

**Investigating the *Hes/Hey* and *odd* family genes
for roles in sequential segment patterning in the
red flour beetle, *Tribolium castaneum***

Jacob David Leese

Submitted in accordance with the requirements for the degree
of Doctor of Philosophy

The University of Leeds

School of Biology

May 2023

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Jacob David Leese to be identified as Author of this work has been asserted by Jacob David Leese in accordance with the Copyright, Designs and Patents Act 1988.

ACKNOWLEDGEMENTS

I would firstly like to thank my supervisor, Andrew Peel, for his expertise and guidance whenever I sought it, and for helping me get started with many of the protocols and experiments that went into this thesis.

I would also like to thank Rahul Sharma for teaching me how to perform the (often painstaking) task of embryonic injections during the first few months of my PhD, and to my fellow PhD lab member, Bradleigh Cocker, who has kept me company in our quiet corner of the lab throughout.

Thank you to the University of Leeds and the BBSRC White Rose DTP for providing me with the necessary facilities and resources to accomplish my PhD. Thanks also to the FBS Bioimaging facility and Ruth Hughes for helping me on the confocal microscope.

I am grateful for all those I've shared the wider lab and office spaces with over the years who have helped to maintain a relaxed and friendly environment. A special thanks to Britt and Rosie for taking me under their wing during my first year and for introducing me to the lunch time gang, through which I made some of my strongest friendships in Leeds. In more recent times, I am thankful to Kane for his company in the office and for making an excellent conference partner on our trip to Naples.

My time in Leeds of course would not have been the same without all the amazing friends I made along the way, including, but not limited to, Chris, Harri, Rosie, Katie, Britt, Georgia and Matt. Thank you for all the memorable Yorkshire adventures and long in-depth discussions about every topic imaginable.

Finally, I would like to thank my family for all their continuous love and support, and for providing me with the perfect countryside base for writing up my thesis. Last but not least, thank you to my partner, Perla, for being there every evening and weekend throughout these last few months. Even if not always in person, our daily chats have helped to keep me sane and switch off after long hours at the computer. Thank you for your love and support, and for putting up with my occasional rambling.

ABSTRACT

Body segmentation is a defining characteristic of arthropods, however, the mechanism by which segments are patterned during development has diverged greatly during evolution. The ancestral mode of patterning, in which segments are added sequentially to the posterior of the embryo during axial elongation, relies on a clock and wavefront mechanism that remains poorly understood and variable in nature between different species. Using the sequentially segmented beetle, *Tribolium castaneum*, I attempt to address some of the underlying complexities of this system and improve our understanding of the evolutionary processes responsible for this diversity in patterning mechanisms across arthropods. Driven by recent data from our lab that question the existing *Tribolium* segmentation clock circuit, I investigate candidates from the *Hes/Hey* and *odd*-related gene families for possible roles in primary segment patterning.

Phylogenetic analyses demonstrate a strong potential for redundancy and developmental system drift within both gene families and provide evidence for concerted evolution among the arthropod *odd*-family genes. I offer a revised understanding of *Tribolium* *odd*-family gene expression patterns which, while inconsistent with roles in primary segment patterning, suggest *Tc-sister-of-bowel* and *Tc-drumstick* may be involved in the late segmentation network. A screen of *Tribolium* *Hes/Hey* gene expression identifies a novel segmentation clock candidate, *Tc-extramacrochaetae* (*Tc-emc*), while multiplex *in situ* hybridisations provide evidence linking oscillations of *Tc-hairy*, *Tc-deadpan* and *Tc-emc* to genes from the primary clock circuit (*Tc-even-skipped*, *Tc-odd-skipped*, *Tc-runt*). Although segmentation roles for *Tc-deadpan* were not identified in this study, knockdown of *Tc-hairy* in *Tc-odd*-null mutants indicates some redundancy in thoracic segment patterning, while *Tc-emc* knockdowns generate phenotypes consistent with a key role in abdominal segment patterning. As these genes are commonly associated with Notch signalling, these findings renew the discussion of Notch function in *Tribolium* segmentation while also highlighting changes to the network that occur across different stages of segment addition.

TABLE OF CONTENTS

Declarations.....	i
Acknowledgements.....	ii
Abstract.....	iii
Table of Contents.....	iv
List of Figures.....	viii
List of Tables.....	x
List of Abbreviations.....	xi
 Chapter 1: General Introduction.....	 1
1.1. Animal segmentation as a case study for Evo-Devo.....	1
1.2. Segmentation in arthropods.....	3
1.3. The <i>Drosophila</i> model and the history of segmentation gene nomenclature.....	6
1.4. Sequential segmentation.....	10
1.5. The clock and wavefront model.....	12
1.6. <i>Tribolium</i> as a model for sequential segmentation research.....	19
1.7. The <i>Hes/Hey</i> family genes: Candidates for secondary oscillations in <i>Tribolium</i> segmentation.....	25
1.8. <i>Odd-skipped</i> family genes: Candidates for partial redundancy with <i>Tc-odd</i>	27
1.9. Aims of this thesis.....	29
 Chapter 2. Materials and Methods.....	 30
2.1. Phylogenetics analysis.....	30
2.2. <i>Tribolium</i> husbandry and embryo fixation.....	31
2.3. Gene fragment amplification and cloning.....	32
2.4. <i>in situ</i> hybridisation probe synthesis.....	34
2.5. Double-stranded RNA synthesis.....	34
2.6. Colorimetric <i>in situ</i> hybridisation.....	35
2.7. Hybridisation Chain Reaction (HCR) multiplex <i>in situ</i> hybridisation.....	35
2.8. Parental and embryonic RNAi.....	37
2.9. Cuticle preparation and imaging.....	38
 Chapter 3. Phylogenetic Analyses of <i>Hes/Hey</i> and <i>odd</i> family genes.....	 39
3.1. Introduction.....	39

3.1.1. The evolution of genes and multigene families.....	39
3.1.2. Concerted evolution of gene families.....	40
3.1.3. Current understanding of <i>Hes/Hey</i> gene family evolution.....	42
3.1.4. Current understanding of <i>odd-skipped</i> family evolution.....	43
3.1.5. Aims of the chapter.....	44
3.2. Results.....	45
3.2.1. Phylogenetic analysis of the arthropod <i>Hes/Hey</i> family genes.....	45
3.2.2. Arthropod <i>hairy/deadpan</i> genes exhibit a conserved intron/exon structure.....	52
3.2.3. Phylogenetic analysis of arthropod <i>odd</i> family genes.....	55
3.2.4. Conserved genomic arrangement of arthropod <i>odd</i> family genes.....	63
3.2.5. <i>odd</i> family gene structure consistent with phylogenetics results.....	66
3.3. Discussion.....	70
3.3.1. Arthropod <i>Hes/Hey</i> family evolution overview.....	70
3.3.2. Placing suspected segmentation genes against the consensus <i>Hes</i> gene tree.....	70
3.3.3. Conserved intron structure of <i>Hes</i> genes: A link to transcription cycle time constraints?.....	72
3.3.4. Arthropod <i>odd</i> family evolution overview.....	73
3.3.5. Co-localisation of arthropod <i>odd</i> family genes may reflect shared regulatory inputs.....	73
3.3.6. Context for subsequent chapters.....	75
Chapter 4. A revised understanding of <i>odd</i> family gene expression in <i>Tribolium</i>.....	76
4.1. Introduction.....	76
4.1.1. <i>odd</i> family expression and conserved function in arthropods.....	76
4.1.2. Arthropod <i>odd</i> family genes are regularly expressed in patterns consistent with roles in segmentation.....	77
4.1.3. HCR as an effective method for evaluating gene co-expression.....	78
4.1.4. Aims of the chapter.....	79
4.2. Results.....	80
4.2.1. Staging with <i>Tc-wingless</i> (<i>Tc-wg</i>) as a marker.....	80
4.2.2. A revised understanding of <i>odd</i> family gene expression in <i>Tribolium</i>	80
4.2.3. <i>odd</i> family gene expression shows no signs of divergence between SB and GA2 strains.....	85
4.3. Discussion.....	87

4.3.1. Previously published <i>Tc-bwl</i> and <i>Tc-sob</i> expression is consistent with cross-reaction of <i>in situ</i> probes with transcripts of other <i>odd</i> family genes.....	87
4.3.2. <i>in situ</i> probes used in <i>odd</i> family expression studies of other arthropods commonly include regions with high sequence similarity.....	91
4.3.3. Revised expression patterns do not support redundancy between <i>odd</i> family genes in the <i>Tribolium</i> segmentation clock network.....	94
Chapter 5. Expression dynamics of <i>Hes/Hey</i> genes in <i>Tribolium</i>.....	96
5.1. Introduction.....	96
5.1.1. <i>Hes/Hey</i> gene family expression and function in arthropods.....	96
5.1.2. Aims of the chapter.....	99
5.2. Results.....	100
5.2.1. <i>Hes/Hey</i> family gene expression screen.....	100
5.2.2. Co-expression analysis of striped <i>Hes</i> genes: <i>Tc-h</i> , <i>Tc-dpn</i> and <i>Tc-emc</i>	106
5.3. Discussion.....	112
5.3.1. Blastoderm expression points to a separate mechanism for mandibular segment patterning in <i>Tribolium</i>	112
5.3.2. Expression dynamics in the SAZ indicate that <i>Tc-h</i> , <i>Tc-dpn</i> and <i>Tc-emc</i> oscillations are linked to the <i>Tc-eve/Tc-odd/Tc-run</i> circuit.....	113
5.3.3. Other <i>Hes/Hey</i> expression patterns are consistent with roles in neurogenesis and posterior signalling.....	117
Chapter 6. Investigating function of candidate <i>Hes</i> genes and <i>Tc-emc</i>.....	121
6.1. Introduction.....	121
6.1.1. An ancestral role for arthropod <i>Hes</i> genes in Notch-mediated segmentation.....	121
6.1.2. Existing functional data for <i>Hes</i> genes and Notch signalling in <i>Tribolium</i>	122
6.1.3. Aims of the chapter.....	123
6.2. Results.....	124
6.2.1. Use of non-overlapping dsRNA fragments for RNAi knockdowns.....	124
6.2.2. Single knockdowns of <i>Tc-h</i> and <i>Tc-dpn</i> corroborate previously published data.....	124
6.2.3. <i>Tc-emc</i> RNAi produces a high penetrance of abnormalities relating to abdominal segmentation, dorsal closure, head, leg, and hindgut development.....	127
6.2.4. <i>Tc-Sidpn1</i> RNAi induces abdominal truncation and anterior head defects.....	134
6.2.5. Double knockdowns fail to reveal additive effects or redundancy between <i>Tc-h</i> , <i>Tc-dpn</i> and <i>Tc-emc</i>	137

6.2.6. Testing for redundancy between <i>Hes</i> genes and <i>Tc-odd</i>	139
6.3. Discussion.....	146
6.3.1. Phenotypes are consistent with a role for <i>Tc-emc</i> as a mediator of Notch signalling in <i>Tribolium</i>	146
6.3.2. <i>Tc-emc</i> phenotypes: A reflection bHLH function in abdominal segmentation?.....	148
6.3.3. A possible role for <i>Tc-Sidpn1</i> in posterior signalling during GBE.....	151
6.3.4. Joint knockdown of <i>Tc-h</i> with <i>Tc-odd</i> exposes redundancy in thoracic segment patterning and/or leg development.....	152
6.3.5. Caveats and future experiments.....	155
Chapter 7. General Discussion.....	157
7.1. <i>Odd</i> family evolution and expression: Candidates for the transition to pair-rule patterning in insects?.....	157
7.2. <i>Tc-emc</i> segmentation phenotype renews the discussion of <i>Hes</i> and Notch signalling involvement in the <i>Tribolium</i> segmentation clock.....	164
7.3. Shifts in patterning mechanisms over the course of <i>Tribolium</i> segment addition.....	168
7.4. General remarks and future directions.....	173
Appendices.....	177
Appendix A.....	177
Appendix B.....	179
Appendix C.....	180
Appendix D.....	181
Appendix E.....	190
Appendix F.....	191
Bibliography.....	192

LIST OF FIGURES

Figure 1.1 - Prevalence of body segmentation among Bilateria.....	2
Figure 1.2 - Divergent mechanisms of segment patterning in arthropods.....	5
Figure 1.3 - The <i>Drosophila</i> “segmentation cascade”	9
Figure 1.4 - Sequential segmentation.....	11
Figure 1.5 - The three-tier clock-and-wavefront model.....	13
Figure 1.6 - Candidate oscillator genes and the divergence of Notch-mediated synchrony in the arthropod segmentation clock.....	16
Figure 1.7 - Proposed <i>Tribolium</i> <i>Tc-eve/Tc-odd/Tc-run</i> oscillator network models.....	21
Figure 1.8 - Weaker <i>Tc-odd</i> ^{RNAi} segmentation phenotype in SB strain confirmed through CRSIPR gene knockout.....	23
Figure 1.9 - PR gene expression in <i>Tc-odd</i> -null SB embryos and hypothetical clock networks.....	24
Figure 1.10 - Highly conserved zinc finger domain among <i>Tribolium</i> Odd family proteins.....	28
Figure 3.1 - Gene conversion can lead to inaccurate phylogenetic inferences.....	41
Figure 3.2 - Arthropod <i>Hes/Hey</i> family gene trees.....	46
Figure 3.3 - Arthropod non- <i>E(spl)</i> <i>Hes</i> gene trees.....	50
Figure 3.4 - Signs of conserved residues in <i>Hes</i> genes that lack orange domains.....	52
Figure 3.5 - Gene structure comparison of non- <i>E(spl)</i> arthropod <i>Hes</i> genes.....	54
Figure 3.6 - Arthropod <i>odd</i> family gene trees.....	57
Figure 3.7 - Nucleotide sequence alignments of conserved ZF coding regions in <i>Parasteatoda</i> , <i>Ischnura</i> , and <i>Oncopeltus</i> <i>odd</i> family genes.....	59
Figure 3.8 - Arthropod <i>odd</i> family gene trees based on nucleotide codon alignments of the zinc finger coding region and surrounding sequence.....	61
Figure 3.9 - Inferred gene conversions in insect <i>odd</i> family genes.....	62
Figure 3.10 - Genomic organisation of the <i>odd</i> family cluster across arthropods.....	65
Figure 3.11 - Arthropod <i>odd</i> family gene structures.....	67
Figure 3.12 - Comparison of the 5th ZF amino acid sequence in selected insect Odd proteins compared to those of other arthropod Odd family proteins.....	68
Figure 4.1 - <i>Tc-wg</i> staging through GBE.....	81
Figure 4.2 - Expression of <i>Tc-drm</i> and <i>Tc-sob</i> in relation to <i>Tc-odd</i> during segment addition.....	82
Figure 4.3 - Post GBE expression of <i>odd</i> family genes in <i>Tribolium</i>	85

Figure 4.4 - Comparison of <i>Tc-drm</i> and <i>Tc-sob</i> expression in SB and GA2 strains during segment addition.....	86
Figure 4.5 - Previously published <i>Tc-bwl</i> and <i>Tc-sob in situ</i> hybridisations used probes overlapping the highly similar zinc finger coding regions.....	88
Figure 4.6 - Comparison of previously published <i>Tc-sob</i> and <i>Tc-bwl</i> expression patterns to those from this study.....	90
Figure 4.7 - Review of <i>odd</i> family <i>in situ</i> probes from published expression studies in arthropods.....	93
Figure 5.1 - <i>Tc-h</i> expression in <i>Tribolium</i>	101
Figure 5.2 - <i>Tc-dpn</i> expression in <i>Tribolium</i>	102
Figure 5.3 - <i>Tc-emc</i> expression in <i>Tribolium</i>	103
Figure 5.4 - <i>Tc-Sidpn1</i> , <i>Tc-Sidpn2</i> , <i>Tc-E(spl)1</i> , <i>Tc-E(spl)3</i> and <i>Tc-Hey1</i> expression during GBE....	105
Figure 5.5 - Early blastoderm expression of <i>Tc-h</i> , <i>Tc-dpn</i> and <i>Tc-emc</i>	107
Figure 5.6 - Co-expression analysis of striped <i>Hes</i> genes and <i>Tc-odd</i> in during GBE.....	110
Figure 5.7 - Expression dynamics of cyclical genes in the <i>Tribolium</i> SAZ compared with orthologues in <i>Drosophila</i>	115
Figure 5.8 - Hypothetical regulatory logic of <i>Tribolium Hes</i> genes in the SAZ.....	116
Figure 6.1 - parental RNAi with <i>Tc-hairy</i> 3' dsRNA confirms published head phenotype.....	125
Figure 6.2 - Impact of <i>Tc-emc</i> eRNAi on abdominal segmentation.....	130
Figure 6.3 - eRNAi reveals role of <i>Tc-emc</i> in <i>Tribolium</i> head development.....	132
Figure 6.4 - <i>Tc-emc</i> leg and hindgut phenotypes.....	133
Figure 6.5 - <i>Tc-Sidpn1</i> RNAi results.....	135
Figure 6.6 - eRNAi results for <i>Tc-emc/Tc-h</i> , <i>Tc-emc/Tc-dpn</i> double knockdowns, compared to GFP control and <i>Tc-emc</i> single knockdowns.....	138
Figure 6.7 - <i>Tc-odd/Tc-h</i> and <i>Tc-odd/Tc-dpn</i> double knockdown results.....	140
Figure 6.8 - <i>Tc-h</i> single and <i>Tc-h/Tc-dpn</i> double knockdowns in <i>Tc-odd</i> -null mutant.....	144
Figure 6.9 - <i>Tc-odd^{-/-}</i> cuticle scoring.....	145
Figure 6.10 - <i>Emc</i> gene regulation.....	149
Figure 7.1 - The evolution of <i>odd</i> -family gene expression in arthropods.....	160
Figure 7.2 - Linking Notch signalling to <i>Tc-emc</i> expression and function.....	167
Figure 7.3 - Differential patterning across <i>Tribolium</i> segmentation.....	173
Figure A.1 - Amino acid multiple sequence alignment of Orange domains from arthropod non- <i>E(spl)</i> HES proteins.....	177

Figure A.2 - Amino acid multiple sequence alignment of zinc finger 5 from arthropod Odd family proteins.....	178
Figure B.2 - Expression dynamics of <i>Tc-h</i> and <i>Tc-odd</i> are consistent in the SAZ throughout GBE.....	179
Figure C.1 - <i>Tc-h</i> , <i>Tc-dpn</i> and <i>Tc-emc</i> all exhibit expression patterns consistent with roles in hindgut and malpighian tubule specification.....	180
Figure E.1 - Additional examples of <i>Tc-emc</i> RNAi induced localised abnormalities.....	190
Figure F.1 - ISH sense controls for <i>Tc-emc</i> , <i>Tc-Sidpn1</i> , <i>Tc-Sidpn2</i> and <i>Tc-Hey1</i>	191

LIST OF TABLES

Table 2.1 - Primers used for probe and dsRNA fragments.....	33
Table 2.2 - HCR probes.....	37
Table 6.1 - Results overview for <i>Tc-dpn</i> parental RNAi.....	127
Table 6.2 - Overview of <i>Tc-emc</i> eRNAi.....	128
Table D.1 - <i>Tc-h</i> parental RNAi results.....	181
Table D.2 - Impact of <i>Tc-emc</i> pRNAi on egg laying.....	182
Table D.3 - <i>Tc-Sidpn1</i> parental RNAi results in a high proportion of empty eggs.....	182
Table D.4 - <i>Tc-emc</i> embryonic RNAi results.....	183
Table D.5 - <i>Tc-Sidpn1</i> embryonic RNAi results.....	184
Table D.6 - eRNAi results for <i>Tc-emc/Tc-h</i> , and <i>Tc-emc/Tc-dpn</i> double knockdowns, compared with <i>Tc-emc</i> singles.....	185
Table D.7 - <i>Tc-h/T-dpn</i> pRNAi double knockdown.....	186
Table D.8 - pRNAi results for <i>Tc-h/Tc-odd</i> and <i>Tc-dpn/Tc-odd</i> double knockdowns.....	187
Table D.9 - <i>Tc-odd</i> mutant pRNAi results.....	188
Table D.10 - <i>Tc-dpn</i> pRNAi egg laying.....	189

LIST OF ABBREVIATIONS

A1-10 – Abdominal segments 1 to 10

AEL – After egg lay

Ant – Antenna

AP – Anterior-posterior

bwl – *bowel*

cwo – *clockwork orange*

dpn – *deadpan*

drm – *drumstick*

dsRNA – double-stranded RNA

emc – *extramacrochaetae*

en – *engrailed*

eRNAi – embryonic RNA interference

E(spl) – *enhancer-of-split*

eve – *even-skipped*

FISH – fluorescence *in situ*

hybridisation

GA2 - Georgia

GBE – Germband elongation

h – *hairy*

Hes – *hairy/enhancer-of-split*

Hey – *Hes-related with YRPW motif*

HCR – Hybridisation Chain Reaction

ISH – *in situ* hybridisation

Lb – Labia

Lbr – Labrum

Mn – Mandible

Mx – Maxilla

odd – *odd-skipped*

pRNAi – parental RNA interference

RNAi – RNA interference

run – *runt*

SAZ – segment addition zone

SB – San Bernadino

Sidpn – *similar to deadpan*

sob – *sister-od-bowel*

T1-3 – Thoracic segments 1 to 3

wg – *wingless*

1. GENERAL INTRODUCTION

1.1. Animal segmentation as a case study for Evo-Devo

One of the fundamental research goals of biology is to understand the mechanisms by which morphological evolution and innovation occurs. Evolutionary developmental biology (Evo-Devo) is a field with a long history in the search for answers to this question, and with the ever-growing wealth of research tools and genomic resources at our disposal, our understanding of this process is constantly improving.

A long-standing case study within Evo-Devo is that of animal body segmentation. This is a defining characteristic of several distinct bilaterian taxa, in which the anterior-posterior axis of the animal is divided into a series of repetitive units, or segments. This trait is observed among Panarthropods (arthropods, Onychophora, Tardigrada), annelid worms, and Chordates (vertebrates, Cephalochordata), and is believed to have attributed to the great success and diversity of some of these groups (**Fig. 1.1**) (Hannibal and Patel, 2013; Dunn et.al., 2014). A segmented body plan offers significant potential for morphological diversification as individual segments can independently evolve to perform unique tasks (Tautz, 2004). This is perhaps most clearly illustrated by arthropods which have evolved a plethora of specialised segments.

Segmentation is of particular interest to evolutionary biologists due to the relative positions of segmented taxa among unsegmented taxa in the metazoan phylogeny (**Fig 1.1**). This raises the question as to whether segmentation has evolved independently in these groups or whether it has just 1 or 2 ancestral origins at the dawn of Bilateria or Protostomia, respectively, having subsequently been lost in multiple lineages. A combination of evidence from developmental, phylogenetics, and fossil studies lends support to independent origins of segmentation as the most parsimonious explanation (Reviewed by Chipman, 2010). Indeed, the widespread loss of segmentation that would have had to occur in the case of any ancestral origin seems unlikely, particularly given the selective advantages that segmentation offers. Another key observation from which the conclusion of independent origins is based, is that segment patterning occurs across

different germ layers between arthropods (ectoderm), annelids (ectoderm and mesoderm), and vertebrates (mesoderm) (Valentin and Oates, 2013). As an example of convergent evolution, we can use comparative developmental research to identify similarities and differences between the underlying genes and molecular mechanisms that have been employed into segment patterning in each case, thereby highlighting common themes underpinning the evolution of this trait.

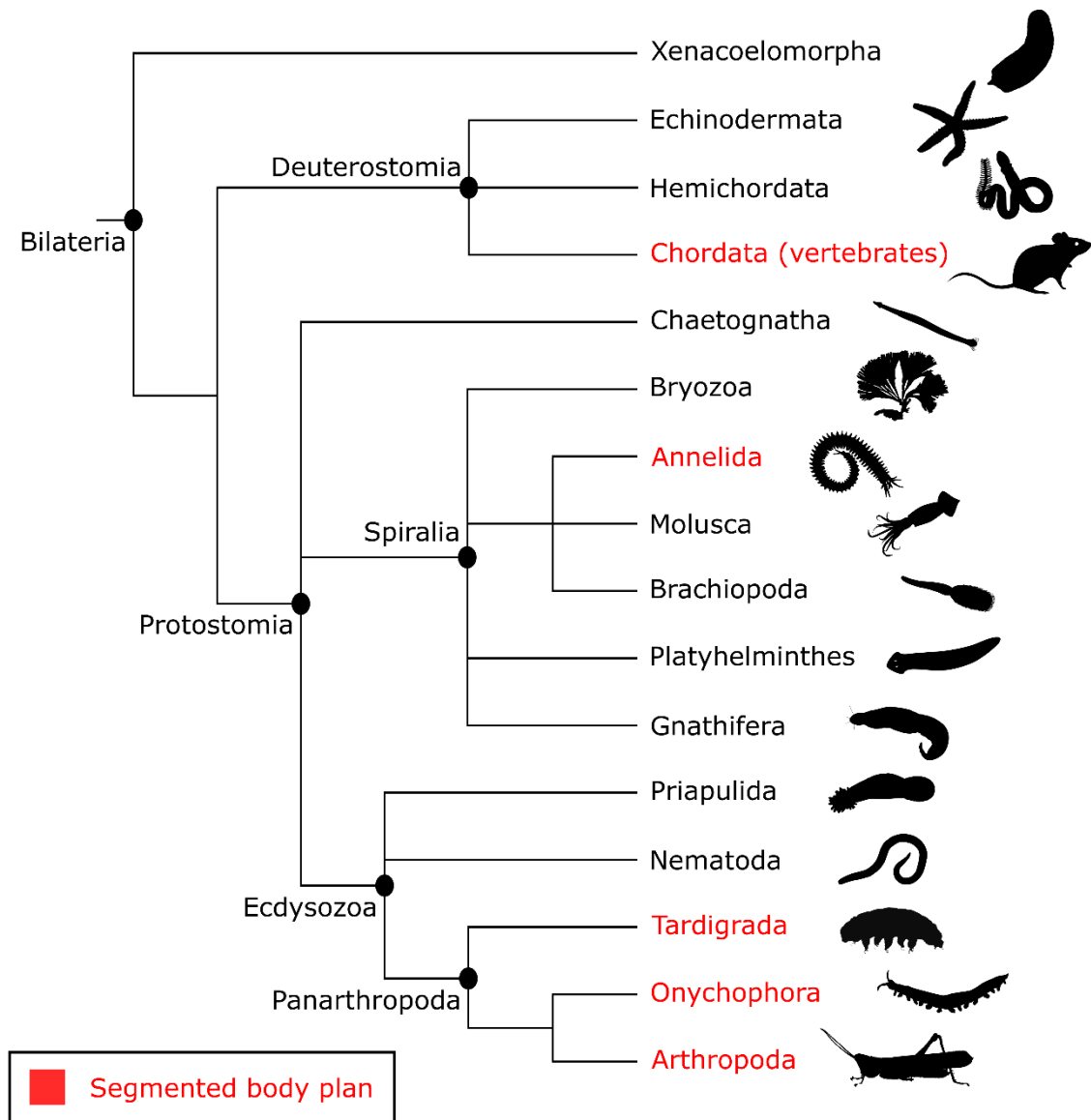


Figure 1.1. Prevalence of body segmentation among Bilateria. Segmented phyla highlighted in red. Figure adapted from Dunn et.al. (2014).

1.2. Segmentation in arthropods

Another reason why segmentation remains a current area of focus for evo-devo research, particularly within arthropods, is that seemingly very different mechanisms for patterning segments have evolved in this phylum. This is of interest, given that the segmented body plan of arthropods represents a core morphological constraint that is shared among all arthropod species.

One of the most notable differences observed in arthropods is between species, such as *Drosophila*, in which body segments appear to be patterned almost simultaneously in the blastoderm, in contrast to those that pattern the majority of their segments sequentially at the posterior of the embryo during germband elongation (GBE) (**Fig. 1.2A**) (Clark et.al., 2019). Note that, even in sequentially segmented arthropods, there are always several pre-gnathal segments patterned in the blastoderm. However, these are established by a different mechanism relative to the more posterior segments that will be the focus of this thesis (Chipman and Edgecombe, 2019). Some species, such as the wasp, *Nasonia vitripennis*, or the milkweed bug, *Oncopeltus fasciatus*, use a mixed mode of segmentation, in which anterior segments are patterning simultaneously in the blastoderm, in a manner reminiscent of *Drosophila*, while posterior segments are added sequentially during GBE (Rosenberg et.al., 2014; Chipman, 2017). These different modes of segmentation are often correlated with different lengths of germ anlage, leading researchers to historically classify simultaneously segmenting arthropods as long germ, while “mixed mode” or sequentially segmenting species would be described as intermediate or short germ arthropods respectively (Davis and Patel, 2002). However, germ size does not always accurately reflect the mode of segmentation (Tidswell, 2021). Therefore, consistent with recent publications, I will instead use the terms “sequential” or “simultaneous” to distinguish between these different modes of segment patterning (Clark et.al., 2019).

Sequential segmentation is widespread among arthropods and represents the ancestral mode of segment patterning in this group, while simultaneous segmentation has only been observed across holometabolous insect lineages (**Fig. 1.2B**) (Davis and Patel, 2002). Despite the seemingly drastic differences between these two modes,

simultaneous segmentation is believed to have evolved independently multiple times in insects, suggesting that the regulatory changes required for this transition are not prohibitively complex (Clark et.al., 2019). Furthermore, recent research has begun to bridge the gap between sequential and simultaneous segmentation, identifying important heterochronic shifts in timing factors, as well as new subtleties in the *Drosophila* model that question the entirely simultaneous nature of this system (Clark, 2017; Clark and Peel, 2018).

Another striking difference between segment patterning across arthropods, is the evolution of pair-rule patterning. All segment patterning is based upon the establishment of periodic stripes of gene expression that ultimately define segment boundaries. In arthropods that exhibit pair-rule patterning, segmentation genes are initially expressed in a double-segment periodicity (i.e. 1 stripe for every 2 segments) before transitioning to a single-segment periodicity (1 stripe per segment) as the segment pattern resolves (Clark et.al., 2019). This appears to have evolved independently in insects and the centipede, *Strigamia maritima*, while there is no sign of double-segment periodicity in spiders, crustaceans, or the millipede, *Glomeris marginata* (**Fig. 1.2B**) (Chipman et.al., 2004; Davis et.al., 2005; Schwager, 2008; Janssen et.al., 2011; Eriksson et.al., 2013). Exactly how such a change occurred during evolution remains a puzzling question with several different hypotheses (reviewed by Clark et.al., 2019).

Most of what we know about arthropod segmentation comes from the extensively studied arthropod model, *Drosophila*. However, as previously mentioned, this species employs a highly derived mode of simultaneous segmentation and is therefore not representative of most arthropods. Research has since turned to other arthropod models that have retained the ancestral mode of sequential segmentation in an effort to rectify this. However, a detailed understanding of the precise molecular mechanisms that control sequential segmentation is still lacking.

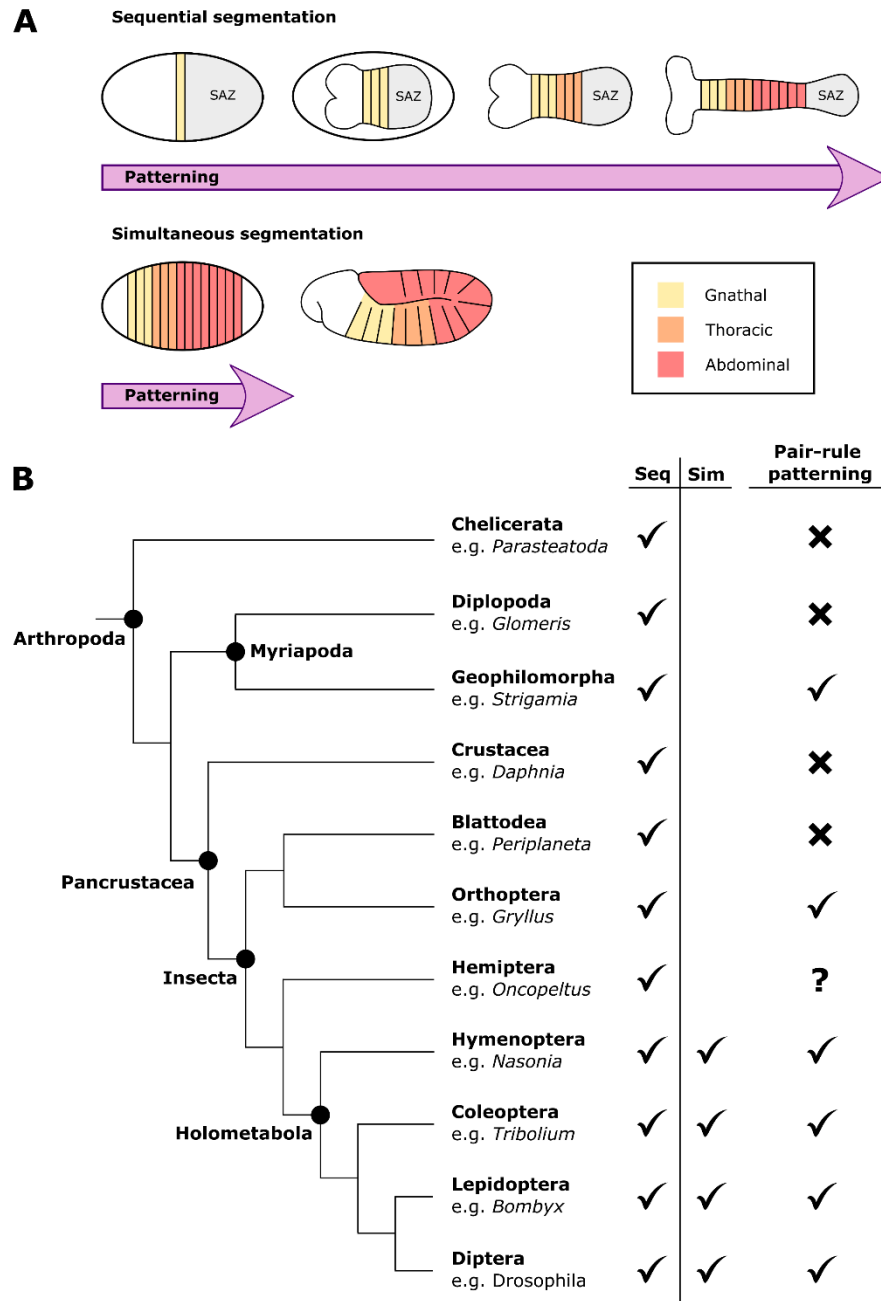


Figure 1.2. Divergent mechanisms of segment patterning in arthropods. (A) Schematic illustrating sequential (e.g. *Tribolium*) vs simultaneous (e.g. *Drosophila*) modes of segmentation. SAZ refers to the segment addition zone, an undifferentiated posterior region where sequential segment patterning occurs. (B) arthropod phylogeny to show the existence of sequential(seq)/simultaneous(sim) segmentation in different taxa. Examples of both simultaneous and sequential segmentation are found in holometabolous insect orders. Presence or absence of pair-rule patterning is based on the example model species given for each taxa. Question mark reflects the unclear nature of pair-rule patterning in *Oncopeltus fasciatus* (Clark et.al., 2019). Figure adapted from Tidswell (2021). Phylogeny based on Misof et.al. (2014).

Using the red flower beetle, *Tribolium castaneum*, as a model for sequential segmentation, this thesis attempts to address aspects of patterning that remain poorly understood within this system. With this, it may then be possible to make more accurate inferences regarding the evolution of segmentation in arthropods, both in terms of its origin, and how patterning mechanisms have subsequently diverged. Furthermore, segment patterning in both annelids and vertebrates also occurs in a sequential manner during periods of axial extension (Chipman, 2010). A deeper understanding of sequential segmentation in arthropods may therefore contribute towards future comparisons between other segmented phyla, and provide us with insights into some of the common mechanisms and genes that have been independently recruited in each case.

1.3. The *Drosophila* model and the history of segmentation gene nomenclature

While *Drosophila* segmentation may not be representative of most arthropods, it was the first species in which segment patterning and segmentation genes were described. As such, the classifications of segmentation genes, which are commonly referred to in comparative studies of other arthropods, are all based upon this model. *Drosophila* segment patterning relies on a hierarchical system of gene expression and regulation, termed the “segmentation cascade”, to ultimately define segment boundaries (**Fig. 1.3**) (reviewed in Nasiadka et.al., 2002). Note that, much of this occurs within a syncytial blastoderm. This allows the diffusion of RNA transcripts and proteins, which subsequently form concentration gradients from their points of synthesis (Peel et.al., 2005). There are four main gene classes that are distinguished by the timing of their expression within the “segmentation cascade”.

Maternal co-ordinate genes are the first of these classes that convey positional information along the egg (**Fig. 1.3A**). The mRNA transcripts of these genes are maternally derived and their positions within the egg are established during oogenesis (Johnston and Nuusslein-Volhard, 1992). After fertilization, the translation of these transcripts and diffusion of their protein products results in broad protein gradients along the egg (Peel et.al., 2005). These include anterior-to-posterior gradients of the

protein products, Bicoid (BCD) and Hunchback (HB), and posterior-to-anterior gradients of Caudal (CAD) and Nanos (NOS).

The next class of segmentation genes are the gap genes (**Fig. 1.3B**), whose expression domains are initially defined by inputs from the maternal co-ordinate genes before being further refined by cross regulation from each other (reviewed in Jaeger, 2011). While the expression patterns of gap genes are aperiodic, the graded concentrations of their protein products, along with those of maternal co-ordinate proteins, are sufficient to convey unique positional identities along the embryo that are interpreted by the next set of segmentation genes, the pair-rule genes (**Fig. 1.3C**) (Schroeder et.al., 2011).

In *Drosophila*, there are 7 pair-rule genes, *hairy (h)*, *even-skipped (eve)*, *runt (run)*, *odd-skipped (odd)*, *fushi tarazu (ftz)*, *paired (prd)*, and *sloppy-paired (slp)*, which were initially identified by mutations resulting in the loss of alternating segments (Nüsslein-Volhard and Wieschaus, 1980). Pair-rule genes are the first genes that convey a periodic pattern along the AP axis, with each initially expressed in 7 transverse stripes that reflect a double-segment periodicity (1 stripe for every 2 prospective segments) (**Fig. 1.3C**) (Clark, 2017). Pair-rule genes can be further categorised into primary and secondary pair-rule genes, based on the order in which they appear and how their striped expression is established (Ingham, 1988). Primary pair-rule genes are the first to be expressed and their stripes are patterned via combinations of maternal and gap inputs that generate periodicity by activating “stripe-specific” enhancer elements (Schroeder et.al., 2011). Stripes of secondary pair-rule genes appear later and instead take their regulatory cues from the already periodic pair-rule genes. These classifications have historically been the source of much debate, however, close examination of the regulatory dynamics underpinning pair-rule stripes has formed the basis for *h*, *eve*, *run*, *odd* and *ftz* being defined as primary pair-rule genes due to their initial “stripe-specific” regulation, while *prd* and *slp* are expressed later and fit the description of secondary pair-rule genes (Schroeder et.al., 2011).

Among the primary pair-rule genes, there is still variation in the extent to which stripes are generated through stripe-specific elements versus cross-regulation from other pair-rule genes. In most cases, except for *h* (which relies solely on stripe-specific elements), combinations of each of these processes is required for the refinement and even

spacing of pair-rule stripes (Schroeder et.al., 2011). The stripes for each pair-rule gene at this point are slightly out of phase with each other, and their overlapping patterns are sufficient to confer positional information at an almost cellular resolution along the AP axis (Clark, 2017). At gastrulation, there is a sudden switch in the regulatory logic of the pair-rule network, triggered by expression of the timing factor, *opa* (Clark and Akam, 2016). This change drives a transition from double- to single-segment periodicity for several pair-rule genes, and the expression of the final class of genes from the *Drosophila* “segmentation cascade”, the segment-polarity genes (**Fig. 1.3D**).

Activated by *opa*, and spatially regulated by pair-rule genes, the segment polarity genes, *wingless* (*wg*) and *engrailed* (*en*), are each expressed in 14-15 alternating periodic stripes along the length of the embryo (Benedyk et.al., 1994; Nasiadka et.al., 2002). These stripes define the boundaries of patterning units known as parasegments which, while equal in size, are offset slightly from the eventual morphological segments. This is believed to reflect the discrepancy between the segmental organisation of the nervous system compared to that of the epidermis and musculature, which themselves are similarly offset (Deutsch, 2004).

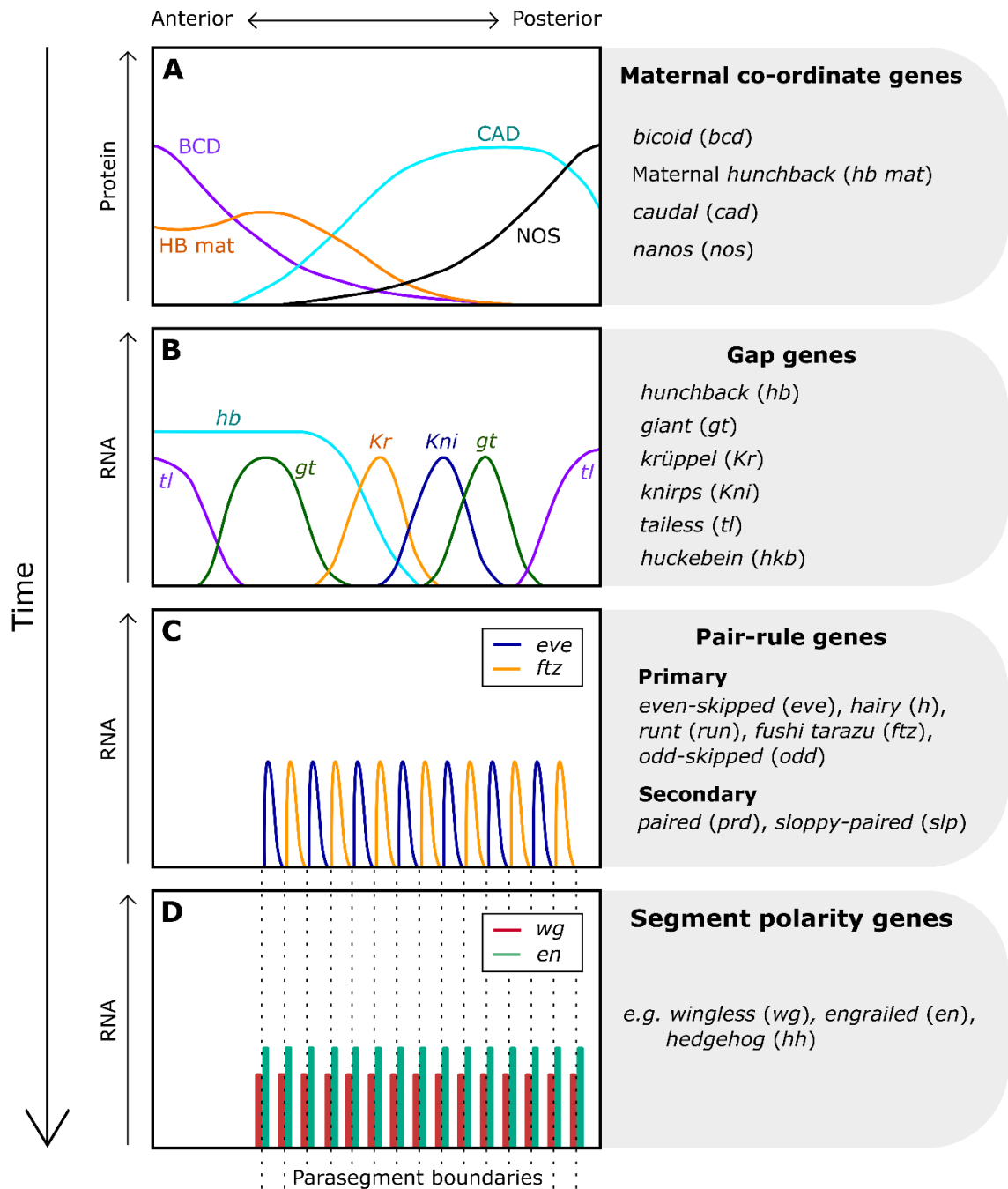


Figure 1.3. The *Drosophila* “segmentation cascade”. An overview of protein and mRNA gradients/domains along the anterior-posterior axis of the developing *Drosophila* blastoderm. Members of each class of segmentation genes are listed to the right of their respective expression profiles. Dotted lines indicate parasegment boundaries defined by adjacent stripes of *wg* and *en*. Note that graphs are for illustrative purposes, and do not reflect real values. Figure adapted from Peel et.al. (2005) with maternal gradients from Jaeger (2011)

1.4. Sequential segmentation

In contrast to *Drosophila*, most other arthropods have retained the ancestral mode of sequential segmentation, in which segments are patterned one or two at a time, from head to tail, during germband elongation (Davis and Patel, 2002). Some anterior segments are usually patterned within the blastoderm, but in general, most of the patterning occurs within a posterior region of the germband, termed the “segment addition zone” (SAZ) (**Fig. 1.4**). Note that this is the updated terminology for what has historically been known as the “growth zone”. This is to avoid the assumption that this region is undergoing continuous cell proliferation (or “growth”), which is not always true, as cell movement and rearrangement also contributes to axial elongation in arthropods, albeit to varying degrees (Mito et.al., 2011; Benton and Pavlopoulos, 2014; Benton et.al., 2016; Auman et.al., 2017; Clark et.al., 2019; Constantinou et.al., 2020; Cepeda et.al., 2021). The SAZ varies in size and shape between species, and even between different stages of GBE in the same species (Nakamoto et.al., 2015). Its boundaries are largely defined by the conserved expression of *Wnt* genes at its posterior, and the emerging periodic stripes of segment polarity gene expression at the anterior (Williams and Nagy. 2017; Clark et.al., 2019; Mundaca-Escobar et.al., 2022). The SAZ can further be divided into the “anterior SAZ”, in which segments are in the process of being patterned, and the “posterior SAZ”, where cells are yet to be assigned to a particular segmental fate (**Fig. 1.4**) (Clark et.al., 2019). In *Tribolium*, these regions correlate with the expression of key developmental timing factors, *caudal* (posterior SAZ) and *odd-paired* (anterior SAZ) (Clark and Peel, 2018).

Candidate gene approaches in sequentially segmenting arthropods commonly find pair-rule homologues expressed in transverse stripes within the SAZ (**Fig. 1.4**). While there are some discrepancies in the relative positions of these stripes to one another, and in the segment periodicity to which they are expressed, these patterns have been reported in species from across the arthropod phylum (Chelicerata: Damen et.al., 2000; Schönauer et.al., 2016; Myriapoda: Green and Akam, 2013; Hughes and Kaufman, 2002; Janssen et.al., 2011; insects: Mito et.al., 2007; Auman and Chipman, 2018; Patel et.al., 1994).



Figure 1.4. Sequential segmentation. Key gene expression patterns and functional regions of a *Tribolium* embryo during germband elongation. Pair-rule genes (blue) are expressed in transverse stripes in the segment addition zone (SAZ), while segment polarity genes (green) define the segmented germband and mark the anterior edge of the SAZ. The SAZ is further divided into two functional regions that correlate with the expression of timing factors such as *Tc-caudal* (posterior SAZ; dark pink) and *Tc-opa* (anterior SAZ; light pink) (Clark and Peel, 2018). Figure adapted from Clark et.al. (2019)

In sequentially segmenting arthropods, we can consider the prepatterned SAZ as analogous to the *Drosophila* blastoderm at the point where pair-rule stripes are being established and refined, while the anterior germband that has already undergone patterning represents a more mature stage of development. Indeed, conserved sequential expression of timing factors, such as *opa* and *caudal*, between *Drosophila* and *Tribolium* are consistent with this comparison (Clark and Peel, 2018). When we apply this logic, we see that, like *Drosophila*, primary pair-rule gene expression in sequentially segmenting arthropods precedes that of the segment polarity genes. Consistent with this, knockdowns of pair-rule genes in the sequentially segmenting beetle, *Tribolium*, results in the disruption of segment polarity gene expression, as well as of other pair-rule genes, supporting a similar regulatory framework to that in the *Drosophila* “segmentation cascade” (Choe et.al., 2006). However, where in *Drosophila*, pair-rule stripes take spatial cues from gap genes and maternal factors, in sequentially segmenting arthropods, these stripes are instead generated by inherently periodic oscillations of gene expression.

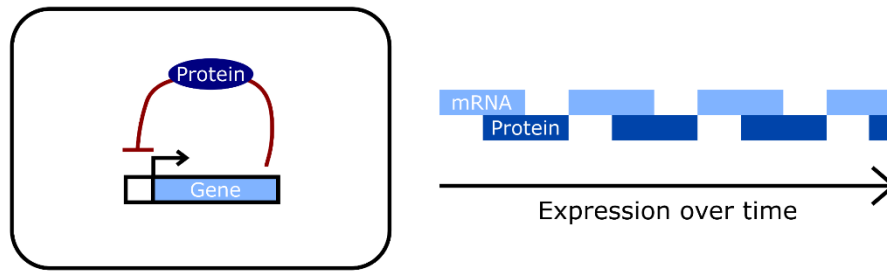
1.5. The clock and wavefront model

Evidence for the oscillatory expression of pair-rule genes was first demonstrated in *Tribolium*, when Sarrazin et.al. (2012) used split-embryo experiments and live cell tracking to show that *Tc-odd* stripes are dynamic in nature, and that individual cells in the SAZ experience temporal changes in expression levels (i.e. oscillations). While the same types of experiment have not been widely conducted in other sequentially segmented arthropods, the variable expression patterns of segmentation genes between carefully staged embryos indicates that similar dynamics most likely exist in the SAZ of other arthropods from across the phylum (Brena and Akam, 2013; Schönauer et.al., 2016).

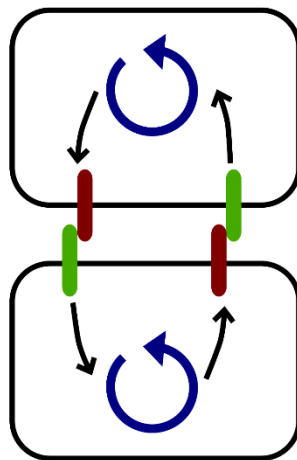
It has been proposed that this oscillatory expression of segmentation genes is part of an arthropod “segmentation clock” analogous to the clock-and-wavefront mechanism that patterns segmental units (somites) in vertebrates (Clark et.al., 2019). While arthropod segments and vertebrate somites are not homologous morphological structures, there is evidence that related genes have been independently co-opted into similar developmental roles during their evolution. Where in arthropods, this patterning occurs along the SAZ, in vertebrates it takes place in an undifferentiated posterior region of the elongating embryo, termed the pre-somitic mesoderm (PSM) (Oates et.al., 2012). The vertebrate clock-and-wavefront model has been described in detail in several different model species (reviewed by Oates et.al., 2012). In arthropods, however, the precise network topology of this system is still unclear.

The clock and wavefront mechanism of segment patterning, first proposed by Cooke and Zeeman (1976), can be divided into 3 tiers of regulatory activity. In its simplest form, it relies on (1) intracellular oscillations of gene expression (**Fig. 1.5A** (cellular oscillator)), (2) Co-ordination of these oscillations across cells (**Fig. 1.5B** (local synchrony)), and (3) posterior-to-anterior regulatory control to activate, slow, and eventually stop oscillations as they reach the anterior PSM/SAZ (**Fig. 1.5C** (global control))), thereby leaving behind fixed periodic stripes of gene expression that convey the segment pattern (Oates et.al., 2012).

A Cellular oscillator



B Local synchrony



C Global control

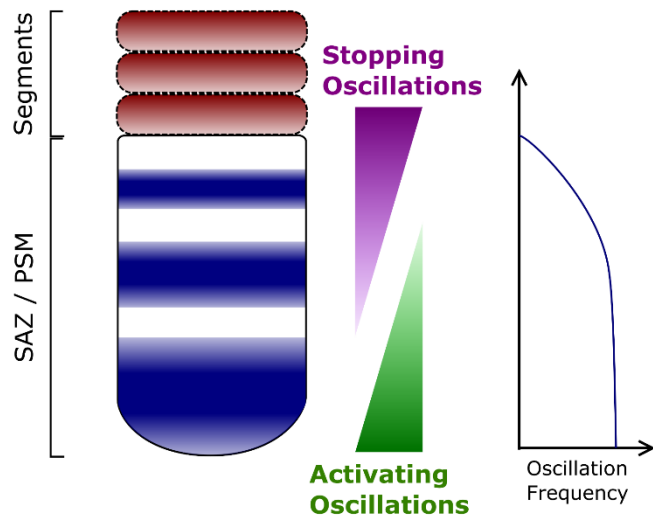


Figure 1.5. The three-tier clock-and-wavefront model. (A) Cellular oscillations driven by a negative feedback loop. In this case, a single auto-repressive gene is shown as an example, however, cellular oscillations can also be driven by multi-gene circuits such as the *Tc-eve/Tc-odd/Tc-run* oscillator in *Tribolium* (Choe et.al., 2006). (B) Intracellular oscillations are synchronised across cells through intercellular signalling (red ligands, green receptors). In vertebrates, and some arthropods, this occurs via the Notch signalling pathway. (C) Opposing anterior-posterior gradients of regulatory factors control the oscillation frequency across the SAZ/PSM. Cellular oscillations are initiated at the posterior end of the embryo. These temporal oscillations are synchronised between neighbouring cells and present themselves spatially in the form of dynamic “waves” that propagate anteriorly. Oscillations are then arrested in the anterior SAZ/PSM, fixing these “waves” in place and laying down the segment pattern. Figure adapted from Liao and Oates (2017)

Cellular oscillations and local synchrony

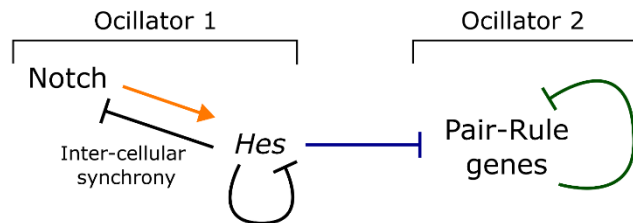
In vertebrates, members of the *hairy/enhancer-of-split (Hes)* gene family are known to be key drivers of intracellular oscillations, achieving this via the auto-inhibitory action of HES proteins on their own expression (**Fig. 1.5A**) (reviewed in Kageyama et.al., 2012). Transcriptome sequencing of the PSMs of zebrafish, mice and chicks, shows that *Hes* genes are the only cyclical genes conserved across all three species, demonstrating their importance (Krol et.al., 2011). As well as regulating their own expression, *Hes* genes also inhibit modulators of the intercellular Notch signalling pathway, which in turn synchronises *Hes* oscillations to neighbouring cells (**Fig. 1.5B**) (Kageyama et.al., 2012).

In some arthropods, there is also evidence for similar roles of *Hes* genes and Notch signalling in segmentation (**Fig. 1.6**). Indeed, the primary pair-rule gene, *hairy*, is itself a member of the *Hes* gene family. In the spider, *Cupiennius salei*, *hairy*, *notch* and *delta*, all exhibit dynamic striped expression in the SAZ, and perturbing the Notch pathway results in malformed segments and irregular *hairy* expression (Stollewerk et.al., 2003). Similar expression patterns for these genes have been observed in several other arthropod models, and knockdowns of key Notch signalling genes and/or *Hes* genes in the crustacean, *Daphnia magna*, the cockroach, *Periplaneta americana*, and the moth, *Bombyx mori*, all confirm important segment patterning roles in these species also (**Fig. 1.6**) (Pueyo et.al., 2008; Eriksson et.al., 2013; Liu, 2013). These results support the idea that Notch signalling synchronises cellular *hairy/Hes* oscillations in sequentially segmented arthropods, as is the case in vertebrates. However, the exact topology of the gene regulatory network(s) that drives *Hes* oscillations, as well as the downstream targets of HES transcription factors, are still unknown. Furthermore, there is contradicting evidence for the roles of these genes and pathways in other sequentially segmenting insects, with expression data and functional studies in the cricket, *Gryllus bimaculatus*, the milkweed bug, *Oncopeltus fasciatus*, and the beetle, *Tribolium*, suggesting that *Hes* genes and/or Notch signalling are not required in the segmentation clocks of these species (**Fig. 1.6**) (Aranda et.al., 2008; Kainz et.al., 2011; Reding et.al., 2019).

In *Tribolium*, the *Hes* genes, *Tc-h* and *Tc-deadpan* (*Tc-dpn*), are both expressed in dynamic stripes in the SAZ, although single and double knockdowns appear to have no effect on segmentation (Aranda et.al., 2008). There also appears to be no striped SAZ expression of Notch signalling genes and *Tc-delta* knockdowns do not result in any obvious segmentation phenotypes (Aranda, 2006). However, knockdown of the ubiquitously expressed receptor coding gene, *Notch*, has been found to result in abdominal segment defects and missing stripes of the segment polarity gene, *Tc-en* (Kux et.al., 2013). Given the lack of segmental expression of Notch signalling genes, along with the highly pleiotropic nature of the pathway, it is unclear whether these phenotypes are directly related to a function in segment patterning in *Tribolium*.

The discrepancies observed across arthropods only raise more questions regarding the gene regulatory networks that underpin the arthropod segmentation clock, especially among species that do not appear to use Notch signalling to synchronise oscillations. The apparent loss of *Hes*-driven cellular oscillations and local synchronisation by Notch signalling appears to have only occurred during the evolution of some insect lineages, although evidence that it has been retained in *Periplaneta* and *Bombyx* make the timing of this evolutionary transition unclear (**Fig. 1.6**). In *Tribolium*, cellular oscillations appear to instead be driven by a pair-rule gene circuit consisting of *Tc-eve*, *Tc-odd* and *Tc-run*, which generates cyclical expression via a negative feedback loop (**Fig. 1.7**) (Choe et.al., 2006). Given that dynamic stripes of primary pair-rule genes have been observed in the SAZ of species from across the arthropod phylum, and essential roles have been identified in several other sequentially segmented insects, it is possible that similar pair-rule networks drive cellular oscillations in other species also (**Fig. 1.6**). A hypothetical model by Clark et.al. (2019) proposes the existence of, and interaction between, two cellular oscillators, one consisting of a pair-rule circuit, and another of autonomous *Hes* oscillations that are linked to intercellular Notch signalling (**Fig. 1.7A**). However, there is still a lack of sufficient experimental evidence to support this model, and it is currently unclear how cellular oscillations are synchronised in the apparent absence of Notch signalling.

A Hypothesis: Linked intracellular Hes and pair-rule oscillators co-ordinated by Notch signalling



Functional evidence:

*Cupiennius*¹, *Daphnia*⁵,
*Periplaneta*⁶

*Tribolium*¹⁴

*Drosophila*¹⁸

B

	Model organism	Cellular Oscillator		Local synchrony
		Hes genes	PR genes	Notch signalling
Chelicerata	<i>Cupiennius salei</i> (Wandering spider)	<i>h</i> ¹	<i>eve, run, pby-3, odd-r1</i> ²	<i>Notch, Delta-1, Delta-2</i> ¹
Myriapoda	<i>Strigamia maritima</i> (Centipede)	<i>hdprl(1, 2, 3&4)</i> ³	<i>eve(A&B), odr(1, 2&3), run</i> ³	<i>Notch, Delta</i> ⁴
Crustacea	<i>Daphnia magna</i> (Water flea)	<i>Hes2, Hes3</i> ⁵	<i>odl, prd(A&B)</i> ⁵	<i>Notch, Delta</i> ⁵
Insecta				
Blattodea	<i>Periplaneta americana</i> (Cockroach)	<i>h</i> ⁶	<i>eve, run</i> ⁷	<i>Notch, Delta</i> ⁶
Orthoptera	<i>Gryllus bimaculatus</i> (Cricket)	no data	<i>eve</i> ⁸	<i>Notch</i> ⁹ , <i>Delta</i> ¹⁰
Hemiptera	<i>Oncopeltus fasciatus</i> (Milkweed bug)	<i>h, dpn</i> ¹¹	<i>eve, odd, run</i> ¹¹	<i>Delta</i> ¹²
Holometabola				
Coleoptera	<i>Tribolium castaneum</i> (Red flour beetle)	<i>h, dpn</i> ¹³	<i>eve, odd, run</i> ¹⁴	<i>Notch, Delta, Serrate, Fringe</i> ¹³
Lepidoptera	<i>Bombyx mori</i> (Silk moth)	no data	<i>eve, odd, run</i> ¹⁵	no expression data ¹⁷

Table key

Dynamic SAZ expression
Anterior SAZ stripes only
Non-segmental expression



Expression data only

Functional data unclear



Functional data for

Functional data against

Figure 1.6. Candidate oscillator genes and the divergence of Notch-mediated synchrony in the arthropod segmentation clock. (A) A hypothetical model by Clark et.al. (2019) based on regulatory relationships with functional evidence from segmentation studies in different arthropod species. In this model, intercellular Notch signalling synchronises intracellular *Hes* oscillations, which themselves are linked to a pair-rule gene oscillator network (e.g. **Fig. 1.7**). As of yet, there is no evidence for the existence of all of these interactions together in a single species. (B) The phylogenetic tree (based on Misof et.al., 2014) shows the relationships of different arthropod taxa, each of which is represented by an experimental model organism in the corresponding table. Proposed cellular oscillators are divided into pair-rule orthologues (excluding *hairy*) and *Hes* genes, based on the hypothetical model in (A). Genes written in black all exhibit dynamic striped expression across the SAZ consistent with oscillatory expression, while those in blue are expressed in the anterior SAZ only, and therefore likely to be involved in a later stage of patterning. Genes written in red are expressed in patterns that appear unrelated to segmentation. Table box colours indicate functional evidence for or against roles in segmentation for the corresponding groups of genes. Gene abbreviations: *pby-3*, *pairberry-3*; *hdntl*, *hairy-deadpan-like*; *odr/odd-r*, *odd-related*; *odl*, *odd-like*. Superscripted numbers are references: 1. Stollewerk et.al., 2003; 2. Damen et.al., 2005; 3. Green and Akam, 2013; 4. Chipman and Akam, 2008; 5. Eriksson et.al., 2013; 6. Pueyo et.al., 2008; 7. Chesebro, 2012; 8. Mito et.al., 2007; 9. Mito et.al., 2011; 10. Kainz et.al., 2011; 11. Reding et.al., 2019; 12. Auman et.al., 2017; 13. Aranda et.al., 2008; 14. Choe et.al., 2006; 15. Kux et.al., 2013; 16. Nakao, 2015; 17. Liu, 2013; 18, Schroeder et.al., 2011

Posterior to anterior control of regulatory logic

In the vertebrate PSM, posterior-to-anterior gradients of FGF and Wnt family activity oppose anterior-to-posterior gradients of retinoic acid expression, which together convey global regulatory control over cellular oscillations (reviewed in Resende et.al., 2014). Perturbing these pathways produces changes in the length of somites, demonstrating their importance for positioning segment boundaries. In arthropods, it is unknown precisely how the regulatory context across the length of the SAZ changes. However, there are several genes, with conserved expression patterns correlating to specific functional regions, that are thought to be involved (reviewed in Clark et.al., 2019 and Diaz-Cuadro et.al., 2021).

Similar to vertebrates, *Wnt* genes show conserved expression in arthropods at the posterior end of the elongating embryo (Mundaca-Escobar et.al., 2022). This similarity may reflect a shared origin of Wnt signalling for posterior patterning in a bilaterian ancestor, rather than the independent co-option of this network into segmentation (Ryan and Baxevanis, 2007). That said, there is evidence of an indirect role for posterior Wnt signalling in promoting pair-rule gene expression in the spider, *Parasteatoda*, and the beetle, *Tribolium* (Oberhofer et.al., 2014; Schönauer et.al., 2016). This is suspected to occur via downstream targets of Wnt signalling, such as *caudal* and *Dichaete* (Oberhofer et.al., 2014; Clark et.al., 2019). In *Tribolium*, the expression domains of these genes roughly coincide with the region of the SAZ where oscillations are occurring (Clark and Peel, 2018). Furthermore, partially knocking down *Tc-caudal* results in a reduced frequency of *Tc-eve* stripes demonstrating its importance upstream of the clock network (El-Sherif et.al., 2014).

The primary candidate gene for changing the regulatory context in the anterior of the SAZ, is *odd-paired (opa)* (Clark et.al., 2019). *Opa* has a conserved expression profile across distantly related arthropod taxa, where its domain extends into the anterior SAZ throughout GBE, immediately preceding nascent segment polarity gene stripes (Janssen et.al., 2011; Clark and Peel, 2018). Its knockdown in *Tribolium* causes phenotypes consistent with a role in segmentation (Clark and Peel, 2018). Interestingly, *cad*, *Dichaete* and *opa*, all exhibit a conserved order of expression between *Drosophila* and *Tribolium* associated with important shifts in the regulatory logic of both systems, further bridging the gap between sequential and simultaneous segmentation (Clark and Peel, 2018). While there is growing evidence that these timing factors play important roles in global control of the arthropod segmentation clock, the precise molecular mechanisms and regulatory relationships with other genes are still poorly understood.

1.6. *Tribolium* as a model for sequential segmentation research

After *Drosophila*, *Tribolium* is perhaps the most established experimental arthropod model. It has a number of traits that make it easy to work with in the laboratory, as well as characteristics of its development and behaviour that make it well suited for research in evolution and ecology (reviewed by Pointer et.al., 2021). As a sequentially segmented holometabolan insect, *Tribolium* is well placed for researching mechanisms of the arthropod segmentation clock, while still providing a useful comparison against other holometabola that have evolved simultaneous segmentation, such as *Drosophila*. It is largely from comparisons between these two species that similarities reconciling sequential and simultaneous segmentation have been identified, and from which possible explanations for this evolutionary transition have been proposed (Clark et.al., 2019; Diaz-Cuadros et.al., 2021).

Tribolium has a chromosome level genome assembly that has been updated and improved through multiple iterations, most recently by Herndon et.al. (2020). Furthermore, the genome has been fully annotated with the aid of comprehensive RNA-seq data, enabling the quick and accurate identification of candidate gene sequences for developmental research. There are several long established and robust experimental techniques in *Tribolium* for analysing gene expression and function, such as *in situ* hybridisation (ISH) and RNAi respectively, while the use of more advanced methods in imaging, gene manipulation, and transgenics are constantly being developed and improved (Reviewed in Klingler and Bucher, 2022).

Current understanding of the *Tribolium* segmentation clock and outstanding questions

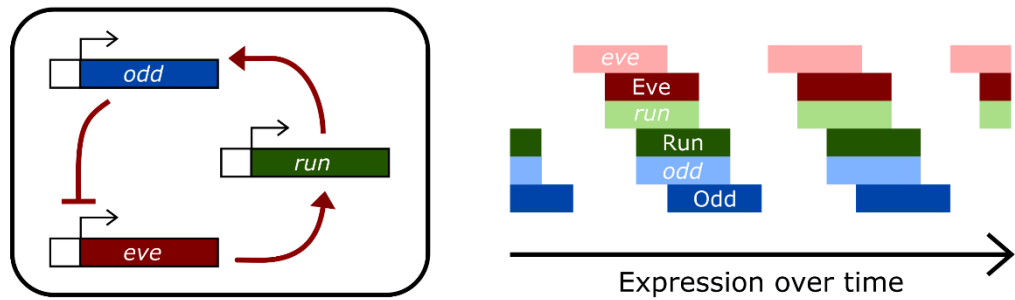
Much of the research regarding sequential segmentation and the clock-and-wavefront model in arthropods has utilized *Tribolium* as a model. Indeed, this was the first species in which evidence for a segmentation clock in arthropods was found (Sarrazin et.al., 2012). That said, there are still aspects of the *Tribolium* clock network that are poorly understood. The pair-rule orthologues, *Tc-eve*, *Tc-odd* and *Tc-run*, are known to play an essential role in segment patterning. These genes are expressed in co-ordinated

dynamic stripes in the SAZ which maintain the same relative order of expression, indicating that their oscillations are linked to each other (Choe et.al., 2006). Single knockdowns of each gene produce almost asegmental phenotypes with just a few head appendages still visible (**Fig. 1.8B**) (Choe et.al., 2006). Note that, as is true for other arthropods, pre-gnathal segments are patterned by a separate mechanism, unrelated to that which patterns more posterior segments (Chipman and Edgecombe, 2019). Therefore, these phenotypes observed by Choe et.al. (2006) are indicative of a complete collapse in sequential segment patterning. Consistent with this, there were no visible stripes of the segment polarity gene, *Tc-en*, in embryos where *Tc-eve*, *Tc-run* or *Tc-odd* had been knocked down (Choe et.al., 2006). In *Tc-odd* knockdown embryos, *Tc-eve* and *Tc-run* are over-expressed in the SAZ, while the knockdown of *Tc-eve* leads to the down-regulation of both *Tc-run* and *Tc-odd*, and *Tc-run* knockdown results in the down-regulation of *Tc-odd* and overexpression of *Tc-eve* (Choe et.al., 2006). These observations led Choe et.al. (2006) to propose a negative feedback loop in which *Tc-eve* activates *Tc-run*, which in turn activates *Tc-odd*, which itself inhibits *Tc-eve* (**Fig. 1.7A**). Based on their precise relative expression patterns, and from what we know of the interactions between these genes in *Drosophila*, an alternative repressor-based network has since been modelled by Clark et.al. (2019) (**Fig. 1.7B**). However, this simple network alone is insufficient to explain the down-regulation of *Tc-odd* observed in *Tc-eve* and *Tc-run* knockdowns by Choe et.al. (2006). Regardless of which hypothesised network topology is correct, these three genes appear to be essential for driving cellular oscillations in the *Tribolium* segmentation clock. However, the mechanism by which these oscillations are co-ordinated across cells remains elusive.

More recent functional data from our lab has raised further questions over the exact nature of the *Tribolium* oscillatory network and even hints at an intraspecies divergence between 2 different lab strains. *Tc-odd* RNAi in the San Bernadino (SB) strain from our lab failed to capture the asegmental phenotype observed by Choe et.al. (2006) (**Fig. 1.8B**), who themselves used the Georgia (GA2) strain for their experiments. Rather than a complete collapse in segmentation that would be expected with the loss of cellular oscillations, *Tc-odd* RNAi in SB beetles generated consistent pair-rule-like phenotypes, in which every other segment was missing (**Fig. 1.8C**) (Sharma, pers comms). This

phenotype has also been published by Jeon et.al. (2020) for the GA2 strain, where they proposed a role for *Tc-odd* in regulating expression of the secondary pair-rule gene, *Tc-prd*, which itself is required for the formation of odd-numbered segments (Choe and Brown, 2007). However, these phenotypes were attributed to a weaker knockdown efficiency, and more severe asegmental cuticles were still observed. Comparisons of *Tc-odd* knockdowns between different strains in our lab also finds a higher penetrance of

A Activation based PR oscillator



B Repression based PR oscillator

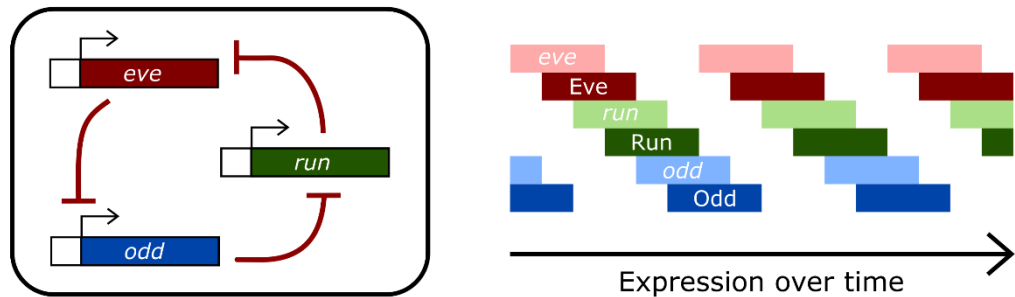


Figure 1.7. Proposed *Tribolium* *Tc-eve/Tc-odd/Tc-run* oscillator network models. (A) Gene circuit hypothesised by Choe et.al. (2006) that involves activator roles for *Tc-eve* and *Tc-run*. This feedback loop is sufficient to generate periodic oscillations and is consistent with expression patterns of *Tc-eve*, *Tc-odd*, and *Tc-run* in knockdown embryos (Choe et.al., 2006). (B) Cellular oscillations based on a repression based negative feedback loop proposed by Clark et.al. (2019). This model more accurately captures relative expression patterns *in situ* and is consistent with the repressive nature of these genes. However, taken on its own, this cannot explain observed expression patterns in knockdown embryos (Choe et.al., 2006). Temporal expression patterns are based on *in silico* models by Clark et.al. (2019) using Boolean logic with equal time delays for protein synthesis and decay (Clark, 2017).

severe segmentation phenotypes in GA2 cuticles compared to SB cuticles, indicating that there is indeed a real strain difference (Sharma pers comms).

To test if poor knockdown efficiency of *Tc-odd* was responsible for the lack of asegmental phenotypes seen in the SB strain, a *Tc-odd* knockout strain was developed using a CRISPR/Cas9 “knock in and replace” approach. A construct containing a GFP marker was inserted in place of part of the *Tc-odd* gene that contains the functional zinc finger coding domain (**Fig. 1.8E**). Its correct placement was confirmed through DNA sequencing and molecular analysis (Sharma pers comms). Interestingly, even in *Tc-odd* knockout embryos, 2 thoracic segments and 2 to 3 abdominal segments were often still patterned, reflecting a pair-rule-like phenotype, albeit slightly more severe than those observed in *Tc-odd* knockdowns (**Fig. 1.8D**). This suggests that, despite the complete loss of *Tc-odd*, there is still sufficient patterning activity to rescue almost half of the body segments in the SB strain.

Analysis of *Tc-eve* expression in *Tc-odd* null embryos shows ubiquitous distribution throughout the SAZ (**Fig. 1.9A**), consistent with that observed in *Tc-odd*^{RNAi} experiments by Choe et.al. (2006). In contrast, poorly resolved stripes of *Tc-run* can still be seen in the *Tc-odd*^{-/-} germbands (**Fig. 1.9A red arrowheads**), indicating that *Tc-run* oscillations are being maintained despite the loss of cyclical *Tc-odd* and *Tc-eve* expression (Sharma pers comms). This would suggest either that there are genes connected to the pair-rule oscillator that share a degree of redundancy with *Tc-odd* (**Fig. 1.9C**), or that *Tc-run* is taking regulatory input from an alternate oscillator network (**Fig. 1.9D**). The latter of these scenarios is consistent with the hypothetical model proposed by Clark et.al. (2019), which points towards direct regulation of *Tc-run* from a secondary *hairy/Hes* based oscillator (**Fig. 1.9D**). The auto-oscillatory behaviour of *h/Hes* genes in this model is also consistent with the preservation of *Tc-h* stripes observed in the SAZ of *Tc-odd*^{-/-} embryos, however, it is unclear how *Tc-eve* is overexpressed given the proposed repression by *Tc-run* (**Fig. 1.9A**). An alternative explanation is that other *Tribolium odd-skipped*-family genes confer partial redundancy with *Tc-odd* (**Fig. 1.9C**). Indeed, *Tc-bowel* (*Tc-bwl*) and *Tc-sister-of-bowel* (*Tc-sob*), are both expressed in similar patterns to *Tc-odd* in the SAZ during early stages of GBE (Thümecke and Schröder, 2021).

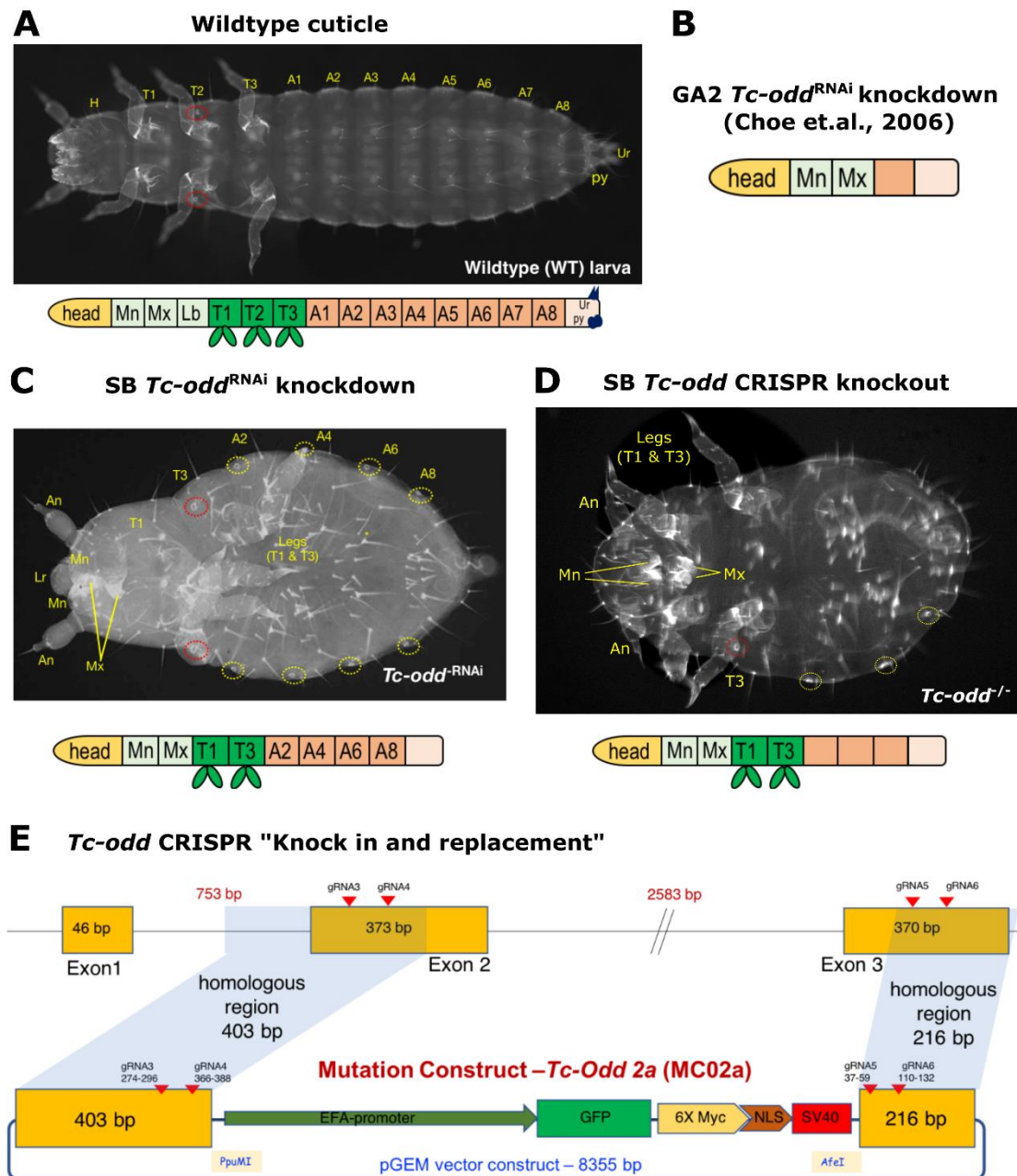


Figure 1.8. Weaker *Tc-odd*^{RNAi} segmentation phenotype in SB strain confirmed through CRISPR gene knockout. (A) Wildtype cuticle and diagram of a normal *Tribolium* body plan. (B) Diagram depicting asegmental *Tc-odd*^{RNAi} phenotype in GA2 strain reported by Choe et.al. (2006). (C) Pair-rule-like phenotype from *Tc-odd* knockdown in SB strain. (D) *Tc-odd* CRISPR knockout cuticle in which 2 thoracic segments and 3 abdominal segments have been patterned. (E) *Tc-odd* CRISPR knockout design. The construct includes a GFP gene which is expressed ubiquitously via an upstream *EFA* (*EF1-alpha*) promoter. Guide RNAs (gRNAs) were designed to target and replace a section of the gene, from exon 2 to exon 3, that encompasses the functional zinc-finger coding domain. Sequencing confirms the full construct has been inserted into the correct position as intended. Images and CRISPR schematic all provided by Sharma and Peel (pers comms)

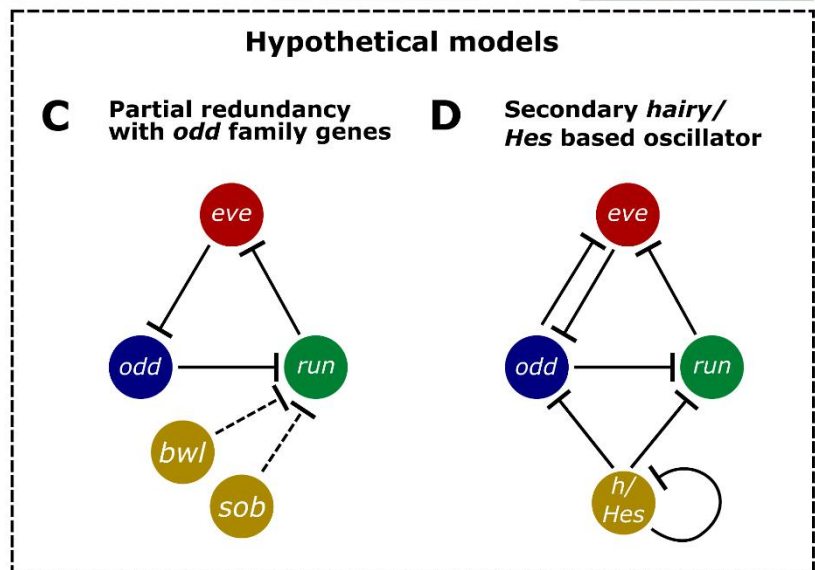
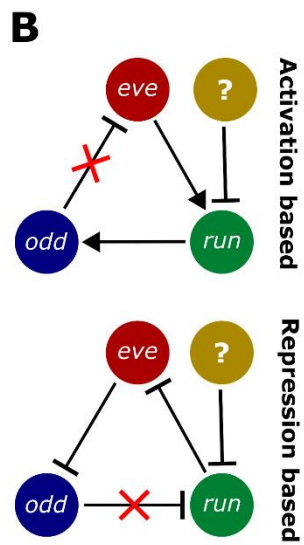
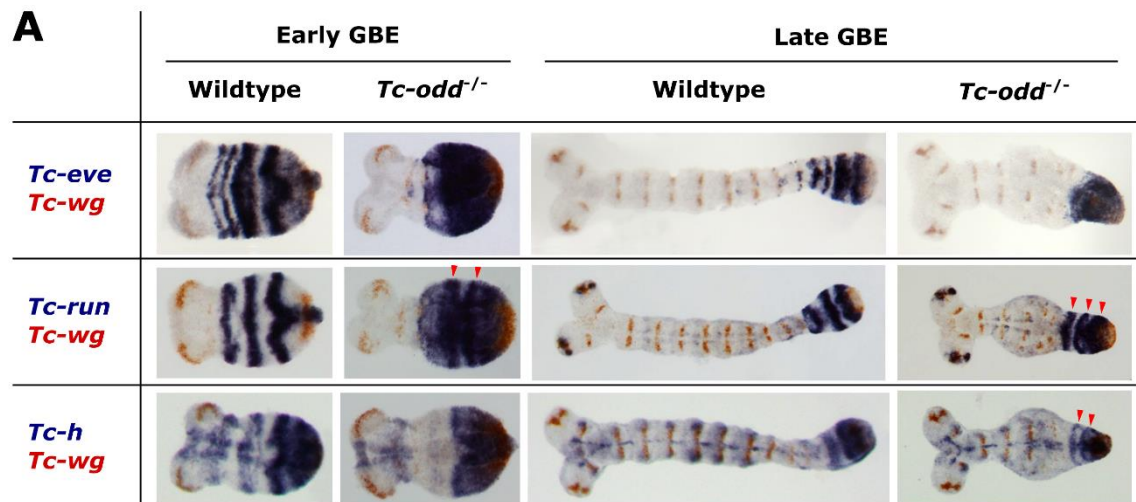


Figure 1.9. PR gene expression in *Tc-odd*-null SB embryos and hypothetical clock networks. (A) Colorimetric *in situ* hybridisations for *Tc-eve*, *Tc-run* and *Tc-h* (blue) alongside the segment polarity gene, *Tc-wg* (red), as a marker. Gene expression is shown for early and late stage germbands in wildtype and *Tc-odd*^{-/-} knockout mutants. Red arrowheads indicate poorly resolved *Tc-run*/*Tc-h* stripes in the SAZ of *Tc-odd*-null germbands. (B) activation- and repression-based *Tc-eve*/*Tc-odd*/*Tc-run* oscillator networks with the loss of *Tc-odd* indicated by the red cross. In either case, an unknown element (yellow) must be inhibiting *Tc-run* to create the stripes seen in (A). The repression-based oscillator is insufficient to explain over-expression of *Tc-eve* in *odd*^{-/-} embryos. (C-D) Hypothetical models to explain how *Tc-run* and *Tc-h* stripes are established in the absence of *Tc-odd*. (C) Repression-based *Tc-eve*/*Tc-odd*/*Tc-run* oscillator, with partially redundant *odd* family genes, *Tc-bwl* and/or *Tc-sob*, inhibiting *Tc-run*. (D) Network proposed by Clark (2017) and tested *in silico* to pattern segments. Built on the repression-based network, *Tc-odd* and *Tc-run* take regulatory input from a secondary oscillator, in this case an autoinhibitory *hairly/Hes* gene. The regulatory relationship between *Tc-odd* and *Tc-eve* here is consistent with over-expression of *Tc-eve* observed in the SAZ of *Tc-odd*^{-/-} germbands. *In situ* images provide by Sharma (pers comms)

1.7. The *Hes/Hey* family genes: Candidates for secondary oscillations in *Tribolium* segmentation

Hes/Hes-related with YRPW motif (Hey) genes encode a group of basic helix-loop-helix (bHLH) transcription factors that function as repressors (reviewed by Kageyama et.al., 2007). They are often described as effectors of Notch signalling; however, they are not core components of the pathway and are not essential for all Notch-dependent processes (Gazave et.al., 2014). Furthermore, there are several cases where *Hes/Hey* expression and/or function appears to be independent of Notch signalling (Carroll et.al., 1988; Leimeister et.al., 2000; Peel et.al., 2005). The relationship between *Hes/Hey* genes with Notch signalling was likely established at the dawn of Metazoan evolution, where together they share an ancient role in regulating neural differentiation (Gazave et.al., 2014; Layden and Martindale, 2014).

Several characteristics of *Hes* genes, such as their repressive nature and short degradation times, make them particularly suitable as genetic oscillators (Kageyama et.al., 2007), and their cyclical expression has been observed across multiple

developmental contexts. In mouse embryos, *Hes1* oscillations and Notch signalling are essential for the proper maintenance of neural progenitors during neurogenesis (Shimojo et.al., 2008), while Notch-coupled *Hes* oscillations also represent a core component of the vertebrate segmentation clock (**Section 1.5**) (Kageyama et.al., 2012). As previously discussed, there is evidence of a similar role for *Hes* genes and Notch signalling in the segmentation of some arthropods, although the precise network topology is still poorly understood, and roles for *Hes* and/or Notch signalling in segmentation appear to have been lost in several insect lineages (**Fig. 1.6**).

In *Tribolium*, two *Hes* genes, *Tc-h* and *Tc-dpn*, are expressed in dynamic stripes in the SAZ, seemingly independent of any core components of the Notch signalling pathway (Aranda et.al., 2008). While their knockdowns do not produce any obvious segmentation phenotypes, we cannot rule out redundancy with other *Hes/Hey* or segmentation genes (Aranda et.al., 2008). An example of this is seen in the zebrafish segmentation clock, where two distinct negative feedback loops, made up of multiple *Hes/Hes-related (Her)* genes, function redundantly with each other to create a robust clock network (Schröter et.al., 2012). Single, and even combinations of double, *Hes/Her* knockdowns in this system fail to generate segmentation defects, as one oscillator network compensates for the loss of the other. As HES/HEY proteins form dimers with either themselves (homodimers) or other bHLH factors (heterodimers) in order to repress transcription, the presence/absence of other dimerization partners can also significantly affect the dynamics of HES/HEY activity. For example, in the zebrafish clock network, some of the HES/HER dimers that form do not bind DNA and therefore do not directly repress transcription, however, the formation of these dimers still plays an important functional role in sequestering monomers and regulating their repressive capacity (Schröter et.al., 2012). In this system, a broad posterior-to-anterior gradient of *Hes6* expression controls oscillation period across the PSM through this principle. Similarly important protein-protein interaction are observed in chick neurogenesis, where ID proteins sequester HES1 proteins through the formation of ID/HES1 heterodimers (Bai et.al., 2007). Due to the lack of a basic DNA binding domain in ID proteins, this inhibits the normal autorepressive activity of HES1, and ultimately releases it from cyclical expression (Bai et.al., 2007). These detailed studies of *Hes/Hey*

network topologies and protein-protein dynamics in vertebrates, demonstrate the potential underlying complexity of oscillator circuits, and highlight the need for an exhaustive screen of *Hes/Hey* gene redundancy in the *Tribolium* segmentation clock. The nature of *Hes/Hey* genes as autonomous oscillators across multiple developmental contexts, along with their functional roles in other sequentially segmented arthropods, makes these genes key candidates for a secondary oscillator in *Tribolium* (**Fig. 1.9D**).

1.8. *odd-skipped* family genes: Candidates for partial redundancy with *Tc-odd*

The primary pair-rule gene, *odd*, is related to several genes that encode a family of C2H2 zinc finger transcription factors. First identified in *Drosophila*, these include *bowel* (*bwl*), *sister-of-bowel* (*sob*), and *drumstick* (*drm*) (Wang and Coulter, 1996; Hart et.al., 1996; Green et.al., 2002). In *Drosophila*, roles for these genes have been described in several different developmental contexts, including leg joint patterning and hindgut specification (Hao et.al., 2003; Iwaki et.al., 2001). *Odd* is the only member of this family so far shown to be involved in body segmentation, however, *sob* and *drm* are also both expressed in similar pair-rule stripes during segment patterning in *Drosophila* (Hart et.al., 1996; Green et.al., 2002). In *Tribolium*, similar functions of *odd* family genes have also been identified in leg joint formation and hindgut specification (Angelini et.al., 2012; Thümecke and Schröder., 2021). During leg patterning, *Tc-sob* and *Tc-drm* are both expressed in rings around the borders of each prospective leg segment and are thought to share partial or border-specific redundancy that rescues alternate leg segments in *Tc-drm* knockdown embryos (Thümecke and Schröder, 2021). However, this has not been explicitly proven experimentally. Striped expression of *Tc-sob* and *Tc-bwl* in the SAZ during *Tribolium* segmentation closely resembles that of *Tc-odd* and therefore also hints at possible redundancy in the context of body segmentation (Thümecke and Schröder, 2021).

In *Drosophila*, the zinc finger (ZF) domains of Odd, Sob, and Bwl proteins exhibit a particularly high degree of sequence conservation (Green et.al., 2002). Comparison of the *Tribolium* Odd family amino acid sequences also reveals a high level of sequence similarity among the ZF regions of Tc-Odd, Tc-Sob and Tc-Bwl (**Fig. 1.10**). Each protein

has 5 C2H2 ZFs, the first 4 of which share an almost identical amino acid sequence (**Fig. 1.10**). Given the functional specificity that the ZF domain is likely to confer, this similarity, along with that of *Tc-sob*, *Tc-bwl* and *Tc-odd* expression patterns, makes the *odd*-family genes highly suitable candidates for partial redundancy with *Tc-odd* in the segmentation clock network (**Fig. 1.9C**).

Tc-Odd	PKRQFI	CKYCN	RQFTK	SYNLL	IHERTH	-TDERPY	SCD	ICGKA	FRRQD	HLLR
Tc-Bwl	PKKQFI	CKFCN	RQFTK	SYNLL	IHERTH	-TDERPY	SCD	ICGKA	FRRQD	HLLR
Tc-Sob	PKKQFI	CKFCN	RQFTK	SYNLL	IHERTH	-TDERPY	SCD	ICGKA	FRRQD	HLLR
Tc-Drm	PKYEFI	CKYCQ	RKFTK	SYNLM	IHERTH	KSPELT	FSCEV	CGRGF	KRQDN	LK
Tc-Odd	DHRYI	HSKEK	PFKCT	ECGKG	FCQSRT	LAVHK	ILHME	ESPHK	CPVCS	RSFN
Tc-Bwl	DHRYI	HSKEK	PFKCT	ECGKG	FCQSRT	LAVHK	ILHME	ESPHK	CPVCS	RSFN
Tc-Sob	DHRYI	HSKEK	PFKCT	ECGKG	FCQSRT	LAVHK	ILHME	ESPHK	CPVCS	RSFN
Tc-Drm	QHRTI	H-----	-----	-----	-----	-----	-----	-----	-----	-----
Tc-Odd	QRSNL	KTHLL	THT	E	-RPLE	CTLCA	QFFAS	YSDL	KTHEA	RHC
Tc-Bwl	QRSNL	KTHLL	THT	D	HKPYE	CNSCG	KVFRN	CDLRR	HALTH	A
Tc-Sob	QRSNL	KTHLL	THT	D	IKPYN	CNSAC	GKVFRR	NCDLRR	HSLTH	N
Tc-Drm	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Figure 1.10. Highly conserved zinc finger domain among *Tribolium* Odd family proteins.

Amino acid sequence alignment of Tc-Odd, Tc-Bwl, Tc-Sob and Tc-Drm, displaying only the conserved domain encompassing C2H2 zinc fingers (blue). Dotted red box highlights a near identical region of the first 4 zinc fingers of Tc-Odd, Tc-Bwl and Tc-Sob. Only 1 residue (in red) is not conserved along this sequence between these 3 proteins.

1.9. Aims of this thesis

In this thesis, I aim to address some of the outstanding questions regarding the molecular mechanisms that underpin sequential segmentation, using *Tribolium castaneum* as a model. Largely driven by recent observations in our lab, along with the inconsistencies observed among sequentially segmenting arthropods, I set out to investigate possible roles of *Hes/Hey* and *odd* family genes in the *Tribolium* segmentation clock network. I hypothesised that redundancy between genes within or across these families may explain some of the divergences that we see in the *Tribolium* segmentation network. Specifically, I aimed to:

1. Perform comprehensive phylogenetic analyses of *Hes* and *odd* family genes in order to better understand their evolutionary history, and to provide additional context for subsequent experimental results and cross-species comparisons (Chapter 3).
2. Screen the expression patterns of *Hes* and *odd* family genes to confirm previously published data, and to identify/rule-out additional candidate genes for which expression patterns have not yet been described in *Tribolium* (Chapters 4 and 5).
3. Use multiplex fluorescent *in situ* hybridisation to analyse the relative positions and overlap of candidate gene domains during segment addition (Chapters 4 and 5).
4. Assess gene functions and redundancy with regards to segmentation, using combinations of single and double knockdowns in wildtype and *Tc-odd*-null beetles (Chapter 6).

With these objectives, I hope to better understand the underlying complexity of the *Tribolium* segmentation clock network and how primary segment patterning can be partially rescued in the absence of *Tc-odd* function in the SB strain. A deeper understanding of this process in *Tribolium* can help us to interpret results from studies in other sequentially segmenting arthropods and to address how segment patterning mechanisms have diverged during arthropod evolution.

2. MATERIALS AND METHODS

2.1. Phylogenetics analysis

Hes/Hey and *odd-skipped*-family genes were identified for arthropod species using NCBI BLASTP searches with known amino acid sequences as queries (Altschul et.al., 1990). Conserved domains were identified for each sequence using the NCBI conserved domain search tool (Lu et.al., 2020). Nucleotide coding sequences for corresponding proteins were taken from NCBI genbank and saved to an online workbench, Benchling (<https://www.benchling.com/>). Here, sequences were annotated with their domains and other details obtained from the NCBI and Ensembl genome browsers, such as intron and exon structures, and genome locations. Multiple sequence alignments (MSA) were built in MEGA X (Kumar et.al., 2018) and alignments were made using MUSCLE (codon) and MUSCLE for nucleotide and amino acid sequences, respectively (Edgar, 2004).

Bayesian phylogenetic analyses were performed using BEAST 2 and its associated packages (Bouckaert et.al., 2014). A model averaging approach was employed using bModelTest (Bouckaert and Drummond, 2017) and OBAMA (Bouckaert, 2020) packages for nucleotide and amino acid MSAs, respectively. All analyses were performed using relaxed clock models and a yule tree prior. For each analysis, 4 parallel Monte-Carlo Markov chains (MCMC) were run using different starting seeds in order to check for adequate convergence and mixing. MCMC outputs were checked using TRACER (Rambaut et.al., 2018), and runs were stopped once combined log scores for each parameter reached an effective sample size (ESS) of at least 200. Using BEAST associated programmes, LogCombiner and TreeAnnotator, tree files were combined and used for generating a consensus maximum clade credibility tree, discarding the first 10% of trees as burn-in. The final consensus trees were formatted and annotated with posterior branch values in FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>), while figures were finalized in the vector graphics editor, Inkscape (<https://inkscape.org/>).

2.2. *Tribolium* husbandry and embryo fixation

Tribolium stock keeping and fixation protocols closely follow those described in The Beetle Book (Bucher, 2009). *Tribolium* San Bernadino (SB) and Georgia (GA2) strains were kept at 30°C on a 500 µm sieved stock keeping mix consisting of: 20 parts organic wholemeal flour (Dove Farm Foods), 1 part yeast and 0.1 part diluted fumagillin-B (12 g fumagillin-B (Medivet Pharmaceuticals; DIN 02231180) with 200 g pre-sieved organic white flour). For egg collections, adult beetles were separated from the flour mix using a 700 µm metal sieve and moved to 250 µm sieved organic plain white flour (Dove Farm Foods) where they were left for 48 hours at 30°C. Eggs were then separated using a 300 µm sieve and transferred to fresh stock keeping mix where they were left to hatch and grow for a period of 30 to 40 days until eclosion. Newly eclosed adult beetles were then placed on to fresh stock keeping mix. Adults were left for at least 7 days after eclosion until they were used for subsequent egg collections. This was to allow time for the onset of female egg production.

To fix embryos, eggs from 48 hour collections were placed in baskets (0.1 mm mesh in 50 mm falcon tube lid). For dechoriation, eggs were first rinsed with ddH₂O to remove flour, before being washed twice with 50% thin bleach (1:1 ddH₂O : Morrisons thin bleach) for 90 seconds. Eggs were then subjected to 5 successive rinses with ddH₂O to ensure all the bleach was washed away. Using a paintbrush, eggs were then transferred to glass vials containing 6 ml heptane, 2 ml PEMS (Bucher, 2009), and 300 µl formaldehyde solution 37%, where they were fixed for 30 minutes on a mixer at 230 RPM. Once finished, the aqueous phase of the fixation mix was removed using a pipette, and embryos were devitellinized by adding 8 ml methanol and shaking vigorously for ~30 seconds. Eggs that were successfully devitellinized fell to the bottom of the glass vials and were transferred to separate Eppendorf tubes. In order to devitellinize the remaining embryos at the interphase, they were drawn up and rapidly forced out through a 0.8 mm (19G) needle attached to a syringe. Fallen embryos were transferred from the bottom and the process was repeated multiple times until no more embryos could be gathered. Fixed embryos were rinsed 8 times in methanol to remove as much debris and egg casings as possible before being stored at -20°C.

2.3. Gene fragment amplification and cloning

Tribolium candidate gene sequences were obtained from the NCBI RefSeq genome assembly and annotation (Tcas5.2). Sequences were saved to the online workbench, Benchling, which was used for designing primers. Primers were selected to amplify fragments for use as *in situ* probes and non-overlapping double-stranded RNA (dsRNA) fragments (**Table 2.1**). Template cDNA for PCR reactions was generated using the SuperScript™ IV First-Strand Synthesis System (Invitrogen, #18091050) on total RNA solutions from multiple stages provided by Rahul Sharma. Separate reactions with random hexamers and Oligo(dT) primers were performed and the cDNA from each was tested using positive control PCRs. Fragments were amplified by mixing primers with Roche PCR nucleotide mix (#11814362001) and Taq DNA polymerase (#11418432001), and running through a touchdown PCR programme.

PCR products were purified using a QIAquick gel extraction kit (QIAGEN, #28706) and ligated into the pGEM®-T Easy Vector with the kit provided (Promega, #A1360). 5 µl of the ligation reaction product was then incubated with 45 µl competent *E. coli* cells (provided by Rahul Sharma) on ice for 30 minutes. Plasmids were transformed into cells by heat shock at 42°C for 45 seconds and returned to ice. Transformed *E. coli* cells were then incubated in LB media at 37°C before being spread on LB Agar plates inoculated with ampicillin and coated in X Gal and IPTG. Blue white colony screening was used to identify cells transformed with ligated plasmids. Positive colonies were cultured in 100 ml LB media overnight at 37°C on a shaker (180 RPM). Plasmids were then extracted from the liquid cultures using the QIAGEN plasmid midiprep kit (#12145). Final plasmid inserts were confirmed with restriction digests using cut sites identified in Benchling. The pGEM®-T Easy Vector and gene fragment sequences were used as for *in silico* restriction digests in Benchling, predicting fragment sizes for forward and reverse facing inserts. These predictions were subsequently compared to gel bands from the real restriction digests to confirm presence and orientation of each fragment. 4 different digests were performed with different restriction enzymes to confirm results for each plasmid. *Tc-hairy*, *Tc-deadpan*, *Tc-odd*, *Tc-wg*, and *GFP* fragments were cloned by Andrew Peel and Rahul Sharma (**Table 2.1**) and were confirmed with DNA sequencing.

Table 2.1. Primers used for probe and dsRNA fragments. *Tc-hairy*, *Tc-deadpan*, *Tc-odd-skipped*, and *GFP* fragments were cloned by Rahul Sharma (RS) and Andrew Peel (AP).

Gene	NCBI Accession number	Fragment		Primers (5'-3')
		Use	Length /bp	
<i>Tc-emc</i>	XM_964922.4	Probe	914	F: CCTCATTGACCATTCGTTTCGC R: CGAAACCGTGACCCCAACAATT
		5' dsRNA	316	F: CTCGTAGTGCATCAGTGTTTCAG R: GCCTCCTGCGACTTCTTCTG
		3' dsRNA	544	F: CAGCCCCAACACCATCCTCTC R: CGAAACCGTGACCCCAACAATT
<i>Tc-Sidpn1</i>	XM_970094.2	Probe	798	F: CAGTTCGCCTCGTCCTCTC R: GCTCAAGTCTTGAAAAAGTGTGACT
		5' dsRNA	372	F: GCCTCTAGTATGACCGCCAAGA R: GACGCAGTTGTCCAGGTGC
		3' dsRNA	273	F: CGACGCTCGAAGAAGTCAACAA R: AGTTAAATCACCACGGCCTCCA
<i>Tc-Sidpn2</i>	XM_008200042.2	Probe	497	F: AGTGACTACCATGAAAAGTGAAACCG R: ACACTTGAATATCACAGCACGAGAC
<i>Tc-E(spl)1</i>	XM_967592.3	Probe	457	F: CTACACCCACTGCGCCAAC R: CAAGCCTTAATCACAGGTACCAACA
<i>Tc-E(spl)3</i>	XM_967400.4	Probe	284	F: CCGGTTTTCACGCAGTGTG R: TCACACTCAGGTACGAAATGCG
<i>Tc-Hey</i>	XM_008192597.2	Probe	1,058	F: GACCATGCCCTGTGCGAGAATA R: TAAACCTCCACTCCCGTGCAA
<i>Tc-cwo</i>	XM_008199468.2	Probe	1,395	F: GTGTTGTCGGTCGGTATCAGT R: CCTCGTCATTCTCGCTCCTCAA
<i>Tc-hairy</i> * (Cloned by RS)	NM_001114293.1	Probe		Provided by RS, primers unknown
		5' dsRNA	459	F: TGTTGCCTACCAAGATGGTTAC R: TCCTTCTGGCCCTGGTT
		3' dsRNA	445	F: CCCAATAACATTATTCTGGGCAAC R: ACCAACTGGAAGACATGAATGA
<i>Tc-deadpan</i> * (Cloned by RS)	XM_962601.3	Probe		Provided by RS, primers unknown
		5' dsRNA	555	F: ACAATGTCGAAAGCCGAACTA R: GGTATCAATTGGATGCTTTGTGG
		3' dsRNA	503	F: GGGAAATCGCCTTGCTTT R: ATTCGACTCCCGAGAAGTG
<i>Tc-wingless</i> * (Cloned by AP)	NM_001114350.1	Probe	1,095	F: CAACGCCAGAAGCAAGAAC R: ACGACTTCCTGGGTACGATA
<i>Tc-odd-skipped</i> (Cloned by AP)	XM_008198532.2	5' dsRNA	343	F: GCTGGACACCGAACCGAT R: CTTGCAGATGAATTGTCGCTTG
<i>GFP</i> (Cloned by RS)		dsRNA	704	F: ATGGGTAGTAAAGGAGAAGAACTTT R: TCCATGCCATGTGTAATCCC

2.4. *in situ* hybridisation probe synthesis

DIG and FITC labelled probes were prepared according to Kosman et.al. (2004; Supporting material). pGEM®-T Easy plasmids containing probe inserts were first linearised using high-fidelity SpeI or NcoI restriction enzymes (New England BioLabs), based on the orientation of the fragment. In the case that inserts had SpeI or NcoI cut sites, SalI or SacII restriction enzymes were used instead, respectively. After incubation, a small sample of the reaction mix was checked on a gel to ensure all plasmids were cleanly cut, after which the plasmids were isolated through phenol/chloroform extractions. The plasmids were then precipitated using 2x volume of ethanol and 0.1x volume 3M sodium acetate (pH 5.2) and left at -20°C overnight. Precipitate was spun down and washed with 70% ethanol before being air dried and resuspended in nuclease free ddH₂O.

RNA probes were synthesised by transcription reactions using Roche T7 (#10881775001) or SP6 (#11487671001) polymerases and DIG (#11277073910) or FITC (#11685619910) labelled NTP mixes. After 2 hours at 37°C, reactions were diluted with RNase-free ddH₂O and 1 µl was kept for gel analysis. Reactions were stopped using carbonate buffer and STOP solution, after which the RNA was precipitated with 0.1x volume lithium chloride and stored overnight at -20°C. RNA quality was checked on a gel (looking for a strong band of approximately the right size). Precipitated RNA was then centrifuged and washed with 70% ethanol, before being resuspended in 200 µl Hyb A solution (Bucher, 2009). Aliquots of RNA probes were stored at -20°C until use. For genes where expression patterns have not previously been published, sense strand RNA probes were also synthesised for use in control experiments.

2.5. Double-stranded RNA synthesis

Plasmids containing fragments for use as dsRNA were linearised as described in section 2.5. Antisense and sense single-stranded RNA (ssRNA) was synthesised using T7 and SP6 MEGAscript™ Transcription Kits (Invitrogen, #AM1334 and #AM1330). A small sample was kept for gel analysis before stopping the reactions and precipitating the RNA with lithium chloride at -20°C. The precipitate was centrifuged, and pellets were

washed with 70% ethanol before being air dried and redissolved in RNase-free H₂O. ssRNA concentrations were measured using a Nanodrop™ spectrophotometer and diluted accordingly. Antisense and sense ssRNA was mixed at a 1:1 molarity ratio in an RNase-free Eppendorf. The tube was then placed in 800 ml H₂O at 95°C and left to cool, allowing the strands to anneal. The final dsRNA product was checked for quality on a gel alongside the ssRNA, before being aliquoted and stored at -20°C until use.

2.6. Colorimetric *in situ* hybridisation

Double whole-mount *in situ* hybridisations were performed as described by Schinko et.al. (2009). The NBT/BCIP (Roche, #11681451001) colour reaction was used to develop staining for the DIG labelled probes, while INT/BCIP (Roche, #11681460001) was used for developing the FITC labelled gene (*Tc-wg*). Staining times varied and the amount of probe used was adjusted in any subsequent experiments where staining was too weak or strong. After the experiment, embryos were stored in 50% glycerol in PBT in the fridge. Before mounting, yolk cells were removed from germbands using an eye lash fashioned to the end of a pipette tip. This implement enabled gentle brushing of embryos without damaging them. Germbands were then transferred to a new slide and mounted under a coverslip in 50% glycerol. Stained embryos were imaged on a Leica M165FC Fluorescence Stereo Microscope using a Q Imaging Retiga EXI colour cooled fluorescence camera operated in Q Capture Pro 7 software.

2.7. Hybridisation Chain Reaction (HCR) multiplex *in situ* hybridisation

Third generation (Choi et.al., 2018) HCR™ probes were ordered from, and designed by, Molecular Instruments (<https://www.molecularinstruments.com>), based off the NCBI accession numbers provided (**Table 2.2**). For *Tribolium odd-skipped*-family genes, instructions were given to avoid the inclusion of the zinc-finger coding domains that share high levels of similarity. HCR probes for *Tc-wg* were provided by the Akam lab (University of Cambridge), and were also produced by Molecular Instruments. Experiments followed the adapted HCR protocol by Bruce et.al. (2021) with the

exception that dextran sulphate concentrations were reduced in the hybridisation and amplification buffers from 10% to 5% (volume). This alteration was inspired by *Tribolium* HCR results from Tidswell et.al. (2021) and served to reduce viscosity and improve germband retention during washes. Due to imaging limitations, only 3 sets of HCRTM amplifiers were used for each experiment. Also ordered from Molecular Instruments, these consisted of hairpins labelled with fluorophores, 514, 546, and 647. After amplification, embryos were stained for ~40 mins in 1 µg/ml DAPI in order to visualise whole embryos. Embryos were stored in DTG (DABCO/Tris/Glycerol; Kosman et.al., 2004) mounting solution and kept in the dark at 4°C.

Prior to mounting in DTG, embryos were cleaned of yolk cells as described in section 2.6. An effort was made to avoid prolonged exposure to light as to reduce fluorophore degradation. Mounted embryos were imaged using the Zeiss LSM880 + Airyscan Upright Confocal Microscope in the Leeds University Bio-imaging and flow cytometry facility. Argon 514 nm, DPSS 561 nm and He-Ne 633 nm lasers were used to image mRNA tagged with 514, 546 and 647 fluorophores, respectively. Images were captured using ZEN imaging software at 40x magnification. For larger embryos, tile scans with 10% overlap were captured to cover the whole embryo, and were stitched together in Zen. Z stack images with a 2.5 µm step size were taken through the entire embryo in order to capture gene expression across different tissues. Post-image processing was performed in the open-source platform, Fiji (image J) (Schindelin, 2012). This was used for maximum intensity projections of Z stacks, adjusting contrast/brightness, changing lookup tables, and merging channels.

Table 2.2. HCR probes. For each gene, NCBI accession numbers were provided to Molecular Instruments as reference. Initiators for each probe are as shown. 5' and 3' ends of the gene sequences excluded from probe designs are provided. These relate to the highly similar *odd*-family zinc finger coding domains.

Gene	NCBI accession number	Initiator(s)	Sequence to avoid (5'... 3')
<i>Tc-odd-skipped</i>	XM_008198532.2	B1, B2, B3	CCCAAGCGACAATTCATCTG... ...CACTTGTTGACCCACACCGA
<i>Tc-bowel</i>	XM_008198533.2	B4	CCCAAGAAACAGTTCATTTG... ...CACCTCCTGACCCACACCGA
<i>Tc-sister-of-bowel</i>	XM_008198531.2	B2	CCGAAAAAACAGTTTATTTG... ...CATCTGCTCACCCATACCGA
<i>Tc-drumstick</i>	XM_966887.3	B3	CCGAAGTACGAATTCATCTG... ...AACAGCACAGGACTATACAC
<i>Tc-hairy</i>	NM_001114293.1	B1	N/A
<i>Tc-deadpan</i>	XM_962601.3	B4	N/A
<i>Tc-emc</i>	XM_964922.4	B2	N/A
<i>Tc-wingless</i> (from Akam lab)	NM_001114350.1	B1, B3, B4	N/A

2.8. Parental and embryonic RNAi

Adult parental RNAi (pRNAi) was performed as described in Posnien et.al. (2009). Needles were made from borosilicate glass capillaries (Harvard Apparatus, #30-0019) using a Narishige needle puller (model PD-5). Pulled needles were then cut and sharpened using a Narishige needle grinder (model EG-45). Needles were filled using a syringe by drawing up dsRNA solution from a RNase-free Eppendorf lid. For genes without published functional data, two separate groups of females were injected with non-overlapping 5' and 3' dsRNA fragments to control for off target effects. For each experiment, a negative control group was also injected with *GFP* dsRNA. Beetles were subdued prior to injections by cooling on ice and were then placed on double sided tape ventral side down. Injections were made in the dorsal abdomen beneath the elytra, taking care to avoid the ventral nervous system, genitals, or gut. Pressure was steadily exerted using a syringe, and stopped once the internal pressure had caused the head to extend and the ovipositor to protrude. After injections, female beetles were

returned to stock keeping flour mix, where they were left to recover for 48 hours. Injected females were then mixed with an equal number of males from the same starting population. Egg collections were taken every 2 days, with 1 day recovery on stock keeping flour mix between.

Embryonic RNAi (eRNAi) was performed as described in Clark and Peel (2018). Eggs (1 to 3 hours after egg lay) were gently washed in 5% thin bleach for 90 seconds. They were then rinsed 5 times in ddH₂O and lined up along the edge of a microscope slide with a fine paintbrush. Eggs were angled with the posterior pole facing towards the capillary holder so that injections could be made in this region perpendicular to the egg axis. Needles were backfilled with dsRNA using MicroloaderTM microcapillary pipette tips (Eppendorf). Injections were carried out on a Zeiss Axiovert 10 inverted microscope using a continuous flow set up. Eggs were filled to the point where they became visibly clearer, and care was taken to avoid bursting them. After the injections, slides with the injected eggs were placed on damp blue roll inside petri dishes. This ensured that there was sufficient humidity for eggs to develop. Petri dishes left at 30°C until hatching (~4 days).

2.9. Cuticle preparation and imaging

Eggs from RNAi experiments were left to develop at 30°C and as larvae hatched they were transferred to a droplet (~40 µl) of 1:1 Hoyer's medium (Stern and Sucena, 2012) : lactic acid mix. Up to 20 larvae were placed in each droplet on a microscope slide, at which point a coverslip was placed over the top and they were left overnight at 60°C to clear. Unhatched eggs were given 6 days from egg lay to allow time for developmental delays. Those from pRNAi experiments were then rinsed and washed in 5% bleach for 90 seconds. For any eggs containing clearly visible embryos, a fine tip brush was used to carefully remove them from the egg casing. Dissected embryos were then treated in the same way as the hatched larvae. The remaining unhatched eggs were mounted whole, and a small amount of pressure was placed on the coverslip to burst them and push out any embryos that may be inside. Cuticles were imaged under GFP light settings on a Leica M165FC Fluorescence Stereo Microscope with a Q Imaging Retiga EXI colour cooled fluorescence camera and Q Capture Pro 7 software.

3. PHYLOGENETIC ANALYSES OF *HES/HEY* AND *ODD* FAMILY GENES

3.1. INTRODUCTION

3.1.1. The evolution of genes and multigene families

Understanding the evolutionary history of genes is an important step in learning how gene functions and developmental systems have diversified over time. One of the key contributors to molecular evolution is the process of gene duplication. There are several different mechanisms by which genes can be duplicated, which range from localised events that produce tandem copies along the same chromosome, to polyploidization, where entire chromosomes and genomes are duplicated (reviewed in Reams and Roth, 2015). The resulting gene copies from a duplication event are referred to as paralogues (Ramos and Ferrier, 2012).

If an entire gene and its regulatory domain is duplicated, this will often result in two daughter genes that are functionally redundant with one another (Birchler and Yang, 2022), in which case, nullifying mutations in one of the duplicates will often be of no consequence to the fitness of the organism, and the new gene is simply lost through genetic drift. Alternatively, paralogues can escape deletion if their functions diverge and develop important unique roles. This can occur either through subfunctionalisation, whereby the ancestral function of the original gene is divided between the two daughter genes, or neofunctionalization, in which a completely novel function is acquired by one of the paralogues (Stolzfus, 1999; Force et.al., 1999). These processes have the capacity to produce diverse multigene families, where functionality can be shared, transferred, gained, and lost over the course of evolution. In the context of developmental biology, these changes in gene functionality can contribute to the divergence of homologous developmental systems in a process termed developmental system drift (True and Haag, 2001).

3.1.2. Concerted evolution of gene families

When trying to infer the relationships between genes across species it is important to consider the phenomenon of concerted evolution, whereby a lower level of divergence between two or more paralogues is observed than would be expected based on the molecular clock hypothesis (Teshima and Innan, 2004). This primarily occurs due to a process called gene conversion, where lengths of DNA sequence are unilaterally copied from one paralogous gene to another, usually as a by-product of recombinational repair mechanisms (**Fig. 3.1A**) (Hastings, 2010). Gene conversion rates tend to be higher when the distance between paralogues is shorter (Fawcett and Innan, 2011), therefore evidence for this phenomenon is commonly found among closely linked genes and gene complexes (Bettencourt and Feder, 2002; Noonan et.al., 2004; Peel et.al., 2006).

The occurrence of gene conversion events can be problematic when trying to infer the evolutionary history of gene paralogues, as this can result in gene trees that overestimate the frequency of gene duplications (Teshima and Innan, 2004) (**Fig. 3.1B**). When investigating the relationships between genes, especially those which have maintained close physical linkage within the genome, it is therefore important to consider the possibility of gene conversion events.

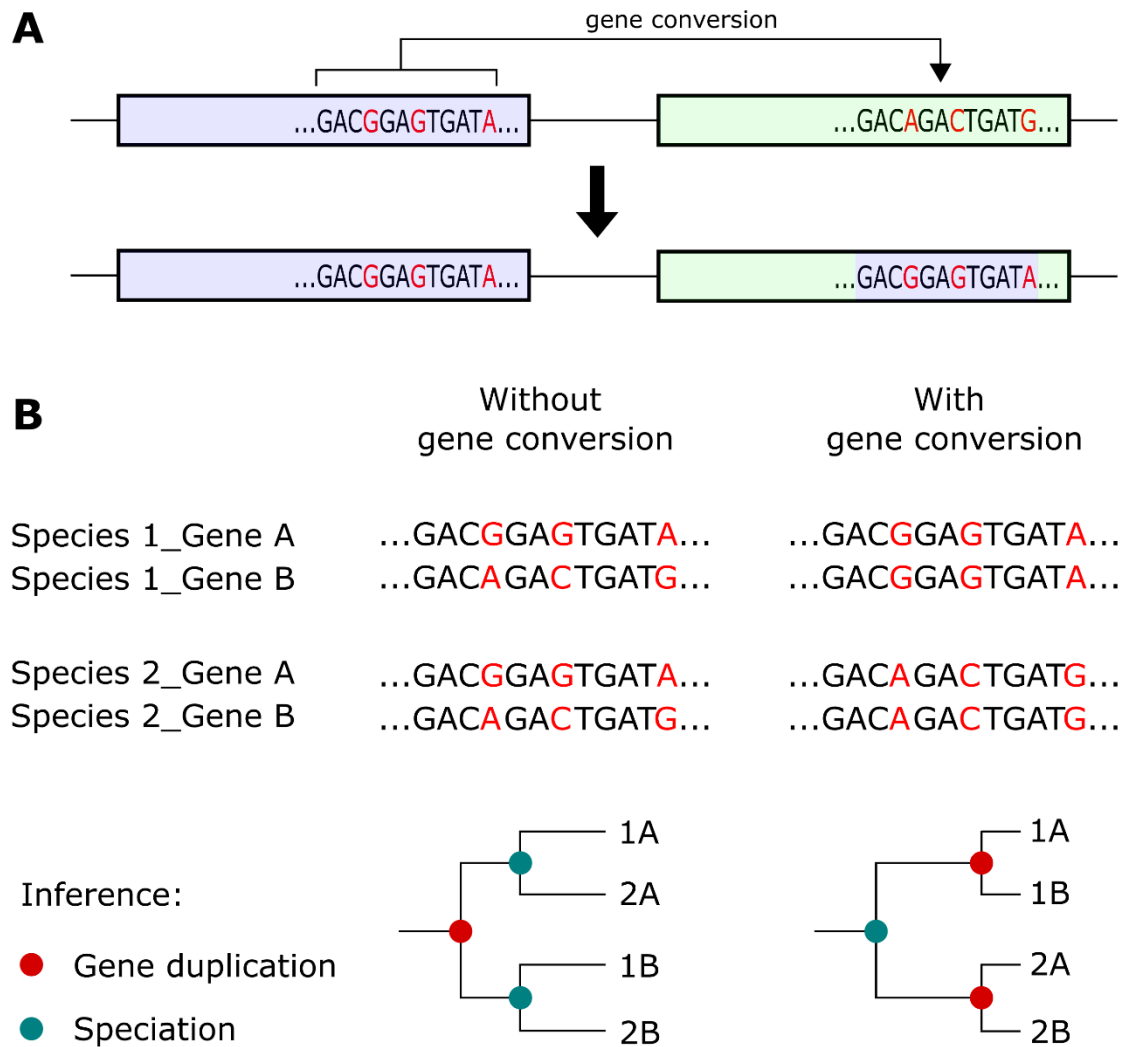


Figure 3.1. Gene conversion can lead to inaccurate phylogenetic inferences. (A) Simple gene conversion model between two adjacent paralogous genes. A fragment is unilaterally copied over from one gene to the other. (B) Different scenarios for two orthologous genes, A and B, across two species, and the gene trees that would be generated from the sequences in each case. Without gene conversion, the tree accurately reflects the evolutionary history of genes A and B, which originated from a single duplication. In a scenario where the genes have undergone conversion events in each species the tree overestimates the number of duplications. Figure adapted from Fawcett and Innan (2011)

3.1.3. Current understanding of *Hes/Hey* gene family evolution

Hes/Hey family genes have been identified in species spanning the entire metazoan kingdom, while no homologous sequences are known to exist in plant, fungi, microbes, or protists (Zhou et.al., 2012). They are likely derived from a bHLH gene that evolved sequence encoding the Orange domain in an ancient metazoan ancestor (Zhou et.al., 2012). The bHLH and Orange domains are characteristic of almost all HES/HEY factors, and are essential for selecting dimer partners and binding DNA (Dawson et.al., 1995; Kageyama et.al., 2007). Phylogenetic analysis of the *Hes/Hey* related genes has identified several distinct classes with representative homologues from all major metazoan lineages, pointing towards an early expansion of these genes during the emergence of the Metazoa (Gazave et.al., 2014). The pan-metazoan classes described in this analysis include *Hey*, *Helt*, and a large clade containing all remaining *Hes* genes. Additional lineage specific classes designated *Stich* (*cwo*) and *Dec* were also identified for protostomes and chordates respectively (Gazave et.al., 2014). Various duplication events in different metazoan lineages are responsible for further expansions of the *Hes* gene family, creating a complex picture of its evolution (Zhou et.al., 2012).

Phylogenetic analyses of *Hes* family genes in arthropods support distinct clades for *Hey* and *cwo* (*Stich*) class genes, for which all species assessed have just one or two copies of the respective orthologues (Duncan and Dearden, 2010; Dearden, 2015). Most of the remaining arthropod *Hes* genes fit into clades based on *Drosophila* *Hes* genes, *hairy/deadpan*, *Similar to deadpan* (*Sidpn*), and *Enhancer-of-split* (*E(spl)*) (Dearden, 2015). Note that several articles refer to *Similar to deadpan* as *Side* (Bitra et.al., 2009; Bitra and Pali, 2010; Duncan and Dearden, 2010; Dearden, 2015), which is instead the symbol for an unrelated gene, *sidestep*. For clarity, I will use *Sidpn* as shorthand for this gene and its orthologues, consistent with online databases such as NCBI and flybase.

The absence of any clear *E(spl)* orthologues in chelicerate species, suggests that these genes arose in the lineage leading to myriapods, crustaceans and insects. In pancrustacea, *E(spl)* genes have formed a conserved genomically linked complex alongside *bearded* class genes (Duncan and Dearden, 2010; Dearden, 2015). How exactly the remaining insect *Hes* genes, *h*, *dpn*, and *sidpn*, are related to *Hes* genes in

non-insect arthropods is still unclear. As this group includes genes found to be involved in segment patterning across arthropods, this knowledge will be of particular value when considering the evolution of this developmental system.

In *Tribolium*, eight *Hes* genes have been identified and named based on their inferred relationships to *Drosophila* *Hes* genes (Bitra et.al., 2009). These include *Tc-hairy*, *Tc-deadpan*, *Tc-Sidpn1*, *Tc-Sidpn2*, *Tc-Hey1*, *Tc-Hey2*, *Tc-E(spl)1* and *Tc-E(spl)3*. These classifications have mostly been corroborated by separate molecular analyses, with the one correction that *Tc-Hey2* is instead a *cwo* orthologue (Duncan and Dearden, 2010; Kux et.al., 2013). *Tc-Sidpn2* also appears to be no more closely related to *Sidpn* genes, as it is to *h* or *dpn* genes. While not strictly a *Hes* gene, due to its lack of an orange domain, *Tc-extramacrochaete* (*Tc-emc*) forms a sister clade to *Hes* family members in a gene tree of *Tribolium* and *Drosophila* bHLH transcription factors (Bitra et.al., 2009).

3.1.4. Current understanding of *odd-skipped* family evolution

Odd-skipped (*odd*) and its paralogues, *bowel* (*bwl*), *sister-of-bowel* (*sob*), and *drumstick* (*drm*), were first characterised in *Drosophila* (Coulter et.al., 1990; Wang and Coulter, 1996; Hart et.al., 1996; Green et.al., 2002). The amino acid sequences reveal conserved C₂H₂ zinc finger (ZF) regions, with a particularly high level of similarity between Odd, Sob and Bwl (Green et.al., 2002). Bwl and Sob contain five of these zinc fingers, while Odd has four, and Drm has two. Outside of these ZF domains, the amino acid sequences of the *Drosophila* Odd family proteins share little sequence similarity (Green et.al., 2002). Direct orthologues of *Drosophila* *odd*-family genes in other holometabolous insects, *Tribolium* and *Nasonia vitripennis*, have been confirmed through small phylogenetic analyses presented alongside comparative functional studies (Angelini et.al., 2009; Rosenberg et.al., 2014; Thümecke and Schröder, 2021).

Three *odd-related* genes have also been identified in the spider, *Cupiennius salei*, and centipede, *Strigamia maritima*, however their relationship to insect *odd* family genes has yet to be distinguished (Damen et.al., 2005; Green and Akam, 2013 (Fig. S1)). In each case, *odd-related* genes from *Cupiennius* and *Strigamia* appear most closely

related to those from the same species, suggesting that they are the product of independent gene duplications in chelicerate and myriapod lineages. In the case of *Cupiennius*, only partial amino acid sequences, comprised primarily of the ZF domain, were used for the phylogenetic analysis (Damen et.al., 2005).

To date, phylogenetic analyses of arthropod *odd* family genes have consisted of small datasets, primarily concerned with identifying orthologues of *Drosophila odd* family genes as part of candidate gene studies. As such, there is currently a lack of comprehensive analyses that can provide us with a good understanding of the evolutionary history of this gene family in arthropods.

3.1.5 Aims of the chapter

In this chapter, I will present arthropod-wide phylogenetic analyses of the *Hes/Hey* and *odd* family genes using both nucleotide and amino acid datasets. For the *Hes/Hey* analysis, I first aimed to confirm whether my results support the broad classifications of *Hes/Hey* genes that have previously been published, before attempting to resolve the relationships between *non-E(spl)* *Hes* genes, particularly those with reported roles in segmentation. For the *odd* family genes, I aimed to establish the first comprehensive phylogenetic analysis of these genes across arthropods, and to glean insights into their evolutionary history.

For each gene family, I investigate gene structures and genomic organisation to search for additional clues regarding their relationships and possible functional ties. Signs of gene conversion in *odd* family genes led me to further evaluate the prevalence of this phenomenon in these genes across arthropod evolution.

3.2. RESULTS

3.2.1. Phylogenetic analysis of the arthropod *Hes/Hey* family genes

Nucleotide and amino acid based phylogenetic analyses support previously published relationships between *Hes/Hey* genes

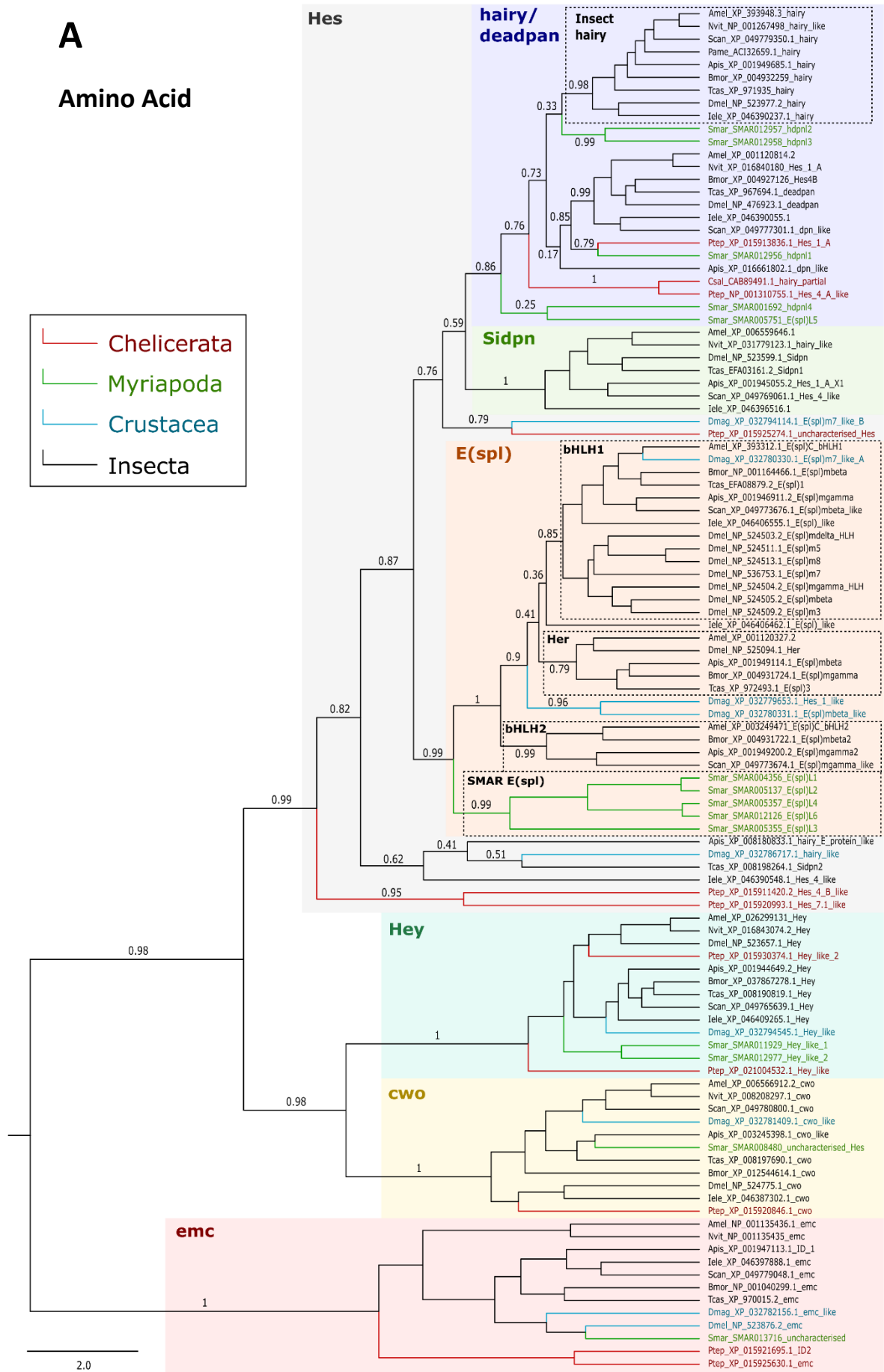
To investigate the evolutionary history of the *Hes/Hey* family gene in arthropods, I first selected a range of arthropod species to represent a wide coverage of the phylum, and to include several model developmental organisms. Species with high quality genome assemblies and annotations were favoured, and *Hes/Hey* related genes were identified by BLASTP searches using known genes from other species. For two species, the spider, *Cuppiennius*, and the cockroach, *Periplaneta americana*, only one sequence, taken from segmentation studies, was included (Stollewerk et.al., 2003; Pueyo et.al., 2008). This is due to limited genome annotations being available for these species.

Initial phylogenetics analyses using both amino acid and nucleotide (codon) sequence alignments, show strong support for *emc* genes as a separate group to the *Hes/Hey* family genes, while *Hey*, *cwo*, and the remaining *Hes* genes form their own well supported clades (**Fig. 1**). This is consistent with metazoan-wide analyses by Gazave et.al. (2014) and Zhou et.al. (2012). Most arthropod species used in these analyses have single copies of *emc*, *Hey* and *cwo* genes, with a few exceptions such as the spider, *Parasteatoda tepidariorum*, and centipede, *Strigamia*, both of which have two *Hey* orthologues.

The remaining *Hes* genes can be divided into several groups with strong posterior support in both trees. This includes a clade containing all myriapod and insect *E(spl)* orthologues, consistent with previously published trees (Duncan and Dearden, 2020; Dearden, 2015). Furthermore, the nucleotide analysis replicates the tree topology and classifications assigned to the arthropod *E(spl)* genes by Dearden (2015) (**Fig. 3.2B**). The amino acid analysis gives a similar, but slightly less clear picture of the relationship between these genes (**Fig. 3.2A**).

A

Amino Acid



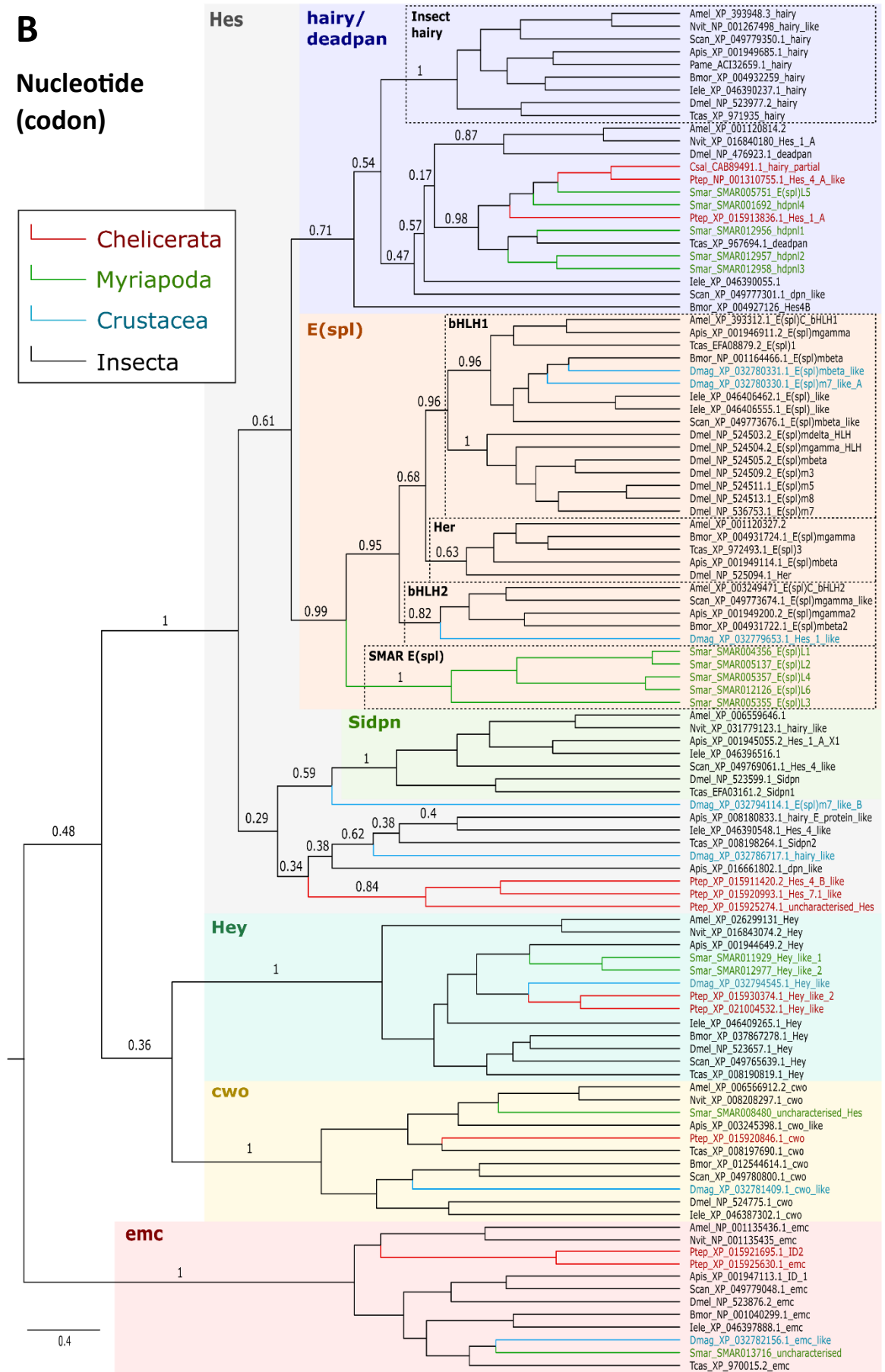
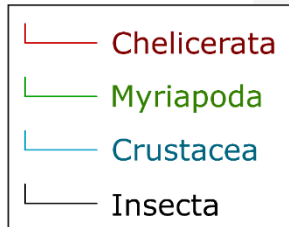
B**Nucleotide
(codon)**

Figure 3.2. Arthropod *Hes/Hey* family gene trees. Based on amino acid (A) and nucleotide (codon)(B) sequence alignments. Phylograms generated by mixed model Bayesian inference. Branch labels indicate posterior probabilities. Well supported clades, reproduced in both analyses, are highlighted and labelled accordingly. Sequence identifiers include species, protein accession number, and gene names (based on NCBI/Ensembl annotations). Species: Ptep, *Parasteatoda tepidariorum*; Csal, *Cupiennius salei*; Smar, *Strigamia maritima*; Dmag, *Daphnia magna*; Iele, *Ischnura elegans*; Scan, *Schistocerca cancellata*; Pame, *Periplaneta americana*; Apis, *Acyrtosiphon pisum*; Nvit, *Nasonia vitripennis*; Amel, *Apis mellifera*; Tcas, *Tribolium castaneum*; Bmor, *Bombyx mori*; Dmel, *Drosophila melanogaster*.

Within the *Hes* superclade, there is a well-supported clade of insect genes related to *Drosophila Sidpn* (**Fig. 3.2**). All insect species examined have a single gene within this clade, except for *Bombyx mori*, where no orthologues could be identified. While there is strong support for this clade from both analyses, its position among the wider *Hes* family genes changes, with the amino acid tree placing it alongside *hairy/deadpan* genes (**Fig 3.2A**), while the nucleotide-based tree places *E(spl)* genes as the sister group to *h/dpn* genes instead (**Fig. 3.2B**).

There is consensus from both gene trees for a clade that contains insect *h* and *dpn* orthologues, alongside several chelicerate and myriapod genes (**Fig. 3.2**). Interestingly, no genes from the crustacean, *Daphnia magna*, are in this group, suggesting that any *hairy/deadpan* related genes in this species have either been lost, or have significantly diverged. In both trees, insect *hairy* orthologues form their own clade with high posterior probabilities. However, the relationships between the remaining *hairy/deadpan* genes are less clear. Consensus clades for a number of arthropod *Hes* genes, including *Tc-Sidpn2*, could also not be established from these trees.

These trees highlight several inaccuracies in the names assigned to genes. To avoid confusion in subsequent analyses, I will rename *Strigamia E(spl)L5* to *hdpnL5*, and *Daphnia E(spl)m7-like-B* to *Hes-1* (as originally named by Eriksson et.al. (2013). Accession numbers will remain unchanged.

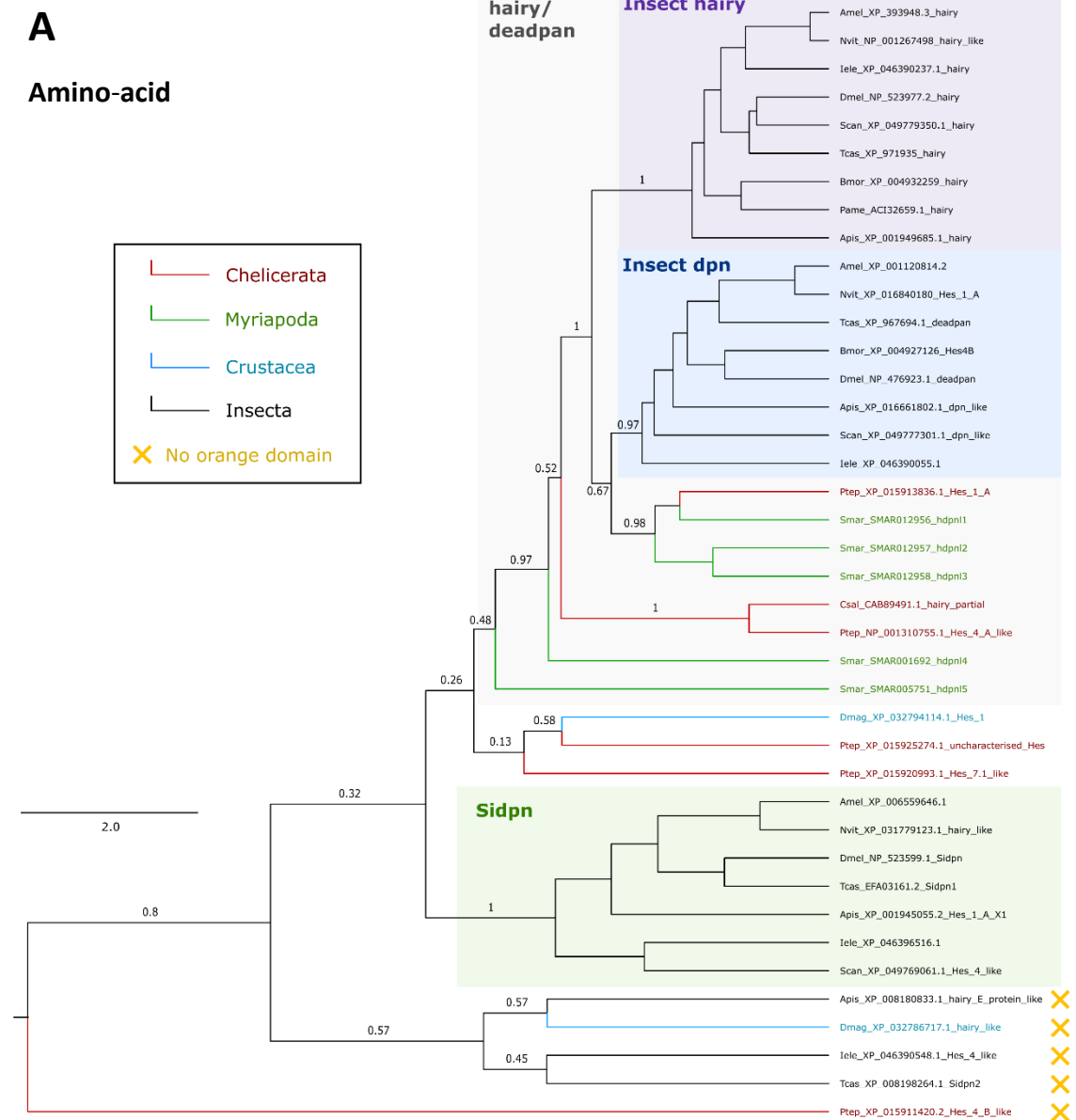
Phylogenetic analyses of non-*E(spl)* *Hes* genes fails to provide a consensus for the timing and number of duplications of *hairy/deadpan-related* genes in arthropods

To better understand the evolutionary history of non-*E(spl)* *Hes* genes in arthropods, I performed a smaller analysis of the same dataset, excluding *emc*, *Hey*, *cwo*, and *E(spl)* orthologues. By removing divergent sequences and reducing the size of the dataset, this should result in more accurate sequence alignments (Thompson et.al., 2011), which in turn will generate more accurate trees.

As in the initial analyses, the same group of insect, chelicerate and myriapod *h/dpn* genes form a clade with consensus from both amino acid and nucleotide sequences (**Fig. 3.3**). The exact topology of this clade varies between analyses, however, both support a close relationship of chelicerate and myriapod *hairy/deadpan-like* genes with insect *deadpan* genes, while insect *hairy* orthologues form their own branch (**Fig. 3.3**). All insects and the spider, *Parasteatoda*, have 2 *h/dpn* genes, while *Strigamia* has 5, suggesting there have been several additional rounds of duplication during myriapod evolution. Consistent with this, *Strigamia hdpn1*, *hdpn2*, and *hdpn3* are all found side by side in the genome within 10-20kb of each other, indicating that these are most likely paralogues. The other 2 genes, *hdpn4* and *hdpn5*, are separated, however, due to the lack of a chromosome level assembly, we cannot rule out that they are still nearby. The *h/dpn* genes of *Parasteatoda* and insects are also found along the same chromosome, and in some cases (*Parasteatoda*, *Schistocerca*, *Apis*, *Nasonia*) are within several hundred kb of each other (data not shown). While there does not appear to be a consensus on the number and timings of *h/dpn* gene duplications, both analyses place insect *h* orthologues away from the *h/dpn*-related genes of other arthropods, which themselves group more readily with insect *dpn* orthologues (**Fig. 3.3**). This might suggest that insect *dpn* genes are more representative of the ancestral state of *h/dpn* genes retained by chelicerates and myriapods, and that insect *h* orthologues have since diverged early in insect evolution.

These analyses again confirm a well-supported clade of insect *Sidpn* orthologues. However, a number of arthropod *Hes* genes still fail to form high confidence clades with support from both trees. NCBI conserved domain searches against these sequences show that several, are missing orange domains, including *Tribolium Tc-Sidpn2* (**Fig. 3.3**

Orange crosses). An alignment of these genes against the orange consensus sequence from other *Hes* genes does, however, reveal that most have retained some conserved residues (**Fig. 3.4**), suggesting that despite having diverged, it is possible that these domains may still offer functionality. There is no evidence of any orange-like domain in *Tc-Sidpn2*, suggesting that this region has been lost entirely from the gene.



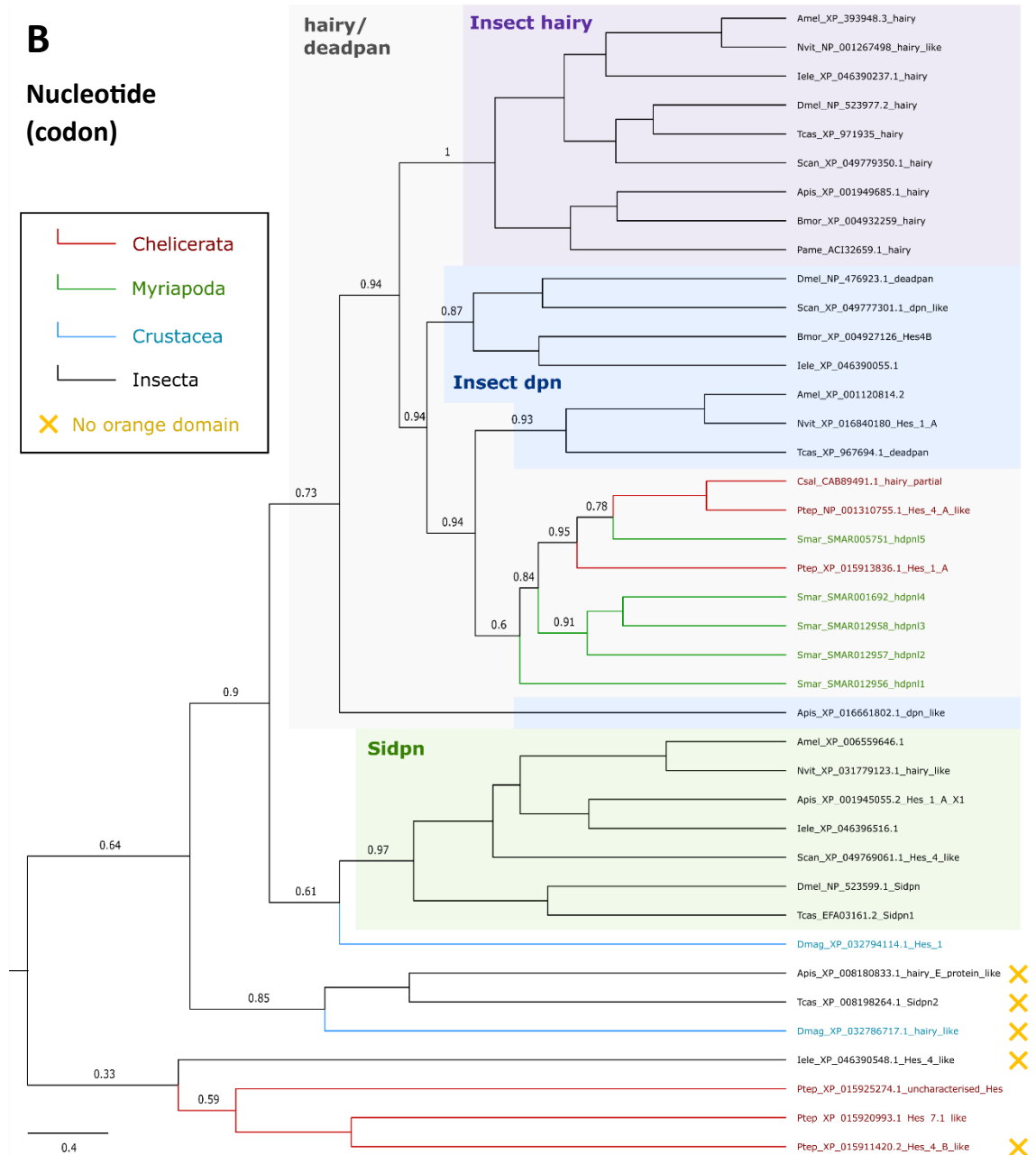


Figure 3.3. Arthropod non-*E(spl)* Hes gene trees. Based on amino acid (A) and nucleotide (codon)(B) sequence alignments. Phylograms generated by mixed model Bayesian inference. Branch labels indicate posterior probabilities. Well supported clades, reproduced in both analyses, are highlighted and labelled accordingly. Sequence identifiers as previously described (**Fig. 3.2**). Note name changes for Smar_E(spl)L5 and Dmag_E(spl)m7-like-B to Smar_hdpnl5 and Dmag_Hes-1 respectively.

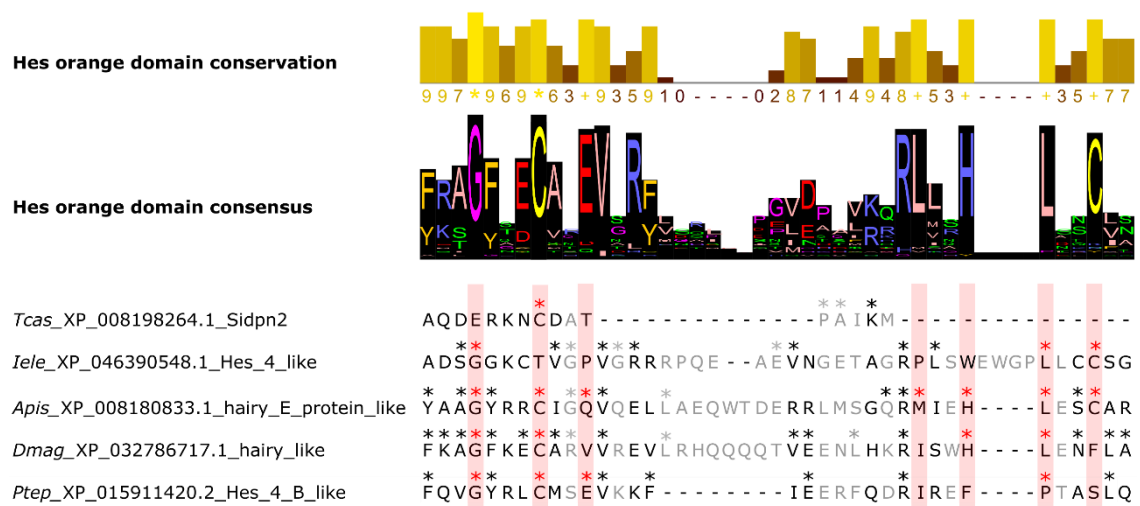


Figure 3.4. Signs of conserved residues in *Hes* genes that lack orange domains. Orange domain conservation values and consensus residues were obtained from a multiple sequence alignment (MSA) of all non-*E(spl)* *Hes* genes for which an Orange domain was identified in NCBI (**Appendix A, Fig. A.1**). An additional MSA, which included *Hes* genes without identifiable Orange domains, was performed and the alignments of these genes are shown below. Conservation values are a measure of conserved physico-chemical properties for each position in the Orange domain MSA, based on Livingstone and Barton (1993). “+” indicates a site where all sequences have residues that share the same physico-chemical properties, while “*” indicates that all sequences have the same amino acid in this position. These are therefore the most highly conserved residues along the Orange domain, and likely functionally important. As such, these sites are highlighted red in the alignment below. The consensus orange sequence is represented by a logo, where overall stack heights represent sequence conservation and individual letter heights reflect the relative frequency of each amino acid. Asterisks mark residues that match the most common amino acids from the consensus logo. Letters are greyed out at positions of low Orange sequence conservation (<5)

3.2.2 Arthropod *hairy/deadpan* genes exhibit a conserved intron/exon structure

In addition to sequence-based phylogenetic analyses, gene intron/exon structures can also inform us of the relationships between genes, as well as providing insights into possible evolutionary constraints affecting intron losses or gains (reviewed in Betts et.al., 2001; and Heyn et.al., 2015). Comparison of non-*E(spl)* *Hes* genes reveals a conserved intron structure across arthropods, with the highest degree of conservation visible in *hairy* and *deadpan* related genes (**Fig. 3.5**). By plotting intron positions against

the amino acid MSA, I identified several homologous introns that are found at precisely the same insert across different genes (**Fig. 3.5 Coloured introns**). 3 homologous introns are found in *Hes* genes from the spider, *Parasteatoda*, and the centipede, *Strigamia* (**Fig. 3.5C-D. Red, green, orange**), while only introns 1 and 2 are also found in insects and the crustacean, *Daphnia*. Given that *Parasteatoda* and *Strigamia* both have at least 1 *hairy/deadpan* related gene with homologous introns 1 and 2, and another with introns 1, 2 and 3, it could be deduced that those with the same number of homologous introns between the two species are orthologues. However, the gene trees do not support this, and instead finds the *Strigamia* genes with 2 introns (*hdpn1-3*) to be more closely related to the *Parasteatoda* gene with 3 introns (*Hes-1-A*) (**Fig. 2A**). This might instead suggest that the ancestral *hairy/deadpan*-like gene(s) had all 3 introns, and the third of these has been independently lost in genes from both lineages. If true, this third intron appears to have also been lost from all non-*E(spl)* *Hes* genes in the crustacean and insect species included in this analysis.

Introns 1 and 2 (**Fig. 3.5 Red, green**) are conserved throughout arthropods, and are seen in almost all *Hes* genes. In most insect *hairy* and *deadpan* orthologues, these comprise the only introns, except for a few independent intron gains in *Ischnura elegans*, *Schistocerca cancellate*, and *Bombyx mori hairy/deadpan* (**Fig. 3.5A-B**). Furthermore, the exon between homologous introns 1 and 2 is exactly 96 bp in length in all *hairy/deadpan* genes, including those in *Strigamia* and *Parasteatoda*. While this is more variable across the remaining *Hes* genes. This conservation of intron/exon structure only goes as far as the positions of introns, and no conserved patterns of intron length were found between any of the *Hes* genes included in this analysis. In contrast to *hairy* and *deadpan*, insect *Sidpn* genes exhibit a large diversity in intron numbers and positions, with conserved introns 1 and 2 being lost in orthologues from holometabolous insects, *Nasonia*, *Tribolium* and *Drosophila* (**Fig. 3.5E**). There also appears to be multiple independent intron gains in these genes in several lineages.

Annotations for *Acyrtosiphon pisum dpn-like* and *Ischnura Hes-4-like* begin at positions that immediately follow conserved introns in other arthropod *Hes* genes (**Fig. 3.5B,F. Asterisks**), suggesting that upstream exons are either missing from their annotation/assembly, or that they have been lost during evolution. Manual searches of

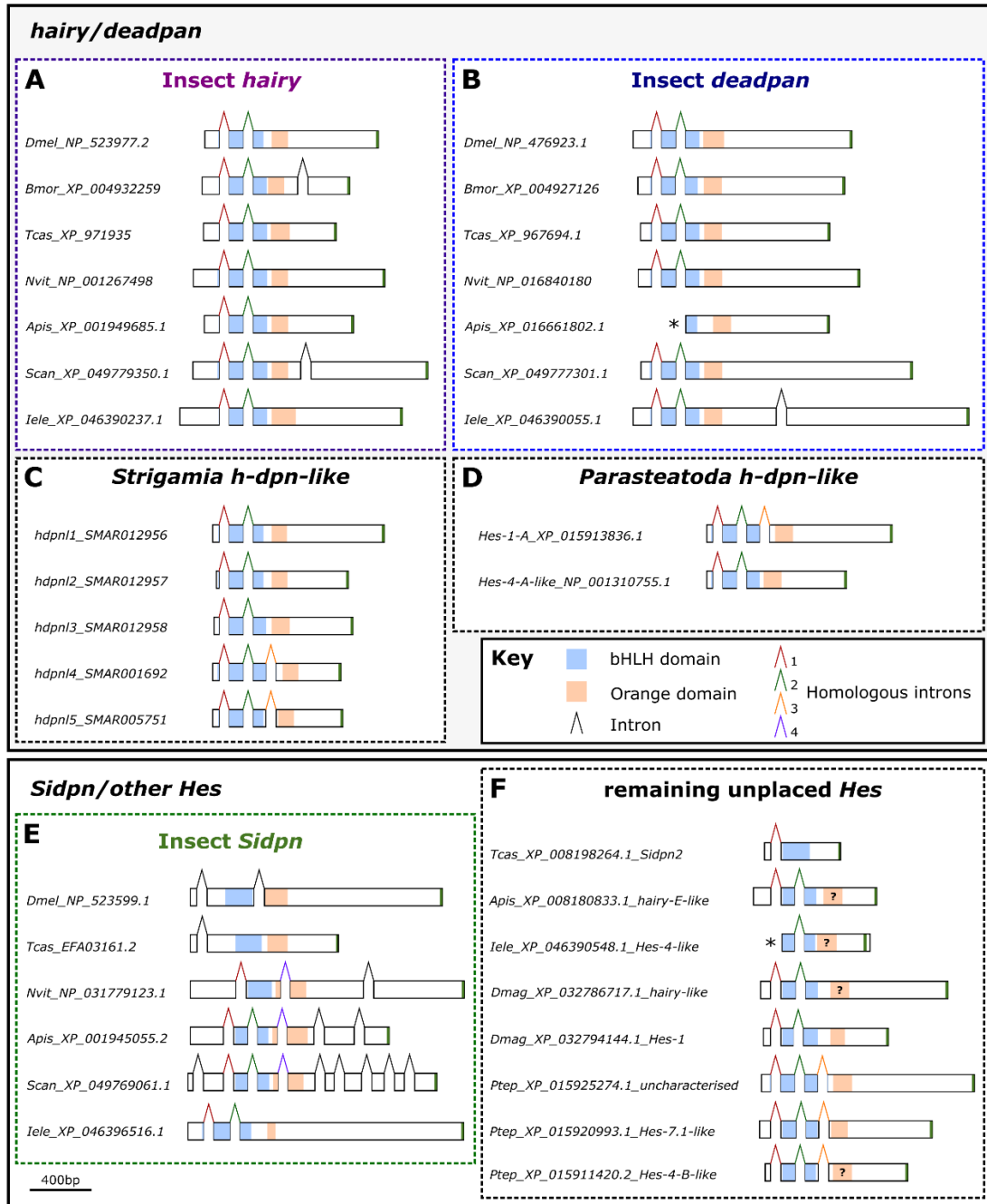


Figure 3.5. Gene structure comparison of non-*E(spl)* arthropod *Hes* genes. Groupings of genes are based on previous phylogenetic analyses (Fig. 3.3), broadly split into *hairy/deadpan* genes (A-D) and the remaining non-*E(spl)* *Hes* genes (E-F). Introns are indicated by gaps. Homologous introns were identified from their positions along the amino acid MSA and are colour coded (numbered 1 to 4 in the key). Asterisks mark genes that appear to be missing upstream exons. “?” indicates unconfirmed orange domains (Fig. 3.4). Species identifiers are the same as those described in Fig. 3.2 legend.

the upstream DNA of these genes failed to find any sign of missing exons. Many of the unplaced *Hes* genes also share conserved intron structures with *hairy/deadpan* genes (**Fig. 3.5F**), however, without support from phylogenetic analysis, this is not sufficient evidence to confidently infer their relationships to other *Hes* genes

3.2.3. Phylogenetic analysis of arthropod *odd* family genes

To investigate the evolutionary history of the *odd* family gene in arthropods, a similar range of species were used as in the *Hes* family analysis, with a few changes to the selection of insects. All species chosen, with the exception of *Oncopeltus fasciatus* (*odd* and *sob* sequences taken from Auman and Chipman (2018)), have well-annotated genome assemblies based on sequenced transcriptomes, and in most cases are chromosome mapped. *Odd* family paralogues were identified for each species based on protein BLAST queries. Four *odd* family genes were identified in all insect species used, while only three were found in crustacean (*Daphnia*), myriapod (*Strigamia*), and chelicerate (*Parasteatoda*) genomes. Accession numbers and gene names are based on NCBI genbank records, except for *Strigamia*, which relates to the Ensembl database.

Gene trees support the radiation of *odd-family* genes in insects, with independent duplications having occurred in myriapods and chelicerates

Phylogenetic analysis based on amino acid sequences produces several well supported clades that group together insect and crustacean genes alongside *Drosophila odd*, *bwl*, *sob*, and *drm* (**Fig. 3.6A**). *Strigamia* and *Parasteatoda odd*-related genes form their own branch that appears most closely related to *sob* genes. Two *odd-like* genes from the aphid, *Aphis gossypii*, produce a separate clade that is distantly branched from other arthropod *odd* genes. Finally, *drm* genes appear to have diverged significantly from other arthropod *odd* family genes, and orthologues could only be identified for crustacean and insect species. In several cases, *odd* family genes from the same species grouped together (**Fig. 3.6A asterisks**), such as *Oncopeltus odd* and *sob*, *Ischnura odd*-

like and *bwl-like*, and *Daphnia bwl-like 1 and 2*, as well as *Strigamia* and *Parasteatoda odd-related* genes.

Phylogenetic analysis based on nucleotide sequences aligned by codons, produced a similar looking tree, but in which *Strigamia* and *Parasteatoda* genes are grouped among insect *sob* genes, supporting a closer relationship between these classes. However, posterior support for this new clade was relatively low (0.55). In the nucleotide tree, there was additional grouping of genes from the same species (**Fig. 3.6 asterisks**), including *Chrysoperla carnea odd* and *sob*, and *Ischnura sob-like* with *odd-like* and *bwl-like*. Furthermore, the nucleotide analysis gave stronger statistical support to these branches, with almost all intra-species clades scoring posterior values of 1.

In both trees, *Daphnia odd* family genes (excluding *drm*) are positioned among insect *bwl* genes, while centipede and spider sequences appear more closely related to insect *sob* genes. Based purely on these phylograms, it could be inferred that an ancestral *bwl* or *sob* like gene has duplicated independently in the lineages leading to chelicerates, myriapods, crustaceans, and insects. Indeed, only one *odd-related* gene could be identified in genomes of arthropod sister phyla, Onychophora (*Epiperipatus broadwayi*) and Tardigrada (*Ramazzottius varleornatus*), supporting the existence of only one *odd-related* gene in the shared ancestor of arthropods. A duplication event likely occurred in the common ancestor of insects and crustaceans to produce *drm*, which itself has diverged rapidly. A further duplication in crustaceans could explain the two closely related *bwl-like* genes seen in *Daphnia*, while duplications in the lineage leading to insects could be responsible for the radiation of *odd*, *sob* and *bwl* genes.

This deduction, however, fails to explain why *odd* family genes for several insects were also found to group together by species, rather than by gene class. This is especially true for non-basal species such as *Oncopeltus* and *Chrysoperla*, which, if we were to interpret the gene grouping as the result of recent lineage-specific duplications, would create a particularly complex picture of *odd* family evolution in insects. Furthermore, the likelihood of so many independent gene duplications occurring in different arthropod taxa seems remote. To better understand the results from these initial phylogenetic analyses, I took a closer look at some of the genes from the same species that have grouped together.

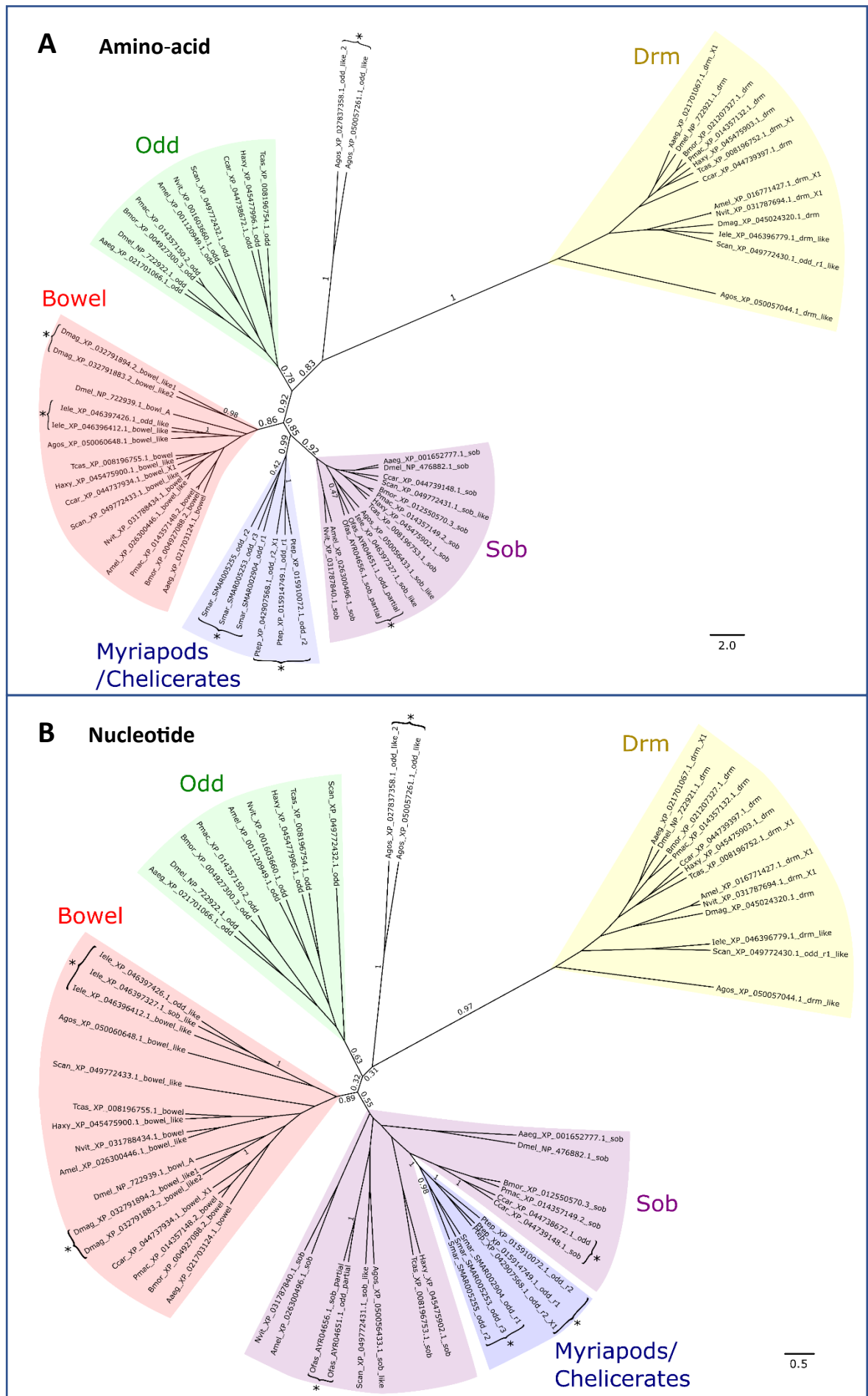


Figure 3.6. Arthropod *odd* family gene trees. Based on amino acid (A) and nucleotide codon (B) alignments. Phylograms generated by mixed model Bayesian inference. Branch labels indicate posterior probabilities. Well supported clades, reproduced in both analyses, are highlighted and labelled based on grouping with already-distinguished insect *odd* family genes, *odd*, *bwl*, *sob* and *drm*. Myriapod and Chelicerate species form a separate clade labelled as such. Curly brackets and asterisks mark genes from the same species that have grouped together. Sequences identifiers include species, protein accession number, and gene names (based on existing annotations). Species: Ptep, *Parasteatoda tepidariorum*; Smar, *Strigamia maritima*; Dmag, *Daphnia magna*; Iele, *Ischnura elegans*; Scan, *Schistocerca cancellata*; Agos, *Aphis gossypii*; Nvit, *Nasonia vitripennis*; Amel, *Apis mellifera*; Ccar, *Chrysoperla carnea*; Tcas, *Tribolium castaneum*; Haxy, *Harmonia axyridis*; Bmor, *Bombyx mori*; Pmac, *Papilio machaon*; Aaeg, *Aedes aegypti*, Dmel, *Drosophila melanogaster*.

The zinc finger coding region exhibits a high degree of sequence similarity among genes from the same species

Alignments of genes from the same species reveal stretches of ZF coding sequence that are highly conserved at the nucleotide level. This varies among species, but in general, those that clustered together in the gene trees, have higher sequence similarities. In some cases, such as *Parasteatoda*, *Ischnura* and *Oncopeltus*, there are regions between 270 bp and 340 bp in length that share over 95% sequence similarity to other *odd* family genes in those species (**Fig. 3.7**). This is despite sharing little similarity outside of the ZF encoding region, suggesting that this is unlikely to be the result of recent gene duplications. In each case, it appears to be the ZF coding region that bears this unusually high level of conservation. One explanation for this observation is that the ZF domains of *odd* family genes are under strong stabilising selection, which while may be true, would still allow for a certain number of synonymous substitutions to occur at the nucleotide level. Given the apparent shortage of these mutations over what appears to be deep evolutionary time (as indicated by the divergence of the surrounding sequences), it is plausible that instead this region in the *odd* family genes has been subjected to gene conversion events during arthropod evolution, and that these events are resulting in the overestimation of duplication frequencies from the gene tree.

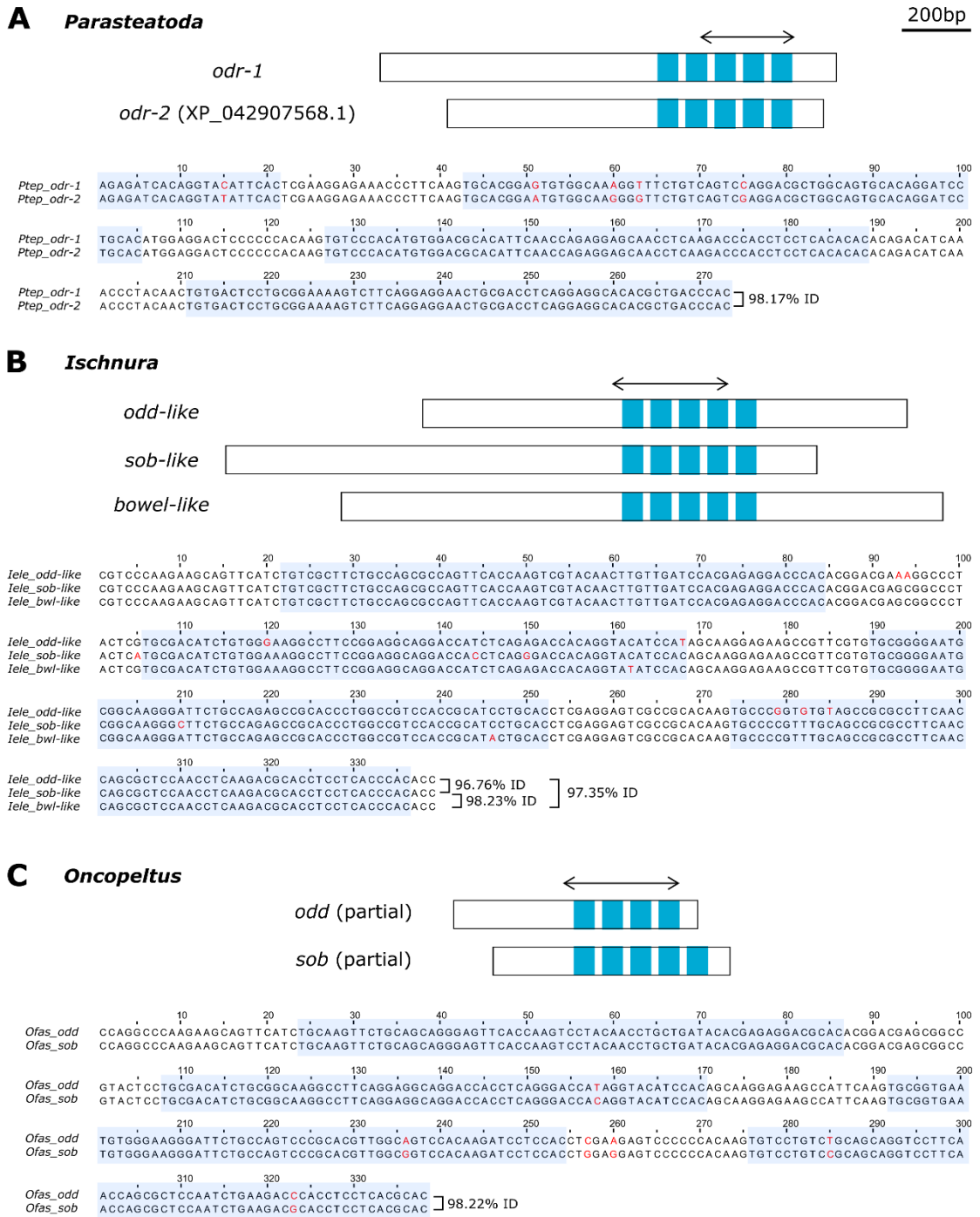


Figure 3.7. Nucleotide sequence alignments of conserved ZF coding regions in *Parasteatoda* (A), *Ischnura* (B) and *Oncopeltus* (C) odd family genes. Diagrams of genes are provided for reference, showing the conserved regions from which the alignments were taken (double headed arrows). Blue squares represent sequence encoding C2H2 ZFs, and are reflected in the alignments by blue shading. Alignment mismatches are highlighted in red, and pairwise sequence similarities are given as % identity (ID). Scale bar refers to all gene diagrams.

Partitioned analyses indicate that multiple gene conversions have occurred between the zinc finger coding regions of *odd*, *sob* and *bwl* genes during insect evolution

To test whether *odd* family genes have undergone concerted evolution, I conducted side-by-side phylogenetics analyses using different gene regions. As the conserved sequence in genes from the same species tends to present itself within the ZF encoding domains (**Fig. 3.7**), I created one alignment using this region alone (from the start of the first ZF to the end of the last), and another based on the surrounding sequence. As the amino acid sequence of the ZF domain is highly conserved across arthropod *odd* family genes, I have used nucleotide sequences, aligned by codons, as the basis for these analyses.

Comparison of the phylograms generated by the different starting sequences reveals a distinct difference in the inferred relationships between arthropod *odd* family genes. When based on the ZF encoding domain, *odd*, *sob* and *bwl* genes commonly group together by species or taxa, rather than by gene class (**Fig. 3.8A**). In contrast, the tree based on the surrounding sequence more closely resembles those already described (**Fig. 3.6**), with distinct clades for each gene, albeit with lower statistical support (**Fig. 3.8B**). This is except for *Schistocerca odd*, *sob* and *bwl* sequences, which group together (**Fig. 3.8B**), despite not doing so in the original nucleotide gene tree (**Fig. 3.8B**). In both ZF and surrounding sequence-based gene trees, *drm* orthologues consistently form their own clade with strong posterior probabilities (0.94 and 0.83 respectively).

These results suggest that the ZF encoding region in *odd*, *sob* and *bwl*, has undergone concerted evolution in insects. This may also be the case for chelicerate, myriapod, and crustacean *odd* family genes, however, the evidence against lineage specific duplications in these groups is still lacking. That said, genes from *Parasteatoda*, *Strigamia* and *Daphnia*, do begin to stop segregating by species in the tree generated from surrounding sequences (**Fig. 3.8B**), although the posterior probabilities for these newly inferred relationships are relatively small.

From the ZF phylogram, clades containing *sob*, *odd*, and *bwl* orthologues from the same species could be interpreted as evidence for gene conversions having occurred in these lineages. This is seen for insect species, *Ischnura*, *Schistocerca*, *Oncopeltus*, *Aphis*, *Apis*,

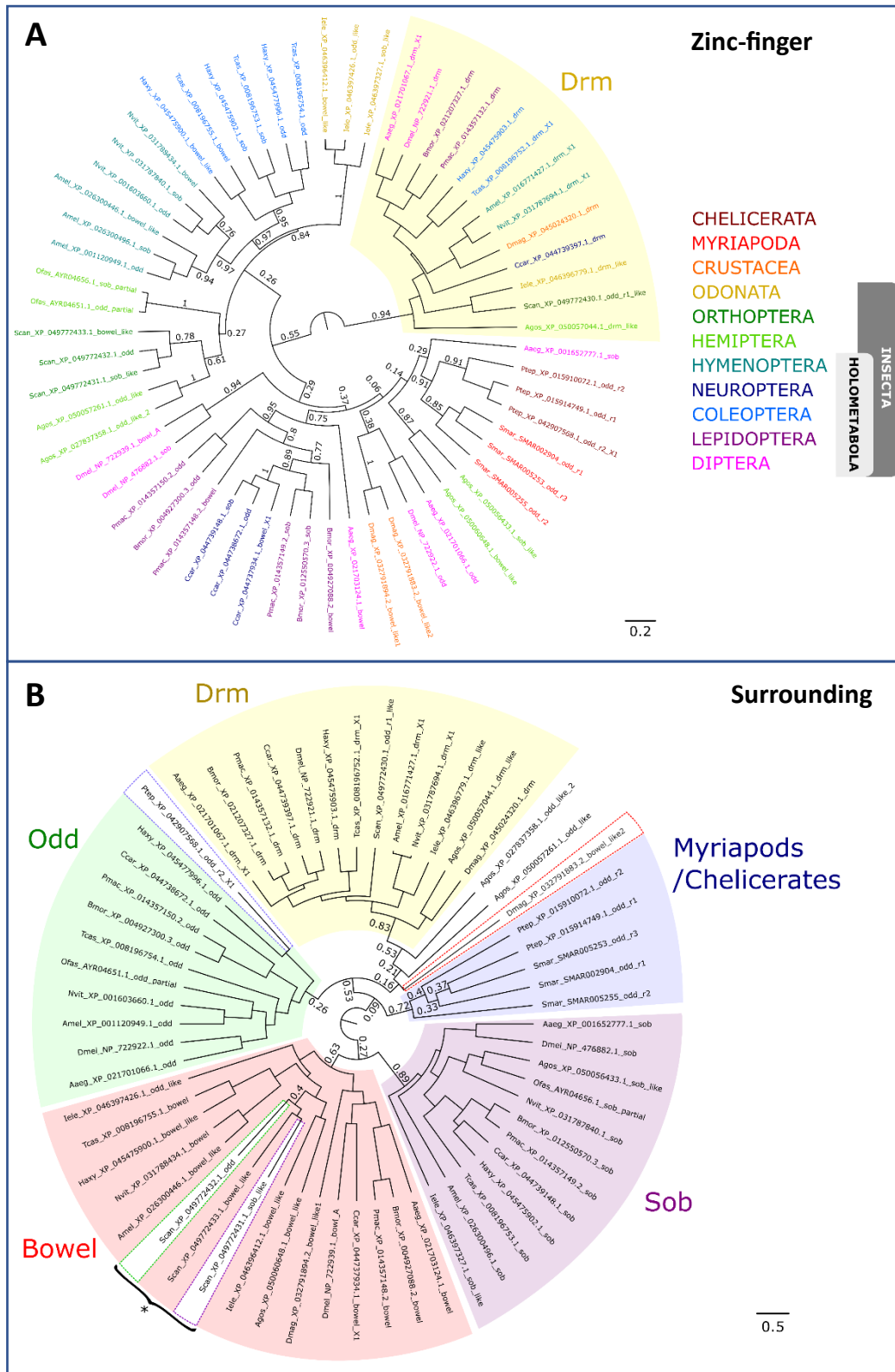


Figure 3.8. Arthropod *odd* family gene trees based on nucleotide codon alignments of the zinc finger coding region (A) and surrounding sequence (B). Phylograms generated by mixed model Bayesian inference. Branch labels indicate posterior probabilities. (A) sequence identifiers are colour coded by arthropod subphyla and insect orders listed to the right. (B) Genes that have changed position from the initial analyses (**Fig. 1**) are outlined with dashed lines coloured according to their original clade. Curly brackets and asterisks mark genes from the same species that have grouped together. Sequence identifiers as described in **figure 3.6**.

Nasonia and *Chrysoperla*. There is also strong statistical support for a clade containing only coleopteran *odd*, *sob* and *bwl* genes, but in which the genes still segregate between the two species, *Tribolium* and *Harmonia*. This could be interpreted as gene conversions having occurred prior to the split of these two coleopteran lineages. Using these inferences, it is possible to map possible gene conversion events to the species phylogeny, demonstrating the widespread occurrence of concerted evolution in the insect *odd* family complex (**Fig. 3.9**).

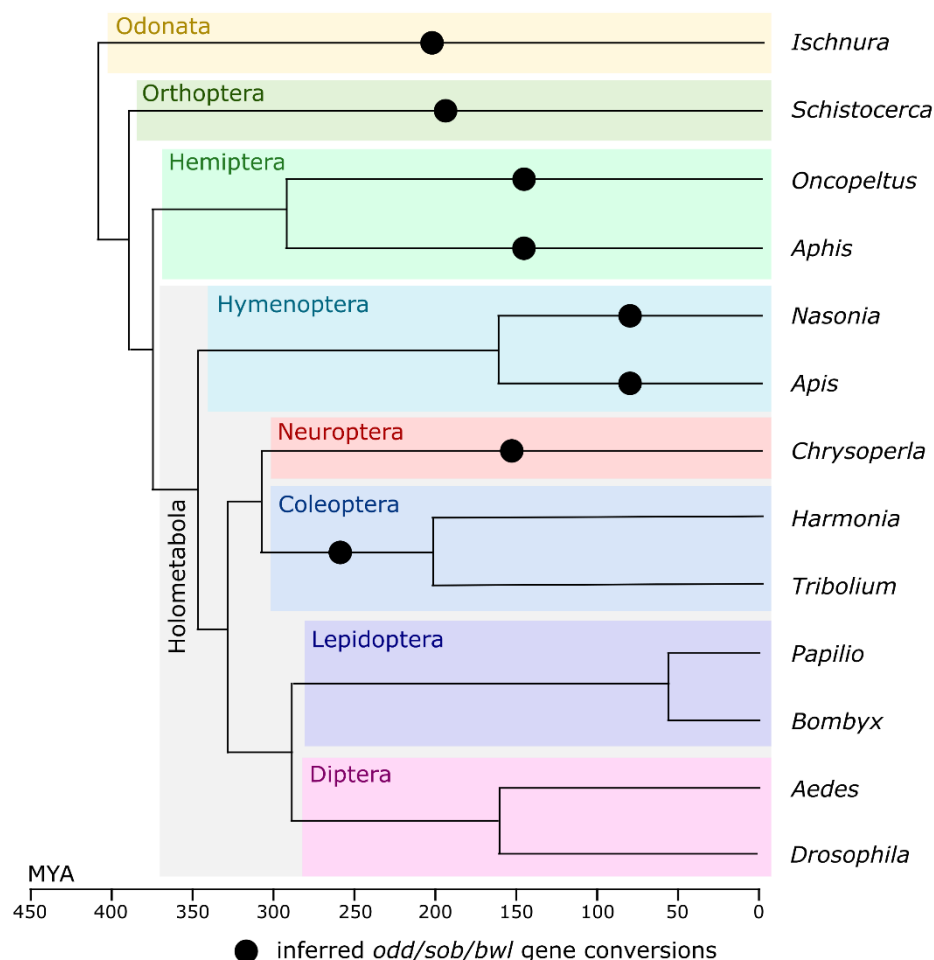


Figure 3.9. Inferred gene conversions in insect *odd* family genes. Gene conversions are inferred by grouping of *odd*, *sob*, or *bwl* sequences from the ZF phylogram (**Fig. 3.8A**). If two or more genes from the same species form their own clade, this is interpreted as a gene convergence in this lineage. Hypothesised gene conversions (black circles) are marked in the centre of branches and do not reflect the precise timing of these events. Similar analyses with more intermediate species would be required to resolve these timings. Tree topology and timescale based on phylogenomic analyses by Misof et.al. (2014). Insect orders highlighted and labelled by colour.

3.2.4. Conserved genomic arrangement of arthropod *odd* family genes

Genomic maps of arthropod *odd* family genes reveal tight linkage across the phylum

As gene conversion rates tend to be higher when the genomic distance between paralogues is shorter (Fawcett and Innan, 2011), close genomic organisation can provide additional support for the occurrence of this phenomenon. In all arthropod species that have chromosome level genome assemblies, at least 2 or more *odd* family genes are found near to each other in the genome (**Fig. 3.10**), suggesting that there is either an evolutionary constraint on the separation of the *odd* family cluster in each case, or that genes have recently duplicated.

The genome of the spider, *Parasteatoda*, used in the phylogenetic analyses, only has a scaffold level assembly. Furthermore, two of the *odd-related* genes in this species, *odr1* and *odr2* (XP_042907568.1), are located at the ends of separate scaffolds, meaning that we cannot rule out genomic linkage between these genes. Instead, a chromosome level assembly for another chelicerate, *Rhipicephalus microplus*, was used for comparison. This species, along with other ticks (*Dermacentor*, *Ixodes*, *Varroa*), appears to possess only 2 *odd-related* genes. In *Rhipicephalus*, these *odd-related* genes exist relatively close to each other, with no other genes between them (**Fig. 3.10**).

Similarly to *Parasteatoda*, *Strigamia* also only has a scaffold level genome assembly. However, at present, there are no annotated, chromosome mapped assemblies for any other myriapod species that can provide a better picture of *odd* complex organisation. Two *Strigamia odd-related* genes, *odr2* and *odr3*, are found nearby on the same

scaffold, while *odr1* is on a separate scaffold. Given the positions of the *Strigamia odr* genes at the end of their respective scaffolds, we cannot rule out that they may still be linked (**Fig. 3.10**).

The genomic arrangement of insect *odd*, *sob* and *drm* genes is highly conserved

The organisation of the *odd* family complex in insects is generally well conserved, with the exception of *bwl*, which appears more susceptible to movement, and has become distanced from the rest of the cluster in hemipteran and dipteran species (**Fig. 3.10**). *Drm*, *sob* and *odd* orthologues are consistently found side by side in the same order along the chromosome and are relatively evenly spaced. In most cases, *drm* is also transcribed in the opposite direction to *sob* and *odd*. While the phylogenetic analysis was unable to distinguish *Ischnura odd* from *bwl*, the genomic organisation supports the initial names assigned from the NCBI refseq annotation, with *lele_odd-like* located next to *lele_sob-like* (**Fig. 3.10**).

The two *Aphis odd-like* genes are extremely close to each other and facing opposite directions (**Fig. 3.10**). Given that in another hemipteran species, *Nilaparvata lugens*, the *odd* family complex appears to be typical of other insects, it is likely that the two *odd-like* genes in *Aphis* are the result of an aphid specific duplication, and that they now share upstream promoters. *Odd-like* genes with the same organisation are seen in other aphids such as the pea aphid (*Acyrtosiphon*) and yellow sugarcane aphid (*Melanaphis sacchari*).

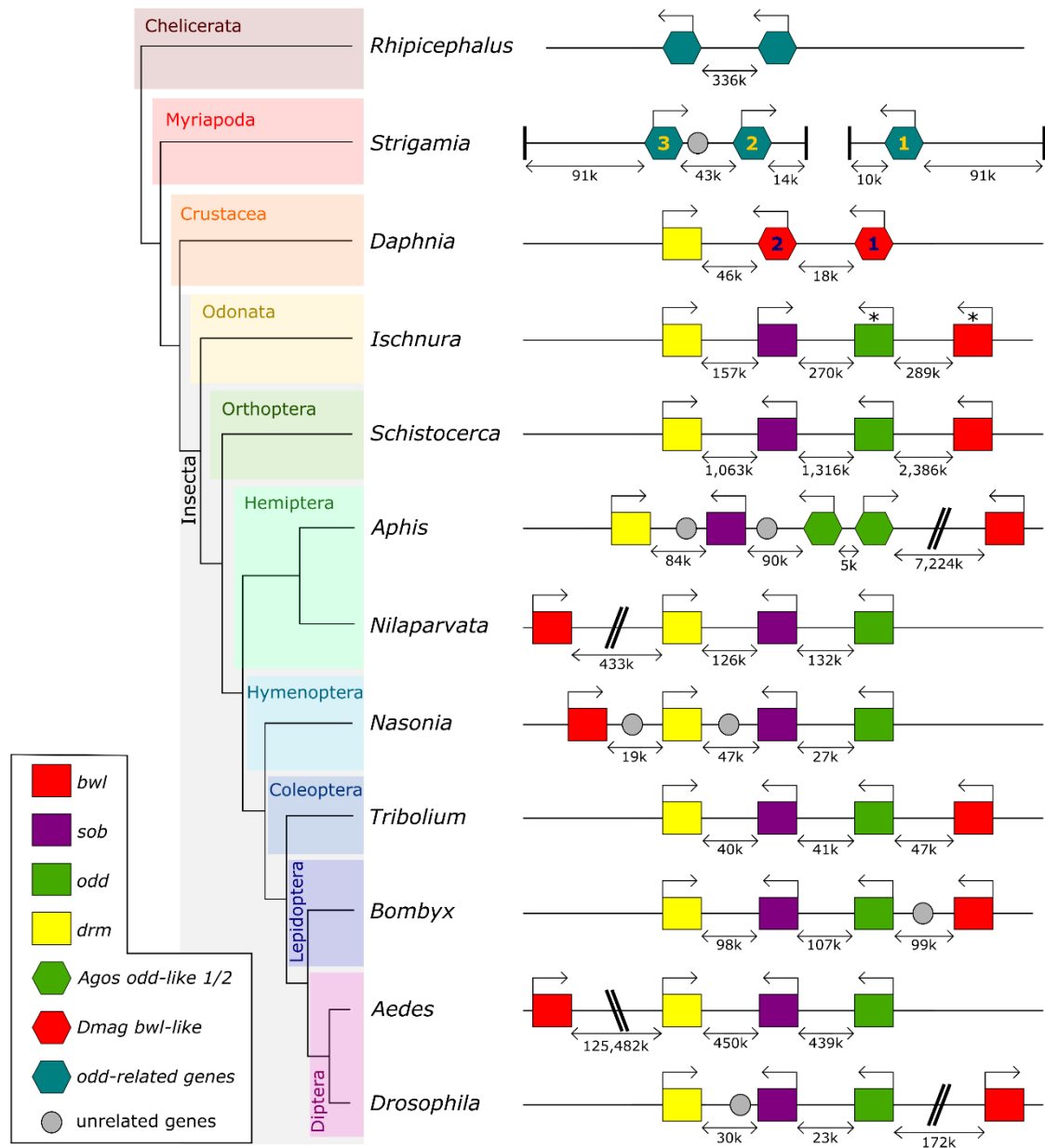


Figure 3.10. Genomic organisation of the *odd* family cluster across arthropods. Orthologues of *bwl*, *sob*, *odd* and *drm* genes confirmed by phylogenetic analysis (Fig. 1, Fig. 3B), except for *Ischnura bwl* and *odd* (asterisks), which have not been distinguished, and are instead based on NCBI genbank entries and their relative positions in the cluster. Transcription direction indicated by arrows above genes. Spacing of genes is not to scale, and real distance is given to the nearest kb (k) in each case. Double slashes mark a greater separation of two genes, between which unrelated genes may lie. Branches of phylogenetic tree not to scale.

3.2.5. *odd* family gene structure consistent with phylogenetics results

When comparing the gene structures of arthropod *odd* family genes, a clear distinction can immediately be made between *drm* orthologues, and the remaining *odd* family genes (**Fig. 3.11**). *Drm* genes are much shorter in length, and possess sequence encoding only 2 ZFs. In Dipteran and neuropteran *Drm* orthologues, the second ZFs have missing histidine residues, although the presence of cysteine residues in their place means that they possibly still function as C2HC (*Drosophila*, *Chrysoperla*) or C2C2 (*Aedes*) ZFs instead. This has already been reported in *Drosophila* when *Drm* was first characterised (Green et.al., 2002), but the observations in *Aedes* and *Chrysoperla* confirm a wider prevalence of this variation.

All *bwl* and *sob* orthologues possess sequence encoding 5 C2H2 ZFs, as do *Daphnia bwl-like* genes, *Parasteatoda odr* genes, and *Strigamia odr-1* and *odr-2* (**Fig. 3.11**). Most insect *odd* orthologues, on the other hand, have 4 ZF coding domains, except for *Ischnura* and *Tribolium* which have 5, and *Aphis odd-like* genes which only have 2 (+ 1 C2C2-like ZF) (**Fig. 3.11C**). Taken alongside the fact that chelicerate, myriapod and crustacean Odd-related proteins all have five ZFs, it is likely that this represents the ancestral state of arthropod Odd family proteins, and that the fifth ZF has subsequently been lost from Odd orthologues during insect evolution. The presence of 5 ZFs in the most basal branching insect, *Ischnura*, may represent the ancestral state of Odd with the fifth ZF lost in other insect lineages, while in coleoptera, a fifth ZF may have subsequently been regained. Supporting this, the fifth ZF of coleopteran Odd orthologues exhibit a large degree of difference from an otherwise highly conserved consensus sequence in other Odd family proteins (**Fig. 3.12**), suggesting that this ZF may not be homologous. In contrast, the fifth ZF of *Ischnura odd* closely matches the consensus sequence (**Fig. 3.12**).

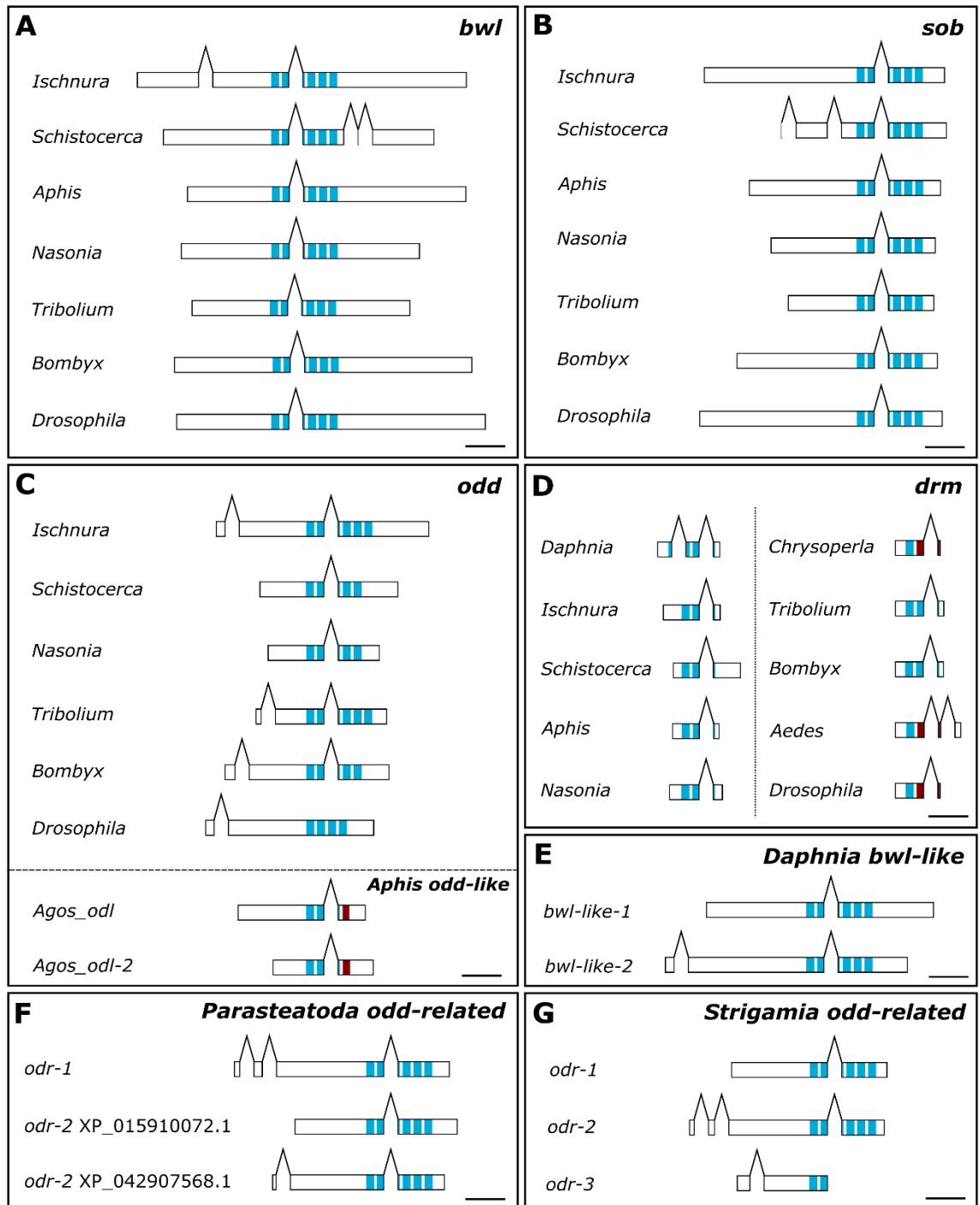


Figure 3.11. Arthropod *odd* family gene structures. To scale representations of coding sequences for *bwl* (A), *sob* (B), *odd* (C), *drm* (D), *Daphnia bwl-like* (E), *Parasteatoda* (F) and *Strigamia* (G) *odd-related* genes. Intron positions are indicated by gaps, while C2H2 and C2HC/C2C2 ZFs are represented by blue and red boxes respectively. Scale bar: 300 bp



Figure 3.12. Comparison of the 5th ZF amino acid sequence in selected insect Odd proteins compared to those of other arthropod Odd family proteins. ZF 5 consensus logo is based on a multiple sequence alignment of all arthropod Odd family proteins with 5 C2H2 ZFs, excluding any insect Odd-skipped orthologues (n=32) (**Appendix A, Fig. A.2**). Overall stack heights represent sequence conservation at that position, while letter heights reflect the relative frequency of each amino acid. Yellow asterisks indicate residues that are present in all ZF 5 sequences used for the consensus. ZF 5 of *Ischnura* Odd-like, *Tribolium* Odd, and *Harmonia* Odd are aligned below the consensus logo for comparison. C2H2 Cysteine (C) and Histidine (H) residues are highlighted in red. Black asterisks mark amino acids that match the consensus sequence.

Arthropod *odd* family genes exhibit conserved intron/exon structure both within and across gene classes

Arthropod *odd* family genes appear to have a highly conserved structural feature, where, in almost all cases, an intron exists within the second zinc finger coding region (**Fig. 3.11**). This intron is always inserted precisely 53 bp along sequence encoding the second zinc finger. Several exceptions exist, such as Diptera *odd* orthologues, and *Drosophila* *drm*. *Strigamia* *odr-3* also lacks this intron, although interestingly the gene stops almost exactly where this intron would usually begin, cutting the second zinc finger short by 3 amino acids. This suggests that any downstream exons have either been missed from the assembly/annotation, or that they have been lost during myriapod evolution, and a shortened *odd* family gene has persisted. A thorough

examination of the sequence downstream of *Strigamia odr3* failed to find any sign of zinc finger coding sequence.

Several other similarities and differences between the overall structures of arthropod *odd* family genes are also apparent. The insect *bwl* genes appear to be the largest from this family, and the zinc finger domain is relatively central (**Fig. 3.11A**). In contrast, the zinc finger coding domain of *sob* genes sit close to the 3' end of the gene (**Fig. 3.11B**). This is more similar to *Parasteatoda* and *Strigamia odd-related* genes (**Fig. 3.11F,G**), which is consistent with the inferred relationship from initial phylogenetic analyses (**Fig. 3.6**). Based on the relative positions of the zinc finger domains, *Daphnia bwl-like-1* appears more similar to insect *bwl* genes, while *bwl-like-2* more closely resembles insect *sob* genes (**Fig. 3.11D**). Interestingly, this is also consistent with their genomic arrangement (**Fig. 3.10**), with *bwl-like-2* sitting between *drm* and *bwl-like-1*, which is the order that *drm*, *sob* and *bwl* are found in most insects. However, given the lack of support from phylogenetic analysis, this is not sufficient evidence to confidently infer relationships between these genes.

3.3. DISCUSSION

3.3.1. Arthropod *Hes/Hey* family evolution overview

Consistent with previously published phylogenetic analyses, *Hes/Hey* genes in arthropods fall into three broad classes, *Hey*, *cwo*, and *Hes*, that have ancient origins in the common ancestor of all major metazoan lineages (Zhou et.al., 2012; Gazave et.al., 2014). In arthropods, the *Hes* superclade can be further categorised into the *E(spl)* genes, which arose in the lineage leading to myriapods and pancrustaceans (Dearden, 2015), insect *Sidpn* orthologues, and *hairy/deadpan* related genes. A number of arthropod *Hes* genes could not be confidently placed within any of these clades, and are possibly the result of lineage specific gene duplications and/or divergences. The *Tribolium* gene, *Tc-Sidpn2*, is an example of one of these genes, and despite its name, does not appear closely related to *Tc-Sidpn1*. Furthermore, it is on the same chromosome as *Tc-h* and *Tc-dpn* (LG9), which would suggest it is more likely to be a paralogue of one of these genes, having since diverged and lost its orange domain. Further analysis of other ill-defined genes may come to similar conclusions, but these would need investigating on a case-by-case basis.

3.3.2. Placing suspected segmentation genes against the consensus *Hes* gene tree

In this chapter, I attempted to resolve the relationships between *Hes* genes that have been reported for roles in segmentation across the arthropod phyla. While functional evidence varies and is still limited in many cases, almost all the genes from the *hairy/deadpan* clade that have so far been the subject of segmentation research, have expression patterns consistent with roles in segment patterning. This is the case for *Cupiennius hairy*, *Strigamia hdpn1-4*, as well as a several insect *hairy* and *deadpan* orthologues (Bier et.al., 1992; Stollewerk et.al., 2003; Aranda et.al., 2008; Pueyo et.al., 2008; Wilson and Dearden, 2012; Green and Akam, 2013; Rosenberg et.al., 2014). This suggests that these genes have an ancient history of involvement in arthropod

segmentation, which has remained the case in at least chelicerates, myriapods, and insects.

Interestingly, the crustacean, *Daphnia*, has no genes within the *hairy/deadpan* class (**Fig.1-2**), despite a study by Eriksson et.al. (2013) demonstrating the function of *Hes* genes and Notch signalling in *Daphnia* segmentation. The two segmentally expressed *Hes* genes from this study, *Hes2* and *Hes3*, are in fact *E(spl)* genes, *E(spl)m7-like-A* and *E(spl)mbeta-like* respectively. Furthermore, the *Daphnia* gene that appears most closely related to the *hairy/deadpan* clade, *Hes-1*, exhibits no striped reiterated expression pattern. Consistent with the conclusion that these segmentation genes are indeed *E(spl)* orthologues, an earlier study in *Daphnia*, found that they regulate neuroblast differentiation (Ungerer et.al., 2012) in a manner reminiscent of *E(spl)* genes in *Tribolium* and *Drosophila* (Culí and Modolell, 1998; Kux et.al., 2013). The role of *E(spl)* genes in *Daphnia* segmentation may also explain why it was not possible to identify any clear *hairy/deadpan* related genes in this species. If *E(spl)* genes were co-opted into segmentation during evolution, this may have made any *hairy/deadpan-like* genes redundant, removing selection pressure, and subsequently opening them up to genetic drift and the acquisition of divergent roles or deletion.

The fact that *Daphnia E(spl)* genes are involved in segmentation demonstrates the interchangeability of *Hes* genes as effectors of Notch signalling across different developmental contexts, and highlights the gaps in our current understanding of this gene family outside of *Drosophila*. While most candidate gene approaches have solely focused on *hairy* orthologues, given its status as a *Drosophila* pair-rule gene, it is clearly important to also consider the wider group of *Hes* family genes when investigating potential roles in segmentation.

3.3.3. Conserved intron structure of *Hes* genes: A link to transcription cycle time constraints?

The intron structures of arthropod *Hes* genes reported here are consistent with those reported for other *Hes* genes across metazoans, with cnidarian, vertebrate and annelid *Hes* genes also possessing homologous introns 1, 2 and 3 (Zhou et.al., 2012; Gazave et.al., 2014). There are also annelid and vertebrate *Hes* genes that only have introns 1 and 2, suggesting that there may be an ancient origin for *Hes* genes with both of these intron structures. However, if this were true, we would expect phylogenetic analyses to produce a clear division between genes with 2 introns from those with 3, which is not the case. It is instead likely that intron 3 has been independently lost in multiple lineages.

Regardless, there is clearly a deep conservation of some aspects of intron/exon structure across metazoan *Hes* genes, which has been maintained in the *hairy/deadpan* genes of arthropods. Conserved introns have, however, been lost from some insect *Sidpn* genes and also from *E(spl)* genes. The *E(spl)* genes of *Daphnia* and the basal branching insect, *Ischnura*, have retained intron 1 only, while in holometabolous insects, such as *Tribolium* and *Drosophila*, all introns have been lost. Unlike *Sidpn* genes, *E(spl)* genes in these species do not appear to have gained additional introns. However, a more comprehensive analysis of *E(spl)* gene structures would be required to confirm if this is true for other species. Whatever the case, there appears to be a varying degree of evolutionary constraint on the intron structure of different arthropod *Hes* genes, with the *hairy/deadpan* class retaining the ancestral form observed across other metazoan groups (Zhou et.al., 2012).

Due to a common functionality of *Hes* family genes as molecular oscillators (Kageyama et.al., 2007), it is feasible that the importance of transcription cycle timing in these systems could be responsible for the conservation of intron structure. Indeed, it has been shown in mice, that removing introns from *Hes7* can accelerate the speed of the segmentation clock, and subsequently increase the number of somites that are patterned (Harima et.al., 2013). While the number of introns significantly impacts the speed of RNA splicing, the length of introns also affects the speed of transcription itself

(Heyn et.al., 2015). No obvious similarities in intron length were observed between arthropod *hairy/deadpan* related genes. However, even if these genes are oscillators that rely on precise transcriptional timing, the exact timings would still likely vary between species via selection and developmental system drift. The increase or decrease of transcription speed, caused by gradual changes to intron length, could even be a mechanism for morphological evolution, whereas gaining or losing an intron may affect timing too drastically to persist. Intron alteration studies of arthropod *Hes* genes, similar to those done in mice, could test this theory.

3.3.4. Arthropod *odd* family evolution overview

In this chapter I have presented strong evidence for the concerted evolution of the arthropod *odd* family complex. Using the extensive genomic data now readily available, it was possible to infer the timing of multiple conversion events between the ZF coding domains of *odd*, *sob*, and *bwl* orthologues during insect evolution. There are signs that this phenomenon may also have occurred in chelicerate, myriapod and crustacean *odd* family genes, given high nucleotide sequence similarity of ZF domains among genes from the same species. However, a more comprehensive species set from these groups would be required to resolve the presence and timing of such events. With the occurrence of concerted evolution in this gene family, it is difficult to confidently establish the number and timings of *odd* family gene duplications in arthropods from the gene trees generated here.

3.3.5. Co-localisation of arthropod *odd* family genes may reflect shared regulatory inputs

One conclusion that can be made from this analysis, is that *drm* orthologues appear to have escaped gene conversion in insects and crustaceans, and have significantly diverged from the other *odd* family genes. This is despite maintaining a close physical linkage to the cluster, in particular with *odd* and *sob* in insects, where a highly conserved genomic structure can be observed across the clade. *Bwl* has switched positions and

even become separated from the complex in some lineages, suggesting that the same constraint on genomic organisation does not exist for this gene.

Perhaps the most parsimonious explanation for the conserved arrangement of *odd*, *sob* and *drm*, is that these genes share important regulatory elements in their local genomic landscape. In many cases, genes that require co-ordinated expression are often co-localised to prevent them from being differentially regulated by chromosomal factors such as DNA methylation and histone modification, and chromosome compaction (Zhang et.al., 2019). Published expression patterns for *Tribolium* and *Drosophila odd*, *sob*, and *drm* genes show, that while still variable, they do indeed appear to be similarly expressed during certain developmental stages (Hart et.al., 1996; Green et.al., 2002; Hao et.al., 2003; Thümecke and Schröder, 2021). Furthermore, in *Tribolium*, *bwl* expression appears similar to *odd*, *sob*, and *drm* (Thümecke and Schröder, 2021), while in *Drosophila*, it is notably different (Hart et.al., 1996). This is consistent with the separation of *bwl* from the *odd* family cluster in dipterans. In the hemipteran, *Oncopeltus*, *odd* and *sob* are also expressed in almost identical patterns during segmentation (Auman and Chipman, 2018; Reding et.al., 2019), further supporting the idea that co-localisation is linked to co-expression.

While we cannot ascertain the exact genomic organisation of *Strigamia odd-related* genes (due to the lack of a chromosome level assembly), all three *odr* genes in this species are expressed in dynamic stripes in the SAZ (Green and Akam, 2013). Each *odr* gene appears to be inactivated at different AP positions, however, all are expressed in the same phase relative to *eve*, suggesting that they are likely co-expressed with each other. In the spider, *Cupiennius*, one *odd-related* gene is expressed in the SAZ, consistent with a role in segment patterning, while two other *odd-related* genes were not observed in this region (Damen et.al., 2005). However, no expression was described elsewhere, meaning further expression analyses would be required to establish whether these genes are co-expressed at other stages. A chromosome level genome assembly would also be required to determine the genomic organisation of these genes in the spider and whether they are located close to each other.

3.3.6. Context for subsequent chapters

Overall, the analyses of these two gene families highlights some important aspects of gene family evolution, and provides us with important additional context for evo devo research around these genes. The story of *E(spl)* co-option in *Daphnia* segmentation informs us about the capacity of ancient developmental systems to change and co-opt new genes for old tasks, and while the emphasis of segmentation research in arthropods has always been on *hairy* orthologues, it is clear that more attention should be paid to other members of the *Hes/Hey* gene family. Meanwhile, the concerted evolution of the sequence encoding the ZF domains of *odd* family genes, along with their conserved genomic arrangement, means that they have great potential to share redundancy with each other. Indeed, this relationship has already been described for *odd*, *sob* and *drm* in the context of leg patterning in *Drosophila* (Hao et.al., 2003). Similar striped expression patterns of *odd*, *sob* and *bwl* in the *Tribolium* germband point towards possible redundancy in the context of segmentation also (Thümecke and Schröder, 2021).

4. A REVISED UNDERSTANDING OF *ODD* FAMILY GENE EXPRESSION IN *TRIBOLIUM*

4.1. INTRODUCTION

4.1.1. *odd* family expression and conserved functions in arthropods

Odd-skipped was the first *odd* family gene to be characterised. It was initially described for its pair-rule function within the *Drosophila* “segmentation cascade” (Coulter et.al., 1990). Other members of the insect *odd* family genes, *sob*, *drm* and *bwl* were identified several years later, when, alongside *odd*, they were discovered to play important roles in *Drosophila* leg development and hindgut specification (Hart et.al., 1996; Iwaki et.al., 2001; Green et.al., 2002; Hao et.al., 2003).

In the context of arthropod leg development, *odd* family genes appear to have an ancestral role as downstream effectors of Notch-mediated limb segmentation and growth, with *odd-r1* in the spider, *Cupiennius*, exhibiting similar expression patterns and regulatory input from Notch signalling as are observed for *odd*, *sob* and *drm* in *Drosophila* (Hao et.al., 2003; Prpic and Damen, 2009). In *Tribolium*, *odd*-family genes appear to function in the same capacity alongside Notch signalling in the patterning of adult legs and mandibulate mouthparts during metamorphosis (Angelini et.al., 2012 (a); Angelini et.al., 2012 (b)), while knockdowns of *Tc-drm* and *Tc-bwl* expression during embryogenesis produce larval phenotypes consistent with embryonic roles for these genes in leg segmentation and hindgut specification (Thümecke and Schröder, 2021). Consistent with these functions, *Tc-sob*, *Tc-drm* and *Tc-bwl* are all expressed in similar patterns in the limb buds, developing head appendages and hindgut anlage of *Tribolium* embryos (Thümecke and Schröder, 2021).

Like their *Drosophila* orthologues, *Tribolium* *Tc-odd*, *Tc-sob* and *Tc-drm* genes are each expressed in rings around prospective joints during leg development. In *Drosophila*, these overlapping domains confer functional redundancy in promoting joint formation,

wherein single gene knockouts of *odd*, *sob* or *drm* do not produce any detectable defects relating to leg segmentation or growth (Hao et.al., 2003). In contrast, single knockdowns of *Tc-drm* in *Tribolium* are sufficient to produce fused leg segments, suggesting that, despite their conserved expression patterns, the same level of redundancy is not found in *Tribolium* as *Drosophila* (Thümecke and Schröder, 2021). That said, 3 of 6 leg joints are still established in the absence of *Tc-drm*, and the retained expression of *Tc-sob* around these joints suggests that there may still be some partial redundancy between *odd*-family genes, although this was never tested (Thümecke and Schröder, 2021). Regardless of the exact molecular mechanisms, observations of *odd* family expression and function in the context of arthropod leg development tells us that these genes have the capacity to function redundantly with each other. Whether this may also be the case with regards to body segmentation is, as of yet, unknown.

4.1.2. Arthropod *odd*-family genes are regularly expressed in patterns consistent with roles in segmentation

Due to its importance in *Drosophila* segmentation, direct orthologues of *odd-skipped* have commonly been the subject of candidate gene approaches for segmentation research in other arthropods, while much less attention has been paid to other members of the gene family. Despite the lack of any functional roles having yet been described for other *odd*-family genes with regards to segmentation, in cases where expression has been investigated, similar expression patterns are often observed during segment patterning.

In *Drosophila*, *Tc-sob* and *Tc-drm* are both expressed in 7 pair-rule stripes at the same time as *Tc-odd* (Hart et.al., 1996; Green et.al., 2002), while in *Tribolium*, *Tc-sob* and *Tc-bwl* are reported to be expressed in stripes in the SAZ during early GBE that resemble those of *Tc-odd* (Choe et.al., 2006; Thümecke and Schröder, 2021). In the milkweed bug, *Oncopeltus*, *Of-sob* is reported to be expressed in a seemingly identical pattern to that of *Of-odd* during segment addition (Auman and Chipman, 2018; Reding et.al., 2019). And, despite their unclear relationships to insect *odd*-family genes (**section 3.2.3**), all 3

odd-related genes of the centipede, *Strigamia*, are also found to be expressed in overlapping stripes in the SAZ during segmentation (Green and Akam, 2013).

As discussed in the previous chapter, these patterns may solely reflect shared regulatory inputs as the result of genomic co-localisation. However, functional redundancy, in the context of segmentation, has not yet been fully tested as an alternative explanation for these observations. Furthermore, there is a need for co-expression analyses to determine if the expression of these genes truly overlap.

4.1.3. HCR as an effective method for evaluating gene co-expression

Colorimetric *in situ* hybridisation, using antigen-labeled nucleotide probes and antibody staining, is a well-established technique in *Tribolium* for visualising patterns of gene expression during development (Brown et.al., 2009; Schinko et.al., 2009). This remains a useful and cost-effective method for screening expression patterns of candidate genes. However, if evaluating co-expression patterns of multiple genes at once, these experiments have several limitations.

In recent years, new multiplex fluorescent *in situ* hybridisation (FISH) techniques have been established and adapted for use in *Tribolium*. One such technique is hybridisation chain reaction (HCR), developed and refined through several iterations (Ikbal et.al., 2015; Bi et.al., 2017; Choi et.al., 2018). The most recent iteration, HCR 3.0, utilises sets of short (~10-20bp) probe pairs that are tagged by complimentary fragments of ssDNA (Choi et.al., 2018). When these pairs bind adjacent sequences of mRNA, their complimentary ssDNA fragments come together to form an active initiator sequence. The advantage this gives is that if just 1 probe without its partner binds non-specifically in the sample, this probe will lack the complete initiator sequence and will be inert. This reduces background “noise” by ensuring that probes binding off-target sites are less likely to generate non-specific signals (Choi et.al., 2018).

Complete initiator sequences trigger the polymerisation of fluorophore-tagged DNA hairpins once added, which amplify a fluorescent signal along the length of the gene transcript of interest. 5 different initiator sequences, combined with 5 DNA hairpins

labelled with distinct fluorophores, enables the potential visualisation of 5 genes in parallel within a single embryo (Choi et.al., 2018). Successful examples of 4 gene multiplex HCR FISH have previously been carried out in *Tribolium* (Tidswell et.al., 2021). However, the availability of confocal microscope lasers and embryo autofluorescence has limited my experiments to just 3 genes at once in this case. Regardless, this still proves a greatly useful tool for accurate visualisation and analysis of any potential overlap in *odd*-family gene expression.

4.1.4. Aims of the chapter

In this chapter, I will aim to corroborate previously published expression patterns for *odd-skipped*-family genes in *Tribolium*, and to evaluate the extent to which their expression domains overlap during segmentation. If there is any functional redundancy in the clock network between these genes, we would expect to see them co-expressed in the *Tribolium* SAZ.

Given the differences in phenotype severity observed between SB and GA2 strains after *Tc-odd* knockdown (**section 1.6**), I will also compare the expression patterns of *odd*-family genes across these two strains. If there has been a genuine intraspecies divergence in the level of redundancy conferred by *odd*-family genes with regards to body segmentation, then this may be evident from their expression dynamics.

4.2. RESULTS

4.2.1. Staging with *Tc-wingless (Tc-wg)* as a marker

For the purpose of staging *Tribolium* embryos, I selected the segment polarity gene, *Tc-wingless (Tc-wg)*, as a marker. *Tc-wg* stripes appear sequentially as new parasegment borders are established throughout the course of segment addition. Each stripe marks the posterior boundary of a parasegment and corresponds to a future segment identity (**Fig. 4.1B**). In keeping with *Tc-wg* staging used by Tidswell et.al. (2021), I will refer to stages based on the posterior-most *Tc-wg* stripe and the parasegment it corresponds to (**Fig. 4.1A**). The ectoderm between the final *Tc-wg* stripe and the posterior *Tc-wg* domain can be considered the segment addition zone (SAZ) (Janssen et.al., 2010), in which primary segment patterning occurs.

4.2.2. A revised understanding of *odd* family gene expression in *Tribolium*

Tc-drm and *Tc-sob* are expressed in late pair-rule stripes during segment addition

In order to investigate possible redundancies between *Tribolium odd* family genes, I used hybridisation chain reaction (HCR) multiplex Fluorescent *in situ* hybridisation (FISH) to see how these genes are expressed relative to each other. Probes were designed to avoid the highly similar ZF coding regions of each gene (**Table 2.2**). *Tc-drm* and *Tc-sob* show highly similar expression patterns during all stages captured in this study up to 48 hours after egg lay. Both are expressed in late pair-rule stripes during segment addition in a phase opposite to the pair-rule stripes of *Tc-odd* expression (**Fig. 4.2. cyan/magenta arrowheads**). These domains abutt the posterior edge of their corresponding *Tc-wg* stripe, without overlapping, and their expression shortly precedes that of *Tc-wg*. *Tc-drm* and *Tc-sob* pair-rule stripes are transiently expressed in newly established segments and turn off as segments mature. However, in the case of head and thoracic segments, *Tc-drm* and *Tc-sob* become expressed in head appendage and limb bud domains immediately after the pair-rule stripes disappear (**Fig. 4.2. wg9-**

wg13). In some cases, faint secondary stripes can also be seen in segments where *Tc-odd* is expressed, and vice versa, with secondary *Tc-odd* stripes overlapping primary *Tc-drm* and *Tc-sob* pair-rule stripes (**Fig. 4.2. faint arrowheads**). Interestingly, in contrast to previously published results for *Tc-sob* (Thümecke and Schröder, 2021), no striped expression was observed in the posterior SAZ. In addition, *Tc-drm* stripes were previously only reported in an early germband stage, whereas here I have observed late pair-rule stripes of expression throughout segment addition.

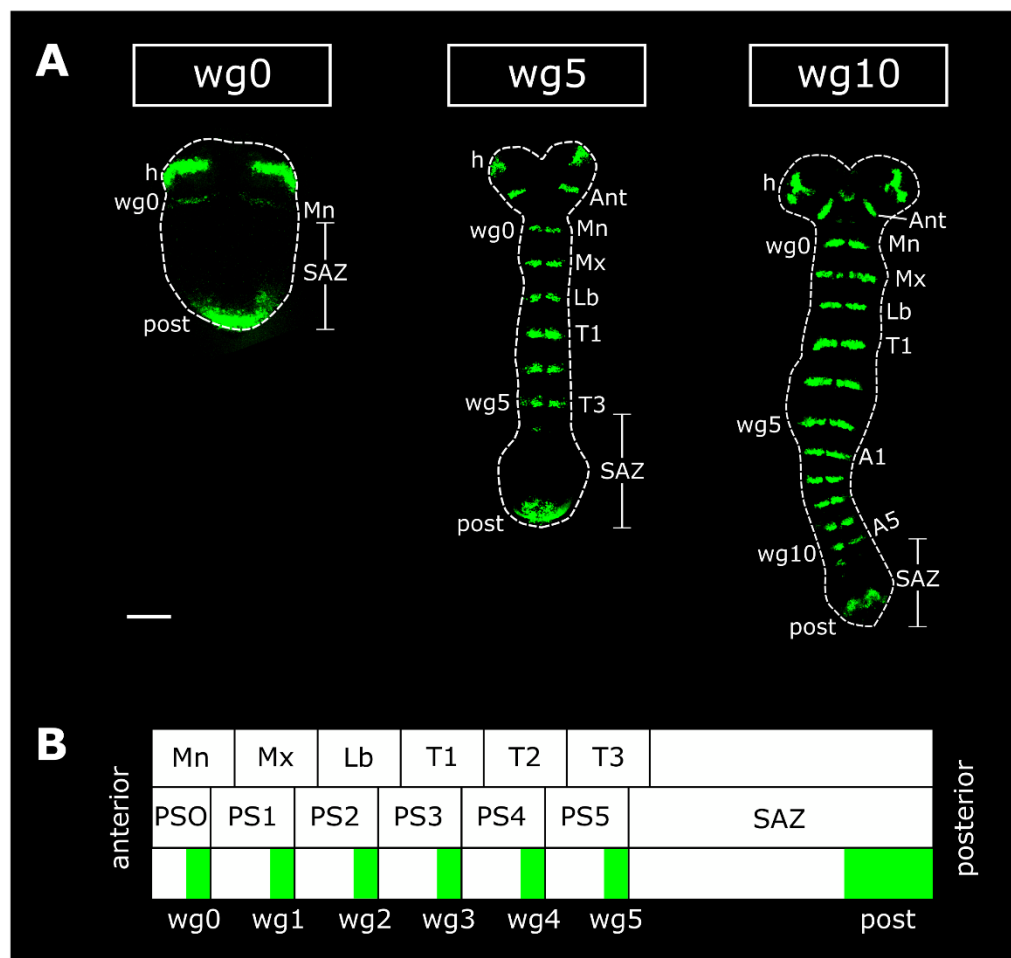


Figure 4.1. *Tc-wg* staging through GBE. (A) Examples of *Tc-wg* expression in 3 embryos at different stages of segment addition (taken from HCR FISH experiments). Stages defined by posterior-most *Tc-wg* stripe. Segment identities for *Tc-wg* stripes are given to the right of each embryo. h = head domain, Ant = Antennal, Mn = Mandibular, Mx = Maxillary, Lb = Labial, T(1-3) = Thoracic, A(1-5) = Abdominal, post = posterior domain. (B) Example schematic illustrating relative positions of *Tc-wg* stripes against parasegment and segment boundaries in a *wg5* stage embryo. SAZ = Segment Addition Zone. Scale bar: 100 μ m

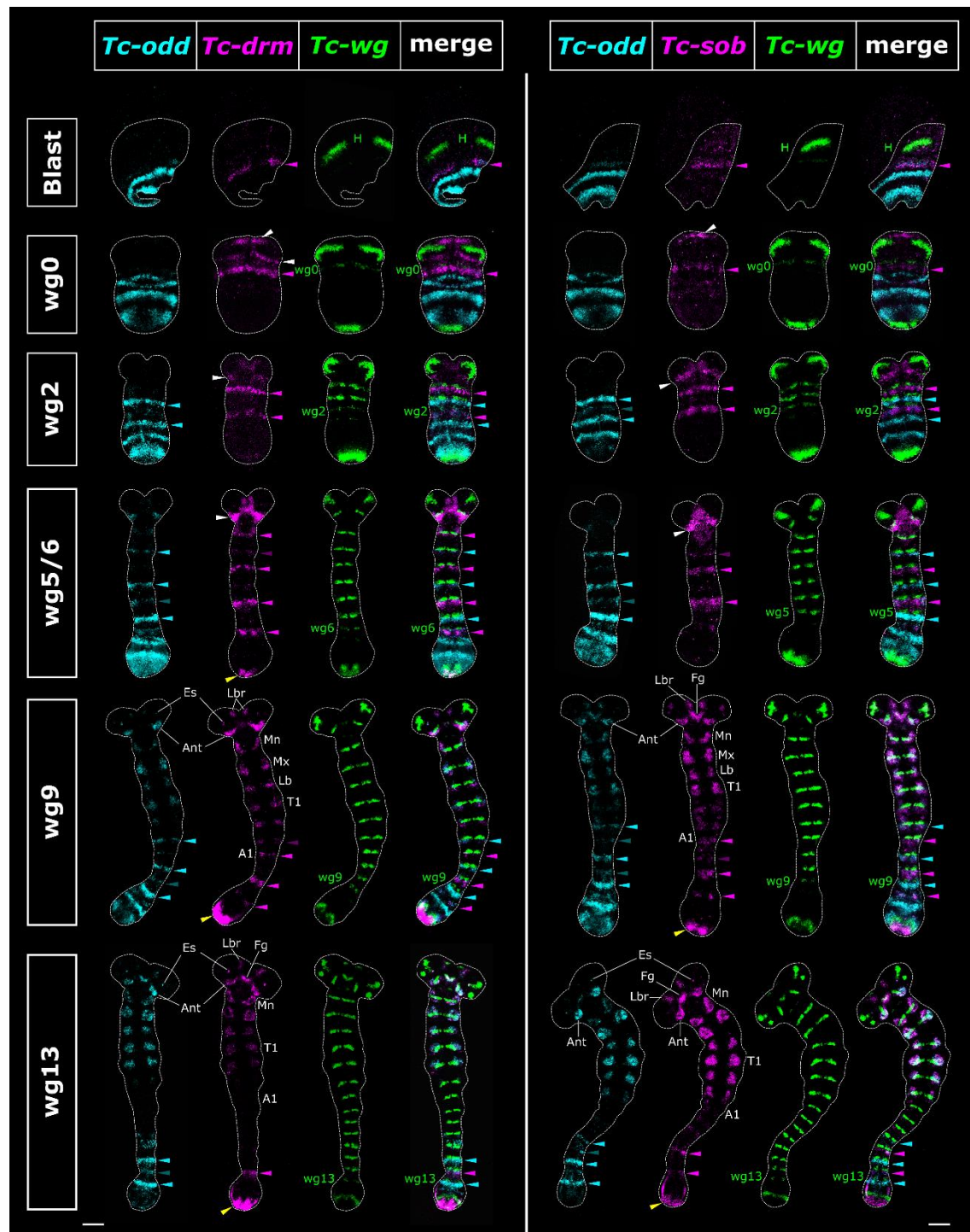


Figure 4.2. Expression of *Tc-drm* and *Tc-sob* in relation to *Tc-odd* during segment addition. SB strain *Tribolium* embryos from HCR FISH experiments showing *Tc-odd/Tc-drm/Tc-wg* (left) and *Tc-odd/Tc-sob/Tc-wg* (right) expression at comparative stages of segment addition. For each embryo, individual channel images are shown for all 3 genes alongside a merged image, as indicated by labels along the top of the figure. Embryos from each experiment were staged based on *Tc-wg* expression (**Fig. 4.1**) and are indicated by the labels down the left-hand side. The posterior-most *Tc-wg* stripe is labelled in green to demonstrate embryo staging. Cyan (*Tc-odd*) and magenta (*Tc-drm* / *Tc-sob*) arrowheads indicate pair-rule stripes, while dimmed arrowheads indicate weaker secondary pair-rule stripes. White arrowheads mark the early head domains of *Tc-sob* and *Tc-drm*, and yellow arrowheads mark the posterior *Tc-sob* and *Tc-drm* domains. Abbreviations: Lbr=Labrum, Ant=Antennal, Mn=Mandibular, Mx=Maxilla, Lb=Labia, Es=Eyespot, Fg=Foregut, T=Thoracic, A=Abdominal. Scale bar: 100 μ m

Tc-drm and *Tc-sob* are also both expressed in the head, first in a transverse stripe along the anterior head of newly formed germbands, and subsequently along the base of the head lobes (**Fig. 4.2. white arrowheads**). In later stage germbands, *Tc-sob* and *Tc-drm* are expressed in several distinct domains within eye spots, labrum and antennae anlage (**Fig. 4.2. wg9-wg13**). Both genes also become expressed in the posterior tip of the germband, behind the posterior *Tc-wg* domain, from around stage wg6 (**Fig. 4.2. yellow arrowheads**). While *Tc-sob* and *Tc-drm* expression patterns are almost identical, *Tc-sob* appears to exhibit stronger expression in head appendages and limb buds, while *Tc-drm* expression is stronger in the posterior domain. Unlike previously published results (Thümecke and Schröder, 2021), no *Tc-bwl* expression was observed during segment addition, and instead, only becomes visible in later stages (see below).

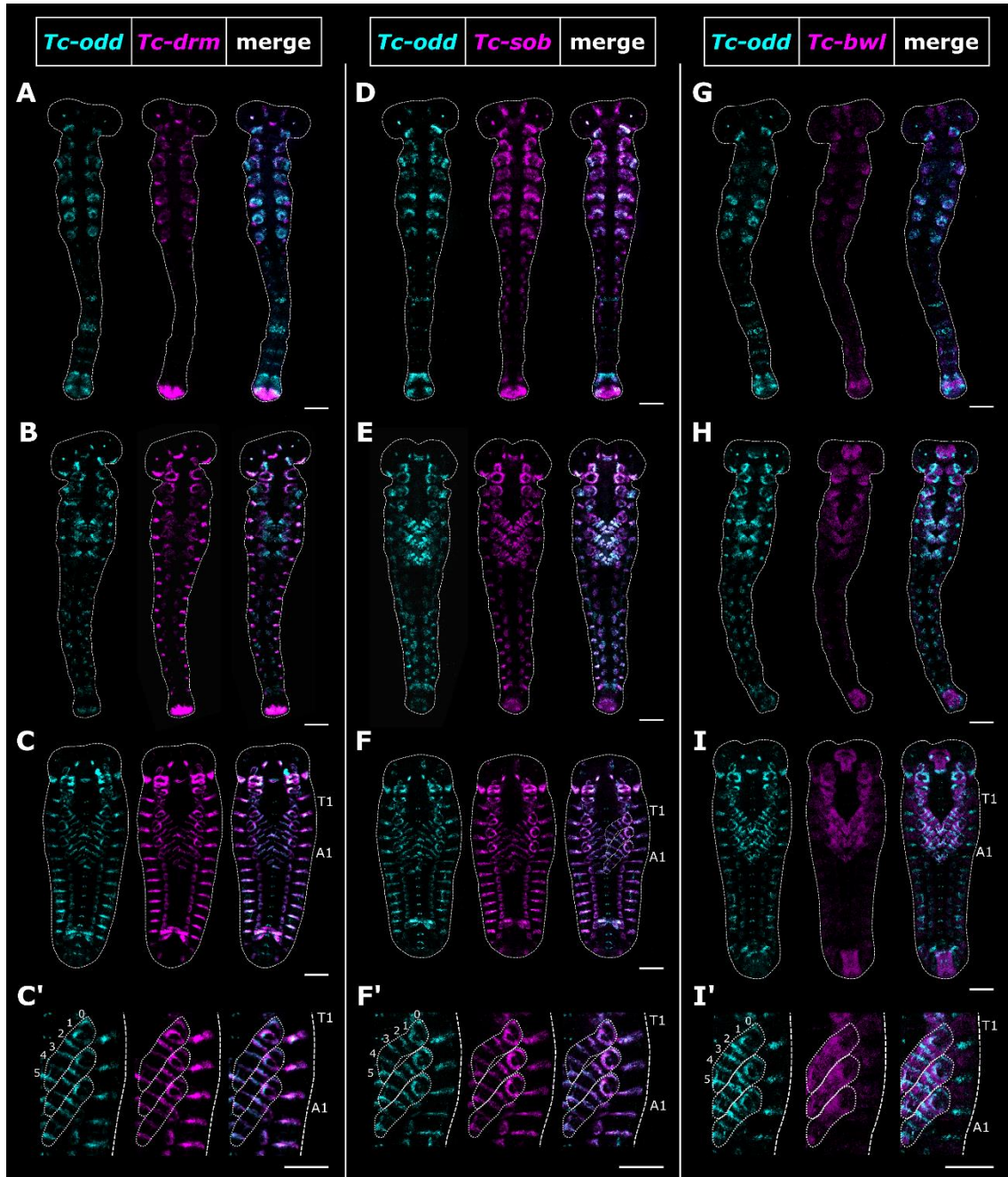
Post GBE expression of *odd*-family genes is consistent with roles in leg development and hindgut specification.

After segment addition has finished, *Tc-odd*, *Tc-sob* and *Tc-drm* are expressed in almost identical overlapping patterns. They continue to be co-expressed in domains within the developing head appendages, limb buds and presumptive hindgut. As trunk segments mature, *Tc-odd*, *Tc-sob* and *Tc-drm* become expressed in transverse stripes along

segment boundaries. In late stages, all three genes display overlapping striped expression in the developing legs, with one ring around the base, and 5 more that correspond to each leg joint (**Fig. 4.3C', F'**) (Thümecke and Schröder, 2021).

Tc-bwl expression appears shortly after the completion of GBE and is also found in the head appendages, limb buds and presumptive hindgut region (**Fig. 4.3G**). Expression continues in these domains and grows stronger as development proceeds. However, the exact domains of *Tc-bwl* are distinct from those of other *odd*-family genes. During hindgut morphogenesis, this is particularly apparent, as *Tc-bwl* is expressed in a broad domain around the rectum (King and Denholm, 2014), while *Tc-sob/Tc-drm/Tc-odd* expression is restricted to an adjacent anterior region of the hindgut (**Fig. 4.3C,F,I**). When other *odd*-family genes become expressed in rings in the legs, *Tc-bwl* is diffusely expressed between *Tc-odd* leg stripe 1 to the tip of the legs (**Fig. 4.3I'**). While this uniform late-stage *Tc-bwl* expression in the leg is generally consistent with that reported by Thümecke and Schröder (2021), no elevated stripes of expression were observed in the prospective joints in this instance.

Figure 4.3. (overleaf) Post GBE expression of *odd* family genes in *Tribolium*. HCR multiplex FISH for *Tc-odd/Tc-drm* (A-C'), *Tc-odd/Tc-sob* (D-F') and *Tc-odd/Tc-bwl* (G-I'). Individual channel images are shown for each gene alongside merged channel images to indicate co-expression. Three similar embryonic stages are shown for each experiment based on morphology and *Tc-odd* expression. Embryos are arranged in increasing age from top to bottom. C', F' and I' are taken from embryos C, F and I respectively to highlight gene expression in developing legs. Dotted lines outline legs and rings of *Tc-odd/Tc-drm/Tc-sob* expression, with 0 corresponding to base of the leg, and 1-5 marking each leg joint (based on inferences by Thümecke and Schröder (2021). T=Thoracic, A=Abdominal. Scale bar: 100 μ m



4.2.3. *odd* family gene expression shows no signs of divergence between SB and GA2 strains

To test for redundancy and developmental system drift between *odd* family genes, and to corroborate the differences in expression patterns observed here compared to those previously published by Thümecke and Schröder (2021), I performed HCR FISH experiments on SB and GA2 strain embryos in parallel. While details of the strain used by Thümecke and Schröder (2021) were not provided, strain variation was considered

one possible explanation for the differences in results. However, there were no obvious differences in *odd* family gene expression between strains at all stages that I captured (up to 48 hours after egg lay). *Tc-drm* and *Tc-sob* were both expressed in the same late pair-rule stripes during segment addition in GA2 and SB embryos (**Fig. 4.4**), while *Tc-bwl* expression was similarly only seen after GBE (data not shown). Importantly, the previously reported stripes of *Tc-bwl* and *Tc-sob* expression in the posterior SAZ (Thümecke and Schröder, 2021) were still not observed in the GA2 strain, ruling out the possibility that these contradicting results are down to strain differences.

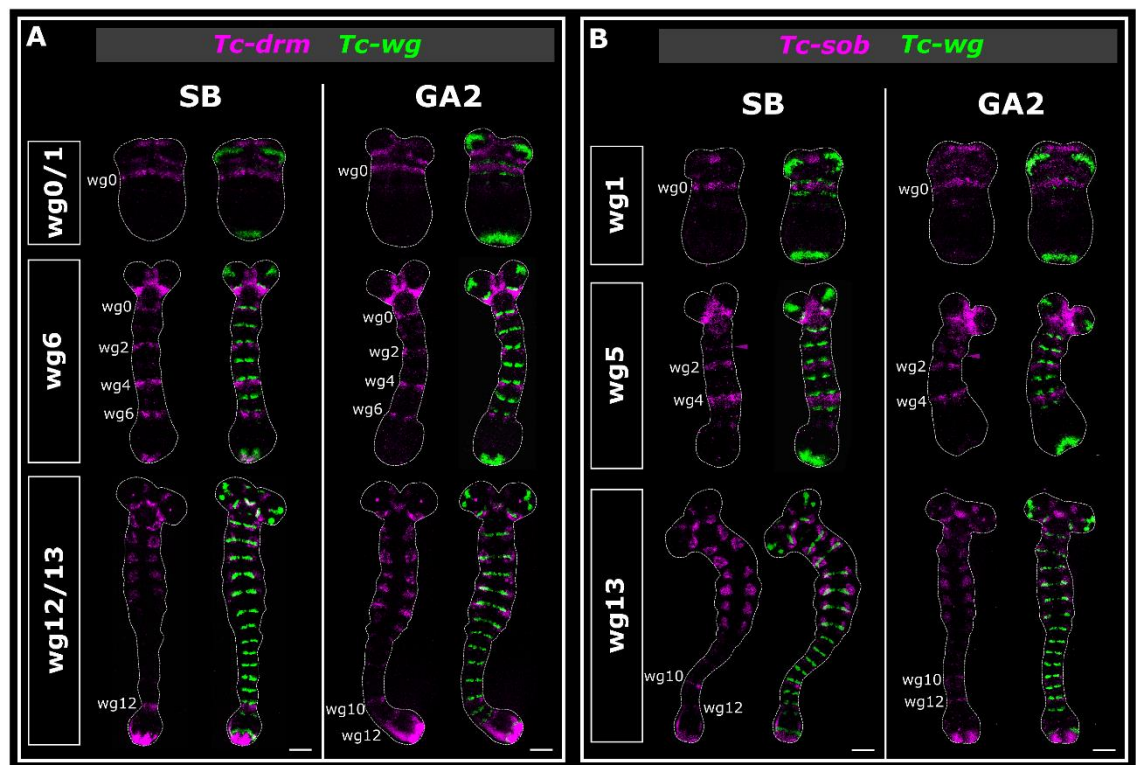


Figure 4.4. Comparison of *Tc-drm* and *Tc-sob* expression in SB and GA2 strains during segment addition. Three comparative germband stages of SB and GA2 embryos are shown from HCR FISH experiments with *Tc-drm* (A) and *Tc-sob* (B) carried out in parallel. *Tc-wg* (green) was included as a marker in each case and is shown alongside single channel images of *Tc-drm*/*Tc-sob* (magenta) expression to indicate positions of pair-rule stripes. Germband stages are labelled down the left-hand side of each panel. Pair-rule stripes are labelled in relation to the corresponding *Tc-wg* stripe that they border. Secondary *Tc-sob* pair-rule stripes are indicated by magenta arrowheads. Scale bar: 100 μm

4.3. DISCUSSION

In this chapter, I have described a revised understanding of *odd* family gene expression during segment addition in *Tribolium*. In contrast to results by Thümecke and Schröder (2021), I have found no evidence of *Tc-sob* or *Tc-bwl* expression in the posterior SAZ during GBE, and instead show that *Tc-sob*, along with *Tc-drm*, is transiently expressed in late pair-rule stripes in opposite phase to those of *Tc-odd*, while *Tc-bwl* expression is not observed until after GBE. Consistent with previously described expression patterns, *Tribolium odd*-family genes are also expressed in domains in the developing head appendages, legs, and the hindgut (Thümecke and Schröder, 2021). Comparison of *odd* family gene expression between SB and GA2 strains showed no divergence and is therefore unable to explain the differences in *Tc-odd* knockdown severity or *odd* family expression between my results and those previously published.

4.3.1. Previously published *Tc-bwl* and *Tc-sob* expression is consistent with cross-reaction of *in situ* probes with transcripts of other *odd* family genes

RNA *in situ* probes used in study by Thümecke and Schröder (2021) encompass a region of the zinc-finger coding domain with a high degree of sequence similarity between *Tc-sob*, *Tc-bwl* and *Tc-odd*

In order to understand why the expression patterns for *Tribolium odd* family genes presented in this study are different to those reported by Thümecke and Schröder (2021), I compared the methods, specifically with regards to probe design. As discussed in the Chapter 3, arthropod *odd* family genes share a highly conserved zinc-finger coding region, which in the case of insect *odd*, *sob* and *bwl* orthologues, has been the subject of gene conversion events during evolution, resulting in an especially high degree of sequence similarity between these domains in genes from the same species. With this knowledge, I specified probes to be designed using mRNA sequence outside of this region (**Table 2.2**). However, mapping of the forward and reverse primers used for *Tc-bwl* and *Tc-sob* probe synthesis by Thümecke and Schröder (2021), reveals that their probes overlap this region of high nucleotide similarity (**Fig. 4.5**).

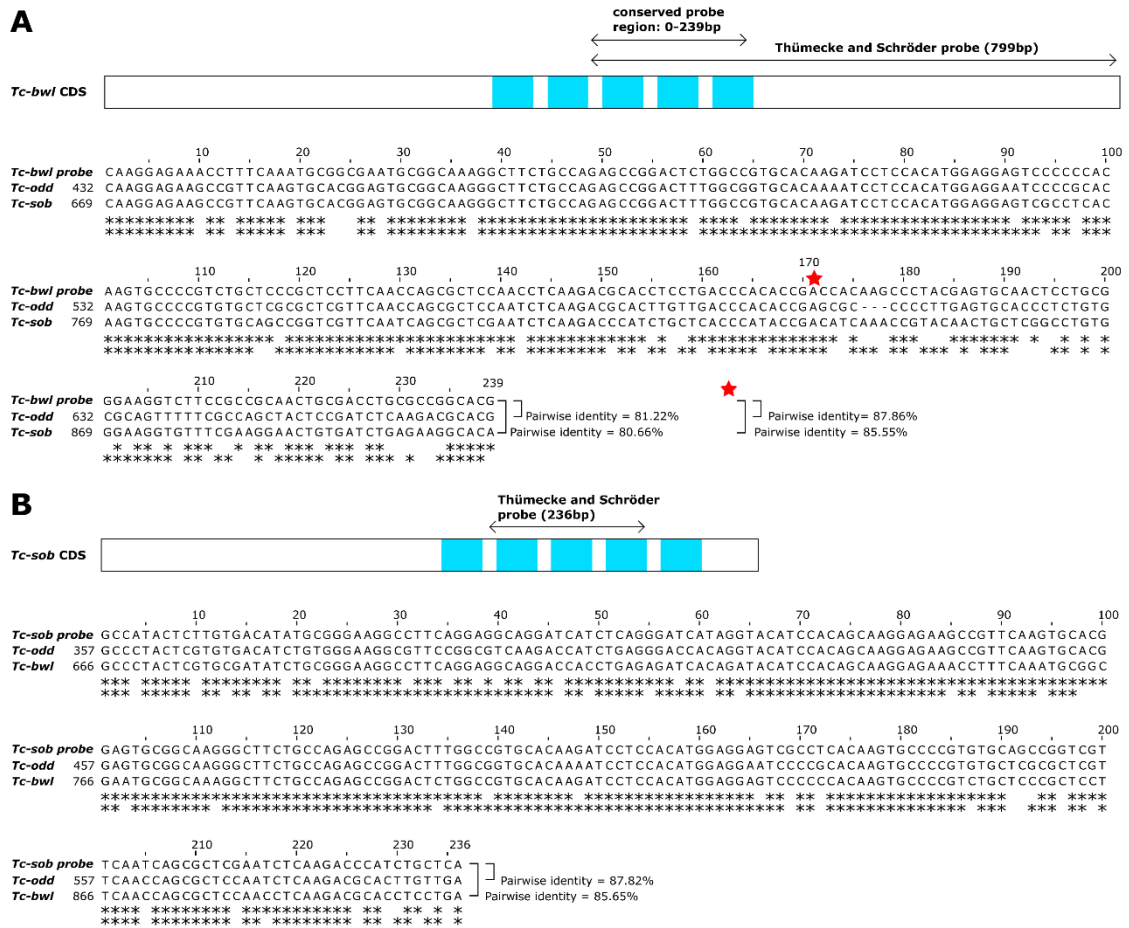


Figure 4.5. Previously published *Tc-bwl* and *Tc-sob in situ* hybridisations used probes overlapping the highly similar zinc finger coding regions. *Tc-bwl* and *Tc-sob* CDS overviews illustrate the regions selected for probe synthesis by Thümecke and Schröder (2021) in relation to ZF coding sequences (blue blocks). A 239 bp region from the *Tc-bwl* probe with high sequence similarity to other *odd*-family genes was used as a template for sequence alignments against *Tc-odd* and *Tc-sob*, while the entire *Tc-sob* probe was used as a template for sequence alignments against *Tc-odd* and *Tc-bwl*. The first 171 nucleotides (red star) of the *Tc-bwl* probe has a particularly high sequence similarity to *Tc-odd* and *Tc-sob* sequences. Asterisks mark matching bases between probes and the gene sequences, with the top and bottom rows of asterisks corresponding to the middle and bottom sequences respectively. Pairwise identities are given as a percentage of matching bases between the probe and selected gene sequence. Numbers at the start of each row for gene sequences indicate the first base's position in the CDS. Note that the sense strand was used for alignments and measuring similarity. This similarity is directly related to how complementary the antisense RNA probes are to the mRNA for each gene, and may confer binding affinity. Probe sequences were deduced from primer sequences provided in Thümecke and Schröder, (2021 (section 4.3.1)).

The *Tc-bwl* probe used in experiments by Thümecke and Schröder (2021) was 799bp long, with a 239 bp region that shares 81.22% and 80.66% pairwise identity and almost no gaps in the alignment with *Tc-odd* and *Tc-sob* respectively (**Fig. 4.5A**). The first 171 nucleotides of the *Tc-bwl* probe have a particularly high similarity with *Tc-odd* and *Tc-sob*, sharing 87.86% and 85.55% sequence identity respectively (**Fig. 4.5A red star**). For reference, the remaining *Tc-bwl* sequence had a pairwise identity of 49.86% and 52% against *Tc-odd* and *Tc-sob* respectively and a fragmented alignment (data not shown).

The *Tc-sob* probe used by Thümecke and Schröder (2021) was 236 nucleotides in length, and was contained entirely within the conserved region encoding the zinc fingers (**Fig. 4.5B**). The whole length of probe had a pairwise identity of 87.82% and 85.65% against sequences from *Tc-odd* and *Tc-bwl* respectively.

Previously reported expression patterns for *Tc-bwl* and *Tc-sob* are consistent with cross-hybridisation against *Tc-odd* and *Tc-sob/Tc-bwl* mRNA

With these levels of sequence similarity, it is possible that the *Tc-bwl* probe used by Thümecke and Schröder (2021) could bind to *Tc-odd* and/or *Tc-sob* transcripts, while the *Tc-sob* probe could bind to *Tc-odd* and/or *Tc-bwl* transcripts in cells expressing these genes. The expression patterns reported for these two genes are consistent with this. Thümecke and Schröder (2021) observed dynamic stripes of *Tc-bwl* and *Tc-sob* in the SAZ during early stages of segment addition, while conversely, this was never detected in the HCR results presented here. These stripes are highly reminiscent of *Tc-odd* expression. When anterior stripes are also accounted for, germband expression patterns for *Tc-sob* and *Tc-bwl* from Thümecke and Schröder (2021, fig. 3C, fig. 4B) accurately reflect that of *Tc-sob/Tc-odd* double FISH embryos of a comparative stage (**Fig. 4.6**). Furthermore, *Tc-sob* leg expression from Thümecke and Schröder (2021, Fig. 3H') includes faint background staining that is reminiscent of *Tc-bwl* leg expression, while their *Tc-bwl* results, in addition to diffuse staining, show stripes of expression in the legs that are reminiscent of *Tc-odd/Tc-sob* leg joint domains at this stage (**Fig. 4.6 asterisks**).

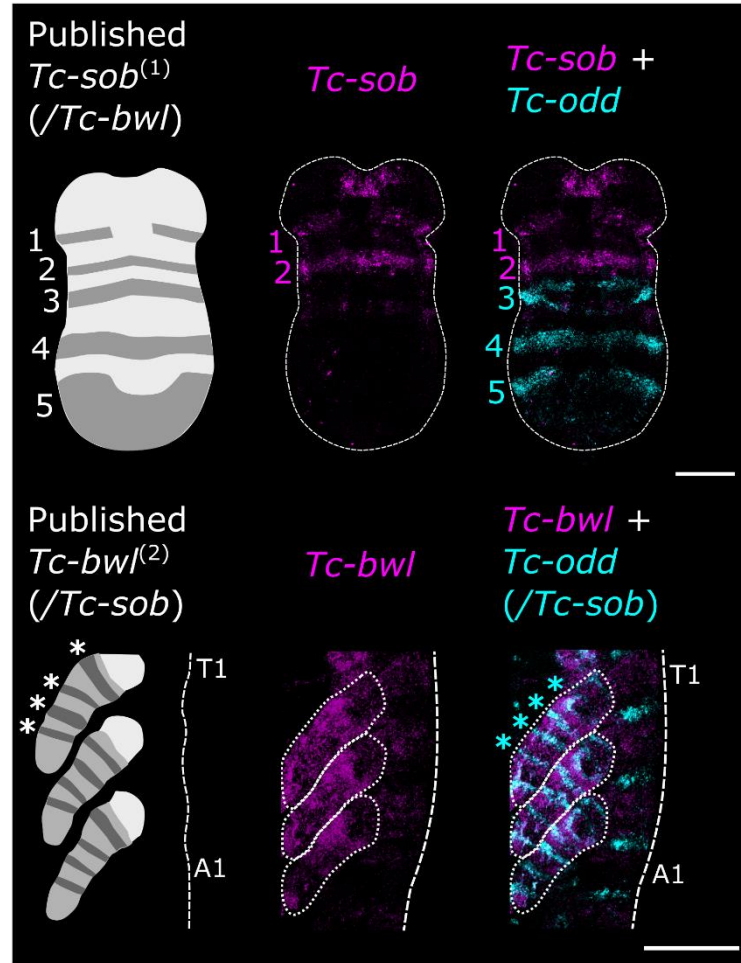


Figure 4.6. Comparison of previously published *Tc-sob* and *Tc-bwl* expression patterns to those from this study. Diagrams were drawn from *in situ* figure images in Thümecke and Schröder (2021) showing *Tc-sob* expression in an early stage germband ((1) Fig. 3C) and *Tc-bwl* expression in the developing legs ((2), Fig. 4I). These patterns are similar to those reported for *Tc-bwl* and *Tc-sob*, respectively, from the same study. Diagrams are shown alongside HCR FISH embryos from my experiments at comparative stages. In both cases, published expression patterns more closely resemble those of *Tc-sob/Tc-odd* and *Tc-bwl/Tc-odd* double FISHs. Stripes 1 and 2 in the early germband are akin to *Tc-sob* stripes, while 3-5 match those of *Tc-odd* expression. Asterisks mark elevated stripes of expression in previously published *Tc-bwl* results that resemble those of *Tc-odd* (or *Tc-sob*) leg joint domains. T=Thoracic, A=Abdominal. Scale bar: 100 μ m

Taken together, there is convincing evidence that the *Tc-bwl* and *Tc-sob* probes designed by Thümecke and Schröder (2021) are both capable of binding to *Tc-bwl*, *Tc-sob* and *Tc-odd* mRNA. *Tc-bwl* expression was reported to be weaker, requiring a longer staining reaction. This might be explained by a lower binding affinity between the *Tc-bwl* probe

and *Tc-odd/Tc-sob* mRNA, consistent with a smaller proportion of probe that overlaps the conserved zinc finger domain. Regardless, these findings cast doubt over the accuracy of *Tc-bwl* and *Tc-sob* expression patterns reported by Thümecke and Schröder (2021) and highlight the importance of careful probe design.

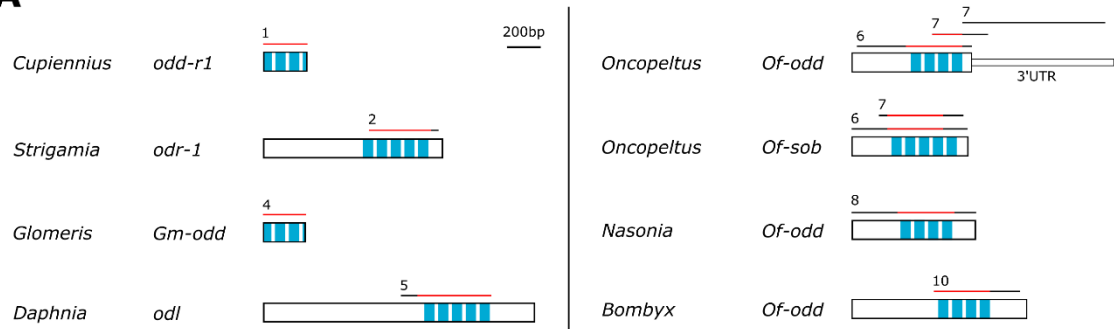
While the *Tc-drm* probe used by Thümecke and Schröder (2021) also overlapped with the zinc finger domains, this region shares a much lower sequence similarity to that of other *odd* family genes in *Tribolium*. Furthermore, there was no evidence from their results to suggest any probe cross-hybridisation with non-target mRNA. However, Thümecke and Schröder (2021) failed to capture the same pair-rule expression of *Tc-drm* observed here, with the exception of one stripe preceding the first *Tc-wg* stripe (*wg0*). In my experiments, the signal of *Tc-drm* pair-rule stripes during segment addition was relatively low in comparison to other domains such as in the hindgut anlage. As Thümecke and Schröder (2021) used colorimetric ISH, it is possible that the staining reaction was stopped before these weaker pair-rule stripes became visible. Furthermore, HCR has a much higher sensitivity than traditional antibody staining techniques to the point that it can detect single molecules within thick tissue samples (Shah et.al., 2016).

4.3.2. *in situ* probes used in *odd* family expression studies of other arthropods commonly include regions with high sequence similarity

In light of the evidence for cross hybridisation of *Tc-sob* and *Tc-bwl* probes used in Thümecke and Schröder (2021), I reviewed the probe designs for expression studies of *odd* family genes in other arthropod species. For almost all publications from which probe sequences could be inferred, significant regions of these probes overlap the zinc finger coding domains, and share significant similarity to the sequences of other *odd* family genes from the same species (**Fig. 4.7**). Given that a 239bp region of the *Tc-bwl* probe used by Thümecke and Schröder (2021) with ~80% similarity to *Tc-odd* and *Tc-sob* was enough to illicit cross hybridisation, it is feasible that many of these other gene expression studies could be reporting expression patterns for more than one *odd*-family gene.

For example, two separate papers have published *Of-sob* expression in the milkweed bug, *Oncopeltus fasciatus*, describing an almost identical expression pattern to that of *Of-odd* (Auman and Chipman, 2018; Reding et.al., 2019). In both cases, probes included a highly conserved 338bp region of the zinc finger domain that shares 98.22% sequence similarity to the corresponding region of *Of-odd* (**Fig. 4.7**). Such a high sequence similarity would likely equate to a high binding affinity between the *Of-sob* probe and *Of-odd* mRNA, and could therefore explain the similarity in expression patterns reported for both genes. These papers also used *odd* probes that overlap this highly similar region. However, Reding et.al. (2019) used an additional control probe outside of this sequence to confirm that cross hybridisation was not influencing their results for *Of-odd* expression. Published expression patterns for other arthropod *odd* family genes where probes include homologous ZF-coding regions with over 80% sequence similarity to paralogues from the same species include: *Strigamia odr1* (with *odr2*), *Daphnia odl* (with *bwl-like-2*), and *Nasonia Nv-odd* (with *Nv-sob* and *Nv-bwl*) (**Fig 4.7B**) (Chipman et.al., 2004; Eriksson et.al., 2013; Rosenberg et.al., 2014). Several other expression studies include probes with regions of 75-80% sequence similarity to paralogues, while in some cases probe sequences could not be inferred from the methods provided, therefore leaving it unclear what capacity these probes have for off target cross-hybridisation (**Fig 4.7B**).

Whatever the case, these findings highlight the importance of careful probe design, and cast a degree of uncertainty over the accuracy of published expression patterns for *odd*-family genes in arthropods. If the expression patterns of alternating pair-rule stripes observed for *Tc-sob* and *Tc-odd* in the segmenting *Tribolium* germband are conserved across other arthropods, cross-hybridisation of probes could theoretically lead to incorrect conclusions about single or double segment periodicity.

A**B**

Species	Gene	Probe(s): Total size	overlapping ZF domain?	ZF region overlap: size, similarity with (%ID)
<i>Cupiennius salei</i>	<i>odd-r1</i>	263bp ¹	Yes	263bp, <i>odd-r2</i> (76.43%) 263bp, <i>odd-r3</i> (77.19%)
<i>Strigamia maritima</i>	<i>odr1</i>	442bp ²	Yes	369bp, <i>odr2</i> (84.01%) 99bp, <i>odr3</i> (74.75%)
	<i>odr2</i>	? ³	-	-
	<i>odr3</i>	? ³	-	-
<i>Glomeris marginata</i>	<i>Gm-odd</i>	255bp ⁴	Yes	**
<i>Daphnia magna</i>	<i>Odl</i>	537bp ⁵	Yes	438bp, <i>bwl-like-2</i> (88.36%)
<i>Oncopeltus fasciatus</i>	<i>Of-odd</i>	681bp, ⁶	Yes	338bp, <i>Of-sob</i> (98.22%)
		~350bp* ⁷	Yes	188bp, <i>Of-sob</i> (96.81%)
		~850bp* ⁷	No	-
	<i>Of-sob</i>	692bp ⁶ 502bp ⁷	Yes Yes	338bp, <i>Of-odd</i> (98.22%) 338bp, <i>Of-odd</i> (98.22%)
<i>Nasonia vitripennis</i>	<i>Nv-odd</i>	747bp ⁸	Yes	339bp, <i>Nv-sob</i> (91.45%) 339bp, <i>Nv-bwl</i> (87.61%)
<i>Dermestes maculatus</i>	<i>Dmac-odd</i>	? ⁹	-	-
<i>Bombyx mori</i>	<i>Bm-odd</i>	512bp ¹⁰	Yes	338bp, <i>Bm-sob</i> (77.51%) 338bp, <i>Bm-bwl</i> (78.99%)

Figure 4.7. Review of *odd* family *in situ* probes from published expression studies in arthropods. (A) Schematics of arthropod *odd*-family genes, for which expression patterns have been described, aligned with the RNA probes used in each case. Blue squares indicate C2H2 ZF coding sequences, while sections of probe overlapping sequence with high similarity to other paralogues from the same species are coloured in red. Numbers relate to the papers from which the probe sequences have been inferred (see references below). (B) Table detailing probe lengths, overlap of high similarity ZF coding domain, and the size and similarity of this region to sequences of paralogues from the same species. In papers where primers or probe sequences are provided, NCBI conserved domain searches were made to identify the presence/absence of the zinc finger (ZF) domain. If probe sequences overlapped the ZF domain, they were aligned against other known *odd* family genes for that species, conserved regions were identified, and pairwise similarities (%ID) were calculated (as in **fig. 4.5**). The conserved region alignments from which similarities are provided share 100% sequence coverage. References: 1 = Damen et.al., 2005, 2 = Chipman et.al., 2004, 3 = Green and Akam, 2013, 4 = Janssen et.al., 2011, 5 = Eriksson et.al., 2013, 6 = Auman and Chipman, 2018, 7 = Reding et.al., 2019, 8 = Rosenberg et.al., 2014, 9 = Xiang et.al., 2017, 10 = Nakao, 2015. “?” indicates where probe sequences could not be inferred from the publication due to exclusion of primer sequences. Single asterisks denote probe sizes with uncertainty. Double asterisks indicate where other *odd*-family gene sequences are unknown for that species. Also note that *Of-odd* and *Of-sob* are the only genes identified for *Ofas*.

4.3.3. Revised expression patterns do not support redundancy between *odd* family genes in the *Tribolium* segmentation clock network

While the late pair-rule stripes of *Tc-sob* and *Tc-drm* expression in *Tribolium* may be related to roles in segmentation for these genes, the lack of any expression within the posterior SAZ during segment addition means that they are unlikely to be involved in primary segment patterning. This role appears to be reserved for *Tc-odd* alone. If any of the other *Tribolium odd*-family genes were acting redundantly in the oscillatory network to drive *Tc-run* stripe expression in the absence of *Tc-odd* (**Fig. 1.9C**), then we would expect to see these genes co-expressed in the SAZ, however, it is now clear that this is not the case.

Despite not being active players in driving primary pair-rule oscillations, *Tc-sob* and *Tc-drm* expression patterns may indicate later roles in segment patterning. Their stripes of expression appear immediately before adjacent stripes of the segment polarity gene, *Tc-wg*. It is therefore possible that *Tc-drm* and *Tc-sob* function falls somewhere within the regulatory network that promotes segment-polarity gene expression. It has been shown from *Tc-odd* knockdowns in *Tribolium*, that *Tc-odd* is required for the activation of *Tc-prd* stripes in odd-numbered segments, which in turn are required for the proper formation of these segments via activation of the segment polarity gene, *Tc-en* (Choe et.al., 2007; Jeon et.al., 2020). This explains the pair-rule-like phenotype that is observed in some *Tc-odd* knockdowns where cellular oscillations and primary segment patterning are still preserved (Jeon et.al., 2020). It is therefore possible, that while *Tc-odd* stripes are required for the proper development of odd-numbered segments, *Tc-sob* and *Tc-drm* may be required for the development of even-numbered segments. However, the primary stripes of *Tc-odd* are expressed in even-numbered segments, while the primary stripes of *Tc-sob* and *Tc-drm* are co-expressed in odd-numbered segments meaning that if such roles do exist, they are likely to be indirect and complex. To date, single *Tc-drm* knockdown experiments have not been reported to illicit any segmentation defects (Thümecke and Schröder, 2021). However, it would be necessary to perform joint knockdowns of *Tc-sob* and *Tc-drm* to fully test this given that their overlapping expression may confer redundancy.

5. EXPRESSION DYNAMICS OF *HES/HEY* GENES IN *TRIBOLIUM*

5.1. INTRODUCTION

5.1.1. *Hes/Hey* gene family expression and function in arthropods

The *Hes/Hey* family of genes encode a group of bHLH-Orange (bHLH-O) transcription factors that have ancient roles in regulating developmental and physiological processes, such as neurogenesis and circadian rhythms, across a diverse range of metazoans. In *Drosophila*, the *Hes* gene, *hairy* (*h*), was first identified for its function as a primary pair-rule gene during segmentation (Holmgren, 1984; Ingham et.al., 1988). Due to its classification as a pair-rule gene, candidate gene approaches in other arthropod models have often focused on investigating expression and function of *h* orthologues, leaving data scarce for other *Hes/Hey* family genes outside of *Drosophila*.

Hes genes

As discussed in chapter 3, a group of arthropod *Hes* genes related to *h/deadpan* (*dpn*) are commonly found to exhibit expression patterns consistent with roles in segmentation, and represent the primary candidates for an alternate oscillator network in the arthropod segmentation clock. Expression patterns for *h/dpn*-related genes in the SAZ of spiders, *Cupiennius* and *Parasteatoda*, and the centipede, *Strigamia*, are consistent with roles in primary segment patterning in non-insect arthropods, and in the case of *Cupiennius*, appear to be regulated by Notch signalling (Stollewark et.al., 2003; Green and Akam, 2013; Paese, 2018). In the cockroach, *Periplaneta*, *Pa-h* exhibits dynamic expression in the SAZ, and functional evidence suggests that it acts in concert with Notch signalling to drive and co-ordinate cellular oscillations during segmentation (Stollewark et.al., 2003; Pueyo et.al., 2008). However, despite having similarly conserved expression patterns, there is contrasting functional evidence for roles of *h* orthologues in the segmentation of other sequentially segmented insects, such as

Tribolium and *Oncopeltus* (Aranda et.al., 2008; Reding et.al., 2019). While no segmentation function has thus far been described for the insect *dpn* genes, orthologues in *Drosophila* and *Tribolium* are both expressed in primary pair-rule stripes during segment patterning (Bier et.al., 1992; Aranda et.al., 2008). In *Drosophila*, *dpn* is also expressed in the neuroectoderm, where it is believed to function as a downstream target of the homeodomain transcription factor, Prospero, in regulating neural stem cell renewal and differentiation (Southall and Brand, 2009).

Another sub-group of arthropod *Hes* genes, the *Enhancer-of-split* (*E(spl)*) genes, are primarily associated with roles in neurogenesis. In *Drosophila* and *Tribolium*, they have a conserved function within a Notch-mediated lateral inhibition mechanism that controls the patterning of neuroblasts (neural stem cells) (Culí and Modolell, 1998; Kux et.al., 2013). As established from phylogenetic analyses in Chapter 3, however, *E(spl)* homologues in the crustacean, *Daphnia*, are interestingly found to be expressed in patterns consistent with roles in primary segment patterning, which in this case appears to be co-ordinated by Notch signalling (Eriksson et.al., 2013). A limited availability of *E(spl)* expression data in other arthropods means we cannot confirm if this phenomenon is widespread.

Other *Hes* genes, such as the insect *Similar to deadpan* (*Sidpn*) genes, have not been the subject of any significant research to date. *In situ* data from the Berkeley *Drosophila* Genome Project shows that in *Drosophila*, *Sidpn* is expressed in the developing brain and ventral nerve cord during neurogenesis (Tomancak et.al., 2007).

Hey and clockwork orange (cwo)

Hes-related bHLH-O genes, *cwo* and *Hey*, are essential for their roles in controlling circadian rhythms and neurogenesis, respectively. In *Drosophila*, *cwo* oscillates via an autoinhibitory feedback loop whilst directly regulating other genes in the circadian clock network to confer behavioural rhythmicity (Matsumoto et.al., 2007; Kadener et.al., 2007). Evidence from several other insect models points to a conserved function of *cwo* in this system (Ingram et.al., 2012; Tomiyama et.al., 2020). With regards to development, there is also some evidence from *Drosophila* for a role of *cwo* in dendrite morphogenesis (Iyer et.al., 2013). *Hey* is expressed in the developing nervous system

and, as a downstream target of Notch, it promotes alternate cell fates in newly born neurons (Monastirioti et.al., 2010). Other arthropod *Hey* orthologues have not yet been the subject of expression or functional studies, and while there is little evidence from *Drosophila* to suggest that they may play a role in segmentation, vertebrate homologues of *Hey* are cyclically expressed in the chick PSM, demonstrating their potential for co-option into an analogous clock-and-wavefront mechanism (Leimeister et.al., 2000).

Extramacrochaetae (*emc*)

While not strictly a member of the *Hes/Hey* gene family, *emc* is an important regulator of bHLH transcription factor activity. It encodes a bHLH protein that lacks the basic domain required for DNA binding, but that has retained the helix-loop-helix domain that enables it to dimerize with other bHLH proteins (Campuzano, 2001). This means that Emc can repress bHLH transcription factor activity through the formation of inactive heterodimers that are unable to bind DNA and regulate transcription. In vertebrates, homologues of Emc, Id proteins, behave in this capacity with HES/HEY monomers in the context of neurogenesis, releasing the chick *Hes1* gene from autorepressive oscillations (Bai et.al., 2007). While *emc* does not appear to function in *Drosophila* segmentation, in the context of eye development there is evidence to suggest that Emc antagonizes E(spl) function, indicating that the Emc protein is capable of dimerising HES/HEY proteins in arthropods (Bhattacharya et.al., 2009). Given the potential of *emc* to influence HES/HEY protein dynamics, this makes this gene a strong candidate for investigating *Hes/Hey* function in relation to sequential segmentation.

Published *Hes/Hey* expression patterns in *Tribolium* development

As described in chapter 2, there are 8 known *Hes/Hey* genes and 1 *emc* gene in *Tribolium*. To date, embryonic expression patterns have been reported for *Tc-h*, *Tc-dpn*, *Tc-E(spl)1* and *Tc-E(spl)3* (Aranda et.al., 2008; Kux et.al., 2013), leaving *Tc-Sidpn1*, *Tc-Sidpn2*, *Tc-Hey*, *Tc-cwo* and *Tc-emc* expression unexplored. *Tc-h* and *Tc-dpn* both exhibit dynamic stripes of expression emanating from the posterior SAZ, while *Tc-E(spl)1* and *Tc-E(spl)3* are expressed in the neuroectoderm during GBE, consistent with their conserved roles in neuroblast patterning (Aanda et.al., 2008; Kux et.al., 2013).

5.1.2. Aims of the chapter

In this chapter, I aim to screen the expression patterns of *Tribolium Hes/Hey* genes and *Tc-emc* during segmentation, in order to corroborate previous studies and to describe novel results in the hope of identifying new candidate genes for segmentation. As an ultimate goal of this thesis is to establish whether an alternate *Hes/Hey*-based oscillator network may be operating in the *Tribolium* segmentation clock, any genes that exhibit dynamic expression in the SAZ are of immediate interest. For genes expressed in this way, such as *Tc-h* and *Tc-dpn*, I will perform multiplex HCR FISH to establish the relative order of expression, and the extent (if any) to which domains overlap. To this end, I aim to generate a detailed analysis of expression dynamics during segmentation as to provide additional context for future research.

5.2. RESULTS

5.2.1. *Hes/Hey* family gene expression screen

For the purpose of screening gene expression patterns, colorimetric *in situ* hybridisation (ISH) experiments were performed. In each case, the gene of interest was stained in blue against a red stained *Tc-wg* marker. Germband staging follows the same logic as described in the previous chapter (**section 4.2.1; Fig. 4.1**). For genes whose expression patterns have not been previously described, sense strand negative controls were performed in parallel to confirm novel results (**Appendix F; Fig. F.1**)

Tc-h, *Tc-dpn*, and *Tc-emc* all display striped expression in the SAZ during GBE.

Colorimetric *in situ* hybridisations (ISH) confirm previously published data for *Tc-h* and *Tc-dpn* expression during segmentation (Aranda et.al, 2008), and provide novel results for *Tc-emc* expression. These three genes all exhibit stripes of gene expression in the SAZ throughout almost all of segment addition, up until around wg12 to wg13 (**Figures 5.1, 5.2 and 5.3 (black arrowheads)**). In all cases, this expression is transient, and stripes are visible only in the SAZ, with the exception of a single *Tc-h* stripe that persists around the future mandibular segment from stages wg0 to wg5 (**Fig. 5.1 green arrowhead**). In addition to this, all three genes are also expressed in other regions during *Tribolium* GBE.

Tc-h expression is visible in the mesoderm, from the head to the tail, for most of axial elongation. Once striped expression of *Tc-h* ceases, from around stage wg12, a posterior *Tc-h* domain is established that overlaps the posterior *Tc-wg* domain, and remains in place post GBE (**Fig. 5.1 red arrowheads**). *Tc-h* also becomes expressed in the limb buds in late-stage embryos (**Fig. 5.1 blue arrowheads**). Strong *Tc-dpn* expression appears in the neuroectoderm from around stage wg6, and spreads as the nervous system develops (**Fig. 5.2**). This expression appears speckled and is reminiscent of neuroblast markers such as *Tc-asense* (Wheeler et.al., 2003). As the posterior *Tc-wg* domain changes shape and moves anteriorly, at around wg6, *Tc-dpn* also becomes expressed in

the posterior tip of the growth zone behind *Tc-wg* (**Fig. 5.2 red arrowheads**). *Tc-dpn* expression persists here throughout the rest of GBE and into later developmental stages.

Tc-emc ISH produced rapid ubiquitous staining, suggesting that it is expressed in all cells of the *Tribolium* germband. However, there are clear stripes of elevated *Tc-emc* expression in the SAZ during GBE (**Fig. 5.3 black arrowheads**). *Tc-emc* also appears to be expressed at higher levels in the posterior mesoderm during the early stages of GBE, until around wg2 (**Fig. 5.3**). Similarly to *Tc-dpn*, speckles of elevated expression were observed in the neuroectoderm in later stage germbands. Strong *Tc-emc* domains also appeared in the centre of the head, the limb buds, and the posterior end of the embryo immediately after the completion of GBE (**Fig. 5.3**).

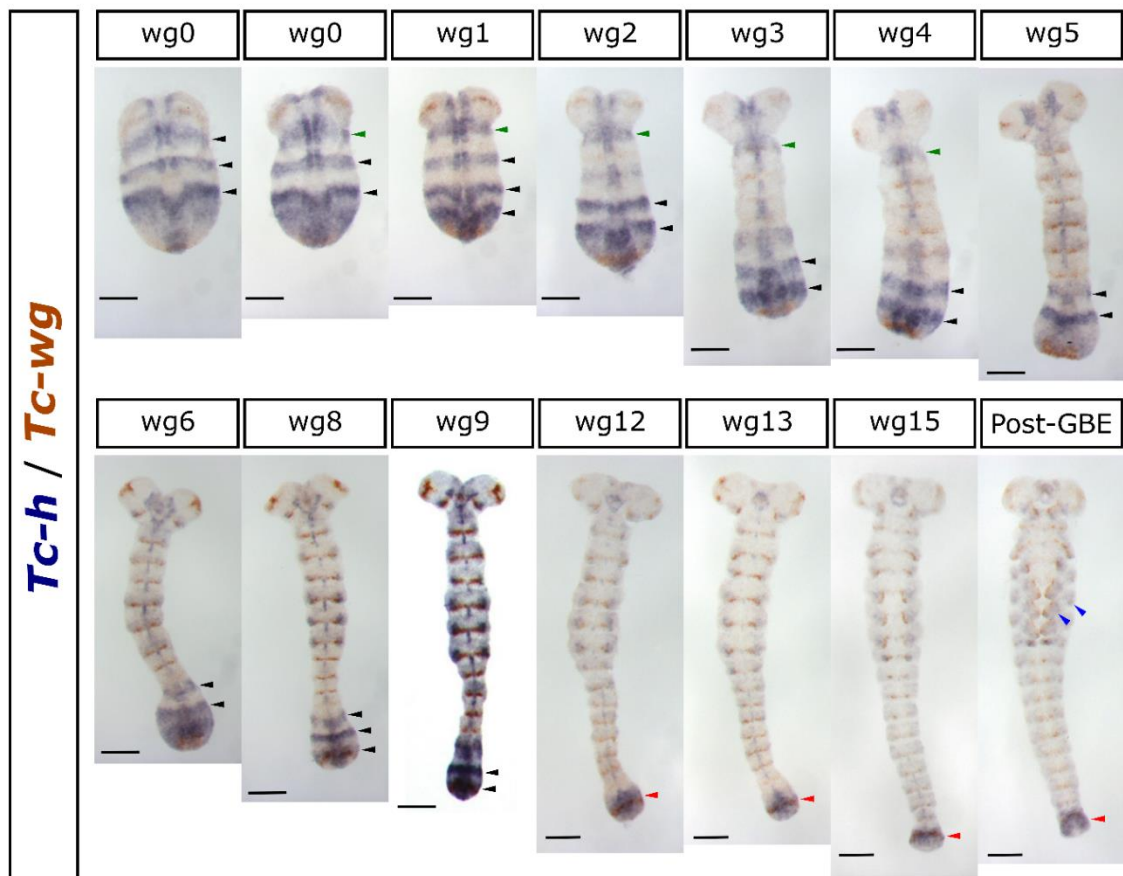


Figure 5.1. *Tc-h* expression in *Tribolium*. Colorimetric ISH showing *Tc-h* (blue) and *Tc-wg* (red) expression in *Tribolium* germbands over the course of segment addition. Arrowheads indicate *Tc-h* stripes in the SAZ (black), a sustained *Tc-h* mandibular domain (green), *Tc-h* posterior domain (red), and *Tc-h* expression in the limb buds (blue). Wg9 image from Sharma (pers. Comms) using the same RNA probes. Scale bar: 100 μ m

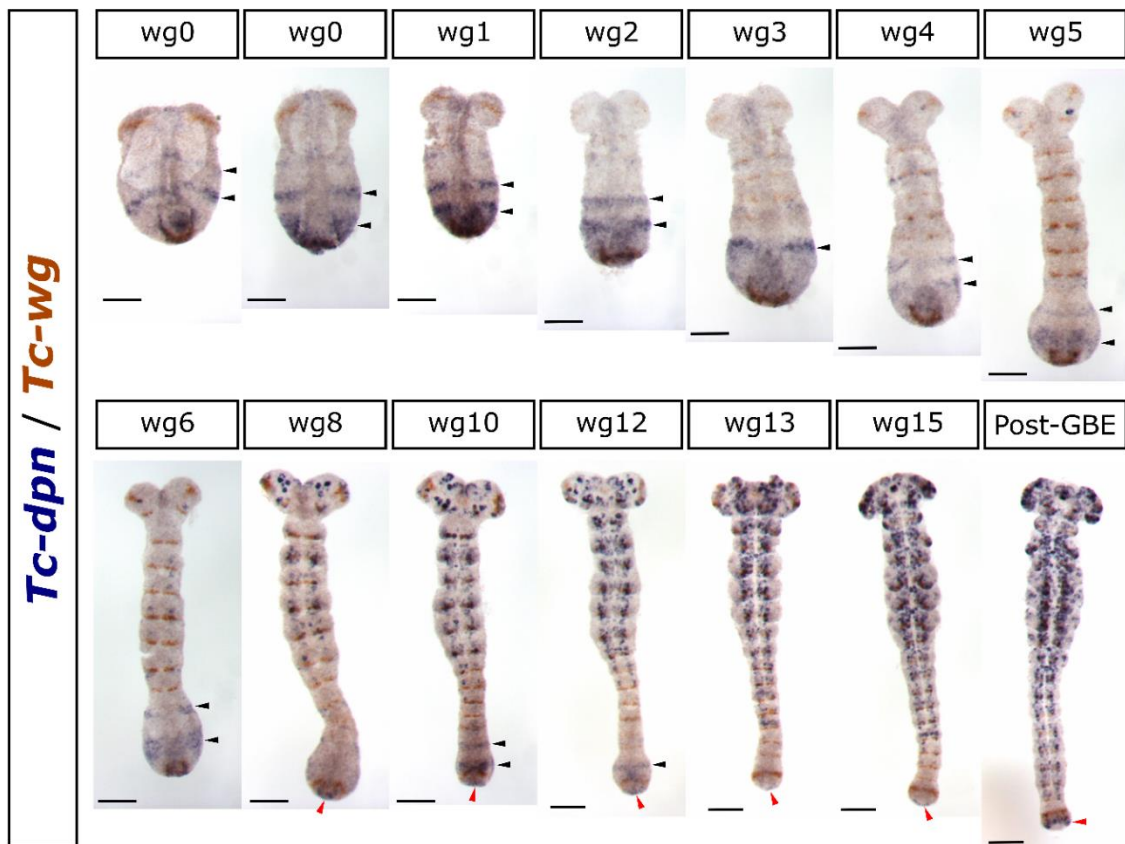


Figure 5.2. *Tc-dpn* expression in *Tribolium*. Colorimetric ISH showing *Tc-dpn* (blue) and *Tc-wg* (red) expression in *Tribolium* germbands over the course of segment addition. Arrowheads indicate *Tc-dpn* stripes in the SAZ (black) and a posterior domain (red). Scale bar: 100 μ m

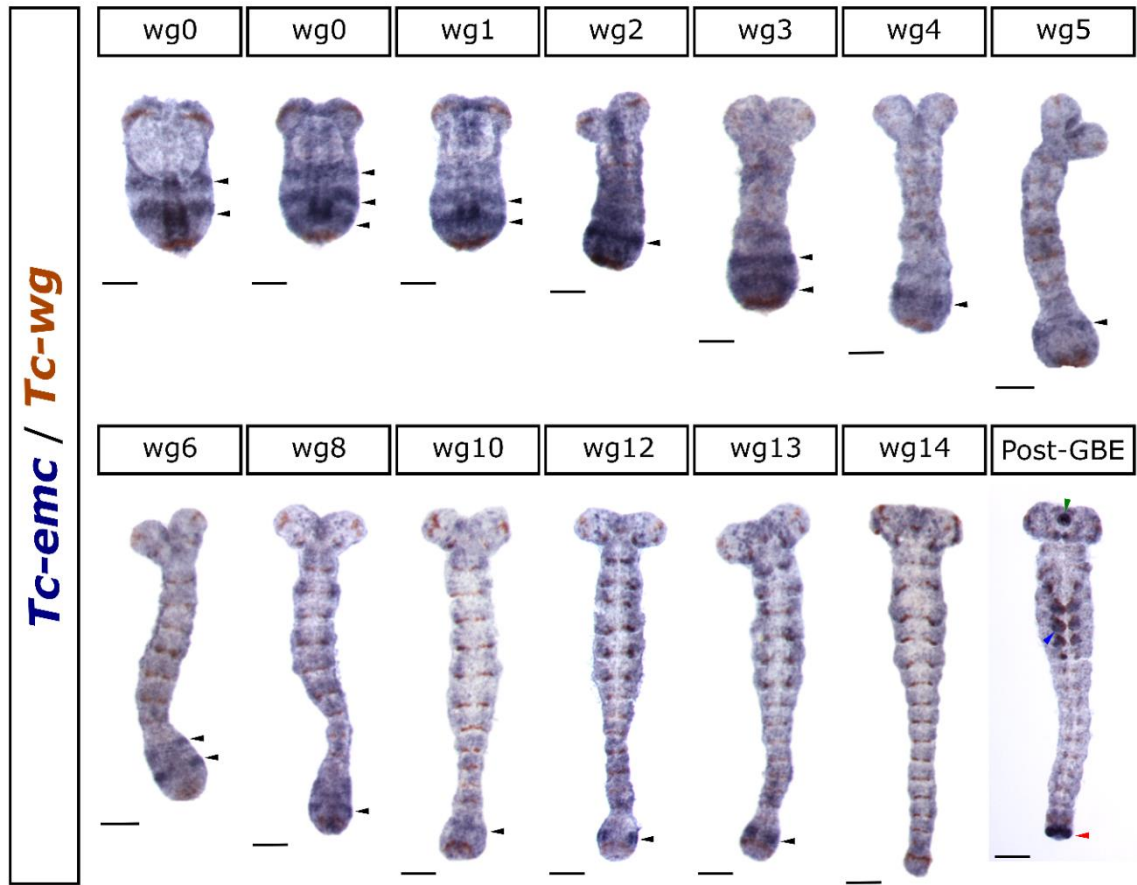


Figure 5.3. *Tc-emc* expression in *Tribolium*. Colorimetric ISH showing *Tc-emc* (blue) and *Tc-wg* (red) expression in *Tribolium* germbands over the course of segment addition. Arrowheads indicate *Tc-h* stripes in the SAZ (black), and elevated *Tc-emc* expression in the centre of the head (green), limb buds (blue), and posterior tip (red) of a post-GBE embryo. Scale bar: 100 μ m

Other *Tribolium* *Hes/Hey* genes are expressed in the neuroectoderm and posterior domains during segment addition

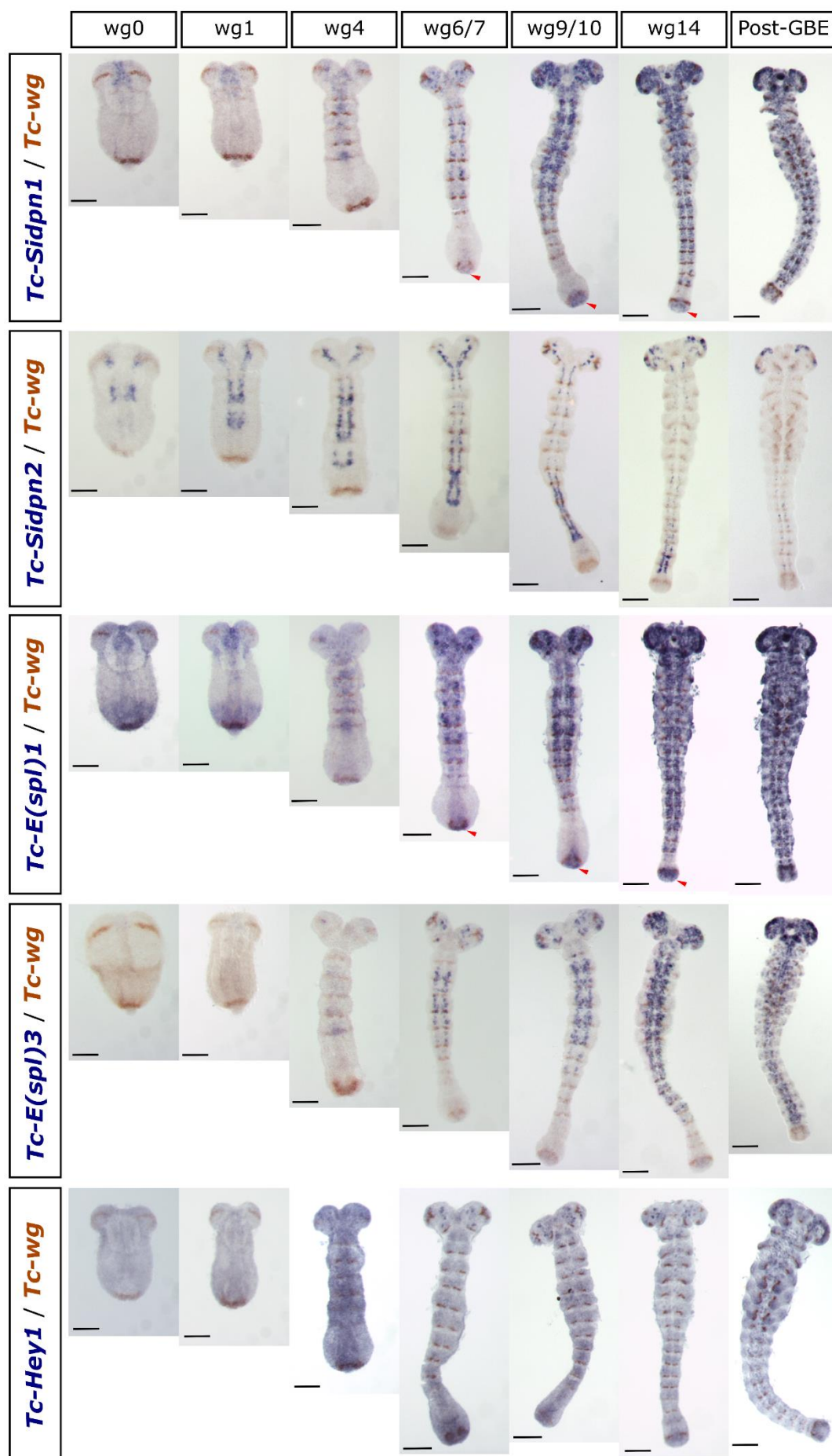
Tc-Sidpn1, *Tc-Sidpn2*, *Tc-Hey*, *Tc-E(spl)1* and *Tc-E(spl)3*, are all expressed in the neuroectoderm at varying stages during GBE (**Fig. 5.4**). *Tc-E(spl)1* and *Tc-E(spl)3* expression patterns reported here corroborate those published by Kux et.al. (2013). *Tc-Sidpn1* and *Tc-E(spl)1* follow similar patterns and timing of expression in the neuroectoderm, first appearing in the head at stage wg0, with progressively more

expression throughout the neuroectoderm as development proceeds (**Fig. 5.4**). Towards the end of GBE, *E(spl)1* becomes almost ubiquitously expressed. At around stage wg6, *Tc-Sidpn1* and *Tc-E(spl)1* expression appears behind the posterior *Tc-wg* domain, where it persists through to the end of GBE (**Fig. 5.4 red arrowheads**). *Tc-E(spl)1* is also transiently expressed in a broad posterior domain in early germband stages wg0 and wg1.

Transient expression of *Tc-Sidpn2* appears in the neuroectoderm of newly patterned segments throughout segment addition (**Fig. 5.4**). While neural expression in the trunk disappears, some *Tc-Sidpn2* expression is retained in the head lobes later into development. *Tc-E(spl)3* follows a very similar pattern and timing of neuroectodermal expression as *Tc-dpn*, first appearing at around wg5/wg6, and then spreading throughout the nervous system in older embryos (**Fig. 5.4**). *Tc-Hey* expression also appears at the wg6/wg7 stage in the neuroectoderm, however, at first is restricted to small dots in the head (**Fig. 5.4**). Expression of *Tc-Hey* only becomes visible in the trunk in later-stage embryos towards the end of GBE. Faint ubiquitous staining was also visible for *Tc-Hey* throughout the embryonic stages that were collected.

Tc-cwo produced faint ubiquitous staining, and no signs of any specific expression were detected in embryos up to 48 hours old (data not shown). Faint ubiquitous staining was also observed with a sense probe control, therefore making it unclear whether this reflects real expression or not.

Figure 5.4 (overleaf). *Tc-Sidpn1*, *Tc-Sidpn2*, *Tc-E(spl)1*, *Tc-E(spl)3* and *Tc-Hey1* expression during GBE. Colorimetric ISH showing the expression of *Hes* genes (blue), *Tc-Sidpn1*, *Tc-Sidpn2*, *Tc-E(spl)1*, *Tc-E(spl)3*, and *Tc-Hey1*, against a *Tc-wg* (red) marker at different stages of *Tribolium* GBE. Images correspond to the stages labelled along the top of the figure, and to the genes labelled along the left-hand side. While comparative, in some cases, the exact same stage was not captured for all genes. Wg6 stages are given for *Tc-Sidpn1*, *Tc-Sidpn2* and *Tc-E(spl)1*, while wg7 stages are given for *Tc-E(spl)3* and *Tc-Hey1*. Wg10 stages were used for all genes except *Tc-E(spl)1*, for which the closest stage captured was wg9. Scale bar: 100 μ m



5.2.2. Co-expression analysis of striped *Hes* genes: *Tc-h*, *Tc-dpn* and *Tc-emc*

Tc-h, *Tc-dpn* and *Tc-emc* stripes first established in blastoderm stages prior to GBE

To find out when stripes of *Tc-h*, *Tc-dpn* and *Tc-emc* expression are first established, I performed hybridisation chain reaction (HCR) fluorescent *in situ* hybridisation (FISH) on *Tribolium* blastoderms with different combinations of *Tc-h*, *Tc-dpn*, or *Tc-emc*, using *Tc-wg* as a marker. While *Tc-wg* was included in these experiments, its expression was not sufficient to distinguish the earliest stages of development. Instead, I followed blastoderm staging described by Benton et.al. (2013), using DAPI stained nuclei to evaluate cell arrangements. The number of *Tc-h* stripes was also informative and allowed for direct comparisons between different experiments.

Hes gene expression is visible from stage 1 of *Tribolium* embryogenesis, when blastoderm cells are still uniformly distributed on the ventral side of the egg (**Fig. 5.5A-B**). *Tc-h* first appears in a transverse stripe in the middle of the blastoderm, while *Tc-dpn* is expressed in a broad posterior domain (**Fig. 5.5A**). At this point, *Tc-emc* is expressed in the anterior pole, and along the midline of the embryo (**Fig. 5.5B**). This midline expression of *Tc-emc* appears highly similar to that of *Tribolium* mesoderm markers, *Tc-twist* and *Tc-snail*, at these stages (Sommer and Tautz, 1994). This, coupled with the fact that *Tc-emc* is expressed in the mesoderm of early germbands (**Fig. 5.3, wg0-wg2**), would suggest that this *Tc-emc* domain is part of the presumptive mesoderm (**Fig. 5.5C**), and unrelated to the later stripes of ectodermal expression.

In stage 2 blastoderms, cells migrate posteriorly and towards the ventral midline (Benton et.al., 2013) (**Fig. 5.5B DAPI**). At this point, a second *Tc-h* stripe appears, overlapping the posterior *Tc-dpn* domain (**Fig. 5.5A,C**). At the same time, *Tc-emc* expression has almost disappeared from the anterior pole but continues to be expressed down the midline of the embryo in the presumptive mesoderm (**Fig. 5.5B**).

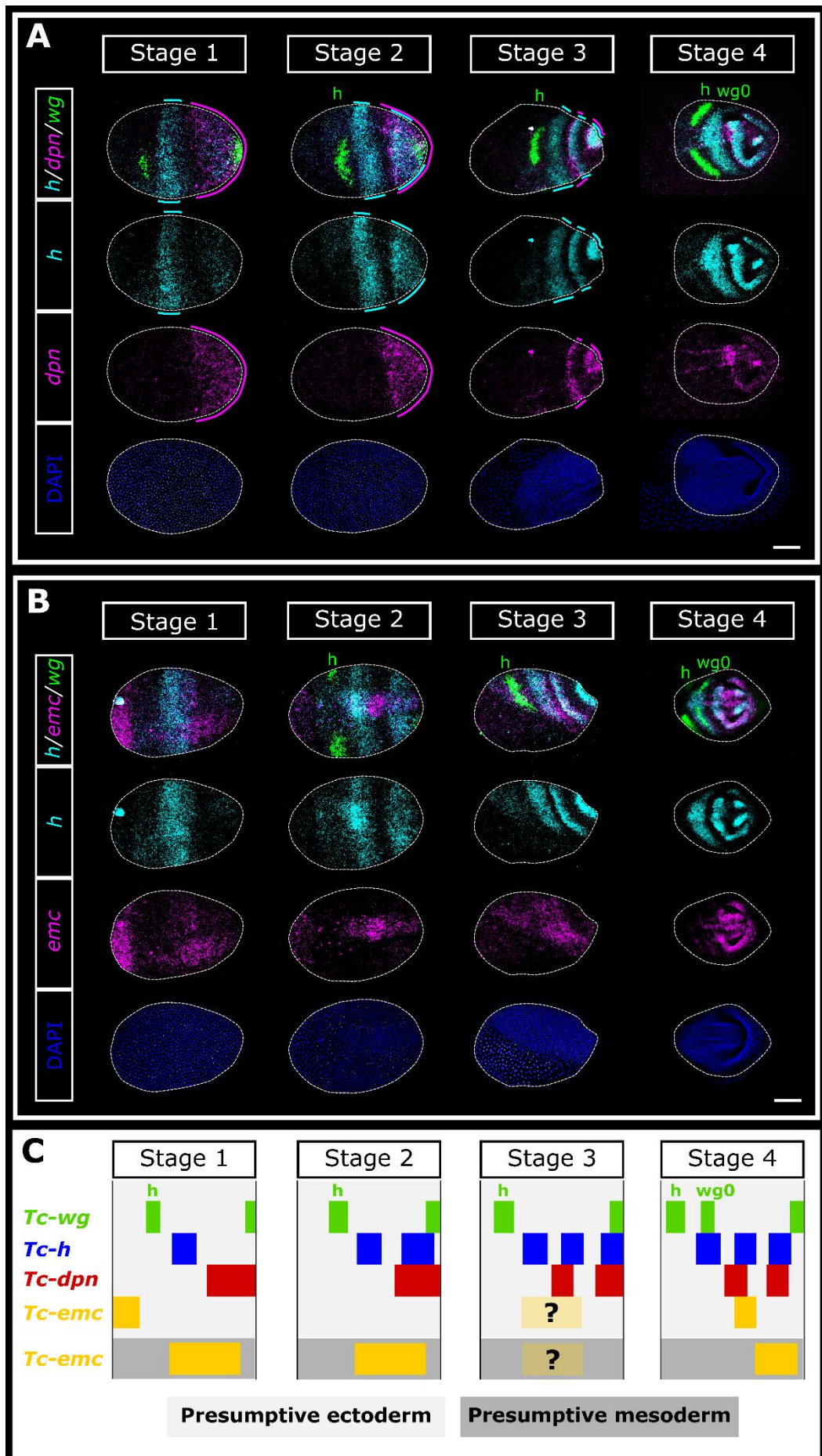


Figure 5.5. Early blastoderm expression of *Tc-h*, *Tc-dpn* and *Tc-emc*. (A) HCR FISH of *Tc-h*, *Tc-dpn*, and *Tc-wg* at blastoderm stages 1 to 3. Composites are given for each stage, along with individual channels for *Tc-h*, *Tc-dpn*, and DAPI nuclei stains. Lines along the edges highlight ranges of expression for each *Tc-h* and *Tc-dpn* stripe, and how they overlap in merged images. (B) HCR FISH of *Tc-h*, *Tc-emc* and *Tc-wg* at blastoderm stages 1 to 4. Composites are provided for each stage, along with individual channel images for *Tc-h*, *Tc-emc*, and DAPI nuclei stains. (C) Simplified schematics to illustrate relative expression of *Tc-wg* (green), *Tc-h* (blue), *Tc-dpn* (magenta), and *Tc-emc* (orange) at the first four stages of *Tribolium* development, based on embryos in (A) and (B). Separate panel along the bottom to distinguish *Tc-emc* expression in the presumptive mesoderm, from the presumptive ectoderm. Question marks indicate where *Tc-emc* expression could not be distinguished between the presumptive mesoderm and presumptive ectoderm at stage 3. Green “h” and “wg0” in (A-C) refer to the *Tc-wg* head and wg0 (PS0) stripes respectively. Embryos and schematics are all orientated with anterior to the left. Scale bar: 100 μ m

Stage 3 blastoderms are characterised by the initiation of germband condensation, and signs of invagination at the posterior pole as amnion folding begins (Benton et.al., 2013). At this stage, additional *Tc-dpn* and *Tc-h* stripes are resolved (**Fig. 5.5A**). These domains overlap, however, the leading edges of *Tc-dpn* stripes are slightly anterior to those of *Tc-h*. This relationship between *Tc-dpn* and *Tc-h* stripes continues for most of segment addition, until striped expression ceases at germband stages wg-12 to wg-13 (see below). *Tc-emc* expression was only captured from a lateral view at stage 3 (**Fig. 5.5B**). Based on expression immediately before and after this point in development, it is likely that *Tc-emc* is still expressed in the presumptive mesoderm. It also appears that *Tc-emc* is expressed in the ectoderm at this stage, overlapping the 2 anterior stripes of *Tc-h* expression. However, from this lateral view, it is difficult to distinguish the precise ectoderm pattern against expression in the presumptive mesoderm. (**Fig. 5.5B-C**). Note that at this stage, a degree of ubiquitous expression was already evident in the embryo proper. Laser power and digital gain was set to remove background “noise” and capture patterns of elevated expression when imaging *Tc-emc* from this point onwards.

At stage 4 of development, the germband begins to take shape, and amnion folding has progressed to produce a horse-shoe shaped fold (Benton et.al., 2013). By this point, a resolved *Tc-emc* stripe has been established that almost perfectly overlaps the second

Tc-h stripe (**Fig. 5.5B-C**). *Tc-emc* is still expressed in the developing mesoderm, but no other expression can be seen overlapping the posterior *Tc-h* stripe. *Tc-dpn* and *Tc-h* stripes retain their relative overlapping positions as described for stage 3.

While expression is dynamic, the relative positions of *Tc-h*, *Tc-dpn*, and *Tc-emc* stripes in the SAZ are consistent throughout segment addition

The striped expression of *Hes* genes in the *Tribolium* SAZ appears to be highly dynamic in nature. HCR FISH of these genes with *Tc-odd*, a known oscillatory gene (Sarrazin et.al., 2012), confirms this to be true. The relative positioning of *Tc-h* stripes to *Tc-odd* stripes in the SAZ remains consistent throughout segmentation (**Appendix B, Fig. B.1**), indicating that oscillations of *Hes* gene expression are synchronised with those of *Tc-odd*, *Tc-run*, and *Tc-eve* in the SAZ. By combining results from combinations of multiplex HCR FISH experiments, I was able to infer the relative positions of *Hes* gene stripes, against *Tc-odd* stripes in the *Tribolium* SAZ (**Fig. 5.6**). Results from varying stages of GBE indicate that the dynamics of gene expression remain consistent up until the point that striped *Tc-h*, *Tc-dpn*, and *Tc-emc* expression ceases at around wg12-wg13 (**Fig. 5.6A-B**).

Both *Tc-h* and *Tc-dpn* stripes begin as broad posterior domains that then resolve into narrow stripes as they propagate anteriorly through the SAZ (**Fig. 5.6A**). With each new cycle, *Tc-dpn* expression precedes that of *Tc-h*, and *Tc-dpn* stripes remain slightly anterior to those of *Tc-h* at all times (**Fig. 5.6A,D**). Due to their broader size, *Tc-dpn* and *Tc-h* stripes in the posterior SAZ appear to share a larger degree of overlap compared to those more anterior (**Fig. 5.6A**). Once *Tc-dpn* and *Tc-h* stripes become more refined towards the anterior SAZ, the extent to which they overlap each other becomes roughly 50%. Elevated stripes of *Tc-emc* expression in the SAZ were found to always be co-expressed with *Tc-h* stripes (**Fig. 5.6B**). The posterior *Tc-emc* stripe often appears narrower than the initial broad *Tc-h* domain, however, resolved stripes of *Tc-h* and *Tc-emc* towards the anterior SAZ are almost perfectly overlapping (**Fig. 5.6B,D**). HCR FISH with *Tc-emc* and *Tc-dpn* confirmed a similar relationship to *Tc-h* and *Tc-dpn* stripes, consistent with the co-expression of *Tc-emc* and *Tc-h* (data not shown).

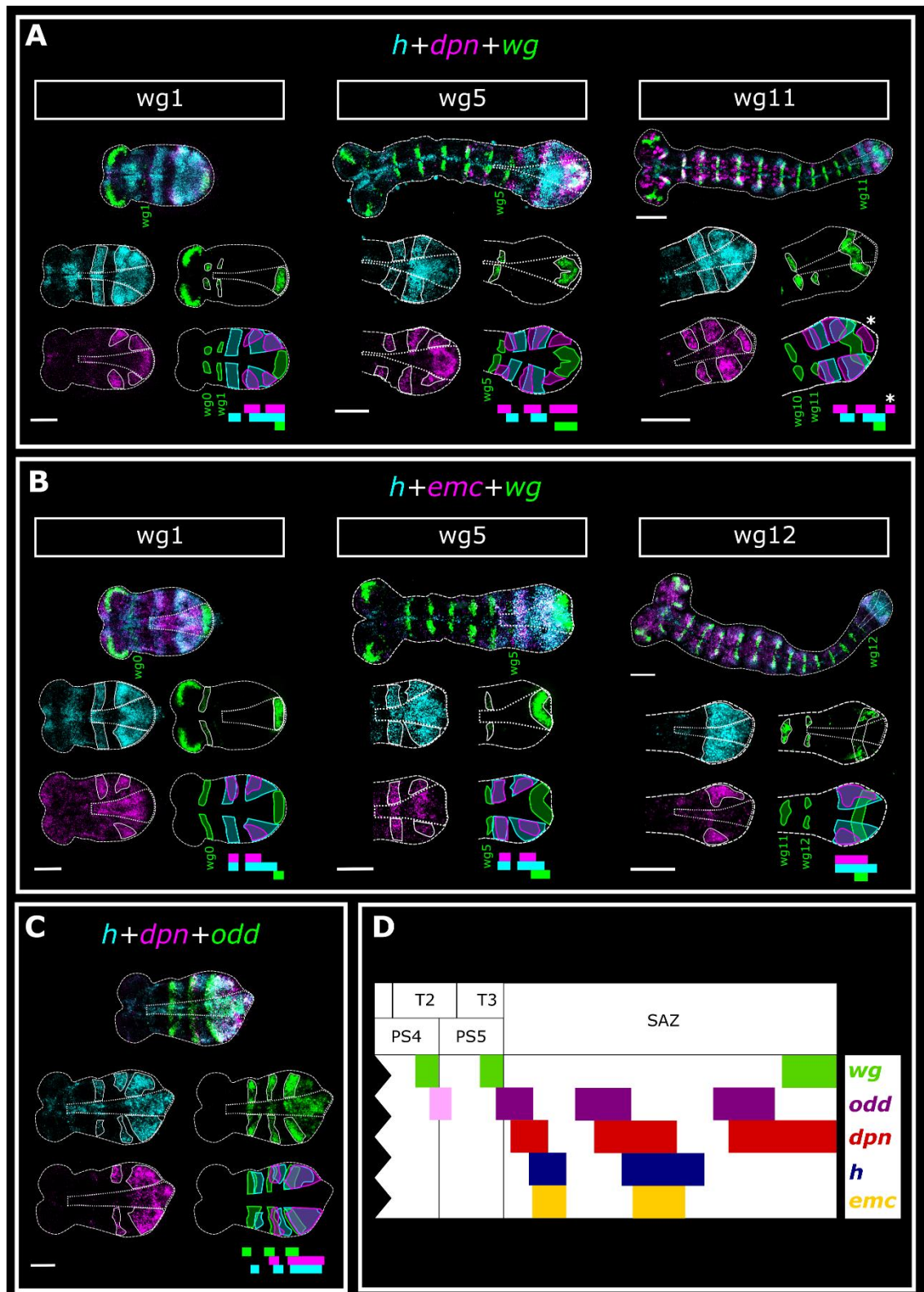


Figure 5.6. Co-expression analysis of striped Hes genes and *Tc-odd* during GBE. (A) HCR FISH of *Tc-h* (cyan), *Tc-dpn* (magenta) and *Tc-wg* (green) in germband stages wg1, wg5 and wg11. (B) HCR FISH of *Tc-h* (cyan), *Tc-emc* (magenta) and *Tc-wg* (green) in germband stages wg1, wg5 and wg12. (C) HCR FISH of *Tc-h* (cyan), *Tc-dpn* (magenta) and *Tc-odd* (green) in early germband. For each embryo in panels (A-C), whole embryo composite images are shown alongside individual channels for each gene. In later stage embryos, individual channel images are cropped to provide a closer look at posterior expression. The outline of each embryo, based on DAPI nuclei stains, is indicated by the dashed white line, while the mesoderm is highlighted by a dotted white line. Boundaries are drawn around the ectoderm expression domains for each gene and overlaid in a separate image. Simple schematics are given for each embryo to further indicate relative patterns of expression in the SAZ. Asterisks in (A) mark posterior *Tc-dpn* domain previously identified from colorimetry ISH. Posterior *Tc-wg* stripes are labelled in (A) and (B). Scale bars: 100 μ m. (D) Hypothetical schematic of relative gene expression in the *Tribolium* SAZ. The relative positions are deduced from combinations of embryos in (A-C). Example stage given is wg5, however, this pattern is representative of the SAZ from any germband stage up to wg12/13.

5.3. DISCUSSION

In this chapter I have confirmed previously published expression patterns for the *Tribolium Hes* genes, *Tc-h*, *Tc-dpn*, *Tc-E(spl)1* and *Tc-E(spl)3* (Aranda et.al., 2008; Kux et.al., 2013), as well as describing for the first time, expression patterns of *Tc-emc*, *Tc-Sidpn1*, *Tc-Sidpn2*, and *Tc-Hey* during GBE. Dynamic stripes of *Tc-emc* expression in the *Tribolium* SAZ provide us with a new candidate gene for arthropod segment patterning. By performing multiplex HCR FISH experiments, I have shown that striped *Hes* genes, *Tc-h*, *Tc-dpn* and *Tc-emc*, are expressed dynamically, yet maintain consistent relative positioning to one another for almost all of segment addition. Furthermore, these stripes are coordinated with *Tc-odd* stripes in the SAZ, suggesting that the oscillations of these genes are tightly linked. In addition to segmental stripes, *Tribolium Hes* genes are found to be expressed in several different tissues at varying stages of development, pointing to conserved roles in neurogenesis, as well as potential functions relating to mesoderm patterning, germ band elongation, leg and gut development.

5.3.1. Blastoderm expression points to a separate mechanism for mandibular segment patterning in *Tribolium*

Tc-h is first expressed in a transverse stripe across the middle of the blastoderm. This stripe corresponds to the mandibular *Tc-h* domain that is observed throughout the early germband stages of *Tribolium*. Unlike subsequent dynamic stripes, the mandibular domain does not originate from the posterior, an observation that is consistent with previously reported blastoderm expression for *Tc-h* (Sommer and Tautz, 1993; Eckert et.al., 2004). This, along with its sustained expression, suggests that the *Tc-h* mandibular stripe may be differentially regulated to posterior stripes. Indeed, lacZ reporter assays using different genomic regions upstream of *Tc-h* has shown that this first mandibular *Tc-h* stripe is activated by separate enhancer elements to those of posterior stripes (Eckert et.al., 2004). The unique behaviour of this *Tc-h* stripe likely relates to an important role for *Tc-h* in the development of the mandibular segment,

something that is evident from phenotypes of previously published *Tc-h* knockdowns (Aranda et.al., 2008).

Unlike *Tc-h*, initial stripes of *Tc-dpn* and *Tc-emc* both appear in the posterior end of the blastoderm and do not appear to be involved in the patterning of the mandibular segment. If *Tc-h*, *Tc-dpn* and/or *Tc-emc* function in any capacity as oscillators in the *Tribolium* segmentation clock, this role is therefore limited to segments posterior to the mandibular segment. The termination of dynamic SAZ stripes for these genes after stages wg13/wg14 also suggests that they are unlikely to be involved in the primary segmentation of the posterior-most trunk segments. Consistent with this, there is evidence from expression of the segment-polarity gene, *Tc-H15*, that these posterior segments are activated by a separate mechanism to other segments (Janssen, 2014). The exact nature of these different patterning mechanisms is yet to be investigated.

5.3.2. Expression dynamics in the SAZ indicate that *Tc-h*, *Tc-dpn* and *Tc-emc* oscillations are linked to the *Tc-eve/Tc-odd/Tc-run* circuit

For the dynamically expressed *Hes* genes, *Tc-h* and *Tc-dpn*, and the bHLH gene, *Tc-emc*, stripes emanating from the posterior SAZ are at first broad, with larger degrees of overlap when compared to the resolved stripes towards the anterior SAZ. This likely reflects a process of refinement by cross-regulation with each other and/or other cyclically expressed genes. Indeed, it is this cross-regulation between *Tc-eve*, *Tc-odd* and *Tc-run* that has so far been attributed to periodic oscillations and the establishment of resolved pair-rule stripes during primary segment patterning in *Tribolium* (Choe et.al., 2006). This process largely occurs across a spatiotemporal framework as dynamic stripes move anteriorly, however, it can be considered analogous to the temporal refinement of pair-rule stripes in the *Drosophila* “segmentation cascade”, which are also patterned by cross-regulation (Schroeder et.al., 2011). Furthermore, this process in *Drosophila* is now considered to involve (subtle) dynamic anterior shifts of pair-rule gene expression also, reconciling this mechanism between sequential and simultaneous modes of segmentation (Clark, 2017).

In this study, I have identified the consistent arrangement of resolved stripes of *Tc-h*, *Tc-dpn* and *Tc-emc* expression against those of *Tc-odd* in the *Tribolium* SAZ. Using previously published expression patterns for *Tc-eve*, *Tc-run* and *Tc-odd* (Choe et.al., 2006 (supporting information)), I have inferred the relative positions of *Tc-dpn*, *Tc-h*, *Tc-emc* and *Tc-odd* stripes with those of *Tc-run* and *Tc-eve* stripes in the mid-SAZ, once stripes are more refined (**Fig. 5.7A**). Comparison of these expression patterns to those of orthologues in *Drosophila* blastoderms at mid-cellularisation (Clark, 2017), reveal a generally conserved arrangement between species at comparative stages of patterning (**Fig. 5.7A-B**). In both cases, the anterior edges of *h* stripes slightly overlap the posterior edges of *odd* stripes. In *Drosophila*, under the dynamic model proposed by Clark (2017), this relative pattern can be explained by the repression of *odd* expression by *Hairy* (Schroeder et.al., 2011). In this model, as *h* stripes shift anteriorly, *odd* expression is turned off from the posterior overlapping region. The slight overlap observed in mRNA expression likely reflects a time lag between translation of *h* mRNA and repression of *odd* by the *h* protein product. Given that the expression of *Tc-odd* is known to be dynamic in the *Tribolium* SAZ (Sarrazin et.al., 2012), the same may also be true for the relationship between *Tc-h* and *Tc-odd* in *Tribolium*. This would be consistent with the previously discussed sequential segmentation network by Clark (2017) (**Fig. 5.7C**). However, in this model, *h* is also proposed to repress *run* expression, which is inconsistent with the inferred relative expression patterns for these two genes in *Tribolium* (**Fig. 5.7A**). From this, it would appear that *Tc-run* stripes have a larger degree of overlap and that *Tc-run* expression turns on in cells that are already expressing *Tc-h*. The phase of *Tc-dpn* stripes against *Tc-run* stripes would instead make *Tc-dpn* a more plausible candidate for this role (**Fig. 5.7D-E**).

These hypotheses are currently based on a mixture of inferred relative expression patterns, computational models, and limited functional data. In addition, published knockdown experiments for *Tc-h* and *Tc-dpn* have so far shown no signs of important roles in relation to *Tribolium* segmentation (Aranda et.al., 2008). More multiplex FISH experiments are required to confirm the relative expression patterns of *Tc-h/Tc-dpn* stripes with those of *Tc-run* and *Tc-eve* in the SAZ, while a comprehensive analysis of *Tc-h/Tc-dpn* function would be necessary to rule out any potential redundancy between

these genes or between these genes and those of the *Tc-eve/Tc-odd/Tc-run* circuit. As Tc-Emc is incapable of binding DNA, it is unlikely to directly regulate other segmentation genes. However, it has the potential to affect protein dynamics and function of bHLH transcription factors such as Tc-h and/or Tc-dpn, and is therefore also still a key candidate in the *Tribolium* clock network.

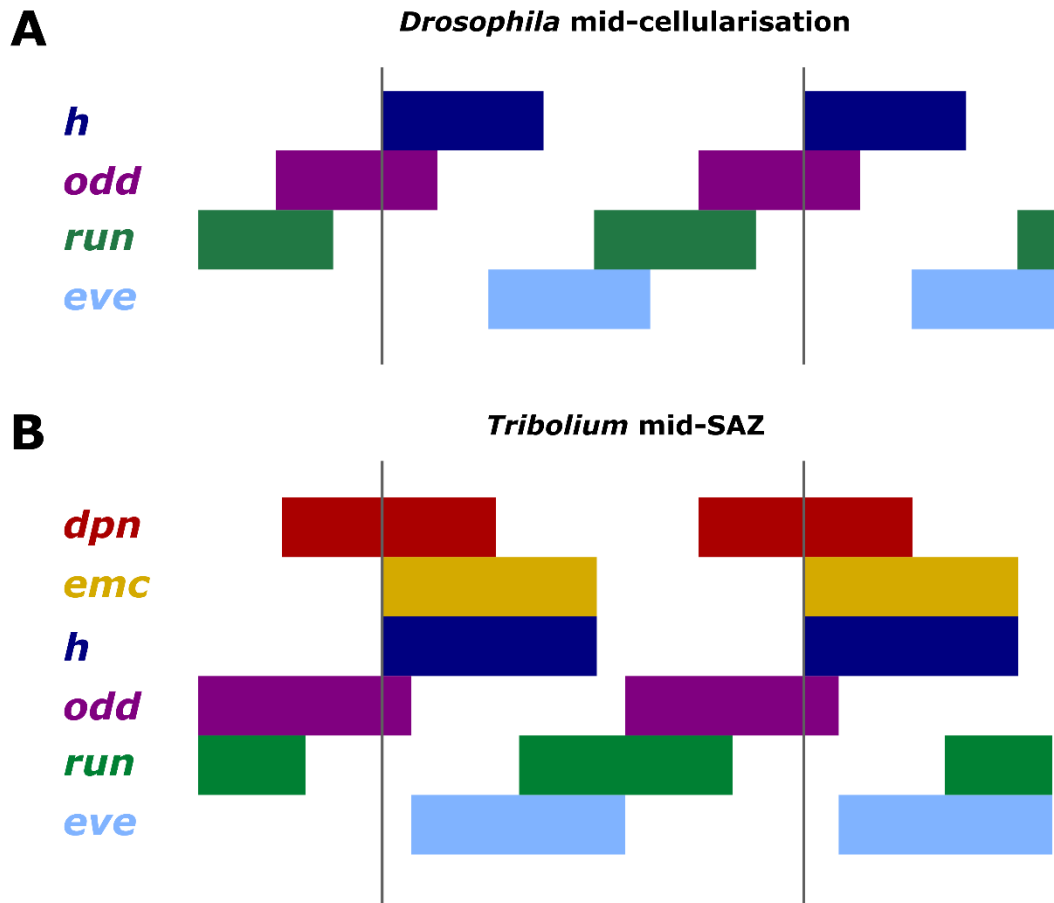


Figure 5.7. Expression dynamics of cyclical genes in the *Tribolium* SAZ compared with orthologues in *Drosophila*. (A) Graphical summary of pair-rule stripes in *Drosophila* mid-cellularisation (from Clark, 2017). (B) A simplified graphical summary of the relative expression patterns of *Tc-h/Tc-dpn/Tc-emc* and primary pair-rule genes in the mid-SAZ of *Tribolium*. This represents a point where stripes of expression are resolved, but where the regulatory logic should still reflect the primary segmentation network. Arrangement of *Tc-dpn*, *Tc-emc*, *Tc-h* and *Tc-odd* stripes are from this study, while *Tc-run* and *Tc-eve* patterns are inferred from double ISH with *Tc-odd* by Choe et.al. (2006, supplementary information). Vertical lines represent the start and end of a repeated pattern, aligned with the anterior boundary of *Dm-h* / *Tc-h* in each case. Anterior is to the left.

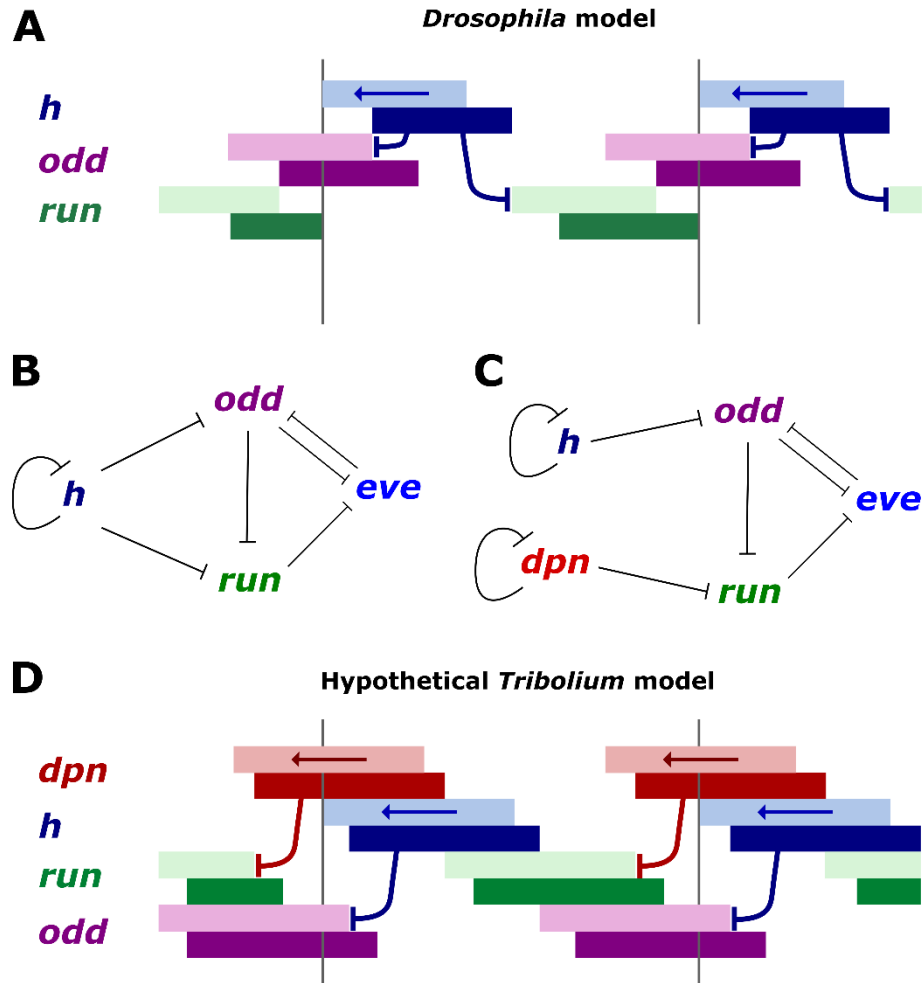


Figure 5.8. Hypothetical regulatory logic of *Tribolium* *Hes* genes in the SAZ. (A) Schematic of *h*, *odd* and *run* expression in *Drosophila* during mid-cellularisation. Relative positioning of stripes are, in part, explained by a dynamic model with negative regulation of *odd* and *run* expression by Hairy Clark (2017). Lighter and darker coloured stripes reflect transcript and protein domains, respectively. Offset transcript and protein patterns are due to translation lags, and are observed between *eve* and Eve, supporting the subtle dynamic nature of *Drosophila* pair-rule genes (Clark, 2017). (B) Part of a hypothetical model by Clark (2017) that replicates sequential segmentation *in silico*. (C) Modified network from (B) that fits the expression dynamics observed in *Tribolium*. (D) Graphical summary of mid-SAZ expression patterns of *Tc-dpn*, *Tc-h*, *Tc-run* and *Tc-odd* in *Tribolium* as inferred in Fig. 5.7. With posteriorly offset proteins domains, as would be expected in a dynamic system such as this, relative positioning of *Tc-dpn*, *Tc-h*, *Tc-run* and *Tc-odd* stripes are shown to fit the hypothetical regulatory relationships proposed in (C). Note that, this is based on inferred relative expression patterns only, and further experiments would be required to confirm these. Furthermore, the translation lags assigned here are arbitrary, and comparisons between transcript and protein expressions would be necessary to deduce these. Anterior is to the left.

5.3.3. Other *Hes/Hey* expression patterns are consistent with roles in neurogenesis and posterior signalling

Tribolium *Hes/Hey* expression in relation to neurogenesis

All *Tribolium* *Hes/Hey* genes, with the exception of *Tc-h* and possibly *Tc-cwo* (unverified ISH), are expressed within the neuroectoderm at varying stages of GBE. These findings are consistent with the expression of various *Hes/Hey* genes in the neuroectoderm of *Drosophila* (Southall et.al., 2009 (*dpn*); Monastirioti et.al., 2010 (*Hey*); Berkeley *Drosophila* Genome Project; Hammonds et.al., 2013), and the spider, *Parasteatoda tepidariorum* (Paese, 2018), suggesting that this gene family has a conserved role in neurogenesis across arthropods. As previously discussed, this likely reflects an ancient function of the *Hes/Hey* gene family in neurogenesis that was established at the dawn of metazoan evolution (Gazave et.al., 2014; Layden and Martindale, 2014).

Early nervous system development in insects involves the establishment of neural stem cells, known as neuroblasts, from proneural cell clusters in the neuroectoderm (Biffar, 2013). Neuroblasts then delaminate and divide asymmetrically, renewing themselves while producing differentiated neurons or intermediate progenitor cells. The expression of neuroblast markers in *Tribolium*, *Tc-asense* (*Tc-ase*), *Tc-achaete-scute-homolog* (*Tc-ash*), and *Tc-snail* indicate the timing of this process (Wheeler et.al., 2003; Kux et.al., 2013).

The expression patterns I obtained for *Tc-E(spl)1* and *Tc-E(spl)3* are consistent with previously published results for these genes in *Tribolium* (Kux et.al., 2013). Co-expression analysis with the early neuroblast marker, *Tc-ash*, as well as loss of function experiments, strongly suggests that *Tc-E(spl)* genes are involved in a conserved Notch-mediated lateral inhibition mechanism that controls neuroblast specification in *Tribolium* (Kux et.al., 2013). Neuroectodermal expression of *Tc-dpn* and *Tc-Sidpn1* appears at a similar time point to *Tc-E(spl)* genes, suggesting that these *Hes* genes may also play important roles in the specification of neuroblasts in *Tribolium*. However, without co-expression data against neuroblast markers, it is hard to infer the precise nature of *Tc-dpn* and *Tc-Sidpn1* expression and function. While there are no functional studies regarding *Tc-Sidpn1* or any orthologues in other insects, *Drosophila* *Sidpn*

expression does not appear in the neuroectoderm until after neuroblast specification (Berkeley *Drosophila* Genome Project; Hammonds et.al., 2013). The difference in timing suggests that there may be some divergence in the precise function of insect *Sidpn* genes. In contrast, *Drosophila dpn* expression follows a similar pattern in the neuroectoderm to its *Tribolium* counterpart. In *Drosophila*, *dpn* is co-expressed with neuroblast genes *asense* and *snail* in the *Drosophila* neuroectoderm (Southall and Brand, 2009), and analysis of Dpn binding sites suggests that it regulates genes involved in neuroblast renewal and differentiation (Southall and Brand, 2009), however, functional data is still lacking.

Interestingly, *Tc-Sidpn2* is expressed in the neuroectoderm before any other *Hes/Hey* gene and appears to precede neuroblast specification. This early expression points toward a possible role in nerve cord patterning, and/or as an upstream promoter of neural precursors alongside genes such as *ventral nerve cord defective (Tc-vnd)* and *intermediate nerve cord defective (Tc-ind)*, which are similarly expressed at this stage in *Tribolium* (Wheeler et.al., 2005; Biffar and Stollewerk, 2015). Phylogenetic analysis failed to identify any clear orthologues of *Tc-Sidpn2* in other arthropods (**Chapter 3**), and there are therefore currently no comparable studies. In contrast to *Tc-Sidpn2*, *Tc-Hey* and *Tc-emc* are both expressed in the neuroectoderm at later stages than the other *Hes/Hey* genes, indicating that these genes do not contribute to neuroblast specification. This later neuroectodermal expression is consistent with that observed in *Drosophila* orthologues, *Hey* and *emc*, where they are responsible for regulating later cycles of neuronal differentiation (Monastirioti et.al., 2010; Cubas et.al., 1994).

Expression patterns of several *Hes/Hey* gene orthologues display a large degree of conservation between *Tribolium* and *Drosophila* in the context of neurogenesis, and functional data from *Drosophila* allows us to make reasonable predictions about their roles in *Tribolium*. However, co-expression analyses with neuroblast markers and functional studies are still necessary to test these.

Posterior *Hes/Hey* expression: Roles in hindgut specification and/or germband elongation?

In this study, I observed posterior expression domains for several *Hes/Hey* genes at various points during development. *Tc-dpn*, *Tc-Sidpn1*, and *Tc-E(spl)1* were all expressed at the posterior end of the *Tribolium* germband, behind the posterior *Tc-wg* domain, from around stage wg6 through to the end of GBE. The exact nature of these domains are varied, with *Tc-dpn* appearing in a small cluster of cells, while *Tc-Sidpn1* and *Tc-E(spl)1* have slightly broader domains. *Tc-h* expression also becomes fixed in this region, overlapping the posterior *Tc-wg* domain once striped expression has ceased, while a posterior *Tc-emc* domain is only visible towards the end of GBE.

Posterior expression of *Tc-dpn*, *Tc-Sidpn1*, and *Tc-E(spl)1* appears at a time when the posterior *Tc-wg* domain changes shape and moves anteriorly. The timing of this coincides with several important developmental processes such as the onset of hindgut patterning, and the transition from thoracic to abdominal segment patterning. It is at this point that several key genes involved in specification of the hindgut, such as *Tc-drm* and *Tc-forkhead* (*Tc-fkh*), first become expressed in this region (Schröder et.al., 2000; Thümecke and Schröder, 2021). Meanwhile, a dramatic increase in the rate of segment addition, alongside a shift in the relative contribution of extensive cell rearrangements to GBE, occurs immediately after the establishment of the wg6 stripe (Nakamoto et.al., 2015). Previous studies in which posterior Wnt signalling has been perturbed, through knockdowns of Wnt ligands or targets of the canonical Wnt pathway, consistently produce *Tribolium* embryos in which head and thoracic segments have developed, but that lack abdominal segments (Bolognesi et.al., 2008; Bolognesi et.al., 2009), suggesting there is a shift in the regulatory control of GBE and/or segmentation during the transition from thoracic to abdominal segment patterning. The appearance of posterior *Tc-dpn*, *Tc-Sidpn1*, and *Tc-E(spl)1* expression domains at this stage are therefore consistent either with roles in hindgut specification, or as elements of the posterior signalling centre that directs the “growth” and patterning of abdominal segments. However, functional studies of *Tc-dpn* and *Tc-E(spl)1* have not reported phenotypes relating to either of these processes (Aranda et.al., 2008; Kux et.al., 2013),

and further investigation is required to confirm which, if any, are true for the posterior expression of *Tc-Sidpn1*.

The posterior domains of *Tc-h* and *Tc-emc* only appear later in development once segment patterning and GBE are nearly complete, suggesting that they are unlikely to be involved in the regulation of these processes. Analysis of post-GBE *Tribolium* embryos shows that *Tc-h*, *Tc-emc* and *Tc-dpn* are expressed in patterns consistent with roles in hindgut and malpighian tubule morphogenesis (**Appendix C; Fig. C1**). The same stages were not captured for other posteriorly expressed *Hes* genes in this study.

6. INVESTIGATING FUNCTION OF CANDIDATE *HES* GENES AND *TC-EMC*

6.1. INTRODUCTION

I have shown in chapter 5 that the dynamic stripes of expression for the *Hes* genes, *Tc-h* and *Tc-dpn*, and the bHLH gene, *Tc-emc*, in the *Tribolium* SAZ are co-ordinated with those of the primary pair-rule genes, *Tc-odd*, *Tc-eve* and *Tc-run*, indicating the potential for a degree of cross-regulation between them. The exact relationships among these genes, and their functional significance in the oscillatory network governing primary segment patterning, is still unclear. However, this finding, along with the pair-rule like phenotype and resolution of *Tc-h* and *Tc-run* stripes observed in *Tc-odd* knockout embryos (Sharma pers comms), make *Tribolium Hes* genes and *Tc-emc* primary candidates for partially rescuing primary segment patterning in the absence of *Tc-odd*. The revised expression patterns described for *Tribolium odd-skipped*-family genes in chapter 4 makes it very unlikely that they act redundantly with *Tc-odd* in the segmentation clock network. For this reason, this chapter will solely focus on investigating the functions of *Hes* family candidate genes and *Tc-emc* for possible roles and redundancy in the *Tribolium* segmentation network.

6.1.1. An ancestral role for arthropod *Hes* genes in Notch-mediated segmentation

In sequentially segmented arthropods where patterning appears to be mediated by Notch signalling, there is multiple evidence that *Hes* oscillations in the SAZ are tied to the Notch signalling pathway. This has been demonstrated in the spider, *Cupiennius*, the crustacean, *Daphnia*, and the cockroach, *Periplaneta*, by experimentally inhibiting Notch signalling (Stollewerk et.al., 2003; Pueyo et.al., 2008; Eriksson et.al., 2013), while expression patterns in the SAZ of the centipede, *Strigamia*, support a similar story (Chipman and Akam, 2008; Green and Akam, 2013). In addition, knockdowns targeting a *Hes* gene, *Pa-h*, in *Periplaneta*, replicate the phenotypes produced by Notch

inhibition, indicating linked functional roles in the segmentation clock of this species (Pueyo et.al., 2008). The prevalence of Notch-mediated segmentation and associated *Hes* oscillations across basally branching arthropod taxa (**Fig. 1.6**), suggests that this mechanism is ancestral. In each case, genes from the Notch signalling pathway, *Notch* and *Delta*, are expressed in dynamic stripes in the SAZ, consistent with their relationship with *Hes* oscillations.

In contrast, members of the *Tribolium* Notch signalling pathway, *Tc-Notch*, *Tc-Delta*, *Tc-Serrate* and *Tc-Fringe*, do not exhibit signs of dynamic SAZ expression during segment addition (Aranda et.al., 2008), suggesting that Notch-mediated segmentation has been lost in this species. This is also true for *Notch* and/or *Delta* expression in several other sequentially segmented insects, such as *Gryllus* and *Oncopeltus* (Mito et.al., 2011; Kainz et.al., 2011; Auman et.al., 2017). In *Oncopeltus*, *Of-h* expression is only observed in the anterior SAZ alongside *Of-Delta*, indicating that, while still linked to Notch signalling, any ancestral role for this gene in primary segment patterning has been lost (Auman et.al., 2017). Furthermore, *Of-dpn* does not exhibit any striped expression, and the knockdown of *Of-h* fails to produce any segment patterning defects (Reding et.al., 2019). While evidence in *Oncopeltus* points towards a complete loss of *Hes* oscillations and Notch-mediated synchrony in the segmentation clock of this species, an exhaustive screen of *Hes/Hey* family genes or other members of the Notch signalling pathway would be required to rule out redundancy and developmental system drift. In *Tribolium*, despite the apparent lack of cyclically expressed genes from the Notch signalling pathway, the *Hes* genes, *Tc-h* and *Tc-dpn* are still both expressed in dynamic stripes in the SAZ, suggesting that they may be acting independently of the Notch pathway in this instance. In addition, *Tc-emc*, whose orthologue is commonly associated with Notch signalling in *Drosophila* (Campuzano, 2001; Bhattacharya and Baker, 2009), has also now been found to exhibit oscillatory expression in the *Tribolium* SAZ (**Chapter 5**).

6.1.2. Existing functional data for *Hes* genes and Notch signalling in *Tribolium*

While expression patterns for members of the *Tribolium* Notch signalling pathway are inconsistent with a role in synchronising cellular oscillations across the SAZ, knockdown

of the ubiquitously expressed receptor, *Tc-Notch*, generates abdominal segmentation defects that suggest Notch signalling may still be required for segmentation in some capacity (Kux et.al., 2013). Meanwhile, published knockdowns for the SAZ expressed *Hes* genes, *Tc-h* and *T-dpn*, have found no evidence to suggest that either gene is required for segmentation (Aranda et.al., 2008). However, for reasons already discussed in this thesis, there is a need to revisit this and to test for functional redundancy with other segmentation genes. As a repressor of bHLH function, the co-expressed *Tc-emc* gene is also of interest for any potential regulatory relationship it may have with these genes. *Tc-emc* has previously been included in a parental RNAi screen of bHLH genes where it was found to be essential for oogenesis, resulting in females that were unable to lay eggs after injection (Bitra and Pali, 2010). As such, any possible roles in embryonic development have not yet been tested.

Expression patterns for the *Hes* gene, *Tc-Sidpn1*, are inconsistent with a role in segment patterning, however, its function during development has not yet been described in *Tribolium*, or for orthologues in any other arthropods. Furthermore, it is possible that the posterior expression domain observed in chapter 5 could be linked to a signalling centre upstream of segment patterning. Like *Tc-emc*, *Tc-Sidpn1* also appears to play an essential role in reproduction (Bitra and Pali, 2010), meaning that embryonic functions have yet to be investigated.

6.1.3. Aims of the chapter

In this chapter, I aim to use combinations of parental and embryonic RNAi to corroborate published functional data for candidate *Hes* genes, while screening for developmental roles of *Tc-emc* and *Tc-Sidpn1*. Using various combinations of double RNAi knockdowns in wildtype and *Tc-odd*-null mutants, I aim to test for functional redundancy among *Tc-h*, *Tc-dpn* and *Tc-emc* in the *Tribolium* clock network, and to investigate redundancy of these genes with *Tc-odd* in the SB strain.

6.2. RESULTS

6.2.1. Use of non-overlapping dsRNA fragments for RNAi knockdowns

Canonical RNA interference (RNAi) relies on the introduction of sequence-specific double-stranded RNA (dsRNA) into the cellular environment during development (reviewed by Svoboda, 2020). For all RNAi experiments performed here, where dsRNA was synthesised or used for the first time, two non-overlapping fragments of dsRNA were used to test reproducibility and to control for possible off-target effects. One fragment would be designed from the 5' region of the gene's mRNA sequence, while the other would be based on the 3' end. For simplicity in this chapter, I will use 5' or 3' to distinguish between groups that have been injected with the corresponding non-overlapping dsRNA fragments for each gene. In addition, a negative control group injected with *GFP* dsRNA was included alongside each experiment.

6.2.2. Single knockdowns of *Tc-h* and *Tc-dpn* corroborate previously published data

Tc-hairy 3' dsRNA, but not 5', generates head defects consistent with published phenotypes

Female beetles injected with *Tc-h* 3' dsRNA produced a high percentage of offspring with abnormalities, compared to those from the *GFP* dsRNA control injections (**Fig.6.1A**). Of the 317 cuticles analysed from the 3' group, 71.6% exhibited some degree of head abnormality. In contrast, only 0.5% of the 435 cuticles analysed from *Tc-h* 5' injected females displayed any such defects (**Fig.6.1B**), suggesting that either the 5' dsRNA was unsuccessful in generating an efficient knockdown, or that the 3' dsRNA is generating off-target effects. A BLAST search of the 3' dsRNA sequence against the *Tribolium* genome assembly found no other sequence with significant similarity, suggesting that the former is most likely to be true. Furthermore, results from the 3' group closely resemble those previously reported in RNAi experiments by Aranda et.al.

(2008). Due to its confirmed effectiveness in knocking down *Tc-h*, the 3' dsRNA is used for all subsequent experiments in which *Tc-h* is targeted.

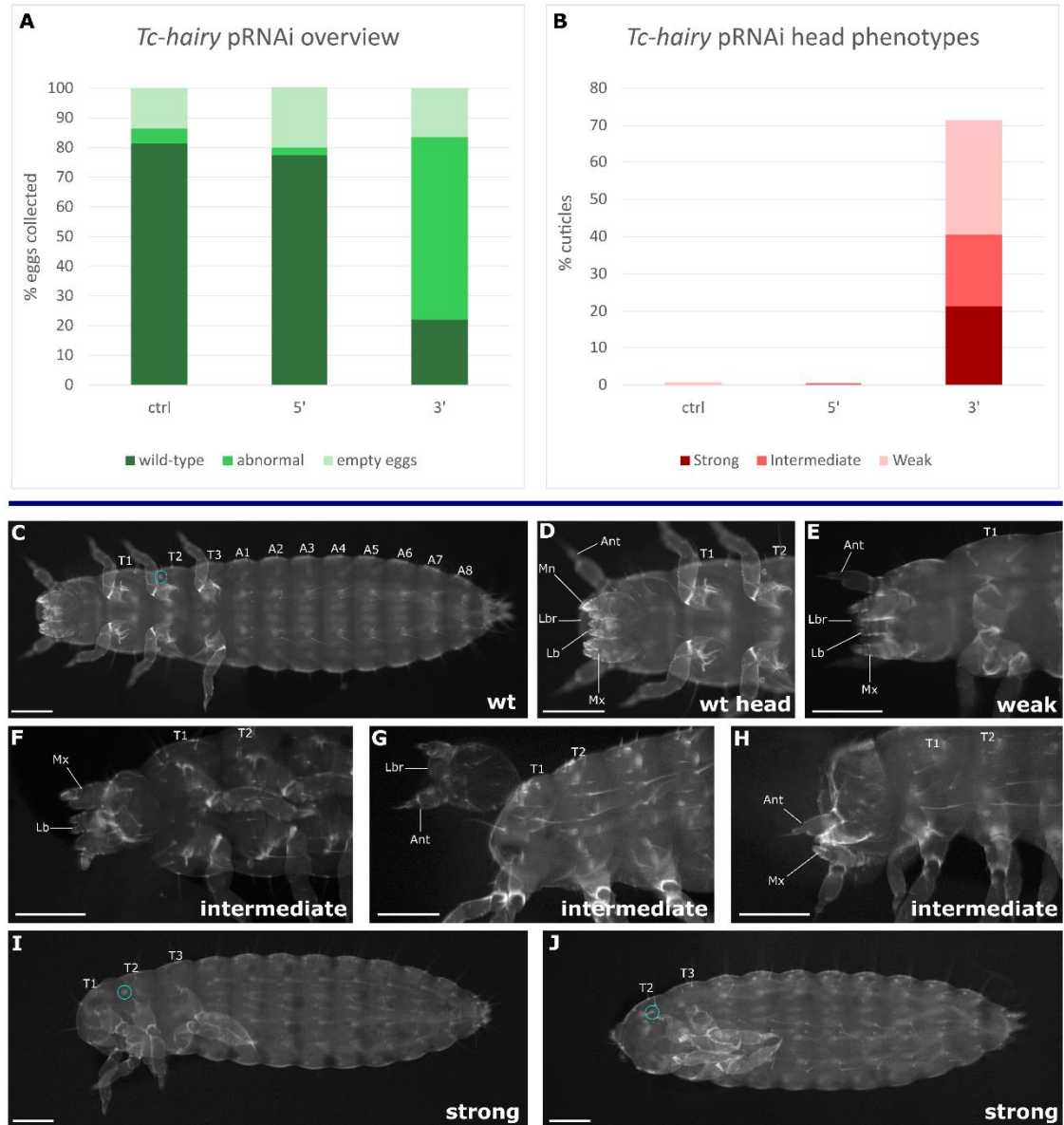


Figure 6.1 parental RNAi with *Tc-hairy* 3' dsRNA confirms published head phenotype.

(A) Summary results for *Tc-h* pRNAi compared with *GFP* dsRNA injected control. (B) Frequencies of weak, intermediate, and strong head phenotypes as described in main text. (C,D) Wildtype cuticles. (E-J) cuticles representing phenotypic series from *Tc-h* 3' injections. Cyan circles highlight T2 spiracles. Injections were performed using 2 $\mu\text{g}/\mu\text{L}$ dsRNA. Heads labelled to highlight missing appendages: Lbr = Labrum; Ant = Antenna; Mn = mandible; Mx = Maxilla; Lb = Labia. Scale bar: 100 μm .

The abnormalities observed in the *Tc-h* 3' group follow a phenotypic series consistent with those described by Aranda et.al. (2008). Weak *Tc-h* knockdown results in the loss of the mandibular segment (**Fig.1E**). Many of the larvae that display this phenotype had hatched successfully. In intermediate cuticles, additional head appendages were lost or deformed. However, the loss of these did not appear to follow any obvious order (**Appendix D; Table D.1**). The entire anterior head could be lost (**Fig.1F**), or just the gnathal segments (**Fig.1G**), and in many cases, parts of both (**Fig.1H**). Despite this variation, the loss of mandibles was always observed. This is consistent with the mandibular stripe of *Tc-h* expression that persists throughout the early stages of GBE (**Fig. 5.1 green arrowheads**) In embryos that experienced the most severe effect of *Tc-h* knockdown, the entire head failed to develop, and in some cases, anterior thoracic segments were also missing (**Fig.1 I-J**).

Tc-dpn RNAi produces no clear developmental abnormalities

Of the eggs collected from females injected with *Tc-dpn* dsRNA, no significant deviations were observed in the frequencies of wild-type cuticles, abnormal cuticles, and empty eggs, when compared with those from *GFP* injected controls (**Table 6.1**). Both 5' and 3' non-overlapping fragments produced similar results. Female survival was low in these experiments, however, a similar proportion of the *GFP* dsRNA injected females also died, suggesting that the starting population may have been unhealthy, readily succumbing to injection related injuries. From the eggs collected from surviving females, results are consistent with previously published *Tc-dpn* pRNAi results for *Tribolium*, which similarly reported a high hatch rate (Aranda et.al., 2008; Bitra & Palli, 2010).

Table 6.1. Results overview for *Tc-dpn* parental RNAi. Females were injected with 2 µg/µl dsRNA and eggs were collected on 4 separate occasions over a 2 week period. Experiments were carried out alongside a *GFP* dsRNA injected control group.

Treatment	females injected	females survived	eggs analysed	% wild-type	% abnormal	% empty eggs
<i>GFP</i> (ctrl)	37	7	277	72.6	6.5	20.9
<i>dpn</i> 5'	42	11	478	79.5	5.4	15.1
<i>dpn</i> 3'	42	10	275	73.5	5.5	21.1

6.2.3. *Tc-emc* RNAi produces a high penetrance of abnormalities relating to abdominal segmentation, dorsal closure, head, leg, and hindgut development.

I first attempted to assess *Tc-emc* function by performing parental RNAi (pRNAi). However, consistent with results from Bitra and Pali (2010), this had a severe impact on egg laying for injected females. Over a 2-week period, in which 4 egg collections were made, females injected with *Tc-emc* 5' dsRNA and *Tc-emc* 3' dsRNA, laid on average 2.2 and 0 eggs per female respectively, compared to 46.1 eggs per female in the *GFP* dsRNA injected control (**Appendix D; Table D.2**). An additional egg collection one month after injections found that egg laying had still not recovered, with 3' injected females failing to lay a single egg. Of the eggs that were laid by *Tc-emc* 5' injected females, the majority hatched, and no consistent developmental abnormalities were observed. In order to identify any roles for *Tc-emc* in *Tribolium* development, it was necessary to circumvent its effect on female reproduction through the use of embryonic RNAi (eRNAi).

Eggs injected with *Tc-emc* 5' and 3' dsRNA yielded a high penetrance of phenotypes related to the head, abdominal segments, legs, hindgut, and dorsal closure, when compared with *GFP* dsRNA injected controls (**Table 6.2**). In addition, *Tc-emc* 5' injections produced a notably higher rate of abnormalities in all contexts except the hindgut

Table 6.2. Overview of *Tc-emc* eRNAi. *Tc-emc* and *GFP* dsRNA were mixed and injected together as these experiments were performed alongside double knockdowns (see 6.2.5). This served as a control for keeping dsRNA concentrations the same between groups. 2 µg/µl *Tc-emc* dsRNA was mixed with an equal volume of 2 µg/µl *GFP* dsRNA to produce a final solution with 2 µg/µl total dsRNA but 1 µg/µl individual concentrations for both *Tc-emc* and *GFP* dsRNA. The *GFP* dsRNA treated control group were injected with 2 µg/µl *GFP* dsRNA. N = number of eggs injected. No. cuticles = total number of injected eggs that formed a cuticle (incl. hatched larvae). The percentage of cuticles exhibiting head, segmentation, leg, hindgut, or dorsal closure phenotypes is listed in the subsequent columns.

Treatment	N	No. cuticles	Abnormalities / % cuticles				
			head	segmentation	leg	hindgut	dorsal closure
<i>GFP</i> dsRNA	227	150	4.0	8.0	2.7	0.7	3.3
<i>Tc-emc</i> 5' + <i>GFP</i> dsRNA	229	147	81.6	83.0	72.8	72.8	54.4
<i>Tc-emc</i> 3' + <i>GFP</i> dsRNA	218	130	56.9	70.8	59.2	78.4	48.5

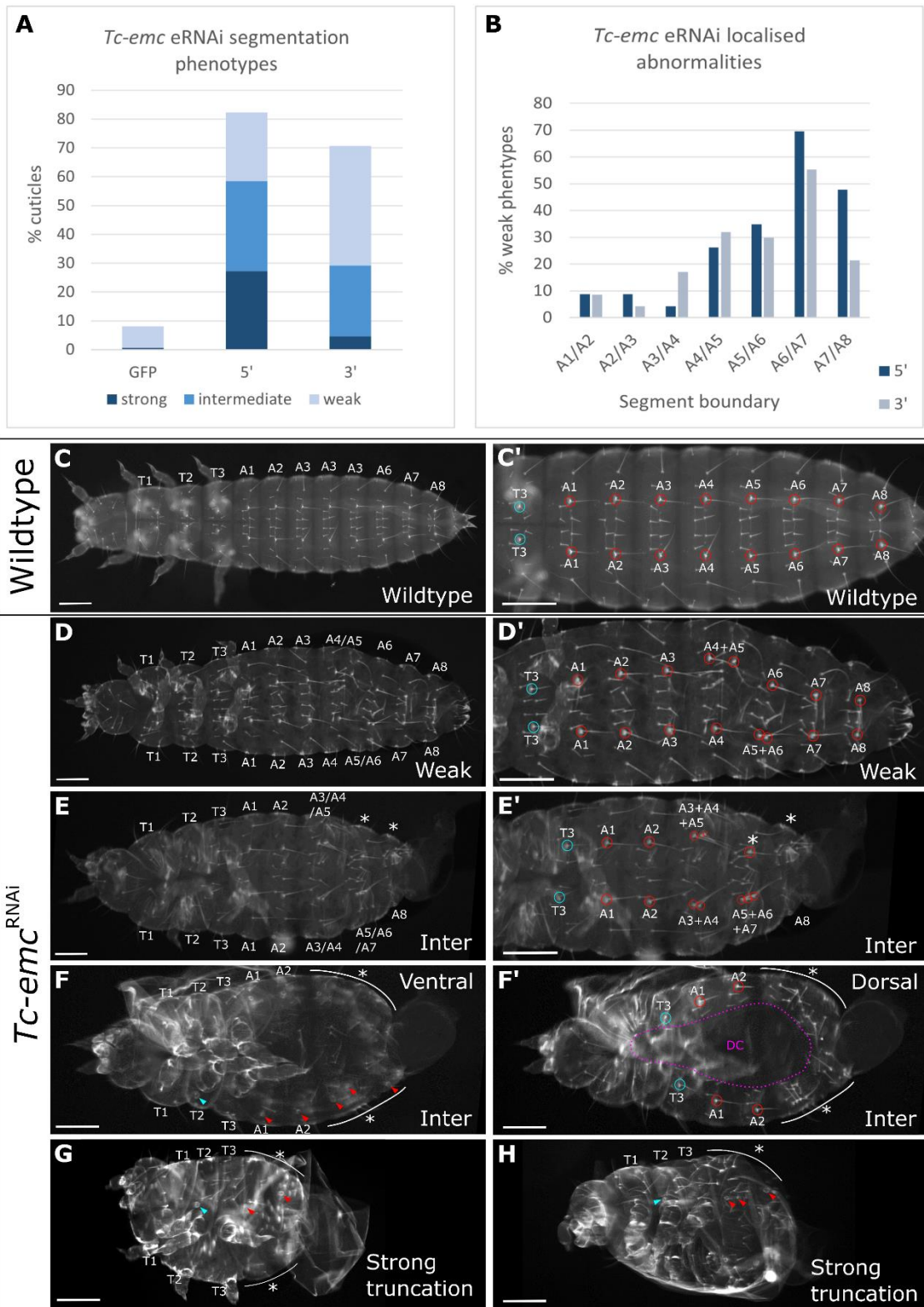
Tc-emc knockdown results in the fusion and loss of abdominal segments

A broad phenotypic series of abdominal segment abnormalities was obtained from *Tc-emc* eRNAi injections and these were classified into weak, intermediate, and strong phenotypes. Weak effects of *Tc-emc* knockdown on *Tribolium* segmentation include localised segment merges and/or irregular bristle patterns (**Fig. 6.2D, D'**). Of these weak phenotypes, the majority of localised abnormalities occurred between segments in the posterior half of the embryo (A4-A8) (**Fig. 6.2B; Appendix E; Fig. E.1**). Intermediate phenotypes were defined as having widespread fusions and/or missing between 2 and 4 segments (**Fig. 6.2E-F'**). In many cases, embryos also exhibited dorsal closure defects (**Fig. 6.2F'**), which made close analysis of abdominal bristle patterns difficult. However, of those that could be analysed, multiple abdominal segments appear to be fusing

together, producing an increasingly disorganised trunk (**Fig. 6.2E, E'**). As more segments fuse with each other from both sides of the embryo, this results in the loss of abdominal segments and shortening of the trunk. In cuticles with dorsal openings, spiracle counts were used to identify the number of missing and/or merged segments (**Fig. 6.2F-F'**). In the strongest phenotypic class, four or more abdominal segments were missing, resulting in severely truncated embryos (**Fig. 6.2G-H**). In most cases, the patterning of thoracic segments did not appear to be affected. This is despite the observation that *Tc-emc* is expressed in stripes in the SAZ throughout segment addition, including the period in which thoracic segments are being established (**Fig. 5.3**).

Penetrance of segment abnormalities was high in both the *Tc-emc* 5' and 3' injected groups, at 83% and 70.8% respectively (**Table 6.2**). However, more strong phenotypes were observed in the 5' group, with 27.2% of cuticles missing 4 or more abdominal segments, as opposed to just 4.6% of embryos injected with 3' dsRNA (**Fig. 6.2A**)

Figure 6.2 (overleaf). Impact of *Tc-emc* eRNAi on abdominal segmentation. (A) Phenotype penetrance from *Tc-emc* eRNAi compared to *GFP* dsRNA injected control. (B) Spread of localised A segment merges and/or abnormalities, expressed as a % of weak segmentation phenotypes (excluding indistinguishable cuticles; 5' n=23, 3' n=47). (C, C') wildtype cuticle. (D, D') Weak phenotype displaying A4/A5 and A5/A6 merges. (E-F') Intermediate (Inter) phenotypes with multiple segment merges and missing segments. (G, H) Strong segmentation phenotypes. Cuticles severely truncated. Cyan (thoracic) and red (abdominal) circles highlight homologous bristles, while thoracic and abdominal spiracles are indicated by cyan and red arrowheads respectively. DC in F' highlights dorsal closure defect. Asterisks mark indistinguishable segments. Scale bar: 100 μ m. See Appendix E for more examples of localised abnormalities.



Tc-emc knockdown results in consistent head defects, with the labrum and antennae exhibiting the highest degree of sensitivity

Tc-emc knockdown resulted in a range of head phenotypes. The mildest of these was a reduced or missing labrum (**Fig. 6.3C, C'**), which was recorded in 27.9% and 36.2% of cuticles in the 5' and 3' groups respectively. In cuticles exhibiting stronger head phenotypes, additional head appendages were deformed or missing (**Fig. 6.3D-E'**). In most cases, this included, but was not limited to, the antennae. In the most severely affected embryos, the head was greatly reduced or even missing entirely (**Fig. 6.3F, F'**). While there was a degree of variation across the phenotypes, one thing that remained consistent was a reduced or missing Lbr. The Lbr was affected in 99.2% and 94.6% of the head phenotypes observed in the *Tc-emc* 5' and 3' groups respectively.

Additional functional roles identified for *Tc-emc* in leg and hindgut development

Tc-emc RNAi induced a high penetrance of leg abnormalities (**Fig. 6.4**) consistent with elevated expression of *Tc-emc* in the limb buds post-GBE (**Fig. 5.3 blue arrowheads**). In mild cases, leg segments were visibly shortened and squashed (**Fig. 6.4D, D'**). Stronger phenotypes appear to have fused or missing leg segments (**Fig. 6.4E, E'**), and in the most severe cases, legs were reduced to indistinguishable nubs (**Fig. 6.4F, F'**).

Another consistent abnormality observed in *Tc-emc* injected embryos, is a deformed posterior, in which urogomphi are missing, and the hindgut is protruding (**Fig. 6.4H, H'**). In strong phenotypes, a large featureless piece of tissue extends from the posterior end of the embryo (**Fig. 6.4I, I'**). As this protrusion was sometimes observed in cuticles with all eight abdominal segments, it appears to be unrelated to the segmentation phenotype. Furthermore, it is always observed in the absence of any visible internal hindgut structures, suggesting it is most likely represents a deformed hindgut. A function of *Tc-emc* in hindgut development would be consistent with the strong posterior *Tc-emc* observed in chapter 5 (**Fig. 5.3 red arrowhead**) While the total penetrance of hindgut phenotypes is slightly higher in the 3' injected group, *Tc-emc* 5' injected embryos exhibit higher proportion of strong hindgut phenotypes (**Fig. 6.4B**), consistent with the higher rate of severity for other phenotypes in this group.

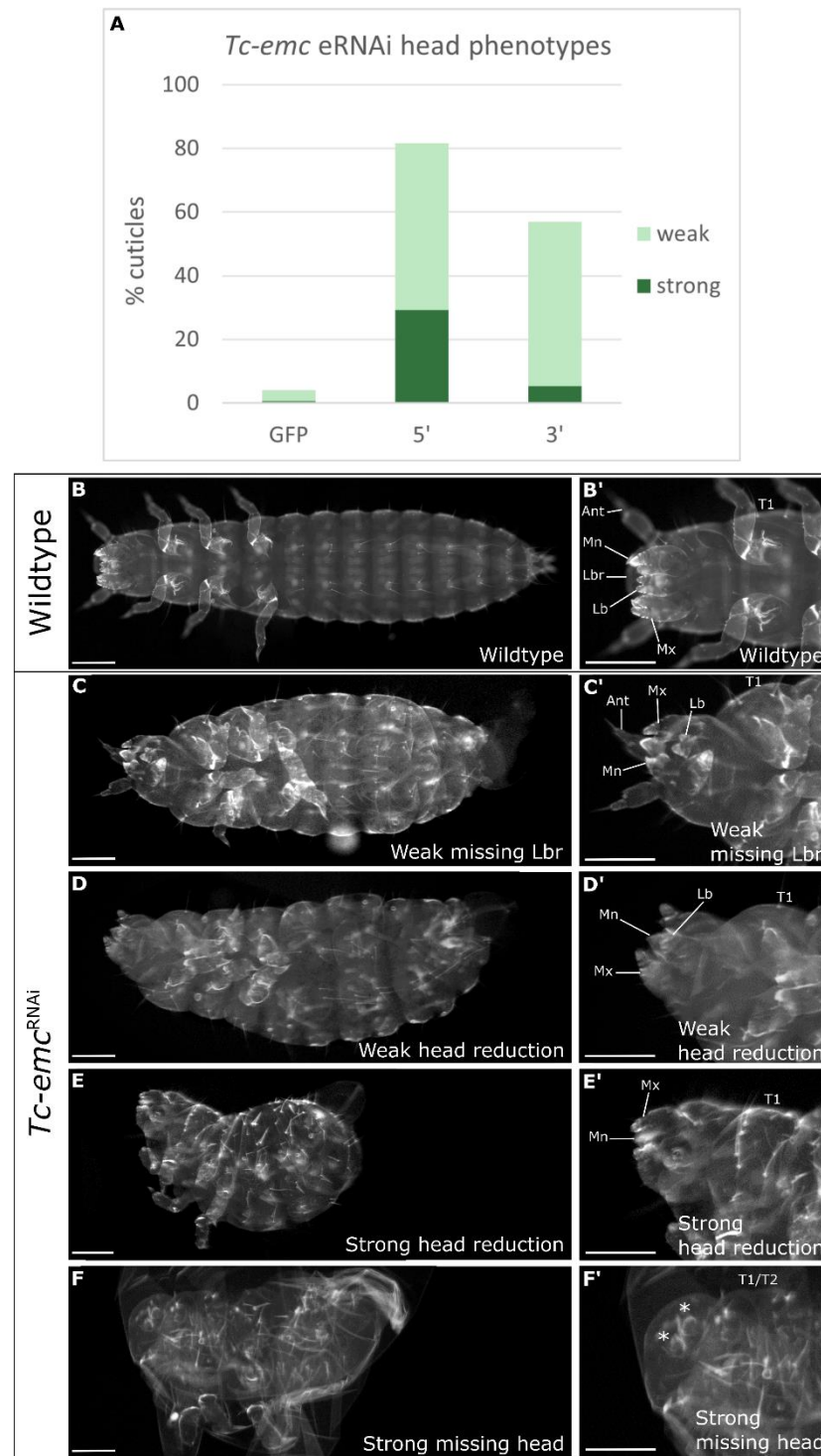


Figure 6.3. eRNAi reveals role of *Tc-emc* in *Tribolium* head development. (A) chart overview of head phenotypes from *Tc-emc* 5' and 3' injections, alongside *GFP* control. (B, B') Wildtype cuticle for reference. (C-F') *Tc-emc*^{RNAi} cuticles. (C-D') Examples of weak head phenotypes where 1 or 2 head appendages are missing/deformed. (E-F') Examples of strong head phenotypes where 3 or more head appendages are missing/deformed. Asterisks mark indistinguishable buds, which are either head appendages (e.g. Mx) or T1 legs. Head parts are labelled to highlight missing appendages (see **Fig. 6.1** for abbreviations). Scale bar: 100 μ m

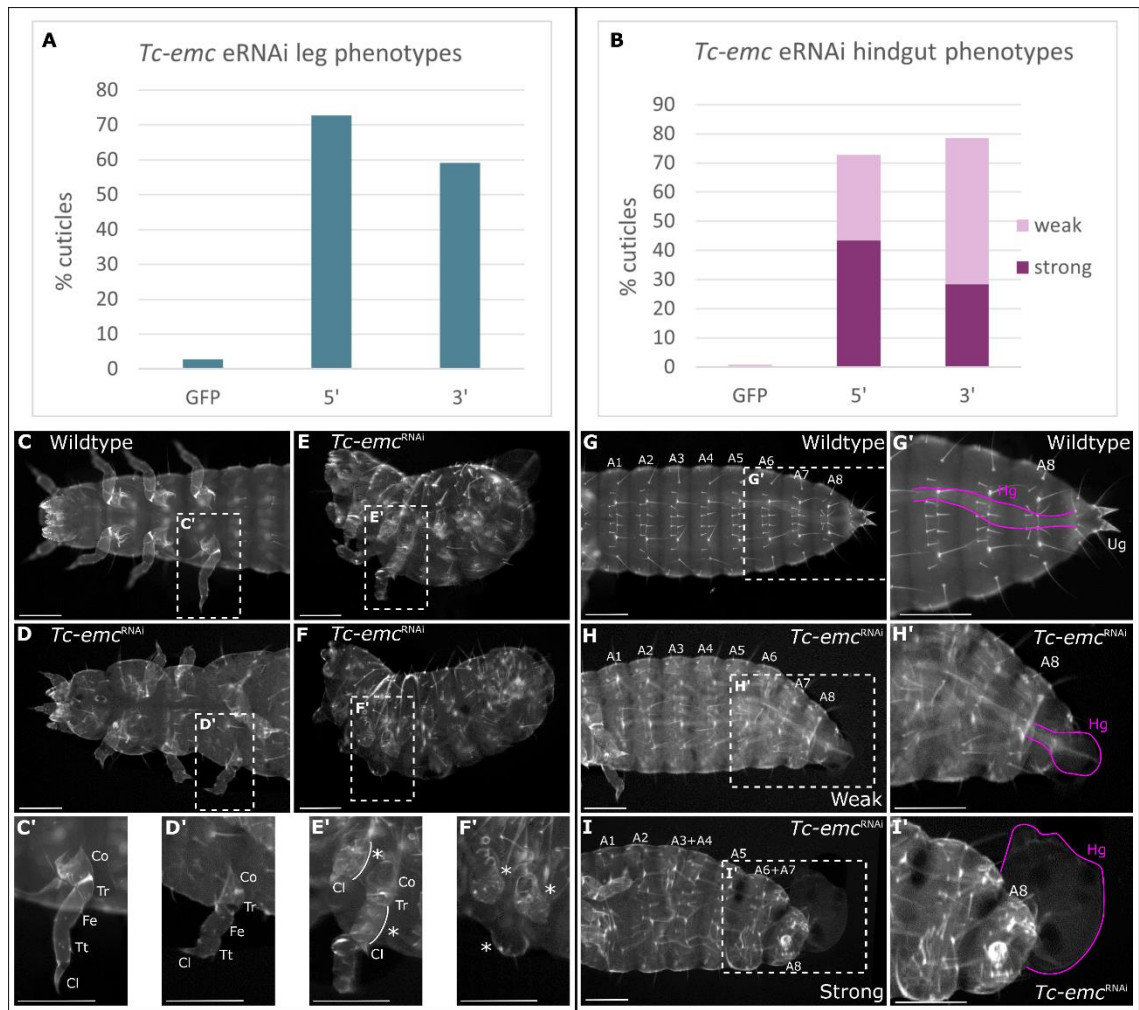


Figure 6.4. *Tc-emc* leg and hindgut phenotypes. (A, B) Overview of leg and hindgut abnormalities from *Tc-emc* eRNAi compared with *GFP* dsRNA injected control. (C, C') Wildtype cuticle with example leg. (D-F') *Tc-emc* RNAi induced leg phenotypes in increasing severity. (G, G') Wildtype cuticle with example hindgut outlined in magenta. (H, H') weak protruding hindgut phenotype. (I, I') Strong ballooned hindgut phenotype. Co = coxa; Tr = tronchanter; Fe = femur; Tt = tibiotarsus; Cl = pretarsal claw; Ug = urogomphi, Hg = hindgut. Asterisks mark indistinguishable leg segments. Scale bar: 100 μ m

6.2.4. *Tc-Sidpn1* RNAi induces abdominal truncation and anterior head defects

Consistent with previously published results (Bitra and Pali, 2010), parental RNAi of *Tc-Sidpn1* resulted in females producing a higher proportion of empty eggs than the control group injected with GFP dsRNA. This effect was particularly evident from the 3' injections, in which 96.5% of eggs collected had no visible signs of embryonic development, compared to just 11% in the control group (**Fig. 6.5A**). *Tc-Sidpn1* 5' injected females also produced more empty eggs than the GFP injected controls, although the difference was smaller.

The proportion of abnormal embryos obtained from the experimental and control groups did not differ significantly, and of the cuticles that were analysed, no consistent phenotypes were observed. A final collection 1 month after injections revealed that the percentage of empty eggs had dropped close to normal levels again in both the 5' and 3' groups (**Appendix D; Table D.3**). However, the proportion of wildtype and abnormal cuticles remained relatively constant as the effect of the knockdown diminished. Due to the severity of the empty egg phenotype, and the lack of intermediate phenotypes from pRNAi, it was necessary to use eRNAi to identify potential roles that *Tc-Sidpn1* may have in development.

Similarly to pRNAi, embryonic injections with *Tc-Sidpn1* 5' and 3' also resulted in a higher proportion of empty eggs, when compared with *GFP* dsRNA injected controls (**Fig. 6.5B**). As was the case in pRNAi, this effect was more pronounced in the 3' group, with 66.1% of injected eggs appearing to be empty, compared to 45.3% and 33.9% in the 5' and control groups respectively. In this case, of the 74 cuticles that developed from the *Tc-Sidpn1* 3' injected eggs, several phenotypes with high penetrance were observed. Most notably, 41.9% of the cuticles exhibited a truncated abdomen (**Fig. 5D, F-J**). In many of these cuticles, a crumpled section of indistinguishable ectoderm can also be seen at the posterior of the embryo (**Fig. 6.5F-H, J (white asterisks)**). While the occurrence of this phenotype was not limited to a specific point along the abdomen, over half of the truncations occurred around the A2 or A3 segments (**Fig. 6.5G-J; Appendix D; Table D.5**). Interestingly, the posterior expression domain of *Tc-Sidpn1* reported in chapter 5 (**Fig. 5.4 red arrowhead**), first appears at around the time that

the A2/A3 *Tc-wg* stripes are being specified, which, taken alongside the truncation phenotype observed here, is consistent with a role in posterior patterning from this stage in development.

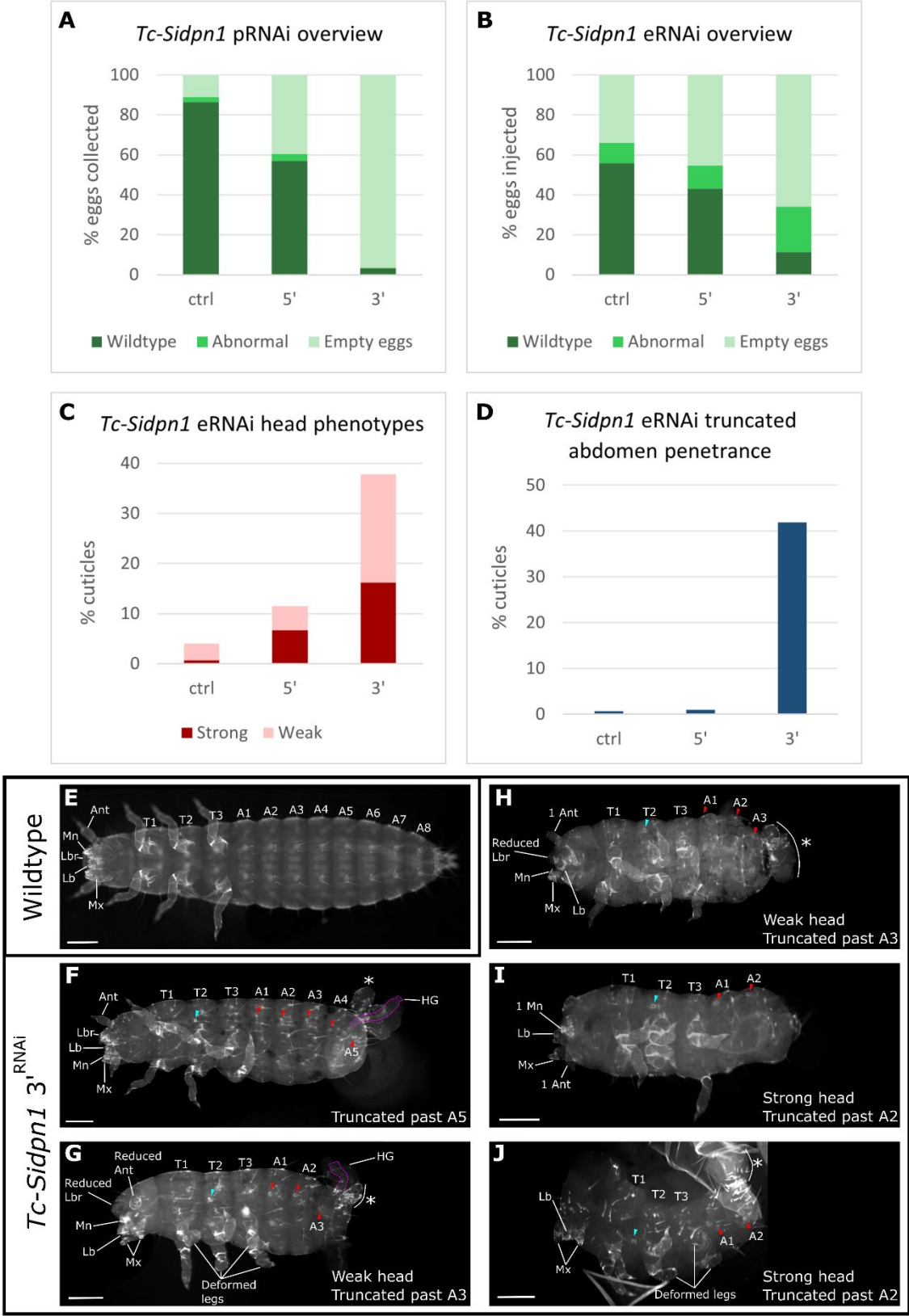


Figure 6.5. *Tc-Sidpn1* RNAi results. Summary results for pRNAi (A) and eRNAi (B), compared with *GFP* dsRNA injected controls. (C) Proportion of cuticles from eRNAi experiments with head abnormalities, split into weak (1-2 deformed/missing head appendages), and strong (3+ deformed/missing head appendages) phenotypes. (D) Proportion of cuticles from eRNAi experiments exhibiting a truncated abdomen. (E) Wildtype cuticle. (F-J) Cuticles from *Tc-Sidpn1* 3' eRNAi experiment displaying truncated phenotypes. Examples of weak and strong head abnormalities can be seen in (G, H) and (I, J) respectively. Thoracic (T1-T3) and abdominal (A1-A8) segments are labelled where possible. Arrowheads highlight thoracic (cyan) and abdominal (red) spiracles. Heads are labelled to highlight missing/reduced parts, and the number specified in instances where 1 of 2 head appendages is missing. White asterisks indicate indistinguishable parts of the cuticle. Hindgut (HG) is highlighted by magenta outline where visible. Scale bar: 100 μ m

In addition to a truncated abdomen, 37.8% of cuticles from the *Tc-Sidpn1* 3' injected eggs had some degree of head deformity, compared to just 4% from the *GFP* control (**Fig. 5C**). 26 out of the 28 head phenotypes recorded from the 3' injections, involved deformed or missing antennae, and in stronger head phenotypes, the labrum and mandibles were most commonly affected (**Fig. 6.5I, J; Appendix D; Table D.4**). Consistent with this observation, *Tc-Sidpn1* exhibits strong expression in several distinct regions of the head towards the end of GBE, including the antennal lobes (**Fig. 5.4**). 20.3% of cuticles from eggs injected with *Tc-Sidpn1* 3' dsRNA also exhibited some degree of leg deformity (**Appendix D; Table D.4**) in which legs were malformed and/or shortened (**Fig. 6.5G,J**). While *Tc-Sidpn1* does not appear to be expressed in the limb buds during, or immediately after GBE (**Fig. 5.4**), these leg phenotypes could be explained by later stage expression that was not captured in this instance. Additional *in situ* hybridisations would be necessary to confirm if this is the case or not.

While the penetrance of head, truncation, and leg phenotypes were relatively high in the 3' injected group, the same phenotypes were not observed in cuticles from eggs injected with *Tc-Sidpn1* 5' dsRNA. Of the few head abnormalities scored from the 5' group (**Fig. 6.5C**), there was no qualitative consistency as was the case for *Tc-Sidpn1* 3' cuticles.

6.2.5. Double knockdowns fail to reveal additive effects or redundancy between *Tc-h*, *Tc-dpn* and *Tc-emc*

To test for redundancy between co-expressed *Hes* genes and *Tc-emc*, dsRNA fragments from multiple genes were combined and injected together. *Tc-h/Tc-dpn* double pRNAi knockdowns resulted in a similar spread of head phenotypes as previously observed in *Tc-h* single knockdowns, although with a slightly lower penetrance, likely due to lower individual dsRNA concentrations (**Appendix D; Table D.7**). Furthermore, no additional abnormalities were recorded as the result of joint *Tc-h* and *Tc-dpn* knockdown.

Due to an essential role of *Tc-emc* in female reproduction, *Tc-emc/Tc-h* and *Tc-emc/Tc-dpn* double knockdowns were performed by eRNAi. *Tc-emc/Tc-h* and *Tc-emc/Tc-dpn* doubles produced very similar results to the *Tc-emc* single eRNAi experiments, and the severity and quality of segmentation phenotypes did not differ significantly (**Fig. 6.6A**). Consistent with the known role of *Tc-h* in head development, there was a slightly higher proportion of cuticles with deformed/missing mandibles, and/or missing heads in *Tc-emc/Tc-h* doubles, than in *Tc-emc* single and *Tc-emc/Tc-dpn* double knockdowns (**Fig. 6.6C, Appendix D; Table D.6**). Overall, however, there was no evidence to suggest redundancy between *Tc-emc*, *Tc-h* or *Tc-dpn*, in any developmental context.

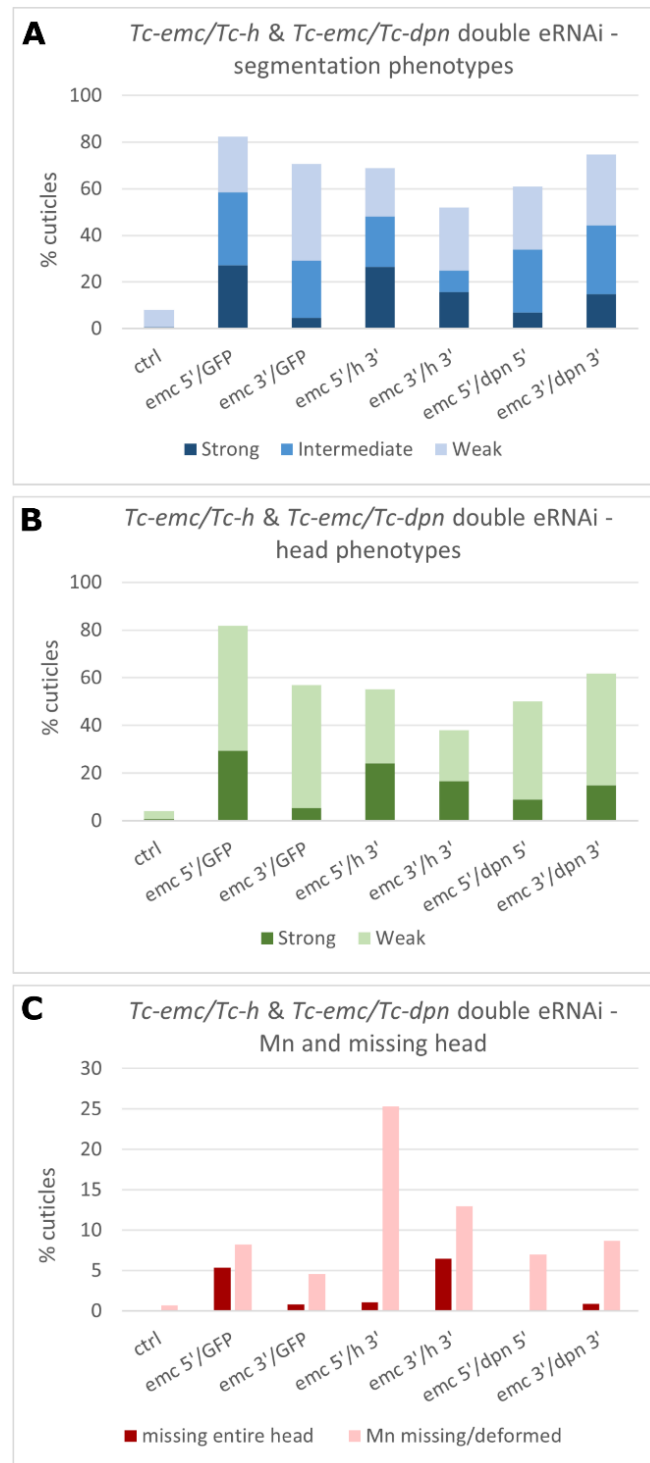


Figure 6.6. eRNAi results for *Tc-emc/Tc-h*, *Tc-emc/Tc-dpn* double knockdowns, compared to *GFP* control and *Tc-emc* single knockdowns. (A) Penetrance of segmentation phenotypes, as classified in **Fig. 6.2**. (B) Penetrance of head phenotypes as classified in **Fig. 6.3**. (C) Proportion of cuticles with missing/deformed Mn or missing entire head. Note that, *Tc-emc* dsRNA was mixed with *GFP* dsRNA at a 1:1 ratio to create an equal individual and total dsRNA concentration to double knockdown dsRNA mixes. This is to control for differences in penetrance that may be the result of dsRNA concentration.

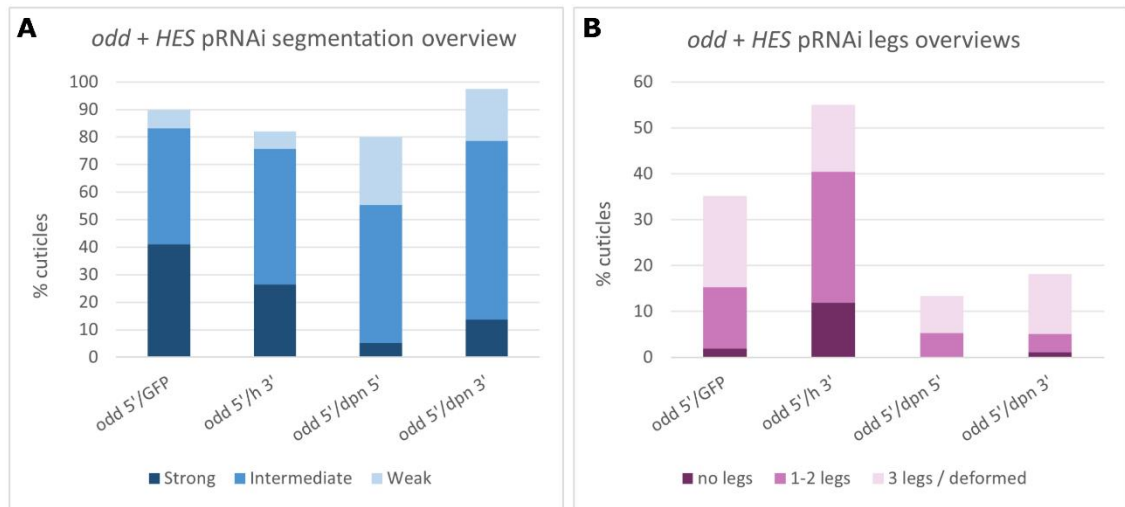
6.2.6. Testing for redundancy between *Hes* genes and *Tc-odd*

Knocking down *Tc-h* with *Tc-odd* results in the loss of additional legs and/or thoracic segments

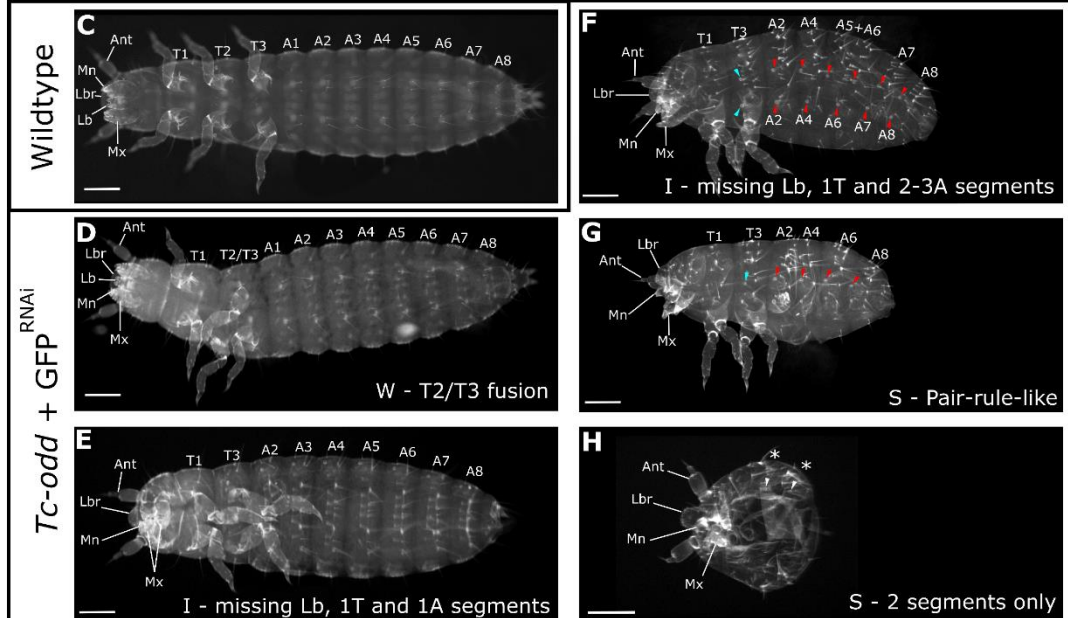
To test if *Tc-hairy* and/or *Tc-deadpan* are acting redundantly with the already established *Tc-odd/Tc-run/Tc-eve* oscillator, I performed double pRNAi knockdowns of *Tc-odd* with *Tc-h* and *Tc-dpn* in parallel with single *Tc-odd* knockdowns for comparison. Measures were taken to keep individual and total dsRNA concentrations the same in each case. To this end, I mixed *Tc-odd* dsRNA with *GFP* dsRNA (for singles) or *Tc-h/Tc-dpn* dsRNA (for doubles) at a 1:1 ratio so that each dsRNA fragment was being injected at a concentration of 1µg/µl, with a combined total concentration of 2µg/µl. Note that, as non-overlapping *Tc-odd* dsRNA fragments had already been confirmed to induce efficient knockdowns (Sharma pers comms), only one dsRNA fragment (5') was used in these experiments.

The *Tc-odd* single knockdowns yielded a consistent progression of segmentation phenotypes. In weak *Tc-odd* knockdowns, the labial, T2 and A1 segments are the first to exhibit signs of malformation, fusion, or loss (**Fig. 6.7D**). In intermediate cuticles, additional abdominal segments become fused or lost in a pair-rule-like fashion from anterior to posterior (**Fig. 6.7E-F**). Strong *Tc-odd* knockdown results in a classic pair-rule-like phenotype, in which every other body segment is missing (**Fig. 6.7G**). This was observed in 27% of the cuticles collected from the *Tc-odd/GFP* injected females. A further 14.2% of cuticles from this control were missing additional thoracic and/or abdominal segments in the most severe examples of *Tc-odd* knockdowns (**Fig. 6.7H**).

When *Tc-h* and *Tc-dpn* were knocked down alongside *Tc-odd*, there were no drastic impacts on segmentation, and in fact, the proportion of strong segmentation phenotypes were slightly lower in these groups compared with the *Tc-odd/GFP* group (**Fig. 6.7A**). Evidence of *Tc-h* related head defects were seen in the *Tc-h/Tc-odd* double, but with lower penetrance than previous experiments. Just 14.3% of cuticles displayed mandibular defects/loss, and only 4% were missing their entire head, still comparatively higher than the 4.9% and 0.5%, respectively, recorded from the *Tc-odd/GFP* injections (**Appendix D; Table D.8**).



odd phenotypic series



odd + h thoracic/leg phenotypes

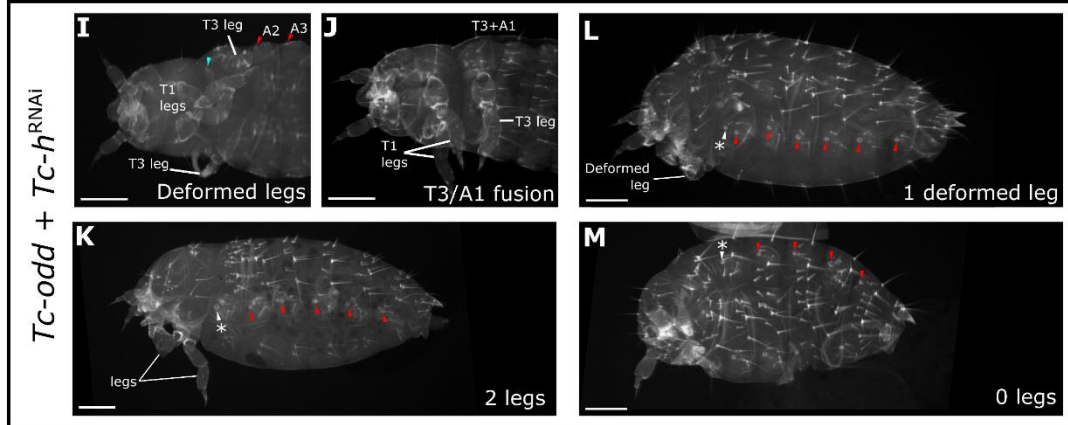


Figure 6.7. *Tc-odd/Tc-h* and *Tc-odd/Tc-dpn* double knockdown results. (A) Proportion of cuticles with weak (segment fusions and/or 1 missing segment), intermediate (2-4 missing segments), and strong segmentation phenotypes (5 or more missing segments). (B) proportion of cuticles with no legs, 1 to 2 legs, and 3 legs / leg deformities. (C) Wildtype cuticle. (D-H) Example cuticles from *Tc-odd* single knockdowns, showing segmentation phenotypes in order of increasing severity. W=Weak, I=Intermediate, S=Strong. (I-M) Example cuticles from *Tc-odd/Tc-h* double knockdown, displaying thoracic/leg phenotypes in addition to missing T2. (I) Deformed T3 legs. (J) Missing T3 leg due to T3/A1 merge. (K) 2 legs. (L) 1 leg. (M) 0 legs. Head appendages are labelled in panels C to E to highlight loss of Labia (Lb). Thoracic (T1-T3) and abdominal (A1-A8) segments labelled where relevant. Arrowheads highlight thoracic (cyan), and abdominal (red) spiracles, and in cases where this distinction can't be made, white arrowheads with an asterisk are used. Scale bar: 100 μ m

The most notable effect of *Tc-h* knockdown with *Tc-odd*, was a higher incidence of missing legs, in addition to those lost from the missing T2 segment (**Fig. 6.7B**). Cuticles showed evidence of leg specific defects, as well as missing legs due to thoracic segment fusions (**Fig. 6.7I,J**). In many cases, it was not possible to tell apart missing legs as the result of severe leg defects, or from the loss of thoracic segments. This is because, without the presence of legs, thoracic and abdominal segments are almost indistinguishable, making it difficult to interpret which segments are missing from cuticle preps alone (**Fig. 6.7K-M**). Furthermore, no clear qualitative differences could be seen between leg phenotypes in *Tc-odd* single and *Tc-odd/Tc-h* doubles, with thoracic segment merges and deformed legs seen at similar frequencies in both (**Appendix D; Table D.8**). However, in the *Tc-odd* pRNAi single, leg phenotypes were more strongly associated with severe segmentation phenotypes, suggesting that missing additional legs or thoracic segments reflects a stronger *Tc-odd* knockdown. From this group, 89.8% of cuticles with only 1 or 2 legs also had a strong segmentation phenotype. In contrast, only 46.1% of cuticles with 1 or 2 legs from the *Tc-odd/Tc-h* double also had a strong segmentation phenotype. The higher number of cuticles with 2 or fewer legs, in light of a lower proportion of strong segmentation phenotypes, suggests that this is the result of a real additive effect from *Tc-h* knockdown with *Tc-odd*.

While there were no clear effects on development from knocking down *Tc-dpn* with *Tc-odd*, the rate of egg lay was reduced slightly, with the number of eggs laid per female being 26.4% and 23.2% lower in the *Tc-odd* 5'/*Tc-dpn* 5' and *Tc-odd* 5'/*Tc-dpn* 3' respectively, when compared with the *Tc-odd* 5'/*GFP* control (**Table S10**). This reduction in egg lay is consistent with previously published results for *Tc-dpn* RNAi (Bitra and Pali, 2010)

Tc-h, but not *Tc-dpn*, knockdowns impact leg and spiracle counts in *Tc-odd*-null embryos

The penetrance of *Tc-odd* and *Tc-h* induced phenotypes are weaker in double knockdowns, compared with singles, either due to lower dsRNA concentrations, or from a limited capacity of RNAi machinery in the developing embryo. To ensure the strongest possible knockdowns of *Tc-h* and *Tc-dpn* in parallel with *Tc-odd*, I carried out *Tc-h* and *Tc-dpn* singles, and *Tc-h/Tc-dpn* double knockdowns in a *Tc-odd* knockout strain.

Hemizygote females were injected and then bred with hemizygote males. Based on mendelian inheritance, this on average should result in 25% of the offspring containing both copies of the *Tc-odd* mutant allele, and therefore exhibiting a complete knockout of *Tc-odd* expression during development. *Tc-odd*-null mutants were easily distinguishable from other abnormalities and were identified by their similarity to the most severe phenotypes from the *Tc-odd* pRNAi (**Fig. 6.7G-H**). Due to the severity of the *Tc-odd*^{-/-} phenotype, accurately scoring segmentation based on bristle patterns was not possible, therefore spiracles were chosen as a proxy measure for the number of body segments that had successfully developed. The final counts were taken from whichever side of the embryo had the most spiracles visible (**Fig. 6.9**). In addition to a spiracle count, I also noted the number of legs and absence of head appendages in each cuticle (**Fig. 6.9**).

The percentage of cuticles that were identified as *Tc-odd*-null did not differ greatly between the experimental and control groups, and in most cases, was slightly below the 25% predicted (**Appendix D; Table D.9**). There was a large degree of variation in the severity of the *Tc-odd*^{-/-} phenotype, even among embryos from *GFP* dsRNA

injections (**Fig. 6.8C; Fig. 6.9A-J**). A range of cuticles with spiracle counts between 0 and 5 were recorded in the GFP injected group, and a range of 0 to 4 spiracles was observed in all other groups. There was also variation in the number of legs counted in *odd*-null cuticles from each group.

In *Tc-dpn* single knockdowns, there were no consistent differences in *Tc-odd*^{-/-} cuticle scores across 5' and 3' injected groups, when compared with the control (**Table S10**). All groups that were treated with *Tc-h* dsRNA, however, displayed a trend towards fewer legs and spiracles, when compared with the control. *Tc-h* single and *Tc-h/Tc-dpn* double knockdowns produced a higher proportion of *Tc-odd*^{-/-} cuticles with 0 visible legs (**Fig. 6.8B**). This effect was most drastic in the *Tc-h* and *Tc-h/Tc-dpn* 3' groups, in which 68.2% and 71.2% of *Tc-odd*^{-/-} cuticles respectively were missing legs, compared to just 24.2% in the control. *Tc-odd*^{-/-} cuticle spiracle counts were distributed across a range of 0 to 5 in the control, and 0 to 4 in the *Tc-h* single and *Tch/Tc-dpn* double knockdowns. However, this distribution is skewed towards a lower average number of spiracles in the *Tc-h* single and *Tc-h/Tc-dpn* double knockdowns, compared with the control (**Fig. 6.8C**). There was also a smaller proportion of *Tc-odd*^{-/-} cuticles with 0 or 1 spiracles from the control group (18.5%), than from *Tc-h* (44.3%), *Tc-h/Tc-dpn* 5' (41.3%), and *Tc-h/Tc-dpn* 3' (56.2%) injected females (**Appendix D; Table D.9**). Consequently, there were also fewer *Tc-odd*^{-/-} cuticles with 3 or more spiracles from the *Tc-h* (19.3%), *Tc-h/Tc-dpn* 5' (25%), and *Tc-h/Tc-dpn* 3' (17.8%) knockdowns compared with the control (52.4%).

Apart from these trends, and known *Tc-h* related head phenotypes (deformed/missing Mn, **Appendix D; Table D.9**), there were no discernible differences between the cuticles from the control and experimental groups (**Fig. 6.9**). There was also no evidence to suggest the knockdown of *Tc-dpn* with *Tc-h* in the *Tc-odd* mutant had any additional effect from what was seen in the *Tc-h* single knockdowns. Any other subtle, additive effects that may exist could be clouded by the large degree of variation and severity of *Tc-odd*^{-/-} development.

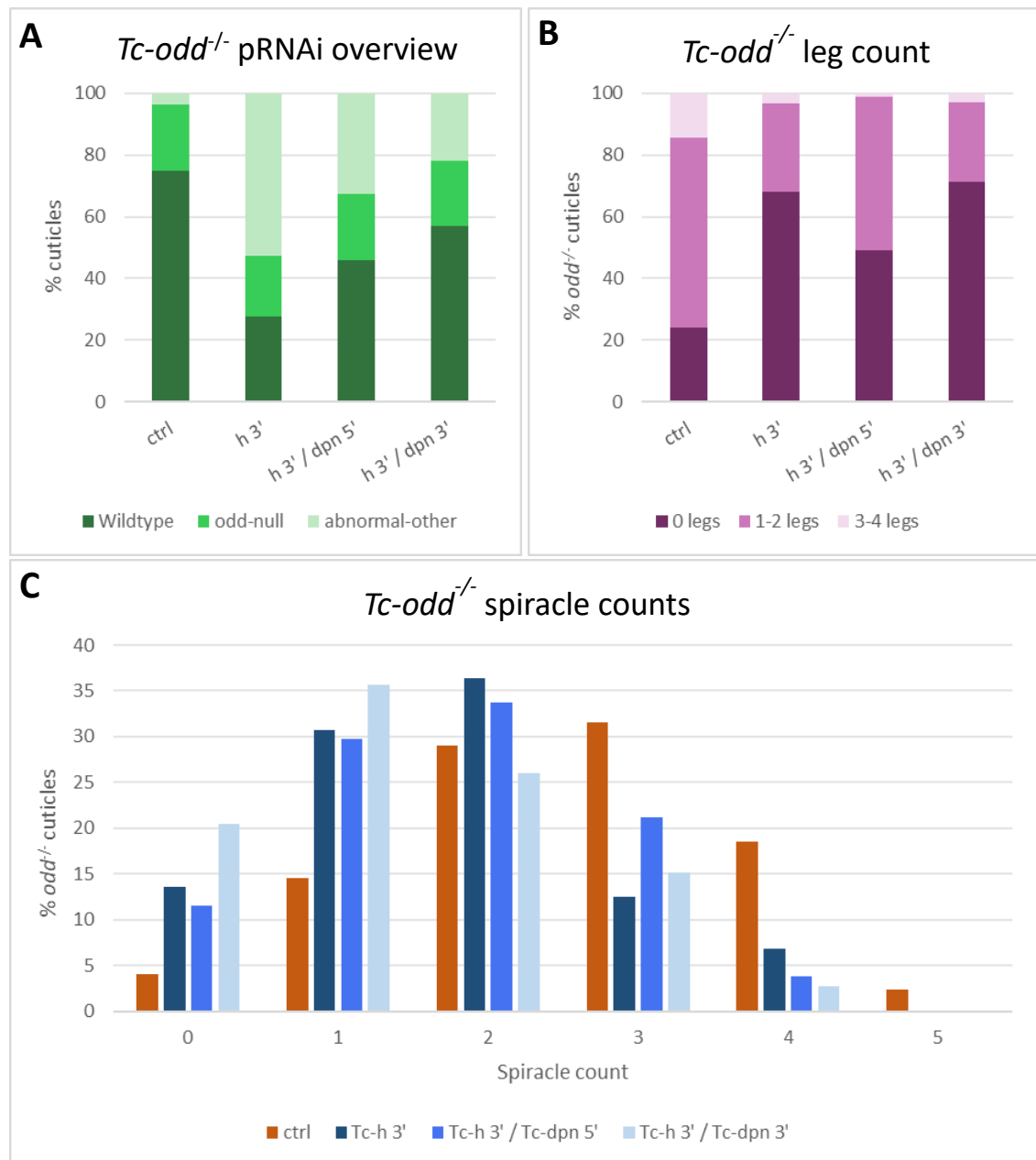


Figure 6.8. *Tc-h* single and *Tc-h/Tc-dpn* double knockdowns in *Tc-odd*-null mutant. (A) Proportions of cuticles that are Wildtype, *odd*-null, and have other abnormalities. (B) Proportions of *odd*-null cuticles with 0 legs, 1-2 legs, and 3-4 legs. (C) Distributions of spiracle counts for *odd*-null cuticles from experimental groups versus *GFP* control. The control group were injected with *GFP* dsRNA. See Table D.9 (**Appendix D**) for more details.

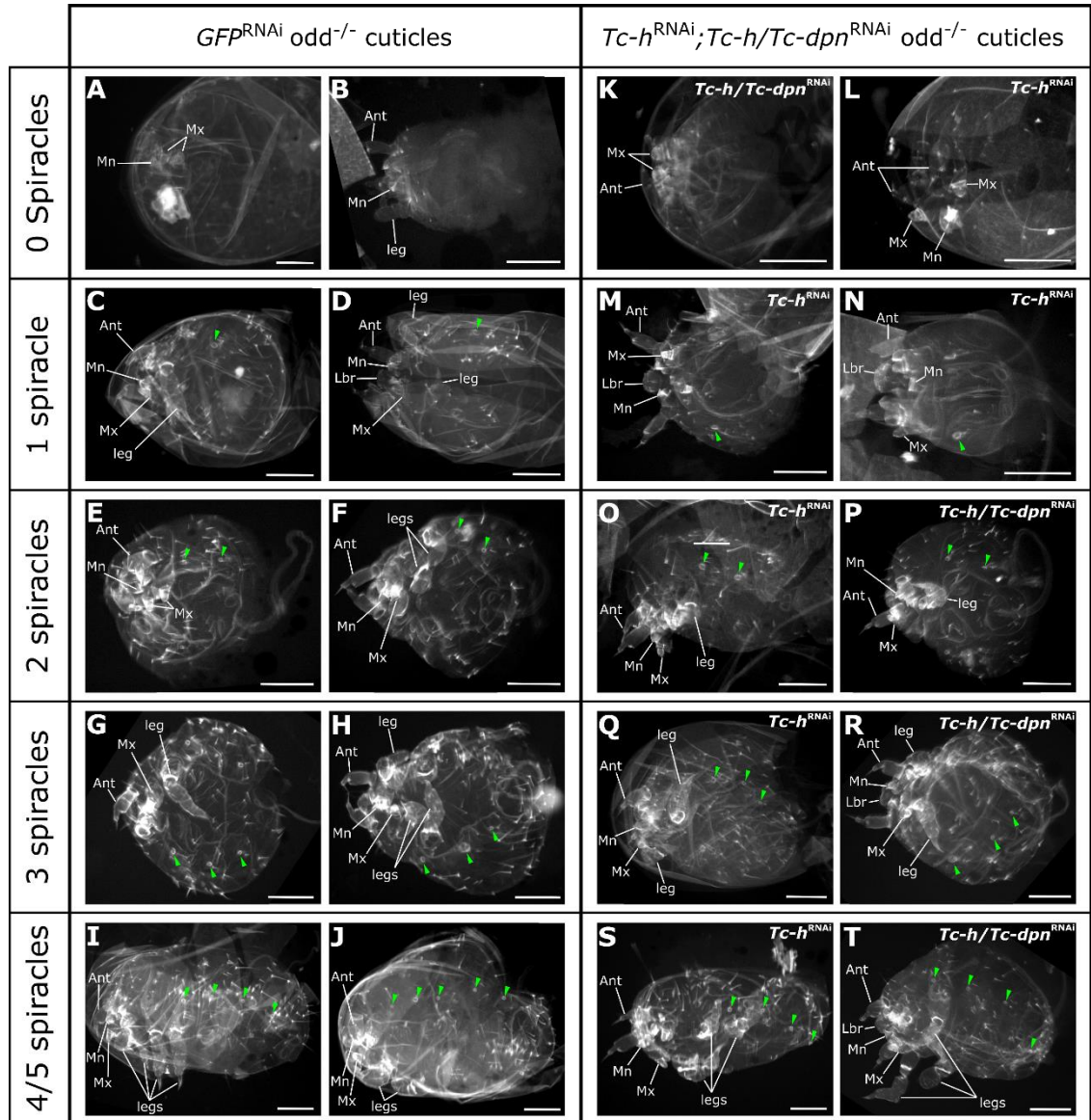


Figure 6.9. *Tc-odd*^{-/-} cuticle scoring. Example cuticles from *GFP* injected controls (A-J), and from *Tc-h* single and *Tc-h/Tc-dpn* double knockdowns (K-T). Spiracle counts were used as a proxy for measuring the severity of the segmentation phenotype. Cuticles are arranged in rows corresponding to these counts: 0 (A,B,K,L), 1 (C,D,M,N), 2 (E,F,O,P), 3 (G,H,Q,R), and 4 (I,S,T) / 5 (J). Spiracles are indicated by green arrowheads. In each case, all visible legs, including those deformed, are labelled to show how these were also counted. Head appendages are also labelled where easily identified. Scale bar: 100 μ m

6.3. DISCUSSION

In this chapter, I have identified novel functions for *Tc-emc* in the patterning of abdominal segments, as well as additional roles in head, leg, hindgut, and dorsal closure development. Embryonic RNAi has also revealed a possible role for *Tc-Sidpn1* in germband extension. Finally, a combination of *Tc-odd/Tc-h* double knockdown and *Tc-h* single knockdowns in a *Tc-odd* knockout mutant, has exposed a degree of redundancy between these two genes in relation to leg development and/or thoracic segmentation.

6.3.1. Phenotypes are consistent with a role for *Tc-emc* as a mediator of Notch signalling in *Tribolium*

To date, the only published *Tc-emc* knockdowns in *Tribolium* have been done using pRNAi, which causes a severe reduction in female egg laying (Bitra and Pali, 2010). As such, its role in *Tribolium* development has not been characterised until now. Using eRNAi, I have been able to bypass *Tc-emc*'s essential role in oogenesis and knockdown the gene during embryogenesis.

Several of the phenotypes I observed are consistent with phenotypes described in *Drosophila* crosses containing an *emc*-null allele (Saito et.al., 2002). In these *Drosophila* mutants, the dorsal head fails to migrate anteriorly, leaving gnathal segments exposed. This is consistent with the missing labrum (and often antennae) phenotype that was seen in a high proportion of *Tribolium* embryos after *Tc-emc* dsRNA injections. *Drosophila* embryos carrying the *emc*-null allele also exhibit problems with dorsal closure and *emc* is found to regulate cell chirality and left-right asymmetry in the *Drosophila* hindgut (Ishibashi et.al., 2018). These results demonstrate that *emc* orthologues have conserved pleiotropic roles in head, gut, and dorsal closure patterning across *Tribolium* and *Drosophila*. In addition to functions already described in *Drosophila*, *Tc-emc* knockdowns in *Tribolium* have also been found to generate segmentation and leg phenotypes. In both cases, segments, either abdominal or of the

leg, become fused or are missing entirely. Regarding body segmentation, this often resulted in cuticles with short and disorganised abdomens.

While the regulatory relationships of *Tc-emc* have not been studied in *Tribolium*, the *Drosophila emc* orthologue is known to act downstream of the Notch signalling pathway during eye development (Bhattacharya and Baker, 2009). In this context, *emc* expression is elevated in cells surrounding ectopically expressed Notch ligands, Delta (Dl), and Serrate (Ser), regardless of location in the eye imaginal discs. Similar regulatory control of *emc* by Notch signalling is also observed during wing vein morphogenesis in *Drosophila* (Baonza et.al., 2000). Furthermore, there is evidence to suggest that vertebrate homologues, *Id* genes, are in some cases also downstream targets of Notch signalling (Reynaud-Deonauth et.al., 2002), and the interplay of ID proteins with Hes factors is an essential part of downstream Notch function in the context of vertebrate neurogenesis (Boareto et.al., 2017), demonstrating an ancient association between these genes and Notch signalling in metazoans. It is therefore interesting that in *Tribolium*, *Tc-emc* knockdown produces several phenotypes that closely resemble those from experiments in which the Notch signalling pathway is perturbed.

Tc-Notch eRNAi by Kuz et.al., (2013) has been reported to result in shortened and twisted abdomens. Importantly, it is the abdominal segments that are most effected, as is also the case in *Tc-emc*^{RNAi} embryos. In the strongest *Tc-Notch* knockdowns, the normal striped expression of the segment polarity gene, *Tc-engrailed* (*Tc-en*), is barely detectable or interrupted, while in weaker knockdowns, the posterior-most stripes are bunched closer together (Kux et.al., 2013). These results closely reflect the range of segmentation defects that were seen in cuticles after *Tc-emc* knockdown, especially as weak phenotypes most often present themselves in the posterior segments (**Fig. 6.2B**). Taken together with what we know about the interaction between Notch signalling and *emc* in *Drosophila*, it is plausible that in *Tribolium*, *Tc-emc* functions as a downstream effector of Notch signalling in the context of segmentation. However, genes from the Notch signalling pathway, *Tc-Notch*, *Tc-Delta*, *Tc-Serrate* and *Tc-fringe*, are not dynamically expressed in the SAZ, as is true for *Tc-emc* (Aranda et.al., 2008).

Tc-Notch is expressed ubiquitously during GBE, while *Tc-Delta* expression is observed in the mesoderm and a posterior domain (Aranda et.al., 2008). Given that *Tc-emc* also

appears to be ubiquitously expressed and elevated in the mesoderm early in development, it is possible that these patterns reflect regulation by Notch signalling. However, this is insufficient to explain the elevated stripes of *Tc-emc* expression in the SAZ. A closer examination of *Tc-Notch* expression or of other members of the pathway would be required to confirm if indeed *Tc-emc* might be a downstream target of Notch signalling in the SAZ.

Knockdown of *Tc-emc* in *Tribolium* embryos also affects leg development, resulting in legs with shortened segments and fusions, and in severe cases legs that are reduced to indistinguishable nubs (**Fig. 6.4D'-F'**). Notch signalling has been proposed to maintain a highly conserved role in the mediation of appendage segmentation across arthropods (Prpic and Damen, 2009). Disrupting the pathway results in leg segments that are shortened and fused in several different insect species, including *Drosophila*, *Gryllus bimaculatus*, and *Tribolium* itself (Rauskolb and Irvine, 1999; Angelini et.al., 2012 (a); Mito et.al., 2011). These phenotypes closely resemble the legs of *Tc-emc*^{RNAi} cuticles, and further support the connection between Notch signalling and *Tc-emc* in *Tribolium*.

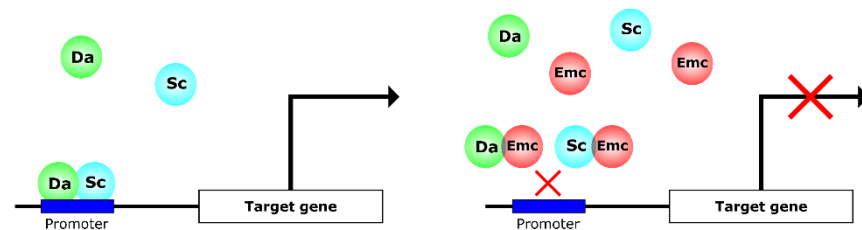
Finally, knockdowns of two *Tribolium* Notch signalling genes, *Tc-Serrate* and *Tc-mind bomb1* (*Tc-mib1*), have been reported to result in a high penetrance of cuticles with missing labrums (Siemanowski et.al., 2015). Further investigation deduced that Notch signalling is an important upstream regulator of cell proliferation in the developing *Tribolium* labrum. Interestingly, *Tc-emc* exhibits strong expression in the labrum anlage similar to that of *Tc-ser* post-GBE (Aranda et.al., 2008). Taken together with the missing labrum phenotype observed in a high proportion of *Tc-emc*^{RNAi} cuticles, this is yet more evidence to suggest that *Tc-emc* is an essential effector of Notch signalling in *Tribolium*.

6.3.2. *Tc-emc* phenotypes: A reflection of bHLH function in abdominal segmentation?

While there is multiple evidence to suggest that *Tc-emc* and *Notch* are closely linked in *Tribolium* development, the precise mechanism of how *Tc-emc* itself controls segment patterning has not been identified in this instance. As discussed in chapter 5, *emc* encodes a bHLH protein that lacks the basic domain necessary for binding DNA, and is

therefore able to inhibit other bHLH protein activity through the formation of inactive heterodimers (Campuzano, 2001). In *Drosophila*, Emc is known to dimerize with proneural genes from the *aecchaete-scute-complex* (AS-C) as well as daughterless, inhibiting their roles as promoters and subsequently repressing their target genes (**Fig. 6.10A**) (Van Doren et.al., 1991; Cabrera et.al., 1994). Through the same mechanism, it is also possible that Emc can promote gene expression by inhibiting transcriptional repressors (**Fig. 6.10B**), making Emc function highly versatile. There is evidence from *Drosophila* eye development, that *emc* mediates Notch function by activating *E(spl)* expression (Battycharya and Baker, 2009), demonstrating its potential as a passive activator as well as a repressor, although the precise molecular mechanism in this case is unknown.

A Passive repression



B Passive activation

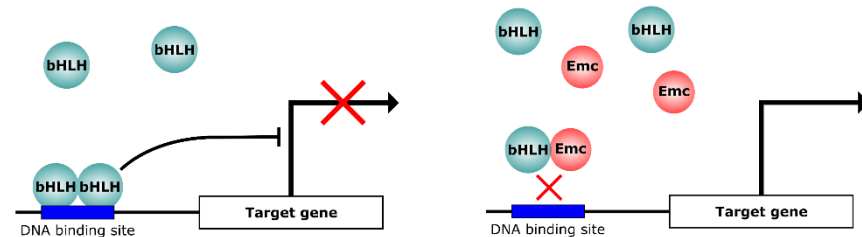


Figure 6.10. Emc gene regulation. (A) The presence of Emc indirectly inhibits the activation of Daughterless(Da)/Scute(Sc) target genes by forming heterodimers that are unable to bind to the promoter region. This is a known example of passive repression by Emc in *Drosophila* (Cabrera et.al, 1994). (B) Emc also has the potential to indirectly activate transcription of a target gene by forming heterodimers with bHLH transcription factors that normally inhibit expression.

It is possible that the segmentation phenotype induced by *Tc-emc* knockdown is related to a bHLH gene with a segmentation function that relies on regulation by Tc-Emc dimerization. Currently, the only bHLH genes known to be co-expressed with *Tc-emc* in the SAZ during segment addition are *Tc-h* and *Tc-dpn*. However, the results from knockdowns of these genes once again failed to capture any evidence for roles relating to segmentation, as previously reported by Aranda et.al. (2008). That said, if a *Hes*-based oscillator is active in the *Tribolium* SAZ, it may be the case that perturbing *Tc-emc* expression is more effective in disrupting the network than targeting the *Hes* genes themselves. This is especially true for *Tc-dpn*, which has been shown to have a particularly low knockdown efficiency from RNAi (~50% reduction in mRNA levels), even with high dsRNA concentrations (5µg/µl) (Bitra et.al., 2009). Indeed, in a scenario where the Tc-Emc protein dimerizes with Tc-H and Tc-Dpn proteins, *Tc-emc* knockdown would have the effect of increasing Tc-H and Tc-Dpn availability, and consequently increasing their regulatory activity. This would be more akin to overexpressing *Tc-h* or *Tc-dpn*, rather than knocking them down, and therefore could explain the differences between the phenotypes or lack thereof.

Alternatively, *Tc-emc* is interacting with other bHLH genes in the SAZ that are, as of yet, unidentified. Of the 53 known bHLH genes in *Tribolium* (Bitra et.al., 2009), there are no other published examples of striped expression in the SAZ. However, the vast majority have not been investigated. *Tribolium* only has two AS-C genes, *achaete-scute-homolog* (*Tc-ASH*), and *asense* (*Tc-ase*), both of which are expressed exclusively in the neuroectoderm during development (Wheeler et.al., 2003). *Tc-da* expression has not yet been described in *Tribolium*, although in *Drosophila*, *da* is expressed ubiquitously throughout development (Cronmiller and Cummings, 1993; Berkley *Drosophila* Genome Project). Single cell transcriptomics across the SAZ during GBE could be used to identify other bHLH genes co-expressed with *Tc-emc* in order to establish any new segmentation candidates.

6.3.3. A possible role for *Tc-Sidpn1* in posterior signalling during GBE

Consistent with previously published pRNAi results for *Tc-Sidpn1*, injected females laid a high proportion of eggs that contained no visible cuticles (Bitra and Pali, 2010). As a result, eRNAi was chosen to elucidate any roles of *Tc-Sidpn1* in embryogenesis. While this also resulted in a higher number of empty eggs than control injections, a high penetrance of head and truncation phenotypes were observed for eggs injected with *Tc-Sidpn1* 3' dsRNA. The failure to replicate these results with the non-overlapping 5' dsRNA casts doubts over the legitimacy of these findings. However, the lower proportion of empty egg phenotypes from the 5' pRNAi group suggests that the 5' fragment elicits a weaker response. It is therefore possible that the knockdown achieved from 5' embryonic injections did not meet the threshold required to induce the same phenotypes observed in the 3' injections. For publications that use non-overlapping dsRNA controls for RNAi in *Tribolium*, there can sometime be a quantitative degree of variation in the results observed for each fragment (when such data is included) (Clark and Peel, 2018; Garcia-Perez et.al., 2021). These highlight the variability in RNAi efficiency that can exist between different sequences from the same gene. 5' and 3' fragments were both confirmed by sequencing, and BLAST of the 3' fragment against the *Tribolium* genome returned no other significant hits, ruling out the likelihood of these phenotypes being caused by off target effects.

To date, there have been no studies on the expression pattern or function of *Tc-Sidpn1* orthologues in any other arthropods. As reported in chapter 5, *Tc-Sidpn1* is predominantly expressed in the developing nervous system, with expression first appearing in the head at the start of GBE. However, a posterior domain also appears at around the point that the *Tc-wg* stripe corresponding to A1 (PS6) is established (**Fig. 5.4**). The timing of this posterior *Tc-Sidpn1* expression correlates to the truncation phenotype observed, in that development is arrested at segments A2 or A3 in the majority of cases (**Fig. 6.5G-J; Appendix D; Table D.5**). This suggests that this posterior *Tc-Sidpn1* domain is directly involved in promoting GBE and/or segmentation.

As a *Hes* gene, it is possible that *Tc-Sidpn1* is also regulated by the Notch signalling pathway, which is known to play an important role in the posterior growth zone of

species from across the arthropod phylum (reviewed in Williams and Nagy, 2017). Consistent with this idea, eRNAi knockdown of the Notch ligand, *Tc-dl*, produces a truncated phenotype in *Tribolium* (Aranda, 2006). In severe cases, these cuticles lack all abdominal segments posterior of A2/A3, as is seen in *Tc-Sidpn1* knockdowns. Interestingly, it is also expressed in a posterior domain during GBE (Aranda, 2006). These observations, along with those made for *Tc-Sidpn1* here, are consistent with a functional relationship between *Tc-Sidpn1* and the Notch signalling pathway in the context of posterior patterning. However, more evidence would be required to confirm if this is true.

6.3.4. Joint knockdown of *Tc-h* with *Tc-odd* exposes redundancy in thoracic segment patterning and/or leg development

Knocking down *Tc-h* or *Tc-dpn* with *Tc-odd* did not result in the complete collapse of segmentation, as might have been expected if these genes were part of a redundant oscillator controlling segmentation alongside the *Tc-odd/Tc-run/Tc-eve* circuit. However, *Tc-h/Tc-odd* double knockdowns did appear to produce an additive effect in which the number of legs that successfully developed was reduced in cuticles from these injections compared to those from single *Tc-odd* knockdowns. These results were further backed up when *Tc-h* knockdowns were performed in a *Tc-odd* knockout mutant. Again, there was a higher proportion of *Tc-odd*-null cuticles with fewer legs from females injected with *Tc-h* dsRNA when compared to the *GFP* dsRNA injected group. In addition to this, the average spiracle count was also lower in *Tc-odd*^{-/-} cuticles in which *Tc-h* was knocked down, suggesting fewer segments were successfully established. However, a large degree of variation was seen for both the *GFP* control and *Tc-h* knockdown groups, and the severity of the *Tc-odd*^{-/-} phenotype made it difficult to closely examine these differences. There are two possible explanations for why the knockdown of *Tc-h* with *Tc-odd* resulted in additional missing legs in these experiments.

Explanation 1: Redundancy between *Tc-h* and *Tc-odd* in leg development

The first possible explanation for a reduced number of legs, is that *Tc-h* has a redundant relationship with *Tc-odd* in the context of leg development. Deformed legs were observed in both the *Tc-h/Tc-odd* double knockdowns, and the *Tc-h* single knockdowns in *Tc-odd*^{-/-} embryos. It is possible that these represent mild knockdowns in a phenotypic series that leads to arrested leg development in more severe cases. In *Drosophila*, *odd* and *drm* act redundantly to mediate Notch function in leg segmentation (Hao et.al., 2003), and in *Tribolium*, members of the *odd* gene family are essential for normal leg segmentation in embryogenesis and metamorphosis (Angelini et.al., 2012 (a); Thümecke and Schröder, 2021). The expression pattern of *Tc-odd* in the leg anlagen is also consistent with a conserved role in leg development in *Tribolium* (**Fig. 4.3**). Furthermore, *Tc-hairy* is co-expressed with *Tc-odd* in the limb buds prior to morphogenesis (**Fig. 5.1**). This, coupled with the relationships that both *Tc-odd* and *Tc-h* may have with Notch signalling, makes redundancy between these two genes plausible in the context of leg development.

While there is certainly an argument to be made here, there are several factors that suggest this is unlikely to be the case. Firstly, if there was redundancy relating to leg development, then we might expect to see a higher proportion of cuticles with intermediate leg phenotypes in *Tc-h/Tc-odd* doubles, than *Tc-odd* singles. However, in both cases, only a small percentage of cuticles had deformed legs (**Appendix D; Table D.9**). Furthermore, if the joint knockdown causes a missing leg phenotype, then it would most likely impact all the legs, rather than leaving one or two normal legs, as is often observed. Finally, redundant roles in leg development do not explain the lower average spiracle counts observed in *Tc-odd* null mutants where *Tc-h* is knocked down.

Explanation 2: *Tc-h* and *Tc-odd* redundancy in thoracic segmentation

Perhaps a better explanation for why fewer legs are observed in embryos when *Tc-h* and *Tc-odd* are knocked down together, is that thoracic segmentation is being disrupted. This would explain the low penetrance of intermediate leg phenotypes in *Tc-*

h/Tc-odd double knockdowns, and the lower average spiracle count in *odd* null mutants with *Tc-h* knocked down.

Tc-odd itself already has a clear role in thoracic segmentation. The T2 segment is usually the first to be lost in weak *Tc-odd* single knockdowns (**Fig. 6.7**). In stronger *Tc-odd* knockdowns, there are often signs of additional thoracic segment merges and losses. In *Tc-odd*^{-/-} embryos from *GFP* dsRNA control injection experiments, there is also a range of phenotypes that includes cuticles with no visible spiracles or legs (**Fig. 6.9A**). Despite this, the majority of *Tc-odd*^{-/-} cuticles from this control group still have some legs, even if those legs are deformed, suggesting that thoracic segments can still be established in the complete absence of *Tc-odd* expression. In contrast, most of the *Tc-odd*^{-/-} cuticles from *Tc-h* single, and *Tc-h/Tc-dpn* double knockdowns had no visible legs (**Fig. 6.8B**), perhaps indicating the loss of all thoracic segments. If this is the case, then the thoracic spiracle might also be lost, which would explain why fewer spiracles were observed on average in these groups, compared to the control.

Tc-odd and *Tc-h* are expressed in the SAZ in nearly opposite phase throughout *Tribolium* segmentation, including the period in which thoracic parasegment boundaries are being established (**Fig. S**). While striped expression of *Tc-eve* is completely lost in *Tc-odd*^{-/-} embryos, and *Tc-run* stripes are poorly resolved, *Tc-h* continues to be expressed in clear stripes in the SAZ (**Fig. 1.9**), suggesting that *Tc-h* oscillations are independent on those of *Tc-odd* or *Tc-eve* (Sharma, pers comms). This, along with the results presented here, suggest that sustained oscillations of *Tc-h* in *Tc-odd*^{-/-} embryos may partially rescues thoracic segment patterning. However, given that some spiracles are often still visible when *Tc-h* is knocked down, this would suggest that abdominal segments are being partially rescued by a separate mechanism. The lack of any clear additive effects on abdominal segmentation in *Tc-h/Tc-odd* double knockdowns also supports the notion that redundancy between these genes is limited to thoracic segment patterning only, or that there is another redundant element controlling abdominal segmentation. Given that the *Tc-emc* segmentation phenotype is limited to the abdomen, this would be a prime candidate for such a function.

6.3.5. Caveats and future experiments

While RNAi is often a robust and effective way of achieving knockdowns in *Tribolium* (Bucher et.al., 2002), knockdown efficiency can vary significantly between different genes (Bitra et.al., 2009), and as demonstrated in this chapter, between different dsRNA sequences targeting the same gene. Competitive inhibition can also occur when multiple dsRNA are injected together, which in the case of double or triple knockdowns can further reduce knockdown efficiencies (Miller et.al., 2012).

Of all the bHLH genes screened by Bitra et.al. (2009), *Tc-dpn* had one of the lowest knockdown efficiencies of around 50% in larval injections, despite using a high dsRNA concentration (5 µg/µL). The exact reason for this is unknown, however, *Tc-cwo* RNAi produced a similar result, and as a highly conserved clock gene that is known to inhibit its own expression in *Drosophila* (Matsumoto et.al., 2007), it could be proposed that knocking down such a gene could inadvertently increase transcription by breaking its autoregulatory feedback loop. If a similar autoregulatory network controls *Tc-dpn* oscillations, then this may also explain the low knockdown efficiency for this gene. Regardless of the reason, a low knockdown efficiency of *Tc-dpn*, along with competitive inhibition in double knockdowns, might explain why no additive effects were detected in any *Tc-dpn* RNAi experiments. Furthermore, consistent with previously published data (Bitra and Pali, 2010), egg laying was marginally lower for *Tc-dpn* injected females compared to controls, which may also have masked any potential function of *Tc-dpn* in embryogenesis. Taken together, RNAi may not be the most suitable method for investigating *Tc-dpn* function in *Tribolium* development, and a transgenic knockout, such as that created for *Tc-odd*, might provide a clearer picture.

While my results from this chapter have identified redundancy between *Tc-h* and *Tc-odd*, the exact nature of this relationship has been hard to establish through cuticle analysis alone. The severity and variation of *Tc-odd*-null mutants makes it difficult to interpret subtle additive effects in RNAi experiments. There are also limitations to using spiracle counts as a proxy measure for segmentation, namely the fact that spiracle development itself could be affected by the knockdown of additional genes. Nevertheless, this was the only segmental feature that could be identified and counted

in severely affected cuticles, and there were no previous signs of spiracle abnormalities in cuticles from *Tc-h* and *Tc-dpn* knockdowns.

To investigate further the impact of *Tc-h* knockdown in *Tc-odd*^{-/-} embryos, *in situ* analyses of other segmentally expressed genes should be compared across controls and *Tc-h* pRNAi germbands. As an important pair-rule gene that still forms poorly resolved stripes in the absence of *Tc-odd*, *Tc-run* will be a primary candidate to investigate. A complete loss of *Tc-run* stripes in *Tc-h* knockdown *odd*^{-/-} germbands would confirm if there is redundancy between *Tc-h* and *Tc-odd* in driving gene oscillations in the *Tribolium* segmentation clock.

7. GENERAL DISCUSSION

In this thesis, I have investigated the evolution of the *Hes/Hey*- and *odd-skipped*-family genes and used *Tribolium castaneum* as a model for researching their potential roles in the arthropod segmentation clock. Through phylogenetic analyses, I have made several important observations with implications for the functional evolution of each family. I have presented a revised understanding of *odd*-family gene expression in *Tribolium*, which while greatly reducing the likelihood of redundancy between *odd* family genes in the segmentation clock gene network, does raise additional questions regarding the roles of *Tc-sob* and *Tc-drm* in the late segmentation network. This left *Hes/Hey* genes as the primary candidates for partially rescuing primary segment patterning in an *odd*-null knockout SB strain. While the expression of *Hes* genes, *Tc-h* and *Tc-dpn*, are consistent with a regulatory relationship with the *Tc-eve/Tc-odd/Tc-run* oscillator, functional knockdowns have once again failed to identify any important roles for these genes in the clock network. However, the expression and knockdown of another bHLH gene, *Tc-emc*, has demonstrated an importance in abdominal segment patterning, and renews the discussion around Notch signalling as a mediator of *Tribolium* segmentation.

In this chapter, I will try to contextualise some of the key findings from this thesis in a broader evolutionary framework while addressing some of the elements of arthropod segmentation that remain poorly understood. Many of the results presented here raise more questions than they answer, and so much of this discussion will revolve around informed hypotheses which need to be tested further.

7.1. *Odd* family evolution and expression: Candidates for the transition to pair-rule patterning in insects?

In this thesis I have explored the evolution of an often-overlooked family of genes and their expression during *Tribolium* development. The arthropod *odd-skipped*-family genes provide us with insights into some of the dynamics that influence gene evolution, and studies of their expression and function demonstrate examples of redundancy,

subfunctionalisation and neofunctionalization across different developmental contexts. Due to the occurrence of regular gene conversion events, it has not been possible to infer accurate relationships between *odd-skipped* family genes between distantly related arthropod species. Therefore, the exact number and timing of *odd*-family gene duplications during arthropod evolution is still unclear. However, in Onychophora, the sister group to arthropods, only one *odd*-skipped homologue could be identified from the transcriptome of *Euperipatoides kanangrensis* (Janssen and Budd, 2013), while the tardigrade, *Ramazzottius varieornatus*, also only possesses one homologue (GAV04140.1), suggesting that the radiation of *odd*-skipped family genes does not predate arthropod evolution. Whether or not this radiation occurred before the divergence of lineages leading to chelicerates, myriapods and pancrustacea, remains unknown.

One thing that is clear from conserved developmental expression patterns in species across the phyla, is that *odd*-family genes have ancestral roles in body segmentation and leg development (Damen et.al., 2005; Prpic and Damen, 2009; Green and Akam, 2013; Eriksson et.al., 2013; Bruce, 2021). This may also be true for other domains of *odd*-family gene expression observed in *Tribolium*, such as in the hindgut anlage or head appendages. However, there is limited expression data outside of insect models for these stages. In keeping with the theme of this thesis, I will discuss what is currently known about *odd*-family expression and function in the context of arthropod segmentation, and along with findings presented in chapters 2 and 3, will propose several scenarios for the involvement of these genes in the evolution and diversification of segment patterning mechanisms.

The recruitment of *odd*-family genes into arthropod segmentation

While expression patterns for *odd*-family genes in arthropods point towards an ancestral role in segment patterning, the exact nature of this ancestral role is unknown. Despite conserved stripes of segment-polarity gene expression, the lack of any segmentation-related expression for the single *odd*-related gene in the onychophoran, *Euperipatoides*, suggests that *odd*-family genes were co-opted into the segmentation

network at the dawn of arthropod evolution (**Fig. 7.1C**) (Eriksson et.al., 2009; Janssen et.al., 2013). In the spider, *Cupiennius*, only 1 *odd*-family gene, *odd-r1*, is expressed in the SAZ during segmentation, while two others, *odd-r2* and *odd-r3* do not exhibit any SAZ expression (Damen et.al., 2005). Interestingly, this expression of *odd-r1* is limited to stripes in the anterior SAZ and segmented germband only, suggesting that it is not involved in primary segment patterning (**Fig. 7.1C**). In contrast, dynamic stripes of posterior SAZ expression are observed for *odd*-family genes in other sequentially segmented species spanning myriapods, crustaceans and insects (**Fig. 7.1C**) (Janssen et.al., 2011; Green and Akam, 2013; Eriksson et.al., 2013; Sarrazin et.al., 2012; Nakao, 2015). In addition, functional studies in insects, such as *Tribolium*, have shown that *Tc-odd* is an essential component of the clock network governing primary segment patterning, while also maintaining a function in the late segmentation network as an upstream regulator of segment polarity genes (**Fig. 7.1A-B**) (Choe et.al., 2006; Choe et.al. 2007; Jeon et.al., 2020). While it has been proposed that *Tc-odd* is ultimately required for positive regulation of *Tc-en* stripes, via *Tc-prd* stripes in odd-numbered parasegments, exactly how this mechanism occurs is unclear (Jeon et.al., 2020). Given that *Tc-odd* itself is strongly expressed in alternate segments to those lost in *Tc-odd*^{RNAi} induced PR-like phenotypes (**Fig. 4.2; Fig. 6.7**), this is contradictory to a direct activator role, and is more consistent with inhibition of genes upstream of odd-numbered *Tc-prd* stripes. Indeed, intermediate knockdowns of *Tc-odd* result in resolved primary stripes of *Tc-eve* expression but unresolved secondary *Tc-eve* stripes, indicating that it is, at least in part, related to its negative regulation of *Tc-eve* in the late segmentation network (Jeon et.al., 2020).

Based on existing data, much of which is limited to expression patterns alone, it could be proposed that, in the common ancestor of arthropods, an *odd*-family gene was at first recruited into a role in the late segmentation network responsible for regulating segment polarity gene expression (**Fig. 7.1**). In the lineage leading to myriapods and Pancrustacea, this *odd*-family gene was then co-opted into the primary segmentation network, where it has subsequently become an integral component of cellular oscillations. This would explain the widespread occurrence of posterior SAZ expression for *odd*-family genes in all lineages except for chelicerates (**Fig. 1.6**). However,

expression patterns for *odd*-related genes in other chelicerates would be necessary to confirm whether *Cupiennius* is truly representative of the clade.

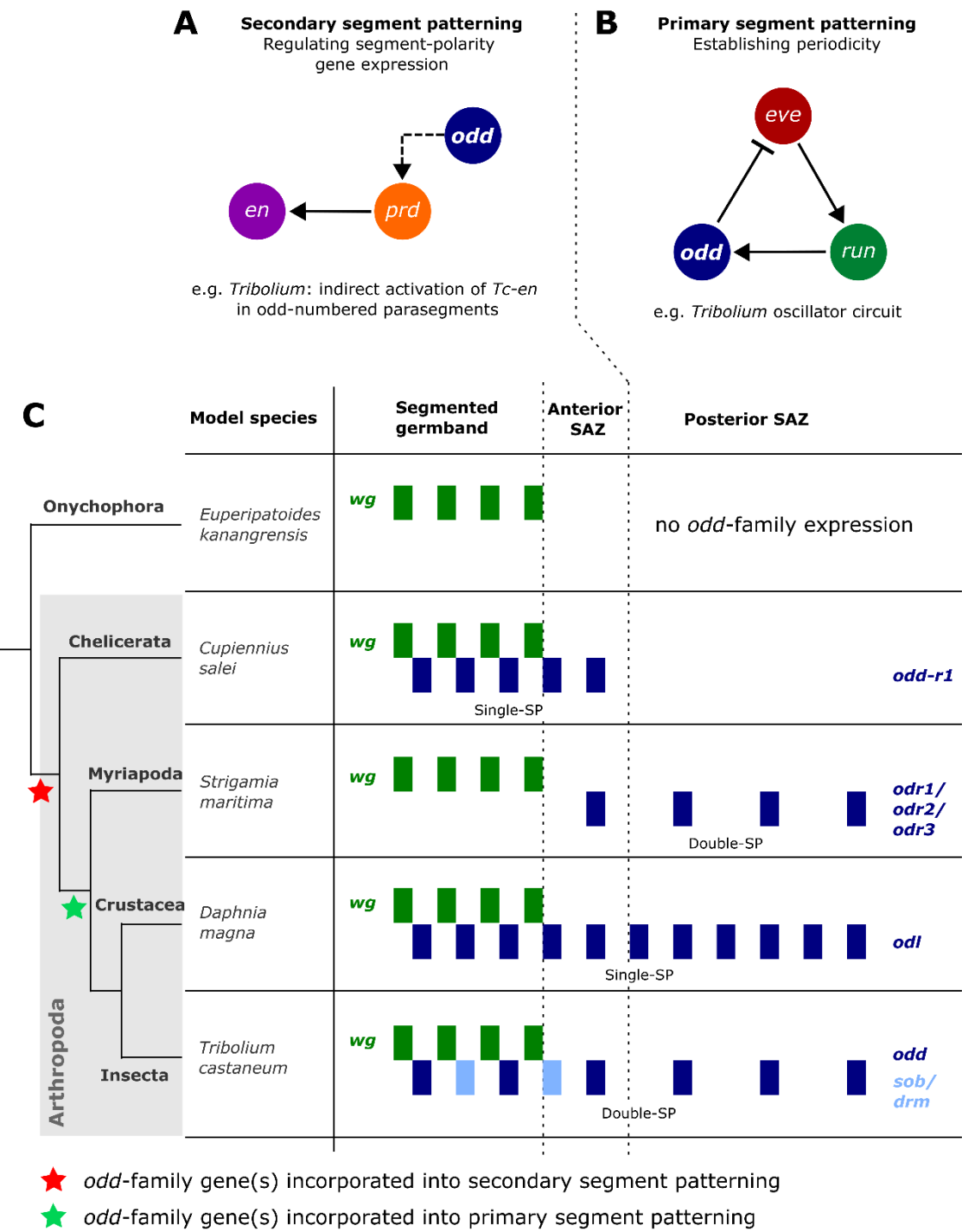


Figure 7.1. The evolution of *odd*-family gene expression in arthropods. (A) Secondary segment patterning: The activation of segment-polarity gene expression at the anterior SAZ. In *Tribolium*, *Tc-odd* stripes are responsible for activating stripes of *Tc-en* via the secondary pair-rule gene, *Tc-prd*. Dashed line indicates indirect activation. Based on results from Jeon et.al. (2020). (B) Primary segment patterning: Generation of periodicity from the posterior SAZ. In *Tribolium*, *Tc-odd* is part of a 3 gene feedback loop driving cellular oscillations within a clock-and-wavefront mechanism (Choe et.al., 2006). (C) Simplified schematics of *odd*-family gene expression across the posterior SAZ, anterior SAZ, and segmented germband of model species from across panarthropods. Stars indicate hypothesised timings of *odd*-family gene(s) being recruited into segment patterning roles, based on expression patterns alone. Note that the number of *odl* stripes in the posterior SAZ of *Daphnia* is to illustrate single-segment periodicity (SP) in contrast to the double-SP of *Strigamia* and *Tribolium*, and does not accurately reflect real observations. Expression patterns based on **Fig. 4.2** and data from Janssen et.al., 2013; Damen et.al., 2005; Chipman and Akam, 2004; Green and Akam, 2013; Eriksson et.al., 2013

Evolution of pair-rule patterning

As previously discussed in chapter 1, pair-rule patterning, the initial expression of segmentation genes in a double-segment periodicity (i.e. 1 stripe per 2 segments), is a feature that appears to have evolved independently in myriapods and insects. In myriapods, this behaviour has only been described for the centipede, *Strigamia*, in relation to *odr1* expression in the posterior SAZ, that seemingly switches from a double- to a single-segment periodicity as it reaches the anterior SAZ (**Fig. 7.1C**) (Chipman et.al., 2004). In the millipede, *Glomeris*, and another centipede species, *Lithobius atkinsoni*, expression patterns for pair-rule gene orthologues display no obvious signs of a double-segment periodicity (Hughes and Kaufman, 2002; Janssen et.al., 2011), suggesting that *Strigamia* pair-rule patterning is independently derived.

The timing of the evolution of pair-rule patterning in insects is not entirely clear. Transitions from single- to double- segment periodicity have been observed for multiple holometabolan models, and the orthopteran, *Gryllus* (Mito et.al., 2007; Nakao, 2010; Rosenberg et.al., 2014). In contrast, expression of pair-rule orthologues in the cockroach, *Periplaneta*, and the milkweed bug, *Oncopeltus*, do not show signs of pair-rule patterning (Chesebro, 2012; Auman and Chipman, 2017; Reding et.al., 2019),

suggesting that there are multiple origins in insects. However, in the case of *Oncopeltus*, the nuclear receptor *E75A*, does exhibit pair-rule expression and function, making the exact nature of pair-rule patterning unclear in this species (Erezyilmaz et.al., 2009). Regardless of when these transitions occurred, the relative expression patterns of *Tc-odd*, *Tc-sob* and *Tc-drm* observed for *Tribolium* (**Fig. 4.2**), may offer an insight into one of the possible mechanisms that enabled segment periodicity to be doubled in holometabolous insects.

In more basal branching (with respect to insects) arthropod species, such as the spider, *Cupiennius*, the millipede, *Glomeris*, and the crustacean, *Daphnia*, *odd*-family orthologues are expressed in a single-segment periodicity in the anterior SAZ (Damen et.al., 2005; Janssen et.al., 2011; Eriksson et.al., 2013). As previously discussed, this expression may reflect an ancestral role for an *odd*-family gene(s) in the late segmentation network regulating segment-polarity stripes, although functional studies are lacking in these species (**Fig. 7.1A**). Given that *Tc-odd* is expressed in opposite phase to its paralogues, *Tc-drm* and *Tc-sob*, the evolution of double segment periodicity observed in *Tribolium* may initially have been made possible by subfunctionalisation of the *odd*-family genes in the anterior SAZ, with *Tc-sob* and *Tc-drm* adopting part of *Tc-odd*'s role in regulating segment-polarity gene expression in alternate segments to those regulated by *Tc-odd*. Consistent with this, anterior SAZ stripes for both *Tc-sob* and *Tc-drm* precede those of *Tc-wg* stripes, suggesting that their activation is upstream of segment-polarity gene expression.

Given that *Tc-odd*, *Tc-sob* and *Tc-drm* are co-expressed in many other areas during *Tribolium* development, they likely share many of the same regulatory inputs, perhaps explained by their conserved genomic arrangement (**Fig. 3.10**). It is therefore possible that at some point during insect evolution, these genes were also co-expressed in the context of segmentation. Indeed, the faint secondary stripes of *Tc-odd* and *Tc-sob/Tc-drm* expression in alternate segments (**Fig. 4.2**) may reflect an ancestral single-segment periodicity that these genes once shared. While expression data is limited for *odd*-family orthologues in other insects, *Of-sob* in *Oncopeltus* is reported to overlap *Of-odd* expression across the SAZ as well as the segmented germband, where they are both expressed in a single-segment periodicity (Auman and Chipman, 2017; Reding et.al.,

2019). Given the phylogenetic position of *Oncopeltus*, this may represent the ancestral pattern of insect *odd*-family gene expression prior to the evolution of pair-rule patterning. However, as previously discussed, it is possible that published *Of-sob* expression is inaccurate, and instead reflects cross-hybridisation with *Of-odd* mRNA, due to the inclusion of a highly similar zinc-finger coding sequence in the *in situ* probes (**Fig. 4.7**). Experiments would need repeating with probes designed from gene-specific sequence to confirm if *Of-odd* and *Of-sob* are indeed co-expressed during segmentation.

More *odd*-family expression data from other insects is also necessary to establish whether there is a connection between the expression of these genes and the evolution of pair-rule patterning. As with *Oncopeltus*, any such study should take careful consideration over their probe designs in order to minimise the chances of cross hybridisation between *odd*-family ZF coding sequences. In *Drosophila*, *odd*, *sob* and *drm* all exhibit pair-rule stripes (Hart et.al., 1996; Green et.al., 2002). However, their co-expression has not been assessed to see if they are overlapping or, like *Tribolium*, are expressed in opposite phase to each other.

The subfunctionalisation of *odd*-family genes only explains how, in the case of *odd-skipped*, a transition of the clock cycle from a single- to double-segment periodicity could still produce the same frequency of segment-polarity stripes, provided anterior SAZ expression of other *odd*-family paralogues and their downstream target genes remain unchanged. This model still leaves many question marks over the exact mechanisms by which pair-rule patterning evolved and assumes that the transition occurred as a result of an altered clock cycle, rather than by a change to the “readout” of the clock (Discussed by Clark et.al., 2019). Furthermore, it does not explain how anterior SAZ stripes of *drm* and *sob* are activated independently of *odd*, which is itself an important consideration. Additional functional studies of *Tc-drm* and *Tc-sob* are necessary to confirm what, if any, roles they play in the late segmentation network in *Tribolium*. Due to their overlapping expression, double knockdowns would be required to rule out any functional redundancy that *Tc-sob* and *Tc-drm* may share.

7.2. *Tc-emc* segmentation phenotype renews the discussion of *Hes* and Notch signalling involvement in the *Tribolium* segmentation clock

One of the most significant changes to the clock and wavefront mechanism underlying sequential segment patterning in arthropods, is the apparent loss of Notch signalling in some insects (**Fig. 1.6**). In arthropods where there is clear functional evidence for Notch-mediated segmentation, such as the spider, *Cupiennius*, the crustacean, *Daphnia*, and the cockroach, *Periplaneta*, oscillatory expression in the SAZ is observed for both *Delta* and *Notch*, genes which encode the ligands and receptors necessary for sending and receiving signals, respectively, between cells (Stollewerk et.al., 2003; Pueyo et.al., 2008; Eriksson et.al., 2013). This expression, along with the loss of synchronised *Hes* expression in the SAZ of embryos where Notch signalling has been inhibited, is consistent with feedback between intracellular *Hes* oscillations and the intercellular Notch signalling pathway. Expression and functional studies of Notch signalling and/or *Hes* genes in several sequentially segmented insects indicates that these roles have been lost in the segmentation clocks of some species, with evidence to suggest this has occurred multiple times during insect evolution (**Fig. 1.6**) (Mito et.al., 2007; Kainz et.al., 2011; Auman et.al., 2017; Reding et.al., 2019).

In *Tribolium*, *Tc-Delta* and *Tc-Notch*, as well as another ligand coding gene, *Tc-Serrate*, and a modulator of receptor-ligand interactions, *Tc-Fringe*, do not exhibit striped expression in the SAZ during segmentation (Aranda et.al., 2008). For this reason, previously described abdominal deformities described by Kux et.al. (2013), as a result of embryonic *Tc-Notch* knockdowns, were not taken as definite evidence for a role of Notch signalling in *Tribolium* segmentation. In this thesis, I have closely replicated these *Tc-Notch*^{RNAi} phenotypes via knockdown of a common effector of Notch signalling, *Tc-emc* (**Fig. 6.2**), providing supporting evidence for Notch-mediated segmentation in *Tribolium*. In addition, *Tc-emc* RNAi-induced labrum (**Fig. 6.3**) and leg deformities (**Fig. 6.4**) also closely resemble those reported from experiments where Notch signalling has been disrupted (Angelini et.al., 2012; Siemanowski et.al., 2015), consistent with *Tc-emc*'s role as a mediator of Notch signalling. Importantly, unlike *Tc-Notch*, *Tc-emc* is expressed in dynamic stripes in the posterior SAZ (**Fig. 5.3**; **Fig. 5.6**), providing a clear

link between the phenotype and a direct role in segment patterning. However, how the oscillatory SAZ expression and apparent Notch-mediated function of *Tc-emc* occurs in the context of abdominal segmentation is unclear, especially given the lack of similarly expressed members of the Notch signalling pathway.

Linking *Tc-emc* expression and function to Notch signalling in *Tribolium*

Tc-Notch is ubiquitously expressed during *Tribolium* segmentation (Aranda et.al., 2008), indicating that all cells, including those in the SAZ where *Tc-emc* is expressed, are primed for activation of the Notch signalling cascade. In the case of canonical Notch signalling, this reaction relies on the Notch receptor making contact with a ligand from a neighbouring cell. This then triggers cleavage and translocation of the Notch intracellular domain (NICD) to the nucleus where it upregulates Notch target genes (**Fig. 7.3A**) (Kovall et.al., 2017). If *Tc-emc* is a downstream target of this process, we would expect to see Notch ligands expressed in similar patterns to *Tc-emc*. This is partly true for *Tc-Delta*, whose mesodermal expression closely resembles that of *Tc-emc* at similar stages (**Fig. 7.3C**), with both genes first appearing in a longitudinal stripe along the blastoderm (**Fig. 5.5**; Aranda, 2006). However, neither *Tc-Delta* nor *Tc-Serrate* are expressed in the SAZ during *Tribolium* development, suggesting that another pathway is regulating *Tc-emc* stripes here (Aranda et.al., 2008). Expression for a third, lesser-known Notch ligand-coding gene, *Tc-weary* (*Tc-wry*), has not yet been investigated. The *Drosophila* orthologue, *wry*, was first identified as a ligand of the Notch signalling pathway in the context of adult heart function (Kim et.al., 2010), although roles in development have not been studied. *Tc-wry* would be a primary candidate for further investigating the possible link between ligand-dependent Notch signalling and *Tc-emc* function with regards to the *Tribolium* segmentation clock. At this point in time, no other ligands of the Notch signalling pathway are known in arthropods, and *Tribolium* does not appear to possess any duplicates of those previously described.

An alternative explanation for the link between *Tc-Notch* and *Tc-emc*, is that *Tc-emc* is activated by the noncanonical Notch signalling pathway which does not require activation by ligands (**Fig. 7.3B**) (reviewed by Zhou et.al., 2022). Given that *Tc-emc*,

along with *Tc-Notch*, appear to be ubiquitously expressed in the *Tribolium* germband (**Fig. 5.3**; Aranda et.al., 2008), it is possible that ligand-independent activation of Notch signalling could be required to maintain a background level of *Tc-emc* expression in all cells during development (**Fig. 7.3C**). This would be consistent with the similarities observed between *Tc-Notch* and *Tc-emc* RNAi phenotypes (Kux et.al., 2013; **Fig. 6.2**). Even the loss of base level *Tc-emc* expression as the result of *Tc-Notch* knockdown could be sufficient to induce a segmentation phenotype. In this scenario, another regulatory input would still be required to explain the elevated stripes of *Tc-emc* expression observed in the *Tribolium* SAZ. Furthermore, this would still leave the question of how cellular oscillations are co-ordinated across cells in the absence of ligand-dependent Notch signalling.

In any case, these hypotheses are solely based upon phenotype comparisons and known regulatory interactions from *Drosophila*. Analysis of *Tc-emc* expression in *Tc-Delta/Tc-Notch* knockdowns would first be necessary to confirm if *Tc-emc* is indeed a downstream target of Notch signalling in *Tribolium*. Regardless of if, and how, *Tc-emc* is connected to Notch signalling, it is still clear that it has an important role in abdominal segment patterning. With what we know about the mechanisms by which *emc* regulates bHLH factor activity in *Drosophila*, it is plausible that bHLH genes are also active in abdominal segment patterning. This leaves open the possibility of the co-expressed *Hes* genes, *Tc-h* and *Tc-dpn*, as candidates for this role, despite the lack of functional data to support this. As previously discussed, RNAi efficiency for these genes may not reach the necessary threshold to induce clear segmentation phenotypes, especially when attempting to perform joint knockdowns testing for redundancy. In contrast, *Tc-emc* knockdowns have the potential to severely impact the protein dynamics of both *Tc-H* and *Tc-Dpn* and thus produce a different output. Other techniques, such as the generation of gene knockout strains, may be required to rule out possible functions of these genes in the *Tribolium* segmentation clock. Beyond this, other bHLH genes that are identified for expression in the SAZ may also provide good candidates for further investigations into the mechanism of *Tc-emc* function in *Tribolium* segmentation.

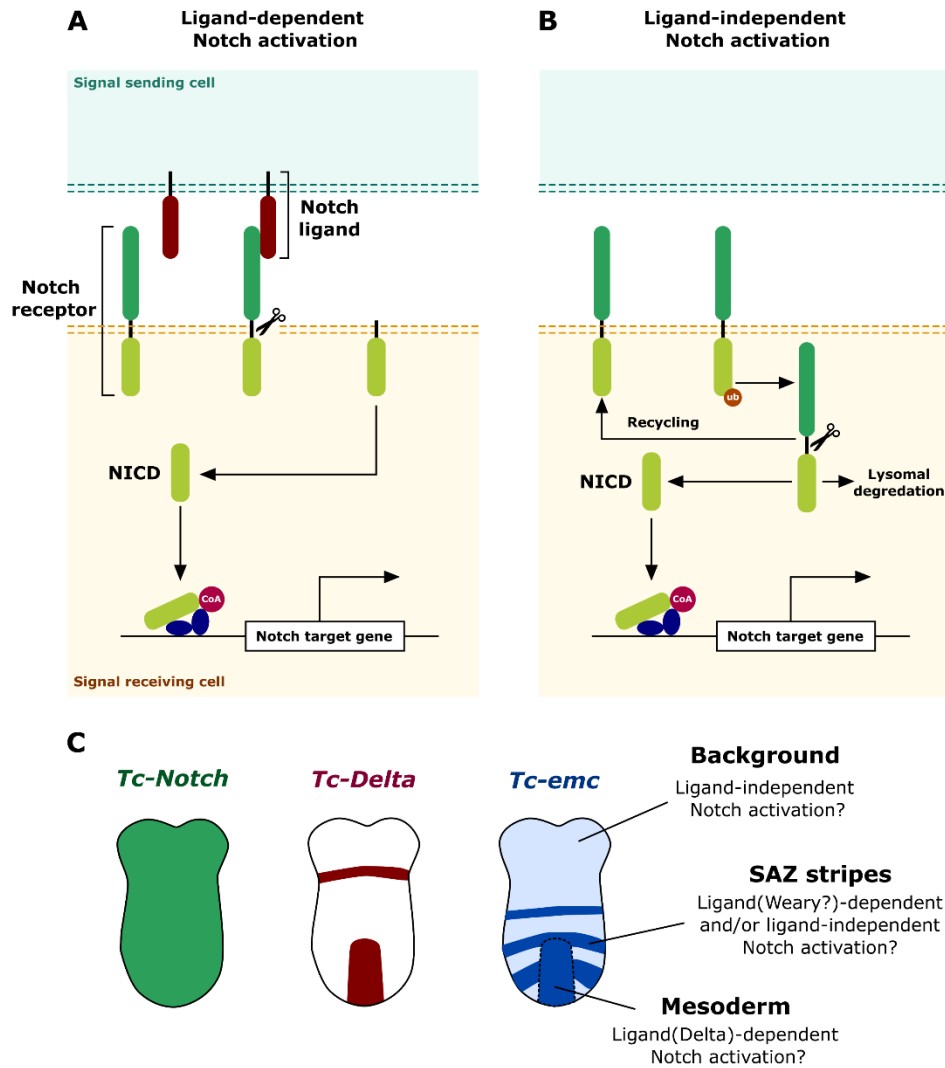


Figure 7.2. Linking Notch signalling to *Tc-emc* expression and function. (A-B) Schematics of ligand-dependent (canonical) Notch signalling and ligand-independent (non-canonical) Notch signalling. (A) relies on signalling from a ligand in a neighbouring cell, which once bound to the Notch receptor, triggers the cleavage and transportation of the Notch intracellular domain (NICD) to the nucleus where it binds to a co-activator (CoA) to promote target gene expression. (B) ubiquitin ligases (ub) help transport some Notch receptors back into the cell where they are either recycled back to the cell membrane, sent for lysosomal degradation, or cleaved. In the latter case, the cleaved NICD can activate target Notch genes as in (A). (C) Expression patterns for *Tc-Notch* (Aranda et.al., 2008), *Tc-Delta* (Aranda, 2006) and *Tc-emc* (**Fig. 5.3**) in an early *Tribolium* germband. Ubiquitous *Tc-Notch* expression could maintain background levels of *Tc-emc* expression through ligand-independent signalling (B), while mesodermal *Tc-Delta* expression could explain elevated *Tc-emc* expression through a ligand-dependent mechanism (A). Regulation of *Tc-emc* stripes in the SAZ is unclear without expression patterns for a remaining Notch ligand, *Tc-Weary*. *Tc-emc* expression may also be activated independently of Notch signalling entirely. Notch signalling pathways adapted from figure by Zhou et.al. (2022)

7.3. Shifts in patterning mechanisms over the course of *Tribolium* segment addition

As previously mentioned, pre-gnathal segments in arthropods are differentially patterned to the remaining posterior segments which have been the focus of this thesis (Chipman and Edgecombe, 2019). In addition to this, further regional changes to patterning mechanisms are observed over the course of segment addition in many species. For example, prosomal and opisthosomal segmentation in the spider, *Parasteatoda*, appears to be regulated by two distinct processes (Schönauer et.al., 2016), while in some insects, such as the milkweed bug, *Oncopeltus*, mixed modes of simultaneous and sequential segment patterning are observed for anterior and posterior trunk segments respectively (Auman and Chipman, 2018). Indeed, an incremental anterior-to-posterior adoption of simultaneous segmentation, through the evolution of stripe specific enhancers, is believed to be how complete simultaneous segmentation may have evolved (Clark et.al., 2019). While all *Tribolium* trunk segments are patterned sequentially during GBE, I have identified gene expression patterns and phenotypes that highlight dynamic shifts in the regulatory landscape of *Tribolium* across different stages of segmentation (**Fig. 7.3**).

Gnathal and thoracic segmentation

The early establishment of a static mandibular stripe of *Tc-h* expression in early blastoderm stages (**Fig. 5.5**), along with the consistent loss of mandibles from cuticles in which *Tc-h* has been knocked down, confirm previous studies linking *Tc-h* to a unique mechanism patterning the mandibular segment in *Tribolium* (Aranda et.al., 2008). A similar stripe of *Tc-Delta* expression at this stage also indicates that this may occur downstream of Notch signalling, especially given the common association with *Hes* genes as Notch effectors (Aranda, 2006). Consistent with this, *Tc-Delta* knockdowns were also reported to result in malformed or absent gnathal appendages (Aranda, 2006). The more severe head phenotypes induced by *Tc-h* RNAi experiments, along with expression patterns for the segment-polarity gene, *Tc-gooseberry* (*Tc-gsb*), in *Tc-h^{RNAi}* embryos, indicate that *Tc-h* function extends to a pair-rule-like role in the gnathal region, with mandibular and labial segments being most sensitive (Aranda et.al., 2008).

Results obtained from this thesis suggest that *Tc-h* has a redundant function in patterning thoracic segments also, something that is apparent from joint knockdowns with *Tc-odd* (**Fig. 6.7**), and that was confirmed through *Tc-h* knockdowns in *Tc-odd*-null mutants (**Fig. 6.9**). The complete loss of visible legs and reduced spiracle counts induced by *Tc-h* RNAi in a high proportion of *Tc-odd*-null cuticles, suggests that there is often a complete failure to pattern thoracic segments in the absence of both *Tc-h* and *Tc-odd*. Without knowledge of gene expression patterns in these embryos, it is not clear exactly how this occurs, whether from a breakdown of cellular oscillations in the SAZ, disruption to the late segmentation network, or from roles of *Tc-h* further downstream in development. While it does appear that additive effects of *Tc-h* loss of function in the absence of *Tc-odd* are limited to thoracic segmentation only, asegmental phenotypes, in which segmentation has failed completely, were also observed. It is therefore possible that these severe phenotypes could reflect a complete collapse of primary segment patterning, as would be expected if *Tc-h* was responsible for rescuing this process in *Tc-odd*-null embryos. However, the prevalence of these phenotypes was not significantly different to the *GFP* dsRNA injected group, meaning we cannot confidently attribute this to *Tc-h* loss of function. Further investigation would be necessary to confirm the mechanism by which *Tc-h* knockdown leads to the loss of thoracic segments in *Tc-odd*^{-/-} embryos, as well as to conclude whether thoracic and abdominal segment patterning is indeed differentially regulated by *Tc-h*.

Abdominal segmentation

One of the primary findings of this thesis is that of *Tc-emc* expression and function in relation to *Tribolium* segmentation. While the dynamic expression of *Tc-emc* in the SAZ begins at the onset of segment addition, its knockdown appears only to affect abdominal segmentation. The reason why the role of *Tc-emc* in segmentation is limited to abdominal segment patterning is unclear. One possibility is that this is related to changes in posterior signalling that are observed around the wg6-wg7 stage of GBE. This stage corresponds to the transition from thoracic to abdominal segmentation and the appearance of new posterior domains are observed for a number of different genes. Experiments from this thesis confirmed posterior expression previously reported at this

stage for *Tc-E(spl)1* and *Tc-drm* (Kux et.al., 2013; Thümecke and Schröder, 2021), while also providing similar novel descriptions for *Tc-dpn*, *Tc-Sidpn1* and *Tc-sob* expression (**Fig.5.2; Fig. 5.4; Fig. 4.2**). In the case of *Tc-drm*, functional evidence suggests that this posterior domain is related to a role in hindgut development (Thümecke and Schröder, 2021), while knockdowns of *Tc-E(spl)1* do not reveal any functions related to posterior expression (Kux et.al., 2013). *Tc-dpn* RNAi from this thesis has failed to identify any function for this gene also, and *Tc-sob* knockdowns have not yet been performed in *Tribolium*. We have, however, found evidence to suggest that the posterior domain of *Tc-Sidpn1* is related to a function in germband elongation.

Tc-Sidpn1^{RNAi} induced truncation phenotypes from around the A2/A3 segment closely correlate to the timing of *Tc-Sidpn1* expression behind the posterior *Tc-wg* domain at stage wg6/wg7. Interestingly, posterior expression of *Tc-Delta* also switches on in this region at a similar time during GBE and embryonic knockdowns produce similar truncation phenotypes (Aranda, 2006), suggesting that *Tc-Sidpn1* may be acting downstream of Notch signalling. This differential regulation of posterior development by Notch signalling in later stages of development is a phenomenon observed in other arthropods from across the phyla. The exact point that this switch occurs varies between species, but *Notch/Delta* RNAi and/or DAPT inhibition of Notch signalling commonly result in truncation phenotypes consistent with posterior signalling in sequentially segmented arthropods (Reviewed by Williams and Nagy, 2017).

In the cockroach, *Periplaneta*, as well as mediating primary segment patterning, the Notch signalling pathway is also responsible for SAZ maintenance via the regulation of posterior *Pa-Wnt1* and *Pa-cad* expression (Chesebro et.al., 2013). As appears to be the case in *Tribolium*, this role is limited to later stages of GBE. Similar dynamics between posterior Notch signalling with the Wnt pathway and *Pt-cad* are observed during opisthosomal segmentation in the spider, *Parasteatoda*, suggesting this is an ancestral mechanism of posterior patterning in arthropods (Schönauer et.al., 2016). In *Parasteatoda*, this posterior signalling network is necessary for the proper activation of pair-rule gene expression in the SAZ. While similar mechanisms of posterior Notch/Delta signalling have been identified in distantly related arthropods, it does not appear to be conserved among all species. In particular, it has been lost from insects

that employ simultaneous modes of segmentation, such as *Drosophila* and the honeybee, *Apis* (Wilson et.al., 2010; Williams and Nagy, 2017). This is perhaps unsurprising given that there is no requirement for SAZ maintenance and that elongation occurs after segment patterning. In the milkweed bug, which employs a mixed mode of segment patterning, *Of-Delta* knockdowns also cause truncations, however, given that the expression of segmentation genes are unaffected in the blastoderm, the authors attribute this to a late acting role of *Of-Delta* in segment development rather than in early patterning (Stahi and Chipman, 2016). Furthermore, *Of-Delta* does not display a posterior expression domain in *Oncopeltus* during its sequential phase of segmentation, suggesting the Notch/Delta pathway is not required for posterior signalling in this species (Auman and Chipman, 2018). To test if *Tc-Delta* and *Tc-Sidpn1* RNAi mediated phenotypes in *Tribolium* are the result of a similar interactions to those observed in *Parasteatoda* and *Periplaneta*, it would be interesting to assess *Tc-cad* and *Tc-wg* expression in knockdown embryos. *Tc-cad* and SAZ expressed genes, *Tc-eve*, *Tc-odd* and *Tc-h*, have already been shown to be downstream targets of posterior Wnt signalling (Oberhofer et.al., 2014), therefore any connection between this pathway and *Tc-Delta/Tc-Sidpn1* would demonstrate roles for these genes in posterior signalling and explain the truncation phenotypes induced by their loss of function.

Regardless, it is possible that the abdominal-specific segmentation phenotypes captured by *Tc-emc* loss of function experiments are somehow related to this switch in posterior signalling. As previously mentioned, *Tc-emc* is likely to function via interactions with a bHLH factor (Campuzano, 2001), therefore any bHLH genes that are switched on during this stage are likely to be key candidates. Alternatively, the lack of anterior segmentation phenotypes observed for *Tc-emc* RNAi cuticles could be the result of low knockdown efficiency from embryonic injections. The phenotypic series from weak to strong severity generally follows a posterior-to-anterior progression, respectively, in relation to where abnormalities manifest. By following this logic, thoracic segments would be the last to be affected by *Tc-emc* loss of function, and might therefore require a stronger knockdown. Unfortunately, as *Tc-emc* is essential for oogenesis (Bitra and Pali, 2010), it is not possible to generate a knockout strain to test

this. Further investigation of expression patterns for segmentation genes in *Tc-emc* loss of function embryos should be carried out to elucidate its exact role in segment patterning.

In addition to the differential regulation of gnathal, thoracic, and abdominal segment patterning, there are several indications that the posterior-most abdominal segments (A8-A10) are also patterned by a separate mechanism. The segmental stripes of a segment-polarity gene, *Tc-H15*, appear sequentially in pairs throughout segment addition, except for those corresponding to segments A8, A9 and A10, which appear one at a time (Janssen, 2014). This suggests that these posterior-most stripes are patterned by a separate mechanism to earlier stripes. Interestingly, the timing of this change corresponds to the disappearance of *Tc-h*, *Tc-dpn* and *Tc-emc* expression from the SAZ in later stage germbands (**Fig. 5.1; Fig. 5.2; Fig. 5.3**). While previously published expression patterns for *Tc-eve* and *Tc-run* have not been described at this late stage in *Tribolium*, HCR FISH for *Tc-odd* shows that stripes of SAZ expression persist until the completion of segment addition (**Fig. S?**). The paired activation of *Tc-H15* in segments anterior to A8 could be explained by parallel inputs from the primary oscillator circuit (*Tc-odd/Tc-eve/Tc-run*) alongside *Hes* genes (*Tc-h/Tc-dpn*), while the single-segmental activation of A8, A9 and A10 stripes are consistent with the loss of input from *Hes* genes at this stage. However, the mechanisms by which stripes of *Tc-H15* are regulated in *Tribolium* is currently unknown. *Tc-H15* is also the only segment polarity gene reported to appear in a double segmental fashion during segmentation, with *Tc-Hedgehog* (*Tc-Hh*), *Tc-wg* and *Tc-en* stripes all appearing one at a time throughout the entire course of segmentation (Nagy and Carroll, 1994; Brown et.al., 1994; Farzana and Brown, 2008). Exactly how anterior and posterior stripes of *Tc-H15* are differentially regulated, and what, if any, significance it has for the wider segmentation network, remains unclear. However, it further highlights the regional variation in segment patterning mechanisms throughout *Tribolium* segmentation.

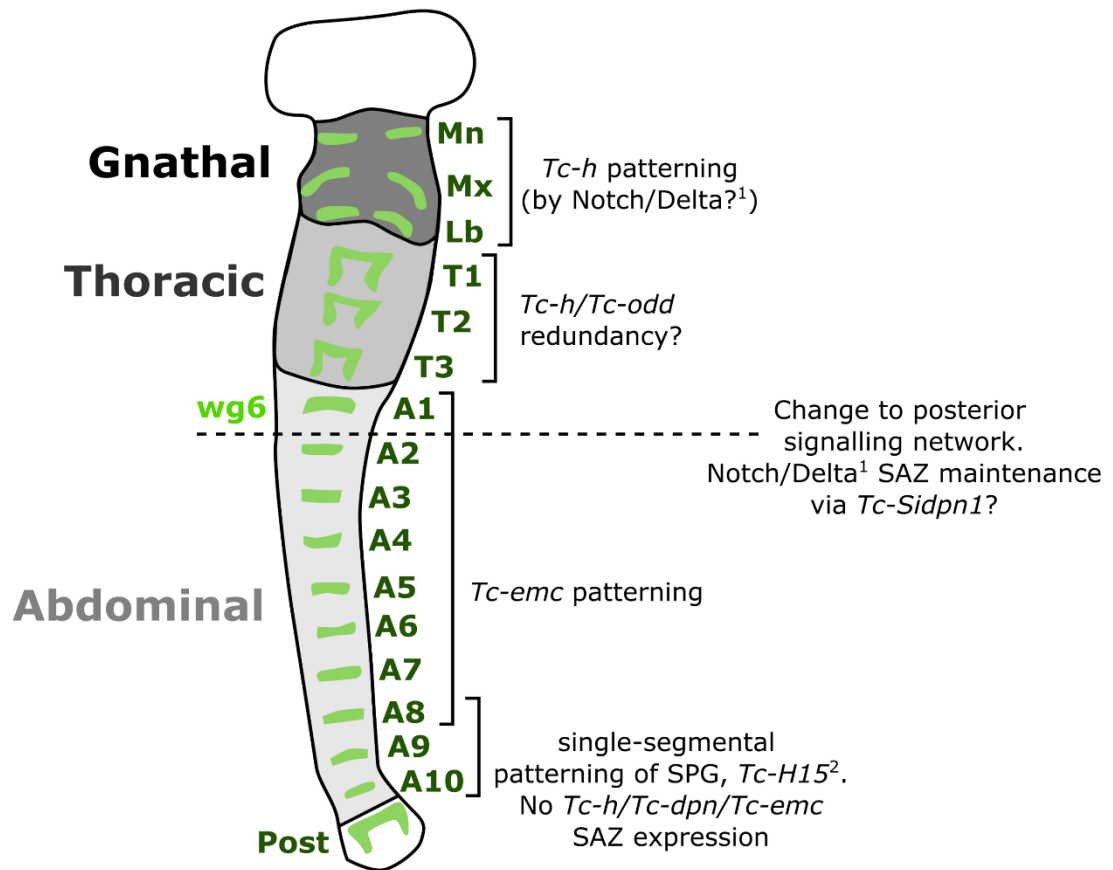


Figure 7.3. Differential patterning across *Tribolium* segmentation. Diagram of a *Tribolium* embryo at the end of segment addition with *Tc-wg* stripes in green and labelled with their corresponding prospective segment. The different patterning mechanisms discussed in the main text are summarised for each region. References: 1, Aranda et.al., 2008; 2, Janssen, 2014

7.4. General remarks and future directions

The primary aims for this thesis were to investigate the underlying complexity of the *Tribolium* segmentation clock, driven largely by recent observations from our lab of primary segment patterning and striped expression of *Tc-run* still being established in a *Tc-odd* knockout strain. This challenged previously published data and models built around the *Tc-eve/Tc-odd/Tc-run* circuit as the sole network driving cellular oscillations. While the results presented and discussed here have not explicitly uncovered the molecular interactions that govern primary segment patterning in *Tribolium*, they have opened up several new lines of research for future investigations.

Analysis of arthropod *odd*-family genes have revealed characteristics that make them prime candidates for redundancy and developmental system drift. However, in the case of insects, these genes are often overlooked, with most segmentation studies in insects solely focusing on direct orthologues of *odd-skipped*. It is important that any future research investigating roles for these genes carefully considers the exclusion of the ZF coding region from probe and dsRNA design, in order to avoid cross-hybridisation and off-target effects, respectively. Revised expression patterns for *Tc-sob*, *Tc-bwl* and *Tc-drm* described for *Tribolium* in this thesis have significantly reduced the likelihood of redundancy between these genes with regards to primary segment patterning. However, anterior SAZ and segmented germband expression of *Tc-sob* and *Tc-drm* stripes prompts the need for investigating potential roles for these genes in the late segmentation network and the regulation of segment polarity genes. It is possible that such studies in *Tribolium* and other insects may offer novel insights into the evolution of pair-rule patterning.

Phylogenetic analyses of arthropod *Hes/Hey* genes have characterised previously identified segmentation genes from *Daphnia* (Eriksson et.al., 2013) as *E(spl)* genes, rather than orthologues of the *hairy/deadpan*-related genes that are commonly associated with Notch-mediated segmentation in other species (Stollewerk et.al., 2005; Pueyo et.al., 2008; Green and Akam, 2013). This highlights the interchangeability of *Hes* genes in evolution, and future studies investigating their expression and function in relation to segmentation should therefore consider the wider gene family in their analyses. While this approach has failed to identify additional *Hes/Hey* genes expressed in patterns consistent with primary segmentation in *Tribolium*, expression and functional data for a related bHLH gene, *Tc-emc*, reveal a new candidate for arthropod segmentation. The precise mechanism by which *Tc-emc* regulates segment patterning has not been identified in this case. In order to better understand its general function, expression patterns for segmentation genes should be analysed in *Tc-emc*^{RNAi} embryos. Based on what we know of *Drosophila emc*, possible interactions with co-expressed bHLH factors should also be investigated. *Tc-h* and *Tc-dpn* still remain candidates for these interactions, however, different approaches, such as gene knockouts or inducing ectopic expression, should be employed to fully assess their involvement. Beyond this,

single cell transcriptomics from the SAZ could be used to identify any other co-expressed candidate bHLH genes, while co-immunoprecipitation experiments can inform us of *Tc-emc*'s ability to dimerize said candidates.

Upstream of *Tc-emc*, and of primary segment patterning in general, there is evidence to suggest that Notch signalling is still involved in *Tribolium* segmentation. Whether this function extends to local synchrony of cellular oscillations in the clock network remains to be seen. Investigation of the Notch ligand, *Tc-weary*, is necessary to rule-out or confirm this. If this is ruled-out, the need to identify an alternate mechanism coordinating cellular oscillations in *Tribolium* is essential for better understanding the segmentation clock and the transition away from Notch signalling in insects. Truncation phenotypes from *Tc-Delta* (Aranda, 2006) and *Tc-Sidpn1* (this thesis) point to a conserved role for Notch signalling in posterior development upstream of segment patterning. To establish whether the Notch/Delta pathway functions in a similar manner to other arthropods (Chesebro et.al., 2013; Schönauer et.al., 2016), posterior domains of *Tc-wg* and *Tc-cad* should be analysed in *Tc-Delta/Tc-Sidpn1* RNAi embryos.

Much of the research in *Tribolium*, including that conducted in this thesis, focuses on the well-established techniques of *in situ* hybridisation and RNAi for determining gene functions. However, this can only tell us so much about regulatory relationships between different genes and does not allow distinctions to be made between direct and indirect interactions. Indeed, even the topology of the primary *Tc-eve/Tc-odd/Tc-run* network is still debated (Choe et.al., 2006; Clark et.al., 2019). In order to build accurate models of gene regulatory networks in *Tribolium* segmentation, It would be beneficial to analyse cis-regulatory sequences of various segmentation genes and the activators/repressors that bind them. Techniques such as ChIPSeq can help identify relevant protein-DNA interactions (Park, 2009), while yeast one-hybrid assays can measure the direct regulatory relationships between genes of interest. The latter of these approaches has been demonstrated in research of the vertebrate segmentation clock, enabling the formulation of accurate network models that replicate experimental observations (Schröter et.al., 2012).

The ongoing research in *Tribolium* and other arthropods continues to highlight the conserved and divergent aspects of patterning mechanisms across the phylum. In

particular, our understanding of segment patterning across a diverse group of insect models, covering Blattodea, Orthoptera, Hemiptera, and a number of Holometabolan orders, points to several significant evolutionary transitions having occurred multiple times, including the loss of Notch-mediated segmentation (both in terms of posterior signalling and synchrony of cellular oscillations), the evolution of pair-rule patterning, and the evolution of simultaneous segment patterning. The independent occurrences of these changes demonstrate a developmental system that is highly plastic and prone to drift. A deeper understanding of the precise mechanisms underpinning sequential segmentation in *Tribolium*, and other insects will be essential for learning exactly how these changes have occurred.

APPENDICES

Appendix A - Multiple sequence alignments used to generate residue conservation values and consensus logos used in Chapter 3

	10	20	30	40
<i>Dmel_NP_523977.2</i>	F K A G F A D C V N E V S R F	- - - - - P G I E P A Q R R R L L Q H L S N C I N		
<i>Dmel_NP_476923.1</i>	F K T G F V E C A E E V N R Y V S Q M	- - D G I D T G V R Q R L S A H L N Q C A N		
<i>Dmel_NP_523599.1</i>	Y R A G Y T D C A R E V A R Y L A T P E P P P M G P G S K A R L L R H L D Q C I A			
<i>Bmor_XP_004927126</i>	F K S G F G D C A E V R R Y L S R L	- - A S V P T G L R H R L G N H L T S C I S		
<i>Bmor_XP_004932259</i>	F R A G Y R Q C V T E V A K F	- - - - - Q G M E P A L R R R L V R H L E T C V N		
<i>Tcas_XP_971935</i>	F R A G F S E C A S E V G R F	- - - - - P G L D P V V K R R L L Q H L A S C L N		
<i>Tcas_XP_967694.1</i>	F K S G F N E C A N E I D R F V S Q S	- - - E V D D G L K D R M R S H L Q K C I N		
<i>Tcas_EFA03161.2</i>	Y K A G Y A D C V R E V Q R Y L D T P D A Q V V D A G V R Q R L L R H L D N C V A			
<i>Amel_XP_393948.3</i>	F R A G F T E C A G E V G R F	- - - - - P G L D A S V K R R L M A H L A S C L G		
<i>Amel_XP_006559646.1</i>	Y K A G Y Q D C S R E V S R Y L D A P D I I P M D P A V K R R L L Q H L D S C V S			
<i>Amel_XP_001120814.2</i>	F R S G F S E C A T E V S R Y V S H L	- - E N V D P V V K Q R L V S H L N N C V S		
<i>Nvit_NP_001267498</i>	F R A G F T E C A N E V G R F	- - - - - P G L E A S M K R R L L A H L S T C L S		
<i>Nvit_XP_031779123.1</i>	Y K A G Y Q D C S R E V S R Y L D A P D I I P L E P A F K Q R L L R H L D S C V S			
<i>Nvit_XP_016840180</i>	F R S G F S E C A S E V S R Y V N H L	- - E N V D P T V K Q R L V S H L N N C V T		
<i>Apis_XP_001949685.1</i>	F K A G F T E C T N E V N R F	- - - - - P G L E P H V R K R L M Q H L N T Y F N		
<i>Apis_XP_001945055.2</i>	Y R A G F Q E C A R E V T R F L L S A G S T V I E P A V K Q R L L R H L D T C S S			
<i>Pame_ACI32659.1</i>	F R A G F T E C A G E V G R F	- - - - - P G L E S P V R K R L L Q H L A N C L N		
<i>Dmag_XP_032794114.1</i>	Y R L G Y Q A C C H D I G R F L D G	- - - P L H E L T K E R L M E H L L E K K Q		
<i>Smar_SMAR001692</i>	F S A G Y G E C A N E V T R Y E G G E	- - G H V D A T A R T H L I N H L S R C M N		
<i>Smar_SMAR012958</i>	Y R A G F G E C A T E V S R Y V S R L	- - D T I D P H M R T G L L N H L S T C L T		
<i>Smar_SMAR012956</i>	F R S G F T E C A S E V S R Y V N R I	- - - E V D S G V K Q R L V A H L A N C V S		
<i>Smar_SMAR012957</i>	Y R A G F G E C V A E V S R F L S G S	- - P A V D T A V Q Q R L I S H L N N N F S		
<i>Smar_SMAR005751</i>	Y L A G F S E C L H Q V S L C M D Q L	- - D G L N K D D R M H L L H H L T K C L S		
<i>Ptep_NP_001310755.1</i>	F R A G F A E C A N E V N R F M G R I	- - D G V D H T I R Q R L M N H L A N C L T		
<i>Ptep_XP_015913836.1</i>	F K A G F N E C A T E V T R Y V N R I	- - E G V D P G L R Q R L V S H L S N C L T		
<i>Ptep_XP_015925274.1</i>	F E I G Y T E C A R E V D L F L S S L V D S T V N S E L H R R L T Q H L S R R L H			
<i>Ptep_XP_015920993.1</i>	Y M A G F S Y C M K E V R H F I E	- - - S N I D D D S K S R V L H Y - L N C F A		
<i>Csal_CAB89491.1</i>	F R A G F A E C A N E V N R F M G S M	- - E G I D H T I R Q R L L N H L A N C L T		
<i>Apis_XP_016661802.1</i>	F Q T G Y Q E C A G E V N R F V D R L	- - D G V D E D I K K R L M S H L D S C V A		
<i>Scan_XP_049779350.1</i>	F R A G F S E C A G E V G R F	- - - - - P G L E A N V R R R L L S H L A S C L N		
<i>Scan_XP_049777301.1</i>	F R T G F S E C A G E V G R Y L G R L	- - S S V E P A V R Q R L S G H L S Q C I A		
<i>Scan_XP_049769061.1</i>	F R A G F Q E C A R E V G R F L E S P	- - - E V V A S L K Q R L L R H L E S C V A		
<i>lele_XP_046390237.1</i>	Y K A G F S E C A A E V S R F N S A S S S P S M D Q H F K R R L L S H L T H C L N			
<i>lele_XP_046390055.1</i>	F R T G F N E C A G E V S R Y I G R I	- - E G V D P A V R S R L L S H L S S C V S		
<i>lele_XP_046396516.1</i>	Y R A G F S E C A R E V A R F V Q R P A R P G Q D S S L K H R L L G H L N A C V A			

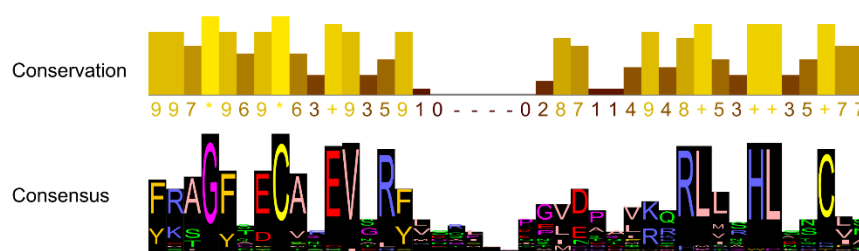


Figure A.1. Amino acid multiple sequence alignment of Orange domains from arthropod non-E(spl) HES proteins. HES proteins where Orange domains could not be identified through NCBI conserved domain search were not included in the analysis. Conservation and Consensus logos were used in **Fig. 3.4** to compare against alignments with non-E(spl) HES proteins missing Orange domains. Species identifiers as described in Chapter 3. Protein accession numbers from NCBI GenBank or Ensembl



Figure A.2. Amino acid multiple sequence alignment of zinc finger 5 from arthropod Odd family proteins. All arthropod Odd family proteins with 5 ZFs, except for any insect Odd-skipped orthologues, were used in this MSA to measure residue conservation across the 5th ZF. Conservation and Consensus logos were used in **Fig. 3.12** to compare against the 5th ZF domain of some insect Odd-skipped proteins. Species identifiers as described in Chapter 3. Protein accession numbers from NCBI GenBank or Ensembl

Appendix B — Comparison of *Tc-h* and *Tc-odd* expression across different germband stages



Figure B.1. Expression dynamics of *Tc-h* and *Tc-odd* are consistent in the SAZ throughout GBE. HCR FISH with *Tc-h* (green) and *Tc-odd* (magenta) at 3 stages of GBE. Single channel images for each embryo are shown alongside merged images. Green and magenta lines indicate the extent of each stripe in the SAZ for *Tc-h* and *Tc-odd* respectively. Domains in the posterior SAZ tend to be broader and more overlapping. Once stripes have resolved, *Tc-h* and *Tc-odd* stripes are in almost opposite phase to each other. A small overlap between the posterior edge of the *Tc-odd* stripe with the anterior edge of the *Tc-h* stripe is seen in each case until expression reaches the anterior SAZ/segmented germband. Scale bar 100 μ m

Appendix C — Late stage *Tc-h*, *Tc-dpn* and *Tc-emc* expression in the hindgut/malpighian tubule

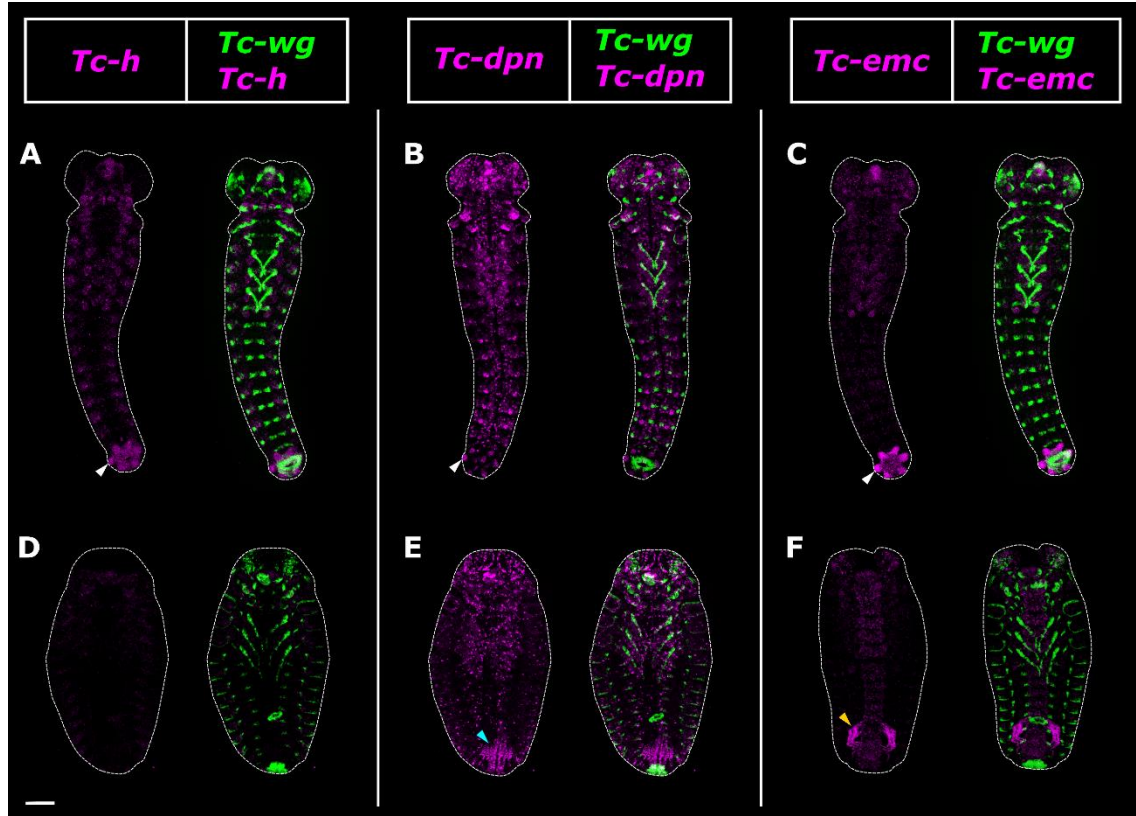


Figure C.1. *Tc-h*, *Tc-dpn* and *Tc-emc* all exhibit expression consistent with roles in hindgut and malpighian tubule specification. HCR FISH for *Tc-h*, *Tc-dpn* and *Tc-emc* with a *Tc-wg* marker. (A-C) Comparative stages of gut development, where each gene is expressed in malpighian tubule (MpT) buds (white arrowhead) after evagination. (D) *Tc-h* expression disappears in later stage embryos. (E) *Tc-dpn* is expressed in squiggly lines that most likely correspond to cryptonephridial tubules surrounding the rectum at this stage (cyan arrowhead). (F) *Tc-emc* is expressed in the free tubules (yellow arrowhead) that do not associate with the rectum, and become freely suspended in the haemolymph. Scale bar: 100µm. Interpretations based off description of MpT development and the expression of a key MpT marker, *Tc-cut*, by King and Denholm, 2014

Appendix D – RNAi results tables for Chapter 6

Table D.1. *Tc-h* parental RNAi results. Percentage values correspond to the number given in the first row of each box. These being eggs examined, cuticles scored, and intermediate head phenotypes. Miscellaneous abnormalities refer to any defects that are considered too inconsistent to be of any significance. Breakdown of head appendages and region affected is not given for the control and *Tc-h* 5' group due to the negligible number of intermediate head phenotypes.

	ctrl	<i>h</i> 5'	<i>h</i> 3'
Eggs examined (n)	264	545	379
Wildtype cuticles (%)	81.1	77.5	21.9
Abnormal cuticles (%)	5.3	2.6	61.7
Empty eggs (%)	13.6	20.3	16.4
Cuticles scored (n)	228	435	317
Wildtype (%)	93.9	96.8	26.2
Abnormal (%)	6.1	3.2	73.8
Head defect (%)	0.9	0.5	71.6
Weak - missing/deformed Mn	0	0	30.9
Intermediate - Additional head appendages missing/deformed	0	0.5	19.2
Strong - Missing entire head	0	0	11.4
Strong - Missing head + T1/T2	0	0	9.8
Miscellaneous (%)	5.2	2.7	5.4
Intermediate head phenotypes (n)	0	2	61
Head appendages affected (%):			
Lbr	N/A	N/A	47.5
Ant	N/A	N/A	36.1
Mn	N/A	N/A	100
Mx	N/A	N/A	41
Lb	N/A	N/A	54.1
Region affected (excluding Mn) (%):			
Anterior (Lbr/Ant) only	N/A	N/A	29.5
Gnathal (Mx/Lb) only	N/A	N/A	37.7
Anterior and Gnathal	N/A	N/A	32.8

Table D.2: Impact of *Tc-emc* pRNAi on egg laying. Females were injected with 2µg/µl conc. dsRNA. Collections were taken over 2 days, with 1 day rests on whole flour in between. Collection 1 was 6 days after injections.

Treatment	females injected	females survived	Collection 1		Collection 2		Collection 3		Collection 4		Total period	
			Eggs laid		Eggs laid		Eggs laid		Eggs laid		Eggs laid	Eggs / female
GFP (ctrl)	29	9	39		101		122		153		415	46.1
<i>Tc-emc</i> 5'	34	26	13		14		15		15		57	2.2
<i>Tc-emc</i> 3'	32	16	0		0		0		0		0	0

Table D.3. *Tc-Sidpn1* parental RNAi results in a high proportion of empty eggs. Females were injected with 3µg/µl concentration dsRNA. Collections 1 to 3 were taken over 2 day periods, with 1 day rests on whole flour in between. Collection 4 was made 2 weeks later to assess lasting effects. Totals were obtained from collections 1 to 3 only.

Treatment	Collection 1			Collection 2			Collection 3			Collection 4 (+ 2 weeks)			Total (collections 1-3)		
	eggs examined (n)	empty eggs (%)		eggs examined (n)	empty eggs (%)		eggs examined (n)	empty eggs (%)		eggs examined (n)	empty eggs (%)		eggs examined (n)	empty eggs (n)	empty eggs (%)
GFP	48	4.2		295	14.2		138	6.5		201	13.4		481	53	11
<i>Sidpn1</i> 5'	25	80		126	61.1		263	25.5		201	25.9		414	164	39.6
<i>Sidpn1</i> 3'	47	100		348	96		527	96.6		436	34.6		922	890	96.5

Table D.4. *Tc-emc* embryonic RNAi results. Percentage values correspond to the numbers provided for eggs injected (top box) and cuticles scored (lower box). A1/A2 – A7/A8 refer to the segment boundaries where localised irregularities/fusions were recorded. Miscellaneous refers to all other cuticles with abnormalities that are not specified in the table. Segmentation phenotypes were classified as follows: Weak – localised segment irregularities/fusions, or 1 missing segment. Intermediate – Missing 1 segment with multiple fusions, missing 2 segments with up to 2 fusions visible, or missing 3 segments only. Strong – Missing 3 segments with additional fusions, or missing 4 or more segments. 1µg/µl concentrations were used for *Tc-emc*, made up to 2µg/µl total with GFP dsRNA. 2µg/µl GFP dsRNA used for control injections.

	ctrl	<i>emc</i> 5'	<i>emc</i> 3'
Eggs injected (n)	227	229	218
Wildtype cuticles (%)	55.9	3.1	9.2
Abnormal cuticles (%)	10.1	61.1	50.5
Empty eggs (%)	33.9	35.8	40.4
Cuticles scored (n)	150	147	130
Wildtype (%)	84.7	4.8	15.4
Abnormal (%)	15.3	95.2	84.6
Head phenotypes (%)	4	81.6	56.9
Weak: 1-2 deformed/missing head appendages (%)	3.3	52.4	51.5
Strong: 3+ deformed/missing head appendages (%)	0.7	29.3	5.4
Head part affected (%):			
Lbr	0.7	74.8	53.1
Ant	0.7	36.1	16.9
Mn	0.7	8.2	4.6
Mx	1.3	11.6	3.1
Lb	2.7	31.3	3.1
entire head missing	0	5.4	0.8
Segmentation phenotypes (%)	8	82.3	70.8
Weak (%)	7.3	23.8	41.5
Intermediate (%)	0	31.3	24.6
Strong (%)	0.7	27.2	4.6
Weak segmentation phenotypes - irregularities/fusions (%):			
A1/A2	0	1.4	3.1
A2/A3	0.7	1.4	1.5
A3/A4	1.3	0.7	6.2
A4/A5	0.7	4.1	11.5
A5/A6	2	5.4	10.8
A6/A7	2	10.9	20
A7/A8	0	7.5	7.7
Indistinguishable (due to dorsal opening)	0	0	1.5
Deformed/missing legs (%)	2.7	72.8	59.2
Weak hindgut/posterior phenotype (%)	0.7	29.3	50
Strong ballooned hindgut (%)	0	43.5	28.5
Dorsal opening	3.3	54.4	48.5
Miscellaneous / indistinguishable (%)	1.4	0	6.8

Table D.5. *Tc-Sidpn1* embryonic RNAi results. Percentage values correspond to the numbers given in the first row of each box. These being the number of eggs injected, and the cuticles that were scored, from each experiment. T3 (thoracic) and A1-A7 (abdominal) refers to the posterior most segment, from where the truncation occurs. Miscellaneous refers to all other cuticles with abnormalities that are not specified in the table.

	ctrl	<i>Sidpn1</i> 5'	<i>Sidpn1</i> 3'
Eggs injected (n)	227	192	218
Wildtype cuticles (%)	55.9	43.2	11.5
Abnormal cuticles (%)	10.1	11.5	22.5
Empty eggs (%)	33.9	45.3	66.1
Cuticles scored (n)	150	105	74
Wildtype (%)	84.7	79	33.8
Abnormal (%)	15.3	21	66.2
Head abnormalities (%)	4	11.4	37.8
Weak - 1 or 2 head appendages deformed/missing (%)	3.3	4.8	21.6
Strong - 3+ head appendages deformed/missing (%)	0.7	6.7	16.2
Head appendage affected (%):			
Lbr	0.7	1.9	21.6
Ant	0.7	5.7	35.1
Mn	0.7	4.8	18.9
Mx	1.3	1.9	0
Lb	2.7	3.8	2.7
Abdominal truncation (%):	0.7	1	41.9
T3	0	0	2.7
A1	0	0	5.4
A2	0	0	13.5
A3	0.7	0	10.8
A4	0	0	2.7
A5	0	0	5.4
A6	0	0	0
A7	0	1	1.4
Legs deformed/missing (%)	2.7	0	20.3
Miscellaneous (%)	10	13.3	17.6

Table D.6. eRNAi results for *Tc-emc/Tc-h*, and *Tc-emc/Tc-dpn* double knockdowns, compared with *Tc-emc* singles. Percentage values correspond to the numbers provided for eggs injected (top box) and cuticles scored (lower box). Miscellaneous refers to all other cuticles with abnormalities that are not specified in the table. Segmentation phenotypes were classified as follows: Weak – localised segment irregularities/fusions, or 1 missing segment. Intermediate – Missing 1 segment with multiple fusions, missing 2 segments with up to 2 fusions visible, or missing 3 segments only. Strong – Missing 3 segments with additional fusions, or missing 4 or more segments. A total of 2µg/µl dsRNA was used for all experiments. 1:1 ratios were used for mixing, meaning each dsRNA fragment was injected at 1µg/µl. 2µg/µl *GFP* dsRNA was used for ctrl

	ctrl	<i>emc</i> 5'/ <i>GFP</i>	<i>emc</i> 3'/ <i>GFP</i>	<i>emc</i> 5'/ <i>h</i> 3'	<i>emc</i> 3'/ <i>h</i> 3'	<i>emc</i> 5'/ <i>dpn</i> 5'	<i>emc</i> 3'/ <i>dpn</i> 3'
Eggs injected (n)	227	229	218	150	154	138	184
Wildtype cuticles (%)	55.9	3.1	9.2	10	17.5	21	6
Abnormal cuticles (%)	10.1	61.1	50.5	48	52.6	51.4	56.5
Empty eggs (%)	33.9	35.8	40.4	42	29.9	27.5	37.5
Cuticles scored (n)	150	147	130	87	108	100	115
Wildtype (%)	84.7	4.8	15.4	17.2	25	29	9.6
Abnormal (%)	15.3	95.2	84.6	82.8	75	71	90.4
Head phenotypes (%)	4	81.6	56.9	55.2	38	50	61.7
Weak: 1-2 deformed/missing head appendages (%)	3.3	52.4	51.5	31	21.3	41	47
Strong: 3+ deformed/missing head appendages (%)	0.7	29.3	5.4	24.1	16.7	9	14.8
Head part affected (%):							
Lbr	0.7	74.8	53.1	51.7	25.9	49	60
Ant	0.7	36.1	16.9	16.1	13.9	14	27.8
Mn	0.7	8.2	4.6	25.3	13	7	8.7
Mx	1.3	11.6	3.1	8	9.3	5	7.8
Lb	2.7	31.3	3.1	19.5	8.3	14	13
entire head missing	0	5.4	0.8	1.1	6.5	0	0.9
Segmentation phenotypes (%)	8	82.3	70.8	69	51.9	61	74.8
Weak (%)	7.3	23.8	41.5	20.7	26.9	27	30.4
Intermediate (%)	0	31.3	24.6	21.8	9.3	27	29.6
Strong (%)	0.7	27.2	4.6	26.4	15.7	7	14.8
Deformed/missing legs (%)	2.7	72.8	59.2	59.8	33.3	52	71.3
Weak hindgut/posterior phenotype (%)	0.7	29.3	50	18.3	25.9	25	42.6
Strong ballooned hindgut (%)	0	43.5	28.5	40.2	21.3	31	31.3
Dorsal opening (%)	3.3	54.4	48.5	23	17.6	30	31.3
Miscellaneous / indistinguishable (%)	1.4	0	6.8	4.6	6.5	5	4.3

Table D.7. *Tc-h/T-dpn* pRNAi double knockdown. Percentage values correspond to the proportions of eggs examined (top box), and cuticles scored (lower box). A total of 2µg/µl dsRNA was used in all experiments. For *Tc-h/Tc-dpn* doubles, a 1:1 ratio was used, meaning each dsRNA fragment was injected at 1µg/µl. Miscellaneous refers to all other cuticles with abnormalities not specified in the table.

	Experiment 1		Experiment 2		
	ctrl	<i>h</i> 3'	ctrl	<i>h</i> 3' / <i>dpn</i> 5'	<i>h</i> 3' / <i>dpn</i> 3'
Eggs examined (n)	264	379	628	421	401
Wildtype cuticles (%)	81.1	21.9	84.2	44.9	67.6
Abnormal cuticles (%)	5.3	61.7	2.9	37.8	15
Empty eggs (%)	13.6	16.4	12.9	17.3	17.5
Cuticles scored (n)	228	317	547	348	331
Wildtype (%)	93.9	26.2	96.7	54.3	81.9
Abnormal (%)	6.1	73.8	3.3	45.7	18.1
Head defect (%)	0.9	71.6	0.4	44.3	16
Weak - missing/deformed Mn	0	30.9	0	32.8	11.5
Intermediate - Additional head appendages missing/deformed	0	19.2	0	8	3
Strong - Missing entire head	0	11.4	0	1.7	1.2
Strong - Missing head + T1/T2	0	9.8	0	0.6	0.3
Miscellaneous (%)	5.2	5.4	2.9	2.3	3

Table D.8. pRNAi results for *Tc-h/Tc-odd* and *Tc-dpn/Tc-odd* double knockdowns.

Percentage values correspond to the proportion of eggs examined (top box), and cuticles scored (lower box). Segmentation phenotypes were classified based on *odd* induced pair-rule-like effects. Weak: localised segment irregularities/fusions, or 1 missing segment. Intermediate: Missing 2 to 4 segments. Strong: Pair-rule-like phenotype or worse, in which 5 or more thoracic or abdominal segments are missing. Miscellaneous/indistinguishable refers to cuticles with abnormalities not specified in the table, or that I was unable to score. dsRNA was mixed at a 1:1 ratio to a total concentration of 2µg/µL.

	Ctrl	<i>odd</i> 5'/ <i>GFP</i>	<i>odd</i> 5'/ <i>h</i> 3'	<i>odd</i> 5'/ <i>dpn</i> 5'	<i>odd</i> 5'/ <i>dpn</i> 3'
Eggs examined (n)	628	436	553	187	334
Wildtype cuticles (%)	84.2	6.9	10.5	13.9	1.5
Abnormal cuticles (%)	2.9	77.3	70.3	66.3	81.4
Empty eggs (%)	12.9	15.8	19.2	19.8	17.1
Cuticles scored (n)	547	367	447	150	277
Wildtype (%)	96.7	8.2	13	17.3	1.8
Abnormal (%)	3.3	91.8	87	82.7	98.2
Head phenotype (%)	0.4	86.1	78.5	56.7	86.3
Weak - 1-2 head appendages missing/deformed (%)	0.4	83.1	68.5	53.3	85.2
Strong - 3+ head appendages missing/deformed (%)	0	3	10.1	3.3	1.1
Head part affected (%):					
Lbr	0	2.5	3.1	2	1.4
Ant	0	6.8	3.1	0.7	2.5
Mn	0	4.9	14.3	4.7	3.2
Mx	0.2	2.5	5.4	4	2.2
Lb	0.2	84.7	73.8	56.7	86.3
entire head missing	0	0.5	3.8	0	0
Segmentation phenotype (%)	0.5	89.9	82.1	80	97.5
Weak (%)	0.5	6.5	6.3	24.7	18.8
Intermediate (%)	0	42.2	49.4	50	65
Strong (%)	0	41.1	26.4	5.3	13.7
Leg count (%):					
0 legs	0	1.9	12.1	0	1.1
1 leg	0	2.2	8.3	0.7	0.7
2 legs	0	11.2	20.4	4.7	3.2
3 legs	0	17.7	13.2	6	9.4
Deformed legs (%)	0	2.5	1.8	3.3	4.3
Miscellaneous / indistinguishable (%)	2.4	1.9	4.5	2	0.4

Table D.9. *Tc-odd* mutant pRNAi results. Counts are presented as percentages of the numbers given in the first row of each box. These being eggs examined, total cuticles, *odd* null cuticles, and cuticles excluding *odd* null. Percentages are given to 1 decimal point. Miscellaneous refers to all other cuticles with abnormalities that are not specified in the table. Single knockdowns were done with 2µg/µl dsRNA, while doubles used a 1:1 mixture with a combined total of 2µg/µl. Results are split into 2 experiments that were done at separate times, each alongside respective GFP injected controls.

	Experiment 1			Experiment 2			
	Ctrl	<i>Tc-dpn</i> 5'	<i>Tc-dpn</i> 3'	Ctrl	<i>Tc-h</i> 3'	<i>Tc-h</i> 3'/ <i>Tc-dpn</i> 5'	<i>Tc-h</i> 3'/ <i>Tc-dpn</i> 3'
Eggs Examined (n)	215	377	585	789	661	670	490
Wildtype cuticles (%)	53	41.4	53.2	55.3	18.6	33	40.2
Abnormal cuticles (%)	26	32.1	23.4	18.5	48.9	38.9	30.4
Empty eggs (%)	20.9	26.5	23.4	26.2	32.5	28.1	29.4
Total cuticles scored (n)	170	277	448	582	446	482	346
Wildtype (%)	67.1	56.3	69.4	74.9	27.6	45.9	56.9
<i>odd</i> null (%)	24.7	26.7	20.1	21.3	19.7	21.6	21.1
abnormal-other (%)	8.2	17	10.5	3.8	52.7	32.6	22
<i>odd</i> null cuticles (n)	42	74	90	124	88	104	73
In addition to Lb:							
1 missing/deformed head appendage (%)	16.7	16.2	26.7	18.5	38.6	23.1	39.7
2 or more missing/deformed head appendages (%)	2.4	8.1	4.4	4	12.5	11.5	8.2
Missing/deformed Mn (%)	4.8	4.1	3.3	8.1	39.8	26.9	26
Spiracle counts (%):							
0	4.8	9.5	5.6	4	13.6	11.5	20.5
1	16.7	20.3	17.8	14.5	30.7	29.8	35.6
2	26.2	33.8	28.9	29	36.4	33.7	26
3	40.5	28.4	34.4	31.5	12.5	21.2	15.1
4	11.9	8.1	13.3	18.5	6.8	3.8	2.7
5	0	0	0	2.4	0	0	0
Leg counts (%):							
0	21.4	32.4	31.1	24.2	68.2	49	71.2
1	31	29.7	35.6	21.8	14.8	37.5	15.1
2	26.2	25.7	28.9	39.5	13.6	12.5	11
3	4.8	8.1	1.1	8.9	2.3	1	1.4
4	11.9	4.1	3.3	5.6	1.1	0	1.4
cuticles excluding <i>odd</i> null (n)	128	203	358	458	358	378	273
Head phenotype (%)	0.8	3	0.8	0.6	58.1	37.3	20.9
weak - missing 1 head appendage (%)	0.8	2	0.8	0.2	32.4	27.5	15.4
Intermediate - missing 2-4 head appendages (%)	0	1	0	0.2	12.8	5.6	4
Strong - missing entire head (%)	0	0	0	0.2	12.9	4.2	1.5
Miscellaneous (%)	10.2	20.2	12.3	4.2	7.5	4.2	7

Table D.10. *Tc-dpn* pRNAi egg laying. The numbers of egg laying females, and the eggs they lay for each consecutive collection after injections, are provided for all experiments that included *Tc-dpn*, along with respective GFP controls. For double knockdowns carried out alongside single knockdowns, data is also given for the other genes involved. Beetles were injected with 2µg/µL total dsRNA in all experiments. In cases where two genes are used, dsRNA was mixed at a 1:1 ratio. The “% Ctrl” column shows the relative eggs/female tally for each experiment compared to their respective controls.

Experiment	Treatment	Females survived	Collection 1		Collection 2		Collection 3		Total	
			Eggs laid	Eggs/female	Eggs laid	Eggs/female	Eggs laid	Eggs/female	Eggs laid	Eggs/female
<i>Tc-h/Tc-dpn</i> doubles	Ctrl	18	235	13.1	393	21.8	-	-	628	34.9
	<i>h 3'/dpn 5'</i>	13	170	13.1	251	19.3	-	-	421	32.4
	<i>h 3'/dpn 3'</i>	14	119	8.5	282	20.1	-	-	401	28.6
<i>Tc-odd/Tc-dpn</i> doubles	Ctrl	18	235	13.1	393	21.8	-	-	628	34.9
	<i>odd 5'/GFP</i>	14	115	8.2	321	22.9	-	-	436	31.1
	<i>odd 5'/dpn 5'</i>	7	58	8.3	102	14.6	-	-	160	22.9
	<i>odd 5'/dpn 3'</i>	14	62	4.4	272	19.4	-	-	334	23.9
odd mutant singles	Ctrl	5	2	0.4	32	6.4	97	19.4	131	26.2
	<i>dpn 5'</i>	23	6	0.3	57	2.5	314	13.7	377	16.4
	<i>dpn 3'</i>	28	19	0.7	112	4	447	16	578	20.6
odd mutant doubles	Ctrl	15	165	11	176	11.7	216	14.4	557	37.1
	<i>h 3'</i>	15	68	4.5	158	10.5	186	12.4	412	27.5
	<i>h 3'/dpn 5'</i>	15	67	4.5	138	9.2	199	13.3	404	26.9
	<i>h 3'/dpn 3'</i>	13	69	5.3	118	9.1	136	10.5	323	24.8

Appendix E – Scoring localised segmentation defects from *Tc-emc* knockdowns

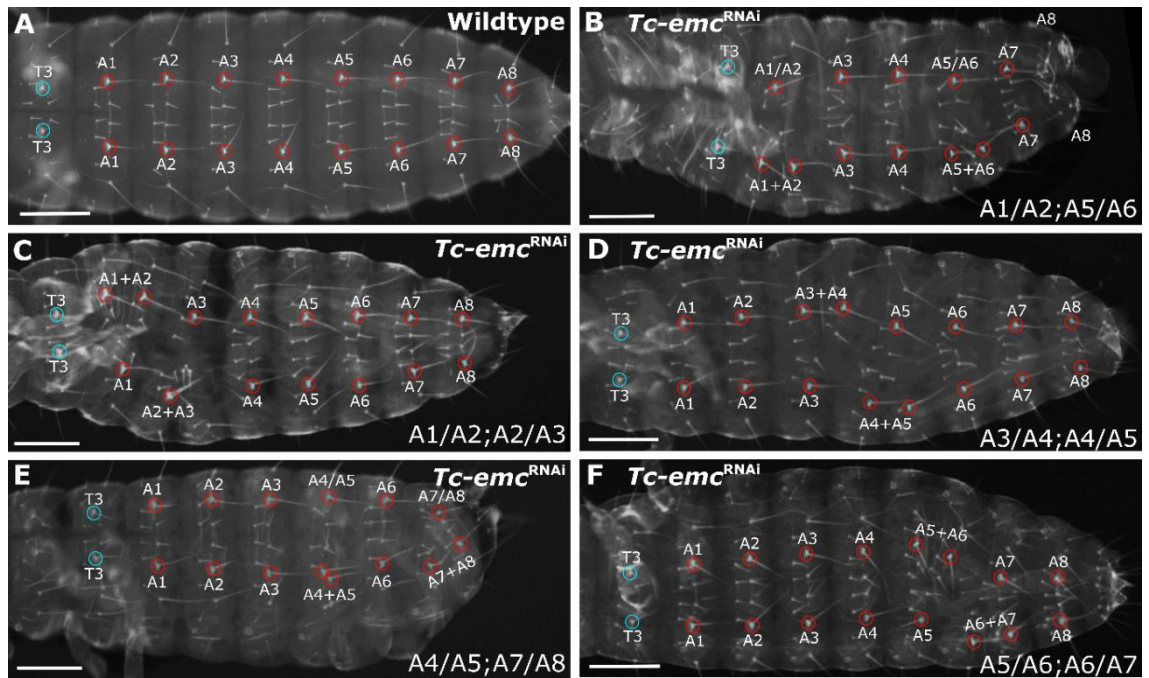


Figure E.1. Additional examples of *Tc-emc* RNAi induced localised abnormalities. (A) wildtype trunk. (B-F) trunks of embryos from *Tc-emc* eRNAi experiments. The abdominal segment boundaries that were scored as abnormal for **Fig. 2B** are labelled in the bottom right of each image. Cyan (thoracic) and red (abdominal) circles highlight the positions of homologous bristles. Scale bar: 100 μ m

Appendix F – Sense controls confirm novel gene expression patterns from colorimetric *in situ* hybridisation

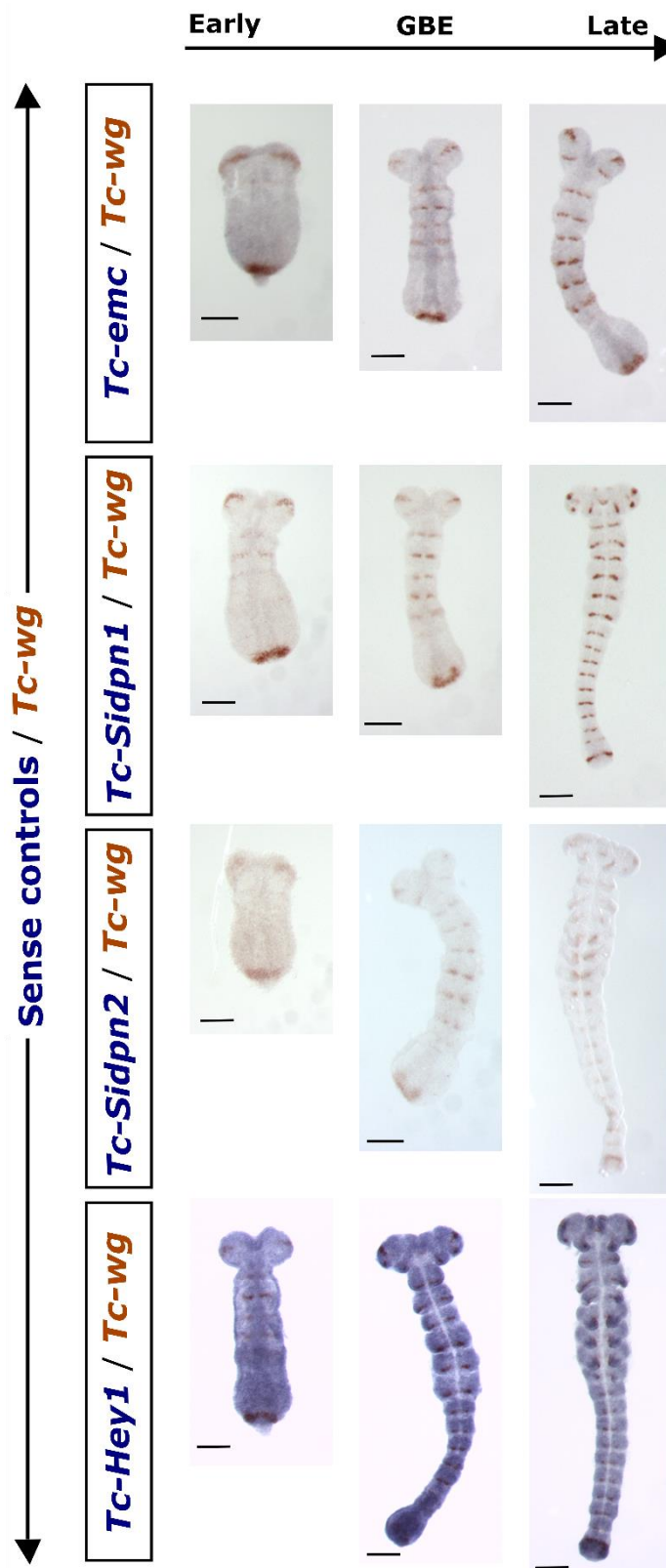


Figure F.1. ISH sense controls for *Tc-emc*, *Tc-Sidpn1*, *Tc-Sidpn2* and *Tc-Hey1*. For genes whose expression patterns have not been previously described, colorimetric ISH experiments using sense strand probes were carried out alongside the corresponding antisense probes. These act as a negative control to confirm the genuine expression patterns reported in chapter 5. No staining is visible in embryos treated with *Tc-emc*, *Tc-Sidpn1*, *Tc-Sidpn2* sense probes. There is ubiquitous staining in embryos treated with the *Tc-Hey1* sense probe, however, the neuroectodermal spots observed for *Tc-Hey1* in **Fig. 5.4** are not present, suggesting these represent real gene expression patterns. As for why ubiquitous non-specific staining was observed in sense strand embryos is unclear. The correct insert was confirmed using restriction digests, while a BLAST query against the *Tribolium* genome assembly failed to return any significant hits to other genes.

BIBLIOGRAPHY

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410.
[https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Angelini, D. R., Kikuchi, M., & Jockusch, E. L. (2009). Genetic patterning in the adult capitulate antenna of the beetle *Tribolium castaneum*. *Developmental Biology*, 327(1), 240–251.
<https://doi.org/10.1016/j.ydbio.2008.10.047>
- Angelini, D. R., Smith, F. W., & Jockusch, E. L. (2012)(a). Extent with modification: Leg patterning in the beetle *tribolium castaneum* and the evolution of serial homologs. *G3: Genes, Genomes, Genetics*, 2(2), 235–248. <https://doi.org/10.1534/g3.111.001537>
- Angelini, D. R., Smith, F. W., Aspiras, A. C., Kikuchi, M., & Jockusch, E. L. (2012)(b). Patterning of the adult mandibulate mouthparts in the red flour beetle, *Tribolium castaneum*. *Genetics*, 190(2), 639–654. <https://doi.org/10.1534/genetics.111.134296>
- Aranda, M. (2006) Functional analysis of a homolog of the pair-rule gene hairy in the short-germ beetle *Tribolium castaneum*. *PhD thesis*, Universität zu Köln
- Aranda, M., Marques-Souza, H., Bayer, T., & Tautz, D. (2008). The role of the segmentation gene hairy in *Tribolium*. *Development Genes and Evolution*, 218(9), 465–477.
<https://doi.org/10.1007/s00427-008-0240-1>
- Auman, T., & Chipman, A. D. (n.d.). *Growth zone segmentation in the milkweed bug Oncopeltus fasciatus sheds light on the evolution of insect segmentation*.
<https://doi.org/10.1186/s12862-018-1293-z>
- Auman, T., Vreede, B. M. I., Weiss, A., Hester, S. D., Williams, T. A., Nagy, L. M., & Chipman, A. D. (2017). *Dynamics of growth zone patterning in the milkweed bug Oncopeltus fasciatus*. <https://doi.org/10.1242/dev.142091>
- Bai, G., Sheng, N., Xie, Z., Bian, W., Yokota, Y., Benezra, R., Kageyama, R., Guillemot, F., & Jing, N. (2007). Id Sustains Hes1 Expression to Inhibit Precocious Neurogenesis by Releasing

- Negative Autoregulation of Hes1. *Developmental Cell*, 13(2), 283–297.
<https://doi.org/10.1016/j.devcel.2007.05.014>
- Baonza, A., De Celis, J. F., & García-Bellido, A. (2000). Relationships between extramacrochaetae and Notch signalling in *Drosophila* wing development. *Development*, 127(11), 2383–2393. <https://doi.org/10.1242/dev.127.11.2383>
- Benedyk, M. J., Mullen, J. R., & DiNardo, S. (1994). Odd-paired: A zinc finger pair-rule protein required for the timely activation of engrailed and wingless in *Drosophila* embryos. *Genes and Development*, 8(1), 105–117. <https://doi.org/10.1101/gad.8.1.105>
- Benton, M. A., Akam, M., & Pavlopoulos, A. (2013). Cell and tissue dynamics during *Tribolium* embryogenesis revealed by versatile fluorescence labeling approaches. *Development (Cambridge)*, 140(15), 3210–3220. <https://doi.org/10.1242/dev.096271>
- Benton, M. A., & Pavlopoulos, A. (2014). *Tribolium* embryo morphogenesis: May the force be with you. *BioArchitecture*, 4(1), 16–21. <https://doi.org/10.4161/bioa.27815>
- Benton, M. A., Pechmann, M., Frey, N., Stamatakis, E., Pavlopoulos, A., & Correspondence, S. R. (2016). Toll Genes Have an Ancestral Role in Axis Elongation. *Current Biology*, 26, 1609–1615. <https://doi.org/10.1016/j.cub.2016.04.055>
- Bettencourt, B. R., & Feder, M. E. (2002). Rapid concerted evolution via gene conversion at the *Drosophila* hsp70 genes. *Journal of Molecular Evolution*, 54(5), 569–586.
<https://doi.org/10.1007/s00239-001-0044-7>
- Betts, M. J., Guigo, R., Agarwal, P. and Russell, R. B. (2001). Exon structure conservation despite low sequence similarity: a relic of dramatic events in evolution? *The EMBO Journal*, 20(19), 5354–5360. <https://doi.org/10.1093/emboj/20.19.5354>
- Bhattacharya, A., & Baker, N. E. (2009). The HLH protein Extramacrochaetae is required for R7 cell and cone cell fates in the *Drosophila* eye. *Developmental Biology*, 327(2), 288–300.
<https://doi.org/10.1016/J.YDBIO.2008.11.037>
- Bi, S., Yue, S., & Zhang, S. (2017). Hybridization chain reaction: A versatile molecular tool for biosensing, bioimaging, and biomedicine. *Chemical Society Reviews*, 46(14), 4281–4298.
<https://doi.org/10.1039/c7cs00055c>

- Bier, E., Vaessin, H., Younger-Shepherd, S., Lily Yeh Jan, & Yuh Nung Jan. (1992). deadpan, an essential pan-neural gene in *Drosophila*, encodes a helix-loop- helix protein similar to the hairy gene product. *Genes and Development*, 6(11), 2137–2151.
<https://doi.org/10.1101/gad.6.11.2137>
- Biffar, L. (2013). Early neurogenesis in the flour beetle *Tribolium castaneum*. *PhD thesis*, Queen Mary, University London
- Biffar, L., & Stollewerk, A. (2015). Evolutionary variations in the expression of dorso-ventral patterning genes and the conservation of pioneer neurons in *Tribolium castaneum*. *Developmental Biology*, 400(1), 159–167. <https://doi.org/10.1016/j.ydbio.2015.01.025>
- Birchler, J. A., & Yang, H. (2022). The multiple fates of gene duplications: Deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. *Plant Cell*, 34(7), 2466–2474.
<https://doi.org/10.1093/plcell/koac076>
- Bitra, K., & Palli, S. R. (2009). The members of bHLH transcription factor superfamily are required for female reproduction in the red flour beetle, *Tribolium castaneum*. *Journal of Insect Physiology*, 56(10), 1481–1489. <https://doi.org/10.1016/j.jinsphys.2010.03.005>
- Bitra, K., Tan, A., Dowling, A., & Palli, S. R. (2013). Functional characterization of PAS and HES family bHLH transcription factors during the metamorphosis of the red flour beetle, *Tribolium castaneum*. *Magn Reson Imaging*, 31(3), 477–479.
<https://doi.org/10.1016/j.immuni.2010.12.017>.Two-stage
- Boareto, M., Iber, D., & Taylor, V. (2017). Differential interactions between Notch and ID factors control neurogenesis by modulating Hes factor autoregulation. *Development (Cambridge)*, 144(19), 3465–3474. <https://doi.org/10.1242/dev.152520>
- Bolognesi, R., Farzana, L., Fischer, T. D., & Brown, S. J. (2008). Multiple Wnt Genes Are Required for Segmentation in the Short-Germ Embryo of *Tribolium castaneum*. *Current Biology*, 18(20), 1624–1629. <https://doi.org/10.1016/j.cub.2008.09.057>
- Bolognesi, R., Fischer, T. D., & Brown, S. J. (2009). Loss of Tc-arrow and canonical Wnt signaling alters posterior morphology and pair-rule gene expression in the short-germ

- insect, *Tribolium castaneum*. *Development Genes and Evolution*, 219(7), 369–375.
<https://doi.org/10.1007/s00427-009-0299-3>
- Bouckaert, R. R. (2020). OBAMA: OBAMA for Bayesian aminoacid model averaging. *PeerJ*, 8, 1–15. <https://doi.org/10.7717/peerj.9460>
- Bouckaert, R. R., & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, 17(1), 1–11.
<https://doi.org/10.1186/s12862-017-0890-6>
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, 10(4), 1–6.
<https://doi.org/10.1371/journal.pcbi.1003537>
- Brown, S. J., Patel, N. H., Denell, R. E. (1994). Embryonic expression of the single *Tribolium* engrailed homolog. *Dev Genet* 15: 7-18.
- Brown, S. J., Shippy, T. D., Miller, S., Bolognesi, R., Beeman, R. W., Lorenzen, M. D., Bucher, G., Wimmer, E. A., & Klingler, M. (2009). The red flour beetle, *Tribolium castaneum* (Coleoptera): A model for studies of development and pest biology. *Cold Spring Harbor Protocols*, 4(8), 1–10. <https://doi.org/10.1101/pdb.emo126>
- Bruce, H. S. (2021). How to align arthropod legs. *BioRxiv*, 2021.01.20.427514.
- Bruce, H. S., Jerz, G., Kelly, S., McCarthy, J., Pomerantz, A., Senevirathne, G., Sherrard, A., Sun, D. A., Wolff, C., & Patel, N. H. (2021). Hybridization Chain Reaction (HCR) In Situ Protocol. *Protocols.io*, September, 1–10.
<https://doi.org/10.17504/protocols.io.bunznvf6>
- Bucher, G. (2009). *The Beetle Book*
- Bucher, G., Scholten, J., & Klingler, M. (2002). Parental RNAi in *tribolium* (coleoptera). *Current Biology*, 12(3), 85–86. [https://doi.org/10.1016/S0960-9822\(02\)00666-8](https://doi.org/10.1016/S0960-9822(02)00666-8)
- Cabrera, C. V., Alonso, M. C., & Huikeshoven, H. (1994). Regulation of scute function by extramacrochaete in vitro and in vivo. *Development*, 120(12), 3595–3603.
<https://doi.org/10.1242/dev.120.12.3595>

- Campuzano, S. (2001). Emc, a negative HLH regulator with multiple functions in Drosophila development. *Oncogene*, 20(58 REV. ISS. 8), 8299–8307.
<https://doi.org/10.1038/sj/onc/1205162>
- Carroll, S. B., Laughon, A., & Thalley, B. S. (1988). Expression, function, and regulation of the hairy segmentation protein in the Drosophila embryo. *Genes & Development*, 2(7), 883–890. <https://doi.org/10.1101/gad.2.7.883>
- Cepeda, R. E., Terraza, J. B., Pardo, R. V., Núñez-Pascual, V., Mundaca-Escobar, M., & Sarrazin, A. F. (123 C.E.). Spatiotemporal variation in cell proliferation patterns during arthropod axial elongation. *Scientific Reports* |, 11, 327. <https://doi.org/10.1038/s41598-020-79373-0>
- Chesebro, J. (2012). Mechanisms of segmentation in the American cockroach, *Periplaneta americana*. *PhD thesis*, University of Sussex
- Chesebro, J. E., Pueyo, J. I., & Couso, J. P. (2013). Interplay between a Wnt-dependent organiser and the Notch segmentation clock regulates posterior development in *Periplaneta americana*. *Biology Open*, 2(2), 227–237.
<https://doi.org/10.1242/bio.20123699>
- Chipman, A. D. (2010). Parallel evolution of segmentation by co-option of ancestral gene regulatory networks. *BioEssays*, 32(1), 60–70. <https://doi.org/10.1002/bies.200900130>
- Chipman, A. D. (2017). *Oncopeltus fasciatus* as an evo-devo research organism. *Genesis*, 55(5). <https://doi.org/10.1002/dvg.23020>
- Chipman, A. D., & Edgecombe, G. D. (2019). Developing an integrated understanding of the evolution of arthropod segmentation using fossils and evo-devo. *Proceedings of the Royal Society B: Biological Sciences*, 286(1912). <https://doi.org/10.1098/rspb.2019.1881>
- Chipman, A. D., & Akam, M. (2008). *The segmentation cascade in the centipede Strigamia maritima: Involvement of the Notch pathway and pair-rule gene homologues*.
<https://doi.org/10.1016/j.ydbio.2008.02.038>

- Chipman, A. D., Arthur, W., & Akam, M. (2004). A Double Segment Periodicity Underlies Segment Generation in Centipede Development. *Current Biology*, 14(14), 1250–1255. <https://doi.org/10.1016/j.cub.2004.07.026>
- Choe, C. P., & 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*. *Developmental Biology*, 302(1), 281–294. <https://doi.org/10.1016/j.ydbio.2006.09.037>
- Choe, C. P., Miller, S. C., & Brown, S. J. (2006). *A pair-rule gene circuit defines segments sequentially in the short-germ insect Tribolium castaneum*. www.pnas.org/cgi/doi/10.1073/pnas.0510440103
- Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., & Pierce, N. A. (2018). Third-generation in situ hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Development (Cambridge)*, 145(12), 1–10. <https://doi.org/10.1242/dev.165753>
- Clark, E. (2017). Dynamic patterning by the *Drosophila* pair-rule network reconciles long-germ and short-germ segmentation. *PLoS Biology*. <https://doi.org/10.1371/journal.pbio.2002439>
- Clark, E., & Akam, M. (2016). Odd-paired controls frequency doubling in *Drosophila* segmentation by altering the pair-rule gene regulatory network. *eLife*, 5(AUGUST), 1–42. <https://doi.org/10.7554/eLife.18215>
- Clark, E., & Peel, A. D. (2018). Evidence for the temporal regulation of insect segmentatytion by a conserved sequence of transcription factors. *Development*, 145.
- Clark, E., Peel, A. D., & Akam, M. (2019). *Arthropod segmentation*. <https://doi.org/10.1242/dev.170480>
- Constantinou, S. J., Duan, N., Nagy, L. M., Chipman, A. D., & Williams, T. A. (2020). Elongation during segmentation shows axial variability, low mitotic rates, and synchronized cell cycle domains in the crustacean, *Thamnocephalus platyurus*. *EvoDevo*, 11(1), 1–18. <https://doi.org/10.1186/s13227-020-0147-0>

- Cooker, J., & Zeeman, E. C. (1976). A Clock and Wavefront Model for Control of the Number of Repeated Structures during Animal Morphogenesis. In *J. theor. Biol* (Vol. 58). https://ac.els-cdn.com/S0022519376801312/1-s2.0-S0022519376801312-main.pdf?_tid=4701c8e9-700e-4f4d-b2c8-61ef5d79a698&acdnat=1546515044_c5568da0f7213f2448b4847aa2c5c053
- Coulter¹, D. E., Swaykus¹, E. A., Beran-Koehn¹, M. A., Goldberg², D., Wieschaus², E., & Schedl², P. (1990). Molecular analysis of odd-skipped, a zinc finger encoding segmentation gene with a novel pair-rule expression pattern. *The EMBO Journal*, 8(12), 3795–3804.
- Cronmiller, C., & Cummings, C. A. (1993). The daughterless gene product in *Drosophila* is a nuclear protein that is broadly expressed throughout the organism during development. *Mechanisms of Development*, 42(3), 159–169. [https://doi.org/10.1016/0925-4773\(93\)90005-I](https://doi.org/10.1016/0925-4773(93)90005-I)
- Cubas, P., Modolell, J., & Ruiz-Gomez, M. (1994). The helix-loop-helix extramacrochaetae protein is required for proper specification of many cell types in the *Drosophila* embryo. *Development*, 120(9), 2555–2566. <https://doi.org/10.1242/dev.120.9.2555>
- Culí, J., & Modolell, J. (1998). Proneural gene self-stimulation in neural precursors: An essential mechanism for sense organ development that is regulated by Notch signaling. *Genes and Development*, 12(13), 2036–2047. <https://doi.org/10.1101/gad.12.13.2036>
- Damen, W. G. M., Weller, M., & 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. *Proceedings of the National Academy of Sciences*, 97(9), 4515–4519. <https://doi.org/10.1073/pnas.97.9.4515>
- Damen, W. G. M., Janssen, R., & Prpic, N.-M. (2005). Pair rule gene orthologs in spider segmentation. *Evolution & Development*, 7(6), 618–628. <https://doi.org/10.1111/j.1525-142X.2005.05065.x>
- Davis, G. K., D'Alessio, J. A., & Patel, N. H. (2005). Pax3/7 genes reveal conservation and divergence in the arthropod segmentation hierarchy. *Developmental Biology*, 285(1), 169–184. <https://doi.org/10.1016/j.ydbio.2005.06.014>

- Davis, G. K., & Patel, N. H. (2002). *SHORT, LONG, AND BEYOND: Molecular and Embryological Approaches to Insect Segmentation*. www.annualreviews.org
- Dawson, S. R., Turner, D. L., Weintraub, H., & Parkhurst, S. M. (1995). Specificity for the Hairy/Enhancer of split Basic Helix-Loop-Helix (bHLH) Proteins Maps outside the bHLH Domain and Suggests Two Separable Modes of Transcriptional Repression †. In *MOLECULAR AND CELLULAR BIOLOGY* (Vol. 15, Issue 12). <http://mcb.asm.org/>
- Dearden, P. K. (2015). Origin and evolution of the enhancer of split complex. *BMC Genomics*. <https://doi.org/10.1186/s12864-015-1926-1>
- Deutsch, J. S. (2004). Segments and parasegments in arthropods: A functional perspective. *BioEssays*, 26(10), 1117–1125. <https://doi.org/10.1002/bies.20111>
- Diaz-Cuadros, M., Pourquié, O., & El-Sherif, E. (2021). Patterning with clocks and genetic cascades: Segmentation and regionalization of vertebrate versus insect body plans. *PLoS Genetics* (Vol. 17, Issue 10). <https://doi.org/10.1371/journal.pgen.1009812>
- Duncan, E. J., & Dearden, P. K. (2010). Evolution of a genomic regulatory domain: The role of gene co-option and gene duplication in the enhancer of split complex. *Genome Research*, 20(7), 917–928. <https://doi.org/10.1101/gr.104794.109>
- Dunn, C. W., Giribet, G., Edgecombe, G. D., & Hejnol, A. (2014). Animal phylogeny and its evolutionary implications. *Annual Review of Ecology, Evolution, and Systematics*, 45, 371–395. <https://doi.org/10.1146/annurev-ecolsys-120213-091627>
- Eckert, C., Aranda, M., Wolff, C., & Tautz, D. (2004). Separable stripe enhancer elements for the pair-rule gene hairy in the beetle *Tribolium*. *EMBO Reports*, 5(6), 638–642. <https://doi.org/10.1038/sj.embor.7400148>
- Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5, 1–19. <https://doi.org/10.1186/1471-2105-5-113>
- El-Sherif, E., Zhu, X., Fu, J., & Brown, S. J. (2014). Caudal Regulates the Spatiotemporal Dynamics of Pair-Rule Waves in *Tribolium*. *PLoS Genetics*, 10(10). <https://doi.org/10.1371/journal.pgen.1004677>

- Erezylmaz, D. F., Kelstrup, H. C., & Riddiford, L. M. (2009). The nuclear receptor E75A has a novel pair-rule-like function in patterning the milkweed bug, *Oncopeltus fasciatus*. *Developmental Biology*, 334(1), 300–310. <https://doi.org/10.1016/j.ydbio.2009.06.038>
- Eriksson, B. J., Ungerer, P., & Stollewerk, A. (2013). The function of Notch signalling in segment formation in the crustacean *Daphnia magna* (Branchiopoda). *Developmental Biology*, 383(2), 321–330. <https://doi.org/10.1016/j.ydbio.2013.09.021>
- Farzana, L., & Brown, S. J. (2008). Hedgehog signaling pathway function conserved in *Tribolium* segmentation. *Development Genes and Evolution*, 218(3–4), 181–192. <https://doi.org/10.1007/s00427-008-0207-2>
- Fawcett, J. A., & Innan, H. (2011). Neutral and non-neutral evolution of duplicated genes with gene conversion. *Genes*, 2(1), 191–209. <https://doi.org/10.3390/genes2010191>
- Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151(4), 1531–1545. <https://doi.org/10.1093/genetics/151.4.1531>
- Garcia-Perez, N. C., Bucher, G., & Buescher, M. (2021). The homeodomain transcription factor shaking hands controls axon outgrowth of central complex neurons in the insect model *Tribolium*. *Development (Cambridge)*, 148(19). <https://doi.org/10.1242/DEV.199368>
- Gazave, E., Guillou, A., & Balavoine, G. (2014). History of a prolific family: The Hes/Hey-related genes of the annelid *Platynereis*. *EvoDevo*, 5(1), 1–33. <https://doi.org/10.1186/2041-9139-5-29>
- Green, J., & Akam, M. (2013). Evolution of the pair rule gene network: Insights from a centipede. *Developmental Biology*, 382(1), 235–245. <https://doi.org/10.1016/j.ydbio.2013.06.017>
- Green, R. B., Hatini, V., Johansen, K. A., Liu, X. J., & Lengyel, J. A. (2002). Drumstick is a zinc finger protein that antagonizes Lines to control patterning and morphogenesis of the *Drosophila* hindgut. *Development*, 129(15), 3645–3656. <https://doi.org/10.1242/dev.129.15.3645>

- Hammonds, A. S., Bristow, C. A., Fisher, W. W., Weizmann, R., Wu, S., Hartenstein, V., Kellis, M., Yu, B., Frise, E., & Celniker, S. E. (2013). Spatial expression of transcription factors in *Drosophila* embryonic organ development. *Genome Biology*, *14*(12), 1–15. <https://doi.org/10.1186/gb-2013-14-12-r140>
- Hannibal, R. L., & Patel, N. H. (2013). What is a segment? *EvoDevo*, *4*(1), 1–10. <https://doi.org/10.1186/2041-9139-4-35>
- Hao, I., Green, R. B., Dunaevsky, O., Lengyel, J. A., & Rauskolb, C. (2003). The odd-skipped family of zinc finger genes promotes *Drosophila* leg segmentation. *Developmental Biology*, *263*(2), 282–295. <https://doi.org/10.1016/j.ydbio.2003.07.011>
- Harima, Y., Takashima, Y., Ueda, Y., Ohtsuka, T., & Kageyama, R. (2013). Accelerating the Tempo of the Segmentation Clock by Reducing the Number of Introns in the *Hes7* Gene. *CellReports*, *3*, 1–7. <https://doi.org/10.1016/j.celrep.2012.11.012>
- Hart, M. C., Wang, L., & Coulter, D. E. (1996). Comparison of the structure and expression of odd-skipped and two related genes that encode a new family of zinc finger proteins in *Drosophila*. *Genetics*, *144*(1), 171–182. <https://doi.org/10.1093/genetics/144.1.171>
- Hastings, P. J. (2010). *Mechanisms of Ectopic Gene Conversion*. 427–439. <https://doi.org/10.3390/genes1030427>
- Herndon, N., Shelton, J., Gerischer, L., Ioannidis, P., Ninova, M., Dönitz, J., Waterhouse, R. M., Liang, C., Damm, C., Siemanowski, J., Kitzmann, P., Ulrich, J., Dippel, S., Oberhofer, G., Hu, Y., Schwirz, J., Schacht, M., Lehmann, S., Montino, A., ... Bucher, G. (2020). Enhanced genome assembly and a new official gene set for *Tribolium castaneum*. *BMC Genomics*, *21*(1), 1–13. <https://doi.org/10.1186/s12864-019-6394-6>
- Heyn, P., Kalinka, A. T., Tomancak, P., & Neugebauer, K. M. (2015). Introns and gene expression: Cellular constraints, transcriptional regulation, and evolutionary consequences. *BioEssays*, *37*(2), 148–154. <https://doi.org/10.1002/bies.201400138>
- Holmgren, R. (1984). Cloning sequences from the hairy gene of *Drosophila*. *The EMBO Journal*, *3*(3), 569–573. <https://doi.org/10.1002/j.1460-2075.1984.tb01849.x>

- Hughes, C. L., & Kaufman, T. C. (2002). Exploring myriapod segmentation: The expression patterns of even-skipped, engrailed, and wingless in a centipede. *Developmental Biology*, 247(1), 47–61. <https://doi.org/10.1006/dbio.2002.0683>
- Ikbal, J., Lim, G. S., & Gao, Z. (2015). The hybridization chain reaction in the development of ultrasensitive nucleic acid assays. *TrAC - Trends in Analytical Chemistry*, 64, 86–99. <https://doi.org/10.1016/j.trac.2014.08.014>
- 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(6185), 25–34.
- Ingram, K. K., Kutowoi, A., Wurm, Y., Shoemaker, D. W., Meier, R., & Bloch, G. (2012). The Molecular Clockwork of the Fire Ant *Solenopsis invicta*. *PLoS ONE*, 7(11), 15–19. <https://doi.org/10.1371/journal.pone.0045715>
- Ishibashi, T., Hatori, R., Maeda, R., Nakamura, M., Taguchi, T., Matsuyama, Y., & Matsuno, K. (2019). E and ID proteins regulate cell chirality and left–right asymmetric development in *Drosophila*. *Genes to Cells*, 24(3), 214–230. <https://doi.org/10.1111/GTC.12669>
- Iwaki, D. D., Johansen, K. A., Singer, J. B., & Lengyel, J. A. (2001). Drumstick, bowl, and lines are required for patterning and cell rearrangement in the *Drosophila* embryonic hindgut. *Developmental Biology*, 240(2), 611–626. <https://doi.org/10.1006/dbio.2001.0483>
- Iyer, E. P. R., Iyer, S. C., Sullivan, L., Wang, D., Meduri, R., Graybeal, L. L., & Cox, D. N. (2013). Functional Genomic Analyses of Two Morphologically Distinct Classes of *Drosophila* Sensory Neurons: Post-Mitotic Roles of Transcription Factors in Dendritic Patterning. *PLoS ONE*, 8(8). <https://doi.org/10.1371/journal.pone.0072434>
- Jaeger, J. (2011). The gap gene network. *Cellular and Molecular Life Sciences*, 68(2), 243–274. <https://doi.org/10.1007/s00018-010-0536-y>
- Janssen, R. (2014). Gene expression suggests double-segmental and single-segmental patterning mechanisms during posterior segment addition in the beetle *Tribolium castaneum*. *International Journal of Developmental Biology*, 58(5), 343–347. <https://doi.org/10.1387/ijdb.140058rj>

- Janssen, R., & Budd, G. E. (2013). Deciphering the onychophoran “segmentation gene cascade”: Gene expression reveals limited involvement of pair rule gene orthologs in segmentation, but a highly conserved segment polarity gene network. *Developmental Biology*, 382(1), 224–234. <https://doi.org/10.1016/j.ydbio.2013.07.010>
- Janssen, R., Budd, G. E., Prpic, N. M., & Damen, W. G. M. (2011). Expression of myriapod pair rule gene orthologs. *EvoDevo*, 2(1), 1–15. <https://doi.org/10.1186/2041-9139-2-5>
- 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., Prpic, N. M., Kosiol, C., Vervoort, M., Damen, W. G., Balavoine, G., & McGregor, A. P. (2010). Conservation, loss, and redeployment of Wnt ligands in protostomes: Implications for understanding the evolution of segment formation. *BMC Evolutionary Biology*, 10(1), 1–21. <https://doi.org/10.1186/1471-2148-10-374>
- Jeon, H., Gim, S., Na, H., & Choe, C. P. (2020). A pair-rule function of odd-skipped in germband stages of *Tribolium* development. *Developmental Biology*, 465(1), 58–65. <https://doi.org/10.1016/J.YDBIO.2020.07.002>
- Johnston, D. S., & Nüsslein-Volhard, C. (1992). The origin of pattern and polarity in the *Drosophila* embryo. *Cell*, 68(2), 201–219. [https://doi.org/10.1016/0092-8674\(92\)90466-P](https://doi.org/10.1016/0092-8674(92)90466-P)
- Kadener, S., Stoleru, D., McDonald, M., Nawathean, P., & Rosbash, M. (2007). *Clockwork Orange is a transcriptional repressor and a new Drosophila circadian pacemaker component*. <https://doi.org/10.1101/gad.1552607>
- Kageyama, R., Ohtsuka, T., & Kobayashi, T. (2007). The Hes gene family: repressors and oscillators that orchestrate embryogenesis. *Development*. <https://doi.org/10.1242/dev.000786>
- Kageyama, R., Niwa, Y., Isomura, A., González, A., & Harima, Y. (2012). Oscillatory gene expression and somitogenesis. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1(5), 629–641. <https://doi.org/10.1002/wdev.46>

- Kainz, F., Ewen-Campen, B., Akam, M., & Extavour, C. G. (2011). Notch/Delta signalling is not required for segment generation in the basally branching insect *Gryllus bimaculatus*. *Development*, 138(22), 5015–5026. <https://doi.org/10.1242/dev.073395>
- Kim, I. M., Wolf, M. J., & Rockman, H. A. (2010). Gene deletion screen for cardiomyopathy in adult drosophila identifies a new notch ligand. *Circulation Research*, 106(7), 1233–1243. <https://doi.org/10.1161/CIRCRESAHA.109.213785>
- King, B., & Denholm, B. (2014). Malpighian tubule development in the red flour beetle (*Tribolium castaneum*). *Arthropod Structure and Development*, 43(6), 605–613. <https://doi.org/10.1016/j.asd.2014.08.002>
- Klingler, M., & Bucher, G. (2022). The red flour beetle *T. castaneum*: elaborate genetic toolkit and unbiased large scale RNAi screening to study insect biology and evolution. *EvoDevo*, 13(1), 1–11. <https://doi.org/10.1186/s13227-022-00201-9>
- Kosman, D., Mizutani, C. M., Lemons, D., Cox, W. G., McGinnis, W., & Bier, E. (2004). Multiplex detection of RNA expression in *Drosophila* embryos. *Science*, 305(5685), 846. <https://doi.org/10.1126/science.1099247>
- Kovall, R.A., Gebelein, B., Sprinzak, D., Kopan, R. (2017). The Canonical Notch Signaling Pathway: Structural and Biochemical Insights into Shape, Sugar, and Force. *Dev Cell*, 8;41(3), 228-241. doi: 10.1016/j.devcel.2017.04.001.
- Krol, A. J., Roellig, D., Dequeant, M.-L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A. C., & Pourquie, O. (2011). Evolutionary plasticity of segmentation clock networks. *Development*, 138(13), 2783–2792. <https://doi.org/10.1242/dev.063834>
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Kux, K., Kiparaki, M., & Delidakis, C. (2013). The two *Tribolium* E(spl) genes show evolutionarily conserved expression and function during embryonic neurogenesis. *Mechanisms of Development*, 130(4–5), 207–225. <https://doi.org/10.1016/j.mod.2013.02.003>

- Layden, M. J., & Martindale, M. Q. (2014). *Non-canonical Notch signaling represents an ancestral mechanism to regulate neural differentiation*. <https://doi.org/10.1186/2041-9139-5-30>
- Leimeister, C., Dale, K., Fischer, A., Klamt, B., Hrabe De Angelis, M., Radtke, F., McGrew, M. J., Pourquié, O., & Gessler, M. (2000). Oscillating expression of c-Hey2 in the presomitic mesoderm suggests that the segmentation clock may use combinatorial signaling through multiple interacting bHLH factors. *Developmental Biology*, 227(1), 91–103. <https://doi.org/10.1006/dbio.2000.9884>
- Liao, B.-K., & Oates, A. C. (2017). Delta-Notch signalling in segmentation. *Arthropod Structure and Development*, 46, 429–447. <https://doi.org/10.1016/j.asd.2016.11.007>
- Liu, W. (2013). Bmdelta phenotype implies involvement of Notch signaling in body segmentation and appendage development of silkworm, *Bombyx mori*. *Arthropod Structure and Development*, 42(2), 143–151. <https://doi.org/10.1016/j.asd.2012.10.002>
- Livingstone, C. D., & Barton, G. J. (1993). Protein sequence alignments: A strategy for the hierarchical analysis of residue conservation. *Bioinformatics*, 9(6), 745–756. <https://doi.org/10.1093/bioinformatics/9.6.745>
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J., & Marchler-Bauer, A. (2020). CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research*, 48(D1), D265–D268. <https://doi.org/10.1093/nar/gkz991>
- Matsumoto, A., Ukai-Tadenuma, M., Yamada, R. G., Houl, J., Uno, K. D., Kasukawa, T., Dauwalder, B., Itoh, T. Q., Takahashi, K., Ueda, R., Hardin, P. E., Tanimura, T., & Ueda, H. R. (2007). A functional genomics strategy reveals clockwork orange as a transcriptional regulator in the *Drosophila* circadian clock. *Genes and Development*, 21(13), 1687–1700. <https://doi.org/10.1101/gad.1552207>
- Mendivil Ramos, O., & Ferrier, D. E. K. (2012). Mechanisms of Gene Duplication and Translocation and Progress towards Understanding Their Relative Contributions to Animal Genome Evolution. *International Journal of Evolutionary Biology*, 2012, 1–10. <https://doi.org/10.1155/2012/846421>

- Miller, S. C., Miyata, K., Brown, S. J., & Tomoyasu, Y. (2012). Dissecting Systemic RNA Interference in the Red Flour Beetle *Tribolium castaneum*: Parameters Affecting the Efficiency of RNAi. *PLoS ONE*, 7(10). <https://doi.org/10.1371/journal.pone.0047431>
- Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., et.al. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210), 763-7. doi: 10.1126/science.1257570.
- Mito, T., Shinmyo, Y., Kurita, K., Nakamura, T., Ohuchi, H., & Noji, S. (2011). Ancestral functions of Delta/Notch signaling in the formation of body and leg segments in the cricket *Gryllus bimaculatus*. *Development*, 138(17), 3823–3833. <https://doi.org/10.1242/dev.060681>
- Mito, T., Kobayashi, C., Sarashina, I., Zhang, H., Shinahara, W., Miyawaki, K., Shinmyo, Y., Ohuchi, H., & 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(1), 202–213. <https://doi.org/10.1016/j.ydbio.2006.11.003>
- Monastirioti, M., Giagtzoglou, N., Koumbanakis, K. A., Zacharioudaki, E., Deligiannaki, M., Wech, I., Almeida, M., Preiss, A., Bray, S., & Delidakis, C. (2010). *Drosophila* Hey is a target of Notch in asymmetric divisions during embryonic and larval neurogenesis. *Development*, 137(2), 191–201. <https://doi.org/10.1242/dev.043604>
- Mundaca-Escobar, M., Cepeda, R. E., & Sarrazin, A. F. (2022). The organizing role of Wnt signaling pathway during arthropod posterior growth. *Frontiers in Cell and Developmental Biology*, 10(August), 1–14. <https://doi.org/10.3389/fcell.2022.944673>
- Nagy, L. M., & Carroll, S. (1994). *short-germ embryos of Tribolium castaneum*. 367(February), 1–4.
- Nakamoto, A., Hester, S. D., Constantinou, S. J., Blaine, W. G., Tewksbury, A. B., Matei, M. T., Nagy, L. M., & Williams, T. A. (2015). ARTICLE Changing cell behaviours during beetle embryogenesis correlates with slowing of segmentation. *Nature Communications*. <https://doi.org/10.1038/ncomms7635>

- Nakao, H. (2015). Analyses of interactions among pair-rule genes and the gap gene Krüppel in *Bombyx* segmentation. *Developmental Biology*, 405(1), 149–157.
<https://doi.org/10.1016/j.ydbio.2015.06.012>
- Nasiadka, A., Dietrich, B. H. and Krause, H. M. (2002). Anterior-posterior patterning in the *Drosophila* embryo. In *Advances in Developmental Biology and Biochemistry, Gene Expression at the Beginning of Animal Development* (ed. M. L. DePamphilis). *Elsevier*, pp. 155-204.
- Noonan, J. P., Grimwood, J., Schmutz, J., Dickson, M., & Myers, R. M. (2004). Gene conversion and the evolution of protocadherin gene cluster diversity. *Genome Research*, 14(3), 354–366. <https://doi.org/10.1101/gr.2133704>
- Nüsslein-Volhard, C., & Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature*, 287(5785), 795–801. <https://doi.org/10.1038/287795a0>
- Oates, A. C., Morelli, L. G., & Ares, S. (2012). Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. *Development*, 139(4), 625–639. <https://doi.org/10.1242/dev.063735>
- Oberhofer, G., Grossmann, D., Siemanowski, J. L., Beissbarth, T., & Bucher, G. (2014). Wnt/ β -catenin signaling integrates patterning and metabolism of the insect growth zone. *Development (Cambridge)*, 141(24), 4740–4750. <https://doi.org/10.1242/dev.112797>
- Paese, C. L. B. (2018). Investigating the roles of Hes and Sox genes during embryogenesis of the spider *Parasteatoda tepidariorum*. *PhD thesis*, Oxford Brookes University
- Park, P. J. (2009). ChIP-seq: Advantages and challenges of a maturing technology. *Nature Reviews Genetics*, 10(10), 669–680. <https://doi.org/10.1038/nrg2641>
- Patel, N. H., Condrón, B. G., & Zinn, K. (1994). Pair-rule expression patterns of even-skipped are found in both short- and long-germ beetles. *Nature*, 367(6462), 429–434.
<https://doi.org/10.1038/367429a0>
- Peel, A. D., Chipman, A. D., & Akam, M. (2005). Arthropod segmentation: Beyond the *Drosophila* paradigm. *Nature Reviews Genetics*, 6(12), 905–916.
<https://doi.org/10.1038/nrg1724>

- Peel, A. D., Telford, M. J., & Akam, M. (n.d.). *The evolution of hexapod engrailed-family genes: evidence for conservation and concerted evolution*.
<https://doi.org/10.1098/rspb.2006.3497>
- Pointer, M. D., Gage, M. J. G., & Spurgin, L. G. (2021). Tribolium beetles as a model system in evolution and ecology. *Heredity*, 126(6), 869–883. <https://doi.org/10.1038/s41437-021-00420-1>
- Posnien, N., Schinko, J., Grossmann, D., Shippy, T. D., Konopova, B., & Bucher, G. (2009). RNAi in the red flour beetle (Tribolium). *Cold Spring Harbor Protocols*, 4(8), 1–8.
<https://doi.org/10.1101/pdb.prot5256>
- Prpic, N. M., & Damen, W. G. M. (2009). Notch-mediated segmentation of the appendages is a molecular phylotypic trait of the arthropods. *Developmental Biology*, 326(1), 262–271.
<https://doi.org/10.1016/j.ydbio.2008.10.049>
- Pueyo, J. I., Lanfear, R., & Couso, J. P. (2008). Ancestral Notch-mediated segmentation revealed in the cockroach *Periplaneta americana*. *Proceedings of the National Academy of Sciences*, 105(43), 16614–16619. <https://doi.org/10.1073/pnas.0804093105>
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901–904. <https://doi.org/10.1093/sysbio/syy032>
- Rauskolb, C., & Irvine, K. D. (1999). Notch-mediated segmentation and growth control of the *Drosophila* leg. *Developmental Biology*, 210(2), 339–350.
<https://doi.org/10.1006/dbio.1999.9273>
- Reams, A. B., & Roth, J. R. (2015). Mechanisms of gene duplication and amplification. *Cold Spring Harbor Perspectives in Biology*, 7(2).
<https://doi.org/10.1101/cshperspect.a016592>
- Reding, K., Chen, M., Lu, Y., Cheatle Jarvela, A. M., & Pick, L. (2019). Shifting roles of *Drosophila* pair-rule gene orthologs: Segmental expression and function in the milkweed bug *Oncopeltus fasciatus*. *Development (Cambridge)*, 146(17), 1–11.
<https://doi.org/10.1242/dev.181453>

- Resende, T. P., Andrade, R. P., & Palmeirim, I. (2014). Timing embryo segmentation: Dynamics and regulatory mechanisms of the vertebrate segmentation clock. *BioMed Research International*, 2014. <https://doi.org/10.1155/2014/718683>
- Reynaud-Deonauth, S., Zhang, H., Afouda, A., Taillefert, S., Beatus, P., Kloc, M., Etkin, L. D., Fischer-Lougheed, J., & Spohr, G. (2002). Notch signaling is involved in the regulation of Id3 gene transcription during *Xenopus* embryogenesis. *Differentiation*, 69(4–5), 198–208. <https://doi.org/10.1046/j.1432-0436.2002.690413.x>
- Rosenberg, M. I., Brent, A. E., Payre, F., & Desplan, C. (2014). Dual mode of embryonic development is highlighted by expression and function of *Nasonia* pair-rule genes. *eLife*, 3. <https://doi.org/10.7554/eLife.01440>
- Ryan, J. F., & Baxevas, A. D. (2007). Hox, Wnt, and the evolution of the primary body axis: Insights from the early-divergent phyla. *Biology Direct*, 2. <https://doi.org/10.1186/1745-6150-2-37>
- Saito, M., Awasaki, T., & Hama, C. (2002). Genetic analyses of essential genes in cytological region 61D1-2 to 61F1-2 of *Drosophila melanogaster*. *Molecular Genetics and Genomics*, 268(4), 446–454. <https://doi.org/10.1007/s00438-002-0770-6>
- Sarrazin, A. F., Peel, A. D., & Averof, M. (2012). A segmentation clock with two-segment periodicity in insects. *Science*, 336(6079), 338–341. <https://doi.org/10.1126/science.1218256>
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, 9(7), 676–682. <https://doi.org/10.1038/nmeth.2019>
- Schinko, J., Posnien, N., Kittelmann, S., Koniszewski, N., & Bucher, G. (2009). Single and double whole-mount in situ hybridization in red flour beetle (*Tribolium*) embryos. *Cold Spring Harbor Protocols*, 4(8). <https://doi.org/10.1101/pdb.prot5258>
- Schönauer, A., Paese, C. L. B., Hilbrant, M., Leite, D. J., Schwager, E. E., Martins Feitosa, N., Eibner, C., Damen, W. G. M., & Mcgregor, A. P. (2016). *The Wnt and Delta-Notch signalling pathways interact to direct pair-rule gene expression via caudal during*

segment addition in the spider Parasteatoda tepidariorum.

<https://doi.org/10.1242/dev.131656>

Schröder, R., Eckert, C., Wolff, C., & Tautz, D. (2000). Conserved and divergent aspects of terminal patterning in the beetle *Tribolium castaneum*. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12), 6591–6596.
<https://doi.org/10.1073/pnas.100005497>

Schroeder, M. D., Greer, C., & Gaul, U. (2011). How to make stripes: deciphering the transition from non-periodic to periodic patterns in *Drosophila* segmentation. *Development*, 138(14), 3067–3078. <https://doi.org/10.1242/dev.062141>

Schröter, C., Ares, S., Morelli, L. G., Isakova, A., Hens, K., Soroldoni, D., Gajewski, M., Jülicher, F., Maerkl, S. J., Deplancke, B., & Oates, A. C. (2012). Topology and dynamics of the zebrafish segmentation clock core circuit. *PLoS Biology*, 10(7), 11.
<https://doi.org/10.1371/journal.pbio.1001364>

Schwager, E. (2008). Segmentation of the spider *Achaearanea tepidariorum* investigated by gene expression and functional analysis of the gap gene hunchback. *PhD thesis*, Universität zu Köln

Shah, S., Lubeck, E., Schwarzkopf, M., He, T. F., Greenbaum, A., Sohn, C. H., Lignell, A., Choi, H. M. T., Gradinaru, V., Pierce, N. A., & Cai, L. (2016). Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. *Development (Cambridge)*, 143(15), 2862–2867. <https://doi.org/10.1242/dev.138560>

Shimojo, H., Ohtsuka, T., & Kageyama, R. (n.d.). *Article Oscillations in Notch Signaling Regulate Maintenance of Neural Progenitors.*
<https://doi.org/10.1016/j.neuron.2008.02.014>

Siemanowski, J., Richter, T., Dao, V. A., & Bucher, G. (2015). Notch signaling induces cell proliferation in the labrum in a regulatory network different from the thoracic legs. *Developmental Biology*, 408(1), 164–177. <https://doi.org/10.1016/j.ydbio.2015.09.018>

Sommer, R. J., & 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(6411), 448–450. <https://doi.org/10.1038/361448a0>

- Southall, T. D., & Brand, A. H. (2009). Neural stem cell transcriptional networks highlight genes essential for nervous system development. *EMBO Journal*, 28(24), 3799–3807. <https://doi.org/10.1038/emboj.2009.309>
- Stahi, R., & Chipman, A. D. (2016). Blastoderm segmentation in oncopeltus fasciatus and the evolution of insect segmentation mechanisms. *Proceedings of the Royal Society B: Biological Sciences*, 283(1840). <https://doi.org/10.1098/rspb.2016.1745>
- Stern, D. L., & Sucena, E. (2012). Rapid mounting of adult Drosophila structures in hoyer's medium. *Cold Spring Harbor Protocols*, 7(1), 107–109. <https://doi.org/10.1101/pdb.prot067371>
- Stollewark, A., Schoppmeier, M., & Damen, W. G. M. (2003). Involvement of Notch and Delta genes in spider segmentation. *Nature*. <https://doi.org/10.1038/nature01682>
- Stoltzfus, A. (1999). On the possibility of constructive neutral evolution. *Journal of Molecular Evolution*, 49(2), 169–181. <https://doi.org/10.1007/PL00006540>
- Svoboda, P. (2020). Key Mechanistic Principles and Considerations Concerning RNA Interference. *Frontiers in Plant Science*, 11(August), 1–13. <https://doi.org/10.3389/fpls.2020.01237>
- Tautz, D. (2004). Segmentation. *Dev Cell*, 7(3), 301–312. <https://doi.org/10.1016/j.devcel.2004.08.008>
- Teshima, K. M., & Innan, H. (2004). The Effect of Gene Conversion on the Divergence between Duplicated Genes. *Genetics*, 166(3), 1553–1560. <https://doi.org/10.1534/genetics.166.3.1553>
- Thompson, J. D., Linard, B., Lecompte, O., & Poch, O. (2011). A comprehensive benchmark study of multiple sequence alignment methods: Current challenges and future perspectives. *PLoS ONE*, 6(3). <https://doi.org/10.1371/journal.pone.0018093>
- Thümecke, S., & Schröder, R. (2021). The odd-skipped related gene drumstick is required for leg development in the beetle Tribolium castaneum. *Developmental Dynamics*. <https://doi.org/10.1002/dvdy.347>

- Tidswell, O.R.A. (2021). Investigating the roles of candidate gap genes from the neuroblast timer series during axial patterning in the beetle *Tribolium*. *PhD thesis*, University of Cambridge
- Tidswell, O., Benton, M. A., & Akam, M. E. (2021). *Long title The neuroblast timer gene nubbin exhibits functional redundancy with gap genes to regulate segment identity in Tribolium*.
- Tomancak, P., Berman, B. P., Beaton, A., Weiszmamm, R., Kwan, E., Hartenstein, V., Celniker, S. E., & Rubin, G. M. (2007). Global analysis of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biology*, 8(7), 1–24. <https://doi.org/10.1186/gb-2007-8-7-r145>
- Tomiyama, Y., Shinohara, T., Matsuka, M., Bando, T., Mito, T., & Tomioka, K. (2020). The role of clockwork orange in the circadian clock of the cricket *Gryllus bimaculatus*. *Zoological Letters*, 6(1), 1–15. <https://doi.org/10.1186/s40851-020-00166-4>
- True, J. R. and Haag, E. S. (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evolution and Development*, 3(2), 109–119. <https://doi.org/10.1046/j.1525-142X.2001.003002109.x>
- Ungerer, P., Eriksson, B. J., & Stollewerk, A. (2012). Unravelling the evolution of neural stem cells in arthropods: Notch signalling in neural stem cell development in the crustacean *Daphnia magna*. *Developmental Biology*, 371(2), 302–311. <https://doi.org/10.1016/j.ydbio.2012.08.025>
- Valentin, G., & Oates, A. C. (2013). *Opening a can of centipedes: new insights into mechanisms of body segmentation*. <http://www.biomedcentral.com/1741-7007/11/112>
- Van Doren, M., Ellis, H. M., & Posakony, J. W. (1991). The *Drosophila* extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. *Development*, 113(1), 245–255. <https://doi.org/10.1242/dev.113.1.245>
- Wang, L., & Coulter, D. E. (1996). Bowel, an odd-skipped homolog, functions in the terminal pathway during *Drosophila* embryogenesis. *EMBO Journal*, 15(12), 3182–3196. <https://doi.org/10.1002/j.1460-2075.1996.tb00681.x>

- Wheeler, S. R., Carrico, M. L., Wilson, B. A., Brown, S. J., & Skeath, J. B. (2003). The expression and function of the achaete-scute genes in *Tribolium castaneum* reveals conservation and variation in neural pattern formation and cell fate specification. *Development*, 130(18), 4373–4381. <https://doi.org/10.1242/dev.00646>
- Wheeler, S. R., Carrico, M. L., Wilson, B. A., & Skeath, J. B. (2005). The *Tribolium* columnar genes reveal conservation and plasticity in neural precursor patterning along the embryonic dorsal-ventral axis. *Developmental Biology*, 279(2), 491–500.
- Williams, T. A., & Nagy, L. M. (2017). Linking gene regulation to cell behaviors in the posterior growth zone of sequentially segmenting arthropods. *Arthropod Structure and Development*, 46(3), 380–394. <https://doi.org/10.1016/j.asd.2016.10.003>
- Wilson, M. J., & Dearden, P. K. (2012). Pair-Rule Gene Orthologues Have Unexpected Maternal Roles in the Honeybee (*Apis mellifera*). *PLoS ONE*, 7(9), e46490. <https://doi.org/10.1371/journal.pone.0046490>
- Wilson, M. J., Mckelvey, B. H., Van Der Heide, S., & Dearden, P. K. (n.d.). *Notch signaling does not regulate segmentation in the honeybee, Apis mellifera*. <https://doi.org/10.1007/s00427-010-0340-6>
- Xiang, J., Reding, K., Heffer, A., & Pick, L. (2017). Conservation and variation in pair-rule gene expression and function in the intermediate-germ beetle *dermestes maculatus*. *Development (Cambridge)*, 144(24), 4625–4636. <https://doi.org/10.1242/dev.154039>
- Zhang, J., Chen, H., Li, R., Taft, D. A., Yao, G., Bai, F., & Xing, J. (2019). Spatial clustering and common regulatory elements correlate with coordinated gene expression. *PLoS Computational Biology*, 15(3), 1–16. <https://doi.org/10.1371/journal.pcbi.1006786>
- Zhou, B., Lin, W., Long, Y., Yang, Y., Zhang, H., Wu, K., & Chu, Q. (2022). Notch signaling pathway: architecture, disease, and therapeutics. *Signal Transduction and Targeted Therapy*, 7(1), 1–33. <https://doi.org/10.1038/s41392-022-00934-y>
- Zhou, M., Yan, J., Ma, Z., Zhou, Y., Abbood, N. N., Liu, J., Su, L., Jia, H., & Guo, A. Y. (2012). Comparative and evolutionary analysis of the HES/HEY gene family reveal exon/intron loss and teleost specific duplication events. *PLoS ONE*, 7(7). <https://doi.org/10.1371/journal.pone.0040649>