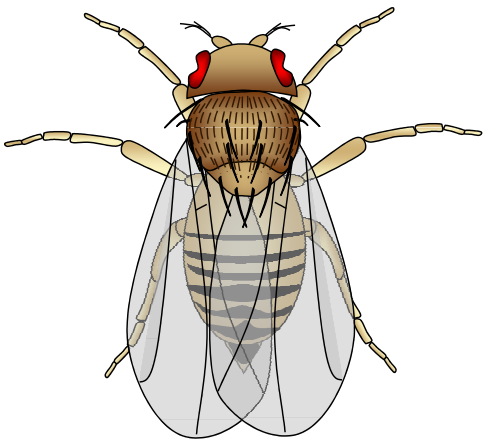
# Post-transcriptional gene regulation by the exoribonuclease Pacman

Christopher Iain Jones

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

May 2011



# Abstract

The gene *pacman* (*pcm*) in *Drosophila melanogaster* encodes the exoribonuclease XRN1, which is highly conserved across eukaryotes and is the only known cytoplasmic exoribonuclease that degrades RNA in the 5' – 3' direction. Hypomorphic mutations to *pacman* have previously been shown cause developmental phenotypes, particularly during wing and thorax development.

The focus of this thesis was twofold. Firstly, to create a null *pacman* allele and associated control lines to further characterise the phenotypes of *pcm*. Two new alleles were created, one of which was amorphic ($pcm^{14}$). $pcm^{14}$ is 100% lethal, and flies die during pupation. The wing imaginal discs of $pcm^{14}$ larvae are less than half the size of those in wild-type larvae at the same stage (3rd instar). It was also found that wing imaginal discs in the hypomorphic mutant $pcm^{5}$ are significantly smaller than wild-type, by almost 20%. Therefore, *pcm* appears to play a role in cell proliferation or apoptosis during the growth of wing imaginal discs. Along with $pcm^{14}$, a new deficiency that includes *pcm* was created using a DrosDel Rearrangement Screen. The 17,963bp *Df(1)ED7452* deficiency is >13 times smaller than the two other publically available deficiencies that include *pcm*.

The second focus of this thesis was to identify RNAs that are likely to be specifically targeted by Pacman. This was achieved using mRNA and miRNA arrays to compare wild-type and $pcm^{5}$ wing imaginal discs. Surprisingly few mRNAs or miRNAs were differentially expressed in $pcm^{5}$ wing imaginal discs, but one with potential significance during later wing development was *simjang*. Simjang (p66) is a highly conserved protein that forms part of the Nucleosome Remodelling and Deacetylase (NuRD) complex, which is recruited to DNA by repressors to downregulate genes by histone deacetylation. s*imjang* mRNA was upregulated 3.8-fold in $pcm^{5}$ wing discs. However, *simjang* pre-mRNA was not upregulated, indicating that the *simjang* mRNA upregulation was due to a post-transcriptional effect, and thus strongly suggesting that *simjang* mRNA is a direct target of Pacman in this tissue. A number of mRNAs and one miRNA were downregulated in $pcm^{5}$ mutant wing discs and the pre-mRNA level of one, *CG32364*, has been tested. *pre-CG32364* was transcriptionally downregulated by the same amount as mature *CG32364* (-8.3-fold), and thus could be a target of *simjang*.

The results from the mRNA and miRNA arrays were also compared to equivalent arrays performed on adult wild-type and $pcm^{5}$ testes. Only two transcripts were differentially regulated in both tissues, showing that the effect of Pacman is tissue specific.

# Candidate's declaration

I declare that the research 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 degree.

Signed:

Dated:

# Acknowledgements

I would like to take this opportunity to thank the following people who have helped and supported me during my time as a PhD student:

Dr. Sarah Newbury, my supervisor, for giving me the opportunity to do this PhD, and for her support and advice from beginning to end.

Dr. Dom Grima, for being a great friend and for all the fly knowledge I have extracted from him over the past few years (I promise I'll put it back when I've finished with it).

Dr. Maria Zabolotskaya, for being a shining example of how everything *should* be done in a lab.

Clare Rizzo-Singh and Dawn Field, for maintaining the steady flow of fly food and consumables to prevent the lab descending into anarchy.

Dr. Robert Ray, for his genetics advice and for knowing everything about *Drosophila*.

Lastly, I would like to thank my wife Zoe, for putting up with my long monologues about my work, and my parents for giving me the opportunity to get this far in life without a real job.

# Contents

# Table of Figures

# Commonly used abbreviations

| Abbreviation | Definition |
|---|---|
| XRN1/Xrn-1/ Xrn1p/Pacman | Exoribonuclease-1. |
| Pcm | Pacman protein. |
| *pcm* | *pacman* gene (gDNA), *pacman* mRNA, or *pacman* cDNA. |
| XRN2/Rat1 | Exoribonuclease-2. |
| *simj* | *simjang*/p66. |
| *Adh* | *Alcohol dehydrogenase.* |
| *EP1526* | *P{w$^{+mC}$=EP}EP1526.* |
| *FM7i* | *FM7i*, *P{w$^{+mC}$=ActGFP}JMR3.* |
| *TY* | *T(1;Y)B92*. |
| sqPCR | Semi-quantitative Reverse Transcriptase PCR. |
| qPCR | Comparative quantitative Reverse Transcriptase PCR. |
| miRNA or miR | microRNA. |
| RISC | RNA-Induced Silencing Complex. |
| NMD | Nonsense Mediated Decay |
| PTC | Premature Termination Codon. |
| EJC | Exon Junction Complex. |
| ARE | AU-Rich Element. |
| PABP | Poly(A)-Binding Protein. |
| NuRD complex | Nucleosome Remodelling and Deacetylase complex. |
| DrosDel RS | DrosDel Rearrangement Screen. |

# 1 Chapter 1: mRNA degradation and post-transcriptional gene regulation

## 1.1 The flow of information from genes to proteins

The central dogma of molecular biology describes the flow of information within all living cells (Crick, 1970). The information required to create life, and allow its reproduction, is stored in the sequence of nucleotides that make up DNA. The information contained in individual genes in DNA is used to produce proteins and RNAs that each form a tiny, yet often essential, part of the cellular machinery. During the production of proteins, the information from DNA is transcribed into a complementary sequence of nucleotides to make messenger RNA (mRNA), and then translated into a defined sequence of amino acids that form the protein. Most genes encode information that ultimately becomes functional when translated into protein, but some also code for RNAs, such as ribosomal RNAs (rRNAs), which themselves function within the cell. The information within DNA can also be passed to new DNA molecules in new cells to allow life to continue. In normal cellular function, information does not flow from RNA to RNA or RNA to DNA, but this can occur in some circumstances, such as infection by a retrovirus. Transfer of information from protein to DNA or protein to RNA does not occur in nature, likely because the process would involve an unavoidable loss of information.

The number of genes in *Homo sapiens* is estimated to be 29,201 in release 62 (April 2011) of the Ensembl database of eukaryotic genomes (Flicek *et al.*, 2010). This includes 20,930 known protein-coding genes and 8,271 RNA genes. There are 13,781 known protein-coding genes in *Drosophila melanogaster* and 993 RNA genes, 14,774 in total. However, simple expression of thousands of genes is not sufficient to make a human, fly or any other organism. The product of almost every gene is under tight spatial and temporal control at multiple levels to ensure it is present at the correct quantity at the correct time, as misexpression of a single gene can have catastrophic consequences and reduce evolutionary fitness to zero. There are many steps in the process of creating a functioning protein from a gene, and there are examples of regulation occurring at almost every level. Figure 1.1 shows the stages that lie between DNA and protein function, and gives examples of processes that can be regulated to alter the ultimate level of protein function. The focus of this thesis will be regulation of mature mRNA levels after they have been exported from the nucleus.

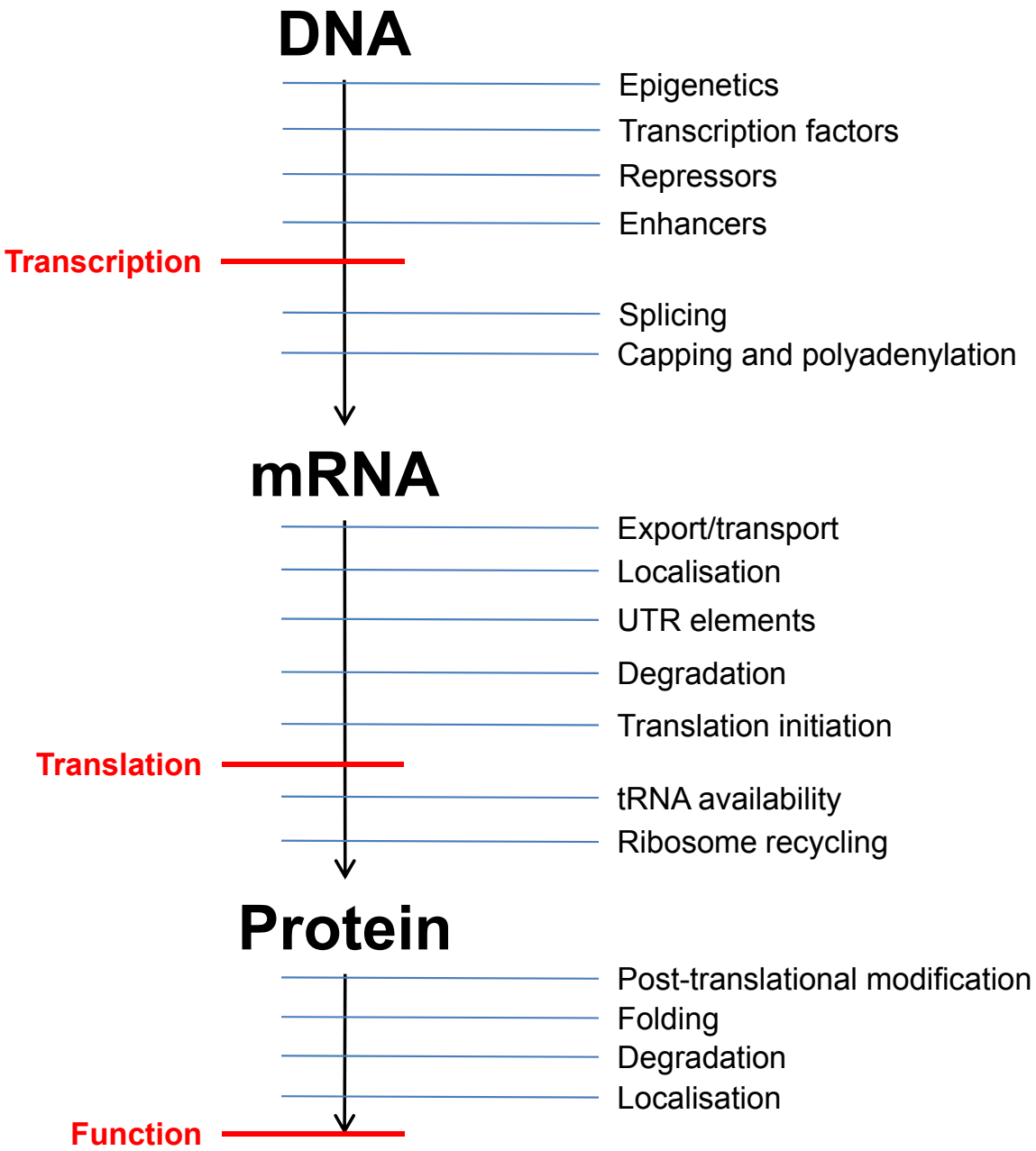



**Figure 1.1**: Examples of potential levels of regulation between genes and their functional protein.

## 1.2 mRNA conformation and translation initiation

Actively translated mRNAs are thought to exist in a circular conformation (Wells *et al.*, 1998). This conformation is caused by interaction between factors that bind to the 3' poly(A) tail and the 5' 7-methylguanosine cap (Figure 1.2), which are attached co-transcriptionally. eIF-4E binds directly to the cap of the mRNA and Poly(A)-binding protein (PABP) binds to the poly(A) tail. eIF-4G binds to both PABP and eIF-4E, bringing the cap and tail into close proximity. The eIF-4F complex (eIF-4E, eIF-4G and other factors such as eIF-4A and eIF-4B) allows binding of the smaller ribosomal subunit, which scans along the mRNA to find the AUG codon that marks the start of translation (Curry *et al.*, 2009, Sonenberg and Hinnebusch, 2009). The circular conformation of the mRNA allows efficient recycling of ribosomes that have completed translation, and confers protection to the transcript. For degradation of the transcript to occur, the 5' cap or the poly(A) tail must be removed, or the transcript must be internally cleaved, to allow access to degradation enzymes.

## 1.3 Preventing translation by degradation of mRNAs

mRNA degradation begins in two main ways (Figure 1.3), either by deadenylation of the transcript or by internal cleavage (Garneau *et al.*, 2007). The enzymes involved in deadenylation vary slightly between organisms, but follow a similar process. The poly(A) tail length of mRNA is dynamic, modulated by enzymes such as PAN2-PAN3 (Brown *et al.*, 1996), but once the decision is made to degrade a transcript, it is removed by the CCR4-NOT complex or PARN. CCR4-NOT is the main deadenylase in *Saccharomyces cerevisiae* (Tucker *et al.*, 2001, Tucker *et al.*, 2002) and has also been shown to deadenylate transcripts in mammalian cells (Yamashita *et al.*, 2005). PARN is a cap-dependent deadenylase, but is inhibited by cap binding proteins (Martınez *et al.*, 2001). PARN has been shown to be important in plants (Reverdatto *et al.*, 2004) and *Xenopus laevis* (Korner *et al.*, 1998) and has been found in mammals and some insects, but is not present in *D. melanogaster* (Temme *et al.*, 2004). The deadenylase activities of PARN and CCR4-NOT are both inhibited by PABP. After deadenylation, the mRNA can be degraded in two ways. Without the poly(A) tail, the 3' of the transcript is directly vulnerable to the exosome, a large multi-subunit complex that degrades RNAs in the 3'-5' direction (Schmid and Jensen, 2008, Lykke-Andersen *et al.*, 2009). After degradation by the exosome, only the 5' cap remains, which is hydrolysed by the scavenger enzyme DcpS (Liu *et al.*, 2002).

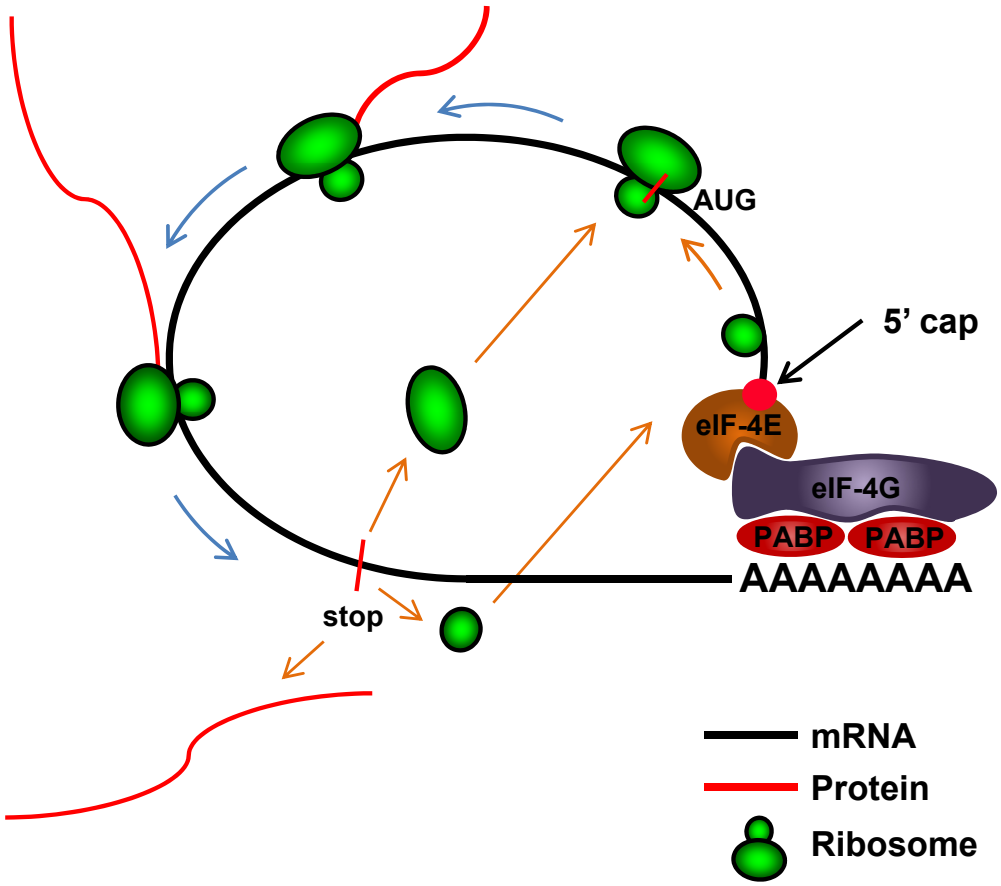



**Figure 1.2**: Simplified cartoon showing the conformation of actively translated mRNAs. eIF-4G brings together eiF-4E and PABP, which are respectively bound to the 5' cap and poly(A) tail of the mRNA. This brings the 5' and 3' ends of the mRNA into close proximity and allows for efficient recycling of ribosomal subunits. Many other factors are involved in initiation, translation and release, but these are not shown.

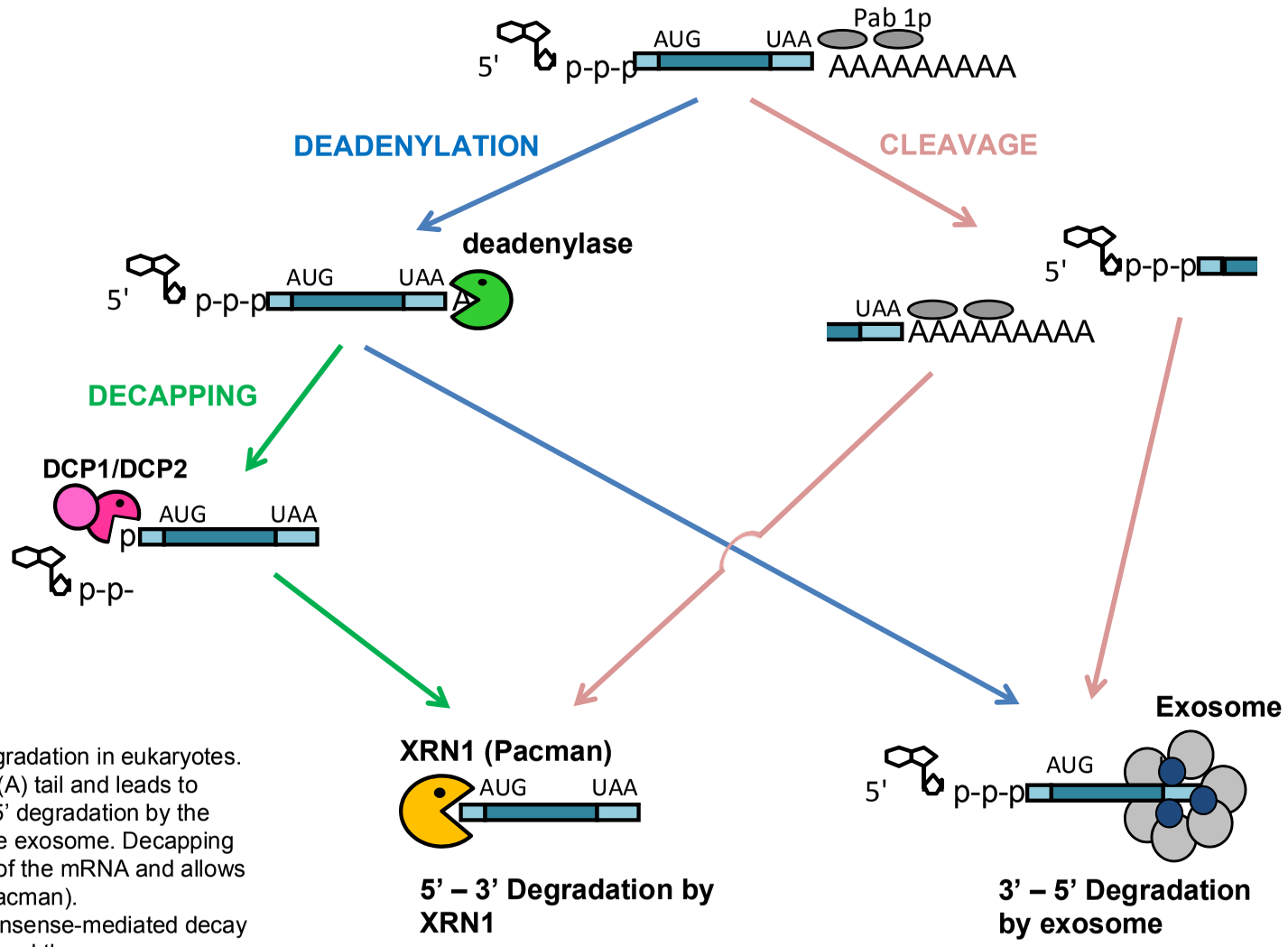



**Figure 1.3**: Overview of mRNA degradation in eukaryotes. Deadenylation removes the 3' poly(A) tail and leads to decapping of the transcript or 3' – 5' degradation by the multi-subunit complex known as the exosome. Decapping by Dcp1/Dcp2 removes the 5' cap of the mRNA and allows for 5' – 3' degradation by XRN1 (Pacman).
Internal cleavage, as caused by nonsense-mediated decay or RNA interference, allows XRN1 and the exosome access to degrade the RNA from both directions.

The deadenylated mRNA can also be degraded in the 5'-3' direction if the 5' cap is removed by a decapping enzyme such as Dcp2, which is found in all eukaryotes, or Nudt16, originally identified in *Xenopus* and recently found in mammalian cells (Song *et al.*, 2010). This exposes the 5' end to the 5'-3' exoribonuclease XRN1 (Pacman in *D. melanogaster*) (Bashkirov *et al.*, 1997, Till *et al.*, 1998). In some cases, decapping can occur independently of deadenylation, such as occurs to begin degradation of the *S. cerevisiae* mRNAs *RPS28B* (Badis *et al.*, 2004) and *EDC1* (Muhlrad and Parker, 2005).

Aberrant mRNA transcripts are detected and targeted for rapid degradation by the nonsense-mediated decay (NMD) pathway. The response to premature termination codons (PTCs) occurs by varying mechanisms in mammals, *D. melanogaster* and *S. cerevisiae* , but the result is the same as rapid degradation of the transcript occurs in a UPF protein dependent manner (Conti and Izaurralde, 2005). In mammalian cells, exon junction complexes (EJCs) are deposited at exon junctions during splicing, and are usually displaced by the ribosome during the first round of translation. If a PTC is present and one or more EJCs are not removed, UPF proteins are recruited by the stalled ribosome and interact with the EJC, triggering decay and release of the ribosome (Silva and Romão, 2009). In *D. melanogaster* and *S. cerevisiae*, NMD is induced if the 3' UTR is too long (Hansen *et al.*, 2009). This has also been shown to be the cause of NMD for mammalian transcripts that contain a PTC in the last exon (and therefore have had all EJCs removed) (Amrani *et al.*, 2004). Aberrant transcripts can also be detected if they lack a stop codon (Frischmeyer *et al.*, 2002, van Hoof *et al.*, 2002) or cause ribosomes to stall during elongation (Doma and Parker, 2006, Passos *et al.*, 2009). NMD is also important to remove many endogenous transcripts, rather than just aberrant transcripts. One example is the sex determination gene *transformer* in *D. melanogaster*, which is alternately spliced in males and females. In males, the transcript produced is very long, contains a PTC and consequently undergoes NMD, whereas the female transcript is short, does not contain a PTC, and is translated (Metzstein and Krasnow, 2006). The requirement for NMD is highlighted by null *Upf1* and *Upf2* mutations in *D. melanogaster*, which are larval lethal, and deletion of *Upf1* (*Rent1*) in *Mus musculus*, which causes embryonic lethality (Medghalchi *et al.*, 2001). Many of the genes involved in NMD were first identified in C. elegans, where mutations cause abnormal genitalia (Hodgkin *et al.*, 1989).

The fate of transcripts targeted to the NMD pathway varies between organisms. In yeast, transcripts are decapped or deadenylated and degraded by XRN1 or the exosome (Cao and Parker, 2003, Mitchell and Tollervey, 2003). In *D. melanogaster*, internal cleavage of the mRNA occurs and each fragment is rapidly degraded by Pacman or the exosome (Gatfield and Izaurralde, 2004). In mammalian cells, there is evidence that both accelerated decapping (Couttet and Grange, 2004) and deadenylation (Chen and Shyu, 2003), and internal cleavage occur (Huntzinger *et al.*, 2008, Eberle *et al.*, 2009, Kashima *et al.*, 2010). The different methods of degradation correlate with the presence or absence of proteins that interact with UPF1. SMG7 increases the rate of decapping by DCP2 and is not present in *D. melanogaster*, which may explain why NMD occurs by internal cleavage in flies. SMG6 is the endonucleolytic protein present in flies that associates with UPF1 and causes nonsense transcripts to be cleaved, but is missing in *S. cerevisiae*, while a homologue of SMG7 is present (Luke *et al.*, 2007), which causes decapping/deadenylation. SMG6 and SMG7 are both present in mammalian cells, where both cleavage and decapping/deadenylation have been shown to occur, and there is much debate about the contribution of each pathway (Bhuvanagiri *et al.*, 2010, Mühlemann and Lykke-Andersen, 2010, Nicholson and Mühlemann, 2010).

### 1.3.1 microRNA induced degradation

Endonucleolytic cleavage is a very effective way of exposing transcripts to the degradation machinery. In all eukaryotes, the RNA-induced silencing complex (RISC), which is directed to specific targets by microRNAs (miRNAs or miRs) to cause their downregulation, often works by cleaving the mRNA. Since the discovery of the first miRNAs (Lee *et al.*, 1993, Reinhart *et al.*, 2000) the importance of this system for 3’ UTR-mediated gene expression control has been realised. miRNAs are expressed as long hairpin-forming precursor RNAs that are cleaved and processed to form the mature miRNA of around 22 nucleotides (Bartel, 2004) (Figure 1.4). The mature miRNA is then incorporated by the Argonaute 1 (Ago1) protein, which is part of the RISC. The specificity of the complex is determined by nucleotides 2-8 of the miRNA (the “seed region”), which bind to complementary sequences usually in the 3’ UTRs of mRNAs. This can lead to either degradation or translational repression of the transcript, the choice of which is thought to depend on the complementarity between the rest of the miRNA and the target site (Valencia-Sanchez *et al.*, 2006, Gu and Kay, 2010).

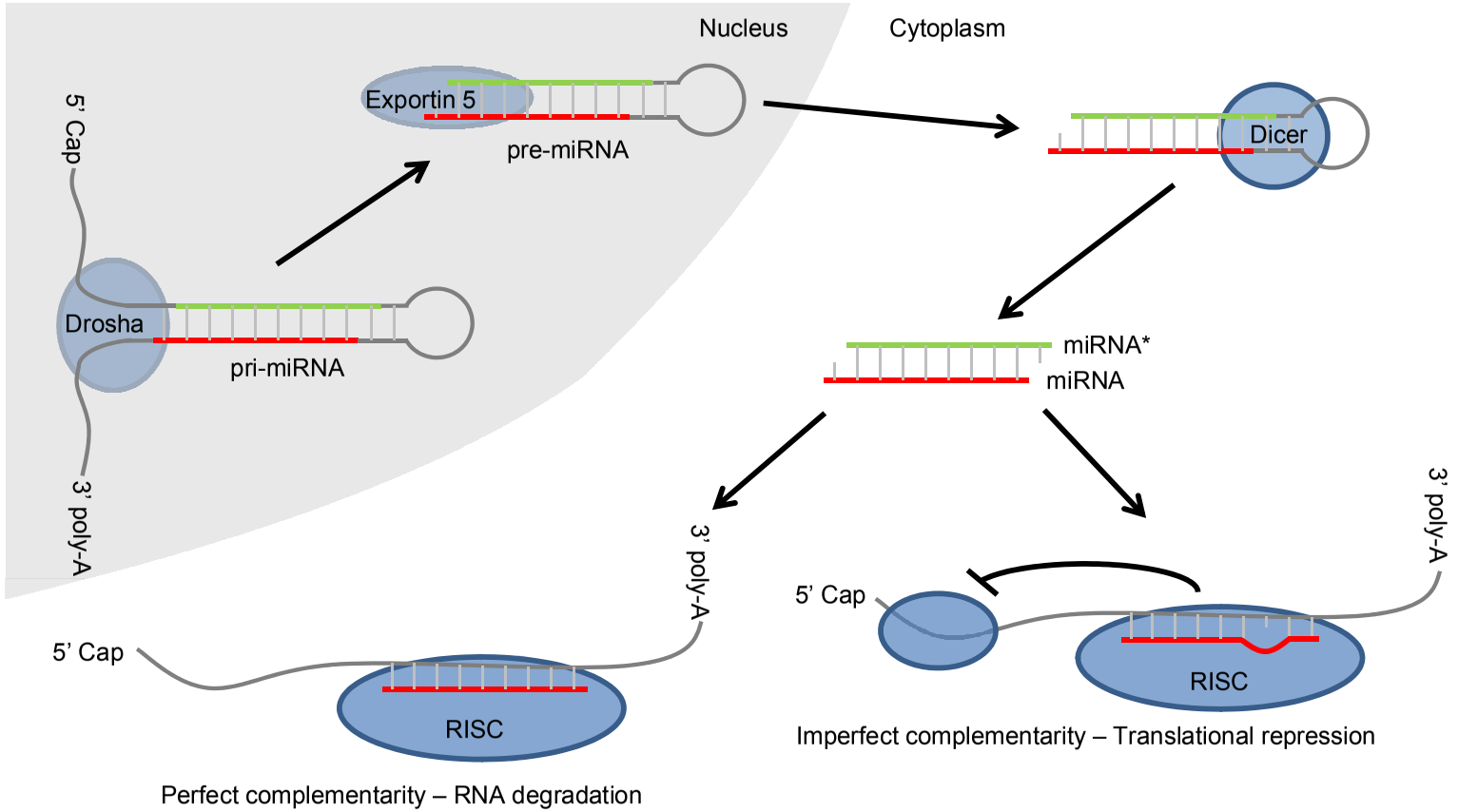



**Figure 1.4**: miRNA biogenesis.
The pri-miRNA is typically transcribed from intergenic or intronic regions of the genome by RNA Polymerase II and is cleaved in the nucleus by Drosha and its partner DGCR8 (Pasha) to form the pre-miRNA. The 2nt overhang at the 3' end is recognised by Exportin 5 and the pre-miRNA is exported to the cytoplasm, where it undergoes hairpin removal by Dicer-1 to produce the miRNA·miRNA* complex. The mature miRNA is separated from the passenger strand (the miRNA*) and associates with Argonaute, part of the RNA-induced silencing complex (RISC). In some instances, the miRNA* can also associate with Argonaute and itself target mRNAs. The 'seed region' (nucleotides 2-8) of the miRNA play the largest part in determining the target mRNAs. Either translational repression or degradation can occur to mRNAs targeted by the RISC (Bartel, 2004, Valencia-Sanchez *et al.*, 2006, Gu and Kay, 2010).

## 1.4 Control of translation and mRNA stability during development

Differential control of stability and translational repression of mRNA transcripts is an effective method for controlling gene expression. Numerous examples of mRNA level control exist, and disruption to this regulation can have significant developmental effects. Specific examples are given below:

### 1.4.1 3' UTR factors

#### 1.4.1.1 Proteins

Factors specific to mRNAs are often involved in stability control and commonly found in the 3' UTR of transcripts. AU-rich elements (AREs) attract a number of proteins that affect turnover. Tristetraprolin (TTP) is an example of a protein that binds to AREs in the 3' UTRs of specific mRNAs, causing them to be destabilised. TTP deficiency in mice has been shown to lead to overexpression of tumour necrosis factor-α (TNF-α) and granulocyte-macrophage colony stimulating factor (GM-CSF), which leads to multiple problems, including growth retardation and autoimmunological effects (Carballo and Blackshear, 2001, Aslam and Zaheer, 2010). mRNA level regulation of TNF-α and numerous other inflammatory factors have also been implicated in the remodelling of the airways of asthmatics (Ammit, 2005), showing a direct link between mRNA stability and a human pathology. Other examples of destabilising ARE binding proteins are AUF1 and KSRP, which recruit the exosome (Chen *et al.*, 2001, Gherzi *et al.*, 2004). HuR (ELAV in *D. melanogaster*) is an example of an ARE binding protein that increases stability of mRNAs (Brennan and Steitz, 2001), possibly by competing with the ARE binding proteins that have a destabilising effect (Linker *et al.*, 2005). Pyrimidine rich 3' UTR regions can also confer stability to mRNAs and are often found in long-lived mRNAs. KH-domain proteins interact with pyrimidine rich regions and PABP and have been shown to increase the stability of a number of transcripts (Makeyev and Liebhaber, 2002).

*D. melanogaster* eggs and embryos have proven to be an ideal system in which to study mRNA regulation, as maternal mRNAs are deposited and localised in an environment where transcription is not occurring. The deposited mRNAs are responsible for gene expression until the zygotic genome becomes active. Smaug is an example of a protein that is responsible for differential abundance of transcripts and translational repression of

transcripts until they are needed. *Hsp83* is an mRNA targeted by Smaug and in embryos <1 hour old, both Smaug and *Hsp83* are present across the embryo. However, by 2-3 hours, maternal *Hsp83* mRNA is localised solely to the posterior pole, due to recruitment of the CCR4-NOT complex to *Hsp83* by Smaug, which leads to deadenylation and degradation of *Hsp83* (Semotok *et al.*, 2005). Smaug exerts the same effect on *nanos* mRNA, expression of which is required at the posterior pole to form a gradient across the embryo. Smaug causes translational repression and degradation of *nanos* everywhere except at the posterior pole, where it is prevented from interacting with *nanos* by Oskar (Zaessinger *et al.*, 2006, Jeske *et al.*, 2011).

#### 1.4.1.2 microRNAs

A good example of the function of a miRNA is that of *Drosophila miR-9a*, which is conserved throughout eukaryotes (*miR-9*). During *D. melanogaster* development, *miR-9a* is expressed in wing imaginal discs and two of its targets have been identified as the mRNAs *senseless* and *Drosophila LIM-only* (*dLMO*) (Li *et al.*, 2006, Biryukova *et al.*, 2009, Bejarano *et al.*, 2010). *miR-9a* mutant flies exhibit an increased number of sensory bristles on the anterior wing margin and notum, and loss of tissue from the posterior wing margin. The increase in sensory bristles is due to increased expression of *senseless*, which has three *miR-9a* binding sites in the 3' UTR. Genetic interaction between *miR-9a* and *senseless* mutants is apparent as the production of ectopic sensory bristles is suppressed. Overexpression of *miR-9a* has the opposite effect to the mutant, reducing the number of sensory bristles (Li *et al.*, 2006). The relationship between *miR-9a* and *dLMO* is similar, as loss of *miR-9a* leads to an increase in *dLMO* and apoptosis of dorsal wing primordium. The same phenotype is seen for dominant *dLMO* mutations where the 3' UTR of *dLMO*, which contains a *miR-9a* binding site, has been removed. When attached to a luciferase reporter, point mutations in the sequence of the *dLMO* 3' UTR that is complementary to the *miR-9a* seed region are sufficient to increase the level of expression (Biryukova *et al.*, 2009, Bejarano *et al.*, 2010).

The most recent release of miRBase (Griffiths-Jones *et al.*, 2006) (release 17, April 2011) contains 19,724 miRNA sequences for 153 organisms. This includes 238 for *D. melanogaster* and 1,424 for *H. sapiens*. There are tens of examples of specific miRNA:mRNA interactions in many organisms, and the number verified is increasing quickly. A number of methods exist to predict miRNA:mRNA interactions *in silico*, such as TargetScanFly (Kheradpour *et al.*, 2007), PicTar (Grün *et al.*, 2005) and Microcosm (Griffiths-Jones *et al.*, 2008). However, thus far it has proved difficult to accurately predict miRNA:mRNA interactions, but as

technology improves and more examples are found, prediction algorithms should improve. For now, predicted interactions require experimental verification.

### 1.4.2 5' UTR factors

Elements that determine stability are not limited to the 3' UTR, as there are a growing number of examples of 5' UTR elements capable of mediating mRNA turnover. Recent examples include the "D5" region in the 5' UTR of Hypoxia-inducible Factor-1α (Wang and Lin, 2009) that promotes stability, and the alternate 5' UTRs of Acyl-coenzyme A:cholesterol acyltransferase, of which the longer promotes decay of the mRNA (Zhao *et al.*, 2009).

The 5' UTR can also greatly impact translational efficiency of different mRNAs due to varying thermal stability of secondary structures that the ribosome must pass through to reach the translation start site. Many proto-oncogene mRNAs have long, structured, GC rich 5' UTRs that require sufficient eIF-4E availability and eIF-4A, a helicase capable of opening up mRNA secondary structure. An increase in eIF-4E expression or loss of its regulation are common features in cancer cells (De Benedetti and Graff, 2004).

### 1.4.3 Iron response elements

RNA binding proteins recognise their targets by the secondary structures formed as a consequence of the mRNA primary sequence. An elegant and well studied example of a RNA binding protein that is known to affect various transcripts via both 3' and 5' UTRs is iron regulatory protein (IRP), which binds to iron response elements (IREs). IREs found in the 5' UTR of transcripts (such as ferritin mRNAs) bind IRPs when iron concentration is low, which physically blocks translation initiation. Conversely, binding of IRPs to IREs in the 3' UTR of the transferrin receptor mRNA leads to increased stability and increased iron uptake. In circumstances when iron concentration is high, transferrin receptor mRNAs (and others) are efficiently deadenylated and degraded, while ferritin mRNAs are translated (Thomson *et al.*, 1999, Leipuviene and Theil, 2007).

## 1.5 XRN1

XRN1 has been most extensively characterised in yeast, where it was identified as being required for numerous functions in various screens (Kim *et al.*, 1990, Heyer *et al.*, 1995, Kim, 2002). It was identified as a 5' - 3' exoribonuclease, and it seems likely that its apparent multiple functions are due to differential degradation of mRNA transcripts by XRN1, rather than being roles fulfilled by a single protein. Differential mRNA degradation by XRN1 has now been shown to be important in a number of contexts in various organisms. In *S. cerevisiae*, mutations in Xrn1p cause a significant reduction of growth rate (Larimer and Stevens, 1990) and a build up of decapped and deadenylated mRNAs (Hsu and Stevens, 1993). In *C. elegans*, downregulation of Xrn-1 by RNAi results in embryonic lethality at the twofold stage of development, due to failure of ventral enclosure (Newbury and Woollard, 2004). In cell lines derived from *H. sapiens*, RNAi knockdown of XRN1 has been shown to increase the half-life of ARE containing mRNAs (Stoecklin *et al.*, 2006) and prevent degradation of histone mRNAs (Mullen and Marzluff, 2008). In *D. melanogaster*, hypomorphic *pacman* mutations lead to a number of developmental phenotypes, including reduced fertility, dull wings and bristle defects (Grima, 2002, Grima *et al.*, 2008, Sullivan, 2008, Zabolotskaya *et al.*, 2008).

### 1.5.1.1 XRN1 and P-bodies

P-bodies are distinct cytoplasmic aggregates of mRNA and factors involved in decapping, NMD, miRNA-mediated decay and deadenylation, and include XRN1, the only known 5'-3' cytoplasmic exoribonuclease (Balagopal and Parker, 2009, Kulkarni *et al.*, 2010). P-bodies are dynamic and often only visible when there is a build up of mRNAs for degradation, such as in decapping factor or XRN1 mutants (Sheth and Parker, 2003, Zabolotskaya *et al.*, 2008). This suggests P-bodies are the sites of 5'-3' decay, but their function does not solely seem to be turnover, as some mRNAs directed to P-bodies are subsequently released for further translation (Brengues *et al.*, 2005, Bhattacharyya *et al.*, 2006) and mRNA degradation is still possible without their formation (Eulalio *et al.*, 2007, Stalder and Mühlemann, 2009).

### 1.5.1.2 Structure of XRN1

The crystal structure of XRN1 has recently been determined in *D. melanogaster* (Jinek *et al.*, 2011) and the yeast *Kluyveromyces lactis* (Chang *et al.*, 2011), which follow an earlier study reporting the crystal structure of the highly related protein XRN2 (Xiang *et al.*, 2009). XRN1 and XRN2 share a common and highly conserved N-terminal, where the $Mg^{2+}$ active site is located, however, XRN2 lacks further domain structures present towards the C-terminal of

XRN1. Jinek *et al*. (2011) crystallised roughly two thirds of XRN1 (residues 1-1141 of 1612) that had been made catalytically inactive, but was still capable of substrate binding. The N-terminal "core" of XRN1 was found to be made up of the XRN catalytic domain and a PAZ/Tudor domain and the C-terminal "arch" included individual KOW, Winged helix and SH3-like domains (Figure 1.5, panel A). The C-terminal was referred to as an arch, as it curls around to interact with the N-terminal, and certain amino acid substitutions in the interacting regions were sufficient to prevent catalytic activity *in vitro*. Figure 1.5, panel B shows the percent sequence similarity between the domains in Pacman and *H. sapiens* XRN1 and *C. elegans* Xrn-1.

## 1.6 Advantages of studying XRN1 in *Drosophila melanogaster*

*D. melanogaster* is an ideal model organism in which to study XRN1 and mRNA decay as much of the mRNA degradation machinery is strongly conserved between eukaryotes (Newbury, 2006, Garneau *et al.*, 2007). Fly husbandry is much cheaper and simpler compared to mammal models such as *M. musculus* or *Rattus norvegicus* and offers considerably quicker turnover of generations whilst retaining the ability to study genes in a developmental context, unlike yeast. The genetics of *D. melanogaster* is probably the most developed of any organism. The research community is very well established and the tools and techniques developed over the past century allow for almost any gene to be mutated or inserted, and knocked down or over expressed constitutively or in very specific cells. Numerous genetic markers and balancer chromosomes ensure tractability in *Drosophila* genetics. For well conserved genes/proteins, results obtained are directly relevant to humans, potentially even more so than results obtained from human cell lines, as cell lines cannot fully represent functioning organisms.

### 1.6.1 Overview of *D. melanogaster* development

*D. melanogaster* development has been intensively studied and can be described in fine detail. A brief overview is given below, and considerably more information can be found in the references used for this section (Ashburner, 1989, Lawrence, 1992, Gilbert, 2000, Wolpert *et al.*, 2002, Greenspan, 2004). The time from egg laying to eclosion of adult flies varies with temperature. The shortest time is 7 days at 28°C, above which the time taken increases due to heat stress. At 25°C development time is 10 days and at 19°C it is 14 days.

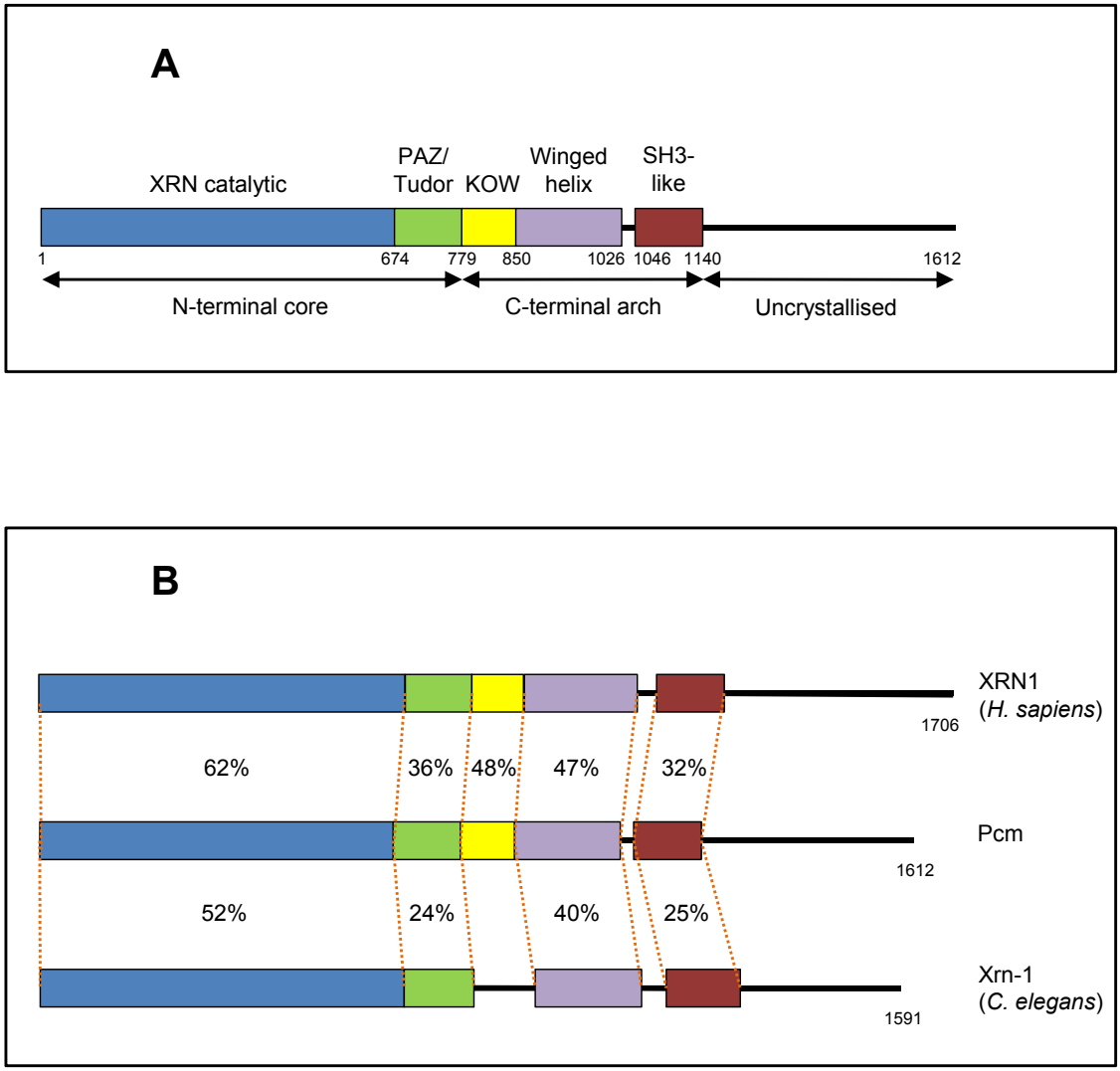



**Figure 1.5**: Pacman protein domains and their conservation in *H. sapiens* XRN1 and *C. elegans* Xrn-1.
**A**. The structural domains identified in the crystal structure of Pacman by Jinek *et al.* (2011).
**B**. Percent similarity of domains between *D. melanogaster* Pacman, *H. sapiens* XRN1 and *C. elegans* Xrn-1, determined using sequence alignments of individual domains in VectorNTI Advance 11.

All times given below are at 25°C. After embryogenesis, larvae hatch from the egg and moult twice before undergoing complete metamorphosis during pupation and eclosing as fully formed adults.

During development of the oocyte, some maternal mRNAs are differentially deposited in the egg to define the anterior-posterior axis. Upon fertilisation, these mRNAs are translated and form protein gradients from the anterior (Hunchback and Bicoid) and posterior (Caudal and Nanos) ends of the oocyte. For the first two hours the egg is a syncytium, in which there is nuclear division in a continuous cytoplasm. During this time, a distinct number of nuclear divisions occur and the majority of the nuclei migrate from the centre of the egg to the surface, while some are left behind to form the egg yolk cells. The nuclei that migrate to the posterior pole of the egg become pole cells, the germline cells of the fly. The rest of the cells that migrate to the edge make up the blastoderm, and the plasma membrane forms to create a monolayer of cells. By gastrulation, the majority of cells are fully enclosed by membrane and no longer directly connected to the yolk. Gastrulation creates three germ layers as the mesoderm and endoderm invaginate at the ventral furrow and the ectoderm remains on the outside. These germ layers exhibit distinct levels of gene expression and their fates become constricted. Cells at the ventral midline elongate and move around the posterior of the egg to form the germ band, which retracts slightly while undergoing segmentation into 14 parasegments. The segments are specified by a hierarchy of genes, beginning with gap genes and followed by pair-rule and segment polarity genes. Most of these cells form only larval structures, but some are retained in small groups and these form the imaginal discs that will go on to construct the adult. After dorsal closure, head involution and the completion of gut development, embryogenesis is complete. Fertilisation to hatching takes about 24 hours and after hatching from the egg, larvae undergo two moults before pupariation. The 1$^{st}$ and 2$^{nd}$ larval instars last 24 hours each and the third lasts 48 hours. 20-hydroxyecdysone (ecdysone) is the hormone responsible for initiating moults and metamorphosis. Pulses of ecdysone that lead to the moults from 1$^{st}$ to 2$^{nd}$ instar and 2$^{nd}$ to 3$^{rd}$ instar larvae are prevented from causing metamorphosis by juvenile hormone (JH). At the end of the 3$^{rd}$ instar, JH secretion is inhibited and existing JH is degraded before the first pulse of ecdysone, which commits cells to pupal development by inactivation of larval genes. A second pulse of ecdysone induces transcription of pupa-specific genes and initiates pupariation, approximately 5 days after the start of embryonic development. Pupariation is the formation of the puparium, as the larval skin darkens and hardens. 6 hours after pupariation, apolysis occurs and the larval epidermis

retracts from the cuticle, which marks completion of the prepupal period, and the animal enters the pupal period, which lasts 24 hours, during which time head eversion occurs (12 hours after pupariation). Pupal to adult ecdysis marks the end of the pupal period, from which point the animal is technically a pharate adult. This is followed by deposition of the adult cuticle and pigmentation of the eyes (Bainbridge and Bownes, 1981). Roughly 10 days after fertilisation, the fully formed adult ecloses from its pupal case. The wings of the fly are fully expanded by 1 hour after eclosion, and body pigmentation is complete by 2 hours later. Females become receptive to mating 8-14 hours after eclosion.

#### 1.6.1.1 Overview of wing imaginal disc development

The largest imaginal discs in the larva are the wing discs, from which the wings and thorax of the adult fly form. During the 1$^{st}$ instar of larval development, the anterior-posterior axis of discs is specified by Engrailed, which is expressed only in the posterior compartment. Engrailed causes expression of Hedgehog in the posterior compartment, but also prevents the posterior cells being responsive to Hedgehog. The immediately adjacent anterior cells respond to Hedgehog and express Decapentaplegic (Dpp), which forms a concentration gradient across both compartments and activates genes dependent on its concentration. During the 2$^{nd}$ larval instar, the dorsal-ventral axis is defined by Apterous, which is only expressed in the dorsal cells and causes the expression of Serrate and Fringe. Serrate is a ligand of the Notch receptor, which is expressed by the ventral cells, so Serrate-Notch binding occurs at the wing margin. This leads to expression of Wingless in cells adjacent to the wing margin. Wingless forms a concentration gradient across the wing and causes expression of different genes dependent on concentration. However, Wingless acts as a morphogen only on the dorsal side, as only cells expressing Apterous are able to respond to the signal. Eversion of the wing imaginal disc occurs during pupariation and involves cell division and rearrangement of cells already present (Taylor and Adler, 2008). Cells in the wing pouch telescope out and form the 3D shape of the wing. The rest of the imaginal disc cells move across the dorsal side (to form the thorax and notum) and fuse to each other, and across the ventral side to fuse with the leg discs (Auerbach, 1936, Ashburner, 1989, Zeitlinger and Bohmann, 1999, Wolpert *et al.*, 2002). Inside the pupal case the wings are highly folded and made up of two layers of cuticle that have been secreted by layers of epidermal cells that are still present. After eclosion, wing expansion occurs within the first few hours, but normally within 10 minutes and takes around 15 minutes. During and after this time, the epidermal cells of the wings degenerate and the dorsal and ventral surfaces gradually come together and appose. By 24 hours after eclosion, the only cells remaining in

the wings are associated with the wing veins (Johnson and Milner, 1987). Figure 1.6 shows the areas of the L3 wing imaginal discs that form each area of the adult wing.

### 1.6.2 Pacman expression during *D. melanogaster* development

Till *et al.* (1998) showed by Northern blotting that *pacman* mRNA is differentially expressed during development (Figure 1.7, panels A and B). The highest levels of *pcm* mRNA was detected in 0-8 hour old embryos and adult flies (relative to *rp49* expression). There was also an increase in the level of *pcm* in early pupae over the level in larvae, which was too low to be detected on the Northern blot. Comparative quantitative RT-PCR experiments presented in this thesis show that *pcm* mRNA is readily detectable in both whole L3 larvae and L3 wing imaginal discs, although the purpose of these experiments was not to quantity *pcm* expression across tissues/developmental stages. Staining for Pacman protein in L3 wing imaginal discs, performed recently by Joe Waldron in the Newbury lab, shows that Pacman is expressed across the whole wing disc (Figure 1.7, panel C).

## 1.7 The link between Pacman and wing development

A number of the phenotypes of *D. melanogaster pacman* mutants show that *pacman* is required for proper development of the structures that form from the wing imaginal discs – the wings and thorax. Dull wings, characterised as a loss of iridescence, is the most common phenotype seen in *pcm* mutants and is often combined with a “ruffling” of the posterior wing margin (Grima, 2002). A less common phenotype reported for some *pcm* mutants is failure of thorax closure. This is characterised by a “cleft” in the thorax where the wing imaginal discs have failed to seal together. Due to similarities of phenotypes, a link has been suggested between Pacman and the JNK pathway, which is active during thorax closure (Grima *et al.*, 2008, Sullivan, 2008).

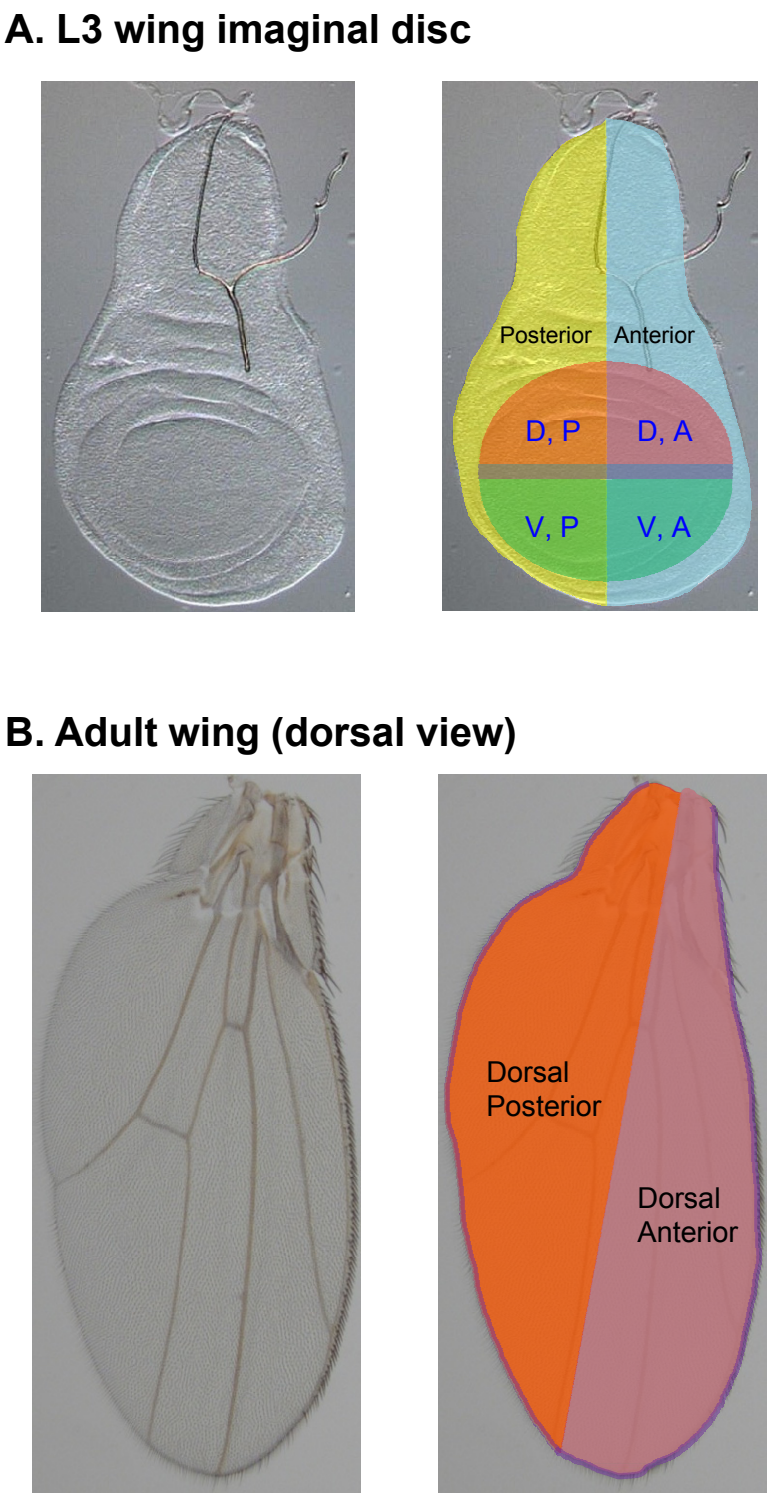



**Figure 1.6**: L3 wing imaginal disc to adult wing.
**A**. During pupation, the cells of the wing pouch (circular region) of the L3 wing imaginal discs telescope out to form the adult wing. Blue text indicates the section of wing that will be formed by the  each section of the wing discs (e.g. D, P = dorsal, posterior). The purple band will form the wing margin. Yellow and blue regions show the anterior and posterior parts of the wing disc.
**B**. Dorsal view of the adult wing. The coloured areas correspond to the area of the wing disc from which they are derived. The ventral regions (coloured green) form the underside of the wing and are not visible from the dorsal side.

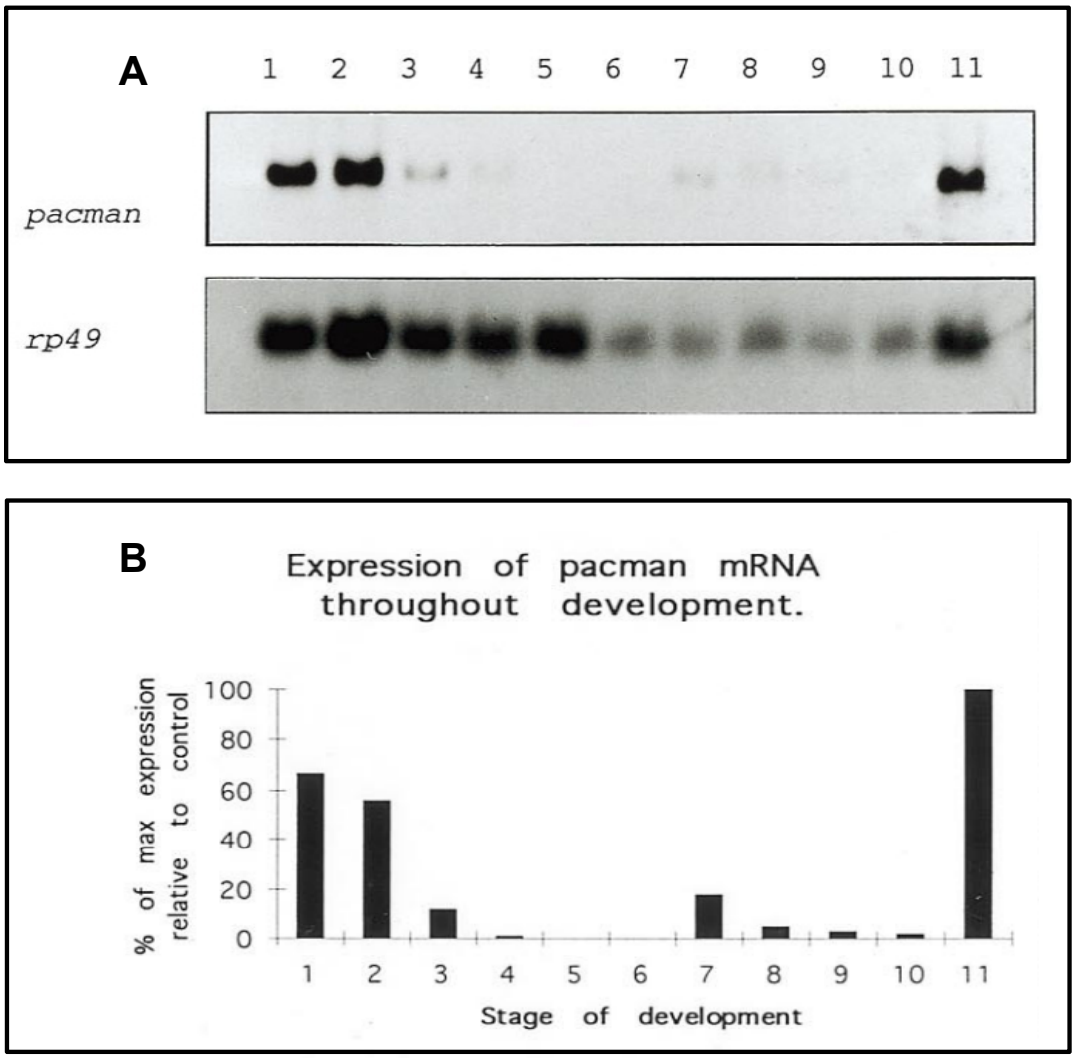



**Figure 1.7**: Pacman mRNA and protein expression.
**A** and **B.** Till *et al.* (1998) quantified *pcm* expression relative to *rp49* at various development stages in *D. melanogaster*: (1) 0-4h embryos, (2) 4-8h embryos, (3) 8-24h embryos, (4) 1st instar larvae, (5) 2nd instar larvae, (6) 3rd instar larvae, (7) early pupae, (8) mid-pupae, (9) mid-late pupae, (10) late pupae and (11) adults.
**C**. Pacman protein expression in wing imaginal discs, stained with rabbit anti-Pacman primary antibody and Cy3 goat anti-rabbit secondary. Secondary only controls fail to stain wing discs (not shown).

## 1.8 Understanding of the function of Pacman before commencement of this project

The *pacman* gene is located in the cytological region 18C7 of the *X* chromosome at 19,379,913bp – 19,385,984bp, on the minus strand. The length of the gene is 6,072bp, which includes a 138bp 5' UTR, a 223bp 3' UTR and 11 exons. There are two annotated transcripts for *pcm* (*pcm-RA* 5,200bp and *pcm-RB* 5,203bp) which differ by 3bp at the end of exon 2. The translations of these transcripts differ only by a single glutamine encoded by the three extra base pairs in *pcm-RB* (Pcm-PA, 1612aa and Pcm-PB, 1613bp). It may be the case that there is a single Pacman protein encoded by a single *pacman* transcript, and the difference above is due to an error made during sequencing. Data retrieved from FlyBase version FB2011_04 (April 2011) (Tweedie *et al.*, 2009).

Pacman was first described by Till *et al.* (1998) and was shown to be differentially expressed during development. Since then, several hypomorphic mutations of *pcm* have been produced by P-element excision. The phenotypes of the mutants suggest a link between Pacman and the JNK pathway at two points during development (dorsal closure and thorax closure) and also in adults as *pcm* mutant flies do not recover from wounds as well as wild-type flies (Grima *et al.*, 2008, Sullivan, 2008). Pacman has also been shown to localise to P-bodies in adult testes and affect male fertility, as mutants have shorter and smaller testes and produce fewer sperm and offspring than wild-type males (Zabolotskaya *et al.*, 2008). *pcm* mutant females are less fertile than wild-type and their ovaries exhibit more, and larger, P-bodies (Lin *et al.*, 2008).

## 1.9 Aims of this project

The overall aim of this project is to determine which RNAs have expression levels that are directly Pacman-dependent, and to attempt to elucidate the direct and secondary effects that cause the phenotypes observed in *pacman* mutants. The approaches employed are described in the four following results chapters and outlined briefly below.

**Chapter 3**

The aims of the experiments reported in Chapter 3 were to create a null *pacman* allele and a suitable control line, and to verify these lines molecularly. Techniques used include P-element excision to induce mutations in *pacman* and PCR and sequencing to determine the extent of created *pcm* deletions.

**Chapter 4**

The aims of Chapter 4 were to verify the lethality and determine the phenotypes of the newly created *pacman* alleles. Techniques used include recombination mapping, complementation with deficiencies and hypomorphic *pcm* alleles, and rescue using a translocation of wild-type *pcm* to the *Y* chromosome. Also described is creation of a new deficiency that includes the *pcm* locus using the DrosDel RS system.

**Chapter 5**

The hypothesis tested in Chapter 5 is that Pacman is responsible for maintaining the level of one or more mRNAs during wing/thorax development. The tissue used for the majority of experiments was L3 wing imaginal discs. Techniques used include Affymetrix 3' arrays and TaqMan comparative quantitative reverse-transcriptase PCR (qPCR) to compare mRNA levels between wild-type and mutants. Custom TaqMan assays were used to determine if observed differences were transcriptional or post-transcriptional, and Western blotting was performed to check that mRNA changes translated into protein level changes.

**Chapter 6**

The hypothesis tested in Chapter 6 is that Pacman is responsible for degradation and maintaining the level of one or more miRNAs. Tissue used was mainly L3 wing imaginal discs. Techniques used include use of Exiqon Other Species miRNA arrays and miRNA qPCR.

## 1.10 Introduction to techniques used in this thesis to quantify gene expression

Chapters 5 and 6 of this thesis present results obtained using various techniques to quantify the relative levels of gene expression in *pcm* mutants. Four methods are used: semi-quantitative PCR (sqPCR), quantitative PCR (qPCR), microarrays and Western blotting. The theory of each technique is outlined below, with particular focus on qPCR as it is the most extensively used, and further technical details can be found in the Materials and methods (Chapter 2). Applications of each technique to test specific hypotheses are described in the relevant results chapters. Advantages and disadvantages of each technique are summarised in Figure 1.8.

### 1.10.1 Microarrays

Microarrays are an established technology used to globally assay genomes and expression of genomes. The most common type is a DNA microarray, where thousands of different oligonucleotides are anchored to a solid supporting structure, such as a glass slide. The oligonucleotides are arranged into “spots” (or “features”) and the defined sequence of the oligonucleotides in a spot determines their specificity for their target molecule. A sample of fluorescently labelled target molecules (e.g. cDNA) is applied to the slide, which allows for complimentary base pairing between probes and target molecules. The chip is then scanned to quantify the amount of fluorescence at each individual spot, which is relative to the abundance of target molecule in the original sample. Several samples can be compared to identify differentially represented targets, which allows for global analysis of changes in gene expression levels under different conditions. Some known limitations of microarrays are that they can underestimate large fold changes, struggle to pick up small expression changes due to variation and may not be able to distinguish between multiple transcripts from the same gene (Kothapalli *et al.*, 2002, Yao *et al.*, 2004, Kim *et al.*, 2010). Two microarray platforms were used for this thesis: Affymetrix GeneChip *Drosophila* Genome 2.0 arrays (Affymetrix, 2004) and Exiqon miRCURY LNA microRNA v.11.0 other species arrays (Exqion, 2008). Figure 1.9 shows an example of a generic microarray experiment work flow.

| | sqPCR | TaqMan qPCR | Microarrays | Western blotting |
|---|---|---|---|---|
| **Platform** | Open | Proprietary | Proprietary | Open |
| **Substrate** | mRNA (cDNA) | mRNA/miRNA (cDNA) | mRNA (cDNA)/miRNA | Protein |
| **Resolution, accuracy and reproducibility** | Low | Very High | Moderate – High | Low |
| **Cost** | Low | High | High | Moderate |
| **Time cost** | Low – High | Low – High | Low (service) – Moderate (DIY) | Moderate |
| **Coverage (per experiment)** | Very low | Low | Complete | Low |
| **Overall transcriptome/ proteome coverage** | Complete | Very High | Complete | Moderate |
| **Equipment requirements** | Low | High | Low (service) – Extreme (DIY) | Moderate |
| **Flexibility/customisability** | High | Moderate | N/A | Moderate |
| **Ease of data analysis (i.e. determination of FC)** | Moderate | Simple | Hard | Moderate |

**Figure 1.8**: Comparison of techniques used to calculate fold changes in gene expression.
Green indicates positive attributes of each technique and red indicates negative attributes. Yellow indicates attributes that are neither positive nor negative. Attributes are rated qualitatively and relatively between techniques.

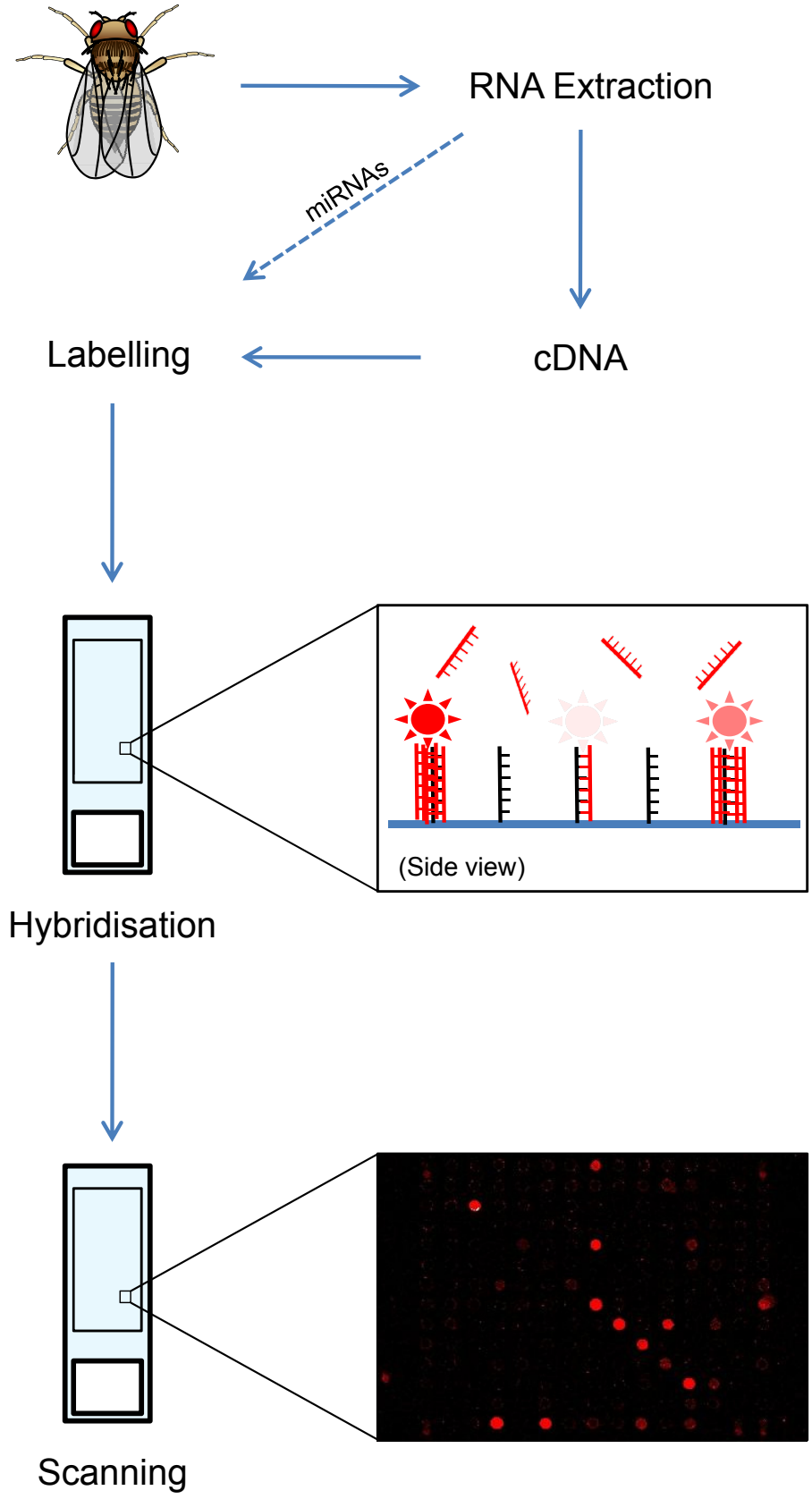



**Figure 1.9**: Overview of a microarray experiment.
RNA is extracted from a biological sample using a method appropriate for the RNA being tested (e.g. mRNA or miRNA). mRNA is converted in cDNA before labelling and miRNAs are labelled directly. During the hybridisation step, cDNAs/miRNAs bind to spots with complementary sequences. The more molecules bound to a spot, the brighter the fluorescence from that spot when the array is scanned. This example shows a single sample/dye microarray, such as the Affymetrix mRNA arrays. The Exiqon miRNA arrays were two colour, meaning two samples were labelled with different dyes and hybridised to the same array. In this case, the ratio of the dyes on each spot shows the relative level of expression between the samples.

#### 1.10.1.1 Affymetrix GeneChip *Drosophila* Genome 2.0 arrays

The Affymetrix arrays were used to globally assay gene expression in wild-type flies and *pcm* mutant flies to find genes differentially expressed in the mutants. Gene expression is measured at the mRNA level with probes targeted to the 3' end of mRNA transcripts. In total, there are 18,880 11µm spots for over 18,500 transcripts on these arrays. The samples were collected and RNA extraction was started in the Newbury lab before being sent to the Sir Henry Wellcome Functional Genomics Facility (SHWFGF, University of Glasgow) for completion of RNA extraction, cDNA creation (using oligo(dT) primers) and to perform the arrays. Bioinformatic analysis for the arrays was performed by the SHWFGF and at Sussex with the assistance of Dr. Sue Jones.

#### 1.10.1.2 Exiqon miRCURY LNA microRNA v.11.0 other species arrays

The Exiqon arrays were used to assay miRNA expression in wild-type flies and *pcm* mutant flies to find miRNAs differentially expressed in the mutants. These arrays contain 10,032 90µm spots including 152 *D. melanogaster* miRNAs. LNA (Locked Nucleic Acid) technology is used to increase the strength of binding between miRNAs and their respective spots on the array. LNA nucleotides contain a modification to the ribose moiety that greatly increases the $T_m$ of complementary base pairing, which is useful for detection of miRNAs, as their short length makes it difficult to increase the $T_m$ between miRNA and probe by simply extending the probe sequence. An RNA extraction method that preserves small miRNAs must be used to prepare RNA from the samples, which are directly labelled with different dyes for wild-type and mutant samples. One wild-type and one mutant sample were hybridised to the same array and the intensities of each dye on each spot were measured to identify differentially expressed miRNAs.

### 1.10.2 Semi-quantitative reverse transcriptase polymerase chain reaction ("sqPCR")

As the semi-quantitative RT-PCR experiments presented in this thesis were all performed on cDNA (hence the "RT" in "RT-PCR"), this method is referred to as sqPCR for simplicity. Normal PCR is used to determine if a DNA sequence (or RNA sequence in the form of cDNA) is present in a sample. At the end point of a PCR, it is only possible to tell if the target is present or absent, but it is not possible to determine the amount of target present in the original sample. It is possible to get an estimate of the relative level of a target present in the original sample by visualising the reaction as it is efficiently occurring and new product is being produced at the maximum rate (doubling every cycle in the exponential phase). To do this, a PCR is set up at sufficient volume to allow samples to be taken at specific cycles and run on a gel to visualise when the band for the desired product first appears. Two or more samples can be compared and any difference can be expressed as a fold change between the samples. As a theoretical example, if the product of primers specific for Gene A cDNA first becomes visible at cycle 22 in wild-type and at cycle 27 in a mutant sample, then there is an obvious decrease in the level of expression of Gene A in the mutant sample. The fold difference can be estimated using the formula 2^cycle difference. In this case, $2^5 = 32$, so Gene A is expressed at -32-fold in the mutant.

The theory of sqPCR assumes 100% reaction efficiency. As this is rarely the case in practice (the efficiency of amplification by any given primer set is typically not known) there are possible sources of error. This makes sqPCR unsuitable for measuring small fold changes, but has less effect on measuring large changes as long as appropriate controls are used. sqPCR is best used as a cheap and quick way to roughly verify expression differences between samples and requires no special reagents or equipment above what is required for normal PCR.

### 1.10.3 TaqMan comparative quantitative real time reverse transcriptase polymerase chain reaction ("qPCR")

Comparative quantitative RT-PCR follows the same principle as sqPCR, but can give much more accurate measurements of expression levels. In this thesis, TaqMan comparative quantitative RT-PCR is referred to as qPCR as it is the only type used. qPCR could also include other assay methods, such as SYBR green, can be absolute rather than comparative and does not have to be performed on cDNA, but these types of qPCR are not relevant to this thesis. The experiments performed for this thesis were broadly congruent with the suggested Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin *et al.*, 2009).

#### 1.10.3.1 TaqMan qPCR on mRNAs

TaqMan qPCR is a proprietary technology developed by Applied Biosystems (Applied Biosystems (Life Technologies), 2010a). It works the same way as sqPCR in that quantification of PCR product at the exponential phase of the reaction is used to compare expression levels between samples. However, TaqMan assays include a fluorescent probe as well as PCR primers. The probe binds within the area amplified downstream of one of the primers. As the primers are extended, the probe is hydrolysed, which separates a fluorophore from a fluorescent quencher, allowing a fluorescent signal to be detected by the qPCR machine (Figure 1.10). The level of fluorescence from each well/reaction is measured at the end of every cycle and the cycle at which the fluorescence is recorded above background level is called the $C_t$ value. $C_t$ values can be compared between samples to calculate the fold difference in expression.

There are many aspects to the design of TaqMan qPCR that ensure it is accurate. Initially, the same amount of RNA is used to make cDNA, and the reverse transcriptase step (using random primers) is duplicated two or three times per RNA sample. The same amount of cDNA is then used in the qPCRs, each of which is duplicated two or three times.

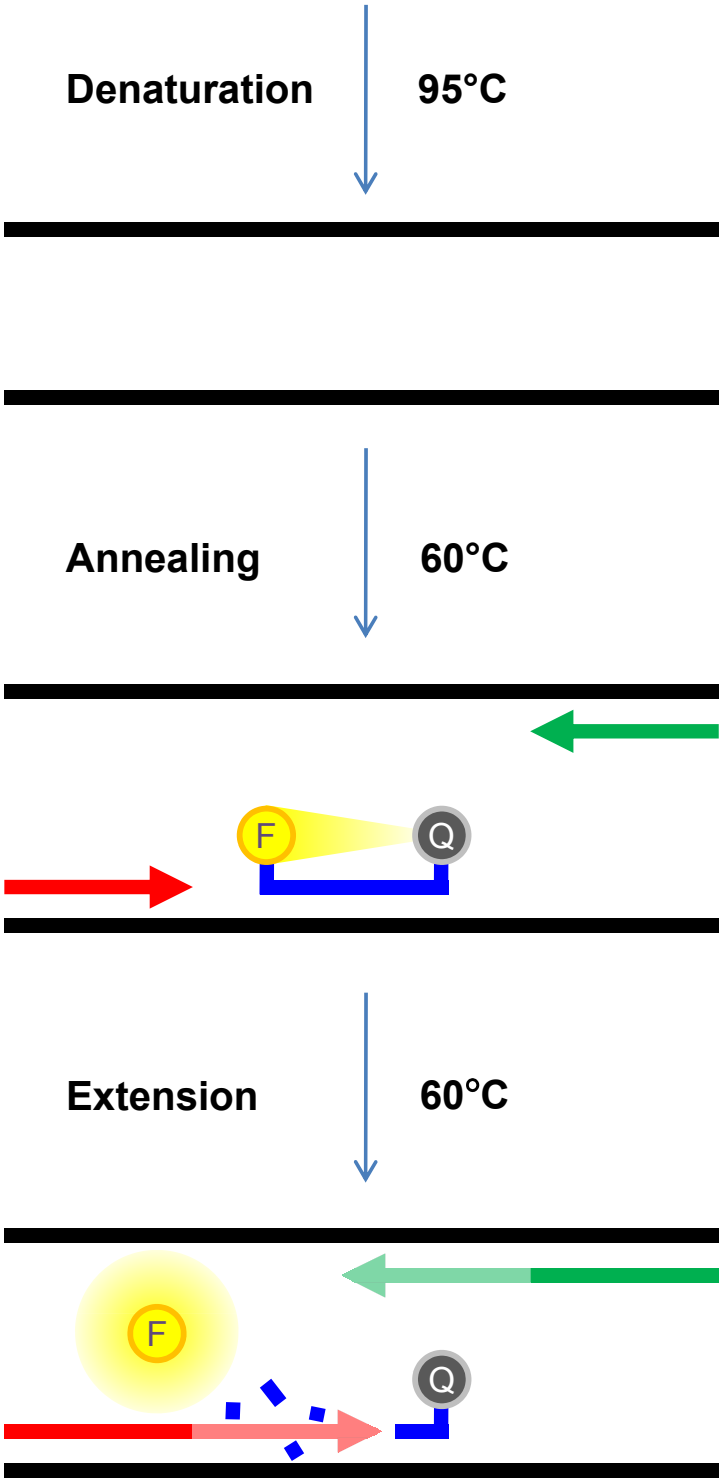



**Figure 1.10**: TaqMan quantitative PCR chemistry.
At the start of the PCR cycle, the template strands are denatured at 95°C. As the temperature is reduced, the forward (red) and reverse (green) primers are able to anneal by complementary base pairing to each template strand. The fluorescent probe (blue) base pairs to a sequence downstream of one of the primers. The probe consists of a specific nucleotide sequence, a fluorescent reporter (F) and a quencher (Q). When the probe is intact, the fluorescence from the reporter is quenched. During the extension phase, the probe is cleaved as the Taq Polymerase extends along the template. As more probe molecules are cleaved in subsequent cycles, the level of fluorescence emitted increases.

Applied Biosystems recommend three reverse transcriptase replicates and three qPCR replicates for each RNA sample/probe combination used, totalling nine technical replicates that should each give the same result. To control for variations in the initial input of RNA or cDNA, reference assays are run alongside test assays. These reference genes are known not to vary between samples (i.e. are the same level in mutant and wild-type). The results obtained for test genes are first normalised to the reference genes, then compared between samples. A well established reference gene used in *D. melanogaster* is *rp49*, a ribosomal protein gene expressed consistently across tissues and at the same level in wild-type flies and *pacman* mutants.

Pre-designed TaqMan assays against almost every *D. melanogaster* gene are available from Applied Biosystems. Typically, a number of assays are available for each gene, and are designed so that primers amplify across the junction between two exons. This, combined with the specific probe, ensures that only the mature mRNA is detected. The specific probe also prevents any non-specific amplification products being detected. Non-specific products can be a severe issue when performing SYBR green qPCR, as the SYBR green dye will fluoresce when bound to any double stranded DNA. Depending on the assays available, it is often possible to select assays to enable discrimination between different transcript variants from the same gene. Custom assays can also be designed to requirements, as were used in this thesis to test pre-mRNAs. In this case, it is essential to perform a DNase step after RNA extraction to ensure no genomic DNA is present when the PCR is performed, as assays designed against pre-mRNAs will also amplify the genomic copy of the gene.

#### 1.10.3.2 TaqMan qPCR on miRNAs

While the theory of performing TaqMan qPCR on miRNAs or mRNAs is essentially the same, there are three technical considerations that must be made for miRNAs (Applied Biosystems (Life Technologies), 2010b). Firstly, as with samples prepared for miRNA microarrays, RNA extraction should be performed that preserves small nucleotide molecules. Often kits for RNA extraction exclude molecules under a certain size, for example the Qiagen RNeasy Mini kit (used for preparation of the RNA used on the mRNA arrays) has a minimum cut-off size of 200 nucleotides. RNA (or DNA) molecules smaller than this will lost, which is desirable when mRNAs are the molecules of interest, but precludes the use of these RNA samples for miRNA analysis. The *mir*Vana miRNA Isolation kit from Ambion allows total RNA extractions with recovery of small RNA molecules and optional enrichment for small RNAs (removal of mRNAs, large rRNAs etc.). All RNA extractions for

miRNA experiments were performed using the *mir*Vana kit for total RNA extraction, which allowed for the mRNAs to also be tested if required.

The second consideration for miRNA qPCR is that cDNA prepared from miRNAs using random primers (as it is for mRNA cDNA using the Applied Biosystems High Capacity Reverse Transcription kit used to produce cDNA from mRNAs) is too small to be used in PCR. miRNAs are about the length of a primer used for normal PCR (around 21 nucleotides), making it impossible to amplify them directly. This is compounded by the fact that TaqMan primers are designed to have a higher $T_m$ than primers for normal PCR typically are, so are often 22-30 nucleotides in length. At least 19bp is also required for the specific probe to bind, which is why the minimum sequence length required for design of custom TaqMan assays is 70bp. To overcome this, the reverse transcription reaction is modified to increase the size of the cDNA derived from miRNA molecules. This is achieved by using proprietary looped RT primers specific for individual miRNAs. This allows a defined sequence to be attached to the miRNA sequence, which can be used during qPCR. During qPCR the sequence from the looped RT-primer opens up due to the higher temperature and one PCR-primer binds to it and the other binds to the miRNA sequence. The probe binds in the region of the interface between the RT-primer sequence and miRNA sequence (Figure 1.11) It is possible that the PCR-primer that binds to the RT-primer sequence is common between all assays, but the sequences for the RT-primers are not given. The practical effect of this system is that all miRNA cDNA reactions have to be set up individually with their specific RT-primer. The amount of total RNA required for each of these cDNA reactions is fortunately considerably smaller than is required for preparation of cDNA from mRNA (around 10ng instead of >1µg).

The final consideration comes from the fact miRNA analysis is in its infancy compared to mRNA analysis, so endogenous controls for use as normalisers are less well defined. *rp49* is commonly used as a normaliser for mRNAs due to its consistent level of expression between *Drosophila* cells, however, when this work began no endogenous controls for miRNAs could be recommended by Applied Biosystems. Initially, we used *rp49* as a normaliser for miRNA experiments and encountered no problems. The possible issue with using *rp49* is that the reverse transcription step differs from the miRNA reverse transcription step. Sometime after this work began, Applied Biosystems began to suggest a number of small RNA molecules for use as miRNA normalisers. We tried two of these, U27 (FlyBase annotation: CR32892, name: snoRNA:U27:54Ea) and snoR442 (FlyBase annotation:

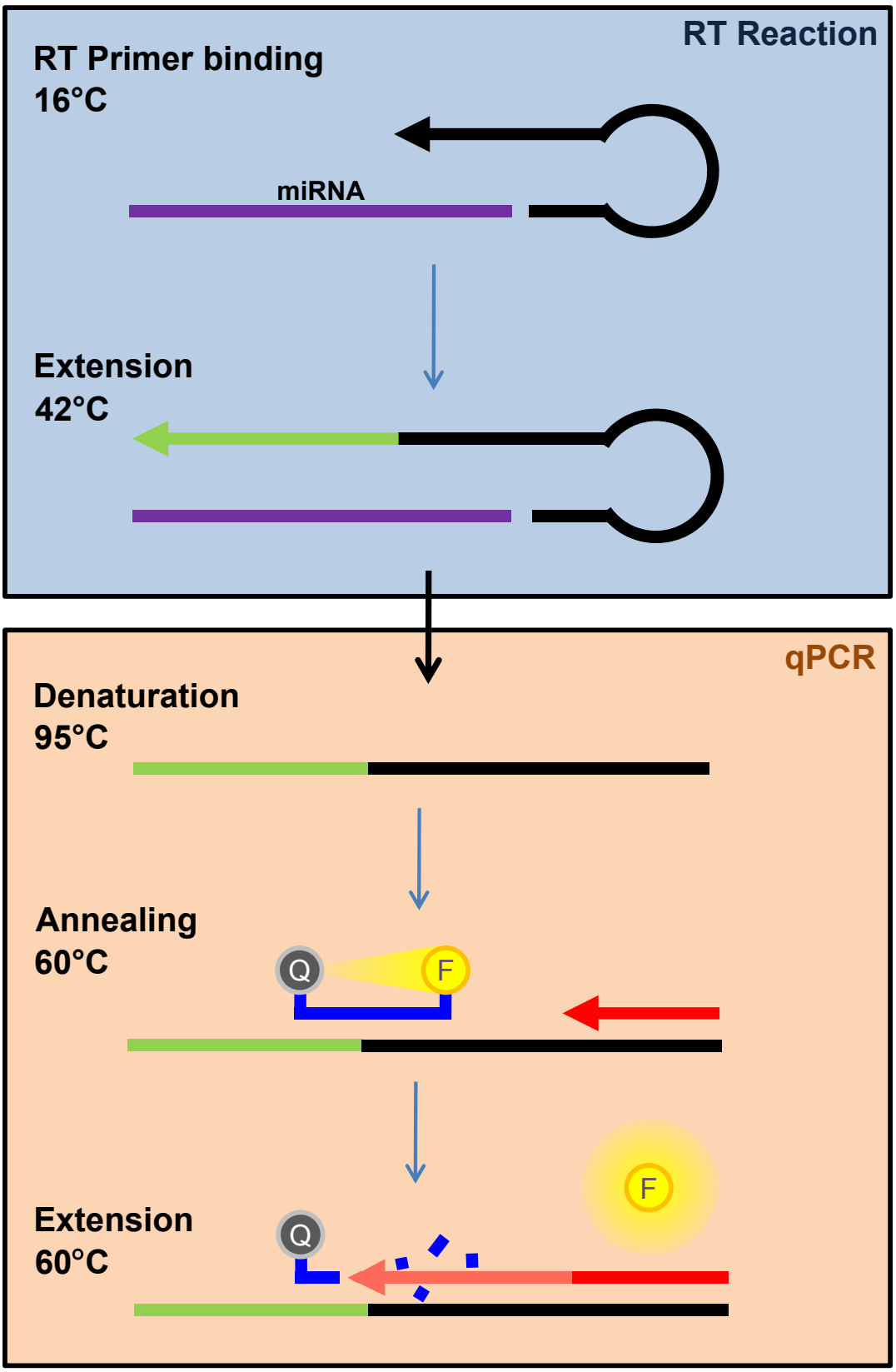



**Figure 1.11**: TaqMan Reverse Transcriptase and quantitative PCR chemistry for miRNAs. During the RT step the looped RT primer binds to the miRNA and is extended. At the start of the PCR cycle, the template strands are denatured at 95°C. As the temperature is reduced, the forward (red) and reverse (not shown) primers are able to anneal by complementary base pairing to each template strand. The fluorescent probe (blue) base pairs at the interface between the primer and miRNA sequences. During the extension phase, the probe is cleaved as the Taq Polymerase extends along the template, separating the fluorophore (F) and the quencher (Q). As more probe molecules are cleaved in subsequent cycles, the level of fluorescence emitted increases.

CR33778, name: snoRNA:Me28S-A2564), both of which are snoRNAs, which are involved in processing of pre-rRNA to mature rRNA. These undergo the same RT reaction as miRNAs, which decreases the chance of technical variance and reduces the amount of total RNA required. U27 and snoR442 were run alongside *rp49* over a number of experiments and both were suitable for use as normalisers. We chose to use U27 as it showed slightly less technical and biological variation than snoR442. When choosing small RNA controls, the method by which they are transcribed should be considered. miRNAs are transcribed by RNA Pol II (as are mRNAs such as *rp49*), while snoRNAs may be transcribed by either RNA Pol II or III (Dieci *et al.*, 2009). This means they may be unsuitable for use as normalisers in situations that differentially affect transcription enzyme efficiency. However, testing U27 and snoR442 against *rp49* showed that this was not an issue for *pacman* mutants.

### 1.10.4 Protein level quantification using Western blotting.

Microarrays, sqPCR and qPCR all measure mRNA expression from genes. If a gene is more highly transcribed, it is likely expression of protein from the mRNA also increases. However, this assumes that mRNA level is a/the factor limiting level of translation, which is not necessarily the case, as regulation at the level of translation is not uncommon (Spriggs *et al.*, 2010). An increase in mRNA level must be verified by testing the protein level. This can be done directly using a Western blot if a suitable antibody is available against the protein of interest. Calculation of the fold difference between wild-type and mutant can be attempted using Western blot films, but achieving a consistent, accurate figure is almost impossible due to practical limitations. This is mostly due to the fact double protein expression ≠ double fluorescence level. Despite this, it is still important to show that the protein level is visibly different between samples when differences to the mRNA level suggest that this might be the case.

### 1.10.5 Statistics to determine significance of gene expression changes

In this thesis, p-values were most commonly calculated for expression differences determined by qPCR. Either one sample or two sample *t*-tests were used depending on the data available. In instances where sufficient data were available to calculate the standard deviations of the wild-type expression level and the mutant expression level, two sample *t*-tests were used. The data required for a two sample *t*-test is mean, standard deviation and n (number) for two populations. In situations where standard deviation was not available for one population, one sample *t*-tests were used with a hypothetical mean of 1. One sample *t*-tests were used for experiments using pooled wild-type samples. In these experiments, the data available is mutant mean, mutant standard deviation and mutant n, and wild-type mean and wild-type n. The wild-type mean is always set to 1, and then the level of each mutant biological replicate is calculated. *t*-tests are a valid method for determining significance for the gene expression level data presented in this thesis, as each expression level is independent and only one variable is being tested (*pcm* mutation or wild-type). Analysis of Variance (ANOVA) (with appropriate post-test to determine p-value) tests are also valid for this data, but are unnecessary as the p-values obtained are the same as those obtained by *t*-tests. $p<0.05$ is taken as statistically significant for all experiments in this thesis. In figures, $p<0.05$ is indicated by * above bars on graphs. $p<0.001$ is indicated by **. It should be remembered that statistical significance ≠ biological significance.

# 2 Chapter 2: Materials and methods

## 2.1 Fly stocks

| Stock | Genotype | Source | Experiment/Use |
|---|---|---|---|
| BL11456 | *$w^{1118}$, P{$w^{+mC}$=EP}EP1526* | Bloomington (BL11456) | P-element excision (P-element source). |
| P-element transposase source | *w ; ; Ly ry/ $TM3^{\Delta 2\text{-}3}$* | | P-element excision (transposase source). |
| OregonR | wild-type | | Wild-type control for sqPCR. |
| 50E | *$w^{1118}$ 50E* | P-element excision (Chapter 3) | Wild-type control for qPCR, microarrays and Western blotting. |
| *$pcm^{3}$* | *$w^{1118}$ $pcm^{3}$ / FM7i, P{$w^{+mC}$=ActGFP}JMR3* | P-element excision (Dom Grima) | sqPCR. |
| *$pcm^{5}$* | *$w^{1118}$ $pcm^{5}$ / FM7i, P{$w^{+mC}$=ActGFP}JMR3* | P-element excision (Dom Grima) | P-element excision (over *FM7c*), sqPCR, qPCR, microarrays, Western blotting. |
| *$pcm^{6}$* | *$w^{1118}$ $pcm^{6}$ / FM7i, P{$w^{+mC}$=ActGFP}JMR3* | P-element excision (Dom Grima) | sqPCR. |
| *$pcm^{13}$* (121C) | *$w^{1118}$ $pcm^{13}$ $Nat1^{121c}$ / FM7i, P{$w^{+mC}$=ActGFP}JMR3* | P-element excision (Chapter 3) | qPCR. |
| *$pcm^{14}$* (16A) | *$w^{1118}$ $pcm^{14}$ / FM7i, P{$w^{+mC}$=ActGFP}JMR3* | P-element excision (Chapter 3) | Recombination, rescue, qPCR. |
| BL1515 | *$y^{1}$ $cv^{1}$ $v^{1}$ $f^{1}$ $car^{1}$* | Bloomington (BL1515) | Recombination. |

| Stock | Genotype | Source | Experiment/Use |
|---|---|---|---|
| *Df(1)JA27* | *Df(1)JA27 / FM7i, P{w$^{+mC}$=ActGFP}JMR3* | Bloomington (BL971) | Rescue, complementation (BL971 balancer is *FM7c*). |
| *Df(1) Exel7468* | *Df(1)Exel7468, w$^{1118}$ / Binsinscy* | Bloomington (BL7768) | Rescue. |
| *Df(1)ED7452* | *Df(1)ED7452, w$^{1118}$ / FM7i, P{w$^{+mC}$=ActGFP}JMR3* | DrosDel RS (Chapter 4) | Recombination, rescue, complementation. |
| *T(1;Y)B92* | *T(1;Y)B92, y$^{1}$ y$^{+}$ B$^{S}$ / In(1)FM7* | Kyoto (101110) | Rescue. |
| *l(1)G0156* mutant | *w$^{67c23}$, P{lacW}l(1)G0156$^{G0156}$ / FM7c* | Bloomington (BL11896) | Rescue (presumed positive control). |
| *vav* mutant | *w$^{67c23}$, P{lacW}l(1)G0147$^{G0147}$ / FM7c* | Bloomington (BL11837) | Rescue (presumed positive control). |
| 5B | *5B, w$^{1118}$ / FM7i, P{w$^{+mC}$=ActGFP}JMR3* | P-element excision (Chapter 3) | Rescue (presumed negative control). |
| 50B | *50B, w$^{1118}$ pcm$^{5}$ / FM7i, P{w$^{+mC}$=ActGFP}JMR3* | P-element excision (Chapter 3) | Recombination, rescue. |
| 5-HA-1900 | *P{RS5}Nat1$^{5\text{-}HA\text{-}1900}$* | Bloomington (BL125461) | DrosDel RS. |
| CB-0310-3 | *P{RS3}Pfrx$^{CB\text{-}0310\text{-}3}$* | Bloomington (BL123119) | DrosDel RS. |
| BL6876 | *w$^{1118}$ ; noc$^{Sco}$/ SM6B, P{70FLP ry$^{+}$}* | Bloomington (BL6876) | DrosDel RS (FLP source). |
| BL6878 | *P{RS3}l(1)CB-6411-3$^{1}$, w$^{1118}$ / FM7h* | Bloomington (BL6878) | DrosDel RS (*FM7h* source). |

## 2.2 Fly husbandry

### 2.2.1 Fly food recipe

Recipe for 7 litres of fly food:

| | | | |
|---|---|---|---|
| Agar (GP) | 81g | Black treacle | 410g |
| Baker's yeast | 81g | Propionic acid | 40ml |
| Oatmeal | 616g | Nipagin | 1 spatula |
| Distilled $H_2O$ | 7000ml | | |

The agar, baker's yeast, oatmeal and water were mixed together in a large saucepan and brought to the boil whilst continuously being stirred to avoid burning. This was then removed from heat and allowed to cool for 5 minutes. The treacle was then added and the mixture was allowed to cool for a further 10 minutes, followed by addition of the nipagin and propionic acid. The food was thoroughly stirred and poured into bottles or vials and covered in muslin to cool overnight. Yeast pellets were sprinkled over the food and foam bungs were used to cap the bottles/vials. The bottles/vials were stored at 4°C until needed.

### 2.2.2 Stock maintenance

Files raised on the bottles containing the above food were kept duplicated in two stock incubators at 19°C. These bottles were turned over every 4 weeks (the incubators were offset from each other by 2 weeks).

Stocks in active use for experiments were kept in bottles or vials containing the above food at 25°C. Bottles were turned over every 2 weeks and discarded within 4 weeks.

### 2.2.3 Cross timings

Stock bottles of flies required for a cross were set up on day 0 (Friday). Adults were tipped off the bottles on day 7 (Friday). Virgin females and males required were collected from day 10-14 (Monday – Friday) and the cross between the collected virgins/males was set up on day 14 (Friday) in bottles or vials. The adults were tipped off the cross bottles/vials on day 21 (Friday). Offspring from the cross began to eclose from day 24 (Monday).

### 2.2.4 Virgin collection

*D. melanogaster* females do not mate within the first 8 hours after eclosion (Ashburner, 1989, Greenspan, 2004). The presence of a meconium in the abdomen, visible from the ventral side, ensures flies are only a few hours old and therefore virgins, as it persists for only a few hours after eclosion. For all crosses, flies were selected as virgins only if the meconium was visible. To maximise recovery of virgins, collections were performed twice a day in the morning and afternoon for 5 days (Monday – Friday). Bottles or vials for virgin collection were kept at 19°C overnight and at 25°C during the day.

### 2.2.5 Larvae collection

L3 wandering larvae were used in several experiments. Stock bottles were set up on day 0 (typically Monday) and L3 wandering larvae were collected for 3-5 days from day 7 (typically Monday). They were gently removed from the side of bottles using forceps and placed in 1.5ml tubes prior to use. To discriminate hemi/homozygous *pcm* mutants from heterozygous siblings, a Leica MZ16F microscope with a longpass autofluorescence GFP filter was used. Flies containing the balancer *FM7i*, *P{w$^{+mC}$=ActGFP}JMR3* (referred to as *FM7i* in the main text) express Actin-GFP which is clearly visible in L3 larvae under UV light, particularly in the gut. Fluorescent larvae were either *FM7i/Y* or *pcm/FM7i*. Non-fluorescent larvae were either *pcm/Y* or *pcm/pcm*. For control lines such as 50E that were hemi/homozygous viable, picking under UV was not required, as no balancer chromosomes were present.

#### 2.2.5.1 L3 wing imaginal disc dissection for RNA extraction

Suitable larvae were picked and placed in a 1.5ml tube. Larvae were dissected in *Drosophila* Ringer's solution using two pairs of forceps; one to grasp the middle of the larva and one to peel open the cuticle from midway between the middle and the head. Wing imaginal discs were transferred to a 1.5ml tube containing around 1ml *Drosophila* Ringer's solution (on ice) until 30 had been collected. The Ringer's and any floating debris were removed, leaving roughly 20µl Ringer's and the imaginal discs in the tube. The discs were frozen in liquid nitrogen and stored at -80°C until required.

***Drosophila* Ringer's solution**: 3mM $CaCl_2$, 182mM KCl, 46mM NaCl, 10mM Tris, adjusted to pH 7.2 with HCl. The solution was autoclaved and passed through a 0.45µm filter prior to use.

#### 2.2.5.2 L3 wing imaginal disc dissection for size comparison

Suitable larvae were picked and placed in a 1.5ml tube. Larvae were dissected in *Drosophila* Ringer's solution using two pairs of forceps. First each larva was torn in half and the head end was inverted. Debris was removed to leave the cuticle with the imaginal discs on the outside. The inverted heads were transferred to 4% formaldehyde to fix for 1 hour. The wing imaginal discs were dissected from the rest of the head directly onto a microscope slide and the rest of the larva was discarded. 85% glycerol was used to mount the discs and spacer cover slips (size 1) were glued to the glass slide using nail varnish. A size 1 cover slip was placed on top of the spacers to cover the discs.

## 2.3 Genetic techniques

### 2.3.1 Mutant generation by P-element excision

Crosses were set up using the timings and food described above. Cross A was set up in bottles. Cross B was set up in individual numbered vials (e.g. 16). Cross C was set up in individual vials. The Cross B number was retained and a letter added to distinguish each vial (e.g. 16A). 450 potential new mutants were created and male offspring from these were tested by PCR, using the quick DNA extraction method below and the primer pairs indicated in the main text (with sequences given in Appendix A). Stocks lacking detectable *pcm* deletions were discarded.

### 2.3.2 Recombination experiments

Crosses were set up using the timings and food described above. The frequency of phenotypes of the male offspring from $y^1\ cv^1\ v^1\ f^1\ car^1$ / *X* mothers (where *X* represents any *X* chromosome tested) were recorded over 5 days.

### 2.3.3 DrosDel deficiency creation

Creation of the DrosDel deficiency *Df(1)ED7452* was performed using the timings and food described above and following the DrosDel cross/heat shock protocol (Ryder *et al.*, 2004, Ryder *et al.*, 2007).

## 2.4 Molecular techniques

### 2.4.1 General methods

#### 2.4.1.1 Quick DNA extraction

Three flies were collected in a 1.5ml tube and frozen in liquid nitrogen. 150µl of Squish Buffer was added to each tube, and the flies were thoroughly homogenised using a disposable pestle (Sigma) before being placed in a heating block at 37°C for 30 minutes, followed by a further 10 minutes at 95°C. 1µl of this solution was used in the subsequent PCRs. Squish preps were used immediately and not stored or reused.

**Squish Buffer**: 10mM Tris (pH 8.2), 1mM EDTA, 25mM NaCl, 200µg/ml Proteinase K (added immediately prior to use).

#### 2.4.1.2 Standard DNA extraction

10 flies or larvae were collected in a 1.5ml tube frozen in liquid nitrogen. 120µl of Grind Buffer was added to each tube and the flies were homogenised using a disposable pestle (Sigma), which was rinsed as 370µl Grind Buffer was added after homogenisation. The tubes were placed in a 65°C heating block for 10 minutes, after which 75µl 8M KAc was added and the tubes were placed on ice for 15 minutes. Phenol extraction was then performed, with the aqueous phase being transferred to a fresh tube containing 1ml 100% EtOH, which was vortexed and left at 20°C overnight. The following day the DNA was pelleted (10 minutes, max speed in a desktop centrifuge), washed with 70% EtOH, dried thoroughly and resuspended in 50µl $H_2O$. To each gDNA prep, 6µl NEB Buffer 2 and 6µl 333µg/ml RNase A was added prior to 30 minutes incubation at 37°C. The samples were then made up to 400µl with $H_2O$ and a second phenol extraction was performed, with the aqueous phase transferred to 1ml 100% EtOH in a fresh tube and left overnight at -20°C. The following day the DNA was pelleted, washed with 70% EtOH, allowed to dry thoroughly and resuspended in 50µl $H_2O$.

**Grind buffer**: 0.2M Sucrose, 0.1M Tris (pH 9.2), 50mM EDTA, 0.5% SDS, 200µg/ml Proteinase K (added immediately prior to use).

#### 2.4.1.3 PCR primer design

PCR primers were designed using VectorNTI Advance v.11 (Invitrogen) or NCBI Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Annealing temperatures between 56°-60°C were used. Primers designed in VectorNTI were checked for specificity using FlyBase BLAST against the *D. melanogaster* genome (http://flybase.org/blast/). NCBI Primer-BLAST includes automatic BLAST against a specified genome (*D. melanogaster*).

A full list of bioinformatic software used in this thesis can be found in Appendix B.

#### 2.4.1.4 PCR on genomic DNA

Each 21µl reaction contained: 16µl $H_2O$, 2.1µl 10x PCR Buffer (AmpliTaq Gold Buffer), 0.8µl 2.5mM dNTPs (Sigma), 0.9µl 5mM primer mix, 1µl DNA and 0.2µl AmpliTaq Gold (Applied Biosystems). The DNA used was prepared by either the quick DNA extraction method or the standard DNA extraction method detailed above. The reactions were performed in a Perkin Elmer 2400 (24 well) or an Applied Biosystems Veriti (96 well).

Typical cycling conditions:

1. Activation step: 95°C – 10 minutes
2. Denaturing step: 95°C – 30 seconds
3. Annealing step: 57°C – 30 seconds
4. Extension step: 72°C – 30 seconds per kilobase of product
5. Extension step: 72°C – 10 minutes
6. Cooling step: 4°C – unlimited

Steps 2-4 repeated 35 times.

#### 2.4.1.5 Agarose gel electrophoresis

All gels were prepared as 1.5%: 2.25g Agarose in 150ml 1x TBE or 0.45g Agarose in 30ml 1x TBE. GelRed (Cambridge Bioscience) was added before the gels were allowed to set (10,000x dilution) to allow for visualisation of DNA. The gels were run submerged in 1x TBE. 4µl loading buffer per 20µl of sample was added to sample tubes prior to loading. 10-12µl of sample + loading buffer was loaded per well and gels were run for 10-60 minutes at 40-120V depending on size of product/amount of separation required.

**1xTBE (Tris-Borate-EDTA) buffer**: 90mM Tris, 90mM Boric acid, 2mM EDTA.

**Loading buffer** (for 5ml): 1.75g Ficoll 70 (Sigma), 6mg Bromophenol blue, 5mg Xylene cyanol, 3.1ml 0.2M EDTA, 1.9ml $H_2O$.

#### 2.4.1.6 DNA sequencing

For each sample to be sequenced, 3x 21µl PCR reactions were performed to amplify the gDNA (prepared using the standard DNA extraction method). Amplified DNA was purified using a PCR Cleanup kit (Qiagen). Quality and concentration of the purified DNA was checked on a NanoDrop 1000 (Thermo Scientific) to ensure they conformed to the requirements before being sent to Eurofins MWG Operon sequencing service. The appropriate sequencing primers at a concentration of 2mM were also sent.

#### 2.4.1.7 Sequence alignments

Sequence alignments, for example sequenced DNA results, were performed using the AlignX component of VectorNTI Advance v.11.

### 2.4.2 Semi-quantitative PCR

#### 2.4.2.1 RNA extraction

RNA extraction for sqPCR was performed on 10 whole L3 larvae using an RNeasy Mini Kit (Qiagen). The RNA samples were DNase treated with a DNA-free kit (Ambion).

#### 2.4.2.2 cDNA production

The RNA samples were converted to cDNA using a Superscript III RT kit (Invitrogen).

#### 2.4.2.3 sqPCR

Each 100µl sqPCR contained: 86µl $H_2O$, 10µl 10x PCR Buffer (Jumpstart Red Taq Buffer), 0.9µl 25mM dNTPs (Sigma), 2.2µl 5mM primer mix, 2µl cDNA, 5.5µl Jumpstart Red Taq (Sigma).

Cycling conditions (Perkin Elmer 2400):

1. Denaturing step: 95°C – 10 minutes
2. Denaturing step: 95°C – 30 seconds
3. Annealing step: 55°C – 30 seconds
4. Extension step: 72°C – 1 minute
5. Extension step: 72°C – 10 minutes
6. Cooling step: 4°C – unlimited

Steps 2-4 were repeated 35 times. Cycling was paused at the end of the extension step (4) to remove 10µl samples at predetermined cycles. The samples were placed on ice until all had been collected, then run on a 1.5% agarose gel.

### 2.4.3 Quantitative PCR (mRNAs and pre-mRNAs)

#### 2.4.3.1 RNA extraction

RNA extraction for mRNA qPCR was performed on 10 larvae or a specified number of L3 wing imaginal discs using an RNeasy Mini Kit (Qiagen). The RNA samples were DNase treated with a DNA-free kit (Ambion) and had their quality tested on a NanoDrop 1000 spectrophotometer (Thermo Scientific).

#### 2.4.3.2 cDNA production

The RNA samples were converted to cDNA using a High Capacity RNA-to-cDNA Kit (Applied Biosystems). This kit uses random primers.

#### 2.4.3.3 qPCR

Each 20µl qPCR contained: 8.2µl $H_2O$, 10µl 2x TaqMan Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems), 0.8µl cDNA, 1µl TaqMan Gene Expression Assay (Applied Biosystems).

Cycling conditions (Stratagene MX3005P):

1. Activation step: 95°C – 10 minutes
2. Denaturing step: 95°C – 15 seconds
3. Annealing/extension step: 60°C – 30 seconds

Steps 2 and 3 were repeated 50 times. Fluorescence data was collected at the end of step 3 of each cycle.

### 2.4.4 Quantitative PCR (miRNAs)

RNA extraction for miRNA qPCR was performed on 10 larvae or a specified number of L3 wing imaginal discs using a *miR*Vana miRNA Isolation Kit (Ambion/Applied Biosystems). The protocol for total RNA extraction was followed and the samples were not enriched for miRNAs. The RNA samples were DNase treated with a DNA-free kit (Ambion) and had their quality tested on a NanoDrop 1000 spectrophotometer (Thermo Scientific).

#### 2.4.4.1 cDNA production

The RNA samples were converted to cDNA using a TaqMan microRNA Reverse Transcription Kit (Applied Biosystems). Individual cDNA reactions were set up for each miRNA using primers specific to that miRNA (provided as part of the TaqMan miRNA assay).

#### 2.4.4.2 qPCR

Each 20µl qPCR contained: 8.2µl $H_2O$, 10µl 2x TaqMan Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems), 0.8µl cDNA, 1µl TaqMan miRNA Assay (Applied Biosystems).

Cycling conditions (Stratagene MX3005P):

1. Activation step: 95°C – 10 minutes
2. Denaturing step: 95°C – 15 seconds
3. Annealing/extension step: 60°C – 30 seconds

Steps 2 and 3 were repeated 50 times. Fluorescence data was collected at the end of step 3 of each cycle.

### 2.4.5 Affymetrix mRNA microarrays

L3 wing imaginal discs were extracted using the method above. 240 discs were used for each RNA extraction (4x wild-type, 4x *pcm*$^5$). The discs for each sample were homogenised in RLT buffer (from an RNeasy Mini Kit (Qiagen), with β-ME added) and pooled together. After homogenisation, the samples were frozen in liquid nitrogen and sent to the Sir Henry Wellcome Functional Genomics Facility (SHWFGF), University of Glasgow. The RNA extraction was completed and the microarrays were performed at the SHWFGF.

### 2.4.6 Exiqon miRNA microarrays

180 L3 wing imaginal discs for each sample (2x wild-type and 2x *pcm*$^5$) were extracted using the method above. RNA was extracted using a *miR*Vana miRNA Isolation Kit (Ambion/Applied Biosystems). The protocol for total RNA extraction was followed and the samples were not enriched for miRNAs. The RNA samples were DNase treated with a DNA-free kit (Ambion), had their quality tested on a NanoDrop 1000 spectrophotometer (Thermo Scientific) and were concentrated using a Concentrator Plus (Eppendorf).

#### 2.4.6.1 Labelling and Array hybridisation

Two arrays were performed at a time following the protocol provided by Exiqon, including the use of spike in controls (miRCURY LNA Array, Spike-in miRNA Kit). *pcm*$^5$ RNA was labelled with Hy5 dye (green, 532nm) and 50E RNA was labelled with Hy3 dye (red, 635nm). Hybridisation was performed in Microarray Hybridisation chambers (Agilent) overnight at 65°C in a rotary HB-1000 hybridisation oven (UVP). After washing and drying of the slides they were kept in nitrogen before being scanned on an Axon GenePix 4000B Microarray Scanner (Molecular Devices) within 20 minutes.

#### 2.4.6.2 Array analysis

Array analysis was performed using GenePix Pro 7 (Molecular Devices). Normalisation was performed using spike-in miRNAs and miRNAs known to be present at the same level in wild-tye and *pcm*$^5$.

### 2.4.7 Western blotting

#### 2.4.7.1 Sample preparation

60 wing imaginal discs were used per sample/lane, extracted using the method given above. Samples were homogenized in 1.5 tubes using disposable pestles (Sigma) in 35µl 2x SDS loading buffer (per tube/30 discs) then placed at 100˚C for 7 minutes. The samples were then centrifuged for 5 minutes at maximum speed. As samples were homogenised in the same tubes used for collection, each tube contained 30 discs and a small amount of *Drosophila* Ringer's solution left over from collection (around 15µl). 35µl 2x SDS was added to each tube containing 30 discs, giving a total volume of about 50µl in each. After centrifugation, the supernatant was transferred into new 1.5ml tubes and the samples were pooled together (2x 30 wild-type into tube A, 2x 30 *pcm*$^5$ into tube B) to give a total volume of around 100µl prepared from 60 discs.

#### 2.4.7.2 Blotting and transfer

Prior to loading 30µl of the above samples and 25µl pre-stained protein marker onto a NuPAGE Novex 7% Bis-Tris precast gel (Invitrogen), the samples and marker were placed into a heat block at 100°C for 2 minutes. The gel was run in NuPAGE Tris-Acetate SDS Running buffer (Invitrogen) for 1 hour at 150V in an X Cell Surelock Novex Mini Cell tank (Invitrogen). Proteins were transferred to Immobilon P Transfer Membrane (Millipore) for 1 hour at 100V (400mA) in Transfer buffer in a Mini Transblot Cell (BioRad). The membrane was washed in Blocking buffer on a rocking platform (protein side up) for 1 hour. The membrane was rinsed briefly in Wash buffer then probed for protein using the primary antibody (diluted in Wash buffer) on the rocking platform for 1 hour. anti-Simjang was diluted 1:1000, anti-Tubulin was diluted 1:2000. This was followed by 2x 5 minute washes of the membrane in Wash buffer on the rocking platform. The membrane was then probed with the secondary antibody (diluted in Wash buffer) for 1 hour on the rocking platform. anti-rat antibody was used for Simjang and anti-mouse was used for Tubulin, both diluted 1:10,000. This was followed by 2x 5 minute washes in Wash buffer and 2x 5 minute washes in 1x PBS-Tween on the rocking platform.

#### 2.4.7.3 Exposure and developing

In a dark room, the membrane was covered with the ECL Western Blotting Detection solutions (GE Healthcare) for 1 minute and exposed to Super RX Medical X-ray Film (Fujifilm) for varying amounts of time until the desired intensities were visible after developing using an Xograph Compact X2 machine.

#### 2.4.7.4 Western blotting buffers

**2x SDS loading buffer** (pH 6.8): 250mM Tris, 4% SDS, 10% Glycerol, 0.006% Bromophenol blue, 2% β-Mercaptoethanol, 2% Protease cocktail inhibitors. The β-Mercaptoethanol and Protease cocktail inhibitors were added immediately prior to use.

**Transfer Buffer**: 25mM Tris, 190mM Glycine, 20% Methanol.

**Blocking buffer** (100ml): 1 Sachet PBS, 0.1ml Tween 20, 5g Powdered milk (Marvel), 99.9ml $H_2O$.

**Wash buffer** (500ml): 1 Sachet PBS, 0.5ml Tween 20, 2.5g Powdered milk (Marvel), 499.5ml $H_2O$.

**1x PBS-Tween** (1l): 1 Sachet PBS, 1ml Tween 20, 999ml $H_2O$.

# 3 Chapter 3: Creation and characterisation of new *pacman* mutations and control lines

## 3.1 Generation of new *pacman* alleles by P-element excision

Prior to this work, imprecise P-element excision had been performed by Dr. Dominic Grima in the Newbury lab to generate *pacman* mutations (Grima, 2002). Eight *pcm* mutants were generated by imprecise excision of *P{w$^{+mC}$=EP}EP1526* (referred to as *EP1526*) which lies 485bp downstream of the *pcm* 3' terminus in stock BL11456 (full genotype: *w$^{1118}$, P{w$^{+mC}$=EP}EP1526*). These lines were named *pcm$^{1,2,3,4,5,6,9}$* and $^{10}$. Of these, all except *pcm$^{5}$* are deletions ranging from 1,070bp to 1,785bp in size extending from the P-element site (Figure 3.1). In order of size: *pcm$^{2}$* (1,070bp), *pcm$^{9}$* (1,378bp), *pcm$^{3/4}$* (1,523bp), *pcm$^{1}$* (1,558bp) and *pcm$^{6/10}$* (1,785bp). *pcm$^{6}$* and $^{10}$ and *pcm$^{3}$* and $^{4}$ appear to be duplications of the same two events. *pcm$^{5}$* is an internal deletion of 516bp that partially deletes exons 8 and 9 and is unlike any other allele as the deletion does not extend from the P-element insertion site. We think this allele may be the result of an incomplete repair event, however, the undeleted/repaired *pacman* sequence after the deletion is out of frame. In total, six different *pcm* alleles were created, five of which are still available (*pcm$^{1}$* was unfortunately lost). All these alleles are hypomorphic and show varying levels of phenotype penetrance. None are null alleles, but Western blot analysis shows varying levels of Pacman protein expression in each strain, below the level in wild-type controls. *pcm$^{5}$* shows the lowest level of protein expression, the greatest phenotype penetrance and the lowest viability of all the extant alleles (Grima *et al.*, 2008).

Identifying a gene of interest and mutating it to observe the effects is a reverse genetics approach that has worked well to identify the phenotypes of *pcm* mutations. However, hypomorphic mutations cannot show the effects of a total loss of *pcm*. For this reason, we decided to perform a second P-element excision of *EP1526* in an attempt to obtain a null *pcm* allele. There were also two additional objectives: firstly to generate one or more precise excision lines to act as controls for the *pcm* mutant lines (these were not kept after the original P-element excision) and secondly to generate more partial *pcm* deletions.

The excision scheme for the second P-element excision only differed from the first in the source of the transposase gene, which was located on the third chromosome balancer *TM3$^{\Delta 2\text{-}3}$*. Figure 3.2 shows the excision scheme, which consisted of three crosses (A-C).

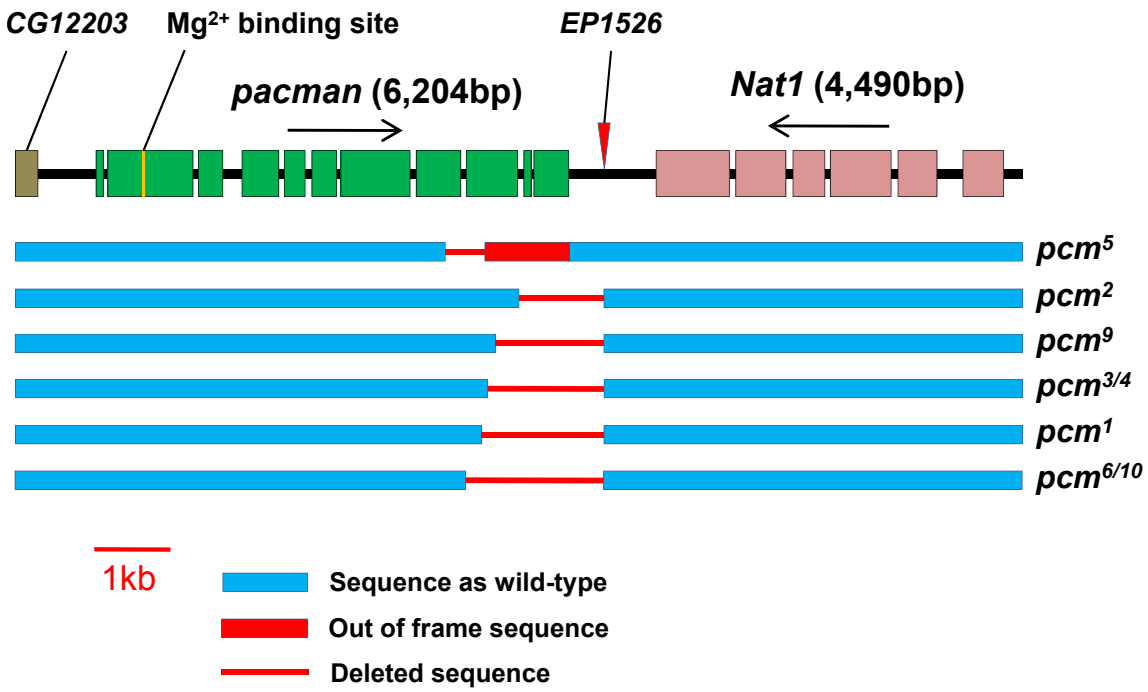



**Figure 3.1**: *pcm* genomic region and alleles previously created by imprecise excision of *P{w*$^{+mC}$*=EP}EP1526* (*EP1256*).
The *pcm* gene consists of 11 exons and expresses a single transcript. The *pcm* promoter region is not known.
Previously created alleles in order of ascending size: *pcm*$^{5}$ (516bp), *pcm*$^{2}$ (1,070bp), *pcm*$^{9}$ (1,378bp), *pcm*$^{3/4}$ (1,523bp), *pcm*$^{1}$ (1,558bp) *pcm*$^{6/10}$ (1,785bp).
The undeleted *pacman* sequence after the deletion in *pcm*$^{5}$ is out of frame. The *pcm*$^{1}$ allele is no longer available.

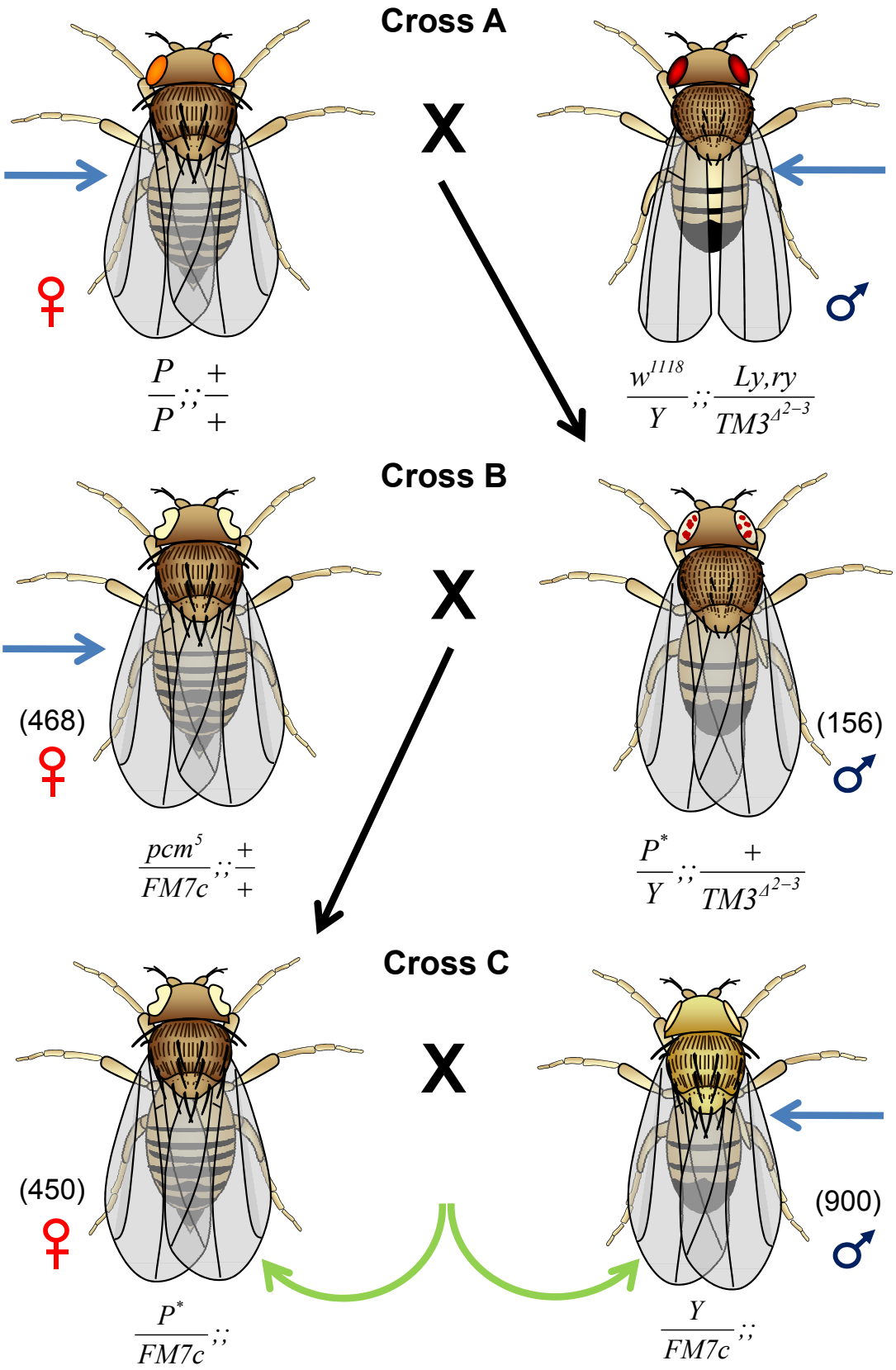



**Figure 3.2**: Crosses used to generate stable lines of flies with excised *EP1526* and potential mutations into *pcm*. Black arrows indicate which offspring from a cross were used in the subsequent cross. Blue arrows indicate adults used that were not offspring from the previous cross. Green arrows indicate a stable stock where genotypes of the offspring are the same as the genotypes of the parents. Numbers in parenthesis indicate the number of adults used in each cross. Crosses B and C were performed in individual vials. *P* refers to the chromosome containing *EP1526*, full annotation $w^{1118}$, $P\{w^{+mC}=EP\}EP1526$. *P** indicates chromosomes from which *EP1526* may have been excised.

### 3.1.1 Cross A

Cross A mobilised *EP1526* by bringing together the *X* chromosome containing *EP1526* (*$w^{1118}$, P{$w^{+mC}$=EP}EP1526*, referred to here as '*P*') and the third balancer chromosome containing the constitutively active transposase gene *$T^{\Delta 2\text{-}3}$* on an immobile P-element insertion (*$TM3^{\Delta 2\text{-}3}$*). This *TM3* chromosome also contains a mutation in the gene *Stubble* which produces the dominant marker phenotype of short hairs. This allows the flies carrying the transposase gene to be easily identified. Any offspring from this cross carrying both *EP1526* and the transposase (including the desired male offspring) were prone to eye colour changes and non-uniform eye colours as shown in Figure 3.3. These eye colours are due to the somatic excision of *EP1526* in independent events. This shows that the *EP1526* has been successfully mobilised, but does not necessarily indicate which flies may contain novel *pcm* alleles as only germ line excision events can be transmitted to offspring, so can only be observed to have occurred in the next generation. Any *P* chromosome present alongside the *$TM3^{\Delta 2\text{-}3}$* chromosome is from this point referred to as *P** to reflect its potential to be different from the original *EP1526* containing chromosome.

### 3.1.2 Cross B

The male offspring from cross A containing both the *P** chromosome and the transposase were crossed to females containing the first chromosome balancer, *FM7c*. These crosses were performed individually in separate vials as each *P** chromosome is potentially unique. The offspring selected to go into the final cross were females with *P** and *FM7c*, but without *$TM3^{\Delta 2\text{-}3}$*. In heterozygous conditions, *FM7c* leads to flies with heart-shaped light yellow (apricot) eyes, so females with heart-shaped eyes and wild-type bristles (indicating no *$TM3^{\Delta 2\text{-}3}$*) were selected.

### 3.1.3 Cross C

The final cross was to set up the stable stocks. Each virgin female with the desired genotype produced from the previous cross was placed in a vial with two males. In the stable stocks, only two non-wild-type chromosomes are present – the balancer *FM7c* and a potentially mutated *P** chromosome. Of the 450 stocks created, around 50% still had yellow eyes, around 40% had white eyes and around 10% had orange or red eyes (Figure 3.3, panels C-F). The white, orange or red eyed stocks were the most likely to show *pcm* mutations as the change in eye colour suggests the P-element has moved or been removed entirely.

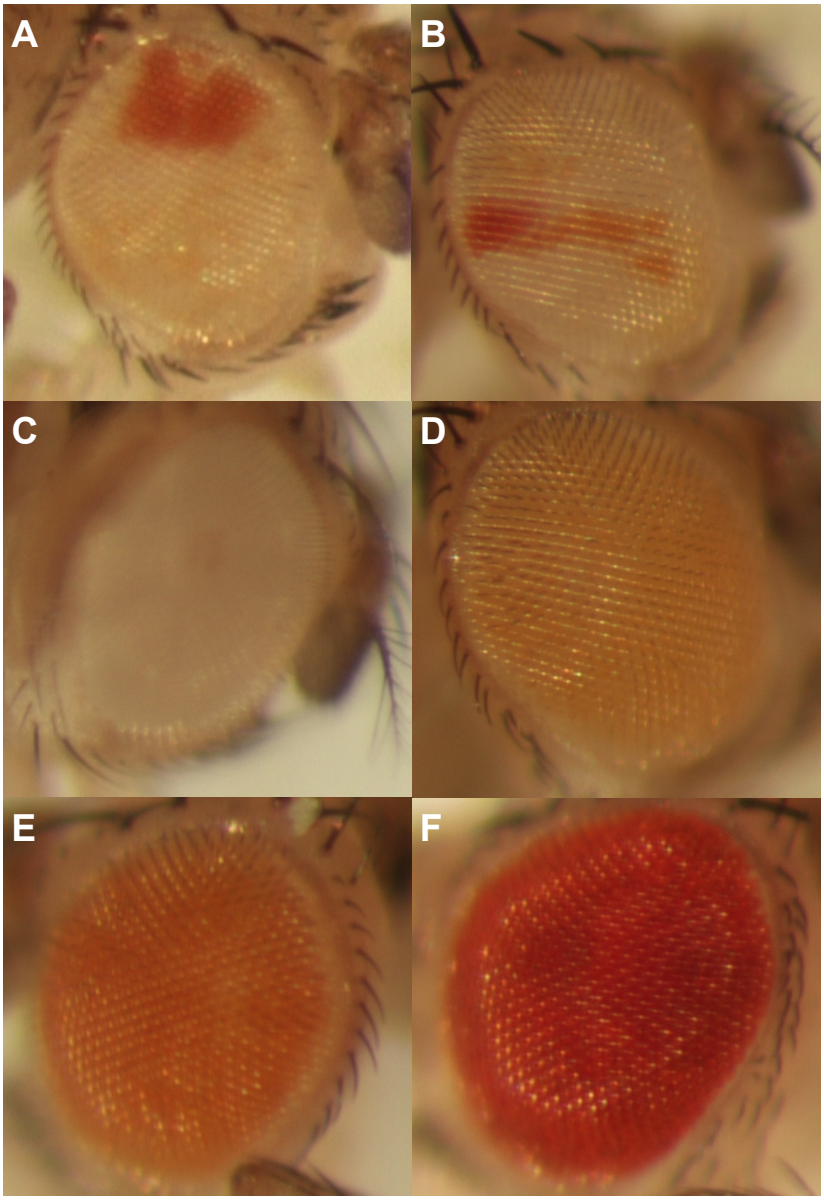



**Figure 3.3**: Eye colours observed during excision of *EP1526*.
**A and B**. Mosaic eye patterns observed in male offspring of Cross A caused by somatic excision of *EP1526*.
**C-F**. Range of eye colours observed in males in the stable stocks set up in Cross C. White (**C**) is the background colour and flies with white eyes most likely no longer contain *EP1526*. Light orange (**D**) is the colour given by *EP1526* in its original location. *EP1526* has most likely not moved in flies with light orange eyes. Colours darker than light orange (**E** and **F**) are caused my increased expression of *EP1526*. In these lines *EP1526* must have been excised and re-inserted into the genome at a point where it is more highly transcribed.

Orange or red eye colour suggests the P-element has been excised and subsequently re-integrated somewhere else in the genome, where it is more highly transcribed. White eye colour suggests the P-element been excised and either been lost entirely or has been re-integrated back into a region of genome where it is not transcribed. The P-element could have been excised in the yellow eyed lines, and re-integrated again in a location where its level of transcription is comparable to its original position, but it is more likely the P-element has not moved. 15 of the 450 stocks were hemi/homozygous lethal.

### 3.1.4 PCR screens to detect *pacman* alleles

Two PCR screens were carried out to identify any *P** chromosomes with deletions into *pcm*. The flies used in the screens were *P*/Y* males. *P*/P** females could also have been used, but would have taken an extra generation to appear in the stocks. *P*/FM7c* flies were unsuitable for use as the *FM7c* chromosome contains a wild-type copy of *pcm*, which would mask any lesions to the *P* pcm* locus in a PCR screen. The 15 hemi/homozygous lethal stocks could not be tested using this method as *P*/Y* and *P*/P** flies were not attainable. Screen 1 (Figure 3.4, panel B) utilised primers designed to amplify an area in the 5' of *pcm* (jls1 and jls2). Failure to amplify this region indicates a deletion into *pcm*. Each PCR was carried out in duplex with a set of primers designed to amplify a distal gene (*dob1*, primers dob1F and dob1B) as an internal control, to eliminate false positives. Screen 2 (Figure 3.4, panel C) was carried out in duplex with two sets of test primers and no internal control, as it was considered less likely that any mutations would be found in this screen that hadn't already been picked up in Screen 1. The two primer sets amplified regions in the 5' (SFN30 and jes6) and middle of the *pcm* gene (DpcmF2 and DpcmR2).

Screen 1 identified 33 lines that failed to amplify the region covered by the test primers. Screen 2 did not identify any lines that failed to amplify the test primers. The 33 lines picked up in Screen 1 were also tested in Screen 2, but amplified both regions successfully, indicating these were not full deletions of *pcm*.

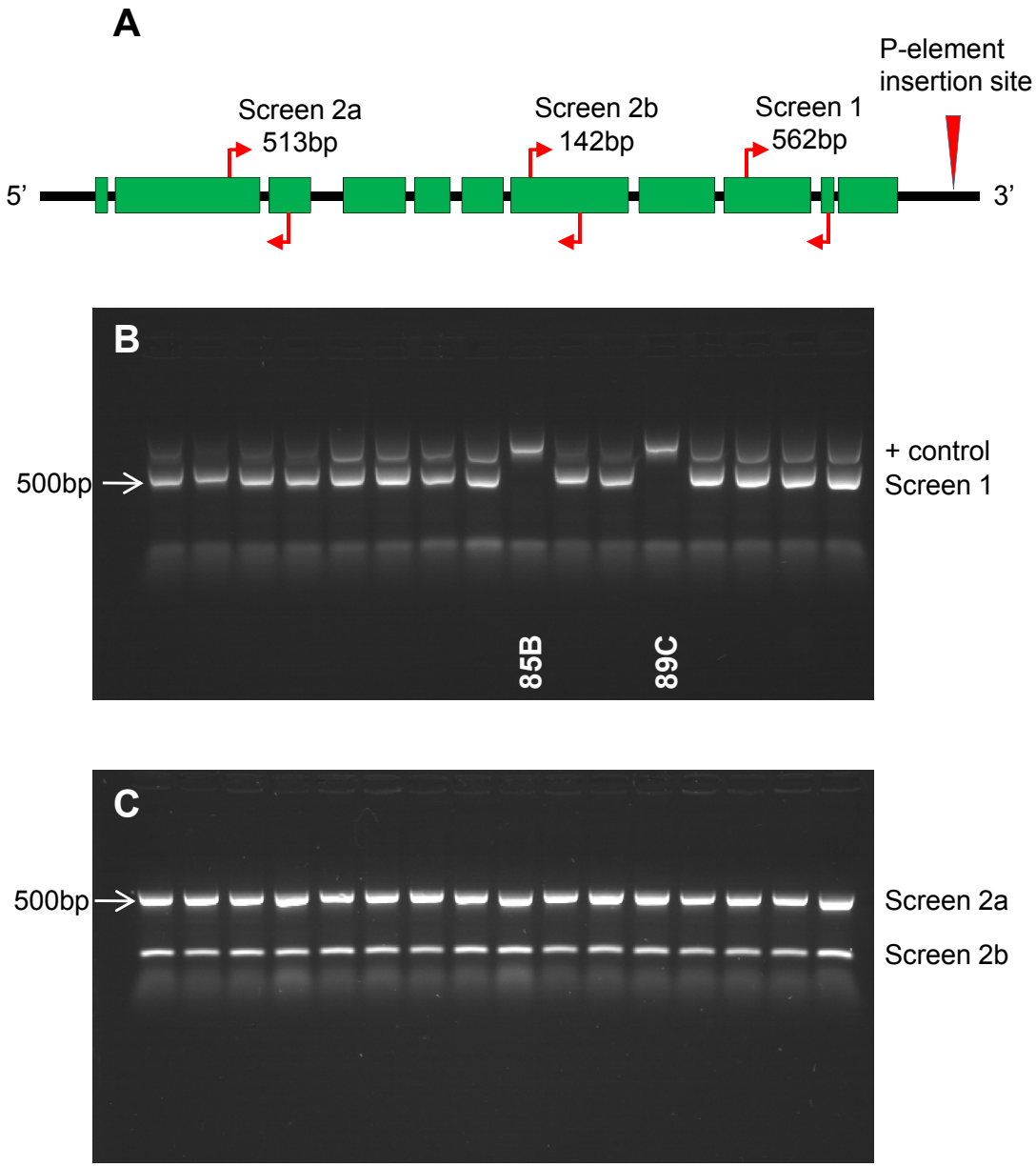



**Figure 3.4**: PCR screens used to detect *pcm* mutations in lines created by excision of *EP1526*.
**A.** Schematic of the *pcm* locus. Green blocks represent exons of *pcm,* locations of primers used for each screen are shown as red arrows and the red triangle shows the original insertion point of *EP1526*.
**B.** Example of a Screen 1 gel showing two positive results (lines carrying a *pcm* mutation). 33 lines failed to produce a band with the Screen 1 primers.
**C.** Example of a Screen 2 gel. No lines were found that did not amplify the regions covered by either the Screen 2a or 2b primers.

### 3.1.5 PCR to eliminate false positives due to non-disjunction or non-virgin females.

Cross B of the excision scheme is flawed in that it provides possibility for contamination of the stocks produced in Cross C by the chromosome containing *pcm*$^5$, because *pcm*$^5$*/FM7c* flies are used as a source of *FM7c*. This would cause false positives - lines that do not amplify the region tested in Screen 1, but are not novel deletions. There are two ways in which contamination could be introduced by the *pcm*$^5$*/FM7c* females used in the cross: the use of non-virgin females or chromosomal non-disjunction (Bridges, 1916). The use of non-virgins is easily avoided as selection of virgin females is a basic fly pushing technique. However, chromosomal non-disjunction is a natural phenomenon that will always occur in a minority of flies. Primary non-disjunction results in the creation of *XXY* female offspring when during egg development, the two *X* chromosomes of the mother fail to disjoin and create an egg containing two *X* chromosomes. Subsequent fertilisation of the *XX* egg by a *Y* containing sperm will produce a viable *XXY* female (*XXX* is lethal). This *XXY* daughter will appear the same as the mother, as they both contain the same two *X* chromosomes. This could have happened at any time in the *pcm*$^5$*/FM7c* stocks, meaning the virgins taken from this stock could have been *pcm*$^5$*/FM7c/Y*. Such *XXY* females will produce a variety of eggs at unequal frequencies. *XX* eggs fertilised by *Y* sperm will produce daughters identical to their *XXY* mother, whereas *XY* eggs fertilised by *X* sperm will produce females that appear as expected due to their maternal and paternal *X* chromosomes, but will be capable of producing daughters identical to themselves in future generations. Either of these two situations could result in the virgins used in Cross C containing a *pcm*$^5$ chromosome, which upon establishment of the stocks could produce *pcm*$^5$*/Y* males, which would have been detected as *pcm* mutants in Screen 1.

To identify and eliminate false positive lines, primers were used to amplify the region spanning the *pcm*$^5$ deletion (Figure 3.5). These primers (pcm5F and pcm5R) give a band of 867bp or 320bp in wild-type and *pcm*$^5$ mutant flies respectively. *P*/Y* flies identified as having a *pcm* mutation by Screen 1 will give a wild-type size band if the deletion does not extend to the area covered by the primers or no band if the deletion extends past the primer lying closest to the 3' of *pcm*. If a *pcm*$^5$ sized band is produced, then the line is likely contaminated by *pcm*$^5$. Of the 33 stocks identified as mutants in Screen 1, only one gave a band of a different size to what would be expected for *pcm*$^5$. This strongly suggests 32 of these lines contain the *pcm*$^5$ lesion and are not novel deletions.

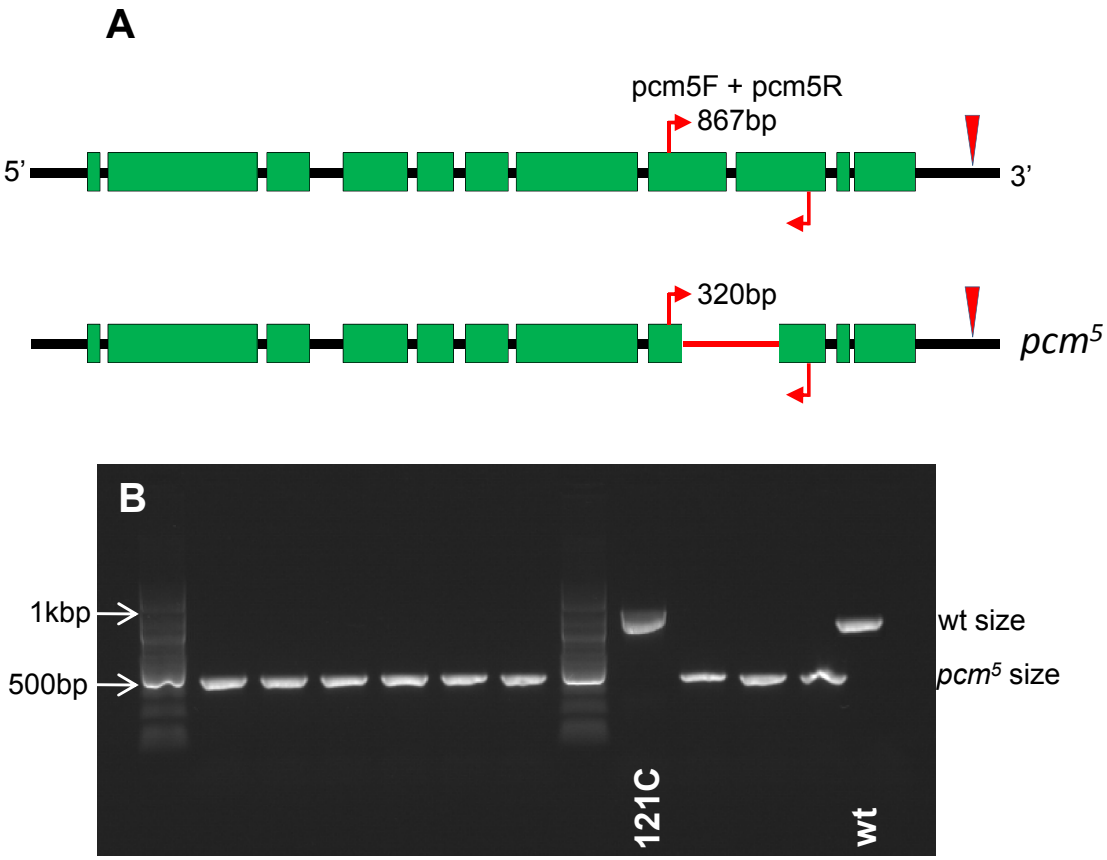



**Figure 3.5**: PCR used to test *pcm* mutations identified in Screen1 for false positives (contamination by the allele *pcm*$^5$ due to the use of non-virgin offspring or chromosomal non-disjunction).
**A.** PCR primers used to detect the *pcm*$^5$ deletion. The wild-type *pcm* locus is shown above and the *pcm*$^5$ locus below. Green blocks represent exons, red arrows show the location of primers and the red line represents deleted sequence in *pcm*$^5$. The red triangle shows the original insertion site of *EP1526*.
**B.** Example gel showing the single line that does not produce a band of the size expected for the *pcm*$^5$ mutation.

### 3.1.6 Sequencing of probable false positives

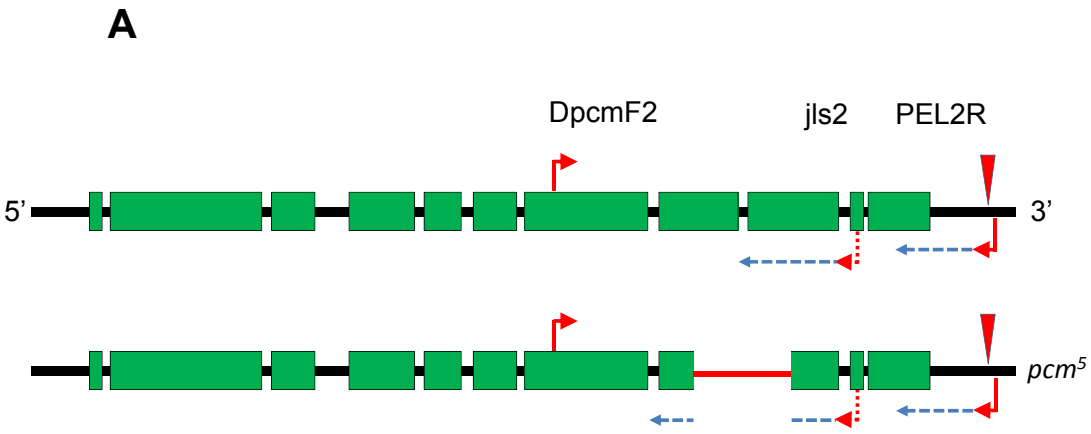
It is possible that the lines that appear to be the same as *pcm*$^5$ are actually either recreations of *pcm*$^5$ or similar but slightly different deletions. To test this possibility, the 3' end of *pacman* was amplified and sequenced at two points (Figure 3.6). The primer jls2 was used to sequence the breakpoints of the deletions and PEL2R was used to sequence the area from which *EP1526* has been excised (referred to as the "footprint"). When P-elements are excised, they leave a short sequence of nucleotides in the position they used to occupy. This sequence is variable, and if it is identical in two lines, it suggests they have a common origin. All the breakpoint regions sequenced were the same as in *pcm*$^5$ (Figure 3.6, panel B shows two examples), indicating that the deletions present are the same as the *pcm*$^5$ deletion. All the footprint regions sequenced were also the same (Figure 3.6, panel C), strongly suggesting that these are not independent recreations of the *pcm*$^5$ deletion. Although it cannot be determined for certain if these 32 lines contain the *pcm*$^5$ mutation due to chromosomal non-disjunction, the use of non-virgin females, recreation of the deletion, or a mixture of the three, it seems most likely that production of false positives on this scale was due to non-disjunction. All these lines were discarded.

## 3.2 Analysis of the new, non-lethal allele 121C

One line, referred to as 121C, appeared to contain a novel deletion into *pacman*. Screen 1 showed there was a deletion into the gene and the test for false positives showed that the deletion could not extend further than the 3' of exon 9.

### 3.2.1 Molecular lesion in 121C

To narrow down the possible lesion in 121C, intronic primers were used to amplify each exon of *pcm* individually to see which were still present. This made it possible to see which exons were fully or partially deleted, and provided further evidence that the lesion detected towards the 3' of *pcm* was the only one present in the gene in this line. For convenience, PCR conditions were used that allowed individual PCRs for each exon to be performed at the same time and run on a gel together. Figure 3.7 shows which exons (numbered 1-11) are present in wild-type, *pcm*$^5$ and 121C. Primer pairs were named PcmEx1F + PcmEx1R to Pcm11ExF + PcmEx11R. The area from the immediate 3' of the *pcm* gene to 91bp past the *EP1526* insertion site was also tested (labelled "PEL", primers PEL1F and PEL3R). All 12 bands are present at their correct sizes for wild-type, and all bands are




## B. Sequence from jls2 (*pcm*$^5$ breakpoint)

```
wt (FlyBase)  ATGCGTTGGCGTGCTGCATGCGCGGGTGCACCCAC//AGTTTTTGTGGGCGGAGCAACAAAAATACTACCGCCC
wt (w1118)    ATGCGTTGGCGTGCTGCATGCGCGGGTGCACCCAC//AGTTTTTGTGGGCGGAGCAACAAAAATACTACCGCCC
85B           ATGCGTTGGCGTGCTGCATGCG-------------//--------------GAGCAACAAAAATACTACCGCCC
109C          ATGCGTTGGCGTGCTGCATGCG-------------//--------------GAGCAACAAAAATACTACCGCCC
pcm5          ATGCGTTGGCGTGCTGCATGCG-------------//--------------GAGCAACAAAAATACTACCGCCC
```

## C. Sequence from PEL2R (footprint)

```
wt (FlyBase)  TTCTGAGCGAGAGAGAACCATG-----------------------------------GAATAATGACCACACTTGAG
wt (w1118)    TTCTGAGCGAGAGAGAACCATG-----------------------------------GAATAATGACCACACTTGAG
85B           TTCTGAGCGAGAGAGAACCATGATGAAATAACATGTTATTTCATCATGGAGAGAACCATGGAATAATGCCACACTTGAG
109C          TTCTGAGCGAGAGAGAACCATGATGAAATAACATGTTATTTCATCATGGAGAGAACCATGGAATAATGCCACACTTGAG
pcm5          TTCTGAGCGAGAGAGAACCATGATGAAATAACATGTTATTTCATCATGGAGAGAACCATGGAATAATGCCACACTTGAG
```

**Figure 3.6**: Areas sequenced to verify the *pcm*$^5$ deletion in the probable false positive lines.
**A.** Primers used to amplify and sequence the breakpoint and footprint regions of the *pcm*$^5$ deletion . The wild-type *pcm* locus is shown above and the *pcm*$^5$ locus below. Green blocks represent exons, red arrows show the location of primers and the red line represents deleted sequence in *pcm*$^5$. The red triangle shows the original insertion site of *EP1526*. Blue dashed lines extending from primers estimate the section sequenced by the primer in each line.
**B.** Two possible *pcm*$^5$ lines (109C and 85B) compared to wild-type sequence from a control line (*w*$^{1118}$), FlyBase and *pcm*$^5$.
**C.** The "footprint" sequence left after excision of *EP1526* in 109C and 85B (blue) compared to *pcm*$^5$ and the wild-type sequence in *w*$^{1118}$.

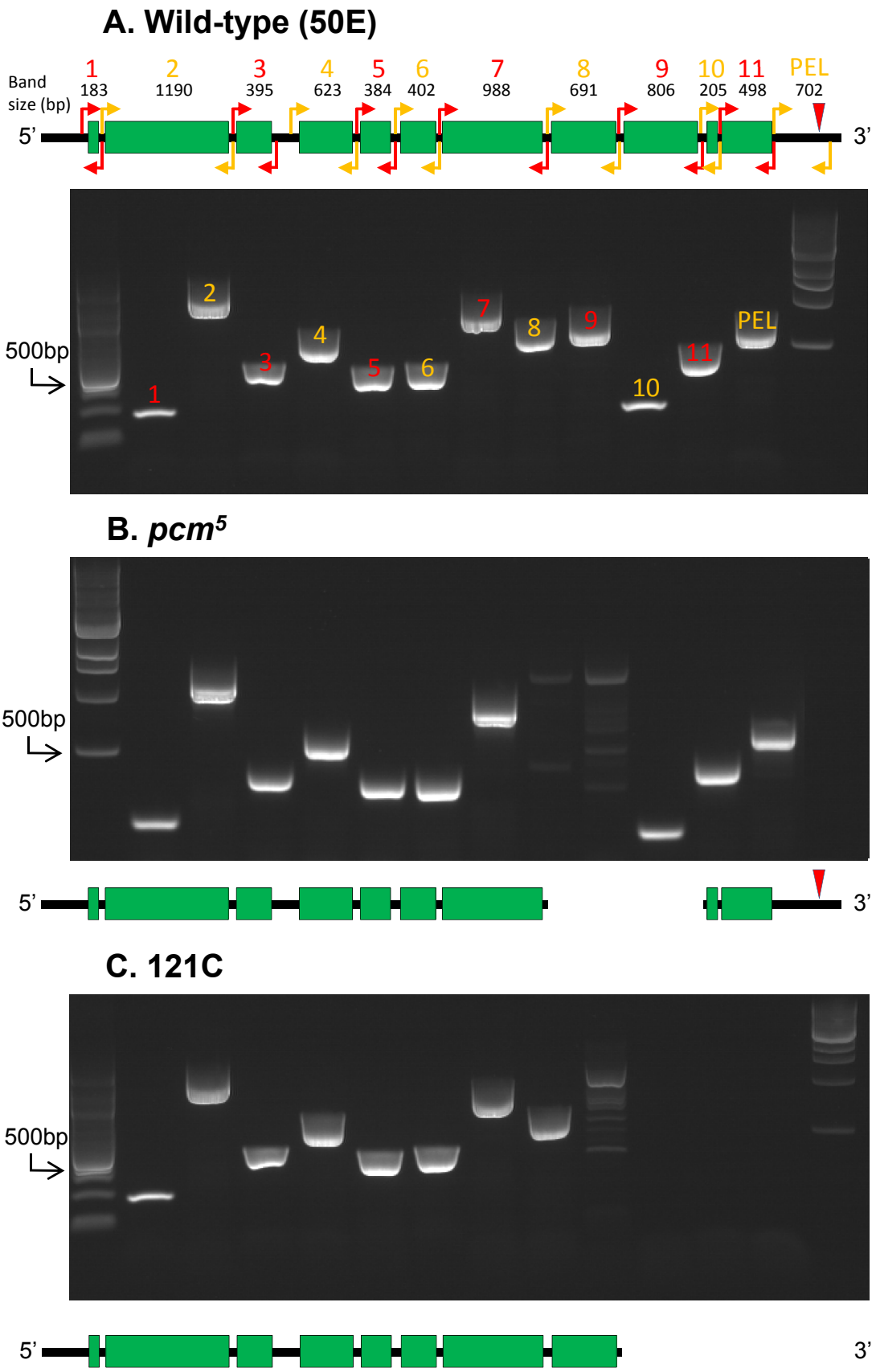



**Figure 3.7**: Amplification of individual *pcm* exons to narrow down the extent of the deletion in 121C.
**A.** Wild-type *pcm* with primers used to amplify each exon and the sequence immediately following *pcm*.
**B.** Exons that can be amplified in the known deletion, $pcm^5$.
**C.** Exons that can be amplified in the line 121C, showing that the deletion extends into exon 9.

present for *pcm*$^{5}$ except the bands representing exons 8 and 9, as the lesion in *pcm*$^{5}$ partially deletes both of these exons. Eight bands are present for 121C, representing exons 1-8, indicating that the 5' of the 121C lesion lies within exon 9 and the rest of the gene is unaffected. Combined with the result in Figure 3.5 where 121C gives a wild-type sized band, it can be inferred that the majority of exon 9 is present.

### 3.2.2 Sequence of the 121C allele

Initially it was assumed that the 3' of the 121C lesion could be inferred from the *pcm* alleles already available. With the exception of *pcm*$^{5}$, the 3' breakpoint for all the other *pcm* alleles is at the *EP1526* insertion site. However, it proved impossible to amplify the region between the primers jls1 and PEL1R, PEL2R or PEL3R (Figure 3.8 panel B). The jls1 site was known to be present and is located at the start of exon 9, which suggests the PEL1R/2R/3R sites are missing. This means that unlike any other *pcm* allele, the lesion in 121C extends in both directions – into *pcm* and possibly into the neighbouring gene *Nat1*. To test this, primers were designed at various points in *Nat1*. Figure 3.8, panels A and C show two such primers, one positioned in the intron before exon 5 (Nat15F) and the in the intron before exon 6 (Nat16F). These were paired with the forward primer for *pcm* exon 8, PcmEx8F. *pcm* and *Nat1* are in opposite orientations to each other in the genome, so primers named as "forward" primers for each gene form valid pairs when used together. PcmEx8F and Nat15F/Nat16F both produced bands smaller than what would be expected for wild-type (just under 2,000bp compared to 4,145bp and around 3,000bp compared to 4,881bp respectively). PcmEx8F and Nat16F were used to amplify and sequence the region, which contained a deletion of 2,222bp. 1,183bp of the deletion extended from the *EP1526* insertion site away from *pcm* towards *Nat1*, deleting 529bp of the final exon (exon 6). 1,039bp of the deletion extended towards *pcm*, deleting 590bp of the gene, including exons 10 and 11, and about two thirds of the intron between exons 9 and 10. 121C is a deletion into *pcm* and *Nat1*, which complicates its use for the investigation of *pcm*.

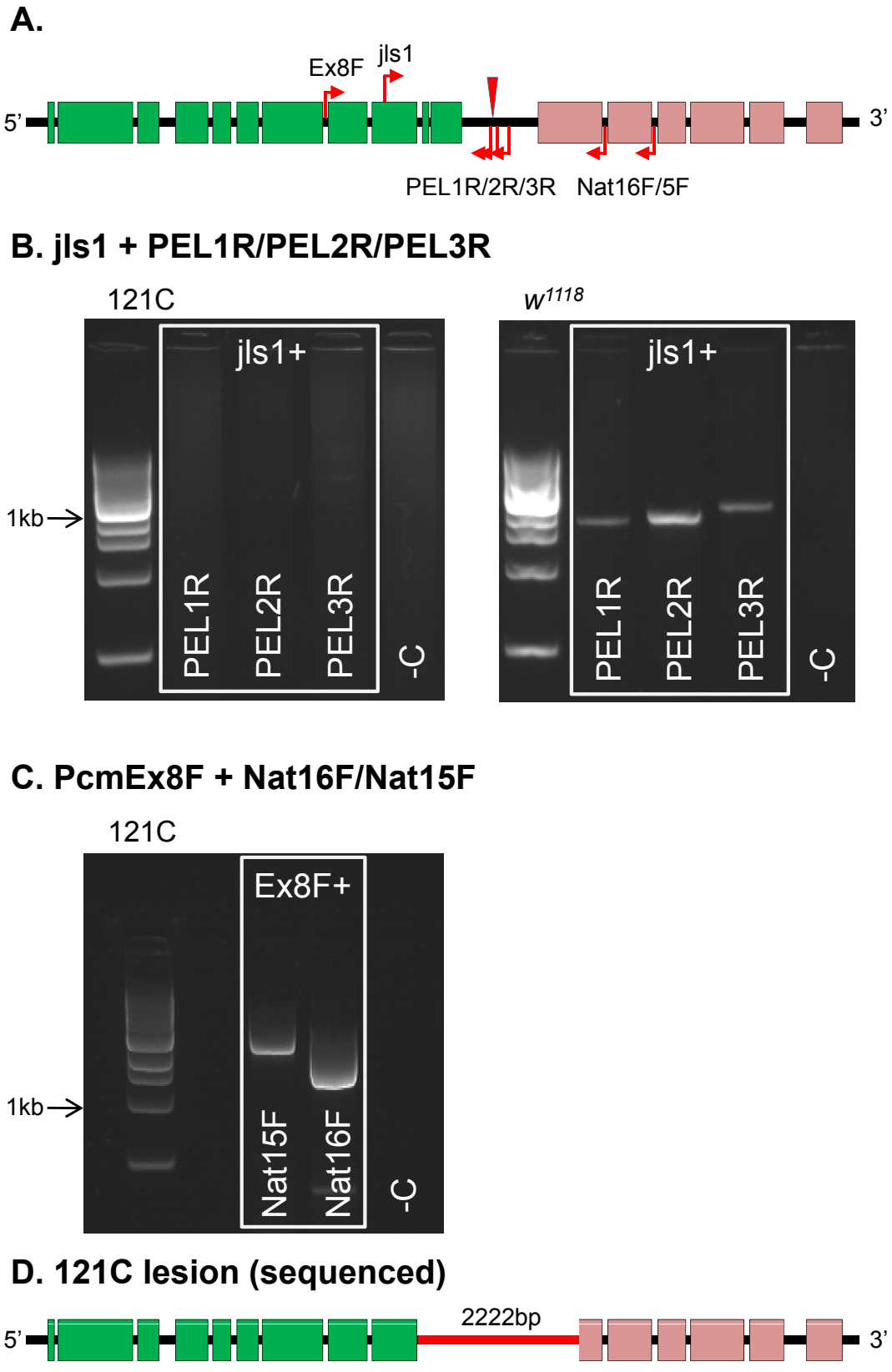



**Figure 3.8**: PCR and sequencing to determine breakpoints of the deletion in 121C.
**A.** Wild-type *pcm* with primers used for PCRs and sequencing. Green blocks represent *pcm* exons, pink blocks represent *Nat1* exons, red arrows show the location of primers and the red triangle shows the original insertion site of *EP1526*.
**B.** PCR showing the 3' breakpoint in 121C does not lie at the *EP1526* insertion site as no bands are produced with jls1 combined with PEL1R, PEL2R or PEL3R.
**C.** PCR with primers PcmEx8F combined with Nat15F and Nat16F showing the deletion does not extend further into *Nat1* than the Nat16F site.
**D.** Extent of the deletion in 121C determined by sequencing.

## 3.3 Identification and characterisation of control lines

The genetic background of stocks of flies can vary depending on their origin and the accumulation of random mutations while they are kept as separate lines. Ideally when testing a gene mutation, only the single gene of interest should be mutated between wild-type and control flies. In practice, genetically related flies provide a good control. In this case, *pcm* mutant flies should be compared to flies that have undergone the same process of P-element excision, but do not contain any deletions. This greatly reduces the chance of secondary mutations being present between the test and control lines that may impact the results of any experiment, as all lines share a common origin and therefore should have the same genetic background.

### 3.3.1 Selection of candidate control lines

The majority of lines created during the P-element excision were white eyed and did not appear as *pcm* deletions in Screen 1, which suggest *EP1526* has been precisely excised. 24 of these lines were selected as possible control lines

### 3.3.2 PCR and sequencing to identify precise excision lines

Primers PEL1F and PEL3R amplify the region immediately 3' of *pcm* to 91bp past the *EP1526* insertion site. In wild-type they produce a band of 702bp. The 24 candidate control lines were tested using these primers. No band was produced for 11 lines, suggesting deletions were present, either towards *Nat1* or towards *pcm* but not far enough to have been detected by Screen 1. These lines were not suitable as controls and were discarded. 13 lines produced bands, seven of which were around the same size as wild-type. Six lines produced bands larger than wild-type (Figure 3.9, panel B). Sequencing with the same primers showed the lines that gave larger bands still contained some of *EP1526* (Figure 3.9, panel C) and these were also unsuitable for use as controls. The seven lines with bands the same size as wild-type were sequenced and each contained a small "footprint" sequence between 28bp and 43bp (Figure 3.9, panel D). 22bp of the footprint was consistent in all seven lines, while the rest varied between lines. Any of these lines is suitable for use a control. 50E was randomly selected as the control to be used in future experiments.

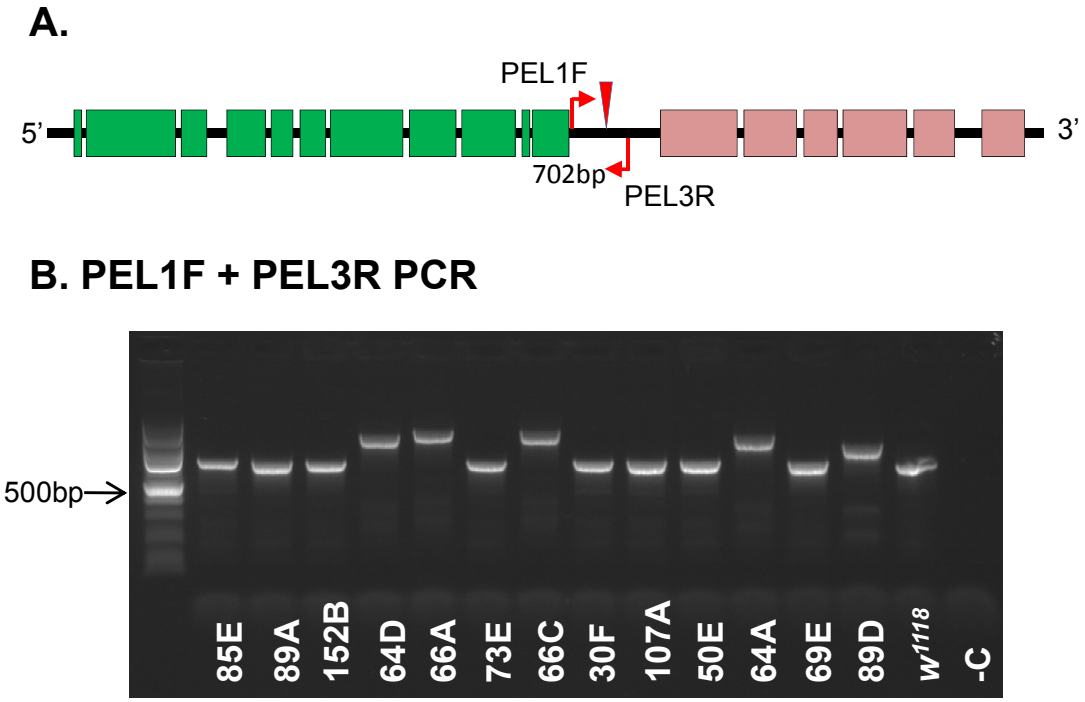



**Figure 3.9**: PCR and sequencing to find precise excision lines for use as controls.
**A.** Wild-type *pcm* with primers used for PCRs. Green blocks represent *pcm* exons, pink blocks represent *Nat1* exons, red arrows show the location of primers and the red triangle shows the original insertion site of *EP1526*.
**B.** PCR to check for complete excision of *EP1526*. Bands larger than wild-type suggest part of EP1526 may remain in these lines.
**C.** Sequences of 64D and 66C which produce bands larger than wild-type showing they align with the wild-type sequence up to the *EP1526* insertion site (bright blue) and with the sequence of *EP1526* after the insertion site (bright green).
**D.** Sequence of each of the precise excision lines compared to wild-type. Each contains a variable "footprint" left after the excision of *EP1526*. Yellow blocks indicate alignment with wild-type sequence, dark blue blocks show the minimal consistent footprint sequence and light blue blocks show the variable region of the footprint sequences.

## 3.4 Characterisation of lethal lines created by P-element excision

During the P-element excision, 15 lines were identified that did not produce any hemi- or homozygous flies carrying the *X* chromosome from which *EP1526* may have been excised (*P**). Lethal *P** chromosomes are referred to as *L* chromosomes. Whatever lesions have occurred in these lines were 100% lethal if no alternative non-mutated region was present. It is possible the lethality could be due to a deletion that includes *pcm*, or due to an unrelated event such and reintegration of *EP1526* into a different gene elsewhere on the chromosome. These flies could not initially be tested using Screen 1 as the flies available containing the *L* chromosomes were heterozygotes that contained a wild-type copy of *pcm* (*L/FM7c*).

### 3.4.1 Screen to identify lethal lines containing *pacman* mutations

To allow the lethal chromosomes to be tested using the primers from Screen 1, *L/FM7c* females were crossed to $pcm^5/Y$ males to produce $L/pcm^5$ offspring. These flies could be tested using Screen 1 as the $pcm^5$ deletion includes the site for the jls1 primer. Using this method, the three lethal lines 16A, 50B and 50C were found to contain deletions into *pcm* (Figure 3.10).

### 3.4.2 Pupal lethality of the lethal *pacman* mutants

$L/pcm^5$ flies can be used to see if there is a mutation into the 3' of *pcm*, but cannot reveal any information about the area outside that covered by the $pcm^5$ deletion. As no non-lethal mutations or deficiencies exist that include a deletion of *pcm*, and as the possible size of the deletion in the *L* chromosomes was unknown, the only direct way of determining the extent of the lesion would be to sequence the genomic DNA of *L/L* or *L/Y* flies. Adults with these genotypes do not exist, but *L/L* or *L/Y* flies may exist up to a certain point before they die. To test this, the *FM7c* balancer was crossed out of the lethal stocks and replaced with a version of the balancer *FM7i* that includes constitutively expressed *actin-GFP* (*FM7i*, $P\{w^{+mC}=ActGFP\}JMR3$, referred to as just *FM7i*). This allows larvae hemi- or homozygous for the *L* chromosome to be picked out under UV light, as they do not fluoresce, unlike their *FM7i* siblings. The three lethal lines containing *pcm* mutations all

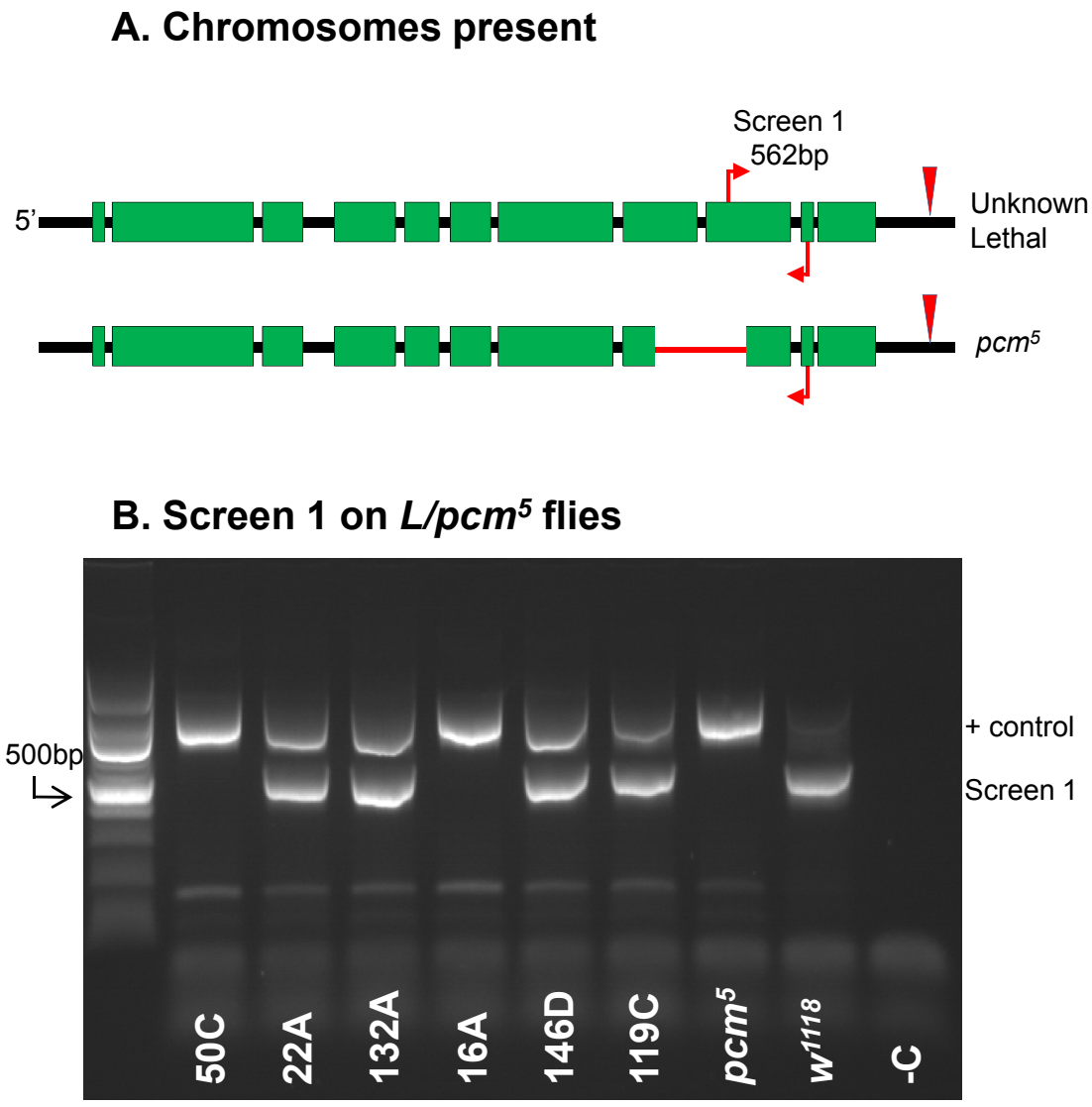



**Figure 3.10**: PCR to check for *pcm* deletions in the lethal chromosomes.
**A.** Chromosomes present in the *lethal/pcm*$^5$ flies showing primers used for PCRs. Green blocks represent *pcm* exons, red arrows show the location of primers and the red triangle shows the original insertion site of *EP1526*. Only the lethal chromosome is capable of producing a band with the Screen 1 primers if it does not contain a deletion in this region.
**B.** Example gel showing the result of the Screen 1 PCR on *lethal/pcm*$^5$ flies. 16A, 50C and 50B (not shown) showed mutations in *pcm*.

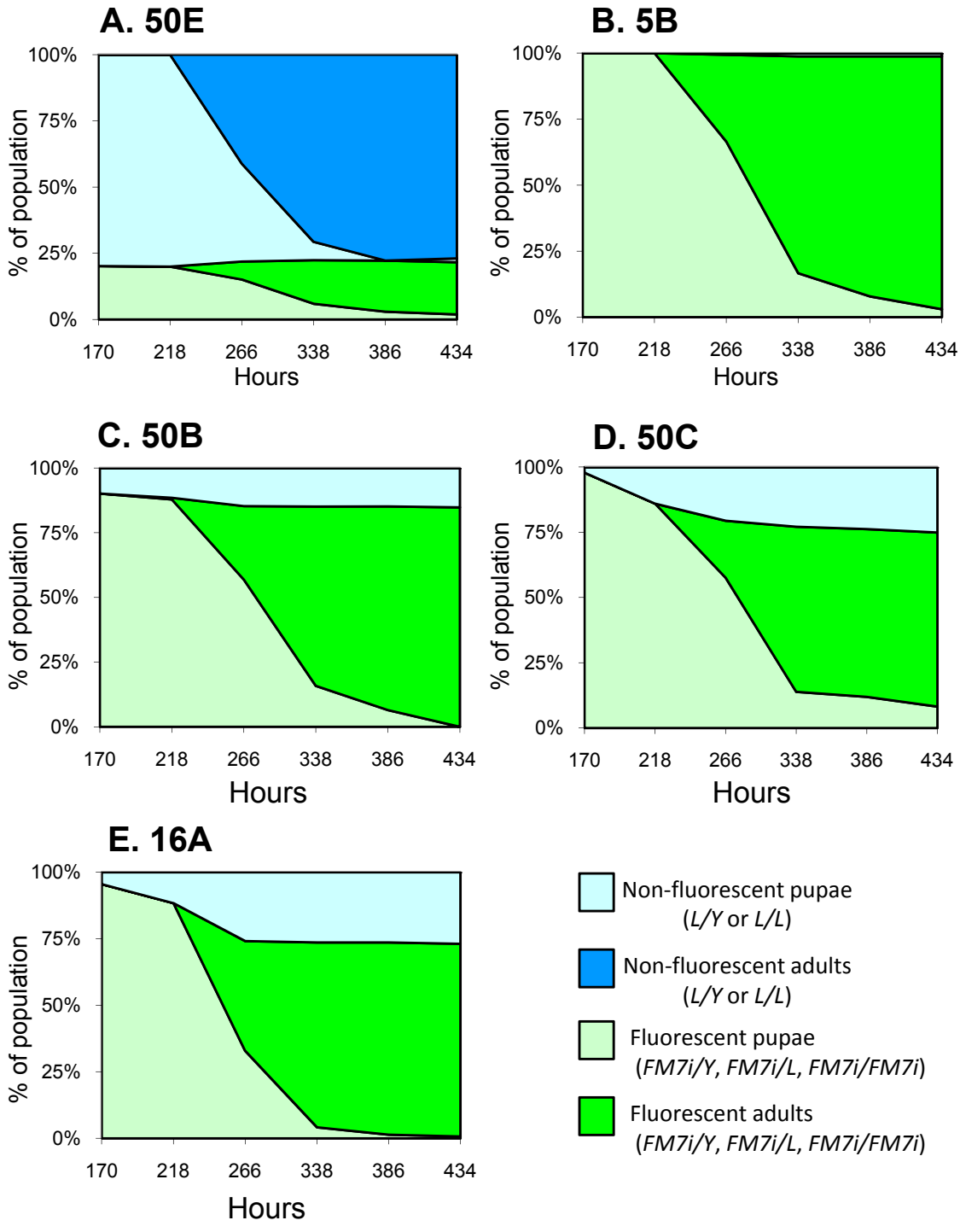



**Figure 3.11**: Survival graphs for lethal lines containing *pcm* mutations.
**A.** Wild-type control showing frequency of fluorescent and non-fluorescent of offspring observed as pupae and adults.
**B.** A non-*pcm* mutant embryonic lethal line showing frequency of fluorescent and non-fluorescent of offspring observed as pupae and adults.
**C-E.** 50B, 50C and 16A lines showing frequency of fluorescent and non-fluorescent of offspring observed as pupae and adults. Non-fluorescent pupae are seen but adults are not, showing that the lethality in these lines occurs during pupation.

produced homozygous larvae that died during pupation. Figure 3.11 shows survival graphs for five lines – 50E as a wild-type control, 5B as an example of an embryonic lethal, and the three *pcm* mutants that were also pupal lethals – 16A, 50B and 50C. No hemi- or homozygous flies are seen for 5B as they die before hatching as larvae. For the three *pcm* mutants, hemi- and homozygous pupae are seen, but no non-fluorescent adults ever eclosed. Both fluorescent and non fluorescent pupae and adults are seen for the non-lethal control 50E. There are two possibilities for the *pcm* mutant lines. Either they contain a *pcm* mutation that causes lethality at pupation, or they contain a viable *pcm* allele alongside another mutation that causes lethality at pupation.

### 3.4.3 Molecular characterisation of the lethal *pacman* mutations

As L3 larvae exist for each of the lethal lines identified as *pcm* mutants, genomic DNA was prepared from these and the lesions present determined molecularly. The same method as used to narrow down the extent of the deletion in 121C was used to find out which exons of *pcm* were deleted in the lethal mutants. Figure 3.12 shows the results for 50B and 50C, the exon profiles of which came out the same as what is seen for $pcm^5$ (only exons 8 and 9 are missing). Figure 3.13 shows the result for 16A, where only exons 1-5 are fully present.

### 3.4.4 Sequencing of the lethal *pacman* mutations (16A, 50B and 50C)

16A and 50B/50C were sequenced using different primers as the nature of the 16A deletion was different to the deletion in 50B and 50C. Figure 3.14, panel A shows the primers used for 16A sequencing (SFN22 in exon 4, and PEL3R, 91bp downstream of the *EP1526* insertion site). Together these primers produce a band slightly larger than 500bp (compared to 4,132bp expected for wild-type). Sequencing showed the deletion to be 3,501bp extending 3,068bp into *pcm* leaving exons 1-5 as wild-type, but deleting the last quarter of exon 6 (76bp of 320bp). In common with the other *pcm* alleles with deletions extending from the *EP1526* site, a small insertion of 9bp of footprint sequence was also present in 16A (Figure 3.14, panel B).

50B and 50C were sequenced with the same primers previously used to sequence the breakpoint and footprint regions in $pcm^5$ (Figure 3.14, panel C). Each sequence was identical in all three lines – a 531bp deletion with the same breakpoints in exons 8 and 9

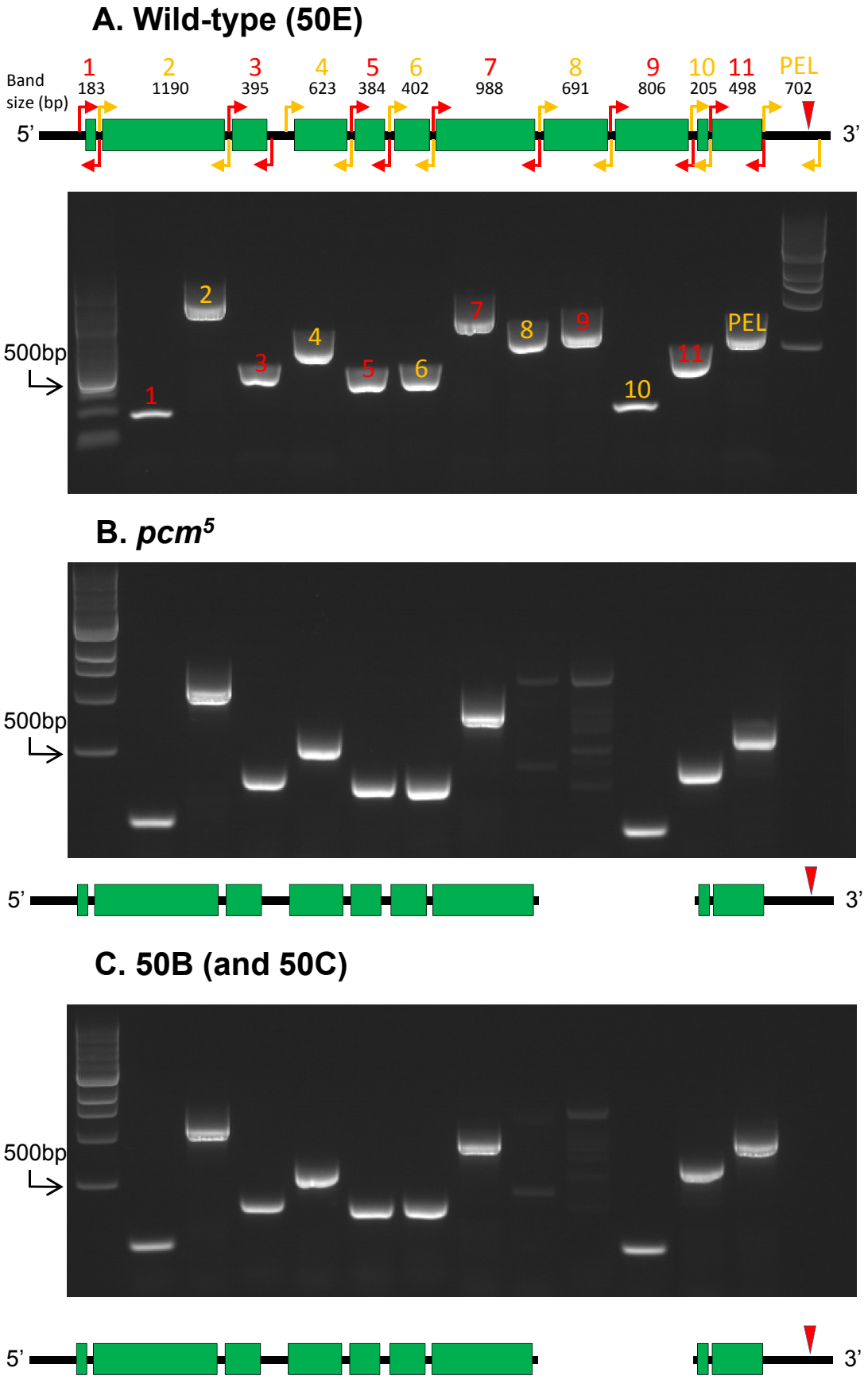



**Figure 3.12**: Amplification of individual *pcm* exons to narrow down the extent of the deletion in 50B and 50C.
**A.** Wild-type *pcm* with primers used to amplify each exon and the sequence immediately following *pcm*.
**B.** Exons that can be amplified in the known deletion, $pcm^5$.
**C.** Exons that can be amplified in the lines 50B and 50C, showing that the deletion extends no further than the deletion in $pcm^5$.

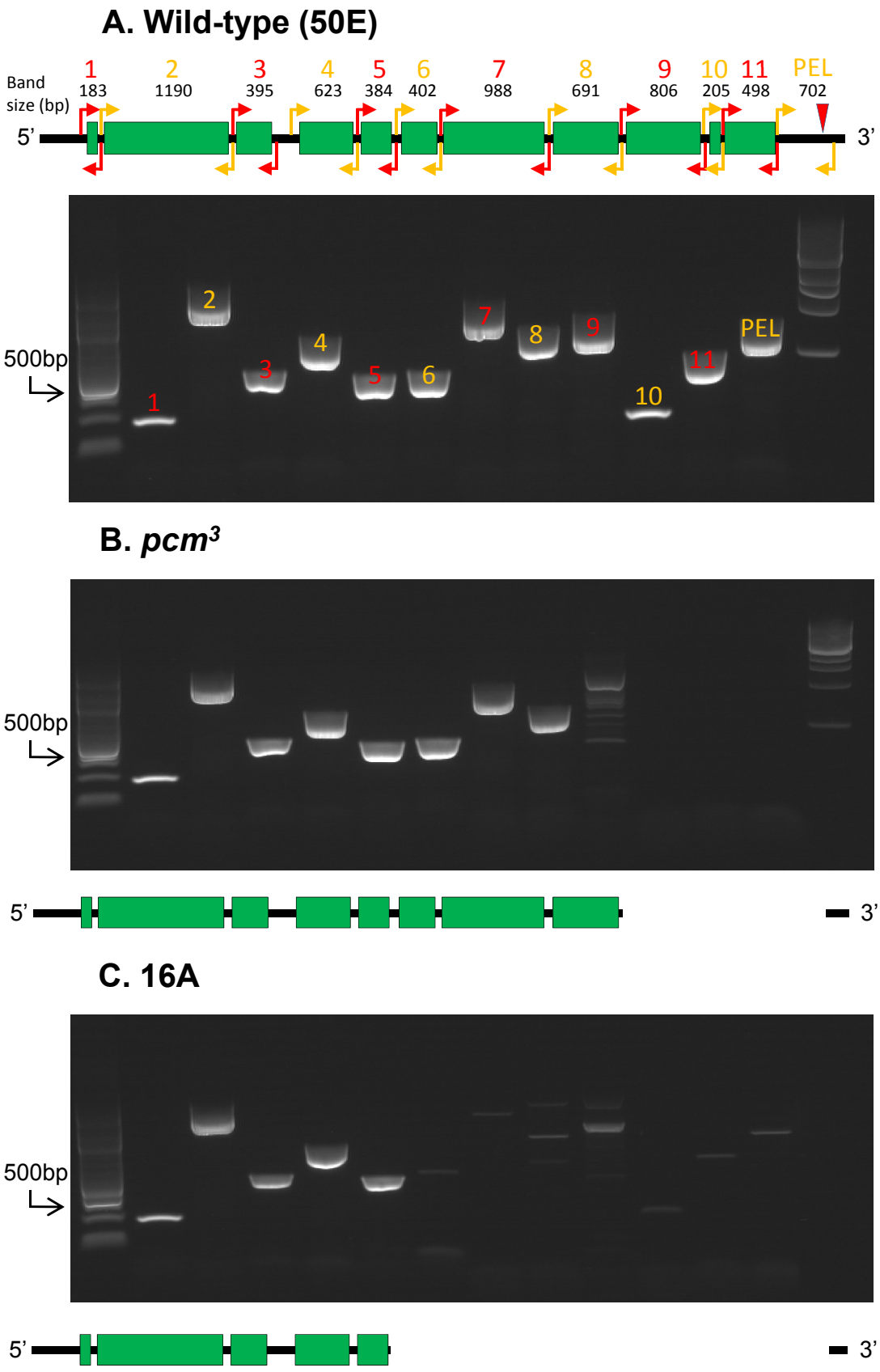
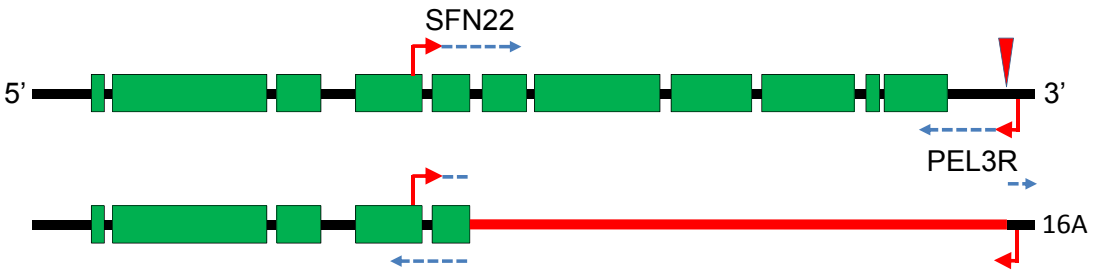



**Figure 3.13**: Amplification of individual *pcm* exons to narrow down the extent of the deletion in 16A.
**A.** Wild-type *pcm* with primers used to amplify each exon and the sequence immediately following *pcm*.
**B.** Exons that can be amplified in the known deletion, $pcm^3$.
**C.** Exons that can be amplified in the line 16A, showing that the deletion extends no further than exon 6.

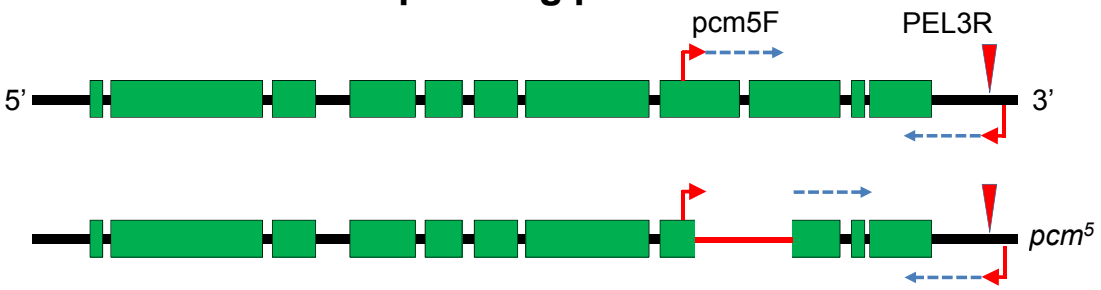
## A. 16A sequencing primers




## B. 16A breakpoint sequences

```
wt (FlyBase)  ACGATCTTCTTGAGAATGCGCAAGAATTTCCTCA//TCCTCAAGTGTGGTCATTATTC---------CATGGTTCTCTC
16A           ACGATCTTCTTG----------------------//----------------------TTATTTCATCATGGTTCTCTC
```

## C. 50B and 50C sequencing primers




## D. 50B and 50C breakpoint sequences

```
wt (FlyBase)  GCGGTAGTATTTTTGTTGCTCCGCCCACAAAAACTCC//GCATGCAGCACGCCAACGCATTCCACGCTGGCGTCAACA
pcm5          GCGGTAGTATTTTTGTTGCTCCG--------------//----------------CGCATTCCACGCTGGCGTCAACA
50B           GCGGTAGTATTTTTGTTGCTCCG--------------//----------------CGCATTCCACGCTGGCGTCAACA
50C           GCGGTAGTATTTTTGTTGCTCCG--------------//----------------CGCATTCCACGCTGGCGTCAACA
```

## E. 50B and 50C footprint sequences

```
wt (FlyBase)  TCCTCAAGTGTGGTCATTATTCC--------------------------------ATGGTTCTCTCTCGCTCAGAACA
pcm5          TCCTCAAGTGTGGTCATTATTCCGTTCTCTCCATGATGAAATAACATGTTATTTCATGGTTCTCTCTCGCTCAGAACA
50B           TCCTCAAGTGTGGTCATTATTCCGTTCTCTCCATGATGAAATAACATGTTATTTCATGGTTCTCTCTCGCTCAGAACA
50C           TCCTCAAGTGTGGTCATTATTCCGTTCTCTCCATGATGAAATAACATGTTATTTCATGGTTCTCTCTCGCTCAGAACA
```

**Figure 3.14**: Sequencing of breakpoints and footprint sequences in 16A, 50B and 50C.
**A.** Wild-type *pcm* and the predicted deletion in 16A with primers used for sequencing. Green blocks represent *pcm* exons, red arrows show the location of primers, the red triangle shows the original insertion site of *EP1526* and dashed blue lines extending from primers estimate the area sequenced by each.
**B.** The breakpoints in 16A. Yellow shows alignment with wild-type sequence, bright blues show were the sequence is only present in wild-type (deleted in 16A) and bright green shows the small footprint insertion left by *EP1526* in 16A.
**C.** Wild-type *pcm* and the predicted deletion (*pcm*$^5$) in 50B and 50C with primers used for sequencing. Green blocks represent *pcm* exons, red arrows show the location of primers, the red triangle shows the original insertion site of *EP1526* and dashed blue lines extending from primers estimate the area sequenced by each.
**D** and **E.** The breakpoints (sequence only present in wild-type shown in bright blue) and footprint sequences (bright green) of 50B and 50C compared to *pcm*$^5$. There are no differences between 50B, 50C or *pcm*$^5$.

and the same 32bp footprint (Figure 3.14, panels D and E). This strongly suggests the *pcm* mutation in 50B and 50C is the same as in $pcm^5$, however the $pcm^5$ deletion is not lethal. For this reason, we decided to continue investigating 50B and 50C for the possibility a second lesion is present in these lines within *pcm*.

### 3.4.5 Comparison between the 50B/50C lethals and $pcm^5$

The same complete exons can be amplified in 50B, 50C and $pcm^5$ but if there is a deletion or insertion into one of these exons in 50B or 50C leading to the lethality, the difference may be visible on a gel if it is large enough. Each exon was amplified from the three lines and run on an agarose gel, but no significant difference in size between any of the exons could be seen (Figure 3.15). A small deletion, insertion or point mutation would not be detectable on a gel, so sections covering the entirety of the *pcm* gene were amplified and sequenced. No difference was seen between $pcm^5$, 50B or the control line 50E, except for the $pcm^5$ deletion and footprint.

As no difference was seen in the coding sequence of *pcm* in 50B and $pcm^5$, it seemed very unlikely the lesion causing the lethality in 50B/50C could be due to an effect on *pcm*. The only remaining possibility was the lesion was in the *pcm* promoter region, however, the genomic sequence that constitutes the *pcm* promoter is unknown.

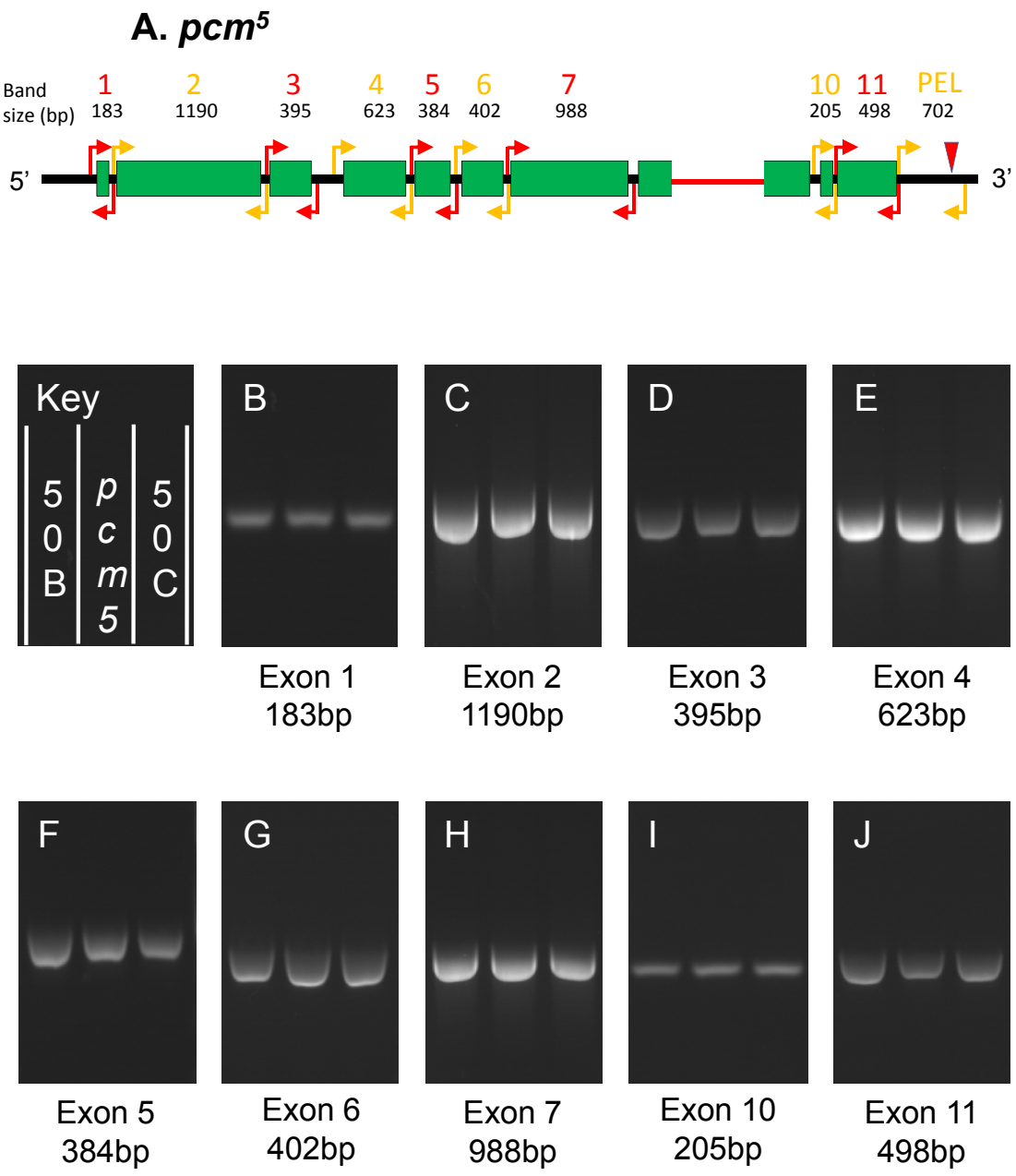



**Figure 3.15**: Comparison of exon sizes in 50B and 50C with *pcm*$^5$.
**A.** *pcm*$^5$ with primers used for amplification of individual exons. Green blocks represent *pcm* exons, red and yellow arrows show the location of primer pairs and the red triangle shows the original insertion site of *EP1526*. The red line shows the area deleted in *pcm*$^5$.
**B-J.** Individual amplifications of exons run next to each other on a gel to check for any variations in size large enough to observe. No differences were obvious.

## 3.5 Chapter summary

The objectives of the experiments detailed in this chapter were to create a null *pcm* allele, new hypomorphic alleles and control lines for future experiments. Three lines have been created that each satisfy or potentially satisfy one these objectives. 50E as a control line, 121C as a new hypomorphic allele and 16A as a possible null allele.

A number of precise excision control lines were produced, from which 50E was randomly selected as control for future experiments. 50E differs from fully wild-type *D. melanogaster* at two loci. It contains the $w^{1118}$ mutation, which gives white eyes, and a 43bp insertion left after the excision of *EP1526* in the intergenic region between *pcm* and *Nat1*. It is unlikely this affects expression of either *pcm* or *Nat1* as it lies downstream of both genes.

Two new *pcm* alleles were created – 121C and 16A. 121C was shown to be a new hypomorphic allele containing a 2,222bp deletion including 590bp of *pcm* and 529bp of the neighbouring gene *Nat1*. While the *pcm* deletion is unlike any previously created, the *Nat1* mutation makes the line difficult to work with, as any phenotypes or effects observed could be due to either gene. 16A shows most promise as a useful *pcm* allele for future experiments, and may satisfy the aim of creating a null allele. The deletion in 16A was determined to be 3,501bp including 3,068bp into the 3' of *pcm*. The line is not a genetic null for *pcm* as the first five exons and the area preceding the *pcm* gene, which mostly likely contains the *pcm* promoter, are intact. However, the allele may still be a protein or functional null. The deletion is larger than any of the previously generated mutations and is the first to be lethal. From this point on, 121C is referred to as $pcm^{13}$ and 16A as $pcm^{14}$. Figure 3.16 summarises the new and previously created *pcm* alleles.

### 3.5.1 Following chapter

This chapter has determined that the $pcm^{14}$ line contains a mutation to *pcm* and is lethal. Confirmation that the lethality is due to the lesion in *pcm,* the phenotypes of this allele and experiments to determine if it is a null, are covered in the next chapter. This includes complementation experiments with deficiencies and inter-allelic crosses between $pcm^{14}$ and the non-lethal *pcm* alleles. Crossover experiments were also performed for the lines $pcm^{14}$ and 50B to determine the section of chromosome from which the lethality stems.

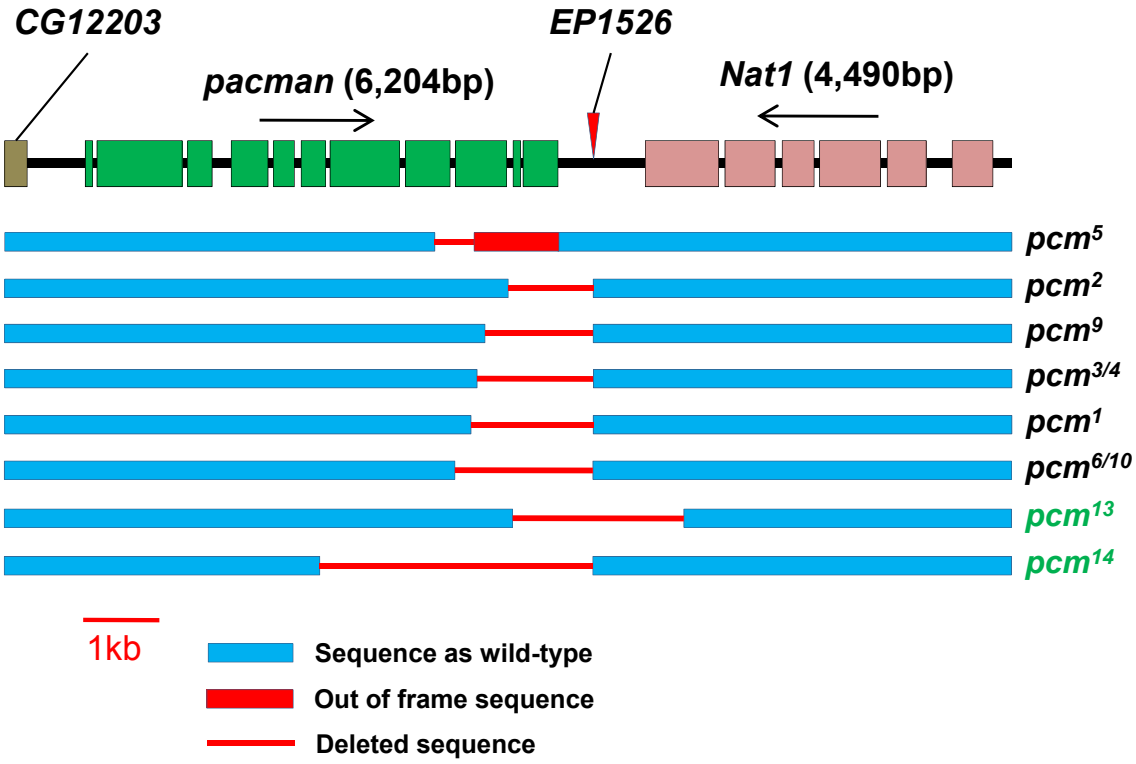



**Figure 3.16**: All *pcm* alleles created by imprecise excision of *P{w*$^{+mC}$*=EP}EP1526* (*EP1256*). Names of alleles created in this chapter shown in green. *pcm*$^{13}$ = 121C, *pcm*$^{14}$ = 16A. In order of ascending size: *pcm*$^{5}$ (516bp), *pcm*$^{2}$ (1,070bp), *pcm*$^{9}$ (1,378bp), *pcm*$^{3/4}$ (1,523bp), *pcm*$^{1}$ (1,558bp) *pcm*$^{6/10}$ (1,785bp), *pcm*$^{13}$ (2,222bp), *pcm*$^{14}$ (3,501bp). The undeleted *pacman* sequence after the deletion in *pcm*$^{5}$ is out of frame. The *pcm*$^{1}$ allele is no longer available.

# 4 Chapter 4: Genetic analysis of *pacman* mutants

## 4.1 Aims of Chapter 4

The previous chapter detailed the creation of two new *pcm* alleles, $pcm^{13}$ and $pcm^{14}$. This chapter will focus on $pcm^{14}$, as it is most likely to be the first null *pacman* allele. This chapter presents work performed to determine the nature and effects of the $pcm^{14}$ mutation.

### 4.1.1 Determination of the source of lethality in $pcm^{14}$ (and 50B/50C)

The first aim of this chapter was to determine whether the lethality in the chromosome containing the $pcm^{14}$ mutation stems from the *pcm* deletion, or from a second mutation elsewhere. This is important as no phenotypes can be ascribed to *pcm* until it is known that no other mutations are present. The location of the lethality was tested in two ways, by chromosomal crossover and by attempted rescue of the lethality by introducing a wild-type copy of *pcm*. The source of the lethality in lines 50B and 50C was also tested by crossover in case it was due to an undetected lesion at the *pcm* locus.

### 4.1.2 Comparison between $pcm^{14}$ and deficiencies

The second aim of this chapter was to observe the effect of combining non-lethal *pcm* alleles with deficiencies and $pcm^{14}$, to see if $pcm^{14}$ acts like a deficiency. Complementation tests with deficiencies normally play an important role in determining whether an observed phenotype is the result of a known mutation (i.e. by attempted generation of $pcm^{14}$*/deficiency* flies). However, in the case of $pcm^{14}$, generation of flies to test for complementation at the *pcm* locus is hindered by *pcm* being located on the *X* chromosome, as male flies containing hemizygous lethal mutations or deficiencies cannot exist. For this reason, comparison of non-lethal *pcm* with a deficiency or $pcm^{14}$ is required.

### 4.1.3 $pcm^{14}$ phenotypes and interallelic crosses

The third aim of this chapter was analysis of the phenotypes of $pcm^{14}$ and comparison to the phenotypes of the pre-existing *pcm* alleles. This includes TaqMan quantitative RT-PCR (qPCR) to measure the level of *pcm* in $pcm^{14}$ and other alleles, and interallelic crosses to examine the phenotypic effects of combining alleles.

### 4.1.4 Deficiencies covering the *pcm* locus

The final section of this chapter presents more detailed information about three deficiencies used in the experiments reported in the chapter. This includes creation of a new deficiency, *Df(1)ED7452* using the DrosDel Rearrangement Screen system.

## 4.2 Crossover experiments to determine the location of lethality in the lethal lines 16A ($pcm^{14}$), 50B and 50C

Many genes encoding phenotypic markers are known in *D. melanogaster* and are commonly used to map the chromosomal location of other genes. Markers can also be used to observe chromosomal crossover which occurs during prophase 1 of meiosis in female flies when the parental chromosomes align and can break and recombine with each other. Chromosomal crossover contributes to variation between individuals in natural populations with diverse gene pools, but the effect is rendered moot in inbred laboratory strains. It can also be prevented from occurring by the use of balancer chromosomes, which are a necessity for experimental crosses where the heredity of genes must be entirely predictable.

Allowing recombination to occur between a chromosome containing an unmapped gene/phenotype, such as a source of lethality, with a chromosome containing numerous recessive marker genes can greatly narrow down the area from which the lethality could stem. This can be determined by observing the combinations of markers present in the living hemizygous (male) progeny. The marker chromosome used to test the location of the lethality on the 16A and 50B chromosomes has the genotype $y^1\ cv^1\ v^1\ f^1\ car^1$ (stock BL1515). Hemi/homozygotes have yellow bodies, (due to $yellow^1$), missing posterior wing crossveins ($crossveinless^1$) and forked bristles ($forked^1$). They have orange eyes due to the combination of $vermillion^1$ and $carnation^1$. $v^1$ combined with wild-type *car* gives bright red eyes and $car^1$ combined with wild-type *v* gives dark red eyes (Figure 4.1). Aside from the *pcm* mutation, the 16A and 50B chromosomes contain one gene mutation – $w^{1118}$, which gives white eyes. This introduces a complication, as any chromosome combination with $w^{1118}$ will have white eyes irrespective of whether the *v* and *car* alleles present are mutant or wild-type. It should also be noted that *yellow* and *white* are close to each other in the genome, <2.5Mbp and around 1.2cM apart (Dexter, 1912, Sturtevant, 1913), so are unlikely to be separated by recombination. *car* and *pcm* are extremely close to each other (<75Kbp apart) and are therefore very unlikely to be separated. This means that a chromosome containing

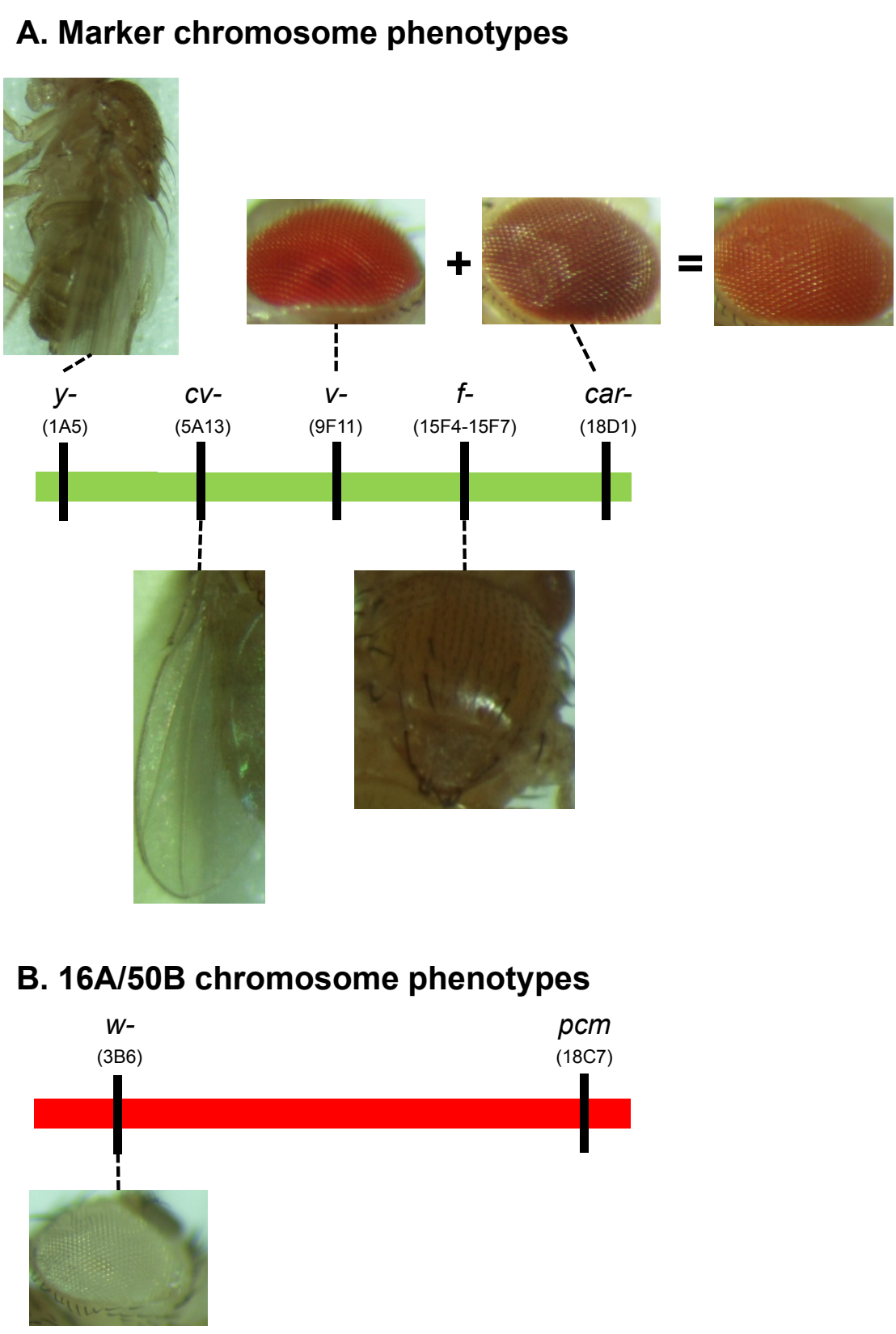



**Figure 4.1**: Phenotypes observed in crossover experiments.
**A.** Phenotypes or recessive markers on marker chromosome $y^1$ $cv^1$ $v^1$ $f^1$ $car^1$ from stock BL1515 and their approximate locations on the chromosome.
**B.** $w^{1118}$ phenotype on the chromosomes containing *pcm* deletions and the approximate locations of $w^{1118}$ and *pcm*.
Cytological locations are given in parentheses.

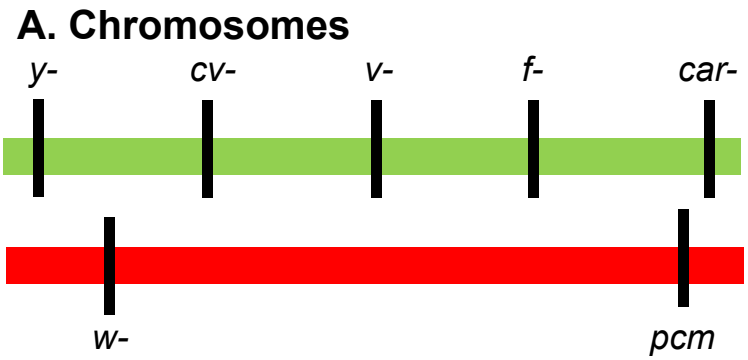
wild-type *car* almost certainly contains mutant *pcm*, and a chromosome containing $car^1$ almost certainly contains wild-type *pcm*.

### 4.2.1 16A ($pcm^{14}$)

16A was the only lethal line from the P-element excision that contained a novel *pcm* mutation. Observation of the male progeny (n=305) of a cross between $16A/y^1\ cv^1\ v^1\ f^1\ car^1$ females and *FM7i/Y* males showed the lethality on the 16A chromosome was localised to the region containing *pcm* (Figure 4.2). Slightly fewer than 50% of the flies observed were the product of chromosomal crossover. No combination of markers was observed that did not include the $car^1$ marker, which means no chromosome containing wild-type *car* (and therefore mutant *pcm*) was able to survive. 19 flies produced contained the entirety of the 16A chromosome up to at least the *f* locus. As these flies contained $w^{1118}$, the presence or absence of $car^1$ could not be directly observed. To ensure these flies were not *pcm* mutants, they were tested by PCR and each returned a band consistent with wild-type *pcm* (and therefore each almost certainly did contain $car^1$).

### 4.2.2 50B/50C

50B and 50C were two lethal lines that were identified as containing *pcm* mutations, but further analysis showed that the *pcm* mutation was likely the same as found in the allele $pcm^5$. As 50B and 50C had appeared identical in all experiments up to this point, and shared a common origin (same parents in Cross B of the P-element excision) it is reasonable to assume there is no difference between them. For this reason, the crossover experiment was only performed on 50B, and the results are assumed valid for 50C. As with the previous crossover experiment on 16A, slightly fewer than 50% of the offspring of the cross between $50B/y^1\ cv^1\ v^1\ f^1\ car^1$ and *FM7i/Y* were observed to have undergone chromosomal crossover (n=427). 88% of these flies contained $car^1$ and wild-type *pcm*. No flies were observed with wild-type *y* and the $w^{1118}$ mutation (from the 50B chromosome). This shows that the lethality in the 50B chromosome is due to an unknown lesion between the start of the chromosome and the *cv* locus, and no significant lethality comes from the region that includes *pcm* (Figure 4.3). With all the evidence gathered the chance that 50B or 50C is any different from $pcm^5$ in any region that could affect the expression of *pcm* is extremely low. Therefore, it is most likely 50B and 50C are examples of $pcm^5$ contamination that have picked up a secondary, lethal mutation unrelated to the P-element excision.

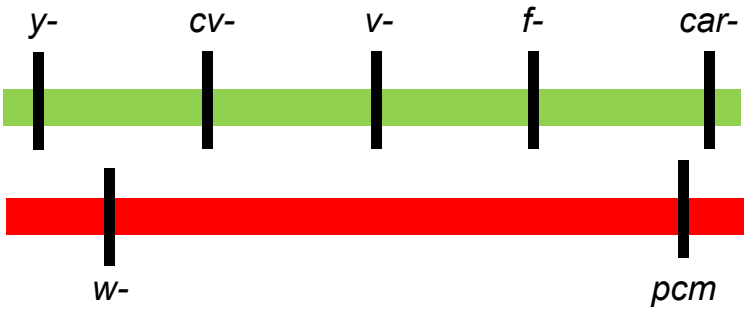



## B. 16A crossover

| Possible marker combinations | No. observed | |
|---|---|---|
| y- w+ cv- v- f- car- | 163 | 53% |
| y+ w- cv- f- | 25 | 8% |
| y+ w- cv+ f- | 79 | 26% |
| y+ w- cv+ f+ | 19 | 6% |
| y- w+ cv+ v+ f+ car+ | 0 | 0% |
| y- w+ cv- v+ f+ car+ | 0 | 0% |
| y- w+ cv- v- f+ car+ | 0 | 0% |
| y- w+ cv- v- f- car+ | 0 | 0% |
| y- w+ cv+ v- f- car- | 1 | <1% |
| y- w+ cv+ v+ f- car- | 8 | 3% |
| y- w+ cv+ v+ f+ car- | 1 | <1% |
| y- w+ cv- v+ f- car- | 4 | 1% |
| y- w+ cv- v+ f+ car- | 5 | 2% |
| y- w+ cv- v- f+ car- | 0 | 0% |
| y+ w- cv- f+ | 0 | 0% |
| y+ w- cv+ v+ f+ car+ | 0 | 0% |
| | 305 | |

**Figure 4.2**: Crossover between 16A and marker chromosomes.
**A.** Approximate location of markers on marker chromosome (green) and 16A chromosome (red).
**B.** Numbers and frequencies of flies with specific marker combinations observed. Grey blocks indicate when the phenotypes of $v^1$ and $car^1$ are not observable, due to the presence of $w^{1118}$. No flies were seen containing the section of 16A chromosome that includes mutant *pcm* and wild-type *car*, indicating the lethality of this chromosome is due to a lesion in this region.

## A. Chromosomes




## B. 50B crossover

| Possible marker combinations | No. observed | |
|---|---|---|
| y- w+ cv- v- f- car- | 231 | 54% |
| y+ w- cv- f- | 0 | 0% |
| y+ w- cv+ f- | 0 | 0% |
| y+ w- cv+ f+ | 0 | 0% |
| y- w+ cv+ v+ f+ car+ | 13 | 3% |
| y- w+ cv- v+ f+ car+ | 53 | 12% |
| y- w+ cv- v- f+ car+ | 82 | 19% |
| y- w+ cv- v- f- car+ | 25 | 6% |
| y- w+ cv+ v- f- car- | 0 | 0% |
| y- w+ cv+ v+ f- car- | 3 | <1% |
| y- w+ cv+ v+ f+ car- | 2 | <1% |
| y- w+ cv- v+ f- car- | 6 | 1% |
| y- w+ cv- v+ f+ car- | 10 | 2% |
| y- w+ cv- v- f+ car- | 2 | <1% |
| y+ w- cv- f+ | 0 | 0% |
| y+ w- cv+ v+ f+ car+ | 0 | 0% |
| | 427 | |

**Figure 4.3**: Crossover between 50B and marker chromosomes.
**A.** Approximate location of markers on marker chromosome (green) and 50B chromosome (red).
**B.** Numbers and frequencies of flies with specific marker combinations observed. Grey blocks indicate when the phenotypes of $v^1$ and $car^1$ are not observable, due to the presence of $w^{1118}$. No flies were seen containing the section of 50B chromosome that includes wild-type *y* and mutant *w*, indicating the lethality of this chromosome is due to a lesion in this region.

## 4.3 Rescue of *pcm*$^{14}$ lethality

A standard way of testing that a phenotype is due to a specific mutation is to show that the phenotype can be rescued by reintroducing a wild-type copy of the mutated gene. This can be achieved in a number of ways in *D. melanogaster* and the two methods employed to attempt to rescue the *pcm*$^{14}$ lethality are described below.

### 4.3.1 Rescue using wild-type *pacman* cDNA

The first set of rescue experiments were attempted using *UAS-pcm* constructs made in the Newbury lab. *UAS* constructs are normally used in combination with insertions that express the yeast transcription factor *GAL4* (Phelps and Brand, 1998, Duffy, 2002). Depending on their location in the genome, *GAL4* constructs express *GAL4* in specific patterns. *GAL4* binds to the *UAS* promoter and drives expression of the gene inserted downstream on the *UAS* construct. The *UAS-pcm* constructs available contain wild-type *pcm* cDNA (exons only) and "nuclease dead" *pcm* cDNA that has had a codon changed in the catalytic domain (in exon 2) to prevent the translated protein degrading RNA. Combined with *GAL4* drivers, these could be used to drive expression of wild-type or nuclease dead *pcm*. However, as *pcm* expression may be required in many tissues to prevent lethality, guessing tissues in which to drive *pcm* expression is unlikely to lead to rescue of the *pcm*$^{14}$ lethality. Instead of specific *pcm* expression, we decided to attempt to rescue *pcm*$^{14}$ using "leaky" expression from the *UAS* constructs. *UAS* constructs are referred to as "leaky" if they are expressed when *GAL4* is not present. Various factors can affect this, such as position in the genome, and the *UAS-pcm* constructs available are thought to each be leaky to varying extents (Grima, unpublished data). Crosses were performed to generate *pcm*$^{14}$*/Y ; UAS-pcm/+* or *pcm*$^{14}$*/Y ; ; UAS-pcm/+* using both wild-type and nuclease dead cDNA, but no constructs were able to rescue the lethality of *pcm*$^{14}$. As leaky expression from these constructs has not previously been shown to be sufficient to rescue lethality from the *pcm* locus, no conclusions can be drawn from this data. It would be unsurprising if low level leaky expression of *pcm* is not a sufficient alternative to endogenous *pcm* expression, especially since *pcm* levels are known to vary during development (Till *et al.*, 1998). As leaky expression of *pcm* cDNA could not rescue *pcm*$^{14}$ lethality, the next approach taken was to attempt rescue using a genomic copy of *pcm*.

### 4.3.2 Rescue of lethality with *T(1;Y)B92*

There are a number of lines listed on FlyBase that should allow the production of males carrying lethal alleles of genes in the *pcm* region. There are two translocations and eight duplications of regions (of varying sizes and including *pcm*) from the *X* to a different chromosome. These would allow a wild-type copy of *pcm* to be present on a chromosome other than *X*, which would compensate for the lack of *pcm* on the *X*. All of the duplications and one of the translocations available were from *X* to *Y*. The other translocation was from *X* to *3*. We decided to test the translocation from *X* to *Y* (*T(1;Y)B92*) as it was the smallest area (cytologically mapped to 18A5--18D1) and the translocation from *X* to *3* (*T(1;3)143-3*, cytologically mapped to 17--20). However, *T(1;3)143-3* did not produce any results to suggest it did actually contain the specified translocation, so this section concentrates on rescue of lethality by *T(1;Y)B92* only. The ability of *T(1;Y)B92* (*In(1)FM7/T(1;Y)B92*, $y^1\ y^+\ B^S$, stock BL101110) to rescue the lethality of a number of mutations/deficiencies was tested.

#### 4.3.2.1 Mutations tested

$pcm^{14}$ was tested along with two more lethal mutations in the same region (stocks BL11837 and BL11896). These unrelated lethal mutations were selected as positive controls (however, rescue of these lethals by *T(1;Y)B92* has not previously been demonstrated). BL11837 ($w^{67c23}$, *P{lacW}l(1)G0147*$^{G0147}$*/FM7c*) contains a P-element insertion into the gene *vav* which results in recessive lethality. *vav* is located 212,601bp away from *pcm* in the area covered by *T(1;Y)B92*. BL11896 ($w^{67c23}$, *P{lacW}l(1)G0156*$^{G0156}$*/FM7c*) contains a P-element insertion in the gene *l(1)G0156* which lies 27,164bp away from the other end of *pcm*, within the region covered by *T(1;Y)B92*.

Two lethal mutations, 5B and 50B, created in the P-element excision were also tested as negative controls. 5B is embryonic lethal and was found to not contain a *pcm* mutation in the previous chapter. The actual location of the lethality in 5B is not known, so there was a small chance 5B could be rescued by *T(1;Y)B92*. The lethality for 50B however was known to be due to a lesion in an area not covered by *T(1;Y)B92*, as shown by the crossover experiments at the beginning of this chapter.

#### 4.3.2.2 Deficiencies tested

Three lethal deficiencies were also tested: *Df(1)JA27* , maximum size 501,546bp (18A5--18D), minimum size 396,337bp (18A5--18D1); *Df(1)Exel7468,* 245,148bp (18B7--18C8) and *Df(1)ED7452*, 17,963bp (18C7--18C8). These deficiencies all include the *pcm* locus (18C7) and are all embryonic lethal. *Df(1)Exel7468* and *Df(1)ED7452* are both entirely covered by *T(1;Y)B92*. *Df(1)JA27* may not be entirely covered as its breakpoints are not known precisely. Further information on these deficiencies is presented later in this chapter, including creation of *Df(1)ED7452*.

#### 4.3.2.3 Rescued lesions

Figure 4.4 shows the frequency of rescued flies (*lethal/T(1;Y)B92*) for each line tested as a percentage of total offspring. As expected, no rescued flies were seen for *50B/T(1;Y)B92* or *5B/T(1;Y)B92* as the translocation does not cover the region that causes the lethality in these files. *Df(1)JA27* is not rescued by *T(1;Y)B92*. This may be due to *T(1;Y)B92* being smaller than the deletion in *Df(1)JA27*, as the exact sizes of each are unknown. *Df(1)Exel7468* is also not rescued by *T(1;Y)B92*, which is unexpected, as *T(1;Y)B92* covers the whole region deleted in *Df(1)Excel7468*. *Df(1)ED7452* is rescued. The insertion in *vav* (*P{lacW}l(1)G0147*$^{G0147}$) is rescued, but the insertion in *l(1)G0156* (*P{lacW}l(1)G0156*$^{G0156}$) is not. *pcm*$^{14}$ is rescued to the expected frequency (25%) by *T(1;Y)B92*.

These results show that *T(1;Y)B92* is capable of rescuing the lethality of *pcm*$^{14}$, which allows for the creation of *pcm*$^{14}$ males, and theoretically, a stock with the *T(1;Y)B92* chromosome instead of wild-type *Y* so *pcm*$^{14}$ males are quickly available when needed. However, the failure to rescue the lethality of the insertion into *l(1)G0156* or the deficiency *Df(1)Exel7468* indicates the translocated region may not act exactly as the wild-type region on *X*.

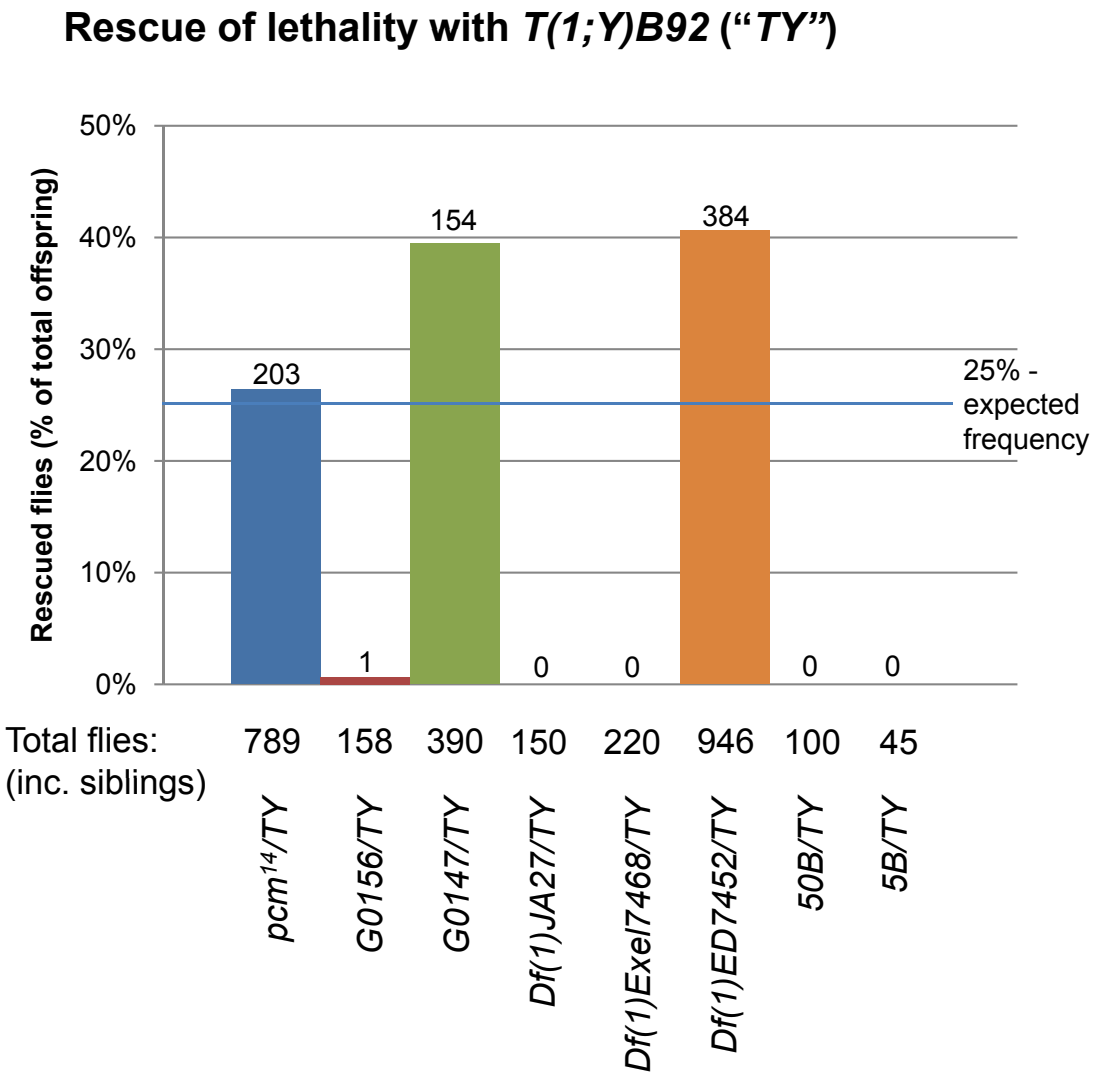



**Figure 4.4**: Lethal lesions rescued by *T(1;Y)B92* (*TY*).
The lethality of *pcm*$^{14}$, the insertion *G0147* and the deficiency *Df(1)ED7452* can be rescued by *T(1;Y)B92*. The expected frequency is 25% as there are four theoretically viable phenotypes from the cross *In(1)FM7/T(1;Y)B92* x *lethal/FM7i*.
One lethal insertion, *G0156*, and the deficiency *Df(1)Exel7468* are not rescued by *T(1;Y)B92*, which is unexpected as *T(1;Y)B92* should carry replacement copies of the *l(1)G0156* gene, and all 30 genes deleted in *Df(1)Exel7468*.
*Df(1)JA27* is also not rescued by *T(1;Y)B92*. It is possible that *Df(1)JA27* is larger than *T(1;Y)B92*, which may explain the lack of rescue, as the size of each has not been defined to the exact base pair.
The lines 50B and 5B contain lethal lesions in an area not covered by *T(1;Y)B92* and were used as negative controls.
Absolute numbers of rescued flies are given above the bars.

#### 4.3.2.4 Fertility of *pcm$^{14}$/T(1;Y)B92* files

It was quickly noticed that *pcm$^{14}$/T(1;Y)B92* males appeared to be infertile. Males collected over a week (up to 5 days old when used) failed to produce any offspring when they were used for further crosses. It was also noticed that these flies were relatively short-lived, expiring by roughly 10 days after eclosion. To investigate this further, 1 day old *pcm$^{14}$/T(1;Y)B92* and *Df(1)ED7452/T(1;Y)B92* males were crossed back to their respective stock females (*pcm$^{14}$/FM7i* or *Df(1)ED7452/FM7i*). The use of 1 day old males allowed some offspring to be produced (Figure 4.5), however, far fewer than expected were produced, and the males again died off quickly. No apparent *pcm$^{14}$* or *Df(1)ED7452* homozygotes were produced in the back crosses, indicating that the rescue of lethality by *T(1;Y)B92* was genuine.

The reason for the extremely poor fertility in these files is not immediately obvious. The lack of fertility is not an inherent issue of *T(1;Y)B92*, as the stock containing *T(1;Y)B92* would struggle to survive if that were the case. It may be that *pcm* expression from *T(1;Y)B92* is not at wild-type levels due to it being located on *Y* instead of *X*, or they may be some mutation to *pcm* on *T(1;Y)B92* that affects fertility but is mild enough to rescue the lethality of *pcm$^{14}$*.

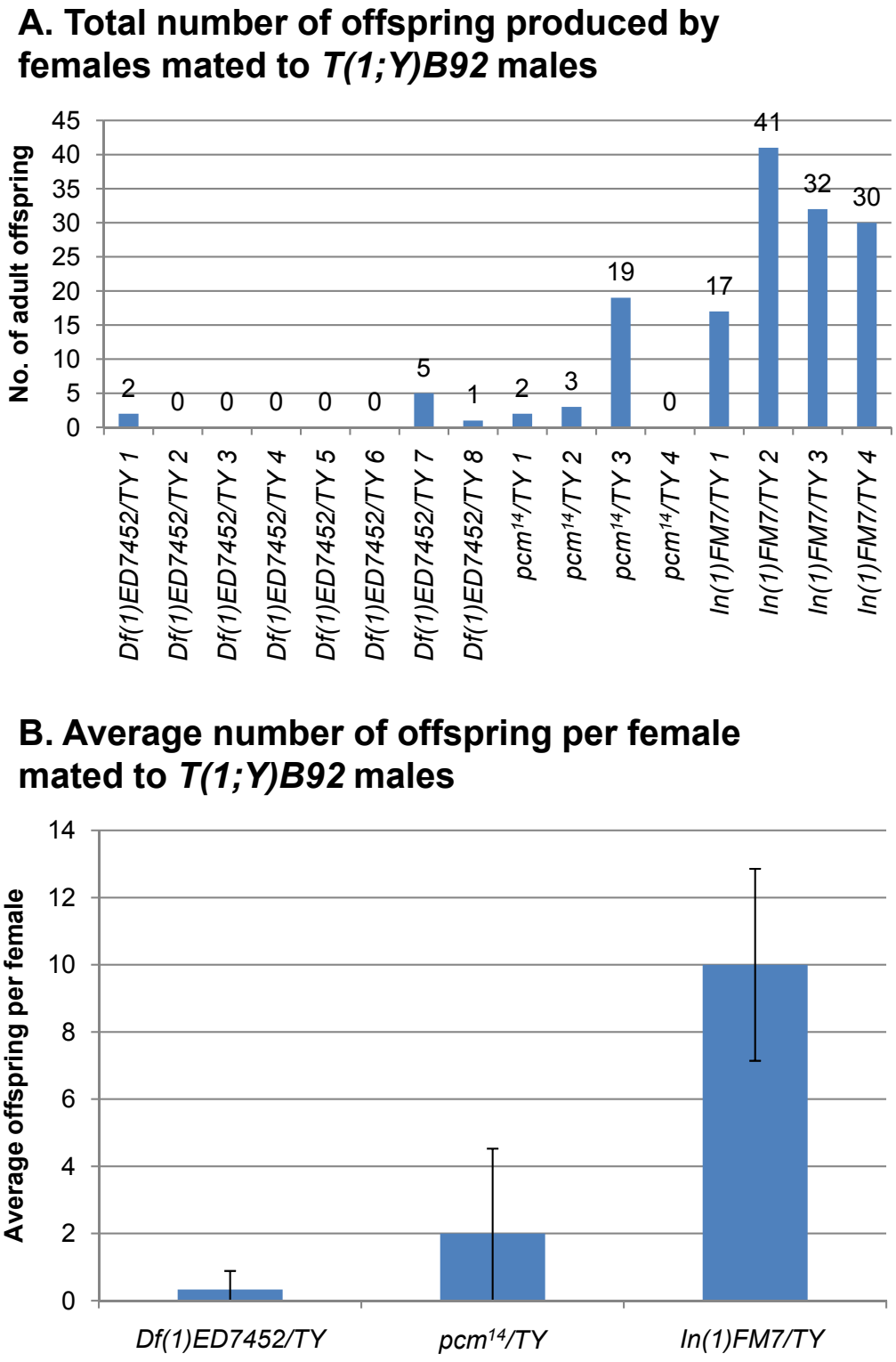



**Figure 4.5**: Fertility of *T(1;Y)B92* (*TY*) over *In(1)FM7*, *Df(1)ED7452* or *pcm*$^{14}$.
Males were crossed back to females from their stock genotypes:
*T(1;Y)B92/Df(1)ED7452* (♂) x *Df(1)ED7452/FM7i* (♀),
*T(1;Y)B92/pcm*$^{14}$ (♂) x *pcm*$^{14}$*/FM7i*) (♀),
*T(1;Y)B92/In(1)FM7* (♂) x *In(1)FM7/In(1)FM7* (♀).
**A**. The number of offspring produced from three females in the above crosses. Numbers for *Df(1)ED7452* and *pcm*$^{14}$ are total offspring and numbers for *In(1)FM7* are total produced after 17 days (eclosion began on day 10).
**B**. Average number of offspring per female fly for each male genotype. Error bars show standard deviation.

## 4.4 *pcm*$^{5}$ complementation with deficiencies to test if *pcm*$^{14}$ is a null allele

A vital test for any new mutation is to test for complementation between the chromosome containing it and a deficiency. If a mutant is lethal as a homozygote, it should still be lethal when placed over a deficiency that includes the gene the mutation is in. If the flies produced are viable, it strongly suggests the mutation in question is not the reason for the lethality. Reciprocally, if a deficiency is used that does not cover the mutated gene, successful complementation should be observed and viable heterozygotes should be produced, just as would be produced if the mutated chromosome was crossed to a wild-type chromosome.

We already know that *pcm*$^{5}$ and *Df(1)JA27* do not complement each other at 19°C, and *pcm*$^{5}$*/Df(1)JA27* flies die during eclosion (Grima, 2002). This is stronger than the phenotype seen for *pcm*$^{5}$/*pcm*$^{5}$ homozygotes at 19°C, which shows the *pcm*$^{5}$ allele is hypomorphic. If *pcm*$^{14}$ is a null allele, *pcm*$^{5}$*/pcm*$^{14}$ flies should act the same way as *pcm*$^{5}$*/Df(1)JA27* flies at 19°C. This was tested using *pcm*$^{14}$, *Df(1)JA27* and *Df(1)ED7452* at 19°C. *pcm*$^{5}$ combined with *pcm*$^{14}$ or either of the deficiencies resulted in the same frequencies of flies stuck in pupal cases at 19°C and viable flies at 25°C (Figure 4.6). This strongly indicates that the *pcm*$^{14}$ allele is a null allele as it is acting in the same was as *Df(1)JA27* and *Df(1)ED7452*, which are known to be null for *pcm*. *pcm*$^{5}$*/pcm*$^{14}$ acting the same as *pcm*$^{5}$*/deficiency* also shows that the phenotype must be due solely to the *pcm* locus.

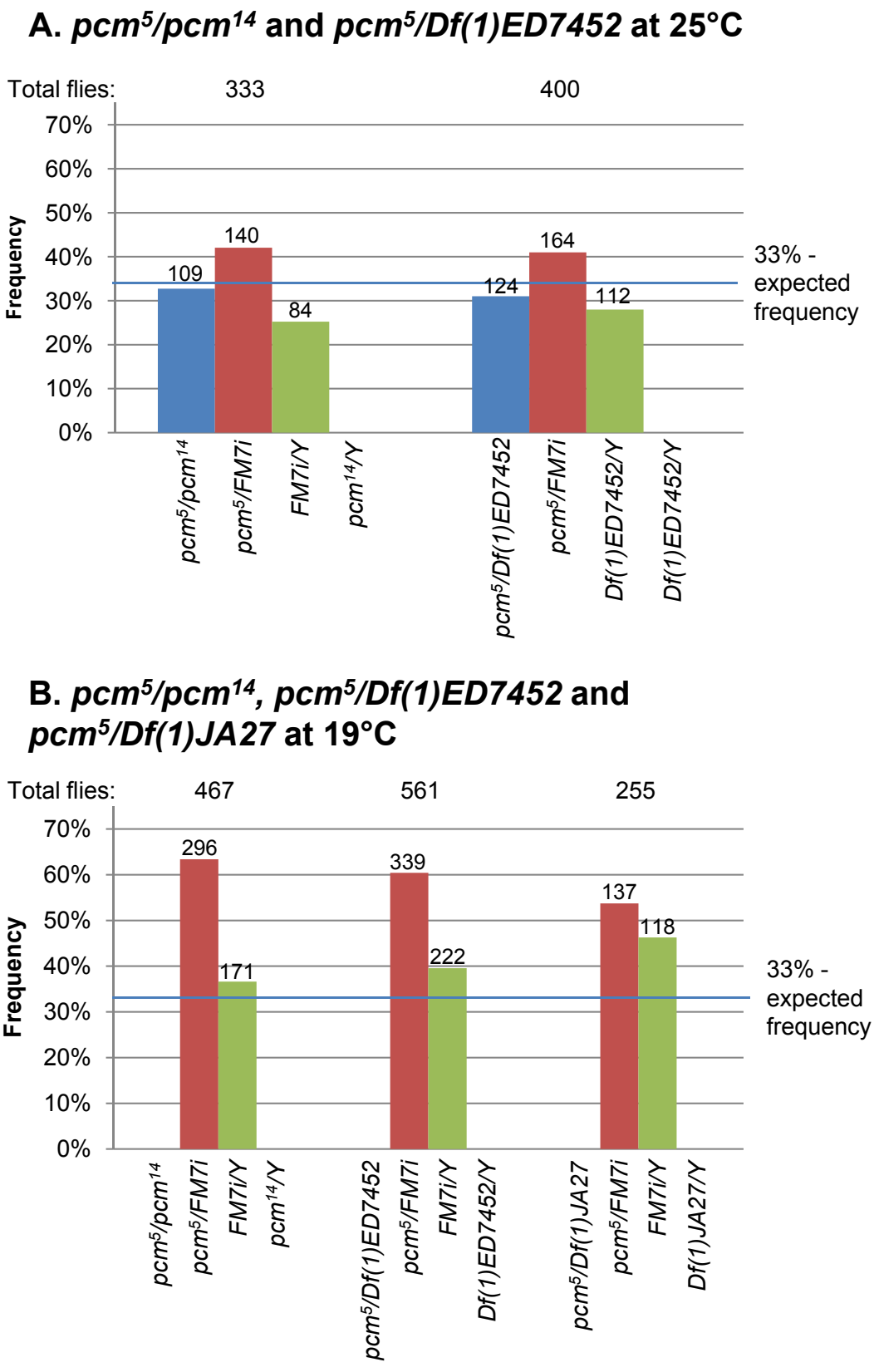



**Figure 4.6**: *pcm*$^{5}$ crossed to *pcm*$^{14}$, *Df(1)ED7452* and *Df(1)JA27*.
**A**. Frequencies of offspring produced when *pcm*$^{5}$*/Y* is crossed with *pcm*$^{14}$*/FM7i* or *Df(1)ED7452/FM7i* at 25°C. Numbers of offspring produced are not significantly different for *pcm*$^{14}$ and *Df(1)ED7452*. The expected frequency for each genotype is 33% as there are three viable genotypes.
**B**. Frequencies of offspring produced when *pcm*$^{5}$*/Y* is crossed with *pcm*$^{14}$*/FM7i*, *Df(1)ED7452/FM7i* or *Df(1)JA27/FM7i* at 19°C. Numbers of offspring produced are similar for *pcm*$^{14}$, *Df(1)ED7452* and *Df(1)JA27*. The expected frequency for each genotype is 33% as there are three theoretically viable genotypes.
Absolute numbers of flies are shown above each bar.

## 4.5 Phenotypes of *pacman* mutations

Many of the phenotypes in *pcm* mutants manifest either in the wing discs, or in structures that grow from the wing discs (described below), which suggests that the wing discs are the tissue most sensitive to mutations of *pcm*. Why this should be is not immediately obvious, but it shows that sensitivity is not equal across tissues. While the descriptions of phenotypes below focus mainly on wing imaginal discs, it should be remembered that some *pcm* phenotypes do manifest in other structures (such as the testes) and this may reflect common mechanisms of post-transcriptional gene regulation in diverse tissues.

### 4.5.1 Phenotypes of non-lethal *pacman* mutants

Each of the non-lethal *pcm* mutants exhibit a number of phenotypes at differing frequencies and these have been described previously (Grima, 2002, Grima *et al.*, 2008, Lin *et al.*, 2008, Sullivan, 2008, Zabolotskaya *et al.*, 2008). One of the most penetrant phenotypes is dull/crumpled wings in *pcm* hemi- and homozygotes. Figure 4.7 shows the frequency of the dull wing phenotype in $pcm^{3}$, and $pcm^{5}$ recorded by Dr. Dominic Grima after creation of these mutants. The penetrance increases at lower temperatures, indicating these *pcm* alleles are temperature sensitive. Other phenotypes previously observed for *pcm* alleles, such as cleft thorax, bent legs and bristle defects are shown in Figure 4.8.

### 4.5.2 Phenotypes of the lethal allele $pcm^{14}$

The previous chapter showed that $pcm^{14}$ hemi/homozygote flies die during pupation. On closer examination, it was clear that most flies die shortly after pupariation, as adult structures are rarely seen for $pcm^{14}/Y$ or $pcm^{14}/pcm^{14}$ larvae (Figure 4.9, panel A), but apolysis appears to have occurred in most cases. In a minority of cases, pharate adults form but do not eclose (Figure 4.9, panel B). Dissection of these flies from their pupal cases revealed no obvious abnormalities (Figure 4.9, panel C). This phenotype is not temperature sensitive.

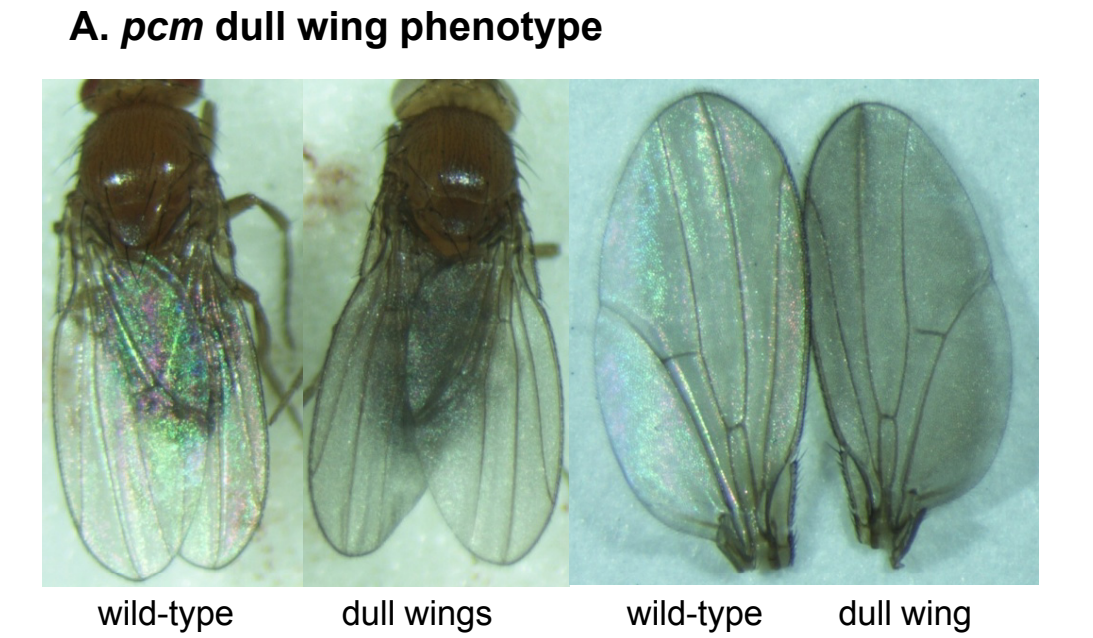



**Figure 4.7**: The dull wing phenotype seen in *pcm* mutants.
**A**. Left: a wild-type fly and a dull winged *pcm*$^{5}$ homozygote under the same light conditions. The wings on the mutant fly are not iridescent. Right: dissected wild-type and *pcm*$^{5}$ homozygote wings.
**B**. Penetrance of the dull wing phenotype in different mutants at four temperatures, as reported after creation of the mutants (Grima, 2002).

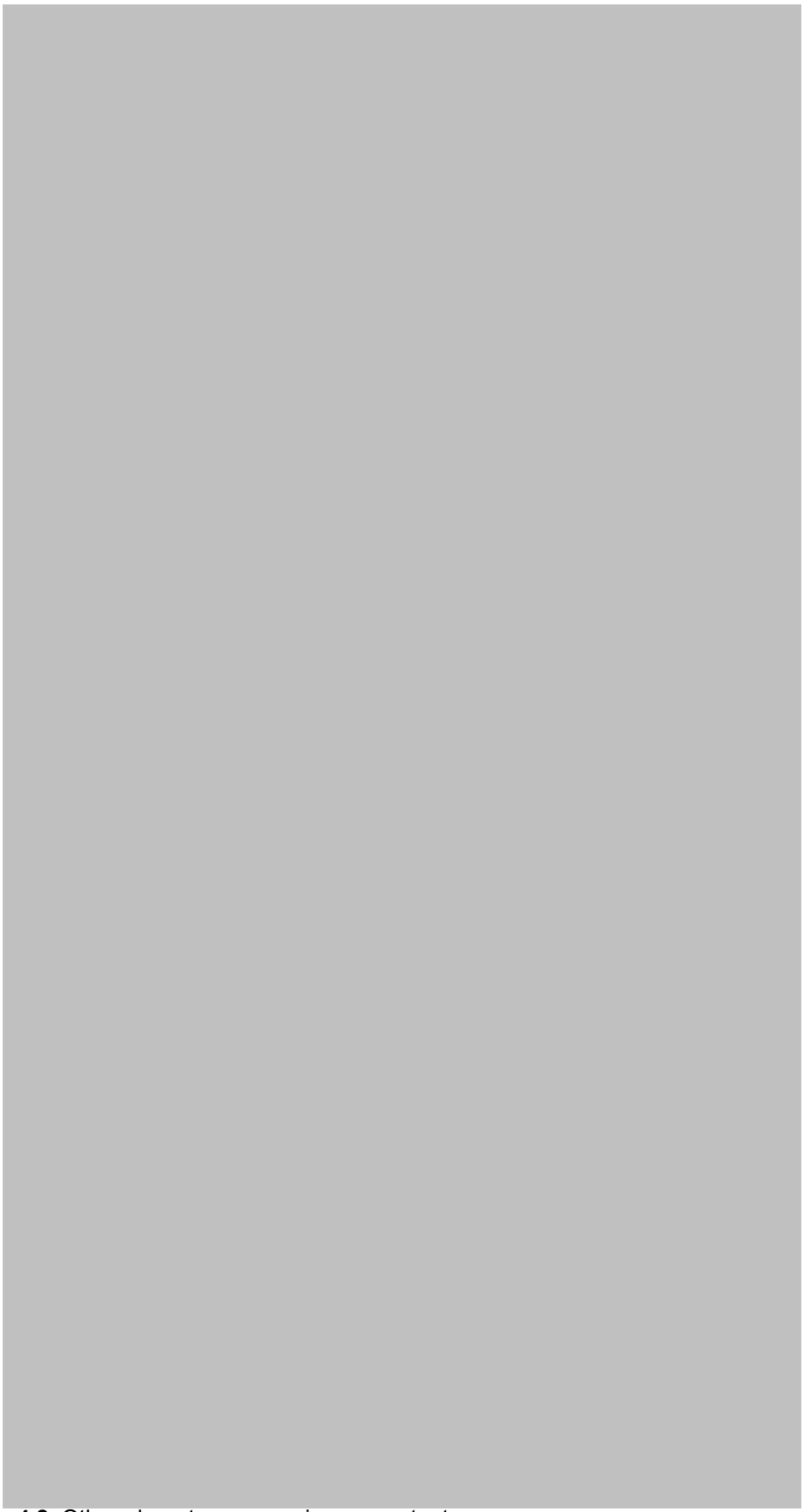


**Figure 4.8**: Other phenotypes seen in *pcm* mutants.
**A** and **B**. Extra bend (white arrow) in the tibia of the third leg in $pcm^{3}$ homozygote (right) compared to wild-type (left).
**C** and **D**. Cleft thorax phenotype of a $pcm^{5}$ homozygote raised at 29°C (right) compared to wild-type (left).
**E** and **F**. Examples of bristle defects seen in *pcm* mutants. Missing head bristles (red arrows), bent bristle (blue arrow) and ectopic macrochaetae (white arrow).
Images **A**-**D** from Grima, 2002.

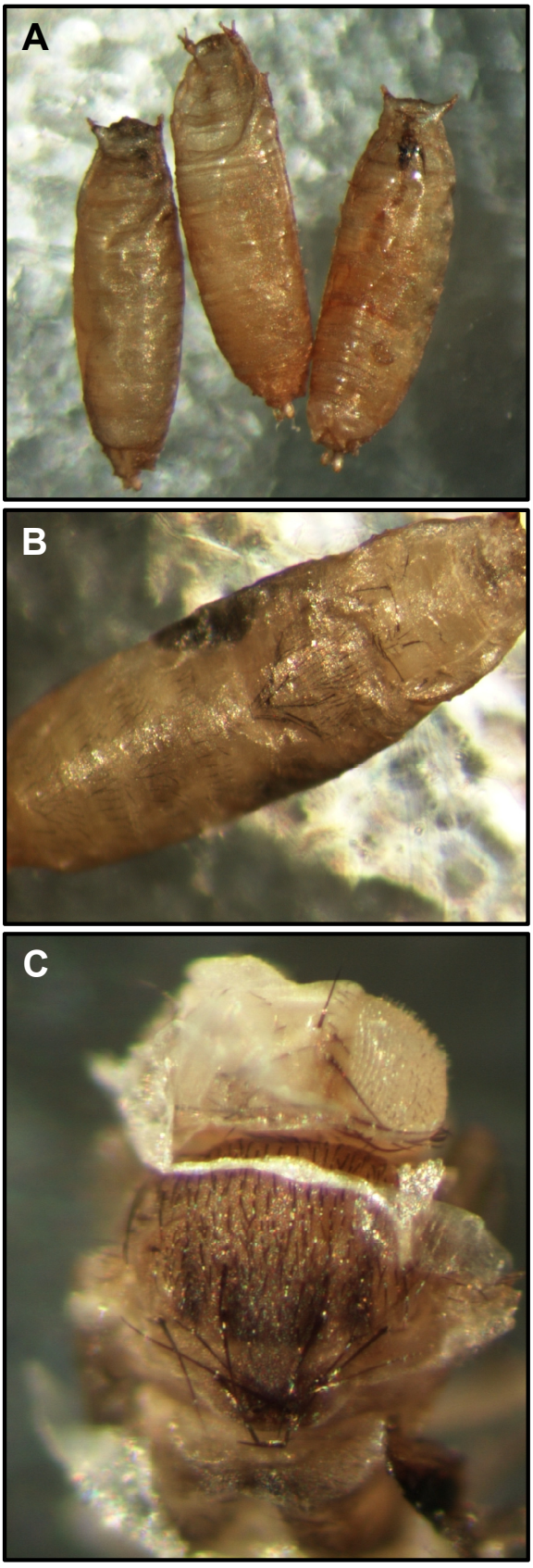



**Figure 4.9**: Lethality in *pcm*$^{14}$ hemi-/homozygotes.
**A**. The overwhelming majority of *pcm*$^{14}$ hemi/homozygotes die early in pupation before any recognisable adult structures form.
**B**. Very rarely, pupae survive until the pharate adult stage, but flies never begin to eclose.
**C**. A dissected *pcm*$^{14}$ pharate adult. No gross morphological defects are obvious in pharate adults excised from their pupal cases.

### 4.5.3 Wing imaginal disc phenotypes in *pcm*$^{5}$ and *pcm*$^{14}$ L3 larvae

Imaginal discs are parts of insect larvae that will go on to form a distinct part of the adult insect. *Drosophila* L3 larvae contain numerous pairs of discs, such as three pairs of leg discs, a pair of haltere discs and a pair of wing discs. Each of these discs undergoes rapid growth and development during metamorphosis to form their defined adult structures. Adult *pcm* mutants most frequently show phenotypes in structures derived from the wing imaginal discs, which form the wings and the thorax of the adult fly. The two most common phenotypes, dull wings and disrupted thorax bristles, are apparent in structures derived from wing imaginal discs, as is the less common cleft thorax phenotype. To determine whether these phenotypes are due to effects during metamorphosis, or earlier, I excised wing imaginal discs from *pcm*$^{5}$ and *pcm*$^{14}$ L3 larvae. Both *pcm*$^{5}$ and *pcm*$^{14}$ discs were measured to determine any size difference from wild-type. Changes to gene and microRNA expression levels were also analysed in *pcm*$^{5}$ discs to search for Pacman mRNA/miRNA targets, which are the subjects of Chapters 5 and 6.

When compared to wild-type (or *pcm*$^{5}$), *pcm*$^{14}$ wing imaginal discs are markedly smaller. Upon initial dissection, *pcm*$^{5}$ wing discs were not subjectively different in appearance to wild-type wing discs, but accurate measurements of their size revealed they were about 82% the size of wild-type (50E) wing discs, $p<0.001$. Measurement of *pcm*$^{14}$ wing discs showed they were 43% the size of wild-type, $p<0.001$ (Figure 4.10). Pictures of excised discs were taken at the same magnification and the areas were measured in arbitrary units using ImageJ. It should be noted that the mean size of *pcm*$^{14}$ wing discs recorded may be an overestimation. Wing discs in *pcm*$^{14}$ larvae are considerably more difficult to identify and excise than those in wild-type or *pcm*$^{5}$ larvae, due to their small size. This may have created a bias towards the recovery of larger examples.

## A. Wild-type and *pcm*$^{14}$ wing imaginal discs

wild-type (100%) *pcm*$^{14}$ (47%)

## B. Wild-type, *pcm*$^{5}$ and *pcm*$^{14}$ wing imaginal discs sizes

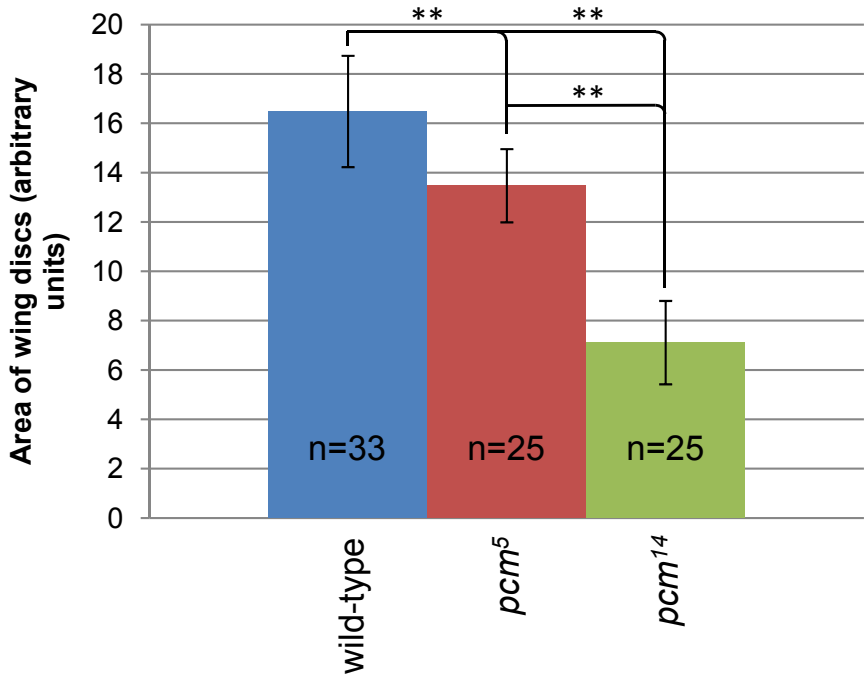



**Figure 4.10**: L3 wing imaginal discs are smaller in *pcm* mutants.
**A**. Wild-type (50E) wing disc compared to a *pcm*$^{14}$ wing disc. The *pcm*$^{14}$ wing disc shown has just under half the area of the wild-type disc.
**B**. Size comparison between wild-type (50E), *pcm*$^{5}$ and *pcm*$^{14}$ wing discs. On average, *pcm*$^{5}$ wing discs are 82% the size of wild-type and *pcm*$^{14}$ wing discs are 43% the size of wild-type. *pcm*$^{14}$ wing discs are 53% the size of *pcm*$^{5}$ wing discs. ** = p-value <0.001.

### 4.5.4 Inter-allelic phenotypes – *pcm*$^{14}$ with other *pcm* mutants

A number of experiments have been performed to create flies with *pcm*$^{14}$ over another *pcm* allele, at both 19°C and 25°C. The objective of these was to determine a scale of severity for *pcm* alleles and to further test if *pcm*$^{14}$ is a null. All followed the same basic scheme (Figure 4.11). The frequency of offspring from each cross was recorded and the phenotypes were recorded from random samples of *pcm*$^{14}$/*pcm*$^{x}$ females (where *pcm*$^{x}$ = *pcm*$^{3}$, *pcm*$^{5}$ *or pcm*$^{6}$). Figure 4.12, panel A shows the frequency of *pcm*$^{14}$/*pcm*$^{x}$ files produced in each cross as a percentage of the total offspring. Each cross can theoretically produce four genotypes of fly, *pcm*$^{14}$/*Y*, *FM7i/Y pcm*$^{x}$/*FM7i* and *pcm*$^{14}$/*pcm*$^{x}$, however, no *pcm*$^{14}$/*Y* adults eclose as *pcm*$^{14}$ is lethal. Consequently, the expected frequency of each genotype of adult flies seen is 33%. *pcm*$^{14}$/*pcm*$^{3}$ and *pcm*$^{14}$/*pcm*$^{6}$ flies occur at roughly the expected frequency at 19°C and 25°C. *pcm*$^{14}$/*pcm*$^{5}$ flies occur at the expected frequency at 25°C, but no adults are seen at 19°C. Adult *pcm*$^{14}$/*pcm*$^{5}$ flies do occur, but are unable to free themselves from the pupal case during eclosion (Figure 4.13). A minority do not begin eclosion, but most only get their head and front legs out of the pupal case before becoming stuck. Some manage to get their whole body out of the pupal case, but remain attached to the inside of it by one or both of their rear legs. The stuck flies survive for some time before dying, most likely of starvation. This lethality has been previously observed to occur at 19°C with *pcm*$^{5}$/*Df(1)JA27* (a large deficiency that includes the *pcm* locus).

Figure 4.12 also shows the frequency of phenotypes observed in adult *pcm*$^{14}$/*pcm*$^{x}$ flies at 25°C (panel B) and the phenotype frequencies for the *pcm*$^{14}$/*pcm*$^{x}$ flies that survive at 19°C (panel C). Dull wings are the most common phenotype, seen in all *pcm*$^{14}$/*pcm*$^{x}$ combinations at both temperatures. 54% of *pcm*$^{14}$/*pcm*$^{5}$ flies have dull wings at 25°C, as do 92% of *pcm*$^{14}$/*pcm*$^{3}$ flies. This increases to 100% of *pcm*$^{14}$/*pcm*$^{3}$ flies at 19°C. The only other phenotype seen at a reasonably high level was extra macrochaetae in 34% of *pcm*$^{14}$/*pcm*$^{5}$ flies at 25°C. Only phenotypes seen are shown in Figure 4.12 panels B and C. Bent legs, cleft thorax and missing macrochaetae were also searched for, but none were seen in any *pcm*$^{14}$/*pcm*$^{x}$ combination at either temperature. The *pcm*$^{5}$/*FM7i* flies in panel B were siblings of the *pcm*$^{5}$/*pcm*$^{14}$ flies in the same panel and were examined as controls. These results are consistent with *pcm*$^{14}$ being a null allele.

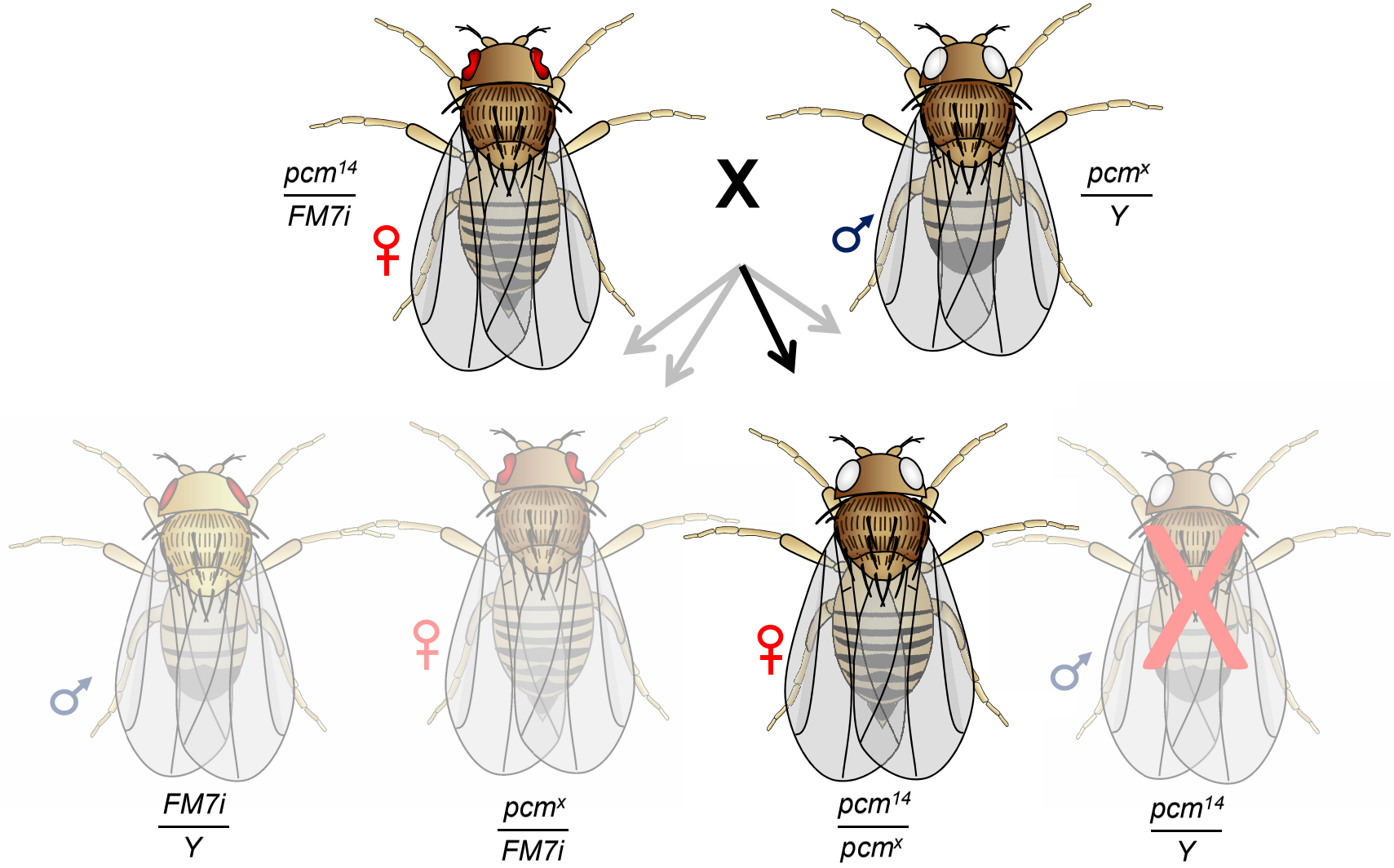



**Figure 4.11**: General scheme used to cross *pcm*$^{14}$ with other *pcm* alleles (*"pcm*$^{x}$*"*).

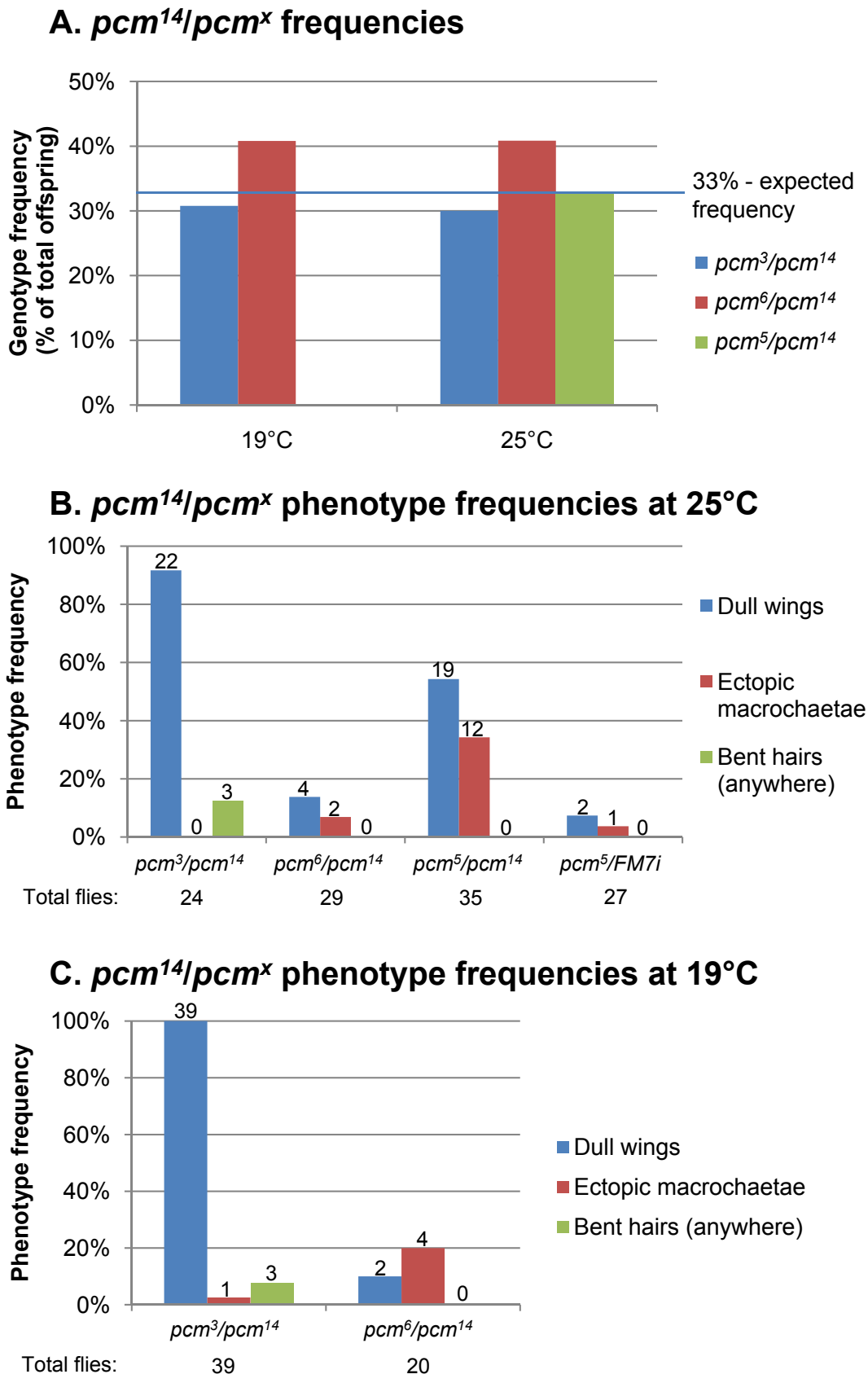



**Figure 4.12**: Interallelic cross frequencies.
**A**. Number of *pcm*$^{14}$/*pcm*$^{3,6\ or\ 5}$ (*pcm*$^{14}$/*pcm*$^{x}$) flies produced at 19°C and 25°C, as percentage of total offspring. As three genotypes can occur, the expected frequency of *pcm*$^{14}$/*pcm*$^{x}$ flies is 33%
**B** and **C**. Frequencies of phenotypes observed in *pcm*$^{14}$/*pcm*$^{x}$ files at 25°C and 19°C, as percentage of total. Absolute numbers of flies exhibiting each phenotype are shown above bars. The frequency of each phenotype in *pcm*$^{5}$/*FM7i* at 25°C was also checked as a control. *pcm*$^{14}$/*pcm*$^{5}$ flies did not survive at 19°C. No cleft thorax or bent leg phenotypes that have previously been reported for *pcm* mutants were observed.



**Figure 4.13**: Phenotype of *pcm*$^{14}$/*pcm*$^{5}$ flies at 19°C.
**A**. Most flies became stuck in their pupal cases during eclosion.
**B**. Flies from **A** after dissection from their pupal cases. Red arrows show material on rear legs that was also attached to the pupal case. Alive flies dissected from their pupal cases were not able to survive.

### 4.5.5 Scale of phenotype severity

Figure 4.14 shows the scale of severity for *pacman* phenotypes. The mildest phenotypes are seen in flies hemi- or homozygous for one of the non-lethal hypomorphic alleles, such as $pcm^{5}$. The most common of these phenotypes is dull/crumpled wings, which is more penetrant at lower temperatures. Lethality is seen in flies heterozygous for $pcm^{5}$ and $pcm^{14}$ or a deficiency at 19°C as the flies are unable to fully eclose from their pupae. The strongest phenotype is seen when no functional copy of *pcm* is present, such as in $pcm^{14}$ homozygotes. These flies survive as L3 larvae and begin pupation, but no recognisable adult structures normally form. $pcm^{14}$*/deficiency* should act the same as $pcm^{14}/pcm^{14}$, but this has not yet been confirmed. *deficiency/deficiency* flies (e.g. *Df(1)ED7452/Df(1)ED7452*) are embryonic lethal, which must be due to one or more of the other genes affected.

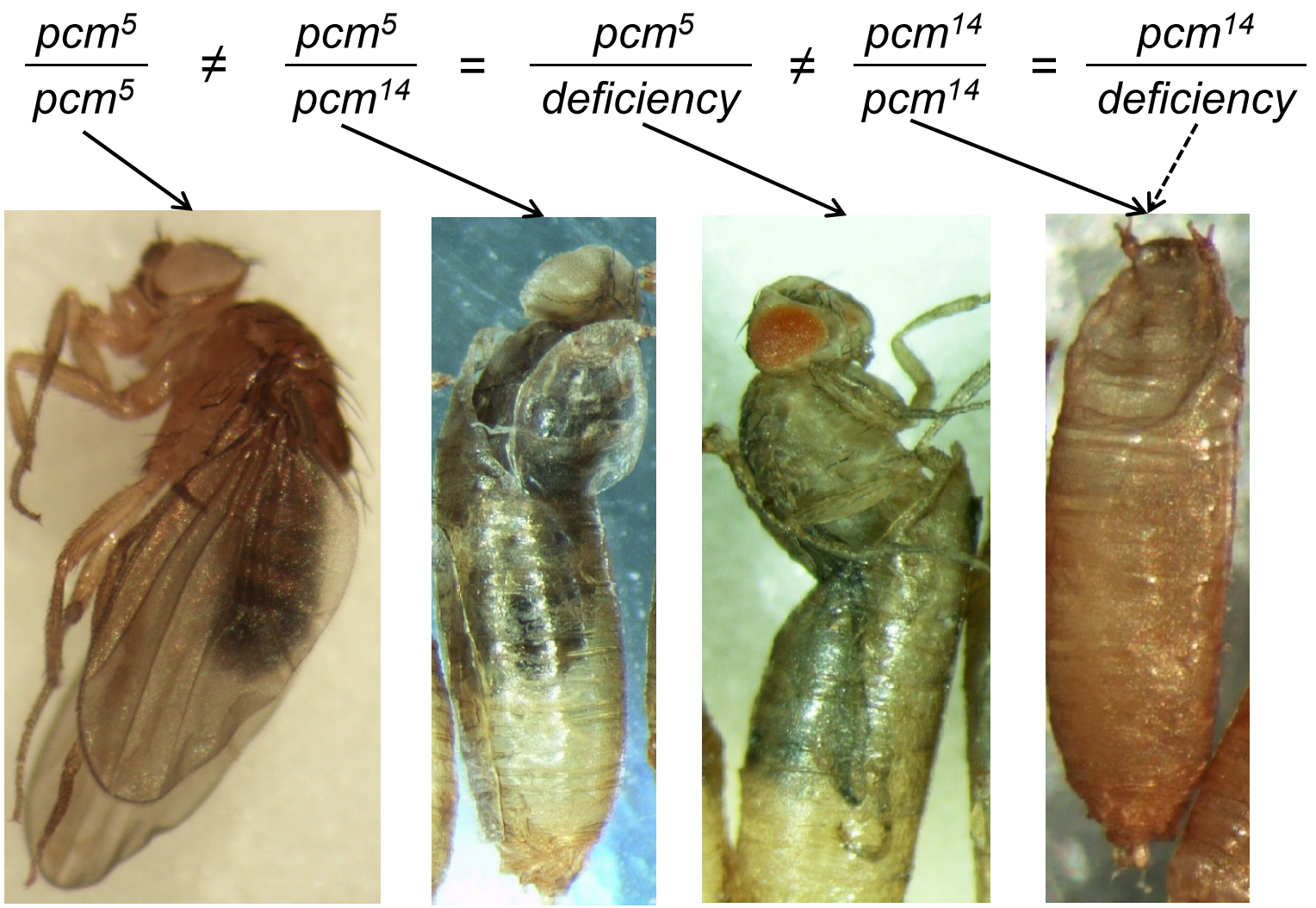



**Figure 4.14**: Scale of severity of *pcm* phenotypes.
The non-lethal *pcm* alleles, such as $pcm^5$ (shown) display a number of phenotypes in adult flies, such as dull wings and reduced fertility, but all are viable and can reproduce. Removal of one copy of $pcm^5$ causes lethality at 19°C, when flies fail to fully eclose, which shows $pcm^5$ is hypomorphic. $pcm^5$ acts the same over a deficiency or $pcm^{14}$, which shows $pcm^{14}$ is a null allele. The earliest lethality is seen in $pcm^{14}$ homozygotes when larvae pupate but do not form adult structures. Although untested, the same phenotype should be observed for $pcm^{14}$/*deficiency* flies.

## 4.6 Measuring Pacman activity in *pacman* mutants

The penetrance of and severity of $pcm^5$ phenotypes increase when one copy of *the* $pcm^5$ allele is removed (e.g. in $pcm^5$*/deficiency* females) and at lower temperatures. These are indicative of a hypomorphic mutation, where the protein produced from the allele is still partially performing the function of the wild-type protein. The evidence presented for $pcm^{14}$ suggests it is a null (amorphic) mutation. The $pcm^{14}$ phenotype does not worsen at lower temperatures. Null mutations are not more detrimental when put over a deficiency (e.g. $pcm^{14}$*/deficiency*), but this has not yet been directly tested with $pcm^{14}$, due to the sex linked lethality. It is impossible to produce the $pcm^{14}$*/Y* or *deficiency/Y* males that would be required for a cross to produce $pcm^{14}$*/deficiency* offspring. Unfortunately, this problem is not fully overcome by using an *X* to *Y* translocation to rescue the lethality of $pcm^{14}$ or the deficiency, as males produced are almost infertile. However, we were able to observe the frequencies of offspring produced when $pcm^5$ is crossed to $pcm^{14}$ or a deficiency ($pcm^5$*/*$pcm^{14}$ or $pcm^5$*/deficiency*), and they are the same in both conditions, with no $pcm^5$*/*$pcm^{14}$ or $pcm^5$*/deficiency* files able to fully eclose at 19°C.

The next step to further the genetic evidence that $pcm^{14}$ is a null would be to use males rescued with *T(1;Y)B92* to produce $pcm^{14}$*/deficiency* flies. However, this would likely provide little information above what is already known and would require considerably more effort than a normal cross, due to the complication of poor fertility. A more efficient approach is to assay the $pcm^{14}$ and other alleles molecularly (e.g. by Western blotting or RT-PCR), as this will provide further evidence for or against $pcm^{14}$ being a null, and will allow us to determine the nature and magnitude of the molecular effects that are behind the observed phenotypes.

### 4.6.1 Pacman protein expression in *pacman* mutants

The levels of Pcm protein in $pcm^5$, $pcm^3$ and $pcm^6$ mutants have been shown previously (Grima, 2002, Zabolotskaya *et al.*, 2008). On a western blot, the $pcm^3$ and $pcm^6$ alleles produce much less intense bands for Pcm than wild-type, and no band for $pcm^5$ is obvious (Figure 4.15, panel A). This fits with the observation that $pcm^5$ is the strongest non-lethal allele, but it should be noted that some protein must be produced for $pcm^5$ as the allele acts as a hypomorph and not a null. Unfortunately the level of protein expression from the $pcm^{14}$ allele cannot be tested by western blotting, as the only good antibody for Pcm binds within the region of protein encoded for by part of the deleted region in $pcm^{14}$.

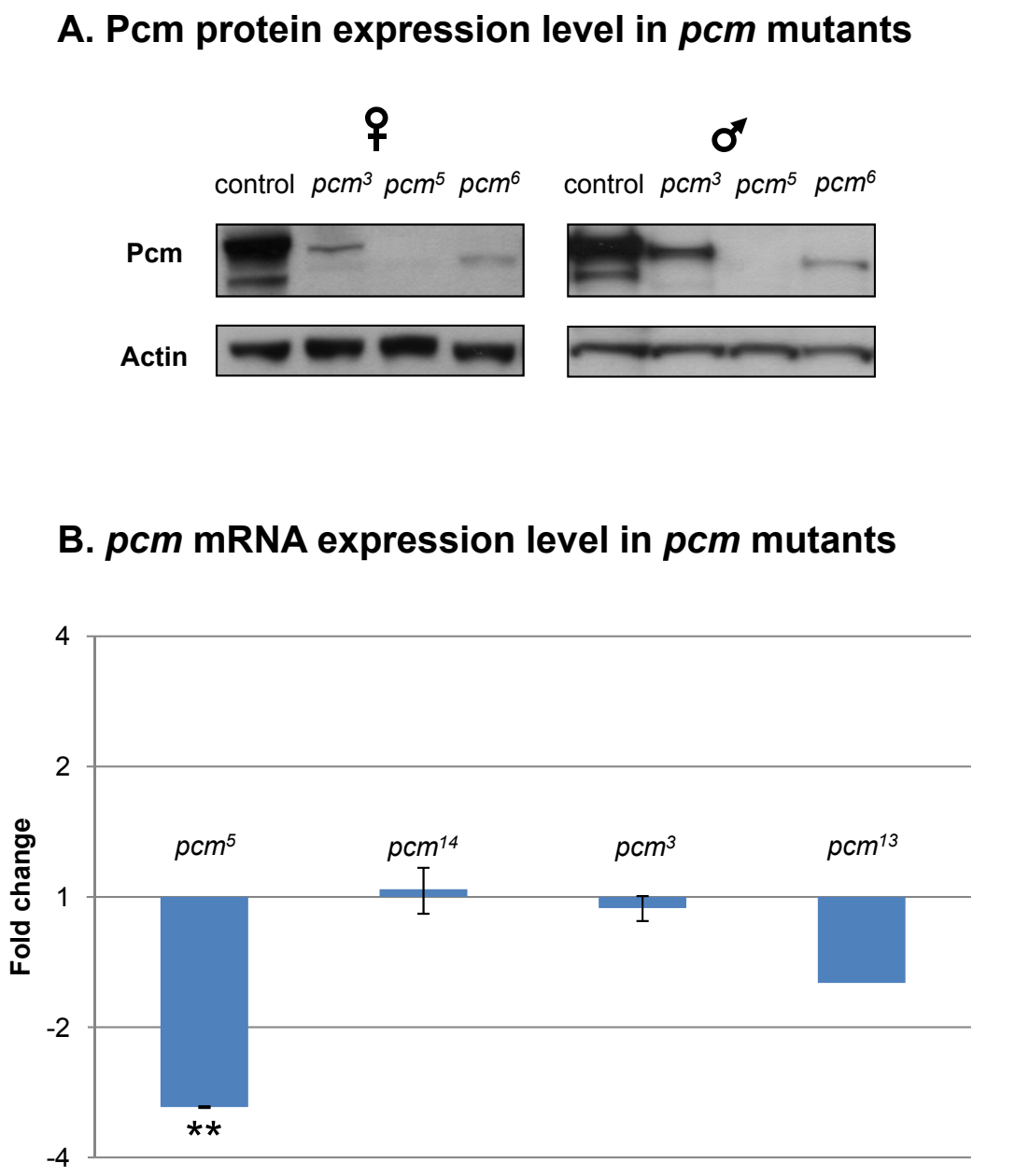



**Figure 4.15**: Expression levels of Pacman protein and *pacman* mRNA in *pcm* mutants.
**A**. Western blot from Grima *et al*., 2008 showing Pacman and Actin loading control performed on male and female adult flies. Pcm expression is visibly higher in the control flies (OregonR) than any of the three *pcm* alleles tested. Of the *pcm* alleles, $pcm^5$ shows the lowest level of Pcm.
**B**. TaqMan comparative quantitative RT-PCR comparing wild-type (50E) *pcm* mRNA expression to $pcm^5$, $pcm^{14}$, $pcm^3$ and $pcm^{13}$ in L3 larvae. Only $pcm^5$ shows a significantly reduced level of *pcm* compared to wild-type. The $pcm^3$ level is the same as wild-type, which means the lower level of Pcm protein must be due to reduced translation of the mRNA, or increased degradation of the protein. The protein level in $pcm^{14}$ cannot currently be measured, but its mRNA level is the same as wild-type. $pcm^{13}$ appears to have a slightly reduced level of *pcm*, but this result needs confirmation.
** = p-value <0.001.

### 4.6.2 *pacman* mRNA expression in *pacman* mutants

The level of Pcm protein in *pcm* alleles varies. This could be due to reduced translation of *pcm* mRNA or reduced levels of *pcm* mRNA. TaqMan comparative quantitative RT-PCR (qPCR) was used to test the relative levels of *pcm* mRNA in $pcm^{5}$, $pcm^{14}$, $pcm^{3}$ and $pcm^{13}$ L3 larvae (Figure 4.15, panel B). This technique compares the level in wild-type with each mutant and gives the result as a fold change in expression. The *pcm* mRNA level in $pcm^{5}$ was found to be consistently downregulated by 3-fold ($p<0.001$). $pcm^{14}$ and $pcm^{3}$ express their versions of *pcm* mRNA at the same level as wild-type. Only one test was performed on $pcm^{13}$ and this gave a level intermediate between that of wild-type and $pcm^{5}$, about 1.5-fold downregulated. The assay used for *pcm* is targeted at the interface between exons 3 and 4, so is capable of detecting wild-type *pcm* and any of the currently available alleles.

These results indicate that less Pcm protein is present in $pcm^{3}$ due to reduced translation of *pcm* mRNA, which is present at the wild-type level, or an increased rate of Pcm protein turnover. Although the mRNA level of $pcm^{6}$ has not been tested, it is likely to produce similar results to $pcm^{3}$ as the deletions present in these alleles are similar in nature. $pcm^{14}$ is not an mRNA null as *pcm* mRNA is expressed at the same level as wild-type. This is not surprising as the transcription start site and any sequence upstream of it (presumably including the *pcm* promoter) are unaffected by the deletion in $pcm^{14}$. As the nature of the $pcm^{14}$ deletion is similar to the $pcm^{3}$ deletion, albeit larger, it seems likely Pcm protein is produced by the $pcm^{14}$ allele. The genetic data acquired show $pcm^{14}$ is a functional null, so if protein is produced, it must be non-functional due to the lack of the deleted region, and/or rapidly degraded. $pcm^{5}$ mRNA is downregulated by 3-fold compared to wild-type. It seems likely this is due to nonsense-mediated decay in response to the aberrant $pcm^{5}$ transcript. The $pcm^{5}$ deletion differs from other *pcm* alleles as it is smaller and internal to the gene. The remainder of the *pcm* gene after the $pcm^{5}$ deletion is out of frame, which may cause the mRNA to be detected as aberrant. This 3-fold downregulation at the mRNA level may account for the >3-fold downregulation of the Pcm protein, depending on how the normal rate of turnover of $pcm^{5}$ Pcm compares to the normal rate of translation. Alternatively, the transcript may be translated at a lower rate than wild-type and/or the protein may be turned over at a higher rate than wild-type, which results in undetectable levels of Pcm on a western blot of $pcm^{5}$ flies. Despite the low levels of Pcm, what is present must be at least partially functional as the $pcm^{5}$ allele is hypomorphic. It may be the case that the $pcm^{5}$ deletion does not actually affect the function of the Pcm protein it produces, but the phenotype comes about due to the reduced level of Pcm protein. $pcm^{13}$ requires

further replicates to confirm the level of downregulation seen in the first test, but this level being different to $pcm^{3/14}$ and $pcm^{5}$ could be due to the fact $pcm^{13}$ is the only *pcm* mutation of its type. The level of Pcm protein in $pcm^{13}$ has not been tested, but it is likely downregulated in a similar way to the other *pcm* mutants.

### 4.6.3 Quantifying Pacman protein function in *pacman* mutants

An experiment has been planned to relatively quantify the level of function of Pacman protein in the mutant lines, by performing qPCR on a transcript degraded by Pacman. Due to the length of time required to produce the flies needed for this experiment (six generations), the full results were not ready in time to be included in this thesis. The full plan for this experiment, and early preliminary results, can be found in the Future work section of Chapter 7.

## 4.7 Genetic deficiencies affecting the pacman locus

This section contains the information known about each of the deficiencies used earlier in this chapter and presents the work performed to create *Df(1)ED7452*. Deficiencies are deletions of large areas of chromosome and typically include many genes, depending on their location. As deficiencies are certain to be null for any genes located in the deleted region, they are ideal for determining the character of new gene mutations. Three deficiencies are available that include the *pcm* locus and these are summarised in Figure 4.16 and below.

### 4.7.1 *Df(1)JA27*

*Df(1)JA27* is the largest deficiency that includes *pcm* and was created by X-ray induced mutagenesis. The breakpoints have been mapped cytologically. The maximum size is 501,546bp (18A5--18D) which covers 75 genes, and the minimum size is 396,337bp (18A5--18D1). Flies hemi- or homozygous for *Df(1)JA27* are embryonic lethal.

### 4.7.1 *Df(1)Exel7468*

*Df(1)Exel7468* was one of a large set of deficiencies created by utilising FRT-bearing insertions and FLP recombinase. The advantage of creating deletions with this method is they are predictable, as the locations of FRT insertions are known (Parks *et al.*, 2004). *Df(1)Exel7468* is a deletion of 245,148bp covering 30 genes (18B7--18C8). Flies hemi- or homozygous for *Df(1)Exel7468* are embryonic lethal.

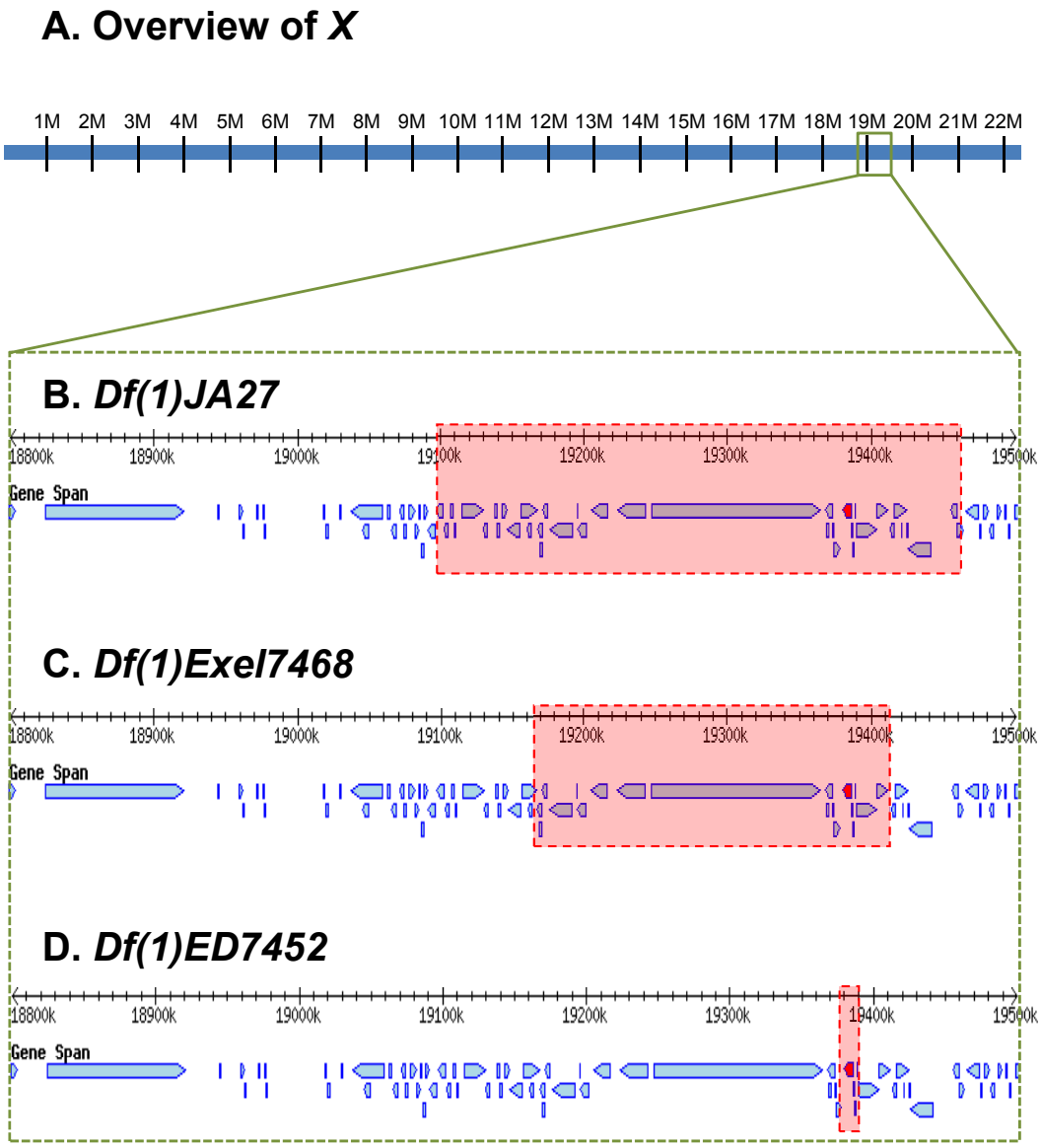



**Figure 4.16**: Available deficiencies that include the *pcm* locus.
**A**. Location of 700kbp expanded area on the *X* chromosome.
**B**. *Df(1)JA27* – minimum size of deficiency shown, 396,337bp (18A5--18D1). Maximum size is 501,546bp (18A5--18D).
**C**. *Df(1)Exel7468* – 245,148bp, 30 genes.
**D**. *Df(1)ED7452* – 17,963bp, 5/6 genes.
*pacman* (6,071bp) is highlighted in red.

### 4.7.2 *Df(1)ED7452*

The DrosDel deficiency collection was created using engineered *p{RS3}* and *p{RS5}* P-elements, in a similar fashion to *Df(1)Exel7468*. The original collection consisted of 3,243 deletions of 35kb on average. Using the collection of insertions, it is possible to generate another >12,000 deletions between 1bp and 1Mb by eye colour selection and 37,000 by molecular screening (Ryder *et al.*, 2004, Ryder *et al.*, 2007). The DrosDel website provides a facility for finding the P-element stocks needed for creating custom deletions. There are 199 possible deletions that include *pcm*, varying in size from 17,963bp to 991,348bp. Of these, only one has previously been created (*Df(1)ED7494*, 397,939bp), but it is not publically available. As we are looking at a single gene, the best deficiency for our purposes is the smallest possible. For this reason we decided to attempt to create *Df(1)ED7452*, which is smallest by almost an order of magnitude. At 17,963bp it is much smaller than *Df(1)Exel7468* and *Df(1)JA27*. The area deleted includes the entirety of *pcm*, *CG12203*, *CG12204* and *kish* (*CG14199*). The 5' breakpoint lies within *Nat1* and all five translated exons of this gene are deleted (only part of the 5' UTR is unaffected). The 3' breakpoint lies within the gene *Pfrx* (*6-phosphofructo-2-kinase*). The deletion includes the alternate first (untranslated) exons of the *Pfrx-RF* and *Pfrx-RA*/*RI* transcripts, but does not affect five other transcript variants. There are only two possible coding sequences (translated regions) from the eight *Pfrx* transcripts and both of these can be produced from the transcripts that are not affected by the deletion. Depending on the location of the *Pfrx* promoter, this gene may still be able to perform its function. Figure 4.17 shows the region of *X* affected in by the *Df(1)ED7452* deficiency, before and after rearrangement.

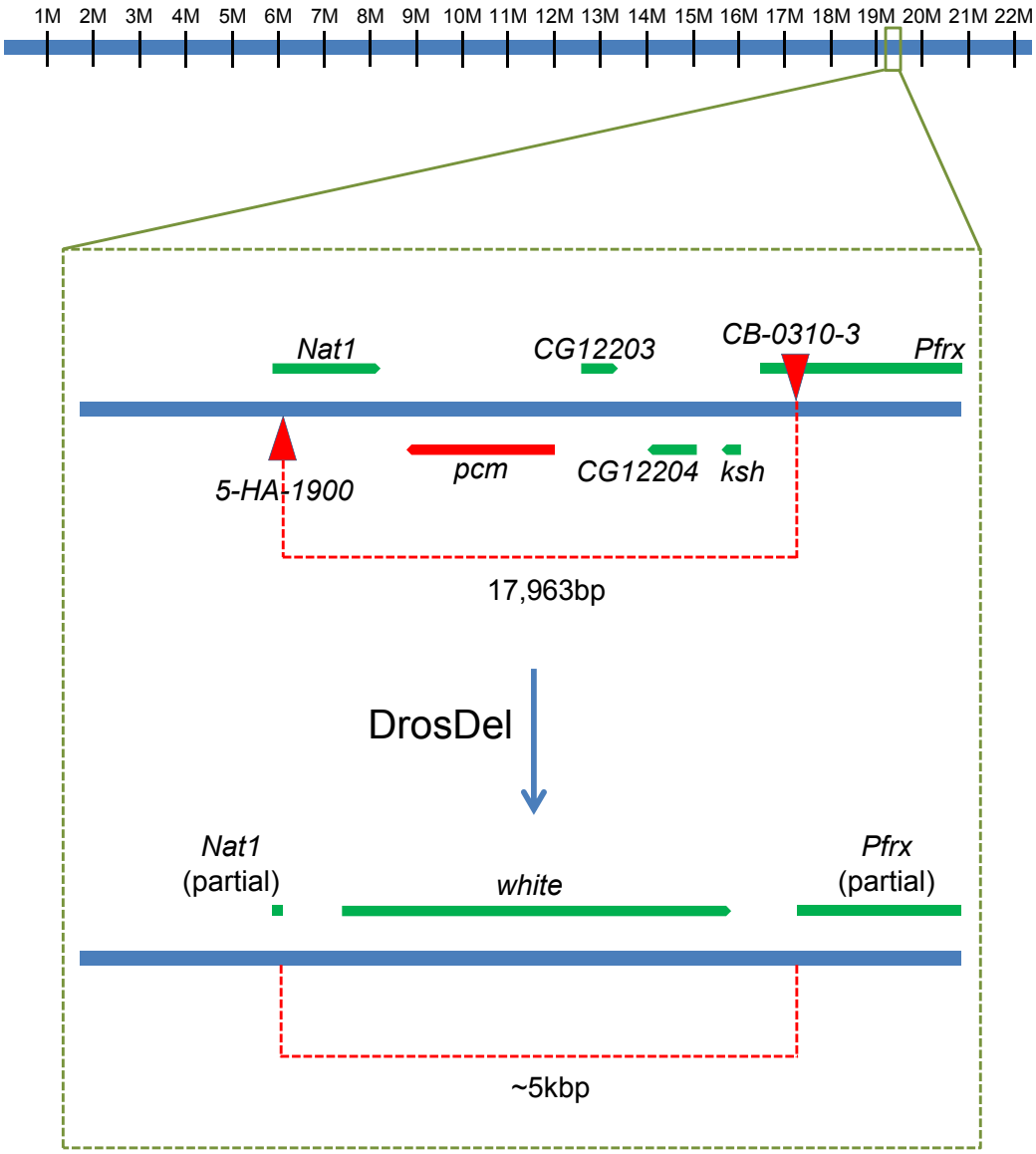



**Figure 4.17**: The region of the *X* chromosome affected by *Df(1)ED7452*, before and after rearrangement.
The two P-elements (red triangles) are utilised in the DrosDel scheme to cause replacement of the intervening region with a functioning allele of *white*. Four genes are deleted, including *pcm*. *Nat1* is almost entirely deleted, with only part of the 5' UTR remaining. A small part of the 5' of *Pfrx* is deleted, but it may still be possible for the gene to function.
P-element positions are shown on a single chromosome in this diagram for simplicity and gene locations are approximate.

## 4.8 Creation of *Df(1)ED7452*

### 4.8.1 DrosDel RS overview

The DrosDel RS (Rearrangement Screen) follows the principles for creating chromosomal rearrangements originally described by (Golic and Golic, 1996). The P-elements required are *5-HA-1900* (*P{RS5}Nat1*$^{5\text{-}HA\text{-}1900}$, stock BL125461) and *CB-0310-3* (*P{RS3}Pfrx*$^{CB\text{-}0310\text{-}3}$, stock BL123119). Other stocks required are the FLP containing stock (*w*$^{1118}$*; noc*$^{Sco}$*/SM6B, P{70FLP ry*$^{+}$*}*, stock BL6876) and a stock for use as a source of *FM7h* (BL6878). There are two main steps to the RS screen, followed by phenotypic selection and molecular verification of lines carrying the *Df(1)ED7452* deletion (Figure 4.18). The chromosomal form for each P element includes a functional *white* gene (*w*$^{hs}$) and two FRT sites. The only differences between the elements are the location of the FRT sites in relation to the 5' and 3' sections of the *white* gene. FLP mediated excision is used to produce the "remnant" form of each P-element (*P{RS3r}* or *P{RS5r}*) which now each contain one FRT site and either the 5' or 3' region of the *white* gene. The 5' and 3' regions are non-functional, so these flies are selected based on loss of eye colour. FLP mediated recombination between the *RS3r* and *RS5r* elements will result in a reconstituted, functional *white* gene with two 3' P-element ends, and deletion of the intervening sequence. These flies are identified by their regained eye colour.

### 4.8.2 DrosDel Crosses

Seven generations are required to produce the final *Df(1)ED7452* carrying flies. The first three "Flip Out" crosses are to generate separate stable stocks of each P-element remnant. The last four "Flip In" crosses are to generate the stable stocks containing *Df(1)ED7452*.

#### 4.8.2.1 "Flip Out" crosses

The first cross using virgins carrying FLP (on *SM6b*, marked with *Curly*) and males with the FRT containing P-element (either *5-HA-1900* or *CB-0310-3*) is followed by heat shock treatment of the offspring to induce excision to produce the P-element remnant. Male offspring are crossed to *FM7h* virgins in the second cross. The eye colour loss resulting from successful excision can be seen in the offspring of this cross, and single white eyed non-*Curly* females (*P{w*$^{-}$*}/FM7h*) are crossed to *FM7h* males to form the stable stock (Figure 4.19).

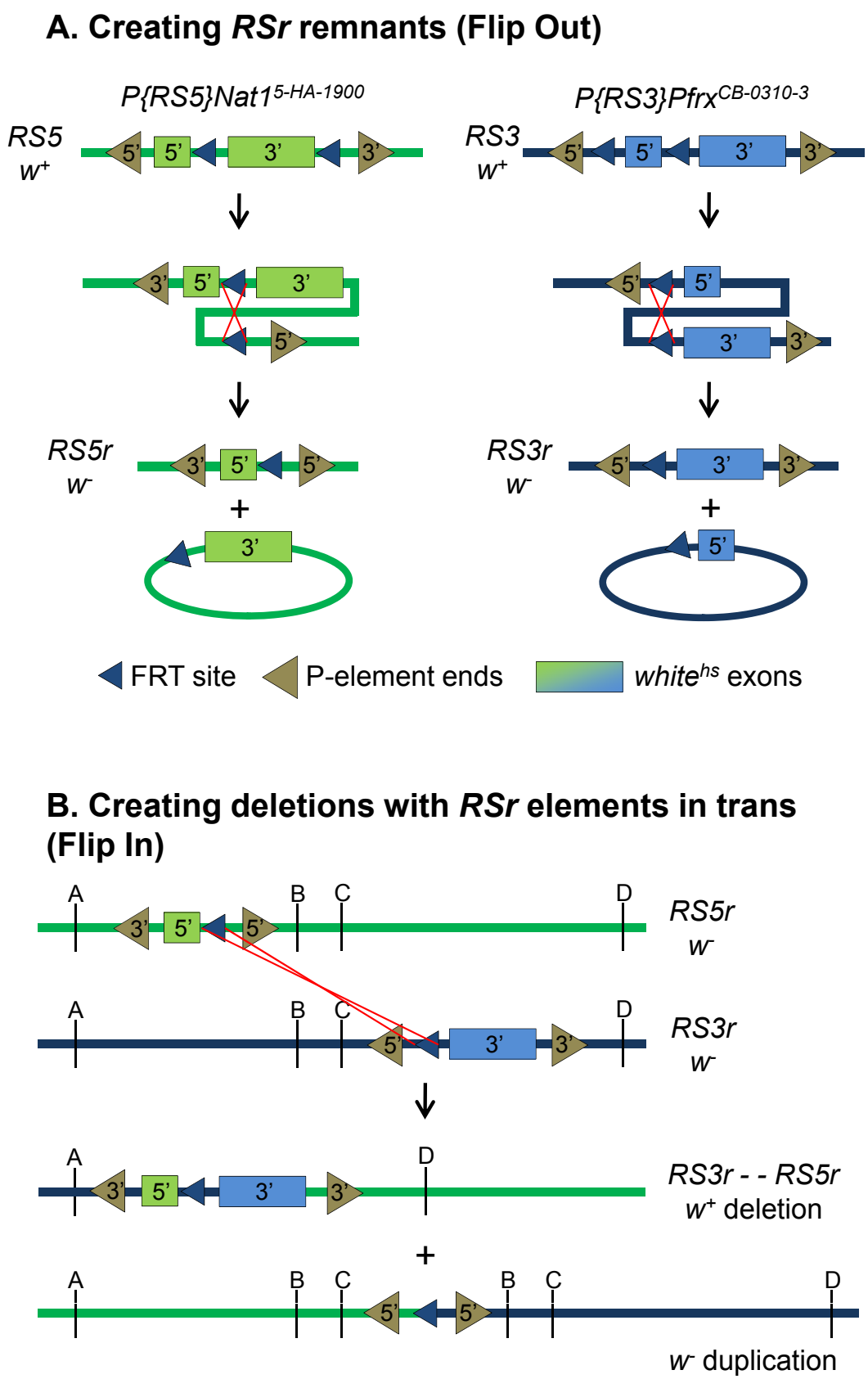



**Figure 4.18**: Using DrosDel RS elements to create deletions.
**A**. "Flip Out" of $w^+$ *RS5* and *RS3* elements by recombination between FRT sites to create $w^-$ *RS3r* and *RS5r* elements.
**B**. "Flip In" between a *RS5r* and a *RS3r* element to produce the $w^+$ deletion (e.g. *Df(1)ED7452*).

Images reproduced from Ryder, 2007.

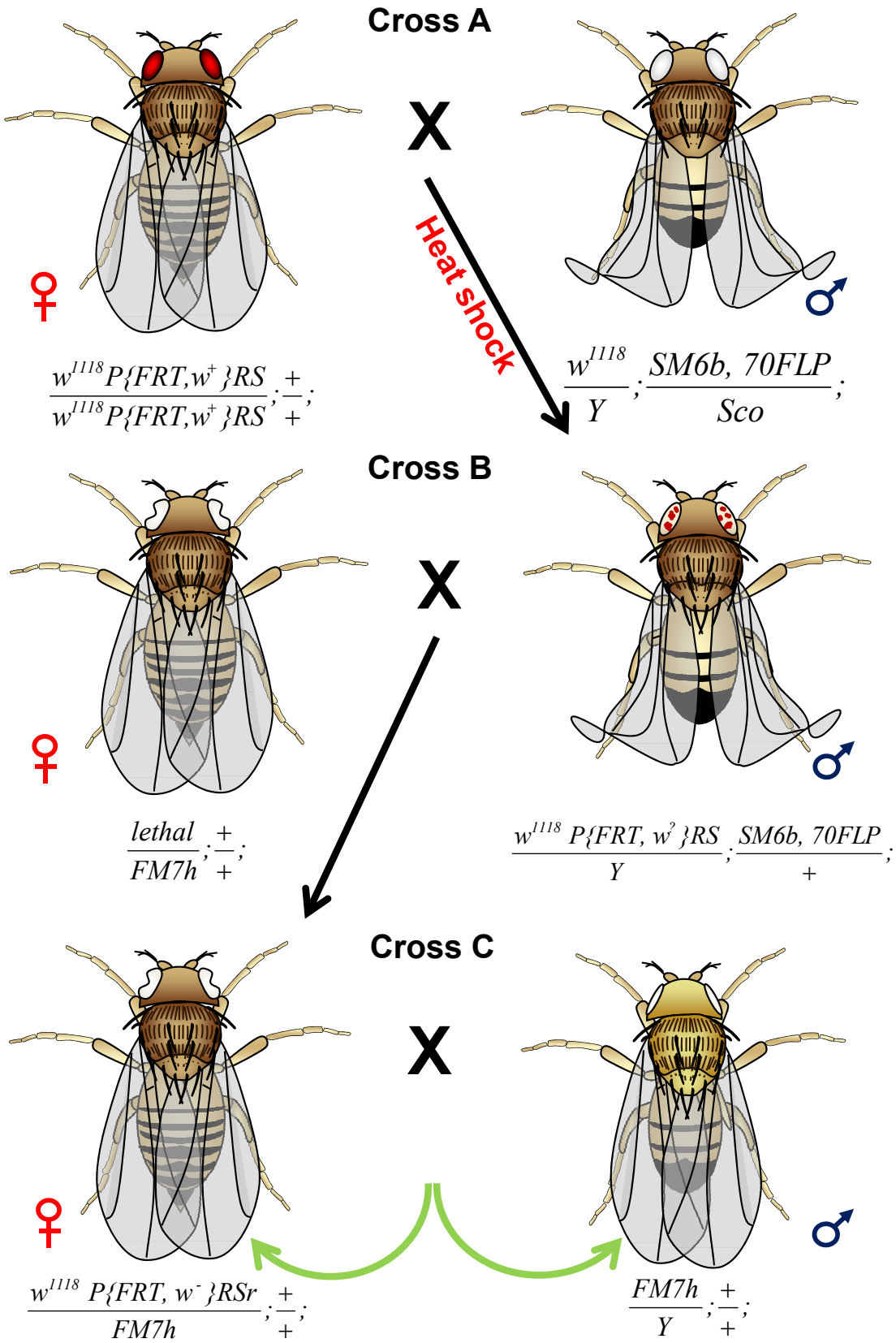



**Figure 4.19**: "Flip Out" crosses to generate stable lines of flies carrying *RS3* and *RS5* remnants (*RS3r* and *RS5r*). Black arrows indicate which offspring from a cross were used in the subsequent cross. Green arrows indicate a stable stock where genotypes of the offspring are the same as the genotypes of the parents. Offspring from Cross A were subjected to heat shock for 1 hour at 37°C 48-72 hours after egg laying to induce expression for the FLP recombinase.

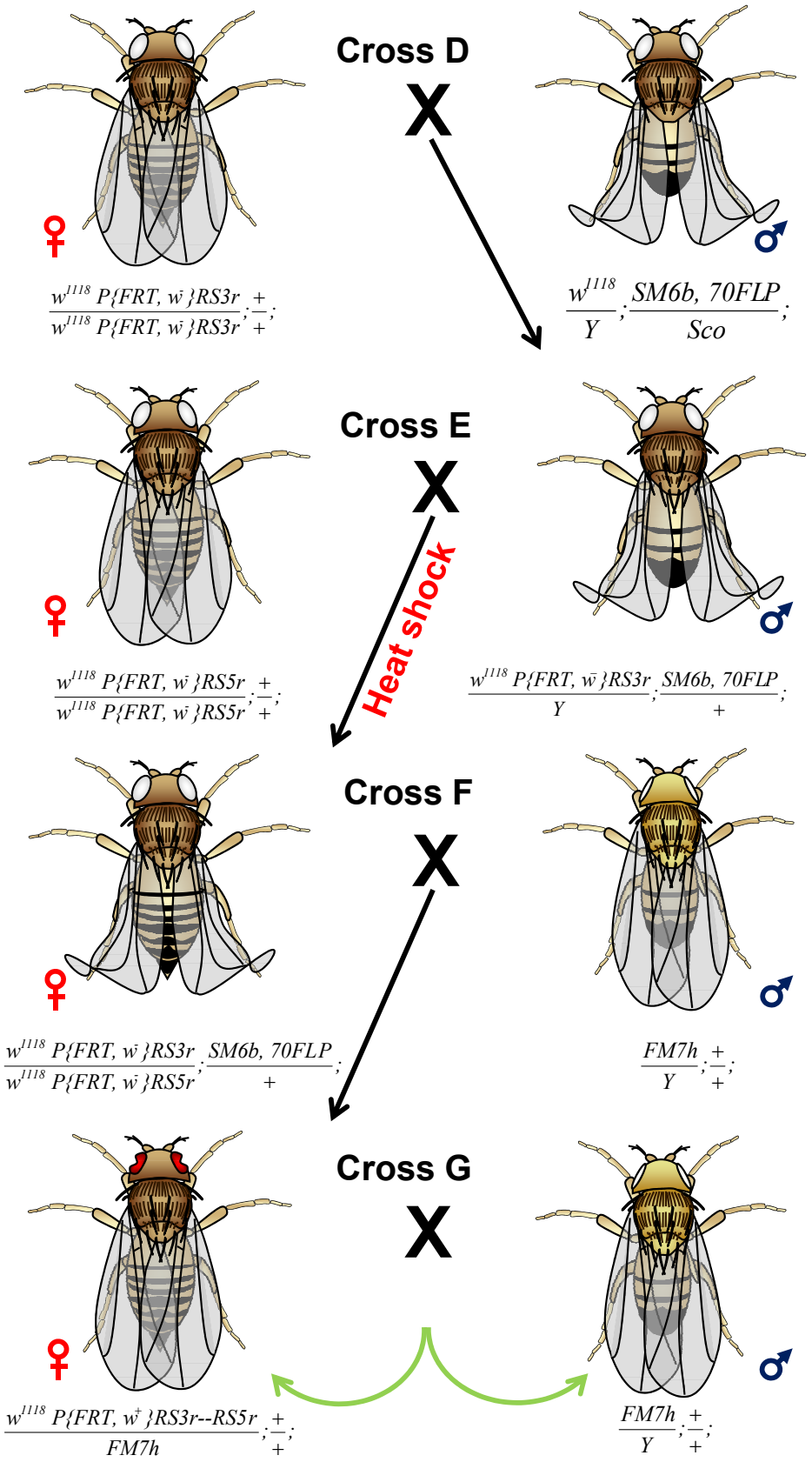



**Figure 4.20**: "Flip In" crosses to generate stable lines of flies with the *Df(1)ED7452* deletion. Black arrows indicate which offspring from a cross were used in the subsequent cross. Green arrows indicate a stable stock where genotypes of the offspring are the same as the genotypes of the parents. Offspring from Cross E were subjected to heat shock for 1 hour at 37 C 48-72 hours after egg laying and again 48 hours later to induce expression for the FLP recombinase. Individual females were used in Cross G and the resulting stocks were tested by PCR to ensure the deletion was as expected.

#### 4.8.2.2 "Flip In" crosses

In the first cross the FLP containing males are crossed to virgins carrying one of the *P{w⁻}* remnants. Resulting males containing both the FLP chromosome and remnant are crossed to virgins carrying the other remnant, and the offspring of this cross are heat shocked to induce FLP expression and recombination between the FRT sites of the two remnants. Virgin females containing both remnants and the FLP chromosome are crossed to *FM7h* males to remove the FLP chromosome. Non-*Curly*, red eyed virgins are selected from the offspring and singly crossed to *FM7h* males to produce the final stocks that should contain *Df(1)ED7452* (Figure 4.20).

### 4.8.3 Identification of *Df(1)ED7452* lines

After the flip out crosses, the two remnant carrying lines should have white eyes instead of the starting eye colour of almost wild-type red. In reality, the remnant from *HA-1900-5* gave white eyes and the remnant from *CB-0310-3* gave light orange eyes. After recombination between the remnants, the eye colour produced was dark orange. Two lines with dark orange eyes were isolated and both were homozygous embryonic lethal (referred to as DD2 and DD3).

### 4.8.4 Molecular verification of *Df(1)ED7452* by PCR

Phenotypes give a good indication that the desired event has occurred, but the deficiency must be verified molecularly. The reconstituted *white* gene and 3' P-element ends which should now exist in place of the 17,963bp of deleted sequence contain four sites suggested for use as primer sites to verify the deletion. Two primers are present in the *white* gene itself, the forward primer (W7500D) is located in the 5' section and the 3' section contains the reverse primer (W11678U). There are also two PRY4 primer sites located close to the 3' P-element ends, in the same orientation (i.e. in opposite orientations to each other) (Figure 4.21, panel A). W7500D and W11678U can be used to ensure the *white* gene has been reconstituted as expected. The two PRY4 sites can be combined with two custom primers to verify the breakpoints of the deficiency. The two lines with dark orange eyes were tested by PCR for the reconstituted *white* gene and the breakpoints were tested using the custom primers ED7452F and ED7452R combined with PRY4. Both lines gave the same result, which was as expected for successful creation of the deficiency (Figure 4.21, panel C). Sequencing of the ED7452F+PRY4 and ED7452R+PRY4 fragments confirmed the exact breakpoints of the deficiency.

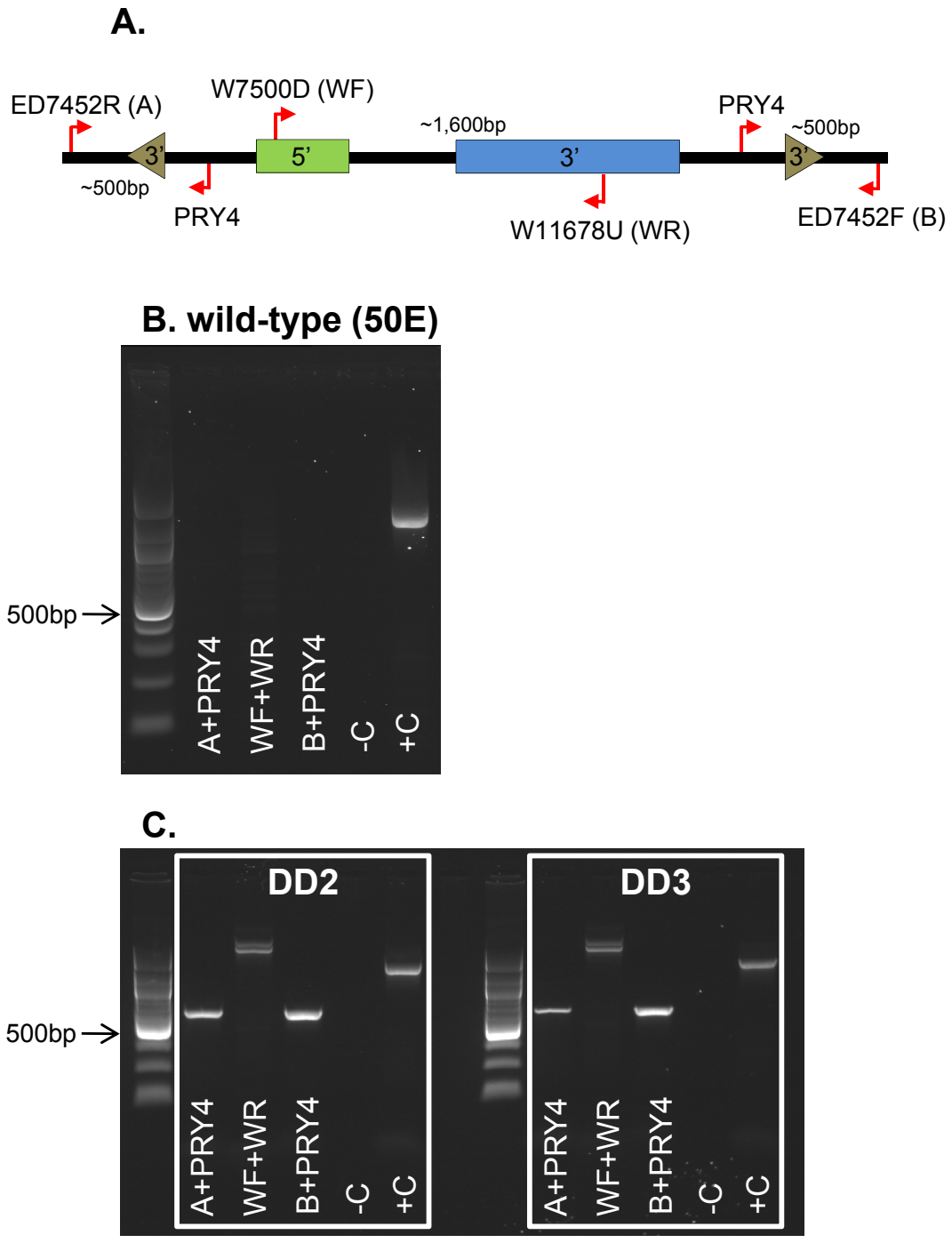



**Figure 4.21**: Molecular verification of the *Df(1)ED7452* deletion.
**A**. Diagram of *white* gene and 3' P-element ends in place of the deleted sequence in *Df(1)ED7452* and location of primers used to verify successful deletion.
**B**. Three primer pairs used on wild-type (50E) gDNA. None produce a band, as expected. The endogenous *white* gene does not produce a band with the WF and WR primers as the allele in 50E is $w^{1118}$, which lacks one of the primer sites.
**C**. PCR verification performed on the two lines with phenotypes consistent with a successful deletion (red eyes, lethality). Both lines give exactly the same result, which is as would be expected for a successful creation of *Df(1)ED7452*. These lines also carry $w^{1118}$.

### 4.8.5 Genetic verification of *Df(1)ED7452* by chromosomal crossover

As the created lines were homozygous lethal, we decided to ensure that the lethality came from the region containing the deficiency and was not due to an unrelated event. Chromosomal crossover between the *Df(1)ED7452* chromosome and the marker chromosome $y^1\ cv^1\ v^1\ f^1\ car^1$ was used to test this, in the same way that the lines $pcm^{14}$ (16A) and 50B were tested in at the start of this chapter. Both *Df(1)ED7452* lines produced results consistent with the lethality coming from the deficiency (Figure 4.22).

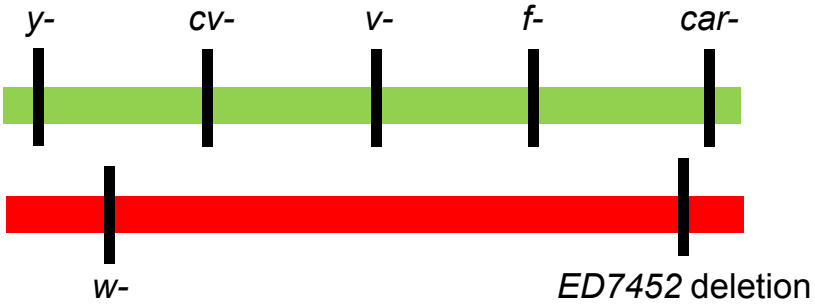
### 4.8.6 Rescue of *Df(1)ED7452* by *T(1;Y)B92*, a translocation from *X* to *Y* including the *pcm* locus

The ability of *T(1;Y)B92* to rescue the lethality of *Df(1)ED7452* was tested at the same time as rescue of $pcm^{14}$ lethality was attempted. *Df(1)ED7452* was successfully rescued by *T(1;Y)B92* (Figure 4.4). The fertility of the rescued males was below the level of *In(1)FM7/T(1;Y)B92,* as was the fertility of $pcm^{14}$*/T(1;Y)B92* males (Figure 4.5).

### 4.8.7 Comparison between *Df(1)JA27* and *Df(1)ED7452*

The only practical difference between *Df(1)JA27* and *Df(1)ED7452* was that the lethality of *Df(1)JA27* could not be rescued by *T(1;Y)B92*, but this may have been due to the coverage of *T(1;Y)B92*. The results did not differ when *Df(1)JA27* and *Df(1)ED7452* combined with the hypomorphic allele $pcm^5$, as both were viable at 25°C and lethal at 19°C. The $pcm^5$*/Df(1)JA27* and $pcm^5$*/Df(1)ED7452* flies at 19°C died after failing to fully eclose (Figure 4.6).

## A. Chromosomes




## B. DD2/DD3 crossover

| Possible marker combinations | No. observed DD2 | | No. observed DD3 | |
|---|---|---|---|---|
| y- w+ cv- v- f- car- | 141 | 48% | 138 | 41% |
| y+ w- cv- f- | 20 | 7% | 30 | 9% |
| y+ w- cv+ f- | 106 | 36% | 131 | 39% |
| y+ w- cv+ f+ | 10 | 3% | 14 | 4% |
| y- w+ cv+ v+ f+ car+ | 0 | 0% | 0 | 0% |
| y- w+ cv- v+ f+ car+ | 0 | 0% | 0 | 0% |
| y- w+ cv- v- f+ car+ | 0 | 0% | 0 | 0% |
| y- w+ cv- v- f- car+ | 0 | 0% | 0 | 0% |
| y- w+ cv+ v- f- car- | 0 | 0% | 1 | <1% |
| y- w+ cv+ v+ f- car- | 5 | 2% | 8 | 2% |
| y- w+ cv+ v+ f+ car- | 4 | 1% | 2 | <1% |
| y- w+ cv- v+ f- car- | 4 | 1% | 10 | 3% |
| y- w+ cv- v+ f+ car- | 3 | 1% | 3 | 1% |
| y- w+ cv- v- f+ car- | 0 | 0% | 0 | 0% |
| y+ w- cv- f+ | 0 | 0% | 0 | 0% |
| y+ w- cv+ v+ f+ car+ | 0 | 0% | 0 | 0% |
| | 293 | | 337 | |

**Figure 4.22**: Crossover between DD2/DD3 (potential *Df(1)ED7452* lines) and marker chromosomes.
**A.** Approximate location of markers on marker chromosome (green) and DD2/3 chromosomes (red).
**B.** Numbers and frequencies of flies with specific marker combinations observed. Grey blocks indicate when the phenotypes of $v^1$ and $car^1$ are not observable, due to the presence of $w^{1118}$. No flies were seen containing the section of DD2/3 chromosome that includes the *Df(1)ED7452* deficiency and wild-type *car*, indicating the lethality of this chromosome is due to a lesion in this region.

## 4.9 Chapter summary

The main aims of this chapter were to verify that the lethality seen in the $pcm^{14}$ mutant is due to the *pcm* deletion and to observe the phenotypes of $pcm^{14}$, including its interactions with other *pcm* alleles and deficiencies, and to determine if $pcm^{14}$ is a null allele.

### 4.9.1 The source of the lethality on the $pcm^{14}$ chromosome localises to the *pcm* locus

The first aim was addressed using crossover, rescue and by comparing $pcm^{14}$ with other *pcm* alleles. The crossover and rescue experiments showed that the lethality in $pcm^{14}$ mapped to the area that included *pcm*. Comparing $pcm^{14}$ homozygotes with $pcm^{14}/pcm^{5}$ heterozygotes at 19°C and 25°C produces a scale of phenotype severity consistent with the *pcm* locus being the source of lethality:
$pcm^{14}$ 19°C (lethal) = $pcm^{14}$ 25°C (lethal) > $pcm^{14}/pcm^{5}$ 19°C (lethal) > $pcm^{14}/pcm^{5}$ 25°C. If the lethality was due to a lesion at another locus on the $pcm^{14}$ containing chromosome, it would be rescued by the $pcm^{5}$ chromosome at 19°C. This is not the case and instead, the temperature sensitive $pcm^{5}$ allele fails to fully rescue the $pcm^{14}$ lethality. One copy of $pcm^{5}$ at 19°C was already known to not be sufficient for survival. At 25°C the single copy of $pcm^{5}$ is sufficient to allow flies to survive, which was also previously known.

### 4.9.2 $pcm^{14}$ is a null allele

The fact that $pcm^{5}/pcm^{14}$ and *$pcm^{5}$/deficiency* flies are indistinguishable by phenotype or by frequency at 19°C shows that $pcm^{14}$ is a null allele. This is supported by the rescue data that shows $pcm^{14}$ and *Df(1)ED7452* flies can both be rescued by *T(1;Y)B92*, but each suffer from almost complete infertility.

#### 4.9.2.1 $pcm^{14}$ is most likely amorphic at the functional level

The level of Pacman protein has previously been shown to be almost undetectable in $pcm^{5}$ mutants and at a lower level than wild-type in $pcm^{3}$ and $pcm^{6}$. The level of protein in $pcm^{14}$ cannot be tested as the region targeted by the Pcm antibody is deleted. However, it seems likely that $pcm^{14}$ mRNA acts in a similar way to $pcm^{3}$ and $pcm^{6}$ mRNA as the nature of the deletion is common between these alleles, and the level of *pcm* mRNA is the same as wild-type in $pcm^{14}$, $pcm^{3}$ and $pcm^{6}$ flies. If the $pcm^{14}$ mRNA is capable of producing protein, it must be non-functional, given the evidence that shows the $pcm^{14}$ allele is amorphic. $pcm^{5}$ is the only allele that shows a significant decrease in *pcm* mRNA expression, by 3-fold. This

is likely due to nonsense-mediated decay and may be a factor that causes *pcm*$^{5}$ to be the most severe hypomorphic allele. The level of *pcm* mRNA in *pcm*$^{13}$ may also be lower, but there is currently insufficient data to determine if there is a significant difference.

### 4.9.3 *pcm*$^{14}$ causes pupal lethality and decreased size of L3 wing imaginal discs

The phenotypes of *pcm*$^{14}$ are the same at 19°C and 25°C (again providing evidence it is a null). *pcm*$^{14}$ flies die during early pupation and rarely form any visible adult structures. The wing imaginal discs of *pcm*$^{14}$ L3 larvae raised at 25°C are significantly smaller than wild-type wing discs of the same age (43% of wild-type size). Consistent with the observations that *pcm*$^{14}$ is a stronger allele than *pcm*$^{5}$, *pcm*$^{5}$ wing discs of the same age are intermediate in size between *pcm*$^{14}$ and wild-type discs (82% of wild-type size). The possible reasons for the reduced disc sizes seen in *pcm* mutants are discussed in Chapter 7.

### 4.9.4 *pcm*$^{14}$/*pcm*$^{3/5}$ heterozygotes show greater phenotype penetrance than *pcm*$^{3/5}$ homozygotes

Crossing *pcm*$^{14}$ to *pcm*$^{5}$ or *pcm*$^{3}$ has clear effects indicative of *pcm*$^{5}$ and *pcm*$^{3}$ being hypomorphic alleles and *pcm*$^{14}$ being a null allele. *pcm*$^{14}$/*pcm*$^{5}$ flies die at 19°C as described above, and at 25°C the frequency of the dull wing phenotype in *pcm*$^{14}$/*pcm*$^{5}$ flies is 54%, compared to 2% in *pcm*$^{5}$ homozygotes. The same trend is seen for *pcm*$^{3}$, at 19°C the frequency of the dull wing phenotype for *pcm*$^{3}$ homozygotes is 99% but drops to 6% at 25°C. For *pcm*$^{14}$/*pcm*$^{3}$ heterozygotes, the frequency is 100% at 19°C and 92% at 25°C. *pcm*$^{14}$/*pcm*$^{6}$ flies do not exhibit any phenotypes at a frequency >20% at 19°C or 25°C, and there is no obvious great increase in frequency of any phenotype when the temperature is decreased. Therefore, the evidence suggests *pcm*$^{6}$ is a weaker allele than *pcm*$^{3}$ or *pcm*$^{5}$.

### 4.9.5 *Df(1)ED7452*, a <18kb deficiency including the *pacman* locus has been successfully created

Using the DrosDel Rearrangement Screen system, *Df(1)ED7452* has been created and verified using genetic and molecular techniques. *Df(1)ED7452* is 17,963bp in size, which fully deletes four genes including *pcm*, almost entirely deletes one gene and deletes a small portion of the upstream sequence of another gene. *Df(1)ED7452* is 22-28 times smaller than *Df(1)JA27*, the deficiency previously most used in the Newbury lab.

### 4.9.6 Following chapters

The first two chapters of this thesis have described the creation and effects of a *pcm* null allele. One of the most developmentally interesting aspects of all the *pcm* alleles is the effect on wings and/or wing imaginal discs. The following two chapters report work performed to identify mRNA and microRNA changes in the wing imaginal discs of $pcm^{5}$ L3 larvae. $pcm^{5}$ wing discs were used instead of $pcm^{14}$ wing discs due to ease of extraction, and at the time extraction was started $pcm^{14}$ had not fully been verified as containing only a *pcm* mutation.

# 5 Chapter 5: Effect of *pacman* mutations on specific messenger RNAs

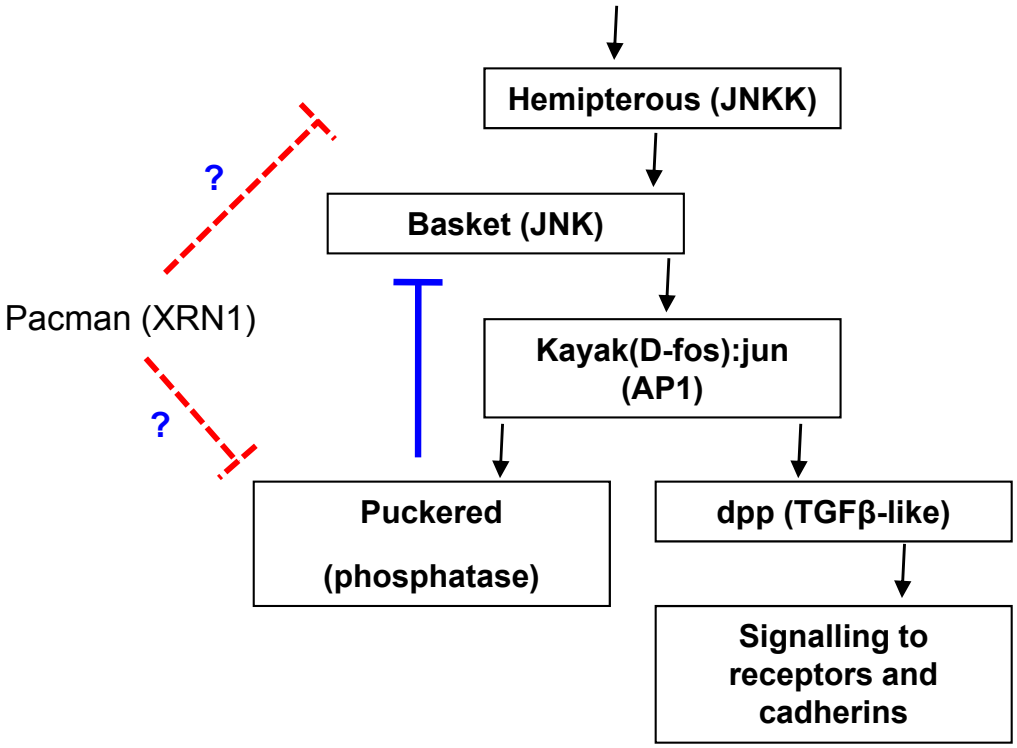
## 5.1 Pacman and degradation of specific mRNAs

As the only cytoplasmic 5' – 3' exoribonuclease, Pacman has high potential to post-transcriptionally regulate mRNAs that undergo 5' – 3' degradation. *pcm* alleles exhibit a phenotype of small L3 wing imaginal discs, which are smallest in the null mutant $pcm^{14}$. Hypomorphic *pcm* mutants exhibit phenotypes in the structures that develop from wing imaginal discs, such as dull wings and bristle defects. This shows that Pacman is required for normal wing/thorax development. The obvious hypothesis is that lack of Pacman leads to loss of post-transcriptional downregulation of one or more mRNA transcripts in wing imaginal discs, which in turn prevents wild-type development of adult structures from the wing discs. The objective of the work presented in this chapter was to test this hypothesis and to determine if any misregulated mRNAs could be responsible for the phenotypes observed.

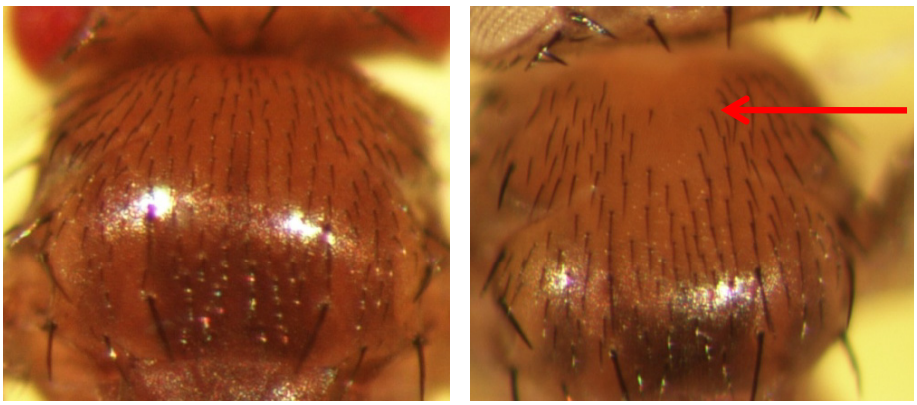
## 5.2 Pacman and the JNK pathway hypothesis

Previous work in the Newbury lab has suggested a link between Pacman and the JNK pathway. The JNK pathway is involved in epithelial sheet sealing in animals, and is active in processes such as dorsal and thorax closure during *Drosophila* development and in wound healing (Zeitlinger *et al.*, 1997, Agnes *et al.*, 1999, Martin-Blanco *et al.*, 2000). Results from the PhD thesis of Dr. Melanie Sullivan suggest Pacman could interact with the JNK pathway via Puckered, which is part of the negative feedback loop initiated by the pathway. Her results show an increase in the protein level of Basket (JNK) and an increase in the mRNA level of *puckered* (MKP1) in *pacman* mutants. They also show a genetic interaction between *pacman* and *puckered* which manifests as a "bald patch" phenotype at the anterior of the thorax of double mutant flies (Figure 5.1) (Grima *et al.*, 2008, Sullivan, 2008). The initial aim of this chapter was therefore to test the hypothesis that Pacman is affecting the JNK pathway.

## A. Pacman and the JNK pathway hypothesis




## B. Genetic interaction between *pcm* and *puc*



*+/Y ; ; puc$^{A251}$/TM6, Tb Sb*  *pcm$^{5}$/Y ; ; puc$^{A251}$/TM6, Tb Sb*

**Figure 5.1**: Hypothesis that Pacman affects the JNK pathway and evidence for an interaction.
**A.** The hypothesis proposed in Dr. Melanie Sullivan's PhD thesis (Sullivan, 2008). Under normal conditions Pacman regulates JNK pathway by degrading the mRNA of one or more of its members.
**B.** The genetic interaction observed between *pacman* and *puckered*. The "bald spot" phenotype, a chaetae-less area at the anterior of the thorax, is exclusive to double mutant flies and occurs at both 19°C and 25°C.

## 5.3 Preliminary TaqMan Quantitative RT-PCR analysis of JNK pathway transcripts

To test the hypothesis that Pacman regulates the JNK pathway via *puckered* we decided to perform TaqMan comparative quantitative Reverse Transcriptase PCR (qPCR) on L3 larvae from four strains: wild-type (50E), $pcm^3$, $pcm^5$ and $pcm^{14}$ and wing imaginal discs from wild-type and $pcm^5$. If Pacman does degrade *puckered* mRNA during wild-type development, we would expect to see an increase in *puckered* mRNA in *pacman* mutants. If this is sufficient to disrupt the JNK pathway, other transcripts should also change in level due to increased or decreased transcription.

TaqMan qPCR is superior to SYBR Green qPCR (the technique used to produce the qPCR results in Dr. Melanie Sullivan's thesis) as a fluorescent probe specific to the target transcript is used instead of a fluorescent reporter than binds any double stranded DNA. This reduces the problem of non-specific amplification, which can hinder the production of good quality results when using SYBR Green.

### 5.3.1 Whole larvae

The JNK pathway is known to be active in the wing imaginal discs of L3 larvae and previous results suggest that Pacman affects the pathway in these tissues. However, the effect may not be limited to the wing imaginal discs, so initially the expression levels of a number of JNK transcripts, including *puckered,* were tested in whole L3 larvae (Figure 5.2). The strains used were wild-type (50E), $pcm^3$, $pcm^5$ and $pcm^{14}$. The results are displayed with the level of each transcript in each mutant strain compared to the wild-type level for each transcript, which is set to 1-fold (100%). Each test was performed on two biological replicates. This shows the relative difference between wild-type and the mutants for each transcript. All transcripts except *puckered* in $pcm^{14}$ show a <2-fold mean difference from wild-type, however, none of the differences are statistically significant. With the tests performed on two biological replicates, the levels of the JNK transcripts tested (*puc*, *bsk* and *tak1*) do not differ significantly from their levels in wild-type in L3 larvae for any of the *pcm* mutants tested ($pcm^5$, $pcm^{14}$ and $pcm^3$). One or more statistically significant differences may become apparent if the number of biological replicates is increased, however these differences will likely be small (<2-fold), so their actual biological significance would be debateable.

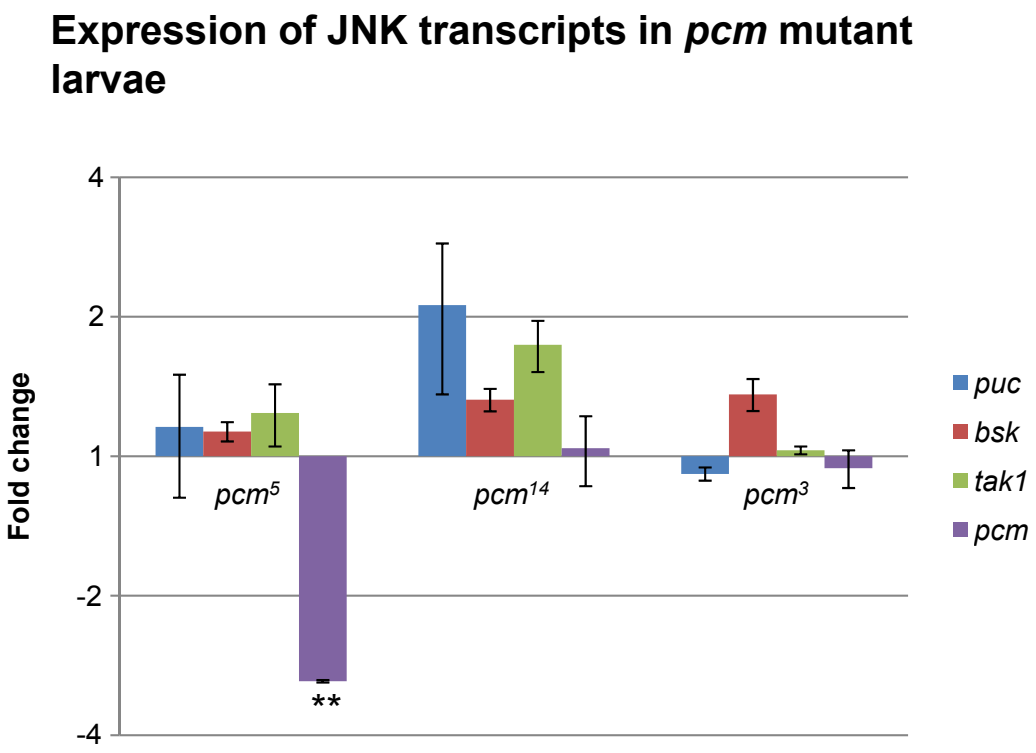



**Figure 5.2**: qPCRs to compare the mRNA levels of members of the JNK pathway between wild-type and *pcm* mutants show no significant differences. Performed on the L3 larvae of wild-type (50E, set to 1) and three *pacman* mutants ($pcm^5$, $pcm^{14}$ and $pcm^3$). *pcm* mRNA in $pcm^5$ mutants was the only transcript found to be significantly different than wild-type.

Error bars show standard deviations between biological replicates, n=2, ** = p-value <0.001. 9 technical replicates were performed per biological replicate.

The level of *pacman* mRNA in each of the alleles was also compared to wild-type. $pcm^5$ mRNA was 3-fold lower than the wild-type *pcm* mRNA ($p<0.001$). The $pcm^{14}$ and $pcm^3$ transcript levels did not differ significantly from wild-type. The most likely explanation for this is the $pcm^5$ transcript may be undergoing NMD due to the nature of the deletion in this allele.

There are two possibilities why Pacman may seem to not affect the JNK pathway. A. The null hypothesis is true (*pacman* mutations do not affect the JNK pathway), or B. The effect is limited to specific tissues, perhaps wing imaginal discs, and using whole larvae 'dilutes out' the effect.

### 5.3.2 Wing imaginal discs

To test whether Pacman is affecting the JNK pathway specifically in wing imaginal discs, RNA was extracted from 125 wild-type (50E) and 125 $pcm^5$ wing imaginal discs. $pcm^5$ wing discs were used rather than $pcm^{14}$ discs for ease of collection.

As these tests were performed on only one biological replicate, no statistics can be performed to determine statistical significance of any transcript differences between mutant and wild-type. Figure 5.3 shows the expression levels of *puc*, *bsk*, *tak1* and *pcm* in $pcm^5$ wing discs alongside their levels in whole $pcm^5$ L3 larvae. None of the transcript levels appear to be very different from the levels in whole larvae. Another transcript, *kayak* (*kay*) was also tested on wing imaginal discs. It has not been tested on whole larvae, but does not differ from the wild type level in wing discs.

As no large differences were seen in the expression levels of the JNK transcripts tested between whole larvae or wing imaginal discs with the first wing disc samples, it seemed unlikely that performing a second extraction of wing discs/RNA for this purpose would be an efficient use of time. These results are not conclusive and more biological replicates would need to be performed on both whole larvae and wing discs, but they appear to agree with the null hypothesis that Pacman is not affecting JNK pathway transcripts. For this reason, we decided to move on to a method that would allow for a broad overview of what transcripts change in *pcm* mutants in an effort to find genuine targets of Pacman.

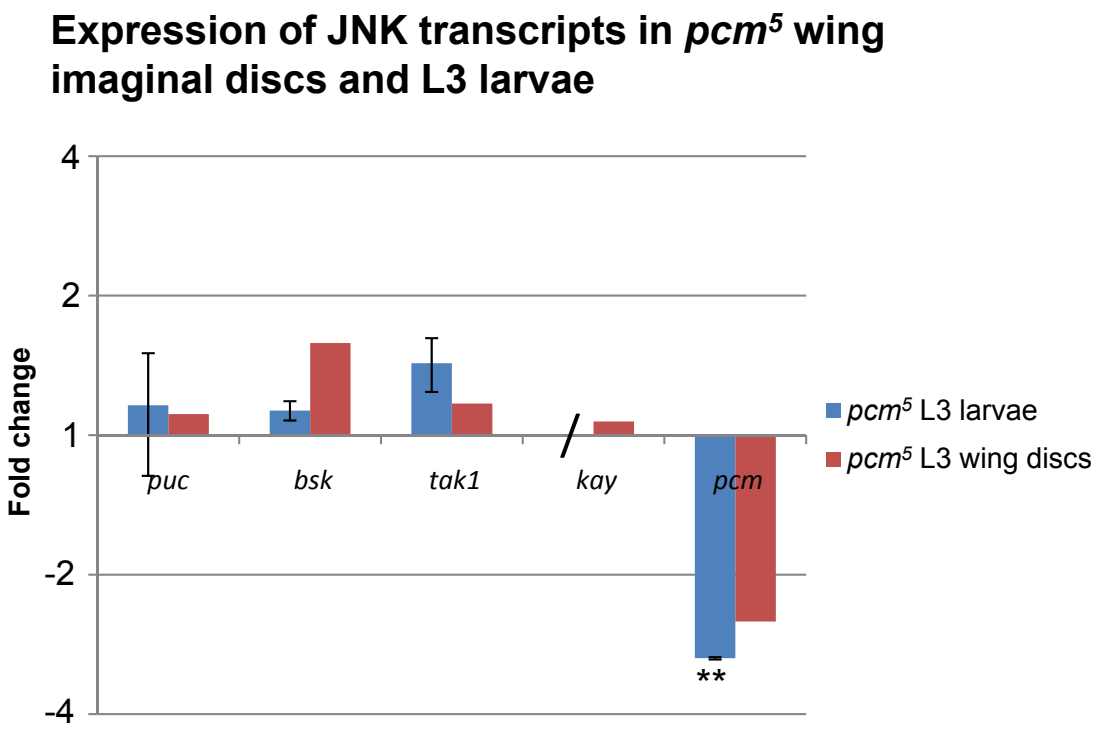



**Figure 5.3**: qPCRs to test the mRNA levels of members of the JNK pathway and *pcm*, in $pcm^5$ wing imaginal discs show no levels greatly different from wild-type. Expression levels in L3 larvae are also shown. The level of *kay* was only tested in wing imaginal discs.
For L3 larvae results, error bars show standard deviations between biological replicates, n=2. ** = p-value <0.001. n=1 for wing imaginal disc data, so standard deviations or p-values can not be calculated. 9 technical replicates were performed per biological replicate.

## 5.4 Global analysis of mRNA expression levels in *pcm*$^5$ wing imaginal discs – Affymetrix 3' *Drosophila* Genome arrays

### 5.4.1 Microarrays

Microarrays are an established technique that can be used to measure expression levels of many thousands of genes at a time. cDNA samples are labelled with a fluorescent reporter and allowed to hybridise to short sequences specific to individual genes that are attached to a chip (such as a glass slide). The level of fluorescence from each spot is proportional to the amount of each RNA present in the original sample (from which the cDNA is made). This allows samples to be compared to test for differences in expression levels.

### 5.4.2 Array details

The Affymetrix GeneChip *Drosophila* Genome 2.0 arrays contain 18,880 11µm spots containing probes for over 18,500 *Drosophila* transcripts. The probes are targeted at the 3' end of each transcript. 2µg of RNA is required for each single colour array to be performed at the Sir Henry Wellcome Functional Genomics Facility (SHWFGF), University of Glasgow.

### 5.4.3 Experimental design

To provide robust results, the SHWFGF recommend performing four biological replicates for each strain. We decided to compare wild-type with *pcm*$^5$ wing imaginal discs, which meant eight arrays in total. 125 wing imaginal discs give 50µl of 75-100ng/µl RNA (3.75-5µg total) and 2µg RNA minimum is required for each array. To ensure enough RNA was collected to allow qPCR to be performed on the remainder of the samples after the arrays, 240 wing imaginal discs were collected for each biological replicate (960 wing discs per strain, 1920 in total).

The wing discs were collected over three months, 30 at a time. The discs were extracted and moved into a 1.5ml tubes containing *Drosophila* Ringer's solution on ice. When 30 had been collected, the majority of the Ringer's was removed from the discs with a Pasteur pipette and they were frozen in liquid nitrogen and stored at -80°C.

Discs were also collected for the miRNA arrays (Chapter 6) at the same time. 180 discs were collected in the same way for each replicate (four wild-type replicates and four *pcm*$^5$). In total, 3360 discs were collected, 1680 wild-type and 1680 *pcm*$^5$. For each strain there were

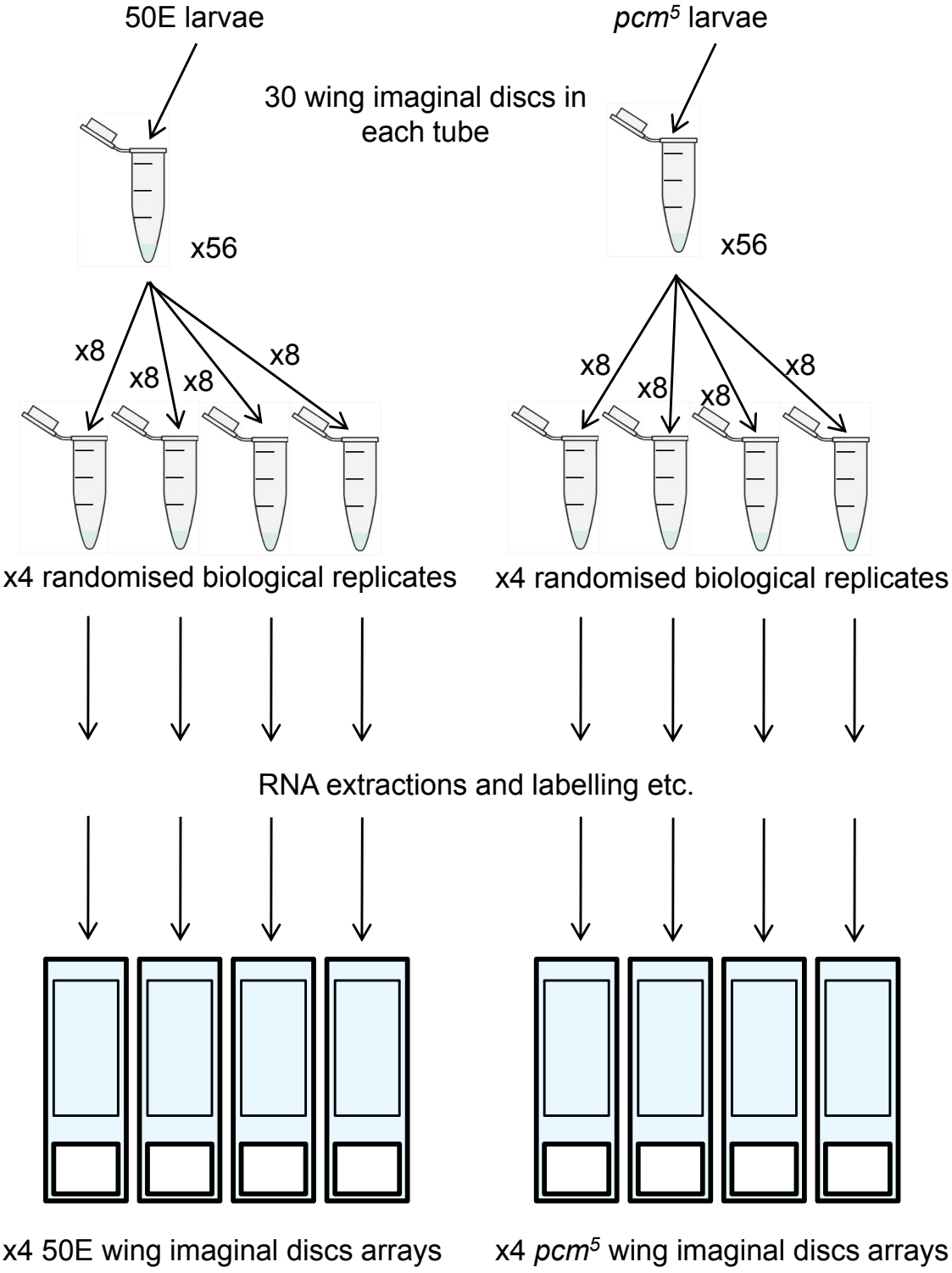



**Figure 5.4**: Design of array experiments to compare mRNA expression levels in wild-type and *pcm*$^5$ wing imaginal discs.
Four biological replicate arrays were performed for each genotype. 1,680 imaginal discs were collected per genotype in 1.5ml tubes containing 30 each. Eight tubes were randomly selected to make up each biological replicate (240 wing discs per replicate). The discs were homogenised and pooled into 1 tube per replicate before being sent to the Sir Henry Wellcome Functional Genomics Facility, University of Glasgow for the RNA extraction, array hybridisation and analysis. Remaining RNA was returned for future experiments.

56 1.5ml tubes containing 30 wing discs. Eight tubes were randomly selected for each biological replicate (to make up four biological replicates for each strain) for use on the mRNA arrays (Figure 5.4).

The RNA extraction and arrays were carried out at the SHWFGF. Each biological replicate was homogenised in RLT buffer containing β-ME from the Qiagen RNeasy Mini Kit, pooled, frozen on liquid nitrogen and sent on dry ice.

### 5.4.4 Quality control

Figure 5.5 shows raw data quality analysis performed using RMAExpress software (Irizarry *et al.*, 2003). Pseudo-images and plots of array intensities showed that each array is of similar quality. No arrays fell outside the main intensity grouping, indicating that no arrays should be discarded from the analysis.

### 5.4.5 Analysis of results (Glasgow)

The SHWFGF microarray service includes analysis of the arrays with their own FunAlyse software. This normalises with RMA and calculates differential expression of genes using their Rank Products method (Breitling *et al.*, 2004b). The genes are annotated and grouped into functional classes using iterative Group Analysis (iGA) (Breitling *et al.*, 2004a). The upregulated and downregulated lists of genes are organised in order of ascending False Discovery Rate (FDR). The FDR gives an estimate of the rate of false positives. If the FDR for a gene at position 50 was 0.05, an estimated 5% (2.5 genes) listed above this position would be false positives. The top 30 up/downregulated genes, ranked by the RPscore calculated using the FunAlyse software, are shown in Figures 5.6 and 5.7 respectively.

The iGA analysis identified numerous genes in the upregulated list as members of stress response groups, including heat, hypoxia and response to unfolded proteins. One group of three genes involved in leucine metabolic process were all in the upregulated list.

The groupings of genes in the downregulated list were more disparate, with more groups represented, but with fewer members from each group affected. Two of the groups included proteins associated with the nucleosome and two groups involved in chromatin assembly were represented.

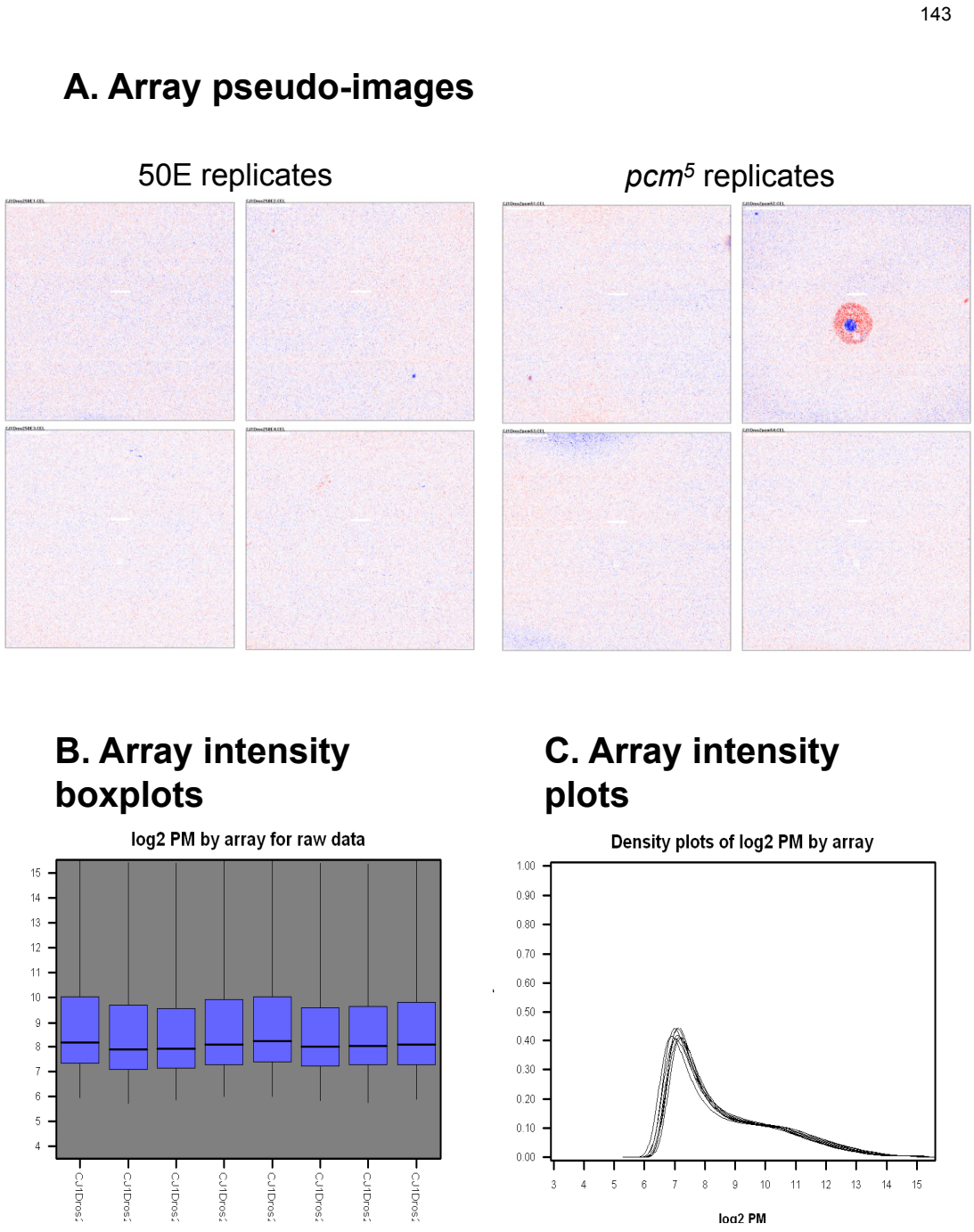



**Figure 5.5**: Array data quality control using RMAExpress software.
**A**. Pseudo-images of the array data generated to qualitatively assess quality. High intensity spots are marked in red and low in blue. Many extreme red or blue spots indicate high noise in the data. The data for each biological replicate show no noticeable bias to either extreme. The only area of low quality is clear in the top right *pcm*$^5$ image.
**B** and **C**. The intensity of the arrays can be compared in two ways. Boxplots or line graphs show the spread of intensities on each chip individually. Good quality data should group closely, as is shown in the plots for these arrays.
Overall, there are no indicators of poor array data for any of the eight arrays. The artefact on one of the *pcm*$^5$ arrays should not have too much effect due to the good data on in the same region of the other *pcm*$^5$ chips.

| Rank | RPscore | FDR | FC_rma | Gene Symbol |
|---|---|---|---|---|
| 1 | 4.62 | 0 | 10.47 | *---///HDC08750* |
| 2 | 8.63 | 0 | 6.36 | *CG31477///CG31477-RA* |
| 3 | 14.64 | 0 | 7.4 | *Hsp70Aa /// Hsp70Ab///CG18743-RA* |
| 4 | 16.12 | 0 | 6.34 | *Hsp70Ba /// Hsp70Bb /// Hsp70Bbb /// Hsp70Bc///CG6489-RA* |
| 5 | 19.32 | 0 | 3.93 | *Hsp67Bc///CG4190-RA* |
| 6 | 20.42 | 0 | 3.63 | *sqz///CG5557-RA* |
| 7 | 23.01 | 0 | 3.39 | *nAcRalpha-30D///CG4128-RA* |
| 8 | 25.25 | 0 | 3.27 | *CG5326///CG5326-RA* |
| 9 | 30.45 | 0 | 2.9 | *Pmi///CG32912-RA* |
| 10 | 32.59 | 0 | 2.77 | *CG5326///CG5326-RB* |
| 11 | 33.89 | 0 | 4.96 | *Hsp70Bc///CG6489-RA* |
| 12 | 35.29 | 0 | 2.7 | *Hsp26///CG4183-RA* |
| 13 | 39.29 | 0 | 2.54 | *CG6357///CG6357-RA* |
| 14 | 39.46 | 0 | 2.76 | *sm///CG9218-RA* |
| 15 | 42.76 | 0 | 2.39 | *CG33099///CG33099-RA* |
| 16 | 45.42 | 0 | 2.29 | *shn///Transposon.24* |
| 17 | 46.82 | 0 | 2.78 | *Hsp22 /// Hsp67Bb///CG32041-RB* |
| 18 | 52.84 | 0 | 2.18 | *Or65a///CG32401-RA* |
| 19 | 52.86 | 0 | 2.2 | *CG5953///CG5953-RA* |
| 20 | 53.03 | 0 | 2.19 | *Ste:CG33236* |
| 21 | 60.72 | 0 | 2.03 | *CG7054///CG7054-RA* |
| 22 | 60.85 | 0 | 2.08 | *scb///CG8095-RB* |
| 23 | 63.72 | 0 | 3.61 | *CG11854///CG11854-RA* |
| 24 | 66.33 | 0 | 1.97 | *CG5687///CG5687-RA* |
| 25 | 69.31 | 0 | 2.02 | *Corin///CG2105-RA* |
| 26 | 73.74 | 0 | 2.36 | *CG2444///CG2444-RA* |
| 27 | 75.42 | 0 | 1.92 | *CG32850///Transposon.97* |
| 28 | 80.47 | 0 | 1.86 | *CkIIalpha-i3///CG3217-RA* |
| 29 | 81.23 | 0 | 2.77 | *Ddc///CG10697-RB* |
| 30 | 81.27 | 0 | 2.06 | *CG4576///CG4576-RA* |

**Figure 5.6**: Top 30 mRNAs upregulated in *pcm*$^{5}$ wing imaginal discs as determined at the Sir Henry Wellcome Functional Genomics Facility, University of Glasgow using RMA normalisation and FunAlyse software. List is ordered by descending RankProduct score (RPscore), a measure of result consistency and rank across replicates.

| Rank | RPscore | FDR | FC_rma | Gene Symbol |
|---|---|---|---|---|
| 1 | 3.67 | 0 | -8.89 | *Hsc70-2///CG7756-RA* |
| 2 | 8.52 | 0 | -6.42 | *CG32364///CG32364-RA* |
| 3 | 8.77 | 0 | -5.97 | *CG31157///CG31157-RA* |
| 4 | 10.29 | 0 | -5.17 | *His2B:CG17949 /// His2B:CG33910///CG17949-RA* |
| 5 | 12.11 | 0 | -4.63 | *His2B:CG17949* |
| 6 | 14.97 | 0 | -3.81 | *---///CG3132-RA* |
| 7 | 18.11 | 0 | -3.57 | *---///Transposon.40* |
| 8 | 18.15 | 0 | -3.59 | *CG17669///CG17669-RA* |
| 9 | 20.29 | 0 | -3.14 | *CG7966///CG7966-RA* |
| 10 | 28.81 | 0 | -2.63 | *pinta///CG13848-RA* |
| 11 | 28.85 | 0 | -2.64 | *CG9238///CG9238-RA* |
| 12 | 30.28 | 0 | -2.49 | *CG42330///CG7060-RA* |
| 13 | 31.22 | 0 | -2.43 | *pcm///CG3291-RA* |
| 14 | 31.31 | 0 | -2.42 | *Hrb87F///CG12749-RA* |
| 15 | 33.95 | 0 | -2.31 | *CG34104 /// CG6934 /// Tequila///Transposon.6* |
| 16 | 36.81 | 0 | -2.31 | *CG10581///CG10581-RA* |
| 17 | 37.21 | 0 | -2.49 | *regucalcin///HDC18536* |
| 18 | 37.36 | 0 | -2.29 | *Cpr64Aa///CG15006-RA* |
| 19 | 38.06 | 0 | -2.22 | *---///Transposon.32* |
| 20 | 38.44 | 0 | -2.22 | *CG5281///CG5281-RA* |
| 21 | 40.65 | 0 | -2.14 | *CG13053///CG13053-RA* |
| 22 | 41.33 | 0 | -2.16 | *CG6045///CG6045-RA* |
| 23 | 41.36 | 0 | -2.15 | *Cpr67Fa2///CG18349-RA* |
| 24 | 45.93 | 0 | -2.09 | *CG10089///CG10089-RB* |
| 25 | 49.09 | 0 | -2 | *spd-2///CG17286-RA* |
| 26 | 49.86 | 0 | -2.02 | *---///S.C2R000018* |
| 27 | 52.16 | 0 | -2.01 | *---///Transposon.38* |
| 28 | 54.47 | 0 | -1.96 | *Cyp6d4///CG12800-RA* |
| 29 | 54.63 | 0 | -1.9 | *CG34247///CT32271* |
| 30 | 62.15 | 0 | -1.83 | *---///Transposon.65* |

**Figure 5.7**: Top 30 mRNAs downregulated in *pcm*$^{5}$ wing imaginal discs as determined at the Sir Henry Wellcome Functional Genomics Facility, University of Glasgow using RMA normalisation and FunAlyse software. List is ordered by descending RankProduct score (RPscore), a measure of result consistency and rank across replicates.

### 5.4.6 Analysis of results using MAS5, RMA and GCRMA

Previous microarray experiments performed by the SHWFGF for Dr. Maria Zabolotskaya in the Newbury lab were analysed by Dr. Sue Jones (Life Sciences, Sussex University). These arrays were performed on the testes of *pcm*$^{5}$ males compared to wild-type (OregonR). They were analysed using three different normalisation methods – MAS5, RMA and GCRMA. MAS5 is the method developed by Affymetrix to take into account probe data and data from mis-matched probes to eliminate non-specific binding. RMA analysis does not rely on the mis-matched probe data as in some instances this can distort the results if the mis-matched probe is still able to bind the actual target (Irizarry *et al.*, 2003). GCRMA uses the same analysis method as RMA, but also takes into account data from the mis-match probes (Cambon *et al.*, 2007). Lists of differentially expressed genes were created after normalisation with each method and compared to identify genes that appeared consistently. Genes that came up in two or more lists were predicted to be the most likely to be due to a real biological difference, rather than false positives. To ensure the two datasets (testes and wing imaginal discs) would be directly comparable, the same method was also used to generate a list of differentially expressed genes for the wing imaginal disc arrays. An important assumption made when using these methods of normalisation is that there is no wholesale increase in mRNA levels. As Pcm degrades mRNAs, this is a possibility, however analysis of the raw data did not suggest any such large scale increase in mRNAs in *pcm*$^{5}$ mutants. This can be seen in Figure 5.5, panels B and C, where the raw data plots for all of the arrays group together, rather than forming two distinct groups of wild-type and *pcm*$^{5}$ arrays. Also, the approach employed here has worked previously and identified numerous differentially expressed mRNAs in *pcm*$^{5}$ mutant testes. The genes shown to be upregulated in *pcm*$^{5}$ wing imaginal discs that appear in two or more lists of differentially expressed genes after application of the three normalisation methods are shown in Figure 5.8. The equivalent downregulated genes are shown in Figure 5.9.

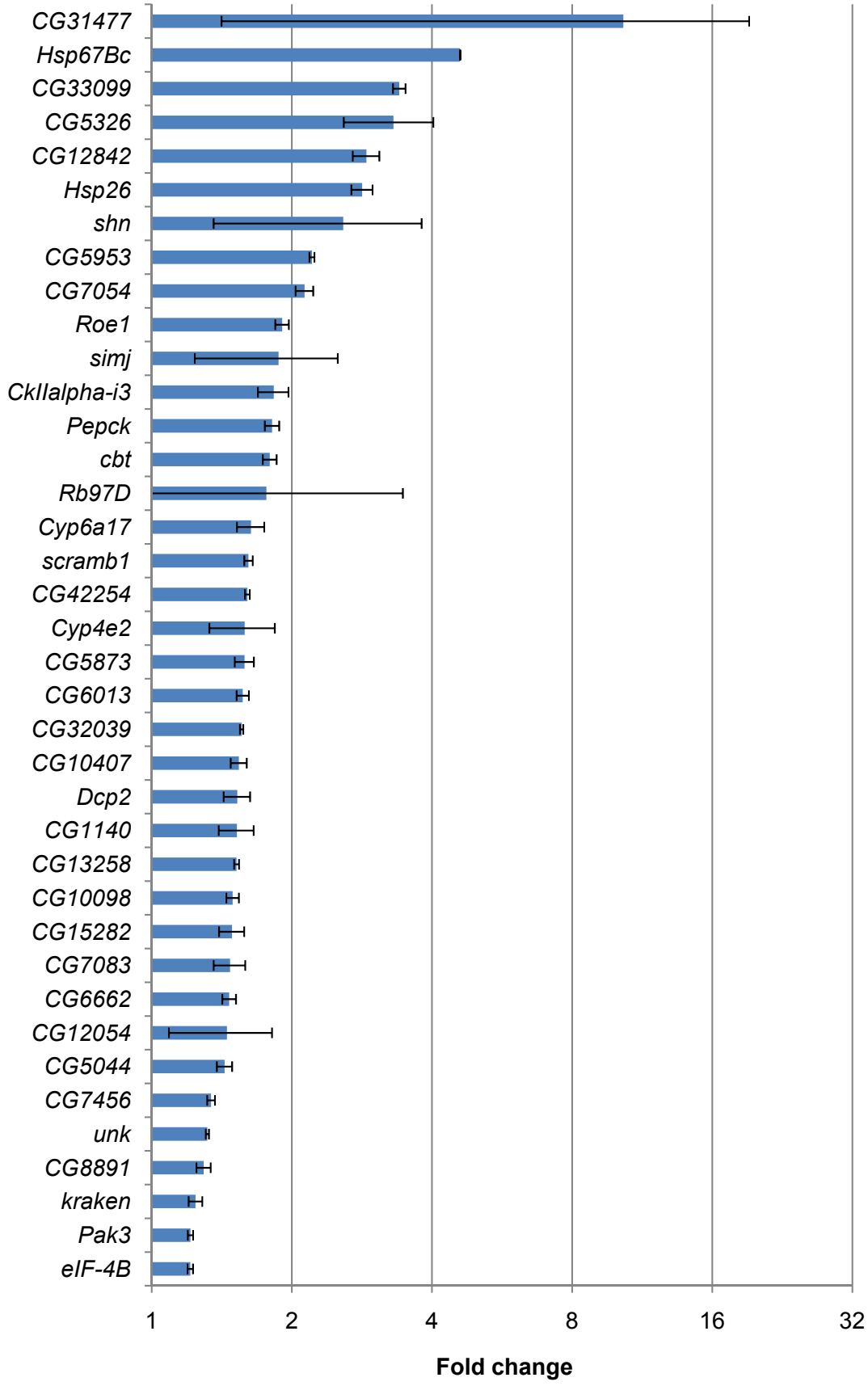



**Figure 5.8**: Upregulated mRNAs in *pcm*$^{5}$ wing imaginal discs as determined by analysis comparing results obtained after normalisation by RMA, GCRMA and MAS5 methods. Only mRNAs consistent between at least two normalisation methods shown. Error bars show standard deviations between mean expression levels given by different normalisation methods.

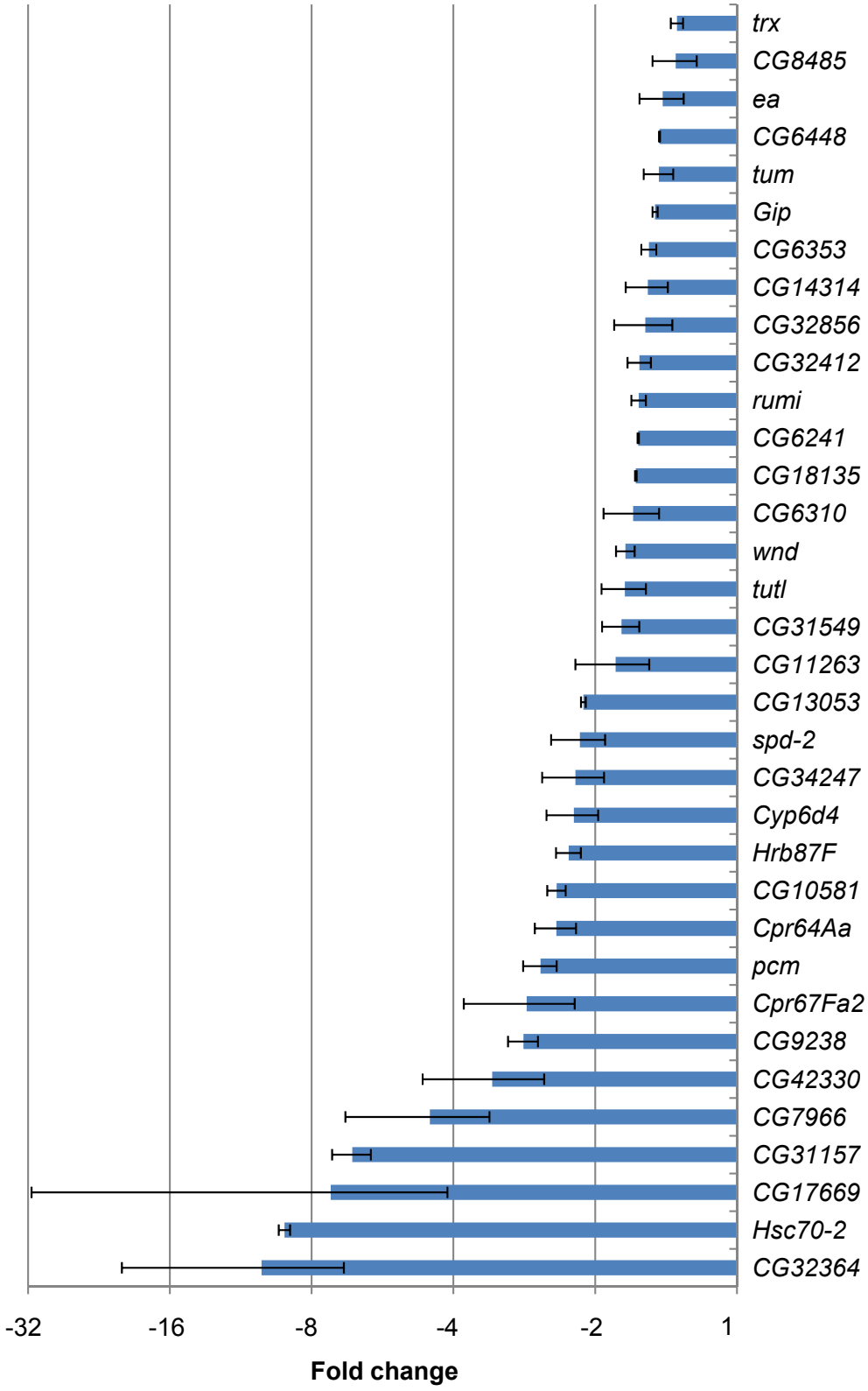



**Figure 5.9**: Downregulated mRNAs in *pcm*$^{5}$ wing imaginal discs as determined by analysis comparing results obtained after normalisation by RMA, GCRMA and MAS5 methods. Only mRNAs consistent between at least two normalisation methods shown. Error bars show standard deviations between mean expression levels given by different normalisation methods.

### 5.4.7 Comparison between arrays on *pcm*$^5$ adult testes and *pcm*$^5$ wing imaginal discs

As the testes and wing imaginal discs arrays were performed and analysed in the same way, they can be compared to see whether *pacman* mutations affect the same mRNAs in separate tissues. Overall, more genes were up or downregulated in *pcm*$^5$ testes than wing imaginal discs, and to a greater extent. 117 genes were upregulated >2-fold in *pcm*$^5$ testes compared to wild-type, whereas only 14 genes were upregulated >2-fold in *pcm*$^5$ wing imaginal discs compared to wild-type. The largest upregulation in testes was 185-fold whereas the largest in wing discs was 10.3-fold. No genes appeared in both upregulated lists for testes and wing discs (including genes with <2-fold difference).

91 genes were downregulated in *pcm*$^5$ testes and 18 were downregulated in *pcm*$^5$ wing discs. The largest downregulation was 74-fold in testes and 10.2-fold in wing discs. Three genes appeared in both downregulation lists for testes and wing discs (including genes with <2-fold difference): *pacman*, *spd-2* and *Hrb87F*. *pacman* has previously been tested in whole larvae and wing discs using qPCR and is known to be downregulated by 3-fold in each. The fold changes in testes and wing discs for *pcm* on the arrays were -2.68-fold and -2.60-fold respectively. The fold changes for the other two genes were also similar on each array: -3.27-fold (testes) and -2.16-fold (wing discs) for *spd-2* and -2.01-fold (testes) and -2.28-fold (wing discs) for *Hrb67F*.

## 5.5 qPCR verification of array results

It is known that the fold changes given by microarrays do not necessarily represent the exact fold-difference in reality, and they tend towards an under-estimation of the difference, especially when the difference is large (Cope *et al.*, 2004). The fold changes can also vary greatly depending on the normalisation method used (see the error bars in Figures 5.8 and 5.9). For these reasons, we decided to consider the fold changes given by the arrays as a rough estimate of the real difference, and to determine the real difference more precisely for interesting genes using TaqMan qPCR. TaqMan qPCR is very specific, has a wide dynamic range and the gene expression levels calculated are typically very consistent between technical and biological replicates. It is also good practice to measure gene expression changes using entirely different methods, to ensure no results are artefacts of one method.

### 5.5.1 qPCR experimental design

The qPCR was initially performed on the same mRNA samples used for the arrays (four wild-type samples and four mutant samples). Each biological replicate is converted to cDNA in duplicate and the qPCR itself is performed in duplicate on each cDNA preparation. This means there are four technical replicates per mRNA sample/probe combination. The TaqMan protocol from Applied Biosystems recommends preparing cDNA in triplicate and performing the qPCR in triplicate on each cDNA preparation (nine technical replicates in total per mRNA/probe combination). However, experience has shown that variation between technical replicates is extremely small, much lower than the variation between biological replicates. As the mRNA samples were limited and performing cDNA preparations in triplicate uses 50% more mRNA than performing them in duplicate, we decided to prepare cDNA in duplicate, and use it more sparingly by also performing the qPCRs in duplicate. Four technical replicates are enough to ensure that there are no wild variations indicative of a problem with either the reverse transcriptase reaction or the qPCR itself (theoretically, each technical replicate should give exactly the same result). Technical replicates are important to ensure the quality of qPCR, but from a biological viewpoint the biological replicates are the only replicates of interest as they show if a difference in expression of a gene is consistent and repeatable.

Figure 5.10 shows the experimental setup using a 96-well Stratagene MX3005P Real-Time PCR system. Each mRNA sample/probe combination (e.g. wild-type/*CG31477*) takes up four

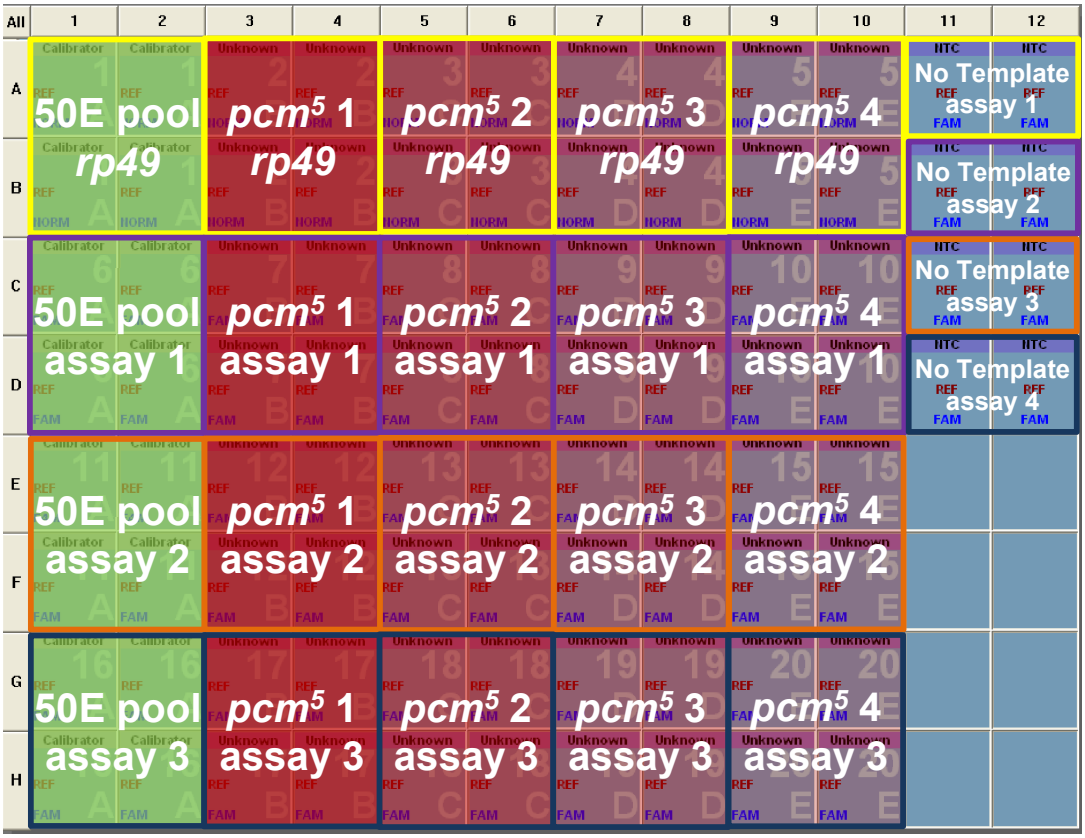



**Figure 5.10**: 96-well plate setup for qPCR verification of array results. The four 50E biological replicates were pooled (green) whereas the four *pcm*$^5$ biological replicates (shades of red) were kept separate. For each RNA sample, two reverse transcriptase replicates were performed, and two qPCRs were performed per cDNA sample, giving four technical replicates per biological mRNA replicate per transcript tested. The normaliser transcript, *rp49* (yellow border), was run on every plate, along with 3 test transcripts (purple, orange and dark blue borders). A no template control (two technical replicates) was run for *rp49* and each test transcript to ensure there was no contamination or fluorescence due to non-specific product amplification or probe degradation.

wells. Each transcript being tested has to be normalised to an internal control before it can be compared between samples. The internal control used is *rp49*, the long-lived, ubiquitously expressed mRNA for a ribosomal protein. For one mRNA sample, normaliser + test transcript takes up eight wells. Each additional mRNA tested increases the number of wells required per mRNA sample by four. On each qPCR plate, three transcripts were tested per mRNA sample, totalling 16 wells including the internal control (4+4+4+4=16 wells).

There were eight mRNA samples (4x wild-type and 4x $pcm^5$). This would require 8x16=128 wells. The maximum number of mRNAs samples that could be run at once to test three different transcripts would be five (5x16=80 wells). To run all eight mRNA samples at the same time, only one transcript could be tested per 96-well plate, which would have greatly increased the amount of mRNA required as the normaliser would have to be repeated on each plate (giving a normaliser:test transcript ratio of 1:1). Performing five mRNA samples with three test transcripts, as calculated above, reduces this ratio to 1:3. We decided to pool the four wild-type mRNA samples into one to give five mRNA samples overall (1x wild-type pool + 4x $pcm^5$). This has the disadvantage of effectively reporting the mean level of each transcript for wild-type but not the standard deviation. The mean and standard deviation can still be calculated for the individual mutant samples. This was a compromise, but without it we would not have been able to verify the level of many transcripts by qPCR, due to the limited amount of mRNA available. Statistics can still be performed on the results.

No template controls were run for each transcript tested, including *rp49*. These control for any level of fluorescence that comes from each primer/probe combination due to non-specific amplification/probe degradation. Two technical replicates were performed for each transcript (4x2=8 wells). The total number of wells in use on each 96-well plate was 88.

### 5.5.2 Transcripts chosen for qPCR verification

Due to cost, time, and availability of mRNA it was not possible to verify the expression level of every transcript that was up or downregulated on the wing disc arrays. Four transcripts were chosen from the upregulated list: *CG31477*, *simjang* (*simj*), *schnurri* (*shn*) and *Dcp2*. *CG31477* was chosen as it was upregulated by the largest amount on the arrays (10.3-fold). *simj* and *shn* were chosen as both are known to be involved in wing disc development and affect transcription of other genes (*simj* as a repressor and *shn* as a transcription factor). *Dcp2* was chosen as it provides an interesting link to Pacman – mRNA transcripts that

undergo 5'-3' degradation by Pacman are first decapped by Dcp1/2. The two genes downregulated by the largest amount were also chosen (*CG32364*, -10.2-fold and *Hsc70-2*, -9.1-fold). Fewer genes were selected from the downregulated list as it was predicted that these were less likely to be direct targets of Pacman. If Pacman normally degrades mRNAs, the level of those it degrades would be expected to increase in a *pcm* mutant.

Applied Biosystems offer numerous TaqMan primer/probe kits for most genes. They are designed across exon boundaries to ensure only mature mRNA transcripts are detected. Some genes can produce a number of possible mRNA transcripts and depending on the exons boundaries crossed, not all TaqMan kits will detect all possible mRNA transcripts. For the purposes of the qPCR verification, TaqMan kits were selected that detect all possible transcripts for any gene that is known or predicted to produce multiple transcripts. It may be the case than only a subset or possible transcripts for each gene are actually expressed in imaginal discs or affected in $pcm^5$ mutants, but we did not want to miss any transcripts or waste reagents by selecting TaqMan kits specific to a few or individual mRNA variants.

### 5.5.3 qPCR results

qPCR gave good results (consistent technical and biological replicates) for all of the transcripts tested, except for *Hsc70-2* (Figure 5.11). *Hsc70-2* showed large variations in level between biological and technical replicates. It seems likely this is due to qPCR being inaccurate at the low level of expression of *Hsc70-2*. Fluorescence for *Hsc70-2* was not detected until after cycle 37 for any $pcm^5$ sample or the wild-type pool and was not detected at all for some technical replicates. The variation between the highest and lowest signals detected was around five cycles (37 min, 42 max, both $pcm^5$ samples). This equates to a difference of 32-fold between two $pcm^5$ samples. For comparison, *rp49* on the same plate is first detected at cycle 18 and the variation between the maximum and minimum is much less than 0.5 cycles (<<1.5-fold difference). Fluorescence for all other transcripts tested was detected before cycle 30 (but always later than *rp49*) and the variation between biological replicates was typically <1 cycle (2-fold). This is the second time we have observed inaccuracy in qPCR when testing a transcript that is expressed at low levels and appears to be a limitation of the method.

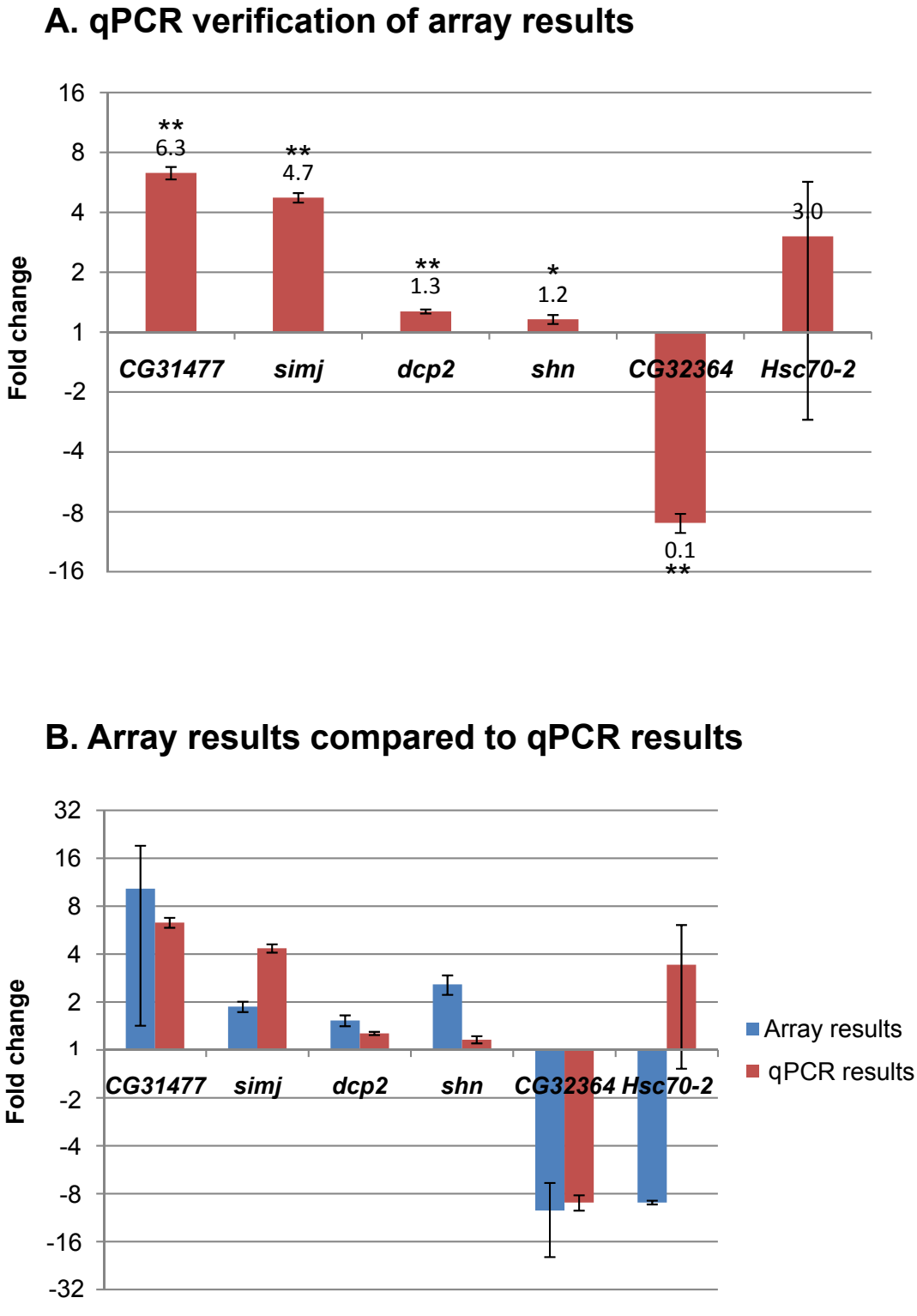



**Figure 5.11**: Verification of expression levels of transcripts selected from array results by qPCR and comparison to array results.
**A.** Expression level of 6 transcripts in *pcm*$^{5}$ wing discs compared to wild-type, determined by qPCR. *Hsc70-2* was expressed at a too low a level for qPCR to give an accurate result. Error bars show standard deviations between biological replicates, n=4. * = p-value <0.05, ** = p-value <0.001. 4 technical replicates were performed per biological replicate.
**B.** Comparison between expression levels determined by arrays and qPCR.
For array data, error bars show standard deviations between results obtained using different normalisation methods. For qPCR data, error bars show standard deviations between biological replicates.

During analysis of comparative quantitative PCR, the wild-type level of each test transcript is set to 1 (100% expression). The level in the mutant is reported as a relative factor of the wild-type level, e.g. 0.5 (50% or -2-fold expression) or 2.0 (200% or +2-fold expression). The relative expression levels for *CG31477*, *simj*, *shn*, *Dcp2* and *CG32364* are shown in Figure 5.11 panel A. p-values were calculated using a one sample *t*-test which compares the observed expression level of transcript with the hypothetical mean of 1 (the level in wild-type). All transcripts tested showed a statistically significant difference from the expected/wild-type level ($p<0.05$).

The qPCR levels and array levels for each transcript are roughly comparable (Figure 5.11, panel B). However, *simj* is expressed at a higher level than the arrays suggest and *shn* is upregulated by a lower magnitude when measured by qPCR. Note that the error bars on the array/qPCR data bars are not comparable. The error bars on the array data show the standard deviation between the results given by the different array normalisation methods. The error bars on the qPCR data show the standard deviation between biological replicates.

## 5.6 Transcriptional vs. post-transcriptional effects on mRNA transcripts in *pcm*$^5$ wing imaginal discs

The array and TaqMan qPCR data show transcripts that change in level in *pcm*$^5$ mutants. However, as both methods only measure the level of mature mRNA, they cannot provide information about which transcripts are altered in level due to being direct targets of Pacman, or which transcripts vary in level as a secondary consequence of the *pacman* mutation. Primary targets of Pacman would be the subjects of post-transcriptional regulation, whereas secondary targets would most likely change in level due to altered transcriptional regulation caused by the change in level of the primary targets. It seems most likely that direct targets of Pacman would be upregulated in a *pacman* mutant, as Pacman is an exoribonuclease, so by its nature is not capable of interacting directly with mRNAs to cause an increase in their level.

### 5.6.1 Primer design to detect pre-mRNAs

Applied Biosystems offer pre-designed TaqMan assays against mature mRNAs, but custom assays can be designed against submitted sequences. Three custom assays were designed for the pre-mRNA of *CG31477*, *simj* and *CG32364*. The sequence submitted for each was 100bp spanning an exon/intron boundary (Figure 5.12). This ensures that the primers

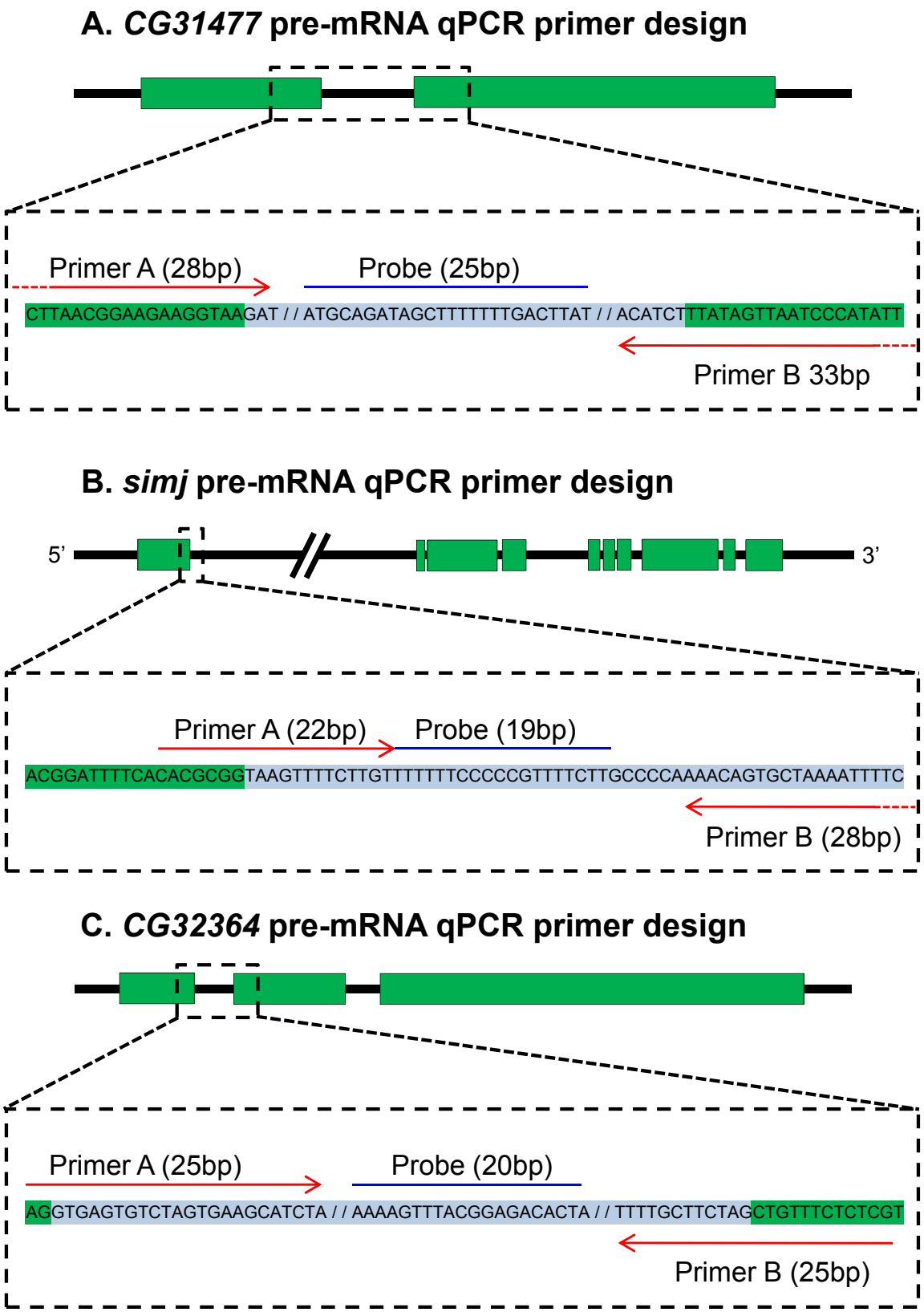



**Figure 5.12**: Method used to design qPCR assays against pre-mRNAs.
**A-C**. Locations of primers and probes for pre-mRNA assays against *pre-CG31477*, *pre-simj* and *pre-CG32364*. 100bp sections covering exon (green)/intron (black/grey) boundaries (expanded region) were submitted to the Applied Biosystems TaqMan custom assay design tool. The probe and primer locations determined by the software are shown.

cannot detect mature mRNA as the introns have been removed. It also ensures that the introns that have been spliced out are not detected.

### 5.6.2 pre-mRNA levels compared to mRNA levels

qPCR experiments were performed alongside repeats of the mature mRNA assays. mRNA from the arrays was used where possible until it ran out. Four more biological samples were prepared (2x 240 wild-type and 2x *pcm*$^5$ wing discs). These four were run on the same plate and the wild-type samples were not pooled.

Figure 5.13, panels A, C and D, show the pre-mRNA compared to mature mRNA level of *CG31477*, *simj* and *CG32364* in *pcm*$^5$ wing discs. *simj* is a clear example of post-transcriptional regulation as the wild-type level of *simj* pre-mRNA is exactly the same as in the *pcm*$^5$ mutant wing discs, whereas the mature mRNA level is significantly (3.8-fold) higher than wild-type. Note that the mean fold change for *simj* is 3.8-fold after the data from the new biological replicates is taken into account (n is now 6). *CG32364* is a clear example of transcriptional regulation as both the pre- and mature *CG32364* mRNAs are significantly lower in *pcm*$^5$ wing discs. For *CG31477*, the pre-mRNA is slightly but statistically significantly higher than the level in wild-type (1.7-fold, p=0.04), however it seems unlikely that an increase in transcription this small would lead to the 6.7-fold increase in the mature *CG31477* mRNA, so the majority of the effect is probably post-transcriptional.

### 5.6.1 Simjang protein level in *pcm*$^5$ wing imaginal discs

An increase in mRNA level does not necessarily lead to an increase in protein level. For the level of protein to increase, the rate of translation must increase. If the mRNA level increases, the rate of translation will only increase if the amount of mRNA is a limiting factor. There must also be no increase in the rate of degradation of the protein, for example triggered by a negative feedback mechanism.

As Simjang has previously been studied in *D. melanogaster*, antibodies against it are available. Rat anti-Simj was kindly sent to us by the Nusse lab, Stanford University School of Medicine. A western blot was performed on 60 wild-type and 60 *pcm*$^5$ wing imaginal discs to check that an increase in *simj* mRNA leads to an increase in Simj protein. The membrane was probed with rat anti-Simj and mouse anti-Tubulin. The bands were quantified in ImageJ and Tubulin was used to normalise the intensity of the Simjang band. The Simjang band was roughly 3-fold brighter for the *pcm*$^5$ wing disc sample (Figure 5.13, panel B).

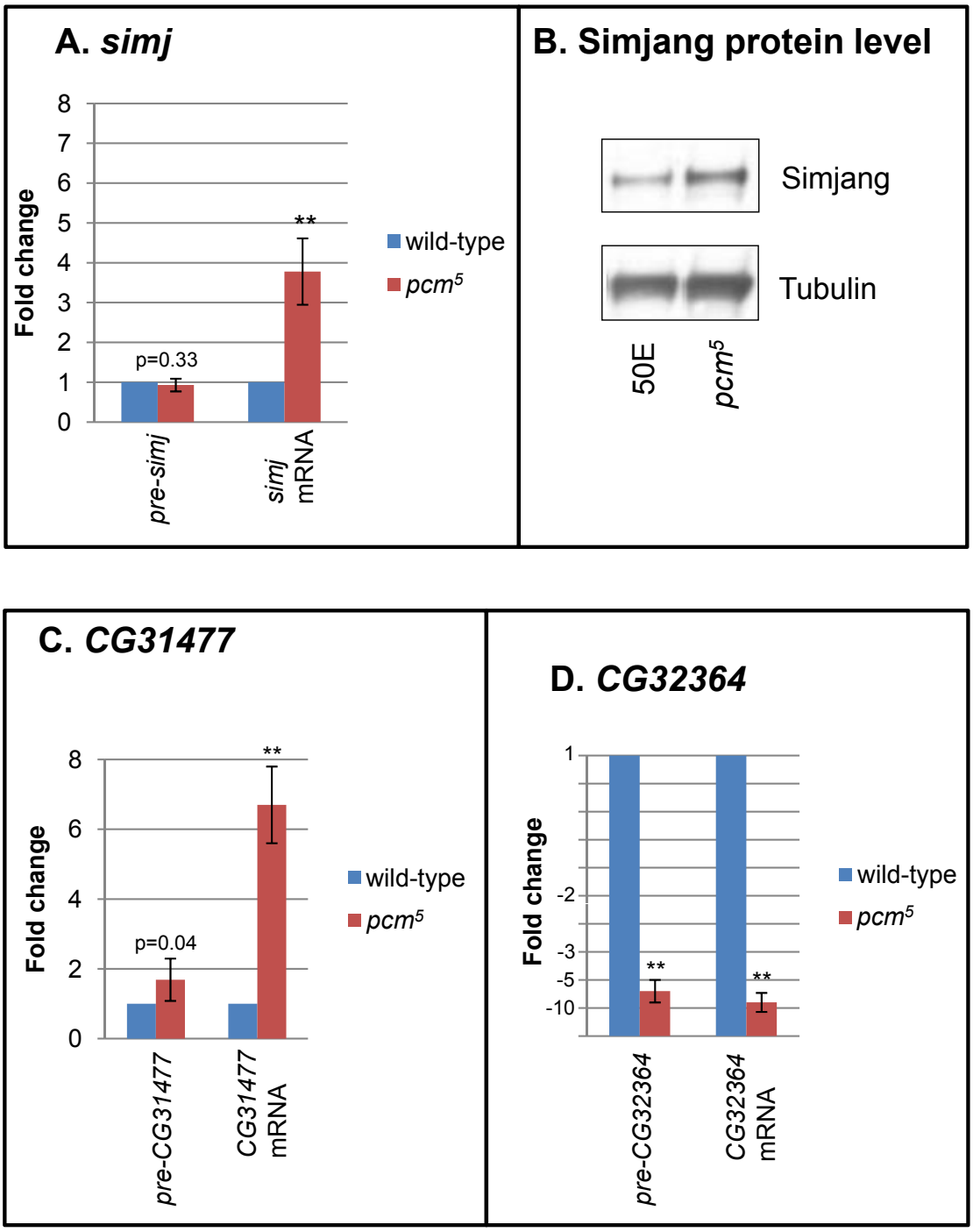



**Figure 5.13**: Change in pre-mRNA expression levels compared to the fold change of the mature mRNA, and expression of Simj protein, in *pcm*$^5$ mutant wing discs.
The large difference between pre- and mature mRNA levels for *simj* suggests it is a direct post-transcriptional target of Pacman (**A**). The increase in *simj* mRNA level translates into an increase in the level of Simj protein of roughly 3-fold (**B**), using rat anti-Simj N-terminal antibody kindly provided by the Nusse lab, University of Stanford.
The large difference between pre- and mature mRNA levels for *CG13477* (**C**) suggests it may also be a direct post-transcriptional targets of Pacman.
The small difference between pre- and mature *CG32364* levels (**D**) suggests *CG32364* is regulated transcriptionally and is not a direct target of Pacman.
Error bars show standard deviations between biological replicates, n=2 for *pre-CG32364* and n=6 for all others. p-values show the significance of the difference between levels in mutant/wild-type. ** = p-value <0.001. 4 technical replicates were performed per biological replicate.

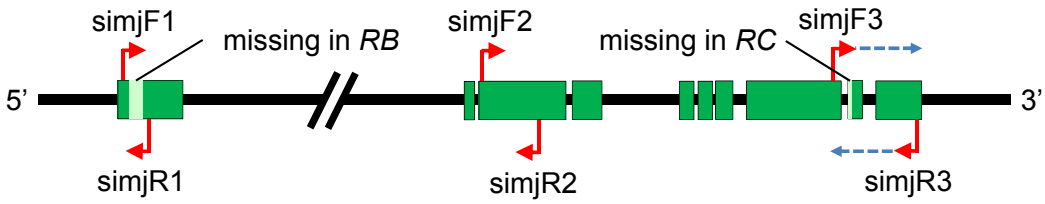
## 5.7 Further analysis of the effect of Pacman on Simjang

### 5.7.1 Determination of *simjang* transcript variant

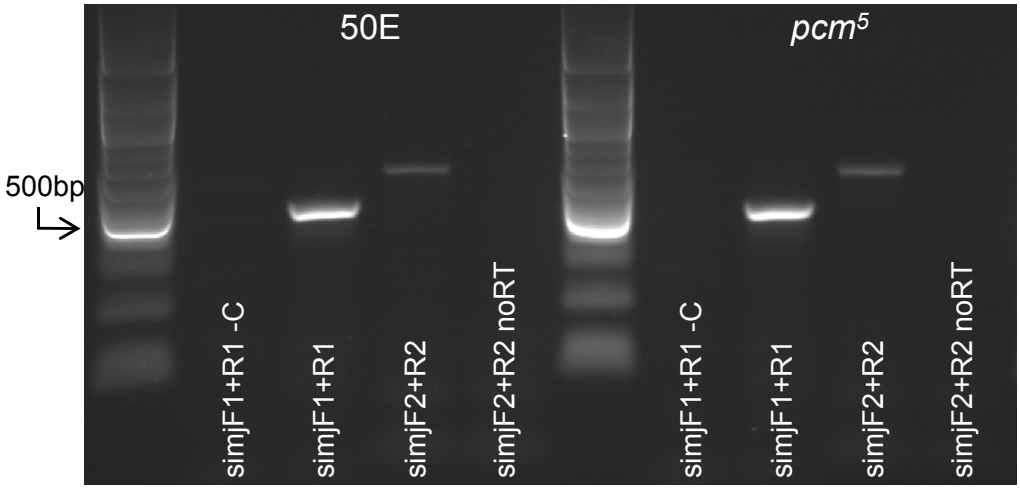
The *simj* gene can produce three *simj* mRNA variants (*RA*, *RB* and *RC*). There are two differences between the three transcripts. *simj-RA* is the longest transcript at 3,689bp. *simj-RB* is identical to RA except for 241bp missing from the 5' UTR (total 3,448bp). *simj-RC* is identical to *RA* except for 12bp missing from the penultimate exon (total 3,677bp). The array and qPCR assays used to measure *simj* expression detect any transcript variant.

To determine which *simj* mRNA variants are expressed in wing discs (both wild-type and $pcm^5$ discs), primers were designed to amplify the region including the gap in RB (simjF1 and simjR1), and to amplify the region covering the 12bp gap in RC (simjF3 and simjR3) (Figure 5.14 panel A). A pair of primers was also designed to detect any *simj* variant (simjF2 and simjR2) for use as a positive control. The products produced from PCR with the simjF1 and simjR1 were visualised to determine if variant *RB* was present. *RA*/*RC* give a band of 400bp with these primers and *RB* would give a band of 159bp. Only a band of 400bp was seen in either wild-type or $pcm^5$ samples (Figure 5.14 panel B) showing that *RB* is not expressed in wing imaginal discs. To determine if *RC* or *RA* is present, the simjF3 and simjR3 primers were used to amplify and sequence the region covering the 12 gap in RC. If only one of the two transcripts was present, the sequence would be clean and could be identified by whether it contains the 12bp missing in *RC*. If both transcripts were present, the sequence would be clear up until the point containing the extra 12bp in *RA* and would be unreadable after that point. The sequence was clear and included the extra 12bp for both the wild-type and mutant samples (Figure 5.14 panel C), indicating that only *simj-RA* is expressed in wing imaginal discs.

## A. *simj* mRNA primers




## B. Test for *RB*




## C. Test for *RC*

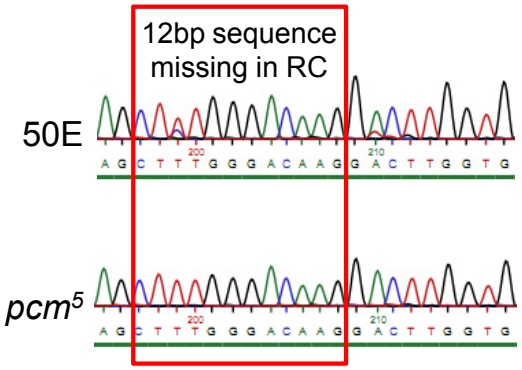



**Figure 5.14**: Determination of *simj* transcript variant expressed wing imaginal discs.
**A**. Location of *simj* primers. Exons shown in green, deviations from transcript *RA* shown in light green. simF1 and simjR1 were designed to distinguish between *RB* and *RA*/*RC*. simjF3 and simjR3 were designed to distinguish between *RC* and *RA*/*RB*. simjF2 and simjR2 were designed for use as positive controls as will pick up any *simj* variant.
**B**. cDNA was prepared from mRNA (from wild-type and *pcm*$^5$ mutant wing discs after DNase treatment to remove genomic DNA) and tested to detect presence of *RB* and/or *RA*/*RC*. simjF1 and simjR1 primers would give a 159bp band if *RB* was present. Only the 400bp band expected for *RA*/*RC* is seen. The simjF2 and simjR2 noRT control shows no gDNA is present.
**C.** Test for *RC* performed on the same mRNA samples. If two variants are present the sequence would be unreadable from the start of the 12bp section that is only present in *RA*. The sequence is clear and corresponds to the *RA* variant, showing that this is the only variant expressed in wild-type or *pcm*$^5$ mutant wing discs

## 5.8 Chapter summary

### 5.8.1 mRNAs that change in level in *pcm*$^5$ wing imaginal discs.

The array experiments performed on wild-type and *pcm*$^5$ wing imaginal discs identified 32 genes (14 up, 18 down) that are up or downregulated by >2-fold when analysed using RMA, GCRMA and MAS5 normalisation methods. The expression levels of another 45 genes (28 up, 17 down) were shown to be different than wild-type by <2-fold. Several genes have had their expression levels determined more accurately using qPCR and three have also had their pre-mRNA level measured to determine if the effect in the mutant wing discs is transcriptional or post-transcriptional. The effect is at the transcriptional level for *CG32364* and at the post-trancriptional level for *CG31477* and *simj*.

#### 5.8.1.1 *simjang*

The level of mature *simj* mRNA is increased by 3.8-fold in *pcm*$^5$ wing imaginal discs. Of the three transcript variants annotated for the *simj* gene, only the largest (*simj-RA*) is expressed in wild-type and *pcm*$^5$ mutant wing discs. The level of *simj* pre-mRNA is not significantly different from the wild-type level. This indicates that *simjang* is regulated at the post-transcriptional level by Pacman in wild-type wing imaginal discs. The increase in *simj* mRNA leads to a 3-fold increase in the Simj protein level.

*simj* does not appear as a gene that changes in level on the arrays performed on adult *D. melanogaster* testes. This could be due to a tissue-specific action of Pacman, or *simj* may not be expressed in testes.

#### 5.8.1.2 *CG31477*

The level of mature *CG31477* mRNA is increased by 6.7-fold in *pcm*$^5$ wing imaginal discs. The level of *CG31477* pre-mRNA is just significantly different from the wild-type level by +1.7-fold. It seems unlikely that the 1.7-fold increase in transcription of *CG31477* can account for the 6.7-fold increase in mature mRNA, unless *CG31477* mRNA is long-lived in *pcm*$^5$ mutants, which may be the case if Pacman is solely responsible for its degradation. It is possible that *CG31477* is also a direct target of Pacman in wild-type wing discs.

*CG31477* does not appear as a gene that changes in level on the arrays performed on adult *D. melanogaster* testes. This could be due to a tissue-specific action of Pacman, or *CG31477* may not be expressed in testes.

#### 5.8.1.3 *CG32364*

The levels of pre- and mature *CG32364* mRNAs are significantly lower in *pcm*$^{5}$ mutant wing discs (-6.3-fold and -8.3-fold respectively). This strongly suggests that *CG32364* is affected at the transcriptional level and is not a direct target of Pacman. *CG32364* may be targeted by one of the direct targets of Pacman, such as Simjang, which is known cause transcriptional repression.

*CG32364* also does not appear as a gene that changes in level on the arrays performed on adult *D. melanogaster* testes. This could be due to a tissue-specific action of Pacman, or *CG32364* may not be expressed in testes.

#### 5.8.1.4 *spd-2* (*spindle defective 2*) and *Hrb87F*

The levels of *spd-2* and *Hrb87F* have not been verified by qPCR in *pcm*$^{5}$ wing imaginal discs or *pcm*$^{5}$ testes, but they are the only genes in common between the arrays two tissues. *spd-2* is downregulated by -3.3-fold in testes and by -2.2-fold in wing discs and *Hrb87F* is downregulated by -2.0-fold in testes and by -2.3-fold in wing discs. As downregulated genes, they are unlikely to be direct targets of Pacman, but their consistent change in levels in disparate tissues may be indicative of a common pathway which is regulated by Pacman.

### 5.8.2 Impact on the JNK hypothesis

One hypothesis tested at the start of this chapter was that Pacman affects the JNK pathway by regulating the mRNA level of one or more members of the pathway, possibly *puckered*. However, qPCR results on JNK pathway transcripts early in the chapter are consistent with the null hypothesis that Pacman does not directly affect the JNK pathway. While the genetic interaction between *pcm* and *puc*, characterised by a bald spot on the thorax, is consistent and unique to *pcm* and *puc* double mutants, it is not strong evidence for a real interaction between Pcm and the JNK pathway without corroborating molecular evidence. It may be that the two mutations independently "sensitise" the wing discs to abnormal development and the presence of the *puc* mutation in a *pcm*$^{5}$ background is enough to cause the bald spot phenotype.

### 5.8.3 *pacman* and *simjang*

If Pacman does not affect the JNK pathway, it must be affecting something else to produce the phenotypes observed. *simjang* appears to be a direct target of Pacman and could be responsible for the observed phenotypes, independent of the JNK pathway. Simjang is well

characterised protein that interacts with the Nucleosome Remodelling and Deacetylase (NuRD) complex. The NuRD complex contains histone deacetylases and is recruited to areas of the genome by transcriptional repressors. The actions of the NuRD complex, phenotypes of *simj* mutants and the implications of the *pcm-simj* interaction are discussed in Chapter 7.

# 6 Chapter 6: Effect of *pacman* mutations on specific microRNAs

## 6.1 Introduction to microRNAs

There are numerous steps that control the level of every protein in every cell of all living organisms, from initial transcription to degradation of the protein. Many opportunities to increase or decrease protein expression exist at the RNA level, as discussed in Chapter 1. Work in this chapter was performed to test the hypothesis that Pacman can affect levels of mRNAs by affecting the levels of miRNAs.

## 6.2 Pacman affecting microRNA levels hypothesis

The overriding hypothesis for this chapter is that *pacman* mutations affect the level of miRNAs in *D. melanogaster*. The work in this chapter was based on a more speculative premise than the others, as no direct links between the mRNA degradation machinery and miRNA expression levels had previously been shown. However, this may be due to the lack of work to elucidate what actually happens to miRNAs. Overwhelmingly, current work on miRNAs focuses on what they do. The biogenesis of miRNAs and their mechanisms of action have been determined in some detail, but degradation of miRNAs has been comparatively neglected. In parallel to the work in this chapter, the question of what happens to miRNAs has begun to be addressed and has been reviewed recently (Kai and Pasquinelli, 2010). It has been shown previously that miRNA levels are dependent in part upon Ago1 levels, presumably because Ago1 confers physical protection to the miRNA bound within (Vaucheret *et al.*, 2004, Diederichs and Haber, 2007). Some evidence now exists showing that miRNAs are degraded by "small RNA degrading nucleases" (SDNs) in the 3'-5' direction in *Arabidopsis* (Ramachandran and Chen, 2008), and in the 5'-3' direction by Xrn-2 (Rat1 in *D. melanogaster*, XRN2 in *H. sapiens*) in *C. elegans* (Chatterjee and Großhans, 2009, Chatterjee *et al.*, 2011). XRN2 is known to be involved in the degradation of aberrant pre-ribosomal RNAs (Wang and Pestov, 2010) and is thought to play a role in transcription termination for both RNA polymerase 1 (El Hage *et al.*, 2008) and RNA polymerase 2 (Luo and Bentley, 2004). However, degradation of miRNAs would be a significant extension of the abilities of XRN2, especially as it is thought to be a nuclear protein (Xiang *et al.*, 2009), while mature miRNAs are cytoplasmic. The mRNA decay machinery, such as XRN1 (Pacman) and the exosome are cytoplasmic, and one paper has recently shown that *miR-382* is degraded mostly by the exosome with a "modest contribution" from XRN1 (Bail *et al.*,

2010). Currently, the fate of miRNAs is unclear, and in accordance with the rest of cellular biology, is unlikely to be straight-forward. The motivation behind the experiments in this chapter was to investigate the possibility that Pacman could play a, perhaps specific, role in miRNA degradation in whole animals. If Pacman does affect miRNA levels, the change should be evident in *pacman* mutants, and may even provide an explanation for one or more of the observed mRNA level changes and phenotypes in *pacman* mutants.

## 6.3 Preliminary TaqMan quantitative RT-PCR analysis of miRNAs predicted to affect the JNK pathway

The hypothesis that *pacman* mutations lead to misexpression of the JNK pathway was outlined in the previous chapter. Preliminary comparative quantitative RT-PCR (qPCR) did not show a difference in the mRNA level of any JNK pathway member tested, which lead us to consider alternative scenarios that could lead to reduction of JNK pathway protein levels, without reduction of the mRNA level. miRNA regulation is one such way that would allow for repression of translation of specific transcripts, as miRNA-mediated regulation does not necessarily lead to transcript degradation. qPCR on mRNAs would not be able to detect this effect, as it gives a measure of transcript level only and not translation.

### 6.3.1 Prediction of JNK pathway miRNAs

There are tens of examples of specific miRNA:mRNA interactions in many organisms, and the number verified is increasing quickly. A number of methods exist to predict miRNA:mRNA interactions *in silico*, such as TargetScanFly (Kheradpour *et al.*, 2007), PicTar (Grün *et al.*, 2005) and Microcosm (Griffiths-Jones *et al.*, 2008). However, thus far it has proved difficult to accurately predict miRNA:mRNA interactions, but as technology improves and more examples are found, prediction algorithms should improve. For now, predicted interactions require experimental verification, and such experiments have shown how important a role miRNAs play during *D. melanogaster* development (Jones and Newbury, 2010). The most recent release of miRBase (Griffiths-Jones *et al.*, 2006) (release 17, April 2011) contains 19,724 miRNA sequences for 153 organisms. This includes 238 for *D. melanogaster* and 1,424 for *H. sapiens*.

Predicting functional miRNA:mRNA interactions is the aim of the prediction methods mentioned above, but currently no system is able to provide much more than a guide as to which miRNAs may affect any given mRNA. As there are no experimentally verified miRNA:JNK pathway interactions, the only option was to attempt to form a consensus

between the prediction methods to produce a list of candidate miRNAs that may affect the JNK pathway. miRNA predictions from TargetScanFly, PicTar-Fly and Microcosm were all considered and a list of potential miRNAs that affect the JNK pathway was generated (Figure 6.1). Levels of miRNA expression in larvae/imaginal discs were also taken into account (Ruby *et al.*, 2007).

Seven miRNAs (*miR-9a*, *miR-31a*, *miR-79*, *miR-92b*, *miR-274*, *miR-289* and *miR-315*), each with predicted binding sites in one or more of *basket*, *puckered* or *kayak* were selected for initial testing. Two others (*miR-303* and *miR-304*) were also selected as they are not predicted to bind to these transcripts. *miR-274* was selected as a negative control, as Ruby *et al*., 2007 showed that it is only expressed in adults and not larvae.

### 6.3.2 Preliminary qPCR on predicted JNK pathway miRs in whole larvae and wing imaginal discs

The initial miRNA qPCRs were performed on RNA extracted from L3 larvae of $pcm^5$ and $pcm^{14}$ mutants and $pcm^5$ wing imaginal discs, compared to wild-type (50E) L3 larvae and wing imaginal discs respectively. The results are shown in Figure 6.2. Two cDNA replicates were performed for each RNA sample and three qPCR replicates were performed per cDNA replicate (six technical replicates total per RNA sample/miRNA combination). Biological replicates were planned for miRNAs that showed a difference in the initial tests.

The majority of the miRNAs tested showed <2-fold difference between the mutants and wild-type. The level of *miR-9a* was 3-fold lower in $pcm^{14}$ whole larvae and *miR-31a* and *miR-315* were just under 2-fold lower. *miR-315* in $pcm^5$ wing imaginal discs appeared to be >10-fold lower than in wild-type wing discs. As this was by far the greatest difference, *miR-315* was investigated further. The levels were tested in three wild-type and six $pcm^5$ biological replicates and found to be extremely inconsistent. The mean level in $pcm^5$ discs was 0.91 (-1.1-fold) overall, with a standard deviation of 0.87 (95% of the mean). The inconsistency was not just between biological replicates, but also between technical replicates. The standard deviations of the technical replicates for each biological replicate of wild-type were 13.9%, 51.6% and 30.9% (average 32.1%), expressed as percentages of the mean. The equivalent standard deviations for the $pcm^5$ replicates were 34.4%, 22.0%, 38.1%, 69.7%, 42.3% and 34.3% (average 40.1%). As a comparison, equivalent technical

## A. Comparison of number of miRs/total sites predicted by each method

| | No. of predicted miRs/total no. of sites | | |
|---|---|---|---|
| mRNA | TargetScan Fly | PicTar | Microcosm |
| *basket* | 1/1 | 0/0 | 4/4 |
| *puckered* | 14/21 | 3/4 | 2/2 |
| *kayak* | 17/18 | 7/7 | 3/3 |
| *tak1* | 6/6 | 0/0 | 3/4 |

## B. miRNAs selected

| | Possible JNK pathway targets | | |
|---|---|---|---|
| miRNA | +++ | ++ | + |
| *miR-9a* | | | ***basket*** |
| *miR-31a* | | ***puckered*** | |
| *miR-79* | | ***puckered*** | |
| *miR-92b* | ***kayak*** | ***puckered*** | |
| *miR-274* | | ***basket*** | |
| *miR-289* | | ***kayak*** | |
| *miR-303* | | | |
| *miR-304* | | | |
| *miR-315* | ***puckered*** | | ***kayak*** |

**Figure 6.1**: miRNAs predicted to affect the JNK pathway.
**A**. The number of individual miRNAs and total number of miRNA sites predicted by TargetScanFly, PicTar and Microcosm for the JNK pathway transcripts shown.
**B**. miRNAs selected for testing with comparative quantitative PCR and the members of the JNK pathway they are predicted to interact with. +++ - + indicate consistency and strength of predicted interactions.

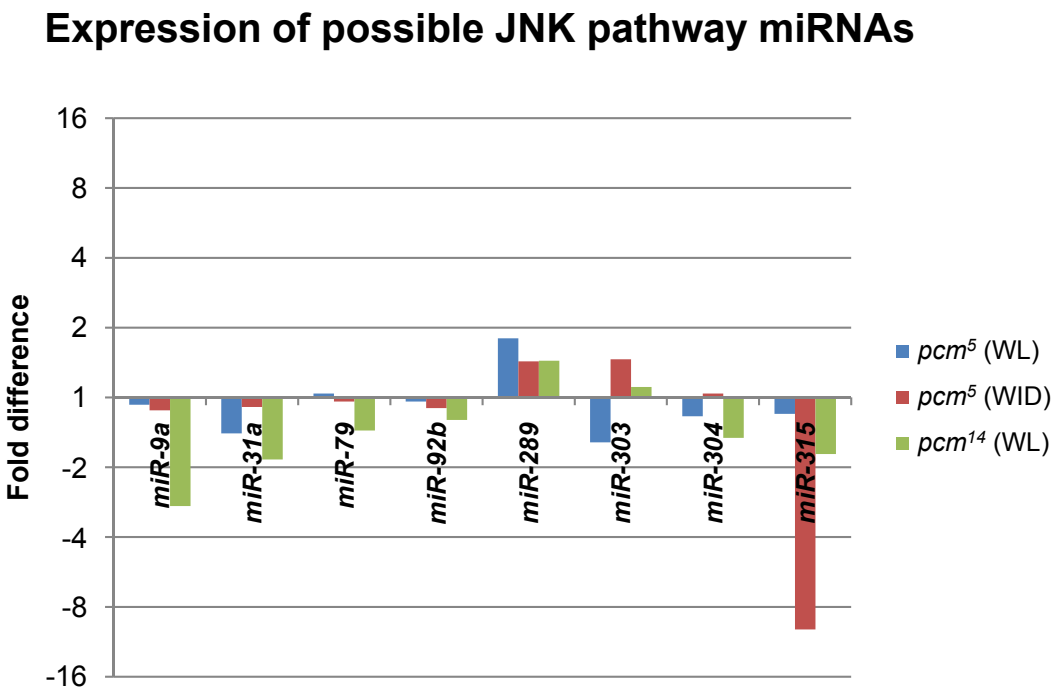



**Figure 6.2**: Expression of predicted JNK pathway miRs in *pcm*$^{5}$ wing imaginal discs (WID) and whole larvae (WL), and *pcm*$^{14}$ whole larvae, compared to wild-type.
*miR-315* and *miR-9a* were the only two miRNAs that appeared to be at a different level from wild-type after the preliminary experiments (n=1 for all samples). 9 technical replicates were performed per miRNA tested.
Further analysis of *miR-315* showed that the level in wild-type and mutant was very low, which lead to inconsistent results.

standard deviations for *miR-34* (taken from a later experiment) were 3.4%, 2.5%, 3.8% and 7.6% (average 4.8%). Such high technical variations are indicative of a qPCR assay that is not working correctly. Amplification of *miR-315* was first detectable in the range of cycles 32-38 (a range of 64-fold), which is much later than when most miRNAs or mRNAs are picked up. This was only the case in wing imaginal discs, as the level of *miR-315* in L3 larvae for mutant or wild-type was considerably higher, coming up at around cycle 26. It seems likely that *miR-315* is not expressed in wing imaginal discs and the amplification observed from the *miR-315* assay is non-specific. There was no non-specific amplification for *miR-274* negative control in any sample, but the specificity of individual assays is likely variable. Overall, only one >2-fold difference was observed in a miRNA expressed at a high enough level to be accurately measured (*miR-9a* in $pcm^{14}$ whole larvae). As *miR-9a* is only weakly predicted to affect the JNK pathway, there is no strong evidence to support the hypothesis that *pacman* is affecting the JNK pathway by miRNAs. It is however still possible that *pacman* could be playing a role in miRNA regulation of miRNAs not tested here.

## 6.4 Preliminary miRNA arrays (human)

As the preliminary experiments on both mRNA and miRNAs failed to produce any data inconsistent with the null hypothesis that Pacman does not affect the JNK pathway, we decided to extend the scope of the search for miRNAs that may be affected in *pcm* mutants. miRNA microarrays are a newer technique than mRNA arrays and suffer from rapid obsolescence due to the rate new miRNAs are being found, but offer a quick method of assaying all known miRNAs at the same time. The lab of Dr. Martin Bushell at Nottingham University (now at the MRC Toxicology Unit, Leicester University) have extensive experience in miRNA microarrays on human samples, and were kind enough to share their arrays and experience at the time we were considering which miRNA array platform to use. We performed five dual-colour miRNA arrays in Nottingham, using their custom arrays with Exiqon probe sets for human/mouse/rat miRNAs. These arrays are not ideal for use with *Drosophila* samples, as the only *Drosophila* miRNAs present are ones also conserved in humans/mice/rats. The arrays performed were comparisons between $pcm^{5}$ wing discs, $pcm^{5}$ whole larvae and $pcm^{14}$ whole larvae and the relevant wild-type (50E) tissue. The arrays showed numerous miRNAs expressed in the *Drosophila* samples, with many apparently at a different level in the mutant samples. As the arrays are not designed for *Drosophila* the results are difficult to interpret, so after learning the technique, we decided to move on to performing experiments on arrays designed for *Drosophila*.

## 6.5 Global analysis of miRNA expression levels – Exiqon "other species" miRNA arrays on *pcm*$^5$ wing imaginal discs

### 6.5.1 Array details

The arrays with the most extensive *Drosophila* miRNA coverage available at the time were Exiqon LNA miRNA v.11.0 "other species" arrays. These arrays contain probe sets designed against miRNAs for organisms that are not human, mouse or rat from miRBase version 11. This includes 152 *D. melanogaster* miRNAs, which at the time was 100% coverage. Each 18mm x 49mm array slide contains 10,032 90µm spots arranged into four replicates of 12 sub-arrays. Each sub array consists of 193 test spots, 10 spike in control spots and five landing light spots arranged in a grid of 15x14 (with one less spot on the last row). This is in contrary to the description given in the manual, which shows a grid of 15x15 spots with the spike in control spots in the wrong place. The GAL file provided by Exiqon containing the annotations of the spot positions fortunately did not contain the errors present in the manual.

### 6.5.2 Experimental design

Figure 6.3 shows the design of the miRNA microarray experiments on *pcm*$^5$ and wild-type wing imaginal discs. The wing discs were collected at the same time as the discs for the mRNA arrays. 180 *pcm*$^5$ and 180 wild-type (50E) wing discs were collected for each array (six groups of 30) as this number would give sufficient total RNA for the arrays (1µg required for each strain per array) and subsequent qPCR to verify the results.

Total RNA was prepared from each sample of 180 wing discs using the Ambion *mir*Vana kit. The samples were not enriched for small RNAs as Exiqon advise that this is unnecessary, and maintaining the larger mRNAs allows them to also be tested later if required. The purity and quantity of the RNA extracted was tested on a NanoDrop 1000 Spectrophotometer and concentrated using an Eppendorf Concentrator Plus. Two arrays were performed at a time as per the instructions from Exiqon, including the use of spike in controls. *pcm*$^5$ RNA was labelled with Hy5 dye (green, 532nm) and 50E RNA was labelled with Hy3 dye (red, 635nm). Hybridisation was performed in Agilent Microarray Hybridisation chambers overnight in a UVP HB-1000 hybridisation oven. After washing and drying of the slides they were kept in

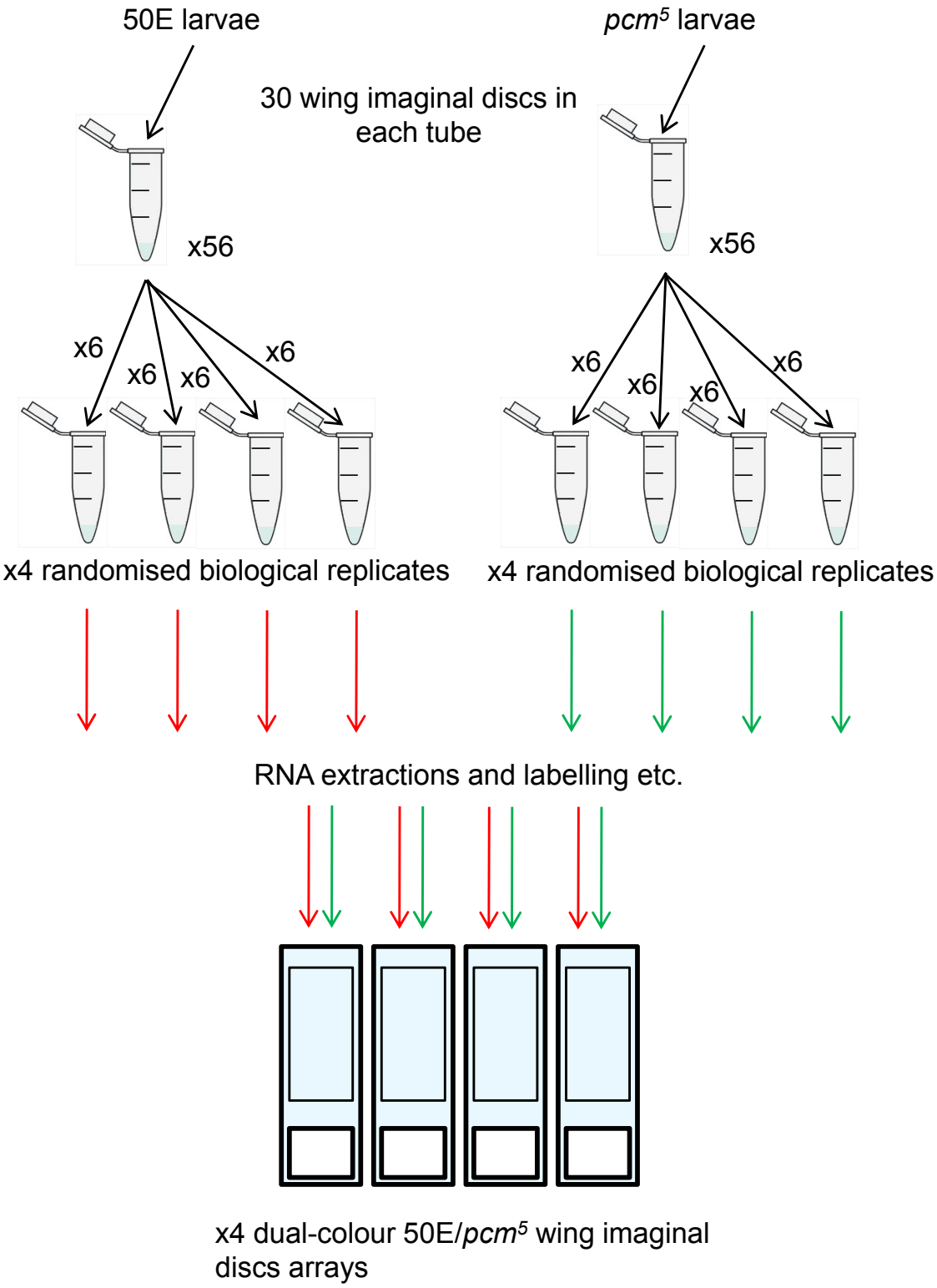



**Figure 6.3**: Design of array experiments to compare miRNA expression levels in wild-type and *pcm*$^5$ wing imaginal discs.
Four biological mutant vs. wild-type replicate arrays were planned. 1,680 imaginal discs were collected per genotype in 1.5ml tubes containing 30 each. Six tubes were randomly selected to make up each biological replicate (180 wing discs per replicate). The total RNA was extracted using the Ambion *mir*Vana kit to preserve small RNAs.

nitrogen before being scanned on an Axon GenePix 4000B Microarray Scanner within 20 minutes.

In total, only two of the four planned 50E vs. *pcm*$^{5}$ wing disc arrays were performed as it seemed unlikely that performing the second two would add any extra information, since potential differentially expressed miRNAs were to be verified with qPCR. Two OregonR vs. *pcm*$^{5}$ testes arrays were also performed for Dr. Maria Zabolotskaya's work on *pcm*$^{5}$ mutant testes. The results of these will not be presented in detail, but a comparison between testes and wing imaginal discs miRNA expression will be made later in the chapter.

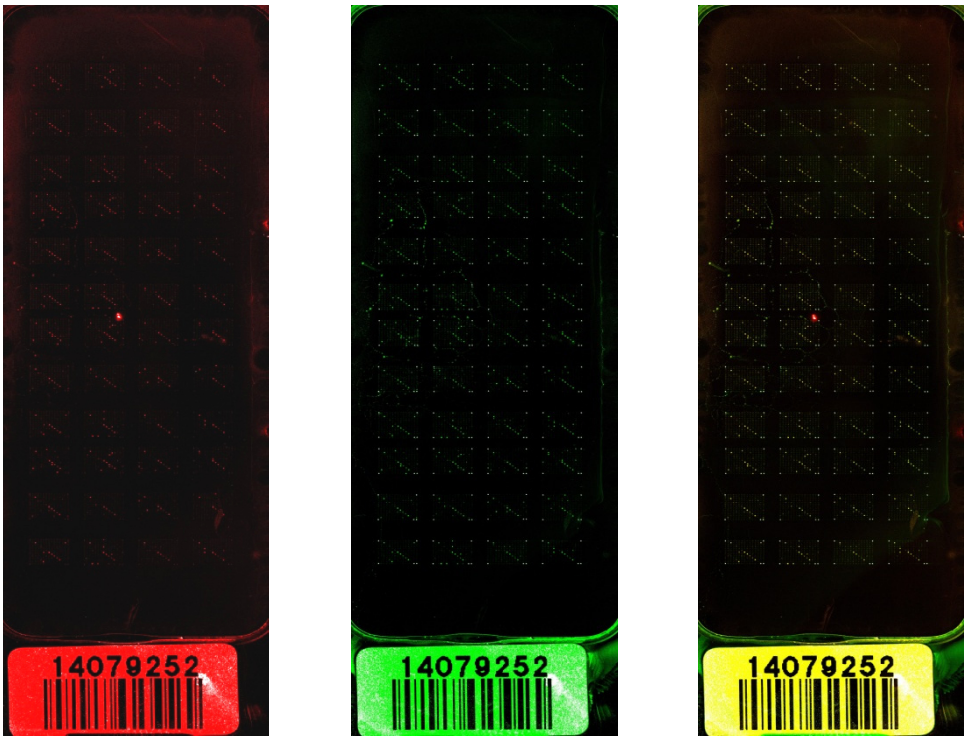
### 6.5.3 Quality control

One benefit of the low density of the miRNA arrays is that there is space for four replicates of each test miRNA, and 48 replicates of each spike in control. As there are 10 spike in controls, there are 480 total spike in control spots spread across the whole array, with each spike in at a different concentration. These spike ins can be checked for consistency to ensure hybridisation across the array is uniform. The arrays can also be checked visually to ensure there are no large anomalies present that may obscure some spots. The low density also allows for each spot to be inspected individually and discounted if it is not of sufficient quality. Figure 6.4 shows images of both arrays with the red and green channels separated and merged. No large biases in hybridisation or artefacts were seen. Few spots were discarded due to low quality. The majority of the spots were blank, which was as expected as only a minority are for *D. melanogaster* miRNAs.

### 6.5.4 miRNA array normalisation

Normalisation of the miRNA arrays is required to compensate for any difference in the level of red and green fluorescence that is not due to a real biological difference. Normalisation is complicated by the fact that there are no established endogenous control miRNAs for *D. melanogaster* wing imaginal discs. These are miRNAs that are present at the same level in the wild-type and mutant samples. As the hypothesis being tested is that Pacman could be affecting the level of any or all miRNAs, no miRNA could initially be used for normalisation. The spike in controls are suitable for normalisation of the arrays, but are not as good as genuine 1:1 ratio miRNAs. The data from the arrays was normalised in GenePix Pro 7 in two ways. Firstly using the spike in controls that were at a level high enough to be detected (spike_control_d, e, g, h, i and j), which account for 284 spots across each array.

## A. 50E vs *pcm*$^5$ miRNA Array 1

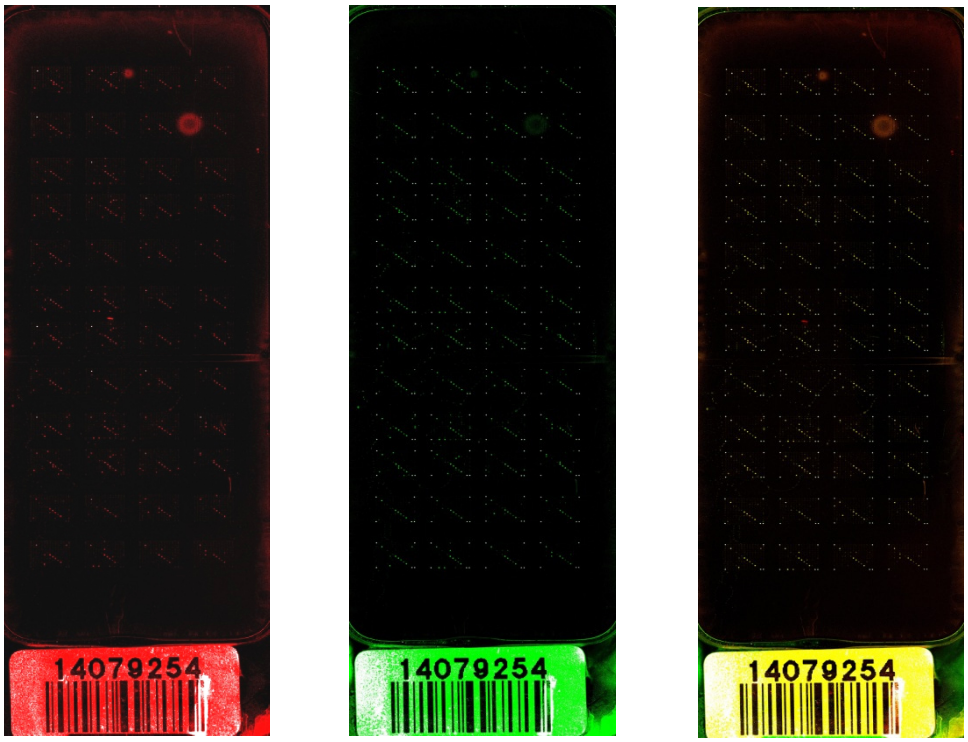



50E
Hy3 - 635nm

*pcm*$^5$
Hy5 - 532nm

Merge

## B. 50E vs *pcm*$^5$ miRNA Array 2




50E
Hy3 - 635nm

*pcm*$^5$
Hy5 - 532nm

Merge

**Figure 6.4**: Red (wild-type) and green (*pcm*$^5$) channel images of both miRNA arrays performed.

miRNAs selected for qPCR verification were selected from the data normalised by spike in controls. Subsequently, three miRNAs (*miR-8*, *miR-79* and *miR-92b*) found to be expressed at a 1:1 ratio by qPCR were used to normalise the array data again. There are four spots per miRNA, so using three as controls means 12 spots per array. More spots would have been desirable. The normalisations using spike ins and miRNA controls did not produce very different results, and miRNA levels given by both methods are presented below.

### 6.5.5 Upregulated miRNAs in *pcm*$^5$ wing imaginal discs

Figure 6.5 shows miRNAs upregulated in *pcm*$^5$ wing imaginal discs compared to wild-type (50E). No miRNA was increased in level by >2-fold after normalisation by either spike ins or 1:1 ratio miRNAs. Four miRNAs (*miR-965*, *let-7*, *miR-13b* and *miR-2a*) were found to be statistically significantly different from the wild-type level for both types of normalisation, but only between 1.6-fold and 1.2-fold. It seems unlikely increases this small would have much biological effect. The lack of increase in the expression of all or specific miRNAs suggests Pacman/XRN1 is not involved in miRNA degradation *in vivo*, at least in *pcm*$^5$ mutant wing imaginal discs. The alternative is that Pacman is redundantly involved in miRNA degradation, or the *pcm*$^5$ allele is sufficient to perform normal Pacman activity in regards to miRNA degradation, but not mRNA degradation (as shown by the previous chapter).

### 6.5.6 Downregulated miRNAs in *pcm*$^5$ wing imaginal discs

Figure 6.6 shows miRNAs downregulated in *pcm*$^5$ wing imaginal discs compared to wild-type. Two miRNAs (*miR-277* and *miR-278*) are downregulated by >2-fold, but only *miR-277* (-3.2-fold) shows statistical significance for both normalisation methods. The average for *miR-278* is 3-fold downregulation, but the result is not statistically significant due to the low number of good quality spots for *miR-278*. Only two out of the eight spots on the two arrays showed fluorescence above background and it is possible that this result is due to chance. Five other miRNAs show significant downregulation for both normalisation methods (*miR-34*, *miR-33*, *miR-7*, *miR-276** and *miR-317*) with fold changes between -1.9-fold and -1.16-fold. It seems one or more miRNAs are downregulated in *pcm*$^5$ wing imaginal discs, but this is unlikely to be a direct effect of Pacman. The most likely explanation for the reduction of a specific miRNA is a transcriptional effect due to an effect Pacman is having on something else. As no global downregulation of miRNAs is seen, it is unlikely Pacman is affecting the mRNAs in the miRNA biogenesis pathway.

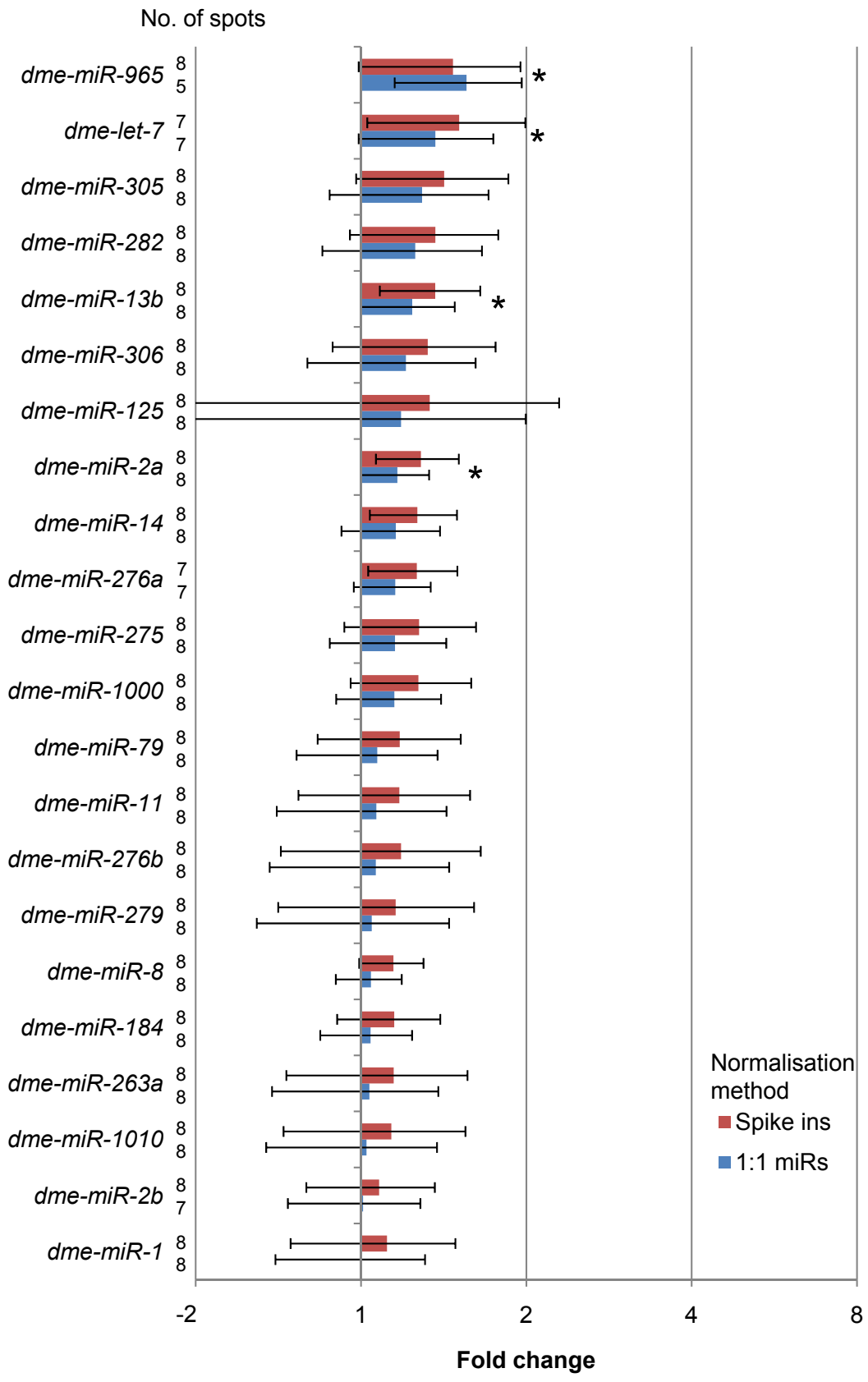



**Figure 6.5**: Upregulated miRNAs in *pcm*$^{5}$ wing imaginal discs.
The data was first normalised using spike in RNAs and re-normalised with 1:1 wild-type:*pcm*$^{5}$ ratio miRNAs identified by qPCR.
Error bars show standard deviation between ratios for of quality spots for each miRNA (max 8).
* = p-value <0.05 .

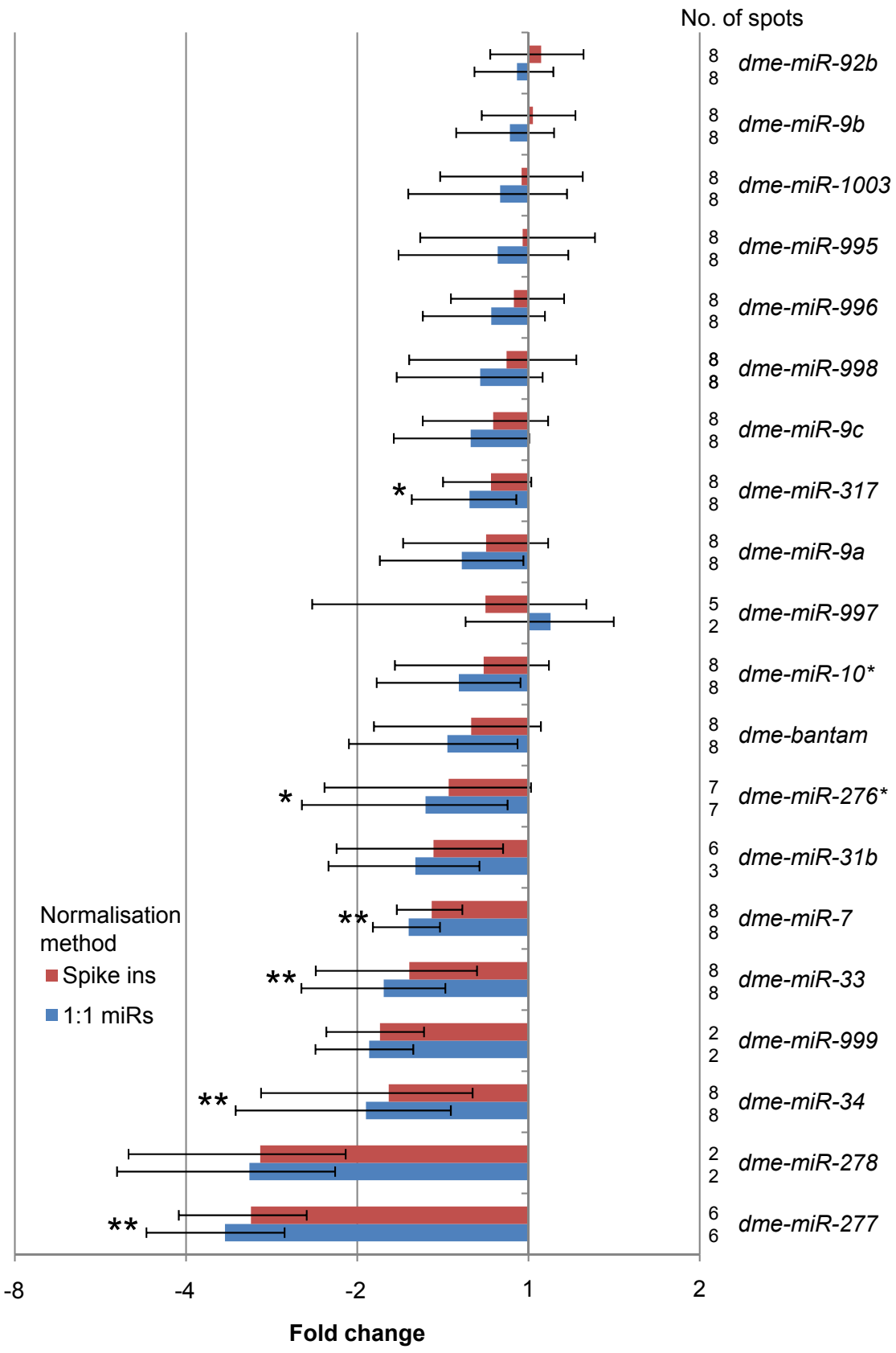



**Figure 6.6**: Downregulated miRNAs in *pcm*$^{5}$ wing imaginal discs.
The data was first normalised using spike in RNAs and re-normalised with 1:1 wild-type:*pcm*$^{5}$ ratio miRNAs identified by qPCR.
Error bars show standard deviation between ratios for of quality spots for each miRNA (max 8).
**= p-value <0.001, * = p-value <0.05 .

### 6.5.7 miRNAs expressed in *D. melanogaster* L3 wing imaginal discs

Of the 238 *D. melanogaster* miRNAs annotated in miRBase release 17, 152 were represented on the Exiqon arrays used. 42 of these 152 were detectable in the wild-type wing imaginal discs under the conditions used here. It is possible that some refinement of the experiment would allow for the detection of more low level miRNAs, as even at full gain settings on the array reader, few spots on the array were saturated. Increasing the amount of total RNA may give better results.

Figure 6.7 shows the mean level of the each miRNA relative to the median level of all 42. As the list contains an equal number of miRNAs, the median is the mean level of the middle two, *miR-995* and *miR-1010*. *miR-14* is the most highly expressed miRNA in wild-type wing imaginal discs, followed by *miR-79*, *miR-2b*, *miR-9a* and *miR-305* (in descending order). *miR-277* is the lowest level miRNA detected on these arrays, followed by *miR-278*, *miR-31b*, *miR-999* and *miR-276** (in ascending order). *miR-14* is expressed at just over 100 times the level of *miR-277*. The four miRNAs significantly upregulated in $pcm^5$ wing discs were spread across the expression level range (ranks 7, 11, 20 and 32). The six significantly downregulated miRNAs are all expressed at below the median level, at ranks 28, 33, 34, 36, 38, and 42. *miR-278* which is also possibly downregulated is at rank 41. Of six miRNAs confirmed as having a 1:1 expression ratio between wild-type and $pcm^5$ by qPCR, four of these appear in the array results (*miR-79*, *miR-9a*, *miR-8* and *miR-92b*) and are spread across expression levels (ranks 2, 4, 12 and 30 respectively). Two 1:1 ratio miRNAs were not picked up on the arrays (*miR-304* and *miR-31a*), possibly because they are at too low levels. This spread of 1:1 miRNAs has examples of high and low level miRNAs, but is not representative of the real distribution as the miRNAs selected for qPCR were not selected randomly.

Overall, it seems that miRNAs at lower expression levels are more likely to be downregulated in $pcm^5$ mutant wing discs. This may be due to how the miRNAs are functioning. Highly expressed miRNAs may be expressed in many cells to globally downregulate their targets and to exclude expression of certain pathways from the wing discs. Low expression miRNAs may be low because they are expressed in specific cells to

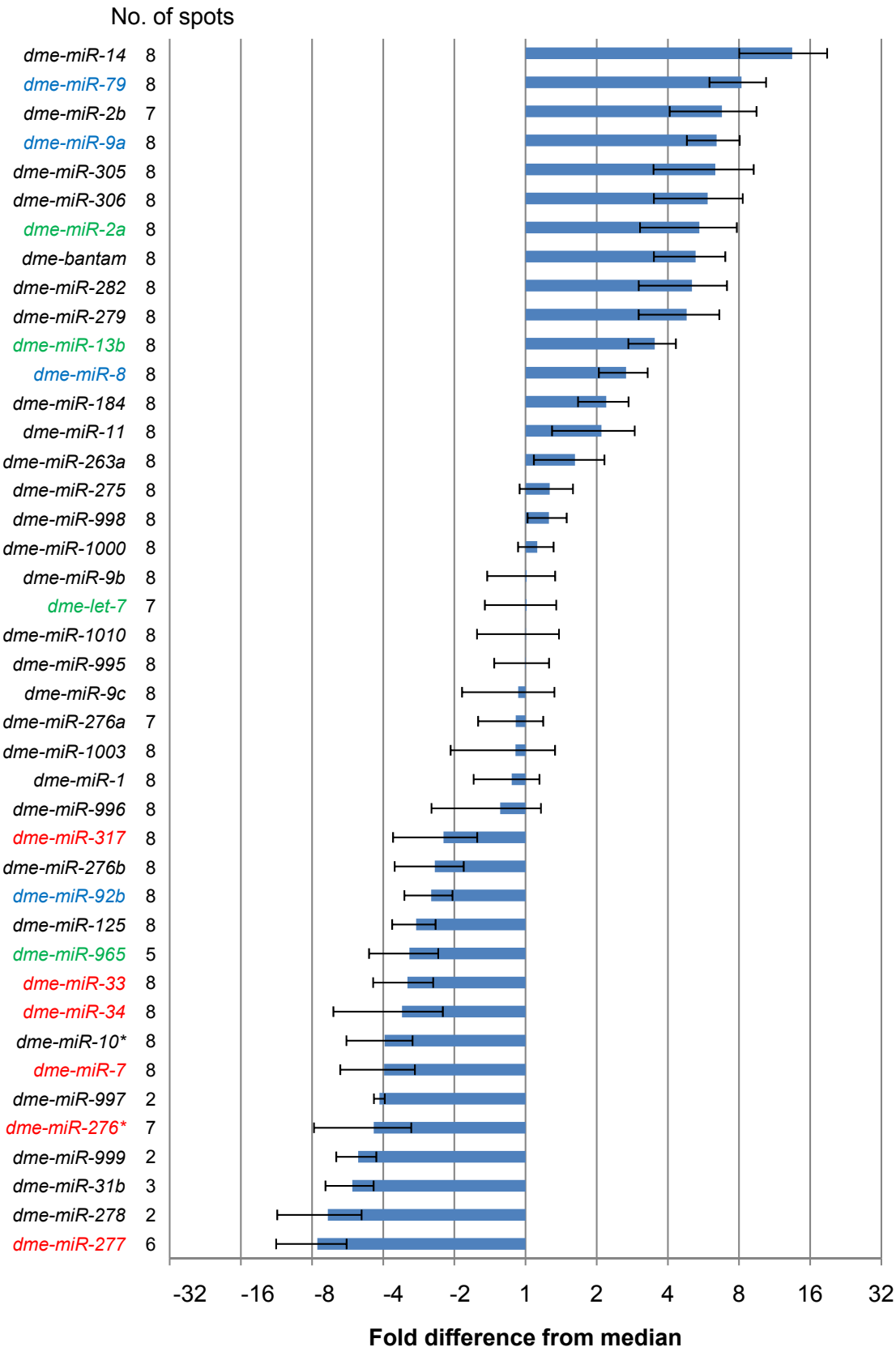



**Figure 6.7**: miRNAs expressed in wild-type wing imaginal discs, relative to the overall median level.
Blue indicates 1:1 ratio miRNAs which are at the same relative level in wild-type as in *pcm*$^{5}$ mutant wing discs. Green indicates miRNAs identified as statistically significantly upregulated in *pcm*$^{5}$ wing discs. Red indicates miRNAs identified as statistically significantly downregulated in *pcm*$^{5}$ wing discs.

knock down certain pathways in a small subset of cells, and as such may themselves be targets of specific regulation. It also seems likely that 42 is an underestimation of the number of miRNAs expressed in wing imaginal discs, as the arrays appear less sensitive than qPCR. Two miRNAs have been found that can be picked up by qPCR but do not appear on the array results.

### 6.5.8 Comparison of miRNA expression in L3 wing imaginal discs and adult testes

At the same time as the wing imaginal disc miRNA arrays were performed, I also did the same experiment on RNA from adult *D. melanogaster* testes for Dr. Maria Zabolotskaya. The arrays were wild-type vs. $pcm^5$ and were performed in the same as the wing disc arrays. Eight miRNAs were statistically significantly different from their wild-type level in the $pcm^5$ testes (six up, two down), but the largest fold difference was only 1.7-fold (Figure 6.8). $pcm^5$ mutants are less fertile than wild-type and a number of mRNAs have been identified as targets of Pacman, but Pacman does not appear to be affecting miRNA levels to any great extent.

Overall, the number of miRNAs detected was lower for testes than wing discs. 31 were detected in testes above background level, compared to 42 in wing discs (Figure 6.9). In common with wing discs, *miR-14* was the most highly expressed miRNA and *miR-277* was at a low level in both tissues. Figure 6.10 shows a comparison of the rank of each miRNA in each tissue. Note that only the rank of each miRNA relative to the other miRNAs expressed in the same tissue can be compared. The levels between tissues are not directly comparable, for example, *miR-14* may rank first for both tissues, but this does not mean its absolute level is the same in each. 20 miRNAs were common to both testes and wing discs, 11 miRNAs were detected only in testes and 22 were detected only in wing discs.

Comparison of the miRNA expression was done to observe any similarities and differences, rather than test a specific hypothesis. It is interesting to note that *miR-14* expression is high in both tissues, despite them being so diverse in function. It is also interesting that *miR-277* is low in both, but appears to be affected by Pacman only in wing discs. The mRNAs Pacman affects in wing discs and testes are different, and the differential effect on miRNA expression fits with the apparent tissue specificity of Pacman function.

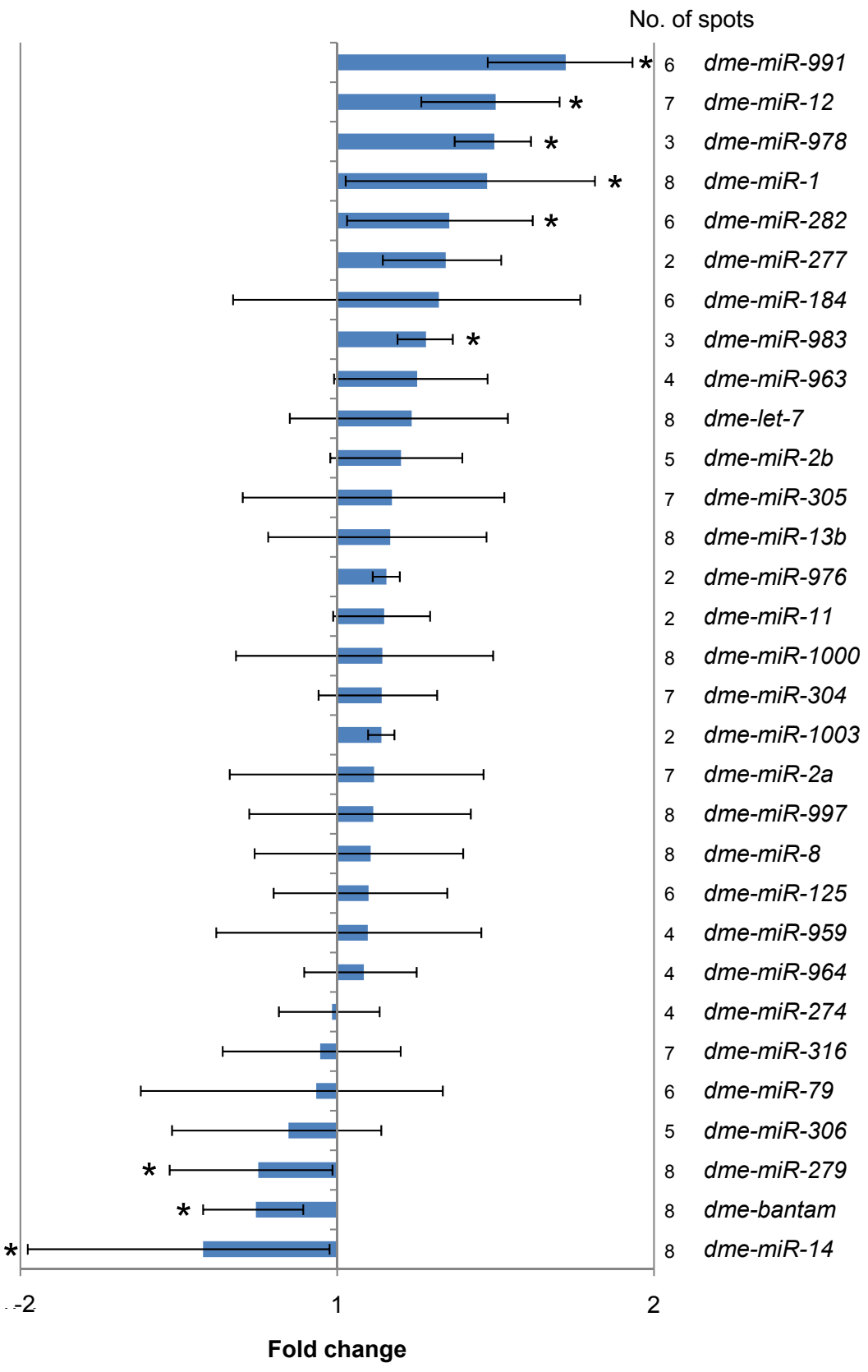



**Figure 6.8**: Up- and downregulated miRNAs in *pcm*$^{5}$ adult testes.
The data was normalised using spike in RNAs. Error bars show standard deviations between ratios for good quality spots for each miRNA (max 8). * = p-value <0.05 .

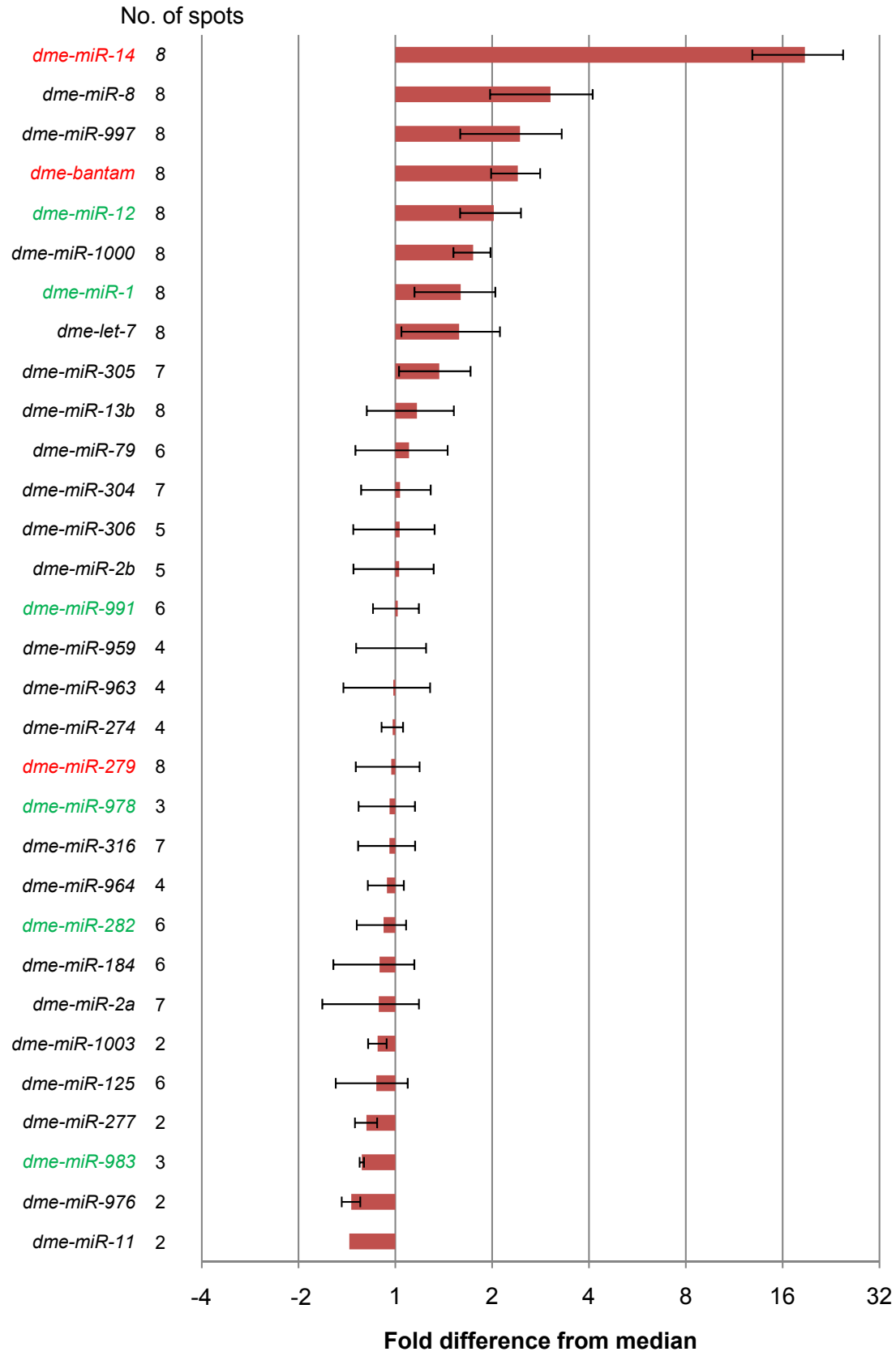



**Figure 6.9**: miRNAs expressed in wild-type adult testes, relative to the overall median level. Green indicates miRNAs identified as statistically significantly upregulated in *pcm*$^{5}$ testes. Red indicates miRNAs identified as statistically significantly downregulated in *pcm*$^{5}$ testes.

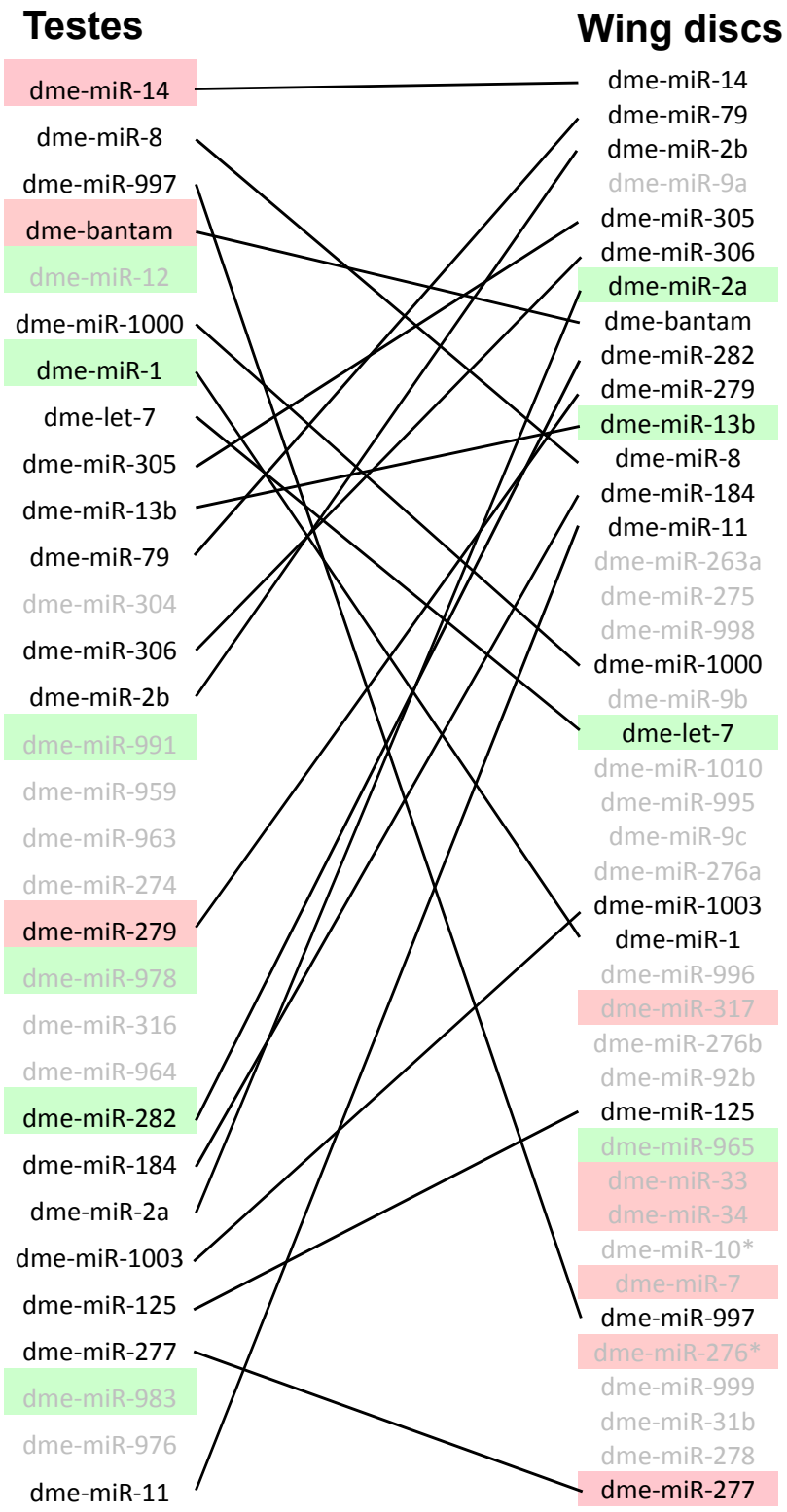



**Figure 6.10**: Comparison of miRNAs expressed in wing imaginal discs and testes, ranked in order of expression level relative to other miRNAs in the same tissue.
Green indicates miRNAs identified as statistically significantly upregulated in *pcm*$^{5}$ mutants.
Red indicates miRNAs identified as statistically significantly downregulated in *pcm*$^{5}$ mutants.
Grey miRNAs are only expressed in one tissue.

## 6.6 qPCR verification of array results

### 6.6.1 miRNAs chosen for verification

As only two miRNA arrays were performed, the results had to be verified with qPCR. The miRNAs chosen for verification after the arrays had been performed were *miR-8*, *miR-14*, *miR-33*, *miR-34* and *miR-277*. These qPCRs were set up in the same way as the qPCRs to verify mRNA fold changes in the previous chapter (Figure 5.10). The four wild-type replicates (the two used on the arrays and the two not used on the arrays) were pooled together and treated as one sample while the four $pcm^5$ replicates were tested separately.

### 6.6.2 qPCR results

Figure 6.11, panel A shows the five miRNAs from the arrays picked for verification alongside the levels of the other miRNAs tested in wing imaginal discs. Five show a significant difference from wild-type but only one, *miR-277* shows a greater than 2-fold difference (-5.9-fold). *miR-289* and *miR-34* show differences of just under 2-fold, +1.73-fold and -1.8-fold respectively. Figure 6.11, panel B compares the array and qPCR results for miRNAs at detectable levels on the array after normalisation with 1:1 ratio miRNAs (*miR-8*, *miR-79* and *miR-92b*). The array and qPCR expression levels are broadly consistent.

## 6.7 Search for potential relationships between the miRNA and mRNA expression levels

If Pacman is having an effect on miRNA expression that causes a biological effect, the relevant target(s) of that miRNA should be identifiable. There are two approaches to predicting miRNA:mRNA interactions. The mRNAs to which a given miRNA could bind can be predicted, or potential binding sites for miRNAs can be identified in the 3' UTR of specific mRNAs. Ideally, the effect of the change in level of a number of miRNAs on the mRNAs expressed in a cell could be predicted, but this is not yet possible as few miRNA:mRNA interactions have been verified *in vivo* or even in cell culture. Attempting to reconcile changes in mRNA levels and miRNA levels with the aim of finding a causal relationship is unlikely to produce useful data that could be separated from the background of false interactions and incorrect predictions. If we were to take the mRNA and miRNA array datasets from this and the previous chapter, attempting to find global links between the data is unlikely to produce any results that could be easily interpreted. An approach that

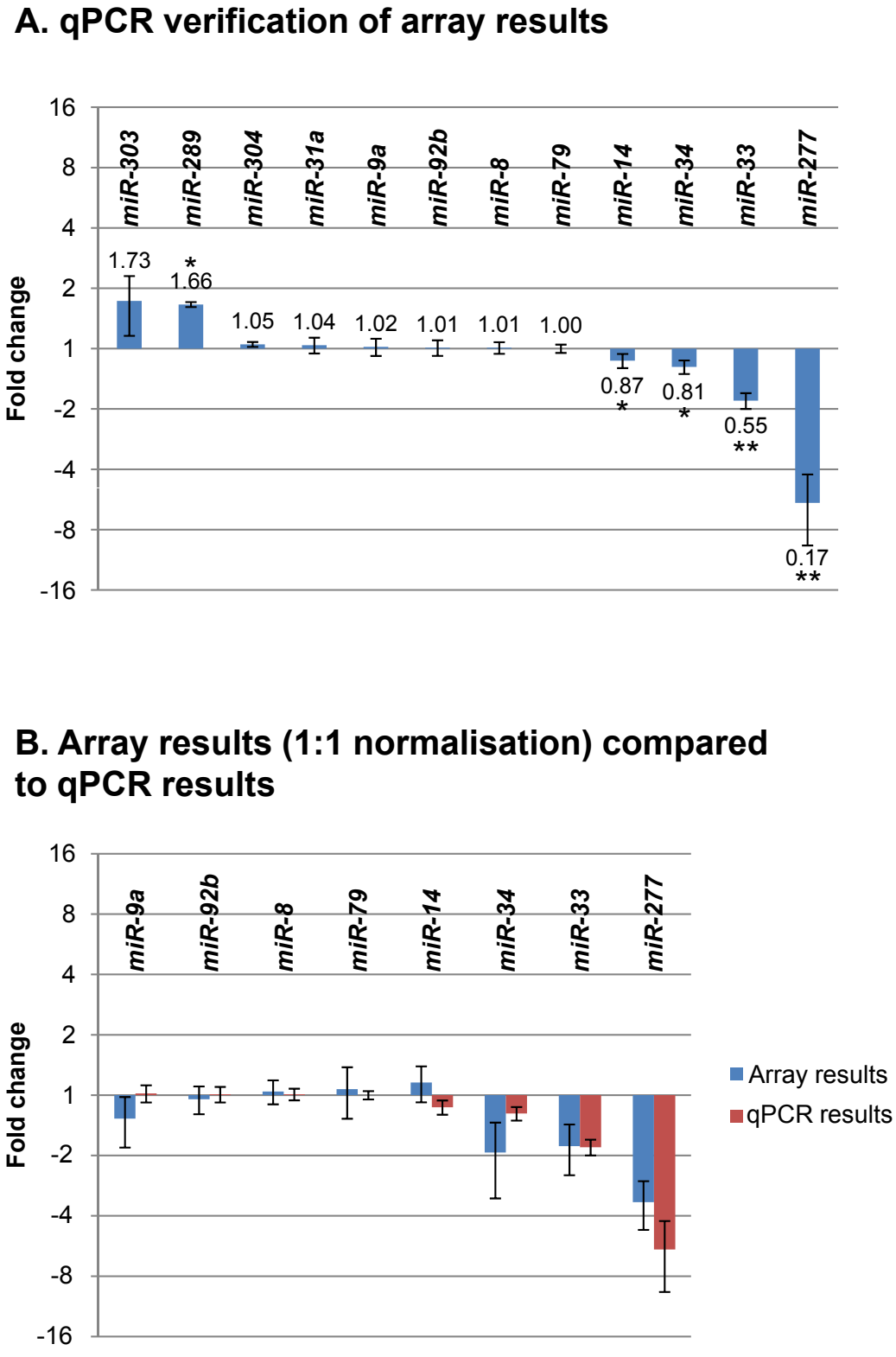



**Figure 6.11**: qPCRs performed on miRNAs in wild-type and *pcm*$^{5}$ wing imaginal discs.
**A**. Summary of all qPCRs performed. n=4 for *miR-8/14/34/33/277* and n=3 for *miR-303/289/304/31a/9a/79/92b*. Error bars show standard deviation between biological replicates. 4 technical replicates were performed per biological replicate. Mutant/wild-type ratio is given above each bar.
** = p-value <0.001, * = p-value <0.05 .
**B**. Comparison of qPCR and microarray levels (microarray data normalised using 1:1 ratio miRNAs *miR-92b*, *miR-8* and *miR-79*).

may highlight real interactions is to look only at miRNAs/mRNAs that are unlikely to be false positive results, i.e. those that have been verified with qPCR and vary from their wild-type level by a large amount (>2-fold).

### 6.7.1 *miR-277* targets

There are three mRNA transcripts and one miRNA that vary from wild-type by >2-fold in Pacman mutant wing discs and have been verified by qPCR. *CG31477* and *simjang* are upregulated by 6.3-fold and 3.8-fold respectively. *CG32364* and *miR-277* are downregulated by 8.3-fold and 5.9-fold respectively. As there is only one miRNA that changes in level, the first step is to see if *miR-277* is predicted to affect any of the three mRNAs. In this case, *CG32364* is unlikely to be a target as it is downregulated in *pcm* mutants and this has been shown to be at the level of transcription. If *miR-277* was having an effect, the level of its targets should increase at the post-transcriptional level, in the way that is seen for *CG31477* and *simj*. TargetScanFly predicts 758 conserved binding sites in 691 mRNAs for *miR-277*, but none within *CG31477*, *simj* or *CG32364*. If the search is widened to include mRNAs that change on the arrays but have not been verified to have a >2-fold difference by qPCR, five conserved sites can be found in four mRNAs in the upregulation list: *CG1140*, *Dcp2* (x2), *CG10098* and *CG504*. Two more conserved sites can be found in the downregulation list, in *wallenda* and *CG31549*. Of these, *Dcp2* has been shown by qPCR to increase by 1.3-fold. Whether or not there is any meaningful relationship between the 5.9-fold downregulation of *miR-277* and the 1.3-fold upregulation of *Dcp2* (or possible upregulation of the other mRNAs) in *pcm* mutants could only be further elucidated experimentally. However, this would be difficult with changes as small as 1.3-fold, and the biological relevance of such a small change is debateable.

### 6.7.2 miRs predicted to target *simj*, *CG31477* and *CG32364*

TargetScanFly predicts one miRNA target site in the 270bp 3' UTR of the *simj* transcript for *miR-1000*. *miR-1000* is expressed in wing discs, but its level does not change in $pcm^5$ mutants. The 100bp 3' UTR of *CG31477* contains no predicted miRNA binding sites. The 1.5kb 3' UTR of *CG32364* contains four conserved sites for *miR-33*, *miR-92*/*310*/*311*/*312*/*313*, *miR-927* and *miR-964*. Of these, *miR-92b* and *miR-33* were detected by the arrays in wing discs and both have been verified by qPCR. *miR-92b* was at wild-type level in *pacman* mutants and *miR-33* was downregulated by 1.8-fold. However, the relationship should be opposite if *miR-33* was causing the downregulation of *CG32364*. Overall it does not appear that the downregulation of *miR-277* is obviously responsible for

the changes seen in the levels of any mRNAs in *pcm*$^5$ mutant wing discs. It is more likely that *miR-277* is downregulated as a result of a change in level of one or more mRNAs. *simjang* is a good candidate, as it is a known repressor of transcription (Kon *et al.*, 2005).

## 6.8 Chapter summary

This chapter investigated the hypothesis that as an exoribonuclease, Pacman may be responsible for degradation of miRNAs and disruption of miRNA expression in *pacman* mutants could be responsible for changes in mRNA expression and the phenotypes observed. Initial experiments did not identify any miRNAs predicted to affect the JNK pathway that changed in level in *pcm* mutants, in either whole larvae or wing imaginal discs. miRNA arrays and subsequent qPCR verification revealed one miRNA, *miR-277*, that is downregulated by 5.9-fold in *pcm* mutant wing imaginal discs. No other miRNAs were identified as being up- or downregulated by >2-fold. The results are consistent with the null hypothesis that Pacman does not degrade miRNAs and the specific effect on *miR-277* is most likely at the transcriptional level as the result of upregulation of one or more mRNAs. *simjang* is a good candidate for this role as it codes for a transcriptional repressor and is upregulated in *pcm*$^5$ wing imaginal discs.

The comparison between testes and wing imaginal disc miRNAs showed that, as expected, the miRNA profiles of these diverse tissues are different. There are some similarities, such as *miR-14* being at a high level and *miR-277* being at a low level in each, but these miRNAs could be functioning differently by targeting different mRNAs. It is interesting though that *miR-277* is not downregulated in *pcm*$^5$ adult testes. This is further evidence that the effect of Pacman on *miR-277* is not direct.

# 7 Chapter 7 – Conclusions and discussion

This discussion is split into three sections. The key results/conclusions from each chapter will be presented first, followed by discussion of the significance of these findings and how they relate to the published literature. Finally, the Future work section will present ideas to test the hypotheses formed during the discussion of the results.

## 7.1 Summary of thesis results

### 7.1.1 New fly stocks created

| Mutation | Full stock genotype | Source |
|---|---|---|
| *pcm$^{13}$* | *w$^{1118}$ pcm$^{13}$ Nat1$^{121c}$ / FM7i, P{w$^{+mC}$=ActGFP}JMR3* | P-element excision (Chapter 3) |
| *pcm$^{14}$* | *w$^{1118}$ pcm$^{14}$ / FM7i, P{w$^{+mC}$=ActGFP}JMR3* | P-element excision (Chapter 3) |
| *pcm$^{14}$* | *y$^{1}$ pcm$^{14}$ / FM7h* | Recombination (Chapter 4) |
| 50E | *w$^{1118}$ 50E* | P-element excision (Chapter 3) |
| *Df(1)ED7452* | *Df(1)ED7452, w$^{1118}$ / FM7i, P{w$^{+mC}$=ActGFP}JMR3* | DrosDel Rearrangement Screen (Chapter 4) |

#### 7.1.1.1 *pcm$^{13}$*

*pcm$^{13}$* is a deletion of 2,222bp including 590bp into the 3' of *pcm* and 529bp into the 3' of the neighbouring gene *Nat1* (Figure 3.16). A preliminary test of the expression level of *pcm* mRNA in *pcm$^{13}$* showed it to be about 1.5-fold lower than the wild-type level (Figure 4.15). The phenotypes observed for *pcm$^{13}$* (such as dull wings and bristle defects) are consistent with other non-lethal *pcm* mutants such as *pcm$^{3}$*, *pcm$^{5}$* and *pcm$^{6}$*. Comparison with other *pcm* mutants suggests *pcm$^{13}$* is a hypomorphic mutation. Expression of *Nat1* is known to be unaffected in *pcm$^{3}$* and *pcm$^{5}$* (Grima *et al.*, 2008) and the deletion into *Nat1* in *pcm$^{13}$* is not obviously responsible for any phenotypes. No lines encountered during the P-element excision that were deletions towards *Nat1* (but not towards *pcm*) showed any obvious adult phenotypes and appeared as wild-type. These stocks were balanced over *FM7c*, which was lost from the stocks after a number of generations, as occurs with wild-type (due to the fact

flies carrying a balancer are slightly less fit than wild-type flies). *Nat1* is annotated on FlyBase as a peptide alpha-N-acetyltransferase. Six alleles are reported (all are insertions into the 5' UTR) but no phenotypes are described. It may be the case that *Nat1* mutations do not have any effect, but there is currently insufficient data to determine this for certain. It could also be that *Nat1* causes phenotypes that have not been observed, for example during early development, or it may be required under specific conditions not encountered by our flies. For these reasons, $pcm^{13}$ is not as useful an allele as the other non-lethal *pcm* alleles, as the *Nat1* deletion is still a complicating factor, despite having no known effects.

It should be noted that there are two genes annotated in FlyBase with the name *Nat1*. The other is *NAT1* (*CG3845*), which is located on chromosome *2* and encodes a translation initiation factor. *NAT1* should not be confused with the *Nat1* (*CG12202*) neighbouring *pcm* on *X*.

#### 7.1.1.2 $pcm^{14}$

$pcm^{14}$ was determined to be a 3,501bp deletion extending from the site of $P\{w^{+mC}=EP\}EP1526$ including 3,068bp into the 3' of *pcm*. Exons 7-11 and the 3' UTR of *pcm* are entirely deleted while exons 1-5 and the 5' UTR are unaffected. Exon 6 is partially deleted (76bp of 320bp) (Figure 3.16). The deleted region is replaced by a 9bp "footprint" left over from the P-element excision (Figure 3.14). This mutation causes lethality during early pupation (Figure 4.9). The lethality maps to the *pcm* locus (Figure 4.2) and can be rescued by *T(1;Y)B92*, a translocation from *X* to *Y* that includes *pcm* (Figure 4.4). $pcm^{14}$ acts as a functional null (amorphic) allele (Figure 4.6). *pcm* mRNA is expressed from what remains of the *pcm* gene at the same level as wild-type (Figure 4.15). We do not know if protein is translated from the $pcm^{14}$ transcript; however any protein translated must be non-functional, as the $pcm^{14}$ allele gives the same phenotypes as a deficiency (such as *Df(1)ED7452*) when combined with the hypomorphic allele $pcm^{5}$ (Figure 4.12).

Two stocks have been generated containing the $pcm^{14}$ deletion. The first was the original created during the P-element excision that includes the $w^{1118}$ allele, meaning these flies would have white eyes if the $pcm^{14}$ mutation was not hemi/homozygous lethal. The second was derived from the original $pcm^{14}$ stock and was recovered during the crossover experiments (Figure 4.2), where recombination has occurred to replace the section containing wild-type *y* and $w^{1118}$ with a wild-type *w* gene and the allele $y^{1}$. These flies would have red eyes and yellow bodies if $pcm^{14}$ mutation was not hemi/homozygous lethal. The balancer used for this stock is *FM7h*, which includes the alleles $w^{1}$ and $y^{31d}$. $y^{1}\ pcm^{14}/FM7h$

heterozygotes are yellow bodied due to the *y* alleles on each chromosome, but red eyed due to the wild-type *w* on the $pcm^{14}$ chromosome.

#### 7.1.1.3 Control lines

From a number of suitable control lines created and sequenced, two were randomly selected to be retained, referred to as 50E and 69E (Figure 3.9). Of these 50E has become the main wild-type control line to which the *pcm* mutants are compared. 50E was created by a precise excision of $P\{w^{+mC}=EP\}EP1526$, i.e. one that did not induce a deletion. A 43bp "footprint" remains at the $P\{w^{+mC}=EP\}EP1526$ excision site, downstream of both *pcm* and *Nat1*. 50E flies do not display any non-wild-type phenotypes, except for white eyes due to the $w^{1118}$ mutation that is also present on the chromosome. The $w^{1118}$ mutation is also present on each chromosome containing a *pcm* allele (such as in $pcm^{3}$, $pcm^{5}$ and $pcm^{6}$ stocks).

#### 7.1.1.4 *Df(1)ED7452*

*Df(1)ED7452* is a 17,963bp deletion that includes the entirety of *pcm*, *CG12203*, *CG12204* and *kish* (*CG14199*) and the majority of *Nat1*. It also includes a small amount of the 5' of the large gene *Pfrx*. The deleted sequence is replaced by roughly 5kb of sequence containing a functional allele of *white* ($w^{hs}$) and an FRT site, bounded by two 3' P-element ends (Figure 4.17). *Df(1)ED7452* has been verified molecularly and genetically and is the smallest available deficiency that includes *pcm* (Figure 4.21).

## 7.2 Effects of *pacman* mutations

### 7.2.1 Pacman is required for *D. melanogaster* development

Data presented in this thesis are consistent with the prior work that shows *pacman* mutations impair proper development. The most common phenotypes observed in non-lethal *pcm* mutants are dull wings and bristle defects (Figures 4.7 and 4.8). In the null mutant ($pcm^{14}$), L3 larvae are able to survive, but die during early pupation (Figure 4.9). Pacman may also be essential at earlier points during development, such as embryogenesis, but *pacman* mRNA is known to be contributed maternally (Till *et al.*, 1998), and this may be sufficient to overcome any earlier lethality.

### 7.2.2 $pcm^{14}$ and $pcm^{5}$ L3 wing imaginal discs are significantly smaller than wild-type

The wing imaginal discs of $pcm^{5}$ L3 larvae are 82% the size of equivalent wild-type (50E) wing discs (Figure 4.10). $pcm^{14}$ wing discs are significantly smaller than $pcm^{5}$ wing discs and 43% the size of wild-type ($n \geq 25$ for each group and $p<0.001$ for all comparisons).

### 7.2.3 *simjang* and *CG31477* are upregulated in $pcm^{5}$ mutant L3 wing imaginal discs

The level of *simj* mRNA is increased by 3.8-fold ($p<0.001$) in $pcm^{5}$ L3 wing imaginal discs and the level of *CG31477* mRNA is increased by 6.7-fold ($p<0.001$) (Figure 5.13). There are three possible mRNA transcripts expressed from the *simj* gene, but the only one expressed in wild-type or $pcm^{5}$ wing discs is the longest, *simj-RA* (Figure 5.14). *CG31477* is a gene that produces a single transcript. The 3.8-fold increase in *simj* mRNA leads to a roughly 3-fold increase in Simj protein (Figure 5.13).

#### 7.2.3.1 *simjang* is a direct target of Pacman

The level of *simj* pre-mRNA does not differ significantly from wild-type ($p=0.33$). This shows that the effect on mature s*imj* mRNA must be at the post-transcriptional level (Figure 5.13). As the only difference between the wild-type and $pcm^{5}$ strains is a *pcm* mutation, the post-transcriptional effect must be a result of the *pcm* mutation. As *pcm* is the only gene capable of post-transcriptionally affecting mRNA levels that is not at its wild-type level of function (no other known mRNA degradation or processing factors were differentially expressed in *pcm* mutants, as shown by the mRNA arrays), the logical conclusion is that *simj*

mRNA is a direct target of Pacman in L3 wing imaginal discs. The effect on *simj* is specific, as a global increase in mRNA levels is not observed.

#### 7.2.3.2 *CG31477* may be a direct target of Pacman

If statistical significance is taken as p≤0.05, then the level of *CG31477* pre-mRNA in *pcm*$^5$ wing discs is just significantly higher than in wild-type (1.7-fold, p=0.04), while the level of mature *CG31477* mRNA is 6.7-fold higher than wild-type (Figure 5.13). There are three hypotheses that could account for this the observed expression levels. 1) The rate of turnover of pre-mRNA (maturation) is greater than the rate of turnover of mRNA (degradation), therefore a small increase in the level of transcription of a gene (e.g. +1.7-fold) could lead to a greater increase in the mRNA level (e.g. +6.7-fold). 2) *CG31477* causes increased transcription of itself, so failure of Pacman to degrade *CG31477* mRNA also leads to an increase in *CG31477* transcription, further increasing the level of *CG31477* mRNA. 3) The +1.7-fold difference may be due to technical reasons rather than represent a real biological difference. Further work is required to test these hypotheses and determine if *CG13477* is a direct target of Pacman.

### 7.2.4 *CG32364* and *miR-277* are downregulated in *pcm*$^5$ mutant L3 wing imaginal discs

The levels of *CG34364* and *miR-277* are both reduced in *pcm*$^5$ L3 wing discs, by 8.3-fold (p<0.001) and 5.9-fold (p<0.001) respectively.

#### 7.2.4.1 *CG32364* is an indirect target of Pacman

The reduction in level of mature *CG32364* mRNA by 8.3-fold (to 12% of wild-type level) is concurrent with a reduction in the level of *CG32364* pre-mRNA by 6.3-fold (to 16% of wild-type level). The difference between these values is not statistically significant (p=0.2), so the reduction in the level of pre-mRNA accounts for the reduction in level of mRNA (Figure 5.13). Therefore, *CG32364* is transcriptionally reduced in level as an indirect effect of the *pcm*$^5$ mutation, but is not directly affected by Pacman.

#### 7.2.4.2 *miR-277* is probably an indirect target of Pacman

The levels of *pri-* or *pre-miR-77* have not yet been measured due to time constraints. However, it seems unlikely that *miR-277* would be found to be not transcriptionally downregulated, as this would mean Pacman in the *pcm*$^5$ mutant would be degrading *miR-277* more efficiently than Pacman in wild-type. There are no significant technical considerations to overcome to measure the level of a pri- or pre-miRNA, but until such a

test has been performed, it seems reasonable to assume *miR-277* is transcriptionally downregulated in a similar manner to *CG32364*.

### 7.2.5 Other direct and indirect mRNA and miRNA targets of Pacman

*simj*, *CG31477*, *CG32364* and *miR-277* have all had their expression levels verified by qPCR. Except for in a minority of cases (*Hsc-70* and *miR-315*), TaqMan qPCR has proven a reliable, consistent and accurate technique. It is also relatively quick to perform extra biological and technical replicates when required, as the sample sizes do not have to be as large as for array analysis (in a pilot for an experiment detailed in the Future work section, *simj* expression has been measured in RNA extractions of around 1µg from as few as 20 L3 wing imaginal discs). For these reasons we favour the accuracy of the qPCR fold changes over those given by the mRNA or miRNA microarrays. The microarrays worked as we intended them to – they provided a number of transcripts or miRNAs that changed in level, from which we picked candidates for further analysis. Precluding results not verified by qPCR from more detailed analysis undoubtedly means we have excluded other direct and indirect targets of Pacman. Figure 7.1 shows a combined list of mRNAs and miRNAs that change in level on arrays by ≥1.6-fold. Those verified with qPCR are marked in green. While some may be false positives, it is likely that there are more targets of Pacman to be found among those that have not been verified.

| CG/CR number | Synonym | Fold change (array) | Fold change (qPCR) | Direction |
|---|---|---|---|---|
| *CG31477* | | 10.3 | 6.3 | Up |
| *CG32364* | | 10.2 | 8.3 | Down |
| *CG7756* | *Hsc70-2* | 9.2 | Not determinable | Down |
| *CG17669* | | 7.3 | | Down |
| *CG31157* | | 6.6 | | Down |
| *CG4190* | *Hsp67Bc* | 4.6 | | Up |
| *CG7966* | | 4.4 | | Down |
| *CG33099* | | 3.4 | | Up |
| *CG5326* | | 3.3 | | Up |
| *CG42330* | | 3.3 | | Down |
| *CR33588* | *dme-miR-277* | 3.2 | | Down |
| *CG12842* | | 2.9 | | Up |
| *CG4183* | *Hsp26* | 2.8 | | Up |
| *CG9238* | | 2.8 | | Down |
| *CG18349* | *Cpr67Fa2* | 2.8 | | Down |
| *CG7734* | *shn* | 2.6 | 1.2 | Up |
| *CG3291* | *pcm* | 2.6 | | Down |
| *CG15006* | *Cpr64Aa* | 2.4 | | Down |
| *CG10581* | | 2.4 | | Down |
| *CG12749* | *Hrb87F* | 2.3 | | Down |
| *CG5953* | | 2.2 | | Up |
| *CG12800* | *Cyp6d4* | 2.2 | | Down |
| *CG34247* | | 2.2 | | Down |
| *CG17286* | *spd-2* | 2.2 | | Down |
| *CG7054* | | 2.1 | | Up |
| *CG13053* | | 2.1 | | Down |
| *CG6155* | *Roe1* | 1.9 | | Up |
| *CG32067* | *simj* | 1.9 | 3.8 | Up |
| *CR33326* | *dme-miR-34* | 1.9 | | Down |
| *CG3217* | *CkIIalpha-i3* | 1.8 | | Up |
| *CG17725* | *Pepck* | 1.8 | | Up |
| *CG4427* | *cbt* | 1.8 | | Up |
| *CG6354* | *Rb97D* | 1.8 | | Up |
| *CR33569* | *dme-miR-33* | 1.8 | | Down |
| *CG11263* | | 1.8 | | Down |
| *CG31549* | | 1.8 | | Down |
| *CG15427* | *tutl* | 1.7 | | Down |
| *CG8789* | *wnd* | 1.7 | | Down |
| *CG6310* | | 1.7 | | Down |
| *CG10241* | *Cyp6a17* | 1.6 | | Up |
| *CG32056* | *scramb1* | 1.6 | | Up |
| *CG42254* | | 1.6 | | Up |
| *CR42959* | *dme-miR-965* | 1.6 | | Up |
| *CG18135* | | 1.6 | | Down |
| *CG6241* | | 1.6 | | Down |
| *CG31152* | *rumi* | 1.6 | | Down |
| *CG32412* | | 1.6 | | Down |
| *CR33042* | *dme-miR-7* | 1.6 | | Down |
| *CR33568* | *dme-miR-31b* | 1.6 | | Down |

**Figure 7.1**: mRNAs and miRNAs that change in level by ≥ ±1.6-fold in *pcm*$^{5}$ L3 wing imaginal discs.
Expression levels of RNAs marked in green have also been measured by qPCR.

## 7.3 Discussion

### 7.3.1 Pacman and Simjang

The most obviously interesting result presented in this thesis is that Pacman directly affects the level of *simjang* mRNA, which leads to an increase in the level of Simjang protein. The post-transcriptional increase in expression level of *simj* in $pcm^5$ L3 wing imaginal discs shows that Pacman must be responsible for degradation of *simj* under wild-type conditions. There are two alternative hypotheses concerning the role Pacman could be playing in *simj* regulation, one active and one passive. 1) Actively: *simj* mRNA is actively targeted either by Pacman or by a factor that causes it to undergo Pacman-dependent degradation.
2) Passively: *simj* mRNA can only be efficiently degraded in the 5' – 3' direction, and not by the 3' – 5' degradation machinery, so loss of Pcm function leads to loss of ability to efficiently degrade *simj* mRNA.

### 7.3.2 Simjang and transcriptional control by histone deacetylation

Simjang is most generically referred to as p66, but the protein is known by a variety of names in different organisms. p66 is a highly conserved protein found in all eukaryotes. A single p66 gene is present in *D. melanogaster* and *C. elegans* (dcp-66) and two paralogues of p66 exist in *M. musculus* (p66α/Gatad2a and p66β/Gatad2b) and *H. sapiens* (GATAD2A and GATAD2B). There are several other genes that share the synonym p66 in mice and humans (named Pold3, Shc1, Eif3d and Hspa9 in mice, with similar names in humans), but these genes are not evolutionarily or functionally related to Simjang. As this section will discuss work on p66 in a number of species, the name p66 will be used instead of Simjang/GATA2D2/dcp-66 etc. Other proteins referred to in this section will be given using *H. sapiens* nomenclature, with the *D. melanogaster* homologue following the first use in parenthesis, if there is a known homologue.

p66 is part of the Nucleosome Remodelling and Deacetylase (NuRD) complex (also known as the Mi-2 complex), which is one of a number of conserved histone deacetylation complexes found in eukaryotes. The mammalian NuRD complex varies in composition (Figure 7.2), but is typically comprised of the chromatin remodeler CHD3/4 (Mi-2), MBD2/3 (MBD-like), MTA1/2/3, (MTA1-like), p66 (Simjang), the histone deacetylases HDAC1 (Rpd3) and HDAC2 (also Rpd3), and two histone binding proteins RBBP4 (Caf1) and RBBP7

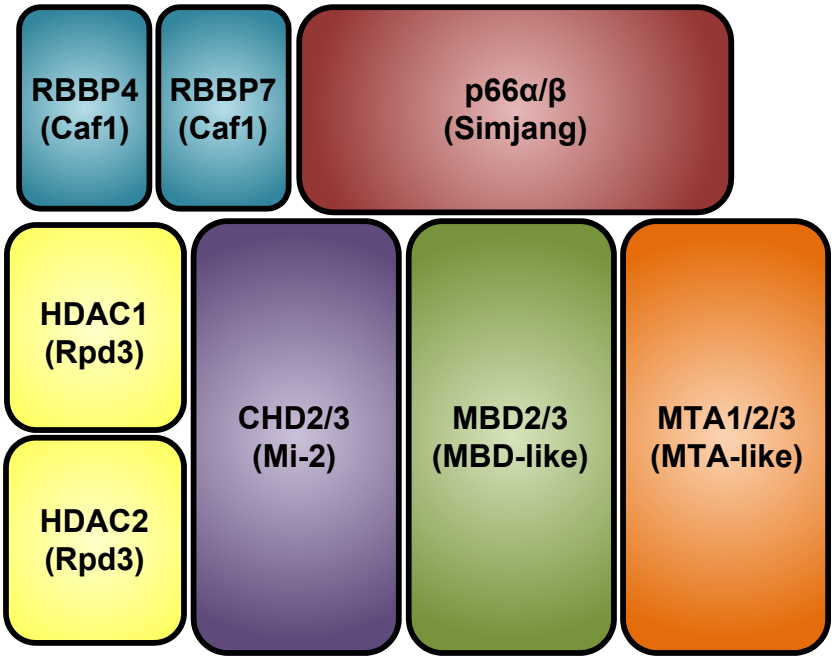



**Figure 7.2**: The mammalian NuRD complex and *Drosophila* homologues.
The mammalian NuRD complex consists of the histone deacetylases HDAC1 and HDAC2, the chromatin remodeler CHD2 or 3, MBD2 or 3 (involved in targeting), MTA1, 2 or 3 (DNA binding), the histone binding proteins RBBP4 and 7, and p66 (Le Guezennec *et al*., 2006, McDonel *et al.,* 2009, Ho and Crabtree, 2010). Two paralogues of p66 exist but may not be mutually exclusive. They appear to act synergistically to increase repression caused by the NuRD complex (Brackertz *et al*., 2006).

(also Caf1) (Le Guezennec *et al.*, 2006, Ho and Crabtree, 2010). Another deacetylation complex, the SIN3 complex, also contains HDAC1, HDAC2, RBBP4 and RBBP7, along with a number of other subunits, but does not include p66 (McDonel *et al.*, 2009). The NuRD and SIN3 complexes cause repression of gene expression by deacetylation of lysine residues in histone tails, which increases the affinity of histone:DNA binding. The SIN3 complex has been implicated in the control of a number of cellular processes, such as cell cycle arrest by repression of Wnt (Wingless) pathway targets (Sampson *et al.*, 2001), differentiation by repression of Myc (Diminutive) (Ayer *et al.*, 1995), neuronal (Nomura *et al.*, 2005) and cardiac (Bingham *et al.*, 2007) specific gene expression and B-/T-lymphocyte development (Koipally *et al.*, 1999). The SIN3 complex is recruited to genes by transcription repressors bound to the DNA (Silverstein and Ekwall, 2005). The NuRD complex operates in a similar manner to the SIN3 complex, but the targets it affects are less well defined. The NuRD complex is known to associate with a number of transcriptional repressors, some of which it shares in common with the SIN3 complex, such as IKZF1 and IKZF3, which are involved in T-cell development (Kim *et al.*, 1999). The developmental effects of mutants of the constituents of the NuRD and SIN3 complexes have been the subjects of numerous research papers (reviewed in McDonel *et al*., 2009), and the developmental requirement for histone deacetylase complexes is clear.

#### 7.3.2.1 p66 in the NuRD complex

p66α and β contain two highly conserved regions, CR1 and CR2. CR1 directly interacts with HDAC1, HDAC2, MTA2, MBD2/3, RBBP4 and RBBP7 (Feng *et al.*, 2002) and CR2 is required for association with histones (Brackertz *et al.*, 2006). The affinities of p66α and β for MBD2/3 differ slightly as p66α binds more strongly, and is also a more potent repressor of transcription (Brackertz *et al.*, 2002). SUMOlation of specific sites in p66α and β increase their affinities for MBD2/3 and increase transcriptional repression (Gong *et al.*, 2006). Unlike MBD2 and MBD3 or CHD3 and CHD4, the presence of one p66 paralogue does not exclude the other, and knock down of either paralogue will reduce repression caused by the other (Brackertz *et al.*, 2006).

#### 7.3.2.2 Targeting of the NuRD complex for repression

The NuRD complex appears to be targeted to specific areas of DNA by sequence-specific transcriptional repressors, but the exact interactions that occur are not fully clear. Repressors/co-repressors known to recruit the NuRD complex include Prospero (Tea *et al.*, 2010), Groucho (Chen *et al.*, 1999, Winkler *et al.*, 2010), Hunchback (Kehle *et al.*, 1998) and IKZF1 and IKZF3 (Kim *et al.*, 1999). The MBD proteins appear to be important in the targeting of the NuRD complex and the choice of MBD2 or MBD3 changes the function of the complex. MBD2, but not MBD3, binds to methylated regions of DNA (Zhang *et al.*, 1999, Le Guezennec *et al.*, 2006). MBD3 has recently been shown to be responsible for the recruitment of the NuRD complex to promoters induced by AP-1, by association with unphosphorylated c-Jun. Phosphorylation of c-Jun by JNK (*basket* in *D. melanogaster*) prevents the c-Jun:NuRD association and allows transcription to occur (Aguilera *et al.*, 2011).

### 7.3.3 Phenotypes of NuRD complex mutants

#### 7.3.3.1 p66 mutations

p66 mutants have been created in *M. musculus* (for p66α), *C. elegans* and *D. melanogaster*. In *M. musculus*, p66α null mutations are lethal as mutant embryos develop more slowly than siblings and suffer various defects before dying around day 10 of development. Array analysis performed on mutant embryos revealed around 45 genes at increased expression levels, including embryo-specific genes and male-specific genes that were being expressed in females (Marino and Nusse, 2007). In *C. elegans*, deletion of *dcp-66* causes the death of worms as small adults, due to failure of osmoregulation. Reports of the action of dcp-66 in *C. elegans* are somewhat contradictory, as it has been reported to inhibit transcription of *lag-2* in some cell types (Poulin *et al.*, 2005) and to increase the expression of *pgp-12* by binding to its promoter in others (Zhao *et al.*, 2005). In *D. melanogaster*, a number of p66 alleles of varying strengths have been created. Flies homozygous for these alleles die earliest as second instar larvae and latest during late pupal stages. The flies that die at later stages exhibit defects such as bent or shortened legs, split notum, bristle defects and wing abnormalities. Some pupae also contain a white fluid that caused pharate adults to stick to the pupal case. Embryonic phenotypes were not observed in p66 mutants, but this may be due to a maternal contribution of wild-type p66 (Kon *et al.*, 2005). An earlier study using RNAi also identified p66 as a gene required for embryonic cardiac development (Kim *et al.*, 2004).

#### 7.3.3.2 Histone Deacetylase mutations

There are numerous HDACs present in mammalian cells, but only two are known to associate with the NuRD complex (HDAC1/2), both of which correspond to the single HDAC Rpd3 in *D. melanogaster.* Flies carrying mutations in *rpd3* often die during larval stages (Mottus *et al.*, 2000) and show that it is required for embryonic segmentation (Mannervik and Levine, 1999) and olfactory dendrite targeting (Tea *et al.*, 2010). The L3 wing imaginal discs in a hypomorphic *rpd3* allele are smaller than wild-type due to lack of cell proliferation, which is in contrast to the effect mutations to the non-NuRD HDAC3, where L3 wing imaginal discs are reduced in size due to apoptosis (Zhu *et al.*, 2008). Phenotypes of HDAC1 and HDAC2 in mice suggest their roles are not redundant, as HDAC1 null mice die as embryos around day 10 due to lack of cell proliferation (Lagger *et al.*, 2002), whereas HDAC2 null mice die within 24 hours of birth due to heart defects (Montgomery *et al.*, 2007).

#### 7.3.3.3 CHD (Mi-2) mutations

Initial isolation of *D. melanogaster Mi-2* alleles showed that a maternal contribution of *Mi-2* mRNA allowed mutant flies to survive past the embryonic stage before dying as 1$^{st}$ or 2$^{nd}$ instar larvae. Generation of embryos lacking the maternal contribution of *Mi-2* mRNA failed as germline cells homozygous for any *Mi-2* alleles did not develop (Kehle *et al.*, 1998). A later study found that a small fraction of *Mi-2* mutants did survive to adulthood, and these exhibited a phenotype of extra Sensory Organ Precursors in the wing imaginal discs, which lead to extra sensory bristles on the thorax (Yamasaki and Nishida, 2006).

#### 7.3.3.4 MBD mutations

A single protein in *D. melanogaster*, MBD-like, is related to the mammalian MBD2 and MBD3 (Hendrich and Tweedie, 2003). There has been some debate as to whether MBD-like acts more like MBD2 or MBD3, largely regarding its ability to bind to methylated DNA, and this may be due to the fact the *MBD-like* gene can express two transcripts, the shorter of which lacks strong affinity for methylated DNA (Ballestar *et al.*, 2001, Marhold *et al.*, 2004). Flies null for *MBD-like* are viable and fertile, despite sometimes producing smaller embryos and showing chromosome segregation defects (Marhold *et al.*, 2004). *Mbd-like* has also been linked with embryonic cardiac development, displaying similar heart patterning phenotypes to *p66* overexpression (Kim *et al.*, 2004). In mice, loss of *Mbd3* causes lethality before day 8 of embryonic development, whereas *Mbd2* null mice are viable and fertile, but

females did display some behavioural defects when caring for their offspring (Hendrich *et al.*, 2001).

### 7.3.4 *D. melanogaster p66* phenotypes compared to *pacman* phenotypes

The phenotypes seen in *D. melanogaster* p66 mutants are similar to those seen for *pacman* mutants. Bent legs, failure of thorax closure and dull wings have been reported previously for $pcm^{5}$ and other hypomorphic alleles (Grima, 2002, Grima *et al.*, 2008). For the null allele $pcm^{14}$, flies die during early pupation (sometime after apolysis has occurred). At 19°C, $pcm^{5}/pcm^{14}$ flies are able to survive to begin eclosing from the pupae, but invariably become stuck to the pupal case and die. Phenotypic indicators are a good starting point for showing genetic relationships, but the information they provide is limited. In this case, the relationship suggested is opposite of what would be expected from the molecular results presented in this thesis. The phenotypes observed in p66 mutants are due to a decrease in p66 function (unless all the mutations lead to increased p66 function, which is unlikely). The phenotypes of *pcm* mutants are due to reduced Pcm function, which has been shown to lead to increased p66 expression. Logically, similar phenotypes would be expected if *pcm* mutants lead to decreased p66 expression. However, in practice the failure of this biological system may have the same end point (i.e. disrupted wing imaginal disc development), even if one failure is due to excessive p66 function and the other is due to lack of p66 function, if a fine balance of protein expression is required. Further work is required to verify exactly how the phenotypes are produced in both *pcm* and p66 mutants to explain this possible discrepancy.

Overexpression of p66 in embryos has been shown to cause lethality in first instar larvae, whereas overexpression in wing imaginal discs (driven by *patched*) causes bristle defects and reduced wing size (Kon *et al.*, 2005). An undergraduate in the Newbury lab, Hannah Parker, has recently shown that $pcm^{5}$ mutant wings are also smaller than wild-type, so in this case, the observed phenotype is as expected.

### 7.3.5 Phenotypes of other NuRD complex components compared to *pacman* phenotypes

It is interesting to note that *rpd3* (HDAC1/2) and *Mi-2* (CHD2/3) mutations produce phenotypes reminiscent of *pacman* mutant phenotypes. The phenotype of extra sensory bristles occurs in both *Mi-2* and *pcm* mutants, and $pcm^{14}$ L3 wing imaginal discs are greatly reduced in size compared to wild-type, as are *rpd3* mutant discs. The reduction in size of *rpd3* mutant discs is due to a reduction in cell proliferation, but it is not yet known why $pcm^{14}$ wing discs are smaller. If complete loss of *pacman* leads to an increase in p66, this should lead to an increase in Rpd3 activity, i.e. the effect should be opposite to mutating *rpd3*. Further work will need to be conducted to elucidate the details of this apparent discrepancy, but the reason for the small discs may not be the same in each mutant line, just as the reason for the same phenotype is different in *rpd3* and *HDAC3* mutants (Zhu *et al.*, 2008). For example, the reason for smaller wing discs in $pcm^{14}$ could be due to apoptosis due to widespread disruption of gene expression.

### 7.3.6 Hypothetical model for the role of Pacman and Simjang in wing imaginal discs

Previous work, outlined above, suggests that p66 plays a facilitative role in the NuRD complex as it is able to bind to many of the other members of the complex, as well as histones. In cell culture, reduction in the level of p66 interferes with the interaction between p66 and MBD2/3, which hinders targeting of the NuRD complex. Conversely, repression by the NuRD complex is increased when p66 expression is increased or when p66 is SUMOlated, which increases the affinity of p66 for MBD2/3. The recent work of Aguilera *et al.*, 2011, shows that the NuRD complex (via MBD3) is recruited to promoters of genes by unphosphorylated c-Jun, to prevent premature transcription. Activation of the JNK pathway leads to phosphorylation of c-Jun and loss of the c-Jun:NuRD interaction and allows transcription to occur. In this situation, the NuRD complex appears to act like a handbrake in a car, which can be removed when transcriptional requirements change. p66 appears to moderate the strength of the brake through its interaction with MBD2/3. In situations where the level of p66 is moderately increased, such as in *D. melanogaster* $pcm^{5}$ L3 wing imaginal discs, development can still progress, but not as efficiently, akin to driving with the handbrake partially on. The level of p66 is post-transcriptionally dependent on Pacman, which shows a novel link between post-transcriptional and epigenetic control of gene expression. Reduction of *pacman* in wing imaginal discs leads to phenotypes in the

wings and thorax (in both mutant and RNAi lines) and this could be due to the increase in Simjang level. Experimental evidence shows Pacman does not act as a simple "shredder" where all used mRNAs go to be destroyed. Pacman acts in concert with many other proteins and its associations change depending on cellular context. The results of this thesis, and other work on XRN1, show that XRN1/Pacman is capable of discriminating between mRNA transcripts. Exactly what targets transcripts for Pacman-dependent degradation is not known, but it is clear that mutations in Pacman/XRN1 lead to post-transcriptional misregulation of some genes, such as p66. p66 is directly targeted by (or to) Pacman for degradation that is required for correct development, and can apparently not be redundantly performed by the exosome.

Hypothetical model for Pacman and Simjang wild-type wing disc development (Figure 7.3): Prior to pupariation, the NuRD complex represses genes that will be expressed during eversion and later wing development. A signal around the start of pupariation (possibly Ecdysone) removes the repression to allow expression of genes required for wing development. Pacman post-transcriptionally controls the level of Simjang pre-pupariation, and may actively reduce Simjang levels during pupariation/pupation.

### 7.3.7 Predictions from the model

#### 7.3.7.1 Gene expression profiles during wing disc development

There are a modest number of genes expressed at different levels in wild-type L3 wing imaginal discs compared to *pcm*$^{5}$ mutant wing discs. If this model is correct, the level and number of genes differentially expressed should increase during metamorphosis, as the wing discs turn into the adult structures. The increase in p66 expression in *pcm*$^{5}$ mutants should prevent the cells fully responding to the signal to remove NuRD repression.

#### 7.3.7.2 *pacman* phenotypes could be rescued by *simjang* mutations

If the phenotypes in *pcm* mutations occur due to an increase in *simj* expression as the model proposes, then the phenotypes may be able to be rescued using a *simj* mutation. As the increase in the level of *simj* in *pcm*$^{5}$ wing imaginal discs is modest, it seems likely that a weak allele of *simj* should be able to rescue the *pcm*$^{5}$ phenotypes. Alternatively, removal of a single copy of *simj* (i.e. *simj/+*) may be sufficient to rescue the *pcm*$^{5}$ phenotype.

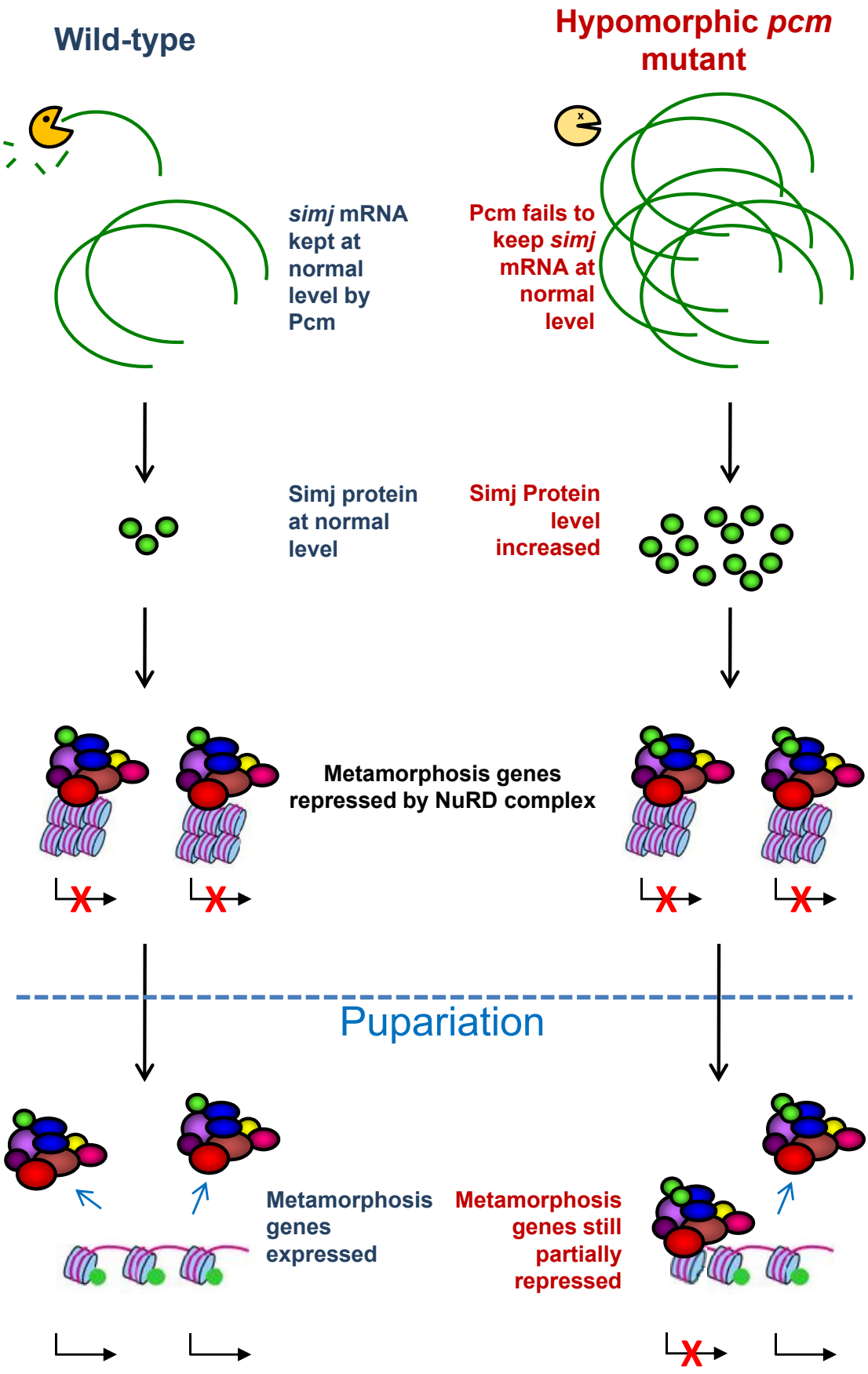



**Figure 7.3**: Hypothetical model for Pacman and Simjang during wing disc development. In wild-type Pcm keeps the level of *simj* low, and repression of genes by the NuRD complex is lifted around pupariation to allow wing development. In the hypomorphic *pcm* mutant, Simj levels increase as Pcm fails to efficiently degrade *simj* mRNA. This leads to extra repression by the NuRD complex that is not fully lifted to allow wing development to progress as normal.

### 7.3.8 The hypothesis that Pacman affects the JNK pathway

Previous work from the Newbury lab has led to the hypothesis that *pcm* mutant phenotypes are produced by misregulation of the JNK pathway due to reduced Pcm function (Grima *et al.*, 2008, Sullivan, 2008). This hypothesis formed a starting point for Chapters 5 and 6 of this thesis and was tested by using qPCR, either looking for a direct effect on JNK mRNAs, or for an indirect effect via miRNAs. However, the results obtained were consistent with the null hypothesis that Pacman does not affect the JNK pathway. The data obtained from the mRNA and miRNA arrays also supported the null hypothesis, as no JNK pathway transcripts or miRNAs were misexpressed in *pcm* mutant wing imaginal discs. However, while it is clear Pacman does not affect the JNK pathway directly or via miRNAs, recent work has shown that the NuRD complex does directly interact with the JNK pathway at the level of transcription (Aguilera *et al.*, 2011). As Pacman directly affects the level of the NuRD complex protein p66, there is an indirect link between Pacman and the JNK pathway that may explain the similar phenotypes seen for JNK pathway/*pcm* mutants. This indirect interaction could be tested by comparing expression of JNK target genes during later wing development in $pcm^5$ and wild-type.

### 7.3.9 Medical significance of the interaction between *pacman* and *simjang*

Components of the NuRD complex have received a great deal of attention in regards to human cancers, particularly the MTA proteins and HDACs. *MTA1* (metastasis-associated gene 1), was first identified due to its overexpression in mouse mammary adenocarcinoma cell lines. Since then it has been found to be highly expressed in many cancer lines and is a potential target for cancer treatment, as downregulation of *MTA1* in cell lines (e.g. by RNAi) inhibits further growth in metastatic cells (Toh and Nicolson, 2009). Inhibition of HDACs has also shown promise for cancer treatment and several compounds are under clinical trial (Hoshino and Matsubara, 2010). However, the mechanism by which HDAC inhibition counteracts the progression of cancer is unclear, especially since not all HDACs target histones (De Ruijter *et al.*, 2003). While the relationship between *pcm* and *simj* presented in this thesis is somewhat removed from the field of human oncology, it does highlight a previously unknown link between post-transcriptional gene regulation and NuRD complex function. This link, like any other biological interaction, may have the potential to be exploited for treatment of relevant pathologies.

## 7.4 Future work

The results from this thesis indicate a number of leads that can be followed to further investigate the effect of Pacman in wing imaginal discs (and other tissues), particularly regarding the relationship between Pacman and Simjang. There is also work that can (and should) be done to further characterise the available *pcm* alleles.

### 7.4.1 Molecular analysis of the function of Pacman in *pacman* mutants

The genetic experiments reported in this thesis provide strong evidence that $pcm^{5}$ is a hypomorphic mutant and $pcm^{14}$ is a null mutant. However, it would be prudent to verify this using a molecular method. Using a method such as qPCR to assay the level of a Pacman target also has the advantage of allowing relative levels of function to be determined. Hypothetically for this experiment, the highest level of function should be the wild-type level. The lowest level of Pacman function should be found in a known null, such as *Df(1)ED7452*. The level of function of $pcm^{14}$ should be indistinguishable from the level in *Df(1)ED7452* and the level of function in $pcm^{5}$ should be intermediate between wild-type and *Df(1)ED7452*.

There are several criteria that need to be fulfilled by the mRNA that is to be quantified to assay Pacman function. 1) It should be an mRNA only degraded only by Pacman, and not by the exosome. 2) For cross simplicity, it should be located on a different chromosome to *pcm* (i.e. not the *X* chromosome). 3) If the mRNA is expressed from an inserted construct or mutant gene, the insertion/mutation should be homozygous viable. 4) Expression of the mRNA should be detectable in L3 larvae to facilitate ease of collection of RNA samples.

An mRNA that fulfils all these criteria is the *Alcohol dehydrogenase* (*Adh*) allele $Adh^{fn6}$. $Adh^{fn6}$ contains a mutation to the splice site at the end of the first exon of *Adh*, which prevents the first intron being spliced out. The first intron contains a premature stop codon and the $Adh^{fn6}$ mRNA undergoes nonsense-mediated decay (Benyajati *et al.*, 1982, Brogna, 1999). As NMD in *D. melanogaster* involves cleavage of the mRNA, two sections are formed, one that undergoes degradation by Pacman and one that undergoes degradation by the exosome. The level of the section degraded by Pacman can be measured and compared between wild-type and mutants. $Adh^{fn6}$ fulfils the other criteria as it is located on chromosome *2*, is homozygous viable and is expressed in L3 larvae. The phenotype of $Adh^{fn6}$

is a significant reduction in the level of ethanol that can be tolerated in the food media by the flies (7% maximum as opposed to 20%).

To create $Adh^{fn6}$ and *pcm* double mutants takes six generations. These flies have been created and are have been molecularly verified with PCR to check the *pcm* mutations, and sequencing to check the *Adh* mutation. The L3 larvae used to for assaying *pcm* function will be $pcm^{14/5}/Y$ (or $pcm^{14/5}$) ; $Adh^{fn6}$, and these will be compared to the stock used as a source of $Adh^{fn6}$ (BL1983, $Adh^{fn6}$ $cn^{1}$ ; $ry^{506}$). Unfortunately no deficiency can be used as an example of a *pcm* null, as all (including *Df(1)ED7452*) are embryonic lethal. This only allows for a comparison between wild-type and $pcm^{5}/pcm^{14}$. This cannot show for certain that $pcm^{14}$ is acting as a null, but it is expected to show that $pcm^{14}$ has a lower level of Pacman function than $pcm^{5}$ as it is a stronger allele. Figure 7.4 shows an example workflow for this experiment.

This proposed experiment is similar to the method used in a paper investigating the decay machinery required for RNA interference (Orban and Izaurralde, 2005). In this paper, reporter constructs (including an *Adh* RNA) were targeted by RNAi in cell lines in which various degradation components, such as XRN1, had been depleted, and the resulting degradation products were visualised and quantified using Northern blotting. The experiment proposed here will be limited to XRN1 function, but has the advantages of being conducted *in vivo* and quantified by sqPCR (or TaqMan qPCR), which should be simpler and more accurate than Northern blotting.

A preliminary pilot experiment performed shortly before completion of the final version of this thesis has indicated that the level of *Adh* mRNA is around 60-fold higher in wild-type larvae compared to the level of $Adh^{fn6}$ mRNA in the mutant. In $pcm^{5}$ ; $Adh^{fn6}$ larvae, the level of the 3' fragment of $Adh^{fn6}$ mRNA is present at a 6-fold higher level than detected in the $Adh^{fn6}$ single mutant, suggesting that the level of Pcm function in the $pcm^{5}$ mutant is around 90% that of wild-type (at 25°C). In $pcm^{14}$ ; $Adh^{fn6}$ larvae, the 3' fragment of the $Adh^{fn6}$ mRNA is present at the same level as the *Adh* mRNA in wild-type, suggesting that there is no Pcm function in the $pcm^{14}$ mutant.

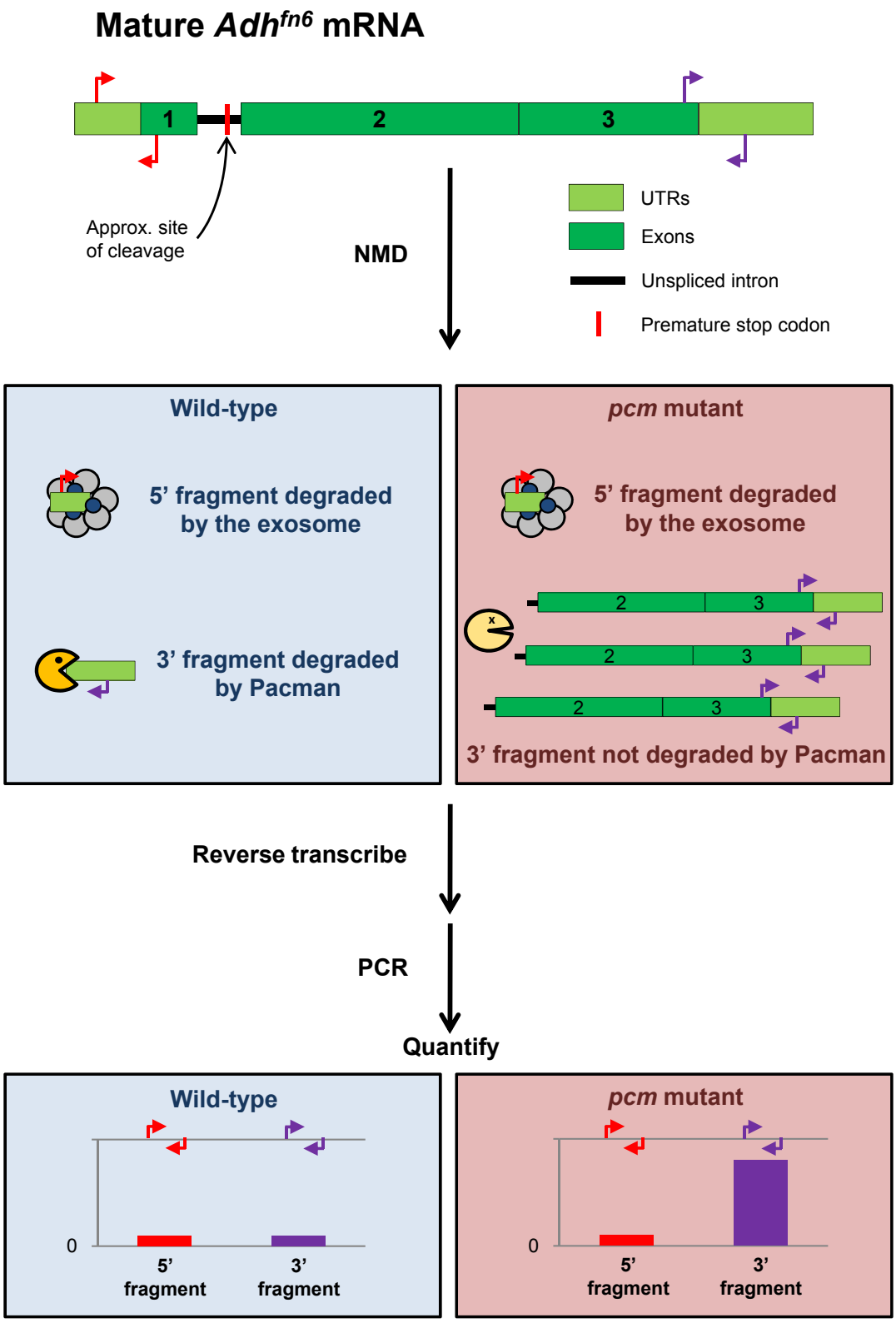



**Figure 7.4**: Relative quantification of Pacman function in *pacman* mutants using an allele of *Alcohol dehydrogenase* that undergoes nonsense-mediated decay.
The *Adh*$^{fn6}$ allele contains a mutation that inactivates the first splice site, and the first intron is not spliced out. As the first intron contains a premature stop codon, the transcript is cleaved to form two fragments. In wild-type, the 5' and 3' fragments are rapidly degraded by the exosome and Pacman respectively. In the *pcm* mutant, the 5' fragment is rapidly degraded by the exosome, but the 3' fragment should build up due to the *pcm* mutation. The relative levels of each fragment can be determined by PCR (quantitative or semi-quantitative) to compare the level of Pacman function between wild-type and the mutant strains. Red arrows indicate primer sites to detect the 5' fragment and purple arrows indicate primer sites to detect the 3' fragment.

## 7.4.2 The relationship between Pacman and Simjang

### 7.4.2.1 Half-life of *simjang* mRNA in *pcm*$^5$ mutants L3 wing imaginal discs

If *simjang* is a direct target of Pacman then when Pacman function is impaired, the half-life of *simjang* mRNA should increase. The half-life of mRNAs can be tested *in vivo* at certain time points in a situation where transcription has been stopped. It is possible to stop transcription in excised L3 wing imaginal discs by incubating them with the toxin α-amanitin, which specifically prevents the action of RNA Pol II (Chafin *et al.*, 1995). Excised imaginal discs can be kept alive for many hours in a suitable medium (Robb, 1969, Fristrom *et al.*, 1973). The plan for this experiment involves extraction of RNA from wing disc samples that have been incubated with and without α-amanitin for 0, 0.5, 1, 2 and 4 hours ($t_0$, $t_{0.5}$ etc.) from wild-type and *pcm*$^5$ mutants. Three time points will be done together (e.g. $t_0$, $t_{0.5}$, $t_1$ and $t_0$, $t_2$, $t_4$) with 20 wing imaginal discs per time point. 60 discs will be collected at the start and be added to three separate 1.5ml tubes so that each tube has the same average time since extraction, i.e. disc 1 to tube 1, disc 2 to tube 2, disc 3 to tube 3, disc 4 to tube 1, disc 5 to tube 2 etc., until each tube contains 20 discs. Extraction of 60 wing discs takes around 70 minutes, so the average time since extraction for discs in each tube will be 35 minutes. α-amanitin will then be added to the samples, to a final concentration of 1ng/µl (or water will be added in to negative control samples). $t_0$ will be immediately frozen in $LN_2$, followed by the other two samples at their specified time points. RNA will then be extracted from each sample at a later time. 20 wing discs produce around 1µg total RNA, which is sufficient to allow qPCR to be performed on at least two mRNA transcripts, which will be *simjang* and the internal control *rp49*. A pilot experiment performed on wild-type samples has shown that α-amanitin added to a concentration of 0.5ng/µl, causes the *simjang* level to drop to 65% after 1 hour, and 55% after 2 hours, with no change to the level of *rp49*. The rate of degradation should be lower in *pcm*$^5$ L3 wing discs treated in the same way.

#### 7.4.2.2 *simjang* mRNA and protein level in *pcm*$^{14}$ L3 wing imaginal discs

The level of *simjang* mRNA has been tested multiple times by qPCR in *pcm*$^{5}$ wing discs to determine that the increase in level compared to wild-type is 3.8-fold. The level has not yet been tested in *pcm*$^{14}$ wing discs, but would be predicted to be increased by a greater amount as *pcm*$^{14}$ is a stronger allele. The level in *pcm*$^{14}$ discs can be tested in the same way as it has been in *pcm*$^{5}$ discs, with the only consideration being that *pcm*$^{14}$ discs are half the size of wild-type wing discs. An RNA extraction from a given number of *pcm*$^{14}$ wing discs will likely give roughly half the amount of RNA that could be extracted from the same number of wild-type or *pcm*$^{5}$ wing discs. While it is possible to do qPCR on samples at different starting concentrations, it is good practice to use the same concentration/amount of RNA for all samples. The samples could be concentrated or diluted to equalise the concentrations, but it is also good practice to keep manipulations to RNA samples as consistent as possible across samples. Therefore, the best way to get roughly the same RNA concentration, which can then be modified by a small amount to equalise samples, would be to extract twice the number of *pcm*$^{14}$ wing discs compared to the number extracted for wild-type or *pcm*$^{5}$. 60 wild-type and 120 *pcm*$^{14}$ wing discs per replicate would give sufficient mRNA for this test.

The level of Simjang protein in L3 wing discs of *pcm*$^{5}$ and wild-type has only been measured once using a Western blot, so requires replication, as quantification of Western blots can be variable. At the same time, the level of Simjang protein could be tested in *pcm*$^{14}$ wing discs, with the same considerations as above – twice the number of *pcm*$^{14}$ wing discs should be used compared to wild-type and *pcm*$^{5}$. 30 wild-type and 60 *pcm*$^{14}$ wing discs per replicate would give sufficient material for a Western blot.

This experiment makes the assumption that *pcm*$^{14}$ wing discs develop far enough to be comparable to *pcm*$^{5}$ L3 wing discs. However, it is not yet known if this is the case and it may be that they have not developed very far, or have undergone apoptosis. If either of these is the case, wing discs at the L3 stage will not be a valid comparison. *pcm*$^{14}$ discs may need to be analysed at an earlier stage to determine when development starts to go awry. Whatever causes the small discs in *pcm*$^{14}$ mutants could conceivably not have anything to do with Simjang, if the defect occurs much earlier than in *pcm*$^{5}$ mutant wing discs.

### 7.4.3 The effect of increased Simjang levels

Both *pcm*$^{5}$ and *pcm*$^{14}$ L3 wing discs are significantly smaller than equivalent wild-type discs, which are made up of around 60,000 cells (Gilbert, 2000). Small wing discs could theoretically occur in a number of ways, such as lack of proliferation, apoptosis or the cells present being smaller than normal. The first step would be to stain the nuclei of cells with a DNA stain such as DAPI, and count the number of cells present in wild-type and mutant discs. The discs could then be stained with anti-activated caspase-3 antibody, which will show any cells undergoing apoptosis. This method was used by Zhu *et al*., 2008 to show that *hdac1* (*rpd3*) mutant wing imaginal discs are small due to lack of proliferation and *hdac3* wing discs are small due to apoptosis. Very little staining by anti-activated caspase-3 was observed in wild-type or *hdac1* mutant discs, while extensive staining was observed in *hdac3* mutant discs. To test if the cells are smaller than normal, the cell membranes can be stained with a suitable dye and the area of cells measured.

### 7.4.4 Targeting of transcripts to Pacman

Typically, proteins that target mRNAs interact with them at certain sequences or via a secondary structure that forms as a result of the primary nucleotide sequence. There is likely some element(s) in *simj* mRNA that causes it to be targeted by or to Pcm. Elucidating what sequence(s) cause *simj* to be degraded by Pcm could be approached bioinformatically by comparing a number of transcripts known to be targeted by Pcm to determine areas of similarity. However, as *simj* is currently the only transcript that is known to be directly targeted by Pcm, there is nothing to compare it to. *CG31477* is a possibility, but as regulation may not be entirely post-transcriptional, comparing *simj* and *CG31477* may not be valid. There are a number of other possibilities from the imaginal discs array results that could be direct targets of Pcm and also many possibilities from the testes array data. Future work as has been performed in this thesis is required to find more direct targets of Pcm, and when they are found, bioinformatics can be used to attempt to find similarities between them.

In parallel, molecular techniques can be employed to identify RNA-stability elements in *simj*. *simj* constructs can be made with sections substituted and measured by qPCR to identify regions that promote or reduce stability. For example, a *simj* construct can be engineered to replace the 3’ UTR of *simj* with the 3’ UTR of a gene not under the influence of Pcm. If the region that targets *simj* to Pcm lies within the 3’ UTR then the level if the new construct should be unaffected in *pcm* mutants. Reciprocally, if the *simj* 3’ UTR contains a

stability element then this should affect the stability of an otherwise unaffected mRNA with the *simj* 3' UTR added to it. This can be repeated with the 5'UTR and sections of the coding sequence.

### 7.4.5 The action of Simjang/p66

Compared to the testes mRNA array results, the overall number and magnitudes of mRNAs that change in level in L3 wing imaginal discs are much lower. It is not surprising that disparate tissues are affected differently by a mutation, but it would be interesting to piece together the pathways that cause the effects in each tissue, and determine how such different effects occur from a common source. This may be easier to achieve in wing imaginal discs, as their further development can be relatively easily monitored. The effect of an increase in p66 expression in L3 wing imaginal discs may be relatively small, but it may also be magnified as development progresses. If the NuRD complex is acting to repress genes required during metamorphosis, this small effect could translate into a large effect on numerous genes. To test this, further developed wing imaginal discs could be extracted from wild-type and $pcm^5$ mutants to compare their mRNA levels by microarray or deep sequencing. This would be a simple way to see if the increase in level of p66 in L3 wing imaginal discs in $pcm^5$ mutants leads to more profound gene expression changes later in wing/thorax development, potentially as a global downregulation of many genes required for development. This could also add significantly to the general knowledge of *D. melanogaster* wing development, as at the time of writing there is only one published microarray study on global comparisons of gene expression between L3 wing discs and developing wings (Ren *et al.*, 2005). Genes may be misexpressed due to the reduction in *pcm* function or the increase in *simj*, but it should be possible to distinguish between the two, as it is possible to distinguish between transcriptional and post-transcriptional changes by comparing pre-mRNA and mature mRNA levels.

## 7.5 Concluding remarks

The results presented in this thesis have shown that the exoribonuclease XRN1 (Pacman) is capable of selectively degrading specific mRNAs to reduce their expression level. The example investigated in most detail was the post-transcriptional increase in expression of the NuRD complex protein Simjang in *pcm* mutant L3 wing imaginal discs. This shows a novel link between post-transcriptional regulation and epigenetic regulation of gene expression. Comparison of microarray results between $pcm^5$ wing imaginal discs and $pcm^5$ testes indicates that the role of Pacman is tissue specific, but whatever system is targeting transcripts to Pacman may well be consistent between tissues.

# 8 References

AFFYMETRIX (2004) Technical Note: GeneChip Expression Platform: Comparison, Evolution, and Performance. Santa Clara, Affymetrix.

AGNES, F., SUZANNE, M. & NOSELLI, S. (1999) The Drosophila JNK pathway controls the morphogenesis of imaginal discs during metamorphosis. *Development,* 126**,** 5453-5462.

AGUILERA, C., NAKAGAWA, K., SANCHO, R., CHAKRABORTY, A., HENDRICH, B. & BEHRENS, A. (2011) c-Jun N-terminal phosphorylation antagonises recruitment of the Mbd3/NuRD repressor complex. *Nature,* 469**,** 231-235.

AMMIT, A. J. (2005) The role of mRNA stability in airway remodelling. *Pulmonary Pharmacology & Therapeutics,* 18**,** 405-415.

AMRANI, N., GANESAN, R., KERVESTIN, S., MANGUS, D. A., GHOSH, S. & JACOBSON, A. (2004) A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature,* 432**,** 112-118.

APPLIED BIOSYSTEMS (LIFE TECHNOLOGIES) (2010a) Introduction to Gene Expression: Getting Started Guide. rev A ed. Carlsbad, Life Technologies.

APPLIED BIOSYSTEMS (LIFE TECHNOLOGIES) (2010b) TaqMan MicroRNA Assays Product Overview. Carlsbad, Life Technologies.

ASHBURNER, M. (1989) *Drosophila A Laboratory Handbook, 1st ed.,* New York, Cold Spring Harbor Laboratory Press.

ASLAM, N. & ZAHEER, I. (2010) The biosynthesis characteristics of TTP and TNF can be regulated through a posttranscriptional molecular loop. *Journal of Biological Chemistry*.

AUERBACH, C. (1936) The development of the legs, wings, and halteres in wild type and some mutant strains of *Drosophila melanogaster*. *Transactions of the Royal Society of Edinburgh,* 58**,** 787-815.

AYER, D. E., LAWRENCE, Q. A. & EISENMAN, R. N. (1995) Mad-max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. *Cell,* 80**,** 767-776.

BADIS, G., SAVEANU, C., FROMONT-RACINE, M. & JACQUIER, A. (2004) Targeted mRNA Degradation by Deadenylation-Independent Decapping. *Molecular Cell,* 15**,** 5-15.

BAIL, S., SWERDEL, M., LIU, H., JIAO, X., GOFF, L. A., HART, R. P. & KILEDJIAN, M. (2010) Differential regulation of microRNA stability. *RNA,* 16**,** 1032-1039.

BAINBRIDGE, S. P. & BOWNES, M. (1981) Staging the metamorphosis of *Drosophila melanogaster*. *Journal of Embryology and Experimental Morphology,* 66**,** 57-80.

BALAGOPAL, V. & PARKER, R. (2009) Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. *Curr. Opin. Cell Biol.,* 21**,** 403-408.

BALLESTAR, E., PILE, L. A., WASSARMAN, D. A., WOLFFE, A. P. & WADE, P. A. (2001) A *Drosophila* MBD family member is a transcriptional corepressor associated with specific genes. *European Journal of Biochemistry,* 268**,** 5397-5406.

BARTEL, D. P. (2004) MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell,* 116**,** 281-297.

BASHKIROV, V. I., SCHERTHAN, H., SOLINGER, J. A., BUERSTEDDE, J.-M. & HEYER, W.-D. (1997) A Mouse Cytoplasmic Exoribonuclease (mXRN1p) with Preference for G4 Tetraplex Substrates. *The Journal of Cell Biology,* 136**,** 761-773.

BEJARANO, F., SMIBERT, P. & LAI, E. C. (2010) miR-9a prevents apoptosis during wing development by repressing *Drosophila* LIM-only. *Dev. Biol.,* 338**,** 63-73.

BENYAJATI, C., PLACE, A. R., WANG, N., PENTZ, E. & SOFER, W. (1982) Deletions at intervening sequence splice sites in the alcohol dehydrogenase gene of *Drosophila*. *Nucleic Acids Research,* 10**,** 7261-7272.

BHATTACHARYYA, S. N., HABERMACHER, R., MARTINE, U., CLOSS, E. I. & FILIPOWICZ, W. (2006) Relief of microRNA-Mediated Translational Repression in Human Cells Subjected to Stress. *Cell,* 125**,** 1111-1124.

BHUVANAGIRI, M., SCHLITTER, A. M., HENTZE, M. W. & KULOZIK, A. E. (2010) NMD: RNA biology meets human genetic medicine. *Biochemical Journal,* 430**,** 365-377.

BINGHAM, A. J., OOI, L., KOZERA, L., WHITE, E. & WOOD, I. C. (2007) The Repressor Element 1-Silencing Transcription Factor Regulates Heart-Specific Gene Expression Using Multiple Chromatin-Modifying Complexes. *Mol. Cell. Biol.,* 27**,** 4082-4092.

BIRYUKOVA, I., ASMAR, J., ABDESSELEM, H. & HEITZLER, P. (2009) *Drosophila* mir-9a regulates wing development via fine-tuning expression of the LIM only factor, dLMO. *Dev. Biol.,* 327**,** 487-496.

BRACKERTZ, M., BOEKE, J., ZHANG, R. & RENKAWITZ, R. (2002) Two Highly Related p66 Proteins Comprise a New Family of Potent Transcriptional Repressors Interacting with MBD2 and MBD3. *Journal of Biological Chemistry,* 277**,** 40958-40966.

BRACKERTZ, M., GONG, Z., LEERS, J. & RENKAWITZ, R. (2006) p66α and p66β of the Mi-2/NuRD complex mediate MBD2 and histone interaction. *Nucleic Acids Research,* 34**,** 397-406.

BREITLING, R., AMTMANN, A. & HERZYK, P. (2004a) Iterative Group Analysis (iGA): A simple tool to enhance sensitivity and facilitate interpretation of microarray experiments. *BMC Bioinformatics,* 5**,** 34.

BREITLING, R., ARMENGAUD, P., AMTMANN, A. & HERZYK, P. (2004b) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Letters,* 573**,** 83-92.

BRENGUES, M., TEIXEIRA, D. & PARKER, R. (2005) Movement of Eukaryotic mRNAs Between Polysomes and Cytoplasmic Processing Bodies. *Science,* 310**,** 486-489.

BRENNAN, C. M. & STEITZ, J. A. (2001) HuR and mRNA stability. *Cell Mol. Life Sci.,* 58**,** 266-77.

BRIDGES, C. B. (1916) NON-DISJUNCTION AS PROOF OF THE CHROMOSOME THEORY OF HEREDITY. *Genetics,* 1**,** 1-52.

BROGNA, S. (1999) Nonsense mutations in the alcohol dehydrogenase gene of *Drosophila melanogaster* correlate with an abnormal 3' end processing of the corresponding pre-mRNA. *RNA,* 5**,** 562-573.

BROWN, C. E., TARUN, S. Z., JR., BOECK, R. & SACHS, A. B. (1996) PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae. *Mol. Cell. Biol.,* 16**,** 5744-5753.

BUSTIN, S. A., BENES, V., GARSON, J. A., HELLEMANS, J., HUGGETT, J., KUBISTA, M., MUELLER, R., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J. & WITTWER, C. T. (2009) The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin Chem,* 55**,** 611-622.

CAMBON, A., KHALYFA, A., COOPER, N. & THOMPSON, C. (2007) Analysis of probe level patterns in Affymetrix microarray data. *BMC Bioinformatics,* 8**,** 146.

CAO, D. & PARKER, R. (2003) Computational Modeling and Experimental Analysis of Nonsense-Mediated Decay in Yeast. *Cell,* 113**,** 533-545.

CARBALLO, E. & BLACKSHEAR, P. J. (2001) Roles of tumor necrosis factor-α receptor subtypes in the pathogenesis of the tristetraprolin-deficiency syndrome. *Blood,* 98**,** 2389-2395.

CHAFIN, D. R., GUO, H. & PRICE, D. H. (1995) Action of α-Amanitin during Pyrophosphorolysis and Elongation by RNA Polymerase II. *Journal of Biological Chemistry,* 270**,** 19114-19119.

CHANG, J. H., XIANG, S., XIANG, K., MANLEY, J. L. & TONG, L. (2011) Structural and biochemical studies of the 5'-3' exoribonuclease Xrn1. *Nat Struct Mol Biol,* 18**,** 270-276.

CHATTERJEE, S., FASLER, M., BÜSSING, I. & GROßHANS, H. (2011) Target-Mediated Protection of Endogenous MicroRNAs in *C. elegans*. *Developmental Cell,* 20**,** 388-396.

CHATTERJEE, S. & GROßHANS, H. (2009) Active turnover modulates mature microRNA activity in *Caenorhabditis elegans*. *Nature,* 461**,** 546-9.

CHEN, C.-Y., GHERZI, R., ONG, S.-E., CHAN, E. L., RAIJMAKERS, R., PRUIJN, G. J. M., STOECKLIN, G., MORONI, C., MANN, M. & KARIN, M. (2001) AU Binding Proteins Recruit the Exosome to Degrade ARE-Containing mRNAs. *Cell,* 107**,** 451-464.

CHEN, C.-Y. A. & SHYU, A.-B. (2003) Rapid Deadenylation Triggered by a Nonsense Codon Precedes Decay of the RNA Body in a Mammalian Cytoplasmic Nonsense-Mediated Decay Pathway. *Mol. Cell. Biol.,* 23**,** 4805-4813.

CHEN, G., FERNANDEZ, J., MISCHE, S. & COUREY, A. J. (1999) A functional interaction between the histone deacetylase Rpd3 and the corepressor Groucho in Drosophila development. *Genes & Development,* 13**,** 2218-2230.

CONTI, E. & IZAURRALDE, E. (2005) Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Current Opinion in Cell Biology,* 17**,** 316-325.

COPE, L. M., IRIZARRY, R. A., JAFFEE, H. A., WU, Z. & SPEED, T. P. (2004) A benchmark for Affymetrix GeneChip expression measures. *Bioinformatics,* 20**,** 323-331.

COUTTET, P. & GRANGE, T. (2004) Premature termination codons enhance mRNA decapping in human cells. *Nucleic Acids Research,* 32**,** 488-494.

CRICK, F. (1970) CENTRAL DOGMA OF MOLECULAR BIOLOGY. *Nature,* 227**,** 561-563.

CURRY, S., KOTIK-KOGAN, O., CONTE, M. R. & BRICK, P. (2009) Getting to the end of RNA: Structural analysis of protein recognition of 5' and 3' termini. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1789**,** 653-666.

DE BENEDETTI, A. & GRAFF, J. R. (2004) eIF-4E expression and its role in malignancies and metastases. *Oncogene,* 23**,** 3189-3199.

DE RUIJTER, A. J. M., VAN GENNIP, A. H., CARON, H. N., KEMP, S. & VAN KUILENBURG, A. B. P. (2003) Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochemical Journal,* 370**,** 737-749.

DEXTER, J. S. (1912) On Coupling of Certain Sex-Linked Characters in *Drosophila*. *Biological Bulletin,* 23**,** 183-194.

DIECI, G., PRETI, M. & MONTANINI, B. (2009) Eukaryotic snoRNAs: A paradigm for gene expression flexibility. *Genomics,* 94**,** 83-88.

DIEDERICHS, S. & HABER, D. A. (2007) Dual Role for Argonautes in MicroRNA Processing and Posttranscriptional Regulation of MicroRNA Expression. *Cell,* 131**,** 1097-1108.

DOMA, M. K. & PARKER, R. (2006) Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature,* 440**,** 561-564.

DUFFY, J. B. (2002) GAL4 system in Drosophila: A fly geneticist's Swiss army knife. *Genesis,* 34**,** 1-15.

EBERLE, A. B., LYKKE-ANDERSEN, S., MUHLEMANN, O. & JENSEN, T. H. (2009) SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat Struct Mol Biol,* 16**,** 49-55.

EL HAGE, A., KOPER, M., KUFEL, J. & TOLLERVEY, D. (2008) Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. *Genes & Development,* 22**,** 1069-1081.

EULALIO, A., BEHM-ANSMANT, I., SCHWEIZER, D. & IZAURRALDE, E. (2007) P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. *Mol. Cell. Biol.,* 27**,** 3970-3981.

EXQION (2008) microRNA Array benchmark study. Vedbaek, Exiqon.

FENG, Q., CAO, R., XIA, L., ERDJUMENT-BROMAGE, H., TEMPST, P. & ZHANG, Y. (2002) Identification and Functional Characterization of the p66/p68 Components of the MeCP1 Complex. *Mol. Cell. Biol.,* 22**,** 536-546.

FLICEK, P., AKEN, B. L., BALLESTER, B., BEAL, K., BRAGIN, E., BRENT, S., CHEN, Y., CLAPHAM, P., COATES, G., FAIRLEY, S., FITZGERALD, S., FERNANDEZ-BANET, J., GORDON, L., GRÃ¤F, S., HAIDER, S., HAMMOND, M., HOWE, K., JENKINSON, A., JOHNSON, N., KÃ¤HÃ¤RI, A., KEEFE, D., KEENAN, S., KINSELLA, R., KOKOCINSKI, F., KOSCIELNY, G., KULESHA, E., LAWSON, D., LONGDEN, I., MASSINGHAM, T., MCLAREN, W., MEGY, K., OVERDUIN, B., PRITCHARD, B., RIOS, D., RUFFIER, M., SCHUSTER, M., SLATER, G., SMEDLEY, D., SPUDICH, G., TANG, Y. A., TREVANION, S., VILELLA, A., VOGEL, J., WHITE, S., WILDER, S. P., ZADISSA, A., BIRNEY, E., CUNNINGHAM, F., DUNHAM, I., DURBIN, R., FERNÃ¡NDEZ-SUAREZ, X. M., HERRERO, J., HUBBARD, T. J. P., PARKER, A., PROCTOR, G., SMITH, J. & SEARLE, S. M. J. (2010) Ensembl's 10th year. *Nucleic Acids Research,* 38**,** D557-D562.

FRISCHMEYER, P. A., VAN HOOF, A., O'DONNELL, K., GUERRERIO, A. L., PARKER, R. & DIETZ, H. C. (2002) An mRNA Surveillance Mechanism That Eliminates Transcripts Lacking Termination Codons. *Science,* 295**,** 2258-2261.

FRISTROM, J. W., LOGAN, W. R. & MURPHY, C. (1973) The synthetic and minimal culture requirements for evagination of imaginal discs of *Drosophila melanogaster in vitro*. *Developmental Biology,* 33**,** 441-456.

GARNEAU, N. L., WILUSZ, J. & WILUSZ, C. J. (2007) The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol,* 8**,** 113-126.

GATFIELD, D. & IZAURRALDE, E. (2004) Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. *Nature,* 429**,** 575-578.

GHERZI, R., LEE, K.-Y., BRIATA, P., WEGMÜLLER, D., MORONI, C., KARIN, M. & CHEN, C.-Y. (2004) A KH Domain RNA Binding Protein, KSRP, Promotes ARE-Directed mRNA Turnover by Recruiting the Degradation Machinery. *Molecular Cell,* 14**,** 571-583.

GILBERT, S. F. (2000) *Developmental Biology, 6th ed.,* Massachusetts, Sinauer Associates.

GOLIC, K. G. & GOLIC, M. M. (1996) Engineering the Drosophila Genome: Chromosome Rearrangements by Design. *Genetics,* 144**,** 1693-1711.

GONG, Z., BRACKERTZ, M. & RENKAWITZ, R. (2006) SUMO Modification Enhances p66-Mediated Transcriptional Repression of the Mi-2/NuRD Complex. *Mol. Cell. Biol.,* 26**,** 4519-4528.

GREENSPAN, R. J. (2004) *Fly Pushing: The Theory and Practice of Drosophila Genetics, 2nd ed.,* New York, Cold Spring Harbor Press.

GRIFFITHS-JONES, S., GROCOCK, R. J., VAN DONGEN, S., BATEMAN, A. & ENRIGHT, A. J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucl. Acids Res.,* 34**,** D140-144.

GRIFFITHS-JONES, S., SAINI, H. K., VAN DONGEN, S. & ENRIGHT, A. J. (2008) miRBase: tools for microRNA genomics. *Nucl. Acids Res.,* 36**,** D154-158.

GRIMA, D. (2002) Analysis of the role of the exoribonuclease *pacman* in *Drosophila* development. Doctor of Philosophy thesis, Biochemistry Department, University of Oxford

GRIMA, D. P., SULLIVAN, M., ZABOLOTSKAYA, M. V., BROWNE, C., SEAGO, J., WAN, K. C., OKADA, Y. & NEWBURY, S. F. (2008) The 5' - 3' exoribonuclease *pacman* is required for epithelial sheet sealing in *Drosophila* and genetically interacts with the phosphatase *puckered*. *Biol. Cell,* 100**,** 687-701.

GRÜN, D., WANG, Y.-L., LANGENBERGER, D., GUNSALUS, K. C. & RAJEWSKY, N. (2005) microRNA Target Predictions across Seven *Drosophila* Species and Comparison to Mammalian Targets. *PLoS Comput. Biol.,* 1**,** e13.

GU, S. & KAY, M. (2010) How do miRNAs mediate translational repression? *Silence,* 1**,** 11.

HANSEN, K. D., LAREAU, L. F., BLANCHETTE, M., GREEN, R. E., MENG, Q., REHWINKEL, J., GALLUSSER, F. L., IZAURRALDE, E., RIO, D. C., DUDOIT, S. & BRENNER, S. E. (2009) Genome-Wide Identification of Alternative Splice Forms Down-Regulated by Nonsense-Mediated mRNA Decay in *Drosophila*. *PLoS Genet,* 5**,** e1000525.

HENDRICH, B., GUY, J., RAMSAHOYE, B., WILSON, V. A. & BIRD, A. (2001) Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes & Development,* 15**,** 710-723.

HENDRICH, B. & TWEEDIE, S. (2003) The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends in Genetics,* 19**,** 269-277.

HEYER, W. D., JOHNSON, A. W., REINHART, U. & KOLODNER, R. D. (1995) Regulation and intracellular localization of Saccharomyces cerevisiae strand exchange protein 1 (Sep1/Xrn1/Kem1), a multifunctional exonuclease. *Mol. Cell. Biol.,* 15**,** 2728-2736.

HO, L. & CRABTREE, G. R. (2010) Chromatin remodelling during development. *Nature,* 463**,** 474-484.

HODGKIN, J., PAPP, A., PULAK, R., AMBROS, V. & ANDERSON, P. (1989) A New Kind of Informational Suppression in the Nematode Caenorhabditis elegans. *Genetics,* 123**,** 301-313.

HOSHINO, I. & MATSUBARA, H. (2010) Recent advances in histone deacetylase targeted cancer therapy. *Surgery Today,* 40**,** 809-815.

HSU, C. L. & STEVENS, A. (1993) Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. *Mol. Cell. Biol.,* 13**,** 4826-4835.

HUNTZINGER, E., KASHIMA, I., FAUSER, M., SAULIÃ¨RE, J. R. M. & IZAURRALDE, E. (2008) SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA,* 14**,** 2609-2617.

IRIZARRY, R. A., BOLSTAD, B. M., COLLIN, F., COPE, L. M., HOBBS, B. & SPEED, T. P. (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research,* 31**,** e15.

JESKE, M., MORITZ, B., ANDERS, A. & WAHLE, E. (2011) Smaug assembles an ATP-dependent stable complex repressing nanos mRNA translation at multiple levels. *Embo J,* 30**,** 90-103.

JINEK, M., COYLE, SCOTT M. & DOUDNA, JENNIFER A. (2011) Coupled 5' Nucleotide Recognition and Processivity in Xrn1-Mediated mRNA Decay. *Molecular Cell,* 41**,** 600-608.

JOHNSON, S. A. & MILNER, M. J. (1987) The final stages of wing development in *Drosophila melanogaster*. *Tissue and Cell,* 19**,** 505-513.

JONES, C. I. & NEWBURY, S. F. (2010) Functions of microRNAs in *Drosophila* development. *Biochemical Society Transactions,* 038**,** 1137-1143.

KAI, Z. S. & PASQUINELLI, A. E. (2010) MicroRNA assassins: factors that regulate the disappearance of miRNAs. *Nat Struct Mol Biol,* 17**,** 5-10.

KASHIMA, I., JONAS, S., JAYACHANDRAN, U., BUCHWALD, G., CONTI, E., LUPAS, A. N. & IZAURRALDE, E. (2010) SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. *Genes & Development,* 24**,** 2440-2450.

KEHLE, J., BEUCHLE, D., TREUHEIT, S., CHRISTEN, B., KENNISON, J. A., BIENZ, M. & MŸLLER, J. (1998) dMi-2, a Hunchback-Interacting Protein That Functions in Polycomb Repression. *Science,* 282**,** 1897-1900.

KHERADPOUR, P., STARK, A., ROY, S. & KELLIS, M. (2007) Reliable prediction of regulator targets using 12 *Drosophila* genomes. *Genome Res.,* 17**,** 1919-1931.

KIM, J. (2002) KEM1 is involved in filamentous growth of Saccharomyces cerevisiae. *Fems Microbiology Letters,* 216**,** 33-38.

KIM, J., LJUNGDAHL, P. O. & FINK, G. R. (1990) kem Mutations Affect Nuclear Fusion in Saccharomyces cerevisiae. *Genetics,* 126**,** 799-812.

KIM, J., SIF, S., JONES, B., JACKSON, A., KOIPALLY, J., HELLER, E., WINANDY, S., VIEL, A., SAWYER, A., IKEDA, T., KINGSTON, R. & GEORGOPOULOS, K. (1999) Ikaros DNA-Binding Proteins Direct Formation of Chromatin Remodeling Complexes in Lymphocytes. *Immunity,* 10**,** 345-355.

KIM, K., ZAKHARKIN, S. O. & ALLISON, D. B. (2010) Expectations, validity, and reality in gene expression profiling. *Journal of Clinical Epidemiology,* 63**,** 950-959.

KIM, Y.-O., PARK, S.-J., BALABAN, R. S., NIRENBERG, M. & KIM, Y. (2004) A functional genomic screen for cardiogenic genes using RNA interference in developing *Drosophila* embryos. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 159-164.

KOIPALLY, J., RENOLD, A., KIM, J. & GEORGOPOULOS, K. (1999) Repression by Ikaros and Aiolos is mediated through histone deacetylase complexes. *Embo J,* 18**,** 3090-3100.

KON, C., CADIGAN, K. M., DA SILVA, S. L. & NUSSE, R. (2005) Developmental Roles of the Mi-2/NURD-Associated Protein p66 in Drosophila. *Genetics,* 169**,** 2087-2100.

KORNER, C. G., WORMINGTON, M., MUCKENTHALER, M., SCHNEIDER, S., DEHLIN, E. & WAHLE, E. (1998) The deadenylating nuclease (DAN) is involved in poly(A) tail removal during the meiotic maturation of Xenopus oocytes. *Embo J,* 17**,** 5427-5437.

KOTHAPALLI, R., YODER, S., MANE, S. & LOUGHRAN, T. (2002) Microarray results: how accurate are they? *BMC Bioinformatics,* 3**,** 22.

KULKARNI, M., OZGUR, S. & STOECKLIN, G. (2010) On track with P-bodies. *Biochemical Society Transactions,* 038**,** 242-251.

LAGGER, G., O'CARROLL, D., REMBOLD, M., KHIER, H., TISCHLER, J., WEITZER, G., SCHUETTENGRUBER, B., HAUSER, C., BRUNMEIR, R., JENUWEIN, T. & SEISER, C. (2002) Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. *Embo J,* 21**,** 2672-2681.

LARIMER, F. W. & STEVENS, A. (1990) Disruption of the gene XRN1, coding for a 5'-->3' exoribonuclease, restricts yeast cell growth. *Gene,* 95**,** 85-90.

LAWRENCE, P. A. (1992) *The Making of a Fly, 1st ed.,* Oxford, Blackwell Science.

LE GUEZENNEC, X., VERMEULEN, M., BRINKMAN, A. B., HOEIJMAKERS, W. A. M., COHEN, A., LASONDER, E. & STUNNENBERG, H. G. (2006) MBD2/NuRD and MBD3/NuRD, Two Distinct Complexes with Different Biochemical and Functional Properties. *Mol. Cell. Biol.,* 26**,** 843-851.

LEE, R. C., FEINBAUM, R. L. & AMBROS, V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell,* 75**,** 843-854.

LEIPUVIENE, R. & THEIL, E. (2007) The family of iron responsive RNA structures regulated by changes in cellular iron and oxygen. *Cellular and Molecular Life Sciences,* 64**,** 2945-2955.

LI, Y., WANG, F., LEE, J.-A. & GAO, F.-B. (2006) MicroRNA-9a ensures the precise specification of sensory organ precursors in *Drosophila*. *Gene. Dev.,* 20**,** 2793-2805.

LIN, M.-D., JIAO, X., GRIMA, D., NEWBURY, S. F., KILEDJIAN, M. & CHOU, T.-B. (2008) Drosophila processing bodies in oogenesis. *Developmental Biology,* 322**,** 276-288.

LINKER, K., PAUTZ, A., FECHIR, M., HUBRICH, T., GREEVE, J. & KLEINERT, H. (2005) Involvement of KSRP in the post-transcriptional regulation of human iNOS expression-complex interplay of KSRP with TTP and HuR. *Nucleic Acids Research,* 33**,** 4813-4827.

LIU, H. D., RODGERS, N. D., JIAO, X. & KILEDJIAN, M. (2002) The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *Embo Journal,* 21**,** 4699-4708.

LUKE, B., AZZALIN, C. M., HUG, N., DEPLAZES, A., PETER, M. & LINGNER, J. (2007) *Saccharomyces cerevisiae* Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay. *Nucleic Acids Research,* 35**,** 7688-7697.

LUO, W. & BENTLEY, D. (2004) A Ribonucleolytic Rat Torpedoes RNA Polymerase II. *Cell,* 119**,** 911-914.

LYKKE-ANDERSEN, S., BRODERSEN, D. E. & JENSEN, T. H. (2009) Origins and activities of the eukaryotic exosome. *Journal of Cell Science,* 122**,** 1487-1494.

MAKEYEV, A. V. & LIEBHABER, S. A. (2002) The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. *RNA,* 8**,** 265-278.

MANNERVIK, M. & LEVINE, M. (1999) The Rpd3 histone deacetylase is required for segmentation of the Drosophila embryo. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 6797-6801.

MARHOLD, J., KRAMER, K., KREMMER, E. & LYKO, F. (2004) The Drosophila MBD2/3 protein mediates interactions between the. *Development,* 131**,** 6033-6039.

MARINO, S. & NUSSE, R. (2007) Mutants in the Mouse NuRD/Mi2 Component P66α Are Embryonic Lethal. *PLoS ONE,* 2**,** e519.

MARTIN-BLANCO, E., PASTOR-PAREJA, J. C. & GARCÍA-BELLIDO, A. (2000) JNK and decapentaplegic signaling control adhesiveness and cytoskeleton dynamics during thorax closure in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 7888-7893.

MARTıNEZ, J., REN, Y.-G., NILSSON, P., EHRENBERG, M. & VIRTANEN, A. (2001) The mRNA Cap Structure Stimulates Rate of Poly(A) Removal and Amplifies Processivity of Degradation. *Journal of Biological Chemistry,* 276**,** 27923-27929.

MCDONEL, P., COSTELLO, I. & HENDRICH, B. (2009) Keeping things quiet: Roles of NuRD and Sin3 co-repressor complexes during mammalian development. *The International Journal of Biochemistry & Cell Biology,* 41**,** 108-116.

MEDGHALCHI, S. M., FRISCHMEYER, P. A., MENDELL, J. T., KELLY, A. G., LAWLER, A. M. & DIETZ, H. C. (2001) Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Human Molecular Genetics,* 10**,** 99-105.

METZSTEIN, M. M. & KRASNOW, M. A. (2006) Functions of the Nonsense-Mediated mRNA Decay Pathway in *Drosophila* Development. *PLoS Genet,* 2**,** e180.

MITCHELL, P. & TOLLERVEY, D. (2003) An NMD Pathway in Yeast Involving Accelerated Deadenylation and Exosome-Mediated 3'-->5' Degradation. *Molecular Cell,* 11**,** 1405-1413.

MONTGOMERY, R. L., DAVIS, C. A., POTTHOFF, M. J., HABERLAND, M., FIELITZ, J., QI, X., HILL, J. A., RICHARDSON, J. A. & OLSON, E. N. (2007) Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility. *Genes & Development,* 21**,** 1790-1802.

MOTTUS, R., SOBEL, R. E. & GRIGLIATTI, T. A. (2000) Mutational Analysis of a Histone Deacetylase in *Drosophila melanogaster*: Missense Mutations Suppress Gene Silencing Associated With Position Effect Variegation. *Genetics,* 154**,** 657-668.

MUHLRAD, D. & PARKER, R. (2005) The yeast *EDC1* mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p. *Embo J,* 24**,** 1033-1045.

MULLEN, T. E. & MARZLUFF, W. F. (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**,** 50-65.

MŸHLEMANN, O. & LYKKE-ANDERSEN, J. (2010) How and where are nonsense mRNAs degraded in mammalian cells? *RNA Biology,* 7**,** 28-32.

NEWBURY, S. & WOOLLARD, A. (2004) The 5'-3' exoribonuclease xrn-1 is essential for ventral epithelial enclosure during C. elegans embryogenesis. *RNA,* 10**,** 59-65.

NEWBURY, S. F. (2006) Control of mRNA stability in eukaryotes. *Biochem. Soc. Trans.,* 34**,** 30-34.

NICHOLSON, P. & MŸHLEMANN, O. (2010) Cutting the nonsense: the degradation of PTC-containing mRNAs. *Biochem Soc Trans,* 38**,** 1615-20.

NOMURA, M., UDA-TOCHIO, H., MURAI, K., MORI, N. & NISHIMURA, Y. (2005) The Neural Repressor NRSF/REST Binds the PAH1 Domain of the Sin3 Corepressor by Using its Distinct Short Hydrophobic Helix. *Journal of Molecular Biology,* 354**,** 903-915.

ORBAN, T. I. & IZAURRALDE, E. (2005) Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA,* 11**,** 459-469.

PARKS, A. L., COOK, K. R., BELVIN, M., DOMPE, N. A., FAWCETT, R., HUPPERT, K., TAN, L. R., WINTER, C. G., BOGART, K. P., DEAL, J. E., DEAL-HERR, M. E., GRANT, D., MARCINKO, M., MIYAZAKI, W. Y., ROBERTSON, S., SHAW, K. J., TABIOS, M., VYSOTSKAIA, V., ZHAO, L., ANDRADE, R. S., EDGAR, K. A., HOWIE, E., KILLPACK, K., MILASH, B., NORTON, A., THAO, D., WHITTAKER, K., WINNER, M. A., FRIEDMAN, L., MARGOLIS, J., SINGER, M. A., KOPCZYNSKI, C., CURTIS, D., KAUFMAN, T. C., PLOWMAN, G. D., DUYK, G. & FRANCIS-LANG, H. L. (2004) Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome. *Nat Genet,* 36**,** 288-292.

PASSOS, D. O., DOMA, M. K., SHOEMAKER, C. J., MUHLRAD, D., GREEN, R., WEISSMAN, J., HOLLIEN, J. & PARKER, R. (2009) Analysis of Dom34 and Its Function in No-Go Decay. *Mol. Biol. Cell,* 20**,** 3025-3032.

PHELPS, C. B. & BRAND, A. H. (1998) Ectopic Gene Expression inDrosophilaUsing GAL4 System. *Methods,* 14**,** 367-379.

POULIN, G., DONG, Y., FRASER, A. G., HOPPER, N. A. & AHRINGER, J. (2005) Chromatin regulation and sumoylation in the inhibition of Ras-induced vulval development in Caenorhabditis elegans. *Embo J,* 24**,** 2613-2623.

RAMACHANDRAN, V. & CHEN, X. M. (2008) Degradation of microRNAs by a family of exoribonucleases in *Arabidopsis*. *Science,* 321**,** 1490-1492.

REINHART, B. J., SLACK, F. J., BASSON, M., PASQUINELLI, A. E., BETTINGER, J. C., ROUGVIE, A. E., HORVITZ, H. R. & RUVKUN, G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature,* 403**,** 901-906.

REN, N., ZHU, C., LEE, H. & ADLER, P. N. (2005) Gene Expression During Drosophila Wing Morphogenesis and Differentiation. *Genetics,* 171**,** 625-638.

REVERDATTO, S. V., DUTKO, J. A., CHEKANOVA, J. A., HAMILTON, D. A. & BELOSTOTSKY, D. A. (2004) mRNA deadenylation by PARN is essential for embryogenesis in higher plants. *RNA,* 10**,** 1200-1214.

ROBB, J. A. (1969) MAINTENANCE OF IMAGINAL DISCS OF DROSOPHILA MELANOGASTER IN CHEMICALLY DEFINED MEDIA. *The Journal of Cell Biology,* 41**,** 876-885.

RUBY, J. G., STARK, A., JOHNSTON, W. K., KELLIS, M., BARTEL, D. P. & LAI, E. C. (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Research,* 17**,** 1850-1864.

RYDER, E., ASHBURNER, M., BAUTISTA-LLACER, R., DRUMMOND, J., WEBSTER, J., JOHNSON, G., MORLEY, T., CHAN, Y. S., BLOWS, F., COULSON, D., REUTER, G., BAISCH, H., APELT, C., KAUK, A., RUDOLPH, T., KUBE, M., KLIMM, M., NICKEL, C., SZIDONYA, J., MAROY, P., PAL, M., RASMUSON-LESTANDER, A., EKSTROM, K., STOCKER, H., HUGENTOBLER, C., HAFEN, E., GUBB, D., PFLUGFELDER, G., DORNER, C., MECHLER, B., SCHENKEL, H., MARHOLD, J., SERRAS, F., COROMINAS, M., PUNSET, A., ROOTE, J. & RUSSELL, S. (2007) The DrosDel Deletion Collection: A Drosophila Genomewide Chromosomal Deficiency Resource. *Genetics,* 177**,** 615-629.

RYDER, E., BLOWS, F., ASHBURNER, M., BAUTISTA-LLACER, R., COULSON, D., DRUMMOND, J., WEBSTER, J., GUBB, D., GUNTON, N., JOHNSON, G., O'KANE, C. J., HUEN, D., SHARMA, P., ASZTALOS, Z., BAISCH, H., SCHULZE, J., KUBE, M., KITTLAUS, K., REUTER, G., MAROY, P., SZIDONYA, J., RASMUSON-LESTANDER, A., EKSTROM, K.,

DICKSON, B., HUGENTOBLER, C., STOCKER, H., HAFEN, E., LEPESANT, J. A., PFLUGFELDER, G., HEISENBERG, M., MECHLER, B., SERRAS, F., COROMINAS, M., SCHNEUWLY, S., PREAT, T., ROOTE, J. & RUSSELL, S. (2004) The DrosDel Collection: A Set of P-Element Insertions for Generating Custom Chromosomal Aberrations in Drosophila melanogaster. *Genetics,* 167**,** 797-813.

SAMPSON, E. M., HAQUE, Z. K., KU, M.-C., TEVOSIAN, S. G., ALBANESE, C., PESTELL, R. G., PAULSON, K. E. & YEE, A. S. (2001) Negative regulation of the Wnt-β-catenin pathway by the transcriptional repressor HBP1. *Embo J,* 20**,** 4500-4511.

SCHMID, M. & JENSEN, T. H. (2008) The exosome: a multipurpose RNA-decay machine. *Trends in Biochemical Sciences,* 33**,** 501-510.

SEMOTOK, J. L., COOPERSTOCK, R. L., PINDER, B. D., VARI, H. K., LIPSHITZ, H. D. & SMIBERT, C. A. (2005) Smaug Recruits the CCR4/POP2/NOT Deadenylase Complex to Trigger Maternal Transcript Localization in the Early Drosophila Embryo. *Current Biology,* 15**,** 284-294.

SHETH, U. & PARKER, R. (2003) Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies. *Science,* 300**,** 805-808.

SILVA, A. L. & ROMÃO, L. (2009) The mammalian nonsense-mediated mRNA decay pathway: To decay or not to decay! Which players make the decision? *FEBS Letters,* 583**,** 499-505.

SILVERSTEIN, R. A. & EKWALL, K. (2005) Sin3: a flexible regulator of global gene expression and genome stability. *Current Genetics,* 47**,** 1-17.

SONENBERG, N. & HINNEBUSCH, A. G. (2009) Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell,* 136**,** 731-745.

SONG, M.-G., LI, Y. & KILEDJIAN, M. (2010) Multiple mRNA Decapping Enzymes in Mammalian Cells. *Molecular Cell,* 40**,** 423-432.

SPRIGGS, K. A., BUSHELL, M. & WILLIS, A. E. (2010) Translational Regulation of Gene Expression during Conditions of Cell Stress. *Molecular Cell,* 40**,** 228-237.

STALDER, L. & MŸHLEMANN, O. (2009) Processing bodies are not required for mammalian nonsense-mediated mRNA decay. *RNA,* 15**,** 1265-1273.

STOECKLIN, G., MAYO, T. & ANDERSON, P. (2006) ARE-mRNA degradation requires the 5[prime]-3[prime] decay pathway. *EMBO Rep,* 7**,** 72-77.

STURTEVANT, A. H. (1913) The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology,* 14**,** 43-59.

SULLIVAN, M. (2008) Investigation of the role of the 5' - 3' exoribonuclease *pacman* in *Drosophila* wound healing. Doctor of Philosophy thesis, Clinical and Laboratory Investigation Division, Brighton and Sussex Medical School

TAYLOR, J. & ADLER, P. N. (2008) Cell rearrangement and cell division during the tissue level morphogenesis of evaginating *Drosophila* imaginal discs. *Developmental Biology,* 313**,** 739-751.

TEA, J. S., CHIHARA, T. & LUO, L. (2010) Histone Deacetylase Rpd3 Regulates Olfactory Projection Neuron Dendrite Targeting via the Transcription Factor Prospero. *J. Neurosci.,* 30**,** 9939-9946.

TEMME, C., ZAESSINGER, S., MEYER, S., SIMONELIG, M. & WAHLE, E. (2004) A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in Drosophila. *Embo J,* 23**,** 2862-2871.

THOMSON, A. M., ROGERS, J. T. & LEEDMAN, P. J. (1999) Iron-regulatory proteins, iron-responsive elements and ferritin mRNA translation. *Int. J. Biochem. Cell Biol.,* 31**,** 1139-1152.

TILL, D. D., LINZ, B., SEAGO, J. E., ELGAR, S. J., MARUJO, P. E., DE LOURDES ELIAS, M., ARRAIANO, C. M., MCCLELLAN, J. A., MCCARTHY, J. E. G. & NEWBURY, S. F. (1998) Identification and developmental expression of a 5'-3' exoribonuclease from Drosophila melanogaster. *Mechanisms of Development,* 79**,** 51-55.

TOH, Y. & NICOLSON, G. L. (2009) The role of the MTA family and their encoded proteins in human cancers: molecular functions and clinical implications. *Clin Exp Metastasis,* 26**,** 215-27.

TUCKER, M., STAPLES, R. R., VALENCIA-SANCHEZ, M. A., MUHLRAD, D. & PARKER, R. (2002) Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. *Embo J,* 21**,** 1427-1436.

TUCKER, M., VALENCIA-SANCHEZ, M. A., STAPLES, R. R., CHEN, J., DENIS, C. L. & PARKER, R. (2001) The Transcription Factor Associated Ccr4 and Caf1 Proteins Are Components of the Major Cytoplasmic mRNA Deadenylase in Saccharomyces cerevisiae. *Cell,* 104**,** 377-386.

TWEEDIE, S., ASHBURNER, M., FALLS, K., LEYLAND, P., MCQUILTON, P., MARYGOLD, S., MILLBURN, G., OSUMI-SUTHERLAND, D., SCHROEDER, A., SEAL, R., ZHANG, H. & THE FLYBASE, C. (2009) FlyBase: enhancing *Drosophila* Gene Ontology annotations. *Nucleic Acids Research,* 37**,** D555-D559.

VALENCIA-SANCHEZ, M. A., LIU, J., HANNON, G. J. & PARKER, R. (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Gene. Dev.,* 20**,** 515-524.

VAN HOOF, A., FRISCHMEYER, P. A., DIETZ, H. C. & PARKER, R. (2002) Exosome-Mediated Recognition and Degradation of mRNAs Lacking a Termination Codon. *Science,* 295**,** 2262-2264.

VAUCHERET, H., VAZQUEZ, F., CRÉTÉ, P. & BARTEL, D. P. (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & Development,* 18**,** 1187-1197.

WANG, M. & PESTOV, D. G. (2010) 5'-end surveillance by Xrn2 acts as a shared mechanism for mammalian pre-rRNA maturation and decay. *Nucleic Acids Research*.

WANG, M. J. & LIN, S. (2009) A Region within the 5'-Untranslated Region of Hypoxia-inducible Factor-1α mRNA Mediates Its Turnover in Lung Adenocarcinoma Cells. *J. Biol. Chem.,* 284**,** 36500-36510.

WELLS, S. E., HILLNER, P. E., VALE, R. D. & SACHS, A. B. (1998) Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell,* 2**,** 135-140.

WINKLER, C. J., PONCE, A. & COUREY, A. J. (2010) Groucho-Mediated Repression May Result from a Histone Deacetylase-Dependent Increase in Nucleosome Density. *PLoS ONE,* 5**,** e10166.

WOLPERT, L., BEDDINGTON, R., JESSELL, T., LAWRENCE, P., MEYEROWITZ, E. & SMITH, J. (2002) *Principles of Development, 2nd ed.,* Oxford, Oxford University Press.

XIANG, S., COOPER-MORGAN, A., JIAO, X., KILEDJIAN, M., MANLEY, J. L. & TONG, L. (2009) Structure and function of the 5' - 3' exoribonuclease Rat1 and its activating partner Rai1. *Nature,* 458**,** 784-788.

YAMASAKI, Y. & NISHIDA, Y. (2006) Mi-2 chromatin remodeling factor functions in sensory organ development through proneural gene repression in Drosophila. *Development, Growth & Differentiation,* 48**,** 411-418.

YAMASHITA, A., CHANG, T.-C., YAMASHITA, Y., ZHU, W., ZHONG, Z., CHEN, C.-Y. A. & SHYU, A.-B. (2005) Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol,* 12**,** 1054-1063.

YAO, B., RAKHADE, S., LI, Q., AHMED, S., KRAUSS, R., DRAGHICI, S. & LOEB, J. (2004) Accuracy of cDNA microarray methods to detect small gene expression changes induced by neuregulin on breast epithelial cells. *BMC Bioinformatics,* 5**,** 99.

ZABOLOTSKAYA, M. V., GRIMA, D. P., LIN, M.-D., CHOU, T.-B. & NEWBURY, S. F. (2008) The 5' - 3' exoribonuclease Pacman is required for normal male fertility and is dynamically localized in cytoplasmic particles in *Drosophila* testis cells. *Biochem. J,* 416**,** 327-335.

ZAESSINGER, S., BUSSEAU, I. & SIMONELIG, M. (2006) Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. *Development,* 133**,** 4573-4583.

ZEITLINGER, J. & BOHMANN, D. (1999) Thorax closure in *Drosophila*: involvement of Fos and the JNK pathway. *Development,* 126**,** 3947-3956.

ZEITLINGER, J., KOCKEL, L., PEVERALI, F. A., JACKSON, D. B., MLODZIK, M. & BOHMANN, D. (1997) Defective dorsal closure and loss of epidermal decapentaplegic expression in Drosophila fos mutants. *Embo Journal,* 16**,** 7393-7401.

ZHANG, Y., NG, H.-H., ERDJUMENT-BROMAGE, H., TEMPST, P., BIRD, A. & REINBERG, D. (1999) Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes & Development,* 13**,** 1924-1935.

ZHAO, X., CHEN, J., LEI, L., HU, G., XIONG, Y., XU, J., LI, Q., YANG, X., CHANG, C. C. Y., SONG, B., CHANG, T. & LI, B. (2009) The optional long 5'-untranslated region of human ACAT1 mRNAs impairs the production of ACAT1 protein by promoting its mRNA decay. *Acta Biochim. Biophys. Sin.,* 41**,** 30-41.

ZHAO, Z., FANG, L., CHEN, N., JOHNSEN, R. C., STEIN, L. & BAILLIE, D. L. (2005) Distinct Regulatory Elements Mediate Similar Expression Patterns in the Excretory Cell of *Caenorhabditis elegans*. *Journal of Biological Chemistry,* 280**,** 38787-38794.

ZHU, C. C., BORNEMANN, D. J., ZHITOMIRSKY, D., MILLER, E. L., O'CONNOR, M. B. & SIMON, J. A. (2008) *Drosophila* Histone Deacetylase-3 Controls Imaginal Disc Size through Suppression of Apoptosis. *PLoS Genet,* 4**,** e1000009.

# 9 Appendix A: PCR Primers

Primers appear in the order they first mentioned. When primers were designed in pairs and both were used the size of the pair in wild-type (unless otherwise stated) is given.

| Name | Sequence | Designed partner and size |
|---|---|---|
| jls1 | TCAAAAAGGCAGTGGCATGAG | jls2 – 562bp |
| jls2 | AGACCCCATCAGATTCGGAC | jls1 – 562bp |
| dob1F | GGAGCGGTGAGGTCGTTAAATACTC | dob1F – 1,006bp |
| dob1R | GACATTGTTCAGGGCAAGGCAG | dob1R – 1,006bp |
| SFN30 | CAGACGAGCACAAATTGG | |
| jes6 | GGAACTCCTTCAGAAGAACC | |
| DpcmF2 | CGTGACCAAGCACAAC | DpcmR2 – 142bp |
| DpcmR2 | GTAACGCTGGCAATAGT | DpcmF2 – 142bp |
| pcm5F | ACGCGTCAGTACGCCAACAA | pcm5R – 867bp (wt)/320bp (*pcm*$^{5}$) |
| pcm5R | TCTTGGGACGGTACGCCTGT | pcm5F – 867bp (wt)/320bp (*pcm*$^{5}$) |
| PEL2R | ATGGCAGGTAGGGAAGAAATC | PEL1F – 537bp |
| PcmEx1F | TCATCTAAATTGCCACTGATCCAGG | PcmEx1R – 183bp |
| PcmEx1R | TGCATGGCAACATATAACGTGACC | PcmEx1F – 183bp |
| PcmEx2F | ATGTTGCCATGCACCGTACGG | PcmEx2R – 1,190bp |
| PcmEx2R | ATGGGAAGGCGGCATTGAGC | PcmEx2F – 1,190bp |
| PcmEx3F | ATTTCCGTCCGCTCAATGCC | PcmEx3R – 395bp |
| PcmEx3R | AGCCTGATGCTAGGAAAGTTATCGG | PcmEx3F – 395bp |
| PcmEx4F | CCAATGGTGGTGCTAGATCATTTGG | PcmEx4R – 623bp |
| PcmEx4R | CATGGCGAGCGTTATGCATGTG | PcmEx4F – 623bp |
| PcmEx5F | CACATGCATAACGCTCGCCAT | PcmEx5R – 384bp |
| PcmEx5R | GGAAGATGGAGCATTAGCAGTGTGA | PcmEx5F – 384bp |
| PcmEx6F | TGCTCCATCTTCCCTTTCCTCC | PcmEx6R – 402bp |
| PcmEx6R | GAAATGATGATGGGGTGTCAAGGA | PcmEx6F – 402bp |
| PcmEx7F | CCATGTCCTTGACACCCCATCA | PcmEx7R – 988bp |
| PcmEx7R | TTGCTATGTATGTGCCTCCTCCGA | PcmEx7F – 988bp |
| PcmEx8F | GCATAGCCTACAGAATCGGAGGAGG | PcmEx8R – 691bp |
| PcmEx8R | TATGTGAACACATGGTTGGGTGGC | PcmEx8F – 691bp |

| Name | Sequence | Designed partner and size |
|---|---|---|
| PcmEx9F | AGCCCAGATGTGCCACCCAA | PcmEx9R – 806bp |
| PcmEx9R | CAGATTCGGACTGCGACGCA | PcmEx9F – 806bp |
| PcmEx10F | CATTGGCTACAAGTCAAAGCCTTGA | PcmEx10R – 205bp |
| PcmEx10R | GCAACAGAAGCGAGTAAATGGAATG | PcmEx10F – 205bp |
| PcmEx11F | CATTCCATTTACTCGCTTCTGTTGC | PcmEx11R – 498bp |
| PcmEx11R | CAAATCACGAACGAGATTGCTCAAT | PcmEx11F – 498bp |
| PEL1F | GAGCAATCTCGTTCGTGATTT | PEL2R – 537bp |
| PEL3R | GGCAGATTTTGACGAAACTC | PEL1F – 702bp |
| Nat15F | CCCACTAGCTCCCTTGCTCGTTCT | |
| Nat16F | CCATTTCTCCACCCACTGACTGG | |
| SFN22 | TTCGACTCGGAATGCAAGGCAC | |
| ED7452F | TGAGCCCATACAGTGAATTGCAAA | PRY4 – ~700bp (*Df(1)ED7452* lines only). |
| ED7452R | GCCAAAGGTAAACAAAACAGAACGG | PRY4 – ~700bp (*Df(1)ED7452* lines only). |
| PRY4 | CAATCATATCGCTGTCTCACTCA | ED7452F, ED7452R – ~700bp (*Df(1)ED7452* lines only). |
| W7500D | GTCCGCCTTCAGTTGCACTT | W11678U – ~1,600bp (*Df(1)ED7452* lines only). |
| W11678U | TCATCGCAGATCAGAAGCGG | W7500D – ~1,600bp (*Df(1)ED7452* lines only). |

# 10 Appendix B: Bioinformatic tools

Online databases and web-based tools.

| Tool | Hyperlink | Function |
|---|---|---|
| FlyBase | http://flybase.org/ | *Drosophila* gene annotation database. |
| Ensembl | http://www.ensembl.org/ | Eukaryotic genome database. |
| miRBase | http://www.mirbase.org/ | miRNA annotation database. |
| MicroCosm | http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/ | miRNA target prediction. |
| TargetScanFly | http://www.targetscan.org/fly_12/ | miRNA target prediction. |
| PicTar | http://pictar.mdc-berlin.de/ | miRNA target prediction. |
| NCBI Primer-BLAST | http://www.ncbi.nlm.nih.gov/tools/primer-blast/ | PCR Primer design. |

Software packages.

| Tool | Developer | Function |
|---|---|---|
| VectorNTI Advance v.11 | Invitrogen | PCR primer design and sequence manipulation. |
| AlignX | Invitrogen | Sequence alignments. |
| MxPro | Stratagene | qPCR experiment setup and analysis. |
| GenePix Pro 7 | Molecular Devices | miRNA array experiment analysis. |
| R 2.10.1 | The R Foundation for Statistical Computing | Affymetrix mRNA array analysis. |
| ImageJ | Wayne Rasband | Imaginal disc measurements and Western blot relative quantification. |