

Characterizing new players in regulating the production of the steroid hormone ecdysone during larval development of *Drosophila melanogaster*.

by

Jie Zeng

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Molecular Biology and Genetics

Department of Biological Sciences
University of Alberta

© Jie Zeng, 2017

Abstract

The major steroid hormone in *Drosophila* is Ecdysone. This hormone triggers developmental transitions such as the molts and the onset of metamorphosis. During the second half of the last (i.e. 3rd) instar, ecdysone biosynthesis is upregulated in the prothoracic gland (PG). This results in a major pulse of ecdysone that will trigger the onset of metamorphosis. In *Drosophila*, the PG is part of the ring gland, the principal neuroendocrine organ in larvae, and is the site of synthesis of a range of insect hormones, including ecdysone. In a search for novel regulators of ecdysone production, the King-Jones lab carried out ring gland-specific microarrays and identified 108 genes that are specifically expressed in this tissue. Surprisingly, the *snail* and *curled* genes were among those identified. The known roles for *snail* were previously associated with embryonic development, while *curled* had been linked to circadian-dependent RNA degradation. I chose to study these two genes in further detail.

PG-specific disruption of *snail* via RNA interference (RNAi) resulted in larval arrest, a phenotype often caused by ecdysone deficiency. PG-specific RNAi of *curled*, on the other hand, caused developmental acceleration, which often results from precocious ecdysone pulses. These results suggested that *snail* and *curled* are novel players in the regulation of ecdysone production and my work focused on characterizing the molecular mechanisms underlying their functions in the PG.

Immunofluorescent staining showed that the Snail protein is present only in a subset of PG nuclei at any given time, which resembled the pattern of PG S-phase cells when visualized by incorporation of 5-ethynyl-2'-deoxyuridine (EdU), a nucleotide analog. The PG undergoes an alternative form of cell cycle called endocycle or endoreplication where cells have only alternating S and G phase without cell division and the endocycle is unsynchronized amongst PG cells. I observed two waves of endocycle progression in the PG, namely one at 17-18 hr in the 2nd instar and one at 10-12 hr in the 3rd instar (L3), which correlated well with two peaks of Snail-positive cells in the PG. A recent study by Ohhara *at al.* (2016) showed that the endocycle progression at 10-12 hr L3 is tightly coupled with the time window of critical weight attainment (CW), a developmental checkpoint that, once bypassed, the animals' commitment to metamorphosis is no longer affected by nutrient conditions. The exact molecular mechanism of CW

attainment remains unclear. However, the recent study showed that nutrient-dependent endoreplication in the PG might be part of the molecular basis of CW attainment. My results demonstrated that with both *sna*-RNAi and *sna* overexpression, the endocycle in the PG was arrested during the time window of the CW checkpoint and the animals failed to pupariate, suggesting that larvae did not receive the appropriate signal for passing the CW checkpoint. Moreover, I showed that Snail levels in the PG are responsive to the nutrient sensor TOR, as well as starvation, suggesting that Snail coordinates nutrient-dependent endoreplication, CW checkpoint and ecdysone production in the PG.

The developmental acceleration that I observed in PG>*curled*-RNAi animals appears to phenocopy *Drosophila* Hormone Receptor 4 (*DHR4*) mutants, which also develop faster than controls. DHR4 is a nuclear receptor that periodically shuttles between cytoplasm and nucleus, and is believed to transcriptionally repress ecdysone biosynthesis when it is in the nucleus. I showed that the function of *DHR4* is genetically dependent on *curled*, raising the possibility that Curled assists nuclear entry of DHR4. Interestingly, a similar system appears to be in place in vertebrates, where entry into the nucleus of the nuclear receptor PPAR γ is dependent on the Curled ortholog Nocturnin. Moreover, Curled/Nocturnin is predicted to function as a deadenylase as part of the CCR4-NOT complex, one of the conserved complexes that shorten the mRNA poly (A)-tail. However, RNAi of several other CCR4-NOT components in the PG caused ecdysone deficiency, a different phenotype from what I observed in *curled*-RNAi animals, suggesting Curled works independently of the CCR4-NOT complex in the PG. My study broadens our current understanding of how ecdysteroidogenesis is regulated, and describes the CCR4-NOT complex as a novel regulator required for ecdysone production in the *Drosophila* PG.

Preface

This thesis is an original work by Jie Zeng. Part of the results discussed in Chapter 3.3.1 and Chapter 3.3.3 have been previously published (Ou, Q *et al.* 2016. The Insect Prothoracic Gland as a Model for Steroid Hormone Biosynthesis and Regulation. *Cell Reports* 16, 247-262. Figure 4D-4G). From this above-mentioned article, only the experiments carried out by Jie Zeng are displayed in this thesis. Previously published results by Jie Zeng have been rewritten for this thesis. The related figures were also reconstructed for this thesis; therefore no figures from the *Journal* were used.

Acknowledgements

I would like to thank many people who have provided me with invaluable assistance during my PhD program. I owe the deepest gratitude to my supervisor, Dr. Kirst King-Jones who gave me the opportunity to join his research group. Throughout the years, Kirst has always been patient and encouraging. Thank you so much for all your guidance, invaluable discussions and constructive ideas on my research projects. Your pursuit of precision and perfection has helped me become a better researcher during my academic pursuit. I would also like to express my gratitude to my committee members Dr. Frank Nargang and Dr. Sarah Hughes for always being supportive and encouraging. Thank you so much for all your good suggestions and insights on my PhD project.

I am thankful for the help and friendship from all the past and present members of the King-Jones lab. I want to give special thanks to Dr. Ran Zhuo and Dr. Adam Magico who helped me learn all the basic techniques when I first joined the lab. Also special thanks to Dr. Ran Zhuo and Dr. Francesca di Cara for all the stimulating discussions as well as being wonderful listeners to my concerns about my work and my personal life. For my thesis, I would like to express my gratitude to Brian Phelps for his great patience, useful comments and valuable suggestions. For the Snail project, special thanks to Dr. Naoki Yamanaka and Dr. Yuya Ohhara for technical assistance and exchanging ideas. Last but not least, I owe a special thank-you to my family, my parents, my husband and my aunt. Without your unconditional love, support and understanding, this thesis would have been impossible.

Table of Contents

1. General Introduction	1
1.1 Ecdysone, a major steroid hormone in <i>Drosophila melanogaster</i> , controls developmental transitions	2
1.2 Neuroendocrine control of ecdysone production	2
1.3 The ecdysone biosynthetic enzymes and their coding genes	3
1.4 The PTTH/MAPK cascade regulates ecdysone production	3
1.5 Prothoracic gland, a CPU-like decision-making center for developmental progression	4
1.6 IIS/TOR signaling and its function in the PG	5
1.7 Other factors known to regulate the expression of ecdysone biosynthetic genes in the PG	6
1.8 Cellular transport and release of ecdysone.....	6
1.9 Feedback regulation of ecdysone production.....	7
1.10 Critical weight coordinates body size and timing of metamorphosis	8
1.11 The prothoracic gland is an endoreplicating tissue	9
1.12 Regulation of endocycle in <i>Drosophila melanogaster</i>	9
1.13 Identifying novel regulators for ecdysone production via RG-specific microarrays and PG-specific RNAi screen	10
1.14 Figures.....	13
Chapter 2. Snail coordinates nutrient-dependent endoreplication with ecdysone production in the prothoracic gland	18
2.1 Introduction	19
2.1.1 Snail family proteins	19
2.1.2 Protein structures and structural motifs.....	19
2.1.3 Cellular functions in development and pathology.....	20
2.1.4 <i>Drosophila sna</i> is essential for embryonic ring gland development	21
2.2 Materials and Methods.....	21
2.2.1 Fly stocks and fly crosses.....	21
2.2.2 Rescue by feeding ecdysone-supplemented diets	22
2.2.3 RNA extraction from dissected tissues and cDNA synthesis.....	22
2.2.4 Quantitative Real-time PCR (qPCR) analysis.....	22
2.2.5 20E titer measurement.....	25
2.2.6 Immunofluorescence with larval tissues	25
2.2.7 Cellular DNA content measurements.....	26
2.2.8 CRISPR-based tissue-specific <i>sna</i> deletion	26
2.2.9 Genomic extraction of ring glands and sequencing	27

2.2.10 Animal staging	28
2.2.11 Starvation protocol for critical weight determination	28
2.2.12 EdU incorporation assays.....	29
2.2.13 Generation of <i>sna</i> -overexpressing clones by FLP-out system	29
2.2.14 Heat shock induction of <i>sna</i> overexpression.....	29
2.2.15 Next-generation RNA sequencing (RNA-Seq) analysis	30
2.3 Results	33
2.3.1 <i>Snail (sna)</i> is specifically and dynamically expressed in larval ring glands	33
2.3.2 Loss-of- <i>sna</i> in the PG affected ecdysone production.....	34
2.3.3 PG> <i>sna</i> -RNAi animals showed endocycle arrest in the PG	38
2.3.4 PG-specific deletion of <i>sna</i> caused pupal lethality and affected the morphology of ring glands	41
2.3.5 Endocycle in the PG was arrested at the critical weight checkpoint when <i>sna</i> was disrupted	43
2.3.6 PG> <i>sna</i> -RNAi phenotype could not be rescued by promoting S-phase entry	45
2.3.7 TOR functions upstream of <i>sna</i> in the PG	46
2.3.8 Sna protein is probably present during G phase of the endocycle	51
2.3.9 Sna protein stability may be regulated by α -ecdysone	52
2.3.10 Overexpressing <i>sna</i> family proteins in the PG caused endocycle arrest in a cell-autonomous manner	53
2.3.11 Identifying Snail target genes in the PG via Next Generation RNA-Seq.....	57
2.3.12 Transcription of <i>ouija board (ouib)</i> is controlled by IIS/TOR signaling in the PG	78
2.3.13 The cell adhesion molecule N-Cadherin is present in the CA of the ring gland.	78
2.4 Discussion and future directions	78
2.4.1 Snail function in the PG is essential for ecdysone biosynthesis.....	78
2.4.2 Snail might directly regulate the transcription of ecdysone biosynthetic genes	79
2.4.3 Sna as a potential regulator of endoreplication in the PG	82
2.4.4 Snail as a candidate of the molecular basis for critical weight checkpoint.....	85
2.4.5 Posttranscriptional regulation of Sna in the PG cells.....	87
2.5 Conclusions and significance	88
2.6 Figures.....	89
Chapter 3. The deadenylase Curled (<i>aka</i> Nocturnin) and the CCR4-NOT deadenylase complex have distinct roles in regulating the ecdysone production in the prothoracic gland	143
3.1 Introduction	144
3.1.1 <i>Curled</i> , a <i>Drosophila</i> homolog of <i>Nocturnin</i>	144
3.1.2 CCR4-NOT deadenylase complex and its components	144
3.1.3 Deadenylase and transcriptional repression	145
3.1.4 The CCR4-NOT deadenylase complex is multifunctional.....	146
3.1.5 DHR4 and its role in regulating the ecdysone synthesis	146

3.2 Materials and Methods	147
3.2.1 Fly stocks and fly crosses.....	147
3.2.2 Constructing an RNA interference-resistant form of <i>cu</i> -cDNA.....	147
3.2.3 Measurement of pupal body mass.....	147
3.2.4 Developmental timing measurement.....	148
3.2.5 qPCR analysis	148
3.2.6 RNA-Seq analysis	149
3.2.7 Sterol rescue by feeding.....	149
3.3 Results and Discussion.....	149
3.3.1 <i>Curled (cu)</i> has a role in the PG during larvae-pupae transition.....	149
3.3.2 Overexpression of <i>cu</i> in the PG caused developmental delay.....	150
3.3.3 The genetic interaction between <i>cu</i> and <i>DHR4</i>	151
3.3.4 The CCR4-NOT complex is essential for ecdysone production	154
3.3.5 RNA-Seq analysis of PG> <i>Pop2</i> -RNAi	155
3.4 Conclusion and significance	162
3.5 Figures.....	163
References	180
Appendix	196

List of figures

Figure 1-1. Schematic representation of whole body ecdysteroid titer throughout the <i>Drosophila</i> life cycle.....	13
Figure 1-2. Ecdysone is produced in the prothoracic gland, a part of an endocrine tissue, the ring gland 14	
Figure 1-3. Known signaling pathways that are essential for ecdysone biosynthesis in the PG.....	15
Figure 1-4. A simplified demonstration of the interconnected IIS/TOR signaling pathway in <i>Drosophila</i>	16
Figure 1-5. Ubiquitous knock down of <i>Oatp74D</i> resulted in 3rd instar lethality	17
Figure 2-1. Schematic demonstration of the main functional domains in four Snail family proteins	89
Figure 2-2. <i>sna</i> is specifically and dynamically expressed in the larval PG	90
Figure 2-3. Disruption of <i>sna</i> function in the PG caused developmental arrest	92
Figure 2-4. <i>sna</i> functions in the larval PG are essential for developmental progression	93
Figure 2-5. Loss-of- <i>sna</i> function in the PG via RNAi affected ecdysone biosynthesis.....	94
Figure 2-6. Developmental arrest caused by PG> <i>sna</i> -RNAi was partially rescued by ecdysone feeding. 95	
Figure 2-7. <i>sna</i> does not function through the PTTH/MAPK signaling pathway in the PG	96
Figure 2-8. <i>Sna</i> levels in the PG are not controlled by PTTH/MAPK signaling	97
Figure 2-9. <i>sna</i> mRNA levels are not dependent on Torso (the receptor for PTTH).....	98
Figure 2-10. PG-specific <i>sna</i> -RNAi affected the size of the PG nuclei as well as the PG cell number	99
Figure 2-11. The GeneSwitch/UAS expression system in <i>Drosophila</i>	100
Figure 2-12. <i>sna</i> -RNAi in the larval PG resulted in small PG nuclei	101
Figure 2-13. Endocycle progression in the PG requires <i>sna</i> function.....	102
Figure 2-14. The <i>Drosophila sna</i> locus organization, target sites of all the existing <i>sna</i> -RNAi lines as well as gRNA target sites for CRISPR-based tissue-specific <i>sna</i> deletion	103
Figure 2-15. Larval tissues where GFP is expressed via the <i>Mai60-Gal4</i> driver	104
Figure 2-16. Developmental phenotypes caused by CRISPR-based PG-specific <i>sna</i> mutations	105
Figure 2-17. PG specific deletion of <i>sna</i> affected the morphology of the ring glands	106
Figure 2-18. Determining the time of critical weight attainment.....	107
Figure 2-19. Endocycle progression was arrested at the time of critical weight checkpoint in PG> <i>sna</i> -RNAi PGs.....	108
Figure 2-20. PG-specific expression of <i>CycE</i> can not rescue the developmental arrest caused by PG> <i>sna</i> -RNAi	109
Figure 2-21. The effect of <i>TOR</i> -RNAi on <i>Sna</i> protein levels in the PG	110
Figure 2-22. <i>Sna</i> levels in the PG are regulated by TORC1 complex	111
Figure 2-23. <i>Sna</i> levels in the PG are partially dependent on IIS signaling pathway	113
Figure 2-24. <i>Sna</i> levels in the PG are partially dependent on <i>tim</i>	114
Figure 2-25. Loss-of-IIS/TOR signaling does not significantly affect <i>sna</i> mRNA levels in the PG.....	115
Figure 2-26. Presence of <i>Sna</i> in PG nuclei is dependent on nutrient conditions	116

Figure 2-27. The presence of Sna in PG nuclei is dependent on nutrient conditions around the CW checkpoint but not after	117
Figure 2-28. Sna is expressed in the IPCs in the brain.....	119
Figure 2-29. Sna ⁺ cells do not largely overlap with S-phase cells in the PG	120
Figure 2-30. Sna protein levels in various <i>ex vivo</i> culture conditions.....	121
Figure 2-31. Overexpression of <i>esg</i> in the PG blocked endocycle progression.....	122
Figure 2-32. Overexpression of <i>sna</i> in the PG blocked endocycle progression and increased PG cell number.....	124
Figure 2-33. Overexpressing <i>CycE</i> in the PG increased the cell number	126
Figure 2-34. Spok-Gal4-GeneSwitch is inducible by RU486 feeding.....	127
Figure 2-35. <i>sna</i> inhibits the endocycle at the CW checkpoint in a cell-autonomous manner	128
Figure 2-36. Loss-of- <i>sna</i> in the PG disrupted the expression of genes essential for ecdysone biosynthesis	130
Figure 2-37. The upregulated genes in PG> <i>sna</i> -RNAi ring glands suggested metabolism-related defects	131
Figure 2-38. Cross comparison of the <i>sna</i> -RNAi RNA-Seq data and the <i>sna</i> -overexpression RNA-Seq data	133
Figure 2-39. Comparison between <i>sna</i> -overexpression RNA-Seq data and published ChIP-on-chip data	134
Figure 2-40. qPCR results showing the expression of the misregulated genes in either PG> <i>sna</i> -RNAi or <i>hs</i> > <i>sna</i> -cDNA ring glands	135
Figure 2-41. <i>ouib</i> mRNA levels are dependent on the IIS/TOR signaling	136
Figure 2-42. The expression profile of two cell-adhesion genes in the ring gland.....	137
Figure 2-43. The schematic view of predicted Sna binding sites in the six major ecdysone biosynthetic genes.....	138
Figure 2-44. The predicted Sna binding sites in the endoreplication related gene <i>double park (dup)</i>	139
Figure 2-45. Multiple protein sequences alignment between several Sna family proteins.....	140
Figure 2-46. Models for Snail function in the PG.....	141
Figure 3-1. The <i>cu</i> expression profile in 3rd instar larval ring glands revealed by microarray analysis .	163
Figure 3-2. CCR4-NOT complex and the relationship between mRNA translation and deadenylation .	164
Figure 3-3. Loss-of- <i>cu</i> function in the PG resulted in small pupae due to accelerated larval development	165
Figure 3-4. Overexpression of <i>cu</i> in the ring gland caused developmental arrest or delay	166
Figure 3-5. <i>cu</i> -dsRNA (VDRC construct GD8898) is functioning effectively.....	167
Figure 3-6. Models showing the putative role of Cu (Nocturnin) in PG cells	168
Figure 3-7. <i>cu</i> -RNAi in the ring gland phenocopied DHR4-RNAi with respect to developmental timing	169
Figure 3-8. <i>cu</i> genetically interacts with <i>DHR4</i>	170
Figure 3-9. Knocking down <i>cu</i> in the PG derepressed <i>Cyp6t3</i> expression.....	171

Figure 3-10. The <i>cu</i> -RD isoform is dependent on Torso (the receptor of the PTH pathway)	172
Figure 3-11. InDA-C treatment (Nugen, InC) significantly reduced the amount of rRNA reads from total RNA samples in RNA-Seq analysis	173
Figure 3-12. The downregulated genes in PG> <i>Pop2</i> -RNAi ring glands suggested protein translation defects.....	174
Figure 3-13. RNAi of <i>Pop2</i> in the PG upregulated the expression of genes encoding RNA polymerase and genes related to amino acid metabolism	176
Figure 3-14. qPCR results for the expression of six major ecdysone biosynthetic genes in PG> <i>Pop2</i> -RNAi	178
Figure 3-15. Larval arrest caused by PG> <i>Pop2</i> -RNAi could be partially rescued by 7DC and cholesterol feeding.....	179

List of tables

Table 2-1. qPCR primers used in the <i>sna</i> study	23
Table 2-2. Covaris S-Series Sonicator system settings	31
Table 2-3. >2-fold downregulated genes found in PG> <i>sna</i> -RNAi that are related to the ecdysone production	59
Table 2-4. Overrepresented GO terms and KEGG pathways for downregulated genes identified in PG> <i>sna</i> -RNAi	61
Table 2-5. Go term enrichment analysis results for upregulated genes identified in PG> <i>sna</i> -RNAi	62
Table 2-6. Go term enrichment analysis results for downregulated genes identified in <i>hs</i> > <i>sna</i> -cDNA ring glands	64
Table 2-7. Go term enrichment analysis results for upregulated genes identified in <i>hs</i> > <i>sna</i> -cDNA ring glands	66
Table 2-8. Genes that are >2-fold downregulated in PG> <i>sna</i> -RNAi and >2-fold upregulated in <i>hs</i> > <i>sna</i> -cDNA ring glands	67
Table 2-9. Genes that are >2-fold upregulated in PG> <i>sna</i> -RNAi and >2-fold downregulated in <i>hs</i> > <i>sna</i> -cDNA ring gland samples	68
Table 2-10. Overlap between genes misregulated in PG> <i>sna</i> -RNAi ring glands and genes associated with Sna binding peaks in the embryo stage	72
Table 2-11. Overlap between genes misregulated in <i>sna</i> -overexpression ring glands and genes associated with Sna binding peaks	73
Table 2-12. Predicted Sna binding sites for the six major ecdysone biosynthetic genes in <i>Drosophila melanogaster</i> using INSECT 2.0	80
Table 2-13. Predicted Sna binding sites for <i>Drosophila melanogaster dup</i> using INSECT 2.0	83
Table 3-1. qPCR primers used in <i>cu</i> and <i>Pop2</i> studies	148
Table 3-2. PG-specific RNAi of CCR4-NOT complex components resulted in ecdysone related phenotypes	155
Table 3-3. Number of genes affected by PG> <i>Pop2</i> -RNAi using various fold change cut-offs	157
Table 3-4. Go term enrichment analysis results for >2-fold downregulated genes in PG> <i>Pop2</i> -RNAi ring glands	158
Table 3-5. Go term enrichment analysis results for >2.5-fold upregulated genes in PG> <i>Pop2</i> -RNAi ring glands	159
Table A-1. Top 100 differentially expressed genes in the PG> <i>sna</i> -RNAi ring glands	196
Table A-2. Top 100 differentially expressed genes in the <i>hs</i> > <i>sna</i> -cDNA ring glands	200
Table A-3. Top 100 differentially expressed genes in the PG> <i>Pop2</i> -RNAi ring glands	204

Abbreviations

20E	20-hydroxyecdysone
7DC	7-dehydrocholesterol
Act5C	Actin 5C
AKH1	Adipokinetic Hormone
AEL	after egg laying
<i>Atet</i>	<i>Atypical topology ecdysone transporter</i>
CA	Corpus allatum
CadN	N-Cadherin
Cas9	CRISPR associated protein 9
CC	Corpora cardiaca
CCR4	yeast carbon catabolite repression 4
<i>Cdk2</i>	<i>Cyclin-dependent kinase 2</i>
cDNA	complementary DNA
ChIP-on-chip	chromatin immunoprecipitation (ChIP) and DNA microarray (chip)
CRISPR	Clustered regularly interspaced short palindromic repeats
CtBP	C-terminal Binding Protein
<i>cu</i>	<i>curled</i>
C-value	Chromatin value
CW	critical weight
<i>CycE</i>	<i>cyclin E</i>
Cyp	cytochrome P450
<i>D.</i>	<i>Drosophila</i>
DAPI	4',6-diamidino-2-phenylindole
DHR4	Drosophila Hormone Receptor 4
<i>dib</i>	<i>disembodied</i>
DN	Dominant negative
DP	DP transcription factor
dsRNA	double-stranded RNA
<i>dup</i>	<i>double parked</i>
E2f1	E2F transcription factor 1
EcR	Ecdysone receptor
EdU	5-ethynyl-2'-deoxyuridine
EGFP	Enhanced Green Fluorescent Protein
<i>egr</i>	<i>eiger</i>
EIA	enzyme immunoassay
EMSA	electrophoretic mobility shift assay

EMT	epithelial to mesenchymal transition
<i>Erk7</i>	<i>Extracellularly regulated kinase 7</i>
<i>esg</i>	<i>escargot</i>
FBS	Fetal Bovine Serum
FC	fold change
FDR	False Discovery Rate
<i>Fdx2</i>	<i>Ferredoxin 2</i>
L1	first instar
FLP	Flippase
FRT	Flippase Recognition Target
G phase	Gap phase
GFP	Green fluorescent protein
GO	Gene Ontology
GOI	Gene of interest
gRNA	guide RNA
GS	GeneSwitch
GSK3 β	Glycogen Synthase Kinase-3 beta
<i>GstE14</i>	<i>Glutathione S transferase E14</i>
<i>h</i>	<i>hairy</i>
<i>hid</i>	head involution defective
hr	hour
hs	heat shock
IIS	Insulin/Insulin-like Growth Factor signaling
ILP	Insulin-like peptide
InDA-C	Insert-Dependent Adaptor Cleavage
InR	Insulin-like Receptor
IPC	Insulin-producing median neurosecretory cells
JH	Juvenile hormone
JHBP	juvenile hormone-binding protein
JNK pathway	c-Jun N-terminal kinase pathway
KEGG	Kyoto Encyclopedia of Genes and Genomes
MAPK	Mitogen-Activated Protein Kinase
min	minute
<i>mld</i>	<i>molting defective</i>
mRNA	messenger RNA
<i>neur</i>	<i>neuralized</i>
NIG	National Institute of Genetics (Japan)
<i>npc1a</i>	<i>Niemann-Pick type C-1a</i>

<i>nvd</i>	<i>neverland</i>
<i>Oatp74D</i>	<i>Organic anion transporting polypeptide 74D</i>
<i>ouib</i>	<i>ouija board</i>
PBS	Phosphate Buffered Saline
PBST	PBS containing 0.3% Triton-X 100
<i>per</i>	<i>period</i>
PG	Prothoracic Gland
<i>phm</i>	<i>phantom</i>
PI3K	phosphoinositide 3-kinase
<i>Pten</i>	<i>Phosphatase and tensin homolog</i>
PTTH	prothoracicotropic hormone
qPCR	Quantitative Real-time
RG	ring gland
RNAi	RNA interference
RNA-Seq	RNA sequencing
RPKM	Reads Per Kilobase Million
<i>rpr</i>	<i>repear</i>
rRNA	Ribosomal RNA
RT	room temperature
S phase	Synthesis phase
<i>sad</i>	<i>shadow</i>
L2	second instar
<i>serp</i>	<i>serpentine</i>
siRNA	small interference RNA
<i>sna</i>	<i>snail</i>
<i>SNCF</i>	<i>Sox Neuron Co-factor</i>
<i>spok</i>	<i>spookier</i>
<i>sro</i>	<i>shroud</i>
<i>TfIIFβ</i>	<i>Transcription factor TFIIIFβ</i>
L3	third instar
Tim	Timeless
TOC1	Target of rapamycin complex 1
TOR	Target of Rapamycin
<i>Traf4</i>	<i>TNF receptor-associated factor 4</i>
TRiP	The Transgenic RNAi Resource Project
<i>trk</i>	<i>trunk</i>
TS	temperature sensitive
<i>tsl</i>	<i>torso-like</i>

<i>tub</i>	<i>tubulin</i>
UAS	upstream activation sequence
VDRC	Vienna <i>Drosophila</i> Resource Center
<i>verm</i>	<i>vermiform</i>
Vvl	Ventral veins lacking

1. General Introduction

1.1 Ecdysone, a major steroid hormone in *Drosophila melanogaster*, controls developmental transitions

Steroid hormones are signaling molecules found in both the animal and plant kingdoms that have crucial roles in multiple physiological processes, such as regulating glucose metabolism, immune functions, salt and water balance as well as controlling development and sexual maturation (1-4). My research focused on how the production of steroid hormones is regulated using *Drosophila melanogaster* as a model organism. The *Drosophila* life cycle consists of several discrete developmental stages, including embryo, three larval stages (L1, L2 and L3), pupae and adult stages. Extensive studies have shown that, in *Drosophila*, all major developmental transitions are triggered by peaks of 20-hydroxyecdysone (20E, the biologically active form of the steroid hormone ecdysone) (5) (Fig. 1-1). In target tissues, ecdysone binds to a nuclear receptor heterodimer composed of the Ecdysone receptor (EcR) and the nuclear receptor Ultraspiracle (USP) (6-8). This triggers the transcription of ecdysone response genes (9-15), leading to specific developmental outputs. This cascade has been well studied for decades and is not the focus of my thesis.

In contrast to vertebrates where growth and maturation are coupled during puberty, in holometabolous insects, growth and maturation are separate events, with growth being restricted to three larval stages and maturation occurring during metamorphosis (16). Pulses of ecdysone control all developmental transitions, raising the question as to how ecdysone pulses are regulated themselves. In fact, the timing of each ecdysone pulse is precisely controlled, resulting in durations for each developmental stage being consistent between animals. For example, the *Drosophila* L3 stage, which immediately precedes pupariation, always lasts for two days under standard conditions. As will be discussed in following sections, the amplitude and duration of each ecdysone pulse is dependent on various factors including rate of synthesis, release, transport and metabolism of ecdysone. The aim of my thesis is to expand knowledge on how ecdysone production is regulated.

1.2 Neuroendocrine control of ecdysone production

Ecdysone is synthesized from suitable sterol precursors such as cholesterol via a series of enzymatic reactions in the larval prothoracic gland (PG). The PG is one of three endocrine glands that together form the ring gland (RG) (Fig. 1-2). It is well-known that changes in ecdysone titers

are - in part - controlled by a neuropeptide, prothoracicotropic hormone (PTTH), which is produced by pairs of neurosecretory cells in each brain hemisphere that directly project to the PG (17). This has striking similarity to the vertebrate hypothalamic-pituitary-adrenal (HPA) axis, in which the peptide adrenocorticotropic hormone (ACTH) produced in the brain, regulates the production of cortisol in the peripheral endocrine gland, adrenal cortex (18,19).

In addition to the production of ecdysone, the RG is also the source of juvenile hormone (JH) and adipokinetic hormone (AKH) (20,21), which are synthesized in two other small glands in the RG: the Corpus allatum (CA; produces JH) and the Corpora cardiaca (CC; makes AKH)(Fig. 1-2A). JH also has important roles during development, where it promotes larval fates by suppressing metamorphosis and determines the ecdysteroid-induced molt as larval or pupal molt (22). AKH is similar to mammalian glucagon and regulates hemolymph sugar homeostasis (23).

1.3 The ecdysone biosynthetic enzymes and their coding genes

Most of the enzymatic steps that convert dietary cholesterol to ecdysone as well as the genes encoding these enzymes have been identified in the past decade. The process starts with Neverland (Nvd), which in turn generates 7-dehydrocholesterol (7DC) (Fig. 1-2B) (24). Also, the last three steps towards ecdysone are well characterized, and the genes encoding these three enzymes are *phantom* (*phm*), *disembodied* (*dib*) and *shadow* (*sad*) (25-28). However, the intermediate steps from 7DC to 5 β -ketodiol are relatively poorly understood and are referred to as the black box. At least three enzymes are known to fall within the black box, Shroud (Sro), Spookier (Spok) and Cyp6t3 (29-31). Once the prohormone α -ecdysone is synthesized in the PG, it is secreted into the hemolymph and converted in peripheral tissues to its active form, 20-Hydroxyecdysone (20E) by Shade (Shd) (32). Except for Nvd and Spok, all of the ecdysone biosynthetic enzymes belong to the Cytochrome P450 family. The term “ecdysone” in my PhD dissertation refers generically to α -ecdysone and 20E, unless 20E is specifically mentioned.

1.4 The PTTH/MAPK cascade regulates ecdysone production

The cellular signaling pathway that mediates PTTH signaling in the PG has also been characterized. In PG cells, PTTH binds to its receptor Torso, which subsequently activates the mitogen-activated protein kinase (MAPK) pathway, which includes the *Drosophila* homologs of Ras (encoded by *Ras oncogene at 85D*, *Ras85D*), Raf (encoded by *Raf oncogene*, *Raf*), MAPK

kinase (MEK, encoded by *Downstream of raf1*, *Dsor1*), and extracellular signal–regulated kinase (ERK, encoded by *rolled*, *rl*) (Fig. 1-3) (33,34). How the MAPK cascade stimulates the production of ecdysone is still unclear since direct targets of the PTTH/MAPK pathway in the PG have yet to be identified. However, it has been shown in other insect systems that expression of some ecdysone biosynthetic genes (including *spok*, *phm* and *dib*) is upregulated upon stimulation with PTTH, and that loss-of-*torso* function or loss-of-PTTH by ablating the PTTH-producing neurons in *Drosophila* results in a failure to upregulate these enzyme genes (27,35). Our lab demonstrated that DHR4 (*Drosophila* Hormone Receptor 4), a nuclear receptor, is a key target of the PTTH pathway (31) (see section 3.1.4 for details). Intriguingly, loss-of-PTTH by ablating the PTTH-producing neurons just delayed the onset of metamorphosis, but did not block development completely, which suggests that PTTH is sufficient, but not necessary, for regulating the formation of ecdysone pulses (17), and therefore, other signaling pathways must be involved in regulating ecdysone production as well.

1.5 Prothoracic gland, a CPU-like decision-making center for developmental progression

Since ecdysone essentially dictates the timing of developmental progression, multiple internal and external cues need to converge at the PG through cellular signaling pathways to control ecdysone biosynthesis, ensuring that developmental transitions only occur when all conditions are met (36). Besides the PTTH/MAPK signaling pathway, other signaling cascades that are known to act on the PG to regulate ecdysone production include: transforming growth factor beta (TGF- β)/Activin signaling, Nitric oxide (NO), circadian machinery, Insulin/IGF signaling (IIS) and Target of Rapamycin (TOR) signaling (Fig. 1-3) (37-44). The Insulin/IGF signaling (IIS) and Target of Rapamycin (TOR) signaling will be explained in more detail in the next section. A specific example of how the animals' internal status is coupled with developmental timing/progression is that when there is impairment in larval tissue growth, *Drosophila* insulin-like peptide 8 (*dilp8*) is secreted to stop ecdysone biosynthesis by delaying PTTH synthesis in the brain (Fig. 1-3). In this way, it allows for extra time the tissue need to repair and grow before the onset of metamorphosis (45,46). Therefore, PG is the decision-making center where all signals converge to designate ecdysone production.

1.6 IIS/TOR signaling and its function in the PG

Interconnected IIS/TOR signaling is required in the PG for ecdysone synthesis (41,43,47). Studies in various organisms including *Drosophila* showed that IIS/TOR signaling is the nutrient-sensing system that couples growth to nutritional conditions (48). Specifically, TOR is the amino acid sensor since mutations in *TOR* phenocopies amino acid deprivation (49). Hence the importance of the IIS/TOR signaling in the PG suggests a potential mechanism coupling the animals' nutrient and metabolic states to developmental timing. This is supported by the fact that developmental delay caused by low nutrient conditions could be rescued by PG-specific activation of TOR (44).

In *Drosophila*, the core IIS signaling starts with binding of ligands to the receptor, insulin-like receptor (InR). The ligands are Insulin-like peptides (ILPs), and seven of them (ILPs 1-7) have been identified in *Drosophila* (50), among which ILPs 1, 2, 3 and 5 are produced in pairs of seven insulin-producing median neurosecretory cells (IPCs) in each brain hemisphere (50-52) during the larval stage. Expression of ILP2 and ILP5 is nutrient-dependent (53-55). For example, the availability of amino acids is sensed via TOR signaling in the fat body, the tissue analogous to the liver and adipose tissue in vertebrates, and consequently affects the release of ILPs from the brain. The IPCs have axons that project to the CC in the ring gland as well as aorta so that ILPs are secreted into the hemolymph. In turn, ILPs function in different tissues through the InR to regulate growth, lipid metabolism, carbohydrate metabolism and longevity (48,56).

In target cells, upon activation, InR phosphorylates insulin receptor substrate, *chico*, in *Drosophila*. Chico then activates the downstream effector, phosphoinositide 3-kinase (PI3K), which phosphorylates the inositol lipid (converting the Phosphatidylinositol 4,5-bisphosphate to Phosphatidylinositol 3,4,5-trisphosphate) and eventually generates a variety of second messengers, including Akt, aka Akt1 in *Drosophila* (Fig. 1-4). In *Drosophila*, PI3K consists of the catalytic subunit (p110 encoded by *Pi3K92E*) and the PI3K adaptor (p60, encoded by *Pi3K21B*) (57,58). Another component of the IIS is the lipid phosphatase PTEN, which antagonizes the kinase activity of PI3K, and thus is a negative regulator of IIS signaling (Fig. 1-4)(59). Akt phosphorylates several target genes, one of which is Glycogen synthase kinase-3 beta (GSK-3 β), encoded by *shaggy* (*sgg*) in *Drosophila*. Upon Akt phosphorylation, GSK-3 β is inactivated. One example of GSK-3 β function is to phosphorylate the "Clock" transcription factor Timeless (Tim)

and in turn regulates its nuclear translocation. As a consequence, disrupting GSK-3 β would affect the length of the circadian period (60). Consistent with a link between GSK-3 β and Tim, the circadian clock is also coupled with IIS signaling for steroidogenesis in the PG (40).

The kinase TOR participates in two distinct protein complexes, TORC1 and TORC2 (61). The core adaptor proteins of TORC1 are Raptor and LST8, whereas, Rictor and Sin1 are the distinct components of TORC2 (Fig. 1-4). The activity of TORC1 is responsive to the availability of glucose, amino acids and oxygen (62). In mammals, TORC1 can also be activated by IIS signaling through Akt, where Akt inhibits the tumor suppressor Tsc2, which normally negatively regulates TORC1 functions. However, this regulation between Akt and Tsc2 in *Drosophila* is not entirely clear (63,64). TORC1 controls cell growth, global translation through ribosomal protein S6 kinase (S6K) and eIF4E-binding proteins (4E-BP, *aka* Thor in *Drosophila*) as well as autophagy (65,66), while TORC2 is less well understood and was implicated in regulating glucose metabolism and the actin cytoskeleton (67). The best studied downstream target of TORC2 is Akt, where TORC2 phosphorylates Akt to modulate its activity (68,69), which adds another layer of complexity between the IIS and TOR pathways.

1.7 Other factors known to regulate the expression of ecdysone biosynthetic genes in the PG

Several transcription factors have been shown to regulate ecdysone production in the PG, some of which directly affect the expression of the ecdysone biosynthetic genes. These factors include Without Children (WOC), a C2C2 zinc finger transcription factor that is thought to regulate the activity of Neverland (70,71). Molting Defective (Mld) is a protein with seven C2H2 zinc fingers and it was found that the transcript levels of *spookier* (*spok*) were decreased in *mld* mutant, suggesting that *spok* is a target of Mld in the PG (72,73). Moreover, Ventral veins lacking (Vvl) and Knirps (Kni) are also required for the proper expression of the ecdysone biosynthetic genes during late larval development and were shown to directly bind to the enhancer region of *phm* (74,75). Lastly, another C2H2 zinc finger transcription factor Ouija Board (Ouib) directly regulates the transcription of *spok* in the PG (76).

1.8 Cellular transport and release of ecdysone

Compared to the regulation of steroidogenesis, relatively little is known about how ecdysone is transported and released from the PG. The conventional view was that steroid

hormones simply enter and exit cells by diffusion across lipid bilayers. However, recently Yamanaka *et al.* (77) reported that calcium signal-mediated vesicle exocytosis is involved in ecdysone release and intracellular calcium release in the PG is regulated by G protein-coupled receptors (GPCRs)-mediated cellular signaling pathway. However, the actual GPCR has yet to be identified (77). Furthermore, an ATP-binding cassette (ABC) transporter encoded by *atypical topology ecdysone transporter (Atet)* was found to be highly expressed in the PG and Atet is located on the vesicle surface to load ecdysone into secretory vesicles (77).

Another important aspect of controlling ecdysone production is the intake of dietary cholesterol and transport into the PG. *Niemann-Pick type C1 (npc1)* or *C2 (npc2)* family genes play important roles in intracellular cholesterol trafficking (78-80). *Drosophila* mutants in *npc1* or *npc2* genes have defects in ecdysone biosynthesis (81-83). One specific member of the family, *npc1a*, is highly expressed in the RG and expressing *npc1a* in the RG alone can rescue the larval arrest phenotype caused by ecdysone deficiency in *npc1a* whole body mutants, establishing *npc1a* as an important regulator of ecdysone production, likely in the aspect of cholesterol transport (82). Another known player is *Start1*, which encodes a cholesterol-binding protein with the Steroid acute regulatory protein (StAR)-related lipid transfer domain (START). Start1 has been shown to express in the PG, and its exact molecular function remains uncharacterized (84).

1.9 Feedback regulation of ecdysone production

For the discrete peaks of ecdysone to form throughout development, there must be forward and feedback mechanisms to rapidly ramp up of ecdysteroidogenesis and PG activities for the ecdysone peaks as well as sufficiently repress ecdysone levels or activities at the troughs. Simply put, EcR is expressed in the PG, which in turn can mediate ecdysone signals and affect expression of ecdysone-responsive genes in the PG (85-88). Meanwhile, several ecdysone-responsive nuclear receptors were shown to regulate or potentiate ecdysteroidogenesis in the PG, thus forming a so-called autoregulatory loop. Indeed, knocking down EcR in the PG disrupted ecdysone biosynthesis (87). One relevant example to my thesis is DHR4 (see section 3.1.4 for details), a nuclear receptor that was first identified as an ecdysone inducible gene (89) and was later shown to be a repressor for ecdysone biosynthesis in the PG (31). Moreover, in target tissues, another Cytochrome P450 enzyme, Cyp18a1, which functions to metabolically inactivate 20E was also inducible by ecdysone (90,91). This represents at least one of the mechanisms by which an ecdysone peak can

eventually decline. Lastly, a classic ecdysone-inducible gene *Early gene at 23 (E23)* codes for another ABC transporter, like *Atet*, which could be a potential cellular ecdysone transporter that is responsive to the level of ecdysone itself (92). However, so far no direct experimental evidence shows that E23 is involved in ecdysone transport (77). Interestingly, E23 also responds to circadian clock signals (93), therefore connecting the ecdysone response and circadian clock in the target tissue in a way that requires further examination.

1.10 Critical weight coordinates body size and timing of metamorphosis

As discussed previously, nutrient conditions feed into the regulatory signals of ecdysone production. During the last larval stage (L3 in *Drosophila*), there is a physiological transition called critical weight (CW) attainment, after which larvae are committed to metamorphosis regardless of nutrient conditions (94). If larvae are starved before CW is attained, further development will be delayed until nutrients become available, at the same time the production of the major ecdysone peak that triggers the onset of metamorphosis will not be engaged. Once conditions improve, larval growth will resume and animals will grow to the target body size. On the contrary, when animals are starved after the critical weight checkpoint, the onset of metamorphosis will not be affected but larvae will pupariate at a smaller size due to limited nutrient uptake (95) (also see Fig. 2-17A). Molecularly, JH titers drop after CW attainment allowing for PTTH to be released to trigger the late ecdysone peak for metamorphosis at least according to the several classic experiments carried out in *Manduca* (96,97). The discovery of the CW checkpoint showed that animals have an intrinsic molecular mechanism to assess their body mass/weight, which is necessary to ensure animals could survive metamorphosis. Some evidence showed that the intrinsic CW is genetically determined. For instance, slowing down the larval growth rate via attenuating insulin signaling delays the attainment of CW, but the actual weight at which the checkpoint is fulfilled does not change (98). However, molecular mechanisms underlying size assessment and the molecular switch before and after CW remain mostly unknown. It is known that before the major ecdysone pulse that triggers pupariation, there are also three minor pulses during the L3 stage (99). The first minor pulse coincides with the time of CW attainment. However, the causal-effect relationship is not entirely clear since the minor ecdysone pulse could happen either right after the CW checkpoint, being the result of CW attainment or right before CW attainment, necessary for animals to bypass the CW checkpoint.

1.11 The prothoracic gland is an endoreplicating tissue

PG cells undergo a modified form of the cell cycle, called endoreplication, where cells replicate the genome without cell divisions. Consequently, cells have multiple copies of each chromosome and the homologous chromosome arms are closely paired, which appear as single polyploid giant chromosomes and this is also called polyploidy. The level of polyploidy is indicated by Chromatin values (C). C-values are based on DNA content as multiples of the haploid genome. For example, an egg nucleus is 1C and a zygote is 2C. Generally, endoreplication is the cellular program chosen when a specific tissue focuses more on increasing cell size/mass instead of an increase in cell number for the tissue's specific functions. In the case of the *Drosophila* larval PG and salivary gland, endoreplication helps to quickly scale up the capacity of hormone production (in the PG) or glue production for pupae adhesion (in the salivary gland) (100). The PG cell number is fixed at around 50 early in development (at the L1 stage), but the C-value can reach 64C by the end of the final larval stage (101). In the case of the salivary gland, the cells can reach a C-value as high as 1024C. Since diploidy is an important quality of genome integrity, endoreplication is not usually favored; however, in the case of the PG and salivary gland, they are specific larval tissues that would not persist into the adult stage. Two other primary purposes of endoreplication are: large nutrient uptake and storage (102,103) and maintaining specific tissue morphology (104,105).

One study suggested that the growth condition/size of the PG could be an internal assessment of the whole organism's growth, thus being part of the molecular basis of CW attainment. Over-activating IIS signaling in the PG by overexpressing p110 decrease the CW, probably because the animals overestimate the body growth due to enlarged PG (43). A more recent study showed that the CW might be assessed via the number of endocycles that the PG went through in *Drosophila*. A C-value of 16 seems to be the threshold of CW attainment and there is one round of endoreplication tightly coupled with the time of the CW checkpoint (101).

1.12 Regulation of endocycle in *Drosophila melanogaster*

In endoreplicating cells, DNA synthesis does not occur continuously. Instead, endoreplication comprises discrete periods of DNA synthesis phase (S phase) interspaced with gaps (G phase), which are also called endocycles (106,107). The genome will only replicate once within a single S phase to ensure genome integrity. To achieve this, the eukaryotic genome harbors

many distinct sites called "origins of replication," where the prereplicative complexes (pre-RC) are assembled during G phase (108). pre-RC would then load the Minichromosome Maintenance (MCM), a protein complex that acts as a DNA unwinding helicase, necessary for the formation and elongation of the DNA replication fork (109). Once DNA synthesis has started at the origins, several mechanisms will act on the protein members of pre-RC complexes to prevent them from reassembling again at the origins until the next G phase, which will not be discussed in detail here (110). One important pre-RC complex member is *Cdc10-dependent transcript 1 (Cdt1)*, also called *double parked (dup)* in *Drosophila* (see chapter 2 discussion) (111). The alternating S and G phase are regulated by a molecular oscillator, the E2f1 transcriptional activator and DP protein, together they form a heterodimer to increase the transcription of *CyclinE (CycE)* in late G phase. CycE, in turn, activates the cyclin-dependent kinase2 (Cdk2) and together with Cdk2 triggers S phase entry, by driving loading of MCM complex and preventing re-replication (reassembly of pre-RC complexes) within S phase (112). Another target of CycE is Cullin 4 (Cul4)-containing E3 ubiquitin ligase complex CRL4^{Cdt2}, which targets E2f1 for degradation during S phase. Consequently, the *CycE* mRNA supply declines at late S phase and in combination with a degradation mechanism of CycE proteins, allows for a period of low CycE/Cdk2 that releases cells into G phase. The low CycE/Cdk2 activities are required to relicense pre-RC assembly to allow successive S phase. Therefore, constitutive expression of CycE arrests the endocycle in salivary glands (113,114). Similarly, stabilizing E2f1 can also inhibit endocycle progression (115). Upstream factors known to promote endocycle progression in *Drosophila* are IIS/TOR signaling and Myc. This makes sense because they are known drivers of general cellular growth (116-118). However, a study has shown that TOR signaling might be a more specific regulator for endoreplication than IIS signaling since TOR mutants affected the growth of the endoreplicating tissues to a larger extent than when IIS was disrupted (119).

1.13 Identifying novel regulators for ecdysone production via RG-specific microarrays and PG-specific RNAi screen

To gain a better understanding of RG function and to explore novel signaling pathways that regulate the formation of ecdysone pulses, and potentially novel regulators of AKH and JH production, our lab performed genome-wide microarray analyses aimed at identifying genes that are specifically expressed in the RG. For this, we compared total RNA extracted from *Drosophila*

RG with RNA isolated from whole larvae at four different time points (4, 8, 24 and 36 hr after the molt to 3rd instar (L3), the last instar stage before pupation). We identified 108 transcripts that are >20-fold enriched in *Drosophila* RG, suggesting that these genes play specific roles in this tissue (120). Previously unidentified players in regulating ecdysone, AKH or JH production and secretion should be enriched in this “RG-specific gene set”.

Since our lab focuses on studying the regulation of ecdysone production, we asked whether the 108 transcripts are functionally important for the PG. Our approach to answer this question was to perform an *in vivo* PG-specific RNAi screen using the UAS/Gal4 system to target the 108 transcripts that we identified in our microarray studies. Thanks to the existence of the *Drosophila* transgenic library carrying hairpin RNAi constructs under the control of the upstream activation sequence (UAS), targeting more than 82% of the *Drosophila* genome (121), this kind of *in vivo* tissue-specific RNAi screen can be carried out. Using a PG-specific Gal4 driver (*phantom22-Gal4*), the UAS-RNAi constructs will only be expressed in the PG. In the end, we identified a total of 25 lines with obvious phenotypes, 85% of which are either larval arrest where developmental transition triggered by ecdysone pulses did not occur or large pupae phenotype caused by prolonged feeding times when the molt from larva to pupa is delayed. Again, the large pupae phenotype is usually due to misregulation of ecdysone production where the ecdysone pulse that triggers the onset of metamorphosis is delayed (120). These phenotypes suggested that these 25 genes have important functions in ecdysteroidogenesis. Among those were zinc finger transcription factor Snail, which has well-characterized roles during *Drosophila* embryogenesis (122) and the fly homolog of the vertebrate circadian rhythm output gene *nocturnin* (also known as *curled; cu*), which encodes an mRNA deadenylase (123,124). The functional roles of Snail and Nocturnin in PG need to be explored in detail and are essentially the main focus of my thesis.

I participated in screening the 108 RNAi lines using the CC-specific Gal4 driver, *Akh-Gal4*, and did not identify any hit. As for the screen in the CA using the *Aug21-Gal4*, I identified a single hit, *Oatp74D*, knocking down of which in the CA resulted in ~52% pupal lethality. More strikingly, RNAi based on a ubiquitous driver *act5C-Gal4* resulted in a specific development defect where the top part of the larvae elongated and attempted to form pupal structures, which made the animals look like a wine bottle, hence called a bottleneck phenotype (Fig. 1-5). So far, the possible role of *Oatp74D* has not been followed up. RNAi studies have possible off-target effects where the small interference RNAs (siRNAs) processed from the long double-stranded

RNAs (dsRNAs) interfere with unknown transcripts instead of, or in addition to, the intended target, owing to a tolerance for mismatches and gaps in base-pairing with targets. A simple way to confirm that the RNAi phenotype is indeed caused by the disruption of the intended transcripts is to use another RNAi line targeting *Oatp74D* transcript at a separate mRNA region to see if the result is repeatable. So far only one RNAi line (VDRC #37295) was tested; therefore, I cannot rule out the possibility that the RNAi phenotype could be an off-target effect. Fly Stocks of National Institute of genetics (NIG) provides another *Oatp74D*-RNAi line (NIG 7571R-1), which targets a complete different region of *Oatp74D* mRNA as the VDRC line. In the future, one can order the RNAi from NIG and confirm the phenotype before further investigating the function of *Oatp74D* in the CA.

As previously stated, my work focuses on characterizing the roles of *snail* and *nocturnin* (*curled*) in the PG and each gene will be discussed in a separate chapter. The mechanism by which *snail* regulates the production of ecdysone in the PG will be demonstrated in Chapter 2 and how *nocturnin* (*curled*) function was related to the regulation of ecdysone production will be explored in Chapter 3.

1.14 Figures

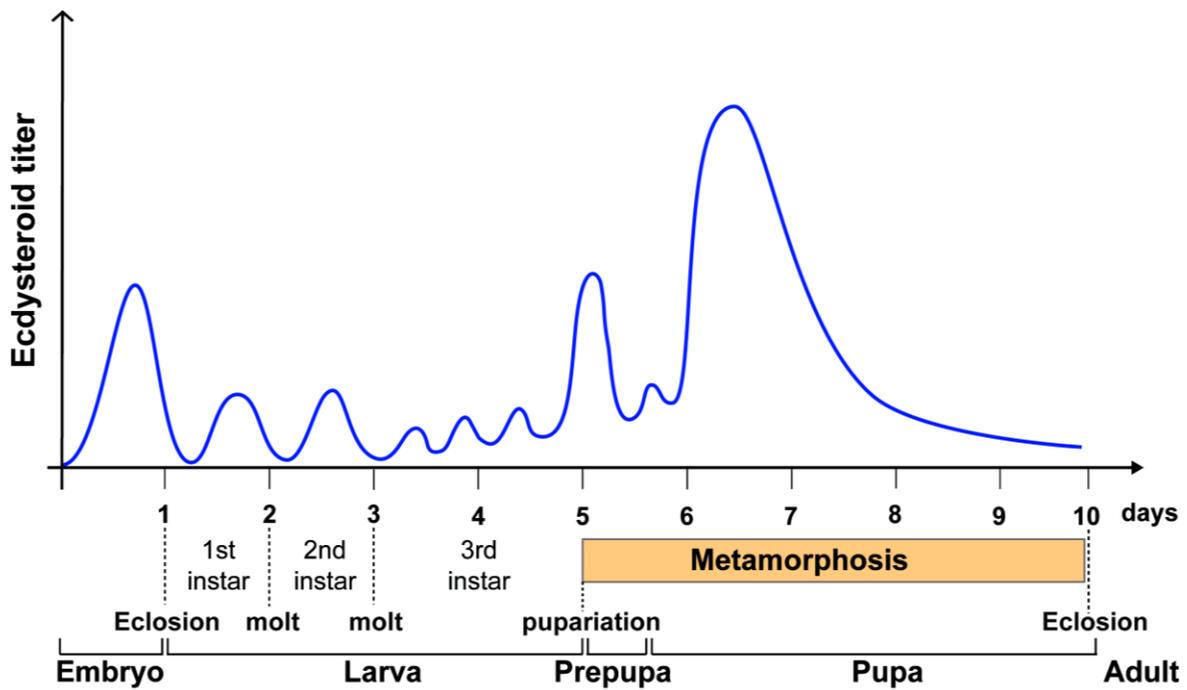


Figure 1-1. Schematic representation of whole body ecdysteroid titer throughout the *Drosophila* life cycle.

The *Drosophila* life cycle consists of several distinct stages that are separated by molts. After the completion of embryogenesis, larvae progress through three instar stages before entering metamorphosis, where animals transform into a sexually mature adult. The duration and onset of each stage are determined by pulses of the molting hormone ecdysone.

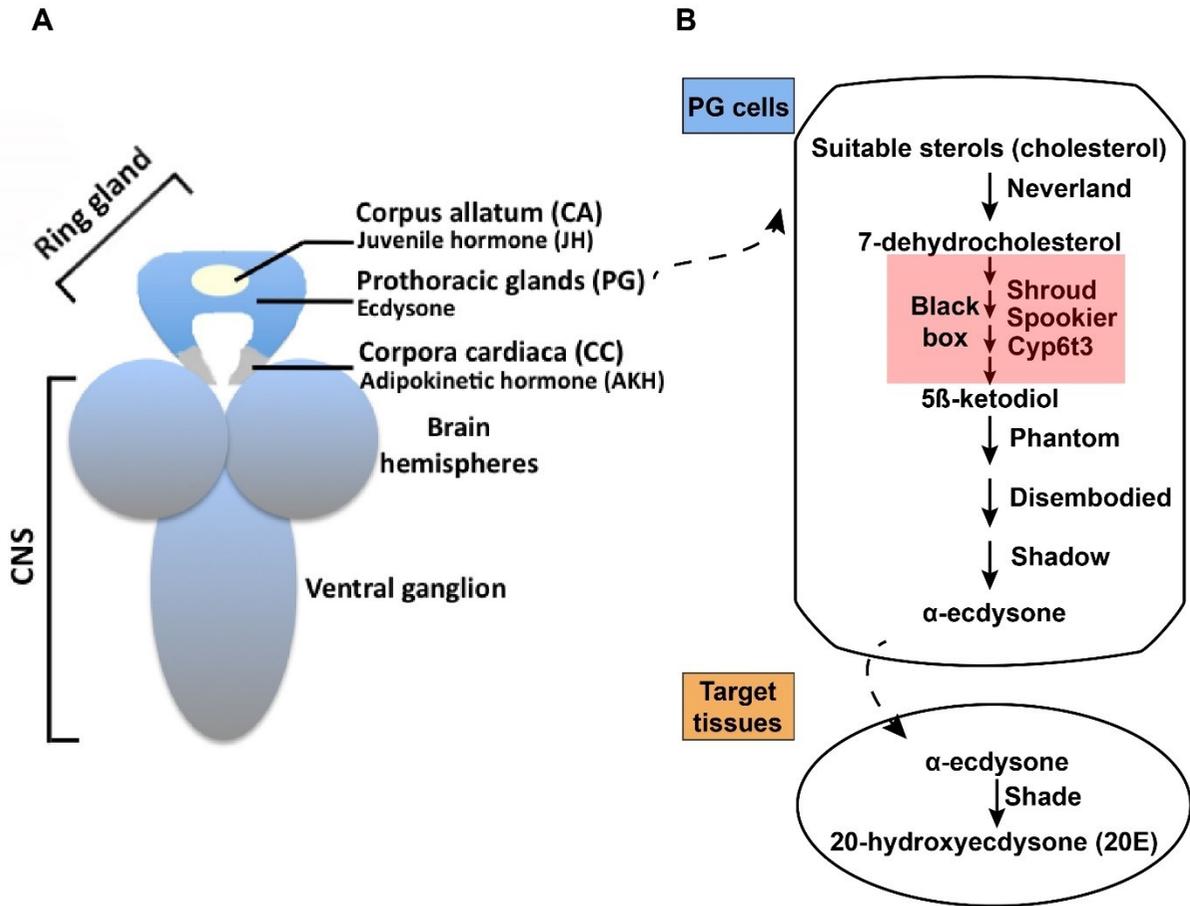


Figure 1-2. Ecdysone is produced in the prothoracic gland, a part of an endocrine tissue, the ring gland. (A) The structure of the brain-RG complex in larvae. The RG consists of three fused glands: the Corpus allatum (CA: produces Juvenile hormone), the Prothoracic Gland (PG: synthesizes ecdysone) and the Corpora cardiaca (CC: synthesizes AKH). The larval RG attaches to the two brain hemispheres. CNS: Central nervous system. (B) Simplified scheme of 20-hydroxyecdysone (20E) biosynthetic pathway. Dietary sterols, commonly cholesterol, serve as the precursors for 20E synthesis. All the known ecdysteroid biosynthesis enzymes are shown to the right of the arrows. For simplicity, only the first two known intermediates, namely 7-dehydrocholesterol (7DC) and 5β-ketodiol, are shown since they are relevant to my dissertation. The conversion steps from 7DC to 5β-ketodiol are not fully characterized (Black Box), however, they involve at least three enzymes, namely Shroud, Spookier and Cyp6t3. The prohormone α-ecdysone is secreted from the PG and converted to the biologically active form 20-hydroxyecdysone (20E) in target tissues.

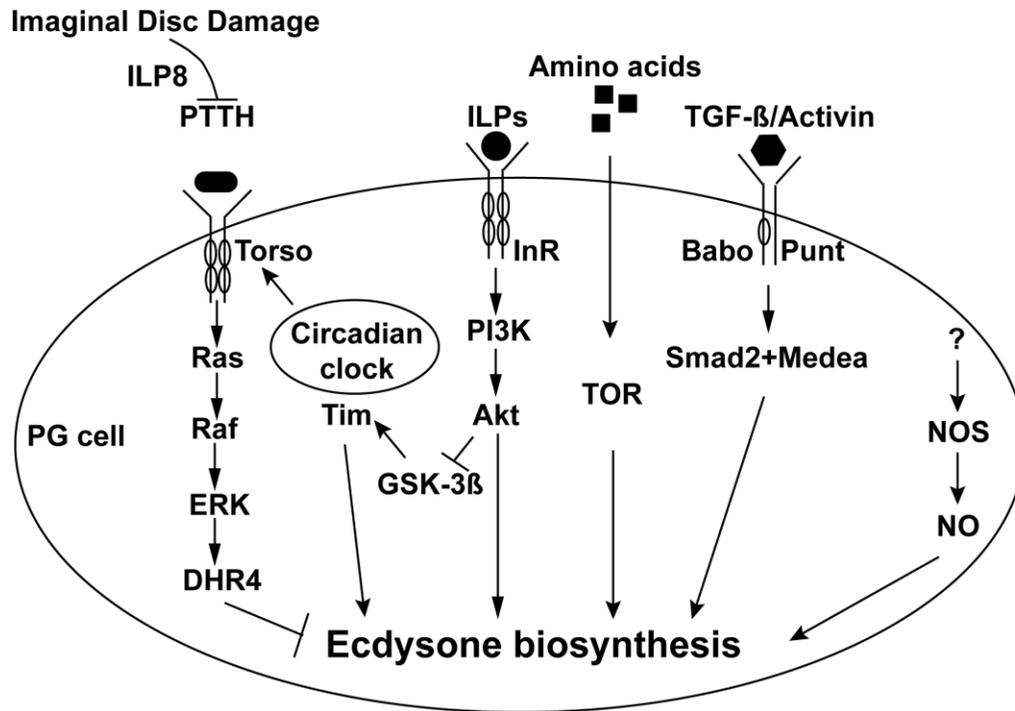


Figure 1-3. Known signaling pathways that are essential for ecdysone biosynthesis in the PG.

Several signaling pathways are involved in regulating ecdysone biosynthesis in the PG. Only the core components of each pathway are shown. Arrows indicate positive regulation and cross lines indicate inhibitory interactions. PTTH: prothoracicotropic hormone. Tim: Timeless. ILP: Insulin-like peptide. InR: insulin receptor. PI3K: phosphoinositide 3-kinase. TOR: Target of rapamycin. TGF- β : transforming growth factor beta. Babo: Baboon. Smad2: mothers against decapentaplegic homolog 2. NOS: Nitric oxide synthase. NO: Nitric oxide.

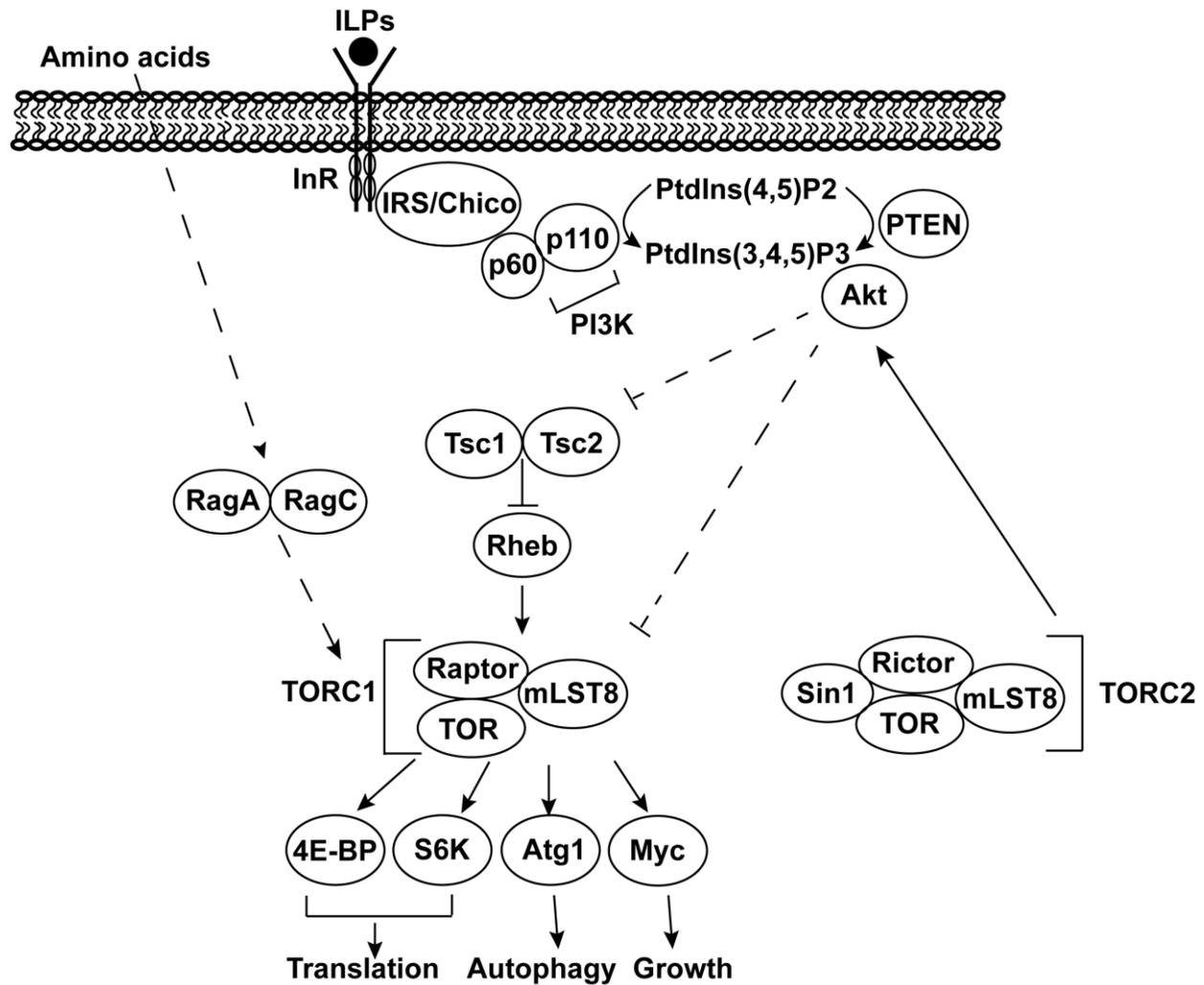


Figure 1-4. A simplified demonstration of the interconnected IIS/TOR signaling pathway in *Drosophila*. Arrows indicate positive regulation and cross lines indicate inhibitory interactions. Dashed lines indicate indirect interactions or interactions requiring further study. ILP: Insulin-like peptide. InR: insulin receptor. IRS: Insulin receptor substrate. PI3K: phosphoinositide 3-kinase. Rheb: Ras homolog enriched in brain. TOR: Target of rapamycin. Atg1: Autophagy-related 1. RagA: Ras-related GTP binding A/B. RagC: Ras-related GTP binding C/D. TORC1: target of rapamycin complex 1, TORC2: target of rapamycin complex 2. 4E-BP: eukaryotic translation initiation factor 4E binding protein. S6K: Ribosomal protein S6 kinase. PtdIns(4,5)P2: Phosphatidylinositol 4,5-bisphosphate. PtdIns(3,4,5)P3: Phosphatidylinositol 3,4,5-trisphosphate. PTEN: Phosphatase and tensin homolog.



Figure 1-5. Ubiquitous knock down of *Oatp74D* resulted in 3rd instar lethality.

A ubiquitous Gal4 driver, *actin5C*-Gal4 was crossed to *UAS-Oatp74D*-RNAi lines (VDRC#37295). Control: *actin5C*> *w¹¹¹⁸* and *Oatp74D*-RNAi: *actin5C*>*Oatp74D*-RNAi.

Chapter 2. Snail coordinates nutrient-dependent endoreplication with ecdysone production in the prothoracic gland

2.1 Introduction

2.1.1 Snail family proteins

The *Snail* gene was first identified in *Drosophila melanogaster*, and its function has been well-studied during embryogenesis (125). Snail is involved in establishing the mesoderm-neuroectoderm boundary (126-128) and promotes ventral cell invagination (129). Homozygous loss-of-function mutants of *snail* have a high penetrance of embryonic lethality (130,131). The name Snail refers to the U-shaped mutant embryo, which is reminiscent of a snail (132). Until now, more than 50 Snail homologs have been found in various species across the metazoans (122,133). The complex evolutionary history of this protein family has been discussed in several reviews and will not be described in my thesis (122,134-136). The general idea is that three *Drosophila* snail genes (*snail*, *escargot* and *worniu*) arose from the metazoan ancestral *snail* by tandem duplications (134,136). In vertebrates, there are also three *snail* genes (*snail1*, *snail2*, aka *slug* as well as *snail3*) as the outcome of unrelated duplication events. Snail1 and Snail2 are probably closer to each other than Snail3 since they split in a later event (136).

2.1.2 Protein structures and structural motifs

The Snail proteins are C2H2 zinc finger transcription factors that contain a conserved C-terminal region where the zinc finger DNA binding domains are located (125,137). The more divergent N-terminus is important for all the regulatory functions, such as mediating fine control of their transcriptional factor activities as well as posttranscriptional regulation of Snail protein levels (138-141). I used the software T-coffee for protein alignments and the webserver Ident and Sim (http://www.bioinformatics.org/sms2/ident_sim.html) for calculating the protein sequence similarity (142) and I found that human Snail1, Snail2 and Snail3 sequences exhibit only 34% , 37% and 32% similarity to *Drosophila* Snail (Sna), respectively. In contrast, the more conserved C-terminal region (zinc fingers) of human Snail1, 2 and 3 proteins have about 62%, 77% and 71% similarities to *Drosophila* Sna, respectively (Fig. 2-1). Putting this into context, the *Drosophila* Sna sequence was only 74% identical to another *Drosophila* Snail family protein, Worniu (143).

Several domains have been identified in vertebrate Snail proteins at the N-terminus. First, the SNAG domain, which is responsible for recruiting co-repressors, is present in all vertebrate Snail proteins (144-147). Consistent with this finding, Snail proteins usually function as transcriptional repressors (139,148-153). Moreover, vertebrate Snail2 contains a specific 28

amino-acid SLUG domain of unknown function (145). However, neither SNAG nor SLUG domains were found in *Drosophila* Sna. Regardless of this fact, *Drosophila* Sna also functions as a repressor and recruits co-repressors Ebi and CtBP at the N-terminus (Fig. 2-1) (132,154,155). A recent study showed that *Drosophila* Sna can also serve as a potentiator of active transcription (156), suggesting that it could also be a transcriptional activator (157).

2.1.3 Cellular functions in development and pathology

Despite the sequence divergence at the N-terminus, the cellular functions of the Snail family proteins are quite conserved. A well-known function of Snail proteins is to regulate the epithelial to mesenchymal transition (EMT) by transcriptionally repressing the expression of the adhesion gene *E-cadherin* (138,146,152). EMT is a process in which the polarized, immotile epithelial cells lose adherent properties and gain migratory abilities. This is an essential process during development (for gastrulation and tube morphogenesis) as well as in pathological conditions such as tumor metastasis (122,133). Therefore, Snail proteins are mainly studied with regard to their roles in embryogenesis, especially in *Drosophila*, as well as in cancer research in humans. Nevertheless, the property/fate of cells changes dramatically during EMT; therefore Snail proteins are also thought to be cell fate determinants. In line with this, *Drosophila* Sna determines the embryonic mesoderm cell fate by repressing some neuroectoderm genes in Sna expressing cells, thus setting the boundary between mesoderm-neuroectoderm (127,128,158). In other words, as transcription factors, Snail proteins essentially influence cell properties by regulating a group of genes that lead to certain cell fates. Therefore, the key to study Snail functions, in my opinion, is to identify its direct and indirect targets in a tissue-specific manner.

Various studies on human and mouse Snail1 and Snail 2 revealed that Snail proteins repress pro-apoptotic and cell cycle genes, including *proliferating cell nuclear antigen (PCNA)*, suggesting a function in cell survival and proliferation (159-168). This is of great importance for oncology since Snail was shown to be upregulated in cancer cells, conferring resistance toward radiotherapy- or chemotherapy-mediated cellular stress (166,169). Hence Snail could be a therapeutic target of cancer. In fact, resistance to p53-mediated apoptosis is one of the capabilities of stem-cell like cells. Indeed, the *Drosophila* snail family gene *escargot (esg)* is commonly used as a molecular marker for intestinal stem cells and is the key regulator for maintenance of stemness (170). Last but not least, *esg* was shown to prevent endoreplication in imaginal disc cells to

maintain diploidy (171,172) and a similar result was found in mouse trophoblast giant cells with a murine homolog of Snail zinc finger transcription factor (173).

2.1.4 *Drosophila sna* is essential for embryonic ring gland development

Sanchez-Higueras and colleagues demonstrated that *Sna* is essential for *Drosophila* embryonic RG development (174). Specifically, *Sna* is expressed in the corpus allatum (CA) and the prothoracic gland (PG) primordium of the embryo. By stage 11, the *Drosophila* CA and PG are two separated cell patches and would eventually migrate and fuse together in about a 2-hr time window by stage 15 to form the RG. The fusion of the two glands requires *Sna*-mediated EMT, evidenced by the fact that loss-of-*sna* in the CA and the PG primordium abolished the coalescence of the two glands (174). More intriguingly, tracheal and endocrine (CA and PG) primordium tissues originate from common precursor cells with the characteristic expression of *vvl*, loss of which resulted in degradation of these common precursor cells. The subsequent activation of *Sna* leads to the CA and PG cell fate, while activation of *Tracheless* (*Trh*) and *Tango* (*Tgo*) determines the fate to a respiratory primordium. In other words, *Vvl* is crucial for RG development during embryonic stages but also has later functions, as it is essential for ecdysone production during larval stage (74,75, also see Chapter 1.7). Similarly, our lab showed that *Sna* is still highly expressed in the PG during the larval stage, suggesting that it might be crucial for ecdysone production later on.

2.2 Materials and Methods

2.2.1 Fly stocks and fly crosses

Drosophila melanogaster was maintained on a standard agar-cornmeal medium at 25°C. *phm22-Gal4* (on 3rd chromosome), *phm22-Gal4/CyO* (on 2nd chromosome), *Smad2*-RNAi and *spok^{GS}-Gal4* were obtained from Dr. Michael B. O'Connor's lab. *UAS-sna*-RNAi (#50003), *UAS-broad*-RNAi (#104648), *UAS-timeless*-RNAi (#101100) and *UAS-Ras85D*-RNAi (#106642) were ordered from the Vienna *Drosophila* Resource Center (VDRC). *UAS-Cas9*-cDNA (#54595), *UAS-Dicer2*-cDNA (#24650), *UAS-TOR^{DN}* (#7013), *UAS-TOR*-RNAi (#33627), *UAS-raptor*-RNAi (#41912), *UAS-Pi3k92E*-RNAi (#27690), *UAS-Akt1*-RNAi (#31701) and *UAS-cycE*-cDNA(#30725) were ordered from Bloomington stock center and *UAS-escargot*-cDNA (#109127), *UAS-sna*-cDNA (#109121) were obtained from Tokyo Stock Center. *UAS-Pten*-

cDNA was obtained from FlyORF (#F001338). *hsFLP; tub-FRT-CD2- FRT-Gal4, UAS-GFP* was a kind gift from Dr. Hwei-Jan Hsu.

2.2.2 Rescue by feeding ecdysone-supplemented diets

For the 20-Hydroxyecdysone (20E) feeding rescue experiments, a 20E stock solution at a concentration of 10 mg/ml was made in ethanol. The 20E solution was then added to the agar-cornmeal medium before it was solidified to reach the final concentration of 20E at 333.3 ng/ml (3.33% ethanol). Control food was prepared similarly with only 3.33% ethanol. Sixty larvae were transferred to each vial, and allowed to develop at 25°C, after which the phenotype of the larvae was scored.

2.2.3 RNA extraction from dissected tissues and cDNA synthesis

Ring glands were dissected from larvae at desired time points in 1x PBS buffer and put immediately into TRizol reagent (Ambion, Life Technologies) in a 1.5 ml Eppendorf tube. The tube was then flash frozen in liquid nitrogen for short-term storage. Upon total RNA extraction, tissues were homogenized with a pestle presoaked in 1% SDS. RNA was then isolated by phenol-chloroform phase separation; the aqueous phase (contains the RNA) was transferred to a fresh RNA free Eppendorf tube with the addition of an equal volume of 70% ethanol. Next, the mix was subjected to the QIAGEN RNeasy spin column (RNeasy mini kit) and further purified following the manufactory instructions. The concentration and integrity of the RNA samples were then assessed using Agilent RNA 6000 Nano chip (RNA 6000 Nano kit) run on a 2100 Bioanalyzer Instrument.

50-100 ng of total RNA (for different experiments) was used for cDNA synthesis using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, ABI, Catalog number: 4368814) following manufacturer's instructions.

2.2.4 Quantitative Real-time PCR (qPCR) analysis

Primer validation

qPCR primers were designed using the web-based Universal ProbeLibrary Assay Design Center (Roche Life Science) and ordered from Integrated DNA Technologies (IDT). All primers used for the Sna study are listed in Table 2-1. The primer mix containing the forward and reverse primers was prepared at a final concentration of 3.2 μ M. To test the efficiency and specificity of

the amplification for each primer pair, the primers were verified using serial dilutions of stock cDNA, reverse-transcribed from 1 µg total RNA of *w¹¹¹⁸* larvae. In brief, the cDNA was diluted by 1:4, 1:16, 1:64, 1:256 and 1:1024 and for each dilution the reactions were set up with 5 µl of 2x SYBR Fast qPCR Master Mix (Kapa biosystem, Catalog# KK4601), 2.5 µl primer mix and 2.5 µl of cDNA in triplicate. The reaction was performed on a QuantStudio™ 6 Flex Real-Time PCR System (ABI) using the standard curve mode. The standard curve and the melting curve were analyzed using the QuantStudio™ software. The slope of the standard curve for the gene of interest (GOI) should be the same as the slope of *rp49*, a housekeeping gene. A single peak in the melting curve would indicate a specific binding of primer to the template.

qPCR reaction

cDNA samples were diluted at a 1:8 ratio for 50-100 ng of RNA input during cDNA synthesis. For all the experiments, three biological replicates were included, and three technical replicates were used for each biological sample. The qPCR reaction was set up in the same manner as the primer validation. The reactions were then run using the comparative C_T (ΔΔC_T) mode with the following thermocycling parameters: Step 1: 95°C for 20 seconds (s); Step 2: 95°C for 1 s and Step 3: 60°C for 20 s. Step 2 and three were performed for a total of 40 cycles.

Fold change determination

Through the qPCR reactions, the C_T (threshold cycle) value was determined for individual samples with each primer pair. The relative expression level of the GOI (represented by ΔC_T) is determined by the difference between its C_T value and the C_T of the endogenous control gene, i.e. *rp49*, for each corresponding cDNA sample. ΔΔC_T was then calculated using the ΔC_T of the experimental condition subtracting by the ΔC_T of the control condition for each every GOI. Since samples for both experimental and control conditions were collected in triplicate, three ΔC_T values were generated for each condition, which resulted in 3 x 3 (=9) ΔΔC_T and the average of the nine values was taken in the end. Finally, the fold change between experimental and control conditions was calculated as $2^{-\Delta\Delta C_T}$.

Table 2-1. qPCR primers used in the *snail* study

Primer Name	Primer Sequence
<i>snail</i> -#39_F	CGACGAGTGCCAGAAGATG
<i>snail</i> -#39_R	GTGGAAGTACGGTGCTTG

<i>snail</i> -NCBI-F	GGATTACCTGCCACCGAAA
<i>snail</i> -NCBI_R	TCACAAAGGCGGACTGGAAG
<i>torso</i> _F	GCCTGCAGAACTTTTTACGTG
<i>torso</i> _R	TGTCCACGTTCTGTTCAAGG
<i>neverland</i> _F	CCCTCACCTAGGAGCCAAC
<i>neverland</i> _R	GGCATATAACACAGTCGTCAGC
<i>spookier</i> _F	GCGGTGATCGAAACAAC
<i>spookier</i> _R	CGAGCTAAATTTCTCCGCTTT
<i>shroud</i> _F	CGAATCGCTGCACATGAC
<i>shroud</i> _R	TAGGCCCTGCAGCAGTTTAG
<i>phantom</i> _F	GGCATCATGGGTGGATTT
<i>phantom</i> _R	CAAGGCCTTTAGCCAATCG
<i>disembodied</i> _F	GTGACCAAGGAGTTCATTAGATTC
<i>disembodied</i> _R	CCAAAGGTAAGCAAACAGGTTAAT
<i>shadow</i> _F	CAAGCGGATATTTGTAGACTTGG
<i>shadow</i> _R	AAGCCCACTGACTGCTGAAT
<i>eiger</i> #158-F	CGACGAGTTCCAAAAGGAGT
<i>eiger</i> #158-R	GTCGTCGTCCTCCTCATCC
<i>Traf4</i> #154-F	CCACTCCTGGGAACAACAAC
<i>Traf4</i> #154-R	ATCGGGACCGGGATAGATAA
<i>Dcp-1</i> #119-F	TGTACGCCGGAGTCTCTTGT
<i>Dcp-1</i> #119-R	TATCGTTCCACGGGCATT
<i>reaper</i> _F	TCGATTTCTACTGCAGTCAAGG
<i>reaper</i> _R	GAGTAAACTAAAATTGGGTGGGTGT
<i>hid</i> _F	TTCTGCCCTCTTTCTTTG
<i>hid</i> _R	GTCCTTATCCGCTTCCTTCC
<i>EcR</i> _F	ACCAGCGTTTACAAAGATACCC
<i>EcR</i> _R	ATCACCTCCGACGAGCAG
<i>escargot</i> _F	CCGGATTGCCAGAAATCTTA
<i>escargot</i> _R	ATGGAAGTCTGATGTTTGGT
<i>damm</i> #114-F	GGCGACCGAAAAGAGAAAAT
<i>damm</i> #114-R	AAGGACGTCATCGTCCAAGT
<i>period</i> -F	TGAGAGCGAGAGCGAGTGTA
<i>period</i> -R	CCATGGTGCTTAGGTTCTCC
<i>timeless</i> _F	CCCTTATACCCGAGGTGGAT
<i>timeless</i> _R	TGATCGAGTTGCAGTGCTTC
<i>erk7</i> #19-F	AATCCTGCCTCCTTTTACTGC

<i>erk7#19-R</i>	TTCTAGTTGGCCCGGTTG
<i>hairy-#104-F</i>	GCGTAACAGCAGCCAACAT
<i>hairy-#104-R</i>	CATGATGGGCTTGTTTCGAC
<i>trunk#134-F</i>	CACTTTGCGCTGGTAGAGAA
<i>trunk#134-R</i>	CTTCTCGTGCGCAGACTCTT
<i>torso-like#132-F</i>	TTTGAGTCTTCGCGCTTGA
<i>torso-like#132-R</i>	TCGCTCAGCATTCAAGGTTA
<i>SNCF_F</i>	AGACGACGCTGTAACCTGTG
<i>SNCF_R</i>	AATAATGCCCCATCCCAGCC
<i>ouiji board-#77-F</i>	TTTCTCTGCTGGCGAACT
<i>ouiji board-#77-R</i>	AAACGGTTGCTCCCCAGTAT

2.2.5 20E titer measurement

Ecdysteroid titers of whole larvae were measured following the Manufacturer's instructions of the 20-Hydroxyecdysone Enzyme Immunoassay kit (Bertin Pharma #A05120.96 wells). In brief, eight larvae were collected per sample and homogenized in 400 μ l of methanol with a motorized pestle to extract the ecdysteroids. Supernatants were transferred into a new tube after centrifuging at the maximal speed (16,100 g). Another round of extraction was done with the old lysate in 400 μ l of methanol followed by one more extraction with 400 μ l of ethanol. All the extracts were pooled together (1.2 ml in total) and dried by Speed-Vacuum. Next, the samples were re-dissolved in 110 μ l of EIA buffer (provided by the kit) for 2 hr at room temperature (RT) or overnight at 4 degrees. The 96 well plate was pre-washed with wash buffer (provided by the kit), after which 50 μ l per sample, as well as control, were loaded onto the plate in duplicate (provided by the kit). The plate is incubated at 4 degrees overnight and the plate was then emptied and washed with 300 μ l wash buffer. At the end the assay was developed with 200 μ l of Ellman's reagent and incubated in the dark at RT on shaker. After 75 min, results were obtained at a wavelength of 410 nm using the Synergy H1 microplate reader (BioTek).

2.2.6 Immunofluorescence with larval tissues

Larvae were dissected at the 1/3 portion towards the head with a pair of Dumont #5 forceps (Fine Science Tools #11252-40) in 1x PBS. The cuticle of the head part was then turned inside-out to expose the tissues. With the ring gland still attached to the brain and the brain attached to the mouthhook, the other imaginal tissues were removed and the remaining tissues were

transferred into 4% formaldehyde (diluted from 16% EM Grade formaldehyde, Electron Microscopy Sciences, Catalog #RT15710) in a 0.6 ml Eppendorf tube. Tissues were fixed for 22 min at RT and then washed three times with PBST (PBS containing 0.3% Triton-X 100) for 15 min each. Next, blocking solution with 5% normal goat serum in PBST was prepared fresh every time and the tissues were blocked for 30 min at RT. Tissues were incubated with primary antibodies either at 4°C overnight or at RT for 4 hr followed by three washes (with PBST). Then the secondary antibody was added for 1 hr at RT followed by three washes (with PBST). Nuclei were stained with DAPI (1:5000). After three washes (with PBST), the ring glands were mounted onto the slides in VECTASHIELD Antifade Mounting Medium. Rabbit anti-Sna is a kind gift from Dr. Zeitlinger and was used at a 1:600 dilution. Rabbit anti-GFP was used 1:20 (Thermo Fisher Scientific, #G10362). Mouse anti-N-cadherin (1:20, Developmental Studies Hybridoma Bank). Secondary antibodies (Goat anti-Rabbit, Alexa Fluor® 488, anti-Rabbit Alexa Fluor® 555 and anti-mouse Fluor® 555) were used at a 1:500 dilution. Images were captured on a Nikon C2+ confocal microscope.

2.2.7 Cellular DNA content measurements

Cellular DNA content was assessed by DAPI staining (1:50,000). Z-stack images of the brain-RGs complexes were taken and the summation of DAPI intensity for each single pixel of all the stacks were obtained using ImageJ. Corrected DAPI intensity (CDI) in the PG area was calculated using the following formula: $CDI = \text{Summation of DAPI intensity of the selected area} - (\text{Area} * \text{Mean intensity of background readings})$. All the parameters in the formula were measured in ImageJ. Next, the CDI for the PG area was normalized to the average DAPI intensity in the brain lobe and the normalized DAPI intensity was divided by PG cell number to obtain relative DNA intensity per cell.

2.2.8 CRISPR-based tissue-specific *sna* deletion

The 20-nucleotide gRNA sequences targeting the *sna* coding region were designed with CRISPR Optimal Target Finder (<http://tools.flycrispr.molbio.wisc.edu/targetFinder>) (175). Two gRNA sequences that were ~100 bp apart from each other in the genome were selected to secure the chance of generating a double-stranded break by Cas9. The two gRNA sequences were cloned into the pCFD4-U6:1_U6:3tandem gRNAs vector (Addgene#49411) according to F. Port *et al.* (176). In brief, the following forward and reverse PCR primers were used:

fwd-sna-gRNA

5'TATATAGGAAAGATATCCGGGTGAACTTCG**TCTGCCACAAACGGAGGCCTGTTTT**
AGAGCTAGAAATAGCAAG3'

rev-sna-gRNA

5'ATTTTAACTTGCTATTTCTAGCTCTAAAAC**GAAACGGGGTCCGACGAGGCGACG**
TTAAATTGAAAATAGGTC3'

The two gRNA sequences were incorporated into the primers (red letters). The bold letters indicate the homology sequence to the pCFD4 backbone and a PCR product of 595bp was amplified using Q5® High-Fidelity DNA Polymerase (NEB) and pCFD4 as the template. Next, the pCFD4 vector was linearized by BbsI-digestion and PCR fragment was cloned into the backbone using Gibson Assembly (NEB#E2611) following the manufacturer's instructions. In principal, the PCR fragments have overlapping sequences to the linearized pCFD4 at both ends (the underlined sequence in both primers), which would allow insertion and repair of the vector during Gibson Assembly. Finally, the transgenic flies expressing the *sna* pCFD4-U6:1_U6:3tandemgRNAs were generated via φC31-mediated integration through the service from Bestgene Inc.

UAS-Cas9 (Bloomington#54595) was stably combined to gRNA transgenic flies (tandem *sna*-gRNAs; UAS-Cas9). Flies were then crossed with the PG-specific-Gal4 lines to drive Cas9 expression only in the PG. The two gRNAs will direct the Cas9 enzyme to the *sna* locus and generate two double stranded breaks close to each other; upon which the cell will attempt to repair the double strand break through NHEJ, which should result in accumulation of small deletions at *sna* locus.

2.2.9 Genomic extraction of ring glands and sequencing

Genomic DNA was extracted from ring glands using the NucleoSpin Tissue kit (MACHEREY-NAGEL#740952.50). 25 ring glands were dissected in 1x PBS and put directly into 180 µl of Buffer T1. The manufacturer's instructions were then followed. DNA concentration was measured using the Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific #Q32854). To test for ring gland-specific CRISPR deletions, a fragment of 854 bp spanning the predicted gRNA

target sites was amplified using the Phusion® High-Fidelity DNA Polymerase (NEB) with the two following primers:

Fwd: TACCTGCCCACCGAAATCC

Rev: ATGTCGCTCCTGAACAAGCA

The PCR fragment was then blunt-end-ligated into the SmaI-digested pUC19 vector. The positive clones were identified by blue-white selection on the carbenicillin-containing plates. In brief, bacteria with self-ligated pUC19 will have functional β -galactosidase. When X-gal (a dye-linked substrate for β -galactosidase) and IPTG (an inducer for β -galactosidase) were added to the transformation plate medium, the bacteria can metabolize X-gal and the colonies appear blue. For this study, 20 μ l of 200 mM X-gal and 40 μ l of 20 mg/ml IPTG solution were spread onto the surface of the medium and let dry in the fume hood for 30 min. After transformation, bacteria were incubated on the plate for 16 hr at 37°C. Colonies with a successful insertion in pUC19 will have a non-functional β -galactosidase and maintain a whitish-cream color. Therefore, white colonies were picked and submitted for sequencing in the Molecular Biology Service Unit (MBSU) in the Department of Biological Science.

2.2.10 Animal staging

Parental lines were mated in cages that were capped with agar-grape juice plates with yeast paste at the center of the cap. On day 0, the cages were left upside down at 25°C for 8 hr to allow egg laying onto the fresh caps. From the caps, 100 eggs were then transferred to each fresh plate with standard medium (35 mm in diameter) and put back to 25°C to allow development. On day three the developmental stages of the larvae were determined by tracheal morphology (177,178) and the unstaged third instar larvae were picked out and discarded. Plates were then put back to 25°C and newly molted L3 were transferred to fresh plates every hour. Those larvae were reared at 25°C until the desired developmental stage (hours after second to third instar molt) for experiments. For staging second instar larvae, the procedure was the same but carried out on day two after egg laying.

2.2.11 Starvation protocol for critical weight determination

Whether a larva had passed the critical weight checkpoint or not was determined by whether the larva could pupariate when deprived of food. Briefly, larvae were staged at the L2/L3

molt as described. A population of larvae with the same developmental age was transferred to medium with only 2% agar in water and the fraction of pupariation was scored in the following two days.

2.2.12 EdU incorporation assays

EdU incorporation assays were carried out using Click-iT EdU 555 Imaging Kit (Life Technologies). Larvae were dissected in 1x Ringer's solution and were incubated for 30 min at RT with 10 μ M EdU. Tissues were then fixed in 4% PFA in PBS for 25 min followed by two brief washes in 0.3% PBST and then washed again twice for 20 min each in 0.3% PBST. Blocking was performed with 1% bovine serum albumin in 0.3% PBST for 30 min, and tissues were then incubated with Click-iT reaction cocktail (prepared according to manufacturer's instructions) for 30 min at RT. Tissues were washed in 0.3% PBST and incubated with Hoechst 33342 (Life Technologies, included in the kit) at a 1:1500 dilution in 0.3% PBST for 5 min. Normal antibody staining could be performed following the EdU incorporation procedure to label other proteins according to the protocol previously described in section 2.2.6. 1x Ringer's solution was prepared by adding 0.33 g $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 13.6 g KCl, 2.7 g NaCl and 1.21 g Tris base to water to make 1 L solution. The pH was adjusted to 7.2 with 1 N HCl and sterilized by autoclaving.

2.2.13 Generation of *sna*-overexpressing clones by FLP-out system

Virgin females of *y, w, hsp70-FLP; tub-FRT-CD2- FRT-Gal4/TM6, Hu, Tb* genotype were crossed to $y^{[1]} w^{[67c23]}$; *sna*-cDNA males (#109121 from Tokyo Stock Center) or $y^{[1]} w^{[67c23]}$ males (control). The progeny of the two crosses was collected and reared on standard agar-cornmeal medium at room temperature (RT) until animals reached the early L1 stage, at 40 hr after egg laying (AEL). Larvae were then subjected to a heat shock at 37°C for 40 min followed by 1 hr of recovery at RT. Recovered animals were allowed to develop at 25°C for four days and the non-*Tb* larvae were dissected at the end of the larval stage to examine the RGs.

2.2.14 Heat shock induction of *sna* overexpression

hs-Gal4 (Bloomington #2077) was crossed with $y^{[1]} w^{[67c23]}$; *sna*-cDNA lines and were kept at RT since the heat shock line had leaky Gal4 expression and resulted in early lethality of animals when *sna* was overexpressed. The progeny was heat-shocked at 17 hr after the L2 to L3 molt for 45 min in a 37°C water bath and allowed to recover for 6 hours at 25°C before RNA

extraction for RNA-Seq experiments (section 2.2.15). *hs-Gal4> y¹¹ w^[67c23]*, which served as controls were treated in the exact same manner.

2.2.15 Next-generation RNA sequencing (RNA-Seq) analysis

60 ring glands were dissected for each sample and ring gland RNA was extracted. Each condition was tested with two biological replicates. The DNase (supplied by the RNAeasy mini kit) digestion step was administered during the RNA extraction according to the manufacturer's instructions to eliminate any genomic DNA. After the extraction, the purity of the RNA was assessed by the ratio of the absorbance at 260 nm and 280 nm using NanoDrop ND-1000. Samples with an A260/A280 value below 1.8 were discarded. In addition, the integrity of the total RNA was assessed by Bioanalyzer as described in section 2.2.4. Lastly, RNA concentration was measured using the Qubit RNA HS Assay Kit (Thermo Fisher Scientific #Q32852) to ensure that the amount of RNA input was between 10-100 ng for cDNA library construction. RNA input outside this range would result in sub-optimal cDNA libraries. In this study, 30 ng of total RNA was used for each sample.

cDNA libraries were constructed using the Ovation® Universal RNA-Seq kit for *Drosophila* (NuGEN#0350) following the manufacturer's instructions. Briefly, first strand cDNAs were synthesized from total RNA using a mixture of random and poly(T) primers (provided by the kit), followed by degradation of RNAs in DNA/RNA heteroduplex by RNase H. Second strand cDNAs were then generated with the addition of degradable nucleotide analogs, in which the sense strand was labeled with the degradable nucleotide analogs and was utilized in the subsequent steps for degradation to achieve the strand-specific RNA-seq data. Next, the double-stranded cDNA was fragmented into a median size of 200 bp using a Covaris S-series System (see Table 2-2 for detailed setting parameters) followed by end repair to generate blunt ends.

The barcoded adaptors were then blunt-ligated to the fragments to allow cluster formation and sequencing on an Illumina platform. In total, 16 unique sequence barcodes were provided which allowed for 16 multiplex sequencing reactions, as long as each sample was labeled with a unique barcode. After adaptor ligation, the labeled cDNA strand was degraded. Note that one strand of the forward adaptors was labeled with the same degradable nucleotide analog to ensure that the inserts between the adaptors were orientated in a uniform direction because inserts in the

other direction would have the forward adaptor degraded and only the inserts (now single stranded) with both the forward and reverse adaptor sequences could be sequenced.

Following strand selection, Insert-Dependent Adaptor Cleavage (InDA-C) step was performed to eliminate unwanted transcripts, in this case, ribosomal RNA. In principal, the library is incubated with gene-specific primers that target rRNA and primer extension into the reverse adaptor generates a cleavage site in reverse adaptor when it becomes double-stranded. In the addition of the cleavage reagent, the reverse adaptors of the specific inserts will be cut and making them non-amplifiable. In the end, the libraries were amplified by PCR.

The final library for each sample was analyzed by an Agilent Bioanalyzer DNA1000 chip and the software can estimate the average of the library size distribution, which ideally should be around 275 bp. Then the libraries were quantified using Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific #Q32850).

Table 2-2. Covaris S-Series Sonicator system settings

Parameter	Value
Duty Cycle	10%
Intensity	5
Cycles/Burst	200
Time (s)	180
Temperature (Water Bath)	6°C-8°C
Power Mode	Sweeping
Frequency	
Degassing Mode	Continuous
Sample Volume	120 µl
Water (FILL/RUN)	S2 – level 12, E210 – level 6
AFA Intensifier	Yes

Adapted from Encore® complete RNA-Seq library systems user guide

Next generation sequencing was performed by Delta Genomics on an Illumina Hi-Seq 2500 platform. Raw data were analyzed by Arraystar 4.0 (DNASTAR) to create individual files for

each sample. Data were analyzed with Arraystar 4.0 (DNASTar) as well as Microsoft Access. Gene ontology statistics was performed with DAVID (179,180).

2.3 Results

2.3.1 *Snail (sna)* is specifically and dynamically expressed in larval ring glands

Based on the microarray analysis conducted in our lab, *sna* is one of 108 ring gland-enriched transcripts (120). Fig. 2-2A summarizes the microarray signals for *sna* where the signals from the whole body stay constantly low (~100, arbitrary units), while the RG signals were >30-fold higher (~3600) at least at two early time points (4 hr and 8 hr L3). Around 24 hr L3, *sna* transcript levels in the RG started to decrease and at 36 hr L3, microarray signals of *sna* in the RG dropped to levels comparable to that in the whole body. This suggested that *sna* mRNA is not only abundantly present in the RG but is also dynamically regulated throughout development which reinforced the hypothesis that *sna* has an important function in this tissue.

I next asked whether Sna protein levels follow the same trend as the mRNA profile. In the microarray analysis, RGs were used, which consist of three small glands (Fig. 1-2) that cannot be separated by dissection. Therefore, I needed to establish in which of the three glands Sna is present. For this purpose, I took advantage of the transgenic line developed in A. Stathopoulos' lab, which contains a 25 Kb genomic construct with a C-terminal GFP-tagged *sna* as well as all known *cis*-regulatory elements of *sna* (181). This genomic construct was shown to be able to rescue embryonic lethality in the *sna* mutant (181), which suggested that *sna-gfp* could recapitulate the essential aspects of endogenous *sna* function. I then performed immunofluorescence using GFP antibody to monitor *sna* expression in the RG throughout larval stages (Fig. 2-2B-D).

There are multiple conclusions I can draw from the results of these experiments. First, during L3 stage, *sna* protein levels indeed follow a similar trend as *sna* mRNA levels revealed by microarrays. This also confirmed the validity of the *sna-gfp* lines. Specifically, the GFP signals start to rise at 4 hr L3; then start to decrease at 24 hr and become almost undetectable around 34 hr L3. The decline of Sna-GFP levels slightly lagged the decline of *sna* mRNA levels probably because: 1) transcription and translation do not happen at the same time. 2) Sna-GFP might have a longer half-life than endogenous Sna (see section 2.3.9). Secondly, there are time points that GFP is exclusively present in the nucleus, which fits well with the fact that Sna is a transcription factor. Thirdly, GFP is not uniformly present in the tissue and this is not an artifact caused by the *sna-gfp* transgene as I verified the expression pattern in an independent experiment using the anti-

Sna antibody in control (*w¹¹¹⁸*) flies (see Fig. 2-8 and Fig. 2-27). It seems that *sna* is selectively turned on in some cells at a time, which I later determined to be related to endoreplication (see section 2.3.3 and 2.3.5). Moreover, Sna is also present during the late L2 stage, starting at 14 hr with a peak at ~18 hr L2; shortly thereafter, Sna becomes undetectable at 21-22 hr L2. Around 18 hr L2, Sna is present in all nuclei in the PG for only a short time window, as opposed to other time points where only some nuclei showed Sna-GFP signals (also see Fig. 2-8A for Sna antibody staining). The only other time point at which the unified expression pattern was observed is around 22 hr L1 stage. However, the biological meaning of the unified expression pattern at these two time points is still not clear. Furthermore, Sna-GFP signals were mainly observed in the PG but rarely captured in the CA (Fig. 2-2) which was later also verified by Sna antibody staining (Fig. 2-8A). In Fig. 2-2, the PG and CA are outlined by the white-dotted line. PG and CA cells could be distinguished by DAPI staining at least during L2 and L3 stages because of the larger size of PG nuclei. The corpora cardiaca (CC) was not included in the outlined area because the *sna-gfp* line also carries a M{3xP3-RFP.attP}ZH-86Fb construct (for transgene insertion purposes). The expression of this construct caused intense RFP expression in the CC that resulted in bleed-through signals in the green channel. Therefore, it is impossible to decide whether there is GFP expression also in the CC. However, I later showed by Sna antibody staining (Fig. 2-8A) that Sna is not present in the CC. Lastly, I observed Sna-GFP signals in other larval tissues, indicating some *sna* expression in larval brain hemispheres and parts of the eye discs (data not shown). However, Sna was not detected in the eye discs by antibody staining while the presence of Sna in the brain was confirmed by Sna antibody staining (see Fig. 2-28). This suggested that Sna may also have function in the larval brain, which could be further investigated in the future (section 2.3.7.5). In summary, these results suggested that *sna* is dynamically expressed in the PG during larval development. Since the PG is the gland where ecdysone is made, *sna* may be a new player in regulating ecdysone production.

2.3.2 Loss-of-*sna* in the PG affected ecdysone production

2.3.2.1 Loss-of-*sna* in the PG via RNAi resulted in larval developmental arrest

To test whether *sna* is indeed functionally important in PG cells, I knocked down *sna* specifically in the PG using the UAS/Gal4 binary system. For this study, PG-specific Gal4 (*phm22-Gal4*) was used and *UAS-Dicer2* (*Dicer-2* encodes the endoribonuclease Dicer, which cleaves double-stranded RNA during RNAi) was also introduced together with the *UAS-sna-*

RNAi (VDRC #50003) construct to enhance the efficiency of the knock down. The controls were *UAS-Dicer2; phm22-Gal4* crossed to *w¹¹¹⁸* (a reference strain). As a result, knocking down *sna* in the PG caused larval arrest where ~20% of the population were arrested as L2 and ~80% arrested as L3 (Fig. 2-3A). A closer inspection of the arrested L2 larvae revealed that they experienced molting defects where the animals developed some characteristics of third instar larvae, e.g. having both a 2nd instar mouth hook and a larger 3rd instar (L3) mouth hook structure (double mouth hooks); however, they did not complete the molt to L3 (Fig. 2-3B). The arrested L3 larvae never triggered wandering behavior and kept feeding, which resulted in continuous growth and giant larvae compared to controls. The above-mentioned phenotypes typically result from the lack of ecdysone pulses during the transitions between different developmental stages (4, 15, 26-28).

As mentioned previously, *Sna* function is required for embryonic RG development. To show that the developmental arrest phenotype was caused by a disruption of *sna* in the larval stage, rather than loss-of-*sna* in the embryonic stage, I used the inducible Gal80 temperature-sensitive (Gal80^{TS}) system to turn on *sna*-RNAi in the PG in post-embryonic stage (182). The general principle of the temperature-inducible system is that Gal80 works as a Gal4 repressor, therefore there would be no *UAS-sna*-RNAi expression when Gal80 is functional. The Gal80^{TS} allele produces a functional Gal80 at 18°C, while a switch to $\geq 29^\circ\text{C}$ would inactivate Gal80, allowing Gal4 to turn on *UAS-sna*-RNAi expression. Therefore, I first raised the *tub-Gal80^{TS}; PG>sna*-RNAi embryos at 18°C. After 48 hr, when all the embryos had hatched as L1 larvae, I switched the temperature to 30°C (Fig. 2-4A). With this procedure, the *sna*-RNAi animals still experienced L3 arrest (Fig. 2-4B). As a control, I showed that when *tub-Gal80^{TS}; PG>sna*-RNAi animals were kept at 18°C throughout the life cycle, no developmental arrest was observed. These animals were able to develop into adults, showing that Gal80 suppressed RNAi expression at 18°C. There was also no effect of temperature shift on development when there was no Gal4 involved (*UAS-sna* RNAi/+; *UAS-Dicer2*/+; *tub-Gal80^{TS}*/+) (Fig. 2-4B). Taken together, my results suggested that *sna* has a role in the larval PG in regulating ecdysone production.

2.3.2.2 Larval arrest phenotypes in PG>*sna*-RNAi animals are caused by ecdysone deficiency

The phenotypes observed in the PG>*sna*-RNAi animals indicated ecdysone deficiency; therefore, I directly measured ecdysteroid levels in whole larvae. The kit I used mainly detects the biologically active form of ecdysone (20E). The developmental time point I chose for the 20E

measurement was 28 hr after L2 to L3 molt, when the 3rd minor ecdysone pulse occurs during L3 in controls (99). Indeed, a reduced 20E titer was observed in PG>*sna*-RNAi animals compared to that in controls (Fig. 2-5A). Next, I asked whether the lack of ecdysone was caused by a disruption in the expression of ecdysone biosynthetic enzymes. For that, I carried out qPCR analysis using RGs of PG>*sna*-RNAi; *Dicer2* and control animals both staged at 24 hr after L2 to L3 molt. First, I measured the expression of *sna* itself. I found that *sna* transcript levels were reduced to 28% of the control, which indicated that the dsRNA construct was functional. For the six ecdysteroidogenic genes I examined (*nvd*, *sro*, *spok*, *phm*, *dib* and *sad*), all of them were drastically downregulated in *sna*-RNAi. It is striking that *nvd*, *spok* and *sad* were more than 10-fold reduced in their expression (Fig. 2-5B). These results suggest that maximal expression of ecdysone biosynthetic genes and subsequently the production of ecdysone is dependent on Sna function in the PG.

If the lethality and developmental arrest seen in PG>*sna*-RNAi animals was merely caused by a lack of ecdysone, I expected to see a rescue when 20E is supplied to the animals. Such rescue experiments are routinely done in ecdysteroids studies by adding 20E into the food. I observed that by the time most of the control animals (either with or without 20E) reached the adult stage, *sna*-RNAi animals without 20E (ethanol, the vehicle alone) were still feeding in the food and they never started wandering behavior and never pupariated (Fig. 2-6A and B). Strikingly, ~50% of *sna*-RNAi L3 started wandering behavior when 20E was provided (Fig. 2-6 A and B). Some of these wandering larvae could even form pupae, which accounted for 15.4% ($\pm 6.5\%$) of the starting population (Fig. 2-6A), although they were not healthy enough to develop into adults (Fig. 2-6C). However, I did not observe any improvement of L2 lethality in *sna*-RNAi animals by 20E feeding (Fig. 2-6A). Therefore, I concluded that 20E partially rescued the PG>*sna*-RNAi phenotype.

2.3.2.3 Developmental arrest caused by PG>*sna*-RNAi cannot be rescued by activating PTTH/MAPK signaling in the PG

As discussed in the introduction, the PTTH/MAPK pathway is the best studied pathway in the PG and has an essential role in triggering ecdysone production in response to a brain-derived signal at the proper developmental time (33). Therefore, I also measured the expression of *torso*, the receptor for PTTH. I found that levels of *torso* were also reduced to 38% of the levels in controls although it was not statistically significant due to the discrepancy among the three

replicates in *sna*-RNAi samples (Fig. 2-5B). However, trend of *torso* downregulation in PG>*sna*-RNAi ring glands was later confirmed by RNA-Seq analysis (Table 2-3, Fig. 2-36).

The disrupted *torso* expression in the *sna*-RNAi animals suggested reduced PTTH/MAPK signaling. Therefore, I wondered if reactivating PTTH/MAPK signaling by overexpressing the *torso* cDNA would rescue the *sna*-RNAi phenotype. However, expression of the *torso*-cDNA along with *sna*-RNAi in the PG resulted in 100% L1 arrest (data not shown). This is because too much Torso in the PG itself affects development, which is evidenced by the fact that *phm22>UAS-torso*-cDNA; *UAS-EGFP* (*UAS-EGFP* is included to make up for the same number of UAS-binding sites for all the conditions) also led to 100% L1 arrest (data not shown). Hence, it was impossible to use *torso*-cDNA in this case. Since Ras acts downstream of Torso, I then tried to express a constitutively active form of Ras, *Ras^{V12}*, to hyperactivate the PTTH pathway in *sna*-RNAi animals to see whether there would be a rescue. As shown in Fig. 2-7 expressing *sna*-RNAi along with *UAS-EGFP* in the PG again caused L3 arrest, with less than 10% of escapers reaching pupal and adult stages. The escapers were observed probably because *UAS-EGFP* has occupied some Gal4 protein, diluting the induction of *UAS-sna*-RNAi. Overexpression of *Ras^{V12}* did not rescue the L3 arrest phenotype and even fewer pupae and adults were seen (less than 1%) (Fig. 2-7). This result confirmed that the function of Sna in regulating ecdysone production is not predominantly through the PTTH/MAPK pathway. Sna is probably involved in other biological processes that are linked to ecdysone production.

2.3.2.4 PTTH is not required for *sna* expression in the PG

My previous results indicated that *sna* function in the PG is essential for proper ecdysone production. Therefore, I investigated how exactly *sna* function is related to ecdysone. My first attempt to answer this question was to test whether *sna* expression in the PG is regulated by pathways or players that are already known to be involved in ecdysone production. I first asked whether *sna* expression in the PG is dependent on the PTTH/MAPK pathway. Therefore, I knocked down *torso*, the receptor of the pathway in the PG and examined Sna levels by antibody staining against endogenous Sna. I started testing Sna levels around 17 hr L2, when Sna is present in all PG nuclei. Consistent with my findings from the *sna-gfp* line, control animals (*phm22>w¹¹¹⁸*) also showed a peak of Sna protein at this time point (Fig. 2-8A). When *torso* was knocked down via RNAi, this peak of *sna* expression was not abolished, but just delayed (Fig. 2-8A). This could be explained by the fact that PG>*torso*-RNAi animals were already developmentally delayed

during L2 stage by ~4 hr (33). With this in mind, I predicted that *sna* expression was not affected by the loss-of-*torso* function *per se*. I then looked at 0 hr L3 since the L2 to L3 molt is a definite developmental hallmark when *torso*-RNAi and control animals are developmentally at the same stage. For this reason, I mainly focused on 0 hr L3 when I examined *sna* expression in experiments in the following sections. At 0 hr L3, I confirmed that loss-of-*torso* function did not appear to affect the presence of Sna in the PG, which was also true when I knocked down another component, *ras85D*, acting in the same pathway downstream of *torso* (Fig.2-8B). These results reinforced the idea that the dynamic expression pattern of *sna* is developmentally relevant. However, it remains unclear as to which developmental event(s) are linked to the Sna peak around 17 hr-18 hr in L2.

Although Sna protein levels in the PG were unaffected when *torso* was knocked down, I wanted to confirm that *sna* mRNA levels were not altered either. I tested this via qPCR using larval brain-RG complexes. I staged the larval populations using the blue gut method (183) instead of absolute hours to compensate for the difference of developmental timing between the RNAi and control animals. The two larval populations tested were the blue gut wandering larvae and clear gut wandering larvae. These two populations correspond to larvae that are about 18 hr and 8 hr before pupariation in controls. As seen in Fig. 2-9, at either time point, *sna* mRNA levels in *torso*-RNAi samples were not significantly downregulated compared to controls. Taken together, I showed that PTTH/MAPK signaling is not upstream of *sna* function in the PG.

2.3.3 PG>*sna*-RNAi animals showed endocycle arrest in the PG

So far, the underlying mechanism by which *sna* regulates ecdysone production was still not clear. I proceeded to tackle this question based on another phenotypic observation namely the fact that when I knocked down *sna* in the PG, the size of tissues was affected. The *Drosophila* larval PG is an endoreplicating tissue where PG cells undergo alternating S (DNA synthesis) and G (gap) phases without cell divisions, thus resulting in large nuclei with polytene chromosomes. An increase in DNA content often correlates with increased nucleus and cell size. Therefore, I hypothesized that the small size of the *sna*-RNAi PG was caused by endocycle arrest. I then assessed the PG DNA content in both PG-specific *sna*-RNAi animals and controls by DAPI staining. As shown in Fig. 2-10A, the nuclei size of *sna*-RNAi PG was comparable to that of controls at 0 hr after L2 to L3 molt. At 24 hr L3, the nuclei in control PG cells grew bigger because of endoreplication. In fact, one can distinguish the CA from the PG because of the larger size of

the PG cells at this time point. In contrast, PG nuclei size in *sna*-RNAi larvae remained unchanged and never caught up even at the end of the larval stage.

I also noticed that PG cell number in *sna*-RNAi animals was reduced (Fig. 2-10B and C). This is consistent with the finding that PG-specific RNAi of *Cdk2*, *CycE*, *cdt1*, *PCNA*, or *cul4*, all the major regulators of endoreplication, not only reduced DNA content, but also reduced cell numbers (101), suggesting that *sna* could be a novel regulator for endoreplication.

I wanted to make sure that the small PG nuclei phenotype was not caused by disrupted Sna function at the embryonic stages. To achieve both the temporal and spatial control of *sna*-RNAi expression, I took advantage of the Gal4-GeneSwitch system, where Gal4 will only be functional in the presence of a ligand RU486 (184). The Gal4-GeneSwitch I used was controlled by the enhancer element of *spookier* (*spok*) (Fig. 2-11). The ligand RU486 was administered to larvae in the food. This, along with the fact that *spok* is shown to be only expressed in the larval PG (72), ensures that *sna*-RNAi is not turned on in the embryonic PG when raised on RU486-containing food. As a result, when I examined the *spok-Gal4^{GS}>sna*-RNAi animals raised on food containing 8 µg/ml of RU486 at the end of the larvae stage right before pupariation, I still observed small PG nuclei compared to controls (*spok-Gal4^{GS}>sna*-RNAi animals raised on food without RU486) (Fig. 2-12). This result suggested that endoreplication in the PG requires Sna function during the larval stage.

Moreover, it is possible that the DNA content in the *sna*-RNAi nuclei is condensed, but the amount of DNA is the same as controls. To further verify that there is indeed decreased DNA content due to the lack of DNA synthesis phase (S phase) in the endocycle, I quantified the percentage of S-phase cells in both control and *sna*-RNAi PGs by EdU (5-ethynyl-2'-deoxyuridine) incorporation, an analog of thymidine. Basically, when cells are undergoing DNA synthesis, EdU will get incorporated and S-phase cells can be visualized using the Click-iT EdU Alexa Fluor 555 Imaging Kit (see material and methods at section 2.2.12). In brief, the Click-iT reaction refers to the process when the Alexa Fluor 555 (fluorophore) labelled Azide stably and specifically bind to EdU under the catalyzation of copper. For the EdU labelling, I first carried out the tests at several developmental stages in controls: a. 17 hr L2 (when there is a peak of Sna presence in the nuclei) as well as the time point right before (15 hr L2) and, after the peak (20 hr and 22 hr L2); b. 12 hr L3 (when *sna* expression is high) as well as 20 hr L3 (when *sna* expression

declined) (see Fig. 2-2, Fig. 2-21 and Fig. 2-27). In my experiments, EdU was incubated with RGS for 30 min. I found that within a 30 min time window, only a few cells in the PG are in S phase. For instance, the average number of S-phase cells in the PG was 11.3% ($\pm 5.4\%$) in controls at 15 hr L2. This demonstrated that the endocycle is unsynchronized in the PG, which is consistent with previous findings (101). More interestingly, I noticed that at time points when *sna* expression was low (i.e. 15 hr L2, 20 hr L2, 22 hr L2 and 20 hr L3), the percentage of S-phase cells in the PG also stayed low (never reaching beyond 13%) (also see Fig. 2-2, Fig. 2-13 and see Fig. 2-27). In contrast, around 17 hr L2 and 12 hr L3, there was a surge of cells that are in S-phase (Fig. 2-13), which correlates well with the high percentage of *sna*-expressing cells around the same developmental time points (see Fig. 2-2, Fig. 2-21 and Fig. 2-27). These results indicated that although the endocycle is not synchronized in the PG, the overall frequency is not stochastic, instead it is developmentally controlled.

On the other hand, when I knocked down *sna* in the PG, the percentage of the EdU positive cells was low at all time points (never beyond 12%). Furthermore, the peak of endoreplication did not occur (at 17 hr L2 and 12 hr L3) (Fig. 2-13). I concluded that *sna* is required for endocycle progression in the PG. Another conclusion is that there is a correlation between the percentage of Sna-positive (Sna⁺) cells and S-phase cells (Fig. 2-2B and C, Fig. 2-13, Fig. 2-22 as well as Fig. 2-27), suggesting that nuclear Sna is cell-cycle-controlled, and is only present in a certain period of the endocycle (see section 2.3.8 for details).

The percentage of S-phase (EdU⁺) cells within the 30 min time window would depend on the rate of endocycle progression (e.g. the length of the S-phase and G-phase of each endocycle). In controls at 17 hr L2, the percentage of S-phase cells in the PG ranged from 4% to 71% (Fig. 2-13B). This big discrepancy could be explained by the possibility that the one round of endoreplication at 17 hr L2 finishes within a very sharp time window, in other words, the length of S-phase is fairly short at that developmental time point. Moreover, developmental timing would not be synchronized within the animal population even though I staged the larvae at the L1 to L2 molt (178). Therefore, there was a smaller chance to capture the peak of S-phase cells around 17 hr L2. However, the *sna* expression period during the endocycle is probably longer than S-phase because almost all the cells are Sna⁺ around 17 hr L2 (Fig. 2-2 and Fig. 2-8A). Around 12 hr L3, the other round of endoreplication probably occurs more gradually and takes a longer time to complete, therefore, I observed lower variations for the percentage of S-phase cells among the

animal population (Fig. 2-13B). So far, it remains unclear what the exact biological significance of the sharp endocycle progression at 17 hr L2 is, but it appears that the biological event that couples with the endoreplication wave around 12 hr L3 is the critical weight checkpoint (see section 2.3.6).

2.3.4 PG-specific deletion of *sna* caused pupal lethality and affected the morphology of ring glands

As previously stated, when using RNAi, one needs to be cautious about off-target effects. A simple way to confirm that the RNAi phenotype is indeed caused by the disruption of the intended transcripts is to use another RNAi line targeting *sna* transcript at a separate region of the mRNA to determine whether the result is repeatable. However, so far I only found two VDRC *sna*-RNAi lines (VDRC #50003 and #50004) that gave us phenotypes. These two transgenic lines actually carry the same dsRNA construct (GD17171 in Fig. 2-14A). Moreover, all other existing dsRNA constructs (GD1536 from VDRC, HMS01252 and JF03094 from TRiP) did not give observable phenotypes even when turned on together with *UAS-Dicer2* in the PG. Hence, I decided to make a PG-specific *sna* deletion using CRISPR/Cas9 system as an independent approach to further verify the RNAi phenotypes. Briefly, the idea is to make transgenic flies ubiquitously expressing two ~20 nt guide RNAs (gRNAs) with homology to the *sna* locus and at the same time drive *UAS-Cas9* expression using *phm22-Gal4*. As a result, the two gRNAs will guide the endonuclease Cas9 to the *sna* locus and generate two double-stranded breaks close to each other in PG cells. The cells will then repair double-stranded breaks through error-prone NHEJ mechanism (176), thus generating small deletions in the *sna* gene (Fig. 2-14B).

Surprisingly, I found that when I expressed *UAS-Cas9* alone using *phm22-Gal4*, it already caused some developmental defects which were interestingly nutrient-dependent. In short, I observed a low percentage of L2 arrest and L2 prepupae (the delayed L2 animals eventually form pupae forgoing the L3 larval stage) when the animals were raised on standard corn-based medium at 25°C (the rest of larvae pupariated and eventually developed into normal adults). More strikingly, if I raised the animals entirely on yeast medium 25°C, 100% of the larvae were arrested as L2 and some of them would eventually become L2 prepupae (data not shown). I speculated that since *phm* is very highly expressed in the PG (26,120,185), there would be extremely high levels of Cas9 proteins when the *phm22-Gal4* was used to drive the *UAS-Cas9* expression; this possibly

had a detrimental effect on the cell independent of the gRNAs through a yet unknown mechanism. I hope my observation will raise caution for future research when others try to achieve CRISPR-based tissue-specific gene deletion using approaches similar to mine.

Therefore, I switched to a weaker driver *Mai60-Gal4*, which is expressed in the PG, but also in the brain, salivary gland and part of the eye disc (Fig. 2-15). I observed that *Mai60>UAS-Cas9* animals developed relatively normally with some degree of larval lethality. When *sna*-gRNAs were expressed together with *UAS-Cas9*, 85% ($\pm 3.6\%$) of pupal lethality was observed (Fig. 2-16A). However, it is unclear whether the lethality was owing to the loss-of-*sna* function in the PG or in other tissues since *Mai60-Gal4* is not exclusively expressed in the PG. When I used another tissue-specific driver that is not related to the PG (*eyeless-Gal4*), I did not observe any developmental defects or lethality (Fig. 2-16A). Moreover, I also showed that small deletions around the *sna*-gRNA target regions were detected by PCR followed by sequencing using genomic DNA extracted exclusively from RGs in the *Mai60>sna*-gRNAs; *UAS-Cas9* larvae. This result confirmed that Cas9 was guided to the *sna* locus (Fig. 2-16B).

Next, I dissected RGs from the PG-specific CRISPR mutants and examined the size of PG nuclei. In contrast to *sna*-RNAi PGs where small nuclei were observed, the conditional mutants had larger nuclei, meaning more DNA content per PG cell compared to *Mai60>UAS-Cas9* controls. Based on the DAPI quantification, the conditional mutants had about one round of endoreplication ahead of the control and this phenotype was already detectable early in the mid-L2 stage (Fig. 2-17A and B). It remains unclear as to why there were opposite phenotypes between the RNAi and the conditional mutants. Although small deletions were detected around the *sna*-gRNA target regions in PG cells of conditional mutants, how these deletions affect expression levels of *sna* is still unclear. In the future, one needs to determine whether CRISPR-based *sna* conditional mutants represent gain-of-function or loss-of-function effects. On the other hand, I did observe defects that were comparable to *sna*-RNAi RGs, because CRISPR-based *sna* conditional mutants showed dramatically reduced PG cell numbers when compared to *Mai60>UAS-Cas9* controls (Fig. 2-17C and D). Taken together, when I disrupted *sna* in the PG using two independent strategies, the cell number was affected and the endocycle progression was misregulated.

2.3.5 Endocycle in the PG was arrested at the critical weight checkpoint when *sna* was disrupted

Interestingly, various tissues during *Drosophila* larval stages undergo endoreplication, but *sna* is highly and quite specifically expressed in the PG. This suggested that *sna* must be involved in a unique aspect of endocycle regulation that is characteristic to the PG, which still remains unclear. I believe that the ecdysone deficiency in *sna*-RNAi animals is not merely a consequence of a general growth problem of the PG. The reason for this is that endocycle progression around 12 hr L3 (Fig. 2-13) is tightly coupled with a physiological event called critical weight attainment.

As previously discussed, ecdysone triggers major developmental transitions including onset of metamorphosis where animals transition from growth to initiation of sexual maturation. During the early half of the last larval stage in holometabolous insects, there is a developmental checkpoint called critical weight (CW) attainment. Once the CW has been passed, animals will commit to metamorphosis regardless of nutrient conditions, meaning starvation no longer stops the initiation of metamorphosis (Fig. 2-18A). The expression of the ecdysone biosynthetic genes increases gradually after attainment of CW in *Drosophila* (101) so that the major ecdysone peak triggering metamorphosis will form. CW attainment is like a "point of no return" regarding upregulation of ecdysone synthesis to initiate metamorphosis (101). At the molecular level, it is thought that the irreversible decision is at least linked to endocycle progression in the PG. Critical weight normally is attained around 9-12 hr L3 for *Drosophila melanogaster* kept at 25°C depending on genetic background and other external conditions (43). This is the time point when there is a wave of endoreplication in the PG as I described earlier (Fig. 2-13), as well as a high percentage of *Sna*⁺ cells in the PG (Fig. 2-2C Fig. 2-22 and Fig. 2-27), suggesting *sna* could be involved in the process of critical weight attainment. Since endocycle progression is blocked in *sna*-RNAi, I hypothesized that the animals did not get the internal signals as to whether they have attained critical weight or not. This may explain why *sna*-RNAi animals have ecdysone deficiency and never initiate metamorphosis.

In more detail, Ohhara *et al.* (101) have shown that during the last larval stage (L3), PG cells in *Oregon R* strains undergo two rounds of endoreplication. At the beginning of the 3rd instar PG cells have a chromatin value (C-values) of 16C and around the time of the critical weight attainment one round of endoreplication occurs, resulting in a C-value of 32C, after that the C-value of the PG will increase once more and stops at 64C. When the larvae are starved before the

CW checkpoint, endoreplication stops and the C-value remains at 16C, which correlates with a failure to turn on ecdysone production for metamorphosis (101). Therefore, I expected that the endocycle of PG cells in *sna*-RNAi larvae was arrested at a C-value of 16C. With this in mind, I genetically labeled the PG cells by *phm22>UAS-EGFP* expression and quantified DAPI intensity of the Z-stacked confocal images to determine the relative DNA content in each single PG cell at several specific time points.

For accuracy, I also tested the exact time of CW attainment in control (*UAS-Dicer2; phm22-Gal4>UAS-EGFP*) larvae. As a result, I found that none of the larvae could initiate puparium formation if they were starved before 6 hr L3. If the larvae were starved just one hour later, ~30% ($\pm 20\%$) larvae were able to form pupae indicating that these larvae had passed the CW by the time I starved them (Fig. 2-18B). After 12 hr into the L3 stage, the entire population had attained CW. Therefore, at least for larvae with the above-mentioned genetic background, the time when 50% of the larval population has attained critical weight was at 9-10 hr after the L2/L3 molt (Fig. 2-18B). The entire larval population did not pass CW within a sharp time window, which suggested that although I had staged the animals at the L2 to L3 molt, the development of the animals still tended to become asynchronized after the staging, which is a common problem when studying *Drosophila* development (178). Another possibility is that the process of CW attainment itself is somewhat stochastic. However, I am confident that before 6 hr L3 the animals are all pre-CW and after 12 hr L3 they are all post-CW according to my analysis.

In terms of endoreplication, I showed that in controls there are indeed two to three rounds of endoreplication during the L3 stage manifested in the relative DAPI intensity per PG increase from ~150 to ~900 (arbitrary units), which is a six-time difference. The log base 2 of 6 is 2.6, which means the DNA content replicated two to three times. Assuming the C-value under normal conditions always reaches 64C by the end of the larval stage in *Drosophila* PGs according to a previous study (101), the C-value in controls remained at 8 to 16C before the CW checkpoint (between 0 hr L3 to 12 hr L3), while the DNA content dramatically increased right after CW attainment (after 12 hr L3) (Fig. 2-19). These results confirmed that reaching 16C in the PG cells might be the molecular threshold determining whether the animals had gone through the CW checkpoint and committed to the production of the major ecdysone pulse that triggers metamorphosis. More importantly, PG cells in *sna*-RNAi animals have the same amount of DNA as that of the control PGs dissected before CW attainment and the DNA content never increased

by the end of L3 stage (Fig. 2-19). In other words, the C-value failed to reach beyond 16C similarly to the larvae that were starved before critical weight attainment (101). These data suggested that knocking down *sna* in the PG via RNAi blocked endocycle progression around the CW checkpoint, such that larvae did not commit to the production of the major ecdysone pulse triggering metamorphosis.

2.3.6 PG>*sna*-RNAi phenotype could not be rescued by promoting S-phase entry

Since knocking down *sna* in the PG caused endocycle arrest as shown by the findings that fewer S-phase cells were detected around CW checkpoint in the PG when *sna* was disrupted, I next asked whether I could rescue the size defect as well as the L3 arrest phenotypes by causing PG cells to reenter the S phase of the endocycle. It is well known that S-phase entry in endocycling cells is initiated by cyclin E/cyclin-dependent kinase 2 (CycE/Cdk2) (186). First, I tried to overexpress *CycE* in the *sna*-RNAi PG. However, continuous CycE/Cdk2 activity will also arrest endocycles in *Drosophila* (113,114), because CycE/Cdk2 needs to be silenced during G-phase to allow the reassembly of pre-Replication Complexes (preRCs) for the next S-phase. Therefore I used the *UAS-CycE.R* construct, which expresses full-length type 1 *cycE*, the level of which is known to oscillate between S and G phase when expressed in the PG (101) and will not affect endocycle progression and the onset of metamorphosis. When I expressed this *CycE*-cDNA construct in *sna*-RNAi animals, I found that it did not rescue L3 arrest (Fig. 2-20). In fact, the phenotype worsened where fewer larvae developed to L3 stages compared to *sna*-RNAi alone (here *UAS-EGFP* was also included to replace the *UAS-CycE*-cDNA to keep the same number of Gal4 binding sites).

Next, I also tried overexpression of *E2f1* in the PG, a key transcription factor known to periodically promote *CycE* expression, along with its obligate partner DP (115,187). However, overexpressing *E2f1/DP* alone in the PG caused 100% L1 arrest and when *E2f1/DP* and *sna*-RNAi were expressed together in the PG, the phenotype remained L1 arrest (data not shown). This could be explained by the finding that ectopic *E2f1* in the eye disc first promoted the cell to enter S-phase, but the cells subsequently died due to apoptosis (188). Since ectopic expression of *E2f1* itself will probably damage cells, it was not feasible to use the *E2f1/DP* construct for my rescue experiments.

Nevertheless, based on the result of *CycE*-cDNA rescue, forcing endocycle-arrested PG cells entering into S-phase could not rescue the larval arrest caused by *sna*-RNAi. This suggested that the normal function of *sna* in the PG is not simply to drive the progression of the endocycle. This again is in line with the idea that *sna* must be involved in a unique aspect of endocycle regulation that is characteristic for the PG, presumably being part of the molecular mechanism by which the endoreplication of PG cells and the assessment of critical weight attainment are coordinated. When *sna* is disrupted in the PG, there is possibly a miscoordination between the endocycle progression of the PG cells and the assessment of critical weight attainment. Animals perceived loss-of-*sna* as a red flag for failing the critical weight checkpoint; therefore, forcing cells to reenter the S-phase by *CycE* overexpression to override the "red flag" will generate more deteriorating damage to the larvae.

2.3.7 TOR functions upstream of *sna* in the PG

2.3.7.1 TOR is required for maximal expression of *sna* in the PG around the time of critical weight attainment

It has been suggested that endocycle activity in the PG is strongly correlated to CW attainment (101). Considering that before the CW checkpoint, growth is nutrient-dependent while it becomes nutrient-independent after CW is attained, it is reasonable to hypothesize that there must be a signaling pathway that bridges nutrient sensing, endoreplication and ecdysone biosynthesis in the PG. *Target of rapamycin* (TOR) is a key component of this signaling bridge, since it is known to function in nutrient sensing, controlling endoreplication and regulating ecdysone biosynthesis in the PG (44,101). More specifically, Ohhara *et al.* (101) showed that when TOR function is disrupted in the PG prior to CW, the C-value of PG stays at 16C while the endocycle is not affected if TOR function is lost after CW attainment, suggesting TOR is required for promoting the endocycle during the CW window. Interestingly, loss-of-*TOR* in the PG prior to CW attainment phenocopied PG-specific *sna*-RNAi where the endocycle was arrested at 16C (Fig. 2-19). Therefore, I hypothesized that *sna* may act in the same pathway as TOR and I asked whether TOR could act upstream of *sna* in the PG. To test this, I knocked down *TOR* in the PG via RNAi and tested the presence of Sna by immunofluorescence using anti-Sna antibodies. I found that Sna fluorescent levels were indeed reduced in *TOR*-RNAi PGs (Fig. 2-21A). Since Sna was only present in a few cells within the PG, I selected three Sna⁺ nuclei in each PG and then

quantified the average Sna intensity for both control and *TOR*-RNAi cells. This approach showed that the difference of immunofluorescent signals was significant (Fig. 2-21B).

Next, I further confirmed the results expressing *TOR* dominant negative (*UAS-TOR^{DN}*) at three different time points during the L3 stage (0 hr L3, 12 hr L3 and 24 hr L3) where 0 hr would represent the time before CW checkpoint, 12 hr is CW attainment and 24 hr after CW attainment. Since the amount of anti-Sna sera was limited, I used the *sna-gfp* line to monitor *sna* expression in the PG and this time I simply quantified the percentage of Sna-positive nuclei per PG. In controls, I again observed an increase of Sna-positive cells in 12 hr-old L3 PGs (Fig. 2-22). In contrast, the percentage of Sna+ cells stayed constantly low at all three time points in *PG>TOR^{DN}*. Taken together, these results suggest that *sna* works downstream of TOR in regulating nutrient-dependent endocycle progression around the time of CW checkpoint as well as ecdysone synthesis.

2.3.7.2 TOR functions through the TORC1 complex to regulate Sna levels in the PG

TOR forms two distinct protein complexes, TORC1 and TORC2 (61), where only the TORC1 complex was shown to couple nutrient-dependent endocycle progression to ecdysone biosynthesis (101). Therefore, I also tested the effect of loss-of-*raptor* on Sna levels, where *raptor* is a key component of the TORC1 complex (Fig. 1-4) (189). Consistent with the *TOR^{DN}* results, the *sna* expression pattern was affected in the same trend in *PG>raptor*-RNAi animals (Fig. 2-22).

Next, I tried to knock down *Akt* (*aka Akt1*) via RNAi, a downstream target of TORC2 complex (Fig. 1-4) (68,69,190) and then examined Sna protein levels in the PG. When I expressed *Akt*-RNAi using *phm22-Gal4*, I observed L2 arrest (data not shown). To keep the experiments at the same developmental stage as my previous studies (i.e. 0 hr L3), I switched to a weaker driver *P0206-Gal4* which is specifically expressed in whole RG instead of just the PG. *P0206>Akt*-RNAi had a weaker phenotype, which was L3 arrest (data not shown). Therefore, I could at least dissect RGs at 0 hr L3 and carried out anti-Sna antibody staining. I found that when *Akt* was knocked down in the RG, Sna distribution in the PG was not affected (Fig. 2-23A and B). Specifically, the P0206 driver also carries the *UAS-EGPF:mCD8* construct which marks the cell membrane. I noticed that the cell size of PGs in *P0206>Akt1*-RNAi became heterogeneous as evidenced by the GFP-labeled cell boundary. There appeared to be no correlation between the presence of Sna in

the nucleus and cell size (Fig. 2-23A). This inferred that the lack of Sna when TOR was disrupted in the PG was probably not caused by the small size of the PG cells that consequently makes less protein in general. In summary, my results suggested that TOR function in the TORC1 complex controls *sna* expression in the PG. Core components of the TORC2 complex (e.g., *ric*) may have to be tested in the future to definitely rule out the possibility that TORC2 function is related to *sna* expression in the PG (69,191).

2.3.7.3 IIS signaling is not dismissible for proper *sna* expression in the PG

The insulin/IGF signaling (IIS) and TOR pathways are interconnected at multiple steps in *Drosophila* (Fig. 1-4) (192,193) and are the core of the nutrient-sensing system that couples growth to nutritional conditions (48). In line with this concept, I expressed a dominant negative form of the insulin receptor (*InR^{DN}*) to test whether proper *sna* expression in the PG also requires IIS activities. Again, I had to use a weaker driver, *P0206* to obtain any L3 larvae that can be dissected for staining. I found that expressing *InR^{DN}* in the RG did not significantly affect *sna* expression (Fig. 2-23C). There is still the possibility that IIS signaling was not disrupted to the point that one could observe an effect on *sna* expression in *P0206>InR^{DN}* animals. More tests could be done in the future using *phm22>InR^{DN}* to examine *sna* expression at multiple time points around the 17 hr L2 stage to see whether the peak of *sna* expression is abolished. I also tested some players further downstream in the IIS signaling pathway, including PTEN, a negative regulator, as well as p110 (the catalytic subunit of PI3K encoded by *Pi3K92E* in *Drosophila*), a positive effector of the pathway. When I overexpressed *Pten* using the *P0206* driver (aimed at suppressing IIS signaling), Sna levels were significantly reduced based on the quantification of the Sna staining signal in the PG (Fig.2-23A and C). When I tried to disrupt the IIS signaling by *Pi3K92E* knock-down (*Pi3K-RNAi* in Fig. 2-23C), I found an inconsistent result where some of the PG had abolished *sna* expression, whereas others appeared to have relatively normal Sna levels (Fig. 2-23C). Taken together, my results showed that disrupting some of the IIS components including PI3K and PTEN, affected *sna* expression but the effect was not as predominant as the loss of TOR. I speculate that proper *sna* expression in the PG directly requires TOR function but is not directly dependent on IIS signaling; however, since there is complex crosstalk between the two pathways, Sna levels could be affected to a certain degree if IIS is disrupted in the PG.

A recent study from our lab showed that the circadian clock in the PG is required for ecdysteriodogenesis and that IIS signaling interacts with the circadian clock in PG (40). When

timeless (tim), a core component of the circadian machinery in insects, was knocked down in the PG, the size of the PG was also affected (40). If *Sna* levels in the PG were indeed dependent on IIS/TOR signaling, then I would expect that knocking down *tim* would also affect *sna* expression to some degree. Based on this idea, I entrained flies in a 12 hr light-dark cycle for three days and then set up immunofluorescent experiments with *phm22>tim-RNAi* and control animals. I found that ~40% of the PG had reduced *Sna* staining signals, which resembled the *Pi3K* knock-down (Fig. 2-24). The control and RNAi RGs were dissected at the same developmental time (0 hr L3), but not at the same Zeitgeber, because of the developmental delay caused by *tim-RNAi*. Therefore, I cannot rule out the possibility that the difference in *Sna* intensity between controls and *tim-RNAi* was due to differential expression of *sna* at different circadian time points since there is so far no evidence showing that *Sna* levels are not circadian regulated. In the future, it may be worthwhile to examine *Sna* levels at the same Zeitgeber between control and *tim-RNAi* and test at multiple Zeitgebers. Taken together, there are several lines of evidence showing that *Sna* levels in the PG are indeed regulated by IIS/TOR signaling.

Finally, in Fig. 2-21, 2-22 and 2-23, I also noticed that *phm22>TOR^{DN}*, *phm22>TOR-RNAi*, *phm22>raptor-RNAi*, *P0206>Pten-cDNA* and *phm22>Pi3K-RNAi* animals all have smaller PG nuclei compared to their control counterparts RNAi, which confirmed the previous findings that IIS/TOR is required for nutrient-dependent endocycle progression in the PG (101).

2.3.7.4 IIS/TOR signaling regulates *sna* expression at the post-transcriptional level

I showed that *Sna* protein levels in the PG are dependent on IIS/TOR pathway and I next asked whether the expression of *sna* mRNA also responded to IIS/TOR signaling. Therefore, I measured mRNA levels via qPCR in *TOR^{DN}*, *TOR-RNAi* and *Pi3K-RNAi* RGs, respectively. I found that, except for PG>*TOR-RNAi* larvae where there is a slight reduction in *sna* expression (but not significant), knocking down IIS/TOR components in the PG did not seem to affect *sna* at the mRNA levels (Fig. 2-25), suggesting that IIS/TOR probably regulates *sna* at a post-transcriptional level.

2.3.7.5 *Sna* protein levels in the PG are nutrient-dependent around the CW checkpoint

TOR is known to be part of a nutrient-sensing machinery in various organisms (194). In *Drosophila melanogaster*, it has been shown that amino acid deprivation phenocopies the *TOR* mutant in multiple aspects (195,196). If *Sna* levels in the PG were indeed dependent on TOR

function, I expected that the expression of *sna* might also be regulated by nutrient conditions. Moreover, my data suggested that *sna* function is essential for endocycle progression especially around the CW checkpoint. A previous study also demonstrated that endoreplication around the CW time point requires TOR activity, while the endocycle becomes TOR-independent after CW attainment (101). Therefore, I wondered whether maximal expression of *sna* around the CW checkpoint would also be sensitive to nutrient status, but not responsive to nutrient conditions after CW attainment. I tested this by food deprivation either starting at 4 hr L3 (before the CW checkpoint, = early starvation) or at 13 hr L3 (after the CW checkpoint, = late starvation) for 9 hr and then examined *sna* expression by immunofluorescent staining. I first used the *sna-gfp* line for a preliminary test. Consistent with previous results, Sna-GFP is present in a few nuclei in the PG at 4 hr L3 (no starvation), while after nine hours of starvation (early starvation) the number of Sna⁺ positive cells were reduced significantly. Even in the one or two Sna-GFP⁺ cells, the signals were very blurry and weak. The larvae that were not starved would developmentally reach 13 hr L3. In striking contrast, at this time point Sna-GFP is still easily detectable in the PG. This result confirmed that Sna levels are indeed relying on nutrient conditions at least around the CW checkpoint. I also found that when larvae were starved at 13 hr L3 for 9 hr (late starvation), the expression of *sna* was also significantly affected comparing to feeding animals at 22 hr L3 (Fig. 2-26).

Next, I carried out similar experiments using anti-Sna antibody in *w¹¹¹⁸* larvae and this time I reduced the starvation time to 6 hr to see whether Sna levels would be sensitive to nutrient deprivation within a shorter period. In the end, I acquired similar results. In particular, 6 hr of starvation started before the CW checkpoint (at 4 hr L3) was sufficient to abolish *sna* expression, while the developmentally matched feeding counterpart (reached 10 hr L3 developmentally) had a high percentage of Sna⁺ cells (Fig. 2-27). The Sna staining pattern at 10 hr and 13 hr L3 were again a confirmation that there is a peak of Sna⁺ cells around the time of CW attainment under feeding conditions (Fig. 2-27). In contrast, the effect of starvation on Sna levels after CW attainment was not as prominent as that prior to CW attainment since the percentage of Sna⁺ positive cells in animals starved at 13 hr L3 for 6 hr was comparable to that in the 19 hr-old feeding L3 larvae (Fig. 2-27B). All these results suggest that *sna* expression is somehow dependent on nutrient conditions around the CW checkpoint. In the future, it would be worthwhile to confirm these results by Western Blot using RG samples. I am aware that nutrient availability affects

general protein translation through TOR signalling (Fig. 1-4) (197). Therefore, I am not sure how specific this starvation effect on Sna protein levels is. It is possible that a range of proteins have this dependency on nutrients around the time of critical weight checkpoint in the PG. However, the Sna expression pattern itself makes it a good candidate as a player in critical weight attainment, since it gradually increases until the point of critical weight attainment and then starts to decline regardless of nutrient availability. This is at least an example for how the levels of a specific protein could be nutrient-dependent before the CW, but nutrient-independent after CW.

Interestingly, I noticed that there was also Sna immunofluorescent signals in parts of the brain where insulin-producing median neurosecretory cells (IPCs) seem to be located, and Sna appeared to be present in the cytoplasm of these IPCs (Fig. 2-28; also see 13 hr L3 feeding sample in Fig. 2-26). IPCs are the main site of insulin-like peptides (ILPs) expression and secretion. In *Drosophila*, ILPs activate the single InR and subsequently the IIS signaling pathway, which plays a key role in coordinating tissue growth in response to changing nutrient conditions (198). In this context, secretion of at least ILP2 and ILP5 are also nutrient-dependent (53,54). Therefore, it will be interesting to investigate whether *sna* also has a function in IPCs regulating ILPs secretion in response to nutrient status. In preliminary work, I did confirm that the Sna⁺ cells in the brain are IPCs (Fig. 2-28). The IPCs were labeled with *EGFP* driven by *ilp3-Gal4* and I found that Sna signals largely overlapped with the EGFP signal; however, it was not a perfect match. It seems that the *ilp3-Gal4* I used did not cover as many IPCs as the one showed by J. Cao *et al.* (199) using a *ilp2-Gal4*. Therefore, one can repeat these experiments using the *ilp2-Gal4* to drive a *UAS-CD8:mGFP* expression, which might provide a clearer view of the IPCs. Nevertheless, I am confident that Sna is also present in the cytoplasm of the IPCs.

2.3.8 Sna protein is probably present during G phase of the endocycle

So far, there have been multiple lines of evidence suggesting that Sna may have a function in controlling endoreplication around the CW checkpoint depending on nutrient status. I also hypothesized that Sna is only present within a small time window during each round of the endocycle (see section 2.3.3). As my previous data suggested that Sna function is probably not simply promoting S-phase entry, revealing in which phase *sna* is expressed during the endocycle may provide insight into how Sna functions are related to endoreplication. Generally, endoreplication has only two key stages: S and G phase. There is no defined G-phase marker, while S phase is defined by DNA synthesis, which can be clearly marked using EdU incorporation.

Therefore, I performed EdU labeling along with Sna antibody staining at 8-12 hr L3 (around CW attainment) as well as 16 hr L2, the time points known to have both a high percentage of Sna⁺ cells and EdU positive cells in the PG, to test whether the Sna⁺ positive cells are also the S-phase cells. For this, I used the transgenic line carrying the C-terminal GFP-tagged *sna* genomic clone described in section 2.3.1.

In general, the *sna*-expressing cells did not overlap with the EdU⁺ cells. On occasion, there are a few cells that were both Sna⁺ and EdU⁺ (Fig. 2-29). When Sna is C-terminal tagged with GFP, it may have slight different stability compared to endogenous Sna, which may explain the occasional overlap of Sna-GFP with the EdU staining. In the future one can co-stain the Sna⁺ cell and EdU⁺ cell using the endogenous Sna antibody to see whether similar results would be obtained. Nevertheless, my results suggested that Sna is probably expressed in Gap phase, either right after or right before S-phase (since there was occasionally overlap between Sna⁺ and S-phase cells). In the future, which target genes are controlled by Sna during Gap phase in endoreplicating PG cells need to be explored by ChIP-Seq analysis using RG samples.

2.3.9 Sna protein stability may be regulated by α -ecdysone

For EdU labeling, larvae have to be dissected and tissues need to be incubated *ex vivo* in saline containing 10 μ M EdU for 30 min for the S-phase cells to incorporate the nucleotide analog. I realized that since Sna levels are nutrient-dependent, Sna became already undetectable by antibody staining after the 30 min of *ex vivo* culture in 1x PBS or 1x Ringer's solution, which made double labelling not feasible. Therefore, I switched to Schneider's insect medium with 10% fetal bovine serum (FBS) during EdU incorporation to see if Sna could be better preserved within a 30 min incubation. It seems that Schneider's medium with 10% FBS was still not sufficient to maintain *in vivo* expression pattern of *sna*. Considering that I made sure RGs were still attached to the brain *ex vivo*, I hypothesized that Sna level in the PG might be responsive to a non-brain derived signal which is missing *ex vivo*. Since Schneider insect medium contains necessary amino acids; therefore, amino acids and serum were not the critical factors or at least they were not sufficient to maintain the Sna protein stability.

I then tried to add insulin or α -ecdysone, two important hormones for insects, respectively to the Schneider's insect medium (with 10% FBS). Interestingly, I found that adding 2 μ g/ml or 4 μ g/ml α -ecdysone could preserve *sna* expression pattern comparable to what I observed *in vivo*

(Fig. 2-30A). More surprisingly, adding 10 $\mu\text{g/ml}$ of insulin did not make a difference compared to Schneider's insect medium (with 10% FBS) alone (Fig. 2-30A). These data suggested that the stability of the Sna protein is in part maintained by α -ecdysone in the PG. Sna-GFP fusion on the other hand has higher stability compared to WT Sna. The protein can be preserved in 1x Ringer's solution (an isotonic solution similar to the body fluids of *Drosophila*) or Schneider's insect medium (with 10% FBS) alone, but not in 1x PBS (Fig. 2-30B). α -ecdysone is the final product of ecdysone synthesis in the PG before it is released into the hemolymph, followed by conversion to the biologically active form, 20E, outside the PG. Given that Sna has a role in regulating ecdysone production in the PG and Sna protein levels in the PG are possibly regulated by ecdysone itself, Sna may be involved in a feedback loop that controls ecdysone levels. Again, it is unclear how wide-spread this effect of α -ecdysone on protein stability in the PG since I only tested one protein (i.e. Sna) in this matter. It is possible that α -ecdysone itself is a general stabilization signal for proteins in the PG, which needs to be further investigated.

2.3.10 Overexpressing *sna* family proteins in the PG caused endocycle arrest in a cell-autonomous manner

2.3.10.1 The effect of *escargot* (*esg*) overexpression on endocycle progression in the PG

My results suggest that *sna* is probably expressed in G-phase of each endocycle and has a crucial role in regulating endocycle progression in the PG. However, the exact biological event that might be controlled by Sna during the G-phase remains unknown. One previous study showed that MDCK (Madin-Darby canine kidney) cells transfected with human *snail* were blocked at G0 or G1 of the cell cycle and in developing mouse embryos; a negative correlation was also observed between cells expressing Sna and cells undergoing S-phase or mitosis, suggesting Sna suppresses cell-cycle progression (160). Moreover, in *Drosophila*, ectopic expression of *escargot* (*esg*, another member of the *snail* family genes) in salivary gland cells inhibits endoreplication (172), while in *esg* mutants, the abdominal histoblasts, normally non-endoreplicative cells, became polyploid and failed to develop to adult abdominal cells, suggesting that the function of *esg* is to repress endoreplication in those cells (171). Thus there appears to be a link between Sna function and the inhibition of cell-cycle or endocycle progression, which is opposite to my observation in the PG that loss-of-*sna* function resulted in endocycle arrest. Interestingly, expression of *esg* is also >20-fold enriched in *Drosophila* larval RGs compared to the whole body

(120). Thus, I hypothesized that the two Snail family proteins, Sna and Esg have opposing roles in the larval PG, with *esg* limiting and *sna* maintaining endocycle progression to balance the rate of endocycle progression with development. With this in mind, I tested whether overexpressing *esg* in the PG indeed inhibits endocycle progression by inducing the expression of *esg*-cDNA at two different time points during larval stages.

For *esg* overexpression, I raised both *spok*^{GS} (on 2nd chromosome) >*w*¹¹¹⁸ (control) and *spok*^{GS}>*esg*-cDNA animals on plain NutriFly food until the beginning of L3 stage. Larvae were synchronized at the L2 and L3 molt and subsequently switched to 8 µg/ml RU486-containing Nutrifly food for 24 hr to induce *esg*-cDNA expression (Fig. 2-31A). I found that DNA content in *spok*^{GS}>*esg*-cDNA PG cells was reduced compared to that of control animals (*spok*^{GS}>*w*¹¹¹⁸) after 24 hr of RU486 feeding, suggesting that endocycle progression was affected during the first half of L3 stage (Fig. 2-31B). *spok*^{GS}>*esg*-cDNA animals raised on RU486-containing food also displayed larval arrest phenotypes, probably caused by endocycle arrest in the PG (Fig. 2-31A).

Next, I asked whether inhibition of endoreplication caused by *esg* overexpression is dependent on developmental stage. With this in mind, I tried to induce the *esg*-cDNA during L2 stage, i.e. larvae were transferred to RU486 food at 0 hr L2 for 23 hr and RGs were dissected and examined subsequently. This time I also examined DNA content at 0 hr L2 before RU486 feeding for both controls and *spok*^{GS}>*esg*-cDNA animals (Fig. 2-31A). Interestingly, DNA content did not increase significantly throughout the entire L2 stage even in *spok*^{GS}>*w*¹¹¹⁸ (control) PG cells when animals were fed RU486. Hence, it remains inconclusive as to whether *esg* overexpression also blocked endoreplication during L2 larval stage. The results found in *spok*^{GS}>*w*¹¹¹⁸ (control) PG cells contradicted the finding that PG cells undergo two rounds of endocycles during L2 stage under normal conditions (101). There could be three possible explanations for this: 1. DAPI staining and intensity quantification might not be sensitive enough to detect changes of DNA content when nuclei are relatively small (for L2 PG cells). 2. The genetic background of the *spok*^{GS} driver is different from *Oregon R* flies used in the other study on endoreplication in the PG during L2 stage (101) and genetic background may affect endocycle progression of the animals. 3. RU486 itself or raising larvae on Nutrifly food may have slowed down endocycle progression during the L2 stage. These could be easily tested in the future using a different food source with or without RU486 in a reference fly strain.

2.3.10.2 The effect of *sna* overexpression on endocycle progression in the PG

Next, I asked what would be the effect of *sna* overexpression on endocycle progression in the PG. I expected that it might result in the opposite effect to loss-of-*sna* function in the PG, i.e. acceleration of endocycle progression in the PG. Based on this idea, I induced *sna* expression in a similar manner as in the *esg* experiment. As mentioned previously, there was some concern that the *spok*^{GS} driver might already have reduced the rate of endoreplication during the L2 stage. Therefore, I switched to another line where *spok*^{GS}-*Gal4* construct was inserted on the 3rd chromosome to see whether I would obtain different results for the control (*spok*^{GS} (2) > *w*¹¹¹⁸ in Fig. 2-32). This time, when larvae were fed RU486-containing food throughout the entire L2 stage, I detected significant changes in DNA content in controls, which represents one round of endoreplication (Fig. 2-32B and C). This still differs slightly from previous findings that during the 2nd instar, there are two rounds of endoreplication (101). This result suggested that the two rounds of endoreplication during the L2 stage are food source-dependent or genetic background-dependent, i.e. not applicable to general conditions. Surprisingly, when *spok*^{GS} (2) > *sna*-cDNA animals were fed with RU486-containing food throughout the entire L2 stage, the one round of endoreplication did not happen (Fig. 2-32C), suggesting that overexpression of *sna* in the PG also blocked endoreplication. After 24 hr of RU486 feeding from the beginning of the L3 stage, DNA content per PG cell was also lower in *spok*^{GS} (2) > *sna*-cDNA animals compared to *spok*^{GS} (2) > *w*¹¹¹⁸ controls, again suggesting endocycle arrest upon *sna* overexpression. *spok*^{GS} (2) > *sna*-cDNA animals raised on RU486-containing food also displayed larval arrest phenotypes, probably caused by endocycle arrest in the PG (Fig. 2-32A).

Interestingly, after only 24 hr of RU486 administration, the number of PG cells in *spok*^{GS} (2) > *sna*-cDNA animals increased compared to the number of PG cells before induction of *sna* expression (at 0 h L3 on plain Nutrifly food). The increase of PG cell number was not observed when *spok*^{GS} (2) > *w*¹¹¹⁸ were fed on RU486 food for 24 hr, which suggested that this result was specifically caused by *sna* overexpression. I speculated that when *sna* was overexpressed in the PG, cells might undergo mitosis instead of endoreplication. This idea needs to be further tested by staining the PG with a mitotic cell cycle marker. More interestingly, overexpressing *CycE* in the PG also caused increased cell numbers (Fig. 2-33), just as in *sna* overexpression (Fig. 2-32B and D). This correlation again tied *Sna* function to the regulation of cell cycle or more specifically endoreplication in the PG.

Until now, I have not shown any evidence that the *spok^{GS}-Gal4* driver I used did not have leaky expression in the absence of RU486. Therefore, I included *spok^{GS} (2)>UAS-EGFP; UAS-sna-cDNA* animals to visualize the turn on of *sna-cDNA* expression using EGFP. Again, animals were reared on Nutrifly food (without RU486) until 0 hr L3 (Fig. 2-34A) and then either transferred again to fresh Nutrifly food or RU486-containing Nutrifly food for 24 hr. As expected, larvae fed on plain Nutrifly food did not express EGFP, while EGFP was turned on only in the PG after 24 hrs of RU486 administration, showing that *spok^{GS}-Gal4* expression was induced as expected. In the end, I obtained a similar result, namely that the *sna*-overexpressing PG cells (GFP positive) had small nuclei compared to its counterpart (24 hr L3 PG without RU486, GFP negative) (Fig. 2-34B and C). However, the inhibition of endoreplication was weaker when the *sna-cDNA* was expressed along with *EGFP*. I did not observe the increase of PG cell number in this case either, which is probably because an extra UAS-binding site in *UAS-EGFP* construct titrated the Gal4 molecules, thus weakening the expression of *UAS-sna-cDNA*.

I was also concerned that the small nuclei I observed when *sna* was overexpressed in the PG might be an indirect effect. Since *sna* overexpression in the PG resulted in the developmental delay and developmental arrest (Fig. 2-32), it was possible that the developmental delay or arrest slowed down endocycle progression within the PG. To show that *sna* overexpression inhibits endoreplication in a cell-autonomous manner, I took advantage of the flip-out-Gal4 system to generate both the wild type and *sna* overexpression cells within the same RG. Basically, *hs-FLP; tubulin-FRT-CD2-FRT-Gal4, UAS-GFP* flies were used where a stop codon flanked by two FRT sites normally prevents Gal4 expression. After a mild heat shock, the stop codon was excised by Flippase in a random manner within the same tissue which, in turn, switches on *sna* expression in a mosaic pattern (the *sna*-overexpressing cells would be marked by the presence of GFP) (Fig. 2-35A). The animals were allowed to develop further following the heat-shock and I dissected the RGs at the end of the larval stage. I found that *sna*-overexpressing cells (GFP+ positive) had small nuclei and DNA content was significantly reduced compared to the endogenous control (non-GFP cells). Since heat shock was applied early on during development (at L1 stage), *sna* was overexpressed throughout almost the entire larval development in those cells; however, based on DAPI quantification, *sna*-overexpressing cells were 2-3 endocycles away from the C-value that a control cell could reach at the end of larval development (Fig. 2-35C), suggesting that endoreplication was blocked before the CW checkpoint (C-value remains ~ 8-16C), a result

similar to *sna* knock-down in the PG. I also included the controls (*hs-FLP; tubulin-FRT-CD2-FRT-GAL4, UAS-GFP>y^[1]w^[67c23]*) where even after Flp-out, cells will only express GFP but not *sna*-cDNA. As a result, GFP+ cells had slightly smaller nuclei compared to non-GFP cells, suggesting that overexpressing GFP had a slight effect on endoreplication as well, however, this was not nearly as substantial as the effect of *sna* overexpression. These results confirmed that overexpressing *sna* specifically inhibits endoreplication in the PG.

In sum, my data suggested that Sna levels in the PG are tightly regulated. Too much or too little Sna protein affects cell cycle (endoreplication). Interestingly, continuous expression of *CycE*, the key driving force of S-phase entry, by the UAS/Gal4 system in the salivary gland (an endoreplication tissue) also inhibited endocycle progression because *CycE* levels normally oscillate during each round of endoreplication, i.e. *CycE* activity needs to be downregulated after S-phase entry to allow the next cycle of S-phase. This is another hint suggesting that *sna* function is linked to endoreplication. It will be worthwhile to show that Sna levels indeed oscillate within a single PG cell in relation to the endocycle using live imaging. In this way, one will be able to understand when exactly during each round of the endocycle is *sna* expressed.

2.3.11 Identifying Snail target genes in the PG via Next Generation RNA-Seq

Considering that Sna is a transcription factor, identification of Sna target genes in the PG would be another approach to reveal roles for Sna. As a first step to identify Sna targets, I carried out RG-specific RNA-Seq analysis with RG samples collected from PG-specific *sna*-RNAi animals, controls, as well as animals with *hs>sna*-cDNA overexpression. The RNA-Seq was performed using the Illumina HiSeq 2500 platform, in which bar-coding cDNA libraries from different genotypes allowed me to sequence RNA from all my samples (12 samples) in a single lane. Gene Ontology (GO) term enrichment analysis of the most differentially expressed transcripts in *sna*-RNAi or *sna*-cDNA PG samples may allow me to identify PG-specific processes that are dependent on Sna function. I found that using *phm22-Gal4* driver to induce *sna*-cDNA expression in the PG was too strong and resulted in 100% L1 lethal, which made it impossible to collect RGs for RNA-Seq experiments. Therefore, the alternative method was to induce *sna* expression using a heat shock driver. A single heat shock (at 37°C for 50 min) was carried out when larvae reached the mid L3 stage and larvae were recovered at 25°C for 6 hr before dissection. Given the short period of Sna overexpression, I expected that secondary effects caused by

overexpression would be limited, and differentially expressed genes identified by RNA-Seq would be primarily directly related to *Sna* function.

For RNA-Seq analysis, raw sequencing reads were mapped to the *Drosophila* transcriptome using Arraystar 4.0 and Reads Per Kilobase of transcript per Million mapped reads (RPKM) were also calculated by Arraystar 4.0, which was used to represent the relative abundance of each transcript. Ultimately, relative fold change of expression was determined by the ratio of RPKMs between the experimental genotypes (PG>*sna*-RNAi or *hs*>*sna*-cDNA) and the control counterpart. An RPKM value greater than one was set as a cut off for a reliable read count for each transcript.

2.3.11.1 Genes that are differentially expressed in PG>*sna*-RNAi ring glands

Initially, I looked at which are the most differentially expressed genes when *sna* is disrupted in the PG. Therefore, I listed the top 50 downregulated genes as well as the top 50 upregulated genes in the *sna*-RNAi samples in Appendix (Table A-1).

Downregulated genes and Gene Ontology (GO) term enrichment analysis

When I knocked down *sna* in the PG, a total of 121 genes were identified as more than 3-fold downregulated. When I lowered the cut-off to a 2-fold change, I found 201 genes that were downregulated. I next carried out Gene Ontology (GO) term enrichment analysis via DAVID, a multifunctional bioinformatic tool that allows one to categorize gene functions by GO_biological process, GO_cellular component and GO_molecular function, KEGG pathway and InterPro proteins families (179,180). DAVID will present enriched GO terms with their *p* value as well as the false discovery rate (FDR) based on the Benjamini–Hochberg method (200). One has to decide the cut-off as statistically significant and the cut-off I used was FDR<0.05 (usually FDR is more stringent than *p* value).

For the Gene Ontology-enriched analysis, I used >2-fold downregulated gene set. My rationale was that if several genes in the same pathway were all differentially expressed in *sna*-RNAi, even if the expression of those genes in the same pathway was only moderately affected, it would still suggest that the process or pathway, in which these differentially expressed genes were involved, was dependent on *Sna* function. The only overrepresented term I found was “Insect hormone biosynthesis” (KEGG_Pathway, P=1.3E-5, FDR=3.5E-4); within this group are *nvd*, *spok*, *phm*, *dib* and *sad*. This finding is consistent with the idea that PG>*sna*-RNAi animals displayed ecdysone deficiency, thus

resulting in developmental arrest (Fig. 2-5). I then manually sorted all the >2-fold downregulated genes previously known to be involved in regulating ecdysone production according to the literature and those ecdysone-related downregulated genes are summarized in Table. 2-3 along with their fold changes as well as the *p* value (based on two-way ANOVA). Since my RNA-Seq was carried out in duplicate, the *p* value is of little relevance. Therefore, all differentially expressed genes in *sna*-RNAi need to be further validated by qPCR. However, the expression of *spok*, *nvd*, *phm*, *sad* and *dib* were already shown to be severely downregulated via qPCR (Fig. 2-5), thus further confirmed the validity of my RNA-Seq procedure.

Table 2-3. >2-fold downregulated genes found in PG>*sna*-RNAi that are related to ecdysone production

Gene name	Fold change	<i>p</i> value	Function annotation	Validated by qPCR	References*
<i>nvd</i>	-9.4	0.17	Ecdysteroid biosynthetic gene	Yes	Yoshiyama <i>et al.</i> , 2006
<i>spok</i>	-8.9	0.11	Ecdysteroid biosynthetic gene	Yes	Ono <i>et al.</i> , 2006
<i>Start1</i>	-8.8	0.03	Cholesterol binding	No	Roth <i>et al.</i> , 2004
<i>CG11762</i>	-6.9	0.09	Zinc finger transcription factor	No	Komura-Kawa <i>et al.</i> , 2015
<i>sad</i>	-5.9	0.17	Ecdysteroid biosynthetic gene	Yes	Petryk <i>et al.</i> , 2003
<i>phm</i>	-5.4	0.10	Ecdysteroid biosynthetic gene	Yes	Warren <i>et al.</i> , 2004
<i>tor</i>	-5.3	0.14	Receptor tyrosine kinases, PTH receptor	Yes	Rewitz <i>et al.</i> , 2009
<i>GstE14</i>	-4.7	0.08	Cholesterol homeostasis	No	Enya <i>et al.</i> , 2014
<i>Npcl1a</i>	-4.4	0.20	Regulation of cholesterol transport	No	Huang <i>et al.</i> , 2005
<i>dib</i>	-4.0	0.15	Ecdysteroid biosynthetic gene	Yes	Warren <i>et al.</i> , 2002
<i>vvl</i>	-2.7	0.15	POU transcription factor	No	Danielsen <i>et al.</i> , 2014
<i>Atet</i>	-2.5	0.29	ABC transporter-like, ecdysone transport	No	Yamanaka <i>et al.</i> , 2015
<i>Fdx2</i>	-2.4	0.01	Ferredoxin 2, positive regulation of ecdysteroid biosynthetic process	No	Palandri <i>et al.</i> , 2015

<i>mld</i>	-2.2	0.30	Zinc finger transcription factor	No	Neubueser <i>et al.</i> , 2005
<i>sro</i>	-2.0	0.32	Ecdysteroid biosynthetic gene	Yes	Niwa <i>et al.</i> , 2010

*References indicate the original publication that demonstrated the role of each gene in ecdysone production

Creating functional association networks for downregulated genes

Another way to test whether deregulated genes were functionally associated was to see whether the gene products, i.e. proteins, interact with each other. The database I used for the functional association networks analysis was STRING, which integrates protein–protein interactions based on physical as well as indirect (functional) interactions (201).

To create functional association networks using STRING database, I took >3-fold downregulated genes in PG>*sna*-RNAi RGs because STRING works better with smaller data sets. As a result, STRING generated a visual overview of the protein-protein interaction network (Fig. 2-36). First of all, the >3-fold downregulated gene set in PG>*sna*-RNAi is predicted by STRING to yield more interactions on the protein level than what would be expected for a random set of proteins of similar size, drawn from the genome. This conclusion is based on the protein–protein interaction enrichment value (PPI) calculated by STRING to be lower than 0.05 (in this case PPI=0, Fig. 2-36). Moreover, several function association subgroups were present within the big interaction network and it turned out that those function association subgroups exactly reflect the results from the term enrichment analysis performed by DAVID (Fig. 2-36, Table 2-3) and (Fig. 2-37 and Table 2-5). The annotation of the function interaction subunits was done by checking the GO term for each node (represent each protein) within the interaction subunit and then annotating the function association subgroup according to the GO term of the nodes. In this case, ecdysone-related downregulated genes are shown to be functionally associated into networks by STRING analysis (Fig. 2-36). Moreover, STRING automatically provides a term enrichment analysis based on GO term, KEGG pathway, InterPro domains and PFAM Protein Domains, which confirmed the findings from the DAVID GO analysis. The statistics used in STRING for term enrichment analysis was also false discovery rate (FDR) (23) and the cut-off was also <0.05. There are several overrepresented terms calculated by STRING that were not found by DAVID when I used the >3-fold downregulated gene set. The results of GO analysis by STRING are shown in Table. 2-4 and all those enriched GO terms are related to ecdysone production. In

conclusion, DAVID and STRING database work equally well with large-scale transcriptome data, while STRING provides more user-friendly and visual output of analysis results.

Table 2-4. Overrepresented GO terms and KEGG pathways for downregulated genes identified in PG>*sna*-RNAi

#Pathway ID	Pathway description	FDR	Gene names
GO biological process			
GO.0045456	Ecdysteroid biosynthetic process	3.0E-07	<i>GstE14, npc1a, dib, nvd, phm, sad, spok</i>
GO.0016125	Sterol metabolic process	2.5E-04	<i>npc1a, dib, phm, sad, spok</i>
GO.0006697	Ecdysone biosynthetic process	8.2E-04	<i>dib, phm, sad, spok</i>
GO.0015918	Sterol transport	4.4E-02	<i>npc1a, npc2e, start1</i>
GO molecular function			
GO.0008395	Steroid hydroxylase activity	1.5E-02	<i>dib, phm, sad</i>
KEGG pathway			
981	Insect hormone biosynthesis	9.9E-06	<i>dib, nvd, phm, sad, spok</i>

>3-fold downregulation was used as the cut-off.
 GO enrichment analysis was carried out using STRING.
 FDR: False Discovery Rate.

Upregulated genes identified in PG>*sna*-RNAi ring glands

I identified 200 genes that were >3-fold upregulated in *sna*-RNAi PGs. When I set the cut-off to >2-fold up, the number of affected genes went up to 370. Similarly, I took the >2-fold upregulated genes and carried out GO term analysis using DAVID and the enriched terms, as well as the genes associated with each term are shown in Table 2-5. Essentially, three main enriched terms were identified, namely, "Alkaline proteases activity" (GO_molecular function, P=7.6E-5, FDR=0.021), "Glutathione metabolism" (KEGG_PATHWAY, P=5.4E-5, FDR=1.6E-3) and "CHK kinase-like domain" (Interpro domain, P=4.6E-05, FDR=0.01) (Table 2-5). Genes under the term "Folate biosynthesis" are essentially the same as the ones under "Alkaline protease activity", while "Metabolism of xenobiotics by cytochrome P450" as well as "Drug metabolism-cytochrome P450" are redundant with "Glutathione metabolism". Moreover, genes under "Protein of unknown function DUF227" overlaps with "CHK kinase-like domain". Furthermore, 33 out of 370 upregulated genes are involved in general metabolic processes (a general term)-"Metabolic Pathway" (KEGG_PATHWAY, P=6.6E-4, FDR=9.6E-3), suggesting that the metabolic homeostasis was disrupted in the PG when *sna* is knocked down. This change may indicate defects

in the cellular homeostasis of the PG cells, which may explain why growth (endoreplication) of PG cells was arrested.

Table 2-5. Go term enrichment analysis results for upregulated genes identified in PG>*sna*-RNAi

Category	Term	p value	FDR	Genes
GO_molecular function	Alkaline phosphatase activity	7.6E-05	2.1E-02	<i>Alp4, CG10592, CG3264, CG3292, CG5150, CG5361</i>
GO_cellular component	Extracellular space	2.9E-04	3.0E-02	31 genes, not shown
GO_cellular component	Cell surface	3.8E-04	1.9E-02	<i>Alp4, CG10592, CG3264, CG3292, CG5150, Dl, Est-6, alpha-Est1, rols</i>
KEGG	Metabolism of xenobiotics by cytochrome P450	4.8E-05	2.8E-03	<i>CG4302, CG5999, GstD2, GstD5, GstD7, GstE1, GstE10, GstE9, Ugt35b</i>
KEGG	Drug metabolism - cytochrome P450	4.8E-05	2.8E-03	<i>CG4302, CG5999, GstD2, GstD5, GstD7, GstE1, GstE10, GstE9, Ugt35b</i>
KEGG	Glutathione metabolism	5.4E-05	1.6E-03	<i>CG4302, CG5999, GstD2, GstD5, GstD7, GstE1, GstE10, GstE9, Ugt35b</i>
KEGG	Folate biosynthesis	1.9E-04	3.6E-03	<i>Alp4, CG10592, CG3264, CG3292, CG5150, CG5361</i>
KEGG	Metabolic pathways	6.6E-04	9.6E-03	34 genes, not shown
InterPro domain	Alkaline phosphatase	1.3E-05	5.5E-03	<i>Alp4, CG10592, CG3264, CG3292, CG5150, CG5361</i>
InterPro domain	Protein of unknown function DUF227	4.6E-05	1.0E-02	<i>CG10513, CG10514, CG10559, CG11892, CG11893, CG32195, CG3351, CG6834, CG9259</i>
InterPro domain	CHK kinase-like	4.6E-05	1.0E-02	<i>CG10513, CG10514, CG10559, CG11892, CG11893, CG32195, CG33510, CG6834, CG9259</i>

>2 fold up-regulation was used as the cut-off.

GO enrichment analysis was carried out using DAVID GO.

FDR: False Discovery Rate.

I also used STRING to generate functional association networks for the >3-fold upregulated gene set. “Alkaline phosphatase activity”, “Glutathione metabolism” and “Metabolic

pathways (general)” functional groups showed up in the big interaction network (Fig. 2-37). These three terms were previously identified by the DAVID GO analysis as overrepresented with the FDR less than 0.05 using the 370 >2-fold upregulated genes (Table 2-5), confirming that these three functions were affected when *sna* was knocked down specifically in the PG. However, the FDR for “Glutathione metabolism” and “Metabolic pathways” went beyond the FDR<0.05 cut-off, when only the 200 >3-fold upregulated genes were used (Fig. 2-37), suggesting some of the genes in these three functional groups are only moderately upregulated. STRING analysis revealed one more functional interaction subgroup “proteolysis” (Fig. 2-37), which was not identified by DAVID GO analysis because it was calculated to have an FDR=0.18.

Alkaline phosphatases are usually glycoproteins that localize to the cell membrane (202,203) and in *Drosophila* have been reported to dramatically increase in activity during L3 stage prior to the secretion of the pupal cuticle, possibly regulated by the RG (204). None of the six Alkaline phosphatases that were upregulated in *sna*-RNAi are highly expressed in the RG according to the previously mentioned RG microarray analysis from our lab (31); therefore their possible functions could have been overlooked in the RG. One can validate the expression of these alkaline phosphatase genes in the PG>*sna*-RNAi RGs via qPCR in the future and further examine the role of alkaline phosphatases in the PG and how they are regulated by *Sna* function. Nine genes that are upregulated in *sna*-RNAi RG are predicted to encode proteins containing the CHK kinase-like domain. CHK kinase-like refers to Zinc finger C4 and HLH domain containing kinases domain subfamily. One interesting member of this protein family is the ecdysteroid 22-kinase, which in silkworms phosphorylates ecdysteroids at position C-22, which inactivates ecdysone by producing ecdysteroid 22-phosphates (205). However, the ecdysteroid 22-kinase was not identified in *Drosophila*. The nine “CHK kinase-like” genes also have not been characterized by any experiments, therefore, could be looked into in the future.

Ultimately I did not identify a set of genes that was directly involved in the regulation of cell cycle or endoreplication. One possible explanation could be that those genes may be affected at a post-transcriptional level, *e.g.* disrupting *sna* could affect the expression of some proteases or kinases (as the term “Alkaline phosphatase activity” and “CHK kinase-like domain” are enriched for the upregulated genes in PG>*sna*-RNAi) and in turn affect the protein levels of the endoreplication-related cell cycle regulators.

2.3.11.2 Genes that are differentially expressed in ring glands after *sna* overexpression

I listed the top 50 downregulated genes as well as the top 50 upregulated genes in the *sna* overexpression samples in Appendix (Table A-2).

Downregulated genes in *hs>sna*-cDNA ring glands and GO term enrichment analysis

In *hs>sna*-cDNA RGs, expression of 245 genes were more than 3-fold downregulated and 528 genes were >2-fold downregulated. For the GO analysis via DAVID I used the smaller set of 245 genes. None of the enriched terms met the cut-off of a FDR<0.05. Hence I listed the four enriched GO terms with the smallest FDR value and these terms are summarized in Table 2-6. Interestingly, the term “Proteolysis” (GO_Biological Process, P=5.8E-4, FDR=0.24) was also picked up in the upregulated gene set in the PG>*sna*-RNAi samples (Fig. 2-36), which is internally consistent, since this term was in the “induced” set when *sna* was disrupted. It is indicative that *Sna* might directly regulate the processes of proteolysis in the PG; however, the expression of these proteolysis genes need to be further validated by qPCR in both *sna*-RNAi and *sna*-overexpressing RGs.

Table 2-6. Go term enrichment analysis results for downregulated genes identified in *hs>sna*-cDNA ring glands

Category	Term	<i>p</i> value	FDR	Genes
GO_Biological Process	Proteolysis	5.8E-04	2.4E-01	<i>CG10764, CG11034, CG11529, CG18478, CG18754, CG3097, CG31219, CG31827, CG33462, CG42335, CG4793, CG8773, CG9737, Fur1, grass, Mmp1, Ser6, scaf</i>
GO_Biological Process	Open tracheal system development	7.3E-04	1.6E-01	<i>Hr78, Mmp1, bnl, esg, hairy, hid, serp, verm</i>
GO_Biological Process	Salivary gland cell autophagic cell death	9.7E-04	1.4E-01	<i>CG3829, Ect3, Hr78, Mmp1, emp, hid, rpr</i>
InterPro domain	Haemolymph juvenile hormone binding	1.6E-04	5.6E-02	<i>CG14258, CG2016, CG33306, CG34316, CG7916, CG7968</i>

>3-fold downregulation was used as the cut-off.

GO enrichment analysis was carried out using DAVID GO.

FDR: False Discovery Rate

Intriguingly, I found that overexpressing *sna* affected expression of genes (8 genes/245) involved in open tracheal system development (one of the important tubular structures). According to our previous RG microarray analysis aiming to identify genes that are specifically expressed in the RG, the term “Tube morphogenesis” was enriched (using the program GOSTAT), suggesting the same gene network crucial for tube development (including trachea, salivary glands, and the heart) was somehow reutilized in the RG. *Sna* itself is one of the genes related to tube morphogenesis (120). *esg* and *hairy* (*h*) that fell into the term “open tracheal system development” (Table. 2-6) have already been shown to be bound by the transcription factor Sna in embryos in a ChIP-chip experiment (206). These results again demonstrated that my RNA-Seq procedure was technically sound and capable of extracting relevant biological information from transcriptional changes, especially if combined with existing ChIP results (see section 3.2.11.3). Two other genes under the “Tube morphogenesis” group, *serpentine* (*serp*) and *vermiform* (*verm*), have been shown to genetically interact with each other in the same process, i.e., controlling tracheal tube length (207). These two genes were previously not included in our RG microarray chip. However, at least according to the average absolute RPKM from my RNA-Seq data, *werp* and *verm* are moderately expressed in the RG with the RPKM of 150 and 172 respectively in control samples. As a reference, *nvd* and *spok* had a RPKM of 373 and 792 respectively and the RPKM of *sna* itself was 190 in control samples. In the future, it is worthwhile to confirm whether *serp* and *vermiform* are specifically expressed in the larval PG and whether their expression is regulated by Sna. Moreover, based on our previous genome-wide PG-specific RNAi screen, knocking down *serp* in the PG resulted in L3 larval arrest (VDRC#15466), suggesting *serp* might play an important role in the PG (208).

Furthermore, under the group “Salivary gland cell autophagic cell death”, two of the three main pro-apoptotic genes in *Drosophila* (209) were present, i.e. *reaper* (*rpr*) and *head involution defective* (*hid*), with a fold change of -3.6 (P=0.0289) and -5.5 (P=0.011), respectively. Human Sna protein confers resistance to apoptosis induced by serum depletion in cultured cells and similar results were also reported for mouse Sna *in vivo* (160). It remains possible that Sna protects cells from apoptosis *in vivo* in the PG because overexpressing *sna* reduced the expression of the pro-apoptotic genes.

The last interesting overrepresented term is “Haemolymph juvenile hormone binding”. Juvenile hormone (JH) is the primary hormone produced in the *Drosophila* CA of the RG, which

has a function of designating a larval-to-larval molt to allow further body growth (20,210), while after CW attainment, the level of JH would decrease to allow the onset of metamorphosis (end of growth period). From CA to target tissues, JH is carried by the haemolymph juvenile hormone-binding protein (JHBP), where JHBP protects JH from hydrolysis in the insect haemolymph (211). Moreover, emerging evidence has demonstrated complex crosstalks between JH and ecdysone functions (20,212,213). However, it remains unclear why expression of genes encoding haemolymph juvenile hormone binding domain containing protein was repressed by *Sna*. Could there be a link with the regulation of ecdysone biosynthesis or JH?

Upregulated genes in *hs>sna*-cDNA ring glands and GO term enrichment analysis

109 genes were more than 3-fold upregulated in *sna*-overexpressing RGs and the number went up to 204 when the >2-fold upregulation was applied. Functional annotation analysis by DAVID (using the 204 upregulated genes) showed that the terms “response to bacterium” and “Neuroactive ligand-receptor interaction” were overrepresented (Table 2-7). However, it remains unclear as to whether this is biologically significant.

Table 2-7. Go term enrichment analysis results for upregulated genes identified in *hs>sna*-cDNA ring glands

Category	Term	<i>p</i> value	FDR	Genes
GO_Biological process	Response to bacterium	1.4E-05	6.5E-03	<i>CecA1, IM14, IM23, IM4, Mtk, Sid, Yp3</i>
GO_Biological process	Defense response to Gram-positive bacterium	4.5E-05	1.0E-02	<i>CecA1, IM14, IM23, IM4, Mtk, Sid, Yp3</i>
KEGG	Neuroactive ligand-receptor interaction	8.8E-05	3.2E-03	<i>AdoR, AstA-R2, CG30031, NPFR, deltaTry, gammaTry</i>

>2-fold upregulation was used as the cut-off.

GO enrichment analysis was carried out using DAVID GO.

FDR: False Discovery Rate.

2.3.11.3 Comparing the transcriptional profile of *sna*-RNAi ring glands to that of the *sna* overexpression

Next, I tried to identify potential direct targets of *Sna* using the RNA-Seq data as I expected that expression of potential direct targets of *Sna* should be upregulated in *PG>sna*-RNAi and at the same time downregulated in *hs>sna*-cDNA RGs, or *vice versa*. For this comparison, I used the cut-off of >2-fold differentially expressed genes in all conditions. Only 11 genes were more than 2-fold downregulated in *PG>sna*-RNAi but upregulated in *sna* overexpression and the results

are presented in the Venn diagram (Fig. 2-38A). The overlap was considered significant ($P=2.7E-05$, χ^2 test) because only three genes were expected to be present in both gene sets by a random chance given the large size of the *Drosophila* genome. The 11 overlapping genes are listed in Table. 2-8. Next, as shown in Fig. 2-38B, with the same criteria, 32 genes were found to be upregulated in PG>*sna*-RNAi (total 370 genes) but downregulated in the PG>*sna*-overexpression (total 528 genes), while only 16 genes would be expected to be present in both conditions by random chance ($P=1.96E-4$). The 32 overlapping genes along with their functional annotation are listed in Table. 2-9. Interestingly, six out of the 32 genes encode membrane transporters (FDR=0.0385 based on STRING Functional enrichment analysis), suggesting that Sna could be involved in regulating cellular homeostasis via membrane transport activity. However, in the future, I need to carry out qPCR validation of these six membrane transporters genes in both PG>*sna*-RNAi and PG>*sna*-cDNA RG samples before making any conclusion.

Table 2-8. Genes that are >2-fold downregulated in PG>*sna*-RNAi and >2-fold upregulated in *hs*>*sna*-cDNA ring glands

Symbol	<i>sna</i> -RNAi FC	<i>P</i> value	<i>sna</i> -cDNA FC	<i>P</i> value	Function	Enrichment in the RG*
<i>CG15528</i>	-3.0	0.18	3.5	0.08	protein tyrosine phosphatase	~2 fold
<i>CG31156</i>	-2.0	0.11	3.6	0.03	nucleic acid binding	~5 fold
<i>CG6660</i>	-11.0	0.08	3.6	0.02	ELO family, functions in embryonic/larval trachea	no
<i>CG7587</i>	-9.2	0.08	34.6	0.12	unknown	no
<i>CG7730</i>	-3.6	0.03	2.7	0.00	domain of unknown function DUF4781, learning/memory	>20 fold
<i>Eig71Ee</i>	-4.6	0.22	20.0	0.06	puparial adhesion	no
<i>loh</i>	-7.1	0.08	2.7	0.03	positive regulation of extracellular matrix assembly, larval heart development	>10 fold
<i>rdgA</i>	-2.0	0.16	2.0	0.01	diacylglycerol kinase, protein kinase C-activating G-protein coupled receptor signaling pathway	n.a.
<i>salr</i>	-3.6	0.03	2.8	0.01	zinc finger C2H2 transcription factor	~5 fold
<i>Sgs1</i>	-3.3	0.19	12.8	0.14	puparial adhesion	no
<i>side</i>	-3.2	0.08	2.3	0.01	immunoglobulin subtype 2, motor neuron axon guidance	>10 fold

FC: fold change.

*Enrichment in the RG: indicates whether the specific transcript is enriched in the ring gland based on the previous microarray analysis comparing gene expression profile of the ring gland to that of the whole body larva.

n.a.: not applicable, meaning the specific gene was not included in the original array (Ou *et al.*, 2016).

Table 2-9. Genes that are >2-fold upregulated in PG>*sna*-RNAi and >2-fold downregulated in *hs*>*sna*-cDNA ring gland samples

Symbol	<i>sna</i> -RNAi FC	<i>p</i> value	<i>sna</i> -cDNA FC	<i>p</i> value	Function	Enrichment in the RG*
<i>CG15534</i>	10.9	0.176	-2.0	0.039	sphingomyelin phosphodiesterase activity	no
<i>CG13160</i>	7.4	0.009	-2.6	0.089	Peptidase M28	no
<i>CG3014</i>	7.2	0.184	-3.2	0.003	Arrestin-like	no
<i>CG6283</i>	5.4	0.099	-2.1	0.062	phosphatidylcholine 1-acylhydrolase activity, lipid catabolic process	no
<i>CG11529</i>	4.5	0.035	-3.8	0.013	neurogenesis, serine-type endopeptidase activity	no
<i>tobi</i>	4.2	0.017	-2.9	0.085	alpha-1,4-glucosidase activity	n.s.
<i>CG14879</i>	4.0	0.092	-4.0	0.153	Galectin, carbohydrate recognition domain	no
<i>GstD7</i>	3.9	0.095	-3.0	0.018	glutathione transferase	no
<i>aay</i>	2.8	0.039	-3.9	0.019	L-serine biosynthetic process	no
<i>CG5999</i>	2.4	0.123	-4.4	0.030	glucuronosyltransferase activity, metabolic process	no
<i>CG31233</i>	2.1	0.106	-2.5	0.041	proteolysis	no
<i>CG43179</i>	2.8	0.105	-3.0	0.075	Serine protease gd, N-terminal domain	n.s.
<i>CG33306</i>	3.8	0.128	-4.1	0.041	Haemolymph juvenile hormone binding	no
<i>CG7916</i>	3.6	0.056	-4.0	0.087	Haemolymph juvenile hormone binding	no
<i>CG7968</i>	3.5	0.202	-3.1	0.131	Haemolymph juvenile hormone binding	no
<i>CG1143</i>	9.8	0.012	-2.2	0.076	unknown	no

<i>CG34301</i>	3.5	0.052	-9.4	0.015	unknown	no
<i>CG42615</i>	3.2	0.280	-3.5	0.108	unknown	n.s.
<i>Cyp304a1</i>	4.9	0.236	-2.8	0.136	Cytochrome P450	no
<i>Cyp4d14</i>	2.7	0.025	-2.0	0.082	Cytochrome P450	no
<i>CG10911</i>	3.0	0.196	-3.6	0.082	Protein of unknown function DUF725	no
<i>CG14132</i>	2.0	0.079	-2.4	0.160	Protein of unknown function DUF725	no
<i>nord</i>	2.9	0.096	-2.1	0.021	Protein of unknown function DUF2369, learning/memory	~2 fold
<i>Picot</i>	2.5	0.045	-4.0	0.082	phosphate ion transport	no
<i>NaPi-III</i>	2.4	0.032	-2.3	0.032	sodium-dependent phosphate transport	n.s.
<i>ine</i>	2.2	0.111	-4.5	0.116	neurotransmitter:sodium symporter activity, perineurial glial growth	no
<i>Slc45-1</i>	2.1	0.057	-6.2	0.033	sucrose:proton symporter activity	no
<i>CG5853</i>	2.1	0.018	-3.2	0.009	ABC transporter-like, phagocytosis	no
<i>CG7442</i>	3.0	0.005	-2.6	0.118	organic cation transmembrane transporter	no
<i>CG14636</i>	2.7	0.055	-2.4	0.069	sensory perception of sound	no
<i>CG8736</i>	2.5	0.044	-2.3	0.104	structural constituent of cuticle	no
<i>CG6739</i>	2.2	0.062	-2.6	0.006	Low-density lipoprotein (LDL) receptor class A	no

FC: fold change.

*Enrichment in the RG indicates whether the specific transcript is enriched in the ring gland based on the previous microarray analysis comparing gene expression profile of the ring gland to that of the whole body larva.

n.a.: not applicable, meaning the specific gene was not included in the original array (Ou *et al.*, 2016)

2.3.11.4 Comparing *sna* RNA-Seq results to existing ChIP-on-chip data

Sna has many functions during *Drosophila* embryogenesis including mesoderm fate determination and neurogenesis (127,128,158,214). Given the importance of *Sna*, several ChIP-on-Chip and ChIP-Seq analyses have been carried out using embryo samples

(156,206,215). The Berkeley *Drosophila* Transcription Network Project (BDTNP) also released ChIP-on-Chip data for 21 transcription factors important for early embryogenesis, including Sna at stage 5. I did not expect my RNA-seq data to significantly overlap with the ChIP-on-Chip results because (1) the general transcriptional programs could be different in embryos versus in the more specified endocrine tissue RGs. (2) RNA-Seq would pick up differentially expressed genes that are not directly related to Sna function while ChIP-on-chip would show Sna binding but may not necessarily turn on or off transcription since the Sna mode of action is also dependent on co-binding of other transcription factors, *e.g.* Twist in the embryo (206,216). However, I still looked for overlaps just to determine which of the differentially expressed genes from my RNA-Seq data have previously reported Sna binding sites and could still be the target genes of Sna in the PG.

For comparison, I used the released data from BDTNP, where they used two anti-Sna antibodies recognizing two distinct yet adjacent peptides in Sna protein (<http://bdtnp.lbl.gov/Fly-Net/browseChipper.jsp>). Using anti-Sna 1 they identified 596 binding regions with the cut-off for false discovery rate (FDR) less than 1%, while anti-Sna 2 gave 2800 binding regions (FDR<1%). The overlap between the two data sets was 100%, and I used the anti-Sna 1 data set, which is more likely to represent genuine Sna targets. Also, I included other known direct *Drosophila* Sna target genes published in the literature (156), which resulted in 574 unique genes that were bound by Sna. Surprisingly, among the 560 downregulated genes (>2-fold change) in *sna*-overexpressing samples, 41 genes have Sna binding sites based on ChIP-on-chip data, which is considered significant ($P=2.85E-05$, χ^2 test) (Fig. 2-39) since the expected overlap between two random gene sets with a similar size was 22. However, significant overlap was not observed in the other RNA-Seq conditions as I expected (data not shown). There are two possible explanations for the significant overlap between 560 >2-fold downregulated genes and 574 genes associated with Sna binding peaks: 1). Since the actual overlap (41 genes) is less than 2-fold enriched than the expected random overlap (22 genes), it is still possible that I observed the overlap by chance. 2). Given that Sna primarily functions as a transcriptional

repressor, perhaps high ectopic expression of Sna in the PG represses some genuine Sna targets in other biological processes that are not primary targets in the PG, thus leading to significant overlap between the downregulated genes in *sna*-overexpression and the ChIP-on-chip results. Table. 2-10 lists all the >2-fold differentially expressed genes in *sna*-RNAi that were shown to be bound by Sna and those genes were ordered according to functional annotation. I noticed that there are Sna-binding regions adjacent to the two ecdysone biosynthetic genes, *sad* and *phm*, as well as two JNK pathway regulators, *Traf4* and *egr* (*egr*), both of which were moderately downregulated when *sna* was knocked down in the PG (Table. 2-10). Later on, I also validated the expression of *Traf4* and *egr* by qPCR (Fig. 2-40). In the future, one can further test whether *Traf4* and *egr* are direct targets of Sna in the PG by ChIP-qPCR using ring gland samples. Interestingly, Eiger is the Tumor necrosis factor (TNF) ligand and Traf4 is the adapter protein binding to the TNF receptor. These two proteins are often involved in regulating the cell polarity during cell morphogenesis as well as regulating apoptosis through JNK pathway (217-220). Interestingly, both *Traf4* and *egr* were moderately enriched in the RG compared to the whole body according to the RG microarrays carried out by our lab (120), suggesting a role in the RG. Some other Sna-bound cell signaling genes such as *frizzled* (*fz*) and *derailed* (*drl*) were moderately upregulated in PG>*sna*-RNAi (Table 2-10). These two genes (*fz* and *drl*) participate in Wnt-signaling and are involved in processes like cell fate determination and guided cell migration, both processes that require polarized distribution of proteins within cells as well as relaying cell signals to neighboring cells (221-224).

The overlap between genes deregulated in *hs>sna*-cDNA RGs and genes associated with Sna-binding peaks is shown in Table. 2-11, among which are several genes involved in cell fate determination (e.g. *esg*, *melt*, *pros* and *Tom*, FDR= 0.021 based on DAVID GO). In summary, my transcriptional profiling analysis suggested that Sna might regulate similar processes (i.e. cell polarity, cell migration and cell fate determination) in the PG and in the embryo. Are these above-mentioned processes critical for ecdysone production/release in the PG? What specific biological processes in the PG require cell-cell communication and cell polarity? These ideas need to be studied in the future.

Table 2-10. Overlap between genes misregulated in PG>*sna*-RNAi ring glands and genes associated with Sna binding peaks in the embryo stage*

Gene symbol	<i>sna</i> -RNAi FC	<i>p</i> value	Note	PG-RNAi phenotype#
GO biological process, Sleep, P=0.00028				
<i>Rab27</i>	3.5	0.01	vesicle-mediated transport	NOP
<i>Dl</i>	2.6	0.15	Notch binding, positive regulation of cell proliferation	NOP
<i>CG10830</i>	2.3	0.09	Potassium channel tetramerisation-type BTB domain	NOP
<i>Sodh-1</i>	3.2	0.05	oxidoreductase activity	NOP
<i>aay</i>	2.8	0.04	L-serine metabolic process	NOP
GO biological process, imaginal disc-derived wing morphogenesis, P=0.019				
<i>fz</i>	2.3	0.03	Wnt-protein binding	NOP
<i>drl</i>	2.1	0.13	Wnt-protein binding	NOP
<i>dve</i>	2.1	0.01	Homeobox domain	NOP
GO biological process, ecdysone biosynthetic process, P=0.034				
<i>sad</i>	-5.9	0.17	Ecdysteroid biosynthetic gene	L1 arrest
<i>phm</i>	-5.4	0.10	Ecdysteroid biosynthetic gene	L1 arrest
GO biological process, asymmetric protein localization involved in cell fate determination, P=0.036				
<i>egr</i>	-2.6	0.18	TNF receptor binding	NOP
<i>Traf4</i>	-2.6	0.00	ventral furrow formation	NOP
<i>T48</i>	-2.2	0.13	ventral furrow formation	minor delayed L3
Other				
<i>CG10176</i>	2.1	0.09	TNF-activated receptor activity	NOP
<i>ImpE2</i>	11.5	0.08	Ecdysone-inducible gene E2	NOP
<i>CG9837</i>	6.4	0.04	unknown	NOP
<i>CG7968</i>	3.5	0.20	haemolymph juvenile hormone binding	NOP
<i>Nek2</i>	3.2	0.19	protein kinase	NOP
<i>Cyp4d20</i>	3.1	0.07	Cytochrome P450	NOP
<i>cad</i>	3.0	0.23	transcription factor	L1, L2 arrest, and large L3
<i>aos</i>	2.9	0.06	antagonist of EGFR signalling	NOP
<i>babos</i>	2.2	0.04	cell adhesion	NOP
<i>CG11123</i>	2.1	0.06	RNA binding	NOP
<i>edl</i>	2.1	0.02	negative regulation of transcription	NOP

<i>Elba2</i>	2.0	0.14	chromatin silencing	NOP
<i>bnb</i>	-7.7	0.18	bangles and beads	NOP
<i>CG3097</i>	-4.6	0.19	proteolysis	NOP
<i>CG30156</i>	-3.1	0.28	heat shock protein 40/Dnaj co-chaperone	minor delay
<i>Muc26B</i>	-2.7	0.13	extracellular matrix structural constituent	NOP
<i>CG3887</i>	-2.7	0.03	selenoprotein T like	NOP
<i>cib</i>	-2.3	0.14	actin monomer binding	NOP
<i>fus</i>	-2.2	0.06	mRNA binding, Egfr signalling pathway	NOP
<i>lace</i>	-2.2	0.09	regulation of Wnt and Notch signaling pathway	NOP
<i>CG31051</i>	-2.1	0.27	unknown	NOP
<i>CG18508</i>	-2.1	0.11	signal transduction	NOP

* Sna binding peaks are based on the ChIP-on-chip results released by the Berkeley *Drosophila* Transcription Network Project (BDTNP). The samples were collected from embryo stage 5.

PG-RNAi phenotypes are based on the results of the genome-wide PG>RNAi screen (Danielsen *et al.*, 2016).

Table 2-11. Overlap between genes misregulated in *sna*-overexpression ring glands and genes associated with Sna binding peaks¹

Gene symbol	<i>sna</i> -cDNA FC	<i>p</i> value	Note	PG-RNAi phenotype
GO biological process, peripheral nervous system development, P=0.000012, FDR=0.0038				
<i>esg*#</i>	-4.3	0.07	transcriptional repressor, maintenance of imaginal histoblast diploidy	n.a.
<i>aay</i>	-3.9	0.02	L-serine biosynthetic process	NOP
<i>shn#</i>	-2.3	0.08	Zinc finger C2H2-type transcription factor	NOP
<i>melt*</i>	-2.1	0.07	sequestering of triglyceride, insulin receptor signaling pathway; cell fate specification	NOP
<i>pros*#</i>	2.2	0.02	Homeodomain-like transcription factor activity, R7 cell fate commitment	NOP
<i>neur</i>	4.4	0.02	germ-line stem cell population maintenance, regulation of Notch signaling pathway, ubiquitin protein ligase	NOP
GO biological process, compound eye retinal cell programmed cell death, P=0.00057, FDR=0.043				

<i>hid</i>	-5.5	0.01	programmed cell death	L1 and L2 arrest
<i>ec</i>	-2.0	0.09	ubiquitin-specific protease activity, retinal cell programmed cell death	NOP
<i>S</i>	2.0	0.04	epidermal growth factor receptor signaling pathway, malpighian tubule morphogenesis, stem cell fate commitment	NOP

***GO biological process, cell fate specification, P= 2.1E-4, FDR= 2.1E-2**

<i>Tom</i>	-3.5	0.14	Notch signaling pathway; cell fate specification	NOP
------------	------	------	--	-----

#GO biological process, regulation of transcription from RNA polymerase II promoter, P=0.000015, FDR=0.0023

<i>bowl</i>	-6.1	0.01	Zinc finger C2H2 transcription factor; hindgut morphogenesis	NOP
<i>h</i>	-5.1	0.01	bHLH transcription repressor, salivary gland morphogenesis	NOP
<i>ci</i>	-2.9	0.12	Zinc finger, C2H2 transcription factor activity, Hedgehog signaling complex	NOP
<i>ovo</i>	-2.2	0.00	Zinc finger C2H2 transcription factor	L2 arrest

Other

<i>CG32392</i>	4.4	0.00	microtubule associated complex	NOP
<i>Mkp3</i>	3.3	0.07	protein tyrosine/serine/threonine phosphatase activity; negative regulation of MAPK cascade	NOP
<i>CG31051</i>	2.8	0.00	unknown	NOP
<i>upd1</i>	2.6	0.04	intestinal stem cell homeostasis, positive regulation of JAK-STAT cascade	NOP
<i>CG9837</i>	2.4	0.09	unknown	NOP
<i>Atx-1</i>	-17.2	0.00	RNA binding, photoreceptor cell maintenance, imaginal disc-derived wing vein specification	NOP
<i>CG12056</i>	-5.7	0.02	Cytochrome b5-like heme/steroid binding domain	NOP
<i>CG3097</i>	-5.5	0.05	proteolysis	NOP
<i>spz5</i>	-3.8	0.07	Toll signaling pathway	L3 arrest
<i>pll</i>	-3.7	0.00	protein phosphorylation; Toll signaling pathway	NOP
<i>CG31523</i>	-3.5	0.03	fatty acid elongase activity, fatty acid biosynthetic process	NOP

<i>Pmp70</i>	-3.4	0.03	ABC transporter-like, peroxisome organization	NOP
<i>gem</i>	-3.4	0.03	transcription factor, pointed domain	NOP
<i>drongo</i>	-3.2	0.00	regulation of GTPase activity	NOP
<i>Nbr</i>	-3.2	0.03	mature miRNA 3'-end processing	NOP
<i>CG7968</i>	-3.1	0.13	Haemolymph juvenile hormone binding	NOP
<i>egl</i>	-2.9	0.00	intracellular mRNA localization, germarium-derived oocyte fate determination	NOP
<i>mas</i>	-2.9	0.03	proteolysis	NOP
<i>Kap3</i>	-2.9	0.03	kinesin complex; microtubule-based movement	NOP
<i>Inx2</i>	-2.8	0.01	foregut morphogenesis, intercellular transport, cell communication	NOP
<i>CG3558</i>	-2.8	0.03	Armadillo-like helical	NOP
<i>Cys</i>	-2.7	0.01	cysteine-type endopeptidase inhibitor activity, multicellular organism reproduction	NOP
<i>comm2</i>	-2.5	0.06	axon guidance	NOP
<i>Cpr</i>	-2.5	0.06	oxidation-reduction process	NOP
<i>uif</i>	-2.4	0.04	regulation of tube architecture, open tracheal system, Notch binding	NOP
<i>veil</i>	-2.4	0.00	nucleotide catabolic process	NOP
<i>CG12581</i>	-2.4	0.08	PH domain-like	NOP
<i>CG10211</i>	-2.4	0.07	Haem peroxidase, oxidation-reduction process	NOP
<i>Mnn1</i>	-2.3	0.06	regulation of JNK cascade, cellular response to DNA damage stimulus	NOP
<i>CG15093</i>	-2.3	0.02	L-valine degradation	NOP
<i>CG11160</i>	-2.3	0.00	histone deacetylase binding; negative regulation of gene expression	NOP
<i>sad</i>	-2.2	0.03	ecdysone biosynthetic process	L1 arrest
<i>Cyp6v1</i>	-2.2	0.07	Cytochrome P450	NOP
<i>CG10188</i>	-2.1	0.01	Rho guanyl-nucleotide exchange factor activity	NOP
<i>CG31342</i>	-2.0	0.05	PH domain-like	NOP

*Genes that are also involved in cell fate specification.

Genes that also encode for transcription (regulation of transcription form RNA polymerase II promoter).

¹-Sna binding peaks are based on the ChIP-on-chip results released by the Berkeley Drosophila Transcription Network Project (BDTNP). The samples were collected from embryo stage 5.

²-PG-RNAi phenotypes are based on the results of the genome-wide PG>RNAi screen (Danielsen *et al.*, 2016). n.a.: not applicable, genes were not included in the original genome-wide RNAi screen.

2.3.11.5 Validation of differentially expressed genes via qPCR

Among the differentially expressed genes identified from PG>*sna*-RNAi samples, ecdysone-related genes *tor*, *nvd*, *sro*, *spok*, *phm*, *dib* and *sad* were already validated to be significantly downregulated (Fig. 2-5). Moreover, I also examined the expression of two known Sna targets *egr* and *Traf4* (Table 2-10 and Fig. 2-40A) in PG>*sna*-RNAi RGs via qPCR and I observed modest down-regulation as observed in RNA-Seq analysis.

For the *sna* overexpression RNA-Seq results, the GO term “Salivary gland cell autophagic cell death” was enriched for the >3-fold downregulated gene set (Table. 2-6). I validated two out of seven genes under this term, i.e. *rpr* and *hid* (Fig. 2-40B). If I expanded the downregulated genes to >2-fold, 12 genes would be under this term (data not shown) and I verified two more genes out the 12, i.e. Ecdysone receptor (*EcR*) and Death caspase-1 (*Dcp-1*). All these four genes (*rpr*, *hid*, *EcR* and *Dcp-1*) gave consistent results between RNA-Seq and qPCR except that the fold change for *Dcp-1* was not significant in the qPCR analysis (Fig. 2-40B).

I also noticed that expression of the ecdysone biosynthetic gene *dib* was reduced 17-fold in *sna* overexpression samples (ranked #17 among the downregulated genes, Table A-2). Considering that two ecdysone biosynthetic genes *sad* and *phm* have potential Sna binding sites (Table 2-10), I validated expression of all six ecdysone biosynthetic genes in the PG>*sna*-overexpression samples. As shown in Figure. 2-40B, there was good consistency between the RNA-Seq and qPCR results with *sro* and *dib* being severely reduced, as well as *phm* and *sad* being moderately but significantly downregulated. This suggested that ecdysone production was disrupted when *sna* is overexpressed. This is consistent with the observation that overexpressing *sna* in the PG resulted in larval developmental arrest (Fig. 2-32), a phenotype often caused by ecdysone deficiency.

Previously, our lab demonstrated that the circadian machinery in the PG is an essential component of the ecdysone production machinery. It is interesting that, when *sna* is overexpressed, the circadian gene *period* (*per*) was 7.6-fold downregulated (ranked #40 among the downregulated set) and that another circadian gene, *tim*, was 4.6-fold upregulated (ranked #48 among the upregulated genes) (Table A-2), suggesting deteriorated circadian rhythm in the PG, which could further perturb ecdysone production. The expression of *per* and *tim* in *sna*-overexpressing RGs was also validated by qPCR (Fig. 2-40).

Next, I validated three known *Sna* targets that were differentially expressed when *sna* was overexpressed in the PG (Table. 2-10, 2-11), which included *neuralized* (*neur*), *esg* (another member of Snail family transcription factor) and *hairy* (*h*), where only the changes in *neur* expression was not significant based on qPCR (Fig. 2-40). Sox Neuron Co-factor (*SNCF*) is another known *Sna* target according to the literature (120,156); however, it missed the cut-off for >3-fold upregulated genes in *sna* overexpression since it has a high *p* value ($P=0.19$) in RNA-Seq. As I examined the expression of *SNCF* by qPCR in *sna* overexpression samples, the change was indeed not significant.

Lastly, since the PTTH/MAPK axis activated through Torso is one of the best understood signals in the PG to regulate ecdysone biosynthesis, it caught my attention that the Extracellularly regulated kinase 7 (*Erk7*), an atypical MAP kinase, was among the top 50 upregulated genes (ranked #17, Table. A-2) after *sna* overexpression. Moreover, *torso like* (*tsl*), an upstream regulator of Torso during the process of anterior/posterior terminal specification in the embryos (225), was 2.7-fold upregulated. *tsl* is also expressed in the *Drosophila* PG and has been shown to regulate developmental timing and onset of metamorphosis independent of Torso function, suggesting a link to ecdysone production (33,226-228). Moreover, Trunk (*Trk*), the presumed Torso ligand in the embryo during posterior terminal specification (opposed to the ligand, PTTH, in the PG) was also ~2.3-fold upregulated upon *sna* overexpression ($P=0.07$) based on the RNA-Seq. Therefore, *Erk7*, *tsl* and *trk* were also included in the qPCR validation. Fig. 2-40 shows that both *Erk7* and *tsl* were significantly upregulated in *hs>sna*-cDNA RGs, but not as profound as the fold change found by RNA-Seq, while *trk* upregulation was not reproducible by qPCR. It remains unclear whether the upregulation of these two genes (*Erk7* and *tsl*) in *sna* overexpression was related to the regulation of ecdysone production.

2.3.12 Transcription of *ouija board* (*ouib*) is controlled by IIS/TOR signaling in the PG

The transcription factor *Ouib* (*aka* CG11762), which specifically controls expression of *spok*, had reduced transcript levels in PG>*sna*-RNAi RGs (Table 2-3, 2-4). A previous study has shown that *ouib* expression is not regulated by PTTH/MAPK signaling in the PG (43). Since *Sna* is responsive to IIS/TOR signaling, I hypothesized that *ouib* expression might also be controlled by IIS/TOR signaling. Therefore, I tested *ouib* levels via qPCR in PG>*TOR*-RNAi, PG>*TOR*^{DN} and PG>*Pi3K*-RNAi RGs respectively. I found that *ouib* expression is consistently affected by loss-of-TOR as well as in *Pi3K*-RNAi (*Pi3K*, a downstream effector of the IIS signaling) (Fig. 2-41). This finding suggested that *ouib* levels are not controlled by PTTH, but rather by IIS/TOR signaling, which is responsive to nutrient input.

2.3.13 The cell adhesion molecule N-Cadherin is present in the CA of the ring gland.

The best-known function of *Sna* is to regulate the process of EMT by controlling the expression of the cell adhesion protein E-Cadherin. I asked if *Sna* also regulates the same target in the PG. However, E-Cadherin (*shotgun*) is only marginally expressed in the RG (120) (Fig. 2-42A), which suggested that it might not be essential for RG function. On the other hand, a different Cadherin, N-Cadherin (*CadN*), is 5-fold enriched in the RG at least at the beginning of the L3 stage and eventually declines in expression at later time points, an expression pattern that is similar to that of *sna* (120) (Fig. 2-42B). *CadN* was also shown to have binding sites for *Sna* (156), raising the idea that *Sna* regulates *CadN* transcription in the PG. First, I needed to determine whether *CadN* was expressed in the PG or in the other two glands. Immunofluorescent staining showed that *CadN* is specifically present at the cell boundary of CA cells (Fig. 2-42B). Multiple time points were tested during the L2 and L3 stages and I observed a subtle reduction of *CadN* levels in the CA throughout development; however, *CadN* signals were never observed in the other part of the RG at all time points tested (data not shown). Therefore, I concluded *CadN* is most likely not related to *Sna* function in the PG.

2.4 Discussion and future directions

2.4.1 Snail function in the PG is essential for ecdysone biosynthesis

Sna is dynamically expressed in the PG, the main endocrine organ for ecdysone biosynthesis during larval stages (Fig. 2-2). My results showed that lack of *Sna* in the PG affected transcript levels of the six main ecdysone biosynthetic genes, consistent with a low ecdysteroid

titer in the larvae (Fig. 2-5). Consequently, the *sna*-RNAi animals displayed developmental arrest since ecdysone triggers each developmental transition (Fig. 2-2). However, high levels of Sna in the PG also appeared to block ecdysone biosynthesis, evidenced by the fact that overexpressing *sna* in the PG caused developmental arrest (Fig. 2-32, also see section 2.3.10, data not shown) and that four out of six ecdysone biosynthetic genes *sro*, *phm*, *dib* and *sad* were downregulated (Fig. 2-40). These results suggested that the levels of Sna need to be tightly controlled in the PG during development for its proper function. It seems that ecdysone biosynthesis requires Sna, but at the same time Sna does not allow for maximal production of ecdysone. This hypothesis would also explain why *sna* levels are drastically reduced from the second half of the last larval stage (L3) (Fig. 2-2) and became almost undetectable at the end of the L3 stage. A major ecdysone peak needs to occur at this time to initiate the onset of metamorphosis. Presumably Sna must be cleared away to allow maximal expression of the ecdysone biosynthetic genes thus, resulting in a high level of ecdysone and another factor must provide the stimulus to trigger the major ecdysone pulse at that time.

2.4.2 Snail might directly regulate the transcription of ecdysone biosynthetic genes

According to the ChIP-chip experiments released by BDTNP, there were Sna binding sites near the *phm* and *sad* genes at least during embryonic stage 5 (Table. 2-10, 2-11), which suggested it could potentially directly regulate the transcription of these two ecdysone biosynthetic genes. I then followed up by carrying out an *in silico* search for potential Sna binding sites in the six major ecdysone biosynthetic genes using the web server, IN-silico SEarch for Co-occurring Transcription factors (INSECT 2.0), based on the consensus Sna binding motifs from multiple databases (229). Using INSECT 2.0, I only searched for transcription binding sites within 2 kb upstream and 1 kb downstream from the transcription start site (TSS). Interestingly, all six ecdysteroidogenic enzyme genes have potential Sna binding sequence (Fig. 2-43). Coincidentally, the previously mentioned Sna binding peaks associated with *phm* started -832 nt relative to the TSS, thus near one of the predicted binding sequences at -917. Similarly, the binding peak associated to *sad* started at +3 relative to the TSS, close to the three predicted Sna binding sequence around 61-75 bp after the TSS (Table 2-12, Fig. 2-43). Some of the predicted binding sites in *sro*, *phm*, *dib* and *sad* (but not in *nvd* and *spok*) were found conserved among several other *Drosophila* species, suggesting that they might be functionally important and represent *bona fide* Sna binding sites (Table 2-12). Moreover, the most confident core consensus sequence CAGGTG

for Sna binding was identified in *sro*, *dib* and *sad* (214). Therefore, I speculated that four ecdysone biosynthetic genes (*sro*, *phm*, *dib* and *sad*) are potentially transcriptionally regulated by Sna (73,75). In the future, one can further confirm these predicted Sna binding region by electrophoretic mobility shift assay (EMSA) or PG-specific ChIP analysis.

Table 2-12. Predicted Sna binding sites for the six major ecdysone biosynthetic genes in *Drosophila melanogaster* using INSECT 2.0

Gene symbol	Predicted binding sequence	Start	End	Conservation	Confidence	Binding Peak
<i>nvd</i>	CCCCCTGCTTC	-1919	-1909	no	medium	
<i>nvd</i>	CCACCTTCTTC	973	983	no	medium	
<i>spok</i>	CCACTTTTTCG	-1371	-1361	<i>D.simulans</i>	medium	
<i>spok</i>	CAACCTGCGAA	657	667	no	medium	
<i>sro</i>	CAACCTGCTGG	522	532	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.erecta</i>	high	
<i>sro</i>	CAAGTG	176	181	<i>D.simulans</i> , <i>D.sechellia</i>	medium	
<i>phm</i>	GAACCTGTTTG	-917	-907	<i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i>	medium	Yes*
<i>dib</i>	CAAACCTGTAA	56	66	<i>D.simulans</i> , <i>D.sechellia</i>	medium	
<i>dib</i>	CAGGTG	444	449	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i>	high	
<i>dib</i>	CAAGTG	495	500	<i>D.simulans</i> , <i>D.yakuba</i> , <i>D.erecta</i> , <i>D.ananassae</i> , <i>D.pseudoobscura</i> , <i>D.persimilis</i> , <i>D.willistoni</i>	medium	
<i>dib</i>	TAACCTGTTTG	738	748	no	medium	
<i>sad</i>	CAACTTTCTGA	-973	-963	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i>	medium	

<i>sad</i>	CAACTTTCTCG	-357	-347	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i>	medium	
<i>sad</i>	GCAGGTGCAGG TGGT	61	75	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i>	medium	Yes*
<i>sad</i>	CAGGTG	62	67	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i>	high	Yes*
<i>sad</i>	CAGGTG	68	73	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i> , <i>D.ananassae</i>	high	Yes*
<i>sad</i>	ACACCTGCTCG	150	160	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i>	medium	
<i>sad</i>	CAACCTTTTGA	493	503	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i>	medium	

Start: start of the predicted Sna binding sequence relative to the transcription start site (TSS).

End: end of the predicted Sna binding sequence relative to the TSS.

High Confidence indicates the actual core consensus sequence with a Positional Weight Matrices (PWM) value of 1. medium: PWM between 0.8 to 0.98.

*Actual binding of Sna to *phm* and *sad* were identified by ChIP-chip analysis (BDTNP). The binding peak for *phm* started at -832 relative to the TSS and the binding peak for *sad* started at +3 relative to the TSS (in close proximity to some predicted Sna binding sites).

Sna has been shown to function as a transcriptional repressor (154,155,230); however, a recent study also showed that it could serve as a potentiator of active transcription. The mode of action (i.e. whether Sna represses or activates gene transcription) depends on, for instance, the co-binding of Twist at the enhancer elements of the target genes, at least during embryogenesis (156). In fact, typically the spatio-temporal regulation of gene transcription is achieved by multiple transcription factors binding to *cis*-regulatory modules (231). Several other transcription factors have been reported to directly bind to some of the ecdysone biosynthetic genes (75,76). Whether Sna represses or promotes active transcription of its target genes probably depends on the genomic

context and co-bound factors. This again explains why both loss of *sna* and overexpression of *sna* reduced transcript levels of the ecdysone production related genes.

In the case of *sna* overexpression, expression of *sro*, *phm*, *dib* and *sad* was repressed, but not *nvd*, *spok*, which correlated with the finding that *sro*, *phm*, *dib* and *sad* contain potential Sna binding sequences found conserved among several other *Drosophila* species, but not *nvd* and *spok*. My explanation is that abnormally high occupancy of Sna on these four genes (*sro*, *phm*, *dib* and *sad*) when overexpressed probably inhibited co-factors binding or possibly recruited alternative or atypical co-factors, thus abolishing normal *sro*, *phm*, *dib* and *sad* expression. In other words, I hypothesize that *sro*, *phm*, *dib* and *sad* are bound and regulated by Sna; however, whether they are repressed or activated by Sna is context-dependent.

To tease out whether a target gene is Sna-activated or Sna-repressed under normal physiological conditions in the PG, both the PG gene expression profile in *sna* loss of function (which I could obtain from my *sna*-RNAi RNA-Seq analysis) and the Sna ChIP binding data for RGs should be taken into consideration. As discussed in section 3.2.11.4, Sna binding sites may differ between the embryo and the PG. Therefore, one may need to carry out a ChIP-Seq analysis using RG samples in the future, probably, around 17 hr L2 stage, when *sna* is maximally expressed. RG ChIP-Seq data may help to identify the core direct targets of Sna in the PG.

2.4.3 Sna as a potential regulator of endoreplication in the PG

Immunofluorescent staining has revealed the peculiar non-uniformed presence of Sna in PG nuclei, which is reminiscent of the pattern of EdU-labeled S-phase cells. The stochastic distribution of S-phase cells in the PG is due to the fact that endocycles in the PG are not synchronized. I showed that the percentage of Sna-positive cells correlated well with the percentage of the S-phase cells (Fig. 2-2, 2-13 and 2-27). Hence, I reasoned that Sna is expressed in a certain time window at each round of endoreplication. Double labelling showed that Sna expressing cells do not largely overlap with S-phase cells (Fig. 2-29), indicating Sna is expressed in the Gap phase of the endocycle, either right before or right after S-phase. Loss of Sna resulted in endocycle arrest in the PG, which suggested that Sna is a potential regulator of endoreplication. However, I have not yet identified any endocycle-related genes that are possibly regulated by Sna based on my *sna* loss-of-function RNA-Seq data. One gene, *CG7910* stood out as being both 6-fold downregulated in PG>*sna*-RNAi (Fig. 2-35) as well as being previously linked to cell cycle

regulation in another RNAi screen (232). Interestingly, knocking down *CG7910* in the PG (Bloomington stock #51702) resulted in L3 arrest, a typical phenotype caused by ecdysone deficiency (data not shown). In the future, one may validate the RNAi phenotype using multiple independent lines and follow up to examine whether the endocycle is affected in the PG in *CG7910* knock-down. Relatively little functional information was available on *CG7910* except the link to the cell cycle, which means it could be a novel regulator in the PG for both the ecdysone biosynthesis and endoreplication processes.

My *in silico* search using INSECT 2.0 revealed that one of the key genes that regulates the process of endoreplication, *double parked* (*dup*), contains the most number of potential Sna binding sites compared to other endocycle-related genes (*E2f1*, *Cult4*, *Cdk2* and *CycE*); therefore only the potential Sna binding sequences in *dup* are shown in Table 2-13 and Fig. 2-44). Four out of nine predicted binding sites were found in almost all 12 *Drosophila* species, suggesting that they are functionally important. Binding of Sna to these sequences could be tested in the future by EMSA. *CycE* has also been reported to have Sna binding peaks in its intron region and proper *CycE* expression pattern in the embryo stage is dependent on Sna (156). However, transcript levels of these two genes were not altered in the *sna*-RNAi RGs according to the RNA-Seq data. One possible reason could be that the time point that I used for the RNA-Seq analysis was 24 hr L3, when endoreplication in the PG was almost finalized (Fig. 2-13 and Fig. 2-19); hence the effect of *sna*-RNAi on the expression of *dup* and *CycE* was no longer as profound. Future work could examine the expression of *dup* and *CycE* level in PG>*sna*-RNAi via qPCR at developmental times when endoreplication were more active (i.e. 17 hr L2 and 12 hr L3). Again, ChIP-Seq data from RG samples would be helpful to identify binding of Sna to the *CycE* and *dup* loci in RGs.

Table 2-13. Predicted Sna binding sites for *Drosophila melanogaster dup* using INSECT 2.0

Binding sequence	Start	End	Conservation	Confidence
CAACCTGTTGT	-1516	-1506	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i> , <i>D.ananassae</i> , <i>D.pseudoobscura</i> , <i>D.persimilis</i>	medium
CAGGTG	-1436	-1431	<i>D.simulans</i> , <i>D.sechellia</i> , <i>D.yakuba</i> , <i>D.erecta</i> , <i>D.ananassae</i> , <i>D.pseudoobscura</i> , <i>D.persimilis</i> , <i>D.willistoni</i> , <i>D.virilis</i>	high

GTACCTGCTGA	-1221	-1211	<i>D.simulans, D.sechellia, D.yakuba, D.erecta</i>	medium
CAGGTG	-872	-867	<i>D.simulans, D.sechellia, D.yakuba, D.erecta</i>	high
CTACTTGCCTC	-577	-567	<i>D.simulans, D.sechellia, D.yakuba, D.erecta, D.ananassae, D.pseudoobscura, D.persimilis, D.virilis</i>	medium
CCACTTGTTCC	-560	-550	<i>D.simulans, D.sechellia, D.erecta, D.pseudoobscura, D.persimilis, D.virilis, D.mojavensis, D.grimshawi</i>	medium
CAAGTG	117	122	<i>D.simulans, D.sechellia, D.yakuba, D.erecta</i>	medium
CAAGTG	223	228	<i>D.simulans, D.sechellia, D.yakuba, D.erecta</i>	medium
CAAGTG	518	523	<i>D.simulans, D.sechellia, D.yakuba, D.erecta</i>	medium

dup: double parked (Flybase_ID FBgn0000996).

Start: start of the predicted Sna binding sequence relative to the TSS.

End: end of the predicted Sna binding sequence relative to the TSS.

High Confidence indicates the actual core consensus sequence with a Positional Weight

Matrices (PWM) value of 1; medium: PWM between 0.8 to 0.98.

Dup is the *Drosophila* ortholog of Cdt1, a conserved component of the pre-replicative complex (pre-RC), which is crucial for DNA replication during the cell cycle. Dup protein is cell cycle-regulated with a decline in protein levels at the transition from G to S phase during each round of cell cycle or endocycle (111,233). Therefore, its expression domain during endocycle potentially matches that of Sna, since Sna is probably present in the nucleus at G phase but not S according to my own results. Moreover, strong *dup* mutant embryos failed to undergo S-phase during postblastoderm divisions. In addition, cell cycle progression was arrested in another aspect that some transcripts induced at the G to S transition, that must be downregulated during S-phase to ensure cell cycle progression, e.g. PCNA, failed to be downregulated in *dup* mutants. If Sna is indeed functionally linked to Dup, it fits my findings that expressing *CycE* (the S-phase cyclin) to push PG cells to reenter the S-phase would not rescue endocycle arrest.

Cell cycle progression and cell survival/death are obviously interconnected, since each phase of the cell cycle has checkpoints allowing cell cycle progression in normal conditions, activation of repair mechanisms if there are any kind of defects or eventually removal of the damaged cell via cell death (234). The apoptosis machinery is a crucial part of cell cycle

checkpoints (235,236). My data demonstrated that overexpressing *sna* reduced the expression of two pro-apoptotic cell death genes, *hid* and *reaper*, suggesting reduced cell death in *sna*-overexpressing RGs. Moreover, loss of *sna* resulted in reduced cell number in the PG and coordinately, overexpressing *sna* increased cell number in the PG (although not yet tested whether it was due to cell survival or proliferation issues in both cases). Decreased cell numbers in the PG were also observed in PG>*CycE*-RNAi RGs (101), while overexpressing *CycE* in the PG resulted in increased cell number (Fig. 2-31). Nevertheless, all these changes in cell numbers and the expression of pro-apoptotic genes might be just byproducts of a misregulated cell cycle/endocycle processes caused by misregulation of Sna. Previous studies have also shown that vertebrate Sna family proteins, Snail1 and Snail 2 regulate cell survival and apoptotic genes (160,163,164,237), which further supports my findings.

2.4.4 Snail as a candidate of the molecular basis for critical weight checkpoint

Various tissues in *Drosophila* larvae undergo endoreplication, but *sna* is quite specific and strongly expressed in the PG (Fig. 2-2). This suggested that *sna* may be involved in a unique aspect of endocycle regulation that is characteristic of the PG. So far I have evidence to suggest that Sna is a strong candidate involved in the underlying molecular mechanism of the CW checkpoint. First, the endocycle progression in the PG is part of the molecular basis of CW attainment where an irreversible decision is made in animals to commit to metamorphosis when PG cells reached a C-value of 16, probably a molecular threshold that determines whether animals have passed the CW checkpoint or not (101). When *sna* was knocked down in the PG, the endocycle was arrested and the C-value of the PG cells failed to progress beyond 16C, meaning the animals may never have made it through the CW checkpoint, consistent with the finding that animals were developmentally arrested at the L3 stage (Fig. 2-46). Second, one round of endoreplication occurs around the CW checkpoint, which was shown by the increased presence of S-phase cells around that time (Fig. 2-13). Likewise, the percentage of Sna⁺ cells in the PG also increased at similar time, suggesting that Sna-nuclear localization is endocycle-related (Fig. 2-27).

Endoreplication in the PG around the CW checkpoint is nutrient-dependent and requires TOR function (101). When animals were physically starved or genetically starved by *TOR^{DN}* before CW attainment (when the C-value did not yet reach 16C), the endocycle was arrested and the C-value stayed below 16C. As a consequence, animals failed to pass through the CW

checkpoint. But when the animals were starved or *TOR^{DN}* was expressed after CW, endoreplication was not affected and was able to reach a final C-value of 64C, followed by metamorphosis (101). Coincidentally, *sna* expression in the PG is dependent on TOR function, as well as nutrient conditions (Fig. 2-21, Fig. 2-22, Fig 2-27 and Fig. 2-46). Interestingly, my results showed that Sna levels were more sensitive to nutrient status before CW attainment than post-CW attainment. Since the percentage of Sna⁺ cells in the PG is rising from the time of pre-CW to the CW checkpoint under normal conditions, the impact of declined Sna levels due to early starvation (before the CW checkpoint) would be significant (Fig. 2-27). In contrast, after CW attainment percentage of Sna⁺ cells declined even in normal conditions; therefore late starvation would not generate a profound effect on the animals (Fig. 2-27). This Sna expression pattern might explain why Sna could be part of the molecular mechanism by which the CW attainment time window is determined. However, there is still not sufficient evidence to show that Sna is the cause of, rather than the downstream effect of, CW attainment. More work should be done to confirm whether Sna is the real molecular determinant of CW attainment. A game-changing experiment would be to starve the physiologically pre-CW animals (for instance, starvation at 4 hr L3), and make the starved animals bypass the failure of CW attainment by expressing Sna in the PG. However, it would not be a trivial task to control the proper level of Sna to push the animals through CW prematurely, especially since Sna function is related to the endocycle where an oscillation pattern of expression is necessary. Nevertheless, all my data at least link the Sna function to the endocycle around the time of CW attainment.

Antibody staining revealed that Sna is also present in IPCs (Fig.2-28). In *Drosophila*, IPCs produce dILPs in the larval brain depending on nutrient status, which activates the single InR in target tissues and subsequently the IIS signaling pathway to coordinate tissue growth with changing nutrient conditions (198). I showed that *sna* expression itself in the PG requires IIS signaling. If *sna* also has a function in controlling the production of ILPs in the IPCs, then Sna definitely has a nutrient-sensing aspect in its function. It is an appealing idea to look into how this possible nutrient-sensing role of Sna relates to the CW checkpoint because the CW checkpoint is a physiological switch where larval growth is no longer coupled with environmental nutrient conditions after the switch.

2.4.5 Posttranscriptional regulation of Sna in the PG cells

My previous experiments found that Sna protein levels, but not transcript levels, were dependent on TOR signaling. Moreover, Sna protein became almost undetectable via immunostaining when RGs were dissected from larvae and left in PBS for 30 min, and this degradation could be reduced by adding α -ecdysone in the buffer. Moreover, a perinuclear localization in the PG was observed in some circumstances after incubating the brain-RG complex in 1x PBS for 30 min (Fig. 2-30). All these results suggested that Sna is a labile protein that is subject to posttranscriptional regulation to fine-tune its levels and subcellular localization for its function in the nucleus. Previous studies have revealed that mouse Snail1 has a nuclear export signal (NES) with a consensus motif of [LX(1-3)LX(2-3)LXL] (green shaded sequence in Fig. 2-45), C-terminal adjacent to the first Zinc finger domain. Near the NES is a serine-rich domain (SRD) and phosphorylation of SRD allows the export of Sna from the nucleus, thus reducing Sna function (238). I carried out a protein sequence alignment using ClustalW (239) and found that the NES is not present in *Drosophila* Sna (Fig. 2-45). Another study showed that both mouse and human Snail1 has two consensus motifs for GSK-3 β phosphorylation (in the serine-rich domain). Phosphorylation of the first motif targeted Sna for ubiquitination and degradation (240) since the first motif overlaps with the destruction motif with a basic sequence of DSGXXS. According to their research, the half-life of Sna protein is about 25 min, which is consistent with my observation. The phosphorylation of the second motif caused Sna to localize to the cytoplasm, thus attenuating its function in the nucleus as a transcription factor. So collectively, GSK-3 β generally represses Sna activity. According to the sequence alignment, *Drosophila* Sna does not have the DSGXXS destruction sequence and does not contain the two phosphorylation motifs.

When I carried out a phosphorylation site prediction analysis using GPS 3.0 (241-243), I identified three clusters of potential GSK-3 β phosphorylation sites on serine or tyrosine in *Drosophila* Sna, one of which is located around the area of the two GSK-3 β phosphorylation motifs in mouse and human Snail1 (Fig. 2-45). Most of the serine and tyrosine residues that are potentially phosphorylated by GSK-3 β are conserved across twelve *Drosophila* species, which are marked by the blue dots, suggesting these sites are potentially of functional importance (Fig. 2-45). Therefore, *Drosophila* Sna is potentially phosphorylated by GSK-3 β , however; it remains unclear how the phosphorylation events would affect protein stability and subcellular localization. Biochemical work needs to be carried out in the future to test this *in vitro*. For instance, expressing

mutated Sna (both the phosphorylation mimic form and the phosphorylation-disabled form) in S2 cells and then examining the consequences on Sna protein levels and its subcellular localization, would allow us to validate the functional importance of these potential phosphorylation sites. I have tried to express the constitutively active form of GSK-3 β (GSk3^{S9A}) in the PG to see how it would affect Sna levels. However, overexpressing GSk3^{S9A} resulted in early larval lethality, which made it difficult to examine Sna levels in the PG.

As mentioned before, the C-terminal sequence of *Drosophila* Sna is more divergent from mouse and human Sna (Fig. 2-1), while the N-terminus (where the zinc finger domains are) is more conserved. All regulatory functions including posttranscriptional regulation, protein-protein interaction/co-factor recruitment are usually limited to the C-terminus of Sna. Therefore, the conclusions drawn in mouse and human Sna regarding these functions may not be directly applicable to *Drosophila* Sna. It would be helpful if one could carry out immunoprecipitation coupled to mass spectrometry (IP-MS) to identify proteins that interact with *Drosophila* Sna. I could then infer which kind of posttranscriptional modification affect Sna and what proteins could act as co-factors mediating Sna's transcription function.

2.5 Conclusions and significance

My work is the first to demonstrate that Snail has a crucial role in *Drosophila* larval prothoracic gland linking endoreplication, critical weight checkpoint and ecdysone production. The discovery that the presence of *Drosophila* Snail in the nucleus was endocycle-regulated is entirely novel. *Snail* family genes were found to be highly expressed in several carcinomas and the expression of Snail in those cells positively correlates with tumor recurrence, metastasis and drug resistance (244-248). Therefore, further studies identifying the novel cell cycle-related direct targets of the Snail transcription factor activity will provide new perspectives on cancer research since the potential ability to manipulate cell cycle or cell survival via Snail function could open new avenues for controlling tumor cell growth and metastasis.

2.6 Figures

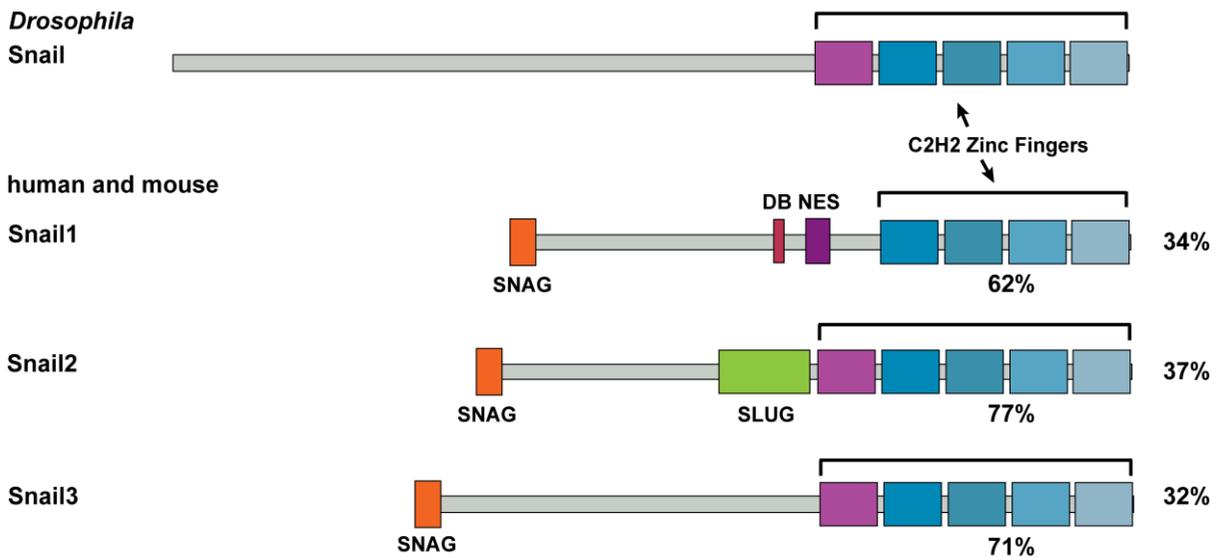


Figure 2-1. Schematic demonstration of the main functional domains in four Snail family proteins.

Main functional domains of the *Drosophila* Snail (Sna) as well as mammalian Snail1, Snail2 and Snail3 are represented by colored boxes. Mammalian Snail proteins contain an N-terminal SNAG domain, which is responsible for recruiting co-repressors. Snail2 has a SLUG domain of unknown function at the N-terminus, therefore, vertebrate Snail2 proteins are also known as Slug. Neither the SNAG nor the SLUG domain was identified in the *Drosophila* Sna. At the C-terminus, there are four to five zinc fingers in all four Snail proteins, which are responsible for DNA binding. DB in Snail1 refers to Destruction Box and NES stands for the nuclear export signal. The numbers underneath the zinc finger domains indicate the protein similarities of the zinc fingers region to that of *Drosophila* Sna, while the numbers to the right of each protein represent the percent similarity of the entire protein sequence to that of the *Drosophila* Sna protein.

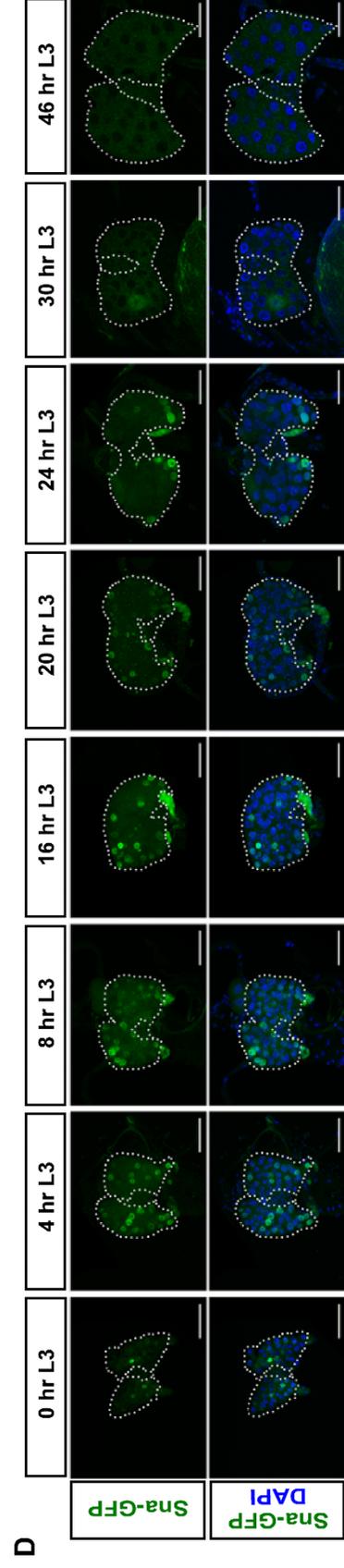
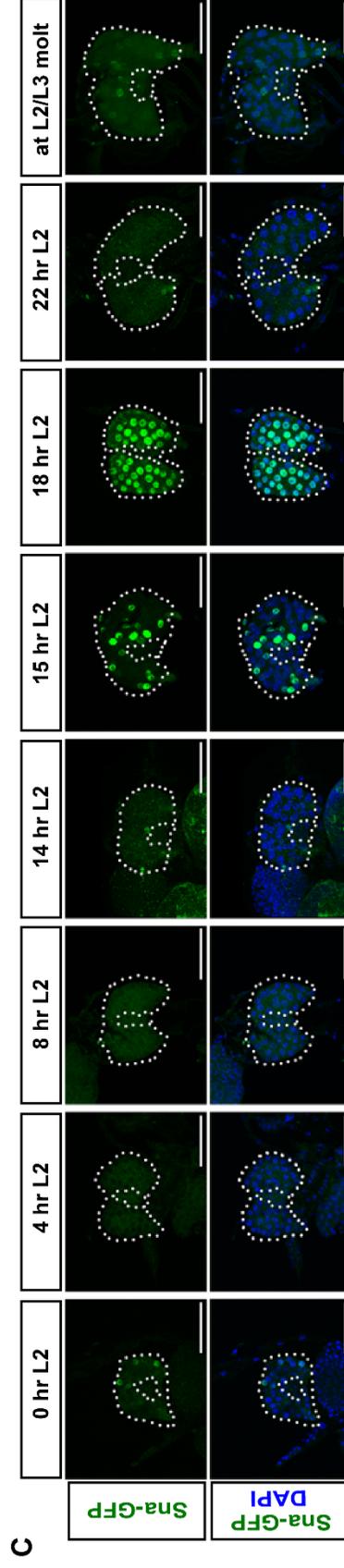
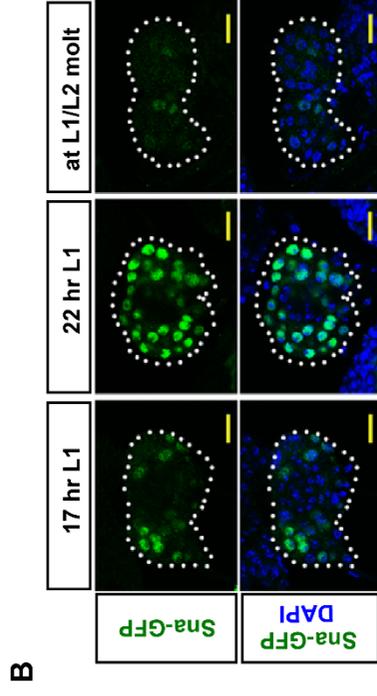
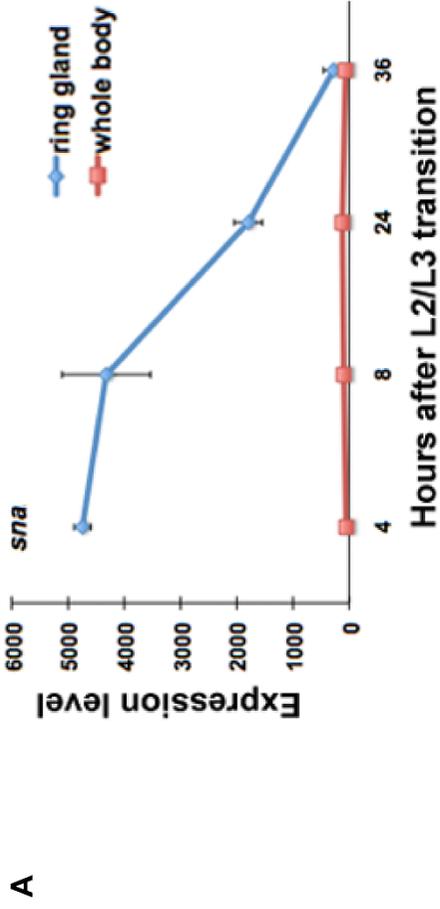


Figure 2-2. *sna* is specifically and dynamically expressed in the larval PG.

(A) *Sna* expression profile of L3 RGs based on microarray analysis. Microarray analyses were carried out using mRNA extracted from RGs of L3 at four different developmental time points (4, 8, 24 and 36 hr after the L2/L3 molt) compared to mRNA extracted from whole larvae at the same time points. The y-axis represents average array signals from three biological replicates (blue line: RG signal and red line: whole body signal). Error bars represent SD. (B-D). Immunofluorescent images of RGs dissected from a transgenic line carrying the *GFP*-tagged genomic *sna* sequence at various time points during larval stages. Upper panel shows the GFP signals after staining with an anti-GFP antibody, while the lower panel shows the merged image of GFP and DAPI staining. (B) RGs isolated from late L1; RG area is outlined with a white dotted line. Scale bar (yellow) represents 10 μm . (C) and (D) RGs from L2 (C) and L3 (D) at different time points. PG and CA are outlined with a white dotted line, where CA usually lies in the middle of the tissue. Scale bar (white) represents 50 μm .

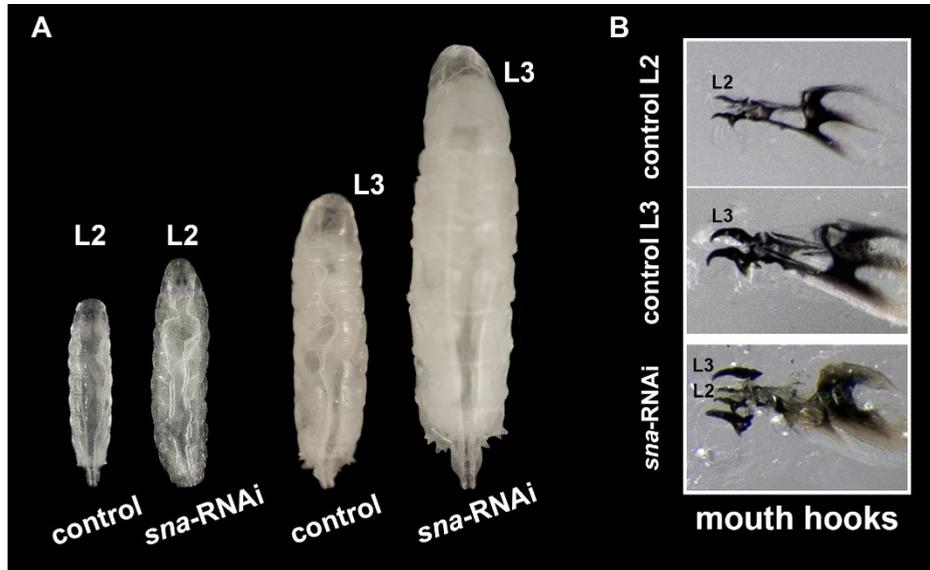


Figure 2-3. Disruption of *sna* function in the PG caused developmental arrest.

(A) PG>*sna*-RNAi larvae compared to control larvae. Control L2 were taken around 64 hr after egg laying (AEL) and the arrested RNAi L2 were taken around 86 hr AEL when the controls already molted to L3. Control L3 taken at 92 hr AEL and the arrested RNAi L3 taken four days later when the controls already pupariated. (B) The PG>*sna*-RNAi larvae have double mouth hooks. (A and B) control: *UAS-Dicer2*; *phm22-Gal4*>*w¹¹¹⁸*. *sna*-RNAi: *UAS-Dicer2*; *phm22-Gal4*>*UAS-sna*-RNAi (VDRC#50003).

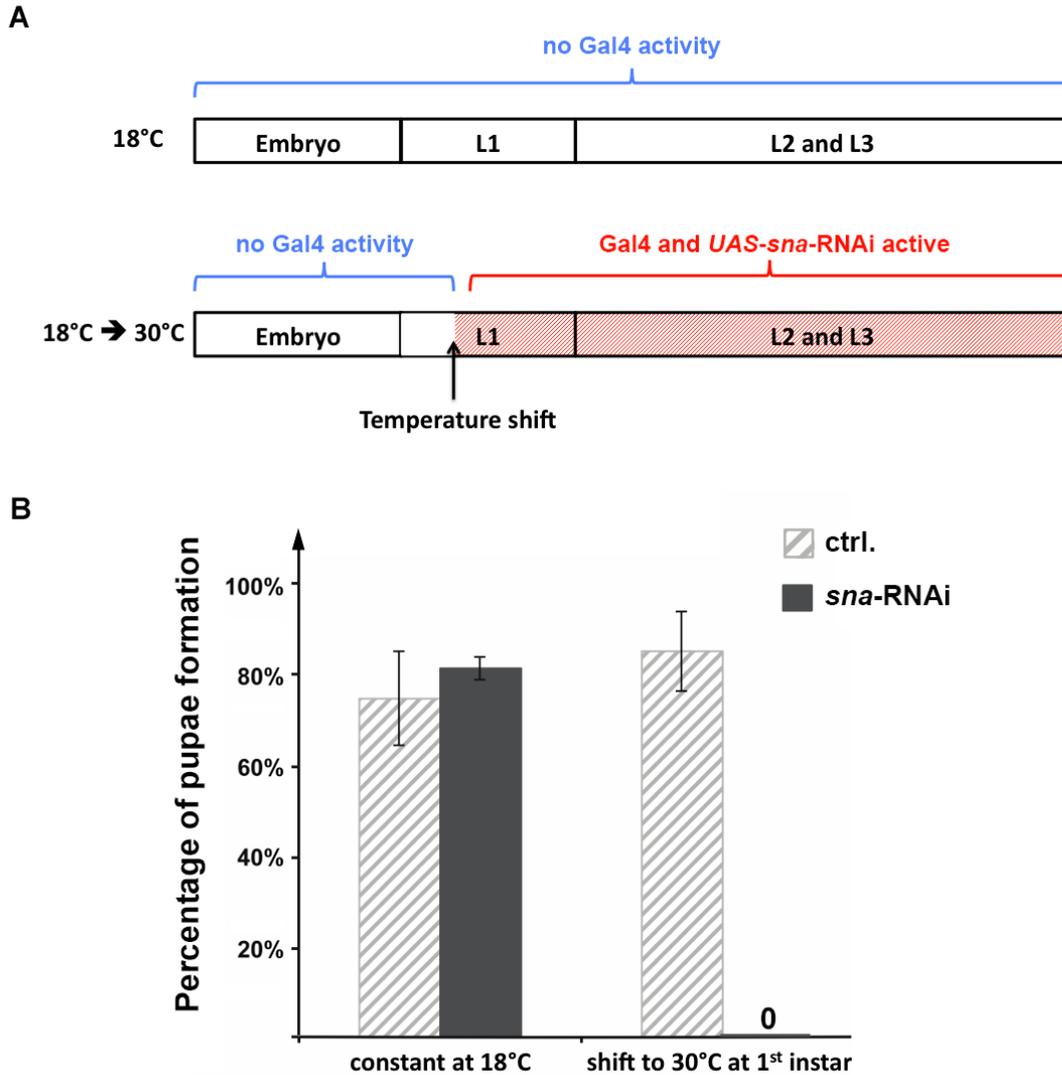


Figure 2-4. *sna* functions in the larval PG are essential for developmental progression.

(A) A schematic view of the Gal80 temperature shift procedure. Animals were either raised at 18°C throughout the entire life cycle or first kept at 18°C and then shifted to 30°C at 48 hr after egg laying (AEL) when all the embryos are already eclosed as first instar larvae. At 18°C, Gal80 represses Gal4 activity, while the repression of Gal4 will be released when the temperature exceeds 29°C, thus turning off the *UAS-sna*-RNAi. The procedure was applied to both the control and RNAi lines. (B) Percentage of embryos survived to the pupal stage under two different temperature schemes. ctrl.: *UAS-sna*-RNAi/+; *UAS-Dicer2*/+; *tub-Gal80^{TS}*/+. *sna*-RNAi: *UAS-sna*-RNAi/+; *UAS-Dicer2*/+; *phm22-Gal4/tub-Gal80^{TS}*. Error bars represent standard deviation.

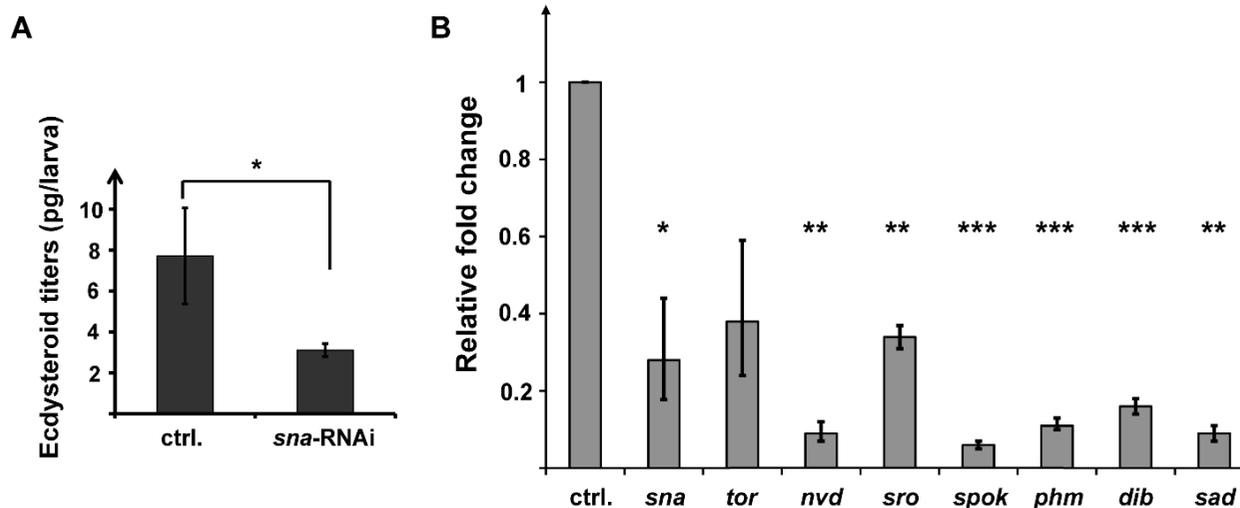


Figure 2-5. Loss-of-*sna* function in the PG via RNAi affected ecdysone biosynthesis.

(A) Ecdysteroid titers in whole-body larvae at the 28 hr L3 stage. (B) I measured the expression levels of six ecdysone biosynthetic genes, as well as the expression of *sna* itself by qPCR. Ring glands were dissected from the *sna*-RNAi and control animals at 24 hr after L2 to L3 molt. Three replicates were included for each condition. I normalized the expression of each gene in the *sna*-RNAi samples to the expression in the control respectively; the expression in the control was shown as 1. * $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$ (based on Student's t-test). (A and B) ctrl.: *UAS-Dicer2; phm22-Gal4 > w¹¹¹⁸*. *sna*-RNAi: *UAS-Dicer2; phm22-Gal4 > UAS-sna*-RNAi.

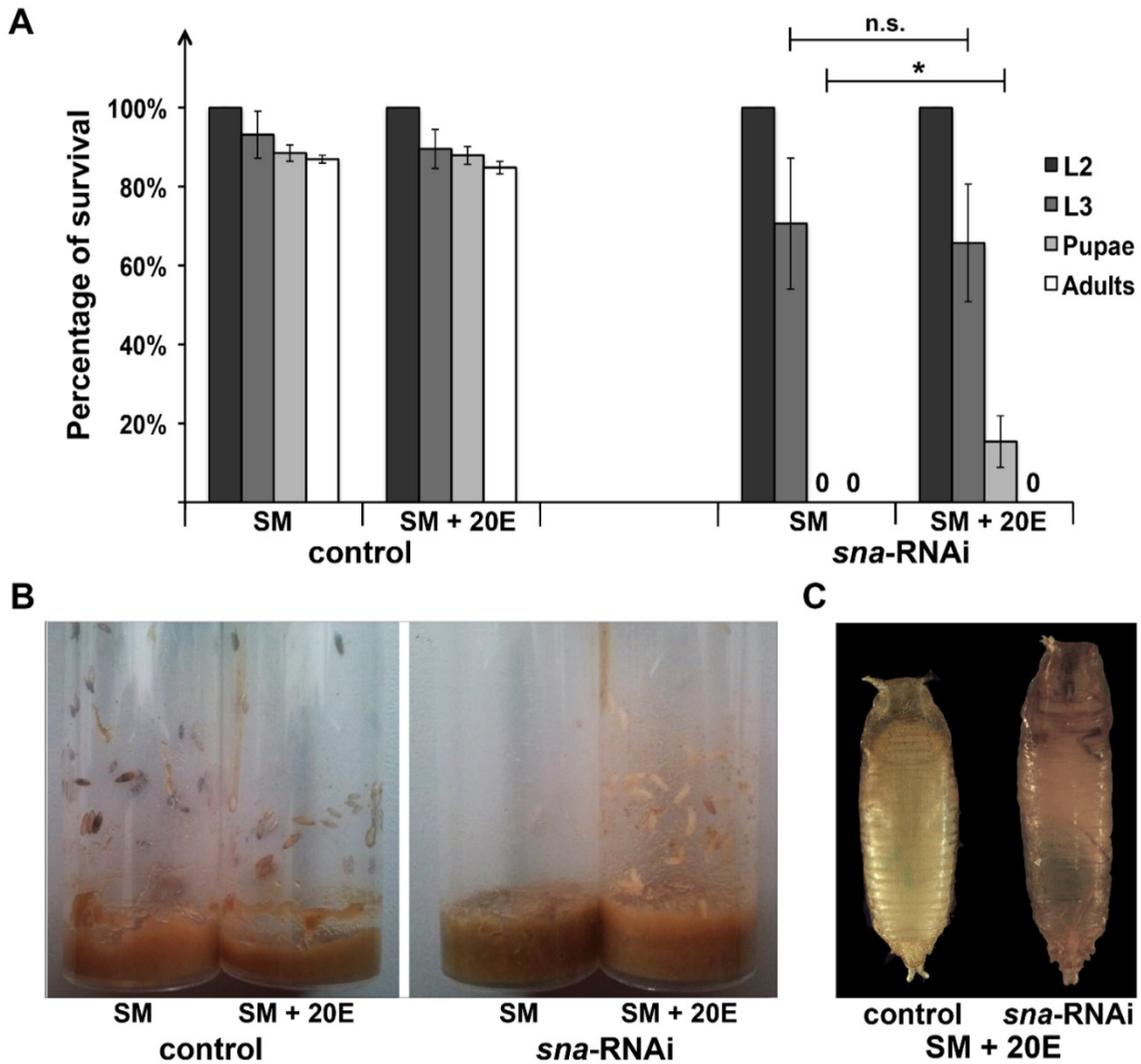


Figure 2-6. Developmental arrest caused by *PG>sna-RNAi* was partially rescued by ecdysone feeding.

(A) Percent of larvae that survived to indicated stages. I used 50 L2 of each genotype as a starting population. Error bars represent standard deviation. * $p < 0.05$; n.s.: not significant (based on Student's *t*-test). (B) Examples of vials taken from the experiments in (A). (C) An example of a pupa that resulted from *UAS-Dicer2; phm22-Gal4>UAS-sna-RNAi* animals when fed with 20E (right) compared to control pupae fed with 20E. (A-C) Control: *UAS-Dicer2; phm22-Gal4>w¹¹¹⁸*. *sna-RNAi*: *UAS-Dicer2; phm22-Gal4>UAS-sna-RNAi*. SM: agar-cornmeal-based standard medium; SM + 20E: standard medium supplied with 20-Hydroxyecdysone.

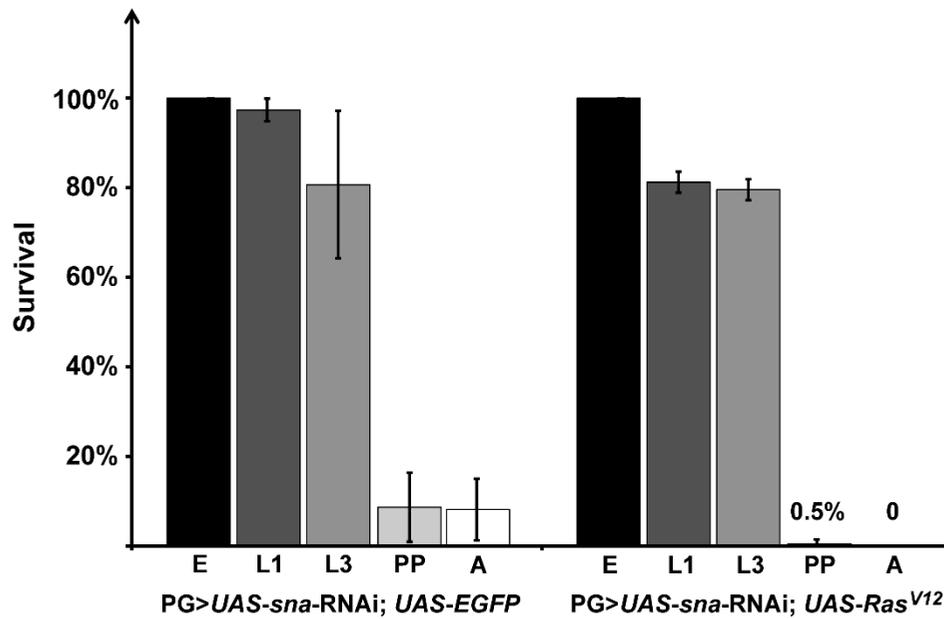


Figure 2-7. *sna* does not function through the PTTH/MAPK signaling pathway in the PG.

The bar graph shows percent survival at each indicated developmental stage in the PG>*sna*-RNAi alone and PG>*sna*-RNAi + *Ras*^{V12} animals. The actual genotype of the PG>*UAS-sna*-RNAi; *UAS-EGFP* animal was: *UAS-sna* RNAi/+; *UAS-EGFP/UAS-Dicer2*; *phm22-Gal4*/+. The full genotype of the PG>*UAS-sna*-RNAi; *UAS-Ras*^{V12} animal was: *UAS-sna*-RNAi/+; *UAS-Dicer2*/+; *phm22-Gal4/UAS-Ras*^{V12}. E: embryos; L1: first instar larvae; L3: third instar larvae; PP: pupae and A: adults.

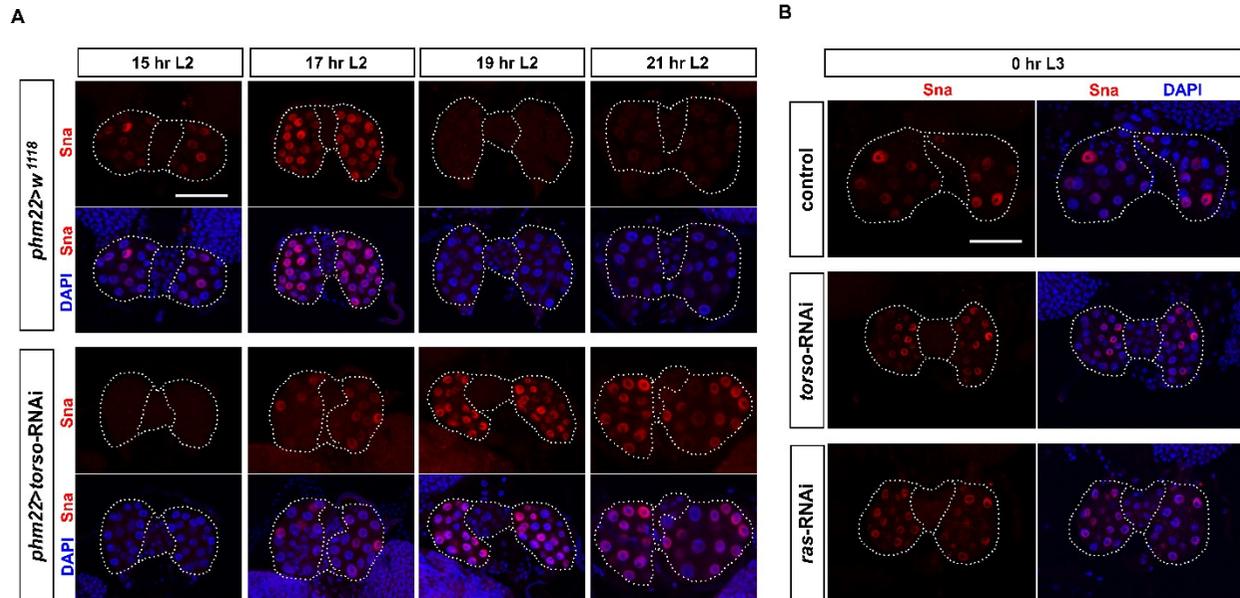


Figure 2-8. *Sna* levels in the PG are not controlled by PTTH/MAPK signaling.

(A) Single-plane confocal images of RGs dissected at different time points around 17 hr after the L1 to L2 molt. Tissues were stained with anti-*Sna* antibody and DAPI. The PG and CA are outlined by white dotted lines. (B) Single-plane confocal images showing RGs dissected at the L2 to L3 molt. Tissues were stained with anti-*Sna* antibody and DAPI. Control: *phm22-Gal4>w¹¹¹⁸*. *torso-RNAi*: *phm22-Gal4>UAS-torso-RNAi*. *ras-RNAi*: *phm22-Gal4>UAS-ras85d-RNAi*. (A-B) Scale bars: 50 μ m applicable to all samples.

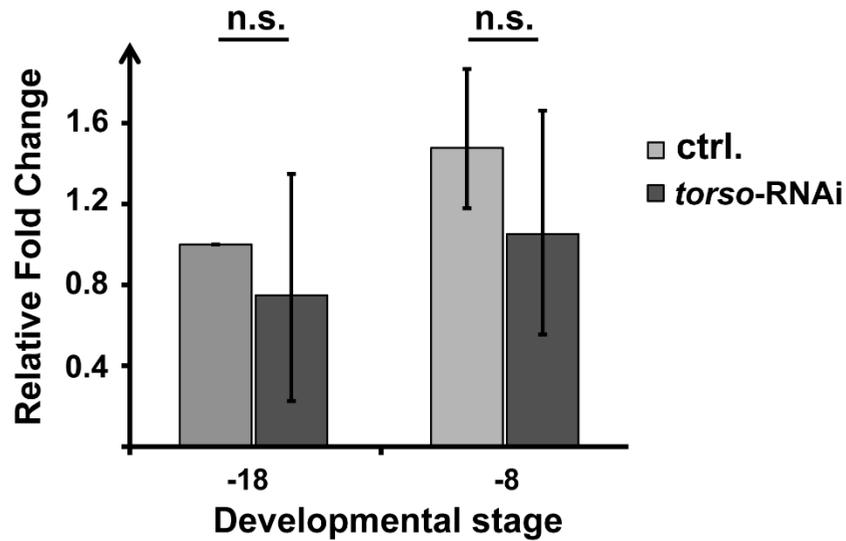


Figure 2-9. *sna* mRNA levels are not dependent on Torso (the receptor for PTTH).

I carried out qPCR analysis to examine transcript levels of *sna* in *phm22>torso*-RNAi larvae. For this, the brain-RG complexes were collected at two different developmental stages. Animals were staged according to the blue gut method: -18 represents 18 hrs before puparium formation (BPF), *aka* blue gut wandering larvae and -8 represents 8 hrs before puparium formation, *aka* partial blue gut wandering larvae. Fold changes are relative to the *phm22>w¹¹¹⁸* control at 18 BPF. The error bars are 95% confidence intervals. n.s.: not significant.

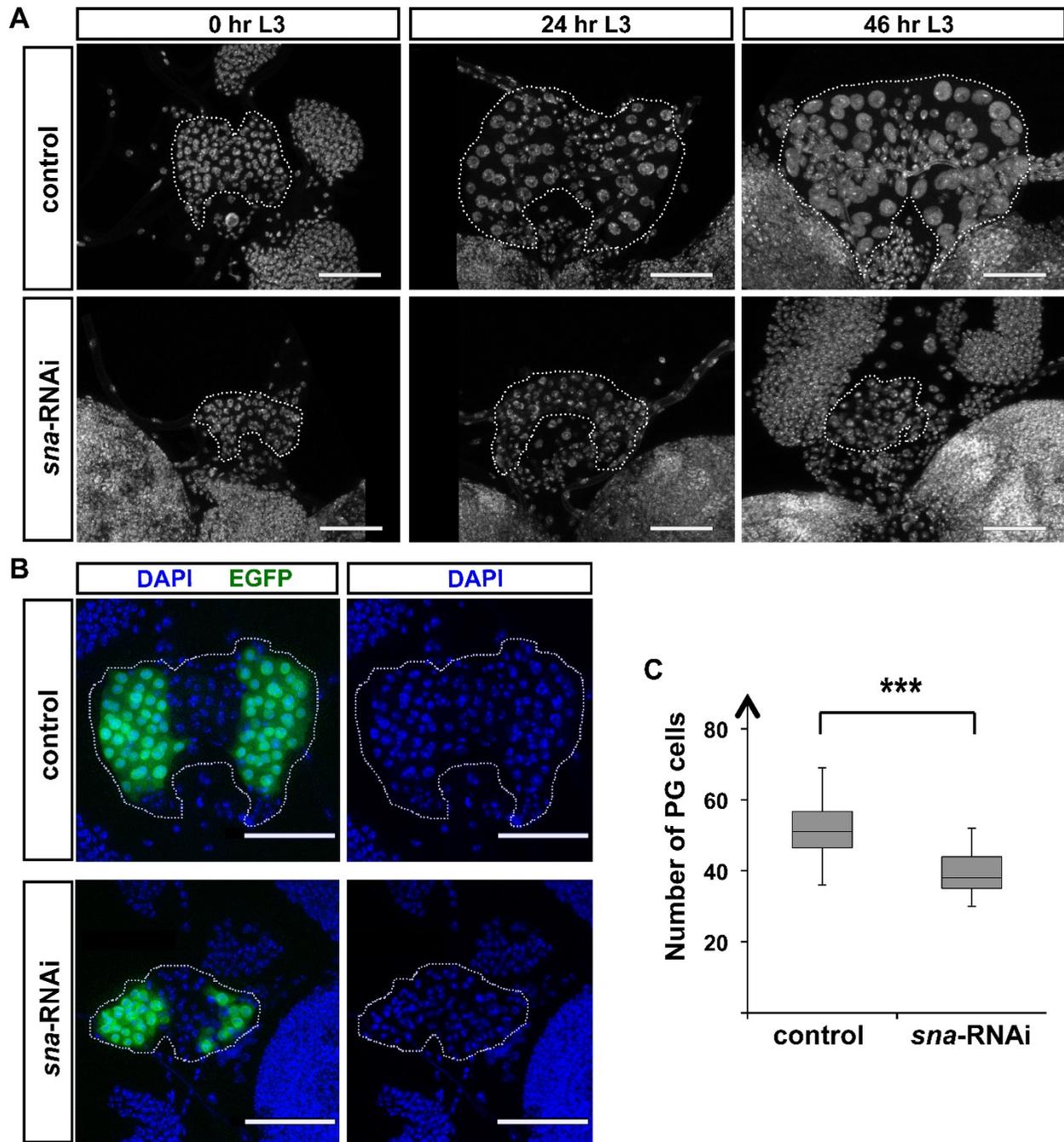


Figure 2-10. PG-specific *sna*-RNAi affected the size of the PG nuclei as well as the PG cell number.

(A) Maximal projection of Z-stack confocal images for RGS dissected at different times relative to the L2/L3 molt. Tissues were stained with DAPI to show nuclei. RGS are outlined by white dotted lines. *sna*-RNAi: *UAS-Dicer2*; *phm22-Gal4*>*UAS-sna*-RNAi and control: *UAS-Dicer2*; *phm22-Gal4*>*w¹¹⁸*. (B) Maximal projection of the Z-stack confocal images for RGS dissected at the L2/L3 molt. *UAS-EGFP* expression was driven by *phm22-Gal4* to label PG cells. Tissues were stained with DAPI. (A-B) Scale bars: 50 μ m. (C) Whisker box plot of PG cell numbers in control and *sna*-RNAi animals. 10–15 ring glands were examined for each condition. *** p <0.0001 (based on Student's t-test). (B and C) *sna*-RNAi: *UAS-Dicer2*; *phm22-Gal4*>*UAS-sna*-RNAi; *UAS-EGFP*. Control: *UAS-Dicer2*; *phm22-Gal4*>*UAS-EGFP*.

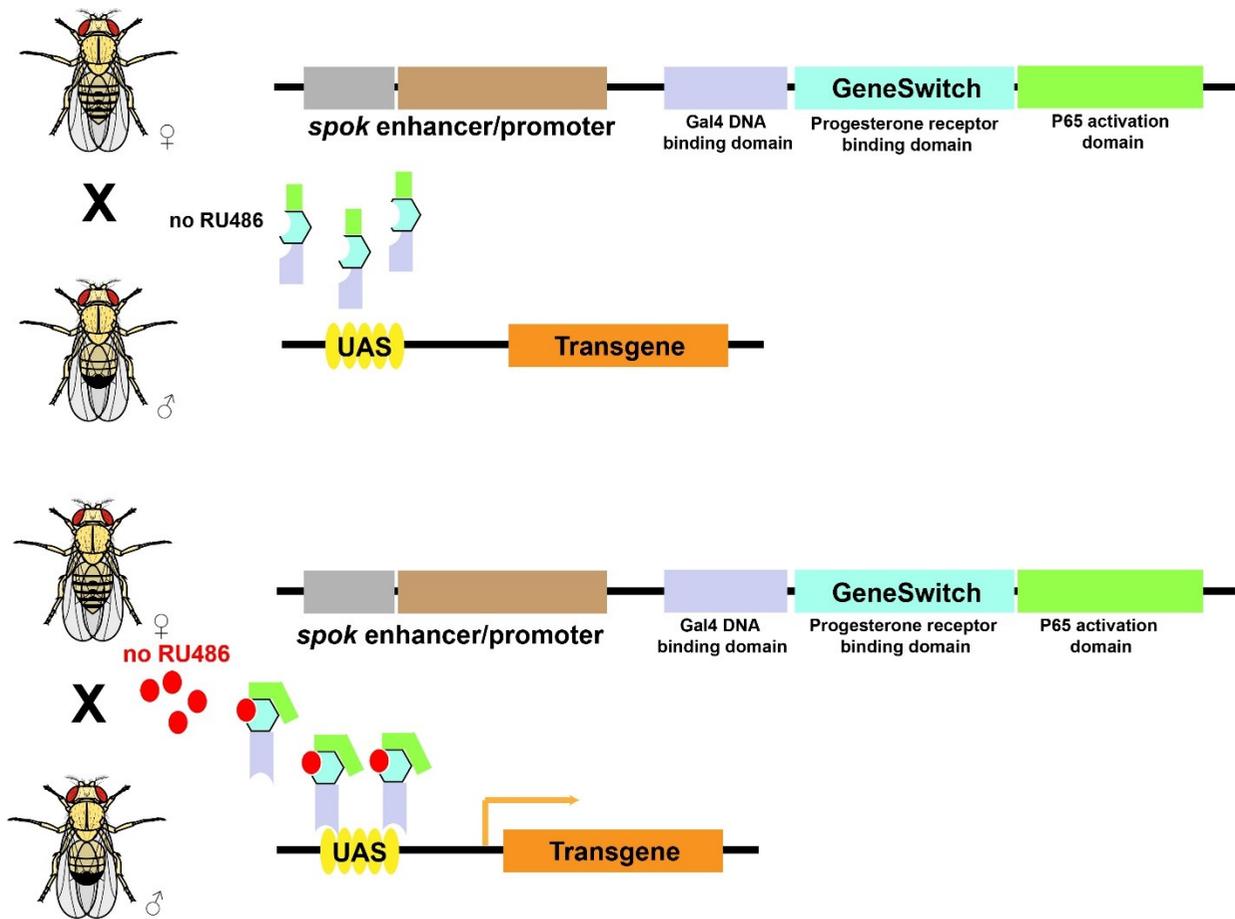


Figure 2-11. The GeneSwitch/UAS expression system in *Drosophila*.

Driver lines expressing the transcriptional activator GeneSwitch-Gal4 in a tissue-specific manner could be crossed to UAS lines with transgene fused to five Gal4-binding sites arrayed in tandem (5× UAS). The GeneSwitch-Gal4 is based on a Gal4-progesterone receptor fusion protein, which contains the Gal4 DNA-binding domain, the p65 activation domain and the ligand-binding domain of the progesterone receptor. This fusion protein functions as a ligand-stimulated transcription factor whose expression is under the control of a tissue-specific enhancer, in this case, the *spookier* (*spok*) enhancer, which is specific for PG-specific expression. In the absence of the ligand mifepristone (RU486), the GeneSwitch-Gal4 protein is expressed in target tissues but remains inactive. In the presence of ligand, the GeneSwitch-Gal4 molecule adopts an active state to turn on the transgene downstream of the UAS sequence.

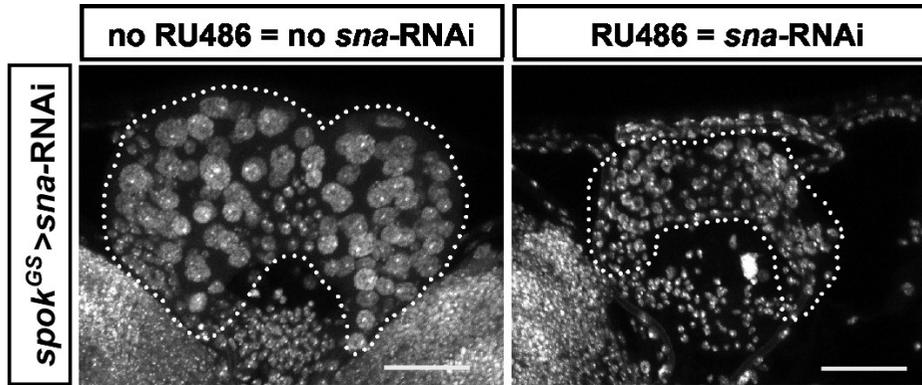


Figure 2-12. *sna*-RNAi in the larval PG resulted in small PG nuclei.

(A) Maximal projection of the Z-stack confocal images for ring glands dissected at the end of larval development before pupae formation. *spok*^{GS}>*sna*-RNAi: *spok-Gal4*^{GS}>*UAS-sna*-RNAi; *UAS-Dicer2*. Animals were either reared on Nutrifly food with 2% ethanol (where no *sna*-RNAi was expressed in the PG) or on Nutrifly food containing 8 µg/ml RU486 with 2% ethanol (where *sna*-RNAi was expressed after hatching, as soon as first instar larvae started feeding). Tissues were stained with DAPI to show the nuclei. Ring glands are outlined by white dotted lines. Scale bars: 50 µm.

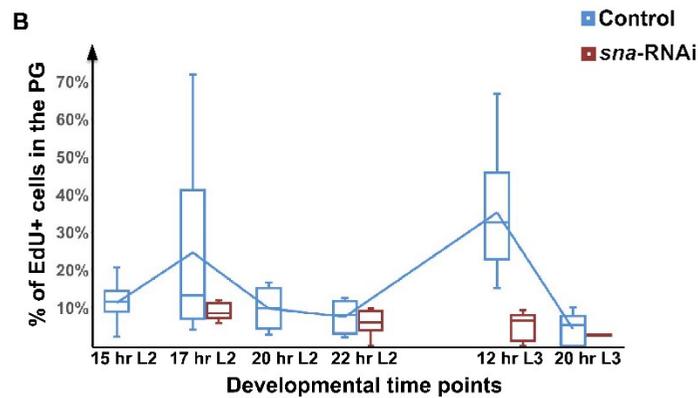
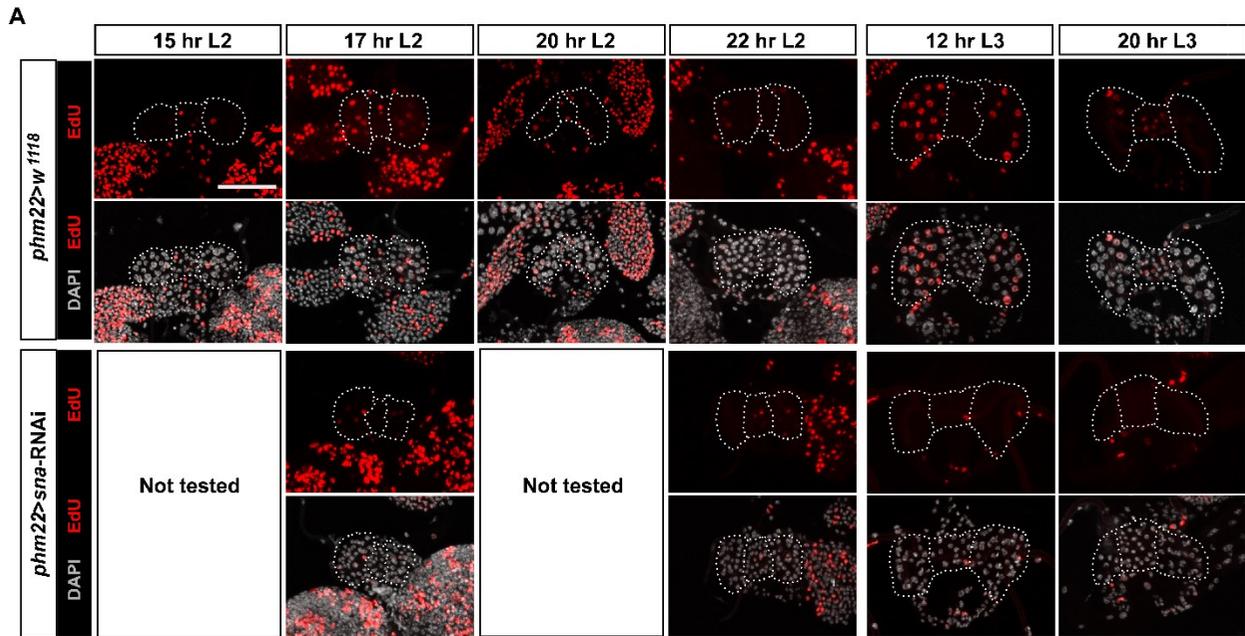


Figure 2-13. Endocycle progression in the PG requires *sna* function.

(A) The PGs of the *sna*-RNAi animals showed a reduction in EdU incorporation at the indicated developmental stages. The PGs are outlined by white dashed lines. Scale bars: 50 μ m. I examined 8-10 ring glands for each condition. (B) Box plots showing the percentages of EdU-positive PG cells in controls and *sna*-RNAi larvae. The average values of controls at each time point are connected by the blue line. (A-B) Control: *UAS-Dicer2*; *phm22-Gal4>w¹¹¹⁸*. *sna*-RNAi: *UAS-Dicer2*; *phm22-Gal4>UAS-sna*-RNAi.

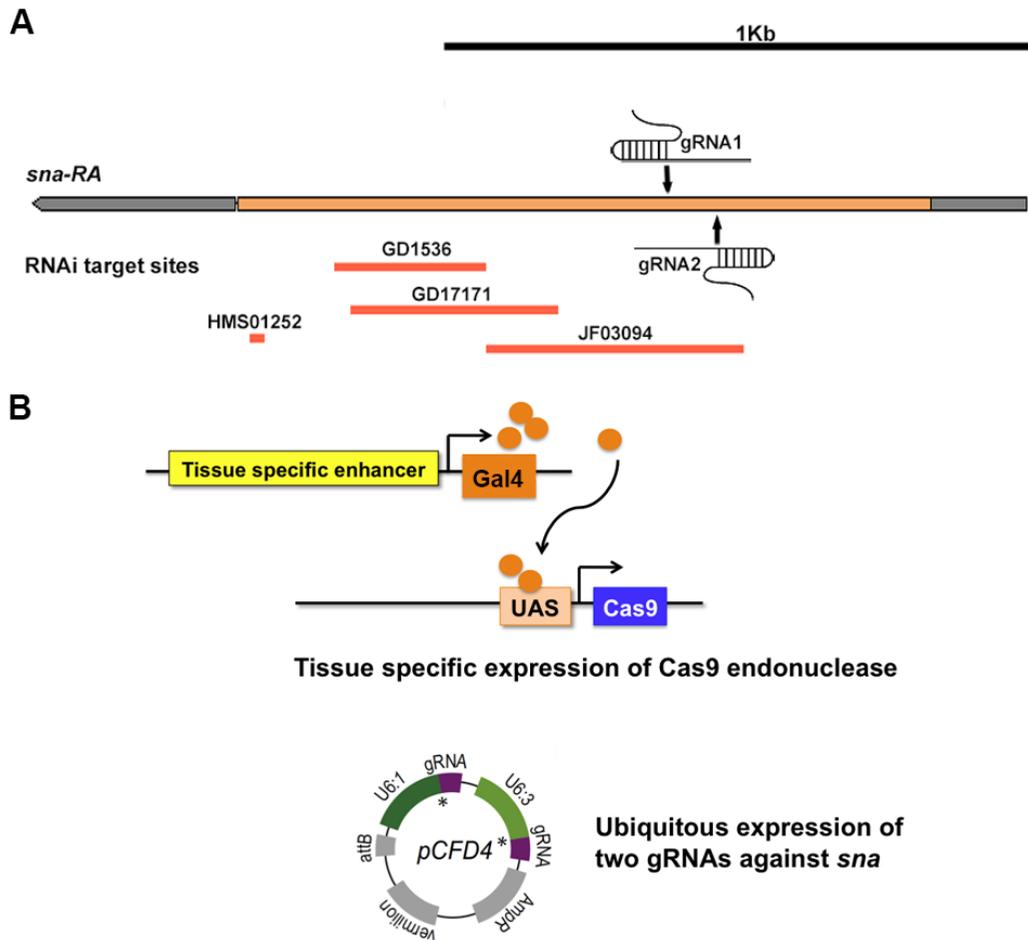


Figure 2-14. The *Drosophila sna* locus organization, target sites of all the existing *sna*-RNAi lines as well as gRNA target sites for CRISPR-based tissue-specific *sna* deletion.

(A) The *Drosophila sna* gene has no intron and encodes only one known transcript: *sna*-RA. Boxes represent non-coding (gray) and coding (orange) exons. There are three transgenic dsRNA fly lines from VDRC, but they represent only two individual dsRNA snapback constructs (GD1536 and GD17171). Transgenic RNAi Resource Project (TRiP) at Harvard medical school built two RNAi lines with two independent constructs: HMS01252 and JF03094. Among these constructs, only GD17171 (transformants #50003) gave phenotypes when used to knock down *sna*. Two gRNAs which are ~70 bp away from each other were designed to target *sna* in the CRISPR-based tissue-specific *sna* deletion. (B) A schematic overview of CRISPR-based tissue-specific *sna* deletion. I generated the transgenic stocks carrying pCFD4-U6:1_U6:3tandemgRNAs, which ubiquitously express two gRNA targeting *sna*. *UAS-Cas9* (Bloomington#54595) was stably combined to the gRNA transgenic flies (tandem *sna*-gRNAs; *UAS-Cas9*). The flies were then crossed with the PG-specific-*Gal4* flies to drive *Cas9* expression only in the PG. The two gRNAs will direct the *Cas9* enzyme to the *sna* locus and generate two double stranded breaks close to each other; upon which the cell will attempt to repair the double strand break through NHEJ, which should result in the accumulation of small deletions in the *sna* gene (the PG is a polyploid tissue and thus contains multiple *sna* loci).

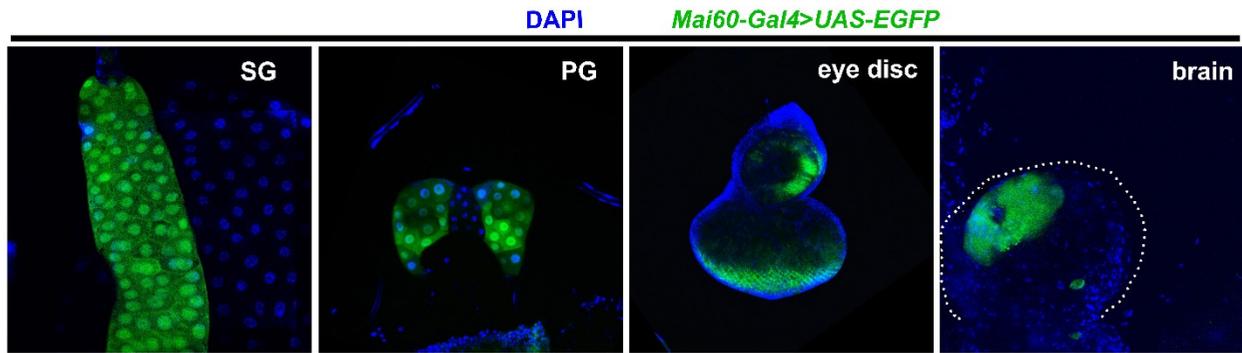


Figure 2-15. Larval tissues where GFP is expressed via the *Mai60-Gal4* driver.

Immunofluorescent images of tissues dissected from *Mai60-Gal4>UAS-EGFP* wandering larvae. Tissues were stained with DAPI (blue). SG: salivary gland. PG: prothoracic gland. The white dotted line marks the boundary of one brain hemisphere.

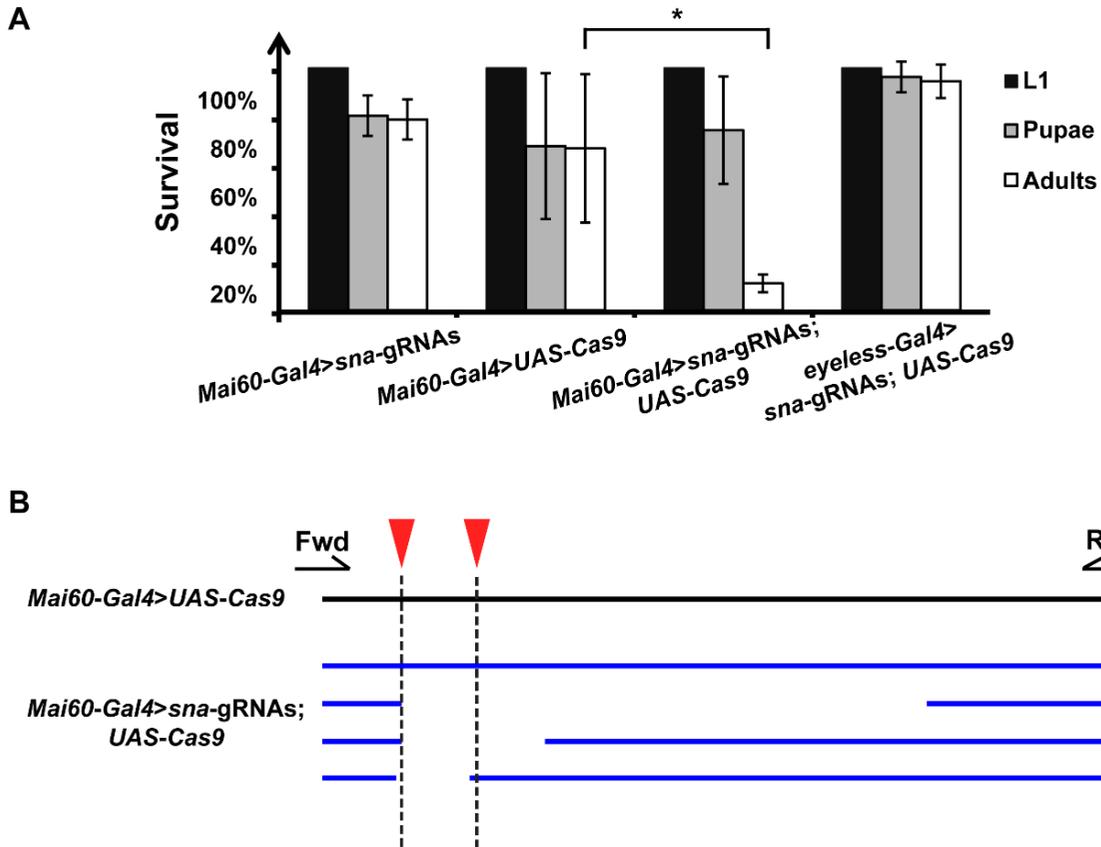


Figure 2-16. Developmental phenotypes caused by CRISPR-based PG-specific *sna* mutations.

(A) Percent of larvae that survived to indicated stages. A defined number of L1 were used for each genotype as a starting population. Error bars represent standard deviation. * $p < 0.05$ (based on Student's t-test). (B) A schematic diagram showing the extent of deletions in the *Mai60-Gal4>sna-gRNAs; UAS-Cas9* larvae determined by sequencing flanking the predicted breakpoints. *Mai60-Gal4>UAS-Cas9* animals served as controls. The genomic DNA was extracted from ring glands in wandering L3 larvae. The red triangle indicated the cleavage sites and primer locations are indicated by black arrows. Fwd: forward primer; Rev: reverse primer for genomic PCR.

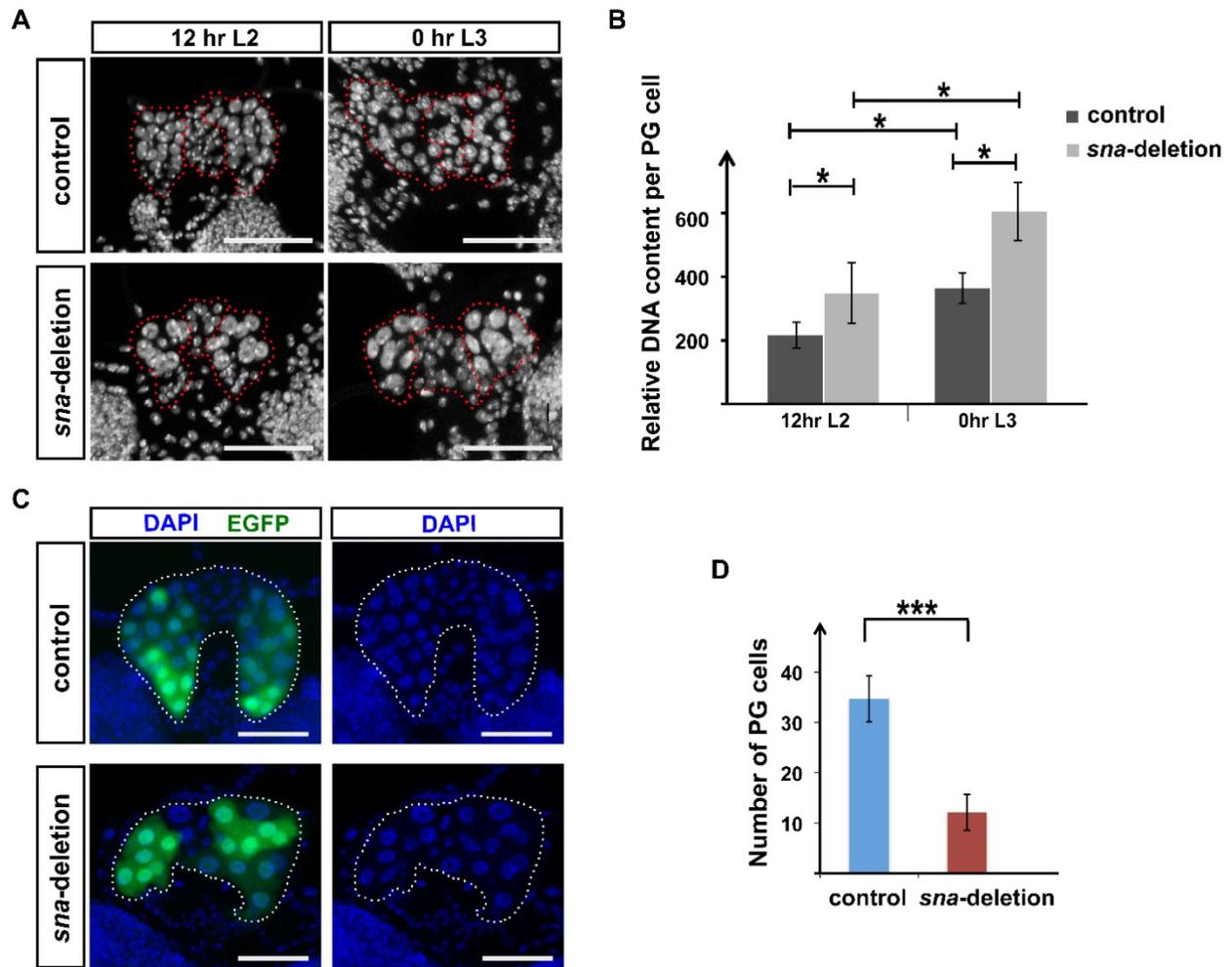


Figure 2-17. PG specific deletion of *sna* affected the morphology of the ring glands.

(A) Maximal projection of Z-stack confocal images for ring glands dissected at 12 hr after the L1/L2 molt and 0 hr after the L2 to L3 molt. I stained tissues with DAPI to show the nuclei. PG and CA are outlined by the red-dotted line. (B) Summation of DAPI intensity per PG cell normalized to average DAPI intensity in the brain hemisphere. (C) Maximal projection of the Z-stack confocal images for ring glands dissected at 24 hr after the L2 to L3 molt. *UAS-EGFP* expression was driven by *Mai60-Gal4* to indicate the cells that expressed *UAS-Cas9*. I stained tissues with DAPI and outlined the PG and CA with a white-dotted line. (D) The number of PG cells in controls and *sna*-deletion animals. (A-D) 5–10 ring glands were examined per condition. (A, B) control: *Mai60-Gal4>UAS-Cas9*; *sna*-deletion: *Mai60-Gal4>sna*-gRNAs; *UAS-Cas9*. (C and D) control: *UAS-EGFP*; *Mai60-Gal4>UAS-Cas9* and *sna*-deletion: *UAS-EGFP*; *Mai60-Gal4>sna*-gRNAs; *UAS-Cas9*. (A and C) Scale bars: 50 μ m. (C and D) Error bars represent standard deviation. * $p < 0.05$; *** $p < 0.0001$ (based on Student's t-test).

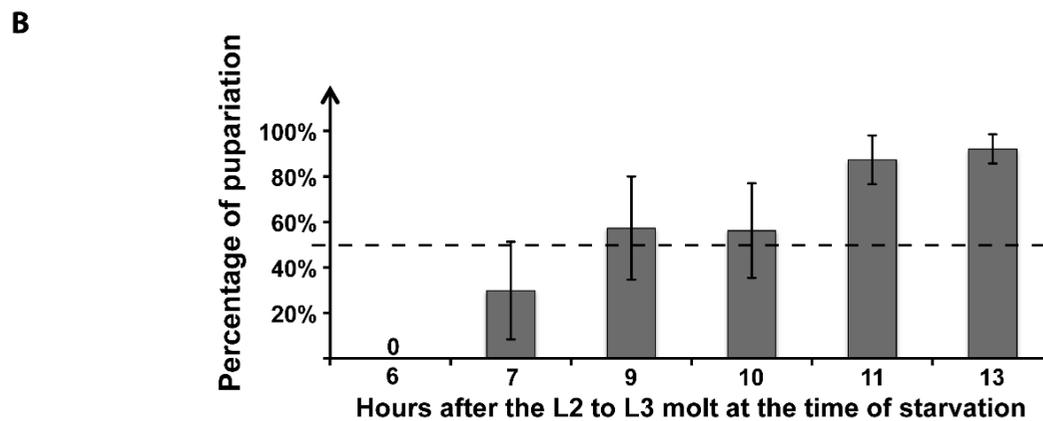
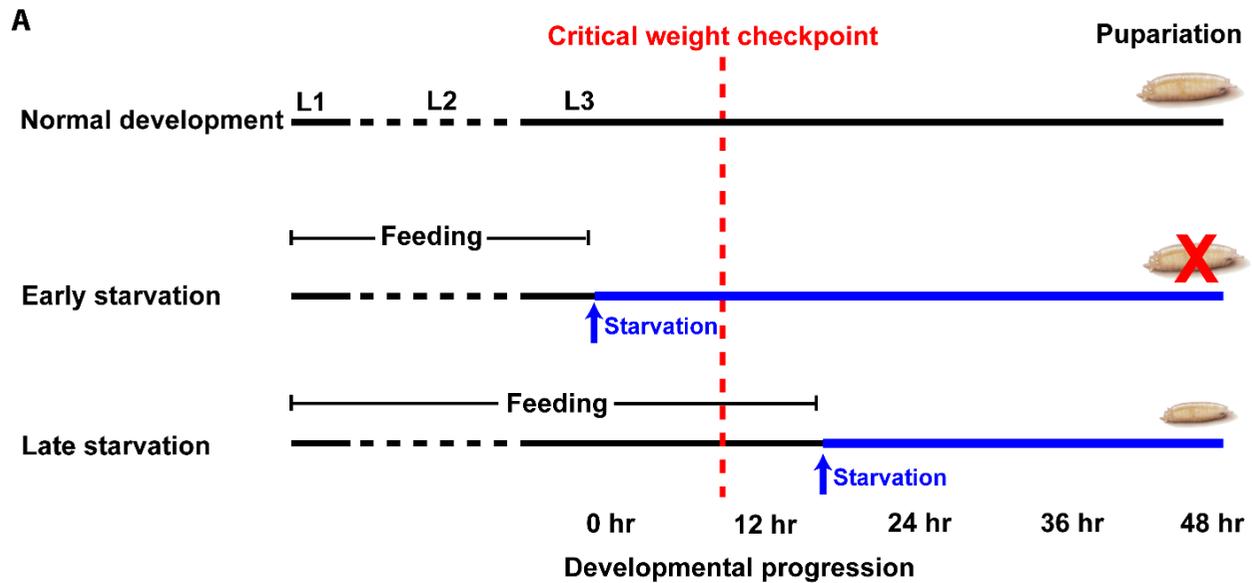


Figure 2-18. Determining the time of critical weight attainment.

(A) A schematic illustration of how to determine whether larvae have attained critical weight. The critical weight for metamorphosis was determined by starving L3 larvae of known developmental time classes. If starvation occurs before the larvae attained critical weight, development will stop and larvae do not form pupae, while when starvation occurs after critical weight checkpoint, larvae could pupariate. (B) Percentage of larvae pupariated after starvation at serial developmental stages during the L3 stage. The genotype tested was *UAS-Dicer2; phm22-Gal4>UAS-EGFP*. The dotted line marks the cut off for 50% pupariation. Error bar represents standard deviation.

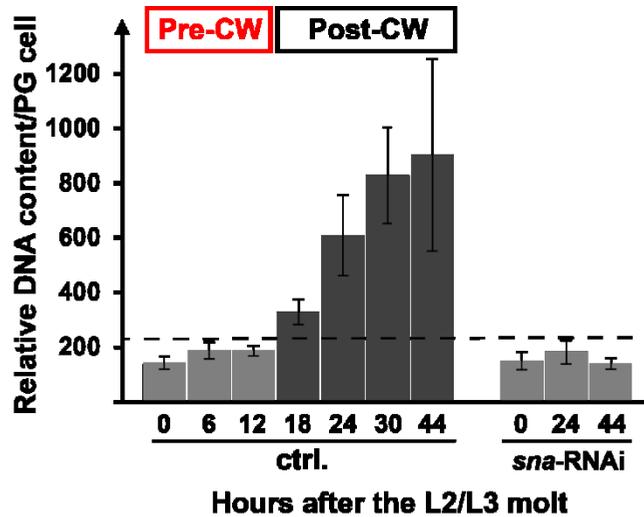


Figure 2-19. Endocycle progression was arrested at the time of critical weight checkpoint in PG>*sna*-RNAi PGs.

Relative DNA content per PG cell at different time points after the second instar to third instar molt was determined by the summation of DAPI intensity in the PG normalized to the mean DAPI intensity of the brain hemisphere. ctrl.: *UAS-Dicer2*; *phm22-Gal4>UAS-EGFP* and *sna*-RNAi: *UAS-Dicer2*; *phm22-Gal4>UAS-sna*-RNAi; *UAS-EGFP*. The dotted line marks the threshold of DNA content per PG cell before the critical weight checkpoint (gray bars are all below the threshold). After 12 hr L3 (post critical weight) the DNA content in controls increased beyond the threshold (black bars). pre-CW: developmental time points before critical weight attainment in controls. post-CW: developmental time points after critical weight attainment in controls. Error bar represents standard deviation.

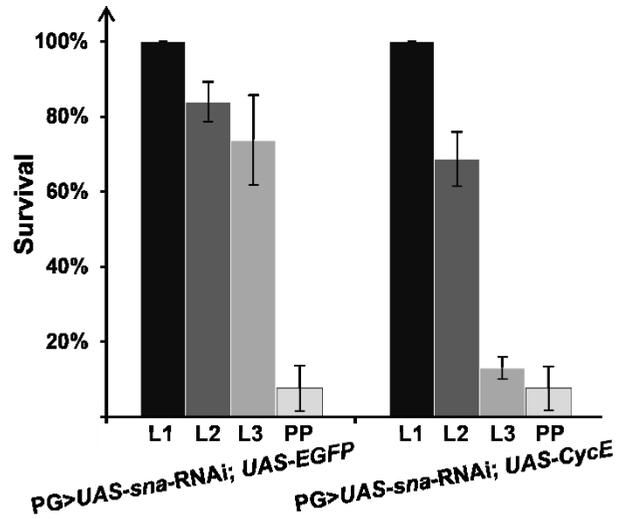


Figure 2-20. PG-specific expression of *CycE* can not rescue the developmental arrest caused by PG>*sna*-RNAi.

The bar graph shows percent survival at each indicated developmental stage in the PG>*sna*-RNAi alone and PG>*sna*-RNAi + *CycE* animals. The full genotype of the PG>*UAS-sna*-RNAi; *UAS-EGFP* animal was: *UAS-sna* RNAi/+; *UAS-EGFP/UAS-Dicer2*; *phm22-Gal4*/. The full genotype of the PG>*UAS-sna*-RNAi; *UAS-CycE* animal was: *UAS-sna*-RNAi/+; *UAS-Dicer2/UAS-CycE*-cDNA; *phm22-Gal4*/. L1: first instar larvae; L2: second instar larvae; L3: third instar larvae and PP: pupae.

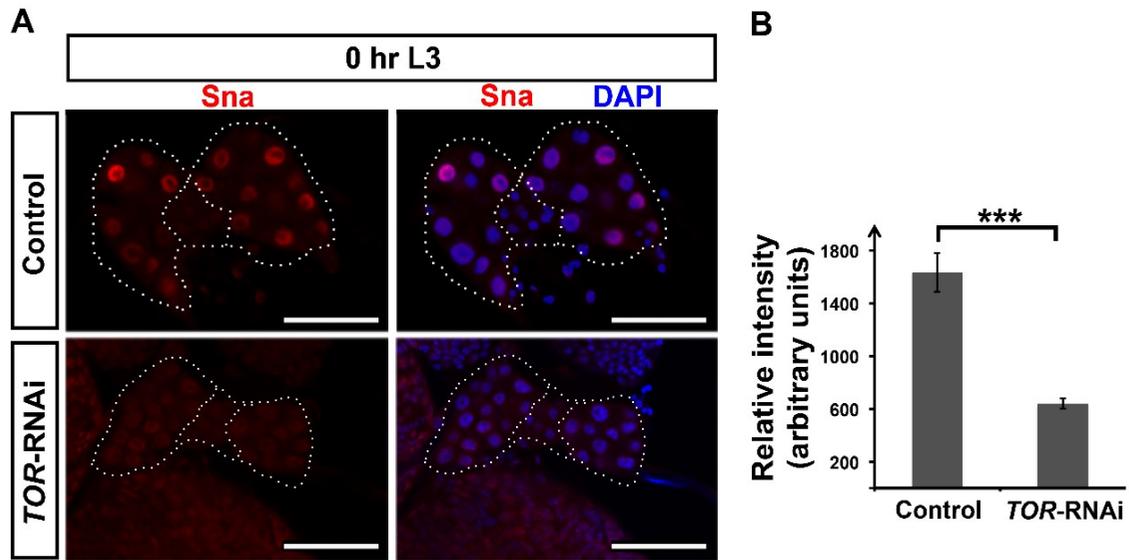


Figure 2-21. The effect of *TOR*-RNAi on *Sna* protein levels in the PG.

Expression of *TOR*-RNAi in the PG affected *Sna* levels. Tissues were stained with anti-*Sna* antibody as well as DAPI to label nuclei. The figures show single-plane confocal images. 8-12 ring glands were examined for each condition. The white dotted line marks the boundary of PG and CA in the ring glands. Scale bar: 50 μ m. (B) Average *Sna* fluorescent intensity per nucleus. Error bars represent standard deviation. *** p <0.0001; (Student's t-test). (A-B) control: *phm22>w¹¹¹⁸*. *TOR*-RNAi: *phm22>TOR*-RNAi.

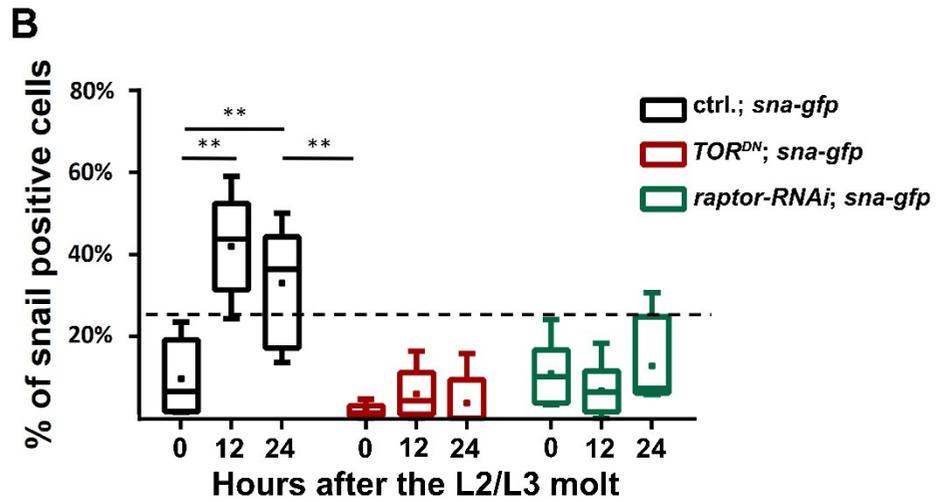
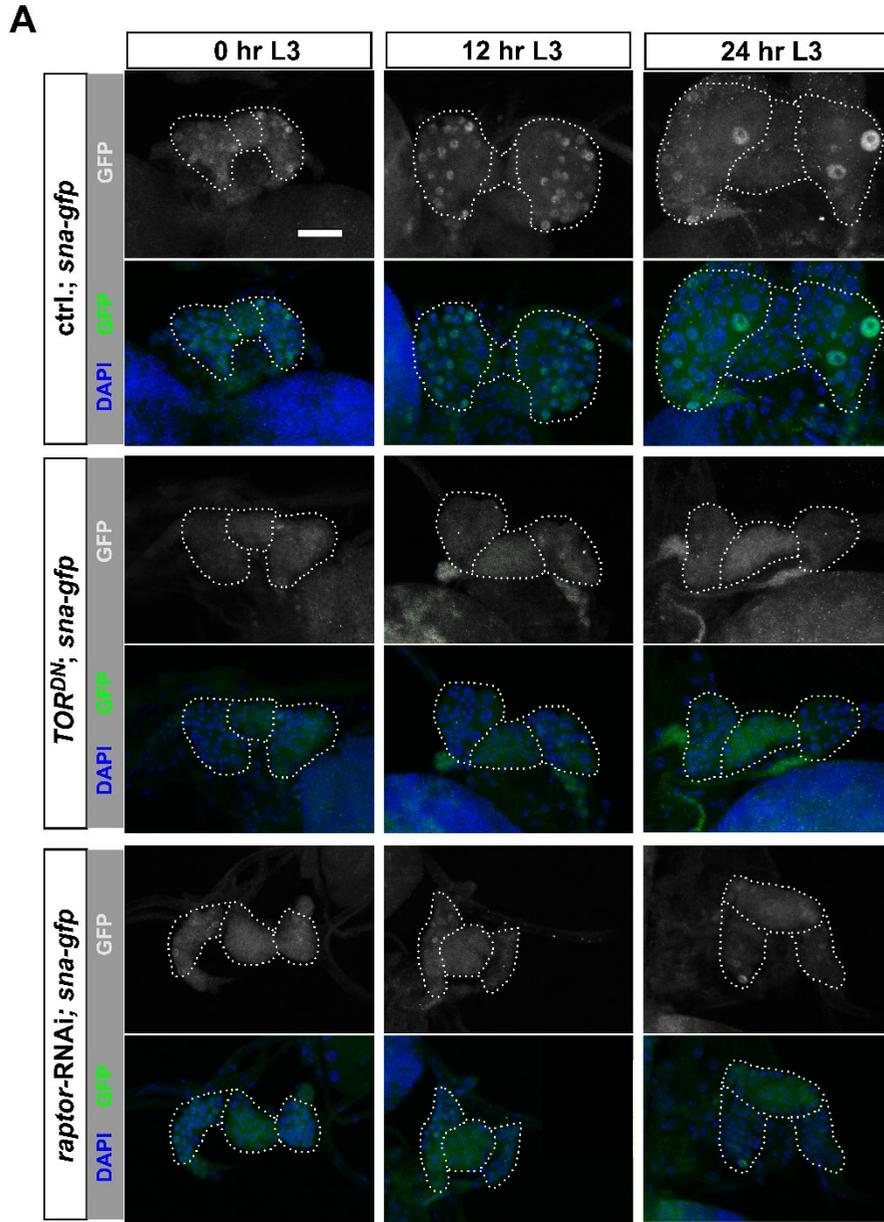


Figure 2-22. Sna levels in the PG are regulated by TORC1 complex.

(A) Immunofluorescence of GFP-tagged Sna in ctrl. (*phm22-Gal4/+; sna-gfp*), *TOR^{DN}* (*phm22-Gal4/UAS-TOR^{DN}; sna-gfp*) and *raptor-RNAi* (*phm22-Gal4/UAS-raptor-RNAi; sna-gfp*) ring glands. *sna-gfp* represents the transgenic fly carrying the *gfp*-tagged *sna* genomic clone. PG and CA are outlined by a white dotted line. Scale bars: 50 μ m applicable to all the samples showed in panel A. 8-12 ring glands were examined for each condition. (B) Box plot of Sna⁺ nuclei in the PG quantified from the ring glands in panel (A).

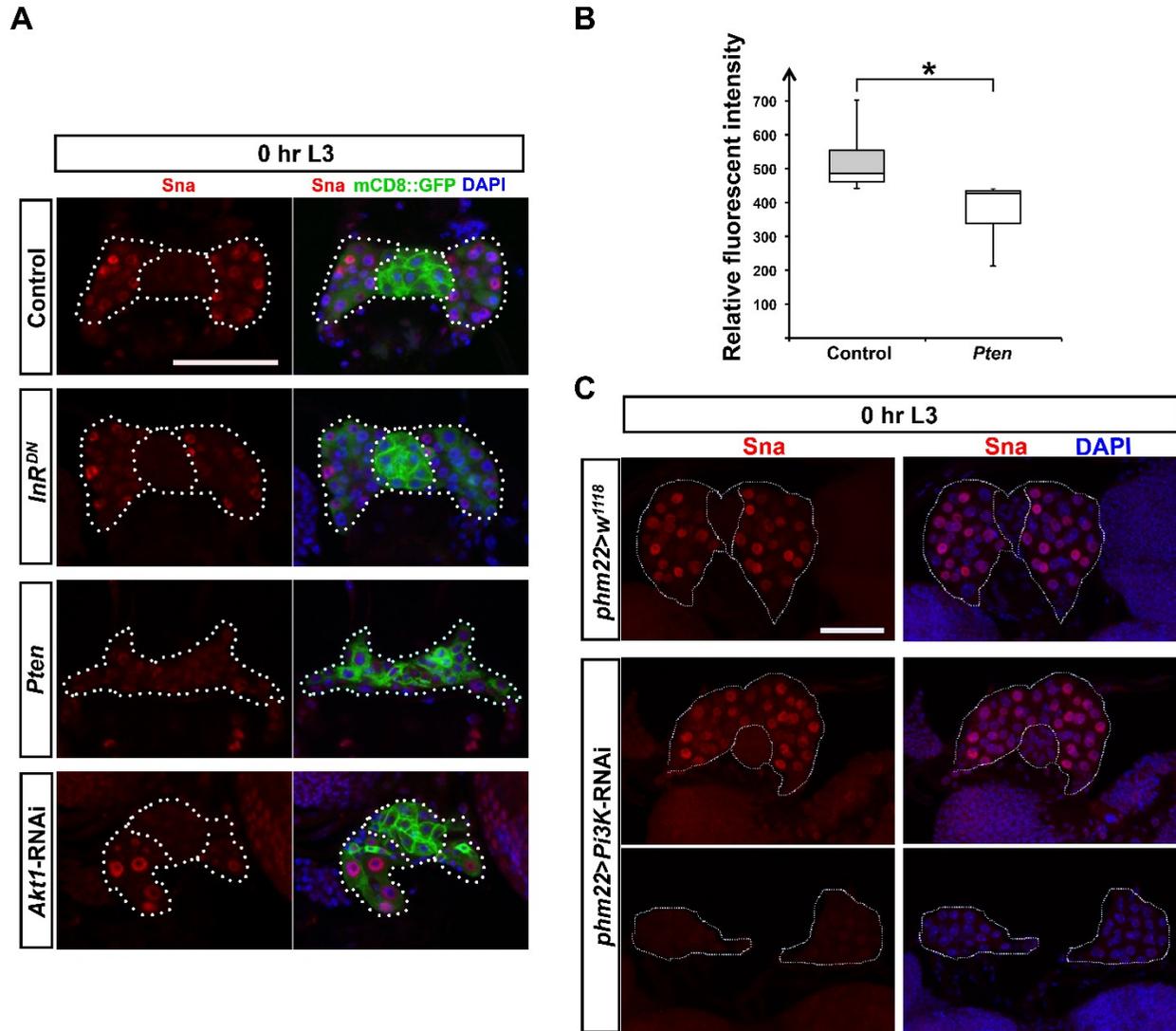


Figure 2-23. Sna levels in the PG are partially dependent on IIS signaling pathway.

(A) Sna levels were affected by *Pten* overexpression, but not by *InR^{DN}* and *Akt1*-RNAi. Control: *P0206-Gal4>w¹¹⁸*; *UAS:mCD8::GFP*. *InR^{DN}*: *P0206-Gal4>UAS:mCD8::GFP*; *UAS-InR^{DN}*. *Pten*: *P0206-Gal4>UAS:mCD8::GFP*; *UAS-Pten-cDNA*. *Akt1*-RNAi: *P0206-Gal4>UAS:mCD8::GFP*; *UAS-Akt1-RNAi*. 5-10 samples were examined for each condition. (B) Box plot of average Sna fluorescent intensity in the nuclei of each PG examined for control (*P0206-Gal4>w¹¹⁸*; *UAS:mCD8::GFP*) and *Pten* overexpression (*P0206-Gal4>UAS:mCD8::GFP*; *UAS-Pten-cDNA*). * $p < 0.05$; (Student's t-test). (C) Expression of *Pi3K*-RNAi in the PG partially affected Sna levels. 11 control samples (*phm22>w¹¹⁸*) and 15 RNAi samples (*phm22>Pi3K-RNAi*) were examined. Nine out of 15 RNAi samples showed relatively normal Sna levels (upper panel) while six out of 15 samples showed reduced Sna levels (lower panel). (A and C) Tissues were stained with anti-Sna antibody as well as DAPI to indicate the nuclei. PG and CA were marked by the white-dotted line. Scale bars: 50 μ m applicable to all samples.

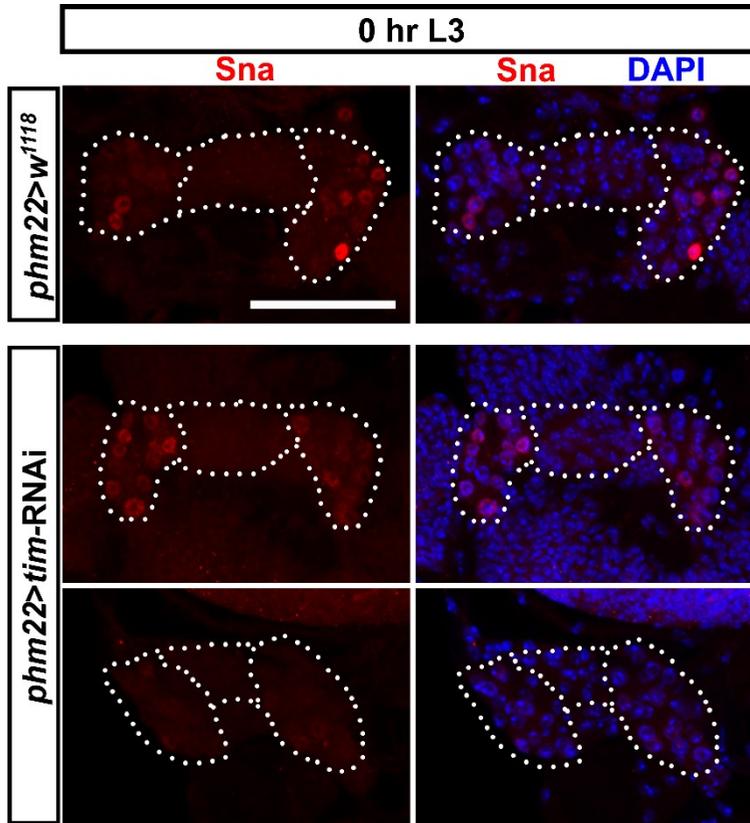


Figure 2-24. Sna levels in the PG are partially dependent on *tim*.

Single plane confocal images showing the anti-Sna antibody staining. Animals were entrained to a 12-hr light-dark cycle for at least three days prior to dissection. 14 control samples (*phm22>w¹¹¹⁸*) and 23 RNAi samples (PG>*tim*-RNAi) were examined. 14 out of 23 RNAi samples showed relatively normal Sna levels (upper panel) while nine out of 23 samples showed reduced Sna levels (lower panel). PG and CA were marked by white-dotted line. Scale bars: 50 μ m applicable to all samples.

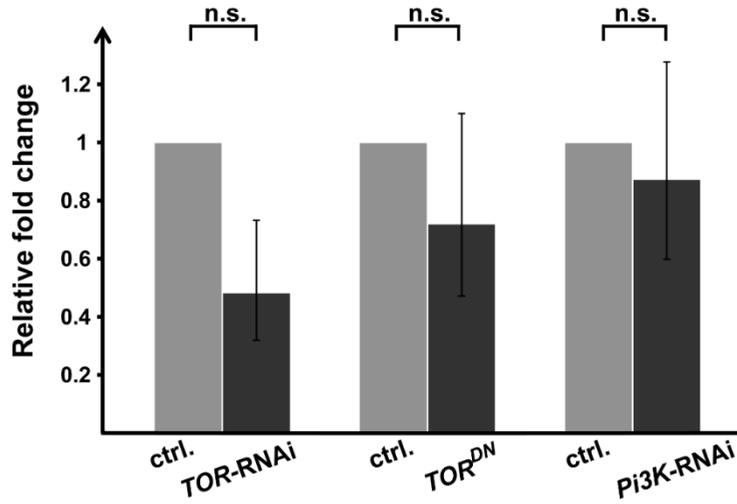


Figure 2-25. Loss-of-IIS/TOR signaling does not significantly affect *sna* mRNA levels in the PG.

The expression of *sna* in ring glands was examined via qPCR at 0 hr after the L2/L3 molt. Three replicates were performed for each condition. n.s.: not significant. ctrl.: *phm22-Gal4* crossed to Bloomington stock #36303 (background control for TRiP RNAi lines). *TOR*-RNAi: *phm22>TOR*-RNAi (TRiP line). *TOR*^{DN}: *phm22>TOR*^{DN} and *Pi3K*-RNAi: *phm22>Pi3K*-RNAi (TRiP line).

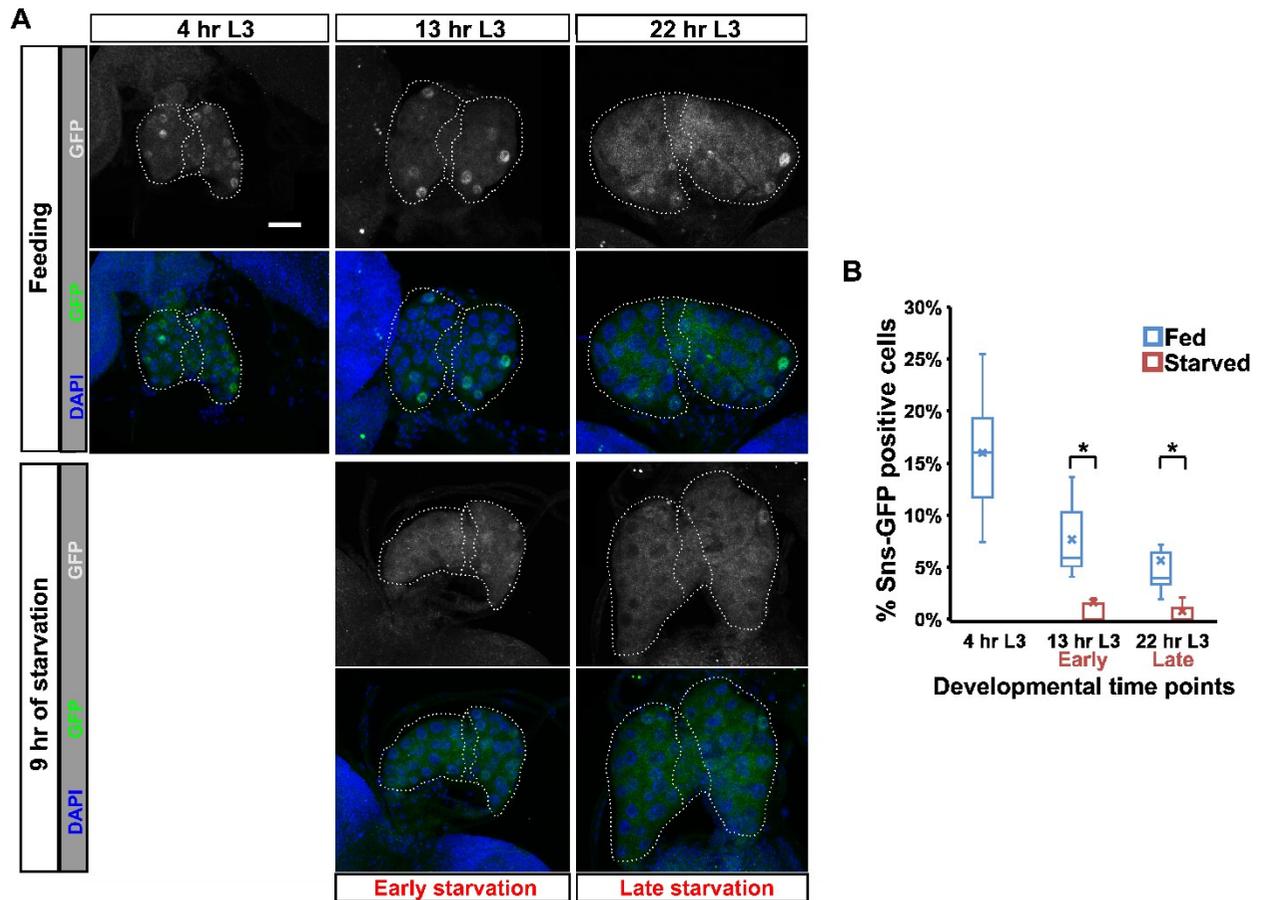


Figure 2-26. Presence of Sna in PG nuclei is dependent on nutrient conditions.

(A) Maximal projection of confocal images showing the Sna-GFP distribution in the PG under fed and starved conditions. Larvae were either fed with yeast paste or starved on 2% agar. Tissues were stained with the anti-GFP antibody and DAPI. The white-dotted line marks the boundary of PG and CA. Scale bar: 50 μ m. Early starvation: starvation started before CW attainment (at 4 hr L3). Late starvation: starvation occurred after CW attainment (at 13 hr L3). 5-10 ring glands were examined for each condition. (B). Box plot showing the percentage of Sna-GFP positive cells in the PG quantified from the results in panel (A). Early: early starvation. Late: late starvation. $*p < 0.05$.

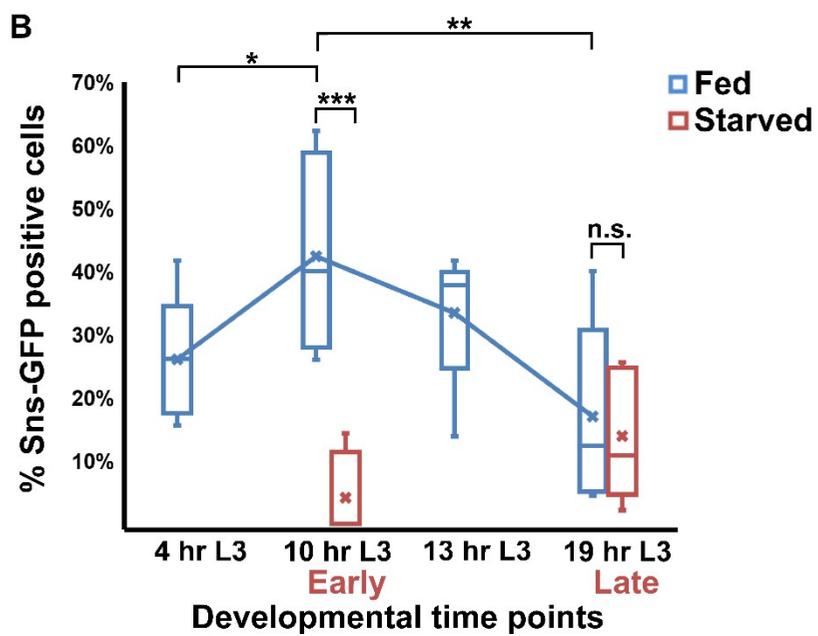
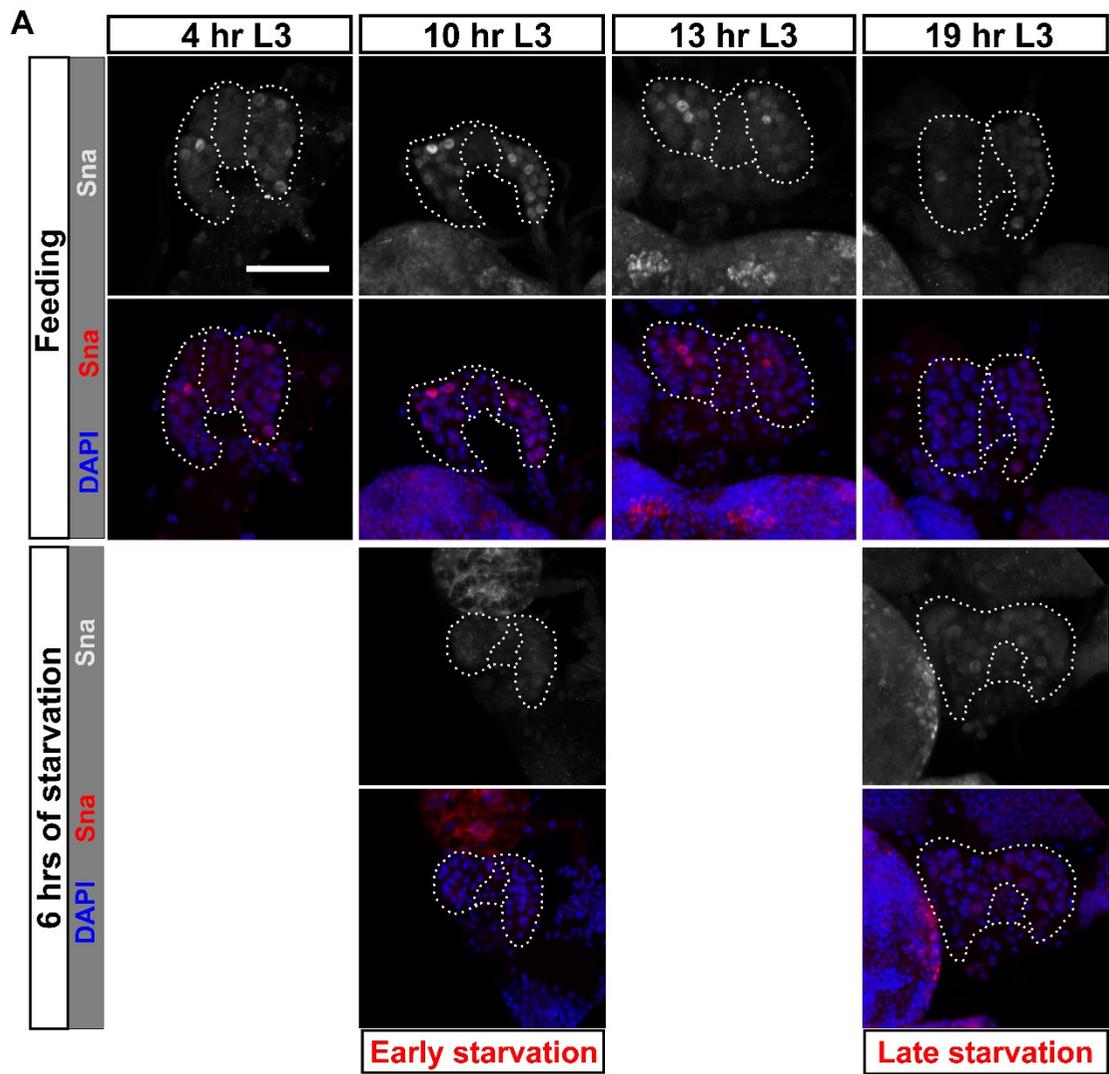


Figure 2-27. The presence of Sna in PG nuclei is dependent on nutrient conditions around the CW checkpoint but not after.

(A) Maximal projection of confocal images showing the Sna distribution in the PG under fed and starved conditions. Larvae were either fed with standard cornmeal-based medium or starved on 2% agar. Tissues were stained with the anti-Sna antibody and DAPI. The white-dotted line marks the boundary of PG and CA. Scale bar: 50 μm . Early starvation: starvation started before CW attainment (at 4 hr L3). Late starvation: starvation occurred after CW attainment (at 13 hr L3). 5-10 ring glands were examined for each condition. (B). Box plot showing the percentage of Sna-positive cells in the PG quantified from the results in panel (A). The average values of controls at each time point are connected by the blue line. Early: early starvation. Late: late starvation. * $p < 0.05$. ** $p < 0.01$, *** $p < 0.001$, n.s.: not significant.

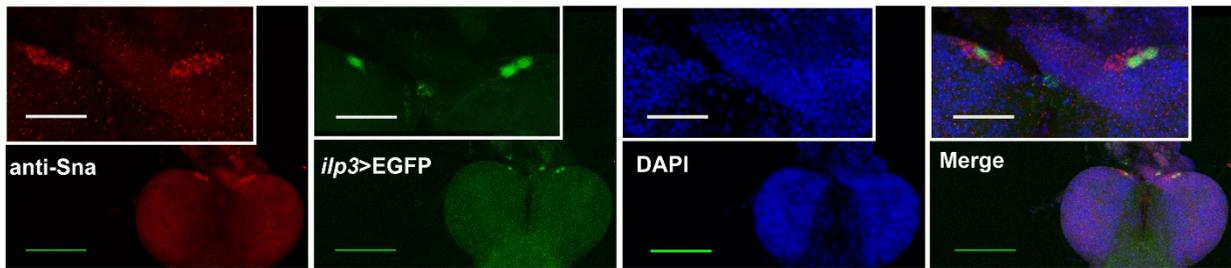


Figure 2-28. Sna is expressed in the IPCs in the brain.

Maximal projection of Z-stack confocal images showing Sna distribution in the brain hemisphere. *ilp3-Gal4* line was crossed to *UAS-EGFP* line to mark the insulin producing cells with EGFP. Tissues were dissected from early 3rd instar larvae and stained with anti-Sna antibody and DAPI. White scale bar: 50 μm . Green scale bar: 100 μm .

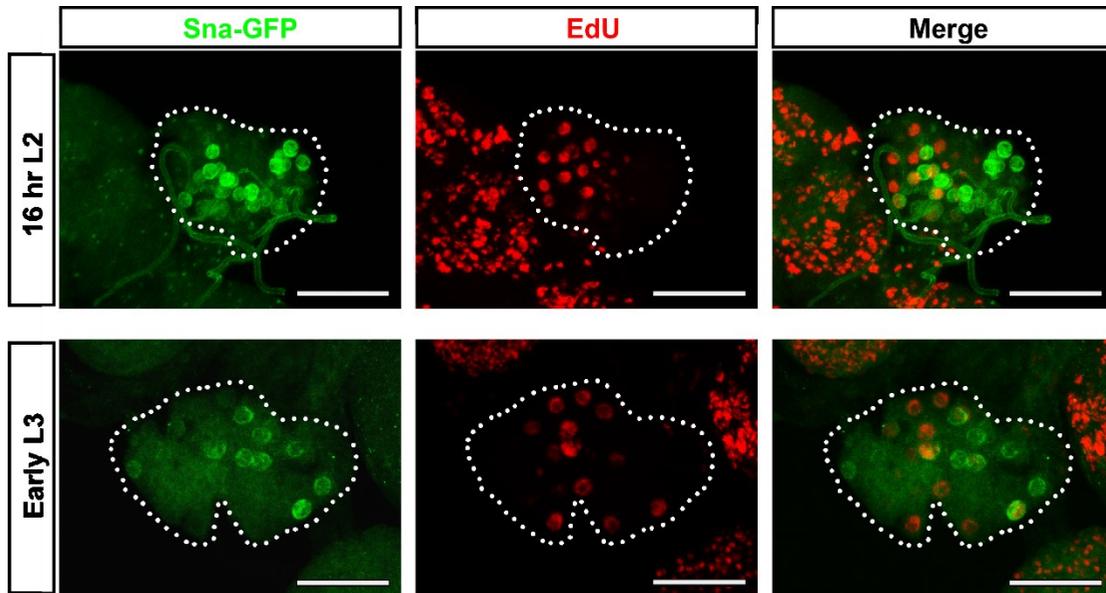


Figure 2-29. *Sna*⁺ cells do not largely overlap with S-phase cells in the PG.

Maximal projection of Z-stack confocal images showing the *Sna*-GFP distribution as well as the S-phase cells by EdU incorporation at two developmental stages. Early L3: 8-12 hr L3 (around the time of CW attainment). Tissues were dissected from *sna-gfp* flies and incubated in 1x Ringer's solution with 10 μ M EdU for 30 min before the anti-GFP antibody staining and EdU Click-iT reaction. Scale bar: 50 μ m.

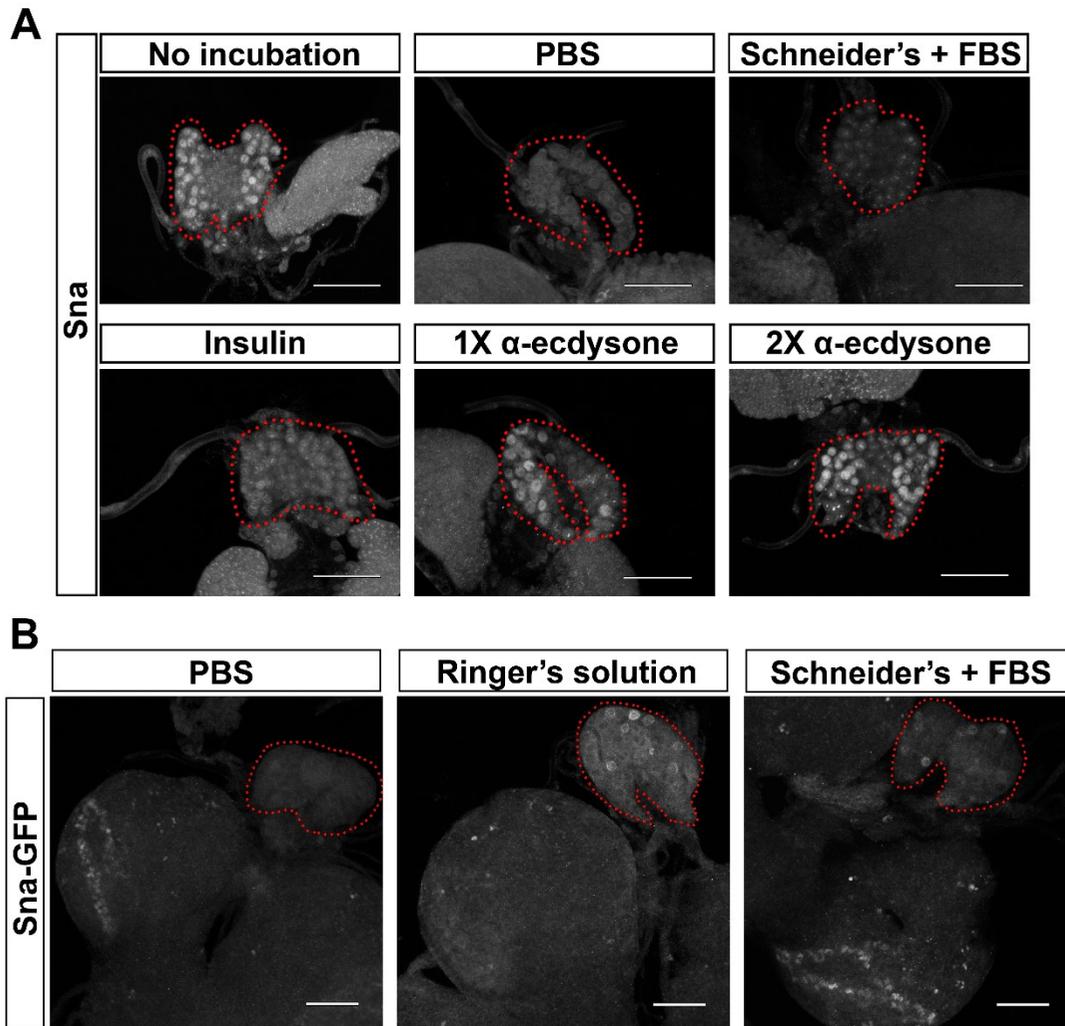


Figure 2-30. Sna protein levels in various *ex vivo* culture conditions.

(A) Maximal projection of Z-stack confocal images showing Sna protein distribution. Prior to anti-Sna antibody staining, tissues were incubated for 30 min *ex vivo* in various conditions as indicated. No incubation: tissues were dissected and directly subjected to antibody staining (as a control). Insulin: Schneider's insect medium with 10% Fetal Bovine Serum (FBS) and 10 $\mu\text{g}/\text{ml}$ insulin. 1x α -ecdysone: Schneider's insect medium with 10% FBS and 2 $\mu\text{g}/\text{ml}$ α -ecdysone. 2x α -ecdysone: Schneider's insect medium with 10% FBS and 4 $\mu\text{g}/\text{ml}$ α -ecdysone. (B) Maximal projection of Z-stack confocal images showing the Sna-GFP. Tissues were incubated *ex vivo* for 30 min in various conditions as indicated before the GFP antibody staining. PBS: 1x PBS. Ringer's solution: 1x Ringer's solution. Schneider's + FBS: Schneider insect medium with 10% Fetal Bovine Serum. (A-B) Images are shown in gray scale. Ring glands are marked with red dotted lines. Scale bar: 50 μm .

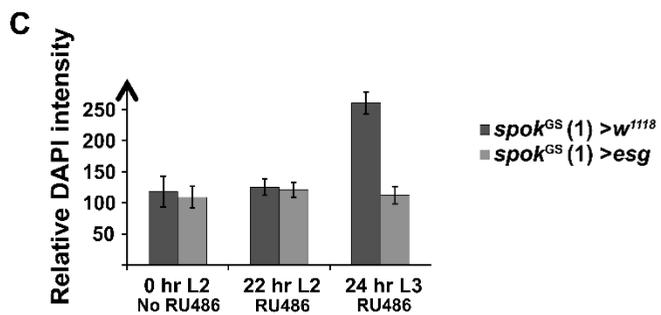
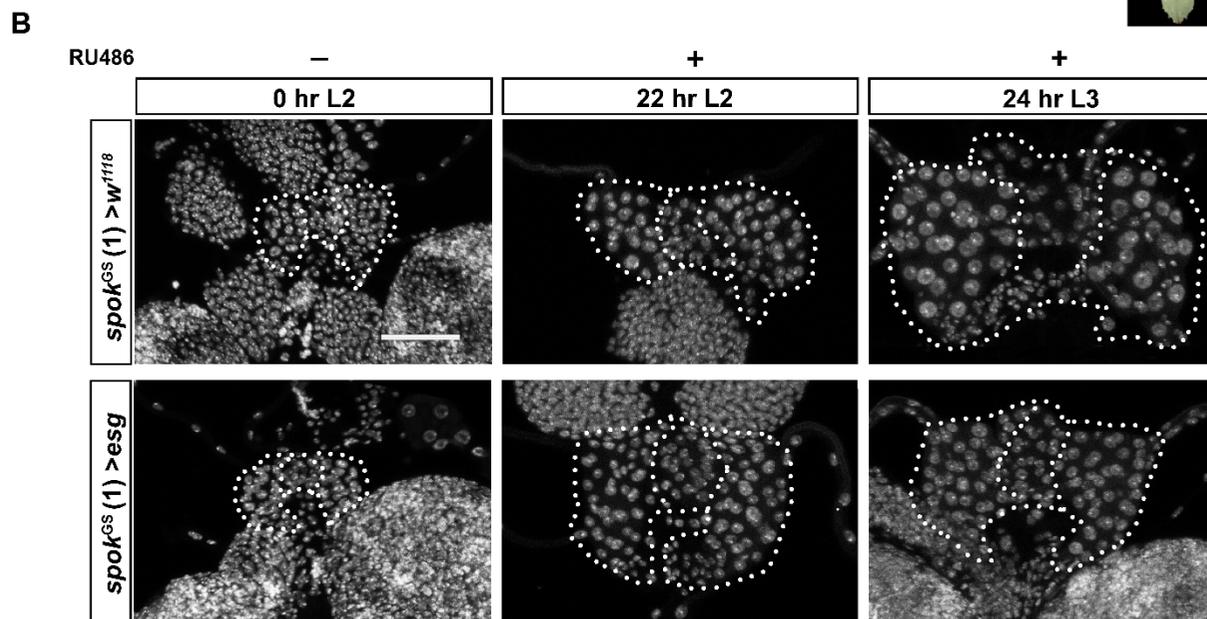
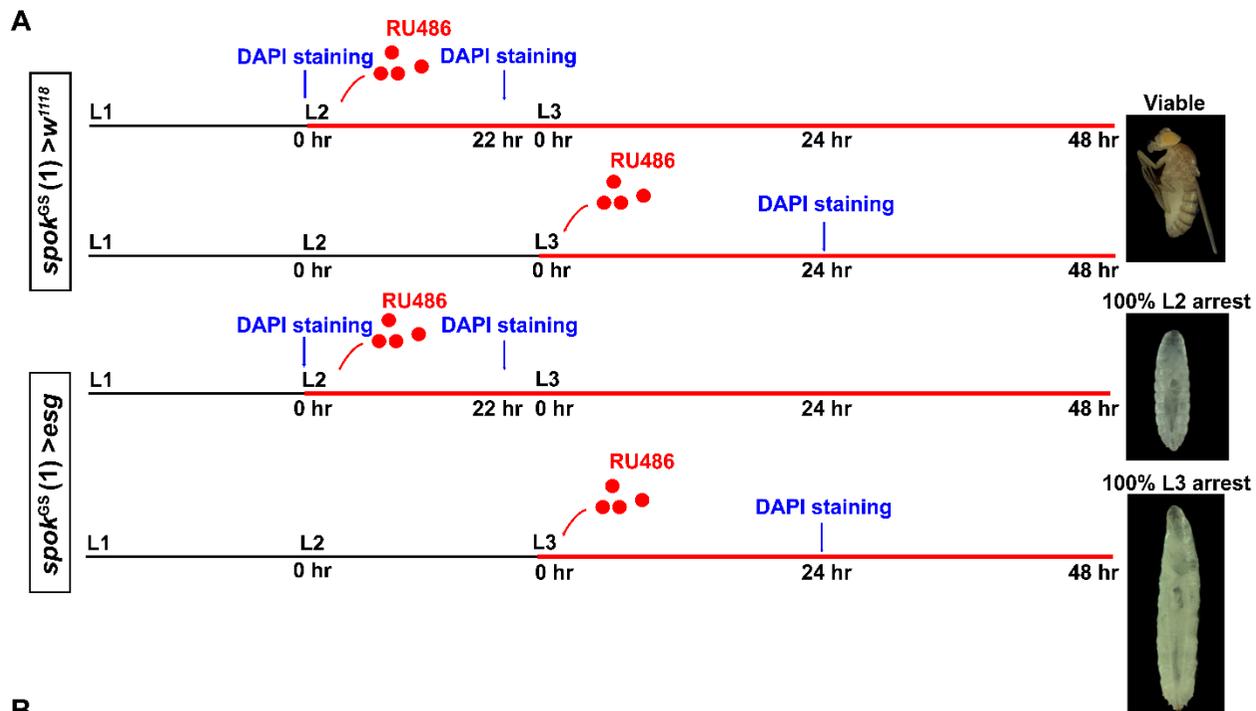


Figure 2-31. Overexpression of *esg* in the PG blocked endocycle progression.

(A) A schematic illustration showing RU486 feeding procedures for temporal control of transgene induction. The black line indicates developmental stages prior to RU486 treatment, and the red line indicates when animals were fed with RU486-containing food. Hours labelled with “DAPI staining” indicate time points when ring glands were dissected and imaged. The images illustrate the outcome of RU486 administration on developmental fates for both genetic backgrounds. (B) Maximal projection of Z-stack confocal images showing the size of PG nuclei. Samples were collected and stained with DAPI according to the procedures shown in panel A. White dotted line marks the boundary of the PG and CA. Scale bar: 50 μm applicable to all the samples shown in panel B. (C) Relative DNA content per PG nucleus calculated with the summation of the Z-stack images shown in panel B. No RU486: before RU486 administration. RU486: after RU486 administration. (A-D) *spok^{GS}* (1): *spok-Gal4-GeneSwitch* on the second chromosome.

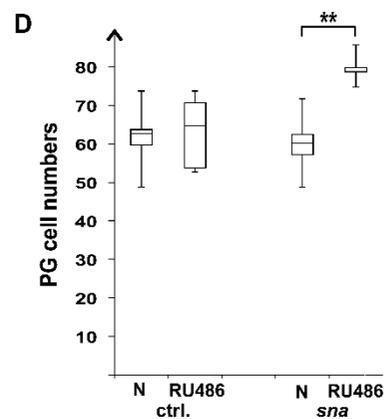
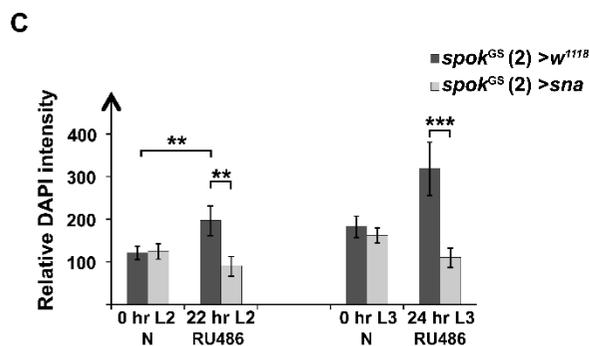
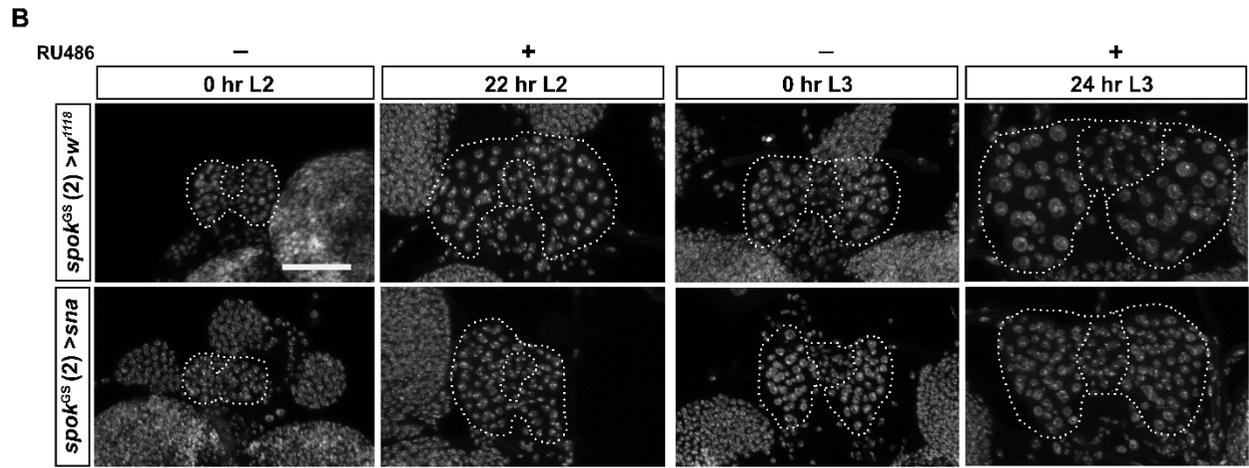
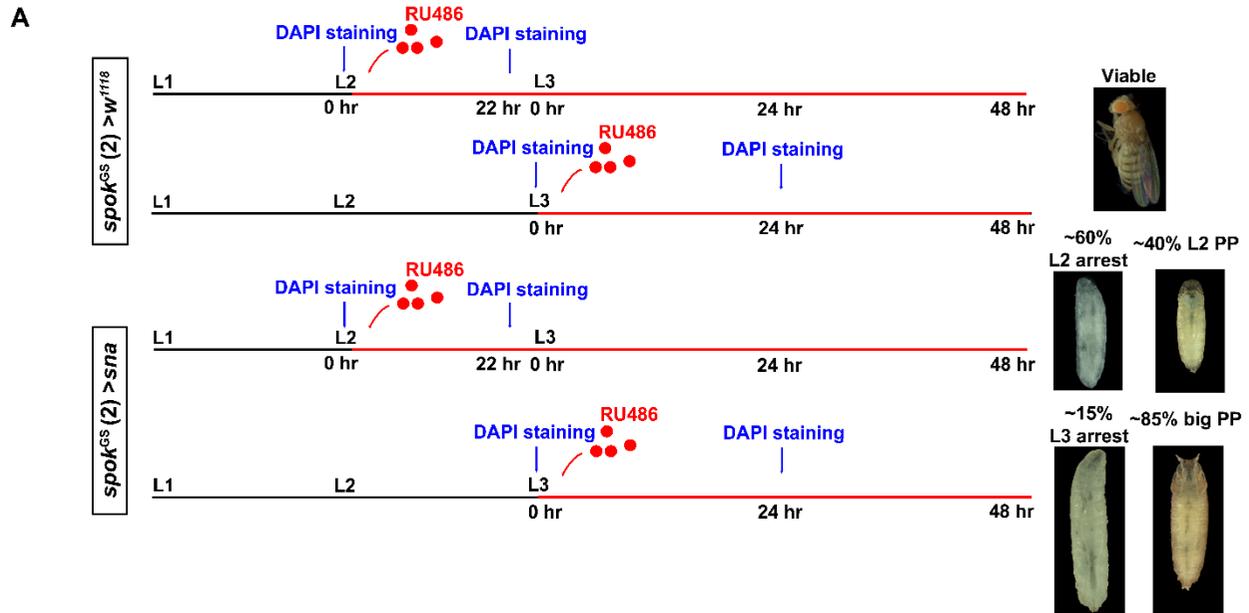


Figure 2-32. Overexpression of *sna* in the PG blocked endocycle progression and increased PG cell number.

(A) A schematic illustration showing RU486 feeding procedures for temporal control of transgene induction. The black line indicates developmental stages prior to RU486 treatment, and the red line

indicates when animals were fed with RU486-containing food. Hours labelled with “DAPI staining” indicate time points when ring glands were dissected and imaged. The images illustrate the outcome of RU486 administration on developmental fates for both genetic backgrounds. (B) Maximal projection of Z-stack confocal images showing the size of PG nuclei. Samples were collected and stained with DAPI according to the procedures shown in panel A. White dotted line marks the boundary of the PG and CA. Scale bar: 50 μm applicable to all the samples shown in panel B. (C) Relative DNA content per PG nucleus calculated with the summation of the Z-stack images shown in panel B. N: before RU486 administration (Nutrifly food alone). RU486: after RU486 administration. (D) PG cell numbers were increased in PG>*sna* overexpression animals. For this, I compared PG cell numbers before (N) and after (RU486) 24 hr of RU486 application during the L3 stage for both genotypes. ctrl.: *spok^{GS} (2) >w¹¹¹⁸*. *sna*: *spok^{GS} (2) >sna-cDNA*. (C and D). * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. (A-D) *spok^{GS} (2); spok-Gal4-GeneSwitch* on the third chromosome.

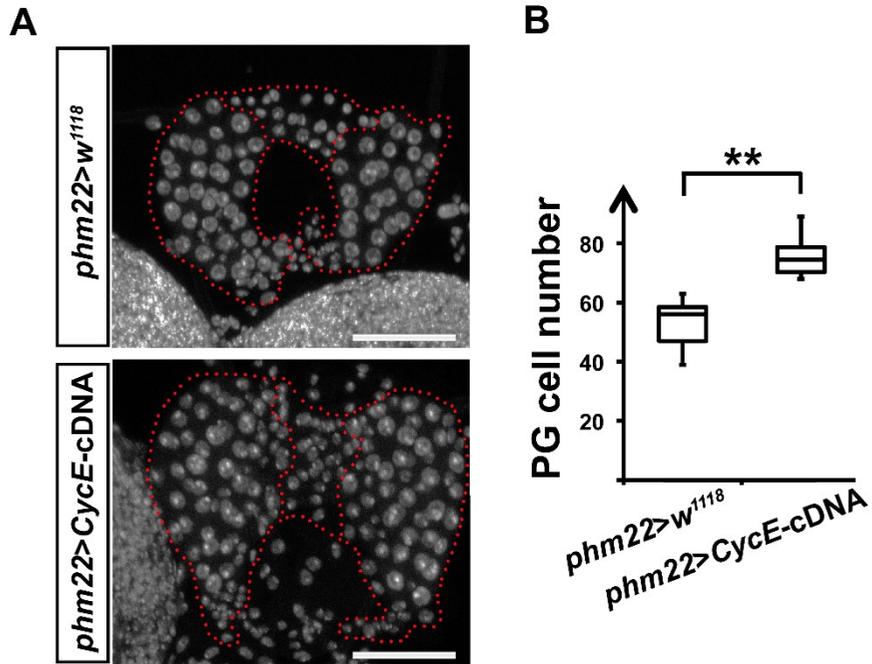


Figure 2-33. Overexpressing *CycE* in the PG increased the cell number.

(A) Maximal projection of Z-stack confocal images showing the number of cells in the PG. Tissues were dissected at the early 3rd instar stage and stained with DAPI. The white dotted line marks the boundary of the PG and CA. Scale bar: 50 μ m. (B) Box plot showing the quantified PG cell number for control (*phm22>w¹¹⁸*) and *CycE* overexpression (*phm22>CycE-cDNA*). ** $p < 0.01$ (based on Student's t-test).

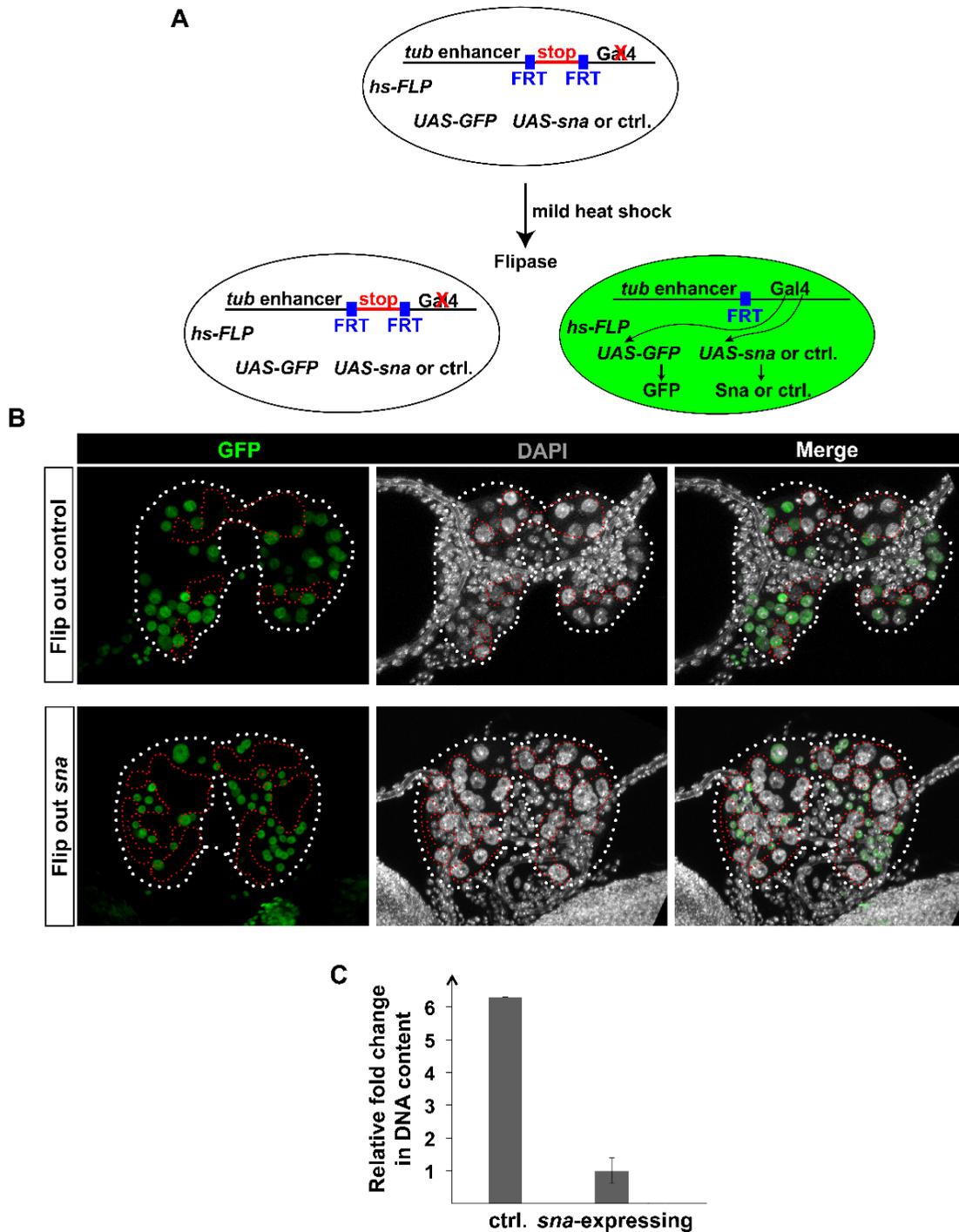


Figure 2-35. *sna* inhibits the endocycle at the CW checkpoint in a cell-autonomous manner.

(A) A schematic illustration of the flip-out-Gal4 system. *hs-FLP*; *tubulin-FRT-CD2-FRT-Gal4*, *UAS-GFP* flies were used where a stop codon that is flanked by two FRT sites in front of the *Gal4* sequence normally prevents the production of Gal4 protein. After a mild heat shock, the stop codon will be excised by Flippase (FLP) in a stochastic manner, which in turn switches on transgene expression in a mosaic pattern (the *sna*-overexpressing cells will be marked by the presence of GFP). *hs-FLP*: Flippase controlled by a heat shock promoter. *tub*: *tubulin*. Blue boxes represent FRT sites. stop: a stop codon. (B). Maximal projection of Z-stack confocal images showing the size of the nuclei. Samples were collected and stained with DAPI. Flip

out control: *hs-FLP; tubulin-FRT-CD2-FRT-Gal4, UAS-GFP>y^[1] w^[67c23]*. Flip out *sna*: *hs-FLP; tubulin-FRT-CD2-FRT-Gal4, UAS-GFP>y^[1] w^[67c23]; sna-cDNA*. The white-dotted line marks the boundary of the PG and CA. The red-dotted line marks the non-GFP cells. Scale bar: 50 μ m. (C). The relative fold change of DNA content per PG cell between *sna*-overexpressing cells and non-overexpressing cells (ctrl.) in *sna*-flip-out animals. ** $p < 0.01$ (based on Student's t-test).

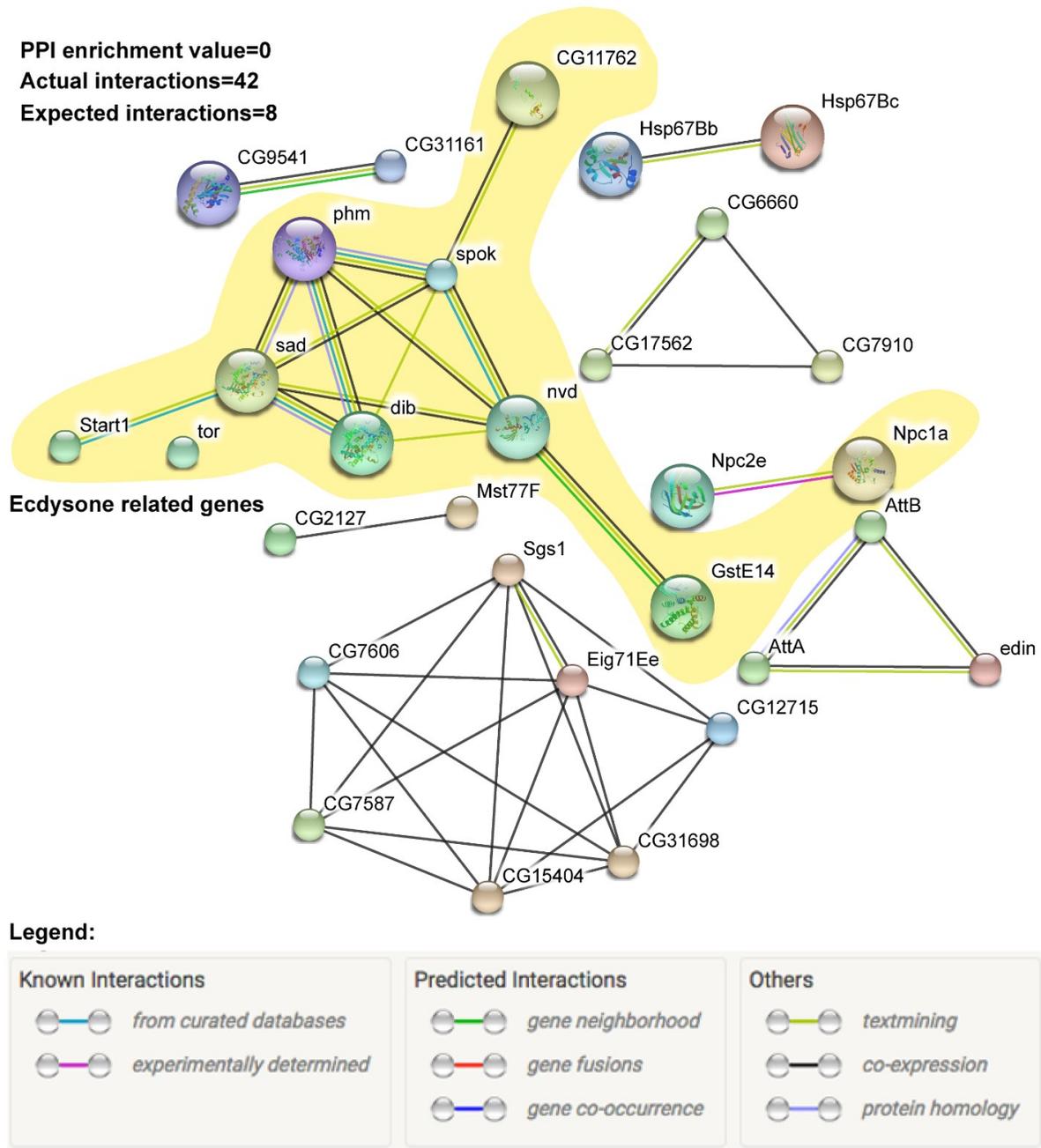


Figure 2-36. Loss-of-*sna* in the PG disrupted the expression of genes essential for ecdysone biosynthesis.

Interaction network of >3-fold downregulated genes in PG>*sna*-RNAi (*phm22-Gal4*>*sna*-RNAi; *UAS-Dicer2*) ring glands compared to controls (*phm22-Gal4*> *w¹¹¹⁸*; *UAS-Dicer2*) generated by the STRING database. The ecdysone-related genes are interconnected with each other and highlighted by yellow shading. The non-associated genes were not included in the figure. The large node size indicates the availability of protein structure information. Node colors have no particular meaning. Line colors indicate different types of evidence for the interactions (see legend). PPI enrichment value: protein-protein interaction enrichment value (the smaller, the more enriched). Actual interactions: number of protein interactions. Expected interactions: number of protein interactions that would be expected for a random set of proteins of similar size, drawn from the genome.

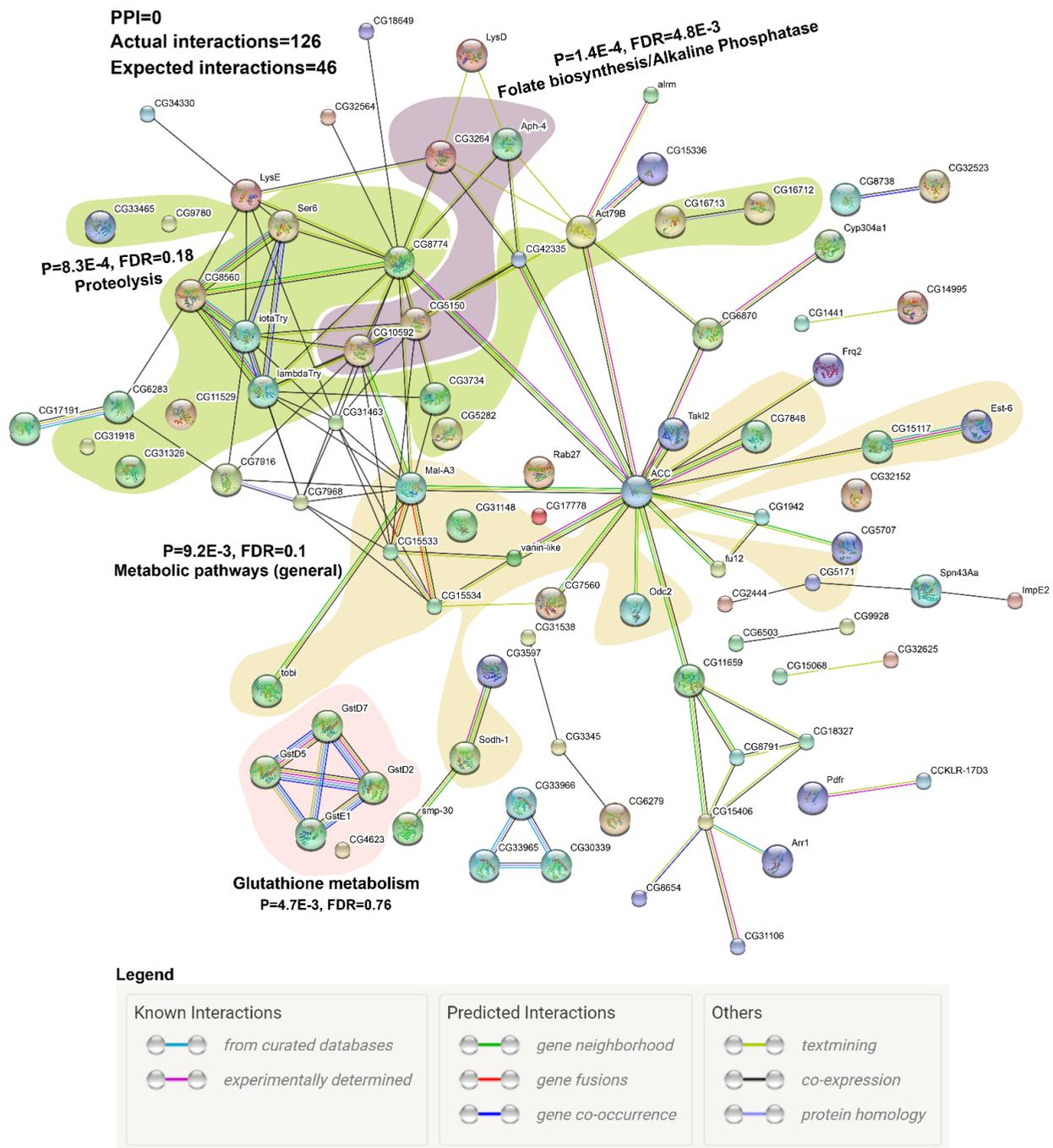


Figure 2-37. The upregulated genes in PG>*sna*-RNAi ring glands suggested metabolism-related defects. Interaction network of >3-fold upregulated genes in PG>*sna*-RNAi (*phm22-Gal4>sna*-RNAi; *UAS-Dicer2*) ring glands compared to controls (*phm22-Gal4>w¹¹¹⁸*; *UAS-dicer2*) generated by the STRING database. Several gene ontology (GO) groups (based on DAVID GO) within the big network were highlighted using different color shading. *p* value: based on modified Fisher exact model (the smaller, the more enriched for the GO term). FDR: False discovery rate based on Benjamini method. The non-

associated genes were not included in the figure. The large node size indicates the availability of protein structure information. Node colors have no particular meaning. Line colors indicate different types of evidence for the interactions (see legend). PPI: protein-protein interaction enrichment value. Actual interactions: number of protein interactions. Expected interactions: number of protein interactions that would be expected for a random set of proteins of similar size, drawn from the genome.

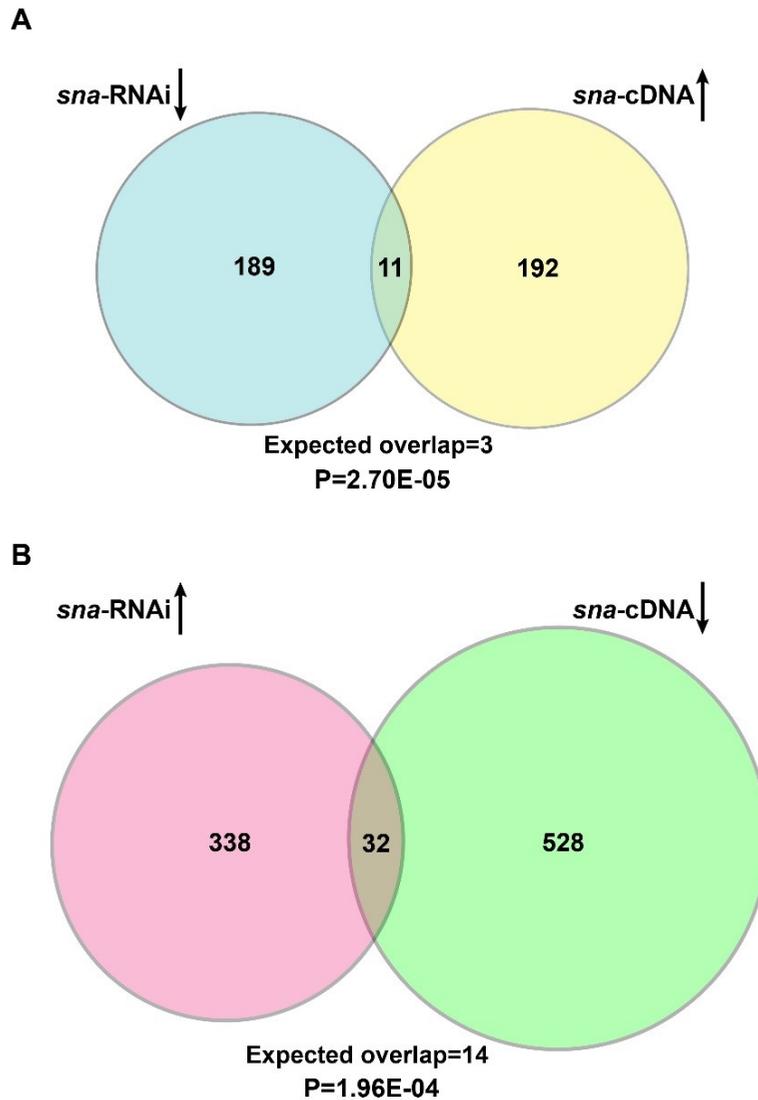


Figure 2-38. Cross comparison of the *sna*-RNAi RNA-Seq data and the *sna*-overexpression RNA-Seq data.

(A) Venn diagram showing the overlap between genes >2-fold downregulated in the PG>*sna*-RNAi ring glands and genes upregulated >2-fold in the ring gland upon *hs*>*sna*-cDNA expression. The eleven overlapping genes are listed in Table 2-8. (B) Venn diagram showing the overlap between genes upregulated >2-fold in the PG>*sna*-RNAi ring gland and genes downregulated >2-fold in the ring gland upon *hs*>*sna*-cDNA. The 32 overlapped genes are listed in Table 2-9. (A-B) The direction of the arrows indicates either the up- or down-regulation. *p* value represents Pearson's chi-squared test (χ^2) results.

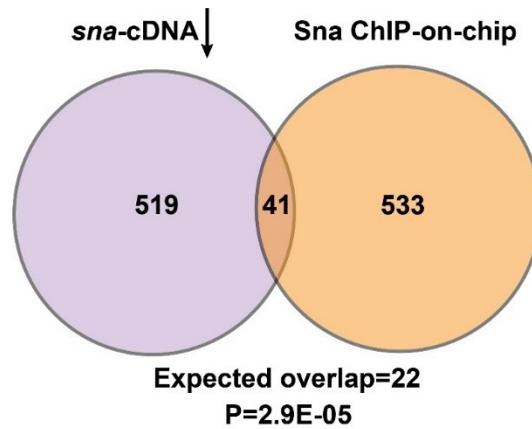


Figure 2-39. Comparison between *sna*-overexpression RNA-Seq data and published ChIP-on-chip data.

(A) Venn diagram showing that genes downregulated >2-fold in the ring gland upon *hs>sna*-cDNA significantly overlapped with genes associated with Sna binding peaks. The Sna binding peaks were determined by ChIP-on-chip data released from the Berkeley Drosophila Transcription Network Project (BDTNP). The 41 overlapped genes are listed in Table 2-11. *p* value represents Pearson's chi-squared test (χ^2) results.

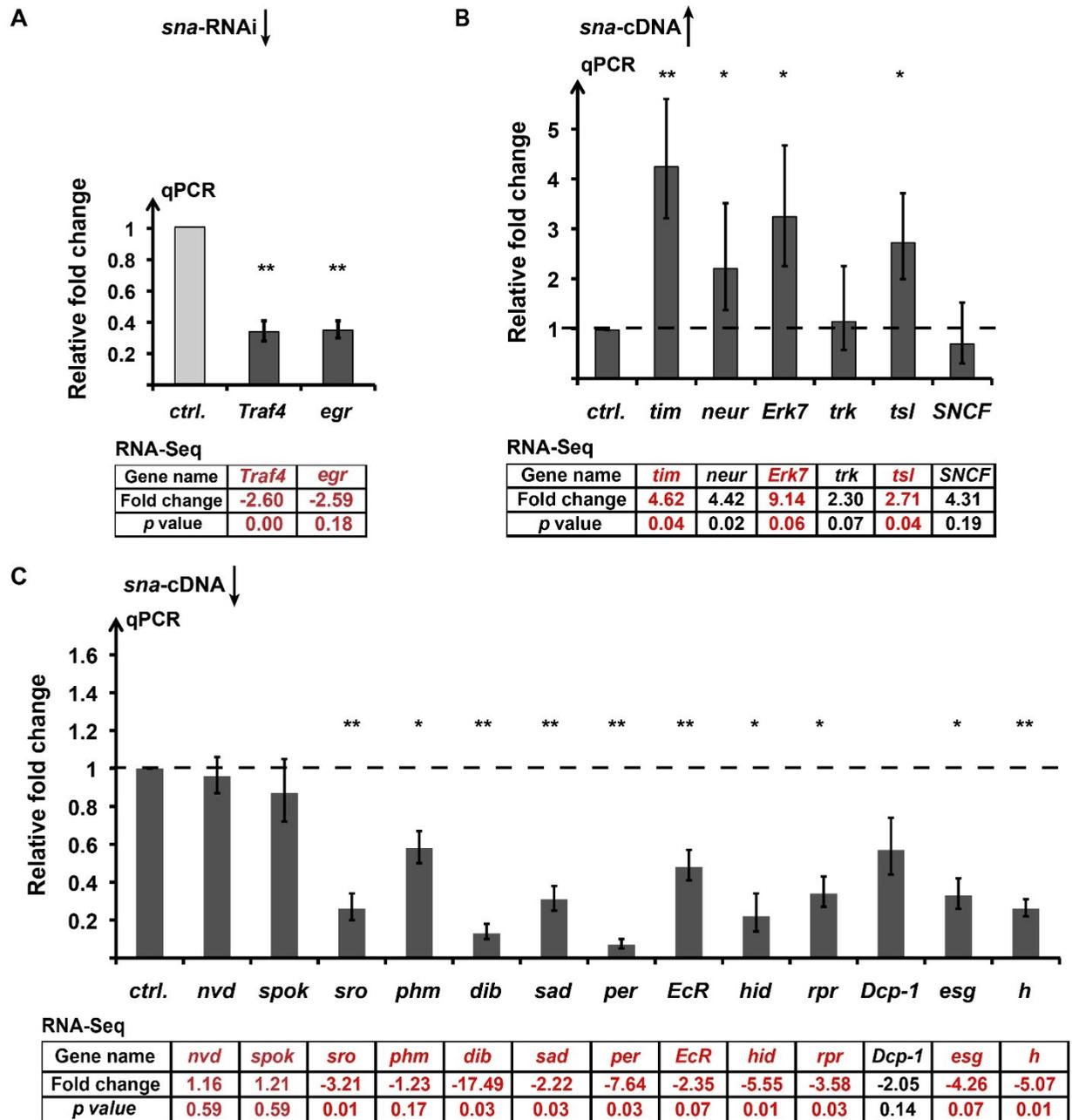


Figure 2-40. qPCR results showing the expression of the misregulated genes in either PG>*sna*-RNAi or *hs*>*sna*-cDNA ring glands.

(A) qPCR results of the two downregulated genes in PG>*sna*-RNAi samples. Ring glands were collected at 24 hr L3. Relative fold change was determined by comparing the expression in the RNAi to that of the ctrl. (*phm22-Gal4*>*sna*-RNAi; *UAS-Dicer2*). (B) qPCR results for some upregulated genes in *hs*>*sna*-cDNA ring glands. (C) qPCR results for some downregulated genes in *hs*>*sna*-cDNA ring glands. (B and C) The expression of each gene in *sna*-overexpression was normalized to the expression in controls (*hs-Gal4*> *y^[1] w^[67c23]*); the expression in controls shown as 1 and indicated by the dotted line. (A-C) Tables show the fold change and *p* value from the RNA-Seq analysis. Red-colored genes have been validated by qPCR. ***p*<0.01 and **p*<0.05 (based on Student's t-test).

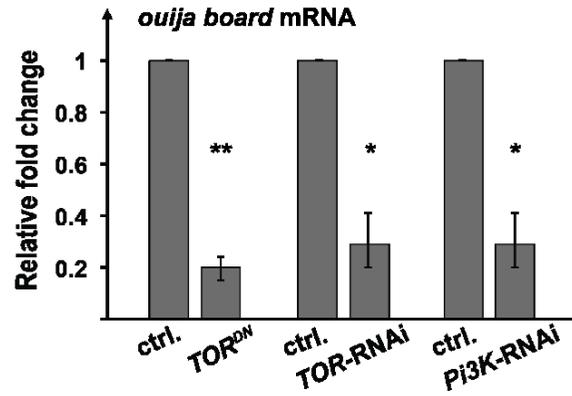


Figure 2-41. *ouib* mRNA levels are dependent on the IIS/TOR signaling.

(A) The expression of *ouija board* (*ouib*, aka *CG11762*) in the ring gland was examined by qPCR at 0 hr after the L2 to L3 molt. Three replicates were included for each condition. * $p < 0.05$ and ** $p < 0.01$. ctrl.: *phm22-Gal4* crossed to Bloomington stock #36303 (background control for TRiP RNAi line). *TOR*-RNAi: *phm22 > TOR-RNAi* (TRiP line). *TOR*^{DN}: *phm22 > TOR*^{DN} and *Pi3K*-RNAi: *phm22 > Pi3K-RNAi* (TRiP line).

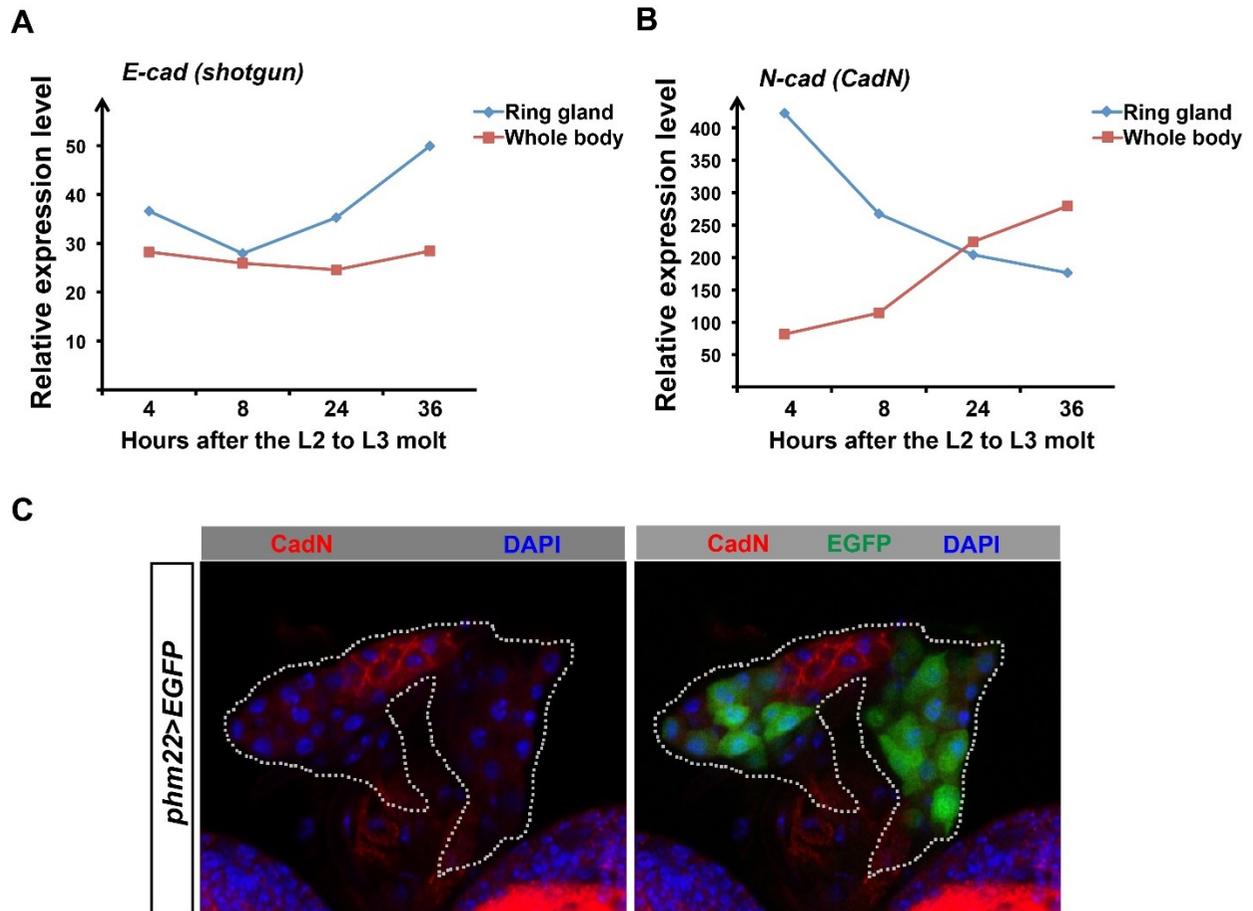


Figure 2-42. The expression profile of two cell-adhesion genes in the ring gland.

(A) Microarray results showing the expression profile of *shotgun* (*aka E-cadherin*) in the ring gland as well as in the whole larva at four different time points after the L2 to L3 molt. (B) Microarray results showing the expression profile of *N-cadherin* (*cadN*) in the ring gland as well as in the whole larva at four different time points after the L2 to L3 molt. (C) Immunofluorescent images of RGs dissected from the *phm22>EGFP* animals (*phm22-Gal4>UAS-EGFP*) at 0 hr L3. Tissues were stained with anti-CadN antibody and DAPI. The white dotted line marks the RG area.

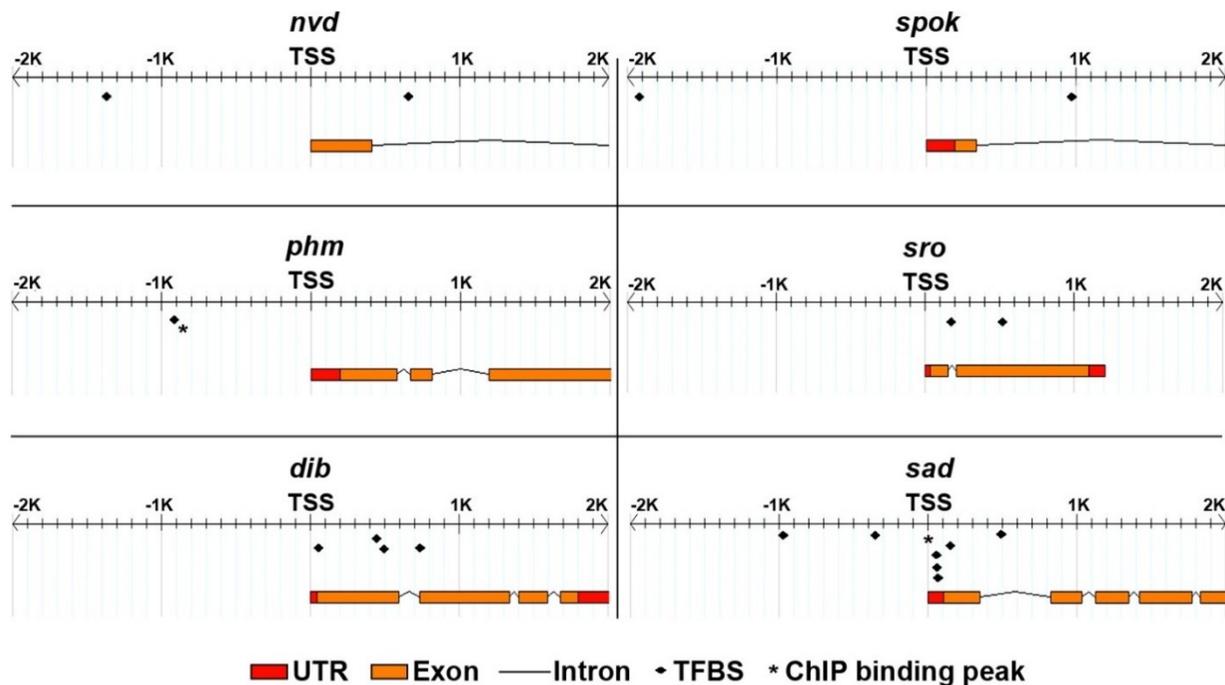


Figure 2-43. The schematic view of predicted *Sna* binding sites in the six major ecdysone biosynthetic genes.

The position of predicted *Sna* binding sequences in the ecdysone biosynthetic genes relative to the transcription start site (TSS). The *in silico* search was carried out using the web server, IN-silico SEarch for Co-occurring Transcription factors (INSECT 2.0), which only examines the sequence 2 kb upstream and 1 kb downstream from the transcription start site (TSS) for each gene. The gene structure is only shown up to 2 kb upstream of the TSS. The detailed position and actual sequence for the predicted binding sites are shown in Table 2-12. UTR: untranslated region. TFBS: *Sna* transcription factor binding site. ChIP binding peaks were based on the *Sna* ChIP-on-chip data released by the Berkeley *Drosophila* Transcription Network Project (BDTNP).

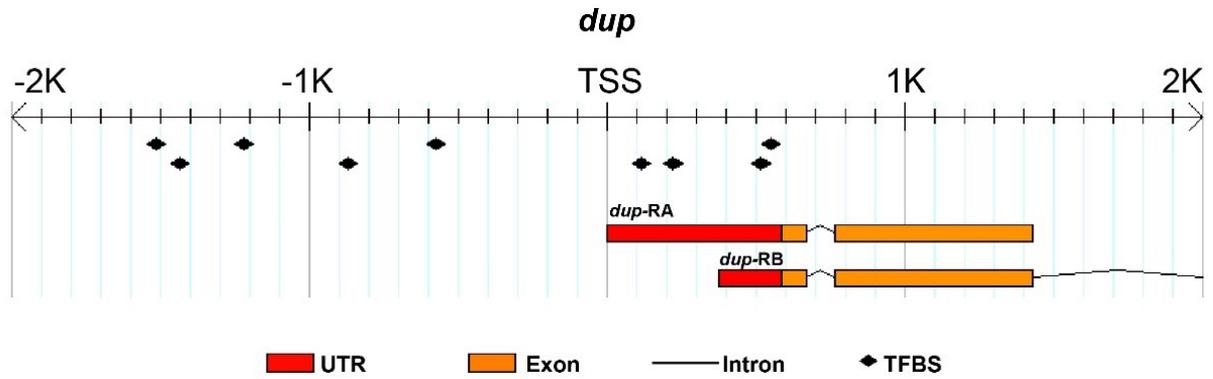


Figure 2-44. The predicted Sna binding sites in the endoreplication related gene *double park (dup)*.

The schematic view showing the position of predicted Sna binding sequences in *dup* relative to the transcription start site (TSS). The *in silico* search was carried out using the web server, IN-silico SEArch for Co-occurring Transcription factors (INSECT 2.0), which only examine the sequence 2 kb upstream and 1 kb downstream from the transcription start site (TSS) for each gene. The structure for both mRNA isoform (*dup-RA* and *dup-RB*) was represented by boxes (exon) and lines (intron) only up to 2 kb upstream of TSS. The detailed position and actual sequence for the predicted binding sites are shown in Table 2-13. UTR: untranslated region. TFBS: Sna transcription factor binding site.

Figure 2-45. Multiple protein sequences alignment between several Sna family proteins.

Multiple protein sequences alignment between Human Snail1 (SNA1_HUMAN), mouse Snail1 (SNA1_MOUSE), *Drosophila melanogaster* Snail protein (SNA_DROMelanogaster), *Drosophila simulans* Snail protein (SNA_DROsimulans) and *Drosophila virilis* Snail protein (SNA_DROvirilis) using ClustalW (1.8.3). *Drosophila simulans* is closely related to *Drosophila melanogaster* while *Drosophila virilis* is distantly related to *Drosophila* species. “*” (asterisk) indicates positions which have a single, fully conserved residue. “:” (colon) indicates conservation between groups of strong similarity. “.” (period) indicates conservation between groups of weakly similar properties. Dark gray shading highlights the four zinc finger domains in Snail1 (human and mouse). Light gray shading highlights the five zinc finger domains in Snail (in three *Drosophila* species). Green shading highlights the nuclear export signal (NES) identified in the mouse Snail1. The red box indicates the serine rich region SRD in Snail1, in which lie the two motifs for GSK-3 β phosphorylation. The actual phosphorylation sites (the serine) are red-coded. Yellow shaded tyrosine or serine residues are predicted GSK-3 β phosphorylation sites in *Drosophila* based on GPS 3.0. The blue dots indicate the predicted phosphorylation sites that are conserved across twelve *Drosophila* species.

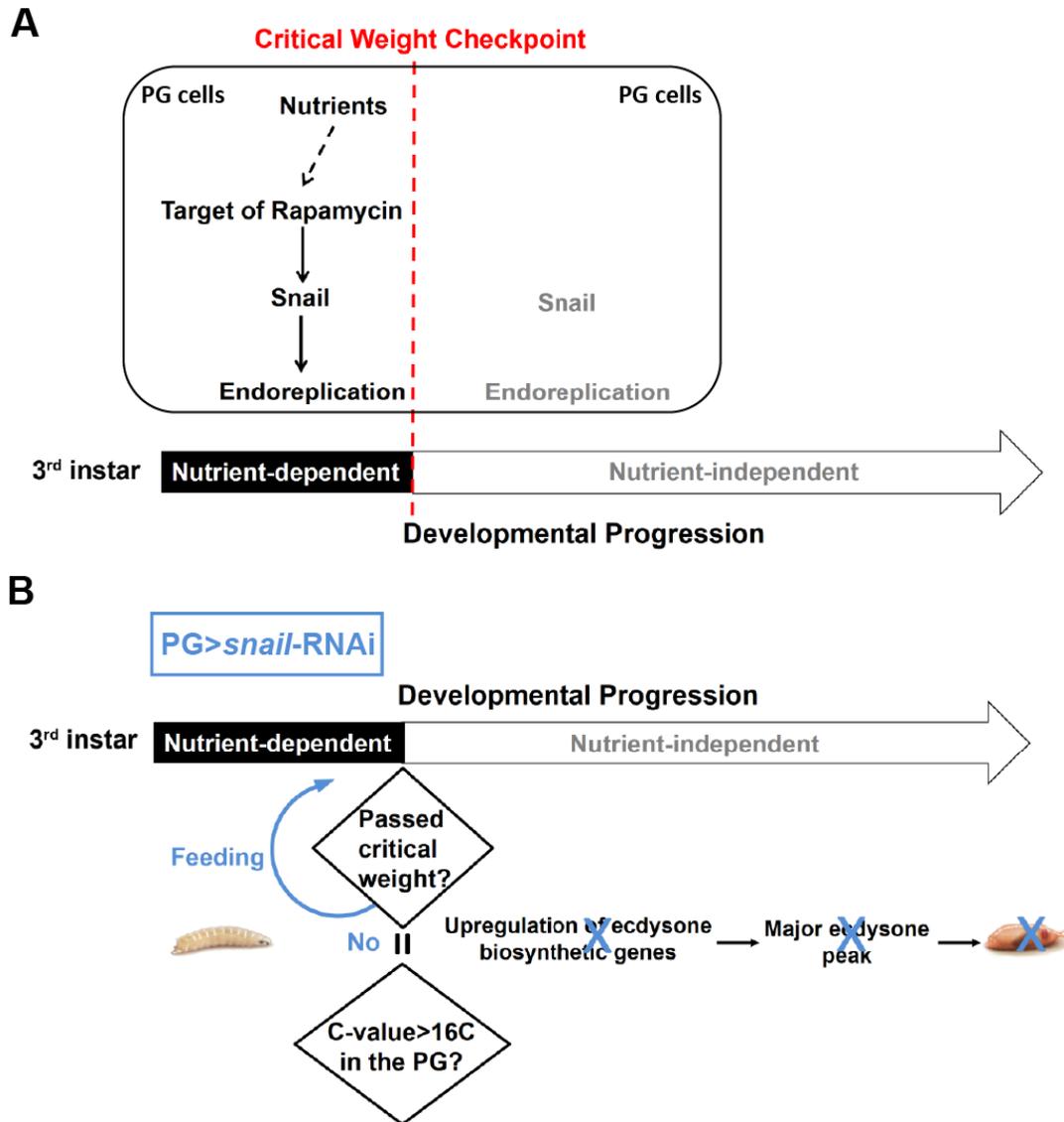


Figure 2-46. Models for Snail function in the PG.

(A) Snail function in regulating endoreplication in the PG. In the early half of the last larval stage, which is 3rd instar in *Drosophila*, there is a developmental switch called critical weight checkpoint. Developmental progression is nutrient-dependent before the CW checkpoint, but nutrient-independent after animals attained CW (94). There is one round of endoreplication in PG cells tightly coupled with critical weight attainment (101). Around the time of critical weight attainment, nutrient conditions mediated by the Target of Rapamycin (TOR) signaling control the levels of Snail to promote endoreplication, which appears to be crucial for animals to pass the critical weight checkpoint. After the critical weight attainment, Snail levels will decline so that nutritional inputs are no longer relevant. Therefore, the endocycle progression is no longer dependent on nutrients after the critical weight checkpoint. (B) Consequence of disrupting Snail function in the PG. When *snail* is knocked down in the PG, the DNA content of PG cells never reach beyond 16C, the putative intrinsic threshold for passing the critical weight checkpoint. Therefore, larvae keep feeding and never commit to metamorphosis. As a consequence, they did not upregulate expression of ecdysone biosynthetic genes to prepare for the major ecdysone pulse for triggering the onset of metamorphosis, a series of events that normally happen after animals passed critical weight in the wide type (101).

**Chapter 3. The deadenylase Curled (*aka* Nocturnin) and the CCR4-
NOT deadenylase complex have distinct roles in regulating the
ecdysone production in the prothoracic gland**

3.1 Introduction

3.1.1 *Curled*, a *Drosophila* homolog of *Nocturnin*

Microarray analyses comparing the gene expression profiles of the RG and the larval whole body at 4 different developmental time points during the L3 have shown that *curled* (*cu*) transcripts are highly enriched in the RG at all four time points we examined (Fig. 3-1). This is consistent with previous findings that *cu* is specifically expressed in the proventriculus and RG during the third instar larval stage (123). *cu* was originally identified as a circadian clock regulated gene with high expression at night in photoreceptors of *Xenopus laevis* (249). Therefore, the gene is also called *nocturnin* in species other than *Drosophila*. Mouse *nocturnin* is also rhythmically expressed in a variety of tissues such as liver, kidney and spleen (250). It encodes a protein homologous to yeast carbon catabolite repression 4 (CCR4), a putative deadenylase (124). Nocturnin's deadenylase activity likely contributes to the circadian control of gene expression by rhythmically controlling transcript degradation. *Drosophila* Cu, like other CCR4 homologs, has an endo/exonuclease/phosphatase domain (EEP), which can hydrolyse RNA poly(A) tails from 3' to 5' in a Mg²⁺-dependent manner, leading to mRNA decay. *cu* null mutants have a curled wing phenotype (123), hence the name *curled*. However, Cu function is relatively uncharacterized in *Drosophila*. There are three reported *cu* transcripts, RC, RD and RE (Fig. 3-1), among which the *cu*-RD isoform is rhythmically expressed in a set of *timeless* (*tim*)-expressing neurons in the adult brain. Consistent with this, knock down of *cu*-RD in these neurons altered light-mediated locomotor rhythm (251). *cu* transcripts are also induced by starvation, which demonstrate a 2.5-fold up-regulation after 6 hr of food deprivation (123). The Cu mammalian homolog, Nocturnin, also has a deadenylase-independent function. In cultured cells, Nocturnin promotes adipogenesis by facilitating translocation of the nuclear receptor PPAR γ into the nucleus (see section 3.3.3 for details) (252).

3.1.2 CCR4-NOT deadenylase complex and its components

In almost all biological systems, CCR4 works in the CCR4-NOT complex to catalyze deadenylation of mRNA, which contributes to mRNA decay (253,254). CCR4-NOT is a multi-component protein complex consisting of at least CCR4 (encoded by *twin* in *Drosophila*), Pop2 (also known as CAF1), NOT1, NOT2 (encoded by *Regena* in *Drosophila*), NOT3, NOT4 and CAF40 as the core subunits (255-259, Fig. 3-2). Other than Twin, there are three more CCR4

paralogs in higher eukaryotes including *Drosophila*, which are Nocturnin (Cu), Angel and 3635 (CG31759). Presumably any of the four proteins could be the actual CCR4 component in the CCR4-NOT complex. Twin and Cu can be co-immunoprecipitated with NOT1 protein, the scaffold component of the complex (260), but not Angel and 3635, which suggested that either Twin or Cu could be the actual CCR4 working in the CCR4-NOT complex but not Angel and 3635. There is direct evidence that *Drosophila* Twin, Pop2, NOT1, NOT2, and NOT3 are involved in mRNA poly(A) tail shortening *in vivo* during embryogenesis (257,261-265), while Cu was shown to have poly(A) shortening ability only *in vitro*. This fact suggested that Cu may not be an essential component of the CCR4-NOT complex like Twin and may only work in the complex in certain tissues or certain circumstances. Other than Twin or Cu, Pop2 also contains the catalytic activities for deadenylation, therefore Twin or Cu and Pop2 are the actual functional enzymatic components within the complex (266). Instead of having an EPP domain as in CCR4, Pop2 is a DEDD-type (Asp-Glu-Asp-Asp) nuclease.

3.1.3 Deadenylase and transcriptional repression

Eukaryotic mRNAs undergo a series of processing events including 5'-capping, splicing and 3'-polyadenylation. After export of mRNAs to the cytoplasm, the optimal translation of mRNA requires mRNA circularization (267). Circularization is achieved when the poly(A)-binding protein (PABP) binds to the tail while the Eukaryotic translation Initiation Factor 4E (eIF4E) binds to the cap. Then, the translation initiation factor eIF4G bridges between eIF4E and PABP to form the mRNA loop for translation (268,269). Therefore, shortening of the poly(A) tail by deadenylation affects the rate of circularization, which is an essential process of translational repression. Moreover, extensive deadenylation will eventually result in mRNA degradation. Some RNA-binding proteins, such as Nanos, Pumilio, Smaug and Bicaudal-C (Bic-C) as well as the microRNA-induced silencing complexes (miRISCs) can recruit the CCR4-NOT complex to specific target mRNAs for degradation (262-264,270,271). Apart from the CCR4-NOT complex, there are two other Poly(A)-specific ribonucleases, the Pan2–Pan3 complex and the poly(A)-specific ribonuclease (PARN), that are involved in mRNA deadenylation (272-274). Interestingly, it is thought that the deadenylation function carried out by the CCR4-NOT complex is not essential for viability in *Drosophila* (257), which is also the case in yeast, probably due to the presence of redundant deadenylase enzymes (275). However, the CCR4-NOT complex is essential for oogenesis and embryogenesis (263,276).

3.1.4 The CCR4-NOT deadenylase complex is multifunctional

Studies have shown that the CCR4-NOT complex was multifunctional and influences biological processes beyond translation and mRNA degradation. The CCR4-NOT complex has been inferred to both positively and negatively regulate transcription initiation, assist transcription elongation, and regulate both mRNA export and chromatin modifications (254,277-279), hence being a so-called “control freak” of mRNA regulation. First off, the CCR4-NOT components were found at sites of transcription and could be cross-linked to promoters (280-282). Loss of yeast CCR4-NOT function caused a dramatic redistribution of the general transcription factor TFIID on promoters, with reduced binding to highly expressed ribosomal genes promoters and induced binding to stress-responsive element (STRE)-controlled genes promoters (283). It has also been shown that overexpression of human CCR4 or CAF1 (Pop2) activated transcription of Estrogen Receptor (ER) responsive genes (284,285). Moreover, the CCR4–NOT complex was recruited to transcribed genes during elongation by RNA polymerase II (RNAPII) in yeast and was able to stimulate transcription elongation of arrested polymerases *in vitro* (279). However, other than its role as a deadenylase (255,257,261-264,276), the diverse functions of the CCR4–NOT complex have not been studied in *Drosophila* (286).

3.1.5 DHR4 and its role in regulating the ecdysone synthesis

In Chapter 1.4 and 1.11, I mentioned that DHR4 is an important target of PTTH/MAPK signaling, which is essential in regulating ecdysone production. Our lab showed that DHR4 oscillates between the nucleus and cytoplasm of PG cells and when DHR4 protein was in PG nuclei, it inhibited ecdysone biosynthesis by directly repressing the expression of *Cyp6t3*, an ecdysone biosynthetic gene. PTTH signaling appeared to promote translocation of DHR4 from the nucleus to the cytoplasm, thereby inactivating DHR4 and thus eliminating the repression of ecdysone biosynthesis. *DHR4^l* mutants showed precocious wandering behavior caused by precocious ecdysone pulses since DHR4 was not able to repress ecdysone biosynthesis in mutants at the developmental time when ecdysone titers should be low. This resulted in larvae that have shorter feeding times, giving rise to small pupae, which is consistent with the observation that larvae with hyperactivated PTTH/MAPK signaling [by expressing a constitutively active form of *Ras* (*Ras^{V12}*) specifically in the PG] also form small, precocious pupae (31).

3.2 Materials and Methods

3.2.1 Fly stocks and fly crosses

Drosophila melanogaster stocks were maintained on standard agar-cornmeal medium at 25°C. *phm22-Gal4* was obtained from Dr. O'Connor's lab (33) and the Gal4 line, *P0206-Gal4*, was previously described in Chapter 2.2.1 (31,287). UAS-*curled (nocturnin)*-RNAi strains (#45441, #45442 and #45443) were ordered from the VDRC stock center. UAS-*Pop2*-RNAi (#52947 and #30492), UAS-NOT1-RNAi (#32836) and UAS-NOT3-RNAi (#34966) were purchased from the Bloomington stock center.

3.2.2 Constructing an RNA interference-resistant form of *cu*-cDNA

The DNA sequence of the *cu*-RNAi target region (#45442) is 354 bp in length. 78 silent mutations were introduced into the RNAi target region to ensure that no more than 14 bp of continuous DNA was aligned between the RNAi target sequence and the RNAi-resistant cDNA sequence. The codon usage preference in *Drosophila* was taken into account according to Moriyama and Powell (288). Finally, the RNAi-resistant sequence was verified at the amino acid level using NCBI blastx to make sure that it encoded for the same protein sequence as Curled/Nocturnin. Additional DNA sequence flanking the RNAi regions was included, which then provided the restriction sites of BsrGI and XbaI near each end. The whole sequence of 574 bp was submitted to Biomatik for custom gene synthesis. The final DNA sequence was delivered in the pBMH vector. Meanwhile, the original *cu* cDNA clone (RE6512) was ordered from the *Drosophila* Genomics Resource Center (DGRC) and the RNAi target region was swapped with the RNAi-resistant sequence by digestion with BsrGI and XbaI restriction enzymes and followed by conventional T4 ligation (invitrogen#15224017). The modified cDNA was subsequently subcloned into the pUAST vector at the KpnI and NotI sites. Finally, transgenic flies were generated via P-element mediated transformation through the service from Bestgene Inc.

3.2.3 Measurement of pupal body mass

Eggs were laid on grape juice plates with a supplement of yeast paste at the center. After 2 hr of collection, eggs were transferred to standard agar-cornmeal medium and reared at 25°C for 5 days. On day 5, the already pupariated animals were cleared away and larvae were allowed to continue to grow. After 10 hr, the newly formed pupae were used for measurements.

12 pupae were randomly sampled from each biological replicate. Pupae were washed in distilled water, dried on a KimWipe. The same 12 pupae were weighed together on a scale for three times and the average reading was taken. Mass per pupa was then calculated. Four biological replicates were carried out for each condition/genotype.

3.2.4 Developmental timing measurement

Flies were allowed to lay eggs twice on the grape juice plates supplemented with yeast paste for 2 hr prior to the actual collection in order to reduce egg retention. Then eggs were collected within a 2 hr time window and were then transferred to standard agar-cornmeal medium and reared at 25°C. After 5 days, pupariation was scored at 2-hr intervals.

3.2.5 qPCR analysis

Tissue RNA extraction and cDNA synthesis were carried out as described in Chapter 2.2.4. Primer validation and calculation of fold changes were also done following the same protocol as in Chapter 2.2.4 expect that the reactions were performed in the StepOne Real-Time PCR System. All primers used for the *cu* (*nocturnin*) and *Pop2* study are listed in Table 3-1.

Table 3-1. qPCR primers used in *cu* and *Pop2* studies

Primer Name	Primer Sequence
<i>cu-common</i> _#86_F	AAGCTGAATTCCGTGTCGAG
<i>cu-common</i> _#86_R	GCCCGAGAGTTTGTGATAGG
<i>cu-RD</i> _#20_F	TCACCATGAGGATGATGGAA
<i>cu-RD</i> _#20_R	GGCAATTCGAACAGGGTATCT
<i>neverland</i> _F	CCCTCACCTAGGAGCCAAC
<i>neverland</i> _R	GGCATATAACACAGTCGTCAGC
<i>spookier</i> _F	GCGGTGATCGAAACAAC
<i>spookier</i> _R	CGAGCTAAATTTCTCCGCTTT
<i>shroud</i> _F	CGAATCGCTGCACATGAC
<i>shroud</i> _R	TAGGCCCTGCAGCAGTTTAG
<i>phantom</i> _F	GGCATCATGGGTGGATT
<i>phantom</i> _R	CAAGGCCTTTAGCCAATCG
<i>disembodied</i> _F	GTGACCAAGGAGTTCATTAGATTTC
<i>disembodied</i> _R	CCAAAGGTAAGCAAACAGGTTAAT

<i>shadow_F</i>	CAAGCGGATATTTGTAGACTTGG
<i>shadow_R</i>	AAGCCCACTGACTGCTGAAT

3.2.6 RNA-Seq analysis

RNA-Seq was performed as described in Chapter 2.2.15 expect that cDNA libraries were constructed using the Encore Complete RNA-Seq DR Multiplex System (Nugen #0333-32 and #0334-32), in which the DNase digestion during RNA extraction was not required and at least 100 ng of total RNA was needed as input. Lastly, there were no Insert-Dependent Adaptor Cleavage (InDA-C) technologies included at that time to further eliminate the abundant unwanted transcripts, such as ribosomal RNA.

3.2.7 Sterol rescue by feeding

Every 17.8 g of Nutri-Fly™ Bloomington Formulation mixture was dissolved in 100 ml of distilled water. The mixture was stirred and brought to boil, then immediately removed from the hot plate and let simmer for 10 minutes when it was still hot. After that, the mixture was cooled down to below 50°C and before it solidified, 450 µl of propionic acid was added per 100 ml of food. Finally, the food was supplied with various sterols dissolved in ethanol to a final concentration of 75 µg/ml for cholesterol, 75 µg/ml for 7-dehydrocholesterol (7DC) and 200 µg/ml for 20E with a final concentration of 2% ethanol. The control food was prepared in a similar manner with only a final concentration of 2% ethanol. Sixty embryos were transferred to each vial, allowed to develop at 25°C and the phenotypes of the larvae were scored.

3.3 Results and Discussion

3.3.1 Curled (*cu*) has a role in the PG during larvae-pupae transition

To study the function of *cu* in the PG, we knocked down *cu* in a PG-specific manner using *phm22-Gal4* and found it caused a small pupae phenotype (Fig. 3-3A). However, the RNAi animals are viable and will eclose as slightly smaller adults. One possible explanation for the small pupae phenotype is that animals have accelerated developmental processes. Since the main purpose of the larval stages is to feed and grow, the animals with accelerated developmental processes will have shorter feeding times, therefore, give rise to small pupae. I tested whether the small pupae phenotype was indeed caused by accelerated larval development by observing the developmental timing of homozygous *phm22-Gal4*, *UAS-cu*-RNAi animals and the *phm22-Gal4*

line as the control. For this, eggs were collected strictly in 2-h intervals and the time of pupariation was then scored 5 days after. As shown in Fig. 3-3B, homozygous *phm22-Gal4, UAS-cu*-RNAi animal pupariated early by 13 hrs compared to controls (~130 hr After Egg Deposition, AED, compared to ~143 hr AED in controls when 50% of the population form pupae), which makes up about 10% of the time from embryos to pupae (Fig. 3-3C). *phm22-Gal4, UAS-cu*-RNAi pupae also had reduced body mass by ~13% compared to controls ($p < 0.05$) (Fig. 3-3B). This further demonstrated that the feeding period of RNAi larvae was probably shortened and they had precocious wandering.

Since the *phm22-Gal4* driver also has some expression in the fat body, I wanted to confirm that the RNAi phenotype was not an effect of *cu* being disrupted in the fat body. Therefore, I used another driver *P0206-Gal4* (highly specific to the RG; the PG is part of the RG) to repeat the timing assay. My result showed that *cu*-RNAi animals still pupariate earlier (Fig. 3-3D), but this time only preceding the control by ~6 hrs. Given that P0206 is a weaker driver compared to *phm22* the results appear to be consistent. In summary, disrupting *cu* in the PG caused larvae to develop faster, resulted in a smaller final body/pupal size. This is potentially due to the occurrence of precocious ecdysone pulses. In the future, the ecdysone titer should be directly measured to show that *cu*-RNAi animals have elevated or precocious ecdysone pulses.

3.3.2 Overexpression of *cu* in the PG caused developmental delay

Originally, I constructed an RNAi-resistant version of *cu*-cDNA (I here refer to as *cu*-RcDNA) to rescue the *cu*-RNAi phenotype as a way to show that the phenotype we observed in *cu*-RNAi was not caused by off-target effect (referring to methods in section 3.2 for detailed procedure). The logic behind that was that if the RNAi construct was knocking down the anticipated target (*cu*) as well as an unknown off-target at the same time, and if the phenotype was caused by the off-target, introducing the native cDNA might show false rescue because the cDNA, after being expressed, might just titrate siRNAs from the off-target, thus alleviating the phenotype (289). Introducing the RNAi-resistant version of the cDNA, however, would not take away the siRNAs, but still produce a functional Cu protein. In this case, if I see a rescue of the PG>*cu*-RNAi phenotype, I would be confident that the phenotype was caused solely by the loss of the indented target (*cu*), but not an off-target.

Unfortunately, overexpressing *cu*-RcDNA in the PG (using the *phm22*-Gal4 driver) itself resulted in larvae developmental arrest and lethality (Fig. 3-4), and is therefore not usable for the rescue experiment. More specifically, most of the animals were arrested in L2 stage with some escapers (~2%) reaching the adult stage. Some of the arrested L2 emerged from the food and formed L2 prepupae (~50%) (Fig. 3-4A). Developmental arrests can be caused by lack of ecdysone pulses where animals cannot get the signal to progress to the next stage. Especially in the case of L2 prepupae when larvae forego the molt to an L3 and directly molt from an L2 to a prepupa. The L2 prepupae phenotype is relatively rare and has been only associated with mutations in *E75* (290), *dre4* (291), *itpr* (292), and *Cyp6t3* (31), all of which have dramatically reduced ecdysone levels. These results agree with the idea that loss of *cu* in the PG will lead to the occurrence of precocious ecdysone pulses (high ecdysone) while overexpressing *cu* in the PG reduced ecdysone levels. The phenotype of *cu*-RcDNA overexpression is also dose-dependent since with a weaker driver (*P0206*), animals formed bigger pupae with prolonged L3 stage (100%) and could eclose as adults (Fig. 3-4), while developmental delay and giant pupae are other very common phenotypes observed in animal with low ecdysone levels. Moreover, I confirmed that overexpressing a GFP-tagged native *cu*-cDNA (123) using the *phm22*-Gal4 driver also resulted in 100% L3 arrest (data not shown), suggesting the *cu*-RcDNA is working properly. In other words, the normal function of *cu* in the PG is to negatively regulate (suppress) ecdysone synthesis.

As for the concern of off-target effects, I still do not have evidence to rule out that the RNAi phenotype was not caused by an off-target effect. Nevertheless, I showed via qPCR that the *cu* transcripts are indeed downregulated by ~3.5 fold in the *phm22*>*cu*-RNAi animals (Fig. 3-5A). Moreover, when Cu function was disrupted in the whole body using *actin5c*>*cu*-RNAi (the same construct that gave the small pupae phenotypes using PG driver), it recapitulated the null mutant phenotype (curled wings) (123) (Fig. 3-5B), suggesting that the RNAi construct was indeed interfering with *cu* function, and was otherwise viable.

3.3.3 The genetic interaction between *cu* and *DHR4*

Notably, the small pupae caused by PG>*cu*-RNAi resemble the *DHR4* mutant phenotypes, in both cases the phenotypes arise from developmental acceleration (Fig. 3-3A, 31). Likewise, *cu*-cDNA overexpression in the PG phenocopied *DHR4* overexpression, with both cDNAs blocking the molts (Fig. 3-4A, 31), suggesting Cu and *DHR4* could possibly function in the same process to regulate ecdysone production.

Recent studies showed that the mammalian homolog of Curled, Nocturnin, also had a non-canonical function other than being a deadenylase. In cultured cells, Nocturnin promotes adipogenesis by increasing the nuclear translocation of PPAR γ , which like DHR4, is a member of the nuclear receptor family, in a deadenylase-independent manner (252). More intriguingly, MAPK signaling cascade controls the nucleo-cytoplasmic shuttling of PPAR γ (293) just like DHR4, which is only the second example where MAPK regulates the nuclear translocation of a nuclear receptor. MAP kinase attenuates PPAR γ 's transactivation function in the nucleus by keeping PPAR γ 's in the cytoplasm, which is exactly the case for DHR4. I hypothesize that *Drosophila* Cu (*aka* Nocturnin) may also help DHR4 translocate to the nucleus in PG cells, similar to the vertebrate Nocturnin/PPAR γ system. My general working model is that in PG cells, DHR4 is able to repress ecdysone biosynthesis when it is in the nucleus. Upon binding to its receptor Torso, PTHH signal will activate the Ras/Raf/MAPK pathway and transfer DHR4 from the nucleus to the cytoplasm to derepress ecdysone biosynthesis at the developmental time points that require high levels of ecdysone (Fig. 3-6). Cu could function as a negative regulator of ecdysone production by promoting the nuclear translocation of DHR4, which would then decrease ecdysone production to terminate a pulse.

To test the hypothesis, I first repeated the results that either RG-*cu*-RNAi or *DHR4*-RNAi caused developmental acceleration (31) and that developmental timing is comparable between the two lines (Fig. 3-7A). RNAi of both *cu* and *DHR4* resulted in more profound effects than single *DHR4*-RNAi (Fig. 3-7B), which could be explained by the fact that DHR4, having already compromised level, could not translocate into the nucleus to exert its function, thus worsened the phenotype.

If Cu function was indeed to help DHR4 translocate to the nucleus to repress ecdysone biosynthesis, we would expect that knocking down *cu* in the PG would rescue the L1 arrest caused by PG>*DHR4* overexpression. Therefore, I did genetic epistasis analysis by comparing developmental progression between the PG>*DHR4*-cDNA larvae and the larvae with PG>*DHR4* overexpression in the *cu*-RNAi background. As expected with *DHR4* overexpression in the PG, only ~3% of the L1 could proceed to the L2 stage. The L1 arrest could be partially rescued by *cu*-RNAi where ~19% of the L1 molted to L2 ($p < 0.05$, Fig. 3-8A). I also confirmed this result by overexpression of *DHR4* in a *cu* mutant background. Here we used the weaker RG driver, *P0206*, because the chromosome location makes it easy to combine the driver and the *DHR4*-cDNA

together in the *cu* mutant background. *P0206>DHR4*-cDNA caused L3 larval arrest with only ~5% of the L3 larvae forming pupae. This L3 arrest phenotype was significantly ($p<0.001$) rescued when the RG-specific *DHR4* cDNA expression was driven in the *cu* mutants background, in which ~83% of the L3 formed pupae (Fig. 3-8B). In the reciprocal approach, *DHR4*-RNAi could also partially rescue the developmental arrest caused by *cu*-RcDNA overexpression (Fig. 3-8C). In more detail, when *cu*-RcDNA was overexpressed in PG cells, larvae were developmentally stuck as L2; however, when *DHR4*-RNAi was introduced, ~20% of the L2 was now able to molt to L3, which was a significant increase ($p<0.05$). These results suggested that *cu* and *DHR4* interact genetically. However, even if Cu and DHR4 work in two independent pathways to regulate the ecdysone pulses, it would still be possible to observe these above-mentioned genetic results.

An additional experiment to establish whether DHR4 and Cu act in the same pathway is based on previous work from our lab, where it was established that an ecdysteroidogenic gene called *Cyp6t3* is a target gene of DHR4. When DHR4 is in the nucleus, *Cyp6t3* is repressed, likely directly by DHR4, thereby throttling ecdysone production (31). According to our model, without Cu, DHR4 cannot move efficiently into the nucleus, and one would predict that *Cyp6t3* would be derepressed. Therefore, I measured the transcript levels of *Cyp6t3* in RGs at 8 hrs after the L2/L3 molt, a time point when DHR4 is in the nucleus and *Cyp6t3* is repressed under the normal conditions. I found that *Cyp6t3* was indeed ~2 fold (and significantly) upregulated in *PG>cu*-RNAi RGs compared to controls (Fig. 3-9). Considering that the *PG>cu*-RNAi animals were developmentally accelerated, I next repeated the experiments by staging larval populations using the blue gut method (183) instead of by the actual time spent feeding, which compensates the difference of developmental timing between the RNAi and control animals. Hence, I collected RGs from the wandering blue gut larvae (represent a developmental stage that is about at least 12 hr before pupa formation) and I saw a trend of *Cyp6t3* upregulation in *cu* knockdown samples; however, this was not statistically significant (Fig. 3-9). One reason might be that, although the populations were developmentally synchronized between the RNAi and controls, the wandering blue gut stage was not a sharp developmental window, which probably spans a time of ~6 hours (183). This probably caused some variations in *Cyp6t3* expression within the population, leading to the poor statistics. Nevertheless, the same trend of upregulation was observed at both time points. Furthermore, as mentioned previously, the L2 prepupae phenotype observed in *PG>cu*-

overexpression (Fig. 3-4) is reminiscent to the *Cyp6t3* loss-of-function phenotype. Collectively, these data suggested that a genetic interaction between *cu* and *DHR4* does exist. In the future, one can test whether there is also a physical interaction between Cu and DHR4 via co-immunoprecipitation (co-IP); probably using *Drosophila* S2 cells as a start.

Since DHR4 oscillation between nucleus and cytoplasm is controlled by PTTH/MAPK in PG cells (Fig. 3-6) (31), I wondered whether *cu* is also a target of PTTH/MAPK signaling, where PTTH/MAPK represses *cu* function, thus keeping DHR4 in the cytoplasm to allow ecdysone biosynthesis. First of all, I tested whether *cu* is transcriptionally repressed by PTTH/MAPK signaling by carrying out qPCR analysis. For this, total RNA was extracted from the brain-RG complexes of the larvae with loss of PTTH/MAPK signaling in the PG (*phm22>torso*-RNAi; Torso is the receptor of PTTH) at two different developmental time points, which represent the stage with low (18 hours before pupal formation, BPF) and high (8 hours before pupal formation) ecdysone levels. Here I specifically looked at the isoform RD (primer pair *cu*-RD_#20 in Table 3-1) because the expression of isoform D in the RG was the only one that fluctuated at the four developmental time points we tested in the RG microarray (Fig. 3-1), suggesting its levels are dynamically regulated. The expression of *cu*-RD declined at 36 hr L3 when the PTTH signaling and ecdysone level in the animals start to increase (Fig. 3-1) (31); therefore representing an isoform that might be controlled by PTTH signaling. As shown in Fig. 3-10, when *torso* was disrupted, *cu*-RD was significantly upregulated at both two time points compared to *phm22>w¹¹¹⁸* controls which suggested that *cu*-RD transcript levels are regulated by Torso. However, a flaw of this approach is that in controls, if PTTH/MAPK signaling can suppress *cu* expression, the *cu*-RD level at 8 BPF should be lower compared to that at 18 BPF since 8 BPF is the developmental stage with high ecdysone titer (hence high PTTH/MAPK signaling). One possible explanation is that there is also *cu*-RD expression in the brain (251), therefore, the qPCR result does not represent the *cu*-RD levels in the RG alone. The same qPCR analysis could be performed using RG samples instead of the brain-RG complexes to obtain more conclusive results. Nevertheless, my results are consistent with the idea that isoform *cu*-RD might be regulated by PTTH/MAPK pathway and could be the isoform that interacts with DHR4 in repressing ecdysone biosynthesis.

3.3.4 The CCR4-NOT complex is essential for ecdysone production

Next, I asked whether Cu still has the canonical function as the deadenylase in the PG, in other words, whether Cu functions in the CCR4-NOT complex. Also, is the CCR4-NOT complex

essential for regulating ecdysone production in the PG? With this in mind, I went back to our genome-wide PG>RNAi results (120,208) and found that PG-specific knock-down of most of the components in the CCR4-NOT complex (including *CCR4 (twin)*, *Pop2*, *NOT1* and *NOT4*) resulted in L3 arrest, a common phenotype caused by ecdysone deficiency, suggesting that the deadenylase complex has a crucial role in ecdysone production (Table 3-2). PG>*twin*-RNAi, the CCR4 paralog like *cu*, also caused developmental arrest as seen in the knock-down of other components (Table 3-2), while PG>*cu*-RNAi caused developmental acceleration, which suggested that *cu* does not likely function in the CCR4-NOT complex in the PG. Nonetheless, it seems that CCR4-NOT complex is essential for regulating the production of ecdysone.

Table 3-2. PG-specific RNAi of CCR4-NOT complex components resulted in ecdysone related phenotypes

Gene name	PG-specific RNAi phenotype	Tested RNAi lines with phenotype	Tested RNAi lines without phenotype
<i>CCR4 (twin)</i>	L3 arrest	V104442	BL32490, BL32901
<i>CCR4 (cu)</i>	developmental acceleration	V25176, V45441, V45442, V45443	V109759
<i>Pop2</i>	L3 arrest	BL52947, BL30492	n.a.
<i>NOT1</i>	L3 arrest	V106587, BL32836	BL31696, BL28681
<i>NOT2 (Regena)</i>	no	n.a.	V20826, BL35460, BL57549
<i>NOT3</i>	L3 arrest	V105990, BL34966*	BL33002
<i>NOT4</i>	L3 arrest	V110472	n.a.
<i>CAF40 (Rcd-1)</i>	no	n.a.	V101462

phm22-Gal4 was used for the knocking down of all genes.

V: VDRC transgenic lines.

BL: transgenic lines from Bloomington stock center

n.a: not applicable.

*Knocking down *NOT3* in the PG using the transgenic line BL34966 resulted in pupal lethality.

3.3.5 RNA-Seq analysis of PG>*Pop2*-RNAi

Since CCR4-NOT complex is involved in mRNA deadenylation and decay (253), I expected that the mRNA homeostasis would be affected when the CCR4-NOT function was disrupted in the PG. Furthermore, the CCR4-NOT complex was shown to be multifunctional in yeast, which also included regulating transcription (254,277-279). However, the diverse functions

of this complex are less studied in *Drosophila* (286), except for its deadenylase aspect (255,257,261-264,276). Interestingly, it was thought that the deadenylation function carried out by CCR4-NOT complex is not essential for viability in *Drosophila* (257), which is also the case in yeast (275). But knocking down components of the CCR4-NOT complex in the PG affected development, suggesting that the CCR4-NOT complex could have crucial functions in the PG. To identify mRNAs that are dependent on CCR4-NOT activity in the PG, I carried out RNA-Seq analysis comparing the transcription profile of PG>CCR4-NOT loss-of-function RGs to that of controls. Specifically, PG>*Pop2*-RNAi animals were chosen because: 1) only CCR4 (*twin*) and Pop2 have the catalytic activities for deadenylation, therefore probably representing the functional components within the complex instead of being just a scaffold component (275). 2) There are already two independent RNAi lines available (#52947 and #30492 from Bloomington stock center) that gave the L3 arrest phenotype, which ensured that *Pop2* has an important role in the PG. The RNAi line #52947 was used for the RNA-Seq analysis and RGs were collected at 24 hr after L2 to L3 molt for both the PG>*Pop2*-RNAi as well as controls (PG>*w¹¹¹⁸*).

3.3.5.1 Overall RNA-Seq quality

RNA-Seq analysis of *Pop2*-RNAi was carried out before the *sna* project. At that time, the cDNA construction kit did not yet include the InDA-C technology to further eliminate rRNA (see Chapter 2.2.15 for details), which was later used for the *sna* project. To see how much the InDA-C step had improved the RNA sequencing reads, I examined the total reads and calculated the percentage of over-presented reads for both projects using the “NGS: QC and manipulation” in the web-based bioinformatic platform galaxy (<https://usegalaxy.org/>). I found that with the original strategy, on average $\sim 26\% \pm 5.4\%$ of the total reads per sample were the overrepresented sequences (*aka* rRNA) in the earlier experiments, leaving 9.6 ± 1.3 million (M) total informative reads per sample (Fig. 3-11). In contrast, the new strategies resulted in, on average, $96\% \pm 2.4$ of the total reads per sample are informative, which means 16.7 ± 4.4 M informative reads (Fig. 3-11). This result showed that the InDA-C step used in *sna* RNA-Seq analysis (see Chapter 2.2.15 and 2.3.11) is indeed working as well as advertised by the manufacturer. Normally, 10 M reads per sample are needed to achieve a similar quality as microarray analysis (<https://genohub.com/next-generation-sequencing-guide>), which means our *Pop2*-RNAi RNA-Seq analysis was marginally suboptimal (with 9.6 ± 1.3 M informative reads). However, the already obtained data is still usable, although low-expressed genes were either not detected or the fold

change for low-expressed genes might not be reliable. However, the low-expressed genes in RGs might not be functionally important for RGs in the first place.

3.3.5.2 PG>Pop2-RNAi ring glands have altered mRNA translation machinery

The fold change of expression for each transcript in RNAi animals was determined using the same method as mentioned in Chapter 2.3.11. I first listed the top 50 downregulated genes as well as the top 50 upregulated genes in the PG>Pop2-RNAi samples in Appendix (Table A-3).

Downregulated genes and GO term enrichment analysis

A total of 204 genes were identified as more than 3-fold downregulated and 338 genes were more than 3-fold upregulated in the PG>Pop2-RNAi RGs when compared to controls. Generally, I noticed that the number of upregulated genes was consistently greater than that of downregulated genes when various cut-offs were used (Table. 3-3), demonstrating that the function of *Pop2* was mainly to negatively regulate mRNA levels, thus likely still function in part in deadenylation and mRNA decay. Next, I examined whether differentially expressed genes had distinct patterns, i.e. whether they fell into similar functional groups. I carried out the term enrichment analysis and generated the functional association networks using DAVID GO and STRING as described in Chapter 2.3.11.

Table 3-3. Number of genes affected by PG>Pop2-RNAi using various fold change cut-offs

Fold change	#of downregulated genes	# of upregulated genes
>5	88	95
>3	204	338
>2.5	242	549
>2	304	965

Among the >3-fold downregulated genes, the Protein-protein interaction (PPI) enrichment *p* value was 3.3E-16, which means those genes have more interactions among themselves at the protein level than what would be expected for a random set of proteins of similar size, drawn from the genome. I observed several protein-protein interaction units within the big network, suggesting genes within the same interaction unit contributing to a common biological purpose. According to the gene ontology (GO), these functional interaction groups were “cytoplasmic translation”/“ribosome”, “mitochondria”, “mRNA splice site selection”/“regulation of

transcriptional start site selection at RNA polymerase II promoter”, “actin filament organization” and “proteolysis” (Fig. 3-12). According to DAVID GO, all these terms have a *p* value less than 0.05, however, with the FDR greater than 0.05 (FDR<0.05 is more stringent than *p*<0.05, see Chapter 2.3.11). Next, I expanded my list to >2-fold downregulated gene sets and carried out the GO enrichment analysis via both DAVID GO and STRING database. The enriched terms from both databases were combined and listed in Table 3-4. It turned out that the term “ribosome” and “mitochondrial protein complex” now became significantly overrepresented (FDR< 0.05, Table 3-4). In summary, my transcriptome analysis of PG>*Pop2*-RNAi suggested that the CCR4-NOT complex or at least *Pop2* itself is related to the negative regulation of mRNA translation and mitochondria function.

Table 3-4. Go term enrichment analysis results for >2-fold downregulated genes in PG>*Pop2*-RNAi ring glands

Category	Pathway ID	Pathway description	FDR	Gene names
GO_biological process	GO:0002181	Cytoplasmic translation	0.015	<i>RpL19, RpL26, RpL34a, RpL37A, RpL8, RpS10b, RpS12, RpS13, RpS26, RpS29, RpS9</i>
GO_cellular component	GO.0022626	Ribosome	0.011	<i>RpL19, RpL26, RpL34a, RpL37A, RpL8, RpS10b, RpS12, RpS13, RpS26, RpS29, RpS9</i>
GO_cellular component	GO.0098798	Mitochondrial protein complex	0.047	<i>CG10219, CG3621, CG7834, CG8199, CG9603, CoVIII, CoVIb, Tim8, levy, mge, mt:ATPase8</i>
GO_cellular component	GO.0005751	Mitochondrial respiratory chain complex IV	0.050	<i>CG9603, CoVIII, CoVIb, levy</i>

FDR: False discovery rate

Upregulated genes and GO term enrichment analysis.

For the >3-fold upregulated genes, several functional groups were also identified, including “RNA polymerase”, “mitochondrial translation”, “intracellular cholesterol transport”, “proteolysis”, “lipid particle”/ “Lipid catabolic process”, “Glycine/serine/threonine metabolism”, and “Galactose metabolism” (Fig. 3-13). All the above-mentioned terms have a *p* value<0.05, but FDR>0.05. Next, I expanded the gene list to >2.5-fold upregulated for the GO enrichment term analysis and the results listed in Table 3-5 show the enriched GO terms as well as the KEGG

pathways with the stringent FDR<0.05 cut off. It appears that genes encoding for RNA polymerase subunits, mitochondrial ribosomal proteins and mitochondrial respiratory chain complex III were significantly overrepresented.

Table 3-5. Go term enrichment analysis results for >2.5-fold upregulated genes in PG>Pop2-RNAi ring glands

Category	Pathway ID	Pathway description	FDR	Gene names
GO_biological process	GO:0032543	mitochondrial translation	0.002	<i>CG12848, CG15390, mRpL15, mRpL20, mRpL27, mRpL30, mRpL32, mRpL33, mRpS18A, mRpS21, mRpS23, mRpS25, mRpS31, mRpS33, mRpS34</i>
GO_cellular component	GO.0005750	mitochondrial respiratory chain complex III	0.012	<i>UQCR-11, UQCR-14, UQCR-6.4, UQCR-Q, ox</i>
GO_cellular component	GO:0005730	nucleolus	0.034	<i>Bka, CG11030, CG11563, CG11583, CG17652, CG2260, CG6712, CG7006, CG8414, CG9004, CG9246, Mys45A, TAF1B, bys, l(3)07882, tbrd-2</i>
GO_cellular component	GO.0005956	protein kinase CK2 complex	0.023	<i>CG33237, CG33238, CG33239, CG33242, Ste:CG33239, Ste12DOR</i>
GO_cellular component	GO.0070013	intracellular organelle lumen	0.011	50 genes, not shown
GO_cellular component	GO.0005576	extracellular region	0.040	38 genes, not shown
GO_molecular function	GO.0008970	phosphatidylcholine 1-acylhydrolase activity	0.025	<i>CG6277, CG6283, CG6295, CG6296, CG8552</i>
KEGG	3020	RNA polymerase	0.010	<i>CG33051, Rp112, Rp118, Rpb10, Rpb11, Rpb12, l(2)37Cg</i>
KEGG	260	Glycine, serine and threonine metabolism	0.050	<i>CG10184, CG11236, CG3999, CG6188, CG6415, ppl</i>

FRD: False discovery rate

In summary, based on my transcriptional data, disrupting the CCR4-NOT complex component *Pop2* in the PG repressed the expression of genes important for cytoplasmic translation but increased the expression of genes related to mitochondrial translation. In the future, one should confirm that the synthesis of mitochondrial proteins is indeed elevated while the synthesis of cytoplasmic proteins is reduced in PG>*Pop2*-RNAi RGs. Another situation when the cytoplasmic translation was repressed while the mitochondrial translation was stimulated was when human embryonic kidney (HEK) cells encountered amino acid starvation (294). However, it remains unclear as what the actual biological meaning of the changes in the expression of translation-related genes when *Pop2* is knocked down in the PG. Moreover, several genes encoding for RNA polymerase subunits had increased expression, which includes two out of 12 genes (*Rpb10* and *Rpb12*) encoding RNA polymerase II subunits, one gene (*CG33051*) encoding a polymerase III subunit as well as two other poorly characterized genes (*l(2)37Cg* and *Rpl12*), which are also predicted to encode RNA polymerase subunits. Furthermore, the expression of *TfIIIF β* , the gene encoding the general transcription factor II F β subunit, was induced in PG>*Pop2*-RNAi samples, suggesting misregulation of transcription initiation. As mentioned previously, loss of yeast CCR4-NOT function causes a dramatic redistribution of the general transcription factor TFIID on promoters, with reduced binding to the highly expressed ribosomal genes promoters and induced binding to stress-responsive element (STRE)-controlled genes promoters (283). My data from *Drosophila* PG also suggest that loss of CCR4-NOT function might affect transcription initiation. Hence, *Drosophila* CCR4-NOT complex also has functions other than its deadenylase activity. Lastly, apart from the induced mitochondrial translation, the expression of some components in mitochondrial complex IV (*levy*, *CG9603* and *CoVIII* in Fig. 3-12) were downregulated while some of the components in mitochondrial complex III were induced (Table. 3-5) in *Pop2* loss-of-function, suggesting dysfunction of the mitochondrial respiration chain and misregulation of energy production. My RNA-Seq result is the first evidence that CCR4-NOT complex is linked to mitochondrial function.

3.3.5.3 Loss-of-*Pop2* in the PG affected expression of ecdysone biosynthetic genes

Despite the possible global effect caused by *Pop2*-RNAi on transcription initiation and mRNA translation, the L3 arrest phenotype (a typical phenotype due to ecdysone deficiency) observed in the PG>*Pop2*-RNAi might be more directly caused by a reduction in the expression of the three ecdysone biosynthetic genes, i.e. *nvd*, *spok* and *sad* (Fig. 3-12). When I tested their

expression using another independent RNAi line (Bloomington #30492) at a different time point (0 hr L3) via qPCR, the expression of *nvd* and *spok* was consistently low (Fig. 3-14).

Nvd catalyzes the first step of ecdysone biosynthesis where it converts cholesterol to 7-dehydrocholesterol (7DC) (24). Spok plays a crucial role in the “Black Box” which consists of several not yet characterized conversion steps, however, the net result is that 5 β -ketodiol (5 β kd) is derived from 7-dehydrocholesterol (7DC) (30). To test our hypothesis that ecdysone deficiency phenotype in PG>*Pop2*-RNAi was caused directly by reduced levels of *nvd* and *spok*, I tried to supply the animals with steroid precursors after the Nvd and Spok enzymatic steps. In theory, 5 β -ketodiol (5 β kd) should be used since it is the precursor more downstream of Spok, however, at that time I did not have access to 5 β kd and there was only 7DC available in the lab. Surprisingly, 7DC alone was able to rescue the arrested L3 animals to the pupal stage. However, it was not sufficient to help the animals to proceed to adult stage (Fig. 3-15), probably because the conversion step from 7DC to 5 β kd was still problematic. Therefore, it is crucial to obtain 5 β kd and complete the rescue experiments in the future. The rescue of L3 arrest phenotype by 7DC was observed in both PG>*Pop2*-RNAi lines (Fig. 3-15), confirming the fidelity of the results. More intriguingly, unlike 7DC, providing animals with the biologically active form of ecdysone-20E did not rescue. Moreover, adding 7DC together with 20E also lowered the rescue effect of 7DC. These results suggested that lacking 7DC and possibly also 5 β kd due to misregulation of *nvd* and *spok* were the more direct cause of the L3 arrest phenotype, while providing 20E to the animals may have some toxic effect on development.

Finally, three genes in Niemann-Pick disease type C (NPC) family were upregulated in the *Pop2*-RNAi RGs (Fig. 3-12). These genes are involved in "intracellular cholesterol transport", although this GO term was not significantly overrepresented (P=0.024, FDR=0.63), however, this indicated that there might be an issue with cholesterol intake into the PG. Therefore, I also carried out rescue experiments with cholesterol and it did rescue the RNAi animals to the pupae stage to a similar degree as 7DC (Fig. 3-15). In conclusion, loss-of-*Pop2* or possibly the function of CCR4-NOT complex in the PG specifically reduced the expression of two ecdysone biosynthetic genes *nvd*, and *spok* as well as genes related to cholesterol transport.

3.4 Conclusion and significance

My data suggests that the function of CCR4-NOT deadenylase complex is essential in the PG for ecdysone biosynthesis. The predicted deadenylase Cu (*aka* Nocturnin) might not function in the complex in the PG. Instead, it may have the non-canonical function to help nuclear receptor DHR4 to translocate into the nucleus to repress ecdysone production at developmental time points when ecdysone levels need to be low. However, biochemistry experiments need to be done in the future to further show the physical interaction between Cu and DHR4.

Moreover, my results for the first time provided *in vivo* evidence that CCR4-NOT has functions beyond its mRNA deadenylase activity and is probably a multifunctional protein complex in *Drosophila* as in the case of yeast. The effect of CCR4-NOT loss-of-function in the PG is likely to be a global one instead of just targeting a certain group of genes since the expression of several genes encoding for RNA polymerases subunits are affected as well as the expression of both mitochondria and cytoplasmic ribosomal genes changed. This is the first time where the consequences of CCR4-NOT loss-of-function was examined through transcriptional profiling *in vivo* in a *Drosophila melanogaster* tissue.

3.5 Figures

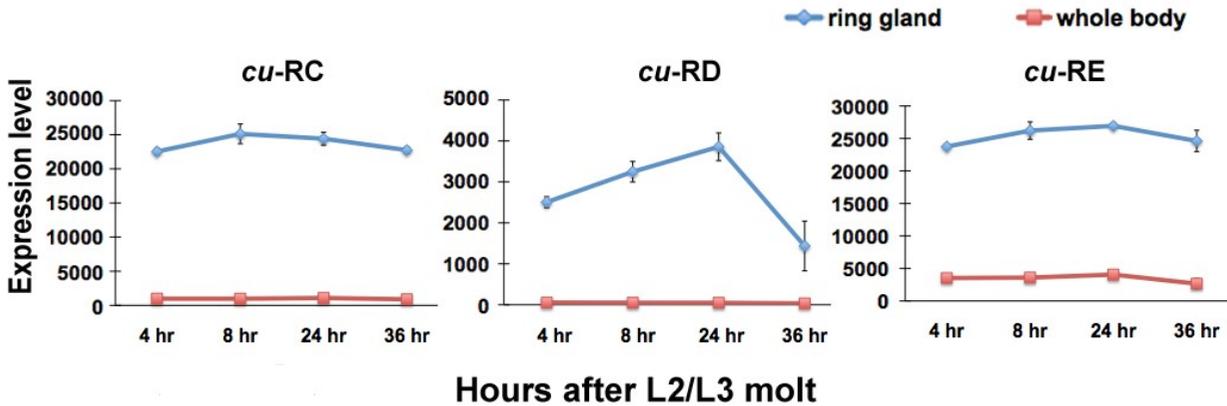


Figure 3-1. The *cu* expression profile in 3rd instar larval ring glands revealed by microarray analysis.

Microarray analyses were carried out with mRNA extracted from ring glands compared to mRNA extracted from whole larvae at four developmental time points in L3 (4, 8, 24 and 36 hr after the molt to L3). The results showed that all the three *cu* (*nocturnin*) isoforms (i.e. *cu-RC*, *cu-RD* and *cu-RE*) are enriched in the ring gland, especially *cu-RC* and -*RD*, which are >20-fold enriched. However, as a circadian output gene, only the mRNA level of *cu-RD* fluctuated at the four time points we examined in the ring gland, consistent with the idea that the RD isoform is under circadian control.

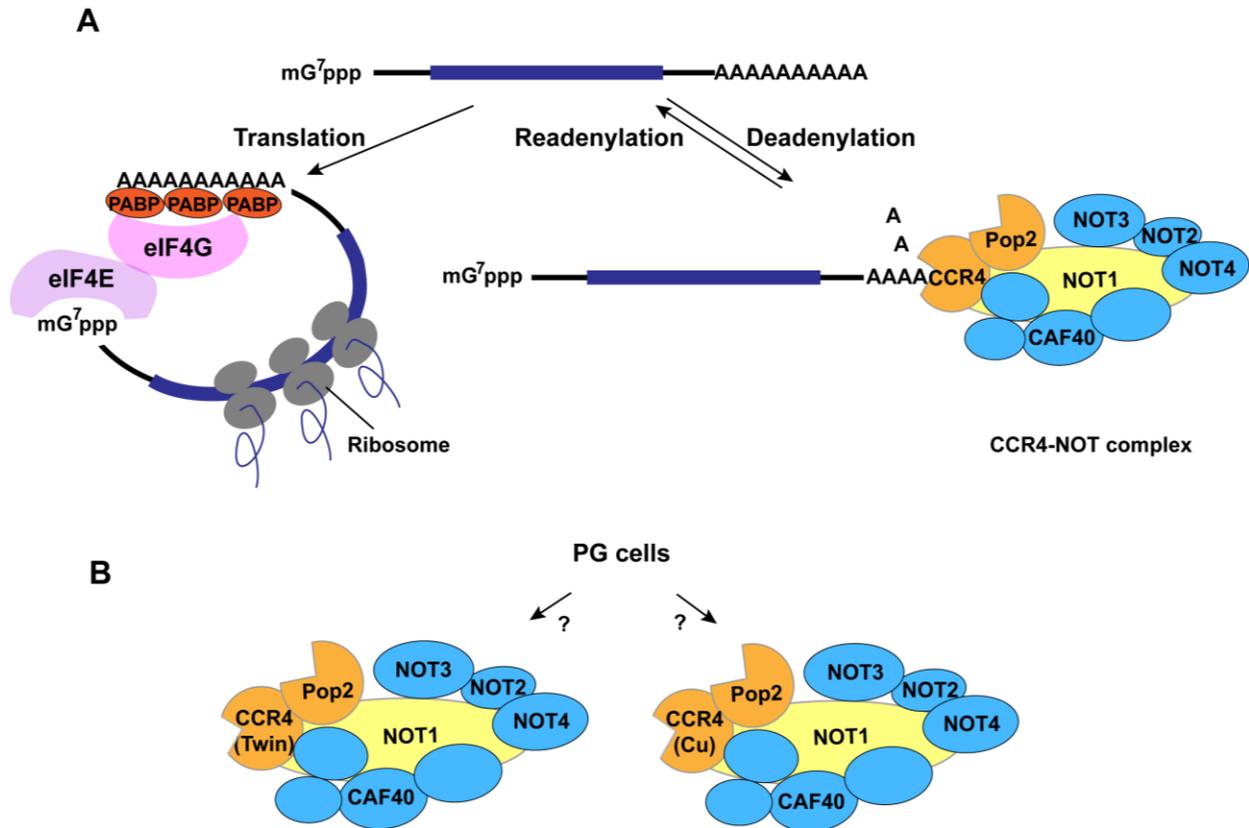


Figure 3-2. CCR4-NOT complex and the relationship between mRNA translation and deadenylation.

(A) The poly(A) tail length and translation. Eukaryotic mRNAs are modified at the 5' end with a 7-methylguanosine cap and at the 3' end with a poly(A) tail. Efficient translation of mRNAs requires mRNA circularization, which is achieved when Poly(A)-bound poly(A) binding protein (PABP) and 5'-cap-bound Eukaryotic Translation Initiation Factor 4E (eIF4E) are brought together by another translation initiation factor eIF4G. The CCR4-NOT complex functions as an mRNA deadenylase, which will shorten the poly(A) tail, thus causing translational repression and mRNA decay. The mRNA deadenylation is balanced with readenylation which will not be the focus of my dissertation. The CCR4-NOT is a multi-protein complex comprising of at least the homolog of yeast carbon catabolite repression 4 (CCR4), Pop2, NOT1, NOT2, NOT3, NOT4 and CAF40 as the core subunits. (B) The potential role of CCR4-NOT complex in the PG. There are four CCR4 paralogs in higher eukaryotes including *Drosophila*, which are Twin, Nocturnin (Cu), Angel and 3635 (CG31759). Twin and Cu can be co-immunoprecipitated with NOT1 protein, the scaffold component of the complex, but not Angel and 3635, which suggests that either Twin or Cu could be the functional CCR4 subunit working in the CCR4-NOT complex. I am interested in identifying which CCR4 (Twin or Cu) is working with the CCR4-NOT complex in the PG.

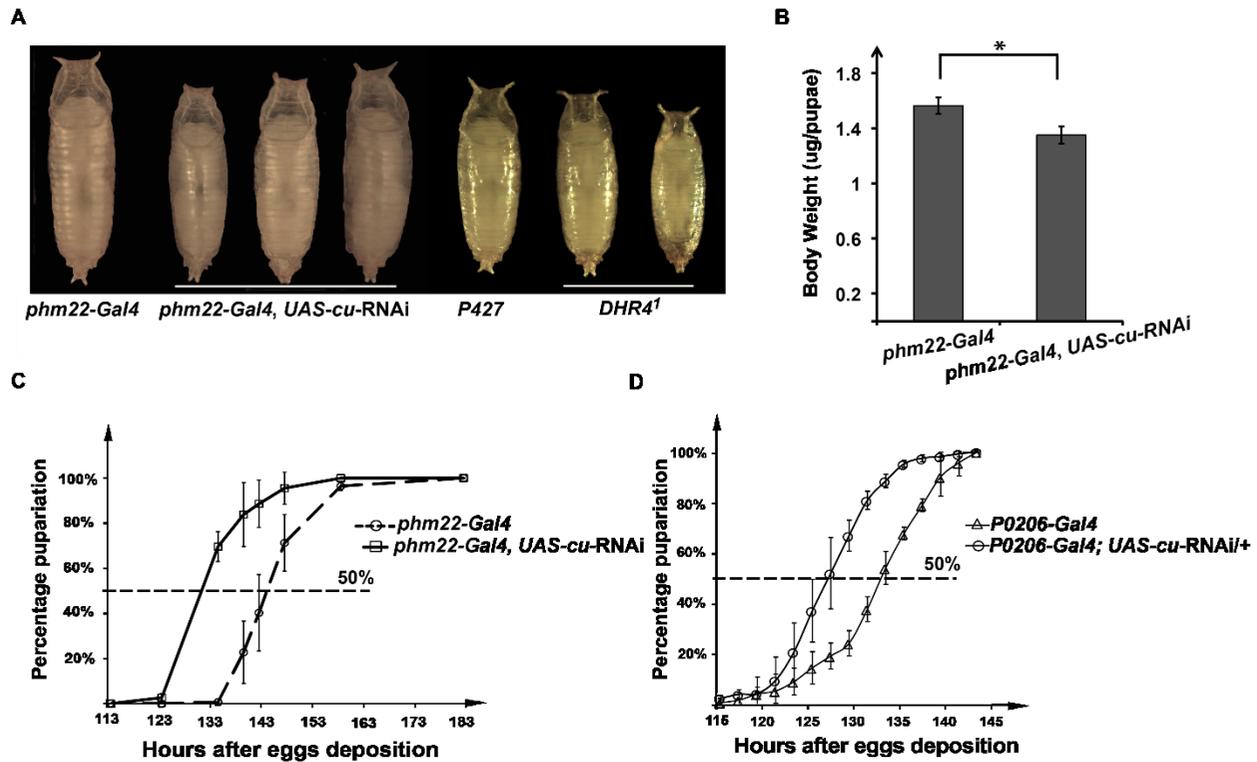


Figure 3-3. Loss-of-*cu* function in the PG resulted in small pupae due to accelerated larval development. (A) Pupae size defects caused by PG-specific knock down of *cu* as well as in *DHR4*¹ mutants. Homozygous *phm22-Gal4* is the control. *P427*: parental control line for *DHR4*¹ mutants. (B) Homozygous *phm22-Gal4, UAS-cu-RNAi* pupae have reduced body weight compared to *phm22-Gal4* (x2) controls. **p*<0.05. (C) Expression of *cu-RNAi* in the RG causes premature pupariation. Y-axis shows the percentages of embryos (staged within a 3-hr interval) that reached pupariation for *phm22-Gal4, UAS-cu-RNAi* (x2) (solid line, N = 240) and *phm22-Gal4* (x2) controls (dashed line, N = 240) at different time, hours are after egg deposition. (D) Disrupting of *cu* in the ring gland (*P0206-Gal4*) causes premature pupariation. Y-axis shows the percentages of embryos that reached pupariation. Homozygous *P0206-Gal4* line is the control (N = 180) and single copy of *cu-RNAi* driven by homozygous *P0206-Gal4* caused developmental acceleration by ~ 6 hr (N=180).

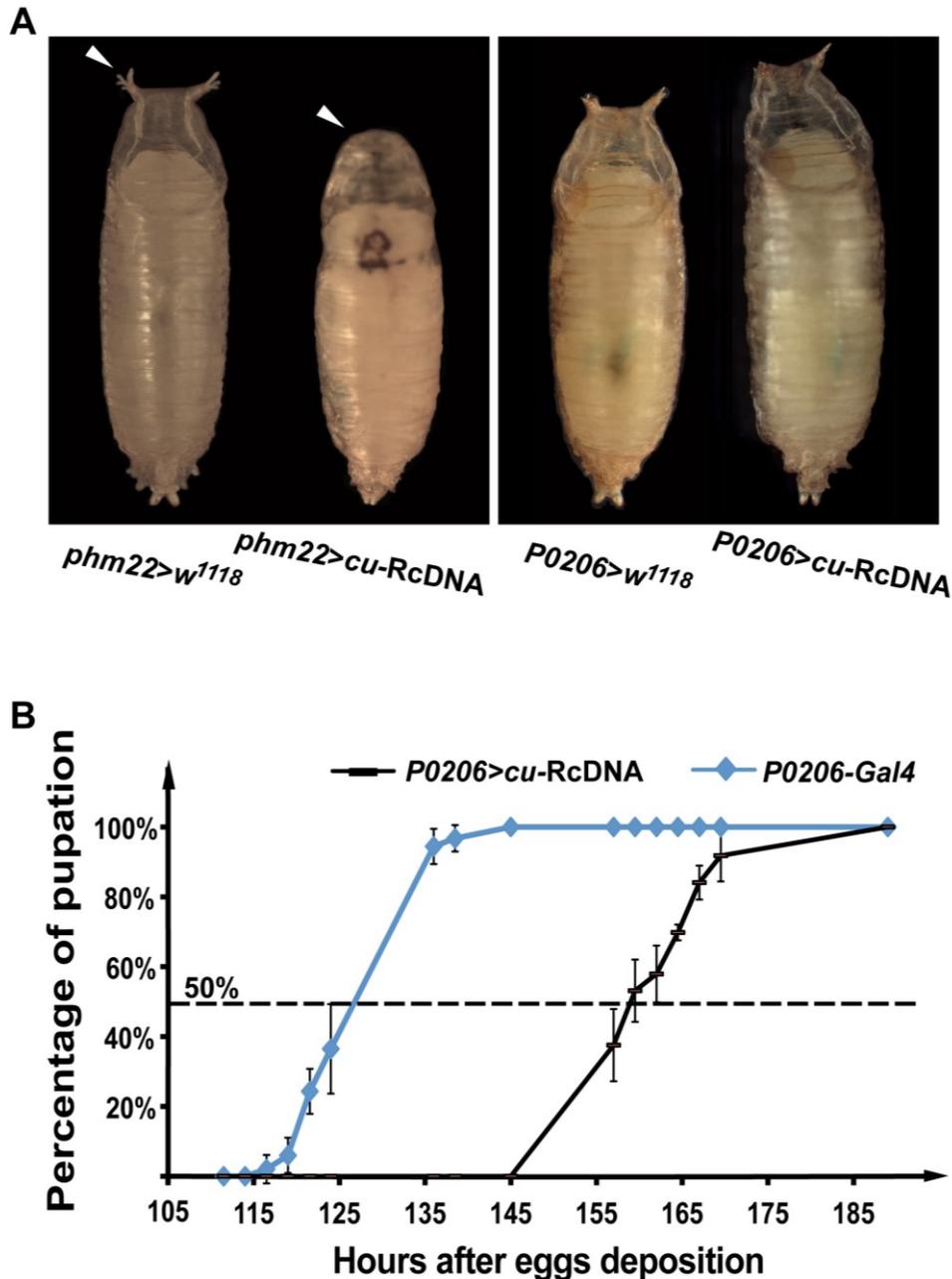


Figure 3-4. Overexpression of *cu* in the ring gland caused developmental arrest or delay.

(A) L2 prepupae and giant pupae observed in *cu*-RcDNA overexpression using the *phm22-Gal4* and *P0206-Gal4* respectively. From left to right: *phm22>w¹¹¹⁸* control L3 pupae, L2 prepupa phenotype caused by PG-specific overexpression of *cu*-RcDNA, *P0206>w¹¹¹⁸* control and giant pupae caused by ring gland specific overexpression of *cu*-RcDNA. *phm22>cu*-RcDNA is lethal/arrested, while *P0206>cu*-RcDNA can make to adult stage (just delayed). White arrowhead indicates the missing everted anterior spiracles in L2 prepupae because they forgo the L3 stage, while the anterior spiracles everted normally in control L3 pupae. (B) Overexpression of *cu* in the ring gland (*P0206-Gal4*, a weaker driver) caused developmental delay (black curve). Controls are *P0206-Gal4* driver alone. Y-axis represents the percentage of larvae (N=180) that formed pupae at a given developmental time points (indicated by hours after eggs deposition).

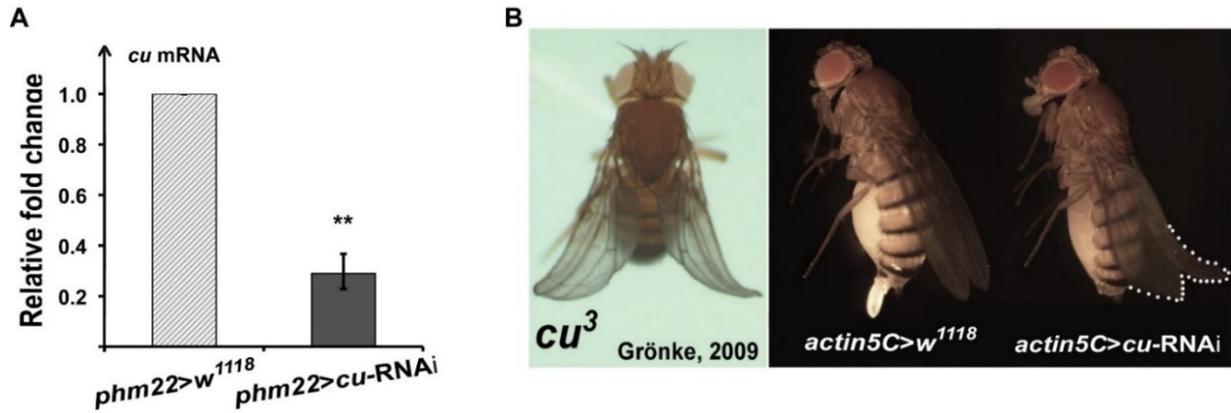


Figure 3-5. *cu*-dsRNA (VDRC construct GD8898) is functioning effectively.

(A) Expression of *cu*-dsRNA transgene (VDRC construct GD8898) in the PG disrupting *cu* transcripts. qPCR analysis of *cu* transcript levels in *phm22>cu-RNAi* L3 wandering larvae. Ring glands were collected at 24 hr L3 and the qPCR primer pair detects all three *cu* isoforms (*cu*-common_#86 in Table 3-5). The fold change is relative to the *phm22>w¹¹¹⁸* control. ***p*<0.01. (B) Ubiquitous knock down of *cu* phenocopied *cu* null mutants. *Cu* null mutants (*cu³*) have a curled wing phenotype (123). Ubiquitous expression of *UAS-cu-RNAi* (construct GD8898) using an *actin5C-Gal4* driver also gave rise to the curled wing phenotype.

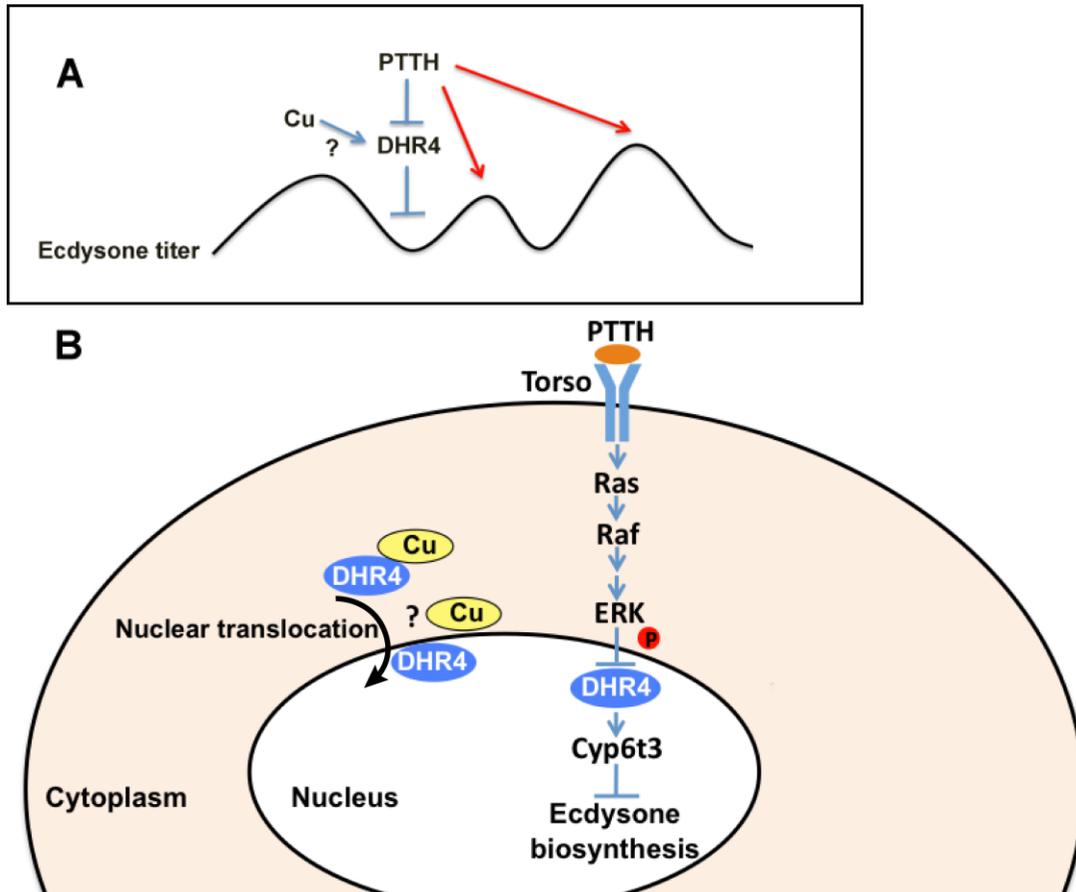


Figure 3-6. Models showing the putative role of Cu (Nocturnin) in PG cells.

In inset (A): a general description of Cu function in the PG. DHR4 functions as a repressor of ecdysone pulses, while PTTH represses DHR4 allowing for ecdysone pulses to occur. Cu probably has a function of negatively regulating ecdysone biosynthesis by promoting DHR4 activity. Lower panel (B): a detailed description of Cu functions. In PG cells, DHR4 is able to repress ecdysone biosynthesis when it is in the nucleus by directly repressing the expression of *Cyp6t3*, an ecdysone biosynthetic gene. Cu probably functions as a negative regulator of ecdysone pulses by promoting the nuclear translocation of DHR4. Upon binding to its receptor *torso*, PTTH will activate the Ras/Raf/MAPK pathway and allow ecdysone pulses to occur by promoting the translocation of DHR4 from the nucleus to the cytoplasm probably via phosphorylation.

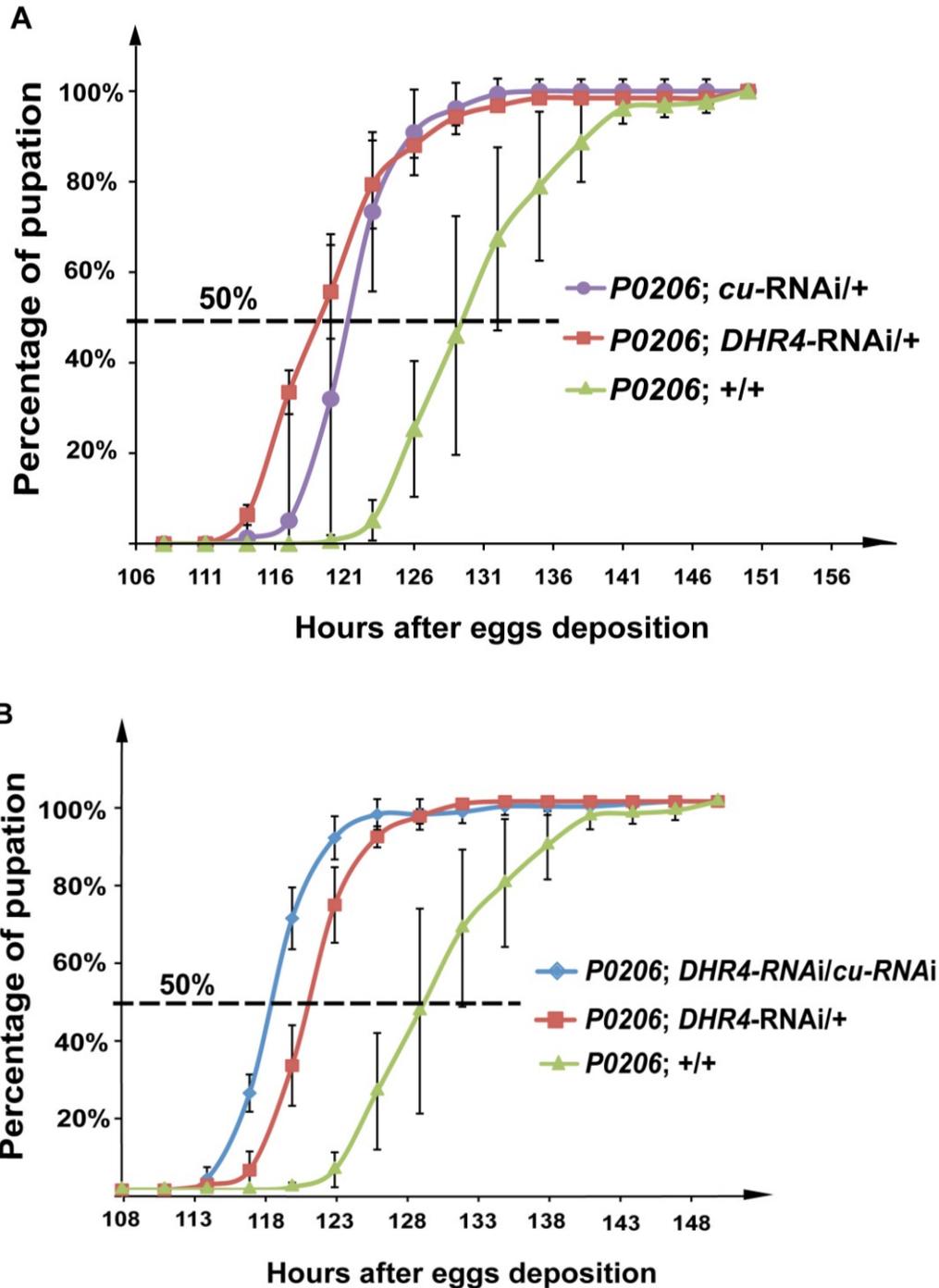


Figure 3-7. *cu*-RNAi in the ring gland phenocopied *DHR4*-RNAi with respect to developmental timing.

(A) Expression of *cu*- or *DHR4*-RNAi in the ring gland resulted in developmental acceleration. Genotypes: *P0206-Gal4* alone (homozygous, x2) (green, N=136), *P0206* (x2)>*DHR4*-RNAi (one copy, x1) (red, N=155) and *P0206* (x2)>*cu*-RNAi (x1) (purple, N=100). (B) *cu*- and *DHR4*-double knock-down in the ring gland further elevated the developmental acceleration. Genotypes: *P0206-Gal4* alone (x2) (green, N=136), *P0206* (x2)>*DHR4*-RNAi (x1) (red, N=155) and *P0206* (x2)>*DHR4*-RNAi/*cu*-RNAi (blue, N=170). (A and B) Percentages indicate the fraction of embryos that developed into prepupae at a given time point, hours are after egg deposition. All populations were tested in triplicate.

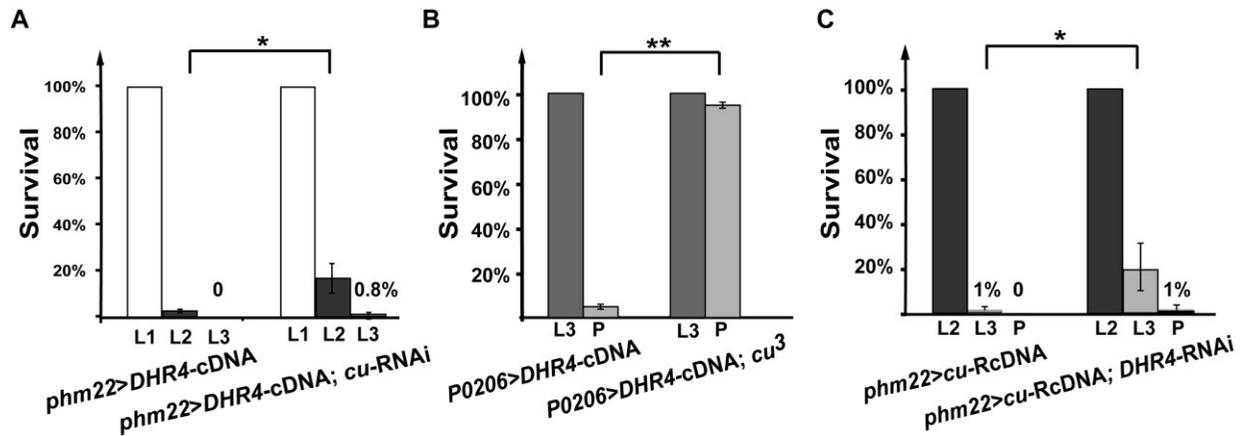


Figure 3-8. *cu* genetically interacts with *DHR4*.

(A) Genetic epistasis analysis examining the developmental arrest of PG>*DHR4*-cDNA animals, or animals with PG>*DHR4*-cDNA in the *cu*-RNAi background. *phm22*>*DHR4*-cDNA: *phm22-Gal4*>*UAS-EGFP*, *UAS-DHR4*-cDNA. *phm22*>*DHR4*-cDNA; *cu*-RNAi: *phm22-Gal4*> *UAS-DHR4*-cDNA; *UAS-cu*-RNAi. *UAS-EGFP* (not related to either *DHR4* or *cu*) is meant to add an extra UAS site to make an equal number of UAS sites within the two genotypes. (B) Comparing the developmental progression (survival) between ring gland-*DHR4*-overexpression and *DHR4*-overexpression in the *cu* mutant background. *cu*³: a the null allele of *cu* (123). (C) Genetic epistasis analysis examining the developmental arrest for transgenic lines expressing *cu*-RcDNA, or both the *cu*-RcDNA and *DHR4*-RNAi in the PG. *phm22*>*cu*-RcDNA: *phm22-Gal4*>*UAS-EGFP*, *UAS-cu*-RcDNA. *phm22*>*cu*-RcDNA; *DHR4*-RNAi: *phm22-Gal4*>*UAS-cu*-RcDNA; *UAS-DHR4*-RNAi. (A-C) Percentages indicate the fraction of larvae that developed to the indicated stages. Error bar represents standard deviation. L1: first instar, L2: second instar, L3: third instar and P: pupae.

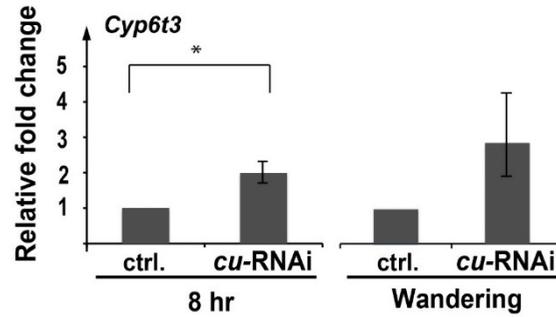


Figure 3-9. Knocking down *cu* in the PG derepressed *Cyp6t3* expression.

qPCR analysis examining the mRNA levels of *Cyp6t3* at two different developmental stages. *Cyp6t3* is a potential target of DHR4, when DHR4 is not in the nucleus of PG cells, *Cyp6t3* will be upregulated. Brain-ring gland complexes are collected at: 1) 8 hr after the L2/L3 molt and 2) the wandering stage. ctrl.: *phm22>w¹¹⁸*. *cu*-RNAi: *phm22>cu*-RNAi. Fold changes are relative to the *phm22>w¹¹⁸* control at two time point respectively. The error bars are 95% confidence intervals and * $p < 0.05$.

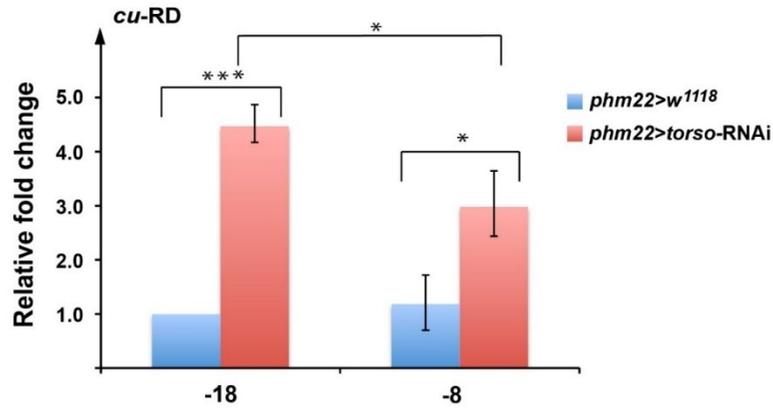


Figure 3-10. The *cu*-RD isoform is dependent on Torso (the receptor of the PTTH pathway).

qPCR analysis was carried out to examine the transcripts levels of *cu*-RD in *phm22>torso*-RNAi larvae. Brain-ring gland complexes were collected at two different developmental stages. Animals were staged according to the blue gut method: -18 represents 18 hrs before pupae formation (BPF), *aka* blue gut wandering larvae and -8 represents 8 hrs before pupae formation (BPF), *aka* partial blue gut wandering larvae. Fold changes are relative to the *phm22>w¹¹¹⁸* control at 18 BPF. The error bars are 95% confidence intervals and *** $p < 0.001$; * $p < 0.05$.

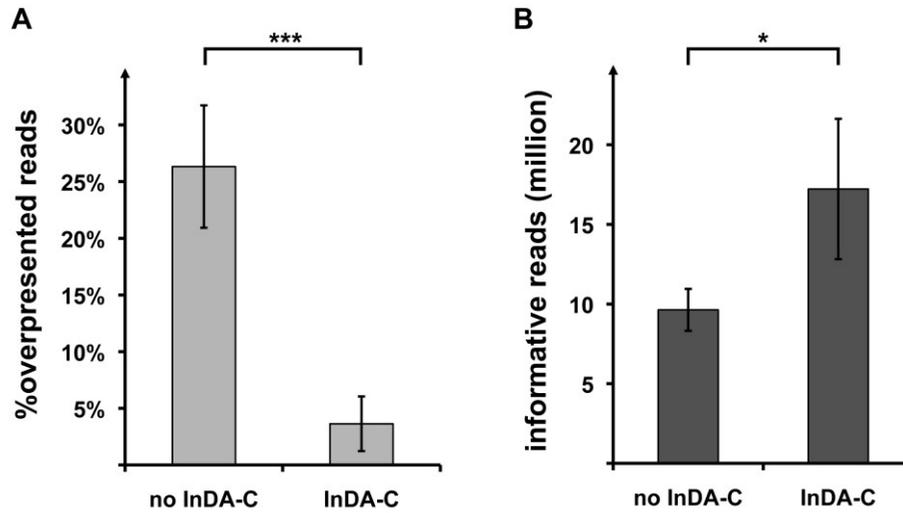
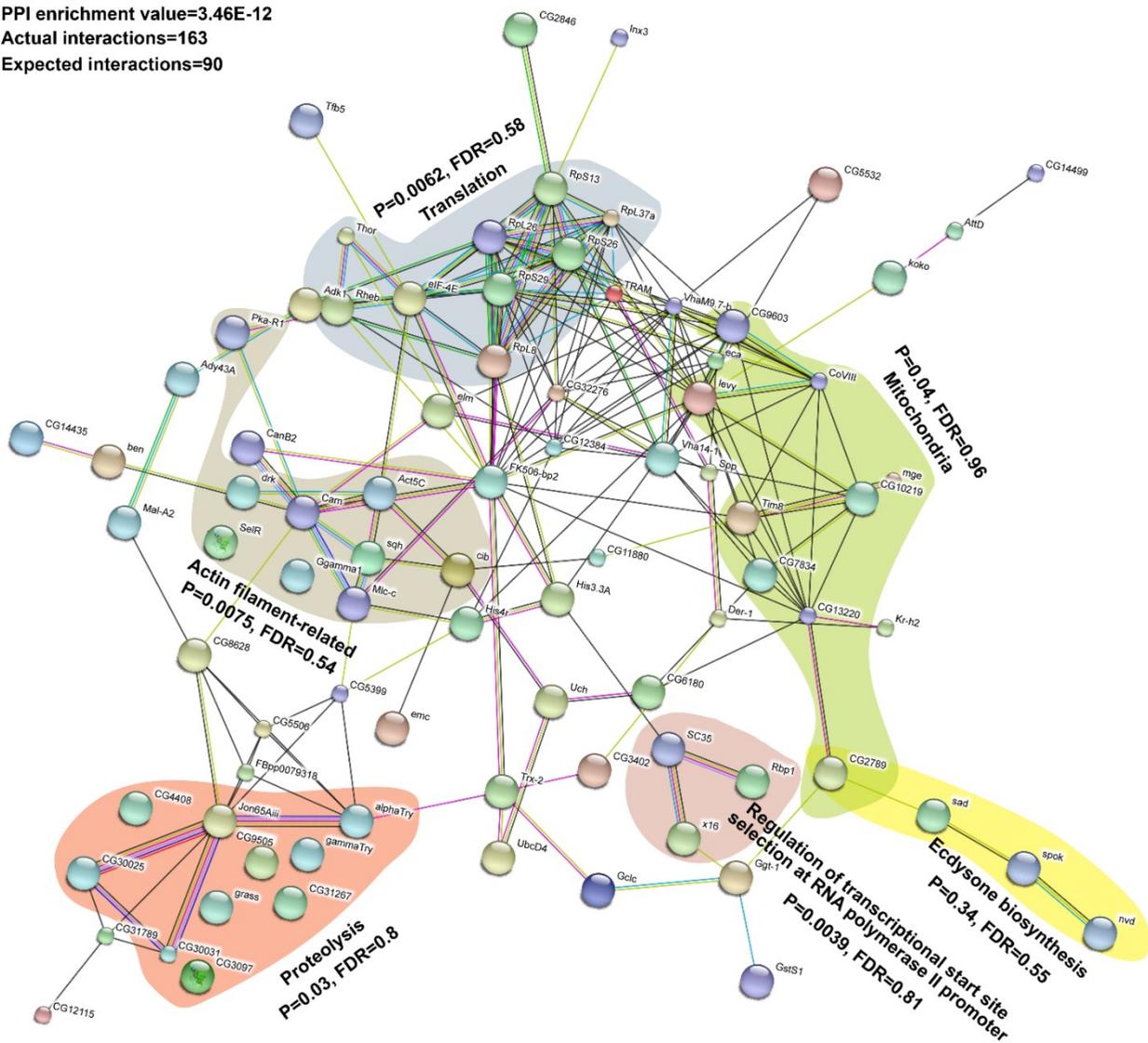


Figure 3-11. InDA-C treatment (Nugen, InC) significantly reduced the amount of rRNA reads from total RNA samples in RNA-Seq analysis.

(A) The average percentage of overrepresented reads (rRNA) with or without InDA-C treatment. (B) Average total informative million reads per sample with or without InDA-C treatment. (A-B) no InDA-C: four cDNA sequencing libraries taken from the *Pop2*-RNAi RNA-Seq analysis prepared without the InDA-C treatments were used for the quantification. InDA-C: four cDNA sequencing libraries taken from the PG>*sna*-RNAi and hs>*sna*-cDNA RNA-Seq analysis (see section 3.2.15 and 3.3.11) prepared with the InDA-C treatment steps were used for the quantification. The error bars represent the standard deviation and *** $p < 0.001$; * $p < 0.05$.

PPI enrichment value=3.46E-12
 Actual interactions=163
 Expected interactions=90



Legend

Known Interactions	Predicted Interactions	Others
from curated databases	gene neighborhood	textmining
experimentally determined	gene fusions	co-expression
	gene co-occurrence	protein homology

Figure 3-12. The downregulated genes in PG>Pop2-RNAi ring glands suggested protein translation defects.

Interaction network of >3-fold downregulated genes in PG>Pop2-RNAi ring glands compared to controls (*phm22-Gal4>w¹¹¹⁸*) generated by STRING database. Several gene ontology (GO) groups (based on DAVID GO) within the big network were highlighted using different color shading. p value: based on modified Fisher exact model (the smaller, the more enriched for the GO term). FDR: False discovery rate based on Benjamini method. The non-associated genes were not included in the figure. The large node size

indicates the availability of protein structure information. Node colors have no particular meaning. Line colors indicate different types of evidence for the interaction (see legend). PPI enrichment value: protein-protein interaction enrichment value (the smaller, the more enriched). Actual interactions: number of protein interactions. Expected interactions: number of protein interactions that would be expected for a random set of proteins of similar size, drawn from the genome.

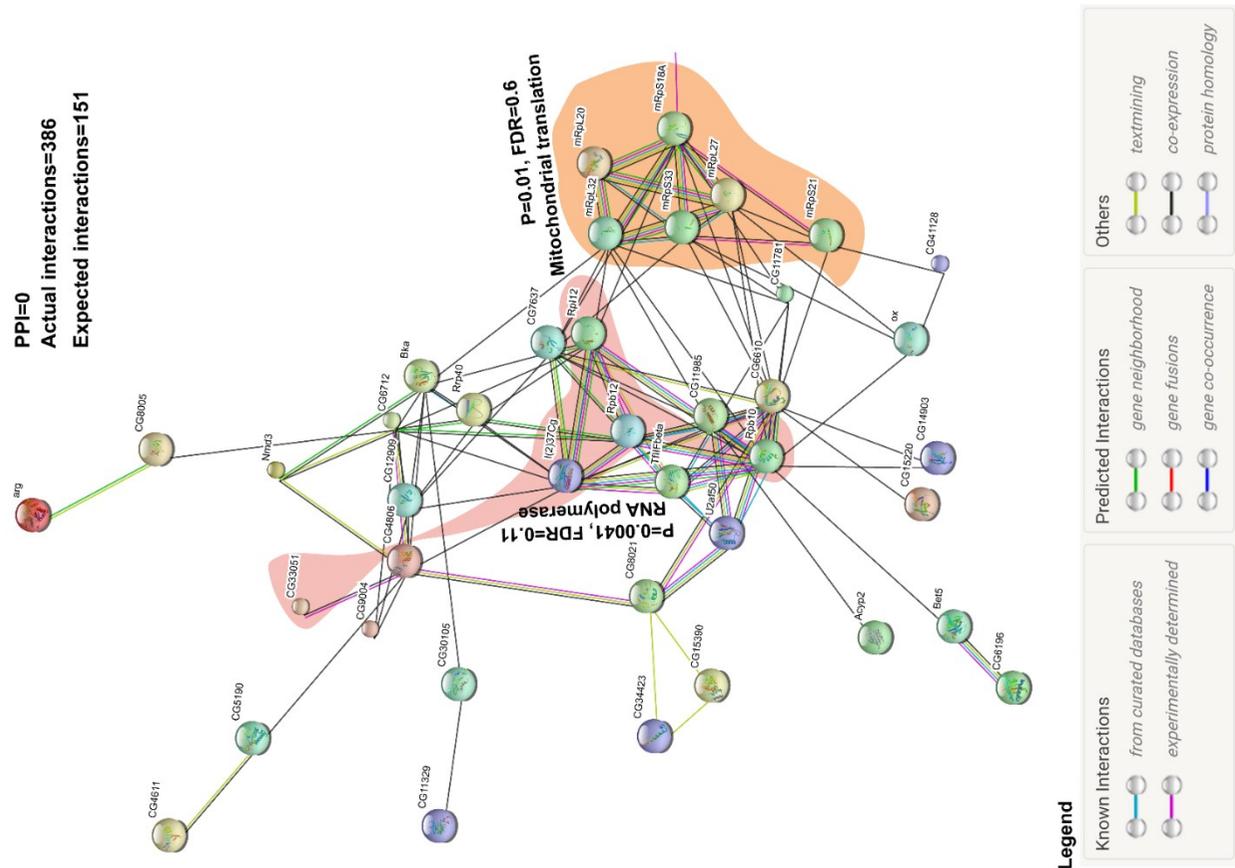


Figure 3-13. RNAi of *Pop2* in the PG upregulated the expression of genes encoding RNA polymerase and genes related to amino acid metabolism.

Interaction network of >3-fold upregulated genes in PG>*Pop2*-RNAi ring glands compared to controls (*phm22-Gal4*> *w¹¹¹⁸*) generated by STRING database. Several gene ontology (GO) groups (based on DAVID GO) within the big network were highlighted using different color shading. *p* value: based on modified Fisher exact model (the smaller, the more enriched for the GO term). FDR: False discovery rate based on Benjamini method. The non-associated genes were not included in the figure. The large node size indicates the availability of protein structure information. Node colors have no particular meaning. Line colors indicate different types of evidence for the interaction (see legend). PPI: protein-protein interaction enrichment value (the smaller, the more enriched). Actual interactions: number of protein interactions. Expected interactions: number of protein interactions that would be expected for a random set of proteins of similar size, drawn from the genome.

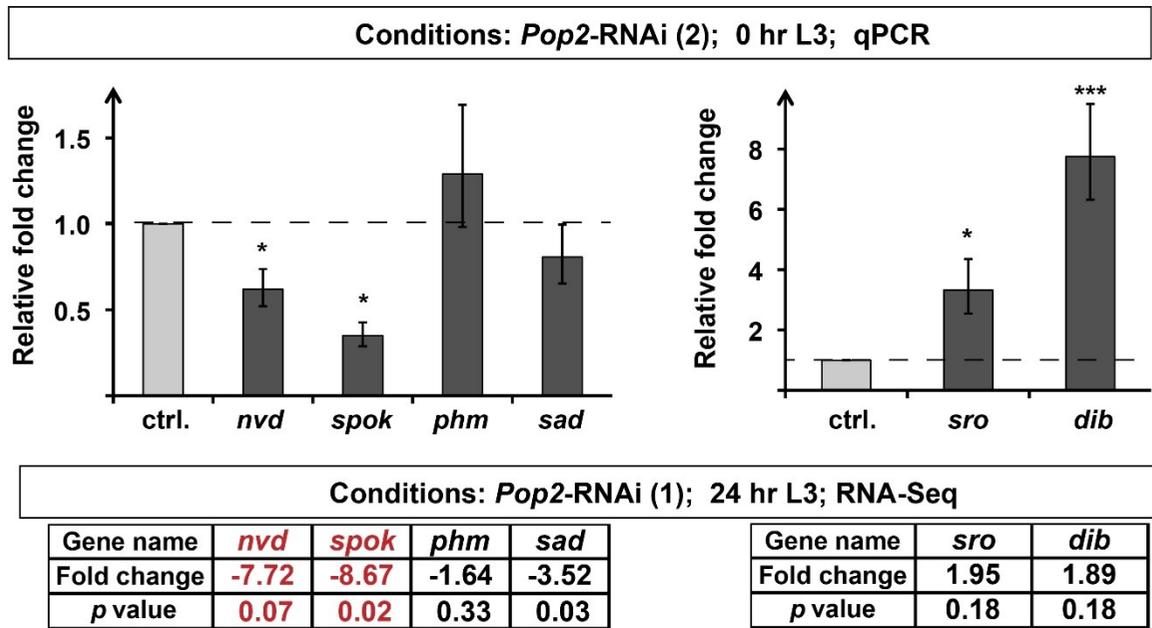


Figure 3-14. qPCR results for the expression of six major ecdysone biosynthetic genes in *PG>Pop2*-RNAi. For the qPCR analysis, ring glands were collected at 0 hr L3. Relative fold change was determined by comparing the expression in the RNAi to that of the ctrl. (*phm22-Gal4* crossed to Bloomington stock #36303, the parental line for the TRiP RNAi line) for each gene tested. *Pop2*-RNAi (2): *phm22>Pop2*-RNAi (Bloomington stock #30492). *** $p < 0.001$ and * $p < 0.05$ (based on Student's *t*-test). Tables show the fold change and *p* value from the RNA-Seq analysis. For RNA-Seq, the ring glands were collected at 24 hr L3. *Pop2*-RNAi (1): *phm22>Pop2*-RNAi (Bloomington stock #52947). Red-coded genes were validated by qPCR.

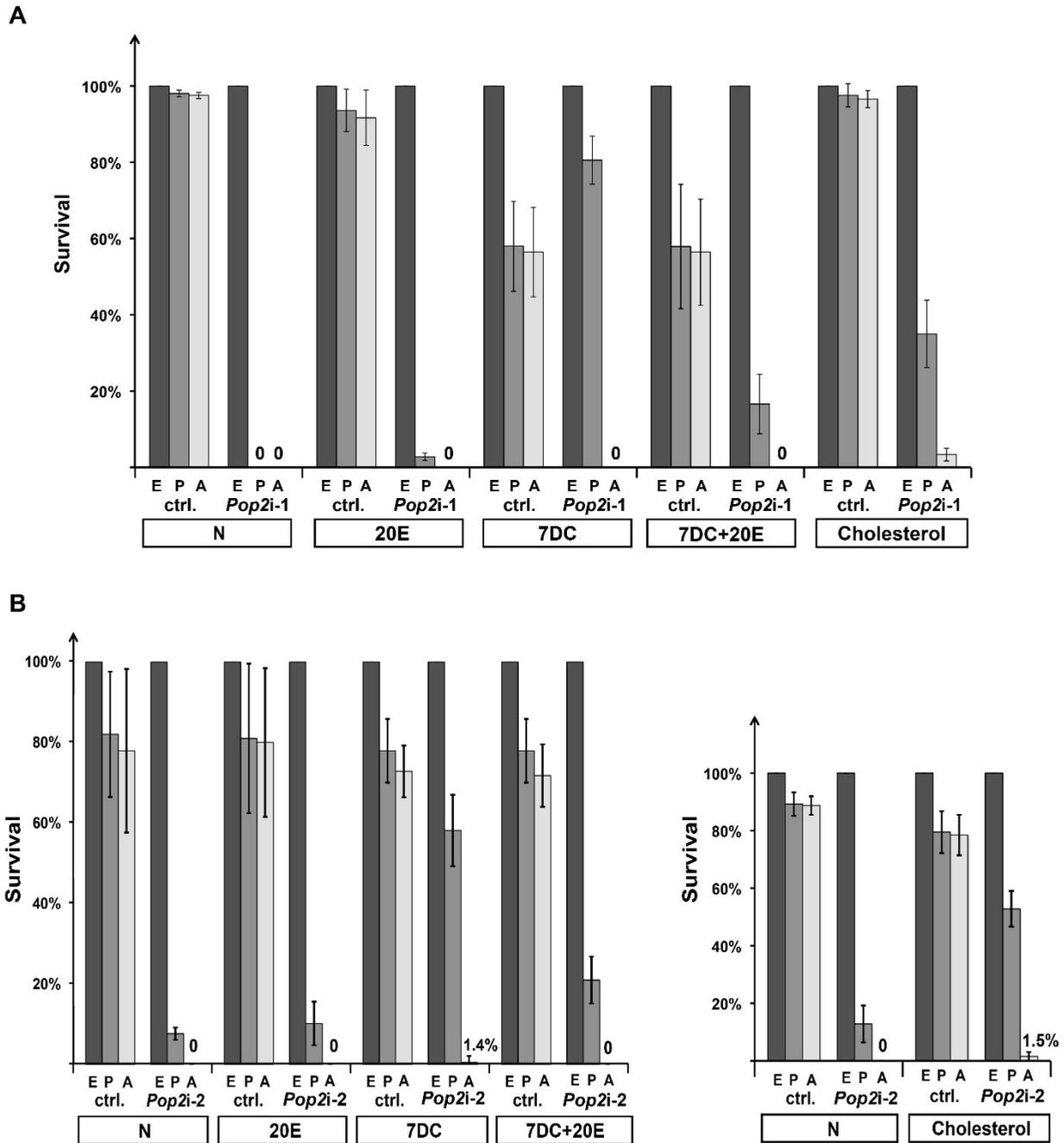


Figure 3-15. Larval arrest caused by PG>Pop2-RNAi could be partially rescued by 7DC and cholesterol feeding.

(A) Percentage of embryos survived to the indicated stages under different conditions and genotypes. ctrl.: *phm22-Gal4* crossed to Bloomington stock #36304 (the parental line for the TRiP RNAi line #52947). *Pop2i-1*: *phm22>Pop2-RNAi* (Bloomington stock #52947). (B) ctrl.: *phm22-Gal4* crossed to Bloomington stock #36303 (the parental line for the TRiP RNAi line #30492). *Pop2i-2*: *phm22>Pop2-RNAi* (Bloomington stock #30492). (A-B) E: embryo, P: pupae, A: adult. N: Nutrifly food without the sterols, 20E: 20-Hydroxyecdysone, 7DC: 7-dehydrocholesterol. Error bars represent standard deviation.

References:

- 1 Fujita, T. Mineralocorticoid receptors, salt-sensitive hypertension, and metabolic syndrome. *Hypertension* **55**, 813-818, (2010).
- 2 Sapolsky, R. M., Romero, L. M. & Munck, A. U. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocr Rev* **21**, 55-89, (2000).
- 3 Sisk, C. L. & Foster, D. L. The neural basis of puberty and adolescence. *Nat Neurosci* **7**, 1040-1047, (2004).
- 4 Kuo, T., McQueen, A., Chen, T. C. & Wang, J. C. Regulation of Glucose Homeostasis by Glucocorticoids. *Adv Exp Med Biol* **872**, 99-126, (2015).
- 5 Riddiford, L. M., Cherbas, P. & Truman, J. W. Ecdysone receptors and their biological actions. *Vitamins and Hormones - Advances in Research and Applications, Vol 60* **60**, 1-73, (2001).
- 6 Koelle, M. R. *et al.* The Drosophila EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* **67**, 59-77, (1991).
- 7 Thomas, H. E., Stunnenberg, H. G. & Stewart, A. F. Heterodimerization of the Drosophila ecdysone receptor with retinoid X receptor and ultraspiracle. *Nature* **362**, 471-475, (1993).
- 8 Yao, T. P. *et al.* Functional Ecdysone Receptor Is the Product of Ecr and Ultraspiracle Genes. *Nature* **366**, 476-479, (1993).
- 9 Burtis, K. C., Thummel, C. S., Jones, C. W., Karim, F. D. & Hogness, D. S. The Drosophila 74ef Early Puff Contains E74, a Complex Ecdysone-Inducible Gene That Encodes 2 Ets-Related Proteins. *Cell* **61**, 85-99, (1990).
- 10 Segraves, W. A. & Hogness, D. S. The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila encodes two new members of the steroid receptor superfamily. *Genes Dev* **4**, 204-219, (1990).
- 11 DiBello, P. R., Withers, D. A., Bayer, C. A., Fristrom, J. W. & Guild, G. M. The Drosophila Broad-Complex encodes a family of related proteins containing zinc fingers. *Genetics* **129**, 385-397, (1991).
- 12 Walker, V. K. & Ashburner, M. The control of ecdysterone-regulated puffs in Drosophila salivary glands. *Cell* **26**, 269-277, (1981).
- 13 Huet, F., Ruiz, C. & Richards, G. Puffs and PCR: the in vivo dynamics of early gene expression during ecdysone responses in Drosophila. *Development* **118**, 613-627, (1993).
- 14 Gonsalves, S. E., Neal, S. J., Kehoe, A. S. & Westwood, J. T. Genome-wide examination of the transcriptional response to ecdysteroids 20-hydroxyecdysone and ponasterone A in Drosophila melanogaster. *BMC Genomics* **12**, 475, (2011).
- 15 King-Jones, K. & Thummel, C. S. Nuclear receptors--a perspective from Drosophila. *Nat Rev Genet* **6**, 311-323, (2005).
- 16 Delanoue, R., Slaidina, M. & Leopold, P. The steroid hormone ecdysone controls systemic growth by repressing dMyc function in Drosophila fat cells. *Dev Cell* **18**, 1012-1021, (2010).
- 17 McBrayer, Z. *et al.* Prothoracicotropic hormone regulates developmental timing and body size in Drosophila. *Dev Cell* **13**, 857-871, (2007).
- 18 Stavreva, D. A. *et al.* Ultradian hormone stimulation induces glucocorticoid receptor-mediated pulses of gene transcription. *Nat Cell Biol* **11**, 1093-1102, (2009).

- 19 Smith, S. M. & Vale, W. W. The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress. *Dialogues Clin Neurosci* **8**, 383-395, (2006).
- 20 Hiruma, K. & Kaneko, Y. Hormonal regulation of insect metamorphosis with special reference to juvenile hormone biosynthesis. *Curr Top Dev Biol* **103**, 73-100, (2013).
- 21 Kim, S. K. & Rulifson, E. J. Conserved mechanisms of glucose sensing and regulation by *Drosophila corpora cardiaca* cells. *Nature* **431**, 316-320, (2004).
- 22 Truman, J. W. & Riddiford, L. M. The morphostatic actions of juvenile hormone. *Insect Biochem Mol Biol* **37**, 761-770, (2007).
- 23 Lee, G. & Park, J. H. Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in *Drosophila melanogaster*. *Genetics* **167**, 311-323, (2004).
- 24 Yoshiyama, T., Namiki, T., Mita, K., Kataoka, H. & Niwa, R. Neverland is an evolutionally conserved Rieske-domain protein that is essential for ecdysone synthesis and insect growth. *Development* **133**, 2565-2574, (2006).
- 25 Warren, J. T. *et al.* Molecular and biochemical characterization of two P450 enzymes in the ecdysteroidogenic pathway of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* **99**, 11043-11048, (2002).
- 26 Warren, J. T. *et al.* Phantom encodes the 25-hydroxylase of *Drosophila melanogaster* and *Bombyx mori*: a P450 enzyme critical in ecdysone biosynthesis. *Insect Biochemistry and Molecular Biology* **34**, 991-1010, (2004).
- 27 Niwa, R. *et al.* The ecdysteroidogenic P450 Cyp302a1/disembodied from the silkworm, *Bombyx mori*, is transcriptionally regulated by prothoracicotropic hormone. *Insect Molecular Biology* **14**, 563-571, (2005).
- 28 Chavez, V. M. *et al.* The *Drosophila* disembodied gene controls late embryonic morphogenesis and codes for a cytochrome P450 enzyme that regulates embryonic ecdysone levels. *Development* **127**, 4115-4126, (2000).
- 29 Niwa, R. *et al.* Non-molting glossy/shroud encodes a short-chain dehydrogenase/reductase that functions in the 'Black Box' of the ecdysteroid biosynthesis pathway. *Development* **137**, 1991-1999, (2010).
- 30 Ono, H. *et al.* Spook and Spookier code for stage-specific components of the ecdysone biosynthetic pathway in Diptera. *Dev Biol* **298**, 555-570, (2006).
- 31 Qiuxiang Ou, A. M. a. K. K.-j. Nuclear Receptor DHR4 controls the timing of steroid hormone pulses during *Drosophila* development. *PLoS Biology*, (2011).
- 32 Petryk, A. *et al.* Shade is the *Drosophila* P450 enzyme that mediates the hydroxylation of ecdysone to the steroid insect molting hormone 20-hydroxyecdysone. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 13773-13778, (2003).
- 33 Rewitz, K. F., Yamanaka, N., Gilbert, L. I. & O'Connor, M. B. The insect neuropeptide PTTH activates receptor tyrosine kinase torso to initiate metamorphosis. *Science* **326**, 1403-1405, (2009).
- 34 Li, W. X. Functions and mechanisms of receptor tyrosine kinase Torso signaling: lessons from *Drosophila* embryonic terminal development. *Dev Dyn* **232**, 656-672, (2005).
- 35 Kataoka, H. *et al.* Differential regulation of ecdysteroidogenic P450 gene expression in the silkworm, *Bombyx mori*. *Bioscience Biotechnology and Biochemistry* **71**, 2808-2814, (2007).
- 36 Yamanaka, N., Rewitz, K. F. & O'Connor, M. B. Ecdysone control of developmental transitions: lessons from *Drosophila* research. *Annu Rev Entomol* **58**, 497-516, (2013).

- 37 Gibbens, Y. Y., Warren, J. T., Gilbert, L. I. & O'Connor, M. B. Neuroendocrine regulation of *Drosophila* metamorphosis requires TGFbeta/Activin signaling. *Development* **138**, 2693-2703, (2011).
- 38 DeLalio, L. J., Dion, S. M., Bootes, A. M. & Smith, W. A. Direct effects of hypoxia and nitric oxide on ecdysone secretion by insect prothoracic glands. *J Insect Physiol* **76**, 56-66, (2015).
- 39 Caceres, L. *et al.* Nitric oxide coordinates metabolism, growth, and development via the nuclear receptor E75. *Genes & Development* **25**, 1476-1485, (2011).
- 40 Di Cara, F. & King-Jones, K. The Circadian Clock Is a Key Driver of Steroid Hormone Production in *Drosophila*. *Current Biology* **26**, 2469-2477, (2016).
- 41 Caldwell, P. E., Walkiewicz, M. & Stern, M. Ras activity in the *Drosophila* prothoracic gland regulates body size and developmental rate via ecdysone release. *Curr Biol* **15**, 1785-1795, (2005).
- 42 Colombani, J. *et al.* Antagonistic actions of ecdysone and insulins determine final size in *Drosophila*. *Science* **310**, 667-670, (2005).
- 43 Mirth, C., Truman, J. W. & Riddiford, L. M. The role of the prothoracic gland in determining critical weight for metamorphosis in *Drosophila melanogaster*. *Curr Biol* **15**, 1796-1807, (2005).
- 44 Layalle, S., Arquier, N. & Leopold, P. The TOR pathway couples nutrition and developmental timing in *Drosophila*. *Dev Cell* **15**, 568-577, (2008).
- 45 Colombani, J., Andersen, D. S. & Leopold, P. Secreted peptide Dilp8 coordinates *Drosophila* tissue growth with developmental timing. *Science* **336**, 582-585, (2012).
- 46 Garelli, A., Gontijo, A. M., Miguela, V., Caparros, E. & Dominguez, M. Imaginal discs secrete insulin-like peptide 8 to mediate plasticity of growth and maturation. *Science* **336**, 579-582, (2012).
- 47 Layalle, S., Arquier, N. & Leopold, P. The TOR Pathway Couples Nutrition and Developmental Timing in *Drosophila*. *Developmental Cell* **15**, 568-577, (2008).
- 48 Britton, J. S., Lockwood, W. K., Li, L., Cohen, S. M. & Edgar, B. A. *Drosophila*'s insulin/PI3-kinase pathway coordinates cellular metabolism with nutritional conditions. *Dev Cell* **2**, 239-249, (2002).
- 49 Zhang, H. B., Stallock, J. P., Ng, J. C., Reinhard, C. & Neufeld, T. P. Regulation of cellular growth by the *Drosophila* target of rapamycin dTOR. *Genes & Development* **14**, 2712-2724, (2000).
- 50 Brogiolo, W. *et al.* An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Current Biology* **11**, 213-221, (2001).
- 51 Ikeya, T., Galic, M., Belawat, P., Nairz, K. & Hafen, E. Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in *Drosophila*. *Current Biology* **12**, 1293-1300, (2002).
- 52 Broughton, S. J. *et al.* Longer lifespan, altered metabolism, and stress resistance in *Drosophila* from ablation of cells making insulin-like ligands. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3105-3110, (2005).
- 53 Ikeya, T., Galic, M., Belawat, P., Nairz, K. & Hafen, E. Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in *Drosophila*. *Curr Biol* **12**, 1293-1300, (2002).
- 54 Rulifson, E. J., Kim, S. K. & Nusse, R. Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. *Science* **296**, 1118-1120, (2002).

- 55 Hasygar, K. & Hietakangas, V. p53-and ERK7-Dependent Ribosome Surveillance Response Regulates Drosophila Insulin-Like Peptide Secretion. *Plos Genetics* **10**, (2014).
- 56 Teleman, A. A. Molecular mechanisms of metabolic regulation by insulin in Drosophila. *Biochem J* **425**, 13-26, (2009).
- 57 Leever, S. J., Weinkove, D., MacDougall, L. K., Hafen, E. & Waterfield, M. D. The Drosophila phosphoinositide 3-kinase Dp110 promotes cell growth. *EMBO J* **15**, 6584-6594, (1996).
- 58 Verdu, J., Buratovich, M. A., Wilder, E. L. & Birnbaum, M. J. Cell-autonomous regulation of cell and organ growth in Drosophila by Akt/PKB. *Nat Cell Biol* **1**, 500-506, (1999).
- 59 Gao, X. S., Neufeld, T. P. & Pan, D. J. Drosophila PTEN regulates cell growth and proliferation through PI3K-dependent and -independent pathways. *Developmental Biology* **221**, 404-418, (2000).
- 60 Martinek, S., Inonog, S., Manoukian, A. S. & Young, M. W. A role for the segment polarity gene shaggy/GSK-3 in the Drosophila circadian clock. *Cell* **105**, 769-779, (2001).
- 61 Wullschleger, S., Loewith, R. & Hall, M. N. TOR signaling in growth and metabolism. *Cell* **124**, 471-484, (2006).
- 62 Grewal, S. S. Insulin/TOR signaling in growth and homeostasis: a view from the fly world. *Int J Biochem Cell Biol* **41**, 1006-1010, (2009).
- 63 Jones, K. A., Jiang, X. Y., Yamamoto, Y. & Yeung, R. S. Tuberin is a component of lipid rafts and mediates caveolin-1 localization: role of TSC2 in post-Golgi transport. *Experimental Cell Research* **295**, 512-524, (2004).
- 64 Schleich, S. & Teleman, A. A. Akt Phosphorylates Both Tsc1 and Tsc2 in Drosophila, but Neither Phosphorylation Is Required for Normal Animal Growth. *PLoS One* **4**, (2009).
- 65 Miron, M. & Sonenberg, N. Regulation of translation via TOR signaling: Insights from Drosophila melanogaster. *Journal of Nutrition* **131**, 2988S-2993S, (2001).
- 66 Grewal, S. S. Insulin/TOR signaling in growth and homeostasis: A view from the fly world. *International Journal of Biochemistry & Cell Biology* **41**, 1006-1010, (2009).
- 67 Cybulski, N. & Hall, M. N. TOR complex 2: a signaling pathway of its own. *Trends in Biochemical Sciences* **34**, 620-627, (2009).
- 68 Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098-1101, (2005).
- 69 Hietakangas, V. & Cohen, S. M. Re-evaluating AKT regulation: role of TOR complex 2 in tissue growth. *Genes Dev* **21**, 632-637, (2007).
- 70 Wismar, J. *et al.* The mutation without children(rgl) causes ecdysteroid deficiency in third-instar larvae of Drosophila melanogaster. *Developmental Biology* **226**, 1-17, (2000).
- 71 Warren, J. T., Wismar, J., Subrahmanyam, B. & Gilbert, L. I. Woc (without children) gene control of ecdysone biosynthesis in Drosophila melanogaster. *Molecular and Cellular Endocrinology* **181**, 1-14, (2001).
- 72 Ono, H. *et al.* Spook and Spookier code for stage-specific components of the ecdysone biosynthetic pathway in Diptera. *Developmental Biology* **298**, 555-570, (2006).
- 73 Neubueser, D., Warren, J. T., Gilbert, L. I. & Cohen, S. M. molting defective is required for ecdysone biosynthesis. *Developmental Biology* **280**, 362-372, (2005).

- 74 Cheng, C. C. *et al.* The POU Factor Ventral Veins Lacking/Drifter Directs the Timing of Metamorphosis through Ecdysteroid and Juvenile Hormone Signaling. *Plos Genetics* **10**, (2014).
- 75 Danielsen, E. T. *et al.* Transcriptional Control of Steroid Biosynthesis Genes in the Drosophila Prothoracic Gland by Ventral Veins Lacking and Knirps. *Plos Genetics* **10**, (2014).
- 76 Komura-Kawa, T. *et al.* The Drosophila Zinc Finger Transcription Factor Ouija Board Controls Ecdysteroid Biosynthesis through Specific Regulation of spookier. *Plos Genetics* **11**, (2015).
- 77 Yamanaka, N., Marques, G. & O'Connor, M. B. Vesicle-Mediated Steroid Hormone Secretion in Drosophila melanogaster. *Cell* **163**, 907-919, (2015).
- 78 Carstea, E. D. *et al.* Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. *Science* **277**, 228-231, (1997).
- 79 Morris, J. A. & Carstea, E. D. Niemann-Pick C disease: cholesterol handling gone awry. *Mol Med Today* **4**, 525-531, (1998).
- 80 Loftus, S. K. *et al.* Murine model of Niemann-Pick C disease: mutation in a cholesterol homeostasis gene. *Science* **277**, 232-235, (1997).
- 81 Fluegel, M. L., Parker, T. J. & Pallanck, L. J. Mutations of a Drosophila NPC1 gene confer sterol and ecdysone metabolic defects. *Genetics* **172**, 185-196, (2006).
- 82 Huang, X., Suyama, K., Buchanan, J., Zhu, A. J. & Scott, M. P. A Drosophila model of the Niemann-Pick type C lysosome storage disease: dnpc1a is required for molting and sterol homeostasis. *Development* **132**, 5115-5124, (2005).
- 83 Huang, X., Warren, J. T., Buchanan, J. A., Gilbert, L. I. & Scott, M. P. Drosophila Niemann-Pick Type C-2 genes control sterol homeostasis and steroid biosynthesis: a model of human neurodegenerative disease. *Development* **134**, 3733-3742, (2007).
- 84 Roth, G. E. *et al.* The Drosophila gene Start1: A putative cholesterol transporter and key regulator of ecdysteroid synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 1601-1606, (2004).
- 85 Talbot, W. S., Swyryd, E. A. & Hogness, D. S. Drosophila tissues with different metamorphic responses to ecdysone express different ecdysone receptor isoforms. *Cell* **73**, 1323-1337, (1993).
- 86 Parvy, J. P. *et al.* Forward and feedback regulation of cyclic steroid production in Drosophila melanogaster. *Development* **141**, 3955-3965, (2014).
- 87 Moeller, M. E., Danielsen, E. T., Herder, R., O'Connor, M. B. & Rewitz, K. F. Dynamic feedback circuits function as a switch for shaping a maturation-inducing steroid pulse in Drosophila. *Development* **140**, 4730+, (2013).
- 88 Xiang, Y., Liu, Z. & Huang, X. br regulates the expression of the ecdysone biosynthesis gene npc1. *Dev Biol* **344**, 800-808, (2010).
- 89 King-Jones, K., Charles, J. P., Lam, G. & Thummel, C. S. The ecdysone-induced DHR4 orphan nuclear receptor coordinates growth and maturation in Drosophila. *Cell* **121**, 773-784, (2005).
- 90 Hurban, P. & Thummel, C. S. Isolation and characterization of fifteen ecdysone-inducible Drosophila genes reveal unexpected complexities in ecdysone regulation. *Mol Cell Biol* **13**, 7101-7111, (1993).
- 91 Guittard, E. *et al.* CYP18A1, a key enzyme of Drosophila steroid hormone inactivation, is essential for metamorphosis. *Dev Biol* **349**, 35-45, (2010).

- 92 Hock, T., Cottrill, T., Keegan, J. & Garza, D. The E23 early gene of *Drosophila* encodes an ecdysone-inducible ATP-binding cassette transporter capable of repressing ecdysone-mediated gene activation. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9519-9524, (2000).
- 93 Itoh, T. Q., Tanimura, T. & Matsumoto, A. Membrane-bound transporter controls the circadian transcription of clock genes in *Drosophila*. *Genes to Cells* **16**, 1159-1167, (2011).
- 94 Tennessen, J. M. & Thummel, C. S. Coordinating Growth and Maturation - Insights from *Drosophila*. *Current Biology* **21**, R750-R757, (2011).
- 95 Rewitz, K. F., Yamanaka, N. & O'Connor, M. B. Developmental Checkpoints and Feedback Circuits Time Insect Maturation. *Animal Metamorphosis* **103**, 1-33, (2013).
- 96 Nijhout, H. F. & Williams, C. M. Control of moulting and metamorphosis in the tobacco hornworm, *Manduca sexta* (L.): cessation of juvenile hormone secretion as a trigger for pupation. *J Exp Biol* **61**, 493-501, (1974).
- 97 Nijhout, H. F. & Williams, C. M. Control of moulting and metamorphosis in the tobacco hornworm, *Manduca sexta* (L.): growth of the last-instar larva and the decision to pupate. *J Exp Biol* **61**, 481-491, (1974).
- 98 Shingleton, A. W., Das, J., Vinicius, L. & Stern, D. L. The temporal requirements for insulin signaling during development in *Drosophila*. *PLoS Biology* **3**, 1607-1617, (2005).
- 99 Warren, J. T. *et al.* Discrete pulses of molting hormone, 20-hydroxyecdysone, during late larval development of *Drosophila melanogaster*: correlations with changes in gene activity. *Dev Dyn* **235**, 315-326, (2006).
- 100 Lee, H. O., Davidson, J. M. & Duronio, R. J. Endoreplication: polyploidy with purpose. *Genes & Development* **23**, 2461-2477, (2009).
- 101 Ohhara, Y., Kobayashi, S. & Yamanaka, N. Nutrient-Dependent Endocycling in Steroidogenic Tissue Dictates Timing of Metamorphosis in *Drosophila melanogaster*. *PLoS Genet* **13**, e1006583, (2017).
- 102 Micchelli, C. A. & Perrimon, N. Evidence that stem cells reside in the adult *Drosophila* midgut epithelium. *Nature* **439**, 475-479, (2006).
- 103 Ohlstein, B. & Spradling, A. The adult *Drosophila* posterior midgut is maintained by pluripotent stem cells. *Nature* **439**, 470-474, (2006).
- 104 Galbraith, D. W., Harkins, K. R. & Knapp, S. Systemic Endopolyploidy in *Arabidopsis-thaliana*. *Plant Physiology* **96**, 985-989, (1991).
- 105 Walker, J. D., Oppenheimer, D. G., Concienne, J. & Larkin, J. C. SIAMESE, a gene controlling the endoreduplication cell cycle in *Arabidopsis thaliana* trichomes. *Development* **127**, 3931-3940, (2000).
- 106 Edgar, B. A. & Orr-Weaver, T. L. Endoreplication cell cycles: More for less. *Cell* **105**, 297-306, (2001).
- 107 Lilly, M. A. & Duronio, R. J. New insights into cell cycle control from the *Drosophila* endocycle. *Oncogene* **24**, 2765-2775, (2005).
- 108 Bell, S. P. & Dutta, A. DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**, 333-374, (2002).
- 109 Blow, J. J. Control of chromosomal DNA replication in the early *Xenopus* embryo. *EMBO J* **20**, 3293-3297, (2001).
- 110 Arias, E. E. & Walter, J. C. Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev* **21**, 497-518, (2007).

- 111 Whittaker, A. J., Royzman, I. & Orr-Weaver, T. L. Drosophila Double parked: a conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. *Genes & Development* **14**, 1765-1776, (2000).
- 112 Su, T. T. & O'Farrell, P. H. Chromosome association of minichromosome maintenance proteins in Drosophila endoreplication cycles. *J Cell Biol* **140**, 451-460, (1998).
- 113 Follette, P. J., Duronio, R. J. & O'Farrell, P. H. Fluctuations in cyclin E levels are required for multiple rounds of endocycle S phase in Drosophila. *Curr Biol* **8**, 235-238, (1998).
- 114 Weiss, A., Herzig, A., Jacobs, H. & Lehner, C. F. Continuous Cyclin E expression inhibits progression through endoreduplication cycles in Drosophila. *Current Biology* **8**, 239-242, (1998).
- 115 Zielke, N. *et al.* Control of Drosophila endocycles by E2F and CRL4(CDT2). *Nature* **480**, 123-127, (2011).
- 116 Britton, J. S. & Edgar, B. A. Environmental control of the cell cycle in Drosophila: nutrition activates mitotic and endoreplicative cells by distinct mechanisms. *Development* **125**, 2149-2158, (1998).
- 117 Saucedo, L. J. *et al.* Rheb promotes cell growth as a component of the insulin/TOR signalling network. *Nature Cell Biology* **5**, 566-571, (2003).
- 118 Pierce, S. B. *et al.* dMyc is required for larval growth and endoreplication in Drosophila. *Development* **131**, 2317-2327, (2004).
- 119 Oldham, S., Montagne, J., Radimerski, T., Thomas, G. & Hafen, E. Genetic and biochemical characterization of dTOR, the Drosophila homolog of the target of rapamycin. *Genes & Development* **14**, 2689-2694, (2000).
- 120 Ou, Q. X. *et al.* The Insect Prothoracic Gland as a Model for Steroid Hormone Biosynthesis and Regulation. *Cell Reports* **16**, 247-262, (2016).
- 121 Dickson, B. J. *et al.* A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* **448**, 151-U151, (2007).
- 122 Nieto, M. A. The snail superfamily of zinc-finger transcription factors. *Nat Rev Mol Cell Biol* **3**, 155-166, (2002).
- 123 Gronke, S., Bickmeyer, I., Wunderlich, R., Jackle, H. & Kuhnlein, R. P. Curled encodes the Drosophila homolog of the vertebrate circadian deadenylase Nocturnin. *Genetics* **183**, 219-232, (2009).
- 124 Baggs, J. E. & Green, C. B. Nocturnin, a deadenylase in *Xenopus laevis* retina: a mechanism for posttranscriptional control of circadian-related mRNA. *Curr Biol* **13**, 189-198, (2003).
- 125 Boulay, J. L., Dennefeld, C. & Alberga, A. The Drosophila developmental gene snail encodes a protein with nucleic acid binding fingers. *Nature* **330**, 395-398, (1987).
- 126 Alberga, A., Boulay, J. L., Kempe, E., Dennefeld, C. & Haenlin, M. The Snail Gene Required for Mesoderm Formation in Drosophila Is Expressed Dynamically in Derivatives of All 3 Germ Layers. *Development* **111**, 983-992, (1991).
- 127 Kosman, D., Ip, Y. T., Levine, M. & Arora, K. Establishment of the mesoderm-neuroectoderm boundary in the Drosophila embryo. *Science* **254**, 118-122, (1991).
- 128 Leptin, M. Twist and Snail as Positive and Negative Regulators during Drosophila Mesoderm Development. *Genes & Development* **5**, 1568-1576, (1991).

- 129 Hemavathy, K., Meng, X. & Ip, Y. T. Differential regulation of gastrulation and neuroectodermal gene expression by Snail in the *Drosophila* embryo. *Development* **124**, 3683-3691, (1997).
- 130 Grau, Y., Carteret, C. & Simpson, P. Mutations and Chromosomal Rearrangements Affecting the Expression of Snail, a Gene Involved in Embryonic Patterning in *DROSOPHILA MELANOGASTER*. *Genetics* **108**, 347-360, (1984).
- 131 Simpson, P. Maternal-Zygotic Gene Interactions during Formation of the Dorsoventral Pattern in *Drosophila* Embryos. *Genetics* **105**, 615-632, (1983).
- 132 Qi, D., Bergman, M., Aihara, H., Nibu, Y. & Mannervik, M. *Drosophila* Ebi mediates Snail-dependent transcriptional repression through HDAC3-induced histone deacetylation. *Embo Journal* **27**, 898-909, (2008).
- 133 Barrallo-Gimeno, A. & Nieto, M. A. The Snail genes as inducers of cell movement and survival: implications in development and cancer. *Development* **132**, 3151-3161, (2005).
- 134 Manzanares, M., Locascio, A. & Nieto, M. A. The increasing complexity of the Snail gene superfamily in metazoan evolution. *Trends in Genetics* **17**, 178-181, (2001).
- 135 Materna, S. C., Howard-Ashby, M., Gray, R. F. & Davidson, E. H. The C2H2 zinc finger genes of *Strongylocentrotus purpuratus* and their expression in embryonic development. *Developmental Biology* **300**, 108-120, (2006).
- 136 Barrallo-Gimeno, A. & Nieto, M. A. Evolutionary history of the Snail/Scratch superfamily. *Trends in Genetics* **25**, 248-252, (2009).
- 137 Choi, S. *et al.* Structural basis for the selective nuclear import of the C2H2 zinc-finger protein Snail by importin beta. *Acta Crystallographica Section D-Biological Crystallography* **70**, 1050-1060, (2014).
- 138 Hou, Z. *et al.* 14-3-3 binding sites in the snail protein are essential for snail-mediated transcriptional repression and epithelial-mesenchymal differentiation. *Cancer Res* **70**, 4385-4393, (2010).
- 139 Hemavathy, K., Guru, S. C., Harris, J., Chen, J. D. & Ip, Y. T. Human slug is a repressor that localizes to sites of active transcription. *Molecular and Cellular Biology* **20**, 5087-5095, (2000).
- 140 Dominguez, D. *et al.* Phosphorylation regulates the subcellular location and activity of the snail transcriptional repressor. *Mol Cell Biol* **23**, 5078-5089, (2003).
- 141 Molina-Ortiz, P. *et al.* Characterization of the SNAG and SLUG domains of Snail2 in the repression of E-cadherin and EMT induction: modulation by serine 4 phosphorylation. *PLoS One* **7**, e36132, (2012).
- 142 Notredame, C., Higgins, D. G. & Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* **302**, 205-217, (2000).
- 143 Ashraf, S. I., Hu, X. D., Roote, J. & Ip, Y. T. The mesoderm determinant Snail collaborates with related zinc-finger proteins to control *Drosophila* neurogenesis. *Embo Journal* **18**, 6426-6438, (1999).
- 144 Lin, Y. W. *et al.* The SNAG domain of Snail1 functions as a molecular hook for recruiting lysine-specific demethylase 1. *Embo Journal* **29**, 1803-1816, (2010).
- 145 Molina-Ortiz, P. *et al.* Characterization of the SNAG and SLUG Domains of Snail2 in the Repression of E-Cadherin and EMT Induction: Modulation by Serine 4 Phosphorylation. *PLoS One* **7**, (2012).

- 146 Peinado, H., Ballestar, E., Esteller, M. & Cano, A. Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. *Molecular and Cellular Biology* **24**, 306-319, (2004).
- 147 Langer, E. M. *et al.* Ajuba LIM proteins are Snail/Slug corepressors required for neural crest development in *Xenopus*. *Developmental Cell* **14**, 424-436, (2008).
- 148 Fujiwara, S., Corbo, J. C. & Levine, M. The snail repressor establishes a muscle/notochord boundary in the *Ciona* embryo. *Development* **125**, 2511-2520, (1998).
- 149 Mayor, R., Guerrero, N., Young, R. M., Gomez-Skarmeta, J. L. & Cuellar, C. A novel function for the Xslug gene: control of dorsal mesendoderm development by repressing BMP-4. *Mechanisms of Development* **97**, 47-56, (2000).
- 150 LaBonne, C. & Bronner-Fraser, M. Snail-related transcriptional repressors are required in *Xenopus* for both the induction of the neural crest and its subsequent migration. *Developmental Biology* **221**, 195-205, (2000).
- 151 Cano, A. *et al.* The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nature Cell Biology* **2**, 76-83, (2000).
- 152 Battle, E. *et al.* The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nature Cell Biology* **2**, 84-89, (2000).
- 153 Hemavathy, K., Hu, X. D., Ashraf, S. I., Small, S. J. & Ip, Y. T. The repressor function of Snail is required for *Drosophila* gastrulation and is not replaceable by Escargot or Worniu. *Developmental Biology* **269**, 411-420, (2004).
- 154 Nibu, Y., Zhang, H. L. & Levine, M. Interaction of short-range repressors with *Drosophila* CtBP in the embryo. *Science* **280**, 101-104, (1998).
- 155 Nibu, Y. *et al.* dCtBP mediates transcriptional repression by Knirps, Kruppel and Snail in the *Drosophila* embryo. *Embo Journal* **17**, 7009-7020, (1998).
- 156 Rembold, M. *et al.* A conserved role for Snail as a potentiator of active transcription. *Genes & Development* **28**, 167-181, (2014).
- 157 Hemavathy, K., Ashraf, S. I. & Ip, Y. T. Snail/Slug family of repressors: slowly going into the fast lane of development and cancer. *Gene* **257**, 1-12, (2000).
- 158 Kasai, Y., Nambu, J. R., Lieberman, P. M. & Crews, S. T. Dorsal-ventral patterning in *Drosophila*: DNA binding of snail protein to the single-minded gene. *Proc Natl Acad Sci U S A* **89**, 3414-3418, (1992).
- 159 Roy, H. K. *et al.* Down-regulation of SNAIL suppresses MIN mouse tumorigenesis: modulation of apoptosis, proliferation, and fractal dimension. *Mol Cancer Ther* **3**, 1159-1165, (2004).
- 160 Vega, S. *et al.* Snail blocks the cell cycle and confers resistance to cell death. *Genes & Development* **18**, 1131-1143, (2004).
- 161 Kajita, M., McClinic, K. N. & Wade, P. A. Aberrant expression of the transcription factors snail and slug alters the response to genotoxic stress. *Mol Cell Biol* **24**, 7559-7566, (2004).
- 162 Tribulo, C., Aybar, M. J., Sanchez, S. S. & Mayor, R. A balance between the anti-apoptotic activity of Slug and the apoptotic activity of *msx1* is required for the proper development of the neural crest. *Developmental Biology* **275**, 325-342, (2004).
- 163 Merino, D. *et al.* Pro-apoptotic Bim suppresses breast tumor cell metastasis and is a target gene of SNAI2. *Oncogene* **34**, 3926-3934, (2015).
- 164 Perez-Mancera, P. A. *et al.* Cancer development induced by graded expression of Snail in mice. *Hum Mol Genet* **14**, 3449-3461, (2005).

- 165 Inoue, A. *et al.* Slug, a highly conserved zinc finger transcriptional repressor, protects hematopoietic progenitor cells from radiation-induced apoptosis in vivo. *Cancer Cell* **2**, 279-288, (2002).
- 166 Kurrey, N. K. *et al.* Snail and slug mediate radioresistance and chemoresistance by antagonizing p53-mediated apoptosis and acquiring a stem-like phenotype in ovarian cancer cells. *Stem Cells* **27**, 2059-2068, (2009).
- 167 Escriva, M. *et al.* Repression of PTEN phosphatase by Snail1 transcriptional factor during gamma radiation-induced apoptosis. *Mol Cell Biol* **28**, 1528-1540, (2008).
- 168 Tseng, C. Y., Kao, S. H. & Hsu, H. J. Snail controls proliferation of Drosophila ovarian epithelial follicle stem cells, independently of E-cadherin. *Dev Biol* **414**, 142-148, (2016).
- 169 Larriba, M. J. *et al.* Snail2 cooperates with Snail1 in the repression of vitamin D receptor in colon cancer. *Carcinogenesis* **30**, 1459-1468, (2009).
- 170 Korzelius, J. *et al.* Escargot maintains stemness and suppresses differentiation in Drosophila intestinal stem cells. *Embo Journal* **33**, 2967-2982, (2014).
- 171 Hayashi, S., Hirose, S., Metcalfe, T. & Shirras, A. D. Control of Imaginal Cell-Development by the Escargot Gene of Drosophila. *Development* **118**, 105-115, (1993).
- 172 Fuse, N., Hirose, S. & Hayashi, S. Diploidy of Drosophila Imaginal Cells Is Maintained by a Transcriptional Repressor Encoded by Escargot. *Genes & Development* **8**, 2270-2281, (1994).
- 173 Nakayama, H., Scott, I. C. & Cross, J. C. The transition to endoreduplication in trophoblast giant cells is regulated by the mSNA zinc finger transcription factor. *Developmental Biology* **199**, 150-163, (1998).
- 174 Sanchez-Higueras, C., Sotillos, S. & Hombria, J. C. G. Common Origin of Insect Trachea and Endocrine Organs from a Segmentally Repeated Precursor. *Current Biology* **24**, 76-81, (2014).
- 175 Gratz, S. J. *et al.* Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila. *Genetics* **196**, 961-+, (2014).
- 176 Port, F., Chen, H. M., Lee, T. & Bullock, S. L. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E2967-E2976, (2014).
- 177 Bodenstein, D. The postembryonic development of Drosophila. *Biology of Drosophila.*, 275-367, (1950).
- 178 Ashburner, M., Golic, K. G. & Hawley, R. S. *Drosophila: a laboratory handbook. Second edition.* (2005).
- 179 Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44-57, (2009).
- 180 Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research* **37**, 1-13, (2009).
- 181 Dunipace, L., Ozdemir, A. & Stathopoulos, A. Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. *Development* **138**, 4075-4084, (2011).
- 182 McGuire, S. E., Roman, G. & Davis, R. L. Gene expression systems in Drosophila: a synthesis of time and space. *Trends Genet* **20**, 384-391, (2004).

- 183 Andres, A. J. & Thummel, C. S. Methods for Quantitative-Analysis of Transcription in
Larvae and Prepupae. *Methods in Cell Biology, Vol 44* **44**, 565-573, (1994).
- 184 Osterwalder, T., Yoon, K. S., White, B. H. & Keshishian, H. A conditional tissue-
specific transgene expression system using inducible GAL4. *Proceedings of the National
Academy of Sciences of the United States of America* **98**, 12596-12601, (2001).
- 185 Niwa, R. *et al.* CYP306A1, a cytochrome P450 enzyme, is essential for ecdysteroid
biosynthesis in the prothoracic glands of *Bombyx* and *Drosophila*. *Journal of Biological
Chemistry* **279**, 35942-35949, (2004).
- 186 Lilly, M. A. & Spradling, A. C. The *Drosophila* endocycle is controlled by cyclin E and
lacks a checkpoint ensuring S-phase completion. *Genes & Development* **10**, 2514-2526,
(1996).
- 187 Royzman, I., Whittaker, A. J. & Orr-Weaver, T. L. Mutations in *Drosophila* DP and E2F
distinguish G1-S progression from an associated transcriptional program. *Genes Dev* **11**,
1999-2011, (1997).
- 188 Asano, M., Nevins, J. R. & Wharton, R. P. Ectopic E2F expression induces S phase and
apoptosis in *Drosophila* imaginal discs. *Genes & Development* **10**, 1422-1432, (1996).
- 189 Hara, K. *et al.* Raptor, a binding partner of target of rapamycin (TOR), mediates TOR
action. *Cell* **110**, 177-189, (2002).
- 190 Jacinto, E. & Lorberg, A. TOR regulation of AGC kinases in yeast and mammals.
Biochemical Journal **410**, 19-37, (2008).
- 191 Sarbassov, D. D. *et al.* Rictor, a novel binding partner of mTOR, defines a rapamycin-
insensitive and raptor-independent pathway that regulates the cytoskeleton. *Curr Biol* **14**,
1296-1302, (2004).
- 192 Benmimoun, B., Polesello, C., Waltzer, L. & Haenlin, M. Dual role for Insulin/TOR
signaling in the control of hematopoietic progenitor maintenance in *Drosophila*.
Development **139**, 1713-1717, (2012).
- 193 Tennessen, J. M. & Thummel, C. S. Coordinating growth and maturation - insights from
Drosophila. *Curr Biol* **21**, R750-757.
- 194 Cuyas, E., Corominas-Faja, B., Joven, J. & Menendez, J. A. Cell cycle regulation by the
nutrient-sensing mammalian target of rapamycin (mTOR) pathway. *Methods Mol Biol*
1170, 113-144, (2014).
- 195 Zhang, H., Stallock, J. P., Ng, J. C., Reinhard, C. & Neufeld, T. P. Regulation of cellular
growth by the *Drosophila* target of rapamycin dTOR. *Genes Dev* **14**, 2712-2724, (2000).
- 196 Britton, J. S. & Edgar, B. A. Environmental control of the cell cycle in *Drosophila*:
nutrition activates mitotic and endoreplicative cells by distinct mechanisms.
Development **125**, 2149-2158, (1998).
- 197 Wang, X. & Proud, C. G. The mTOR pathway in the control of protein synthesis.
Physiology (Bethesda) **21**, 362-369, (2006).
- 198 Hietakangas, V. & Cohen, S. M. Regulation of Tissue Growth through Nutrient Sensing.
Annual Review of Genetics **43**, 389-410, (2009).
- 199 Cao, J. *et al.* Insight into Insulin Secretion from Transcriptome and Genetic Analysis of
Insulin-Producing Cells of *Drosophila*. *Genetics* **197**, 175-192, (2014).
- 200 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate - a Practical and
Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series
B-Methodological* **57**, 289-300, (1995).
- 201 Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over
the tree of life. *Nucleic Acids Res* **43**, D447-452, (1990).

- 202 Trowsdale, J., Martin, D., Bicknell, D. & Campbell, I. Alkaline phosphatases. *Biochem Soc Trans* **18**, 178-180, (1990).
- 203 Yang, M. Y., Wang, Z., MacPherson, M., Dow, J. A. & Kaiser, K. A novel *Drosophila* alkaline phosphatase specific to the ellipsoid body of the adult brain and the lower Malpighian (renal) tubule. *Genetics* **154**, 285-297, (2000).
- 204 Schneiderman, H., Young, W. J. & Childs, B. Patterns of alkaline phosphatase in developing *Drosophila*. *Science* **151**, 461-463, (1966).
- 205 Ito, Y. & Sonobe, H. The role of ecdysteroid 22-kinase in the accumulation of ecdysteroids in ovary of silkworm *Bombyx mori*. *Ann N Y Acad Sci* **1163**, 421-424, (2009).
- 206 Zeitlinger, J. *et al.* Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev* **21**, 385-390, (2007).
- 207 Luschnig, S., Batz, T., Armbruster, K. & Krasnow, M. A. serpentine and vermiform encode matrix proteins with chitin binding and deacetylation domains that limit tracheal tube length in *Drosophila*. *Current Biology* **16**, 186-194, (2006).
- 208 Danielsen, E. T. *et al.* A *Drosophila* Genome-Wide Screen Identifies Regulators of Steroid Hormone Production and Developmental Timing. *Developmental Cell* **37**, 558-570, (2016).
- 209 Goyal, L., McCall, K., Agapite, J., Hartweg, E. & Steller, H. Induction of apoptosis by *Drosophila* reaper, hid and grim through inhibition of IAP function. *Embo Journal* **19**, 589-597, (2000).
- 210 Smykal, V. *et al.* Importance of juvenile hormone signaling arises with competence of insect larvae to metamorphose. *Dev Biol* **390**, 221-230, (2014).
- 211 Palli, S. R. *et al.* A nuclear juvenile hormone-binding protein from larvae of *Manduca sexta*: a putative receptor for the metamorphic action of juvenile hormone. *Proc Natl Acad Sci U S A* **91**, 6191-6195, (1994).
- 212 Mirth, C. K. *et al.* Juvenile hormone regulates body size and perturbs insulin signaling in *Drosophila*. *Proc Natl Acad Sci U S A* **111**, 7018-7023, (2014).
- 213 Maki, A. *et al.* Juvenile hormones antagonize ecdysone actions through co-repressor recruitment to EcR/USP heterodimers. *Biochemical and Biophysical Research Communications* **320**, 262-267, (2004).
- 214 Mauhin, V., Lutz, Y., Dennefeld, C. & Alberga, A. Definition of the DNA-Binding Site Repertoire for the *Drosophila* Transcription Factor Snail. *Nucleic Acids Research* **21**, 3951-3957, (1993).
- 215 He, Q. *et al.* High conservation of transcription factor binding and evidence for combinatorial regulation across six *Drosophila* species. *Nat Genet* **43**, 414-420, (2011).
- 216 Lin, S. L., Ewen-Campen, B., Ni, X. C., Housden, B. E. & Perrimon, N. In Vivo Transcriptional Activation Using CRISPR/Cas9 in *Drosophila*. *Genetics* **201**, 433-+, (2015).
- 217 Mathew, S. J., Rembold, M. & Leptin, M. Role for Traf4 in Polarizing Adherens Junctions as a Prerequisite for Efficient Cell Shape Changes. *Molecular and Cellular Biology* **31**, 4978-4993, (2011).
- 218 Willsey, H. R. *et al.* Localized JNK signaling regulates organ size during development. *Elife* **5**, (2016).
- 219 Ruan, W., Srinivasan, A., Lin, S., Kara k, I. & Barker, P. A. Eiger-induced cell death relies on Rac1-dependent endocytosis. *Cell Death Dis* **7**, e2181, (2016).

- 220 Shklover, J., Levy-Adam, F. & Kurant, E. The role of Drosophila TNF Eiger in developmental and damage-induced neuronal apoptosis. *FEBS Lett* **589**, 871-879, (2015).
- 221 Bastock, R. & Strutt, D. The planar polarity pathway promotes coordinated cell migration during Drosophila oogenesis. *Development* **134**, 3055-3064, (2007).
- 222 Katanaev, V. L. & Tomlinson, A. Dual roles for the trimeric G protein Go in asymmetric cell division in Drosophila. *Proc Natl Acad Sci U S A* **103**, 6524-6529, (2006).
- 223 Yao, Y. *et al.* Antagonistic roles of Wnt5 and the Drl receptor in patterning the Drosophila antennal lobe. *Nat Neurosci* **10**, 1423-1432, (2007).
- 224 Lahaye, L. L., Wouda, R. R., de Jong, A. W., Fradkin, L. G. & Noordermeer, J. N. WNT5 interacts with the Ryk receptors doughnut and derailed to mediate muscle attachment site selection in Drosophila melanogaster. *PLoS One* **7**, e32297, (2012).
- 225 Stevens, L. M., Frohnhof, H. G., Klingler, M. & Nusslein-Volhard, C. Localized requirement for torso-like expression in follicle cells for development of terminal anlagen of the Drosophila embryo. *Nature* **346**, 660-663, (1990).
- 226 Johnson, T. K. *et al.* Torso-like functions independently of Torso to regulate Drosophila growth and developmental timing. *Proc Natl Acad Sci U S A* **110**, 14688-14692, (2013).
- 227 Duncan, E. J., Johnson, T. K., Whisstock, J. C., Warr, C. G. & Dearden, P. K. Capturing embryonic development from metamorphosis: how did the terminal patterning signalling pathway of Drosophila evolve? *Current Opinion in Insect Science* **1**, 45-51, (2014).
- 228 Grillo, M., Furriols, M., de Miguel, C., Franch-Marro, X. & Casanova, J. Conserved and divergent elements in Torso RTK activation in Drosophila development. *Sci Rep* **2**, 762, (2012).
- 229 Parra, R. G., Rohr, C. O., Koile, D., Perez-Castro, C. & Yankilevich, P. INSECT 2.0: a web-server for genome-wide cis-regulatory modules prediction. *Bioinformatics* **32**, 1229-1231, (2016).
- 230 Ip, Y. T., Maggert, K. & Levine, M. Uncoupling gastrulation and mesoderm differentiation in the Drosophila embryo. *EMBO J* **13**, 5826-5834, (1994).
- 231 Potier, D. *et al.* Identification of cis-regulatory modules encoding temporal dynamics during development. *BMC Genomics* **15**, 534, (2014).
- 232 Guest, S. T. *et al.* A protein network-guided screen for cell cycle regulators in Drosophila. *BMC Syst Biol* **5**, 65, (2011).
- 233 Thomer, M., May, N. R., Aggarwal, B. D., Kwok, G. & Calvi, B. R. Drosophila double-parked is sufficient to induce re-replication during development and is regulated by cyclin E/CDK2. *Development* **131**, 4807-4818, (2004).
- 234 Barnum, K. J. & O'connell, M. J. Cell Cycle Regulation by Checkpoints. *Cell Cycle Control: Mechanisms and Protocols, 2nd Edition* **1170**, 29-40, (2014).
- 235 Pucci, B., Kasten, M. & Giordano, A. Cell cycle and apoptosis. *Neoplasia* **2**, 291-299, (2000).
- 236 Alenzi, F. Q. B. Links between apoptosis, proliferation and the cell cycle. *British Journal of Biomedical Science* **61**, 99-102, (2004).
- 237 Hatzold, J. & Conradt, B. Control of apoptosis by asymmetric cell division. *PLoS Biology* **6**, 771-784, (2008).
- 238 Dominguez, D. *et al.* Phosphorylation regulates the subcellular location and activity of the snail transcriptional repressor. *Molecular and Cellular Biology* **23**, 5078-5089, (2003).

- 239 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948, (2007).
- 240 Zhou, B. H. P. *et al.* Dual regulation of Snail by GSK-3 beta-mediated phosphorylation in control of epithelial-mesenchymal transition. *Nature Cell Biology* **6**, 931+, (2004).
- 241 Xue, Y. *et al.* GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. *Mol Cell Proteomics* **7**, 1598-1608, (2008).
- 242 Xue, Y. *et al.* GPS: a comprehensive www server for phosphorylation sites prediction. *Nucleic Acids Res* **33**, W184-187, (2005).
- 243 Xue, Y. *et al.* GPS 2.1: enhanced prediction of kinase-specific phosphorylation sites with an algorithm of motif length selection. *Protein Eng Des Sel* **24**, 255-260, (2010).
- 244 Jin, H. Y. *et al.* Snail is critical for tumor growth and metastasis of ovarian carcinoma. *International Journal of Cancer* **126**, 2102-2111, (2010).
- 245 Yanagawa, J. *et al.* Snail Promotes CXCR2 Ligand-Dependent Tumor Progression in Non-Small Cell Lung Carcinoma. *Clinical Cancer Research* **15**, 6820-6829, (2009).
- 246 Kosaka, T. *et al.* Expression of Snail in Upper Urinary Tract Urothelial Carcinoma: Prognostic Significance and Implications for Tumor Invasion. *Clinical Cancer Research* **16**, 5814-5823, (2010).
- 247 Yang, M. H. *et al.* Comprehensive Analysis of the Independent Effect of Twist and Snail in Promoting Metastasis of Hepatocellular Carcinoma. *Hepatology* **50**, 1464-1474, (2009).
- 248 Moody, S. E. *et al.* The transcriptional repressor snail promotes mammary tumor recurrence. *Cancer Cell* **8**, 197-209, (2005).
- 249 Green, C. B. & Besharse, J. C. Identification of a novel vertebrate circadian clock-regulated gene encoding the protein nocturnin. *Proc Natl Acad Sci U S A* **93**, 14884-14888, (1996).
- 250 Wang, Y. *et al.* Rhythmic expression of Nocturnin mRNA in multiple tissues of the mouse. *BMC Dev Biol* **1**, 9, (2001).
- 251 Rosbash, M. *et al.* Dissecting differential gene expression within the circadian neuronal circuit of Drosophila. *Nature Neuroscience* **13**, 60-U220, (2010).
- 252 Kawai, M. *et al.* A circadian-regulated gene, Nocturnin, promotes adipogenesis by stimulating PPAR-gamma nuclear translocation. *Proc Natl Acad Sci U S A* **107**, 10508-10513, (2010).
- 253 Wahle, E. & Winkler, G. S. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* **1829**, 561-570, (2013).
- 254 Goldstrohm, A. C. & Wickens, M. Multifunctional deadenylase complexes diversify mRNA control. *Nat Rev Mol Cell Biol* **9**, 337-344, (2008).
- 255 Temme, C. *et al.* Subunits of the Drosophila CCR4-NOT complex and their roles in mRNA deadenylation. *Rna-a Publication of the Rna Society* **16**, 1356-1370, (2010).
- 256 Temme, C., Simonelig, M. & Wahle, E. Deadenylation of mRNA by the CCR4-NOT complex in Drosophila: molecular and developmental aspects. *Front Genet* **5**, 143.
- 257 Temme, C., Zaessinger, S., Meyer, S., Simonelig, M. & Wahle, E. A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in Drosophila. *Embo Journal* **23**, 2862-2871, (2004).
- 258 Sakai, A., Chibazakura, T., Shimizu, Y. & Hishinuma, F. Molecular analysis of POP2 gene, a gene required for glucose-derepression of gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **20**, 6227-6233, (1992).

- 259 Liu, H. Y. *et al.* DBF2, a cell cycle-regulated protein kinase, is physically and functionally associated with the CCR4 transcriptional regulatory complex. *Embo Journal* **16**, 5289-5298, (1997).
- 260 Petit, A. P. *et al.* The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Research* **40**, 11058-11072, (2012).
- 261 Barckmann, B. & Simonelig, M. Control of maternal mRNA stability in germ cells and early embryos. *Biochim Biophys Acta* **1829**, 714-724, (2013).
- 262 Joly, W., Chartier, A., Rojas-Rios, P., Busseau, I. & Simonelig, M. The CCR4 deadenylase acts with Nanos and Pumilio in the fine-tuning of Mei-P26 expression to promote germline stem cell self-renewal. *Stem Cell Reports* **1**, 411-424, (2013).
- 263 Zaessinger, S., Busseau, I. & Simonelig, M. Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. *Development* **133**, 4573-4583, (2006).
- 264 Chicoine, J. *et al.* Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Dev Cell* **13**, 691-704, (2007).
- 265 Morris, J. Z., Hong, A., Lilly, M. A. & Lehmann, R. twin, a CCR4 homolog, regulates cyclin poly(A) tail length to permit Drosophila oogenesis. *Development* **132**, 1165-1174, (2005).
- 266 Niinuma, S., Fukaya, T. & Tomari, Y. CCR4 and CAF1 deadenylases have an intrinsic activity to remove the post-poly(A) sequence. *Rna* **22**, 1550-1559, (2016).
- 267 Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* **2**, 135-140, (1998).
- 268 Zhao, J., Hyman, L. & Moore, C. Formation of mRNA 3' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiology and Molecular Biology Reviews* **63**, 405-+, (1999).
- 269 Gorgoni, B. & Gray, N. K. The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: a developmental perspective. *Brief Funct Genomic Proteomic* **3**, 125-141, (2004).
- 270 Doidge, R., Mittal, S., Aslam, A. & Winkler, G. S. Deadenylation of cytoplasmic mRNA by the mammalian Ccr4-Not complex. *Biochem Soc Trans* **40**, 896-901, (2012).
- 271 Kadyrova, L. Y., Habara, Y., Lee, T. H. & Wharton, R. P. Translational control of maternal cyclin B mRNA by nanos in the Drosophila germline. *Development* **134**, 1519-1527, (2007).
- 272 Virtanen, A., Henriksson, N., Nilsson, P. & Nissbeck, M. Poly(A)-specific ribonuclease (PARN): an allosterically regulated, processive and mRNA cap-interacting deadenylase. *Crit Rev Biochem Mol Biol* **48**, 192-209, (2013).
- 273 Boeck, R. *et al.* The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. *Journal of Biological Chemistry* **271**, 432-438, (1996).
- 274 Brown, C. E., Tarun, S. Z., Boeck, R. & Sachs, A. B. PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **16**, 5744-5753, (1996).
- 275 Tucker, M. *et al.* The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell* **104**, 377-386, (2001).

- 276 Morris, J. Z., Hong, A., Lilly, M. A. & Lehmann, R. twin, a CCR4 homolog, regulates cyclin poly(A) tail length to permit *Drosophila* oogenesis. *Development* **132**, 1165-1174, (2005).
- 277 Collart, M. A. The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdisciplinary Reviews-Rna* **7**, 438-454, (2016).
- 278 Miller, J. E. & Reese, J. C. Ccr4-Not complex: the control freak of eukaryotic cells. *Critical Reviews in Biochemistry and Molecular Biology* **47**, 315-333, (2012).
- 279 Shirai, Y. T., Suzuki, T., Morita, M., Takahashi, A. & Yamamoto, T. Multifunctional roles of the mammalian CCR4-NOT complex in physiological phenom. *Frontiers in Genetics* **5**, (2014).
- 280 Deluen, C. *et al.* The Ccr4-not complex and yTAF1 (yTaf(II)130p/yTaf(II)145p) show physical and functional interactions. *Molecular and Cellular Biology* **22**, 6735-6749, (2002).
- 281 Swanson, M. J. *et al.* A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo. *Molecular and Cellular Biology* **23**, 2800-2820, (2003).
- 282 Venters, B. J. *et al.* A Comprehensive Genomic Binding Map of Gene and Chromatin Regulatory Proteins in *Saccharomyces*. *Molecular Cell* **41**, 480-492, (2011).
- 283 Lenssen, E. *et al.* The Ccr4-not complex independently controls both Msn2-dependent transcriptional activation - via a newly identified Glc7/Bud14 type I protein phosphatase module - and TFIID promoter distribution. *Molecular and Cellular Biology* **25**, 488-498, (2005).
- 284 Prevot, D. *et al.* Relationships of the antiproliferative proteins BTG1 and BTG2 with CAF1, the human homolog of a component of the yeast CCR4 transcriptional complex - Involvement in estrogen receptor alpha signaling pathway. *Journal of Biological Chemistry* **276**, 9640-9648, (2001).
- 285 Morel, A. P. *et al.* BTG2 antiproliferative protein interacts with the human CCR4 complex existing in vivo in three cell-cycle-regulated forms. *Journal of Cell Science* **116**, 2929-2936, (2003).
- 286 Okada, H., Schittenhelm, R. B., Straessle, A. & Hafen, E. Multi-functional regulation of 4E-BP gene expression by the Ccr4-Not complex. *PLoS One* **10**, e0113902, (2015).
- 287 Janning, W. FlyView, a *Drosophila* image database, and other *Drosophila* databases. *Semin Cell Dev Biol* **8**, 469-475, (1997).
- 288 Moriyama, E. N. & Powell, J. R. Codon usage bias and tRNA abundance in *Drosophila*. *J Mol Evol* **45**, 514-523, (1997).
- 289 Schulz, J. G., David, G. & Hassan, B. A. A novel method for tissue-specific RNAi rescue in *Drosophila*. *Nucleic Acids Res* **37**, e93, (2009).
- 290 Bialecki, M., Shilton, A., Fichtenberg, C., Segraves, W. A. & Thummel, C. S. Loss of the ecdysteroid-inducible E75A orphan nuclear receptor uncouples molting from metamorphosis in *Drosophila*. *Dev Cell* **3**, 209-220, (2002).
- 291 Sliter TJ, G. L. Developmental arrest and ecdysteroid deficiency resulting from mutations at the dre4 locus of *Drosophila*. *Genetics*, 555-568, (1992).
- 292 Venkatesh K, H. G. Disruption of the IP3 receptor gene of *Drosophila* affects larval metamorphosis and ecdysone release. *Curr. Biol.* **7**, 500-509, (1997).
- 293 Burgermeister, E. & Seger, R. MAPK kinases as nucleo-cytoplasmic shuttles for PPAR gamma. *Cell Cycle* **6**, 1539-1548, (2007).
- 294 Johnson, M. A. *et al.* Amino Acid Starvation Has Opposite Effects on Mitochondrial and Cytosolic Protein Synthesis. *PLoS One* **9**, (2014).

Appendix

Table A-1. Top 100 differentially expressed genes in the PG>*sna*-RNAi ring glands

Top 50 downregulated				
Flybase ID	Gene symbol	Note	Fold change	<i>p</i> value
<i>FBgn0266261</i>	<i>CG44956</i>		-27.7	0.145
<i>FBgn0031512</i>	<i>CG15404</i>		-19.9	0.079
<i>FBgn0038114</i>	<i>CG11670</i>		-16.2	0.013
<i>FBgn0040849</i>	<i>Ir41a</i>	<i>Ionotropic receptor 41a;</i>	-14.0	0.056
<i>FBgn0035673</i>	<i>CG6602</i>		-13.9	0.160
<i>FBgn0041581</i>	<i>AttB</i>	<i>Attacin-B</i>	-13.0	0.094
<i>FBgn0030260</i>	<i>CG1537</i>		-12.7	0.056
<i>FBgn0030259</i>	<i>CG1545</i>		-12.4	0.099
<i>FBgn0026878</i>	<i>CG4325</i>		-11.8	0.128
<i>FBgn0039030</i>	<i>CG6660</i>		-11.0	0.079
<i>FBgn0040565</i>	<i>CG7606</i>		-10.3	0.178
<i>FBgn0038449</i>	<i>CG17562</i>		-10.3	0.235
<i>FBgn0259697</i>	<i>nvd</i>	<i>neverland</i>	-9.4	0.168
<i>FBgn0033702</i>	<i>CG8854</i>		-9.2	0.003
<i>FBgn0063368</i>	<i>Gpb5</i>	<i>Glycoprotein hormone beta 5 ortholog (H. sapiens)</i>	-9.2	0.045
<i>FBgn0038523</i>	<i>CG7587</i>		-9.2	0.083
<i>FBgn0264478</i>	<i>CG43886</i>		-9.0	0.111
<i>FBgn0086917</i>	<i>spok</i>	<i>spookier</i>	-8.9	0.106
<i>FBgn0035028</i>	<i>Start1</i>		-8.8	0.026
<i>FBgn0085261</i>	<i>CG34232</i>		-8.2	0.210
<i>FBgn0038239</i>	<i>CG14850</i>		-7.7	0.189
<i>FBgn0001090</i>	<i>bnb</i>	<i>bangles and beads</i>	-7.7	0.180
<i>FBgn0039073</i>	<i>CG4408</i>		-7.4	0.121

FBgn0032083	CG9541		-7.2	0.097
FBgn0032252	loh	lonely heart	-7.1	0.078
FBgn0001229	Hsp67Bc	Heat shock gene 67Bc	-7.1	0.009
FBgn0032084	CG13101		-7.0	0.160
FBgn0037618	CG11762		-6.9	0.090
FBgn0262722	CG43166		-6.8	0.043
FBgn0052405	Cpr65Av	Cuticular protein 65Av	-6.8	0.030
FBgn0261612	CngA	Cyclic nucleotide-gated ion channel subunit A	-6.7	0.026
FBgn0037547	CG7910		-6.7	0.196
FBgn0029821	CG4020		-6.3	0.089
FBgn0035813	CG8492		-6.0	0.105
FBgn0262004	CG42822		-6.0	0.106
FBgn0003312	sad	shadow	-5.9	0.172
FBgn0029838	CG4666		-5.9	0.244
FBgn0261681	CG42728		-5.7	0.000
FBgn0001228	Hsp67Bb	Heat shock gene 67Bb	-5.6	0.007
FBgn0052368	CG32368		-5.5	0.149
FBgn0032280	CG17105		-5.5	0.102
FBgn0004959	phm	phantom	-5.4	0.099
FBgn0003733	tor	torso	-5.3	0.142
FBgn0050479	CG30479		-5.0	0.120
FBgn0033817	GstE14	Glutathione S transferase E14	-4.7	0.083
FBgn0032153	CG4537		-4.7	0.018
FBgn0085428	Nox	NADPH oxidase	-4.7	0.126
FBgn0029804	CG3097		-4.6	0.192
FBgn0004592	Eig71Ee	Ecdysone-induced gene 71Ee	-4.6	0.216
FBgn0052249	CG32249		-4.6	0.056

Top 50 upregulated

Flybase ID	Gene symbol	Note	Fold change	p value
FBgn0039435	TwdlP	TwedleP	35.4	0.112

<i>FBgn0039438</i>	<i>TwdlO</i>	<i>TweedleO</i>	30.2	0.028
<i>FBgn0031678</i>	<i>CG31918</i>		29.9	0.178
<i>FBgn0036157</i>	<i>CG7560</i>		25.2	0.004
<i>FBgn0011283</i>	<i>Obp28a</i>	<i>Odorant-binding protein 28a</i>	18.8	0.066
<i>FBgn0033820</i>	<i>CG4716</i>		16.5	0.144
<i>FBgn0046878</i>	<i>Obp83cd</i>	<i>Odorant-binding protein 83cd</i>	15.0	0.020
<i>FBgn0004427</i>	<i>LysD</i>	<i>Lysozyme D</i>	12.7	0.121
<i>FBgn0032472</i>	<i>CG9928</i>		12.3	0.055
<i>FBgn0037230</i>	<i>CG9780</i>		11.7	0.015
<i>FBgn0001254</i>	<i>ImpE2</i>	<i>Ecdysone-inducible gene E2</i>	11.5	0.083
<i>FBgn0041180</i>	<i>Tep4</i>	<i>Thioester-containing protein 4</i>	11.0	0.010
<i>FBgn0039769</i>	<i>CG15534</i>		10.9	0.176
<i>FBgn0035359</i>	<i>CG1143</i>		9.8	0.012
<i>FBgn0052574</i>	<i>Twldalpha</i>	<i>Tweedlealpha</i>	8.8	0.036
<i>FBgn0035620</i>	<i>CG5150</i>		8.8	0.077
<i>FBgn0040553</i>	<i>CG14374</i>		8.6	0.018
<i>FBgn0034335</i>	<i>GstE1</i>	<i>Glutathione S transferase E1</i>	8.5	0.025
<i>FBgn0004428</i>	<i>LysE</i>	<i>Lysozyme E</i>	8.5	0.204
<i>FBgn0051463</i>	<i>CG31463</i>		8.2	0.122
<i>FBgn0033720</i>	<i>CG13160</i>		7.4	0.009
<i>FBgn0038148</i>	<i>CG14377</i>		7.2	0.080
<i>FBgn0036607</i>	<i>CG13059</i>		7.1	0.033
<i>FBgn0036056</i>	<i>CG6709</i>		7.0	0.003
<i>FBgn0033222</i>	<i>CG12824</i>		6.9	0.166
<i>FBgn0053966</i>	<i>CG33966</i>		6.7	0.045
<i>FBgn0000592</i>	<i>Est-6</i>	<i>Esterase 6</i>	6.5	0.046
<i>FBgn0037635</i>	<i>CG9837</i>		6.4	0.043
<i>FBgn0029765</i>	<i>CG16756</i>		6.2	0.042
<i>FBgn0034479</i>	<i>CG8654</i>		5.9	0.012
<i>FBgn0260756</i>	<i>CG42554</i>		5.9	0.077
<i>FBgn0033207</i>	<i>CG12826</i>		5.8	0.148

<i>FBgn0033464</i>	<i>CG1441</i>		5.8	0.023
<i>FBgn0034712</i>	<i>CG3264</i>		5.7	0.067
<i>FBgn0039474</i>	<i>CG6283</i>		5.4	0.099
<i>FBgn0038257</i>	<i>smp-30</i>	<i>Senescence marker protein-30</i>	5.3	0.001
<i>FBgn0051148</i>	<i>CG31148</i>		5.2	0.054
<i>FBgn0004777</i>	<i>Ccp84Ag</i>		5.2	0.143
<i>FBgn0013308</i>	<i>Odc2</i>	<i>Ornithine decarboxylase 2</i>	5.2	0.145
<i>FBgn0053965</i>	<i>CG33965</i>		5.2	0.248
<i>FBgn0030009</i>	<i>CG15336</i>		5.2	0.131
<i>FBgn0261563</i>	<i>wb</i>	<i>wing blister</i>	5.1	0.039
<i>FBgn0040606</i>	<i>CG6503</i>		5.0	0.116
<i>FBgn0031865</i>	<i>Nha1</i>	<i>Na^[+]/H^[+] hydrogen antiporter 1</i>	5.0	0.164
<i>FBgn0031560</i>	<i>CG16713</i>		4.9	0.025
<i>FBgn0085325</i>	<i>CG34296</i>		4.7	0.163
<i>FBgn0036469</i>	<i>CG18649</i>		4.7	0.192
<i>FBgn0262683</i>	<i>CG43153</i>		4.6	0.020
<i>FBgn0031232</i>	<i>CG11617</i>		4.6	0.020
<i>FBgn0036264</i>	<i>CG11529</i>		4.5	0.035

Table A-2. Top 100 differentially expressed genes in the *hs>sna*-cDNA ring glands

Top 50 downregulated

Flybase ID	Gene symbol	Note	Fold change	<i>p</i> value
FBgn0259977	<i>Tdc1</i>	<i>Tyrosine decarboxylase 1</i>	-49.2	0.001
FBgn0261990	<i>CG42808</i>		-47.0	0.004
FBgn0033033	<i>scaf</i>	<i>scarface</i>	-28.1	0.016
FBgn0036106	<i>CG6409</i>		-27.6	0.010
FBgn0032184	<i>CG13135</i>		-24.5	0.018
FBgn0028855	<i>CG15282</i>		-23.8	0.053
FBgn0085223	<i>CG34194</i>		-23.1	0.005
FBgn0037227	<i>TwdlV</i>	<i>TweedleV</i>	-22.3	0.032
FBgn0000449	<i>dib</i>	<i>disembodied</i>	-17.5	0.028
FBgn0029907	<i>Atx-1</i>	<i>Ataxin 1</i>	-17.2	0.004
FBgn0053117	<i>Victoria</i>		-15.8	0.064
FBgn0029681	<i>CG15239</i>		-14.8	0.026
FBgn0034756	<i>Cyp6d2</i>		-14.5	0.016
FBgn0034883	<i>CG17664</i>		-14.0	0.012
FBgn0037534	<i>CG2781</i>		-12.3	0.020
FBgn0039178	<i>CG6356</i>		-12.2	0.001
FBgn0036461	<i>Zip71B</i>	<i>Zinc/iron regulated transporter-related protein 71B</i>	-11.6	0.039
FBgn0033307	<i>CG14752</i>		-9.9	0.005
FBgn0032283	<i>CG7296</i>		-9.8	0.107
FBgn0085330	<i>CG34301</i>		-9.4	0.015
FBgn0036393	<i>CG17362</i>		-8.8	0.125
FBgn0265512	<i>mlt</i>	<i>mulet</i>	-8.7	0.020
FBgn0050115	<i>GEFmeso</i>	<i>Guanine nucleotide exchange factor in mesoderm</i>	-8.1	0.009
FBgn0029723	<i>Proc-R</i>	<i>Proctolin receptor</i>	-8.1	0.010
FBgn0035084	<i>CG15861</i>		-8.0	0.036
FBgn0050345	<i>CG30345</i>		-7.7	0.015

FBgn0003068	<i>per</i>	<i>period</i>	-7.6	0.032
FBgn0085345	<i>CG34316</i>		-7.6	0.043
FBgn0003980	<i>Vm26Ab</i>	<i>Vitelline membrane 26Ab</i>	-7.6	0.059
FBgn0030260	<i>CG1537</i>		-7.5	0.137
FBgn0035089	<i>Phk-3</i>	<i>Pherokine 3</i>	-7.4	0.005
FBgn0039758	<i>CG9737</i>		-7.2	0.002
FBgn0260429	<i>CG42524</i>		-7.2	0.003
FBgn0035049	<i>Mmp1</i>	<i>Matrix metalloproteinase 1</i>	-6.7	0.002
FBgn0038160	<i>CG9759</i>		-6.5	0.056
FBgn0036596	<i>CG13045</i>		-6.4	0.074
FBgn0011225	<i>jar</i>	<i>jaguar</i>	-6.4	0.016
FBgn0035546	<i>CG11345</i>		-6.4	0.182
FBgn0034162	<i>CG6426</i>		-6.3	0.010
FBgn0035551	<i>CG7465</i>		-6.3	0.169
FBgn0261560	<i>Thor</i>		-6.2	0.011
FBgn0036732	<i>Oatp74D</i>	<i>Organic anion transporting polypeptide 74D</i>	-6.2	0.005
FBgn0031011	<i>CG8034</i>		-6.2	0.001
FBgn0035968	<i>Slc45-1</i>	<i>Slc45 ortholog 1</i>	-6.2	0.033
FBgn0030060	<i>CG2004</i>		-6.1	0.001
FBgn0003060	<i>CG9757</i>		-6.1	0.037
FBgn0004893	<i>bowl</i>	<i>brother of odd with entrails limited</i>	-6.1	0.011
FBgn0250839	<i>CG2016</i>		-6.1	0.024
FBgn0052820	<i>CG32820</i>		-6.1	0.020
FBgn0052819	<i>CG32819</i>		-6.1	0.020

Top 50 upregulated

Flybase ID	Gene symbol	Note	Fold change	p value
FBgn0052071	<i>CG32071</i>		32.2	0.144
FBgn0003374	<i>Sgs4</i>	<i>Salivary gland secretion 4</i>	28.0	0.083
FBgn0003375	<i>Sgs5</i>	<i>Salivary gland secretion 5</i>	26.6	0.089
FBgn0003373	<i>Sgs3</i>	<i>Salivary gland secretion 3</i>	26.4	0.141

FBgn0004592	<i>Eig71Ee</i>	<i>Ecdysone-induced gene 71Ee</i>	20.0	0.062
FBgn0034328	<i>IM23</i>	<i>Immune induced molecule 23</i>	17.0	0.011
FBgn0037114	<i>Cpr78E</i>	<i>Cuticular protein 78E</i>	14.4	0.074
FBgn0000639	<i>Fbp1</i>	<i>Fat body protein 1</i>	14.3	0.120
FBgn0040582	<i>CG5791</i>		13.9	0.018
FBgn0003372	<i>Sgs1</i>	<i>Salivary gland secretion 1</i>	12.8	0.139
FBgn0052625	<i>CG32625</i>		12.2	0.027
FBgn0003377	<i>Sgs7</i>	<i>Salivary gland secretion 7</i>	10.9	0.106
FBgn0067905	<i>IM14</i>	<i>Immune induced molecule 14</i>	9.9	0.080
FBgn0003378	<i>Sgs8</i>	<i>Salivary gland secretion 8</i>	9.8	0.104
FBgn0040653	<i>IM4</i>	<i>Immune induced molecule 4</i>	9.2	0.042
FBgn0052703	<i>Erk7</i>	<i>Extracellularly regulated kinase 7</i>	9.1	0.060
FBgn0085319	<i>CG34290</i>		8.6	0.017
FBgn0032652	<i>CG6870</i>		8.6	0.020
FBgn0020445	<i>E23</i>	<i>Early gene at 23</i>	7.9	0.010
FBgn0265266	<i>CG13639</i>		7.5	0.074
FBgn0040743	<i>CG15919</i>		7.5	0.003
FBgn0035711	<i>CG8519</i>		7.4	0.014
FBgn0032835	<i>CG16772</i>		7.1	0.035
FBgn0030102	<i>CG12119</i>		6.9	0.006
FBgn0037503	<i>CG14598</i>		6.7	0.101
FBgn0030105	<i>CG15369</i>		6.7	0.055
FBgn0032452	<i>CG15484</i>		6.0	0.060
FBgn0034317	<i>CG14499</i>		5.9	0.154
FBgn0031559	<i>CG3513</i>		5.9	0.167
FBgn0261570	<i>CG42684</i>		5.9	0.005
FBgn0263621	<i>CG43630</i>		5.8	0.049
FBgn0028400	<i>Syt4</i>	<i>Synaptotagmin 4</i>	5.7	0.019
FBgn0038641	<i>CG7708</i>		5.7	0.001
FBgn0039438	<i>TwldO</i>	<i>TweedleO</i>	5.7	0.071

FBgn0033137	<i>Tsp42Ep</i>	<i>Tetraspanin 42Ep</i>	5.4	0.065
FBgn0004872	<i>piwi</i>		5.3	0.157
FBgn0001145	<i>Gs2</i>	<i>Glutamine synthetase 2</i>	5.3	0.043
FBgn0004876	<i>cdi</i>	<i>center divider</i>	5.3	0.022
FBgn0034331	<i>CG15067</i>		5.3	0.100
FBgn0014865	<i>Mtk</i>	<i>Metchnikowin</i>	5.2	0.096
FBgn0029824	<i>CG3726</i>		4.9	0.010
FBgn0035952	<i>CG5280</i>		4.8	0.034
FBgn0028537	<i>CG31775</i>		4.8	0.142
FBgn0260954	<i>CG42586</i>		4.8	0.142
FBgn0010620	<i>CG10939</i>		4.7	0.024
FBgn0085285	<i>CG34256</i>		4.7	0.044
FBgn0014396	<i>tim</i>	<i>timeless</i>	4.6	0.043
FBgn0262794	<i>CG43175</i>		4.6	0.140
FBgn0038631	<i>CG7695</i>		4.6	0.029

Table A-3. Top 100 differentially expressed genes in the PG>Pop2-RNAi ring glands

Top 50 downregulated

Flybase ID	Gene symbol	Note	Fold change	p value
<i>FBgn0052557</i>	<i>CG32557</i>		-33.3	0.015
<i>FBgn0261990</i>	<i>CG42808</i>		-29.1	0.002
<i>FBgn0032805</i>	<i>CG10337</i>		-19.0	0.050
<i>FBgn0013954</i>	<i>FK506-bp2</i>	<i>FK506-binding protein 2</i>	-16.0	0.043
<i>FBgn0023550</i>	<i>CG18031</i>		-15.9	0.024
<i>FBgn0035084</i>	<i>CG15861</i>		-15.6	0.003
<i>FBgn0264743</i>	<i>CG44001</i>		-14.8	0.100
<i>FBgn0036766</i>	<i>CG5506</i>		-14.5	0.154
<i>FBgn0038645</i>	<i>CG7714</i>		-14.4	0.228
<i>FBgn0035528</i>	<i>CG15012</i>		-13.7	0.010
<i>FBgn0261599</i>	<i>RpS29</i>	<i>Ribosomal protein S29</i>	-13.5	0.060
<i>FBgn0038331</i>	<i>Ccm3</i>	<i>Cerebral cavernous malformation 3 ortholog</i>	-12.2	0.004
<i>FBgn0003514</i>	<i>sqh</i>	<i>spaghetti squash</i>	-12.1	0.071
<i>FBgn0026602</i>	<i>Ady43A</i>		-11.7	0.006
<i>FBgn0266490</i>	<i>CG45089</i>		-11.5	0.047
<i>FBgn0085261</i>	<i>CG34232</i>		-11.3	0.017
<i>FBgn0036511</i>	<i>CG6498</i>	<i>drop out</i>	-10.7	0.008
<i>FBgn0029804</i>	<i>CG3097</i>		-10.6	0.072
<i>FBgn0034440</i>	<i>CG10073</i>		-10.3	0.214
<i>FBgn0039486</i>	<i>CG6074</i>		-10.3	0.103
<i>FBgn0028663</i>	<i>VhaM9.7-b</i>	<i>Vacuolar H[+] ATPase M9.7 subunit b</i>	-10.3	0.023
<i>FBgn0040636</i>	<i>CG13255</i>		-10.1	0.040
<i>FBgn0260747</i>	<i>CG5010</i>		-10.0	0.038
<i>FBgn0004921</i>	<i>Ggamma1</i>	<i>G protein gamma 1</i>	-10.0	0.016
<i>FBgn0038804</i>	<i>CG10877</i>		-9.6	0.001
<i>FBgn0031805</i>	<i>CG9505</i>		-9.0	0.007
<i>FBgn0029838</i>	<i>CG4666</i>		-8.9	0.086

FBgn0086917	<i>spok</i>	<i>spookier</i>	-8.7	0.020
FBgn0265274	<i>Inx3</i>	<i>Innexin 3</i>	-8.4	0.046
FBgn0026878	<i>CG4325</i>		-8.4	0.034
FBgn0003448	<i>sna</i>	<i>snail</i>	-8.3	0.174
FBgn0031260	<i>Spp</i>	<i>Signal peptide peptidase</i>	-8.3	0.052
FBgn0031261	<i>nAChRbeta3</i>	<i>nicotinic Acetylcholine Receptor beta3</i>	-8.2	0.004
FBgn0051901	<i>Mur29B</i>	<i>Mucin related 29B</i>	-8.1	0.185
FBgn0034902	<i>CG5532</i>		-8.1	0.001
FBgn0036563	<i>CG13075</i>		-8.0	0.177
FBgn0051821	<i>CG31821</i>		-7.9	0.034
FBgn0259697	<i>nvd</i>	<i>neverland</i>	-7.7	0.074
FBgn0014184	<i>Oda</i>	<i>Ornithine decarboxylase antizyme</i>	-7.3	0.081
FBgn0031263	<i>CG2789</i>	<i>Translocator protein</i>	-7.2	0.170
FBgn0004638	<i>drk</i>	<i>downstream of receptor kinase</i>	-7.2	0.064
FBgn0040565	<i>CG7606</i>		-7.0	0.130
FBgn0040658	<i>CG13516</i>		-6.9	0.090
FBgn0039494	<i>grass</i>	<i>Gram-positive Specific Serine protease</i>	-6.9	0.009
FBgn0041707	<i>7B2</i>		-6.8	0.057
FBgn0261989	<i>CG42807</i>		-6.8	0.223
FBgn0001090	<i>bnb</i>	<i>bangles and beads</i>	-6.7	0.017
FBgn0263911	<i>CoVIII</i>	<i>Cytochrome c oxidase subunit 8</i>	-6.6	0.096
FBgn0040718	<i>CG15353</i>		-6.6	0.062
FBgn0031910	<i>CG15818</i>		-6.6	0.176

Top 50 upregulated

Flybase ID	Gene symbol	Note	Fold change	p value
FBgn0085246	<i>CG34217</i>		34.1	0.006
FBgn0031141	<i>CG1304</i>		34.0	0.041
FBgn0040764	<i>CG13230</i>		25.9	0.057
FBgn0004428	<i>LysE</i>	<i>Lysozyme E</i>	16.3	0.016
FBgn0053267	<i>CG33267</i>		15.8	0.206

<i>FBgn0010295</i>	<i>ng3</i>	<i>new glue 3</i>	14.2	0.202
<i>FBgn0039719</i>	<i>CG15515</i>		13.8	0.071
<i>FBgn0053237</i>	<i>Ste,CG33237</i>		13.5	0.026
<i>FBgn0085771</i>	<i>CR40741</i>	<i>28S ribosomal RNA pseudogene:CR40741</i>	12.9	0.050
<i>FBgn0004427</i>	<i>LysD</i>	<i>Lysozyme D</i>	12.6	0.042
<i>FBgn0085320</i>	<i>CG34291</i>		12.4	0.044
<i>FBgn0044048</i>	<i>Iip5</i>	<i>Insulin-like peptide 5</i>	12.3	0.194
<i>FBgn0054003</i>	<i>NimB3</i>	<i>Nimrod B3</i>	12.1	0.035
<i>FBgn0031561</i>	<i>CG16712</i>	<i>Immune induced molecule 33</i>	11.9	0.097
<i>FBgn0039474</i>	<i>CG6283</i>		10.9	0.127
<i>FBgn0053241</i>	<i>Ste,CG33241</i>		10.8	0.030
<i>FBgn0004426</i>	<i>LysC</i>	<i>Lysozyme C</i>	10.5	0.028
<i>FBgn0053238</i>	<i>Ste,CG33238</i>		10.4	0.159
<i>FBgn0053242</i>	<i>Ste,CG33242</i>		10.0	0.021
<i>FBgn0028534</i>	<i>CG7916</i>		10.0	0.124
<i>FBgn0030105</i>	<i>CG15369</i>		9.5	0.127
<i>FBgn0262577</i>	<i>CG43117</i>		9.4	0.177
<i>FBgn0039471</i>	<i>CG6295</i>		9.3	0.141
<i>FBgn0040606</i>	<i>CG6503</i>		8.9	0.044
<i>FBgn0085244</i>	<i>CG34215</i>		8.5	0.040
<i>FBgn0262574</i>	<i>CG43114</i>		7.3	0.040
<i>FBgn0042118</i>	<i>Cpr65Ax2</i>	<i>Cuticular protein 65Ax2</i>	7.2	0.038
<i>FBgn0086900</i>	<i>Cpr65Ax1</i>	<i>Cuticular protein 65Ax1</i>	7.2	0.038
<i>FBgn0028396</i>	<i>TotA</i>	<i>Turandot A</i>	7.1	0.035
<i>FBgn0266598</i>	<i>Kmn2</i>	<i>kinetochore Mis12-Ndc80 network component 2</i>	7.0	0.020
<i>FBgn0262683</i>	<i>CG43153</i>		6.9	0.004
<i>FBgn0046687</i>	<i>Tre1</i>	<i>Trapped in endoderm 1</i>	6.9	0.021
<i>FBgn0039154</i>	<i>Npc2f</i>	<i>Niemann-Pick type C-2f</i>	6.8	0.100
<i>FBgn0031926</i>	<i>CG6739</i>		6.7	0.021
<i>FBgn0035926</i>	<i>CG5804</i>		6.6	0.009
<i>FBgn0262721</i>	<i>CG43165</i>		6.4	0.041

<i>FBgn0262036</i>	<i>CG42847</i>		6.3	0.026
<i>FBgn0010425</i>	<i>epsilonTry</i>	<i>epsilonTrypsin</i>	6.2	0.166
<i>FBgn0003187</i>	<i>qua</i>	<i>quail</i>	6.2	0.000
<i>FBgn0032824</i>	<i>CG13962</i>		6.2	0.022
<i>FBgn0034511</i>	<i>CG13422</i>	<i>GNBP-like 3</i>	6.1	0.000
<i>FBgn0053307</i>	<i>CG33307</i>		6.1	0.064
<i>FBgn0023495</i>	<i>Lip3</i>		6.1	0.040
<i>FBgn0030262</i>	<i>Vago</i>		6.0	0.003
<i>FBgn0010387</i>	<i>Dbi</i>	<i>Diazepam-binding inhibitor</i>	6.0	0.029
<i>FBgn0036146</i>	<i>CG14141</i>		5.9	0.122
<i>FBgn0039475</i>	<i>CG6277</i>		5.8	0.186
<i>FBgn0262035</i>	<i>CG42846</i>		5.8	0.090
<i>FBgn0262563</i>	<i>CG43103</i>		5.8	0.147
<i>FBgn0038795</i>	<i>CG4335</i>		5.7	0.049

