Understanding the importance of RNA export factor, Nxt1, during the metamorphosis and in ovaries in *Drosophila melanogaster*

Kevin van der Graaf

March 2018

This thesis is submitted
for the Degree of Doctor of Philosophy

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (candidate) Date ……..……..…

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed ……..……..…… (candidate) Date …….

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ……..……..…… (candidate) Date …….

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ……..….. (candidate) Date ……

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ……..……..…… (candidate) Date ……..……

Acknowledgements

I would like to thank my supervisor Helen White-Cooper for her guidance, patience and enthusiasm throughout my PhD period. It has been a great experience in the past four years and I am grateful for all the help I have received during my projects. I would also like to thank our collaborator Julius Brennecke for providing a great insight in the small RNA sequencing process and the bioinformatics. I would also like to thank the other PIs in the *Drosophila* facility, Sonia Lopez de Quinto, Mike Taylor, Wynand van der Goes van Naters and Joaquin de Navascués Melero for all their valuable feedback during my PhD.

A big thanks to all my colleagues in the fly labs for the past four years for all their support and their friendship. It made my PhD much more enjoyable.

Thanks to my Mom, Dad and Brother for their support and visits throughout the years.

My PhD was funded by the BBSRC.

Summary

The RNA export pathway is vital for export-competent mRNA to be exported from the nucleus to the cytoplasm. Without this pathway, RNA is unable to be translated into proteins for their function. A hetero-dimer comprised of Nxt1-Nxf1 is required to bind to the export-competent mRNA for a functional interaction between Nxf1 and the nuclear pore complex. Without Nxt1, Nxf1 interacts less effectively with the complex.

In this thesis, the importance of Nxt1 is further examined during metamorphosis and in ovaries. I have shown that a reduction of functional Nxt1 during metamorphosis affects the muscle integrity in third instar larvae. This, in turn, affects the migration of the air bubble, which moves from the posterior to the anterior end, leading to dead pupae with no head eversion. Increasing the expression of a single gene, *abba*, rescues the muscle phenotype, but not the pupa viability. Further examination of differentially expressed genes revealed that genes with many and long introns were the most affected.

In the ovaries, I have shown that Nxt1 is important in the piRNA pathway. Microarray and qRT-PCR have shown that transposons are de-repressed with less Nxt1, leading to sterility and death. Two rounds of small RNA-sequencing were performed, however, discrepancies between the two rounds made it difficult to analyse the data. Either Nxt1 influences the de-repression of transposons or Nxt1 has a mild effect on de-repression of transposons and is more involved in the silencing process. Together with our collaborators we also got a first insight into which potential Nxf protein is interacting with Nxt1 in the piRNA pathway.

Table of Contents

Abbreviations

AEL	After egg laying
CBC	Cap binding complex
CAG	Cyo Actin GFP
DEOM	Dorsal external oblique muscle
DFM	Dorsal flight muscle
DIOM	Dorsal internal oblique muscle
DLM	Dorsal longitudinal muscle
Dpc	Days post coitum
Dpp	Days post partum
DVM	Dorso-Ventral muscle
Eig	Ecdysone induced genes
EJC	Exon-junction complex
FC	Founder cell
FCM	Fusion competent myoblast
FPKM	Fragments per kilobase of exon per million mapped fragments
Fps	Frame per second
GSC	Germline stem cell
GTSF1	Gametocyte-specific factor 1
hnRNP	Heterogeneous nuclear ribonucleoprotein
HuR	Human antigen R
IFM	Indirect flight muscle
IP	Immunoprecipitation
iTF	Identity transcription factor
LLR	Leucine-rich repeat
MD	Muscular dystrophies
Mef2	Myocyte enhancer factor-2
Mib2	Mind bomb 2
mRNP	Messenger ribonucleoprotein
NES	Nuclear export signal

NPC	Nuclear pore complex
NTF2L	Nuclear transport factor 2-like domain
Nxt1	NTF2-related export protein 1
ORF	Open reading frame
OSC	Ovarian somatic cell
piRNA	Piwi-interacting RNA
Piwi	P-element induced wimpy testis
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
RdRP	RNA-dependent RNA polymerase
RIP	RNA immunoprecipitation
RRM	RNA recognition motif
SEM	Scanning electron microscope
SR	Sarcoplasmic reticulum
SSCR	Signal sequence coding region
TE	Transposable element
tMAC	Testis-specific meiotic arrest complex
TREX	Transcription-export complex
TSS	Transcriptional start site
UBA	Ubiquitin-associated
UPS	Ubiquitin proteasome
UTR	Untranslated region
WT	Wild type

List of Figures

List of Tables

Chapter 1.

Introduction

The process of formation of a complex multicellular animal from a single cell, a fertilised egg, requires the coordination of many different processes. During development, this cell divides repeatedly and produces many different cells in a final pattern, which is very complex and precise. Ultimately, the genome determines this pattern. Cells are different due to the different set of genes it expresses. This selective gene expression controls four essential processes: 1) Cell proliferation, 2) cell specialization, 3) cell interactions and 4) cell movement. The cells in an organism – neuron, skin cells, gut cells, muscle cells, etc. – maintain their specialization because they retain a record of signal their ancestors received in early embryonic development.

Again, all these events are underpinned by regulation of gene expression, at both the transcriptional and translational level. There has been extensive research into this, but several outstanding questions remain. One is how defects in function of ubiquitously expressed proteins that are essential for fundamental functioning of cells can lead to specific defects in individual cellular or organ systems rather than non-specific lack of cell viability. This project addresses one such situation using the genetically tractable system *Drosophila melanogaster*.

1.1 The *Drosophila* life cycle

Drosophila melanogaster is a holometabolous insect that has a larval and a pupal stage prior to the adult stage. Fertilization occurs internally, and sperm is stored within the female's body in a seminal receptacle and the paired spermathecae. Females lay approximately 50-70 eggs per day during their peak, which is between four and seven days after mating.

The eggs are about one-half mm in length, white, oval and slightly flattened. The ovum has an inner, thin vitelline envelope and an outer, stronger coat called a chorion. Two small extensions of the chorion, extended from the dorsal surface, are the respiratory filaments that serve for gas exchange. Eggs are laid in rotten fruit or in medium in culture jars.

Normally, eggs hatch after 22-24 hours at 25^{o}C. The larva emerges and is called the first instar lava. The larva feeds on the substrate that the eggs were laid in to molt into a larger second

instar larva about 25 hours later. The larva continues to eat for another 24 hours to develop into a third instar larva. The larva feeds for another few hours before it crawls away from the substrate and moves to a suitable place to undergo pupariation. After 30 hours in the third instar phase the larva molts into a prepupa. The prepupa is stationary and has a yellowish-white colour that becomes progressively darker. During this stage, an air bubble is formed at the far posterior and moves to the far anterior by muscular contractions. The air bubble increases in size and forces all the abdomen content against the body wall (Shatoury 1959). Within four hours, prepupae have become separated from the puparium, however, they remain connected with the mouth, anus and posterior spiracles (Shatoury 1959).When the air bubble reaches the anterior it allows space for head eversion. Head eversion is the prepupa to pupa stage transition. During the pupal stage, which lasts for 3-4 days, the larva is metamorphosing into an adult fly.

1.1.1 Larval growth

After 22-24 hours eggs are hatching and the larval emerges. During this phase of the *Drosophila* life cycle, larvae will eat food to grow and molt into second and finally third instar. The larva uses its mouth hooks to extend and eat the food. Besides the differences in size between first and second instar, a difference in mouth hooks can be observed as well. In first instar mouth hooks are two pairs of tiny black spots, whereas second instar have larger mouth hooks and a much clearer structure. The most obvious morphological differences between second and third instar lies in the spiracles. Second instar has club-like anterior spiracles, while third instar has branched spiracles. Another difference between the two instars is at the end of the posterior spiracle. Third instar larvae have a dark orange ring at their tip, which is lacking or weakly present in second instar. These morphological features help for staging larvae as some mutants or growing conditions stunt growth (Conlon and Raff 1999).

1.1.2 Metamorphosis

The pupa development can be subdivided into 15 different stages (P1-P15) based on their morphological changes (Figure 1.1) (Bainbridge and Bownes 1981). There are a total of 46 events visible during the pupa phase, only the most significant will be described here. P1

marks the start of the pupa phase, where pupa has a white colour and slowly turns into a brown colour within the first hour. During stage P2, an air bubble becomes visible within the abdomen at the posterior of the prepupa and moves slowly to the anterior. In stage 4 the puparium starts to separate from the underlying epidermis beginning at the anterior end. The final larval/pupal cuticle separation is completed after the air bubble allows to evert the head. The resulting stage is termed the *Phanerocephalic* pupa. By stage P4 (i) the bubble within the abdomen disappears and is moved further to the anterior in stage P4 (ii) to cause a head eversion marking the transition of the prepupa-pupa stage. P5 (i) pupa achieves full extension along the abdomen and white malpighian tubules become visible dorsally in the abdomen. These malpighian tubules become prominent and green in P5 (ii). At P6 a 'yellow body' appears between the anterior end of the two malpighian tubule segments and moves back between the malpighian tubules at P7. At the same time, the eyes become yellow at its perimeter. Over the next few stages the eye colour changes to yellow-eyed (P8), amber (P9) and red (P10). At P11 the head and thoracic bristles become visible and the tips of the folded wings become grey. During P12 the wings become grey and darken to black. The sex comb darkens, and abdominal bristles become visible. The tarsal bristles darken, and claws become black (P13), whereas the meconium appears dorsally at the posterior tip of the abdomen (P14). Finally, in stage P15, legs start to twitch and the pouch on the head expands by hydrostatic pressure that opens the operculum.

In normal development, the larva has three phases of contraction. The first is observed at the puparium formation that associates with the biggest reduction in length of the larva. The second is observed at the anterior where an appearance of space becomes visible which is later occupied by the head, and a surrounding space into which the everted appendages could grow. And finally, the third contraction is the escape of the air bubble that is important for the eversion of the head (Shatoury 1959). The larval muscles carry out all these contractions.

1.1.2.1 Air bubble

Three hour-old prepupa contract away from the anterior end partially withdrawing the anterior tracheal lining (Shatoury 1959). At this stage an air bubble forms in the abdominal cavity (Figure 1.1, arrow). The bubble gradually increases in size and all the abdomen content is forced against the body wall. One hour later, prepupa becomes separated from the

puparium due to the secretion of the prepupal cuticle. The prepupa however remains connected to the cuticle with the mouth, anus and posterior spiracles. The air bubble migrates to the anterior by muscular contractions to evert the head. It is then preceded by contraction of the tip of the abdomen and the disappearance of the air bubble via the spiracles.

This figure (Figure 1.1) was removed because of copyright restrictions.

Figure 1.1: **Overview of different stages during metamorphosis.** The metamorphosis can be staged into 15 different stages. P1 marks the start of the metamorphosis, where the pupa has a white colour that slowly becomes darker. During P2 the air bubble becomes visible within the abdomen and moves anteriorly during P3 (arrow) and disappears at P4. At P5, the pupa gets it adult shape. From P8 to P12 several events occur with eye colour development being quite prominent. Finally, the fly ecloses after 3-4 days. Picture taken and adapted from http://www.devbio.net/images/metam.gif. Copyright permission pending.

1.1.2.2 Ecdysone

The steroid hormone 20-hydroxyecdysone (hereafter referred to as ecdysone) coordinates tissue-specific morphogenetic changes in *Drosophila melanogaster*. There are at least six pulses of ecdysone throughout the life cycle from embryonic through the end of the pupa stage (Ashburner et al. 1974). During the first and second instar, a weak pulse of ecdysone is observed to trigger the molting process of the cuticle. In the third instar, a strong pulse of ecdysone is observed to trigger the onset of metamorphosis of the larva. Twelve hours later,

another pulse is observed to trigger head eversion, marking the prepupa-pupa transition. Then, another strong and longer pulse of ecdysone is observed later during the pupa stage (Richards 1981; Handler 1982). Ecdysone affects all tissues during metamorphosis to reorganize the animal, where most tissue break down and are replaced by adult structures.

In the salivary gland, polytene chromosomes have a few active "intermoult puffs". Puffs are caused by high levels of transcription giving the chromosomes a very open, decondensed conformation. After a pulse of ecdysone, some of these "intermoult puffs" regress and a few "early" ecdysone-inducible puffs appear triggering the puparium formation. Several hours later, more than 100 "late" puffs have appeared. Ashburner and colleagues have been able to reproduce the late larval puffing response to ecdysone by performing a series ecdysone dose-response, hormone removal and addition, and drug inhibition (Ashburner et al. 1974). This resulted in a model where ecdysone, bound to its receptor protein, had three main regulatory effects: (1) it regressed the intermoult puffs, (2) it induced the early puffs, and (3) it repressed the late puffs. The early puffs encode transcripts that are translated into early proteins that have a dual-regulatory function: (1) it represses the translation of their own transcripts and (2) it induces the late puffs (Figure 1.2).

The three ecdysone-inducible early puffs best defined at a molecular level are *Broad-Complex* (*BR-C*; 2B5 puff)(DiBello et al. 1991), *E74* (74EF puff) and *E75* (75B puff)(Ashburner 1990; Thummel 1990). All three genes are repressed by ecdysone-induced proteins as predicted by the Ashburner model. Mutations in *BR-C, E74* or *E75* result in lethality during late larval, prepupal or pupal development indicating the importance of these early puff genes for metamorphosis (Thummel 1990).

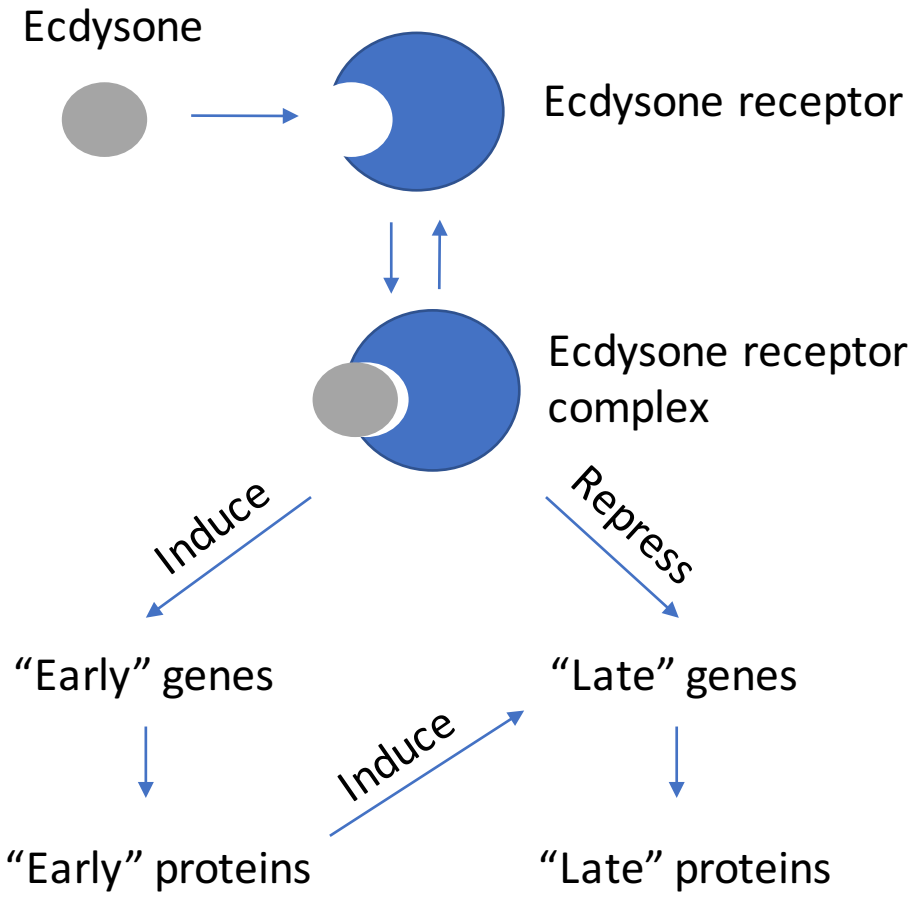


Figure 1.2: **Ecdysone “Ashburner model”.** Ecdysone when bound to its receptor has two functions: 1) Induce ‘Early’ genes and 2) repress ‘Late’ genes. ‘Early’ genes are translated into ‘Early’ proteins that have a dual function: 1) Repress ‘Early’ genes and 2) induce ‘Late’ genes for translation.

1.2 Muscle development

1.2.1 Embryonic muscle development

At embryo gastrulation (3 hours after egg laying), mid-ventral cells invaginate and spread in a layer under the ectoderm to form the mesoderm, which give rise to several different cell types, including various muscles (Bate 1993). During gastrulation these invaginating cells are fate restricted to be mesodermal derivatives, but are not yet committed to any cell type (Beer, Technau, and Campos-Ortega 1987). Over the next few hours, these cells proliferate and diversify, followed by fate commitment. Somatic, visceral and heart are the three different embryonic muscle types. At 5 hours after egg laying (AEL), two cell layers are created by the separation of the mesoderm: the somatic muscles from the external layer and the visceral from the internal layer (Bate 1993). The heart derives from the most dorsal external cells. For the somatic muscles, two or three fused myoblasts appear and abut the central nervous system or epidermis 8 hours AEL. Founder cells recruit and fuses with neighbouring myoblasts to form the precursor of a myotube (Bate 1990). Together they pre-figure the final somatic muscle pattern.

Each muscle is seeded by a single founder cell (FC). FCs undergo multiple rounds of fusion with fusion competent myoblasts (FCMs) (Rushton et al. 1995). The muscle identity reflects the expression by each FC of a specific combination of identity transcription factors (iTFs) (Baylies, Bate, and Ruiz Gomez 1998; Frasch 1999). FCs come from asymmetric division of progenitor cells, selected by Notch-mediated lateral inhibition from promuscular clusters specified at fixed positions within the somatic mesoderm (Carmena et al. 1998; Corbin et al. 1991). Abdominal segments A2-A7 and Thoracic segments T2-T3 have a set of 30 muscles, whereas segment A1 has a modified pattern with the absence of muscle 25 and 17, and presence of 31 (associated with mouthparts) (Bate 1990). There are other specialized patterns observed in A8 and in T1. The diversification of the muscle pattern is regulated by homeotic gene expression in the mesoderm (Greig and Akam 1993; Michelson 1994). Hox proteins have a role in establishing the pattern of *Drosophila* muscles by implementing the expression of iTFs at the progenitor stage (Enriquez et al. 2010). For example, the number of myoblasts

allocated to LT1-4 muscles, which express the iTF Apterous, is under Hox control (Capovilla, Kambris, and Botas 2001).

1.2.2 Larval muscle development

During the larval phase, the completed muscle pattern in the embryo undergo a ~50-fold increase in fiber size in only 5 days, via sarcomere assembly and the addition of novel myofibrils (Demontis and Perrimon 2009). The fiber growth is necessary, but not sufficient on nutrients, and sensing of these via the Insulin/Akt/TOR pathway is required (Demontis and Perrimon 2009). Maximal muscle growth requires additional signals such as mitochondriogenesis, ribosome biogenesis, sarcomere assembly and other anabolic responses (Teleman et al. 2008; Bai, Hartwig, and Perrimon 2007). Overexpression of the Insulin receptor inhibits FOXO, which increases *Myc* expression to promote nucleolar biogenesis and muscle growth. Muscle growth can be inhibited by FOXO that repress *Myc* expression and its transcriptional activity (Demontis and Perrimon 2009).

In mammals, muscle atrophy is the result of loss in muscle mass and a decrease in protein synthesis. Two systems are activated by protein degradation: Ubiquitin proteasome (UPS) and autophagy/lysosome systems. Both of the systems are activated during muscle atrophy and key components are induced by transcription factors of the FOXO family (Sandri et al. 2004; Zhao et al. 2007) and NF-kB (Cai et al. 2004; Hunter et al. 2002). Loss of Insulin-like Growth Factor 1 (IGF-1), through both endocrine and autocrine mechanisms, contributes to muscle atrophy by signaling in the IGF-1/Akt pathway. IGF-1 is released by contractile activity, which functions as in important autocrine growth factor. Reduction of Akt signaling decreases the protein synthesis in muscle and other cells via activation of the FOXO transcription factors. Several other transcription factors, such as Smad 2 and 3, NF-kB, JunB and Runx1 have been shown to be involved in the muscle mass. Smad transcription factors mediate the catabolic effects of myostatin that inhibits normal muscle growth (Lee 2004). Myostatin binds to the Activin RIIB receptors and activates transcription by Smad 2 and 3, and inhibition causes muscle hypertrophy (Sartori et al. 2009). NF-kB is an important factor in muscle atrophy, but its exact mechanism is currently unclear (Cai et al. 2004). JunB is a rapid-response gene that triggers cell proliferation, determines the growth rate of muscle fibers in adult mammals, and

overproduction can both promote hypertrophy and inhibit muscle atrophy (Raffaello et al. 2010). The transcription factor, Runx1, limits denervation-induced muscle wasting by regulating expression of muscle structural proteins and ion channels (Wang et al. 2005).

1.2.3 Pupal muscle development

The thoracic somatic muscles are grouped into the Direct Flight Muscles (DFMs) and Indirect Flight Muscles (IFMs). The latter have two sub-groups: Dorsal Longitudinal Muscles (DLMs) and Dorso-Ventral Muscles (DVMs).

By 8 hours into the pupa phase most muscles in the head and thoracic segments undergo histolysis while the abdominal segments are unaffected (Fernandes, Bate, and Vijayraghavan 1991). The larval oblique muscles escape histolysis and are used as a template to form the DLMs, which are part of the IFMs in the adult flies (Farrell, Fernandes, and Keshishian 1996; Roy and VijayRaghavan 1998; Dutta et al. 2004). Later in pupa development, most larval abdominal muscles undergo histolysis. Dorsal External Oblique Muscles (DEOMs) start degenerating late in the prepupal stage (Wasser, Bte Osman, and Chia 2007). Other muscles loose muscle mass (atrophy), increase in muscle mass again (hypertrophy) and form the Dorsal Internal Oblique Muscles (DIOMs). DIOMs are important for eclosion and degenerate after the adult has emerged (Kimura and Truman 1990). Besides DLMs, all other thoracic muscles are thought to be developed de novo (Bate 1993). Research has shown that adult abdominal muscle cells are corresponding to founder myoblasts that are required for myotube formation. (Dutta et al. 2004; Dutta and Vijayraghavan 2006).

1.2.4 Muscle structure

Muscles are composed of tandem arrays of sarcomeres containing thick and thin filaments (Figure 1.3). When a muscle twitches, the filaments slide past each other in response to calcium release from the sarcoplasmic reticulum (SR) resulting in force generation (Taylor 2006).

In humans, muscle fibers can be categorized into three types. Type 1 slow twitch oxidative, Type 2A fast twitch oxidative and Type 2B fast twitch glycolytic. The direct and indirect flight

muscles in flies promote wing motion indirectly by compressing the thorax, are primarily oxidative. Body wall muscles in the larva and leg muscles of adult flies rely mainly on glycolysis (Taylor 2006). The muscle energy metabolism resembles the difference between type I and type IIb fiber mammalian muscles. Type I are non-fatiguing, burn fatty acids and glucose oxidatively, and are dark in colour due to their enrichment in mitochondria, myoglobin and blood supply. Type IIb are easily fatigued and have low mitochondrial content and capillarity density (Taylor 2006).

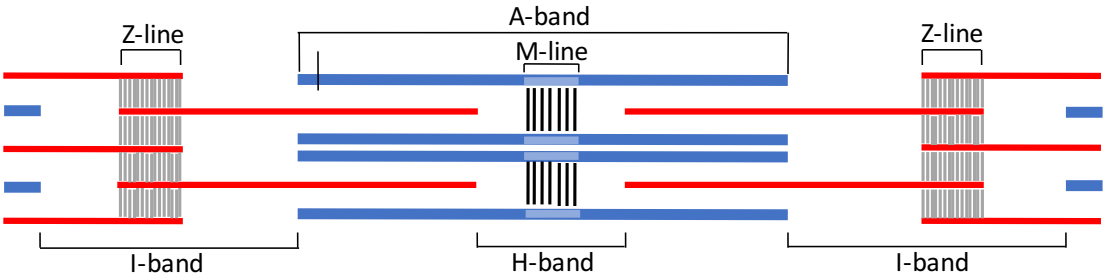


Figure 1.3: **The arrangement of thick and thin filaments in vertebrate striated muscle sarcomere.** The two Z lines, composed of mostly a-actinin, define the edges of the sarcomere. The thin filaments attached to each Z line extend to the center of the sarcomere. The region only containing thin filaments is the I band. The thick filaments are located in the center of the sarcomere. The region only containing thick filaments is the H band with the region defined by their extent is the A band. At the center of the sarcomere is the region called the M line.

1.2.5 Muscle dysfunction and diseases

The myocyte enhancer factor-2 (Mef2) transcription factors are a key factor in skeletal/somatic muscle differentiation (Taylor 1995; Black and Olson 1998). Mef2 is able to regulate muscle gene transcription through a binding site found in the promoter of the majority of muscle specific genes (Black and Olson 1998). In *Drosophila*, a single *Mef2* gene exists opposed to four in mice. In Mef2 mutants, somatic muscle precursor cells are produced in the correct pattern, but cells fail to fuse and differentiate further (Taylor 1995; Bour et al. 1995; Lilly et al. 1995; Ranganayakulu et al. 1995). *Mef2* was also the first gene shown to control differentiation in multiple muscle cell types (Taylor 1995). Mef2 proteins have been involved in several disorders. A transcription factor associated with Mef2D is Gtf3. Gtf3 shares

a high homology with TFII-1, which suggests a role in the integration of muscle-specific transcription factor activity (O'Mahoney et al. 1998). Eleven out of 24 known splice variants are detectable only in skeletal muscle, and human Gtf3 is localized to a region that is deleted in Williams-Beuren syndrome, a disease associated with muscle weakness and atrophy. Of the genes deleted in this syndrome, Gtf3 is the only one associated with skeletal muscle function, suggesting a role of Gtf3 in human skeletal muscle (Tassabehji et al. 1999).

In humans, muscular dystrophies (MD) are a group of inherited genetic conditions that cause muscles to weaken, leading to an increasing level of disability. MD is typically caused by mutations in the genes involved in the muscles. There are several types of MD, with Duchenne MD one of the most common and severe one. Other examples are myotonic dystrophy, Becker MD and limb-girdle MD. Most of these MDs can start developing in early childhood and so far, no cure has been found for any MDs. However, treatments exist that can help with physical disabilities and problems that may develop. It is important to understand muscle genes and its targets to understand their involvement in these and many other muscles diseases.

1.3 RNA export pathway

1.3.1 Introduction to mRNA export

Eukaryotic gene expression is an intricate pathway that is controlled by multiple mechanisms. Regulation of gene expression is important for the normal functioning of organelles, cells, tissues and the whole organism. Gene expression starts with various regulatory proteins binding to DNA that allows the ability of RNA polymerase to bind to the promoter site and carry out transcription. The transcribed RNA can also be controlled by altering the splicing of exons and by controlling access to or efficiency of transport channels. The passage of the mRNA transcripts to the cytoplasm requires the transcript to be recognized by receptors in the interior of the pores. In the cytoplasm, enzymes are able to degrade the mRNA, whereas different proteins can take part in the translation process. After translation, post-translational modification can occur that can alter the activity of a protein (Losos and Raven 2008).

Part of this gene expression controls is the involvement of transporting mRNA transcripts from the nucleus to the cytoplasm where they are translated into proteins. There are distinct stages of the nuclear export of mRNA transcript: first, a transcript is transcribed into a precursor mRNA (pre-mRNA), which is then packed into messenger ribonucleoprotein (mRNP) complexes; second, the mRNPs are transported to and translocate through the nuclear pore complexes (NPCs); and third, the mRNPs are released in the cytoplasm for translation or assembly into other mRNPs (Figure 1.4).

The export-competent mRNPs starts with the transcription of protein coding genes. During the transcriptional elongation, different proteins will bind, which are part of the family of heterogeneous nuclear ribonucleoproteins (hnRNPs). hnRNPs are RNA-binding proteins essential for the mRNA process, including packaging, export and translation (Dreyfuss, Kim, and Kataoka 2002). Some of these hnRNPs are retained during the whole mRNA export process while some others are removed prior to export or replaced by other hnRNPs. The main steps that occur during RNA processing to generate an mRNA are 5' capping, splicing, 3'-end cleavage and polyadenylation. 5' capping is the first step where a 7-methylguanosine cap is added to the 5' end to protect the pre-mRNA from degradation (Shatkin and Manley 2000). 5' capping has also an important effect on maturation, promotion on translation and increase in stability by preventing degradation by exonucleases (Shatkin 1976; Gao et al. 2000). With the 5' cap, the cap binding complex (CBC) binds, which is then recognized by the NPC (Gonatopoulos-Pournatzis and Cowling V 2014). Splicing is an event where introns in a transcript are spliced out by the spliceosome. Proteins defined as exon-junction complex (EJC) are then deposited at the splicing site. 3'-end cleavage is the process where the 3'-end is cleaved downstream, by the recognition of a polyadenylation site, located in the 3'-untranslated region (UTR). Then, finally, a poly(A)$^+$ tail is added by a poly(A)$^+$ polymerase to mark the completion of the mRNA process.

1.3.1.1 Recruitment of RNA export factors

During transcription, the THO complex, in *S. cerevisiae*, consisting of *Tho2*, *Hpr1*, *Mft1* and *Thp2*, binds transcripts as they emerge from the RNA polymerase to promote transcription

elongation and 3' end processing (Jimeno et al. 2002). Sub2 (UAP56 in *Drosophila*) and Yra1 (Ref1 in *Drosophila*) interacts with the THO complex to form a super complex called the transcription-export complex (TREX) (Strasser et al. 2002). In metazoans, the THO complex lacks *Mft1* and *Thp2*, but contains Thoc5, Thoc6 and Thoc7 (Masuda et al. 2005; Rehwinkel et al. 2004). In transcripts where splicing occurs, EJC, consisting of SRM160, DEK, RNPS1, REF1 and Y14, in *S. cerevisiae* (Le Hir et al. 2000), is deposited near the 5' of the splice site (Le Hir et al. 2001). After EJC is bound to the transcript, a dimer comprised of Nxf1 (metazoans; also known as TAP; MEX67 in *S. cerevisiae*) and Nxt1 (metazoans; also known as P15; MTR2 in *S. cerevisiae*), is recruited and binds to the mRNA, via an interaction with Ref1 and Nxf1, to make it competent for nuclear export (Strasser and Hurt 2000)(Figure 1.4). In intron-less genes, however, no EJC is deposited on the transcript. Interestingly, those transcripts can still interact with Ref1 and therefore also recruit Nxt1-Nxf1. Thus, the EJC increases the affinity of depositing Nxt1-Nxf1 to the transcript (Le Hir et al. 2001).

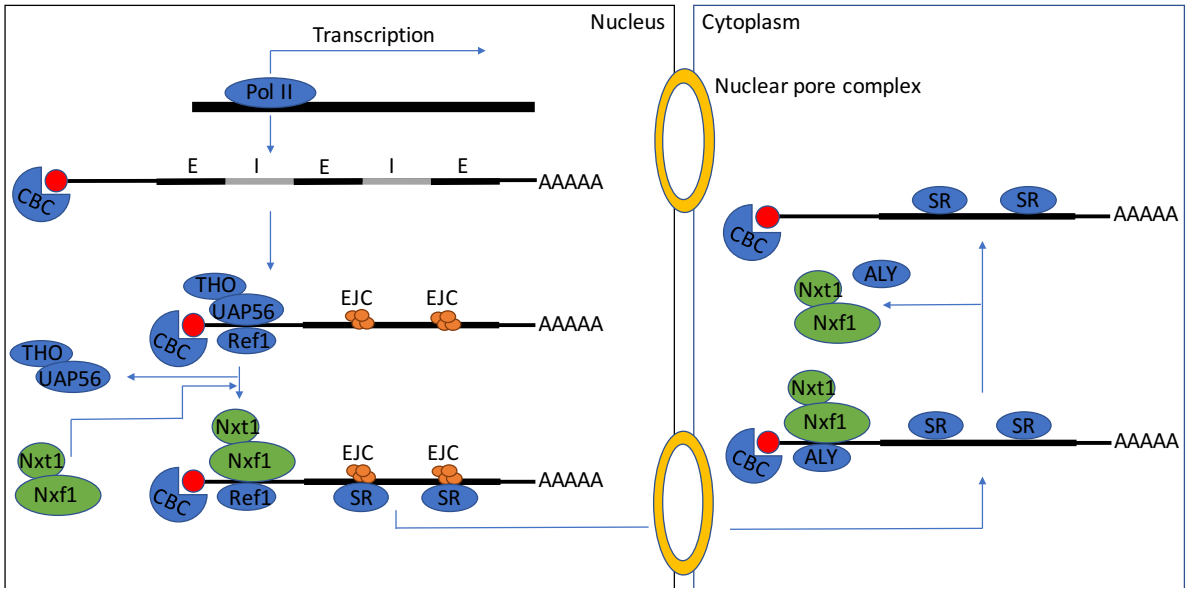


Figure 1.4: **RNA export pathway in mammals.** DNA is transcribed by polymerase II into an RNA transcript with introns and exons. During transcription, CBC caps the 5' end and the TREX complex binds to the RNA. The EJC is recruited for splicing of introns while THO and UAP56 are displaced from the RNA. RNA-binding proteins Ref1 binds to the RNA transcripts and recruits the heterodimer Nxf1(TAP)/Nxt1(p15). The Nxf1/Nxt1 binds to the RNA and renders the mRNA in a mature state. The complex goes through the nuclear pore complex by the interaction with Nxf1 and the nuclear pore proteins. In the cytoplasm the heterodimer is released and re-imported to the nucleus. SR proteins act as adaptors for mRNA export and as regulators for translation in the cytoplasm. E = exon, I = Intron.

1.3.1.2 Nxt1-Nxf1 dimer in the mRNA export pathway

In *Drosophila*, NTF2-related export protein 1 (Nxt1) is a small protein related to the Ran-GDP-binding nuclear transport factor NTF2 and is known to bind to Nxf1 (encoded in Drosophila by the gene *small bristles*, *sbr*) to form a heterodimer to facilitate the export of poly(A)$^+$ RNA by interacting to nucleoporins (Katahira et al. 1999; Fribourg et al. 2001; Wiegand et al. 2002; Wilkie et al. 2001). Nxf1 is a protein comprised of four domains: an N-terminal RNA recognition motif (RRM), a leucine-rich repeat (LRR), a nuclear transport factor 2-like domain (NTF2L) and an ubiquitin-associated (UBA) domain. Nxt1 and Nxf1 form a heterodimer via their NTF2L domains (Herold, Klymenko, and Izaurralde 2001; Katahira et al. 2002; Fribourg and Conti 2003; Korey et al. 2001). NTF2L and UBA domains have hydrophobic surface depressions to which phenylalanine-glycine (FG) motifs present in nuclear pore complex (NPC) proteins (or nucleoporins) can bind and this interaction is crucial for mRNA transportation (Valkov et al. 2012). In the absence of Nxt1, Nxf1 protein is unable to interact effectively with these nucleoporins (Valkov et al. 2012). Formation of the Nxt1-Nxf1 heterodimer enhances nucleoporin binding both *in vitro* and *in vivo*, and induces formation of Nxt1-Nxf1-nucleoporin ternary complex (Wiegand et al. 2002). The core mRNA export pathway is extremely well conserved as demonstrated by the finding that human TAP (Nxf1) can rescue mRNA export in *S. cerevisiae* MEX67/MTR2P double-mutant deficient cells when co-expressed with human P15 (Nxt1)(Katahira et al. 1999).

1.3.1.3 Non-canonical roles of Nxt1 and other RNA export factors

A role for Nxt1 beyond the core the mRNA export pathway has been discovered. A link has been made between Nxt1 and the testis-specific meiotic arrest complex (tMAC) that regulates the transcriptional program in testes in the *Drosophila melanogaster* (Caporilli et al. 2013). The hypomorph $Nxt1^{z2\text{-}0488}$ homozygote males showed a meiotic arrest phenotype where spermatocytes are arrested and unable to differentiate into spermatids and sperm. The null allele $Nxt1^{DG05102}$ is embryonic lethal in homozygous condition and trans-heterozygotes $Nxt1^{z2\text{-}0488l}$/ $Nxt1^{DG05102}$ were semi-lethal and male and female sterile (Caporilli et al. 2013). The Nxt1 trans-heterozygotes had a meiotic arrest phenotype. Most previously described meiotic arrest genes can be classified into two classes according to their phenotypes, via RNA

in-situ hybridization, which are *aly*-class (*aly, comr, tomb, topi* and *achi+vis*) and *can*-class (*can, mia, nht, rye* and sa) with Nxt1 being identified as a new meiotic arrest class (Caporilli et al. 2013). About 78% of *Nxt1*-dependent transcripts were testis-specific, however, the majority of testis-specific transcripts were not *Nxt1*-dependent (Caporilli et al. 2013). Loss of function of Nxt1 results in failure of many transcripts to accumulate in primary spermatocytes. Strikingly, transcripts that were sensitive to the loss of Nxt1 were also dependent on tMAC for their transcription, which indicated a link between the testis-specific transcript machinery and the core mRNA export pathway (Caporilli et al. 2013). In tMAC dependent genes, however, a dose response for intron number, specifically, has a higher expression in Nxt1 trans-heterozygotes (Caporilli et al. 2013). Their proposed mechanism is that feedback from the export pathway to increase tMAC activity, the presence of the EJC and the increased affinity for Nxt1-Nxf1 dimer, results in higher transcript levels from intron-containing genes than intron-less genes in Nxt1 trans-heterozygotes (Caporilli et al. 2013).

In *Drosophila*, Thoc5 is part of the THO complex important for binding RNA transcripts for further processing (Jimeno et al. 2002). However, loss of Thoc5 showed male sterility and a meiotic arrest phenotype (Moon et al. 2011) This new data showed that the THO complex is involved in a specific cell differentiation program, the *Drosophila* spermatogenesis, most likely participating in the establishment or maintenance of nucleolar integrity in spermatocytes (Moon et al. 2011) .

1.3.2 CRM1-dependent RNA export

rRNAs, U snRNAs, and a subset of mRNAs are exported from the nucleus via a CRM1-dependent route (Grosshans et al. 2001; Fornerod, Ohno, et al. 1997). CRM1 (encoded in *Drosophila* by the gene *embargoed*, *emb*) is a member of the importin-beta family, most commonly known as protein export receptors. The function of CRM1 as an RNA export receptor was first shown when CRM1 was able to export unspliced mRNA of HIV in a mechanism dependent on Rev, a viral RNA binding protein (Fornerod, Ohno, et al. 1997; Neville et al. 1997; Wolff, Sanglier, and Wang 1997). To function in RNA nuclear export CRM1 needs an adaptor protein, as it cannot bind directly to RNA. Three adaptor proteins have so far been identified to interact with CRM1 during mRNA export: RNA-binding protein human

antigen R (HuR), leucine-rich pentatricopeptide repeat protein (LRPPRC), and nuclear export factor 3 (NXF3; Figure 1.5). RNAs with AU-rich elements (ARE) bind to HuR and CRM1 binds pp32 and APRIL in addition to HuR for export (Brennan, Gallouzi, and Steitz 2000). LRPPRC interacts with CRM1, eIF4E and RNA element 4E-SE (Topisirovic et al. 2009). Finally, NXF3 possesses a CRM1-dependent export signal and can tether RNA for exportation (Yang et al. 2001). In *Drosophila*, Nxf3 has low similarities and low identity compared to human NXF3. Nxf3 is similar to Nxf1 but lacks the C-terminal domain that is required for interaction with the NPC. There are several RNAs known to be exported via the CRM1-dependent route for which the adaptor proteins are still unknown (Hutten and Kehlenbach 2007).

U snRNAs, transported via the CRM1-dependent route, are transcribed by RNA pol II, have a m^7G-cap and a CBC (Kohler and Hurt 2007). The adaptor protein PHAX can bind to the CBC and near the cap, followed by recruitment of CRM1-RanGTP (Ohno et al. 2000).

There are several factors involved in the mRNA transport via CRM1, including GTP binding RNA (Ran-GTP), RanBP1, RanBP2 (Nup358), RanBP3, and RanGAP. CRM1 can interact with the nuclear export signal (NES) of its cargo and RanBP3 supports this interaction (Englmeier et al. 2001). Like Nxf1, CRM1 interacts with FG repeat-containing Nups to pass the cargo through the central channel of the NPC. RanBP2 is associated with the cytoplasmic face of the NPC, whereas RanGAP and RanBP1 are soluble in the cytoplasm. The hydrolysis of GTP is required to dissociate RNA from the export factors and RanBP1 and RanGAP promotes this decoupling process (Kehlenbach et al. 1999). RanBP2 has been shown to have binding capacity with Nup88, Nup214, RanGAP, RanGTP, RanGDP, and Nxt1-Nxf1. RanBP2 uses a similar function to RanBP1 in dissociating mRNP from RanGAP (Kehlenbach et al. 1999; Hutten and Kehlenbach 2006; Fornerod, van Deursen, et al. 1997; Askjaer et al. 1999; Bernad et al. 2004; Forler et al. 2004).

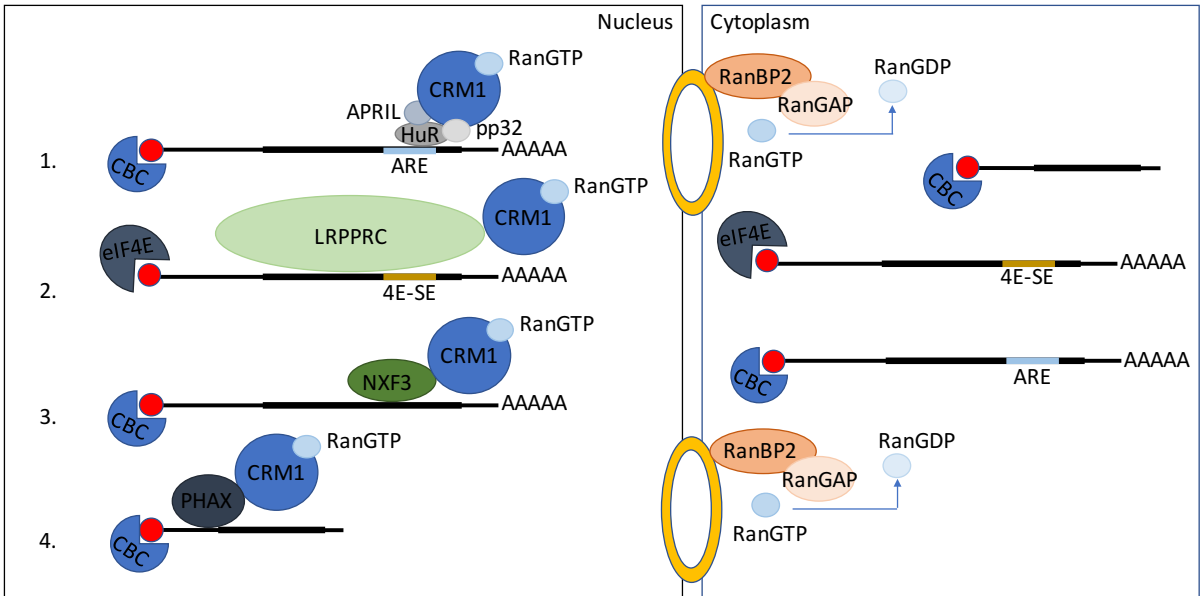


Figure 1.5: **CRM1-dependent RNA export pathway.** Several types of RNAs are transported via the CRM1-dependent pathway. These are (1) HuR-mediated; (2) LRPPRC-mediated; (3) NXF3-mediated; and (4) U snRNA export. Each of the pathways require binding of CRM1 with RanGTP. The hydrolysis of RanGTP to RanGDP is required to dissociate RNA from the export factors.

1.4 piRNA pathway in *Drosophila*

1.4.1 The piwi-interacting RNA pathway in the ovaries

Many organisms such as plants, fungi, and animals utilize small RNA silencing pathways to silence transposable elements (TEs) to prevent any harmful impact on genome integrity (Slotkin and Martienssen 2007). In animals, the TE silencing system is the gonad-specific piwi-interacting RNA (piRNA) pathway (Malone and Hannon 2009; Siomi et al. 2011). A specific set of Argonaute proteins - the Piwi proteins, are the core of this silencing pathway. These Piwi proteins are Piwi, Aubergine (Aub) and Argonaute 3 (AGO3) and are loaded with ~23-29nt long piRNAs to silence TEs. The length distribution of piRNAs is different than other small RNAs, such as siRNA (~21nt) or miRNAs (21-23nt).

In *Drosophila*, piRNA populations are highly enriched in TE sequences and most piRNAs originate from discrete genomic loci composed of TE fragments (Senti and Brennecke 2010). These so-called piRNA clusters are composed of many inactive TE sequences. These TEs are

unable to transposase and thus do not have mutagenic capacity. However, the piRNAs derived from these piRNA clusters can, together with Piwi proteins, and many other proteins repress active TEs.

1.4.2 Processes in the piRNA pathway

1.4.2.1 Two distinct pathways in the piRNA pathway

The piRNA pathway in the germline can be divided into a primary pathway and secondary pathway. The primary pathway is active in both somatic and germ cells, whereas the secondary pathway is solely active in the germ cells. Both pathways have their unique ways of effectively silencing transposons in the germline.

Primary pathway

In the primary pathway P-element induced Wimpy testis (Piwi) is the only protein present that can bind to piRNAs. Piwi only has the capability of binding antisense oriented piRNAs. The somatic cells only have the primary pathway available (Figure 1.6). piRNAs derived from *flamenco* (*flam*) precursor RNAs are mostly antisense oriented and can thus bind to Piwi to form a Piwi pi-RISC. This RISC complex is then imported to the nucleus. In the nucleus Piwi identifies targets that are complementary to the piRNA to which it is bound, most likely co-transcriptionally. Binding of the Piwi pi-RISC to the target results in deposition of the repressive marker H3K9me3, reduction of RNA pol II occupancy at promoters and a decrease in transcription at TE loci (Sienski, Donertas, and Brennecke 2012; Le Thomas et al. 2013; Rozhkov, Hammell, and Hannon 2013).Gametocyte-specific factor 1 (GTSF1) is a nuclear protein that interacts with Piwi and is part of the Piwi piRISC-mediated transcriptional silencing (Ohtani et al. 2013). Loss of GTSF1 decreases the level of H3K9me3 on sequence targeted by Piwi pi-RISC, leading to de-repression of transposons and female sterility (Ohtani et al. 2013). HP1a, interacting *in vitro* with Piwi, and histone methyltransferase Su(var)3.9, are also recruited to the Piwi pi-RISC target site, presumably to play an important role in H3K9me3 loading and spreading on TE DNA (Huang et al. 2013). Lastly, Maelstrom (Mael), is believed to function either downstream or in parallel with the H3K9 trimethylation step (Sienski, Donertas, and Brennecke 2012). Loss in Mael results in de-repression of transposons,

however, the amount of piRNAs loaded onto Piwi and the level of H3K9me3 on TE loci remain unchanged in *mael* mutants (Sienski, Donertas, and Brennecke 2012).

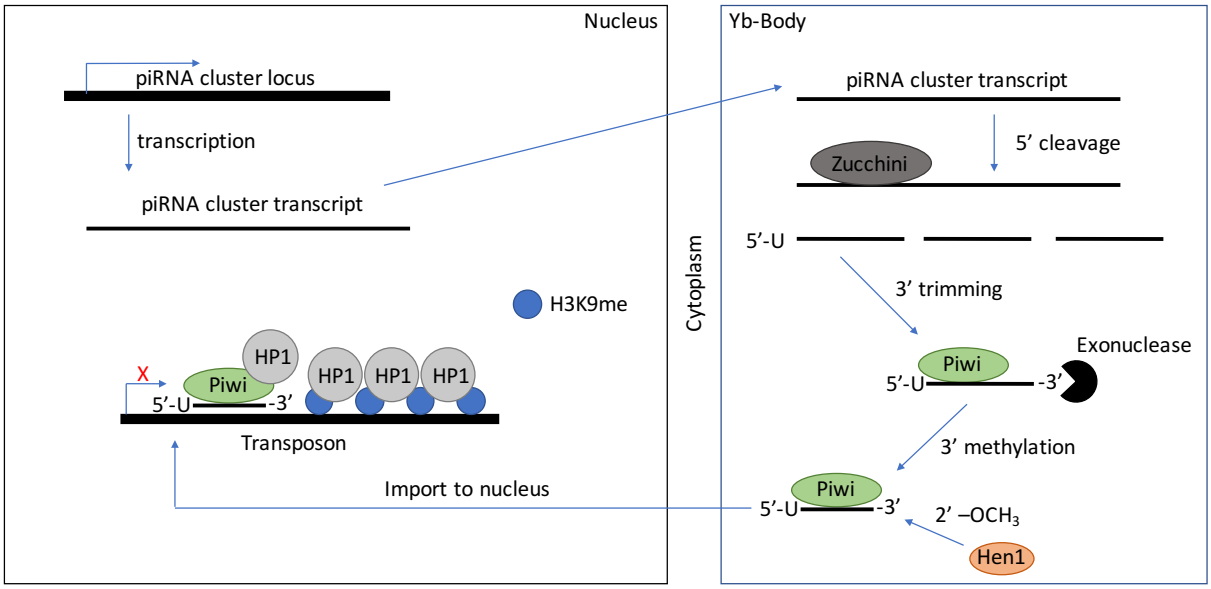


Figure 1.6: **Primary piRNA processing in ovarian somatic cells**. Primary processing starts with the transcription of piRNA cluster loci. These precursor transcripts are exported to the Yb-body. The precursor transcripts are 5' end cleaved by Zucchini followed by loading onto Piwi. An exonuclease trims the 3'end of the piRNA followed by 3' methylation before the Piwi/piRNA complex is imported into the nucleus. The Piwi/piRNA complex targets a transposon and silences it by recruiting the repressive marker H3K9me3 to the target.

Secondary pathway

The secondary pathway (also known as ping-pong amplification loop; Figure 1.7) is important to amplify the pool of primary (piRNAs originally derived from piRNA precursor RNAs) piRNAs via a ping-pong loop. This process requires the other two PIWI proteins, Aub and AGO3. This amplification process is functional only in the female germline, and the processing occurs in a specific subcellular compartment termed the nuage (Brennecke et al. 2007; Lim and Kai 2007; Gunawardane et al. 2007). Aub-associated piRNAs are mainly derived from antisense strands of TEs while AGO3-associated piRNAs are predominantly from sense strands. Aub pi-RISC targets mRNA produced from active TEs and slice the target with its slicing activity. With this slicing activity, Aub defines the 5' end of new, secondary, piRNAs that can then be loaded onto AGO3 and form an AGO3 pi-RISC. After maturation of this secondary piRNA, AGO3 pi-RISC cleaves complementary target transcripts coming from piRNA clusters. The cleavage

products from AGO3 pi-RISC can also be processed into secondary piRNAs, which can then be loaded onto Aub. piRNAs can also be maternally inherited and initiate secondary piRNA biogenesis, offering an alternative source of piRNAs to start the ping-pong cycle (Brennecke et al. 2008).

Piwi/Aub-associated piRNAs have a strong preference for 5'U (uracil) at the first nucleotide and AGO3-associated piRNAs prefer an adenine (A) at the tenth nucleotide from the 5' end. The slicer activity of Aub and AGO3 targets cleavage between the tenth and eleventh positions. This leads to a perfect complementarity along the first ten bases between primary and secondary piRNAs (Brennecke et al. 2007).

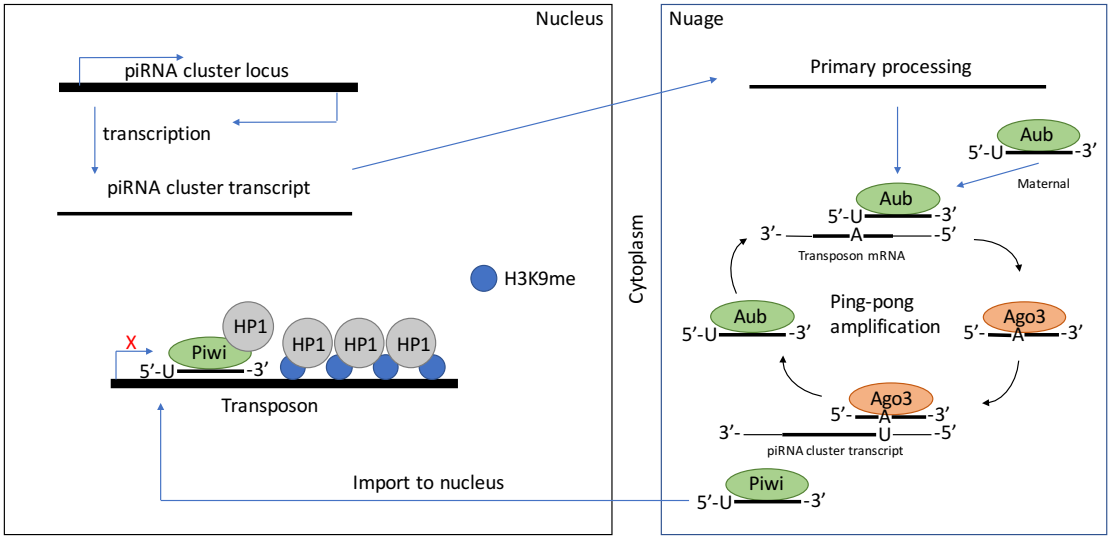


Figure 1.7: **Secondary piRNA processing in ovarian germ cells**. Transcription of piRNA cluster loci, often from both strands, is transcribed and exported to the nuage for primary processing. Secondary piRNA processing requires the input of Aub/piRNA complex from primary processing or maternally inherited piRNAs. During the secondary piRNA processing Aub/piRNA complex target a transposon mRNA for cleavage. The remnant of the mRNA is used by AGO3 to form a AGO3/piRNA complex. The AGO3/piRNA complex targets a piRNA cluster transcript for cleavage. The remnant from that can be used by Aub again to form an Aub/piRNA complex. This reinforces the antisense piRNAs bound to Aub. Primary piRNAs bound to Piwi are imported to the nucleus for transcriptional repression.

1.4.2.2 piRNA cluster biology

piRNA clusters are mostly found within heterochromatin or in its close proximity. Some of these piRNA clusters can reach lengths of up to several hundred kilobases with most of their sequence corresponding to TE fragments or other repeats. piRNA cluster loci can be transcribed as long single-stranded precursor RNAs that are substrates for the piRNA pathway (Siomi et al. 2011; Senti and Brennecke 2010). Unfortunately, little is known how these loci are defined and transcribed and their mRNA is protected against degradation in the nucleus and instead exported to the piRNA processing centres.

The piRNA clusters are identified via piRNAs mapping uniquely to the genome. This resulted in two types of clusters: Uni-strand clusters where only one genomic strand is transcribed with only piRNAs arising from that strand and dual-strand clusters in which both genomic strands are transcribed with piRNAs being generated from both strands.

Uni-strand clusters

In *Drosophila*, *flam* and *cluster 20A* are the most well-known uni-strand piRNA clusters. *Flam* is localized at the pericentromeric region of the X-chromosome and is strongly enriched in retrotransposons mostly inserted in the same orientation. *Flam* is only expressed in the somatic follicle cells and with most of its retrotransposons copies in antisense orientation the piRNAs produced from the precursor RNAs are antisense oriented (Brennecke et al. 2007).

Like most other piRNA clusters expressed in the follicle cells, *flam* is transcribed from a polymerase II promoter as a long single-stranded precursor RNA. Some of these derived precursor RNAs undergo differential alternative splicing (Figure 1.8, left). This could possibly generate diverse RNA precursors that all share the first exon at their 5' end before being processed into piRNAs (Goriaux et al. 2014). Although the biological role of these alternative splicing events is still unclear one can predict that this can contribute to a high diversity of *flam* precursors. These *flam* precursor RNAs are then transferred to cytoplasmic structures, called Yb-bodies where mature piRNAs are formed (Murota et al. 2014). Before precursor RNAs from *flam* and other clusters are transferred to Yb-bodies they are clustered to a single nuclear structure adjacent to Yb-bodies, called Dot COM (Figure 1.8, left) (Dennis et al. 2013).

It is believed that the formation of Dot COM occurs upstream of the cytoplasmic processing of transcripts as normal distribution and morphology of Dot COM is not changed with disruption in Yb-body components (Yb and armi) or piRNA processing (Zuc) (Dennis et al. 2013).

Dual-strand clusters

In *Drosophila*, cluster *42AB* is the largest dual-strand cluster with a length of about 240 kilobases. This is the most studied of the clusters active in germ cells. The processing of clusters in the germ cells is quite different than in somatic cells. Dual-strand clusters do not have a clear transcriptional start site (TSS), and instead their transcription is initiated at neighbouring genes. Their transcription depends on a protein complex named RDC, comprising Rhino (Rhi), Deadlock (Del) and Cutoff (Cuff) (Figure 1.8, right)(Mohn et al. 2014; Klattenhoff et al. 2009; Zhang et al. 2014; Pane et al. 2011). Rhi is a germline-specifically expressed HP1 homolog that specifically binds H3K9me3 residues on dual-strand clusters. Rhi can directly interact with Del through its chromoshadow domain and Del can physically interact with Cuff. This complex can repress RNA pol II termination leading to a read-through transcription. The RDC complex most likely also suppresses splicing or destabilizes spliced transcripts from these loci (Mohn et al. 2014). Cuff is also thought to compete with the CBC to bind uncapped precursor RNAs. This Cuff capping is believed to prevent RNA capping and splicing and possibly be a signal to the nucleus to transport these precursor RNAs to the cytoplasmic piRNA machinery (Mohn et al. 2014). It has been suggested that UAP56 is involved in transporting the precursor RNAs to the cytoplasmic nuage (Zhang et al. 2012). UAP56 colocalizes with Cuff and Rhi and is thought to bind nuclear piRNA precursor RNAs and escort them to the nuclear pores. From there, Vasa, localized at the nuclear membrane, likely take over and transfer the piRNA precursor RNAs from nuclear pores to cytoplasmic nuage (Zhang et al. 2012).

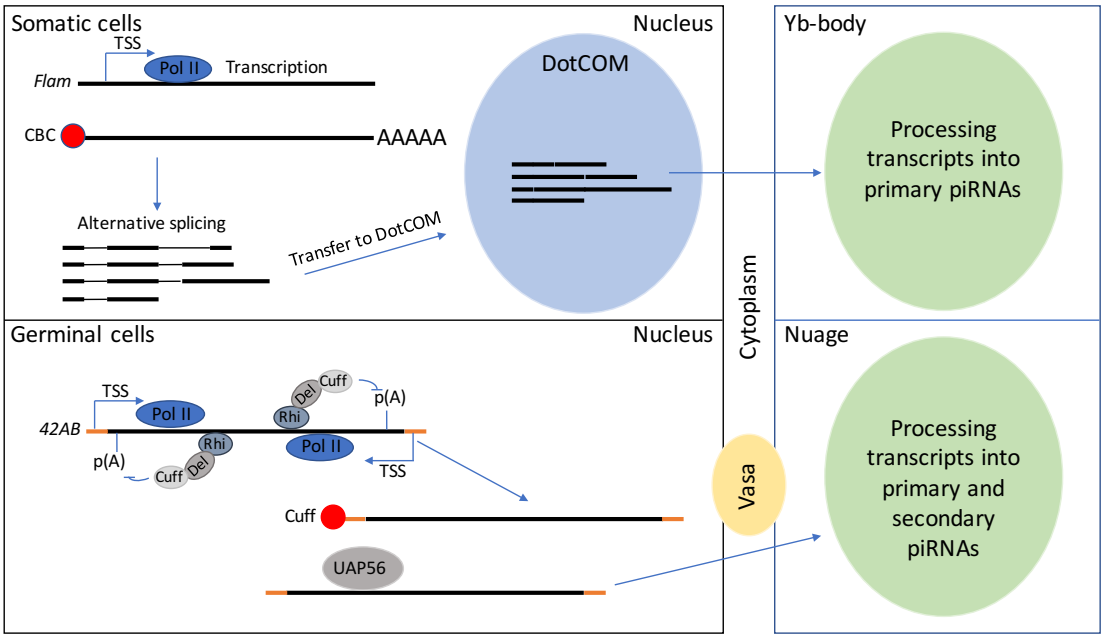


Figure 1.8: **Transportation of precursor transcript RNAs to the processing centres**. In somatic cells (top), piRNA clusters (*flam* depicted in figure) are mostly uni-strand and have a defined TSS. The precursor transcript is capped by CBC and alternatively spliced before moved to DotCOM foci. From there, the precursor transcripts are transported to the Yb body. In germline cells (bottom), piRNA clusters (*42AB* depicted in figure) are predominantly dual-strand. These clusters do not have a defined TSS and most likely uses one from neighboring genes. The Rhi/Del/Cuff complex interacts with the polyadenylation termination site leading to a precursor transcript with extensions at both the 3' and 5' end. Cuff competes with CBC to bind as a 5' cap to presumably prevent the transcript from being spliced. UAP56 binds to the precursor transcripts to help during transportation and Vasa take over later to get the precursor transcript to the nuage. Orange bar = neighbouring gene.

1.4.2.3 piRNA processing bodies

The precursor RNAs are transported to the piRNA processing bodies where the transcripts are processed into mature piRNAs. These processing bodies are different foci in somatic and germ cells and have a slight difference in their protein composition. In the somatic cells Yb-bodies process precursor transcript RNAs into primary piRNAs, whereas in germ cells, the nuage processes precursor transcript RNAs into primary and secondary piRNAs (Malone et al. 2009; Saito et al. 2010).

One of the first steps occurring in both Yb-bodies and nuages is the 5'end cleavage by Zucchini (Zuc), a member of the phospholipase-D family of phosphodiesterases, which includes both phospholipases and nucleases (Ipsaro et al. 2012; Nishimasu et al. 2012). Crystal structure of *Drosophila* DmZuc revealed a positively charged, catalytic groove that could accommodate a single –stranded, but not a double-stranded RNA (Nishimasu et al. 2012). Both DmZuc and the mouse homolog MmZuc exhibits endoribonuclease activity for single-stranded RNAs in vitro (Nishimasu et al. 2012). The cleavage of RNA results in a 5'-monophosphate group that is characteristic of mature piRNAs, instead of a 5'-phosphate group in siRNA and miRNA (Nishimasu et al. 2012; Elbashir, Lendeckel, and Tuschl 2001).

Somatic Yb-bodies

In somatic cells, several components are required in the Yb-bodies to process precursor transcript RNAs. The components include Tudor domain and helicase domain factor Yb, and the Yb-related protein, Sister of Yb (SoYb), co-chaperone Shutdown (Shu), Tudor domain containing protein Vreteno (Vret) and RNA helicase Armitage (Armi) (Handler et al. 2011; Olivieri et al. 2012; Olivieri et al. 2010; Saito et al. 2010; Zamparini et al. 2011; Preall et al. 2012). Mutations in any of these components lead to TE de-repression. Among these proteins, Yb is the only one expressed solely in the somatic follicle cells, whereas the others are also expressed in the germ cells. These proteins are all co-localized in the Yb-bodies and have a genetic hierarchy that reveals information about the assembly of the Yb body: (Yb -> Armi -> Vret -> Shu). Localization of Armi is dependent on Yb protein (Olivieri et al. 2010; Saito et al. 2010), whereas Vret localization is dependent on Armi and Yb (Handler et al. 2011; Zamparini et al. 2011), with Shu being dependent on Vret, Armi, Yb and Piwi for its localization (Olivieri et al. 2012; Preall et al. 2012). More recently, the *Drosophila* Gasz protein, a homolog of the mouse Gasz, has been postulated to be a mitochondrial transmembrane protein that serves as an adaptor to recruit Armi to mitochondria in both somatic and germ cells (Czech et al. 2013; Muerdter et al. 2013; Handler et al. 2013). Yb-bodies are often attached to the mitochondria to which Zuc localizes at the outer membranes for endoribonuclease processing (Murota et al. 2014).

Germ cell nuage

A perinuclear electron-dense structure called the nuage is the piRNA processing body specific to the germline. The piRNA process is similar in somatic Yb bodies: 5' cleavage, loading pre-piRNAs onto PIWI protein and maturation (3' trimming and 2'-O-methylation). Proteins involved in the nuage are also very similar to those in the Yb-bodies with the germline protein paralogue Brother of Yb (BoYb) believed to replace the soma-specific Yb. In the nuage both primary pathway and secondary pathway are active unlike the Yb-body with only the primary pathway active. This secondary pathway requires the other two PIWI proteins, Aub and AGO3. Additional proteins also used in the germline include Vasa, Spindle-E (Spn-E), Papi, Tejas (Tej), Tapas (Tap), Krimper (Krimp) and Qin/Kumo (Malone et al. 2009; Liu, Qi, et al. 2011; Vagin et al. 2006; Lim and Kai 2007; Anand and Kai 2012; Zhang et al. 2011; Patil et al. 2014; Patil and Kai 2010).

1.4.2.4 piRNA maturation

After the precursor transcript RNAs are processed in the processing bodies (nuage or Yb-bodies) the pre-piRNA, piRNAs slightly longer than their mature form, will bind to the PIWI proteins to from the RISC complex. At that stage, the pre-piRNA is processed to become a full mature piRNA. The two most important changes are the 3' trimming by an exonuclease followed by 2'-O-methylation. After these final steps, the RISC complex is ready to silence transposons.

3' trimming

5' processing of piRNAs has been known for some time, whereas the details of 3' processing have been unknown. Recently, more has been discovered about this 3' formation. Zucchini, a critical player in the 5' processing of piRNAs, has also been shown to be involved in 3' maturation (Han, Wang, Li, et al. 2015; Mohn, Handler, and Brennecke 2015). Secondary piRNAs (loaded into Aub/AGO3) are specified via piRNA-guided slicing. In wild type, these piRNAs display the characteristic 1U (Aub piRNAs) and 10A (AGO3 piRNAs), but also exhibit a downstream U bias at their 3' end. In Zucchini-depleted ovaries, piRNA 3' ends have no U bias suggesting that Zucchini leaves a downstream U-bias fingerprint at the 3'end. Altogether, this

shows that Zucchini is the central nuclease in piRNA biogenesis (Mohn, Handler, and Brennecke 2015).

It has been suggested that a 3' to 5' exonuclease would trim 3' ends of piRNAs, but only recently more insight have been revealed. *Drosophila* Nibbler (Nbr) is a 3' to 5' exonuclease known to trim 3' ends of several miRNAs (Han et al. 2011; Liu, Abe, et al. 2011). However, loss of *nbr* also impacts the length of piRNAs (Feltzin et al. 2015). A subset of piRNAs have a higher abundance of piRNA reads with longer length than in wild type. Interestingly, Feltzin and colleagues showed that piRNAs generated from the *flam* locus were not affected and that Nbr only acts in the germline cells. Later research has shown that Nbr had a more global effect in the piRNA pathway and acts both in germline and somatic cells (Wang et al. 2016). Unlike Feltzin and colleagues, Wang and colleagues calculated the average piRNA length derived from both *42AB* and *flam* piRNA clusters. This showed a slight increase in piRNA length (*42AB* forward 0.62nt, *42AB* reverse 0.40nt and *flam* forward 0.23nt), clearly indicating a mild affect in *Nbr* mutant in both germline and somatic cells (Wang et al. 2016). Hayashi and colleagues found that there are two parallel pathways with varying contributions towards 3' end formation for Aub/AGO3-bound piRNAs: Zucchini generates most Aub-bound piRNAs, whereas Nibbler generates most AGO3-bound piRNAs (Hayashi et al. 2016). The length of piRNAs binding to the Piwi proteins can be distinguished with Piwi-bound piRNAs around 26nt, Aub-bound piRNAs are 25nt and AGO3-bound piRNAs are 24nt long (Brennecke et al. 2007).

Methylation

The final step in the process of maturing piRNAs is 2'-O-methylation at the 3'ends. This is catalyzed by Hen1, a methyl transferase (Saito et al. 2007; Horwich et al. 2007). In *Arabidopsis,* methylation by Hen1 has a protective effect (Li et al. 2005; Yu et al. 2005). All fly piRNAs are 2'-O-methylated, and knockdown of *Hen1* shortened the piRNAs dramatically, revealing the importance of this modification in Drosophila and provided a first insight of Hen1 at the genome level (Wang et al. 2016).

1.4.3 The piwi-interacting RNA pathway in the testis

In the *Drosophila* testis, the most well-known piRNA-mediated pathway is the interaction between *crystal* (*cry* or *Su(Ste)*) and *Stellate* (*Ste*) (Li, Vagin, et al. 2009; Vagin et al. 2006; Nagao et al. 2010; Bozzetti et al. 2012). In wild type testis, piRNAs are mainly produced from the Y-linked *cry* locus to repress the X-linked *Ste* RNA. Males that lack *cry,* e.g. XO males that lack a Y chromosome, have defects in chromosome condensation and segregation and accumulate cytoplasmic crystalline aggregates, due to production of the Ste protein (Bozzetti et al. 1995; Palumbo et al. 1994). All the *cry-Ste* modifiers are involved in the piRNA pathway and also mutants show both sterility and germline activation of transposable elements (Li, Vagin, et al. 2009; Nagao et al. 2010; Vagin et al. 2006; Zhang et al. 2011; Specchia et al. 2010). The piRNAs from the *cry* locus are mostly antisense oriented and bound to Aub (Vagin et al. 2006; Aravin et al. 2004).

Another large class of piRNAs, associated with Aub, is derived from a short-repeated region, termed *AT-chX,* on the X chromosome (Nishida et al. 2007). These piRNAs are mostly antisense oriented, bind to Aub and has strong complementarity to *vasa* mRNA.

The two largest classes of piRNAs present in testis, *cry* and *AT-chX*, do not bind to Piwi (Nishida et al. 2007) and have no clear ping-pong signature. However, Aub and AGO3 are present in the testis and low functional ping-pong amplification might occur, similarly as shown in the ovaries (Nagao et al. 2010). *AGO3* mutants showed that AGO3 is required for the accumulation of both *cry* and *AT-chX* piRNAs in fly testes (Li, Vagin, et al. 2009).

1.4.4 The piwi-interacting RNA pathway in other species

Piwi proteins were first identified in *Drosophila melanogaster* in a screen to identify factors involved in germline stem cell (GSC) maintenance (Carmell et al. 2002; Lin and Spradling 1997). After that, Piwi proteins were shown to also be involved in GSCs in other organisms (Cox et al. 1998). Several years later, multiple studies identified piRNAs from mouse and rat germ cells (Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Lau et al. 2006; Watanabe

et al. 2006). Here I discuss further the differences and similarities of the piRNA pathway in *Mus musculus* and *Caenorhabditis elegans* (Figure 1.9).

1.4.4.1 *M. musculus*

The mouse genome contains three Piwi proteins, MIWI (PIWI1), MILI (PWIL2) and MIWI2 (PWIL4), which are all required for male but not female fertility (Carmell et al. 2007; Deng and Lin 2002; Kuramochi-Miyagawa et al. 2004). The Piwi proteins are expressed at different stages in testes: MILI expression starts at embryonic day 12.5 and persists into adulthood. MIWI2 expression is between 15 days post coitum (dpc) to 3 days post partum (dpp), whereas MIWI is expressed in the adult testes from 14 dpp. *Mili* and *Miwi2* knockout have phenotypes during spermatocyte meiosis and *Miwi* has a phenotype at the round spermatid stage (Kuramochi-Miyagawa et al. 2004; Deng and Lin 2002). The length of the piRNAs binding to the PIWI proteins can be distinguished with MILI-bound piRNAs around 26 or 27nt, MIWI2-bound around 28nt and MIWI-bound peak at 30nt (Aravin et al. 2006; Aravin et al. 2008; Girard et al. 2006).

There are two types of piRNAs present during male germline differentiation: pre-pachytene piRNAs that bind to MILI and MIWI2 and pachytene piRNA that bind to MILI and MIWI. Pre-pachytene piRNAs are present in the germ cells of foetal and newborn mice and are derived from a wide variety of loci covering ~0.2% of the mouse genome (Aravin et al. 2007). Pachytene piRNAs originate from intergenic loci and are depleted of repeat sequences (Aravin et al. 2008; Li et al. 2013). For the pachytene piRNAs, a transcriptional master regulator A-MYB (MYBL1) has been identified that induces pol II-mediated transcription of long pachytene piRNA precursors and pathway proteins (Li et al. 2013). PLD6, the mouse homolog of the *Drosophila melanogaster* Zuc, is involved as a nuclease, paralleling the situation in *Drosophila* (Ipsaro et al. 2012). The 3' end of piRNAs is trimmed by an unidentified exonuclease, then the 3' end becomes 2'-O-methylated by the mouse homolog of HEN1 (HENMT1), followed by binding to the Piwi protein via the PAZ domain (Kirino and Mourelatos 2007a, 2007b).

As in *Drosophila melanogaster*, a secondary amplification of piRNAs exists in mouse. However, this ping-pong cycle only occurs between MILI and MIWI2 bound pre-pachytene piRNAs (Figure 1.9B). This cycle will allow for amplification of the piRNA pool.

1.4.4.2 *C. elegans*

The piRNAs in *C. elegans* are often referred to as 21U-RNAs due to their 21nt length and 5'U bias. All piRNAs are bound to PRG-1, which is the only functional *C. elegans* Piwi homolog (Batista et al. 2008; Wang and Reinke 2008). PRG-1 is required for the stability of 21U-RNAs and is unlikely to have a role anywhere else in the piRNA biogenesis (Das et al. 2008). There is another Piwi gene encoded in the *C. elegans* genome, *prg-2*, but it has no defects in 21U-RNA expression in the piRNA pathway (Das et al. 2008; Batista et al. 2008).

Most of the piRNA genes of *C. elegans* are found in two clusters on chromosome 4, where each piRNA is downstream of a motif referred to as the Ruby motif (Ruby et al. 2006). This sequence acts as an autonomous promoter for the individual piRNA precursors (Billi et al. 2013; Cecere et al. 2012). The 21U-RNAs precursor transcripts carry a 5' 7-methylguanylate cap and 21U-RNAs are likely generated by pol II with transcription 2nt upstream of the 5'U of the mature piRNA (Figure 1.9C). Forkhead proteins, including UNC-130, FKH-3, FKH-4 and FKH-5, are required for the transcription of piRNAs from the Ruby motif. Depleting any of the Forkhead proteins results in decreased levels of mature piRNAs (Cecere et al. 2012). piRNA defective 1 (PRDE-1) has been shown to be essential for the generation of Ruby-motif dependent piRNA precursors and is expressed in pachytene germ cells. PRDE-1 is either involved in stabilizing the precursors after transcription or generating the precursors by direct or indirect interaction with the Ruby motif (Weick et al. 2014). In an RNAi based genome-wide screen, several factors were identified to be required for the *C. elegans* piRNA pathway, which are referred to as "twenty-one U fould up" (TOFU) (Goh et al. 2014). Several of these TOFUs have been shown to be required for precursor production, precursor processing and chromatin remodeling. As has been described in other organisms, piRNAs are 2'-O 3'end methylated by the *C. elegans* HEN1 ortholog HENN-1 (Kamminga et al. 2012; Montgomery et al. 2012).

The *C. elegans* has a unique mechanism of the piRNA biogenesis and stability. To date, no piRNA biogenesis factor that is conserved between *C. elegans* and *Drosophila melanogaster* or mice have been identified. Furthermore, *C. elegans* have no ping-pong mechanism that amplifies piRNAs (Das et al. 2008). More interestingly, worms have an analogous signal amplification mechanism that leads to generation of secondary downstream siRNAs upon piRNA:target RNA interaction. The generation of these RNA-dependent RNA polymerases (RdRPs) secondary siRNAs, also known as 22G-RNAs, are incorporated downstream into an Argonaute protein that mediates target silencing of pre-mRNA (Figure 1.9C). The *C. elegans* piRNA pathway also mediates silencing via RdRP-dependent generation of secondary 22G-RNAs (Bagijn et al. 2012; Lee et al. 2012; Das et al. 2008). The piRNA pathway mediates gene silencing at the pre-mRNA level and silencing depends on several chromatin factors, including the *C. elegans* homolog of H3K9me3 HP1, HPL-2, and several histone methyltransferases (Ashe et al. 2012; Shirayama et al. 2012). This nuclear silencing downstream of piRNAs is similar to the well-characterized nuclear RNAi pathway acting in somatic tissues. A germline-specific Argonaute protein, heritable RNAi-deficient 1 (HRDE-1), binds secondary 22G-RNAs, shuttling into the nucleus and initiating H3K9me3 methylation and pol II stalling.

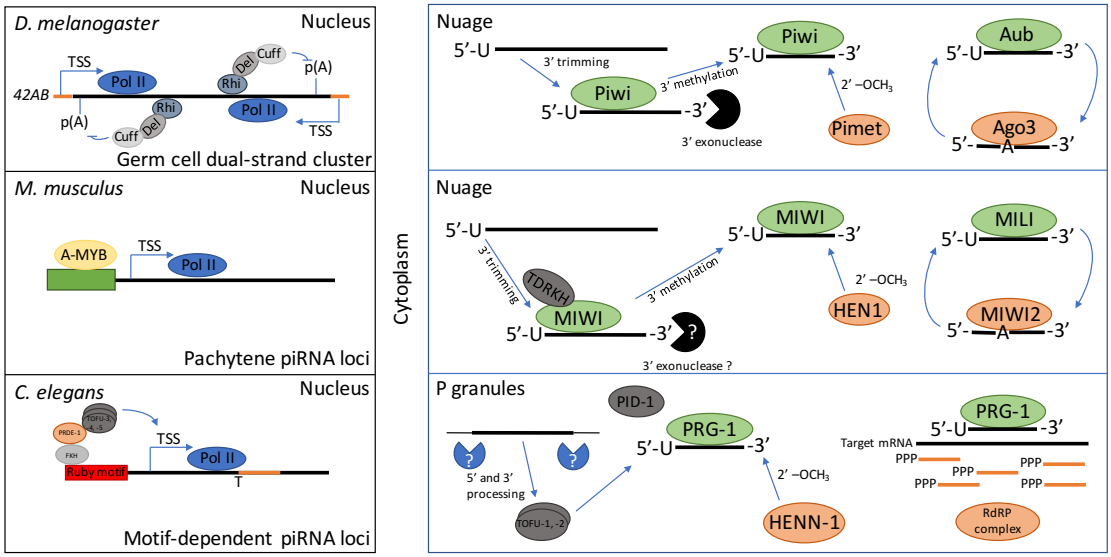


Figure 1.9: **piRNA pathway in *D. melanogaster, M. musculus* and *C. elegans.*** In *D. melanogaster*, piRNAs can be transcribed from dual-stranded clusters by Pol II from neighboring genes. Rhi, Del and Cuff are required during this process to protect the precursor transcript from splicing and degradation. UAP56 is believed to transport the transcript to the nuage for further processing. In the nuage, precursor transcripts are loaded onto Piwi or Aub, 3’ trimmed and 2’-O-methylated. In *M. musculus*, piRNAs produced from pachytene piRNA loci require A-MYB for recruitment of Pol II for transcription. In the nuage, MIWI bind to the precursor transcript for further processing into a mature piRNA by 3’ trimming and modification. In *C. elegans*, a conserved Ruby motif is bound by Forkhead proteins and most likely additional factors. Transcription and/or stability are dependent on several TOFU proteins. In P granules, precursor transcripts are further processed and piRNAs are bound to PRG-1 for targeting. PRG-1:piRNA:target RNA leads to downstream 22G-RNAs carrying a 5’triphosphate (indicated by PPP) by a multi-protein machinery containing RdRPs. These small RNAs are then incorporated into AGO protein and mediate target silencing. Secondary amplification processing is only available in *D. melanogaster* and *M. musculus*, but not *C. elegans*.

1.4.5 Link between mRNA export and piRNA pathway

In 2012, Czech and colleagues published a transcriptome-wide RNAi screen that identified several new genes potentially involved in repressing transposons (Czech et al. 2013). Some of these are involved in the mRNA export pathway, such as CBP80 and CBP20, part of the CBC complex, Thoc5, Thoc6, Thoc7, Thoc2 and Hpr1, part of the THO complex or Nxt1, Nxf1 and UAP56, transient factors of the EJC (Czech et al. 2013).

UAP56 is a ubiquitously expressed DEAD box protein involved in splicing and RNA export (Shen 2009). However, UAP56 mutants disrupt germline piRNA production, transposon silencing and perinuclear localization of Vasa (Zhang et al. 2012). The proposed function of UAP56 is to associate with nascent transcripts from dual-strand piRNA clusters, which bind the HP1 homolog Rhi. UAP56 potentially interact with clusters through a domain defined by residue 245E (Zhang et al. 2012). UAP56-cluster transcript complexes interact with the nuclear pore, while Vasa binds cluster transcripts within the pore or as they emerge from the pore to deliver these RNAs to the processing centres (Zhang et al. 2012).

The conserved THO/TREX complex is required for pre-mRNA nuclear export. In metazoans, association of TREX with pre-mRNA depends on splicing and requires the interaction of several TREX subunits with the nuclear cap-binding proteins CBP20 and CBP80 (Cheng et al. 2006). Another subunit of the TREX complex is UAP56, and as mentioned earlier, is required for dual-strand piRNA cluster export (Zhang et al. 2012). Thoc5 is part of the THO complex and Thoc5 mutants are viable but have spermatogenesis defects (Moon et al. 2011). However, in *Drosophila* germline, Thoc5 is also required for biogenesis of piRNAs (Hur et al. 2016). Similarly, to UAP56, Thoc5 only has an effect in dual-strand piRNA clusters, where it leads to accumulation of piRNA precursors by enabling efficient transcription of dual-strand cluster loci (Hur et al. 2016). Thoc5 forms a complex with UAP56 in an RNA-independent manner (Hur et al. 2016). More importantly, a novel splicing-independent mechanism has been proposed of TREX loading in which Rhi and Cuff result in locus-specific, but sequence-independent, loading of this complex on nascent transcripts (Hur et al. 2016).

The Nxt1-Nxf1 dimer, required for mRNA export (see section 1.3.1.2), has also been shown to affect transposon de-repression (Czech et al. 2013). Given its known canonical role of binding to mRNA transcripts, there is a high possibility that the Nxt1-Nxf1 dimer is also necessary for transportation of piRNA precursor transcripts from the nucleus to the piRNA processing centres.

1.4.6 Transposons

Transposition of transposons causes insertions, deletions and chromosomal rearrangements that could lead to premature lethality or offspring sterility in *Drosophila.* Eukaryotic genomes have a wide variety of transposons, which spread and maintain in the genome by colonizing new genomic locations in germline cells. Transposons can be categorized as retrotransposons (class I), which move via an RNA intermediate, or DNA transposons (class II), which use a "cut-and-paste" mechanism (Figure 1.10). Most of these transposons are repressed in the germline cells while only a few of them are exclusively repressed in the somatic cells such as *gypsy, ZAM* and *idefix* (Mevel-Ninio et al. 2007; Desset et al. 2008; Prud'homme et al. 1995) who are all derived from the *flamenco* locus.

1.4.6.1 Class I retrotransposons

Retrotransposons have an RNA intermediate that encodes a reverse transcriptase that is used for their movement mechanism. Retrotransposons can be divided into Long terminal repeat (LTR) and non-LTR retrotransposons. LTR retrotransposons are similar to retroviruses and can contain several hundreds of nucleotide terminal repeats at both 5' and 3' ends. A typical LTR retrotransposon has an open reading frame (ORF) in the internal region flanked by two LTRs. The ORF itself is divided into two regions: *gag* and *pol*. *gag* encodes an RNA-binding protein gag, and *pol*, encodes for several proteins, such as protease, integrase, reverse transcriptase, and RNase H, which are required for cDNA synthesis and integration of cDNA into the host chromosomes (Boeke and Corces 1989). LTR retrotransposons can be divided into different subfamilies: Ty1-copia-like (Pseudoviruses / Hemiviruses / Sireviruses), Ty3-gypsy-like (Metaviruses / Errantiviruses) and Pao-BEL-like (Semotiviruses). These subfamilies are based on the amino-acid sequences of the RT protein, which is the most highly conserved of the

retrotransposons proteins. The three major retrotransposons families can also be distinguished by the order of the coding regions within their *pol* genes.

Non-LTR retrotransposons do not have terminal repeats but have conserved reverse transcriptases. Both LTR and non-LTR have an internal promoter in the 5'UTR and a polyadenylation signal at the 3'UTR (Russo, Harrington, and Steiniger 2016; Gogvadze and Buzdin 2009). The transposition is an RNA pol II dependent transcription of the entire element followed by translation of each independent ORF in different reading frames.

1.4.6.2 Class II DNA transposons

Class II DNA transposons can only move via a cut-and-paste mechanism in which the transposon is excised from its current position and pasted at a different position within the genome. The main differences between class I and class II transposons are that almost all class II transposons move through a non-replicative mechanism, while all class I transposons move through a replicative mechanism. DNA transposons have a transposase gene that is flanked by two terminal inverted repeats (TIRs). The transposase can recognize these sequences and excise the transposons and insert it into a new location.

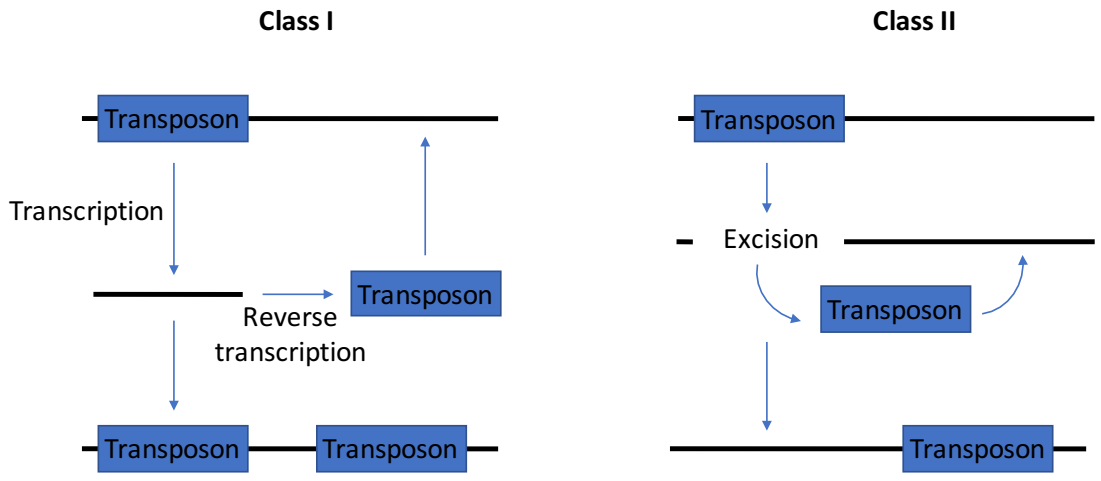


Figure 1.10: **Main two classes of transposable elements**. Class I transposable elements are transcribed and reverse transcription occurs before inserted back into the genome site. This results in multiplication of the transposon. Class II transposable elements are excised by the transposase and inserted into a new location into the genome. TEs can move to other chromosomes within the same cell.

1.5 Aims and objectives

Current research has revealed a lot about the canonical role of Nxt1 in the RNA export pathway. The discovery of Nxt1 as a meiotic arrest gene has expanded the role of Nxt1 in *Drosophila melanogaster*.

The overall aim of this project is to understand the semi-lethality in pupa and the sterility in female flies. I will look more closely in the processes that occurs during the larva and pupa phase. I will also look at the gene expression in ovaries, since the sterility in males cannot be explained in females due to the lack of the meiotic arrest complex.

This research is divided into two separate projects and the following objectives were set:

1) **Understanding why an RNA export factor, Nxt1, leads to a specific air bubble phenotype causing semi-lethality in pupa.**

To better understand the importance of Nxt1 during the metamorphosis I will use whole larvae RNA-seq from wandering larvae, stationary larvae and white prepupae to analyse the differential gene expression and the potential involvement of ecdysone in the air bubble phenotype. If ecdysone is not involved, then I will focus on the larval muscles that are required for the air bubble to migrate to the anterior end.

2) **Understanding how Nxt1 is involved in the sterility of female flies.**

Previous microarray data of 7 days old ovaries have shown that transposons are differentially expressed in Nxt1 trans-heterozygotes. I will use small RNA-seq to map piRNA reads to the transposons and piRNA precursor transcripts to understand how big an effect of loss of Nxt1 has in Nxt1 trans-heterozygotes. With these data, I will continue to look at different parts of the piRNA pathway and where Nxt1 plays a role.

Chapter 2.

Materials & Methods

2.1 Molecular Biology

2.1.1 Primer Design

Primers for real-time quantitative PCR analysis were designed by using PrimerBLAST. Primers were designed to amplify products between 70-150 nucleotides, contain 55-70% GC content, low self-complementarity, low self 3' complementarity and to have a melting temperature of around 60 degrees. All primer pairs were tested by PrimerBLAST to ensure that it is specific for the target of interest. Finally, during qRT-PCR the melting curve data was used to confirm only one product was produced.

2.1.2 Polyacrylamide Gel

Small RNA fragments were separated by size using 15% polyacrylamide/urea gels. Bio-rad gel cast system was used to prepare and run polyacrylamide gels. The 15% separating gel was prepared with 15% acrylamide/bis (Sigma), 0.38 M Tris, pH 8.8, 0.1% ammonium persulfate (APS), 8M urea and 10 μl TEMED. The 5% stacking gel was prepared with 5% acrylamide/bis (Sigma), 0.38 M Tris, pH 6.8, 0.1% APS and 10 μl TEMED. Samples were prepared accordingly (see section 2.1.3.1) and denatured at 95^{o}C for 1 minute and put on ice before loading. Samples were loaded on the gel with at least 1 empty lane in between to prevent any contamination. ZR small-RNA^{tm} Ladder (Bioscience) was used to size the small RNA fragments. Gels were analysed using either the Dark Reader Transilluminator (Clare Chemical Research) or a GelDoc-It®-310 documentation system (UVP) using UV light. RNA fragments were excised using scalpel blades.

For 12% and 10% polyacrylamide/urea gels, the final percentage of acrylamide/bis (Sigma) was changed to 12% and 10%, respectively.

2.1.3 RNA Extraction

2.1.3.1 Small RNA Isolation

30 pairs of ovaries were collected in 1x PBS per replicate for 4, 14 and 20 days old females wild type and *Nxt1* ^{DG05102}/ *Nxt1*$^{z2\text{-}0488}$, flash frozen in liquid nitrogen and stored at -80ºC. Total RNA (including micro RNA) was purified with miRNeasy Micro Kit (Qiagen) following manufacturer's protocol.

The RNA was quantified on a nanodrop ND-1000 and 40µg of total RNA, mixed with RNA Gel Loading Dye (2x) (ThermoFisher Scientific) was run on a 15% polyacrylamide gel for 30-45 min at 100V together with 20ng of ZR small-RNA^{tm} Ladder (Bioscience).

The bands ranging 21-29nt were cut from the gel using a scalpel blade and purified by the ZR small-RNA^{tm} PAGE Recovery Kit (Bioscience) following manufacturer's protocol. The purified small RNA was sent to the sequencing facility at the University of Liverpool for library preparation and sequencing. Libraries were prepared with the NEBNext® Small RNA Library Prep Set for Illumina (NEB).

2.1.3.2 Isolation of Adult Ovaries (Sequencing)

Seven days old females wild type and *Nxt1* ^{DG05102}/ *Nxt1*$^{z2\text{-}0488}$ were dissected for ovaries in 1x PBS, flash frozen in liquid nitrogen and stored at -80ºC. A total of 500µl ovaries (settled volume = approx. 500 flies) were collected for both genotypes. For RNA extraction, 50µl ovaries were used. 500µl of Trizol was added and the ovaries were smashed with a smasher. After 5 minutes, 100µl chloroform was added and samples were mix vigorously for 20-30 seconds. Samples were spun for 15 minutes at 4ºC at 13.000 rpm (Heraeus Biofuge Pico). The upper layer, containing RNA, was pipetted into a new tube and 160µl 100% isopropanol (around 80% of total) was added and centrifuged for 1 hour at 4ºC at max speed for precipitation. The pellet was washed with 70% ethanol and resuspended in 130µl RNAse-free water. The extracted RNA was quantified with a nanodrop ND-1000 and stored at -80ºC until further use.

2.1.3.3 Isolation of Whole Larvae (Sequencing)

RNA was extracted for wandering larvae, stationary larvae and white prepupae. Fly crosses for both *w^{1118}* and *Nxt1 ^{DG05102}/Cyo Actin GFP (CAG) x Nxt1$^{z2\text{-}0488}$ /CAG* were placed in food vials containing 0.05% bromophenol blue. Wandering larvae have blue guts, whereas stationary larvae have no blue or slightly light blue-ish guts. White prepupae were selected within an hour into the pupa stage. Gender was not determined during selection. For RNA sequencing, 30 animals were collected per replicate in triplicates. Total RNA was extracted using a combination of Trizol (ThermoFisher Scientific) and the RNeasy Mini Kit (Qiagen) according to manufacturer's protocol. The DNaseI step was included. RNA was quantified with a nanodrop ND-1000 (ThermoFisher Scientific) and stored at -80^{o}C.

2.1.3.4 Isolation of Larval Carcasses

Fly crosses for both *w^{1118}* and *Nxt1 ^{DG05102}/CAG x Nxt1$^{z2\text{-}0488}$/CAG* were put in food vials mixed with 0.05% bromophenol blue. The distinguishing between wandering and stationary larvae was done as explained in section 2.1.3.3.

The number of larval carcasses used depended on the experiment. For RNA sequencing 30 carcasses were used per replicate in triplicates. For qRT-PCR single carcass or a mix of up to 10 carcasses were used. RNA was prepared with Trizol and chloroform as described in 2.1.3.2.

Total RNA was extracted using a combination of Trizol (ThermoFisher Scientific) and the RNAqueous®-Micro Total RNA Isolation Kit (ThermoFisher Scientific) following manufacturer's protocol. The extracted RNA was treated with DnaseI (Life Technologies), quantified with a nanodrop ND-1000 (ThermoFisher Scientific) and stored at -80^{o}C.

2.1.4 cDNA Synthesis

cDNA synthesis was performed using the superscript III First Strand synthesis system (ThermoFisher Scientific). 100 ng or 1µg of total RNA was added to a 1.5 ml tube with 3.85 ng of random primers and 0.77mM dNTPs, made to a final volume of 13µl with dH_2O. The tubes were heated to 65^{o}C for 5 min and on ice for at least 1 min. 50mM Trish-HCl (pH 8.3), 75mM KCl and 3mM $MgCl_2$ (1x First-Strand Buffer), 40 units RNaseOUT^{tm} and 200 units Superscript

III RT (ThermoFisher Scientific) were added for a total volume of 20µl. The tubes were incubated at 50°C for 60 min, followed by inactivation by heating to 70°C for 15 minutes. cDNA was stored at -80°C until further use.

For qRT-PCR, the final volume was typically further diluted to 60µl (ovaries and whole larvae) or 120µl (larval carcasses).

2.1.5 Real Time Quantitative PCR

The cDNA reaction was diluted to 60 or 120µl with dH_2O, and 1µl of this cDNA was further diluted with 7µl dH_2O to use as a template in the qRT-PCR reactions. 10µl PowerSybr reagent (ABI) with 1µl of each target-specific forward and reverse primers (**full list shown in Appendix 1**) were added for a total reaction of 20µl. The qRT-PCR was performed on a Chromo4 instrument (MJR) using the Opticon Monitor 3 software. Expression of the Nxt1 trans-heterozygotes transcript was compared to the wild type transcript with *Rp49* used as normalization. All reactions were performed in triplicate.

2.2 *Drosophila* cultures and genetics

2.2.1 General Fly Work

Flies were maintained in standard food vials containing cornmeal, yeast, dextrose and plant agar. Proprionic acid and Nipagen (10% p-hydroxy benzoic acid methyl ester) were added as anti-microbial and anti-fungal respectively.

Virgin female and male flies for experiments were collected by anaesthetization with carbon dioxide emitted from a pad. Flies were collected in food vials and stored at 25°C until required.

Setting up a cross, approximately 1:3 male to female ratio was used with a minimum of 3 females per cross. Dried yeast was added if females did not lay enough eggs. After 3-4 days, flies were transferred into a fresh new tube. The progeny was used to establish new crosses. For larva/pupa experiments, larva and/or pupa were picked out, screened for absence or presence of the green fluorescent protein (GFP) marker, and used for experiments.

2.2.2 Fly Stocks

A complete list of fly stocks can be found in **Appendix 2**. A brief description for some stocks can be found below. The w^{1118} fly strain was used as a wild type.

2.2.2.1 Nxt1 Trans-Heterozygotes

To acquire the phenotype observed in Nxt1 mutants, two alleles were used. y^{1} w^{67c23}; $P[wHy]Nxt1^{DG05102}$/SM6a (Bloomington Drosophila Stock Center), which carries a P-element with a transposition-component hobo element and is presumed to be a null allele as no protein can be detected (Caporilli et al. 2013). The other allele, $Nxt1^{z2\text{-}0488}$, is a hypomorph allele with a prediction of low function Nxt1 protein. Both stocks have been crossed to get the mutant alleles over a CyoActinGFP (CAG) balancer that includes a mini-white marker and a fluorescent marker. Crossing the stocks together provides the Nxt1 trans-heterozygotes used for the experiments.

2.2.2.2 UAS Stocks

The pUAST-eGFP-Nxt1 M1 and M2A lines were generated in our lab for Nxt1 localization in the testis by PCR and subcloning of the Nxt1 ORF (Caporilli et al. 2013). w;UAS-$abba^{fl}$, kindly provided by Hanh T. Nguyen, uses the EST GH06739 (GenBank accession number AY121620) as template to amplify and clone the entire *abba* coding or defined regions of *abba* into the pUAST vector. The 3' end of the construct encodes a hemagglutinin tag. All UAS-stocks were in heterozygous condition.

2.2.2.3 RNAi Stocks

Two VDRC Nxt1 RNAi lines, P{KK107745}VIE-260B (Nxt1 RNAi 103146) and w^{1118};P{GD17336}v52631 (Nxt1 RNAi 52631), were used to analyse muscle integrity in the third instar larvae. 103146 has a 286bp hairpin length with predictions of 1 ON and 1 OFF targets, whereas 52631 has a 340bp length with 1 ON and 0 OFF targets predicted.

2.2.2.4 Deficiency Stocks

To test the gene dosage hypothesis in the larval muscles, deficiency lines for *abba* and *mib2* were used. Df(2R)Exel6068, $abba^{df}$, spans ~194kb deleting 16 genes. Df(2L)Exel8030, $mib2^{df}$, spans ~48kb deleting 9 genes. Each deficiency was crossed into the Nxt1 trans-heterozygotes background for muscle analysis.

2.2.2.5 Duplication Stocks

To test the gene dosage hypothesis in the larval muscles, a duplication line for *abba* was used. Dp(2;3)GV-CH321-38H19, spanning ~90kb, has been inserted into the PBac{y[+]-attP-3B}VK00031 docking site on the third chromosome (GenetiVision). This duplication was crossed into the Nxt1-trans-heterozygotes background for muscle analysis.

2.2.3 Recombination

Four stocks were recombined with the $Nxt1^{DG05102}$ allele to investigate the gene dosage hypothesis and for the *Nxt1* rescue. Virgin females yw;$Nxt1^{DG05102}$/CAG were crossed with males w; Df(2R)Exel6068/Cyo. The virgin females from the progeny with the genotype yw;$Nxt1^{DG05102}$/ Df(2R)Exel6068 were crossed with males w;if/Cyo. yw;$Nxt1^{DG05102}$// Df(2R)Exel6068/Cyo single males were picked out and tested whether recombination had occurred. At first, a yw;$Nxt1^{DG05102}$// Df(2R)Exel6068/Cyo male was crossed with yw; $Nxt1^{DG05102}$/CAG females. After four days, the yw;$Nxt1^{DG05102}$// Df(2R)Exel6068/Cyo male was crossed with w; Df(2R)Exel6068/Cyo females. If recombination occurred, both crosses would not show non-Cyo progeny, since both $Nxt1^{DG05102}$ and Df(2R)Exel6068 are homozygous lethal. This experiment was also performed for Df(2L)Exel8030, w;pUAST-eGFP-Nxt1-M1/Cyo;Sb/Tm3 and w;pUAST-eGFP-Nxt1-M2A /Cyo;Sb/Tm3 stocks.

2.2.4 Crosses

2.2.4.1 *Abba* Rescue Expression

yw;*Nxt1*^{DG05102}/CAG;UAS-*abba*^{fl} flies were crossed with either yw;*Nxt1*$^{z2\text{-}0488}$/CAG;mef2gal4 or yw;*Nxt1*$^{z2\text{-}0488}$/CAG;armgal4 to rescue *abba* expression in muscles specific or ubiquitously, respectively.

2.2.4.2 Nxt1 Rescue Expression

The Gal4-UAS system was used to look at the muscle integrity in the larva and pupa survivability. To rescue Nxt1 expression in the Nxt1 trans-heterozygote background w;pUAST-eGFP-Nxt1-M1,*Nxt1*^{DG05102}/CAG flies were crossed with w;*Nxt1*$^{z2\text{-}0488}$/CAG;mef2gal4 to get w;pUAST-eGFP-Nxt1-M1,*Nxt1*^{DG05102}/*Nxt1*$^{z2\text{-}0488}$; mef2gal4/+ for muscle specific rescue or crossed with w;*Nxt1*$^{z2\text{-}0488}$/CAG;armgal4 to get w;pUAST-eGFP-Nxt1-M1,*Nxt1*^{DG05102}/*Nxt1*$^{z2\text{-}0488}$; armgal4/+ for moderate ubiquitous rescue. Similar crosses were made with w;pUAST-eGFP-Nxt1-M2A, *Nxt1*^{DG05102}/CAG to get w;pUAST-eGFP-Nxt1-M2A,*Nxt1*^{DG05102}/*Nxt1*$^{z2\text{-}0488}$; mef2gal4/+ and w;pUAST-eGFP-Nxt1-M2A,*Nxt1*^{DG05102}/*Nxt1*$^{z2\text{-}0488}$; armgal4/+.

2.2.4.3 RNAi in Muscles

Both RNAi lines were used to investigate the phenotype in the larval muscles. UAS-Dicer was used to provide extra Dicer for the long hairpin lengths. Mef2gal4 driver was used to specifically shutdown the expression of Nxt1 in the muscles while not affecting other tissues. Genotypes and crosses used for RNAi can be found in **Appendix 2**.

2.2.4.4 Gene Dosage in Larval Muscles

To increase *mib2* expression in muscles, yw;*Nxt1*^{DG05102}/CAG;UAS-*mib2* flies were crossed with yw;*Nxt1*$^{z2\text{-}0488}$/CAG;mef2gal4 to get yw; *Nxt1*^{DG05102}/*Nxt1*$^{z2\text{-}0488}$;mef2gal4>UAS-*mib2*.

Two deficiency lines were used to reduce *abba* and *mib2* expression. Both deficiencies were recombined to get w;*Nxt1*^{DG05102}, Df(2R)Exel6068/CAG and w;*Nxt1*^{DG05102}, Df(2L)Exel8030/CAG. Both stocks were crossed with yw;*Nxt1*$^{z2\text{-}0488}$/CAG to get w;*Nxt1*^{DG05102},

Df(2R)Exel6068/*Nxt1*$^{z2\text{-}0488}$ and w;*Nxt1*^{DG05102}, Df(2R)Exel8030/*Nxt1*$^{z2\text{-}0488}$ to investigate the muscle integrity.

The Dp(2;3)GV-CH321-38H19 line was used to increase *abba* expression by adding an extra allele of *abba* on the 3rd chromosome. Specifically, yw;*Nxt1*^{DG05102}/CAG; Dp(2;3)GV-CH321-38H19 flies were crossed with yw;*Nxt1*$^{z2\text{-}0488}$/CAG to get yw;*Nxt1*^{DG05102}/*Nxt1*$^{z2\text{-}0488}$; Dp(2;3)GV-CH321-38H19.

2.2.5 Scanning Electron Microscopy

Young pupae (<4h) were picked, cleaned with water, and air dried overnight. The following day up to 6 pupae were put on an aluminium backing plate of a SC500 sputter target. A BIO-RAD SC500 sputter coater was used to coat the non-conducting pupae with ~20nm thick layer of Au90Pd10 according to manufacturer's protocol (Quorum Technologies). After coating, pins were put in a FEI-XL30 Field Emission Gun Environmental Scanning Electron Microscope. For imaging, a 30 μm diameter final aperture with a beam current <1nA was used to take pictures from each pupa at the posterior/anterior end and the middle section.

2.2.6 Magnetic Chamber Dissection and larva staining

For larval dissections, a magnetic chamber was used (Figure 2.1)(Budnik, Gorczyca, and Prokop 2006). A magnetic strip, with a 30 mm diameter hole in the middle, was glued on a 76 mm x 51 mm slide with 10 mm of the strip sticking out at all sides. Each magnetic chamber uses 2 center and 4 corner pins. For constructing all pins, see figure 1 in (Budnik, Gorczyca, and Prokop 2006). All pins were glued to a vintage metal index tab.

A larva is put in the middle of the hole with a drop of low Ca^{2+} Saline, HL-3 (Stewart et al. 1994). Larva are put ventral side and two center pins are used to prevent the larva from moving by pinning at the most anterior and posterior side. Vannas Spring Scissors – 3mm blades (RS-5618; FineScience) were used for cutting. A dorsal incision cut is made at the posterior end with short shallow cuts between the two trachea until the anterior end is reached.

The larva was cleaned on the magnetic chamber by removing all internal organs carefully. All Ca^{2+} Saline, HL-3 solution was removed with a P-100 pipette and fresh drops were added two more times while cleaning the larval carcass. The larval carcass was fixated by adding fresh 4% paraformaldehyde for 2 minutes. All the pins were removed from the larval carcass and the sample was transferred to a glass well with 100µl 4% paraformaldehyde for another hour. The fixative was removed and the larval carcass was washed twice with 100µl PBS-T (0.1% Triton X-100) for 5 minutes each. Two drops of Alexa Fluor™ 488 Phalloidin (ThermoFisher) was added to 1ml of PBS-T. 100µl Alexa Fluor™ 488 Phalloidin mix was added to the well to stain F-actin in larval muscles for 1 hour. Phalloidin solution was removed and the larval carcass was washed with PBS-T two times for 5 minutes each. The larva carcass was put on a microscope slide and mounted with Vectashield and covered with a cover slip. Images were made on an Olympus BX 50 (Olympus) microscope and Hamamatsu ORCA-05G digital camera. For a close-up of the larval muscle sarcomeres pattern, images were taken with a Leica DM6000B upright microscope with HC PL Fluotar 20x/0.50 and HCX PL APO 40x/1.25 oil objectives.

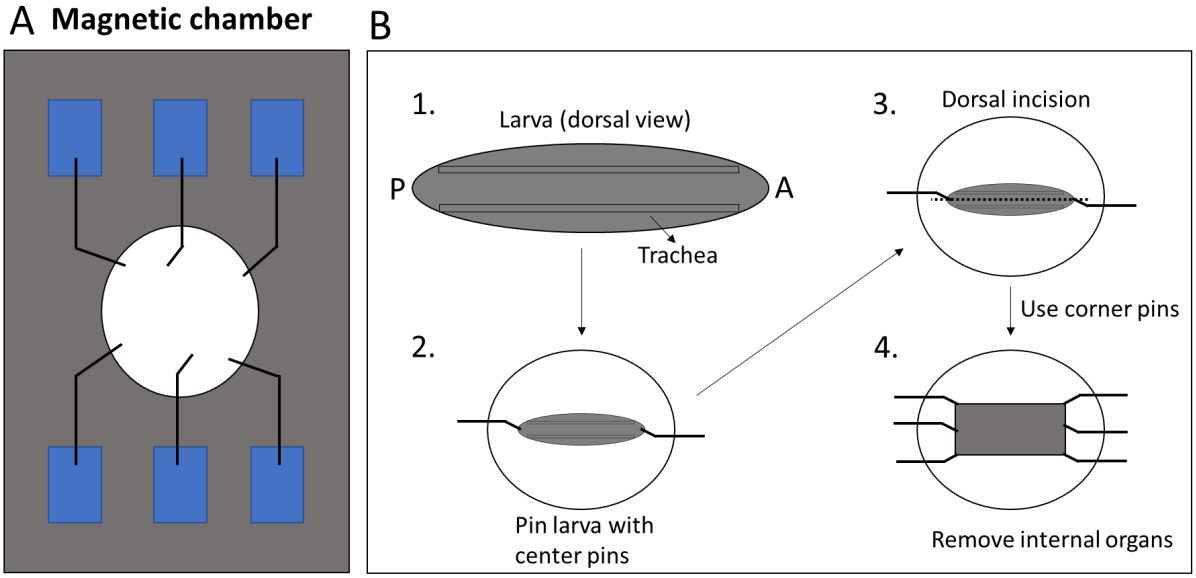


Figure 2.1: **Set-up for larval dissections.** A) View of the magnetic chamber. B) Dissection set-up for a larva. 1) Dorsal view of the larva, including the position of the two trachea. 2) Put larva in the circle with a drop of HL-3 saline and use center pins to pin both end of the larva. 3) Make a dorsal incision from posterior to anterior. 4) Use corner pins to open the larva and remove all internal organs. P = posterior. A = anterior.

2.2.7 Mobility Assay

First, second and third instar larvae were used for mobility analysis. Larvae (up to 10) were put in the middle on a 1% agar dish with 1% paraffin oil on one side and an odor (1% 2-propanol) on the other side. Larvae were filmed with a Samsung SDN-550 camera using the micro manager 1.4 program for 200 frames, with 1 frame per second (fps) in the dark under red light. Movies were analysed with MtrackJ (Meijering, Dzyubachyk, and Smal 2012) via Fiji. Larvae were tracked per 20 frames (1st instar) or 10 frames (2nd and 3rd instars) for at least 5 points while larvae were continuously crawling on the plate. Mean speed is calculated in mm/sec.

2.2.8 Viability Assay

Third instar larvae were taken from the vial and put in a new vial with low quantity of food for the viability assay. After 5 days of incubation at 25^{o}C vials were taken out and pupa lethality was backtracked by looking at 6 different points: 1) larvae that did not pupate (0h APF), 2) no visible development in pupae (24h APF), 3) developing of the eye (48h APF), 4) bristles on back and shoulders (72h APF), 5) complete fly development in pupa (72h APF) and 6) emerging adults.

2.2.9 Axial Ratios

Young pupae (less than 12 hours old) were taken from the vial, cleaned in dH_2O and measured in length and width with a paper ruler. Posterior and anterior spiracles were excluded for measuring the length. The width was measured at around half the length of the pupa. The axial ratio was calculated by dividing the length by the width.

2.2.10 Muscle Integrity

The integrity of larva muscles was calculated as a percentage from a total of 8 hemisegments. The hemisegments A2-A5 were in the abdominal area and were not damaged by any of the pins. Each hemisegment contained 30 different muscles, so a total of 240 muscles were inspected. The integrity of a muscle is compromised if the muscle is damaged in any way, such as being torn, thin, loss of sarcomeric structure, missing muscles, etc. The muscle percentage is the total number of damaged muscles divided by 240 and multiplied by 100.

2.2.11 Fertility

Flies (for genotypes see **appendix 2**) were tested for fertility by crossing three mutant males with six female WT flies and vice versa. Crosses were put at 25°C for several days and checked each day for egg hatches. If no eggs were hatched after 5 days, the flies with the mutant genotype were considered sterile.

2.3 Microarray

Five pairs of 7 days old ovaries from *Nxt1* ^{DG05102}/ *Nxt1*$^{z2\text{-}0488}$ trans-heterozygotes and w^{1118} females, both raised at 25°C were used for Affymetrix microarray analysis Microarray was performed with three replicates per genotype on an affymetrix array v2 by the Glasgow polyomics facility. Sample preparation, processing and data analysis were as described previously (Doggett et al. 2011) and performed by Helen White-Cooper. Processed data of normalized intensity for each probe was provided in an Excel file and used for this project.

2.4 Sequencing

2.4.1 RNA Sequencing (Whole Larvae)

For whole larvae sequencing, a total of 18 samples were required. w^{1118} and Nxt1 trans-heterozygotes were the genotypes used. Each genotype had samples of three different stages (wandering larvae, stationary larvae and white prepupae; see section 2.1.3.3 for selection) in triplicates. RNA extraction was performed using a combination of Trizol (ThermoFisher Scientific) and the RNeasy Mini Kit (Qiagen).

RNA samples were checked for quality on an Agilent Bioanalyzer 2300 (Agilent Technologies) using the Agilent RNA 600 Pico Kit. Each sample had an RNA concentration between 200 – 5000pg/µl, quantified via the Nanodrop ND-1000 (ThermoFisher Scientific). The Bioanalyzer was used according to manufacturer's protocol. In short, a gel-dye mix was prepared and 9µl was added in three designated wells. 9µl RNA conditioning solution was added in another designated well. In the remaining wells, 5µl RNA marker was added. For each RNA sample, 1µl was added in the wells. 1µl heat denatured ladder was added in the final designated well.

The chip was run in the Agilent 2300 Bioanalyzer and the provided software was used for analysis of RNA quality.

The samples were send to the University of Exeter to perform the library preparation (ScriptSeq RNA-Seq Library Preparation Kit (Illumina)) and sequencing. All samples were 100bp paired-end sequenced on an Illumina HiSeq 2500 using standard mode.

2.4.2 RNA Sequencing (Larval Carcasses)

2.4.2.1 Qubit

RNA samples were quantified on a Qubit® 2.0 Fluorometer using either the Qubit® RNA BR Assay Kit or Qubit® RNA HS Assay Kit (ThermoFisher Scientific) using manufacturer's protocol. In short, Qubit® Working Solution is prepared for each standard and sample by diluting the Qubit® RNA BR (or HS) Reagent 1:200 in Qubit® RNA BR (or HS) Buffer. Two assay tubes were set up each with 190µl Qubit® Working Solution and 10µl standard. Each sample was set up with 199µl Qubit® Working Solution and 1µl sample. The RNA Broad Range (or High Specificity) Assay option was selected and calibration was set up with standards followed by reading each sample.

The dsDNA libraries (see section 2.4.2.3 for library preparation) were quantified using the Qubit® dsDNA BR Assay Kit and the Qubit® dsDNA HS Assay Kit (ThermoFisher Scientific) using manufacturer's protocol. The set-up is similar as described previously for the RNA samples.

2.4.2.2 Bioanalyzer and TapeStation

Before library preparation, RNA samples were checked for quality on an Agilent Bioanalyzer 2300 (Agilent Technologies) using the Agilent RNA 6000 Pico Kit. Each sample had an RNA concentration between 200 – 5000pg/µl, quantified via the Nanodrop ND-1000. The Bioanalyzer was used according to manufacturer's protocol.

After libraries were prepared, libraries were checked for quality using the 2200 TapeStation (Agilent Technologies). The TapeStation worked like the Bioanalyzer but uses a gel strip

instead of a chip. The Broad Range D1000 and the High Sensitivity D1000 ScreenTape Assay (Agilent Technologies) were used according to manufacturer's protocol. The libraries were deemed successful if it showed the expected amplicon of 180-200bp.

2.4.2.3 mRNA Libraries Preparation

Six libraries were prepared from total RNA for RNA sequencing. Both w^{1118} and Nxt1 trans-heterozygotes were performed in triplicates. Below is a general explanation of the important steps used to generate libraries using the TruSeq Stranded mRNA Library Prep (Illumina).

Purify and Fragment mRNA

Total RNA of each sample was diluted to 10ng/µl in dH_2O to a final volume of 50µl in a 0.3ml tube. 50µl of RNA Purification Beads (Agilent) were added to the samples to bind the polyA RNA to the oligo dT magnetic beads for 5 minutes. The RNA was denatured at 65°C for 5 minutes. To wash the beads, tubes were put in a magnetic stand. After 3 minutes, supernatant was removed, and Bead Washing Buffer was added to wash the beads by pipetting up and down several times. Supernatant was removed again and 50µl of Elution Buffer is added. The mRNA was eluted by incubating at 80°C for 2 minutes. 50µl of Bead Binding Buffer was added to allow mRNA to rebind to the beads. The beads were washed as explained previously and 19.5µl of Fragment, Prime, Finish Mix was added to the beads. The RNA was eluted, fragmented and primed by incubating at 94°C for 8 minutes.

First Strand cDNA Synthesis

The tubes were placed on a magnetic stand and 17µl of the supernatant was transferred to a new 0.3ml tube. 8µl of First Strand Synthesis Act D Mix and SuperScript II mix were added to each tube. The cDNA synthesis was performed on a PCR machine using the following program:

1. Pre-heat lid to 100°C
2. 25°C for 10 minutes
3. 42°C for 15 minutes
4. 70°C for 15 minutes
5. Hold at 4°C

Second Strand cDNA Synthesis

5µl Resuspension Buffer and 20µl Second Strand Marking Master Mix were added to each tube. Samples were incubated at 16°C for 1 hour. 90µl of AMPure XP beads were added to separate the ds cDNA from the second strand reaction mix. The beads were washed, using the magnetic stand, by removing all supernatant and adding 200µl 80% EtOH. Washes were repeated for two more times. All supernatant was removed and 17.5µl Resuspension Buffer was added. The beads were mixed and put back on a magnetic stand to transfer 15µl supernatant (ds cDNA) to a new 0.3ml tube.

Adenylate 3' Ends

Before ligation, a single 'A' nucleotide was added to the 3' ends of the blunt fragments to prevent them from ligating to one another and allowing the adapters to ligate. 2.5µl Resuspension Buffer and 12.5µl A-Tailing Mix were added to each tube. The adenylation was performed on a PCR machine using the following program:

1. Pre-heat lid to 100°C
2. 37°C for 30 minutes
3. 70°C for 5 minutes
4. Hold at 4°C

Ligate Adapters

Add 2.5µl Resuspension Buffer and 2.5µl Ligation Mix to each tube. To each tube, 2.5µl of RNA Adapter Index was added. Each sample used a different adapter index explained in 2.4.2.4. The samples were incubated at 30°C for 10 minutes. Next, 5µl Stop Ligation Buffer and 42µl AMPure XP Beads were added and tubes were incubated for 15 minutes, before it was put on a magnetic stand. The beads were washed with 80% EtOH as explained previously for three times. 52.5µl Resuspension Buffer was added to mix the AMPure XP beads, followed by separating it on a magnetic stand, and 50µl of supernatant was transferred to a new 0.3ml tube. 50µl of AMPure XP beads was added for a second cleanup. After the second cleanup, 22.5µl Resuspension Buffer was added, separated on a magnetic stand, and 20µl of supernatant was transferred to a new 0.3ml tube.

PCR Amplification

To each tube, 5µl PCR Primer Cocktail and 25µl PCR Master Mix was added. The following program on the PCR machine was used:

1. Pre-heat lid to 100°C
2. 98°C for 30 seconds
3. 15 cycles of:
    a. 98°C for 10 seconds
    b. 60°C for 30 seconds
    c. 72°C for 30 seconds
4. 72°C for 5 minutes
5. Hold at 4°C

50µl of AMPure XP beads were added, incubated at RT for 15 minutes, and put on a magnetic stand. The beads were washed as explained previously with 80% EtOH for three times. 32.5µl Resuspension Buffer was added, separated on a magnetic stand, and 30µl of supernatant was transferred to a new 0.3ml tube.

Validating Libraries

The libraries were validated using the 2200 TapeStation (Agilent technologies) with the Broad Range D1000 ScreenTape Assay. Most samples had a smear at around 300bp.

Normalization and Pooling Libraries

The six libraries were normalized to 1.5nM each by using Qubit to determine the concentration of the libraries in ng/µl and converting to nM by multiplying bp (from the 2200 TapeStation) with concentration in ng. The pooled libraries were quantified on the Qubit before loading into the sequencer.

2.4.2.4 Sequencing

Six samples were used to prepare libraries. To pool libraries together, different adapter indexes were used. Adapter index 2, 4, 5, 6, 7, and 12 were used for w^{1118} rep 1-3 and Nxt1 trans-heterozygotes rep 1-3, respectively. The adapter index sequences can be found in **Appendix 3**. The samples were sequenced using 2x75bp paired end with the NextSeq 500/550

Mid Output v2 kit (150 cycles; Cat. No. FC-404-2001) on an Illumina NextSeq500 Sequencer (Illumina).

2.4.3 Small RNA Sequencing

First round of small RNA sequencing

Forty micrograms of RNA from 4, 14 and 21 days old ovaries for both Nxt1 trans-heterozygotes and w^{1118} were used. Small RNA extraction was performed as described in section 2.1.3.1. All samples were sent to the Centre for Genomic Research, University of Liverpool for library preparation (NEBNext® Small RNA Library Prep Set for Illumina) and 50 bp single-end sequencing on an Illumina HiSeq 2500.

Second round of small RNA sequencing

Forty micrograms of RNA from 7 days old ovaries for both Nxt1 trans-heterozygotes and w^{1118} wild type were used. Small RNA library preparation and sequencing was performed in collaboration with Julius Brennecke group at the Institute for Molecular Biotechnology in Vienna, Austria. A brief overview of the library preparation is explained below.

2.4.3.1 small RNA Libraries Preparation (non-oxidized)

For small RNA sequencing, w^{1118} and Nxt1 trans-heterozygotes genotypes were used. Below is a general overview of generating small RNA libraries for small RNA sequencing using a combination of the TruSeq Small RNA Library Kit (Illumina) for library preparation and Kapa Biosystems for library amplification. The mixes for ambion decade marker, spikes, 3' linker ligation, 5' linker ligation, KAPA PCR and PmeI digestion can be found in **Appendix 4**.

RNA extraction and 2s rRNA depletion

RNA was extracted using Trizol (ThermoFisher Scientific) as explained previously in 2.1.3.2. DynaBeads-Streptavidin (Thermofisher Scientific) were used for 2S rRNA depletion. 100µl beads were used per 10µg total RNA. For both genotypes, 30µg total RNA was used.

Beads were washed 2x with 0.5x SSC on a magnetic stand. 600µl 0.5x SSC + 6µl in-house biotinylated 2S rRNA were added and put on ice for 30 minutes. Beads were washed once more with 0.5x SSC on a magnetic stand and resuspended in 600µl 0.5x SSC. Beads were incubated at 65°C and RNA samples at 80°C for 5 minutes. 300µl beads were added to each sample (containing 30µg total RNA). Samples were incubated at 50°C for 1 hour and RNA was transferred into a new tube by using a magnetic stand. The samples were precipitated O/N with 40µl 3M NaAc pH 5.2, 1µl glycogen and 2.5x EtOH 100%.

Purification of Small RNAs

The spikes mix was used to spike all samples with ^{32}P-labeled 19bp and 28bp oligos (~200 counts per second). Total RNA was heated at 95°C for minutes, then chilled more than 5 minutes on ice. A 1mm 15% polyacrylamide/urea gel (see section 2.1.2) was prepared and samples were loaded plus 10ng of ^{32}P-labeled decade marker mix. The gel was run at constant 35W for ~1 hour, followed by 30 minutes exposure with a phosphor imaging screen (Molecular Dynamics). Small RNAs were excised ranging between ~19-28nt and put in 1.5 ml tubes. RNA was eluted and precipitated by adding 500µl 0.4M NaCl with O/N agitation at RT. Gel slices were spun and eluated through micropore filter (Ultrafree-MC, 0.22 mm, Millipore) for 5 minutes at 12.000 x g at RT. RNA was precipitated by adding 1µl glycogen and 2.5 volume of 100% EtOH. Samples were incubated at -20°C for 3 hours, spun at 4°C for 30 minutes, washed with 75% EtOH, and spun again at 4°C for 30 minutes. All EtOH was removed and air dried for 5-10 minutes. Pellet was resuspended in 10.5µl nuclease free H_2O and 0.5µl 3' adapter (100µM).

3' Adapter Ligation

RNA+3' adapter mix was denatured for 5 minutes at 65°C and cooled on ice immediately afterwards. 3' Linker ligation mix was added and incubated at 16°C O/N. 20µl 2x RNA UREA loading buffer (ThermoFisher Scientific) was added and heat inactivated at 95°C for minutes and chilled on ice afterwards.

A 0.75mm 12% polyacrylamide/urea gel was prepared (see section 2.1.2). Samples and 10ng of ^{32}P-labeled decade marker were loaded and gel was run at constant 30W for ~1 hour. Gel was exposed with a phosphor imaging screen (Molecular Dynamics) for ~1 hour and bands

were excised between ~45-54nt and put into 1.5ml tubes. RNA was eluted and precipitated by adding 500µl 0.4M NaCl with O/N agitation at RT. Pellet was resuspended in 8µl nuclease free H_2O and 0.5µl 5' adaptor (100µM).

5' Adapter Ligation

RNA+5' adapter was denatured for 5 minutes at 70°C and cooled on ice immediately afterwards. 5' Linker ligation mix was added and incubated at 16°C O/N. 20µl 2x UREA loading buffer (ThermoFisher Scientific) was added and heat inactivated at 95°C for 5 minutes.

A 0.75mm 10% polyacrylamide/urea gel was prepared (see section 2.1.2). Samples were loaded and 10ng of ^{32}P-labeled decade marker and gel was run at constant 30W for ~1 hour. Gel was exposed for ~1-2 hours with a phosphor imaging screen (Molecular Dynamics) and band was excised between ~82-91nt and put into 1.5 ml tubes. RNA was eluted and precipitated by adding 500µl 0.4M NaCl with O/N agitation at RT. Pellet was resuspended in 9µl nuclease free H_2O and 2.5µl Solexa RT rev (10µM).

Reverse Transcription

11.5µl RNA-lib and primer was incubated and mixed at 65°C for 5 minutes and cooled on ice for at least 2 minutes. The following was added to each sample:

1.0µl dNTPs (10mM)
4.0µl 5x FS buffer
2.0µl 0.1M DTT
0.5µl RiboLock
1.0µl SuperScript II

The mix was incubated at 42°C for 1 hour, heated to 70°C for 15 minutes and stored at -20°C until further use.

PCR Amplification with KAPA

KAPA PCR mix (Kapa Biosystems) was set up and the following PCR program was used:

1. Pre-heat lid to 100°C
2. 98°C for 15 seconds

3. 30 cycles of:
    a. 98°C for 15 seconds
    b. 65°C for 30 seconds
    c. 72°C for 25 seconds
    d. Picture (for fluorescent signal)
    e. 72°C for 10 seconds

PCR was stopped for each individual sample when the fluorescent signal reached 3000-4000 units.

PCR samples were purified on columns with the Wizard SV kit (Promega) according to manufacturer's protocol.

PmeI Digestion of Radiolabeled Spike-In Oligos

The radiolabeled spike-in oligos were digested with PmeI using the PmeI digestion mix. All samples were incubated at 37°C for 3 hours.

Gel Purification of Amplified cDNA

A 2% low-melt agarose gel (ThermoFisher Scientific) was prepared and all samples were loaded with 6X Orange Loading Dye (KAPA box) and 30µl 50bp DNA ladder (0.1µg/µl). Gel was run at constant 100V until orange band was at 2/3 of the gel length. The DNA band was excised using a Dark Reader Transilluminator (Clare Chemical Research) using UV light and transferred into a 1.5ml tube. The Zymoclean Gel DNA Recovery Kit (Thermofisher Scientific) was used and each sample was eluted in 20µl nuclease free H_2O.

2.4.3.2 small RNA Libraries Preparation (oxidized)

Oxidized samples followed the same protocol for non-oxidized samples. Before purification of small RNAs, sodium perodiate was used to oxidize the RNA. 15µl RNA was used and mixed with 12µl dH_2O, 8µl 0.5M 5x Borate buffer (ThermoFisher Scientific), 5µl 50mM sodium perodiate and incubated for 30 minutes at 25°C. 229µl dH_2O, 30µl 3M NaAc (pH 5.2), 1µl glycogen and 900µl 100% EtOH was added for O/N precipitation.

2.4.3.3 Immunoprecipitation sequencing

The immunoprecipitation (IP) for w^{1118} and $Nxt1^{DG05102}/Nxt1^{z2\text{-}0488}$ with Aub, Piwi and AGO3 antibodies were done by Rippei Hayashi (Brennecke lab) according to their in-house protocol (Brennecke et al. 2007).

2.4.3.4 Sequencing

To pool libraries together, different adapter indexes were used. Adapter index 1, 2, 3 and 4 were used for w^{1118} and Nxt1 trans-heterozygotes for non-oxidized and oxidized, respectively. Adapter sequences can be found in **Appendix 3**. The libraries prepared from small RNA were sent to the sequencing facility for 50bp single-end sequencing on an Illumina HiSeq 2500 at the Institute for Molecular Biotechnology (IMBA), Vienna.

2.5 Image and Data Analysis

2.5.1 Image processing

Images from the confocal and fluorescence microscope were processed using FIJI. The image contrast was adjusted, if muscles in image were too bright, using Image; Adjust; Window/Level. Scale bars were added using Analyse; Tools; Scale Bar.

Whole third instar larva pictures were a combination of six images. The images were overlapped carefully in Microsoft PowerPoint and grouped together. The edges were filled with black bars to get the image in a rectangular shape. The scale bar was added in on one image via FIJI as explained previously.

2.5.2 Microarray Analysis

Microarray data used in this project was analysed by Helen White-Cooper using RMA normalization in R using the affy package. Normalized data was provided in an Excel file and was used for further analyses for this project.

2.5.3 Whole larvae Sequencing Analysis

The data analysis was done via Galaxy and GenePattern (Reich et al. 2006). The reads were mapped to the reference genome dm3 (UCSC) with Bowtie v1.1.1 (Langmead et al. 2009). The tuxedo suite v2.0.2 (Trapnell et al. 2012) was used to process the mapped reads through Cufflinks, Cuffmerge and Cuffdiff. The data from Cuffdiff was imported to excel and contained the significant changes in transcript expression, splicing, and promoter use. All programs used their default settings for analysis.

2.5.4 Larval carcasses Sequencing Analysis

The data analysis was done with self-made scripts using bash and awk (Aho et al 1978; Bourne 1978). Reads were mapped to reference genome r6.13 (FlyBase) with BDGP6.87 as GTF file using TopHat v2.1.1 (Trapnell, Pachter, and Salzberg 2009) with Bowtie v1.1.1 (Langmead et al. 2009). The mapped reads were then processed with the tuxedo suite v2.2.1 (Trapnell et al. 2012) through Cufflinks, Cuffmerge, CuffQuant and Cuffdiff. The output of Cuffdiff was imported to excel for further analysis. All programs used their default settings for analysis.

2.5.5 Ovaries Sequencing Analysis

First round of small RNA sequencing

PiPipes (Han, Wang, Zamore, et al. 2015) was used to analyse the data. The small RNA-seq pipeline reports the abundance, length distribution, nucleotide composition and 5'-to-5' distance of piRNAs assigned to genomic annotations, including individual piRNA clusters and transposon families.

The general workflow for analysing small RNA data via piPipes is as follows. All reads aligned to ribosomal RNA sequences were removed. The remaining reads were mapped to miRNA hairpin sequences for 5'- and 3'-end heterogeneities of mature miRNAs. The remaining reads were mapped to the reference genome dm3 (UCSC) via Bowtie 1.2.0 (Langmead et al. 2009). Reads were then separated to different genomic features by their coordinates. piPipes drew graphs of length distribution, nucleotide composition and 'Ping-Pong' amplification for each

genomic feature. The piRNA read coverage was normalized per 1 million mapped miRNA reads.

Second round of small RNA sequencing

For the second round of small RNA-seq data, self-made scripts using bash and awk were used to analyse the data (Aho et al 1978; Bourne 1978). The fastQ files were processed through the annotation pipeline provided by Julius Brennecke group. These data provided information about the quality, read coverage and normalization depth. The collapsed annotation fasta file was used for further analysis with self-made scripts and processed the following:

The fastx-toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/) was used to convert the files to Fasta, remove artifacts, clip adapters and trim sequence lengths.

Bowtie v1.1.1 (Langmead et al. 2009) was used to map reads to different genomic features such as rRNA, miRNA, transposons, drosophila genome r6.13 (FlyBase), tRNAs, exons, introns, 5' UTR and 3' UTR. Samtools (Li, Handsaker, et al. 2009) and Bedtools v2.26.0 (http://quinlanlab.org) were used to convert the mapped data into a .bed file.

Reads mapping to transposons were counted by manipulating the .bed file from earlier analyses. In short, reads were grouped by length, strand type, transposon and count. Length distributions from different genomic features were produced by manipulating the .bed file with awk and bash. Both read mappings and length distributions were imported to Excel to calculate the normalized reads.

The .bed file from reads mapping uniquely to the *Drosophila* genome was used to produce BigWig files to load tracks on the UCSC genome browser (www.genome.ucsc.edu). These tracks contained the unique reads mapping across the genome.

Oxidized small RNA sequencing

The same script used for non-oxidized samples were used for the oxidized samples. The length distribution was calculated with reads mapping across all known transposons with a length of 23-29 nt. All data was imported into Excel for read normalization.

IP-sequencing

The same script was also used to analyse IP-sequencing data. Length distribution was calculated with reads mapping across all known transposons with a length of 23-29nt. All data was imported into Excel for read normalization.

Reads Normalization

Different normalization methods were used for the different small RNA sequencing data. Non-oxidized data was normalized to 1 million mapped miRNA reads, whereas oxidized and IP-seq data were normalized to 10 million uniquely mapped reads due to the absence of miRNAs.

2.5.6 FlyMine as a data analysis tool

Lists of differentially expressed genes from sequencing were uploaded on FlyMine (Lyne et al. 2007) to look at the gene expression in *Drosophila*, gene ontology and pathway enrichment. A selection of genes based on the data was further analysed by qRT-PCR for double confirmation.

FlyMine was also used for the genes with introns analysis. A python script, kindly provided by Rachel Lyne, was used to filter out genes with introns and provide information based on total intron length, number of introns, smallest intron and largest intron. Total intron length is the sum of all introns of all isoforms for each gene. The number of introns is the sum of all unique introns of all isoforms for each gene. The script was adjusted by Rachel Lyne to also provide the gene length (sum of all isoforms, including introns), number of transcripts, shortest transcript and longest transcript for each provided list. Microsoft Excel was used to calculate the median values from all the lists.

2.5.7 Statistical Analysis

The error bars from the mobility assay were the standard error of the mean. The statistical test performed for this assay was the student's t-test. The graphs were generated either via Microsoft Excel or R. Statistical tests were performed in Microsoft Excel. Microsoft PowerPoint was used to add in the asterisks.

The median of total intron length and gene length of the down-regulated genes and up-regulated genes in the 1.5-, 2-, 4-, and 16-fold categories were compared with the non-differentially expressed genes with the Mann-Whitney test. P-value lower than 0.05 indicates statistically significant.

Chapter 3.

Loss of Nxt1 causes muscle degeneration defects in third instar larva

Chapter 3 – Results

During this project, I focus on understanding the importance of Nxt1 during the metamorphosis. To do this, I will use Nxt1 trans-heterozygotes, which still have some functional Nxt1. I will use whole larvae RNA-seq from wandering larvae, stationary larvae and white prepupae to analyse the differential gene expression and the potential involvement of ecdysone in the air bubble phenotype. If ecdysone is not involved, then I will focus on the larval muscles by dissecting WT and Nxt1 trans-heterozygotes larvae to understand whether the muscle integrity is compromised.

3.1 Nxt1 as an RNA export factor and novel meiotic arrest gene

Nxt1 is mostly known as an RNA export factor that binds to Nxf1, which binds to mRNA for nuclear export transportation. The line *Nxt1*$^{z2\text{-}0488}$ was identified in an EMS-screen as a male sterile mutant using phase-contrast microscopy of live testis squash preparations (Koundakjian et al. 2004). Homozygote *Nxt1*$^{z2\text{-}0488}$ showed a typical meiotic arrest where mutant testes had stages up to and including mature primary spermatocytes (Caporilli et al. 2013). Positional cloning showed that *z2-0488* has an amino acid substitution (D126N) in Nxt1 (Figure 3.1). This residue is conserved in metazoans (D131 in human Nxt1). The D131 sidechain, with its peptide backbone on a surface beta-sheet, is towards the core of the protein and participates in a hydrogen-bonding network with Y24, Y39 and Q111 residues. Substitution of D131N disrupts the H-bonding network, which decreases the protein stability. A different allele, *Nxt1*^{DG05102}, has a P-element insertion in the coding region and is expected to be a null. *Nxt1*^{DG05102} homozygotes are embryonic lethal. *Nxt1*$^{Z2\text{-}0488}$ homozygotes and *Nxt1*$^{Z2\text{-}0488}$/*Nxt1*^{DG50102} trans-heterozygotes were semi-lethal and male and female sterile (Caporilli et al. 2013).

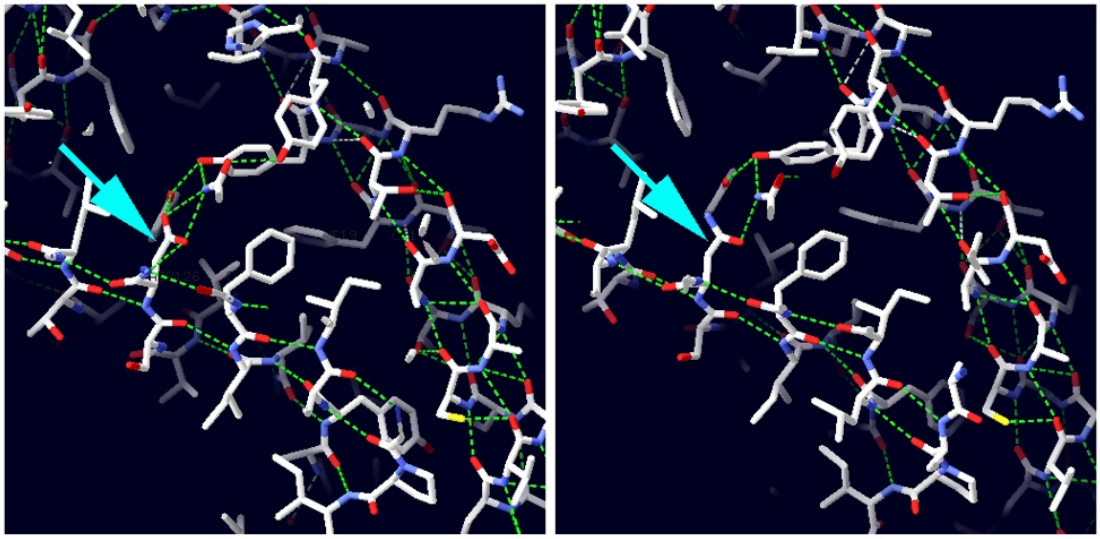


Figure 3.1: **D126N mutation revealed in Nxt1 with *z2-0488* allele.** *Z2-0488* allele has a mutation that disrupts the H-bonding network (arrow) between the Y24, Y39, Q111 and D131 (D126 in *Drosophila*) residues in human. Figure taken from (Caporilli et al. 2013).

3.2 Nxt1 pupae have a distinctive curved shape, uneverted spiracles and fail head eversion

Previously, our research group worked on Nxt1 in the testis in the *Drosophila* and found that there was a link between the RNA export factor and the tissue specific transcriptional regulation (Caporilli et al. 2013). Several phenotypes were observed throughout the life cycle of the mutants, including an air bubble defect, a curved shape and uneverted spiracles during the pupa phase (Figure 3.2A). Before I investigated the air bubble defect, the pupa itself was inspected in more detail. I measured the axial ratios from the pupa by measuring the length and width, while excluding the posterior and anterior spiracles. I noticed that the ratios between w^{1118} and Nxt1 trans-heterozygotes were different (Figure 3.2B). This revealed that the mutant pupae were slightly taller and thinner than wild type. Scanning electron microscopy (SEM) revealed that the structure of the pupa case was not different by looking at the carcass and posterior/anterior ends. (Figure 3.2C-E). However, Nxt1 trans-heterozygotes can have uneverted anterior spiracles for one or both of their spiracles (Figure 3.2E'').

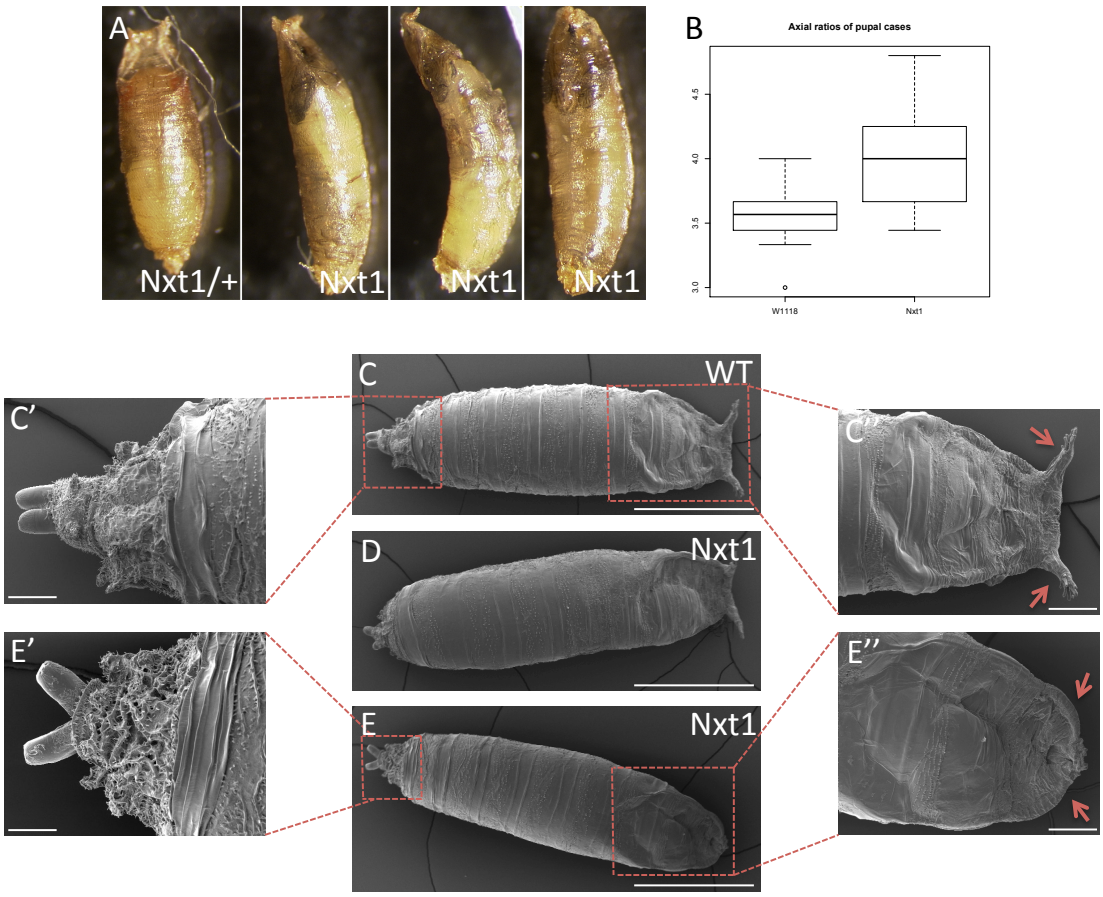


Figure 3.2: **Nxt1 trans-heterozygote have a distinctive curved pupa shape, uneverted spiracles, but have no differences in overall pupa structure.** A) Nxt1 mutant pupae are thinner, curved, and have uneverted spiracles. B) Axial ratios of Nxt1 mutants are different due to being thinner and larger. C-E) Scanning electron microscopy images of wild type and Nxt1 mutant pupae skin and close-up. Figure in A was taken from (Caporilli et al. 2013). SEM pictures were taken on a FEI-XL30 Field Emission Gun Environmental SEM.

Failure of proper air bubble migration has an effect on head eversion (Supplementary Figure 1) and therefore the viability of the pupa. I took a w^{1118} and Nxt1 trans-heterozygote pre-pupa and filmed for 15 hours to observe the air bubble (Figure 3.3). The air bubble is faintly visible from 0h APF (Figure 3.3A white arrows), but stalls in the mutant after migrating around 8h APF (Figure 3.3I white arrows). The air bubble continues to stay halfway up to 11h APF (Figure 3.3J-M white arrows) before disappearing and leaving an empty space at the posterior of the pupa body (Figure 3.3N-P). In wild type controls head eversion occurs at around 12 hours into the pupa phase (Figure 3.3M blue arrow). No head eversion was seen in this mutant pupa

(Figure 3.3M). Before the head eversion occurs, the posterior end of the larva body retracts to the anterior by “wiggling” around (Supplementary Figure 2A-E). Immediately after head eversion the whole larval body “wiggles” back to reach the posterior end (Figure 3.3M; Supplementary Figure 2F-H). In Nxt1 mutant, the air bubble stays halfway down the pupa at the 12-hour mark (Figure 3.3M). One hour later, the air bubble has disappeared, no head eversion occurred, and the posterior larval body has retracted to the anterior (Figure 3.3N-P). No noticeable differences were observed at later time points (data not shown).

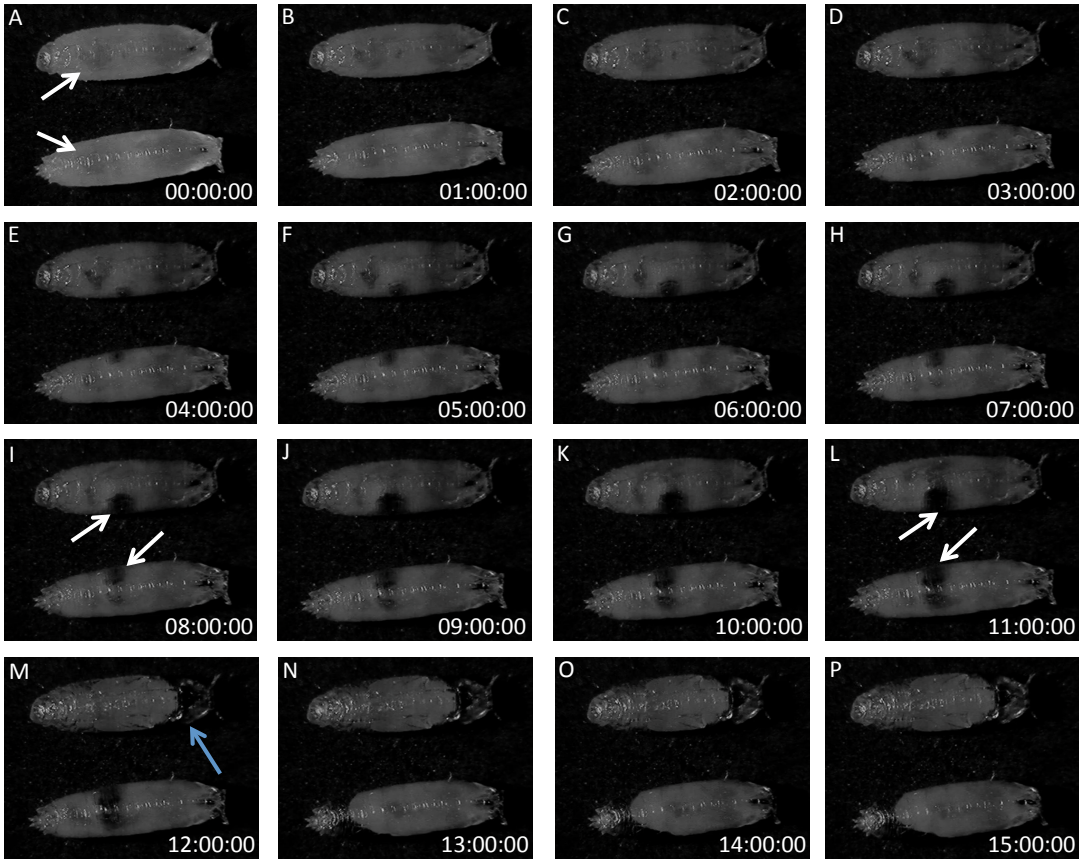
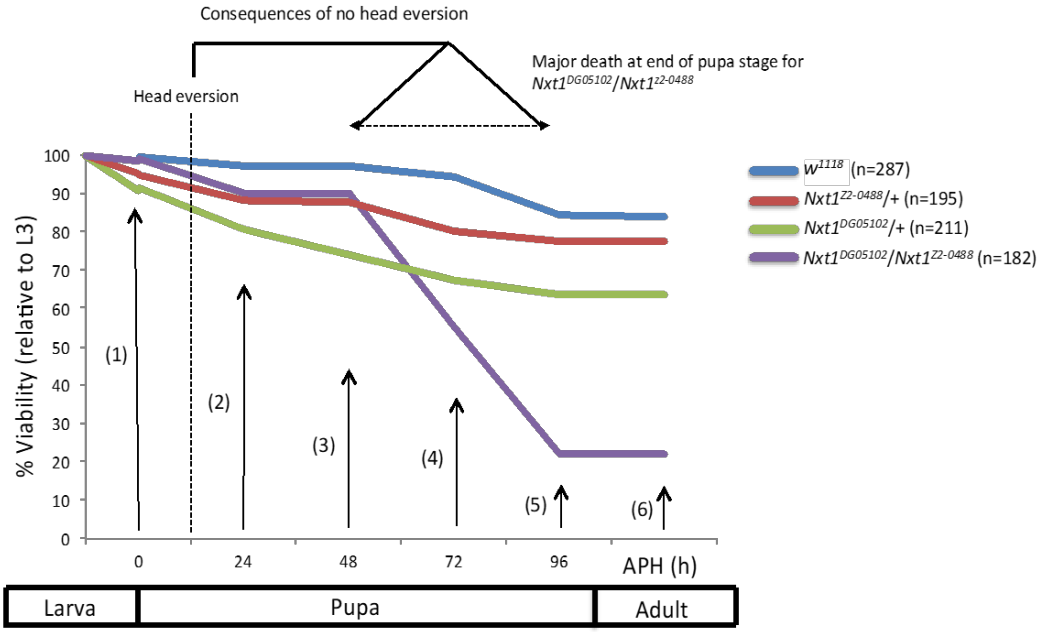


Figure 3.3: **Air bubble not fully migrating to the anterior end in Nxt1 trans-heterozygotes.** A) Faint visibility of air bubble in both genotypes (white arrows). B-D) Air bubble becomes more prominent. E-G) Air bubble is clearly visible and reaches halfway through the pupa. I-L) Air bubble expands further (white arrows). M) Air bubble disappeared and caused head eversion for wild type (blue arrow), whereas Nxt1 mutant still show air bubble. N) Larva body in Nxt1 mutant "wiggles" at posterior end and retracts to anterior. O-P) No further changes observed in both genotypes. Pictures were taken with a Samsung SDN-550 camera.

Both Nxt1 alleles were balanced over *Cyo Actin GFP*, which helped during the larva stage for genotyping. Larva without GFP, under a fluorescent microscope, were Nxt1 trans-heterozygotes. Crosses of $Nxt1^{DG05102}/CAG$ to $Nxt1^{Z2\text{-}0488}/CAG$ did not reveal the expected Mendelian ratios of 1:1:1 genotypes (*CAG* homozygotes are lethal). Instead, the observed ratio was 1:5:5 or roughly 1 in 11 progenies were Nxt1 trans-heterozygotes. I set up an assay to look at the viability of the pupa at four different stages (24h, 48h, 72h and 96h pupa), not including the larva check and adults emerging check. Only about 20% (N=182) of the Nxt1 trans-heterozygote pupae

survived through to adulthood, compared to 90% (N= 287) viability for control w^{1118} pupae (Figure 3.4). The $Nxt1^{DG05102}$/+ showed a slight reduction to pupa viability (60% (N= 211); Figure 3.4).



(1): Larvae that did not pupate
(2): No visible development in pupa
(3): Eye starts developing
(4): Bristles on back and shoulders
(5): Complete fly development in pupa
(6): Adults that emerged

Figure 3.4: **Semi-lethality observation in Nx1 trans-heterozygotes.** Pupae were examined at six different time points. Nxt1 mutants have no proper head eversion, which becomes visible from 48 hours into the pupa phase. Normal development in these mutants still occur and lethality occurs somewhere between 48-96h into the pupa phase. $Nxt1^{DG05102}$/+ genotype showed a slight reduction in pupa viability.

3.3 Ecdysone is not involved in the Nxt1 trans-heterozygotes pupa phenotype

The specific air bubble phenotype suggests that there are defects in the function of the larval/pupal abdominal muscles. Air bubble migration occurs in response to the first pre-pupal pulse of ecdysone. Ecdysone coordinates tissue-specific morphogenetic changes by several pulses throughout the *Drosophila* development. A high concentration pulse is observed immediately preceding the larva-pupa transition.

To understand if genes responding to ecdysone are differentially expressed in our mutant, we performed RNA sequencing. Three different stages were selected: before, during, and after the pulse of ecdysone. These stages are referred to as wandering larvae, stationary larvae and white prepupae, respectively. Each stage was performed in triplicates, and a sequence depth between 6.4M – 11.1M was achieved with sequencing all libraries. The data was analysed with the Tuxedo suite (Trapnell et al. 2012) via GenePattern browser (Reich et al. 2006). The lists of genes and their Fragments per Kilobase of Exon per Million Mapped Fragments (FPKM) values were compared and statistically tested between the genotypes with Cuffdiff and were imported to excel and divided into more than 2-fold down regulated, more than 2-fold up regulated and non-differentially expressed lists. The differentially expressed genes were assigned statistically significant and had a p-value of <0.05.

Initial analysis of the lists of differentially expressed genes revealed more than 50% of those down-regulated in Nxt1 trans-heterozygotes compared to wild type were testis-specific or testis-enriched in expression. To filter out these genes, and concentrate on the somatic function of Nxt1, a list, containing genes that are testis-specific or testis-enriched, was cross-referenced with the differentially expressed lists to remove these genes using FlyMine (Lyne et al. 2007).

Several hundreds of genes were more than 2-fold down regulated in each of the three stages (Figure 3.5A), with 229 genes more than 2-fold down regulated in all the stages (Figure 3.5B). Between the different stages, 163-351 genes were more than 2-fold up regulated, with stationary larvae the most affected and only 39 were more than 2-fold up regulated in all stages (Figure 3.5C and D). To determine whether transcriptional regulation of ecdysone-responsive genes could underlie the air bubble defect I extracted the expression data for all the ecdysone-responsive genes (Fletcher and Thummel 1995). Each of the genes were analysed with respect to three different criteria to (1) be differentially expressed, (2) minimum of 2-fold up or down regulated and (3) a minimum FPKM value of 10 in at least one condition. I noticed that only four genes were mildly down regulated (Table 3.1). It has been shown previously that none of these genes had an air bubble phenotype when down regulated (Wright et al. 1996;

Vaskova et al. 2000). Therefore, I concluded that ecdysone-responsive genes were not responsible for the air bubble phenotype in Nxt1 trans-heterozygotes.

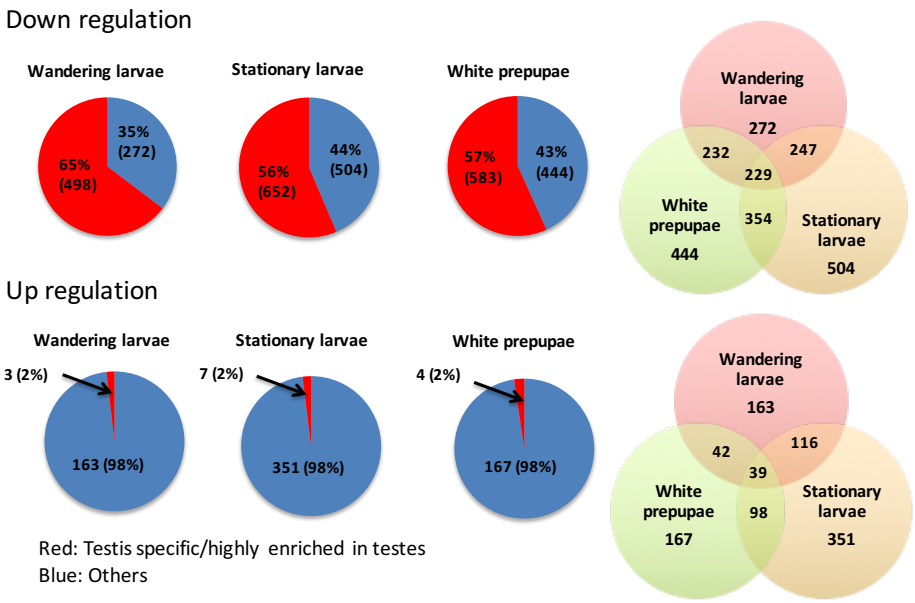


Figure 3.5: **More than 50% of down regulated genes across all three stages were either testis-specific or highly expressed in testis.** RNA sequencing showed that between 56-65% of higher than 2-fold down regulated genes had a high expression or were specifically expressed in the testis (shown in red). Only a few up regulated genes were highly or specifically expressed in the testis (shown in red). Venn diagram showing 229 genes were down regulated in all three stages, but only 39 of the up regulated genes were consistently up in all three stages.

Table 3.1: **List of all known ecdysone-responsive genes.** All ecdysone responsive genes (Fletcher and Thummel 1995) were examined by looking their differential expression status from cuffdiff (P-value <0.05), a minimum fold change of 2 and a FPKM threshold of at least 10 FPKM in one of the two genotypes.

Gene (wandering)	Differentially expressed?	Fold change (log2)	FPKM threshold?	Gene (stationary)	Differentially expressed?	Fold change (log2)	FPKM threshold?	Gene (prepupae)	Differentially expressed?	Fold change (log2)	FPKM threshold?
Br	NO	-0.10	YES	***Br***	NO	-0.20	YES	***Br***	NO	+0.02	YES
DHR3	NO	-0.31	NO	***DHR3***	NO	-0.87	YES	***DHR3***	NO	-0.35	YES
EcR	NO	-0.08	YES	***EcR***	NO	+0.44	YES	***EcR***	NO	+0.05	YES
Usp	NO	+0.54	YES	***Usp***	NO	+0.33	YES	***Usp***	NO	-0.14	YES
Ecd	NO	+0.28	YES	***Ecd***	NO	+0.24	YES	***Ecd***	NO	+0.36	YES
Edg78E	NO	+0.64	NO	***Edg78E***	NO	+2.15	NO	***Edg78E***	NO	+1.05	NO
Edg84A	NO	-	-	***Edg84A***	NO	-	-	***Edg84A***	NO	-	-
Edg91	NO	-0.05	YES	***Edg91***	NO	+0.46	YES	***Edg91***	NO	-0.16	NO
Eip55E	NO	+0.38	YES	***Eip55E***	NO	+0.30	YES	***Eip55E***	NO	0.00	YES
Eip63E	NO	0.00	YES	***Eip63E***	NO	-0.04	YES	***Eip63E***	NO	+0.36	YES
Eip63F-1	NO	-0.74	YES	***Eip63F-1***	YES	-0.97	YES	***Eip63F-1***	YES	+1.00	YES
Eip63F-2	NO	+0.26	NO	***Eip63F-2***	NO	+0.06	NO	***Eip63F-2***	NO	+0.85	NO
Eip71CD	YES	-0.87	YES	***Eip71CD***	NO	-0.47	YES	***Eip71CD***	NO	-0.18	YES
Eip74EF	NO	+0.03	YES	***Eip74EF***	NO	0.00	YES	***Eip74EF***	NO	+0.42	YES
Eip75B	NO	-0.37	YES	***Eip75B***	NO	-0.49	YES	***Eip75B***	NO	-0.30	YES
Eip78C	NO	-0.31	YES	***Eip78C***	NO	-0.37	YES	***Eip78C***	YES	+0.68	YES
Eip93F	NO	+0.35	NO	***Eip93F***	NO	-0.43	YES	***Eip93F***	NO	+0.29	YES
Ftz-f1	NO	+0.46	NO	***Ftz-f1***	NO	+0.66	NO	***Ftz-f1***	NO	-0.47	NO
Eig71Ea	NO	-1.18	YES	***Eig71Ea***	YES	-0.57	YES	***Eig71Ea***	NO	+0.32	YES
Eig71Eb	NO	-1.16	YES	***Eig71Eb***	YES	-1.62	YES	***Eig71Eb***	YES	-0.78	YES
Eig71Ec	NO	-2.16	NO	***Eig71Ec***	NO	-0.73	YES	***Eig71Ec***	NO	0.00	YES
Eig71Ed	NO	-1.27	YES	***Eig71Ed***	YES	-0.93	YES	***Eig71Ed***	YES	-0.42	YES
Eig71Ee	NO	-0.28	YES	***Eig71Ee***	NO	-0.08	YES	***Eig71Ee***	YES	+0.85	YES
Eig71Ef	NO	-2.44	NO	***Eig71Ef***	YES	-1.14	YES	***Eig71Ef***	NO	-0.34	YES
Eig71Eg	NO	-1.99	YES	***Eig71Eg***	YES	-1.55	YES	***Eig71Eg***	YES	-0.97	YES
Eig71Eh	NO	-	-	***Eig71Eh***	NO	-3.34	NO	***Eig71Eh***	YES	-1.48	YES
Eig71Ei	NO	-	-	***Eig71Ei***	NO	-2.50	NO	***Eig71Ei***	NO	-0.75	YES
Eig71Ej	NO	-	-	***Eig71Ej***	NO	-	-	***Eig71Ej***	NO	-0.92	NO
Eig71Ek	NO	-	-	***Eig71Ek***	NO	-	-	***Eig71Ek***	NO	+1.21	NO

While the ecdysone-responsive genes were not affected I double-checked the RNA-seq data if genes known to be important for biosynthesis of ecdysteroids are affected in the Nxt1 mutant. The biosynthesis of ecdysteroids is encoded by the terminal cytochrome P450 hydroxylases also known as Halloween genes (Grieneisen, Warren, and Gilbert 1993). Halloween genes catalyze most of the process from cholesterol to 20-hydroxyecdysone. Null mutants of any of the Halloween genes lead to morphogenetic abnormalities such as failure of head involution and cuticle formation, leading to embryonic death (Chavez et al. 2000; Niwa et al. 2004; Petryk et al. 2003; Warren et al. 2002; Warren et al. 2004). I searched for these genes in our RNA sequencing data and found that the majority of the P450 enzymes were not differentially expressed (Table 3.2). Only *Cyp307a2* (*spok*) was less than 2-fold down regulated. *Spok* has been speculated to be involved in the synthesis in the so-called "blackbox" between 7-dehydrocholesterol and Δ4-diketol early in the ecdysone synthesis (Ono et al. 2006). It is only mildly down regulated, and since the ecdysone-responsive genes are expressed normally, it is not expected that it would have a big impact on the synthesis of ecdysone. Both analyses strongly suggest that ecdysone has no effect in the air bubble phenotype.

Table 3.2: **List of cytochrome P450 hydroxylases enzymes**. Log2 fold change of six genes in the three different stages. *More than 1.5-fold down regulated. ΨLess than 10 FPKM in at least one condition.

Gene	Log2 fold change – Wandering	Log2 fold change – Stationary	Log2 fold change - Prepupae
***dib* (*Cyp302a1*)**	-0.17	+1.59$^{\Psi}$	+0.07$^{\Psi}$
***sad* (*Cyp315a*)**	+0.23	+0.33	+0.53
***shd* (*Cyp314a1*)**	+0.13	+0.07	+0.11
***phm* (*Cyp306a1*)**	-0.33	+0.54	-0.13
***spo* (*Cyp307a1*)**	+0.88$^{\Psi}$	0.00$^{\Psi}$	-0.10$^{\Psi}$
***spok* (*Cyp307a2*)**	-0.22	-0.60	-0.89*

3.4 Nxt1 trans-heterozygotes third instar larvae muscles show signs of early degeneration

The analysis of the ecdysone signaling pathway by RNA-seq indicated that the air bubble phenotype is almost certainly not caused by defects in ecdysone and the ecdysone-responsive genes, hence the underlying defect must be elsewhere. The air bubble migrates through the pupa via abdominal muscular contractions, and so I investigated muscle function in the Nxt1 trans-heterozygotes mutant animals. A larval movement assay was used to track 1st, 2nd, and 3rd instar larvae to compare their speed between *w^{1118}* and Nxt1 trans-heterozygotes. Larvae were tracked on a 1% agar plate with a control and odor on each side (Figure 3.6A and B). These experiments showed that there was no significant difference in the average speed of mutant larvae compared to wild type at either 1st or 2nd instar stage. However, mutant 3rd instars were significantly slower than control animals (Figure 3.6C-G). Normally there is a dramatic increase in migration speed at this developmental transition, wild type 3rd instar larvae travel up to 5x faster than 2nd instars (Figure 3.6F). In contrast, Nxt1 trans-heterozygous 3rd instar larvae were no faster than 2nd instar larvae (Figure 3.6G), despite being much larger. Larvae move around through muscular contraction and these data indicate something affects the movement.

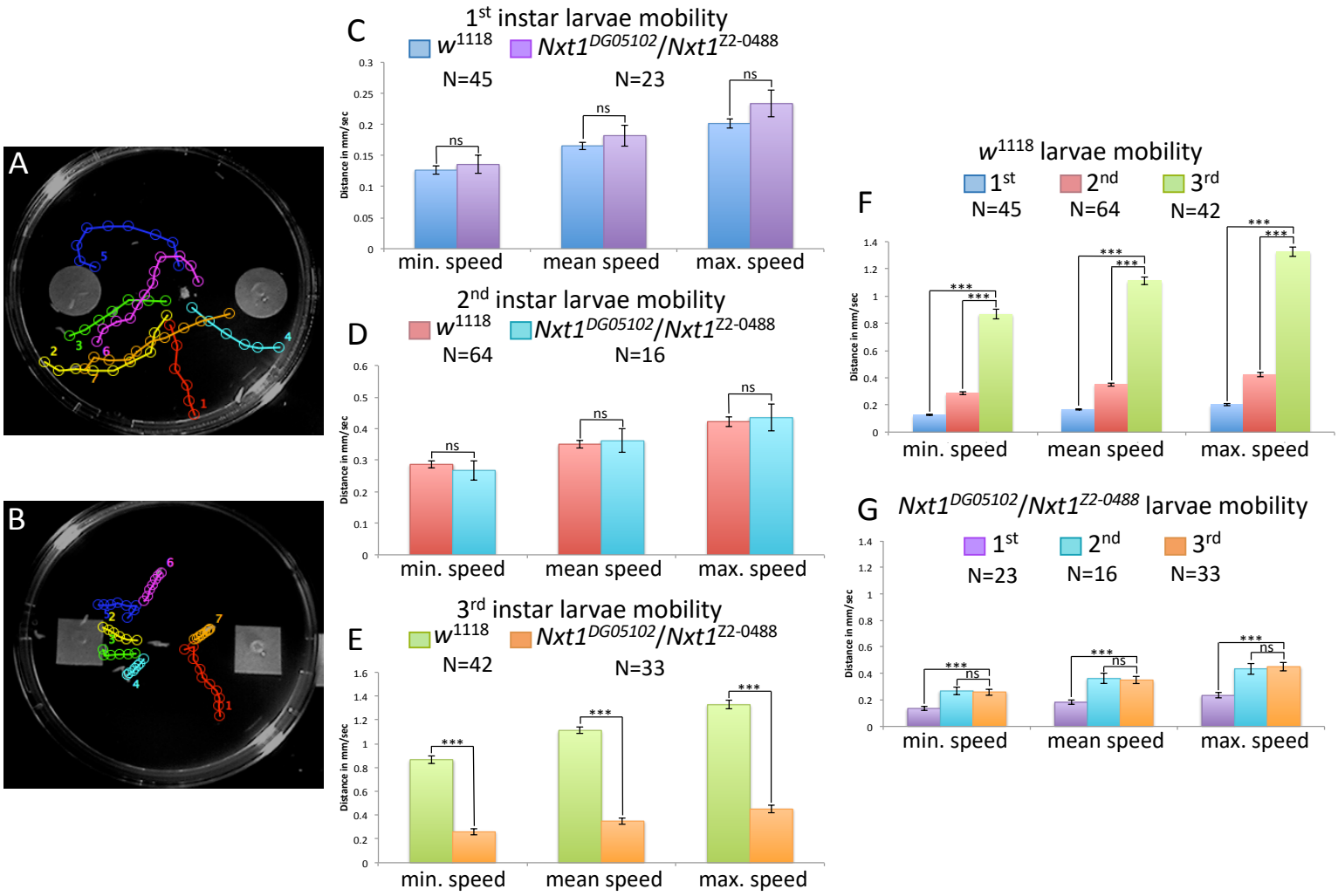


Figure 3.6: **Nxt1 trans-heterozygotes 3rd instar larvae have a reduced mobility.** A-B) Image of agar plate with control substance on the left and odor on the right. Larvae tracking was done via MtrackJ on ImageJ. C-E) Mobility analysis of 1st, 2nd and 3rd instars of wild type and Nxt1 trans-heterozygotes. F-G) All 1st, 2nd and 3rd instar data combined for each genotype. Student's t-test *** p value = <0.001. n.s = not significant. N = number of samples.

Muscle contractions are required for the air bubble to move to the anterior for head eversion. I examined directly at the larval muscles of 3rd instar larvae stained with F-actin marker phalloidin, which stains the thin filaments in the sarcomere. In *Drosophila* larva, each of the abdominal hemisegments A2-A7 has a stereotypical pattern of 30 different muscles. This pattern was clearly observed in the wild type (Figure 3.7A), however, our mutant showed clear signs of muscle degeneration (Figure 3.7B). Eight hemisegments (30 muscles per hemisegment) in each of the six Nxt1 trans-heterozygotes stationary larvae (240 muscles per larva) were examined in detail.

All muscles in the hemisegments were examined by muscle mass, loss of sarcomeric structure, degeneration and torn. Defects seen were quite variable, and 77% of all observed muscles in the six larvae were degenerating, 22% had a loss of sarcomeric structure, 15% were thinner and 8% were torn (Figure 3.7F-K). Degenerating muscles, typically, also showed loss of sarcomeric structure and were thinner. Degenerating muscles in third instars are likely to be the cause of the effect on larva movement observed previously. The air bubble migration occurs around 12-hour into the pupa phase, and with degenerating muscles in Nxt1 trans-heterozygotes, air bubble movement is greatly impaired (Figure 3.3).

The Nxt1 mutant animals also show defects in muscle structure. The normal muscle sarcomere structure consists of thick and thin filaments. Staining for F-actin with phalloidin shows the thin filaments in the sarcomere (Figure 3.8A). For Nxt1 trans-heterozygotes, the sarcomere structure was often found to be compromised (Figure 3.8C). In wild type, F-actin staining of muscles shows ribbed lines, indicating the thin filaments (Figure 3.8A). In Nxt1 trans-heterozygotes, these ribbed lines can be compromised (i.e. missing or faint) throughout the muscle (Figure 3.8C). However, Nxt1 trans-heterozygotes have muscles that did not show any degeneration and those typically had normal sarcomere structure (Figure 3.8B).

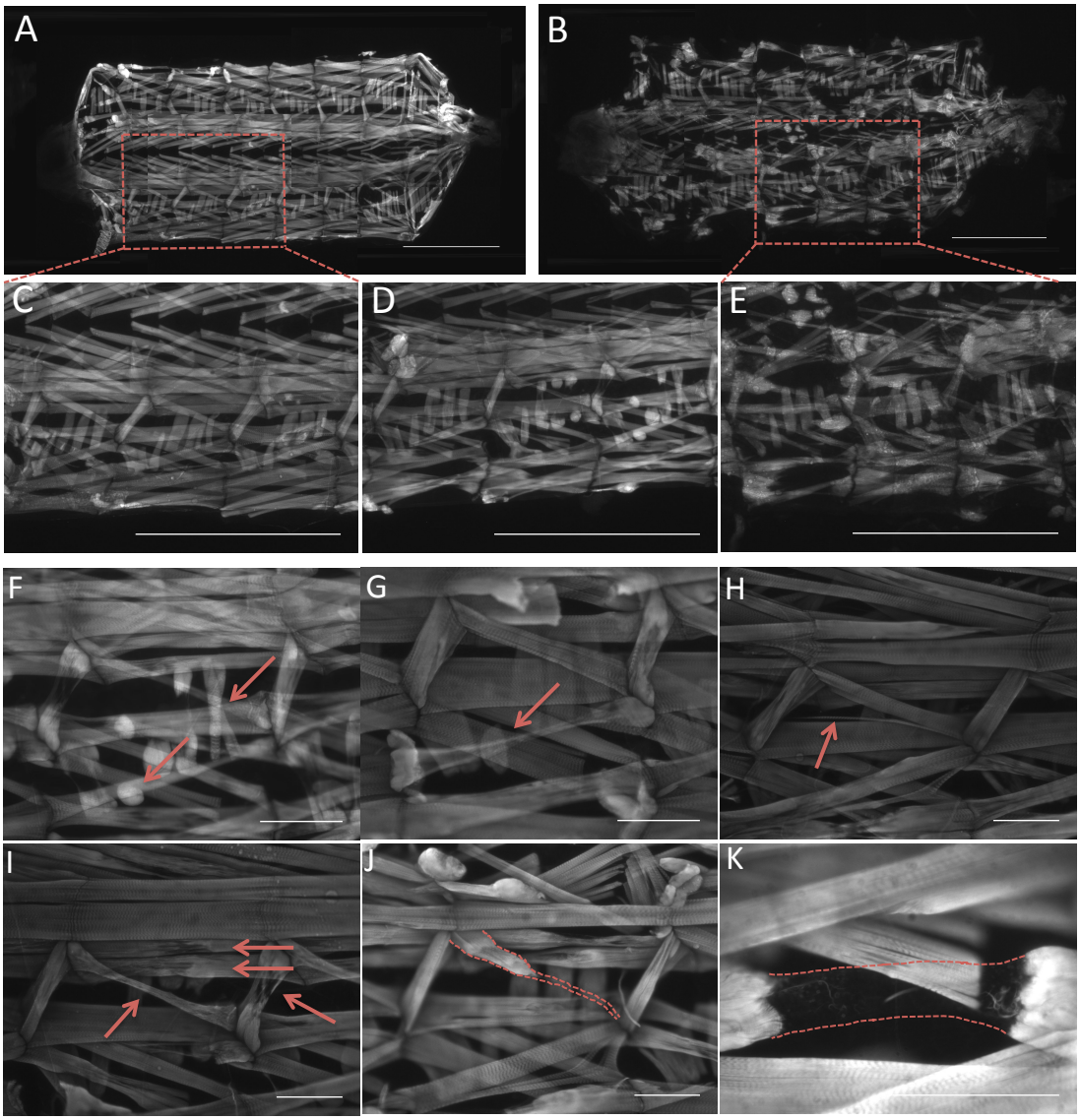


Figure 3.7: **Nxt1 trans-heterozygotes have muscle degeneration in 3rd instars.** A) Overview of wild type 3rd instar larva muscles. B) Overview of *Nxt1*^{DG05102}/*Nxt1*$^{Z2-0488}$ 3rd instar larva muscles. C) Close-up of larva hemisegments from (A). D) Hemisegments from *Nxt1*^{DG05102}/*Nxt1*$^{Z2-0488}$ with mild muscle degeneration. E) Close-up of larva hemisegments from (B). F) Lateral Transverse (LT) 1 and 2 crossover (short arrow). F-G) Short LT4 muscle (long arrow). H) Lateral Oblique (LO) 1 fiber split (red arrow). I) Ventral longitudinal (VL) 1 and 2, Segment Border Muscle (SBM) and LO 1 degenerating (indicated by arrows) I) J) LO 1 partially degenerated into small bundle of fibers connected to the SBM. K) Dorsal Oblique (DO) 2 showing strings of fibers (red dotted lines) attached to each end of the muscle. Samples orientated from posterior (left) to anterior (right). A-E) Scale bar = 1mm. F-K) Scale bar = 200µm. Figures were taken on Olympus BX50.

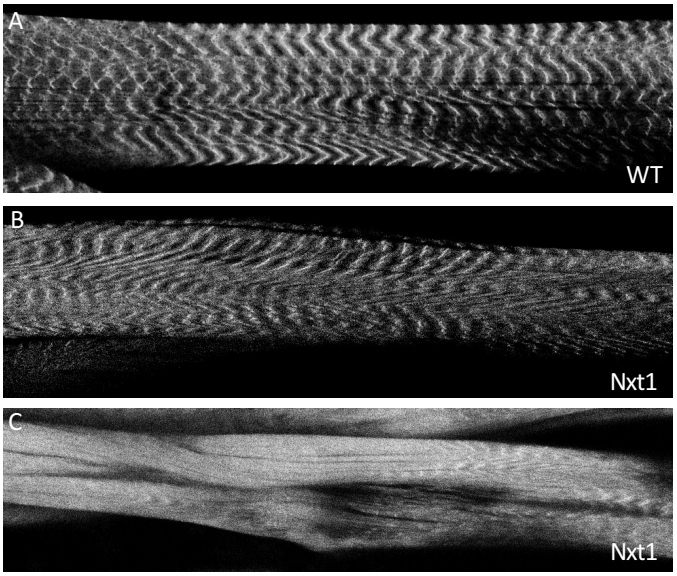


Figure 3.8: **Sarcomere structure compromised in degenerating muscles.** A) Wild type muscles stained for F-actin with phalloidin show normal, ribbed lines, thin filaments. B) Nxt1 trans-heterozygotes non-degenerating muscles with normal, ribbed lines, thin filaments. C) Nxt1 trans-heterozygotes degenerating muscles with compromised thin filaments. Figures were taken on a Leica DM6000B upright microscope.

I used two Nxt1 RNAi lines to confirm that the muscle phenotype observed was due to defects in Nxt1 rather than being a non-specific effect. RNAi line 103146 is on the 2nd chromosome, whereas RNAi line 52631 is on the 3rd chromosome (VDRC). Both RNAi lines were validated in spermatocytes and phenocopied the *z2-0488* allele in testis (Caporilli et al. 2013). The hairpin length for both 103146 (286nt) and 52631(340nt) was long and can be problematic for cleavage with endogenous dicer, so therefore UAS-dicer was crossed with both RNAi lines to increase the probability of hairpins to be cleaved into siRNAs. The crosses were put at 29^{o}C throughout development with an induction at the egg laying phase. The larvae were dying early 2nd instar with both RNAi lines. Dissecting early second instars and staining them with phalloidin showed that most of the muscles were degenerating (Figure 3.9). In another experiment, both RNAi lines were put at 25^{o}C and shifted to 29^{o}C starting at 1st instar larvae which revealed that RNAi line 103146 had 3rd instar larvae that were able to pupate before dying, whereas RNAi line 52631 showed early 2nd instar lethality. These

3rd instar larvae also showed degeneration of muscles and loss of muscle mass (Figure 3.10). Therefore, the larval muscle degeneration was caused due to a reduction of Nxt1.

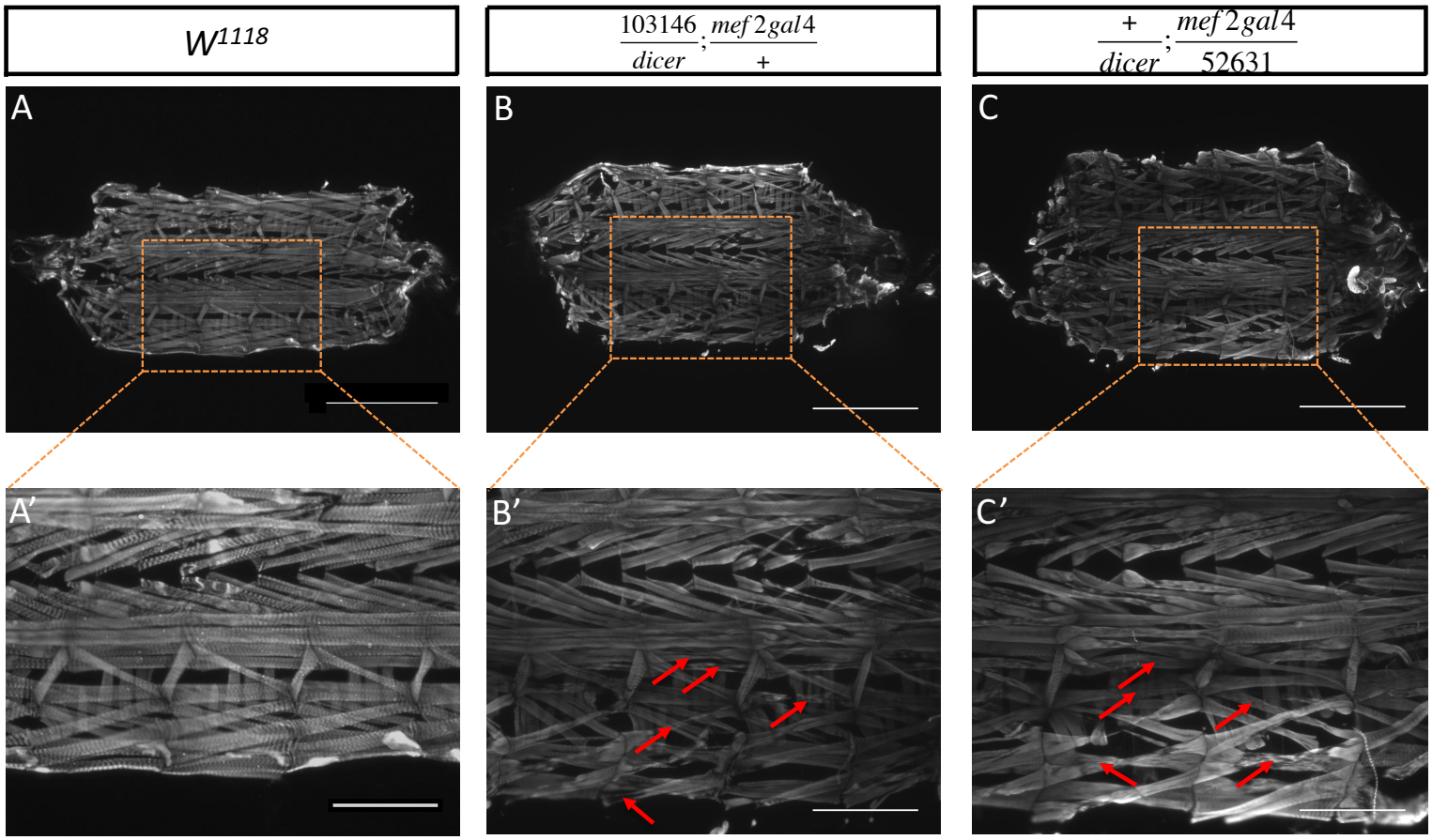


Figure 3.9: **Early 2nd instar show muscle degeneration with Nxt1 RNAi.** A-A') Wild type 2nd instar larva muscles, stained with phalloidin, showing normal muscle composition and no damage. B-B') Signs of muscle degeneration with the RNAi line 103146 visible in early 2nd instars with thinner muscles and fiber damage (see insert B'). C-C') Similar degeneration defects shown for the RNAi line 52631. A-C) Scale bar = 1mm. A'-C') Scale bar = 0.5mm. Red arrows indicate examples of degenerating muscles. Figures were taken on an Olympus BX 50.

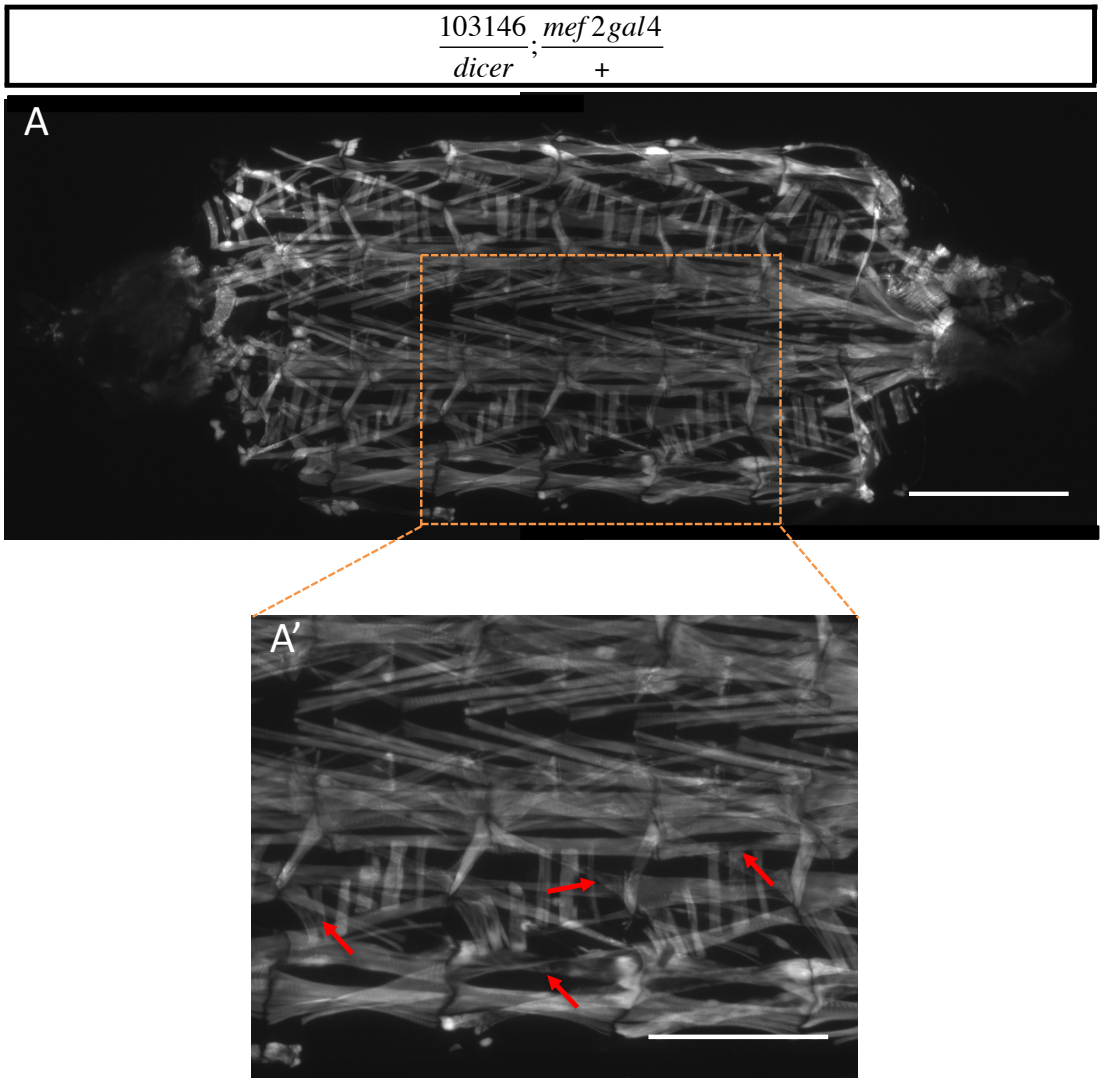


Figure 3.10: **Nxt1 RNAi 103146 have 3rd instar muscle degeneration after 1st instar induction.** A) RNAi induced from 1st instars showed lethality at pupa stage for Nxt1 RNAi 103146. 3rd instars show signs of early degeneration. A') Close-up of hemisegments from (A). A) Scale bar = 1mm. A') Scale bar = 500µm. Red arrows indicate examples of degenerating muscles. Figures were taken on an Olympus BX 50.

3.5 Muscle degeneration in Nxt1 trans-heterozygotes is caused by growth of the muscles in the third instar phase

During the larva phase, larvae eat food to grow and molt, first to become a 2nd instar, then to become a 3rd instar larva. During the last instar phase, larvae will grow substantially compared to earlier instars. Larvae removed from the food 70 hours after egg laying do not grow, but still move up to the normal time for pupation (~112 hours

AEL) and many survive through to adulthood (Beadle, Tatum, and Clancy 1938). This is also known as the "seventy-hour change" (Beadle, Tatum, and Clancy 1938). Any larvae deprived of food any time up to the "70-hour change" are retarded in development, the retardation being a function of the length of the starvation period (Beadle, Tatum, and Clancy 1938). Larvae around 70 hours AEL are towards the end of 2nd, or early 3rd instars. Nxt1-transheterozygotes larvae at 70 hour AEL have WT-like muscles, since the degeneration occurs later when the 3rd instar larva grows. To test whether growth or use (movement) is implicated in the muscle degeneration I fed ~60 WT larvae up to 70 hours before removing all larvae from the food. However, only 7% of the w^{1118} larvae survived and pupated, however, none of the pupae survived through the end of the stage (Figure 3.11A). I extended the experiment by feeding WT larvae up to 71, 72 and 73 hours (Figure 3.11A). 14-25% of the larvae were able to pupate and less than 5% of those pupae were able to emerge as adults (Figure 3.11A). I simultaneously looked at ~60 of the Nxt1 trans-heterozygote larvae that were stopped feeding 70 hours AEL and found, similar to w^{1118}, that only a few larvae pupated (Figure 3.11B), and none emerged as adults (data not shown). Interestingly, larvae that did not pupate were able to survive for several days before dying. Ten larvae were dissected of Nxt1 trans-heterozygotes four days after they were pulled out of the food and were stained with phalloidin. For all ten larvae, no muscle damage was observed (Figure 3.11C). Normally these larvae should have been towards the end of the pupa phase, if they pupated, but instead were still late 2nd or early 3rd instars (based on the orange tip at the posterior spiracle). This showed that the muscle degeneration occurs during the growth phase in the 3rd instar larvae.

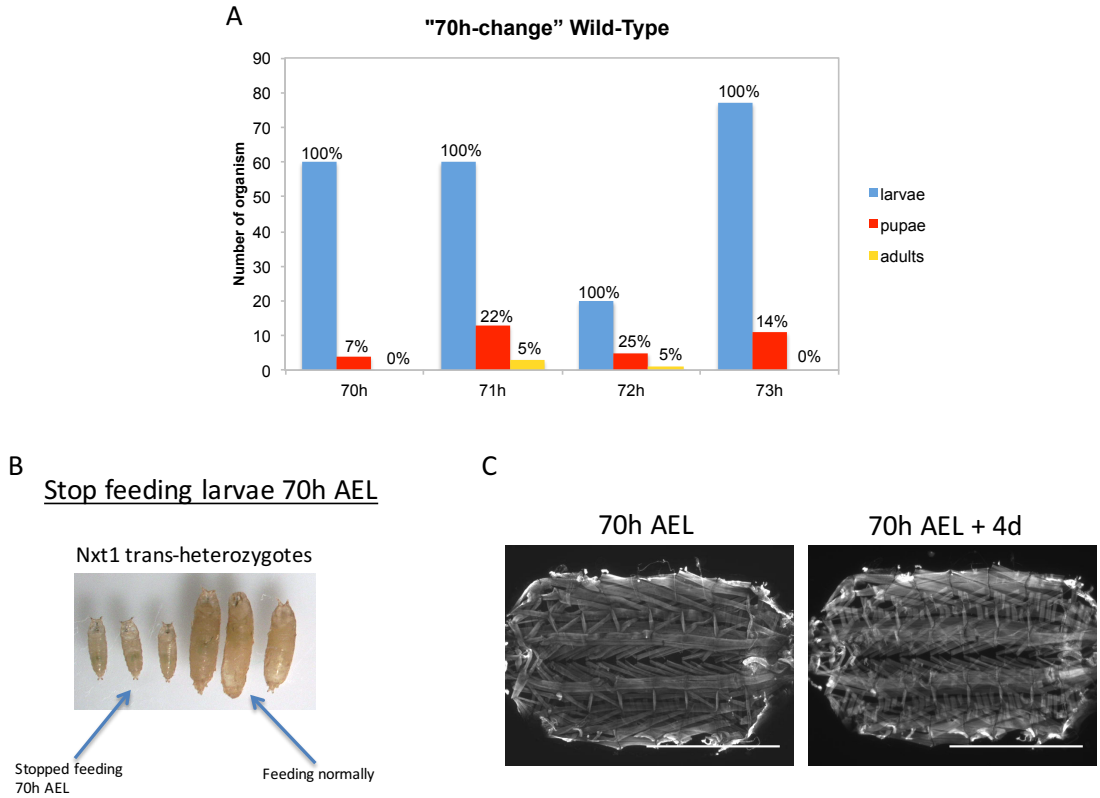


Figure 3.11: **Growth of muscles causing degeneration defects.** A) Wild type larvae were stopped feeding after 70, 71, 72 and 73h AEL. B) Nxt1 trans-heterozygote pupae were smaller if stopped feeding 70h AEL. C) Nxt1 trans-heterozygote 70h AEL were examined for muscle degeneration. Four days after no feeding and not pupating showed muscles were still intact (N= 10). Figures were taken on an Olympus BX 50.

The Insulin/Akt/TOR pathway is involved in the nutrient supply for the growth of larva muscles. The muscles during larval development undergo ~50-fold increase in fiber size in only five days (Demontis and Perrimon 2009). Transcription factors such as *FOXO*, *dMyc* and *Mnt* are necessary to achieve normal muscle growth. Luciferase assays revealed that FOXO inhibit *dMyc* transcriptional activity, and reduction of *dMyc* inhibits the muscle growth so muscles will not go into a hypertrophy state (Demontis and Perrimon 2009). I used qRT-PCR to assay the expression of some components of the Insulin/Akt/TOR pathway to identify potential defects in the muscle growth system during wandering, stationary and white prepupae phase (Figure 3.12A). Of the genes analysed, a reduction of ~3 to 5-fold *FOXO* and ~2 to 3-fold *CG11658* during the wandering and white prepupae phase, respectively (Figure 3.12B). A decrease in *FOXO*

would lead to less reduction of *dMyc*, which in turn, would promote protein synthesis. Not much is known about the exact role of *CG11658* in *Drosophila* and what its role is in inducing muscle atrophy, however, a reduction of *CG11658* is expected to lead to less induction of muscle atrophy (Figure 3.12A; (Demontis and Perrimon 2009)). The experiment revealed that the Insulin/Akt/TOR pathway is mostly unaffected at the level of mRNA expression. The observed muscle phenotype in Nxt1 trans-heterozygotes 3rd instar stationary larvae is unlikely to be a result of muscle atrophy via the Insulin/Akt/TOR pathway and rather be muscle degeneration of unknown origin.

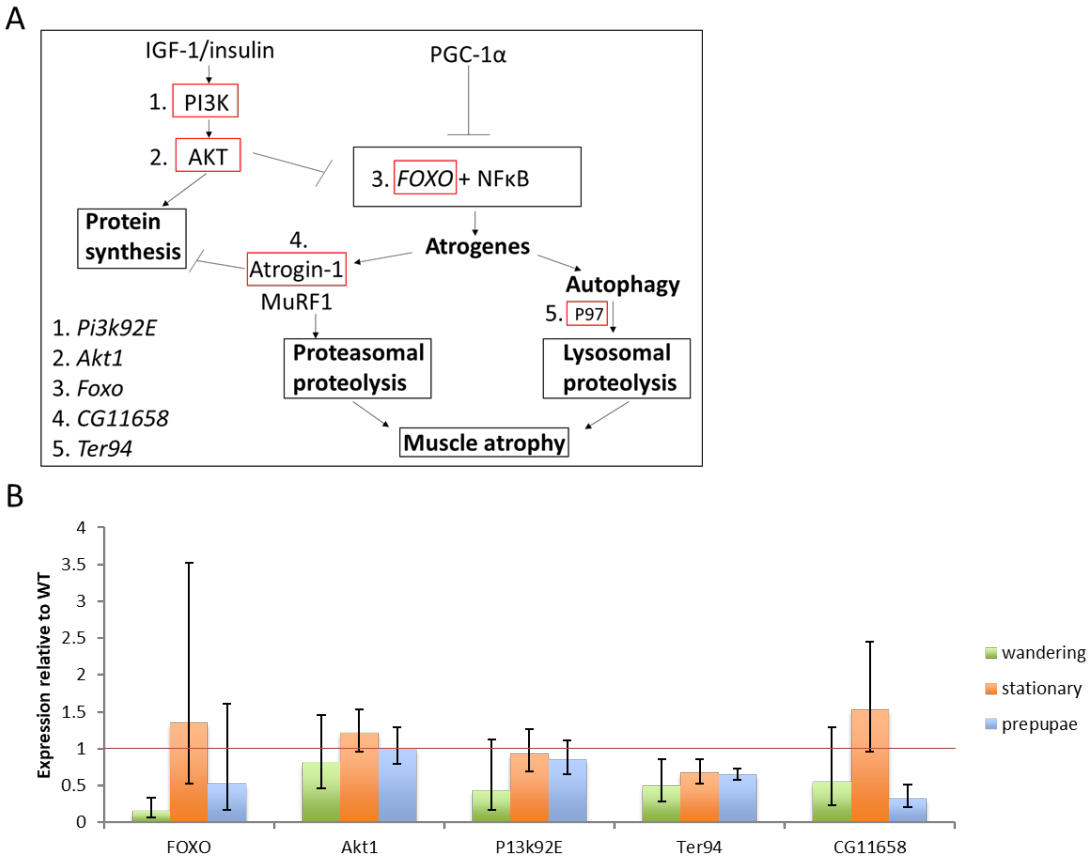


Figure 3.12: **Insulin/Akt/TOR muscle nutrient pathway mostly unaffected in Nxt1 trans-heterozygotes in wandering, stationary and prepupae phase.** A) Overview of the Insulin/Akt/TOR pathway that controls muscle growth and the muscle atrophy process. B) qRT-PCR of several genes (red box in A; *Drosophila* gene name in italics) involved in Insulin/Akt/TOR pathway in wandering larvae, stationary larvae, and white prepupae.

3.6 Increasing *abba* rescues muscle phenotype, but not semi-lethality with a muscle-specific driver

The known role of Nxt1 in RNA export and transcriptional control suggested that the muscle phenotype could be due to defects in gene expression. To test this possibility in detail the expression of known muscle-specific or –enriched genes was analysed in the whole larval RNA-seq data. Three lists (more than 1.5-fold down regulated, more than 1.5-fold up regulated and non-differentially expressed) were created and provided to FlyMine to identify muscle genes. From these lists, a selection of 13 different muscle genes known to have a role in larval muscles were selected (Table 3.3). Nine genes were more than 1.5-fold up regulated, with only one down regulated in the mutant larvae. *abba* (also known as *thin* (*tn*)), the only gene down regulated in our list, is an essential TRIM/RBCC protein to maintain integrity of sarcomeric cytoarchitecture (Domsch, Ezzeddine, and Nguyen 2013). *abba* is a large gene with several long introns and has a RING finger, B-Box & CC-domain and several NHL repeats (Figure 3.13A). Performing a qRT-PCR in 10 individual stationary larvae and a mix of 30 stationary larvae, compared to a mix of 10 and 30 wild type stationary larvae, respectively, showed that *abba* was significantly down regulated in 9 out of 10 larvae (Figure 3.13B). To determine whether the increase in the expression of *abba* could modify the mutant phenotype I used a UAS-*abba*$^{\text{full length (fl)}}$ construct (kindly gifted by Hanh T. Nguyen; hereafter referred to as UAS-*abba*) and expressed it under the muscle specific driver mef2gal4 in the Nxt1 trans-heterozygote background. Dissecting 3rd instars, and staining muscles with phalloidin, showed that the muscle degeneration was rescued (Figure 3.13C). qRT-PCR of 10 individual stationary larvae, compared to a mix of 10 wild type stationary larvae, showed that the 8 out of 10 had a rescue in expression to normal level of *abba* (Figure 3.13D). It appears that the UAS-*abba*$^{\text{fl}}$ construct does not work with the same efficiency in all larvae, which is why only 8 out of 10 had a normal of even higher expression of *abba*. Importantly, the increase of expression of *abba* rescues the muscle phenotype observed in the third instar larvae.

Table 3.3 **Overview of larval muscle genes from RNA sequencing.** Thirteen larval muscle genes with log2 fold change from RNA sequencing showing majority of genes up regulated with only one down regulated. Red = more than 1.5-fold down regulated. Green = more than 1.5-fold up regulated

Gene name	Log2 fold change
abba	-0.75
up	+1.54
Arc1	+1.84
Prm	+1.62
Mlc2	+2.16
Mhc	+3.02
Act57B	+2.04
mib2	-0.48
Mlp84B	+0.88
Actn	+0.33
Mp20	+2.08
wupA	+0.90
sls	+0.16

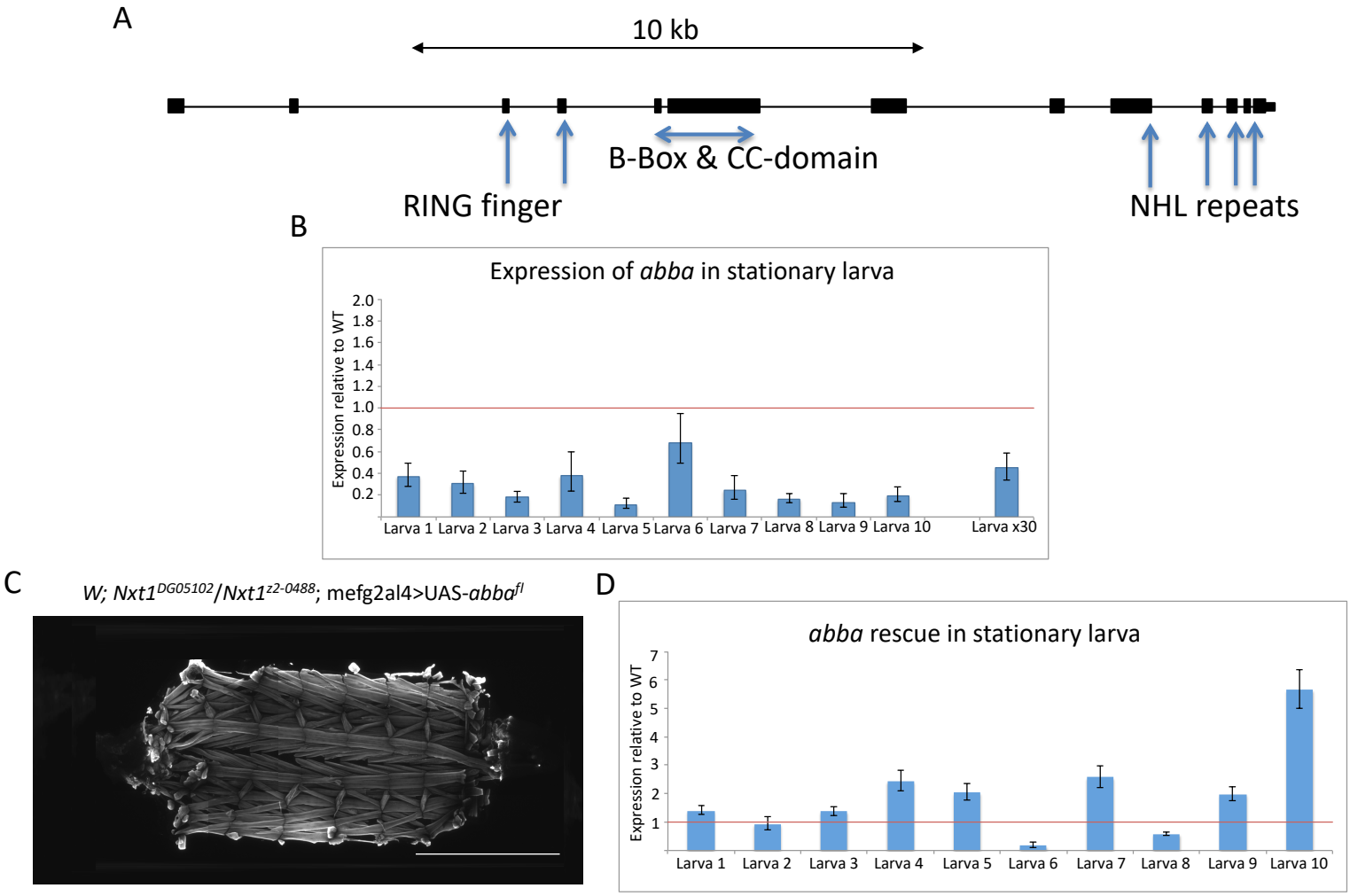


Figure 3.13: **Increase of *abba* expression rescues muscle degeneration.** A) Overview of longest *abba* transcript including all known domains. B) Expression of *abba* in 10 individual stationary larvae plus a mix of 30 stationary larvae in Nxt1 trans-heterozygotes, compared to a mix of 10 and 30 stationary wild type larvae, respectively. C) Increasing *abba* with the UAS/Gal4 system specifically in muscles rescues muscle degeneration. D) Expression of *abba* in 10 individual larvae from the rescue construct, compared to a mix of 10 stationary wild type larvae. Figure was taken on an Olympus BX 50.

I performed a pupa viability assay for the *abba* rescue genotype using the same approach as described in section 3.2. While increasing mef2gal4>UAS-*abba* in Nxt1 trans-heterozygotes rescues the muscle phenotype, semi-lethality at the pupa phase still occurred at a similar survival rate as Nxt1 trans-heterozygotes (Figure 3.14A). Therefore, semi-lethality in Nxt1 trans-heterozygotes cannot be rescued by increasing *abba* in the muscles.

The experiment was extended by expressing UAS-*abba* with a moderate ubiquitous promoter armgal4 in the Nxt1 trans-heterozygote background. The phalloidin staining and viability assay were compared to the mef2gal4>UAS-*abba*. The muscles in 8 hemisegments were examined and any damage, such as muscles torn, thin, loss of sarcomeric structure, missing muscles, were counted to calculate muscle damage (see 2.2.10) The mef2gal4>UAS-*abba* larvae muscles were mostly intact with three outliers between 15-29% muscle damage (N= 16), whereas the armgal4>UAS-*abba* larvae muscles had between 2-11% muscle damage (N= 10) in Nxt1 trans-heterozygotes background (Figure 3.14B). However, armgal4>UAS-*abba* gave a much better pupa survivability than mef2gal4>UAS-*abba* (70% (N= 210) vs. 23% (N= 187), respectively; Figure 3.14A). Interestingly, rescuing *abba* across all tissues had a big impact in the pupa survivability, indicating that Abba might had roles elsewhere in the larva.

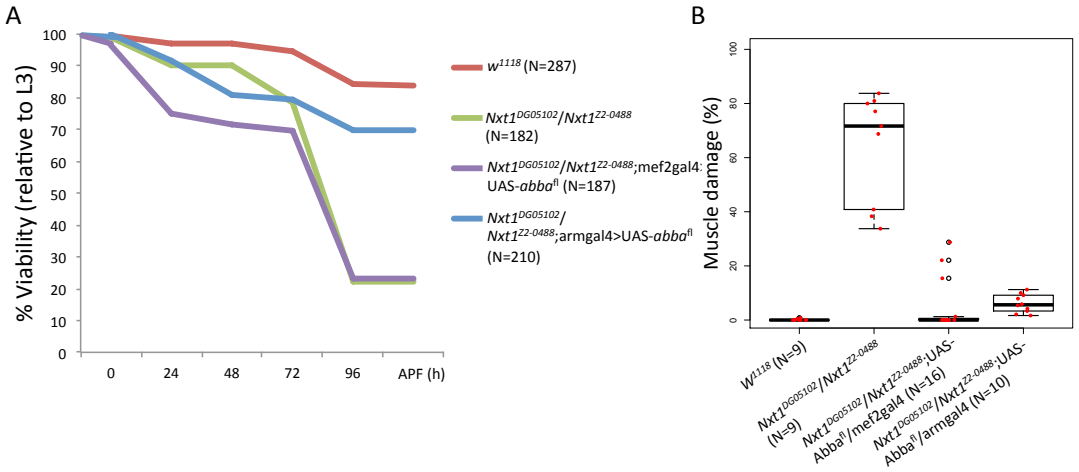


Figure 3.14: **Increasing *abba* with armgal4 increases pupa survivability.** A) Pupa viability using mef2gal4 and armgal4 drivers (23% and 70% survivability, respectively). B) Muscle damage observed from eight hemisegments per larva from all genotypes (see 2.2.10 for calculation). Muscle damage was reduced with both mef2gal4 and armgal4 drivers.

---

abba has six different isoforms with 13 exons in its longest transcript and contains introns up to 4 kb in length (FlyBase). With the reduction of Nxt1 in Nxt1 trans-heterozygotes I compared nascent versus spliced transcript expression to determine whether transcription of the gene was altered in mutants, and more interestingly,

whether different regions of the transcript were affected. Four different regions were selected that would target as many isoforms as possible (Figure 3.15A). Total RNA from stationary larvae (see 2.1.3.3 for extraction method) was used for this experiment. For spliced transcript expression, primer pairs were in adjacent exons. For nascent expression, primers were designed within the intron flanked by the two exons used for spliced expression. Interestingly, all regions throughout *abba* had lower spliced transcript expression than nascent transcript expression (Figure 3.15B). For *abba*, loss of Nxt1 affects its process to splice out introns during transcription.

It has previously been shown that genes, with particularly large introns, were sensitive to the disruption of the EJC (Ashton-Beaucage and Therrien 2011). The spliceosome recognizes and removes intronic sequences and is guided in this activity by several splicing regulatory factors that determine the splice sites that are recognized and used for intron removal. The EJC was identified to regulate splicing of *mapk* transcripts, and mutations of the EJC components resulted in several *mapk* exons to be skipped (Ashton-Beaucage and Therrien 2011). I examined the four genes to determine the expression of the EJC with the sequencing data and qRT-PCR. RNA sequencing data showed that none of these genes were differentially expressed (data not shown), however qRT-PCR confirmed that that all four genes were mildly (less than 1.5-fold) down regulated (Figure 3.15C). These data revealed that the mild down regulation of the EJC components might have an effect of the splicing in *abba* and potentially other genes.

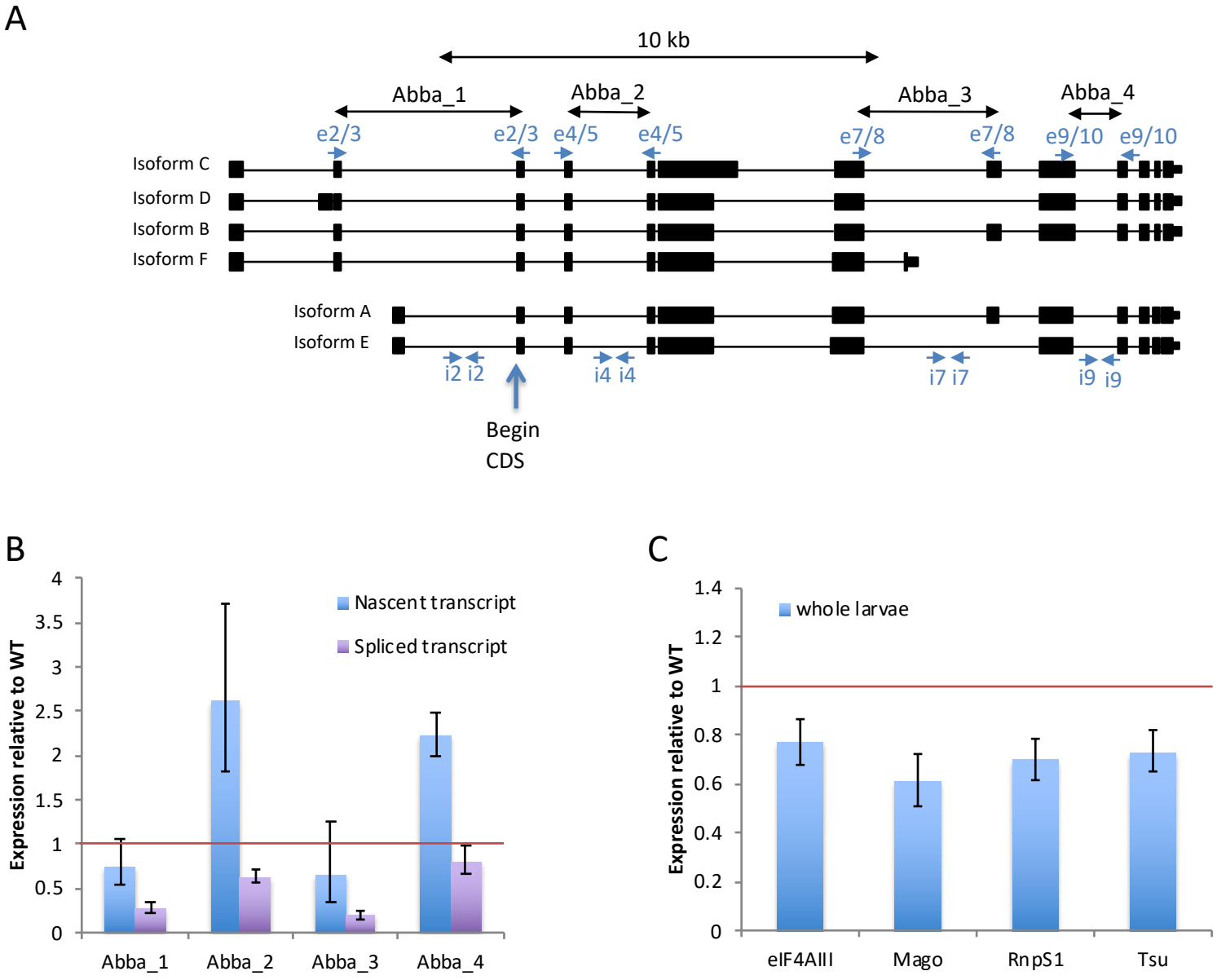


Figure 3.15: ***abba* has a reduced expression of spliced transcripts.** A) Overview of six different *abba* isoforms, including the four regions used for nascent vs. spliced transcript analysis. B) qRT-PCR of four different regions in *abba* comparing nascent vs. spliced expression. C) qRT-PCR of four genes composing the EJC. Short blue arrows indicate positions of primer pairs used to examine nascent and spliced transcript expression. Large blue arrows indicate beginning of coding sequence. I = introns, E= exons.

3.7 Expression balance between *abba* and *mib2* not required for muscle integrity

Increasing *abba* rescued the muscle phenotype with a muscle specific driver and the pupa viability with a moderate ubiquitous driver. To understand the mechanism, I investigated the physical and genetic interactions between *abba* and other important muscle genes (Figure 3.16). Abba has been identified in a yeast two-hybrid screen with Mind Bomb 2 (Mib2) as bait (Domsch, Ezzeddine, and Nguyen 2013). Mib2 maintains the integrity of fully differentiated muscles and prevents their apoptotic degeneration (Nguyen et al. 2007). RNA sequencing data showed that *mib2* was not differentially expressed and this was also confirmed with a qRT-PCR (data not shown). I hypothesized that muscle integrity may depend on the balance between *abba* and *mib2*. In this scenario increasing *abba* or decreasing *mib2* would rescue the muscle integrity, whereas decreasing *abba* and increasing *mib2* would be detrimental for the muscles (Figure 3.17). To test this hypothesis, I designed gain- and loss- of function experiments for both *abba* and *mib2* in Nxt1 trans-heterozygotes and analysed the muscle integrity, axial ratios, pupa viability and larval mobility (Figure 3.18).

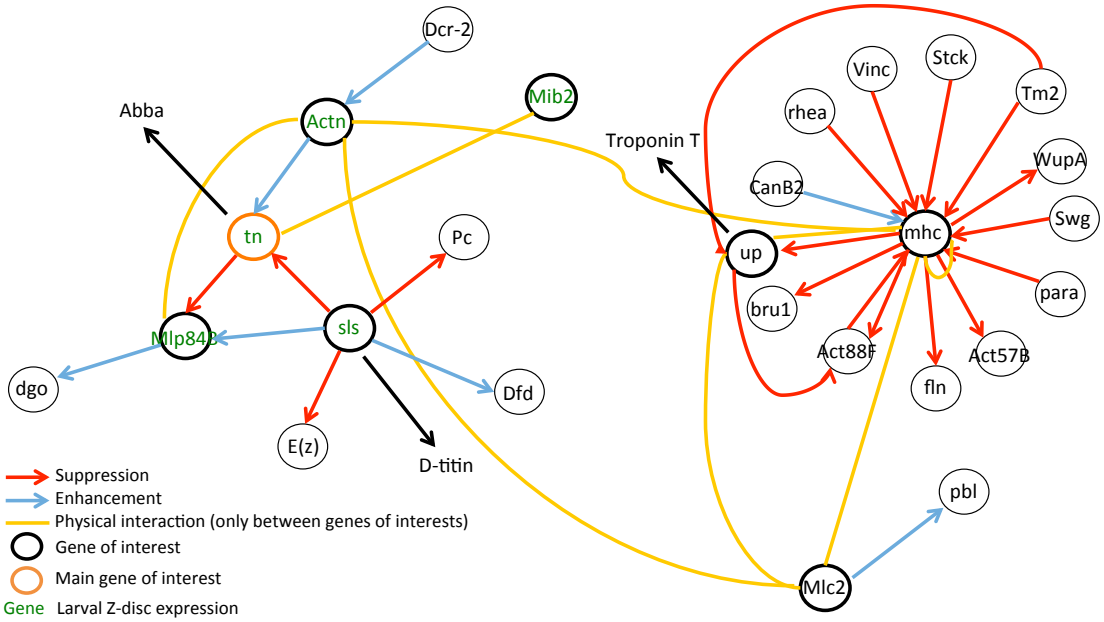


Figure 3.16: **Physical and genetic interactions between Abba and other larval muscle genes.** Abba has several genetic interactions with other genes expressed in the larval Z-discs (Domsch, Ezzeddine, and Nguyen 2013). Abba has only one known physical interaction with Mib2, which is also expressed in larval Z-discs. See esyN network in FlyBase for all publications regarding the interactions with the muscle genes.

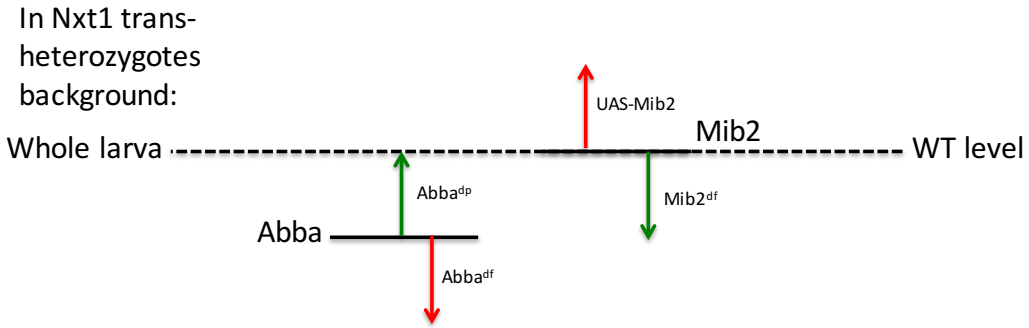


Figure 3.17: **Visual representation of the proposed expression balance effects on muscle integrity.** Hypothesis of genetic interactions between Abba and Mib2 maintaining muscle integrity. If the balance between the expression of the two genes is altered, muscle integrity is affected. Increasing *abba* or decreasing *mib2* should rescue muscle integrity in the Nxt1 mutant background, whereas decreasing *abba* or increasing *mib2* should enhance muscle integrity defects.

Increasing abba or decreasing mib2

If the balance between *abba* and *mib2* is important for muscle integrity, then increasing *abba* or decreasing *mib2* in the Nxt1 trans-heterozygotes should restore this balance, and result in better muscle integrity and possibly higher pupa viability. Indeed, increasing *abba* with the UAS and duplication constructs in Nxt1 trans-heterozygotes showed less muscle damage than Nxt1 trans-heterozygotes (Figure 3.18C). However, mef2gal4>UAS-*abba* in Nxt1 trans-heterozygotes did not rescue the pupa viability (Figure 3.18A). Pupal viability was slightly increased in Nxt1 trans-heterozygotes also bearing an *abba* duplication or a *mib2* deficiency, (from 22% (N= 63) to 43% (N= 197); Figure 3.18A). The axial ratios were like wild type for mef2gal4>UAS-*abba* and Abba duplication, but was less wild type like with a lower median for *mib2* deficiency in Nxt1-trans-heterozygotes (Figure 3.18B). These pupae were slightly thinner than wild type. Finally, Nxt1 trans-heterozygotes in combination with mef2Gal4>UAS-*abba* and *mib2* deficiency both resulted in an increase in mobility compared to Nxt1 trans-heterozygotes but were still significantly slower than wild type (Student's t-test * $p<0.05$; Figure 3.18D). For Nxt1 trans-heterozygotes bearing *mib2* deficiency, there was no difference in mobility compared to the Nxt1 trans-heterozygotes. When measuring the axial ratios and while looking at the pupa viability I also looked at how many pupae had normal spiracles and normal head eversion, respectively. While the actual head eversion occurs 12 hours into the pupa phase, it can be seen most clearly towards the end of the pupa phase (48h-96h) after the pigmentation has developed. Wild type shows no irregular spiracles and no defects in head eversion. For Nxt1 trans-heterozygotes, however, only 50% (N= 36) of the pupae have normal spiracles and only 26% (N= 103) have normal head eversion (Table 3.4). Increasing a*bba* with both UAS and duplication constructs and decreasing *mib2* influenced everting spiracles in Nxt1 trans-heterozygotes (mef2gal4>UAS-*abba*: 93% (N= 40), $abba^{dp}$: 73% (N= 37), $mib2^{df}$:92% (N= 48); Table 3.4). However, none of these manipulations were able to rescue head eversion (mef2gal4>UAS-*abba*: 38% (N= 87), $abba^{dp}$: 21% (N= 28), $mib2^{df}$: 32%;(N= 65); Table 3.4). These data revealed that all these experiments support the gene balance hypothesis. The focus on this experiment was between a gene balance between *abba* and *mib2*, however, it Is interesting that

reducing the expression of a gene, whether it is important for a particular gene balance or not, can increase muscle integrity.

Decreasing abba or increasing mib2

According to the hypothesis that the balance between *abba* and *mib2* is critical for muscle integrity, decreasing *abba* or increasing *mib2* should enhance the Nxt1 trans-heterozygote phenotype. In contrast to our predictions, decreasing *abba* with a deficiency allele was able to partially rescue the Nxt1 muscle phenotype with about 5-20% (N= 11) of the muscles damaged versus 35-80% (N=9) in the Nxt1 trans-heterozygotes (Figure 3.18C). Nxt1 trans-heterozygotes bearing mef2gal4>UAS-*mib2* or *abba* deficiency did not modify the axial ratios (Figure 3.18B) and, however their mobility was significantly increased compared to Nxt1 trans-heterozygotes but were still significantly slower than wild type (Student's t-test * $p<0.05$; Figure 3.18D). Decreasing *abba* expression in Nxt1 trans-heterozygotes showed more pupae were able to evert their spiracles (84% (N= 50); Table 3.4), however, Nxt1 trans-heterozygotes bearing mef2gal4>UAS-*mib2* showed that only 37% (N= 49) of the pupae can evert their spiracles. Decreasing *abba* expression in Nxt1 trans-heterozygotes also showed that more pupae were able to have a normal head eversion (71% (N= 68); Table 3.4). Finally, mef2gal4>UAS-*mib2* in the Nxt1 trans-heterozygotes had a very strong effect on viability: as many pupae died early in the pupa phase (24h-48h). Those pupae that did survive towards the end of the pupa phase had 89% (N= 9) head eversion success (Table 3.4). All the gene manipulations data between *abba* and *mib2* do not generally support the balance hypothesis. The muscle integrity does not only rely between the gene expression balance of *abba* and *mib2*, but potentially many more.

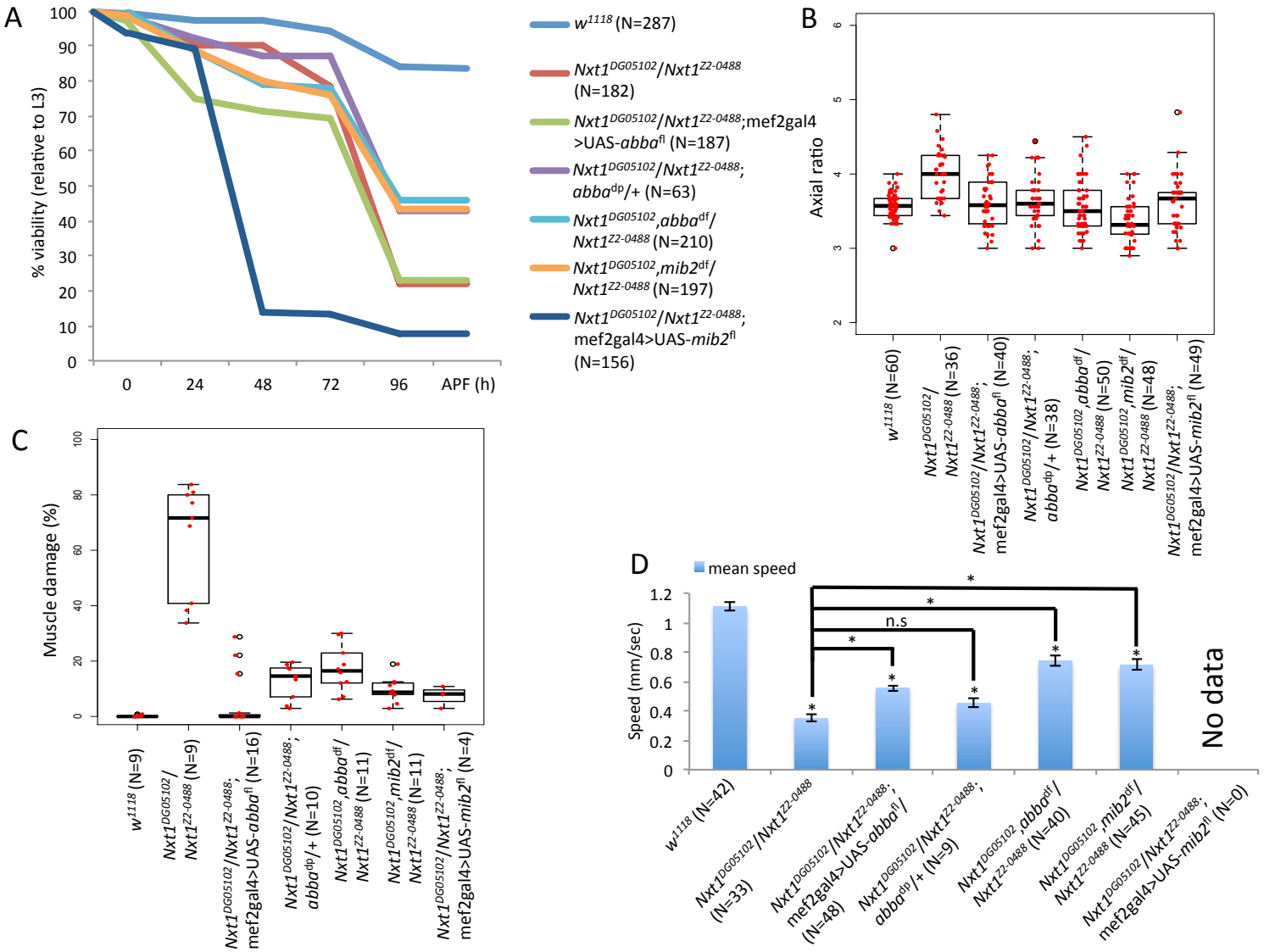


Figure 3.18: **Gain- and loss of function for both *abba* and *mib2* rescues integrity of muscles.** A) Pupa viability assay with various viability between different genotypes. B) Axial ratios between different genotypes. C) Muscle damage observed from eight hemisegments per larva from all genotypes. D) Mobility assay of 3rd instars from all genotypes. B-C) red dots = individual larva/pupa. N= numbers of larvae/pupae examined. Student's t-test * p value = <0.05. n.s = not significant.

3.8 UAS-Nxt1 rescues semi-lethality with a ubiquitous driver, but not with a muscle-specific driver

I increased the expression of Nxt1 via the UAS-gal4 system to rescue the pupa semi-lethality by mef2gal4>UAS-*Nxt1* and armgal4>UAS-*Nxt1* in Nxt1 trans-heterozygotes. At the same time, I also investigated the muscle integrity, axial ratios and larva mobility. Two lines were used, pUAST-eGFP-*Nxt1*-M1 and pUAST-eGFP-*Nxt1*-M2A, to assess the experiments (hereafter referred to a UAS-*Nxt1*-M1 and -M2A). Both lines were generated by PCR and sub cloning of *Nxt1* ORF from $Nxt1^{z2\text{-}0488}$/CyO flies (Parker, Gross, and Alphey 2001; Caporilli et al. 2013). Both lines were inserted on the 2nd chromosome and two independent insertions, M1 and M2A lines, were used for the analysis.

Increasing UAS-Nxt1 with mef2gal4

Both mef2gal4>UAS-*Nxt1*-M1 and mef2gal4>UAS-*Nxt1*-M2A lines were able to rescue muscle integrity in Nxt1 trans-heterozygotes (M1: ~13-26% with two 47% outliers (N= 10), M2A: ~3-35% with one 51% outlier (N= 10); Figure 3.19C). The muscle damage between individual larvae was quite variable, especially for the M2A line (Figure 3.19C). Line M2A had a worse pupa viability (~18% (N= 175)), whereas line M1 had a slightly better pupa viability (~40% (N= 186); Figure 3.19A). On average, the axial ratios were like wild type with some more variability (Figure 3.19B). Finally, mobility was increased compared to Nxt1 trans-heterozygotes, but was still significantly reduced compared to wild type (Student's t-test * $p<0.05$; Figure 3.19D). For both lines, pupae with normal spiracles was increased (M1: 69% (N= 52), M2A: 60% (N= 45); Table 3.4) and similarly for normal head eversion (M1: 46% (N= 52), M2A: 62% (N= 71); Table 3.4). Increasing expression of *Nxt1* with mef2gal4 rescued the muscle degeneration phenotype, however, pupa lethality still occurred at a similar rate in Nxt1 trans-heterozygotes.

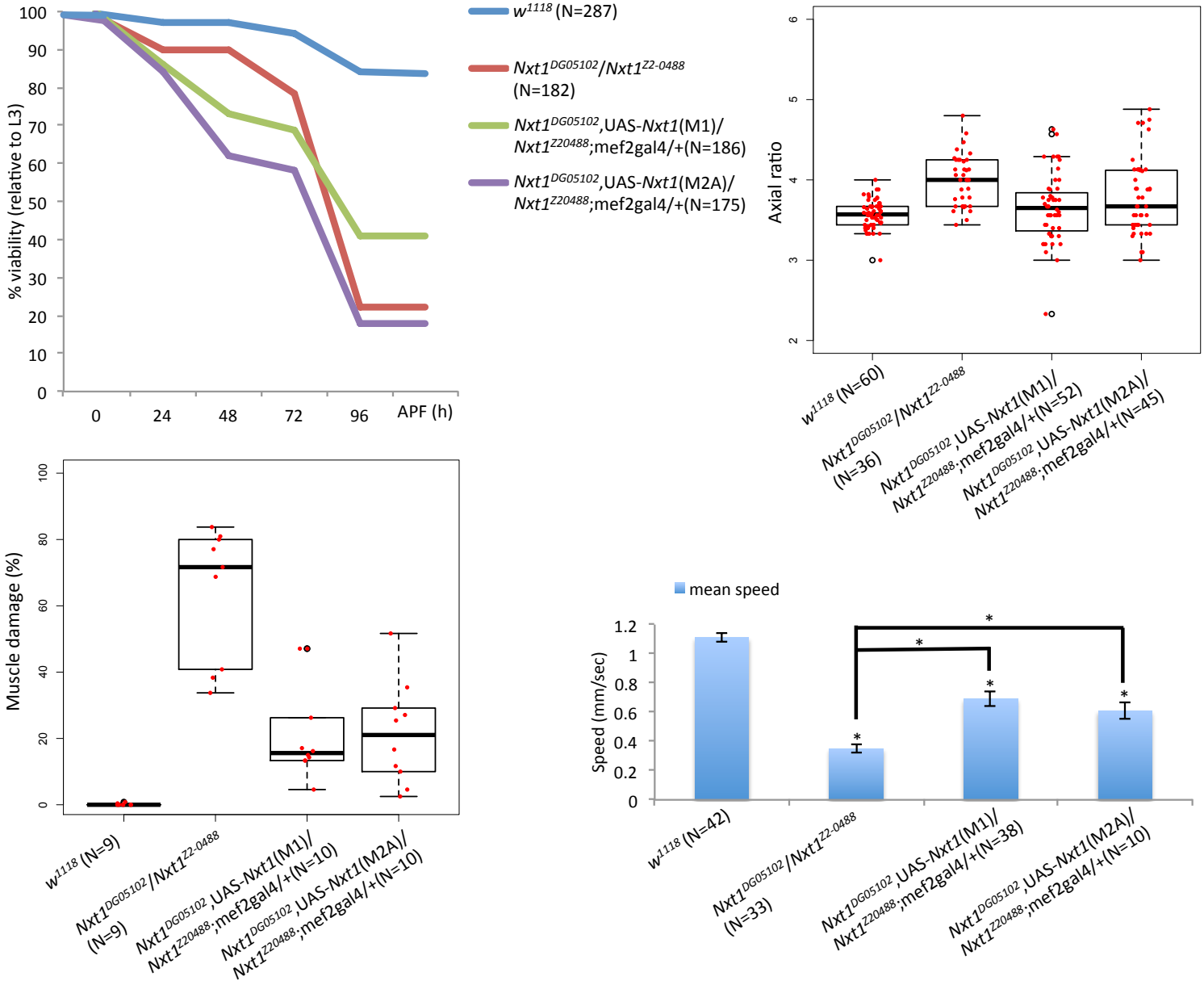


Figure 3.19: **Increasing *Nxt1* with mef2gal4>UAS-*Nxt1* rescues muscle degeneration partially, but not pupa viability.** A) Pupa viability assay with various viability between different genotypes. B) Axial ratios between different genotypes. C) Muscle damage observed from eight hemisegments per larva from all genotypes. D) Mobility assay of 3rd instars from all genotypes. B-C) red dots = individual larva/pupa. N= numbers of larvae/pupae examined. Student's t-test * p value = <0.05. n.s = not significant.

Increasing UAS-Nxt1 with armgal4

Armgal4>UAS-*Nxt1*-M1 and armgal4>UAS-*Nxt1*-M2A in Nxt1 trans-heterozygotes showed similar muscle integrity (M1: ~3-33% with one 73% outlier (N= 9), M2A: ~6-34% with one 84% outlier (N= 10); Figure 3.20C). More interestingly, both lines increased the pupa viability (M1: ~75% (N= 163), M2A: ~70% (N= 188); Figure 3.20A). The axial ratios were like wild type (Figure 3.20B). Line M1 had an increased mobility, whereas line M2A was not significantly different than Nxt1 trans-heterozygotes (Student's t-test * $p<0.05$; Figure 3.20D). More interestingly, none of the examined pupae had uneverted spiracles for either lines and normal head eversion was higher when compared to both mef2gal4>UAS-*Nxt1*-M1 and mef2gal4>UAS-Nxt1-M2A (M1: 57% (N= 23), M2A: 69% (N= 35); Table 3.4). Increasing expression of *Nxt1* with armgal4 rescued the muscle degeneration phenotype, and more interestingly, pupa lethality was mostly rescued in both lines.

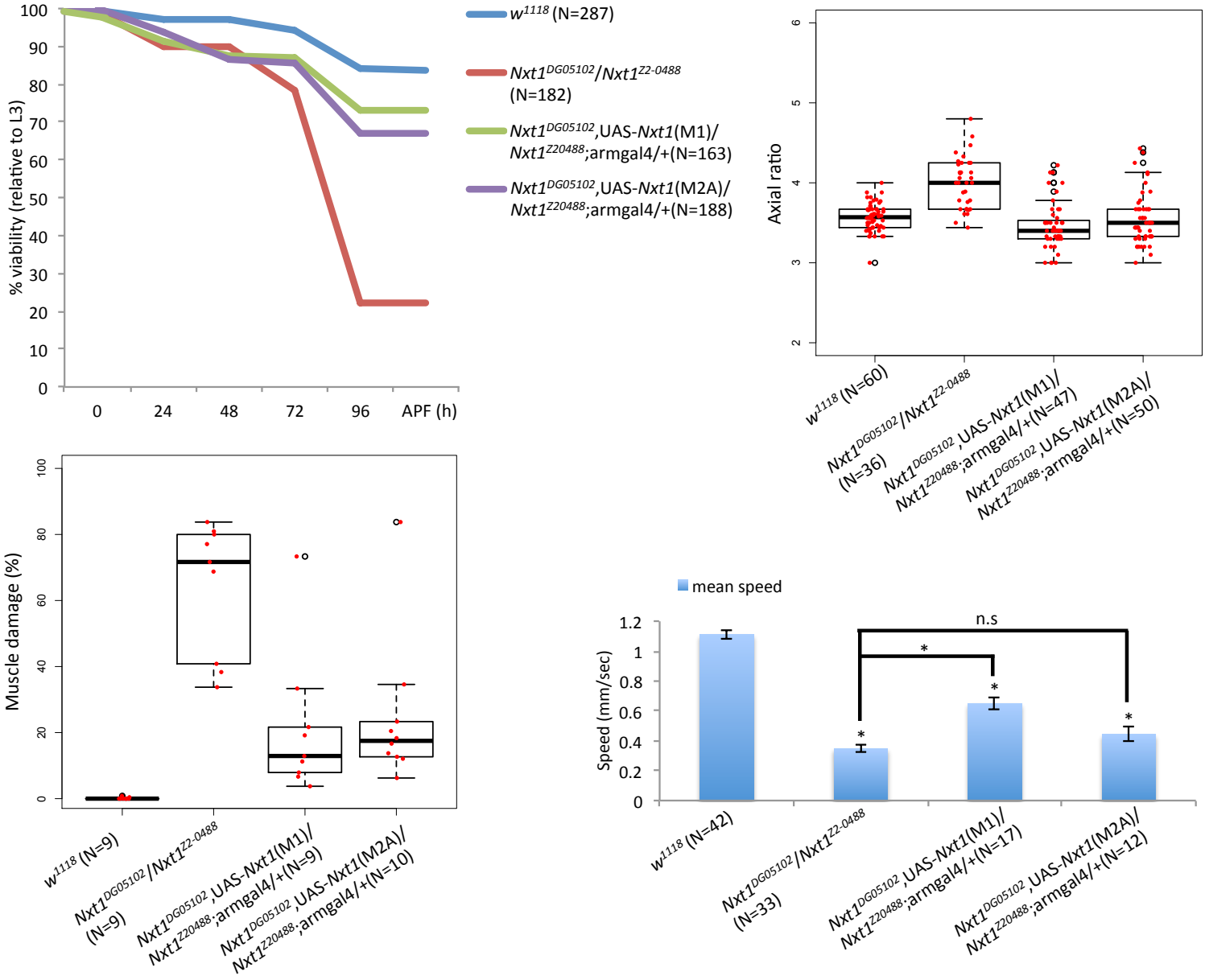


Figure 3.20: **Increasing *Nxt1* with armgal4>UAS-*Nxt1* rescues muscle degeneration partially and significantly increases pupa viability.** A) Pupa viability assay with various viability between different genotypes. B) Axial ratios between different genotypes. C) Muscle damage observed from eight hemisegments per larva from all genotypes. D) Mobility assay of 3rd instars from all genotypes. B-C) red dots = individual larva/pupa. N= numbers of larvae/pupae examined. Student's t-test * p value = <0.05. n.s = not significant.

Table 3.4: **Pupae spiracle and head eversion observation.** For all the different genotypes, pupae were observed for everting their spiracles and for having a normal head eversion.

	Normal spiracles (in %)	**No. of pupae examined (spiracles)**	**Head eversion (in %)**	**72-96h pupae (died at this stage)**	**No. of pupae examined (viability assay)**
W^{1118}	100%	60	100%	29	287
$Nxt1^{DG05102}/Nxt1^{Z2\text{-}0488}$	50%	36	26.21%	103	182
$Nxt1^{DG05102}/Nxt1^{Z2\text{-}0488}$;UAS-$Abba^{fl}$;mef2gal4	92.5%	40	37.93%	87	187
$Nxt1^{DG05102}/Nxt1^{Z2\text{-}0488}$; $Abba^{dp}$/+	72.97%	37	21.43%	28	63
$Nxt1^{DG05102}$,$Abba^{df}/Nxt1^{Z2\text{-}0488}$	84%	50	70.59%	68	210
$Nxt1^{DG05102}$,$Mib2^{df}/Nxt1^{Z2\text{-}0488}$	91.67%	48	32.31%	65	197
$Nxt1^{DG05102}/Nxt1^{Z2\text{-}0488}$;UAS-$mib2^{fl}$/mef2gal4	36.73%	49	88.89%	9	156
$Nxt1^{DG05102}$,UAS-Nxt1(m1)/$Nxt1^{Z2\text{-}0488}$;mef2gal4/+	69.23%	52	46.15%	52	186
$Nxt1^{DG05102}$,UAS-Nxt1(m2a)/$Nxt1^{Z2\text{-}0488}$;mef2gal4/+	60%	45	61.98%	71	175
$Nxt1^{DG05102}$,UAS-Nxt1(m1)/$Nxt1^{Z2\text{-}0488}$;armgal4/+	100%	47	56.52%	23	163
$Nxt1^{DG05102}$,UAS-Nxt1(m2a)/$Nxt1^{Z2\text{-}0488}$;armgal4/+	100%	50	68.57%	35	188

3.9 Gene with many and long introns are more sensitive in Nxt1 trans-heterozygotes

The expression balance between *abba* and *mib2* did not provide insight in muscle integrity. Nxt1 is important for mRNA exportation and is recruited indirectly by EJC to spliced transcripts (Le Hir et al. 2001). Mutations in the EJC can affect transcripts with long introns and components of the EJC were mildly down regulated in Nxt1 trans-heterozygotes. *abba* is a long transcript with introns up to 4 kb in length. I went back to the whole larva, mixed sex, RNA sequencing and analysed genes with introns. A python script (provided by Rachel Lyne; (Lyne et al. 2007)) identified genes with introns and calculated total intron length of all isoforms of each gene, the unique number of introns of all isoforms of each gene, and the smallest/longest intron of each gene. The lists were separated into non-differentially, 1.5-, 2-, 4- and 16-up and down regulated genes. Upon analysis, the more down regulated the genes were, the shorter the total intron length (median decrease of 718nt to 120nt) and the lower the total number of introns was (median of 4 to 2; Table 3.5). This decline was very noticeable of genes more than 16-fold down regulated. Up regulated genes did not show any trend with regards to intron length or number (Table 3.5).

In Nxt1 trans-heterozygote testis, short and intron-less genes were particularly sensitive to the lack of *Nxt1* function (Caporilli et al. 2013). I used a testis-specific and testis-enriched genes list (provided by Helen White-Cooper) to remove these genes before RNA sequencing analyses as I did not sex the larvae, which would interfere with analysing the data. However, when the data was analysed more in depth I noticed that the remaining down regulated genes were still highly testis-enriched. To overcome this problem and generate a data set that documents gene expression in the tissue in which the major defect is detected (i.e. muscle) I performed new RNA sequencing from third instar larval carcasses. This sample was made from a mixed-sex population, but the dissection removes the gonads, and thus avoids the problems associated with highly expressed testis-specific transcripts. Each genotype (w^{1118} and Nxt1 trans-heterozygotes) was performed in triplicates. Each library achieved a read depth between 21M to 32M reads and the data was analysed through the Tuxedo suite

(Trapnell et al. 2012). The list of genes with FPKM values provided by Cuffdiff was used to separate the lists in more than 1.5-, 2-, 4- and 16-fold down regulated, more than 1.5-, 2-, 4- and 16-fold up regulated and non-differentially expressed. The differentially expressed genes were assigned statistically significant and had a p-value of <0.05. With the new data set I analysed again the total intron length, number of introns and smallest/largest intron (Table 3.6). This revealed that all four categories of down regulated genes had more introns (median increase of 9 to 11 introns) and a higher total intron length (median increase of ~5.7kb to ~15kb), while all four categories of up regulated genes had less introns (median decrease of 3 to 2 introns) and a lower total intron length (median decrease of ~370bp to ~240bp) than the genes that were not differentially expressed (median of ~500bp and 4 introns; Mann-Whitney p<0.05; Table 3.6). I extended the analysis by also looking at the gene length, number of transcripts and shortest/longest transcripts (Table 3.7). This was performed with an adapted version of the python script to calculate gene length (with introns), total number of transcripts of each gene, and shortest/longest transcript of each gene (Rachel Lyne; (Lyne et al. 2007). Again, a clear trend showed that all four categories of down regulated genes had longer gene length (median increase of ~8.4kb to ~14.2kb) and all four categories of up regulated genes had shorter gene length (median decrease ~1.5kb to ~1.4kb) when compared to non-differentially expressed genes (median of ~2.4kb; Mann-Whitney p<0.05; Table 3.7). A full list with statistical testing for all the features is shown in supplementary table 1. qRT-PCR was performed for 12 genes with high total intron length to confirm the sequencing data (Figure 3.21). Indeed, these genes were significantly down in Nxt1 trans-heterozygotes. *abba*, with a total intron length of approximately 41,000bp, was down regulated as expected and *Nxt1*, with a total intron length of 184bp, was only mildly affected (Figure 3.21). Data analysis has confirmed that down-regulated genes in Nxt1 trans-heterozygotes were mostly genes with a high total intron length and had 9 or more introns.

Table 3.5: **Total intron length observation for genes down and up regulated from whole larvae sequencing.** No noticeable trend visible in total intron length from genes up or down regulated. Similar observation for number of introns and smallest/largest introns.

Median	No. of genes	Total intron length	Number of introns	Smallest intron	Largest intron
Down-regulated >1.5 fold	653	718	4	59	321
Down-regulated >2 fold	404	374	3	58	152
Down-regulated >4 fold	215	219	3	57	72
Down-regulated >16 fold	42	120	2	59	65
Median	**No. of genes**	**Total intron length**	**Number of introns**	**Smallest intron**	**Largest intron**
Up-regulated >1.5 fold	406	463	3	62	209
Up-regulated >2 fold	274	307	3	61	123
Up-regulated >4 fold	106	294	3	61	110
Up-regulated >16 fold	10	676	4	59	308
Median	**No. of genes**	**Total intron length**	**Number of introns**	**Smallest intron**	**Largest intron**
Non-differentiated	9995	604	4	60	287

Table 3.6: **Total intron length observation for genes down and up regulated from larval carcass sequencing.** Genes down regulated have longer total intron length compared to up regulated and non-differentially expressed genes. The total intron length becomes longer with the more highly down regulated genes. Down regulated genes have more introns and their largest intron is significantly longer. Up regulated genes have a lower total intron length compared to non-differentially expressed genes. For these, a reduction in number of introns is observed the more a gene is up regulated.

Median	No. of genes	Total intron length	Number of introns	Smallest intron	Largest intron
Down-regulated >1.5 fold	1,821	5,769	9	59	2,341
Down-regulated >2 fold	1,340	7,594	9	59	2,931
Down-regulated >4 fold	567	10,201	11	59	3,642
Down-regulated >16 fold	32	15,254	10	63	8,081
Median	**No. of genes**	**Total intron length**	**Number of introns**	**Smallest intron**	**Largest intron**
Up-regulated >1.5 fold	1,339	368	3	62	160
Up-regulated >2 fold	572	352	3	62	155
Up-regulated >4 fold	89	237	2	62	133
Up-regulated >16 fold	15	238	2	61	180
Median	**No. of genes**	**Total intron length**	**Number of introns**	**Smallest intron**	**Largest intron**
Non-differentiated	4,754	497	4	60	228

Table 3.7: **Gene length observation from down and up regulated genes.** Genes that are down regulated also have a longer gene length compared to up regulated and non-differentially expressed genes. For down regulated genes, the number of transcripts is higher and their shortest/longest transcripts are larger. Up regulated genes have a reduction in gene length and number of transcripts compared to non-differentially expressed genes.

Median	No. of genes	Gene length	No. of transcripts	Shortest transcript	Longest transcript
Down-regulated >1.5 fold	1,821	8,440	3	3,220	4,263
Down-regulated >2 fold	1,340	9,706	3	3,301	4,533
Down-regulated >4 fold	567	12,162	3	3,722	5,168
Down-regulated >16 fold	32	14,213	4	3,486	5,761
Median	**No. of genes**	**Gene length**	**No. of transcripts**	**Shortest transcript**	**Longest transcript**
Up-regulated >1.5 fold	1,339	1,584	2	1.034	1,186
Up-regulated >2 fold	572	1,411	2	901	1,027
Up-regulated >4 fold	89	1,353	1	885	994
Up-regulated >16 fold	15	1,411	1	917	1,020
Median	**No. of genes**	**Gene length**	**No. of transcripts**	**Shortest transcript**	**Longest transcript**
Non-differentiated	4,754	2,496	2	1,642	1,856

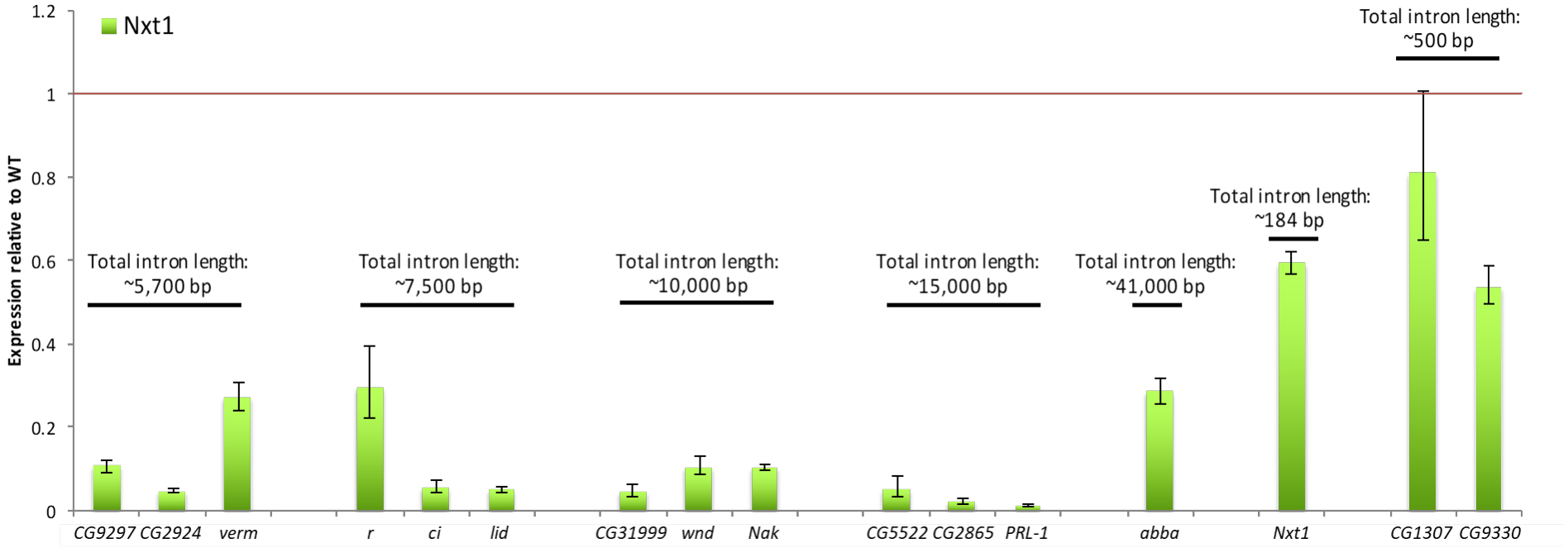


Figure 3.21: **Genes with high total intron length are significantly down regulated in Nxt1 trans-heterozygotes.** Genes with an average total intron length of 5700, 7500, 10000 and 15000 were examined. Expression of *abba* and *Nxt1* are also shown. *CG1307* and *CG9330* were used as control genes with low total intron length. All genes examined have expression in the carcass. *Rp49* was used for normalization.

3.10 *abba* non-differentially expressed in larval carcass RNA-seq

Larval carcass RNA-seq was used to investigate the intron sizes of down regulated genes. I used this RNA-seq to look at the gene ontology enrichment and larval muscle genes. In total, 1968 genes were more than 1.5-fold down regulated in Nxt1 trans-heterozygotes larval carcasses. These genes were put into FlyMine, which recognized 1924 genes and used that for analysis. For Gene Ontology enrichment analysis, FlyMine identified 1527 genes and used the Holm-Bonferroni test correction, with a max p-value of 0.05 to look at the top 10 biological processes (Table 3.8). Ontology enrichment was found in development, such as system (678 counts), anatomical structure (609 counts), multicellular organism (771 counts) and animal organ (504 counts) or processes, such as cellular development (638 counts), single-organism developmental (861 counts) and single-multicellular organism (843 counts; Table 3.8). Similarly, 1581 genes were more than 1.5-fold up regulated and FlyMine identified 1574 genes for analysis. FlyMine identified 1163 genes for Gene Ontology enrichment and used Holm-Bonferroni test correction, with a max p-value of 0.05 looking at the top 10 biological processes (Table 3.9). Nine out of 10 processes had less than 100 counts each and 8 out of 10 processes were enriched in different metabolic processes (63-65 counts). GO enrichment showed that genes down regulated in Nxt1 trans-heterozygotes larval carcasses had an effect in the development process, whereas genes up regulated had an increase in metabolic processes.

Table 3.8: **Top 10 Gene Ontology Enrichment of genes more than 1.5-fold down regulated in larval carcass RNA-seq**. 1527 genes were analysed for gene ontology enrichment, using Holm-Bonferroni as test correction with a p-value of <0.05.

GO Term	Counts	Test Correction	P-value
System development	678	Holm-Bonferroni	1.440756e-138
Anatomical structure morphogenesis	609	Holm-Bonferroni	5.249362e-121
Single-multicellular organism process	843	Holm-Bonferroni	1.046095e-120
Multicellular organism development	771	Holm-Bonferroni	8.755146e-119
Anatomical structure development	833	Holm-Bonferroni	1.296775e-116
Single-organism developmental process	861	Holm-Bonferroni	3.351494e-116
Developmental process	870	Holm-Bonferroni	1.750122e-115
Animal organ development	504	Holm-Bonferroni	3.293350e-109
Cell differentiation	629	Holm-Bonferroni	5.931754e-103
Cellular developmental process	638	Holm-Bonferroni	1.113612e-102

Table 3.9: **Top 10 Gene Ontology Enrichment of genes more than 1.5-fold up regulated in larval carcass RNA-seq**. 1163 genes were analysed for gene ontology enrichment, using Holm-Bonferroni as test correction with a p-value of <0.05.

GO Term	Counts	Test Correction	P-value
Cytoplasmic translation	70	Holm-Bonferroni	4.424644e-32
Oxidation-reduction process	169	Holm-Bonferroni	5.350025e-18
ATP metabolic process	62	Holm-Bonferroni	3.714191e-16
Nucleoside triphosphate metabolic process	64	Holm-Bonferroni	8.758239e-16
Purine nucleoside triphosphate metabolic process	63	Holm-Bonferroni	1.018484e-15
Purine ribonucleoside triphosphate metabolic process	63	Holm-Bonferroni	1.018484e-15
Ribonucleoside triphosphate metabolic process	63	Holm-Bonferroni	1.474921e-15
Purine nucleoside monophosphate metabolic process	65	Holm-Bonferroni	3.095007e-15
Purine ribonucleoside monophosphate metabolic process	65	Holm-Bonferroni	3.095007e-15
Ribonucleoside monophosphate metabolic process	65	Holm-Bonferroni	4.532684e-14

With the whole larvae RNA-seq data, I examined 13 different muscle genes where *abba* was the only gene down regulated in that list. This time, I compared the same 13 genes with the new larval carcass RNA-seq data to identify similarities and differences in gene expression (Table 3.10). The expression profile was different in larval carcass than whole larvae. Most noticeably, *abba* in larval carcass is not differentially expressed, whereas it is down regulated in whole larvae. This speculates that *abba* has roles in other tissues rather than larval muscles only. This could also explain why armgal4>UAS-*abba*^{fl} in Nxt1 trans-heterozygotes rescues pupa lethality at a 70% success rate. Aside the difference in *abba* expression, *Up, Mhc, Act57B, Actn* and *wupA* were not differentially expressed in larval carcass, while all these genes were up regulated in whole larvae. *Sls*, which was not differentially expressed in whole larvae is more than 2-fold down regulated in larval carcass (Table 3.10). Sls, also known as Kettin, links myosin thick filaments to the Z-line structure and is required for the maintenance of normal sarcomere structures (Hakeda, Endo, and Saigo 2000). This analysis revealed that Abba might have roles in other tissues but increase of *abba* in whole larvae does rescue the muscle phenotype in Nxt1 trans-heterozygotes, even if *abba* is not differentially expressed. *Sls* is down regulated in larval carcass and its role in linking thick filaments to the Z-line structure could potentially affect the sarcomeric integrity.

Table 3.10: **Comparison of 13 muscle genes between whole larvae and larval carcass RNA-seq**. RNA-seq data were used to look at 13 different muscle genes. Red colour indicates more than 1.5-fold down regulated and green colour indicates more than 1.5-fold up regulated. Not sig. indicates genes had p-value of >0.05.

Gene	Whole larvae (log2 fold change)	Larval carcass (log2 fold change)
abba	-0.75	+0.05
up	+1.54	+0.69 (not sig.)
Arc1	+1.84	+5.88
Prm	+1.62	+2.14
Mlc2	+2.16	+1.61
Mhc	+3.02	+0.28
Act57B	+2.04	+0.27
mib2	-0.48	-0.54 (not sig.)
Mlp84B	+0.88	+0.83
Actn	+0.33	+0.17
Mp20	+2.08	+2.54
wupA	+0.90	+0.002
sls	+0.16	-1.18

Chapter 3 - Discussion

3.11 Reduction in Nxt1 function causes muscle degeneration

Nxt1 trans-heterozygotes pupae have a high semi-lethality with a clear curved shape. I tried to understand why reducing Nxt1 during the fly development results in a high pupa semi-lethality. Nxt1 is mostly known for its role in the RNA export pathway. Nxt1 binds to Nxf1, which interacts with the nuclear pore complex for exporting mRNA to the cytoplasm (Herold, Klymenko, and Izaurralde 2001). Reducing Nxt1 reduces the Nxt1-Nxf1 dimer and without Nxt1, Nxf1 can interact less effectively with the nuclear pore complex (Wiegand et al. 2002). Since most mRNAs are exported via the Nxt1-Nxf1 route, it is no surprise that many genes will be affected. However, the only clear phenotypes that were observed was the air bubble defect, muscle degeneration and the curved pupa shape. Not all Nxt1 trans-heterozygote pupae have the air bubble defect or a curved shape and these, more likely, have a higher survivability, since only 20% survives through adulthood (Figure 3.4). Pupae that have either one of the phenotypes or both generally do not survive at all. With this part of the project I focused on understanding the air bubble phenotype.

3.11.1 No defects in ecdysone signaling in *Nxt1* mutant larvae

A defect in air bubble movement and subsequent head eversion has been observed in mutants of trf2, a TATA-box binding protein paralogue, and Eip74EF, an ets family transcription factor (Bashirullah et al. 2007; Fletcher and Thummel 1995). This let to our hypothesis that ecdysone and ecdysone-responsive genes were responsible for the air bubble defect. I chose to use RNA-seq to give an unbiased view of gene expression at the wandering larvae, stationary larvae and white prepupae. The hypothesis predicted that ecdysone-responsive genes would be under-expressed, however, RNA-seq data revealed that this was generally not the case. I looked at the genes that were responsible for the production of ecdysone. Only one gene, *Spok*, was mildly down regulated. *Spok* has been speculated to be involved in the synthesis in the so-called “blackbox” between the 7-dehydrocholesterol and Δ4-diketol in the early ecdysone synthesis (Ono et al. 2006). This could potentially mean that the ecdysone synthesis is mildly affected. However, most ecdysone-responsive genes were

not affected which indicated that the produced ecdysone was enough for regulating the ecdysone-responsive genes. I also looked at all the known ecdysone-responsive genes but noticed that only four genes were mildly down regulated (Table 3.1). *71Ea*, *71Ef* and *71Eg* are within the Ecdysone Induced Genes (Eig) cluster of five divergently transcribed pairs of genes with short intergenic regions (~260bp) are part of the 71E late puff (Wright et al. 1996). Many of these late genes (*L71* genes) encode for proteins related to mammalian defensins that may function as secreted antimicrobial polypeptides involved in host defense during metamorphosis (Wright et al. 1996). However, these genes exhibit late developmental expression, starting from 2-4 hours into the pupa phase, which is after the white prepupa stage (Wright et al. 1996). RNA-seq data also revealed that *L71* genes were low in expression (data not shown) and several not expressed during wandering and stationary larvae (Table 3.1). Eip63F-1 was the fourth gene down regulated and deficiency analysis has shown that both Eip63F-1 and Eip63F-2 were not essential for viability or fertility (Vaskova et al. 2000). I concluded that neither a reduction in ecdysone production nor response in Nxt1 trans-heterozygotes cause the air bubble defect.

3.11.2 Degeneration of larval muscles affect air bubble movement

An alternative hypothesis that could explain the air bubble movement phenotype is that the larval muscles are defective, since the air bubble is moved by muscle contractions. Mobility analysis showed that 3rd instar Nxt1 trans-heterozygote larvae crawled significantly slower than wild type (Figure 3.6). Interestingly, Nxt1 trans-heterozygotes 2nd instar had the same mobility as 3rd instars. This showed that during the final growth muscles were growing, but defects arise during this growth. Staining 3rd instar larvae with phalloidin showed a clear degeneration of the muscles (Figure 3.7). Most of the muscles were thinner than wild type, which explains why the air bubble cannot move all the way to the anterior since the muscles presumably have less strength. Both Nxt1 RNAi lines also showed thinner muscles in 3rd instars when driven in the muscles, confirming that the phenotype is due to reduction of *Nxt1* function. However, the muscle degeneration with the RNAi lines was apparent in early 2nd instars when the RNAi induction was during embryonic development (Figure 3.9). It is not surprising that the RNAi has a greater effect since Nxt1 could be more efficiently depleted, and genetic experiments have confirmed that Nxt1 trans-heterozygotes still have some

functional protein (Caporilli et al. 2013). I tried to understand why the muscles degenerate during the end of the larva phase. A classic paper from the 1930s showed that larvae, who were deprived from food 70h AEL, survived through adulthood, but did not increase in size (Beadle, Tatum, and Clancy 1938). I tried to replicate this experiment, since 70h AEL, larvae are in their late 2nd instar – early 3rd instar phase. Nxt1 trans-heterozygotes had no muscle degeneration at that point, so this experiment would reveal if the degeneration was caused by growth or by something else, for example muscle use or atrophy over time. The experiment was not fully conclusive since wild type larvae which stopped feeding 70h, 71h, 72h and 73h AEL mostly died, and only a small percentage pupated; few adults emerged (Figure 3.11). However, many Nxt1 trans-heterozygotes larvae survived up to several days before dying. This permitted the examination of the muscle phenotype in Nxt1 trans-heterozygous larvae 70h AEL and 70h AEL + 4 days without any food. This showed that the muscles were still intact in both cases. With these data, I could confirm that the muscle degeneration occurred during the final growth in the larva phase.

The Insulin/Akt/TOR pathway is involved in the nutrient supply for the growth of the larval muscles. Since the muscle degeneration occurs during the final growth I used the RNA-seq data to examine expression of several genes involved in this pathway and performed qRT-PCR (Figure 3.12). RNA-seq did not indicate a role for Nxt1 in regulation of transcript levels for genes in this pathway and qRT-PCR only showed that *FOXO* was down in wandering larvae and *CG11658* was down in white prepupae. A reduction in *FOXO* would result into less induction of the Atrogenes that are required for the muscle atrophy process. Similarly, *CG11658* is required for the proteasomal proteolysis, and a reduction would again indicate less induction for muscle atrophy. The most interesting part would be *FOXO* and *CG11658* during the stationary larvae phase. The stationary larvae phase in Nxt1 trans-heterozygotes showed muscle degeneration. Both *FOXO* and *CG11658* seem to have a great variability in expression (Figure 3.12B). An increase in *FOXO* and *CG11658* would mean that muscle atrophy is induced at a potential higher level that would break down the muscles and therefor show thinner muscles. However, it is technically very challenging to accurately stage the larvae to a particular time point. All larvae have a slight difference in development time and some can develop several hours faster than others. I only staged the larvae into wandering and stationary phase. The wandering phase is roughly 6-24h BPF, whereas stationary phase is

between 1-6h BFP. Expression variability is likely to be a consequence of this staging issue. Therefor it is hard to tell if genes like *FOXO* and *CG11658* were expressed normally or potentially up regulated. For RNA-seq data I tried to compensate for the variability by using 30 larvae for each replicate to get an average expression. Similarly, big changes in variability were observed when examining expression levels for several muscle genes in individual larva with qRT-PCR (data not shown). Altogether, expression variability should be kept in mind when analysing the data and qRT-PCR should be performed for double confirmation.

3.11.3 Molecular basis of curved pupa shape remains unclear

One of the phenotypes that I did not focused on was the curved pupa shape. This phenotype has been observed in other mutants such as *Broad-Complex* and *E74* (Fletcher and Thummel 1995), both induced by ecdysone, but the molecular basis remained unclear. Another gene called *Fondue* has been found that might be required for the sclerotization and hardening of the pupal case (Scherfer et al. 2006). I looked at the pupal case with scanning electron microscopy and found no obvious differences in the case pattern (Figure 3.2C-E). Also, the sclerotization of the pupae seems normal in terms of the colour of the pupae. It remains unclear what exactly is causing the mis-shaping of the pupae. It is possible that when a larva tries to shorten as they pupate, and the muscles on one side being more damaged than those on the other side, the larva shortens on one side only and therefore ends up both long and curved. Degeneration of muscles can therefore lead to this asymmetry.

3.12 Muscle degeneration can be rescued by increasing *abba*

I used the RNA-seq data to identify potential candidates to understand the muscle degeneration. A list of muscle genes involved in the larval muscles was used (Table 3.3). From all the candidates only one gene had an interest. *abba* was slightly down regulated, and qRT-PCR data confirmed that *abba* was indeed down regulated in 9 out of 10 stationary larvae (Figure 3.13B). Abba, is a TRIM/RBCC protein involved in maintaining the integrity of sarcomeric cytoarchitecture (Domsch, Ezzeddine, and Nguyen 2013). Abba mutant, formally known as *thin* (*tn*), has a muscle phenotype with thinner muscles. Nxt1 trans-heterozygotes had thinner muscles due to the muscle degeneration so I decided to get further into Abba as the candidate.

3.12.1 Increasing *abba* rescues muscle degeneration, but not pupa semi-lethality

I used the UAS/gal4 system to understand if Abba is involved in the muscle degeneration in Nxt1 trans-heterozygotes. The UAS-*abba*^{fl} construct was kindly gifted by Hanh T. Nguyen and contained the entire *abba* coding region (isoform C) from cDNA source, cloned into the pUAST vector. An HA epitope tag was added at the 3' end of the coding sequence, to tag the protein at the C-terminus. I crossed flies carrying this construct into Nxt1 trans-heterozygotes background and expressed the construct with a muscle specific driver, mef2gal4. Surprisingly, muscle phenotype was rescued, and qRT-PCR confirmed that the larvae had increased *abba* expression (Figure 3.13C and D). It is interesting how a pleiotropic factor such as Nxt1 leads to a specific muscle phenotype that can be rescued by increasing expression of *abba*. A similar phenomenon has been observed before where the lethality of *flapwing* (*flw*), a protein involved in dephosphorylation of proteins, was rescued by reducing the gene dosage of non-muscle myosin II heavy chain *zipper* (Vereshchagina et al. 2004). The non-muscle form of myosin regulatory light chain *spaghetti squash* (*sqh*), is a target of the enzyme, and lethality was due to mis-regulation of a phosphorylation event (Vereshchagina et al. 2004).

The pupa semi-lethality was not rescued by increasing *abba* in the muscles. This is consistent with Nxt1 being required in other tissues in addition to muscle, since Abba is involved in the muscles predominantly and no other known processes. However, armgal4>UAS-*abba*^{fl} in Nxt1 trans-heterozygotes showed an increase in pupa viability from ~20% to ~70% (Figure 3.14A). This indicates that up regulation of *abba* in other tissues allows for an increase in pupa viability. Abba is known for maintaining muscle integrity in larval Z-discs (Domsch, Ezzeddine, and Nguyen 2013). For future experiments, expression of *abba* should be measured in other tissues to understand where in the larva Abba has a potential role. This could then be used to study Abba in said tissue(s) to understand the larva survivability.

3.12.2 Gene dosage between *abba* and *mib2* cannot explain muscle integrity

To understand how one gene can rescue the muscle phenotype I looked further into the physical and genetic interactions (Figure 3.16). *mib2* is the only physical interactor with *abba* and I speculated that the ratio of expression and activity of these the two genes could be required for muscle integrity. Mib2 maintains the integrity of fully differentiated muscles and

prevents apoptotic degenerations (Nguyen et al. 2007). RNA-seq data showed that *mib2* was not differentially expressed and I proposed a hypothesis about the muscle integrity between *abba* and *mib2* via a gene dosage effect (Figure 3.17). However, UAS-*abba*^{fl}, UAS-*mib2*, *abba*^{dp}, *mib2*^{df} and *abba*^{df} all showed a rescue in the muscle phenotype. UAS-*mib2* was slightly different, since 3rd instar larvae seemed slightly shorter than wild type. If the larvae do not grow as large as intended, then it might influence the muscle degeneration. Our data has shown that the final growth between 2nd and 3rd instar induces muscle degeneration. With these data, both increasing and decreasing *abba* has a rescue in muscle degeneration making it more complicated to understand the muscle integrity. Similarly, decreasing *mib2* also partly rescues muscle integrity. All four constructs used for this experiment showed a rescue in muscle integrity, while only two were predicted by the model to rescue. Indeed, I predicted that reduction in *abba* dose (*abba*^{df}) should enhance the phenotype, but instead this also suppressed. These data shows that the muscle integrity is a lot more complicated than just the gene dosage between two genes.

Both *abba* and *mib2* are expressed in the Z-line in muscles. The Z-line is part of the sarcomere where thin filaments extend to the center to each Z-line. The Z-line is a three-dimensional electron dense structure that separates neighboring sarcomeres and serves to anchor and organize thin filaments and to mechanically link thin filaments from one sarcomere to the adjacent sarcomere (Clark et al. 2002). Titan and nebuline/nebulette overlap within and form important parts of the Z-line (Clark et al. 2002). Titan spans half sarcomeres between the M-band and Z-line, whereas Nebulin runs along the thin filament and also forms the template of thin filament assembly (McElhinny et al. 2003). Titan Z-repeats domain spans the width of the Z-line and it is believed that two Z-repeats could be the template for the Z-line (Gregorio et al. 1998; Atkinson et al. 2000). Nebulin plays an important role in the assembly, structure and function of the Z-line in skeletal muscle (McElhinny et al. 2003). Another protein, CapZ, caps the barbed end of the actin filaments and interacts with α-actinin and nebulin (Papa et al. 1999; Pappas et al. 2008). The Z-line houses many more proteins and have involvement with stretch sensing and signaling (Epstein and Davis 2003; Pyle and Solaro 2004). Stretch sensors are required for muscle fibers, which respond to stretch. The identity and mechanism of stretch sensors are not known but are believed to be located in the Z-line and M-band (Knoll et al. 2002). There are many proteins in the Z-line and mutations of many of these proteins

are associated with defects in signaling and with diseases (Clark et al. 2002; Faulkner, Lanfranchi, and Valle 2001; Frank et al. 2006; Pyle and Solaro 2004; Sheikh et al. 2007). For diseases, mutations of Cypher, a protein maintaining cytoskeletal structural integrity during contraction, in human leads to cardiomyopathy, hypertrophic cardiomyopathy and skeletal muscle myopathies (Sheikh et al. 2007). Mutations in telethonin, a substrate of titin for sarcomere assembly, lead to autosomal recessive limb-girdle muscular dystrophy type 2G (Moreira et al. 2000). Like telethonin, a missense mutation in myotilin lead to autosomal dominant limb-girdle muscular dystrophy type 1A (Hauser et al. 2000). The *abba* mutant phenotype shares similarities with the phenotype by TRIM32-deficient mice and limb-girdle muscular dystrophy type 2H patients with mutations in the *TRIM32* gene (Kudryashova et al. 2009).

In larval carcasses, RNA-seq data suggested that *abba* was not differentially expressed, however, by qRT-PCR, I found that *abba* is slightly down regulated in larval carcasses (Figure 3.21). With the current data, it looks that Abba alone is responsible for the muscle integrity in larval muscles in Nxt1 trans-heterozygotes. However, larval carcass RNA-seq in Nxt1 trans-heterozygotes revealed that genes with several and/or long total intron length were more sensitive towards the loss of Nxt1 (Table 3.6). These genes also had larger transcripts (Table 3.7). However, there might be a possibility that other, possibly unknown, muscles genes are affected by the loss of Nxt1 and affect the muscle integrity. To fully understand the muscle integrity in Nxt1 trans-heterozygotes, a more extensive search to muscles genes, particularly proteins known to localize to the Z-line, should be examined for their expression by qRT-PCR and role in the larval muscles in the mutant background.

Interestingly, Nxf4, one of the four Nxf proteins in *Drosophila* shows a strong muscle defect when RNAi against it is driven by mef2gal4 (Schnorrer et al. 2010). In adult IFM, Nxf4 RNAi showed a “fuzzy z”, broadening of the Z-line, phenotype (Schnorrer et al. 2010). Nxf4 has a high expression in the testis, and its missing NTFL2-like domain makes it unlikely to bind to Nxt1. It would be interesting to understand the role of Nxf4 in the muscles and if Nxt1 is involved with the observed phenotype.

3.12.3 Is *Nxt1* particularly important for expression of transcripts from genes with long introns?

abba is a long transcription unit that contains several long introns with a total intron length of ~41.000bp. Transcripts with many and large introns could be more vulnerable towards the loss of Nxt1, based on the finding that genes with large introns were sensitive to the loss of the EJC (Ashton-Beaucage and Therrien 2011). Interestingly, in testes, genes without introns were particularly sensitive to loss of Nxt1 (Caporilli et al. 2013). The four genes that compose the EJC were very mildly reduced in Nxt1 trans-heterozygotes (Figure 3.15C). With the RNA-seq data from whole larvae I used the down and up regulated genes and looked at the total intron length for each gene (Table 3.5). The data did not reveal any differential expression of genes with long or short introns. The RNA-seq data set came from whole larvae, which were not sexed. Most down-regulated genes were highly testis-enriched. This bias towards the testis interfered with the intron analysis, since this would not help understanding the semi-lethality during the pupa phase. Therefore, I repeated the experiment again by using new RNA-seq data from larval carcasses to have no interference with highly expressed/testis-specific genes and specifically look at the carcass, which includes the larval muscles. This time, a significant trend was observed for both down and up regulated genes (Mann-Whitney test p-value <0.05, except for 16-fold up regulated (p –value of 0.16)). For down regulated genes, the more dramatic the down regulation, the longer the total intron length. Similar for up regulated genes, the more up regulated the genes, the shorter the total intron length (Table 3.6). How can down regulated genes with high total intron length be more sensitive towards the loss of Nxt1 with no big disruption in the EJC? Nxt1 loading partly depends on the recruitment of EJC (Le Hir et al. 2001). If EJC loading is less efficient on transcripts that have many introns and a long total intron length, then they might be extra sensitive in the Nxt1 trans-heterozygotes where the availability of Nxt1 protein is limiting. To better understand this, RNA immunoprecipitation (RIP) sequencing can be performed using antibodies to the EJC proteins (Cloonan et al. 2008). With this method, RNA-protein complexes are immunoprecipitated with antibodies targeted to the protein of interest. RNA is digested, and RNA protected by the protein binding is extracted and reverse-transcribed to cDNA. Using this method for wild type and Nxt1 trans-heterozygotes larval carcasses allows to understand if

EJC loading is affected of transcripts with many introns and long total intron length in Nxt1 trans-heterozygotes.

Are the introns in *abba* spliced? When I compared nascent with spliced transcript expression at four different regions I found that there was less spliced transcript in whole larva and in larval carcasses (Figure 3.15B and data not shown). There have been reports that spliced transcripts require distinct factors for rapid and efficient export that is not observed with unspliced transcripts and are exported faster to the cytoplasm than their unspliced counterparts (Ryu and Mertz 1989; Luo and Reed 1999). Later research showed that the EJC complex is recruited after the spliceosome spliced out the introns (Le Hir et al. 2001). This would indicate the association between splicing machineries and nuclear export, however, there are some contradictory studies that showed antibodies specific to RNA binding protein Ref1 prevent interaction with RNA and inhibit mRNA nuclear export, resulting in inhibition for both spliced and unspliced transcripts (Rodrigues et al. 2001). Another study showed that removing introns only affects mRNA accumulation and translational yield, not nucleocytoplasmic distribution (Nott, Meislin, and Moore 2003). However, fluorescent in situ hybridization has also showed how intron-bearing and transcripts without any introns are distributed differently across the nucleus and cytoplasm, with intron-bearing transcripts (6 to 10-fold nucleus/cytoplasm ratio increase) preferentially located in the cytoplasm (Valencia, Dias, and Reed 2008). Transcripts containing introns in their 5'UTR has been shown to be regulated by the transcription export complex and the serine/arginine-rich (SR) proteins, whereas transcripts lacking introns in their 5'UTR were regulated by signal sequences located in the ORFs (Palazzo et al. 2007). The signal sequence coding region (SSCR), an RNA element, can promote the export of mRNAs that lack intron or a function cap (Palazzo et al. 2007). Increase in adenine content in the SSCR can also inhibit the nuclear export activity (Palazzo et al. 2007)

Another phenotype that I found when analysing the larval carcass RNA-seq data was that genes highly down regulated also had longer mature mRNAs. This led us to another question: Why are larger and long/many intron-containing transcripts more sensitive towards the loss of Nxt1? In spliced transcript, after the spliceosome removes the intron, the EJC is recruited to the mRNA (Le Hir et al. 2001). The THO complex is recruited to the mRNA, upstream of the

EJC, which also recruits UAP56 and Ref1 to form the TREX complex (Strasser et al. 2002). UAP56 is displaced by the Nxt1-Nxf1 dimer, binding to Ref1, then Ref1 is removed from the mRNP and Nxt1-Nxf1 binds directly to the mRNP (Strasser and Hurt 2000). This process occurs at every intron, and therefore, transcripts with many introns have more Nxt1-Nxf1 dimer recruitment. The RNA-seq data indicates that genes non-differentially expressed had a median gene length of 2.5kb, 2 isoforms, 4 introns, and total intron length of ~500bp (Table 3.6 and 3.7). Genes sensitive towards the loss of Nxt1 had longer gene transcripts (median of 8.4kb or more), which had a median of 9 or more introns, and 5.7 kb or more total intron length (Table 3.6 and 3.7). Genes that were not sensitive, and in fact, up regulated had shorter gene transcripts (median of less than 1.6kb), which had a median of less than 3 introns, and less than 370bp total intron length (Table 3.6 and 3.7). In general, loss of Nxt1 affected genes with more isoforms (median of 3 or more), than up regulated (median of less than 2) compared to non-differentially expressed (median of 2). Without Nxt1, Nxf1 is unable to effectively interact to components of the NPC for transportation of mRNPs to the cytoplasm (Wiegand et al. 2002). Since transcripts with more introns recruits more Nxt1-Nxf1 dimers, reduction of Nxt1 could potentially affect the transportation of these transcripts. Adding this to the fact that *abba* and potentially other transcripts have difficulty in splicing properly, which can also affect the efficiency of transportation, these data shows the importance of Nxt1 and long/many intron-containing genes. However, why are the larger transcripts affected? While this is unclear, there might be the possibility that larger transcripts need more Nxt1-Nxf1 dimers to be exported completely. Less Nxt1-Nxf1 dimers for larger transcripts could potentially lead to errors for its transportation (Figure 3.22). Since the loss of Nxt1 affects the long/many intron-containing and larger genes, it is important to study this further to fully understand the effects of Nxt1 in the RNA export pathway. It would be interesting to understand if all spliced introns require the loading of the EJC and recruitment of Nxt1-Nxf1 dimer or if one successful recruitment of EJC and Nxt1-Nxf1 dimer to a spliced intron is sufficient for transportation. There is a possibility that larger transcripts with several introns require a minimum amount of Nxt1-Nxf1 dimer recruitment for successful exportation.

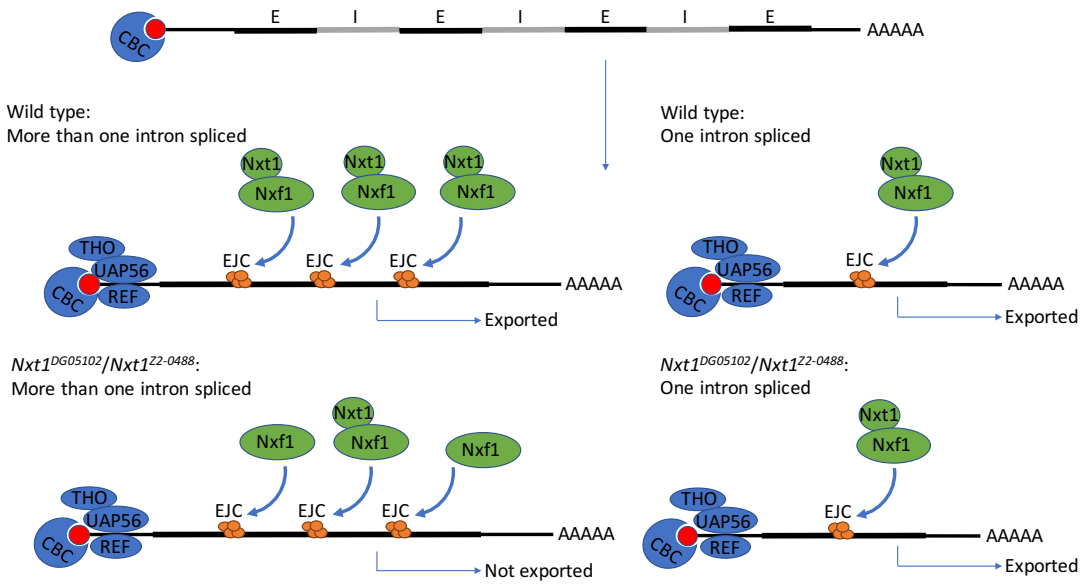


Figure 3.22: **Proposed hypothesis of Nxt1-Nxf1 dimer recruitment in Nxt1 trans-heterozygotes.** After splicing of introns, Nxt1-Nxf1 is recruited through a series of events, and is required for transportation of mRNPs to the cytoplasm. More splicing events lead to more recruitment of the Nxt1-Nxf1 dimer. In Nxt1 trans-heterozygotes, less function protein of Nxt1 is expected (Caporilli et al. 2013). The current hypothesis is that transcripts with more introns, and therefore more recruitment of Nxt1-Nxf1 dimer, could lead to incomplete recruitment of the dimer (only Nxf1) or potentially no dimer recruitment. No / incomplete recruitment might have an effect of transportation of the mRNP complex through the NPC to the cytoplasm and instead be degraded.

However, while these results are intriguing, why did I not see this phenotype for the whole larva RNA-seq? The RNA-seq differentially expressed genes were mostly specifically expressed in the testis. This was not an entire surprise, since many genes in adult testis are differentially expressed with the loss of Nxt1 (Caporilli et al. 2013). In *Drosophila*, roughly 22% of the annotated genes lack introns, and genes with no introns are enriched for testis-specific expression (Caporilli et al. 2013). The overabundance of testis-specific genes in the RNA-seq data therefore interferes with our data making it impossible to use for gene intron analysis. The carcass RNA-seq was used specifically to look at genes important for muscles. This identified a correlation between the total intron length and number of introns of all down regulated genes. However, it is unclear if this phenotype is specific to the carcass or also observed in other tissues. RNA-seq of different tissues of larvae for wild type and Nxt1 trans-heterozygotes should be performed, followed by bioinformatics analysis of all differentially

expressed genes. Or a new whole larvae RNA-seq can be performed again by sexing the larvae to avoid the abundance of testis-specific genes and perform bioinformatics analysis of all differentially expressed genes. The differentially expressed genes can also be sorted by tissue-specific expression and the same bioinformatics analysis can be performed to look at the total intron length, number of introns and more.

Finally, up regulated genes have smaller transcripts than the non-differentially expressed genes (Table 3.7). If we believe that larger transcripts need more Nxt1-Nxf1 dimers for transportation, then smaller transcripts need fewer Nxt1-Nxf1 dimers for transport (Figure 3.22). This, in turn, with the loss of Nxt1, has a greater chance of successful export than larger transcripts. If there is less Nxt1 available, smaller transcripts that require less Nxt1-Nxf1 dimer might be more prioritized, since they are potentially more successful for transport, and therefore have an increased expression than normal. However, one could equally argue that these transcripts would be more susceptible to loss of Nxt1 because they have fewer potential loading sites and might not get any Nxt1-Nxf1 dimer at all. Finally, there is also the possibility that the transcripts increased in expression are due to a count normalization artifact. A case study of chickpea transcriptome revealed expression artifacts that can bias expression and differential expression analysis due to calculation of counts from transcripts encoding multiple genes (Chakraborty 2016).

3.13 Increasing Nxt1 expression in Nxt1 trans-heterozygotes rescues lethality

3.13.1 Increasing Nxt1 with UAS/gal4 system can rescue lethality with ubiquitous driver

We had two UAS-Nxt1 lines that I used to rescue the expression of Nxt1 with the UAS/gal4 system. I tried two different gal4 drivers, muscle-specific, mef2gal4 and moderate ubiquitous, armgal4. These were crossed into the Nxt1 trans-heterozygotes background and the fly development was monitored and 3rd instar larva muscles were analysed. Increasing Nxt1 via mef2gal4 partially rescued the muscle degeneration (Figure 3.19C). As expected, pupa semi-lethality still occurred at high levels, even though line M1 had a slight improvement in semi-lethality. These data indicate that rescuing Nxt1 in muscles only is sufficient enough to partly

restore muscle integrity. Data also showed that larval mobility is slightly rescued which corresponds to the partial rescue of the muscle phenotype (Figure 3.19D).

Rescuing Nxt1 ubiquitously showed some interesting data. Again, muscle degeneration was partially rescued with some outliers that showed no rescue at all (Figure 3.20C). This is quite surprising since mef2gal4 is a muscle-specific driver, yet the muscle rescue is very similar to a moderate ubiquitous driver armgal4. I did not look at the *Nxt1* levels via qRT-PCR, but so far, higher expression of *Nxt1* in muscles does not necessarily mean a better rescue of the muscle phenotype.

Most interestingly, rescuing Nxt1 with the armgal4 driver reduces the semi-lethality significantly (Figure 3.20A). The semi-lethality is reduced from 80% to 30% (M2A) and 25% (M1). With armgal4, Nxt1 is rescued across all tissues rather than just the muscles. However, both lines partially rescue the muscle phenotype, yet, only armgal4 can rescue the semi-lethality. The semi-lethality is partly caused by improper air bubble migration that results in no head eversion. I looked closely at the pupae between the mef2gal4 and armgal4 drivers to better understand the results (Table 3.4). About 50% of the Nxt1 trans-heterozygotes have uneverted or missing spiracles. Spiracles can be seen at both anterior and posterior ends. Spiracles at the anterior end are affected in the Nxt1 trans-heterozygotes. The spiracles allow air to enter the trachea and can also be opened and closed in an efficient manner to reduce water loss. Contracting muscles close the spiracles and this is controlled by the central nervous system. I examined pupae spiracles for the Nxt1 rescue experiments for both mef2gal4 and armgal4 drivers and found that with mef2gal4 roughly 60-70% of the pupae had normal spiracles, whereas with armgal4, all pupae had normal spiracles (Table 3.4). Pupae with uneverted or missing spiracles can lead to death if air is not supplied to the organs. Mef2gal4 expresses Nxt1 specifically in the larval muscles and it is possible that the muscles controlling the spiracles are not affected with this driver, hence I see no rescue of this phenotype. Therefore, with mef2gal4, a small percentage of pupae can be affected due to this. I also looked at the head eversion that is clearly visible towards the end of the pupa phase. With mef2gal4 more pupae were dying compared to with armgal4 at the end of the pupa phase (Table 3.4). Interestingly, from the pupa that did not survive there were no big changes seen in the head eversion between mef2gal4 and armgal4 (mef2gal4 M1/M2A;

46%/62% and armgal4 M1/M2A; 57%/69%), but it was still a significant increase compared to Nxt1 trans-heterozygotes (26%). Thus, with armgal4, pupae were surviving through adulthood more significantly compared to mef2gal4, but from the few pupae that did not survive, the head eversion still occurred at a relatively high level, rather than lethality due to other vital processes required for survivability. So, why are fewer pupae dying with the armgal4 driver?

3.13.2 *Nxt1* rescue with ubiquitous driver indicates that Nxt1 affects lots of processes in pupae required for viability

A rather simple answer to the question above would be that armgal4 is a ubiquitous driver and has an effect on all tissues of the pupa. But what processes would be rescued with armgal4?

The pupae that were dying close before emerging had most of the organs developed. Eye development, for example, starts towards the middle of the pupa phase and slowly darkens towards the rest of the phase. I could see that pupae with head eversion had the correct eye color for their genotype. Eye development for pupae with no head eversion also sometimes occurs, albeit abnormally as the heads are within the thorax (data not shown). For pupae with head eversion, the shape of the fly inside the pupa case (head, thorax and abdomen) is developed and bristles on back, humeri and legs were visible. For pupae with no head eversion the body shape is slightly different since only the thorax and abdomen were visible with the head hidden inside the body. Wing development also occurs as expected. Pupae dying due to no head eversion is self-explanatory, but how about pupae that do have head eversion, visible organ development and still die towards the end of the phase?

This is most likely due to pathways and / or processes affected in the Nxt1 trans-heterozygotes. To fully understand the semi-lethality issues in the pupa, RNA-seq should be performed at differently staged pupae, mostly towards the end of the phase. The larvae must be sexed to avoid the genes highly enriched in testis. This allows looking closely at the differentially expressed genes in pathways and other processes that are important in the pupa. With this project, I only focused on the air bubble phenotype in the early pupa. In the future, experiments can be performed to understand the effects of Nxt1 in the late pupa.

3.14 Gal4 driven without an UAS construct increases expression of many genes

Larval carcass RNA-seq data revealed that in Nxt1 trans-heterozygotes highly down regulated genes had high total intron length and many more introns. The qRT-PCR revealed, that indeed, genes with high total intron length were reduced in expression (Figure 3.21). Interestingly, qRT-PCR data of mef2gal4>UAS-*abba*^{fl} in Nxt1 trans-heterozygotes larval carcasses showed that expression of *abba* was actually not rescued, but instead, rescue of gene expression was observed from several genes that were supposedly down in Nxt1 trans-heterozygotes (Supplementary figure 4). The failure in *abba* rescue might be due to the sample size (10 vs. 90 larval carcass samples for WT and Nxt1 trans-heterozygotes) due to the high variability that can be observed in larvae. However, when I looked at the gene expression in Nxt1 trans-heterozygotes with mef2gal4 or armgal4 without an UAS-construct, gene expression was rescued for most of the genes examined (Supplementary figure 4). Surprisingly, larval carcasses from armgal4; Sb/Tm3 also showed a rescue in expression of all the genes (Supplementary figure 4). Two control genes, *CG1307* and *CG9330*, were not affected in any of the genotypes used. There is a possibility that the mef2gal4>UAS-*abba*^{fl} in Nxt1 trans-heterozygotes larval carcass rescues the expression of other genes, due to the presence of Gal4, that contribute to the integrity of the larval muscles. These data, so far, revealed that gal4 alone affects the expression of genes. This has been shown before where gal4 causes developmental defects and apoptosis when expressed in the developing eye (Kramer and Staveley 2003). Another research has shown that expressing gal4 leads to a genomic response leading to genes being up and down regulated (Liu and Lehmann 2008). However, the qRT-PCR data only showed a few genes with only 2 control genes, using small number of samples. To better understand the effect of gal4 in Nxt1 trans-heterozygotes larva and wild type like larva, RNA-seq should be performed for Nxt1 trans-heterozygotes with armgal4 alone and armgal4; Sb/Tm3 larva. It is important to keep in mind the effects of gal4 alone on gene expression and how it could impact analysis and result interpretation of your data.

Chapter 4.

Understanding the importance of Nxt1 in the piRNA pathway

Chapter 4 – Results

In this chapter, I will focus on the importance of Nxt1 in adult ovaries. Microarray data of 7 days old ovaries have shown that loss of Nxt1 up regulates the expression of transposons. Small RNA-seq will be performed to understand how big the effect of loss of Nxt1 is in Nxt1 trans-heterozygotes. With the small RNA-seq data, I will continue to look at different parts of the piRNA pathway, such as piRNA precursor transportation, piRNA 3' trimming, piRNA methylation, to understand where in the piRNA pathway Nxt1 is involved.

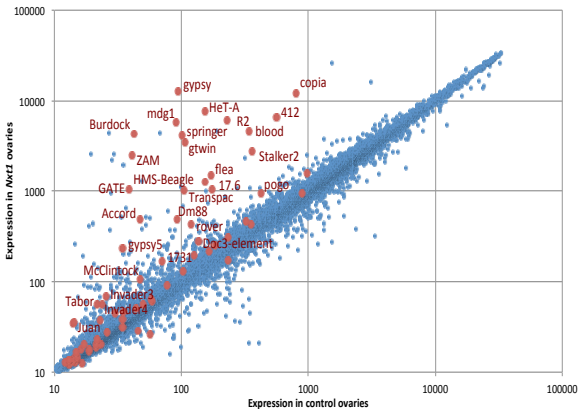
4.1 Transposons are up regulated in Nxt1 trans-heterozygotes ovaries

In Nxt1 trans-heterozygotes testis, many genes were dramatically under-expressed compared to wild type controls (Caporilli et al. 2013). Adult Nxt1 trans-heterozygote females are, like males, sterile. Affymetrix microarray of 7 days old ovaries from w^{1118} and Nxt1 trans-heterozygotes, each genotype in triplicates, was performed on an affymetrix array v2. Data was processed in Excel by normalized intensity for each probe. These data revealed that 198 genes were more than 2-fold down regulated and 104 genes were more than 2-fold up regulated. The list of differentially expressed genes showed that 31 transposons were differentially expressed. Thirty transposons were more than 2-fold up regulated and one more than 2-fold down regulated (Figure 4.1A and 4.2). Forty-one transposons were not differentially expressed. The piRNA pathway is involved in actively repressing transposons in the *Drosophila* and other organisms (Brennecke et al. 2007). Without the piRNA pathway, transposons jump throughout the genome, leading to sterility (Siomi et al. 2011). To validate this result, I performed a qRT-PCR on several up regulated transposons for extra confirmation. These data showed that, indeed, transposons are up regulated up to a 100-fold difference (Figure 4.1B).

The piRNA pathway is responsible for actively repressing transposons and protecting the stability of the genome (Siomi et al. 2011). Microarray data showed that transposons repressed in germline cells (HeT-A) and somatic cells (ZAM) were up regulated. These data

revealed that the piRNA pathway could be compromised and that transposon repression in both germline and somatic cells are affected in Nxt1 trans-heterozygotes.

A



B

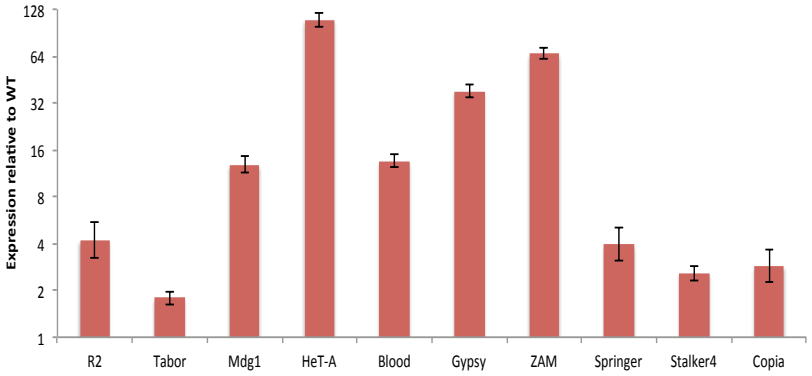


Figure 4.1: **Thirty transposons severely de-repressed in *Nxt1* trans-heterozygotes.** A) Scatter plot showing normalized expression levels in 7 days old w^{1118} ovaries and Nxt1 trans-heterozygotes mutant ovaries. B) qRT-PCR on independent samples of 7 days old w^{1118} ovaries and Nxt1 trans-heterozygotes mutant ovaries for 10 transposons shown to be up regulated in the microarray data. Error bars show the confidence interval of the difference between means. *Rp49* was used as a control.

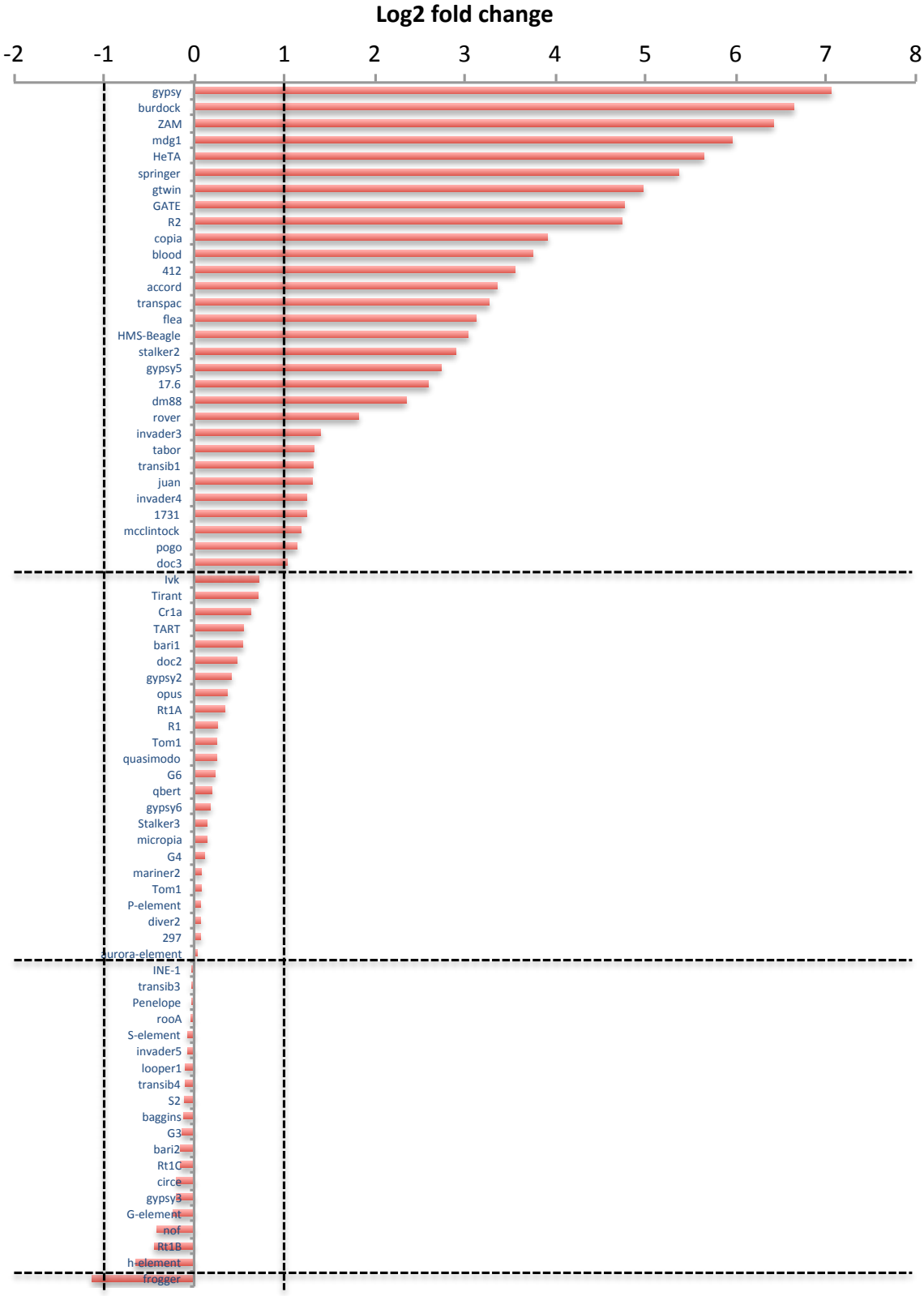


Figure 4.2: **Many transposons detected by microarray highly increased in expression.** List of all transposons detected by microarray and its log2 fold change value in Nxt1 trans-heterozygotes.

4.2 Three major piRNA precursor RNAs are expressed at normal levels in Nxt1 trans-heterozygotes ovaries

Microarray data suggested that Nxt1 is involved in the active repression of transposons, which is known to be regulated by the piRNA pathway. In this pathway piRNA clusters are transcribed to generate long single stranded transcripts that are processed into piRNAs in the nuage or Yb bodies in germline or somatic cells, respectively (Murota et al. 2014; Malone et al. 2009). To understand where Nxt1 could potentially be involved in this process, I examined the transcript levels of three major piRNA clusters via qRT-PCR. *flam* is a major piRNA source and is only active in the somatic cells. *42AB* (also known as *Cluster 1)* is only active in germline cells. *20A* (also known as *Cluster 2*) is active in both germline and somatic cells. Each piRNA cluster transcript expression was examined with qRT-PCR using primer pairs targeted against two different places within the transcript, as performed previously by (Goriaux et al. 2014; Klattenhoff et al. 2009). The transcript expression levels from all three piRNA clusters were similarly expressed in wild type and Nxt1 trans-heterozygous ovaries at both 7 and 14 days of age (Figure 4.3).

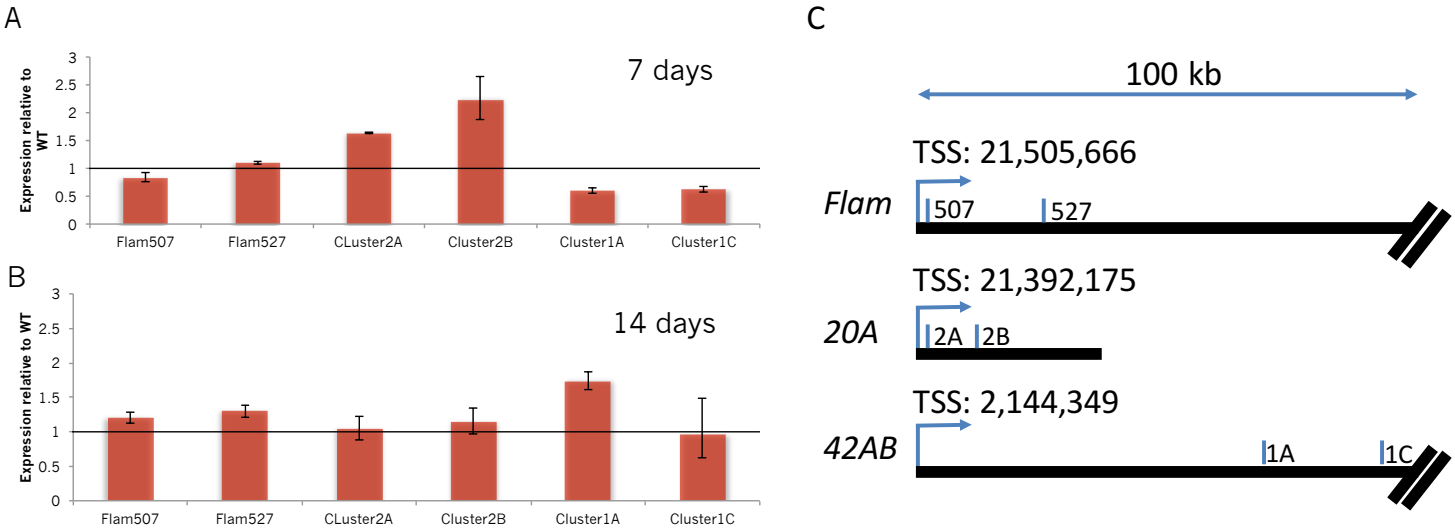


Figure 4.3: **Three major piRNA precursor RNAs are expressed at normal levels in Nxt1 trans-heterozygotes.** qRT-PCR of piRNA transcript RNAs from 7 days (A) and 14 days (B) are at relative levels as wild type for *Flam*, *42AB* (*Cluster 1*), and *20A* (*Cluster 2*). C) Locations of primers for each piRNA cluster transcript. Each piRNA transcript is detected at two different regions. Error bars show the confidence interval of the difference between means.

4.3 None of the genes involved in the piRNA pathway are affected

Microarray data was used again from seven days old Nxt1 trans-heterozygotes ovaries to look at all the genes known or suspected to be involved in the piRNA pathway (Czech et al. 2013). Interestingly, none of these genes were differentially expressed (Figure 4.4) revealing that all the genes identified so far to be required for a functional piRNA pathway were present at normal transcript levels in Nxt1 trans-heterozygote ovaries, yet transposons were up regulated. Even though the transcript levels were normal in the Nxt1 trans-heterozygotes, it is unknown if the protein expression levels were affected.

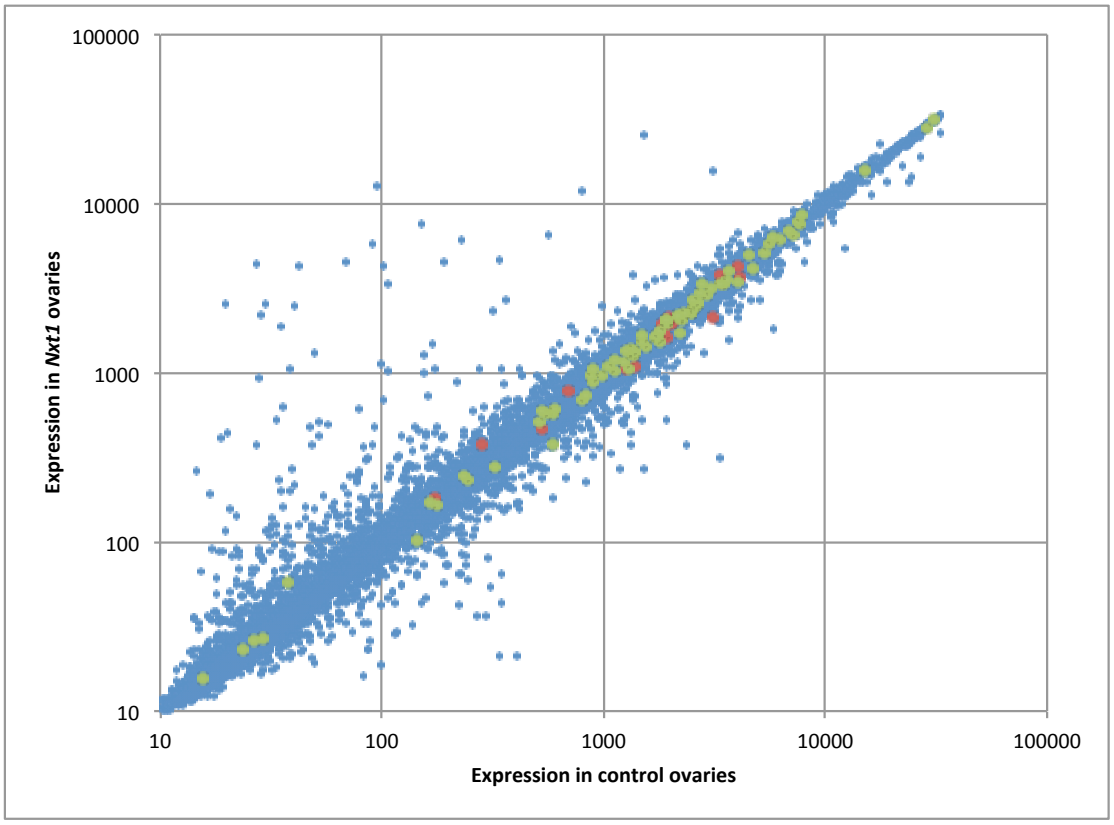


Figure 4.4: **Known and suspected genes involved in piRNA pathway are not significantly expressed.** Scatter plot showing normalized expression levels in 7 days old w^{1118} ovaries and Nxt1 trans-heterozygotes mutant ovaries. Microarray data from seven days old wild type and Nxt1 trans-heterozygote ovaries with known genes highlighted in red and suspected genes highlighted in green (See Appendix 5 for full list of gene names).

4.4 piRNA levels and ping-pong amplification are reduced, but U-bias is maintained

To further understand the role of Nxt1 in the piRNA pathway, small RNAs between 21-29nt were isolated from w^{1118} and Nxt1 trans-heterozygote ovarian total RNA from 4, 14 and 20 days old females. The small RNA-seq was performed at different time points to examine whether the degenerating ovaries of 20 days old Nxt1 trans-heterozygotes were due to defects in the piRNA pathway. Libraries were made from the purified small RNAs and sequenced by the University of Liverpool. Each genotype was performed in triplicates and all data was analysed by piPipes (Han, Wang, Zamore, et al. 2015). To use piPipes dual mode, one replicate per time point for each genotype was selected (Supplementary table 2). The selection for each replicate was dependent on the total read depth, rRNA mapping and miRNA hairpin mapping (Supplementary table 2). For each time point of both genotypes, replicates were selected that were similar in read depth, rRNA mapping and miRNA mapping. FlyBase dm3 was used for genome mapping and the piRNA reads were normalized per 1 million miRNA reads. The uniquely mapping piRNA reads between 23-29 nucleotides were mapped to all known piRNA clusters and length distribution plots, ping-pong analysis and nucleotide composition were performed. This analysis revealed that at 4 days, there was a slight reduction of piRNAs mapping to both the top and bottom strand at all known piRNA clusters in mutant ovaries (Figure 4.5A). The piRNA levels were more dramatically reduced (>50%) at 14 days in Nxt1 trans-heterozygotes (Figure 4.5B). Interestingly at 20 days, the piRNA levels in both wild type and Nxt1 trans-heterozygotes were significantly reduced (Figure 4.5C) compared to the earlier time points. This revealed that the piRNA pathway becomes less active at an older age in *Drosophila* females.

The ping-pong analysis compares the nucleotide overlap between two piRNAs of opposing strands and calculates the relative offset. piRNAs involved in the ping-pong amplification loop have a 10 nucleotide overlap due to Aub/AGO3 cleavage. This analysis was performed in all three time points, and the ping-pong amplification loop was severely reduced at 14 days in Nxt1 trans-heterozygotes, whereas no reduction of piRNA pairs is seen at 4 and 20 days (Figure 4.5D-F). At 20 days, a higher background, with piRNAs overlapping at ~20nt and ~23nt, is observed in both wild type and Nxt1 trans-heterozygotes (Figure 4.5F).

Antisense oriented piRNAs have a U-bias at the first nucleotide at the 5'end. This bias comes from 5'end cleavage by Zucchini (Ipsaro et al. 2012; Nishimasu et al. 2012). A nucleotide composition of 30 nucleotides upstream and downstream of the 5' end was performed by piPipes to examine the 5' U-bias. I examined the first nucleotide of the 5' end of all piRNAs mapping to all piRNA clusters and showed that no significant change of the U-bias at all three time points in Nxt1 trans-heterozygotes was observed (Figure 4.5G-I). This revealed that Zucchini is likely to perform 5' end cleavage of the piRNA precursor transcripts.

As I observed a reduction of piRNA levels at all piRNA clusters, a closer analysis was performed for *flam*, *42AB*, *20A* and all transposons (Figure 4.6, 4.7, 4.8 and Supplementary figure 3). *flam* and *20A* are uni-strand and only provide piRNAs derived from the top strand, while 42AB is a dual strand cluster transcribed from both strands (Brennecke et al. 2007) (Figure 4.6A-C and 4.8A-C). From these three clusters the most severe reductions in piRNA abundance were observed for *flam* (-75%) and *42AB* (-63%) in 14 days old Nxt1 trans-heterozygotes (Figure 4.6B and 4.7B). In *20A*, piRNA levels were reduced by 35% in Nxt1 trans-heterozygotes (Figure 4.8B). Also, in *20A*, the piRNA level in wild type at 14 days of age is lower than 4 days, while *flam* and *42AB* wild type piRNA levels are similar between 4 and 14 days of age (Compare Figure 4.8A-B with 4.6A-B and 4.7A-B). Reduced piRNA levels in wild type was observed for all three clusters at 20 days of age compared to wild type at younger ages. In all three piRNA clusters, wild type piRNA levels were severely reduced compared to earlier stages. (Figure 4.6A-C, 4.7A-C and 4.8A-C). However, piRNA levels in Nxt1 trans-heterozygotes at 20 days were lower than wild type (Figure 4.6C, 4.7C and 4.8C).

flam piRNA cluster is only active in somatic cells, and piRNAs do not have a 10nt overlap with other piRNAs as these piRNAs are antisense oriented with no ping-pong amplification availability in somatic cells. Analysis of the small RNA sequence dataset confirmed that, indeed, no 10nt overlap was observed (Figure 4.6D-F). Conversely, piRNAs from the *42AB* cluster have a clear 10nt overlap between opposing strands as expected, however, the 10nt overlap was reduced in Nxt1 trans-heterozygotes at 14 days, while more background (longer overlaps) was observed at 20 days for both wild type and Nxt1 trans-heterozygotes (Figure 4.7D-F). However, *20A*, active in both germline and somatic cells, showed a 10nt overlap at 4 days, which became absent at 14 and 20 days in wild type and Nxt1 trans-heterozygotes

(Figure 4.8D-F). In 14 and 20 days ovaries, piRNA levels from *20A* cluster were reduced for both wild type and Nxt1 trans-heterozygotes when compared to 4 days (Figure 4.8A-C). This revealed that the *20A* cluster in 14 and 20 days is potentially only active in somatic cells where the ping-pong amplification is absent, whereas at 4 days, *20A*-derived piRNAs are involved in the ping pong amplification.

At 4, 14 days and 20 days, no difference was observed in the 5' end first nucleotide U-bias in *flam*, *42AB*, and *20A* (Figure 4.6G-I, 4.7G-I and 4.8G-I) between wild type and Nxt1 trans-heterozygotes, which indicated that the 5' trimming occurred as expected in both genotypes.

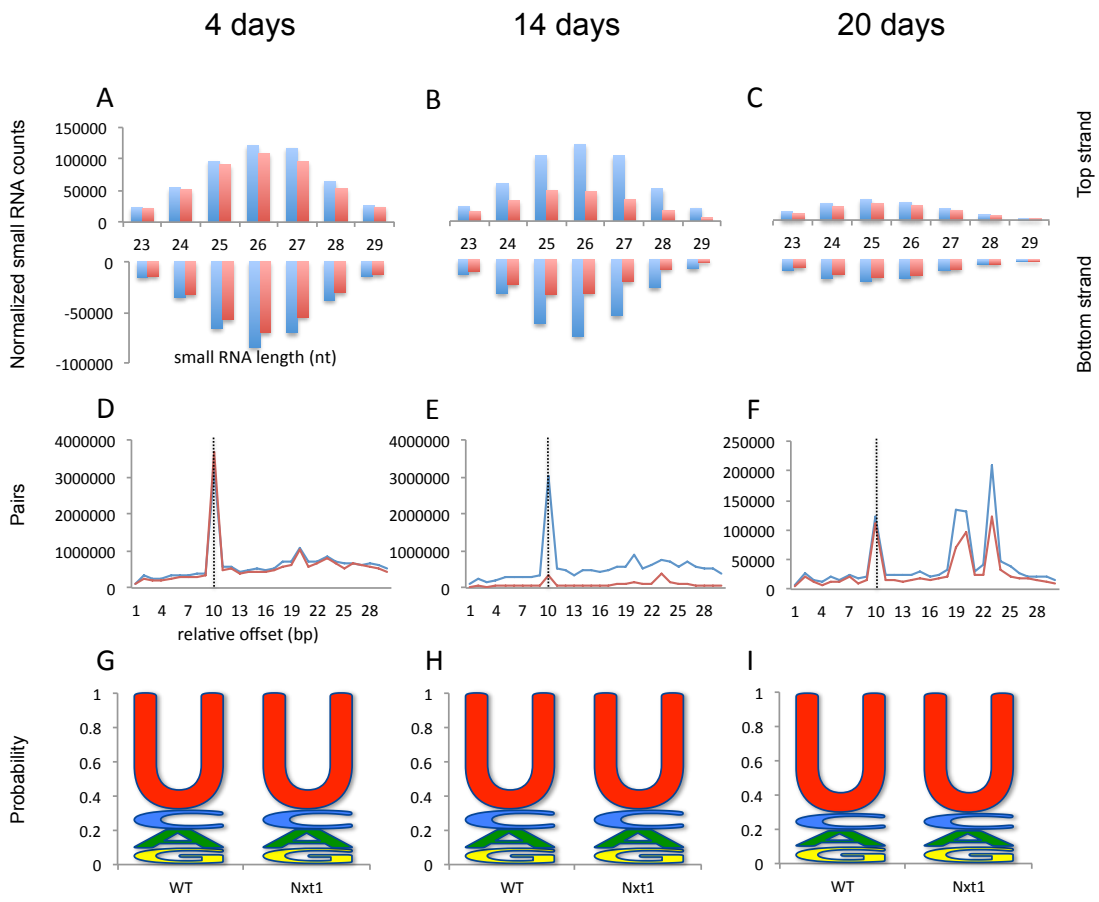


Figure 4.5: **Reduced piRNA levels and ping-pong amplification but maintained U-bias in Nxt1 trans-heterozygotes.** Unique piRNA reads mapping to all known piRNA clusters. Length distribution plots from (A) 4 days, (B) 14 days and (C) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red) with reads mapping to top strand (top) and bottom strand (bottom). Ping-pong analysis of (D) 4 days, (E) 14 days and (F) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red). Ten nucleotide overlaps indicated with dashed lines. First nucleotide 5' U-bias of (G) 4 days, (H) 14 days and (I) 20 days. piRNA reads were normalized per 1 million miRNA reads.

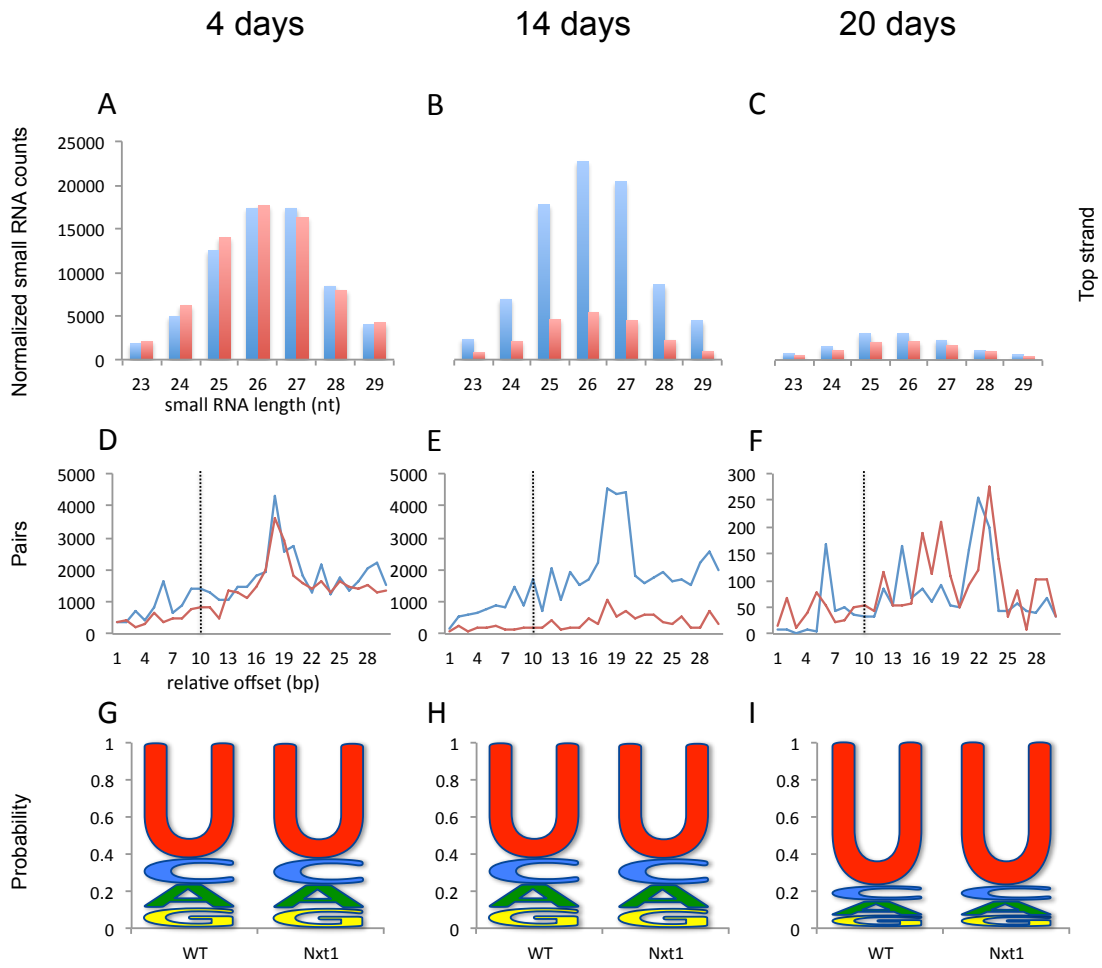


Figure 4.6: **Unique piRNA reads mapping to *flam* piRNA cluster**.

Length distribution plots from (A) 4 days, (B) 14 days and (C) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red) with reads mapping to top strand (top). Ping-pong analysis of (D) 4 days, (E) 14 days and (F) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red). Ten nucleotide overlaps indicated with dashed lines. First nucleotide 5' U-bias of (G) 4 days, (H) 14 days and (I) 20 days. piRNA reads were normalized per 1 million miRNA reads.

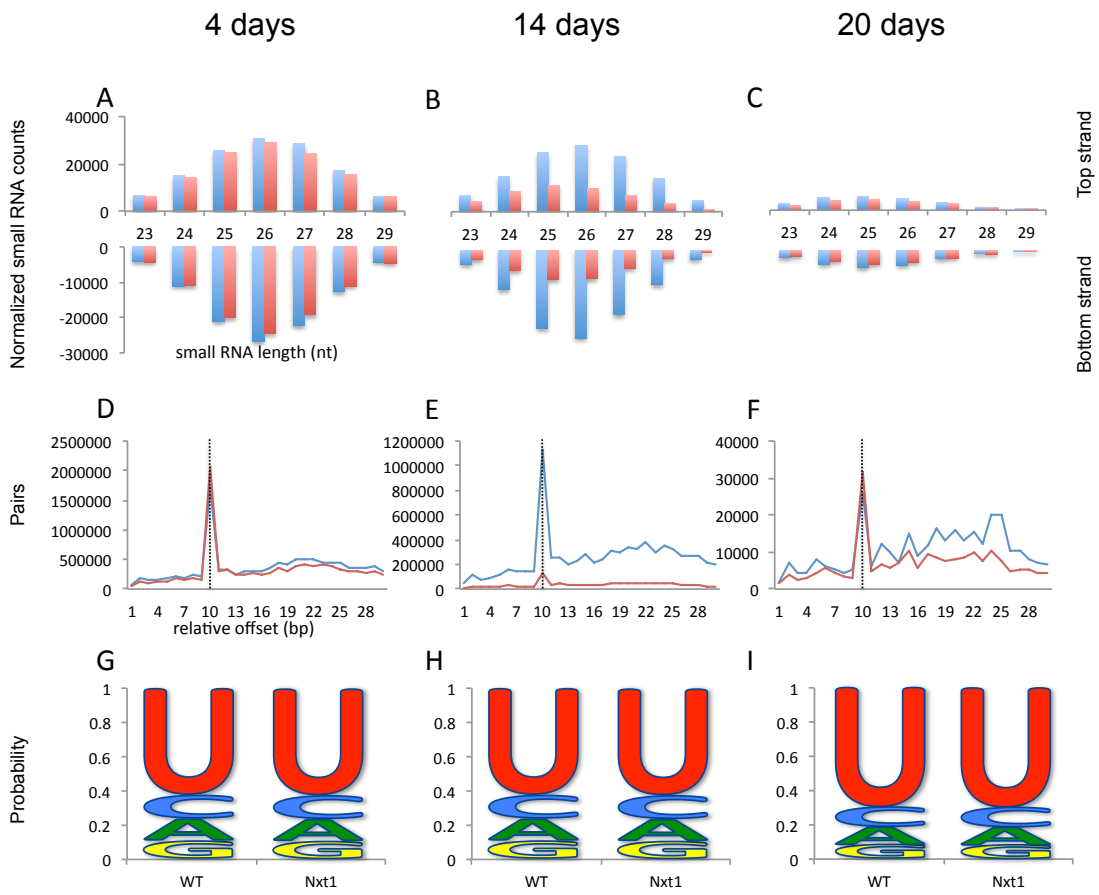


Figure 4.7: **Unique piRNA reads mapping to *42AB* piRNA cluster**. Length distribution plots from (A) 4 days, (B) 14 days and (C) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red) with reads mapping to top strand (top) and bottom strand (bottom). Ping-pong analysis of (D) 4 days, (E) 14 days and (F) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red). Ten nucleotide overlaps indicated with dashed lines. First nucleotide 5' U-bias of (G) 4 days, (H) 14 days and (I) 20 days. piRNA reads were normalized per 1 million miRNA reads.

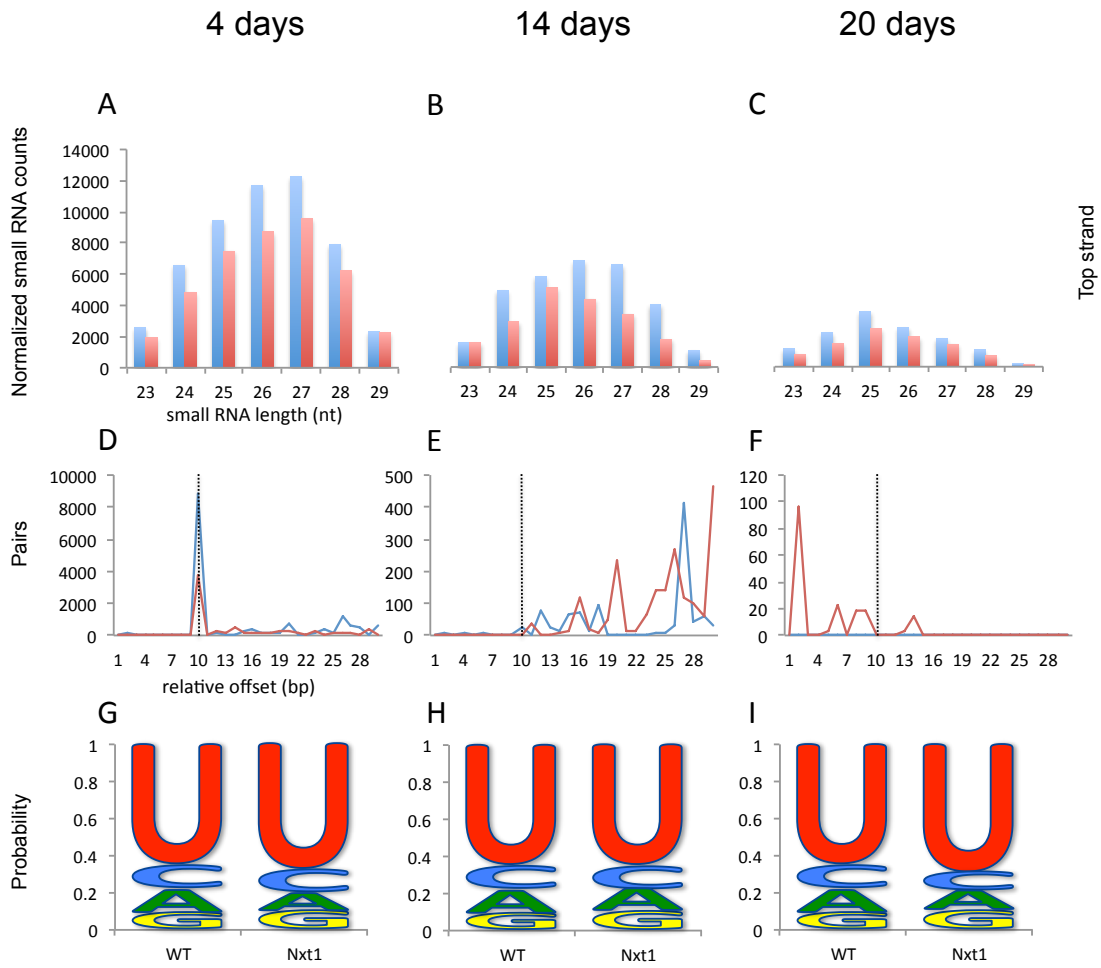


Figure 4.8: **Unique piRNA reads mapping to *20A* piRNA cluster**. Length distribution plots from (A) 4 days, (B) 14 days and (C) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red) with reads mapping to top strand (top). Ping-pong analysis of (D) 4 days, (E) 14 days and (F) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red). Ten nucleotide overlaps indicated with dashed lines. First nucleotide 5' U-bias of (G) 4 days, (H) 14 days and (I) 20 days. piRNA reads were normalized per 1 million miRNA reads.

4.5 Sense and antisense piRNA reads mapping to piRNA clusters are reduced in Nxt1 trans-heterozygotes

The analysis presented thus far showed that the difference between wild type and Nxt1 trans-heterozygotes piRNA levels is reduced at all three time points, while 14 days showed the largest difference. However, that data only showed the piRNAs mapping to the plus and minus strand and no additional information about sense and antisense orientation of the piRNAs. In piPipes dual mode, wild type sense piRNAs mapping to all piRNA clusters were plotted against Nxt1 trans-heterozygotes sense piRNAs mapping to all piRNA clusters. The same was done for antisense piRNAs. This revealed a shift towards wild type, indicating a reduction of piRNAs in Nxt1 trans-heterozygotes at 4 days for both sense and antisense piRNAs mapping to a majority of the piRNA clusters (Figure 4.9A-B). At 14 days, a larger shift is observed towards wild type with more piRNA clusters affected in Nxt1 trans-heterozygotes (Figure 4.9C-D). At 20 days, there is no shift observed between wild type and Nxt1 trans-heterozygotes (Figure 4.9E-F). From the three different time points, 14 days showed the biggest shift in Nxt1 trans-heterozygotes for both sense and antisense piRNA reads mapping to all known piRNA clusters.

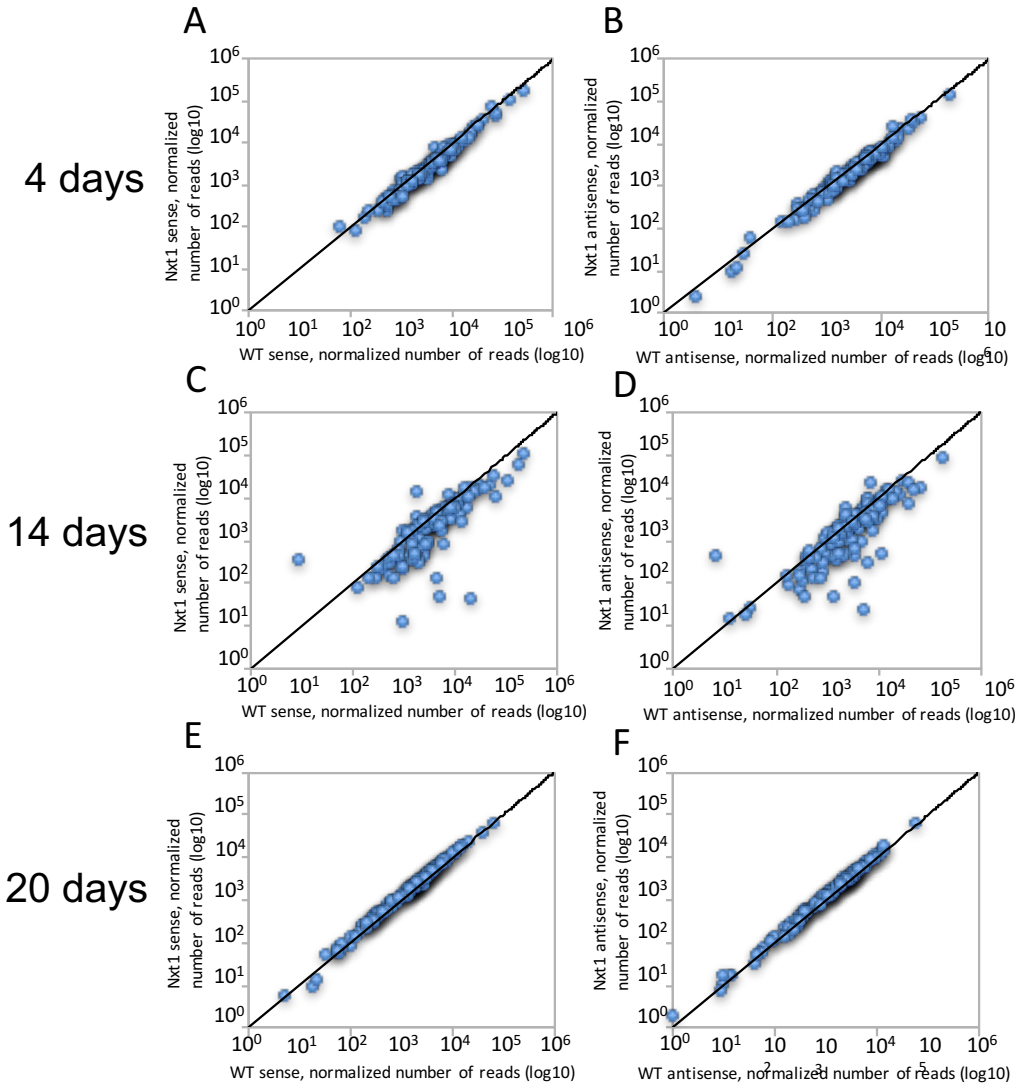


Figure 4.9: **Both sense and antisense oriented piRNAs show reduced level at 4 and 14 days old ovaries.** WT sense piRNAs plotted against Nxt1 trans-heterozygotes sense piRNAs (A, C and E) and WT antisense piRNAs plotted against Nxt1 trans-heterozygotes antisense piRNAs (B, D and F) at 4 (A and B), 14 (C and D) and 20 days (E and F). piRNA reads were normalized per 1 million miRNA reads.

4.6 Mature piRNAs are shorter in Nxt1 trans-heterozygotes

piPipes calculates the average length of all piRNAs mapping to each transposon in the *Drosophila* genome. Interestingly, the median length of both sense and antisense piRNAs were reduced by ~0.5nt at 4 and 14 days, and by ~0.3nt at 20 days in Nxt1 trans-heterozygotes (Figure 4.10). Additionally, in the wild type 20-day old ovaries, the median piRNAs length was

reduced, by ~0.5nt, compared to wild type at earlier time points. The reduction in length could be explained due to increased activity of Nbr observed in aging ovaries resulting in shorter piRNAs (Wang et al. 2016). Nbr is a 3' to 5' exonuclease that trims the 3' end of piRNAs (Feltzin et al. 2015). However, it is unclear from the small RNA-seq how Nxt1 could be involved in the process of piRNA length and whether the Nbr protein activity is increased. This analysis, however, did indicate that the 3' trimming might be affected in Nxt1 trans-heterozygotes.

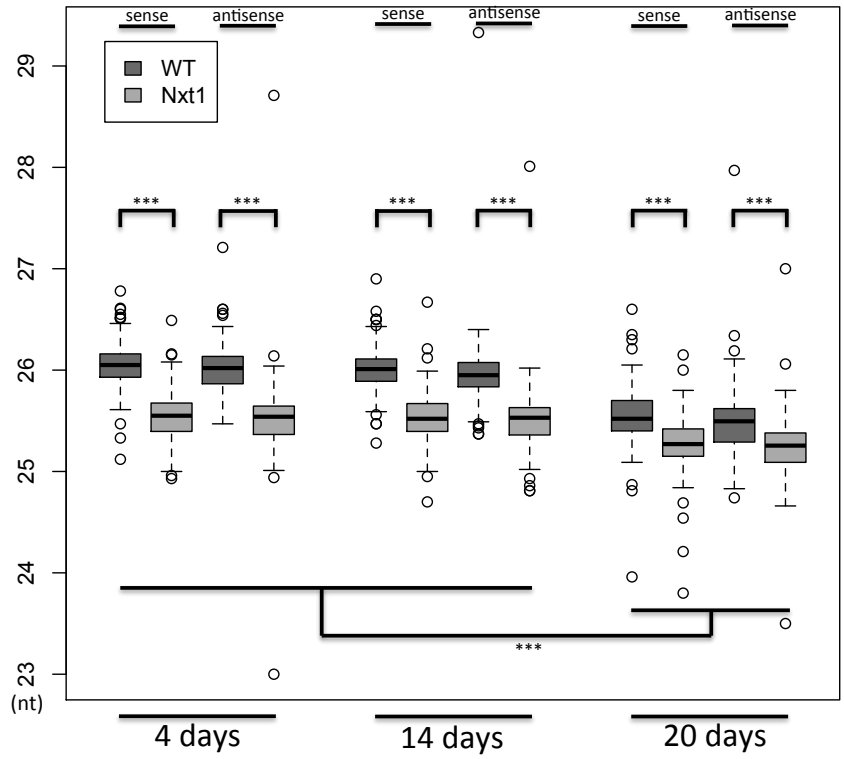


Figure 4.10: **piRNAs have a shorter length in Nxt1 trans-heterozygotes.** Box plot of 4, 14 and 20 days wild type (dark gray) and Nxt1 trans-heterozygotes (light gray) plotting piRNA length separated by sense and antisense direction. Both sense and antisense oriented piRNAs have a shorter median piRNA length at 4 and 14 days, by ~0.5nt, when compared to Nxt1 trans-heterozygotes. At 20 days, piRNAs length in wild type is shorter, by ~0.5nt, than 4 and 14 days wild type. Nxt1 trans-heterozygotes were shorter by ~0.3nt in 20 days compared to wild type. Student's t-test *** p value = <0.005.

4.7 New small RNA-seq data revealed differences observed in the expression of the majority of transposons

To further investigate the role of Nxt1 in piRNA function, a collaboration was set up with Julius Brennecke, a group heavily focused on piRNAs in *Drosophila* to focus on the role of both Nxt1 and Nxf1 in the piRNA pathway. We decided to re-do the small RNA sequencing according to their protocols to directly compare the data from their lab. The libraries were prepared with a combination of the TruSeq small RNA library prep kit and their own components compared to the NEB small RNA kit from the previous round. Using their protocol allowed us to use adapters with 4 random nucleotides for ligation to small RNAs. This prevented spikes to appear in small RNA-seq due to the reduction of preferred substrates with the presence of 4Ns and the variability of the adapter sequence (personal communication with Dominik Handler). I looked at the old sequencing data and compared it with a wild type small RNA data set from their group and found indeed that the NEB small RNA kit showed more spikes across the piRNA clusters (Figure 4.11 top arrows). The increase in spikes interferes when using small RNA seq for looking into the piRNA processing patterns. Their small RNA libraries did not show these spikes (Figure 4.11 bottom).

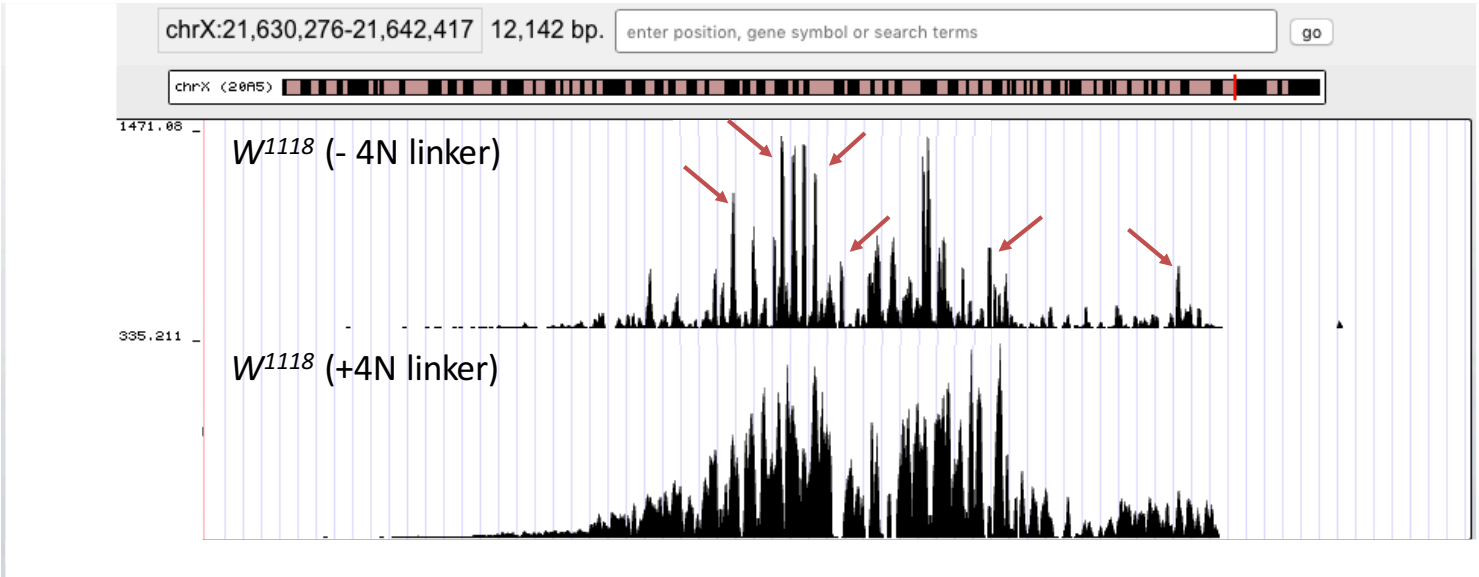


Figure 4.11: **More spikes observed without using 4N linkers.** Small portion of *flamenco* region of two wild type datasets were compared with -4N linker on top and +4N linker on bottom. Linkers without the 4N show more peaks (red arrows).

Seven days old ovaries were used for the new set of small RNA-seq to look at transposons that were repressed specifically in the germline (Figure 4.12A), both germline and soma (Figure 4.12B) and specifically in the soma (Figure 4.12C). However, looking at the data I noticed some discrepancies from the 1st round of small RNA seq and microarray data. Most of the transposons had an increase in reads mapping to the sense strand in Nxt1 trans-heterozygotes (*rover*, *297*, *idefix*, *blood*, *Stalker2* and *mdg1*; Figure 4.12A-B). Also, some of the soma-specific transposons highly up regulated in our microarray data had more reads mapping to the antisense strand in Nxt1 trans-heterozygotes (*gtwin*, *tabor* and *gypsy*; Figure 4.12C). Length distribution across all transposons revealed that more piRNA reads were mapped in Nxt1 trans-heterozygotes (Figure 4.13A). Separating the reads by antisense and sense strand showed that Nxt1 trans-heterozygotes had more reads mapping to both sense and antisense strands (Figure 4.13B). Interestingly, the 2nd round of small RNA-seq data revealed that the majority of transposons were not up regulated, and had in fact, more piRNA reads mapping in Nxt1 trans-heterozygotes, which could indicate that Nxt1 is involved in silencing of transposons rather than piRNA precursor transportation.

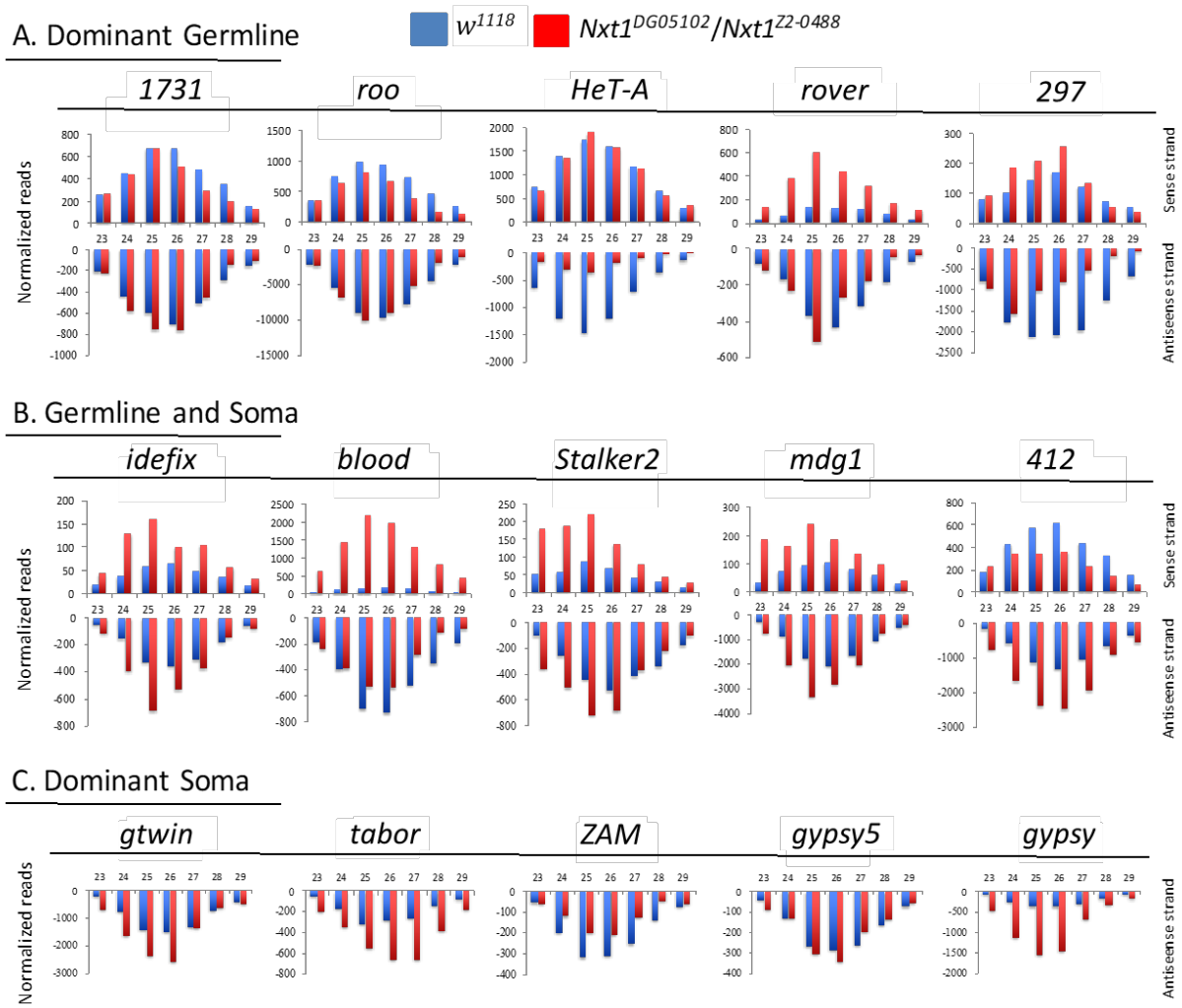


Figure 4.12: **Increase in piRNA reads mapping to transposons in Nxt1 mutant.** piRNA reads mapped to (A) Dominant Germline, (B) Germline and Soma, and (C) Dominant Soma transposons of wild type (blue) and Nxt1 trans-heterozygotes (red). Mapped piRNA reads are separated by sense strand (top) and antisense strand (bottom). piRNA reads were normalized per 1 million mapped miRNA reads.

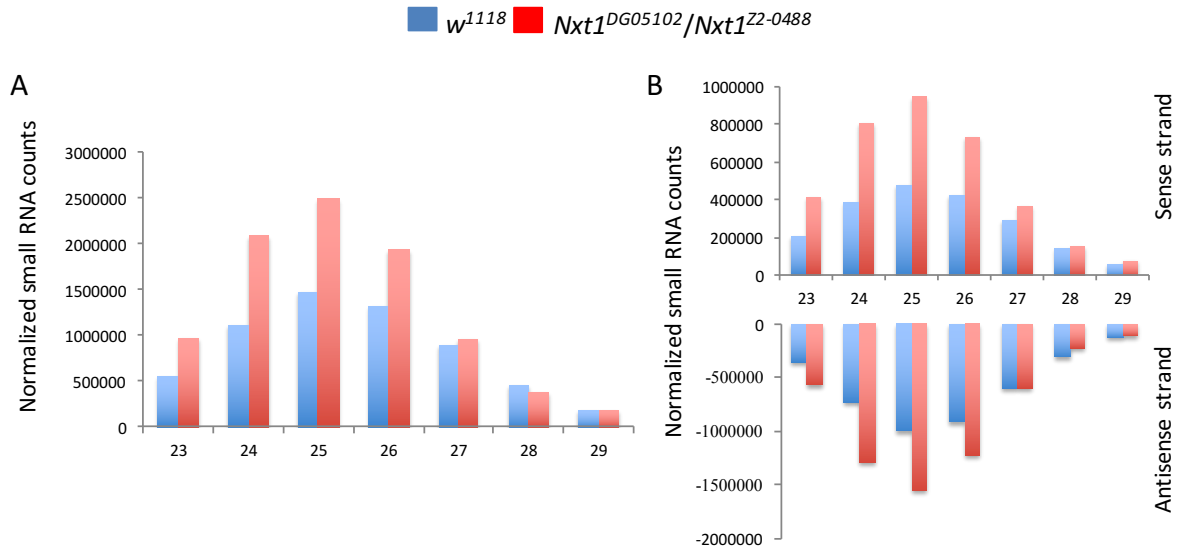


Figure 4.13: **2nd round of small RNA data showed increase in piRNA reads mapped to transposons.** A) Non-oxidized data showed an increase in reads mapping to Nxt1 trans-heterozygotes (red) compared to wild type (blue). B) piRNA reads separated by sense strand (top) and antisense strand (bottom). piRNA reads were normalized per 1 million mapped miRNA reads.

The UCSC genome browser was used to visualize the unique piRNAs on the plus and minus strand of three major piRNA clusters *flam*, *42AB* and *20A* (Figure 4.13). As expected from the transposon data, all three clusters showed more piRNAs mapping to Nxt1 trans-heterozygotes.

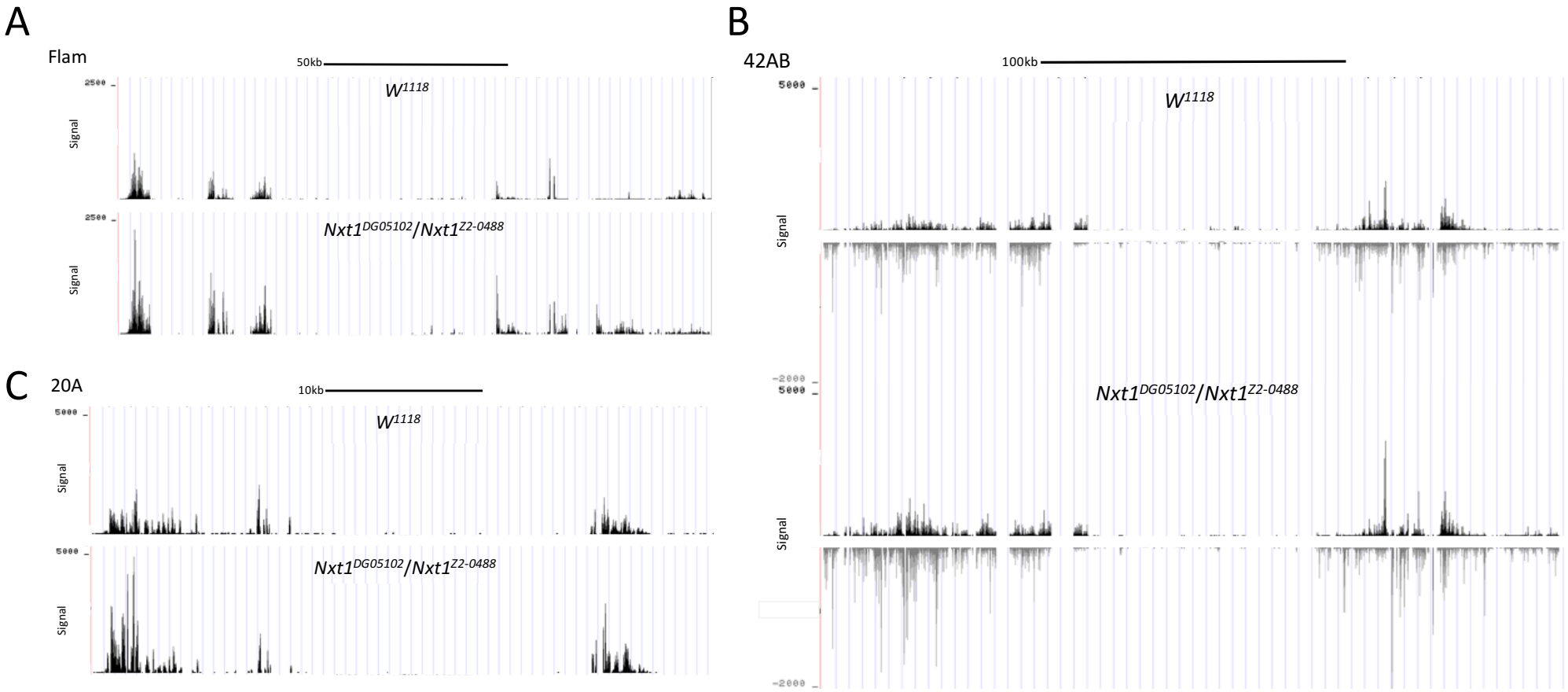


Figure 4.14: **Higher density of unique piRNA reads across three major piRNA clusters in Nxt1 trans-heterozygotes.** BigWig files from 2nd round of small RNA-seq were uploaded to UCSC genome browser for visualization of unique piRNA reads mapping to (A) *flam*, (B) *42AB*, and (C) *20A*. In Nxt1 trans-heterozygotes, a higher signal was observed in *flam*, *42AB* and *20A*.

To better interpret the data between the microarray and the both rounds of small RNA sequencing I looked at the top 20 most up regulated transposons, based on the microarray data, and compared to the normalized reads mapping of the 1st and 2nd small RNA rounds (Table 4.1). The 1st round of small RNA-seq showed that almost all transposons had reduced reads on both sense and antisense strands in Nxt1 trans-heterozygotes, which was resembling the transposon up regulation in the microarray data (Table 4.1). However, the 2nd round of small RNA sequencing identified that most transposons had an increase in reads on both sense and antisense strands in Nxt1 trans-heterozygotes (Table 4.1). This revealed that both small RNA-seq data sets were contradicting to each other based on the piRNA reads mapping to transposons. While the exact cause is unclear, the differences in preparation of small RNA-seq libraries might have an effect. As a result, it is unclear whether Nxt1 is involved in the transportation of precursor RNAs or silencing of transposons.

Table 4.1: **Top 20 up regulated transposons from microarray compared to 1st and 2nd round of small RNA-seq data**. piRNA reads were mapped to sense, antisense and both strands in wild type and compared to Nxt1 trans-heterozygotes. Red = less piRNA reads mapping to transposons in Nxt1 trans-heterozygotes. Green = more piRNA reads mapping to transposons in Nxt1 trans-heterozygotes.

Transposon name	Microarray (log2 fold change)	1st round of sequencing (4 days old)			2nd round of sequencing (7 days old)		
		Sense strand	Antisense strand	Both strands	Sense strand	Antisense strand	Both strands
(Log2 fold change in Nxt1)		(Reads mapping to Nxt1 vs. WT)			Reads mapping to Nxt1 vs. WT)		
gypsy	7.06	0.74x	0.75x	0.75x	4.26x	3.60x	3.76x
Burdock	6.65	0.77x	0.62x	0.66x	3.06x	1.16x	2.01x
ZAM	6.42	0.70x	0.80x	0.80x	1.59x	0.63x	0.73x
mdg1	5.97	0.73x	0.71x	0.71x	1.50x	1.08x	1.11x
HeT-A	5.65	0.87x	0.84x	0.85x	1.10x	0.16x	0.67x
springer	5.37	0.79x	0.86x	0.84x	1.95x	3.05x	2.65x
gtwin	4.98	0.67x	0.69x	0.69x	7.40x	0.97x	1.13x
gate	4.77	0.86x	0.82x	0.83x	1.77x	2.10x	1.96x
R2	4.74	0.97x	1.00x	1.00x	0.79x	0.80x	0.80x
copia	3.92	0.90x	0.75x	0.81x	0.65x	0.72x	0.69x
blood	3.76	0.85x	0.80x	0.80x	9.44x	0.70x	2.63x
412	3.56	0.52x	0.86x	0.82x	0.66x	1.95x	1.61x
accord	3.36	1.02x	1.06x	1.05x	0.87x	0.83x	0.84x
Transpac	3.27	0.88x	0.79x	0.81x	7.19x	2.33x	4.26x
flea	3.10	0.91x	0.89x	0.89x	7.87x	1.52x	3.09x
HMS-Beagle	3.04	0.84x	0.85x	0.85x	2.29x	1.63x	1.88x
Stalker2	2.90	0.87x	0.95x	0.94x	1.73x	1.10x	1.16x
gypsy5	2.74	0.86x	0.74x	0.74x	1.11x	0.68x	0.69x
17.6	2.60	0.99x	1.13x	1.11x	1.46x	1.36x	1.39x
Dm88	2.35	0.61x	0.73x	0.69x	3.13x	3.68x	3.46x

4.8 piRNAs in Nxt1 trans-heterozygotes are methylated

The 1st round of small RNA sequencing showed that there were less piRNAs mapping to transposons in Nxt1 trans-heterozygotes. One of the possibilities could be that piRNAs were not methylated, which could lead to potential degradation of piRNAs (Li et al. 2005; Yu et al. 2005). All piRNAs are methylated by Hen1 at their 3' end with an addition of 2' $-OCH_3$ (Saito et al. 2007). With the 2nd round of small RNA sequencing we oxidized our libraries. Oxidization includes the addition of sodium perodiate ($NaIO_4$) to the small RNAs. Small RNAs methylated at their 3' end are protected from reaction with $NaIO_4$, leaving only siRNAs and piRNAs (Seitz, Ghildiyal, and Zamore 2008). Length distribution across all transposons revealed no major

differences of the mapped piRNAs between WT and Nxt1 trans-heterozygotes (Figure 4.15A). Slightly reduced amount of piRNAs were mapped to Nxt1 trans-heterozygotes (Figure 4.15A). Reads separated to sense and antisense strands showed slightly less reads mapping to the antisense strand of piRNAs and slightly more reads mapping to the sense strand of piRNAs in Nxt1 trans-heterozygotes (Figure 4.15B). The methylation of piRNA indicated that piRNAs were unlikely to be degraded due to their protection status.

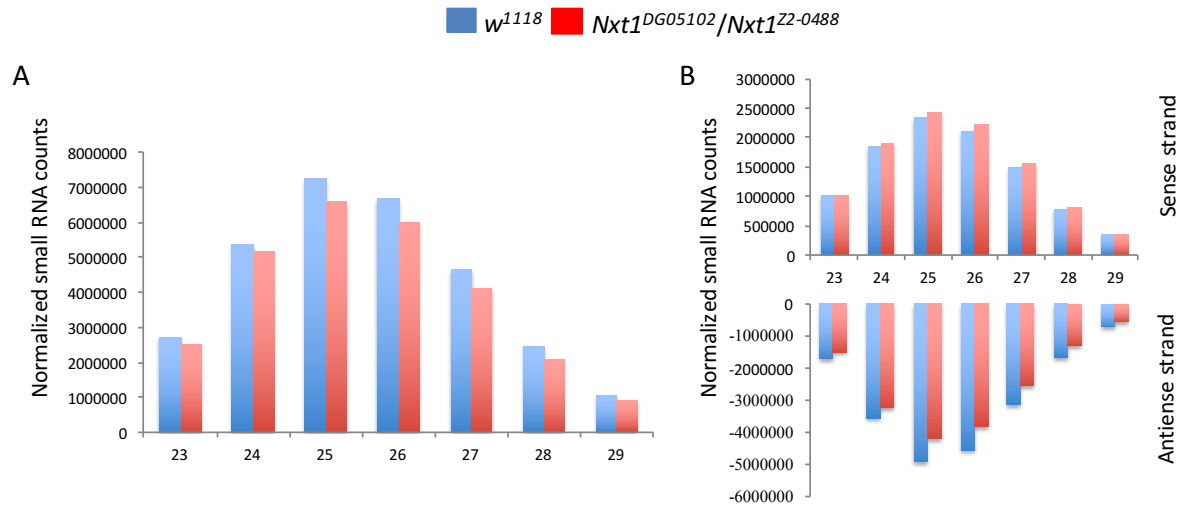


Figure 4.15: **2nd round of oxidized small RNA-seq data showed piRNAs were methylated in both genotypes.** A) piRNA reads mapping to transposons in Nxt1 trans-heterozygotes (red) compared to wild type (blue). B) piRNA reads separated by sense strand (top) and antisense strand (bottom). piRNA reads were normalized per 10 million uniquely mapped reads.

4.9 No length shifts observed with Piwi, Aub and AGO3 bound piRNAs

The first round of small RNA-seq showed a reduction in the length of total piRNAs in Nxt1 trans-heterozygotes (Figure 4.10). An alternative approach is to use immunoprecipitation (IP) sequencing to specifically isolate and identify all piRNAs binding to the PIWI proteins, Piwi, Aub and AGO3. Antibodies to each of these proteins are used to purify its associated piRNAs, which are then sequenced. It is known that on average piRNAs with a length of 26nt, 25nt and 24nt bind to Piwi, Aub and AGO3, respectively (Brennecke et al. 2007). IP sequencing for w^{1118} and Nxt1 trans-heterozygotes was performed by Rippei Hayashi (Brennecke lab). Length

distribution showed that Piwi had an average piRNA length of 25.87nt in wild type and 25.90nt in Nxt1 trans-heterozygotes (Figure 4.16A), Aub with an average piRNA length of 25.20nt in wild type and 25.07nt in Nxt1 trans-heterozygotes (Figure 4.16B), whereas AGO3 had an average piRNA length of 24.69nt in wild type and 24.65nt in Nxt1 trans-heterozygotes (Figure 4.16C). Thus, piRNA length was not affected in Nxt1 trans-heterozygotes. While the 1st round of small RNA-seq indicated that the piRNA length might be mildly affected, IP-sequencing provided a more in-depth analysis of piRNA lengths binding to Piwi, Aub and AGO3 and no clear differences in Nxt1 trans-heterozygotes were observed.

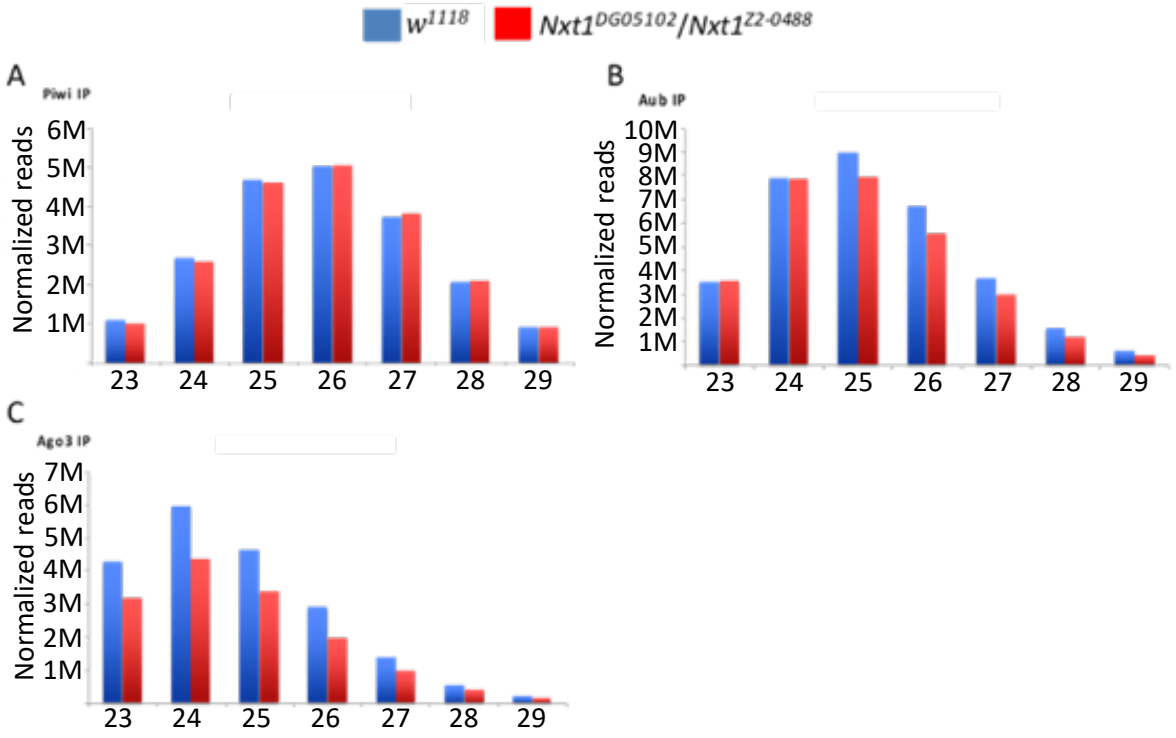


Figure 4.16: **No difference in piRNA length observed in Nxt1 trans-heterozygotes.** A) Piwi IP, B) Aub IP and C) AGO3 IP between wild type (blue) and Nxt1 trans-heterozygotes (red) showed no differences in the piRNA length. piRNA reads were normalized per 10 million uniquely mapped reads.

Chapter 4 - Discussion

4.10 Nxt1 affects the repression of transposons in the piRNA pathway

In the testis, expression of genes exclusively in the primary spermatocytes, are dependent on the testis-specific meiotic arrest complex (tMAC) (White-Cooper 2010). In Nxt1 trans-heterozygotes many genes were differentially expressed, and this was due to the testis-specific transcriptional regulation (Caporilli et al. 2013). Transcripts sensitive to the loss of Nxt1 were also dependent on tMAC for transcription, revealing a link between the testis-specific transcription regulation and the mRNA export pathway (Caporilli et al. 2013). The role in transcription of Nxt1 in testis and the sterility observed in female flies allowed us to focus at gene expression in the ovaries. We performed microarray data to look at the gene expression in 7 days old ovaries. In the testis, Nxt1 was linked with tMAC, however this complex is not present in the ovaries. In the Nxt1 trans-heterozygotes ovaries microarray, only 302 genes were more than 2-fold differentially expressed in ovaries, whereas in Nxt1 trans-heterozygotes testis, 485 genes were more than 16-fold down regulated (Caporilli et al. 2013). With this microarray, I also looked at the intron length, number of introns and gene length (data not shown). While the median intron length and number of introns was increased for both down and up regulated genes, no clear trend was observed as the carcass RNA-seq data where a correlation with the intron length and number of introns and the down regulation was observed (data not shown; Table 3.6). The median gene length was higher in down regulated genes, but again, no correlation like carcass RNA-seq data was observed (data not shown; Table 3.7). Looking more carefully, 30 transposons were found to be up regulated (Figure 4.1A). This was a first indication that loss of Nxt1 in the ovaries could have an effect in the piRNA pathway. The piRNA pathway is important in both ovaries and testis to prevent transposons to hop around the genome, which can lead to sterility. I looked at the up regulated transposons and found that transposons repressed in both germline and soma were affected (Figure 4.2). This concludes that Nxt1, in the ovaries, is involved in the piRNA pathway in both soma and germ cells. Microarray of the 0-1 days old testis in Nxt1 trans-heterozygotes revealed that some transposons were down regulated (data not shown). Many transposons are expressed in germ cells, where their movement can lead to heritable expansions in

numbers, whereas some transposons are either exclusively or additionally expressed in the somatic cells of the ovary (Mevel-Ninio et al. 2007; Brennecke et al. 2008).

4.10.1 Where in the piRNA pathway is Nxt1 involved?

The protein coding genes, which are known or suspected to take part in the piRNA pathway, were not differentially expressed, indicating that the role of Nxt1 is not in regulation of transcription of piRNA pathway protein coding genes (Figure 4.4). Nxt1 is involved in the RNA export processing, so I started to look at the piRNA precursor transcript. piRNA precursor transcripts are transcribed from large piRNA clusters. These transcripts are transported to the nuage in germ cells or Yb-bodies in soma cells (Malone et al. 2009; Saito et al. 2010). Expression of the three major piRNA clusters, *flam*, *42AB* and *20A* by qRT-PCR showed that all three transcripts were not differentially expressed (Figure 4.3). It was clearly shown that the precursor transcripts were transcribed at similar levels to wild type in Nxt1 trans-heterozygotes. These data indicated that, in Nxt1 trans-heterozygotes, three major piRNA precursor transcripts were transcribed at normal levels.

To further explore the involvement of Nxt1 in the piRNA pathway, small RNA sequencing was performed to look at the piRNA levels. In Nxt1 trans-heterozygotes piRNA levels were reduced, which also resulted in a reduction of the ping-pong piRNA pairs, while the U-bias at the 5'end of antisense piRNAs was maintained (Figure 4.5). Data also showed that both sense and antisense piRNAs were reduced in Nxt1 trans-heterozygotes (Figure 4.9). However, most interestingly, piRNAs were slightly shorter in our mutant (Figure 4.10). The 2nd round of small RNA-seq, however, revealed that transposons that were up regulated in 7 days old ovaries microarray and 4 days old ovaries from the 1st round of small RNA-seq were mostly more repressed in the 7 days old ovaries from the 2nd round of small RNA-seq data (Table 4.1). The discrepancies between the 1st and 2nd round of small RNA made it difficult to understand whether piRNA levels were affected or not. Data from the 1st round indicates that, given the role of Nxt1, the export of piRNA precursor transcripts is affected, whereas the 2nd round indicates that silencing of transposons is affected while piRNA levels are like wild type. This could mean that Nxt1 either interact with Nxf1 (export) or with Nxf2 (silencing) depending on which data set is true. The differences in both rounds of small RNA-seq data could be due to

the different kits used for library preparation (Figure 4.11). The presence of 4Ns at the adapters and the variability of the adapter sequence would reduce the spikes that can interfere with piRNA processing patterns (Dominik Handler, personal communication).

Shorter piRNAs are a specific phenotype caused by defects that occur towards the end the piRNA processing. Normally, pre-piRNAs, a longer version of mature piRNAs, are trimmed at the 3'end by Zucchini or Nibbler (Han et al. 2011; Liu, Abe, et al. 2011; Hayashi et al. 2016; Mohn, Handler, and Brennecke 2015), followed by methylation by Hen1 (Li et al. 2005; Yu et al. 2005). I looked at the microarray data for Zucchini and Nibbler and both genes were not differentially expressed. However, I did not look at their protein levels or if both proteins influenced piRNAs. If Hen1 does not methylate piRNAs, piRNAs are also dramatically shortened due to disruption of a delicate balance in the interplay between Nbr and Hen1 (Wang et al. 2016). With the 2nd round of small RNA-seq we oxidized the RNA to remove all small RNAs that are not methylated. In the normal situation of all short RNA species only piRNAs and siRNAs are methylated. If, in the mutant ovaries, piRNAs are not methylated, the levels will appear dramatically reduced compared to wild type. Interestingly, no big differences were seen when I compared our WT with Nxt1 trans-heterozygotes (Figure 4.13). Based on the oxidized data, and microarray data, I concluded that both piRNA trimming and methylation processes were probably occurring normally in the Nxt1 trans-heterozygous ovaries.

Nxt1 is an RNA export factor, and as a small protein with only a NTF2-like domain, it is unlikely, but not impossible, that Nxt1 has a moonlighting role outwith the RNA export pathway. It seems more than likely that Nxt1, with an Nxf protein (more explained in later), has a role in the transport of precursor transcripts. qRT-PCR confirmed that depletion of Nxt1 does not alter the expression levels of the three major piRNA precursor transcripts (Figure 4.3). How are the transposon levels affected in Nxt1 trans-heterozygotes? Nxt1 is involved in transporting RNAs from the nucleus to the cytoplasm. Potentially, Nxt1 could be involved in transporting the precursor transcripts. Since the precursor transcripts are transcribed in the Nxt1 trans-heterozygotes, it could be possible that the transcripts are not transported and get stuck in the nucleus. Single cell fluorescent in-situ hybridization of *flam* and *42AB* precursor transcripts have been performed that clearly showed the localization of these

transcripts near the nuclear pore complex (Mohn et al. 2014). Similarly, this experiment could be performed in the Nxt1 trans-heterozygotes to understand if the precursor transcripts are transported to the nuclear pore complex. Since *Nxt1* expression is lowered in the Nxt1 trans-heterozygotes, it is likely that precursor transcripts are not transported at a full potential, like mRNAs, if Nxt1 is indeed involved in this transportation.

4.10.2 The impact of genetic background differences in the piRNA pathway

The biggest challenge is to understand the discrepancies between the two rounds of small RNA-seq data, which generated different results, and thus made it difficult to interpret the data. Besides the discrepancies, Nxt1 trans-heterozygotes are in a different background than wild type.

4.10.2.1 The effects of different genomic backgrounds and piRNA sequencing

The wild type I used is w^{1118}, whereas our Nxt1 trans-heterozygotes are in a y^{14},w^{1118} background. To date, not much is known about the effect of piRNAs in different genomic backgrounds. w^{1118} background is a highly inbred line, which should be homozygous at most loci, whereas the Nxt1 trans-heterozygotes is a result of a cross between two lines, neither of which is inbred. As a result, these lines will have a different history of transposon activity, such as differences at the piRNA clusters, copy numbers and insertion sites of the transposons being silenced. Using two different *flamenco* alleles, ISO1A and Rev, revealed that the expression of *ZAM* and *Idefix* was affected in the Rev line due to a deletion that included the *ZAM* and *Idefix* copies (Zanni et al. 2013). To fully understand these differences is to perform whole genome sequencing of different genetic backgrounds and look whether there are any differences within the different piRNA clusters.

4.10.2.2 The effects of library preparation with different kits

Besides the differences in the genomic backgrounds, the differences in library preparation might also influence the data. The 1st round of small RNA-seq was prepared with the NEB small RNA kit and the 2nd round with the TruSeq small RNA library prep kit. As a result, more spikes were observed across the piRNA clusters with the NEB small RNA kit (Figure 4.11). These spikes were due to the two different approaches of the adaptors. With the 2nd round

of small RNA-seq, adaptors with 4 random nucleotides on the ends were used for ligation to the small RNAs. With small RNA-seq discrepancies can occur between results from sequencing, microarrays and qRT-PCR (Baker 2010; Git et al. 2010). Biases in small RNA-seq would undermine the sensitivity and accuracy that can be achieved by deep sequencing. For piRNAs, the T-bias at their 5'end is inferred through profiling (Brennecke et al. 2007). Discrepancies can arise only in the ligation or amplification steps and T4-RNA ligases are predominantly the cause of these discrepancies and can mediate sequence-specific ligations (Jayaprakash et al. 2011). The distinguishing feature of piRNAs is the bias for a T at position 1 and changes in the bias can indicate processing defects in the piRNA pathway. However, changes in the T-bias can be observed with different 3'ends of the 5'adapter (Jayaprakash et al. 2011) For the 2nd round of small RNA-seq we used a 5'adapter with 4Ns at the 3'end and a 3'adapter with 4N at the 5' end to minimize any biases as not only the substrate, but also the adapter sequence is now variable (Jayaprakash et al. 2011).

So, what can I conclude from both small RNA-seq rounds? The data from the 1st round and the qRT-PCR suggests that the export of precursor transcripts is affected (Nxt1/Nxf1), leading to less piRNAs and therefore an increase in TE expression. If the 2nd round of small RNA-seq is to be believed, then the effects of the piRNA pathway are with silencing (Nxt1/Nxf2) and not so much on piRNA populations.

4.10.3 With which Nxf protein is Nxt1 interacting?

Drosophila has four annotated Nxf proteins and three of them have an NTF2-like domain to which Nxt1 is predicted to bind (Figure 4.17). So, this leads to a very simple question: Does Nxt1 interacts with Nxf1 in the piRNA pathway or are any of the other Nxf proteins involved? In *C. elegans* two Nxf proteins exist and four in *Drosophila* (Izaurralde 2002) . In humans, four genes related to TAP (NXF1) and one pseudogene has been reported, but only two that encode ORFs (NXF2 and NXF3) (Herold et al. 2000). Human NXF1 is, like *Drosophila* Nxf1, involved in the RNA export pathway (Herold, Klymenko, and Izaurralde 2001). Human NXF2 can stimulate RNA export, which was observed with an assay using human 293 cells that were co-transfected with pDM138 reporter plasmid and plasmids encoding NXF2, p15-1 (NXT1) and p15-2 (NXT2) (Herold et al. 2000). NXF3 has been shown to activate the nuclear export of

tethered RNA molecules by recruitment of the Crm1 nuclear export factor (Yang et al. 2001). Both proteins are expressed in testis, which indicates that they have the potential to act as tissue-specific export factors (Herold, Klymenko, and Izaurralde 2001; Yang et al. 2001). *Drosophila* Nxf2 and Nxf3 are expressed ubiquitously in embryos using different cell lines, and depletion of Nxf2 or Nxf3 does not influence cell growth when compared to untreated cells or cells transfected with GFP dsRNA (Herold, Klymenko, and Izaurralde 2001).

In *Drosophila*, out of the four Nxf proteins, Nxf1-4, Nxf1-3 have a NTF2-like domain required for binding to Nxt1. Nxf1 and Nxf2 have UBA domains; this domain is important for binding to FG-repeat-containing nuclear pore proteins and helps to mediate nuclear shuttling (Valkov et al. 2012). Nxf4 is the only Nxf protein in *Drosophila* who lacks either of these domains (Figure 4.17). Nxf1 is the only Nxf protein with extensive data, well known to play a part in the RNA export pathway (Herold, Klymenko, and Izaurralde 2001; Corey et al. 2001; Wilkie et al. 2001). While Nxf2 and Nxf3 have an ubiquitous expression during the embryonic phase, Nxf2 has an observable expression in the ovaries in adult stage. Both were proposed to have a more specific role, rather than involved in a broad process (Herold, Klymenko, and Izaurralde 2001; Corey et al. 2001; Wilkie et al. 2001). Nxf4 lacks several domains and is unlikely to bind to Nxt1. FlyAtlas data showed that Nxf4 is only expressed in the testis (FlyAtlas Anatomy Microarray).

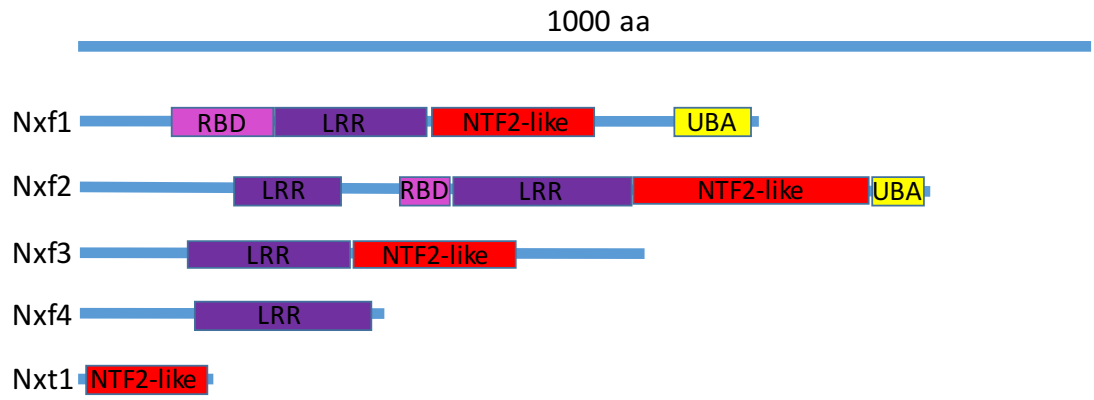


Figure 4.17: **Overview of domains in Nxt/Nxf proteins.** Overview of all the domains in the Nxt/Nxf proteins shows that Nxf1-3 are predicted to bind to Nxt1 via the NTF2-like domain. Nxf1 and Nxf2 have UBA domains that are required to mediate nuclear shuttling. Nxf3 has the ability to bind Nxt1 but lacks the UBA domain. Nxf4 is the smallest Nxf protein and lacks most of the domains. RBD= RNA binding domain, UBA= ubiquitin-associated, LRR= leucine-rich repeat, NTF2-like= nuclear transport factor 2-like.

Our collaborator, Julius Brennecke, tagged Nxt1 and Nxf1-3 with GFP and performed FISH to look at the localization in germ line ovaries together with Rhi for defining dual strand piRNA clusters, Wheat Germ Agglutinin (WGA) for nuclear pore complex and DAPI for DNA (Figure 4.18). Nxt1 has a strong expression in the germ line, which co-localizes with Rhi and WGA (Figure 4.18A-E). This co-localization indicates that Nxt1 localizes to the nuclear pore complex as expected with its role in the RNA export, but also localizes with Rhi which defines the Rhino-Deadlock-Cutoff complex required for dual strand piRNA cluster transportation (Mohn et al. 2014). Nxf1, known to interact with the NPC has a clear co-localization with WGA (Figure 4.18F-I). Nxf2 is strongly expressed in the soma and has a weak ubiquitous expression in the germ line. Nxf2 has a low expression in the germ line and has a higher expression observed in the somatic follicle cells (Figure 4.18J). Nxf3 co-localizes, like Nxt1, with Rhi and weakly with WGA (Figure 4.18O-S). The co-localization with Rhi was expected as Nxf3 has a potential role in dual strand piRNA cluster transportation (Julius Brennecke, personal communication). Our collaborators have focused on Nxf2 and 3, and based on their data, we have been able to speculate if these Nxf proteins have a role in the piRNA pathway. This also includes data of both small RNA-seq rounds that I have from Nxt1.

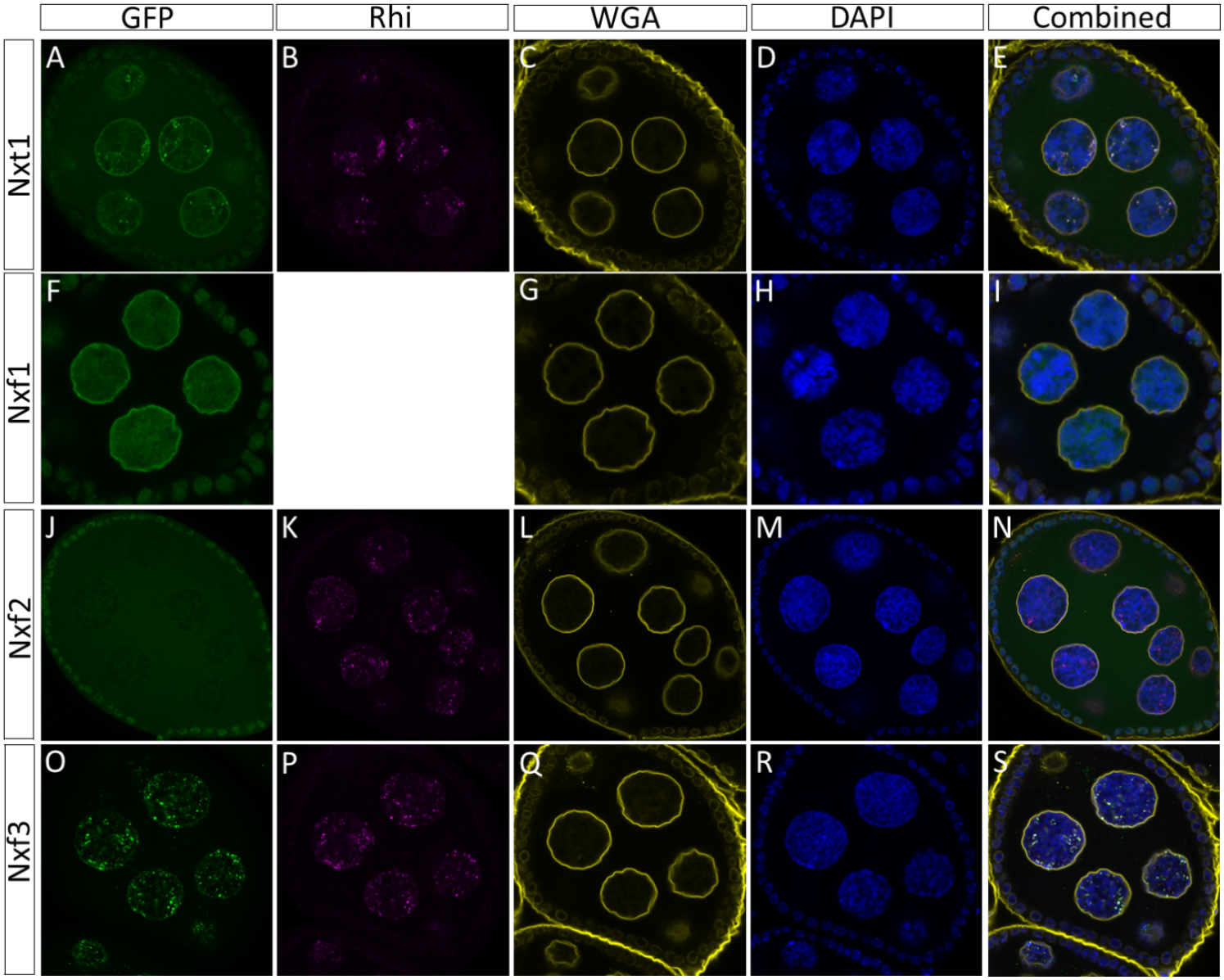


Figure 4.18**: Nxt1 and Nxf3 showed co-localization with Rhi revealing potential role in dual strand precursor transportation**. All samples were stained with Rhi for defining dual strand piRNA clusters in magenta, WGA for nuclear pore complex in yellow, DAPI for DNA in blue and GFP for Nxt1, Nxf1-3 in green. A-E) Nxt1 co-localizes to the nuclear pore complex with WGA staining, but also with Rhi foci. F-I) Nxf1 highly co-localizes with WGA, J-N) Nxf2 is highly localized to the soma and has a weak expression in the germ line. In the germ line, Nxf2 is more ubiquitously expressed and showed no co-localization with Rhi. O-S) Nxf3 has co-localization with WGA, and with Rhi. Images were kindly provided from the Brennecke lab for visualization.

The 2nd round of small RNA-seq was quite different than the 1st round. For the 2nd round we prepared the small RNA libraries ourselves, whereas small RNA libraries of the 1st round where prepared by the centre for genomic research at the university of Liverpool. For now, the focus

is set on the 2nd round of small RNA-seq data, since I had full control on the library preparation process. Based on these data, I must think about the effect of the different genetic backgrounds. Our wild type is w^{1118}, whereas our Nxt1 trans-heterozygotes is y^{14},w^{1118}. As explained earlier, different *flam* alleles influence TE silencing and differences in genetic background will have a different history of transposon activity (Zanni et al. 2013). The small RNA-seq data showed that piRNA levels were increased in Nxt1 trans-heterozygotes. With the 2nd round of small RNA-seq data, several transposons (*rover*, *297*, *idefix*, *blood*, *Stalker2* and *mdg1*) had an increase in reads mapping to the sense strand, whereas some transposons had an increase in reads mapping to the antisense strand (*gtwin*, *tabor* and *gypsy*) in Nxt1 trans-heterozygotes (Figure 4.12). Most transposons up regulated in the 1st round of sequencing had more reads mapping to sense and/or antisense strand in Nxt1 trans-heterozygotes (Table 4.1). The differences in antisense piRNA might be dictated by genetic background, which affects mostly *flam* piRNA cluster (Julius Brennecke, personal communication). If piRNA population is not affected in both soma and germ line, then it seems more likely that the defect is in silencing, which is like the loss of Nxf2 that has no impact on piRNA populations in the soma and only mild ones in the germ line (Julius Brennecke, personal communication). The possibility of TEs that are de-repressed could force more ping-pong processing, resulting in more sense piRNAs.

Nxt1 interacting with Nxf1 in the piRNA pathway

Most experiments have been focused on Nxt1 interacting with Nxf1 in the RNA export pathway. Nxf1 knockdowns in ovarian somatic cells (OSCs), *flam* derived piRNAs, but not piRNAs from 3' UTR were strongly reduced, which likely indicates the importance of Nxf1 for efficient flamenco export (Julius Brennecke, personal communication). In somatic cells, the majority of piRNAs is derived from *flamenco*. More recently, Dennis and colleagues reported the involvement of Nxt1 and Nxf1 in *flam* RNA export (Dennis et al. 2016). Depletion of *Nxt1* and *Nxf1* in OSC led to TE de-repression and loss of Piwi nuclear localization (Dennis et al. 2016). Interestingly, *Nxt1* and *Nxf1* knockdowns resulted in an increase in proportion of *flam* precursor localization in the nucleus than in wild type *ISO1A* line (Dennis et al. 2016).

Nxt1 interacting with Nxf2 in the piRNA pathway

The role of Nxf2 is potentially involved in the transcriptional silencing branch downstream of Piwi in somatic as well as in germ line cells (Julius Brennecke, personal communication). A western blot was performed against Nxf2 and Nxf3 and it was found that Nxf2 was reduced in Nxt1 trans-heterozygotes (Julia Batki, personal communication). This would suggest that in the germ line, a reduction of Nxf2 has effects in the silencing, but not cluster export. While the 1st round of small RNA-seq data showed a reduction in the piRNAs, the 2nd round clearly showed the opposite. This makes it difficult to understand the role of Nxt1, however, if piRNA levels are not altered in Nxt1 trans-heterozygotes, then Nxt1 might have an important link with Nxf2. As explained earlier, Nxf2 has a role in silencing, and has no impact on piRNA populations in the soma, and only very mild on a few ones in the germ line (for example *HeT-A*; Julius Brennecke, personal communication). If you now de-silence the transposable elements, it might be possible that much more ping-pong amplification is forced; hence you can get more sense piRNAs for those elements that are strongly de-repressed. If this hypothesis holds, it could possibly explain the increase in sense piRNAs in Nxt1 trans-heterozygotes (Figure 4.12).

Nxt1 interacting with Nxf3 in the piRNA pathway

So far, evidence has suggested that Nxf3 is a dedicated exporter for dual-strand piRNA precursor transcripts (Julius Brennecke, personal communication). Nxf3 null mutants (generated by CRISPR) have a selective loss of dual strand cluster derived piRNAs, which resembles one to one loss of Rhino (Julius Brennecke, personal communication). Nxf3 is highly expressed in the ovaries and co-localizes with Rhino in the nucleus (Figure 4.18O-S).

4.11 Aging of ovaries and the affects in the piRNA pathway

In Nxt1 trans-heterozygotes females, old ovaries at 20 days were degenerating at lot, while the young ovaries looked normal early on (Simona Caporilli, personal communication). During the first round of small RNA sequencing, I looked at three different time points (4, 14 and 20 days) to understand the effects of transposon repression in the piRNA pathway during aging. If transposons are de-repressed, genome instability occurs, leading to sterility. The wild type

piRNA read levels between 4 and 14 days were similar, however, Nxt1 trans-heterozygotes showed a slight decline at 4 days and a massive decline in 14 days (Figure 4.5). In 20 days, however, the wild type piRNA reads were significantly lower than earlier time points with slightly less reads observed in the Nxt1 trans-heterozygotes (Figure 4.5). The observation was interesting that at first, piRNA reads in Nxt1 trans-heterozygotes were slowly declining over age, and second, in wild type, the piRNA reads were declining significantly between 14 and 20 days of age. The phenotype was not published before until in 2016, Wang and colleagues looked at 3 (newly eclosed adults), 15 (young), 30 (mid-age) and 45 (old) days old ovaries and found that in older flies the piRNA levels were slightly reduced and piRNAs became shorter over time (Wang et al. 2016). Interestingly, the reduction in piRNA counts in 30 and 45 days was much milder compared to our 20 days small RNA-seq data ((Wang et al. 2016); Figure 4.5). The reduction in piRNA length from Wang and colleagues is very mild and like our observed data ((Wang et al. 2016); Figure 4.10).

The reduction in piRNA length in Nxt1 trans-heterozygotes was observable in all three time points, and for wild type the piRNA length was reduced only at 20 days of age. With the IP-seq data, I looked at the piRNAs binding to Piwi, Aub and AGO3 (Figure 4.15). For this, we used 7 days old ovaries and found no significant shift in the piRNA length in Piwi, Aub and AGO3. Piwi showed an average piRNA length of 25.87nt in wild type and 25.90nt in Nxt1 trans-heterozygotes, while Aub had an average of 25.20nt in wild type and 25.07nt in Nxt1 trans-heterozygotes and AGO3 had an average length of 24.69nt in wild type and 24.65nt in Nxt1 trans-heterozygotes (Figure 4.16). With these data, I can confirm that piRNA length is not shorter from the IP-seq data. The reduction that I observed in the 1st round of small RNA-seq data might be from an unknown artefact. Since libraries of the 1st small RNA-seq round were prepared differently than the 2nd round of small RNA-seq and IP-seq, I do not know exactly how much of an effect this could have on the data analysis. However, it is still possible that the piRNA lengths are affected in the older ovaries, especially 20 days of age where the piRNA length in wild type is reduced significantly. IP-seq from 14 days and 20 days of age should be performed to understand if piRNA length reduction is occurring over age.

The other phenotype observed was the significant reduction of piRNA levels at 20 days of age for both wild type and Nxt1 trans-heterozygotes (Figure 4.5). This phenotype was published

in 2016 but was not as significant as our data (Wang et al. 2016). Our data clearly showed a big reduction of piRNA levels in wild type and an even bigger reduction in Nxt1 trans-heterozygotes (Figure 4.5). Wang and colleagues showed that over age piRNAs undergo a progressive shortening from the 3'end by an increase of Nbr activity (Wang et al. 2016). At 7 days of age, the IP-seq data showed no significant shifts in piRNA lengths binding to Piwi, Aub and AGO3 (Figure 4.15), whereas the RNA-seq data of 4 days of age showed a slight decrease in piRNA length (Figure 4.10). Since I do not have IP-seq data from 14 and 20 days of age, it is unclear if piRNA length is really affected. It is also unclear if Nbr activity is increased at older age in w^{1118} and Nxt1 trans-heterozygotes. These data, however, gives a further insight about the connection between adult fertility and the piRNA pathway.

Conclusions.

1) **Reduction of Abba in Nxt1 trans-heterozygotes affect larva muscle integrity**

- RNA-seq data has revealed that the lack of head eversion is unlikely to be the caused by ecdysone and ecdysone-responsive genes. Only four ecdysone-responsive genes were affected, and its current roles would not explain the head eversion phenotype.

- Third instar larvae Nxt1 trans-heterozygotes were slower than wild type and was caused by muscle degeneration. The muscle degeneration occurs throughout the whole larva, and the air bubble is unlikely to migrate to the anterior end by the contracting muscles.

- Nxt1 trans-heterozygote larvae deprived from food towards the end of the second instar have shown that the muscle degeneration defect only occurs during the final growth stage in the third instar larvae.

- Rescuing the expression of *abba* with the UAS/gal4 system in muscles specifically in third instar Nxt1 trans-heterozygote larvae rescues the muscle degeneration in third instar larvae, but not the semi-lethality in pupae. The absence of head eversion leads to lethality to most of these pupae.

- The ratio of expression and activity between *abba* and *mib2* in Nxt1 trans-heterozygote larvae showed that the muscle integrity is more complicated in the larva than just the gene dosage between two genes.

- Rescuing *abba* via the UAS/gal4 system throughout the whole larvae rescued muscle integrity, but also increased pupa viability from ~20% to ~70%. This indicates that Abba most like has roles in other tissues required for pupa processes and its survivability.

2) **Observed correlation between total intron length and down regulated genes in larva carcass**

- RNA-seq data from larvae carcass revealed that genes down regulated were genes with several introns, high total intron length and have longer gene transcripts. EJC, who has been shown to affect genes with large introns, was mildly down regulated in Nxt1 trans-heterozygotes.

- The sensitivity of transcript expression from genes with long introns was not seen in whole larvae, most likely, due to the bias of testis-specific genes. Similarly, in ovaries, no clear trend was observed, as in the carcass RNA-seq data.

- qRT-PCR of different regions in *abba* confirmed that loss of Nxt1 affects splicing activity across the whole transcript. Unclear if *abba* is the only exception or if it is specifically to genes with large and many introns.

3) **Increasing *Nxt1* in Nxt1 trans-heterozygotes rescues semi-lethality in pupa significantly**

- Rescuing *Nxt1* via the UAS/gal4 system revealed that with the muscle-specific driver, mef2gal4, muscle degeneration is partially rescued, but not the pupa viability. Using the moderate ubiquitous driver, armgal4, rescues the muscle degeneration and the pupa viability significantly.

4) **Nxt1 plays an important role in the piRNA pathway**

- Microarray, qRT-PCR and 1st round of small RNA-seq have shown that loss of Nxt1 affected the transposon repression in the piRNA pathway. However, the 2nd round of small RNA-seq showed that most transposons affected in the microarray, qRT-PCR and 1st round of small RNA-seq were not affected.

- The different genetic backgrounds between wild type and Nxt1 trans-heterozygotes made it difficult to analyse the small RNA-seq data. Two rounds of small RNA-seq were performed and data analysis between the two rounds was quite to the contrary. Based

on the current data, Nxt1 either affects the transport of precursor transcript with Nxf1 and/or Nxf3 or affects silencing with Nxf2.

- Nxt1 is known as an RNA export factor and most likely affects the piRNA pathway during the transportation of the precursor transcripts to the processing centres in the nuage or Yb-bodies. It is unclear with which Nxf protein Nxt1 interacts. Data from collaborators has shown that Nxf1, 2 and 3 are all likely to interact with Nxt1 and, most likely, also in the piRNA pathway. Nxf1 for flamenco export, Nxf2 for silencing and Nxf3 for dual-strand export.

There is still more to be researched about Nxt1 and its role in the RNA export pathway and piRNA pathway in *Drosophila* and other organisms. This thesis gave more detail into the loss of Nxt1 during metamorphosis that showed a specific defect in the muscles and in the ovaries where the piRNA pathway was highly affected. In the future, more research of Nxt1 can be done to further understand its importance in *Drosophila melanogaster* and other organisms.

References.

Anand, A., and T. Kai. 2012. 'The tudor domain protein kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila', *Embo j*, 31: 870-82.
Aravin, A. A., M. S. Klenov, V. V. Vagin, F. Bantignies, G. Cavalli, and V. A. Gvozdev. 2004. 'Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line', *Mol Cell Biol*, 24: 6742-50.
Aravin, A. A., R. Sachidanandam, D. Bourc'his, C. Schaefer, D. Pezic, K. F. Toth, T. Bestor, and G. J. Hannon. 2008. 'A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice', *Mol Cell*, 31: 785-99.
Aravin, A. A., R. Sachidanandam, A. Girard, K. Fejes-Toth, and G. J. Hannon. 2007. 'Developmentally regulated piRNA clusters implicate MILI in transposon control', *Science*, 316: 744-7.
Aravin, A., D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, N. Iovino, P. Morris, M. J. Brownstein, S. Kuramochi-Miyagawa, T. Nakano, M. Chien, J. J. Russo, J. Ju, R. Sheridan, C. Sander, M. Zavolan, and T. Tuschl. 2006. 'A novel class of small RNAs bind to MILI protein in mouse testes', *Nature*, 442: 203-7.
Ashburner, M. 1990. 'Puffs, genes, and hormones revisited', *Cell*, 61: 1-3.
Ashburner, M., C. Chihara, P. Meltzer, and G. Richards. 1974. 'Temporal control of puffing activity in polytene chromosomes', *Cold Spring Harb Symp Quant Biol*, 38: 655-62.
Ashe, A., A. Sapetschnig, E. M. Weick, J. Mitchell, M. P. Bagijn, A. C. Cording, A. L. Doebley, L. D. Goldstein, N. J. Lehrbach, J. Le Pen, G. Pintacuda, A. Sakaguchi, P. Sarkies, S. Ahmed, and E. A. Miska. 2012. 'piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans', *Cell*, 150: 88-99.
Ashton-Beaucage, D., and M. Therrien. 2011. 'The exon junction complex: a splicing factor for long intron containing transcripts?', *Fly (Austin)*, 5: 224-33.
Askjaer, P., A. Bachi, M. Wilm, F. R. Bischoff, D. L. Weeks, V. Ogniewski, M. Ohno, C. Niehrs, J. Kjems, I. W. Mattaj, and M. Fornerod. 1999. 'RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase', *Mol Cell Biol*, 19: 6276-85.
Atkinson, R. A., C. Joseph, F. Dal Piaz, L. Birolo, G. Stier, P. Pucci, and A. Pastore. 2000. 'Binding of alpha-actinin to titin: implications for Z-disk assembly', *Biochemistry*, 39: 5255-64.
Bagijn, M. P., L. D. Goldstein, A. Sapetschnig, E. M. Weick, S. Bouasker, N. J. Lehrbach, M. J. Simard, and E. A. Miska. 2012. 'Function, targets, and evolution of Caenorhabditis elegans piRNAs', *Science*, 337: 574-8.
Bai, J., J. H. Hartwig, and N. Perrimon. 2007. 'SALS, a WH2-domain-containing protein, promotes sarcomeric actin filament elongation from pointed ends during Drosophila muscle growth', *Dev Cell*, 13: 828-42.
Bainbridge, S. P., and M. Bownes. 1981. 'Staging the metamorphosis of Drosophila melanogaster', *J Embryol Exp Morphol*, 66: 57-80.
Baker, M. 2010. 'MicroRNA profiling: separating signal from noise', *Nat Methods*, 7: 687-92.
Bashirullah, A., G. Lam, V. P. Yin, and C. S. Thummel. 2007. 'dTrf2 is required for transcriptional and developmental responses to ecdysone during Drosophila metamorphosis', *Dev Dyn*, 236: 3173-9.
Bate, M. 1990. 'The embryonic development of larval muscles in Drosophila', *Development*, 110: 791-804.
———. 1993. *The mesoderm and its derivatives* (Cold Spring Harbor Laboratory Press).

Batista, P. J., J. G. Ruby, J. M. Claycomb, R. Chiang, N. Fahlgren, K. D. Kasschau, D. A. Chaves, W. Gu, J. J. Vasale, S. Duan, D. Conte, Jr., S. Luo, G. P. Schroth, J. C. Carrington, D. P. Bartel, and C. C. Mello. 2008. 'PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans', *Mol Cell*, 31: 67-78.

Baylies, M. K., M. Bate, and M. Ruiz Gomez. 1998. 'Myogenesis: a view from Drosophila', *Cell*, 93: 921-7.

Beadle, G. W., E. L. Tatum, and C. W. Clancy. 1938. 'Food level in relation to rate of development and eye pigmentation in Drosophilha melanogaster', *Bio. Bull.*, 75: 447-62.

Beer, J., G. M. Technau, and J. A. Campos-Ortega. 1987. 'Lineage analysis of transplanted individual cells in embryos of Drosophila melanogaster : IV. Commitment and proliferative capabilities of mesodermal cells', *Roux Arch Dev Biol*, 196: 222-30.

Bernad, R., H. van der Velde, M. Fornerod, and H. Pickersgill. 2004. 'Nup358/RanBP2 attaches to the nuclear pore complex via association with Nup88 and Nup214/CAN and plays a supporting role in CRM1-mediated nuclear protein export', *Mol Cell Biol*, 24: 2373-84.

Billi, A. C., M. A. Freeberg, A. M. Day, S. Y. Chun, V. Khivansara, and J. K. Kim. 2013. 'A conserved upstream motif orchestrates autonomous, germline-enriched expression of Caenorhabditis elegans piRNAs', *PLoS Genet*, 9: e1003392.

Black, B. L., and E. N. Olson. 1998. 'Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins', *Annu Rev Cell Dev Biol*, 14: 167-96.

Boeke, J. D., and V. G. Corces. 1989. 'Transcription and reverse transcription of retrotransposons', *Annu Rev Microbiol*, 43: 403-34.

Bour, B. A., M. A. O'Brien, W. L. Lockwood, E. S. Goldstein, R. Bodmer, P. H. Taghert, S. M. Abmayr, and H. T. Nguyen. 1995. 'Drosophila MEF2, a transcription factor that is essential for myogenesis', *Genes Dev*, 9: 730-41.

Bozzetti, M. P., L. Fanti, S. Di Tommaso, L. Piacentini, M. Berloco, P. Tritto, and V. Specchia. 2012. 'The "Special" crystal-Stellate System in Drosophila melanogaster Reveals Mechanisms Underlying piRNA Pathway-Mediated Canalization', *Genet Res Int*, 2012: 324293.

Bozzetti, M. P., S. Massari, P. Finelli, F. Meggio, L. A. Pinna, B. Boldyreff, O. G. Issinger, G. Palumbo, C. Ciriaco, S. Bonaccorsi, and et al. 1995. 'The Ste locus, a component of the parasitic cry-Ste system of Drosophila melanogaster, encodes a protein that forms crystals in primary spermatocytes and mimics properties of the beta subunit of casein kinase 2', *Proc Natl Acad Sci U S A*, 92: 6067-71.

Brennan, C. M., I. E. Gallouzi, and J. A. Steitz. 2000. 'Protein ligands to HuR modulate its interaction with target mRNAs in vivo', *J Cell Biol*, 151: 1-14.

Brennecke, J., A. A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, and G. J. Hannon. 2007. 'Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila', *Cell*, 128: 1089-103.

Brennecke, J., C. D. Malone, A. A. Aravin, R. Sachidanandam, A. Stark, and G. J. Hannon. 2008. 'An epigenetic role for maternally inherited piRNAs in transposon silencing', *Science*, 322: 1387-92.

Budnik, V., M. Gorczyca, and A. Prokop. 2006. *Selected methods for the anatomical study of drosophila embryonic and larval neuromuscular junctions* (International Review of Neurobiology).

Cai, D., J. D. Frantz, N. E. Tawa, Jr., P. A. Melendez, B. C. Oh, H. G. Lidov, P. O. Hasselgren, W. R. Frontera, J. Lee, D. J. Glass, and S. E. Shoelson. 2004. 'IKKbeta/NF-kappaB activation causes severe muscle wasting in mice', *Cell*, 119: 285-98.

Caporilli, S., Y. Yu, J. Jiang, and H. White-Cooper. 2013. 'The RNA export factor, Nxt1, is required for tissue specific transcriptional regulation', *PLoS Genet*, 9: e1003526.

Capovilla, M., Z. Kambris, and J. Botas. 2001. 'Direct regulation of the muscle-identity gene apterous by a Hox protein in the somatic mesoderm', *Development*, 128: 1221-30.

Carmell, M. A., A. Girard, H. J. van de Kant, D. Bourc'his, T. H. Bestor, D. G. de Rooij, and G. J. Hannon. 2007. 'MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline', *Dev Cell*, 12: 503-14.

Carmell, M. A., Z. Xuan, M. Q. Zhang, and G. J. Hannon. 2002. 'The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis', *Genes Dev*, 16: 2733-42.

Carmena, A., S. Gisselbrecht, J. Harrison, F. Jimenez, and A. M. Michelson. 1998. 'Combinatorial signaling codes for the progressive determination of cell fates in the Drosophila embryonic mesoderm', *Genes Dev*, 12: 3910-22.

Cecere, G., G. X. Zheng, A. R. Mansisidor, K. E. Klymko, and A. Grishok. 2012. 'Promoters recognized by forkhead proteins exist for individual 21U-RNAs', *Mol Cell*, 47: 734-45.

Chakraborty, S. 2016. *RNA-seq assembler artifacts can bias expression counts and differential expression analysis - case study on the chickpea transcriptome emphasizes importance of freely accessible data for reproducibility [version 2; referees: 2 not approved]*.

Chavez, V. M., G. Marques, J. P. Delbecque, K. Kobayashi, M. Hollingsworth, J. Burr, J. E. Natzle, and M. B. O'Connor. 2000. 'The Drosophila disembodied gene controls late embryonic morphogenesis and codes for a cytochrome P450 enzyme that regulates embryonic ecdysone levels', *Development*, 127: 4115-26.

Cheng, H., K. Dufu, C. S. Lee, J. L. Hsu, A. Dias, and R. Reed. 2006. 'Human mRNA export machinery recruited to the 5' end of mRNA', *Cell*, 127: 1389-400.

Clark, K. A., A. S. McElhinny, M. C. Beckerle, and C. C. Gregorio. 2002. 'Striated muscle cytoarchitecture: an intricate web of form and function', *Annu Rev Cell Dev Biol*, 18: 637-706.

Cloonan, N., A. R. Forrest, G. Kolle, B. B. Gardiner, G. J. Faulkner, M. K. Brown, D. F. Taylor, A. L. Steptoe, S. Wani, G. Bethel, A. J. Robertson, A. C. Perkins, S. J. Bruce, C. C. Lee, S. S. Ranade, H. E. Peckham, J. M. Manning, K. J. McKernan, and S. M. Grimmond. 2008. 'Stem cell transcriptome profiling via massive-scale mRNA sequencing', *Nat Methods*, 5: 613-9.

Conlon, I., and M. Raff. 1999. 'Size control in animal development', *Cell*, 96: 235-44.

Corbin, V., A. M. Michelson, S. M. Abmayr, V. Neel, E. Alcamo, T. Maniatis, and M. W. Young. 1991. 'A role for the Drosophila neurogenic genes in mesoderm differentiation', *Cell*, 67: 311-23.

Cox, D. N., A. Chao, J. Baker, L. Chang, D. Qiao, and H. Lin. 1998. 'A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal', *Genes Dev*, 12: 3715-27.

Czech, B., J. B. Preall, J. McGinn, and G. J. Hannon. 2013. 'A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway', *Mol Cell*, 50: 749-61.

Das, P. P., M. P. Bagijn, L. D. Goldstein, J. R. Woolford, N. J. Lehrbach, A. Sapetschnig, H. R. Buhecha, M. J. Gilchrist, K. L. Howe, R. Stark, N. Matthews, E. Berezikov, R. F. Ketting, S. Tavare, and E. A. Miska. 2008. 'Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline', *Mol Cell*, 31: 79-90.

Demontis, F., and N. Perrimon. 2009. 'Integration of Insulin receptor/Foxo signaling and dMyc activity during muscle growth regulates body size in Drosophila', *Development*, 136: 983-93.

Deng, W., and H. Lin. 2002. 'miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis', *Dev Cell*, 2: 819-30.

Dennis, C., E. Brasset, A. Sarkar, and C. Vaury. 2016. 'Export of piRNA precursors by EJC triggers assembly of cytoplasmic Yb-body in Drosophila', *Nat Commun*, 7: 13739.

Dennis, C., V. Zanni, E. Brasset, A. Eymery, L. Zhang, R. Mteirek, S. Jensen, Y. S. Rong, and C. Vaury. 2013. '"Dot COM", a nuclear transit center for the primary piRNA pathway in Drosophila', *PLoS One*, 8: e72752.

Desset, S., N. Buchon, C. Meignin, M. Coiffet, and C. Vaury. 2008. 'In Drosophila melanogaster the COM locus directs the somatic silencing of two retrotransposons through both Piwi-dependent and -independent pathways', *PLoS One*, 3: e1526.

DiBello, P. R., D. A. Withers, C. A. Bayer, J. W. Fristrom, and G. M. Guild. 1991. 'The Drosophila Broad-Complex encodes a family of related proteins containing zinc fingers', *Genetics*, 129: 385-97.

Doggett, K., J. Jiang, G. Aleti, and H. White-Cooper. 2011. 'Wake-up-call, a lin-52 paralogue, and Always early, a lin-9 homologue physically interact, but have opposing functions in regulating testis-specific gene expression', *Dev Biol*, 355: 381-93.

Domsch, K., N. Ezzeddine, and H. T. Nguyen. 2013. 'Abba is an essential TRIM/RBCC protein to maintain the integrity of sarcomeric cytoarchitecture', *J Cell Sci*, 126: 3314-23.

Dreyfuss, G., V. N. Kim, and N. Kataoka. 2002. 'Messenger-RNA-binding proteins and the messages they carry', *Nat Rev Mol Cell Biol*, 3: 195-205.

Dutta, D., S. Anant, M. Ruiz-Gomez, M. Bate, and K. VijayRaghavan. 2004. 'Founder myoblasts and fibre number during adult myogenesis in Drosophila', *Development*, 131: 3761-72.

Dutta, D., and K. Vijayraghavan. 2006. *Metamorphosis and the formation of the adult musculature* (Landes Bioscience).

Elbashir, S. M., W. Lendeckel, and T. Tuschl. 2001. 'RNA interference is mediated by 21- and 22-nucleotide RNAs', *Genes Dev*, 15: 188-200.

Englmeier, L., M. Fornerod, F. R. Bischoff, C. Petosa, I. W. Mattaj, and U. Kutay. 2001. 'RanBP3 influences interactions between CRM1 and its nuclear protein export substrates', *EMBO Rep*, 2: 926-32.

Enriquez, J., H. Boukhatmi, L. Dubois, A. A. Philippakis, M. L. Bulyk, A. M. Michelson, M. Crozatier, and A. Vincent. 2010. 'Multi-step control of muscle diversity by Hox proteins in the Drosophila embryo', *Development*, 137: 457-66.

Epstein, N. D., and J. S. Davis. 2003. 'Sensing stretch is fundamental', *Cell*, 112: 147-50.

Farrell, E. R., J. Fernandes, and H. Keshishian. 1996. 'Muscle organizers in Drosophila: the role of persistent larval fibers in adult flight muscle development', *Dev Biol*, 176: 220-9.

Faulkner, G., G. Lanfranchi, and G. Valle. 2001. 'Telethonin and other new proteins of the Z-disc of skeletal muscle', *IUBMB Life*, 51: 275-82.

Feltzin, V. L., M. Khaladkar, M. Abe, M. Parisi, G. J. Hendriks, J. Kim, and N. M. Bonini. 2015. 'The exonuclease Nibbler regulates age-associated traits and modulates piRNA length in Drosophila', *Aging Cell*.

Fernandes, J., M. Bate, and K. Vijayraghavan. 1991. 'Development of the indirect flight muscles of Drosophila', *Development*, 113: 67-77.

Fletcher, J. C., and C. S. Thummel. 1995. 'The Drosophila E74 gene is required for the proper stage- and tissue-specific transcription of ecdysone-regulated genes at the onset of metamorphosis', *Development*, 121: 1411-21.

Forler, D., G. Rabut, F. D. Ciccarelli, A. Herold, T. Kocher, R. Niggeweg, P. Bork, J. Ellenberg, and E. Izaurralde. 2004. 'RanBP2/Nup358 provides a major binding site for NXF1-p15 dimers at the nuclear pore complex and functions in nuclear mRNA export', *Mol Cell Biol*, 24: 1155-67.

Fornerod, M., M. Ohno, M. Yoshida, and I. W. Mattaj. 1997. 'CRM1 is an export receptor for leucine-rich nuclear export signals', *Cell*, 90: 1051-60.

Fornerod, M., J. van Deursen, S. van Baal, A. Reynolds, D. Davis, K. G. Murti, J. Fransen, and G. Grosveld. 1997. 'The human homologue of yeast CRM1 is in a dynamic subcomplex with CAN/Nup214 and a novel nuclear pore component Nup88', *Embo j*, 16: 807-16.

Frank, D., C. Kuhn, H. A. Katus, and N. Frey. 2006. 'The sarcomeric Z-disc: a nodal point in signalling and disease', *J Mol Med (Berl)*, 84: 446-68.

Frasch, M. 1999. 'Controls in patterning and diversification of somatic muscles during Drosophila embryogenesis', *Curr Opin Genet Dev*, 9: 522-9.

Fribourg, S., I. C. Braun, E. Izaurralde, and E. Conti. 2001. 'Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor', *Mol Cell*, 8: 645-56.

Fribourg, S., and E. Conti. 2003. 'Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15', *EMBO Rep*, 4: 699-703.

Gao, M., D. T. Fritz, L. P. Ford, and J. Wilusz. 2000. 'Interaction between a poly(A)-specific ribonuclease and the 5' cap influences mRNA deadenylation rates in vitro', *Mol Cell*, 5: 479-88.

Girard, A., R. Sachidanandam, G. J. Hannon, and M. A. Carmell. 2006. 'A germline-specific class of small RNAs binds mammalian Piwi proteins', *Nature*, 442: 199-202.

Git, A., H. Dvinge, M. Salmon-Divon, M. Osborne, C. Kutter, J. Hadfield, P. Bertone, and C. Caldas. 2010. 'Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression', *Rna*, 16: 991-1006.

Gogvadze, E., and A. Buzdin. 2009. 'Retroelements and their impact on genome evolution and functioning', *Cell Mol Life Sci*, 66: 3727-42.

Goh, W. S., J. W. Seah, E. J. Harrison, C. Chen, C. M. Hammell, and G. J. Hannon. 2014. 'A genome-wide RNAi screen identifies factors required for distinct stages of C. elegans piRNA biogenesis', *Genes Dev*, 28: 797-807.

Gonatopoulos-Pournatzis, T., and H. Cowling V 2014. 'Cap-binding complex (CBC)', *Biochem J*, 457: 231-42.

Goriaux, C., S. Desset, Y. Renaud, C. Vaury, and E. Brasset. 2014. 'Transcriptional properties and splicing of the flamenco piRNA cluster', *EMBO Rep*, 15: 411-8.

Gregorio, C. C., K. Trombitas, T. Centner, B. Kolmerer, G. Stier, K. Kunke, K. Suzuki, F. Obermayr, B. Herrmann, H. Granzier, H. Sorimachi, and S. Labeit. 1998. 'The NH2

terminus of titin spans the Z-disc: its interaction with a novel 19-kD ligand (T-cap) is required for sarcomeric integrity', *J Cell Biol*, 143: 1013-27.
Greig, S., and M. Akam. 1993. 'Homeotic genes autonomously specify one aspect of pattern in the Drosophila mesoderm', *Nature*, 362: 630-2.
Grieneisen, M. L., J. T. Warren, and L. I. Gilbert. 1993. 'Early steps in ecdysteroid biosynthesis: evidence for the involvement of cytochrome P-450 enzymes', *Insect Biochem Mol Biol*, 23: 13-23.
Grivna, S. T., E. Beyret, Z. Wang, and H. Lin. 2006. 'A novel class of small RNAs in mouse spermatogenic cells', *Genes Dev*, 20: 1709-14.
Grosshans, H., K. Deinert, E. Hurt, and G. Simos. 2001. 'Biogenesis of the signal recognition particle (SRP) involves import of SRP proteins into the nucleolus, assembly with the SRP-RNA, and Xpo1p-mediated export', *J Cell Biol*, 153: 745-62.
Gunawardane, L. S., K. Saito, K. M. Nishida, K. Miyoshi, Y. Kawamura, T. Nagami, H. Siomi, and M. C. Siomi. 2007. 'A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila', *Science*, 315: 1587-90.
Hakeda, S., S. Endo, and K. Saigo. 2000. 'Requirements of Kettin, a giant muscle protein highly conserved in overall structure in evolution, for normal muscle function, viability, and flight activity of Drosophila', *J Cell Biol*, 148: 101-14.
Han, B. W., J. H. Hung, Z. Weng, P. D. Zamore, and S. L. Ameres. 2011. 'The 3'-to-5' exoribonuclease Nibbler shapes the 3' ends of microRNAs bound to Drosophila Argonaute1', *Curr Biol*, 21: 1878-87.
Han, B. W., W. Wang, C. Li, Z. Weng, and P. D. Zamore. 2015. 'Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production', *Science*, 348: 817-21.
Han, B. W., W. Wang, P. D. Zamore, and Z. Weng. 2015. 'piPipes: a set of pipelines for piRNA and transposon analysis via small RNA-seq, RNA-seq, degradome- and CAGE-seq, ChIP-seq and genomic DNA sequencing', *Bioinformatics*, 31: 593-5.
Handler, A. M. 1982. 'Ecdysteroid titers during pupal and adult development in Drosophila melanogaster', *Dev Biol*, 93: 73-82.
Handler, D., K. Meixner, M. Pizka, K. Lauss, C. Schmied, F. S. Gruber, and J. Brennecke. 2013. 'The genetic makeup of the Drosophila piRNA pathway', *Mol Cell*, 50: 762-77.
Handler, D., D. Olivieri, M. Novatchkova, F. S. Gruber, K. Meixner, K. Mechtler, A. Stark, R. Sachidanandam, and J. Brennecke. 2011. 'A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors', *Embo j*, 30: 3977-93.
Hauser, M. A., S. K. Horrigan, P. Salmikangas, U. M. Torian, K. D. Viles, R. Dancel, R. W. Tim, A. Taivainen, L. Bartoloni, J. M. Gilchrist, J. M. Stajich, P. C. Gaskell, J. R. Gilbert, J. M. Vance, M. A. Pericak-Vance, O. Carpen, C. A. Westbrook, and M. C. Speer. 2000. 'Myotilin is mutated in limb girdle muscular dystrophy 1A', *Hum Mol Genet*, 9: 2141-7.
Hayashi, R., J. Schnabl, D. Handler, F. Mohn, S. L. Ameres, and J. Brennecke. 2016. 'Genetic and mechanistic diversity of piRNA 3'-end formation', *Nature*, 539: 588-92.
Herold, A., T. Klymenko, and E. Izaurralde. 2001. 'NXF1/p15 heterodimers are essential for mRNA nuclear export in Drosophila', *Rna*, 7: 1768-80.
Herold, A., M. Suyama, J. P. Rodrigues, I. C. Braun, U. Kutay, M. Carmo-Fonseca, P. Bork, and E. Izaurralde. 2000. 'TAP (NXF1) belongs to a multigene family of putative RNA export factors with a conserved modular architecture', *Mol Cell Biol*, 20: 8996-9008.

Horwich, M. D., C. Li, C. Matranga, V. Vagin, G. Farley, P. Wang, and P. D. Zamore. 2007. 'The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC', *Curr Biol*, 17: 1265-72.

Huang, X. A., H. Yin, S. Sweeney, D. Raha, M. Snyder, and H. Lin. 2013. 'A major epigenetic programming mechanism guided by piRNAs', *Dev Cell*, 24: 502-16.

Hunter, R. B., E. Stevenson, A. Koncarevic, H. Mitchell-Felton, D. A. Essig, and S. C. Kandarian. 2002. 'Activation of an alternative NF-kappaB pathway in skeletal muscle during disuse atrophy', *Faseb j*, 16: 529-38.

Hur, J. K., Y. Luo, S. Moon, M. Ninova, G. K. Marinov, Y. D. Chung, and A. A. Aravin. 2016. 'Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in Drosophila', *Genes Dev*, 30: 840-55.

Hutten, S., and R. H. Kehlenbach. 2006. 'Nup214 is required for CRM1-dependent nuclear protein export in vivo', *Mol Cell Biol*, 26: 6772-85.

———. 2007. 'CRM1-mediated nuclear export: to the pore and beyond', *Trends Cell Biol*, 17: 193-201.

Ipsaro, J. J., A. D. Haase, S. R. Knott, L. Joshua-Tor, and G. J. Hannon. 2012. 'The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis', *Nature*, 491: 279-83.

Izaurralde, E. 2002. 'A novel family of nuclear transport receptors mediates the export of messenger RNA to the cytoplasm', *Eur J Cell Biol*, 81: 577-84.

Jayaprakash, A. D., O. Jabado, B. D. Brown, and R. Sachidanandam. 2011. 'Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing', *Nucleic Acids Res*, 39: e141.

Jimeno, S., A. G. Rondon, R. Luna, and A. Aguilera. 2002. 'The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability', *Embo j*, 21: 3526-35.

Kamminga, L. M., J. C. van Wolfswinkel, M. J. Luteijn, L. J. Kaaij, M. P. Bagijn, A. Sapetschnig, E. A. Miska, E. Berezikov, and R. F. Ketting. 2012. 'Differential impact of the HEN1 homolog HENN-1 on 21U and 26G RNAs in the germline of Caenorhabditis elegans', *PLoS Genet*, 8: e1002702.

Katahira, J., K. Straesser, T. Saiwaki, Y. Yoneda, and E. Hurt. 2002. 'Complex formation between Tap and p15 affects binding to FG-repeat nucleoporins and nucleocytoplasmic shuttling', *J Biol Chem*, 277: 9242-6.

Katahira, J., K. Strasser, A. Podtelejnikov, M. Mann, J. U. Jung, and E. Hurt. 1999. 'The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human', *Embo j*, 18: 2593-609.

Kehlenbach, R. H., A. Dickmanns, A. Kehlenbach, T. Guan, and L. Gerace. 1999. 'A role for RanBP1 in the release of CRM1 from the nuclear pore complex in a terminal step of nuclear export', *J Cell Biol*, 145: 645-57.

Kimura, K. I., and J. W. Truman. 1990. 'Postmetamorphic cell death in the nervous and muscular systems of Drosophila melanogaster', *J Neurosci*, 10: 403-1.

Kirino, Y., and Z. Mourelatos. 2007a. 'The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs', *Rna*, 13: 1397-401.

———. 2007b. 'Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini', *Nat Struct Mol Biol*, 14: 347-8.

Klattenhoff, C., H. Xi, C. Li, S. Lee, J. Xu, J. S. Khurana, F. Zhang, N. Schultz, B. S. Koppetsch, A. Nowosielska, H. Seitz, P. D. Zamore, Z. Weng, and W. E. Theurkauf. 2009. 'The

Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters', *Cell*, 138: 1137-49.
Knoll, R., M. Hoshijima, H. M. Hoffman, V. Person, I. Lorenzen-Schmidt, M. L. Bang, T. Hayashi, N. Shiga, H. Yasukawa, W. Schaper, W. McKenna, M. Yokoyama, N. J. Schork, J. H. Omens, A. D. McCulloch, A. Kimura, C. C. Gregorio, W. Poller, J. Schaper, H. P. Schultheiss, and K. R. Chien. 2002. 'The cardiac mechanical stretch sensor machinery involves a Z disc complex that is defective in a subset of human dilated cardiomyopathy', *Cell*, 111: 943-55.
Kohler, A., and E. Hurt. 2007. 'Exporting RNA from the nucleus to the cytoplasm', *Nat Rev Mol Cell Biol*, 8: 761-73.
Koundakjian, E. J., D. M. Cowan, R. W. Hardy, and A. H. Becker. 2004. 'The Zuker collection: a resource for the analysis of autosomal gene function in Drosophila melanogaster', *Genetics*, 167: 203-6.
Kramer, J. M., and B. E. Staveley. 2003. 'GAL4 causes developmental defects and apoptosis when expressed in the developing eye of Drosophila melanogaster', *Genet Mol Res*, 2: 43-7.
Kudryashova, E., J. Wu, L. A. Havton, and M. J. Spencer. 2009. 'Deficiency of the E3 ubiquitin ligase TRIM32 in mice leads to a myopathy with a neurogenic component', *Hum Mol Genet*, 18: 1353-67.
Kuramochi-Miyagawa, S., T. Kimura, T. W. Ijiri, T. Isobe, N. Asada, Y. Fujita, M. Ikawa, N. Iwai, M. Okabe, W. Deng, H. Lin, Y. Matsuda, and T. Nakano. 2004. 'Mili, a mammalian member of piwi family gene, is essential for spermatogenesis', *Development*, 131: 839-49.
Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. 'Ultrafast and memory-efficient alignment of short DNA sequences to the human genome', *Genome Biol*, 10: R25.
Lau, N. C., A. G. Seto, J. Kim, S. Kuramochi-Miyagawa, T. Nakano, D. P. Bartel, and R. E. Kingston. 2006. 'Characterization of the piRNA complex from rat testes', *Science*, 313: 363-7.
Le Hir, H., D. Gatfield, E. Izaurralde, and M. J. Moore. 2001. 'The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay', *Embo j*, 20: 4987-97.
Le Hir, H., E. Izaurralde, L. E. Maquat, and M. J. Moore. 2000. 'The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions', *Embo j*, 19: 6860-9.
Le Thomas, A., A. K. Rogers, A. Webster, G. K. Marinov, S. E. Liao, E. M. Perkins, J. K. Hur, A. A. Aravin, and K. F. Toth. 2013. 'Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state', *Genes Dev*, 27: 390-9.
Lee, H. C., W. Gu, M. Shirayama, E. Youngman, D. Conte, Jr., and C. C. Mello. 2012. 'C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts', *Cell*, 150: 78-87.
Lee, S. J. 2004. 'Regulation of muscle mass by myostatin', *Annu Rev Cell Dev Biol*, 20: 61-86.
Li, C., V. V. Vagin, S. Lee, J. Xu, S. Ma, H. Xi, H. Seitz, M. D. Horwich, M. Syrzycka, B. M. Honda, E. L. Kittler, M. L. Zapp, C. Klattenhoff, N. Schulz, W. E. Theurkauf, Z. Weng, and P. D. Zamore. 2009. 'Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies', *Cell*, 137: 509-21.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25: 2078-9.

Li, J., Z. Yang, B. Yu, J. Liu, and X. Chen. 2005. 'Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis', *Curr Biol*, 15: 1501-7.

Li, X. Z., C. K. Roy, X. Dong, E. Bolcun-Filas, J. Wang, B. W. Han, J. Xu, M. J. Moore, J. C. Schimenti, Z. Weng, and P. D. Zamore. 2013. 'An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes', *Mol Cell*, 50: 67-81.

Lilly, B., B. Zhao, G. Ranganayakulu, B. M. Paterson, R. A. Schulz, and E. N. Olson. 1995. 'Requirement of MADS domain transcription factor D-MEF2 for muscle formation in Drosophila', *Science*, 267: 688-93.

Lim, A. K., and T. Kai. 2007. 'Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster', *Proc Natl Acad Sci U S A*, 104: 6714-9.

Lin, H., and A. C. Spradling. 1997. 'A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary', *Development*, 124: 2463-76.

Liu, L., H. Qi, J. Wang, and H. Lin. 2011. 'PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition', *Development*, 138: 1863-73.

Liu, N., M. Abe, L. R. Sabin, G. J. Hendriks, A. S. Naqvi, Z. Yu, S. Cherry, and N. M. Bonini. 2011. 'The exoribonuclease Nibbler controls 3' end processing of microRNAs in Drosophila', *Curr Biol*, 21: 1888-93.

Liu, Y., and M. Lehmann. 2008. 'A genomic response to the yeast transcription factor GAL4 in Drosophila', *Fly (Austin)*, 2: 92-8.

Losos, J., and P.H. Raven. 2008. *Biology* (McGraw-Hill Higher Eduction, Boston).

Luo, M. J., and R. Reed. 1999. 'Splicing is required for rapid and efficient mRNA export in metazoans', *Proc Natl Acad Sci U S A*, 96: 14937-42.

Lyne, R., R. Smith, K. Rutherford, M. Wakeling, A. Varley, F. Guillier, H. Janssens, W. Ji, P. McLaren, P. North, D. Rana, T. Riley, J. Sullivan, X. Watkins, M. Woodbridge, K. Lilley, S. Russell, M. Ashburner, K. Mizuguchi, and G. Micklem. 2007. 'FlyMine: an integrated database for Drosophila and Anopheles genomics', *Genome Biol*, 8: R129.

Malone, C. D., J. Brennecke, M. Dus, A. Stark, W. R. McCombie, R. Sachidanandam, and G. J. Hannon. 2009. 'Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary', *Cell*, 137: 522-35.

Malone, C. D., and G. J. Hannon. 2009. 'Small RNAs as guardians of the genome', *Cell*, 136: 656-68.

Masuda, S., R. Das, H. Cheng, E. Hurt, N. Dorman, and R. Reed. 2005. 'Recruitment of the human TREX complex to mRNA during splicing', *Genes Dev*, 19: 1512-7.

McElhinny, A. S., S. T. Kazmierski, S. Labeit, and C. C. Gregorio. 2003. 'Nebulin: the nebulous, multifunctional giant of striated muscle', *Trends Cardiovasc Med*, 13: 195-201.

Meijering, E., O. Dzyubachyk, and I. Smal. 2012. 'Methods for cell and particle tracking', *Methods Enzymol*, 504: 183-200.

Mevel-Ninio, M., A. Pelisson, J. Kinder, A. R. Campos, and A. Bucheton. 2007. 'The flamenco locus controls the gypsy and ZAM retroviruses and is required for Drosophila oogenesis', *Genetics*, 175: 1615-24.

Michelson, A. M. 1994. 'Muscle pattern diversification in Drosophila is determined by the autonomous function of homeotic genes in the embryonic mesoderm', *Development*, 120: 755-68.
Mohn, F., D. Handler, and J. Brennecke. 2015. 'Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis', *Science*, 348: 812-7.
Mohn, F., G. Sienski, D. Handler, and J. Brennecke. 2014. 'The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in Drosophila', *Cell*, 157: 1364-79.
Montgomery, T. A., Y. S. Rim, C. Zhang, R. H. Dowen, C. M. Phillips, S. E. Fischer, and G. Ruvkun. 2012. 'PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1', *PLoS Genet*, 8: e1002616.
Moon, S., B. Cho, S. H. Min, D. Lee, and Y. D. Chung. 2011. 'The THO complex is required for nucleolar integrity in Drosophila spermatocytes', *Development*, 138: 3835-45.
Moreira, E. S., T. J. Wiltshire, G. Faulkner, A. Nilforoushan, M. Vainzof, O. T. Suzuki, G. Valle, R. Reeves, M. Zatz, M. R. Passos-Bueno, and D. E. Jenne. 2000. 'Limb-girdle muscular dystrophy type 2G is caused by mutations in the gene encoding the sarcomeric protein telethonin', *Nat Genet*, 24: 163-6.
Muerdter, F., P. M. Guzzardo, J. Gillis, Y. Luo, Y. Yu, C. Chen, R. Fekete, and G. J. Hannon. 2013. 'A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila', *Mol Cell*, 50: 736-48.
Murota, Y., H. Ishizu, S. Nakagawa, Y. W. Iwasaki, S. Shibata, M. K. Kamatani, K. Saito, H. Okano, H. Siomi, and M. C. Siomi. 2014. 'Yb integrates piRNA intermediates and processing factors into perinuclear bodies to enhance piRISC assembly', *Cell Rep*, 8: 103-13.
Nagao, A., T. Mituyama, H. Huang, D. Chen, M. C. Siomi, and H. Siomi. 2010. 'Biogenesis pathways of piRNAs loaded onto AGO3 in the Drosophila testis', *Rna*, 16: 2503-15.
Neville, M., F. Stutz, L. Lee, L. I. Davis, and M. Rosbash. 1997. 'The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export', *Curr Biol*, 7: 767-75.
Nguyen, H. T., F. Voza, N. Ezzeddine, and M. Frasch. 2007. 'Drosophila mind bomb2 is required for maintaining muscle integrity and survival', *J Cell Biol*, 179: 219-27.
Nishida, K. M., K. Saito, T. Mori, Y. Kawamura, T. Nagami-Okada, S. Inagaki, H. Siomi, and M. C. Siomi. 2007. 'Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad', *Rna*, 13: 1911-22.
Nishimasu, H., H. Ishizu, K. Saito, S. Fukuhara, M. K. Kamatani, L. Bonnefond, N. Matsumoto, T. Nishizawa, K. Nakanaga, J. Aoki, R. Ishitani, H. Siomi, M. C. Siomi, and O. Nureki. 2012. 'Structure and function of Zucchini endoribonuclease in piRNA biogenesis', *Nature*, 491: 284-7.
Niwa, R., T. Matsuda, T. Yoshiyama, T. Namiki, K. Mita, Y. Fujimoto, and H. Kataoka. 2004. 'CYP306A1, a cytochrome P450 enzyme, is essential for ecdysteroid biosynthesis in the prothoracic glands of Bombyx and Drosophila', *J Biol Chem*, 279: 35942-9.
Nott, A., S. H. Meislin, and M. J. Moore. 2003. 'A quantitative analysis of intron effects on mammalian gene expression', *Rna*, 9: 607-17.
O'Mahoney, J. V., K. L. Guven, J. Lin, J. E. Joya, C. S. Robinson, R. P. Wade, and E. C. Hardeman. 1998. 'Identification of a novel slow-muscle-fiber enhancer binding protein, MusTRD1', *Mol Cell Biol*, 18: 6641-52.

Ohno, M., A. Segref, A. Bachi, M. Wilm, and I. W. Mattaj. 2000. 'PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation', *Cell*, 101: 187-98.

Ohtani, H., Y. W. Iwasaki, A. Shibuya, H. Siomi, M. C. Siomi, and K. Saito. 2013. 'DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the Drosophila ovary', *Genes Dev*, 27: 1656-61.

Olivieri, D., K. A. Senti, S. Subramanian, R. Sachidanandam, and J. Brennecke. 2012. 'The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in Drosophila', *Mol Cell*, 47: 954-69.

Olivieri, D., M. M. Sykora, R. Sachidanandam, K. Mechtler, and J. Brennecke. 2010. 'An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila', *Embo j*, 29: 3301-17.

Ono, H., K. F. Rewitz, T. Shinoda, K. Itoyama, A. Petryk, R. Rybczynski, M. Jarcho, J. T. Warren, G. Marques, M. J. Shimell, L. I. Gilbert, and M. B. O'Connor. 2006. 'Spook and Spookier code for stage-specific components of the ecdysone biosynthetic pathway in Diptera', *Dev Biol*, 298: 555-70.

Palazzo, A. F., M. Springer, Y. Shibata, C. S. Lee, A. P. Dias, and T. A. Rapoport. 2007. 'The signal sequence coding region promotes nuclear export of mRNA', *PLoS Biol*, 5: e322.

Palumbo, G., S. Bonaccorsi, L. G. Robbins, and S. Pimpinelli. 1994. 'Genetic analysis of Stellate elements of Drosophila melanogaster', *Genetics*, 138: 1181-97.

Pane, A., P. Jiang, D. Y. Zhao, M. Singh, and T. Schupbach. 2011. 'The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline', *Embo j*, 30: 4601-15.

Papa, I., C. Astier, O. Kwiatek, F. Raynaud, C. Bonnal, M. C. Lebart, C. Roustan, and Y. Benyamin. 1999. 'Alpha actinin-CapZ, an anchoring complex for thin filaments in Z-line', *J Muscle Res Cell Motil*, 20: 187-97.

Pappas, C. T., N. Bhattacharya, J. A. Cooper, and C. C. Gregorio. 2008. 'Nebulin interacts with CapZ and regulates thin filament architecture within the Z-disc', *Mol Biol Cell*, 19: 1837-47.

Parker, L., S. Gross, and L. Alphey. 2001. 'Vectors for the expression of tagged proteins in Drosophila', *Biotechniques*, 31: 1280-2, 84, 86.

Patil, V. S., A. Anand, A. Chakrabarti, and T. Kai. 2014. 'The Tudor domain protein Tapas, a homolog of the vertebrate Tdrd7, functions in the piRNA pathway to regulate retrotransposons in germline of Drosophila melanogaster', *BMC Biol*, 12: 61.

Patil, V. S., and T. Kai. 2010. 'Repression of retroelements in Drosophila germline via piRNA pathway by the Tudor domain protein Tejas', *Curr Biol*, 20: 724-30.

Petryk, A., J. T. Warren, G. Marques, M. P. Jarcho, L. I. Gilbert, J. Kahler, J. P. Parvy, Y. Li, C. Dauphin-Villemant, and M. B. O'Connor. 2003. 'Shade is the Drosophila P450 enzyme that mediates the hydroxylation of ecdysone to the steroid insect molting hormone 20-hydroxyecdysone', *Proc Natl Acad Sci U S A*, 100: 13773-8.

Preall, J. B., B. Czech, P. M. Guzzardo, F. Muerdter, and G. J. Hannon. 2012. 'shutdown is a component of the Drosophila piRNA biogenesis machinery', *Rna*, 18: 1446-57.

Prud'homme, N., M. Gans, M. Masson, C. Terzian, and A. Bucheton. 1995. 'Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster', *Genetics*, 139: 697-711.

Pyle, W. G., and R. J. Solaro. 2004. 'At the crossroads of myocardial signaling: the role of Z-discs in intracellular signaling and cardiac function', *Circ Res*, 94: 296-305.

Raffaello, A., G. Milan, E. Masiero, S. Carnio, D. Lee, G. Lanfranchi, A. L. Goldberg, and M. Sandri. 2010. 'JunB transcription factor maintains skeletal muscle mass and promotes hypertrophy', *J Cell Biol*, 191: 101-13.

Ranganayakulu, G., B. Zhao, A. Dokidis, J. D. Molkentin, E. N. Olson, and R. A. Schulz. 1995. 'A series of mutations in the D-MEF2 transcription factor reveal multiple functions in larval and adult myogenesis in Drosophila', *Dev Biol*, 171: 169-81.

Rehwinkel, J., A. Herold, K. Gari, T. Kocher, M. Rode, F. L. Ciccarelli, M. Wilm, and E. Izaurralde. 2004. 'Genome-wide analysis of mRNAs regulated by the THO complex in Drosophila melanogaster', *Nat Struct Mol Biol*, 11: 558-66.

Reich, M., T. Liefeld, J. Gould, J. Lerner, P. Tamayo, and J. P. Mesirov. 2006. 'GenePattern 2.0', *Nat Genet*, 38: 500-1.

Richards, G. 1981. 'The radioimmune assay of ecdysteroid titres in Drosophila melanogaster', *Mol Cell Endocrinol*, 21: 181-97.

Rodrigues, J. P., M. Rode, D. Gatfield, B. J. Blencowe, M. Carmo-Fonseca, and E. Izaurralde. 2001. 'REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus', *Proc Natl Acad Sci U S A*, 98: 1030-5.

Roy, S., and K. VijayRaghavan. 1998. 'Patterning muscles using organizers: larval muscle templates and adult myoblasts actively interact to pattern the dorsal longitudinal flight muscles of Drosophila', *J Cell Biol*, 141: 1135-45.

Rozhkov, N. V., M. Hammell, and G. J. Hannon. 2013. 'Multiple roles for Piwi in silencing Drosophila transposons', *Genes Dev*, 27: 400-12.

Ruby, J. G., C. Jan, C. Player, M. J. Axtell, W. Lee, C. Nusbaum, H. Ge, and D. P. Bartel. 2006. 'Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans', *Cell*, 127: 1193-207.

Rushton, E., R. Drysdale, S. M. Abmayr, A. M. Michelson, and M. Bate. 1995. 'Mutations in a novel gene, myoblast city, provide evidence in support of the founder cell hypothesis for Drosophila muscle development', *Development*, 121: 1979-88.

Russo, J., A. W. Harrington, and M. Steiniger. 2016. 'Antisense Transcription of Retrotransposons in Drosophila: An Origin of Endogenous Small Interfering RNA Precursors', *Genetics*, 202: 107-21.

Ryu, W. S., and J. E. Mertz. 1989. 'Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm', *J Virol*, 63: 4386-94.

Saito, K., H. Ishizu, M. Komai, H. Kotani, Y. Kawamura, K. M. Nishida, H. Siomi, and M. C. Siomi. 2010. 'Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila', *Genes Dev*, 24: 2493-8.

Saito, K., Y. Sakaguchi, T. Suzuki, T. Suzuki, H. Siomi, and M. C. Siomi. 2007. 'Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends', *Genes Dev*, 21: 1603-8.

Sandri, M., C. Sandri, A. Gilbert, C. Skurk, E. Calabria, A. Picard, K. Walsh, S. Schiaffino, S. H. Lecker, and A. L. Goldberg. 2004. 'Foxo transcription factors induce the atrophy-related ubiquitin ligase atrogin-1 and cause skeletal muscle atrophy', *Cell*, 117: 399-412.

Sartori, R., G. Milan, M. Patron, C. Mammucari, B. Blaauw, R. Abraham, and M. Sandri. 2009. 'Smad2 and 3 transcription factors control muscle mass in adulthood', *Am J Physiol Cell Physiol*, 296: C1248-57.

Scherfer, C., M. R. Qazi, K. Takahashi, R. Ueda, M. S. Dushay, U. Theopold, and B. Lemaitre. 2006. 'The Toll immune-regulated Drosophila protein Fondue is involved in hemolymph clotting and puparium formation', *Dev Biol*, 295: 156-63.

Schnorrer, F., C. Schonbauer, C. C. Langer, G. Dietzl, M. Novatchkova, K. Schernhuber, M. Fellner, A. Azaryan, M. Radolf, A. Stark, K. Keleman, and B. J. Dickson. 2010. 'Systematic genetic analysis of muscle morphogenesis and function in Drosophila', *Nature*, 464: 287-91.

Seitz, H., M. Ghildiyal, and P. D. Zamore. 2008. 'Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies', *Curr Biol*, 18: 147-51.

Senti, K. A., and J. Brennecke. 2010. 'The piRNA pathway: a fly's perspective on the guardian of the genome', *Trends Genet*, 26: 499-509.

Shatkin, A. J. 1976. 'Capping of eucaryotic mRNAs', *Cell*, 9: 645-53.

Shatkin, A. J., and J. L. Manley. 2000. 'The ends of the affair: capping and polyadenylation', *Nat Struct Biol*, 7: 838-42.

Shatoury, H. H. 1959. 'The Development of a Pupal Lethal Mutant in
Drosophila', *Caryologia*, 12:1: 104-09.

Sheikh, F., M. L. Bang, S. Lange, and J. Chen. 2007. '"Z"eroing in on the role of Cypher in striated muscle function, signaling, and human disease', *Trends Cardiovasc Med*, 17: 258-62.

Shen, H. 2009. 'UAP56- a key player with surprisingly diverse roles in pre-mRNA splicing and nuclear export', *BMB Rep*, 42: 185-8.

Shirayama, M., M. Seth, H. C. Lee, W. Gu, T. Ishidate, D. Conte, Jr., and C. C. Mello. 2012. 'piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline', *Cell*, 150: 65-77.

Sienski, G., D. Donertas, and J. Brennecke. 2012. 'Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression', *Cell*, 151: 964-80.

Siomi, M. C., K. Sato, D. Pezic, and A. A. Aravin. 2011. 'PIWI-interacting small RNAs: the vanguard of genome defence', *Nat Rev Mol Cell Biol*, 12: 246-58.

Slotkin, R. K., and R. Martienssen. 2007. 'Transposable elements and the epigenetic regulation of the genome', *Nat Rev Genet*, 8: 272-85.

Specchia, V., L. Piacentini, P. Tritto, L. Fanti, R. D'Alessandro, G. Palumbo, S. Pimpinelli, and M. P. Bozzetti. 2010. 'Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons', *Nature*, 463: 662-5.

Stewart, B. A., H. L. Atwood, J. J. Renger, J. Wang, and C. F. Wu. 1994. 'Improved stability of Drosophila larval neuromuscular preparations in haemolymph-like physiological solutions', *J Comp Physiol A*, 175: 179-91.

Strasser, K., and E. Hurt. 2000. 'Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export', *Embo j*, 19: 410-20.

Strasser, K., S. Masuda, P. Mason, J. Pfannstiel, M. Oppizzi, S. Rodriguez-Navarro, A. G. Rondon, A. Aguilera, K. Struhl, R. Reed, and E. Hurt. 2002. 'TREX is a conserved complex coupling transcription with messenger RNA export', *Nature*, 417: 304-8.

Tassabehji, M., M. Carette, C. Wilmot, D. Donnai, A. P. Read, and K. Metcalfe. 1999. 'A transcription factor involved in skeletal muscle gene expression is deleted in patients with Williams syndrome', *Eur J Hum Genet*, 7: 737-47.

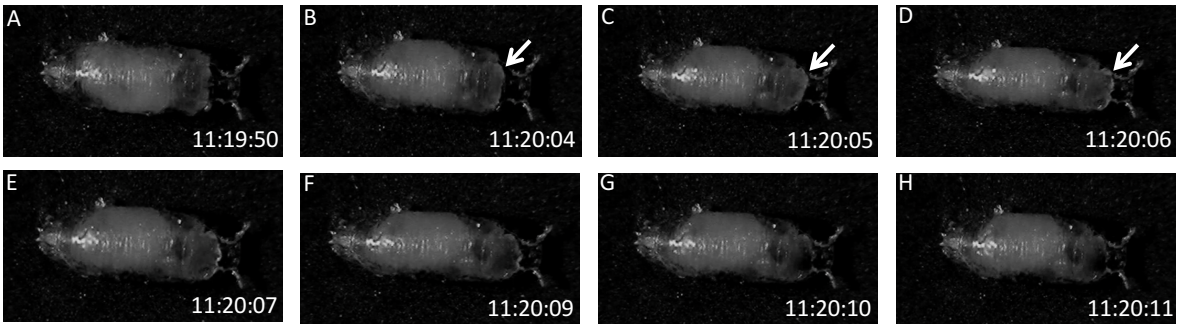
Tata, J. R. 2002. 'Signalling through nuclear receptors', *Nat Rev Mol Cell Biol*, 3: 702-10.

Taylor, M. V. 1995. 'Muscle development. Making Drosophila muscle', *Curr Biol*, 5: 740-2.

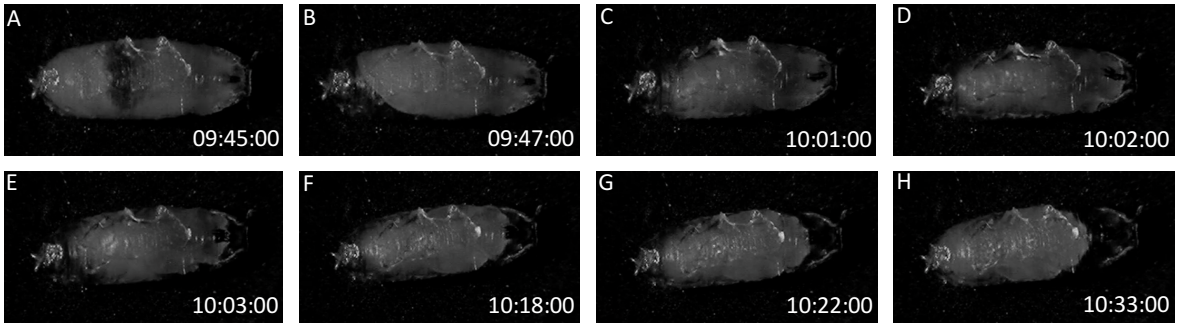
Taylor, M.V. 2006. *Comparison of muscle development in Drosophila and vertebrates* (Landes Bioscience/Springer).
Teleman, A. A., V. Hietakangas, A. C. Sayadian, and S. M. Cohen. 2008. 'Nutritional control of protein biosynthetic capacity by insulin via Myc in Drosophila', *Cell Metab*, 7: 21-32.
Thummel, C. S. 1990. 'Puffs and gene regulation--molecular insights into the Drosophila ecdysone regulatory hierarchy', *Bioessays*, 12: 561-8.
Topisirovic, I., N. Siddiqui, V. L. Lapointe, M. Trost, P. Thibault, C. Bangeranye, S. Pinol-Roma, and K. L. Borden. 2009. 'Molecular dissection of the eukaryotic initiation factor 4E (eIF4E) export-competent RNP', *Embo j*, 28: 1087-98.
Trapnell, C., L. Pachter, and S. L. Salzberg. 2009. 'TopHat: discovering splice junctions with RNA-Seq', *Bioinformatics*, 25: 1105-11.
Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, and L. Pachter. 2012. 'Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks', *Nat Protoc*, 7: 562-78.
Vagin, V. V., A. Sigova, C. Li, H. Seitz, V. Gvozdev, and P. D. Zamore. 2006. 'A distinct small RNA pathway silences selfish genetic elements in the germline', *Science*, 313: 320-4.
Valencia, P., A. P. Dias, and R. Reed. 2008. 'Splicing promotes rapid and efficient mRNA export in mammalian cells', *Proc Natl Acad Sci U S A*, 105: 3386-91.
Valkov, E., J. C. Dean, D. Jani, S. I. Kuhlmann, and M. Stewart. 2012. 'Structural basis for the assembly and disassembly of mRNA nuclear export complexes', *Biochim Biophys Acta*, 1819: 578-92.
Vaskova, M., A. M. Bentley, S. Marshall, P. Reid, C. S. Thummel, and A. J. Andres. 2000. 'Genetic analysis of the Drosophila 63F early puff. Characterization of mutations in E63-1 and maggie, a putative Tom22', *Genetics*, 156: 229-44.
Vereshchagina, N., D. Bennett, B. Szoor, J. Kirchner, S. Gross, E. Vissi, H. White-Cooper, and L. Alphey. 2004. 'The essential role of PP1beta in Drosophila is to regulate nonmuscle myosin', *Mol Biol Cell*, 15: 4395-405.
Wang, G., and V. Reinke. 2008. 'A C. elegans Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis', *Curr Biol*, 18: 861-7.
Wang, H., Z. Ma, K. Niu, Y. Xiao, X. Wu, C. Pan, Y. Zhao, K. Wang, Y. Zhang, and N. Liu. 2016. 'Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in Drosophila', *Development*, 143: 530-9.
Wang, X., C. Blagden, J. Fan, S. J. Nowak, I. Taniuchi, D. R. Littman, and S. J. Burden. 2005. 'Runx1 prevents wasting, myofibrillar disorganization, and autophagy of skeletal muscle', *Genes Dev*, 19: 1715-22.
Warren, J. T., A. Petryk, G. Marques, M. Jarcho, J. P. Parvy, C. Dauphin-Villemant, M. B. O'Connor, and L. I. Gilbert. 2002. 'Molecular and biochemical characterization of two P450 enzymes in the ecdysteroidogenic pathway of Drosophila melanogaster', *Proc Natl Acad Sci U S A*, 99: 11043-8.
Warren, J. T., A. Petryk, G. Marques, J. P. Parvy, T. Shinoda, K. Itoyama, J. Kobayashi, M. Jarcho, Y. Li, M. B. O'Connor, C. Dauphin-Villemant, and L. I. Gilbert. 2004. 'Phantom encodes the 25-hydroxylase of Drosophila melanogaster and Bombyx mori: a P450 enzyme critical in ecdysone biosynthesis', *Insect Biochem Mol Biol*, 34: 991-1010.
Wasser, M., Z. Bte Osman, and W. Chia. 2007. 'EAST and Chromator control the destruction and remodeling of muscles during Drosophila metamorphosis', *Dev Biol*, 307: 380-93.
Watanabe, T., A. Takeda, T. Tsukiyama, K. Mise, T. Okuno, H. Sasaki, N. Minami, and H. Imai. 2006. 'Identification and characterization of two novel classes of small RNAs in the

mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes', *Genes Dev*, 20: 1732-43.
Weick, E. M., P. Sarkies, N. Silva, R. A. Chen, S. M. Moss, A. C. Cording, J. Ahringer, E. Martinez-Perez, and E. A. Miska. 2014. 'PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans', *Genes Dev*, 28: 783-96.
White-Cooper, H. 2010. 'Molecular mechanisms of gene regulation during Drosophila spermatogenesis', *Reproduction*, 139: 11-21.
Wiegand, H. L., G. A. Coburn, Y. Zeng, Y. Kang, H. P. Bogerd, and B. R. Cullen. 2002. 'Formation of Tap/NXT1 heterodimers activates Tap-dependent nuclear mRNA export by enhancing recruitment to nuclear pore complexes', *Mol Cell Biol*, 22: 245-56.
Wolff, B., J. J. Sanglier, and Y. Wang. 1997. 'Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA', *Chem Biol*, 4: 139-47.
Wright, L. G., T. Chen, C. S. Thummel, and G. M. Guild. 1996. 'Molecular characterization of the 71E late puff in Drosophila melanogaster reveals a family of novel genes', *J Mol Biol*, 255: 387-400.
Yang, J., H. P. Bogerd, P. J. Wang, D. C. Page, and B. R. Cullen. 2001. 'Two closely related human nuclear export factors utilize entirely distinct export pathways', *Mol Cell*, 8: 397-406.
Yu, B., Z. Yang, J. Li, S. Minakhina, M. Yang, R. W. Padgett, R. Steward, and X. Chen. 2005. 'Methylation as a crucial step in plant microRNA biogenesis', *Science*, 307: 932-5.
Zamparini, A. L., M. Y. Davis, C. D. Malone, E. Vieira, J. Zavadil, R. Sachidanandam, G. J. Hannon, and R. Lehmann. 2011. 'Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in Drosophila', *Development*, 138: 4039-50.
Zanni, V., A. Eymery, M. Coiffet, M. Zytnicki, I. Luyten, H. Quesneville, C. Vaury, and S. Jensen. 2013. 'Distribution, evolution, and diversity of retrotransposons at the flamenco locus reflect the regulatory properties of piRNA clusters', *Proc Natl Acad Sci U S A*, 110: 19842-7.
Zhang, F., J. Wang, J. Xu, Z. Zhang, B. S. Koppetsch, N. Schultz, T. Vreven, C. Meignin, I. Davis, P. D. Zamore, Z. Weng, and W. E. Theurkauf. 2012. 'UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery', *Cell*, 151: 871-84.
Zhang, Z., J. Wang, N. Schultz, F. Zhang, S. S. Parhad, S. Tu, T. Vreven, P. D. Zamore, Z. Weng, and W. E. Theurkauf. 2014. 'The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing', *Cell*, 157: 1353-63.
Zhang, Z., J. Xu, B. S. Koppetsch, J. Wang, C. Tipping, S. Ma, Z. Weng, W. E. Theurkauf, and P. D. Zamore. 2011. 'Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains', *Mol Cell*, 44: 572-84.
Zhao, J., J. J. Brault, A. Schild, P. Cao, M. Sandri, S. Schiaffino, S. H. Lecker, and A. L. Goldberg. 2007. 'FoxO3 coordinately activates protein degradation by the autophagic/lysosomal and proteasomal pathways in atrophying muscle cells', *Cell Metab*, 6: 472-83.

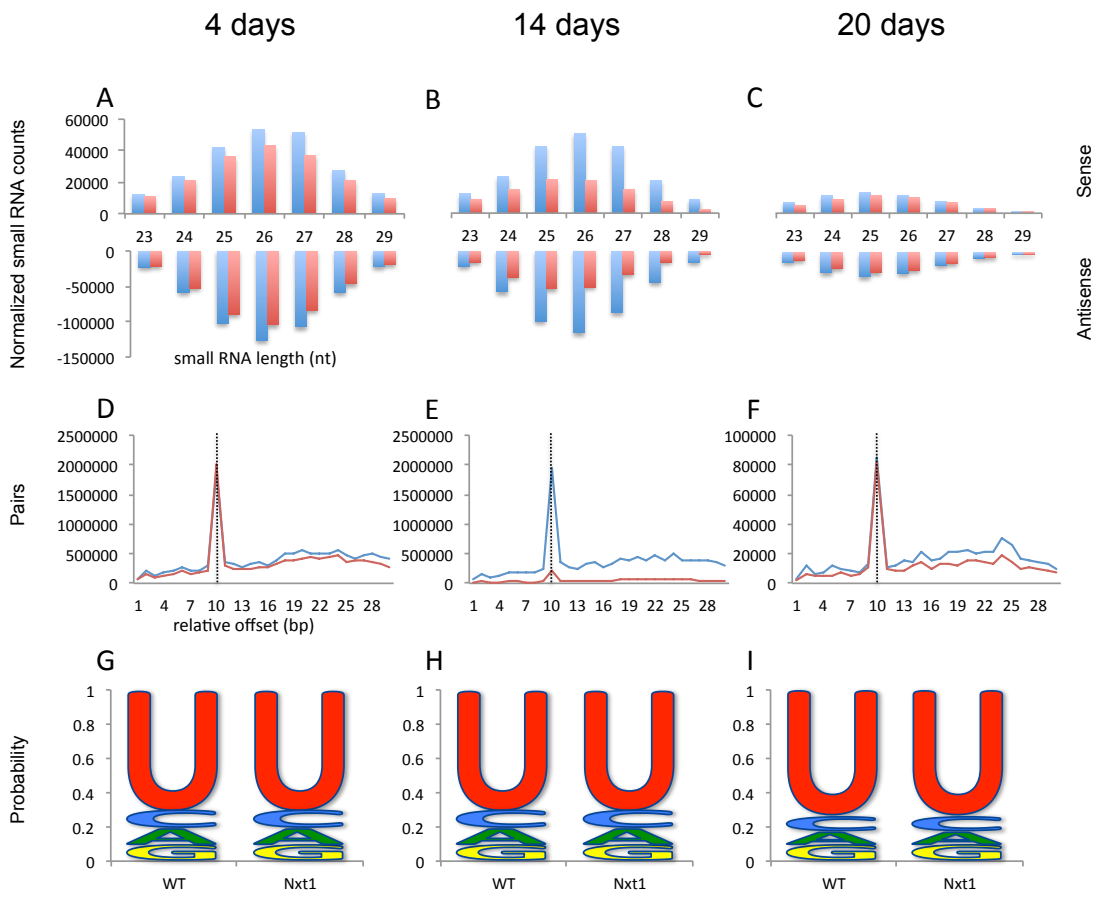
Supplementary.



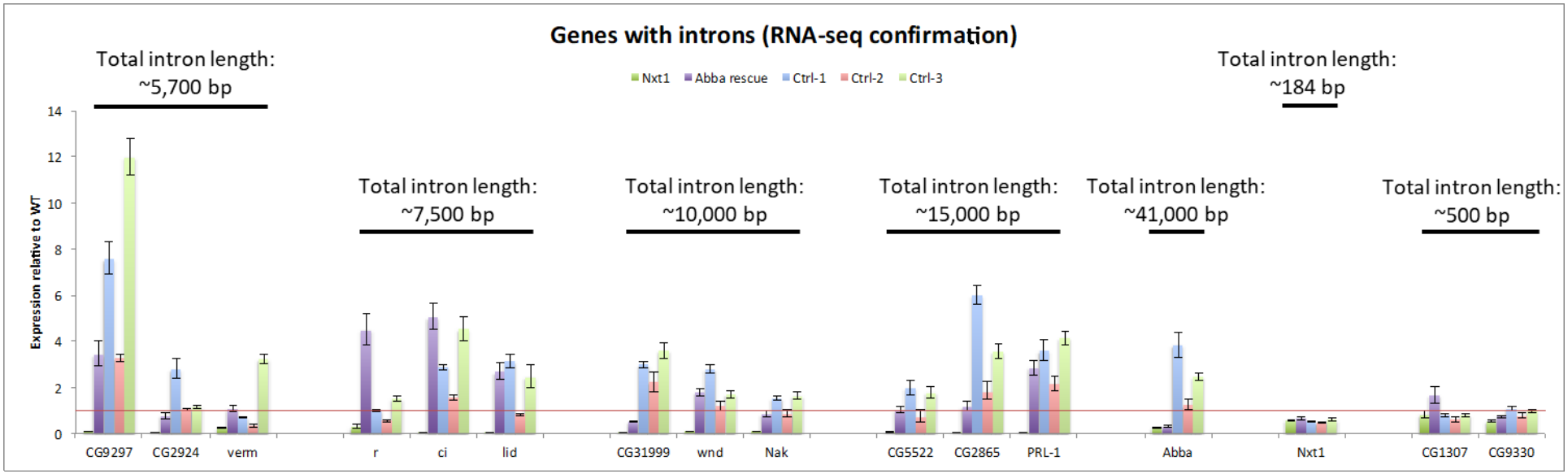
Supplementary figure 1: **Observation of the head eversion in *w*^{1118} pupa.** A) After 11 hours, pupa starts the process of head eversion. B-E) Head eversion occurs in a two seconds window (white arrows). F-H) Head eversion has occurred. Figures were taken with a Samsung SDN-500 camera.



Supplementary figure 2: **Body movement of pupa before head eversion.** A) Close to the 10-hour mark the pupa body starts to move within the pupa. B) The posterior end of the body starts to "wiggle". C-E) The "wiggling" continues, while the anterior end starts to loosen. F-H) The posterior end of the body moves back to the posterior end of the pupa case, whereas the anterior end retracts towards the posterior end. Figures were taken with a Samsung SDN-500 camera.



Supplementary figure 3: **Unique piRNA reads mapping to all transposons**. Length distribution plots from (A) 4 days, (B) 14 days and (C) 20 days of WT (blue) and Nxt1 (red) with sense (top) and antisense (bottom) strands. Ping-pong analysis of (D) 4 days, (E) 14 days and (F) 20 days of WT (blue) and Nxt1 trans-heterozygotes (red). Ten nucleotide overlaps indicated with dashed lines. First nucleotide 5' U-bias of (G) 4 days, (H) 14 days and (I) 20 days. Small RNA data was normalized per million miRNA hairpins.



Supplementary Figure 4: **Gal4 driven without an UAS construct increases the expression of many genes.** Genes with an average total intron length of 5700, 7500, 10000 and 15000 were examined. Expression of *abba* and *Nxt1* are also shown. *CG1307* and *CG9330* were used as control genes with low total intron length. All genes examined have expression in the carcass. *Rp49* was used for normalization. Majority of the genes with long total intron length have an increased expression when Gal4 (mef2gal4 or armgal4) is driven with either an UAS construct (UAS-$abba^{fl}$) or without in the Nxt1 trans-heterozygotes background (Abba rescue, Ctrl-1 and Ctrl-2). Majority of the genes with long total intron length examined also have increased expression when Gal4 (armgal4) is driven without an UAS construct (Ctrl-3). Nxt1 = $Nxt1^{DG05102}/Nxt1^{Z2\text{-}0488}$. Abba rescue = $Nxt1^{DG05102}/Nxt1^{Z2\text{-}0488}$; UAS-$abba^{fl}$>mef2gal4. Ctrl-1 = $Nxt1^{DG05102}/Nxt1^{Z2\text{-}0488}$; mef2gal4/+. Ctrl-2 = $Nxt1^{DG05102}/Nxt1^{Z2\text{-}0488}$; armgal4/+. Ctrl-3 = armgal4; Sb/Tm3.

Supplementary table 1: **Read mappings information from small RNA seq.** Total amount of reads sequenced from each sample. Total amount of reads mapping to rRNA and miRNA hairpins and reads mapping to FlyBase dm3 genome. Samples highlighted in orange were used for piPipes analysis.

				Flybase dm3 genome mapping*		
	Total reads	**rRNA mapping**	**miRNA hairpin**	**Total**	**Unique**	**Multiple**
4 days WT rep 1	9.108.651	1.269.783	495.807	4.488.959	995.813	3.493.146
4 days WT rep 2	12.362.835	2.312.084	729.465	5.769.709	1.342.018	4,427.691
4 days WT rep 3	17.377.413	3.351.109	1.216.393	8.290.148	1.914.472	6.375.676
14 days WT rep 1	14.071.178	2.302.363	1.278.427	7.138.305	1.541.725	5.596.580
14 days WT rep 2	12.827.732	1.492.361	1.011.997	7.704.512	1.700.626	6.003.886
14 days WT rep 3	11.191.934	1.309.629	824.080	5.509.281	1.209.892	4.299.389
20 days WT rep 1	7.213.418	3.593.595	673.516	1.973.190	411.400	1.561.790
20 days WT rep 2	6.820.800	2.611.113	818.161	2.360.198	497.978	1.862.220
20 days WT rep 3	10.356.573	4.520.654	872.366	3.449.759	723.257	2.726.502
4 days Nxt1 rep 1	11.938.123	1.796.977	893.274	6.026.333	1.334.797	4.691.536
4 days Nxt1 rep 2	12.589.807	1.976.676	1.178.248	5.949.935	1.265.965	4.683.970
4 days Nxt1 rep 3	11.297.093	1.193.724	651.532	5.990.260	1.293.400	4.696.860
14 days Nxt1 rep 1	10.649.255	1.943.443	1.653.428	5.099.192	1.133.038	3.966.154
14 days Nxt1 rep 2	9.597.141	1.294.773	1.254.899	4.571.802	1.001.707	3.570.095
14 days Nxt1 rep 3	11.990.063	3.269.155	1.223.201	5.643.207	1.334.676	4.308.531
20 days Nxt1 rep 1	9.828.896	5.708.394	610.605	2.475.353	514.816	1.960.537
20 days Nxt1 rep 2	8.377.702	3.486.840	1.256.671	2.217.598	497.911	1.719.687
20 days Nxt1 rep 3	7.116.800	3.810.437	732.912	1.813.143	394.659	1.418.484

Appendices.

Appendix 1 - Primers

Primer	Sequence	Reference
RP49_F	5’-TGTCCTTCCAGCTTCAAGATGACCATC-3’	N. A
RP49_R	5’-CTTGGGCTTGCGCCATTTGTG-3’	N. A
FOXO_F	5’-GGCATCATATTTGGGATAAAGGGT-3’	N. A
FOXO_R	5’-GTGCGGTGCAAACTGCTAAT-3’	N. A
Akt1_F	5’-CTCCATAGCCGAAGAGCCG-3’	N. A
Akt1_R	5’-CTGTGGCGGTGAAACCAAAA-3’	N. A
P13k92E_F	5’-CGAGCGCGTACGTCTATCAA-3’	N. A
P13k92E_R	5’-AATAAGGCAGCCGTGTTTGC-3’	N. A
Ter94_F	5’-TCAGGTGAGAATTGAAAGGCA-3’	N. A
Ter94_R	5’-TTTGACCGGGGAATCTGCAA-3’	N. A
CG11658_F	5’-CGTACCTGTGGAAGGAGCAC-3’	N. A
CG11658_R	5’-CGATCCTTACCTCGCGGATT-3’	N. A
Abba_F	5’-ACTGGTCCCTTGGGCTTTTC-3’	N. A
Abba_R	5’-CGGAAGGTGCAGGAAGTTCA-3’	N. A
eIF4AIII_F	5’-CTTCTGCAACACGAAACGCA-3’	N. A
eIF4AIII_R	5’-GCTCCTTTTGCGGCATATCG-3’	N. A
Mago_F	5’-TGGACGTGAACCGGTCAAAA-3’	N. A
Mago_R	5’-GAAATGCAGGCCGATGAGTG-3’	N. A
RnpS1_F	5’-TACGGTATCCCCGGTTGTCT-3’	N. A
RnpS1_R	5’-CCTCATGGGTGACTGACGAC-3’	N. A
Tsu_F	5’-GAGAGGCGATCCACAGCTAC-3’	N. A
Tsu_R	5’-CTGCGCCTCCTCATGGATAG-3’	N. A
Abba-intron_2_F	5’-ACAATTAGCGGTGGGCATCA-3’	N. A
Abba-intron_2_R	5’-TTGCTGGCAATATGCATCGC-3’	N. A
Abba-intron_4_F	5’-GGACAACGAGCACTCGTACA-3’	N. A
Abba-intron_4_R	5’-GAGATGGTGGGAAAGGGGTG-3’	N. A
Abba-intron_7_F	5’-TGCTGTATTCGCCCCCATTT-3’	N. A
Abba-intron_7_R	5’-TTGCCTGTGTGTCAGTGTGT-3’	N. A
Abba-exon_2/3_F	5’-AAGGAACAGGCCCCTTTCAA-3’	N. A
Abba-exon_2/3_R	5’-TCCATGCAGAAGGAGTGCTG-3’	N. A
Abba-exon_1/2_F	5’-CAGCAACAGCAAGGATTGCC-3’	N. A
Abba-exon_1/2_R	5’-ACGTAGTCCACCAGTCCCTC-3’	N. A
Abba-exon_7/8_F	5’-AGCGAAGATCGCTACGGAAG-3’	N. A
Abba-exon_7/8_R	5’-ATGGCCAAGTATCTGGAGCG-3’	N. A
CG9297_F	5’-AACAGATGGTGTCAGGGATGG-3’	N. A
CG9297_R	5’-ATCAACCGCAGAATGCCCAC-3’	N. A
CG2924_F	5’-GCCACTCTCGTCAAGGGAAA-3’	N. A
CG2924_R	5’-CGGGGTGAACCATCCATTCT-3’	N. A
verm_F	5’-GTGAAGTCCTGCGATGTGCT-3’	N. A
verm_R	5’-GACTTGCCGTTGCAGAAGAG-3’	N. A
r_F	5’-CGGTCTCCCATCGTTTGTCA-3’	N. A
r_R	5’-TCCTTGGCGGAACGAATCAA-3’	N. A

ci_F	5’-GTACACGCGGTGAAAAACCA-3’	N. A
ci_R	5’-AGCAGCCTTCAAACGTGCAT-3’	N. A
lid_F	5’-ACTTGGTGCCCACTGAGATG-3’	N. A
lid_R	5’-CCAGAGCCGTGATCCATTGT-3’	N. A
CG31999_F	5’-CCCAAATATGTCCCAGCGGT-3’	N. A
CG31999_R	5’-AGCAGTGGTATCCGCCATTC-3’	N. A
wnd_F	5’-ACCGACAATCTGCGAAGGAG-3’	N. A
wnd_R	5’-GAGAGTCCAGGTCGGAGGAT-3’	N. A
Nak_F	5’-CTGCCGCTGTGTCTCCTTAC-3’	N. A
Nak_R	5’-CACCAGATGAAACGCCGATG-3’	N. A
CG5522_F	5’-CAACAAGCGTTGCAGATGCC-3’	N. A
CG5522_R	5’-TCTAGCAGGTGGTGACGAGG-3’	N. A
CG2865_F	5’-TAACCGAAAACGCCACCTGT-3’	N. A
CG2865_R	5’-TAGCAGCATGGGTGCATTGA-3’	N. A
PRL-1_F	5’-AGGCATTACCGTCAAGGACC-3’	N. A
PRL-1_R	5’-CCAGACCAGCCACACAATGA-3’	N. A
Nxt1_F	5’-ACCGACGCCAACAAATTGGA-3’	N. A
Nxt1_R	5’-GTTTGAGGATGGCAGCTCCT-3’	N. A
CG1307_F	5’-CGCAGGAATCTTCGGTGTCT-3’	N. A
CG1307_R	5’-TTGCCCTTGCCCATCTTGAG-3’	N. A
CG9330_F	5’-GAGACTGTGGTCACGCTCAA-3’	N. A
CG9330_R	5’-GAGTCCGATGTGGCTGAGAG-3’	N. A
R2_F	5’-TGTCGTGCAGGATGTGATGC-3’	N. A
R2_R	5’-TACTACGCAGTTGTGTCTAGC-3’	N. A
Tabor_F	5’-TATAACTACCAGAACAGAGGC-3’	N. A
Tabor_R	5’-GACGTGCCTTGTGCATGT-3’	N. A
Mdg1_F	5’-CGCCAGCAACAGCATTTG-3’	Basquin et al. 2014
Mdg1_R	5’-GCATACTCATTTCCGTTTTCTGATC-3’	Basquin et al. 2014
HeT-A_F	5’-GGCAGCATATTAGCGCGTACA-3’	Basquin et al. 2014
HeT-A_R	5’-TTTGCCGCCAGCTTTTGT-3’	Basquin et al. 2014
Blood_F	5’-AGCGCTCACCATCATCAAACT-3’	Basquin et al. 2014
Blood_R	5’-AGTTCGGTGGCTTCTTTCTATTGT-3’	Basquin et al. 2014
Gypsy_F	5’-CAAGGATTGGAAATGGTTAGGC-3’	Basquin et al. 2014
Gypsy_R	5’-TCGTAGTATTGCATTAGTGGTAGGTCTC-3’	Basquin et al. 2014
ZAM_F	5’-AAAGCACCCCCTTACCGC-3’	Basquin et al. 2014
ZAM_R	5’-CGTTGCCCGAATCCATATTT-3’	Basquin et al. 2014
Springer_F	5’-CGGCTTACCCGAAATCGAA-3’	Basquin et al. 2014
Springer_R	5’-TTGACCCTTTGCCTTTTTGG-3’	Basquin et al. 2014
Stalker4_F	5’-AATCAAGCGGCGCTGAGT-3’	Basquin et al. 2014
Stalker4_R	5’-CGCTTCCTTCTAGCGTTTCAAT-3’	Basquin et al. 2014
Copia_F	5’-TCATGTGAAATTTGTGAACCCTG-3’	Basquin et al. 2014
Copia_R	5’-AGGAAGTCTTGCCTGTTTACCATT-3’	Basquin et al. 2014
Flam507_F	5’-TTGGCTATGAGGATCAGACA-3’	Goriaux et al. 2014
Flam507_R	5’-CTTCAAAGCGATTCATTCCT-3’	Goriaux et al. 2014
Flam527_F	5’-GTGGCTTCACAAAACACGAC-3’	Goriaux et al. 2014

Flam527_R	5’-GCCGGTCCTAAATATCTTCTC-3’	Goriaux et al. 2014
Cluster2A_F	5’-GCCTACGCAGAGGCCTAAGT-3’	Klattenhoff et al. 2009
Cluster2A_R	5’-CAGATGTGGTCCAGTTGTGC-3’	Klattenhoff et al. 2009
Cluster2B_F	5’-CTGCTTTGTGCTTGGAGATG-3’	Klattenhoff et al. 2009
Cluster2B_R	5’-TCTGCACAGATTCTGAAATTGAA-3’	Klattenhoff et al. 2009
Cluster1A_F	5’-CGTCCCAGCCTACCTAGTCA-3’	Klattenhoff et al. 2009
Cluster1A_R	5’-ACTTCCCGGTGAAGACTCCT-3’	Klattenhoff et al. 2009
Cluster1C_F	5’-CGCTGTTGAAAGCAAATTGA-3’	Klattenhoff et al. 2009
Cluster1C_R	5’-GAGACCTTCGCTCCAGTGTC-3’	Klattenhoff et al. 2009

Appendix 2 – Stocks and Crosses

Genotype	From	Reference
yw;*Nxt1*^{DG05102}	Bloomington stock	Caporilli et al. (2013)
yw;*Nxt1*^{DG05102}/CAG	Generated in HWC-lab	Caporilli et al. (2013)
yw;*Nxt1*$^{Z2-0488}$/CAG	Generated in HWC-lab	Caporilli et al. (2013)
w;*abba*^{df}/Cyo	Bloomington stock 7550	
w;*mib2*^{df}/Cyo	Bloomington stock 7846	
w;*abba*^{dp}/Tm3,sb	GenetiVision Corporation P6-B8	
w; if/cyo;(16.4)UAS-*mib2*	Gift from Hanh T. Nguyen	
w;;UAS-*abba*$^{fl\ \#\ 5}$	Gift from Hanh T. Nguyen	
w;UAS-dicer;mef2gal4	Gift from Mike Taylor	
w;UAS-eGFP-Nxt1-m1	Generated in HWC-lab	Caporilli et al. (2013)
w;UAS-eGFP-Nxt1-m2a	Generated in HWC-lab	Caporilli et al. (2013)
w;P(GD17336)v52631 (Nxt1 RNAi 52631)	VDRC (ID: 52631)	
P(KK107745)VIE-260B (Nxt1 RNAi 103146)	VDRC (ID:103146)	
Crosses	**Used for**	**Reference**
yw; *Nxt1*^{DG05102}/ *Nxt1*$^{Z2-0488}$	Analysing Nxt1 phenotype	Tested in this project
yw; *Nxt1*^{DG05102}/ *Nxt1*$^{Z2-0488}$;UAS-*abba*$^{fl\#\ 5}$;mef2gal4	Abba gene rescue	Tested in this project
yw; *Nxt1*^{DG05102}/ *Nxt1*$^{Z2-0488}$;UAS-*abba*$^{fl\#\ 5}$;armgal4	Abba gene rescue	Tested in this project
yw;UAS-eGFP-Nxt1-m1,*Nxt1* ^{DG05102}/*Nxt1*$^{Z2-0488}$;mef2gal4/+	Nxt1 rescue in muscles	Tested in this project
yw;UAS-eGFP-Nxt1-m2a,*Nxt1* ^{DG05102}/*Nxt1*$^{Z2-0488}$;mef2gal4/+	Nxt1 rescue in muscles	Tested in this project
yw;UAS-eGFP-Nxt1-m1,*Nxt1* ^{DG05102}/*Nxt1*$^{Z2-0488}$;armgal4/+	Nxt1 rescue in all cells	Tested in this project
yw;UAS-eGFP-Nxt1-m2a,*Nxt1* ^{DG05102}/*Nxt1*$^{Z2-0488}$;armgal4/+	Nxt1 rescue in all cells	Tested in this project
Nxt1 RNAi 103146/UAS-dicer;mef2gal4/+	Nxt1 RNAi in muscles	Tested in this project
UAS-dicer/+;Nxt1 RNAi 52631/mef2gal4	Nxt1 RNAi in muscles	Tested in this project
yw; *Nxt1*^{DG05102}/ *Nxt1*$^{Z2-0488}$;UAS-*mib2*/mef2gal4	Gene dosage experiment	Tested in this project
yw; *Nxt1*^{DG05102},*abba*^{df}/ *Nxt1*$^{Z2-0488}$	Gene dosage experiment	Tested in this project
yw; *Nxt1*^{DG05102},*mib2*^{df}/ *Nxt1*$^{Z2-0488}$	Gene dosage experiment	Tested in this project
yw; *Nxt1*^{DG05102}/ *Nxt1*$^{Z2-0488}$;*abba*^{dp}/+	Gene dosage experiment	Tested in this project

Appendix 3 – Adapter sequences

The index sequence of each ScriptSeq Index PCR primers used for RNA sequencing is:

5’- CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTCAGACGTGTGC TCTTCCGATCT-3’

Contents	**Index sequence**
Index 2 PCR Primer	5’-CGATGT-3’
Index 4 PCR Primer	5’-TGACCA-3’
Index 5 PCR Primer	5’-ACAGTG-3’
Index 6 PCR Primer	5’-GCCAAT-3’
Index 7 PCR Primer	5’-CAGATC-3’
Index 12 PCR Primer	5’-CTTGTA-3’

The index sequence of each TruSeq Small RNA Index PCR primers used for small RNA sequencing is:

5’-CAAGCAGAAGACGGCATACGAGAT**NNNNNN**GTGACTGGAGTTCCTTGGCACCCGA GAATTCCA-3’

Contents	**Index sequence**
Index 1 PCR Primer	5’-CGTGAT-3’
Index 2 PCR Primer	5’-ACATCG-3’
Index 3 PCR Primer	5’-GCCTAA-3’
Index 8 PCR Primer	5’-TCAAGT-3’

Appendix 4 – Small RNA library mixes and oligos

Ambion decade marker mix

0.5µl Decade Marker

6.5µl Nuclease free H_2O

1.0µl 10x Kinase Reaction Buffer

1.0µl Gamma-32P-ATP

1.0µl T4 PNK

Incubate 1 hour at 37°C

8.0µl Nuclease-free H_2O

2.0µl 10x Cleavage reagent

Incubate 5 minutes at RT

20.0µl Gel loading buffer II

Incubate 5 minutes at 95°C

Spikes mix

1.0µl Oligo 10µM

1.0µl PNK buffer

4.5µl Nuclease-free H_2O

1.0µl yATP

0.5µl PNK

Incubate at 37°C for 30 minutes

Incubate at 90°C for 1 minute

Dilute to 50µl with nuclease-free H2O and purify labeled oligos on Microspin G25 columns (Illustra^{tm}) according to manufacturer's protocol.

3' Linker ligation mix

2.0µl T4 RNA ligase buffer (10x)

5.0µl PEG8000 50%

2.0µl T4 RNA Ligase 2, truncated K227Q

5’ Linker ligation mix

2.0µl T4 RNA ligase buffer

5.0µl PEG8000 50%

2.0µl T4 RNA ligase

2.0µl ATP 10mM

0.5µl RNase inhibitor

KAPA PCR mix (KAPA Real Time library amplification kit)

25.0µl KAPA Hot Start ReadyMix 2x MM

2.0µl EvaGreen Dye 20x

0.25µl Solexa FW primer (100µM)

0.25µl Solexa RV primer (100µM)

10µl cDNA

12.5µl nfH_2O

PmeI digestion mix

33.0µl PCR product

4.0µl CutSmart

3.0µl Pme I

3’ linker
/5rApp/NNNNAGATCGGAAGAGCACACGRCR/3ddC/

5’linker
ACACUCUUUCCCUACACGACGCUCUUCCGAUCUNNNN

Solexa RT rev
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

Solexa PCR fw
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

TruSeq Barcode primers were used as reverse primers

Appendix 5 – Known and proposed genes in piRNA pathway

Known genes in piRNA pathway

Shu	*AGO3*	*Vas*	*Rhi*	*Squ*
Piwi	*Tej*	*Mael*	*Cuff*	*Vret*
Zuc	*Armi*	*spn-E*	*BoYb*	*CG5508*
Aub	*qin*	*wde*		

Proposed genes in piRNA pathway

del	*CG2181*	*CG9754*	*CG13741*	*CG3893*
CtBP	*Ars2*	*CG12721*	*Nxf2*	*TfIIA-S*
RnpS1	*CF14438*	*CG4415*	*CG10880*	*Nxf3*
His2Av	*pie*	*CG5694*	*Nurf-38*	*CG7504*
Acn	*Tho2*	*Zn72D*	*CG7946*	*Trn*
CG1737	*eff*	*CG7429*	*YL-1*	*nej*
CG33169	*CG4936*	*Ef1alpha48D*	*trc*	*Nup54*
eIF-2alpha	*L(2)37Bb*	*Actr13E*	*snapin*	*hang*
Vps4	*mun*	*CG7845*	*PCID2*	*CG17293*
Thoc5	*CG9536*	*mof*	*Pp4-19C*	*CG7728*
Woc	*Thox7*	*CG9548*	*Fib*	*Zif*
Su(var)205	*Uba2*	*Tfb4*	*RpS18*	*CG6066*
Aly	*hay*	*Ndc80*	*CG17801*	*CG5585*
CG13924	*Hsc70-5*	*Csl4*	*Bip1*	*Bor*
Ntf-2	*W*	*fra*	*CG2678*	*CG8108*
Rcd1	*eIF3-S10*	*CG5706*	*CG1677*	*Taf2*
Prosalpha7	*FLASH*			

End