

**The Role of 3'-5' Exoribonucleases in  
Growth and Development of *Drosophila  
melanogaster***

Benjamin Paul Towler

A thesis submitted in partial fulfilment of the requirements of  
the University of Brighton and the University of Sussex for the  
degree of Doctor of Philosophy at the Brighton and Sussex  
Medical School.

June 2016

## Abstract

Post-transcriptional control is one of many layers of regulation of gene expression within the cell. Within this, RNA stability is known to play a critical role in determining the period of time an RNA species is able to elicit its function; such as an mRNA to be translated into protein or a miRNA to bind and repress its targets. The stability of a cytoplasmic RNA is dependent upon its degradation which can occur from either the 5' end by Pacman or from the 3' end by a family of related exoribonucleases; the Dis3 family.

The aim of this thesis was to investigate the role 3'-5' exoribonucleases play in post-transcriptional control of gene expression and how these relate to the development of a model tissue. *Drosophila melanogaster* have two highly conserved members of the Dis3 family, Dis3 and Dis3L2, both of which have been implicated in human disease. Dis3 has been recurrently mutated in multiple myeloma and acute myeloid leukaemia whilst Dis3L2 has been shown to be mutant in the human overgrowth syndrome, Perlman syndrome.

Due to their role in developmental diseases the function of both enzymes was investigated in the growth and development of the wing imaginal disc. A reverse genetics approach was used by driving RNA interference specifically in the wing imaginal disc which was directed by the highly specific *GAL4-UAS* system. The results within this thesis show that Dis3 is critical for organism viability together with the correct development of the wing with its loss resulting in large scale apoptosis. Using miRNA-sequencing has allowed the identification of a novel role for Dis3 in the regulation miRNA stability with *miR-252-5p* presenting as a specific target of Dis3.

Unlike, Dis3, Dis3L2 was shown to be dispensable for organism viability and the development of the adult wing identifying a more specific function. This was identified to be in the control of developmental proliferation. Loss of *dis3L2* within the wing imaginal disc resulted in increased proliferation of the wing imaginal discs cells resulting in the overgrowth of the wings and the wing imaginal discs by 20% compared to parental controls. The observed overgrowth was consistent with the human disease of Perlman syndrome which itself is characterised by foetal overgrowth thus identifying a conserved function for Dis3L2 in the control of tissue growth.

RNA sequencing was performed in a global, unbiased, approach in order to identify targets of Dis3L2 which become misregulated following its knockdown and may therefore drive the overgrowth phenotype. This revealed 2 potential targets of Dis3L2, *pyrexia* and *CG2678* which were shown to significantly increase in expression at the post-transcriptional level. Neither of these targets are known to be involved in proliferation, however, further investigation would be required to elucidate the mechanism by which Dis3L2 regulates proliferation within a developing tissue.

The role identified for Dis3L2 in regulating proliferation is in stark contrast to that of the 5'-3' exoribonuclease Pacman which has previously been shown to regulate apoptosis. Work in this thesis has built upon the relationship between Dis3L2 and Pacman and shown that they indeed function to regulate distinct pathways through the regulation of largely discrete sets of transcripts.

## Acknowledgements

I would first like to thank my supervisor, Sarah Newbury for giving me the opportunity to pursue this PhD. The support she has given both personally and scientifically has been outstanding. Her enthusiasm and genuine interest throughout the project has been a critical driving force.

I would also like to thank all the members of the lab both past and present who have provided guidance and entertainment along the way. Including the undergraduate students who were given the tedious jobs so that I didn't have to do them!

Specifically I would like to thank Chris Jones for his constant support and putting up with my endless questions. I think he was secretly happy I gave him something to moan about!

I would also like to thank Amy Pashler for both her interest and input into my work in addition to keeping me entertained with 'cheesy lab music'. I would also like to give her a massive thank you for putting up with me during the writing of this thesis!

Sophie Robinson also deserves a thank you for her mutual interest in my work and valuable input and discussions.

I would also like to thank Joe Waldron for his support during my PhD together with our numerous discussions about football! Between us we turned Chris into a football fan.

A very important thank you also goes to Clare Rizzo-Singh and Karen Scruby whose largely overlooked work allowed the lab to run smoothly. I am extremely grateful for their constant willingness to put in my countless orders and ensuring the fly food never ran out!

A large thank you also goes to the members of the Juan Pablo Couso lab for their valuable input during the *Drosophila* retreats.

Finally, and most importantly, I would like to thank all of my family and close friends. Specifically my parents, Tina and Paul, for their constant support and encouragement throughout this journey. Without them this would not have been possible.

## **Candidate's declaration**

I declare that the research contained in this thesis, unless otherwise formally indicated within the text, is the original work of the author. The thesis has not been previously submitted to this or any other university for a degree, and does not incorporate any material already submitted for a degree.

Signed:

Dated:

## Commonly used abbreviations

<u>Abbreviation</u>	<u>Definition</u>
Dis3	<u>D</u> efective in <u>s</u> ister chromatid rejoining <u>3</u>
Dis3L2	<u>D</u> efective in <u>s</u> ister chromatid rejoining <u>3</u> like <u>2</u>
XRN1/Pacman	<u>E</u> xo <u>r</u> ibo <u>n</u> uclease- <u>1</u>
Dcp1/2	<u>D</u> ecapping protein <u>1/2</u>
mRNA	<u>m</u> essenger <u>R</u> NA
pre-mRNA	preliminary <u>m</u> essenger <u>R</u> NA
miRNA	<u>m</u> icro <u>R</u> NA
pri/pre-miRNA	<u>p</u> ri <u>m</u> ary/ <u>p</u> re <u>m</u> i <u>n</u> ary <u>m</u> icro <u>R</u> NA
5' UTR	<u>5'</u> <u>U</u> n <u>t</u> ranslated <u>r</u> egion
3' UTR	<u>3'</u> <u>U</u> n <u>t</u> ranslated <u>r</u> egion
CDS	<u>C</u> oding <u>s</u> equence
ARE	<u>A</u> U <u>r</u> ich <u>e</u> lement
cDNA	<u>c</u> omplementary <u>D</u> NA
PCR	<u>P</u> olymerase <u>c</u> hain <u>r</u> eaction
qRT-PCR	<u>q</u> uantitative <u>r</u> everse <u>t</u> ranscriptase <u>P</u> CR
UAS	<u>U</u> pstream <u>a</u> ctivation <u>s</u> equence
RNAi	<u>R</u> NA <u>i</u> nterference
RNA-seq	<u>R</u> NA- <u>s</u> equencing
CRISPR	<u>C</u> lustered <u>r</u> egularly <u>i</u> nterspaced <u>s</u> hort <u>p</u> alindromc <u>r</u> epet
RPKM	<u>R</u> eads <u>p</u> er <u>k</u> ilobase of transcript per <u>m</u> illion mapped reads
FPKM	<u>F</u> ragments <u>p</u> er <u>k</u> ilobase of transcript per <u>m</u> illion mapped reads

## Table of Contents

Chapter 1: Introduction .....	18
1.1 Regulation of gene expression.....	18
1.1.1 Transcriptional control of gene expression .....	18
1.1.2 Post-transcriptional control of gene expression.....	19
1.1.3 Regulation by small RNAs .....	20
1.2 RNA species.....	21
1.2.1 messenger RNA.....	21
1.2.2 non-coding RNAs.....	21
1.2.2.1 Ribosomal RNA.....	22
1.2.2.2 Small nuclear/nucleolar RNAs.....	22
1.2.2.3 Long non-coding RNAs .....	22
1.2.2.4 tRNAs.....	23
1.2.3 microRNAs.....	23
1.3 RNA stability.....	27
1.3.1 The importance of RNA stability in gene expression .....	27
1.3.2 Pathways of RNA decay .....	27
1.3.3 An overview of 3'-5' decay.....	30
1.3.3.1 Deadenylation .....	30
1.3.3.2 The Dis3 family.....	31
1.3.3.3 Other core factors involved in 3'-5' decay.....	35
1.3.4 An overview of 5'-3' decay.....	36
1.3.4.1 Decapping .....	36
1.3.4.2 XRN1.....	37
1.3.5 Mechanisms of regulating mRNA stability.....	38
1.3.5.1 AU rich elements (AREs) .....	38
1.3.5.2 Regulation by 3' tailing.....	38
1.3.5.3 Factors promoting RNA stability .....	39
1.3.5.4 miRNA mediated regulation. ....	39
1.3.6 Mechanisms of regulating miRNA stability.....	41
1.3.6.1 Modifications to mature miRNAs.....	41
1.3.6.2 Degradation of miRNAs by exoribonucleases.....	42
1.3.6.3 Target mediated miRNA regulation. ....	42
1.3.6.4 <i>cis</i> -acting stability elements.....	43

1.4 Dis3.....	45
1.4.1 Dis3 structure and conservation.....	45
1.4.2 Cellular functions of Dis3 and phenotypes of Dis3 mutants.....	46
1.4.2.1 Cell cycle regulation.....	46
1.4.2.2 Dis3 in RNA processing and degradation.....	47
1.4.3 Expression pattern of Dis3 in <i>Drosophila melanogaster</i> . ....	51
1.5 Dis3L2.....	53
1.5.1 Dis3L2 structure and conservation. ....	53
1.5.2 Cellular functions of Dis3L2 and phenotypes of Dis3L2 mutants. ....	55
1.5.3 Expression pattern of Dis3L2 in <i>Drosophila melanogaster</i> .....	56
1.6 RNA stability in disease.....	58
1.6.1 Dis3.....	58
1.6.2 Dis3L2.....	59
1.6.3 XRN1.....	59
1.6.4 miRNAs in human disease.....	60
1.7 <i>Drosophila</i> as a model organism to study RNA stability.....	60
1.7.1 Why use <i>Drosophila</i> ? .....	60
1.7.2 <i>Drosophila</i> life cycle and development.....	62
1.7.3 Using the wing imaginal discs as a model system.....	63
1.8 An introduction to the techniques used during this study. ....	66
1.8.1 Using RNA interference to investigate protein function. ....	66
1.8.2 RNA-sequencing vs Microarrays. ....	66
1.8.3 Quantitative RT-PCR on mRNA and miRNA. ....	67
1.8.4 CRISPR. ....	71
1.9 Aims.....	74
Chapter 3.....	74
Chapter 4.....	74
Chapter 5.....	75
Chapter 6.....	75
Chapter 7.....	75
Chapter 2: Methods.....	76
2.1 <i>Drosophila</i> stocks used during the project. ....	76
2.2 <i>Drosophila</i> husbandry. ....	78
2.2.1 Fly food and crosses.....	78
2.2.2 Balancer chromosomes.....	78

2.2.3	Developmental staging. ....	79
2.2.4	Measurement of developmental delay. ....	79
2.2.5	Wing imaginal disc dissection for RNA extraction. ....	79
2.2.6	Wing imaginal disc dissection for measurement. ....	79
2.2.7	Wing dissection for measurement. ....	80
2.2.8	Counting the number of cells in the wing. ....	82
2.2.9	Photographing fly phenotypes. ....	82
2.2.10	Starvation experiment. ....	82
2.3	Genetic techniques. ....	82
2.3.1	The <i>GAL4-UAS</i> system. ....	82
2.3.2	<i>GAL80<sup>ts</sup></i> addition and experimental design. ....	83
2.3.3	P-element mutagenesis. ....	83
2.3.4	Recombination crosses. ....	84
2.3.5	CRISPR. ....	87
2.3.5.1	gRNA design. ....	87
2.3.5.2	Annealing and phosphorylation of gRNA oligos. ....	87
2.3.5.3	Digestion and dephosphorylation of pCFD3-dU6:3gRNA ....	87
2.3.5.4	Ligation of gRNA into pCFD3. ....	88
2.3.5.5	Transformation of pCFD3 into DH5 $\alpha$ cells. ....	88
2.3.5.6	Colony picking and inoculation. ....	88
2.3.5.7	Minipreps and sequencing. ....	89
2.4	Molecular techniques. ....	89
2.4.1	Polymerase Chain Reaction. ....	89
2.4.1.1	Squish prep. ....	89
2.4.1.2	PCR conditions. ....	89
2.4.1.3	Primers used and their design. ....	90
2.4.1.4	Agarose gel electrophoresis. ....	90
2.4.1.5	PCR purification and sequencing. ....	90
2.4.2	quantitative RT-PCR. ....	92
2.4.2.1	RNA extraction on whole adults/L3 larvae. ....	92
2.4.2.2	RNA extraction on dissected wing imaginal discs. ....	92
2.4.2.3	cDNA preparation from total RNA. ....	92
2.4.2.4	cDNA preparation from miRNAs. ....	93
2.4.2.5	qRT-PCR for total RNA and miRNA. ....	93
2.4.2.6	qRT-PCR quantification. ....	94

2.4.3 Western blotting .....	94
2.4.4 Immunocytochemistry .....	95
2.4.5 Mitotic index calculation.....	95
2.4.6 Transformation of DH5 $\alpha$ with <i>dis3L2</i> cDNA.....	96
2.4.6.1 Bacterial transformation.....	96
2.4.6.2 Transformation diagnostics. ....	96
2.5 Next Generation Sequencing.....	97
2.5.1 Small RNA sequencing. ....	97
2.5.2 Analysis of small RNA sequencing.....	97
2.5.2.1 Identification of known miRNAs. ....	97
2.5.2.2 Identification of known and novel miRNAs using miRDeep2. ....	97
2.5.3 RNA sequencing.....	98
2.5.4 Analysis of RNA sequencing.....	98
2.6 Statistical analysis.....	99
Chapter 3: The role of Dis3 in the development of and regulation of miRNA stability in the wing imaginal discs. ....	100
3.1 Introduction.....	100
3.2 Project background and aims.....	103
3.3 Dis3 is required for <i>Drosophila melanogaster</i> viability.....	104
3.4 Loss of Dis3 in the wing imaginal disc results in severe phenotypic defects.....	105
3.4.1 Dis3 is required throughout the wing disc for viability.....	105
3.4.2 Dis3 knockdown within the wing pouch results in an absence of wing development.....	107
3.4.3 Dis3 is required for the correct development of the thorax as well as the wing. ....	108
3.5 Dis3 knockdown results in smaller wing imaginal discs caused by increased apoptosis....	112
3.6 Does Dis3 play a role in the regulation of miRNA stability? .....	115
3.6.1 Sample collection and miRNA-sequencing.....	115
3.6.2 Global analysis of miRNA seq data. ....	119
3.6.3 Analysis of Dis3-sensitive miRNAs.....	124
3.6.4 <i>miR-252-5p</i> and <i>miR-982-5p</i> are regulated at the post-transcriptional level. ....	130
3.6.5 Loss of the cytoplasmic function of Dis3 is responsible for the increase in <i>miR-252</i> expression.....	132
3.7 <i>miR-252</i> is required for correct wing development.....	136
3.8 Analysis of predicted <i>miR-252-5p</i> targets.....	138
3.9 3' editing of miRNAs.....	140
3.10 Identification of a novel miRNA in the <i>Drosophila</i> wing imaginal disc.....	143
3.11 Chapter Summary .....	145

3.12 Chapter Discussion.....	148
3.12.1 Differences between stages of lethality with published data when Dis3 is knocked down ubiquitously. ....	148
3.12.2 Why is the loss of Dis3 so dramatic? .....	149
3.12.3 A novel role for Dis3 in the regulation of mature miRNAs. ....	149
3.12.4 How might Dis3 be targeted to specific miRNAs? .....	152
3.12.5 Issues with 2S contamination in the miRNA seq limits the number of miRNA reads.	153
3.13 Acknowledgements.....	154
Chapter 4: A novel role for the 3'-5' exoribonuclease Dis3L2 in cell proliferation.....	155
4.1 Introduction .....	155
4.2 Aims.....	158
4.3 Validation of the expression of the 2 predicted isoforms of Dis3L2 in <i>Drosophila</i> . ....	158
4.4 Dis3L2 is not required for <i>Drosophila</i> viability.....	162
4.5 Loss of Dis3L2 results in an increased number of cells in the wing.....	166
4.5.1 Successful knockdown of <i>dis3L2</i> in the wing imaginal disc. ....	166
4.5.2 Knockdown of <i>dis3L2</i> results in overgrowth of the wing.....	167
4.5.3 Knockdown of <i>dis3L2</i> results in hyperplasia of the wing.....	170
4.5.4 Wing overgrowth is specific to the loss of <i>dis3L2</i> . ....	175
4.6 Overgrowth occurs early in development. ....	178
4.7 Loss of <i>dis3L2</i> results in increased proliferation in the wing imaginal discs.....	180
4.7.1 Loss of <i>dis3L2</i> results in significantly larger wing imaginal discs. ....	180
4.8 Loss of <i>dis3L2</i> results in enhanced starvation survival .....	186
4.9 Chapter Summary .....	187
4.10 Chapter Discussion.....	189
4.10.1 Why does the loss of <i>dis3L2</i> result in proliferation? .....	189
4.10.2 Why is the overgrowth only observed when <i>dis3L2</i> is lost during early developmental stages? .....	192
4.10.3 Do Dis3L2 and Pacman function in distinct pathways? .....	193
4.11 Acknowledgements.....	194
Chapter 5: Using RNA-sequencing to identify potential targets of Dis3L2.....	195
5.1 Introduction .....	195
5.2 Aims.....	197
5.3 Experimental plan and sample preparation .....	198
5.4 Library preparation and RNA-sequencing performed by Oxford Gene Technology .....	201
5.5 Read mapping and differential expression analysis.....	204
5.5.1 Assessing read quality using FastQC. ....	204

5.5.2	Adaptor removal and quality trimming using Sickle and Scythe .....	204
5.5.3	Mapping quality controlled reads to the <i>Drosophila</i> genome using TopHat2 .....	206
5.5.4	Using Cufflinks to assemble the reads successfully mapped by TopHat2 .....	210
5.5.5	Using the Cuffdiff pipeline to identify differential expression .....	211
5.6	Global overview of the RNA-sequencing data. ....	215
5.6.1	Global expression distributions identify differential expression between control and <i>dis3L2</i> knockdown samples.....	215
5.7	Selection of confidently misexpressed transcripts. ....	219
5.8	Validation and analysis of selected misexpressed transcripts.....	222
5.8.1	Gene ontology analysis of consistently misexpressed transcripts. ....	222
5.8.2	Validation of misexpressed genes by qRT-PCR. ....	227
5.8.3	<i>CG2678</i> and <i>pyrexia</i> show the characteristics of direct Dis3L2 targets.....	229
5.9	Chapter Summary .....	235
5.10	Chapter Discussion.....	236
5.10.1	A lack of confidence in statistical analysis by Cuffdiff. ....	236
5.10.2	Why do so few mRNAs change in expression? .....	237
5.10.3	<i>pyx</i> and <i>CG2678</i> appear to be direct targets of Dis3L2. ....	239
5.11	Acknowledgements.....	241
Chapter 6: Pacman and Dis3L2 function in distinct pathways to regulate wing disc development .....		242
6.1	Introduction. ....	242
6.2	Aims.....	244
6.3	Knockdown of <i>dis3L2</i> and <i>pacman</i> results in an additive effect on <i>Drosophila</i> viability....	244
6.3.1	Simultaneous knockdown of <i>dis3L2</i> and <i>pcm</i> throughout the organism results in lethality. ....	245
6.3.2	Simultaneous knockdown of <i>dis3L2</i> and <i>pcm</i> within the wing imaginal discs results in additive effects.....	245
6.4	<i>dis3L2</i> knockdown rescues small wing/wing disc phenotypes of <i>pcm</i> mutants.....	249
6.4.1	<i>dis3L2</i> knockdown in a hypomorphic <i>pcm</i> mutant background results in phenotypic rescue.....	249
6.4.2	<i>dis3L2</i> knockdown in a null <i>pcm</i> mutant background results in partial phenotypic rescue.....	252
6.4.3	<i>dis3L2</i> knockdown does not rescue the <i>pcm</i> <sup>14</sup> induced developmental delay.....	256
6.5	Dis3L2 does not function through the Hid, Grim and Reaper mediated apoptotic pathway to regulate disc growth.....	258
6.6	Dis3L2 and Pacman target specific RNAs to regulate their opposing pathways. ....	261
6.6.1	Dis3L2 knockdown in a <i>pcm</i> <sup>14</sup> mutant background does not affect <i>rpr</i> expression. ....	261

6.6.2 Comparative analysis of RNA sequencing data reveals transcripts showing specific sensitivity to Pacman and Dis3L2.....	264
6.7 Chapter Summary .....	270
6.8 Chapter Discussion.....	271
6.8.1 Why are there partial differences in the phenotypes observed between double knockdown and knockdown in mutant backgrounds? .....	271
6.8.2 Are Pacman and Dis3L2 involved in regulating opposing pathways?.....	273
6.8.3 How might the pathways specificity be achieved?.....	274
6.9 Acknowledgements.....	275
Chapter 7: P-element mutagenesis screens and Dis3L2 antibody production.....	276
7.1 Introduction .....	276
7.2 Aims.....	277
7.3 Attempted production of a <i>dis3L2</i> mutant .....	277
7.3.1 An annotated P-element stock with potential for use in <i>dis3L2</i> mutagenesis does not contain the specified P-element .....	277
7.3.2 Sourcing and testing of a second P-element containing stock .....	279
7.3.3 P-element mutagenesis using <i>GS6090</i> in an attempt to produce a <i>dis3L2</i> mutant. ....	281
7.4 Removal of the lethality causing factor by using recombination. ....	282
7.5 P-element mutagenesis using the rescued P-element stock.....	287
7.5.1 Initial mutagenesis experiment using the rescued P-element stock.....	287
7.5.2 Repeat mutagenesis using the rescued P-element stock. ....	291
7.6 Use of clustered regularly interspaced short palindromic repeats (CRISPR) technology to produce a <i>dis3L2</i> mutant. ....	296
7.6.1 CRISPR as a mutagenesis tool. ....	296
7.6.2 Identification of target and design of the gRNA. ....	297
7.6.3 Cloning of the gRNA into the pCFD3-dU63gRNA vector.....	297
7.7 Design and production of a Dis3L2 antibody.....	301
7.8 Production of a polyclonal anti-Dis3L2 antibody .....	304
7.9 Chapter Summary .....	308
7.10 Chapter Discussion.....	309
7.10.1 Why were the mutagenesis experiments unsuccessful?.....	309
7.10.2 Using CRISPR should be a more efficient way of creating a <i>dis3L2</i> mutant. ....	310
7.10.3 Limitations of the peptide antibody design.....	311
7.11 Acknowledgements.....	311
Chapter 8: Discussion.....	312
8.1 Summary of main findings. ....	312

8.1.1 Dis3 is essential for cell viability. ....	312
8.1.2 Dis3 plays a role in miRNA regulation.....	312
8.1.3 A novel role for Dis3L2 in regulating proliferation within the wing disc. ....	312
8.1.4 Dis3L2 and Pacman affect opposing but related processes. ....	313
8.2 Does Dis3 play a role in the regulation of miRNA stability? .....	313
8.2.1 The function of Dis3 on miRNAs is specific.....	313
8.2.2 A role for Dis3 in regulating mature miRNA stability.....	314
8.2.3 A potential role for Dis3 in regulating miRNA biogenesis.....	316
8.3 A novel role for Dis3L2 in controlling developmental proliferation? .....	317
8.3.1 Why does Dis3L2 knockdown result in hyperplasia of the wing?.....	317
8.3.2 Is the proliferative function of Dis3L2 tissue specific? .....	319
8.4 Are <i>pyrexia</i> and <i>CG2678</i> direct targets of Dis3L2? .....	321
8.4.1 <i>pyrexia</i> and <i>CG2678</i> are unlikely to be the only Dis3L2 targets. ....	321
8.4.2 Could Pyrexia and CG2678 contribute to the overgrowth phenotype? .....	322
8.4.3 Methods to identify additional Dis3L2 targets. ....	324
8.5 A role for Dis3L2 in starvation response? .....	326
8.6 Co-ordination between apoptosis and proliferation by the 5'-3' and 3'-5' pathways. ....	327
8.6.1 Are Pacman and Dis3L2 involved in regulating opposing pathways?.....	327
8.6.2 How might the pathway specificity be achieved? .....	330
8.7 Implications in human disease.....	332
8.8 Concluding remarks and Future work.....	334
References .....	335
Appendix: Comparative analysis of Pcm and Dis3L2 sequencing. ....	352

## Table of Figures

Figure 1.1: Classical and alternative pathways of miRNA biogenesis.....	26
Figure 1.2: mRNA degradation pathways.....	29
Figure 1.3: The Dis3 family of 3'-5' exoribonucleases.....	33
Figure 1.4: Model of association between Dis3 and the exosome core.....	34
Figure 1.5: Mechanisms of regulating RNA stability.....	40
Figure 1.6: Factors affecting miRNA degradation.....	44
Figure 1.7: Summary of phenotypes observed in Dis3 loss of function studies.....	50
Figure 1.8: The expression profile of <i>dis3</i> mRNA in <i>Drosophila melanogaster</i> .....	52
Figure 1.9: Structure of Dis3L2 and an RNA substrate.....	54
Figure 1.10: The expression profile of <i>dis3L2</i> mRNA in <i>Drosophila melanogaster</i> .....	57
Figure 1.11: The <i>Drosophila</i> wing imaginal disc.....	65
Figure 1.12: How TaqMan probes work for qRT-PCR on mRNAs.....	69
Figure 1.13: How TaqMan probes work for qRT-PCR on miRNAs.....	70
Figure 1.14: Using CRISPR to generate mutants.....	73
Figure 2.1: Guidelines used for the measurement of <i>Drosophila</i> wings.....	81
Figure 2.2: The <i>GAL4-UAS</i> system and the <i>GAL80<sup>ts</sup></i> addition.....	85
Figure 2.3: Wing imaginal disc GAL4 drivers used during this project.....	86
Figure 3.1: Cross schemes to achieve ubiquitous knockdown of Dis3.....	106
Figure 3.2: Knockdown of Dis3 in the wing imaginal disc results in severe developmental phenotypes.....	109
Figure 3.3: The no wing phenotype observed following Dis3 knockdown in the wing pouch was not influenced by <i>UAS-dcr2</i> .....	110
Figure 3.4: Dis3 knockdown in the thoracic region of the wing imaginal disc results in a severe cleft thorax phenotype.....	111
Figure 3.5: Dis3 knockdown in the wing pouch of the wing imaginal discs results in small wing discs and extensive apoptosis.....	114
Figure 3.6: The RNA sent for miRNA-sequencing was of high quality.....	117
Figure 3.7: Summary of miRNA-sequencing read details.....	118
Figure 3.8: miRNA sequencing identifies 109 mature miRNAs in L3 wing imaginal discs.....	121
Figure 3.9: Expression patterns of miRNAs in wing imaginal discs.....	122
Figure 3.10: Hierarchical clustering of replicates in miRNA-seq.....	123

Figure 3.11: Effect of Dis3 knockdown on all wing imaginal disc miRNAs detected above threshold levels.....	126
Figure 3.12: Graphical representation of the miRNAs selected for validation.....	127
Figure 3.13: qRT-PCR validation of misexpressed miRNAs from the miRNA-seq data.....	128
Figure 3.14: Summary of the miRNA-seq fold changes and the qRT-PCR fold changes for each of the selected miRNAs.....	129
Figure 3.15: Table summarising the fold changes as determined by miRNA-seq of the partner miRNAs to those miRNAs validated by qRT-PCR.....	134
Figure 3.16: Knockdown of the exosome subunit Rrp40 results in similar phenotypes as to when Dis3 is knocked down.....	134
Figure 3.17: Unsuccessful knockdown of Ski2.....	135
Figure 3.18: Overexpression of <i>miR-252</i> in the wing pouch of the imaginal disc results in severe wing phenotypes.....	137
Figure 3.19: Knockdown of Dis3 or overexpression of <i>miR-252</i> does not affect the expression of predicted <i>miR-252-5p</i> targets.....	141
Figure 3.20: Terminal nucleotide additions to the Dis3 regulated miRNAs.....	142
Figure 3.21: The methods used by miRDeep2 to determine potential novel miRNA precursors.....	146
Figure 3.22: The detection of a novel miRNA in the <i>Drosophila</i> wing imaginal discs by miRNA-seq.....	147
Figure 4.1: Blast search using the human Dis3L2 peptide sequence to confirm the <i>Drosophila</i> orthologue.....	160
Figure 4.2: Dis3L2 isoform information.....	160
Figure 4.3: Both <i>dis3L2</i> isoforms are expressed in <i>Drosophila</i> .....	161
Figure 4.4: Cross schemes to achieve ubiquitous knockdown of <i>dis3L2</i> performed at 25°C...	164
Figure 4.5: Ubiquitous knockdown of <i>dis3L2</i> does not affect organism viability.....	165
Figure 4.6: Successful knockdown of <i>dis3L2</i> in the wing imaginal disc.....	171
Figure 4.7: Knockdown of <i>dis3L2</i> in the wing imaginal disc results in specific overgrowth.....	172
Figure 4.8: Ubiquitous knockdown also results in wing overgrowth but does not cause total overgrowth of the fly.....	173
Figure 4.9: Wing overgrowth is due to an increase in cell number rather than an increased cell size.....	174
Figure 4.10: <i>dis3L2</i> knockdown induced overgrowth is specific.....	177
Figure 4.11: Experimental plan to knockdown <i>dis3L2</i> during different developmental stages.....	183

Figure 4.12: Wing overgrowth is only observed when <i>dis3L2</i> is knocked down during early larval stages. ....	184
Figure 4.13: Knockdown of <i>dis3L2</i> results in increased proliferation within the wing imaginal disc leading to overgrowth of the disc.....	185
Figure 4.14: Ubiquitous knockdown of <i>dis3L2</i> leads to an improved starvation response.....	188
Figure 5.1: RNA samples sent for RNA-sequencing were of high integrity. ....	200
Figure 5.2: RNA-seq library preparation work flow using the Illumina TruSeqRNA sample preparation kit. ....	203
Figure 5.3: Analysis of RNA-seq 100 cycle output using FastQC.....	208
Figure 5.4: Summary of read alignment following the application of TopHat2. ....	209
Figure 5.5: Overview of the programs and algorithms used in the RNA-seq analysis pipeline.	213
Figure 5.6: The RNA-seq analysis pipeline. ....	214
Figure 5.7: RNA-seq differential expression overview.....	218
Figure 5.8: Graphical representation of the consistently misexpressed transcripts in <i>dis3L2<sup>KD</sup></i> wing imaginal discs. ....	221
Figure 5.9: Strip plot showing replicate consistencies for the top 30 upregulated transcripts.	224
Figure 5.10: Strip plot showing replicate consistencies for the top 30 downregulated transcripts. ....	225
Figure 5.11: Gene ontology analysis of the 229 misexpressed transcripts in <i>dis3L2<sup>KD</sup></i> wing imaginal discs. ....	226
Figure 5.12: RNA-seq validation using qRT-PCR.....	231
Figure 5.13: Summary of RNA-seq validation by qRT-PCR.....	232
Figure 5.14: Sequences of custom TaqMan assays.....	233
Figure 5.15: Transcription factors predicted to bind <i>cyt-c-d</i> and <i>CG31808</i> do not show expression changes in <i>dis3L2<sup>KD</sup></i> wing imaginal discs. ....	234
Figure 6.1: Simultaneous knockdown of <i>pacman</i> and <i>dis3L2</i> using the ubiquitous driver <i>tub-GAL4</i> .....	247
Figure 6.2: Simultaneous knockdown of <i>pacman</i> and <i>dis3L2</i> in the wing imaginal disc results in additive effects.....	248
Figure 6.3: Knockdown of <i>dis3L2</i> in a <i>pcm</i> mutant background results in phenotypic rescue.	251
Figure 6.4: Cross scheme to achieve the knockdown of <i>dis3L2</i> in a <i>pacman</i> null mutant background. ....	255
Figure 6.5: Knockdown of <i>dis3L2</i> does not affect the <i>pcm<sup>14</sup></i> induced upregulation of <i>dilp8</i> ....	257
Figure 6.6: <i>dis3L2</i> knockdown induced overgrowth is unlikely to be caused by the inhibition of apoptosis.....	260

Figure 6.7: Knockdown of <i>dis3L2</i> in <i>pcm<sup>14</sup></i> mutant wing discs has no effect on the post-transcriptional upregulation of <i>rpr</i> .	263
Figure 6.8: 3 classes of transcripts are identified following comparative analysis of RNA-seq data.	267
Figure 6.9: Summary of the number of transcripts in each of the 3 classes presented in Figure 6.8.	268
Figure 6.10: Gene ontology (GO) analysis of the transcripts identified in each of the 3 classes presented in Figure 6.8.	268
Figure 6.11: Comparison of the fold changes between Dis3L2 and Pacman sequencing data.	269
Figure 7.1: Checking the presence of <i>P{KG02284}</i> before proceeding with P-element mutagenesis.	278
Figure 7.2: The <i>P{GS6090}</i> insertion is detected in the new P-element stock.	280
Figure 7.3: P-element mutagenesis cross scheme.	284
Figure 7.4: Examples of mosaic eye phenotypes observed in the presence of the transposase indicating P-element mobilisation.	285
Figure 7.5: Cross scheme used to rescue the lethality inducing factor on the <i>P{GS6090}</i> containing chromosome.	286
Figure 7.6: PCR to check the presence of the P-element in the repaired stocks.	289
Figure 7.7: P-element mutagenesis using <i>P{GS6090}</i> .	290
Figure 7.8: PCR screening of the potential mutant stocks produced during the repeat mutagenesis experiment.	294
Figure 7.9: qRT-PCR analysis of the levels of <i>dis3L2</i> mRNA in all potential mutant stocks.	295
Figure 7.10: Integration of the gRNA sequence into the pCFD3-dU6:3gRNA vector.	299
Figure 7.11: Preparation of the injection vector for CRISPR.	300
Figure 7.12: Design of an anti-Dis3L2 peptide antibody.	302
Figure 7.13: Testing the anti-Dis3L2 peptide antibody.	306
Figure 7.14: Transformation of <i>dis3L2</i> cDNA in preparation for antibody production.	307
Figure 8.1: How might Dis3L2 control proliferation in <i>Drosophila</i> wing imaginal discs?	320
Figure 8.2: <i>pyrexia</i> is the only member of the TRP channel family to be upregulated following <i>dis3L2</i> knockdown.	325
Figure 8.3: Hypothesis behind Dis3L2 and Pacman regulating proliferation and apoptosis respectively.	329

# Chapter 1: Introduction

## 1.1 Regulation of gene expression

Within almost every cell lies the genetic code which is fundamental to life, encoded in DNA. DNA contains all the information responsible for directing cellular processes in simple unicellular organisms as well as more complex processes in multicellular organisms such as humans. These processes are dependent on the level to which each DNA component, a gene, is read and made into either a protein or RNA structure. Each cell contains exactly the same DNA and therefore it is the control over the expression of specific genes that gives each cell a specific fate; such as a neural or muscle cell. It is therefore critical to maintain tight regulation throughout the journey from a DNA sequence of individual nucleotide bases to the final, functional product. A single error encountered during this journey can have drastic consequences and therefore there are a number of regulatory mechanisms devoted to ensuring a safe transition from code to product.

### 1.1.1 Transcriptional control of gene expression

For the DNA code to be made into its functional product it must first be transcribed. This involves the reading of the code by a family of enzymes named RNA polymerases. These copy the message encoded in the DNA into RNA. There are a number of RNA species within the cell which are made by specific members of the RNA polymerase family. These RNAs carry a variety of functions which will be discussed in a later section. For those genes encoding proteins the DNA is transcribed into a pre-messenger RNA (pre-mRNA) which undergoes further nuclear processing into the mature mRNA which is ultimately exported into the cytoplasm to be made into protein. The amount of mRNA that is made from the DNA therefore must be tightly controlled to ensure that excess protein is not subsequently made. There are a number of well characterised factors which are involved in regulating the amount of mRNA that is transcribed from each genetic locus. These regulatory events can occur at a number of places during the process of transcription. A major factor involved in transcriptional regulation is the access of RNA polymerase to the specific locus. The DNA exists in either lightly packed, usually transcriptionally active, euchromatin or in tightly packed structures known as heterochromatin which represses transcription. However, modifications to the proteins involved in regulating the packing can cause the DNA to unwind allowing access of RNA polymerase and the subsequent transcription of the DNA. These proteins are called histones and their activity is modulated through acetylation. Acetylation of the histones causes their unwinding and

thereby promotes transcription whereas the removal of the acetyl groups represses transcription through stimulating the dense packing of the DNA (Eberharter and Becker 2002).

In addition to controlling the access of RNA polymerase through the DNA structure both *cis* and *trans*-acting elements regulate the level of transcription of specific genes. An example of a *cis* acting factor is a promoter which as its names suggests initiates transcription of specific genes. There are also enhancer elements which bind *trans*-acting elements such as transcription factors which in turn promote transcription. There are vast array of transcription factors known which are able to promote or inhibit transcription; these include key oncogenes such as Myc and Fos (Lee and Young 2013).

### **1.1.2 Post-transcriptional control of gene expression**

The process of transcription results in a pre-mRNA construct which contains both exons and introns. Therefore, the introns require removal by splicing which occurs co-transcriptionally (Neugebauer 2002; Görnemann *et al.* 2005). Splicing is a critical event which is also highly regulated by factors such as the small nuclear RNAs (snRNAs) U1, 2, 4, 5 and 6 (Lührmann *et al.* 1990; Görnemann *et al.* 2005). A single pre-mRNA may give rise to a number of different mature mRNAs (isoforms) this is achieved through alternative splicing. This involves the use of different exons within the pre-mRNA and is a process used by at least 74% of human multi-exon genes (Johnson *et al.* 2003). Alternative splicing is a phenomenon which ultimately generates proteomic diversity allowing the expression of a number of different proteins, with potentially differing functions, from a single DNA code. The correct balance of alternative splicing is critical as two protein isoforms of the Fas receptor can have opposing effects on apoptosis (Cascino *et al.* 1995) and therefore an error in this procedure can seal the cells fate. The importance of alternative splicing is shown through the fact that errors in alternative splicing have been implicated in a number of diseases such as Myotonic Dystrophy and Amyotrophic Lateral Sclerosis (Tazi *et al.* 2009; Cieply and Carstens 2015). Terminal additions are then made to the successfully spliced mRNAs which are required for their protection when they reach the cytoplasm. A methylguanosine (m7G) cap is added to the 5' end and a stretch of adenines are added to the 3' end (poly(A) tail) which both function to protect the mRNA from exoribonucleases. The spliced, terminally protected mRNAs are then exported to the cytoplasm where they are subject to further regulatory events. These pre-mRNA processing events are nicely reviewed in Moore and Proudfoot 2009 (Moore and Proudfoot 2009).

Once an mRNA reaches the cytoplasm is it subject to further regulatory events by translational or degradation machinery. In its normal conformation the capped and polyadenylated mRNA is circularised which allows for efficient production into protein, termed translation. Circularising

of the mRNA is achieved through protein-protein interaction between the poly(A) binding protein (PABP) and the cap binding protein eIF4E. This not only allows for efficient translation but it also acts to protect the mRNA from degradation. Translation is initiated through the recruitment of initiation factors to the 5' end of the transcript. These include eIF4G which acts as a scaffold for eIF4A and eIF4E forming the eIF4F initiation complex (Schütz *et al.* 2008). The helicase activity of eIF4A allows the unwinding of the secondary structure (Rogers Jr *et al.* 2002) and subsequent recruitment of the 40S ribosomal subunit together with other initiation factors. This pre-initiation complex scans the transcript until it reaches the AUG start codon which triggers the binding of the 60S ribosomal subunit and dissociation of the initiation factors. This results in the formation of the functional 80S ribosome which translates the RNA message into protein.

A major level of translational regulation occurs through the control of the eIF4F initiation complex. This occurs mainly through phosphorylation events such as phosphorylation of eIF4E binding proteins. Phosphorylation of these specific binding proteins allow their binding to eIF4E which occurs in competition with its binding to the eIF4G scaffold (Gingras *et al.* 1999; Clemens 2001). A reduction of eIF4E-4G interaction results in global downregulation of translation. Similarly eIF4E itself is regulated by phosphorylation an event which has been shown to promote resistance to stresses (Wang *et al.* 1998; Martinez *et al.* 2015).

### **1.1.3 Regulation by small RNAs**

The level of translation of an mRNA can also be regulated by machinery involved in controlling RNA stability. One example of this is microRNA-mediated translational regulation; microRNAs (miRNAs) are small sequences of RNA 18-22 nucleotides in length which bind to complementary sites in the untranslated regions (UTRs). miRNA binding can result in translational repression and/or degradation of the bound transcript. One theory of how translational inhibition is achieved is through direct interaction between the 5' cap and Argonaute 2 (Ago2), a member of the RNA induced silencing complex (RISC) (Kiriakidou *et al.* 2007). This is thought to be in direct competition with eIF4E, therefore, in the absence of eIF4E binding the eIF4F complex cannot be formed preventing the initiation of cap-dependent translation. However, it is clear that there are still multiple hypotheses on how translational repression is achieved (Gu and Kay 2010; Wilczynska and Bushell 2015).

Binding of a miRNA to the 3'UTR can also promote degradation of the transcript; a process better understood than how they repress translation. This occurs through the interaction of GW182 (part of the RISC) with the PABP. GW182 has also been shown to recruit deadenylase complexes (Ccr4-Not and Pan2-Pan3) which remove the poly(A) tail allowing

exoribonucleolytic attack from the 3' end (Braun *et al.* 2011). Similarly, the RISC is also able to recruit decapping machinery which remove the 5' cap allowing mRNA decay to proceed from the 5' end (Nishihara *et al.* 2013). In addition to miRNAs, RNA binding proteins can also be recruited to the mRNA through association with specific *cis* sequences within the UTR's producing another layer of regulation. These include HuR or TTP which bind to AU rich elements (AREs) in the 3' UTR to regulate transcript stability (Carballo and Blackshear 2001; Akamatsu *et al.* 2005) and Bruno which also binds within the 3' UTR and competes for cap binding with eIF4E (Richter and Sonenberg 2005) thereby repressing translation.

## **1.2 RNA species**

There are a number of RNA species many of which have only come to light in the past 2 decades. These include RNAs which simply act as messages to produce protein (mRNAs) or those that elicit their function as the RNA such as miRNAs and snoRNAs. The different species all have distinct functions within the cell and are responsible for the regulation of a number of critical cellular processes. The following sections will introduce the various RNA species and their functions within the cell.

### **1.2.1 messenger RNA**

A messenger RNA (mRNA) is an intermediate between the gene encoded within the DNA and the functional protein it contains the information for. mRNAs are produced following transcription by RNA polymerase II, they are spliced, capped and polyadenylated within the nucleus as outlined in previous sections. Once successfully processed and protected they are exported to the cytoplasm where they are translated into protein and subsequently degraded through a variety of processes which will be discussed in latter sections. An mRNA consists of 3 sections, the 5' UTR, coding sequence and 3' UTR. The 5' and 3' UTRs are structured, regulatory elements which contain binding sites for a number of *trans*-acting factors such as miRNAs or RNA-binding proteins which function to regulate both translation and RNA degradation. The coding sequence is read in triplicate, beginning with the start codon (usually AUG) and contains the genetic code to produce the functional protein. These are of the most well studied members of the RNA family.

### **1.2.2 non-coding RNAs**

In addition to the protein coding mRNAs there are a variety of non-protein coding RNAs many of which have only come to light in recent years. It has been shown that there are four times the number of non-coding RNAs than of mRNAs (Kapranov *et al.* 2007). The functions of these non-coding RNAs are vast. The best characterised of the non-coding RNAs are the ribosomal

rRNAs (rRNAs) which are of the most stable RNAs encoded within the genome. rRNAs associate with and aid the construction of the functional 80S ribosome. 4 rRNAs associate with the two ribosomal subunits; a single 18S, 5.8S and 5S rRNA associate with the 40S subunit whilst the 28S rRNA binds the 60S subunit and are required for translation to occur.

#### **1.2.2.1 Ribosomal RNA**

The 28S, 18S and 5.8S rRNAs are co-transcribed by RNA polymerase I initially into a single, larger rRNA precursor (47S pre-rRNA). This precursor then undergoes a number of regulated stages of cleavage and trimming resulting in the biogenesis of the three final rRNA structures of the correct size (Granneman and Baserga 2004). These are bound by other ribosomal proteins and exported to the cytoplasm where they are able to elicit their function through interacting with the large and small ribosomal subunits. Unlike the 28S, 18S and 5.8s rRNA, the smaller 5S rRNA is transcribed separately by RNA polymerase III. It is also transcribed as a precursor which also undergoes similar processing before the final 5S rRNA is exported to the cytoplasm. Due to their essential role in translation rRNAs are found in every species known and for this reason they are often used to build phylogenetic trees.

#### **1.2.2.2 Small nuclear/nucleolar RNAs**

The biogenesis of these rRNAs requires another class of non-coding RNAs; small nucleolar RNAs (snoRNAs). As their name suggests these are small RNA sequences found in the nucleolus and mainly function to guide modifications, such as 2'-O-methylation, to other RNA structures (Maden and Hughes 1997; Dunbar and Baserga 1998). Along with their role in pre-rRNA processing they are also involved in the regulation of splicing (Lührmann *et al.* 1990; Görnemann *et al.* 2005) and chromatin structure (Schubert *et al.* 2012). There are two main classes of snoRNAs; C/D box and H/ACA box snoRNAs. The classification of snoRNAs is dependent upon specific, conserved sequence motifs and secondary structures.

#### **1.2.2.3 Long non-coding RNAs**

An interesting class of RNAs which has only recently been getting appropriate attention are long non-coding RNAs (lncRNAs). These RNAs have historically been referred to as 'junk' however more recently evidence is being presented for their critical function in maintaining cellular homeostasis. lncRNAs are generally transcribed by RNA polymerase III and have been identified to carry a number of functions (Wilusz *et al.* 2009) such as; regulating transcription (Vance and Ponting 2014) and acting as molecular sponges (Ebert *et al.* 2007; Xia *et al.* 2014). The most famous example is the activity of the lncRNA Xist in X-chromosome inactivation. Xist is a lncRNA 17Kb in length in Humans, which orchestrates the silencing of the X-chromosome

in females. Xist is known to physically associate with and coat one of the X-chromosomes, specifically within gene-dense regions, leading to its inactivation (Engreitz *et al.* 2013).

#### **1.2.2.4 tRNAs**

tRNAs are non-coding RNAs of approximately 80 nucleotides in length that function to deliver specific amino acids to the ribosome during translation. tRNAs are transcribed by RNA polymerase III as pre-tRNAs which then undergo regulated processing into the mature, structured RNA containing 3 hairpin loops and a 3' CCA tail. Before nuclear export the CCA tail is charged with the amino acid by an aminoacyl tRNA synthetase. Each amino acid has specific tRNAs which bear an anti-codon region which base pairs with the 3 nucleotide codon sequence within the mRNA when delivered to the T site in the translating ribosome. The only non-coding RNAs investigated in this thesis were miRNAs which are discussed in greater detail below.

#### **1.2.3 microRNAs**

miRNAs are small RNAs 18-22nt in length which generally function to repress translation and/or promote mRNA decay. The first miRNA, *let-7*, was discovered in *Caenorhabditis elegans* (*C. elegans*) in 1993 (Lee *et al.* 1993; Wightman *et al.* 1993) and since then a vast family of miRNAs have been observed in eukaryotes. As with all RNA families the total number of miRNAs within each organism varies massively as does their conservation. However, to date a total of 2585 mature miRNAs have been identified in humans whilst 462 are known to be present in *Drosophila melanogaster* (miRBase release 21 (June 2014)). Each of these miRNAs has the potential to regulate specific targets which is mainly achieved through binding to complementary sites in the UTRs (mainly 3') of mRNAs. Indeed, miRNAs have been shown to have regulatory roles in nearly all processes within the cell.

miRNA biogenesis has been well characterised and involves a number of key regulatory steps (Figure 1.1). The primary miRNA (pri-miRNA) is transcribed by RNA polymerase II and contains a stem loop within its structure which is subsequently cleaved by the microprocessor complex by an enzyme named Drosha into a pre-miRNA hairpin. The pre-miRNA is then exported to the cytoplasm by Exportin 5 and undergoes further rounds of processing. In the cytoplasm the looped region is cleaved by Dicer-1 leaving a double stranded RNA structure of 18-22 nucleotides in length. Contrary to initial belief both strands are able to behave as functional miRNAs with the miRNA from the 5' side of the cleaved hairpin named the 5p and the strand originally 3' of the hairpin the 3p. Due to their different sequences these miRNAs from the same pre-miRNA have distinct targets and have the potential to control completely different cellular pathways. It is however, important to note that it is often the case the 5p and 3p miRNAs are expressed at different levels which can be tissue or developmental stage specific (García-López

and del Mazo 2012). In addition to conventional biogenesis miRNAs can also be produced from introns which removes the requirement of Drosha. Introns removed by splicing were initially thought to be cellular 'junk', however, it is now clear that a number of critical non-coding RNAs can be processed from introns. The pre-miRNA hairpin can be cleaved from the intronic sequence and subsequently processed into the mature miRNA sequence as outlined above (Figure 1.1).

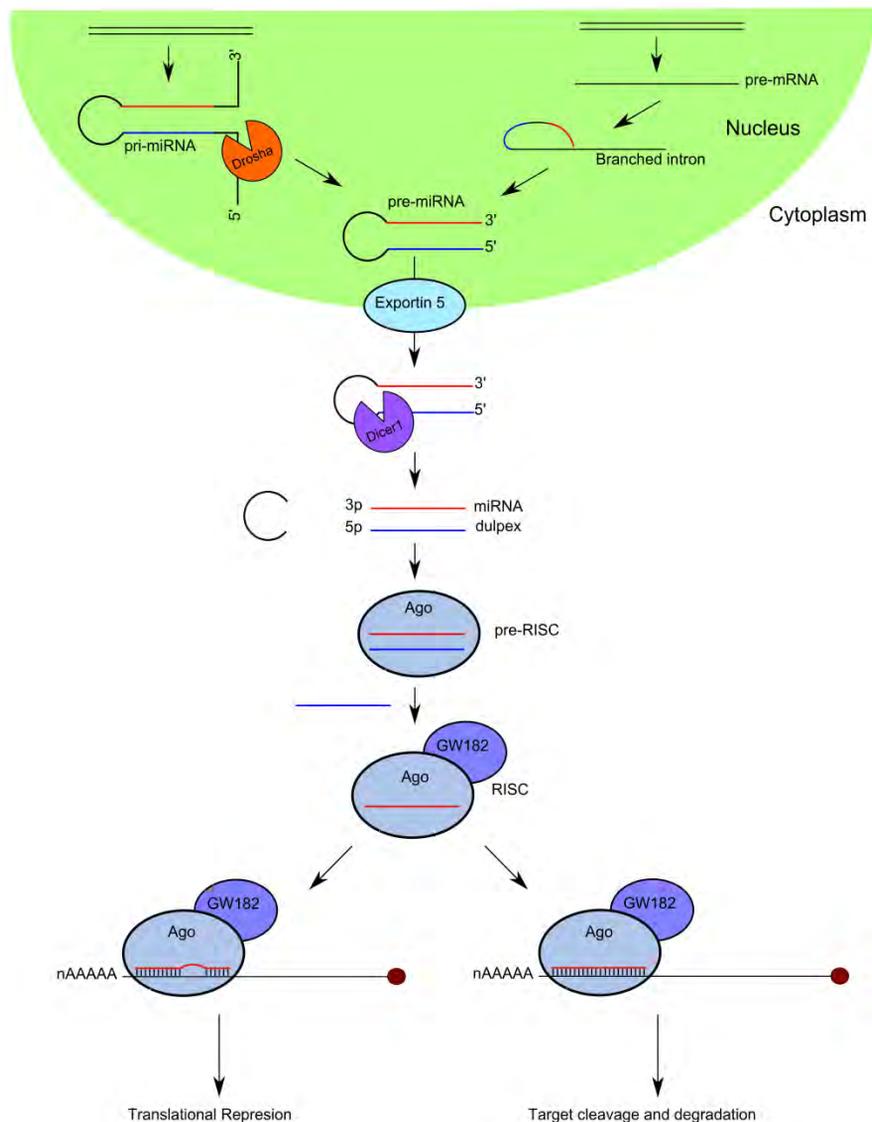
The miRNA duplex is loaded into an Argonaute protein using the molecular chaperones HSP70 and HSP90 (Iwasaki *et al.* 2010) and the dsRNA binding protein TRBP in an ATP-dependent manner (Chendrimada *et al.* 2005; Wang *et al.* 2009). Following its loading into Argonaute, the duplex is unwound by Argonautes' PAZ domain (Gu *et al.* 2012; Kwak and Tomari 2012) and single strand is retained and further processed with the other being emitted from the Argonaute complex. These processing steps include 3' end tailing and trimming (Han *et al.* 2011; Liu *et al.* 2011; Heo *et al.* 2012; Ameres and Zamore 2013). This is particularly important for miRNAs derived from introns (miRtrons) as they often have extended tails outside of the mature sequence after Dicer-1 cleavage. Similarly Dicer-1 processing can produce dsRNAs of 24nt in length which also require trimming to the final mature length. A few exoribonucleases have been found responsible for these trimming processes with Nibbler identified as a 3'-5' exoribonuclease responsible for trimming around 40% of Ago1 bound miRNAs in *Drosophila* (Han *et al.* 2011; Liu *et al.* 2011; Westholm *et al.* 2012). In addition to Nibbler, the exosome has also been implicated in terminal trimming of miRNAs derived from introns in *Drosophila*, specifically *miR-1017* (Flynt *et al.* 2010).

Once a mature miRNA has been successfully processed it elicits its function through binding to complementary regions within target mRNAs which occurs mainly through the seed region (nucleotides 2-8 of the miRNA). These target sites tend to reside in the UTR's with the majority found in the 3' UTR. In order to target specific mRNAs the miRNA is incorporated into a targeting complex named the RNA Induced Silencing Complex (RISC). The major components of miRISC are Argonaute 1 (Ago1) into which the miRNA is pre-loaded during the final stages of biogenesis and GW182 (Zhang *et al.* 2007). It is widely accepted that the fate of the miRNA-targeted mRNA depends the level of complementarity between the miRNA and its target sequence within the mRNA. High complementarity has been shown to simulate cleavage of the mRNA transcript in a similar mechanism to siRNAs; resulting in the decay of both the mRNA and miRNA.

However, recently this has been shown to be the less common result of miRNA-target association as there are often mismatches within the more heterogeneous 3' end of the

mature miRNA. It is more widely believed that miRNA binding promotes translational repression which can subsequently lead to the degradation of the target mRNA from either the 5' or 3' end. A number of mechanisms behind translational repression have been proposed but most reside around the function of the Glycine-Tryptophan repeat containing protein GW182. For example, GW182 has been shown to recruit the deadenylase complexes Ccr4-Not and Pan2-Pan3 which stimulates the release of the protective PABP. This disrupts the PABP-eIF4E interaction resulting in the de-circularisation of the transcript resulting in a loss of the conformation required for efficient translation (Braun *et al.* 2011; Chekulaeva *et al.* 2011; Fabian *et al.* 2011; Zekri *et al.* 2013). The recruitment of deadenylases to the miRNA bound target by GW182 has also been shown to result in the association of the helicase eIF4A2 which can act as a clamp in the 5' UTR preventing ribosome progression (Meijer *et al.* 2013); similar to the role played by eIF4A3 in the formation of the exon junction complex (Chan *et al.* 2004; Shibuya *et al.* 2004). Ultimately the recruitment of deadenylases result in the removal of the protective poly(A) tail allowing exoribonucleolytic decay of the translationally repressed transcript in the 3'-5' direction (Braun *et al.* 2011; Meijer *et al.* 2013). The specific exoribonucleases responsible for this have not been experimentally determined and are likely to vary depending on the target mRNA. Similarly, the RISC is also able to recruit decapping machinery which remove the 5' cap allowing mRNA decay to proceed from the 5' end (Nishihara *et al.* 2013).

Ago1 and GW182 are required and sufficient for miRNA-mediated repression (Fabian and Sonenberg 2012). The impact of miRNAs on mRNA fate and ultimately the cell is demonstrated by the fact that mRNAs have evolved through negative selecting to avoid seed matches with miRNAs present in the same cells (Farh *et al.* 2005; Stark *et al.* 2005).



**Figure 1.1: Classical and alternative pathways of miRNA biogenesis.** A primary miRNA sequence (pri-miRNA) is transcribed containing a hairpin loop which is cleaved by the microprocessor complex containing the nuclease Drosha. The cleaved hairpin, or precursor miRNA (pre-miRNA) is then exported into the cytoplasm by Exportin 5. In the cytoplasm the loop of the pre-miRNA is cleaved by Dicer-1 resulting in the generation of a miRNA duplex. This duplex is then loaded in Argonaute 2 (Ago2) where a single strand of the miRNA duplex is selected and the other is removed from the complex. The Ago2-miRNA together with GW182 and other accessory proteins from the RNA induced silencing complex (RISC) which is guided by the miRNA to its targets. The effect of the RISC-target association is dependent upon the degree of complementarity. Perfect complementarity stimulates target cleavage by Ago2 and subsequent degradation of the fragments. In-perfect base pairing promotes translational repression which is the more common mechanism of miRNA-mediated repression. Alternatively a miRNA can be generated from an intron derived from pre-mRNA splicing. This is processed by a debanching enzyme, removing the requirement of Drosha processing, and is then exported into the cytoplasm where it then follows the same biogenesis pathway as the classical method.

## 1.3 RNA stability

### 1.3.1 The importance of RNA stability in gene expression

As discussed previously there are multiple RNA species within the cell all with specific and critical functions. Therefore it is essential that these RNAs are maintained at the correct level. For example, too much of specific mRNAs such as those encoding Myc and Fos could result in the overproduction of the oncogenic proteins. Similarly, if miRNAs are present for too long in the cell then their negative effects on mRNA expression may have drastic effects on the cell. Ultimately, the level of any specific RNA within the cell is a delicate balance between the level of its production (transcription) and its stability in terms of its rate of degradation. It is therefore critical to maintain both these processes and there are a number of factors responsible for this. Specifically, the loss of control over RNA degradation has been shown to have drastic effects which will be discussed in greater detail below (Carballo and Blackshear 2001; Akamatsu *et al.* 2005; Grima *et al.* 2008; Kiss and Andrulis 2010; Astuti *et al.* 2012; Jones *et al.* 2013; Waldron *et al.* 2015).

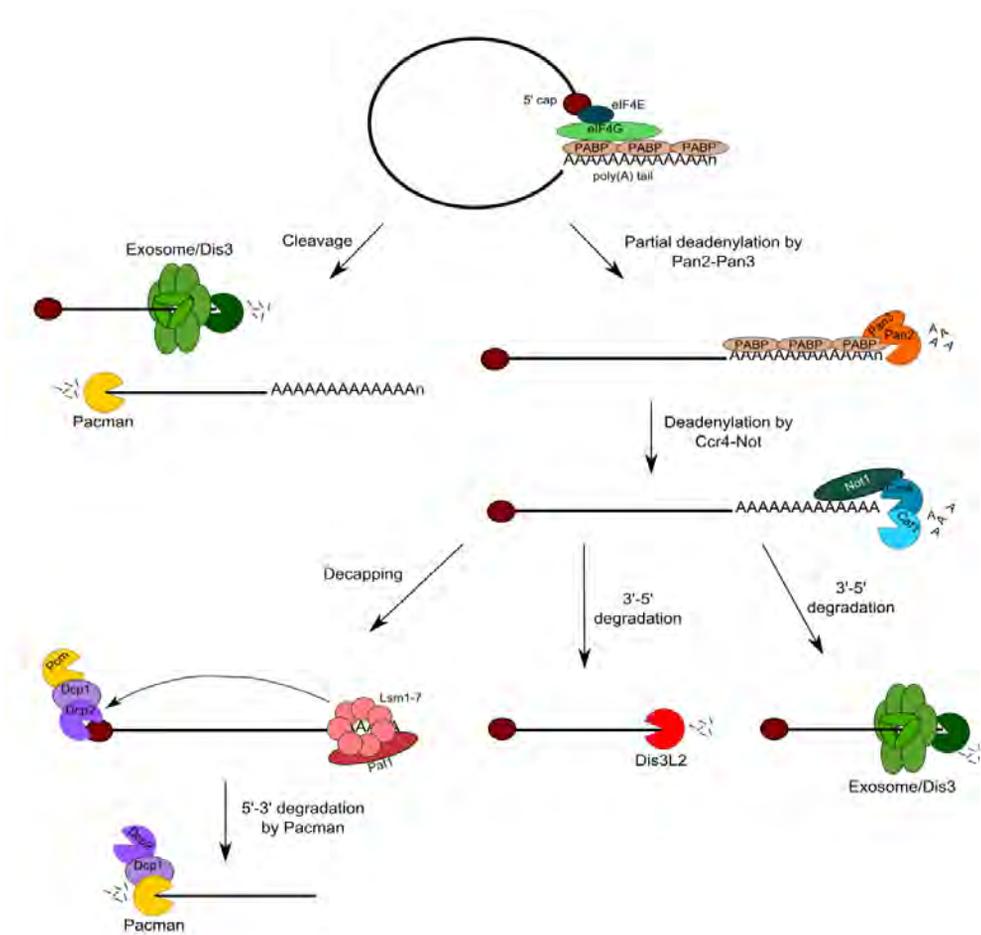
### 1.3.2 Pathways of RNA decay

In the cytoplasm an mRNA is normally protected from exoribonucleolytic decay by the 5' methylguanosine cap and the 3' poly(A) tail; additions which are also required for efficient translation. Therefore, for an mRNA to be degraded these protective additions must be removed to allow access of the exoribonucleases. It is now widely accepted that, in the majority of cases, the first step of mRNA decay involves the removal of the poly(A) tail. Deadenylase complexes are recruited to the mRNA resulting in the dissociation of the PABP and remove the 3' stretch of adenines. Following deadenylation by the Ccr4-Not or Pab2-Pab3 complexes the mRNA can be degraded in the 3'-5' direction by a family of enzymes related to bacterial RNase II/RNase R (Figure 1.2). Alternatively, following deadenylation further processing steps can occur involving the removal of the 5' cap. This is performed by the decapping complex which is comprised of two enzymes; Dcp1 and Dcp2. Following decapping the mRNA can be degraded in the 5'-3' direction by XRN1, known as Pacman in *Drosophila* (Figure 1.2).

In addition to these conventional pathways of mRNA decay an mRNA can also be internally cleaved by endoribonucleases which results in the generation of two fragments which can be degraded in either direction (Figure 1.2). This internal cleavage is normally a result of quality control mechanisms such as non-sense mediated decay (NMD) which functions to ensure that mRNAs containing errors are not translated. Endoribonucleolytic cleavage also occurs

following highly complementary miRNA binding or in the presence of small interfering RNAs (siRNAs).

In addition to mRNAs the other RNA species within the cell are also regulated at the level of their stability. A number of long non-coding RNAs are also polyadenylated and therefore would also require deadenylation prior to exoribonucleolytic decay from the 3' end. Unlike mRNAs, miRNAs do not have protective caps or poly(A) tails and therefore they rely on alternative mechanisms to avoid decay; these include 2'-O-methyl modifications (Ameres *et al.* 2010) together with being protected within the Argonaute complex. The mechanisms behind miRNA decay will be discussed in greater detail below.



**Figure 1.2: mRNA degradation pathways in *Drosophila*.** A translationally competent mRNA is protected from exonucleolytic decay by the protective 3' poly(A) tail and 5' cap which also mediate mRNA circularisation. The first step of mRNA decay is deadenylation which normally occurs in a two-step process. Partial deadenylation by Pan2-Pan3 followed by the completion of deadenylation by the Ccr4-Not complex. This then provides a vulnerable 3' end which can be degraded in the 3'-5' direction by either Dis3L2 or the cytoplasmic exosome complex, which in *Drosophila*, associates with Dis3. Alternatively, further processing can occur stimulated by the binding of the Lsm1-7-Pat1 complex which binds to the 3' end, blocking 3'-5' decay and promoting decapping. Decapping is performed by the Dcp1/Dcp2 decapping complex which also physically associates with the 5'-3' exoribonuclease Pacman (Pcm). Following decapping, Pacman is able to access the 5' end of the mRNA resulting in its degradation in the 5'-3' direction. Alternatively, the circularised mRNA can be cleaved by endoribonucleolytic cleavage which generates two fragments. One of which can be degraded in the 5'-3' direction by Pacman whilst the other is degraded 3'-5' by the exosome/Dis3.

### 1.3.3 An overview of 3'-5' decay.

#### 1.3.3.1 Deadenylation

As outlined above, it is generally accepted that the first, and rate limiting step in an mRNA's decay is the removal of the poly(A) tail (Muhlrad and Parker 1992; Decker and Parker 1993; Yamashita *et al.* 2005; Zheng *et al.* 2008). This occurs in a two-step mechanism following the recruitment of deadenylase complexes and displacement of the protective PABP. The length of the poly(A) tail is critical in determining the stability of an mRNA (Eckmann *et al.* 2011), in which the pab1p-dependent poly(A) nuclease (PAN) plays a critical role (Brown and Sachs 1998). It is likely that the shorter poly(A) tails promote instability as they are unable to provide binding sites of sufficient size for protective proteins, such as PABP, to bind.

Initially, in the cytoplasm, mRNAs carry stabilising poly(A) tails of around 80 nucleotides in yeast and 250 nucleotides in mammals (Wahle 1995; Viphakone *et al.* 2008). These are then shortened in a stepwise manner involving the two main deadenylase complexes. The first step is thought to involve distributive trimming by the Pan2-Pan3 complex (Brown and Sachs 1998; Yamashita *et al.* 2005). The deadenylase activity is provided by the Pan2 protein and, unlike other deadenylation mechanisms, is stimulated via Pan3-PABP interaction (Jonas *et al.* 2014).

Following the initial trimming by PAN the processive Ccr4-Not pentameric complex completes the remaining steps of deadenylation and subsequent RNA degradation. The nuclease activity of this complex is provided by two enzymes in humans and *Drosophila*; Caf1 and Ccr4 (Temme *et al.* 2010; Maryati *et al.* 2015). Interestingly, in *Saccharomyces cerevisiae* (*S. cerevisiae*), Caf1 has lost its deadenylase activity (Goldstrohm *et al.* 2007) although the reason for this remains unknown. Along with the catalytic proteins the Not proteins function as a scaffold from which the deadenylases elicit their function. Unlike the Pan2/3 complex, the Ccr4-Not complex is inhibited by PABP and therefore its dissociation is required before its activity.

Interestingly, although Pan2-Pan3 does not participate in bulk deadenylation (Temme *et al.* 2010) it is able to compensate for the loss of the Ccr4-Not complex (Wahle and Winkler 2013). This indicates a level of redundancy between the two complexes which is an important fail-safe mechanism to maintain control over poly(A) tail length and mRNA stability. In *Drosophila*, the Pan and Ccr4-Not complexes are the only known deadenylases to function in mRNA decay; however, the potential diversity of deadenylases is demonstrated by the fact that *Arabidopsis thaliana* is thought to have up to 26 proteins capable of functioning as deadenylases (Goldstrohm and Wickens 2008). Similarly, humans (along with most metazoans) contain an extra deadenylase complex in poly(A) specific ribonuclease (PARN), these different deadenylase complexes may function to allow specific targeting (Lee *et al.* 2012).

Deadenylation also stimulates decapping and 5'-3' degradation which is discussed in latter sections.

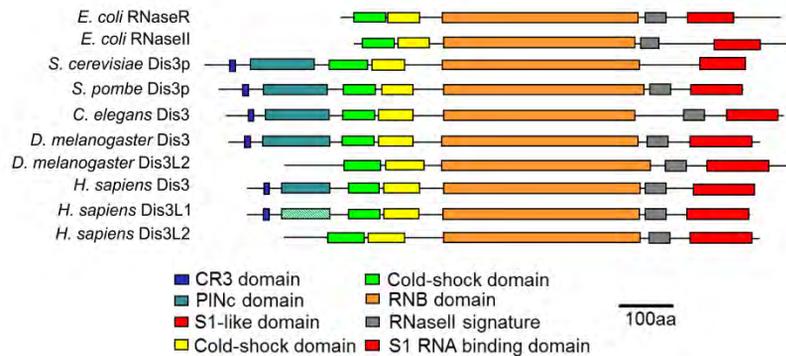
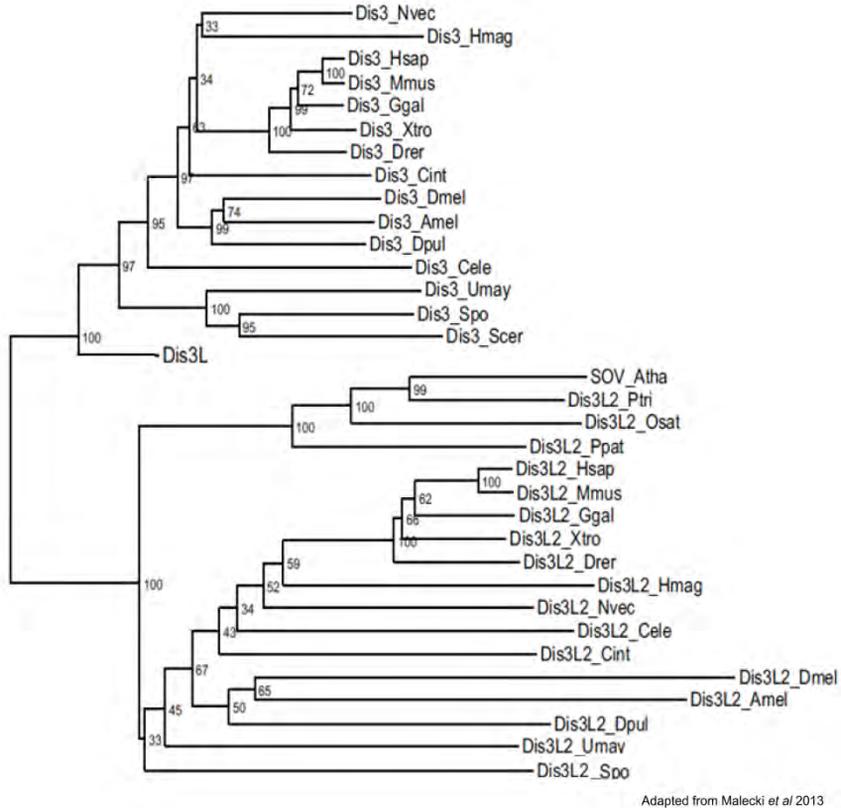
### 1.3.3.2 The Dis3 family

The process of deadenylation leaves the 3' end vulnerable to exoribonucleolytic attack by a family 3'-5' exoribonucleases. This highly conserved family of exoribonucleases are homologous to the bacterial RNaseII/RNaseR family of ribonucleases (Figure 1.3). Between them these hydrolytic, processive exoribonucleases are responsible for the bulk of 3'-5' RNA decay. In *S. cerevisiae* there is a single member of this family, Dis3. However, in *Schizosaccharomyces pombe* (*S. pombe*) and all higher eukaryotes, including *Drosophila*, a second member, Dis3L2, is also present. Finally, in higher eukaryotes a third paralogue is present, namely Dis3L1.

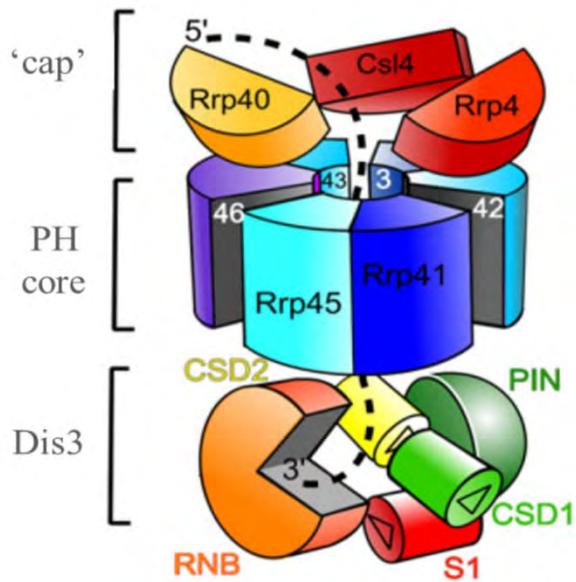
Dis3, the only member within all eukaryotes, provides catalytic activity to the multi subunit complex named the exosome (Figure 1.4) (Liu *et al.* 2006; Wasmuth and Lima 2012; Makino *et al.* 2013). The exosome core consists of nine catalytically inert subunits which form a barrel like structure, similar to bacterial PNPase (Liu *et al.* 2006) which functions to channel the RNA substrate into the exoribonuclease active site of Dis3 (Bonneau *et al.* 2009). Six of these subunits are RNasePH domain containing proteins (Rrp41, 42, 43, 45, 46 and Mtr3 which together form barrel-like structure. The remaining three subunits (Rrp4, 40 and Csl4) are RNA binding proteins which form a cap to which RNA associated before being thread through the channel (Symmons and Luisi 2009) (Figure 1.4). This structure is thought to act to modulate the highly processive activity of Dis3 to ensure that RNAs are present for long enough to elicit their function (Liu *et al.* 2006; Drazkowska *et al.* 2013). In all eukaryotes that do not encode Dis3L1, Dis3 associates with the exosome in both the cytoplasm and the nucleus and therefore has a wide variety of target substrates, however, in higher eukaryotes its localisation is mainly restricted to the nucleus (Tomecki *et al.* 2010) where it functions mainly in quality control.

Dis3 has both exo- and endoribonucleolytic activity provided by the RNB and PiLT N-terminal (PIN) domains respectively (Schneider *et al.* 2007; Lebreton *et al.* 2008) meaning it has the capability to degrade a variety of substrates. The PIN domain is also responsible for its association with the exosome. Dis3L1, a paralogue of Dis3, also associates with the exosome, but its localisation is strictly cytoplasmic (Staals *et al.* 2010; Tomecki *et al.* 2010). Unlike Dis3, the PIN domain of Dis3L1 has lost endoribonuclease activity but is still critical for its interaction with the exosome. Of the Dis3 family, Dis3L1 is the least well studied and no specific targets have been identified.

The final member of the Dis3 family is Dis3L2. Dis3L2 differs from the other members in that it does not contain the PIN domain and therefore is the most similar to its bacterial ancestor RNasell. As a result Dis3L2 does not carry endoribonucleolytic activity and functions independently of the exosome (Lubas *et al.* 2013). Like Dis3L1, Dis3L2 expression has been shown to be strictly restricted to the cytoplasm (Lubas *et al.* 2013; Malecki *et al.* 2013). Dis3 and Dis3L2 are the two enzymes on which this thesis is focused and therefore will be discussed in greater detail below.



**Figure 1.3: The Dis3 family of 3'-5' exoribonucleases.** The Dis3 family is a highly conserved family of 3'-5' exoribonucleases (family tree adapted from Malecki *et al* 2013). The domain organisation within the Dis3 family is conserved back to the RNaseII/R family of prokaryotic RNases. The exoribonuclease activity is provided by the RNB domain (orange). The cold-shock domains (green and yellow) and the S1 domain provide RNA binding capabilities. Eukaryotic Dis3 has an N-terminal extension containing a PIN domain (light blue) which gives endoribonuclease activity and a CR3 motif (dark blue) which together with the PIN domain mediates exosome interaction. Dis3L2 does not contain these N-terminal domains whilst Dis3L1 contains an inactive PIN domain (blue hash).



**Figure 1.4: Model of association between Dis3 and the exosome core.** Dis3 associates with the base of the exosome core through its PIN domain and CR3 motif to provide catalytic activity. The exosome core is a barrel consisting of 9 subunits 6 of which form the barrel whilst the remaining 3 form a cap and are responsible for substrate binding. The RNA substrate (black dash line) is fed through the central channel into the catalytic RNB domain which is positioned at the base of the channel (Adapted from Luisi *et al* 2009).

### 1.3.3.3 Other core factors involved in 3'-5' decay

In addition to the Dis3 family a number of other proteins and protein complexes are involved in controlling 3'-5' RNA decay. The first of these is another exoribonuclease, Rrp6. Unlike the members of the Dis3 family, Rrp6 is a distributive exoribonuclease which is conserved throughout eukaryotes and associates with the nuclear exosome with enrichment in the nucleolus (Briggs *et al.* 1998; van Dijk *et al.* 2007; Tomecki *et al.* 2010). Rrp6 associates towards the top of the exosome core (Makino *et al.* 2013; Wasmuth *et al.* 2014) and has been shown to be required for the correct processing of the 5.8S rRNA (Mitchell *et al.* 1997; Briggs *et al.* 1998) together with a role in the nuclear degradation of miRNA precursors, with distinct substrate preferences between it and Dis3 (Liu *et al.* 2014). Unlike Dis3, there are also examples of Rrp6 showing exosome-independent targets and functions in algae (Ibrahim *et al.* 2010) although there is greater literature on its exosome-dependent roles.

Together with the exoribonucleases themselves there are cofactors which aid the exosome in its activity. For example, the nuclear exosome requires targeting by the TRAMP complex. TRAMP consists of a distributive nuclear poly(A) polymerase named Trf4/5, an RNA binding protein Air1/2 and a helicase in Mtr4 (LaCava *et al.* 2005; Vanáčová *et al.* 2005). It has been shown that the addition of adenine residues to the end of exosome targets in the nucleus such as cryptic unstable transcripts (CUTs), rRNAs and snoRNA precursors, by the TRAMP complex is required for their degradation (LaCava *et al.* 2005; Wyers *et al.* 2005; Slomovic and Schuster 2011). Additionally, the helicase activity of Mtr4 also plays an important role in allowing the unwinding of structured substrates prior to their degradation by the exosome (LaCava *et al.* 2005). Similarly, another accessory complex is present in the cytoplasm which is required for the cytoplasmic activity of the exosome (Anderson and Parker 1998). This is the Ski complex which is comprised of Ski3, Ski8 and the core helicase Ski2 (Twister in *Drosophila*) (Anderson and Parker 1998; Brown *et al.* 2000). The Ski complex associates with the top of the cytoplasmic exosome together with another cofactor, Ski7, and functions to unwind structured RNA before it is fed into the central channel of the exosome (Halbach *et al.* 2013). These factors are all associated with exosome-mediated 3'-5' decay; this is because as yet there is no strong evidence for cofactors involved in Dis3L2-mediated decay (discussed below).

## 1.3.4 An overview of 5'-3' decay.

### 1.3.4.1 Decapping

Following deadenylation (see section 1.3.3.1) instead of being degraded 3'-5' an mRNA can be decapped and degraded 5'-3'. Deadenylation dependent decapping requires the displacement of the PABP protein which occurs following the Ccr4-Not mediated digestion of the poly(A) tail. Once the poly(A) tail is 10nt in length it acts as an association site for the Pat1/Lsm1-7 complex (Chowdhury *et al.* 2007). The effects of Pat1/Lsm1-7 binding to the target are twofold; it protects the transcript from 3'-5' mediated degradation whilst promoting subsequent decapping and 5'-3' degradation (Tharun and Parker 2001; Tharun *et al.* 2005).

The binding of Pat1/Lsm1-7 to the oligo(A) tail results in the recruitment of the decapping machinery to remove the protective 5' methylguanosine cap. The catalytic activity of the decapping complex is provided by the protein Dcp2. Dcp2 is stimulated by a variety of other decapping factors, of which Dcp1 is required for efficient removal of the cap (She *et al.* 2008). This occurs through the physical association of Dcp1-Dcp2 which triggers a conformational change resulting in Dcp2 adopting its active, closed conformation (She *et al.* 2008). In addition to Dcp1 other decapping activators are also known such as Edc1, 2, 3 and 4 (Hatfield *et al.* 1996; Dunckley *et al.* 2001; Kshirsagar and Parker 2004); of which Edc3 has been shown to physically bind Dcp1/2 (Tritschler *et al.* 2007; Harigaya *et al.* 2010).

Alternative stimulatory mechanisms of decapping have also been shown. For example, for transcripts undergoing miRNA mediated silencing the decapping complex (Dcp1/2) is recruited by GW182, a member of RISC (Rehwinkel *et al.* 2005). Other examples of deadenylation independent decapping have also been shown by Badis *et al.*, Muhlrud *et al.* and more recently by Rissland *et al.* (Badis *et al.* 2004; Muhlrud and Parker 2005; Rissland *et al.* 2007; Rissland and Norbury 2009). This latter example, in *S. pombe*, was shown to be stimulated by the addition of uridine nucleotides to the 3' end of the poly(A) tail by a nucleotidyl transferase *Cid1* (Rissland and Norbury 2009). This uridylation recruits the Lsm1-7 complex which in turn promotes decapping and 5'-3' decay (Rissland and Norbury 2009). Recent work in mammalian cells has revealed widespread uridylation 3' to the poly(A) tail confirming that the findings in yeast are likely to be conserved (Chang *et al.* 2014).

#### 1.3.4.2 XRN1

After a transcript has been decapped it is vulnerable to attack from the 5' end. In the cytoplasm the only 5'-3' exoribonuclease is XRN1 which is responsible for the decay of decapped transcripts. XRN1 itself, known as Pacman in *Drosophila*, has been shown to directly bind to decapping factors through its unstructured C-terminal resulting in its co-recruitment and allowing efficient decay following cap removal (Bouveret *et al.* 2000; Braun *et al.* 2012). It appears that recruitment of XRN1 with decapping factors is conserved; however, the factor to which it is bound can be species specific (Jonas and Izaurralde 2013).

XRN1 itself is highly conserved throughout eukaryotes with the catalytically active N-terminal in *Drosophila* showing 70% homology to the same domain in humans (Jones *et al.* 2012). It is the only known 5'-3' exoribonuclease in the cytoplasm and therefore is responsible for the degradation of a variety of RNA substrates including more structured substrates due to its ability to unwind secondary structure (Jinek *et al.* 2011). The critical activity of XRN1 has been shown in a number of studies with it being required for adult viability in *Drosophila* (Waldron *et al.* 2015). Mutations in *pacman* in *Drosophila* have also been shown to result in delayed development and reduced tissue growth due to an increase in apoptosis (Grima *et al.* 2008; Jones *et al.* 2013; Waldron *et al.* 2015). Similarly, mutations in yeast *Xrn1*, although viable, result in an accumulation of deadenylated and decapped transcripts showing its vital role in RNA clearance (Brown *et al.* 2000).

All the components required for 5'-3' degradation, from the deadenylases through to XRN1 have been shown to localise to processing bodies (P bodies) (Zheng *et al.* 2008; Temme *et al.* 2010). Consistent with this deadenylation, the process that functions upstream of decapping, was shown to be a prerequisite for P body formation (Zheng *et al.* 2008). P bodies are cytoplasmic foci which act as a storage unit for translationally repressed mRNAs; they are either subsequently degraded or re-enter translation if required. It was in fact XRN1s distinct localisation within foci that led to their discovery (Bashkurov *et al.* 1997). This is in some contrast the 3'-5' enzymes which although have been shown to be present in P bodies (Malecki *et al.* 2013) generally show a more diffuse expression pattern. It is possible that the target specificity of XRN1 is achieved through the localisation of specific transcripts to these P bodies at particular times. In addition to a role in general mRNA decay XRN1 also plays a role in the clearance of endonucleolytically cleaved in the cytoplasm. For example, XRN1 is required for the removal of the 3' fragment of transcripts that do not pass quality control (NMD) (Conti and Izaurralde 2005; Isken and Maquat 2007) or for transcripts cleaved by RISC following miRNA targeting (Orban and Izaurralde 2005).

## 1.3.5 Mechanisms of regulating mRNA stability

### 1.3.5.1 AU rich elements (AREs)

The historical view of RNA decay was that it was a general housekeeping mechanism with little specificity. However, the half-lives of mRNAs vary dramatically and it has since been shown that there are a number of neat mechanisms which aid in targeting specific transcripts to the decay machinery (Figure 1.5). The most extensively studied of these are AU rich elements (AREs) which are found in the 3' UTR of a number of transcripts. These AREs were identified as early as 1986 (Shaw and Kamen 1986), where they were identified to promote instability of *GM-CSF*. It has now been shown that between 5-8% of human genes encode ARE containing mRNAs which, although encode proteins of varying functions, encode a number of oncogenes and cytokines known to have short half-lives (Bakheet *et al.* 2001; Bakheet *et al.* 2003; Schoenberg and Maquat 2012).

AREs promote instability of their mRNAs through recruiting specific RNA binding proteins which in turn modulate the RNA decay machinery. Although nearly all of the major cytoplasmic exoribonucleases have been implicated in the decay of ARE containing RNAs (Chen and Shyu 1995; Chen *et al.* 2001; Stoecklin *et al.* 2006; Lubas *et al.* 2013), the exosome remains the best characterised in its targeting of AREs. It has been shown that a 266 amino acid domain (RNase PH) within the core exosome subunits Rps41, 43 and 45 is responsible for its binding to the AU stretches (Anderson *et al.* 2006). There are a number of ARE binding proteins which also play critical roles in regulating transcript stability; these include TTP, Auf1 and HuR. The crucial role TTP plays, for example, is exemplified by the fact that its loss in mice results in the increased stability of *TNF- $\alpha$*  and *GM-CSF* subsequently resulting in debilitating inflammation (Carballo and Blackshear 2001). TTP itself is tightly regulated by phosphorylation by MAP-kinase activated protein kinase 2 (MK2) which shows the intricacy of these targeting mechanisms (Mahtani *et al.* 2001). Finally, together with promoting instability proteins binding to AREs can promote translation and mRNA stability such as the Hu family of proteins (Abdelmohsen and Gorospe 2010).

### 1.3.5.2 Regulation by 3' tailing

Another emerging mechanism of targeting RNAs for cytoplasmic degradation is through 3' nucleotide tagging. 3' tail tagging has been shown to promote both the 5'-3' and 3'-5' pathways of RNA decay (Mullen and Marzluff 2008). Although the knowledge of how addition of adenines to nuclear transcripts promote exosome-mediate decay or trimming (LaCava *et al.* 2005) the extent to which the same mechanism is used in the cytoplasm is not fully understood. Recent work identified that, although the frequency varies dramatically, around

80% of all mammalian transcripts show some level of terminal tagging (Chang *et al.* 2014). A good example of tagging comes in the regulation of histone RNAs. This involves the oligouridylation of the histone transcripts which stimulates the binding of the Lsm1-7 complex and subsequent decapping and degradation in both the 5'-3' and 3'-5' direction (Mullen and Marzluff 2008). This uridylation has since been shown to be performed by TUT4 (Schmidt *et al.* 2011). Another example of uridylation directing decay is during the biogenesis of *let-7a*. During the maturation process of *let-7a* the *pre-let7a* hairpin can either be processed by Dicer2 into the mature *let-7a* or degraded; this decision is dependent upon the extent of 3' uridylation. Monouridylation promotes Dicer2 processing and maturation (Heo *et al.* 2012) whilst the addition of multiple uridines results in its degradation by Dis3L2 (Chang *et al.* 2013; Ustianenko *et al.* 2013). The uridylyltransferases TUT4 and TUT7 have been shown to be responsible for this oligouridylation (Chang *et al.* 2013; Ustianenko *et al.* 2013).

### **1.3.5.3 Factors promoting RNA stability**

In addition to those mechanisms of promoting RNA decay other mRNAs can also hold elements that inhibit their degradation. A neat example of this is observed in a family of viruses which contain structural elements within their 3' UTR used to override the cells decay machinery. These function to stall XRN1 which allows the expression of small flaviviral non-coding RNAs (sfRNAs). Together with inhibiting XRN1 activity these sfRNAs accumulate in the cytoplasm, promote mRNA stabilisation and drive viral pathogenesis (Chapman *et al.* 2014; Chapman *et al.* 2014; Moon *et al.* 2015). There are also examples of RNA binding proteins that block the decay machinery such as the previously described poly(A) binding protein. Another example is in the case of Rpm1 which has been shown to bind to specific motifs and promote the stability of transcripts encoding mitochondrial components in *S. pombe* (Hasan *et al.* 2014).

### **1.3.5.4 miRNA mediated regulation.**

miRNA mediation of RNA stability provides a final example of a mechanism of regulating mRNA stability. As discussed previously miRNAs are small non-coding RNAs of 18-22nt in length and function to regulate miRNA stability. miRNAs mainly function by binding to complementary sequences within the UTRs of their target mRNAs, resulting in translational repression and/or degradation of the target mRNA (section 1.2.3). Most miRNAs show only partial complementarity and therefore normally promote translational repression which may be followed by degradation. Interestingly, some studies have also shown miRNAs are able to promote translation of their targets (Vasudevan *et al.* 2007). For example *miR-122* has been shown to stimulate translation of Hepatitis C RNA (Henke *et al.* 2008; Masaki *et al.* 2015) whilst *miR-10a* promotes translation of ribosomal protein mRNAs through binding their 5' UTR (Ørom *et al.* 2008).

Mechanism of regulating RNA stability	Enzymes responsible for modification	Decay enzyme(s) recruited	Organism	Reference(s)
AU rich elements (AREs)	N/A	XRN1 Dis3/exosome Dis3L2	Sp, Dm, Mm, Hs	Shaw and Kamen 1986, Chen and Shyu 1995, Stoecklin <i>et al</i> 2006, Lubas <i>et al</i> 2013.
Nuclear adenylation	TRAMP complex	Dis3/exosome Rrp6/exosome	Sp, Dm, Mm, Hs.	Vanacova <i>et al</i> 2005, LaCava <i>et al</i> 2005, Wyers <i>et al</i> 2005, Slomovic and Schuster 2011.
Uridylation of pre-miRNAs	TUT4, TUT7, Tailor	Dis3/exosome Rrp6/exosome Dis3L2	Dm, Mm, Hs	Liu <i>et al</i> 2014, Chang <i>et al</i> 2013, Ustianenko <i>et al</i> 2013
mRNA uridylation	Unknown	Dis3L2	Sp	Malecki <i>et al</i> 2013
Histone mRNA uridylation	TUT4	XRN1 Dis3/exosome	Hs	Mullen and Marzluff 2008, Schmidt <i>et al</i> 2011,
miRNA 3' monoadenylation	GLD2	Stabilisation effect	Mm	Katoh <i>et al</i> 2009
miRNA 2'-O-methylation	Hen1	Stabilisation effect	Dm	Ameres <i>et al</i> 2010

**Figure 1.5: Mechanisms of regulating RNA stability.** There are a number of mechanisms known to regulate RNA stability and target specific degradation enzymes. Sp = *Schizosaccharomyces pombe*, Dm = *Drosophila melanogaster*, Mm = *Mus musculus*, Hs = *Homo sapiens*.

### 1.3.6 Mechanisms of regulating miRNA stability

The biogenesis and biological function of miRNAs have been extensively studied; however, the factors regulating the stability of miRNAs themselves have been somewhat overlooked. miRNAs have long been thought to be one of the more stable members of the RNA family. A study using *dicer1* knockout mouse embryonic fibroblasts estimated the average half-life of miRNAs to be 119 hours (Gantier *et al.* 2011). However, the critical roles that miRNAs play in developmental timing and tissue growth implies there needs to be tight control over their stability. A prime example of this is with the *miR-16* family which are highly regulated during the cell cycle (Rissland *et al.* 2011). Similarly *miR-29b* is maintained at a low level during all stages except during mitosis, when its level increases dramatically due to enhanced stability rather than increased transcription (Hwang *et al.* 2007). There are a number of mechanisms through which this regulation may occur which are discussed below (Figure 1.6)

#### 1.3.6.1 Modifications to mature miRNAs.

A number of studies have uncovered an elegant yet simple mechanism for miRNA regulation through nucleotide additions to their 3' ends. For example, 3' monoadenylation of *miR-122* by the poly(A) polymerase GLD2 (PAPD4) results in the stabilisation of the miRNA in the liver (Katoh *et al.* 2009). In addition to GLD2, the enzymes MTPAP, PAPOLG, and TUT1 have also been demonstrated to be involved in miRNA 3' adenylation in human cells (Burroughs *et al.* 2010; Wyman *et al.* 2011). Together with adenylation, 3' uridylation has emerged as a mechanism for controlling miRNA activity. The importance of which is shown through the fact that the loss of the uridylyltransferases TUT4 and TUT7 in zebrafish results in developmental defects. This has been shown to result in the aberrant expression of miRNAs predicted to regulate key Hox genes (Thornton *et al.* 2014); suggesting that uridylation promotes instability of the miRNAs. A final modification observed is that of 2'-O methylation in *Drosophila*. This modification has been shown to increase the stability of miRNAs at least partially through the inhibition of 3' tailing by nucleotidyl transferases (Ameres *et al.* 2010). 2'-O methylation is performed by Hen1 and has also been shown to aid Ago loading, which in turn may also protect miRNAs from degradation (Cairrão *et al.* 2005; Czech *et al.* 2009; Okamura *et al.* 2009).

### 1.3.6.2 Degradation of miRNAs by exoribonucleases

As yet the exoribonucleases responsible for the degradation of specific miRNAs largely remain elusive. Identification of these enzymes is key to the understanding of how miRNA stability is regulated and has been somewhat neglected. However, there are some examples in the literature of specific miRNA degradation by either 5'-3' or 3'-5' decay machinery. An example of this is in the regulation of the key developmental miRNA let-7 by XRN2 in *C. elegans* (Chatterjee and Groszhans 2009). It is interesting that a nuclear exoribonuclease is able to specifically regulate a substrate that is thought to be present in the cytoplasm, although miRNAs such as *miR-29b*, which contain a hexanucleotide sequence directing nuclear import, have been observed in the nucleus (Hwang *et al.* 2007). Similarly, XRN1 was shown to have a modest effect on the stability of *miR-382* in human cells (Bail *et al.* 2010).

In addition to the 5'-3' pathway, the 3'-5' degradation pathway has also been observed to play a part in the regulation of miRNA stability. In the study on *miR-382* described above, knockdown of the exosome component Rrp41 in HEK293 cells resulted in stabilisation of *miR-382*, although the catalytic exoribonuclease responsible for this degradation was not identified (Bail *et al.* 2010). Another exoribonuclease, the IFN-inducible 3'-5' exoribonuclease polynucleotide phosphorylase (PNP) has been shown to target *miR-221* for degradation in the human HO-1 cell line (Das *et al.* 2010). Finally, the 3'-5' exoribonuclease ERI1 has been implicated in negatively regulating miRNA expression as natural killer and CD4<sup>+</sup> T cells mutant for *Eri1* showed a significant increase in global miRNA abundance (Thomas *et al.* 2012). The limited number of examples shows the inadequate level of knowledge of the exoribonucleases responsible for miRNA degradation and further indicates the importance of their identification if we are to fully understand how miRNAs function within the cell.

### 1.3.6.3 Target mediated miRNA regulation.

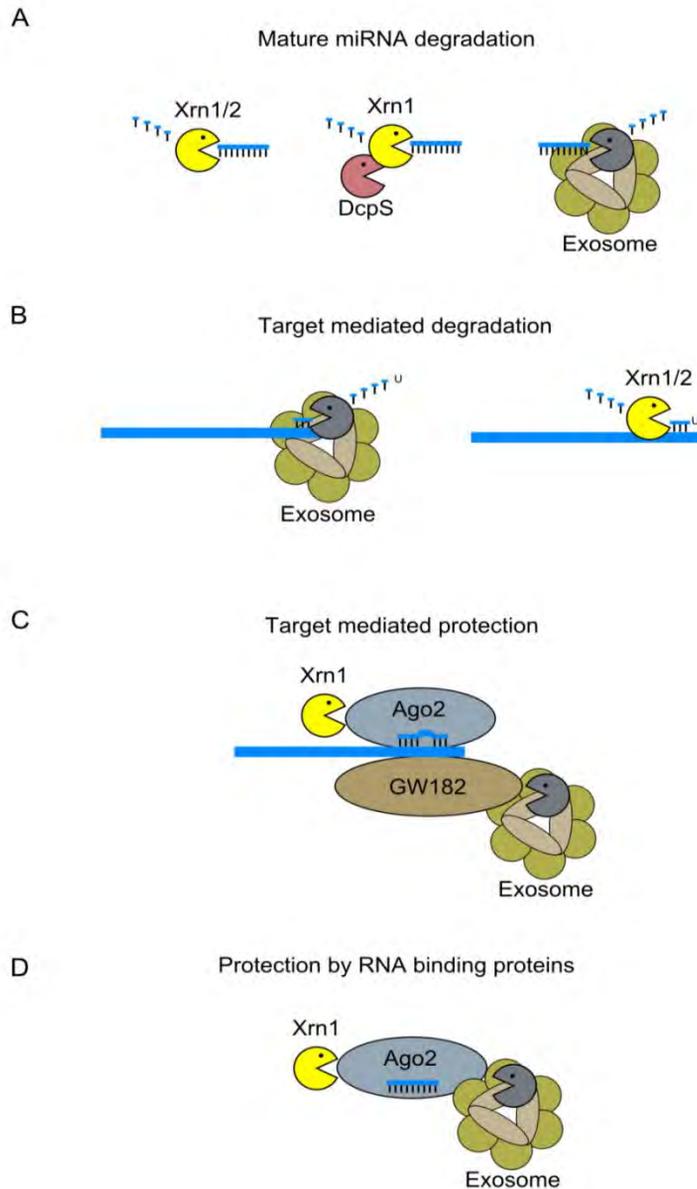
A further method to control miRNA abundance appears to be through co-degradation with its mRNA target. In *Drosophila*, for example, degradation of *let-7*, *miR-34* and *bantam* has been observed when extensive complementarity exists between miRNAs and their targets (Ameres *et al.* 2010). This phenomenon, termed target-mediated miRNA decay (TMMD), has also been observed in human cells, with *miR-223* a specific example (Baccarini *et al.* 2011), and in rodent neurones (de la Mata *et al.* 2015).

In contrast, miRNA-target interaction can also promote miRNA stability, although this tends to be the case with a lower level of complementarity is present. Target association prevents the release of the miRNA leading to its protection within the RISC (Chatterjee *et al.* 2011). Consistent with this GW182 and Argonaute, the major components of the RISC, promote

miRNA stabilisation. Loss of GW182 results in the destabilisation of *miR-146a* and *miR-155* (Yao *et al.* 2012) whilst the stabilities of *miR-21*, *miR-16*, *miR-20a*, *miR-92* and *let-7a* were significantly reduced following Ago2 knockdown in mouse embryonic fibroblasts (Diederichs and Haber 2007; Winter and Diederichs 2011).

#### **1.3.6.4 *cis*-acting stability elements.**

Together with the specific mechanisms discussed previously there are additional mechanisms inherent to miRNA stability. Interestingly, it appears that like mRNAs, miRNAs contain intrinsic instability elements within the mature miRNA sequence. Similar to the function of AREs described previously, miRNAs containing high AU/UA dinucleotide densities show the shortest half-lives in human neuronal cells (Sethi and Lukiw 2009). *miR-382* in HEK293 cells presents another example of intrinsic stability elements, as specific sequences at positions 16-22 have been shown control its stability (Bail *et al.* 2010). A final, corroborating finding is by Rissland *et al* who showed that the instability of *miR-503* was dependent upon the seed region and nucleotides in the 3' end (Rissland *et al.* 2011).



**Figure 1.6: Factors affecting miRNA degradation.** **(A)** Mature miRNAs have been identified to be degraded by exoribonucleases such as the XRN1/2 family and members of the exosome. **(B)** High complementarity between a miRNA and its target(s) can result in destabilisation of the mature miRNA termed target-mediated miRNA destabilisation (TMMD). The mechanisms by which the mature miRNAs become available for decay and how high complementarity stimulates miRNA decay remain unknown. **(C)** Alternatively, lower complementarity of miRNA-target association has been shown to result in stabilisation of mature miRNAs through protecting the miRNA sequence from exoribonucleases. **(D)** RNA binding proteins such as Ago2 (shown), GW182 and Translin associate with and protect miRNAs from exoribonucleolytic decay through making the 3' and 5' ends of the miRNA inaccessible to exoribonucleases.

## 1.4 Dis3

### 1.4.1 Dis3 structure and conservation

As mentioned in previous sections Dis3 is a 3'-5' exoribonuclease which associates with the exosome core to provide its catalytic activity (Mitchell *et al.* 1997; Liu *et al.* 2006; Wasmuth and Lima 2012; Makino *et al.* 2013). The exosome core essentially functions to 'tame the beast' that is Dis3 and provide its specificity which otherwise would degrade anything it encountered in the cell with high processivity (Liu *et al.* 2006; Wasmuth and Lima 2012). Dis3 itself is a member of a highly conserved family of 3'-5' exoribonucleases homologous to the bacterial RNase II/R family (Zuo and Deutscher 2001). Its domain structure is highly conserved (Figure 1.3) with the only difference in eukaryotes being the addition of an N-terminal PIN (PiIT-N-terminal) domain together with a cysteine repeat (CR3) motif. The N-terminal PIN domain has been shown to provide Dis3 with endoribonucleolytic activity for which the amino acids D171 and D198 are essential (Lebreton *et al.* 2008). In addition to carrying endoribonucleolytic activity the PIN domain, along with the CR3 motif, is required for the interaction between Dis3 and the exosome (Mamolen *et al.* 2010; Schaeffer *et al.* 2012; Makino *et al.* 2013).

Together with endoribonuclease activity, Dis3 also contains exoribonuclease activity which resides within the RNB domain and has been shown to demonstrate a preference for substrates with 5' phosphates (Schaeffer *et al.* 2009). The RNB domain is conserved from prokaryotes through to humans with D551 essential for catalytic activity in *S. cerevisiae* (Schneider *et al.* 2007). This catalytic activity is kept in check by the exosome core (Liu *et al.* 2006; Wasmuth and Lima 2012) which binds to its substrates and with the help of co-factors (see section 1.3.3.3) feeds the RNA into its RNB domain. Although most of Dis3's activity is dependent upon the unwinding of structured RNA by helicase co-factors it has been shown to be capable of unwinding secondary structures alone. As long as the substrate contains 4-5nt of unstructured 3' sequence, Dis3 is able to use energy harvested from previous hydrolytic steps to unwind 4nt in a single burst (Lee *et al.* 2012). Although interesting, this activity may be largely redundant as the structured RNA would require unwinding before entry into the exosome central channel.

At the base of the exosome core the RNB domain of Dis3 degrades the RNA in a highly processive, hydrolytic manner releasing a single nucleotide at a time. To reach the catalytic site of Dis3 through the exosome central channel a single stranded RNA must be at least 7nt in length (Lorentzen *et al.* 2008). Dis3 also contains two cold-shock domains responsible which provide non-specific RNA binding. The conservation of Dis3 is shown at the peptide level as

human Dis3 is 55% identical to *Drosophila* Dis3 with the greatest homology within the catalytic RNB domain. This functional homology was demonstrated by classical complementation experiments where human Dis3 was able to complement for the loss of Dis3 in yeast (Mitchell *et al.* 1997; Allmang *et al.* 1999).

The cellular localisation of Dis3 varies between organisms. In *Drosophila melanogaster* Dis3 associates with both the cytoplasmic and nuclear exosome (Graham *et al.* 2006) whilst in humans its localisation is predominantly nuclear (Tomecki *et al.* 2010). This differential localisation is due to the presence of an additional Dis3 paralogue encoded within the human genome, namely Dis3L1. In human cells Dis3L1 has been shown to be localised to the cytoplasm and to associate with the cytoplasmic exosome (Staals *et al.* 2010; Tomecki *et al.* 2010). It is likely that in *Drosophila* Dis3 performs the role Dis3L1 is responsible for in human cells.

## **1.4.2 Cellular functions of Dis3 and phenotypes of Dis3 mutants.**

### **1.4.2.1 Cell cycle regulation**

The first identified biological function of Dis3 was in the regulation of sister chromatid separation which led to its name; defective in sister chromatid rejoining 3 (Ohkura *et al.* 1988). This implies that Dis3 has a function in regulating a key process within the cell cycle. Consistent with this founding study, subsequent work has shown that the loss of Dis3 in a variety of organisms results in the loss of cell cycle control and growth deficiencies (Figure 1.7). For example, *dis3* mutants in *S. pombe* show mitotic arrest (Kinoshita *et al.* 1991; Murakami *et al.* 2007). These mutants, which show reduced exoribonuclease activity, present with elongated metaphase-like spindles and are arrested in mitosis in a Mad2-dependent checkpoint arrest. The requirement of the exoribonuclease activity of Dis3 for spindle assembly was corroborated by Smith *et al* where they showed that *S. cerevisiae* with *dis3* mutations accumulate pre/anaphase spindles which are also mislocalised (Smith *et al.* 2011). Additional work in yeast has shown that loss of the endo- or exoribonuclease activity of Dis3 results in slow growth phenotypes which are likely to occur as a result of these mitotic defects (Schaeffer *et al.* 2011; Smith *et al.* 2011; Schaeffer *et al.* 2012; Reis *et al.* 2013).

Similar roles for Dis3 in the cell cycle and growth have also been shown in multicellular organisms. Loss of Dis3 in *Drosophila melanogaster*, using RNA interference, also results in profound growth deficiencies and ultimate larval lethality (Hou *et al.* 2012). Loss of Dis3 in these flies resulted in early larval size to reduce by 400%, a severe growth deficiency which

ultimately resulted in lethality at the 2<sup>nd</sup> instar larval stage (L2). Complementary work in *Drosophila* embryo derived S2 cells showed that knockdown of Dis3 resulted in a 50% reduction in proliferation (Kiss and Andrusis 2010). Finally, although siRNA directed Dis3 depletion in human cells showed no effect on growth rate, specific mutations (D387N or R780K) in HEK293 cells cause a reduced rate of proliferation (Tomecki *et al.* 2014). Taken together these corroborating studies indicate a key requirement for Dis3 to maintain correct growth in terms of both single cellular and multicellular models.

#### **1.4.2.2 Dis3 in RNA processing and degradation**

The above data show the phenotypic effects of the loss of Dis3 to result in mitotic deficiencies. These have largely been shown to result from a loss of the ribonuclease activity of Dis3 rather than being a more structural role. This implies, therefore, that Dis3 is responsible for the control over specific transcripts which in turn feedback to regulate spindle assembly and ultimately mitotic progression. A number of studies have been performed in an attempt to identify specific transcripts that Dis3 is involved in the processing of which have unveiled a variety of substrates.

Dis3 has been implicated in the processing of a number of RNA species at various stages of their biogenesis. For example, the first Dis3 substrate to be identified was a 5.8S rRNA precursor. Dis3 was shown to be required for trimming of the 7S rRNA which is required for the production of the 5.8S rRNA (Mitchell *et al.* 1997; Schneider *et al.* 2012). Interestingly, using cross-linking experiments Schneider *et al.* have shown that the trimming activity of Dis3 appears to be independent of exosome interaction. This puts forward the hypothesis that the trimming events are achieved through direct access to the active site rather than through the central channel which may facilitate complete degradation. It is possible that in accessing the site directly the RNA structure is not unwound and therefore Dis3 stalls when it reaches more structured regions causing the trimmed RNA to dissociate. In the absence of redundancy deficient ribosomal processing alone could be sufficient to slow cell growth as ribosome function is critical for a cell to grow.

In addition to the processing of rRNAs, Dis3 has been implicated in the processing of pre-tRNAs (Kadowaki 1994; Kadaba *et al.* 2004; Gudipati *et al.* 2012; Schneider *et al.* 2012). The Dis3/exosome complex has been identified as being responsible for the specific removal of hypomodified tRNAs in *S. cerevisiae* (Kadaba *et al.* 2004; Schneider *et al.* 2007); these hypomodified tRNAs do not undergo the essential step of methylation which is required for their correct folding. These aberrant pre-tRNAs are signalled for degradation by polyadenylation by the TRAMP complex (Kadaba *et al.* 2004). In a similar mechanism

structurally unstable tRNAs have been shown to be marked with a 3' CCACCA motif which promotes their clearance by the Dis3/exosome complex (Wilusz *et al.* 2011). In addition to the quality control over pre-tRNAs, it has been shown that mature tRNAs also accumulate in *S. cerevisiae* cells mutant for *dis3* indicating a more widespread role for Dis3 in tRNA regulation (Gudipati *et al.* 2012). This study revealed that nearly half of all transcribed tRNAs do not make it into the cytoplasm.

Together with a quality control role for Dis3 in tRNA biogenesis a similar role has been reported for pre-mRNAs (Bousquet-Antonelli *et al.* 2000; Andrulis *et al.* 2002; Lemieux *et al.* 2011; Gudipati *et al.* 2012; Schneider *et al.* 2012). The exosome has been shown to be involved in the removal of mRNAs with retained introns to ensure incorrectly spliced mRNAs do not reach the cytoplasm (Gudipati *et al.* 2012; Schneider *et al.* 2012). Consistent with a role in pre-mRNA removal/processing *dis3* mutants in both *S. pombe* and *S. cerevisiae* show a modest accumulation of mRNAs and pre-mRNAs (Bousquet-Antonelli *et al.* 2000; Lemieux *et al.* 2011). This function, however, was not specific for Dis3 as *S. pombe rrp6* mutants also shown accumulation mRNAs and pre-mRNAs (Lemieux *et al.* 2011). It has been proposed that the quality control machines (i.e the exosome) and the splicing machinery compete for their pre-mRNA substrates (Bousquet-Antonelli *et al.* 2000). In line with this, direct interactions have been observed between the splicing and decay machinery (Nag and Steitz 2012; Zhou *et al.* 2015). Interestingly the role of the Dis3/exosome complex in processing pre-mRNAs appears to occur in a co-transcriptional manner as is the case for pre-mRNA splicing (Lemay *et al.* 2014). For example, the exosome co-purifies with the RNA pol II elongation complex indicating it is present at the site of transcription in *Drosophila* (Andrulis *et al.* 2002). This could be a mechanism of increasing the efficiency of the pre-mRNA surveillance machinery.

A final role for Dis3 is in the regulation of miRNA biogenesis. Both Dis3 and Rrp6 have been shown to have preference for the removal of a distinct subset of defective pre-miRNAs in the nucleus (Liu *et al.* 2014). These defective or unwanted pre-miRNAs are specifically targeted to the exosome through uridylation by either TUT4 or TUT7 which allows target discrimination between the healthy and defective precursors. The specificity of Dis3 or Rrp6 was inferred to be dependent upon the 3' overhang with Dis3 showing a preference for pre-miRNAs with >2nt overhang whilst Rrp6 favoured <2nt overhangs (Liu *et al.* 2014). This could be consistent with previous data showing that Dis3 requires 4nt 3' overhangs to process more structured subunits (Lee *et al.* 2012). Together with pre-miRNA clearance, Dis3 has also been implicated in the trimming and subsequent maturation of pre-miRNAs. An example of this was presented by Xue *et al.* who showed a role for the exosome/Dis3 in the biogenesis of miRNA-like RNAs in *Neurospora crassa* (Xue *et al.* 2012). Complementary work has shown a similar role for the

exosome in the biogenesis of miRNAs derived from introns in *Drosophila* (Flynt *et al.* 2010). This may be regulated in a similar mechanism to the regulation of pre-mRNAs discussed previously.

In addition to roles in the regulation of RNA precursors, Dis3 has also been implicated in the regulation of RNAs which are at their functional level. For example, a global study by Schneider *et al* revealed that *dis3* mutants in *S. cerevisiae* presented with an accumulation of RNA pol III transcripts such as cryptic unstable transcripts (CUTs), snRNAs and snoRNAs (Schneider *et al.* 2012). Consistent with a role in regulating snoRNAs, the exosome had previously been implicated in the 3' trimming of snoRNAs during their biogenesis (Kim *et al.* 2006). CUTs are unstable sequences of RNA which are produced from introns after they have been spliced out during mRNA maturation. Similar sequences are known in higher organism such as promoter upstream transcripts (PROMPTs) in humans. These PROMPTs have also been shown to be targeted by the exosome complex with both Dis3 and Rps6 responsible for their decay (Preker *et al.* 2008). Consistent with this, specific mutations in Dis3 in human cells results in the stabilisation of a number of PROMPTs (Tomecki *et al.* 2014).

Dis3 and the exosome have also been implicated in the turnover of mRNAs in the cytoplasm in a function distinct to that in the nucleus (van Hoof *et al.* 2000; Schaeffer *et al.* 2009). For example, the Dis3/exosome complex has been shown to target mRNAs containing AU rich elements (AREs) in the 3' UTR to control the expression of specific and in some cases oncogenic transcripts (Chen *et al.* 2001; Schoenberg and Maquat 2012). Consistent with the UTRs playing a role in exosome targeting, transcripts affected by *dis3* knockdown in *Drosophila* S2 cells show longer than average UTRs (Kiss and Andrulis 2010). The exosome has also been shown to degrade meiotic mRNAs during mitotic growth in the fission yeast *S. pombe*. These meiotic mRNAs contain specific destabilising motifs known as 'determinants of selective removal' providing another example by which the exosome can be targeted to specific substrates (Harigaya *et al.* 2006). This work has been shown in yeast, which do not contain the extra member of the Dis3 family, Dis3L1. It is quite possible that in higher eukaryotes Dis3L1 takes over the role of regulating mRNA stability in the cytoplasm.

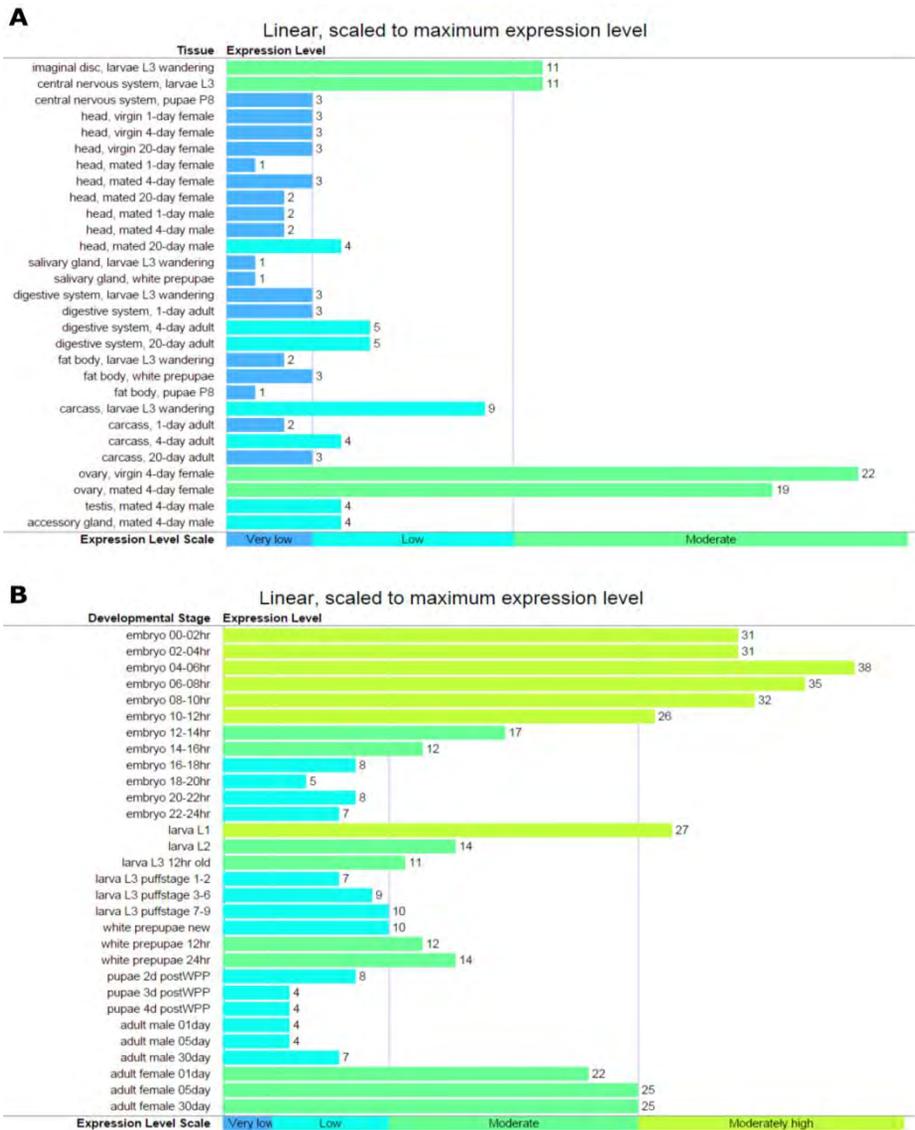
Phenotype	Method	Organism	Reference(s)
Defective in sister chromatid separation	Mutant (P509L)	<i>S. pombe</i>	Ohkura <i>et al</i> 1988
Mitotic arrest	Mutant (P509L)	<i>S. pombe</i>	Murakhami <i>et al</i> 2007 Kinoshita <i>et al</i> 1991
Accumulation of mislocalised anaphase spindles	Mutant ( <i>mtr17-1</i> )	<i>S. cerevisiae</i>	Smith <i>et al</i> 2011 Kadowaki <i>et al</i> 1994
Slow growth phenotypes	Mutant (D171A , D551N, G883R and R847K, Y595A)	<i>S. cerevisiae</i>	Schaeffer <i>et al</i> 2011 Schaeffer <i>et al</i> 2012 Tomecki <i>et al</i> 2014 Reis <i>et al</i> 2013
Slow growth and larval lethality	RNA interference	<i>D. melanogaster</i> (whole organism)	Hou and Andrusis 2012
Reduced proliferation	RNA interference	<i>D. melanogaster</i> (S2 cells)	Kiss and Andrusis 2010
Reduced proliferation	Mutant(D487N and R780K)	<i>H. sapiens</i> (HEK293 cells)	Tomecki <i>et al</i> 2014
Aberent class switch recombination	RNA interference	<i>H. sapiens</i> (B cell lymphoma and HEK293 cells)	Basu <i>et al</i> 2011

**Figure 1.7: Summary of phenotypes observed in Dis3 loss of function studies.**

### **1.4.3 Expression pattern of Dis3 in *Drosophila melanogaster*.**

The level of transcript expression for every coding RNA in *Drosophila* is available due to RNA sequencing experiments performed by the modENCODE project (Graveley *et al.* 2011). Although this database does not show protein, it allows the identification of the expression level of an RNA of interest at each developmental stage and within each tissue. *dis3* mRNA is detected in every tissue which would be consistent with its critical role in maintaining cell viability (Figure 1.8A). It is expressed at highest levels in the ovary which would be consistent with its maternal contribution. It is also expressed at moderate levels in the imaginal discs and the CNS with its lowest level of expression in the pupal fat body.

In terms of the expression of *dis3* at the RNA level during development it is at its highest level during early embryogenesis which again is consistent with its maternal contribution. Its expression is shown to drop off during the late stages of embryogenesis before a peaking during the 1<sup>st</sup> instar larval stage. During the latter developmental stages *dis3* mRNA is still expressed but at moderate to low levels (Figure 1.8B). This pattern of expression is confirmed at the RNA and protein level by a study by Cairrao *et al* who showed that Dis3 protein and RNA is at its highest level during embryogenesis with L1 showing the greatest expression of all the larval stages (Cairrão *et al.* 2005). The ubiquitous expression of Dis3 is consistent with its crucial role in RNA processing and degradation.



**Figure 1.8: The expression profile of *dis3* mRNA in *Drosophila melanogaster*.** (A) The expression of *dis3* mRNA in various larval and adult tissues as determined through RNA-sequencing by the modENCODE project. The values represent the normalised mean level of expression for each tissue. (B) The expression of *dis3* mRNA throughout the developmental life cycle as determined through RNA sequencing by the modENCODE project. The values represent the normalised mean level of expression for each stage. Images taken from Flybase.

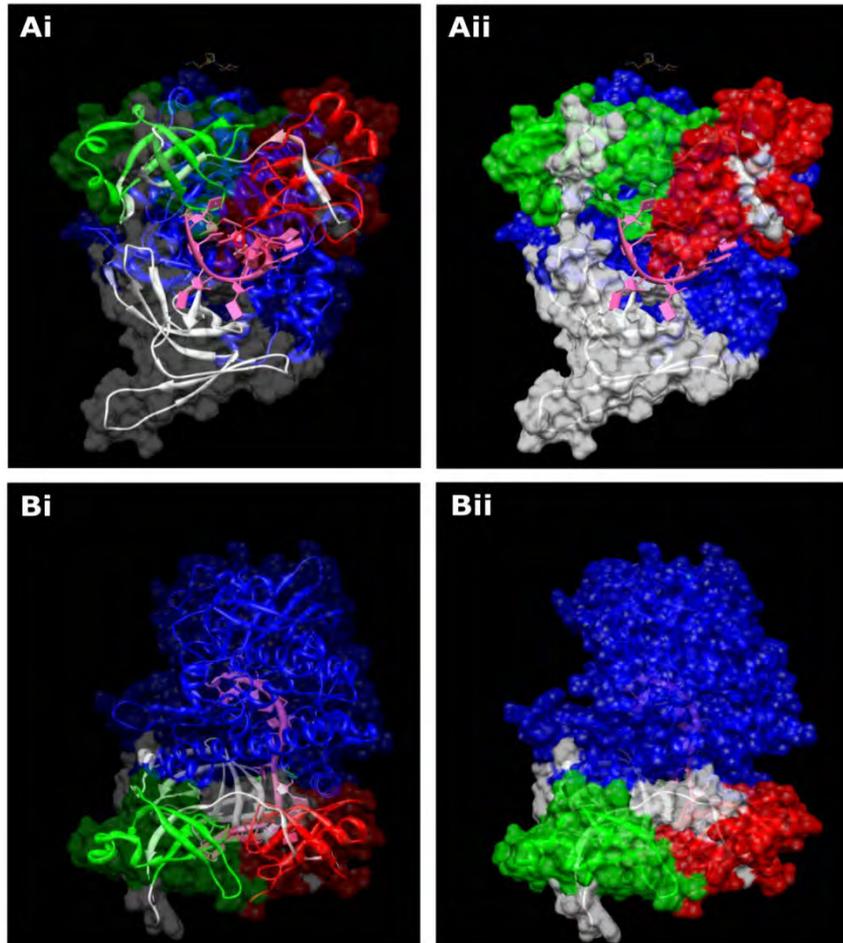
## 1.5 Dis3L2

### 1.5.1 Dis3L2 structure and conservation.

Dis3L2 is another member of the highly conserved RNase II/R family of 3'-5' exoribonucleases. Like Dis3 it is a highly processive exoribonuclease that degrades a variety of RNA substrates with high efficiency (Lubas *et al.* 2013; Malecki *et al.* 2013). It contains very similar domain organisation to its paralogues, however, it is more similar to the bacterial homologue in that it does not contain the N-terminal PIN domain found in Dis3 and Dis3L1 (Figure 1.3). The PIN domain provides endoribonuclease activity together with being critical for exosome interaction (Lebreton *et al.* 2008; Mamolen *et al.* 2010; Schaeffer *et al.* 2012; Makino *et al.* 2013). This infers that Dis3L2 does not contain endoribonuclease activity together with acting independently of the exosome; indeed this has been confirmed in subsequent studies (Lubas *et al.* 2013; Malecki *et al.* 2013). The exoribonuclease domain (RNB) and the three RNA binding domains are, however, present in Dis3L2.

Dis3L2 has only recently been characterised with this work restricted to *S. pombe*, humans and plants (where it is called SOV) where it has been shown to be localised to the cytoplasm (Zhang *et al.* 2010; Lubas *et al.* 2013; Malecki *et al.* 2013). Dis3L2 is found in most eukaryotes, although it is not present in *S. cerevisiae* where much of the founding work on RNA decay has been performed and therefore until recently has eluded attention. Arabidopsis Dis3L2 (SOV) is unable to complement the loss of Dis3 in *S. cerevisiae*, whilst Arabidopsis Dis3 (AtRRp44A) is able to showing how the two proteins are functionally different (Kumakura *et al.* 2013). This is not surprising however, as mentioned previously Dis3L2 is unable to interact with the exosome core.

The structure of Dis3L2, bound to RNA has been solved for mouse Dis3L2. This elegant work has shown that the RNA substrate is bound by the two N-terminal cold-shock domains and is then routed up through the RNB domain into the catalytic site of Dis3L2 (Faehnle *et al.* 2014) (Figure 1.9). Once in the active site the RNA is hydrolysed in a high processive manner. Dis3L2 has been implicated in the degradation of both coding and non-coding RNAs and appears to use uridylation as a targeting mechanism (Chang *et al.* 2013; Malecki *et al.* 2013; Ustianenko *et al.* 2013). This will be discussed in greater detail below.



**Figure 1.9: Structure of Dis3L2 and an RNA substrate. (A)** A base up view of the protein structure of mouse Dis3L2 associated with a poly(U) RNA substrate (pink). The RNA is bound by the two cold-shock domains (red and green) at the base of the structure and fed into the catalytically active RNB domain (blue) where it is hydrolysed in a highly processive manner. **(B)** As in **A** but a side view. Adapted from Faehnle *et al* 2014.

### 1.5.2 Cellular functions of Dis3L2 and phenotypes of Dis3L2 mutants.

Due to its recent discovery Dis3L2 has attracted the attention of many RNA-decay labs around the world. However, unlike Dis3, little is known about its specific targets and if it plays a regulatory role in specific cellular pathways and processes. However, it has been shown that its role in cytoplasmic RNA decay is critical as *S. pombe* double mutant for *xrn1* and *dis3L2* are not viable whilst *lsm1/dis3L2* mutants show a strong inhibition of RNA decay greater than when either is mutated alone (Malecki *et al.* 2013). Consistent with a role in cytoplasmic RNA decay human Dis3L2 has been shown to degrade ARE-containing reporter constructs (Lubas *et al.* 2013) placing it in a similar role of cytoplasmic decay as XRN1 and the exosome.

Interestingly, an RNA-dependent interaction was identified between XRN1 and Dis3L2 which suggests that there may be a level of target redundancy together which indicating that a transcript can be degraded in both directions simultaneously (Lubas *et al.* 2013). Like XRN1, the loss of Dis3L2 in *S. pombe* cells has been shown to affect P bodies; although this results in an increase in P body number (Malecki *et al.* 2013) rather than in their size as was shown following *xrn1* mutations in *Drosophila* (Zabolotskaya *et al.* 2008). This may suggest that in the absence of Dis3L2, at least some of its targets are relocated to additional P. bodies to be degraded by XRN1.

This recent work, together with others, have also identified a potential mechanism through which Dis3L2 is recruited to its targets. For example, *dis3L2* mutants in *S. pombe* and humans have been shown to accumulate transcripts with uridine tails (Malecki *et al.* 2013; Thomas *et al.* 2015). Consistent with this, two studies have shown that during its biogenesis *pre-let-7* is polyuridylylated by TUT4 and TUT7 which results in Dis3L2 recruitment and subsequent degradation in human and mouse cells (Chang *et al.* 2013; Ustianenko *et al.* 2013). This suggests that uridylation may be a conserved mechanism used to direct specific transcripts for Dis3L2-mediated degradation. This mechanism would be similar to one used in regulating histone RNA degradation (Mullen and Marzluff 2008).

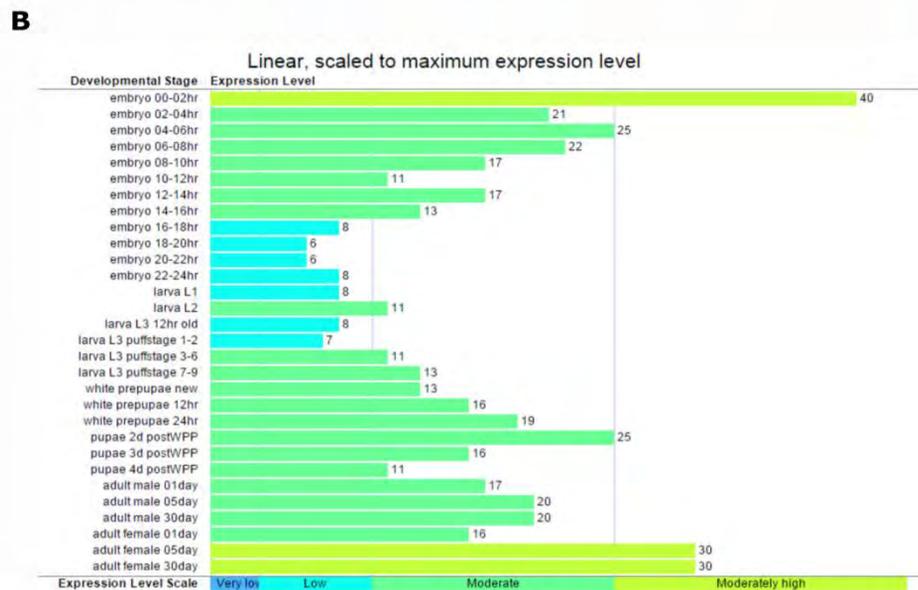
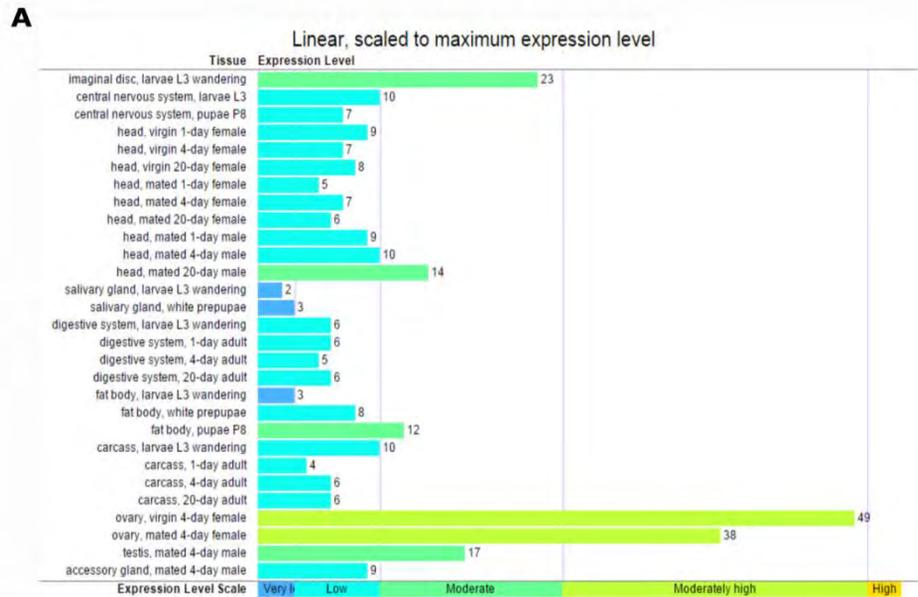
The specific cellular processes Dis3L2 is involved in regulating are yet to be resolved; however, there may be some clues within the literature. For example, it has been shown that Dis3L2 is required for efficient clearance of mRNAs following apoptotic stimuli in human cells (Thomas *et al.* 2015). Although this role does not stimulate apoptosis, it has been shown to be required for efficient cell death with *DIS3L2* knockdown resulting in reduced cell death. In accordance with previous findings, the mRNA transcripts that accumulate in the absence of Dis3L2 present with uridine-rich tails which are TUT4/TUT7 dependent.

Finally, a role for Dis3L2 in maintaining tissue growth has been suggested by the fact that an overgrowth syndrome, named Perlman syndrome, has been shown to result from *DIS3L2* mutations (discussed in greater detail in later section) (Astuti *et al.* 2012). This could be in line with the findings above, however, at yet no work has been performed in whole organisms to investigate the role for Dis3L2 in tissue growth. This was one of the mains objectives of this thesis.

### **1.5.3 Expression pattern of Dis3L2 in *Drosophila melanogaster***

As mentioned previously, the modENCODE data allows analysis of the mRNA expression pattern of *dis3L2* throughout *Drosophila melanogaster*. The expression pattern for *dis3L2* mRNA is similar to that of *dis3* mRNA in that its highest expression is observed in the ovary followed by the imaginal discs (Figure 1.10A). It is likely that *dis3L2* is also maternally contributed which again would be consistent with high expression in the ovary. Overall *dis3L2* mRNA is observed in all examined tissues at low to moderate levels. Similar ubiquitous expression is observed throughout *Drosophila melanogaster* development with highest levels seen during the first 2 hours of embryogenesis. Unlike *dis3*, *dis3L2* mRNA expression is relatively consistent from mid embryogenesis through to adult flies (Figure 1.10B).

As yet an antibody is not available to *Drosophila* Dis3L2 and therefore experiments like those performed by Cairrao *et al* to assess the protein expression are not possible.



**Figure 1.10: The expression profile of *dis3L2* mRNA in *Drosophila melanogaster*.** (A) The expression of *dis3L2* mRNA in various larval and adult tissues as determined through RNA-sequencing by the modENCODE project. The values represent the normalised mean level of expression for each tissue. (B) The expression of *dis3L2* mRNA throughout the developmental life cycle as determined through RNA sequencing by the modENCODE project. The values represent the normalised mean level of expression for each stage. Images taken from Flybase.

## 1.6 RNA stability in disease

The importance of maintaining control over RNA stability is shown through that fact that a number of crucial factors involved in RNA stability have been shown to be mutated, depleted or overexpressed in a number of human diseases. The diseases associated with the major players of RNA decay are discussed below.

### 1.6.1 Dis3

As mentioned previously, Dis3 is a 3'-5' exoribonuclease which is required to maintain cell cycle progression and ultimately organism viability (Murakami *et al.* 2007; Smith *et al.* 2011; Hou *et al.* 2012). Interestingly, Dis3 mutations have been identified in 11% of multiple myeloma (MM) cases (Chapman *et al.* 2011). The most frequent mutations identified have been shown to affect the catalytic activity of Dis3 suggesting a deficiency in its exoribonuclease in these conditions may contribute to the progression of MM (Tomecki *et al.* 2014). Similarly, *DIS3* mutations have been observed in other blood cancers such as acute myeloid leukaemia (AML), at 4% frequency (Ding *et al.* 2012), and chronic lymphocytic leukaemia (CLL) (Ng *et al.* 2007). This poses an intriguing question; if the loss of Dis3 results in reduced cell growth (Murakami *et al.* 2007; Schaeffer *et al.* 2011; Smith *et al.* 2011; Hou *et al.* 2012; Schaeffer *et al.* 2012) then how do loss of function (lof) mutations (Tomecki *et al.* 2014) result in advanced cell survival leading to blood cancers? It is more than likely that the *DIS3* mutations occur alongside a number of other mutations that may, when in collaboration, enhance cell survival. For instance, if a key oncogene becomes overexpressed, such as Ras then the lof *DIS3* mutations may aid tumour survival. Interestingly, Ras has also been shown to be commonly overexpressed in multiple myeloma (Steinbrunn *et al.* 2011). Alternatively, these tumours may have mutations in tumour suppressors such as p53 thus reducing their ability to enter apoptosis. In this case it is possible that the mass misregulation resulting from *DIS3* mutation results in the increased abundance of transcripts which may function to drive tumourigenesis. The frequency of mutation, especially in the case of myeloma, would suggest a strong role for Dis3 in tumour progression. This is an interesting possibility and one that would require further work.

In addition to the blood cancers, *DIS3* overexpression has been observed primary tumour samples. For instance, in primary and metastatic colorectal tumours *DIS3* expression was shown to be up 38-fold higher than control samples (Lim *et al.* 1997; Liang *et al.* 2007). Similar findings show overexpression of *DIS3* in colorectal carcinomas (de Groen *et al.* 2014). Finally, overexpression of *DIS3* was also observed in nodular melanoma cells (Rose *et al.* 2011). The differing expression profiles indicate that *DIS3* is neither a classical oncogene nor a tumour

suppressor and therefore it is likely that it is dependent on the surrounding conditions and the other mutations within the developing tumour. However, further work would be required to identify if DIS3 is recurrently mutated or overexpressed alongside other mutations to begin to piece together an interaction network.

### **1.6.2 Dis3L2**

Dis3L2, like Dis3, has also been implicated in human disease. For example, its mutation has been implicated in the human overgrowth condition Perlman syndrome (Astuti *et al.* 2012; Morris *et al.* 2013). Perlman syndrome is a rare, autosomal recessive condition which is characterised by fetal gigantism, visceromegaly, macrocephaly and neurological deficits (Schilke *et al.* 2000; Neri *et al.* 2013). In all the examined Perlman syndrome samples homozygous deletions of either exon 6, 9 or 19 were identified which were shown to result in an absence of Dis3L2 exoribonuclease activity (Astuti *et al.* 2012). A role for Dis3L2 in controlling the cell cycle was hypothesised as the knockdown of Dis3L2 in HeLa cells was shown to result in mitotic errors and affect the expression of key cell cycle proteins including Cyclin D and Aurora B kinase (Astuti *et al.* 2012).

Interestingly, Perlman syndrome patients show a high predisposition to Wilms tumour, a kidney tumour, with pancreatic hyperplasia also common (Astuti *et al.* 2012; Morris *et al.* 2013). In line with this, 30% (6 of 20) sporadic Wilms tumour samples were identified to contain either partial or complete deletion of Dis3L2 (Astuti *et al.* 2012). This, taken with the Perlman syndrome phenotypes suggests that Dis3L2 may play a role in controlling cell and ultimately tissue growth. This exciting possibility created a major objective of this thesis; to investigate the role of Dis3L2 in tissue growth using *Drosophila melanogaster* as a model system as all previous work has focused on single cells which have little/no developmental potential.

### **1.6.3 XRN1**

Mutations in the 5'-3' pathway have also been implicated in disease. For example, XRN1 expression has been shown to be lower at the RNA level in osteosarcoma cell lines and 6 out of 9 primary tumours (Zhang *et al.* 2002). The importance of XRN1 in maintaining cellular homeostasis is further shown by Chapman *et al.* where a specific family of viruses (flaviviruses) have evolved mechanisms to develop XRN1 resistance allowing their pathogenesis (Chapman *et al.* 2014; Moon *et al.* 2015).

## 1.6.4 miRNAs in human disease

Together with the major RNA decay enzymes themselves other factors involved in regulating the stability of an RNA have also been implicated in disease. For example, misexpression of miRNAs have been identified in cardiovascular disease (Thum *et al.* 2008), arthritis (Stanczyk *et al.* 2008), neurodegeneration (Lukiw 2007) and a number of cancers (Jansson and Lund 2012). The first example in cancer was the frequent deletion of *miR-15* and *miR-16* in CLL which results in increased expression of the anti-apoptotic protein BCL2 (Calin *et al.* 2002). The most famous miRNA, *let-7*, is also known to have tumour suppressor function (Johnson *et al.* 2005; Esquela-Kerscher *et al.* 2008). Finally, tumours have been shown to evade regulation by carrying mutations in the 3' UTR of specific mRNAs resulting in changes in a miRNA target site and consequent insensitivity to miRNA regulation (Ziebarth *et al.* 2012).

In addition to their potential involvement in the progression of disease, miRNAs have also been shown to be effective biomarkers in specific situations. For instance, Jones *et al.* showed that *miR-720*, *miR-1308* and *miR-1246* could be used to differentiate patients with multiple myeloma from those with the pre-cancerous MGUS and health controls (Jones *et al.* 2012). The role of miRNAs in human disease and particularly cancer is perhaps most strongly shown through the fact that miRNAs mimics or inhibitors are being developed for use as therapeutics. For example, the first miRNA mimic (*miR-34*) was used in clinics for cancer therapy in 2013 (Bouchie 2013); with the emerging importance of miRNA-mediated regulation in cancer biology it is highly likely that many more will follow.

## 1.7 *Drosophila* as a model organism to study RNA stability

### 1.7.1 Why use *Drosophila*?

Much of the founding research into how RNA stability is controlled and regulated has been conducted in single cells; either in yeast or in tissue culture cells. Neither of these conditions, although useful in understanding specific mechanisms, allow for the full understanding of the role of RNA stability in terms of development and disease. For this multicellular organisms with full developmental potential are critical. *Drosophila melanogaster* was first introduced for use in lab science by Thomas Hunt Morgan in 1909 and since then has contributed enormously towards our current understanding of developmental and cellular biology.

*Drosophila melanogaster* is an excellent model system in which to study fundamental pathways and processes within the cell in a developmental context. This is because a number of the complex machinery involved in these pathways are highly conserved. Together with this approximately 75% of all disease contributing genes in humans have a *Drosophila* homologue

(Reiter *et al.* 2001) and therefore findings in *Drosophila* are often directly relevant to humans and human disease. This is indeed the case in the context of RNA stability. The key players in RNA degradation and stability are highly conserved (Newbury 2006; Garneau *et al.* 2007), some even from prokaryotes (Kaberdin *et al.* 2011) therefore *Drosophila* provides an ideal model system for their study.

Along with being relatively cheap and easy to keep using *Drosophila* as a model system has a number of more technical advantages. The first of these is the availability of a wide variety of genetic tools such as the *GAL4-UAS* system and balancer chromosomes. The *Drosophila* genome is also more streamlined with only 4 pairs of chromosomes, of which only 3 are generally used, meaning genetic screening is more straightforward. Moreover there is less pathway redundancy within *Drosophila* and therefore dissecting individual pathways can be more straightforward. Balancer chromosomes are a key element in a *Drosophila* geneticists 'swiss army knife'. They allow for genetic tractability due to carrying phenotypic markers which are easily identifiable. Additionally, they also ensure chromosome integrity as their nature prevents chromosomal crossover which if occurred could result in the loss of a mutation. In line with this they allow the keeping of stocks carrying mutations which would otherwise be homozygous lethal.

Together with balancer chromosomes the *GAL4-UAS* system is an elegant system through which the expression of a specific construct can be restricted to almost any region of the fly (Brand and Perrimon 1993; Duffy 2002). This system involves the use of a yeast transcription factor, GAL4, and a series of GAL4 specific binding sites (UAS). The specificity of the system is provided by the promoter through which the expression of the GAL4 gene is controlled. For instance, a *tubulin* promoter would result in the ubiquitous expression of GAL4 whilst a wing specific promoter, such as *nubbin*, would provide an element of specificity. The expression of GAL4, which in turn drives the expression of a desired construct under the control of the UAS, can be restricted to just a few cells.

Another benefit of using *Drosophila melanogaster* is that they produce a high number of offspring in a short generation time. Generating high numbers of offspring allows the collection of sufficient samples for stringent statistical analysis. Finally, and importantly, the entire genome has been sequenced. Work in *Drosophila* has played a large contribution to our current understanding of a number of key cellular and developmental processes. For example, the mechanism by which the decapping enzymes and XRN1 are recruited and interact was first reported in *Drosophila* (Braun *et al.* 2012). Together with this, the relative ease of mutagenesis in *Drosophila* allowed the identification of a number of proteins involved controlling

proliferation which has aided further understanding (Bryant and Schmidt 1990; Baonza and Freeman 2005).

### **1.7.2 *Drosophila* life cycle and development**

The *Drosophila* life cycle has been extensively characterised with the time it takes for a fertilised egg to develop into an adult heavily temperature dependent. For example, at 25°C this process takes 10 days, however, decreasing the temperature to 19°C the development will take closer to 18 days. Similarly, increasing the temperature to 28°C accelerates the developmental time to just 7 days (Ashburner 1989). The temperature the *Drosophila* are cultured under not only affects their developmental time but also their size. *Drosophila* grown at 19°C are much larger than those growth at 28°C for example. The outline of the *Drosophila* life cycle is discussed below and all timing at for flies cultured at 25°C.

Upon fertilisation embryogenesis takes only 24 hours at 25°C. During embryogenesis the critical developmental axes are determined by the deposition of maternal transcripts in specific regions of the oocyte. The anterior-posterior boundary is governed by the expression of 4 genes; Bicoid and Hunchback in the anterior and Nanos and Caudal in the posterior of the embryo. Bicoid and Nanos are two morphogens which set up a concentration gradient from anterior to posterior and posterior to anterior respectively. This in turn allows the regulation of *hunchback mRNA* which is repressed by Nanos in the posterior whilst expressed in the anterior region of the embryo.

These boundaries are set up using maternally provided transcripts; however, further axis specification is achieved following the mid-blastula transition which occurs at cycle 14. This is the stage at which the control of gene expression is 'handed over' to the zygote and zygotic transcription commences. It has been recently shown that early zygotic gene expression is largely regulated by the transcription factor Zelda (Harrison *et al.* 2011).

After 24 hours of embryogenesis the 1<sup>st</sup> instar larvae (L1) hatch out and remain on the surface of the food before their first moult. This occurs after a further 24 hours where they become 2<sup>nd</sup> instar larvae (L2) and begin to bury into the food. The 2<sup>nd</sup> instar larval stage also lasts for 24 hours after which they undergo another moult into 3<sup>rd</sup> instar larvae (L3). L3 lasts longer (48 hours), most of which is spent buried into the food although towards the end they emerge from the food and enter a 'wandering' stage. Larvae at this stage of development were used for the RNA-sequencing experiments performed in Chapter 5. These larval stages are crucial in the development as they allow for rapid growth of the imaginal discs to a sufficient size to enter the pupal stage where they undergo metamorphosis. For example, during the larval stages the wing imaginal disc grows in size from 50 to 50,000 cells.

After 48 hours (~120 hours after egg deposition) the larvae enter the pupal stage. Prior to entering the pupal stage a 'checkpoint' has to be passed, similar to those encountered during the cell cycle, for example. This is to ensure that the imaginal discs, the tissues containing all the cells that will form the adult fly, have reached the 'critical size' required for metamorphosis (Shingleton 2010). The imaginal discs have been shown to have a 'target size' which under normal conditions they will not grow larger than (Martín and Morata 2006). The checkpoint controlling pupariation initiation is largely governed by hormones such as Ecdysone and insulin-like peptides (Shingleton 2010). Ecdysone also promotes the two larval moult stages. The activity of Ecdysone is inhibited by Juvenile Hormone (JH); however, JH secretion is inhibited when the larvae are correctly prepared for metamorphosis, allowing Ecdysone signalling and the onset of pupariation. The importance of these hormones on pupariation initiation is shown by the fact that increased expression of *Drosophila* Insulin-like peptide 8 (Dilp8) can delay the larvae entering pupariation by 2-3 days (Colombani *et al.* 2012). During the pupal stage Imaginal discs change in cell shape and movement and differentiate into the cells required for the formation of the adult. All the larval tissue, such as the fat body, is degraded through controlled cell death. After ~3.5 days adult flies eclose and are sexually receptive after 12-14 hours.

### **1.7.3 Using the wing imaginal discs as a model system**

The wing imaginal discs are of the most highly characterised tissues within *Drosophila* with pathways critical for their developmental highly characterised and well understood. For these reasons it makes them incredibly good tissues to use as a model system to investigate fundamental cellular processes. The wing imaginal disc itself is an epithelial sac of approximately 50 cells specified early in development. During the larval stages these cells proliferate profusely to produce made up of around 50,000 cells suitable to enter metamorphosis (Widmann and Dahmann 2009). The wing imaginal disc is divided into specific compartments with the cells within each compartment fated so form a specific part of the adult fly (Figure 1.11). For example, the wing pouch region of late L3 wing imaginal discs contains all the cells required to form with adult wing blade. The middle region of the disc contains the cells responsible to forming the hinge which attaches the wing blade to the notum whilst the tip (most dorsal) of the disc is fated to form part of the notum (Klein and Arias 1998).

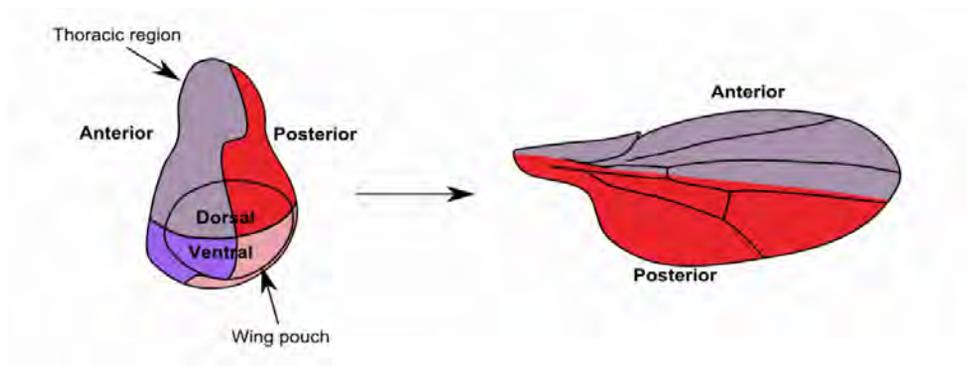
There are a number of signalling pathways that are critical for the correct development and patterning of the wing imaginal disc, most of which are highly conserved. For example, the anterior-posterior axis is determined by Dpp (BMP in humans) and Wingless (Wnt in humans) expression (Teleman 2000), whilst the dorsal-ventral axis is determined by Wingless and

Apterous (LHX2 in humans) (Diaz-Benjumea and Cohen 1993; Ng *et al.* 1996). These pathways are known to regulate similar processes during human development.

In addition to their axis patterning the control over wing disc growth is highly co-ordinated. The Hippo pathway is one signalling pathway heavily involved in the regulation of tissue growth (Neto-Silva *et al.* 2009; Halder and Johnson 2011). Loss of function mutations in kinases within the Hippo pathway results in overgrowth of the imaginal disc tissues (Harvey *et al.* 2003). Like the Wingless and Dpp pathways, Hippo is also conserved in humans and consistent with its role in *Drosophila* in restricting tissue growth is frequently misregulated in cancers (Zhao *et al.* 2010). In addition to inter/intra-cellular signalling the growth of the wing disc is also controlled by mechanical forces. Aegerter-Wilmsen *et al.* propose a model whereby the growth of the disc revolves around a compression force in the center of the disc (Aegerter-Wilmsen *et al.* 2012). In this model the growth of the disc is terminated when the compression in the center of the disc reaches a certain threshold where the growth factors can no longer overcome the repressive compression force.

As outlined above the development and patterning of the wing imaginal disc is of the best characterised for investigating growth defects. This is of immediate interest in this project as a major aim was to explore the regulatory role of Dis3L2 in tissue growth. In this regard the wing imaginal discs are ideal tissues to use as a model system as they also have a number of well established genetic tools. For example, there are a number of GAL4 drivers which are specific to the wing imaginal disc including one which specifically affects the wing fates cells (*nubbin-GAL4*).

Therefore, due to the extensive conservation of the factors responsible for growth regulation, findings in the wing imaginal disc are often applicable to higher organisms, including humans. They are well characterised, easy to work with and contain limited redundancy allowing more straightforward dissection of the mechanisms involved. This places them as an ideal model system in which to study the role of exonucleases in the growth of developmental tissues.



**Figure 1.11: The *Drosophila* wing imaginal disc.** The *Drosophila* wing imaginal disc is divided into distinct developmental axes such as the anterior-posterior and dorsal-ventral axes which are specified by the expression of specific factors. The wing blade is produced from the wing pouch region whilst the tip of the disc form part of the thorax. The dorsal side of a wing is represented with its anterior-posterior boundary.

## **1.8 An introduction to the techniques used during this study.**

### **1.8.1 Using RNA interference to investigate protein function.**

This project used a reverse genetics approach to investigate the function of two exoribonucleases, Dis3 and Dis3L2, in the control of cellular and tissue growth in a developmental system. This approach has historically been used to investigate the function of specific proteins or transcripts in the absence of an available mutant.

This involved the knockdown of the RNA encoding each of the proteins resulting in a reduction in the total level of functional protein. To achieve this knockdown the aforementioned *GAL4-UAS* system was used to drive a hairpin RNA interference (RNAi) construct of ~360bp complementary to the specific mRNA. This approach is viable in *Drosophila* as they do not contain an interferon response which would otherwise remove the dsRNA. This also neatly demonstrates another advantage of working in *Drosophila*. The hairpin RNAi constructs were controlled by an upstream activation sequence (UAS) which meant that their expression, and therefore knockdown, would only be induced by the presence the yeast GAL4 transcription factor. The expression of GAL4 was in turn controlled by tissue specific promoters which allowed the investigation to be restricted to the tissues of interest, the wing imaginal disc. In each case the knockdown was confirmed at the RNA level using quantitative RT-PCR (qRT-PCR) and in the case of Dis3 was also confirmed at the protein level by Western blotting.

### **1.8.2 RNA-sequencing vs Microarrays.**

The project used RNA-sequencing (RNA-seq) in an attempt to identify the specific RNA targets of Dis3 and Dis3L2. RNA-sequencing is a relatively new technology which allows the screening of the entire transcriptome to identify differential expression between experimental conditions. Provided sufficient sequencing depth is achieved RNA-sequencing can also be used to identify novel transcripts. Due to its unbiased nature where the entire transcriptome can be identified it is generally preferred over the dated technique of microarrays.

Before the emergence of RNA-sequencing microarrays were used to study the transcriptome, however, there are many limitations to the microarray technology. For example, the major limiting factor of a microarray is the space upon the hybridisation plate. For eukaryotic genomes it is not possible to screen the entire genome on a single plate meaning the user must select the transcripts they wish to screen. By definition this indicates how microarrays are dependent upon the current knowledge of the transcriptome. The requirement upon target selection also means that potentially interesting transcripts may be missed together with introducing an element of bias. RNA-sequencing overcomes this factor by enabling the

user to screen the entire genome in an unbiased manner. This also introduces another advantage of sequencing in that, provided enough read depth is used, it presents an opportunity to identify novel transcripts which is not possible with microarrays.

An additional major advantage of RNA-sequencing is in its sensitivity of detecting transcripts expressed at low levels. RNA-sequencing has been shown to be much more reliable in determining differential expression of these lowly expressed transcripts than microarrays (Marioni *et al.* 2008; Zhao *et al.* 2014). The issue with microarrays in detecting lowly expressed transcripts comes from high levels of background hybridisation which can also prove problematic with quantification of large fold changes. A study by Fu *et al.* also shows how RNA-sequencing better correlates with proteomic data than microarrays and is therefore more likely to lead to the identification of functional changes (Fu *et al.* 2009).

However, there are limitations to RNA-sequencing. For example, the determination of differential expression between isoforms is dependent upon an 'educated assumption' made by the alignment algorithms based upon reads that map across exon-exon junctions. Similarly, for highly repetitive genomes, such as the human genome which encodes multiple isoforms it is much more difficult to accurately determine the specific origin of reads (Rehrauer *et al.* 2013). As a result it is still important to validate expression changes observed in an RNA-seq experiment using an alternative approach such as qRT-PCR.

Therefore, although there are still limitations to the bioinformatical analysis and factors that need ironing out; RNA-seq still provides a more thorough and modern method to analyse changes in gene expression than the relatively dated method of microarrays.

### **1.8.3 Quantitative RT-PCR on mRNA and miRNA.**

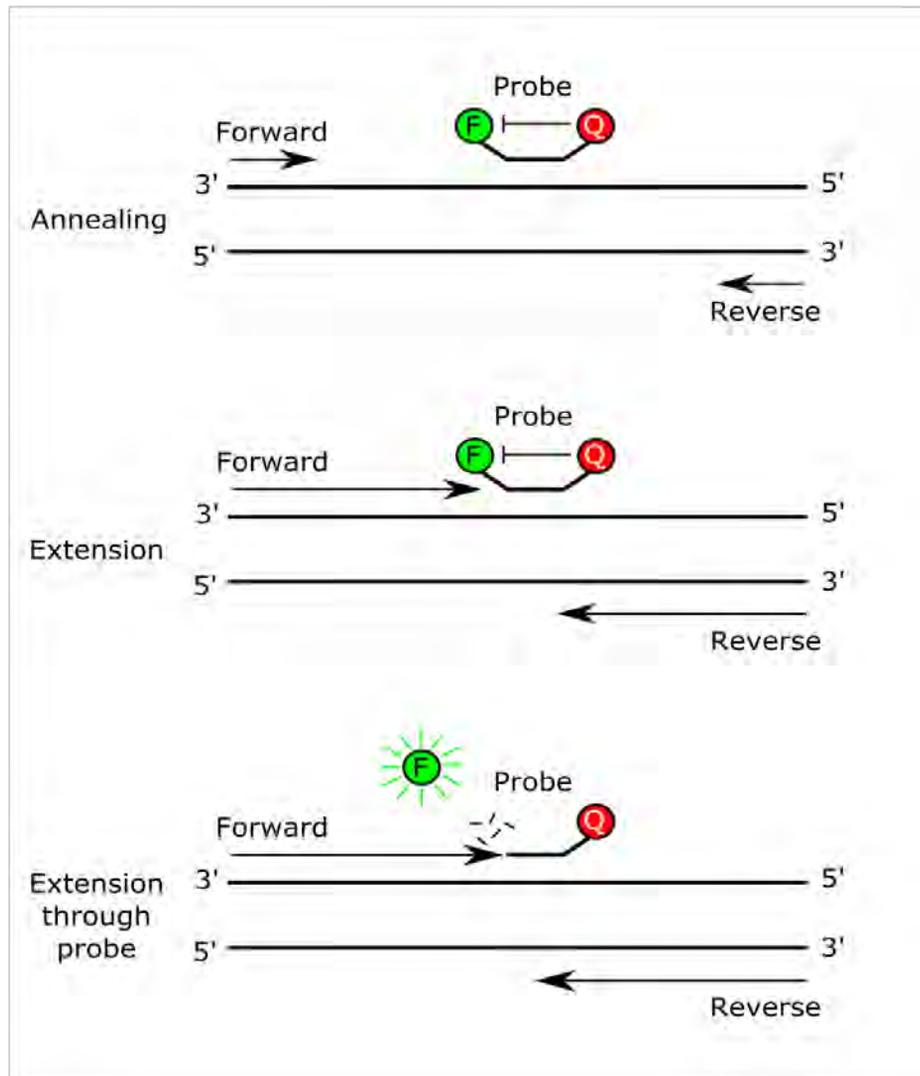
Much of the work performed during this PhD revolved around the technique of quantitative RT-PCR, such as the conformation of RNAi induced knockdown and validation of RNA-sequencing data. qRT-PCR allows an accurate quantification of relative expression of specific RNAs within the sample of interest providing a significant improvement on the previously used techniques of semi-quantitative PCR. Throughout this thesis TaqMan chemistry was used, other qRT-PCR chemistries are available such as SYBR green, however, TaqMan was used due generally showing greater sensitivity and reproducibility.

TaqMan qRT-PCR was developed by Applied Biosystems and contains 3 factors all of which are transcript specific; a forward primer, a reverse primer and a probe. The probe binds in between the forward and reverse primer and contains a fluorophore and a quencher. On annealing the quencher is in close proximity to the fluorophore resulting in the inhibition of

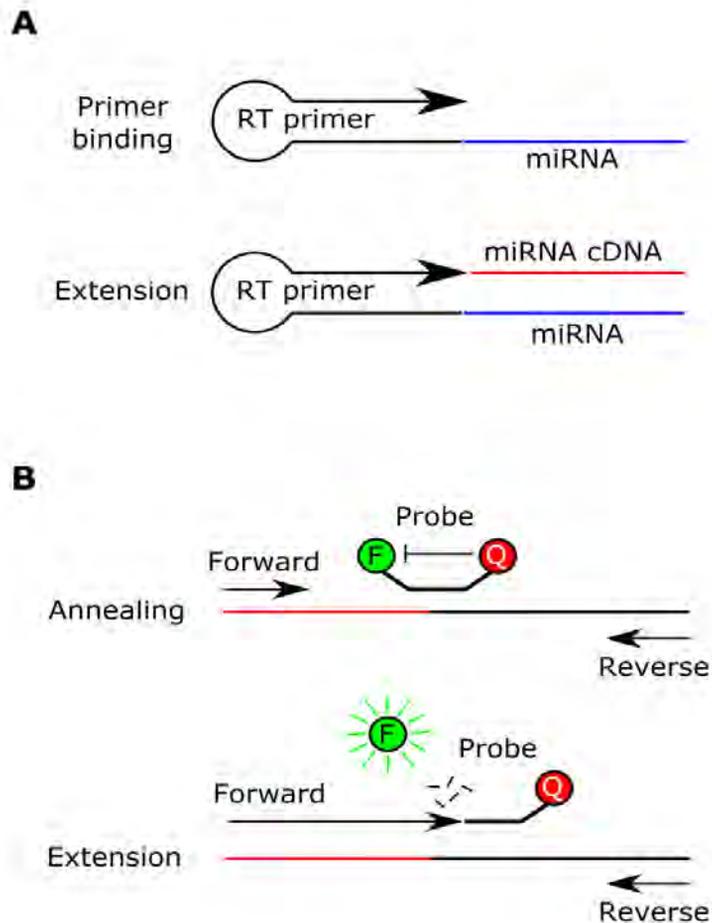
emission. However, on extension the Taq polymerase extends through the probe resulting in the release of the fluorophore away from the repressive quencher (Figure 1.12). This results in a fluorescent signal which is detected by the machine. Multiple cycles of annealing and extension are performed with the fluorescent signal captured at the end of each one; as the reaction progresses through the cycles more product is made meaning more fluorophore is realised and more signal is detected. The cycle at which the fluorescent signal (normalised against the background) crosses a threshold (cycle threshold (Ct)) is recorded and this value is used in differential expression analysis. This can be performed because the level of signal is directly proportional to the expression of the target. For example, a lowly expressed target would take more cycles to amplify a sufficient amount for the fluorescence to cross the threshold and therefore would have a higher Ct than a target expressed at higher levels.

TaqMan qRT-PCR was also performed on miRNAs during this project. Due to their small size (18-22nt) the technology required an adaptation which was achieved by Chen *et al* (Chen *et al.* 2005). This involved the use of a specific 'looped' reverse transcription (RT) primer which was specific to the target miRNA. This is then used to prime the cDNA reaction resulting in a hybrid containing the RT primer and a stretch of miRNA specific sequence (Figure 1.13A). The subsequent qRT-PCR is performed using a forward primer specific for the miRNA and a reverse primer that binds to the RT-primer. The same chemistry is used as outline previously with a probe which binds at the RT-primer-miRNA interface containing a fluorophore and a quencher. The extension of the Taq polymerase causes the separation of the fluorophore from the quencher leading to a detectable fluorescence signal (Figure 1.13B).

The analysis of qRT-PCR used in this thesis is discussed in methods section 2.4.2.5, but briefly a normaliser or 'housekeeper' is used to calculate the relative expression of the target of interest. It is essential that the expression (or Ct) of the housekeeper remains consistent across all the samples and is not affected by the treatments performed. This accounts for any technical error in the total amount of RNA put into the cDNA reaction or the total amount of cDNA put into the qRT-PCR. In order to easily assess this the same concentration of RNA was always used for the cDNA reaction and therefore subsequently used in the qRT-PCR. The expression level (Ct) of the target of interest is then normalised (subtracted) from the Ct of the housekeeper to give a relative expression, known as the delta Ct ( $\Delta Ct$ ). This can then be used to compare differential expression of the target of interest across the experimental conditions.



**Figure 1.12: How TaqMan probes work for qRT-PCR on mRNAs.** TaqMan qRT-PCR is a highly specific mechanism of quantifying relative expression of a target of interest. The system consists of 3 parts including forward and reverse primers which are specific to the target. The 3<sup>rd</sup> part of the system is a fluorescently labelled (F) probed which binds to the specific target between the forward and reverse primers. The probe also contains a quencher (Q) which when in close proximity inhibits the fluorophore. During extension (60°C) the Taq polymerase extends through the probe causing cleavage and the separation of the fluorophore resulting in a fluorescent signal which is detected by the machine.



**Figure 1.13: How TaqMan probes work for qRT-PCR on miRNAs. (A)** For use with miRNAs an adaptation to the TaqMan system using specific, looped reverse transcription (RT) primers was used. The RT primer associates with the specific miRNA of interest and is extended by reverse transcriptase resulting in the generation of an RT-primer-miRNA cDNA sequence. **(B)** The RT-primer-miRNA cDNA sequence is probed by qRT-PCR by using forward and reverse primers one specific to the miRNA sequence and the other to the RT-primer. A fluorescently labelled (F) probe also anneals at the interface between the RT-primer and miRNA sequence which also contains a quencher (Q) which when in close proximity inhibits the fluorophore. During extension (60°C) the Taq polymerase extends through the probe causing cleavage and the separation of the fluorophore resulting in a fluorescent signal which is detected by the machine.

#### 1.8.4 CRISPR.

The work presented in this thesis was carried out using RNA interference to reduce the amount of functional protein within the cell. Although effective, using a null mutant is highly beneficial as it would ensure there is no background activity, which there would always be in a knockdown. A portion of the work performed during this PhD was in an attempt to make a mutant by inducing the excision of a transposon. However, CRISPR has also been performed in an attempt to make a knockout, null mutant.

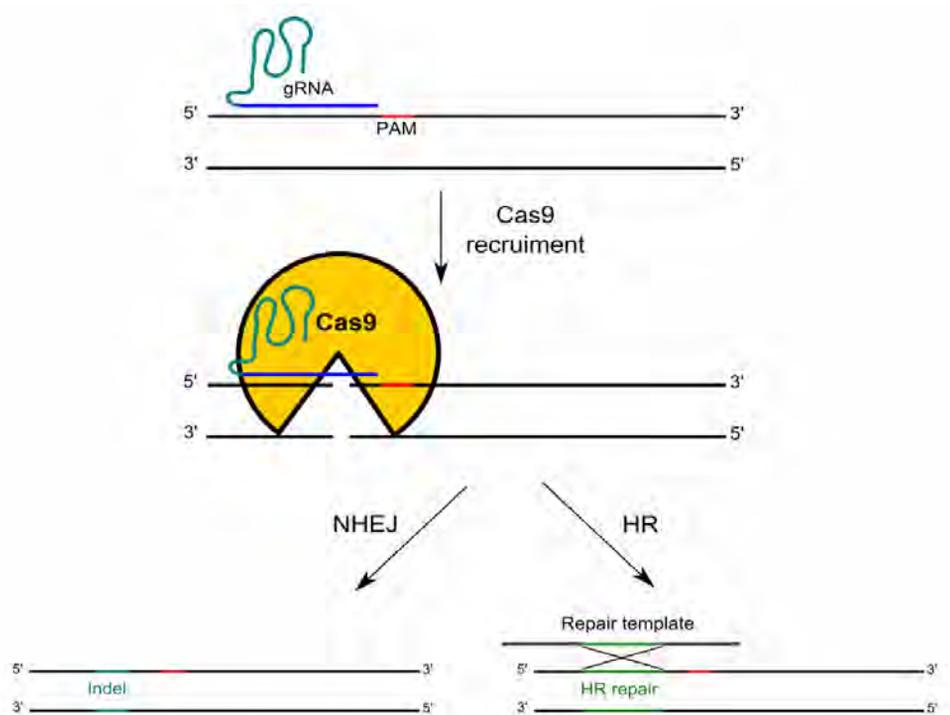
CRISPR or clustered regularly interspaced short palindromic repeats was developed from the adaptable mechanism by which bacteria respond to foreign DNA. It is an elegant system through which the nuclease Cas9 is targeted to complementary sequences in foreign DNA resulting in its cleavage. Cas9 has two nuclease domains (HNH and RuvC) which each cut a single strand of the DNA resulting in a double strand break (Jinek *et al.* 2012). The CRISPR technique has since been developed as a highly efficient method of creating mutants (Ran *et al.* 2013; Sander and Joung 2014). The user provides the targeting sequence, which guides Cas9 to the specific site in the gene they wish to mutate. The only major restriction for RNA targeting is the requirement of the Cas9 recognition motif (PAM) which is a trinucleotide sequence of 'NGG' that needs to be immediately 3' of the targeting RNA sequence (Ran *et al.* 2013). Within the 20nt guide sequence it has been shown that the final 12 nucleotides of the target sequence are essential for cleavage with a single mismatch enough to prevent the formation of a DSB. However, mismatches in the first 8 seem to have little effect.

CRISPR had been effectively used to produce mutants in a number of organisms including *Drosophila*, mice and humans (Bassett *et al.* 2013; Ran *et al.* 2013). On recruitment to the specific sequence in the target gene Cas9 induces a double strand break 3-4 nucleotides upstream of the PAM (Jinek *et al.* 2012). This break is then repaired by the error prone method of DNA repair known as non-homologous end joining (NHEJ) (Figure 1.14). This can result in the introduction of small insertions or deletions (indels) within the target gene which in turn may result in frameshift mutations leading to a premature stop codon.

In addition to the knockouts CRISPR has also been developed for knock in experiments including inducing point mutations or tagging the endogenous gene of interest. This is performed in essentially the same way although it relies on the presence of a repair template which is used by homologous recombination (HR) to repair the break. The genomic change desired is supplied on the repair template and thus incorporated into the genome when HR takes place (Ran *et al.* 2013) (Figure 1.14).

Adaptations to the system have also been made to improve specificity such as the D10A mutation in the RuvC domain of Cas9. This inhibits the activity of the RuvC domain meaning that the targeted Cas9 only cuts or 'nicks' one strand of the DNA; this has been renamed 'nickase' activity. This nicked strand can be repaired by HR. Alternatively, the user can supply two targeting guide RNA sequences in close proximity which results in two nicks that are in turn processed like a double strand break. In providing two targeting sequences it drastically reduces the chance of off target effects.

CRISPR has revolutionised mutagenesis and although other techniques have been used previously none of them are as elegant, easy to perform or as quick in generating mutants. From start to finish CRISPR has the potential to generate mutations (in cell lines) in as little as 1-2 weeks (Ran *et al.* 2013).



**Figure 1.14: Using CRISPR to generate mutants.** A 20 nucleotide guide RNA (dark blue) is designed to target a specific sequence followed by the PAM (NGG) sequence (red) and is coupled to a structured sequence (light blue). Cas9 is recruited to the target site by the structured region of the guide RNA and induces a double strand break 3-4bp upstream of the PAM sequence. This can then be repaired by non homologous end joining (NHEJ) which can result in the generation of small insertions or deletions (indels). Alternatively specific editing can be achieved through supplying the system with a repair template which allows the designed genomic change to be incorporated by homologous recombination (HR).

## **1.9 Aims.**

The overall aim of this project was to explore the roles of two related exoribonucleases, Dis3 and Dis3L2, in control of cell and tissue growth in a developmental tissue and to identify the specific transcripts they degrade. Much of the previous work into these enzymes has been performed in immortalised tissue culture cells or in single cell organisms which have little or no developmental potential. Therefore to investigate the role they may play in human overgrowth disorders it is critical to use a model system with developmental potential. As outlined above a reverse genetics approach was used to explore their functions in the development of the wing imaginal disc. The specific aims of each chapter are outlined below.

### **Chapter 3**

The main aims of the work in Chapter 3 were twofold. The first was to investigate the role of the 3'-5' exoribonuclease Dis3 within the development of the wing imaginal disc. This was to corroborate previous findings in *Drosophila* (Hou *et al.* 2012) and to look further into the roles of Dis3 in the development of the wing imaginal disc. A reverse genetics approach was used to knock down Dis3 and characterise any phenotypes observed. The ultimate aim was to add to the mechanisms by which Dis3 elicits its function. The second main aim of this chapter was to determine whether Dis3 plays a role in the maintenance of miRNA stability. To achieve this, a miRNA-sequencing experiment was performed on the wing imaginal discs. Experiments were then performed to determine if differentially expressed miRNAs are directly or indirectly affected by the loss of Dis3.

### **Chapter 4**

The aims of this chapter were to explore a potential role of Dis3L2 in the regulation of cell proliferation and tissue growth. Due to its implication in Perlman syndrome the hypothesis was that the loss of *dis3L2*, either by mutation or using RNA interference would result in increased cell and/or tissue growth. Previous work has been restricted to single cells either in human cell culture or in yeast cells and therefore this work aims to characterise a role in a developmental model; the wing imaginal disc. To achieve this, the *GAL4-UAS* system was used to restrict *dis3L2* knockdown to the wing imaginal discs to determine if Dis3L2 plays a role in the regulation of tissue growth as is indicated by the Perlman Syndrome phenotypes. This was followed by the characterisation of the observed phenotypes.

## **Chapter 5**

Chapter 5 aimed to identify specific RNA targets of Dis3L2 which when misexpressed contribute to the phenotypes observed in Chapter 4. This was achieved in a global and unbiased manner through performing a large scale RNA-sequencing experiment. The mapping and data analysis was performed using the Sussex High Performance Cluster and through stringent filtering potential Dis3L2 target candidates were selected and validated.

## **Chapter 6**

The work in Chapter 6 set out to explore the interactions and specificities of the cytoplasmic 3'-5' and 5'-3' pathways of RNA decay to test the hypothesis that the two opposing pathways of cytoplasmic RNA decay regulate distinct cellular processes. The specific aims were to explore the phenotypic effects of knocking down *dis3L2* and *pacman* simultaneously within the wing imaginal discs. In addition *dis3L2* was knocked down in both hypomorphic and null *pacman* mutant background and characterise any phenotypic effects. Finally, the specificity and/or redundancy of the two pathways were assessed through cross analysis of RNA-sequencing data sets.

## **Chapter 7**

The aims of the final chapter were of a more technical nature. These involved attempts to create a null or hypomorphic *dis3L2* mutant using P-element mutagenesis and CRISPR. The work also aimed to produce an antibody to *Drosophila* Dis3L2 as an antibody was not commercially available.

## Chapter 2: Methods

### 2.1 *Drosophila* stocks used during the project.

Full Genotype	Used for	Source
$y^1 w^{67c23} ; ; P\{y^{+mDint2} w^{BR.E.BR} = SUPor- P\}CG16940^{KG02284} ry^{506}$	P-element in <i>dis3L2</i> . Initial stock used.	BL13307
$w^* ; ; P\{w^{+mC} = GSV3\}GS6090 / TM6$	P-element upstream of <i>dis3L2</i> . Second stock used.	DGRC 200902
$ln^1 w^{m4h} ; ; Df(3L)Ly, sens^{Ly-1} / TM3, ry^{RK} Sb^1 Ser^1 P\{ry^{[+t7.2]} = Delta2-3\}99B$	Transposase stock	BL2030
$w^* ; ; UAS-pcm27^{ND} / TM6, Sb^1, Tb^1$	Used for its balancer in P-element mutagenesis crosses	Newbury lab
$w^{1118} ; ; Df(3L)Exel6084, P\{w^{+mC} = XP-U\}Exel6084 / TM6B, Tb^1$	Deficiency in <i>dis3L2</i> stock	BL7563
$y^1 w^* ; ; P\{w^{+mC} = tubP-GAL4\}LL7 / TM6b, GFP$	<i>tubulin-GAL4</i> driver	BL5138
$y^1 w^* ; ; P\{w^{+mC} = Act5C-GAL4\}25FO1 / CyO, y^+ ;$	<i>actin-GAL4</i> driver	BL4414
$w^{1118} ; ; da-GAL4$	<i>daughterless-GAL4</i> driver	Gift from Barry Denholm
$w^* ; ; P\{w^{+mW.hs} = GawB\}69B$	<i>69B-GAL4</i> driver	BL1774
$; en-GAL4 UAS-GFP-actin / CyO ;$	<i>enrailed-GAL4</i> driver	Gift from Paul Martin
$P\{w^{+mC} = UAS-Dcr-2.D\}1, w^{1118} ; P\{w^{+mW.hs} = GawB\}nubbin-AC-62 ;$	<i>nubbin-GAL4</i> driver	BI25754
$P\{UAS-Dcr-2.D\}1, w^{1118} ; P\{GawB\}pnrMD237 / TM3, Ser^1$	<i>pannier-GAL4</i> driver	BL25728
$w^* ; ; P\{tubP-GAL80^{ts}\}20 ; TM2 / TM6$	Used to make <i>tub-GAL80<sup>ts</sup></i> ; <i>69B-GAL4</i> stock	BL7019
$w^* ; ; P\{tubP-GAL80^{ts}\}20 / CyO-GFP ; P\{GawB\}69B / TM6$	<i>tub-GAL80<sup>ts</sup></i> ; <i>69B-GAL4</i> stock used to assess developmental timing	Newbury lab
$; ; M\{UAS-miR-252.S\}ZH-86Fb$	Used to overexpress <i>miR-252</i>	FlyORF F002053
$; ; UAS-GFPactin ;$	GFP under the control of an UAS	BL1522
$; ; P\{VALIUM22-EGFP.shRNA.1\}attP40 ;$	<i>EGFP<sup>RNAi</sup></i> construct under control of a UAS.	BL41557
$w^{1118} ; ; P\{GD11917\}v35090$	<i>dis3<sup>RNAi</sup></i> construct under control of a UAS. Initial RNAi used.	Vienna 35090

Full Genotype	Used for	Source
; P{KK101473}VIE-260B ;	<i>dis3</i> <sup>RNAi</sup> construct under control of a UAS. Second RNAi used.	Vienna 108013
<i>w</i> <sup>1118</sup> ; P{GD9240}v51854 ;	<i>dis3L2</i> <sup>RNAi</sup> construct under control of a UAS. Initial RNAi used.	Vienna 51854
<i>w</i> <sup>1118</sup> ; ; P{GD9240}v51853	<i>dis3L2</i> <sup>RNAi</sup> construct under control of a UAS. Second RNAi used.	Vienna 51853
; P{KK105902}VIE-260B ;	<i>dis3L2</i> <sup>RNAi</sup> construct under control of a UAS.	Vienna 100322
<i>w</i> <sup>1118</sup> ; P{GD6902}v38356 ;	<i>Ski</i> <sup>RNAi</sup> construct under control of a UAS.	Vienna 3856
<i>w</i> <sup>1118</sup> ; P{GD10414}v40306 ;	<i>Rrp40</i> <sup>RNAi</sup> construct under control of a UAS.	Vienna 40306
<i>w</i> <sup>1118</sup> ; ; P{GD109216}v21677	<i>pcm</i> <sup>RNAi</sup> construct under control of a UAS.	Vienna 21677
<i>w</i> <sup>*</sup> ; ; Df(3L)H99 <i>kni</i> <sup>ri-1 p</sup> /TM6B, <i>Tb</i> <sup>1</sup> , <i>Hu</i> <sup>1</sup>	Deficiency stock used to inhibit apoptosis	BL1576
<i>pcm</i> <sup>5</sup> /FM7i ; ;	Hypomorphic mutation in <i>pacman</i>	Newbury lab
<i>pcm</i> <sup>14</sup> /FM7i ; ;	Null mutation in <i>pacman</i>	Newbury lab
50E/FM7i ; ;	Isogenic control for <i>pcm</i> <sup>14</sup>	Newbury lab
<i>pcm</i> <sup>5</sup> /FM7i ; UAS- <i>dis3L2</i> <sup>RNAi</sup> /CyO-GFP ;	Knockdown <i>dis3L2</i> in <i>pcm</i> <sup>5</sup> mutant background	Newbury lab
<i>pcm</i> <sup>14</sup> /FM7i ; UAS- <i>dis3L2</i> <sup>RNAi</sup> /CyO-GFP ;	Knockdown <i>dis3L2</i> in <i>pcm</i> <sup>14</sup> mutant background	Newbury lab
50E/FM7i ; UAS- <i>dis3L2</i> <sup>RNAi</sup> /CyO-GFP ;	Knockdown <i>dis3L2</i> in 50E control background	Newbury lab
<i>w</i> <sup>*</sup> ; <i>Kr</i> <sup>rf-1</sup> /CyO-GFP ; MKRS/TM6B <i>Tb</i> <sup>1</sup> , <i>Hu</i> <sup>1</sup>	Doubly balanced stock used to create stocks	Newbury lab

## 2.2 *Drosophila* husbandry.

### 2.2.1 Fly food and crosses.

*Drosophila* stocks were cultured on standard media containing Agar (GP) (81g), Bakers yeast (81g), Oatmeal (616g), Black treacle (410g), propionic acid (40ml), Nupagin ( 1 spatula), Distilled H<sub>2</sub>O (up to 7Litres). Stocks were maintained at 25°C and turned over into fresh food every 2 weeks. For all genetic crosses virgin females were mated to male flies. Virgin females were selected using the presence of the meconium in the abdomen (the remainder of the last larval meal) which ensures the females are only a few hours old as at 25°C *Drosophila melanogaster* females do not mate for the first 8-10 hours of adult life. Virgins were collected every 2-3 hours and kept in vials until crossing; these vials were kept for a few days after the crosses had been set up to ensure no larval activity which would indicate the presence of a non-virgin.

### 2.2.2 Balancer chromosomes.

Balancer chromosomes are a key tool in a fly geneticist's armoury; they allow the tracking of chromosomal inheritance together with preventing recombination in female flies. In preventing recombination they ensure the stability of genetic variants allowing stable stocks to be maintained. Balancer chromosomes such as CyO and TM6 carry dominant observable phenotypes which enables the identification of inheritance. For example, CyO gives the flies curly wings, an easily identifiable phenotype. There are a variety of balancers across the 3 main chromosomes X, 2 and 3 and those used in this thesis are outlined in the table below.

Balancer	Chromosome	Phenotype
<i>FM7i</i>	X	Heart/Bar eye ( $B^1$ ), Yellow body ( $y^{93j}$ ).
<i>CyO</i>	2	Curly wings ( $Cy^1$ ).
<i>lf</i>	2	Black speckled eye ( $Kr^{lf-1}$ ).
<i>MKRS</i>	3	Thoracic hairs short ( $Sb^1$ )
<i>TM3</i>	3	Thoracic hairs short ( $Sb^1$ ), Portions of wing missing ( $Ser^1$ ).
<i>TM6B</i>	3	Pupae and adults tubby ( $Tb^1$ ). Cluster of hairs on 'shoulder' ( $Antp^{Hu}$ ).

### **2.2.3 Developmental staging.**

For larval aging egg lays were performed. Virgin females were crossed to males 2 days prior to the required cross day. These were then tipped into fresh vials and left to lay for between 1-8 hours depending on the period of egg lay required. Larvae were left to develop until 96, 120 hours old where they were dissected. Alternatively egg lays were performed on food containing 0.05% bromophenol blue which allows visualisation of the food in the larval gut. Late L3 were staged by the absence of food from the gut as late L3 larvae clear the food from their gut immediately prior to pupariation.

For wing measurement experiments flies were collected in the first few hours of eclosure and kept in fresh vials at the required temperature (that of the original cross) for 24 hours to ensure they were between 1-2 days old. Wings were then removed and mounted for measurement.

The work in Chapter 6 required the selection of *pacman*<sup>14</sup> mutant larvae. This was achieved by selecting against the GFP present on the FM7i balancer chromosome. The vials containing the developing larvae were examined under a Leica MZ16F microscope equipped with a GFP filter. Those carrying the FM7i balancer were easily identified through a strong GFP signal throughout the larva. The male, non-GFP, larvae were then further selected for by examining them under a Nikon SMZ800 dissection microscope using the larval testis as a male marker.

### **2.2.4 Measurement of developmental delay.**

Crosses were set up at 25°C with flies left for 8 hours to lay eggs (8hr egg lay) with time 0 starting after the 8 hours. Times were recorded when progeny reached wandering L3, pupal and adult stages of development.

### **2.2.5 Wing imaginal disc dissection for RNA extraction.**

Wing imaginal discs were dissected from wandering 3<sup>rd</sup> instar larvae (L3) in Ringers solution (3mM CaCl<sub>2</sub>, 182mM KCl, 46mM NaCl, 10mM Tris (pH7.2)) under a Nikon SMZ800 dissection microscope. The discs were dissected in batches of 30 in a single Eppendorf tube containing 150µl of Ringers solution on ice. Excess Ringers was removed and samples were snap frozen in liquid nitrogen and stored at -80°C until required.

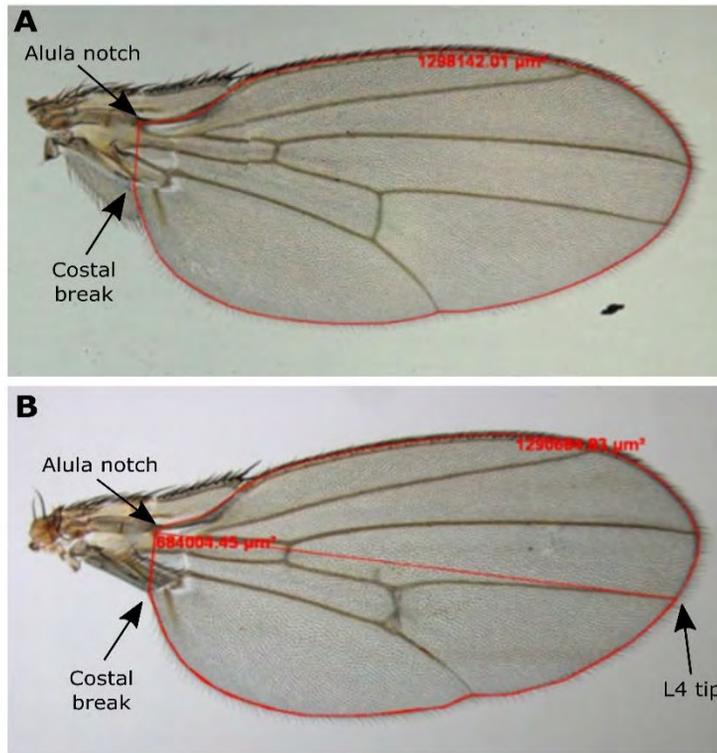
### **2.2.6 Wing imaginal disc dissection for measurement.**

To ensure the larvae were all at the same developmental stage (Late L3) the flies were grown on media containing bromophenol blue (section 2.2.3). Late L3 wing imaginal discs were dissected in PBS under a Nikon SMZ800 dissection microscope and mounted on Ploy-l-lysine microscope slides (no more than 15 per slide) in 15µl of 85% glycerol. Two 22mm x 22mm glass

coverslips were placed either side and a single 22mm x 40mm glass coverslip was placed over the top to ensure the discs were not squashed. Slides were sealed using nail varnish. The measurement of the mounted discs was achieved using the program Axiovision 4.7 on an Axioplan microscope (Carl Zeiss). The area of each disc was measured by drawing around the disc using the 'outline' tool.

### **2.2.7 Wing dissection for measurement.**

Flies were aged to 1-2 day old at the same temperature at which the initial cross was carried out (section 2.2.3). A single wing ('left hand' wing) was then cut from each fly using microscissors and placed in wells of isopropanol for at least 1 hour. This step was to prevent the formation of bubbles in the wing veins. Wings were then mounted onto microscope slides in DPX (Sigma, cat. No. 44581) and covered with a 22mm x 40mm cover slip and left to dry overnight. Wings were then imaged and measured using the program Axiovision 4.7 on an Axioplan microscope (Carl Zeiss). The area of the each wing was measured by drawing around the disc using the 'outline' tool. Figure 2.1 shows the markers used to measure the total area and posterior area of the wing.



**Figure 2.1: Guidelines used for the measurement of *Drosophila* wings.** (A) The total area of the adult wing was measured by tracing around the outside of the wing using the wing margins together with the alula notch and costal break as reference points at the hinge. (B) The posterior area of the wing was measured by tracing a straight line from the costal break to the alula notch followed by a second straight line from the alula notch to the tip of L4 passing through the anterior cross vein. The anterior area was calculated by taking the posterior area away from the total area of the wing measured as in A.

### **2.2.8 Counting the number of cells in the wing.**

To estimate the number of cells in each measured wing the number of hairs were counted in different 0.1mm<sup>2</sup> sections of the wing in both the anterior and posterior regions of the wing. Each cell in the wing protrudes a single hair; therefore each hair represents a single cell. A mean was taken across the different location and multiplied up using the total wing area to give an estimation of the number of cells in the wing. To calculate the area of each cell the total area of the wing was divided by the number of cells.

### **2.2.9 Photographing fly phenotypes.**

Flies were frozen at -20°C for 15 minutes. After checking freezing had not affected the phenotype images were taking using a Nikon D5200 camera mounted on a Nikon SMZ800 dissecting microscope. A series of images were taken through the focal plane which were subsequently stacked and merged in Photoshop (CS6) producing the final image.

### **2.2.10 Starvation experiment.**

Flies were cultured under normal conditions with no overcrowding and collected in the first few hours of adult life and aged to 1-2 days old as outlined in section 2.2.3. Flies were transferred into new vials containing no food and scored over a period of 55 hours. Mating schemes were designed to produce both knockdown and sibling control flies to account for any variation in development.

## **2.3 Genetic techniques.**

### **2.3.1 The *GAL4-UAS* system.**

The *GAL4-UAS* system is a vital tool used by *Drosophila* geneticists (Figure 2.2A). Derived from yeast it allows control over the expression domains of genomic constructs. The *GAL4-UAS* system is a bipartite system using the yeast transcription factor GAL4. The GAL4 gene is coupled to an upstream promotor which controls the expression domain (see Figure 2.3 for drivers used). In the absence of the Upstream Activation Sequence (UAS) section of the system the GAL4 function is repressed due to a lack of binding sites. The *UAS* contains multiple GAL4 binding sites as is inserted upstream of the desired expression construct, for example an RNAi construct. *UAS-RNAi* and *UAS-cDNA* constructs were used to drive knockdowns and overexpression respectively.

An additional arm to the system is the GAL80 temperature sensitive (GAL80<sup>ts</sup>) repressor of GAL4 which adds a temporal control over the activity of GAL4. At low temperatures (19°C) GAL80 is active and inhibits the activity of GAL4 thus preventing the activation of the UAS-

Construct. However, on moving flies to a higher temperature (29°C) results in the inhibition of GAL80 allowing the *GAL4-UAS* system to function as normal, resulting in the controlled special expression of the desired construct. This system was used to investigate the stage in *Drosophila* development in which Dis3L2 is critical to control wing disc growth.

### 2.3.2 GAL80<sup>ts</sup> addition and experimental design.

An additional arm to the system is the GAL80 temperature sensitive (GAL80<sup>ts</sup>) repressor of GAL4 which adds a temporal control over the activity of GAL4. At low temperatures (19°C) GAL80 is active and inhibits the activity of GAL4 thus preventing the activation of the *UAS-Construct*. However, on moving flies to a higher temperature (29°C) results in the inhibition of GAL80 allowing the *GAL4-UAS* system to function as normal, resulting in the controlled special expression of the desired construct (Figure 2.2B). This system was used to investigate the stage in *Drosophila* development in which Dis3L2 is critical to control wing disc growth.

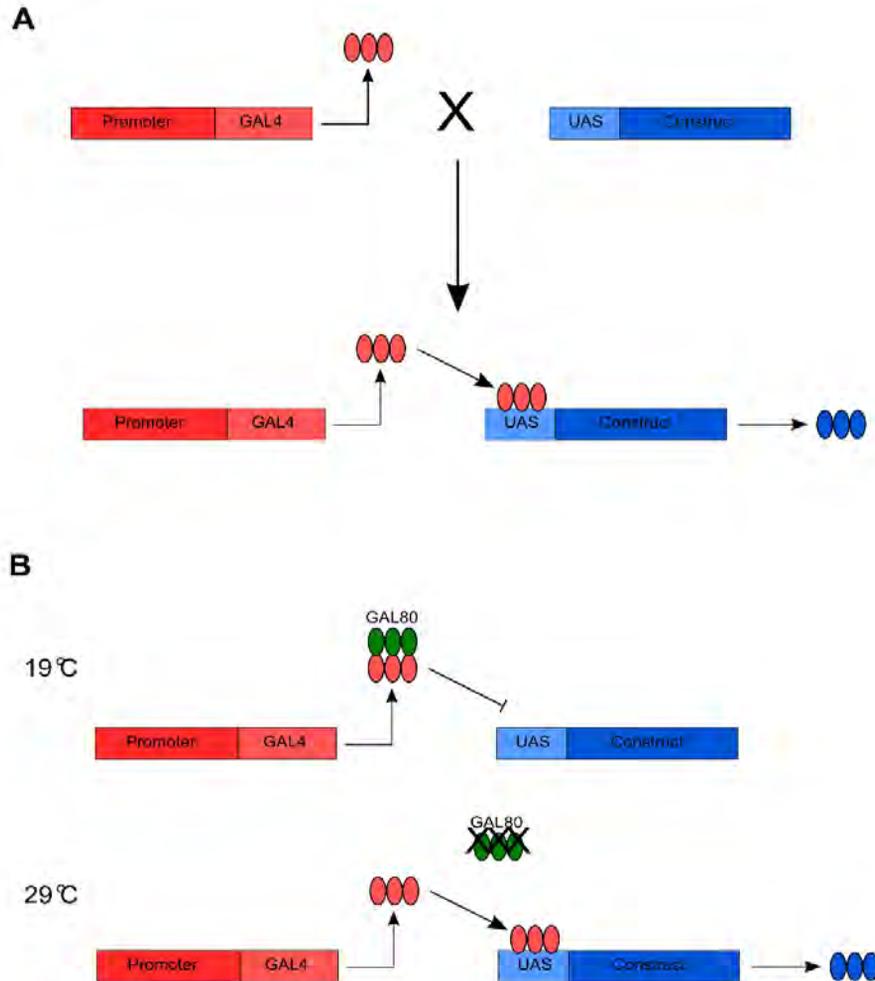
To use the GAL80<sup>ts</sup> system to investigate the requirement of Dis3L2 throughout development ;*tub-GAL80<sup>ts</sup>*;69B-*GAL4* homozygous flies were crossed to ;*UAS-dis3L2<sup>RNAi</sup>*; and left to mate for 2 days at the cross temperature in addition to parental controls for all time points (19°C for all developmental timings except the positive control at 29°C). Flies were tipped into new vials for a 6 hour egg lay. Vials for both the parental controls and knockdowns were moved from 19°C to 29°C at the required times for each developmental stage taking into account a 24-32hr perdurance. The developmental stages selected were; start of L2, start of L3, mid-L3 and the start of pupariation. Each comparison between knockdown and controls were made specifically between those for each time point. See Figure 4.11 (Chapter 4) for the experimental plan.

### 2.3.3 P-element mutagenesis.

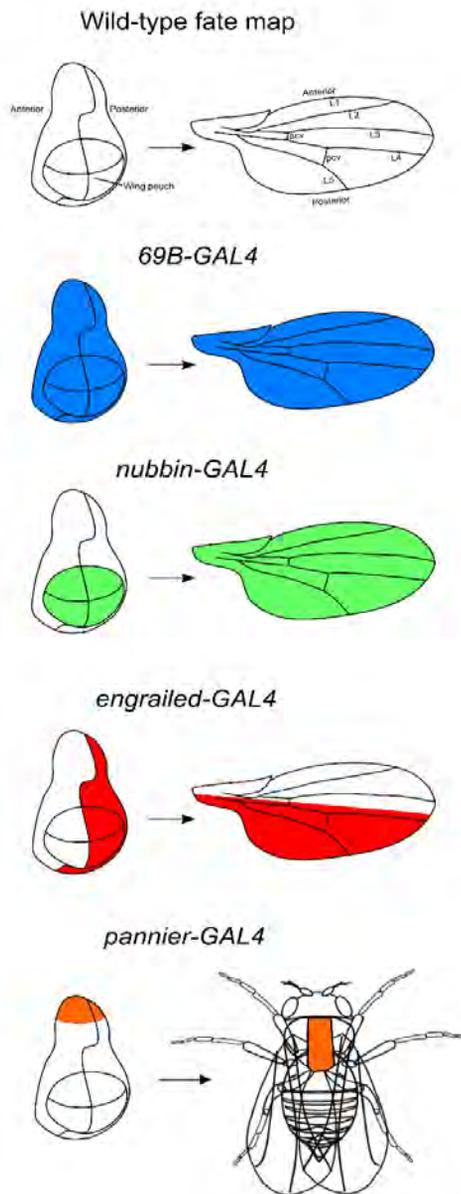
Crosses for the mutagenesis experiments were carried out at 25°C. The stock *P{GSV3}GS6090* (Kyoto 200902) contained a P-element 11bp upstream of *dis3L2* and was used in an attempt to create a *dis3L2* mutant. The stock originally contained a factor causing homozygous lethality; however, as the factor was outside of the *dis3L2* locus (shown using a deficiency) it was removed using recombination (section 2.3.4). The transposase was crossed into the P-element carrying flies resulting in its random excision. Successful excision was identified by an eye colour change from orange to white. Each white eyed fly was then isolated and a stable stock was produced. Each stock was subsequently screened by PCR to identify any potential mutations. Cross scheme shown in Figure 7.3 (Chapter 7).

#### **2.3.4 Recombination crosses.**

The P-element containing *P{GSV3}GS6090* stock carried a homozygous lethal mutation on the 3<sup>rd</sup> chromosome which was outside of the *dis3L2* locus. A series of crosses were performed to induce recombination in attempt to repair the damaged region (Figure 7.5 in Chapter 7). The damaged chromosome was crossed over a wild type chromosome and allowed to recombine. Each potential recombination event was isolated and stable stocks were produced. The presence of homozygous viable lines showed successful recombination. PCR was used to confirm the presence of the P-element in the rescued stocks. Stock 5 was selected to use in the subsequent P-element mutagenesis experiments.



**Figure 2.2: The *GAL4-UAS* system and the *GAL80<sup>ts</sup>* addition. (A)** The *GAL4-UAS* system is a bipartite system consisting of the yeast transcription factor, GAL4 and a series of GAL4 specific binding sites (UAS). The sequence encoding GAL4 is coupled to a specific promoter, the nature of which depends on the desired expression domain. The activity of the promoter drives the production of GAL4 in the expression domain of the gene from which the promoter is taken. The expression of the desired construct is achieved through crossing the UAS- and GAL4 constituents together resulting in progeny containing both sections of the system. In the presence of the UAS, the GAL4 protein binds and drives the expression of the desired construct. **(B)** The temperature sensitive GAL80 (*GAL80<sup>ts</sup>*) functions as a GAL4 inhibitor and is active at 19°C. *GAL80<sup>ts</sup>* mediated inhibition of GAL4 results in inhibition of the expression of the construct. At 29°C *GAL80<sup>ts</sup>* is inactive which restores the activity of GAL4 and subsequent expression of the desired construct.



**Figure 2.3: Wing imaginal disc GAL4 drivers used during this project.** 4 wing disc specific GAL4 drivers were used during this project. *69B-GAL4* (blue) is expressed throughout the wing imaginal disc resulting in the expression of the desired construct throughout the adult wing in addition to the thorax. *nubbin-GAL4* (green) is expressed in the wing pouch and therefore drives the expression of UAS-constructs specifically within the wing blade committed cells. *engrailed-GAL4* (red) is specifically expressed in the posterior compartment of the wing imaginal disc whilst *pannier-GAL4* (orange) is expressed in the thoracic region of the wing disc.

## **2.3.5 CRISPR.**

### **2.3.5.1 gRNA design.**

The key to a successful CRISPR experiment is in the selection of the gRNA targeting sequence. The coding sequencing of *dis3L2* was submitted to two independent online gRNA identification tools; Cas9 Target Finder found on the National Institute of Genetics (NIG-FLY) website (<http://www.shigen.nig.ac.jp/fly/nigfly/cas9/cas9TargetFinder.jsp>) and the CRISPR optimal target finder found on the flyCRISPR website. The CDS for both isoforms was submitted to ensure the gRNA would target both. After running a BLAST search to identify potential off target effects the most 5' target sequence (GGAAACGTCAAGCGTCAACG) was selected.

### **2.3.5.2 Annealing and phosphorylation of gRNA oligos.**

Complementary gRNA oligos were resuspended to 100µM and annealed and phosphorylated in the following reaction mix:

1µl sense oligo (100µM)  
1µl anti-sense oligo (100µM)  
1µl 10X T4 Ligation buffer (NEB cat. no. M020S)  
0.5µl T4 PNK (NEB cat. no. M0201S)  
6.5µl H<sub>2</sub>O

The mix was incubated under the following conditions:

1. Anneal and phosphorylate 37°C – 30 mins
2. Inactivate PNK 95°C – 5 mins
3. Cooling – cool down to 25°C at 5°C/min.

### **2.3.5.3 Digestion and dephosphorylation of pCFD3-dU6:3gRNA**

The CRISPR vector selected was the pCFD3-dU6:3gRNA (pCFD3) vector. This vector contained the core structured gRNA sequence which required the cloning of the targeting gRNA sequence immediately upstream. pCFD3 was digested with Bbs1 in the following reaction:

1µl units Bbs1 (10 units)  
1µl pCFD3 (1µg)  
5µl 10x NEB buffer 2.1 (cat no)  
43µl H<sub>2</sub>O.

This mix was incubated for 30 mins at 37°C then 1µl of calf intestinal alkaline phosphatase was added to prevent re-ligation. This was followed by a further 30 mins at 37°C. The digested

product was gel purified using a Qiagen Gel Extraction kit (Qiagen, cat. no. 28704 ) and followed the manufacturer's instructions.

#### **2.3.5.4 Ligation of gRNA into pCFD3.**

The annealed and phosphorylated oligos (section 2.3.5.2) were then ligated into the gel purified, digested, pCFD3 using the following mix:

4.5µl Bbs1 digested pCFD3 (50ng)  
1µl Annealed and phosphorylated gRNA oligos (diluted 1:200)  
1.5µl 10x T4 DNA ligase buffer (NEB cat. no. B0202S)  
1µl T4 DNA ligase (NEB cat. no. M0202S)  
7µl H<sub>2</sub>O

The above mix was incubated for 1 hour at room temperature. Successful ligation was confirmed by PCR using a gRNA forward primer (CRISPR-42F) and a primer in the vector backbone (T3).

#### **2.3.5.5 Transformation of pCFD3 into DH5α cells.**

The gRNA-pCFD3 plasmid was transformed into competent DH5α cells. Two control samples were run alongside; pCFD3 without the gRNA (positive control) and an untransformed negative control. 10ng of gRNA-pCFD3 or pCFD3 were incubated with 50µl of competent DH5α cells for 30 minutes on ice. After 15 minutes the mix was vortexed for 1 second and returned to the ice for the remaining 15 minutes. The transformation mix was then heat shocked at 42°C for 1 minute in a water bath before returning to the ice for 2 minutes.

500µl of LB-broth (without antibiotic) (Sigma, cat. no. L3022) was added to the transformation mix which was incubated for 1 hour at 37°C in a shaking incubator. This was to allow recovery and allow the antibiotic resistance to develop. Finally, each mix was spread onto 2 LB-Agar (Sigma, cat. no. 52062) plates containing 100µg/ml ampicillin. 100µl was spread onto the first and 300µl was spread onto the second. The plates were incubated overnight at 37°C.

#### **2.3.5.6 Colony picking and inoculation.**

Surviving colonies were picked using a 0.1-10µl pipette tip which was then placed into individual 5ml of LB-broth containing 100µg/ml ampicillin. 5 colonies were picked per gRNA-pCFD3 plate and 2 colonies were picked from each pCFD3 plate. No colonies were present on the untransformed control. These were grown overnight at 37°C in a shaking incubator.

### 2.3.5.7 Minipreps and sequencing.

To extract the plasmid DNA minipreps were performed using 1.5ml of the overnight inoculation mix. This was achieved using a QIAprep Spin Miniprep kit (Qiagen, cat. no. 27106) following the manufacturer's instructions. The isolated DNA was quantified and its purity was assessed using a NanoDrop 1000. 15µl of 75ng/µl samples were sent to MWG Eurofins for DNA sequencing mixed with 2µl of the T3 reverse sequencing primer (10nM). 20µg of isolated DNA was sent to The BestGene.inc for embryo injection into a *vasa-Cas9* ;; line.

## 2.4 Molecular techniques.

### 2.4.1 Polymerase Chain Reaction.

#### 2.4.1.1 Squish prep.

'Quick' DNA was prepared from 3 male flies using 150µl of squish buffer (10mM Tris (pH 8.2), 1mM EDTA, 25mM NaCl) and 0.75µl Proteinase K. Flies were homogenised using a pestle then incubated at 37°C for 30 minutes followed by 5 minutes at 95°C to denature the Proteinase K.

#### 2.4.1.2 PCR conditions.

The mastermix and reaction conditions for each 25µl PCR were:

12.5µl Amplitaq Gold 360 Master Mix

10.5µl H<sub>2</sub>O

1µl 5µM mix of forward and reverse primer

1µl Squish prep product

PCR protocol:

- 1) Activation – 95°C for 10 minutes
  - 2) Denaturing - 95°C for 1 minute
  - 3) Annealing - 57°C for 30 seconds
  - 4) Extension - 72°C for 90 seconds
  - 5) Final Extension - 72°C for 10 minutes
- } Steps 2-4 for 35 cycles

#### **2.4.1.3 Primers used and their design.**

Primers were designed using NCBI primer blast. A genomic region of between 700 and 1500bp was submitted with restrictions on annealing temperature (57-62°C). The resulting primer pairs were blasted against the *Drosophila* genome to ensure specificity. The suitable primers were then ordered from Sigma-Aldrich. See Table for the primer sequences and their use in this project.

#### **2.4.1.4 Agarose gel electrophoresis.**

PCR products were run on 1% agarose gels. The 1% agarose gel was made by adding 1.5g agarose to 150µl 1xTBE (90mM Tris, 90mM Boric acid, 2mM EDTA). 1.5µl of Gel Red was added prior to microwaving to allow UV visualisation of DNA. The solution was then microwaved for 90 seconds and poured into a Gel cassette where it was left to set. 5µl of loading buffer was added to the 25µl PCR product with 10µl subsequently loaded into the 1% gel. Gels were run at 120V submerged in 1xTBE then visualised on a UV illuminator.

#### **2.4.1.5 PCR purification and sequencing.**

For sequencing of PCR products, as used in the mutant diagnostic PCR, products were run on 1% gels as above to ensure only a single band was present. If only one band was present the PCR product was 'cleaned up' using a Qiagen 'PCR Purification' kit following the manufacturer's instructions. If more than one band was present the correct band was cut out of the gel followed by extractions and purification of the DNA using a Qiagen 'PCR Gel extraction' kit following the manufacturer's instructions.

DNA concentration and integrity was examined using a NanoDrop 100. 15 µl of purified PCR product (15ng/µl) was pre-mixed with 2µl of 2µM of either the forward or reverse primer and sent to MWG Eurofins for DNA sequencing. The DNA sequencing results were examined for purity (one clear base call per base) then aligned to the expected sequence using the Align X tool in Vector NTI.

Primer	Sequence	Use
IsoF1	ACCCTTTATATCCCGCCCAT	Detecting the two isoforms of <i>dis3L2</i>
IsoR1	GGAGGCCCTCAATCCTTCTG	Detecting the two isoforms of <i>dis3L2</i>
Dis3L2 F1	TCCTACTGTA CTTGTCCTATCATTT	Primer binding site 2bp downstream of GS6090 insertion site.
Dis3L2 R1	GTTGCGGATGTGCTTCTGAC	Partner for Dis3L2 F1
Thoc7 F1	GAGATCGTCGCTTCATCGCT	Primer pair in <i>thoc7</i> used to ensure any deletion did not go into <i>thoc7</i> . Also used in P-element checking.
Thoc7 R1	ACTGCGTGGTATGTCGGA	Partner for Thoc7 F1
Dis3L2 ex1 F1	GACATACCACGCAGTACGGT	Binds in the first exon of <i>dis3L2</i> . Used for deletion screening
Dis3L2 ex1 R1	GAGCCCGCTTCGACTTCTTT	Partner for Dis3L2 ex1 F1
Dis3 ex1 F1	GGCAGTCCGTAGACTCAAAC	Binding site in the first exon of <i>dis3</i> . Used as a positive control throughout
Dis3 ex1 R1	GCAGGTGATCATCGTACCAC	Partner for Dis3 ex1 R1
Pry2	CTTGCCGACGGGACCACCTTATGTTATT	Specific binding site in the <i>GS6090</i> insertion. Used to validate its presence.
Pry 4	CAATCATATCGCTGTCTCACTCA	Specific binding site in the <i>GS6090</i> insertion. Used to validate its presence. (Did not work)
Spep1	GCACTCAGAATACTATTC	Specific binding site in the <i>GS6090</i> insertion. Used to validate its presence. (Did not work)
Sp1	ACACAACCTTTCCTCTCAACAA	Specific binding site in the <i>GS6090</i> insertion. Used to validate its presence. (Did not work)
CRISPR-42F	GTCGGAAACGTCAAGCGTCAACG	gRNA sequence used in the CRISPR sequence. Also used to validate its successful ligation.
CRISPR-42R	AAACCGTTGACGCTTGACGTTTCC	Complementary sequence to which CRISPR-42F was annealed before ligation into pCFD3.
T3	GCAATTAACCCTCACTAAAGG	Sequencing primer used to check the successful cloning of the gRNA into the correct site of pCFD3.

## **2.4.2 quantitative RT-PCR.**

### **2.4.2.1 RNA extraction on whole adults/L3 larvae.**

RNA extractions were performed on 3 whole adults or L3 larvae by homogenising in Qiazol lysis buffer (Qiagen, cat. no. 217084). The RNA and protein was separated using a phenol-chloroform spin at 4°C. The aqueous layer was removed and added to 1.5 volume of 100% ethanol and spun at 4°C to pellet the RNA. The RNA pellet was subsequently washed in 70% ethanol and spun once more at 4°C. The RNA pellet was left to air dry (max 10 minutes) before resuspension in 25µl of RNase free H<sub>2</sub>O.

### **2.4.2.2 RNA extraction on dissected wing imaginal discs.**

RNA extractions were performed using a miRNeasy RNA extraction kit by the manufacturers instructions (Qiagen, cat. no. 217084), with an on-column DNase digestion (Qiagen, cat. no. 79254). RNA was eluted in 14µl RNase free H<sub>2</sub>O and its concentration was measured on a NanoDrop1000 spectrophotometer (Thermo Scientific).

### **2.4.2.3 cDNA preparation from total RNA.**

All RNA samples were diluted to a consistent concentration (normally 50ng/µl). cDNA was then prepared in duplicate together with a no RT control where reverse transcriptase was replaced by H<sub>2</sub>O. cDNA was made using Life Technologies High Capacity cDNA Reverse Transcription kit (Life Technologies, cat. no. 4368814) with the following composition:

10µl reaction

1µl Buffer

1µl Primer (either Random hexamers or oligo(dT))

0.5µl Reverse Transcriptase

0.4µl dNTPs

2.1µl H<sub>2</sub>O

5µl Diluted RNA

Samples were then run in a thermal cycler under the following conditions:

- 1) Annealing 25°C – 10 mins
- 2) Incubation 37°C - 2 hours
- 3) Inactivation 85°C – 5 mins
- 4) Hold 4°C - ∞

#### **2.4.2.4 cDNA preparation from miRNAs.**

RNA was diluted to 2ng/ $\mu$ l then cDNA was prepared in duplicate together with a no RT control where the reverse transcriptase was replaced by H<sub>2</sub>O. miRNA cDNA reactions are specifically primed therefore a specific reaction must be prepared for each miRNA to be examined. Reverse transcription priming was achieved using looped primers. Life Technologies miRNA Reverse Transcription kit (Life Technologies, cat. no. 4366596) was used to convert the RNA to cDNA using the following composition:

1.5 $\mu$ l 10x Buffer  
0.19 $\mu$ l RNase inhibitor  
1 $\mu$ l Reverse Transcriptase  
0.15 $\mu$ l DNTPs  
4.16 $\mu$ l H<sub>2</sub>O  
5 $\mu$ l Diluted RNA  
3 $\mu$ l miRNA specific primers.

Samples were then run in a thermal cycler under the following conditions:

1. Incubation 16°C – 30 mins
2. Incubation 42°C - 30 mins
3. Inactivation 85°C – 5 mins
4. Hold 4°C -  $\infty$

#### **2.4.2.5 qRT-PCR for total RNA and miRNA.**

qRT-PCR was performed on each cDNA replicate in duplicate resulting in 4 technical replicates per biological sample. When custom TaqMan assays were required such as the pre-mRNA/pre-miRNA assays, 100bp of sequence (spanning intron/exon region for pre-mRNA) was submitted to Life Technologies custom design tool. For each assay a no template control was run to ensure specificity of the assay to target cDNA. 10 $\mu$ l reactions were performed on an applied biosystems Viia7 machine:

12.5 $\mu$ l TaqMan Universal PCR master mix no AmpErase UNG (Life Technologies, cat. no. 4324018)  
9.25 $\mu$ l H<sub>2</sub>O  
1.25 $\mu$ l TaqMan assay (Life technologies)  
2 $\mu$ l cDNA

The samples were run under following conditions:

- 1) Activation 95°C - 10 mins
- 2) Denature 95°C - 15 seconds
- 3) Extension 60°C - 60 seconds (Fluorescence captured after each extension step)

#### **2.4.2.6 qRT-PCR quantification.**

Normalisation was carried out using housekeeper genes, *Rp49* for Pre/mRNA/pre-miRNA reactions and *snoR442* for miRNA reactions. For miRNA reactions *snoR442* and *U27* was examined for suitability as housekeepers with *snoR442* proving the most consistent across the biological samples. Average Cts of the gene of interest was normalised to the appropriate housekeeper with the fold change being calculated as  $2^{(\text{average change in Cts between housekeeper and gene of interest in control sample} - \text{change in Cts between housekeeper and gene of interest in test samples})}$ .

#### **2.4.3 Western blotting.**

Western blotting was performed on samples containing 60 wing imaginal discs or 1 3<sup>rd</sup> instar larva (L3). Tubulin was used as an internal control. Tissue/larval samples were homogenised using a pestle in loading buffer (250mM Tris, 4% SDS, 10% Glycerol, 2%  $\beta$ -mercaptoethanol, 2% Protease cocktail inhibitors), heated at 100°C for 7 minutes and spun at max speed for 5 minutes. The supernatant was removed, 0.05% bromophenol blue was added and 15 $\mu$ l of sample was loaded onto the gel. 7% pre-cast Novex gels were used for all Western blotting which were run in Tris Acetate SDS running buffer at 150 volts for 1 hour.

Proteins were transferred to either Immobilon P transfer membrane (when using HRP conjugated secondary antibodies) or Immobilon F (when using fluorescent secondary antibodies). The transfer was conducted at 100 volts for 1 hour in a BioRad Mini Transblot Cell using in transfer buffer (25mM Tris, 190mM Glycine, % SDS, 20% Methanol, 800ml H<sub>2</sub>O). Following transfer the membrane was washed briefly in 0.1% PBS-Tween (PBS-T) and subsequently blocked for 1 hour at room temperature in blocking buffer (1 Sachet PBS, 0.1% Tween 20, 5% milk powder, 100ml H<sub>2</sub>O).

After blocking the membrane was briefly washed in wash buffer (1 Sachet PBS, 0.1% Tween 20, 0.5% milk powder, 500ml H<sub>2</sub>O) and incubated in primary antibody overnight at 4°C in PBS-T. The next morning the membrane was washed 4 times for 10 minutes in wash buffer and subsequently incubated in the appropriate secondary antibodies for 1 hour at room temperature in wash buffer. See the table below for antibodies used and their dilutions.

The membrane was then washed 4 times for 10 minutes in wash buffer followed by a single 10 minute wash in PBS-T. HRP conjugated antibody binding was detected using Amersham ECL detection reagents (GE Healthcare, cat. no. RPN2209). Relative quantification of bands was performed in ImageJ. Fluorescent antibody binding was detected directly using an Odyssey Fc Imager (LICOR). Quantification was achieved using Image Studio Lite version 3.1 (LICOR).

#### **2.4.4 Immunocytochemistry.**

Developmentally staged wing imaginal discs were dissected in PBS, mounted on Poly-L-lysine coated slides and fixed in 4% paraformaldehyde for 15 minutes. Discs were blocked in PBTA (PBS, 1% Foetal Calf Serum, 0.3% Tween) and incubated in primary antibody at 4°C overnight. Mounted discs were then incubated in secondary antibody diluted in PBTA for 2 hours at room temperature. Finally, discs were mounted in 10µl of vectorshield (containing DAPI) and sealed with nail varnish. See the table below for antibodies used and their dilutions.

Slides were imaged using confocal microscopy. Caspase staining images were taken with a Zeiss Axiovert confocal microscope equipped with a LSM520 Meta. Phosphohistone H3 images were taken using a Leica SP8 confocal microscope.

#### **2.4.5 Mitotic index calculation.**

Z-stacks of Phosphohistone H3 stained discs were taken using a Lecia SP8 confocal microscope with constant slice sizes. These Z-stacks were processed using Image J and the DeadEasy MitoGlia plug in. Default settings were used with a minimum threshold of 20.

<b>Antibody</b>	<b>Host</b>	<b>Dilution</b>	<b>Use</b>	<b>Source</b>
<i>Anti-dDis3</i>	Rabbit	1:1,500	Detect the level of Dis3 protein in <i>Drosophila</i> .	<i>Newbury Lab</i>
<i>Anti-dDis3L2</i>	Sheep	1:500-1:2,000	Detect the level of Dis3L2 protein in <i>Drosophila</i> (did not work).	Made during this PhD.
<i>Anti-hDis3L2</i>	Rabbit	1:500	Attempt to detect the level of Dis3L2 in <i>Drosophila</i> .	Sigma, cat. no. HPA0357497
<i>Anti-hDis3L2</i>	Rabbit	1:1000	Attempt to detect the level of Dis3L2 in <i>Drosophila</i> .	Thermo, cat. No. PA5-31273
<i>Anti-mouse HRP</i>	Goat	1:80,000	HRP conjugated antibody used to visualise mouse primary antibody binding.	Sigma, cat. no. A2304
<i>Anti-rabbit HRP</i>	Goat	1:80,000	HRP conjugated antibody used to visualise rabbit primary antibody binding.	Sigma, cat. no. A2304
<i>Anti-Sheep HRP</i>	Donkey	1:25,000-1:50,000	HRP conjugated antibody used to visualise sheep primary antibody binding.	Abcam, cat. no. 97125
<i>Anti-mouse 800</i>	Goat	1:20,000	Fluorescently labelled antibody used to visualise mouse primary antibody binding.	LICOR, cat. no. 926-32210
<i>Anti-rabbit 680</i>	Goat	1:20,000	Fluorescently labelled antibody used to visualise rabbit primary antibody binding.	LICOR, cat. no. 926-68071
<i>Anti-Phosphohistone H3</i>	Rabbit	1:300	Detect mitotic cells.	Cell Signalling, cat. no. 9701
<i>Anti-Caspase3</i>	Rabbit	1:400	Detect apoptotic cells.	Cell Signalling, cat. no. 9661
<i>Anti-rabbit Cy3</i>	Donkey	1:400	Fluorescently labelled antibody used to visualise rabbit primary antibody binding.	Jackson ImmunoResearch 715-165-151

## 2.4.6 Transformation of DH5 $\alpha$ with *dis3L2* cDNA.

### 2.4.6.1 Bacterial transformation.

Transformation of DH5 $\alpha$  cells with *dis3L2* cDNA was performed as described in sections 2.3.5.5-2.3.5.7. The only difference being the cDNA was removed from paper discs

### 2.4.6.2 Transformation diagnostics.

Potentially transformed DH5 $\alpha$  cultures were tested for the presence of the *dis3L2* cDNA using both restriction enzyme and PCR diagnostics. The restriction enzyme Hpa1 was used for the diagnostic digest and was incubated with DNA from each culture for 1 hour at 37°C. Digest products were run on a 1% agarose gel (section 2.4.1.4).

To confirm the restriction digest, PCR (section 2.4.1) was performed on DNA from each culture with specific primers spanning the site containing the extra 59bp in *dis3L2-RA* therefore allowing confirmation of the presence of *dis3L2* cDNA in addition to validation of the isoform. The PCR products were then sequenced (section 2.4.1.5) to ensure the isoform present and the lack of mutations.

## 2.5 Next Generation Sequencing.

### 2.5.1 Small RNA sequencing.

RNA extraction was performed as in section 2.4.2.2 from 60 wandering L3 wing imaginal discs for each genotype (Knockdown (*nub-GAL/+ ; UAS-Dis3<sup>RNAi</sup>/+*) and both parental controls ( *; nub-GAL4 ;* and *;; UAS-Dis3<sup>RNAi</sup>*) in duplicate). RNA concentration was quantified using a Nanodrop and its integrity was checked using an Agilent 2100 Bioanalyser using the peaks for the 28S and 18S rRNA. 1µg total RNA was sent to ARK genomics for sample preparation. Libraries were prepared using an Illumina TruSeq small RNA kit and size selected to specifically contain RNAs of 18-22bp in length. Samples were run across 2 lanes of an Illumina HiSeq 2500 producing between 12 and 22 million reads.

### 2.5.2 Analysis of small RNA sequencing.

#### 2.5.2.1 Identification of known miRNAs.

Analysis was completed by ARK Genomics. Adaptors were removed using Cutadapt. The remaining reads were mapped to miRBase release 19 using Bowtie2 (Langmead and Salzberg 2012). Mapped reads were normalised per million reads (RPKM).

#### 2.5.2.2 Identification of known and novel miRNAs using miRDeep2.

Fastq files were converted to fasta files using the FASTX toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). Reads were simultaneously clipped and mapped using the mapper.pl script included in miRDeep2 (Friedlander *et al.* 2008). Reads were mapped to the *Drosophila melanogaster* genome, version BDGP5 (Ensembl) (Settings:-c -j -k -l 17 -m). To identify novel miRNAs miRDeep2.pl was used. *Drosophila simulans*, *Drosophila sechellia*, *Drosophila erecta*, *Drosophila yakuba* and *Drosophila ananassae* were used as similar species to aid the identification of novel miRNAs. This was completed by Dr. Sarah Smalley.

Potential novel miRNAs were selected for further analysis if they were found in at least 5 of the 6 replicates, including at least one of each genotype. The only selected potential novel miRNA was validated using custom TaqMan assays to the identified sequence

(CACAU CGAUGUUUUUCCACCUC). qRT-PCR was completed with the custom assay as above. The secondary structure of the novel miRNA was predicted using RNAfold WebServer (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) with default settings.

### 2.5.3 RNA sequencing.

Wing imaginal discs were dissected from wandering L3 larval for RNA extraction as previously outlined (section 2.2.5). RNA was extracted from samples of 60 wing imaginal discs for each genotype (Knockdown (*UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4+*) and both parental controls ( ; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *69B-GAL4*) in duplicate) as in section 2.4.2.2. RNA concentration and purity was examined using a NanoDrop 1000 spectrophotometer whilst its integrity was tested on an Agilent Bioanalyser using a nanochip. 3µg of total RNA was sent to Oxford Gene Technology (OGT) for library preparation using an Illumina TruSeq kit. Libraries were poly(A) selected and run on a single lane of an Illumina HiSeq 2500 sequencer resulting in read coverage of between 13 and 15 million reads per sample.

### 5.5.4 Analysis of RNA sequencing.

Raw reads (Fastq files) from the RNA sequencing were received from OGT and adaptors were removed using the algorithm Scythe v0.993b (<https://github.com/vsbuffalo/scythe>) with the probability of their occurrence provided in each run ( e.g -p 0.00658). The adaptors for each sample were identified using FastQC v0.11.2 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) which identifies over represented sequenced within the Fastq files together with their frequencies. Following adaptor removal reads were filtered for quality using the algorithm Sickle v1.29 (<https://github.com/najoshi/sickle>) using default Sanger preferences. Sickle is a 'sliding window' algorithm which removes low quality bases one at a time. A minimum read length of 15 bases was used.

Following the above filtering the remaining reads were mapped to the *Drosophila melanogaster* genome, Flybase release 6.01 (St Pierre *et al.* 2014) using the algorithm TopHat2 v2.0.12 (Kim *et al.* 2013). All 4 *Drosophila* chromosomes were used including both ribosomal and mitochondrial genes. The index file to which the reads were aligned was built by Dr. Chris Jones using Bowtie v2.2.3 (Langmead and Salzberg 2012). Default parameters were used apart from the minimum and maximum intron length was altered to relate to *Drosophila* intron length (-i 20, -l 150000). Reads that mapped to >20 locations within the genome were removed during the alignment.

Aligned reads were subsequently assembled using Cufflinks (Trapnell *et al.* 2012) using corrections for fragment bias (-b) and multiple reads (-u). In addition a normalisation step was used to improve the robustness in the determination of the expression levels of lowly expressed genes (-N). Assembled transcripts were merged into a single file using Cuffmerge which was then quantified using Cuffquant. Finally, differential expression between samples was assessed using Cuffdiff resulting in a normalised read count (FRKM) for each transcript. FPKM or Fragments Per Kilobase of transcript per Million mapped reads is a normalised expression value which takes into account the total number of reads each sample initially received together with normalising for transcript length which can vary greatly and therefore would affect the probability of a read mapping. See Chapter 5 for a more in detail description of the analysis pipeline.

## **2.6 Statistical analysis.**

All statistical analyses were performed in GraphPad Prism 6 or using R. All data analysed were compatible with parametric tests. Two-sided two-sample t-tests were used to compare the means of single test groups to single control groups. If multiple comparisons were required, a one-way ANOVA was performed with a post-test to compare the means of each possible pair of samples. To assess the differences between global transcript expression profiles a Kolmogorov-Smirnoff test was performed in R. This is a specific test designed to assess statistical differences between kernel density plots (expression profiles). Finally, to assess the differences in starvation survival two different statistical tests were used. A Mantel-Cox (Log-Rank) test was used to assess the starvation survival as it treats all test samples equally throughout the time course. This is in contrast to the other possible test, the Gehan-Breslow-Wilcoxon test which requires that one group have a consistently higher risk than the other; a fact that was unknown within this experiment.

# Chapter 3: The role of Dis3 in the development of and regulation of miRNA stability in the wing imaginal discs.

## 3.1 Introduction.

Post-transcriptional control is one of many layers of regulation of gene expression within the cell. Within this, RNA stability is known to play a critical role in determining the period of time an RNA species is present in the cell to elicit its function; such as an mRNA to be translated into protein or a miRNA to bind and repress its targets. The stability of a cytoplasmic RNA is dependent upon its degradation which can occur from either the 5' end by Pacman or from the 3' end by a family of related exoribonucleases; the Dis3 family.

The Newbury lab has previously focused on the activity of the 5'-3' exoribonuclease Pacman identifying roles in wound healing, fertility, apoptosis and developmental timing (Grima *et al.* 2008; Zabolotskaya *et al.* 2008; Jones *et al.* 2013; Waldron *et al.* 2015; Jones *et al.* 2016). Therefore the aim of this work was to begin to investigate the enzymes involved in the 3'-5' RNA degradation pathway with this chapter focusing on Dis3. Dis3 is a highly conserved processive 3'-5' exoribonuclease which associates with the exosome core to provide its catalytic activity (Mitchell *et al.* 1997). The cellular localisation of Dis3 varies between organisms. In *Drosophila melanogaster* Dis3 associates with both the cytoplasmic and nuclear exosome (Graham *et al.* 2006) whilst in humans its localisation is mainly nuclear (Tomecki *et al.* 2010). This differential localisation is due to the presence of an additional Dis3 paralogue in the human genome, Dis3L1, which has been shown to be localised to the cytoplasm and associate with the cytoplasmic exosome (Staals *et al.* 2010). It is likely that in *Drosophila* Dis3 performs the role Dis3L1 is responsible for in humans.

Studies in a variety of organisms have shown that correct growth and cell cycle progression is dependent upon Dis3 activity. Both *S. cerevisiae* and *S. pombe* *dis3* mutants also display cell

cycle defects (Ohkura *et al.* 1988; Murakami *et al.* 2007; Smith *et al.* 2011). These cell cycle defects could potentially be the cause of the observed growth deficiencies, due to either cell cycle arrest preventing proliferation and/or resulting in apoptosis due to prolonged arrest. The cell cycle deficiencies are concordant with work in *S. cerevisiae* that has shown loss of exoribonuclease activity or mutation in the CR3 motif to result in a slow growth phenotypes (Schaeffer *et al.* 2012). Complementary work in *Drosophila* has also shown growth and developmental delay when Dis3 was knocked down throughout the fly. Loss of Dis3 in these flies resulted in early larval size to reduce by 400%, a severe growth deficiency which ultimately resulted in lethality at the 2<sup>nd</sup> instar larval stage (L2) (Hou *et al.* 2012). Complementary work in *Drosophila* embryo derived S2 cells showed that knockdown of Dis3 resulted in a 50% reduction in proliferation (Kiss and Andrusis 2010). Taken together these corroborating studies in both yeast and *Drosophila* indicate a key requirement for Dis3 to maintain correct growth.

Although these studies have identified various functions of Dis3, very few have explored its role within a developmental context. This is due to the fact that most of the founding work on Dis3 has been performed in yeast or human tissue culture cells both of which have little developmental potential. Therefore, this chapter aimed to extensively characterise the phenotypes caused when *dis3* is lost in a developing organism.

In addition to the phenotype characterisation this chapter also asks if Dis3, as a member of the cytoplasmic exosome, functions to regulate the stability of mature miRNAs. Previous studies have implicated Dis3 in the processing of vast range of targets such as; 5.8S rRNA (Mitchell *et al.* 1997), pre-tRNA, pre-mRNA (Gudipati *et al.* 2012) together with other non-coding RNA species such as snRNAs, snoRNAs, and cryptic unstable transcripts (CUTs) (Preker *et al.* 2008; Schneider *et al.* 2012; Tomecki *et al.* 2014). Additional studies have implicated Dis3 and the exosome in the biogenesis of specific miRNAs. Xue *et al.* have shown a requirement for the exosome in the biogenesis of miRNA-like RNAs in *Neurospora crassa* (Xue *et al.* 2012).

Complementary work has shown a similar role for the exosome in the biogenesis of miRNAs in *Drosophila* (Flynt *et al.* 2010) in addition to a responsibility for pre-miRNA clearance in mouse embryonic fibroblasts (Liu *et al.* 2014). However, a role for Dis3 in mature miRNA regulation has thus far been overlooked.

Dis3 has clear clinical importance as it has been implicated in Multiple Myeloma (MM) and Acute Myeloid Leukaemia (AML) with mutations in *DIS3* observed in 11% (Chapman *et al.* 2011) and 4% (Ding *et al.* 2012) of cases respectively. With respect to MM the mutations have been shown to result in loss of function (Tomecki *et al.* 2014). This is intriguing as the previous work in yeast and *Drosophila* has shown loss of function mutations or Dis3 knockdown to result in slow growth phenotypes. This begs the question as to how a loss of function mutation, in the case of MM, results in overgrowth. This is a question yet to be answered and one that will presumably depend upon the simultaneous mutations of other key tumour suppressors.

Due to the absence of a mutant and the lack of a stock containing an appropriately located P-element, RNA interference, controlled by the *GAL4-UAS* system, was used. Using the *GAL4-UAS* system allowed the expression of the RNAi hairpin to be specifically restricted to the wing imaginal disc, which was used as a specific developmental model (discussed in introduction section 1.7.3). This chapter will explore the requirement of Dis3 for the development of the *Drosophila* wing followed by an in depth approach to investigating if Dis3 plays a role in the regulation of miRNA stability.

## 3.2 Project background and aims.

The Newbury lab previously investigated the expression profile of Dis3 throughout *Drosophila* development and identified Dis3 expression throughout embryogenesis and early larval stages (Cairrão *et al.* 2005). However, there was no further work into the role Dis3 plays during the regulation of these stages or within specific tissues. Preliminary work in the lab showed that Dis3 is essential for *Drosophila* viability as ubiquitous knockdown results in complete lethality prior to the third instar larval stage of development (L3) (Baldock 2012). This is similar to a previous study by Hou *et al* who also showed that ubiquitous knockdown of Dis3 in *Drosophila* results in lethality (Hou *et al.* 2012). Additionally, previous work looking at Dis3 within the wing imaginal discs has shown that knockdown of Dis3 specifically in the dorsal region of the wing pouch results in severe wing malformation (Baldock 2012). Therefore this work aimed to build upon and look further into the role of Dis3 within a developing tissue using the wing imaginal disc as a model system.

The main aims of the work presented in this chapter were twofold:

1. Investigate the role of Dis3 within the development of the wing imaginal disc.
  - Corroborate previous findings in *Drosophila*.
  - Characterise any phenotypes observed following Dis3 knockdown.
  - Identify a potential mechanism or pathway through which Dis3 elicits its function
2. Determine whether Dis3 plays a role in the maintenance of miRNA stability.
  - Conduct a miRNA-sequencing experiment on the wing imaginal discs
  - Determine if differentially expressed miRNAs are directly or indirectly affected by the loss of Dis3.
  - If suitable, suggest a mechanism by which Dis3 regulates miRNA stability

### 3.3 Dis3 is required for *Drosophila melanogaster* viability.

In the absence of a Dis3 mutant along with there being no suitable P-element insertion stock to use to create mutations, RNA interference was used to assess the developmental role of Dis3 in a reverse genetics approach. The *GAL4-UAS* system is a powerful genetic tool used in *Drosophila* which gives both spatial and temporal control over the expression of specific constructs. Here, the *GAL4-UAS* system was used to drive RNA interference, in the form of a hairpin RNA, specific to Dis3 (See methods section 2.3.1).

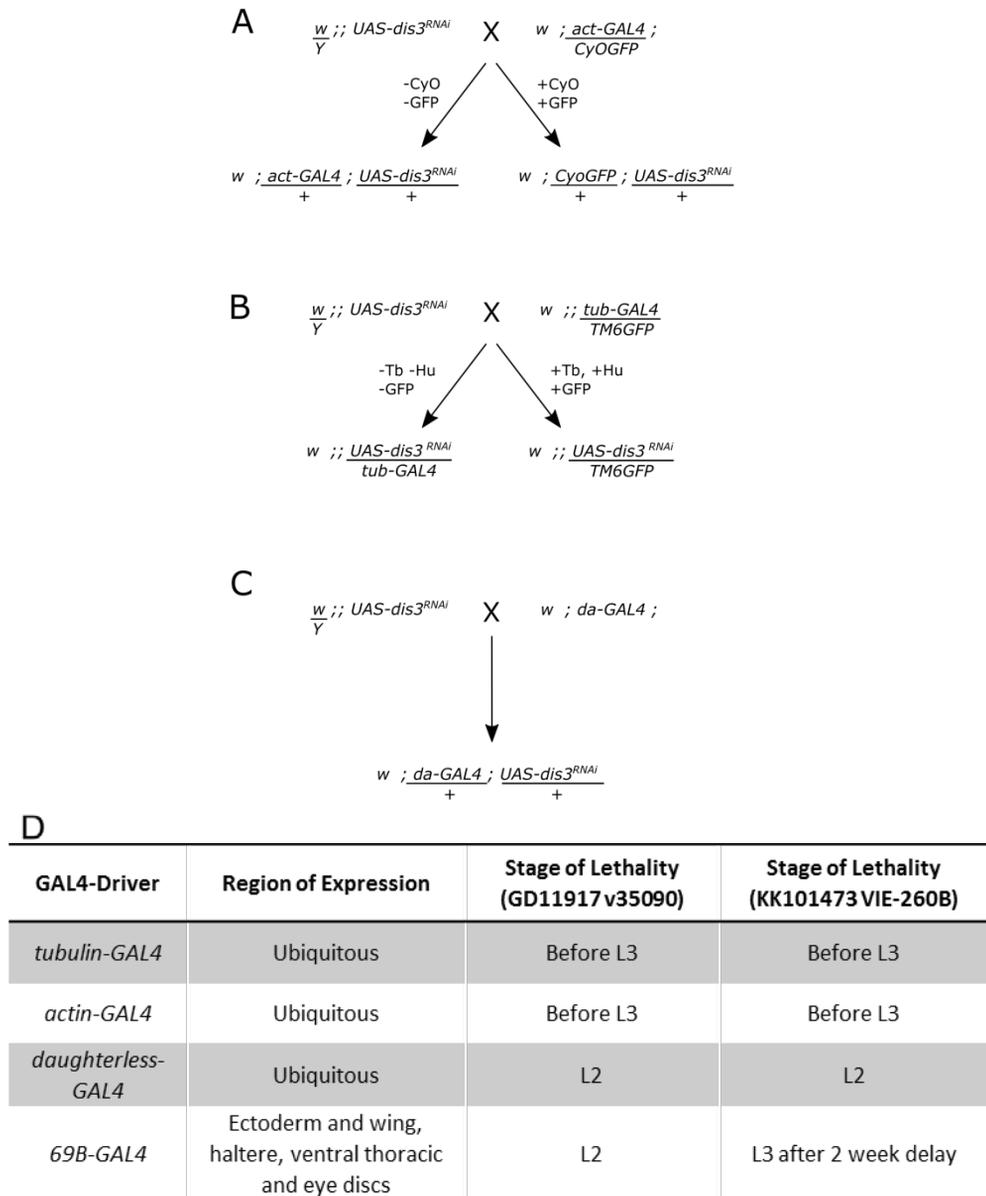
To validate the findings by Hou *et. al.* 2012 and those previously obtained in the Newbury lab *UAS-dis3<sup>RNAi</sup>* was crossed to a collection of ubiquitous *GAL4* drivers. *UAS-dis3<sup>RNAi</sup>* males were crossed to virgin females of either of the two strong ubiquitous *GAL4* drivers used; *actin(act)-GAL4/CyOGFP* and *tubulin(tub)-GAL4/TM6GFP* (Figure 3.1A/B). In both of these crosses 50% of the progeny would have Dis3 knocked down whilst the other half would retain wild-type levels of Dis3 due to the absence of the *GAL4* drivers. In each case the knockdown progeny were identifiable by the lack of the marker carrying balancer chromosome. *act-GAL4/+ ; UAS-dis3<sup>RNAi</sup>/+* progeny were identifiable at the larval and adult stages by the absence of GFP and CyO respectively. *tub-GAL4/UAS-dis3<sup>RNAi</sup>* progeny were identifiable at the larval and adult stages by the absence of GFP/tubby and humeral respectively.

For both the strong *GAL4* drivers knockdown of Dis3 resulted in complete lethality with no knockdown larvae observed (Figure 3.1D). An additional, slightly weaker, ubiquitous *GAL4* driver (*daughterless(da)-GAL4*) is homozygous for the *GAL4* insertion and therefore when virgins were crossed to *UAS-dis3<sup>RNAi</sup>* males 100% of the progeny had Dis3 knocked down (Figure 3.1C). *da-GAL4* driven knockdown of Dis3 resulted in 100% lethality at the 2<sup>nd</sup> instar larval stage (L2). L2 larvae were observed to crawl out of the media and subsequently die on the side of the vials within 48 hours of emerging. No further development towards the later 3<sup>rd</sup> instar larval (L3) stage was observed during this time. Two *UAS-dis3<sup>RNAi</sup>* lines were used with each *GAL4* driver with each showing identical phenotypes (Figure 3.1D).

### **3.4 Loss of Dis3 in the wing imaginal disc results in severe phenotypic defects.**

#### **3.4.1 Dis3 is required throughout the wing disc for viability.**

Due to ubiquitous knockdown of Dis3 resulting in lethality the wing imaginal discs were used as a specific model to investigate the developmental role of Dis3. The wing imaginal discs contain all the information required for the correct development of the wing and part of the thorax. The critical pathways required for correct development of the wing discs, such as the Wingless and Notch pathways, are well characterised and extremely well conserved making them an ideal model system. To restrict the knockdown of Dis3 to the wing imaginal discs the *GAL4-UAS* system was again used with wing disc specific *GAL4* drivers. The *69B-GAL4* driver expresses throughout the wing imaginal disc in addition to expression within the ventral-thoracic disc and the ectoderm (Brand 1997) (methods Figure 2.3). *UAS-dis3<sup>RNAi</sup>* males were crossed to the homozygous *69B-GAL4* virgins resulting in all progeny having Dis3 knocked down throughout the wing imaginal disc. This resulted in lethality similar to that observed for the *da-GAL4* driven knockdown with all knockdown progeny showing lethality at L2 (Figure 3.1D). Knockdown of Dis3 using the second *UAS-dis3<sup>RNAi</sup>* line also results in lethality but at L3 rather than during L2. *UAS-dis3<sup>RNAi</sup>/69B-GAL4* L3 larvae die after a 2 week delay where no further growth is observed. Interestingly, when dissected these L3 larvae showed minimal disc development indicating Dis3 is required for the development of imaginal discs.



**Figure 3.1 Cross schemes to achieve ubiquitous knockdown of Dis3.** (A) Cross scheme used to drive knockdown of Dis3 throughout the *actin* expression domain. *UAS-dis3<sup>RNAi</sup>* males were crossed to *act-GAL4* females which were balanced over CyOGFP. Knockdown adult progeny (*; act-GAL4/+ ; UAS-dis3<sup>RNAi</sup>/+*) were identified by the absence of curly wings (CyO) whilst the knockdown larvae were detected by the absence of GFP. (B) Cross scheme used to drive knockdown of Dis3 throughout the *tubulin* expression domain. *UAS-dis3<sup>RNAi</sup>* males were crossed to *tub-GAL4* females which were balanced over TM6GFP. Knockdown adult progeny (*UAS-dis3<sup>RNAi</sup>/tub-GAL4*) were identified by the absence of the humeral phenotype whilst the knockdown larvae were detected by the absence of the tubby phenotype. (C) Cross scheme used to drive knockdown of Dis3 throughout the *daughterless* expression domain. *UAS-dis3<sup>RNAi</sup>* males were crossed to *da-GAL4* females in a cross where 100% of the progeny would carry Dis3 knockdown. (D) Summary of the phenotypes observed when Dis3 was knocked down using the 3 ubiquitous GAL4 drivers. GD11917 v35090 and KK101473 VIE-260B refer to the two different *UAS-dis3<sup>RNAi</sup>* lines used.

### 3.4.2 Dis3 knockdown within the wing pouch results in an absence of wing development.

Due to the lethality induced by Dis3 knockdown throughout the wing imaginal disc, knockdown was further restricted to the wing pouch region of the wing disc using *nubbin(nub)-GAL4* (methods Figure 2.3). Constraining the knockdown to the wing pouch should have allowed the development of viable adults in addition to the presence of the discs themselves in the L3 larvae. Using the *nub-GAL4* driver was therefore the optimal method to investigate the role of Dis3 in wing and wing disc development. To ensure the RNAi was working efficiently in the wing imaginal discs, they were dissected from control and knockdown wandering L3 larvae and a western blot was performed. This revealed that the Dis3 knockdown discs contained only 19.7% the amount of Dis3 observed in the parental control discs (Figure 3.2C). Knockdown of Dis3 within the wing pouch did indeed result in viable adult progeny, however, they presented with a drastic 'no wing' phenotype where there is a complete absence of wing development (Figure 3.2B).

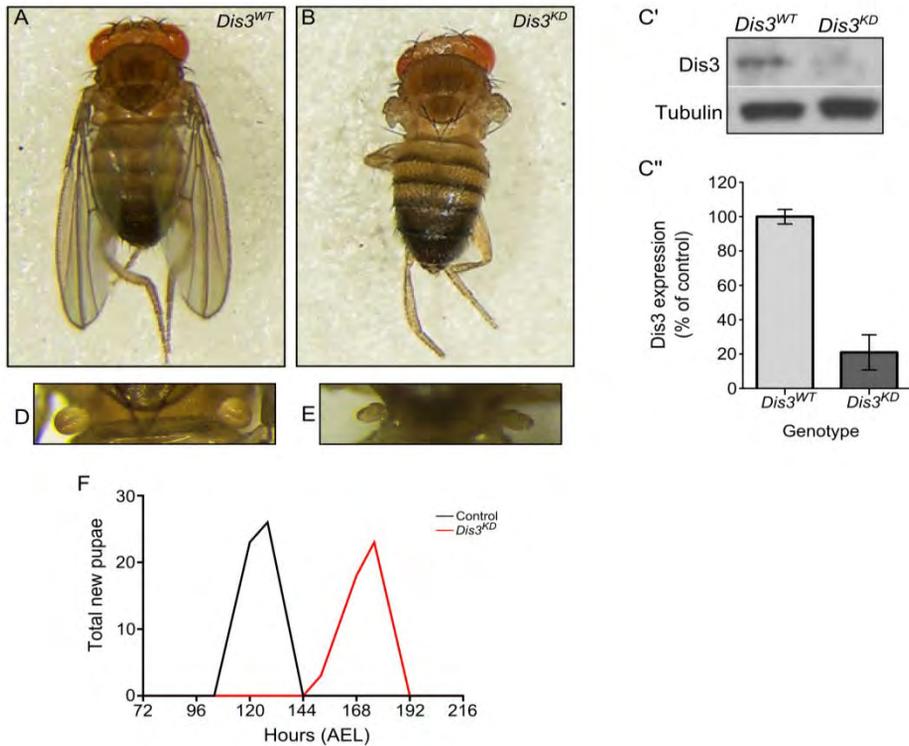
The *nub-GAL4* stock also contained a *UAS-dicer2 (UAS-dcr2)* insertion which would be expected to enhance the activity of the RNAi. To ensure the presence of the *UAS-dcr2* was not contributing to the observed phenotype *nub-GAL4* males were crossed to *UAS-dis3<sup>RNAi</sup>* virgin females (Figure 3.3B). As the *UAS-dcr2* is on X this would ensure that all resulting knockdown males would not contain the additional insertion. These males also showed the no wing phenotype with 100% penetrance showing the activity of the *UAS-dcr2* construct was not contributing to the phenotype (Figure 3.3C).

In addition to the lack of wing development, abnormal haltere development was also observed in the *nub-GAL4* driven Dis3 knockdown flies which indicates an additional expression domain of *nub-GAL4* (Figure 3.2E). To test this *UAS-GFP* was driven by *nub-GAL4* which showed expression in the haltere imaginal discs; this is also consistent with other work showing the

halter disc as an expression domain for *nub-GAL4* (Estella *et al.* 2003). This suggests that Dis3 is not only critical for the development of the wing but also the haltere.

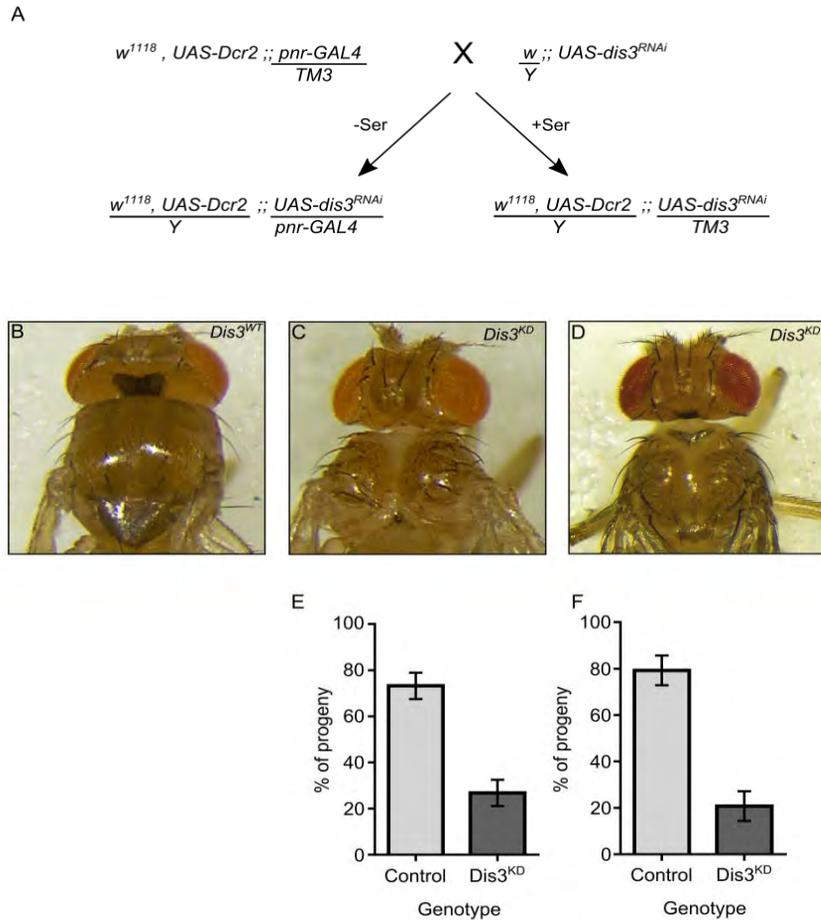
### **3.4.3 Dis3 is required for the correct development of the thorax as well as the wing.**

To investigate if Dis3 is only required for the development the wing and haltere it was knocked down in a region of the wing disc destined to form part of the thorax. *Pannier(pnr)-GAL4* expression is restricted to the tip of the wing imaginal disc and was therefore used to drive the knockdown of Dis3. Homozygous *pnr-GAL4* flies are not viable therefore heterozygous *pnr-GAL4/TM3* virgin females were crossed to *UAS-dis3<sup>RNAi</sup>* males with knockdown progeny being identified by the absence of the wing marker serrate (Figure 3.4A). Non-serrate knockdown progeny showed a severe cleft thorax phenotype with 100% penetrance when either *UAS-dis3<sup>RNAi</sup>* line was used (Figure 3.4C/D). In addition to the cleft thorax phenotype it was also apparent that knockdown of Dis3 within the *pannier* expression domain resulted in partial pupal lethality. Of all progeny eclosing only 20% were of knockdown genotype (Figure 3.4E/F) where as one would expect a 50:50 ratio thus indicating an element of lethality. Taken together these data indicate that Dis3 is not only required for the development of the wing blade but is also required for development of all other tissues tested. This is consistent with previous data showing that Dis3 is required throughout the organism for viability indicating a global requirement for its activity.



**Figure 3.2: Knockdown of Dis3 in the wing imaginal disc results in severe developmental phenotypes.** Knockdown of Dis3 in the wing pouch of the wing imaginal disc results in a "no wing" phenotype at 100% penetrance (*Dis3<sup>KD</sup>*) (B) when compared to a *UAS-dis3<sup>RNAi</sup>* parental control male fly (*Dis3<sup>WT</sup>*) (A). (C') Knockdown of Dis3 protein in *nub-GAL4/+ ; UAS-dis3<sup>RNAi</sup>/+* wing imaginal discs by Western blotting. *Dis3<sup>WT</sup>* genotype = ; *nub-Gal4*; (C'') Dis3 expression is knocked down by 80% compared to parental controls in wing imaginal discs when *UAS-dis3<sup>RNAi</sup>* is driven by *nub-GAL4*. '*Dis3<sup>WT</sup>*' includes parental genotypes ; *nub-Gal4* ; and ; *UAS-dis3<sup>RNAi</sup>*.  $n \geq 3$ ,  $p = 0.0009$ , error bars represent standard error. (D and E) *nub-GAL4/+ ; UAS-dis3<sup>RNAi</sup>/+* flies also show deficient haltere development (E) (100% penetrance) compared to parental control flies (D). (F) Dis3 knockdown progeny (*nub-GAL4/+ ; UAS-dis3<sup>RNAi</sup>/+*) are delayed in larval development by 40 hours (red) when compared to parental controls (black – *UAS-dis3<sup>RNAi</sup>* and *nub-GAL4*).  $n \geq 23$ .





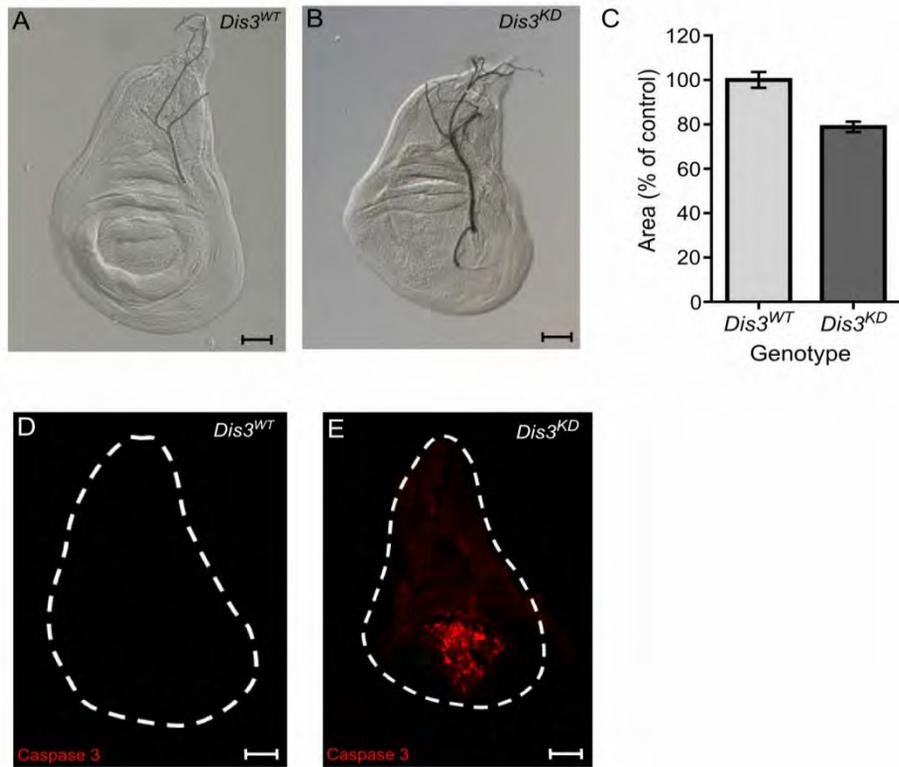
**Figure 3.4: Dis3 knockdown in the thoracic region of the wing imaginal disc results in a severe cleft thorax phenotype. (A)** Cross scheme used to drive the knockdown of Dis3 throughout the *pannier* expression domain. *UAS-dis3<sup>RNAi</sup>* males were crossed to *pnr-GAL4* females which were balanced over TM3. Knockdown adult progeny (*; UAS-dis3<sup>RNAi</sup>/pnr-GAL4*) were identified by the absence of the serrate phenotype (TM3). **(B-D)** Knockdown of Dis3 in the thoracic region of the wing imaginal disc using either *UAS-dis3<sup>RNAi</sup>* line (v35090 **(C)** or VIE260B **(D)**) resulted in a strong cleft thorax phenotype with 100% penetrance. Comparatively no cleft thorax was observed in the parental control lines **(B)**. **(E-F)** Knockdown of Dis3 in the *pannier* domain also resulted in an element of lethality as only ~25% of the eclosing progeny were of the knockdown genotype. If no lethality had occurred a 50:50 ratio would have been expected. This was consistent between both the v35090 **(E)** and VIE260B **(F)** *UAS-dis3<sup>RNAi</sup>* lines. Error bars represent 95% confidence limits, n=329 and 332 flies for **E** and **F** respectively.

### **3.5 Dis3 knockdown results in smaller wing imaginal discs caused by increased apoptosis.**

Due to the absence of wing development the wing imaginal discs themselves were examined. Due to the lethality and absence of disc development using ubiquitous or whole wing disc drivers Dis3 was again knocked down using the *nub-GAL4* driver. Wing imaginal discs were dissected from knockdown and parental control late L3 larvae and measured. Parental control stocks were used as controls due to them containing the most similar genetic background to the knockdown progeny. Larvae were staged as late L3 through the addition of 0.05% bromophenol blue to the fly food which allows visualisation of the food within the larval gut. Late L3 larvae clear their gut just prior to pupation and therefore those larvae with an absence of blue food within the gut were selected for dissection. Late L3 wing imaginal discs were significantly smaller than parental control discs when Dis3 was knocked down with the Dis3-depleted discs 78% the size of parental controls (Figure 3.5A-C).

A reduction in the total area of the disc could result from decreased proliferation or increased apoptosis. The Dis3 knockdown discs appeared to be 'flatter' in the wing pouch region; this together with no wing phenotype suggested an apoptotic phenotype and was therefore the first avenue investigated. Using an antibody to the pro-apoptotic activated Caspase 3 allows staining of cells that are actively undergoing apoptosis. Larvae were again staged to late L3 using the bromophenol blue food and their discs were dissected and stained with anti-activated Caspase 3. In Dis3 knockdown discs there was a large amount of apoptosis occurring specifically in the wing pouch of the wing disc (Figure 3.5E) whereas minimal apoptosis was observed in parental discs (Figure 3.5D). The apoptosis was restricted specifically to the domain where Dis3 was knocked down and therefore it shows that the loss of Dis3 in the wing disc results in large scale apoptosis. The apoptosis is the likely cause of the reduction in disc area together with the eventual lack of wing development.

Previous work in the lab had observed that smaller wing imaginal discs in *pcm*<sup>14</sup> mutants resulted in a developmental delay during the larval stages. To investigate if a similar delay was induced upon Dis3 knockdown the *nub-GAL4* virgins were crossed to *UAS-dis3*<sup>RNAi</sup> males 2 days prior to the experimental start point. Flies were then tipped into new vials and left for 8 hours to lay (8 hour egg lay) to ensure all progeny were of a similar age. Developing progeny were then scored for the time it took to reach the pupal stage. A significant developmental delay is clearly shown in Figure 3.2F with knockdown larvae taking 40 hours longer to reach the pupal stage than parental controls. There was no difference in the developmental timing of the two parental lines and therefore they were grouped to give a control. The Dis3 knockdown induced developmental delay is likely to occur in an attempt to give the wing imaginal discs additional time to reach a critical size for pupation.



**Figure 3.5: Dis3 knockdown in the wing pouch of the wing imaginal discs results in small wing discs and extensive apoptosis. (A and B)** Dis3 knockdown (*nub-GAL4/+ ; UAS-dis3<sup>RNAi</sup>/+*) late L3 wing imaginal discs (**B**) are smaller than parental control (*Dis3<sup>WT</sup> = ;;UAS-dis3<sup>RNAi</sup>*) wing imaginal discs (**A**) (scale bar=50µm). (**C**) *nub-GAL4* driven Dis3 knockdown wing imaginal discs are 23% smaller than parental control discs (*Dis3<sup>WT</sup>* includes parental genotypes ; *nub-GAL4* ; and *;;UAS-dis3<sup>RNAi</sup>*).  $n \geq 21$ ,  $p < 0.0001$ , error bars represent standard error. (**D and E**) Knockdown of Dis3 in the wing pouch results in apoptosis specifically localised within the wing pouch region (**E**) compared parental control discs that show very little apoptotic activity (**D**).  $n \geq 26$ , scale bar=50µm.

### 3.6 Does Dis3 play a role in the regulation of miRNA stability?

Dis3 has been identified to be involved in the regulation of mRNAs and non coding RNAs such as (PROMTs), cryptic unstable transcripts (CUTs) and ribosomal RNAs (Mitchell *et al.* 1997; Wyers *et al.* 2005; Preker *et al.* 2008; Schneider *et al.* 2012). Additionally Dis3 has been shown to be involved during the maturation of miRNAs in both *Drosophila* (Flynt *et al.* 2010) and mouse embryonic fibroblasts (Liu *et al.* 2014). However, it is yet to be seen if Dis3 and the exosome have a role in the regulation of mature miRNA stability. Therefore to address this miRNA sequencing (miRNA-seq) was used in a global, unbiased approach.

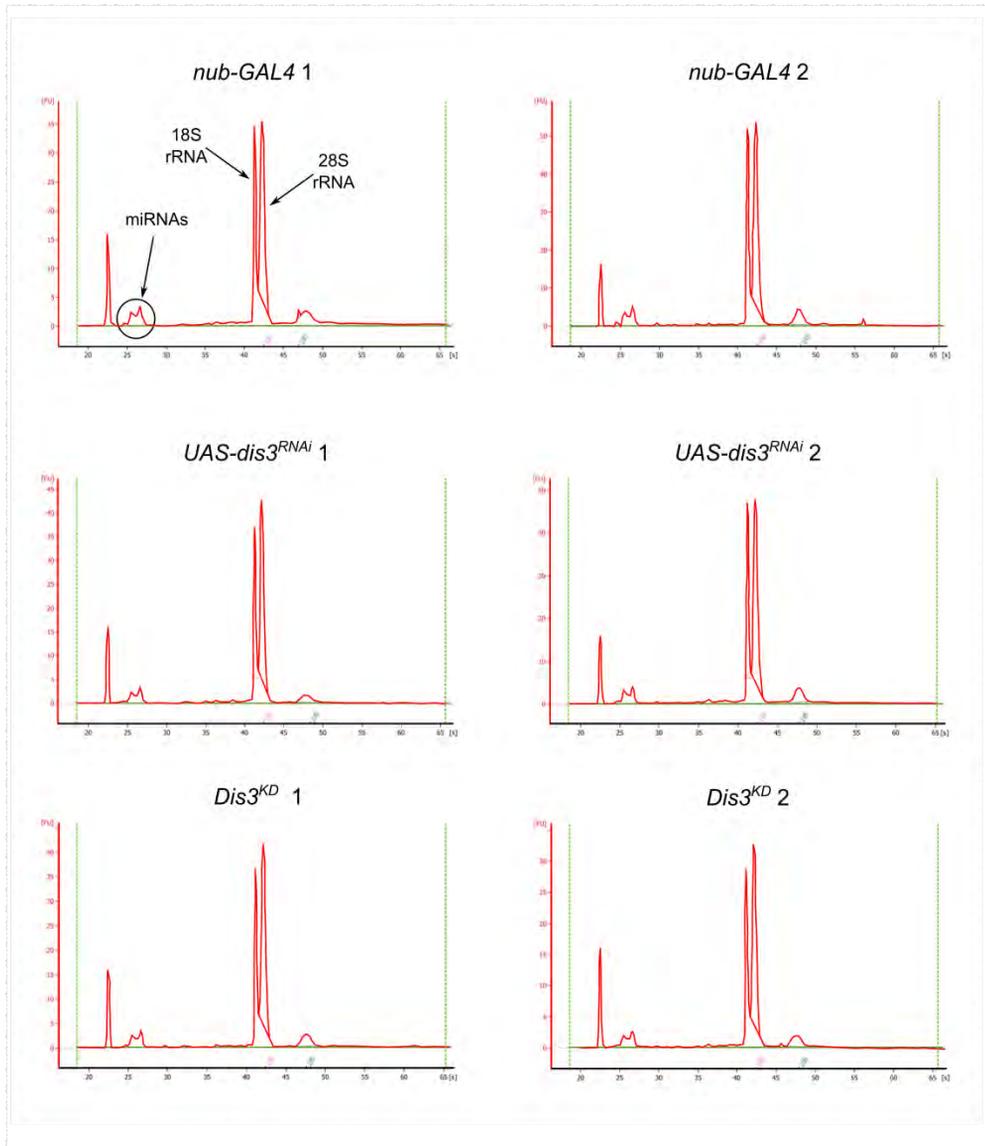
#### 3.6.1 Sample collection and miRNA-sequencing.

Due to the lethality induced when Dis3 is knocked down throughout the wing imaginal disc the *nub-GAL4* driver was used to direct Dis3 knockdown to the wing pouch of the disc. Due to RNA being extracted from the whole disc rather than the wing pouch alone this meant that half of the sample would have had wild type activity of Dis3. This would mean that any small changes would be 'diluted' and therefore were more unlikely to be detected. Despite this any large changes would still be detectable and were the most likely to be due to direct effects from the loss of Dis3.

To ensure all progeny carried the extra copy of *UAS-dcr2*, ; *nub-GAL4* ; virgin females were crossed to *UAS-dis3<sup>RNAi</sup>* males to produce the knockdown offspring. Wing imaginal discs were dissected from wandering L3 larvae from the knockdown progeny in addition to the two parental control lines ( *nub-GAL4* and *UAS-dis3<sup>RNAi</sup>*). For each sample RNA was extracted from 60 wing discs using the miRNeasy micro kit (Qiagen) which retains small RNAs such as miRNAs. To ensure the RNA was of high enough quality for RNA sequencing its integrity was examined on a Bioanalyser. The traces showed the RNA was indeed of high quality as two distinct peaks were observed representing the 18S and 28S rRNAs (Figure 3.6). In addition to the Bioanalyser the 260/280 and 260/230 ratios were used to ensure the RNA was pure. Both ratios were

>2.09 for all samples indicating that there was minimal contamination with organic solvents (260/230) or protein (260/280) which could cause issues with the subsequent sequencing.

1µg of total RNA of each sample was sent to ARK Genomics for library preparation and sequencing. Libraries were prepared using an Illumina TruSeq small RNA kit. The 3' and 5' adaptors were ligated to the RNA, both of which are designed to ligate to miRNAs using the 3' hydroxyl and 5' phosphate acquired during their biogenesis. Adaptor bound RNA was then reverse transcribed into single stranded cDNA using RT primers specific for adaptor sequences. Following reverse transcription a short round of PCR amplification was carried out to introduce the specific index sequences which act as a 'barcode' allowing the identification of each individual sample. Products were then gel purified and further size selected to ensure a specific insert size of 18-22nt (total size = ~147nt (18-22nt insert + adaptors + index)). The strict size selection of RNAs between 18-22nt in length was to ensure that only miRNAs were sequenced to enhance the read coverage. Resulting libraries were run across 2 lanes of a HighSeq 2500 sequencer obtaining between 12 and 22 million reads per sample (Figure 3.7).



**Figure 3.6: The RNA sent for miRNA-sequencing was of high quality.** Bioanalyser traces of the 6 samples sent to ARK Genomics for miRNA sequencing. Traces were performed on total RNA samples with two strong peaks representing the 18S and 28S rRNA (highlighted for *nub-Gal4* 1) indicating the samples were of high integrity. A peak for representing small RNAs such as miRNAs is also present (circled). Due to the difference in size in the 18S and 28S rRNAs between *Drosophila* and Humans the two peaks are much closer for *Drosophila* RNA samples. The software used did not contain a *Drosophila* calculation and therefore the generation of a RIN was not possible. *nub-GAL4* =  $w^{1118}$ , *UAS-dcr2*; *nub-GAL4* ; *UAS-dis3<sup>RNAi</sup>* =  $w^{1118}$  ; *UAS-dis3<sup>RNAi</sup>*, *Dis3<sup>KD</sup>* =  $w^{1118}$ , *UAS-dcr2*; *nub-GAL4/+*; *UAS-dis3<sup>RNAi</sup>/+*.

Bioinformatic analysis was performed by ARK Genomics. Adaptors were identified and removed using Cutadapt (Martin 2011). The remaining, high quality reads were mapped to miRBase release 20 using Bowtie 2. Reads were mapped to miRBase to identify any changes in expression to known miRNAs. Total mapped reads were normalised to the total number of reads each sample obtained to give a normalised read count (RPKM) which was used for the differential expression analysis. An initial observation was that the reads mapping to miRNAs were much lower than would be expected (~1.12% of total reads per sample) therefore FastQC was used to identify any overrepresented sequences. On average 83% of the total reads for each sample mapped to the 2S rRNA, a highly abundant RNA specific to *Drosophila* (Figure 3.7). It is 30nt in length and therefore should have been excluded during size selection; however it is clear that a large proportion was included in the final libraries. Despite this one to two hundred thousand reads still mapped to miRNAs therefore allowing sufficient coverage of the 426 known *Drosophila* miRNAs (miRBase release 20).

Sample	Sample ID	Total reads	miRNA mapped reads (%)	Unmapped (not miRNA)	2S rRNA mapped (%)
<i>nubbin-GAL4</i> 1	1024N0019	22690957	178566 (0.79%)	22436785	17865788 (78.7%)
<i>nubbin-GAL4</i> 2	1072N0020	17720964	218661 (1.23%)	17464795	15437784 (87.1%)
<i>UAS-dis3<sup>RNAi</sup></i> 1	1072N0021	12090092	115734 (0.96%)	11933336	10152688 (84.0%)
<i>UAS-dis3<sup>RNAi</sup></i> 2	1072N0022	12300174	162659 (1.32%)	12111417	10749590 (87.4%)
<i>Dis3<sup>KD</sup></i> 1	1072N0023	15697390	175180 (1.12%)	15483329	12844652 (81.8%)
<i>Dis3<sup>KD</sup></i> 2	1072N0024	13614220	176023 (1.29%)	13400007	10761736 (79.0%)

**Figure 3.7: Summary of miRNA-sequencing read details.** Table showing a breakdown of the total number of reads received by each sample. The percentage of reads mapping to mature miRNA sequences are presented together with the number of reads mapping to other RNA species. The majority of reads mapping to RNA species other than miRNAs mapped to the 2S rRNA sequence. The proportion of reads taken up by the 2S rRNA are presented for each sample.

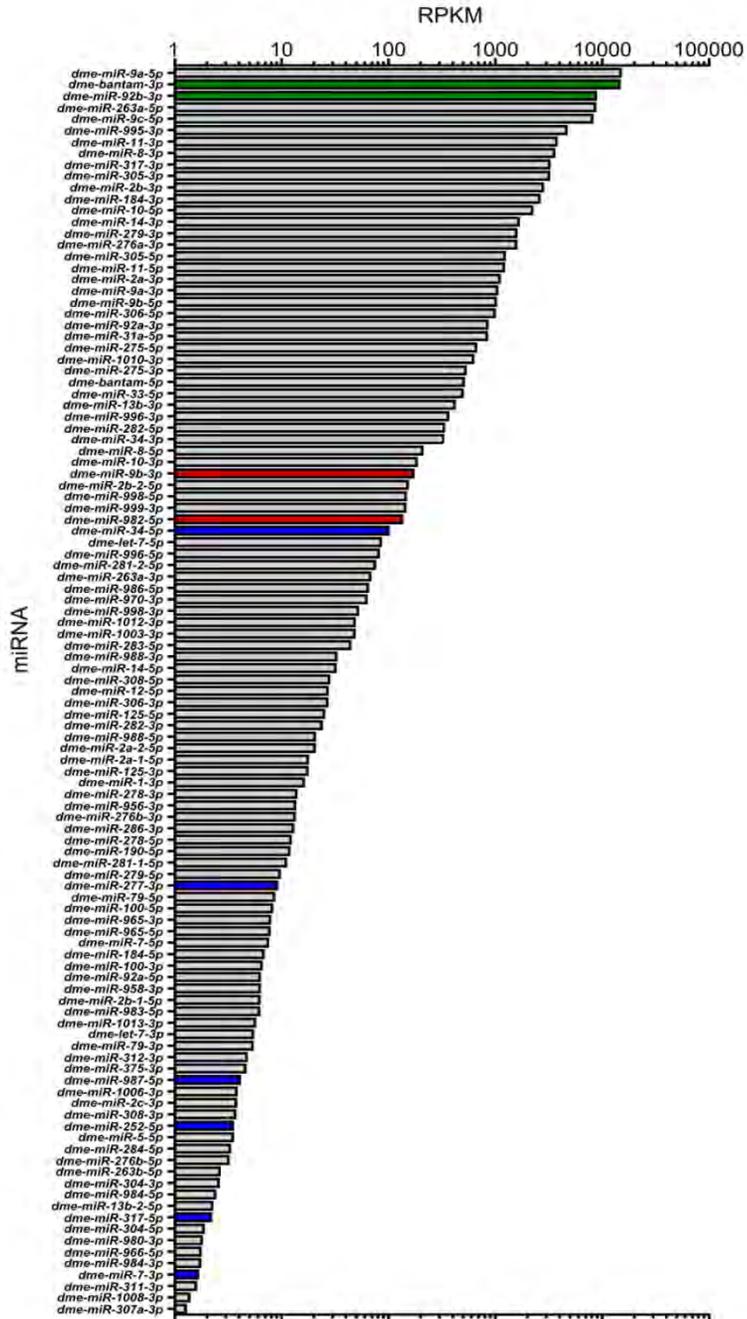
### 3.6.2 Global analysis of miRNA seq data.

For the first time the miRNAs that are expressed specifically in the wing imaginal discs themselves were identified as previous sequencing efforts have included all imaginal discs. The criteria used to determine whether a miRNA was expressed in the wing discs was that it had to be detected in all 6 samples; using this criteria 109 miRNAs were identified as expressed in the wing imaginal discs (Figure 3.8). These ranged from the most highly expressed (*miR-9a-5p*) with an average wild type normalised read count (RPKM) of 14865 to the lowest in expression (*miR-307a-3p*) with an average wild type RPKM of 1.26. To ensure the two parental controls were similar to each other the differential expression of each miRNA between parental replicates was analysed. Any miRNAs showing differential expression between the parental controls was excluded from any further downstream analysis as the changes in expression may be due to the nature of the *UAS-dis3<sup>RNAi</sup>* or *nub-GAL4* insertions.

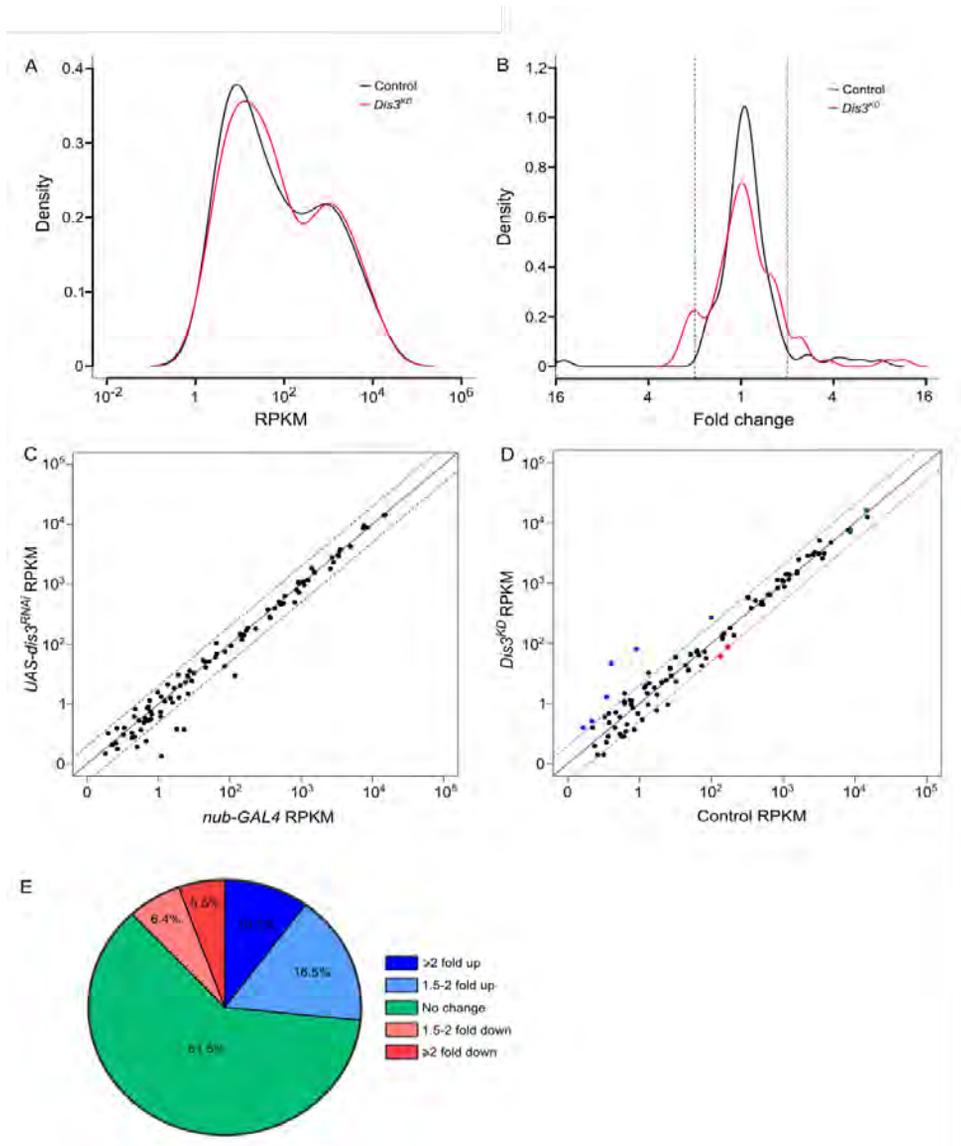
Of the 109 miRNAs, 101 (92.7%) showed differences of less than 2 fold between the two parental lines, 7 of the 8 that showed differences of more than 2 fold between parentals were discarded (points outside the dotted lines in Figure 3.9C). One miRNA, *miR-987-5p*, would not have passed the initial filtering for parental consistency as a large difference in expression was observed between the *nub-GAL4* and *UAS-dis3<sup>RNAi</sup>* parentals. However, it showed the largest fold change following Dis3 knockdown (average of 11.56 fold increase) versus all parental replicates and was therefore retained for further analysis. The overall similarity between the parental lines is consistent with a hierarchical clustering analysis showing the two parental controls are more similar to each other than they are to the knockdown replicates (Figure 3.10). Additionally the two biological replicates of each sample cluster tightly. The global difference between knockdown and parental clustering indicates an overall effect on miRNA expression following Dis3 knockdown.

The difference between the parental controls and Dis3 knockdown discs shown by the hierarchical clustering (Figure 3.10) indicates an overall difference in the miRNA expression

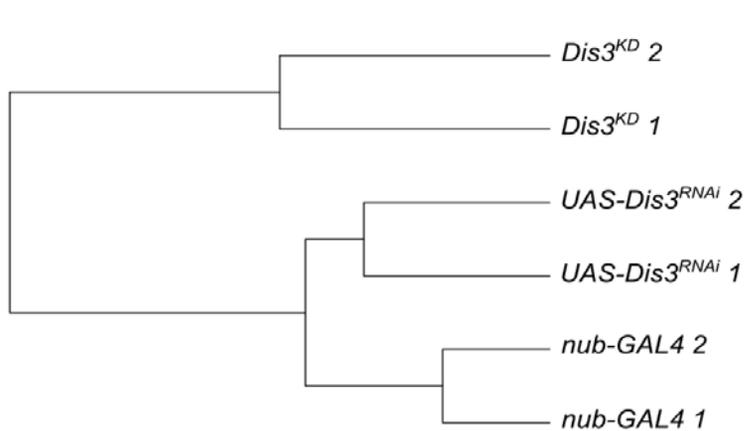
profiles. Kernel density plots were used to visualise the overall expression profiles of the samples. Due to the similarity between the two parental controls, as shown by the scatter plots (Figure 3.9C) and hierarchical clustering (Figure 3.10), they were grouped together to give a control group. There appears to be two main peaks of miRNA expression, those expressed at moderate levels (RPKM ~50) and those expressed at a high level (RPKM ~1000). Plotting the miRNA expression profiles of the control and knockdown discs shows a difference in expression distribution (Figure 3.9A). The major difference observed in the Dis3 knockdown discs appears to be a shift in the moderately expressed miRNAs to a level of slightly higher expression (shift and widening of the left hand peak in red). This suggests that in general a subset of miRNAs were at a higher level of expression following Dis3 knockdown; a trend that might be expected if Dis3 plays a role in the degradation of miRNAs. Additionally, when the distributions of the fold changes were plotted it was again clear that the knockdown of Dis3 resulted in differential expression of a set of miRNAs through the emergence of density peaks around +/- 2 fold change (Figure 3.9B).



**Figure 3.8: miRNA sequencing identifies 109 mature miRNAs in L3 wing imaginal discs.** 109 miRNAs were detected in L3 wing imaginal discs by miRNA-seq. They are presented in descending order of abundance (by normalised read counts (RPKM)) in control samples. *miR-9a-5p* was the most highly abundant whilst *miR-307a-3p* was the least abundant confidently detected miRNA. Those selected for further analysis are highlighted. Green = no change between conditions. Red = miRNAs that decreased following Dis3 knockdown. Blue = miRNAs that increased  $\geq 2$ -fold in expression following Dis3 knockdown.



**Figure 3.9 : Expression patterns of miRNAs in wing imaginal discs. (A)** A kernel density plot of the general distribution of the 109 detected miRNAs indicates 2 ‘peaks’ of miRNA expression. Knockdown of Dis3 results in a shift in the poorly expressed miRNAs (left hand peak) to a higher level of expression compared to the more highly expressed miRNAs which do not change (right hand peak). **(B)** Distribution of fold changes between both parents (black, Control) and grouped parents vs Dis3 knockdown (red, *Dis3<sup>KD</sup>*). Knockdown of Dis3 results in the emergence of peaks around +/- 2-fold change. Dotted vertical lines represent +/- 2-fold change. **(C)** Comparison of miRNA RPKMs between the two parental controls shows high similarity. Dotted lines show +/- 2-fold change, miRNAs lying outside these lines were excluded from further study. **(D)** Comparison of miRNA RPKM between grouped parental controls (Control) and Dis3 knockdown (*Dis3<sup>KD</sup>*) wing imaginal discs. Blue and red dotted lines represent +2-fold change and -2-fold change respectively. Selected miRNAs that increase or decrease in expression following Dis3 knockdown are coloured in blue and red respectively. Selected miRNAs that remain unchanged are highlighted in green. **(E)** 61.5% of miRNAs in the wing imaginal disc do not change in expression following the knockdown of Dis3. 26.6% of miRNAs increase in expression  $\geq 1.5$ -fold (10.1%  $\geq 2$ -fold). 11.9% of miRNAs decrease in expression  $\geq 1.5$ -fold (5.5%  $\geq 2$ -fold).



**Figure 3.10: Hierarchical clustering of replicates in miRNA-seq.** The overall expression profiles of the two parental controls (*UAS-dis3<sup>RNAi</sup>* and *nub-GAL4*) are more similar to each other than they are to the *Dis3<sup>KD</sup>* knockdown replicates (*Dis3<sup>KD</sup>* full genotype: *nub-Gal4/+* ; *UAS-dis3<sup>RNAi</sup>/+*). Additionally the biological replicates within each genotype are more similar to each other than any replicate from either of the two genotypes.

### 3.6.3 Analysis of Dis3-sensitive miRNAs.

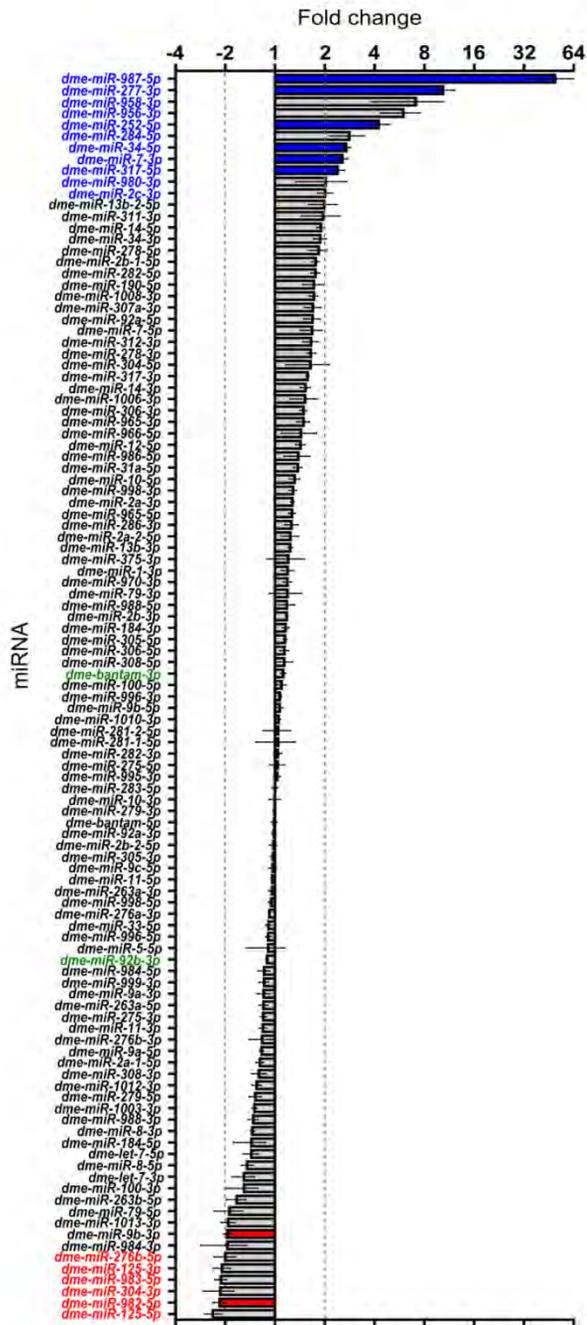
To identify those miRNAs that are sensitive to Dis3 knockdown and therefore may be regulated by Dis3 stringent selection criteria were set. To be selected as misexpressed a miRNA had to show a change in expression  $\geq 2$  fold when each knockdown replicate was compared to each control replicate (8 comparisons in total). Additionally, the miRNA had to be expressed in all 6 samples. Following this criteria 11 miRNAs that increase and 6 miRNAs that decrease in expression  $\geq 2$  fold in Dis3-depleted wing imaginal discs were identified (Figure 3.11). This in itself is interesting as it seems that only a small subset of miRNAs were sensitive to Dis3 regulation in the wing discs which is consistent with the kernel density plots. To confirm the expression changes 6 upregulated and 2 downregulated miRNAs in addition to 2 miRNAs that did not change in expression were validated by quantitative RT-PCR (qRT-PCR). The selected miRNAs are represented in Figures 3.9D, 3.11 and 3.12. The consistency between the replicates of the selected miRNAs can be seen in Figure 3.13A.

To identify a suitable normaliser for the validation by qRT-PCR *snoR442* and *U27* were trialled. Each assay was used on 10ng of the samples to be used for the validation. Of the two, *snoR442* showed an average Ct variation of 0.22Cts, whilst *U27* showed slightly greater variation of 0.33Cts. Due to showing greater consistency across all the samples *snoR442* was selected as a normalised for the validation experiments.

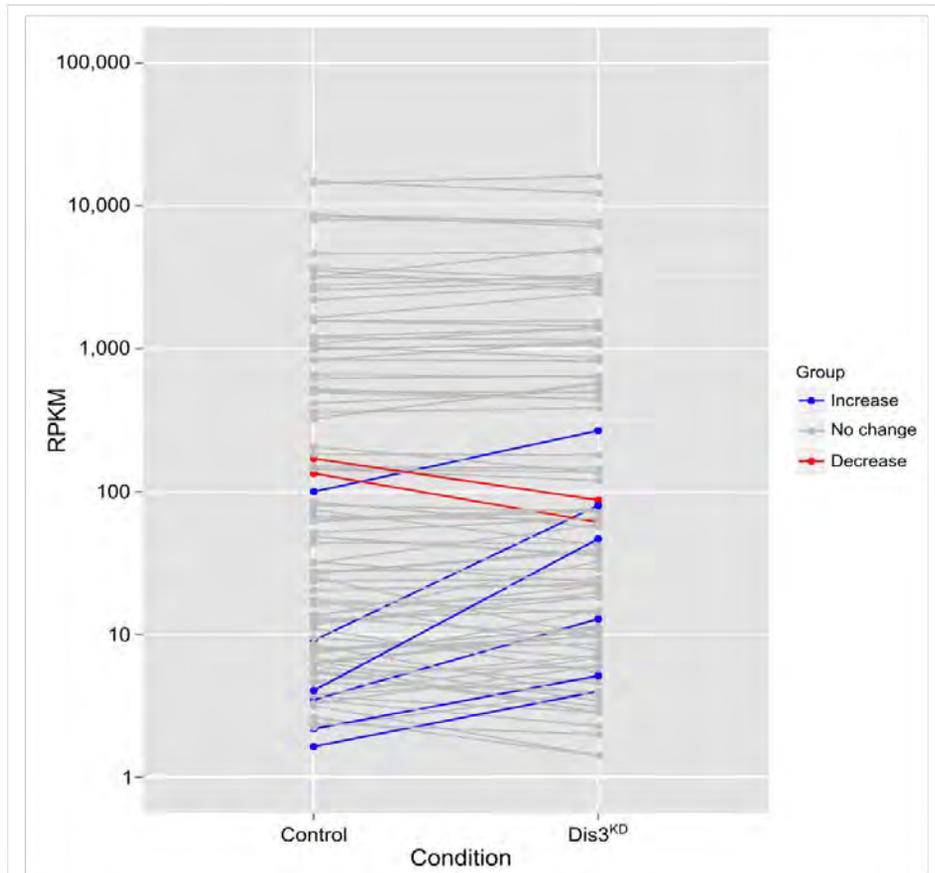
Using TaqMan chemistry and miRNA-specific looped RT primers all but one miRNA showed similar trends in their expression changes between the miRNA-seq and qRT-PCR. *miR-7-3p* was shown to increase 2.43 fold by the miRNA-seq, however, by qRT-PCR its expression is shown to decrease significantly (-1.49 fold) between control and knockdown discs (Figure 3.13B) indicating the differential expression in this case cannot be trusted. The other 5 miRNAs that showed an increase in Dis3 knockdown discs by miRNA-seq also showed significant increases by qRT-PCR (summarised in Figure 3.14). For example, *miR-987-5p* shows the largest fold change by both techniques with an average increase of 11.56 fold and 16.844 fold by miRNA-

seq and qRT-PCR. Similarly *miR-277-3p* increases 8.82 fold by miRNA-seq and 3.74 fold by qRT-PCR. Although there is a difference in the magnitude of the fold change the direction of change is consistent and therefore the actual change most likely lies somewhere between the two. Another miRNA, *miR-252-5p*, increases 3.68 fold and 2.11 fold by sequencing and qRT-PCR; whilst *miR-34-5p* shows significant increases of 2.68 and 1.37 fold by miRNA-seq and qRT-PCR respectively. The final miRNA selected for showing an increase in expression following Dis3 knockdown (2.35 fold) was *miR-317-3p* which again showed a significant fold change of 2.05 fold by qRT-PCR (Figure 3.13B).

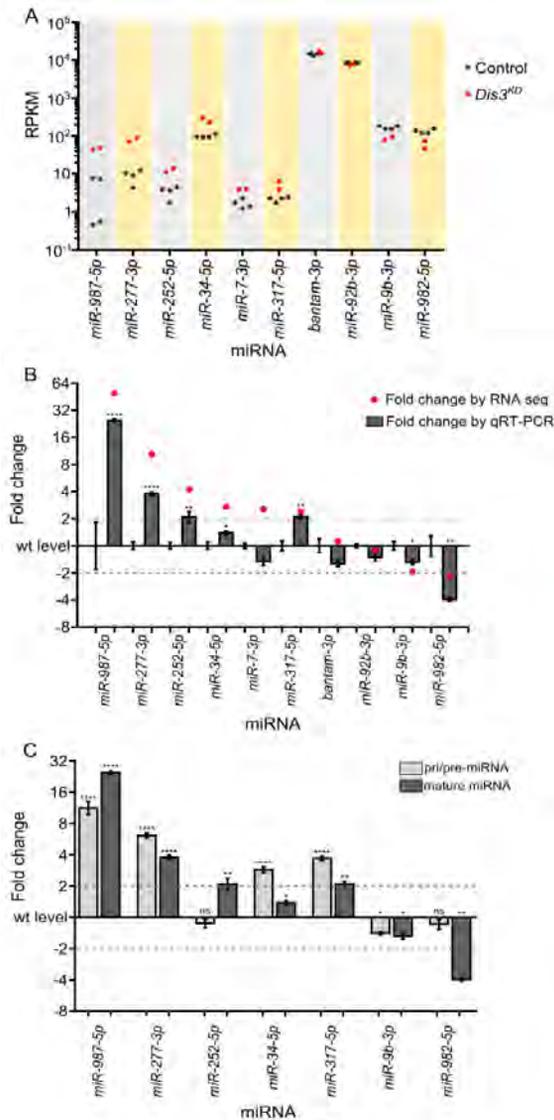
Along with those that showed an increase in expression following Dis3 knockdown two miRNAs that showed a consistent decrease in expression were also validated. Both *miR-9b-3p* and *miR-982-5p* showed similar fold changes by qRT-PCR. *miR-9b-3p* showed a decrease of 1.95 and 1.54 fold by the two techniques; whilst *miR-982-5p* showed decreases of 2.19 fold and 4.35 fold by miRNA-seq and qRT-PCR respectively. These data show that both increases and decreases in expression as determined by the miRNA-seq experiment are largely repeatable using the alternative technique of qRT-PCR. In addition the two selected miRNAs that showed no change in the miRNA-seq data (*bantam-3p* and *miR-92b-3p*) also showed no change by qRT-PCR. The comparison between miRNA-seq and qRT-PCR fold changes is presented in Figure 3.14.



**Figure 3.11: Effect of Dis3 knockdown on all wing imaginal disc miRNAs detected above threshold levels.** Fold changes (*Dis3<sup>KD</sup>* vs grouped parental controls) from miRNA-seq data of the 109 detected miRNAs in the wing imaginal disc. *miR-987-5p* shows the greatest increase in expression (49.5-fold) whilst *miR-125-5p* shows the greatest decrease in expression (-2.4-fold). miRNAs that increase  $\geq 2$ -fold are highlighted in blue font and those selected for further analysis are shown with blue bars. miRNAs that decrease  $\geq 2$ -fold are highlighted in red font and those selected for further analysis shown with red bars. Selected miRNAs that remain unchanged are highlighted in green font. Dotted lines represent  $\pm 2$ -fold changes.  $n \geq 2$ , error bars represent standard error taken from 8 fold change comparisons for each mature miRNA.



**Figure 3.12: Graphical representation of the miRNAs selected for validation.** Plot visually representing the increase (blue) or decrease (red) in normalised read count (RPKM) for those miRNAs selected for verification by qRT-PCR. The remaining miRNAs not selected for verification are shown in grey. Control = average RPKM of the 4 parental samples (*nub-GAL4* ; and ; *UAS-dis3<sup>RNAi</sup>*), *Dis3<sup>KD</sup>* = ; *nub-GAL4/+* ; *UAS-dis3<sup>RNAi/+</sup>*.



**Figure 3.13: qRT-PCR validation of misexpressed miRNAs from the miRNA-seq data.** (A) Grouping of RPKM of selected miRNAs for all replicates is highly consistent between control (black dots) and knockdown replicates (red dots). (B) Validation of miRNA-seq fold changes (red dots) using qRT-PCR. All selected miRNAs that change in expression upon Dis3-depletion in the miRNA-seq data also significantly change in expression when using qRT-PCR (except *miR-7-3p*). Dotted lines show +/- 2-fold changes. Stars represent levels of significance calculated using a two sample t-test. (\*\*\*\* $p < 0.0001$ , \*\* $p < 0.01$ , \* $p < 0.05$ , ns  $p > 0.05$ ,  $n \geq 3$ , error bars represent standard error). (C) The pri/pre-miRNA levels were measured for the miRNAs that significantly change in expression in both the miRNA-seq and qRT-PCR data. The expression of *miR-252-5p* and *miR-982-5p* change post-transcriptionally following Dis3 knockdown. All other miRNAs alter in a transcriptional manner as the pri/pre-mRNA (light grey) also changes significantly in expression, similar to the levels observed mature miRNAs (dark grey). Dotted lines show +/- 2-fold changes. Stars represent levels of significance calculated using a two sample t-test. \*\*\*\* $p < 0.0001$ , \*\* $p < 0.01$ , \* $p < 0.05$ , ns  $p > 0.05$ ,  $n \geq 3$ , error bars represent standard error.

miRNA	Fold change seq	Fold change qPCR	p-value	Validation status	Pre-miR fold change	p-value	Trans/post-trans
<i>miR-987-5p</i>	11.56	16.84	<0.0001	OK	11.03	<0.0001	Trans
<i>miR-277-3p</i>	8.82	3.74	<0.0001	OK	5.94	<0.0001	Trans
<i>miR-252-5p</i>	3.68	2.11	0.0074	OK	-1.14	0.3058	Post-trans
<i>miR-34-5p</i>	2.68	1.37	0.0316	OK	2.81	<0.0001	Trans
<i>miR-317-5p</i>	2.35	2.05	0.0023	OK	3.54	<0.0001	Trans
<i>miR-7-3p</i>	2.43	-1.49	0.0236	NO	N/A	N/A	N/A
<i>miR-9b-3p</i>	-1.95	-1.54	0.0424	OK	-1.44	0.0406	Trans
<i>miR-982-5p</i>	-2.19	-4.35	0.045	OK	-1.16	0.2826	Post-trans
<i>miR-92b-3p</i>	-1.12	-1.33	0.0628	OK	N/A	N/A	N/A
<i>bantam-3p</i>	1.11	1.66	0.1495	OK	N/A	N/A	N/A

**Figure 3.14: Summary of the miRNA-seq fold changes and the qRT-PCR fold changes for each of the selected miRNAs.** For the qRT-

PCR analysis the p-values are included and a validation status as to whether the mature miRNA shows a consistent change in expression between the two techniques. Only *miR-7-3p* had a negative validation status. Where appropriate the fold change of the *pre-miRNA* and respective p-values are included. Trans/post-trans represents whether the fold change detected by miRNA-seq and qRT-PCR occurs at the transcriptional (trans) or post-transcriptional level (post-trans). For all statistical analysis  $n \geq 3$ .

### **3.6.4 *miR-252-5p* and *miR-982-5p* are regulated at the post-transcriptional level.**

With the initial hypothesis being that Dis3 plays a role in the degradation of mature miRNAs, those that increase in expression were of immediate interest as they showed the expected trend; an increase in stability following the loss of Dis3. However, Dis3 may also be directly involved in regulating those miRNAs that show a decrease in expression, for example, through playing a role in their maturation which would be consistent to the function observed by Flynt *et al.*

If Dis3 directly regulates the stability of the identified miRNAs then one would expect regulation to occur at a post-transcriptional level. In comparison those that are affected indirectly following Dis3 knockdown in the wing imaginal disc may be expected to show changes of a transcriptional nature. To assess the nature of the fold changes the pre-miRNA hairpin expression was used as a 'transcriptional read out'. For example, a post-transcriptional change would be represented by a mature miRNA that changed in expression whilst the pre-miRNA hairpin remains unchanged. Alternatively, a change in pre-miRNA expression similar to that of the mature miRNA would signify a transcriptional change in expression. Custom TaqMan assays were designed to the pre-miRNA hairpins using Life Technologies online custom assay design tool. The full sequence of the pre-miRNA was obtained from miRBase release 20 and submitted to the online design tool. The sequences of the assays used in these experiments are shown in Methods section. The reverse transcription reaction was primed using random hexamers and *rp49* was used as a normaliser. Of the miRNAs that showed an increase in expression only one, *miR-252-5p*, showed a post-transcriptional change. *pre-miR-252* showed no significant change in expression following Dis3 knockdown in the wing disc ( $p=0.3058$ ) (Figure 3.13C), identifying *miR-252-5p* as a strong candidate for a Dis3 target. In contrast, *pre-miR-987*, *pre-miR-277*, *pre-miR-34* and *pre-miR-317* all showed significant increases in expression similar to those observed for their respective mature miRNAs

indicating transcriptional changes in expression. These transcriptional changes are likely to be indirect effects of Dis3 knockdown (Figure 3.13C).

Those miRNAs that show a transcriptional increase show a significant increase in the pre-miRNA hairpin. This could therefore result in an increase in expression in both the 3p and 5p miRNAs. If this was the case then one would expect the partner miRNA to that previously examined to also show an increase in the miRNA-seq data which may show slightly lower fold changes or may have been missed due to the read coverage. Of the miRNAs that show a transcriptional change *miR-277-5p*, *miR-34-3p* and *miR-317-5p* show similar trends to those observed for their partners providing further evidence of a transcriptional change (Figure 3.15). *miR-277-5p* would have been excluded from the analysis as it is only present on 5 of the 6 samples and therefore the observed changes are speculative. *miR-34-3p* increases 1.79 fold (compared to a 2.68 increase in *miR-34-5p*) and therefore would not cross the 2 fold criteria.

Of the two miRNAs that decreased in expression, *miR-9b-3p* showed a transcriptional change (*pre-miR-9b* showed a significant decrease of 1.44 fold), whilst *miR-982-5p* showed a post-transcriptional change with *pre-miR-982* showing no significant change in expression ( $p=0.2826$ ) (Figure 3.13C). The post-transcriptional nature of the decrease in *miR-982-5p* suggests that Dis3 may have a role in regulating its expression. This could be during its biogenesis, for example, Dis3 may be involved in trimming the precursor before it is processed by Dicer as was previously observed in *Drosophila* (Flynt *et al.* 2010) and similar to the function observed for Nibbler (Liu *et al.* 2011). Unlike the increases, in this sense one may expect both the 5p and the 3p to decrease in expression if the effect of Dis3 on biogenesis is direct. This appears to be the case as *miR-982-3p* shows a similar decrease in the sequencing (4.6 fold), however, it is only detected in one of the knockdown replicates and therefore would require further validation (Figure 3.15). Nevertheless, it appears that Dis3 may play a role in the regulation of *miR-982-5p* biogenesis.

### **3.6.5 Loss of the cytoplasmic function of Dis3 is responsible for the increase in *miR-252* expression.**

If Dis3 was involved in the degradation of *miR-252-5p* then one would expect the increase in expression to be due to inhibition of the cytoplasmic function of the exosome. Dis3L1, the catalytic subunit of the cytoplasmic exosome in humans, is not present in the *Drosophila melanogaster* genome; therefore Dis3 is presumably responsible for both the nuclear and cytoplasmic activity of the exosome. Therefore there was a requirement to separate the cytoplasmic and nuclear activities of the exosome. Rrp40, is a core exosome subunit that has been shown to be predominantly present in the cytoplasm in *Drosophila* (Graham *et al.* 2006) and therefore in knocking down Rrp40 only the cytoplasmic function of the exosome would be inhibited.

To initially explore the requirement of the cytoplasmic activity of the exosome, Rrp40 was knocked down ubiquitously using *tubulin-GAL4*. This resulted in lethality predominantly at the 2<sup>nd</sup> instar larval stage; however, a few larvae reached L3 before lethality whilst no knockdown progeny reached the pupal stage. The few that reached L3 showed melanotic masses (Figure 3.16D) as were observed for total Dis3 knockdown by Hou *et al.* In these data it appears that inhibition of the cytoplasmic exosome, although still critical, is less severe than the inhibition of Dis3 which would inhibit the activities of both the cytoplasmic and nuclear exosome. The similarity between the Rrp40 phenotypes and the Dis3 phenotype presented by Hou *et al* suggests that maybe their knockdown was less efficient and this could be due to their use of 'room temperature' as a cross temperature compared to the constant temperature of 25°C used in the experiments presented here.

The ubiquitous knockdown of Rrp40 results in a phenotype slightly less severe than that observed for the Dis3 knockdown presented in previous sections. This would be consistent with the hypothesis the nuclear exosome retains at least some of the nuclear function. To ask if the cytoplasmic function of the exosome was responsible for the changes in *miR-252-5p*

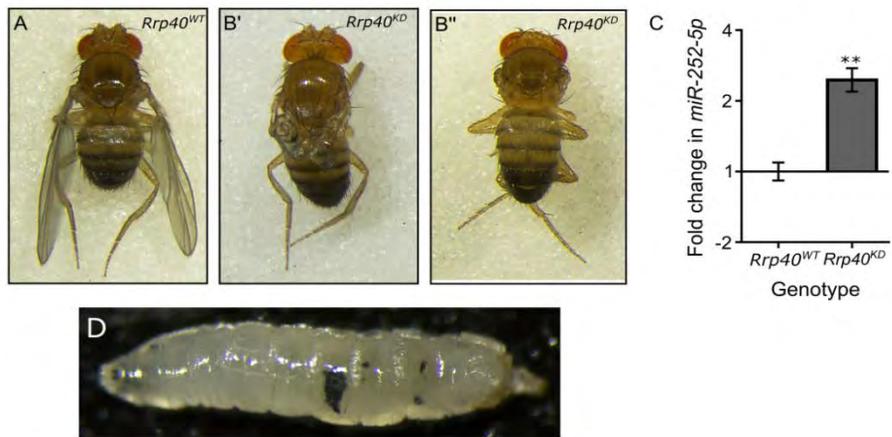
expression in the wing disc Rrp40 was knocked down using *nub-GAL4*. Rrp40 knockdown adults presented with a severe wing phenotype, although similarly this was less severe than when Dis3 was knocked down in the same compartment. 90% of Rrp40 knockdown flies showed severe crumpling of the wing (Figure 3.16B') whilst the remaining 10% showed the no wing phenotype observed when Dis3 is knocked down (Figure 3.16B''). This again suggests that the cytoplasmic function of the exosome/Dis3 is critical for correct wing development.

To see if the cytoplasmic function of the exosome/Dis3 was responsible for the increased expression of *miR-252-5p*, as would be expected if it were a direct target, the effect of Rrp40 knockdown on the expression of *miR-252-5p* was assessed. Knockdown of Rrp40 specifically in the wing pouch resulted in a 2.5 fold increase in *miR-252-5p* expression comparable to the 2.1 fold increase observed when Dis3 is knocked down in the same region (Figure 3.16C). This therefore implies a role of the cytoplasmic activity of the exosome/Dis3 in the regulation of *miR-252-5p* expression.

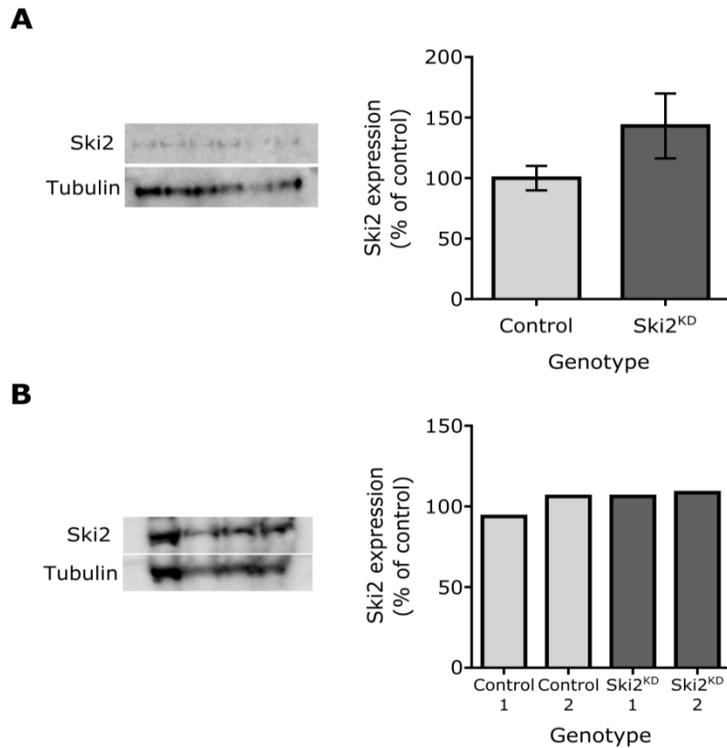
In addition to inhibiting the cytoplasmic exosome through targeting cytoplasmic subunits the Ski complex was also targeted. The Ski complex acts as a cytoplasmic helicase to unwind structured RNAs and feed them into the central channel of the exosome. It has previously been shown that the Ski complex is required for activity of the cytoplasmic exosome (Anderson and Parker 1998; Halbach *et al.* 2013). Ski2 (Twister) was targeted by driving a *UAS-ski2<sup>RNAi</sup>* in the wing imaginal disc using the *69B-GAL4* and *nub-GAL4* drivers. However, no knockdown was observed at the protein level for either driver (Figure 3.17) and therefore it was determined that the RNAi was not working as expected. As a result no further investigation was carried out using Ski2.

Validated miRNA	Fold change (RNA-seq/qRT-PCR)	Partner miRNA	Fold change (RNA-seq)
<i>miR-987-5p</i>	11.56/16.84	<i>miR-987-3p</i>	N/A
<i>miR-277-3p</i>	8.82/3.74	<i>miR-277-5p</i>	11.1
<i>miR-252-5p</i>	3.68/2.11	<i>miR-252-3p</i>	N/A
<i>miR-34-5p</i>	2.68/1.37	<i>miR-34-3p</i>	1.79
<i>miR-317-5p</i>	2.35/2.05	<i>miR-317-3p</i>	1.57
<i>miR-9b-3p</i>	-1.95/-1.54	<i>miR-9b-5p</i>	1.06
<i>miR-982-5p</i>	-2.19/-4.35	<i>miR-982-3p</i>	-4.60

**Figure 3.15:** Table summarising the fold changes as determined by miRNA-seq (RNA-seq) of the partner miRNAs to those miRNAs validated by qRT-PCR. N/A is shown when the miRNA was only present in one of the knockdown replicates.



**Figure 3.16: Knockdown of the exosome subunit Rrp40 results in similar phenotypes as to when Dis3 is knocked down.** Knockdown of the exosome subunit Rrp40 using the *nubbin-GAL4* wing pouch driver (*;nub-Gal4/UAS-Rrp40<sup>RNAi</sup>*) results in severe wing development phenotypes (**B'** and **B''**) when compared to parental controls (**A**). 90% of the eclosing progeny present a severe phenotype with crumpled wings (**B'**) whereas 10% of eclosing progeny show a similar phenotype to Dis3 knockdown with a complete absence of wing development (**B''**). (**C**) Knockdown of Rrp40 in the wing pouch (*nub-GAL4*) results in a similar increase in *miR-252-5p* expression as observed following Dis3 depletion. *Rrp40<sup>WT</sup>* = *;nub-GAL4* ; and *;UAS-Rrp40<sup>RNAi</sup>* ; *Rrp40<sup>KD</sup>* = *;nub-GAL4/UAS-Rrp40<sup>RNAi</sup>* ;  $n \geq 3$ ,  $**p < 0.01$ , error bars represent standard error. (**D**) Ubiquitous knockdown of Rrp40 driven by the *tubulin-GAL4* driver results in large scale lethality at the 2<sup>nd</sup> instar larval stage. However those few surviving to L3 showed melanotic masses (black spots).

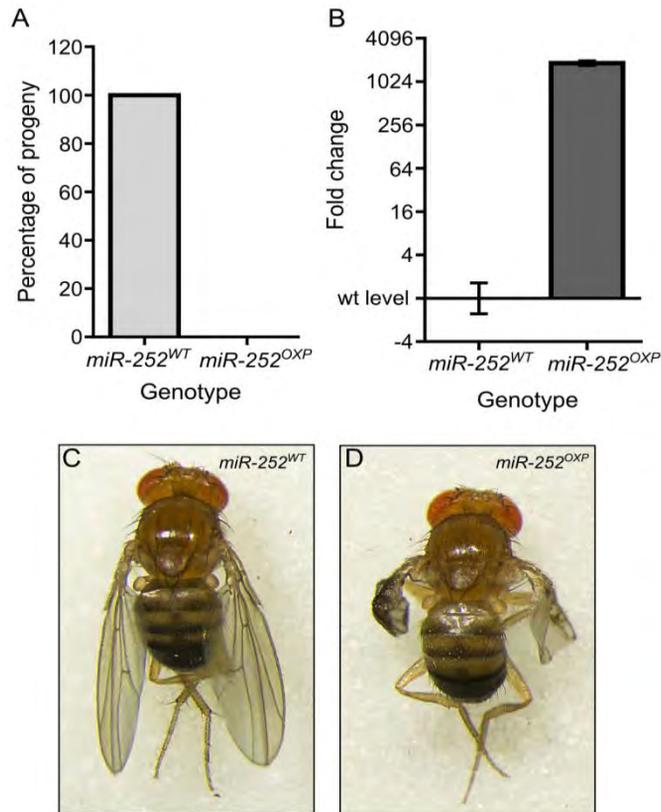


**Figure 3.17: Unsuccessful knockdown of Ski2. (A)** Attempted knockdown of Ski2 in the wing pouch of the imaginal disc by driving *UAS-ski2* with *nub-GAL4*. No significant difference was observed between knockdown and parental control wing imaginal discs with each replicate consisting for 30 wing imaginal discs. Control = ; *UAS-ski2<sup>RNAi</sup>* ; *Ski2<sup>KD</sup>* = ; *nub-Gal4/UAS-ski2<sup>RNAi</sup>* ;  $n \geq 3$ , error bars represent standard error,  $p=0.2061$ . **(B)** Attempted knockdown of Ski2 by driving *UAS-ski2* with *69B-GAL4*. No significant difference was observed between knockdown and parental controls with each replicate consisting for 3 whole wandering L3 larvae. Control = ; *UAS-ski2<sup>RNAi</sup>* ; *Ski2<sup>KD</sup>* = ; *UAS-ski2<sup>RNAi</sup>/+* ; *69B-GAL4/+*.  $n=2$ , error bars represent standard error.

### **3.7 *miR-252* is required for correct wing development.**

*miR-252-5p* has been identified as a potential Dis3 target. Therefore it was of interest to explore the developmental importance of *miR-252* in terms of its requirement within the development of the wing, something Dis3 is critical for (section 3.4). In an attempt to mimic the conditions in which *miR-252-5p* is overexpressed following Dis3 knockdown, *miR-252* was overexpressed in the wing pouch using the *GAL4-UAS* system. The *miR-252* hairpin lies downstream of the UAS allowing *miR-252* to be ectopically expressed under the control a *GAL4* promoter. The global effect of *miR-252* overexpression was first examined using the *tubulin-GAL4* promoter. Ubiquitous overexpression of *miR-252* was confirmed in whole larvae by qRT-PCR (Figure 3.18B) and results in lethality, similar to that observed following ubiquitous Dis3 knockdown with no overexpression progeny reaching L3 (Figure 3.18A). This therefore indicates that it is critical to maintain *miR-252* at the correct level to maintain organism viability.

More specific overexpression of *miR-252* in the wing imaginal disc using the *nubbin-GAL4* promoter was performed to ask if control over the expression of *miR-252* is important for correct wing development. Ectopic expression of *miR-252* in the wing pouch of the wing imaginal disc results in severe wing crumpling (Figure 3.18C/D). The similarity between the knockdown of Dis3 and the overexpression of *miR-252* suggests that their functions are critical for correct wing development. It therefore signifies a potential, developmentally important, role for Dis3 in its regulation of *miR-252-5p* expression.



**Figure 3.18: Overexpression of *miR-252* in the wing pouch of the imaginal disc results in severe wing phenotypes. (A)** Overexpression of *miR-252* using the ubiquitously expressed *tub-GAL4* driver results in 100% lethality compared to controls (*UAS-miR-252/TM6*).  $n=184$ . **(B)** Overexpression of *UAS-miR-252* in the wing pouch using the *nubbin-GAL4* driver results in a 1056-fold increase in *miR-252-5p* expression in the wing imaginal disc.  $n \geq 3$ ,  $p < 0.0001$ , error bars represent standard error. **(C and D)** Overexpression of *miR-252* in the wing pouch of the wing imaginal disc using the *nub-GAL4* driver (*nub-Gal4/+ ; UAS-miR-252/+*) results in severe wing crumpling **(D)** compared to a parental control (*miR-252<sup>WT</sup> = ; ; UAS-miR-252*) **(C)** with 100% penetrance.

### 3.8 Analysis of predicted *miR-252-5p* targets.

Dis3 appears to regulate the expression of *miR-252-5p* with Dis3-depletion resulting in its overexpression at the post-transcriptional level. *miR-252* levels are also critical for organism viability and correct development of the wing. What mRNAs does *miR-252-5p* normally target that may become misregulated themselves following its overexpression? There are multiple target prediction sites available online of which TargetScan Fly 6.2 (Lewis *et al.* 2005) was selected to predict potential *miR-252-5p* targets. TargetScan provides an output of conserved sites and poorly conserved sites in addition to the total number of sites within that specific UTR. Using TargetScan 263 potential targets of *miR-252-5p* were identified with each containing between 1 and 3 complementary sites within the 3' UTRs. No *miR-252-5p* targets have been previously validated and it is highly unlikely that all 263 mRNAs are targeted by *miR-252-5p*; however, those showing the highest complementarity were the most likely targets.

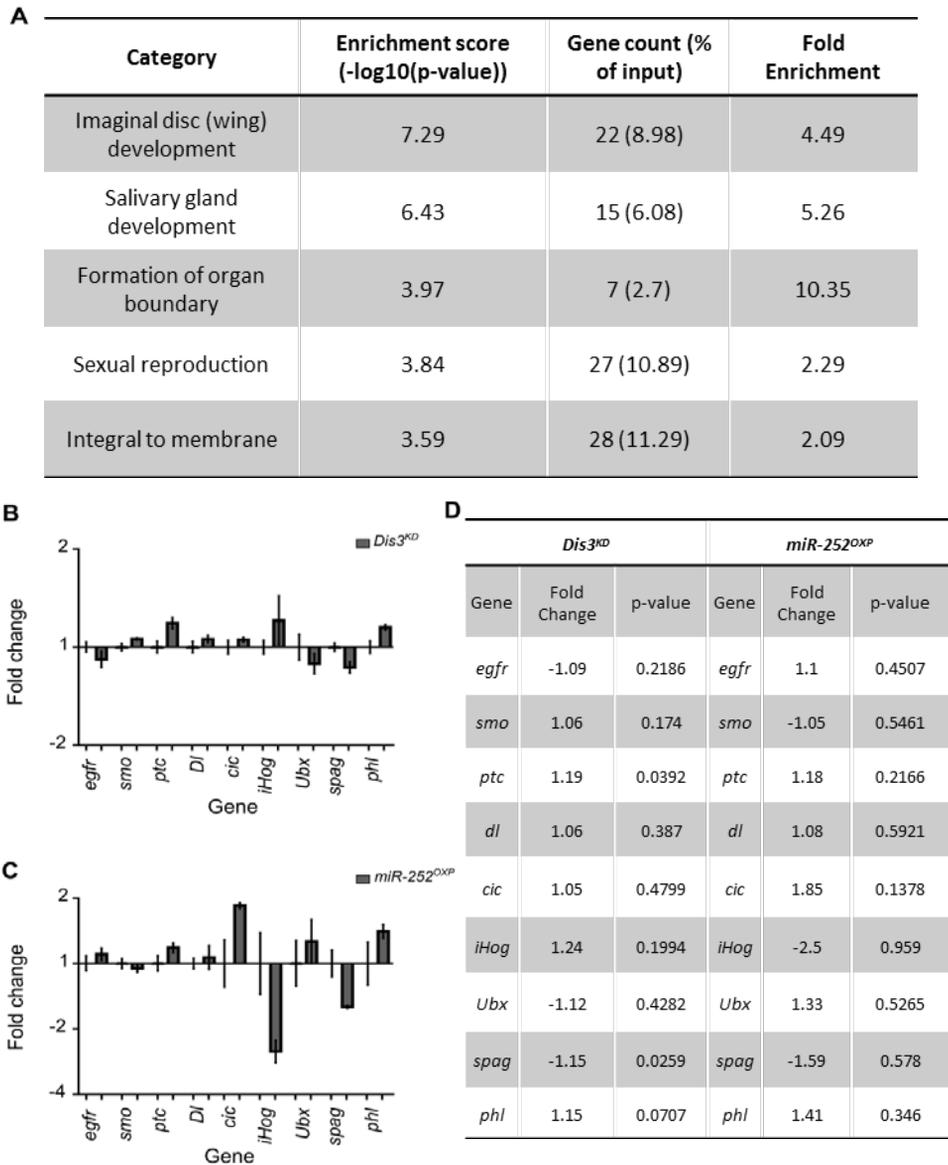
If *miR-252-5p* is involved in the regulation of wing development, as was suggested by the observed wing phenotypes, one would expect that some of the predicted targets would be involved in wing development themselves. To assess the functions of the predicted targets gene ontology analysis was performed using all 263 predicted targets as there is no evidence why to exclude certain predictions. GO analysis was performed using DAVID release 6.7. The Flybase transcript ID of all predicted targets was submitted and functional classification was performed using the DAVID default settings. The highest enrichment of target function (score of 7.29) was indeed in imaginal disc morphogenesis, with 8.98% of the predicted targets being involved in imaginal (wing) disc development ( $p=5.13 \times 10^{-8}$ ) (Figure 3.19A). This therefore suggests the leading function of the predicted *miR-252-5p* targets was to regulate disc development and morphogenesis which correlates with the overexpression phenotypes.

In an attempt to validate *miR-252-5p* targets 9 genes which were within the 'wing disc development' category identified by GO analysis, in addition to having antibodies available (Figure 3.19D), were selected. Changes in expression of the selected genes either when Dis3

was knocked down, or when *miR-252-5p* was overexpressed in the wing imaginal disc was surveyed using qRT-PCR. If *miR-252-5p* normally targets and directs degradation of the selected genes reductions in their expression may be expected when *miR-252-5p* is overexpressed. Two genes showed significant fold changes following Dis3 knockdown, however, these were very modest. For example, *patched (ptc)* only increased 1.19 fold whilst *spaghetti (spag)* decreased 1.15 fold (Figure 3.19 B/D). These modest changes, although statistically significant, are unlikely to carry biological significance. Unfortunately none of the genes showed any significant changes in expression at the RNA level when *miR-252* was overexpressed (Figure 3.19 C/D). Although this may suggest that the target predictions were incorrect, it could be the case that *miR-252-5p* binding to its targets may not result in degradation but instead translational repression. If this was indeed the case the RNA levels of the predicted targets would remain unchanged. Unfortunately the antibodies to the selected genes did not work well enough to investigate this further. The level of buffering within the developmental system may also make it difficult to detect *in vivo* changes following *miR-252* overexpression. As a result this work was unable to validate predicted targets of *miR-252-5p* and make a link to the phenotypes observed when Dis3 is knocked down or *miR-252* is overexpressed.

### 3.9 3' editing of miRNAs.

Recent discoveries have identified a role for 3' terminal modifications in regulating miRNA abundance and activity. For example 3' oligouridylation of *pre-let7a* promotes its degradation by Dis3L2 (Chang *et al.* 2013; Ustianenko *et al.* 2013), whilst its monouridylation is required for Dicer processing and subsequent biogenesis (Heo *et al.* 2012). Additionally 3' adenylation of *miR-122* in humans has been shown to result in the stabilisation of the miRNA (Kato *et al.* 2009). Do the miRNAs identified as potential Dis3 targets show any sign of 3' modifications which may be used as a method to target Dis3 activity? To this end read variations for *miR-982-5p* and *miR-252-5p* were collated and analysed for non-template bases at the 3' end. However, for both examples, in the pre-miRNA the base following the 3' most nucleotide of the mature sequence is a U making it impossible to differentiate between template and non-template U residues. This was made even more difficult for *miR-252-5p* as U's make up the 2 bases downstream of the mature sequence in the precursor. However there are examples of non-template Uridine, Guanine and Adenine residues at the 3' termini at low frequencies (Figure 3.20). The frequencies were also too low to make any substantial conclusion on if the 3' modifications vary between the control and knockdown samples. In addition to the low frequencies the size selection during the library preparations was highly specific to include miRNAs of 18-22nt in length. Both *miR-252-5p* and *miR-982-5p* are 22nt in length and therefore any further modifications may be missed as a result of the size selection cut off. Additionally, if Dis3 is indeed involved in the biogenesis of *miR-982-5p* then one may expect the terminal additions to occur to the precursor which would not be detectable in this data set. As a result, although there appears to be some 3' modifications, any confident conclusions about 3' editing of the Dis3 targets were not possible.



**Figure 3.19: Knockdown of *Dis3* or overexpression of *miR-252* does not affect the expression of predicted *miR-252-5p* targets. (A)** Summary of gene ontology analysis of the targets of *miR-252-5p* as predicted by TargetScanFly 6.2. The greatest enrichment was of genes involved in wing imaginal disc development. **(B)** Knockdown of *Dis3* in the wing pouch using *nub-GAL4* does not affect the expression of predicted *miR-252-5p* targets at the RNA level. *Dis3<sup>KD</sup>* = ; *nub-GAL4/+* ; *UAS-dis3<sup>RNAi</sup>/+*.  $n \geq 3$ , error bars represent standard error. **(C)** Overexpression of *miR-252* in the wing pouch using *nub-GAL4* does not affect the expression of predicted *miR-252-5p* targets at the RNA level. *miR-252<sup>OXP</sup>* = ; *nub-GAL4/+* ; *UAS-miR-252/+*.  $n \geq 3$ , error bars represent standard error. **(D)** Summary of the fold change (shown in **B** and **C**) and respective p-value for each selected predicted *miR-252-5p* targets when either *Dis3* is knocked down (*Dis3<sup>KD</sup>*) or when *miR-252* is overexpressed (*miR-252<sup>OXP</sup>*).

A

Control	Dis3 <sup>KD</sup>	<i>miR-252-5p</i>
0.0%	2.6%	UAAGUACUAGUGCCGCAGGAG
0.0%	2.6%	UAAGUACUAGUGCCGCAGGAG <sup>A</sup>
2.8%	7.0%	UAAGUACUAGUGCCGCAGGAG <sup>U</sup>
12.5%	0.0%	UAAGUACUAGUGCCGCAGGAG <sup>UUU</sup>
1.8%	0.0%	UACGUACUAGUGCCGCAGGAG <sup>UU</sup>
24.4%	14.8%	CUAAGUACUAGUGCCGCAGGAG <sup>U</sup>
22.6%	8.7%	CUAAGUACUAGUGCCGCAGGA
35.9%	64.3%	CUAAGUACUAGUGCCGCAGGAG
		UCCUAAGUACUAGUGCCGCAGGAG <sup>UUU</sup>

B

Control	Dis3 <sup>KD</sup>	<i>miR-982-5p</i>
0.0%	2.5%	UCCUGGACAAAUAUGAAGUA
0.1%	5.0%	UCCUGGACAAAUAUGAAGUAA
1.9%	0.8%	UCCUGGACAAAUAUGAAGUAAA
0.0%	0.4%	UCCUGGACAAAUAUGAAGUAAA <sup>U</sup>
0.3%	0.4%	UCCUGGACAAAUAUGAAGUAAA <sup>UU</sup>
0.1%	0.0%	UCCUGGACAAAUAUGAAGUAAA <sup>UG</sup>
82.0%	45.3%	UCCUGGACAAAUAUGAAGUAAA <sup>U</sup>
15.6%	45.6%	UCCUGGACAAAUAUGAAGUAAA
		GAUCCUGGACAAAUAUGAAGUAAA <sup>UGU</sup>

**Figure 3.20: Terminal nucleotide additions to the Dis3 regulated miRNAs. (A)** *miR-252-5p* shows 2 examples of non-template (red) additions to the 3' end of the mature miRNA however at low frequency. Nucleotides outside the mature sequence but that matched the following precursor bases were also observed (blue). Frequencies of each read variant are presented for the control of Dis3 knockdown samples. Due to low read counts confident determination of nucleotide additions were not possible. **(B)** *miR-982-5p* shows 2 examples of non-template additions to the 3' end of the mature miRNA however at low frequency. Nucleotides outside the mature sequence but that matched the following precursor bases were also observed (blue). Frequencies of each read variant are presented for the control of Dis3 knockdown samples. Due to low read counts confident determination of nucleotide additions were not possible.

### 3.10 Identification of a novel miRNA in the *Drosophila* wing imaginal disc.

The global, unbiased nature of miRNA-sequencing not only allows assessment of known miRNAs but gives the potential for the identification of novel miRNAs. miRDeep2 (Friedlander *et al.* 2008; Friedländer *et al.* 2012) is a programme designed to identify potential novel miRNAs using the unique structural and sequential characteristics of miRNAs. For example the reads mapping to a potential precursor must map to one of the three main features; mature miRNA on either arm or the loop sequence which is cleaved during Dicer processing. For example, if multiple reads span the potential mature sequence and the loop, the precursor and potential novel miRNA would be excluded (Figure 3.21). miRDeep2 also includes a folding prediction (RNAfold) where the potential precursor must be predicted to fold into a unbifurcated hairpin to be retained.

In addition to the predicted novel miRNA sequences miRDeep2 also gives an output of the number of reads mapped to known miRNAs. This allowed cross comparison between this and the analysis performed by ARK Genomics. The selected miRNAs showed similar trends between the two methods of analysis. Using other *Drosophalids* to model species conservation as a method of increasing the reliability of novel miRNA prediction 9 novel sequences were predicted by miRDeep2 across the 6 samples. The initial mapping and identification of potentially novel miRNA sequences was performed by Dr. Sarah Smalley. 5 of the 6 samples showed hits for potential novel miRNAs with no novel sequences being identified in *UAS-dis3<sup>RNAi</sup> 1*. This is most likely due to the reduced coverage as *UAS-dis3<sup>RNAi</sup> 1* received the fewest reads of all the samples (115734 reads).

To identify the most likely candidates only those sequences that were identified in 4 of the 5 samples which showed hits for novel miRNAs, including at least one replicate of all 3 genotypes, were selected (Figure 3.22A). Filtering resulted in the identification of 2 sequences (preliminarily named *miR-B1* and *miR-B2*) which were taken further for validation. One

characteristic miRDeep2 uses is the formation of a hairpin structure by the predicted precursor from which the mature miRNA is processed. To assess the predicted secondary structure of the suggested pre-miRNA the RNA fold web server was used with default parameters (Lorenz *et al.* 2011). *pre-miR-B1* was predicted to form a traditional hairpin loop with *miR-B1* situated on the 3' arm (Figure 3.22B) *pre-miR-B2* was also predicted to form a hairpin-like structure with a few mismatches causing a slight bulge (Figure 3.22B) in the centre of the secondary structure. However both predicted pre-miRs form structures similar to those formed by known pre-miRNAs which increases their likelihood of being novel miRNAs.

To determine whether the novel sequences were conserved a BLAST search was conducted. A BLAST of *miR-B1* identifies a red flour beetle miRNA, *tca-miR-3882-5p*, showing only 2 mismatches (score 63, evalue 5.1) suggesting *miR-B1* may be conserved which often indicates biological relevance. A BLAST search of the mature sequence of *miR-B2* does not identify any potential homologues; therefore it is unlikely to be conserved.

To validate the presence of the predicted novel mature miRNAs custom TaqMan assays were designed by submitting the sequence provided by miRDeep2 to the Life Technologies small RNA custom design tool. RNA was then reverse transcribed using primers specific to the novel miRNAs followed by qRT-PCR to validate their presence. *snoR442* was used as a normaliser. Both sequences were present in the mature form showing strong traces. No signal was observed in the no reverse transcriptase control indicating the signal was specific. *miR-B1* showed an average Ct of 31.7 whilst *miR-B2* was more abundant with an average Ct of 27.6. Neither *miR-B1* nor *miR-B2* showed differential expression between the control and knockdown samples indicating that Dis3 is unlikely to play a role in their regulation (Figure 3.22C/D).

To identify their location within the genome the hairpin sequences from miRDeep2 were mapped to the *Drosophila* genome. *pre-miR-B1* is located at 2R:14067655-2R:14067710 and is derived from the 5' UTR of *CG5726* in the opposing orientation. On the other hand *pre-miR-B2*

is located in the 3' end of tRNA:CR30407. It is possible that *miR-B2* is a product of a cleaved tRNA therefore due to this area of doubt *miR-B2* was not submitted to miRBase as a novel miRNA. However, *miR-B1* was accepted into miRBase and officially named *miR-11182-3p*.

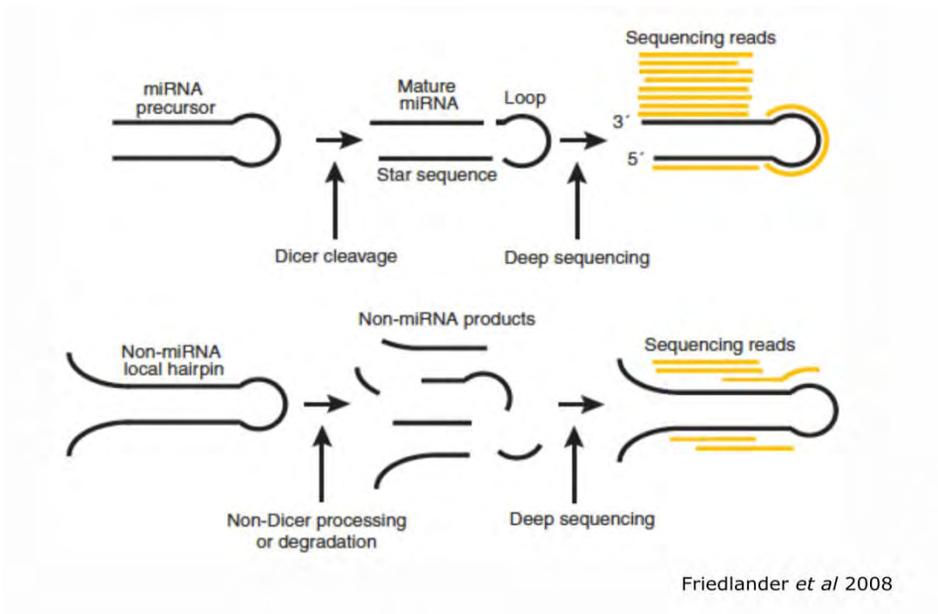
### 3.11 Chapter Summary

The objectives of the work within this chapter were to characterise a role for Dis3 in the development of the *Drosophila* wing imaginal disc together with investigating if Dis3 plays a role in the regulation of miRNA stability. A critical role for Dis3 during the development of the wing disc has been shown with its depletion resulting in a complete lack of wing development. This appears to occur due to large scale apoptosis within the wing pouch resulting in the death of the wing blade fated cells.

A global, unbiased screen has been conducted which has allowed the determination of miRNA expression profile of wing imaginal disc. Using miRNA sequencing the expression of 109 miRNAs within the wing imaginal disc were confidently identified. This is the first time that wing imaginal discs (rather than all imaginal discs) had been analysed using RNA-seq to identify changes in levels of miRNAs. The miRNA-seq experiment also revealed a small pool of Dis3-sensitive miRNAs in addition to the identification of a potential target of Dis3, *miR-252-5*. Taken together the work presented in this chapter achieved the main objectives.

The bulk of this work was published in RNA Biology in 2015:

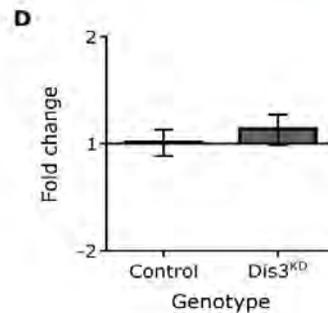
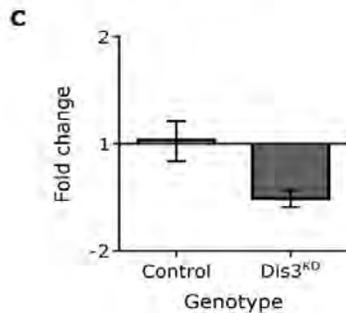
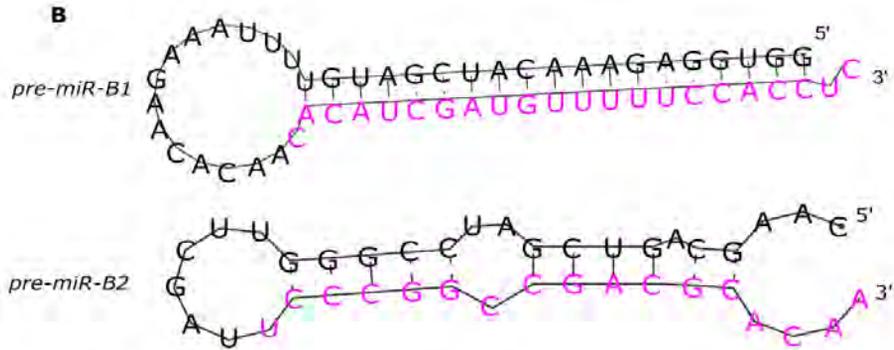
Towler, B. P., C. I. Jones, S. C. Viegas, P. Apura, J. A. Waldron, *et al.* (2015). "The 3'-5' exonuclease Dis3 regulates the expression of specific microRNAs in *Drosophila* wing imaginal discs." RNA Biol **12**(7): 728-741.



**Figure 3.21: The methods used by miRDeep2 to determine potential novel miRNA precursors. Image taken from Friedlander *et al* 2008.**

**A**

Unique Sequences	Mature Read Count per sample						miRBase Blast (common name)
	<i>nub-GAL4</i> 1	<i>nub-GAL4</i> 2	<i>UAS-dis3<sup>RNAi</sup></i> 1	<i>UAS-dis3<sup>RNAi</sup></i> 2	<i>Dis3<sup>KD</sup></i> 1	<i>Dis3<sup>KD</sup></i> 2	
cagcgucagucgagcggaagg							none
<b>cacaucgauguuuuuccaccuc</b>	23	24		17	19	14	<i>tca-miR-3882-5p</i> (Red Flour beetle)
uaccuuucgaacguuuuuuucugu	3						<i>dme-miR-4964-3p</i> , <i>ppe-miR-6286</i> (peach)
ccgcguuuuguuggccu	468						<i>gga-miR-1651-3p</i> (red junglefowl)
<b>ucccgccgagcaca</b>		478		337	2224	2095	none
uucggugcgagagguccgggu		33			73	108	<i>ppt-miR1041</i> (moss)
cuguucgugcguaggu				1			none
ucugauuucgggcuucca					1		none
aucucuagaacguuuuuccag					1		none



**Figure 3.22: The detection of a novel miRNA in the *Drosophila* wing imaginal discs by miRNA-sequencing. (A) Summary of the potential novel miRNA sequences in each sample together with a miRBase blast for potential conservation. The sequences of *miR-B1* (red) and *miR-B2* (blue) are highlighted. (B) Predicted secondary structure of *pre-miR-B1* and *pre-miR-B2* using the RNAfold webserver. The mature sequences of *miR-B1* and *miR-B2* are highlighted in pink. (C) *miR-B1* expression is not affected by *Dis3* knockdown in the wing imaginal disc.  $n \geq 3$ , error bars represent standard error,  $p > 0.05$ . (D) *miR-B2* expression is not affected by *Dis3* knockdown in the wing imaginal disc.  $n \geq 3$ , error bars represent standard error,  $p > 0.05$ .**

## 3.12 Chapter Discussion

### 3.12.1 Differences between stages of lethality with published data when Dis3 is knocked down ubiquitously.

Previous work examining the effect of ubiquitous knockdown of Dis3 in *Drosophila melanogaster* observed lethality at the 2<sup>nd</sup> instar larval stage (L2) when 3 ubiquitous GAL4 drivers were used to drive 2 independent *UAS-dis3<sup>RNAi</sup>* lines (Hou *et al.* 2012). In each case melanotic masses were observed and the lethality was preceded by a delay of 2 weeks. In this work, knockdown of Dis3 using the *da-GAL4* driver resulted in similar lethality at L2, however there was an absence of melanotic masses and pre-death delay. 100% of the *da-GAL4* driven *dis3L2* knockdown larvae has died within 48 hours of emerging from the media. The slight differences between the data continue when the stronger GAL4 drivers were used (*tub-GAL4* and *act-GAL4*). There was no evident development to L2 suggesting lethality prior to that stage. These observed differences are likely due to the efficiency of the knockdown as Hou *et al.* carried out their crosses as 'room temperature' which for a system dependent on temperature is substandard. Room temperature is likely to vary but even if it remained constant it is likely to be lower than the constant 25°C used in these experiments. As a result the *GAL4-UAS* system would be less efficient resulting in a reduced knockdown in their experiments compared to the data presented here.

In line with a less efficient knockdown, melanotic masses were only observed when Rrp40, a cytoplasmic exosomal subunit, was knocked down. Overall, knockdown of Rrp40 resulted in phenotypes of reduced severity, which is most likely due to the retained function of the nuclear exosome. Therefore, although the overall information gained from the two data sets, that Dis3 is essential for organism viability, is consistent there are some discrepancies in the details which are likely to be caused by different knockdown efficiencies.

### **3.12.2 Why is the loss of Dis3 so dramatic?**

The data presented in this chapter show that Dis3 expression is critical to maintain organism viability and tissue development. The ubiquitous loss of Dis3 leads to lethality whilst wing restricted knockdown results in a complete lack of wing development. This clearly shows how the functions of Dis3 within the cell are critical to maintain homeostasis. The loss of Dis3 in the wing imaginal disc has been shown to result in a large and specific increase in apoptosis which is likely to be the cause of the deficient tissue/organism development.

Why does the loss of Dis3 result in the dramatic increase in apoptosis? There could be many mechanistic reasons with a likely one being related to its role in ribosome biogenesis. Dis3 has been shown to be required for the biogenesis of the 5.8S rRNA (Mitchell *et al.* 1997) therefore in its absence the critical activity of the ribosome would be reduced resulting in a growth deficiency (as seen by the discs being 23% smaller) and/or ultimately cell death. Dis3 has also been shown to be essential for correct cell cycling with its mutation resulting in microtubule defects (Smith *et al.* 2011) and deficiencies in mitotic progression (Murakami *et al.* 2007). The apparent role in cell cycle regulation is likely to contribute to the apoptosis observed in the wing imaginal disc. For example, it is known that deficiencies such as cell cycle stalling often drive the cell towards apoptosis. Moreover, Dis3 has been implicated in a number of other cellular processes such as the regulation of non-coding RNAs such as miRNAs (Flynt *et al.* 2010; Liu *et al.* 2014), PROMTs/cryptic unstable transcripts (Wyers *et al.* 2005; Preker *et al.* 2008) together with snoRNAs (Schneider *et al.* 2012; Szczepinska *et al.* 2015) and tRNAs (Gudipati *et al.* 2012). It is therefore unlikely that a single role for Dis3 is responsible for the observed phenotypes but instead the vast range of functions for Dis3 and the exosome is likely to collate to enormous dysregulation within the cell leading to stress and subsequent apoptosis.

### **3.12.3 A novel role for Dis3 in the regulation of mature miRNAs.**

The above data show a relatively small subset of miRNAs are susceptible to changes in Dis3 activity which indicates a level of specificity. Using RNA sequencing allowed a global, unbiased

approach to identify Dis3-regulated miRNAs. However, the RNA-sequencing alone does not show the regulatory level at which the expression changes occur. For example, RNA-sequencing will show a miRNA to increase 4 fold in expression; however the final change may result from an increase in transcription and/or an increase in transcript stability. Differentiating between those miRNAs showing transcriptional and post-transcriptional changes was critical to determine if a specific miRNA may be targeted by Dis3.

Those differentially expressed miRNAs showing transcriptional changes are likely to be induced by indirect effects of Dis3 knockdown. For instance, Dis3 depletion in the disc may result in the increased expression of a transcription factor which in turn results in their transcriptional increase. Interestingly, *miR-34* and *miR-277* are both encoded within a noncoding RNA CR43459, the function of which is unknown. It is therefore possible that they are under the same transcriptional regulation and processing factors which is why transcriptional changes were observed for both. Additionally, another miRNA which shows a transcriptional increase, *miR-317-5p*, is situated 3Kb upstream of CR43459 and is therefore potentially controlled by similar regulatory elements as *miR-34* and *miR-277*. Interestingly miRBase classes the three miRNAs as 'clustered' and these results may also provide evidence of their co-regulation.

Two miRNAs were shown to change in a post-transcriptional manner with one increasing and the other decreasing in expression in Dis3-depleted wing imaginal discs. The changes in expression of both of these miRNAs occur at the correct 'level' within the central dogma to be directly regulated by Dis3. Due to the initial hypothesis that Dis3 does play a role in the regulation of mature miRNA stability the focus was upon *miR-252-5p* which showed an increase in stability following the loss of Dis3. The post-transcriptional increase in stability suggested that *miR-252-5p* was directly targeted by Dis3, an effect that would be expected to occur in the cytoplasm. The effect was indeed shown to result from the inhibition of the cytoplasmic exosome as knockdown of the cytoplasmic specific subunit Rrp40 (Graham *et al.* 2006) resulted in a similar increase indicating the cytoplasmic function of the Dis3/exosome

complex was responsible. This was the first example of Dis3 being implicated in the regulation of mature miRNA stability.

*miR-252-5p* was the only miRNA shown to increase at the post-transcriptional level following the loss of Dis3 in the wing imaginal disc. This indicates that the role Dis3 plays in the regulation of miRNA stability is specific rather than general. However, it is also unlikely that *miR-252-5p* is the only miRNA vulnerable to Dis3-mediated degradation. For example, for technical reasons it is likely changes in other miRNAs that are normally of lower expression are likely to have been missed due to the 2S rRNA library contamination. Additionally, a 2 fold cut off was used as a threshold to determine those miRNAs that changed in expression. Due to the knockdown being restricted to half the dissected tissue, those showing 2 fold changes were likely to actually increase by a larger amount. Therefore, this would also mean that those miRNAs showing 2 fold changes within the whole tissue were likely to be missed due dilution by the Dis3-wild-type portion of the disc. Unfortunately this technical factor was unavoidable as the knockdown of Dis3 across the whole disc resulted in L2 lethality.

In addition to *miR-252-5p*, *miR-982-5p* also shows a post-transcriptional change in expression, although in this case it was a decrease. It is therefore unlikely to be regulated by Dis3 at the level of RNA stability. However, it is possible that Dis3 may play a role in the regulation of *miR-982-5p* biogenesis such as 3' trimming of its precursor during its maturation. Similar functions are known for Dis3 in the 3' trimming of the 5.8S rRNA precursor (Mitchell *et al.* 1997) in addition to a role in the biogenesis of miRNAs derived from introns in *Drosophila* (Flynt *et al.* 2010). If Dis3 was to play a role in the maturation of *miR-982-5p* then one may expect the level of its partner to decrease to a similar extent. Unfortunately *miR-982-3p* was only detected in one knockdown replicate and therefore confident conclusions about its expression cannot be drawn. However, in comparing the remaining knockdown replicate to the 4 control replicates a fold decrease of -4.6 fold is observed which is comparable to the -4.53 fold decrease shown for *miR-982-5p*. For further confirmation this would have to be followed up by qRT-PCR. However

the information presented strongly suggests that that Dis3 is involved in the maturation process of *miR-982*.

### **3.12.4 How might Dis3 be targeted to specific miRNAs?**

The RNA seq data suggests that only a subset of *Drosophila* miRNAs are sensitive to the activity of Dis3. This infers mechanisms of specificity that may come through physical association between Dis3 and targeting co-factors or through nucleotide tagging of the miRNAs. Terminal tagging is known to promote the activity of the exosome in the nucleus. This is illustrated by TRAMP mediated terminal adenylation targeting aberrant 23S rRNA intermediates (LaCava *et al.* 2005), unspliced pre-tRNAs (Wlotzka *et al.* 2011; Gudipati *et al.* 2012) and CUTs (Wyers *et al.* 2005) for exosome-mediated decay (Schmidt and Butler 2013). A similar mechanism may exist in the cytoplasm to target the exosome/Dis3 to specific miRNAs. For example, Liu *et al.* identified a targeting mechanism by which Argonaute bound pre-miRNAs were 3' uridylated which directed their degradation (Liu *et al.* 2014); a similar mechanism for *pre-let-7a* degradation also exists where its oligouridylation stimulates decay by Dis3L2 (Chang *et al.* 2013; Ustianenko *et al.* 2013). It is possible that 3' terminal tagging of *pre-miR-982* may recruit Dis3 if it is indeed involved in its biogenesis. This is more likely to be through monouridylation which has been shown to stimulate *pre-let-7a* maturation (Heo *et al.* 2012). This is opposed to oligouridylation which has been shown to stimulate pre-miRNA clearance rather than its maturation.

Modifications of the 3' ends of mature miRNAs have also been shown to affect their stability (Kato *et al.* 2009). It is possible that *miR-252-5p* is terminally tagged resulting in the specific recruitment of Dis3. This tagging could be performed by Tailor which has recently been identified to be a 3' uridyl transferase in *Drosophila* (Reimao-Pinto *et al.* 2015). Unfortunately, due to the low read count it was not possible to make any conclusions whether *miR-252-5p* was terminally tagged. *miR-252-5p* had a relatively low read count and as the incidence of tailing may not be high this would have been insufficient to make any confident conclusions. In

addition to the low read count, the 2 nucleotides immediately downstream of the 5' end of *miR-252-5p* in the pre-miRNA are uridines and therefore it would be impossible to differentiate between template and non-template additions. It would only become clear if the uridine tail consistent of more than 2 Us. Despite this the read sequences were analysed (presented in Figure 3.20). There was a possibility of terminal tagging but to draw any confident conclusions the experiment would require repeating with greater depth and to include longer RNAs (up to 30nt).

In addition to nucleotide tagging another possibility for Dis3 targeting could be through intrinsic instability elements. For example, AU rich elements are known to be involved in exosome mediated mRNA turnover (Shaw and Kamen 1986; Chen and Shyu 1995; Chen *et al.* 2001) and similarly miRNAs with high AU/UA dinucleotide densities show the shortest half lives in primary human neuronal cells (Sethi and Lukiw 2009). Additionally specific seed regions (Rissland *et al.* 2011) or the most 3' bases (Bail *et al.* 2010) have been shown to control the stability of specific miRNAs. It is possible that specific sequences within *miR-252-5p* are responsible for its Dis3-mediated degradation as *miR-252-5p* contains 3 UA dinucleotides in the first 10 bases. However, a greater number of potential Dis3 targets or in depth biochemical analysis would be required to identify sequence elements.

### **3.12.5 Issues with 2S contamination in the miRNA seq limits the number of miRNA reads.**

Although the miRNA-sequencing experiments were successful in identifying a potential target of Dis3 in *miR-252-5p* it had its limitations. The major limitation being that only 1.12% of all reads mapped to mature miRNAs which 83% of the reads mapping to the 2S rRNA. Due to the relatively small number of miRNAs in *Drosophila* (currently 426) the coverage was sufficient to determine reproducible fold changes in specific miRNAs. Only 109 of the 426 known miRNAs were confidently detected as being expressed in this data and although tissue specific expression is highly likely one may have expected more to be detected. It is likely that the

miRNAs expressed at low levels within the wing imaginal disc were missed as a result of the reduced coverage. These miRNAs may have also been susceptible to Dis3 regulation and therefore there additional Dis3 targets may have been missed.

To identify further Dis3-sensitivie miRNAs the major contaminant, the 2S rRNA, would have to be depleted. All other ribosomal and long RNAs were removed during the library preparation through size selection. However, the 2S rRNA is only 30nt in size so a proportion was retained during the size selection. Due to its abundance it mopped up 83% of the sequencing reads. It has been shown that the 2S rRNA can be removed using specific oligo coupled beads prior to the small RNA library preparation (Aspden *et al.* 2014) and therefore this should be used to minimise the level of 2S rRNA contamination should the experiment be repeated.

### **3.13 Acknowledgements.**

ARK Genomics conducted the library preparation, miRNA-sequencing and initial data analysis up to the generation of the normalised read counts for each miRNA. Dr. Sarah Smalley carried out the initial mapping of sequencing reads using miRDeep2 and the identification of the potential novel miRNA sequences. These sequences were then analysed and taken forward for validation by myself.

## Chapter 4: A novel role for the 3'-5' exoribonuclease

### Dis3L2 in cell proliferation

#### 4.1 Introduction

In the previous chapter the role of Dis3 in *Drosophila* wing development was investigated and a function in regulating miRNA stability was identified. The *Drosophila* genome encodes a second Dis3 paralogue in Dis3L2. Dis3L2 has been largely in the shadow of Dis3 and therefore there has been no investigation into its function within a developing organism. Dis3L2 has come to interest recently following its mutation in Perlman syndrome, a human condition characterised by foetal overgrowth (Astuti *et al.* 2012; Morris *et al.* 2013). This therefore presented an excellent opportunity to investigate the specific pathways in which Dis3L2 functions in an attempt to generate a model system in which to study the effects observed in Perlman syndrome.

The few studies have looked into the basal activity of Dis3L2 and have confirmed that it is a 3'-5' exoribonuclease which shows strict cytoplasmic localisation in both *S. pombe* (Malecki *et al.* 2013) and human tissue culture cells (Lubas *et al.* 2013). It has also been shown to function independently from the exosome as co-immunoprecipitation experiments show an absence of any exosome subunits, which is in complete contrast to its paralogue, Dis3 (Lubas *et al.* 2013). These experiments also failed to identify any co-factors that directly interact with Dis3L2, suggesting at least in these conditions that it acts alone to degrade target substrates. However, high-throughput mapping of *Drosophila* protein interactions identify Tailor and Aluminium tubes as potential interaction partners (Guruharsha *et al.* 2011), although this is yet to be experimentally verified. Interestingly Tailor has been shown to be a uridyl transferase (Reimao-Pinto *et al.* 2015). With recent work showing that 3' terminal uridylation may play a role in targeting *pre-let-7a* to Dis3L2 (Chang *et al.* 2013; Ustianenko *et al.* 2013) together with the

increase of uridylated transcripts in *dis3L2* deficient cells (Malecki *et al.* 2013; Thomas *et al.* 2015); this provides an exciting potential interaction which should be investigated further.

Consistent with the phenotypes observed in Perlman syndrome the loss of Dis3L2 in human tissue culture cells has been shown to increase cell number (Astuti *et al.* 2012). GO analysis performed on misexpressed genes detected by RNA-sequencing in Dis3L2-depleted HeLa cells revealed an enrichment of RNAs involved in the cell cycle (Lubas *et al.* 2013). In addition to an apparent role in cell cycle regulation Dis3L2 has also been shown to clear truncated mRNA intermediates which are generated during apoptosis (Thomas *et al.* 2015). Interestingly *dis3L2*-depleted human cells survive apoptotic stimuli better than control cells whilst overexpression of *dis3L2* was shown to enhance apoptosis (Thomas *et al.* 2015). Taken together this work identified potential functions for Dis3L2 in the control of the cell cycle and apoptosis.

The control of cell proliferation and tissue growth is a conserved, co-ordinated and highly regulated process. It is critical to ensure that the correct level of proliferation is maintained so that ultimately the organs grow to the correct sizes and maintain symmetry between the left and right sides of the organism (Conlon and Raff 1999; Weinkove and Leivers 2000). This is demonstrated in *Drosophila* where rapid proliferation of the wing imaginal discs is required during early development; however as the tissues reach their critical size in the late stage of larval development proliferation usually slows as the cells accumulate in G2 phase in preparation for metamorphosis (Neufeld *et al.* 1998). The 'green light' for metamorphosis is only given when the tissues have proliferated sufficiently to reach their critical size (Shingleton 2010). Hormonal signalling, including insulin-like growth factors, function to delay the initiation of metamorphosis to allow the tissues sufficient time to reach their required size (Colombani *et al.* 2012). This provides a neat example of how the control of proliferation, tissue growth and organism development are tightly co-ordinated.

It is also essential to proliferative control as those cells that evade the regulatory stimuli and therefore become capable of uncontrolled proliferation imitate tumour development. In line

with this, a number of genes involved in growth and proliferation have been implicated in cancer progression (Cantor and Sabatini 2012). Proliferation is also required to repair tissue damage or to replace neighbouring apoptotic cells in a process termed 'compensatory proliferation' (Baehrecke 2006; Fan and Bergmann 2008). The process of proliferation is controlled by an intricate network of cell signalling which is still being explored today. Some of these mechanisms of control hold true for all cell types whilst others are cell type specific. These involve a number of checkpoints during the cell cycle together with signalling through growth factors and kinases. Numerous studies have shown that loss of or gain of function of a single one of the members involved in these pathways can be sufficient to lose control over proliferation. An aim of the work in this thesis was to ask if Dis3L2 functions to control proliferation and tissue growth through one of these highly conserved pathways.

The current understanding of the cellular functions of Dis3L2 is based on work performed in immortalised tissue culture cells or single cell organisms; systems in which it is much more difficult to elucidate a regulatory mechanism by which Dis3L2 may co-ordinate tissue growth. The use of multicellular organisms with full developmental potential is critical in investigating protein function, especially as the human condition suffered by those carrying *dis3L2* mutations is associated with foetal development and tissue growth. The work presented in this chapter used the wing imaginal discs in an attempt to develop a model system in which to study the role Dis3L2 plays in the control of tissue growth. This system could ultimately be used to begin understand the defects observed in Perlman syndrome and how the loss of Dis3L2 contributes towards them.

## 4.2 Aims

The aims of this chapter were to explore a potential role of Dis3L2 in the regulation of cell proliferation and tissue growth. Due to its implication in Perlman syndrome the hypothesis was that the loss of *dis3L2*, directed by RNA interference would result in increased cell and/or tissue growth. Previous work has been restricted to single cells either in human cell culture or in yeast cells and therefore this work aimed to characterise the function of Dis3L2 in a developmental model; the wing imaginal disc.

The specific aims of this chapter were to:

- To investigate if Dis3L2, like its paralogue Dis3, is critical for *Drosophila* viability.
- To use the *GAL4-UAS* system in the wing imaginal discs to determine if Dis3L2 plays a role in the regulation of tissue growth as is indicated by the Perlman Syndrome phenotypes
- To characterise any phenotypes should they be observed.

## 4.3 Validation of the expression of the 2 predicted isoforms of Dis3L2 in *Drosophila*.

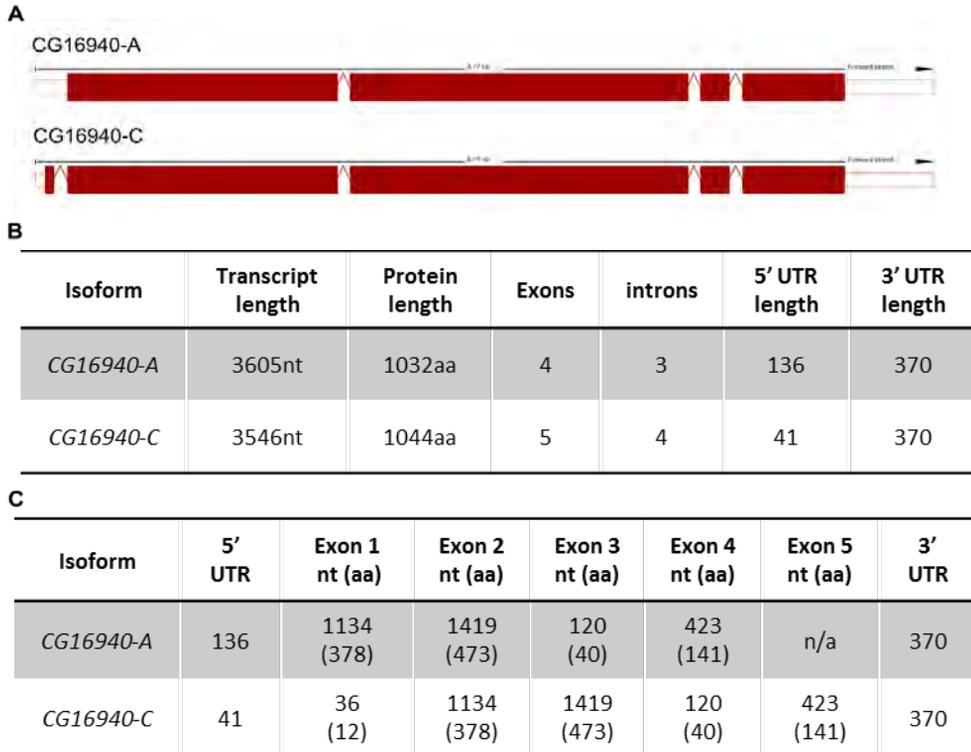
A Basic Local Alignment Search Tool (BLAST) search of the full length human DIS3L2 peptide sequence from transcript 201 identified 2 genes that showed extensive homology; *dis3* and *CG16940*. Due to Dis3 being a paralogue of Dis3L2 this is likely to identify *CG16940* as the gene encoding *Drosophila* Dis3L2. *CG16940* showed the highest BLAST score of 393.275 and an E value of 4.8626e-109 indicating a highly significant alignment. The final 600 amino acids show the highest level of conservation, therefore this region is likely to include the RNB domain which is highly conserved throughout the Dis3 family. *CG16940* showed 39% identity and 57.6% positive matches (Figure 4.1). Additionally, DIS3L2 is signified as the human orthologue of *CG16940* in the Flybase database and therefore *CG16940* was named Dis3L2.

Two isoforms of Dis3L2 are predicted to be expressed which occur due to alternative start codons. These are predicted to differ by 12 amino acids resulting in a 1.4kDa difference in the size of the protein (Figure 4.2). The extra 12 amino acids are encoded by an additional small exon at the 5' end of the RNA for *dis3L2-RC* (Figure 4.2A lower). Although Dis3L2-PC is larger in molecular weight due to the additional exon *dis3L2-RA* is the longer of the two RNAs due to a longer 5' UTR (Figure 4.2B-C). This could potentially lead to differential regulation of this isoform.

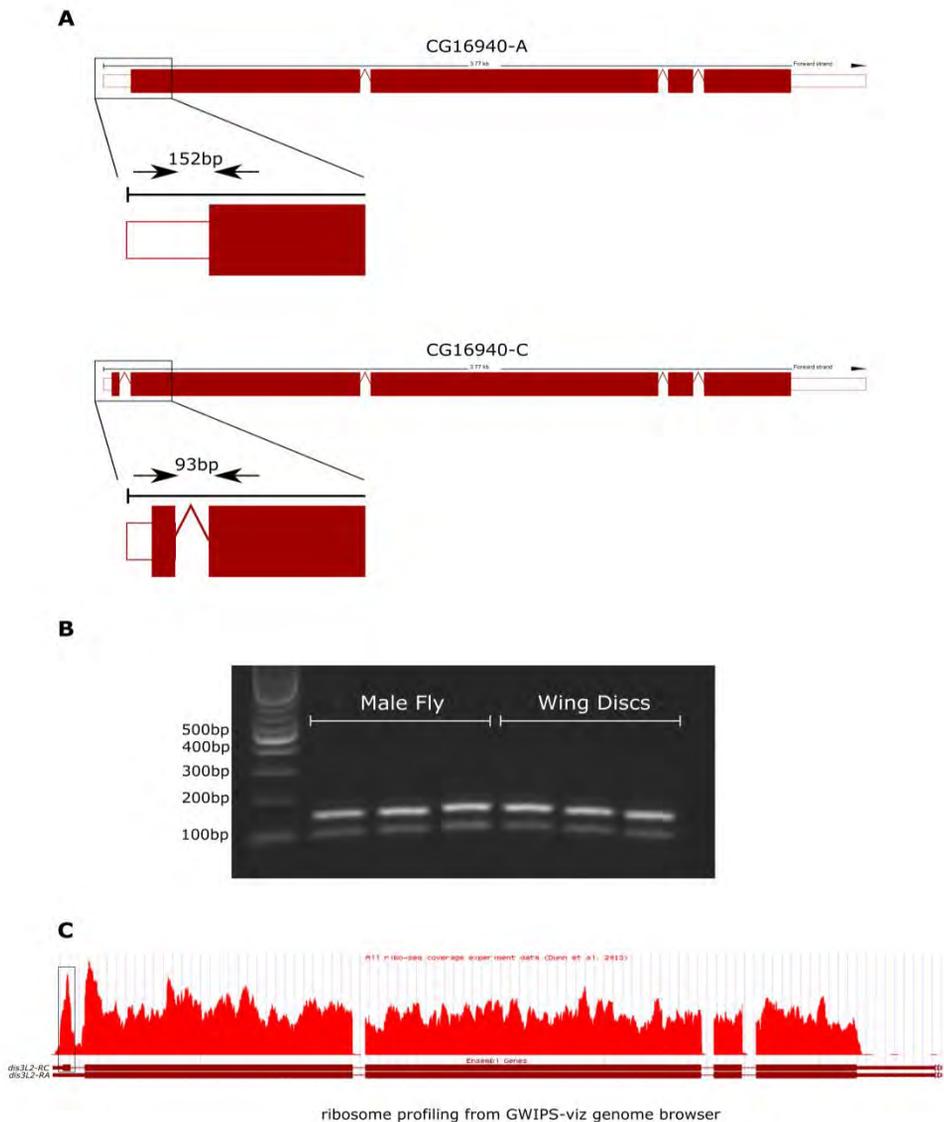
To investigate if the two isoforms are indeed expressed at the RNA level PCR primers were designed to span the unique 59bp region in *dis3L2-RA* meaning if both isoforms are indeed expressed two products would be expected (Figure 4.3A). RNA was extracted from both whole flies and wing imaginal discs, reverse transcribed into cDNA and subsequently put into a polymerase chain reaction with the isoform primers. Running the PCR products on an agarose gel identified two products of the expected size for each isoform (Figure 4.3B). Products of ~152bp and ~93bp were present relating to *dis3L2-RA* and *-RC* respectively. This therefore confirmed that both isoforms were indeed expressed at the RNA level. The absence of a suitable antibody unfortunately means it was not possible to identify both isoforms are expressed at the protein level. However, ribosome profiling (Dunn *et al.* 2013) shows a strong peak for ribosome presence over the *dis3L2-RC* specific exon indicating it is indeed translated (Figure 4.3C). Whether the two isoforms have different activities or functions is thus far unknown.

Isoform	Score	E value	Identities in matched region	Positives in matched region
CG16940-PA	393.275	4.86x10 <sup>-109</sup>	221/556 (39%)	326/566 (57.6%)
CG16940-PC	392.889	7.3x10 <sup>-109</sup>	221/556 (39%)	326/566 (57.6%)
Dis3-PB	349.362	9.4x10 <sup>-96</sup>	229/660 (34.7%)	344/660 (52.1%)
Dis3-PA	349.362	9.4x10 <sup>-96</sup>	229/660 (34.7%)	344/660 (52.1%)

**Figure 4.1: Blast search using the human Dis3L2 peptide sequence to confirm the *Drosophila* orthologue.**



**Figure 4.2: *dis3L2* isoform information. (A)** The two *dis3L2* (CG16940) isoforms differ by the presence of an additional exon at the 5' end of CG16940-C caused by the use of an alternative start codon proximal to the start codon used by CG16940-A. The alternative start codons are present in both isoforms therefore this represents alternate translation. The increase in transcript size results from a retained intron (59bp in size) in CG16940-A which is spliced out in isoform CG16940-C. **(B)** Overall transcript information for the two isoforms. **(C)** Detailed information about the two isoforms including the size of each exon and the number of amino acids encoded by it.



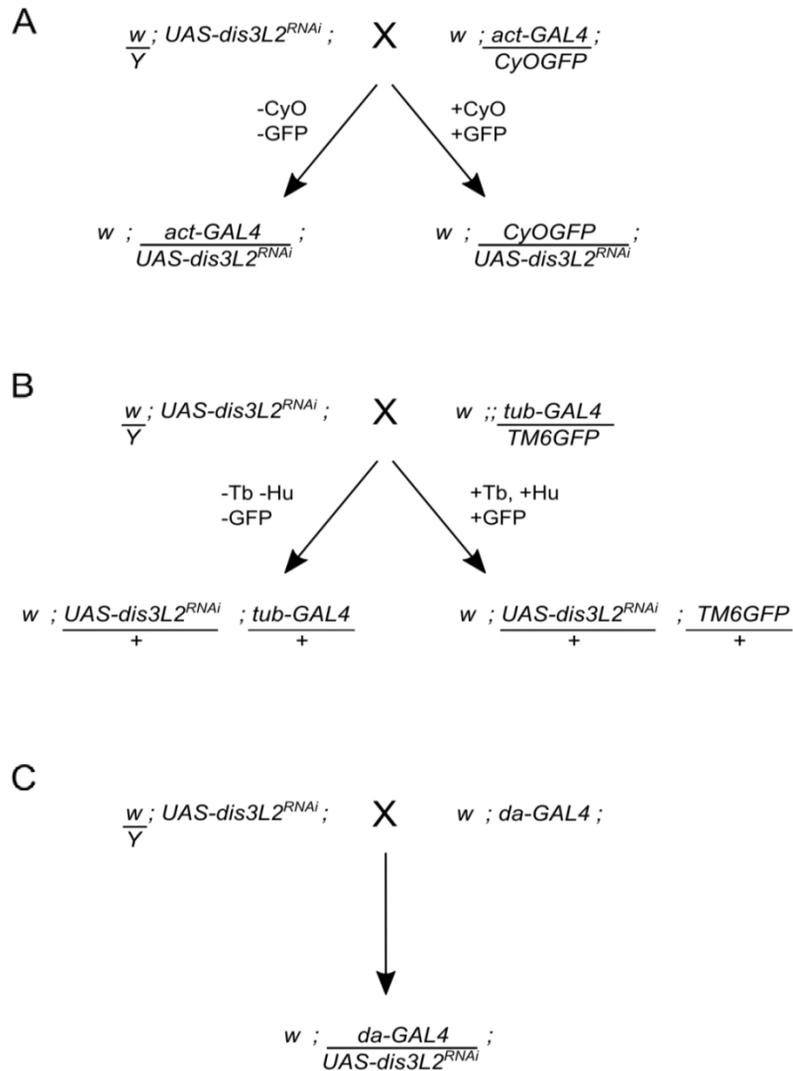
**Figure 4.3: Both *dis3L2* isoforms are expressed in *Drosophila*.** (A) Visual representation of the primers used to identify the expression of both *dis3L2* isoforms at the RNA level. (B) PCR in triplicate using the isoform differentiating primer pairs (A) on cDNA produced from both whole flies (Male Fly) and wing imaginal discs (Wing Discs). A 100bp ladder was loaded along side for size determination. (C) Ribosome profiling information for the GWIPS-viz genome browser showing ribosome presence on the additional exon of *dis3L2-RC* indicating it is indeed translated.

#### 4.4 Dis3L2 is not required for *Drosophila* viability

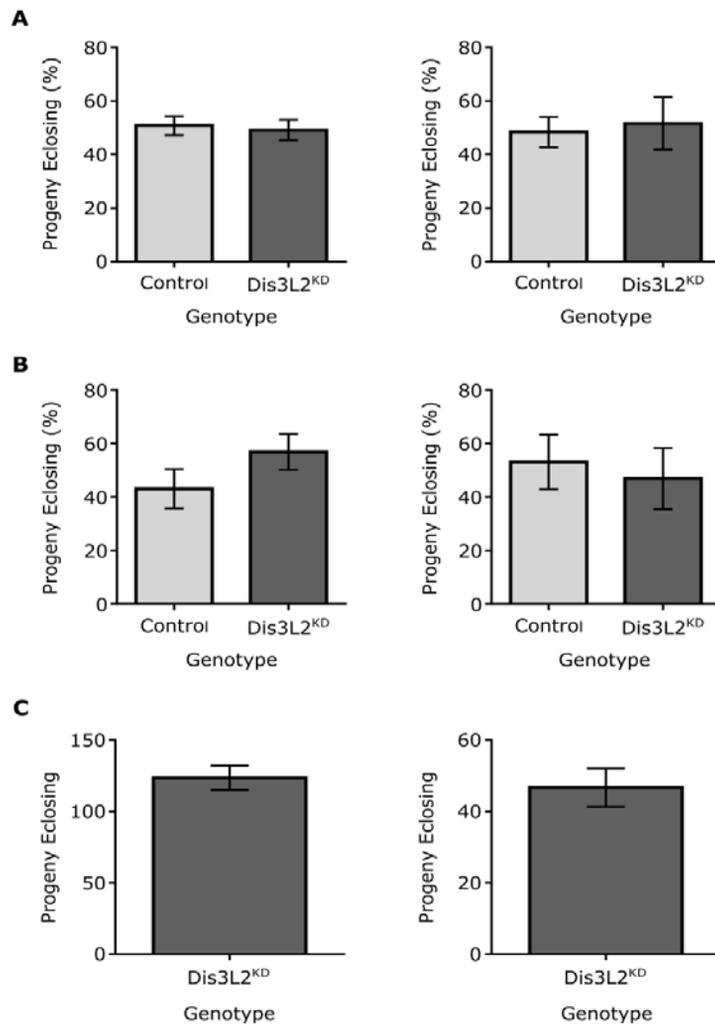
In the previous chapter a complete requirement for Dis3 for organism viability was identified, but it was unknown if its paralogue, Dis3L2, was also critical for *Drosophila* survival. Previous work in *S. pombe* and human cell lines showed no sign of cell death in *dis3L2* mutant or knockdown lines but this work was on single cells rather than a developing organism. To test whether Dis3L2 is required for viability, ubiquitous GAL4 drivers were used to knockdown *dis3L2* throughout the multicellular organism. Two separate *UAS-dis3L2<sup>RNAi</sup>* lines, which both target the two isoforms, were used for each GAL4 driver. *tubulin(tub)-GAL4*, *actin(act)-GAL4* and *daughterless(da)-GAL4* were selected as they had been used successfully in previous work (Chapter 3). The GAL4 insertions for both *tub-GAL4* and *act-GAL4* were homozygous lethal and therefore were used as heterozygotes over identifiable balancers. If any lethality occurs following the knockdown of *dis3L2* the progeny inheriting the balancer, rather than the GAL4 driver should be viable. This therefore provided an internal control to ensure the lack of knockdown progeny would not simply be due to the parents not mating. These simple, one generation crosses were performed at 25°C and are shown in Figure 4.4.

Knockdown of *dis3L2* throughout the organism (90% efficiency) using either *tub-GAL4* or *act-GAL4* resulted in no observed lethality or immediately observable phenotypes. Both expected genotypes were present; those of the internal control (containing the balancer) and the knockdowns with no significant difference between their proportions in each case when either *UAS-dis3L2<sup>RNAi</sup>* line was used (Figure 4.5A/B,  $p > 0.05$ ). To confirm these findings a third ubiquitous GAL4 driver was used to drive the knockdown of *dis3L2*. Knockdown of *dis3L2* using *da-GAL4* again resulted in viable knockdown offspring (Figure 4.5C) with the number of offspring within the expected range for the cross conditions. This therefore indicates that unlike Dis3, Dis3L2 is not required for organism viability. These knockdown progeny also showed no immediately observable phenotypes giving the impression that Dis3L2 is not critical for the development of any particular tissues. These findings were consistent with the work in

*S. pombe* (Malecki *et al.* 2013) and human cell lines (Lubas *et al.* 2013; Thomas *et al.* 2015) showing no effect on cell viability following *dis3L2* depletion.



**Figure 4.4: Cross schemes to achieve ubiquitous knockdown of *dis3L2* performed at 25°C.** (A) Mating scheme used to drive knockdown of *dis3L2* throughout the *actin* expression domain. *UAS-dis3L2<sup>RNAi</sup>* males were crossed to *act-GAL4* females which were balanced over *CyOGFP*. Knockdown adult progeny (*; act-GAL4/UAS-dis3L2<sup>RNAi</sup>*;) were identified by the absence of curly wings (CyO) whilst the knockdown larvae were detected by the absence of GFP. (B) Cross scheme used to drive knockdown of *dis3L2* throughout the *tubulin* expression domain. *UAS-dis3<sup>RNAi</sup>* males were crossed to *tub-GAL4* females which were balanced over TM6GFP. Knockdown adult progeny (*; UAS-dis3L2<sup>RNAi</sup>/+ ; tub-GAL4/+*) were identified by the absence of the humeral phenotype whilst the knockdown larvae were detected by the absence of the tubby phenotype. (C) Mating scheme used to drive knockdown of *dis3L2* throughout the *daughterless* expression domain. *UAS-dis3L2<sup>RNAi</sup>* males were crossed to *da-GAL4* females in a cross where 100% of the progeny would carry *dis3L2* knockdown.



**Figure 4.5: Ubiquitous knockdown of *dis3L2* by 90% does not affect organism viability.** Ubiquitous knockdown of *dis3L2* using either *act-GAL4* (A) or *tub-GAL4* (B) has no effect on organism viability when either *UAS-dis3L2<sup>RNAi</sup>* line is used (left = v51854 , right = v100322). (A) Control = ; *CyOGFP/UAS-dis3L2<sup>RNAi</sup>*; Dis3L2<sup>KD</sup> = ; *act-GAL4/UAS-dis3L2<sup>RNAi</sup>*; n≥157, error bars represent standard error. (B) Control = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *TM6/+*, Dis3L2<sup>KD</sup> = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *tub-GAL4/+*. n≥102, error bars represent standard error. (C) The ubiquitous knockdown of *dis3L2* using *da-GAL4* shows an expected number of progeny when either *UAS-dis3L2<sup>RNAi</sup>* was used. Dis3L2<sup>KD</sup> = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *da-GAL4/+*.

## 4.5 Loss of Dis3L2 results in an increased number of cells in the wing

Having shown that Dis3L2 is not critical for *Drosophila* viability or the overall development of adult structures the wing imaginal discs were then used to study the potential role for Dis3L2 specifically in regulating tissue growth. The wing imaginal discs were ideal tissues in which to study the role of a specific protein as the critical pathways for their growth and development are well characterised and highly conserved. The *GAL4-UAS* system was used to restrict the knockdown of *dis3L2* specifically to the wing imaginal disc. Two wing disc specific *GAL4* drivers were used; *69B-GAL4* and *nubbin(nub)-GAL4* which drive throughout the disc and specifically in the wing pouch respectively (methods Figure 2.3).

### 4.5.1 Successful knockdown of *dis3L2* in the wing imaginal disc.

To ensure *dis3L2* RNA was depleted in the knockdown wing imaginal discs qRT-PCR was used to assess the levels of *dis3L2* compared to parental controls. Knockdown was assessed in the *69B-GAL4* driven knockdown discs (at 25°C) as *dis3L2* will be affected throughout the whole tissue, whereas the *nub-GAL4* driven knockdown would occur in the wing pouch and therefore using the whole tissue dilutes any effects giving a misleading result. RNA was extracted from 30 wandering L3 wing imaginal discs per biological replicate and reverse transcribed using random primers. Both parental lines (*UAS-dis3L2<sup>RNAi</sup>* and *69B-GAL4*) were used as controls as they contained the most similar genetic background to the knockdown progeny. qRT-PCR showed a significant reduction in the ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+* wing imaginal discs when compared to both parental controls (Figure 4.6A). There was no significant difference in the levels of *dis3L2* between the two parental lines and therefore they were grouped together to give a general 'control' group. The *dis3L2* knockdown wing imaginal discs showed a significant 80% reduction in the level of *dis3L2* compared to the control group (Figure 4.6B). Unfortunately due to the absence of a *Drosophila* anti-Dis3L2 antibody it was not possible to confirm this knockdown at the protein level. However, similar knockdown efficiency of the 5'-

3' exonuclease Pacman translated to a comparable knockdown at the protein level (Waldron 2014). Similarly a knockdown (80%) at the protein level for Dis3 was shown in the previous chapter.

#### **4.5.2 Knockdown of *dis3L2* results in overgrowth of the wing.**

Having shown a successful knockdown of *dis3L2* in the wing imaginal discs by qRT-PCR a reverse genetics approach was used to investigate a potential role for Dis3L2 in regulating wing development. Knockdown of *dis3L2* was driven by the two independent wing disc drivers; *69B-GAL4* and *nub-GAL4*. Two *UAS-dis3L2<sup>RNAi</sup>* lines were again used with each GAL4 driver. Crosses were set up alongside control crosses of each parent stock at 25°C in uncrowded conditions. To ensure vials remained uncrowded a maximum of 5 virgin females were used per cross and only allowed 48 hours to lay. Eclosing adults were tipped into new vials and aged to between 1 and 2 days of age at 25°C. Wings were then cut off the males and females using microscissors, stored separately in isopropanol for ≥1 hour and mounted in DPX. Male and female wings were kept separate at all times due to the difference in size between them; female wings are on average around 25% larger than male wings. To avoid duplication of data a single wing was cut from each fly; for consistency this was the "left-hand" wing each time. An initial control experiment was performed where both the left and right wing were cut and measured and no significant difference was observed between them (Figure 4.7B) meaning there would be no bias by taking a single wing.

Knockdown of *dis3L2* throughout the wing imaginal disc using the *69B-GAL4* driver resulted in wings that were 120% the size of parental controls (Figure 4.7Diii). There was no significant difference in wing area between the *UAS-dis3L2<sup>RNAi</sup>* and *69B-GAL4* parents and therefore they were grouped to give a 'control' group. The level of overgrowth was consistent between males and females showing wings 120.7% and 124% the area of control wings respectively. However, the *69B-GAL4* driver also drives in the epidermis, together with the eye and ventral thoracic imaginal discs (Brand 1997). Therefore to ensure the observed overgrowth was due to a

specific effect within the wing disc a more specific GAL4 driver was used. *nub-GAL4* drives specifically in the wing pouch of the wing imaginal disc and therefore *dis3L2* knockdown will be restricted to only those cells fated to form the adult wing blade. Knockdown with *nub-GAL4* also resulted in a significant overgrowth of both male (Figure 4.7Eiii) and female wings indicating that the cause of the overgrowth was indeed within the cells of the wing pouch and therefore specific.

Although the control and parental crosses were kept in uncrowded conditions and on constant food the size of the fly can vary. If the size of the fly varies then the wing area itself may also subsequently vary in area. Therefore to ensure the observed overgrowth was not simply due to the knockdown flies being larger than their parents the mass of individual flies was measured. Individual 1-2 day old flies were weighed in pre-weighed Eppendorf's and the mass of each was calculated. The wing area for each genotype was then normalised to the average mass of the genotype to give a 'wing to weight' ratio ( $\text{mm}^2/\text{mg}$ ) to take into account the potential variation in size across cross replicates. On normalising to fly mass significant overgrowth was still observed for both the *69B-GAL4* and *nub-GAL4* driven knockdown, remaining at 120% overgrowth (Figure 4.7Div/Eiv). This therefore provided strong evidence that the loss of *dis3L2* within the wing imaginal disc results in the overgrowth of the adult wing.

In *Drosophila*, unlike tissue culture cells, scrambled shRNAs are rarely used, however, to ensure the observed effects were not due simply to the expression of a shRNA in the wing disc a shRNA to GFP was driven using the same drivers. *UAS-EGFP<sup>RNAi</sup>* was driven by both *69B-GAL4* and *nub-GAL4* and when normalised to fly mass they were not significantly larger the parental controls (Figure 4.7C). These data further show the overgrowth is due to the loss of *dis3L2* rather than an indirect response to RNAi. Finally, the parental flies were homozygous for the *UAS* and *GAL4* insertions whilst the knockdown progeny were heterozygous. To ensure this was not the driving factor of the wing overgrowth the wings of ; *nub-GAL4/+* ; *69B-GAL4/+*

were measured. Again no significant difference in normalised wing area was observed compared to the homozygous parental controls ( $p>0.05$ ). These additional control measurements were included together with the parental measurements in the 'control' group.

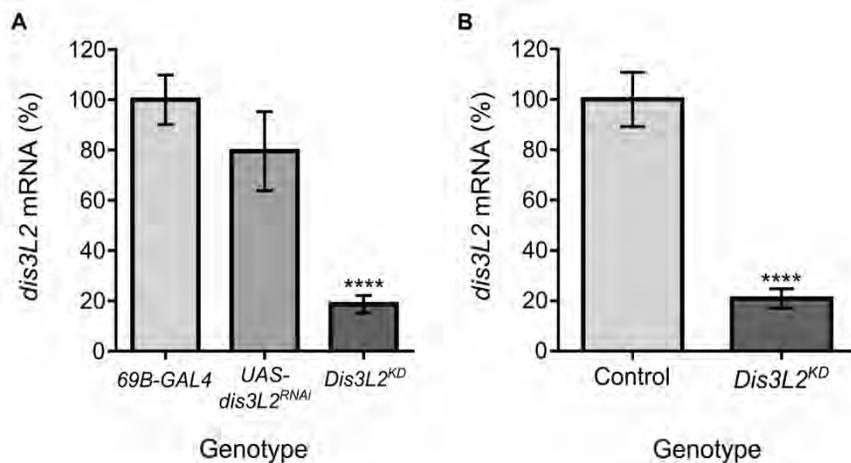
Having observed wing overgrowth following the knockdown of *dis3L2* using the wing specific drivers it was of interest to see if the ubiquitous knockdown resulted in an overall increase in fly size. To confirm that the wing overgrowth was still present when *dis3L2* is knocked down ubiquitously, as would be expected, flies were aged, their wings cut and subsequently measured for the three ubiquitous drivers. For the *tub-GAL4* driven knockdown only the UAS-*dis3L2<sup>RNAi</sup>* parent was used due to the *GAL4* parent containing the Tubby containing TM6 balancer which changes the shape of the fly. Similarly the *act-GAL4* parent could not be used due to containing the CyO balancer which makes the wings unmeasurable. Adult *Drosophila* with ubiquitous *dis3L2* knockdown driven by any of the three drivers show significantly larger wings (Figure 4.8 A-C). Knockdown of *dis3L2* by *act-GAL4* resulted in 36% overgrowth (Figure 4.8A) whilst *tub-GAL4* driven knockdown resulted in 26% overgrowth (Figure 4.8B). Finally the *da-GAL4* driven knockdown resulted in wings that were overgrown by 54%, the largest overgrowth observed in this study (Figure 4.8C). The ; UAS-*dis3L2<sup>RNAi</sup>*/+ ; *da-GAL4*/+ flies were the smallest within the study and therefore this is likely to account for the increased normalised measurement.

To assess if loss of *dis3L2* throughout the organism results in overgrowth of the entire organism the *act-GAL4* was used to drive ubiquitous knockdown. *act-GAL4* was chosen due to the stock containing the CyO balancer chromosome which give sibling controls an easily observable phenotype (curly wings) which does not affect the shape or size of the fly (Figure 4.4A). This is in contrast to the *tub-GAL4* stock which contains the TM6 chromosome containing the Tubby phenotype and therefore make the sibling controls unsuitable for use. When directly comparing the mass of the flies it was important to ensure they were cultured under exactly the same conditions, using the sibling controls allows this. Eclosing adults were

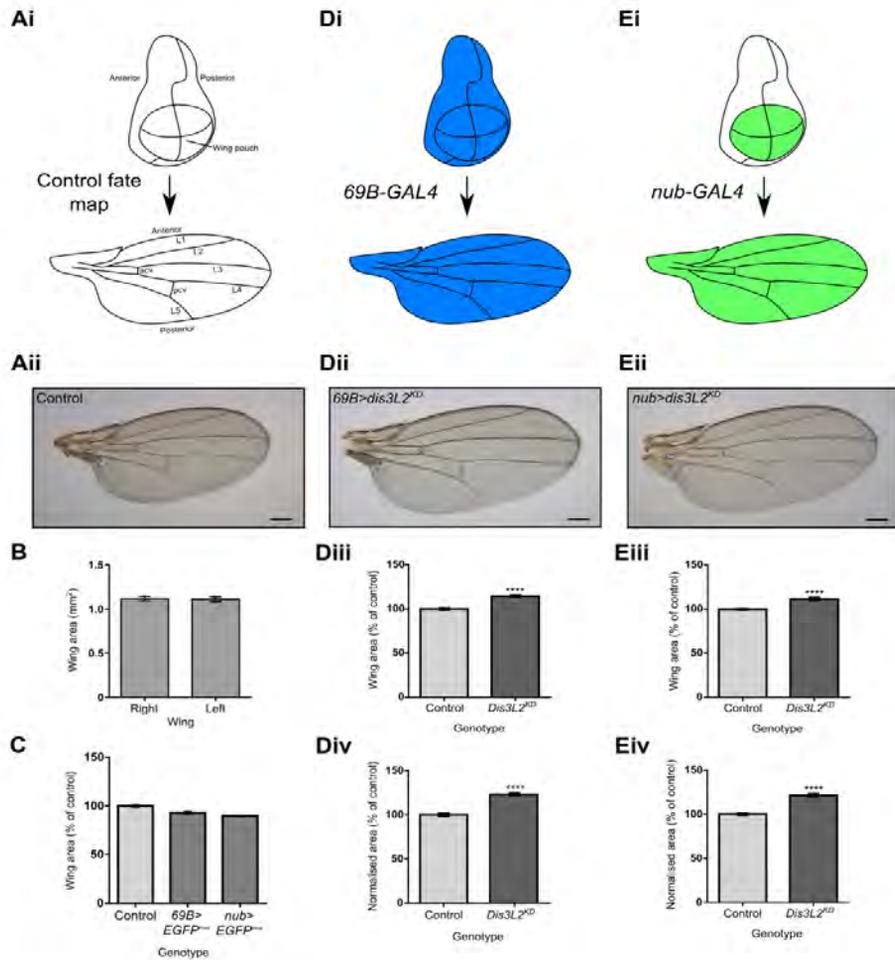
aged to between 1 and 2 days of age at 25°C and weighed using the same method as used previously. Knockdown flies were not significantly larger than their sibling controls (Figure 4.8D  $p > 0.05$ ) suggesting that loss of *dis3L2* throughout the organism does not induce overgrowth of the whole fly. This therefore indicates a level of tissue specificity in the activity of Dis3L2 which would be consistent with the phenotypes of Perlman syndrome; showing overgrowth of specific organs such as the liver and kidney.

#### **4.5.3 Knockdown of *dis3L2* results in hyperplasia of the wing.**

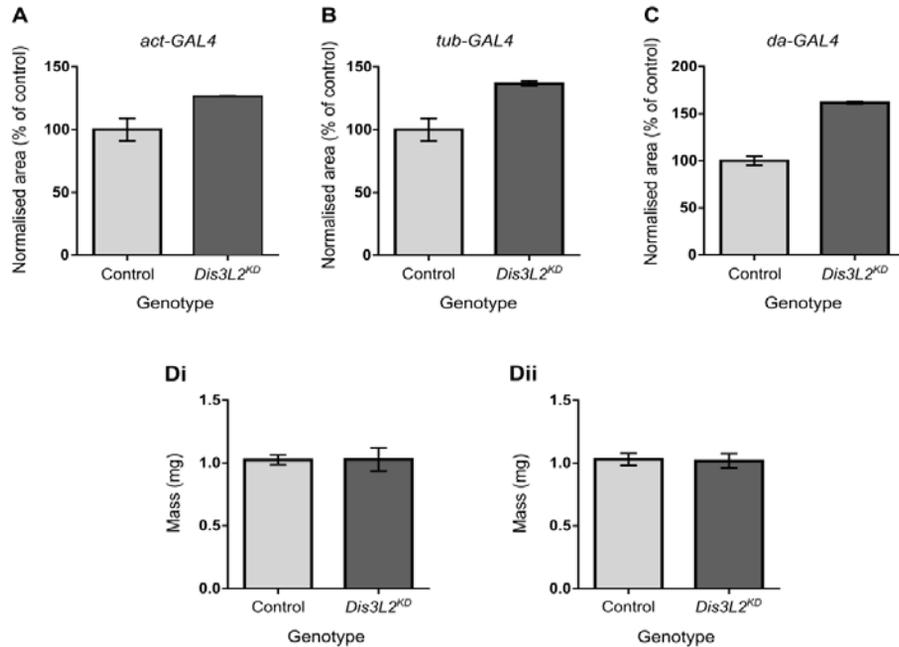
The increased area of the adult wing following *dis3L2* depletion could be caused by one of two factors; hyperplasia (an increase in cell number) or hypertrophy (increase in cell size). To identify which was the cause the number of cells and cell area was calculated for the parental control, *EGFP<sup>RNAi</sup>* and *dis3L2* knockdown wings. Each cell in the adult wing protrudes a single hair which can be accurately counted when wings are photographed under a 20x objective. The number of hairs (and therefore cells) within specific anterior and posterior 10,000 $\mu\text{m}^2$  regions (Figure 4.9A) were counted and the average was taken. This value was then used together with the area of the wing to estimate the total number of cells within the adult wing. Additionally the total number of cells within the defined regions can be used to calculate the cell density and therefore an estimate of average cell size. This indicated a significant increase in the number of cells within the wing when *dis3L2* was knocked down in the wing using either the *nub-GAL4* (Figure 4.9E) or *69B-GAL4* (Figure 4.9C) driver compared to the parental and *EGFP<sup>RNAi</sup>* controls. However, there was no significant difference in the density of cells within the restricted regions indicating there is no difference in the size of the cells (Figure 4.9B/D). Overall this strongly indicates a proliferative (hyperplastic) phenotype rather than an expansion (hypertrophic) phenotype; giving an indication that Dis3L2 is normally required to prevent over proliferation of the wing cells.



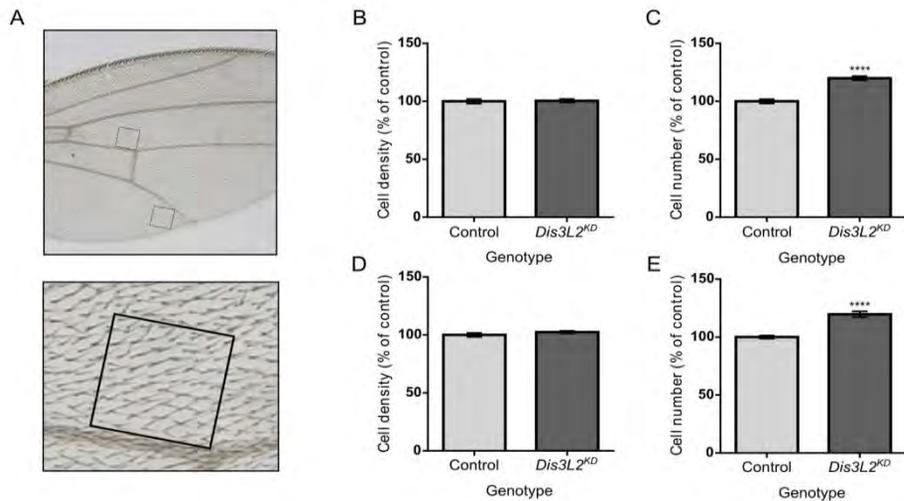
**Figure 4.6: Successful knockdown of *dis3L2* in the wing imaginal disc.** *dis3L2* knockdown wing imaginal discs (*UAS-dis3L2*<sup>RNAi</sup>/+ ; *69B-GAL4*/+) show a significant reduction in *dis3L2* expression by 82% compared to the *69B-GAL4* parent and 77% compared to the *UAS-dis3L2*<sup>RNAi</sup> parent (A) whilst the parents showed no significant difference between them. As a result the parents were grouped to give a general control (*69B-GAL4* and *UAS-dis3L2*<sup>RNAi</sup>) resulting in the final determination of an 80% knockdown (B) when *dis3L2* was knocked down using *69B-GAL4*. n≥7, error bars represent standard error, \*\*\*\* = p<0.0001.



**Figure 4.7: Knockdown of *dis3L2* in the wing imaginal disc results in specific overgrowth.** (A) Wild type representation of the developmental axis and regions of a *Drosophila melanogaster* wing and wing imaginal disc together with a parental control (*UAS-dis3L2<sup>RNAi</sup>*). Scale bar = 200µm. (B) There is no significant difference between the area of the right and left wing on control flies meaning no bias is derived from taking a single wing of each fly. n=26, error bars represent 95% confidence limits, p=0.6698. (C) Driving an RNAi to *egfp* with either the *69B-GAL4* or *nub-GAL4* drivers does not result in wing overgrowth when compared to controls. n≥36, error bars show 95% confidence limits. (D) Knockdown of *dis3L2* throughout the wing imaginal disc results in significant overgrowth of the wing (Diii) which is not due to an increase in fly mass (Div). Representative image of a knockdown male wing is shown, scale bar = 200µm (Dii). Control genotypes (;*UAS-dis3L2<sup>RNAi</sup>*; *UAS-EGFP<sup>RNAi</sup>/+*; *69B-GAL4/+* and ;;*69B-GAL4*), *Dis3L2<sup>KD</sup>* genotype = ;*UAS-dis3L2<sup>RNAi</sup>/+*; *69B-GAL4/+*). n≥18, error bars represent 95% confidence limits, \*\*\*\*=p<0.0001. (E) Knockdown of *dis3L2* within the wing pouch of the wing imaginal disc results in significant overgrowth of the wing (Eiii) which is not due to an increase in fly mass (Eiv). A Representative image of a knockdown male wing is shown, scale bar = 200µm (Eii). Control genotypes (;*UAS-dis3L2<sup>RNAi</sup>*; ;*UAS-EGFP<sup>RNAi</sup>/nub-GAL4*; and ;*nub-GAL4*); *Dis3L2<sup>KD</sup>* genotype = ;*UAS-dis3L2<sup>RNAi</sup>/nub-GAL4*);. n≥22, error bars represent 95% confidence limits, \*\*\*\*=p<0.0001.



**Figure 4.8: Ubiquitous knockdown also results in wing overgrowth but does not cause total overgrowth of the fly.** Knockdown of *dis3L2* driven by *act-GAL4* (A), *tub-GAL4* (B) or *da-GAL4* (C) results in significant overgrowth of the wing when normalised to fly mass. (A) Control = ; *UAS-dis3L2<sup>RNAi</sup>* ; , *Dis3L2<sup>KD</sup>* = ; *act-GAL4/UAS-dis3L2<sup>RNAi</sup>* ; , n≥12, error bars represent 95% confidence intervals, p<0.0001. (B) Control = ; *UAS-dis3L2<sup>RNAi</sup>* ; , *Dis3L2<sup>KD</sup>* = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *tub-GAL4/+*, n≥20, error bars represent 95% confidence intervals, p<0.0001. (C) Control = ; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *da-GAL4* , *Dis3L2<sup>KD</sup>* = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *da-GAL4/+*, n≥18, error bars represent 95% confidence intervals, p<0.0001. (D) *act-GAL4* driven ubiquitous knockdown of *dis3L2* using either *UAS-dis3L2<sup>RNAi</sup>* line (Di and Dii) does not result in an increase in fly mass. n≥16, error bars represent 95% confidence intervals, p>0.05 for both.



**Figure 4.9: Wing overgrowth is due to an increase in cell number rather than an increased cell size.** (A) Representation of the 0.1mm<sup>2</sup> regions of the wing used to calculate the number of cells within the wing. Enlarged image shows the single hairs protruding from each cell which were counted to give a representation of cell number. (B-C) Knockdown of *dis3L2* throughout the wing imaginal disc results in an increase in cell number (C) rather than an increase in cell size (B). Control genotypes = ;*UAS-dis3L2<sup>RNAi</sup>*; ; *UAS-EGFP<sup>RNAi</sup>/+* ; *69B-GAL4/+* and ;;*69B-GAL4*. *Dis3L2<sup>KD</sup>* genotype = ;*UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+*). n≥10, error bars represent SEM, \*\*\*\*=p<0.0001. (D-E) Consistent with B/C, knockdown of *dis3L2* within the wing pouch results in an increase in cell number (E) rather than an increase in cell size (D). Control genotypes = ;*UAS-dis3L2<sup>RNAi</sup>*; ; *UAS-EGFP<sup>RNAi</sup>/nub-GAL4*; and ;*nub-GAL4*; *Dis3L2<sup>KD</sup>* genotype = ;*UAS-dis3L2<sup>RNAi</sup>/nub-GAL4*; . n≥17, error bars represent 95% confidence limits, \*\*\*\*=p<0.0001.

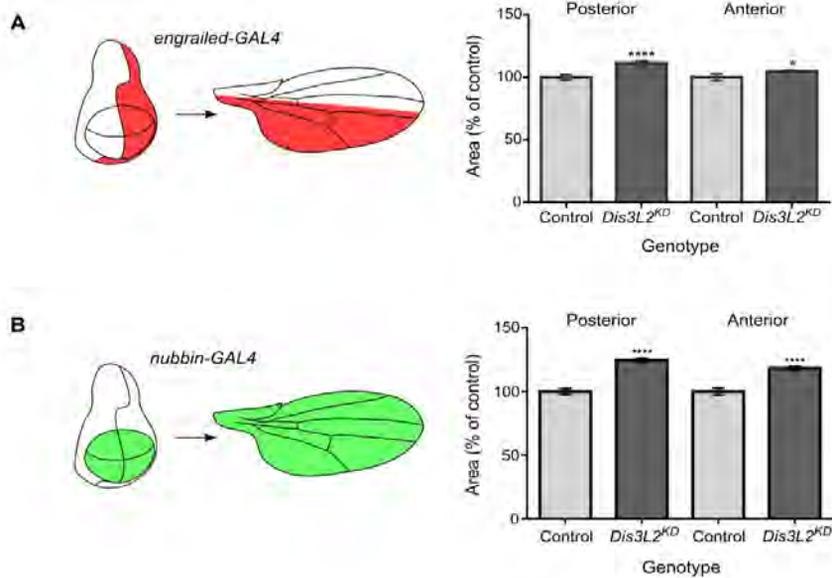
#### 4.5.4 Wing overgrowth is specific to the loss of *dis3L2*.

The above data give strong evidence that Dis3L2 is required to control the proliferation of the cells fated to form the adult wing. However, the wing area also depends on the size of the fly which can be variable. Therefore an ideal experiment would contain an internal control within the fly so that the potential variation in mass can be easily accounted for. To achieve this an additional *GAL4* driver *engrailed(en)-GAL4* was used. The expression of *en-GAL4* within the wing imaginal disc is restricted to the posterior compartment as the *engrailed* gene itself is critical for the anterior-posterior patterning of the disc (Kornberg *et al.* 1985). Using *en-GAL4* ensured that only the cells in the posterior region would have *dis3L2* knocked down whilst levels of *dis3L2* remained unchanged in the anterior cells. The restriction to the posterior compartment allowed the anterior area to serve as an internal control as if the overgrowth was specific to *dis3L2* knockdown its area should be unaffected. Like the previous experiments flies were aged to between 1 and 2 days of age after culturing at 25°C in uncrowded conditions prior to wing cutting.

The anterior/posterior boundary of the wing was determined by drawing a straight line in the imaging software from the L1-L2 meeting point to the distal tip of L4 (Methods Figure 2.1). Knockdown of *dis3L2* specifically in the posterior compartment of the wing disc resulted in the overgrowth of the posterior region of the wing. The posterior area of *en-GAL4/UAS-dis3L2* wings was 111.3% the size of control wings (Figure 4.10A). Similar to the *act-GAL4* crosses only the *UAS-dis3L2<sup>RNAi</sup>* parent was suitable for use as a control as the *en-GAL4* stock was also balanced over CyO. Unexpectedly the anterior area of knockdown wings also showed a very slight, although significant, overgrowth. The anterior area was 102.7% the size of controls (Figure 4.10A); therefore the overgrowth was much less than that observed for the posterior area. The anterior overgrowth was most likely caused by a knock-on effect of the overgrowth in the posterior area following cross axis signalling. In comparison the *nub-GAL4* driven knockdown wings show comparable overgrowth of the posterior (125%) and anterior (119%)

respectively (Figure 4.10B) showing it is not only the posterior region that is susceptible to overgrowth. Together these data show the specificity of the overgrowth induced following the loss of *dis3L2* in the wing imaginal disc.

In addition to the compartment specific overgrowth 46.6% of the *en-GAL4* driven *dis3L2* knockdown wings showed blisters specifically localised to the posterior compartment of the wing. Wing venation phenotypes were also observed, specifically abnormalities in the posterior cross vein and L5. Their restriction to the posterior region indicated an effect induced by the loss of *dis3L2*; however, it may also be caused by the expression of GAL4 in the *engrailed* domain. Due to the presence of the CyO marker in the *en-GAL4* stock it was difficult to see blistering and wing vein phenotypes therefore the *en-GAL4* stock was crossed over a wild type chromosome (*en-GAL4/+*). These wings showed a similar proportion of blistering and wing vein phenotypes indicating they are caused by the expression of *GAL4* in the *engrailed* domain. These phenotypes were therefore not examined any further.



**Figure 4.10: *dis3L2* knockdown induced overgrowth is specific. (A)** Restricting knockdown of *dis3L2* to the posterior region of the wing results in specific posterior overgrowth (111% of control). Minor, but significant overgrowth (102% of control) is also observed, however this is likely to be in response to the overgrowth in the posterior compartment. Control genotype = ;*UAS-dis3L2<sup>RNAi</sup>*;, *Dis3L2<sup>KD</sup>* genotype = ;*UAS-dis3L2<sup>RNAi</sup>/en-GAL4*;.  $n \geq 20$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ , \*= $p < 0.05$ . **(B)** In contrast the overgrowth observed when *dis3L2* was knocked down using *nub-GAL4* resulted in comparable overgrowth of the posterior (124% of control) and anterior (118% of control) compartments. Control genotype = ;*UAS-dis3L2<sup>RNAi</sup>*;, *Dis3L2<sup>KD</sup>* genotype = ;*UAS-dis3L2<sup>RNAi</sup>/nub-GAL4*;).  $n \geq 18$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ .

## 4.6 Overgrowth occurs early in development.

The loss of *dis3L2* within the wing imaginal disc results in the overgrowth of the *Drosophila* wing by causing an increase in cell number. However, to help identify the pathways responsible for the overgrowth it was important to know the developmental time during which the effects take place. The temperature sensitive *GAL80* (*GAL80<sup>ts</sup>*) provides a temporal control in addition to the spatial control given by the *GAL4-UAS* system. At the permissive temperature of 19°C the *GAL80<sup>ts</sup>* inhibits *GAL4*, blocking its binding to the *UAS* and subsequently preventing the knockdown of *dis3L2*. However, at 29°C *GAL80<sup>ts</sup>* is non-functional, relieving its inhibition of *GAL4* thus allowing the knockdown of *dis3L2* (Methods Figure 2.2B). Cross vials were moved from the permissive temperature (19°C) to the active temperature (29°C) at different stages of development allowing the identification of the developmental stages during which *Dis3L2* activity is critical to prevent wing overgrowth (Figure 4.11).

For this type of experiment it is important to take the perdurance into account; that is the time it takes after moving to the active temperature for the current protein to be degraded and the knockdown to take effect. In the absence of a *Dis3L2* antibody it was not possible to calculate this therefore previous data in the lab for the 5'-3' exoribonuclease Pacman was used as it provided the most suitable example. A previous PhD student, Joe Waldron, calculated the perdurance for this system using RNAi to knockdown Pacman to be between 24 and 32 hours (Waldron 2014). This time was taken into account when calculating the times referring to each development stage.

Figure 4.11 shows the times at which cross and control vials were moved from 19°C to 29°C and the developmental stages they represent. To ensure that the progeny were all of the same age a short egg lay period of 6 hours was used. This involved mating the flies 2 days prior to the start of the experiment, transfer into new vials and leaving for 6 hours to lay sufficient eggs before the adults were removed. Due to the effect temperature has on the size of the adult structures of the fly (Azevedo *et al.* 2002) control crosses of both parental stocks were set up

in parallel alongside the knockdown crosses; these were also moved from the permissive to active temperature resulting in specific controls for each developmental stage. The developmental stages selected were the start of L2, the start and mid stage of L3 and the start of the pupal stage. Due to the perdurance L1 was not able to be included as a developmental stage.

Together with the vials that were moved across for each developmental stage, cross and control vials were kept at 19°C and 29°C to serve as negative and positive controls respectively. The level of *dis3L2* RNA was examined within these samples to assess if the GAL80<sup>ts</sup> system was functioning as expected. *69B-GAL4* in conjunction with GAL80<sup>ts</sup> was used to drive *dis3L2* knockdown throughout the wing imaginal disc (Figure 4.11). At 19°C wild-type levels of *dis3L2* was observed in the ; *tubGAL80<sup>ts</sup>/UAS-dis3L2<sup>RNAi</sup> ; 69B-GAL4/+* larvae indicating the GAL80<sup>ts</sup> was functioning as expected and the knockdown of *dis3L2* was inhibited (Figure 4.12A (*Dis3L2<sup>KD</sup>* 19°C)). However, at 29°C a significant knockdown of *dis3L2* was observed in the ; *tubGAL80<sup>ts</sup>/UAS-dis3L2<sup>RNAi</sup> ; 69B-GAL4/+* larvae showing the GAL80<sup>ts</sup> was inhibited allowing the normal activity of the *GAL4-UAS* system (Figure 4.12A *Dis3L2<sup>KD</sup>* 29°C). This showed that the system was working as expected technically and was therefore suitable to carry out the desired investigation. In addition to the expected pattern of *dis3L2* expression we measured the wings of 1-2 day old males cultured at either 19°C or 29°C. As expected there was no significant difference between control and knockdown wings of flies cultured at 19°C (-ve) whilst an overgrowth of (17.4%) was observed in the positive control (+ve) samples (Figure 4.12B) which is consistent with the 20% overgrowth previously observed.

Having identified the system was working as expected both technically and phenotypically the effect of knocking down *dis3L2* during different developmental stages was examined. The progeny having *dis3L2* knockdown switched on at the start of L2 and start of L3 showed significant overgrowth of the wing. Knockdown wings were 111.1% and 110.6% the area of their respective parental controls (Figure 4.12B). However, when *dis3L2* knockdown was

inhibited until the mid L3 stage (101.8% of control) or the start of the pupal stage (101.7% of control) no overgrowth was observed (Figure 4.12B). This therefore indicates that *Dis3L2* is required during early larval development to prevent the overgrowth of the wing and is thus involved in regulating pathways active during these stages. This can be seen as consistent with the Perlman syndrome condition which is characterised by overgrowth during foetal development.

## **4.7 Loss of *dis3L2* results in increased proliferation in the wing imaginal discs.**

The GAL80<sup>ts</sup> experiments identified a role for *Dis3L2* in the regulation of tissue growth during early development which may be consistent with the Perlman syndrome disorder. Therefore the natural progression was to look at the wing precursor tissues, the wing imaginal discs themselves, which are developing in the larva during these stages. If the cause of the overgrowth was occurring during the early developmental stages the working hypothesis was that the late L3 larvae would also show an overgrowth phenotype. To address this the area of late L3 wing imaginal discs were measured. Larvae were staged to late L3 using bromophenol blue food which can be observed in the larval gut. Just prior to pupariation the larva empties its gut and therefore only those with an absence of blue food in the gut were selected for disc dissection. This ensures that the larvae are all of the same age and therefore any difference in disc area would be due to the knockdown of *dis3L2* rather than them being at different stages of development.

### **4.7.1 Loss of *dis3L2* results in significantly larger wing imaginal discs.**

Similar to the wing experiments *dis3L2* knockdown was achieved using the same two wing disc *GAL4* drivers; *69B-GAL4* and *nub-GAL4*. As mentioned previously crosses were cultured at 25°C and late L3 larvae were selected for dissection on the absence of blue food in their gut. There was no significant difference between the parental controls for either cross. In contrast,

knockdown of *dis3L2* using either *69B-GAL4* or *nub-GAL4* results in a significant overgrowth of the late L3 wing imaginal discs (Figure 4.13A-B). *69B-GAL4* driven knockdown discs were 132% the size of their respective controls whilst *nub-GAL4* driven knockdown discs were 121.6% the size of controls. The difference in the level of overgrowth between the two drivers is most likely due to their expression domains. For example *69B-GAL4* drives throughout the wing disc and therefore the whole disc would be susceptible to *dis3L2*-depletion induced overgrowth. On the other hand *nub-GAL4* expresses specifically within the pouch region which makes up around half of the disc and therefore only this region would be expected to show the overgrowth phenotype. This is in comparison to the earlier wing data which the cells forming the wing would have been equally affected by the two drivers, thus very similar levels of overgrowth were observed.

The previously characterised wing overgrowth was due to an increase in cell number rather than an increase in cell size which suggested an increased level of proliferation within these specific cells. However, it is also possible that a decreased level of apoptosis within the disc could also result in an increased number of cells. Therefore to test this actively proliferating cells were stained with an antibody to Phosphohistone H3 (PH3). The histone H3 is phosphorylated specifically during mitosis and therefore using a phospho-specific antibody allowed the identification of cells actively going through mitosis. The *69B-GAL4* driver was selected for use as it drives throughout the wing imaginal disc and therefore the effect of *dis3L2* on proliferation throughout the disc could be assessed. Late L3 discs were dissected from control and knockdown larvae, fixed and stained with anti-PH3.

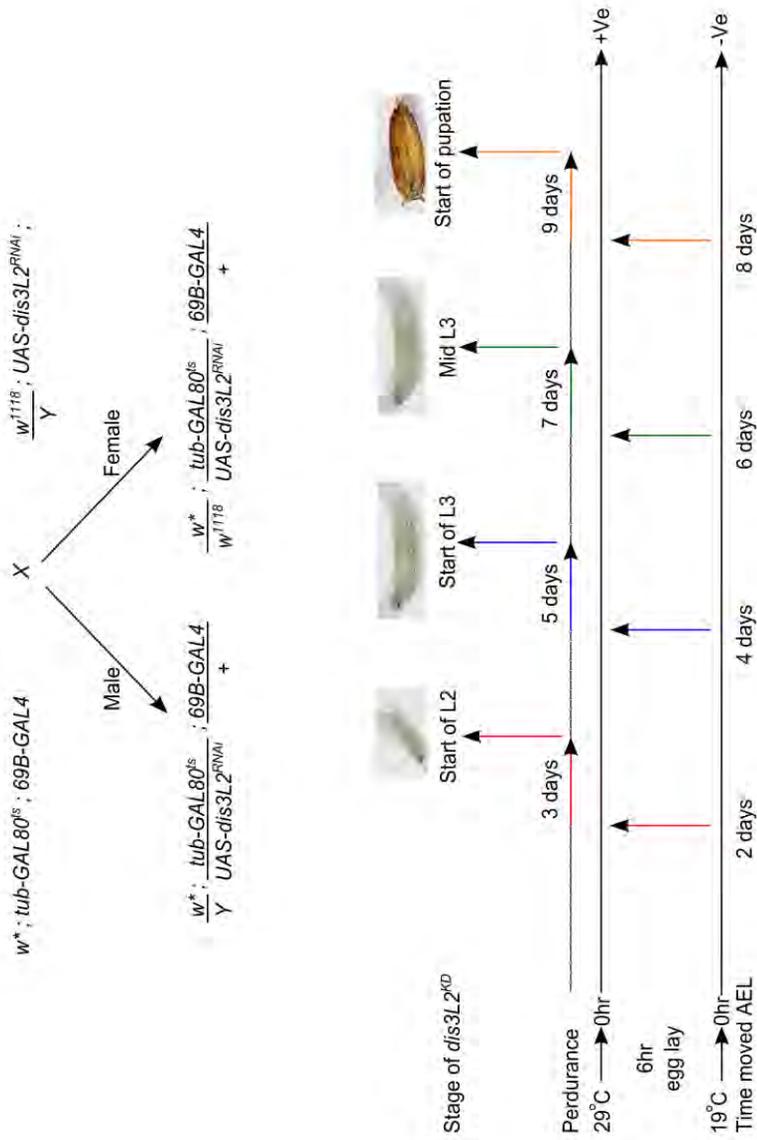
Z-stacks of the stained discs were taken using a Leica SP8 confocal microscope using consistent settings across all slides. To calculate the number of proliferating cells in each disc the 'DeadEasy MitoGlia' imageJ plugin was used (Forero *et al.* 2010). This plugin, designed to count cells labelled with PH3 in *Drosophila* embryos, uses user defined specific parameters such as thresholding and minimum volume. For each disc different thresholds were trialled and

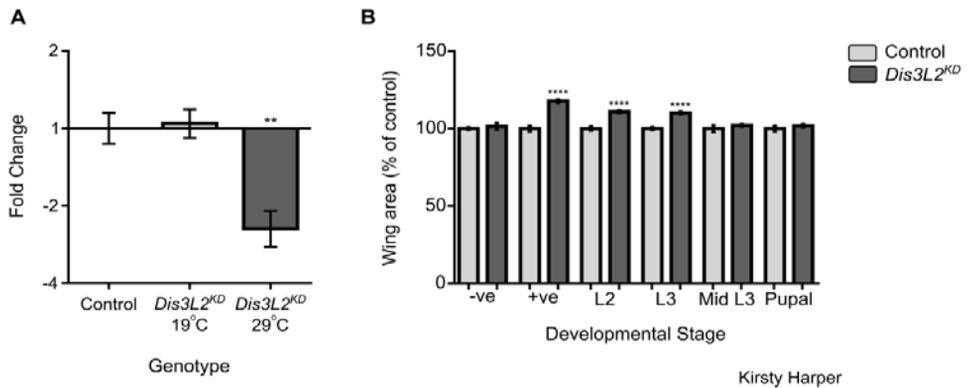
compared to the original image to work out the most suitable threshold that best represents the stained image. A threshold of 20 was selected for the analysis. Each disc was put through the DeadEasy MitoGlia plugin resulting in an output of the number of stained cells.

The number of proliferating cells was normalised to the total area of the disc to account for the difference in size between the control and knockdown discs resulting in a mitotic index. There was no significant difference between the parental controls and so they were again grouped. The mitotic index, representing the number of proliferating cells, in the *dis3L2<sup>KD</sup>* wing discs was 20.5% higher than control discs (Figure 4.13C). These data therefore show that the loss of *dis3L2* results in an increased rate of cell division within the wing imaginal disc which in turn leads in an increase in the total area of the tissue and larger adults wings as observed in 4.5.2.

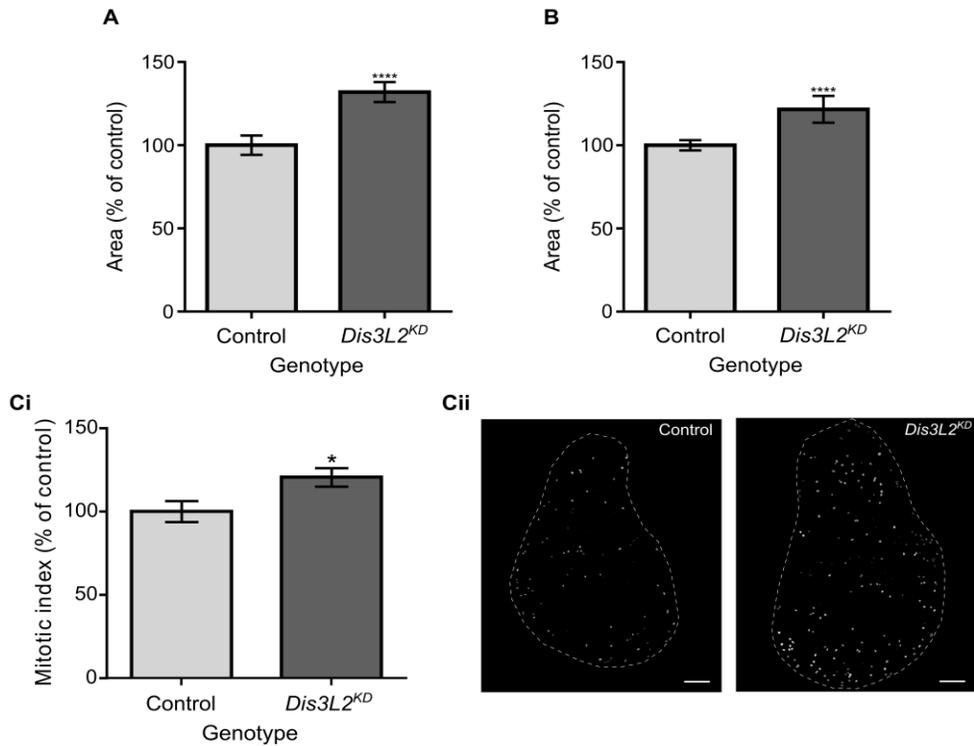
**Figure 4.11: Experimental plan to knockdown *dis3L2* during different developmental stages.** A simple one generation cross was performed to allow the knockdown of *dis3L2* to be controlled by temperature using the temperature sensitive *GAL4* inhibitor *GAL80<sup>ts</sup>*. A 6 hour egg lay was

performed at the required temperatures. Vials were then moved from the permissive temperature (19°C) to the active temperature (29°C) at various stages of development. The perdurance (time delay before the knockdown takes effect) of 24–32 hours is shown along the dotted line and was taken into account for the developmental timings. A positive control (+ve) and negative control (-ve) was kept at the active and permissive temperatures respectively throughout development.





**Figure 4.12: Wing overgrowth is only observed when *dis3L2* is knocked down during early larval stages.** RNA used for qRT-PCR was extracted from a single L3 larva so the knockdown efficiency would be diluted by wild type tissues. **(A)** The GAL80<sup>ts</sup> addition was working as expected as no *dis3L2* knockdown was observed in larvae retained at the permissive temperature (19°C) whilst a significant knockdown of 60% was observed at the active temperature (29°C). Control = ; *tub-GAL80<sup>ts</sup>* ; *69B-GAL4* and ; *UAS-dis3L2<sup>RNAi</sup>* ; Both *Dis3L2<sup>KD</sup>* = ; *tub-GAL80<sup>ts</sup>/UAS-dis3L2<sup>RNAi</sup>* ; *69B-GAL4/+*. n≥3, error bars represent standard error, \*\*p=0.0064. **(B)** Wing overgrowth only occurs when *dis3L2* knockdown is initiated at the start of L2 (L2) or the start of L3 (L3). The positive control (+ve), start of L2 and start of L3 timings show a significant increase in adult wing area (\*\*\*\*=p<0.0001) whilst the negative control (-ve), mid L3 and start or pupal timings so no significant increase in wing area of knockdown flies (*Dis3L2<sup>KD</sup>* genotype = ; *UAS-dis3L2<sup>RNAi</sup>/GAL80<sup>ts</sup>* ; *69B-GAL4/+*) compared to both parental controls (; *UAS-dis3L2<sup>RNAi</sup>*; and ; *GAL80<sup>ts</sup>* ; *69B-GAL4*). n≥37 for all developmental timings, error bars represent 95% confidence limits.



**Figure 4.13: Knockdown of *dis3L2* results in increased proliferation within the wing imaginal disc leading to overgrowth of the disc.** Knockdown of *dis3L2* in the wing imaginal disc results in significant overgrowth of the disc when driven by either *69B-GAL4* (A) or *nub-GAL4* (B). (A) *Dis3L2<sup>KD</sup>* (genotype = ;*UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+*) late L3 wing discs are 131.9% the size of control discs (; *UAS-dis3L2<sup>RNAi</sup>* ; and ; ;*69B-GAL4*).  $n \geq 25$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ . (B) *Dis3L2<sup>KD</sup>* (genotype = ; *UAS-dis3L2<sup>RNAi</sup>/nub-GAL4* ; ) late L3 wing discs are 121.6% the size of control discs (; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *nub-GAL4* ;).  $n \geq 21$ , error bars represent 95% confidence limits, \*\*\*\* =  $p < 0.0001$ . (Ci) *Dis3L2<sup>KD</sup>* wing discs have a significantly higher mitotic index (120.5% the number of cells undergoing mitosis) than control discs.  $n \geq 13$ , \* $p = 0.0241$ , error bars represent 95% confidence limits. (Cii) Representative images of the DeadEasy MitoGlia output following staining of late L3 wing imaginal discs with anti-Phosphohistone H3. Scale bar = 50 $\mu$ m.

## 4.8 Loss of *dis3L2* results in enhanced starvation survival

The previous work in this chapter has focused on the characterisation of the overgrowth phenotype induced by *dis3L2* knockdown. However, as an aside the effect of the loss of *dis3L2* on stress response was investigated. Dis3L2 has previously been identified to play a role in mRNA clearance following stress-induced apoptosis (Thomas *et al.* 2015). Therefore it was of interest to ask if the loss of *dis3L2* affects the *Drosophila's* response or resistance to stress.

An easy method of inducing stress is to put the flies under starvation conditions which simply involved their transfer from a normal vial to a vial containing no food (Figure 4.14A). For the experiment 3 control groups were used; (I) the *UAS-dis3L2<sup>RNAi</sup>* and (II) *tub-GAL4/TM6* parents together with the ; *UAS-dis3L2<sup>RNAi</sup>/+ ; TM6/+* sibling controls (III). The siblings controlled for the conditions under which the cross flies were cultured whilst the parentals provided additional controls for the genetic background. Additionally, the siblings and the *tub-GAL4* parent carried the TM6 balancer which may affect the starvation survival itself, therefore in using the *UAS-dis3L2<sup>RNAi</sup>* parent it was possible to assess the effect the TM6 balancer had. Males and females were scored separately.

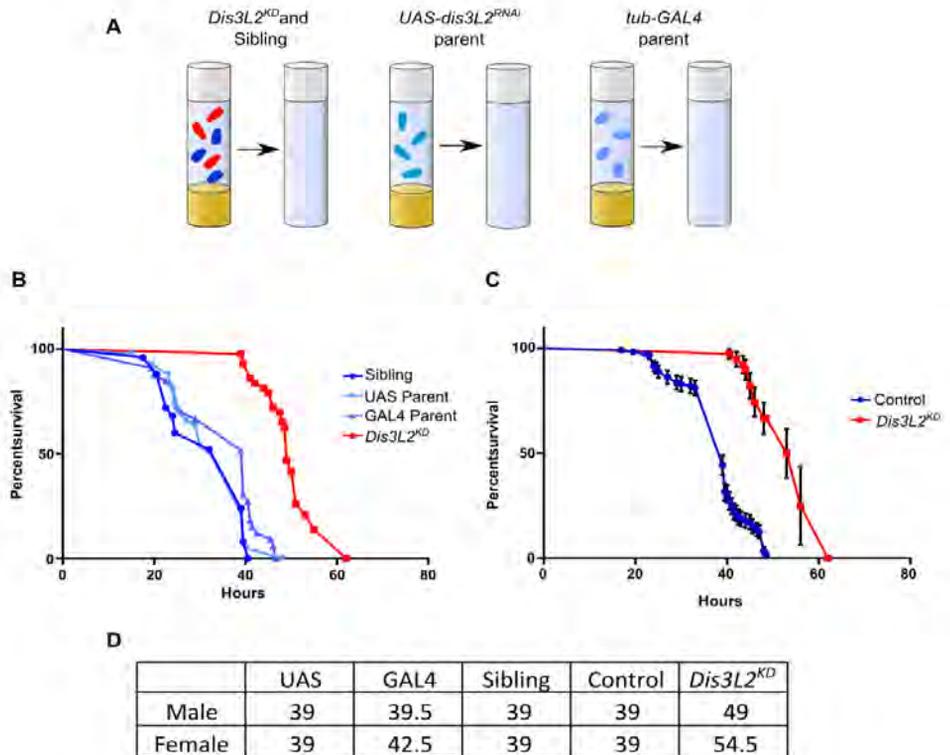
The first observation was that the three control groups showed similar survival times under starvation conditions (Figure 4.14B/D). These were not significantly different from each other for both the males and females suggesting the survival times are generally consistent and the TM6 balancer has no effect on starvation survival. However, the *dis3L2* knockdown flies were shown to survive significantly longer (Figure 4.14). Knockdown flies survived on average 49 and 54.5 hours for males and females respectively. This represents a 26% and 36% increase in survival time for males and females over the controls which both survive for an average of 39 hours following transfer to an empty vial.

This experiment indicates that the loss of *dis3L2* enhances the adult *Drosophila's* ability to resist starvation induced stress for a significant period of time. This indicates that Dis3L2 may function in pathways involved in the stress response. Could this be linked to the overgrowth

phenotype? For example are *dis3L2*-depleted flies/tissues able to source energy at a higher efficiency and therefore are able to grow at a greater rate and/or survive for longer under starvation conditions?

## 4.9 Chapter Summary

The work in this chapter has identified a role for Dis3L2 in regulating the rate of proliferation within the developing wing imaginal disc. Loss of *dis3L2* has been shown to result in increased proliferation, leading to an overgrowth of the wing disc and subsequent overgrowth of the wing resulting from an increase in the total number of cells. These findings mimic the human condition of Perlman syndrome, a rare condition characterised by foetal tissue overgrowth thus potentially identifying a conserved function of Dis3L2. In addition to the overgrowth phenotype an increase in starvation survival was also observed following the loss of *dis3L2* indicating a potential role for Dis3L2 within stress response pathways.



**Figure 4.14: Ubiquitous knockdown of *dis3L2* leads to an improved starvation response.** **(A)** Experimental set up for the starvation experiments. *Dis3L2<sup>KD</sup>* (red, genotype = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *tub-GAL4/+*) and sibling (dark blue, genotype = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *TM6/+*) cross progeny, cultured in the same vial, were tipped into an empty vial containing no food or water. The same was performed for both the parental control lines. The number of surviving progeny was scored over a period of 60 hours. **(B)** The survival of male flies of all three control genotypes is consistent at ~39 hours after transfer to an empty vial. Siblings ( ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *TM6/+*) show no difference in survival to the parentals (UAS parent = ; *UAS-dis3L2<sup>RNAi</sup>* ; GAL4 parent = ; *tub-GAL4/TM6*) indicating changes in survival are not due to the different culturing conditions. *Dis3L2<sup>KD</sup>* males ( ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *tub-GAL4/+*) survive on average 10 hours longer than controls.  $n \geq 25$ , Mantel-Cox test  $p < 0.0001$  for *Dis3L2<sup>KD</sup>* vs each control genotype. **(C)** Due to showing no significant difference the controls were grouped. Control includes ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *TM6/+* ; ; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *tub-GAL4*.  $n \geq 28$ , error bars represent standard error, Mantel-Cox test  $p < 0.0001$ . **(D)** Summary of the survival times (in hours) for controls and *Dis3L2<sup>KD</sup>* flies for both males (shown in B/C) and females. Genotypes are the same as those outlined in A and B.

## 4.10 Chapter Discussion

### 4.10.1 Why does the loss of *dis3L2* result in proliferation?

The work presented in this chapter has identified a role for Dis3L2 in controlling developmental proliferation and ultimately tissue size. The loss of *dis3L2* specifically in the wing imaginal disc results in the overgrowth of the disc; an overgrowth phenotype consistent with that observed in the human disorder Perlman syndrome. This finding builds on the current knowledge that *dis3L2* knockdown in human HEK293 cells results in increased cell growth (Astuti *et al.* 2012). However, this work has, for the first time, identified an increase in proliferation throughout a developmental tissue resulting in the overgrowth of the specific tissue mimicking the Perlman syndrome phenotype of internal organ overgrowth (Baehrecke 2006; Astuti *et al.* 2012).

The wing overgrowth outlined in section 4.5 has been shown to result from an increase in the number of cells in the adult wing (hyperplasia) rather than an increase in the size of the cells (hypertrophy). Further exploration in the wing precursor tissues revealed the hyperplasia resulted from an increase in proliferation. An effect on apoptosis was not directly assessed using Caspase staining (as was used in Chapter 3); however, the inhibition of apoptosis alongside *dis3L2* knockdown results in an additive effect suggesting an alternative pathway of hyperplasia (Chapter 6 Figure 6.6). The levels of apoptosis within the wing imaginal disc during normal development is relatively low for a developmental tissue (Waldron *et al.* 2015) and therefore it would have been difficult to assess a change in the level of apoptosis. A more suitable tissue to study the role of Dis3L2 in apoptosis would have been the eye/eye imaginal disc as developmentally regulated apoptosis is essential for correct eye development (Wolff and Ready 1991). Previous work has shown a role for Dis3L2 in mRNA clearance during apoptosis (Thomas *et al.* 2015) with knockdown of *dis3L2* enhancing cell survival. However, with the additive effects of apoptotic inhibition and *dis3L2* knockdown on wing area together

with the PH3 staining revealing a role in proliferation any further investigation into an apoptotic function was not performed.

Interestingly, the overgrowth observed following the knockdown of *dis3L2* appears to show at least an element of tissue specificity. Knockdown in the wing results in overgrowth of the wing; whilst although ubiquitous knockdown of *dis3L2* also causes wing overgrowth the total size of the fly is unaffected. The wing imaginal discs were the only specific tissue to be examined in this way and therefore it would be of interest to use specific GAL4 drivers to identify other tissues that are susceptible to *dis3L2* knockdown induced overgrowth. This element of tissue specificity would also be consistent with the symptoms of Perlman syndrome, as although the liver and kidney are overgrown other organs such as the lungs do not show such an extreme increase in size. This raises the question, why are specific tissues susceptible to overgrowth? It is noteworthy that the liver and the kidney, like the wing imaginal discs, are regenerative tissues whilst the lungs for example, are not. It is therefore possible that Dis3L2 functions to control proliferation when regeneration is required, such as following tissue damage. However, while intriguing this would require further study.

Why does the loss of *dis3L2* result in increased proliferation in the wing imaginal disc? The mechanism behind the overgrowth phenotype could be a direct or indirect effect of loss of Dis3L2 function. The fact that a proliferative phenotype was also observed in human tissue culture cells (Astuti *et al.* 2012) supports these findings and suggests a conserved function in controlling cell proliferation. A likely cause of proliferation could be that Dis3L2 is normally responsible for the degradation of specific RNA targets which become misregulated in its absence. A selection of these RNA targets may be mRNAs encoding oncogenic proteins; an increase in their RNA stability may in turn increase the level of these proteins which drive the cell through proliferation.

Together with mRNAs, Dis3L2 has been shown to function in the regulation of miRNA biogenesis, specifically *let-7a* in mouse embryonic fibroblasts (Chang *et al.* 2013; Ustianenko *et*

*al.* 2013). Dis3L2 functions to degrade *pre-let-7a* following its oligouridylation. However, the knockdown of *dis3L2* in these cells resulted in an increase in the levels of *pre-let-7a* whilst the levels of mature *let-7a* remained unchanged. Despite this, *let-7a* is known to be a tumour suppressor and therefore places Dis3L2 in a pathway known to affect cell growth and proliferation. It is entirely possible that Dis3L2 functions to regulate other miRNAs, either at the level of biogenesis or their stability to elicit its control over cell growth. If Dis3L2 is involved in the degradation of specific miRNAs they may become overexpressed in *dis3L2* knockdown discs. Overexpression of the miRNA *bantam* in *Drosophila* results in wing overgrowth (Brennecke *et al.* 2003) through stimulating proliferation; whilst overexpression of *miR-278* results in overgrowth of the eye (Nairz *et al.* 2006). Alternatively, a function for Dis3L2 in the biogenesis of specific miRNAs, similar to that of *let-7a* may also be possible. In this case the loss of *dis3L2* may result in a decrease in levels of the mature miRNAs. If these specific miRNAs such as *miR-15a* and *miR-16-1* (Calin *et al.* 2002) or *miR-21* and *miR-24* (Cheng *et al.* 2005) have anti-proliferative functions they may contribute to the observed increase in proliferation. Therefore the regulation of the cell cycle by miRNAs is clear and a role for Dis3L2 within this pathway is a strong possibility which would require further investigation.

In the following chapter RNA sequencing has been carried out to explore how of the loss of *dis3L2* in the wing imaginal disc affects the levels of mRNAs in an attempt to start to answer this question.

#### **4.10.2 Why is the overgrowth only observed when *dis3L2* is lost during early developmental stages?**

The mechanistic cause of wing overgrowth has been shown to occur during mid larval development through the use of the *GAL80<sup>ts</sup>* addition to the *GAL4-UAS* system. This allowed successful control over the developmental periods during which *dis3L2* was knocked down. Simulating *dis3L2* knockdown at the start of the 2<sup>nd</sup> instar larval and the start of the 3<sup>rd</sup> instar larval stages resulted in wing overgrowth consistent with the findings outlined previously. However, when *dis3L2* knockdown was driven from the mid 3<sup>rd</sup> instar larval and pupal stages the wings were wild-type in area. This showed that the cellular misregulation, induced by the loss of *dis3L2*, which leads to the overgrowth occurs during the mid larval stages. The data also demonstrate that the wing overgrowth was not caused by an issue during the metamorphosis occurring during the pupal stages where the wing blade develops from the wing imaginal disc. It has been previously demonstrated that the inhibition of cell cycle during the pupal stage does not affect the size of the pupal wing, but instead results in an increase in cell size (Weigmann *et al.* 1997) which has been shown to be unaffected by *dis3L2* knockdown. This does not mean, however, that the function of Dis3L2 is redundant during the latter stages as modENCODE data shows the expression of *dis3L2* mRNA at a consistent level throughout *Drosophila* development (Graveley *et al.* 2011).

During the larval stages of development the cells of the wing imaginal disc proliferate at a great rate leading to a tissue originally of 50 cells increasing to one of 50,000 cells in size in only a few days with the proliferation occurring ubiquitously throughout the disc (Garcia-Bellido and Merriam 1971; Gonzalez-Gaitan *et al.* 1994). During L2 and early L3 the cells of the wing disc proliferate exponentially (Nienhaus *et al.* 2012) leading to an increased rate of division than during the final stages of larval development where cells in late L3 wing discs accumulate in G2 prior to pupariation (Neufeld *et al.* 1998). This therefore may mean that the wing imaginal disc cells are more sensitive to the over-proliferative stimuli during L2/early L3

following the loss of *dis3L2*. Additionally, knocking down *dis3L2* earlier in development may give the cells more time to over-proliferate, whereas if *dis3L2* is lost during the final larval stages the cells may have the time to over-proliferate sufficiently to result in wing overgrowth.

Interestingly, despite an increase in the total area of the wing as a result of an increase in cell division, the patterning and shape of the wing remains normal with no evidence of wing venation phenotypes. This is consistent with previous findings showing that cell cycle alteration does during the mid larval stages does not interrupt disc patterning (Weigmann *et al.* 1997).

#### **4.10.3 Do Dis3L2 and Pacman function in distinct pathways?**

The data presented here indicate a novel role for Dis3L2 in the control of developmental proliferation. The phenotypes observed show a specific requirement for Dis3L2 to maintain homeostasis and show there is a lack of redundancy between the pathways of RNA decay. Moreover the phenotypes previously described for Dis3 (Chapter 3) plus the previous work in the Newbury lab showing a role for Pacman in the regulation of apoptosis (Waldron *et al.* 2015) show distinct functions in controlling tissue development and cell growth. Although it is evident that Pacman and Dis3L2 do share some targets as they have been shown to co-immunoprecipitate in an RNA dependent manner (Lubas *et al.* 2013), their contrasting roles in apoptosis and proliferation suggest they are able to regulate specific targets.

Knockdown of DIS3L2 and XRN1 in human cells has also been shown to affect the cell in different manners. For example, knockdown of Dis3L2 results in an increase in the total number of P-bodies within the cell whilst XRN1 depletion results in an increase in P-body size (Lubas *et al.* 2013) consistent with previous findings in *Drosophila* (Zabolotskaya *et al.* 2008). This suggests that the cell responds differently to the knockdown of the two proteins, again suggesting differential overall functions within the cell.

How Pacman and Dis3L2 are targeted to their specific substrates is an interesting question and one which largely remains unanswered. It is likely that the specific RNAs are either tagged by additional nucleotides, bound by RNA binding proteins or interaction with miRNAs which results in recruitment of either Pacman or Dis3L2 leading to RNA degradation. For example, as mentioned previously, Dis3L2 has been shown to target uridylated *pre-let-7a*; taken together with in-vitro data showing Dis3L2 has a preference for uridylated substrates (Malecki *et al.* 2013) this suggests that terminal nucleotide tagging is a Dis3L2 recruitment mechanism. Similarly, cis-acting factors in the UTRs of the target RNAs may also play a role in the specific targeting. For instance, AU rich elements (ARE's) have been shown to target RNAs for exosome-mediated decay (Shaw and Kamen 1986; Chen and Shyu 1995; Chen *et al.* 2001) together with further work showing that Dis3L2 (Lubas *et al.* 2013) and XRN1 (Stoecklin *et al.* 2006) are also able to degrade ARE containing RNAs. However, the crossover here is unlikely to provide the substrate specificity therefore further work would be required to identify specific sequence or structural motifs which direct specific degradation.

The interaction between Pacman and Dis3L2 and their substrate specificities is further explored in subsequent chapters.

## 4.11 Acknowledgements

The measurements of the ; *nub-GAL4/+* ; *UAS-dis3L2<sup>RNAi</sup>/+* wings and respective controls were carried out by Yasmeen Mashhadi Ahvazi. This was using the alternative RNAi which showed the same phenotypes as the work carried out by myself. The GAL80<sup>ts</sup> experiments were carried out by Kirsty Harper. Both these undergraduate students carried out the experiments under my supervision.

# Chapter 5: Using RNA-sequencing to identify potential targets of Dis3L2

## 5.1 Introduction

The previous chapter has outlined the exciting finding of a conserved function for Dis3L2 in controlling the level of proliferation within a developmental tissue. However, the data, although strong, is phenotypic and therefore an additional approach was required in an attempt to elucidate a mechanism by which Dis3L2 elicits its control over proliferation. In an attempt to identify the misregulated transcripts which may be responsible for the observed phenotypes RNA-sequencing was used in a global unbiased approach.

In this chapter RNA-sequencing was used to identify those transcripts that change in expression in the wing imaginal disc following the loss of *dis3L2*. This would allow the identification of those transcripts that are sensitive to Dis3L2 activity and therefore may be directly regulated by Dis3L2. Within the misregulated transcripts those that are direct Dis3L2 targets would be expected to increase in expression as a result of an increase in stability. It would therefore be hypothesised that the transcripts upregulated at a post-transcriptional level, rather than a transcriptional level, are the likely Dis3L2 targets. Using RNA-sequencing may also identify transcripts that encode proteins involved in specific pathways and therefore may give evidence towards a particular pathway through which Dis3L2 regulates proliferation in the developing wing imaginal disc.

RNA-sequencing is a relatively new technology which allows extensive screening of the entire transcriptome to identify differential expression between experimental conditions. Provided enough sequencing depth is achieved, RNA-sequencing can also be used to identify novel

transcripts. Due to its unbiased nature where the entire transcriptome can be examined it is generally preferred over the dated technique of microarrays. Although RNA-sequencing requires more complex bioinformatics it allows for the analysis of every transcript expressed in the tissue/cell of study. A direct quantification of expression is achievable which is normalised for the length of the transcript together with the number of sequencing reads each sample receives. This can be useful not only to investigate differential expression between experimental conditions but also to determine the level at which transcripts of interest are expressed between different tissues/cells in wild-type conditions. This also introduces another advantage of sequencing in that, provided enough read depth is achieved, it presents an opportunity to identify novel transcripts which is not possible with microarrays.

Before the emergence of RNA-sequencing microarrays were used to study the transcriptome, however, there are many limitations to the microarray technology. For example, the major limiting factor of a microarray is the space upon the hybridisation plate. For eukaryotic genomes it is not possible to screen the entire genome on a single plate meaning the user must select the transcripts they wish to screen. By definition this indicates how microarrays are dependent upon the current knowledge of the transcriptome. The requirement upon target selection means that potentially interesting and critical transcripts may be missed together with introducing an element of bias. RNA-sequencing overcomes this factor by enabling the user to screen the entire genome in an unbiased manner.

An additional major advantage of RNA-sequencing is in its sensitivity of detecting transcripts expressed at low levels. RNA-sequencing has been shown to be much more reliable in determining differential expression of these lowly expressed transcripts than microarrays (Marioni *et al.* 2008; Zhao *et al.* 2014). The issue with microarrays in detecting transcripts expressed at low levels comes from high levels of background hybridisation which will vary between transcripts due to the nature of the specific oligos on the microarray plate. A further issue with microarrays is the dynamic range of the detector; this not only affects the accuracy

in the detection of transcripts expressed at low levels but also provides a problem in the accurate quantification of large fold changes. The increased dynamic range together with the improved sensitivity of fold change determination in RNA-sequencing means that RNA-sequencing correlates much better with proteomic data than microarrays and are therefore more likely to lead to the identification of functional changes at the protein level (Fu *et al.* 2009).

This chapter discusses the experimental plan, sample preparation, read mapping and subsequent data analysis resulting in the identification of a pool of Dis3L2-sensitive genes, two of which are shown to be strong candidates as direct Dis3L2 targets.

## 5.2 Aims

The work outlined in the previous chapter has shown a novel role for Dis3L2 in the regulation of proliferation within the developing wing imaginal discs. In line with this the aims of this chapter were:

- To successfully perform and analyse a large scale RNA-sequencing experiment.
- To identify any pathways that may be affected by *dis3L2* depletion.
- To select suitable candidates for being Dis3L2 targets.
- To determine whether if the above candidates are regulated and the transcriptional or post-transcriptional level and thereby classify direct or indirect effects.

### 5.3 Experimental plan and sample preparation

Depletion of *dis3L2* throughout the wing imaginal disc (using *69B-GAL4*) results in increased proliferation within the disc which subsequently results in the overgrowth of both the wing disc and the adult wing blade. In an attempt to identify the mechanisms whereby the overgrowth is stimulated, RNA-sequencing was used in a global, unbiased approach to detect misregulated RNAs following the loss of *dis3L2* in the disc. Using RNA sequencing would allow the identification of RNAs that are both directly or indirectly affected by *Dis3L2*; misregulation of either or both of these groups could be responsible for the overgrowth phenotype.

The knockdown of *dis3L2* was achieved using *69B-GAL4* as it has previously been shown to result in the overgrowth phenotype. Additionally, the whole discs were used for the RNA extraction and therefore the *nub-GAL4* driver was less suitable to use as only the wing pouch region would show *dis3L2* knockdown. In this case the portion of the disc with wild-type *dis3L2* levels would dilute out the overall effects meaning any small changes would be missed. Crosses were set up at 25°C as previously carried out and wing imaginal discs were dissected from wandering L3 of knockdown and parental control larvae. As before the parental stocks were used as controls as they contained the most similar genetic background and any effects the specific insertions have on the levels of specific RNAs would be accounted for. RNA was extracted using a miRNeasy kit (Qiagen) from a total of 60 wandering L3 wing imaginal discs for each biological replicate of which there were 2 for each sample.

To obtain accurate, interpretable and reproducible results from RNA sequencing experiments the RNA used must be of high quality therefore two methods were used to assess the RNA integrity. The first was the 260nm/280nm and 260nm/230nm ratios which were measured using a NanoDrop 1000 and are summarised in Figure 5.1A. The 260/280 ratio represents the RNA to protein ratio and therefore will show any protein contamination. The 260/230 ratio was used to identify any contamination with organic solvents which normally results from carryover during the phenol-chloroform step. For an RNA sample to be of high enough quality

for sequencing both these ratios should be greater than 1.8. All the sampled showed a ratio of greater than 1.8 (Figure 5.1A). These ratios give an indication of the purity of the RNA sample; however, its integrity is crucial for sequencing experiments as degraded RNA could skew any data. Therefore to ensure the RNA was not degraded an Agilent Bioanalyser was used. The Bioanalyser works through the separation of RNAs by size in a process similar to gel electrophoresis. The RNA products are detected following the intercalation of a dye with the RNA structures resulting in a detectable fluorescent signal. Different sized RNAs pass the detector at a different speed allowing for the differentiation of RNA by size. A ladder is run alongside which contains RNAs of a known length used to identify the ribosomal RNAs which in turn are used to determine RNA integrity.

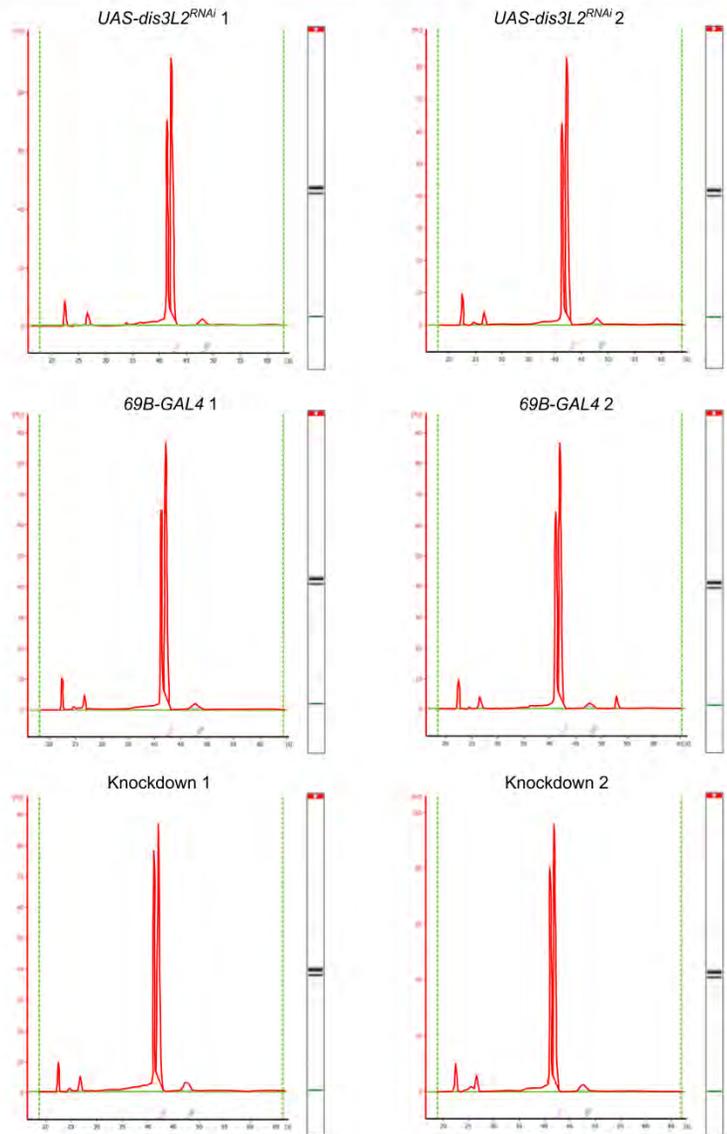
RNA of high integrity is shown by two clear peaks representing the 18S and 28S ribosomal RNAs with the ratio of these normally giving an RNA integrity number (RIN). However, the *Drosophila* 28S rRNA differs slightly from human and mouse in that it is 4.1Kb in size compared to 5Kb in humans. In addition, when heated the 28S rRNA splits in two and therefore when it runs on the gel it runs much closer to the 18S (which is 2Kb in size) and therefore the peaks appear much closer together than would be expected (Krupp 2005). This feature is consistent across a variety of insects (Winnebeck *et al.* 2010). The Bioanalyser used did not contain the specific settings required to calculate the RIN for *Drosophila* samples. Despite this the appearance of two clear, close peaks (together with two distinct bands) for each sample (Figure 5.1B) showed the RNA was suitable for sequencing experiments. Additionally, small peaks were observed around 25nt which represent miRNAs. If the RNA had been degraded these peaks would be larger due to the small fragments produced during the degradation process.

These tests showed that the RNA extracted from 60 wandering L3 wing imaginal discs was suitable for RNA sequencing and therefore 3µg of each sample was sent to Oxford Gene Technology for library preparation and RNA sequencing on an Illumina HiSeq 2500 machine.

A

Sample	260/280	260/230	Yield ( $\mu\text{g}$ in 12 $\mu\text{l}$ )	RNA sent for seq ( $\mu\text{g}$ )
<i>UAS-dis3L2<sup>RNAi</sup></i> 1	2.24	2.45	3.33	3.00
<i>UAS-dis3L2<sup>RNAi</sup></i> 2	2.24	2.52	3.25	3.00
<i>69B-GAL4</i> 1	2.21	4.82	5.70	3.00
<i>69B-GAL4</i> 2	2.27	6.59	4.70	3.00
Knockdown 1	2.10	2.20	5.21	3.00
Knockdown 2	2.13	2.36	4.97	3.00

B



**Figure 5.1: RNA samples sent for RNA-sequencing were of high integrity. (A)** Purity of the RNA samples prepared from 30 wing imaginal discs extracted using a Qiagen miRNeasy kit. All 260/280 and 260/230 ratios are above 2.10 showing that there is minimal contamination with organic solvents or protein respectively. **(B)** Bioanalyser traces of each sample show two distinct peaks representing the 18S and 28S rRNA.

## 5.4 Library preparation and RNA-sequencing performed by Oxford Gene Technology

The preparation of cDNA libraries and the RNA sequencing was outsourced to Oxford Gene Technology, a company previously used for a similar experiment. The data obtained from this experiment was of high quality and is published in Jones *et al* 2016 (Jones *et al.* 2016). The abundance of ribosomal RNAs in any RNA sample would result in them taking up a high proportion of the sequencing reads. Therefore to avoid this the libraries were poly(A) selected using oligo(dT) beads. This not only removes the highly abundant ribosomal RNAs but also isolates the mRNAs within the sample which were the RNA species of interest in this experiment.

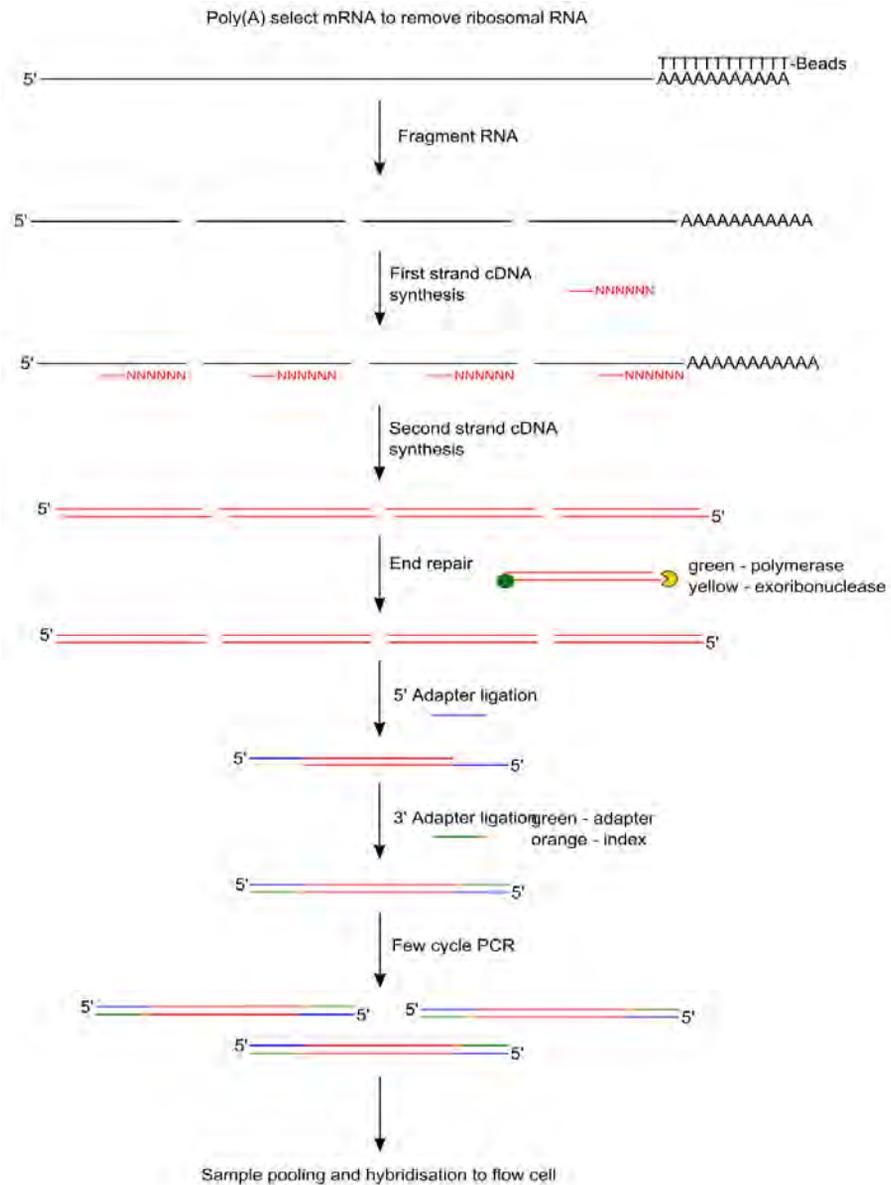
Library preparation was performed by Oxford Gene Technology using an Illumina TruSeq kit with the general work flow shown in Figure 5.2. RNA was first purified using oligo(dT) beads and fragmented by sonication into 200bp fragments. Following fragmentation samples were primed for first strand cDNA synthesis using random hexamers. This was then followed by synthesis of the second cDNA strand which removed the template RNA and results in double stranded cDNA.

The fragmentation of the initial mRNA molecules generates overhangs which are undesirable for the following adaptor ligation steps. Therefore the overhangs in the resulting cDNA were converted to blunt ends using a 3'-5' exonuclease to remove 3' overhangs and a polymerase to repair the 5' ends. Following end repair the 3' and 5' adaptors were ligated; however, prior to this the 3' ends were adenylated to prevent self-ligation and provide a required overhang for the ligation of the adaptor. The ligated adaptors have two functions; the first being for hybridisation to the flow cell during the cluster generation and the second being the addition of a priming site for the sequencing. The ligated 3' adaptor also contained the specific index sequences which allow for sample identification following the pooling the samples prior to the sequencing. Using the adaptors act as priming sites a few subsequent rounds of PCR were

performed to enrich the cDNA fragments that contain both the 3' and 5' adaptors. Minimal rounds of PCR were used to avoid generating errors which would subsequently be misrepresented in the sequencing.

Prior to cluster generation and sequencing all the RNA samples were pooled together. The libraries were pooled in equal volume with all being 10nM in concentration. After pooling the sample mix was made single stranded and passed across the flow cell where one of the ligated adaptors hybridises with complementary oligos bound to the flow cell. The cDNA strands that hybridise to the flow cell act as templates for the bridge amplification which subsequently results in the generation of clusters containing multiple, identical sequences. Optimal cluster densities are between 700-820K clusters/mm<sup>2</sup>.

The flow cell was then subjected to the sequencing reaction. A single, fluorescently labelled nucleotide is incorporated to the complementary strand during each cycle. Over the cycles the complementary strand is built for each hybridised cDNA molecule and each successfully incorporated base is identified or 'called' using base calling software. Each base carries a unique dye which makes base identification possible. Due to this the number of cycles is directly proportional to read length; in these experiments 100 cycles were used generating reads of 100 nucleotides in length. 6 additional, post-sequencing 'index' reads were then carried out to identify the index sequence which allows sample separation.



**Figure 5.2: RNA-seq library preparation work flow using the Illumina TruSeqRNA sample preparation kit.** RNA is poly(A) selected by using oligodT bound beads and subsequently fragmented into sequences of around 200nt. Fragmented RNA is then reverse transcribed into cDNA using random hexamers to prime the reaction. The second strand of cDNA is then synthesised generating double stranded cDNA (ds cDNA). Purified dscDNA contains overhangs which are repaired using either a polymerase (green) or exonuclease (yellow) to clean up the 5' and 3' ends respectively. Sequencing adapters (blue) are then added to the 5' end of the cDNA fragments followed by the ligation of the 3' adapters (green) containing the unique 6nt index sequences (orange) which allow multiplexing of the samples. A few rounds of PCR are then performed using adapter specific primers to positively select those fragments with successfully ligated adapters. The resulting cDNA library is complete and ready for hybridisation to the flow cell for the sequencing reaction. This protocol was completed by Oxford Gene Technology.

## **5.5 Read mapping and differential expression analysis.**

The sequencing run produces fastq files which contain information on the base calls and the quality of each base which is represented by a character (example shown in Figure 5.3A). These are the initial files used in an extensive mapping and analysis pipeline to identify differentially expressed mRNAs between the control and knockdown samples.

### **5.5.1 Assessing read quality using FastQC.**

The fastq files obtained from the HiSeq 2500 sequencing machine were used to assess the quality of the reads. A fastq file is produced for the forward and reverse reads from each sample. These were loaded into the program FastQC v0.11.2 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) which reads the fastq files giving a visual representation of base quality (example on fig 5.3Bi). In general all bases with a quality score of >30 (Q30) are regarded as high quality; bases falling below this threshold are removed in subsequent processing procedures. In addition to the quality of the sequencing FastQC was used to identify the total number of reads each sequence received together with overrepresented sequences. In each of the forward samples only one overrepresented sequence was identified; in each case the overrepresented sequence was an Illumina TruSeq adapter which occurred on average in 0.58% of the samples (Figure 5.3C). The identification of these sequences was critical for subsequent processing. During this initial analysis no sequences were flagged up as being 'poor quality'.

### **5.5.2 Adaptor removal and quality trimming using Sickle and Scythe**

Having observed the sequences were of high quality the adapters and any reads or sections of reads that were of insufficient quality required removal. Two algorithms were used to achieve adaptor removal and quality trimming of the reads.

Before the reads can be mapped to the *Drosophila* genome the adapter sequences must be removed. The algorithm Scythe v0.993b (<https://github.com/vsbuffalo/scythe>) was used to

remove the adapter sequences. Scythe uses base quality and a user-input sequence and probability of finding an adapter sequence to identify the adapter derived bases. Base quality aids Scythe in the identification of the 3' adapters as the quality tends to get poorer towards the 3' end and therefore the 3' adapter reads tend to be of poorer quality. In addition to base quality the identified adapter sequences and their probability of occurrence was input into the Scythe run. FastQC was previously used to identify any overrepresented sequences (5.5.1), in each case the only sequence corresponded to an Illumina adapter. This sequence was therefore taken as the adapter used for that specific sample. The sequences identified by FastQC were provided to Scythe in an adapters.fa file. Finally the probability of the adapter sequence occurring as identified by FastQC was supplied to the Scythe script. The default probability is 0.05%; however, FastQC analysis identified the average adapter occurrence as 0.58%. Therefore the probability identified for each sample by FastQC was input in the Scythe runs for the corresponding sample. Both the forward and reverse fastq files were submitted together with the same adapter sequence and probability. The resulting files were saved as <sample\_name>trim.fastq.

After Scythe mediated removal of the adapter sequences the quality of the remaining reads of each base were assessed and those that were substandard were removed. The bases at the 5' and 3' end of reads tend to be of lower quality, especially those at the 3' end which can be seen in Figure 5.3 Bi. The algorithm Sickle v1.29 (<https://github.com/najoshi/sickle>) was used to quality trim the read sequences so that only high quality bases remained for the mapping stages. Sickle is a 'sliding-window' algorithm using a window which is 0.1x the length of the read (10 bases in the 100 nucleotide reads) using stringent quality criteria to identify low quality bases. The window progresses along the read from 5' to 3' until a threshold for quality is passed; at this stage the bases upstream are removed. The window then continues throughout the read until the base quality drops below the threshold at which point all bases 3' of the site are cleaved. This leaves a read in which all the bases are of sufficient quality to ensure more reliable mapping. Any reads below 15 nucleotides after quality trimming were

removed entirely. As with Scythe both the forward and reverse reads for each sample were submitted together. The quality trimmed Sickle output files were saved as <sample\_name>qual.fastq. The base quality before and after quality trimming is shown in Figure 5.3B which indicates the increase in average base quality at all positions across the read. The only base position showing any example of quality below Q30 was at position 100, the last base which is always of the lowest quality. A comparison of the number of remaining reads together with the number of overrepresented samples and the read lengths for each sample is shown in Figure 5.3C. The adapter removal was confirmed by the absence of overrepresented sequences from all samples.

### **5.5.3 Mapping quality controlled reads to the *Drosophila* genome using TopHat2**

Following adapter removal and quality trimming the remaining high quality reads were mapped to the *Drosophila* genome using TopHat2 v2.0.12 (Kim *et al.* 2013). The file to which the reads are mapped is referred to as the 'index file' a .gff or .gtf file. This index file was previously built by a post-doc in the lab (Dr Chris Jones) using Bowtie v2.2.3 (Langmead and Salzberg 2012). The index file was built using Flybase version r6.01 (St Pierre *et al.* 2014) of the *Drosophila* genome using chromosomes X, Y, 2, 3 and 4 including ribosomal and mitochondrial genes.

TopHat2 works by breaking the reads down into smaller fragments and mapping them to the genome supplied in the index file. However, the default settings for TopHat2 are built in for sequencing human samples and therefore certain parameters were adapted to better suit *Drosophila* samples such as the minimum and maximum intron length. The minimum intron length (-i) was set to 20 whilst the maximum (-l) was set to 150000, these values were selected using Flybase. This enhances the accuracy of TopHat2 in mapping reads that span exons. Reads mapping to multiple locations within the genome can be problematic and therefore those reads mapping to more than 20 locations within the genome were discarded by TopHat2.

Some multiple mapping was allowed to take into account paralogues which are likely to contain partially identical sequences and therefore reads covering these regions would be expected to map to multiple sites.

The total number of reads that were successfully mapped as a pair to the *Drosophila* genome was output by TopHat2 which are summarised in Figure 5.4. The major output from TopHat2 was a file containing the genomic location of each read and therefore identifying the mRNA from which it was derived. This file was named `accepted_hits.bam` and was used in the subsequent stages of analysis.



Sample	Number of reads	Mapped reads (% input)	Multiple alignments (% input)	Aligned pairs (concordant alignment rate)
69B1	13415016	12785248 (95.3%)	20000 (1.6%)	12409992 (91.9%)
69B2	13013021	12412194 (94.4%)	201167 (1.6%)	12062191 (92.0%)
UAS1	15056303	14343531 (95.3%)	229376 (1.6%)	13922153 (92.0%)
UAS2	14785002	13949276 (94.3%)	1193629 (8.6%)	13577094 (86.4%)
KD1	12316130	11740542 (95.3%)	170801 (1.5%)	11392866 (92.0%)
KD2	14039460	13357783 (95.1%)	224153 (1.7%)	12957806 (91.6%)

**Figure 5.4: Summary of read alignment following the application of TopHat2.** The number of input reads is presented for each sample with 95% being successfully mapped to the genome. Only a small fraction of reads map to multiple places within the genome; those that mapped to >20 positions were discarded. On average 91% of the read pairs mapped to the same location in the genome as would be expected (concordant alignment rate) which shows a highly successful TopHat2 run.

## 5.5.4 Using Cufflinks to assemble the reads successfully mapped by TopHat2

Read alignment is a critical step, however it does not quantify the total number of reads each transcript receives. For this to occur the aligned reads must be assembled and normalised to both the total number of reads each sample received and the length of the transcript. The Cufflinks package was used to assemble, merge and quantify the number of reads for each transcript (Trapnell *et al.* 2012).

Cufflinks is an algorithm designed to identify the smallest possible set of transcripts that explains the mapping of all the read pairs. To aid it in its determination of this the same index file as used by TopHat2 was supplied which allows cufflinks to quantify transcript abundance after their assembly. To improve the accuracy of transcript abundance a number of optional parameters were run such as 'frag bias correct' (-b) which runs an algorithm to correct for any read bias. As mentioned previously those multiple mapping reads which mapped to less than 20 locations within the genome were allowed. To improve accuracy of the weighting of these reads the 'multi read correct' (-u) function was used. There can also be issues with accuracy in the estimation of the abundance of genes expressed at lower levels, therefore the 'upper quartile norm function' (-N) was used to improve this. This function uses a slightly different method of normalisation in that it uses the upper quartile of the number of reads mapping to individual loci rather than the total number of reads a sample receives. Finally to ensure only those fragments that map to the reference transcript are used in the abundance estimates the '-compatible-hits-norm' function was used. Cufflinks and TopHat2 are able to make estimations of isoform abundance through the mapping of reads which cross splice sites. The pile up of reads crossing exon junctions unique for each isoform is used to generate an estimation of the abundance of each isoform.

Cufflinks was run with these settings on all 6 samples individually; however, it is important to pool the samples together for downstream analysis for both consistency and to save computer

hours. The merging of samples was achieved using Cuffmerge but it was important to merge the samples after Cufflinks had assembled the transcripts as merging before this can generate complex splice site decisions which are prone to error. Cuffmerge performs a reference annotation-based transcript (RABT) assembly generating a single .gtf file containing the merged assemblies from each Cufflinks run.

The “merged.gtf” file produced by Cuffmerge was then used with Cuffquant which further quantifies transcript abundances and provides a pre-calculation of gene expression levels. Like Cufflinks, Cuffquant requires the “accepted\_hits.bam” file produced by TopHat2 for each sample and therefore has to be run independently for each sample. As used in Cufflinks the ‘frag bias correct’ (-b) function was again included to improve the accuracy of the estimate of transcript abundance. Cuffquant produces an abundance.cxb file for each sample which was then used for the differential expression.

### **5.5.5 Using the Cuffdiff pipeline to identify differential expression**

Using Cufflinks, Cuffmerge and Cuffquant resulted in the generation of files required to assess the differential expression between the control and *dis3L2* knockdown samples. To achieve this Cuffdiff was used. The abundance.cxb files for each sample and their condition (i.e. control or knockdown) was supplied in a single text file whilst the comparisons to be made were provided in an additional text file which defined the conditions. The “merged.gtf” file created by Cuffmerge was also required together with the ‘frag bias correct’ (-b) function which was again used to enhance accuracy. Finally due to only having biological replicates rather than triplicates the ‘-min-reps-for-js-test’ had to be set to 2, however the confidence in the statistical output using an n of 2 is very low. The higher the replicate count the more accurate Cufflinks can be due to learning the natural variation of gene expression within replicate groups.

The differential expression was assessed in two ways; the first being a comparison between 3 'conditions' with the *UAS-dis3L2<sup>RNAi</sup>* and *69B-GAL4* parental controls kept separate in addition to the *dis3L2* knockdown sample. This allowed the comparison between the two parental controls to assess major differences within their transcript profiles. The other comparison contained the grouping of the 4 parental replicates as a 'control' condition against the *dis3L2* knockdown samples. The results and rationale behind these differential expression analyses are discussed below. Cufflinks provides an output showing the fold change of each transcript between the conditions in addition to the raw p-value and the p-value corrected for multiple comparisons (q-value) which allows the differential analysis between the defined conditions.

An overview of the programmes used in the analysis pipeline and the pipeline itself are shown in Figures 5.5 and 5.6.

Programme	Used for	Input	Output
FastQC	Initial quality analysis and identification of adapter sequences	<ul style="list-style-type: none"> <li>Raw fastq files obtained from sequencer</li> </ul>	<ul style="list-style-type: none"> <li>Average quality score for each base position</li> <li>Overrepresented sequence (adapter)</li> </ul>
Scythe	Adapter removal	<ul style="list-style-type: none"> <li>Text file containing Adapter sequences identified by FastQC.</li> <li>Probability of adapter occurrence.</li> <li>Forward and reverse fastq files for sample.</li> </ul>	<ul style="list-style-type: none"> <li>Individual files for each forward and reverse fastq files containing adapter trimmed reads &lt;sample_name&gt;.trim.fastq</li> </ul>
Sickle	Quality trimming	<ul style="list-style-type: none"> <li>Forward (-f) and reverse (-r) adapter trimmed fastq files &lt;sample_name&gt;.trim.fastq</li> </ul>	<ul style="list-style-type: none"> <li>Individual files for each forward and reverse fastq files containing quality trimmed reads &lt;sample_name&gt;.qual.fastq</li> </ul>
TopHat2	Alignment to reference genome	<ul style="list-style-type: none"> <li>Forward and reverse quality trimmed files &lt;sample_name&gt;.qual.fastq</li> <li>Reference genome (index file(.gff)) and file location</li> </ul>	<ul style="list-style-type: none"> <li>Folder containing aligned reads in an accepted_hits.bam file for each sample</li> </ul>
Cufflinks	Assembly of transcripts	<ul style="list-style-type: none"> <li>Accepted_hits.bam file</li> <li>Reference genome (.gff file)</li> <li>Fasta genome reference file (.fa) for bias detection (-b)</li> </ul>	<ul style="list-style-type: none"> <li>Transcript assembly file - transcripts.gtf for each sample</li> </ul>
Cuffmerge	Merging of cufflinks assembled transcripts	<ul style="list-style-type: none"> <li>Text file ("assemblies.txt") containing transcripts.gtf file for each sample.</li> <li>Reference genome (.gff file)</li> <li>Fasta genome reference file (.fa) for bias detection (-b)</li> </ul>	<ul style="list-style-type: none"> <li>Merged.gtf file containing the collation of cufflinks built assemblies for all samples</li> </ul>
Cuffquant	Final quantification of transcript abundance	<ul style="list-style-type: none"> <li>Merged.gtf file</li> <li>Fasta genome reference file (.fa) for bias detection (-b)</li> <li>Accepted_hits.bam (from TopHat2) for the specific sample</li> </ul>	<ul style="list-style-type: none"> <li>Files containing the quantification of each transcript – "abundances.cxb"</li> </ul>
Cuffdiff	Perform differential expression between specific conditions	<ul style="list-style-type: none"> <li>Text file containing the abundances.cxb for each sample being used "cdsample_sheet_&lt;samples&gt;.txt" and the conditions under which they fall ( e.g control/knockdown)</li> <li>Text file containing the conditions to be compared as outlined in sample sheet "contrasts_&lt;samples&gt;.txt"</li> <li>Fasta genome reference file (.fa) for bias detection (-b)</li> <li>Merged.gtf file from Cuffmerge</li> </ul>	<ul style="list-style-type: none"> <li>Multiple files outlining the differential expressions between: <ul style="list-style-type: none"> <li>Genes</li> <li>Isoforms</li> <li>Promoter usage</li> <li>Transcriptional start sites</li> </ul> </li> </ul>

**Figure 5.5: Overview of the programs and algorithms used in the RNA-seq analysis pipeline including the input and output files for each one.**

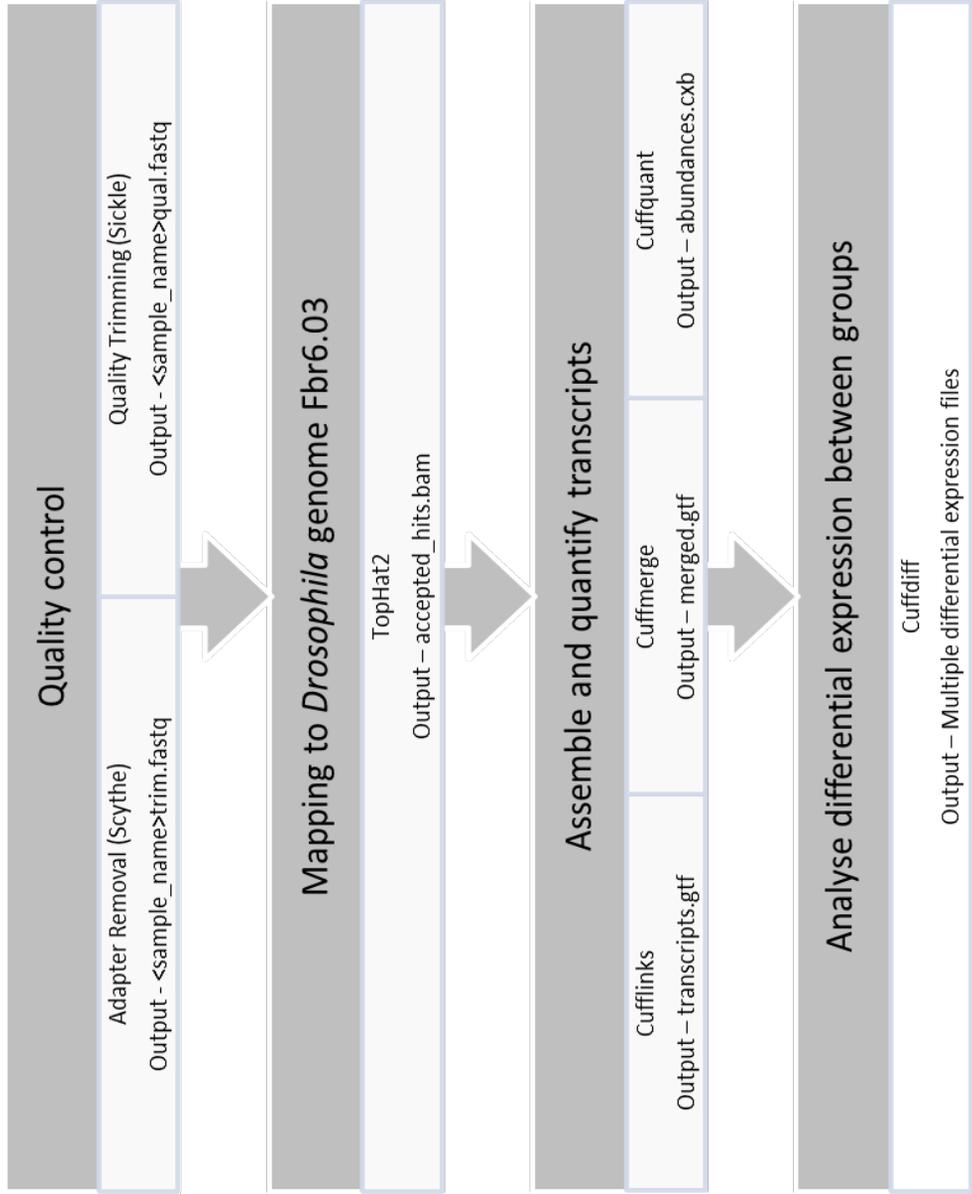


Figure 5.6: Flow diagram showing the RNA-seq analysis pipeline and the output file for each stage.

## 5.6 Global overview of the RNA-sequencing data.

The successful application of the above analysis pipeline resulted in the generation of differential expression analyses which were used to evaluate the effect of *dis3L2* knockdown on the RNA expression profile of wing imaginal discs. In addition to the potential identification of differentially expressed transcripts the output from the pipeline also allowed the visualisation of the expression distribution and the identification of what genes were expressed in the wing imaginal discs.

### 5.6.1 Global expression distributions identify differential expression between control and *dis3L2* knockdown samples.

Having achieved normalised read counts for each sample the patterns of expression were analysed between the control and *dis3L2* knockdown conditions. 8,837 genes of the 16,258 expressed in *Drosophila* had a normalised read count in all 6 samples and therefore could be confidently determined as expressed within the wing imaginal disc. The normalised read count or 'fragments per kilobase of transcript per million mapped reads (FPKM)' is calculated during the cufflinks analysis and normalises against the differing lengths of transcripts and for the number of reads each sample received. The levels of expression within the control wing imaginal discs ranged from an FPKM of 0.0066 (*CG42534*) to 23147 (*mt: ND1*) indicating the vast range of gene expression within the tissues. The global expression distributions were analysed using kernel density plots which generate plots representing the density of genes in terms of their expression levels. The density score is a representation of the probability of expression per  $\log_{10}$ (FPKM).

Cuffdiff was initially run keeping all 3 genotypes separate to identify any differences between the two parental groups and each parental group and the *dis3L2* knockdown sample. The first comparisons were made between the *UAS-dis3L2<sup>RNAi</sup>* and *69B-GAL4* parental control stocks to assess any global changes in expression as large differences between control groups would have to be taken into account in further analyses. The kernel density plots of the two parental

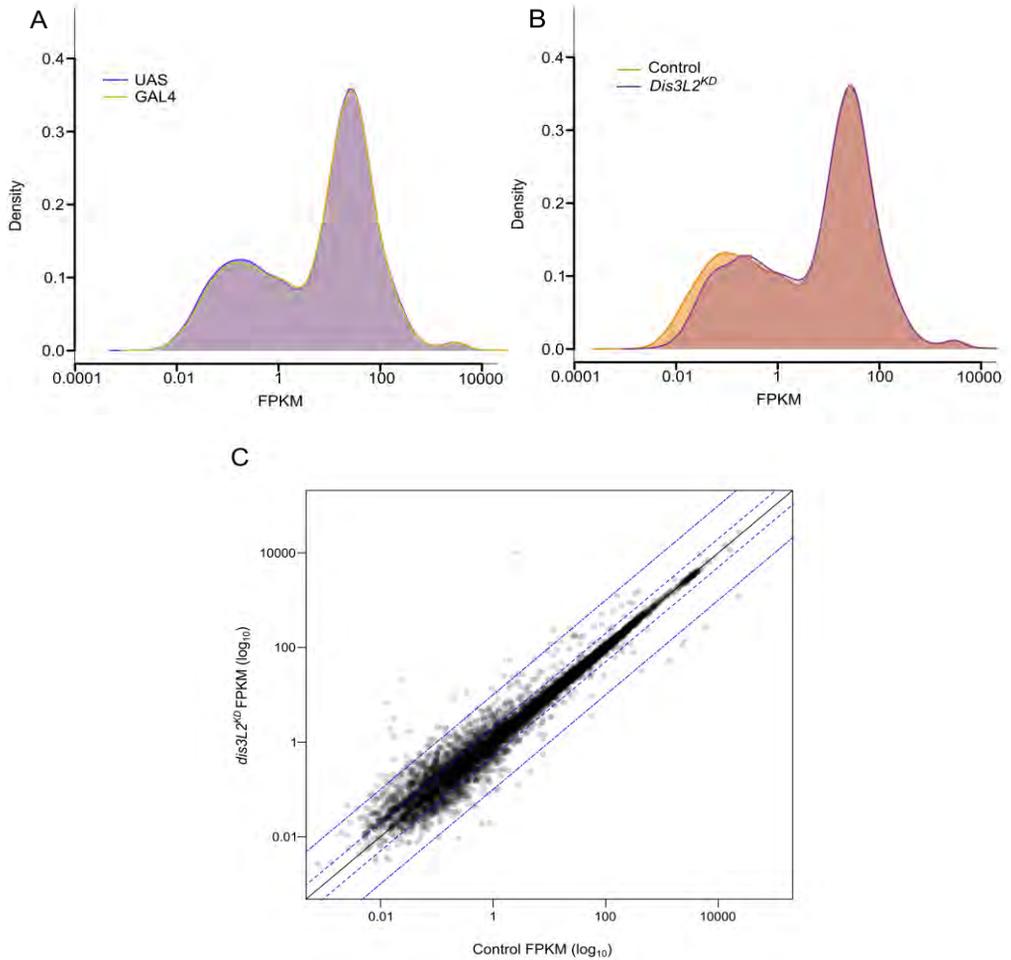
controls appear to neatly overlap indicating minimal differences in the global expression profiles (Figure 5.7A/B). In these control plots there were three major 'peaks' of gene expression; the first, furthest left, represents genes at a low level of expression (FPKM <1), the middle peak represents genes which are moderately expressed (FPKM ~ 100) and finally the small peak shows transcripts expressed at very high levels (FPKM >5000). This expression distribution within the wing discs was consistent with the other sequencing experiment within the lab published in Jones et al 2016 (Jones *et al.* 2016). To ensure there was no statistical significance between the control expression profiles a Kolmogorov-Smirnov (KS) test was performed. A KS test is a statistical test designed to compare probability distributions such as those in kernel density plots. The KS test between the two parental distributions showed there was indeed no significant difference ( $p=0.56$ ). As a result the two parental controls groups were pooled to give a 'control' this also increased the number of control samples to 4 allowing greater confidence in the differential expression.

A second Cuffdiff run was performed including only 2 'conditions'; a control condition including all 4 parental control samples and a knockdown condition containing the 2 *dis3L2* knockdown samples. A comparison of the overall expression distributions between the control and knockdown samples showed a statistically significant difference (KS test -  $p=1.041 \times 10^{-6}$ ). This difference is evident in the kernel density plots where the genes expressed at low levels appear to shift right indicating a higher level of expression. This shift could be expected in samples in which an exonuclease had been knocked down. This was the only difference observed in the kernel density plots as the peaks representing the genes at moderate and high levels of expression were unchanged between the control and knockdown profiles (Figure 5.7B).

In addition to the kernel density plots as a means of viewing overall changes in gene expression scatter plots were also used. Kernel density plots give an idea of a general trend in expression changes; however, the scatter plots allowed a visual overview of transcripts that

both increase and decrease in expression. A number of transcripts appear to be differentially expressed between the control and *dis3L2* knockdown samples (Figure 5.7C) with a similar number lying above and below the  $y=x$  line (solid black line). A few genes showed very large increases or decreases in expression which are represented by the points lying outside the 10 fold lines (blue dot-dash lines), however, most genes that were differentially expressed showed changes between 2 (blue dotted line) and 10 fold. The majority of the genes showing differential expression were within the lower region of expression (FPKMs  $<1$ ) which was consistent with the expression distribution plots. The dense region around the  $y=x$  line indicated that most genes remain unchanged in between the two conditions.

It was therefore clear that some genes were differentially expressed following the knockdown of *dis3L2*. However, their consistency and reliability required further assessment to identify those whose misexpression may be culpable for the overgrowth phenotypes previously observed.



**Figure 5.7: RNA-seq differential expression overview. (A)** A kernel density plot, built on the Cuffdiff differential expression analysis, shows that there is no overall difference in the expression profiles between the two parental control lines. UAS genotype (blue) = ; *UAS-dis3L2<sup>RNAi</sup>* ; GAL4 genotype (orange) = ; *69B-GAL4*. A Kolmogorov-Smirnov (KS) test shows no significant difference between the profiles ( $p=0.5555$ ). Density represents the probability of a transcripts expression falling within that bracket of normalised expression. **(B)** A Cuffdiff differential expression analysis shows that there is a significant difference between the grouped control (containing both parentals in **A**) and *dis3L2<sup>KD</sup>* expression profiles (KS test  $p=1.041e^{-6}$ ). Control (orange) = ; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *69B-GAL4*. *Dis3L2<sup>KD</sup>* (purple) = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+*. Density represents the probability of a transcripts expression falling within that bracket of normalised expression. **(C)** Scatterplot showing the comparison in normalised read count between Control (; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *69B-GAL4*) and *dis3L2<sup>KD</sup>* (; *UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+*) wing imaginal discs. Solid black line is  $y=x$  where the majority of transcripts are located representing no differential expression. Blue dotted lines represent +/- 2-fold change whilst blue dot-dash lines represent +/- 10-fold change.

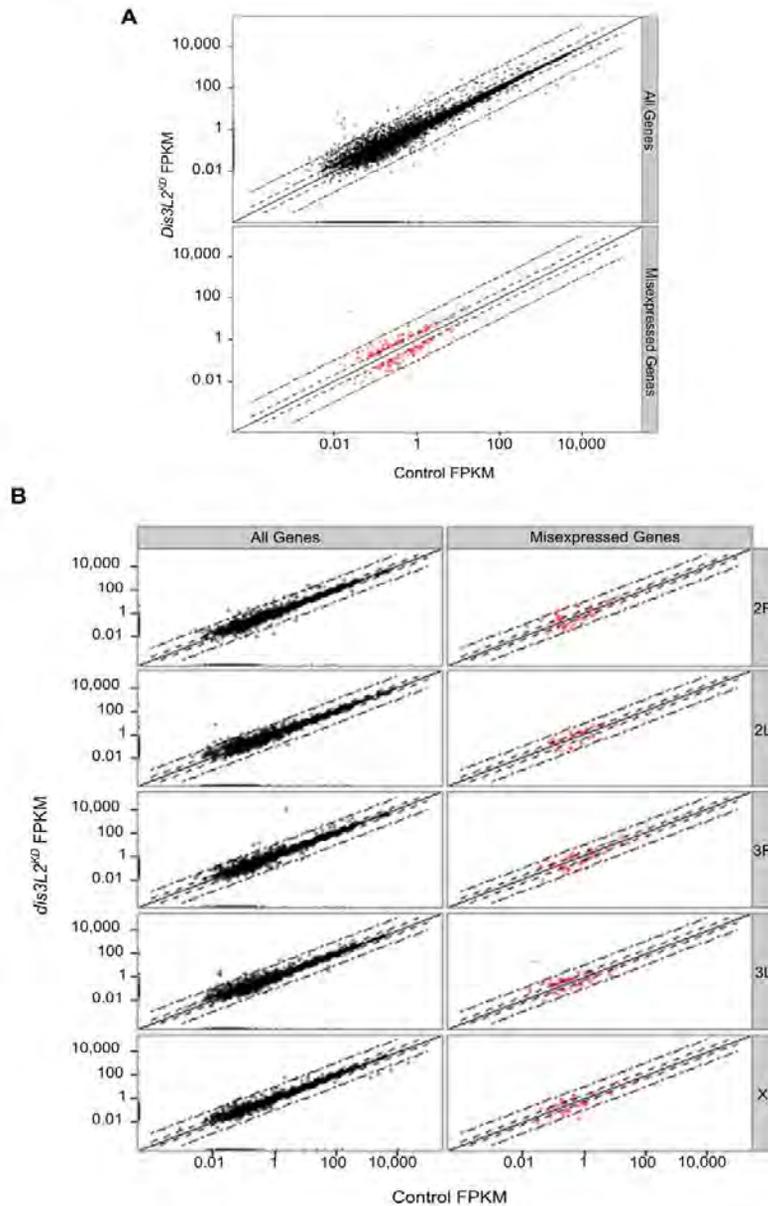
## 5.7 Selection of confidently misexpressed transcripts.

The transcripts that change in expression following the knockdown of *dis3L2* were largely of genes expressed at low levels in the control discs. However, this brought with it an issue of reliability. It is known that those genes with low FPKMs tend to be more variable and the fold changes less reproducible than those expressed at higher levels (McIntyre *et al.* 2011). Therefore it was critical to identify those genes that change consistently across all the replicates as they were most likely to be the 'real' changes in expression. A number of parameters were set for the selection of genes that were able to be confidently assigned as misexpressed.

As mentioned previously it was of great importance to identify those genes that show reproducible fold changes following the loss of *dis3L2* within the wing disc. To achieve this, a strict set of filtering parameters were used. For example, the fold change was assessed between each individual knockdown sample and each control replicate resulting in 8 fold change values. 6 of these 8 comparisons had to show a fold change of >1.5 fold to be taken forward for further analysis. Multiple comparisons across the individual replicates were used as the statistical analysis performed by Cuffdiff was unreliable due to only having knockdown samples in duplicate. In performing the multiple comparisons it was possible to filter out those genes that were showing an 'overall' fold change which was predominated by a single replicate. Those transcripts showing infinite up or down regulation were excluded from the data set as this implies the transcript was either not detected in any control sample (infinite up) or in either knockdown samples (infinite down) and therefore the reliability would be questionable. The filtering resulted in the number of upregulated transcripts (showing an average fold change of >1.5) reducing from 1345 to 125. Although the initial interest was in the identification of direct *Dis3L2* targets this filtering process was performed on both the up and down regulated transcripts. The filtering again lead to a reduction of the number of

misexpressed transcripts from 717 to 114. Within the 114 downregulated genes was *dis3L2* (*CG16940*) as would be expected.

The selected transcripts are represented in Figure 5.8A where they are 'pulled out' of the scatter showing all the identified genes (misexpressed genes). Figure 5.8A shows that the majority of the transcripts that either increase or decrease in expression were fairly modest in their changes mainly lying between the 2 (dash) and 10 (dot-dash) fold increase lines. It is also clear that the majority of the differentially expressed transcripts were in the low expression bracket. This does indicate that fold changes in transcripts with a lower FPKM can be consistent. To ensure there was no bias within the selection, in addition to no technical bias during the read mapping the selected transcripts were plotted in respect to the chromosome on which they are situated. This showed that there was no chromosome bias as the consistently misexpressed transcripts were evenly distributed across X and the right and left arms of 2 and 3 (Figure 5.8B). Male specific genes were not of interest in this study.



**Figure 5.8: Graphical representation of the consistently misexpressed transcripts in *dis3L2<sup>KD</sup>* wing imaginal discs. (A)** Scatterplot showing the transcript expression difference between Control (; *UAS-dis3L2<sup>RNAi</sup>* ; and ;*69B-GAL4*) and *dis3L2<sup>KD</sup>* (; *UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+*) wing imaginal discs. The top panel represents all those transcripts detected in the discs (All Genes, black) whilst the bottom represents the 229 transcripts determined as consistently differentially expressed using stringent filtering procedures (Misexpressed Genes, red). Solid black line is  $y=x$ , blue dotted lines represent +/- 2-fold change whilst blue dot-dash lines represent +/- 10-fold change. **(B)** The differentially expressed genes are located across all chromosomes showing there is no chromosome bias in either the mapping (All Genes, black) or differential expression determination (Misexpressed Genes, red). Solid black line is  $y=x$ , blue dotted lines represent +/- 2-fold change whilst blue dot-dash lines represent +/- 10-fold change.

## **5.8 Validation and analysis of selected misexpressed transcripts.**

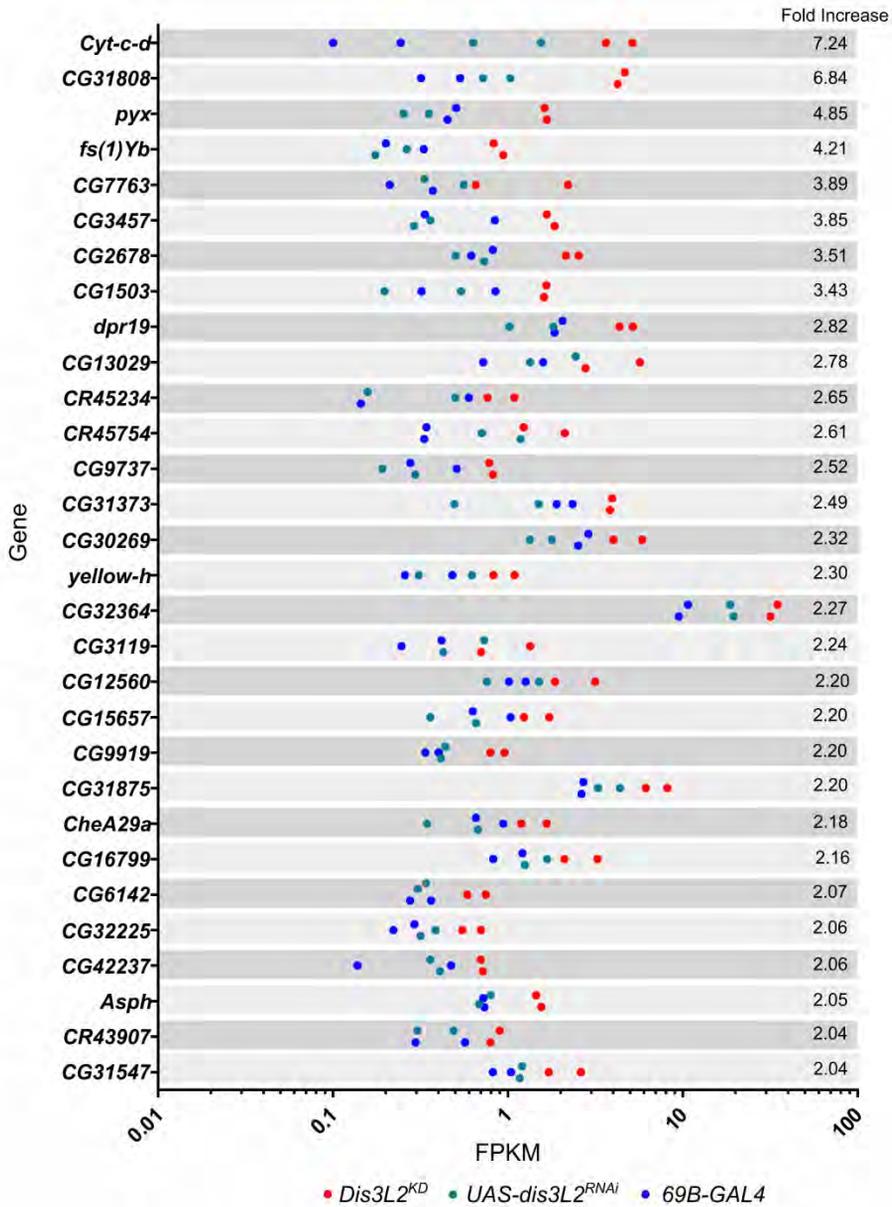
The stringent filtering criteria resulted in the confident identification of differentially expressed transcripts in *dis3L2* knockdown wing imaginal discs. 125 transcripts showed a consistent increase in expression with an average fold increase of 9.74 fold. However this fold change was heavily dominated by a single transcript, *Hsp70Bb* which showed an increase of 879 fold. The average fold change of those transcripts increased in expression excluding *Hsp70Bb* was a more modest 2.73 fold. Comparatively the average fold change of the 114 transcripts which consistently decreased in expression following *dis3L2*-depletion was 3.57 fold. It was therefore apparent that the knockdown of *dis3L2* within the wing imaginal disc resulted in a similar number of increased and decreased transcripts with similar magnitudes of change. To visualise the consistency of the selected transcripts strip plots were used showing the individual replicates for the top 30 up- (Figure 5.9) and downregulated (Figure 5.10) transcripts. The efficiency of the filtering is evident in the clustering of the knockdown (red) and control (blue) replicates with the clear distinction between them.

### **5.8.1 Gene ontology analysis of consistently misexpressed transcripts.**

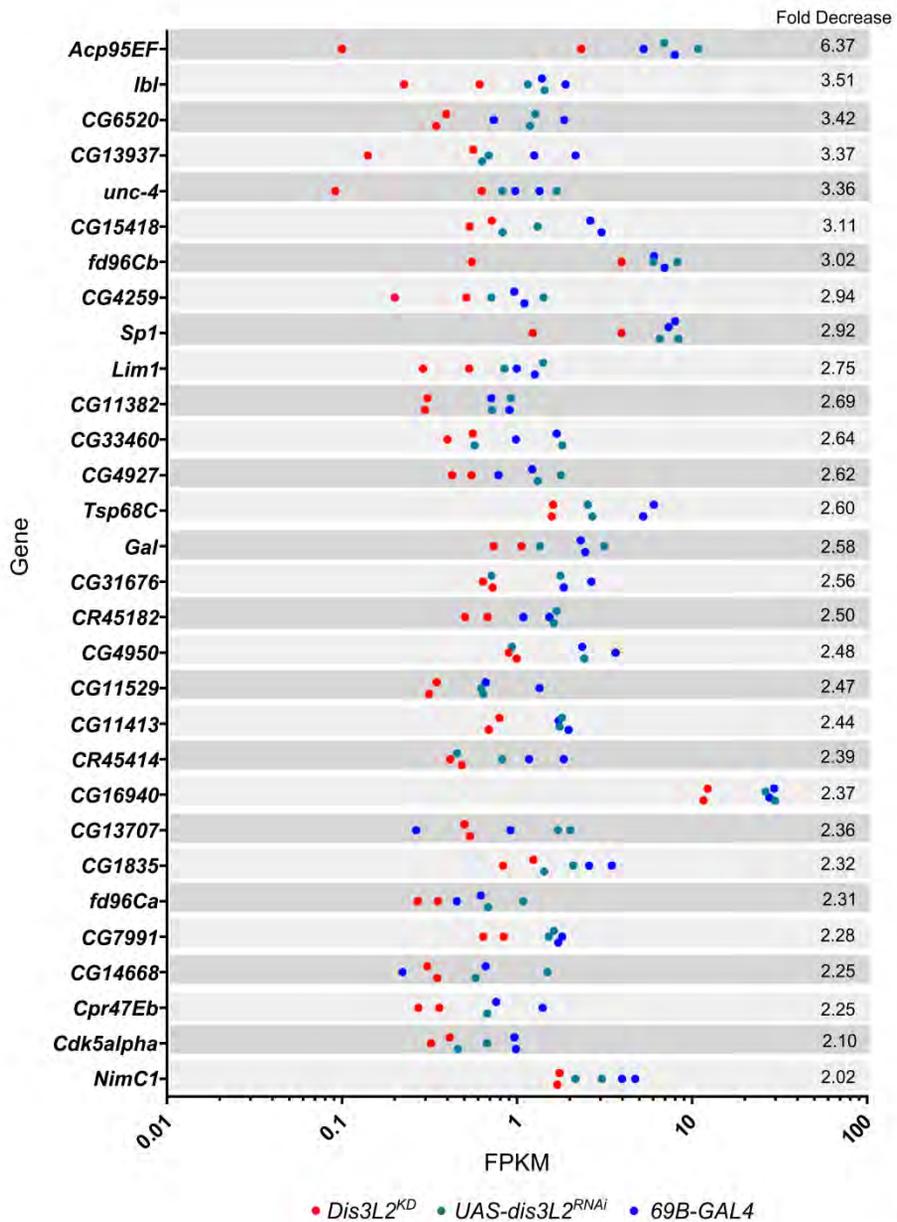
Having filtered out likely false positives it was of interest to identify any pathways that were enriched within the differentially expressed transcripts as it may give an idea towards the pathways in which *Dis3L2* is involved in regulating. Unfortunately the low number of input transcripts made it more difficult to identify large scale enrichment. The 239 misregulated transcripts were submitted to DAVID, an online gene ontology (GO) algorithm. DAVID searches against the background genome for the number of genes which, relative to the size of the input list, would be expected to be identified within each category. An enrichment score is calculated as  $-\log_{10}(\text{p-value})$  and therefore any score  $>1.3$  represents a significant enrichment when using 0.05 (5%) as a significance cut off. The GOtermBPFAT setting was used which groups genes on their annotated biological function.

A limitation of DAVID is that it relies on previous annotation of a gene's biological function which for a high proportion of the misregulated transcripts was unknown. However, two significant categories were identified which are summarised in Figure 5.11. The first of these was transcriptional regulation which was represented by 8.37% of the total input whilst the other was cell morphogenesis which was represented by 5.44% of the total input. Within the transcriptional regulation group, which had an enrichment score of 1.91, were a group of homeobox transcription factors. However, 80% of the genes within the group showed downregulation following *dis3L2* knockdown and are therefore unlikely to be direct Dis3L2 targets. In contrast the cell morphogenesis category consists of a similar number of up and downregulated genes.

Therefore, overall there does not appear to be substantial enrichment of a particular pathway as the enrichment scores of even the two significant categories are relatively low. This is likely to be due to the low number of input transcripts.



**Figure 5.9: Strip plot showing replicate consistencies for the top 30 upregulated transcripts.** Graphical representation of the replicate consistency for the 30 RNAs showing the greatest upregulation in *dis3L2* knockdown discs with an average control FPKM of more than 0.3. Knockdown replicates (*Dis3L2<sup>KD</sup>* = ; *UAS-dis3L2<sup>RNAi</sup>*/+ ; *69B-GAL4*/+ shown in red whilst control replicates (; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *69B-GAL4*) are shown in shades of blue. Fold change is also presented for each upregulated RNA.



**Figure 5.10: Strip plot showing replicate consistencies for the top 30 downregulated transcripts.** Graphical representation of the replicate consistency for the 30 mRNAs showing the greatest downregulation in *dis3L2* knockdown discs with an average control FPKM of more than 0.3. Knockdown replicates (*Dis3L2<sup>KD</sup>* = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+*) shown in red whilst control replicates (; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *69B-GAL4*) are shown in shades of blue. Fold change is also presented for each downregulated RNA.

Category	Enrichment score (-log <sub>10</sub> (p-value))	Gene count (% of input)
Regulation of Transcription	1.91	20 (8.37%)
Cell morphogenesis	1.31	13 (5.44%)

**Figure 5.11: Gene ontology analysis of the 229 misexpressed transcripts in *dis3L2<sup>KD</sup>* wing imaginal discs.**

## 5.8.2 Validation of misexpressed genes by qRT-PCR.

Being an exoribonuclease the direct targets of Dis3L2 are most likely to be those genes that are upregulated following its depletion. New samples of imaginal discs were dissected and used for the validation. If the misregulated transcripts were true hits then their expression would be expected to change in similar manners across independent *dis3L2* knockdown crosses. However, any non-specific changes caused for example by sample handling would not be expected to be consistent in the fresh samples. Each fresh sample consisted of RNA extracted from 30 wandering L3 wing imaginal discs for each parental genotype and *dis3L2* knockdown (n=4 for each). To mimic the conditions used for the library preparation oligo(dT) primers were used to prime the reverse transcription reaction rather than the random primers used previously. Oligo(dT) primers were used at the same concentration as the random primers in the high capacity reverse transcription kit used previously (see methods section 2.4.2.3). The successful knockdown of *dis3L2* within these new samples was first confirmed by qRT-PCR. *dis3L2* was successfully knocked down by 75% which was consistent with samples initially sent for sequencing. For a gene to be validated by qRT-PCR it would have to be expressed at a level detectable by the technique resulting in a Ct of <35. To ensure this was the case an average control FPKM cut off was set to 0.3. This value was set using an initial round of validation where *fs(1)Yb* (an average control FPKM of 0.26 but was retained due to showing reliable validation) was detected with an average control Ct of 34 with consistent results between replicates. However, *bam* (an average control FPKM of 0.05) was not detectable in either the control or knockdown samples. An FPKM cut off of 0.3 has also been used previously in other RNA-seq studies and therefore this was determined a suitable cut off (Ramsköld *et al.* 2009; Hart *et al.* 2013).

Transcripts showing the largest fold changes together with an average control FPKM of  $\geq 0.3$  were selected for validation. 6 of the top 7 genes were selected for validation by qRT-PCR; *CG7763* was not taken forward due to showing high variation between the two knockdown

replicates despite showing upregulation against all controls evident in Figure 5.9. 5 of the 6 selected transcripts showed comparative fold changes between the RNA-seq and qRT-PCR experiments following *dis3L2* knockdown. For example *cytochrome-c-distal (cyt-c-d)* showed the greatest upregulation using the above cut offs with an increase of 7.24 fold in the RNA-seq data; a comparative fold increase of 7.36 was detected by qRT-PCR. Similarly *CG31808* showed consistency between the two techniques with fold increases of 6.84 and 5.97 by RNA-seq and qRT-PCR respectively. *fs(1)Yb* and *pyrexia (pyx)* both also showed significant upregulation by qRT-PCR although the magnitude of increase was half that observed in the RNA-seq data. The final gene to show reproducible upregulation was *CG2678* which showed an increase of 3.51 fold in the RNA-seq data. A significant upregulation of *CG2678* of 2.11 fold was observed by qRT-PCR. However, despite showing a moderate increase of 3.85 fold by RNA-seq, *CG3457* did not show a significant increase by qRT-PCR. A large difference was observed between the two parental genotypes for *CG3457* with the knockdown samples showing an increase 6.16 of fold against the *UAS-dis3L2<sup>RNAi</sup>* parent whilst no change was observed against the *69B-GAL4* parent. The cause of the variation between the parental controls is not clear, however, it is enough to remove *CG3457* from any further investigation. The comparisons between the RNA-seq and qRT-PCR can be seen in Figures 5.12 and 5.13.

In addition to those genes outlined above that show reproducible increases in expression following the knockdown of *dis3L2* an additional gene was validated. *gastrulation defective (gd)* showed a modest, but consistent 1.97 fold increase. *gd* was increased in another RNA-seq data set produced by the lab and therefore to investigate if the same was true for *dis3L2* knockdown discs it too was validated by qRT-PCR. Indeed, *gd* showed a significant increase of 1.79 fold by qRT-PCR indicating it too was a reproducibly differentially expressed. However, its increase in expression in *pacman (pcm)* mutants is much larger (22.5 fold, (Jones *et al.* 2016)) and therefore the change in the *dis3L2* knockdown discs may not be biologically relevant. Overall with 6 of the 7 selected transcripts showing reproducible increases by RNA-seq and qRT-PCR it is clear that the filtering out of false positives was largely successful and those

remaining within the list were likely to represent true changes in expression following the loss of *dis3L2* within the wing imaginal discs.

### **5.8.3 *CG2678* and *pyrexia* show the characteristics of direct Dis3L2 targets.**

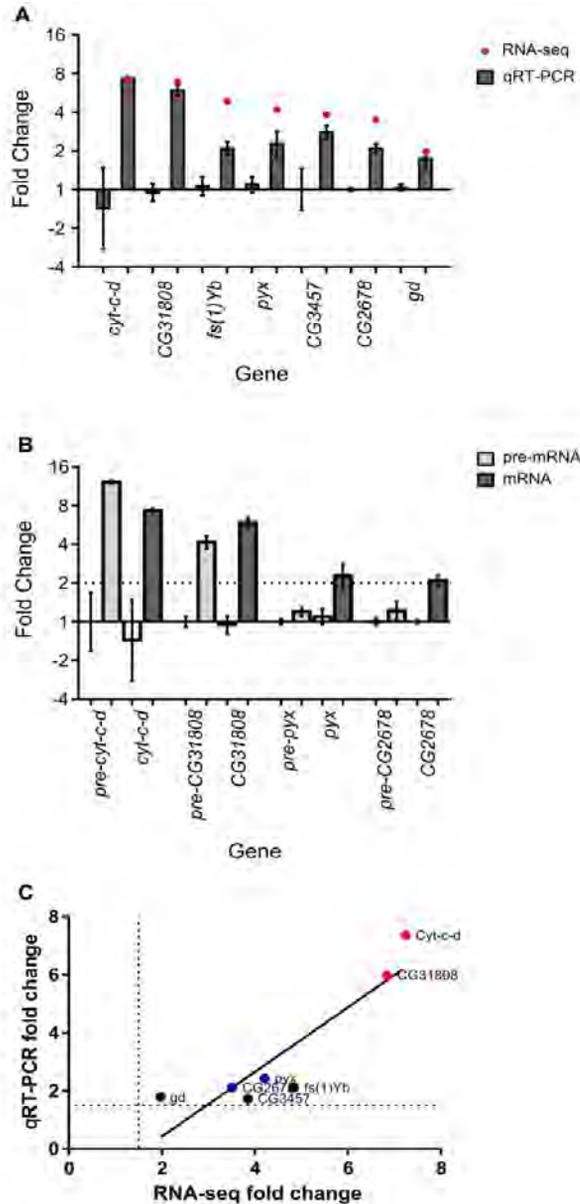
The previous section outlined the reproducibility of the RNA-seq data set and identified those genes that show the largest fold increase following *dis3L2* knockdown within the wing imaginal disc. These are the most likely direct targets of Dis3L2. Dis3L2 functions at the post-transcriptional level to regulate gene expression and therefore Dis3L2 targets would be expected to change in expression at the post-transcriptional level. In contrast, those that alter in expression as an indirect effect of *dis3L2* knockdown may be expected to change at the transcriptional level. To separate the potential direct Dis3L2 targets from those likely to be indirect affects, in terms of post-transcriptional and transcriptional changes, qRT-PCR was again used. To achieve this the pre-mRNA was used as a 'transcriptional read out'. Custom TaqMan assays were designed using the online tool from life technologies. A 100 nucleotide sequence containing 50 nucleotides either side of an intron/exon boundary was submitted for 4 upregulated transcripts. *cyt-c-d*, *CG31808*, *pyx* and *CG2678* were investigated in this manner (Figure 5.14).

Using the above assay those mRNAs that were regulated at the post-transcriptional level and therefore potential direct Dis3L2 targets, would not be expected to show a change in the levels of their respective pre-mRNA. However, those that that show a comparable pre-mRNA increase would represent a transcriptional change which is likely to be a result of an indirect effect of *dis3L2* knockdown. The two transcripts showing the greatest upregulation, *cyt-c-d* and *CG31808*, both showed transcriptional regulation. The *pre-cyt-c-d* precursor displayed a 6.56 fold increase compared to the 7.36 fold increase in the mature transcript. Similarly *pre-CG31808* showed a 4.25 fold increase which again was comparative to the 5.97 fold upregulation in mature *CG31808* (Figure 5.12). This therefore indicates that both *cyt-c-d* and

*CG31808* were upregulated at the transcriptional level following *dis3L2*-depletion in the wing imaginal disc.

Interestingly both *cyt-c-d* and *CG31808* neighbour each other on the left arm of chromosome 2 with their 5'UTRs overlapping in opposing directions. It could therefore be possible that the cause of their increased expression is due to co-regulation, for example by a transcription factor which itself may be sensitive to Dis3L2 activity. Chromatin immunoprecipitation experiments from the modENCODE consortium (Negre *et al.* 2011) has identified binding sites for transcription factors such as Trithorax-like (*Trl*), Dichaete (*D*), Distal-less (*Dll*) and Dorsal (*dl*) within the overlapping region of *cyt-c-d* and *CG31808*. It is possible that their activity is altered following the loss of Dis3L2 in the disc which in turn drives the observed increase in expression of *cyt-c-d* and *CG31808*. The RNA-seq data shows no change in expression of *Dll* and *dl* following the *dis3L2* knockdown (Figure 5.15). *D* shows an average downregulation of 2.23 fold, however, there is large variation between the two knockdown replicates as shown in (Figure 5.15) and therefore the confidence within this fold change is limited. *Trl* expression was not detected within the sequencing data set. Despite the limited/no changes in expression at the RNA level there may still be functional changes of these factors following the loss of *dis3L2* such as phosphorylation events.

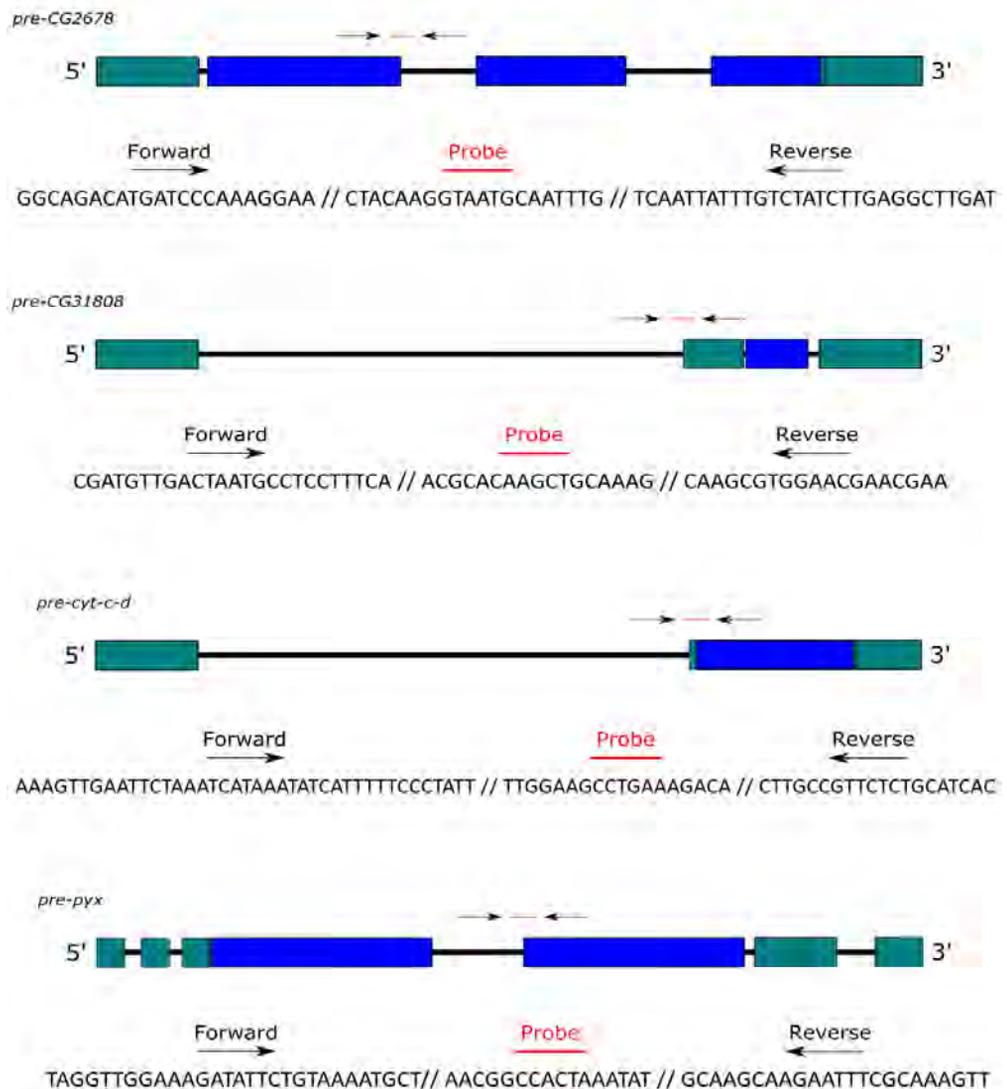
In contrast to *cyt-c-d* and *CG31808* post-transcriptional changes were observed for both *pyx* and *CG2678*. Both *pre-pyx* and *pre-CG2678* do not change significantly in expression following *dis3L2* knockdown which therefore indicates that there is no increase in the level of their transcription (Figure 5.12). This instead implies a change in the stability of the mature mRNAs; a likely cause of this increased stability is due to the loss of Dis3L2 function. It is therefore a strong possibility that both *pyx* and *CG2678* are normally targeted by Dis3L2 for degradation.



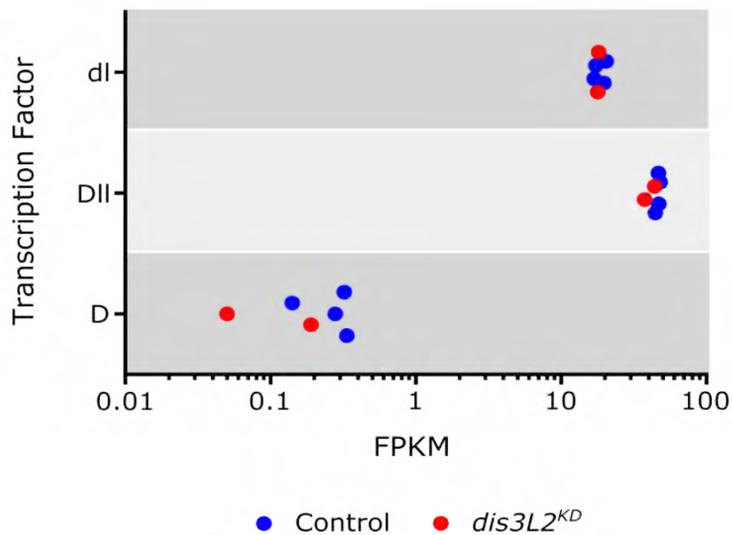
**Figure 5.12: RNA-seq validation using qRT-PCR. (A)** Comparison between the fold changes of transcripts by RNA-seq (red dots) and qRT-PCR (grey bars). Transcripts showed expression increases by both techniques although the magnitude of fold change varied slightly. Dotted line shows 2-fold increase.  $n \geq 4$ , error bars represent standard error. All except CG3457 validated. See Figure 5.13 for p-values. **(B)** *pyrexia* (*pyx*) and CG2678 show post-transcriptional changes in gene expression as only the mature mRNA shows a significant increase (dark grey). *cyt-c-d* and CG31808 show transcriptional changes as both the pre- (light grey) and mature (dark grey) mRNA increase in expression. Dotted line shows 2-fold increase.  $n \geq 4$ , error bars represent standard error. See Figure 5.13 for p-values. **(C)** Differential expression status is consistent between RNA-seq and qRT-PCR excluding CG3457. Those transcripts showing transcriptional changes highlighted in red whilst those showing post-transcriptional changes are highlighted in blue. Transcriptional vs post-transcriptional nature of the expression increases for *fs(1)Yb* and *gd* was not investigated.

Gene	RNA-seq fold change	qRT-PCR fold change	mRNA qRT-PCR p-value	Pre-mRNA qRT-PCR p-value
<i>cyt-c-d</i>	7.25	7.36	0.0423	0.0052
<i>CG31808</i>	6.84	5.97	<0.0001	<0.0001
<i>fs(1)Yb</i>	4.84	2.43	0.0412	N/A
<i>pyx</i>	4.21	2.11	0.0175	0.0709
<i>CG3457</i>	3.85	1.72	0.927	N/A
<i>CG2678</i>	3.51	2.11	<0.0001	0.2611
<i>gd</i>	1.97	1.79	0.0072	N/A

Figure 5.13: Summary of RNA-seq validation by qRT-PCR.



**Figure 5.14: Sequences of custom TaqMan assays used to determine transcriptional vs post-transcriptional changes in expression.** Solid black line represents introns, turquoise blocks represent UTR's and blue blocks represent exons. Assays were designed to span intron/exon boundaries. The positions of the forward and reverse primers (black arrows) and the fluorescent probe (red line) are shown.



**Figure 5.15: Transcription factors predicted to bind *cyt-c-d* and *CG31808* do not show expression changes in *dis3L2<sup>KD</sup>* wing imaginal discs.** There is no difference in normalised expression (FPKM) of *Dichaete* (*D*), *dorsal* (*dl*) or *distal-less* (*dll*) in *dis3L2<sup>KD</sup>* wing imaginal discs (red dots, genotype = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+*) compared to control discs (blue dots, genotypes = ; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *69B-GAL4*).

## 5.9 Chapter Summary

In using a global, unbiased approach, the work presented in this chapter has identified a number of genes that are sensitive to Dis3L2 expression. The raw sequencing reads have been successfully aligned to the *Drosophila* genome and analysed using the Cufflinks pipeline to determine differentially expressed transcripts in *dis3L2* knockdown wing imaginal discs. Within this list of differentially expressed genes were those that were both directly and indirectly affected by *dis3L2* knockdown. Using qRT-PCR it was possible to take a set of these differentially expressed transcripts and determine if there were likely to be directly or indirectly affected by the loss of *dis3L2*. This identified two potential Dis3L2 targets in *pyrexia* and *CG31808*. There are likely to be a number of additional Dis3L2 targets in the wing imaginal disc, however, due to technical and financial limitations, which are discussed below, they were the only two detected in this work.

## **5.10 Chapter Discussion.**

### **5.10.1 A lack of confidence in statistical analysis by Cuffdiff.**

The identification of consistently misexpressed genes is the essential factor behind the identification of potential Dis3L2 targets. In the analysis presented here this was achieved through using individual replicate comparisons with an arbitrary fold change and an FPKM cut off. This procedure was time consuming although it did produce a list of differentially expressed genes, which in general, was reproducible by qRT-PCR. Using stringent filtering was critical as following the filtering the number of misexpressed transcripts (those showing a <1.5 fold change) reduced from 2,062 to 239. This significant reduction in the number of misexpressed transcripts to those that could be determined as confidently misexpressed shows the high level of potential false positives that are often present in global data sets such as sequencing and emphasises the importance of filtering. On the other hand it is likely that a number of potential Dis3L2-targeted transcripts were excluded due to over-stringent cut offs.

Cuffdiff, the algorithm used to determine transcript differential expression, outputs a statistical analysis together with the FPKM and fold changes. Differential expression is determined as statistically significant if the p-value is lower than the false discovery rate of 0.05. This would normally be an incredibly useful tool to drive the selection of consistently misexpressed genes as was shown by a recent paper published from the lab (Jones *et al.* 2016). However, even though the parental genotypes were grouped to give an n of 4 there were only duplicates of the knockdown samples. This therefore minimises the statistical power and renders it almost useless. The Cuffdiff parameters were altered to allow statistical analysis on a condition containing only 2 replicates; however the subsequent statistical output was not used in further analysis. The statistics performed by Cuffdiff relies upon a larger sample size to learn the parameters of natural variation across each condition and a sample size of n is simply not sufficient. This is shown as only 34 genes are determined to show a significant difference between knockdown and control samples. Of these 34 the vast majority show large variations

between the two knockdown replicates with many being heavily dominated by a single replicate.

This therefore removes any confidence in the statistical analysis on the small sample size and was thus not used in the downstream analysis. Any future experiments of a similar nature should use a larger sample size such that the Cuffdiff statistical analysis can be made full use of. For example Jones et al used a sample size of 6 and the statistical analysis was strong and reliable; this would aid in the determination of both the fold change and FPKM cut offs.

### **5.10.2 Why do so few mRNAs change in expression?**

As outlined in the previous section the low replicate number means that stringent criteria had to be set in order to have the upmost confidence in the determination of differentially expressed genes. Unfortunately there were differences between the two *dis3L2* knockdown replicates which meant that a number of genes may have been missed. If a larger sample size was used initially it would have made it easier to rule out a particular sample should it have behaved differently to all others.

The reasons for the differences could be many; for instance it could just represent natural variation and therefore present real biological differences. Alternatively, it could have been due to sample treatment during the tissue dissection or RNA extraction. Precautions were taken to ensure the conditions under which the samples were collected were as consistent as possible. The differences are unlikely to be due to differential knockdown efficiencies as the *dis3L2* FPKM in both knockdown replicates is very consistent. Furthermore both knockdown samples showed comparable levels of knockdown by qRT-PCR. A final possibility for the apparent differences between the samples could be due to the read coverage as knockdown replicate 1 received the lowest read coverage across the six samples. Although the level of expression (in terms of FPKM) is normalised to the total number of reads received per sample the reduced depth may mean that genes expressed at low levels are not detected as reliably as in knockdown replicate 2.

The slight variation between the two knockdown replicates is therefore likely to contribute to the low number of genes differentially expressed following *dis3L2* knockdown. However, another reason so few are detected could be the nature of using RNA interference. Although the knockdown, as shown by qRT-PCR, is efficient at 80% the remaining 20% of functioning Dis3L2 would be present therefore masking some of the effects. To overcome this using a null mutant would be preferable, however, during this work a suitable mutant was not available.

The low FPKMs of the majority of the differentially expressed genes also introduced reliability issues in the determination of misexpressed genes. Genes with low FPKMs are known to show increased variability (McIntyre *et al.* 2011) and are therefore less likely to validate by qRT-PCR. Additionally those genes showing very low expression are difficult to detect at a reliable cycle threshold when using qRT-PCR unless a high concentration of RNA is used. Due to the nature of this work very high RNA yields would require a large number of wing imaginal discs and therefore would prove to be very time consuming. As a result an arbitrary FPKM cut off of 0.3 was used as genes at this level of expression were shown to be detectable by qRT-PCR. This does however mean that any genuine Dis3L2 targets of lower expression would have been missed. Unfortunately, the majority of genes showing expression changes showed FPKMs of <1 with a number falling beneath the 0.3 cut off resulting in a relatively low number detected potential Dis3L2 targets.

Alternatively, the low number of misregulated mRNAs could represent the real biological conditions and that Dis3L2 is responsible for the regulation of a discrete set of mRNAs. Similar work using *pacman* mutants in the lab showed again a relatively small number of misexpressed transcripts (although 10 fold greater than the number in this work) which may further provide evidence for transcript regulation by specific exoribonucleases. The relatively low expression of the potential Dis3L2-sensitive mRNAs may represent the biological function of Dis3L2 – to maintain those RNAs at low levels to prevent pro-proliferative stimuli. This experiment focused on mRNAs through using Poly(A) selection during the library preparation

and therefore if Dis3L2 functions to regulate other non-coding classes of RNA they would not be detected here. Either way the low number of misexpressed genes also makes pathway analysis difficult as a small sample input often leads to a lack of specific category enrichment.

To validate these findings in addition to identifying further Dis3L2 targets another RNA-sequencing experiment should be performed using ribo-depleted RNA, ideally from a null mutant, with a greater number of replicates to account for technical variation.

### **5.10.3 *pyx* and *CG2678* appear to be direct targets of Dis3L2.**

Despite the limitations outlined in the previous sections (5.10.1/5.10.2) the RNA-sequencing experiment has successfully identified 2 likely Dis3L2 targets in *pyrexia* and *CG2678*. Both these transcripts have been shown to increase significantly in expression following the knockdown of *dis3L2* in the *Drosophila* wing imaginal disc. This change occurs at the post-transcriptional level as the levels of their respective pre-mRNAs remain unchanged showing that the increase in expression is not due to an increase in their transcription. This is in contrast to *cyt-c-d* and *CG31808* which both increase in expression following *dis3L2* knockdown at the transcriptional level and therefore the increase is likely to be a result of indirect effects of *dis3L2*-depletion on transcription.

How might Dis3L2 specifically recognise *pyrexia* and *CG2678* for degradation and not other transcripts? There are a number of possibilities involved 3' terminal tagging and the association and specific targeting through co-factors. Work in human tissue culture cells has completed co-immunoprecipitation experiments and failed to identify specific co-factors which may aid Dis3L2 in specific RNA targeting (Lubas *et al.* 2013). However, this work used FLAG-tagged Dis3L2 and therefore the tag may disrupt potential protein-protein interactions. The same work saw an association of Dis3L2 with translating ribosomes and therefore this may be a method of specificity (Lubas *et al.* 2013). Additional work has identified an alternative recruitment mechanism. Mutation in *dis3L2* in *S. pombe* cells or human tissue culture cells results in the accumulation of uridylylated transcripts thus identifying 3' terminal tagging as a

potential recruitment mechanism (Malecki *et al.* 2013; Thomas *et al.* 2015). This is in accordance with the recruitment mechanism identified in Dis3L2 mediated *pre-let-7a* degradation (Chang *et al.* 2013; Ustianenko *et al.* 2013). It is likely that this mechanism of Dis3L2 recruitment is conserved and with the recent discovery of Tailor, a terminal uridylyl transferase (Reimao-Pinto *et al.* 2015), makes it a possibility in the system used here.

It would have been interesting to look for a potential increase in uridylylated transcripts in the RNA-seq data especially in terms of the likely Dis3L2 targets. Terminal uridylation often occurs either 3' of the poly(A) tail or following deadenylation. Therefore, unfortunately the poly(A) selection used in the library preparation denied this method of analysis. This was an oversight as if they were retained it may have been possible to begin to elucidate a targeting mechanism. It would have been more suitable to include ribosome depletion using beads containing rRNA specific oligos which would have removed the issue of rRNAs taking up the majority of the read coverage removing the requirement of poly(A) selection. Given more time it may have been possible to follow up the Dis3L2-sensitive transcripts in an aim to identify terminal uridylation using techniques such as cRACE (Rissland *et al.* 2007; Rissland and Norbury 2009) or tail-seq (Chang *et al.* 2014).

The post-transcriptional changes in *pyrexia* and *CG2678* represent an increase in their stability and therefore identifies them as strong Dis3L2 target candidates, however further work would be required to confirm this. For example, Dis3L2 in physical association with the targets is required to confirm their nature as *bona fide* Dis3L2 targets. In order to achieve this, experiment such as CLIP would need to be performed where Dis3L2 would be crosslinked to its targets, pulled out and subsequently sequenced or probed by qRT-PCR against the potential targets. This currently is not possible as an antibody to Dis3L2 is not available; however, the production of an antibody is currently in process.

As opposed to the direct targets of Dis3L2, its knockdown resulted in the increased expression of *cyt-c-d* and *CG31808*. These changes were at the transcriptional level and are therefore

most likely to be a result of indirect effects of Dis3L2 knockdown. Interestingly, both *cyt-c-d* and *CG31808* neighbour each other on the left arm of chromosome 2 with their 5'UTRs overlapping. It could therefore be possible that the cause of their increase expression is due to co-regulation, for example by a transcription factor which itself may be sensitive to Dis3L2 activity. Chromatin immunoprecipitation experiments from the modENCODE consortium (Negre *et al.* 2011) has identified binding sites for transcription factors such as Trithorax-like (Trl), Dichaete (D), Distal-less (Dll) and Dorsal (dl) within the overlapping region of *cyt-c-d* and *CG31808*. It is possible that their activity is altered following the loss of Dis3L2 in the disc which in turn drives the observed increase in expression of *cyt-c-d* and *CG31808*. Since their expression at the RNA levels remain unchanged following *dis3L2* knockdown, it is possible that they are regulated by post-translational modifications which could also be indirectly affected by *dis3L2* knockdown.

## 5.11 Acknowledgements

The library preparation and sequencing was performed by Oxford Gene Technology. The index file used in TopHat2 was built by Dr Chris Jones.

# Chapter 6: Pacman and Dis3L2 function in distinct pathways to regulate wing disc development

## 6.1 Introduction.

Within the cytoplasm there are a variety of mechanisms of RNA decay. In addition to conventional 3'-5' decay by Dis3L2 or the exosome an RNA can be decapped and degraded in a 5'-3' direction or it can be endonucleolytically cleaved. Cleavage of the RNA generates two fragments which are degraded in the 3'-5' and 5'-3' directions. Degradation from the 5' end in the cytoplasm is carried out by a single processive 5'-3' exoribonuclease named Pacman/Xrn1. The previous 2 chapters have focused on a role for Dis3L2 in proliferation and in the identification of potential Dis3L2 targets. However, it is likely that a level of co-operation and/or interplay occurs between the two pathways.

Dis3L2 has been shown in previous chapters to play a role in controlling proliferation and growth of the wing imaginal disc which is in keeping with the effect loss of *dis3L2* has in the human overgrowth condition of Perlman syndrome. However, thus far the mechanisms by which it exerts this function is unknown. Interestingly, previous work in the Newbury lab has identified a role for the 5'-3' exoribonuclease Pacman in the regulation of apoptosis through controlling the stability of the pro-apoptotic genes *hid*, *grim* and *reaper* at the post-transcriptional level (Waldron *et al.* 2015). An additional role for Pacman in the regulation of the insulin-like peptide *dilp8* was more recently identified which may indicate a role in the control of developmental timing (Jones *et al.* 2016).

Apoptosis is a critical process where a cell sacrifices itself for the good of the organism. This can either be triggered by the identification of DNA damage, the onset of cellular stress or it can also be a developmentally controlled process. For example, apoptosis has been well characterised to play a critical role in the development of the *Drosophila* eye (Wolff and Ready

1991). Apoptosis, or programmed cell death, is a highly conserved process (Twomey and McCarthy 2005; Crawford *et al.* 2012) and has been extensively studied in a number of organisms with its deregulation a key factor in cancer development. In *Drosophila* the onset of apoptosis is promoted by the activity of three 'tumour suppressor' proteins, Hid, Grim and Reaper (Rpr) which inhibit the in activity of the *Drosophila* inhibitor of apoptosis protein (DIAP1) (Berthelet and Dubrez 2013). This allows the formation of the apoptosome including the initiator caspase (Dronc) which in turn cleaves and activates the effector caspases (Dcp-1 and Dr-ICE) (Steller 2008; Pang *et al.* 2015). This process is similar in mammals; however, the onset of apoptosis is largely promoted by the release of Cytochrome c from the mitochondria, although it can also be stimulated by the activity of ARTS, SMAC and HtrA2 which function in a similar role to Hid, Grim and Reaper. Both these pathways results in the activation of the Dronc homologue, Caspase 9 and subsequent activation of the effector caspases, Caspase 3 and 7 (Salvesen and Abrams 2004).

The apparent role of Pacman in apoptotic regulation is in striking contrast to that of Dis3L2 in proliferation (Chapter 4). This led to the hypothesis that the two opposing pathways of RNA degradation may be used at least in part to control opposing pathways of cell proliferation and survival. The co-operation between the two pathways may therefore be critical in controlling tissue growth and developmental. Therefore this chapter set out to explore the interactions between cytoplasmic 3'-5' and 5'-3' decay within a developmental tissue.

## 6.2 Aims.

The work in this chapter aimed to explore the interactions and specificities of the cytoplasmic 3'-5' and 5'-3' pathways of RNA decay to test the hypothesis that the two opposing pathways of cytoplasmic RNA decay regulate distinct cellular processes. The specific aims were:

- To explore the phenotypic effects of knocking down *dis3L2* and *pcm* simultaneously within the wing imaginal discs.
- To knockdown *dis3L2* in both hypomorphic and null *pacman* mutant background and characterise any phenotypic effects.
- To analyse the specificity and/or redundancy of the two pathways through cross analysis of RNA-sequencing data sets.

## 6.3 Knockdown of *dis3L2* and *pacman* results in an additive effect on *Drosophila* viability.

In this thesis it has been shown that the loss of *dis3L2*, although not affecting viability, results in the overgrowth of the wing imaginal discs and subsequently the adult wing. In contrast the knockdown of *pacman* (*pcm*) in the wing disc using RNA interference results in smaller wings together with an increase in apoptosis specifically within the wing pouch region (Waldron *et al.* 2015). Together with this, previous work in *S. cerevisiae* has shown that loss of function mutations in *dis3L2* and *xrn1* is synthetically lethal (Malecki *et al.* 2013). This work however, was in single cell organisms therefore it was of interest to examine the effect of the simultaneous knockdown of *dis3L2* and *pcm* both ubiquitously and throughout the wing imaginal discs.

### **6.3.1 Simultaneous knockdown of *dis3L2* and *pcm* throughout the organism results in lethality.**

Ubiquitous knockdown of *dis3L2* has no effect on organism viability indicating that Dis3L2 is dispensable for *Drosophila* development. In contrast ubiquitous knockdown of *pcm* results in pupal lethality (Figure 6.1) which mimics the null *pcm*<sup>14</sup> phenotype (Waldron *et al.* 2015). To investigate if the ubiquitous loss of *dis3L2* alongside ubiquitous *pcm* knockdown results in an additive effect they were knocked down simultaneously (Figure 6.1A) using the *tub-GAL4* driver. The *tub-GAL4* driver is balanced over TM6 and therefore sibling controls would be identifiable by either the Tubby (Tb) phenotype during the larval/pupal stages or the Humeral (Hu) phenotype in the adults. Simultaneous ubiquitous knockdown of both *dis3L2* and *pcm* results in lethality prior to the 3<sup>rd</sup> instar larval stage of development as no non-Tb L3 or pupae (Figure 6.1B) were observed. This was confirmed in the adults as no non-Hu adults are observed (Figure 6.1C). This indicates an additive effect following the simultaneous knockdown as no adults are observed when *pcm* is knocked down alone (Figure 6.1C) although *pcm* knockdown pupae are observed. This therefore shows that the other degradation pathways are unable to completely compensate for the loss of Dis3L2-mediated 3'-5' cytoplasmic decay. Additionally the pupal lethality observed when *pcm* is knocked down alone shows that Dis3L2 or the cytoplasmic exosome are unable to fully compensate for its loss and thus shows some pathway specificity exists.

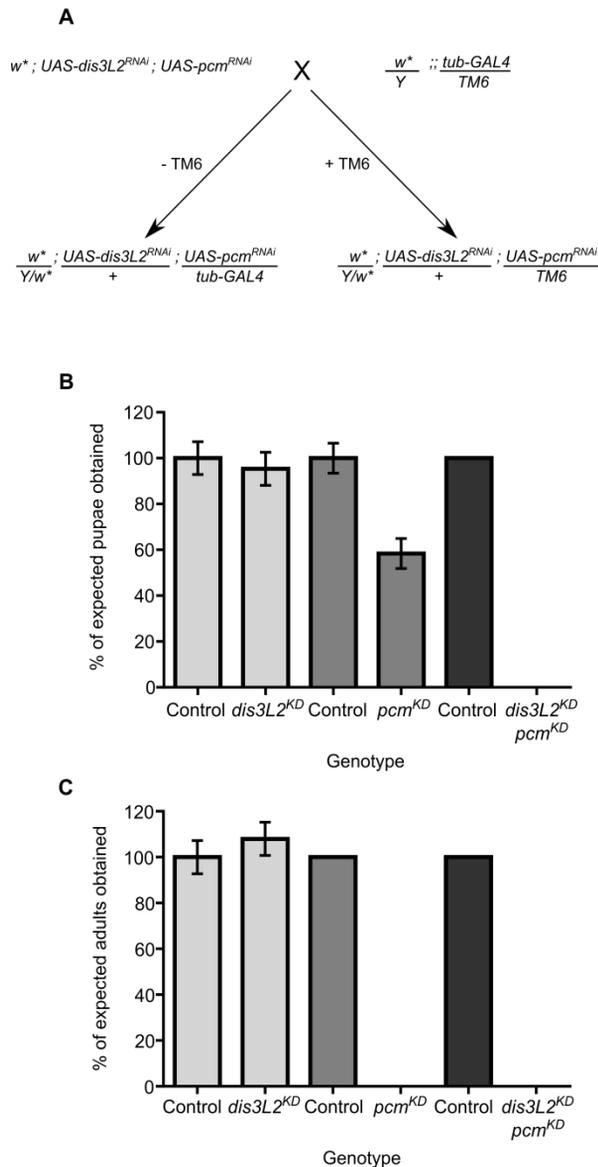
### **6.3.2 Simultaneous knockdown of *dis3L2* and *pcm* within the wing imaginal discs results in additive effects.**

Having determined that enhanced lethality was observed when *dis3L2* and *pcm* are knocked down together in the same organism it was of interest to explore this in more detail using the wing imaginal discs as a model system. Similar crosses were used as in 6.4.1 although the wing disc specific drivers, *69B-GAL4* and *nub-GAL4* were used and therefore no sibling controls were produced.

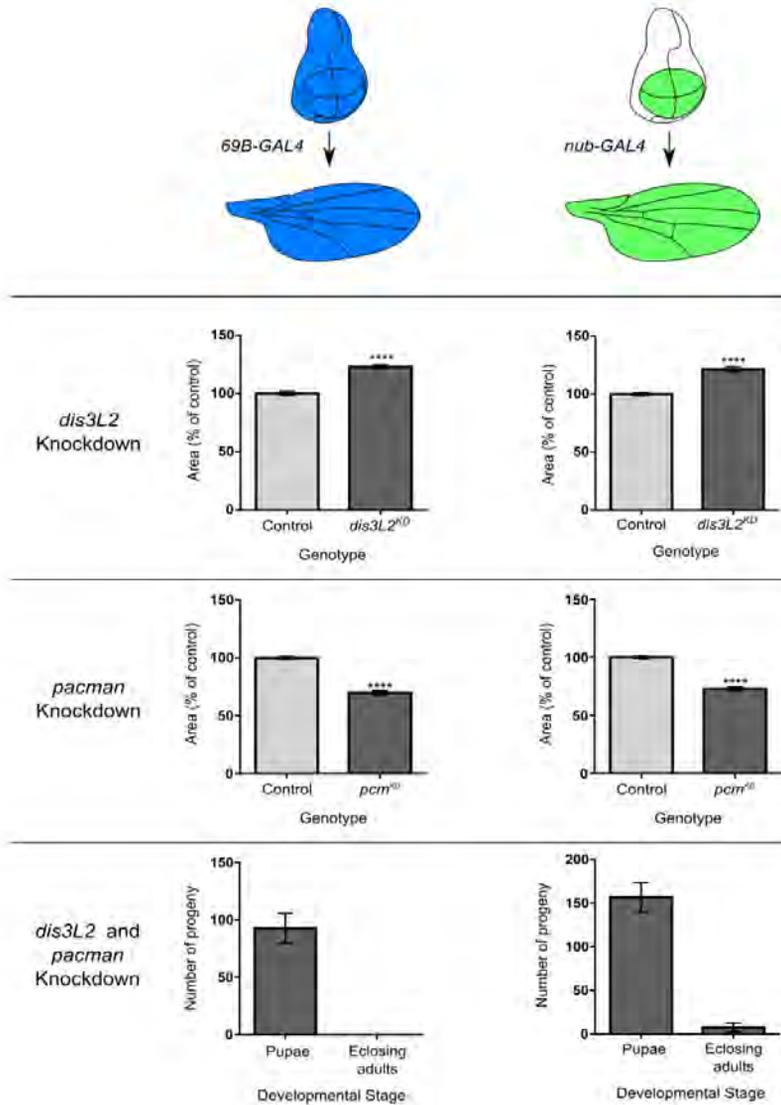
In Chapter 4 of this thesis it was shown that knockdown of *dis3L2* alone with either *69B-GAL4* or *nub-GAL4* results in significant overgrowth of the wing with wings being 20% larger than controls. In contrast loss of *pcm* using RNA interference results in smaller, slightly crumpled wings with wings 72% the size of controls (Waldron 2014). However, knockdown of both *dis3L2* and *pcm* throughout the wing disc using *69B-GAL4* results in complete pupal lethality (Figure 6.2). This therefore again demonstrated that the two pathways are not completely redundant together with indicating the requirement of the pathways for the viability of the organism.

*69B-GAL4* also drives in the epidermis and therefore the pupal lethality could be caused by knockdown of the proteins within regions other than the wing imaginal disc. To determine if the dual knockdown of *dis3L2* and *pcm* within the wing imaginal disc alone is a cause of the pupal lethality the more specific *nub-GAL4* driver was used. Similar to the findings with *69B-GAL4* the knockdown of both *dis3L2* and *pcm* using *nub-GAL4* results in large scale pupal lethality (Figure 6.2) as only 5% of pupae eclosed with these adults presenting with severely crumpled wings. This therefore suggests that the loss of the two proteins within the wing pouch of the disc in the *69B-GAL4* driven cross at least contributes to the pupal lethality. This was slightly unexpected as the loss of developmental control over the wing has not previously led to lethality, for example when *Dis3*, which is essential for viability, was knocked down within the wing pouch (Chapter 3). However it could be a result of stress signals being passed throughout the rest of the disc resulting in the termination of development and subsequent lethality.

These experiments therefore show that the findings in *S. cerevisiae* are consistent with multicellular organisms in that in the absence of either *Dis3L2* or *Pacman* the other is absolutely required to maintain viability.



**Figure 6.1: Simultaneous knockdown of *pacman* (*pcm*) and *dis3L2* using the ubiquitous driver *tub-GAL4*.** (A) Cross scheme to simultaneously knockdown *dis3L2* and *pcm* throughout the organism. Virgin females containing *UAS-RNAi* constructs to both *dis3L2* and *pcm* were crossed to *tub-GAL4/TM6* males. Sibling offspring were identified as larvae/pupae by the presence of the Tubby phenotype whilst adults were identifiable through the presence of the Humeral phenotype. (B) Ubiquitous knockdown of *dis3L2* and *pcm* simultaneously results in lethality prior to the larval stages. Ubiquitous knockdown of *dis3L2* alone (light grey) has no effect on viability, Knockdown of *pcm* alone (grey) results in a slight reduction in the number of progeny reaching the pupal stage. However, knockdown of both simultaneously results in complete lethality (dark grey).  $n \geq 80$  progeny, error bars represent 95% confidence limits. (C) Lethality induced by the simultaneous knockdown of both *pcm* and *dis3L2* is confirmed by the absence of non-Humeral containing adults. Ubiquitous knockdown of *pcm* alone (grey) results in pupal lethality consistent with previous findings (Waldron 2014).  $n \geq 96$  progeny, error bars represent 95% confidence limits.



**Figure 6.2: Simultaneous knockdown of *pacman* (*pcm*) and *dis3L2* in the wing imaginal disc results in additive effects.** Knockdown of *dis3L2* throughout the wing imaginal disc using *69B-GAL4* results in significant overgrowth of the wing. Control genotypes = ;*UAS-dis3L2<sup>RNAi</sup>*, ;*UAS-EGFP<sup>RNAi</sup>/+* ; *69B-GAL4/+* and ;;*69B-GAL4, Dis3L2<sup>KD</sup>* genotype = ;*UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+*.  $n \geq 18$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ . In contrast knockdown of *pcm* in the same region results in a significant reduction in wing area. Control genotype = ;;*UAS-pcm<sup>RNAi</sup>*, *pcm<sup>KD</sup>* genotype = ;;*UAS-pcm<sup>RNAi</sup>/69B-GAL4*.  $n \geq 49$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ . Simultaneous knockdown of *dis3L2* and *pacman* throughout the wing imaginal disc results in 100% pupal lethality. Genotype = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *UAS-pcm<sup>RNAi</sup>/69B-GAL4*.  $n = 555$  pupae over 6 replicate vials, error bars represent 95% confidence limits. Similarly knockdown of *dis3L2* in the wing pouch using *nub-GAL4* results in significant overgrowth of the wing. Control genotypes = ;*UAS-dis3L2<sup>RNAi</sup>*, ;;*UAS-EGFP<sup>RNAi</sup>/nub-GAL4*; and ;*nub-GAL4, Dis3L2<sup>KD</sup>* genotype = ;*UAS-dis3L2<sup>RNAi</sup>/nub-GAL4* ;.  $n \geq 22$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ . In contrast knockdown of *pcm* in the same region results in a significant reduction in wing area. Control genotype = ;;*UAS-pcm<sup>RNAi</sup>*, *pcm<sup>KD</sup>* genotype = ; *nub-GAL4/+* ; *UAS-pcm<sup>RNAi</sup>/+*.  $n \geq 33$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ . Simultaneous knockdown of *dis3L2* and *pacman* throughout the wing imaginal disc results in large scale pupal lethality with only 5% of pupae eclosing. Genotype = ; *UAS-dis3L2<sup>RNAi</sup>/nub-GAL4* ; *UAS-pcm<sup>RNAi</sup>/+*.  $n = 1566$  pupae over 10 replicate vials, error bars represent 95% confidence limits.

## **6.4 *dis3L2* knockdown rescues small wing/wing disc phenotypes of *pcm* mutants.**

Using RNA interference to knockdown both *dis3L2* and *pcm* showed additive effects resulting in lethality indicating that although some redundancy exists the different pathways of RNA decay must be involved in specific processes. However, using RNA interference is never 100% efficient and therefore some background activity would remain. To better examine the pathway interactions within the wing imaginal discs *pcm* mutants were used in collaboration with *dis3L2* targeted RNA interference.

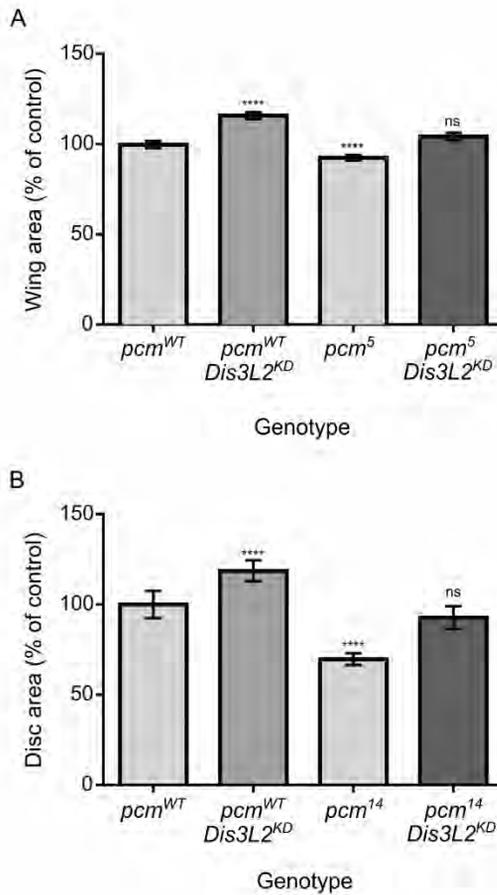
### **6.4.1 *dis3L2* knockdown in a hypomorphic *pcm* mutant background results in phenotypic rescue.**

Previous work in the Newbury lab resulted in the generation of a hypomorphic *pcm* mutant named *pcm*<sup>5</sup> (Grima *et al.* 2008). In this work it was shown that the *pcm*<sup>5</sup> mutant was 66% active and therefore the phenotypes were not as severe as described in the previous section (6.4.1) (Grima *et al.* 2008; Jones *et al.* 2013). The *pcm*<sup>5</sup> mutants presented with smaller wings (80% the size of control wings) that had lost their iridescence (Jones *et al.* 2013). Therefore to investigate if knocking *dis3L2* down in the hypomorphic mutant background resulted in the same phenotypes observed using RNA interference to *pcm*, *dis3L2* was knocked down in a *pcm*<sup>5</sup> mutant.

To ensure that the progeny of interest were homozygous for the *pcm*<sup>5</sup> mutation it was crossed into both the GAL4 and UAS lines balanced over FM7i. FM7i is detected in the adults by the presence of heart/bar shaped eyes whilst it also contains an *actin-GFP* allowing its detection in larvae under a fluorescent light source. *pcm*<sup>5</sup> homozygotes are viable which allowed *pcm*<sup>5</sup>/*Y*; *UAS-dis3L2*<sup>RNAi</sup>; males to be used and crossed to *pcm*<sup>5</sup>/*FM7i*; ; *69B-GAL4* females to produce the desired *pcm*<sup>5</sup>/*Y*; *UAS-dis3L2*<sup>RNAi</sup>/*+*; *69B-GAL4*/*+* (*pcm*<sup>5</sup>*dis3L2*<sup>KD</sup>) males. In addition to these males internal control males were also produced (*FM7i*/*Y*; *UAS-dis3L2*<sup>RNAi</sup>/*+*; *69B-GAL4*/*+*)

which would act as a positive control for the *dis3L2* knockdown induced overgrowth. These males were identifiable from the *pcm*<sup>5</sup>/Y offspring due to being yellow bodied.

To ensure the progeny were all of the same age the freshly eclosing flies were tipped into fresh vials and aged to 1-2 days old at the cross temperature of 25°C. A single wing from each fly was cut and mounted and measured using the wing mounting protocol (methods section 2.2.7). The *FM7i/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* positive control flies presented with wings that were significantly larger (115.7%) than controls indicating the stocks were correct and the overgrowth phenotype caused by *dis3L2* knockdown was reproducible in a different genetic background (Figure 6.3A). Unlike the double knockdown of *pcm* and *dis3L2* in the wing disc *pcm*<sup>5</sup>*dis3L2*<sup>KD</sup> progeny were viable. Strikingly their wings were not significantly different in area to the *UAS-dis3L2<sup>RNAi</sup>* or *69B-GAL4* parental controls (102%) (Figure 6.3A) and therefore were significantly smaller than *dis3L2*<sup>KD</sup> but significantly larger than *pcm*<sup>5</sup> homozygote wings. This indicated that knocking down *dis3L2* in the *pcm*<sup>5</sup> mutants is able to 'rescue' the small wing phenotype to the area of control wings whilst not reaching the 115.7% overgrowth observed in the positive controls. As a result this suggested that Dis3L2 and Pacman do indeed function in separate pathways which is likely to be achieved through their targeting to specific RNAs.



**Figure 6.3: Knockdown of *dis3L2* in a *pcm* mutant background results in phenotypic rescue. (A)** Knockdown of *dis3L2* in the hypomorphic *pcm*<sup>5</sup> mutant background results in a compensatory phenotype where wing area returns to wild type. *pcm*<sup>WT</sup> = ;*UAS-dis3L2*<sup>RNAi</sup>; and ;;69B-GAL4, *pcm*<sup>WT</sup> *Dis3L2*<sup>KD</sup> = ;*UAS-dis3L2*<sup>RNAi</sup>/+ ; 69B-GAL4/+ , *pcm*<sup>5</sup> = *pcm*<sup>5</sup>; , *pcm*<sup>5</sup> *Dis3L2*<sup>KD</sup> = *pcm*<sup>5</sup>; *UAS-dis3L2*<sup>RNAi</sup>/+ ; 69B-GAL4/+. n≥15, error bars represent 95% confidence limits, \*\*\*\*=p<0.0001. **(B)** Knockdown of *dis3L2* in the null *pcm*<sup>14</sup> mutant background results in a compensatory phenotype where wing imaginal disc area returns to wild type. *pcm*<sup>WT</sup> = 50E ; *UAS-dis3L2*<sup>RNAi</sup>; , *pcm*<sup>WT</sup> *Dis3L2*<sup>KD</sup> = 50E ; *UAS-dis3L2*<sup>RNAi</sup>/+ ; 69B-GAL4/+ , *pcm*<sup>14</sup> = *pcm*<sup>14</sup>; *UAS-dis3L2*<sup>RNAi</sup>; , *pcm*<sup>14</sup> *Dis3L2*<sup>KD</sup> = *pcm*<sup>14</sup>; *UAS-dis3L2*<sup>RNAi</sup>/+ ; 69B-GAL4/+. n≥15, error bars represent 95% confidence limits, \*\*\*\*=p<0.0001.

## 6.4.2 *dis3L2* knockdown in a null *pcm* mutant background results in partial phenotypic rescue.

Along with the hypomorphic *pcm*<sup>5</sup> mutant a null mutant (*pcm*<sup>14</sup>) was also made by Dr Chris Jones and was published in Waldron *et al* 2015 (Waldron *et al.* 2015). However, unlike the *pcm*<sup>5</sup> mutant, *pcm*<sup>14</sup> is a null mutation which results in pupal lethality. As a result the wings are unable to be used to investigate an interaction; however, due to their survival to the larval stage the wing imaginal discs were used. As presented in Waldron *et al* the *pcm*<sup>14</sup> wing imaginal discs are 45% the size of control discs at the same developmental stage. The *pcm*<sup>14</sup> larvae are also significantly delayed in their development which allows the wing discs time to increase in size to ~75% the size of control discs. In a similar experiment to that outlined in 6.4.1 *dis3L2* was knocked down using RNA interference in *pcm*<sup>14</sup> mutants. The hypothesis being that unlike in 6.4.1 the phenotypes would mimic those observed when RNAi was used to knockdown both *pcm* and *dis3L2* simultaneously.

Unlike the *pcm*<sup>5</sup> mutants *pcm*<sup>14</sup> homozygotes are not viable as adults therefore heterozygotes had to be used in the crosses. The *pcm*<sup>14</sup> mutation was crossed into the *UAS-dis3L2*<sup>RNAi</sup> again balanced over FM7i. *pcm*<sup>14</sup>/FM7i ; *UAS-dis3L2*<sup>RNAi</sup> virgin females were crossed to *69B-GAL4* males which produced the desired *pcm*<sup>14</sup>/Y ; *UAS-dis3L2*<sup>RNAi</sup>/+ ; *69B-GAL4*/+ male larvae (Figure 6.4A). These larvae were identified from the *FM7i*/Y ; *UAS-dis3L2*<sup>RNAi</sup>/+ ; *69B-GAL4*/+ larvae by the use of the *actin-GFP* from the *FM7i* chromosome which was visible under a fluorescent light source. It was essential that only male larvae were dissected for wing disc measurement as the non-GFP female larvae were only heterozygous for *pcm*<sup>14</sup> (*w*<sup>\*</sup>/*pcm*<sup>14</sup>). This was achieved through selection on the presence of the larval testes which appear as 'bubble-like' structures in the rear third of the wandering L3 larvae.

In addition to the *pcm*<sup>14</sup> mutant the isogenic control chromosome *50E* was crossed into the *UAS-dis3L2*<sup>RNAi</sup> stock to serve as the most relevant control for the *pcm*<sup>14</sup> mutation (Figure 6.4B). The *50E* chromosome was produced by the clean excision of the P-element which was used in

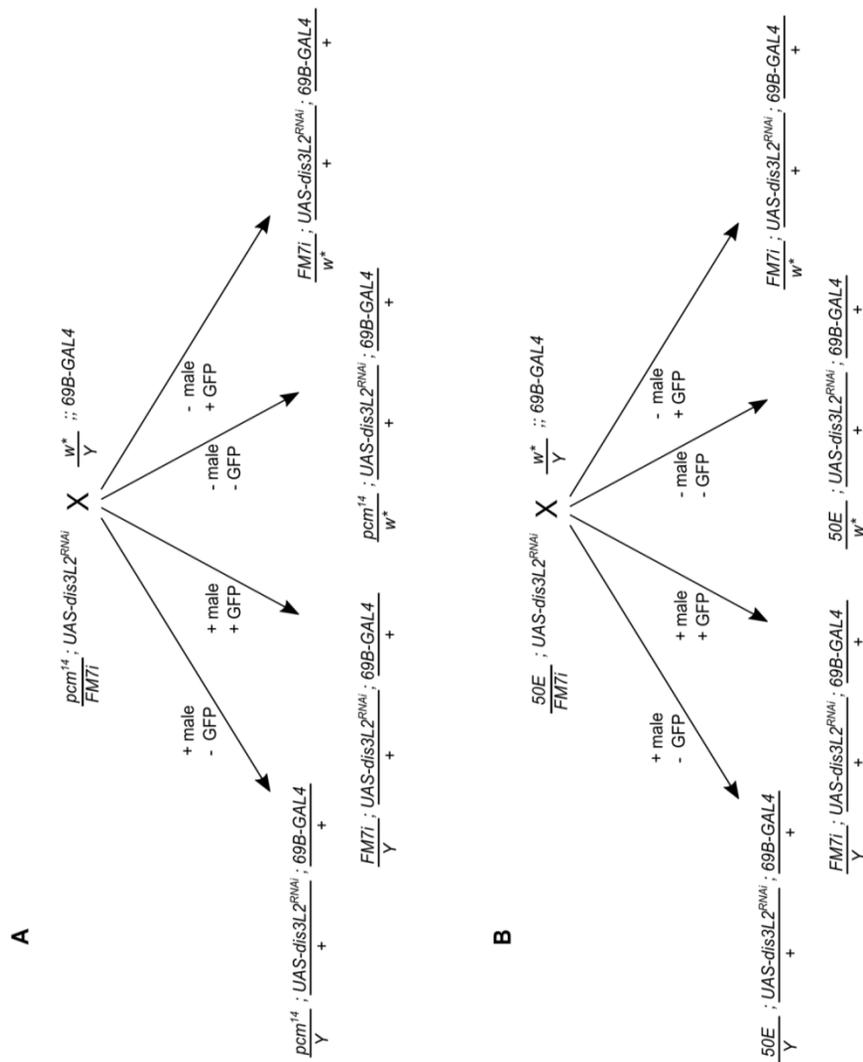
the creation of the *pcm<sup>14</sup>* mutation. Like in the *pcm<sup>14</sup>* crosses *50E/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* male larvae were selected through the absence of GFP and the presence of the larval testes.

Late wandering L3 larvae were specifically selected for dissection through the use of the bromophenol blue food. The blue dye within the fly food is visible in the larval gut, however, in the period immediately prior to pupation (late wandering L3) the larva empties its gut and therefore the blue food is no longer visible. Only those larvae showing an absence of blue food in their gut were selected for dissection as this ensured the measured discs are all from larvae of the same developmental stage. Late wandering L3 wing imaginal discs were dissected, mounted and measured as to the disc dissection protocol (methods section 2.2.6).

To check that the stocks were correct the late L3 wing discs were dissected from *pcm<sup>14</sup>/Y ; UAS-dis3L2<sup>RNAi</sup>* ; larvae as a positive control. These discs were 74.7% the size of *50E/Y ; UAS-dis3L2<sup>RNAi</sup>* ; wing discs which were the most suitable control for the whole experiment (Figure 6.3B). This was consistent with previous findings that *pcm<sup>14</sup>* late L3 wing discs are 75% the size of control discs (Grima *et al.* 2008). *50E/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* wing imaginal discs were significantly larger being 118.6% the size of control discs (Figure 6.3B) which again was consistent with previous findings outlined in Chapter 4 where *dis3L2* knockdown results in significantly larger wing discs. Interestingly the same phenotypic rescue was observed for *pcm<sup>14</sup>/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* wing imaginal discs as was seen in the *pcm<sup>5</sup>/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* adult wings. Late L3 *pcm<sup>14</sup>/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* wing imaginal discs were not significantly different in area to the *50E/Y ; UAS-dis3L2<sup>RNAi</sup>* control discs being 92.7% the area of these discs (Figure 6.3B). Therefore, this indicates that knocking down *dis3L2* in either a null (*pcm<sup>14</sup>*) or hypomorphic (*pcm<sup>5</sup>*) mutant background is able to rescue the area of the discs or wings respectively. The likely mechanism behind the rescue is that the increased proliferation induced by *dis3L2*-depletion in co-operation with the

compensatory proliferation in the *pcm*<sup>14</sup> discs is sufficient to replace the cells lost by *pcm*<sup>14</sup> induced apoptosis.

The *dis3L2* knockdown is not able, however, to rescue the pupal lethality seen in *pcm*<sup>14</sup> mutants, although this is not surprising given that the knockdown of both *dis3L2* and *pcm* in the wing disc results in pupal lethality (section 6.4.2). Unlike the *pcm* mutations, knockdown of *pcm* within the wing imaginal disc does not have an effect on the area of the discs (Waldron 2014) and therefore looking at the discs in the simultaneous knockdown larvae was not worthwhile. Taken together these data therefore provide further evidence that the 3'-5' and 5'-3' pathways act on specific targets to regulate opposing developmental pathways.

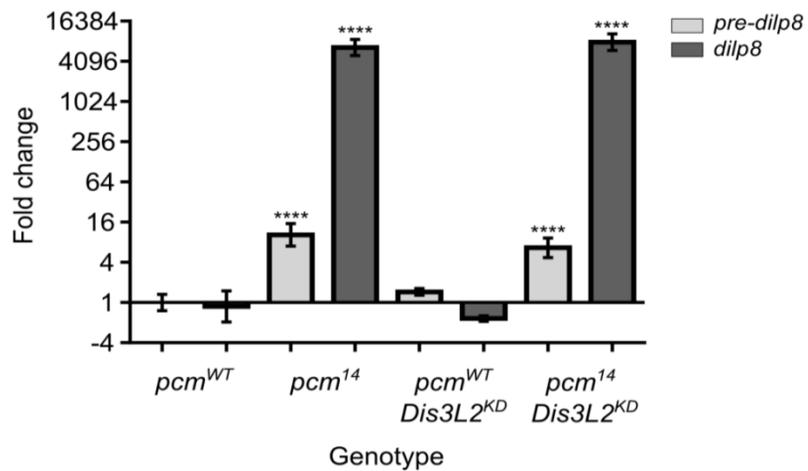


**Figure 6.4: Cross scheme to achieve the knockdown of *dis3L2* in a *pacman* null mutant background. (A) *pcm<sup>14</sup>* heterozygous virgin females containing the *UAS-dis3L2<sup>RNAi</sup>* insertion were crossed to *69B-GAL4* males resulting in offspring of 4 potential genotypes. Only *pcm<sup>14</sup>* hemizygous male larvae were suitable which were selected for by the absence of GFP (from FM71) and the presence of the larval testis. (B) The same cross was performed but using the isogenic control for *pcm<sup>14</sup>*, *50E*. The same selection criteria were used as in A.**

### 6.4.3 *dis3L2* knockdown does not rescue the *pcm<sup>14</sup>* induced developmental delay.

The previous section has shown how the loss *dis3L2* is able to rescue the small wing phenotype of *pcm<sup>14</sup>* mutants. This could be due to the *dis3L2*-depletion induced proliferation being sufficient to compensate for the loss of cells in the *pcm<sup>14</sup>* induced apoptosis. However, knockdown of *dis3L2* is unable to rescue other *pcm<sup>14</sup>* phenotypes. For example, the *pcm<sup>14</sup>* larvae are delayed in their development by 32 hours which is most likely to give the imaginal discs more time to reach a suitable size to enter pupation. The *pcm<sup>14</sup>/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* larvae show a similar developmental delay indicating that the delay initially observed is not simply due to the smaller wing imaginal discs.

The loss of Pacman in the wing imaginal discs has been shown to result in an increase in the expression of *insulin-like peptide 8 (dilp8)* RNA and protein (Jones *et al.* 2016). Ectopic expression of Dilp8 has previously been shown to result in a developmental delay (Colombani *et al.* 2012) and therefore to see if the levels remained elevated in the *pcm<sup>14</sup>/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* wing imaginal discs qRT-PCR was used to assess the levels of *dilp8*. RNA was extracted from all 4 control genotypes used in section 6.4.2 and reverse transcribed using random primers. qRT-PCR showed that an approximate 6600 fold increase in *dilp8* expression was observed in the *pcm<sup>14</sup>/Y ; UAS-dis3L2<sup>RNAi</sup> ; positive control discs* indicating that the increase observed in Jones *et al* is reproducible (Figure 6.5). This was not significantly different from the approximate 8000 fold increase in expression in the *pcm<sup>14</sup>/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* wing imaginal discs (Figure 6.5) showing that the loss of *dis3L2* has no effect on the expression of *dilp8* in the *pcm<sup>14</sup>* mutant tissue. This therefore explains why, although the area of the wing discs are rescued, the *pcm<sup>14</sup>/Y ; UAS-dis3L2<sup>RNAi</sup>/+ ; 69B-GAL4/+* larvae are still delayed in their development.



**Figure 6.5: Knockdown of *dis3L2* does not affect the *pcm*<sup>14</sup> induced upregulation of *dilp8* and therefore does not rescue the developmental delay.** Knockdown of *dis3L2* in *pcm*<sup>14</sup> mutant wing discs (*pcm*<sup>14</sup> *Dis3L2*<sup>KD</sup>) has no effect on the post-transcriptional upregulation of *dilp8*. *pre-dilp8* (light grey) used as a transcriptional output. *pcm*<sup>WT</sup> = 50E ; *UAS-dis3L2*<sup>RNAi</sup>;; *pcm*<sup>WT</sup> *Dis3L2*<sup>KD</sup> = 50E ; *UAS-dis3L2*<sup>RNAi</sup>/+ ; 69B-GAL4/+ , *pcm*<sup>14</sup> = *pcm*<sup>14</sup>; *UAS-dis3L2*<sup>RNAi</sup>;; *pcm*<sup>14</sup> *Dis3L2*<sup>KD</sup> = *pcm*<sup>14</sup>; *UAS-dis3L2*<sup>RNAi</sup>/+ ; 69B-GAL4/+. n≥3, error bars represent SEM, \*\*\*\*=p<0.0001.

## 6.5 Dis3L2 does not function through the Hid, Grim and Reaper mediated apoptotic pathway to regulate disc growth.

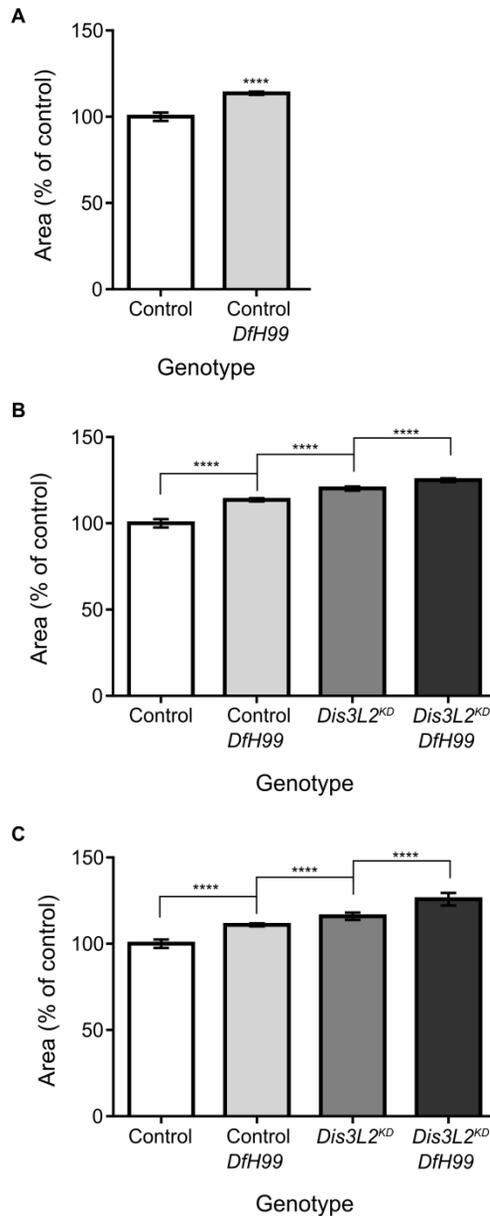
The previous work in Chapter 4 indicates a pro-proliferation environment when *dis3L2* is knocked down in the wing imaginal disc. This would also be consistent with the previous work in this chapter (section 6.4) showing that the loss of *dis3L2* is able to rescue the growth deficiency induced by a mutation in *pacman*. However, it is possible that in addition to stimulating an increase in proliferation the loss of *dis3L2* could also result in less apoptosis within the wing imaginal disc. A link to apoptosis has recently been reported by Thomas *et al* where it was shown that Dis3L2 plays a role in the clearance of mRNAs following pro-apoptotic stimuli (Thomas *et al.* 2015).

A classical genetic approach was taken to investigate if the loss of *dis3L2* results in an alteration in the balance of apoptosis. Apoptosis in *Drosophila* is mediated by 3 key proteins; Hid, Grim and Reaper which function to block the activity of DIAP1 (inhibitor of apoptosis protein) to ultimately promote the activity of the apoptosome. A single deletion (*Df(3L)H99*) removes all three genes encoding the pro-apoptotic proteins and can therefore be used in apoptotic studies (Abbott and Lengyel 1991; Waldron *et al.* 2015). *Df(3L)H99* homozygotes are not viable and therefore *Df(3L)H99* heterozygotes need to be used. As a heterozygote the *Df(3L)H99* deletion still had an apparent effect on apoptosis as the wings of ; *UAS-dis3L2<sup>RNAi</sup>* ; *Df(3L)H99/TM3* flies are significantly larger than *UAS-dis3L2<sup>RNAi</sup>* wings which also indicates that a loss of apoptotic control can stimulate wing overgrowth (Figure 6.6A). ; *UAS-dis3L2<sup>RNAi</sup>* ; *Df(3L)H99/TM3* males were crossed to *69B-GAL4* or *nub-GAL4* virgins to produce progeny with *dis3L2* knockdown in an anti-apoptotic background. If the wing overgrowth observed when *dis3L2* is knocked down was due to a decrease in apoptosis then the *Df(3L)H99* deletion would not be expected to have an additional effect. However, if the overgrowth was indeed due to an increase in proliferation (as shown in Chapter 4) then an additive effect might be expected

where the reduction in apoptosis induced by the *Df(3L)H99* deletion further enhances the pro-proliferative environment induced by *dis3L2*-depletion.

Knockdown of *dis3L2* in the anti-apoptotic background results in wings that are significantly larger than when *dis3L2* is depleted alone (Figure 6.6B/C). As mentioned previously, the *Df(3L)H99* deletion results in a slight overgrowth of the wing and therefore the ; *UAS-dis3L2<sup>RNAi</sup>* ; *Df(3L)H99/TM3* parental flies were used as controls. The siblings ( ; *UAS-dis3L2<sup>RNAi</sup>* ; *69B-GAL4/TM3* or ; *UAS-dis3L2<sup>RNAi</sup>/nub-GAL4 ; TM3/+*) acted as positive controls, and as expected showed significant overgrowth ( $p < 0.0001$ ) of the wing when compared to the parental stock containing the *Df(3L)H99/+* deletion (wings were 6%/5% larger *69B* and *nub* driven knockdowns respectively). However, the progeny containing *dis3L2* knockdown in the anti-apoptotic background presented with wings that were significantly larger than both the parental stock and the sibling positive controls (111% increase in wing area compared to parental control,  $p < 0.0001$  for both comparisons). The additional overgrowth therefore indicates that the *dis3L2* knockdown induced overgrowth is not due to deregulation of the Hid, Grim and Reaper apoptotic pathway. These data show that Dis3L2 functions through a pathway distinct to that of Pacman. This could be the cause of the phenotypic rescue such that the increased apoptosis induced by the loss of Pacman function is compensated for by the increased proliferation that occurs when *dis3L2* is depleted.

These data, therefore, show for the first time that two exoribonucleases function through discrete pathways to regulate cell growth and tissue development



**Figure 6.6: *dis3L2* knockdown induced overgrowth is unlikely to be caused by the inhibition of apoptosis. (A)** The wing area of controls (*;UAS-dis3L2<sup>RNAi</sup>;*) are significantly smaller than the same control wings with the *Df(3L)H99* deletion (*; UAS-dis3L2<sup>RNAi</sup> ; Df(3L)H99/+*).  $n \geq 30$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ . **(B-C)** *dis3L2* knockdown (*Dis3L2<sup>KD</sup>*) has a larger effect on wing overgrowth than *Df(3L)H99* as a heterozygote (Control *DfH99*). However, knocking down *dis3L2* in a heterozygous *DfH99* background with either *69B-GAL4* **(B)** or *nub-GAL4* **(C)** (*Dis3L2<sup>KD</sup> DfH99*) results in an additive effect where the wings are significantly larger than *dis3L2* knockdown wings alone.  $n \geq 18$ , error bars represent 95% confidence limits, \*\*\*\*= $p < 0.0001$ .

## 6.6 Dis3L2 and Pacman target specific RNAs to regulate their opposing pathways.

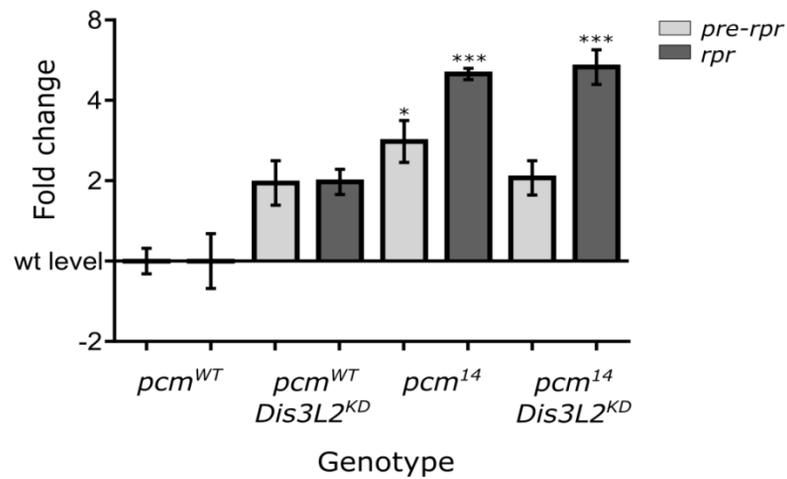
The previous sections have outlined how Dis3L2 and Pacman appear to function to regulate opposing pathways to control cell growth and tissue development. If this is indeed the case then one would hypothesise that Dis3L2 and Pacman target specific RNAs which are involved in the distinct pathways.

### 6.6.1 Dis3L2 knockdown in a *pcm<sup>14</sup>* mutant background does not affect *rpr* expression.

The work in section 6.5 suggests that Dis3L2 plays no role in the Hid, Grim or Reaper pathway of apoptotic regulation whilst Pacman has been previously shown to directly regulate their expression (Waldron *et al.* 2015). To show that Dis3L2 does not affect the levels of *rpr* expression qRT-PCR was again used. The level of *rpr* expression was assessed in wing imaginal discs with *dis3L2* knockdown alone, *pcm<sup>14</sup>* alone and *dis3L2* knockdown with the *pcm<sup>14</sup>* mutation. RNA was produced from L3 wing imaginal discs dissected from each of the required genotypes. The RNA was then reverse transcribed using random primers before being subject to qRT-PCR. As was observed in previous work (Waldron *et al.* 2015) *rpr* expression was increased by 5 fold at the RNA level in *pcm<sup>14</sup>; UAS-dis3L2<sup>RNAi</sup>*; wing imaginal discs when compared to control discs (*50E; UAS-dis3L2<sup>RNAi</sup>*;). This was a largely post-transcriptional effect as the levels of the pre-mRNA (*pre-rpr*) only increased 2 fold when compared to the control discs (Figure 6.7). This showed that the same conditions were present in these experiments as were in previous work in the lab.

Although *rpr* expression was seen to increase by 2 fold in wing discs where *dis3L2* is knocked down alone (*50E; UAS-dis3L2<sup>RNAi</sup>/+; 69B-GAL4/+*), this was not statistically significant. A similar increase which again was not significant was observed for *pre-rpr* (Figure 6.7). These data were confirmed by the RNA-seq data (Chapter 5) as *rpr* expression also remains unchanged in the *dis3L2* knockdown wing imaginal discs used. Finally, the *pcm<sup>14</sup>* induced increase in reaper

expression was not affected by the knockdown of *dis3L2* in *pcm<sup>14</sup>* mutant discs (*pcm<sup>14</sup>/Y*; *UAS-dis3L2<sup>RNAi</sup>/+*; *69B-GAL4/+*) (Figure 6.7). Levels of *rpr* mRNA increase by 5 fold and 5.4 fold when *pcm<sup>14</sup>* is present alone or alongside *ds3L2* knockdown respectively; these increases are not significantly different from one another. The levels of *pre-rpr* also remain constant between the two conditions therefore indicating Dis3L2 does not play a role in regulating *rpr* expression.



**Figure 6.7: Knockdown of *dis3L2* in *pcm*<sup>14</sup> mutant wing discs (*pcm*<sup>14</sup> *Dis3L2*<sup>KD</sup>) has no effect on the post-transcriptional upregulation of *rpr*.** *pre-rpr* (light grey) used as a transcriptional output. *pcm*<sup>WT</sup>= 50E ; *UAS-dis3L2*<sup>RNAi</sup>;; *pcm*<sup>WT</sup> *Dis3L2*<sup>KD</sup> = 50E ; *UAS-dis3L2*<sup>RNAi</sup>/+ ; 69B-GAL4/+ , *pcm*<sup>14</sup> = *pcm*<sup>14</sup>; *UAS-dis3L2*<sup>RNAi</sup>;; *pcm*<sup>14</sup> *Dis3L2*<sup>KD</sup>= *pcm*<sup>14</sup>; *UAS-dis3L2*<sup>RNAi</sup>/+ ; 69B-GAL4/+. n≥3, error bars represent SEM, \*\*\*=p<0.001, \*=p<0.05.

## 6.6.2 Comparative analysis of RNA sequencing data reveals transcripts showing specific sensitivity to Pacman and Dis3L2.

Having determined that, unlike Pacman, Dis3L2 does not have a function in regulating the expression levels of *rpr* mRNA it was of interest to explore specific targets. As outlined in Chapter 5 RNA sequencing has been performed on *dis3L2* knockdown wing imaginal discs to identify potential targets of Dis3L2. Additionally, the lab has previously carried out RNA sequencing on *pcm*<sup>14</sup> mutant discs (Jones *et al.* 2016) which provided an opportunity for comparative analysis of the data sets. As mentioned in previous sections, 2 Pacman targets, *dilp8* and *rpr*, are unaffected in *dis3L2* knockdown discs showing pathway specificity.

To assess if those 233 mRNAs that were affected by *dis3L2* knockdown were specific to the *dis3L2* condition or if they were similarly affected by *pacman* mutation the expression changes were compared between the two data sets. Similarly to the Dis3L2 sequencing analysis a fold change cut off of >1.5 was used. This revealed three categories; those mRNAs that change specifically when *dis3L2* is knocked down (Dis3L2 specific), those that change in opposing directions (opposite) or those that change in the same manner when either *dis3L2* is knocked down or when *pcm* is mutant (Same)(Figure 6.8). Of the 233 analysed mRNAs 20.6% (48 mRNAs) changed in opposing directions, 45.9% (107 mRNAs) showed the same direction of change in expression and 33.5%(78 mRNAs) only showed expression changes when *dis3L2* was knocked down. Figure 6.9 shows the breakdown of the number of up and downregulated genes in each category in respect to *dis3L2* knockdown. From this it is clear that the majority (79.2%) of those mRNAs that change in opposing directions are downregulated in the *dis3L2* knockdown samples. In contrast, those mRNAs that show an expression change only in the *dis3L2* knockdown discs show both upregulation (51.3%) and downregulation (48.7%).

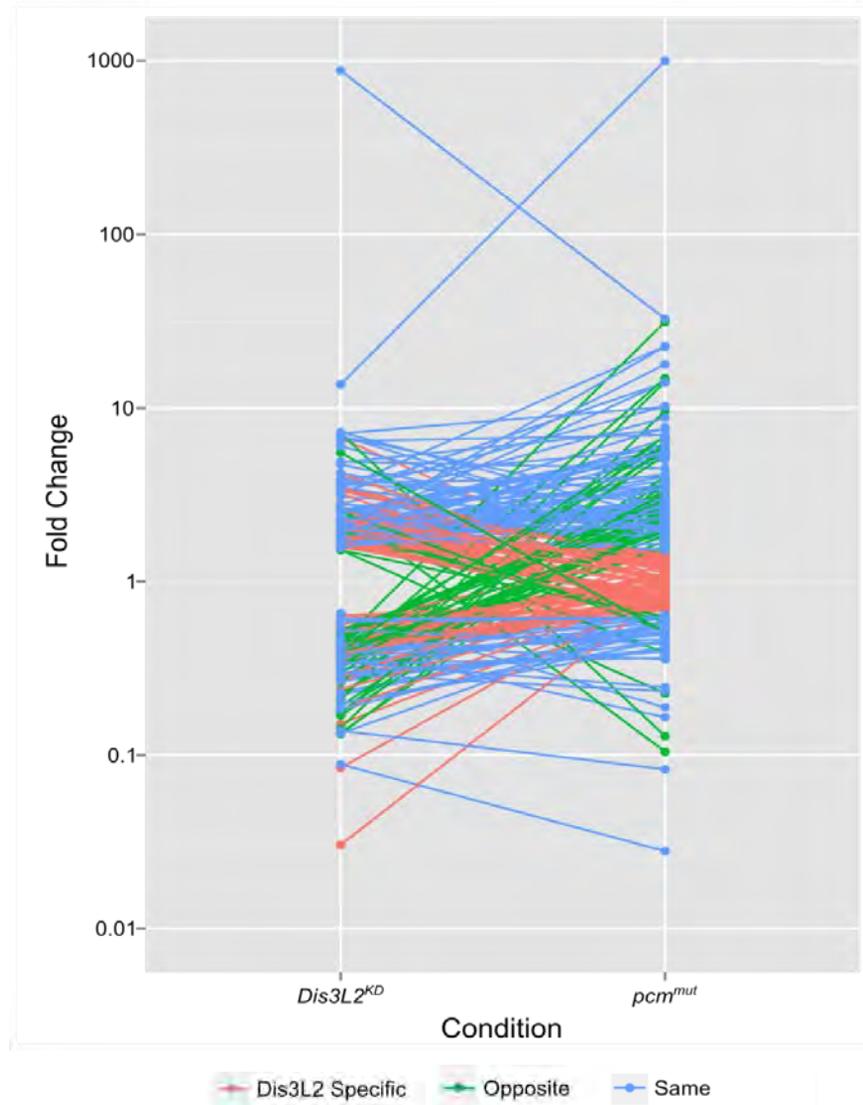
To identify if the transcripts in each category were involved in any particular pathways Gene Ontology (GO) analysis, using the online DAVID tool, was performed. The number of input genes into the GO was low for each category and therefore the pathway analysis is limited but

it still revealed a few pathways. The enrichment score is calculated by the  $-\log_{10}(\text{p-value})$  with the p-value itself being derived from a comparison between the proportion of transcripts within that category and the expected proportion in a background genome. The transcripts within the 'opposite', 'same' and 'dis3L2 specific' groups were submitted separately and analysed with the default settings including moderate stringency. The background to which the lists were compared was the inbuilt background for the *Drosophila* genome.

GO analysis on those genes showing opposite trends of expression changes between the two data sets revealed a strong enrichment of transcripts classified as peptidases. 16.4% of the transcripts in the submitted list fell within this category representing a 6.6 fold enrichment. This was the only significantly enriched category within the 'opposite' data set. No enrichment of specific pathways/functions was observed in the 'same' group of transcripts suggesting they are a diverse group of mRNAs. Finally GO analysis on those transcripts that only change when *dis3L2* is knocked down showed the enrichment of two categories. Transcripts involved in 'metal ion binding' were enriched with a score of 1.54 representing a 2.35 fold enrichment over the expected proportion. These 'metal ion binding' transcripts are likely to encode enzymes. The second category contained transcripts involved in oxidation-reduction processes which were 6.12 fold enriched (score of 1.47). The GO analysis (summarised in Figure 6.10) suggest that those transcripts differentially expressed between the two data sets were involved in distinct processes and therefore provided further evidence towards pathway specificity.

The mRNAs selected for further analysis in Chapter 5 were assessed to see into which category they fell. If the changes in *CG2678* and *pyx* were induced by a lack of Dis3L2 activity upon them they would be expected to fall into the Dis3L2 specific category. *CG2678* was indeed only increased in expression in the Dis3L2 sample as a non-significant 1.3 fold increase was observed in the *pcm*<sup>14</sup> mutant data. Despite *pyx* falling into the 'Same' category with a 1.9 fold increase in the *pcm*<sup>14</sup> mutant data this increase was not significant and therefore it again looks

to be specific to the loss of Dis3L2. Alternatively, the two mRNAs that were shown to increase at the transcriptional level show increases, although of a lower magnitude than in *dis3L2* knockdown discs, in the *pcm<sup>14</sup>* mutant data. *CG31808* and *cyt-c-d* show fold increase of 2.79 and 2.32 in *pcm<sup>14</sup>* mutant discs. This therefore shows that the loss of *dis3L2* or *pacman* can affect either similar or specific transcripts. This is in keeping with the hypothesis that although partial redundancy exists, the two exoribonucleases are able to target specific substrates to elicit their control over wing disc growth and development. Figure 6.11 summarises the difference in fold changes between the validated genes from the *dis3L2* sequencing data set and the *pcm<sup>14</sup>* data.



**Figure 6.8: 3 classes of transcripts are identified following comparative analysis of RNA-seq data.** Some transcripts only show fold changes in *dis3L2<sup>KD</sup>* wing imaginal discs (Dis3L2 specific, shown in red). Other transcripts change in opposing directions in *dis3L2<sup>KD</sup>* and *pcm* null mutant (*pcm<sup>mut</sup>*) wing imaginal discs (Opposite, shown in green). Finally a set of transcripts show fold changes in the same direction but often of different magnitude in *dis3L2<sup>KD</sup>* and *pcm* null mutant wing imaginal discs (Same, shown in blue). Fold changes presented for *Dis3L2<sup>KD</sup>* samples is the change between the grouped parental control (; *UAS-dis3L2<sup>RNAi</sup>* ; and ;; *69B-GAL4*) and knockdown discs (; *UAS-dis3L2<sup>RNAi/+</sup>* ; *69B-GAL4/+*). Fold changes presented for *pcm<sup>mut</sup>* samples is the change between the two null *pacman* mutants (*pcm<sup>14</sup>*;; and *pcm<sup>15</sup>*;;) and the grouped control containing their respective isogenic control lines (*50E*;; and *166V*;;). *pcm* RNA-seq and analysis was performed by Dr. Chris Jones.

	Opposite	Same	Dis3L2 specific
Total	48	107	78
% total	20.60	45.92	33.48
Upregulated	10	74	40
Downregulated	38	33	38
% upregulated	20.83	69.16	51.28
% downregulated	79.17	30.84	48.72
Av fold change Dis3L2	3.34	3.13	2.99
Av fold change Pcm	4.43	4.34	1.25

**Figure 6.9: Summary of the number of transcripts in each of the 3 classes presented in Figure 6.8 and their direction of change.** The transcripts in each class are presented in Appendix.

Category (Opposite)	Enrichment score (-log <sub>10</sub> (p-value))	Gene count (% of input)	Fold enrichment
Peptidase	2.89	7 (16.38)	6.60

Category (Dis3L2 Specific)	Enrichment score (-log <sub>10</sub> (p-value))	Gene count (% of input)	Fold enrichment
Metal ion binding	1.53	10 (16.32)	2.35
Oxidation reduction	1.47	5 (7.94)	6.11

**Figure 6.10: Gene ontology (GO) analysis of the transcripts identified in each of the 3 classes presented in Figure 6.8.** GO analysis identified one enriched function of transcripts that change in opposing directions in *dis3L2<sup>KD</sup>* and *pcm<sup>mut</sup>* wing imaginal discs. Whilst two functional categories were identified for those transcripts that change specifically in *dis3L2<sup>KD</sup>* wing imaginal discs. No enrichment was observed for those showing the same direction of expression change between the two data sets.

	Dis3L2 <sup>KD</sup>		pcm <sup>Mut</sup>	
	Fold Change	Significant? (qPCR)	Fold Change	Significant? (seq)
mRNA				
<i>cyt-c-d</i>	7.24	Y	2.32	N
CG31808	6.84	Y	2.79	Y
<i>pyx</i>	4.2	Y	1.9	N
<i>fs(1)Yb</i>	4.84	Y	7.54	Y
CG2678	3.51	Y	1.3	N
<i>gd</i>	1.97	Y	22.54	Y

**Figure 6.11: Comparison of the fold changes between Dis3L2 and Pacman sequencing data for those transcripts selected for validation in Chapter 5.**

## 6.7 Chapter Summary

The work in this chapter provides evidence towards the specificity of the cytoplasmic 5'-3' and 3'-5' pathways of degradation. It has been shown that the knockdown of both *dis3L2* and *pacman* using RNA interference results in increased phenotypic severity compared to the situation when either is knocked down alone. This showed that, in the absence of one, the other is required to maintain viability together with showing the cytoplasmic exosome is unable to compensate for the loss of Dis3L2 activity. Interestingly, the use of both hypomorphic and null *pacman* mutants revealed a compensatory effect where reduced growth in the *pacman* mutants is rescued by the knockdown of *dis3L2* in the wing imaginal disc. However, the developmental delay and eventual pupal lethality caused by the *pcm<sup>14</sup>* null mutation are not rescued. Therefore, this suggests that although the increased proliferation induced by *dis3L2*-depletion is sufficient to rescue the growth phenotype Dis3L2 acts in a different pathway and is therefore unable to compensate for the loss of Pacman.

In addition to the phenotypic analysis further evidence has been provided to support the hypothesis of Dis3L2 and Pacman targeting specific pathways. For example, two Pacman targets, *rpr* and *dilp8* (Waldron *et al.* 2015; Jones *et al.* 2016) are unaffected by Dis3L2 activity as their levels remain constant between *pcm<sup>14</sup>* mutant discs and *pcm<sup>14</sup> dis3L2* knockdown discs. Similarly the post-transcriptional increase of *CG2678* and *pyx* observed following the loss of Dis3L2 activity is specific as neither change significantly in the *pacman* mutant sequencing data. Consistent with this, comparative analysis of the two sequencing data sets reveals a set of transcripts which show an increase in levels specifically when *dis3L2* is knocked down together with a group showing opposing expression changes when either *pacman* is mutant or *dis3L2* is depleted. Taken together this chapter therefore shows how certain transcripts appear to be specifically targeted for degradation by either Pacman or Dis3L2; it is likely this specific targeting allows the regulatory function of Pacman and Dis3L2 on apoptosis and cell

proliferation respectively. Further work is required to elucidate the mechanisms behind the specific targeting and a wider set of Dis3L2 and Pacman specific targets.

## 6.8 Chapter Discussion

### 6.8.1 Why are there partial differences in the phenotypes observed between double knockdown and knockdown in mutant backgrounds?

When *pacman* and *dis3L2* are simultaneously knocked down a substantial increase in phenotypic severity was observed where even restriction of the knockdown specifically to the wing pouch resulted in large scale pupal lethality. This finding demonstrated that when Pacman activity is inhibited the activity of Dis3L2 is required to maintain organism viability. This is consistent with previous findings in *S. pombe* (Malecki *et al.* 2013) where cells mutant for both *xrn1* and *dis3L2* are not viable. It further shows that although some redundancy does exist between the 5'-3' and 3'-5' pathways the exosome is unable to compensate for the loss of Dis3L2-mediated 3'-5' decay.

The knockdown of both *pacman* and *dis3L2* throughout the wing imaginal disc results in complete pupal lethality; however, this lethality may have also been induced by knockdown elsewhere as the *69B-GAL4* driver also expresses in the epidermis. Unexpectedly, knockdown of *dis3L2* and *pacman* in the wing pouch of the wing disc also resulted in pupal lethality. The expression domain of *nubbin-GAL4* is much more specific than the *69B-GAL4* driver and therefore this suggested that at least some of the pupal lethality observed when *69B-GAL4* was used was due to the loss of *dis3L2* and *pacman* in the wing disc. The cells affected by the *nub-GAL4* driven knockdown of *dis3L2* and *pacman* were solely those that were destined to form the wing blade and therefore it was surprising that the loss of developmental control over these cells specifically induced lethality. The observed lethality may have been as a result of a lack of signalling between the wing pouch cells and the rest of the wing disc which may ultimately affect signalling between the different developmental tissues within the larvae. If

this was indeed the case then it may affect the development of the tip of the disc, destined to form part of the thorax. Deficient development of the thorax could result in pupal lethality.

Using a mutant is preferable over using RNA interference as the knockdown achieved is never 100% effective and therefore some wild type activity will remain. Therefore hypomorphic and null *pacman* mutants were used to see if the same phenotypes were observed. To achieve this *dis3L2* was knocked down in both the hypomorphic *pcm<sup>5</sup>* and the null *pcm<sup>14</sup>* mutant backgrounds. *pcm<sup>5</sup>* is 66% functional and therefore Pacman should have been more functional in these tissues than in the previous RNAi tissues. However, *pcm<sup>14</sup>* null tissues have no Pacman activity and therefore the loss of function would be more severe than in the RNAi tissues. However, interestingly the same phenotypic rescue was observed following *dis3L2*-depletion in either the hypomorphic or null mutant backgrounds.

Unlike the simultaneous RNAi flies the *pcm<sup>5</sup> dis3L2* knockdown flies survived to adulthood which is likely to be as a result of a higher level of Pacman function in these tissues than would have been present in the RNAi tissues (which had an 80% knockdown (Waldron *et al.* 2015). Interestingly, knockdown of *dis3L2* in the *pcm<sup>5</sup>* mutant background resulted control wing area showing a phenotypic rescue. This is in contrast to lethality observed when both *pcm* and *dis3L2* are knocked down simultaneously throughout the wing imaginal discs. This could have been due to the increased activity of Pacman in the hypomorphic mutant compared to the RNAi. However, a similar phenotypic rescue was observed when *dis3L2* was knocked down in the *pcm<sup>14</sup> null* mutant background and therefore it is unlikely to be due to the difference in Pacman activity. Previous work has shown that the knockdown of *pacman* by RNAi has no effect on the overall size of the wing imaginal discs (Waldron 2014) and therefore a similar phenotypic rescue study was not possible.

Knockdown of *dis3L2* in the *pcm<sup>14</sup>* null mutant background, however, did not rescue the *pcm<sup>14</sup>* induced developmental delay and pupal lethality was also observed. This shows that knockdown of *dis3L2* does not simply compensate for the pathways misregulated following the

loss of Pacman function. This has been shown further through genetic analysis where the levels of the developmental delay inducing *dilp8* remain highly overexpressed in the *pcm<sup>14</sup>* *dis3L2* knockdown tissues at a level not significantly different from the *pcm<sup>14</sup>* discs alone. Additionally, the pro-apoptotic gene *rpr*, which also increases in the *pcm<sup>14</sup>* mutant and is likely to contribute to the pupal lethality, is also unaffected by Dis3L2 knockdown.

In general it is difficult to compare RNA interference and mutant experiments due to the fundamental differences in the techniques and their overall result on gene expression. However, it is clear that the two pathways of RNA decay appear to have distinct functions in regulating tissue growth and development. Additionally, the additive effects observed in the RNAi experiments where the activity of one is essential for viability if the other is impaired. This also shows that even within the 3'-5' pathways of decay specificity remains as the exosome is clearly unable to compensate for the loss of Dis3L2 activity.

### **6.8.2 Are Pacman and Dis3L2 involved in regulating opposing pathways?**

The data presented in this chapter indicate functions for Dis3L2 and Pacman in the regulation of opposing pathways. Previous published work from the lab has shown Pacman to be involved in the regulation of apoptosis through its control over the levels of *hid*, *grim* and *reaper*. The increase in *reaper* expression has been repeated in this work and it has been shown to be unaffected by the additional knockdown of *dis3L2*. Similarly *hid*, *grim* and *reaper* remain unchanged in the *dis3L2* sequencing data (Chapter 5) confirming that the activity of Pacman on the mRNA encoding the three pro-apoptotic proteins is specific. This suggests that Dis3L2 does not function in regulating the onset of apoptosis which was further established through the use of the *Df(3L)H99* deletion. The *Df(3L)H99* induced inhibition of apoptosis resulted in an additive effect on the *dis3L2*-depletion induced wing overgrowth. *Drosophila* containing the *Df(3L)H99* deletion together with *dis3L2* knocked down presented with wings that were significantly larger than either those with the deletion or knock down alone. These data

indicate that although Dis3L2 has been implicated in the clearance of mRNAs during apoptosis (Thomas *et al.* 2015) it is not involved in the regulation of the onset of apoptosis.

This work also strengthens the hypothesis proposed in Chapter 4 where Dis3L2 is proposed to be regulating proliferation in the developing wing imaginal disc. The work outlined in Chapter 4 showed strong evidence for a role for Dis3L2 in the control of proliferation. Its knockdown directly results in an increased level of proliferation which subsequently leads to an increase in the number of cells in a wing and significantly larger wings.

It is therefore intriguing that two pathways of RNA decay appear to be involved in the regulation of opposing pathways through targeting specific RNAs for degradation. As discussed in 6.8.1 it appears that the increased proliferation induced following the loss of *dis3L2* is able to compensate for the *pcm*<sup>14</sup> induced apoptosis at least in terms of tissue area. This further shows pathway specificity which may be achieved through a number of mechanisms.

### **6.8.3 How might the pathways specificity be achieved?**

The work presented in this thesis leads towards the hypothesis that Pacman and Dis3L2 are involved in the regulation of apoptosis and proliferation respectively. Pacman appears to function to prevent apoptosis whilst Dis3L2 functions to inhibit proliferation. Sequencing experiments have been performed by Jones *et al* together with the experiment outlined in Chapter 5 to identify transcripts that may be involved in these pathways. Unfortunately, as discussed in Chapter 5, although the RNA-seq experiment revealed two potential targets of Dis3L2 their role in the regulation of proliferation was not immediately apparent. However, the phenotypic data is incredibly strong.

The data build upon the hypothesis in the Newbury lab and shared by others that RNA decay is targeted and specific enzymes are responsible for the degradation of specific transcripts which elicits their regulation of specific cellular pathways. How might the specificity be achieved? Recent work has identified a potential mechanism for targeting RNAs to Dis3L2. For example,

mutation of *dis3L2* in *S. pombe* results in the accumulation of uridylated transcripts. This taken with the previous finding that *pre-let-7* is degraded by Dis3L2 following the addition of uridines to the 3' end implicates a role for uridylation in the targeting of specific transcripts for degradation by Dis3L2. Interestingly *let-7* is known as a tumour suppressor through the inhibition of proliferation (Esquela-Kerscher *et al.* 2008); therefore Dis3L2 may be eliciting part of its role in controlling proliferation through the biogenesis of *let-7*. Until very recently a uridyl transferase had not been identified in *Drosophila*, however, in July 2015 Reimao-Pinto *et al* identified Tailor (CG1091) as a protein responsible for uridylation in *Drosophila* in both S2 cells and adult flies (Reimao-Pinto *et al.* 2015). It would therefore be of interest to investigate any potential interaction between Dis3L2 and Tailor in the future.

Alternatively there has been little evidence for regulatory elements or terminal additions that stimulate the activity of Pacman. Is it therefore likely that these are distinct from those identified for Dis3L2 and thus would provide a mechanism for specificity.

In addition to terminal nucleotide additions the specificity may also be provided by association with other RNA binding proteins. For instance, Pacman has been shown to form a complex with the decapping proteins Dcp1 and Dcp2 (Braun *et al.* 2012) which brings it to the transcripts for degradation. However, this does not answer how the specific transcripts are signalled for degradation in the first place. Further biochemical work is required to identify the features involved in the specific targeting of transcripts to either Pacman or Dis3L2.

## 6.9 Acknowledgements

The RNA-sequencing on *pacman* null mutants was performed and analysed by Dr. Chris Jones.

# Chapter 7: P-element mutagenesis screens and Dis3L2 antibody production.

## 7.1 Introduction

Previous work in this thesis has focused on using RNA interference (RNAi) to knockdown *dis3L2* to study its role in the control of tissue growth and development. Although an efficient knockdown of 80% was achieved using RNAi; 20% wild type Dis3L2 would remain which may be sufficient to mask some of its roles within the cell. Therefore the optimal approach would be to use a null mutant which would ensure that no Dis3L2 activity was present. Having observed that the ubiquitous knockdown of *dis3L2* has no effect on organism viability, one would hypothesise that there would be no issue with the viability of a null *dis3L2* mutant.

A *Drosophila dis3L2* mutant was not available and therefore P-element mutagenesis was used in an attempt to produce a null mutant. P-element mutagenesis is a commonly used method for making mutants in *Drosophila* through randomly inducing mutations (normally deletions) into a gene of interest. It requires the insertion of a P-element in close proximity to the gene of interest. Introducing a transposase causes random hop out of the P-element which may or may not remove a region of the gene of interest resulting in a mutation. The effect the induced mutation has is dependent upon the nature of the genetic aberration itself. Previous work in the lab has produced a number of hypomorphic and null mutations in the 5'-3' exoribonuclease Pacman using this method (Grima *et al.* 2008; Jones *et al.* 2013; Waldron *et al.* 2015; Jones *et al.* 2016). This has led to the findings that Pacman regulates apoptosis through the regulation of *hid*, *grim* and *rpr* expression (Waldron *et al.* 2015). A similar mutant for *dis3L2* would be highly beneficial for further studies into its function in co-ordinating tissue growth.

In addition to the P-element mutagenesis experiments, this chapter will also outline an attempt to make a peptide antibody to Dis3L2 as no antibody was commercially available throughout the project. Having a specific antibody to Dis3L2 would open up a number of techniques which would be critical for understanding the mechanisms behind the cellular processes Dis3L2 is involved in. For example, to identify the extent to which Dis3L2 was knocked down using Western blotting or to identify those RNA targets specifically bound to Dis3L2 using cross-linking immunoprecipitation experiments. This would be particularly important to fully validate the findings presented in Chapter 5.

## 7.2 Aims

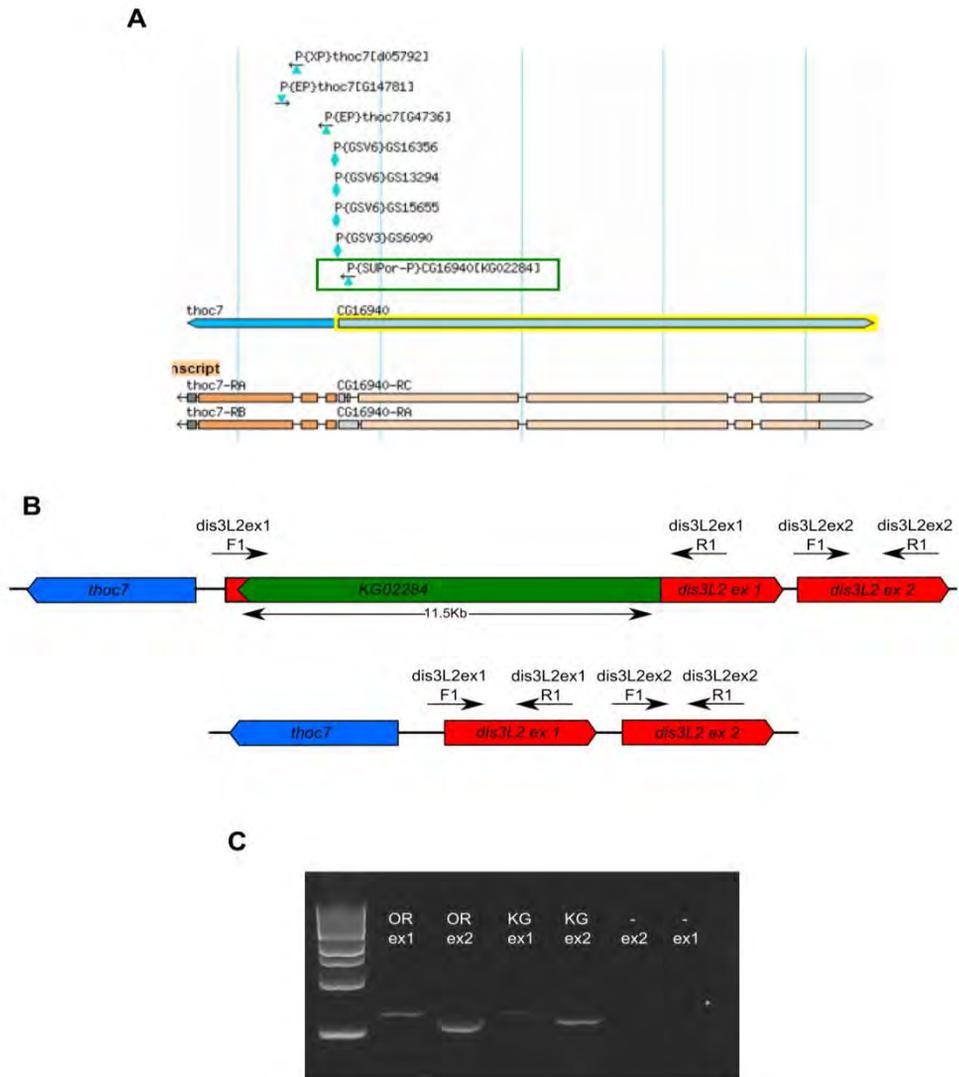
The aims of this chapter were:

- To create a null *dis3L2* mutant using P-element mutagenesis.
- To characterise the produced mutant both genetically and phenotypically.
- To design and produce a Dis3L2 antibody.

## 7.3 Attempted production of a *dis3L2* mutant

### 7.3.1 An annotated P-element stock with potential for use in *dis3L2* mutagenesis does not contain the specified P-element

Flybase contains a feature showing the random insertion sites, and where known, direction of P-elements across the *Drosophila* genome (Figure 7.1A). Using this it was possible to select lines containing suitably located P-element stocks which upon removal might result in a genetic deletion in *dis3L2*. The first stock selected was Bloomington stock 13307 ( $y^1$ ;; *P{SUPor-P}CG16940<sup>KG02284</sup>*) which was defined as having an 11.467Kb insertion 64bp into the *dis3L2* gene. To check the presence of the P-element primers were designed to span the insertion region (Figure 7.1B). If the P-element was present no band would be expected in P-element homozygotes due to the inability to amplify a PCR produced over 11Kb in length. However, if a wild-type sized product was observed this would imply the P-element was absent and therefore make the stock redundant for the required purpose. On screening the P-element stock alongside control stocks, wild-type products were observed for both the control and P-element containing stock (Figure 7.1C). This showed that the P-element was not present in the stated location within the genome and thus meant it was inappropriate to use in a P-element mutagenesis screen.



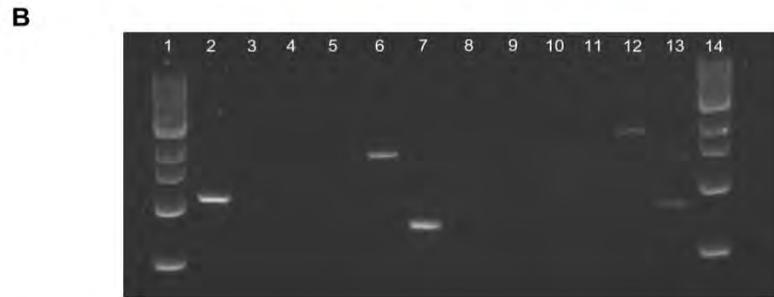
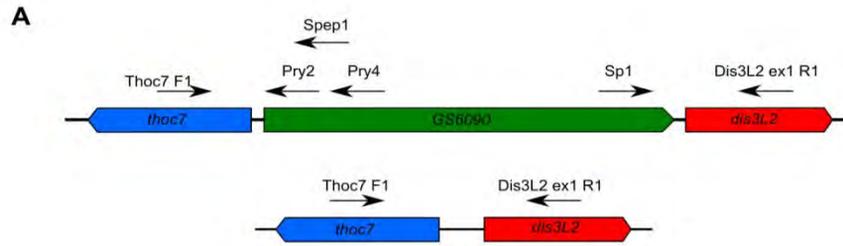
**Figure 7.1 Checking the presence of  $P\{KG02284\}$  before proceeding with P-element mutagenesis. (A)** A screenshot from Flybase showing the P-element insertions around the *dis3L2* (*CG16940*) gene.  $P\{KG02284\}$  is highlighted using a green box. **(B)** Cartoon showing the location of  $P\{KG02284\}$  64bp into the *dis3L2* gene (upper) together with the normal genetic arrangement (lower). Also shown are the primer binding sites for the primers used to check its presence. **(C)** PCR showing that  $P\{KG02284\}$  is absent. A band of wild-type size is observed for both the wild-type (OR ex1) and the theoretical P-element containing stock (KG ex1) showing successful amplification across the insertion site. A positive control (ex2) band is present for both OR and KG. A 1Kb ladder was used as a marker with the lower band representing 500bp. ex1 = primer pair *dis3L2ex1* F1 and *dis3L2ex1* R1, ex2 = primer pair *dis3L2ex2* F1 and *dis3L2ex2* R1.

### 7.3.2 Sourcing and testing of a second P-element containing stock

Having identified that BL13307 was not suitable for the P-element mutagenesis experiments another stock was sourced from the Drosophila Genetic Resource Center in Kyoto (stock 200902: *w\** ;; *P{GSV3}GS6090/TM6*, now referred to as GS6090). GS6090 was defined as containing a 4.828Kb insertion 8bp upstream of *dis3L2*. This was the closest P-element to the *dis3L2* locus excluding BL13307. Unlike the previous stock, the direction of the insertion was unknown. PCR was again used to determine the presence and orientation of the insertion.

The GS6090 stock was homozygous lethal and therefore was balanced over TM6. This meant that using primers spanning the P-element would not allow identification of the presence of GS6090 due to the wild-type *dis3L2* copy of the TM6 chromosome. Therefore specific primers were designed complementary to the GS6090 sequence itself. These were used together with a primer either upstream (Thoc7 F1) or downstream (Dis3L2 ex1 R1) of the insertion site (Figure 7.2A/C). This approach also allowed the determination of the P-element orientation. The spanning primer pair (Thoc7 F1 + Dis3L2 ex1 R1) served as a positive control for both the GS6090 and control stocks. Finally, as an additional positive control unrelated to the *dis3L2* locus a primer pair in the first exon of *dis3* was used (Dis3 ex1 F1 + Dis3 ex1 R1).

Both positive controls gave the expected PCR products in both stocks (Figure 7.2B lanes 6,7,12 and 13). Although most of the primer pairs showed no product, the pair Thoc7 F1 + Pry2, showed a band of 1.4Kb specifically in the GS6090 sample (Figure 7.2B compare lanes 2 and 8). The other 3 primer pairs tested gave no product in either sample suggesting they were not compatible as a pair under the reaction conditions used. However, the presence of the specific product for Thoc7 F1 and Pry2 gave the required information and therefore no further optimisation of the other primer pairs was required. This screen therefore confirmed the correct P-element sequence had been annotated together with the orientation. The binding site of Pry2 is situated in the 5' end of GS6090 (Figure 7.2A) which shows it is orientated in same direction as *dis3L2*. Having validated its presence and identified its orientation, GS6090 was a suitable insertion stock to use in a P-element mutagenesis screen.



**C**

Primer 1	Primer 2	Product size	Band in GS6090?	Band in UAS-EGFP <sup>RNAi</sup> ?	Lane number (GS0690/UAS)	Information gained
Pry2	Thoc7 F1	~875bp	Yes	No	2/8	P-present in same orientation as dis3L2
Pry4	Thoc7 F1	~970bp	No	No	3/9	Primer pair not compatible
Spep1	Thoc7 F1	~937bp	No	No	4/10	Primer pair not compatible
Sp1	Dis3L2 ex1 R1	~680	No	No	5/11	Primer pair not compatible
Thoc7 F1	Dis3L2 ex1 R1	1418bp	Yes	Yes	6/12	As expected as would be present on the TM6 chromosome. (6246bp in presence of P)
Dis3 ex1 F1	Dis3 ex1 R1	647bp	Yes	Yes	7/13	As expected. Positive control.

**Figure 7.2 The  $P\{GS6090\}$  insertion is detected in the new P-element stock ( $GS6090$ ).** (A) Cartoon showing the location of  $P\{GS6090\}$  8bp upstream of the  $dis3L2$  gene (upper) together with the normal genetic arrangement (lower). Also shown are the primer binding sites for the primers used to check its presence. (B) PCR showing the presence and orientation of  $P\{GS6090\}$ . Lanes 2 to 7 are PCR samples on DNA produced from  $;;P\{GS6090\}/TM6$  whilst lanes 8 to 13 are produced from a line not containing the P-element insertion ( $;;UAS-EGFP^{RNAi}$ ). The primer pairs used were: lanes 2 and 8 = Thoc7F1 + Pry2, lanes 3 and 9 = Thoc7F1 + Pry4, lanes 4 and 10 = Thoc7F1 + Spep1, lanes 5 and 11 = Dis3L2ex1 R1 + Sp1, lanes 6 and 12 = Thoc7F1 + Dis3L2ex R1, lanes 7 and 13 Dis3ex1 F1 + dis3ex1 R1 (used as a positive control). 1Kb ladders were run in lanes 1 and 14. (C) Summary of the primer pairs used together with the expected and observed products.

### 7.3.3 P-element mutagenesis using *GS6090* in an attempt to produce a *dis3L2* mutant.

To mobilise the P-element, isolate the individual excision events and stably balance a potential *dis3L2* mutant stock required a 4 generation cross (Figure 7.3). The first cross (cross A) involved crossing *GS6090* to a stock containing a constitutively active transposase (*In(1)<sup>m4h</sup>;; Df(3L)Ly,sens<sup>Ly-1</sup>/TM3p<sup>Δ2-3</sup>99B*) to induce P-element mobilisation. Mobilisation of a P-element often results in a mosaic eye phenotype as observed in Figure 7.4. Those offspring containing the mobilised P (referred to from now on as P\*) and the transposase were selected by the presence of Stubble (Sb) which gives short thoracic hairs together with the absence of Lyra (Ly).

The second cross (cross B) was designed to remove the transposase (*TM3p<sup>Δ2-3</sup>99B*) to prevent any further mobilisation or genetic rearrangement. This was achieved through crossing to a *UAS-pcm<sup>27ND</sup>/TM6* stock. This stock was selected due to carrying a balancer (TM6) which contains both a larval marker (Tubby (Tb)) and an adult marker (Sb). In addition to this the *UAS-pcm<sup>27ND</sup>* insertion gives the flies a yellow eye colour which can also be used in the selection process. From this cross P\*/TM6 were selected by the presence of Sb, Tb and the absence of yellow eyes. Successful P-element excision (in the absence of the transposase) was shown by a white eye phenotype. 6% of offspring showed the white eye phenotype; however, this was much lower than previously observed in the lab (Jones 2011) suggesting a reduced 'hop out' efficiency. 79.3% of the offspring showed an orange eye phenotype similar to the parental P-element stock suggesting unsuccessful excisions. A smaller percentage (14.7%) of offspring showed a darker orange/red eye phenotype which was indicative of the re-insertion of the P-element in a region of the genome where it was expressed at greater levels.

To isolate each potential unique excision event a third cross (cross C) was used where each P\*/TM6 male or virgin female was crossed to the *UAS-pcm<sup>27ND</sup>/TM6* stock in 140 individual crosses. Using individual males or virgin females ensured that different excision events were not mixed which could result in their repair by recombination. From this cross the final stocks were made by crossing P\*/TM6 males and virgin females within each individual vial from cross C, producing a final balanced, stable stock which could be screened for mutations in *dis3L2* (cross D).

On the generation of the final mutant lines all 140 showed homozygous lethality where only Tubby larvae or pupae were observed. The initial P-element stock was also homozygous lethal, however, it was believed to be caused by the P-element itself and therefore those stocks that show excision would have been expected to lose the lethality. To assess if the lethality in the

parental P-element stock was due to the P-element a deficiency covering the *dis3L2* locus was used. If the lethality was indeed due to the P-element one would not expect the lethality to be rescued by the deficiency. However, if the lethality was rescued then this would show it was caused by a factor elsewhere on the chromosome. When GS6090 was crossed to the deficiency (Df) the resulting *PGS6090/Df* flies were viable indicating the observed lethality was caused by a factor elsewhere on the chromosome not covered by the 160Kb deletion. It was therefore assumed that the lethality in all the potential mutation stocks was also driven by this unknown factor and thus the stocks were no longer of any use and were discarded.

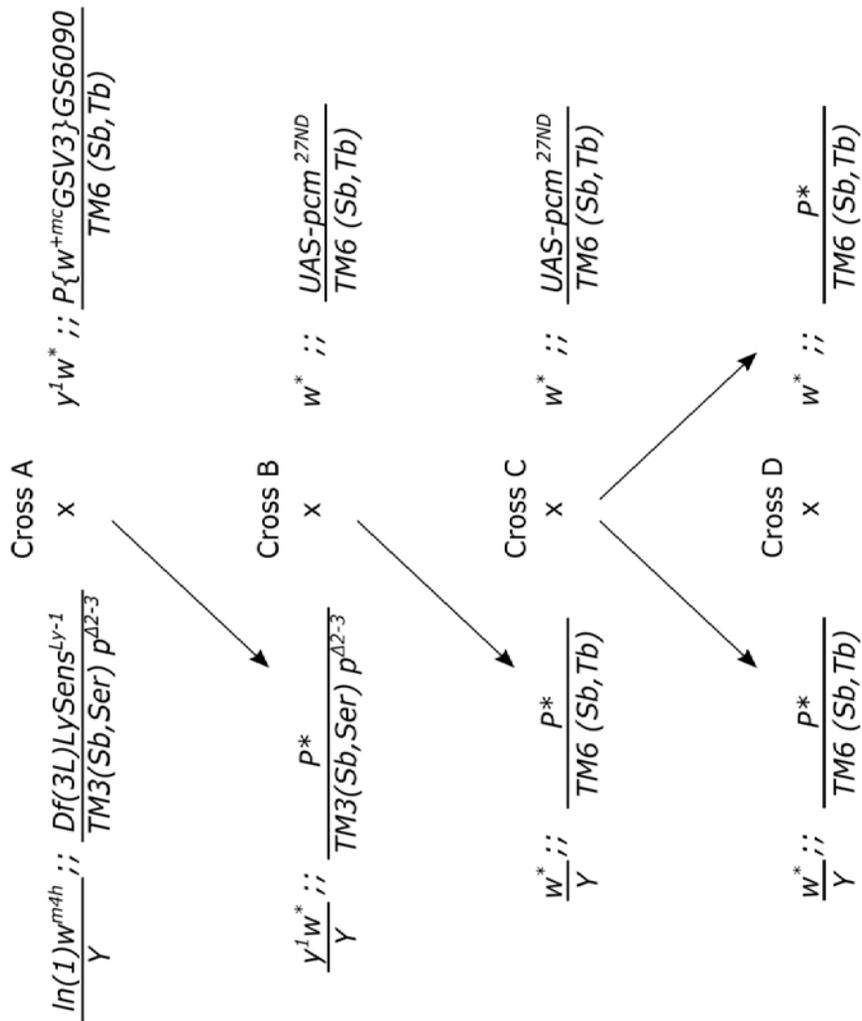
#### **7.4 Removal of the lethality causing factor by using recombination.**

The factor inducing the lethality in GS6090 had to be removed before any repeat of the mutagenesis experiment. Recombination was used to repair the region causing the lethality in the parental P-element stock. To allow the desired recombination to occur GS6090 was crossed to a stock containing two wild-type 3<sup>rd</sup> chromosomes (Figure 7.5 cross A). This allowed recombination to occur between the chromosome containing the lethality inducing factor and the wild-type chromosome which, in some cases, would result in the repair of the region inducing lethality (P\*). P-element containing flies show orange eyes, therefore a white eyed stock containing the wild-type 3<sup>rd</sup> chromosome was used which allowed the use of the orange eye phenotype to track the presence of the P-element.

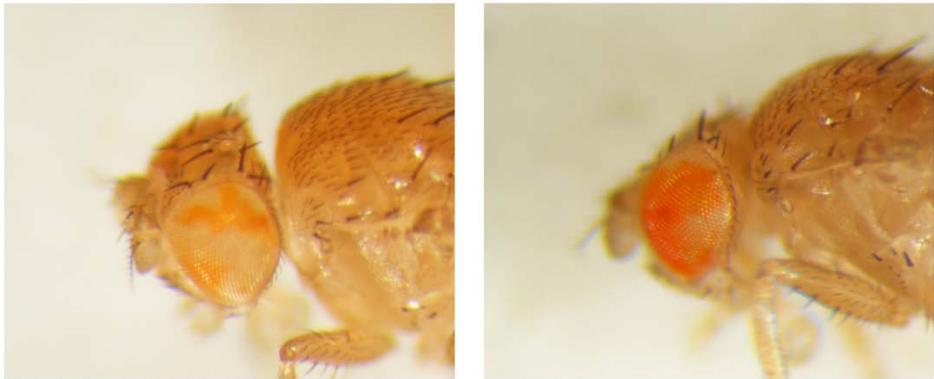
Female progeny from cross A which contained the randomly recombined 3<sup>rd</sup> chromosome were crossed to males containing a double balanced 3<sup>rd</sup> chromosome to balance the stocks and ensure any further recombination was inhibited (Figure 7.5 cross B). Resulting males were then individually back-crossed to GS6090 as it carried a suitable 3<sup>rd</sup> chromosome balancer with an adult (Sb) and larval marker (Tb). Crosses were set up individually to isolate each unique recombination event (Figure 7.5 cross C). Resulting, orange eyed, P\*/TM6 (Sb.Tb) males and virgin females were crossed together to produce a stable stock of each potential unique recombination event (Figure 7.5 cross D). Those stocks containing non-tubby (P\*/P\*) larvae would show a successful rescue of the lethality. 75 individual stocks were set up, of which 19 showed non-tubby progeny indicating a 25% success rate.

As previously mentioned the P-element was tracked using orange eye colour. However, PCR was performed on the selected lines showing non-tubby progeny to confirm its presence. The P-element specific Thoc7 F1 + Pry2 primer pair was used to confirm its presence. 4 lines were screened using these primers, all of which showed a P-element specific product confirming the

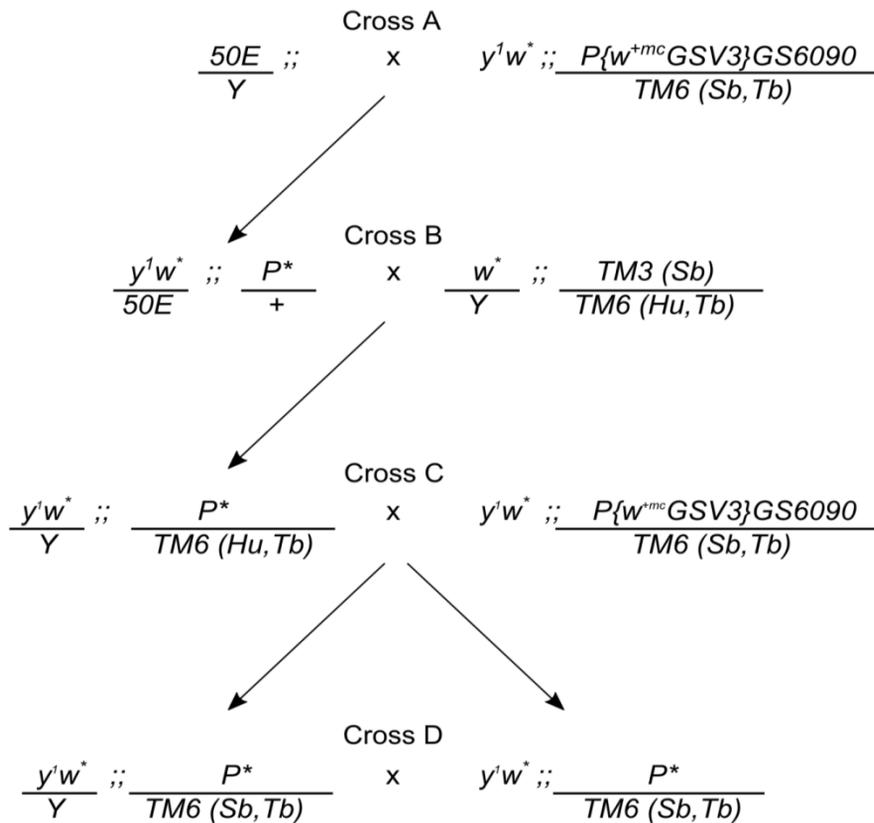
P-element was indeed still present in the rescued lines (Figure 7.6). To ensure the full length P-element remained the primer spanning the insertion site which if the full length P-element was present no product should be produced in the P/P homozygotes. No product was observed in any of the 4 lines indicating the full length P-element remained (Figure 7.6B). As a control, heterozygous flies from the same stocks were also screened using the same two primer pairs. Due to the presence of the balancer, unlike the homozygous flies, products should be observed for both primer pairs; this was indeed the case (Figure 7.6C). In contrast, a control stock which did not contain the P-element was negative when screened by PCR using the P-element specific primers whilst the spanning product was present (Figure 7.6B/C). Stock 5 was selected to use in future mutagenesis crosses.



**Figure 7.3: P-element mutagenesis cross scheme.** To excise  $P\{GS6090\}$  and produce balanced stocks a 4 generation cross scheme was required. The first cross (Cross A) involved crossing in a constitutively active transposase ( $p^{\Delta 2-3}$ ) to induce the excision of  $P\{GS6090\}$ . The desired progeny were selected for by the absence of Lyra (Ly) and the presence of Stubble (Sb). The selected progeny were crossed to a stock carrying a suitable balancer which allowed the removal of the transposase together with retaining the integrity of the chromosome (Cross B). The selected progeny from Cross B were then individually backcrossed (Cross C) to the balancer containing stock used in Cross B. Setting up these crosses individually maintained the purity of each potential unique excision event. Finally selected male and virgin female progeny (using eye colour as the  $UAS-pcm^{27ND}$  insertion gives a yellow eye) from Cross C were crossed together to create a stable stock for each potentially unique excision event. These stocks were then ready to be screened by PCR.



**Figure 7.4: Examples of mosaic eye phenotypes observed in the presence of the transposase indicating P-element mobilisation. These were progeny from Cross A in Figure 7.3.**



**Figure 7.5: Cross scheme used to rescue the lethality inducing factor on the *P{GS6090}* containing chromosome.** Recombination was used to repair the damaged chromosome and restore homozygote viability. *P{GS6090}/TM6* females were crossed to *50E* males. *50E* males were used as they have white eyes allowing the P-element to be tracked through eye colour as it gives orange eyes. The resulting *50E/y<sup>1</sup>w<sup>\*</sup>;; P<sup>\*</sup>/+* females have the potential for the *P<sup>\*</sup>* and wild-type (+) chromosomes to recombine and therefore potentially repair the damaged chromosome. These females were then crossed to a suitable balancer containing male to prevent any further recombination and retain the P-element (Cross B). Individual male progeny from Cross B which carried the potentially repaired chromosome was then crossed to virgin females carrying a desirable balancer in individual vials to isolate each repair event. The resulting males and virgin females carrying the potentially repaired chromosome and desirable balancer were crossed (Cross D) to produce a stable stock of each isolated repair event. The presence of non-Tubby progeny in this stock would indicate chromosomal repair.

## 7.5 P-element mutagenesis using the rescued P-element stock.

### 7.5.1 Initial mutagenesis experiment using the rescued P-element stock.

Having rescued the lethality in GS6090 a new round of P-element mutagenesis was performed in an attempt to make a *dis3L2* mutant. The same cross scheme (Figure 7.3) was used resulting in 67 lines carrying potential mutations. Unfortunately only 6% of all offspring from cross B had white eyes showing low efficiency of P-element excision as was observed in the previous round (Section 7.3.3). Additionally, of the 10 bottles of cross B set up only 4 produced any progeny meaning the overall numbers were severely depleted.

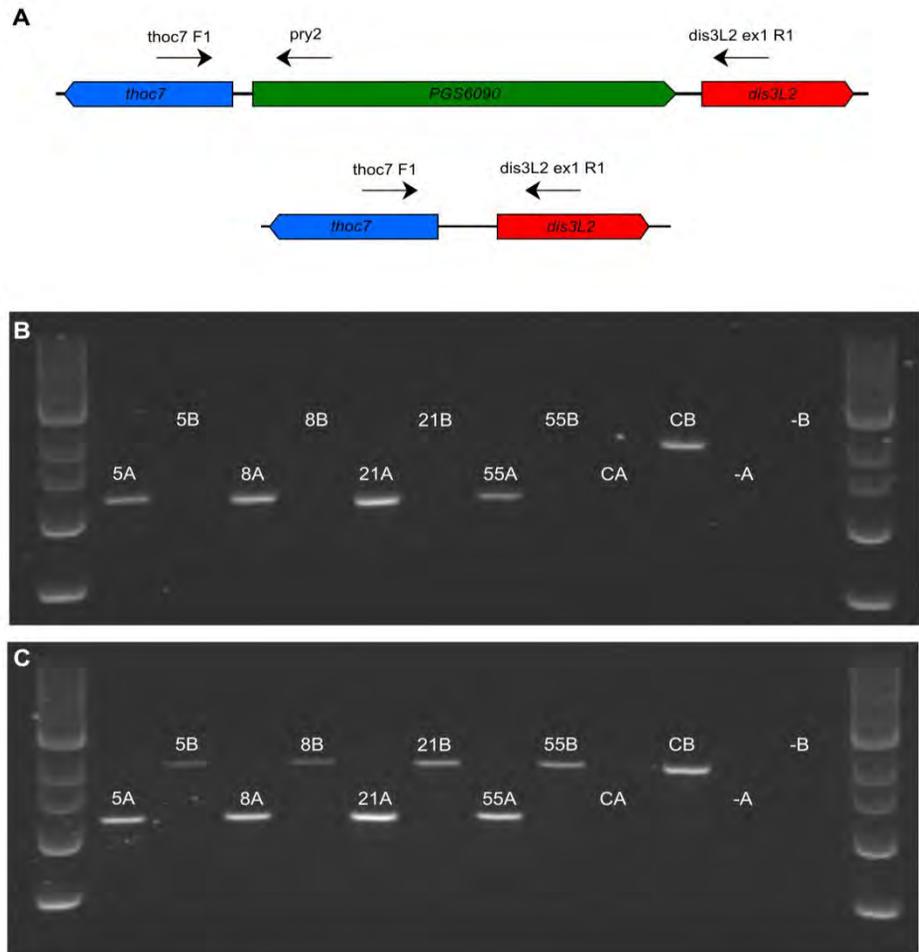
Unlike the previous cross only 5 lines showed homozygous lethality. To assess the nature of the lethality *P\*/TM6* males of the 5 lines were crossed to virgin females carrying the 160Kb deletion covering *dis3L2* (*Df*). The progeny of each cross were then scored for *P\*/Df* viability. In the case of all 5 lines *P\*/Df* larvae were observed indicating the lethality was caused by a factor outside of the locus covered by the deletion. This was most likely due to the re-insertion of the P-element resulting in the loss of expression of a critical gene/genes. These stocks were discarded from any further investigation.

The remaining 62 lines were screened by PCR for the presence of the P-element and any deletions around the excision site. To achieve this specific primer pairs were used including those outlined previously (Figure 7.7A). Thoc7 F1 and Pry2 were used to detect the P-element whilst Thoc7 F1 and Dis3L2 ex1 R1 were used as they spanned the insertion site and therefore would give information on any large deletion. Finally as a positive control a primer pair specific for the first exon of *dis3* was used. Of the 62 lines only 6 showed a complete excision of the P-element, all of which were white eyed. However, all 6 showed bands of wild-type size for the spanning primer pair suggesting no large deletion had occurred. An example of the PCR screen is shown in Figure 7.7B. To investigate if a small deletion or other mutation had occurred, PCR products produced by the spanning primer pair were cleaned up and sent for DNA sequencing with MWG Eurofins. These sequences were aligned using AlignX in vector NTI which showed stocks 5, 6, 9 and 11 had a wild-type sequence indicating a clean excision but unfortunately a complete lack of mutation. However, stock 93 showed a 36bp insertion which mapped to the P-element sequence indicating a small fragment of the P-element remained (Figure 7.7C). Although no deletion or mutation was observed the insertion may have an influence on *dis3L2* expression. As a result the level of *dis3L2* mRNA in stock 93 *P\*/P\** larvae was assessed;

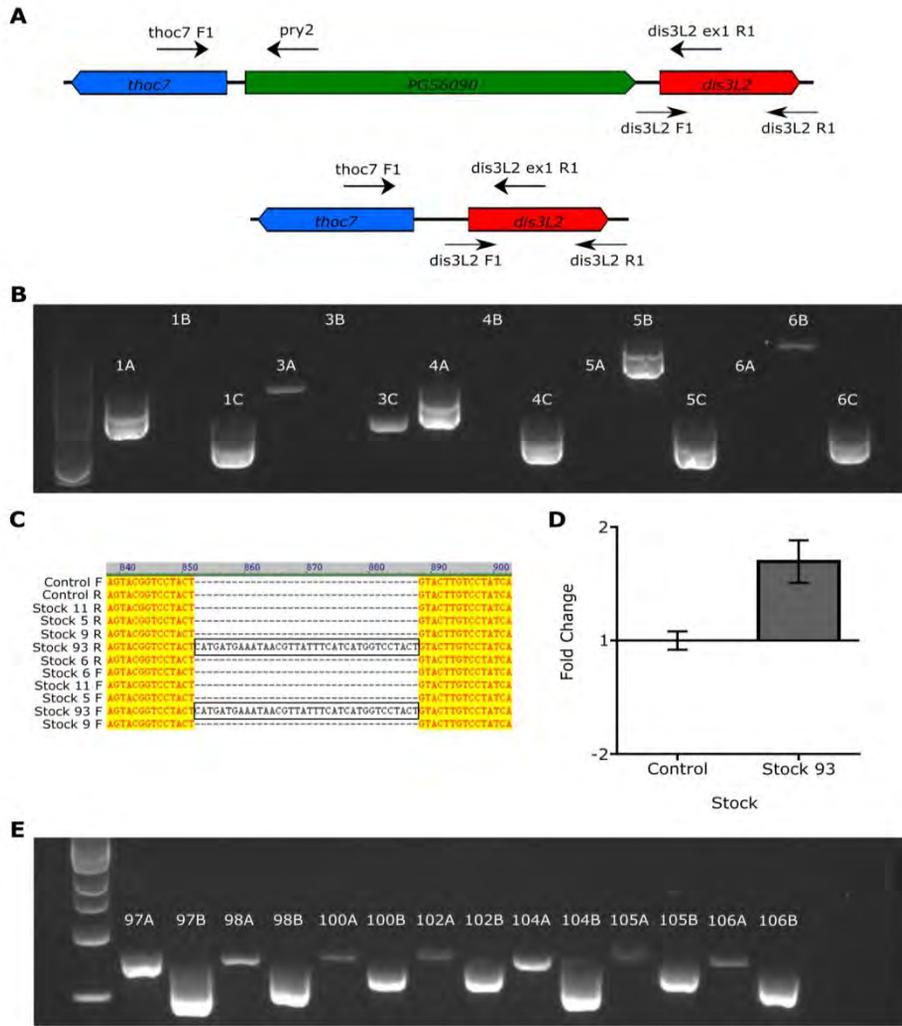
however, no significant difference in *dis3L2* mRNA expression was observed between stock 93 and control larvae (Figure 7.7D).

The other 56 lines showed a PCR product for the P-element. Along with showing a PCR product for the P-element the stocks were all negative for a product from the spanning primer pair. This could have been caused by two factors; first, all or a large fragment of the P-element remained meaning the spanning product would be too large to amplify. Alternatively a small fragment of the 5' end of the P-element remained containing the Pry2 binding site but a deletion following the removal of most of the P-element removed the *dis3L2* ex1 F1 primer binding site. To identify if any line contained a deletion removing the primer binding site the remaining lines were screened further with primer pairs in the first exon of *dis3L2* (*dis3L2* F1 + R1). All 56 lines showed a PCR product (Figure 7.7E) indicating a large region of the P-element remained in each case and therefore were highly unlikely to be *dis3L2* mutants.

Therefore no line resulting from the P-element mutagenesis crosses contained a *dis3L2* mutation. However, as previously stated only a small number of final lines were able to be set up due to the loss of over half of the crosses. As a result the cross scheme was repeated on a larger scale to account for any potential losses and to maximise the chance of inducing a mutation in *dis3L2*.



**Figure 7.6: PCR to check the presence of the P-element in the repaired stocks. (A)** Cartoon showing the location of *P{GS6090}* 8bp upstream of the *dis3L2* gene (upper) together with the normal genetic arrangement (lower). Also shown are the primer binding sites for the primers used to check its presence. **(B)** TM6 negative flies show the homozygous presence of the P-element *P{GS6090}*. Numbers represent the rescued stock number apart from C which is a control stock (*;UAS-EGFP<sup>RNAi</sup>*). 2 primer pairs were used: A = thoc7F1 + pry2 (*P{GS6090}* specific pair) and B = thoc7F1 + dis3L2ex1 R1 (spanning the insertion site of *P{GS6090}*). Presence of band A in each rescued stock and the absence of band B confirmed the presence of the P-element. A 1Kb ladder was loaded at either end of the gel. **(C)** The same experiment as B but using heterozygous *P{GS6090}/TM6* flies as a positive control for the spanning primer pair (pair B).



**Figure 7.7: P-element mutagenesis using  $P\{GS6090\}$ .** (A) Cartoon showing the location of  $P\{GS6090\}$  8bp upstream of the *dis3L2* gene (upper) together with the normal genetic arrangement (lower). Also shown are the primer binding sites for the primers used to check its presence. (B) Example of a PCR screening gel. 3 primer pairs were used the locations of 2 of which are shown in A: A = thoc7F1 + pry2, B = thoc7F1 + dis3L2ex1 R1 and C = dis3ex1 F1 + dis3ex1 R1 used as a positive control. Successful excision of  $P\{GS6090\}$  is observed in samples 5 and 6 through the absence of a product for band A but the presence of band B. (C) DNA sequencing over the insertion site of 5 stocks showing successful excision. Only stock 93 shows a non-wild-type sequence with an insertion left over from  $P\{GS6090\}$ . (D) The 40bp insertion in stock 93 does not cause a reduction in *dis3L2* mRNA expression.  $n=3$ , error bars represent standard error. (E) Example of a PCR to screen for any deletion removing the dis3L2ex1 R1 binding site. A forward primer was used which bound immediately downstream of the insertion site (dis3L2 F1) with the reverse also in the first exon of *dis3L2* (dis3L2 R1). This is referred to as primer pair A. Bands for every stock were identified showing no deletion into *dis3L2*. Primer pair B is the dis3ex1 F1 + dis3ex1 R1 positive control pair. A 1Kb ladder was loaded in the first lane.

### 7.5.2 Repeat mutagenesis using the rescued P-element stock.

Due to the thus far failed attempts to create a *dis3L2* mutant a final round of P-element mutagenesis was performed. Previous rounds showed the low efficiency of P-element hop out and therefore the cross was carried out on a much larger scale to enhance the probability of inducing a mutation. The same cross scheme was used as in the previous two rounds (Figure 7.3). Following P-element removal, balancing and isolation 160 lines were set up, of which 134 were white eyed. This was more than 4 times the number of white eyed stocks previously achieved. The other 26 lines had either the parental orange eye or an eye colour other than white or parental orange. Similar to the previous mutagenesis the overall efficiency of the P-element excision was again very low; only 6.1% of the progeny had a successful excision event indicated by white eyes. However, the actual proportion of successful excisions may be even lower still as the previous experiments showed that the white eye phenotype did not mean a complete excision.

Of the 160 potential mutant lines 66 showed some level of homozygous lethality. 17 of those 66 were lethal at the pupal stage as homozygous larvae and pupae were observed, however, homozygous adults were not viable. The remaining 49 were lethal before the 3<sup>rd</sup> instar larval stage which was shown by an absence of homozygous larvae or pupae. All 66 homozygous lethal stocks were crossed to the 160Kb deficiency stock to assess if the lethality was due to a factor within the *dis3L2* locus. All 17 pupal lethal stocks were rescued by the deficiency stocks indicating the lethality was due to a factor elsewhere on the chromosome; most likely due to the re-insertion of the P-element in a damaging location. However, 14 of the 49 remaining lethal stocks were not rescued by the deficiency showing the cause of the lethality lied within the 160Kb region covered by the deficiency, including *dis3L2*. As shown in previous chapters ubiquitous knockdown of *dis3L2* does not affect organism viability, therefore it would be unlikely that a strong mutant of *dis3L2* would have this affect. Furthermore it was not possible to rule out a large deletion removing other nearby genes which may be critical for viability. A crossover approach was used to determine if the lethality inducing factors within these 14 stocks were the same. Non-rescued lethal lines were crossed to each other; lethality rescue would show different causes of lethality whilst no rescue would be indicative of the same lethality causing factor in the two lines. 13 of the 14 stocks showed no rescue when they were crossed to each other indicating that the cause of the lethality within these lines were consistent. However, the lethality in the remaining line was rescued when it was crossed to each of the other lines suggesting a unique cause of lethality.

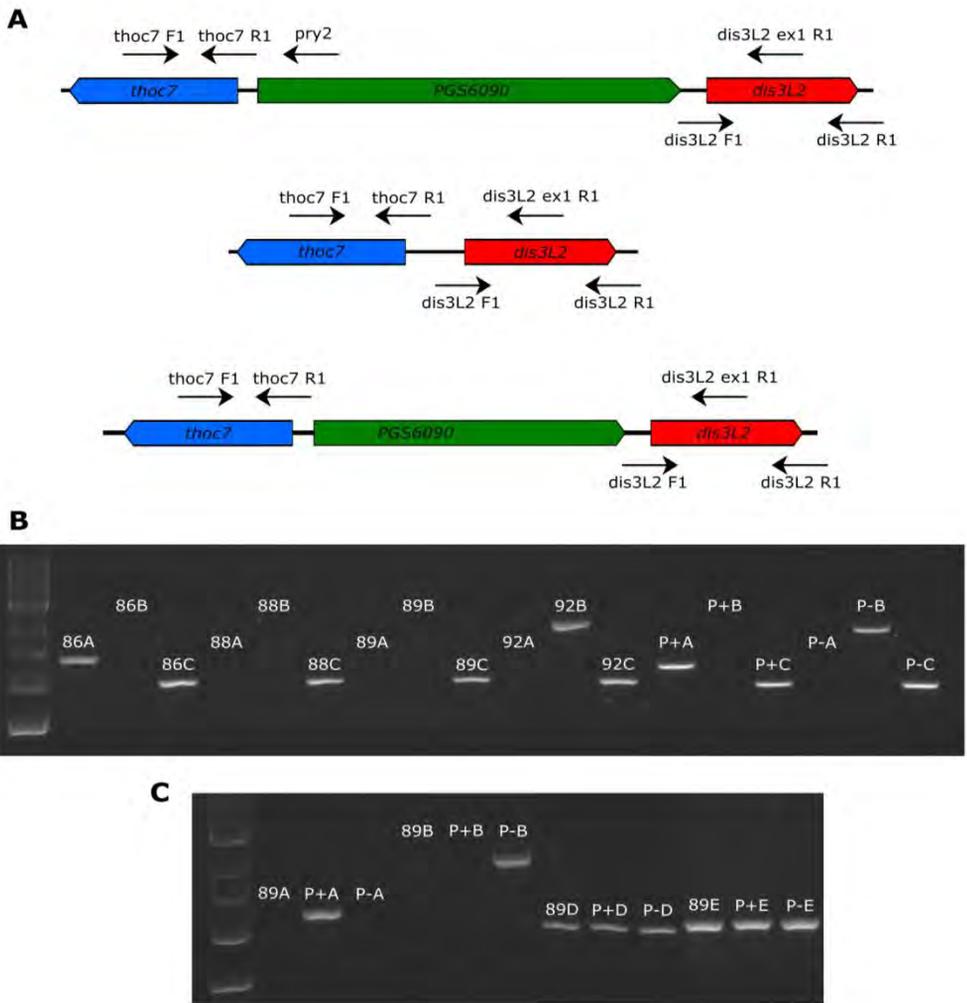
The remaining 94 lines were homozygous viable making them suitable for the same PCR screening as was used previously. Each potential mutant line was screened with the PCR

primers used in the previous mutagenesis experiments (Figure 7.8A). A much higher proportion than the previous experiment showed a clean excision of the P-element with 51.1% (48 lines) showing no P-element specific product (Example shown in Figure 7.8B (see 92A/B)). 26.6% (25 lines) of the screened lines showed the presence of the P-element with no product for the spanning primer pairs indicating a large fragment of the P-element remained (Example shown in Figure 7.8B (see 86A/B)). Unlike the previous experiment, 7.4% (7 lines in total) showed no product for either the P-element or the spanning band (Example shown in Figure 7.8B (see 88A/B)). This pattern suggested the excision of the P-element alongside a potential deletion removing one of the spanning primer binding sites which was indicative of a mutation. These 7 lines were further screened using additional primers designed to bind either side of the P-element insertion to assess the direction of the potential deletion (Figure 7.8A). Primers binding in the first exon of *thoc7* (*thoc7F1+R1*) and the same pair as used previously binding in at the 5' end of *dis3L2* (*dis3L2F1+R1*) were used. The forward primer for the *dis3L2* pair binds 2bp downstream of the P-element insertion, therefore if a deletion occurred into *dis3L2* it would be expected to remove at least the forward primer binding site. However, products for both the *thoc7F1+R1* and *dis3L2F1+R1* primer pairs were observed suggesting no deletion had occurred (example in Figure 7.8C (see 89D/E)). It was therefore concluded that in the absence of a deletion the 5' end containing the *Pry2* binding site had been removed but a large enough 3' fragment remained to prevent amplification across the spanning site. Therefore unfortunately these 7 lines were shown not to contain a deletion into *dis3L2*.

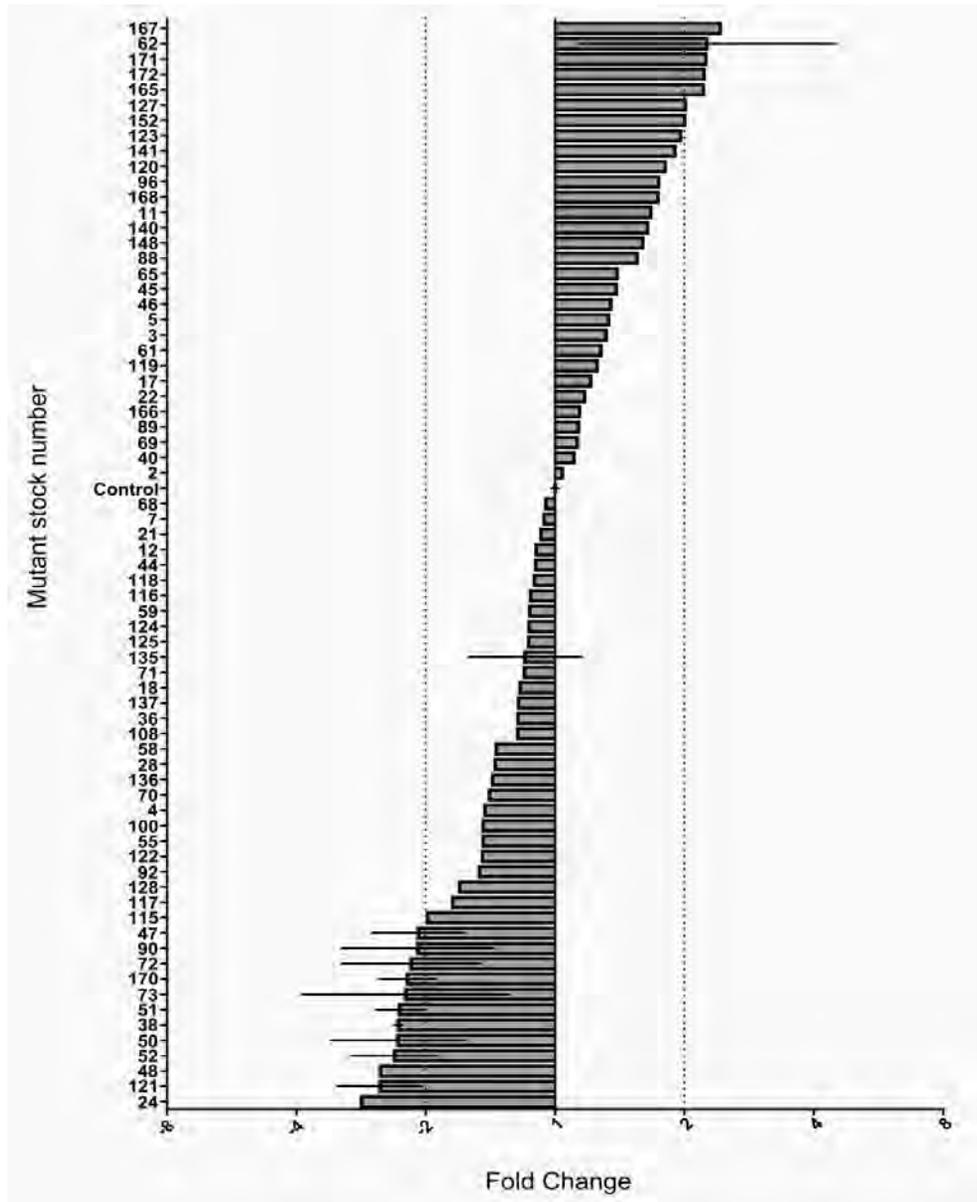
Due to the high proportion of lines showing an excision of the P-element there was still a chance one of the remaining lines contained a small mutation which perturbs *dis3L2* expression. A large deletion was, however, highly unlikely due to the product obtained for the spanning band. To investigate if any of the lines had reduced *dis3L2* expression qRT-PCR was performed on each line to assess the levels of *dis3L2* mRNA. RNA was extracted from whole larvae using the method outlined in (2.4.2.1) and *rp49* was used as a housekeeper. The initial P-element containing stock was the most suitable control due to having the most similar genetic background and had previously been shown to have wild-type levels of *dis3L2*. The level of *dis3L2* varied relative to the initial P-element stock across all the lines, however, no stock showed a large decrease in *dis3L2* expression with the changes ranging from a 2.4 fold increase to a 2.9 fold decrease (Figure 7.9). 9 lines showed a decrease in *dis3L2* mRNA expression greater than 2 fold. However, unfortunately lines 24 and 48 were lost and could not be repeated. The 5 lines showing the greatest decrease in *dis3L2* expression were sent for sequencing. The spanning primer pair (*thoc7F1+dis3L2ex1R1*) was again used to generate the PCR product which was cleaned up using a PCR purification kit (Qiagen, cat. no. 28104) and

subsequently sent for sequencing. Unfortunately, these showed that no deletion had occurred and were therefore not *dis3L2* mutants.

Unfortunately the multiple mutant screening experiments were unsuccessful in producing a *dis3L2* mutant. The inefficiency of P-element hop out (only 6% of flies showed the white eye phenotype implying an excision event) was the major contributing factor to this. Although the crosses were set up on a large scale in an attempt to overcome this issue only 138 potential mutant stocks (with white eyes) were set up which is still a sub-optimal number for an experiment on this nature.



**Figure 7.8: PCR screening of the potential mutant stocks produced during the repeat mutagenesis experiment. (A)** Cartoon showing the location of *P{GS6090}* 8bp upstream of the *dis3L2* gene (upper) together with the normal genetic arrangement (middle). Also shown are the primer binding sites for the primers used to check its presence. **(B)** An example of a gel in the PCR screen showing examples of successful excision (92A/B) and unsuccessful excision (86A/B). Also shows potential deletion mutants which were negative for both A and B (89A/B). P+ was using DNA produced from *P{GS6090}* flies whilst P- was using DNA produced from a *P{GS6090}* negative stock (*;UAS-pcm<sup>27ND</sup>/TM6*). 3 primer pairs were used the locations of 2 of which are shown in **A**: A = thoc7F1 + pry2, B = thoc7F1 + dis3L2ex1 R1 and C = dis3ex1 F1 + dis3ex1 R1 used as a positive control. **(C)** In-depth screening by PCR shows no deletion into *dis3L2* indicating a partial but not complete excision of *P{GS6090}* (lower section in panel **A**). Those stocks showing an absence of PCR product for pairs A and B (in panel **B** of this figure) were screened with primer pairs binding either side of the insertion site both of which are shown in **A**. Primer pair D = thoc7F1 + thoc7 R1, E = dis3L2 F1 + dis3L2 R1 whilst pairs A and B are the same as in panel **B**.



**Figure 7.9: qRT-PCR analysis of the levels of *dis3L2* mRNA in all potential mutant stocks.** Initial screening of all stocks showing a successful excision event to investigate if an excision affected the level of *dis3L2* mRNA. Control = *;;P{GS6090}*.  $n \geq 1$ , error bars represent standard error. Dotted lines show +/- 2 fold change.

## **7.6 Use of clustered regularly interspaced short palindromic repeats (CRISPR) technology to produce a *dis3L2* mutant.**

### **7.6.1 CRISPR as a mutagenesis tool.**

Due to the inefficiency of the P-element mutagenesis experiments it was inappropriate to repeat them any further. Therefore clustered regularly interspaced short palindromic repeats (CRISPR) technology was used in an attempt to produce a *dis3L2* mutant. CRISPR is an extremely efficient method in producing mutants ((Sander and Joung 2014) See Introduction Figure 1.14). In addition to knockout mutants using CRISPR it would also be advantageous to use the technique to create a catalytically dead *dis3L2* mutant which would allow the separation of structural and enzymatic function. Design of the CRISPR experiment and the first steps for the mutant production had been achieved before the writing of this thesis.

CRISPR/Cas9 technology was developed from a bacterial defence against foreign DNA to a highly efficiency system to induce mutations within target genes. It has been developed for use and proved highly successful in multiple organisms such as *Drosophila*, mice and humans (Bassett *et al.* 2013; Ran *et al.* 2013). It is widely accepted to be more specific and efficient development upon the previously used TALENS system. It comprises two components; Cas9 an enzyme which induces double strand breaks (DSB) at targeted regions and a guide RNA (gRNA) which is used to target Cas9 to the site where a DSB is desired. The gRNA itself is made up of a targeting region and a scaffold RNA region which forms stem loop structures which recruit the Cas9 to the targeted site. The 20 nucleotide targeting RNA is specific to the desired gene and must be followed by the PAM recognition motif (NGG). The NGG is essential for Cas9 induced cleavage (Ran *et al.* 2013) with the DSB made 3-4 nucleotides upstream of the PAM. It is critical to ensure off target effects are minimised it has been shown that the final 12 nucleotides of the target sequence are essential for cleavage with a single mismatch enough to prevent the formation of a DSB. However, mismatches in the first 8 seem to have little effect.

### 7.6.2 Identification of target and design of the gRNA.

The key element to a successful CRISPR experiment is in the selection of a suitable gRNA target sequence. Multiple online tools are available to search the given sequence for suitable targets followed by the PAM motif. Two of these tools were used; The Cas9 Target Finder found on the National Institute of Genetics (NIG-FLY) website (<http://www.shigen.nig.ac.jp/fly/nigfly/cas9/cas9TargetFinder.jsp>) and the CRISPR optimal target finder found on the flyCRISPR website. Submission of the coding sequence for both isoforms produced a number of hits for each isoform all of which were present in both isoforms. The two most 5' targets were selected on that basis that targeting the 5' end of the coding sequence is preferable due to the reduced chance of producing a truncated protein of unknown function.

The two selected target sites were validated for their off target potential. Blast searches of both sequences had hits; however, they were not followed by the PAM and therefore would not induce Cas9 cleavage. Additionally the CRISPR optimal target finder contains a tool to search for off target effects. For the most 5' site, using standard settings no off target effects were observed. Therefore the most 5' target sequence of GGAAACGTCAAGCGTCAACG was selected as the gRNA target sequence. The first base of this sequence is at position 42 of *dis3L2-RC* and position 6 of *dis3L2-RA*.

### 7.6.3 Cloning of the gRNA into the pCFD3-dU63gRNA vector.

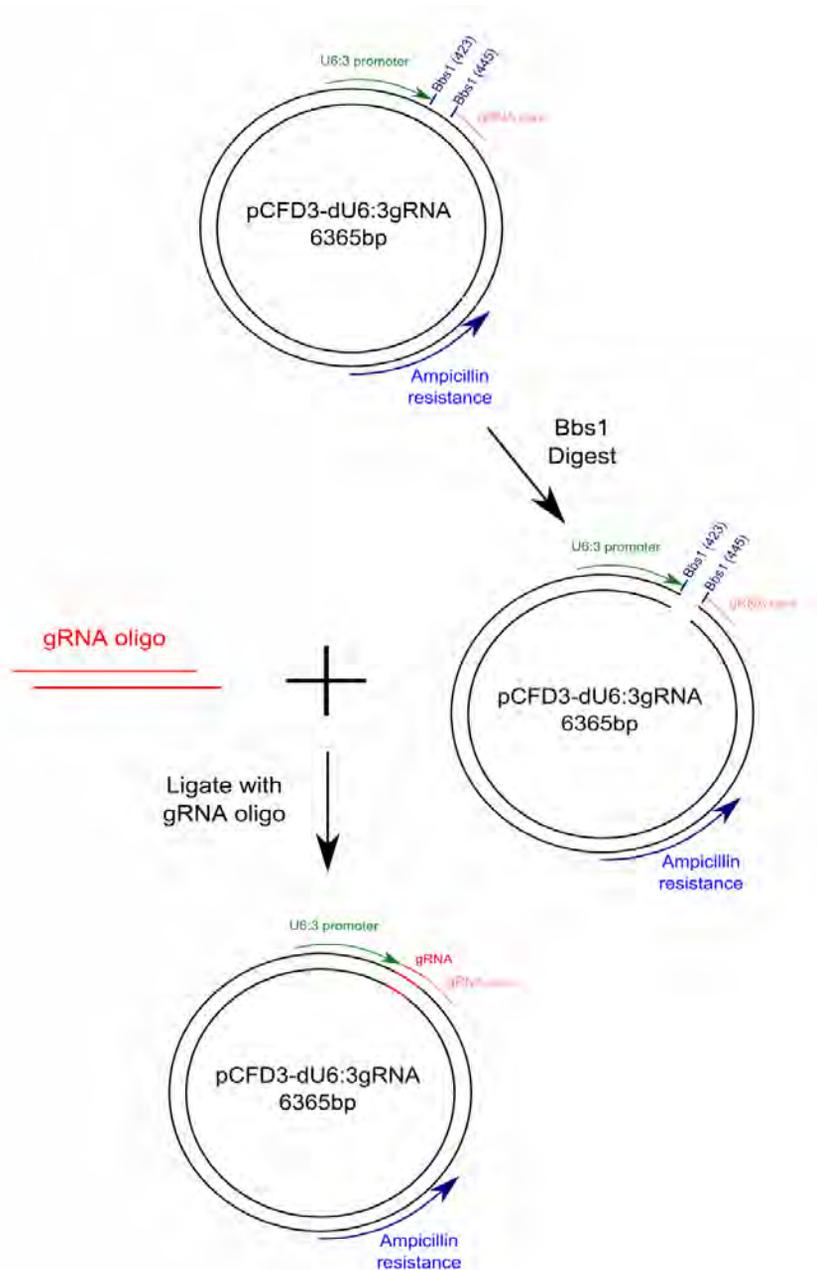
In addition to the 20 nucleotide target sequence the gRNA also consists of the 83 nucleotide scaffold RNA (or gRNA core) sequence which is required to recruit Cas9 to the target site. An RNA 103 nucleotides in length is quite long for *in-vitro* transcription and due to the instability of RNA, its postage and injection is not ideal. Instead of using *in-vitro* transcription the 20 nucleotide gRNA target sequence was cloned into the pCFD3-dU63gRNA (pCFD3); a specific vector used in *Drosophila* CRISPR which already contained the gRNA core sequence. pCFD3 contained a Bbs1 cloning site immediately upstream of the gRNA core sequence allowing the incorporation of the target sequence following Bbs1 digestion of the vector (Figure 7.10). The vector also contains a U6:3 promoter which is ubiquitously expressed in the *Drosophila* embryo. Therefore injection of the vector into a *Drosophila* embryo would result in the expression of the 103 nucleotide gRNA sequence.

To clone the *dis3L2* specific target sequence into the pCFD3-dU63gRNA vector two complementary oligos were annealed; the first containing the target sequence and the other the reverse complement. In addition to the target sequence, vector specific overhangs were incorporated to the 5' end of each oligo to allow ligation into the Bbs1 digested vector (Figure

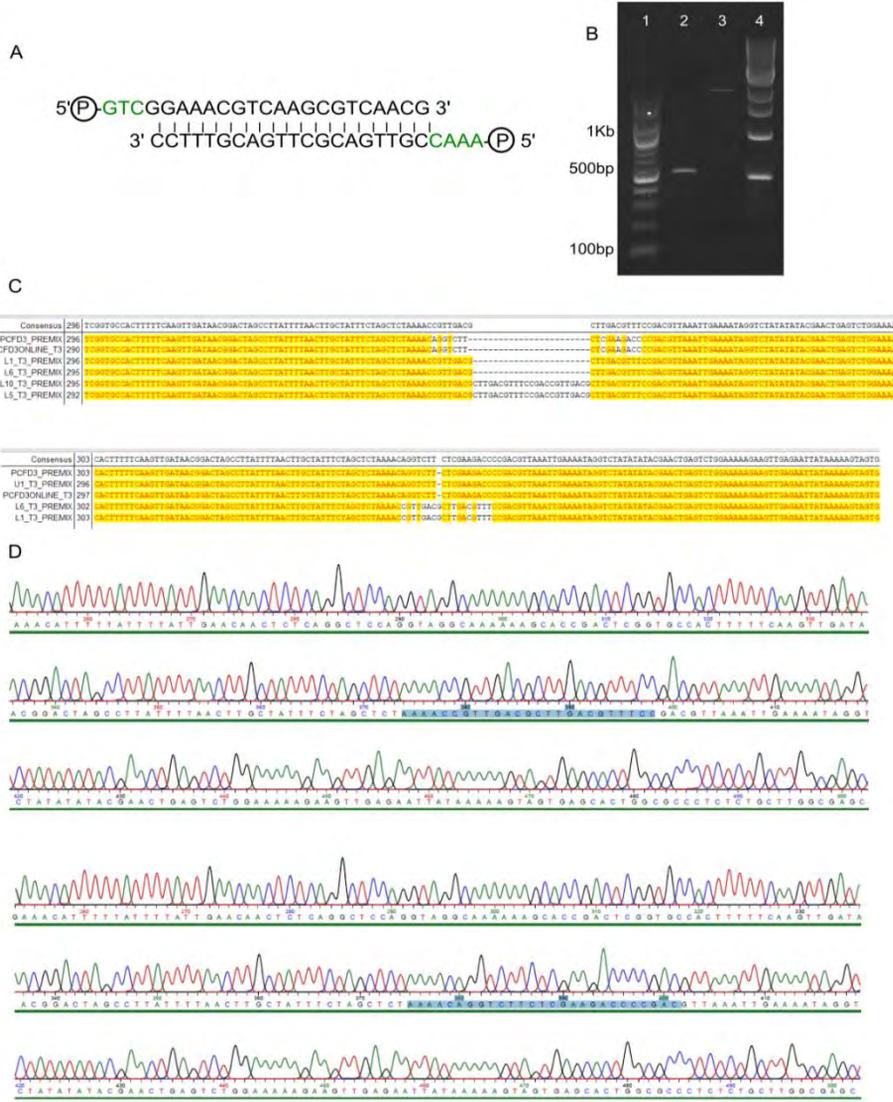
7.11 (green)). During the annealing reaction the oligos were also phosphorylated, using T4 polynucleotide kinase, which was required for the subsequent ligation (Methods 2.3.5). The pCFD3-dU63gRNA vector was digested with Bbs1 for an hour at 37°C. The digested vector was gel purified and 50ng was subsequently incubated with the phosphorylated annealed oligos in the presence of T4 ligase for an hour at room temperature (Figure 7.10). Successful ligation was confirmed by PCR using the gRNA forward oligo in addition to the T3 sequencing primer which has a complementary site in the vector backbone. A product of ~480bp was observed indicating the presence of the insert in the vector (Figure 7.11, lane 2). In contrast no band of the correct size was observed in the undigested control (Figure 7.11, lane 3).

The ligation mix was then transformed into competent DH5 $\alpha$  cells using the method outlined in methods 2.3.5. Briefly, 2 $\mu$ l (6.67ng) of the ligation mix was incubated with competent DH5 $\alpha$  cells for half an hour followed by a 1 minute heat shock and subsequent hour incubation at 37°C in LB-broth. The sample was then plated on agar plates containing 100 $\mu$ g/ml ampicillin for selection of the successful transformants. These plates were incubated overnight at 37°C. Successful antibiotic resistance was confirmed by the absence of colony growth on parallel untransformed control plates. To grow up cultures for DNA extraction individual colonies were picked and incubated overnight in LB-broth containing 100 $\mu$ g/ml ampicillin. Minipreps were performed (achieving consistent yields of ~10 $\mu$ g) on the cultures and the DNA was sent for sequencing to confirm the presence of the insert.

Two selection plates were used with 5 individual colonies selected from each one and subsequent minipreps were performed on each. The DNA from 2 colonies from each plate was sent of sequencing together with the initial pCFD3 vector which served as a control. The T3 sequencing primer was used to prime to sequencing reaction. Sequencing revealed all the selected colonies had successfully taken up the pCDF3 vector containing the gRNA targeting sequence. DNA from colony 1 was purified and sent for embryo injection at BestGene. This is as far as the CRISPR mutagenesis experiments got during the writing of this thesis. Subsequent work is required to identify any *dis3L2* mutant in the resulting progeny.



**Figure 7.10: Integration of the gRNA sequence into the pCFD3-dU6:3gRNA (pCFD3) vector.** The pCFD3 vector was digested with Bbs1 resulting in the removal of the Bbs1 spacer located between the U6:3 (green) promoter and the gRNA core sequence (pink). The phosphorylated double stranded gRNA oligo (red) was ligated into the region vacated by the excised Bbs1 spacer. pCFD3 also contains ampicillin resistance (blue) which was used to positively select for those bacteria that had been successfully transformed with the vector.



**Figure 7.11: Preparation of the injection vector for CRISPR. (A)** The phosphorylated double stranded oligo to be cloned into the pCFD3 injection vector (using the method outlined in Figure 7.10). **(B)** Successful re-ligation of pCFD3 with the gRNA insert. PCR performed with an insert specific primer and a reverse primer in the backbone which only gives a 500bp product if the ligation is successful. Lane 1 = 100bp ladder, lane 2 = re-ligated vector with diagnostic primer pair, lane 3 = undigested vector with diagnostic primer pair, lane 4 = 1Kb ladder. **(C)** Sequencing of selected colonies shows the successful incorporation of the gRNA (L1 and L6). Two colonies, L5 and L10, showed an insertion of duplexed insert with a duplication of the gRNA and were therefore discarded. Undigested pCFD3 (U1) was used as a control in the sequencing and was identical to the sequence published online (PCFD3ONLINE). **(D)** Sequencing peaks for L1 (upper) and control (lower) shows strong single peaks for each base call. Insert sequence is highlighted in the upper panel whilst the sequence replaced is highlighted in the bottom panel.

## 7.7 Design and production of a Dis3L2 antibody.

Many molecular biology techniques such as immunocytochemistry and western blotting are dependent on a good antibody. However, an antibody to *Drosophila melanogaster* Dis3L2 was not commercially available and therefore these techniques could not be used to detect the levels or localisation of *Drosophila* Dis3L2. For example, although an 80% knockdown has been shown at the RNA level with multiple RNAi constructs driven by multiple GAL4 drivers it has not been possible to show that this leads to a subsequent knockdown at the protein level. Additionally, Dis3L2 has been shown to be localised strictly to the cytoplasm in yeast, mouse and human cells, however this has not been confirmed in *Drosophila*; although it can be assumed due to the extensive homology. Furthermore, to take the Dis3L2 target work further it would be crucial to perform cross-linked immune precipitation (CLIP) experiments followed by either RNA sequencing or qRT-PCR to identify those RNAs that are directly bound by Dis3L2. However, an antibody is also required for this.

Due to the extensive conservation of Dis3L2 throughout higher eukaryotes two different human anti-Dis3L2 antibodies were tested on *Drosophila* samples. The two antibodies were selected due to the epitope being in a region of greater conservation. These were obtained from Sigma-Aldrich (HPA0357497) and ThermoFisher Scientific (PA5-31273). A western blot would be the primary use of the antibodies within this work and therefore it was the technique used to trial the two human anti-Dis3L2 antibodies. Protein samples were prepared from whole flies (Oregon R) in duplicate whilst human RPMI-8226 (RPMI cells were used as a positive control).

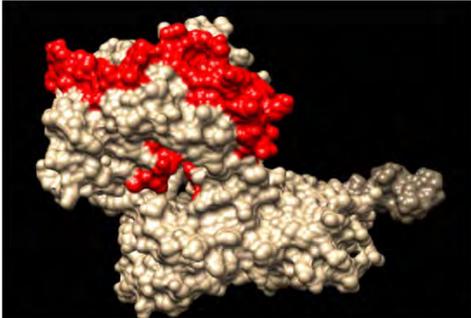
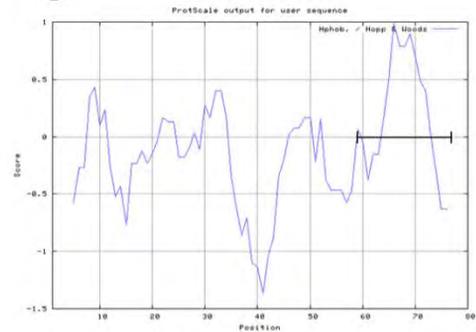
On incubation with the ThermoFisher Scientific antibody a ladder effect was observed in the RPMI cells showing a lack of specificity. A band was produced at the correct molecular weight (100kDa) however the level of non-specificity would make it unsuitable for work in human cells. Bands of a very high molecular weight (245kDa) are observed in both the *Drosophila* samples but these are much too large to be Dis3L2 and therefore are likely to be non-specific, similar to those observed in the RPMI cells. The Sigma-Aldrich antibody however, gave a single, specific band at 100kDa corresponding to human Dis3L2 showing it is a more suitable antibody than the one purchased from ThermoFisher Scientific. Despite the high specificity within the human sample unfortunately no band was observed in either *Drosophila* sample (results summarised in Figure 7.12A). This therefore suggests that the human Dis3L2 antibodies, although the structure is well conserved, were not suitable for use in *Drosophila*.

**A**

Company (Identifier)	Target sequence	Result
Sigma-Aldrich (HPA0357497)	SELFKRKYALFSPSDHRVPRIYVPLKDCPQDFV ARPKDYANTLFICRIVDWKEDCNFALGQLAK SLGQAGEIEPETEGILTEYGVDFSDFFS (amino acids 249-359)	Does not work for <i>Drosophila</i> samples but a single band of the correct MW is observed in the RPMI control sample.
Thermo Scientific (PA5-31273)	Target region is defined as being within amino acids 288-528.	Does not work for <i>Drosophila</i> samples and gives a 'ladder-like' result in the RPMI control sample.

**B**

	238	240	250	260	270	280	290	300	310	320	330									
DrosSeqPA	1	-----	QLYKRPYVLE	YVPMVYVFKDA	AAHI	GNQ	QIDVSI	LLYL	AHL	ETC	CNGHC	IL	JOPV	QRV	NI	DEL	IL	ENGL	-----	
DrosSeqPC	1	-----	QLYKRPYVLE	YVPMVYVFKDA	AAHI	GN	QIDVSI	LLYL	AHL	ETC	CNGHC	IL	JOPV	QRV	NI	DEL	IL	ENGL	-----	
RNase2	171	-----	EKEAPDGVATEML	EGVYED	TALDFVTID	ASTER	MDD	AL	AK	LPD	KLQ	LIV	AT	PTAW	EGSK	AKA	IRAE	TV	PGFNIPHL	
RNaseR	235	YIMPQAVEQQVAGL	KEEV	EEA	AGVDRDL	PLVTID	EDAR	DFDD	VY	CEK	RRGG	NRL	WVA	IA	SYV	VPSTP	LR	EA	NRGTSVYV	PSQVIPHML
Sigma	1	-----	SELFKR	KYALFSPS	DRV	VPRIYV	PLKDCP	QDFV	ARPK	DYANTL	FICR	IVDWK	EDCN	FALG	QLAK	SLGQ	AGEI	EPETE	GILTEY	GVDFSDFFS

**C****D**

**Figure 7.12: Design of an anti-Dis3L2 peptide antibody. (A)** Summary of the results of the attempt to use antibodies against human Dis3L2 to detect Dis3L2 in *Drosophila* samples. **(B)** Alignment of the sequence used to raise the Sigma-Aldrich human anti-Dis3L2 antibody with both *Drosophila* Dis3L2 isoforms and bacterial homologues. The sequence selected for the final peptide is highlighted with a black box. **(C)** The 3D structure of the bacterial homologue RNase II. The sequence homologous to the sequence used to produce the Sigma-Aldrich antibody is highlighted in red showing its major presence on the surface of the protein. **(D)** A Hopp and Woods hydrophobicity plot of the *Drosophila* sequence homologous to the sequence used to produce the Sigma-Aldrich antibody. The most hydrophilic region (marked with a black line) was selected for the final peptide sequence.

As a result of the unsuccessful trails of the human antibodies an attempt was made to design and produce an antibody to *Drosophila* Dis3L2. Due to its high level of specificity in the human RPMI samples the antibody from Sigma-Aldrich was used as a template. The peptide sequence to which the Sigma-Aldrich antibody is designed is encoded by exons 7 and 8 of human Dis3L2 and forms part of the highly conserved RNB domain. Conducting a BLAST search of the peptide used to design the Sigma-Aldrich antibody across the *Drosophila* genome results in the identification of a homologous region in *dis3L2* from amino acid position 440-519 which is identical in both Dis3L2 isoforms. The identified region contains 30.6% identity with the Sigma-Aldrich sequence with 51.8% amino acids showing the same 'nature' referred to as 'positive' hits. To confirm this region the *Drosophila* peptide sequences for both isoforms were aligned to the Sigma-Aldrich peptide sequence using the AlignX tool in Vector NTI (Figure 7.12B). The same region was shown to align indicating it was the most appropriate region to design the Dis3L2 antibody against.

In addition to identifying the homologous sequence it is critical that an antibody epitope is situated on the outside of the protein structure. Therefore to predict the location of the identified sequence the bacterial homologue, RNase II, was used as its structure has been solved previously. The bacterial amino acid sequence was aligned against the Sigma-Aldrich peptide and the two *Drosophila* Dis3L2 isoforms. This identified a homologous region, although the homology is less extensive than the human-*Drosophila* which would be expected. Using Chimera this sequence was then modelled onto the solved structure of RNase II which revealed its presence largely on the surface of the protein (Figure 7.12C) which would be optimal for an antibody epitope. This is modelled on predicted homologous regions; however, to assess the likelihood of the *Drosophila* peptide of being on the surface of the protein a Hopp and Woods hydrophobicity plot was performed. This showed a strong hydrophilic region which implies a surface localisation. Due to 20 amino acids being optimal for peptide design the last 22 amino acids were selected as they made up the most hydrophilic region (Figure 7.12D) within the homologous sequence and therefore the most likely to be on the surface of the Dis3L2 protein. The amino acid sequence selected was: QPVGRVGNLDDELKAILFHNGL.

The selected sequence was provided to Altabioscience for peptide synthesis and subsequent immunisation in sheep. 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> bleeds in addition to the affinity purified antibody were received for testing on *Drosophila* samples. The affinity purified antibody was trialed at 4 dilutions (1:500, 1:1000, 1:1500, 1:2000) on protein extracted from whole flies with and without ubiquitous knockdown of Dis3L2. Ubiquitous knockdown of *dis3L2* was achieved by driving *UAS-dis3L2<sup>RNAi</sup>* with *tubulin-GAL4* which has been shown previously to result in a knockdown efficiency of 90%. A product of 118kDa would be expected in the control flies

whilst a much weaker band, if any, should be observed in the knockdown flies. Unfortunately a high degree of non-specificity was observed for all samples and dilutions as shown by a ladder-like effect (Figure 7.13). In addition no clear band was observed around 118kDa in the control samples. Together these make the antibody unsuitable for use and therefore thus far this aim has not been achieved.

## 7.8 Production of a polyclonal anti-Dis3L2 antibody

Following the failure of the previous peptide antibody production a new approach was taken to produce a polyclonal antibody against recombinant Dis3L2. This required the *dis3L2* cDNA within an expression vector allowing the production of the full length Dis3L2 peptide. As mentioned in previous chapters there are two isoforms of Dis3L2, PA and PC. PC is 12 amino acids larger than PA resulting in a protein 1.4kDa larger. cDNA for both isoforms were sourced and ordered from the Berkeley Drosophila Genome Project (BDGP). The selected cDNA clones were BS10470 (*dis3L2-RA*) and BS27503 (*dis3L2-RC*) both of which were shown to be full length and contain the native stop codon. Each cDNA clone was transformed into competent DH5 $\alpha$  cells with varying efficiencies. Successfully transformed cells were selected using the ampicillin resistance gene within the clone vector. BS27503 plates showed a reasonable number of transformed colonies across multiple plates indicating a successful transformation. However, only one plate showed any colony growth for BS10470 transformed cells indicating less efficient transformation. In addition to the *dis3L2* cDNA clones, a 'test' clone was also sent which transformed with slightly greater efficiency than the BS27503; these colonies would be used in future experiments as controls.

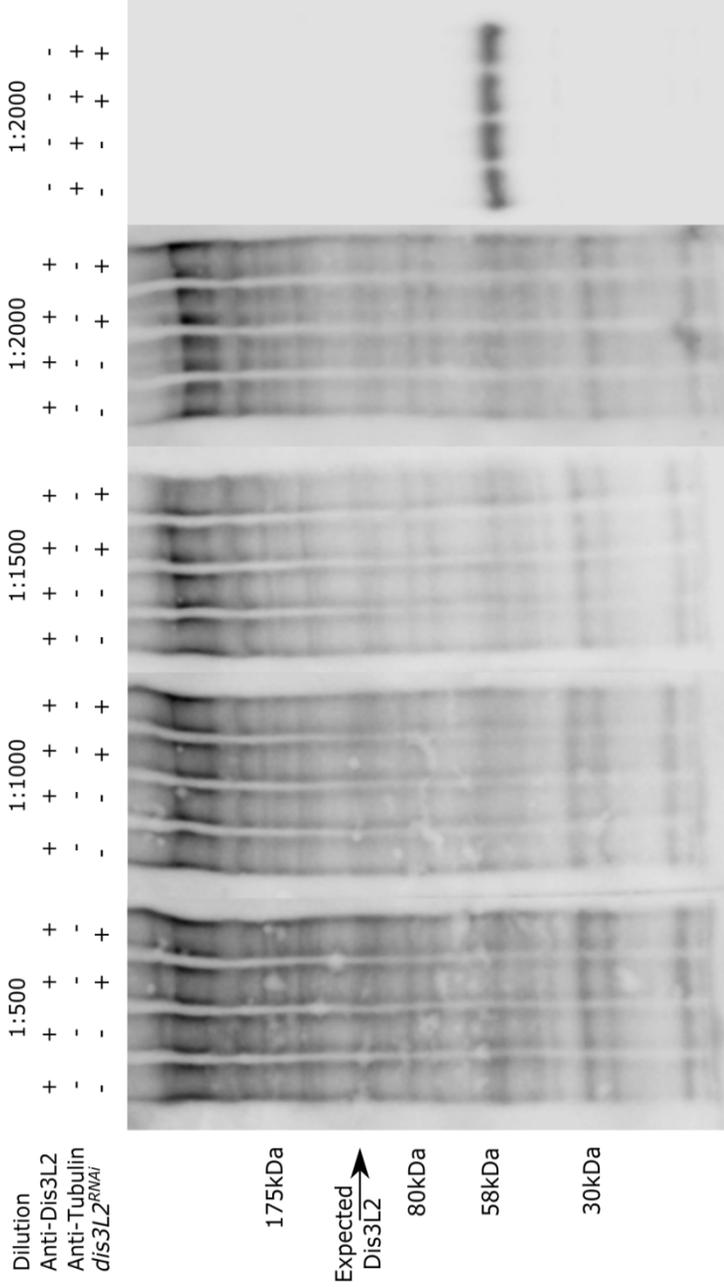
10 BS27503 colonies were selected along with 5 colonies for BS10470 and 'test'. Only those colonies that were individual and not surrounded by other colonies were selected to minimise the risk of contamination between individual colonies. Each colony was incubated overnight (in 5ml of LB broth + 100 $\mu$ g/ml ampicillin) at 37°C. DNA from 1.5ml of each colony was extracted using a Qiagen MiniPrep kit (see methods section 2.3.5.7). The minipreps for each BS27503 and 'test' colonies were successful with DNA yields of ~10 $\mu$ g. However, only BS10470 colony 3 showed a successful miniprep with a yield of 10.2 $\mu$ g; the other 4 showed yields of <250ng and so were discarded.

To ensure the transformed cells contained the correct *dis3L2* cDNA two approaches were used; a restriction digest and PCR. First a diagnostic restriction digest was performed making use of the fact that two Hpa1 cut sites are found in the sequences of both *dis3L2* isoforms. Cut sites were identified and selected using the online tool 'restriction mapper'

(<http://www.restrictionmapper.org/>). The coding sequence for each *dis3L2* isoform was submitted and searched for all cut sites. Hpa1 was selected as it cuts at two sites within the *dis3L2* coding sequence therefore allows a digest with only a single restriction enzyme. Hpa1 cuts at positions 514 and 931 in RA whilst it cuts at 550 and 967 in RC producing a single product 417bp in size. Hpa1 does not cut anywhere else in the plasmid and therefore any product should be specific to the *dis3L2* sequence. 500ng of purified plasmid DNA was used in each restriction digest which were performed for 1 hour at 37°C. All BS27503 colonies were tested alongside 2 'test' colonies and the single BS10470 colony. Undigested samples were run as negative controls where the purified DNA was incubated in the digest mix in the absence of Hpa1 for 1 hour at 37°C.

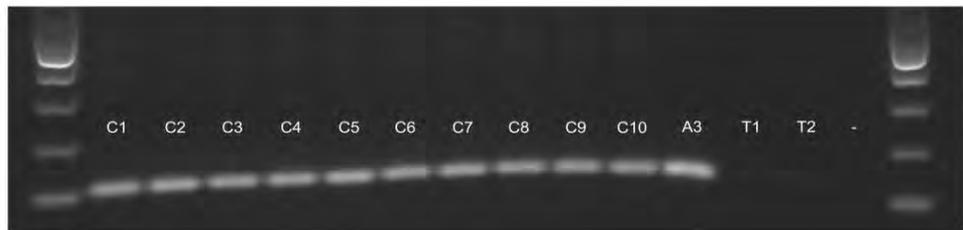
All BS27503 and BS10470 colonies showed a digestion product ~400bp representing a successful cleavage and therefore indicating the presence of the *dis3L2* sequence within the plasmid. No cleavage product was observed for the undigested control showing the products in the digested lanes are indeed specific to the Hpa1 cut sites in the *dis3L2* sequence. Finally, neither the digested nor the undigested 'test' colonies showed digestion products giving further evidence for the observed products being specific for the *dis3L2* sequence.

The restriction digest confirmed the presence of the *dis3L2* sequence within the transformed cells, however, due to cut sites being present in both isoforms it was unable to show if the two isoforms were as expected. Therefore PCR was used with primers spanning the differential region between the two isoforms (Chapter 4, Figure 4.3A). A PCR was performed using 1µl of the same purified DNA as was used for the restriction digest. Unfortunately the PCR products for both BS10470 and BS27503 were the same size, the expected size for *dis3L2-RC* (Figure 7.14A). However, the presence of products further confirms the presence of a *dis3L2* sequence in the transformed cells. The specificity was shown by the absence of PCR product in the 'test' samples or the no template control (Figure 7.14A). The cDNA clone BS10470 was sold as *dis3L2-RA*, however, the PCR results indicate it actually contained the sequence for *dis3L2-RC*. To confirm the sequence of both the cDNA clones selected colonies were sent for DNA sequencing. BS27503 colony 4 was selected as it showed the greatest DNA yield, whilst BS10470 colony 3 was the only remaining colony for that clone. Upon sequencing it became clear that the clones for both BS27503 and BS10470 were in fact the same sequence, that of *dis3L2-RC*. The unique 59bp sequence in *dis3L2-RA* was not present in either sample (Figure 7.14B). Therefore, unfortunately it was not possible to express both isoforms. However, the sequencing does confirm the sequence is correct for the *dis3L2-RC* clones.



**Figure 7.13: Testing the anti-Dis3L2 peptide antibody.** A serial dilution of anti-Dis3L2 antibody from 1:500 to 1:2000 was used to test the specificity of the new antibody. Anti-Tubulin (final panel) was used as loading control at 1:2000. The antibody showed a high degree of non-specificity with no obvious bands at the correct size showing a difference between the control and *dis3L2<sup>RNAI</sup>* samples. Control (*dis3L2<sup>RNAI</sup>* -) genotype = ;*tub-GAL4/TM6, dis3L2<sup>RNAI</sup>* genotype = ; UAS-*dis3L2<sup>RNAI</sup>*/+ ; *tub-GAL4/+*. All samples were produced from 2 male flies.

**A**



**B**

	46	50	60	70	80	90	100	110	120	130	140								
Dis3L2RA	46	CTTACCC	TTTATA	TC	CCGGCC	ATAG	CAAGT	G	TAGTGCCG	TACACT	TTCCGGAC	ATCAGCGAATCC	CATTAACTAAGTGC	ACTTCTTGCAG	AA	TG	CG	AA	CG
Dis3L2RC	46	CTTACCC	TTTATA	TC	CCGGCC	ATAG	CAAGT	G	TAGTGCCG	TACACT	TTCCGGAC	ATCAGCGAATCC	CATTAACTAAGTGC	ACTTCTTGCAG	AA	TG	CG	AA	CG
RADis3L2Ex1R1	1	---	CCCTTT	AT	TC	CCGGCC	ATAG	CAAGT	---	---	---	---	---	---	AA	TG	CG	AA	CG
RCDis3L2ex1R1	1	---	CCCTTT	AT	TC	CCGGCC	ATAG	CAAGT	---	---	---	---	---	---	AA	TG	CG	AA	CG

**Figure 7.14: Transformation of *dis3L2* cDNA in preparation for antibody production. (A)** PCR using primers to differentiate between the two isoforms. T1 and T2 are test plasmids which do not contain the *dis3L2* cDNA sequence which were used as negative controls. A different sized product should be observed for isoforms A (A3) and C (C1-10), however, they are the same. This was confirmed by sequencing **(B)** which showed they were all isoform C.

## 7.9 Chapter Summary

This chapter has outlined the attempts to produce a Dis3L2 mutant and antibody; two tools that would be incredibly useful in dissecting the role of Dis3L2 in the control of tissue growth and development. Unfortunately, to date efforts have been unsuccessful in both.

The production of a mutant using P-element mutagenesis is dependent on random events and therefore is based on chance. Regrettably many problems were encountered along the way; these have been detailed throughout this chapter. In addition to complete absence of a P-element from one stock and the lethality causing factor, the P-element GS6090 showed a very low frequency of hop out. This low frequency was observed in both mutagenesis experiments using the stock meaning the final number of potential mutants was limited. Due to the lack of success using this method the CRISPR/Cas9 technique is currently being used; this has been successfully used to induce mutations in a number of organisms including *Drosophila* (Bassett *et al.* 2013).

Similarly, the production of a Dis3L2 antibody has thus far proved unsuccessful. A thorough and logical approach was taken in designing a peptide antibody; however, the resulting antibody showed extensive non-specificity and was therefore not suitable for use. Steps have been taken to produce a polyclonal antibody to the recombinant protein which may prove more successful. However, to date the aims set out at the beginning of this chapter have not been achieved.

## 7.10 Chapter Discussion

### 7.10.1 Why were the mutagenesis experiments unsuccessful?

The experiments outlined in this chapter set out to create *dis3L2* mutants to use to investigate the role of Dis3L2 in cell proliferation and tissue growth. The phenotypic and genetic analysis presented in this thesis uses RNAi due to the absence of the *dis3L2* mutant. Unfortunately all mutagenesis attempts proved unsuccessful and therefore the main aim of this chapter was not fulfilled. There are multiple potential reasons for the failure of the mutagenesis experiments. The major limiting factor was the low efficiency of the P-element excision which immediately restricts the number of potential deletions. The efficiency was consistent for each repeat being ~6%. This is much lower than the *pacman* mutagenesis experiments performed previously in the Newbury lab which were closer to 40% (Jones 2011). To account for the low excision efficiency the initial crosses were set up on a large scale, however, this still proved to be too small to produce a mutant. P-elements can excise at different efficiencies and it appears that GS6090 used in these experiments was a highly inefficient example. Therefore with the low hop-out frequency the chances of generating a *dis3L2* mutant were greatly reduced.

In addition to the inefficient hop out the initial mutagenesis screen with GS6090 was affected by the lethality causing factor on the P-element containing chromosome. Regrettably, assumptions were made that the P-element insertion itself was the cause of the lethality; however, attempts should have been made to confirm this. This was only performed after the conclusion of the 4 generation cross scheme when it was noted that even those stocks showing a clean excision were also homozygous lethal. As a result all the stocks generated during this cross scheme were discarded and therefore any potential mutant would have been lost. It would have been possible to repair the chromosome of the potential mutants and screen for a mutation but this would have brought with it the possibility of losing the unique mutations and therefore the decision was made to repair the chromosome in the P-element containing stock and repeat the experiments again.

Successful chromosome repair was achieved using homologous recombination which showed a repair rate of 25%. The newly repaired P-element containing chromosome was then used in the subsequent rounds of mutagenesis but unfortunately, as outlined above, these were also unsuccessful in the generation of a *dis3L2* mutant. Other P-element stocks are available with insertion sites in close proximity to *dis3L2*, however, both *P{GSV6}GS16356* and *P{GSV6}GS13294* are inserted within the neighbouring gene *thoc7*. Therefore any deletion into *dis3L2* resulting from their excision would affect *thoc7* as well as *dis3L2* and therefore could

prove problematic. As a result CRISPR was used as an alternative method of generating the highly desirable *dis3L2* mutant.

### **7.10.2 Using CRISPR should be a more efficient way of creating a *dis3L2* mutant.**

The attempts to produce a *dis3L2* mutant using P-element mutagenesis proved unsuccessful therefore an alternative approach was begun at the end of this PhD. Although, like P-element mutagenesis, CRISPR induced mutations are random it has been shown to be highly efficient (Bassett *et al.* 2013; Ran *et al.* 2013). CRISPR induced mutations rely on the error prone mechanism of DNA repair known as non-homologous end joining (NHEJ). Due to the CRISPR enzyme Cas9 inducing double strand breaks the cell is unable to use homologous recombination to repair the damage. Instead NHEJ is used which often results in insertions or deletions (indels) which can affect the normal translation of the protein resulting in hypomorphic or null mutants. The efficiency of CRISPR in *Drosophila melanogaster* has been shown to range from 4% to 88% (Bassett *et al.* 2013). The efficiencies are dependent upon the gRNA targeting sequence; however, using this method should greatly increase the chance of inducing a mutation in *dis3L2* compared to the previous P-element mutagenesis methods.

The gRNA was designed to maximise the chance of mutation together with minimising the chances of off-target effects. The selected guide sequence was shown to have no potential off-target effects by an independent BLAST search in addition to an inbuilt tool on the CRISPR optimal target finder site. The selected sequence was the most 5' option being only 6bp into the coding sequence of isoform RA. This should minimise the chance of a truncated protein being produced which may have a damaging function. The cloning strategy was also successful in producing an insertion vector which contained the 20nt target sequence cloned immediately downstream of the U6 promoter and immediately upstream of the scaffold RNA. Using this method allows the whole gRNA sequence to be made *in-vivo* which removes the issue of the stability of RNAs. The injection vector was sent to BestGene for embryo injection and during the writing of this thesis the screening for the identification of a mutant had not been performed.

### **7.10.3 Limitations of the peptide antibody design.**

A major limiting factor throughout this project has been the absence of an anti-Dis3L2 antibody. The lack of such a valuable reagent meant that Western blotting, localisation and immunoprecipitation studies were not possible. To overcome this, an attempt was made to produce a peptide antibody. Despite attempts to maximise the likelihood of success this proved unsuccessful which could have been for a number of reasons. The antibody was designed to a section of a sequence homologous to that recognised by a human anti-Dis3L2 antibody which was shown to have high specificity on human samples. This seemed the most appropriate approach; however, the sequence given for the human antibody was longer than was suitable for the antibody production. Therefore only a section of the homologous sequence could be used in the *Drosophila* antibody design. This small sequence may have been different to the sequence actually used in the human antibody. This, taken with the fact that the sequences were not completely homologous meant that the epitope to which the antibody would have been made was likely to differ.

The major issue with the design of an antibody against a small peptide is that the epitope that amino acid sequence forms may be completely different to that it forms in the whole protein or even a larger sequence due to the interaction of the amino acid side chains. If the structure formed by the short peptide differs from the structure in the total protein structure then the antibody is unlikely to work. Unfortunately it looks as if this was the case as the resulting tests showed a high level on non-specificity together with no evident band that showed a difference between the control and *dis3L2* knockdown samples.

To overcome the issue of the differential epitope formation the decision was made to produce an antibody designed against the purified protein. The initial steps have been performed for this where the cDNA sequence of *dis3L2* has been transformed into competent bacteria and set to a collaborating lab in Portugal for protein purification. The process of Dis3L2 purification and antibody production was in progress during the writing of this thesis.

### **7.11 Acknowledgements**

BestGene inc. performed the embryonic injection for the CRISPR experiments. The peptide to which the anti-Dis3L2 antibody was synthesised by Altabioscience who also carried out the immunisation. Finally the purification of recombinant Dis3L2 for the creation on the new anti-Dis3L2 antibody was performed by Dr. Rute Matos and Dr Sandra Viegas in the research lab of our collaborator Prof. Cecilia Arriano in ITQB, Lisbon.

## Chapter 8: Discussion

### 8.1 Summary of main findings.

#### 8.1.1 Dis3 is essential for cell viability.

The work presented in Chapter 3 showed that the 3'-5' exoribonuclease is required for *Drosophila* viability. Knockdown of Dis3 throughout the organism using RNA interference resulted in complete lethality during the early larval stages showing its activity is critical. Further investigation using a specific model tissue, the wing imaginal disc, demonstrated that the loss of Dis3 results in a specific increase in apoptosis. This apoptosis is likely to result from massive cellular dysregulation as Dis3 has been previously shown to be involved in the regulation of RNA species which are in turn involved in a variety of cellular pathways (Mitchell *et al.* 1997; Wyers *et al.* 2005; Preker *et al.* 2008; Gudipati *et al.* 2012; Schneider *et al.* 2012).

#### 8.1.2 Dis3 plays a role in miRNA regulation.

In addition to the above section Chapter 3 also began to elucidate a role for Dis3 in the regulation of miRNA stability. Using miRNA-seq in a global, unbiased approach a set of Dis3-sensitive miRNAs were identified. This pool only represented a small proportion of the miRNAs expressed in the wing disc showing an element of specificity behind Dis3-mediated regulation. Within this pool two miRNAs were shown to be excellent candidates for direct Dis3 regulation. *miR-252-5p* expression was shown to increase at the post-transcriptional level following Dis3 knockdown therefore indicating Dis3 may be responsible for its degradation. An alternative role for Dis3 in *miR-982* biogenesis was also presented; however, additional work would be required to further characterise this.

#### 8.1.3 A novel role for Dis3L2 in regulating proliferation within the wing disc.

The experiments in Chapter 4 aimed to explore a potential role for Dis3L2 in regulating tissue growth in an attempt to illuminate the mechanisms behind the human overgrowth condition; Perlman Syndrome. This work uncovered a role for Dis3L2 in the regulation of developmental proliferation within the wing imaginal disc. Knockdown of *dis3L2* within the wing disc resulted in the overgrowth of the adult wing which resulted from an increase in cell number. This was subsequently shown to result from increased proliferation in the wing imaginal disc. This work therefore identified a conserved function for Dis3L2 in controlling tissue growth. RNA-seq was performed in Chapter 5 in an attempt to identify differentially expressed transcripts which may

contribute to the overgrowth phenotype. 2 potential Dis3L2 targets, *pyrexia* and *CG2678*, were identified but further work is required to determine a mechanism by which Dis3L2 regulates proliferation and tissue growth.

#### **8.1.4 Dis3L2 and Pacman affect opposing but related processes.**

Dis3L2 has been shown in this work to result in increased proliferation whilst previous work in the Newbury lab identified a role for the 5'-3' exoribonuclease Pacman in apoptosis. Therefore the aim of Chapter 6 was to investigate this further in terms of pathway specificity. Using RNA interference it was shown that the knockdown of both Pacman and Dis3L2 results in lethality which is in line with previous findings (Malecki *et al.* 2013). Interestingly, using *pacman* mutants and wing disc specific RNAi to *dis3L2* it was shown that their simultaneous inhibition in the wing imaginal disc rescues the tissue growth phenotypes observed when either is present independently. This shows, for the first time, that the 5'-3' and 3'-5' cytoplasmic decay pathways can function to regulate specific cellular processes. Comparative analysis of sequencing data showed examples of transcripts regulated specifically by each pathway which, given time, would be very interesting to investigate further.

## **8.2 Does Dis3 play a role in the regulation of miRNA stability?**

Previous work has identified a role for Dis3 in the regulation of a variety of RNA substrates such as 5.8S rRNA (Mitchell *et al.* 1997), pre-tRNA, pre-mRNA (Kadaba *et al.* 2004; Schneider *et al.* 2007; Gudipati *et al.* 2012) together with other non-coding RNA species such as snRNAs, snoRNAs, PROMPTs and cryptic unstable transcripts (CUTs) (Preker *et al.* 2008; Schaeffer *et al.* 2012; Tomecki *et al.* 2014). However, a role for Dis3 in regulating mature miRNA stability had not previously been explored. Therefore the work presented in Chapter 3 outlined a novel role for Dis3 in regulating the stability of specific miRNAs.

### **8.2.1 The function of Dis3 on miRNAs is specific.**

The role of Dis3 in miRNA regulation had not previously been explored. miRNA-seq was used in a global, unbiased approach to investigate the effect of Dis3 knockdown on the miRNA expression profile of the wing imaginal discs. Of the 109 miRNAs confidently detected in the wing imaginal disc only 16.5 percent showed a change in expression of more than 2 fold. This therefore showed that whilst a proportion of miRNAs were sensitive to Dis3 activity, it does not play a global role in regulating miRNA stability but instead is more likely to have a more specific, targeted role.

This work identified a total of 18 miRNAs that changed in expression (>2 fold) following Dis3 knockdown in the wing imaginal disc. However, it is possible that miRNAs were missed in this

study due to the low number of miRNA mapped reads. The miRNAs that showed differential expression had relatively low read counts and were therefore within the low expression bracket which would be affected by the 2S rRNA contamination. It is therefore possible that Dis3 regulates additional miRNAs that were expressed at levels too low for the achieved read coverage to reliably identify. Either way, the work has shown that only a subset of miRNAs are sensitive to Dis3 activity and therefore elements of target selection and specificity must be present to allow this.

### **8.2.2 A role for Dis3 in regulating mature miRNA stability.**

The work in Chapter 3 identified *miR-252-5p* as a strong candidate as a mature miRNA directly regulated by Dis3. In *Drosophila* Dis3 associates with both the nuclear and cytoplasmic exosome (Graham *et al.* 2006) the activities of which required separation to validate this work. Knockdown of Rrp40, an exosome subunit predominantly in the cytoplasm in *Drosophila* (Graham *et al.* 2006), allowed the cytoplasmic activity of Dis3 to be inhibited whilst retaining a large proportion of nuclear activity. Knockdown of Rrp40 resulted in a comparable increase in *miR-252-5p* expression indicating the cytoplasmic activity of Dis3 was responsible, as would be expected. This was further characterised by assessing the level at which the expression increase was driven. *miR-252-5p* was shown to increase at the post-transcriptional level as the precursor remained unchanged. This indicated that the observed increase in *miR-252-5p* expression was direct and not simply due to an off target effect of Dis3 resulting in its increased biogenesis and transcription.

This work was the first to reveal a role for Dis3 in the regulation of mature miRNA stability, a field which has been largely overlooked (reviewed in (Towler *et al.* 2015)). It is however unlikely that *miR-252-5p* is the only mature miRNA that is directly regulated by Dis3. As mentioned in section 8.2.1 some Dis3-sensitive miRNAs were likely to be missed as a result of low miRNA read coverage. Additionally, Dis3 could only be knocked down in the wing pouch of the disc as its knockdown across the disc resulted in lethality and therefore other Dis3-sensitive miRNAs may also have been missed.

The miRNA-seq data suggests that only a subset of *Drosophila* miRNAs are sensitive to the activity of Dis3. This infers mechanisms of specificity that may come through physical association between Dis3 and targeting co-factors or through nucleotide tagging of the miRNAs. Terminal tagging is known to promote the activity of the nuclear exosome as TRAMP mediated terminal adenylation targets aberrant 23S rRNA intermediates (LaCava *et al.* 2005), unspliced pre-tRNAs (Wlotzka *et al.* 2011) and CUTs (Wyers *et al.* 2005) for exosome-mediated decay (reviewed in (Schmidt and Butler 2013)). A similar mechanism may exist in the

cytoplasm to target Dis3/the exosome to specific miRNAs; such modifications of the 3' ends of mature miRNAs have previously been shown to affect their stability (Kato *et al.* 2009). It is possible that *miR-252-5p* is terminally tagged resulting in the specific recruitment of Dis3. This tagging could be performed by nucleotidyl transferases such as Gld2 or Tailor which has recently been identified to be a 3' uridyl transferase in *Drosophila* (Reimao-Pinto *et al.* 2015). Unfortunately due to the low read count it was not possible to make any conclusions about terminal tagging of *miR-252-5p*. The determination of *miR-252-5p* terminal tagging was also made difficult as the 2 nucleotides immediately downstream of the 5' end of *miR-252-5p* in the pre-miRNA are uridines. This therefore made it impossible to differentiate between template and non-template additions. It would only become clear if the uridine tail consisted of more than 2 Us. Despite this the read sequences still underwent initial analysis (presented in Figure 3.20). There was a possibility of terminal tagging but to draw any confident conclusions the experiment would require repeating with greater depth and to include longer RNAs (up to 30nt).

In addition to nucleotide tagging another possibility for Dis3 targeting could be through intrinsic instability elements. For example, AU rich elements are known to be involved in exosome mediated mRNA turnover (Shaw and Kamen 1986; Chen and Shyu 1995) and similarly miRNAs with high AU/UA dinucleotide densities show the shortest half lives in primary human neuronal cells (Sethi and Lukiw 2009). Additionally specific seed regions (Rissland *et al.* 2011) or the most 3' bases (Bail *et al.* 2010) have been shown to control the stability of specific miRNAs. It is possible that specific sequences within *miR-252-5p* are responsible for its Dis3-mediated degradation as *miR-252-5p* contains 3 UA dinucleotides in the first 10 bases. However, a greater number of potential Dis3 targets or in depth biochemical analysis would be required to identify such *cis*-acting elements.

A final mechanism through which *miR-252-5p* may be targeted for degradation by Dis3 is through its association with its targets. It has been shown that target mediated miRNA degradation (TMMMD) is a phenomenon known to control the levels of specific miRNAs; for example *let-7* and *miR-34* have been shown to be degraded along with their targets following extensive complementarity (Ameres *et al.* 2010). Similar mechanisms have also been observed in human cells (Baccarini *et al.* 2011) and in rodent neurones (de la Mata *et al.* 2015). It is possible that Dis3 elicits at least part of its regulation on *miR-252-5p* levels through its recruitment to the *miR-252-5p* associated-targets resulting in the simultaneous decay of both RNA species.

### 8.2.3 A potential role for Dis3 in regulating miRNA biogenesis.

The miRNA-sequencing experiments have also identified a potential role of Dis3 in the biogenesis pathway of *miR-982*. The knockdown of Dis3 results in the post-transcriptional downregulation of the 5' arm of *miR-982*. Additionally, the 3' arm may also show downregulation although it was only detected in one of the knockdown samples and as a result would need validating by qRT-PCR. The levels of the *pre-miR-982* hairpin remained unchanged suggesting that Dis3 may play a role in the final stages of *miR-982* biogenesis. For example, Dis3 may be responsible for its 3' trimming prior to Dicer cleavage in a function similar to that observed for another 3'-5' exoribonuclease, Nibbler (Liu *et al.* 2011). Functions for Dis3 have been previously identified within pathway of miRNA biogenesis both in *Drosophila* and mouse embryonic fibroblasts. Dis3 was shown to be involved of the biogenesis of miRtrons, miRNAs produced from introns, in *Drosophila* whilst it was shown to degrade Argonaute bound pre-miRNAs in mouse embryonic fibroblasts.

A similar question of specificity can be asked to that addressed in the previous section; how is Dis3 targeted to the specific miRNA precursors? A likely mechanism is again through 3' terminal tagging which has been previously been demonstrated for pre-miRNAs. For example, Liu *et al* identified a targeting mechanism by which Argonaute bound pre-miRNAs were 3' uridylated which directed their degradation by Dis3 (Liu *et al.* 2014). A similar mechanism for *pre-let-7a* degradation also exists where its oligouridylation stimulates decay by Dis3L2 (Chang *et al.* 2013; Ustianenko *et al.* 2013). It is possible that 3' terminal tagging of *pre-miR-982* may recruit Dis3 if it is indeed involved in its biogenesis. However, this is more likely to be monouridylation which is known to stimulate *pre-let-7a* maturation (Heo *et al.* 2012) rather than the oligouridylation mediated degradation.

Further work would be required to identify how Dis3 is targeted to specific precursors. Due to the strict size selection during the library preparation it is not possible to identify pre-miRNAs within the sequencing data. Therefore additional RNA-sequencing experiments may be performed to look at terminal tagging of miRNA precursors. This could be achieved through using an experiment such as tail-seq (Chang *et al.* 2014) a sequencing protocol designed and optimised specifically for the identification of nucleotide additions to RNAs. Alternatively, a more specialised approach using cRACE could be employed to look as specific pre-miRNAs such as *pre-miR-982*. This may be a useful method to build upon the information gained in this miRNA-seq data.

## 8.3 A novel role for Dis3L2 in controlling developmental proliferation?

The work outlined in Chapter 4 characterised a novel role for Dis3L2 in controlling proliferation within the wing imaginal disc. Knockdown of *dis3L2* in the wing imaginal disc was shown to result in increased proliferation resulting in the overgrowth of the tissue. This translated to a 20% overgrowth of the adult wing resulting from an increase in cell number. Considering mutations in human *DIS3L2* have been shown to result in Perlman syndrome, a condition characterised by foetal overgrowth, these findings may have uncovered a conserved function for Dis3L2 in controlling tissue growth.

### 8.3.1 Why does Dis3L2 knockdown result in hyperplasia of the wing?

There are a number of well characterised pathways known to drive proliferation in the wing imaginal disc. These include the Wingless and Decapentaplegic (Dpp) pathways (Aegerter-Wilmsen *et al.* 2007). An obvious potential cause of overgrowth could be through the increased expression of members of these pathways. However, key members of the Wingless and Dpp pathways remain unchanged at the RNA level following the loss of *dis3L2* in the tissue. Alternatively, increased expression of cell cycle regulatory factors such as Cyclin D and CDK4 may have been responsible but these also remain unchanged in the RNA-seq data. Somewhat surprisingly overexpression of Cyclin D and CDK4 in the wing disc has previously been shown not to cause substantial overgrowth (Aegerter-Wilmsen *et al.* 2007). In the absence of a strong link at the RNA level it is likely that Dis3L2 regulates proliferation either through the control of transcripts that are not yet characterised to play a role in proliferation in *Drosophila*; or through the regulation of transcripts encoding proteins involved, for instance, in post-translational modifications.

It has previously been shown that changes in the level of proliferation are often offset by changes in cell size to ensure that the development of the final tissue remains constant (Potter and Xu 2001). However, it is clear from these data that the compensatory mechanism is disrupted by the loss of *dis3L2* in the wing disc. Furthermore, there is also evidence that altered growth in a defined region of the disc leads to a complementary change in the peripheral region showing an element of non-autonomous control (Baena-Lopez and Garcia-Bellido 2006). This is largely not the case in the terms of the *engrailed-GAL4* experiment as the overgrowth is largely restricted to the posterior region. This provides further evidence for the loss of control over inter-tissue growth compensatory mechanisms upon *dis3L2* knockdown. An element of crosstalk may still exist as the anterior area does increase significantly, although only 2% compared to the 11% overgrowth observed in the posterior compartment of the wing.

Either way this apparent loss of control in the compensatory mechanisms may aid in future determination of critical pathways controlled by Dis3L2.

Interestingly, although overexpression of cell cycle proteins do not always drive proliferation (Aegerter-Wilmsen *et al.* 2007) the founding Dis3L2 study by Astuti *et al.* saw changes in protein expression of a collection of cell cycle related proteins. For example, the pro-proliferative Cyclin B1 was increased in expression whilst the anti-proliferative Aurora B kinase and phospho-CDC25C were reduced in expression in *DIS3L2* knockdown HeLa cells. Interestingly the change in expression of Aurora B kinase did not occur at the RNA level (Astuti *et al.* 2012) and therefore this could also be the case in the data presented here. The absence of clear proliferation-related transcripts in the sequencing data suggests that the effect on proliferation could be occurring at the translational or post-translational level.

A number of cell cycle events are regulated by post-translational modifications such as phosphorylation of Retinoblastoma (Rb) resulting in its dissociation with E2F allowing it to drive pro-mitotic transcriptional changes driving the G1/S transition (Chellappan *et al.* 1991; Dyson 1998; Polager and Ginsberg 2008). Another example of post-translational modification is linked to cell metabolism. O-GlcNAc transferase mediated glycosylation of PFK1 has been shown to promote proliferation in cultured cells (Yi *et al.* 2012) through altering their metabolism to allow sustained growth. It is possible that this could also provide a link to the starvation phenotype observed in Chapter 4. For example, if a similar mechanism is involved in the wing imaginal disc then an increase in metabolic efficiency to allow sustained growth could also provide an advantageous environment under starvation conditions. Glycosylation is also known to influence the activity of Wnt (Wingless orthologue) (Komekado *et al.* 2007) and Notch (Shi and Stanley 2003) signalling both of which are critical in the growth of the wing disc. It is therefore possible that the loss of Dis3L2 activity in the wing imaginal disc results in changes at both the post-transcriptional and post-translational level which combine to result in dysregulation of proliferation.

The RNA-seq data showed a discrete set of transcripts that appear to be sensitive to Dis3L2 activity. It is noteworthy that the majority of these differentially expressed transcripts were expressed at low levels in control wing imaginal discs during the wandering 3<sup>rd</sup> instar larval stage. It is possible that at least some of these transcripts are maintained at low levels by Dis3L2 to prevent over proliferation within the tissue. Two of the identified transcripts have been shown to be potential Dis3L2 targets; the predicted transcription factor *CG2678* and the cation channel *pyrexia*. An increase in Pyrexia may also contribute to the overgrowth phenotype as an increase in cation influx has been shown to be promote proliferation in a

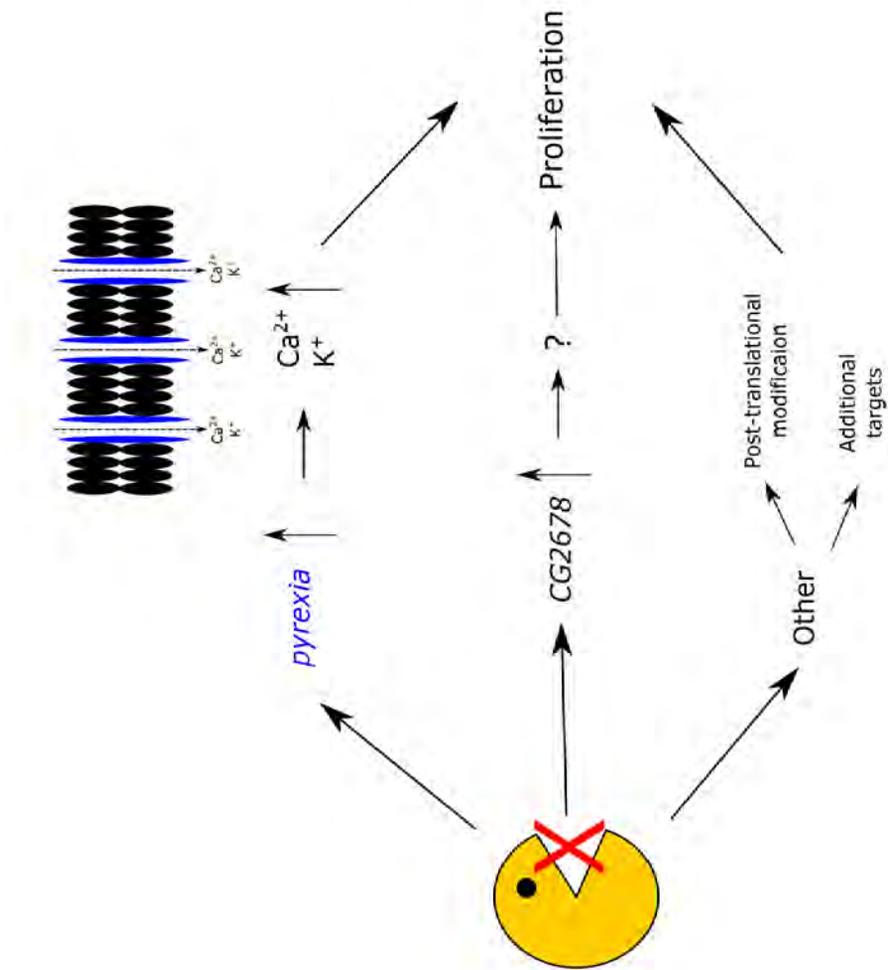
number of examples (DeCoursey *et al.* 1984; Berridge 1995; Hooper *et al.* 2005; Urrego *et al.* 2014) (Figure 8.1).

The potential effects of increased *CG2678* and *pyrexia* expression will be discussed in greater detail in section 8.4.2 below.

### **8.3.2 Is the proliferative function of Dis3L2 tissue specific?**

Are the effects of Dis3L2 on tissue growth tissue specific or a more global mechanism of regulation? Taking Perlman syndrome as an example it appears that the major tissues affected are the kidney and liver whilst tissues such as the lung are not reported to be affected. A clue here may reside in the nature of the affected tissues; both the kidney and liver are regenerative tissues whilst the lung is not a naturally regenerating tissue. Similarly, in the *Drosophila* system used here a global overgrowth is not observed following ubiquitous knockdown of *dis3L2*. However, the wing/wing discs which are also regenerative tissues are receptive to the overgrowth signals. Interestingly a study into wing imaginal disc regeneration showed that they lose their regenerative potential after the start of the third instar larval stage (Smith-Bolton *et al.* 2009). This would be concordant with the GAL80<sup>ts</sup> experiments in showing that the loss of Dis3L2 does not stimulate overgrowth if it is induced after the start of L3. Additionally, this regeneration was stimulated by inducing apoptosis in the wing pouch; this is effectively mimicked by the conditions in the *pcm<sup>14</sup>/dis3L2<sup>KD</sup>* wing imaginal discs which show size rescue compared to the mutant conditions alone (Chapter 6).

A working hypothesis, therefore, is that Dis3L2 normally functions to control the level of proliferation in naturally regenerating tissues. If this is the case then it may be possible to look for specific pathways involved during the regeneration process. It would be likely that at least some of those transcripts regulated by Dis3L2 function within the regeneration pathway. To test this hypothesis further work would be required to investigate if the knockdown of *dis3L2* in other *Drosophila* tissues also results in overgrowth.



**Figure 8.1 How might Dis3L2 control proliferation in *Drosophila* wing imaginal discs?** The loss of *dis3L2* results in the upregulation of specific transcripts of which *pyrexia* and *CG2678* are strong candidates for Dis3L2 targets. *Pyrexia* is a cation channel permeable to calcium and potassium which have both been shown to promote proliferation. The protein encoded by *CG2678* is predicted to be a transcription factor, therefore its upregulation may drive proliferation through the upregulation of unknown transcriptional targets. There are also likely to be targets that have been missed in this study that may also contribute towards the increase in proliferation within the wing imaginal discs.

## 8.4 Are *pyrexia* and *CG2678* direct targets of Dis3L2?

The RNA-sequencing data put forward in Chapter 5 identified a pool of 239 Dis3L2-sensitive RNAs which were mainly expressed at low levels in the control discs. This therefore suggests that Dis3L2 has a high degree of specificity together with functioning to keep critical mRNAs at low levels within the tissue. This relatively low number of differentially expressed genes is consistent to similar profiling experiments performed by Lubas *et al* where DIS3L2 knockdown in HeLa cells resulted in 799 transcripts showing fold changes of >2 fold (Lubas *et al.* 2013). There was no further attempt to identify direct or indirect targets of Dis3L2 within this experiment. However, transcripts involved in cell-cycle regulation were enriched which may be consistent with the phenotypes observed here. The pool of misexpressed transcripts in the RNA-seq experiment in this thesis consisted of 125 transcripts that increased in expression following *dis3L2* knockdown within the wing imaginal disc; it was this sub-pool of transcripts that was likely to contain the novel Dis3L2 targets.

### 8.4.1 *pyrexia* and *CG2678* are unlikely to be the only Dis3L2 targets.

The knockdown of *dis3L2* within the wing imaginal disc resulted in the increased expression of 125 transcripts (by >1.5 fold). This pool of transcripts was likely to contain both those directly and indirectly affected by the loss of *dis3L2*. To address this, a similar approach was taken to those previously performed in the lab (Jones *et al.* 2013; Towler *et al.* 2015; Waldron *et al.* 2015; Jones *et al.* 2016). This involved using the pre-mRNA as a transcriptional read out. Due to financial limitations only 4 of the previously validated transcripts were probed in this manner; these were *cytochrome-c-distal (cyt-c-d)*, *CG31808*, *pyrexia (pyx)* and *CG2678*. Both *cyt-c-d* and *CG31808* showed transcriptional changes and therefore were likely to be driven by indirect effects. In contrast, *pre-pyx* and *pre-CG2678* showed no change in expression following the knockdown of *dis3L2* and therefore the mature transcripts change at the post-transcriptional level.

Although there was clearly some technical limitations of these experiments the two control genotypes were highly concordant therefore showing consistency within the data. In line with this 6 of the 7 validated upregulated RNAs showed complementary changes by qRT-PCR showing reliability in the filtering methods used to identify the misexpressed transcripts. Both the post-transcriptionally regulated transcripts identified in this study show at least a degree of pathway specificity as neither increase significantly in expression in *pacman* null mutant discs. This shows that there are unlikely to be regulated by the 5'-3' pathway of mRNA decay; at least in the wing imaginal discs. It is therefore likely that within this dataset *pyx* and *CG2678* are the two strongest candidates as Dis3L2 targets.

However, it is highly unlikely that Dis3L2 only targets 2 mRNAs for degradation within the wing imaginal discs. Why were more Dis3L2-target candidates not identified in this data? There could be multiple reasons, for example, there were technical faults with the experiment such as the low replicate number which meant that it was not possible to use the statistical output by cufflinks. This in turn meant that stringent, user-defined filtering had to be performed. It is likely that many potential Dis3L2 targets were missed due to the strict parameters set during the filtering such as the arbitrary FPKM cut off. Moreover, RNA interference was used to knockdown *dis3L2* due to the absence of a suitable mutant; although the absence of a mutant was not through a lack of trying (see Chapter 7). Although the knockdown achieved, as shown by qRT-PCR, in the knockdown samples was 80% this would mean that 20% of the wild-type Dis3L2 activity remained. The 20% background activity could therefore mean that some Dis3L2 targets are being sufficiently degraded to avoid detection using these methods. Therefore a null mutant would be desirable to identify further *dis3L2* targets which has proved highly successful in identifying Pacman targets in the Newbury lab (Waldron *et al.* 2015; Jones *et al.* 2016).

#### **8.4.2 Could Pyrexia and CG2678 contribute to the overgrowth phenotype?**

Both *pyx* and *CG2678* are strong Dis3L2-target candidates. What is known about these two genes at the protein level and is there a possibility they may contribute to the overgrowth phenotype? *pyrexia* is conserved in humans and encodes a cation channel which has been shown to be involved in thermotolerance in *Drosophila* (Lee *et al.* 2005). *Pyrexia* mutants have also been shown to be gravataxis defective and become paralytic after heat stress. It is one of 13 members of a family of transient receptor potential (TRP) channels of which 6 are reliably detected in the wing imaginal discs by RNA-seq. Of the 6 expressed TRP channels only *pyrexia* shows a change in expression following *dis3L2* knockdown showing an element of specificity (Figure 8.2). The overall expression of *pyrexia* is shown on Flybase to be fairly ubiquitous but in all tissues is of very low expression potentially suggesting it is tightly controlled.

The potential contribution of *pyrexia* to the overgrowth phenotype is not immediately obvious and there was not time during this project to investigate this further. However, it is possible to draw a few initial hypotheses. For example, *pyrexia* has been shown to be regulated by the Myb-E2f2-RB transcriptional repressor complex which is known to function to inhibit mitotic progression. This could suggest a contribution for *pyrexia* towards cell cycle progression. Furthermore, a number of studies have shown that cation channels can play an important role cell cycle progression. For example, Ca<sup>2+</sup> influxes are observed during G1 and the G2/M transition (Berridge 1995). Additionally, stimulating Ca<sup>2+</sup> signalling in microglial cells stimulates

their proliferation (Hooper *et al.* 2005). Interestingly, this was shown to be mediated by phospholipase C signalling which functions downstream of Pyrexia (Kwon *et al.* 2008). In addition to the calcium signalling, potassium signalling has also been shown to stimulate proliferation (DeCoursey *et al.* 1984; Urrego *et al.* 2014). Together with this K<sup>+</sup> channels have been shown to mediate efficient Ca<sup>2+</sup> uptake (Lepple-Wienhues *et al.* 1996; Lallet-Daher *et al.* 2009). It is therefore possible that the increase in *pyrexia* expression may in turn result in an increase in the number of cation channels in the cell membrane. This would potentially increase the cells ability to take up cations such as Ca<sup>2+</sup> and K<sup>+</sup> which may under certain conditions act as a pro-proliferative signal (Figure 8.1). Such a mechanism has not been described in flies and would therefore require further investigation for which there was not time during this project.

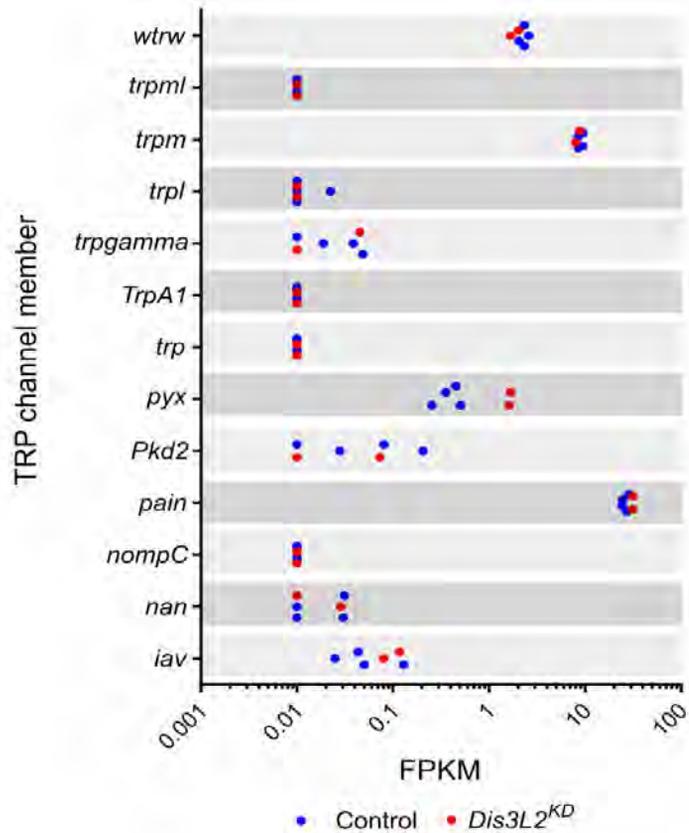
The other post-transcriptionally upregulated transcript was *CG2678*, however, unlike *pyrexia*, little is known about its function and therefore without further work it is not possible to draw any proliferative hypotheses. Unlike *pyrexia*, *CG2678* is only conserved within *Drosophila*. In terms of its biological function, *CG2678* contains a Zinc-finger domain suggesting it acts as a transcription factor, however its targets are unknown. *CG2678* was also a hit in a screen to identify genes involved in dendrite morphogenesis (Parrish *et al.* 2006), however, no further information was described. It is again possible that *CG2678* plays a role in the control of cell growth through regulating the expression of target genes (Figure 8.1); however, at this point any strong hypotheses cannot be made.

It is highly likely that other factors are involved in stimulating the increased proliferation observed in the *dis3L2* knockdown wing imaginal discs. For example, *pyrexia* and *CG2678* will not be the only factors directly regulated by Dis3L2. Others are likely to have been missed as explained in the previous section (8.4.1). There was no detection of obvious pro-proliferatory targets as shown by a lack of proliferation- related enrichment in the gene ontology analysis. It is therefore possible that *dis3L2* knockdown affects the expression of proteins involved in post-translational regulation of pro- or anti-proliferative proteins. For instance, a number of key cell cycle regulatory proteins are controlled at the post-translational level by phosphorylation. It is possible that the loss of Dis3L2 results in the increased expression of a kinase this could in turn result in the increased phosphorylation and activation of proteins that function to drive cell cycle progression (Figure 8.1). As the RNA-seq data is only looking on a transcriptome level these changes would not be observed. In line with this some transcripts encoding proteins containing kinase domains were misexpressed, however, further work would be required to see if they influence cell cycle progression.

### 8.4.3 Methods to identify additional Dis3L2 targets.

The RNA-seq experiments outlined in Chapter 5 identified two likely targets of Dis3L2 in *pyrexia* and *CG2678*. However, as outlined previously Dis3L2 will be responsible for the degradation of other mRNAs within the cytoplasm. Further work would be required in order to identify additional *in vivo* Dis3L2 targets. To achieve this it would be more suitable to use a *dis3L2* null mutant, the production of which was in process during the writing of this thesis (Chapter 7). Using a null mutant carries many advantages over using RNAi; the main one being the fact that there would be no background activity as there would be in the *dis3L2* knockdown samples. It would also be important to run additional samples which would aid in the statistical analysis. A recent study from the lab used two null mutants each with their own isogenic control with 3 replicates of each; these were subsequently grouped giving a sample size of 6 for the mutant and control condition. This was shown to be sufficient to give strong statistical analysis (Jones *et al.* 2016). If the production of the mutant was successful it would be good to follow a similar experimental plan. The CRISPR experiment may produce multiple mutants in which case 2 independent mutants could be used which would act an additional layer of control.

To definitively identify potential targets of Dis3L2 one would need to capture Dis3L2 in the process of their decay. To achieve this experiments such as cross-linked immunoprecipitation (CLIP) could be performed. This would involve the pull down of Dis3L2 followed by either RNA sequencing to assess all the transcripts to which Dis3L2 is bound or qRT-PCR using probes against the previously identified targets. The limiting factor would be the presence of an antibody of sufficient quality to obtain specificity. The production of an anti-Dis3L2 antibody is currently underway (Chapter 7) and therefore if successful could be used. If not it might be possible to tag the endogenous Dis3L2, using CRISPR and use a highly efficient and specific anti-tag antibody. Another possible issue this experiment might encounter would be the short period of time Dis3L2 would be associated with the degraded transcript. This may be difficult to capture and therefore an adaptation using a catalytically dead Dis3L2 mutant may be required. This catalytically dead mutant would be made using CRISPR. A CLIP experiment such as this would allow identification of those transcripts physically associated with Dis3L2 and therefore those in the process of being degraded.



**Figure 8.2: *pyrexia* is the only member of the TRP channel family to be upregulated following *dis3L2* knockdown.** Graphical representation of the replicate consistency for all 13 members of the TRP channel family. Knockdown replicates (*Dis3L2<sup>KD</sup>* = ; *UAS-dis3L2<sup>RNAi</sup>/+* ; *69B-GAL4/+*) shown in red whilst control replicates ( ; *UAS-dis3L2<sup>RNAi</sup>* ; and ; *69B-GAL4*) are shown in blue.

## 8.5 A role for Dis3L2 in starvation response?

In addition to the overgrowth phenotype an enhanced stress response was observed upon ubiquitous *dis3L2* knockdown. This was a thought-provoking and striking result which thus far remains unexplained. 1-2 day old adults with ubiquitous *dis3L2* knockdown were shown to survive 25% longer than both sibling and parental controls under food-free conditions. However, due to it being slightly tangential from the rest of the work there was not much time to delve deeper into the underlying mechanism behind it.

Why could the loss of Dis3L2 result in an enhanced response to starvation? It is possible that the absence of Dis3L2 results in the increased expression of specific RNAs which are involved in the stress response. For example, proteins such as PTEN (Petiot *et al.* 2000) known to enhance the starvation response may become overactive when *dis3L2* is depleted therefore allowing increased survival under starvation conditions. Alternatively, knockdown of *dis3L2* may result in changes in resource allocation and enhance the efficiency in sourcing nutrients. Flies showing starvation resistance have been previously shown have higher levels of triglycerides (Slocumb *et al.* 2015); it is possible that the loss of *dis3L2* results in an increased rate of triglyceride synthesis. Slocumb *et al.* also identified that starvation resistant *Drosophila* lines were larger than those that are affected more by starvation. However, as described in Chapter 4 ubiquitous knockdown of *dis3L2* does not affect the overall size of the adult fly and thus an increase in fly size is unlikely to be a driving factor.

Upon starvation *Drosophila* source the required nutrients for survival using autophagy (Scott *et al.* 2004). Autophagy is a process whereby lysosomal cytoplasmic degradation is used to recover nutrients during stress (Mizushima 2007) often resulting in cell shrinking. Autophagy has been shown to be inhibited by TOR signalling (Blommaert *et al.* 1995; Scott *et al.* 2004) whilst it is stimulated by the negative regulator of the insulin/phosphoinositide-3-kinase pathways, PTEN (Petiot *et al.* 2000). Knockdown of *dis3L2* may disrupt this balance allowing more efficient autophagy resulting in an increase of nutrient availability under the starvation conditions. TOR signalling is also known to be involved in the control of cell growth (Fingar and Blenis 2004) and therefore may provide a connection between the starvation and overgrowth phenotypes. However, how this may occur is not immediately obvious as the functions are contrasting; whilst TOR signalling promotes cell growth it inhibits autophagy. Furthermore, an increase in proliferation may be expected to decrease starvation survival as the greater number of cells would require increased nutrients and therefore have a negative effect on starvation resistance. On the other hand the increased proliferation may be achieved as a result of the increased nutrient availability which also contributes to the starvation resistance. Finally, work in human tissue culture cells show that *dis3L2* knockdown aids the cells in

apoptotic survival (Thomas *et al.* 2015) and therefore perhaps an increase in cell survival under the starvation conditions increases starvation resistance.

## **8.6 Co-ordination between apoptosis and proliferation by the 5'-3' and 3'-5' pathways.**

The work presented in Chapters 4 to 6 and previous work from the Newbury lab have uncovered pathway specificity for the 5'-3' and 3'-5' pathways of RNA decay. The 5'-3' exoribonuclease Pacman has been shown to be involved in controlling the level of apoptosis through the modulation of *hid*, *grim*, and *rpr* expression (Waldron *et al.* 2015; Jones *et al.* 2016). In contrast the work presented here shows a distinct role for Dis3L2 in the regulation of developmental proliferation, although the mechanisms underlying this are yet to be elucidated. The intriguing question is how the pathway specificity might be achieved? The current hypothesis is that specificity is gained through specific targeting mechanisms which are discussed below.

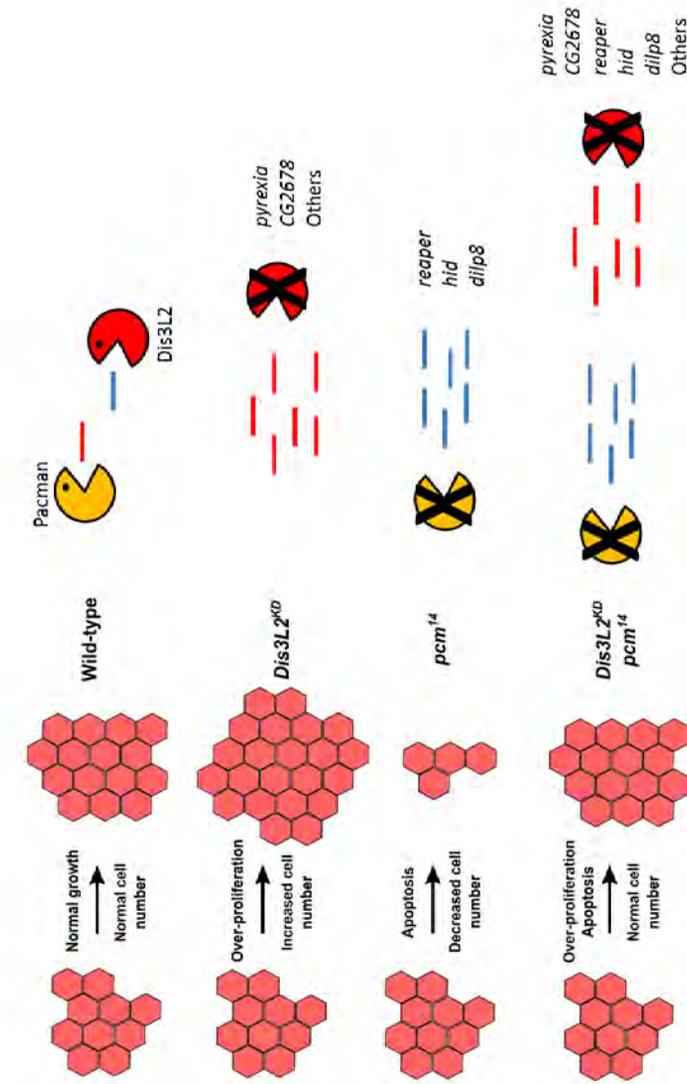
### **8.6.1 Are Pacman and Dis3L2 involved in regulating opposing pathways?**

The work presented in this thesis gives evidence towards the hypothesis that Pacman and Dis3L2 are involved in the regulation of apoptosis and proliferation respectively. Previous published work from the lab has shown Pacman to be involved in the regulation of apoptosis through its control over the levels of *hid*, *grim* and *reaper* (Waldron *et al.* 2015; Jones *et al.* 2016). The increase in *reaper* expression has been repeated in this work and it has been shown to be unaffected by the additional knockdown of *dis3L2*. Similarly *hid*, *grim* and *reaper* remain unchanged in the *dis3L2* knockdown RNA-seq data (Chapter 5) confirming that the activity of Pacman on the three apoptotic proteins is specific. This suggests that Dis3L2 does not function in regulating the onset of apoptosis through this pathway which was further established through the use of the *Df(3L)H99* deletion. The *Df(3L)H99* induced inhibition of apoptosis resulted in an additive effect on the *dis3L2*-depletion induced wing overgrowth. *Drosophila* containing the *Df(3L)H99* deletion together with *dis3L2* knocked down presented with wings that were significantly larger than either those with the deletion or knockdown alone. These data together indicates that although Dis3L2 has been implicated in the clearance of mRNAs during apoptosis (Thomas *et al.* 2015) it is not involved in the regulation of the onset of apoptosis.

This work also strengthens the hypothesis proposed in Chapter 4 where Dis3L2 is involved instead in the regulation of proliferation in the developing wing imaginal disc. The work outlined in Chapter 4 shows strong evidence for a role for Dis3L2 in the control of proliferation. Its knockdown resulted in an increased level of mitotic cells which subsequently led to an increase in the number of cells in the wing and significantly larger wings.

It is therefore intriguing that two pathways of RNA decay appear to be involved in the regulation of opposing pathways through targeting specific RNAs for degradation. As discussed in Chapter 6 it appears that the increase proliferation induced following the loss of *dis3L2* is able to compensate for the *pcm<sup>14</sup>* induced apoptosis at least in terms of tissue area (summarised Figure 8.3). This further shows pathways specificity which may be achieved through a number of mechanisms.

**Figure 8.3: Hypothesis behind Dis3L2 and Pacman regulating proliferation and apoptosis respectively.** Under normal



conditions Dis3L2 (red) and Pacman (yellow) regulate the expression of their specific RNA targets allowing normal tissue growth. The loss of *dis3L2* results in increased proliferation as a result of the upregulation of its specific targets (*pyrexia*, *CG2678* and those not yet detected) together with indirect changes in the expression of transcripts. Similarly, the loss of *pacman* results in the specific upregulation of pro-apoptotic RNAs *reaper* and *hid* resulting in increased apoptosis and subsequent loss of tissue. Consistent with this, here it is shown that the loss of both *dis3L2* and *pacman* results in apparent 'normal' tissue size which is likely to result from the increased proliferation as a result of the upregulation of Dis3L2 targets compensating for those cells lost as a result of the misregulation of Pacman targets. Although tissue area is rescued the lethality induced by *pacman* mutation, nor the developmental delay (induced by *dilp8* upregulation) are rescued.

## 8.6.2 How might the pathway specificity be achieved?

The outstanding question from this work is how do the 5'-3' and 3'-5' pathways of cytoplasmic RNA decay gain their specificity to elicit their control over the opposing developmental processes. This is likely to occur through RNA binding proteins or terminal tagging recruiting the specific decay factors. There may also be the possibility of instability elements within the UTR's which may elicit further pathway specificity. These various mechanisms are discussed below.

How might Dis3L2 be recruited to its specific targets? A few clues can be taken from the literature in that recent studies have identified the potential of uridylation as a Dis3L2 targeting mechanism. For instance, oligouridylation of *pre-let-7a* by TUT4 has been shown to result in its degradation by Dis3L2 (Chang *et al.* 2013; Ustianenko *et al.* 2013). In line with this, *dis3L2* mutants in *S. pombe* and human tissue culture cells contain an increased abundance of uridylated transcripts suggesting they have been tagged for degradation by Dis3L2 (Malecki *et al.* 2013; Thomas *et al.* 2015). This is interesting as uridylation has also been shown to promote degradation of mRNAs in the 5'-3' direction following the association of Lsm1-7/Pat1 to the uridylated 3' end (Rissland and Norbury 2009). This is consistent with work showing that XRN1 physically interacts with the Lsm1-7/Pat1 complex (Nissan *et al.* 2010) in yeast. An intriguing question is therefore how the specificity is gained following uridylation. It is possible that Lsm1-7/Pat1 and Dis3L2 compete for the binding of the uridine tag. However, if this was the case one may not expect to see the increase in uridylated transcripts in *dis3L2 S. pombe* mutants; in this case Lsm1-7/Pat1 would be able to bind and promote decay from the 5' end. It is thus likely that the uridylation signal aids specificity through an alternative mechanism. Unfortunately due to the poly(A) selection of the cDNA libraries it was not possible to analyse the extend of uridylation in the Dis3L2 knockdown discs which was an oversight. The potential of uridylation in *Drosophila* has become more likely through the very recent identification of an uridyl transferase named Tailor (CG1091) (Reimao-Pinto *et al.* 2015). Interestingly, an incredibly large interaction screen in *Drosophila* S2 cells identified a potential interaction between Tailor and Dis3L2. When an antibody becomes available it would therefore be of great interest to validate this potential interaction by co-immunoprecipitation experiments.

This introduces another way in which the decay enzymes may gain their specificity; through RNA binding proteins which may either recruit Dis3L2 to targets or may be in direct association with Dis3L2 to function as a targeting complex. A similar mechanism is observed for the exosome where it is known that its targeting is mediated by other RNA binding proteins such as Tristetraprolin (TTP), MMP6 and the TRAMP complex (Chen *et al.* 2001; LaCava *et al.* 2005). Co-immunoprecipitation experiments performed in HEK293 cells were unable to identify any

direct protein-protein interactions (Lubas *et al.* 2013). These cells did however use a FLAG-tagged Dis3L2 and therefore the presence of the tag may disrupt potential interactions. It would be of great interest to perform similar co-immunoprecipitation experiments in this *Drosophila* system to attempt to identify any potential cofactors which may aid in the targeting of Dis3L2. However, this work would require the presence of a high quality antibody. A recent study has also shown an RNA-dependent interaction between Dis3L2 and Ago2 and Dis3L2 and TUT1 (Haas *et al.* 2016). These interactions were proposed to be due to target mediated miRNA trimming following highly complementary miRNA-target base pairing. This therefore suggests that uridylation by TUT1, in addition to TUT4 (Chang *et al.* 2013), may be involved in the recruitment of Dis3L2 in specific circumstances.

If the pathway specificity is driven by specific RNA-binding proteins then one would hypothesise that different RNA binding proteins would recruit the different decay enzymes. Evidence of this could be provided through specific interactions shown through co-immunoprecipitation experiments. For example, XRN1 is known to have a long, unstructured C-terminal which acts as a binding site for the decapping activators Dcp1 and EDC4 in both human and *Drosophila* cells (Braun *et al.* 2012). It is possible that other RNA binding proteins recruit XRN1 to its apoptotic targets through interaction with the C-terminal tail which are unable to stimulate Dis3L2-mediated decay.

The study by Lubas *et al.* also observed that Dis3L2 is associated with translating ribosomes in HeLa cells so it may also be possible that Dis3L2 degrades at least some of its targets co-translationally or immediately after translation. Interestingly XRN1 has also been found on translating ribosomes and therefore this method of recruitment may be used by both exoribonucleases. Co-translational Dis3L2 mediated decay seems unlikely as Dis3L2 requires an open 3' end to start its processive decay which would not be accessible during standard translation. The translating ribosome/s would be processing in the 5'-3' direction and thus would run into Dis3L2 resulting in incomplete decay. XRN1 may function to degrade transcripts undergoing cap-independent translation as it would be able to access the unprotected 5' end. This could be used as a method to ensure only a small amount of potentially harmful proteins are produced as rapid decay would prevent the over association of ribosomes. It is possible that in recruiting Dis3L2 to the polysomes allows ideally positions it to degrade the transcript immediately after translation thus ensuring excess protein is not produced.

On top of the trans-acting factors another way in which specificity may be obtained is through intrinsic instability elements, for instance AU rich elements (AREs) in the 3' UTR. A number of proteins such as TTP, KSRP, and Auf1 (Chen *et al.* 2001) have been shown to bind AREs resulting

in changes in the stability of the bound mRNA. AREs have been shown to result in the destabilisation of many transcripts due to degradation by Dis3L2 (Lubas *et al.* 2013), XRN1 (Stoecklin *et al.* 2006) and the exosome (Chen *et al.* 2001; Stoecklin *et al.* 2006). As AREs appear to stimulate decay by all three major processive exoribonucleases within the cell it is unlikely to provide enzyme specificity. However, other instability factors, specific for Dis3L2 or XRN1 may be present; these may act as specific binding sites for targeting RNA binding proteins mentioned previously. There was no enrichment in specific motifs within the list of upregulation transcripts from the Dis3L2 RNA-seq data, however this contained transcripts affected both directly and indirectly by Dis3L2. This is also the case for *pacman* mutant sequencing data and therefore to gain further knowledge into such motifs a larger set of direct targets would be required.

A final way in which the pathway specificity may be achieved is through subcellular localisation. For example, XRN1 is known to be located in P-bodies and therefore its apoptotic specific targets may also be located there. Studies into the localisation of Dis3L2 have not shown specific P-body localisation and therefore it is possible that some of the specificity is gained in this manner.

## 8.7 Implications in human disease

What are the overriding implications of the findings presented in this thesis? Both of the 3'-5' exoribonucleases have been shown to be mutated in human diseases. For example, deletions of exons 6, 9 or 19 in human *DIS3L2* have been shown to result in Perlman syndrome (Astuti *et al.* 2012; Morris *et al.* 2013). Perlman syndrome is a rare autosomal recessive condition characterised by foetal overgrowth including macrocephaly and visceromegaly (enlargement of abdominal organs) (Alessandri *et al.* 2008). Perlman syndrome sufferers have a poor prognosis whilst they also show a high predisposition to Wilms' tumour (Astuti *et al.* 2012; Morris *et al.* 2013). Interestingly, 6 (30%) of 20 sporadic Wilms tumour samples analysed showed either complete or partial deletion of Dis3L2 (Astuti *et al.* 2012).

Its apparent role in Perlman syndrome and Wilms' tumour therefore places Dis3L2 in a clinical setting and the findings presented here indicate a conserved function for Dis3L2 in the control of developmental proliferation and ultimately tissue growth. The apparent position for Dis3L2 in the proliferation pathway also raises the question is it involved in other overgrowth disorders? In general there is no current evidence for loss of Dis3L2 expression in other cancers; however, in the age of whole exome and next generation sequencing it is possible that new studies may uncover a role of Dis3L2 in the progression of specific cancers. Extending

these findings would be critical to understand the mechanism behind Dis3L2s regulation of proliferation which may aid us in understanding the development of Perlman syndrome and other overgrowth conditions.

Similarly to Dis3L2, Dis3 has also been implicated in human disease. For example it has been shown to be mutated in 11% of multiple myeloma (Chapman *et al.* 2011) and 4% of acute myeloid leukaemia cells (Ding *et al.* 2012). This poses an intriguing question; if the loss of Dis3 results in widespread cell death and apoptosis (Chapter 3) then how do loss of function (lof) mutations (Tomecki *et al.* 2014) result in advanced cell survival? It is more than likely that the *DIS3* mutations occur alongside a number of other mutations that may, when in collaboration, enhance cell survival. For instance, if a key oncogene becomes overexpressed, such as Ras then the lof *DIS3* mutations may aid tumour survival. Interestingly, Ras has also been shown to be commonly overexpressed in multiple myeloma (Steinbrunn *et al.* 2011). Alternatively, these tumours may have mutations in tumour suppressors such as p53 thus reducing their ability to enter apoptosis. In this case it is possible that the mass misregulation resulting from *DIS3* mutation results in the increased abundance of transcripts which may function to drive tumourigenesis. The frequency of mutation, especially in the case of myeloma, would suggest a strong role for Dis3 in tumour progression. This is an interesting possibility and one that would require further work.

## 8.8 Concluding remarks and Future work

The extensive work presented in this thesis has focused on the role of Dis3L2 in proliferation using the wing imaginal disc as a model developmental tissue. This has identified a conserved role for Dis3L2 in controlling developmental proliferation and subsequently an ideal model system in which this role. This work has also uncovered the intriguing finding that, although some redundancy will exist, Dis3L2 and Pacman are responsible for the regulation of transcripts in discrete cellular pathways. Therefore further understanding of how the pathway specificity is achieved is critical. As alluded to throughout this thesis additional work is required to uncover such mechanisms. The availability of a mutant and antibody of Dis3L2 would open a number of new doors for experiments such as repeating RNA-sequencing and CLIP. If RNA-sequencing was to be performed on *dis3L2* mutants it should be performed on ribo-depleted RNA rather than the poly(A) selection previously used. This would allow the analysis of Dis3L2's effect on non-coding RNAs together with the ability to assess 3' terminal tagging of potential Dis3L2 targets.

Identifying a larger pool of *bona fide* Dis3L2 targets and how they are signalled for Dis3L2-mediated decay is critical to further understand the role Dis3L2 plays in regulating proliferation which ultimately may help in therapeutics for overgrowth disorders.

## References

- Abbott, M. K. and J. A. Lengyel (1991). "Embryonic Head Involution and Rotation of Male Terminalia Require the *Drosophila* Locus Head Involution Defective." *Genetics* **129**(3): 783-789.
- Abdelmohsen, K. and M. Gorospe (2010). "Post-transcriptional regulation of cancer traits by HuR." *Wiley interdisciplinary reviews. RNA* **1**(2): 10.1002/wrna.1004.
- Aegerter-Wilmsen, T., C. M. Aegerter, E. Hafen and K. Basler (2007). "Model for the regulation of size in the wing imaginal disc of *Drosophila*." *Mech Dev* **124**(4): 318-326.
- Aegerter-Wilmsen, T., M. B. Heimlicher, A. C. Smith, P. B. de Reuille, R. S. Smith, *et al.* (2012). "Integrating force-sensing and signaling pathways in a model for the regulation of wing imaginal disc size." *Development* **139**(17): 3221-3231.
- Akamatsu, W., H. Fujihara, T. Mitsushashi, M. Yano, S. Shibata, *et al.* (2005). "The RNA-binding protein HuD regulates neuronal cell identity and maturation." *Proceedings of the National Academy of Sciences of the United States of America* **102**(12): 4625-4630.
- Alessandri, J. L., F. Cuillier, D. Ramful, S. Ernould, S. Robin, *et al.* (2008). "Perlman syndrome: report, prenatal findings and review." *Am J Med Genet A* **146A**(19): 2532-2537.
- Allmang, C., E. Petfalski, A. Podtelejnikov, M. Mann, D. Tollervey, *et al.* (1999). "The yeast exosome and human PM-Scl are related complexes of 3' → 5' exonucleases." *Genes & Development* **13**(16): 2148-2158.
- Ameres, S. L., M. D. Horwich, J.-H. Hung, J. Xu, M. Ghildiyal, *et al.* (2010). "Target RNA-Directed Trimming and Tailing of Small Silencing RNAs." *Science* **328**(5985): 1534-1539.
- Ameres, S. L. and P. D. Zamore (2013). "Diversifying microRNA sequence and function." *Nat Rev Mol Cell Biol* **14**(8): 475-488.
- Anderson, J. R., D. Mukherjee, K. Muthukumaraswamy, K. C. M. Moraes, C. J. Wilusz, *et al.* (2006). "Sequence-specific RNA binding mediated by the RNase PH domain of components of the exosome." *RNA* **12**(10): 1810-1816.
- Anderson, J. S. and R. P. Parker (1998). "The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex." *EMBO J* **17**(5): 1497-1506.
- Andrulis, E. D., J. Werner, A. Nazarian, H. Erdjument-Bromage, P. Tempst, *et al.* (2002). "The RNA processing exosome is linked to elongating RNA polymerase II in *Drosophila*." *Nature* **420**(6917): 837-841.
- Ashburner, M. (1989). *Drosophila, A Laboratory Handbook*, Cold Spring Harbour Laboratory Press.
- Aspden, J. L., Y. C. Eyre-Walker, R. J. Phillips, U. Amin, M. A. Mumtaz, *et al.* (2014). "Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq." *eLife* **3**: e03528.
- Astuti, D., M. R. Morris, W. N. Cooper, R. H. Staals, N. C. Wake, *et al.* (2012). "Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility." *Nat Genet* **44**(3): 277-284.
- Azevedo, R. B., V. French and L. Partridge (2002). "Temperature modulates epidermal cell size in *Drosophila melanogaster*." *J Insect Physiol* **48**(2): 231-237.
- Baccarini, A., H. Chauhan, T. J. Gardner, A. Jayaprakash, R. Sachidanandam, *et al.* (2011). "Kinetic analysis reveals the fate of a microRNA following target regulation in mammalian cells." *Current biology : CB* **21**(5): 369-376.
- Badis, G., C. Saveanu, M. Fromont-Racine and A. Jacquier (2004). "Targeted mRNA degradation by deadenylation-independent decapping." *Mol Cell* **15**(1): 5-15.
- Baehrecke, E. H. (2006). "Growth Control: p53, the Guardian Angel of Compensatory Proliferation." *Current Biology* **16**(19): R840-R842.
- Baena-Lopez, L. A. and A. Garcia-Bellido (2006). "Control of growth and positional information by the graded vestigial expression pattern in the wing of *Drosophila melanogaster*." *Proc Natl Acad Sci U S A* **103**(37): 13734-13739.

- Bail, S., M. Swerdel, H. Liu, X. Jiao, L. A. Goff, *et al.* (2010). "Differential regulation of microRNA stability." *RNA* **16**(5): 1032-1039.
- Bakheet, T., M. Frevel, B. R. Williams, W. Greer and K. S. Khabar (2001). "ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins." *Nucleic Acids Res* **29**(1): 246-254.
- Bakheet, T., B. R. G. Williams and K. S. A. Khabar (2003). "ARED 2.0: an update of AU-rich element mRNA database." *Nucleic Acids Research* **31**(1): 421-423.
- Baldock, R. A. (2012). *The Role of the 3'-5' Exoribonuclease Tazman (Dis3) on Wing Development in D.melanogaster*. BSc., University of Sussex.
- Baonza, A. and M. Freeman (2005). "Control of Cell Proliferation in the Drosophila Eye by Notch Signaling." *Developmental Cell* **8**(4): 529-539.
- Bashkurov, V. I., H. Scherthan, J. A. Solinger, J.-M. Buerstedde and W.-D. Heyer (1997). "A Mouse Cytoplasmic Exoribonuclease (mXRN1p) with Preference for G4 Tetraplex Substrates." *The Journal of Cell Biology* **136**(4): 761-773.
- Bassett, Andrew R., C. Tibbit, Chris P. Ponting and J.-L. Liu (2013). "Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System." *Cell Reports* **4**(1): 220-228.
- Berridge, M. J. (1995). "Calcium signalling and cell proliferation." *Bioessays* **17**(6): 491-500.
- Berthelet, J. and L. Dubrez (2013). "Regulation of Apoptosis by Inhibitors of Apoptosis (IAPs)." *Cells* **2**(1): 163-187.
- Blommaert, E. F., J. J. Luiken, P. J. Blommaert, G. M. van Woerkom and A. J. Meijer (1995). "Phosphorylation of ribosomal protein S6 is inhibitory for autophagy in isolated rat hepatocytes." *J Biol Chem* **270**(5): 2320-2326.
- Bonneau, F., J. Basquin, J. Ebert, E. Lorentzen and E. Conti (2009). "The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation." *Cell* **139**(3): 547-559.
- Bouchie, A. (2013). "First microRNA mimic enters clinic." *Nat Biotech* **31**(7): 577-577.
- Bousquet-Antonelli, C., C. Presutti and D. Tollervey (2000). "Identification of a Regulated Pathway for Nuclear Pre-mRNA Turnover." *Cell* **102**(6): 765-775.
- Bouveret, E., G. Rigaut, A. Shevchenko, M. Wilm and B. Séraphin (2000). "A Sm-like protein complex that participates in mRNA degradation." *The EMBO Journal* **19**(7): 1661-1671.
- Brand, A. (1997). P{GawB} insertions.
- Brand, A. H. and N. Perrimon (1993). "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes." *Development* **118**(2): 401-415.
- Braun, J. E., E. Huntzinger, M. Fauser and E. Izaurralde (2011). "GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets." *Mol Cell* **44**(1): 120-133.
- Braun, J. E., V. Truffault, A. Boland, E. Huntzinger, C. T. Chang, *et al.* (2012). "A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation." *Nat Struct Mol Biol* **19**(12): 1324-1331.
- Brennecke, J., D. R. Hipfner, A. Stark, R. B. Russell and S. M. Cohen (2003). "bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila." *Cell* **113**(1): 25-36.
- Briggs, M. W., K. T. Burkard and J. S. Butler (1998). "Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation." *J Biol Chem* **273**(21): 13255-13263.
- Brown, C. E. and A. B. Sachs (1998). "Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation." *Mol Cell Biol* **18**(11): 6548-6559.
- Brown, J. T., X. Bai and A. W. Johnson (2000). "The yeast antiviral proteins Ski2p, Ski3p, and Ski8p exist as a complex in vivo." *RNA* **6**(3): 449-457.
- Brown, J. T., X. Yang and A. W. Johnson (2000). "Inhibition of mRNA turnover in yeast by an xrn1 mutation enhances the requirement for eIF4E binding to eIF4G and for proper capping of transcripts by Ceg1p." *Genetics* **155**(1): 31-42.
- Bryant, P. J. and O. Schmidt (1990). "The genetic control of cell proliferation in Drosophila imaginal discs." *J Cell Sci Suppl* **13**: 169-189.

- Burroughs, A. M., Y. Ando, M. J. L. de Hoon, Y. Tomaru, T. Nishibu, *et al.* (2010). "A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness." Genome Research **20**(10): 1398-1410.
- Cairrão, F., C. Arraiano and S. Newbury (2005). "Drosophila gene tazman, an orthologue of the yeast exosome component Rrp44p/Dis3, is differentially expressed during development." Developmental Dynamics **232**(3): 733-737.
- Calin, G. A., C. D. Dumitru, M. Shimizu, R. Bichi, S. Zupo, *et al.* (2002). "Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia." Proc Natl Acad Sci U S A **99**(24): 15524-15529.
- Cantor, J. R. and D. M. Sabatini (2012). "Cancer Cell Metabolism: One Hallmark, Many Faces." Cancer discovery **2**(10): 881-898.
- Carballo, E. and P. J. Blakeshear (2001). "Roles of tumor necrosis factor- $\alpha$  receptor subtypes in the pathogenesis of the tristetraprolin-deficiency syndrome." Blood **98**(8): 2389-2395.
- Cascino, I., G. Fiucci, G. Papoff and G. Ruberti (1995). "Three functional soluble forms of the human apoptosis-inducing Fas molecule are produced by alternative splicing." J Immunol **154**(6): 2706-2713.
- Chan, C. C., J. Dostie, M. D. Diem, W. Feng, M. Mann, *et al.* (2004). "eIF4A3 is a novel component of the exon junction complex." RNA **10**(2): 200-209.
- Chang, H., J. Lim, M. Ha and V. N. Kim (2014). "TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications." Mol Cell **53**(6): 1044-1052.
- Chang, H. M., R. Triboulet, J. E. Thornton and R. I. Gregory (2013). "A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway." Nature **497**(7448): 244-248.
- Chapman, E. G., D. A. Costantino, J. L. Rabe, S. L. Moon, J. Wilusz, *et al.* (2014). "The Structural Basis of Pathogenic Subgenomic Flavivirus RNA (sfRNA) Production." Science **344**(6181): 307-310.
- Chapman, E. G., S. L. Moon, J. Wilusz and J. S. Kieft (2014). "RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA." eLife **3**: e01892.
- Chapman, M. A., M. S. Lawrence, J. J. Keats, K. Cibulskis, C. Sougnez, *et al.* (2011). "Initial genome sequencing and analysis of multiple myeloma." Nature **471**(7339): 467-472.
- Chatterjee, S., M. Fasler, I. Büssing and H. Großhans (2011). "Target-Mediated Protection of Endogenous MicroRNAs in *C. elegans*." Developmental Cell **20**(3): 388-396.
- Chatterjee, S. and H. Großhans (2009). "Active turnover modulates mature microRNA activity in *Caenorhabditis elegans*." Nature **461**(7263): 546-549.
- Chekulaeva, M., H. Mathys, J. T. Zipprich, J. Attig, M. Colic, *et al.* (2011). "miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs." Nat Struct Mol Biol **18**(11): 1218-1226.
- Chellappan, S. P., S. Hiebert, M. Mudryj, J. M. Horowitz and J. R. Nevins (1991). "The E2F transcription factor is a cellular target for the RB protein." Cell **65**(6): 1053-1061.
- Chen, C., D. A. Ridzon, A. J. Broomer, Z. Zhou, D. H. Lee, *et al.* (2005). "Real-time quantification of microRNAs by stem-loop RT-PCR." Nucleic Acids Research **33**(20): e179-e179.
- Chen, C. Y., R. Gherzi, S. E. Ong, E. L. Chan, R. Raijmakers, *et al.* (2001). "AU binding proteins recruit the exosome to degrade ARE-containing mRNAs." Cell **107**(4): 451-464.
- Chen, C. Y. and A. B. Shyu (1995). "AU-rich elements: characterization and importance in mRNA degradation." Trends Biochem Sci **20**(11): 465-470.
- Chendrimada, T. P., R. I. Gregory, E. Kumaraswamy, J. Norman, N. Cooch, *et al.* (2005). "TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing." Nature **436**(7051): 740-744.
- Cheng, A. M., M. W. Byrom, J. Shelton and L. P. Ford (2005). "Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis." Nucleic Acids Res **33**(4): 1290-1297.

- Chowdhury, A., J. Mukhopadhyay and S. Tharun (2007). "The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs." *RNA* **13**(7): 998-1016.
- Cieply, B. and R. P. Carstens (2015). "Functional roles of alternative splicing factors in human disease." *Wiley Interdisciplinary Reviews: RNA* **6**(3): 311-326.
- Clemens, M. J. (2001). "Initiation factor eIF2 alpha phosphorylation in stress responses and apoptosis." *Prog Mol Subcell Biol* **27**: 57-89.
- Colombani, J., D. S. Andersen and P. Léopold (2012). "Secreted Peptide Dilp8 Coordinates Drosophila Tissue Growth with Developmental Timing." *Science* **336**(6081): 582-585.
- Conlon, I. and M. Raff (1999). "Size Control in Animal Development." *Cell* **96**(2): 235-244.
- Conti, E. and E. Izaurralde (2005). "Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species." *Curr Opin Cell Biol* **17**(3): 316-325.
- Crawford, E. D., J. E. Seaman, A. E. Barber, D. C. David, P. C. Babbitt, *et al.* (2012). "Conservation of caspase substrates across metazoans suggests hierarchical importance of signaling pathways over specific targets and cleavage site motifs in apoptosis." *Cell Death Differ* **19**(12): 2040-2048.
- Czech, B., R. Zhou, Y. Erlich, J. Brennecke, R. Binari, *et al.* (2009). "Hierarchical rules for Argonaute loading in Drosophila." *Molecular Cell* **36**(3): 445-456.
- Das, S. K., U. K. Sokhi, S. K. Bhutia, B. Azab, Z.-z. Su, *et al.* (2010). "Human polynucleotide phosphorylase selectively and preferentially degrades microRNA-221 in human melanoma cells." *Proceedings of the National Academy of Sciences of the United States of America* **107**(26): 11948-11953.
- de Groen, F. L. M., O. Krijgsman, M. Tijssen, L. E. M. Vriend, B. Ylstra, *et al.* (2014). "Gene-dosage dependent overexpression at the 13q amplicon identifies DIS3 as candidate oncogene in colorectal cancer progression." *Genes, Chromosomes and Cancer* **53**(4): 339-348.
- de la Mata, M., D. Gaidatzis, M. Vitanescu, M. B. Stadler, C. Wentzel, *et al.* (2015). "Potent degradation of neuronal miRNAs induced by highly complementary targets." *EMBO Reports* **16**(4): 500-511.
- Decker, C. J. and R. Parker (1993). "A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation." *Genes Dev* **7**(8): 1632-1643.
- DeCoursey, T. E., K. G. Chandy, S. Gupta and M. D. Cahalan (1984). "Voltage-gated K<sup>+</sup> channels in human T lymphocytes: a role in mitogenesis?" *Nature* **307**(5950): 465-468.
- Diaz-Benjumea, F. J. and S. M. Cohen (1993). "Interaction between dorsal and ventral cells in the imaginal disc directs wing development in Drosophila." *Cell* **75**(4): 741-752.
- Diederichs, S. and D. A. Haber (2007). "Dual Role for Argonautes in MicroRNA Processing and Posttranscriptional Regulation of MicroRNA Expression." *Cell* **131**(6): 1097-1108.
- Ding, L., T. J. Ley, D. E. Larson, C. A. Miller, D. C. Koboldt, *et al.* (2012). "Clonal evolution in relapsed acute myeloid leukemia revealed by whole genome sequencing." *Nature* **481**(7382): 506-510.
- Drazkowska, K., R. Tomecki, K. Stodus, K. Kowalska, M. Czarnocki-Cieciura, *et al.* (2013). "The RNA exosome complex central channel controls both exonuclease and endonuclease Dis3 activities in vivo and in vitro." *Nucleic Acids Res* **41**(6): 3845-3858.
- Duffy, J. B. (2002). "GAL4 system in Drosophila: a fly geneticist's Swiss army knife." *Genesis* **34**(1-2): 1-15.
- Dunbar, D. A. and S. J. Baserga (1998). "The U14 snoRNA is required for 2'-O-methylation of the pre-18S rRNA in Xenopus oocytes." *RNA* **4**(2): 195-204.
- Dunkley, T., M. Tucker and R. Parker (2001). "Two related proteins, Edc1p and Edc2p, stimulate mRNA decapping in Saccharomyces cerevisiae." *Genetics* **157**(1): 27-37.
- Dunn, J. G., C. K. Foo, N. G. Belletier, E. R. Gavis and J. S. Weissman (2013). "Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster." *eLife* **2**.

- Dyson, N. (1998). "The regulation of E2F by pRB-family proteins." Genes & Development **12**(15): 2245-2262.
- Eberharter, A. and P. B. Becker (2002). "Histone acetylation: a switch between repressive and permissive chromatin: Second in review series on chromatin dynamics." EMBO Reports **3**(3): 224-229.
- Ebert, M. S., J. R. Neilson and P. A. Sharp (2007). "MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells." Nature methods **4**(9): 10.1038/nmeth1079.
- Eckmann, C. R., C. Rammelt and E. Wahle (2011). "Control of poly(A) tail length." Wiley Interdiscip Rev RNA **2**(3): 348-361.
- Engreitz, J. M., A. Pandya-Jones, P. McDonel, A. Shishkin, K. Sirokman, *et al.* (2013). "The Xist lncRNA Exploits Three-Dimensional Genome Architecture to Spread Across the X Chromosome." Science **341**(6147).
- Esquela-Kerscher, A., P. Trang, J. F. Wiggins, L. Patrawala, A. Cheng, *et al.* (2008). "The let-7 microRNA reduces tumor growth in mouse models of lung cancer." Cell Cycle **7**(6): 759-764.
- Estella, C., G. Rieckhof, M. Calleja and G. Morata (2003). "The role of buttonhead and Sp1 in the development of the ventral imaginal discs of *Drosophila*." Development **130**(24): 5929-5941.
- Fabian, M. R., M. K. Cieplak, F. Frank, M. Morita, J. Green, *et al.* (2011). "miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT." Nat Struct Mol Biol **18**(11): 1211-1217.
- Fabian, M. R. and N. Sonenberg (2012). "The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC." Nat Struct Mol Biol **19**(6): 586-593.
- Faehnle, C. R., J. Walleshauser and L. Joshua-Tor (2014). "Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway." Nature **514**(7521): 252-256.
- Fan, Y. and A. Bergmann (2008). "Apoptosis-induced compensatory proliferation. The Cell is dead. Long live the Cell!" Trends in cell biology **18**(10): 467-473.
- Farh, K. K.-H., A. Grimson, C. Jan, B. P. Lewis, W. K. Johnston, *et al.* (2005). "The Widespread Impact of Mammalian MicroRNAs on mRNA Repression and Evolution." Science **310**(5755): 1817-1821.
- Fingar, D. C. and J. Blenis (2004). "Target of rapamycin (TOR): an integrator of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression." Oncogene **23**(18): 3151-3171.
- Flynt, A. S., J. C. Greimann, W. J. Chung, C. D. Lima and E. C. Lai (2010). "MicroRNA biogenesis via splicing and exosome-mediated trimming in *Drosophila*." Mol Cell **38**(6): 900-907.
- Forero, M. G., A. R. Learte, S. Cartwright and A. Hidalgo (2010). "DeadEasy Mito-Glia: automatic counting of mitotic cells and glial cells in *Drosophila*." PLoS One **5**(5): e10557.
- Friedlander, M. R., W. Chen, C. Adamidi, J. Maaskola, R. Einspanier, *et al.* (2008). "Discovering microRNAs from deep sequencing data using miRDeep." Nat Biotech **26**(4): 407-415.
- Friedländer, M. R., S. D. Mackowiak, N. Li, W. Chen and N. Rajewsky (2012). "miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades." Nucleic Acids Res **40**(1): 37-52.
- Fu, X., N. Fu, S. Guo, Z. Yan, Y. Xu, *et al.* (2009). "Estimating accuracy of RNA-Seq and microarrays with proteomics." BMC Genomics **10**(1): 1-9.
- Gantier, M. P., C. E. McCoy, I. Rusinova, D. Saulep, D. Wang, *et al.* (2011). "Analysis of microRNA turnover in mammalian cells following Dicer1 ablation." Nucleic Acids Research.
- García-Bellido, A. and J. R. Merriam (1971). "Parameters of the wing imaginal disc development of *Drosophila melanogaster*." Dev Biol **24**(1): 61-87.
- García-López, J. and J. del Mazo (2012). "Expression dynamics of microRNA biogenesis during preimplantation mouse development." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1819**(8): 847-854.

- Garneau, N. L., J. Wilusz and C. J. Wilusz (2007). "The highways and byways of mRNA decay." Nat Rev Mol Cell Biol **8**(2): 113-126.
- Gingras, A.-C., S. P. Gygi, B. Raught, R. D. Polakiewicz, R. T. Abraham, *et al.* (1999). "Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism." Genes & Development **13**(11): 1422-1437.
- Goldstrohm, A. C., D. J. Seay, B. A. Hook and M. Wickens (2007). "PUF protein-mediated deadenylation is catalyzed by Ccr4p." J Biol Chem **282**(1): 109-114.
- Goldstrohm, A. C. and M. Wickens (2008). "Multifunctional deadenylase complexes diversify mRNA control." Nat Rev Mol Cell Biol **9**(4): 337-344.
- Gonzalez-Gaitan, M., M. P. Capdevila and A. Garcia-Bellido (1994). "Cell proliferation patterns in the wing imaginal disc of *Drosophila*." Mech Dev **46**(3): 183-200.
- Görnemann, J., K. M. Kotovic, K. Hujer and K. M. Neugebauer (2005). "Cotranscriptional Spliceosome Assembly Occurs in a Stepwise Fashion and Requires the Cap Binding Complex." Molecular Cell **19**(1): 53-63.
- Graham, A. C., D. L. Kiss and E. D. Andrulis (2006). "Differential Distribution of Exosome Subunits at the Nuclear Lamina and in Cytoplasmic Foci." Mol Biol Cell **17**(3): 1399-1409.
- Granneman, S. and S. J. Baserga (2004). "Ribosome biogenesis: of knobs and RNA processing." Experimental Cell Research **296**(1): 43-50.
- Graveley, B. R., G. May, A. N. Brooks, J. W. Carlson, L. Chervas, *et al.* (2011). The *D. melanogaster* transcriptome: modENCODE RNA-Seq data for dissected tissues.
- Grima, D. P., M. Sullivan, M. V. Zabolotskaya, C. Browne, J. Seago, *et al.* (2008). "The 5'-3' exoribonuclease pacman is required for epithelial sheet sealing in *Drosophila* and genetically interacts with the phosphatase puckered." Biology of the Cell **100**(12): 687-701.
- Gu, S., L. Jin, Y. Huang, F. Zhang and Mark A. Kay (2012). "Slicing-Independent RISC Activation Requires the Argonaute PAZ Domain." Current Biology **22**(16): 1536-1542.
- Gu, S. and M. A. Kay (2010). "How do miRNAs mediate translational repression?" Silence **1**: 11.
- Gudipati, R. K., Z. Xu, A. Lebreton, B. Séraphin, L. M. Steinmetz, *et al.* (2012). "Extensive degradation of RNA precursors by the exosome in wild type cells." Mol Cell **48**(3): 409-421.
- Guruharsha, K. G., J. F. Rual, B. Zhai, J. Mintseris, P. Vaidya, *et al.* (2011). "A protein complex network of *Drosophila melanogaster*." Cell **147**(3): 690-703.
- Haas, G., S. Cetin, M. Messmer, B. Chane-Woon-Ming, O. Terenzi, *et al.* (2016). "Identification of factors involved in target RNA-directed microRNA degradation." Nucleic Acids Research.
- Halbach, F., P. Reichelt, M. Rode and E. Conti (2013). "The Yeast Ski Complex: Crystal Structure and RNA Channeling to the Exosome Complex." Cell **154**(4): 814-826.
- Halder, G. and R. L. Johnson (2011). "Hippo signaling: growth control and beyond." Development (Cambridge, England) **138**(1): 9-22.
- Han, B. W., J. H. Hung, Z. Weng, P. D. Zamore and S. L. Ameres (2011). "The 3'-to-5' exoribonuclease Nibbler shapes the 3' ends of microRNAs bound to *Drosophila* Argonaute1." Curr Biol **21**(22): 1878-1887.
- Harigaya, Y., B. N. Jones, D. Muhlrud, J. D. Gross and R. Parker (2010). "Identification and Analysis of the Interaction between Edc3 and Dcp2 in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **30**(6): 1446-1456.
- Harigaya, Y., H. Tanaka, S. Yamanaka, K. Tanaka, Y. Watanabe, *et al.* (2006). "Selective elimination of messenger RNA prevents an incidence of untimely meiosis." Nature **442**(7098): 45-50.
- Harrison, M. M., X.-Y. Li, T. Kaplan, M. R. Botchan and M. B. Eisen (2011). "Zelda Binding in the Early *Drosophila melanogaster* Embryo Marks Regions Subsequently Activated at the Maternal-to-Zygotic Transition." PLoS Genet **7**(10): e1002266.

- Hart, T., H. K. Komori, S. LaMere, K. Podshivalova and D. R. Salomon (2013). "Finding the active genes in deep RNA-seq gene expression studies." *BMC Genomics* **14**(1): 1-7.
- Harvey, K. F., C. M. Pflieger and I. K. Hariharan (2003). "The *Drosophila* Mst Ortholog, hippo, Restricts Growth and Cell Proliferation and Promotes Apoptosis." *Cell* **114**(4): 457-467.
- Hasan, A., C. Cotobal, C. D. S. Duncan and J. Mata (2014). "Systematic Analysis of the Role of RNA-Binding Proteins in the Regulation of RNA Stability." *PLoS Genetics* **10**(11): e1004684.
- Hatfield, L., C. A. Beelman, A. Stevens and R. Parker (1996). "Mutations in trans-acting factors affecting mRNA decapping in *Saccharomyces cerevisiae*." *Mol Cell Biol* **16**(10): 5830-5838.
- Henke, J. I., D. Goergen, J. Zheng, Y. Song, C. G. Schüttler, *et al.* (2008). "microRNA-122 stimulates translation of hepatitis C virus RNA." *The EMBO Journal* **27**(24): 3300-3310.
- Heo, I., M. Ha, J. Lim, M. J. Yoon, J. E. Park, *et al.* (2012). "Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs." *Cell* **151**(3): 521-532.
- Hooper, C., D. L. Taylor and J. M. Pockock (2005). "Pure albumin is a potent trigger of calcium signalling and proliferation in microglia but not macrophages or astrocytes." *J Neurochem* **92**(6): 1363-1376.
- Hou, D., M. Ruiz and E. D. Andrulis (2012). "The ribonuclease Dis3 is an essential regulator of the developmental transcriptome." *BMC Genomics* **13**: 359.
- Hwang, H.-W., E. A. Wentzel and J. T. Mendell (2007). "A Hexanucleotide Element Directs MicroRNA Nuclear Import." *Science* **315**(5808): 97-100.
- Ibrahim, F., L. A. Rymarquis, E. J. Kim, J. Becker, E. Balassa, *et al.* (2010). "Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in *Chlamydomonas*." *Proc Natl Acad Sci U S A* **107**(8): 3906-3911.
- Isken, O. and L. E. Maquat (2007). "Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function." *Genes Dev* **21**(15): 1833-1856.
- Iwasaki, S., M. Kobayashi, M. Yoda, Y. Sakaguchi, S. Katsuma, *et al.* (2010). "Hsc70/Hsp90 Chaperone Machinery Mediates ATP-Dependent RISC Loading of Small RNA Duplexes." *Molecular Cell* **39**(2): 292-299.
- Jansson, M. D. and A. H. Lund (2012). "MicroRNA and cancer." *Molecular Oncology* **6**(6): 590-610.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, *et al.* (2012). "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity." *Science* **337**(6096): 816-821.
- Jinek, M., S. M. Coyle and J. A. Doudna (2011). "Coupled 5' nucleotide recognition and processivity in Xrn1-mediated mRNA decay." *Molecular Cell* **41**(5): 600-608.
- Johnson, J. M., J. Castle, P. Garrett-Engele, Z. Kan, P. M. Loerch, *et al.* (2003). "Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays." *Science* **302**(5653): 2141-2144.
- Johnson, S. M., H. Grosshans, J. Shingara, M. Byrom, R. Jarvis, *et al.* (2005). "RAS Is Regulated by the let-7 MicroRNA Family." *Cell* **120**(5): 635-647.
- Jonas, S., M. Christie, D. Peter, D. Bhandari, B. Loh, *et al.* (2014). "An asymmetric PAN3 dimer recruits a single PAN2 exonuclease to mediate mRNA deadenylation and decay." *Nat Struct Mol Biol* **21**(7): 599-608.
- Jonas, S. and E. Izaurralde (2013). "The role of disordered protein regions in the assembly of decapping complexes and RNP granules." *Genes Dev* **27**(24): 2628-2641.
- Jones, C. I. (2011). *Post-Transcriptional Gene Regulation by the Exoribonuclease Pacman*. Doctor of Philosophy, Brighton and Sussex Medical School.
- Jones, C. I., D. P. Grima, J. A. Waldron, S. Jones, H. N. Parker, *et al.* (2013). "The 5'-3' exoribonuclease Pacman (Xrn1) regulates expression of the heat shock protein Hsp67Bc and the microRNA miR-277-3p in *Drosophila* wing imaginal discs." *RNA Biology* **10**(8): 1345-1355.

- Jones, C. I., A. L. Pashler, B. P. Towler, S. R. Robinson and S. F. Newbury (2016). "RNA-seq reveals post-transcriptional regulation of *Drosophila* insulin-like peptide dilp8 and the neuropeptide-like precursor Nplp2 by the exoribonuclease Pacman/XRN1." Nucleic Acids Res **44**(1): 267-280.
- Jones, C. I., M. V. Zabolotskaya, A. J. King, H. J. S. Stewart, G. A. Horne, *et al.* (2012). "Identification of circulating microRNAs as diagnostic biomarkers for use in multiple myeloma." British Journal of Cancer **107**(12): 1987-1996.
- Jones, C. I., M. V. Zabolotskaya and S. F. Newbury (2012). "The 5' → 3' exoribonuclease XRN1/Pacman and its functions in cellular processes and development." Wiley Interdiscip Rev RNA **3**(4): 455-468.
- Kaberdin, V. R., D. Singh and S. Lin-Chao (2011). "Composition and conservation of the mRNA-degrading machinery in bacteria." Journal of Biomedical Science **18**(1): 1-12.
- Kadaba, S., A. Krueger, T. Trice, A. M. Krecic, A. G. Hinnebusch, *et al.* (2004). "Nuclear surveillance and degradation of hypomodified initiator tRNA(Met) in *S. cerevisiae*." Genes & Development **18**(11): 1227-1240.
- Kadowaki, T. C., S. Hitomi, M. Jacobs, E. Kumagai, C. Liang, S. Schneiter, R. Singleton, D. Wisniewska, J. and Tartakoff, A. (1994). "Isolation and characterization of *Saccharomyces cerevisiae* mRNA transport-defective (mtr) mutants [published erratum appears in *J Cell Biol* 1994 Sep;126(6):1627]." The Journal of Cell Biology **126**(3): 649-659.
- Kapranov, P., J. Cheng, S. Dike, D. A. Nix, R. Duttagupta, *et al.* (2007). "RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription." Science **316**(5830): 1484-1488.
- Katoh, T., Y. Sakaguchi, K. Miyauchi, T. Suzuki, S. Kashiwabara, *et al.* (2009). "Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2." Genes Dev **23**(4): 433-438.
- Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, *et al.* (2013). "TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions." Genome Biol **14**(4): R36.
- Kim, M., L. Vasiljeva, O. J. Rando, A. Zhelkovsky, C. Moore, *et al.* (2006). "Distinct Pathways for snoRNA and mRNA Termination." Molecular Cell **24**(5): 723-734.
- Kinoshita, N., M. Goebel and M. Yanagida (1991). "The fission yeast *dis3+* gene encodes a 110-kDa essential protein implicated in mitotic control." Molecular and Cellular Biology **11**(12): 5839-5847.
- Kiriakidou, M., G. S. Tan, S. Lamprinaki, M. De Planell-Saguer, P. T. Nelson, *et al.* (2007). "An mRNA m7G cap binding-like motif within human Ago2 represses translation." Cell **129**(6): 1141-1151.
- Kiss, D. L. and E. D. Andrulis (2010). "Genome-wide analysis reveals distinct substrate specificities of Rps6, Dis3, and core exosome subunits." RNA **16**(4): 781-791.
- Klein, T. and A. M. Arias (1998). "Interactions among Delta, Serrate and Fringe modulate Notch activity during *Drosophila* wing development." Development **125**(15): 2951-2962.
- Komekado, H., H. Yamamoto, T. Chiba and A. Kikuchi (2007). "Glycosylation and palmitoylation of Wnt-3a are coupled to produce an active form of Wnt-3a." Genes Cells **12**(4): 521-534.
- Kornberg, T., I. Sidén, P. O'Farrell and M. Simon (1985). "The engrailed locus of *Drosophila*: In situ localization of transcripts reveals compartment-specific expression." Cell **40**(1): 45-53.
- Krupp, G. (2005). "Stringent RNA quality control using the Agilent 2100 bioanalyzer." Agilent Technologies.
- Kshirsagar, M. and R. Parker (2004). "Identification of Edc3p as an enhancer of mRNA decapping in *Saccharomyces cerevisiae*." Genetics **166**(2): 729-739.
- Kumakura, N., H. Otsuki, M. Tsuzuki, A. Takeda and Y. Watanabe (2013). "Arabidopsis AtRRP44A Is the Functional Homolog of Rrp44/Dis3, an Exosome Component, Is

- Essential for Viability and Is Required for RNA Processing and Degradation." PLoS One **8**(11): e79219.
- Kwak, P. B. and Y. Tomari (2012). "The N domain of Argonaute drives duplex unwinding during RISC assembly." Nat Struct Mol Biol **19**(2): 145-151.
- Kwon, Y., H.-S. Shim, X. Wang and C. Montell (2008). "Control of thermotactic behavior via coupling of a TRP channel to a phospholipase C signaling cascade." Nat Neurosci **11**(8): 871-873.
- LaCava, J., J. Houseley, C. Saveanu, E. Petfalski, E. Thompson, *et al.* (2005). "RNA degradation by the exosome is promoted by a nuclear polyadenylation complex." Cell **121**(5): 713-724.
- Lallet-Daher, H., M. Roudbaraki, A. Bavencoffe, P. Mariot, F. Gackiere, *et al.* (2009). "Intermediate-conductance Ca<sup>2+</sup>-activated K<sup>+</sup> channels (IKCa1) regulate human prostate cancer cell proliferation through a close control of calcium entry." Oncogene **28**(15): 1792-1806.
- Langmead, B. and S. L. Salzberg (2012). "Fast gapped-read alignment with Bowtie 2." Nat Methods **9**(4): 357-359.
- Lebreton, A., R. Tomecki, A. Dziembowski and B. Seraphin (2008). "Endonucleolytic RNA cleavage by a eukaryotic exosome." Nature **456**(7224): 993-996.
- Lee, G., M. A. Bratkowski, F. Ding, A. Ke and T. Ha (2012). "Elastic coupling between RNA degradation and unwinding by an exoribonuclease." Science (New York, N.Y.) **336**(6089): 1726-1729.
- Lee, J. E., J. Y. Lee, J. Trembly, J. Wilusz, B. Tian, *et al.* (2012). "The PARN deadenylase targets a discrete set of mRNAs for decay and regulates cell motility in mouse myoblasts." PLoS Genet **8**(8): e1002901.
- Lee, R. C., R. L. Feinbaum and V. Ambros (1993). "The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*." Cell **75**(5): 843-854.
- Lee, Tong I. and Richard A. Young (2013). "Transcriptional Regulation and Its Misregulation in Disease." Cell **152**(6): 1237-1251.
- Lee, Y., J. Lee, S. Bang, S. Hyun, J. Kang, *et al.* (2005). "Pyrexia is a new thermal transient receptor potential channel endowing tolerance to high temperatures in *Drosophila melanogaster*." Nat Genet **37**(3): 305-310.
- Lemay, J.-F., M. Larochelle, S. Marguerat, S. Atkinson, J. Bähler, *et al.* (2014). "The RNA exosome promotes transcription termination of backtracked RNA polymerase II." Nat Struct Mol Biol **21**(10): 919-926.
- Lemieux, C., S. Marguerat, J. Lafontaine, N. Barbezier, J. Bähler, *et al.* (2011). "A Pre-mRNA Degradation Pathway that Selectively Targets Intron-Containing Genes Requires the Nuclear Poly(A)-Binding Protein." Molecular Cell **44**(1): 108-119.
- Lepple-Wienhues, A., S. Berweck, M. Bohmig, C. P. Leo, B. Meyling, *et al.* (1996). "K<sup>+</sup> channels and the intracellular calcium signal in human melanoma cell proliferation." J Membr Biol **151**(2): 149-157.
- Lewis, B. P., C. B. Burge and D. P. Bartel (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell **120**(1): 15-20.
- Liang, L., L. Qu and Y. Ding (2007). "Protein and mRNA Characterization in Human Colorectal Carcinoma Cell Lines with Different Metastatic Potentials." Cancer Investigation **25**(6): 427-434.
- Lim, J.-t., T. Kuroki, K. Ozaki, H. Kohsaki, T. Yamori, *et al.* (1997). "Isolation of Murine and Human Homologues of the Fission-Yeast *dis3+* Gene Encoding a Mitotic-Control Protein and Its Overexpression in Cancer Cells with Progressive Phenotype." Cancer Research **57**(5): 921-925.
- Liu, N., M. Abe, Leah R. Sabin, G.-J. Hendriks, Ammar S. Naqvi, *et al.* (2011). "The Exoribonuclease Nibbler Controls 3' End Processing of MicroRNAs in *Drosophila*." Current Biology **21**(22): 1888-1893.

- Liu, Q., J. C. Greimann and C. D. Lima (2006). "Reconstitution, activities, and structure of the eukaryotic RNA exosome." *Cell* **127**(6): 1223-1237.
- Liu, X., Q. Zheng, N. Vrettos, M. Maragkakis, P. Alexiou, *et al.* (2014). "A MicroRNA precursor surveillance system in quality control of MicroRNA synthesis." *Mol Cell* **55**(6): 868-879.
- Lorentzen, E., J. Basquin, R. Tomecki, A. Dziembowski and E. Conti (2008). "Structure of the Active Subunit of the Yeast Exosome Core, Rrp44: Diverse Modes of Substrate Recruitment in the RNase II Nuclease Family." *Molecular Cell* **29**(6): 717-728.
- Lorenz, R., S. H. Bernhart, C. Höner zu Siederdisen, H. Tafer, C. Flamm, *et al.* (2011). "ViennaRNA Package 2.0." *Algorithms for Molecular Biology* **6**(1): 1-14.
- Lubas, M., C. K. Damgaard, R. Tomecki, D. Cysewski, T. H. Jensen, *et al.* (2013). "Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA." *EMBO J* **32**(13): 1855-1868.
- Lührmann, R., B. Kastner and M. Bach (1990). "Structure of spliceosomal snRNPs and their role in pre-mRNA splicing." *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1087**(3): 265-292.
- Lukiw, W. J. (2007). "Micro-RNA speciation in fetal, adult and Alzheimer's disease hippocampus." *NeuroReport* **18**(3): 297-300.
- Maden, B. E. and J. M. Hughes (1997). "Eukaryotic ribosomal RNA: the recent excitement in the nucleotide modification problem." *Chromosoma* **105**(7-8): 391-400.
- Mahtani, K. R., M. Brook, J. L. E. Dean, G. Sully, J. Saklatvala, *et al.* (2001). "Mitogen-Activated Protein Kinase p38 Controls the Expression and Posttranslational Modification of Tristetraprolin, a Regulator of Tumor Necrosis Factor Alpha mRNA Stability." *Molecular and Cellular Biology* **21**(19): 6461-6469.
- Makino, D. L., M. Baumgartner and E. Conti (2013). "Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex." *Nature* **495**(7439): 70-75.
- Malecki, M., S. C. Viegas, T. Carneiro, P. Golik, C. Dressaire, *et al.* (2013). "The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway." *EMBO J* **32**(13): 1842-1854.
- Mamolen, M., A. Smith and E. D. Andrulis (2010). "Drosophila melanogaster Dis3 N-terminal domains are required for ribonuclease activities, nuclear localization and exosome interactions." *Nucleic Acids Research* **38**(16): 5507-5517.
- Marioni, J. C., C. E. Mason, S. M. Mane, M. Stephens and Y. Gilad (2008). "RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays." *Genome Research* **18**(9): 1509-1517.
- Martín, F. A. and G. Morata (2006). "Compartments and the control of growth in the Drosophila wing imaginal disc." *Development* **133**(22): 4421-4426.
- Martin, M. (2011). "Cutadapt removes adapter sequences from high-throughput sequencing reads." *2011* **17**(1).
- Martinez, A., M. Sese, J. H. Losa, N. Robichaud, N. Sonenberg, *et al.* (2015). "Phosphorylation of eIF4E Confers Resistance to Cellular Stress and DNA-Damaging Agents through an Interaction with 4E-T: A Rationale for Novel Therapeutic Approaches." *PLoS One* **10**(4): e0123352.
- Maryati, M., B. Airhihen and G. S. Winkler (2015). "The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module." *Biochem J* **469**(1): 169-176.
- Masaki, T., Kyle C. Arend, Y. Li, D. Yamane, David R. McGivern, *et al.* (2015). "miR-122 Stimulates Hepatitis C Virus RNA Synthesis by Altering the Balance of Viral RNAs Engaged in Replication versus Translation." *Cell Host & Microbe* **17**(2): 217-228.
- McIntyre, L. M., K. K. Lopiano, A. M. Morse, V. Amin, A. L. Oberg, *et al.* (2011). "RNA-seq: technical variability and sampling." *BMC Genomics* **12**(1): 1-13.
- Meijer, H. A., Y. W. Kong, W. T. Lu, A. Wilczynska, R. V. Spriggs, *et al.* (2013). "Translational Repression and eIF4A2 Activity Are Critical for MicroRNA-Mediated Gene Regulation." *Science* **340**(6128): 82-85.

- Mitchell, P., E. Petfalski, A. Shevchenko, M. Mann and D. Tollervey (1997). "The Exosome: A Conserved Eukaryotic RNA Processing Complex Containing Multiple 3'→5' Exoribonucleases." *Cell* **91**(4): 457-466.
- Mizushima, N. (2007). "Autophagy: process and function." *Genes Dev* **21**(22): 2861-2873.
- Moon, S. L., B. J. T. Dodd, D. E. Brackney, C. J. Wilusz, G. D. Ebel, *et al.* (2015). "Flavivirus sfRNA suppresses antiviral RNA interference in cultured cells and mosquitoes and directly interacts with the RNAi machinery." *Virology* **485**: 322-329.
- Moore, M. J. and N. J. Proudfoot (2009). "Pre-mRNA processing reaches back to transcription and ahead to translation." *Cell* **136**(4): 688-700.
- Morris, M. R., D. Astuti and E. R. Maher (2013). "Perlman syndrome: overgrowth, Wilms tumor predisposition and DIS3L2." *Am J Med Genet C Semin Med Genet* **163C**(2): 106-113.
- Muhrad, D. and R. Parker (1992). "Mutations affecting stability and deadenylation of the yeast MFA2 transcript." *Genes Dev* **6**(11): 2100-2111.
- Muhrad, D. and R. Parker (2005). "The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p." *EMBO J* **24**(5): 1033-1045.
- Mullen, T. E. and W. F. Marzluff (2008). "Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'." *Genes & Development* **22**(1): 50-65.
- Murakami, H., D. B. Goto, T. Toda, E. S. Chen, S. I. Grewal, *et al.* (2007). "Ribonuclease activity of Dis3 is required for mitotic progression and provides a possible link between heterochromatin and kinetochore function." *PLoS One* **2**(3): e317.
- Nag, A. and J. A. Steitz (2012). "Tri-snRNP-associated proteins interact with subunits of the TRAMP and nuclear exosome complexes, linking RNA decay and pre-mRNA splicing." *RNA Biology* **9**(3): 334-342.
- Nairz, K., C. Rottig, F. Rintelen, E. Zdobnov, M. Moser, *et al.* (2006). "Overgrowth caused by misexpression of a microRNA with dispensable wild-type function." *Dev Biol* **291**(2): 314-324.
- Negre, N., C. D. Brown, L. Ma, C. A. Bristow, S. W. Miller, *et al.* (2011). "A cis-regulatory map of the Drosophila genome." *Nature* **471**(7339): 527-531.
- Neri, G., M. E. Martini-Neri, B. E. Katz and J. M. Opitz (2013). "The perlman syndrome: Familial renal dysplasia with Wilms tumor, fetal gigantism and multiple congenital anomalies." *American Journal of Medical Genetics Part A* **161**(11): 2691-2696.
- Neto-Silva, R. M., B. S. Wells and L. A. Johnston (2009). "Mechanisms of growth and homeostasis in the Drosophila wing." *Annual review of cell and developmental biology* **25**: 197-220.
- Neufeld, T. P., A. F. A. de la Cruz, L. A. Johnston and B. A. Edgar (1998). "Coordination of Growth and Cell Division in the Drosophila Wing." *Cell* **93**(7): 1183-1193.
- Neugebauer, K. M. (2002). "On the importance of being co-transcriptional." *J Cell Sci* **115**(Pt 20): 3865-3871.
- Newbury, S. F. (2006). "Control of mRNA stability in eukaryotes." *Biochemical Society Transactions* **34**(1): 30-34.
- Ng, D., O. Toure, M.-H. Wei, D. C. Arthur, F. Abbasi, *et al.* (2007). "Identification of a novel chromosome region, 13q21.33-q22.2, for susceptibility genes in familial chronic lymphocytic leukemia." *Blood* **109**(3): 916-925.
- Ng, M., F. J. Diaz-Benjumea, J.-P. Vincent, J. Wu and S. M. Cohen (1996). "Specification of the wing by localized expression of wingless protein." *Nature* **381**(6580): 316-318.
- Nienhaus, U., T. Aegerter-Wilmsen and C. M. Aegerter (2012). "In-vivo imaging of the Drosophila wing imaginal disc over time: novel insights on growth and boundary formation." *PLoS One* **7**(10): e47594.
- Nishihara, T., L. Zekri, J. E. Braun and E. Izaurralde (2013). "miRISC recruits decapping factors to miRNA targets to enhance their degradation." *Nucleic Acids Res* **41**(18): 8692-8705.

- Nissan, T., P. Rajyaguru, M. She, H. Song and R. Parker (2010). "Decapping activators in *Saccharomyces cerevisiae* act by multiple mechanisms." *Molecular Cell* **39**(5): 773-783.
- Ohkura, H., Y. Adachi, N. Kinoshita, O. Niwa, T. Toda, *et al.* (1988). "Cold-sensitive and caffeine-supersensitive mutants of the *Schizosaccharomyces pombe* dis genes implicated in sister chromatid separation during mitosis." *The EMBO Journal* **7**(5): 1465-1473.
- Okamura, K., N. Liu and E. C. Lai (2009). "Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes." *Molecular Cell* **36**(3): 431-444.
- Orban, T. I. and E. Izaurralde (2005). "Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome." *RNA* **11**(4): 459-469.
- Ørom, U. A., F. C. Nielsen and A. H. Lund (2008). "MicroRNA-10a Binds the 5'UTR of Ribosomal Protein mRNAs and Enhances Their Translation." *Molecular Cell* **30**(4): 460-471.
- Pang, Y., X.-c. Bai, C. Yan, Q. Hao, Z. Chen, *et al.* (2015). "Structure of the apoptosome: mechanistic insights into activation of an initiator caspase from *Drosophila*." *Genes & Development* **29**(3): 277-287.
- Parrish, J. Z., M. D. Kim, L. Y. Jan and Y. N. Jan (2006). "Genome-wide analyses identify transcription factors required for proper morphogenesis of *Drosophila* sensory neuron dendrites." *Genes Dev* **20**(7): 820-835.
- Petiot, A., E. Ogier-Denis, E. F. Blommaert, A. J. Meijer and P. Codogno (2000). "Distinct classes of phosphatidylinositol 3'-kinases are involved in signaling pathways that control macroautophagy in HT-29 cells." *J Biol Chem* **275**(2): 992-998.
- Polager, S. and D. Ginsberg (2008). "E2F – at the crossroads of life and death." *Trends in cell biology* **18**(11): 528-535.
- Potter, C. J. and T. Xu (2001). "Mechanisms of size control." *Curr Opin Genet Dev* **11**(3): 279-286.
- Preker, P., J. Nielsen, S. Kammler, S. Lykke-Andersen, M. S. Christensen, *et al.* (2008). "RNA exosome depletion reveals transcription upstream of active human promoters." *Science* **322**(5909): 1851-1854.
- Ramsköld, D., E. T. Wang, C. B. Burge and R. Sandberg (2009). "An Abundance of Ubiquitously Expressed Genes Revealed by Tissue Transcriptome Sequence Data." *PLoS Computational Biology* **5**(12): e1000598.
- Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, *et al.* (2013). "Genome engineering using the CRISPR-Cas9 system." *Nat. Protocols* **8**(11): 2281-2308.
- Rehrauer, H., L. Opitz, G. Tan, L. Sieverling and R. Schlapbach (2013). "Blind spots of quantitative RNA-seq: the limits for assessing abundance, differential expression, and isoform switching." *BMC Bioinformatics* **14**: 370.
- Rehwinkel, J. A. N., I. Behm-Ansmant, D. Gatfield and E. Izaurralde (2005). "A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing." *RNA* **11**(11): 1640-1647.
- Reimao-Pinto, M. M., V. Ignatova, T. R. Burkard, J. H. Hung, R. A. Manzenreither, *et al.* (2015). "Uridylation of RNA Hairpins by Tailor Confines the Emergence of MicroRNAs in *Drosophila*." *Mol Cell* **59**(2): 203-216.
- Reis, F. P., A. Barbas, A. A. Klauer-King, B. Tsanova, D. Schaeffer, *et al.* (2013). "Modulating the RNA Processing and Decay by the Exosome: Altering Rps44/Dis3 Activity and End-Product." *PLoS One* **8**(11): e76504.
- Reiter, L. T., L. Potocki, S. Chien, M. Gribskov and E. Bier (2001). "A Systematic Analysis of Human Disease-Associated Gene Sequences In *Drosophila melanogaster*." *Genome Research* **11**(6): 1114-1125.
- Richter, J. D. and N. Sonenberg (2005). "Regulation of cap-dependent translation by eIF4E inhibitory proteins." *Nature* **433**(7025): 477-480.
- Rissland, O. S., S. J. Hong and D. P. Bartel (2011). "MicroRNA destabilization enables dynamic regulation of the miR-16 family in response to cell-cycle changes." *Mol Cell* **43**(6): 993-1004.

- Rissland, O. S., A. Mikulasova and C. J. Norbury (2007). "Efficient RNA Polyuridylation by Noncanonical Poly(A) Polymerases." Molecular and Cellular Biology **27**(10): 3612-3624.
- Rissland, O. S. and C. J. Norbury (2009). "Decapping is preceded by 3[prime] uridylation in a novel pathway of bulk mRNA turnover." Nat Struct Mol Biol **16**(6): 616-623.
- Rogers Jr, G. W., A. A. Komar and W. C. Merrick (2002). eIF4A: The godfather of the DEAD box helicases. Progress in Nucleic Acid Research and Molecular Biology, Academic Press. **Volume 72**: 307-331.
- Rose, A. E., L. Polisenio, J. Wang, M. Clark, A. Pearlman, *et al.* (2011). "Integrative genomics identifies molecular alterations that challenge the linear model of melanoma progression." Cancer Research **71**(7): 2561-2571.
- Salvesen, G. S. and J. M. Abrams (2004). "Caspase activation - stepping on the gas or releasing the brakes? Lessons from humans and flies." Oncogene **23**(16): 2774-2784.
- Sander, J. D. and J. K. Joung (2014). "CRISPR-Cas systems for editing, regulating and targeting genomes." Nat Biotech **32**(4): 347-355.
- Schaeffer, D., A. Clark, A. A. Klauer, B. Tsanova and A. van Hoof (2011). "Functions of the cytoplasmic exosome." Adv Exp Med Biol **702**: 79-90.
- Schaeffer, D., F. P. Reis, S. J. Johnson, C. M. Arraiano and A. van Hoof (2012). "The CR3 motif of Rrp44p is important for interaction with the core exosome and exosome function." Nucleic Acids Research **40**(18): 9298-9307.
- Schaeffer, D., B. Tsanova, A. Barbas, F. P. Reis, E. G. Dastidar, *et al.* (2009). "The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities." Nature structural & molecular biology **16**(1): 56-62.
- Schilke, K., F. Schaefer, R. Waldherr, W. Rohrschneider, C. John, *et al.* (2000). "A case of Perlman syndrome: Fetal gigantism, renal dysplasia, and severe neurological deficits." American Journal of Medical Genetics **91**(1): 29-33.
- Schmidt, K. and J. S. Butler (2013). "Nuclear RNA surveillance: role of TRAMP in controlling exosome specificity." Wiley Interdiscip Rev RNA **4**(2): 217-231.
- Schmidt, M.-J., S. West and C. J. Norbury (2011). "The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation." RNA **17**(1): 39-44.
- Schneider, C., J. T. Anderson and D. Tollervey (2007). "The Exosome Subunit Rrp44 Plays a Direct Role in RNA Substrate Recognition." Molecular Cell **27**(2): 324-331.
- Schneider, C., G. Kudla, W. Wlotzka, A. Tuck and D. Tollervey (2012). "Transcriptome-wide analysis of exosome targets." Mol Cell **48**(3): 422-433.
- Schoenberg, D. R. and L. E. Maquat (2012). "Regulation of cytoplasmic mRNA decay." Nature Reviews. Genetics **13**(4): 246-259.
- Schubert, T., Miriam C. Pusch, S. Diermeier, V. Benes, E. Kremmer, *et al.* (2012). "Df31 Protein and snoRNAs Maintain Accessible Higher-Order Structures of Chromatin." Molecular Cell **48**(3): 434-444.
- Schütz, P., M. Bumann, A. E. Oberholzer, C. Bieniossek, H. Trachsel, *et al.* (2008). "Crystal structure of the yeast eIF4A-eIF4G complex: An RNA-helicase controlled by protein-protein interactions." Proceedings of the National Academy of Sciences of the United States of America **105**(28): 9564-9569.
- Scott, R. C., O. Schuldiner and T. P. Neufeld (2004). "Role and Regulation of Starvation-Induced Autophagy in the Drosophila Fat Body." Developmental Cell **7**(2): 167-178.
- Sethi, P. and W. J. Lukiw (2009). "Micro-RNA abundance and stability in human brain: specific alterations in Alzheimer's disease temporal lobe neocortex." Neurosci Lett **459**(2): 100-104.
- Shaw, G. and R. Kamen (1986). "A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation." Cell **46**(5): 659-667.
- She, M., C. J. Decker, D. I. Svergun, A. Round, N. Chen, *et al.* (2008). "Structural basis of Dcp2 recognition and activation by Dcp1." Molecular Cell **29**(3): 337-349.
- Shi, S. and P. Stanley (2003). "Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways." Proc Natl Acad Sci U S A **100**(9): 5234-5239.

- Shibuya, T., T. O. Tange, N. Sonenberg and M. J. Moore (2004). "eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay." Nat Struct Mol Biol **11**(4): 346-351.
- Shingleton, A. W. (2010). "The regulation of organ size in *Drosophila*: Physiology, plasticity, patterning and physical force." Organogenesis **6**(2): 76-87.
- Slocumb, M. E., J. M. Regalado, M. Yoshizawa, G. G. Neely, P. Masek, *et al.* (2015). "Enhanced Sleep Is an Evolutionarily Adaptive Response to Starvation Stress in *Drosophila*." PLoS One **10**(7): e0131275.
- Slomovic, S. and G. Schuster (2011). "Exonucleases and endonucleases involved in polyadenylation-assisted RNA decay." Wiley Interdiscip Rev RNA **2**(1): 106-123.
- Smith-Bolton, R. K., M. I. Worley, H. Kanda and I. K. Hariharan (2009). "Regenerative growth in *Drosophila* imaginal discs is regulated by Wingless and Myc." Dev Cell **16**(6): 797-809.
- Smith, S. B., D. L. Kiss, E. Turk, A. M. Tartakoff and E. D. Andrulis (2011). "Pronounced and extensive microtubule defects in a *Saccharomyces cerevisiae* DIS3 mutant." Yeast **28**(11): 755-769.
- St Pierre, S. E., L. Ponting, R. Stefancsik and P. McQuilton (2014). "FlyBase 102--advanced approaches to interrogating FlyBase." Nucleic Acids Res **42**(Database issue): D780-788.
- Staals, R. H., A. W. Bronkhorst, G. Schilders, S. Slomovic, G. Schuster, *et al.* (2010). "Dis3-like 1: a novel exoribonuclease associated with the human exosome." EMBO J **29**(14): 2358-2367.
- Stanczyk, J., D. M. L. Pedrioli, F. Brentano, O. Sanchez-Pernaute, C. Kolling, *et al.* (2008). "Altered expression of MicroRNA in synovial fibroblasts and synovial tissue in rheumatoid arthritis." Arthritis & Rheumatism **58**(4): 1001-1009.
- Stark, A., J. Brennecke, N. Bushati, R. B. Russell and S. M. Cohen (2005). "Animal MicroRNAs Confer Robustness to Gene Expression and Have a Significant Impact on 3'UTR Evolution." Cell **123**(6): 1133-1146.
- Steinbrunn, T., T. Stuhmer, S. Gattenlohner, A. Rosenwald, A. Mottok, *et al.* (2011). "Mutated RAS and constitutively activated Akt delineate distinct oncogenic pathways, which independently contribute to multiple myeloma cell survival." Blood **117**(6): 1998-2004.
- Steller, H. (2008). "Regulation of apoptosis in *Drosophila*." Cell Death Differ **15**(7): 1132-1138.
- Stoecklin, G., T. Mayo and P. Anderson (2006). "ARE-mRNA degradation requires the 5'-3' decay pathway." EMBO Rep **7**(1): 72-77.
- Symmons, M. F. and B. F. Luisi (2009). "Through Ancient Rings Thread Programming Strings." Structure(London, England:1993) **17**(11): 1429-1431.
- Szczepinska, T., K. Kalisiak, R. Tomecki, A. Labno, L. S. Borowski, *et al.* (2015). "DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts." Genome Res **25**(11): 1622-1633.
- Tazi, J., N. Bakkour and S. Stamm (2009). "Alternative splicing and disease." Biochim Biophys Acta **1792**(1): 14-26.
- Temme, C., L. Zhang, E. Kremmer, C. Ihling, A. Chartier, *et al.* (2010). "Subunits of the *Drosophila* CCR4-NOT complex and their roles in mRNA deadenylation." RNA **16**(7): 1356-1370.
- Tharun, S., D. Muhrad, A. Chowdhury and R. Parker (2005). "Mutations in the *Saccharomyces cerevisiae* LSM1 gene that affect mRNA decapping and 3' end protection." Genetics **170**(1): 33-46.
- Tharun, S. and R. Parker (2001). "Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs." Mol Cell **8**(5): 1075-1083.
- Thomas, M. F., S. Abdul-Wajid, M. Panduro, J. E. Babiarz, M. Rajaram, *et al.* (2012). "Eri1 regulates microRNA homeostasis and mouse lymphocyte development and antiviral function." Blood **120**(1): 130-142.

- Thomas, M. P., X. Liu, J. Whangbo, G. McCrossan, K. B. Sanborn, *et al.* (2015). "Apoptosis Triggers Specific, Rapid, and Global mRNA Decay with 3' Uridylated Intermediates Degraded by DIS3L2." Cell Rep **11**(7): 1079-1089.
- Thornton, J. E., P. Du, L. Jing, L. Sjekloca, S. Lin, *et al.* (2014). "Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4)." Nucleic Acids Research.
- Thum, T., D. Catalucci and J. Bauersachs (2008). "MicroRNAs: novel regulators in cardiac development and disease." Cardiovascular Research **79**(4): 562-570.
- Tomecki, R., K. Drakowska, I. Kucinski, K. Stodus, R. J. Szczesny, *et al.* (2014). "Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target." Nucleic Acids Research **42**(2): 1270-1290.
- Tomecki, R., M. S. Kristiansen, S. Lykke-Andersen, A. Chlebowski, K. M. Larsen, *et al.* (2010). "The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L." The EMBO Journal **29**(14): 2342-2357.
- Towler, B. P., C. I. Jones and S. F. Newbury (2015). "Mechanisms of regulation of mature miRNAs." Biochem Soc Trans **43**(6): 1208-1214.
- Towler, B. P., C. I. Jones, S. C. Viegas, P. Apura, J. A. Waldron, *et al.* (2015). "The 3'-5' exoribonuclease Dis3 regulates the expression of specific microRNAs in *Drosophila* wing imaginal discs." RNA Biol **12**(7): 728-741.
- Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, *et al.* (2012). "Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks." Nat. Protocols **7**(3): 562-578.
- Tritschler, F., A. Eulalio, V. Truffault, M. D. Hartmann, S. Helms, *et al.* (2007). "A Divergent Sm Fold in EDC3 Proteins Mediates DCP1 Binding and P-Body Targeting." Molecular and Cellular Biology **27**(24): 8600-8611.
- Twomey, C. and J. V. McCarthy (2005). "Pathways of apoptosis and importance in development." Journal of Cellular and Molecular Medicine **9**(2): 345-359.
- Urrego, D., A. P. Tomczak, F. Zahed, W. Stuhmer and L. A. Pardo (2014). "Potassium channels in cell cycle and cell proliferation." Philos Trans R Soc Lond B Biol Sci **369**(1638): 20130094.
- Ustianenko, D., D. Hrossova, D. Potesil, K. Chalupnikova, K. Hrazdilova, *et al.* (2013). "Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs." RNA **19**(12): 1632-1638.
- van Dijk, E. L., G. Schilders and G. J. Pruijn (2007). "Human cell growth requires a functional cytoplasmic exosome, which is involved in various mRNA decay pathways." RNA **13**(7): 1027-1035.
- van Hoof, A., P. Lennertz and R. Parker (2000). "Yeast Exosome Mutants Accumulate 3'-Extended Polyadenylated Forms of U4 Small Nuclear RNA and Small Nucleolar RNAs." Molecular and Cellular Biology **20**(2): 441-452.
- Vanáčová, t. p., J. Wolf, G. Martin, D. Blank, S. Dettwiler, *et al.* (2005). "A New Yeast Poly(A) Polymerase Complex Involved in RNA Quality Control." PLoS Biol **3**(6): e189.
- Vance, K. W. and C. P. Ponting (2014). "Transcriptional regulatory functions of nuclear long noncoding RNAs." Trends in Genetics **30**(8): 348-355.
- Vasudevan, S., Y. Tong and J. A. Steitz (2007). "Switching from Repression to Activation: MicroRNAs Can Up-Regulate Translation." Science **318**(5858): 1931-1934.
- Viphakone, N., F. Voisinet-Hakil and L. Minvielle-Sebastia (2008). "Molecular dissection of mRNA poly(A) tail length control in yeast." Nucleic Acids Res **36**(7): 2418-2433.
- Wahle, E. (1995). "Poly(A) tail length control is caused by termination of processive synthesis." J Biol Chem **270**(6): 2800-2808.
- Wahle, E. and G. S. Winkler (2013). "RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes." Biochim Biophys Acta **1829**(6-7): 561-570.
- Waldron, J. A. (2014). The role of the exoribonuclease Pacman/Xrn1 in wing development in *Drosophila*. Doctor of Philosophy, Brighton and Sussex Medical School.

- Waldron, J. A., C. I. Jones, B. P. Towler, A. L. Pashler, D. P. Grima, *et al.* (2015). "Xrn1/Pacman affects apoptosis and regulates expression of hid and reaper." *Biol Open* **4**(5): 649-660.
- Wang, H.-W., C. Noland, B. Siridechadilok, D. W. Taylor, E. Ma, *et al.* (2009). "Structural insights into RNA processing by the human RISC-loading complex." *Nat Struct Mol Biol* **16**(11): 1148-1153.
- Wang, X., A. Flynn, A. J. Waskiewicz, B. L. Webb, R. G. Vries, *et al.* (1998). "The phosphorylation of eukaryotic initiation factor eIF4E in response to phorbol esters, cell stresses, and cytokines is mediated by distinct MAP kinase pathways." *J Biol Chem* **273**(16): 9373-9377.
- Wasmuth, E. V., K. Januszyk and C. D. Lima (2014). "Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA." *Nature* **511**(7510): 435-439.
- Wasmuth, E. V. and C. D. Lima (2012). "Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel." *Mol Cell* **48**(1): 133-144.
- Weigmann, K., S. M. Cohen and C. F. Lehner (1997). "Cell cycle progression, growth and patterning in imaginal discs despite inhibition of cell division after inactivation of *Drosophila* Cdc2 kinase." *Development* **124**(18): 3555-3563.
- Weinkove, D. and S. J. Leivers (2000). "The genetic control of organ growth: insights from *Drosophila*." *Current Opinion in Genetics & Development* **10**(1): 75-80.
- Westholm, J. O., E. Ladewig, K. Okamura, N. Robine and E. C. Lai (2012). "Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs." *RNA* **18**(2): 177-192.
- Widmann, T. J. and C. Dahmann (2009). "Wingless signaling and the control of cell shape in *Drosophila* wing imaginal discs." *Dev Biol* **334**(1): 161-173.
- Wightman, B., I. Ha and G. Ruvkun (1993). "Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*." *Cell* **75**(5): 855-862.
- Wilczynska, A. and M. Bushell (2015). "The complexity of miRNA-mediated repression." *Cell Death Differ* **22**(1): 22-33.
- Wilusz, J. E., H. Sunwoo and D. L. Spector (2009). "Long noncoding RNAs: functional surprises from the RNA world." *Genes Dev* **23**(13): 1494-1504.
- Wilusz, J. E., J. M. Whipple, E. M. Phizicky and P. A. Sharp (2011). "tRNAs marked with CCACCA are targeted for degradation." *Science (New York, N.Y.)* **334**(6057): 817-821.
- Winnebeck, E. C., C. D. Millar and G. R. Warman (2010). "Why Does Insect RNA Look Degraded?" *J Insect Sci* **10**.
- Winter, J. and S. Diederichs (2011). "Argonaute proteins regulate microRNA stability: Increased microRNA abundance by Argonaute proteins is due to microRNA stabilization." *RNA Biol* **8**(6): 1149-1157.
- Wlotzka, W., G. Kudla, S. Granneman and D. Tollervey (2011). "The nuclear RNA polymerase II surveillance system targets polymerase III transcripts." *EMBO J* **30**(9): 1790-1803.
- Wolff, T. and D. F. Ready (1991). "Cell death in normal and rough eye mutants of *Drosophila*." *Development* **113**(3): 825-839.
- Wyers, F., M. Rougemaille, G. Badis, J. C. Rousselle, M. E. Dufour, *et al.* (2005). "Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase." *Cell* **121**(5): 725-737.
- Wyman, S. K., E. C. Knouf, R. K. Parkin, B. R. Fritz, D. W. Lin, *et al.* (2011). "Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity." *Genome Research* **21**(9): 1450-1461.
- Xia, T., Q. Liao, X. Jiang, Y. Shao, B. Xiao, *et al.* (2014). "Long noncoding RNA associated-competing endogenous RNAs in gastric cancer." *Scientific Reports* **4**: 6088.
- Xue, Z., H. Yuan, J. Guo and Y. Liu (2012). "Reconstitution of an Argonaute-Dependent Small RNA Biogenesis Pathway Reveals a Handover Mechanism Involving the RNA Exosome and the Exonuclease QIP." *Molecular Cell* **46**(3): 299-310.

- Yamashita, A., T. C. Chang, Y. Yamashita, W. Zhu, Z. Zhong, *et al.* (2005). "Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover." Nat Struct Mol Biol **12**(12): 1054-1063.
- Yao, B., L. B. La, Y.-C. Chen, L.-J. Chang and E. K. L. Chan (2012). "Defining a new role of GW182 in maintaining miRNA stability." EMBO Reports **13**(12): 1102-1108.
- Yi, W., P. M. Clark, D. E. Mason, M. C. Keenan, C. Hill, *et al.* (2012). "PFK1 Glycosylation Is a Key Regulator of Cancer Cell Growth and Central Metabolic Pathways." Science **337**(6097): 975-980.
- Zabolotskaya, Maria V., Dominic P. Grima, M.-D. Lin, T.-B. Chou and Sarah F. Newbury (2008). "The 5'–3' exoribonuclease Pacman is required for normal male fertility and is dynamically localized in cytoplasmic particles in *Drosophila* testis cells." Biochemical Journal **416**(3): 327-335.
- Zekri, L., D. Kuzuoğlu-Öztürk and E. Izaurralde (2013). "GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation." EMBO J **32**(7): 1052-1065.
- Zhang, K., N. Dion, B. Fuchs, T. Damron, S. Gitelis, *et al.* (2002). "The human homolog of yeast SEP1 is a novel candidate tumor suppressor gene in osteogenic sarcoma." Gene **298**(2): 121-127.
- Zhang, L., L. Ding, T. H. Cheung, M.-Q. Dong, J. Chen, *et al.* (2007). "Systematic identification of miRISC proteins, miRNAs, and their mRNA targets in *C. elegans* by their interactions with GW182 family proteins AIN-1 and AIN-2." Molecular Cell **28**(4): 598-613.
- Zhang, W., C. Murphy and L. E. Sieburth (2010). "Conserved RNasell domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes." Proceedings of the National Academy of Sciences of the United States of America **107**(36): 15981-15985.
- Zhao, B., L. Li, Q. Lei and K.-L. Guan (2010). "The Hippo–YAP pathway in organ size control and tumorigenesis: an updated version." Genes & Development **24**(9): 862-874.
- Zhao, S., W.-P. Fung-Leung, A. Bittner, K. Ngo and X. Liu (2014). "Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells." PLoS One **9**(1): e78644.
- Zheng, D., N. Ezzeddine, C. Y. Chen, W. Zhu, X. He, *et al.* (2008). "Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells." J Cell Biol **182**(1): 89-101.
- Zhou, Y., J. Zhu, G. Schermann, C. Ohle, K. Bendrin, *et al.* (2015). "The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome." Nature Communications **6**: 7050.
- Ziebarth, J. D., A. Bhattacharya and Y. Cui (2012). "Integrative Analysis of Somatic Mutations Altering MicroRNA Targeting in Cancer Genomes." PLoS One **7**(10): e47137.
- Zuo, Y. and M. P. Deutscher (2001). "Exoribonuclease superfamilies: structural analysis and phylogenetic distribution." Nucleic Acids Research **29**(5): 1017-1026.

## Appendix: Comparative analysis of Pcm and Dis3L2 sequencing.

<u>Opposite direction</u>			<u>Same direction</u>		<u>Dis3L2 specific</u>	
<i>Acp36DE</i>	CG6018	<i>Acer</i>	CG31547	<i>gsb-n</i>	<i>Acp95EF</i>	CG7777
CG11529	CG6125	<i>b6</i>	CG31808	<i>Hsp70Bb</i>	<i>AsnS</i>	CG7991
CG11670	CG8292	<i>bam</i>	CG31875	<i>iav</i>	<i>Asph</i>	CG9626
CG13707	CG8398	<i>btd</i>	CG32040	<i>ktub</i>	<i>beat-Ila</i>	CG9760
CG13937	CG9413	<i>cas</i>	CG32248	<i>Lim1</i>	<i>c(2)M</i>	<i>CheA29a</i>
CG14591	CG9743	<i>Cdk5α</i>	CG32364	<i>mGluR</i>	CG10738	CR43423
CG1461	<i>Cpr47Eb</i>	CG10178	CG32783	<i>Mhc</i>	CG11382	CR43478
CG15418	CR43252	CG11261	CG33178	<i>mthl14</i>	CG11413	CR43609
CG18598	CR43907	CG11380	CG3457	<i>mtsh</i>	CG13229	CR43653
CG1941	CR45182	CG11842	CG42237	<i>noe</i>	CG13239	CR44093
CG2837	CR45700	CG11843	CG42867	<i>nvd</i>	CG13875	CR44165
CG2854	<i>dysc</i>	CG1231	CG5527	<i>Or22b</i>	CG14451	CR44662
CG30088	<i>eater</i>	CG12493	CG5867	<i>png</i>	CG14668	CR45001
CG32793	<i>ey</i>	CG12560	CG6026	<i>Pxt</i>	CG15657	CR45234
CG33460	<i>ham</i>	CG12950	CG6142	<i>pyx</i>	CG15880	CR45327
CG4133	<i>lectin-46Cb</i>	CG12990	CG6520	<i>santa-maria</i>	CG15912	CR45414
CG42369	<i>Nep4</i>	CG13029	CG7365	<i>Scr</i>	CG16940	CR45604
CG4259	<i>NimC1</i>	CG13050	CG7560	<i>sna</i>	CG1695	CR45754
CG42711	<i>scro</i>	CG13405	CG7763	<i>Sp1</i>	CG18467	<i>cuff</i>
CG43446	<i>sev</i>	CG13577	CG8773	<i>Syt7</i>	CG18622	<i>Cyp4g1</i>
CG4927	<i>tn</i>	CG13618	CG9259	<i>Tace</i>	CG2678	<i>Cyp6d4</i>
CG5062	<i>upd2</i>	CG14079	CG9692	<i>Tep1</i>	CG3085	<i>dpr19</i>
CG5280	<i>yellow-k</i>	CG14234	CG9737	<i>Ubx</i>	CG3119	<i>eg</i>
CG5890		CG14280	CG9919	<i>unc-4</i>	CG31676	<i>Faa</i>
CG5984		CG14636	<i>Cpr72Ea</i>	<i>wrapper</i>	CG31913	<i>Gyc88E</i>
		CG14787	CR18228	<i>yellow-h</i>	CG32225	<i>Hsp23</i>
		CG14868	CR43460	<i>Yp3</i>	CG3523	<i>Jhl-26</i>
		CG1503	CR44367		CG4017	<i>lbl</i>
		CG15096	CR44939		CG42269	<i>Mst36Fa</i>
		CG15545	CR45549		CG44437	<i>PH4alphaMP</i>
		CG16799	CR45683		CG4950	<i>Rab3</i>
		CG16800	<i>Cyt-c-d</i>		CG5078	<i>Rim</i>
		CG16884	<i>Cyt-c-p</i>		CG5321	<i>sing</i>
		CG17190	<i>del</i>		CG6126	<i>SoxN</i>
		CG17599	<i>Drep-2</i>		CG6696	<i>Trf4-2</i>
		CG1835	<i>fd96Ca</i>		CG7173	<i>Tsp68C</i>
		CG2145	<i>fd96Cb</i>		CG7214	<i>whip</i>
		CG30269	<i>fs(1)Yb</i>		CG7349	
		CG31053	<i>Gal</i>		CG7368	
		CG31373	<i>gd</i>		CG7607	