From Gametogenesis to Oncogenesis: The Many Roles of DDX1

by

Devon Roy Germain

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

Doctor of Philosophy

in

Experimental Oncology

Department of Oncology

University of Alberta

ABSTRACT

DEAD box proteins are a family of RNA helicases implicated in every aspect of RNA biogenesis and metabolism. Modification of RNA secondary structure by DEAD box proteins has widespread effects on numerous cellular processes due to the ubiquitous nature of RNA. While the biochemical activity of several DEAD box proteins has been described, there is limited information into the biological functions of most family members. DEAD box 1 (DDX1) is over-expressed in some cancers, and has *in vitro* RNA/RNA and RNA/DNA unwinding activity. *In vivo* analyses have implicated DDX1 in RNA maturation, double-stranded DNA damage response and viral maturation. To date, there have been no reported DDX1-negative cell lines, with neuronal and germline cells showing the highest DDX1 levels.

Here, we describe three novel facets of DDX1 biology and function. First, we report that DDX1 is a novel and independent prognostic indicator in breast cancer. Both increased *DDX1* mRNA levels and cytoplasmic DDX1 protein localization are negative prognostic markers. Analysis of treatment-specific subgroups revealed that this effect is elevated in patients who receive adjuvant therapy, indicating that DDX1 may be predictive of patient response to treatment. Second, we report the generation of a novel *Ddx1*-null allele in *Drosophila melanogaster. Ddx1*-null flies are viable, but have phenotypes consistent with reduced metabolism and display aberrant gametogenesis. We also describe an interaction between Ddx1 and *Sirup* mRNA, which may underlie the metabolic disruption observed in mutant *Ddx1* flies. Third, we have found that *Ddx1*$^{-/-}$ mice

die at the pre-blastocyst stage. We also observed that mice carrying one *Ddx1*-*a*llele display transgenerational inheritance, with heterozygous *Ddx1* mice generated from heterozygous intercrosses producing significantly fewer wild-type progeny than expected. We attribute this to a novel modification of the wild-type allele in heterozygous mice and demonstrate that transgenerational inheritance of *Ddx1* is distinct from previously described cases of transgenerational epigenetic inheritance.

These combined studies address different aspects of DDX1 function, providing new insight into the spectrum of biological roles associated with DDX1. DDX1 appears to be under strict regulation, as evidenced by the fact that Ddx1 levels in mice do not vary between heterozygous and wild-type animals. Strict regulation of DDX1 levels may explain why prognostic effects are observed when DDX1 levels are elevated in breast cancer. Although we did not attempt to determine why high levels of DDX1 are associated with early recurrence, the phenotypes observed in our *Ddx1*-null flies suggest a role for Ddx1 in regulating metabolism. We propose that DDX1 is an integral regulator of metabolism and growth, and is itself tightly regulated. As such, deregulation of DDX1 can either promote tumourigenesis (when over-expressed) or death (when absent). This model of DDX1 action indicates that it may be possible to target DDX1 to improve patient outcome in those cancer patients with mis-expressed DDX1.

PREFACE

The research presented in this thesis was conducted with assistance or in collaboration as noted below.

A version of chapter 2 has been published as Germain DR, Graham K, Glubrecht DD, Hugh JC, Mackey JR, Godbout R. DEAD box 1: a novel and independent prognostic marker for early recurrence in breast cancer. Breast Cancer Res Treat. 2011 127(1):53-63. Immunohistochemistry and pathology was performed by Darryl Glubrecht. Gene expression microarrays were prepared by Dr. John Mackey and Dr. Kathryn Graham. Ki67 scoring was performed by Dr. Judith Hugh. I was responsible for data analysis and interpretation, pathology, imaging, and writing the manuscript. Dr. Roseline Godbout was involved in all stages of the project and in writing the manuscript.

A version of chapter 3 has been prepared for publication as Germain DR, Li L, Hildebrandt MR, Simmonds AJ, Hughes SC, and Godbout R. Loss of the *Drosophila* melanogaster RNA binding protein Ddx1 leads to reduced size and aberrant gametogenesis. RNA co-immunoprecipitation and PCR amplification of immunoprecipitated RNAs were performed by Dr. Lei Li. Phospho-specific western blot analysis was performed by Dr. Matthew Hildebrandt. Dr. Sarah Hughes and Dr. Andrew Simmonds provided materials and technical assistance. I was responsible for all other experimental procedures, conceptualizing novel experimental protocols, data analysis and interpretation, imaging and writing the manuscript. Dr. Roseline Godbout was involved in all stages of the project and in writing the manuscript.

A version of chapter 4 has been published as Hildebrandt MR*, Germain DR*, Monckton EA, Brun M, and Godbout R. Ddx1 knockout results in transgenerational wild-type lethality in mice. Sci. Rep. 5, 9829; DOI:10.1038 /srep09829 (2015). *These authors contributed equally to this work. Mouse genotyping and collection was performed by Dr. Matthew Hildebrandt and Elizabeth Monckton. Southern blot analysis was performed by Elizabeth Monckton. Bisulfite sequencing was performed by Dr. Matthew Hildebrandt. Quantitative real-time PCR was performed by Miranda Brun. I was responsible for model conceptualization, experimental design, data analysis and interpretation, and writing the manuscript. Dr. Matthew Hildebrandt assisted with data analysis and writing. Dr. Roseline Godbout was involved in all stages of the project and in writing the manuscript. This work received research ethics approval from the Cross Cancer Institute Animal Care Committee, #BC11185.

ACKNOWLEDGMENTS

The work comprising this thesis would not have been possible without the support of a number of people throughout my time at the University of Alberta. Above all else, I have received an enormous amount of support, guidance, and advice from my mentor, Roseline Godbout. In creating a welcoming, talented research group, Roseline has built a wonderful environment for graduate students to develop their skills and mature as researchers. Roseline's patience and dedication to her students has been a key factor in my success, and the success of numerous students before me. It is my hope that later in my career I will be able to emulate the compassionate nature in which she trains new scientists.

I have been fortunate to be part of a talented group of researchers throughout my graduate studies. In particular, Lei, with whom I have had the pleasure of working closely with for the entirety of my studies and who has been a valuable source for technical support and thoughtful consideration. Several of my lab mates have been extremely helpful over the years; Darryl, who's immunohistochemistry expertise was essential for my work in breast cancer; Matt, Liz, Ken and Stanley, the rest of the core DDX1 team; and Miranda and Rong-Zong, who provided much appreciated insight and helpful conversation. Many other lab members have also contributed over the years; Mike, Raja, Tina, Lucy, Saket, Indrani, David, Elizabeth, Amit, Kevin, Erin, Jack and Xia.

I am also privileged to have the support of my committee, Michael Hendzel and Sarah Hughes, who provided insight and direction whenever it was needed. Many other people on campus also aided my work; Andrew Simmonds graciously

allowed me extensive use of his laboratory space; Xuejun Sun and Gerry Barron provided excellent training and technical support at the Cross Cancer Institute Imaging facility; and John Mackey and Kathryn Graham enabled my breast cancer study.

Finally, this work would not have been possible without the love and support of my wife, Shaylyn. She has remained a bastion of strength for me through thick and thin, and I would not be the man I am today without her. I am also thankful for the understanding and support I received from my mother throughout the years of my perpetual studenthood and the boundless interest in my work from my father. While this thesis is presented as the accumulation of my work, it represents the product of efforts from numerous colleagues, friends and family, and for that I am immensely grateful.

TABLE OF CONTENTS

LIST OF TABLES

Title **Page**

LIST OF FIGURES

Title **Page**

LIST OF ABBREVIATIONS

Abbreviation	Full name
4EBP	eukaryotic translation initiation factor 4E binding protein 1
aa	amino acid
ADP	adenosine diphosphate
AGO3	Argonaute 3
APOBEC	apolipoprotein B mRNA editing enzyme
Akt	v-akt murine thymoma viral oncogene homolog 1
ATM	ataxia telangiectasia mutated
ATP	adenosine triphosphate
bp	base pair
BRCA1	breast cancer 1, early onset
BRCA2	breast cancer 2, early onset
CAGE	cancer-associated antigen gene
CAS9	CRISPR associated protein 9
cDNA	complimentary DNA
CRISPR	clustered regularly interspaced short palindromic repeats
CyO	Curly of Oster
DAPI	4',6-diamidino-2-phenylindole
dATF-2	*Drosophila* activating transcription factor-2
Dbp	DEAD box protein
BDSC	Bloomington *Drosophila* Stock Center
DDX	DEAD box protein
Dlnr	*Drosophila* insulin receptor
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide
Dpp	decapentaplegic
Dr	Drop
DSB	double strand break
DSN	duplex specific nuclease
dsRNA	double stranded RNA
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
eIF4	eukaryotic initiation factor 4
EJC	exon junction complex
ER	estrogen receptor
ESC	embryonic stem cell
foxo	forkhead box, subgroup O
Gapdh	glyceraldehyde-3-phosphate dehydrogenase
GFP	green fluorescent protein
GSK-3β	glycogen synthase kinase 3 beta
HDAC1	histone deacetylase 1
HER2	erb-b2 receptor tyrosine kinase 2
HIV	human immunodeficiency virus

hnRNP K	heterogeneous nuclear ribonucleoprotein K
HR	homologous recombination
IP	immune-precipitation
KSRP	KH-type splicing regulatory protein
MAPK	mitogen activated kinase-like protein
MBNL1	muscleblind-like splicing regulator 1
Me31b	maternal expression at 31B
Mg	magnesium
miRNA	micro RNA
mRNA	messenger RNA
mth	methuselah
mthl	methuselah-like
MYCN	v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog
MZT	maternal-to-zygotic transition
NA	numerical aperture
NAG	neuroblastoma amplified sequence
ncRNA	non-coding RNA
NHEJ	non-homologous end joining
NICHD	National Institute of Child Health and Human Development
NIH	National Institute of Health
NMD	nonsense-mediated mRNA decay
NSE1	non-SMC element 1 homolog
nt	nucleotide
PA	polyadenylation signal
PALB2	partner and localizer of BRCA2
PARP	poly-(ADP-ribose) polymerase
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDK1	pyruvate dehydrogenase kinase, isoform 1
PDK2	pyruvate dehydrogenase kinase, isoform 2
PI3K	phosphatidylinositol-4,5bisphosphate 3-kinase
PIP_2	phosphatidylinositol-4,5bisphosphate
piRNA	piwi-interacting RNA
PK	proteinase K
PMSF	phenylmethanesulfonylfluoride
Pol III	RNA polymerase III
PR	progesterone receptor
pS6k	phospho 70-S6 kinase
PTEN	phosphatase and tensin homolog
PVA	polyvinyl alcohol
PVDF	polyvinylidene fluoride
qPCR	quantitative real-time PCR
RB1	retinoblastoma protein
RFP	red fluorescent protein
Rheb	ras homolog enriched in brain

RIN	RNA integrity number
RNA	ribonucleic acid
RpS6	ribosomal protein S6
rRNA	ribosomal RNA
RtcB	RNA 2`,3`-cyclic phosphate and 5`-OH ligase
RT-PCR	reverse transcriptase polymerase chain reaction
S6k	RpS6-p70-protein kinase
SA	splice acceptor
SDS	sodium dodecyl sulfate
SF	super family
sgs	salivary gland secretion
siRNA	small interfering RNA
Sirup	Starvation-upregulated protein
SMN	survival of motor neuron
SMRT	single molecule real time
SnRNA	small nuclear RNA
ssRNA	single stranded RNA
TMA	tissue microarray
TORC1	target of rapamycin complex 1
TORC2	target of rapamycin complex 2
TP53	tumour protein p53
tRNA	transfer RNA
Tsc1	tuberous sclerosis 1
Tsc2	tuberous sclerosis 2
VDRC	Vienna *Drosophila* Resource Center
VEGF	vascular endothelial growth factor
Xist	inactive X specific transcripts
YB-1	Y box binding protein 1

CHAPTER 1

INTRODUCTION

1.1 RNA HELICASES

Unlike common illustrations of RNA, which are presented as a linear string of nucleotides, *in vivo* RNA molecules exist with complex secondary and tertiary structure (Figure 1.1). In addition, RNA molecules are bound by numerous proteins during each stage of RNA biogenesis. Complex structure and protein binding partners allow for intricate regulation of RNA and RNA-dependent processes within a cell. As a result, cells express numerous RNA binding and RNA modifying proteins. Amongst these are helicases, which unwind double-stranded nucleic acid strands. RNA and DNA helicases are broadly categorized into two major groups, with non-ring forming helicases comprising helicase superfamily (SF) 1 and 2, and ring forming helicases comprising SF 3 through 6 (Gorbalenya *et al.* 1988, Singleton *et al.* 2007).

The biochemical actions of ringed versus non-ringed helicases are distinct, though both traditionally use ATP as an energy source (Singleton *et al.* 2007). Ringed helicases generally act as hexamers, encircling DNA or RNA nucleotide strands and progressively translocating along the nucleotide strand as they unwind. Non-ringed helicases predominantly act by translocating along a single strand of the RNA molecule, displacing the up/downstream complementary strand as they relocate. Of note, ringed RNA helicases are only found in bacteria and viruses, with all known eukaryotic RNA helicases belonging to SF 1 and 2 (non-ringed).

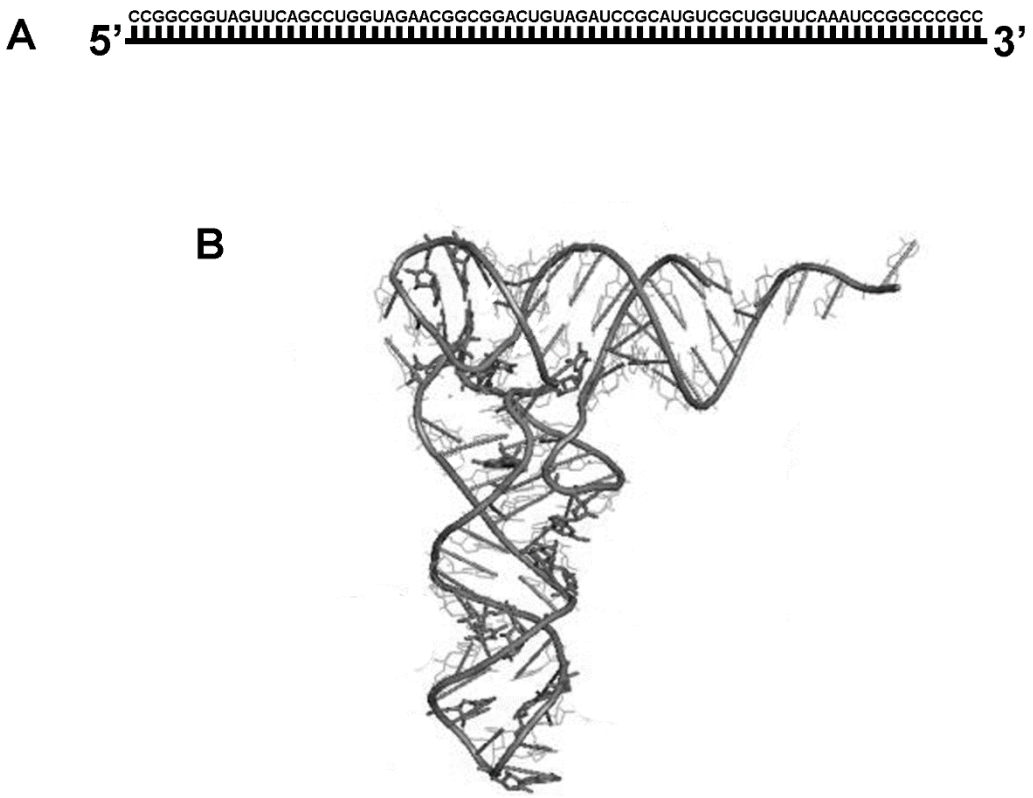


Figure 1.1 RNA structure. (A) A simple depiction of the sequence of a tRNA molecule. (B) 3d modelled structure of a tRNA molecule, adapted with permission from Jackman et al. (2013).

1.2 DEAD BOX PROTEINS

DEAD box proteins form the largest component of SF 2 (Jankowsky *et al.* 2007). They are characterized by 12 conserved motifs, including their namesake motif II: D(asp)-E(glu)-A(ala)-D(asp) (Linder *et al.* 2011, Linder *et al.* 1989). DEAD box proteins have been described in eukaryotes, prokaryotes and archaea (Anantharaman *et al.* 2002). Structurally, DEAD box proteins are composed of two globular domains that can exist in an open or closed configuration (Cordin *et al.* 2006). In contrast to other non-ringed RNA helicases, DEAD box proteins generally unwind by binding to and destabilizing local RNA structure, and do not translocate or act in a progressive manner (Mallam *et al.* 2012, Sengoku *et al.* 2006). As a result, DEAD box proteins generally only unwind shorter lengths (generally 6-17 bp) of double stranded RNA (dsRNA) (Chen *et al.* 2008, Rogers *et al.* 1999), though notably, p68 and p72 have been shown to unwind longer duplexes in some cases (Hirling *et al.* 1989, Rossler *et al.* 2001).

Though ATP hydrolysis is a property displayed by most DEAD box proteins (Andersen *et al.* 2006, Chen *et al.* 2008, Collins *et al.* 2009), it is not necessarily required for unwinding activity (Del Campo *et al.* 2007). However ATP hydrolysis may be essential for efficient release of the nucleotide strand post unwinding (Liu *et al.* 2008). Structural modelling of the conserved region of DEAD box proteins has shown that the interaction between RNA molecules and DEAD box proteins is almost entirely through the RNA backbone (Bono *et al.* 2006, Sengoku *et al.* 2006). This binding location precludes sequence specificity for most DEAD box proteins, suggesting that they are targeted to RNA molecules through other proteins, or bind

RNA non-specifically. Of note, some functions of DEAD box proteins do not require unwinding, with RNA-bound DEAD box proteins acting as an anchor for recruiting and interaction between the RNA molecule and other proteins (Andersen *et al.* 2006, Bowers *et al.* 2006, Tran *et al.* 2007).

In eukaryotic cells DEAD box proteins have been implicated in all stages of RNA biogenesis and metabolism, from transcription, splicing and trafficking to translation and degradation (reviewed in Linder *et al.* 2011). As a consequence of the ubiquitous nature of RNA, DEAD box proteins play a wide spectrum of roles in the cell. While there is a basic understanding of the role or biochemical activity of many DEAD box proteins, the scope of DEAD box protein function and impact on cellular processes is still limited. Below, the functions of a few of the well-studied DEAD box proteins are described.

1.2.1 eIF4A

eIF4A (eukaryotic initiation factor 4A) was the first DEAD box helicase described and is required for efficient translation in eukaryotes (Grifo *et al.* 1982). In conjunction with a number of other translation initiation factors (eIF4B, eIF4E and eIF4G, composing the eIF4F complex) (Conroy *et al.* 1990, Jackson *et al.* 2010), eIF4A facilitates the interaction between the small ribosomal subunit and the 5’ cap of mature mRNA molecules (Jackson *et al.* 2010). The loading rate of ribosomal complexes onto mRNA is inversely related to secondary structure complexity of the 5` terminal region, with simple unstructured mRNA not requiring eIF4A for efficient translation (Koromilas *et al.* 1992, Pelletier *et al.* 1985). eIF4A

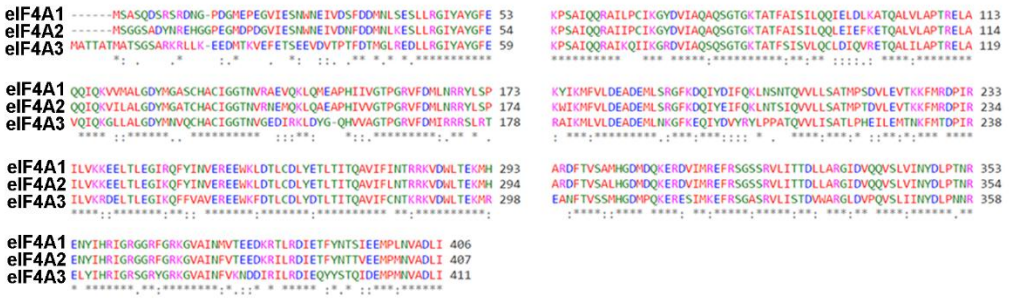
unwinds RNA secondary structures in this region, allowing for more efficient loading of ribosomes onto mRNA molecules and increasing translational rates (Rogers *et al.* 1999). Of interest, despite its name, eIF4A is not wholly restricted to eukaryotic cells, as an orthologue has been described in the archaea species *Methanococcus jannaschii* (Story *et al.* 2001).

Mammalian cells express three eIF4A homologues. eIF4A1 (DDX2A) and eIF4A2 (DDX2B) are closely related, with both localizing to the cytoplasm, whereas eIF4A3 (DDX48) is somewhat divergent and localizes to the nucleus (Figure 1.2A) (Chan *et al.* 2004). While *eIF4A1* and *eIF4A2* are 90% identical, they play distinct roles with differing expression patterns during the cell cycle (Williams-Hill *et al.* 1997). Both eIF4A1 and eIF4A2 bind to the eIF4F complex and play the canonical role of regulating ribosomal loading onto mRNA molecules (Jackson *et al.* 2010). In addition to translation initiation, eIF4A2 acts in the miRNA repression pathway (Meijer *et al.* 2013). Intriguingly, knock-down of eIF4A1 results in increased expression of eIF4A2, but the resulting higher levels of eIF4A2 do not compensate for the loss of eIF4A1 (Galicia-Vazquez *et al.* 2012). Consistently, over-expression of eIF4A1 does not rescue the effect of eIF4A2 depletion (Meijer *et al.* 2013).

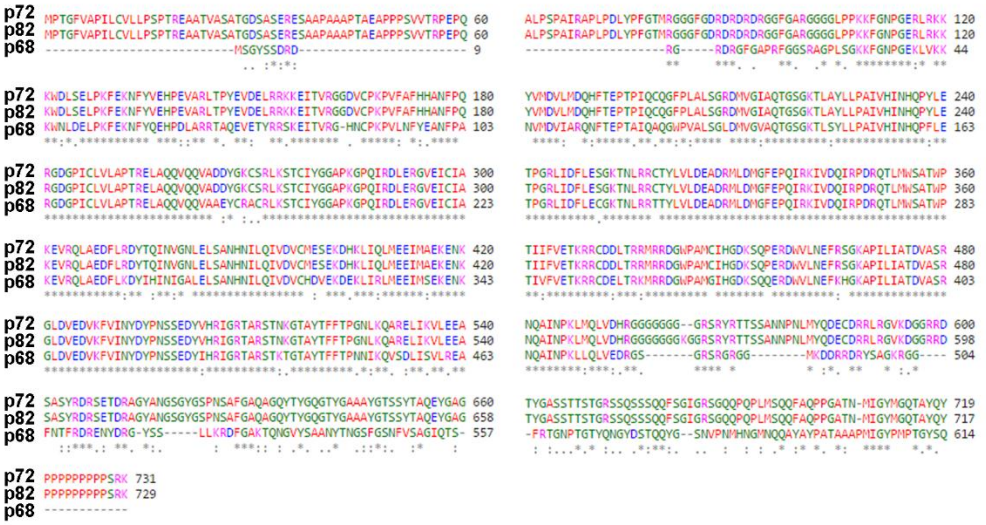
The least similar member of the eIF4A family, eIF4A3, is primarily involved in the exon junction complex (EJC) regulating nonsense mediated mRNA decay (NMD) (Chan *et al.* 2004, Palacios *et al.* 2004). The proposed function for eIF4A3 in EJC is not as an unwinding enzyme, but rather as an RNA clamp that serves as the anchor for the EJC (Zhang *et al.* 2007). Though eIF4A3 exhibits dsRNA

Figure 1.2 DEAD box protein alignment. Amino acid sequences of related DEAD box proteins were aligned using ClustalW2. (A) Human DEAD box proteins eIF4A1, eIF4A2 and eIF4A3 are aligned. (B) Human DEAD box proteins p72, the p72 variant p82 and p68 are aligned. (C) *Drosophila* DEAD box proteins Vasa and Belle are aligned.

A



B



C

unwinding activity *in vitro*, it seems unlikely that its role in EJC is dependent on unwinding activity, but there is some evidence that eIF4A3 may function in translation initiation in yeast (Alexandrov *et al.* 2011).

1.2.2 p68 and p72

p68 (*DDX5*) and *p72* (*DDX17*) are two closely related multifunctional DEAD box genes (Figure 1.2B). Together, they are known to modulate transcription, mRNA splicing, and processing of rRNAs and miRNAs (Caretti *et al.* 2006, Fukuda *et al.* 2007, Honig *et al.* 2002, Wilson *et al.* 2004). p68 and p72 have high sequence homology, interact with each other to form heterodimers and are generally implicated in cellular processes in tandem (Ogilvie *et al.* 2003). Of note, an alternatively spliced variant of p72 has been described that encodes 2 extra amino acids (p82), but it is not yet clear what functional difference, if any, exists between the two isoforms (Uhlmann-Schiffler *et al.* 2002). While p68 and p72 share some redundant functions, they also perform distinct functions as well and display different expression patterns (Dardenne *et al.* 2012, Honig *et al.* 2002, Jalal *et al.* 2007, Lamm *et al.* 1996).

As with the eIF4A family of proteins, many p68 and p72 functions rely on helicase activity, while others do not. Helicase-dependent functions include p68 modulation of a stem loop structure present in tau mRNA that in turn modulates alternative splicing (Kar *et al.* 2011). Similarly, the helicase activity of p68 and p72 has been implicated in microRNAs processing (Liu 2002, Salzman *et al.* 2007) and mRNA splicing (Guil *et al.* 2003, Honig *et al.* 2002, Lee 2002). In contrast, the role

of p68 and p72 as transcriptional modulators has generally been found not to be dependent on their helicase activity. p68 and p72, individually or in tandem, interact with and co-activate androgen receptor (AR), the vitamin D receptor, p53, MyoD and Runx2 (Bates *et al.* 2005, Caretti *et al.* 2006, Clark *et al.* 2008, Nicol *et al.* 2013, Wagner *et al.* 2012, Wong *et al.* 2009). p68 and p72 have also been implicated in transcriptional repression through interaction with the histone deacetylase HDAC1 (Wilson *et al.* 2004). In *Drosophila*, a single orthologue of p68 and p72 (Rm62) is expressed. Rm62 has been implicated in both transcriptional repression, through regulation of chromatin insulators, and in clearance of nascent RNA transcripts (Boeke *et al.* 2011, Buszczak *et al.* 2006).

1.2.3 Vasa

Similar to *eIF4A*, *p68* and *p72, VASA* (DDX4) has been implicated in many biological processes (Lasko 2013). The D*rosophila* orthologue of *Vasa*, *vas*, was originally identified in a screen for mutations that affect female fertility (Schupbach *et al.* 1986) and later identified as a component of polar granules in *Drosophila* (Hay *et al.* 1988). VASA is now recognized as a marker of germ line cells in numerous animals (Gustafson *et al.* 2010). Vasa protein is localized to polar granules within the pole plasm throughout early *Drosophila* development through an interaction with *osk* mRNA (Breitwieser *et al.* 1996).

While the function of Vasa in polar granules has not been fully delineated, it is known to mediate translation of both *Gurken* and *mei-P26* (meiotic P26) mRNA in an eIF4B dependent manner (Liu *et al.* 2009). Gurken is a TGF-α like growth

factor that plays essential roles in oogenesis at distinct developmental stages, whereas Mei-P26 regulates germ cell maintenance, both through BMP signalling and by regulating miRNA expression (Li *et al.* 2012, Neumuller *et al.* 2008). Given that mutations of both *gurken* and *mei-P26* can result in sterility, it is logical to propose that Vasa`s role in oogenesis is mainly due to its role in regulating the translation of these two, and possible other, mRNAs. Though also expressed in *Drosophila* testes, mutation of v*asa* has no effect on male fertility (Lasko *et al.* 1990). Intriguingly, though *vasa* mutation only affects female fertility in *Drosophila*, mouse *Vasa* is required for spermatogenesis (Tanaka *et al.* 2000) and hypermethylation of the promoter of the human *VASA* gene is associated with spermatogenesis defects (Sugimoto *et al.* 2009).

In addition to its role in gametogenesis, *Drosophila* Vasa has also been shown to play a role in chromatin condensation during female germline mitosis by facilitating the localization of components of the condensing I complex (Pek *et al.* 2011b). Of note, this role is not dependent on eIF5B suggesting that it cannot be attributed to Vasa-mediated translational regulation of other transcripts. Instead, this function is dependent on two genes involved in the piRNA (Piwi-associated RNA) pathway (*maelstrom* and *spindle-E*). piRNAs are small non-coding RNA molecules that repress transposon activity through post-transcriptional and epigenetic gene silencing (Brennecke *et al.* 2007). Under normal conditions, Vasa accumulates at nuage particles with several proteins required for piRNA generation (Aubergine and AGO3), but in *vasa* mutants these piRNA-related proteins do not localize to nuage particles, and piRNA generation is prominently

reduced (Malone *et al.* 2009). While a direct role linking piRNAs to chromatin condensation has not yet been described, there is evidence of non-coding RNAs (ncRNAs) modifying chromatin and playing a role in chromosome segregation (Claycomb *et al.* 2009, Pek *et al.* 2011a, Savic *et al.* 2014, van Wolfswinkel *et al.* 2009). A relationship between Vasa and Dicer, an essential protein for miRNA and siRNAs generation, has been described in flies, mice and *C. elegans*, suggesting a more general role for Vasa in small ncRNA biosynthesis (Beshore *et al.* 2011, Kotaja *et al.* 2006, Megosh *et al.* 2006).

Drosophila express a second DEAD box gene closely related to *vas*, *bel* (Figure 1.2C). *Belle (bel)* mutations can result in male and female sterility or larval lethality (Johnstone *et al.* 2005) and mitotic segregation defects are observed in the wing discs of *bel* mutants (Pek *et al.* 2011a). Belle has also been shown to localize to nuage bodies with Vasa (Johnstone *et al.* 2005). Though similar, Belle and Vasa differ significantly in several ways. In contrast to *Vasa, Belle* expression is not enriched in pole plasm. Furthermore, *vasa* mutant progeny die during early embryogenesis, whereas *bel* mutants die during larval development (Johnstone *et al.* 2005). While the specific cause of lethality in *bel* mutants has not been elucidated, Belle has been shown to regulate ecdysone signalling through translational repression of an ecdysone induced transcription factor (Ihry *et al.* 2012). De-regulation of ecdysone signalling could plausibly be the cause of the observed larval lethality in *bel* mutants.

Amongst related DEAD box genes, high sequence similarity does not appear to be indicative of overlapping function, and individual genes often have

multiple, seemingly unrelated, roles. The ubiquitous nature of RNA in or as a precursor to all cellular processes means that individual DEAD box proteins may be involved in multiple processes by modifying different RNA targets' secondary structure. In addition, the non-helicase dependent roles of DEAD box proteins allow for an even greater range of functions, from serving as RNA clamps to facilitate protein binding to localization of RNA through interaction with other proteins to modulate translation. The widely divergent roles of the above described DEAD box genes in addition to the unrelated functions of related genes provide an example of the wide scope of DEAD box protein function.

1.2.4 DEAD box genes in mice

To date there has been limited work into the role of *Ddx1* in mice. siRNA knock-down of *Ddx1* in a mouse spermatogonial cell line resulted in reduced levels of *cyclin-D2*, *CD9*, and GDF3 and reduced cell proliferation (Tanaka *et al.* 2009). While cell lines represent a useful tool for understanding gene function within a cellular context, animals carrying mutations allow for more insightful analysis of the effect of genes of interest. Currently, no previous *Ddx1* mutant mouse lines have been reported. However, knock-out of other DEAD box genes have been described in the literature.

As previously mentioned, knock-out of *Vasa* results in male sterility in mice (Tanaka *et al.* 2000). Knock-out of *Ddx25* results in a phenotype similar to Vasa, with males being infertile, though the disruption to spermatid development appears at an earlier stage than in *Vasa* knock-out males (Tsai-Morris *et al.* 2004b). DDX25

plays a role in mRNA transcript trafficking and storage, in turn regulating expression of genes integral to spermatogenesis (Sheng *et al.* 2006, Tsai-Morris *et al.* 2004a). While both *Vasa* and *Ddx25* knock-outs are viable, most of the other DEAD box genes which have been knocked-out in mice are not.

Knock-out of *ChlR1* (*Ddx11*) has been implicated in sister chromatid cohesion, and results in embryonic lethality due to aneuploidy (Inoue *et al.* 2007, Parish *et al.* 2006). The earliest described lethality associated with a DEAD box gene in mice is cause by knock-out of *DP103* (*Ddx20*), resulting in lethality prior to the four-cell stage, potentially due to disruption of the maternal to zygotic transition, in which zygotic genes become active and maternally supplied protein and RNA undergo widespread degradation (Mouillet *et al.* 2008). Finally two knock-out strains have been generated for *Ddx58*, one of which is embryonic lethal, while the other survive to adulthood but develops gastrointestinal problems (Kato *et al.* 2005, Wang *et al.* 2007). It is not clear why these two different *Ddx58* knock-out mice lines have different phenotypes, but it is likely due to differences between the lines used to generate the individual mutations.

1.2.5 DEAD box genes in flies

A number of DEAD box genes in addition to the previously described *vasa* and *bel,* a number of other DEAD box proteins have been studied in flies. Several of these have also been implicated in germ line development. Whereas *Gemin3*, a member of the SMN (survival of motor neurons) complex, is essential, targeted disruption of *Gemin3* in ovaries results in aberrant oogenesis (Cauchi 2012). In

Gemin3 mutants prior to death, or in flies with less severe disruption of *Gemin3*, motor defects are observed, in keeping with the role of the SMN complex in motor neuron development and maintenance (Cauchi *et al.* 2008, Shpargel *et al.* 2009). Me31B (DDX6) interacts with Vasa in polar granules, and has been implicated in piRNA generation (Liu *et al.* 2011, Thomson *et al.* 2008). During oogenesis, nurse cells generate mRNA transcripts intended to be loaded into the oocyte Me31b forms RNP complexes to repress inappropriate translation of these maternal transcripts (Nakamura *et al.* 2001). Me31b has also been implicated in regulating translation in neuronal cells in a similar manner (Hillebrand *et al.* 2010). Finally, while little is known about Dbp73d, it is expressed in germ line cells, indicating a potential role in gametogenesis (Patterson *et al.* 1992).

Other *Drosophila* DEAD box proteins have described roles in a variety of biological processes. *Drosophila* express two *eIF4A* genes, *eIF4A1* and e*IF4A3*. Similar to its human orthologue, *Drosophila* eIF4a plays additional roles outside of translation initiation, including a role in early oogenesis and regulating Dpp signalling (Jagut *et al.* 2013, Li *et al.* 2006a, Li *et al.* 2005, Shen *et al.* 2009a). eIF4A3 also mirrors the function of its human orthologue, playing a role in EJC (Zhang *et al.* 2007). Another DEAD box gene, *UAP56*, plays a role in RNA export and EJC (Gatfield *et al.* 2002, Gatfield *et al.* 2001, Meignin *et al.* 2008). UAP56 has also been implicated in piRNA biogenesis in conjunction with Vasa (Zhang *et al.* 2012). Two DEAD box genes, *abstrakt* and *Dbp21E2*, are involved in retinal development (Hibbard *et al.* 2012, Schmucker *et al.* 2000). *abstrakt*, an essential gene, is required for axonal outgrowth and the regulation of asymmetrical mitotic

divisions during development in the retina and other neuronal tissues, (Irion *et al.* 1999, Irion *et al.* 2004), while *Dbp21E2* is required to maintain Rhodopsin levels (Hibbard *et al.* 2012). Lastly, Pitchoune has been implicated in the regulation of proliferation and is potential target of dMyc (Zaffran *et al.* 1998).

1.3 DDX1

DDX1 was originally identified as being over-expressed in a subset of retinoblastoma and neuroblastoma cell lines and tumours (Godbout *et al.* 1998, Godbout *et al.* 1993a, Squire *et al.* 1995). Retinoblastomas can occur as multifocal (heritable) or unifocal (sporadic) disease (Ries *et al.* 1999). Multifocal retinoblastomas are caused by inherited or novel germline mutations in the retinoblastoma gene (RB1) while unifocal cases are associated with sporadic RB1 mutations (Ries *et al.* 1999). While there is a general consensus that inactivation of RB1 is the major cause of retinoblastoma, several other genomic aberrations are commonly observed (Theriault *et al.* 2014). Amplification of *MYCN* is observed in approximately 3% of cases (Bowles *et al.* 2007, Corson *et al.* 2007), and more recently, *MYCN* amplification has also been implicated as a possible cause of a subset of retinoblastoma tumours which still retain functional RB protein (Rushlow *et al.* 2013).

The *DDX1* gene locus resides less than half a megabase from the *MYCN* locus *(*Figure 1.3A*).* Amplification of this region can include *DDX1* as well as two other genes at a lower rate (*NAG* and *NSE1*) (Beheshti *et al.* 2003, Squire *et al.* 1995, Wimmer *et al.* 1999). In retinoblastoma, *DDX1* is almost always co-amplified

with *MYCN* (Godbout *et al.* 2007, Godbout *et al.* 1993a, Rushlow *et al.* 2013). In contrast co-amplification of *DDX1* occurs in approximately one-third of *MYCN* amplified neuroblastomas (Godbout *et al.* 1998, Squire *et al.* 1995). *MYCN/DDX1* amplification has also been reported in alveolar rhabdomyosarcoma, medulloblastoma, glioblastoma multiforme, and rat uterine endometrial carcinoma (Adamovic *et al.* 2005, Barr *et al.* 2009, Fruhwald *et al.* 2000, Hodgson *et al.* 2009). While MYCN is a well-documented oncogene (Freier *et al.* 2006, Heine *et al.* 2010, Teitz *et al.* 2011), it is not entirely clear what role, if any, *DDX1* amplification plays in *MYCN/DDX1* amplified tumours. A competitive advantage for tumour cells over-expressing DDX1 can be inferred though, as the amplified copies of *DDX1* produce increased levels of DDX1 protein (Godbout *et al.* 1998). Of note, DDX1 is predominantly nuclear in most cells; however when over-expressed in retinoblastoma and neuroblastoma cells, a significant proportion of DDX1 localizes to the cytoplasm (Godbout *et al.* 1998).

The amino acid sequence of DDX1 is highly conserved across the DEAD box core when compared other DEAD box proteins. Unique to DDX1 is the inclusion of a SPRY domain (Figure 1.3B). The presence of extra domains within the helicase core of DEAD box proteins is rare, suggesting unique properties for DDX1. There are over 100 SPRY domain containing proteins in the human genome, with SPRY domains found throughout all three domains of life (Rhodes *et al.* 2005). Structurally modular, SPRY domains are believed to be involved in protein-protein interactions and/or protein-RNA interactions (Perfetto *et al.* 2013). DDX1 orthologues have been identified in all animals, and a few other eukaryotes,

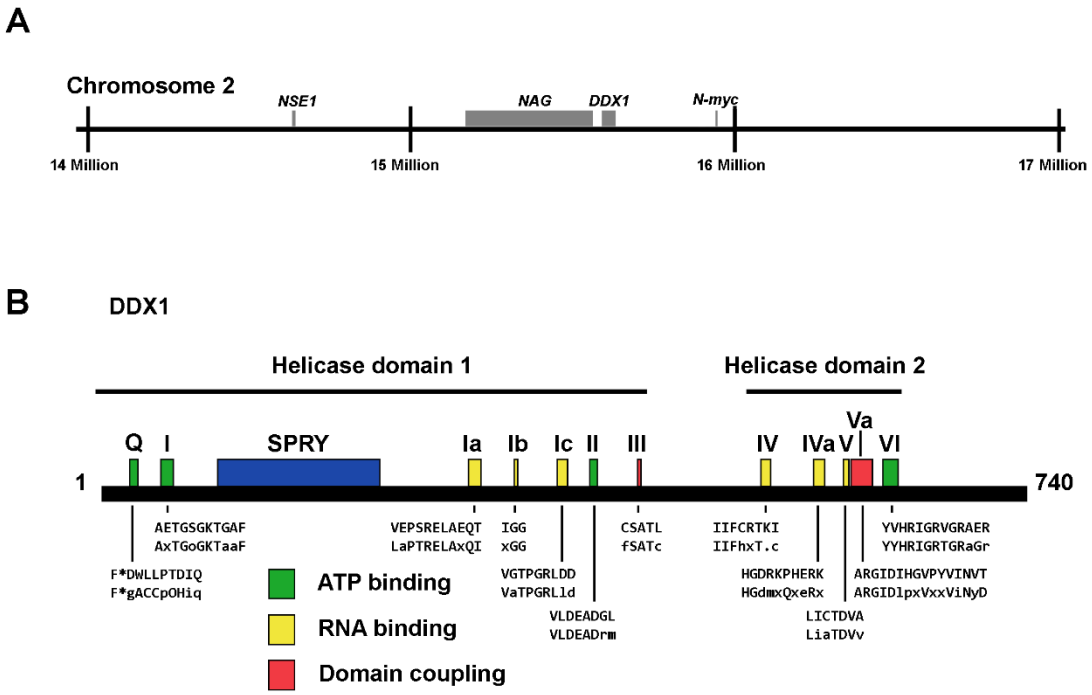


Figure 1.3 DDX1 genomic location and protein domains. (A) Human *DDX1* is located at chromosome 2p24. This region can be co-amplified with *MYCN* in cancers such as retinoblastoma. Nearby genes which are sometimes co-amplified are noted. (B) The 12 conserved domains within DDX1 are shown to scale. DDX1 sequence is indicated above the conserved DEAD box helicase sequence. Domain function is indicated by colour. * designates a 16 amino acid span.

but is not expressed in prokaryotes or fungi. Gene expression analysis has shown ubiquitous expression of DDX1 in all cell lines tested (Godbout *et al.* 1998, Godbout *et al.* 1993a). In chicken, *DDX1* mRNA is found at highest levels in retina, brain, heart, liver and kidney (Godbout *et al.* 2002). In *Drosophila,* highest *DDX1* levels are observed in neuronal and germ cells (Rafti et al. 1996). In both of these studies, DDX1 was more highly expressed in early developmental stages.

1.3.1 Biochemical activity

In vitro analysis has shown that human DDX1 can unwind both RNA/RNA and RNA/DNA duplexes (Li *et al.* 2008). RNA/DNA unwinding is uncommon in DEAD box proteins, but a small number of DEAD box proteins have been shown to have this activity (Rozen *et al.* 1990). DDX1 unwinding activity is further differentiated from the majority of DEAD box proteins, as DDX1 is activated by ADP, as opposed to ATP, and DDX1 is able to unwind duplexes up to 29 bp in length (as opposed to most other DEAD box proteins that unwind 6-17 bp) (Kellner *et al.* 2015, Li *et al.* 2008). In a recent report, DDX1 has been shown to bind ADP with 500X more affinity than other DEAD box proteins (Kellner *et al.* 2015). It is not yet clear if DDX1 hydrolyzes ADP, or if ADP binding is sufficient for its unwinding activity. In addition to unwinding RNA/RNA and RNA/DNA duplexes, DDX1 also degrades ssRNA (Li *et al.* 2008). This activity is distinct from RNAse A activity and energy independent, but requires Mg^{2+} and is heat sensitive.

1.3.2 Role in RNA biogenesis

Generation of mature functional RNA molecules requires numerous cellular processes. In eukaryotic cells, following transcription, RNA molecules are modified by addition of a 5` cap (7-methylguanosine) and/or a 3' poly(A) tail, splicing, RNA editing and cleavage. Furthermore, RNA function can also be modified through RNA stability/degradation and RNA subcellular localization, processes that underly RNA trafficking and transport. Thus, a single locus can generate transcripts with different roles depending on how the RNA is processed. Given that RNA biogenesis and maturation can be regulated by RNA secondary structure, it is not surprising that DEAD box proteins are essential for each of these steps.

In human cells, DDX1 has been implicated in several steps of RNA biogenesis and regulation. Under normal growth conditions DDX1 displays a strong nuclear signal and forms a small number of foci that co-localize with or reside next to cleavage bodies (associated with 3' mRNA processing and polyadenylation) and Cajal bodies and gems (associated with rRNA and histone mRNA processing and snRNA and snoRNA biogenesis) (Bleoo *et al.* 2001, Li *et al.* 2006b). DDX1 was also shown to co-localize with MBNL1 and YB-1 in stress granules upon arsenite treatment (Onishi *et al.* 2008). Stress granules are densely packed structures which selectively hold mRNA molecules in a translationally stalled state to modify gene expression patterns in response to stress stimuli (reviewed in Kedersha *et al.* 2013).

DDX1 is a component of neuronal RNA trafficking granules, RNA/protein complexes that transit RNA molecules along axons for localized translation (Kanai

et al. 2004, Perez-Gonzalez *et al.* 2014). Consistent with these roles DDX1 has also been shown to interact with heterogeneous nuclear ribonucleoprotein K (hnRNP K), a component of heterogeneous nuclear ribonucleoprotein complexes that are involved in regulating RNA transcription, transport and translation (Chen *et al.* 2002). DDX1 also interacts with KSRP, which regulates mRNA decay and microRNA maturation (Chou *et al.* 2013). More recently, DDX1 has been found to be part of the tRNA splicing complex and is required for efficient turnover of an RtcB-guanylate intermediary required for splicing (Popow *et al.* 2011, Popow *et al.* 2014). Finally, there is some evidence that DDX1 may promote maturation of a subset of miRNAs (Han *et al.* 2014).

1.3.3 Viral role

DDX1 is not limited to regulating endogenous RNA molecules. Several viruses have been shown to modulate DDX1 localization or require DDX1 for efficient replication. The initial discovery that linked DDX1 to viral replication identified an interaction between DDX1 and Rev, a viral protein responsible for the nuclear export of unspliced HIV RNA genomes (Edgcomb *et al.* 2012, Fang *et al.* 2005, Fang *et al.* 2004, Lin *et al.* 2014). RNA viruses initially require processing of RNA transcripts in order to facilitate translation of viral proteins, but at later stages require unprocessed whole genomic RNA molecules for the generation of new viruses. Intriguingly, several other DEAD box proteins (DDX3, DDX5, DDX17, DDX21 and DDX56) also can interact with Rev and promote RNA export (Naji *et al.* 2012, Yasuda-Inoue *et al.* 2013, Yedavalli *et al.* 2004).

In addition to HIV, DDX1 has been shown to play a role in the replication of several other viruses. DDX1 acts as a transactivator of the JC viral gene promoter and promotes viral proliferation (Sunden *et al.* 2007a, Sunden *et al.* 2007b). DDX1 also enhances coronavirus replication, and notably, was observed to localize to the cytoplasm following coronavirus infection (Wu *et al.* 2014, Xu *et al.* 2010). While it is still not clear what effect cytoplasmic localization has on DDX1, it is intriguing that this cytoplasmic localization has been previously observed in cancer cells that over-express DDX1.

1.3.4 Role in DNA repair

Even under normal conditions, cells accumulate DNA damage. Amongst the different types of DNA damage, the most deleterious and dangerous to a cell are DNA double strand breaks (DSBs). Left unrepaired, DSBs can lead to genetic instability and ultimately cell death or cancer. DSBs are generally repaired through one of two major pathways, non-homologous end joining (NHEJ) or homologous recombination (HR) (reviewed in Aparicio *et al.* 2014). Two less common DSB repair pathways have also been described (microhomology-mediated end joining and single strand annealing), but will not be discussed here (Bennardo *et al.* 2008, Heyer *et al.* 2010). NHEJ is an error prone form of repair, which acts by degrading the DNA immediately surrounding the DSB and directly ligating the two ends. This process generates a repaired DSB, but often results in loss of genetic information. HR is an error free form of repair, which requires 5’ end resection of both ends of the DSB. The resulting single stranded DNA is then able to invade a homologous

sequence (generally a sister chromatid). The homologous sequence is used as a template to repair the damaged DNA strand. HR rates are enhanced when homologous sequences are available (S, G2 and M phase), while NHEJ is predominant when there are no homologous sequences present (G1) (Mao *et al.* 2008). While this simple cell cycle-dependent model describes the underlying bases for DSB repair by HR versus NHEJ, these two forms of DNA repair can both occur at any point during the cell cycle, and it is not entirely clear how one pathway is chosen over the other (Mao *et al.* 2008).

In addition to its roles in RNA biogenesis, DDX1 has been implicated in the cellular response to DNA DSBs. Following exposure to ionizing radiation (IR), DDX1 forms foci that co-localize with a subset of γH2AX foci, a histone variant that quickly localizes to sites of DSBs (Li *et al.* 2008). DDX1 localization to DSBs is dependent on ATM, a kinase that plays an integral role in the signal cascade activated upon DSB formation. As well, DDX1 localization to DSBs requires active transcription and its retention at DSBs requires RNA/DNA duplexes (Li *et al.* 2008). Given the importance of RNA/DNA duplexes for the co-localization of DDX1 with γH2AX, it has been proposed that the presence of DDX1 and RNA at DSBs marks these sites for repair by HR. Thus, DSBs near sites of active transcription may preferentially be repaired by HR, thereby utilizing an error-free mechanism to repair genetically important regions, while leaving the repair of DSBs in untranscribed regions to the more efficient, but error prone, NHEJ.

1.3.5 DDX1 in cancer

DDX1 is co-amplified with *MYCN*, a well-established oncogene, in some tumours and cell lines (Godbout *et al.* 1998, Squire *et al.* 1995). As such, it is not immediately obvious if DDX1 is playing a significant role in tumour biogenesis or is simply being co-amplified as the result of its proximity to *MYCN*. However, as previously mentioned, increased DDX1 protein levels are observed in *DDX1/MYCN*-amplified cancer cells, suggesting that there is an active selective pressure to express high levels of DDX1 protein (Godbout *et al.* 1998). To date there have been 5 studies to determine the effect of DDX1 amplification in neuroblastoma patients, with the largest study suggesting a better outcome for patients with *MYCN/DDX1* amplification compared to *MYCN* amplification alone. No clear conclusions were derived from the other studies because of sample size (De Preter *et al.* 2005, de Souza *et al.* 2011, George *et al.* 1996, Kaneko *et al.* 2007, Weber *et al.* 2004). DDX1 has also been implicated in testicular and ovarian cancers, with low DDX1 levels being associated with a poor outcome in ovarian cancer patients (Han *et al.* 2014, Tanaka *et al.* 2009).

1.4 DROSOPHILA MELANOGASTER

Drosophila melanogaster has served as a model system for genetic analysis since the early 1900's (Castle 1906, Morgan 1910). Their high reproductive rate in conjunction with their short life-cycle allows for large scale analyses, while not requiring excessive resources. While other simpler and more economical model organisms can be used, the fruit fly model represents a valuable balance between

convenience and complexity, allowing analysis of both basic cellular processes and the more complicated processes involved in developmental patterning and organ function. Over the past century, researchers have accumulated numerous useful tools for molecular analysis in flies. In addition, the creation of several large scale consortias who freely share generated material have resulted in flies becoming arguably the most effective invertebrate eukaryotic model system for investigating cellular processes and gene function.

1.4.1 Development

Drosophila development can be divided into four stages: embryonic, larval, pupal, and adult. Mature *Drosophila e*ggs are fertilized internally just prior to being laid. Early rounds of mitosis are highly choreographed, with nuclei undergoing simultaneous divisions. Notably, cytokinesis does not occur during this period, resulting in a syncytial embryo. During this stage of development, a subset of nuclei are segregated to the posterior end of the embryo and cellularize precociously, becoming pole cellsthat are the precursors to the eventual germ line cells of the adult fly (Mahowald 2001). After approximately two hours, the embryo contains roughly 6000 nuclei (Foe *et al.* 1983). At this point, cellularization occurs and gastrulation begins. Over the next 20 hours, the unpatterned embryo develops into a 1st instar and hatches out of the egg.

The 3 larval stages (1st, 2nd and 3rd instar stadium) are periods of growth. Over a period of 4-5 days the larvae grow from ~0.5 mm to more than 3 mm long. Larval stages are punctuated with 2 moltings, where larvae shed their cuticle after

the 1st and 2nd instar stages to allow further expansion of body size. Growth and molting are carefully regulated by ecdysone, a major developmental hormone that cycles throughout development to control timing of key events, and juvenile hormone (Di Cara *et al.* 2013). 3rd instar which have consumed sufficient nutrients enter a wandering stage, where they leave the food source they have been inhabiting and climb upwards to find an appropriate site to initiate metamorphosis.

Once a suitable site has been found, larvae undergo pupariation. As with previous developmental events, pupariation begins with the release of ecdysone from the brain (Di Cara *et al.* 2013). The larval cuticle undergoes tanning, becoming a puparium. The pupae then retract from the puparium in a process called apolysis. Once free from the puparium, the pupae undergo metamorphosis, with major body segments forming and imaginal discs developing into adult tissues. The last steps of metamorphosis include neuronal development and adult cuticle development. Finally, an adult fly ecloses from the puparium.

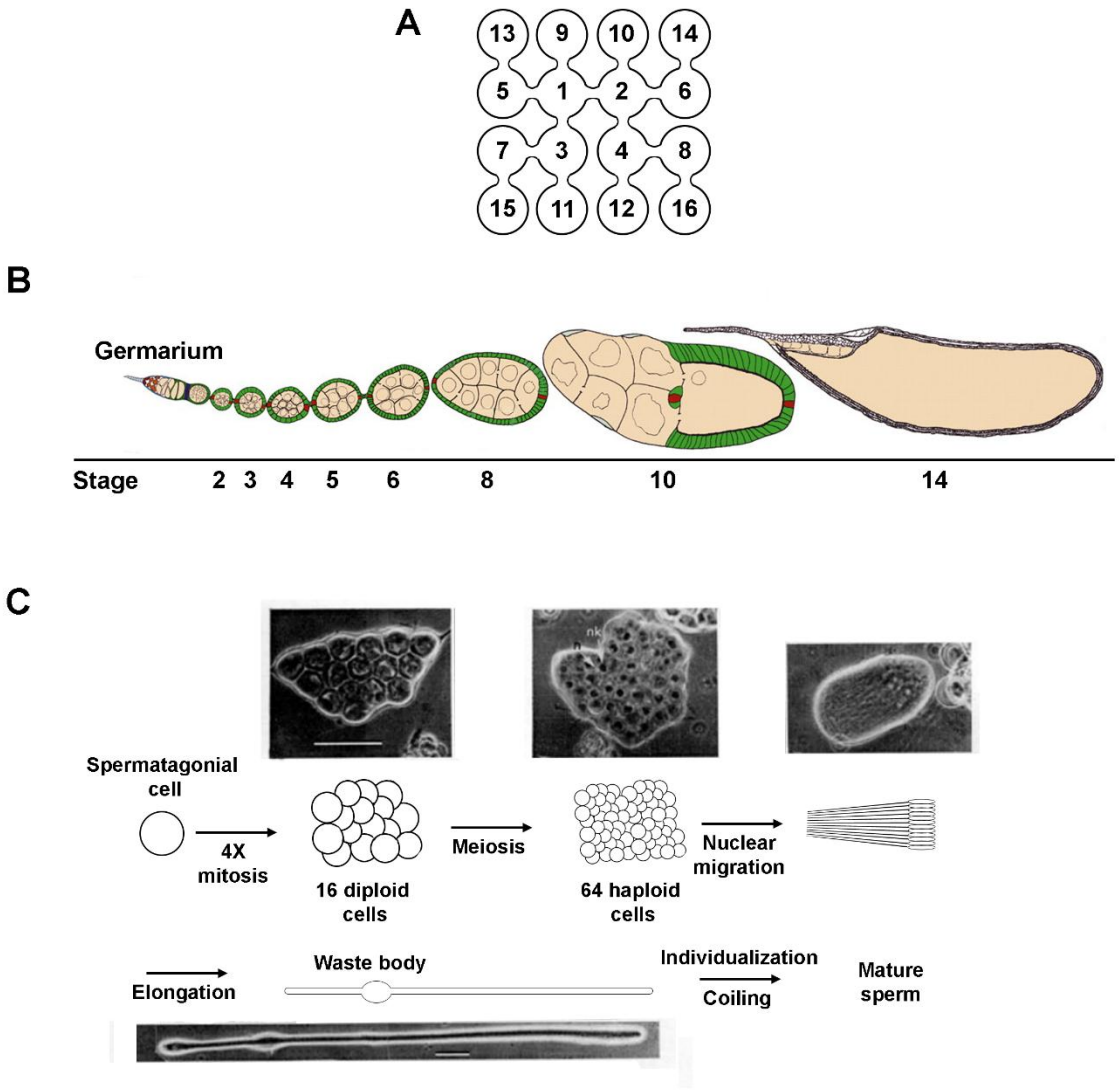
1.4.2 Gametogenesis

While all animals produce ova and sperm through meiosis, the development and structure of germ cell-producing gonads vary. The development of *Drosophila* ovaries has been well characterized (Bastock *et al.* 2008). Briefly *Drosophila* ovaries develop from pole cells, and consist of a bundle of ~16-18 ovarioles that are connected through a single oviduct. Each ovariole independently generates egg chambers from a group of germ line stem cells at its apical tip. Following asymmetrical division of one of the germ line stem cells, a single celled cystoblast

is released into the ovariole. The cytoblast then undergoes four rounds of mitosis to generate 16 cells interconnected by ring canals (Figure 1.4A). Of the 16 cells, 1 will undergo meiosis, becoming the oocyte, while the remaining 15 undergo DNA replication with no cytokinesis, becoming polyploid nurse cells. These nurse cells generate RNA, protein and cellular organelles that are transferred to the oocyte. In addition to nurse cells, the egg chamber is surrounded by a population of follicle cells. Egg chambers grow in size as they travel along the length of the ovariole until they reach maturity (Figure 1.4B). Mature eggs are held by the fly until an appropriate site to lay is found, at which point the mature egg passes through the oviduct and is fertilized by sperm held in the spermatotheca.

Spermatogenesis differs significantly from oogenesis, but both processes begin with a germ-line stem cell population in the apical tip of the gonad. As with oogenesis, asymmetrical mitosis of a germ-line stem cell generates a primary spermatogonial cell that moves along the length of the testis. The spermatogonial cell undergoes four rounds of mitosis with incomplete cytokinesis to generate a cyst of 16 interconnected cells (Figure 1.4C). At this point, each of the 16 cells within the spermatocyte undergoes meiosis, generating a total of 64 spermatids that remain connected by cytoplasmic bridges. The developmental period following meiosis is known as spermiogenesis (Fabian *et al.* 2012). During spermiogenesis, a marked morphological transformation takes place, with mitochondria within spermatids fusing into two large mitochondria that

Figure 1.4 *Drosophila* gametogenesis. (A) The single cell cystoblast undergoes 4 rounds of mitosis with incomplete cytokinesis, resulting in 16 interconnected cells as depicted. Cell number one or two will become the oocyte and undergo meiosis, while the remaining cells will become polyploid nurse cells. (B) *Drosophila* oogenesis within an ovariole. A stem cell population resides in the anterior tip of the ovariole (left). Egg chambers develop as they move along the length of the ovariole. The oocyte and nurse cells are tan, with follicle cells colored green. Adapted from Frydman *et al.* (2001). (C) *Drosophila* spermatogenesis. A single spermatogonial cell undergoes 4 rounds of mitosis and meiosis to generate 64 haploid spermatids. Spermatids then undergo nuclear migration, elongation, individualization and coiling to generate mature sperm. Scale bars equal 50 μm. Phase contrast micrographs adapted with permission from Cross *et al* (1979).


A
13
9
10
14
5
1
2
6
7
3
4
8
15
11
12
16
B
Germarium
Stage
2
3
4
5
6
8
10
14
C
Spermatagonial cell
4X mitosis
16 diploid cells
Meiosis
64 haploid cells
Nuclear migration
Elongation
Waste body
Individualization
Coiling
Mature sperm

interweave to form a Nebenkern, referred to as the onion stage due to the onion-like structure of the Nebenkern (Tokuyasu 1975). Following this transformation, spermatids begin to assemble flagellar axonemes and elongate. The two large mitochondria separate and elongate alongside the axoneme, eventually resulting in a structure that extends the length of the spermatid tail. During this period the spermatid nucleus also undergoes a morphological change, from rounded to a thin needle shape. The final steps in the maturation of spermatid are individualization and coiling (Tokuyasu *et al.* 1972a, Tokuyasu *et al.* 1972b). Beginning at the spermatid head, cytoplasm is moved into a waste body that travels along the length of the spermatid. As the cytoplasm is expelled, intercellular cytoplasmic bridges are removed, resulting in individual sperm. Following individualization, sperm tails coil into a tightly packed structure. Mature sperm are then held in the seminal vesicle until mating.

1.4.3 Tor signalling

The *Drosophila* life cycle can be divided into two major periods, the first involving a high growth rate (pre-pupation) and the second involving very little growth (post-pupation). As with all animals, regulation of growth is essential, as inappropriate growth signalling during times of limited resources can result in damage or even death to the organism. In animals, the insulin/Tor pathway acts to match growth and metabolism with resource intake. The *Drosophila* insulin/Tor pathway has been well characterized and is similar to the mammalian pathway (Grewal 2009). The *Drosophila* insulin receptor (DInr) is a receptor tyrosine kinase

that, in response to *Drosophila* insulin-like peptides (DILPS), is activated by auto-phosphorylation (Fernandez *et al.* 1995, Gronke *et al.* 2010). Several proteins are then recruited to activated DInr, instigating a signalling cascade that enhances PI3K-based production of PIP_3 and MAPK signalling (Bohni *et al.* 1999, Maehama *et al.* 1999, Poltilove *et al.* 2000, Scanga *et al.* 2000). Binding of PIP_3 recruits PDK1, PDK2 and Akt to the cell membrane (Oldham *et al.* 2003). PDK1 and PDK2 can then directly phosphorylate S6k (Flynn *et al.* 2000), while Akt phosphorylates Tsc2, GSK-3β and foxo (Cross *et al.* 1995, Potter *et al.* 2002, Puig *et al.* 2003). Phosphorylation of these targets stimulates metabolism and growth by up-regulating Tor signalling (through repression of Tsc2) (Yang *et al.* 2006), up-regulating *dMyc* mRNA levels (by activating GSK-3β) (Parisi *et al.* 2011) and localizing foxo to the cytoplasm, repressing expression of catabolic genes (Hay 2011) (Figure 1.5).

Tor exists in two distinct protein complexes, Tor complex 1 (TORC1) and Tor complex 2 (TORC2) (Inoki *et al.* 2006). TORC1 is a key intermediary in regulating cellular growth in response to nutrient and energy levels (Oldham 2011). During low resource conditions, TORC1 is repressed by Tsc1/Tsc2 inhibition of Rheb, an essential activator of TORC1 (Nobukini *et al.* 2004, Zhang *et al.* 2003). In response to high levels of amino acids, TORC1 activity is up-regulated through rag GTPases recruiting Tor to Rheb (Kim *et al.* 2008). Activated TORC1 then stimulates cell growth though regulation of mRNA translation initiation via three main effector molecules, S6k, 4EBP and TIF-1A (Grewal *et al.* 2007, Miron *et al.* 2003, Zhang *et al.* 2000). S6k functions by phosphorylating target proteins to promote protein

synthesis, the best characterized of which is a subunit of the small ribosomal subunit, RpS6, phosphorylation of which increases initiation of protein synthesis (Volarevic *et al.* 2001). 4EBP binds to and represses eIF4E, but phosphorylation of 4EBP disassociates it from eIF4E, allowing more efficient translation initiation (Beretta *et al.* 1996). Finally, TIF-1A is essential for ribosomal synthesis and phosphorylation of TIF-1A by Tor promotes the nuclear localization of TIF-1A (Grewal *et al.* 2007). In addition, TORC1 has recently been shown to up-regulate tRNA levels, through inhibition of the Pol III inhibitor Maf1 (Figure 1.5) (Rideout *et al.* 2012). Relatively little is known about TORC2, which has been primarily implicated in regulating the actin cytoskeleton in several different organisms (Cybulski *et al.* 2009). However, there is recent evidence linking TORC2 to cellular growth in *Drosophila* (Hietakangas et al. 2007).

1.5 BREAST CANCER

Cancer is the leading cause of death in Canada (Statistics Canada 2011). More than 40% of Canadians are expected to develop some form of cancer over their lifetime and an estimated 30% of all Canadians will die due to cancer (Canadian Cancer Society 2014). In women, breast cancer is the most common cancer, comprising more than a quarter of all diagnosed cancers, and the second most common cause of cancer related death (American Cancer Society 2013). Breast cancer does occur in men as well, though rarely, representing less than 1% of new cancers diagnosed in men (American Cancer Society 2010). While survival rates for breast cancer patients have seen significant increases in recent decades,

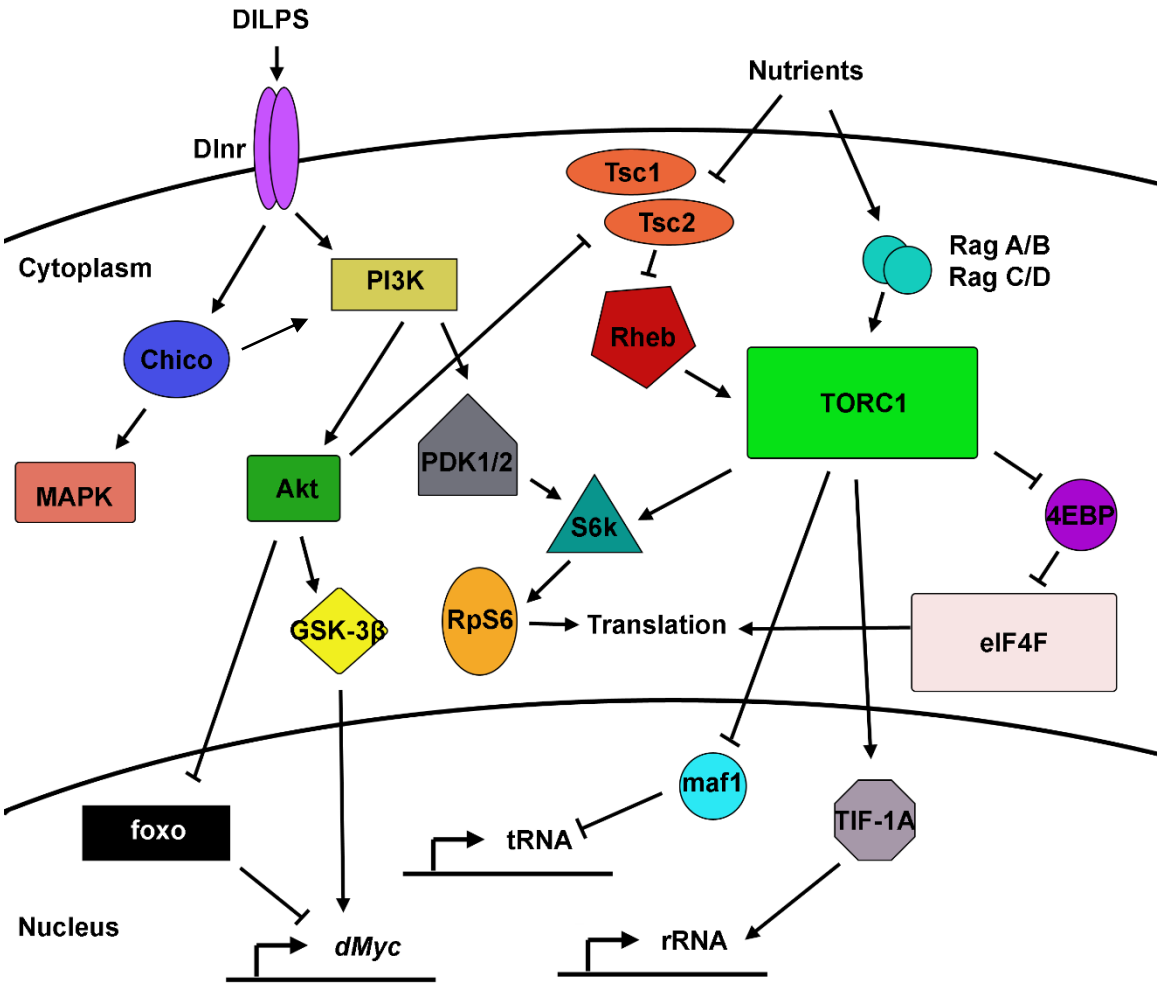


Figure 1.5 Schematic depiction of a simplified version of insulin and Tor signalling in *Drosophila*. In response to DILPS and nutrient signalling, DInr and Tor generate signal cascades that promote growth and proliferation through increased expression of *dMyc,* and increased levels of tRNA and rRNA levels, accompanied by increased translational rates.

with an 88% 5-year survival rate, approximately 30% of patients relapse (Jones 2008). Recurrent breast cancer patients often present with treatment resistant disease and have a much worse prognosis with an estimated 5-year survival of 21% (Hayat *et al.* 2007).

While the majority of breast cancer cases are caused by somatic mutations, an estimated 5-10% are heritable, 90% of which are associated with mutations in the breast cancer 1, early onset *(BRCA1*) and breast cancer 2, early onset (*BRCA2*) genes (Gage *et al.* 2012). BRCA1 and BRCA2 play essential roles in DNA repair (Aparicio *et al.* 2014). Reduced capacity for DNA repair due to these mutations results in the accumulation of additional mutations, which can eventually lead to uncontrolled growth and tumourigenesis. More recently, a number of other genes have been identified as causing inheritable breast cancer susceptibility, with most being involved in DNA repair (e.g. *ATM* and *PALB2*), or cell cycle regulation (e.g. *TP53* and *PTEN*) (Economopoulou *et al.* 2015).

1.5.1 Molecular subtypes

Cancer is by nature a heterogeneous disease, and while it has been long understood that individual tumours differ from one another, historically, treatment for “individual cancer types” (i.e. breast, lung, colon) has been uniform. This was in part due to limitations in identifying and elucidating the significance of molecular markers. However, with recent advances in molecular biology, including whole genome and transcriptome sequencing, it is becoming increasingly clear that each

major cancer type is made up of numerous sub-types, opening the door to a more personalized approach to the treatment of cancer.

In 2000, a large scale study by Perou *et al.* revealed that breast cancers can be clustered into subgroups based on expression profiles, response to treatment and prognosis (Perou *et al.* 2000). Breast cancers are now classified into four major subtypes: luminal A, luminal B, *HER2*-amplified or triple-negative/basal-like. Luminal A and luminal B are both estrogen receptor (ER) and/or progesterone receptor (PR) positive, and are differentiated based on proliferative rates (Ki67 scoring) and *HER2* status (Cheang *et al.* 2009). *HER2*-amplified are ER- and PR-negative and, as the name suggests, *HER2*-amplified. The triple-negative/basal-like cancers, while commonly considered a single subtype, represent a much wider variety of breast cancers (Badve *et al.* 2011). While triple-negative cancers are by definition ER, PR and *HER2* negative, a proportion of basal-like breast cancers are not (Rakha *et al.* 2007). Moreover, though the majority of triple-negative tumours display basal-like expression profiles, there is a sub-population that deviates from these profiles (Bertucci *et al.* 2008). Finally, a fifth subtype is recognized in the literature, normal-like tumours, though it has been proposed that the gene expression profile in these cases is biased by higher concentrations of non-tumourigenic cells (Parker *et al.* 2009).

1.5.2 Treatment and prognosis

Thanks to advances in screening and treatment options, breast cancer patients have seen tremendous improvement in survival (Sledge *et al.* 2014). This

is at least partially due to increased awareness resulting in increased screening, as patients who present with early stage disease have much better predicted outcome (Berry *et al.* 2005). In addition, the development of a number of new targeted drugs in conjunction with efficient screening for these targets in patients has led to the beginnings of a more personalized approach to the treatment of cancer.

Breast cancer treatment includes surgical resection, radiation therapy, chemotherapy, hormone therapy, and targeted drug therapy. Surgical resection is the primary treatment option for solid tumours, and often effectively removes the majority of the tumour mass. In non-invasive cases, surgery can be sufficient as treatment, but where a portion of the main tumour mass is left behind, or invasion and metastasis has occurred, additional treatment is required. Depending on individual circumstances, resection can be either a lumpectomy (removal of the tumour and some surrounding tissue), or a mastectomy (removal of the entire affected breast). Sentinel lymph nodes or axillary lymph nodes are commonly removed to determine if tumour cells have migrated outside of the localized area (Giuliano *et al.* 1994).

Adjuvant therapies are given following surgical resection in order to kill any remaining non-localized tumour cells. Adjuvant therapy consists of two major treatment avenues, radiotherapy and systemic therapy. Radiotherapy utilizes ionizing radiation to kill cells and is targeted to areas that are suspected to contain residual cancer cells (tumour peripheries or lymph nodes). Careful shaping of

radiation beams delivered from multiple angles allows high doses of radiation to target areas, with surrounding healthy tissue receiving a much lower dose.

Systemic therapies are treatments delivered system wide, and are either targeted to tumour cells, or have increased effects in tumour cells compared to normal cells. Chemo-, hormone, and targeted drug therapy fall into this category. Chemotherapeutic agents are cytotoxic drugs that kill cells through a variety of mechanisms, but generally have stronger effects in rapidly dividing cells. This characteristic ensures preferential killing of tumour cells, although some rapidly dividing normal cells also exhibit sensitivity to common chemotherapeutic agents (e.g. hair follicles and components of the immune system). As such, administering chemotherapy needs to be a balance between delivering a tumour-killing dose while not causing too much damage to normal cells. In spite of significant side effects, chemotherapy remains a key tool in cancer treatment.

The ability to sustain unregulated growth and division is one of the key hallmarks of cancer (Hanahan *et al.* 2011). Although tumours have intrinsic properties leading to unrestrained growth, it's important to keep in mind that tumours exist within complex microenvironments and sometimes rely on external growth factors. Many types of tumours express hormone receptors, including breast cancers, prostate cancers, endometrial cancers and adrenal cancers receptors (Shen *et al.* 2009b, Sissung *et al.* 2014, Werner *et al.* 2014). ER-positive breast cancers (luminal A and B) rely, at least in part, on estrogen to promote proliferation, and high estrogen levels are correlated with increased breast cancer risk (Hankinson *et al.* 2004). Therapeutics which reduce estrogen levels

(aromatase inhibitors) or antagonize ER activity (e.g. tamoxifen) have been shown to significantly improve patient outcome in ER positive breast cancer patients with fewer side effects than standard chemotherapeutic agents (Schiavon *et al.* 2014).

Targeted drug therapy utilizes small molecules or monoclonal antibodies that interact with a specific protein target. These treatments commonly take advantage of over- or uniquely-expressed tumour proteins to reduce side effects in normal cells. In breast cancer, several targeted therapies are in use or currently undergoing clinical trials for treatment. Monoclonal antibodies, the most prominent being trastuzumab which binds to HER2, have been successfully used to improve outcome in *HER2*-amplified breast cancer patients (Teplinsky *et al.* 2014). VEGF and the VEGF receptor are also targets for treatment, though the efficacy of bevacizumab (a monoclonal antibody that targets the VEGF ligand), used to treat metastatic breast cancer, has been brought into question (Schneider *et al.* 2011).

Although proteins that are unique or over-expressed in cancer represent valuable targets for drug development, there are other aspects of tumour biology that can be exploited. BRCA1 and BRCA2 mutations are common in hereditary breast cancers, resulting in reduced DNA repair by HR (Roy *et al.* 2012). Alternative DNA repair pathways can compensate for defects in HR; however, as these pathways are error-prone, the cells will begin to accumulate mutations. While a moderate mutational rate can lead to tumourigenesis and cancer progression, a much higher rate of DNA damage will outright kill cells. Inhibitors of PARP, a protein essential for the repair of single strand breaks, have little effect on healthy cells with normal HR rates, but in BRCA1/2 mutants result in epistatic or synthetic

lethality (Dedes *et al.* 2011). This is caused by unrepaired single strand breaks that are converted to double strand breaks upon DNA replication, which in turn cannot be repaired by HR due to BRCA deficiency. There are currently several ongoing clinical trials testing the efficacy of PARP inhibitors in the treatment of HR-deficient cancers (Lee *et al.* 2014).

Prognostic outcome varies between the different molecular subtypes, with luminal A having the best prognosis, followed by luminal B *HER2*-amplified and triple negative/basal-like breast cancers (Kreike *et al.* 2007). Treatment regimens are also tailored to each molecular subtype, with luminal cancers treated with hormonal therapy, and *HER2*-amplified breast cancers treated with HER2 targeted antibodies or drugs. However, many breast cancers do not respond to targeted treatment and others become resistance to treatment (Gonzalez-Angulo *et al.* 2007). This problem is compounded in patients with advanced or metastatic disease. Thus, while significant progress has been made in both the understanding of the molecular biology of breast cancer and the development of new treatments, breast cancer remains a major cause of death for women. New avenues need to be explored to find effective treatment for all breast cancers.

1.6 CHAPTER SUMMARIES

1.6.1 Chapter 2

Breast cancer is a heterogeneous disease characterized by diverse molecular signatures and a variable response to therapy. Clinical management of breast cancer is guided by the expression of estrogen and progesterone receptors

and HER2 amplification. New prognostic and predictive markers, as well as additional targets for therapy, are needed for more effective management of this disease. We used both gene expression microarrays and tissue microarrays to investigate the significance of DDX1 expression on the prognosis of breast cancer patients. For these analyses, we examined a 176-patient cohort, half of which had been selected based on early relapse despite standard adjuvant therapy, but were otherwise matched for estrogen receptor and HER2 status, stage and duration of follow-up. We identified *DDX1* RNA overexpression as an independent prognostic marker for early recurrence in primary breast cancer, with a hazard ratio of 4.31 based on logrank analysis of Kaplan–Meier curves. Elevated levels of DDX1 protein in the cytoplasm were also independently correlated with early recurrence with a hazard ratio of 1.90. Our data indicate a strong and independent association between poor prognosis and deregulation of the DEAD box protein DDX1. We propose that elevated levels of *DDX1* RNA or the presence of DDX1 in the cytoplasm could serve as an effective prognostic biomarker for early recurrence in primary breast cancer.

1.6.2 Chapter 3

Mammalian DDX1 has been implicated in RNA trafficking, DNA double-strand break repair and RNA processing; however little is known about its role during development. We report phenotypes associated with a *Ddx1*-null mutation generated in *Drosophila melanogaster*. *Ddx1*-null flies are viable but significantly smaller than their control counterparts. Female *Ddx1*-null flies show reduced

fertility with egg chambers undergoing autophagy, whereas males are sterile due to disrupted spermatogenesis. Comparative RNA sequencing of control and *Ddx1*-null third instar larvae identified several transcripts affected by Ddx1 inactivation. One of these was *Sirup* mRNA, shown previously to be overexpressed under starvation conditions and implicated in mitochondrial function. We demonstrate that *Sirup* is a direct binding target of Ddx1 and that *Sirup* mRNA is differentially spliced in the presence or absence of Ddx1. Combining *Ddx1*-null with *Sirup* knock-down causes epistatic lethality not observed in either single mutant. Our data suggest that the role of *Drosophila* Ddx1 includes stress-induced regulation of RNA splicing.

1.6.3 Chapter 4

To further investigate the role of DDX1 during development, mice carrying a constitutive Ddx1 knock-out allele were generated. $Ddx1^{+/-}$ mice have no obvious phenotype and express similar levels of DDX1 as wild-type mice indicating compensation from the intact *Ddx1* allele. Heterozygote matings produce no viable knock-out progeny, with $Ddx1^{-/-}$ embryos dying prior to embryonic day E3.5. Intriguingly, the number of wild-type progeny is significantly decreased in heterozygote-heterozygote crosses, with two different heterozygote populations identified based on parental genotype: (i) normal $Ddx1^{+/-}$ mice which generate the expected number of wild-type progeny and (ii) $Ddx1^{*/-}$ mice (with * signifying an altered allele) which generate a significantly reduced number of wild-type progeny. The transgenerational inheritance of wild-type lethality observed upon crossing

Ddx1$^{*/-}$ mice is independent of gender and occurs in cis through a mechanism that is different from other types of previously reported transgenerational epigenetic inheritance.

CHAPTER 2

DDX1: A NOVEL AND INDEPENDENT PROGNOSTIC MARKER FOR EARLY RECURRENCE IN BREAST CANCER

A version of Chapter 2 has been published as Germain DR, Graham K, Glubrecht DD, Hugh JC, Mackey JR, Godbout R. DEAD box 1: a novel and independent prognostic marker for early recurrence in breast cancer. Breast Cancer Res Treat. 2011 127(1):53-63. Immunohistochemistry and pathology was performed by Darryl Glubrecht. Gene expression microarrays were prepared by Dr. John Mackey and Dr. Kathryn Graham. Ki67 scoring was performed by Dr. Judith Hugh. I was responsible for data analysis and interpretation, pathology, imaging, and writing the manuscript. Dr. Roseline Godbout was involved in all stages of the project and in writing the manuscript.

2.1 INTRODUCTION

Breast cancer is the second highest cause of cancer-related death in women, with approximately one million new cases diagnosed each year worldwide (Porter 2008). While there have been significant advances in the development of endocrine and chemotherapy-based therapies for the treatment of breast cancer, approximately 30% of women with early stage disease will eventually relapse (Jones 2008), and those with distant metastases have less than a three percent chance of long term survival (Barnadas *et al.* 2008, Bergh *et al.* 2001, Fossati *et al.* 1998, Lopez-Tarruella *et al.* 2009). The molecular pathways and events underlying recurrence in breast cancer are poorly understood. To compound this problem, breast cancer represents a conglomerate of many different clinical and pathological diseases characterized by different genetic alterations, growth properties and responses to therapy.

A number of clinical and molecular parameters have traditionally been used to classify breast cancers including stage, grade (number of mitoses, nuclear architecture and tubule formation), estrogen receptor (ER) and progesterone receptor (PR) status, and HER2 (ERBB2) amplification. Recent molecular profiling based on hormone receptor status, HER2 amplification, and proliferation rates have resulted in the widely-accepted classification of breast cancer into four major subtypes: luminal A (ER+ve, PR+ve, low proliferation, and HER2-ve); luminal B (ER+ve, PR+ve, with either a higher proliferative index or HER2+ve); HER2-amplified (ER-ve, PR-ve and HER2+ve); and triple-negative or basal-like (ER-ve, PR-ve, and HER2-ve) (Cheang *et al.* 2009, Fan *et al.* 2006, Hugh *et al.* 2009,

Perou *et al.* 2000). Of these four subtypes, luminal A breast cancers have the best prognosis, with tumours responding well to adjuvant hormone therapy. In the absence of a specific target for therapy, triple negative tumours have the worst prognosis (Kreike *et al.* 2007). While this molecular classification allows for more precise prognosis and treatment recommendations, there is still considerable response variation within each subtype (Brenton *et al.* 2005).

Similarly, genome wide transcriptome analysis has defined multi-gene signatures reflecting breast cancer subtypes. Several multi-gene signatures with varying prognostic significance have been reported (Chiuri *et al.* 2007, Oakman *et al.* 2009, Stadler *et al.* 2009). The 21-gene recurrence score assay and the 70-gene signature MammaPrint are currently being marketed as prognostic tools for breast cancer (Albain *et al.* 2010, Slodkowska *et al.* 2009). Recent reports suggest that these multigene assays help identify which patients will benefit from chemotherapy (Mook *et al.* 2010). In spite of these advances, it is clear that we need: (i) a better understanding of the events underlying early relapse in breast cancer, (ii) novel prognostic markers that can independently predict recurrence, and (iii) new approaches to the treatment of breast cancers with a poor prognosis.

DEAD box 1 (DDX1) is a member of the D(Asp)-E(Glu)-A(Ala)-D(Asp) box protein family of RNA unwinding proteins (Linder 2006). *DDX1* is amplified and over-expressed in a subset of retinoblastoma and neuroblastoma tumours (George *et al.* 1996, Godbout *et al.* 1993a, Manohar *et al.* 1995, Squire *et al.* 1995) and has recently been reported to be involved in the development of testicular tumours (Tanaka *et al.* 2009). *DDX1* is widely expressed in different cell types and tissues,

albeit at different levels (Godbout *et al.* 1993a), and appears to be essential for embryonic development as mutation of the *DDX1* gene results in early embryonic lethality in both mice (our unpublished data) and *Drosophila melanogaster* (Zinsmaier *et al.* 1994). A number of roles have been proposed for DDX1 including RNA processing, RNA transport from the nucleus to the cytoplasm and RNA clearance at sites of double strand breaks (Bleoo *et al.* 2001, Fang *et al.* 2004, Kanai *et al.* 2004, Li *et al.* 2008, Li *et al.* 2006b). Although DDX1 is predominantly a nuclear protein, it is also found in the cytoplasm of *DDX1*-amplified neuroblastomas and retinoblastoma cells (Godbout *et al.* 1998).

Here, we examine DDX1 expression and subcellular location in gene expression microarrays and tissue microarrays designed to identify biomarkers associated with early recurrence in primary breast cancer. We show that over-expression of *DDX1* RNA (by as little as 40%) and elevated levels of DDX1 protein in the cytoplasm can both serve as prognostic markers of recurrence and death. Correlation of DDX1 with recurrence is independent of the commonly used breast cancer markers ER, PR, HER2 amplification, grade and stage, thus identifying DDX1 as a novel prognostic marker.

2.2 MATERIALS AND METHODS

2.2.1 Patient selection

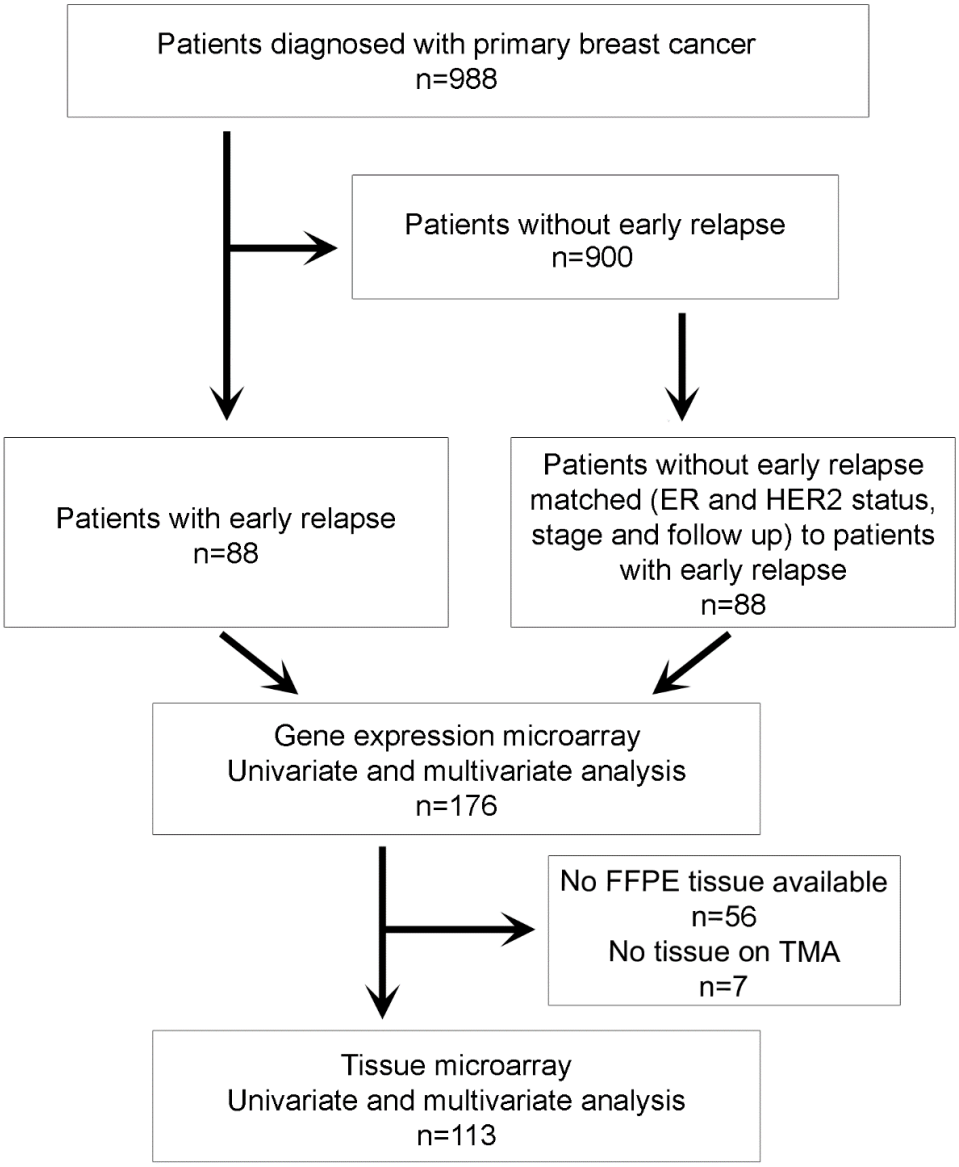
Gene expression microarray analysis was performed on 176 primary, treatment- naive breast cancer samples and ten normal breast tissue samples acquired from reduction mammoplasties through the Canadian Breast Cancer Foundation Tumour Bank. A flow chart depicting patient selection criteria is presented in Figure 2.1. Patient information was collected under Research Ethics Board Protocol ETH-02-86-17. The tumour samples, collected at surgery, were frozen in liquid nitrogen within 20 min of devitalization. Evaluation of histology slides from tissue adjacent to the frozen samples indicated that at least 70% of the cells present were invasive tumour cells.

2.2.2 Gene expression analysis

Total RNA was isolated from the frozen samples using Trizol and QIAGEN RNeasy columns. The RNA was quantified using a NanoDrop 1000 spectrophotometer and its integrity evaluated using a Bioanalyzer 2100. RNA samples with RNA Integrity Numbers (RIN) greater than 7.0 were used.

The RNA was subjected to linear amplification and Cy3 labeling, then hybridized to Agilent Whole Human Genome Arrays using Agilent Technologies kits (One Color Low RNA Input Linear Amplification Kit Plus, One Color RNA Spike-In Kit and Gene Expression Hybridization Kit). The arrays were scanned using an Agilent scanner. The data were extracted and quality-evaluated using Feature Extraction Software 9.5, and normalized and analyzed using GeneSpring GX 7.3 (Agilent Technologies).

Figure 2.1 Flow chart depicting patient selection for the study. Of the original population of 988 consented patients with treatment-naïve primary breast cancer, 88 had suffered an early relapse by September 30, 2009 when the data were locked. The women whose tumors were selected for this study received standardized guideline-based chemo- and hormonal therapies. These treatment guidelines recommend anthracycline chemotherapy for high risk node-negative disease, anthracycline with taxane chemotherapy for node-positive disease, hormonal therapy for all patients with ER+ve disease, and trastuzumab for those with HER2 positive tumors. Two groups of patients were selected for analysis, the first consisting of 88 patients who experienced an early relapse (less than 5 years after the initial treatment), and the second consisting of 88 patients who did not relapse. The two groups were matched for ER and HER2 status, stage and time of follow-up. The median duration of follow-up for surviving patients was 4.5 years. Abbreviation: FFPE, formalin fixed paraffin embedded.


Patients diagnosed with primary breast cancer
n=988
Patients without early relapse
n=900
Patients with early relapse
n=88
Patients without early relapse matched (ER and HER2 status, stage and follow up) to patients with early relapse
n=88
Gene expression microarray
Univariate and multivariate analysis
n=176
No FFPE tissue available
n=56
No tissue on TMA
n=7
Tissue microarray
Univariate and multivariate analysis
n=113

2.2.3 Tissue microarray construction and immunohistochemical analysis

The TMA included three 0.6 mm cores from each of the samples and was constructed using a TMArrayer (Pathology Devices, Westminster, Maryland). TMAs were deparaffinized in xylene, re-hydrated and microwaved for 20 min in epitope retrieval buffer (10 mM citrate, 0.05% Tween-20; pH 6). TMAs were immunostained with rabbit anti-DDX1 antibody (1:2000) (batch 2910) (Godbout *et al.* 1998) or mouse anti-Ki67 antibody (clone MIB-1; proliferation marker) (DakoCytomation, Carpinteria, California).

2.2.4 Scoring and quantification of immunohistochemical staining

Ki67 scoring to measure proliferative index was performed by a single pathologist (JH) blinded to outcomes, using the MIB1 antibody and dichotomized at 15% nuclear staining (Hugh *et al.* 2009). DDX1 protein was scored separately for nuclear and cytoplasmic levels. Each score was based on the average staining intensity throughout the tumour tissue on a scale of 0 to 3. Cytoplasmic staining was considered high if the score was 2 to 3 and low if the score was 0 to 1, while nuclear staining was considered high if the score was 3 and low if the score was 1 or 2 (no tumours were scored 0 as all tumours had some DDX1 in the nucleus). With few exceptions, staining intensity was consistent in all tumour cells throughout a single core. DDX1 staining was scored by DG with 63 random samples independently scored by JH. The agreement between the two sets had a Cohen's kappa value of 0.69 (substantial agreement) with complete agreement in 85% of cases for cytoplasmic intensity and a Cohen's kappa value of 0.57 (moderate

agreement) with complete agreement in 78% of cases for nuclear intensity. Acquisition of images was performed using an Axioskop2 plus microscope with a 20x or 40x lens, a ZeissAxiocam and AxioVision software, version 4.7.1.0 (Carl Zeiss MicroImaging, Jena, Germany).

2.2.5 Statistical analysis

Statistical analyses were performed using MedCalc for Windows, version 11.1.0.0 (MedCalc Software, Mariakerke, Belgium). Rank correlation was performed to determine Spearman's Rho. Clinical/pathological variables as a function of DDX1 scores were assessed for both gene expression microarrays and TMAs using the Students *t* test (continuous variables), Fisher's exact test (2 category variables) and chi square test (3 or more category variables). Survival and recurrence-free survival was analyzed using the logrank test on Kaplan-Meier survival curves. Cox proportional-hazard regression was performed for univariate analysis using an enter model for survival and recurrence-free survival. Multivariate analysis was performed using a backward enter model with variable removal at $p>0.10$ to test statistical significance and independence of factors shown to be significant by univariate analysis for survival or recurrence-free survival. This study complies with the Reporting Recommendation for Tumour Marker Prognostic Studies guidelines (McShane *et al.* 2006).

2.3 RESULTS

2.3.1 Gene expression analysis of DDX1 in breast cancer

Gene expression microarrays were hybridized using RNAs isolated from samples obtained from 176 primary treatment-naive breast cancer patients (45 stage I, 117 stage IIA/IIB and 14 stage IIIA/IIIB). Eighty-eight of the 176 patients experienced early relapse (recurrence within five years) and 57 patients had died when the study was locked (September 30 2009). Of the 176 tumours analysed, 31 were classified as luminal A, 45 as luminal B, 8 as HER2+ve and 56 as triple negative. Thirty-six samples were simply classified as luminal as their Ki67 status was not available (Table 2.1).

Relative *DDX1* RNA levels in the 176 tumours ranged from 0.497 – 3.437. In comparison, relative *DDX1* RNA levels in 10 normal breast tissue samples ranged from 0.804 to 1.094. ROC curve analysis in relation to recurrence defined a relative RNA level of more than 1.365 as the most appropriate cut-off point (sensitivity 23%, selectivity 91%). Of 176 patients, 28 (16%) had relative *DDX1* RNA levels of >1.365, with the remaining 148 (82%) having relative RNA levels of ≤1.365. Univariate Cox regression analysis showed a significant correlation to both recurrence and death for negative ER status, negative PR status, high grade (defined as grade 3) and elevated *DDX1* RNA levels. There was no correlation to death or recurrence for HER2 status or stage (Table 2.1).

			Univariate Analysis			
			Recurrence		Death	
Number of patients		176				
Age at diagnosis	Median	52 years				
	Range	26-89 years				
Recurrence	Events	88 (50%)				
	Average time to	818 days				
Death	Events	57 (32%)				
	Average time to	1056 days	HR	p	HR	p
ER status	Positive	112 (64%)				
	Negative	64 (36%)	0.61	**0.03**	0.35	**<0.001**
PR status	Positive	94 (53%)				
	Negative	82 (47%)	0.55	**<0.01**	0.43	**<0.01**
HER2 status	Amplified	30 (17%)				
	Non-amplified	146 (83%)	1.02	0.95	0.98	0.96
Grade	3	120 (68%)				
	1-2	56 (32%)	1.68	**0.04**	1.84	**0.05**
Stage	I	45 (26%)				
	IIA/IIB	117 (66%)				
	IIIA/IIIB	14 (8%)	1.17	0.46	1.38	0.21
Subtype	Luminal A	31 (18%)				
	Luminal B	45 (26%)				
	Luminal undefined	36 (20%)				
	HER2 Amplified	8 (5%)				
	Triple Negative	56 (32%)				
Gene expression microarray score for *DDX1*	Range	0.497 - 3.437				
	≤1.365	148 (84%)				
	>1.365	28 (16%)	2.74	**<0.0001**	2.09	**0.02**

Table 2.1 Clinicopathologic features of the patients included in the gene expression microarray analysis. p values correspond to univariate Cox regression analysis. Percentages may not equal 100% due to rounding. Abbreviation: HR, hazard ratio.

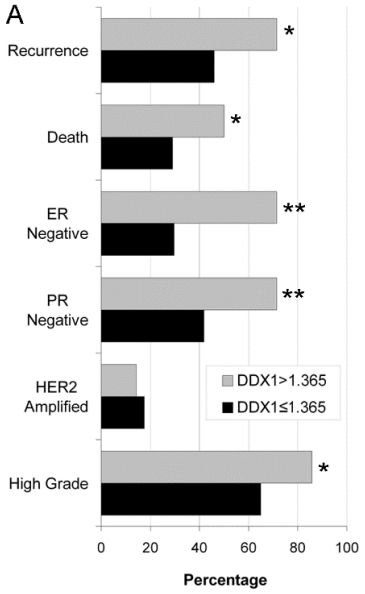
2.3.2 DDX1 RNA levels correlate with death, recurrence, ER negative status, PR negative status and high grade

Fisher's exact tests, chi square tests or Student's t tests were performed to determine if elevated *DDX1* RNA levels correlated with known prognostic indicators and clinical outcomes. High *DDX1* RNA levels were found to correlate with recurrence, death, negative ER status, negative PR status, and high grade (Figures 2.2A-B). There was no correlation between relative *DDX1* RNA levels of >1.365 and HER2 amplification, breast cancer family history, menopause status, stage and tumour size (Figure 2.2B).

Kaplan Meier survival curve analysis of high *DDX1* RNA levels showed a higher risk of recurrence with a hazard ratio of 4.31 (95% CI 2.22 – 9.19, $p<0.0001$) (Figure 2.2C) and a higher risk of death with a hazard ratio of 2.58 (95% CI 1.22 – 5.61, $p=0.014$) (Figure 2.2D). Similar data were obtained upon analysis of recurrence within systemic therapy subgroups (+ or – adjuvant chemotherapy; + or - adjuvant hormone therapy). High *DDX1* RNA levels were significantly associated with recurrence in patients who received chemotherapy, with a hazard ratio of 8.45 (95% CI 3.38 – 21.05, $p<0.0001$) and in patients who received hormone therapy, with a hazard ratio of 14.68 (95% CI 3.49 – 61.65, $p=0.0002$) (Figure 2.3) Although not significant, there was also a trend towards increased recurrence in patients who did not receive chemotherapy and in patients who did not receive hormone therapy.

Figure 2.2 Clinical/pathological features of the expression microarray patient population. (A-B) Occurrence of clinical/pathological features in patients with relative *DDX1* RNA levels of >1.365 (n=28) compared to patients with relative *DDX1* RNA levels of ≤1.365 (n=148). High *DDX1* RNA levels were found to correlate with recurrence [71% of cases with elevated *DDX1* RNA (>1.365) showed recurrence compared to 46% of cases with low *DDX1* RNA (≤1.365)], death (50% compared to 29%), negative ER status (71% compared to 29%), negative PR status (71% compared to 42%), and high grade (86% compared to 65%). There was no correlation between relative *DDX1* RNA levels of >1.365 and HER2 amplification (14% compared to 18%), breast cancer family history (39% compared to 45%), menopause status (39% pre-, 50% post-, and 11% peri-menopausal compared to 33% pre-, 59% post-, and 8% peri-menopausal), stage (29% stage I, 61% stage IIA/B and 11% stage IIIA/B compared to 25% stage I, 68% stage IIA/B, and 7% stage IIIA/B) and tumor size (average size of 2.8 cm compared to 2.6 cm) * indicates $p < 0.05$, ** indicates $p < 0.01$. (C-D) Kaplan-Meier survival curves of patients with relative *DDX1* RNA levels >1.365 (n=28) compared to patients with relative *DDX1* RNA levels ≤1.365 (n=148). (C) Recurrence-free survival. (D) Survival. Abbreviation: HR, hazard ratio.

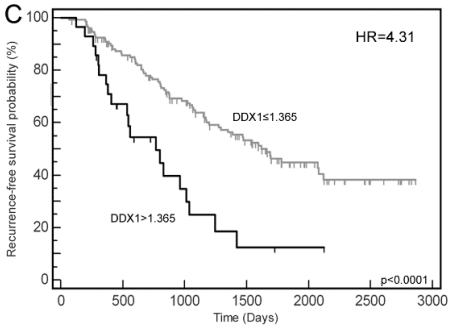
A



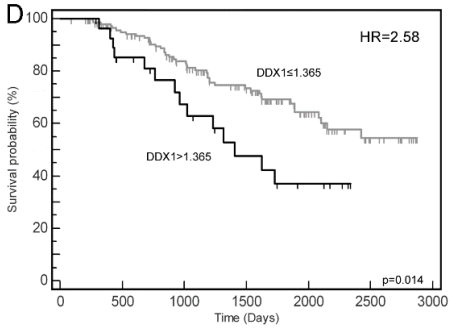
B

DDX1 RNA	>1.365	≤1.365	p
Total	28	148	
Recurrence	20	68	**p=0.011**
Death	14	43	**p=0.028**
Negative ER status	20	44	**p<0.001**
Negative PR status	20	62	**p<0.01**
High grade	24	96	**p=0.024**
HER2 amplified	4	26	p=0.45
Breast cancer family history	11	66	p=0.38
Pre-menopausal	11	49	
Peri-menopausal	3	12	p=0.98
Post-menopausal	14	87	
Stage I	8	37	
Stage IIA/B	17	100	p=0.74
Stage IIIA/B	3	11	
Average tumor size	2.78 cm	2.62 cm	p=0.62

C



D



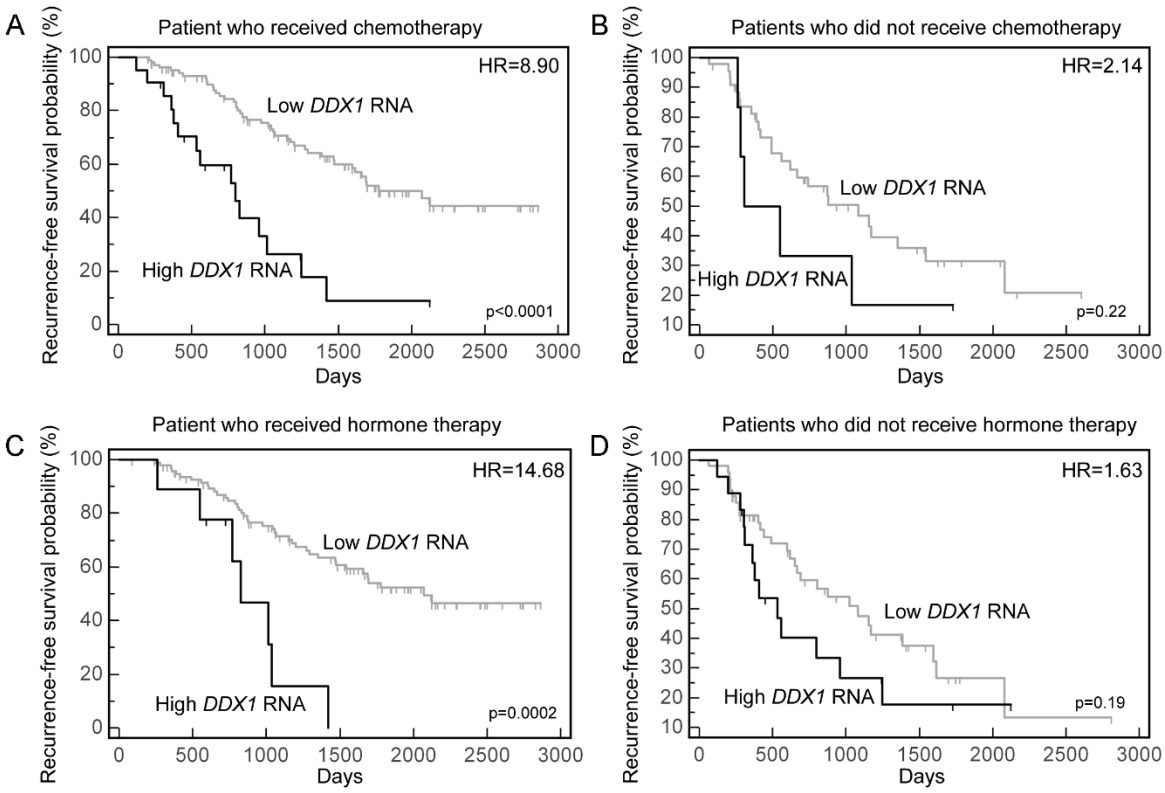


Figure 2.3 Kaplan-Meier survival curves of recurrence-free survival comparing patients with high versus low *DDX1* RNA levels sorted by systemic therapy subgroups. (A-B) Patients classified by receiving or not receiving chemotherapeutic agents. (C-D) Patients classified by receiving or not receiving hormone therapy. Abbreviation: HR, hazard ratio.

2.3.3 DDX1 protein subcellular localization in breast cancer tissue

A TMA was generated using breast cancer tissue samples from 120 (of the original 176) patients, of which seven were discarded because of insufficient tissue left on the TMA. The TMA also included cores from six normal breast tissue samples. TMAs were immunostained with anti-DDX1 antibody. Sixty-two of the 113 patients represented in the TMAs had recurred at the time of analysis, and 32 patients had died (Table 2.2).

As different levels of DDX1 protein were observed in the cytoplasm and nucleus, cytoplasmic and nuclear DDX1 protein were individually scored (Figures 2.4B-G). Nuclear staining was scored on a relative scale of 1 to 3, while cytoplasmic staining was scored on a relative scale of 0 to 3. There was a non-random distribution of nuclear to cytoplasmic staining intensity ($p<0.001$), with a general inverse relationship between cytoplasmic and nuclear levels ($rho=-0.28$, $p=0.0027$) (Figure 2.4A). Thirty-seven of 113 (33%) tumours had a cytoplasmic score of 2 or 3. In contrast, all six normal breast tissues had cytoplasmic scores of 0 or 1 and nuclear scores of 3 (Figure 2.4G).

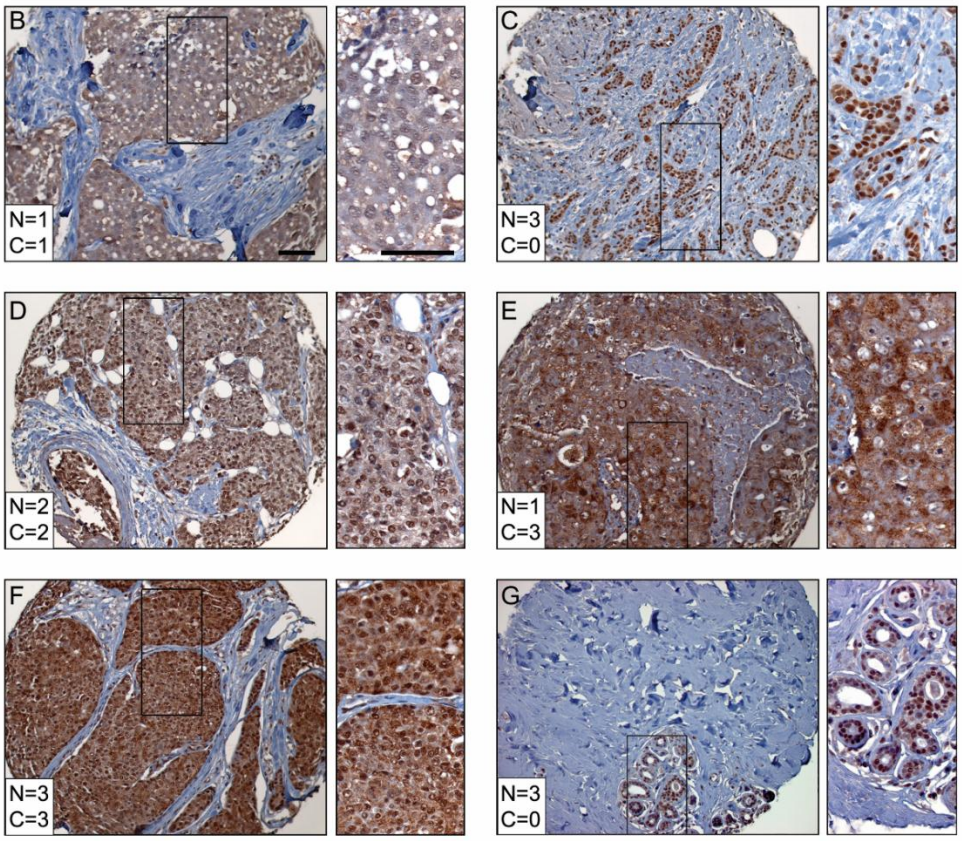
Univariate analysis of the 113 tumour samples showed a significant correlation with both recurrence and death for: (i) negative PR status and (ii) elevated levels of DDX1 in the cytoplasm. A significant correlation was also observed between grade and recurrence, but not grade and death. A negative ER status was correlated with death only, whereas decreased levels of DDX1 protein in the nucleus was correlated with recurrence only (Table 2.2).

Figure 2.4 DDX1 protein subcellular location in breast cancer TMAs. (A) The intensity of cytoplasmic DDX1 protein is plotted against the intensity of nuclear DDX1 protein. The distribution of DDX1 nuclear and cytoplasmic intensities is non-random with a p value of <0.001 based on chi square distribution analysis. (B) Breast cancer tissue from patient MT861 had a nuclear DDX1 score of 1 (N=1) and a cytoplasmic DDX1 score of 1 (C=1). The boxed area is magnified in the right panel. (C-F) The nuclear and cytoplasmic DDX1 scores for patient GT178 (C), patient MT340 (D), patient GT226 (E), and patient MT604 (F) are indicated at the bottom left of each figure. (G) Normal breast tissue had a nuclear DDX1 score of 3 and a cytoplasmic DDX1 score of 0. Scale bars = 60 μm.

A

Nuclear Intensity	Cytoplasmic Intensity			
	0	**1**	**2**	**3**
1	2	7	12	1
2	3	19	8	1
3	23	22	8	7

$p < 0.001$



Number of patients		113				
Age at diagnosis	Median	51 years				
	Range	26-89 years				
Recurrence	Events	62 (55%)				
	Average time to	755 days	Univariate Analysis			
Death	Events	32 (29%)	Recurrence		Death	
	Average time to	910 days	HR	p	HR	p
ER status	Positive	70 (62%)				
	Negative	43 (38%)	0.65	0.10	0.35	**<0.01**
PR status	Positive	58 (51%)				
	Negative	55 (49%)	0.53	**0.01**	0.42	**0.02**
HER2 status	Amplified	21 (19%)				
	Non-amplified	92 (81%)	0.90	0.77	1.00	0.99
Grade	3	78 (69%)				
	1-2	35 (31%)	1.88	**0.03**	2.36	0.06
Stage	I	29 (26%)				
	IIA/IIB	74 (65%)				
	IIIA/IIIB	10 (9%)	0.96	0.84	1.41	0.25
Subtype	Luminal A	34 (30%)				
	Luminal B	36 (32%)				
	HER2 Amplified	6 (5%)				
	Triple Negative	37 (33%)				
Cytoplasmic DDX1 intensity	0	28 (25%)				
	1	48 (42%)				
	2	28 (25%)	0-1 vs 2-3		0-1 vs 2-3	
	3	9 (8%)	1.79	**0.03**	2.51	**<0.01**
Nuclear DDX1 intensity	1	22 (19%)				
	2	31 (27%)	1 vs 2 vs 3		1 vs 2 vs 3	
	3	60 (53%)	0.72	**0.04**	0.75	0.18

Table 2.2 Clinicopathologic features of the patients included in the TMA analysis. p values correspond to univariate Cox regression analysis. Percentages may not equal 100% due to rounding. Abbreviation: HR, hazard ratio.

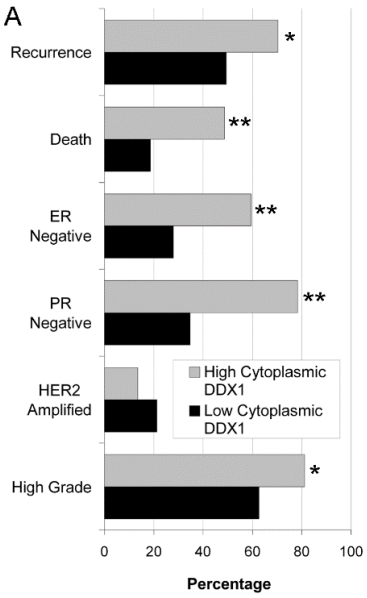
2.3.4 Subcellular DDX1 protein localization correlates with death, recurrence, ER negative status, and PR negative status

Statistical analysis was performed as described for *DDX1* RNA to determine if elevated levels of DDX1 protein in the cytoplasm (defined by a score of 2 or 3) compared to low levels in the cytoplasm (defined by a score of 0 or 1) correlated with known prognostic indicators and clinical outcomes. Elevated levels of DDX1 protein in the cytoplasm were found to correlate with recurrence, death, negative ER status, negative PR status and high grade (Figures 2.5A-B). There was no correlation between elevated levels of cytoplasmic DDX1 protein and HER2 amplification, breast cancer family history, menopause status, stage and tumour size (Figure 2.5B). Kaplan Meier survival curve analysis of cytoplasmic DDX1 localization showed a higher risk of recurrence with a hazard ratio of 1.90 (95% CI 1.09 – 3.34, p=0.0237) (Figure 2.5C) and death with a hazard ratio of 2.79 (95% CI 1.32 – 5.89, p=0.0073) (Figure 2.5D).

There was also correlation between low levels of nuclear DDX1 protein (score of 1 or 2 in the nucleus) and known prognostic markers (negative ER status, negative PR status and high grade) but not clinical outcome based on the number of events (data not shown). However, we did observe a non-significant trend between low levels of DDX1 protein in the nucleus and recurrence, but not death, based on logrank analysis of Kaplan-Meier curves (HR=1.65, 95% CI 0.99 – 2.76, p=0.055) and univariate Cox regression analysis (HR=1.62, 95% CI 0.99 – 2.68, p=0.058).

Figure 2.5 Clinical/pathological features of the tissue microarray patient population. (A-B) Occurrence of clinical/pathological features in patients with high levels of DDX1 protein in the cytoplasm (DDX1 scores of 2 or 3) (n= 37) compared to patients with low levels of DDX1 in the cytoplasm (DDX1 scores of 0 or 1) (n=76). Elevated levels of DDX1 protein in the cytoplasm were found to correlate with death (46% of cases with elevated cytoplasmic DDX1 protein levels compared to 20% of cases with low cytoplasmic DDX1 protein levels,), recurrence (68% compared to 49%), negative ER status (59% compared 28%), negative PR status (78% compared to 34%, $p<0.0001$) and high grade (81% compared to 63%). There was no correlation between elevated levels of cytoplasmic DDX1 protein and HER2 amplification (14% compared to 21%), breast cancer family history (49% compared to 38%), menopause status (32% pre-, 59% post-, and 8% peri-menopausal compared to 36% pre-, 55% post-, and 9% peri-menopausal), stage (22% stage I, 62% stage IIA/B and 16% stage IIIA/B compared to 28% stage I, 67% stage IIA/B, and 5% stage IIIA/B) and tumor size (average size of 3.3 compared to 2.7 cm) * indicates $p < 0.05$, ** indicates $p < 0.01$. (C-D) Kaplan-Meier survival curves of patients with cytoplasmic DDX1 scores of 2 and 3 (n=37) compared to patients with cytoplasmic DDX1 scores of 0 and 1 (n=76). (C) Recurrence-free survival. (D) Survival. Abbreviation: HR, hazard ratio.

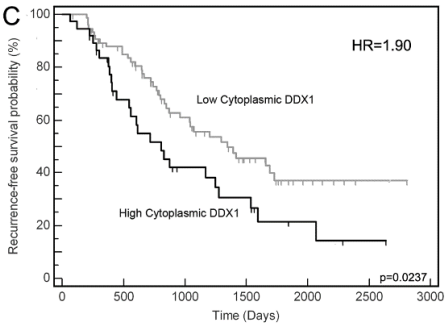
A



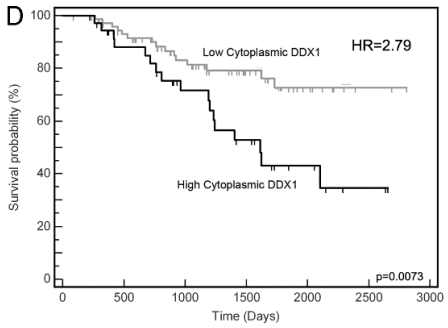
B

Cytoplasmic DDX1	2 - 3	0 - 1	p
Total	37	76	
Recurrence	25	37	**p=0.045**
Death	17	15	**p<0.01**
Negative ER status	22	21	**p<0.01**
Negative PR status	29	26	**p<0.0001**
High grade	30	48	**p=0.041**
HER2 amplified	5	16	p=0.24
Breast cancer family history	18	29	p=0.21
Pre-menopausal	12	27	
Peri-menopausal	22	42	p=0.91
Post-menopausal	3	7	
Stage I	8	21	
Stage IIA/B	23	51	p=0.15
Stage IIIA/B	6	4	
Average tumor size	3.31 cm	2.71 cm	p=0.17

C



D



2.3.5 DDX1 RNA levels and protein localization predict recurrence independently of common markers

Multivariate Cox regression analysis of factors shown to be significant in our univariate analysis (*DDX1* RNA levels, DDX1 localization, ER status, PR status and grade) was performed to determine if either *DDX1* RNA levels or protein localization, or both parameters, were independently predictive of death and recurrence (Table 2.3).

First, we carried out multivariate analysis on the four variables found to be significantly associated with survival or recurrence in the 176-patient gene expression microarray study: relative *DDX1* RNA levels, ER status, PR status and grade. We used the backward stepwise method to remove variables at each step based on a 0.1 level of significance. Only one variable was retained when survival was modeled: ER status (HR=0.35 95% CI 0.21 – 0.60, p=0.0001). Upon modelling recurrence-free survival in the same cohort, all four factors were retained, *DDX1* RNA levels (HR=2.61 95% CI 1.50 – 4.54, p=0.0007), ER status (HR=1.92 95% CI 0.89 – 4.12, p=0.10), PR status (HR=0.43 95% CI 0.23 – 0.83, p=0.01) and grade (HR=1.67 95% CI 0.96 – 2.90, p=0.07). Three of the four variables retained in our model for recurrence had hazard ratios similar to those calculated using univariate analysis (*DDX1* RNA level, PR status and grade) suggesting that they have independent prognostic value. The hazard ratio for ER status is significantly different in the univariate analysis (0.61) compared to the multivariate analysis (1.92), suggesting that ER status does not confer an independent prognostic value.

		Survival			Recurrence free survival		
	Factor	HR	95% CI	p	HR	95% CI	p
Microarray n=176	DDX1	N/S	N/S	N/S	2.61	1.50-4.54	<0.001
	ER	0.35	0.21-0.60	0.0001	1.92	0.89-4.12	0.10
	PR	N/S	N/S	N/S	0.43	0.23-0.83	0.01
	Grade	N/S	N/S	N/S	1.67	0.96-2.90	0.07
TMA n=113	DDX1	1.97	0.96-4.06	0.07	1.73	1.04-2.88	0.04
	ER	0.43	0.21-0.89	0.02	N/S	N/S	N/S
	PR	N/S	N/S	N/S	N/S	N/S	N/S
	Grade	N/S	N/S	N/S	1.82	1.02-3.27	0.04

Table 2.3 Cox multivariate analysis of survival and recurrence-free survival based on the gene expression microarray data and TMA data. Abbreviations: HR, hazard ratio; CI, confidence interval.

This is expected as our patient cohort (relapsed versus non-relapsed) was controlled for ER status.

Second, we carried out multivariate analysis with the four variables found to be significantly associated with survival or recurrence in the 113-patient TMA study: cytoplasmic DDX1 protein, ER status, PR status and grade. Upon modelling survival, only cytoplasmic DDX1 protein (HR= 1.97 95% CI 0.96 – 4.06 p=0.067) and ER status (HR=0.43 95% CI 0.21 – 0.89, p=0.024) were retained. The model for recurrence-free survival retained only cytoplasmic DDX1 protein (HR=1.73 95% CI 1.04 – 2.88, p=0.036) and grade (HR=1.82 95% CI 1.02 – 3.27, p=0.045). Cytoplasmic DDX1 protein and grade were retained with hazard ratios similar to those generated by univariate analysis suggesting that they both provide independent prognostic significance.

2.4 DISCUSSION

Breast cancer is increasingly managed on the basis of molecular classification. There is widespread consensus that ER+/PR+/low proliferation tumours are associated with a good outcome while HER2-positive and triple-negative tumours are associated with a poor outcome. Adjuvant hormonal therapy in ER+ve breast cancers, which constitute ~70% of breast cancers in developed countries, reduces the relative risk of death by approximately 22% and the risk of recurrence by 42% (2005). Nonetheless, a significant number of ER+ve tumours will relapse. At issue are the diverse nature of breast cancer and the complexity and multitude of events leading to tumour formation and progression.

Here, we use gene expression and immunohistochemical analysis to investigate DDX1 expression in 176 primary breast cancers, half of which were selected for early recurrence. We demonstrate a highly significant correlation between recurrence and increases in *DDX1* RNA levels, with a hazard ratio of 4.31. We also observe a significant correlation between recurrence and elevated DDX1 protein in the cytoplasm, and a non-significant trend between recurrence and low levels of DDX1 in the nucleus. Furthermore, analysis of systemic therapy subgroups suggests that elevated levels of *DDX1* RNA is a prognostic factor for all treatment subpopulations. Assessment of DDX1's predictive value for treatment outcome in primary breast cancer will require evaluation of DDX1 in the context of a randomized clinical trial.

Cox multivariate analysis of high *DDX1* RNA levels and DDX1 cytoplasmic localization indicate that both *DDX1* RNA levels and cytoplasmic localization are

independent markers of recurrence. In both cases, the hazard ratio remained relatively unchanged between univariate and multivariate analysis, demonstrating that additional factors (ERα, grade) did not significantly modulate the effect of *DDX1* RNA or protein localization. Cytoplasmic DDX1 localization, but not *DDX1* RNA levels, was also independently correlated with death. These results suggest that DDX1 analysis refines prognostic assessments using standard clinicopathologic parameters (stage, grade, hormone receptor and HER2 status) in a population receiving guideline-based standardized adjuvant therapy.

It is not clear to what extent increased *DDX1* RNA levels correlates with increased DDX1 protein levels in the breast cancer tissues analysed. Although there was considerable overlap between those recurrences characterized by elevated *DDX1* RNA levels and those characterized by elevated levels of DDX1 protein in the cytoplasm, it seems unlikely that the relatively small increases in *DDX1* RNA levels (>40%) detected by gene expression microarray analysis could account for the considerable increases in cytoplasmic DDX1 protein levels observed by TMA analysis. We postulate that deregulation of DDX1, be it at the expression or subcellular distribution level, is at the heart of its association with recurrence. In support of this idea, analyses of a wide variety of tissues and cell lines demonstrate that DDX1 is primarily a nuclear protein (Bleoo *et al.* 2001, Godbout *et al.* 1998), with the exception of *MYCN/DDX1*-amplified retinoblastoma and neuroblastoma tumour cells that show equal distribution of DDX1 protein in the nucleus and cytoplasm (Bleoo *et al.* 2001, Godbout *et al.* 1998). Furthermore, we have not been able to stably alter DDX1 protein levels in either cell lines or

transgenic mice (our unpublished data), and mutation of *DDX1* in fruit flies can lead to an embryonic lethal phenotype (Zinsmaier *et al.* 1994). Combined, these data suggest that levels of DDX1 and its subcellular distribution are tightly controlled and that it is only when cells become tumourigenic that this regulation is relaxed.

DDX1 is a DEAD box protein that can bind and unwind DNA/RNA and RNA/RNA duplexes *in vitro* (Li *et al.* 2008). Roles proposed for DDX1 include RNA processing (Bleoo *et al.* 2001, Sunden *et al.* 2007b), transcription regulation (Tanaka *et al.* 2009), DNA double-strand break repair (Li *et al.* 2008) and RNA transport (Fang *et al.* 2004, Kanai *et al.* 2004). While the first three roles are strictly dependent on the presence of DDX1 protein in the nucleus, RNA transport involves shuttling of molecules between the nucleus and cytoplasm, and to specific regions of the cytoplasm. Deregulation of DDX1 could result in altered subcellular localization of RNAs, that in turn could affect the availability of specific RNAs for translation. Thus, breast cancer cells with elevated levels of cytoplasmic DDX1 protein may exhibit alterations in their complement of translated proteins. As increases in cytoplasmic DDX1 protein are associated with a worse prognosis, reduction of its extraneous cytoplasmic activity represents an attractive therapeutic option. One possibility is to target DDX1 with small molecule inhibitors, as reported for two other DEAD box proteins (Bordeleau *et al.* 2005, Erkizan *et al.* 2009).

The mechanisms of action of several members of the DEAD box protein family implicated in cancer have been investigated. For example, p68 (DDX5) and p72 (DDX17) have been shown to interact with ERα and to alter ERα transcription

activity in breast cancer cells (Fuller-Pace *et al.* 2008, Wortham *et al.* 2009). Expression of p72 in ERα-positive breast cancers is associated with longer recurrence-free survival and overall survival, and is inversely correlated with HER2 expression. DDX6 (RCK/p54), is over-expressed in colorectal cancer, and may be deregulating proliferation by activating the Wnt pathway (Lin *et al.* 2008). Finally, DDX53 (CAGE), normally specific to the testis, is expressed in a variety of cancers, including lung, cervical, and colon (Cho *et al.* 2002). The wide spectrum of associations between DEAD box proteins and cancer define this family of proteins as an attractive target for future therapies.

In summary, we show that increased *DDX1* RNA levels and cytoplasmic localization of DDX1 protein both correlate with increased risk of recurrence in breast cancer, independently of commonly used markers such as ERα and grade. Future work will involve determining whether DDX1 can serve as a prognostic marker for all subtypes of breast cancer, and to assess DDX1's potential as a predictive biomarker and breast cancer therapeutic target.

2.5 ACKNOWLEDGMENTS

We are grateful to Lillian Cook and Cheryl Santos for their excellent technical assistance, Adrian Driga for his help with database analysis and Sunita Ghosh for her assistance with statistical analyses. This work was supported by grants from the Alberta Cancer Board – Research Initiative Program (R.G.), Alberta Cancer Research Institute (J.R.M.) and an Alberta Cancer Foundation studentship (D.R.G.).

CHAPTER 3

LOSS OF THE DROSOPHILA MELANOGASTER RNA BINDING PROTEIN DDX1 LEADS TO REDUCED SIZE AND ABERRANT GAMETOGENESIS

Chapter 3 has been prepared for publication as Germain DR, Li L, Hildebrandt MR, Simmonds AJ, Hughes SC, and Godbout R. Loss of the *Drosophila* melanogaster RNA binding protein Ddx1 leads to reduced size and aberrant gametogenesis. RNA co-immunoprecipitation and PCR amplification of immunoprecipitated RNAs were performed by Dr. Lei Li. Phospho-specific western blot analysis was performed by Dr. Matthew Hildebrandt. Dr. Sarah Hughes and Dr. Andrew Simmonds provided materials and technical assistance. I was responsible for all other experimental procedures, conceptualizing novel experimental protocols, data analysis and interpretation, imaging and writing the manuscript. Dr. Roseline Godbout was involved in all stages of the project and in writing the manuscript.

3.1 INTRODUCTION

DEAD box proteins are a family of RNA helicases implicated in virtually every aspect of RNA metabolism (Cordin *et al.* 2006). These proteins are characterized by 12 conserved domains, including the D-E-A-D motif for which they are named. DEAD box proteins function by modifying RNA secondary structure in an ATP-dependent manner (Linder *et al.* 2013), and play a role in RNA trafficking (Linder *et al.* 2001, Montpetit *et al.* 2012). There are >30 DEAD box genes in the *Drosophila melanogaster* genome (Activities at the Universal Protein Resource (UniProt), 2014).

Several *Drosophila* DEAD box genes are known to play a role in early development and gametogenesis. For example, the *vasa* gene encodes a multifunctional DEAD box protein that localizes to the posterior pole in oocytes and is required for completion of oogenesis (Lasko 2013). Vasa has also been implicated in chromatin condensation and generation of small non-coding RNAs (Pek *et al.* 2011b, Zhang *et al.* 2012). Belle, which is closely related to Vasa, is essential for larval development and required for both male and female fertility (Johnstone *et al.* 2005). Mutation of *pitchoune*, encoding another DEAD box protein, results in developmental arrest during the first instar larval stage. Pitchoune has been implicated in regulating cell growth and proliferation (Zaffran *et al.* 1998).

Only one mutation of *Ddx1* has been previously described in *Drosophila melanogaster* (Zinsmaier *et al.* 1994). This mutation was deemed to be recessive lethal; however, the nature of the mutation was not determined and the mutant line

is no longer available. Rafti *et al.* described the expression of Ddx1 in *Drosophila* in 1996, and reported elevated levels in early embryos, and expression throughout development (Rafti *et al.* 1996). More recently, publically available large scale studies using gene expression microarrays and RNA deep sequencing analysis have revealed widespread expression of *Ddx1* in all tissues and cell lines tested to date (St Pierre *et al.* 2014). These screens show elevated *Ddx1* levels in the nervous system, testes and ovaries with the highest levels observed at early embryonic stages, similar to the results obtained by *Rafti et al.* (Rafti *et al.* 1996).

Human DDX1, which is 68% similar to *Drosophila* Ddx1, is amplified and overexpressed in a subset of *MYCN*-amplified neuroblastoma and retinoblastoma cell lines and tumours (Godbout *et al.* 1998, Godbout *et al.* 1993a, Manohar *et al.* 1995, Squire *et al.* 1995). *DDX1* is also a prognostic marker in breast cancer (Balko *et al.* 2011, Germain *et al.* 2011), and plays a role in testicular tumourigenesis (Tanaka *et al.* 2009). Large scale screens for disease associated genes have identified DDX1 as a potential gene of interest in cervical cancer and chronic obstructive pulmonary disease (Johanneson *et al.* 2014, Smolonska *et al.* 2014). As well, DDX1 is significantly down-regulated in Down syndrome fetal brains (Kircher *et al.* 2002). Clues to DDX1's role in these diseases come from *in vitro* analysis of DDX1. For example, DDX1 was identified in a subset of RNA transport granules involved in the subcellular localization of RNA molecules in neuronal axons (Kanai *et al.* 2004). DDX1's role in RNA trafficking is not limited to endogenously coded genes as HIV replication requires DDX1 for efficient export of unspliced viral genomic RNA from the nucleus to the cytoplasm (Edgcomb *et al.*

2012, Fang *et al.* 2005, Robertson-Anderson *et al.* 2011). This effect is mediated through an interaction between DDX1, and two virally encoded proteins, Rev and Tat, that are essential for RNA export (Lin *et al.* 2014). DDX1 is also required for efficient replication of coronavirus (Wu *et al.* 2014, Xu *et al.* 2010), and has been shown to transactivate hepatitis C and JC viral genes (Sunden *et al.* 2007a, Sunden *et al.* 2007b, Tingting *et al.* 2006).

Under normal conditions, DDX1 forms granules that co-localize with (or reside adjacent to) cleavage bodies, gems and Cajal bodies, nuclear organelles associated with mRNA processing (Bleoo *et al.* 2001, Li *et al.* 2006b). When cells are treated with ionizing radiation, DDX1 is recruited to a subset of DNA double-strand breaks (Li *et al.* 2008). Biochemical analysis has shown that DDX1 can unwind RNA/RNA and RNA/DNA duplexes *in vitro* in an ADP-dependent manner and can efficiently digest single-stranded RNA (Li *et al.* 2008). Finally, DDX1 has also been implicated as a tRNA splicing factor (Popow *et al.* 2011, Popow *et al.* 2014).

While it is clear that DDX1 is a multifunctional protein, we only have a limited understanding of its biological role in the cell and during development. To gain insight into DDX1's role during development, we generated a *Ddx1* mutant *Drosophila* line. Here, we report that *Ddx1*-null flies are viable, with reduced fertility and body size. *Ddx1*-null flies also display aberrant gametogenesis in both testes and ovaries. We also describe a direct interaction between *Sirup* mRNA, previously described as up-regulated during starvation conditions and implicated in mitochondrial function (Erdi *et al.* 2012, Van Vranken *et al.* 2014), and Ddx1

protein. Finally, we found alternative splicing patterns in *Sirup* mRNA contingent on the presence or absence of Ddx1 *in vivo*, and an epistatic lethal effect in *Ddx1*-null/*Sirup* knock-down flies.

3.2 MATERIALS AND METHODS

3.2.1 Drosophila stocks and husbandry

All crosses were performed at 25˚C on Bloomington recipe media. Fly stocks were obtained from the Vienna *Drosophila* RNAi Center (VDRC) and the BDSC. The following fly stocks were used:

w^{1118} - Control line.

y^1 w^*; *Ly/TM3, Sb* – Balancer for potential mutant allele.

y^1 w^{67c23}; *P{EPgy2}EY12792* – P-element upstream of *Ddx1.*

w;* Dr^1*/TMS, P{ry[*∆*2-3]}99B* – Expresses P-element transposase.

w;* Sb^1*/TM3, P{ActGFP}JMR2,* Ser^1- Mutant allele balancer.

w^{1118}*; Df(3L)ED230, P{3'.RS5+3.3'}ED230/TM6C,* cu^1 Sb^1- Deficiency encompassing *Ddx1*. Referred to as *Df(3L)ED230.*

y^1 w^**; P{Act5C-GAL4}25FO1/CyO,* y^+ - Expresses GAL4 under the control of the actin promoter.

w^{1118}*; P{GD14644}v36437 – Sirup* RNAi transgene.

3.2.2 Generation of potential Ddx1 mutant alleles by P-element excision

Potential mutations of the *Ddx1* locus were generated using BDSC stock #21389 (y^1 w^{67c23}*; P{EPgy2}EY12792*), that contains a P-element inserted 43 bp upstream of the *Ddx1* transcriptional start site*. P{EPgy2}EY12792* was excised by crossing virgin *P{EPgy2}EY12792* females with ∆2-3 transposase expressing males (*w*;* Dr^1*/TMS, P{ry[*∆*2-3}99B*, BDSC #1610). F1 single female virgin ∆*P{EPgy2}EY12792/ TMS, P{ry[*∆*2-3}99B* flies were mated to *Ly/TM3, Sb* males.

F2 Δ*P{EPgy2}EY12792/TM3, Sb* females and males were crossed to establish balanced lines.

3.2.3 Characterization of potential mutant lines

Lines displaying white eye color, indicating removal of *P{EPgy2}EY12792,* were analyzed for deletion by sequential PCR reactions using staggered primers. One line (AX) was identified as containing an ~2 kb deletion. Subsequent genomic sequencing revealed that a 1733 bp region spanning the P-element insertion site and most of the *Ddx1* gene locus had been removed and replaced with the 15 bp sequence 5'-CATGATGAAATAACA-3'. This 15 bp sequence does not correspond to any part of *Ddx1* or *P{EPgy2}EY12792.* This allele was designated $Ddx1^{AX}$. The following primers were used for PCR amplification: 5'-CCAGAAGCCGTGCATG-3' (forward primer ~400 bp upstream of *Ddx1* transcription start site*)*, 5'-ATGAGTGTTGGCCAGCG-3' (forward primer ~500 bp downstream of *Ddx1* transcription start site), 5'-AGCTGGTGGAATTGCAC-3' (reverse primer ~400 bp downstream of *Ddx1* transcription start site), 5'-ACCATCTGCAGACGG-3' (reverse primer ~1.4 kb downstream of *Ddx1* transcription start site), 5'-GAGCTCCGACTTCCTAC-3' (reverse primer located in *Ddx1* 3' UTR). The following primer was used for genomic DNA sequencing: 5'-CTCATAAAGTCAAGTAAC-3' (forward primer ~200 bp upstream of *Ddx1* transcription start site).

3.2.4 Western blot analysis and antibodies

Cell lysates were prepared from adult flies or larvae by grinding with a pestle in lysis buffer (1% sodium deoxycholate, 1% Triton-X-100, 0.2% SDS, 150 mM NaCl, 50 mM Tris-HCl pH 7.4 and 1X Complete® [Roche] protease inhibitors). Samples that were analyzed for phosphor-S6k were prepared in lysis buffer supplemented with 1x PhosStop (Roche). Lysates were electrophoresed in 10% SDS-polyacrylamide gels and transferred to nitrocellulose membranes. The following antibodies were used: rabbit anti-pS6k (Cell Signaling, #9209), rabbit anti-Ddx1 (Genscript custom polyclonal antibody, antigen CQKNLRTGSGYEDHV) and mouse anti-β-Tubulin (DSHB, E7). The E7 β-Tubulin antibody developed by Michael Klymkowsky was obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa, Department of Biology, Iowa City, IA 52242. Protein detection was using the Immobilon (Millipore) reagent.

3.2.5 RNA purification and RT-PCR

RNA was purified using an RNeasy Plus Universal Mini Kit (Qiagen) as per the manufacturer's directions. Briefly, whole flies or larvae were crushed in QIAzol reagent using plastic pestles and purified RNA isolated using a mini column. Reverse transcription was carried out using SuperScript II (Life Technologies) reverse transcriptase and oligo dT primers as per the manufacturer's directions. The following primers were used for RT-PCR analysis: *Sirup* , forward primer 5'-CCTGCGAGATTGCAATTCAG-3', reverse primer 5'-AGTGGTTCCTTCTCC

TGGTACG-3', *Sirup* splice specific transcript , forward primer 5'-CAAATGGGCAAACAA*GTGA-3' (asterisk indicates splice junction site), reverse primer 5'-GAATTCTTTAATAGTTTCTGCCC-3', *Actin* 5'-AATCCAGAGACAC CAAACCG-3', reverse primer 5'-GAACGATACCG GTGGTACGA-3'. The forward primer for *Sirup* splice-specific product amplification consisted of the 15 nt sequence upstream of the splice junction followed by the 4 nt sequence downstream of the splice junction.

3.2.6 Viability, fertility, size and developmental delay assays and larval collections

For viability and fertility assays, single virgin females were mated with two males and left to lay eggs in standard culture vials. After ten days, the parental flies were removed and individual pupae were counted and moved to a new tube daily to ensure no cross generational contamination. For crowded conditions, twenty females were mated to twenty males in a single vial. For size analysis, newly eclosed adults or pupae were genotyped and photographed using an Olympus SZX12 fluorescence dissecting microscope. Whole length, for pupae, or thorax length, for adults were measured using Photoshop. For developmental delay assays, three *Ddx1^{AX}/TM3, GFP, Ser* virgins were crossed with two *Ddx1^{AX}/TM3, GFP, Ser* males and allowed to lay eggs for 24 hours. Individual pupae were removed as above and scored for *Ddx1* genotype based on GFP status. For larval collection, parental flies were placed in collection cages on apple juice agar plates with yeast paste for two hours. Plates were then incubated at

25˚C for 24, 48 or 72 hours at which point larvae were manually collected and scored for GFP status using an Olympus SZX12 fluorescent dissecting microscope. Student's t-test was used to compare differences between genotypes.

3.2.7 Immunofluorescence, microscopy and measurement

Ovaries and testes were dissected from virgin male and female flies that had been collected and held in isolation for 3 days. Testes were fixed in 4% paraformaldehyde for 20 minutes and washed three times in PBS. Tissues were then incubated with RFP conjugated-phalloidin (1/500) for 90 minutes, washed in PBS, and mounted in PVA with DAPI. Ovaries were incubated in LysoTracker Red (1/50, Life Technologies) for 10 minutes. Samples were then fixed in 4% paraformaldehyde for 20 minutes. Ovaries were then dissected into individual ovarioles and mounted in PVA with DAPI.

Confocal images were captured on a Zeiss LSM 710 confocal laser scanning microscope with a plan-Apochromat 63x (NA 1.4) oil immersion lens, a plan-Apochromat 40x (NA 1.3) oil immersion lens, or a plan-Apochromat 10x (NA 0.45) lens, and Zen software. An Olympus SZX12 microscope was used to photograph adult flies and gonads. Exported images were saved as TIFF files and measurements made using Photoshop.

3.2.8 Northern blot analysis

RNA samples were isolated from 3rd instar larvae as described above. For each sample 5 µg of total RNA was resolved in a 10% denaturing urea

polyacrylamide gel. A small RNA ladder (NEB) was used for size determination. RNA was then transferred to Hybond-N+ membranes (GE Healthcare) and baked for 1 hour at 80ºC. DNA probes were generated by ^{32}P end-labeling of single-stranded oligonucleotides obtained from IDT. The following probes were used: $tRNA^{tyr}$ 5'-CTACAGTCCACCGCTCTACCAACTGAGCTATCGAAGG-3', $tRNA^{ala}$ 5'-TGCTAAGCGAGCGCTCTACCATCTGAGCTACATCCCC-3', 5s rRNA 5'-CACTCGGCTCATGGGTCGATGAAGAACGCAGCAAACTG-3'. Probe hybridization was carried out in 5X SSC, 50 mM Na_2HPO_4 pH 6.5, 0.5X Denhardt's solution and 0.25 mg/mL salmon sperm DNA at 42ºC. Blots were then washed in 2X SSC, 0.1% SDS, followed by 0.1X SSC, 0.1% SDS. The signal was visualized using X-ray film.

3.2.9 RNA deep sequencing

RNA was isolated from three independent preparations of wandering 3rd instar larvae for both *Ddx1*-null and control genotypes. RNA libraries were prepared by first removing rRNAs using a Ribo-Zero™ rRNA Removal Kit (Human/Mouse/Rat) (Epicentre) followed by TruSeq Stranded Total RNA Sample Prep Kit (Illumina). Paired-end 100 nt RNA sequencing was performed on an Illumina HiSeq 2000 platform and processed using CASAVA v1.8.2 by the UBC Biodiversity Research Centre NextGen Sequencing Facility. Alignment, splice site identification was carried out using Bowtie v2.1.0 and Tophat v2.0.13 with standard parameters. Differential gene expression was calculated using Cufflinks v2.1.1 using standard parameters with a p value of < 0.005 considered significant. Splice

junctions analysis was limited to junctions identified a minimum of 100 times averaged across either Ddx1-null or control samples. Student's t-test was used to determine significance for junctions present in both *Ddx1*-null and control samples.

3.2.10 RNA co-immunoprecipitations

Whole cell lysates were prepared from S2 cells by resuspending cell pellets in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% sodium deoxycholate, 1% NP-40, and 1X Complete (Roche) protease inhibitors. Four hundred micrograms of lysate was first cleared with Protein A agarose beads (GE Healthcare), followed by incubation with 5 μl of rabbit anti-Ddx1 antibody or rabbit IgG for 2 hours at 4°C. Protein A agarose beads were then added and incubation continued for 1 hour at 4°C. Co-immunoprecipitates were washed three times in lysis buffer and extracted with water saturated phenol. An aliquot taken from of the co-immunoprecipitates was saved for western blot analysis to check Ddx1 immunoprecipitation efficiency. Co-immunoprecipitated RNA was dissolved in RNase-free water and stored at -80°C until use. Reverse transcription was carried out as above, using *Sirup*-specific (5'-AGTGGTTCCTTCTCCTGGTACG-3') or *Ddx1*-specific (5'-TCATCG GGCAGCGTCAC-3') reverse primer with the immunoprecipitated RNA serving as template. Following reverse transcription, PCR amplification was carried out using the *Sirup* primers described previously and with the *Ddx1* RT reverse primer above and 5'-GCATGCATTTGAGGTGAAG-3' for *Ddx1*.

3.2.11 Sirup knock-down in Ddx1 modified flies

y^1 *w^*; P{Act5C-GAL4}25FO1/CyO; $Ddx1^{AX}$/TM3, Sb* virgin females were crossed to *w^{1118}; P{GD14644}v36437; $Ddx1^{AX}$/TM3, P{ActGFP}JMR2, Ser^1* males and allowed to lay eggs for 3 days. Adults were scored for *Ddx1* status by the presence of *Sb* or *Ser* alleles, and *Sirup* knock-down was scored based on the absence of *CyO*. Chi square analysis was performed to determine the significance of observed outcomes compared with expected genotype distribution.

3.2.12 Duplex specific nuclease treatment of single stranded Sirup DNA annealed to total RNA to identify possible anti-sense RNA molecules

Full length *Sirup* cDNA was generated by PCR using primers 4F (5'-CCTCGCAGATTGCAATTCAG-3') and 795R (5'-GAATTCTTTAATAGTTTCTG CCC-3') primers. Asymmetrical PCR was then performed using only primer 4F to generate a solution consisting of primarily a single stranded sense DNA copy of *Sirup*. Serial dilutions followed by PCR using the above primers was used to determine that a $1x10^8$ dilution represented the minimal amount of product which could generate a robust visible signal following 40 cycles of PCR amplification. 2.5 µg of total RNA, isolated as previously described, was mixed with 1 µl of $2x10^7$ dilution of single stranded *Sirup* DNA in annealing buffer (20 Mm Tris-HCl, pH 7.5, 200mM potassium acetate, 0.1 mM EDTA), boiled and allowed to cool to room temperature, precipitated and resuspended in water. The annealed mix was split into two samples, one of which was treated with duplex specific nuclease, which degrades DNA from DNA/DNA and RNA/DNA duplexes, as per the manufacturer's

instructions (Evrogen). Treated and untreated samples were precipitated and resuspended in water. PCR was performed using five primer sets that amplify sub-fragments of *Sirup*: primer 4F (see above) and 201R (5'-TTTGTCTGGTCACGGATTG-3'), 91F (5'-GTGTAACTGCGACTAAGAAGCG-3') and 332R (5'-CGTCGATAGATAGCTCACTG-3'), 325F (5'-GTGAGCTATCTATCGACGAG-3') and 481R (5'-AGTGGTTCCTTCTCCTGGTACG-3'), 454F (5'-GCAAGCTGGATGAATTCTC-3') and 618R (5'-ACTAGAAATCGGAGACGC-3'), 601F (5'-GCGTCTCCGATTTCTAGT-3') and 795R (as above).

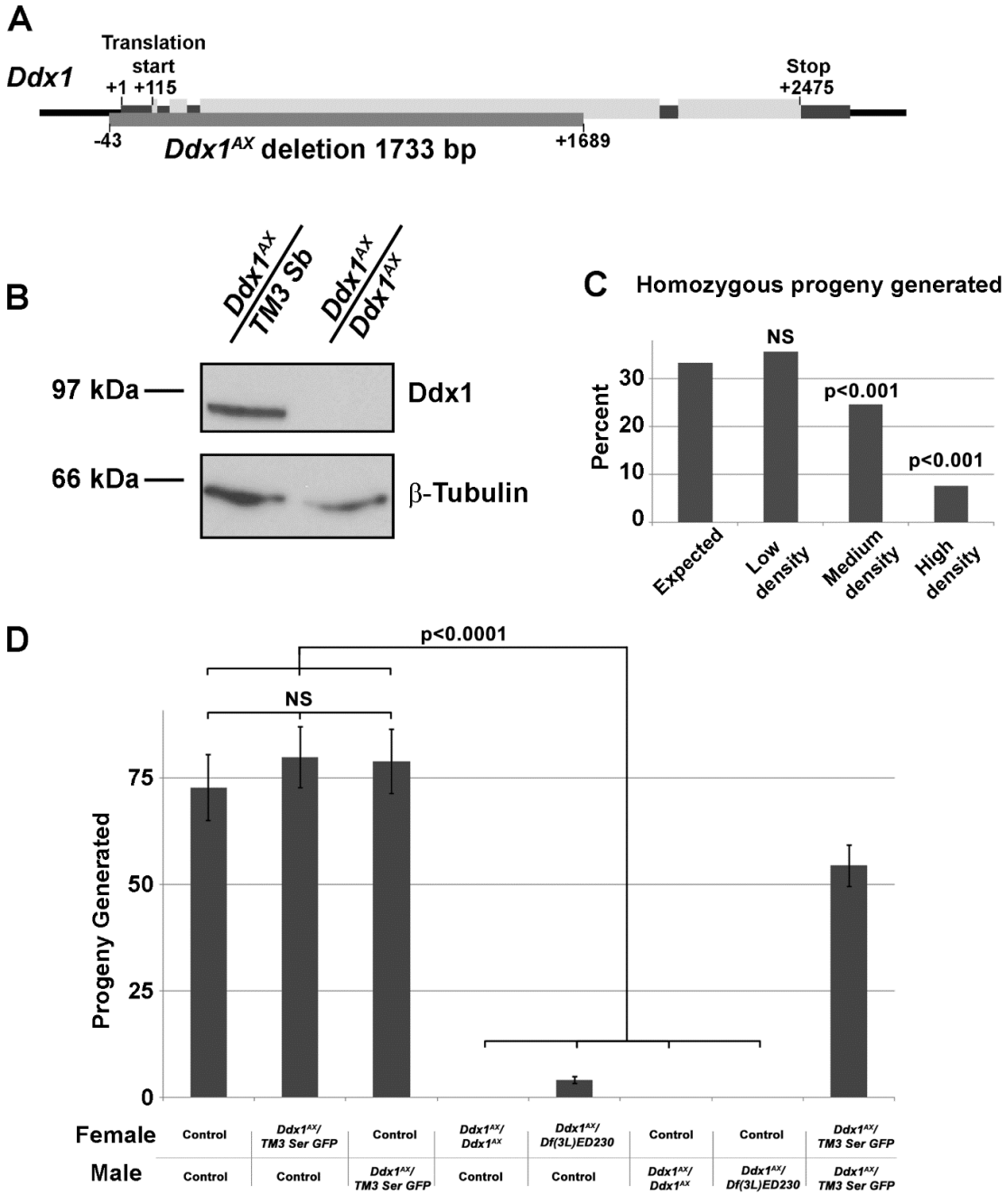
3.3 RESULTS

3.3.1 Ddx1-null flies are viable

Imprecise excision of a P-element located immediately upstream of *Ddx1* (*y^1 w^{67c23}; P{EPgy2}EY12792*) was used to screen for novel *Ddx1* mutations. From this we isolated a *Ddx1* allele with a 1733 bp deletion encompassing the majority of the *Ddx1* open reading frame. This allele was designated *$Ddx1^{AX}$* (Figure 3.1A). In contrast to a previously described *Ddx1* mutation which showed early embryonic lethality (Zinsmaier *et al.* 1994), *$Ddx1^{AX/AX}$* flies reached adulthood. Western blot analysis of control (w^{1118}) and *$Ddx1^{AX/AX}$* adult flies showed a complete absence of Ddx1 in mutant flies (Figure 3.1B). Although a small portion of the 3' *Ddx1* open reading frame is retained in *$Ddx1^{AX/AX}$* flies, no novel bands were detected using an anti-Ddx1 antibody generated against the C-terminus of Ddx1. These results suggest that there is either no translation of the retained 3'-end of *Ddx1* or the resulting protein product is unstable.

Although *$Ddx1^{AX/AX}$* flies were viable, we noticed that they were consistently outcompeted under high density culturing conditions. When single eggs from heterozygote-heterozygote crosses were raised in isolation, we observed the expected 2:1 ratio of heterozygous to homozygous mutant progeny (heterozygous mutants were maintained over a recessive lethal balancer chromosome [*TM3*, *P{ActGFP}JMR2, Ser^1*], so no homozygous non-mutant *Ddx1* progeny were generated). When a single female was allowed to lay eggs for 10 days in a standard collection vial, the ratio dropped to approximately 3:1. When 20 females

Figure 3.1 *Ddx1*-null flies are viable but show reduced fertility. (A) The location of the *Ddx1*^{AX} deletion. (B) Western blot analysis shows no detectable signal for Ddx1 protein in *Ddx1*$^{AX/AX}$ adult flies. (C) Survival of homozygous adult flies generated from heterozygote-heterozygote crosses were counted. As the *Ddx1*^{AX} mutation is carried over a recessive lethal balancer chromosome, the expected rate of homozygous progeny generated is 33%. At low density, homozygous flies were generated at the expected rate. At medium density and high density, a significant reduction in the number of homozygous flies was observed; n = 45 adults (low density), 1165 adults (medium density) and 499 adults (high density). (D) Progeny generated from single virgin females mated with two males (genotypes are indicated) and allowed to lay eggs for 10 days. Pupae were removed and counted daily. Homozygous mutant flies generated very low or no progeny. Heterozygous flies generated progeny at the expected rate. n ≥ 20 crosses for all samples.


A
Ddx1
Translation start
+1 +115
Stop
+2475
-43
+1689
Ddx1AX deletion 1733 bp
B
Ddx1AX/TM3 Sb
Ddx1AX/Ddx1AX
97 kDa
Ddx1
66 kDa
β-Tubulin
C
Homozygous progeny generated
Percent
NS
p<0.001
p<0.001
Expected
Low density
Medium density
High density
D
p<0.0001
NS
Progeny Generated
Female
Male
Control
Ddx1AX/TM3 Ser GFP
Ddx1AX/Ddx1AX
Ddx1AX/Df(3L)ED230

were allowed to lay eggs in a standard collection tube for 10 days, the ratio of heterozygous to homozygous mutant was 13:1 (Figure 3.1C).

3.3.2 Ddx1-null flies have reduced fertility, size and delayed development

Heterozygous *Ddx1^{AX}/TM3 Ser* GFP flies generated the expected number of progeny compared to control flies (Figure 3.1D). Both male and female *Ddx1$^{AX/AX}$* flies were sterile. To confirm that the observed infertility was due to inactivation of *Ddx1*, and not a line-specific effect, we crossed *Ddx1^{AX}* to a line containing a deficiency that encompasses the *Ddx1* locus (*Df[3L]ED230*). *Ddx1^{AX}/Df(3L)ED230* males were completely sterile, while females were able to generate a small number of viable progeny that survive to adulthood (Figure 3.1D). Heterozygote-heterozygote crosses resulted in approximately 75% as many progeny as control. This was expected, as 25% would be homozygous for the recessive lethal balancer chromosome.

We noted that *Ddx1*-null progeny generally eclosed later than control and heterozygous animals. To better define the extent of the developmental delay in *Ddx1*-null flies, we set up heterozygote-heterozygote crosses that were allowed to lay eggs for a period of 24 hours. Pupae were removed from each collection tube at 24 hour intervals and transferred to a secondary collection vial. Secondary collections were checked for eclosed adults at 24 hour intervals. Time to pupation revealed a non-significant trend ($p=0.069$), with heterozygous and homozygous mutant progeny having mean pupation times of 6.07 and 6.38 days, respectively (Figure 3.2A, left). A significant delay was observed for time to eclosion ($p<0.001$),

with heterozygous and homozygous mutant progeny having mean eclosion times of 10.78 and 11.70 days, respectively (Figure 3.2A, right).

In order to determine if maternally loaded Ddx1 protein may be compensating for the mutant allele during early development, we performed western blot analysis of larvae at 24, 48 and 72 hours post egg laying. Significant levels of Ddx1 protein were observed in *Ddx1*$^{AX/AX}$ larvae at 24 hours, with a weak signal detected as late as 48 hours post egg laying (Figure 3.2B). These results suggest a significant amount of maternally deposited protein with a long half-life. This may explain why *Ddx1*$^{AX/AX}$ flies survive early development and only display phenotypes at later developmental stages.

Comparison of control, heterozygous and mutant pupal lengths revealed a significant reduction in size in both *Ddx1*$^{AX/AX}$ and *Ddx1*^{AX}*/Df(3L)ED230* flies (Figure 3.3A). Of note, though both null strains were significantly smaller than control or heterozygous animals, the *Ddx1*^{AX}*/Df(3L)ED230* pupae were slightly, though significantly ($p<0.001$), larger than *Ddx1*$^{AX/AX}$ pupae. Similar results were observed upon measuring adult length (Figure 3.3B). Adult *Ddx1*-null flies were smaller than control or heterozygotes, and *Ddx1*^{AX}*/Df(3L)ED230* adults were slightly, though significantly ($p<0.05$), larger than *Ddx1*$^{AX/AX}$ adults.

3.3.3 Gametogenesis is disrupted in Ddx1-null flies

To investigate the cause of reduced fertility in *Ddx1*-null flies, we dissected ovaries and testes from adult flies which had been held in isolation for 3 or 10 days following eclosion. At 3 days post-eclosion, heterozygous ovaries appeared

Figure 3.2 *Ddx1*-null flies show delayed development and maternally contributed Ddx1 protein in larvae. (A) Heterozygote-heterozygote crosses were allowed to lay eggs for 24 hours. Pupae were removed at 24 hour intervals and scored for *Ddx1* genotype (left, n=233 for heterozygous pupae and n=104 for homozygous mutant pupae). Adults were also counted at 24 hour intervals (right, n=162 for heterozygous adults and n=44 for homozygous mutant adults). A non-significant trend was observed for pupation time, and a significant difference was observed for eclosion time. (B) Control and $Ddx1^{AX}$ heterozygote-heterozygote crosses were allowed to lay eggs on apple juice plates for a period of two hours. Protein lysates were prepared from larvae collected at 24, 48 or 72 hours post-egg laying. Western blot analysis was carried out using anti-Ddx1 (Genscript) and anti-β-tubulin antibodies (E7, DSHB). A faint Ddx1 signal was observed at both 24 and 48 hours post-egg laying, indicating that maternally loaded Ddx1 is still present at these times. Ddx1 was no longer detected at 72 hours.

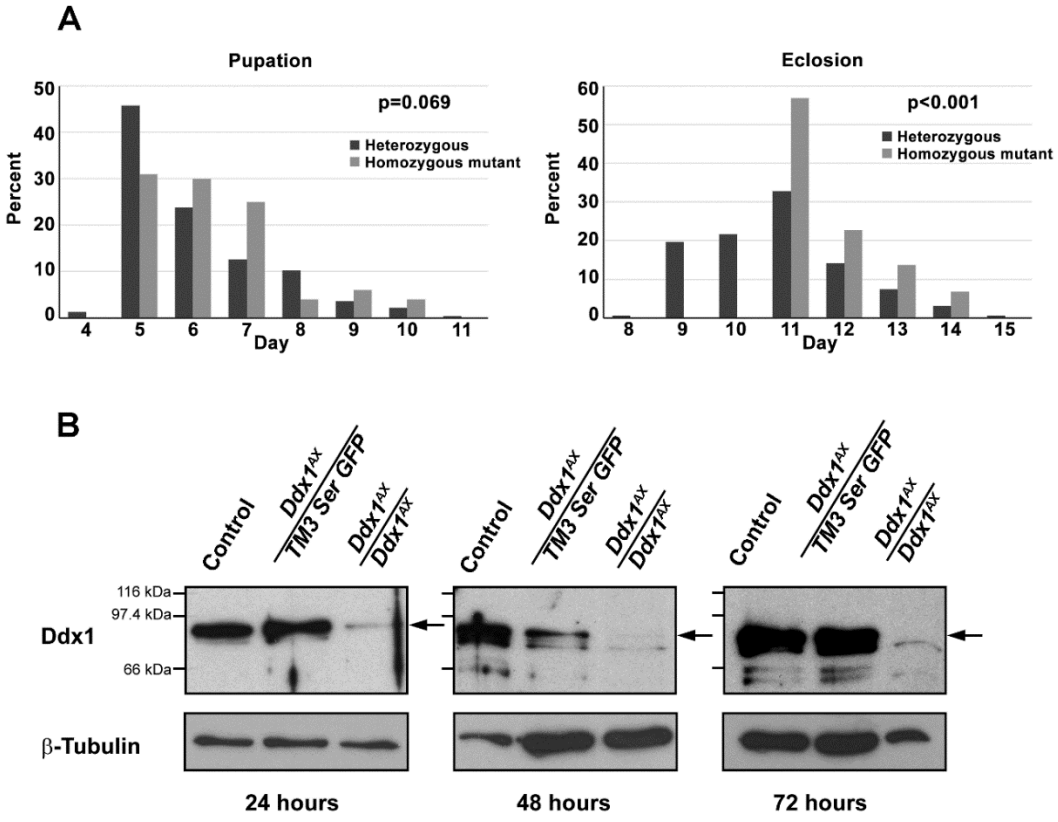
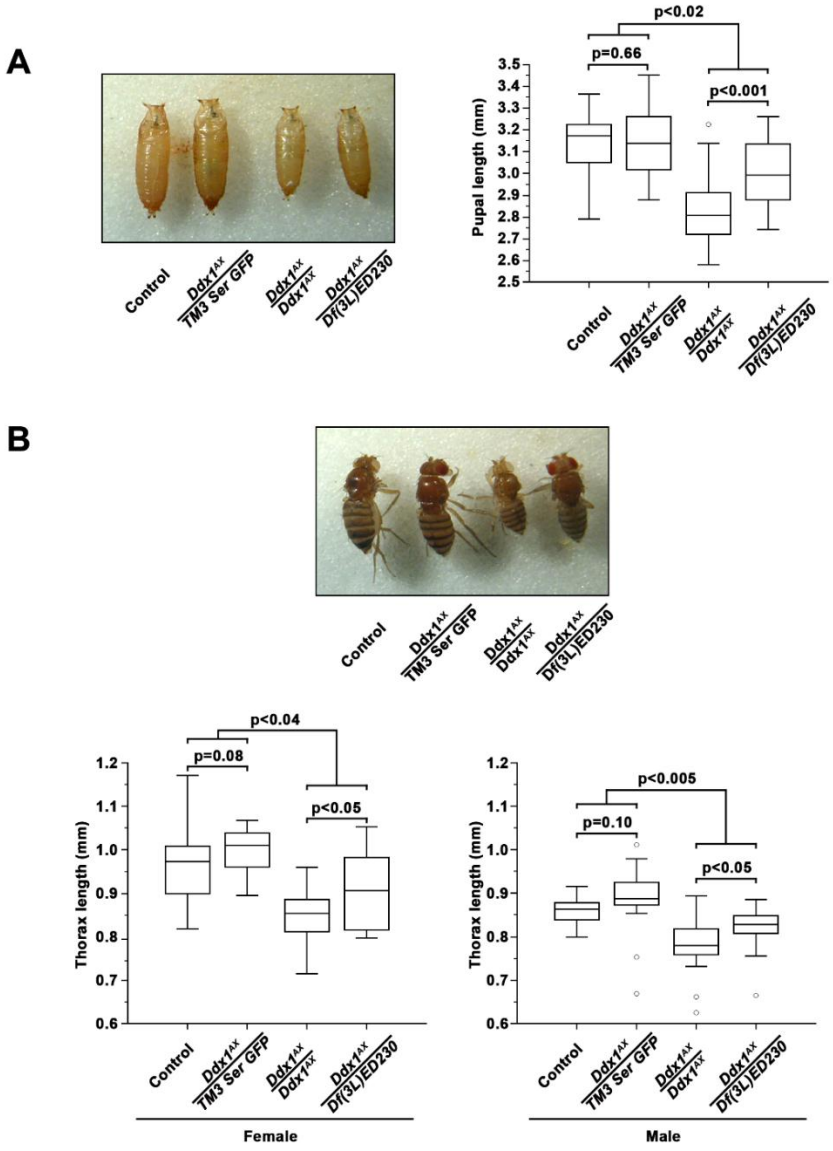

A
Pupation
p=0.069
Heterozygous
Homozygous mutant
Percent
Day
Eclosion
p<0.001
Heterozygous
Homozygous mutant
Percent
Day
B
Control
Ddx1AX/TM3 Ser GFP
Ddx1AX/Ddx1AX
116 kDa
97.4 kDa
66 kDa
Ddx1
β-Tubulin
24 hours
48 hours
72 hours

Figure 3.3 *Ddx1*-null flies are smaller than control. (A) Control, *Ddx1* heterozygous and *Ddx1* homozygous mutant pupae were collected and total pupal length was measured, n ≥ 20 pupae for each sample. (B) Control, *Ddx1* heterozygous and Ddx1 homozygous mutant one day old adults were collected and thorax length was measured, n ≥ 19 adults for each sample. At both pupal and adult stages, *Ddx1*-null animals were significantly smaller than control animals, and $Ddx1^{AX}/Df(3L)ED230$ animals were slightly larger than $Ddx1^{AX/AX}$.


A
Control
Ddx1Δx TM3 Ser GFP
Ddx1Δx Ddx1Δx
Ddx1Δx Df(3L)ED230
p<0.02
p=0.66
p<0.001
Pupal length (mm)
3.5
3.4
3.3
3.2
3.1
3.0
2.9
2.8
2.7
2.6
2.5
B
Control
Ddx1Δx TM3 Ser GFP
Ddx1Δx Ddx1Δx
Ddx1Δx Df(3L)ED230
p<0.04
p=0.08
p<0.05
Thorax length (mm)
1.2
1.1
1.0
0.9
0.8
0.7
0.6
Female
p<0.005
p=0.10
p<0.05
Male

essentially identical to control, containing many mature eggs. Ovaries from both *Ddx1*$^{AX/AX}$ and *Ddx1*^{AX}*/Df(3L)ED230* flies were much smaller, with *Ddx1*^{AX}*/Df(3L)ED230* ovaries occasionally containing a small number of mature eggs and *Ddx1*$^{AX/AX}$ ovaries containing no eggs (Figure 3.4A). Similar results were observed with ovaries isolated from adults held in isolation for 10 days post eclosion, with *Ddx1*^{AX}*/Df(3L)ED230* ovaries containing few mature eggs and some *Ddx1*$^{AX/AX}$ ovaries containing a small number of abnormal eggs that were approximately the size of a mature egg, but lacked dorsal appendages (Figure 3.4B). These observations mirror the relative fertility of *Ddx1*$^{AX/AX}$ and *Ddx1*^{AX}*/Df(3L)ED230* females as compared to heterozygous and control females.

The reduced size of *Ddx1*-null flies is possibly related to metabolism dysfunction, and if this is the case we would expect to observe other phenotypes associated with reduced metabolism. When under metabolic stress or starvation conditions, developing egg chambers undergo autophagy in order to conserve energy (Barth *et al.* 2011, McCall 2004). To determine if developing egg chambers in *Ddx1*-null flies were undergoing autophagy, we stained dissected ovarioles with LysoTracker Red, a dye that is commonly used to identify autophagic cells (DeVorkin *et al.* 2014). In *Ddx1*-null ovaries, at both 3 and 10 days post eclosion, we observed autophagic egg chambers corresponding to approximately stage 7/8 (Figure 3.4, arrows). Autophagic egg chambers were not observed in control and heterozygous ovaries.

Immunofluorescence imaging of testes isolated from adult males revealed aberrant sperm development in *Ddx1*-null males (Figure 3.5A). Early

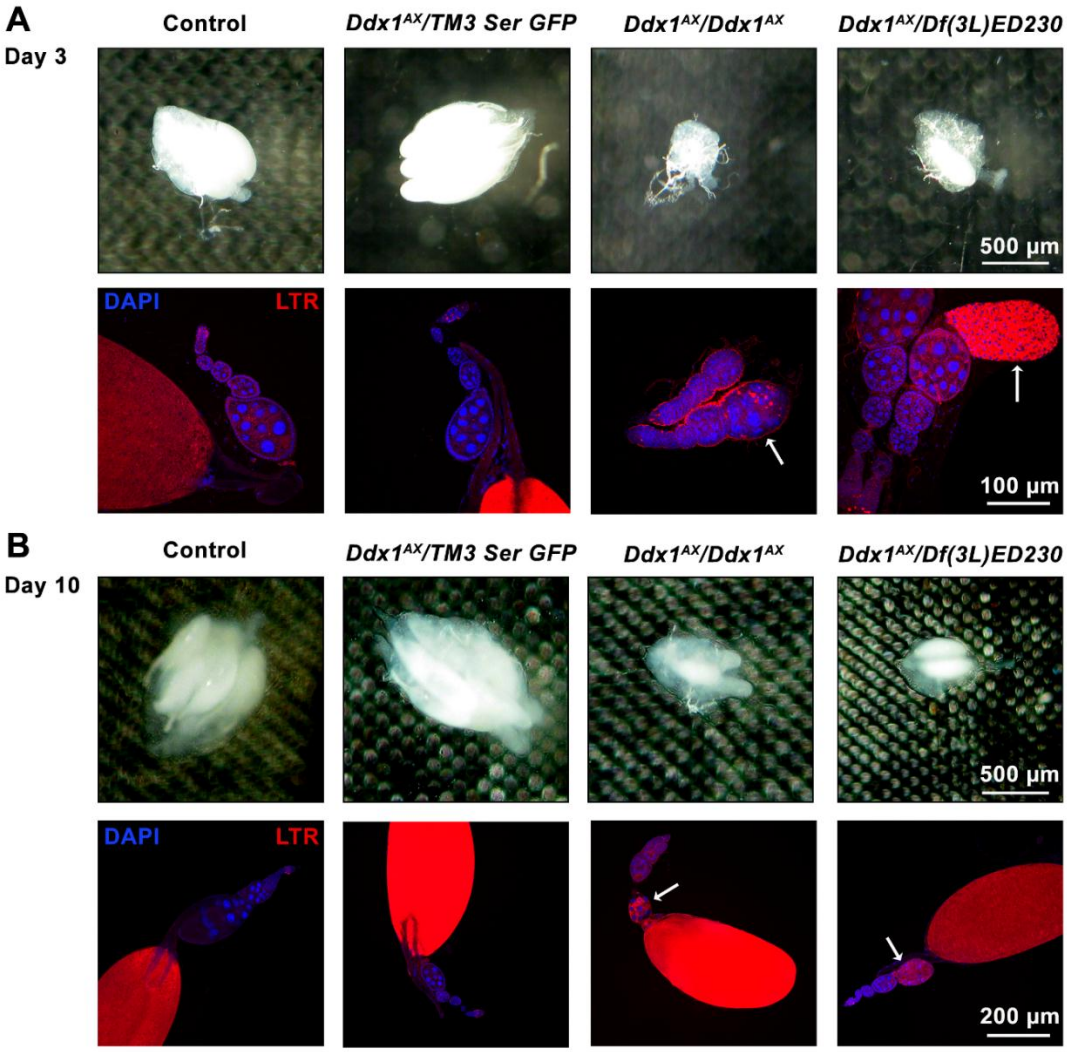


Figure 3.4 Aberrant egg development in *Ddx1*-null fly ovaries. Ovaries collected from virgin females held in isolation for 3 days (A) or 10 days (B). Top – *Ddx1*-null flies have much smaller ovaries, with few (*Ddx1*$^{AX/}$*Df(3L)ED230*) or no (*Ddx1*$^{AX/AX}$) mature eggs present. Bottom – immunofluorescence imaging of LysoTracker Red-treated ovaries reveals developing egg chambers undergoing autophagy (white arrows) in *Ddx1*-null ovaries. Nuclei are stained with DAPI.

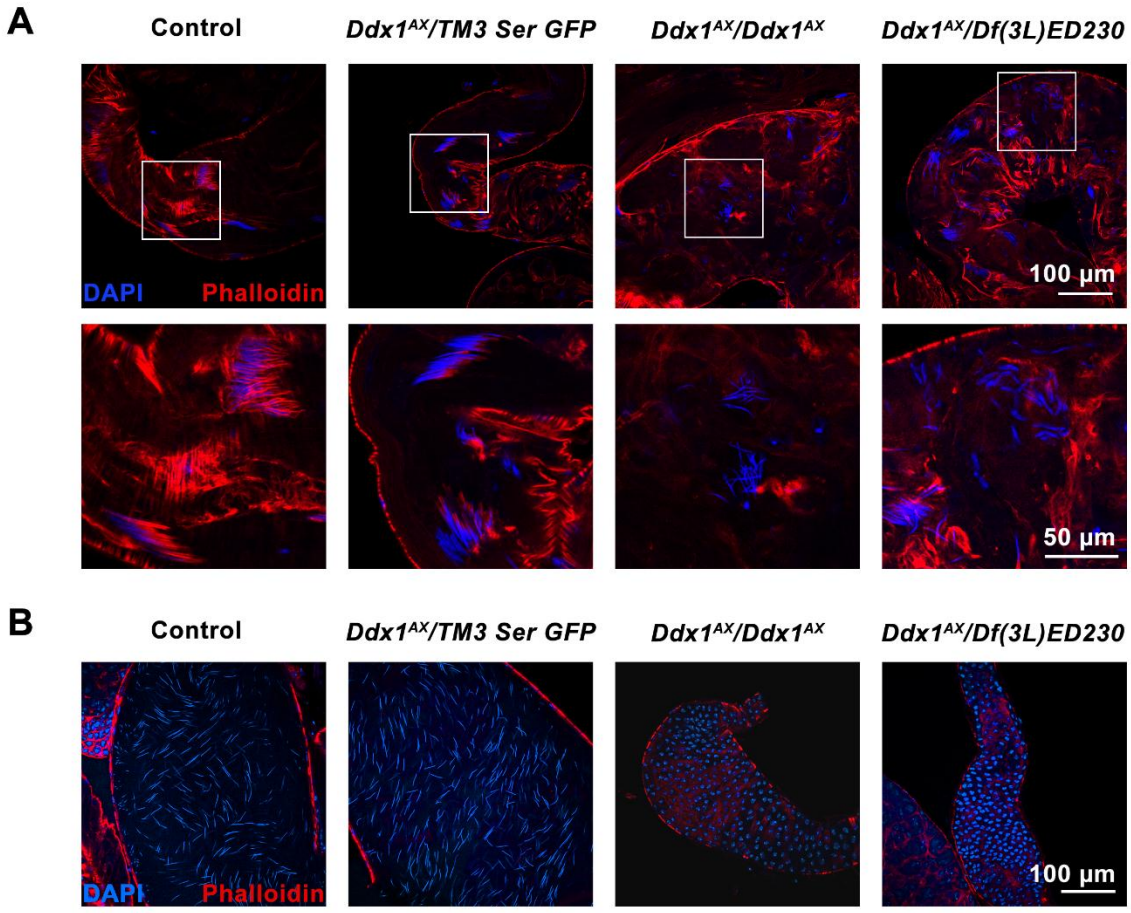


Figure 3.5 Aberrant spermiogenesis in *Ddx1*-null testes. Testes and seminal vesicles collected from male flies held in isolation for 10 days. Immunofluorescence imaging of testes (A, bottom increased magnification of outlined area) and seminal vesicles (B) using RFP conjugated-phalloidin and DAPI. *Ddx1*-null developing spermatid cysts become disordered during spermiogenesis and no mature sperm are observed in *Ddx1*-null seminal vesicles.

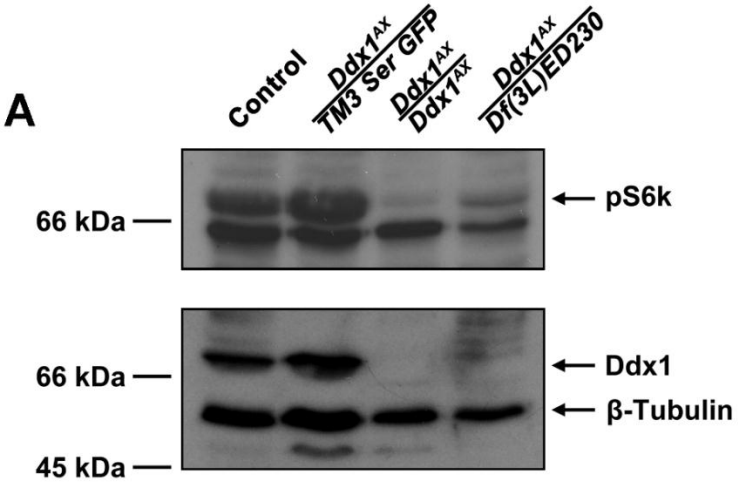
spermatogenesis appeared unaffected in *Ddx1*-null males, with developing spermatids undergoing nuclear elongation. However, spermatids at later stages of development lacked actin cones and were disrupted and scattered. To determine if mature sperm were being produced, seminal vesicles were isolated and imaged (Figure 3.5B). In contrast to control and heterozygous seminal vesicles, which contained mature sperm, we did not observe any mature sperm in the seminal vesicles of homozygous mutant males.

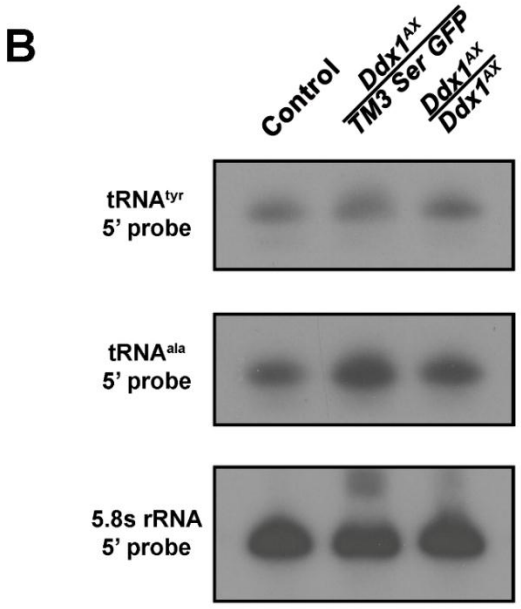
3.3.4 Ddx1-null flies have reduced pS6k levels, but normal tRNA levels

As both reduced body size and autophagy of developing oocysts are phenotypes associated with reduced metabolism (Barth *et al.* 2011, Edgar 2006), we performed western blot analysis for the phosphorylated form of S6k, a downstream effector molecule of TOR and a common marker of active growth (Montagne *et al.* 1999). Phospho-S6k is down-regulated in starvation conditions (Hara *et al.* 1998, Zhang *et al.* 2000). pS6k levels were notably reduced in *Ddx1*-null flies at the 3rd instar larval stage compared to control and heterozygous animals (Figure 3.6A). Notably, the pS6k signal observed in $Ddx1^{AX}/Df(3L)ED230$ larvae was stronger than that in $Ddx1^{AX/AX}$ larvae. This is consistent with our previous results demonstrating that the $Ddx1^{AX}/Df(3L)ED230$ phenotype is slightly attenuated compared to $Ddx1^{AX/AX}$ flies.

As human DDX1 has recently been implicated in tRNA splicing, and pS6k levels are associated with RNA Pol III activity (Marshall *et al.* 2012), we performed

Figure 3.6 Reduced pS6k levels in *Ddx1*-null flies. (A) Western blot analysis of cell lysates prepared from control, *Ddx1* heterozygous and *Ddx1*-null 3rd instar larvae. pS6k levels are reduced in the *Ddx1*-null lines, but slightly elevated in *Ddx1^{AX}/Df(3L)ED230* larvae as compared to *Ddx1$^{AX/AX}$*. (B) Northern blot analysis of tRNA and rRNA levels in control, heterozygous and homozygous mutant 3rd instar larvae. No difference is observed for spliced tRNA (tRNA^{tyr}), unspliced tRNA (tRNA^{ala}) or 5.8s rRNA.


A
Control
Ddx1AX / TM3 Ser GFP
Ddx1AX / Ddx1AX
Ddx1AX / Df(3L)ED230
66 kDa
pS6k
66 kDa
Ddx1
β-Tubulin
45 kDa


B
Control
Ddx1AX / TM3 Ser GFP
Ddx1AX / Ddx1AX
tRNAtyr 5' probe
tRNAala 5' probe
5.8s rRNA 5' probe

northern blot analysis of $tRNA^{tyr}$ (an intron containing tRNA) and $tRNA^{ala}$ (which contains no intron) levels in control, heterozygous and homozygous mutants. Our analysis showed no difference in the relative amounts of a mature tRNA that requires splicing ($tRNA^{tyr}$) compared to one that does not require splicing ($tRNA^{ala}$). As well, there was no difference in relative tRNA to 5S rRNA levels (Figure 3.6B).

3.3.5 Ddx1-null flies display widespread changes in mRNA levels and splicing

DEAD box proteins have been widely implicated in the generation, maturation and degradation of RNA molecules. Therefore, we undertook RNA deep sequencing of control and *Ddx1*$^{AX/AX}$ 3rd instar larvae to determine the effect of knocking-out *Ddx1* on the transcriptome. We identified 72 significantly down-regulated and 261 significantly up-regulated transcripts, using a cut-off of $p<0.01$ (Table 3.1). Of note, we found that RNA is expressed from the remaining portion of the *Ddx1*^{AX} allele, albeit at significantly reduced levels. We also analyzed transcripts for the presence of differentially spliced variants between *Ddx1*$^{AX/AX}$ and control flies (Table 3.2). We limited our analysis to the small number of transcripts that displayed unique splice variants with high levels in all three replicates of one genotype and absence in all three replicates of the other genotype. It is important to note that splice site analysis must be considered within the context of total RNA levels, as a gene that is only expressed in one genotype will necessarily have splice junctions that are unique to this genotype. Using these criteria, we identified

Table 3.1 Genes with differential expression in control and *Ddx1*-null 3rd instar larvae.

Gene	Region	*Ddx1*-null average	Control average	Fold change (log2)	p
CG10960	2-18205	36.90	48.86	0.41	6.66E-16
Cpr47Eg	11-489	1317.71	2.48	-9.05	1.31E-14
snoRNA:2R:9445205	0-150	10790.10	13.53	-9.64	1.95E-14
Lcp65Ag1	6-548	362.31	1.40	-8.01	1.23E-13
Lcp65Ag3	12-577	349.66	1.82	-7.58	5.95E-13
Lcp3	0-710	10976.30	80.81	-7.09	5.97E-13
CG42713	0-487	2.82	179.33	5.99	7.09E-13
Jon65Aii	1-895	49.50	848.22	4.10	3.25E-12
CG12522	0-498	117.02	0.40	-8.19	5.42E-12
Cpr65Ax2	16-792	1467.11	5.47	-8.07	1.20E-11
CG15741	0-484	81.45	0.46	-7.48	1.83E-11
CG14205	4-2940	126.77	1.45	-6.45	5.16E-11
CG32694	15-17187	39.52	7.91	-2.32	2.15E-10
LysX	0-555	36.23	544.61	3.91	3.65E-10
fon	1-3051	50.32	0.21	-7.93	6.14E-10
CR43105	4-664	0.58	52.29	6.48	7.10E-10
TwdlO	0-802	32.80	0.14	-7.88	2.53E-09
CG3397	0-1390	13.53	212.53	3.97	2.78E-09
CG44014	0-940	82.88	3.92	-4.40	3.60E-09
CR32205	537-1254	33.51	0.60	-5.80	7.77E-09
CG16898	9-1362	88.78	0.02	-12.09	7.88E-09
CG11350	0-1547	97.11	1.30	-6.22	1.77E-08
CG31226	0-463	176.17	423.08	1.26	1.81E-08
pip	6-38967	23.29	9.42	-1.31	1.90E-08
CG43153	1-1008	46.37	0.57	-6.34	7.92E-08
CG4753	37-2595	63.38	10.07	-2.65	8.07E-08
CG34451	0-3403	22.81	0.90	-4.67	9.25E-08
CG3292	8-1900	1.62	45.65	4.82	1.17E-07
Cpr65Av	0-639	201.94	6.03	-5.07	1.25E-07
CG44040	17-309	91.62	0.89	-6.69	1.51E-07
CG11617	41-2540	1.68	0.65	-1.38	1.68E-07
lectin-22C	0-857	41.07	0.76	-5.76	1.71E-07
CG10814	15-1521	8.07	0.43	-4.22	1.79E-07
CG11384	16-2159	5.62	0.10	-5.87	1.89E-07
Cyp6a8	0-1806	14.44	0.35	-5.36	3.95E-07
Cyp6a17	1-1527	58.78	1.84	-5.00	6.60E-07
CG5278	14-4174	13.92	0.45	-4.95	7.50E-07
CG7214	11-1158	54.91	1.94	-4.82	7.78E-07
Lcp65Ae	0-492	69.38	2.03	-5.10	1.09E-06
GILT3	1-1064	80.43	8.92	-3.17	1.16E-06
prc	4890-7252	5.59	0.14	-5.32	1.63E-06
CG13068	8-519	29.96	0.43	-6.14	2.03E-06
CG15067	13-246	300.45	2.03	-7.21	2.75E-06
CG16995	3-889	19.14	3.19	-2.58	3.12E-06
GstD4	14-702	13.57	0.46	-4.87	3.50E-06
CG15107	17-1219	5.56	13.16	1.24	3.69E-06
CG5724	2-1773	60.57	2.82	-4.43	3.94E-06
CG3106	6-3809	89.05	6.92	-3.68	4.01E-06
CG8630	61-4770	28.11	2.14	-3.71	4.34E-06
GstD7	19-748	51.54	1.76	-4.88	4.48E-06
CG4830	12-2202	94.41	5.43	-4.12	4.81E-06
Zasp66	8-3360	14.68	41.44	1.50	4.91E-06
lectin-28C	3-2294	5.34	45.26	3.08	5.19E-06
CG40198	2-1051	28.72	1.07	-4.74	5.43E-06
RYa-R	23252-26103	17.18	138.78	3.01	5.78E-06

Continued next page.

Gene	Region	*Ddx1*-null average	Control average	Fold change (log2)	p
CG9928	11-422	28.10	0.56	-5.64	6.00E-06
Gpdh	21-5279	235.97	63.19	-1.90	6.41E-06
CG15515	9-976	170.59	4.21	-5.34	6.85E-06
CG15597	10-739	17.74	0.39	-5.49	9.32E-06
CG11034	1-3669	1.92	19.81	3.37	9.34E-06
CG9463	12-3435	3.18	25.12	2.98	1.02E-05
CG10562	16-1634	11.52	0.81	-3.83	1.05E-05
CG5550	6-926	34.73	1.68	-4.37	1.11E-05
CG16775	0-697	22.44	1.10	-4.35	1.41E-05
CG32071	0-453	809.39	42.87	-4.24	1.43E-05
CG15404	4-470	123.52	704.86	2.51	1.53E-05
Nplp2	51-848	809.98	66.52	-3.61	1.84E-05
CG3835	9-3446	51.53	12.13	-2.09	1.93E-05
CG4842	7-1073	12.21	0.68	-4.16	2.20E-05
Cpr76Bc	213-6314	0.09	3.61	5.32	2.32E-05
Cpr47Eb	10-817	59.37	4.31	-3.78	2.44E-05
CG31698	0-544	134.05	719.02	2.42	2.74E-05
CG9813	26-6644	14.55	2.84	-2.36	3.00E-05
lambdaTry	12-950	27.08	0.92	-4.88	3.21E-05
CG9498	14-1642	17.36	0.96	-4.17	3.21E-05
Nha2	3-17060	35.53	4.10	-3.11	3.62E-05
CG42821	2-466	26.81	0.85	-4.98	3.63E-05
CG34031	8-11303	21.09	0.25	-6.41	3.87E-05
Lcp65Af	11-480	37.50	1.56	-4.59	4.15E-05
CG2681	16-2967	3.63	0.82	-2.15	4.23E-05
lectin-24A	5-967	10.67	0.41	-4.72	4.38E-05
Syt1	13-27650	15.21	32.76	1.11	4.44E-05
Syt7	398-26597	40.16	2.31	-4.12	5.33E-05
CG12607	4-1484	23.66	0.97	-4.61	5.36E-05
Ugt37c1	9-1722	4.66	0.21	-4.45	5.87E-05
CG15533	13-2543	35.62	2.47	-3.85	6.97E-05
galectin	0-3833	53.64	33.58	-0.68	7.25E-05
CG15829	11-771	413.68	44.94	-3.20	7.69E-05
Cyp309a1	4-2141	44.55	5.86	-2.93	8.04E-05
CG33926	90-3154	260.82	32.14	-3.02	8.13E-05
CG11659	11-1819	26.41	1.64	-4.01	8.15E-05
CG31769	0-1143	76.68	6.60	-3.54	8.47E-05
CG4563	13-1957	31.48	2.31	-3.77	8.83E-05
HmgZ	2812-10371	72.82	121.61	0.74	8.94E-05
CG14062	0-1250	0.20	6.16	4.92	9.33E-05
CG33267	186-8291	14.72	0.65	-4.51	9.85E-05
CG32751	314-2075	8.36	0.56	-3.91	1.03E-04
CG6484	13-2006	7.60	0.53	-3.83	1.07E-04
CG8736	1-965	8.72	0.45	-4.26	1.08E-04
CG15255	11-1457	38.72	3.18	-3.61	1.09E-04
CG11796	5-1707	12.17	2.20	-2.47	1.10E-04
TwdlP	6-749	14.71	0.67	-4.46	1.11E-04
CG14237	0-366	192.58	7.56	-4.67	1.21E-04
CG18417	6-1584	0.17	3.59	4.44	1.22E-04
Muc55B	0-2063	331.19	37.78	-3.13	1.34E-04
CG8629	0-352	675.65	70.99	-3.25	1.35E-04
CG13325	12381-15215	6.86	0.97	-2.82	1.36E-04
Acox57D-d	4-3035	39.76	4.16	-3.26	1.43E-04
IntS12	0-1183	44.38	5.81	-2.93	1.79E-04
CG34054	1-620	10.98	0.55	-4.32	1.79E-04

Continued next page.

Gene	Region	*Ddx1*-null average	Control average	Fold change (log2)	p
CG34454	0-653	42.48	4.13	-3.36	2.02E-04
CG9396	0-1208	40.96	3.57	-3.52	2.07E-04
Cpr11A	16-2479	57.94	5.47	-3.41	2.21E-04
Ace	15-36556	60.68	2.80	-4.44	2.24E-04
Ugt36Ba	3-1782	4.24	0.28	-3.92	2.47E-04
Pebp1	0-646	1801.10	207.19	-3.12	2.51E-04
CG13077	13-1497	5.84	0.82	-2.84	2.59E-04
CG32564	12-1448	15.67	0.93	-4.08	2.67E-04
Gasp	10-12089	132.64	51.27	-1.37	2.87E-04
CG11585	0-1235	7.39	0.50	-3.89	2.88E-04
mwh	7647-31888	0.61	6.37	3.37	3.27E-04
Fmo-2	7-2681	17.38	1.39	-3.64	3.46E-04
Mal-A1	8-2232	82.96	13.68	-2.60	3.49E-04
CR44709	3-486	13.41	0.97	-3.79	3.64E-04
GstE6	10-746	40.69	4.37	-3.22	3.77E-04
CG15661	20-2315	3.26	0.31	-3.40	3.80E-04
CG1773	1-1314	26.03	82.15	1.66	3.81E-04
CG8160	15-1117	10.13	0.62	-4.03	3.94E-04
CG11854	4-1077	153.42	900.49	2.55	3.98E-04
CR44344	64-1702	0.04	2.89	6.03	4.10E-04
Cyp9b1	11-1966	23.84	2.30	-3.37	4.27E-04
Ugt35a	6-1843	104.41	16.57	-2.66	4.36E-04
CG13893	3-4915	43.79	179.76	2.04	4.50E-04
CG32603	9-1247	71.19	8.36	-3.09	5.05E-04
CG1441	5-4202	1.08	0.10	-3.48	5.14E-04
prom	11030-12958	1.07	4.32	2.01	5.22E-04
CG12766	25-1498	3.94	0.23	-4.11	5.64E-04
CR44177	10-12608	0.17	2.96	4.11	6.23E-04
BG642312	8-913	10.69	1.08	-3.31	6.80E-04
CG43134	13-411	9.94	0.45	-4.47	6.82E-04
MtnC	0-327	117.83	13.34	-3.14	7.14E-04
mthl8	0-3881	11.09	0.00	N/A	7.24E-04
Dro	5-344	170.48	14.68	-3.54	7.38E-04
Or19b	8-1361	1.71	0.07	-4.56	7.47E-04
Cyp6w1	1-2310	74.88	10.31	-2.86	7.77E-04
wry	37-22012	0.83	1.33	0.68	7.99E-04
Ser6	4-900	28.08	2.68	-3.39	7.99E-04
CG5023	2-5370	184.46	32.88	-2.49	8.11E-04
TpnC25D	1357-2351	88.04	11.62	-2.92	8.16E-04
CG33301	10-1345	35.12	4.35	-3.01	8.19E-04
IM1	13-395	60.16	2.96	-4.35	8.41E-04
CG31041	0-831	44.95	4.89	-3.20	8.51E-04
GstD5	1-803	3.85	0.24	-4.00	8.99E-04
ppk13	1-1854	1.87	0.12	-3.94	9.04E-04
CG6330	14-8207	34.21	15.32	-1.16	9.06E-04
Cpr62Bc	17-1677	35.08	4.33	-3.02	9.13E-04
CG16712	2-351	1206.52	179.89	-2.75	9.28E-04
Acp1	12-957	3.08	0.07	-5.49	9.34E-04
CG9757	13-830	34.67	3.53	-3.29	9.41E-04
su(r)	15-5013	28.85	5.50	-2.39	9.60E-04
fa2h	10-5364	18.27	1.84	-3.31	9.67E-04
alpha-Est7	9-4130	5.50	23.57	2.10	9.80E-04
CG8907	1-4258	20.38	3.72	-2.45	9.91E-04
CG34330	0-699	229.78	41.90	-2.46	9.91E-04
Ugt86Dd	1-1726	9.51	1.03	-3.21	1.02E-03

Continued next page.

Gene	Region	*Ddx1*-null average	Control average	Fold change (log2)	p
lcs	16-590	68.22	6.83	-3.32	1.06E-03
AttD	12-804	0.39	5.31	3.76	1.07E-03
CG15534	0-3076	152.82	23.68	-2.69	1.11E-03
igl	13-34854	4.81	2.28	-1.08	1.12E-03
GstD6	0-766	11.16	0.89	-3.64	1.15E-03
prc	11-2028	3.73	0.11	-5.14	1.16E-03
CG14445	13-3212	8.90	1.78	-2.32	1.18E-03
CG12826	0-770	1.00	12.13	3.60	1.19E-03
CG14893	9-1908	3.08	0.29	-3.40	1.21E-03
CR44030	134-3017	14.62	4.70	-1.64	1.21E-03
CG42500	0-440	1414.26	8508.62	2.59	1.22E-03
CG9444	2-2850	14.96	58.40	1.96	1.27E-03
vkg	758-16047	2.35	0.09	-4.69	1.27E-03
CR44674	4-760	3.94	0.41	-3.27	1.27E-03
CG8774	194-3565	39.14	7.04	-2.47	1.29E-03
Ddx1	20-2645	3.55	16.16	2.19	1.33E-03
vanin-like	0-2360	28.70	4.52	-2.67	1.35E-03
CG14567	4-723	15.93	78.04	2.29	1.35E-03
Mal-A3	9-1984	63.05	11.80	-2.42	1.42E-03
CG43861	0-48546	2.56	4.66	0.86	1.55E-03
Eip74EF	3-60593	33.59	48.05	0.52	1.57E-03
CG13640	13-503	25.67	2.01	-3.68	1.57E-03
Hr38	24369-31076	1.87	0.46	-2.02	1.59E-03
Cht4	8-1700	11.50	1.54	-2.90	1.62E-03
Cyp12c1	0-1912	9.45	26.44	1.48	1.66E-03
Cyp12a5	2-2046	43.72	7.54	-2.54	1.69E-03
sv	23168-24518	0.42	3.71	3.16	1.72E-03
CG4716	0-860	39.71	4.82	-3.04	1.75E-03
mex1	0-2114	178.76	114.15	-0.65	1.78E-03
CG45784	2-303453	0.18	1.51	3.03	1.79E-03
CG14566	0-829	241.18	48.27	-2.32	1.82E-03
CR45658	4-841	2.45	0.07	-5.11	1.83E-03
CG15221	8-12369	7.94	0.94	-3.07	1.91E-03
CG33281	14-2401	42.16	7.57	-2.48	1.94E-03
CG18327	7-1252	7.21	0.89	-3.02	2.10E-03
CG11131	333-1496	84.67	12.60	-2.75	2.12E-03
Smvt	7-2171	19.16	2.77	-2.79	2.19E-03
path	11-10767	80.12	21.47	-1.90	2.25E-03
CG8087	3-483	233.12	31.94	-2.87	2.31E-03
CG11321	6-10229	2.02	0.32	-2.67	2.31E-03
CG30457	10-702	9.59	1.07	-3.16	2.40E-03
TpnC47D	0-1217	552.55	106.95	-2.37	2.41E-03
Mipp1	696-9690	35.29	6.89	-2.36	2.46E-03
CG9287	1281-3013	0.63	0.03	-4.35	2.60E-03
CG17167	14-12835	16.47	5.18	-1.67	2.62E-03
NT5E-2	1-8140	11.42	36.63	1.68	2.67E-03
CG13041	16-501	1.85	16.83	3.19	2.68E-03
CG3119	7-2925	6.13	1.65	-1.90	2.70E-03
CG7299	12-810	42.24	4.92	-3.10	2.79E-03
CG13060	148-599	0.14	3.39	4.61	2.80E-03
IM4	11-276	79.44	5.09	-3.97	2.84E-03
CG3706	0-1536	15.80	2.42	-2.71	2.87E-03
CG32444	0-2027	117.04	25.49	-2.20	2.88E-03
Est-6	15-1886	8.43	1.22	-2.79	2.90E-03
CG15615	2-3288	3.20	0.32	-3.33	2.95E-03

Continued next page.

Gene	Region	*Ddx1*-null average	Control average	Fold change (log2)	p
SNF4Agamma	72525-73530	0.74	5.68	2.94	3.04E-03
CG32248	87-680	5.40	0.35	-3.97	3.21E-03
PGRP-SB1	0-636	47.11	206.25	2.13	3.24E-03
Oseg5	168-2992	0.24	2.05	3.07	3.38E-03
CR45248	0-773	7.12	0.75	-3.24	3.38E-03
CG31259	1-2343	10.04	1.55	-2.69	3.43E-03
Cpr47Ec	6-858	17.46	2.07	-3.08	3.49E-03
Cyp6a18	13-3895	10.48	2.11	-2.31	3.55E-03
Ccp84Ag	17-965	2.51	0.15	-4.08	3.55E-03
CG5945	17-1703	1.24	0.06	-4.25	3.60E-03
Sgs3	0-1182	286.87	1305.19	2.19	3.61E-03
CG10311	0-6228	0.34	4.19	3.64	3.61E-03
CG4982	12-559	4.27	0.30	-3.83	3.61E-03
CG16713	3-373	654.43	114.66	-2.51	3.67E-03
CG14105	10-783	2.55	0.15	-4.11	3.75E-03
CG17475	16-1088	65.38	13.60	-2.27	3.76E-03
CG45073	0-461	0.13	3.39	4.71	3.76E-03
CG15068	11-277	139.45	775.27	2.47	3.81E-03
kappaTry	4-937	65.95	11.85	-2.48	3.81E-03
CG10560	11-1552	35.21	6.16	-2.52	3.90E-03
Faa	14-1918	16.76	3.10	-2.43	3.90E-03
CG33509	47-1899	1.67	0.18	-3.25	3.98E-03
CG33110	98-10201	5.22	1.30	-2.00	3.99E-03
CG43129	1076-4464	1.96	0.48	-2.02	4.07E-03
CG14958	0-644	18.28	2.77	-2.72	4.12E-03
Cyp6a21	9-1681	25.80	4.91	-2.39	4.15E-03
CR44404	0-228	281.97	26.95	-3.39	4.18E-03
Hsp70Bc	1251-2277	93.68	15.11	-2.63	4.20E-03
CG32241	0-1526	150.59	32.55	-2.21	4.21E-03
CG10433	31-2432	98.50	20.99	-2.23	4.22E-03
Obp83cd	7-1596	11.89	1.21	-3.30	4.26E-03
CG4962	4-792	135.50	27.48	-2.30	4.35E-03
CG32055	104-1661	0.35	2.65	2.92	4.39E-03
ETHR	2-18031	2.84	1.63	-0.80	4.40E-03
Cyp313a3	97-2127	4.99	0.62	-3.02	4.43E-03
CG34367	28-6051	2.74	8.14	1.57	4.52E-03
Ady43A	0-2254	24.32	76.40	1.65	4.52E-03
Ugt36Bb	13-1678	0.73	0.05	-3.92	4.58E-03
Mlp60A	331-1381	1567.51	324.35	-2.27	4.63E-03
CG12512	14-3093	9.29	1.35	-2.78	4.77E-03
CG10232	17-2276	3.69	0.65	-2.51	4.82E-03
Ccp84Ab	10-807	2.72	0.25	-3.46	4.94E-03
CG6839	0-1679	124.09	28.99	-2.10	4.97E-03
CG7203	27-985	15.73	2.16	-2.86	0.005063
CR44906	17-1668	1.97	0.26	-2.93	0.005326
Cpr31A	23-1721	7.78	1.07	-2.87	0.005343
CG40298	1-1161	4.29	0.53	-3.01	0.005411
CG5157	174-1379	8.83	1.37	-2.69	0.005415
CG2254	2-3253	118.94	27.95	-2.09	0.005427
CG34291	11-492	10.12	0.83	-3.60	0.005474
CG11399	6-5848	8.36	15.11	0.85	0.005568
Apoltp	5-21847	19.50	4.86	-2.01	0.00558
Mur18B	0-1681	254.97	77.12	-1.73	0.00566
Cyp12d1-p	791-1974	35.91	7.25	-2.31	0.005671
mwh	7647-31888	0.05	1.82	5.18	0.005721

Continued next page.

Gene	Region	*Ddx1*-null average	Control average	Fold change (log2)	p
CG18765	1-1406	2.75	0.27	-3.33	0.005738
CR46006	96-37168	4.49	8.82	0.97	0.005749
CG15293	31-1382	67.25	14.86	-2.18	0.005817
CG9400	22-1153	0.96	0.03	-4.82	0.005988
CG10553	3-1601	3.48	0.37	-3.21	0.006078
Pdxk	0-1644	176.22	48.92	-1.85	0.006143
Cyp4d8	0-2623	38.23	8.71	-2.13	0.006163
CG30272	17-2234	6.21	0.35	-4.14	0.006228
GstD2	0-734	29.31	4.12	-2.83	0.006335
CG13102	0-1958	40.25	10.00	-2.01	0.006473
CG8791	12-3270	7.13	1.87	-1.93	0.006486
CG6762	0-1193	279.53	493.74	0.82	0.00649
CG31021	4-1804	0.03	0.76	4.62	0.006534
CG17974	1-1634	9.75	1.43	-2.77	0.006597
CG17018	45532-60005	3.65	0.81	-2.18	0.006614
CG2187	535-3275	4.48	0.70	-2.67	0.006891
Sgs1	3028-4006	10.35	41.12	1.99	0.006982
CG15254	15-904	24.79	3.63	-2.77	0.007004
CG7432	414-5739	0.95	0.26	-1.88	0.007185
CG11841	0-1191	0.21	2.10	3.29	0.007243
CG31288	4-1598	11.24	1.79	-2.65	0.007246
Cht8	0-1567	0.11	1.17	3.38	0.007367
CG12560	0-1169	16.50	3.09	-2.42	0.007403
GstD10	0-933	49.51	10.98	-2.17	0.007442
CG10516	0-1699	21.90	67.80	1.63	0.007507
CG42819	13-603	3.05	0.25	-3.63	0.007508
Cpr65Ay	0-554	49.83	8.20	-2.60	0.007619
CG10943	6-1172	64.15	17.71	-1.86	0.007919
mwh	7647-31888	5.99	1.29	-2.22	0.007979
CG31948	20-1354	12.96	46.96	1.86	0.007988
BBS1	3-2214	8.52	1.55	-2.46	0.007999
CG15554	21-1186	2.23	0.28	-2.97	0.008046
Zip99C	2-5195	22.91	48.07	1.07	0.008082
CG11300	10-568	25.29	2.96	-3.09	0.008188
sim	9-20362	0.55	0.90	0.70	0.008222
CR33319	3-699	21.65	0.00	N/A	0.008464
CG9527	2-21657	16.56	4.34	-1.93	0.008534
Npc1b	0-4538	13.23	3.36	-1.98	0.008558
CG13954	47-1305	1.00	0.11	-3.24	0.008723
c12.2	0-5097	4.10	1.90	-1.11	0.008807
CG4091	8-3458	31.48	76.95	1.29	0.008818
Vha100-4	1-2965	19.81	4.77	-2.06	0.008864
AcCoAS	2-9027	71.73	19.40	-1.89	0.008871
CG13170	79-377	10.08	0.61	-4.05	0.009067
CG30265	167-2108	0.86	0.10	-3.14	0.009242
sun	0-1332	961.32	457.91	-1.07	0.009252
CG3523	12-11577	222.35	38.12	-2.54	0.009334
CG7560	5-1219	1.15	0.10	-3.48	0.009356
arg	4-18593	25.41	5.33	-2.25	0.0094
CG14253	1865-7134	36.55	21.16	-0.79	0.009597
CG31265	16-905	19.82	3.35	-2.57	0.009803
CG4098	0-1210	13.92	4.64	-1.59	0.009827
CG8028	40-3166	9.49	1.86	-2.35	0.009862
Trpgamma	52-8482	2.18	0.23	-3.27	0.009934
CG4377	8-892	57.28	12.92	-2.15	0.009964
CG9743	10-6639	14.47	3.06	-2.24	0.009966

Table 3.2 Genes with modified splicing in control and *Ddx1*-null larvae.

Gene name	Splice donor site	Splice acceptor site	*Ddx1*-null larvae				Control Larvae				p
			Sample 1	Sample 2	Sample 3	Average	Sample 1	Sample 2	Sample 3	Average	
CG30456	5963	6025	218	215	222	218	157	162	164	161	3.91E-05
AGO1	8240	8302	252	259	241	251	425	429	418	424	4.55E-05
phm	813	1192	79	82	75	79	175	174	176	175	1.75E-04
FKBP59	542	662	748	796	828	791	316	315	278	303	3.02E-04
scra	343	449	45	62	57	55	137	137	123	132	3.61E-04
Fili	62563	71222	27	57	41	42	191	170	181	181	3.71E-04
Npc1a	3617	3777	436	485	436	452	201	229	174	201	3.87E-04
CG3831	496	581	63	72	72	69	119	119	113	117	4.12E-04
nuf	21363	22219	13	24	24	20	132	150	154	145	4.36E-04
CG42813	455	581	0	0	26	9	144	122	140	135	4.43E-04
CG40160	5419	5845	34	37	30	34	67	68	75	70	4.53E-04
Idgf3	401	469	133	121	140	131	428	476	452	452	5.04E-04
Reps	1448	1516	190	193	193	192	172	176	172	173	5.41E-04
CG32164	1672	1728	99	86	100	95	33	30	23	29	5.66E-04
Chd64	6666	6742	35	51	74	53	194	233	226	218	5.81E-04
CG6067	1652	1721	179	173	163	172	62	84	63	70	6.47E-04
CG14764	570	1571	193	204	208	202	133	143	139	138	6.68E-04
CR33013	1173	1227	6	18	10	11	183	198	214	198	6.81E-04
CG2224	1030	1522	52	27	28	36	183	196	158	179	7.56E-04
endos	900	1286	258	225	307	263	616	611	571	599	7.96E-04
CG14967	3623	3685	86	75	90	84	15	23	16	18	8.16E-04
CG4729	3576	3800	609	600	597	602	468	456	483	469	8.28E-04
mfas	6481	6543	58	30	52	47	135	153	162	150	9.09E-04
CG5149	3204	3262	436	353	408	399	62	118	97	92	9.29E-04
CG10585	743	801	89	83	85	86	42	48	52	47	9.56E-04
CG5149	2994	3113	254	225	245	241	94	130	120	115	9.58E-04
wkd	264	1738	47	41	48	45	111	121	106	113	9.59E-04
CG33303	610	666	413	463	496	457	1426	1527	1331	1428	9.69E-04
CG34056	601	657	58	36	64	53	187	155	186	176	9.73E-04
CG43780	147	328	213	292	276	260	625	590	698	638	9.74E-04
CG43781	147	328	213	292	276	260	625	590	698	638	9.74E-04
CG4849	387	447	141	143	157	147	95	77	86	86	1.09E-03
Fbp1	179	245	183	405	46	211	1205	1460	1493	1386	1.16E-03
CG18278	1515	1574	60	65	70	65	102	95	97	98	1.17E-03
CG18600	148	211	27	38	41	35	86	86	76	83	1.20E-03
Cyp12c1	451	628	78	67	42	62	247	257	303	269	1.26E-03
Nrt	8295	8358	84	76	52	71	181	222	211	205	1.28E-03
Rga	2625	2683	75	92	91	86	155	142	148	148	1.30E-03
Tao	7131	7195	193	202	200	198	225	232	228	228	1.30E-03
CG9304	749	1052	60	93	74	76	209	176	211	199	1.31E-03
CR33013	840	896	0	15	2	6	97	119	125	114	1.31E-03
lark	352	696	293	242	277	271	87	108	48	81	1.33E-03
CG12075	25936	29226	119	123	107	116	44	38	44	42	1.34E-03
mdy	133	3452	323	386	396	368	115	160	92	122	1.36E-03
CG32264	82821	83354	40	58	52	50	146	132	161	146	1.38E-03
CG5850	2039	2153	40	110	79	76	313	268	307	296	1.46E-03
CG5958	111	305	1367	1353	1095	1272	2389	2530	2493	2471	1.46E-03
CG5027	1225	1282	36	29	51	39	104	109	127	113	1.47E-03
NtR	1074	1146	33	35	37	35	162	148	167	159	1.47E-03
syd	7652	7716	8	29	26	21	86	109	90	95	1.60E-03
dsd	628	3909	92	82	88	87	59	53	61	58	1.63E-03
Tep2	2223	2285	94	96	106	99	275	309	311	298	1.63E-03
CG32165	1615	1674	2	58	50	37	272	206	228	235	1.67E-03
Cyp6d2	792	865	42	131	111	95	379	397	325	367	1.69E-03
CG13876	185	242	12	8	26	15	100	93	98	97	1.71E-03
mr	1012	1087	74	83	88	82	45	39	35	40	1.74E-03
qtc	176	1926	145	126	156	142	41	61	60	54	1.77E-03
CG17652	477	540	100	110	127	112	18	42	37	32	1.78E-03
ft	12647	12707	5	26	22	18	123	93	103	106	1.81E-03
Rapgap1	78868	78926	137	146	158	147	49	78	60	62	1.86E-03
CG10289	2533	2588	169	152	147	156	84	81	95	87	1.88E-03
or	147	211	164	179	157	167	261	278	298	279	1.99E-03
CG30185	139	207	31	53	73	52	207	203	188	199	2.01E-03
rtet	1749	1806	87	62	75	75	188	158	191	179	2.03E-03
Tpll	1646	1705	278	277	241	265	411	468	424	434	2.04E-03

Continued next page.

Gene name	Splice donor site	Splice acceptor site	*Ddx1*-null larvae				Control Larvae				p
			Sample 1	Sample 2	Sample 3	Average	Sample 1	Sample 2	Sample 3	Average	
capu	15877	18342	113	91	115	106	24	22	45	30	2.05E-03
aralar1	6822	6880	1011	915	1086	1004	483	614	478	525	2.06E-03
CG7185	2492	2752	119	103	124	115	199	193	223	205	2.10E-03
Fak	5468	5522	91	98	105	98	49	59	63	57	2.12E-03
CG7967	334	418	809	862	953	875	452	411	543	469	2.15E-03
MRP	11107	18914	90	98	92	93	61	60	58	60	2.15E-03
CG9304	1753	1817	36	28	14	26	114	89	115	106	2.23E-03
Mal-A5	821	878	659	628	569	619	978	1017	1120	1038	2.30E-03
Nedd8	431	491	238	206	251	232	380	395	344	373	2.32E-03
pav	155	214	130	120	149	133	41	59	64	55	2.37E-03
CG42324	34761	34851	95	85	100	93	17	31	40	29	2.39E-03
Pi3K92E	2955	3017	45	74	82	67	197	158	195	183	2.48E-03
CG42663	44290	44351	67	79	58	68	131	141	135	136	2.55E-03
CG11710	1623	1679	41	46	51	46	73	67	76	72	2.74E-03
CG5355	2373	2427	286	483	419	396	977	935	816	909	2.74E-03
CG7741	1847	1904	52	44	42	46	78	71	70	73	2.75E-03
CG6498	8459	8523	40	51	48	46	82	72	74	76	2.76E-03
CG3967	455	7406	66	62	83	70	224	187	186	199	2.79E-03
CG42671	13584	14608	116	115	136	122	189	191	171	184	2.83E-03
CG5850	1450	1708	223	280	267	257	423	500	446	456	2.88E-03
Ncc69	18090	18162	56	61	72	63	118	102	104	108	2.90E-03
CG7420	1808	1868	101	99	101	100	61	67	58	62	3.06E-03
Pez	5188	5250	91	99	87	92	57	61	65	61	3.08E-03
CG7185	431	826	51	37	37	42	89	79	82	83	3.14E-03
LM408	1860	2222	67	123	73	88	259	247	218	241	3.20E-03
sgg	37280	37356	280	316	322	306	450	455	510	472	3.26E-03
cenB1A	2155	2217	110	134	119	121	11	11	8	10	3.40E-03
CG4747	3517	3589	78	114	106	99	185	231	217	211	3.58E-03
gish	28900	29353	96	102	105	101	122	122	127	124	3.62E-03
CG6509	7058	7117	93	100	104	99	66	73	59	66	3.64E-03
Drp1	261	1029	285	333	327	315	161	144	143	149	3.68E-03
mthl3	3334	3396	538	438	472	483	107	133	136	125	3.70E-03
oys	369	8770	824	1033	993	950	1512	1365	1534	1470	3.73E-03
CG6356	2773	2837	165	197	150	171	17	36	34	29	3.74E-03
scrib	18590	18883	24	20	41	28	307	390	334	344	3.83E-03
exo70	3810	3865	146	136	131	138	90	91	96	92	3.86E-03
mr	139	200	120	95	119	111	56	38	41	45	3.98E-03
e(r)	744	864	6	8	8	7	122	116	142	127	4.09E-03
CG17652	213	276	87	112	102	100	32	54	41	42	4.10E-03
holn1	447	512	40	35	47	41	92	91	75	86	4.10E-03
CG15525	239	301	10	60	38	36	370	370	275	338	4.17E-03
Tsp42Ea	4192	4255	738	714	625	692	349	329	457	378	4.21E-03
mRpL43	253	315	356	313	370	346	240	201	198	213	4.32E-03
Treh	6794	10475	61	91	38	63	161	189	199	183	4.33E-03
CG30344	2836	2893	206	246	143	198	432	560	477	490	4.33E-03
svp	41900	42148	147	151	146	148	93	102	108	101	4.36E-03
Plap	452	518	153	155	179	162	236	236	213	228	4.44E-03
Klp31E	1392	1453	15	11	20	15	94	90	115	100	4.49E-03
Hrb27C	2123	3080	2213	2230	2578	2340	1330	1393	1186	1303	4.51E-03
CG42868	185	254	35	60	51	49	190	152	201	181	4.52E-03
dlg1	30152	35609	84	93	97	91	56	61	64	60	4.58E-03
Pex3	181	314	36	80	50	55	203	161	154	173	4.63E-03
mp	50122	51722	67	75	55	66	105	121	106	111	4.63E-03
Cdk4	5410	5482	166	158	178	167	91	86	115	97	4.70E-03
CG9485	2787	2837	37	18	57	37	111	136	154	134	4.73E-03
Fak	3356	3417	156	189	177	174	104	86	113	101	4.78E-03
CG10289	2653	2711	162	138	164	155	84	101	85	90	4.79E-03
Cpsf73	359	421	48	28	27	34	77	95	86	86	4.79E-03
vir-1	117	1610	75	70	152	99	357	443	504	435	4.80E-03
fwd	13056	13115	202	218	235	218	142	153	133	143	4.80E-03
CG33156	1133	3324	32	35	36	34	80	72	86	79	4.86E-03
Nhe3	3300	3366	121	163	139	141	250	223	225	233	4.94E-03
CG3437	904	989	11	36	26	24	106	78	83	89	4.99E-03
FKBP59	1566	1627	786	822	785	798	534	518	433	495	5.06E-03
CG30020	3533	3603	31	11	34	25	68	84	92	81	5.17E-03
CG42593	15749	15822	79	97	88	88	123	138	123	128	5.19E-03
esn	19296	32101	58	87	112	86	223	199	180	201	5.21E-03
Mbs	24690	25023	15	18	18	17	120	96	105	107	5.24E-03

Continued next page.

Gene name	Splice donor site	Splice acceptor site	*Ddx1*-null larvae				Control Larvae				p
			Sample 1	Sample 2	Sample 3	Average	Sample 1	Sample 2	Sample 3	Average	
Cortactin	1046	1137	90	113	108	104	197	242	208	216	5.24E-03
shark	741	803	85	66	73	75	142	166	135	148	5.29E-03
Sgs5	543	602	6219	18177	15602	13333	53147	75583	65735	64822	5.32E-03
CG4398	657	775	43	27	40	37	78	80	85	81	5.37E-03
Pp1alpha-96A	923	1326	92	89	106	96	153	150	155	153	5.39E-03
Exn	22537	22604	130	135	165	143	62	49	73	61	5.48E-03
CG17646	13506	15168	102	138	140	127	17	44	50	37	5.53E-03
CG12006	153	211	201	251	281	244	77	57	35	56	5.59E-03
Pl31	1376	1636	345	296	335	325	193	231	228	217	5.62E-03
lin	484	541	67	68	77	71	36	49	37	41	5.70E-03
rictor	4180	4243	47	44	53	48	74	88	77	80	5.74E-03
Hel25E	132	338	1073	1056	895	1008	402	463	434	433	5.74E-03
Mio	22847	23432	140	146	171	152	74	77	100	84	5.76E-03
pio	15466	15519	315	535	513	454	860	1052	951	954	5.79E-03
mon2	5107	5161	315	304	326	315	260	270	261	264	5.82E-03
CG3940	5897	5968	288	278	355	307	109	159	119	129	5.84E-03
fwd	12903	12967	246	242	267	252	181	176	172	176	5.88E-03
CG31414	1677	1737	121	96	101	106	171	148	167	162	5.89E-03
Mal-A5	303	360	485	730	743	653	1145	1453	1376	1325	5.93E-03
robo	2519	2579	48	58	36	47	104	94	94	97	5.93E-03
casp	2027	2094	133	133	154	140	190	221	217	209	5.95E-03
CG42533	22360	22418	47	52	45	48	63	66	62	64	6.00E-03
G-salpha60A	3284	3357	424	406	441	424	341	336	368	348	6.03E-03
Lasp	26237	26300	191	149	160	167	68	54	75	66	6.08E-03
CG5001	5778	5842	520	614	565	566	375	340	405	373	6.10E-03
N	35688	35763	73	89	88	83	134	115	122	124	6.10E-03
CG34133	2859	2919	99	95	106	100	56	59	39	51	6.16E-03
Rox8	3908	3975	209	313	295	272	550	484	461	498	6.19E-03
rut	3403	3466	200	154	162	172	55	57	79	64	6.20E-03
eIF-5A	868	1108	9803	9665	10742	10070	7154	7381	7751	7429	6.21E-03
cher	26828	27031	34	226	123	128	621	627	442	563	6.23E-03
CG12121	758	820	149	192	233	191	382	382	338	367	6.38E-03
CG31522	12412	14929	29	84	82	65	170	218	239	209	6.40E-03
aralar1	8456	8519	238	198	239	225	111	136	135	127	6.44E-03
aralar1	6061	6659	369	390	439	399	204	223	232	220	6.47E-03
CG18542	212	274	101	90	113	101	173	166	202	180	6.65E-03
CG32939	213	275	101	90	113	101	173	166	202	180	6.65E-03
scra	2959	3020	14	13	13	13	124	100	101	108	6.66E-03
mthl3	2280	2336	556	631	772	653	155	182	289	209	6.67E-03
NT5E-2	130	5935	13	66	50	43	146	145	196	162	6.67E-03
CG12173	434	743	219	245	263	242	138	166	153	152	6.69E-03
heph	13756	17313	202	538	377	372	985	1176	952	1038	6.75E-03
CG4020	1773	1843	115	143	107	122	190	235	209	211	6.75E-03
eIF3-S10	1379	1429	892	850	948	897	608	675	722	668	6.76E-03
Msh6	3137	3205	70	92	101	88	37	16	20	24	6.76E-03
fax	6514	6669	203	502	259	321	1172	1318	909	1133	6.92E-03
TM9SF4	160	250	70	64	78	71	104	96	97	99	7.00E-03
Snap25	172125	172188	180	143	176	166	67	60	76	68	7.02E-03
CG9616	451	514	147	174	196	172	83	52	92	76	7.02E-03
Abi	1726	1948	305	259	316	293	148	165	140	151	7.05E-03
Aats-cys	1567	1626	132	137	160	143	67	80	75	74	7.06E-03
sec10	2373	2445	95	105	107	102	78	82	72	77	7.09E-03
grsm	17305	17385	465	788	653	635	2038	2714	2144	2299	7.11E-03
ABCB7	2141	2505	91	72	98	87	16	26	20	21	7.18E-03
CG33275	44985	45066	114	110	106	110	49	65	68	61	7.19E-03
NT5E-2	6162	6213	56	124	84	88	212	230	288	243	7.23E-03
CG3732	665	721	234	241	230	235	161	187	176	175	7.29E-03
CG8306	4830	4886	92	55	100	82	228	275	319	274	7.30E-03
Ncc69	122	11800	44	36	41	40	75	76	91	81	7.33E-03
HDAC4	19484	19554	150	146	145	147	109	117	122	116	7.34E-03
CG5292	634	690	81	92	95	89	56	52	67	58	7.47E-03
chic	1200	1314	90	149	102	114	219	279	230	243	7.47E-03
scrib	252	1021	42	72	61	58	102	126	127	118	7.54E-03
Psn	240	299	44	54	66	55	92	99	105	99	7.55E-03
Aac11	1979	2046	113	143	127	128	197	179	185	187	7.56E-03
RhoBTB	3928	3993	235	221	229	228	250	261	265	259	7.60E-03
CG7337	25528	47434	65	68	60	64	81	82	77	80	7.60E-03
kto	5619	5678	66	62	56	61	35	42	45	41	7.60E-03

Continued next page.

Gene name	Splice donor site	Splice acceptor site	*Ddx1*-null larvae				Control Larvae				p
			Sample 1	Sample 2	Sample 3	Average	Sample 1	Sample 2	Sample 3	Average	
CG8785	6185	6241	136	152	161	150	86	79	85	83	7.68E-03
dac	10724	12403	51	51	43	48	75	69	64	69	7.77E-03
CG12090	9610	9677	92	79	74	82	34	39	41	38	7.78E-03
Cortactin	177	363	22	28	33	28	178	139	139	152	7.85E-03
CG8090	1566	1618	97	132	119	116	56	35	52	48	7.91E-03
Cortactin	574	642	260	267	288	272	389	425	457	424	7.93E-03
Slh	1939	1993	90	103	98	97	119	135	131	128	7.94E-03
pncr013:4	498	1660	70	84	87	80	36	43	43	41	8.03E-03
Hydr1	868	1119	97	75	82	85	49	31	23	34	8.04E-03
Mef2	39895	40010	313	334	286	311	235	228	196	220	8.05E-03
Cnot4	2961	3052	71	71	61	68	91	102	108	100	8.06E-03
CG8317	314	1248	17	106	51	58	284	411	293	329	8.09E-03
CanB2	530	1224	299	297	358	318	422	445	475	447	8.14E-03
CG6067	245	1423	94	125	131	117	36	60	32	43	8.18E-03
ed	48388	48449	125	170	179	158	321	279	366	322	8.22E-03
CG33774	382	461	327	204	264	265	535	526	435	499	8.32E-03
rhea	23744	23810	260	284	352	299	65	87	104	85	8.32E-03
pav	1581	1641	44	39	50	44	97	87	76	87	8.34E-03
emb	3283	3351	211	202	202	205	241	229	246	239	8.37E-03
CG8726	2135	2234	63	66	65	65	80	79	86	82	8.37E-03
dally	61558	61659	59	34	55	49	89	98	110	99	8.41E-03
CG6426	2039	2109	1610	1537	1451	1533	928	1139	1143	1070	8.42E-03
CG11089	90	2012	85	106	123	105	0	5	10	5	8.45E-03
bchs	13628	13695	35	23	19	26	84	96	67	82	8.45E-03
CG3726	10642	13593	181	163	207	184	56	61	65	61	8.49E-03
Cad87A	54021	54081	99	103	83	95	48	62	57	56	8.56E-03
mRNA-cap	445	509	74	83	82	80	103	119	109	110	8.57E-03
CG9384	1087	1176	128	126	122	125	81	97	89	89	8.62E-03
GNBP1	1281	1345	158	126	127	137	52	78	79	70	8.63E-03
CG17660	214	318	268	228	293	263	95	89	107	97	8.73E-03
Tsp42Ea	4678	4941	858	1028	957	948	414	324	590	443	8.78E-03
Mvl	2785	2849	189	230	226	215	295	328	356	326	8.81E-03
CG33123	1115	1182	66	73	86	75	30	44	40	38	8.88E-03
CG5463	1364	1433	22	36	44	34	66	85	74	75	8.89E-03
CG32473	7308	7498	122	152	131	135	53	54	48	52	8.90E-03
Rox8	4069	4138	109	190	146	148	269	331	287	296	8.94E-03
cnc	8833	16828	47	84	71	67	147	151	134	144	8.99E-03
cv-d	4862	4920	50	57	55	54	66	70	67	68	9.03E-03
Wwox	4010	4077	306	232	288	275	112	129	166	136	9.08E-03
CG44009	2956	3016	184	134	185	168	266	343	321	310	9.29E-03
Dsp1	802	2655	55	70	65	63	97	115	122	111	9.32E-03
Nmda1	1179	1520	677	805	897	793	149	174	170	164	9.34E-03
CG7523	241	358	500	517	614	544	261	331	367	320	9.39E-03
RfC3	113	169	34	36	30	33	79	80	64	74	9.43E-03
cher	26828	27131	27	302	108	146	895	948	600	814	9.47E-03
Sgs7	63	129	32354	85565	81129	66349	170712	250014	204234	208320	9.47E-03
CG42353	109	4563	249	222	245	239	170	181	174	175	9.48E-03
CG42354	109	4563	249	222	245	239	170	181	174	175	9.48E-03
CR44169	2970	4072	92	94	93	93	61	67	71	66	9.51E-03
CG8519	3620	3680	91	118	90	100	152	191	188	177	9.51E-03
CG8317	314	1287	256	216	463	312	5889	7829	8163	7294	9.51E-03
CanB2	259	490	312	309	364	328	419	440	460	440	9.52E-03
Nmda1	542	1084	509	566	681	585	126	170	143	146	9.53E-03
vir-1	13984	15705	2092	2105	2440	2212	2750	2984	3011	2915	9.53E-03
CG8785	4348	4840	334	360	386	360	230	136	165	177	9.55E-03
CG10555	1064	1177	90	110	100	100	130	132	145	136	9.62E-03
CG1021	12164	12936	42	56	47	48	69	81	72	74	9.65E-03
holn1	195	254	27	27	47	34	85	72	67	75	9.72E-03
TBPH	1407	1469	38	61	57	52	87	96	104	96	9.73E-03
Med	3402	3458	70	72	71	71	58	63	60	60	9.83E-03
CG13046	112	174	44	10	52	35	153	132	101	129	9.86E-03
CG4098	649	705	227	155	170	184	0	20	0	7	9.88E-03
Cat	239	3734	3098	2673	2842	2871	2136	2111	1691	1979	9.90E-03
srp	636	6968	51	112	88	84	207	189	263	220	9.93E-03
CR44133	607	6939	51	112	88	84	207	189	263	220	9.93E-03
tomosyn	5720	5833	92	86	114	97	34	48	54	45	9.95E-03

Table 3.3 Genes with unique splice variants in *Ddx1*-null larvae.

Gene name	Splice donor site	Splice acceptor site	*Ddx1*-null larvae				Control Larvae			
			Sample 1	Sample 2	Sample 3	Average	Sample 1	Sample 2	Sample 3	Average
RNaseMRP:RNA	167	231	3212	5499	11	2907	0	0	0	0
Sirup	234	315	2009	314	2670	1664	0	0	0	0
CG5770	80	135	1560	741	1440	1247	0	0	0	0
CG9021	85	215	52	2949	141	1047	0	0	0	0
RpS21	499	561	237	559	2327	1041	0	0	0	0
CG16898	1041	1100	951	27	1125	701	0	0	0	0
CG10365	6868	6932	581	1273	9	621	0	0	0	0
Jon65Aiv	298	377	20	1590	9	540	0	0	0	0
Muc55B	1074	1182	1185	8	324	506	0	0	0	0
CG10311	216	4656	519	157	646	441	0	0	0	0
CG13705	894	1029	522	21	768	437	0	0	0	0
CG14300	56	112	461	118	547	375	0	0	0	0
Muc55B	1074	1224	867	3	221	364	0	0	0	0
Muc55B	1074	1266	682	4	170	285	0	0	0	0
CG44085	10965	11226	278	136	433	282	0	0	0	0
CG13705	351	486	278	21	419	239	0	0	0	0
Muc55B	1074	1308	493	2	109	201	0	0	0	0
Cyp12d1-p	695	759	3	114	483	200	0	0	0	0
Cyp12d1-d	692	756	3	114	483	200	0	0	0	0
Myd88	3489	3562	20	252	320	197	0	0	0	0
spen	16416	24729	133	168	281	194	0	0	0	0
mop	1372	1438	177	228	149	185	0	0	0	0
CG7299	284	401	222	23	301	182	0	0	0	0
Nipped-A	31913	31976	190	113	204	169	0	0	0	0
CG9894	396	927	181	174	136	164	0	0	0	0
CG9682	1402	1477	209	24	255	163	0	0	0	0
CG7231	1488	1607	163	145	128	145	0	0	0	0
Muc55B	1074	1350	329	3	81	138	0	0	0	0
CG16727	263	473	151	123	124	133	0	0	0	0
CG33774	73	261	299	51	43	131	0	0	0	0
CG17752	592	661	245	15	112	124	0	0	0	0
CG7300	546	720	264	1	102	122	0	0	0	0
CG13705	462	867	151	5	209	122	0	0	0	0
Mocs2	429	490	90	145	129	121	0	0	0	0
arr	26694	26756	107	122	123	117	0	0	0	0
ATPsyn-beta	468	1423	150	73	108	110	0	0	0	0
Muc55B	1074	1392	269	2	56	109	0	0	0	0
sala	67	474	136	94	87	106	0	0	0	0
Pex3	195	314	90	103	111	101	0	0	0	0
Taf12L	312	372	27	112	164	101	0	0	0	0
spen	3970	16329	70	109	123	101	0	0	0	0

Table 3.4 Genes with unique splice variants in control larvae.

Gene name	Splice donor site	Splice acceptor site	*Ddx1*-null larvae				Control Larvae			
			Sample 1	Sample 2	Sample 3	Average	Sample 1	Sample 2	Sample 3	Average
CG17834	227	6777	0	0	0	0	134	107	67	103
CG5455	120	5058	0	0	0	0	124	89	129	114
CG10311	220	4670	0	0	0	0	58	87	168	104
Hsromega	14855	18301	0	0	0	0	433	405	514	451
CG33203	2428	5339	0	0	0	0	143	166	67	125
ldh	114	2191	0	0	0	0	129	172	66	122
Atet	527	2537	0	0	0	0	150	165	130	148
CG9701	255	876	0	0	0	0	110	120	196	142
CG5096	185	744	0	0	0	0	168	276	468	304
CG9894	393	927	0	0	0	0	161	224	280	222
CG43166	91	527	0	0	0	0	2360	3089	3036	2828
Cpr67B	125	493	0	0	0	0	98	184	245	176
CG9682	1423	1783	0	0	0	0	240	164	182	195
dom	6884	7186	0	0	0	0	131	111	109	117
CREG	105	374	0	0	0	0	907	1271	768	982
yellow-f	219	454	0	0	0	0	106	178	82	122
yellow-f	219	439	0	0	0	0	150	288	114	184
CG11309	369	567	0	0	0	0	113	126	163	134
Mur29B	2091	2283	0	0	0	0	205	257	56	173
grsm	10771	10954	0	0	0	0	79	111	136	109
Npc1a	3605	3777	0	0	0	0	326	430	331	362
CG9682	1651	1783	0	0	0	0	216	203	208	209
Sgs3	363	483	0	0	0	0	852	1773	2653	1759
CG9682	1684	1783	0	0	0	0	326	316	251	298
alpha-Est5	1589	1674	0	0	0	0	286	410	525	407
Mur18B	631	715	0	0	0	0	135	74	187	132
CG12826	347	428	0	0	0	0	105	176	196	159
Fbp2	399	474	0	0	0	0	1673	34	2811	1506
Lcp2	310	385	0	0	0	0	13	3	2449	822
CG5973	7368	7435	0	0	0	0	157	258	451	289
CG5792	8401	8468	0	0	0	0	503	466	409	459
CG18812	17077	17139	0	0	0	0	492	469	274	412
CG2246	6100	6161	0	0	0	0	197	158	146	167
CG32388	688	749	0	0	0	0	661	203	98	321
CG5322	3266	3326	0	0	0	0	235	259	277	257
Fst	593	653	0	0	0	0	2	210	279	164
CG9463	122	180	0	0	0	0	240	45	360	215
Pxd	11000	11057	0	0	0	0	592	711	2102	1135
EloA	1274	1329	0	0	0	0	278	254	227	253
CR43105	36	91	0	0	0	0	1013	701	688	801
Ddx1	130	184	0	0	0	0	210	178	172	187
CG9682	1369	1423	0	0	0	0	120	144	113	126
Ddx1	236	289	0	0	0	0	201	142	143	162
CG8317	87	140	0	0	0	0	227	4	295	175
CG31226	369	422	0	0	0	0	1907	667	227	934

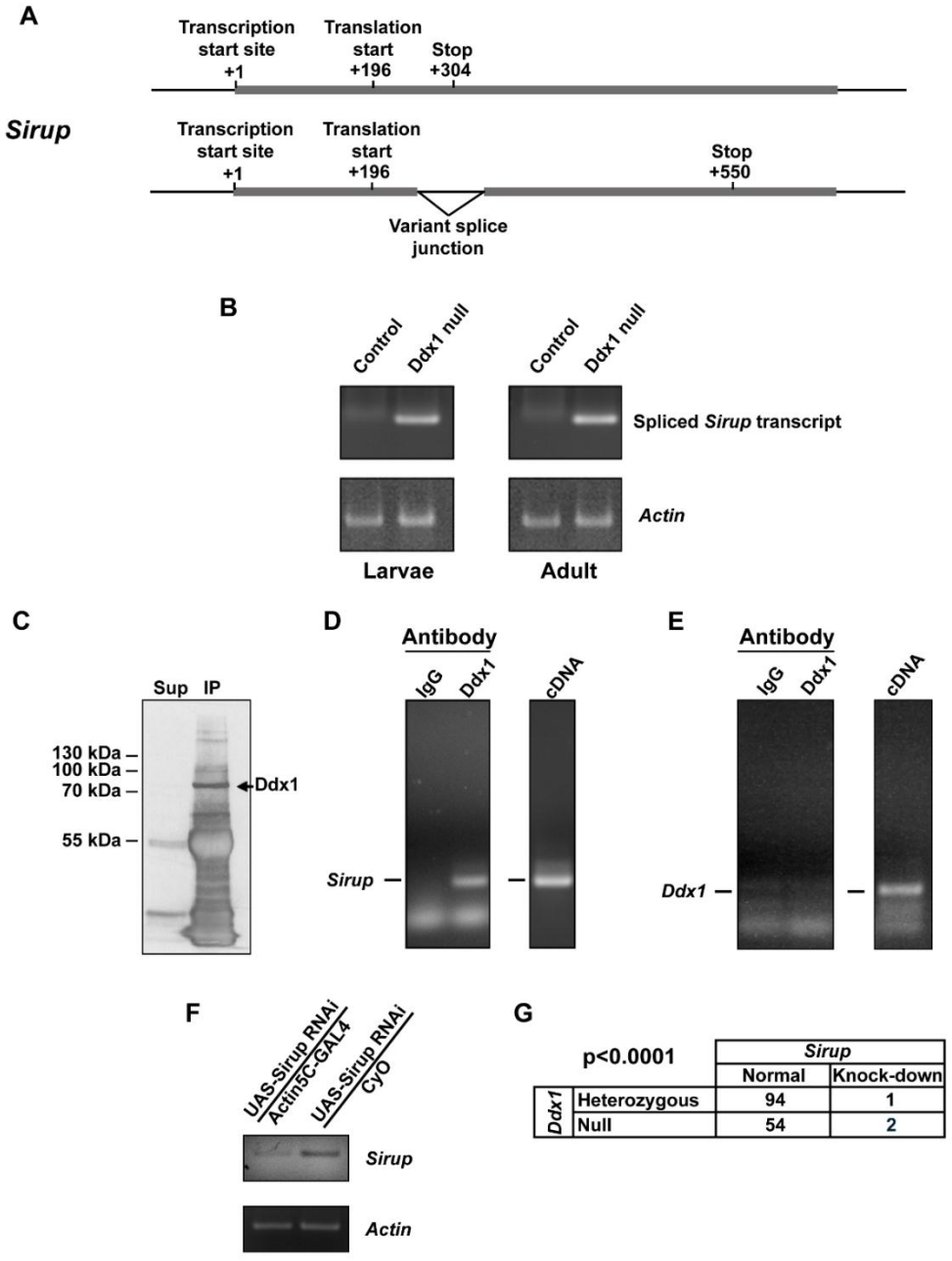
a number of splice variants uniquely found in either control or *Ddx1*$^{AX/AX}$ flies, including *Sirup* (Table 3.3 and 3.4).

3.3.6 Ddx1 interacts with Sirup mRNA

Sirup RNA has previously been shown to be up-regulated in response to starvation conditions (Erdi *et al.* 2012) and is required for proper mitochondrial function (Van Vranken *et al.* 2014). We also observed up-regulation of *Sirup* (by 3-fold) in *Ddx1*$^{AX/AX}$ versus control flies (Table 3.1). However, the biggest effect of *Ddx1*-null on *Sirup* appears to be in terms of splicing, with sequences spanning the 234-315 bp junction found uniquely and at high levels in *Ddx1*-null flies. *Sirup* has an unconventional gene structure, being expressed as a single exon or as an alternatively spliced product characterized by the removal of a small intron located downstream of the start codon (Figure 3.7A). Using a primer that spans the junction splice site, we confirmed that *Sirup* is indeed alternatively spliced in control and null flies, with *Ddx1* control larvae and adult flies expressing unspliced *Sirup*, and *Ddx1*$^{AX/AX}$ larvae and adults expressing spliced *Sirup* (Figure 3.7B).

The *Sirup* splice variant observed in control flies results in a longer transcript (810 nt) that retains the gene's single intron, but a shorter open reading frame (36 aa) due to a stop codon within the intron. In *Ddx1*$^{AX/AX}$ flies, the intron is removed, generating a shorter transcript (662 nt) but a longer open reading frame (118 aa). Intriguingly, this splice-specific regulation may be conserved in humans, as the human *Sirup* orthologue C6orf57 expresses two splice variants, one encoding a 108 aa protein and the other being non-coding, mirroring what we observe in

Figure 3.7 Interactions between *Ddx1* and *Sirup*. (A) Pictograph showing the structure of *Sirup* and the variable splice junction site (top – transcript expressed in control flies; bottom – transcript expressed in $Ddx1^{AX/AX}$ flies). (B) RT-PCR analysis of control and null larvae and adult flies. A spliced *Sirup* product is only observed in *Ddx1*-null animals. (C) IP using anti-Ddx1 antibody demonstrating that almost all Ddx1 protein is retained in the immunoprecipitate. (D, E) Left panels – RT-PCR analysis of RNA co-immunoprecipitated with Ddx1 or IgG. Right panels – RT-PCR of cDNA generated from control flies used as a positive control for the PCR reaction. *Sirup* mRNA results indicate that *Sirup* RNA, but not *Ddx1* RNA, is pulled down with Ddx1 protein. (F) RT-PCR analysis shows reduced *Sirup* RNA levels in *Sirup* knock-down adult flies. (G) Progeny generated from *Actin5C-Gal4/CyO;* $Ddx1^{AX}$*/TM3, Sb* x *UAS-Sirup-RNAi;* $Ddx1^{AX}$*/TM3, Ser GFP* crosses. Expected ratios of 2:1 for heterozygous to homozygous mutant *Ddx1* and 1:1 for *Sirup* knock-down to CyO. A significant reduction in the number of *Sirup* knock-down flies was observed. Chi square analysis comparing observed distribution to expected distribution was used to determine significance.



		Sirup	
		Normal	Knock-down
Ddx1	Heterozygous	94	1
	Null	54	2

Drosophila.

As Ddx1 has previously been shown to bind target RNAs, *Sirup* RNA may be a direct target of Ddx1. On the other hand alternative splicing of *Sirup* may represent an indirect consequence of homozygous mutant *Ddx1.* To distinguish between these possibilities, we performed RNA co-immunoprecipitations (co-IPs) using a Ddx1-specific antibody to immunoprecipitate DDX1-binding RNAs from Schneider2 cell lysates. Using this approach, we were able to efficiently co-precipitate Ddx1 protein (identified by western blot) and *Sirup* mRNA (identified by RT-PCR) but not *Ddx1* mRNA (Figures 3.7C-E). The latter control was included to ensure that the interaction between *Sirup* mRNA and Ddx1 protein was specific, and not the result of non-specific RNA interaction with a DEAD box protein.

Next, we tested a genetic interaction between *Ddx1* and *Sirup*. We generated flies that expressed a *Sirup* dsRNA construct using the UAS-GAL4 system (*Actin5C-GAL4* and *UAS-Sirup* RNAi [*P{GD14644}v36437*]) in *Ddx1* normal, *Ddx1* heterozygous and *Ddx1*-null flies. *Sirup* knock-down in a control background reduced *Sirup* RNA levels by approximately 75% (Figure 3.7F), and generated viable, fertile flies that displayed no observable phenotypes. *Sirup* knock-down in both *Ddx1* heterozygous and null backgrounds resulted in lethality in the majority of progeny. Although 50% of the progeny are expected to express the *Sirup* knock-down, we observed that the majority of *Ddx1* mutant progeny generated from these crosses did not contain the *Act5C-GAL4* transgene (3 containing both the *Actin5C-GAL4* driver and the *UAS-Sirup* transgene versus 148 containing only the *UAS-Sirup* transgene). These results indicate an epistatic lethal

effect between the mutant *Ddx1* locus and *Sirup* knock-down (Figure 3.7G). Of note, the two *Ddx1*-null/*Sirup* knock-down progeny survived for less than 1 day post eclosion, whereas the single *Ddx1* heterozygous/*Sirup* knock-down progeny survived 14 days post eclosion.

3.3.7 Identification of anti-sense Sirup RNA molecules

There are two plausible mechanisms for splicing regulation by an RNA helicase. Modification of secondary structures with an mRNA molecule could suppress or enhance splicing, or alternatively, there has been recent work describing a relationship between splicing and antisense RNA molecules (Morrissy *et al.* 2011). As Ddx1 has well established dsRNA unwinding activity we strove to determine if *Sirup* anti-sense RNA molecules are expressed (Li *et al.* 2008). We generated single stranded sense *Sirup* DNA and annealed total RNA from control 3rd instar larvae. A portion of the annealed product was subjected to duplex specific nuclease treatment (DSN), which selectively degrades DNA from DNA/DNA and DNA/RNA duplexes. DSN treatment results in the degradation of sections of the *Sirup sense* DNA that have been bound by complimentary anti-sense RNA. PCR amplification of sub-fragments of *Sirup* in non-treated samples generates a robust signal for each primer pair (Figure 3.8). Amplification of DSN-treated samples results in a relatively equivalent signal in four of the five sub-fragments. Notably, sub-fragment two, which spans the alternatively spliced region of *Sirup*, exhibited sensitivity to DSN treatment following RNA annealing, indicating the presence of an anti-sense RNA molecule within that region.

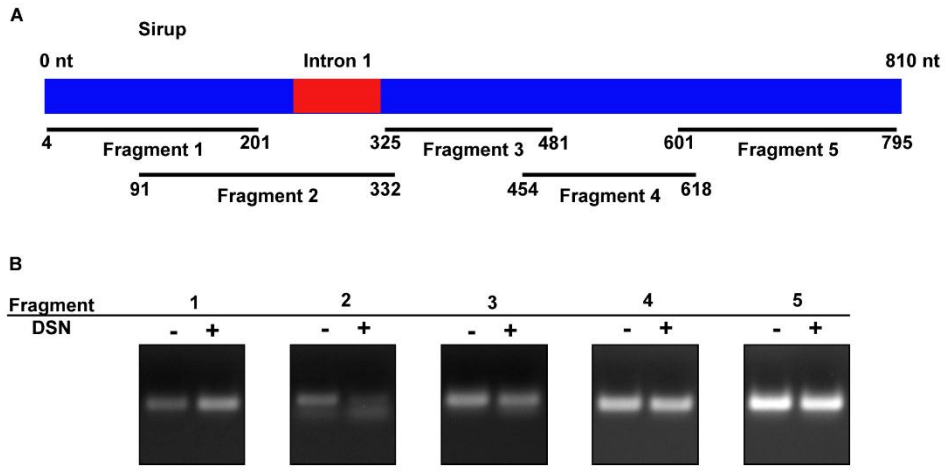


Figure 3.8 Identification of anti-sense *Sirup* RNA. (A) Pictograph indicating *Sirup* mRNA structure and the location of amplified fragments. (B) Single stranded sense *Sirup* DNA was incubated with total RNA from 3rd instar larvae. Left: no treatment, right: treatment with DSN. PCR amplification of sub-fragments was performed on both samples for each region. Sections of the *Sirup* DNA bound by anti-sense *Sirup* RNA will be degraded by DSN and a reduced signal will be observed. Fragment 2, which spans the single intron in *Sirup*, displays reduced signal following DSN treatment indicating that anti-sense RNA may bind to that region.

3.4 DISCUSSION

While human DDX1 has been implicated in numerous biological processes in cultured cells, we still have a poor understanding of its role in animal development. Towards this end, we generated a *Ddx1*-null mutant in *Drosophila*. In contrast to a report listing a previous *Ddx1* mutation as lethal (Zinsmaier *et al.* 1994), we found that *Ddx1*-null flies were viable. While this may suggest that Ddx1 is not crucial in the embryo, we detected significant levels of maternal Ddx1 protein up to the 2nd instar larval stage. As Ddx1 levels are highest in early embryogenesis (Rafti *et al.* 1996), these contributions may be sufficient to fulfill an essential need for Ddx1 in early development, postponing the appearance of *Ddx1* phenotypes until later stages in *Ddx1*$^{AX/AX}$ flies. We are limited in our ability to specifically address the role of maternally contributed Ddx1, as the location of the *Ddx1* gene locus is refractory to the generation of clonal populations, preventing the generation of mosaic females that would produce embryos with no maternal Ddx1 protein.

Additionally, the differing phenotypes between the two described *Ddx1* alleles may be the result of the nature of each mutation. Although the previously described *Ddx1* mutant allele was not characterized, the position of the P-element used to generate this mutant (downstream of the 3' end of *Ddx1)* suggests the possibility that a truncated Ddx1 protein may be produced. Expression of truncated proteins is tightly regulated in the cell as partial proteins often lead to dominant effects (Wagner *et al.* 2002). Aberrant activity of a truncated *Ddx1* protein, compared to animals where Ddx1 is lost completely, could explain the distinct

phenotypes observed in the two lines of Ddx1 mutations. Unfortunately, since the original stock is no longer available, this cannot be tested experimentally.

Although *Ddx1*-null flies survive to adulthood, it is clear that Ddx1 is involved in a wide spectrum of developmental functions. *Ddx1*-null flies are significantly reduced in size, developmentally delayed and out-competed by heterozygous animals when raised in crowded conditions. In addition, *Ddx1*-null flies are infertile due to aberrant gametogenesis, with both oogenesis and spermatogenesis being affected. Reduction in body size and autophagic egg chambers are phenotypes associated with metabolic stress and starvation conditions (Barth *et al.* 2011, Edgar 2006, McCall 2004). While our attempts to generate a transgenic Ddx1 over-expression line to rescue this phenotype were not successful, we were able to confirm comparable phenotypes in flies carrying the null $Ddx1^{AX}$ allele over a deficiency that encompasses the *Ddx1* locus. Thus, the phenotypes observed in *Ddx1*-null flies suggest a role for Ddx1 in normal metabolic function and/or regulation, with absence of Ddx1 phenocopying nutrient deficiency.

In keeping with a role for Ddx1 in metabolism, we found reduced pS6k levels in *Ddx1*-null flies. pS6k levels are indicative of TOR signalling, which is modified by a number of pathways including insulin signalling and nutrient sensing (Edgar 2006). Signalling through TOR and pS6k drives protein synthesis and cell growth (Zhang *et al.* 2000). Our RNA deep sequencing data indicate that neither *TOR* nor *S6k* mRNA levels are affected in $Ddx1^{AX/AX}$ flies. RNA levels of major upstream regulators of the TOR pathway, including *Rheb*, *Tsc1* and *gigas* (dTSC2) are also unchanged. The consistent levels of these key signalling factors suggest that the

absence of Ddx1 does not directly affect the transcripts encoding proteins that participate in TOR signalling. Rather, absence of Ddx1 may play an indirect role in the down-regulation of the TOR signalling pathway mediated through the function of mRNA targets.

In contrast to other DEAD box proteins that have been directly implicated in *Drosophila* oogenesis (Cauchi 2012, Johnstone *et al.* 2005, Styhler *et al.* 1998, Tomancak *et al.* 1998), we propose that the infertility observed in *Ddx1*-null females is due to disrupted metabolism. Altered metabolism also provides a possible explanation for the small number of progeny generated by *Ddx1*^{AX}*/Df(3L)ED230* females, as we observed a slightly higher pS6k level and slightly larger body size in these mutants compared to the completely sterile *Ddx1*$^{AX/AX}$ females. Thus, Ddx1^{AX}*/Df(3L)ED230* flies may have slightly higher metabolic function than *Ddx1*$^{AX/AX}$ flies. This may be due to second-site alleles that are different between the *Ddx1*^{AX} and *Df(3L)ED230* genetic background.

Spermatogenesis can also be affected by starvation conditions, with a noted reduction in the number of germinal stem cells in starved males (McLeod *et al.* 2010). In contrast to the phenotype observed in *Ddx1*-null ovaries, aberrant sperm individualization, the phenotype observed in *Ddx1*-null males, has not been associated with metabolic disruption. Following meiosis, spermiogenesis, the last stage of spermatogenesis, involves reshaping developing interconnected spermatids into individual mature sperm (Fabian *et al.* 2012). Spermiogenesis is characterized by gross morphological changes to the cell and mitochondria in particular. Developing spermatids in *Ddx1*-null males appear to undergo spermatid

elongation, but fail to individualize. Instead, they fall out of the developing sperm bundle prematurely and seminal vesicles remain devoid of mature sperm.

Several DEAD box proteins have been associated with defects in sperm development. RecQ5 and Belle are both required for early spermatogenesis (Johnstone *et al.* 2005, Sakurai *et al.* 2014) and mutation of *Rm62* has also been shown to cause male sterility (Buszczak *et al.* 2007). While the cause of sterility in *Rm62* mutants has not been identified, mutation of *Blanks*, a binding partner for Rm62, results in spermiogenesis defects similar to those observed in *Ddx1*-null males (Gerbasi *et al.* 2011). As disrupted sperm individualization has not been linked to metabolic function, we postulate that this aspect of the *Ddx1*-null phenotype is the result of disruption of distinct RNA targets from those involved in metabolic function.

It has been previously established that tRNA synthesis is regulated by TOR signalling (Ciesla *et al.* 2008) and human DDX1 has been recently identified as a tRNA splicing factor (Popow *et al.* 2011, Popow *et al.* 2014). As our analysis did not show changes in levels of spliced and unspliced tRNAs in *Ddx1*-null flies, it would appear that *Drosophila* Ddx1 is not essential for tRNA splicing. However, our northern blot analysis is only a snapshot of tRNA levels at the time of RNA isolation. Popow *et al.* showed that human DDX1 is necessary for efficient cycling of the RtcB-guanylate intermediate required for tRNA splicing *in vitro* (Popow *et al.* 2014). It is possible that tRNA generation efficiency is reduced in $Ddx1^{AX/AX}$ flies, but either another compensatory mechanism is coming into play, or the reduction in efficiency of tRNA splicing is not sufficient to reduce overall steady state levels.

In either case, it appears that though *Ddx1*-null flies have reduced pS6k signalling, they retain normal tRNA levels relative to whole RNA content.

Our RNA deep sequencing analysis revealed significant changes in the levels of 333 RNA molecules in $Ddx1^{AX/AX}$ 3rd instar larvae. In addition, we identified a number of transcripts that were differentially spliced in *Ddx1* control versus null flies. We verified that *Sirup* mRNA is uniquely spliced in $Ddx1^{AX/AX}$ larvae and up-regulated. We were also able to show that Ddx1 protein binds *Sirup* mRNA, which suggests that this splice site modification is due to a direct interaction with Ddx1, rather than being an indirect downstream consequence of the loss of *Ddx1.* Previous work has demonstrated that loss of *Sirup* results in shortened life span and neurodegeneration (Van Vranken *et al.* 2014). In agreement with our results, *Sirup* has also previously been shown to be up-regulated in response to starvation conditions (Erdi *et al.* 2012). We propose a role for Sirup in limiting metabolism during stress conditions. *Sirup* has recently been identified as the *Drosophila* homolog of yeast *Sdh8* (Van Vranken *et al.* 2014). Similar to its yeast homologue, *Drosophila* Sirup may be required to stabilize the succinate dehydrogenase holocomplex, and enhance succinate dehydrogenase activity.

We observed spliced *Sirup* mRNA only in the absence of Ddx1, and propose that Ddx1 is acting to repress splicing under normal conditions. This is in line with observations that human DDX1 is required for efficient trafficking of unspliced viral RNA genomes (Edgcomb *et al.* 2012, Fang *et al.* 2005, Robertson-Anderson *et al.* 2011). Recently, anti-sense RNA has been implicated in the regulation of alternative splicing (Morrissy *et al.* 2011). There are several possible mechanisms

for splicing repression by Ddx1, but the identification of anti-sense *Sirup* RNA molecules in conjunction with Ddx1's RNA/RNA unwinding activity suggest that Ddx1 is modulating splicing through the regulation of the interaction between *Sirup* mRNA and an anti-sense RNA. The location of this binding (within the region that is alternatively spliced) also supports this possibility. It is still not clear whether ncRNA binding promotes or represses specific splice sites, but it is possible that it is acting in either fashion, depending on where they bind to target RNAs.

Although the relationship between Ddx1-dependent splice regulation of *Sirup* mRNA and the metabolic disruption observed in *Ddx1*-null flies remains a matter of speculation, *Sirup*'s known role in mitochondrial function provides a possible explanation for the epistatic lethality observed in *Ddx1*-null/*Sirup* knock-down flies. We propose that Sirup's unspliced form is required for steady state metabolic function. The unspliced form of *Sirup* retains an early stop codon that produces a transcript that encodes a short 36 aa product missing conserved domains required for succinate dehydrogenase activity. However, there is evidence suggesting that this stop codon undergoes read-through by ribosomes, thereby generating a full length protein in spite of the stop codon (Dunn *et al.* 2013). Under stress conditions, modification of *Sirup* mRNA by Ddx1 would generate a shorter variant protein. Variant Sirup protein is predicted to be less efficient than normally expressed Sirup product, resulting in reduced succinate dehydrogenase activity, which in turn slows down mitochondrial function. The outcome of this scenario is reduced resource usage by the cell, allowing it to survive periods of nutrient scarcity. Under this model the observed epistatic lethality in *Ddx1*-

null/*Sirup* knock-down flies would be the result of a reduction in the amount of *Sirup* mRNA expressed in combination with only the less efficient form of Sirup protein being generated.

In conclusion, the phenotypes generated by a loss of *Ddx1* expression in Drosophila are in keeping with Ddx1 being a multifunctional protein involved in a variety of biological processes. Our work suggests a previously unreported role for Ddx1 in metabolism regulation that occurs through interaction and modification of transcripts such as *Sirup*, which can directly modulate metabolic activity.

3.5 ACKNOWLEDGMENTS

We are grateful to Dr. Xuejun Sun and Gerry Barron for their technical assistance with cell imaging. We thank Dr. Michael Hendzel and Dr. Zhigang Jin for helpful discussions. This study was supported by a grant from the Alberta Cancer Foundation.

CHAPTER 4

KNOCK-OUT OF DDX1 RESULTS IN TRANS-GENERATIONAL WILD-TYPE LETHALITY IN MICE

A version of Chapter 4 has been published as Hildebrandt MR*, Germain DR*, Monckton EA, Brun M, and Godbout R. Ddx1 knockout results in transgenerational wild-type lethality in mice. Sci. Rep. 5, 9829; DOI:10.1038/srep09829 (2015). *These authors contributed equally to this work. Mouse genotyping and collection was performed by Dr. Matthew Hildebrandt and Elizabeth Monckton. Southern blot analysis was performed by Elizabeth Monckton. Bisulfite sequencing was performed by Dr. Matthew Hildebrandt. Quantitative real-time PCR was performed by Miranda Brun. I was responsible for model conceptualization, experimental design, data analysis and interpretation, and writing the manuscript. Dr. Matthew Hildebrandt assisted with data analysis and writing. Dr. Roseline Godbout was involved in all stages of the project and in writing the manuscript. This work received research ethics approval from the Cross Cancer Institute Animal Care Committee, #BC11185.

4.1 INTRODUCTION

DEAD box proteins are RNA unwinding proteins that are characterized by 12 conserved motifs, including the signature motif, D(asp)-E(glu)-A(ala)-D(asp) which is involved in ATP hydrolysis. These proteins have been implicated in all aspects of RNA metabolism including transcription, transport, translation, and degradation (Fuller-Pace 1994, Jankowsky *et al.* 2007, Linder *et al.* 2013, Montpetit *et al.* 2012). Most DEAD box proteins unwind RNA-RNA duplexes *in vitro* through localized strand destabilization rather than processive unwinding (Del Campo *et al.* 2009, Yang *et al.* 2007). DEAD box proteins have been shown to be modulators of ribonucleoprotein complexes by displacing or recruiting different proteins to these complexes (Del Campo *et al.* 2009, Tanner *et al.* 2001).

DEAD box 1 (DDX1) was first identified by differential screening of a retinoblastoma cDNA library, and subsequently found to be amplified and overexpressed in a subset of retinoblastoma and neuroblastoma tumours and cell lines (George *et al.* 1996, Godbout *et al.* 1993b, Manohar *et al.* 1995, Squire *et al.* 1995, Weber *et al.* 2004). DDX1 expression is ubiquitous, with proliferating cells and cells derived from neuroectodermal tissues expressing the highest levels of DDX1 (Godbout *et al.* 2002, Godbout *et al.* 1993b). DDX1 is predominantly located in the nucleus of non-*DDX1*-amplified normal and cancer cells (Bleoo *et al.* 2001). However, when amplified and overexpressed, elevated levels of DDX1 are observed in both the nucleus and cytoplasm (Godbout *et al.* 2007). In breast cancer, DDX1 is a negative prognostic indicator when overexpressed or mis-localized to the cytoplasm (Balko *et al.* 2011, Germain *et al.* 2011).

DDX1 has been associated with a number of biological processes both in the nucleus and the cytoplasm. In the nucleus, DDX1 forms foci that co-localize with cleavage bodies and reside adjacent to Cajal bodies and gems, three spatially related RNA processing bodies (Bleoo *et al.* 2001, Li *et al.* 2006b). When cells are exposed to ionizing radiation, DDX1 is recruited to sites of DNA double-strand breaks where it co-localizes with DNA damage response proteins (Li *et al.* 2008). DDX1 is also part of the tRNA ligase complex involved in pre-tRNA processing, and the pri-miRNA microprocessor complex involved in the processing of miRNAs (Gregory *et al.* 2004, Han *et al.* 2014, Popow *et al.* 2011, Popow *et al.* 2014). In the cytoplasm, DDX1 is found in RNA containing granules involved in the transport of RNAs in neurons, as well as stress granules (Kanai *et al.* 2004, Miller *et al.* 2009, Onishi *et al.* 2008).

Although it is possible to knock-down DDX1 in immortalized cancer cell lines and normal fibroblast cultures (Li *et al.* 2008), to date there have been no reports of DDX1-null cell lines. Furthermore, our attempts to knockout *DDX1* in HeLa cells using CRISPR/Cas9 technology have been unsuccessful. In fact, whereas 2-3 rounds of CRISPR/Cas9 transfection resulted in a 50% reduction in DDX1 levels, repeated rounds of transfection (up to 10) generated cells with near normal levels of DDX1. Based on these data, it appears that HeLa cells have a compensatory mechanism in place to prevent long-term reduction in DDX1 levels.

We have generated *Ddx1* heterozygous mice that contain a constitutive gene-trapped allele. Here, we show that *Ddx1*$^{-/-}$ embryos die during the pre-blastocyst stage of development. Intriguingly, the ratio of wild-type to heterozygote

mice is significantly reduced in heterozygote intercrosses, with wild-type progeny dying between E3.5 and E6.5. By tracing parental lineages, we identified a subpopulation of heterozygous mice that generate significantly reduced numbers of wild-type progeny. This phenotype is observed in both FVB and C57BL/6 backgrounds, and is transmitted through both sexes. Analysis of the methylation status of the *Ddx1* gene revealed no differences between the heterozygous and wild-type mice.

4.2 MATERIALS AND METHODS

4.2.1 Generation and genotyping of Ddx1 mice

The mouse embryonic stem cell line (RRT447) containing an intronic gene trap within intron 14 of the *Ddx1* gene was purchased from BayGenomics. Chimeric *Ddx1* mice were generated by microinjecting RRT447 ES cells into C57BL/6 blastocysts. Male chimeric mice were mated to C57BL/6 females to obtain germ line transmission of the *Ddx1*$^{Gt(RRT447)RG}$ allele (abbreviated as *Ddx1*$^{-}$). Two independent lines were obtained and characterized. To confirm *Ddx1* gene disruption at exon 14 and to ensure that there was a single insertion site of the *β-geo* reporter gene in our two lines, Southern blot analyses were carried out using ^{32}P-labeled *β-geo* or *Ddx1* (exons 10-18) cDNAs. The *Ddx1* probe was generated by restriction endonuclease digestion of *Ddx1* cDNA with *Eco*RI and *Hind*III. The *β-geo* probe was generated with *β-geo* specific primers (5': 5'-TTATCGATGAGCGTGGTGGTTATGC-3' paired with 3': 5'-GCGCGTACAT CGGGCAAATAATATC-3').

All animal work was approved by the Cross Cancer Institute Animal Care Committee (protocol BC11185) and followed the guidelines established by both the Cross Cancer Institute and University of Alberta Animal Care Committees. To generate timed pregnancies, female mice were naturally mated to males. Females were examined for the presence of vaginal plugs over the course of 10 days. Mice with a plug were deemed to be at gestational stage E0.5. Plugged females were sacrificed at E3.5 and 6.5-10.5 to isolate embryos, which were subjected to genotyping by PCR as described below**.**

Genomic DNA was extracted from ear punches of weaned mice using the E.Z.N.A Tissue DNA Kit (Omega) according to the manufacturer's instructions. Genomic DNA was collected from tails of P1 mice or from whole E6-10 embryos by digesting the tissue overnight in 100 µl Tris-EDTA-NaCl (TEN) buffer containing 40 µg/ml proteinase K (PK). The following day, genomic DNA was extracted using phenol/chloroform and precipitated with ethanol. E3.5 embryos were collected in 20 µl PCR buffer supplemented with 40 µg/ml PK. The embryos were digested at 55˚C for 1 hour followed by 10 minutes at 90˚C to inactivate PK.

Genotypes of E6 and older mice were determined by multiplex PCR in a 20 µl reaction volume containing 1 µl DNA template, 2 µl 10X PCR buffer (GE Healthcare), 0.4 µM of each primer (RGo60: 5'-CTGGGGTTCGTGTCCTACAA-3', RGo63: 5'-ATTAGGAACTGGGCATGTATC-3', and RGo65: 5'-AGCACTAG TAAGTACCTACAC-3'), 250 µM dNTP mix and 0.2 µl Taq polymerase. The reaction was PCR-amplified under the following conditions: 94˚C for 5 minutes followed by 35 cycles at 94˚C for 1 minute, 60˚C for 1 minute and 72˚C for 1 minute followed by a final extension at 72˚C for 10 minutes. The reaction mixture was separated on a 1.0% agarose gel in 1X Tris acetate-EDTA buffer.

Genotypes of blastocysts were analyzed by nested PCR. For the first round, we used 1 µl DNA template, 2 µl 10X PCR buffer, 0.8 µM of the following primers: RGo62: 5'-GATGGAGACAGTCCTGGTT-3' paired with RGo66: 5'-CCAAGCTCCACTATTATCCC-3' or RGo62 paired with RGo60, 250 µM dNTP mix and 0.2 µl Taq polymerase using the same amplification protocol described above. For the second round, we used 1 µl from the first round reaction, 2 µl 10X

PCR buffer, 0.4 µM primers (RGo63/RGo65 for the RGo62/66 template or RGo63/60 for the RGo62/60 template), 250 µM dNTP mix and 0.2 µl Taq polymerase using the same amplification protocol described above.

4.2.2 Statistical analysis

Expected groups were defined by applying the normal genotype ratio to the total number of progeny collected at each stage of development. Individual Fisher's exact tests were performed between each expected group and the observed values to determine significant differences between the two groups.

4.2.3 Western blot analysis

Protein was isolated from P1 brain tissue that had been previously flash frozen and stored at -80˚C. Chilled lysis buffer (PBS containing 1% TX-100, 0.1% SDS, 1X Complete (Roche), 1 mM PMSF, and 1 mM DTT) was added to each sample. The samples were homogenized and centrifuged at 14,000 g for 10 minutes at 4˚C before collecting the supernatant. Cell lysates (50 µg per lane) were electrophoresed in an 8% SDS-polyacrylamide gel. The proteins were transferred to PVDF membranes. Membranes were blocked with 10% milk in TBST (0.01% Tween-20) for 1 hour, then sequentially immunostained with anti-DDX1 (batch 2910; 1:5,000 dilution) and anti-actin (Sigma; 1:100,000 dilution) in 5% milk in TBST at 4˚C overnight. The blots were subjected to anti-rabbit (for DDX1) and anti-mouse (for actin) secondary antibodies conjugated to HRP (Molecular Probes;

1:50,000 dilution) in 5% milk in TBST for 4 hours, followed by incubation with ECL reagent (GE) and exposure to X-ray film.

4.2.4 Semi-quantitative RT-PCR

RNA was isolated from P0-3 mouse brains by homogenization in 1 ml Trizol (Life Technologies) as per the manufacturer's protocol. Complementary DNA (cDNA) was generated using Superscript II (Life Technologies) following the manufacturer's protocol using either oligo(dT)$_{12\text{-}18}$ or random hexamer primers and 5 µg RNA. Semi-quantitative RT-PCR was performed in a 20 µl reaction containing 1 µl cDNA, 2 µl 10X PCR buffer (GE Healthcare), 0.4 µM of each primer pair (3' *Ddx1*: sense, 5'-AGAATTATGTGCACCGGATC-3', antisense, 5'-GCACCAGAGG GTTAGAGT-3'; *β-geo*: sense, 5'-CCTGTCCGGTGCCCTGAATG-3', antisense, 5'-GAAGAACTCGTCAAGAAGGCG-3'; *β-geo-Ddx1* fusion:, sense, 5'-CTGAAG AGCTTGGCGGCGAAT-3', antisense, 5'-TTTGGATCCATGTACATCATCAGTT CTAAT-3'; *Gapdh*: sense, 5'-ACGGCAAATTCAACGGCAC-3', antisense, 5'-GA GAGCAATGCCAGCCCC-3'), 250 µM dNTP mix and 0.2 µl Taq polymerase. The reaction was amplified using the following conditions: an initial heating to 94˚C for 5 minutes followed by 25 cycles (*Gapdh*) or 29 cycles (*Ddx1* or *β-geo*, or *β-geo-Ddx1 fusion*) of 94˚C for 1 minute, 55˚C for 30 seconds and 72˚C for 1 minute followed by a final extension for 10 minutes at 72˚C and a hold at 4˚C. The reactions were electrophoresed in a 1% agarose gel to separate the amplified DNA.

4.2.5 Quantitative real-time PCR

Total RNA was isolated from P0-3 brain and first-strand cDNA synthesized as above. The cDNA was amplified using TaqMan Fast Universal PCR Master Mix and gene-specific oligonucleotides (*Ddx1*, Mm01270541_m1; *Gapdh*, Mm99999915_g1) labeled at the 5′ end with the fluorescent reporter dye FAM (Life Technologies) (ABI 7900HT Fast Real-Time PCR System). The *Ddx1* oligonucleotide is 3' to the LacZ insert. All cDNAs were run in triplicate, and the data were normalized using *Gapdh*.

4.2.6 Bisulfite sequencing

1 µg genomic DNA prepared from *Ddx1*$^{+/+}$, *Ddx1*$^{+/-}$ and *Ddx1*$^{*/-}$ mice was treated with sodium bisulfite using the EpiTect Bisulfite kit (Qiagen) using the manufacturer's protocol with an additional cycle of denaturation for 5 minutes at 95^{o}C followed by 2 hours at 60^{o}C to ensure complete conversion. The converted DNA was amplified using 1 µl template, 10X PCR buffer (GE), 0.4 µM of each primer (sense, 5'-AAGTTTATAGGTTTTGAGTGAATTATT-3', antisense, 5'-CCAA ACAAAACAACATCATCTTTAC-3'), 250 µM dNTP mix and 1 µl Taq polymerase in a 100 µl volume. The PCR reaction was electrophoresed in a 6% native acrylamide gel. The expected 700 bp band was cut out and electroeluted on dialysis tubing. The DNA was extracted with phenol and ethanol-precipitated. The purified DNA was ligated into the pGEM-T Easy (Promega) vector using the manufacturer's protocol with overnight ligation at 16˚C. *E. coli* DH5α competent cells were transformed with the ligated products and colonies selected by

blue/white color selection. White colonies were selected for analysis and plasmid DNA purified using the QiaPrep Spin Mini plasmid kit (Qiagen) (Sambrook *et al.* 1989). Plasmid DNA containing inserts were sequenced using the M13 reverse sequencing primer (5'-CAGGAAACAGCTATGAC-3'). DNA sequences were then subjected to analysis by Bisulfite Sequencing DNA Methylation Analysis (BISMA) using default parameters and displayed using Methylation plotter (Mallona *et al.* 2014, Rohde *et al.* 2010).

4.3 RESULTS

4.3.1 Ddx1$^{-/-}$ embryos die pre-implantation

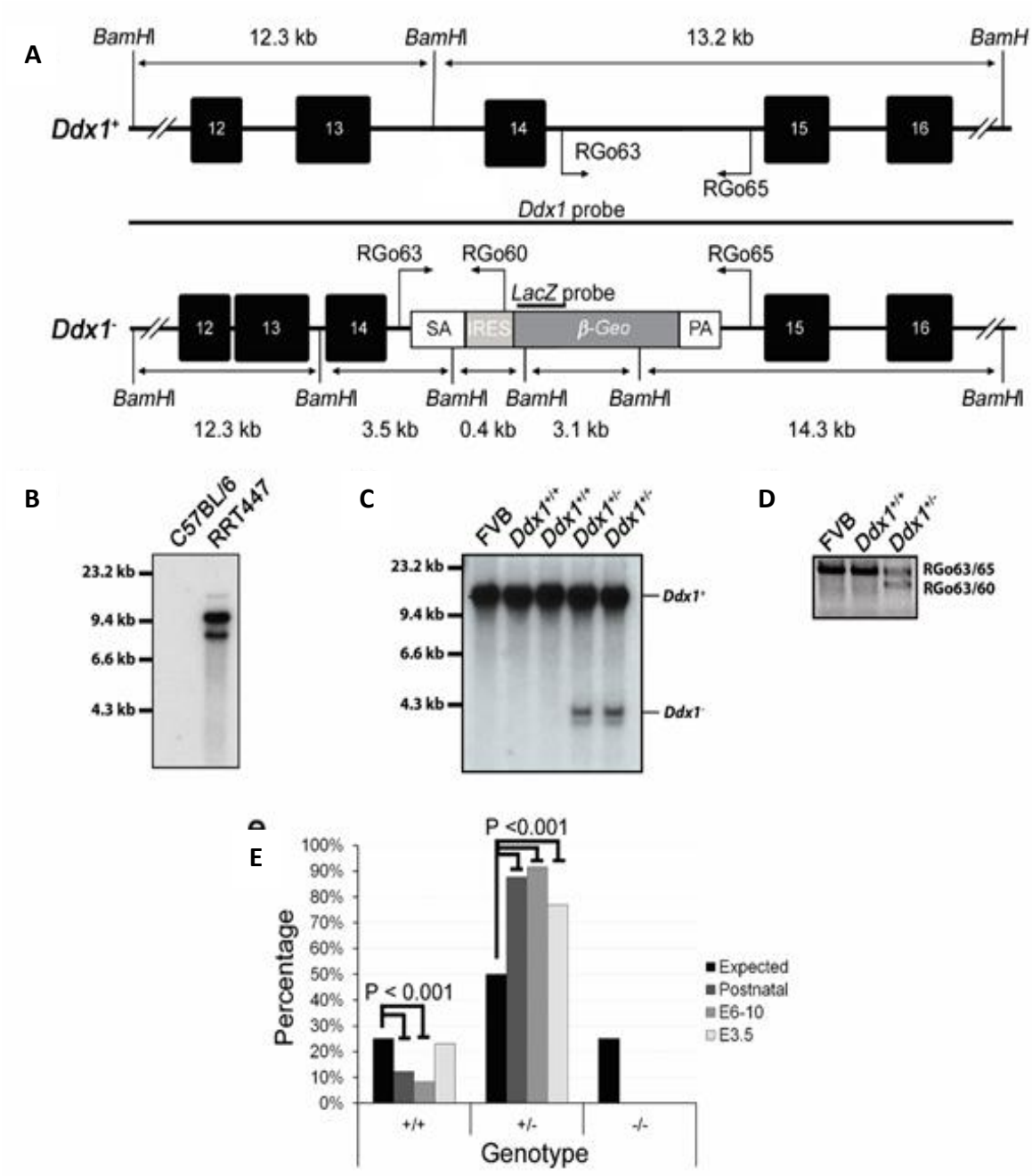
A mouse embryonic stem cell (ESC) line containing an intronic gene trap in the *Ddx1* gene [*GT*(RRT447)Byg; abbreviated as RRT447] was obtained from BayGenomics. Insertion of the gene-trap was in intron 14 of *Ddx1*. RRT447 ES cells were microinjected into blastocysts to generate male chimeras that were mated to C57BL/6 females to obtain germ-line transmission of the *Ddx1*$^{GT(RRT447)Byg}$ allele, designated *Ddx1*$^{-}$ (Figure 4.1A). Southern blot analysis with cDNA probes to *β-geo* and *Ddx1* exons 10-17 showed the presence of a single gene-trap in the RRT447 ESCs and in the *Ddx1*$^{+/-}$ mice generated from the ESCs (Figures 4.1B-C). *Ddx1*$^{+/-}$ mice showed no phenotypic abnormalities and produced phenotypically normal pups.

Offspring produced by heterozygote intercrosses were genotyped by PCR to identify both wild-type and gene-trap *Ddx1* alleles. Of 408 weaned pups analyzed, no *Ddx1*$^{-/-}$ pups were identified (Figure 4.1E). Next, we genotyped embryos from both pre-implantation (E3.5) and post-implantation (E10) stages. Again, no *Ddx1*$^{-/-}$ embryos were identified at either stage, indicating that *Ddx1*$^{-/-}$ embryos die pre-implantation.

4.3.2 Two distinct populations of Ddx1$^{+/-}$ mice produce differing ratios of wild-type to heterozygous progeny

The expected ratio of wild-type to heterozygote animals in *Ddx1* heterozygous intercrosses is 1 wild-type to 2 heterozygotes, as no *Ddx1* knock

Figure 4.1 Genomic map of the gene-trap insertion site. ESCs containing a single gene-trap insertion in *Ddx1* were purchased from BayGenomics. (A) The insertion containing a *β-geo* gene, splice acceptor (SA) and a polyadenylation signal (PA) is located between exons 14 and 15 of *Ddx1*. The insertion generates a truncated DDX1 protein fused to LacZ. Locations of primers (RGo) and Southern blot probes used for genotyping are also shown. (B) Southern blot analysis of RRT447 cell line using a ^{32}P-labeled cDNA probe specific to *β*-geo. (C) Southern blot analysis of wild-type and *Ddx1*$^{+/-}$ mice using a ^{32}P-labeled cDNA probe specific to *Ddx1*. (D) PCR amplification of genomic DNA for routine genotyping using primers shown in (A). (E) Progeny from heterozygous intercrosses (*Ddx1*$^{+/-}$ or *Ddx1*$^{+/-}$) were collected and genotyped at different developmental stages. No *Ddx1* $^{-/-}$ progeny were observed out of a total of 758 postnatal offspring, 225 E6-10 embryos and 91 E9.5 blastocysts genotyped. A significant decrease in the percentage of wild-type mice was observed post E3.5 ($P < 0.001$). Fisher's exact tests were performed to determine significant differences between the expected and observed ratios of *Ddx1*$^{+/+}$ to *Ddx1*$^{+/-}$ mice.


A
BamHI
12.3 kb
BamHI
13.2 kb
BamH
Ddx1+
12
13
14
15
16
RGo63
RGo65
Ddx1 probe
RGo63
RGo60
RGo65
LacZ probe
Ddx1-
SA
IRES
β-Geo
PA
BamHI
BamHI
BamHI
BamHI
BamHI
BamHI
12.3 kb
3.5 kb
0.4 kb
3.1 kb
14.3 kb
B
C57BL/6
RRT447
23.2 kb
9.4 kb
6.6 kb
4.3 kb
C
FVB
Ddx1+/+
Ddx1+/+
Ddx1+/-
Ddx1+/-
23.2 kb
9.4 kb
6.6 kb
4.3 kb
Ddx1+
Ddx1-
D
FVB
Ddx1+/+
Ddx1+/-
RGo63/65
RGo63/60
E
P <0.001
P < 0.001
Percentage
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
+/+
+/-
-/-
Genotype
Expected
Postnatal
E6-10
E3.5

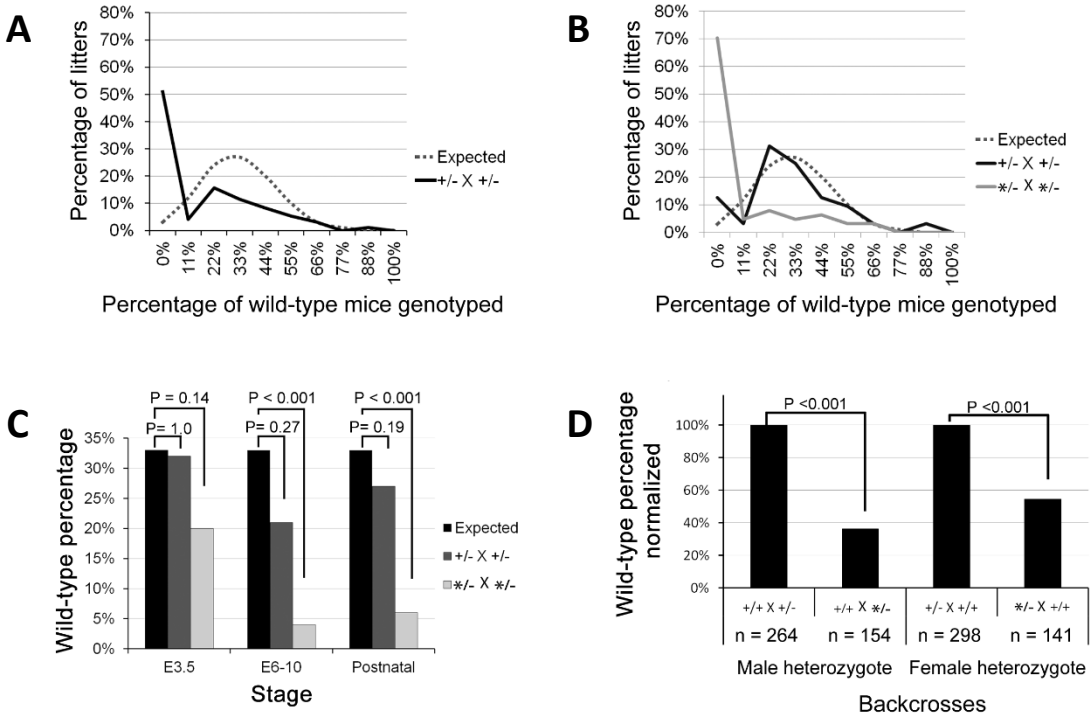
-outs survive to E3.5. Intriguingly, analysis of all heterozygote matings revealed a considerable deviation from the expected 1:2 ratio, with an observed ratio of 1:9 (Table 4.1). To rule out the possibility of a recessive lethal mutation linked to the wild-type *Ddx1* allele in the C57BL/6 background, *Ddx1*$^{+/-}$ mice were backcrossed for six generations to wild-type FVB mice. When FVB *Ddx1*$^{+/-}$ mice were intercrossed, we obtained a similar genotype ratio as in the C57BL/6 background (Table 4.1). In total, 292 weaned pups in the FVB background were genotyped by PCR analysis, with an observed ratio of 1 wild-type to 7 heterozygous mice. As both the FVB and C57BL/6 strains generated a wild-type lethality phenotype, subsequent analyses were carried out using both the FVB and C57BL/6 *Ddx1* lines.

Analysis of *Ddx1*$^{+/+}$ *to Ddx1*$^{+/-}$ progeny ratios in individual litters of *Ddx1* heterozygous intercrosses revealed a bimodal distribution, suggesting the possibility of two distinct heterozygote populations (Figure 4.2A). Upon more detailed examination of individual litters, we discovered that a normal ratio of wild-type to heterozygous progeny was consistently observed when *Ddx1*$^{+/-}$ animals were derived from *Ddx1*$^{+/+}$ X *Ddx1*$^{+/-}$ backcrosses (Figure 4.2B). In contrast, *Ddx1*$^{+/-}$ animals derived from *Ddx1*$^{+/-}$ X *Ddx1*$^{+/-}$ intercrosses generated significantly fewer wild-type progeny. To distinguish the two *Ddx1*$^{+/-}$ populations, we designated the *Ddx1*$^{+}$ allele inherited from *Ddx1*$^{+/-}$ X *Ddx1*$^{+/-}$ intercrosses as *Ddx1** and heterozygous mice derived from these crosses as *Ddx1**$^{/-}$.

Strain	Total	Genotype by PCR		
		+/+	+/-	-/-
C57BL/6/*Ddx1*$^{+/-}$	408	42 (10%)	366 (90%)	0
FVB/*Ddx1*$^{+/-}$	292	34 (12%)	258 (88%)	0

Table 4.1 Genotypes of weaned progeny from heterozygous matings.

Figure 4.2 Heterozygous mice generate a bimodal distribution of progeny genotypes. (A) Litters from heterozygous intercrosses (*Ddx1*$^{+/-}$ or *Ddx1*$^{+/-}$) that contained at least 5 pups were plotted as a percentage of wild-type mice generated (n = 178). A normal random distribution plotted around the expected value of 33% wild-type is included for comparison. (B) *Ddx1*$^{+/-}$ and *Ddx1*$^{*/-}$ intercrosses were separated (n = 32 and n = 146, respectively) and the percentage of wild-type mice generated was plotted. (C) The percentage of wild-type mice at ages E3.5, E6-10 and P0 from *Ddx1*$^{+/-}$ (n = 22, 61, 229, respectively) or *Ddx1*$^{*/-}$ (n = 69, 164, 529, respectively) intercrosses were plotted against the expected percentage. (D) Backcrosses (wild-type X heterozygote) from both FVB and C57BL/6 (combined) mice were separated by genotype and gender of the heterozygote. The percentage of wild-type genotyped was normalized to the *Ddx1*$^{+/-}$ backcross. Fisher's exact tests were performed to determine significant differences.


A
Percentage of litters
80%
70%
60%
50%
40%
30%
20%
10%
0%
0%
11%
22%
33%
44%
55%
66%
77%
88%
100%
Expected
+/- X +/-
Percentage of wild-type mice genotyped
B
Percentage of litters
Expected
+/- X +/-
*/- X */-
Percentage of wild-type mice genotyped
C
Wild-type percentage
35%
30%
25%
20%
15%
10%
5%
0%
P = 0.14
P= 1.0
P < 0.001
P= 0.27
P < 0.001
P= 0.19
E3.5
E6-10
Postnatal
Stage
Expected
+/- X +/-
*/- X */-
D
Wild-type percentage normalized
100%
80%
60%
40%
20%
0%
P <0.001
P <0.001
+/+ X +/-
+/+ X */-
+/- X +/+
*/- X +/+
n = 264
n = 154
n = 298
n = 141
Male heterozygote
Female heterozygote
Backcrosses

4.3.3 Ddx1-associated lethality occurs between E3.5 and E6.5*

To further characterize *Ddx1**-associated lethality, we carried out heterozygote intercrosses using: (i) heterozygote male and female mice generated from *Ddx1*$^{+/+}$ X *Ddx1*$^{+/-}$ matings (Ddx1$^{+/-}$), and (ii) heterozygote male and female mice generated from *Ddx1*$^{+/-}$ X *Ddx1*$^{+/-}$ matings (Ddx1$^{*/-}$). Genotyping the progeny of heterozygote intercrosses at different stages of development revealed reduced numbers of *Ddx1*$^{*/*}$ progeny at E6.5 and later (Figure 4.2C). Only ~5% of the progeny generated at E6.5 in *Ddx1*$^{*/-}$ intercrosses were wild-type (*Ddx1*$^{*/*}$). As no further reduction in wild-type (*Ddx1*$^{*/*}$) progeny numbers were observed after E6.5, we conclude that the lethality observed in *Ddx1*$^{*/*}$ embryos is occurring pre-E6.5.

To further define when *Ddx1*$^{*/*}$ mice die, E3.5 blastocysts were genotyped. Ratios of both *Ddx1*$^{+/+}$ to *Ddx1*$^{+/-}$ and *Ddx1*$^{*/*}$ to *Ddx1*$^{*/-}$ were normal at E3.5, suggesting that *Ddx1*$^{*/*}$ lethality occurs during the post-blastocyst stages of development. The most likely causes for the observed lethality are therefore failure to implant or failure to continue development post-implantation. As no reabsorbed embryos were observed in *Ddx1*$^{*/-}$ intercrosses, lethality is likely due to a failure to implant.

In the heterozygote intercrosses described above, wild-type lethality was observed in *Ddx1*$^{*/*}$ progeny. In order to address whether a single *Ddx1** allele can give rise to lethality, both *Ddx1*$^{+/-}$ and *Ddx1*$^{*/-}$ mice were backcrossed to *Ddx1*$^{+/+}$ mice. As expected, *Ddx1*$^{+/-}$ backcrosses (producing *Ddx1*$^{+/+}$ and *Ddx1*$^{+/-}$ offspring) yielded the expected number of wild-type progeny. In contrast, *Ddx1*$^{*/-}$ backcrosses (producing *Ddx1*$^{*/+}$ and *Ddx1*$^{+/-}$ offspring) yielded approximately 40%

of the expected number of wild-type mice, indicating reduced viability in *Ddx1*$^{*/+}$ animals.

4.3.4 Inheritance of the Ddx1 allele is parental sex independent*

As *Ddx1*$^{+/-}$ mice can be generated from *Ddx1*$^{*/-}$ x *Ddx1*$^{+/+}$ crosses, we can infer that the modification responsible for the observed lethality must be linked with the specific *Ddx1* allele, rather than transmitted in *trans*. In addition, since inheritance of the modified *Ddx1** allele occurs in both heterozygous (*Ddx1*$^{*/-}$) and homozygous wild-type (*Ddx1*$^{*/*}$) progeny, the allele must be transgenerationally maintained. The most common form of epigenetic transgenerational modification in mice is genomic imprinting. Genomic imprinting involves methylation based silencing that occurs during gamete formation, and is sex asymmetric.

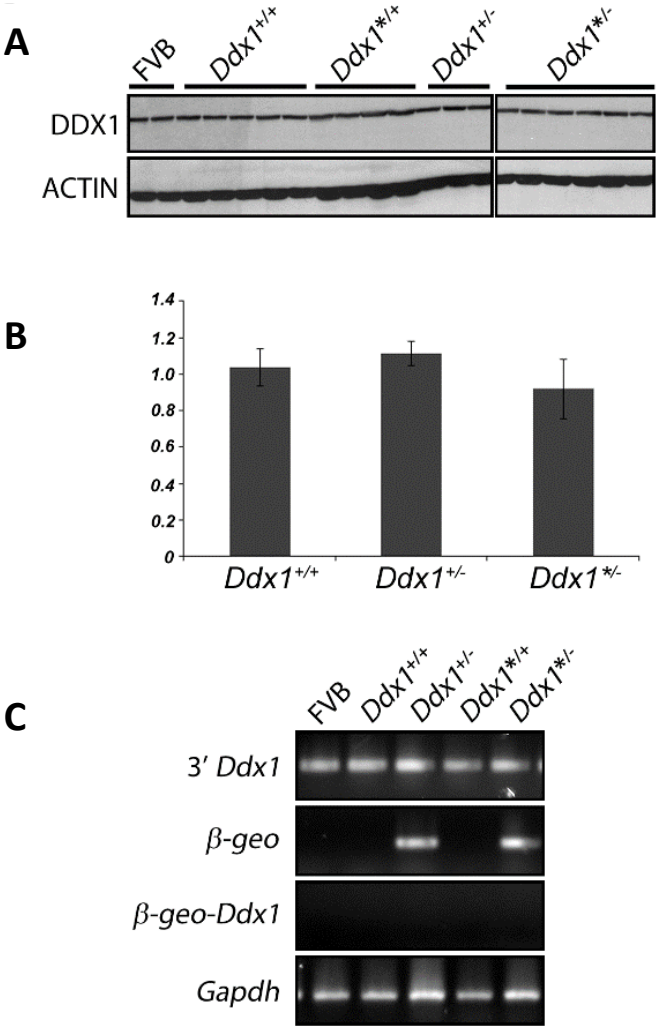
While *Ddx1* has not previously been reported to undergo genomic imprinting, the observed lethality could be explained by an imprinting mechanism of inheritance. In order to determine if genomic imprinting is responsible for modulating *Ddx1*, progeny generated from heterozygote female or male backcrosses were analyzed. Analysis of 295 progeny from *Ddx1*$^{*/-}$ backcrosses (154 offspring from male *Ddx1*$^{*/-}$ mice and 141 from female *Ddx1*$^{*/-}$ mice) revealed altered wild-type to heterozygote ratios in progeny generated from both male and female *Ddx1*$^{*/-}$ mice (Figure 4.2D). The lack of a sex-specific effect indicates that traditional genomic imprinting is not responsible for the observed lethality.

4.3.5 Expression compensation at the Ddx1 locus

Western blot analysis of brain tissues using anti-DDX1 antibody showed similar levels of DDX1 in all progeny irrespective of genotype (Figure 4.3A). The absence of truncated DDX1 products in heterozygote mouse brain using an antibody prepared against the N-terminus of DDX1 suggests that stable DDX1 protein is not produced from the *Ddx1* gene-trap allele. In agreement with western blot data, qPCR analysis of $Ddx1^{+/+}$, $Ddx1^{+/-}$ and $Ddx1^{*/-}$ mice showed similar *Ddx1* mRNA levels (Figure 4.3B).

To further address expression from the wild-type *Ddx1* and gene-trap alleles, we carried out RT-PCR using mouse brain RNA isolated from each genotypic group ($Ddx1^{+/+}$, $Ddx1^{+/-}$, $Ddx1^{*/*}$, $Ddx1^{*/-}$, and FVB control). PCR amplification of *Ddx1* transcripts (exons 22-26) generated a positive signal for *Ddx1* in all samples, and β-*geo* transcripts were detected in all samples containing the gene-trapped *Ddx1* allele. These results indicate that *Ddx1* is bi-allelically expressed in heterozygous mice (Figure 4.3C). RT-PCR analysis using a 5' primer specific to the gene trap and a 3' primer specific to *Ddx1* (exon 21), downstream of the gene trap region, failed to produce a signal, indicating that the gene trap transcript is not being spliced into the downstream region of *Ddx1* (Figure 4.3C). These results indicate compensation from the one functional *Ddx1* allele in heterozygous mice, resulting in similar levels of DDX1 in heterozygous and wild-type mouse brain. Similar results were obtained in liver (data not shown).

Figure 4.3 *Ddx1* mRNA and protein expression levels are similar in wild-type and heterozygous animals. (A) Western blot analysis of 50 µg of whole brain lysates from P0-3 mice of the indicated genotypes. Blots were immunostained with anti-DDX1 antibody (top) and anti-actin antibody (bottom). (B) Quantitative real-time PCR of P0-3 mouse brain from the indicated genotypes. qPCR was carried out with 3' *Ddx1* primers and *Gapdh* primers as a control ($n \geq 4$ for each sample). Expression levels for *Ddx1* are plotted relative to wild-type. Error bars show standard error of the mean. (C) Semi-quantitative RT-PCR analysis of cDNAs generated from P0-3 mouse brain RNA. cDNA samples were amplified with *Ddx1* primers 3' to the gene-trap (top panel), primers specific to *β-geo* (second panel), primers to the 3' end of *β-geo* and the 3' region of *Ddx1* (third panel), and primers to *Gapdh* as a control (bottom panel).


A
FVB
Ddx1+/+
Ddx1*/+
Ddx1+/-
Ddx1*/-
DDX1
ACTIN
B
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
Ddx1+/+
Ddx1+/-
Ddx1*/-
C
FVB
Ddx1+/+
Ddx1+/-
Ddx1*/+
Ddx1*/-
3′ Ddx1
β-geo
β-geo-Ddx1
Gapdh

4.3.6 DNA methylation is not altered in the Ddx1 allele*

Ddx1 compensation in heterozygous mice likely arises from changes in gene transcription as *Ddx1* RNA levels are similar in wild-type and heterozygous mice. While we previously showed that genomic imprinting is not likely to be responsible for the phenotypes observed, it remains possible that DNA methylation is the mechanism by which the *Ddx1** allele is modified. CpG methylation of promoter regions is commonly associated with alterations in gene expression. Low levels of transcription are generally associated with increased methylation. Importantly, altered methylation patterns can potentially be inherited, leading to the observed transgenerational nature of genomic imprinting.

Using MethPrimer prediction software we identified a single CpG island in the *Ddx1* gene (Li *et al.* 2002). This CpG island contains 55 CpGs and flanks the *Ddx1* transcriptional start site from –156 to +487 bp (Figure 4.4A). Using bisulfite conversion of genomic DNA followed by DNA sequencing, we analyzed DNA methylation patterns in $Ddx1^{+/+}$, $Ddx1^{+/-}$ and $Ddx1^{*/-}$ mice. At least 4 clones from each group were sequenced. No differences in methylation patterns were observed between the three different groups indicating that DNA methylation is likely not the mechanism regulating *Ddx1* gene compensation or *Ddx1** allele modification (Figure 4.4B).

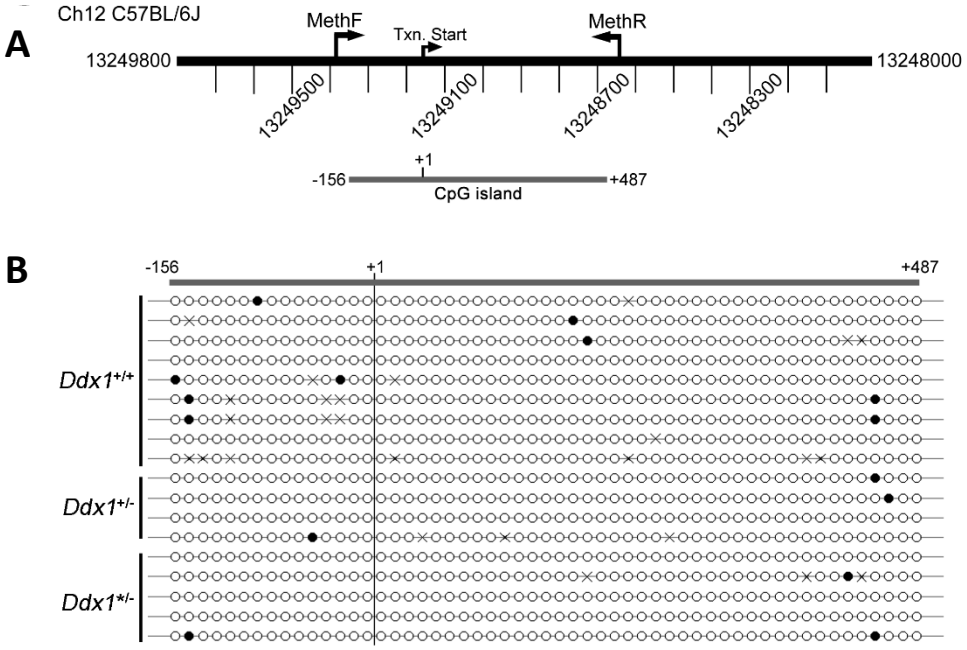


Figure 4.4 Methylation analysis at the *Ddx1* transcription start site. (A) A CpG island consisting of 55 CpG sites was predicted flanking the transcription (txn) start site of *Ddx1* from -156 to +487. MethF and MethR indicate binding sites of primers used to amplify the region following bisulfite conversion. (B) A lollipop diagram shows the methylation status of each of the 55 CpGs, where a white circle indicates no methylation and a black circle indicates methylation. At least 4 clones from each genotype (*Ddx1*$^{+/+}$, *Ddx1*$^{+/-}$, and *Ddx1*$^{*/-}$) were analyzed for their methylation patterns. A cross indicates indeterminate methylation.

4.4 DISCUSSION

Germ-line knockout of a number of DEAD box genes, including *Ddx11*, *Ddx20*, and *Ddx58,* results in embryonic lethality in mice (Inoue *et al.* 2007, Kato *et al.* 2005, Mouillet *et al.* 2008). Other DEAD box gene knockout mice are viable but have defects in gametogenesis; e.g., germ-line knockout of *Ddx4* (*Vasa*) and *Ddx25* both result in spermatid maturation defects (Tanaka *et al.* 2000, Tsai-Morris *et al.* 2004b). The earliest stage lethality upon knockout of a DEAD box gene was observed in *Ddx20* (*DP103, Gemin)* knockout mice. *Ddx20*$^{-/-}$ mice die at the 2-cell stage when zygotic gene expression is activated after rapid degradation of maternal RNAs (referred to as maternal to zygote transition or MZT). DDX20 is up-regulated in the 2-cell stage embryo and has been postulated to be involved in the reprogramming that occurs during maternal to zygote transition (Mouillet *et al.* 2008, Zeng *et al.* 2004).

Ddx1$^{-/-}$ mice die pre-E3.5 suggesting an essential role for DDX1 in early embryonic development. In light of DDX1's demonstrated roles in RNA binding, RNA/RNA unwinding and RNA transport (Kanai *et al.* 2004, Li *et al.* 2008, Miller *et al.* 2009), loss of DDX1 may affect the secondary structure, stability, degradation, subcellular localization and/or translation of RNAs. It is therefore possible that DDX1 plays a similar role to that proposed for DDX20 in the reprogramming from maternal RNA utilization to active transcription from the zygote genome. Lethality could result from disruption of maternal RNA degradation that would interfere with zygote genome activation. Alternatively, deregulation of newly-synthesized zygotic transcripts could have lethal consequence for the developing embryo. The early

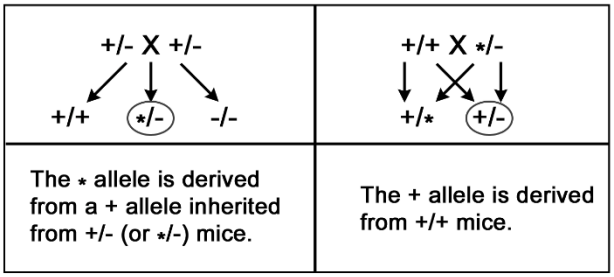
embryonic lethality associated with *Ddx1* and *Ddx20* knock-out suggests distinct roles for these two genes, as expression of DDX1 at early embryonic stages does not compensate for *Ddx20*$^{-/-}$ lethality and vice versa.

Unexpectedly, we observed significantly reduced numbers of wild-type mice when genotyping the progeny of *Ddx1* heterozygote crosses. Reduced numbers of wild-type mice were noted as early as the peri-implantation stage of development that occurs between E4.5 and E5.5 and remained constant at later stages of development suggesting stage-specific lethality. Through analysis of parental genotypes, we were able to identify two distinct populations of heterozygous mice: “abnormal” heterozygote mice (*Ddx1*$^{*/-}$) which arose from heterozygote intercrosses (*Ddx1*$^{+/-}$ X *Ddx1*$^{+/-}$ or *Ddx1*$^{+/-}$ X *Ddx1*$^{*/-}$ or *Ddx1*$^{*/-}$ X *Ddx1*$^{*/-}$) and yielded reduced ratios of wild-type to heterozygote progeny, and “normal” heterozygous mice (*Ddx1*$^{+/-}$) which arose from backcrosses (*Ddx1*$^{+/-}$ X *Ddx1*$^{+/+}$ or *Ddx1*$^{*/-}$ X *Ddx1*$^{+/+}$) and yielded the expected ratios of wild-type to heterozygote progeny (Figure 4.5A). Importantly, the wild-type lethality is not strain-specific as it was observed in both the FVB and C57BL/6 backgrounds. Thus, the presumably genetically identical heterozygous animals are able to distinctly and permanently modulate *Ddx1* expression at a very early developmental stage based on parental genotype. Although the mechanism of *Ddx1*$^{+}$ to *Ddx1** transition is unknown, it may be associated with the epigenetic reprogramming that takes place following MZT as the embryo proceeds to gastrulation (Messerschmidt *et al.* 2014).

Figure 4.5 Inheritance model of the *Ddx1 allele.** (A) Depiction of the two types of wild-type alleles as determined by parental crosses. (B) *Ddx1*$^{+/-}$ intercrosses produce the expected ratio of wild-type to heterozygote progeny, whereas *Ddx1**$^{/-}$ mice intercrosses produce an abnormal ratio of wild-type to heterozygote progeny. (C) *Ddx1*$^{+/-}$ backcrosses produce the expected ratio of wild-type to heterozygote progeny, whereas partial wild-type lethality is observed in *Ddx1**$^{/-}$ backcrosses. (D) Proposed mechanism for wild-type lethality. Under normal conditions, each *Ddx1* allele produces 1X *Ddx1* RNA, resulting in a total of 2X DDX1 RNA and protein. *Ddx1** alleles generate ~2X *Ddx1* RNA to compensate for inactivation of the mutant *Ddx1 allele*. *Ddx1**$^{/+}$ and *Ddx1**$^{/}$* mice are predicted to produce ~3X and 4X *Ddx1* RNA, respectively. This increase results in early embryonic lethality, with higher penetrance observed with increased levels of DDX1.

A

Defining the *Ddx1*⁺ and *Ddx1** alleles



B

Heterozygous intercrosses

Parents	+/- X +/-			*/- X */-		
Progeny	+/+	*/-	-/-	*/*	*/-	-/-
Expected ratio	1	2	0‡	1	2	0‡
Observed ratio	1	2	0	1	9	0

‡ -/- lethality

C

Homozygous X heterozygous backcrosses

Parents	+/+ X +/-		+/+ X */-	
Progeny	+/+	+/-	+/*	+/-
Expected ratio	1	1	1	1
Observed ratio	1	1	1	3

D

Proposed mechanism

Genotype	+/+	+/*	*/*
DDX1 expression levels relative to normal conditions	1X + 1X	~2X + 1X	~2X + ~2X
Total DDX1	2X	~3X	~4X
Observed lethality	None	Partial	High

Two major modes of epigenetic inheritance have been described: paramutation inheritance and genomic imprinting. Paramutations occur when one allele modifies a second locus in a heritable manner. RNA mediated paramutations were first identified in plants, but have also been described in mice (Brink 1958, Chandler 2007, Cuzin *et al.* 2010). The first example of a paramutation in mice was at the *Kit* locus (Rassoulzadegan *et al.* 2006). $Kit^{+/-}$ mice have a white-tail phenotype that is caused by loss of one copy of the Kit tyrosine kinase receptor gene. It was discovered that the white-tail phenotype could be maintained in wild-type (paramutant) $Kit^{+/+}$ offspring derived from $Kit^{+/-}$ heterozygote mice and all $Kit^{+/-}$ mice could generate paramutant $Kit^{+/+}$ offspring. Furthermore, the white-tail phenotype could be transmitted to the next generation when paramutant $Kit^{+/+}$ mice were mated with wild-type mice. Upon further investigation, it was discovered that miRNAs (miR-221 and -222) were being generated at high levels and inherited in subsequent generations through the oocyte or sperm, indicating *trans* rather than *cis* inheritance (Rassoulzadegan *et al.* 2006). These abnormally high levels of miRNAs were responsible for modifying Kit levels from one generation to the next over the course of three generations, resulting in the white-tail phenotype. Two other paramutations were subsequently found to also be induced by microRNAs: *Cdk9* (miR-1) and *Sox9* (miR-124) (Grandjean *et al.* 2009, Wagner *et al.* 2008). While the phenotype associated with the *Ddx1** allele shares some similarities with paramutations, the *Ddx1** phenotype is limited to progeny that inherit the *Ddx1** allele from $Ddx1^{+/-}$ intercrosses. Furthermore, in contrast to *Kit* paramutants which can be generated from $Kit^{+/-}$ backcrosses in addition to heterozygote intercrosses,

mice with the *DDX1*$^{*/-}$ *genotype* are only observed in *Ddx1*$^{+/-}$ intercrosses and subsequent *Ddx1*$^{*/-}$ intercrosses. Thus, our data indicate that, unlike RNA-mediated paramutations, the transgenerational phenotype associated with the *Ddx1** allele is physically associated with the allele.

Genomic imprinting represents a non-conventional form of gene regulation and epigenetic inheritance that is *cis*-acting. Genomic imprinting is characterized by sex-specific changes to DNA methylation that occur during gametogenesis. Imprinted genes display mono-allelic expression, as one of the genes is silenced by methylation. As *Ddx1* expression is bi-allelic and the phenotype associated with the *Ddx1** allele is sex-independent, genomic imprinting is not the mechanism regulating the modification of *Ddx1*. In an attempt to determine whether methylation marks might explain the *Ddx1** phenotype independent of genomic imprinting, we sequenced bisulfite converted genomic DNA from wild type, *Ddx1*$^{+/-}$ and *Ddx1*$^{*/-}$ mice. There were no changes in the methylation status of the single CpG island in the region surrounding *Ddx1*. Thus, we have yet to determine by what mechanism the *Ddx1** phenotype is first generated and then maintained in order to be inherited by subsequent generations.

While we were able to clearly delineate the inheritance pattern underlying the lethality associated with the *Ddx1*$^{*/*}$ genotype, we can only speculate as to the underlying cause of lethality in *Ddx1*$^{*/*}$ embryos (Figures 4.5B-C). We propose that DDX1 protein levels are narrowly regulated in the developing embryo, such that deviations from normal levels are lethal (Figure 4.5D). In support of this idea, attempts to generate lines of transgenic mice overexpressing DDX1 have been

unsuccessful even in mice carrying multiple copies of the *Ddx1* gene (our unpublished data). Compensation in levels of DDX1 RNA and protein in heterozygous mice also indicates that DDX1 levels are tightly regulated. We propose that while heterozygous mice can easily compensate for reduced DDX1 RNA and protein levels by up-regulating DDX1 expression, downward compensation from *Ddx1** alleles that are overexpressing *Ddx1* RNA does not occur. Thus, mice who inherit two compensating *Ddx1* (i.e. *DDX1**) alleles die because of DDX1 overexpression (Figure 4.5D). It is still not clear if DDX1 over-expression is inherently lethal or causing aberrant development during early embryogenesis. The fact that some cancer cell lines can tolerate over-expression of DDX1 (George *et al.* 1996, Godbout *et al.* 1993b, Manohar *et al.* 1995, Squire *et al.* 1995) is in keeping with disruption of developmental processes being the cause of lethality. Based on our data, modification of the wild-type allele in heterozygous mice is flexible for one generation, indicating that the "*cis*" mark is only added after fertilization in the second generation. As some lethality is observed in $Ddx1^{*/+}$ offspring, we attribute this effect to a moderate increase in DDX1 levels that approaches the lethal threshold, such that embryos with acceptable variations in DDX1 levels survive, and embryos that surpass the threshold die.

In summary we found that DDX1 expression is essential for early mouse development, with $Ddx1^{-/-}$ embryos failing to develop to the blastocyst stage. In the process of analyzing the progeny of heterozygote matings, we found that wild-type mice also die during development albeit at a later developmental stage than *Ddx1*

$^{-/-}$ mice (pre E.6.5). In particular, our genotyping analyses indicate that the wild-type allele from *Ddx1*$^{+/-}$ intercrosses is physically marked through an unknown mechanism after the first generation of intercrosses. Our data indicate that DDX1 expression is tightly regulated during embryonic development, and that transcription of the wild-type *Ddx1* gene is up-regulated in *Ddx1*$^{+/-}$ mice thereby compensating for loss of transcription from the mutant allele. We propose a model whereby inheritance of two wild-type *Ddx1* overexpressing alleles leads to embryonic lethality. While we have yet to establish the mechanism causing death during embryonic development, the transgenerational wild-type lethality phenomenon reported here does not appear to have been previously described in the literature and may represent a novel form of epigenetic inheritance.

4.5 ACKNOWLEDGEMENTS

We thank Dr. Peter Dickie for generating the knockout mice, and Mrs. Gail Hipperson, Mr. Dan McGinn and Ms. Daming Li for animal handling and technical support. We are grateful to Dr. Heather McDermid for technical advice and members of the Godbout lab for helpful suggestions.

CHAPTER 5

DISCUSSION AND FUTURE DIRECTIONS

5.1 DISCUSSION

The processes regulating RNA generation, maturation and functional activity are immensely complex. A number of fundamental roles for RNA in protein synthesis were described prior to the 1970's: e.g., mRNAs serving as the intermediary between DNA sequences and protein translation, rRNA involvement in ribosome biogenesis and function, and tRNA involvement in transferring amino acids to elongating peptides (Geiduschek *et al.* 1969, Rich *et al.* 1976, Schweet *et al.* 1966). Over the last 10-15 years, new roles for currently known RNA species, and entirely novel RNA species have been regularly described and continue to be discovered to this day (Girard *et al.* 2006, Lee *et al.* 2009). In addition to the ever increasing number of distinct RNA species expressed in cells, our comprehension of the complexity of the processes that regulate RNA generation and function has also grown. However, while significant progress has been made into the study of RNA biology, there is still much to learn. Beyond a purely biochemical perspective, the nature and consequence of RNA-protein interactions are integral to our understanding of development and disease.

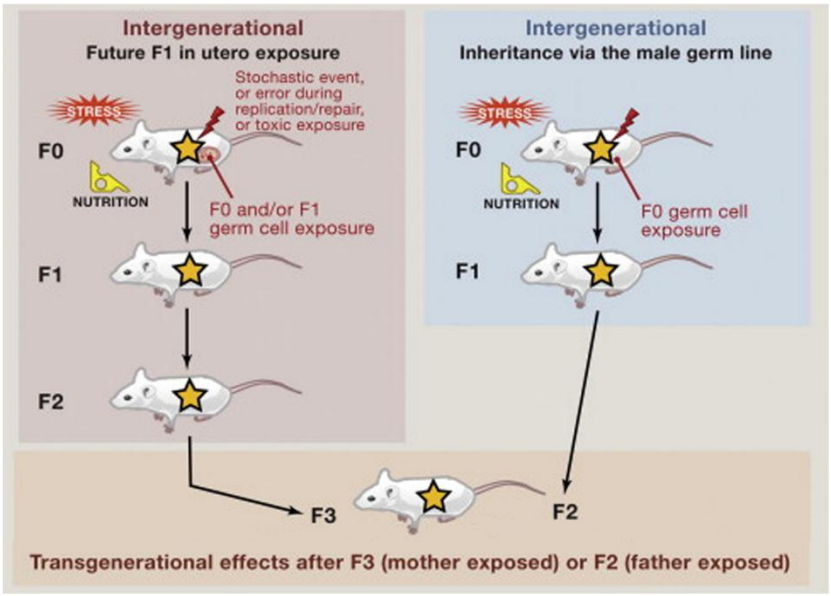
The role of RNA helicases in binding to RNA and modifying their secondary structure allows for intricate regulation of RNA molecules. DEAD box proteins, as a family, have been implicated in every step of RNA biogenesis and metabolism (Linder *et al.* 2011). While a few DEAD box proteins have been extensively described, the majority remain poorly characterized. Here, we present three studies describing novel facets of the function and regulation of the DEAD box protein *DDX1.*

5.1.1 Non-traditional forms of inheritance

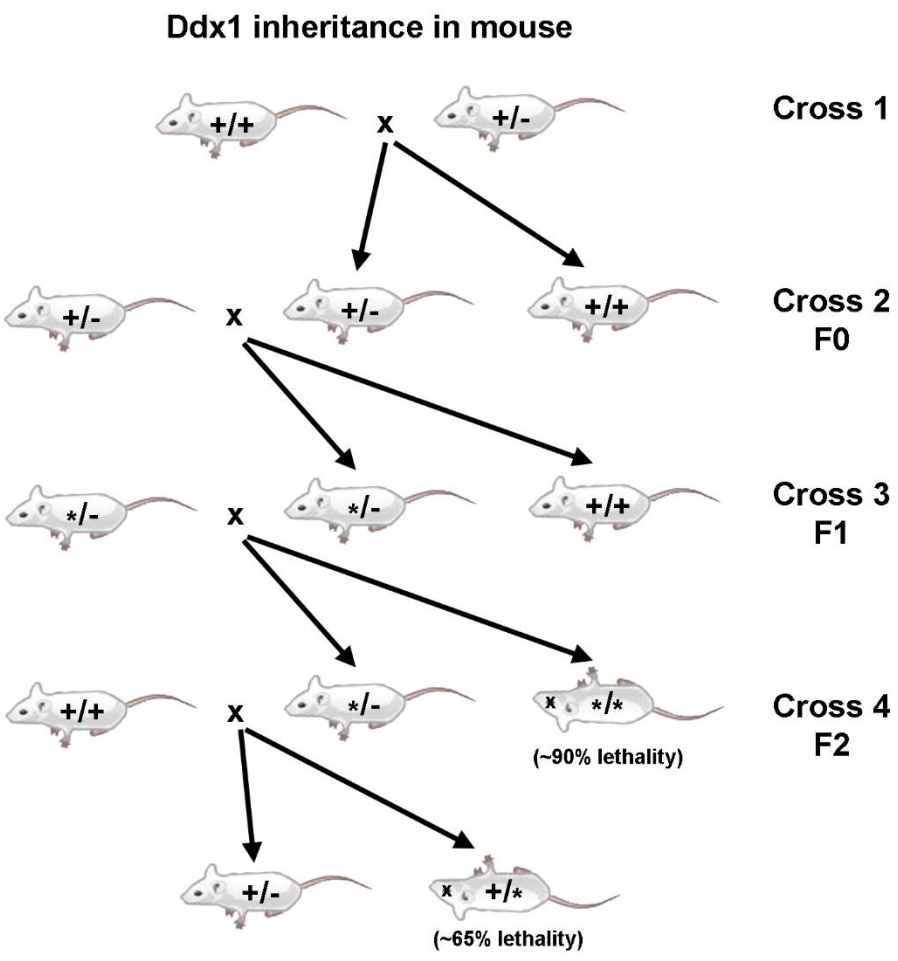
Traditional Mendelian genetic inheritance is dictated by random segregation of DNA encoded genes. An early proposed alternative to Mendelian inheritance was described by Lamarck, based on the idea that parents could pass down traits developed during their lifetime. While the basic tenets of Mendelian genetics have certainly been proven to be the general rule of inheritance, exceptions exist. Several mechanisms have been described whereby heritable information has been transmitted without being directly coded in the genome, or in a non-random fashion. These include well described phenomena such as genomic imprinting, which regulates gene expression based on parental allele-specific DNA methylation (Bartolomei *et al.* 2011), and maternal effects, caused by proteins and RNA molecules produced by maternal cells during zygote generation, which can cause phenotypes associated with parental genotype instead of the progeny's (Mousseau *et al.* 1998). Histones also provide an avenue for modified inheritance, though these effects are generally associated with DNA methylation (Probst *et al.* 2009). Such modifications are often referred to as epi-alleles and although these processes certainly don't fit the traditional description of Mendelian inheritance, they are not truly transgenerational, and should be more appropriately considered inter-generational, as the effects are passed on through only one or two generations (Heard *et al.* 2014) (Figure 5.1). This is mainly due to natural half-life and small quantity of maternally loaded components and the extensive epigenetic reprogramming that occurs during gametogenesis.

Figure 5.1 Transgenerational inheritance. (A) Gametes present in F0 animals are exposed to parental stimuli while undergoing gametogenesis. As such, single generation transmission should not be considered transgenerational. For maternal inheritance in mammals, the developing fetus (F1) and its germ cells (F2) are also exposed to the same stimuli as the F0 animal. In males, F0 to F2 transmission can be considered transgenerational, as the F2 animal was not directly exposed to stimuli. Adapted from Heard *et al.* (2014). (B) Inheritance patterns in *Ddx1* mutant mice. Cross 1: +/+ x +/- generate normal progeny ratios. Cross 2: +/- x +/- generate normal progeny ratios, but resulting heterozygous progeny are modified (indicated by *). Cross 3: */- x */- generate abnormal progeny ratios, with most */* progeny dying. Note that this effect is dependent on the phenotype of the mice in Cross 2. If we designate Cross 2 as F0, we then see the resulting phenotype in F2 animals. As this effect occurs from either parental source, it appears to be a true transgenerational effect. Cross 4: +/+ x */- generate abnormal progeny ratios, but heterozygous mice generated from this cross do not contain a modified *Ddx1* allele. As such we can ascertain that the allele is inherited in *cis* fashion only (as a non-modified allele is inherited from the +/+ wild type animal and future crosses show normal progeny ratios).

A


Intergenerational
Future F1 in utero exposure
Stochastic event, or error during replication/repair, or toxic exposure
STRESS
F0
NUTRITION
F0 and/or F1 germ cell exposure
F1
F2
F3
Intergenerational
Inheritance via the male germ line
STRESS
F0
NUTRITION
F0 germ cell exposure
F1
F2
Transgenerational effects after F3 (mother exposed) or F2 (father exposed)

B


Ddx1 inheritance in mouse
+/+
x
+/-
Cross 1
+/-
x
+/-
+/+
Cross 2
F0
*/-
x
*/-
+/+
Cross 3
F1
+/+
x
*/-
x */*
(~90% lethality)
Cross 4
F2
+/-
x +/*
(~65% lethality)

Multigenerational epigenetic inheritance has been long described in plants, and more recently has been shown across a number of animal species and termed “paramutation” (Chandler 2007). Paramutations differ from the previously described effects, mostly in that they are able to self-propagate. Paramutations include RNA and protein products with self-regulating feedback loops, enduring DNA methylation marks that evade epigenetic reprogramming, and proteins that replicate through structural templating (prions). As these effects are able to amplify during developmental growth of an organism, they can be passed on for an indefinite number of generations. The far-reaching effect of these observations is that stimuli affecting an animal could result in the generation of a paramutation allowing the animal to adapt to environmental conditions and pass that adaption on to its progeny, essentially fulfilling a Lamarckian form of inheritance.

To date, there have been many well-characterized paramutations in plants, but few in animals (Heard *et al.* 2014). Of the studies that have identified epigenetic inheritance, only a minority have been able to determine the mechanism at work. As previously discussed, heterozygous *Kit* mice are able to transmit a white tail phenotype through multiple generations by miRNAs passed on through gametes, that self-perpetuate in the following generation (Rassoulzadegan *et al.* 2006). *Agouti*^{vy} female mice can transmit coat colour and disease susceptibility regardless of progeny genotype, but the effect only lasts a few generations (Daxinger *et al.* 2012). When exposed to heat shock over the course of multiple generations, *Drosophila* develop a modified chromatin state that can persist for multiple

generations without further stimuli due to modifications to dATF-2 (Seong *et al.* 2011).

There is also evidence for transgenerational effects in humans, though our understanding of the underlying process is very limited, making it difficult to differentiate between inter-generational effects and genuine transgenerational inheritance. In either case, a number of studies have shown that mother and grandmother nutritional states can increase the likelihood of children`s metabolic disorders (Hales *et al.* 2013, Lussana *et al.* 2008, Painter *et al.* 2008). Similar analysis in mice has shown transmission of metabolic phenotypes from both maternal and paternal sources (Carone *et al.* 2010, Ferguson-Smith *et al.* 2011, Radford *et al.* 2012). While it is not clear if these effects are truly transgenerational, they set a precedent for environmental stimuli generating inheritable gene regulation in mammals.

5.1.2 Transgenerational inheritance of a modified Ddx1 allele in mice

We have described a novel transgenerational inheritable effect for mouse *Ddx1*. The heterozygous *Ddx1* mouse population is composed of two phenotypically different, but genotypically identical groups. They are differentiated by parental genotype, and their phenotype is a lethal effect in wild-type progeny. As this effect can be maternally or paternally inherited and occurs across two generations, it appears to be transgenerational, though at this time the analysis to prove that it can be inherited indefinitely has not been completed. While we have not directly confirmed the nature of the modification to *Ddx1* alleles in heterozygous

mice, our evidence suggests that the primary effect is to up-regulate Ddx1 levels through increased transcription and/or possibly translation. The combination of our findings in *Drosophila* indicating a role for Ddx1 in metabolism and Ddx1`s tightly controlled expression suggest that Ddx1 may be an important metabolic regulator. If this is the case, then *Ddx1* expression regulation may be modified in response to stress and there could be an evolutionary advantage to pass on this modification in order to prepare progeny for the environment the parent has experienced.

Equally interesting is the currently inexplicable nature of the *Ddx1* modification. As we have observed the phenotype associated with the modified *Ddx1* allele in two distinct mouse strains, and the phenotype is tracked based on the presence of the gene-trapped *Ddx1* allele, the effect must be associated with either the *Ddx1* locus or the nearby surrounding region. In addition, the nature of the phenotype is such that the allele associated with the phenotype originates from a wild-type mouse, as alleles that undergo modification (the + allele in +/- mice) originate from a +/+ parent. The strains utilized are highly inbred, to a point where they are homozygous at virtually all alleles, indicating that the lethal phenotype associated with the allele cannot be the result of a traditional permanent mutation carried in our two control stocks. The likely remaining possibility is that the phenotype results from the heterozygous state of the intermediate generation.

We observe reduced lethality when the modified *Ddx1** allele is carried with a normal wild-type allele as compared to animals homozygous for the modified allele, which we have interpreted as demonstrating that the effect is dose-dependent. As a small number of $^{*/*}$ mice survive, we postulate that there is some

variation in the relative effect of the modification. As such, on occasion mice carrying one or two modified alleles will survive due to stochastic variation. We have proposed that the nature of the modification is to modulate *Ddx1* expression through epigenetic modification. This modification would occur in response to reduced Ddx1 protein levels in heterozygous animals, and acts to normalize Ddx1 levels. When this modification in passed on to a second generation heterozygous mouse, the effect becomes more permanent, and can then be passed on to progeny with two functional *Ddx1* genes. Natural variation in relative over-expression of Ddx1 from the modified allele would then result in *Ddx1*$^{*/+}$ animals, and on occasion *Ddx1*$^{*/*}$ mice, that express Ddx1 within the range associated with viability. While we have no direct evidence for modulation of Ddx1 levels, it is a plausible explanation for the effect observed. Though we cannot dismiss the idea that the observed wild-type lethality is due to changes to other genes in the region surrounding the *Ddx1* locus, the fact that wild-type lethality occurs in response to a *Ddx1* mutation is in keeping with modulation of *Ddx1* levels being the underlying cause of embryonic death.

Current epi-alleles are associated with DNA methylation or maternally/paternally contributed RNA and proteins (Heard *et al.* 2014). The *cis* inheritance of the modified *Ddx1* allele argues against the possibility of parental contributed materials. DNA methylation is similarly unlikely, as the only local CpG island does not show modified methylation and DNA methylation is associated with gene silencing, while we observe an increase in *Ddx1* gene expression. Furthermore, the possibility of a normally methylated region becoming

unmethylated is inconsistent with our observations, as Ddx1 is expressed in all tested conditions. While it remains possible that there is a methylation site distal to the *Ddx1* gene locus, it would still not explain the increased expression observed. As a final scenario, a distal methylation site could be silencing a gene that normally down-regulates *Ddx1* expression, resulting in higher levels of *Ddx1* when the site becomes methylated; however, for this to be the case the site would have to be located in close proximity to *Ddx1*, as we see only a *cis* effect, and the likelihood of a distal, but still linked CpG island silencing a gene, which in turn down-regulates *Ddx1* seems too unlikely to be considered as a valid option.

If *Ddx1* is being epigenetically modified, but not through currently described mechanisms, what alternatives could explain the phenotypes we observe? There are a number of DNA modifications that occur in normal cells about which very little is known. Though the term DNA methylation is commonly used to solely describe 5-methylcytosine, other types of DNA methylation, including 5-hydroxymethylcytosine, 4-methylcytosine and 6-methyaladenine, are also present under normal conditions (Ratel *et al.* 2006). Modifications outside of methylation also occur, including 5-carboxylcytosine, 5-formylcytosine and 8-oxoguanine. Unfortunately, the roles and even existence of biologically relevant levels of these modifications in eukaryotes remains very much unclear.

Another possibility worth consideration is a *cis* based form of gene regulation by an RNA molecule. Barr bodies are formed by *Xist* RNA suppressing a single X chromosome in normal female mammalian cells (Froberg *et al.* 2013). While the randomly selected X chromosome bound by *Xist* RNA is reset every

generation, it persists from gastrulation until death. Theoretically, a modified version of this repression could be responsible for trangenerational epigenetic inheritance, but this model runs into similar issues as previous possible explanations of the cause of *Ddx1* modification. This system of regulation is only known to occur in an “on” or “off” state, which is not consistent with our expression data, and it is not clear if such a system could function on the scale of an individual gene.

A final possible explanation for the observed *Ddx1* mouse phenotype could be DNA editing. Though not an epigenetic modification, a permanent change to the *Ddx1* gene sequence or the regulatory regions of *Ddx1* could result in the observed phenotypes. Apolipoprotein B mRNA editing (APOBECs) enzymes are known to edit RNA in mammalian cells and there is evidence of DNA editing by APOBECs in bacteria (Harris *et al.* 2002). Remarkably, human DDX1 has been shown to interact with APOBEC3G, which could represent a self-regulating feedback system (Chiu *et al.* 2006). Furthermore, the most highly altered transcript observed upon *DDX1* knock-down in HeLa cells is *APOBEC3H* (our unpublished data). Regardless of the exact mechanism, the transgenerational regulation of *Ddx1* appears to be a novel process, and a deeper understanding of the mechanism at work will shed light on both *Ddx1* function as well as inheritable gene regulation.

5.1.3 Ddx1 and metabolism

We report that flies homozygous for a *Ddx1*-null allele display phenotypes previously associated with reduced metabolism. These phenotypes include reduced body size, delayed development and oocysts undergoing autophagy. Given the proposed roles for Ddx1 in modifying RNA secondary structure, we believe that Ddx1 is indirectly modulating metabolism through its interaction with RNA targets. Our deep sequencing analysis identified a number of RNA targets with either modified expression levels and/or altered splice variants in *Ddx1*-null larvae. Gene ontology analysis of genes with increased or decreased expression provides little insight, as the smaller up-regulated group generates no significant ontology terms, and the only specific terms identified in the down-regulated group are uninformative (cuticle development and glutathione S-transferase metabolism). Likewise, the group of genes that display modified splicing could only be associated with general processes.

In addition to the genes that display modified relative levels of splice variants, we identified a population of genes that display unique variants in *Ddx1*-null flies. Specifically, we show that *Sirup* mRNA is unspliced in wild-type flies, and spliced in *Ddx1* knock-out flies. *Sirup* is required for efficient succinate dehydrogenase activity, a step linking the citric acid cycle to the electron transport chain (Van Vranken *et al.* 2014). As *Ddx1*-null flies display metabolically-related phenotypes, we propose that Ddx1 regulates *Sirup* by repressing its splicing under normal conditions. Under stressful conditions (e.g. low nutrient availability), Ddx1 activity would be repressed, allowing for *Sirup* mRNA to be spliced, generating a

variant Sirup protein. In our model, this variant protein has lower activity, and results in the slowing down of cellular metabolism, matching metabolic activity to available nutrient levels.

Perhaps the most fascinating phenotype observed in *Ddx1*-null flies is that they are viable. Previously, *Ddx1* has been considered essential (Zinsmaier *et al.* 1994), and we show that *Ddx1* knock-out mice die at an extremely early developmental stage. Maternally contributed Ddx1 protein can be detected until the 2nd instar larval stage, leading to the suggestion that this protein is sufficient to fulfill the essential role of Ddx1 until later stages when phenotypes emerge. While this model applies to *Drosophila*, it is important to note that *Ddx1* knock-out mice die prior to a period of growth, as mice embryos remain essentially the same size during early development. This discrepancy implies that a different facet of Ddx1 function may be responsible for Ddx1-associated lethality in mice. Given the multitude of roles that Ddx1 is known to play, this is not an unreasonable assertion. Indeed, the miscellaneous nature of the transcripts identified in our RNA deep sequencing screen of *Ddx1*-null *Drosophila* speak to the widespread effects that Ddx1 may be playing.

5.1.4 Other Ddx1 targets in Drosophila

Though our primary focus following our RNA transcriptome analysis was dissecting the relationship between Ddx1 protein and *Sirup* mRNA, other modified RNAs are also of interest. *methuselah* (*mth*) is a G protein-coupled receptor, mutation of which can lead to a significant increase in lifespan (35%), increased

resistance to starvation, but reduced fertility (Araujo *et al.* 2013). Little is known about the function of the family of 15 methuselah-like genes, but RNA expression analysis has shown spatially and temporally restricted expression patterns (Patel *et al.* 2012). We identified *mthl8* as uniquely expressed and *mthl3* as expressing a unique splice variant in *Ddx1*-null 3rd instar larvae. *mthl8* and *mthl3* have been shown to be expressed in 3rd instar imaginal discs and nervous tissue, but are not expressed in embryonic tissue. We did not detect *mthl8* transcripts in wild-type 3rd instar larvae, but we attribute this to low levels of expression under normal conditions. In any case, a significant increase in *mthl8* RNA levels occurred in *Ddx1*-null flies. In the case of *mthl3*, our data indicate that an alternative splice donor in 1st exon is utilized in *Ddx1*-null larvae. This splice variation does not modify the open reading frame, but results in the inclusion of an additional ~300 nt in the 5' untranslated region. Unfortunately, because there is very little information as to the function of the methuselah-like gene family, it is difficult to postulate what the outcome of these changes would be.

Another group of genes that displayed modified expression levels is the family of salivary gland secretion (*sgs*) genes. *sgs* genes are only expressed in the salivary glands during the second half of the 3rd instar larval stage (Lehmann 1996). While their function has not been described in *Drosophila*, they have been shown to be a major component of saliva in insects. We observed a significant down-regulation of *sgs1* (~4X) and *sgs3* (~5X), and a non-significant down-regulation trend for *sgs5* (~2.5X), *sgs7* (~2.5X) *and sgs8* (~2x) in *Ddx1*-null flies. *sgs* genes are regulated by ecdysone levels (Lehmann *et al.* 1995), which means

that as an alternative to direct regulation by Ddx1, this change may be a downstream result of larger changes to developmental regulation and metabolism.

5.1.5 DDX1, driving treatment resistance and tumour growth?

We showed a strong prognostic effect for *DDX1* mRNA levels and protein localization in our breast cancer cohort, but verification in additional cohorts is essential for confirmation of the validity of this finding. Following the publication of our results, two additional analyses for *DDX1* in breast cancer have been published. Balko *et al.* reviewed our results and performed analysis of patients outcome based on *DDX1* mRNA levels using a publicly available database (Balko *et al.* 2011). Their results were consistent with ours, with patients in the top 50 percentile of *DDX1* mRNA levels having a worse outcome. Far more interesting though, was that they showed that this effect is lost in patients who refuse any form of treatment. This mirrors our observation that elevated levels of *DDX1* are associated with early recurrence, specifically in the case of patients who receive either chemotherapy or hormone therapy. Together, these results suggest a role for DDX1 in conveying treatment resistance in tumours. Intriguingly, this effect seems to convey resistance to a wide range of treatment regimens. Li *et al* have previously published that DDX1 is involved in the repair of double strand breaks (Li *et al.* 2008). A partial explanation could therefore be that breast cancer cells with elevated levels of DDX1 are able to repair DNA damage caused by chemotherapeutic drugs that induce DNA damage more effectively than breast cancer cells with low levels of DDX1 (also see below).

One other study of DDX1 status in breast cancer has been reported. Taunk *et al.* looked at DDX1 protein levels in a cohort of 282 early stage node-negative breast cancer patients (Taunk *et al.* 2012). They observed a positive prognostic outcome for patients whose tumours expressed DDX1. While these results conflict with our observations, there are a number of possible reasons for this. First, different patient populations were analyzed, though with Balko *et al.'s* results in line with our observations, it seems unlikely that the inconsistency can be attributed simply to cohort variation. Second, a far more likely explanation is that the anti-DDX1 antibody used by Taunk *et al.*, a previously uncharacterized commercial anti-DDX1 antibody, may not be reliable. This could explain the abnormal localization patterns observed. These authors report that approximately 50% of the tumours analysed were DDX1-positive, with the majority displaying predominantly cytoplasmic localization of DDX1. This is in contrast to our observations using a well-characterized anti-DDX1 antibody showing that virtually every breast cancer expresses nuclear DDX1. Widespread expression of DDX1 in breast cancers and a predominantly nuclear localization is in keeping with the generally accepted expression patterns for DDX1. However it is worth noting that a study in ovarian cancer also reported low DDX1 levels being associated with a worse outcome (Han *et al.* 2014). Though there is still some disagreement as to the effect of DDX1 mis-expression in cancer, at least in breast cancer, the evidence suggests an association between elevated levels of *DDX1* and a worse prognosis.

As alluded to earlier, many of the processes that DDX1 has been implicated in could affect tumour growth and/or response to treatment. As we have evidence

that DDX1 regulates metabolism and growth in *Drosophila*, it is appealing to propose that mis-regulation of DDX1 could also affect the growth properties and metabolism of breast cancer cells. As DDX1 has been implicated in the repair of DNA double strand breaks, overall elevated levels of DDX1 could increase the DNA repair capacity of cancer cells exposed to chemotherapeutic agents that cause damage to the DNA. Alternatively, DDX1 appears to regulate the expression of many different RNA targets, any number of which could promote tumourigenesis when mis-expressed. Each of these possibilities could be responsible for resistance to some forms of treatment. Promoting growth could abrogate the effects of hormone therapies, while increased DNA repair could provide resistance to DNA damaging chemotherapeutic agents. The fact that DDX1 appears to provide resistance to a range of treatment options suggests either a general resistance mechanism, perhaps by reducing cell death in response to different forms of stimuli, or multiple resistance mechanisms, which in combination, provide a general treatment resistant phenotype.

5.2 FUTURE DIRECTIONS

5.2.1 Understanding and targeting DDX1 in cancer

While there is provocative evidence for *DDX1* playing a role in breast cancer, there is still a need for additional studies. In particular, the studies to date have only included patients with variable treatment regimens. It would be very informative to look at the effect of *DDX1* levels and/or localization in clinical trial cohorts with groups receiving controlled treatment regimens. This would allow us to elucidate which treatments *DDX1* mis-expression confers resistance to. In turn, this information could be taken into consideration when determining appropriate treatment regimens for individual patients, potentially reducing harm caused by therapeutic agents that do not provide substantial benefits due to *DDX1* status. In conjunction with additional breast cancer studies, it would also be of interest to determine the effect of *DDX1* in other types of cancers. Such studies could provide valuable insight into the scope of *DDX1* mis-expression and DDX1's prognostic value in cancer.

In addition to utilizing patient tissues, the development of a panel of cancer cell lines with high/low *DDX1* mRNA levels and high/low cytoplasmic DDX1 protein could provide useful tools to further our understanding of the role of *DDX1* in tumour progression and treatment resistance. Particularly, the pathological characteristics (i.e. invasiveness, growth rates) of these cell lines could be measured as a function of their *DDX1* status. These data, in conjunction with experiments designed to modify DDX1 levels (RNAi knock-down, over-expression) could help to clarify the effects of *DDX1* in cancer. A *DDX1* modified panel could

also be utilized to determine more precisely if *DDX1* status confers treatment resistance to cells. Again, by employing methods to modify *DDX1* levels, characterizing responses to treatment agents could be measured within the context of *DDX1* high vs low backgrounds, or in cytoplasmic vs non-cytoplasmic DDX1 conditions.

While we already have some indication of DDX1's importance in breast cancer, the above experiments could validate the use of DDX1 as a prognostic and/or predictive marker, and as a potential drug target. Development of a small molecule inhibitor would be useful, not only for research purposes, but also as a possible adjuvant therapy that may sensitize tumour cells to other agents or directly cause death of tumour cells. To date, small molecule inhibitors have been developed for only a few DEAD box proteins (Cencic *et al.* 2013). So far, different types of compounds have been shown to inhibit DEAD box activity, including nucleoside analogues and small molecules that disrupt the interaction between a target helicase and a co-activating binding protein. Additionally, a small bacterial protein (*Burkholderia* lethal factor 1) has been shown to inactivate eIF4A (Cruz-Migoni *et al.* 2011). Though the use of DEAD box proteins as drug targets in disease is still very much in its infancy, these proteins represent a potentially useful group of targets for new drug development.

5.2.2 Functional analysis of Ddx1 in Drosophila

Previous to our newly generated *Ddx1*^{AX} homozygous flies, no *Ddx1*-null cells had been reported. As such, *Ddx1*-null flies are a valuable tool for functional

analysis. The first objective would be transgenically express *Ddx1* in our *Ddx1*-null background to rescue observed phenotypes, with the objective of conclusively demonstrating that the effects observed are due to the $Ddx1^{AX}$ allele. Following this, a number of different constructs could be tested. For example, expressing human or mouse *DDX1* will determine if DDX1 retains conserved functions across species. Introducing *Ddx1* constructs encoding truncated proteins, or proteins with deleted domains (e.g. SPRY domain) could further reveal which domains are essential for function. In addition, if distinct roles are responsible for separate phenotypes, we may be able to distinguish the roles with truncated protein products.

As we believe *Ddx1*-null phenotypes to be the result of deregulation of downstream RNA targets of Ddx1, it would be of interest to express or knock-down identified targets to recapitulate phenotypes observed in out *Ddx1* mutant flies. For example, we propose that the spliced form of *Sirup* results in reduced metabolism. As such, expressing the alternative *Sirup* transcript should phenocopy at least a subset of the *Ddx1*-null phenotypes. By over-expressing or knocking-down targets, as appropriate, we should be able to link specific targets to specific downstream effects observed in *Ddx1*-null flies. Since we believe Ddx1 to be multi-functional, this analysis may provide insight into the different processes regulated by Ddx1 in *Drosophila.*

We have already shown a direct interaction between Ddx1 and *Sirup* mRNA, and have proposed that Ddx1 regulates *Sirup* splicing through modulation of an anti-sense *Sirup* RNA molecule (Morrissy *et al.* 2011). While we have some

evidence of the binding location of a putative anti-sense *Sirup* RNA molecule, the length, sequence and origin of this RNA is unknown. Further characterization of this RNA could provide new insight into how anti-sense RNA molecules regulate splicing. While somewhat technically difficult, small scale directional RNA deep sequencing could also be used to determine the sequence of the anti-sense *Sirup* RNA. Once this sequence has been identified, it will be possible to measure the binding and unwinding activity of Ddx1 using *in vitro* transcribed *Sirup* mRNA and anti-sense RNA by gel shift and unwinding assays.

It would also be of interest to undertake *in vitro* splicing assays using *Sirup* pre-mRNA. Though it's not possible to generate *Ddx1*-null splicing extracts from traditional sources (due to maternally contributed Ddx1 in embryos, and the absence of a *Ddx1*-null *Drosophila* cell line), we could deplete normal extracts using our anti-Ddx1 antibody, which has already been shown to immunoprecipitate *Drosophila* Ddx1 with high efficiency. If we are able to establish a viable *in vitro* splicing assay for *Sirup,* we could then use Ddx1 depleted extracts and reintroduce Ddx1 variants as described above. This would allow to us to dissect specific domain function and confirm any functions that are conserved across Ddx1 orthologues. Once established, this splicing assay could also be used to determine the role of Ddx1 in regulating splice variant expression of other targets identified in our RNA deep sequencing analysis.

5.2.3 Investigating the regulation and function of Ddx1 mice

The nature of the modification of *Ddx1* alleles in heterozygous mice remains a mystery. As we have already shown that local DNA methylation does not appear to be responsible, investigating alternative modes of modification may reveal the mechanism at work. In particular, DNA editing could be confirmed by sequencing the entire *Ddx1* locus. In addition, a novel sequencing technique, SMRT sequencing, has been developed that can identify the presence of uncommon DNA modifications (Roberts *et al.* 2013). This technique could be used to determine whether the observed inheritable effects are due to uncommon epigenetic DNA modifications.

We have characterized *Ddx1*-related lethality at two distinct stages, with knock-out mice dying prior to E3.5 and *Ddx1*$^{*/*}$ mice dying between E3.5 and E6.5. Further analysis of the cause of death is still required. Live cell imaging of developing embryos isolated directly post fertilization could provide some insight into the cause of death in knock-out mice. Genotyping embryos imaged in this fashion is extremely difficult, limiting our ability to generate definite conclusions; however we could obtain some clues based on trends observed. Lethality in *Ddx1*$^{*/*}$ mice may be easier to assess, as entire uteruses from pregnant mice can be fixed and sectioned. This would allow us to discern morphological phenotypes potentially identifying gross effects of the *Ddx1** allele in affected mice. Immunohistochemistry could be used for genotyping as well as looking at expression patterns for genes of interest that may be playing a role in *Ddx1** related lethality.

The early lethality observed in *Ddx1* knock-out mice precludes most analyses of gene function. However, methods exist for the generation of developmentally and/or spatially targeted gene knock-out in mice. Techniques involving Cre recombinase and LoxP recognition sites can be used for targeted gene deletion, recombination or inversion (Kos 2004). Generation of a mouse line with *Ddx1* flanked by LoxP sites would allow analysis of *Ddx1* disruption at later developmental stages, or in specific tissues, to determine if Ddx1 is ubiquitously essential in mice, or only required during early development. Such an approach could provide valuable insight into the role of *Ddx1* in mice.

5.3 SIGNIFICANCE

Like many other DEAD box proteins, DDX1 is a multi-functional protein involved in a number of cellular processes. DDX1 has also been implicated in several cancers and viral replication. While significant progress has been made towards understanding DDX1's biochemical properties and roles in intracellular processes, there has been little effort towards understanding the role of DDX1 in developmental and the effect of DDX1 in disease remains unclear. This project helped to expand our understanding of the role of DDX1 in development, as well as breast cancer. We have characterized phenotypes caused by novel mutations of *Ddx1* in *Drosophila* and mice. We have identified a novel form of transgenerational epigenetic inheritance and implicated *Ddx1* in early development and metabolism regulation. We have also identified *Sirup* mRNA as a target of Ddx1, with modified splicing of *Sirup* mRNA resulting from Ddx1 depletion. Finally, we have shown that DDX1 is a prognostic factor and have evidence that DDX1 may be a predictive factor in breast cancer. Together, our data have contributed to a better understanding of the many roles of the multifunctional DDX1 in development and cancer, and implicate DDX1 as a possible target for the treatment of breast cancer.

REFERENCES

Activities at the Universal Protein Resource (UniProt). Nucleic Acids Res, 2014. **42**(Database issue): p. D191-8.

Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet, 2005. **365**(9472): p. 1687-717.

Adamovic, T., F. Trosso, L. Roshani, L. Andersson, G. Petersen, S. Rajaei, K. Helou, and G. Levan, *Oncogene amplification in the proximal part of chromosome 6 in rat endometrial adenocarcinoma as revealed by combined BAC/PAC FISH, chromosome painting, zoo-FISH, and allelotyping.* Genes Chromosomes Cancer, 2005. **44**(2): p. 139-53.

Albain, K.S., W.E. Barlow, S. Shak, G.N. Hortobagyi, R.B. Livingston, I.T. Yeh, P. Ravdin, R. Bugarini, F.L. Baehner, N.E. Davidson, G.W. Sledge, E.P. Winer, C. Hudis, J.N. Ingle, E.A. Perez, K.I. Pritchard, L. Shepherd, J.R. Gralow, C. Yoshizawa, D.C. Allred, C.K. Osborne, and D.F. Hayes, *Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial.* Lancet Oncol, 2010. **11**(1): p. 55-65.

Alexandrov, A., D. Colognori, and J.A. Steitz, *Human eIF4AIII interacts with an eIF4G-like partner, NOM1, revealing an evolutionarily conserved function outside the exon junction complex.* Genes Dev, 2011. **25**(10): p. 1078-90.

American Cancer Society, *Breast Cancer Facts & Figures 2013-2014*. 2013: Atlanta.

American Cancer Society, *Breast Cancer in Men*. 2010: Atlanta, GA.

Anantharaman, V., E.V. Koonin, and L. Aravind, *Comparative genomics and evolution of proteins involved in RNA metabolism.* Nucleic Acids Res, 2002. **30**(7): p. 1427-64.

Andersen, C.B., L. Ballut, J.S. Johansen, H. Chamieh, K.H. Nielsen, C.L. Oliveira, J.S. Pedersen, B. Seraphin, H. Le Hir, and G.R. Andersen, *Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA.* Science, 2006. **313**(5795): p. 1968-72.

Aparicio, T., R. Baer, and J. Gautier, *DNA double-strand break repair pathway choice and cancer.* DNA Repair (Amst), 2014. **19**: p. 169-75.

Araujo, A.R., M. Reis, H. Rocha, B. Aguiar, R. Morales-Hojas, S. Macedo-Ribeiro, N.A. Fonseca, D. Reboiro-Jato, M. Reboiro-Jato, F. Fdez-Riverola, C.P. Vieira, and J. Vieira, *The Drosophila melanogaster methuselah gene: a novel gene with ancient functions.* PLoS One, 2013. **8**(5): p. e63747.

Badve, S., D.J. Dabbs, S.J. Schnitt, F.L. Baehner, T. Decker, V. Eusebi, S.B. Fox, S. Ichihara, J. Jacquemier, S.R. Lakhani, J. Palacios, E.A. Rakha, A.L. Richardson, F.C. Schmitt, P.H. Tan, G.M. Tse, B. Weigelt, I.O. Ellis, and J.S. Reis-Filho, *Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists.* Mod Pathol, 2011. **24**(2): p. 157-67.

Balko, J.M. and C.L. Arteaga, *Dead-box or black-box: is DDX1 a potential biomarker in breast cancer?* Breast Cancer Res Treat, 2011. **127**(1): p. 65-7.

Barnadas, A., M. Gil, P. Sanchez-Rovira, A. Llombart, E. Adrover, L.G. Estevez, J. de la Haba, and L. Calvo, *Neoadjuvant endocrine therapy for breast cancer: past, present and future.* Anticancer Drugs, 2008. **19**(4): p. 339-47.

Barr, F.G., F. Duan, L.M. Smith, D. Gustafson, M. Pitts, S. Hammond, and J.M. Gastier-Foster, *Genomic and clinical analyses of 2p24 and 12q13-q14 amplification in alveolar rhabdomyosarcoma: a report from the Children's Oncology Group.* Genes Chromosomes Cancer, 2009. **48**(8): p. 661-72.

Barth, J.M., J. Szabad, E. Hafen, and K. Kohler, *Autophagy in Drosophila ovaries is induced by starvation and is required for oogenesis.* Cell Death Differ, 2011. **18**(6): p. 915-24.

Bartolomei, M.S. and A.C. Ferguson-Smith, *Mammalian genomic imprinting.* Cold Spring Harb Perspect Biol, 2011. **3**(7).

Bastock, R. and D. St Johnston, *Drosophila oogenesis.* Curr Biol, 2008. **18**(23): p. R1082-7.

Bates, G.J., S.M. Nicol, B.J. Wilson, A.M. Jacobs, J.C. Bourdon, J. Wardrop, D.J. Gregory, D.P. Lane, N.D. Perkins, and F.V. Fuller-Pace, *The*

DEAD box protein p68: a novel transcriptional coactivator of the p53 tumour suppressor. EMBO J, 2005. **24**(3): p. 543-53.

Beheshti, B., I. Braude, P. Marrano, P. Thorner, M. Zielenska, and J.A. Squire, *Chromosomal localization of DNA amplifications in neuroblastoma tumors using cDNA microarray comparative genomic hybridization.* Neoplasia, 2003. **5**(1): p. 53-62.

Bennardo, N., A. Cheng, N. Huang, and J.M. Stark, *Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair.* PLoS Genet, 2008. **4**(6): p. e1000110.

Beretta, L., A.C. Gingras, Y.V. Svitkin, M.N. Hall, and N. Sonenberg, *Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation.* EMBO J, 1996. **15**(3): p. 658-64.

Bergh, J., P.E. Jonsson, B. Glimelius, and P. Nygren, *A systematic overview of chemotherapy effects in breast cancer.* Acta Oncol, 2001. **40**(2-3): p. 253-81.

Berry, D.A., K.A. Cronin, S.K. Plevritis, D.G. Fryback, L. Clarke, M. Zelen, J.S. Mandelblatt, A.Y. Yakovlev, J.D. Habbema, and E.J. Feuer, *Effect of screening and adjuvant therapy on mortality from breast cancer.* N Engl J Med, 2005. **353**(17): p. 1784-92.

Bertucci, F., P. Finetti, N. Cervera, B. Esterni, F. Hermitte, P. Viens, and D. Birnbaum, *How basal are triple-negative breast cancers?* Int J Cancer, 2008. **123**(1): p. 236-40.

Beshore, E.L., T.J. McEwen, M.C. Jud, J.K. Marshall, J.A. Schisa, and K.L. Bennett, *C. elegans Dicer interacts with the P-granule component GLH-1 and both regulate germline RNPs.* Dev Biol, 2011. **350**(2): p. 370-81.

Bleoo, S., X. Sun, M.J. Hendzel, J.M. Rowe, M. Packer, and R. Godbout, *Association of human DEAD box protein DDX1 with a cleavage stimulation factor involved in 3'-end processing of pre-MRNA.* Mol Biol Cell, 2001. **12**(10): p. 3046-59.

Boeke, J., I. Bag, M.J. Ramaiah, I. Vetter, E. Kremmer, M. Pal-Bhadra, U. Bhadra, and A. Imhof, *The RNA helicase Rm62 cooperates with SU(VAR)3-9 to re-silence active transcription in Drosophila melanogaster.* PLoS One, 2011. **6**(6): p. e20761.

Bohni, R., J. Riesgo-Escovar, S. Oldham, W. Brogiolo, H. Stocker, B.F. Andruss, K. Beckingham, and E. Hafen, *Autonomous control of cell and organ size by CHICO, a Drosophila homolog of vertebrate IRS1-4.* Cell, 1999. **97**(7): p. 865-75.

Bono, F., J. Ebert, E. Lorentzen, and E. Conti, *The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA.* Cell, 2006. **126**(4): p. 713-25.

Bordeleau, M.E., J. Matthews, J.M. Wojnar, L. Lindqvist, O. Novac, E. Jankowsky, N. Sonenberg, P. Northcote, P. Teesdale-Spittle, and J. Pelletier, *Stimulation of mammalian translation initiation factor eIF4A activity by a small molecule inhibitor of eukaryotic translation.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10460-5.

Bowers, H.A., P.A. Maroney, M.E. Fairman, B. Kastner, R. Luhrmann, T.W. Nilsen, and E. Jankowsky, *Discriminatory RNP remodeling by the DEAD-box protein DED1.* RNA, 2006. **12**(5): p. 903-12.

Bowles, E., T.W. Corson, J. Bayani, J.A. Squire, N. Wong, P.B. Lai, and B.L. Gallie, *Profiling genomic copy number changes in retinoblastoma beyond loss of RB1.* Genes Chromosomes Cancer, 2007. **46**(2): p. 118-29.

Breitwieser, W., F.H. Markussen, H. Horstmann, and A. Ephrussi, *Oskar protein interaction with Vasa represents an essential step in polar granule assembly.* Genes Dev, 1996. **10**(17): p. 2179-88.

Brennecke, J., A.A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, and G.J. Hannon, *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila.* Cell, 2007. **128**(6): p. 1089-103.

Brenton, J.D., L.A. Carey, A.A. Ahmed, and C. Caldas, *Molecular classification and molecular forecasting of breast cancer: ready for clinical application?* J Clin Oncol, 2005. **23**(29): p. 7350-60.

Brink, R.A., *Paramutation at the R locus in maize.* Cold Spring Harb Symp Quant Biol, 1958. **23**: p. 379-91.

Buszczak, M., S. Paterno, D. Lighthouse, J. Bachman, J. Planck, S. Owen, A.D. Skora, T.G. Nystul, B. Ohlstein, A. Allen, J.E. Wilhelm, T.D. Murphy, R.W. Levis, E. Matunis, N. Srivali, R.A. Hoskins, and A.C. Spradling, *The*

carnegie protein trap library: a versatile tool for Drosophila developmental studies. Genetics, 2007. **175**(3): p. 1505-31.

Buszczak, M. and A.C. Spradling, *The Drosophila P68 RNA helicase regulates transcriptional deactivation by promoting RNA release from chromatin.* Genes Dev, 2006. **20**(8): p. 977-89.

Canadian Cancer Society, *Canadian Cancer Statistics 2014.* 2014: Toronto, ON.

Caretti, G., R.L. Schiltz, F.J. Dilworth, M. Di Padova, P. Zhao, V. Ogryzko, F.V. Fuller-Pace, E.P. Hoffman, S.J. Tapscott, and V. Sartorelli, *The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation.* Dev Cell, 2006. **11**(4): p. 547-60.

Carone, B.R., L. Fauquier, N. Habib, J.M. Shea, C.E. Hart, R. Li, C. Bock, C. Li, H. Gu, P.D. Zamore, A. Meissner, Z. Weng, H.A. Hofmann, N. Friedman, and O.J. Rando, *Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals.* Cell, 2010. **143**(7): p. 1084-96.

Castle, W.E., *Inbreeding, Cross-Breeding and Sterility in Drosophila.* Science, 1906. **23**(578): p. 153.

Cauchi, R.J., *Conserved requirement for DEAD-box RNA helicase Gemin3 in Drosophila oogenesis.* BMC Res Notes, 2012. **5**: p. 120.

Cauchi, R.J., K.E. Davies, and J.L. Liu, *A motor function for the DEAD-box RNA helicase, Gemin3, in Drosophila.* PLoS Genet, 2008. **4**(11): p. e1000265.

Cencic, R. and J. Pelletier, *Throwing a monkey wrench in the motor: targeting DExH/D box proteins with small molecule inhibitors.* Biochim Biophys Acta, 2013. **1829**(8): p. 894-903.

Chan, C.C., J. Dostie, M.D. Diem, W. Feng, M. Mann, J. Rappsilber, and G. Dreyfuss, *eIF4A3 is a novel component of the exon junction complex.* RNA, 2004. **10**(2): p. 200-9.

Chandler, V.L., *Paramutation: from maize to mice.* Cell, 2007. **128**(4): p. 641-5.

Cheang, M.C., S.K. Chia, D. Voduc, D. Gao, S. Leung, J. Snider, M. Watson, S. Davies, P.S. Bernard, J.S. Parker, C.M. Perou, M.J. Ellis, and T.O. Nielsen, *Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer.* J Natl Cancer Inst, 2009. **101**(10): p. 736-50.

Chen, H.C., W.C. Lin, Y.G. Tsay, S.C. Lee, and C.J. Chang, *An RNA helicase, DDX1, interacting with poly(A) RNA and heterogeneous nuclear ribonucleoprotein K.* J Biol Chem, 2002. **277**(43): p. 40403-9.

Chen, Y., J.P. Potratz, P. Tijerina, M. Del Campo, A.M. Lambowitz, and R. Russell, *DEAD-box proteins can completely separate an RNA duplex using a single ATP.* Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20203-8.

Chiu, Y.L., H.E. Witkowska, S.C. Hall, M. Santiago, V.B. Soros, C. Esnault, T. Heidmann, and W.C. Greene, *High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition.* Proc Natl Acad Sci U S A, 2006. **103**(42): p. 15588-93.

Chiuri, V.E., G. Leo, and V. Lorusso, *Clinical and therapeutic perspectives of gene expression profiling for breast cancer.* Ann Oncol, 2007. **18 Suppl 6**: p. vi58-62.

Cho, B., Y. Lim, D.Y. Lee, S.Y. Park, H. Lee, W.H. Kim, H. Yang, Y.J. Bang, and D.I. Jeoung, *Identification and characterization of a novel cancer/testis antigen gene CAGE.* Biochem Biophys Res Commun, 2002. **292**(3): p. 715-26.

Chou, C.F., W.J. Lin, C.C. Lin, C.A. Luber, R. Godbout, M. Mann, and C.Y. Chen, *DEAD box protein DDX1 regulates cytoplasmic localization of KSRP.* PLoS One, 2013. **8**(9): p. e73752.

Ciesla, M. and M. Boguta, *Regulation of RNA polymerase III transcription by Maf1 protein.* Acta Biochim Pol, 2008. **55**(2): p. 215-25.

Clark, E.L., F.V. Fuller-Pace, D.J. Elliott, and C.N. Robson, *Coupling transcription to RNA processing via the p68 DEAD box RNA helicase androgen receptor co-activator in prostate cancer.* Biochem Soc Trans, 2008. **36**(Pt 3): p. 546-7.

Claycomb, J.M., P.J. Batista, K.M. Pang, W. Gu, J.J. Vasale, J.C. van Wolfswinkel, D.A. Chaves, M. Shirayama, S. Mitani, R.F. Ketting, D. Conte, Jr., and C.C. Mello, *The Argonaute CSR-1 and its 22G-RNA cofactors are*

required for holocentric chromosome segregation. Cell, 2009. **139**(1): p. 123-34.

Collins, R., T. Karlberg, L. Lehtio, P. Schutz, S. van den Berg, L.G. Dahlgren, M. Hammarstrom, J. Weigelt, and H. Schuler, *The DEXD/H-box RNA helicase DDX19 is regulated by an {alpha}-helical switch.* J Biol Chem, 2009. **284**(16): p. 10296-300.

Conroy, S.C., T.E. Dever, C.L. Owens, and W.C. Merrick, *Characterization of the 46,000-dalton subunit of eIF-4F.* Arch Biochem Biophys, 1990. **282**(2): p. 363-71.

Cordin, O., J. Banroques, N.K. Tanner, and P. Linder, *The DEAD-box protein family of RNA helicases.* Gene, 2006. **367**: p. 17-37.

Corson, T.W. and B.L. Gallie, *One hit, two hits, three hits, more? Genomic changes in the development of retinoblastoma.* Genes Chromosomes Cancer, 2007. **46**(7): p. 617-34.

Cross, D.A., D.R. Alessi, P. Cohen, M. Andjelkovich, and B.A. Hemmings, *Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B.* Nature, 1995. **378**(6559): p. 785-9.

Cross, D.P. and D.L. Shellenbarger, *The dynamics of Drosophila melanogaster spermatogenesis in in vitro cultures.* J Embryol Exp Morphol, 1979. **53**: p. 345-51.

Cruz-Migoni, A., G.M. Hautbergue, P.J. Artymiuk, P.J. Baker, M. Bokori-Brown, C.T. Chang, M.J. Dickman, A. Essex-Lopresti, S.V. Harding, N.M. Mahadi, L.E. Marshall, G.W. Mobbs, R. Mohamed, S. Nathan, S.A. Ngugi, C. Ong, W.F. Ooi, L.J. Partridge, H.L. Phillips, M.F. Raih, S. Ruzheinikov, M. Sarkar-Tyson, S.E. Sedelnikova, S.J. Smither, P. Tan, R.W. Titball, S.A. Wilson, and D.W. Rice, *A Burkholderia pseudomallei toxin inhibits helicase activity of translation factor eIF4A.* Science, 2011. **334**(6057): p. 821-4.

Cuzin, F. and M. Rassoulzadegan, *Non-Mendelian epigenetic heredity: gametic RNAs as epigenetic regulators and transgenerational signals.* Essays Biochem, 2010. **48**(1): p. 101-6.

Cybulski, N. and M.N. Hall, *TOR complex 2: a signaling pathway of its own.* Trends Biochem Sci, 2009. **34**(12): p. 620-7.

Dardenne, E., S. Pierredon, K. Driouch, L. Gratadou, M. Lacroix-Triki, M.P. Espinoza, E. Zonta, S. Germann, H. Mortada, J.P. Villemin, M. Dutertre, R. Lidereau, S. Vagner, and D. Auboeuf, *Splicing switch of an epigenetic regulator by RNA helicases promotes tumor-cell invasiveness.* Nat Struct Mol Biol, 2012. **19**(11): p. 1139-46.

Daxinger, L. and E. Whitelaw, *Understanding transgenerational epigenetic inheritance via the gametes in mammals.* Nat Rev Genet, 2012. **13**(3): p. 153-62.

De Preter, K., F. Speleman, V. Combaret, J. Lunec, J. Board, A. Pearson, A. De Paepe, N. Van Roy, G. Laureys, and J. Vandesompele, *No evidence for correlation of DDX1 gene amplification with improved survival probability in patients with MYCN-amplified neuroblastomas.* J Clin Oncol, 2005. **23**(13): p. 3167-8; author reply 3168-70.

de Souza, D.R., S.S. Sanabani, A.C. de Souza, V. Filho Odone, S. Epelman, and I. Bendit, *Prognostic impact of MYCN, DDX1, TrkA, and TrkC gene transcripts expression in neuroblastoma.* Pediatr Blood Cancer, 2011. **56**(5): p. 749-56.

Dedes, K.J., P.M. Wilkerson, D. Wetterskog, B. Weigelt, A. Ashworth, and J.S. Reis-Filho, *Synthetic lethality of PARP inhibition in cancers lacking BRCA1 and BRCA2 mutations.* Cell Cycle, 2011. **10**(8): p. 1192-9.

Del Campo, M., S. Mohr, Y. Jiang, H. Jia, E. Jankowsky, and A.M. Lambowitz, *Unwinding by local strand separation is critical for the function of DEAD-box proteins as RNA chaperones.* J Mol Biol, 2009. **389**(4): p. 674-93.

Del Campo, M., P. Tijerina, H. Bhaskaran, S. Mohr, Q. Yang, E. Jankowsky, R. Russell, and A.M. Lambowitz, *Do DEAD-box proteins promote group II intron splicing without unwinding RNA?* Mol Cell, 2007. **28**(1): p. 159-66.

DeVorkin, L. and S.M. Gorski, *LysoTracker staining to aid in monitoring autophagy in Drosophila.* Cold Spring Harb Protoc, 2014. **2014**(9): p. 951-8.

Di Cara, F. and K. King-Jones, *How clocks and hormones act in concert to control the timing of insect development.* Curr Top Dev Biol, 2013. **105**: p. 1-36.

Dunn, J.G., C.K. Foo, N.G. Belletier, E.R. Gavis, and J.S. Weissman, *Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster.* Elife, 2013. **2**: p. e01179.

Economopoulou, P., G. Dimitriadis, and A. Psyrri, *Beyond BRCA: new hereditary breast cancer susceptibility genes.* Cancer Treat Rev, 2015. **41**(1): p. 1-8.

Edgar, B.A., *How flies get their size: genetics meets physiology.* Nat Rev Genet, 2006. **7**(12): p. 907-16.

Edgcomb, S.P., A.B. Carmel, S. Naji, G. Ambrus-Aikelin, J.R. Reyes, A.C. Saphire, L. Gerace, and J.R. Williamson, *DDX1 is an RNA-dependent ATPase involved in HIV-1 Rev function and virus replication.* J Mol Biol, 2012. **415**(1): p. 61-74.

Erdi, B., P. Nagy, A. Zvara, A. Varga, K. Pircs, D. Menesi, L.G. Puskas, and G. Juhasz, *Loss of the starvation-induced gene Rack1 leads to glycogen deficiency and impaired autophagic responses in Drosophila.* Autophagy, 2012. **8**(7): p. 1124-35.

Erkizan, H.V., Y. Kong, M. Merchant, S. Schlottmann, J.S. Barber-Rotenberg, L. Yuan, O.D. Abaan, T.H. Chou, S. Dakshanamurthy, M.L. Brown, A. Uren, and J.A. Toretsky, *A small molecule blocking oncogenic protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma.* Nat Med, 2009. **15**(7): p. 750-6.

Fabian, L. and J.A. Brill, *Drosophila spermiogenesis: Big things come from little packages.* Spermatogenesis, 2012. **2**(3): p. 197-212.

Fan, C., D.S. Oh, L. Wessels, B. Weigelt, D.S. Nuyten, A.B. Nobel, L.J. van't Veer, and C.M. Perou, *Concordance among gene-expression-based predictors for breast cancer.* N Engl J Med, 2006. **355**(6): p. 560-9.

Fang, J., E. Acheampong, R. Dave, F. Wang, M. Mukhtar, and R.J. Pomerantz, *The RNA helicase DDX1 is involved in restricted HIV-1 Rev function in human astrocytes.* Virology, 2005. **336**(2): p. 299-307.

Fang, J., S. Kubota, B. Yang, N. Zhou, H. Zhang, R. Godbout, and R.J. Pomerantz, *A DEAD box protein facilitates HIV-1 replication as a cellular co-factor of Rev.* Virology, 2004. **330**(2): p. 471-80.

Ferguson-Smith, A.C. and M.E. Patti, *You are what your dad ate.* Cell Metab, 2011. **13**(2): p. 115-7.

Fernandez, R., D. Tabarini, N. Azpiazu, M. Frasch, and J. Schlessinger, *The Drosophila insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential.* EMBO J, 1995. **14**(14): p. 3373-84.

Flynn, P., M. Wongdagger, M. Zavar, N.M. Dean, and D. Stokoe, *Inhibition of PDK-1 activity causes a reduction in cell proliferation and survival.* Curr Biol, 2000. **10**(22): p. 1439-42.

Foe, V.E. and B.M. Alberts, *Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis.* J Cell Sci, 1983. **61**: p. 31-70.

Fossati, R., C. Confalonieri, V. Torri, E. Ghislandi, A. Penna, V. Pistotti, A. Tinazzi, and A. Liberati, *Cytotoxic and hormonal treatment for metastatic breast cancer: a systematic review of published randomized trials involving 31,510 women.* J Clin Oncol, 1998. **16**(10): p. 3439-60.

Freier, K., C. Flechtenmacher, F. Devens, W. Hartschuh, C. Hofele, P. Lichter, and S. Joos, *Recurrent NMYC copy number gain and high protein expression in basal cell carcinoma.* Oncol Rep, 2006. **15**(5): p. 1141-5.

Froberg, J.E., L. Yang, and J.T. Lee, *Guided by RNAs: X-inactivation as a model for lncRNA function.* J Mol Biol, 2013. **425**(19): p. 3698-706.

Fruhwald, M.C., M.S. O'Dorisio, L.J. Rush, J.L. Reiter, D.J. Smiraglia, G. Wenger, J.F. Costello, P.S. White, R. Krahe, G.M. Brodeur, and C. Plass, *Gene amplification in PNETs/medulloblastomas: mapping of a novel amplified gene within the MYCN amplicon.* J Med Genet, 2000. **37**(7): p. 501-9.

Frydman, H.M. and A.C. Spradling, *The receptor-like tyrosine phosphatase lar is required for epithelial planar polarity and for axis determination within drosophila ovarian follicles.* Development, 2001. **128**(16): p. 3209-20.

Fukuda, T., K. Yamagata, S. Fujiyama, T. Matsumoto, I. Koshida, K. Yoshimura, M. Mihara, M. Naitou, H. Endoh, T. Nakamura, C. Akimoto, Y. Yamamoto, T. Katagiri, C. Foulds, S. Takezawa, H. Kitagawa, K. Takeyama, B.W. O'Malley, and S. Kato, *DEAD-box RNA helicase subunits*

of the Drosha complex are required for processing of rRNA and a subset of microRNAs. Nat Cell Biol, 2007. **9**(5): p. 604-11.

Fuller-Pace, F.V., *RNA helicases: modulators of RNA structure.* Trends Cell Biol, 1994. **4**(8): p. 271-4.

Fuller-Pace, F.V. and S. Ali, *The DEAD box RNA helicases p68 (Ddx5) and p72 (Ddx17): novel transcriptional co-regulators.* Biochem Soc Trans, 2008. **36**(Pt 4): p. 609-12.

Gage, M., D. Wattendorf, and L.R. Henry, *Translational advances regarding hereditary breast cancer syndromes.* J Surg Oncol, 2012. **105**(5): p. 444-51.

Galicia-Vazquez, G., R. Cencic, F. Robert, A.Q. Agenor, and J. Pelletier, *A cellular response linking eIF4AI activity to eIF4AII transcription.* RNA, 2012. **18**(7): p. 1373-84.

Gatfield, D. and E. Izaurralde, *REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export.* J Cell Biol, 2002. **159**(4): p. 579-88.

Gatfield, D., H. Le Hir, C. Schmitt, I.C. Braun, T. Kocher, M. Wilm, and E. Izaurralde, *The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in Drosophila.* Curr Biol, 2001. **11**(21): p. 1716-21.

Geiduschek, E.P. and R. Haselkorn, *Messenger RNA.* Annu Rev Biochem, 1969. **38**: p. 647-76.

George, R.E., R.M. Kenyon, A.G. McGuckin, A.J. Malcolm, A.D. Pearson, and J. Lunec, *Investigation of co-amplification of the candidate genes ornithine decarboxylase, ribonucleotide reductase, syndecan-1 and a DEAD box gene, DDX1, with N-myc in neuroblastoma. United Kingdom Children's Cancer Study Group.* Oncogene, 1996. **12**(7): p. 1583-7.

Gerbasi, V.R., J.B. Preall, D.E. Golden, D.W. Powell, T.D. Cummins, and E.J. Sontheimer, *Blanks, a nuclear siRNA/dsRNA-binding complex component, is required for Drosophila spermiogenesis.* Proc Natl Acad Sci U S A, 2011. **108**(8): p. 3204-9.

Germain, D.R., K. Graham, D.D. Glubrecht, J.C. Hugh, J.R. Mackey, and R. Godbout, *DEAD box 1: a novel and independent prognostic marker for*

early recurrence in breast cancer. Breast Cancer Res Treat, 2011. **127**(1): p. 53-63.

Girard, A., R. Sachidanandam, G.J. Hannon, and M.A. Carmell, *A germline-specific class of small RNAs binds mammalian Piwi proteins.* Nature, 2006. **442**(7099): p. 199-202.

Giuliano, A.E., D.M. Kirgan, J.M. Guenther, and D.L. Morton, *Lymphatic mapping and sentinel lymphadenectomy for breast cancer.* Ann Surg, 1994. **220**(3): p. 391-8; discussion 398-401.

Godbout, R., L. Li, R.Z. Liu, and K. Roy, *Role of DEAD box 1 in retinoblastoma and neuroblastoma.* Future Oncol, 2007. **3**(5): p. 575-87.

Godbout, R., M. Packer, and W. Bie, *Overexpression of a DEAD box protein (DDX1) in neuroblastoma and retinoblastoma cell lines.* J Biol Chem, 1998. **273**(33): p. 21161-8.

Godbout, R., M. Packer, S. Katyal, and S. Bleoo, *Cloning and expression analysis of the chicken DEAD box gene DDX1.* Biochim Biophys Acta, 2002. **1574**(1): p. 63-71.

Godbout, R. and J. Squire, *Amplification of a DEAD box protein gene in retinoblastoma cell lines.* Proc Natl Acad Sci U S A, 1993a. **90**(16): p. 7578-82.

Godbout, R. and J. Squire, *Amplification of a DEAD box protein gene in retinoblastoma cell lines.* Proceedings of the National Academy of Sciences, 1993b. **90**(16): p. 7578-7582.

Gonzalez-Angulo, A.M., F. Morales-Vasquez, and G.N. Hortobagyi, *Overview of resistance to systemic therapy in patients with breast cancer.* Adv Exp Med Biol, 2007. **608**: p. 1-22.

Gorbalenya, A.E., E.V. Koonin, A.P. Donchenko, and V.M. Blinov, *A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination.* FEBS Lett, 1988. **235**(1-2): p. 16-24.

Grandjean, V., P. Gounon, N. Wagner, L. Martin, K.D. Wagner, F. Bernex, F. Cuzin, and M. Rassoulzadegan, *The miR-124-Sox9 paramutation: RNA-*

mediated epigenetic control of embryonic and adult growth. Development, 2009. **136**(21): p. 3647-55.

Gregory, R.I., K.P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj, N. Cooch, and R. Shiekhattar, *The Microprocessor complex mediates the genesis of microRNAs.* Nature, 2004. **432**(7014): p. 235-40.

Grewal, S.S., *Insulin/TOR signaling in growth and homeostasis: a view from the fly world.* Int J Biochem Cell Biol, 2009. **41**(5): p. 1006-10.

Grewal, S.S., J.R. Evans, and B.A. Edgar, *Drosophila TIF-IA is required for ribosome synthesis and cell growth and is regulated by the TOR pathway.* J Cell Biol, 2007. **179**(6): p. 1105-13.

Grifo, J.A., S.M. Tahara, J.P. Leis, M.A. Morgan, A.J. Shatkin, and W.C. Merrick, *Characterization of eukaryotic initiation factor 4A, a protein involved in ATP-dependent binding of globin mRNA.* J Biol Chem, 1982. **257**(9): p. 5246-52.

Gronke, S., D.F. Clarke, S. Broughton, T.D. Andrews, and L. Partridge, *Molecular evolution and functional characterization of Drosophila insulin-like peptides.* PLoS Genet, 2010. **6**(2): p. e1000857.

Guil, S., R. Gattoni, M. Carrascal, J. Abian, J. Stevenin, and M. Bach-Elias, *Roles of hnRNP A1, SR proteins, and p68 helicase in c-H-ras alternative splicing regulation.* Mol Cell Biol, 2003. **23**(8): p. 2927-41.

Gustafson, E.A. and G.M. Wessel, *Vasa genes: emerging roles in the germ line and in multipotent cells.* Bioessays, 2010. **32**(7): p. 626-37.

Hales, C.N. and D.J. Barker, *Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. 1992.* Int J Epidemiol, 2013. **42**(5): p. 1215-22.

Han, C., Y. Liu, G. Wan, H.J. Choi, L. Zhao, C. Ivan, X. He, A.K. Sood, X. Zhang, and X. Lu, *The RNA-binding protein DDX1 promotes primary microRNA maturation and inhibits ovarian tumor progression.* Cell Rep, 2014. **8**(5): p. 1447-60.

Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.

Hankinson, S.E., G.A. Colditz, and W.C. Willett, *Towards an integrated model for breast cancer etiology: the lifelong interplay of genes, lifestyle, and hormones.* Breast Cancer Res, 2004. **6**(5): p. 213-8.

Hara, K., K. Yonezawa, Q.P. Weng, M.T. Kozlowski, C. Belham, and J. Avruch, *Amino acid sufficiency and mTOR regulate p70 S6 kinase and eIF-4E BP1 through a common effector mechanism.* J Biol Chem, 1998. **273**(23): p. 14484-94.

Harris, R.S., S.K. Petersen-Mahrt, and M.S. Neuberger, *RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators.* Mol Cell, 2002. **10**(5): p. 1247-53.

Hay, B., L.Y. Jan, and Y.N. Jan, *A protein component of Drosophila polar granules is encoded by vasa and has extensive sequence similarity to ATP-dependent helicases.* Cell, 1988. **55**(4): p. 577-87.

Hay, N., *Interplay between FOXO, TOR, and Akt.* Biochim Biophys Acta, 2011. **1813**(11): p. 1965-70.

Hayat, M.J., N. Howlader, M.E. Reichman, and B.K. Edwards, *Cancer statistics, trends, and multiple primary cancer analyses from the Surveillance, Epidemiology, and End Results (SEER) Program.* Oncologist, 2007. **12**(1): p. 20-37.

Heard, E. and R.A. Martienssen, *Transgenerational epigenetic inheritance: myths and mechanisms.* Cell, 2014. **157**(1): p. 95-109.

Heine, V.M., M. Priller, J. Ling, D.H. Rowitch, and U. Schuller, *Dexamethasone destabilizes Nmyc to inhibit the growth of hedgehog-associated medulloblastoma.* Cancer Res, 2010. **70**(13): p. 5220-5.

Heyer, W.D., K.T. Ehmsen, and J. Liu, *Regulation of homologous recombination in eukaryotes.* Annu Rev Genet, 2010. **44**: p. 113-39.

Hibbard, K.L. and J.E. O'Tousa, *A role for the cytoplasmic DEAD box helicase Dbp21E2 in rhodopsin maturation and photoreceptor viability.* J Neurogenet, 2012. **26**(2): p. 177-88.

Hietakangas, V. and S.M. Cohen, *Re-evaluating AKT regulation: role of TOR complex 2 in tissue growth.* Genes Dev, 2007. **21**(6): p. 632-7.

Hillebrand, J., K. Pan, A. Kokaram, S. Barbee, R. Parker, and M. Ramaswami, *The Me31B DEAD-Box Helicase Localizes to Postsynaptic Foci and Regulates Expression of a CaMKII Reporter mRNA in Dendrites of Drosophila Olfactory Projection Neurons.* Front Neural Circuits, 2010. **4**: p. 121.

Hirling, H., M. Scheffner, T. Restle, and H. Stahl, *RNA helicase activity associated with the human p68 protein.* Nature, 1989. **339**(6225): p. 562-4.

Hodgson, J.G., R.F. Yeh, A. Ray, N.J. Wang, I. Smirnov, M. Yu, S. Hariono, J. Silber, H.S. Feiler, J.W. Gray, P.T. Spellman, S.R. Vandenberg, M.S. Berger, and C.D. James, *Comparative analyses of gene copy number and mRNA expression in glioblastoma multiforme tumors and xenografts.* Neuro Oncol, 2009. **11**(5): p. 477-87.

Honig, A., D. Auboeuf, M.M. Parker, B.W. O'Malley, and S.M. Berget, *Regulation of alternative splicing by the ATP-dependent DEAD-box RNA helicase p72.* Mol Cell Biol, 2002. **22**(16): p. 5698-707.

Hugh, J., J. Hanson, M.C. Cheang, T.O. Nielsen, C.M. Perou, C. Dumontet, J. Reed, M. Krajewska, I. Treilleux, M. Rupin, E. Magherini, J. Mackey, M. Martin, and C. Vogel, *Breast cancer subtypes and response to docetaxel in node-positive breast cancer: use of an immunohistochemical definition in the BCIRG 001 trial.* J Clin Oncol, 2009. **27**(8): p. 1168-76.

Ihry, R.J., A.L. Sapiro, and A. Bashirullah, *Translational control by the DEAD Box RNA helicase belle regulates ecdysone-triggered transcriptional cascades.* PLoS Genet, 2012. **8**(11): p. e1003085.

Inoki, K. and K.L. Guan, *Complexity of the TOR signaling network.* Trends Cell Biol, 2006. **16**(4): p. 206-12.

Inoue, A., T. Li, S.K. Roby, M.B. Valentine, M. Inoue, K. Boyd, V.J. Kidd, and J.M. Lahti, *Loss of ChlR1 helicase in mouse causes lethality due to the accumulation of aneuploid cells generated by cohesion defects and placental malformation.* Cell Cycle, 2007. **6**(13): p. 1646-54.

Irion, U. and M. Leptin, *Developmental and cell biological functions of the Drosophila DEAD-box protein abstrakt.* Curr Biol, 1999. **9**(23): p. 1373-81.

Irion, U., M. Leptin, K. Siller, S. Fuerstenberg, Y. Cai, C.Q. Doe, W. Chia, and X. Yang, *Abstrakt, a DEAD box protein, regulates Insc levels and*

asymmetric division of neural and mesodermal progenitors. Curr Biol, 2004. **14**(2): p. 138-44.

Jackman, J.E. and J.D. Alfonzo, *Transfer RNA modifications: nature's combinatorial chemistry playground.* Wiley Interdiscip Rev RNA, 2013. **4**(1): p. 35-48.

Jackson, R.J., C.U. Hellen, and T.V. Pestova, *The mechanism of eukaryotic translation initiation and principles of its regulation.* Nat Rev Mol Cell Biol, 2010. **11**(2): p. 113-27.

Jagut, M., L. Mihaila-Bodart, A. Molla-Herman, M.F. Alin, J.A. Lepesant, and J.R. Huynh, *A mosaic genetic screen for genes involved in the early steps of Drosophila oogenesis.* G3 (Bethesda), 2013. **3**(3): p. 409-25.

Jalal, C., H. Uhlmann-Schiffler, and H. Stahl, *Redundant role of DEAD box proteins p68 (Ddx5) and p72/p82 (Ddx17) in ribosome biogenesis and cell proliferation.* Nucleic Acids Res, 2007. **35**(11): p. 3590-601.

Jankowsky, E. and M.E. Fairman, *RNA helicases--one fold for many functions.* Curr Opin Struct Biol, 2007. **17**(3): p. 316-24.

Johanneson, B., D. Chen, S. Enroth, T. Cui, and U. Gyllensten, *Systematic validation of hypothesis-driven candidate genes for cervical cancer in a genome-wide association study.* Carcinogenesis, 2014. **35**(9): p. 2084-8.

Johnstone, O., R. Deuring, R. Bock, P. Linder, M.T. Fuller, and P. Lasko, *Belle is a Drosophila DEAD-box protein required for viability and in the germ line.* Dev Biol, 2005. **277**(1): p. 92-101.

Jones, S.E., *Metastatic breast cancer: the treatment challenge.* Clin Breast Cancer, 2008. **8**(3): p. 224-33.

Kanai, Y., N. Dohmae, and N. Hirokawa, *Kinesin transports RNA: isolation and characterization of an RNA-transporting granule.* Neuron, 2004. **43**(4): p. 513-25.

Kaneko, S., M. Ohira, Y. Nakamura, E. Isogai, A. Nakagawara, and M. Kaneko, *Relationship of DDX1 and NAG gene amplification/overexpression to the prognosis of patients with MYCN-amplified neuroblastoma.* J Cancer Res Clin Oncol, 2007. **133**(3): p. 185-92.

Kar, A., K. Fushimi, X. Zhou, P. Ray, C. Shi, X. Chen, Z. Liu, S. Chen, and J.Y. Wu, *RNA helicase p68 (DDX5) regulates tau exon 10 splicing by modulating a stem-loop structure at the 5' splice site.* Mol Cell Biol, 2011. **31**(9): p. 1812-21.

Kato, H., S. Sato, M. Yoneyama, M. Yamamoto, S. Uematsu, K. Matsui, T. Tsujimura, K. Takeda, T. Fujita, O. Takeuchi, and S. Akira, *Cell type-specific involvement of RIG-I in antiviral response.* Immunity, 2005. **23**(1): p. 19-28.

Kedersha, N., P. Ivanov, and P. Anderson, *Stress granules and cell signaling: more than just a passing phase?* Trends Biochem Sci, 2013. **38**(10): p. 494-506.

Kellner, J.N., J. Reinstein, and A. Meinhart, *Synergistic effects of ATP and RNA binding to human DEAD-box protein DDX1.* Nucleic Acids Res, 2015.

Kim, E., P. Goraksha-Hicks, L. Li, T.P. Neufeld, and K.L. Guan, *Regulation of TORC1 by Rag GTPases in nutrient response.* Nat Cell Biol, 2008. **10**(8): p. 935-45.

Kircher, S.G., S.H. Kim, M. Fountoulakis, and G. Lubec, *Reduced levels of DEAD-box proteins DBP-RB and p72 in fetal Down syndrome brains.* Neurochem Res, 2002. **27**(10): p. 1141-6.

Koromilas, A.E., A. Lazaris-Karatzas, and N. Sonenberg, *mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E.* EMBO J, 1992. **11**(11): p. 4153-8.

Kos, C.H., *Cre/loxP system for generating tissue-specific knockout mouse models.* Nutr Rev, 2004. **62**(6 Pt 1): p. 243-6.

Kotaja, N., S.N. Bhattacharyya, L. Jaskiewicz, S. Kimmins, M. Parvinen, W. Filipowicz, and P. Sassone-Corsi, *The chromatoid body of male germ cells: similarity with processing bodies and presence of Dicer and microRNA pathway components.* Proc Natl Acad Sci U S A, 2006. **103**(8): p. 2647-52.

Kreike, B., M. van Kouwenhove, H. Horlings, B. Weigelt, H. Peterse, H. Bartelink, and M.J. van de Vijver, *Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas.* Breast Cancer Res, 2007. **9**(5): p. R65.

Lamm, G.M., S.M. Nicol, F.V. Fuller-Pace, and A.I. Lamond, *p72: a human nuclear DEAD box protein highly related to p68.* Nucleic Acids Res, 1996. **24**(19): p. 3739-47.

Lasko, P., *The DEAD-box helicase Vasa: evidence for a multiplicity of functions in RNA processes and developmental biology.* Biochim Biophys Acta, 2013. **1829**(8): p. 810-6.

Lasko, P.F. and M. Ashburner, *Posterior localization of vasa protein correlates with, but is not sufficient for, pole cell development.* Genes Dev, 1990. **4**(6): p. 905-21.

Lee, C.G., *RH70, a bidirectional RNA helicase, co-purifies with U1snRNP.* J Biol Chem, 2002. **277**(42): p. 39679-83.

Lee, J.M., J.A. Ledermann, and E.C. Kohn, *PARP Inhibitors for BRCA1/2 mutation-associated and BRCA-like malignancies.* Ann Oncol, 2014. **25**(1): p. 32-40.

Lee, Y.S., Y. Shibata, A. Malhotra, and A. Dutta, *A novel class of small RNAs: tRNA-derived RNA fragments (tRFs).* Genes Dev, 2009. **23**(22): p. 2639-49.

Lehmann, M., *Drosophila Sgs genes: stage and tissue specificity of hormone responsiveness.* Bioessays, 1996. **18**(1): p. 47-54.

Lehmann, M. and G. Korge, *Ecdysone regulation of the Drosophila Sgs-4 gene is mediated by the synergistic action of ecdysone receptor and SEBP 3.* EMBO J, 1995. **14**(4): p. 716-26.

Li, J. and W.X. Li, *A novel function of Drosophila eIF4A as a negative regulator of Dpp/BMP signalling that mediates SMAD degradation.* Nat Cell Biol, 2006a. **8**(12): p. 1407-14.

Li, J., W.X. Li, and W.M. Gelbart, *A genetic screen for maternal-effect suppressors of decapentaplegic identifies the eukaryotic translation initiation factor 4A in Drosophila.* Genetics, 2005. **171**(4): p. 1629-41.

Li, L., E.A. Monckton, and R. Godbout, *A role for DEAD box 1 at DNA double-strand breaks.* Mol Cell Biol, 2008. **28**(20): p. 6413-25.

Li, L., K. Roy, S. Katyal, X. Sun, S. Bleoo, and R. Godbout, *Dynamic nature of cleavage bodies and their spatial relationship to DDX1 bodies, Cajal bodies, and gems.* Mol Biol Cell, 2006b. **17**(3): p. 1126-40.

Li, L.C. and R. Dahiya, *MethPrimer: designing primers for methylation PCRs.* Bioinformatics, 2002. **18**(11): p. 1427-31.

Li, Y., J.Z. Maines, O.Y. Tastan, D.M. McKearin, and M. Buszczak, *Mei-P26 regulates the maintenance of ovarian germline stem cells by promoting BMP signaling.* Development, 2012. **139**(9): p. 1547-56.

Lin, F., R. Wang, J.J. Shen, X. Wang, P. Gao, K. Dong, and H.Z. Zhang, *Knockdown of RCK/p54 expression by RNAi inhibits proliferation of human colorectal cancer cells in vitro and in vivo.* Cancer Biol Ther, 2008. **7**(10): p. 1669-76.

Lin, M.H., H. Sivakumaran, A. Jones, D. Li, C. Harper, T. Wei, H. Jin, L. Rustanti, F.A. Meunier, K. Spann, and D. Harrich, *A HIV-1 Tat mutant protein disrupts HIV-1 Rev function by targeting the DEAD-box RNA helicase DDX1.* Retrovirology, 2014. **11**(1): p. 121.

Linder, P., *Dead-box proteins: a family affair--active and passive players in RNP-remodeling.* Nucleic Acids Res, 2006. **34**(15): p. 4168-80.

Linder, P. and F.V. Fuller-Pace, *Looking back on the birth of DEAD-box RNA helicases.* Biochim Biophys Acta, 2013. **1829**(8): p. 750-5.

Linder, P. and E. Jankowsky, *From unwinding to clamping - the DEAD box RNA helicase family.* Nat Rev Mol Cell Biol, 2011. **12**(8): p. 505-16.

Linder, P., P.F. Lasko, M. Ashburner, P. Leroy, P.J. Nielsen, K. Nishi, J. Schnier, and P.P. Slonimski, *Birth of the D-E-A-D box.* Nature, 1989. **337**(6203): p. 121-2.

Linder, P. and F. Stutz, *mRNA export: travelling with DEAD box proteins.* Curr Biol, 2001. **11**(23): p. R961-3.

Liu, F., A. Putnam, and E. Jankowsky, *ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding.* Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20209-14.

Liu, L., H. Qi, J. Wang, and H. Lin, *PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition.* Development, 2011. **138**(9): p. 1863-73.

Liu, N., H. Han, and P. Lasko, *Vasa promotes Drosophila germline stem cell differentiation by activating mei-P26 translation by directly interacting with a (U)-rich motif in its 3' UTR.* Genes Dev, 2009. **23**(23): p. 2742-52.

Liu, Z.R., *p68 RNA helicase is an essential human splicing factor that acts at the U1 snRNA-5' splice site duplex.* Mol Cell Biol, 2002. **22**(15): p. 5443-50.

Lopez-Tarruella, S. and M. Martin, *Recent advances in systemic therapy: advances in adjuvant systemic chemotherapy of early breast cancer.* Breast Cancer Res, 2009. **11**(2): p. 204.

Lussana, F., R.C. Painter, M.C. Ocke, H.R. Buller, P.M. Bossuyt, and T.J. Roseboom, *Prenatal exposure to the Dutch famine is associated with a preference for fatty foods and a more atherogenic lipid profile.* Am J Clin Nutr, 2008. **88**(6): p. 1648-52.

Maehama, T. and J.E. Dixon, *PTEN: a tumour suppressor that functions as a phospholipid phosphatase.* Trends Cell Biol, 1999. **9**(4): p. 125-8.

Mahowald, A.P., *Assembly of the Drosophila germ plasm.* Int Rev Cytol, 2001. **203**: p. 187-213.

Mallam, A.L., M. Del Campo, B. Gilman, D.J. Sidote, and A.M. Lambowitz, *Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p.* Nature, 2012. **490**(7418): p. 121-5.

Mallona, I., A. Diez-Villanueva, and M.A. Peinado, *Methylation plotter: a web tool for dynamic visualization of DNA methylation data.* Source Code Biol Med, 2014. **9**: p. 11.

Malone, C.D., J. Brennecke, M. Dus, A. Stark, W.R. McCombie, R. Sachidanandam, and G.J. Hannon, *Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary.* Cell, 2009. **137**(3): p. 522-35.

Manohar, C.F., H.R. Salwen, G.M. Brodeur, and S.L. Cohn, *Co-amplification and concomitant high levels of expression of a DEAD box*

gene with MYCN in human neuroblastoma. Genes Chromosomes Cancer, 1995. **14**(3): p. 196-203.

Mao, Z., M. Bozzella, A. Seluanov, and V. Gorbunova, *DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells.* Cell Cycle, 2008. **7**(18): p. 2902-6.

Marshall, L., E.J. Rideout, and S.S. Grewal, *Nutrient/TOR-dependent regulation of RNA polymerase III controls tissue and organismal growth in Drosophila.* EMBO J, 2012. **31**(8): p. 1916-30.

McCall, K., *Eggs over easy: cell death in the Drosophila ovary.* Dev Biol, 2004. **274**(1): p. 3-14.

McLeod, C.J., L. Wang, C. Wong, and D.L. Jones, *Stem cell dynamics in response to nutrient availability.* Curr Biol, 2010. **20**(23): p. 2100-5.

McShane, L.M., D.G. Altman, W. Sauerbrei, S.E. Taube, M. Gion, and G.M. Clark, *REporting recommendations for tumor MARKer prognostic studies (REMARK).* Breast Cancer Res Treat, 2006. **100**(2): p. 229-35.

Megosh, H.B., D.N. Cox, C. Campbell, and H. Lin, *The role of PIWI and the miRNA machinery in Drosophila germline determination.* Curr Biol, 2006. **16**(19): p. 1884-94.

Meignin, C. and I. Davis, *UAP56 RNA helicase is required for axis specification and cytoplasmic mRNA localization in Drosophila.* Dev Biol, 2008. **315**(1): p. 89-98.

Meijer, H.A., Y.W. Kong, W.T. Lu, A. Wilczynska, R.V. Spriggs, S.W. Robinson, J.D. Godfrey, A.E. Willis, and M. Bushell, *Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation.* Science, 2013. **340**(6128): p. 82-5.

Messerschmidt, D.M., B.B. Knowles, and D. Solter, *DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos.* Genes Dev, 2014. **28**(8): p. 812-28.

Miller, L.C., V. Blandford, R. McAdam, M.R. Sanchez-Carbente, F. Badeaux, L. DesGroseillers, and W.S. Sossin, *Combinations of DEAD box proteins distinguish distinct types of RNA: protein complexes in neurons.* Mol Cell Neurosci, 2009. **40**(4): p. 485-95.

Miron, M., P. Lasko, and N. Sonenberg, *Signaling from Akt to FRAP/TOR targets both 4E-BP and S6K in Drosophila melanogaster.* Mol Cell Biol, 2003. **23**(24): p. 9117-26.

Montagne, J., M.J. Stewart, H. Stocker, E. Hafen, S.C. Kozma, and G. Thomas, *Drosophila S6 kinase: a regulator of cell size.* Science, 1999. **285**(5436): p. 2126-9.

Montpetit, B., M.A. Seeliger, and K. Weis, *Analysis of DEAD-box proteins in mRNA export.* Methods Enzymol, 2012. **511**: p. 239-54.

Mook, S., M.K. Schmidt, B. Weigelt, B. Kreike, I. Eekhout, M.J. van de Vijver, A.M. Glas, A. Floore, E.J. Rutgers, and L.J. van 't Veer, *The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age.* Ann Oncol, 2010. **21**(4): p. 717-22.

Morgan, T.H., *Sex Limited Inheritance in Drosophila.* Science, 1910. **32**(812): p. 120-2.

Morrissy, A.S., M. Griffith, and M.A. Marra, *Extensive relationship between antisense transcription and alternative splicing in the human genome.* Genome Res, 2011. **21**(8): p. 1203-12.

Mouillet, J.F., X. Yan, Q. Ou, L. Jin, L.J. Muglia, P.A. Crawford, and Y. Sadovsky, *DEAD-box protein-103 (DP103, Ddx20) is essential for early embryonic development and modulates ovarian morphology and function.* Endocrinology, 2008. **149**(5): p. 2168-75.

Mousseau, T.A. and C.W. Fox, *The adaptive significance of maternal effects.* Trends Ecol Evol, 1998. **13**(10): p. 403-7.

Naji, S., G. Ambrus, P. Cimermancic, J.R. Reyes, J.R. Johnson, R. Filbrandt, M.D. Huber, P. Vesely, N.J. Krogan, J.R. Yates, 3rd, A.C. Saphire, and L. Gerace, *Host cell interactome of HIV-1 Rev includes RNA helicases involved in multiple facets of virus production.* Mol Cell Proteomics, 2012. **11**(4): p. M111 015313.

Nakamura, A., R. Amikura, K. Hanyu, and S. Kobayashi, *Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis.* Development, 2001. **128**(17): p. 3233-42.

Neumuller, R.A., J. Betschinger, A. Fischer, N. Bushati, I. Poernbacher, K. Mechtler, S.M. Cohen, and J.A. Knoblich, *Mei-P26 regulates microRNAs and cell growth in the Drosophila ovarian stem cell lineage.* Nature, 2008. **454**(7201): p. 241-5.

Nicol, S.M., S.E. Bray, H.D. Black, S.A. Lorimore, E.G. Wright, D.P. Lane, D.W. Meek, P.J. Coates, and F.V. Fuller-Pace, *The RNA helicase p68 (DDX5) is selectively required for the induction of p53-dependent p21 expression and cell-cycle arrest after DNA damage.* Oncogene, 2013. **32**(29): p. 3461-9.

Nobukini, T. and G. Thomas, *The mTOR/S6K signalling pathway: the role of the TSC1/2 tumour suppressor complex and the proto-oncogene Rheb.* Novartis Found Symp, 2004. **262**: p. 148-54; discussion 154-9, 265-8.

Oakman, C., S. Bessi, E. Zafarana, F. Galardi, L. Biganzoli, and A. Di Leo, *Recent advances in systemic therapy: new diagnostics and biological predictors of outcome in early breast cancer.* Breast Cancer Res, 2009. **11**(2): p. 205.

Ogilvie, V.C., B.J. Wilson, S.M. Nicol, N.A. Morrice, L.R. Saunders, G.N. Barber, and F.V. Fuller-Pace, *The highly related DEAD box RNA helicases p68 and p72 exist as heterodimers in cells.* Nucleic Acids Res, 2003. **31**(5): p. 1470-80.

Oldham, S., *Obesity and nutrient sensing TOR pathway in flies and vertebrates: Functional conservation of genetic mechanisms.* Trends Endocrinol Metab, 2011. **22**(2): p. 45-52.

Oldham, S. and E. Hafen, *Insulin/IGF and target of rapamycin signaling: a TOR de force in growth control.* Trends Cell Biol, 2003. **13**(2): p. 79-85.

Onishi, H., Y. Kino, T. Morita, E. Futai, N. Sasagawa, and S. Ishiura, *MBNL1 associates with YB-1 in cytoplasmic stress granules.* J Neurosci Res, 2008. **86**(9): p. 1994-2002.

Painter, R.C., C. Osmond, P. Gluckman, M. Hanson, D.I. Phillips, and T.J. Roseboom, *Transgenerational effects of prenatal exposure to the Dutch famine on neonatal adiposity and health in later life.* BJOG, 2008. **115**(10): p. 1243-9.

Palacios, I.M., D. Gatfield, D. St Johnston, and E. Izaurralde, *An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay.* Nature, 2004. **427**(6976): p. 753-7.

Parish, J.L., J. Rosa, X. Wang, J.M. Lahti, S.J. Doxsey, and E.J. Androphy, *The DNA helicase ChlR1 is required for sister chromatid cohesion in mammalian cells.* J Cell Sci, 2006. **119**(Pt 23): p. 4857-65.

Parisi, F., S. Riccardo, M. Daniel, M. Saqcena, N. Kundu, A. Pession, D. Grifoni, H. Stocker, E. Tabak, and P. Bellosta, *Drosophila insulin and target of rapamycin (TOR) pathways regulate GSK3 beta activity to control Myc stability and determine Myc expression in vivo.* BMC Biol, 2011. **9**: p. 65.

Parker, J.S., M. Mullins, M.C. Cheang, S. Leung, D. Voduc, T. Vickery, S. Davies, C. Fauron, X. He, Z. Hu, J.F. Quackenbush, I.J. Stijleman, J. Palazzo, J.S. Marron, A.B. Nobel, E. Mardis, T.O. Nielsen, M.J. Ellis, C.M. Perou, and P.S. Bernard, *Supervised risk predictor of breast cancer based on intrinsic subtypes.* J Clin Oncol, 2009. **27**(8): p. 1160-7.

Patel, M.V., D.A. Hallal, J.W. Jones, D.N. Bronner, R. Zein, J. Caravas, Z. Husain, M. Friedrich, and M.F. Vanberkum, *Dramatic expansion and developmental expression diversification of the methuselah gene family during recent Drosophila evolution.* J Exp Zool B Mol Dev Evol, 2012. **318**(5): p. 368-87.

Patterson, L.F., M. Harvey, and P.F. Lasko, *Dbp73D, a Drosophila gene expressed in ovary, encodes a novel D-E-A-D box protein.* Nucleic Acids Res, 1992. **20**(12): p. 3063-7.

Pek, J.W. and T. Kai, *DEAD-box RNA helicase Belle/DDX3 and the RNA interference pathway promote mitotic chromosome segregation.* Proc Natl Acad Sci U S A, 2011a. **108**(29): p. 12007-12.

Pek, J.W. and T. Kai, *A role for vasa in regulating mitotic chromosome condensation in Drosophila.* Curr Biol, 2011b. **21**(1): p. 39-44.

Pelletier, J. and N. Sonenberg, *Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency.* Cell, 1985. **40**(3): p. 515-26.

Perez-Gonzalez, A., A. Pazo, R. Navajas, S. Ciordia, A. Rodriguez-Frandsen, and A. Nieto, *hCLE/C14orf166 associates with DDX1-HSPC117-*

FAM98B in a novel transcription-dependent shuttling RNA-transporting complex. PLoS One, 2014. **9**(3): p. e90957.

Perfetto, L., P.F. Gherardini, N.E. Davey, F. Diella, M. Helmer-Citterich, and G. Cesareni, *Exploring the diversity of SPRY/B30.2-mediated interactions.* Trends Biochem Sci, 2013. **38**(1): p. 38-46.

Perou, C.M., T. Sorlie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, C.A. Rees, J.R. Pollack, D.T. Ross, H. Johnsen, L.A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S.X. Zhu, P.E. Lonning, A.L. Borresen-Dale, P.O. Brown, and D. Botstein, *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.

Poltilove, R.M., A.R. Jacobs, C.R. Haft, P. Xu, and S.I. Taylor, *Characterization of Drosophila insulin receptor substrate.* J Biol Chem, 2000. **275**(30): p. 23346-54.

Popow, J., M. Englert, S. Weitzer, A. Schleiffer, B. Mierzwa, K. Mechtler, S. Trowitzsch, C.L. Will, R. Luhrmann, D. Soll, and J. Martinez, *HSPC117 is the essential subunit of a human tRNA splicing ligase complex.* Science, 2011. **331**(6018): p. 760-4.

Popow, J., J. Jurkin, A. Schleiffer, and J. Martinez, *Analysis of orthologous groups reveals archease and DDX1 as tRNA splicing factors.* Nature, 2014. **511**(7507): p. 104-7.

Porter, P., *"Westernizing" women's risks? Breast cancer in lower-income countries.* N Engl J Med, 2008. **358**(3): p. 213-6.

Potter, C.J., L.G. Pedraza, and T. Xu, *Akt regulates growth by directly phosphorylating Tsc2.* Nat Cell Biol, 2002. **4**(9): p. 658-65.

Probst, A.V., E. Dunleavy, and G. Almouzni, *Epigenetic inheritance during the cell cycle.* Nat Rev Mol Cell Biol, 2009. **10**(3): p. 192-206.

Puig, O., M.T. Marr, M.L. Ruhf, and R. Tjian, *Control of cell number by Drosophila FOXO: downstream and feedback regulation of the insulin receptor pathway.* Genes Dev, 2003. **17**(16): p. 2006-20.

Radford, E.J., E. Isganaitis, J. Jimenez-Chillaron, J. Schroeder, M. Molla, S. Andrews, N. Didier, M. Charalambous, K. McEwen, G. Marazzi, D. Sassoon, M.E. Patti, and A.C. Ferguson-Smith, *An unbiased assessment*

of the role of imprinted genes in an intergenerational model of developmental programming. PLoS Genet, 2012. **8**(4): p. e1002605.

Rafti, F., D. Scarvelis, and P.F. Lasko, *A Drosophila melanogaster homologue of the human DEAD-box gene DDX1.* Gene, 1996. **171**(2): p. 225-9.

Rakha, E.A., D.S. Tan, W.D. Foulkes, I.O. Ellis, A. Tutt, T.O. Nielsen, and J.S. Reis-Filho, *Are triple-negative tumours and basal-like breast cancer synonymous?* Breast Cancer Res, 2007. **9**(6): p. 404; author reply 405.

Rassoulzadegan, M., V. Grandjean, P. Gounon, S. Vincent, I. Gillot, and F. Cuzin, *RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse.* Nature, 2006. **441**(7092): p. 469-74.

Ratel, D., J.L. Ravanat, F. Berger, and D. Wion, *N6-methyladenine: the other methylated base of DNA.* Bioessays, 2006. **28**(3): p. 309-15.

Rhodes, D.A., B. de Bono, and J. Trowsdale, *Relationship between SPRY and B30.2 protein domains. Evolution of a component of immune defence?* Immunology, 2005. **116**(4): p. 411-7.

Rich, A. and U.L. RajBhandary, *Transfer RNA: molecular structure, sequence, and properties.* Annu Rev Biochem, 1976. **45**: p. 805-60.

Rideout, E.J., L. Marshall, and S.S. Grewal, *Drosophila RNA polymerase III repressor Maf1 controls body size and developmental timing by modulating tRNAiMet synthesis and systemic insulin signaling.* Proc Natl Acad Sci U S A, 2012. **109**(4): p. 1139-44.

Ries, L., M. Smith, J. Gurney, M. Linet, T. Tamra, J. Young, and G. Bunin, *Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975-1995.* SEER Monograph, 1999.

Roberts, R.J., M.O. Carneiro, and M.C. Schatz, *The advantages of SMRT sequencing.* Genome Biol, 2013. **14**(7): p. 405.

Robertson-Anderson, R.M., J. Wang, S.P. Edgcomb, A.B. Carmel, J.R. Williamson, and D.P. Millar, *Single-molecule studies reveal that DEAD box protein DDX1 promotes oligomerization of HIV-1 Rev on the Rev response element.* J Mol Biol, 2011. **410**(5): p. 959-71.

Rogers, G.W., Jr., N.J. Richter, and W.C. Merrick, *Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A.* J Biol Chem, 1999. **274**(18): p. 12236-44.

Rohde, C., Y. Zhang, R. Reinhardt, and A. Jeltsch, *BISMA--fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences.* BMC Bioinformatics, 2010. **11**: p. 230.

Rossler, O.G., A. Straka, and H. Stahl, *Rearrangement of structured RNA via branch migration structures catalysed by the highly related DEAD-box proteins p68 and p72.* Nucleic Acids Res, 2001. **29**(10): p. 2088-96.

Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection.* Nat Rev Cancer, 2012. **12**(1): p. 68-78.

Rozen, F., I. Edery, K. Meerovitch, T.E. Dever, W.C. Merrick, and N. Sonenberg, *Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F.* Mol Cell Biol, 1990. **10**(3): p. 1134-44.

Rushlow, D.E., B.M. Mol, J.Y. Kennett, S. Yee, S. Pajovic, B.L. Theriault, N.L. Prigoda-Lee, C. Spencer, H. Dimaras, T.W. Corson, R. Pang, C. Massey, R. Godbout, Z. Jiang, E. Zacksenhaus, K. Paton, A.C. Moll, C. Houdayer, A. Raizis, W. Halliday, W.L. Lam, P.C. Boutros, D. Lohmann, J.C. Dorsman, and B.L. Gallie, *Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies.* Lancet Oncol, 2013. **14**(4): p. 327-34.

Sakurai, H., S. Takai, K. Kawamura, Y. Ogura, Y. Yoshioka, and K. Kawasaki, *Drosophila RecQ5 is involved in proper progression of early spermatogenesis.* Biochem Biophys Res Commun, 2014. **452**(4): p. 1071-7.

Salzman, D.W., J. Shubert-Coleman, and H. Furneaux, *P68 RNA helicase unwinds the human let-7 microRNA precursor duplex and is required for let-7-directed silencing of gene expression.* J Biol Chem, 2007. **282**(45): p. 32773-9.

Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning : a laboratory manual*. 2nd ed. 1989, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.

Savic, N., D. Bar, S. Leone, S.C. Frommel, F.A. Weber, E. Vollenweider, E. Ferrari, U. Ziegler, A. Kaech, O. Shakhova, P. Cinelli, and R. Santoro, *lncRNA maturation to initiate heterochromatin formation in the nucleolus is required for exit from pluripotency in ESCs.* Cell Stem Cell, 2014. **15**(6): p. 720-34.

Scanga, S.E., L. Ruel, R.C. Binari, B. Snow, V. Stambolic, D. Bouchard, M. Peters, B. Calvieri, T.W. Mak, J.R. Woodgett, and A.S. Manoukian, *The conserved PI3'K/PTEN/Akt signaling pathway regulates both cell size and survival in Drosophila.* Oncogene, 2000. **19**(35): p. 3971-7.

Schiavon, G. and I.E. Smith, *Status of adjuvant endocrine therapy for breast cancer.* Breast Cancer Res, 2014. **16**(2): p. 206.

Schmucker, D., G. Vorbruggen, P. Yeghiayan, H.Q. Fan, H. Jackle, and U. Gaul, *The Drosophila gene abstrakt, required for visual system development, encodes a putative RNA helicase of the DEAD box protein family.* Mech Dev, 2000. **91**(1-2): p. 189-96.

Schneider, B.P. and G.W. Sledge, Jr., *Anti-vascular endothelial growth factor therapy for breast cancer: can we pick the winners?* J Clin Oncol, 2011. **29**(18): p. 2444-7.

Schupbach, T. and E. Wieschaus, *Germline autonomy of maternal-effect mutations altering the embryonic body pattern of Drosophila.* Dev Biol, 1986. **113**(2): p. 443-8.

Schweet, R. and R. Heintz, *Protein synthesis.* Annu Rev Biochem, 1966. **35**: p. 723-58.

Sengoku, T., O. Nureki, A. Nakamura, S. Kobayashi, and S. Yokoyama, *Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa.* Cell, 2006. **125**(2): p. 287-300.

Seong, K.H., D. Li, H. Shimizu, R. Nakamura, and S. Ishii, *Inheritance of stress-induced, ATF-2-dependent epigenetic change.* Cell, 2011. **145**(7): p. 1049-61.

Shen, R., C. Weng, J. Yu, and T. Xie, *eIF4A controls germline stem cell self-renewal by directly inhibiting BAM function in the Drosophila ovary.* Proc Natl Acad Sci U S A, 2009a. **106**(28): p. 11623-8.

Shen, X.C., C.X. Gu, Y.Q. Qiu, C.J. Du, Y.B. Fu, and J.J. Wu, *Estrogen receptor expression in adrenocortical carcinoma.* J Zhejiang Univ Sci B, 2009b. **10**(1): p. 1-6.

Sheng, Y., C.H. Tsai-Morris, R. Gutti, Y. Maeda, and M.L. Dufau, *Gonadotropin-regulated testicular RNA helicase (GRTH/Ddx25) is a transport protein involved in gene-specific mRNA export and protein translation during spermatogenesis.* J Biol Chem, 2006. **281**(46): p. 35048-56.

Shpargel, K.B., K. Praveen, T.K. Rajendra, and A.G. Matera, *Gemin3 is an essential gene required for larval motor function and pupation in Drosophila.* Mol Biol Cell, 2009. **20**(1): p. 90-101.

Singleton, M.R., M.S. Dillingham, and D.B. Wigley, *Structure and mechanism of helicases and nucleic acid translocases.* Annu Rev Biochem, 2007. **76**: p. 23-50.

Sissung, T.M., D.K. Price, M. Del Re, A.M. Ley, E. Giovannetti, W.D. Figg, and R. Danesi, *Genetic variation: effect on prostate cancer.* Biochim Biophys Acta, 2014. **1846**(2): p. 446-56.

Sledge, G.W., E.P. Mamounas, G.N. Hortobagyi, H.J. Burstein, P.J. Goodwin, and A.C. Wolff, *Past, present, and future challenges in breast cancer treatment.* J Clin Oncol, 2014. **32**(19): p. 1979-86.

Slodkowska, E.A. and J.S. Ross, *MammaPrint 70-gene signature: another milestone in personalized medical care for breast cancer patients.* Expert Rev Mol Diagn, 2009. **9**(5): p. 417-22.

Smolonska, J., G.H. Koppelman, C. Wijmenga, J.M. Vonk, P. Zanen, M. Bruinenberg, I. Curjuric, M. Imboden, G.A. Thun, L. Franke, N.M. Probst-Hensch, P. Nurnberg, R.A. Riemersma, C.P. van Schayck, D.W. Loth, G.G. Brusselle, B.H. Stricker, A. Hofman, A.G. Uitterlinden, L. Lahousse, S.J. London, L.R. Loehr, A. Manichaikul, R.G. Barr, K.M. Donohue, S.S. Rich, P. Pare, Y. Bosse, K. Hao, M. van den Berge, H.J. Groen, J.W. Lammers, W. Mali, H.M. Boezen, and D.S. Postma, *Common genes underlying asthma and COPD? Genome-wide analysis on the Dutch hypothesis.* Eur Respir J, 2014. **44**(4): p. 860-72.

Squire, J.A., P.S. Thorner, S. Weitzman, J.D. Maggi, P. Dirks, J. Doyle, M. Hale, and R. Godbout, *Co-amplification of MYCN and a DEAD box gene (DDX1) in primary neuroblastoma.* Oncogene, 1995. **10**(7): p. 1417-22.

St Pierre, S.E., L. Ponting, R. Stefancsik, and P. McQuilton, *FlyBase 102--advanced approaches to interrogating FlyBase.* Nucleic Acids Res, 2014. **42**(Database issue): p. D780-8.

Stadler, Z.K. and S.E. Come, *Review of gene-expression profiling and its clinical use in breast cancer.* Crit Rev Oncol Hematol, 2009. **69**(1): p. 1-11.

Statistics Canada, *CANSIM Table 102-0522*. 2011.

Story, R.M., H. Li, and J.N. Abelson, *Crystal structure of a DEAD box protein from the hyperthermophile Methanococcus jannaschii.* Proc Natl Acad Sci U S A, 2001. **98**(4): p. 1465-70.

Styhler, S., A. Nakamura, A. Swan, B. Suter, and P. Lasko, *vasa is required for GURKEN accumulation in the oocyte, and is involved in oocyte differentiation and germline cyst development.* Development, 1998. **125**(9): p. 1569-78.

Sugimoto, K., E. Koh, H.S. Sin, Y. Maeda, K. Narimoto, K. Izumi, Y. Kobori, E. Kitamura, H. Nagase, A. Yoshida, and M. Namiki, *Tissue-specific differentially methylated regions of the human VASA gene are potentially associated with maturation arrest phenotype in the testis.* J Hum Genet, 2009. **54**(8): p. 450-6.

Sunden, Y., S. Semba, T. Suzuki, Y. Okada, Y. Orba, K. Nagashima, T. Umemura, and H. Sawa, *DDX1 promotes proliferation of the JC virus through transactivation of its promoter.* Microbiol Immunol, 2007a. **51**(3): p. 339-47.

Sunden, Y., S. Semba, T. Suzuki, Y. Okada, Y. Orba, K. Nagashima, T. Umemura, and H. Sawa, *Identification of DDX1 as a JC virus transcriptional control region-binding protein.* Microbiol Immunol, 2007b. **51**(3): p. 327-37.

Tanaka, K., S. Okamoto, Y. Ishikawa, H. Tamura, and T. Hara, *DDX1 is required for testicular tumorigenesis, partially through the transcriptional activation of 12p stem cell genes.* Oncogene, 2009. **28**(21): p. 2142-51.

Tanaka, S.S., Y. Toyooka, R. Akasu, Y. Katoh-Fukui, Y. Nakahara, R. Suzuki, M. Yokoyama, and T. Noce, *The mouse homolog of Drosophila Vasa is required for the development of male germ cells.* Genes Dev, 2000. **14**(7): p. 841-53.

Tanner, N.K. and P. Linder, *DExD/H box RNA helicases: from generic motors to specific dissociation functions.* Mol Cell, 2001. **8**(2): p. 251-62.

Taunk, N.K., S. Goyal, H. Wu, M.S. Moran, S. Chen, and B.G. Haffty, *DEAD box 1 (DDX1) expression predicts for local control and overall survival in early stage, node-negative breast cancer.* Cancer, 2012. **118**(4): p. 888-98.

Teitz, T., J.J. Stanke, S. Federico, C.L. Bradley, R. Brennan, J. Zhang, M.D. Johnson, J. Sedlacik, M. Inoue, Z.M. Zhang, S. Frase, J.E. Rehg, C.M. Hillenbrand, D. Finkelstein, C. Calabrese, M.A. Dyer, and J.M. Lahti, *Preclinical models for neuroblastoma: establishing a baseline for treatment.* PLoS One, 2011. **6**(4): p. e19133.

Teplinsky, E. and F. Muggia, *Targeting HER2 in ovarian and uterine cancers: challenges and future directions.* Gynecol Oncol, 2014. **135**(2): p. 364-70.

Theriault, B.L., H. Dimaras, B.L. Gallie, and T.W. Corson, *The genomic landscape of retinoblastoma: a review.* Clin Experiment Ophthalmol, 2014. **42**(1): p. 33-52.

Thomson, T., N. Liu, A. Arkov, R. Lehmann, and P. Lasko, *Isolation of new polar granule components in Drosophila reveals P body and ER associated proteins.* Mech Dev, 2008. **125**(9-10): p. 865-73.

Tingting, P., F. Caiyun, Y. Zhigang, Y. Pengyuan, and Y. Zhenghong, *Subproteomic analysis of the cellular proteins associated with the 3' untranslated region of the hepatitis C virus genome in human liver cells.* Biochem Biophys Res Commun, 2006. **347**(3): p. 683-91.

Tokuyasu, K.T., *Dynamics of spermiogenesis in Drosophila melanogaster. VI. Significance of "onion" nebenkern formation.* J Ultrastruct Res, 1975. **53**(1): p. 93-112.

Tokuyasu, K.T., W.J. Peacock, and R.W. Hardy, *Dynamics of spermiogenesis in Drosophila melanogaster. I. Individualization process.* Z Zellforsch Mikrosk Anat, 1972a. **124**(4): p. 479-506.

Tokuyasu, K.T., W.J. Peacock, and R.W. Hardy, *Dynamics of spermiogenesis in Drosophila melanogaster. II. Coiling process.* Z Zellforsch Mikrosk Anat, 1972b. **127**(4): p. 492-525.

Tomancak, P., A. Guichet, P. Zavorszky, and A. Ephrussi, *Oocyte polarity depends on regulation of gurken by Vasa.* Development, 1998. **125**(9): p. 1723-32.

Tran, E.J., Y. Zhou, A.H. Corbett, and S.R. Wente, *The DEAD-box protein Dbp5 controls mRNA export by triggering specific RNA:protein remodeling events.* Mol Cell, 2007. **28**(5): p. 850-9.

Tsai-Morris, C.H., S. Lei, Q. Jiang, Y. Sheng, and M.L. Dufau, *Genomic organization and transcriptional analysis of gonadotropin-regulated testicular RNA helicase--GRTH/DDX25 gene.* Gene, 2004a. **331**: p. 83-94.

Tsai-Morris, C.H., Y. Sheng, E. Lee, K.J. Lei, and M.L. Dufau, *Gonadotropin-regulated testicular RNA helicase (GRTH/Ddx25) is essential for spermatid development and completion of spermatogenesis.* Proc Natl Acad Sci U S A, 2004b. **101**(17): p. 6373-8.

Uhlmann-Schiffler, H., O.G. Rossler, and H. Stahl, *The mRNA of DEAD box protein p72 is alternatively translated into an 82-kDa RNA helicase.* J Biol Chem, 2002. **277**(2): p. 1066-75.

Van Vranken, J.G., D.K. Bricker, N. Dephoure, S.P. Gygi, J.E. Cox, C.S. Thummel, and J. Rutter, *SDHAF4 promotes mitochondrial succinate dehydrogenase activity and prevents neurodegeneration.* Cell Metab, 2014. **20**(2): p. 241-52.

van Wolfswinkel, J.C., J.M. Claycomb, P.J. Batista, C.C. Mello, E. Berezikov, and R.F. Ketting, *CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs.* Cell, 2009. **139**(1): p. 135-48.

Volarevic, S. and G. Thomas, *Role of S6 phosphorylation and S6 kinase in cell growth.* Prog Nucleic Acid Res Mol Biol, 2001. **65**: p. 101-27.

Wagner, E. and J. Lykke-Andersen, *mRNA surveillance: the perfect persist.* J Cell Sci, 2002. **115**(Pt 15): p. 3033-8.

Wagner, K.D., N. Wagner, H. Ghanbarian, V. Grandjean, P. Gounon, F. Cuzin, and M. Rassoulzadegan, *RNA induction and inheritance of epigenetic cardiac hypertrophy in the mouse.* Dev Cell, 2008. **14**(6): p. 962-9.

Wagner, M., R. Rid, C.J. Maier, R.H. Maier, M. Laimer, H. Hintner, J.W. Bauer, and K. Onder, *DDX5 is a multifunctional co-activator of steroid hormone receptors.* Mol Cell Endocrinol, 2012. **361**(1-2): p. 80-91.

Wang, Y., H.X. Zhang, Y.P. Sun, Z.X. Liu, X.S. Liu, L. Wang, S.Y. Lu, H. Kong, Q.L. Liu, X.H. Li, Z.Y. Lu, S.J. Chen, Z. Chen, S.S. Bao, W. Dai, and Z.G. Wang, *Rig-I-/- mice develop colitis associated with downregulation of G alpha i2.* Cell Res, 2007. **17**(10): p. 858-68.

Weber, A., P. Imisch, E. Bergmann, and H. Christiansen, *Coamplification of DDX1 correlates with an improved survival probability in children with MYCN-amplified human neuroblastoma.* J Clin Oncol, 2004. **22**(13): p. 2681-90.

Werner, H.M. and H.B. Salvesen, *Current status of molecular biomarkers in endometrial cancer.* Curr Oncol Rep, 2014. **16**(9): p. 403.

Williams-Hill, D.M., R.F. Duncan, P.J. Nielsen, and S.M. Tahara, *Differential expression of the murine eukaryotic translation initiation factor isogenes eIF4A(I) and eIF4A(II) is dependent upon cellular growth status.* Arch Biochem Biophys, 1997. **338**(1): p. 111-20.

Wilson, B.J., G.J. Bates, S.M. Nicol, D.J. Gregory, N.D. Perkins, and F.V. Fuller-Pace, *The p68 and p72 DEAD box RNA helicases interact with HDAC1 and repress transcription in a promoter-specific manner.* BMC Mol Biol, 2004. **5**: p. 11.

Wimmer, K., X.X. Zhu, B.J. Lamb, R. Kuick, P.F. Ambros, H. Kovar, D. Thoraval, S. Motyka, J.R. Alberts, and S.M. Hanash, *Co-amplification of a novel gene, NAG, with the N-myc gene in neuroblastoma.* Oncogene, 1999. **18**(1): p. 233-8.

Wong, H.Y., J.A. Demmers, K. Bezstarosti, J.A. Grootegoed, and A.O. Brinkmann, *DNA dependent recruitment of DDX17 and other interacting proteins by the human androgen receptor.* Biochim Biophys Acta, 2009. **1794**(2): p. 193-8.

Wortham, N.C., E. Ahamed, S.M. Nicol, R.S. Thomas, M. Periyasamy, J. Jiang, A.M. Ochocka, S. Shousha, L. Huson, S.E. Bray, R.C. Coombes, S. Ali, and F.V. Fuller-Pace, *The DEAD-box protein p72 regulates ERalpha-/oestrogen-dependent transcription and cell growth, and is associated with improved survival in ERalpha-positive breast cancer.* Oncogene, 2009. **28**(46): p. 4053-64.

Wu, C.H., P.J. Chen, and S.H. Yeh, *Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to continuous transcription.* Cell Host Microbe, 2014. **16**(4): p. 462-72.

Xu, L., S. Khadijah, S. Fang, L. Wang, F.P. Tay, and D.X. Liu, *The cellular RNA helicase DDX1 interacts with coronavirus nonstructural protein 14 and enhances viral replication.* J Virol, 2010. **84**(17): p. 8571-83.

Yang, Q., M. Del Campo, A.M. Lambowitz, and E. Jankowsky, *DEAD-box proteins unwind duplexes by local strand separation.* Mol Cell, 2007. **28**(2): p. 253-63.

Yang, Q., K. Inoki, E. Kim, and K.L. Guan, *TSC1/TSC2 and Rheb have different effects on TORC1 and TORC2 activity.* Proc Natl Acad Sci U S A, 2006. **103**(18): p. 6811-6.

Yasuda-Inoue, M., M. Kuroki, and Y. Ariumi, *Distinct DDX DEAD-box RNA helicases cooperate to modulate the HIV-1 Rev function.* Biochem Biophys Res Commun, 2013. **434**(4): p. 803-8.

Yedavalli, V.S., C. Neuveut, Y.H. Chi, L. Kleiman, and K.T. Jeang, *Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function.* Cell, 2004. **119**(3): p. 381-92.

Zaffran, S., A. Chartier, P. Gallant, M. Astier, N. Arquier, D. Doherty, D. Gratecos, and M. Semeriva, *A Drosophila RNA helicase gene, pitchoune, is required for cell growth and proliferation and is a potential target of d-Myc.* Development, 1998. **125**(18): p. 3571-84.

Zeng, F., D.A. Baldwin, and R.M. Schultz, *Transcript profiling during preimplantation mouse development.* Dev Biol, 2004. **272**(2): p. 483-96.

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, *UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery.* Cell, 2012. **151**(4): p. 871-84.

Zhang, H., J.P. Stallock, J.C. Ng, C. Reinhard, and T.P. Neufeld, *Regulation of cellular growth by the Drosophila target of rapamycin dTOR.* Genes Dev, 2000. **14**(21): p. 2712-24.

Zhang, Y., X. Gao, L.J. Saucedo, B. Ru, B.A. Edgar, and D. Pan, *Rheb is a direct target of the tuberous sclerosis tumour suppressor proteins.* Nat Cell Biol, 2003. **5**(6): p. 578-81.

Zhang, Z. and A.R. Krainer, *Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components.* Proc Natl Acad Sci U S A, 2007. **104**(28): p. 11574-9.

Zinsmaier, K.E., K.K. Eberle, E. Buchner, N. Walter, and S. Benzer, *Paralysis and early death in cysteine string protein mutants of Drosophila.* Science, 1994. **263**(5149): p. 977-80.

APPENDIX

Modulation of Ddx1 levels in Drosophila

The following fly lines were utilized in an attempt to modulate Ddx1 RNA and protein levels: *w1118; P{GD14106}v29017* (Ddx1 specific RNAi), *w[*]; P{w[+m*]=Ubi-GAL4.U}2/CyO* (GAL4 expressed in all cells, *Ubi*-driver), *P{w[+mC]=GAL4-elav.L}2/CyO* (GAL4 expressed in nervous tissue cells, *elav*-driver), *w[1118]; P{y[+t7.7] w[+mC]=GMR44E09-GAL4}attP2* (GAL4 expressed in the dorsal half of the wing, *ap*-driver) *and y[1] w[*]; P{w[+mC]=tubP-GAL4}LL7/TM3, Sb[1] Ser[1]* (expresses GAL4 ubiquitously, *tub*-driver). Transgenic *Ddx1* over-expression lines were generated by cloning *Ddx1* cDNA into *Drosophila* Gateway Vectors pPWG (C-terminal GFP tag), pPMW (C-terminal HA tag) and pHS (heast-shock promoter), with Bestgene Inc. (Chino Hills, US) microinjecting the plasmid into w^{1118} embryos along with a Δ2-3 plasmid to facilitate insertion. Of the lines generated, only strain 1G displayed phenotypes when over-expressed. Strain 1G was generated using the *Ddx1*-pPWG construct and contained a transgenic insertion on the X chromosome; however, sequencing revealed a point mutation that introduced a stop codon in the linker region between the Ddx1 ORF and the GFP tag resulting in expression of the full length *Ddx1* ORF followed by a short (< 10 aa) sequence.

Modulation of *Ddx1* levels was attempted by crossing the transgenic over-expresser line (1G) or the RNAi knockdown line to GAL4 drivers. A number of phenotypes were observed. For *Ddx1* knock-down, phenotypes were only observed in conjunction with the *ap*-GAL4 driver, resulting in loss of a portion of

the thoraxic bristles, and upward curving of the wing. Over-expression driven by the *ap*-GAL4 driver resulted in downward curving of the wing. Aberrant wing vein formation, and gross morphological changes to imaginal wing discs were also observed in *ap*-GAL4 driven over-expression flies. Both knockdown and over-expression of Ddx1 under the *ap*-GAL4 driver resulted in reduced wing size. Over-expression in neuronal tissue (*elav*-GAL4 driver) or at higher levels ubiquitously (*tub*-GAL4 driver) resulted in 100% lethality. Over-expression at a moderate level ubiquitously (*ubi*-GAL4 driver) resulted in male-specific death and failure to eclose in a sub-population of female pupae. It should be noted that modulated *Ddx1* levels were not confirmed via RT-PCR or western blot analysis. As a result, we cannot state with certainty that the effects are the result of *Ddx1* modulation as opposed to off-target effects.