 12

A) Type-A *tal* peptides

DIPTERA					
Drosophila melanogaster		MAAYLDPTGQY	MAAYLDPTGQY	MSHDLDPTGTY	
Drosophila mojavensis		MAAYLDPTGQY	MAAYLDPTGQY	MSLALDPTGTY	
Anopheles gambiae		MEILDPTGYY	MAKKLDPTGHY	MARKLDPTGHY	MAPEILDPTGYY
Aedes aegypti		MEILDPTGYY	MEKKLDPTGHY	MAFKLDPTGHY	MALEILDPTGYY
Lutzomyia longipalpis	(1)	MTGILDPTEVY	MATELDPTGHY		
	(2)	MASTLDPTGHY	MERSLDPTGMY		
Teleopsis dalmanni		MANYLDPTGQY	MDFALDPTGTY		
Cochliomyia hominivorax		MDFALDPTGTY			
Polypedilum vanderplanki		MALKKLDPTGSY	MESSASRRLDPTGHY		
LEPIDOPTERA					
Heliconius herato		MGKVLDPTGIY	MLGLDPTGVY	MFSLDPTGVY	MFVLDPTGVY
		MFVLDPTGVY	MGLDPTNVY	MAIYLDPTEVYYNS	
Helcoverpa armigera		MDIKVLDPTGIY	MFGLDPTGVY	MFVLDPTGVY	MFGLDPTGVY
		MVVLDPTNVY	MALGLDPTGVY		
Bombyx mori	(1)	MDIVTLDPTGLY	MELTLDPTGQY	MLKVLDPTGQY	MTGLDPTEVY
	(2)	MDVKVLDPTGIY	MLGLDPTGVY	MYVLDPTGVY	MLVLDPTGVY
		MGLDPTGVY	MFILDPTNVY	MFRGLDPTGVY	
Bicyclus anynana		MFVLDPTGVY	MFVLDPTGVY	MDLDPTNVY	MAISLDPTGVY
Striacosta albicosta		MFVLDPTGVY	MFVLDPTGVY	MFILDPTGV.	
Samia cynthia		MNIATLDPTGLY	MDSTLDPTGQY	MGKVLDPTGQY	MIGLDPTEIY
Choristoneura fumiferana		MDIKVLDPTGIY	MFVLDPTGVY	MFVLDPTGVY	MFVLDPTGVY
		MFVLDPTGVY	MIVLDPTNVY		
Heliothis virescens		MDILTLDPTGLY	MDSTLDPTGQY	MDTKVLDPTGQY	MIGLDPTEIY
		MFGLDPTGVY	MFVLDPTGVY	MFGLDPTGVY	MFGLDPTGVY
		MALGLDPTGVY			
COLEOPTERA					
Tribolium castaneum		MSGLDPTGLY	MDGGKLDPTGQY	MKLNGGKSLDPTGLY	
Diaprepes abbreviates		..TGQY	MAKIGGKGLDPTGLY		
Diabrotica virgifera		MFGGLDPTGLY	MDGDKLDPTGQY	MQKTGGKGLDPTGLY	
HYMENOPTERA					
Apis mellifera		MARQLDPTGQY	MATGLDPTGLY	MAGLDPTGQY	MAASTGLDPTGQY
Bombus terrestris		MARQLDPTGQY	MAAGLDPTGLY	MAGLDPTGQY	MAASTGLDPTGQY
Nasonia vitripennis		MAVQLDPTGVY	MAFGLDPTELY	MEILDPTNQY	MAAILDPTNQY
HEMIPTERA					
Homalodisca vitripennis		MFPPTLDPTGLY	MTSDSILDPTGLY		
Maconellicoccus hirsutus		MDLDPTGLY	MQTDSKSQLKRDPKPGISGLDPTEVY		
Acyrthosiphon pisum		MGLDPTGLY	MCGGGPKLDPTGLY		
Aphis gossypii		MGLDPTGLY	MCGGGPKLDPTGLY		
Nilaparvata lugens		MGPKTLDPTGLY	MTATSSARTAPSVEQLDPTGLY		
Rhopalosiphum padi		MGLDPTGLY			
Pemphigus spyrothecae		MGLDPTGLY			
BLATTODEA					
Periplaneta americana		MDPKTLDPTGLY	MDPTHLDPTGLY		
ORTHOPTERA					
Gryllus bimaculatus		MGPKTLDPTGLY	MDPKQLDPTGLY		
Locusta migratoria		MGPKTLDPTGLY	MDTPKQLDPTGLY		
CLADOCERA					
Daphnia pulex		MSPKLDPTGLY			

Appendix 12

B) Type-AA *tal* peptides (Diptera only)

Drosophila melanogaster MLDPTGTYRRPRDTQDSRQKRRQDC LDPTGQY
Drosophila mojavensis MFDLDPTGTYRRPRESRDTRHKQRQT LDPTGQY
Anopheles gambiae MNRKLDPTGMYRRPGSGASSSQRSPFHHHHQQLQNHQRMPHHHQQQQQHQVKCHYLDPTGLY
Lutzomyia longipalpis MTSTDDKLDPTGMYVRPKIEIECH LDPTEYY
Cochliomyia hominivorax MACTKLILDPTGTYLRPTSSTTTSNASSDRRLANSF LDPTGQY

C) Type-B *tal* peptides (Holometabola only)

DIPTERA
Aedes aegypti MGQRRNFWLTVRGREETSSCRRRRKLPIRAVGTRWNP
Anopheles gambiae MALRWWTAPQARWVRSREETSSSRRKRKFPAPGTRQRWHRLAATVAERMNPRSGRSRMNLGVIVAYY
Cochliomyia hominivorax MLGINKLLKLFEPLWLEVRGREETSSCRKRRKIKMFLKTNLIFFTQII
Drosophila melanogaster MIGGARWLRVRGREETSSCRRRRKLGIGASPSDLGEPCDGDFCIYVVFA
Lutzomyia longipalpis MDPCDRGAVKDLAIDTTIKAWRSVPGREETSSCRRRRKSPKMCAPTLSPATNILSSSLNK

LEPIDOPTERA
Bicyclus anynana MPGCTHRHKLISATRNEVSSNRKKRKVFHYT
Bombyx mori MCCGRRRICTGSSPAETSSCRKRRFHTDFPLC
Heliothis virescens MCCGRRRIFTGTSPRETSSCRKRRFSSEHH
Samia cynthia MCCGRRRIFTGCSPAETSSCRKRRFAEEFLIKTD

COLEOPTERA
Tribolium castaneum MWHRNRGDGGRPETSSGRRRRLR

HYMENOPTERA
Nasonia vitripennis MTGGEGEKSAVVVRGREETSSTNLRRSVWSINIKRGRHRKLPEDEDEAPAAR
Bombus terrestris MEHVRRREETSTSYLRRPEWSIHIRKGRHRPPFSLEETHQ