DIFFERENTIAL ACTIVITIES AND GENE REGULATION BY THE DROSOPHILA RETINOBLASTOMA PROTEINS

By

Rima Mouawad

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Cell and Molecular Biology – Doctor of Philosophy

2019

ABSTRACT

DIFFERENTIAL ACTIVITIES AND GENE REGULATION BY THE DROSOPHILA RETINOBLASTOMA PROTEINS

By

Rima Mouawad

The Retinoblastoma proteins are a family of transcriptional co-repressors that play important roles in regulating various cellular processes, including cell proliferation, differentiation and apoptosis. There are three retinoblastoma family proteins in humans; Rb, p107 and p130, which perform overlapping yet distinct roles in regulating the transcription of genes involved in diverse processes. The molecular mechanisms underlying the distinct functions of the retinoblastoma proteins are not fully understood, but may include differences in complex formation with E2f factors as well as interactions with distinct chromatin modifying and remodeling factors. Gene duplication is rare in the retinoblastoma gene family; most metazoan genomes encode a single retinoblastoma protein. Interestingly, Drosophila has undergone an independent duplication event that yielded an ancestral *Rbf1* gene and a more derived *Rbf2* gene, which makes it a good system to study paralogy in this family. Structurally, Rbf1 resembles p107 and p130 most closely, and bears a conserved regulatory C-terminal domain (IE) that is critical for stability and activity of these proteins. The IE is lost in Rb and Rbf2, which are the more derived forms of the protein, suggesting that the presence or absence of the IE may contribute to the different functions of these proteins. Here, using genomic approaches, I provide new insights on the function of Rbf2, which were not apparent from prior cell-based assays. I show that Rbf2 regulates a set of cell growth related genes, and has an impact on fertility and lifespan of flies. I define cis regulatory features of *CycB* gene that allow preferential repression by Rbf2, indicating that it is not merely a weaker version of *Rbf1* as previously thought, but a highly effective repressor in certain contexts. I furthermore show using transcriptomic studies

that the IE of Rbf1 is critical for the repression function of this protein in a gene specific manner. Mutation of specific conserved residues within the IE have a distinct impacts on subsets of genes, indicating that the IE is an important regulatory element for specific sets of genes. The specialization of retinoblastoma function in Drosophila may reflect a parallel evolution found in vertebrates, and raises the possibility that control of cell growth control is equally important to cell cycle function for this conserved family of transcriptional corepressors.

This thesis is dedicated to my parents and my husband.
Thank you for always believing in me.

ACKNOWLEDGMENTS

I would like to thank my remarkable mentor, Dr. David Arnosti for his guidance, advice and support. David taught me how to be a critical thinker and a good scientist, and he inspired me to be more ambitious and motivated to continue my research career. David is a great scientist, a perfect example of a leader, a friend, and working with him was one of the most rewarding experiences I have ever had.

I would like to thank Dr. Bill Henry for his advice and support. Bill played a significant role in my PhD and in my career pursuit. He is a wonderful person and a great scientist. I would like to thank my committee members Dr. Eran Andrechek and Dr. George Mias for their advice and guidance. I would also like to thank my former committee member Dr. Michelle Fluck who opened the door for me to get into the PhD program at MSU. She was a wonderful person and an excellent scientist.

I would like to thank former and current members of the Arnosti lab especially my friend Dr. Sandhya Payankaulam who helped me with everything since I joined the lab. I would also like to thank all the undergrads who contributed significantly to my work.

I would like to thank Dr. Susan Conrad, the CMB director, for her guidance and support. I would like to thank Jon Rennhack (Andrechek lab) and Nicholas Panchy (Shui lab) for their help in the bioinformatics part of my research.

Finally, I would like to thank my parents Samir and Nina Mouawad, my siblings Ramzi, Rami and Rita, my husband Michel Abou Rjeileh and his family, and my friends and family for their constant love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

aPKC	Atypical protein kinase C
Arp53D	Actin-related protein 53D
Bcl-2	B-cell lymphoma 2
BRG1	Brahma-related gene-1
Cdk	Cyclin-dependent Kinase
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation sequencing
CRISPR	Cluster regularly interspaced short palindromic repeats
dCAP-D3	Drosophila chromosome associated protein D3
Debcl	Death executioner Bcl-2
Diap1	Death-associated inhibitor of apoptosis 1
DNMT	DNA methyltransferase
DP	E2f dimerization partner
DREF	DNA replication-related element factor
Drp1	Dynamin related protein 1
E1A	Early region 1A
E2F	E2 promoter binding factor
E3	Ubiquitin ligase enzyme 3
E7	Early region 7
FOXJ2	Forkhead Box J2
Gal4	Galactose 4

GO	Gene ontology
hBRM	Human Brahma
HDAC	Histone deacetylase complex
Hid	Head involution defective
Hif1A	Hypoxia inducible factor 1, alpha subunit
HP1	Heterochromatin Protein 1
IE	Instability element
InR	Insulin-like receptor
Ldh	Lactate dehydrogenase
MAST	Motif alignment & search tool
MCM	Minichromosome maintenance
MDM2	Murine double minute 2
ORC	Origin recognition complex
p55/dCAF-1	Chromatin assembly factor 1, p55 subunit
PCNA	Proliferating cell nuclear antigen
Pi3K68D	Phosphotidylinositol 3 kinase 68D
Polα	Polymerase alpha
qRT-PCR	Quantitative real-time PCR
Rb	Retinoblastoma protein
Rbf1	Retinoblastoma family protein 1
Rbf2	Retinoblastoma family protein 2
RNA-seq	RNA sequencing
RNR	Ribonucleotide reductase

ROS	Reactive oxygen species
SCF	Skp/Cullin/F-box complex
Skp	S-phase kinase-associated protein
SUV	Suppressor of variegation
SV40	Simian virus 40
SWI2/SNF2	Switch/Sucrose nonfermenting complex
tko	Technical knockout
UAS	Upstream activating sequence
Vang	Van Gogh
yw	yellow white

CHAPTER 1

Introduction

The retinoblastoma protein (Rb) is an evolutionarily conserved transcriptional co-repressor involved in the regulation of major cellular processes including proliferation, differentiation, metabolism and apoptosis. The best known function of Rb is regulation of cell cycle progression through repression of E2F target genes that are important for G1 to S phase transition (Giacinti and Giordano, 2006; Khidr and Chen, 2006). Rb was the first tumor suppressor to be identified, based on observations of retinoblastoma, a malignant tumor of the retina that mainly affects young children (Knudson, 1971; Benedict et al., 1983; Cavenee et al., 1983). Deregulation of Rb has been associated with various types of human cancer, involving either direct deletion or mutation of the *RB* gene, or through altered activity of upstream regulators (Giacinti and Giordano, 2006).

The Rb/E2F pathway is evolutionarily conserved in virtually all eukaryotes, from metazoans to plants and fungi (Cao et al., 2010). The human genome encodes three retinoblastoma pocket proteins, Rb, p107 and p130, which perform overlapping and unique functions in gene regulation. Genetic experiments and chromatin immunoprecipitation studies indicate that they are required for repression of different sets of genes (Dyson, 1998; Takahashi et al., 2000; Black et al., 2003). With respect to human disease, Rb is the main tumor suppressor in this family; mutations in p107 and p130 have been described in cancer genomes, but the significance of these lesions is unclear. At a molecular level, many questions remain regarding the specific biochemical activities of the retinoblastoma proteins, and how cellular tasks are apportioned between them. These proteins are not expressed in identical manners; thus, part of the different activities may lie with divergent regulation. At the same time, different members of this family have divergent promoter targeting

activities, a property that structural studies have characterized as due to specific distinctions in the binding of E2F factors by different pocket proteins (Rubin et al., 2005; Liban et al., 2017)). In addition, the protein complexes associated with these pocket proteins contain distincttypes of regulatory proteins, and it is possible that at the gene regulation level, different pocket proteins have distinct targets within the basal transcriptional machinery and/or the chromatin. Thus, the significance of the multiplicity of retinoblastoma proteins is still undetermined.

Despite the conservation of Rb, p107, and p130 in vertebrates, virtually all other metazoa encode a single retinoblastoma pocket family protein in their genomes. Duplication of this family of proteins is thus a rare event in evolutionary terms, unlike the case for numerous other transcription factors. Thus, it is of particular interest that all Drosophila species characterized to date encode two retinoblastoma proteins, Rbf1 and Rbf2, which makes Drosophila a powerful tool to study paralogy in this gene family. In this chapter, I will provide background on the retinoblastoma family of proteins in mammals and Drosophila. I will discuss the evolutionary conservation of retinoblastoma proteins with emphasis on conserved regulatory domains of the protein.

Mammalian retinoblastoma proteins

General characteristics

Studies of a childhood malignant tumor, retinoblastoma, led to the identification of the *RB* gene, the first tumor suppressor to be characterized (Knudson, 1971; Benedict et al., 1983; Cavenee et al., 1983). Studies of retinoblastoma indicated that this cancer is initiated by independent lesions affecting both copies of the *RB* gene, leading to the famous two-hit hypothesis (Knudson, 1971). Since Rb was identified, various studies have shown the association of Rb with a number of human cancers including lung, prostate, bladder and breast cancer (Burkhart and Sage, 2008). The best

studied function of Rb is its role in regulation of the cell cycle through interactions with E2F transcription factors, to which it binds by means of its pocket domain.

Rb is a target of oncoviral proteins, including the adenovirus E1A, human papillomavirus E7 and polyomavirus large T antigen proteins, which play a role in cellular transformation and cancer development. These proteins inhibit Rb interactions with E2F factors by binding to the Rb pocket domain (Felsani et al., 2006). These viral oncoproteins share a common peptide motif called LxCxE, which binds to a conserved region on the B-pocket domain. This specific LxCxE motif is also present in proteins contained in chromatin modifying complexes, including histone deacetylases (e.g. HDAC1 and HDAC2), histone methyl transferases (e.g. Suv39h1), histone binding proteins (e.g. HP1) and others (Henley and Dick, 2012). Therefore, Rb can repress E2F target genes in two major ways: binding to E2Fs and directly inhibiting their transcriptional activation activity, as well as recruiting chromatin remodeling factors that modify chromatin structure leading to repression.

p107 and p130 are pocket proteins related to Rb; they were identified based on their associations with oncoviral proteins (E1A, SV40 large T antigen) and sequence homology to Rb (Harlow et al., 1986; Dyson et al., 1989; Ewen et al., 1991; Li et al., 1993; Mayol et al., 1993). These proteins along with Rb are known as pocket proteins because they contain the conserved pocket domain responsible for most of the protein-protein interactions (Du and Pogoriler, 2006). p107 and p130 are structurally more similar to each other than to Rb (Figure 1-1).

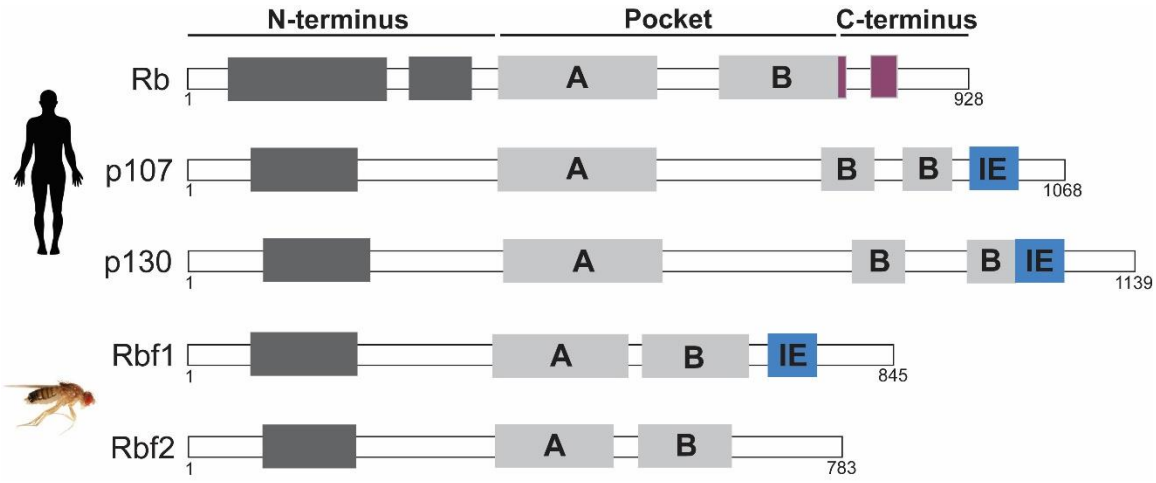


Figure 1-1: Schematic representation of the mammalian and the Drosophila retinoblastoma proteins. The dark grey box represents the cyclin fold domain in the N-terminus. Pocket A and B domains are shown in light grey. The C-terminal instability elements (IE) of p107, p130 and Rbf1 are shown in blue. The RbC^{nter} and RbC^{core} regions of Rb that function in E2F1/DP interactions are shown in magenta. (Figure adapted from Sengupta et al. 2015)

Interaction with E2Fs

There are at least eight E2F transcription factors in the mammalian system, including E2Fs mainly involved in transcriptional activation (E2F1-3) and others mainly involved in repression (E2F4-8) (Du and Pogoriler, 2006). E2F1–E2F5 factors bind pocket protein by means of a region within the transactivation (TA) domains, whereas E2F6-E2F8 contain divergent C-termini that lack pocket protein binding regions. These proteins are thought to be Rb-independent repressors. E2F1-E2F6 act as heterodimers; they bind to dimerization partner (DP) proteins through the leucine zipper (LZ) and marked box (MB) domains (Trimarchi and Lees, 2002). Rb binds to both activator and repressor E2Fs (E2F1-4), whereas p107 and p130 bind to repressor E2Fs: E2F4 and E2F4-5

respectively. The E2F TA binds to the cleft between the A and B pocket domains (Lee et al., 2002; Xiao et al., 2003).

In addition to the interactions between the pocket domains and the E2F TA domains, the C-terminal domains of Rb proteins mediate an additional interaction which is, in part, responsible for differential E2F binding. Studies from the Rubin lab showed that the p107 C-terminal domain has higher affinity to the coiled-coil and marked-box domains (CMs) of E2F4-DP, than to E2F1 (Liban et al., 2017). All the interactions made by the C-terminal domain of p107 are contained within $p107C^{994–1031}$ which is termed the $p107C^{core}$ (Liban et al., 2017). On the other hand, the Rb C-terminal domain exhibits a strong affinity to E2F1-DP CM. Two regions in the C-terminus of Rb interact with E2F1-DP CM: residues 786 - 801, termed the RbC^{nter} and residues 829 - 864, termed the RbC^{core} (Rubin et al., 2005).Interestingly, this unique Rb-E2F1 interaction mediated by Rb C terminus persists in the presence of adenoviral E1A protein, which usually disrupts Rb pocket binding to E2F1 TA (Seifried et al., 2008). Such stable Rb-E2F1 interactions enable Rb-E2F1 complexes to persist when Rb is phosphorylated in the cell cycle, when most other E2Fs are released (Cecchini and Dick, 2011). Therefore, this specific interaction mediated by the Rb C-terminus may be important for regulation E2F1-unique activities such as apoptosis (Dick and Rubin, 2013), a function known to be uniquely attributed to Rb.

Interaction with chromatin remodeling factors and enzymes

In vitro biochemical studies show that the pocket proteins interact with various chromatin remodeling factors and modifying enzymes. Rb, p107 and p130 interact with BRG1 and hBRM, which are the human homologs of the yeast SWI2/SNF2 chromatin remodeling factor. Interaction of Rb with BRG1 and hBRM was shown to be important for Rb repression activity (Dunaief et al.,

1994; Strober et al., 1996; Trouche et al., 1997). In addition, Rb interacts with histone deacetylase 1 (HDAC1), and this interaction potentiates its repression function on certain genes (Brehm et al., 1998; Luo et al., 1998; Magnaghi-Jaulin et al., 1998). Similarly, p107 and p130 interact with HDAC1, and inhibition of the deacetylase affects their repression activity (Ferreira et al., 1998). The retinoblastoma proteins have been also shown to interact with histone methyltransferases SUV39H1, Suv4-20h1 and Suv4-20h2 (Nielsen et al., 2001; Vandel et al., 2001; Nicholas et al., 2003; Gonzalo et al., 2005).

In addition to these chromatin modifying activities, Rb has been shown to interact with the DNA methyltransferase DNMT1, an interaction that enhances the Rb repression function (Robertson et al., 2000). Most of these studies were done in cultured cells or in purified systems, thus it is still unclear whether these properties are the essential functions acting in the intact organism. Interestingly, Isaac et al., 2006 showed that mouse embryonic fibroblasts (MEFs) with mutations in Rb LxCxE binding motif, so that it cannot interact with chromatin remodelers, have no changes in cell cycle progression but display mitotic defects (Isaac et al., 2006). Therefore, the physiological impact of Rb's interaction with chromatin remodeling and modifying enzymes on regulating cellular processes is not fully understood *in vivo*.

Regulation of retinoblastoma proteins

During the cell cycle, the three pocket proteins are regulated through phosphorylation (Henley and Dick, 2012). When Rb is hypophosphorylated, it binds to E2F and induces the repression of genes required for DNA synthesis and cell cycle progression. Upon mitogenic stimuli, CyclinD/Cdk4/6 and CyclinE/Cdk2 phosphorylate Rb during early and late G1 respectively. Phosphorylation leads to the release of E2F transcription factors and activation of E2F target genes leading to S-phase

entry (Giacinti and Giordano, 2006; Du and Pogoriler, 2006). p107 and p130 are also regulated by phosphorylation in a similar fashion to Rb, leading to release of the associated E2F factors (Classon and Dyson, 2001).

Recent studies showed that in early G1, Rb is monophosphorylated by CyclinD/Cdk4/6 (Narasimha et al., 2014). Thirteen conserved Cdk phosphorylation sites are present in Rb, and different phosphorylation events may induce different conformational changes that may alter Rb function in a different way. Some models have been proposed regarding specific phosphorylation events and how they may impact Rb function. For example, phosphorylation of a specific threonine residue (T373) at the Rb N-terminus induces docking between the RbN and the pocket, which inhibits the E2F-TA and LxCxE binding domains, whereas phosphorylation of two threonine residues in the Rb C-terminus (T821 and T826) induces conformational changes that blocks the E2F-DP MB and LxCxE binding domains (Dick and Rubin, 2013). Therefore, phosphorylation of Rb is not uniformly inactivating, but may alter the function of Rb in specific ways until Rb is hyperphorylated in late G1 and totally inactivated.

At the end of mitosis, the pocket proteins are reactivated by dephosphorylation (Henley and Dick, 2012). Dephosphorylation of Rb is mediated through protein phosphatase 1 (PP1) (Kolupaeva and Janssens, 2013). The PP1 binding site overlaps with the cyclin/cdk binding site on the C-terminus of Rb, indicating competition of phosphatases and kinases to bind and modify Rb (Henley and Dick, 2012).

Other post-translational modifications that impact Rb regulation include methylation and acetylation (Burkhart and Sage, 2008). The C-terminal lysine 873 of Rb is subject to acetylation and methylation. Acetylation of this residue, and lysine 874, prevents Rb inactivation by

phosphorylation. Methylation of K873 leads to Rb association with heterochromatin binding protein HP1, which increases repression activity. Methylation of another residue, K810, leads to inhibition of Cdk dependent phosphorylation (Macdonald and Dick, 2012). Interestingly, acetylation of lysine residue (K1079) in p130 increases its response to phosphorylation (Saeed et al., 2012) indicating differential impact of post-translational modifications on Rb protein regulation.

In addition to regulation by phosphorylation, levels of the pocket proteins are controlled by the ubiquitin-proteasome system. Specific E3 ligases are involved in turnover of Rb (MDM2), p107 and p130 (SCFSkp2) (Tedesco et al., 2002; Uchida et al., 2005; Sengupta and Henry, 2015).

Levels of retinoblastoma proteins during cell cycle

p130 is the most abundant pocket protein in G0 phase, whereas p107 is the least abundant. As cells enter the cell cycle, the levels of p130 decline, and the levels of Rb and p107 increase. When the levels of p107 increase, it replaces p130 at E2F responsive promoters (Takahashi et al., 2000). Rb is moderately expressed in G0, and in G1, Rb levels become more abundant where it binds to activator E2Fs and prevents activation of transcription of cell cycle genes (Grana et al., 1998; Henley and Dick, 2012). Therefore distinct retinoblastoma/E2F complexes are present in different stages of the cell cycle: p130/E2F4-5 is found mainly in G0, p107/E2F4 mainly in S phase cells but can also be found in $G_{1,}$ and Rb/E2F1-4 is present mainly as cells progress from G1 to S phase, but can also be found in G0 (Dyson, 1998).

Mutation phenotypes and tumor suppressor functions

Mutations in retinoblastoma proteins produce a variety of phenotypes, indicating distinct roles in development. Loss of Rb in mice is embryonic lethal; *RB1* null mice show defects in neurogenesis and hematopoiesis (Lee et al., 1992; Clarke et al., 1992; Jacks et al., 1992). This indicates that Rb is indispensable for normal development and its loss cannot be compensated by the other pocket proteins. Embryonic lethality in $RB1^{-/-}$ mice has been shown to be associated with abnormal placenta, and supplying a normal placenta, via tetraploid aggregation or by genetic approaches, rescues embryonic lethality, but mice die soon after birth (Wu et al., 2003). p107 and p130 can compensate for each other, as studies indicated that loss of either gene results in viable mice with no developmental defects (Lee et al., 1996; Cobrinik et al., 1996). However, loss of both p107 and p130 is lethal, and mice die shortly after birth due to breathing defects (Cobrinik et al., 1996).

In mice, partial loss of Rb increases predisposition to pituitary and thyroid tumors, and not retinoblastoma as in humans (Jacks et al., 1992; Hu et al., 1994). In contrast to Rb, loss of either p107 or p130 in mice does not lead to any tumor formation (Cobrinik et al., 1996; Lee et al., 1996). However, additional loss of p107 and p130 in an *RB1* mutant background leads to formation of retinoblastoma and other types of tumors, and loss of Rb and p130 in a *Trp53* null background increases the predisposition to small cell lung carcinoma. These results indicate tissue-specific tumor suppressor function of the retinoblastoma proteins (Robanus-Maandag et al., 1998; Jacks et al., 1992; Dannenberg et al., 2004; MacPherson et al., 2007). However, Rb is the major tumor suppressor of this family, and it is the only one frequently mutated in human cancers (Wirt and Sage, 2010; Burd and Sharpless, 2010).

Pleiotropic functions of Rb

Rb is best known for inhibiting E2F transcription factors and downregulating genes that contribute to cell-cycle progression (Dimova and Dyson, 2005; Henley and Dick, 2012). However, Rb also plays important roles in regulating other cellular processes that contribute to its tumor suppressor functions. One of the processes is apoptosis, and interestingly, Rb can have an anti-apoptotic or pro-apoptotic function depending on the cellular context. Through binding E2F1, Rb inhibits pro-apoptotic genes, and it was shown that loss of Rb function triggers the p53 apoptotic pathway (Harbour and Dean, 2000). However, upon DNA damage, Rb and E2F1 form a complex that results in activation of specific pro-apoptotic genes (Dick and Rubin, 2013).

Another function of Rb is regulating cellular differentiation by its interaction with lineage-specific transcription factors that are required for development of tissues such as bone, muscle, pancreas and erythroid cells (Viatour and Sage, 2011). In addition, Rb is essential for proper mitosis and genomic stability. Loss of Rb results in abnormal spindle formation, impaired cohesion, aneuploidy and genomic instability (Hayashi and Takahashi, 2015).

Rb has been shown to regulate cellular metabolism as well. Loss of Rb was shown to be associated with various metabolic defects including reduced mitochondrial respiration, reduced activity of the electron transport chain, and alteration in glutamine metabolism (Dyson, 2016). An additional function of Rb appears to be the control of mitochondrial biogenesis; loss of Rb in erythrocytes leads to mitochondrial defects and impaired erythropoiesis (Viatour and Sage, 2011).

Recent studies uncovered a newly discovered role for Rb in regulation of cell polarity (Parisi et al., 2018). Loss of Rb in mouse keratocytes is associated with polarity defects and upregulation of core polarity genes such as *Par3* and *aPKC*, and Rb inactivation causes tissue closure defects

during mouse development (Parisi et al., 2018). As discussed below, similar involvement of Rbf1 in Drosophila polarity has been described.

Differential roles of pocket proteins in gene regulation

Genetic and molecular studies demonstrate that retinoblastoma proteins have overlapping yet distinct functions in gene regulation. Using reporter assays and 3T3 fibroblast cells lacking different retinoblastoma family proteins, Classon et al. (2000) showed that E2F-responsive promoters are differentially regulated by specific pocket proteins; the *B-myb* reporter gene was found to be derepressed in absence of p107 and p130 and not affected by loss of Rb, however the *p107* reporter gene was derepressed by loss of Rb and not affected by loss of p107 and p130. In agreement with this result, using *RB1*$^{-/-}$, *p107*$^{-/-}$, *p130*$^{-/-}$, and *p107*$^{-/-}$*p130*$^{-/-}$ mouse embryonic fibroblasts, Hurford et al. (1997) showed that Rb, p107 and p130 are required for regulation of different sets of genes. For example, *Cyclin E* and *p107* were derepressed in *RB1*$^{-/-}$ cells and unaffected in *p107*$^{-/-}$*p130*$^{-/-}$ cells, however, *B-myb*, *cdc2*, *cyclin A2* and *E2F1* were derepressed in *p107*$^{-/-}$*p130*$^{-/-}$ cells and affected in *RB1*$^{-/-}$ cells. Interestingly, in both studies, *RB1*$^{-/-}$ and *p107*$^{-/-}$ *p130*$^{-/-}$ cells displayed higher percentage of S-phase in comparison to control cells, indicating that Rb and both p107 and p130 are essential for regulating cell cycle progression in this particular cell type.

In contrast, using a different approach, Stengel et al. (2009) showed that transfection of Saos-2 cells with Rb, p107 or p130 results in similar levels of repression on *TS* and *Cyclin A* reporter genes, and all the expressed pocket proteins were able to bind to endogenous *TS* and *RNR2* genes as shown by ChIP assays. This study suggests that these promoters respond to different pocket

proteins in an identical manner and emphasizes that pocket proteins do have overlapping functions in specific contexts.

Because regulatory regimes change as cells enter quiescence and senescence, the roles of Rb, p107, and p130 in regulation of endogenous genes were explored in growing, quiescent and senescent human lung fibroblasts using shRNAs targeting *RB*, *p107* and *p130* (Chicas et al., 2010). This study identified classes of genes that were either similarly or differentially regulated by the individual pocket proteins. During quiescence, there was minimal impact of suppressing individual retinoblastoma proteins on E2F target gene expression, indicating overlapping functions in this growth state. However, during senescence, Rb knockdown resulted in upregulation of many genes, indicating unique role of Rb in gene regulation during this state. Using ChIP technology, these authors found that Rb and p130 bind to different types of genes in different growth states (growing, quiescence and senescence). A general pattern was that binding of p130 to many promoters increased in quiescent cells knocked down for Rb, but not in senescent cells indicating that Rb's role in gene regulation during senescence is not redundant. The replacement of Rb by p130 in the quiescent state indicates that there may be antagonism between the Rb family proteins on specific target genes in certain cell types. Collectively, these studies indicate that pocket proteins have common and different characteristics related to gene targeting and repression activity, and the molecular mechanisms underlying these characteristics are still not fully understood. Importantly, the unique contributions of the individual retinoblastoma family members on gene regulation haven't been studied *in vivo*.

Drosophila retinoblastoma proteins

General characteristics

The conserved Rb/E2F pathway is simpler in Drosophila, involving fewer components than in mammals, which makes it a good model for Rb and E2F studies. The Drosophila Rb/E2F pathway consists of two pocket proteins, (Rbf1 and Rbf2), two E2fs (E2f1 and E2f2) and one Dp protein (Dimova and Dyson, 2005). Similar to the mammalian system, E2fs dimerize with Dp and are divided into activators and repressors; E2f1 is a potent activator of transcription, whereas E2f2 is associated with repression (Frolov et al., 2001). Co-immunoprecipitation assays showed that Rbf1 interacts with both E2f1 and E2f2, whereas Rbf2 interacts only with E2f2 (Stevaux et al., 2002). Rbf1 and Rbf2 proteins have extensive similarity to the mammalian retinoblastoma proteins, with a defined N-terminus, pocket and C-terminal regions (Figure 1-1), and they are regulated by phosphorylation during the cell cycle in a fashion similar to the mammalian pocket proteins (Stevaux et al., 2002; Xin et al., 2002).

Expression patterns of Rbf1 and Rbf2

Rbf1 and Rbf2 have different patterns of expression during development. During embryonic stages, Rbf1 protein levels are relatively constant with a slight decrease in late embryogenesis. In contrast, Rbf2 is absent from the early oocyte. The levels of this protein increase and peak in 4-10 hr. embryos, and gradually drop until reaching undetectable levels in the late stage embryo (Keller et al., 2005). The patterns of expression of *Rbf1* and *Rbf2* are overlapping in early stages of the embryo, but at stage 13, *Rbf2* expression is confined to the brain and the ventral nerve cord, whereas *Rbf1* is also expressed in the gut. Both *Rbf1* and *Rbf2* are expressed in imaginal discs in third instar larvae including wing, leg and eye (Keller et al., 2005). In adults, Rbf1is expressed in

both males and females, whereas Rbf2 expression is undetectable in males, and confined to ovaries of females, suggesting a potential role in oogenesis (Stevaux et al., 2002).

Rbf1 and Rbf2 mutant phenotypes

Zygotic Rbf1 null mutants (*Rbf1*^{14}) die before the pupal stage and display developmental delays in the third instar larval stage (Du and Dyson, 1999). Stevaux et al. (2005) showed that Rbf2 mutant flies are viable and "do not display any observable phenotypes". However, these authors also reported that homozygous mutant females had an elevated egg laying rate (Stevaux et al., 2005). However subsequent analysis in our laboratory suggests that this effect may have been due to genetic background effects rather than the lesion in *Rbf2*. The *Rbf2* mutant allele that was generated still produces a 30 kD Rbf2 amino-terminal fragment of Rbf2, which may be partially functional. The possible activity of the N-terminus of Rbf2 is still not fully understood, however a conserved cyclin fold domain resides in the N-terminus of mammalian and Drosophila N-terminal regions (Figure 1-1).

Flies with a partial loss of function *Rbf1* allele (*Rbf1*^{120a}) display normal eye morphology (Du and Dyson, 1999), but additional loss of Rbf2 leads to rough eye phenotype, which is absent in *Rbf2* null flies only. Flies that carry null alleles of both *Rbf1* and *Rbf2* display poorer viability and longer developmental delay than *Rbf1* null flies, indicating that loss of Rbf2 enhances the phenotypes observed due to Rbf1 loss (Stevaux et al., 2005).

Diverse functions of Rbf1

Similar to mammalian Rb, Rbf1 regulates various cellular processes in Drosophila. When Rbf1 was first identified based on homology to the mammalian pocket proteins, it was shown to regulate

cell cycle genes (*PCNA* and *RNR2*), which are E2f regulated genes, and therefore regulate G1-S phase transition (Du and Dyson, 1999). In addition to cell cycle regulation, Rbf1 has been shown to regulate apoptosis. Overexpression of Rbf1 in wing imaginal discs using the UAS-Gal4 system results in apoptosis, and notches along the wing margin, and this phenotype was reversed by co-expression of E2F1 (Milet et al., 2010; Elenbaas et al., 2015). Rbf1 overexpression causes apoptosis by downregulating two anti-apoptotic genes: *Buffy* and *Diap1*. *Buffy* encodes an anti-apoptotic protein member of the Bcl-2 family and is directly repressed by Rbf1 at a transcriptional level, whereas *Diap1,* which encodes a caspase inhibitor, is repressed by Rbf1 post-transcriptionally through regulation of *held out wings,* a gene that encodes an RNA binding protein involved in *Diap1* mRNA degradation (Clavier et al., 2014). In addition, Rbf1 overexpression induces mitochondrial fragmentation, ROS production and apoptosis through Debcl, a Bcl-2 family member, and Drp1, which is a large GTPase of the dynamin family required for mitochondrial fission (Clavier et al., 2015). Another study indicated that loss of Rbf1 results in apoptosis in eye imaginal discs through upregulation of the pro-apoptotic gene *hid* (Moon et al., 2006).

Through analysis of loss of function mutants, Rbf1 has been shown to play a role in metabolism; metabolomic analysis of third instar *Rbf1* mutant larvae revealed altered nucleotide synthesis and glutathione metabolism. Under fasting conditions, Rbf1 mutants showed an increase in flux of glutamine toward glutathione synthesis (Nicolay et al., 2013). The impact of a partial loss of function in Rbf1 was investigated in the eye imaginal discs of third instar larvae; scRNA-seq data revealed that this mutation leads to increased intracellular acidification though upregulation of *Ald*, *HIF1a* and *Ldh*, and eventually leading to apoptosis (Ariss et al., 2018).

Rbf1 is also important in maintaining genomic stability. Rbf1 promotes chromatin condensation through association with dCAP-D3, a component of the Condensin II complex, and promoting efficient association of dCAP-D3 with chromatin (Longworth et al., 2008). Another recently discovered role of Rbf1 in Drosophila is the regulation of cell polarity. Knockdown of Rbf1 results in polarity defects in eyes, wings, and notum, and results in upregulation of core polarity genes such as *aPKC, Vang* and *par-6* (Payankaulam et al., 2016). Collectively, studies from Drosophila Rbf1 show that regulation of cell cycle, apoptosis, metabolism, genomic stability and polarity are all conserved features of the retinoblastoma family.

Unlike the functional roles ascribed to Rbf1, the roles of Rbf2 have not been as thoroughly studied. Initial characterization of this protein showed that increased expression of Rbf2 and E2f2 in wing and eye imaginal discs resulted in blocked entry into S-phase, suggesting a role in cell cycle regulation, however Rbf2 exhibits only weak regulation of some well-characterized cell cycle promoters (Stevaux et al., 2002). The unique expression of Rbf2 in female ovaries suggests a role in oogenesis, which hasn't been studied yet.

Gene targets of Rbf1 and Rbf2

Korenjak et al. (2012) performed a ChIP-chip analysis to study Rbf1, E2f1 and E2f2 genome occupancy in Drosophila third instar larvae. Binding sites were enriched around transcription start sites for the three factors. There was a strong enrichment for cell cycle genes among the genes targeted by Rbf1. Rbf1 and E2f2 were found to bind more sites than E2f1, and there was a significant overlap between Rbf1 and E2f2 binding sites. The majority of E2f1 binding sites were also bound by Rbf1. Some of the Rbf1 target genes (14%) lacked a consensus E2F motif, although these genes are bound by E2f2. Using *Dp* mutant larvae, binding of Rbf1 and E2f2 was diminished

on selected genes that lack a consensus E2F motif, indicating that Rb/E2f complexes can be present at sites that don't contain a classical E2F site (Korenjak et al., 2012).

Acharya et. al. (2012) studied Rbf1 occupancy in the Drosophila embryo using ChIP-seq technology. In addition to the canonical cell cycle genes, many targets of Rbf1 in the embryo belonged to signaling pathways including insulin, Notch, Hippo and JAK/STAT. The functional significance of the targeting of these diverse signaling pathways is yet to be elucidated, however Rbf1 appears to regulate the insulin receptor gene through the promoter proximal E2F site. In agreement with the previous study by Dyson lab, most of the peaks were promoter proximal. In addition to the E2F motif, transcription factor motifs including DREF and FOXJ2 were enriched in Rbf1 bound regions (Acharya et al., 2012).

An independent ChIP-seq of Rbf1 and Rbf2 was carried out by Wei et al. (2015) in the Drosophila embryo. Strikingly, Rbf2 was found to bind to twice as many genes as Rbf1, and most of the Rbf1 bound genes were also bound by Rbf2. Genes that were uniquely bound by Rbf2 were not enriched in E2F motifs, unlike genes that were bound by both Rbf1 and Rbf2. Genes that were bound by both Rbf1 and Rbf2 were enriched with cell cycle genes. Interestingly, genes that were uniquely bound by Rbf2 were enriched with ribosomal protein genes, and Rbf2 is able to mediate repression of a *tko* reporter gene, which encodes a cytosolic ribosomal protein. Both Rbf1 and Rbf2 were bound on many genes belonging to signaling pathways. The physiological significance Rbf1/Rbf2 binding to ribosomal protein and signaling pathway genes is still not fully understood, however, coregulating the expression of these genes with cell cycle control may be of importance for proper coordination of cell growth and cell division.

Mechanisms of gene regulation by Rbf1 and Rbf2

To identify regulatory targets of Rb and E2f factors, Dimova et al. (2003) depleted Rbf1, Rbf2, E2f1, E2f2 and Dp in cultured S2 cells using double-stranded RNA (dsRNA). They functionally categorized the genes by response: Groups A and B include genes involved in cell cycle regulation and DNA replication, groups C, D and E include genes that have development and differentiation functions. Interestingly, depletion of Rbf2 alone had little impact on gene expression in these cells, however, when depleted in conjunction with Rbf1, many genes of class C, D, and E were upregulated, indicating that Rbf2 has a redundant function. In contrast, depletion of Rbf1 alone was sufficient to induce upregulation of genes of group A and most of group B, including *PCNA*, *Cyclin E*, *Orc1* and other cell cycle genes. Those Rbf1-dependent genes also were downregulated by E2f1, but no change was seen with E2f2 depletion. Interestingly, groups C, D, and E, which are upregulated by E2f2 depletion, are also affected by joint depletion of Rbf1 and Rbf2, indicating that requirement for Rbf1 or Rbf2 depends on regulation by either E2f1 or E2f2. These genes include oogenesis related genes such as *vasa*, *spn-E* and *bng*, and male specific genes such as *Arp53D*. This study shows that E2fs and Rbfs have different functions in gene regulation, where Rbf1/E2f1 are important for regulating cell cycle genes, whereas both Rbf1 and Rbf2 in complex with E2f2 are important for differentiation related genes (Dimova et al., 2003).

An earlier study showed that on E2f regulated reporter genes (*PCNA*, *MCM3* and *Polα*), Rbf2 can repress E2f-dependent transcription, but is considerably weaker than Rbf1. Unlike Rbf1, Rbf2 failed to block E2f1 activation of *Polα* reporter, consistent with its ability to form complexes with E2F2 only, unlike Rbf1, which can interact with both E2f1 and E2f2 (Stevaux et al., 2002). However, combined expression of E2f2 and Rbf2 was able to block E2f1 activation of a *PCNA*

reporter and inhibit DNA synthesis and entry into S-phase, indicating cooperation between the two factors to antagonize E2f1 activity.

Loss of Rbf2 in ovaries had no impact on cell cycle genes but affected differentiation specific genes, including testis differentiation markers. Strikingly, a significant number of genes were upregulated in Rbf2 mutant ovaries, but unchanged in E2f2 mutant ovaries, suggesting that Rbf2 may play a role in gene regulation independent of E2f2. Interestingly, on these genes, Rbf2 and E2f2 were present on the promoters regardless whether the genes are impacted by loss of E2f2 or Rbf2. There was very minimal overlap between genes affected by loss of Rbf2 in ovaries, S2 cells, and embryos indicating specific roles of Rbf2 in different cell type and different stages of development (Stevaux et al., 2005).

The unique binding of Rbf2 to ribosome protein genes raises the possibility that Rbf2 may be an important regulator of this gene family. Overexpression of Rbf2 in Drosophila S2 cells resulted in modest repression of *tko* reporter and a stronger repression with co-expression of E2f2. Rbf1 failed to repress this ribosomal protein gene, indicating a potential unique role for Rbf2 in regulating this class of genes. The physiological significance of this regulation is yet not fully understood (Wei et al. 2015).

Rbf1 and chromatin remodeling factors

The interaction of retinoblastoma proteins with chromatin remodeling and modifying enzymes, and their impact on gene regulation, is poorly studied in Drosophila. In one study, p55/dCAF-1, a chromatin assembly factor, was shown to be important for repression of a set of E2f2 regulated genes in cultured S2 cells (Taylor-Harding et al., 2004). Rbf1 was shown to bind p55 which is a histone-binding protein and a component of several chromatin complexes. Depletion of p55 in

Drosophila S2 cells resulted in upregulation of differentiation specific genes such as *Arp53D* (group E genes from Dimova et al. 2003) that are known to be repressed by Rbf1 in S2 cells, with no impact on group A genes (e.g. *RNR2*) that are normally coupled with cell proliferation. These results indicate that mechanisms of Rbf1 regulation of different classes of genes are different. Interestingly, unlike mammalian studies which showed requirement of HDAC and SWI/SNF chromatin-modifying complexes in Rb-mediated repression, this study showed that Rbf1 repression of group A and E genes generally does not require HDAC or SWI/SNF proteins (Taylor-Harding et al., 2004).

However, in disagreement with the previous study, another study showed that HDAC activity is essential for repression of the developmentally regulated genes (group D/E) including *Arp53D* and others. In addition, this study showed that the Polycomb group (PcG) protein, Enhancer of zeste (E(Z)), which is an H3k27 methyltransferase in Drosophila, is important for silencing group D and E genes through dimethylation of histone H3 Lys27 (H3K27me2) (Lee et al., 2010). It is not known whether Rbf2 represses target genes by associating with chromatin modifying or remodeling factors. Using mass spectrometric analysis, Ullah et al. (2007) showed that Rbf2 associates with TRRAP, a component of several histone acetyltransferase complexes, and chromatin remodeling factors such as Moira, BAF53 and Caf1/P55. The significance of these protein interactions is yet to be elucidated.

The dREAM complex

Rbf1 and Rbf2 are part of a conserved gene regulatory complex termed dREAM (Drosophila Rbf, E2f, and Myb-interacting proteins). Biochemical studies have shown that two dREAM complexes exist in flies each containing the Myb-interacting proteins (Mip120, Mip130 and Mip40), dMyb

transcription factor, CAF1p55, E2f2/Dp in addition to either Rbf1 or Rbf2 (Korenjak et al., 2004). In proliferating cells, the dREAM complex is important for repression of differentiation-specific genes, which are only targeted by E2f2, but not cell cycle genes, which are bound by both E2f1 and E2f2 (Korenjak et al., 2004; Georlette et al., 2007; Lee et al., 2012). dREAM is also required for E2f2 binding at differentiation-specific genes, which remain repressed during the S-phase, but does not affect its binding on cell cycle genes. E2f1 does not bind to differentiation-specific genes even in the absence of E2f2 and the entire dREAM complex, indicating that there is no competition between the E2fs for binding on these promoters (Lee at al., 2012; Dimova et al., 2003).

As part of dREAM, Rbf1 is hypophosphorylated, and dREAM is required for maintaining the hypophosphorylated form of Rbf1 during the S-phase where the complex binds to differentiation-specific genes (Lee et al., 2012). Therefore, Rbf1 function in the cell cycle is partly dependent on whether it is part of the dREAM complex. In the mammalian system, E2F4 in addition to p107 or p130 are part of the dREAM complex, which plays a role in repressing cell cycle genes during quiescence and has been shown to be deregulated in various human cancers (Sadasivam and DeCaprio, 2003).

The instability element (IE)

In addition to regulation by Cyclin/Cdk phosphorylation, turnover by the ubiquitin-proteasome pathway is another level of control of retinoblastoma family members. Previous studies from Arnosti and Henry labs revealed many aspects of turnover regulation of Rbf1 and the mammalian retinoblastoma family members (Acharya et al., 2010; Raj et al., 2012; Zhang et al., 2014; Sengupta et al., 2015). A conserved C-terminal region (58 residues), termed the instability element (IE), acts as an autonomous degron and is important for degradation of Rbf1, p107 and p130 by

the ubiquitin-proteasome pathway (Acharya et al., 2010; Sengupta et al., 2015). The Rb C-terminal regions (RbC^{nter} and RbC^{core}) that were described by the Rubin lab to mediate additional specific interactions with E2F/DP MB domains, also contribute to Rb turnover by the proteasome (Figure 1-1) (Sengupta et al., 2015; Sengupta and Henry, 2015).

Interestingly, the IE of Rbf1, p107 and p130 also affects repression function as tested by reporter assays for E2F target genes (Acharya et al., 2010; Sengupta et al., 2015). These studies demonstrate a tight link between proteasome-mediated protein turnover and repression function. In reporter assays, Rbf1-ΔIE (deletion of the IE) mutant, which is a more stable than full length WT-Rbf1 protein, showed impaired repression on canonical E2F target genes (*PCNA*, *Polα* and *MCM7*) but not on non-canonical E2F genes (*InR*, *wts* and *Pi3K68D*) suggesting that the IE affects gene-specific repression (Acharya et al., 2010; Raj et al., 2012). Forced ubiquitination of Rbf1-ΔIE enhanced repression on the *PCNA* reporter but not *InR* reporter gene indicating an interesting link between ubiquitination and gene-specific repression (Raj et al., 2012).

Specific lysine and serine residues within the IE affect Rbf1 stability and activity in different manner (Acharya et al., 2010; Zhang et al., 2014). Rbf1 bearing mutations in three, four or the six lysine residues within the IE exhibited higher stability in comparison to wild type protein. Interestingly, lysine to alanine and not lysine to arginine (charge conserving) mutations increased Rbf1 stability, indicating that ubiquitination of these lysines is not the reason for the instability. Interestingly, single mutations in the IE lysine residues did not affect stability of Rbf1 but exhibited hypomorphic and hypermorphic functions, indicating that the IE has positive and negative transcription regulatory elements (Acharya et al., 2010). One interesting hypermorphic mutation, K774A/R, enhanced repression of Rbf1 on a *PCNA* reporter and induced severe phenotypes when overexpressed in fly eyes and wings (Acharya et al., 2010; Zhang et al., 2014). Interestingly, K774

is conserved in mouse (K1079) and human (K1083) p130, and acetylation of mouse K1079 is important for in vitro Cdk4-mediated phosphorylation (Saeed et al., 2012). Whether specific mutations affecting the IE are relevant for cancer is unknown, however, one study showed that this mutation was reported in human lung cancer (Claudio et al., 2000).

Cyclin/Cdk overexpression stabilizes Rbf1 and inactivates it, indicating a tight link between regulation of stability and activity. Three conserved serine residues (S728, S760 and S771) within the IE are important for Cyclin/Cdk control of Rbf1 (Zhang et al., 2014). When these serines were mutated to alanines (3SA), Rbf1-3SA was not stabilized by Cyclin/Cdk overexpression and repression on *PCNA* was not elevated. Overexpression of Rbf1-3SA in wings and eyes resulted in severe phenotypes which were similar to those resulting from the K774A mutant. Interestingly, Rbf1-K774A mutant is also not stabilized nor inactivated by Cyclin-Cdk overexpression. This indicates that this residue may play a role in Cyclin/Cdk control of Rbf1 stability and activity (Zhang et al., 2014).

The evolutionary conservation of the IE and this functional characterization of mutants bearing lesions in this domain of retinoblastoma proteins indicate that the IE is a critical regulatory region. This element is modified or lost in the more derived forms of the pocket proteins, Rb in mammals and Rbf2 in Drosophila, which may be a key to understanding the different functions of the pocket proteins. Importantly, various point mutations and deletions that map to p107 and p130 IE region have been reported in cancer patients (Forbes et al., 2011; Gao et al., 2013).

Thesis preview

A critical gap in our knowledge about the retinoblastoma proteins is their different biochemical activities and functional roles in development, in an *in vivo* setting. Therefore, to answer these

questions, I used Drosophila as a model to study the impact of Rbf1 and Rbf2 proteins on gene regulation in Drosophila embryos. I showed that Rbf1 and Rbf2 regulate different sets of genes, and Rbf2 mainly affects ribosomal and mitochondrial genes hinting to potential role in regulating cell growth. We had questions about the role of Rbf2 in fly development, since it was neglected and hasn't been studied as much as Rbf1. We believed that conservation of *Rbf2* in all Drosophila genomes points to necessary roles in flies. Therefore, to understand the role of Rbf2, we created CRISPR knockout flies and showed that Rbf2 is important for regulating ovary function in females and is necessary for longevity in both females and males.

One interesting feature that is shared by the ancestral retinoblastoma proteins (p107, p130 and Rbf1) is the presence of the IE in the C-terminus, which impacts both stability and repression activity of these proteins. This element is lost in the derived forms, Rb and Rbf2, and the impact of losing this region on the functional diversification of these proteins is still not understood. To answer this question, I studied the function of the IE and how it impacts repression activity of Rbf1 during development. I expressed various Rbf1 isoforms bearing specific mutations in the IE, in wing tissues and in embryos, and I studied how these mutants impact the transcriptome using RNA-seq analysis. I found that the IE is an important regulatory domain, and specific residues within the IE impact Rbf1 function in a distinct manner. Therefore, divergence in this C-terminal domain could be driving the functional diversity of the retinoblastoma proteins.

REFERENCES

REFERENCES

Acharya, P., Negre, N., Johnston, J., Wei, Y., White, K. P., Henry, R. W. and Arnosti, D. N. (2012). Evidence for autoregulation and cell signaling pathway regulation from genome-wide binding of the Drosophila retinoblastoma protein. *G3 (Bethesda).* **2**, 1459–72.

Acharya, P., Raj, N., Buckley, M. S., Zhang, L., Duperon, S., Williams, G., Henry, R. W. and Arnosti, D. N. (2010). Paradoxical Instability–Activity Relationship Defines a Novel Regulatory Pathway for Retinoblastoma Proteins. *Mol. Biol. Cell* **21**, 3890–3901.

Ariss, M. M., Islam, A. B. M. M. K., Critcher, M., Zappia, M. P. and Frolov, M. V (2018). Single cell RNA-sequencing identifies a metabolic aspect of apoptosis in Rbf mutant. *Nat. Commun.* **9**, 5024.

Benedict, W. F., Murphree, A. L., Banerjee, A., Spina, C. A., Sparkes, M. C. and Sparkes, R. S. (1983). Patient with 13 chromosome deletion: evidence that the retinoblastoma gene is a recessive cancer gene. *Science* **219**, 973–5.

Black, E. P., Huang, E., Dressman, H., Rempel, R., Laakso, N., Asa, S. L., Ishida, S., West, M. and Nevins, J. R. (2003). Distinct gene expression phenotypes of cells lacking Rb and Rb family members. *Cancer Res.* **63**, 3716–23.

Brehm, A., Miska, E. A., McCance, D. J., Reid, J. L., Bannister, A. J. and Kouzarides, T. (1998). Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**, 597–601.

Burd, C. E. and Sharpless, N. E. (2010). What's so special about RB? *Cancer Cell* **17**, 313–4.

Burkhart, D. L. and Sage, J. (2008). Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat. Rev. Cancer* **8**, 671–82.

Cao, L., Peng, B., Yao, L., Zhang, X., Sun, K., Yang, X. and Yu, L. (2010). The ancient function of RB-E2F Pathway: insights from its evolutionary history. *Biol. Direct* **5**, 55.

Cavenee, W. K., Dryja, T. P., Phillips, R. A., Benedict, W. F., Godbout, R., Gallie, B. L., Murphree, A. L., Strong, L. C. and White, R. L. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature* **305**, 779–84.

Cecchini, M. J. and Dick, F. A. (2011). The biochemical basis of CDK phosphorylation-independent regulation of E2F1 by the retinoblastoma protein. *Biochem. J.* **434**, 297–308.

Chicas, A., Wang, X., Zhang, C., McCurrach, M., Zhao, Z., Mert, O., Dickins, R. A., Narita, M., Zhang, M. and Lowe, S. W. (2010). Dissecting the unique role of the retinoblastoma tumor suppressor during cellular senescence. *Cancer Cell* **17**, 376–87.

Clarke, A. R., Maandag, E. R., van Roon, M., van der Lugt, N. M., van der Valk, M., Hooper, M. L., Berns, A. and te Riele, H. (1992). Requirement for a functional Rb-1 gene in murine development. *Nature* **359**, 328–30.

Classon, M. and Dyson, N. (2001). p107 and p130: versatile proteins with interesting pockets. *Exp. Cell Res.* **264**, 135–47.

Classon, M., Salama, S., Gorka, C., Mulloy, R., Braun, P. and Harlow, E. (2000). Combinatorial roles for pRB, p107, and p130 in E2F-mediated cell cycle control. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10820–5.

Claudio, P. P., Howard, C. M., Pacilio, C., Cinti, C., Romano, G., Minimo, C., Maraldi, N. M., Minna, J. D., Gelbert, L., Leoncini, L., et al. (2000). Mutations in the retinoblastoma-related gene RB2/p130 in lung tumors and suppression of tumor growth *in vivo* by retrovirus-mediated gene transfer. *Cancer Res.* **60**, 372–82.

Clavier, A., Ruby, V., Rincheval-Arnold, A., Mignotte, B. and Guenal, I. (2015). The Drosophila retinoblastoma protein, Rbf1, induces a Debcl- and Drp1-dependent mitochondrial apoptosis. *J. Cell Sci.* **128**, 3239–3249.

Clavier, A., Baillet, A., Rincheval-Arnold, A., Coléno-Costes, A., Lasbleiz, C., Mignotte, B. and Guénal, I. (2014). The pro-apoptotic activity of Drosophila Rbf1 involves dE2F2-dependent downregulation of diap1 and buffy mRNA. *Cell Death Dis.* **5**, e1405.

Cobrinik, D., Lee, M. H., Hannon, G., Mulligan, G., Bronson, R. T., Dyson, N., Harlow, E., Beach, D., Weinberg, R. A. and Jacks, T. (1996). Shared role of the pRB-related p130 and p107 proteins in limb development. *Genes Dev.* **10**, 1633–44.

Dannenberg, J.-H., Schuijff, L., Dekker, M., van der Valk, M. and te Riele, H. (2004). Tissue-specific tumor suppressor activity of retinoblastoma gene homologs p107 and p130. *Genes Dev.* **18**, 2952–62.

Dick, F. A. and Rubin, S. M. (2013). Molecular mechanisms underlying RB protein function. *Nat. Rev. Mol. Cell Biol.* **14**, 297–306.

Dimova, D. K., Stevaux, O., Frolov, M. V and Dyson, N. J. (2003). Cell cycle-dependent and cell cycle-independent control of transcription by the Drosophila E2F/RB pathway. *Genes Dev.* **17**, 2308–20.

Dimova, D. K. and Dyson, N. J. (2005). The E2F transcriptional network: old acquaintances with new faces. *Oncogene* **24**, 2810–26.

Du, W. and Dyson, N. (1999). The role of RBF in the introduction of G1 regulation during Drosophila embryogenesis. *EMBO J.* **18**, 916–25.

Du, W. and Pogoriler, J. (2006). Retinoblastoma family genes. *Oncogene* **25**, 5190–200.

Dunaief, J. L., Strober, B. E., Guha, S., Khavari, P. A., Alin, K., Luban, J., Begemann, M., Crabtree, G. R. and Goff, S. P. (1994). The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. *Cell* **79**, 119–30.

Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes Dev.* **12**, 2245–62.

Dyson, N., Howley, P. M., Münger, K. and Harlow, E. (1989). The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* **243**, 934–7.

Dyson, N. J. (2016). *RB1*: a prototype tumor suppressor and an enigma. *Genes Dev.* **30**, 1492–1502.

Elenbaas, J. S., Mouawad, R., Henry, R. W., Arnosti, D. N. and Payankaulam, S. (2015). Role of Drosophila retinoblastoma protein instability element in cell growth and proliferation. *Cell Cycle* **14**, 589–97.

Ewen, M. E., Xing, Y. G., Lawrence, J. B. and Livingston, D. M. (1991). Molecular cloning, chromosomal mapping, and expression of the cDNA for p107, a retinoblastoma gene product-related protein. *Cell* **66**, 1155–64.

Felsani, A., Mileo, A. M. and Paggi, M. G. (2006). Retinoblastoma family proteins as key targets of the small DNA virus oncoproteins. *Oncogene* **25**, 5277–85.

Ferreira, R., Magnaghi-Jaulin, L., Robin, P., Harel-Bellan, A. and Trouche, D. (1998). The three members of the pocket proteins family share the ability to repress E2F activity through recruitment of a histone deacetylase. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10493–8.

Forbes, S. A., Bindal, N., Bamford, S., Cole, C., Kok, C. Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945-50.

Frolov, M. V, Huen, D. S., Stevaux, O., Dimova, D., Balczarek-Strang, K., Elsdon, M. and Dyson, N. J. (2001). Functional antagonism between E2F family members. *Genes Dev.* **15**, 2146–60.

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* **6**, pl1–pl1.

Georlette, D., Ahn, S., MacAlpine, D. M., Cheung, E., Lewis, P. W., Beall, E. L., Bell, S. P., Speed, T., Manak, J. R. and Botchan, M. R. (2007). Genomic profiling and expression studies reveal both positive and negative activities for the Drosophila Myb MuvB/dREAM complex in proliferating cells. *Genes Dev.* **21**, 2880–96.

Giacinti, C. and Giordano, A. (2006). RB and cell cycle progression. *Oncogene* **25**, 5220–7.

Gonzalo, S., García-Cao, M., Fraga, M. F., Schotta, G., Peters, A. H. F. M., Cotter, S. E., Eguía, R., Dean, D. C., Esteller, M., Jenuwein, T., et al. (2005). Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nat. Cell Biol.* **7**, 420–8.

Graña, X., Garriga, J. and Mayol, X. (1998). Role of the retinoblastoma protein family, pRB, p107 and p130 in the negative control of cell growth. *Oncogene* **17**, 3365–83.

Harbour, J. W. and Dean, D. C. (2000). Rb function in cell-cycle regulation and apoptosis. *Nat. Cell Biol.* **2**, E65-7.

Hayashi, N. and Takahashi, C. (2015). Pleiotropic functions of RB protein in tumor suppression. *Arch can res.* 3, 3:22.

Henley, S. A. and Dick, F. A. (2012). The retinoblastoma family of proteins and their regulatory functions in the mammalian cell division cycle. *Cell Div.* **7**, 10.

Hu, N., Gutsmann, A., Herbert, D. C., Bradley, A., Lee, W. H. and Lee, E. Y. (1994). Heterozygous Rb-1 delta 20/+mice are predisposed to tumors of the pituitary gland with a nearly complete penetrance. *Oncogene* **9**, 1021–7.

Hurford, R. K., Cobrinik, D., Lee, M. H. and Dyson, N. (1997). pRB and p107/p130 are required for the regulated expression of different sets of E2F responsive genes. *Genes Dev.* **11**, 1447–63.

Isaac, C. E., Francis, S. M., Martens, A. L., Julian, L. M., Seifried, L. A., Erdmann, N., Binné, U. K., Harrington, L., Sicinski, P., Bérubé, N. G., et al. (2006). The retinoblastoma protein regulates pericentric heterochromatin. *Mol. Cell. Biol.* **26**, 3659–71.

Jacks, T., Fazeli, A., Schmitt, E. M., Bronson, R. T., Goodell, M. A. and Weinberg, R. A. (1992). Effects of an Rb mutation in the mouse. *Nature* **359**, 295–300.

Keller, S. A., Ullah, Z., Buckley, M. S., Henry, R. W. and Arnosti, D. N. (2005). Distinct developmental expression of Drosophila retinoblastoma factors. *Gene Expr. Patterns* **5**, 411–21.

Khidr, L. and Chen, P.-L. (2006). RB, the conductor that orchestrates life, death and differentiation. *Oncogene* **25**, 5210–9.

Kolupaeva, V. and Janssens, V. (2013). PP1 and PP2A phosphatases--cooperating partners in modulating retinoblastoma protein activation. *FEBS J.* **280**, 627–43.

Korenjak, M., Anderssen, E., Ramaswamy, S., Whetstine, J. R. and Dyson, N. J. (2012). RBF Binding to both Canonical E2F Targets and Noncanonical Targets Depends on Functional dE2F/dDP Complexes. *Mol. Cell. Biol.* **32**, 4375–4387.

Korenjak, M., Taylor-Harding, B., Binné, U. K., Satterlee, J. S., Stevaux, O., Aasland, R., White-Cooper, H., Dyson, N. and Brehm, A. (2004). Native E2F/RBF complexes contain Myb-interacting proteins and repress transcription of developmentally controlled E2F target genes. *Cell* **119**, 181–93.

Lee, C., Chang, J. H., Lee, H. S. and Cho, Y. (2002). Structural basis for the recognition of the E2F transactivation domain by the retinoblastoma tumor suppressor. *Genes Dev.* **16**, 3199–212.

Lee, E.Y., Chang, C.Y., Hu, N., Wang, Y.C., Lai, C.C., Herrup, K., Lee, W.H. and Bradley, A. (1992). Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. *Nature*. **359,** 288-94.

Lee, H., Ragusano, L., Martinez, A., Gill, J. and Dimova, D. K. (2012). A Dual Role for the dREAM/MMB Complex in the Regulation of Differentiation-Specific E2F/RB Target Genes. *Mol. Cell. Biol.* **32**, 2110–2120.

Lee, H., Ohno, K., Voskoboynik, Y., Ragusano, L., Martinez, A. and Dimova, D. K. (2010). Drosophila RB proteins repress differentiation-specific genes via two different mechanisms. *Mol. Cell. Biol.* **30**, 2563–77.

Lee, M. H., Williams, B. O., Mulligan, G., Mukai, S., Bronson, R. T., Dyson, N., Harlow, E. and Jacks, T. (1996). Targeted disruption of p107: functional overlap between p107 and Rb. *Genes Dev.* **10**, 1621–32.

Li, Y., Graham, C., Lacy, S., Duncan, A. M. and Whyte, P. (1993). The adenovirus E1A-associated 130-kD protein is encoded by a member of the retinoblastoma gene family and physically interacts with cyclins A and E. *Genes Dev.* **7**, 2366–77.

Liban, T. J., Medina, E. M., Tripathi, S., Sengupta, S., Henry, R. W., Buchler, N. E. and Rubin, S. M. (2017). Conservation and divergence of C-terminal domain structure in the retinoblastoma protein family. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4942–4947.

Longworth, M. S., Herr, A., Ji, J.-Y. and Dyson, N. J. (2008). RBF1 promotes chromatin condensation through a conserved interaction with the Condensin II protein dCAP-D3. *Genes Dev.* **22**, 1011–24.

Macdonald, J. I. and Dick, F. A. (2012). Posttranslational modifications of the retinoblastoma tumor suppressor protein as determinants of function. *Genes Cancer* **3**, 619–33.

MacPherson, D., Conkrite, K., Tam, M., Mukai, S., Mu, D. and Jacks, T. (2007). Murine bilateral retinoblastoma exhibiting rapid-onset, metastatic progression and N-myc gene amplification. *EMBO J.* **26**, 784–94.

Magnaghi-Jaulin, L., Groisman, R., Naguibneva, I., Robin, P., Lorain, S., Le Villain, J. P., Troalen, F., Trouche, D. and Harel-Bellan, A. (1998). Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature* **391**, 601–5.

Mayol, X., Graña, X., Baldi, A., Sang, N., Hu, Q. and Giordano, A. (1993). Cloning of a new member of the retinoblastoma gene family (pRb2) which binds to the E1A transforming domain. *Oncogene* **8**, 2561–6.

Milet, C., Rincheval-Arnold, A., Mignotte, B. and Guénal, I. (2010). The Drosophila retinoblastoma protein induces apoptosis in proliferating but not in post-mitotic cells. *Cell Cycle* **9**, 97–103.

Moon, N.-S., Di Stefano, L. and Dyson, N. (2006). A gradient of epidermal growth factor receptor signaling determines the sensitivity of rbf1 mutant cells to E2F-dependent apoptosis. *Mol. Cell. Biol.* **26**, 7601–15.

Narasimha, A. M., Kaulich, M., Shapiro, G. S., Choi, Y. J., Sicinski, P. and Dowdy, S. F. (2014). Cyclin D activates the Rb tumor suppressor by mono-phosphorylation. *Elife* **3**,.

Nicolas, E., Roumillac, C. and Trouche, D. (2003). Balance between acetylation and methylation of histone H3 lysine 9 on the E2F-responsive dihydrofolate reductase promoter. *Mol. Cell. Biol.* **23**, 1614–22.

Nicolay, B. N., Gameiro, P. A., Tschöp, K., Korenjak, M., Heilmann, A. M., Asara, J. M., Stephanopoulos, G., Iliopoulos, O. and Dyson, N. J. (2013). Loss of RBF1 changes glutamine catabolism. *Genes Dev.* **27**, 182–96.

Nielsen, S. J., Schneider, R., Bauer, U. M., Bannister, A. J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R. E., et al. (2001). Rb targets histone H3 methylation and HP1 to promoters. *Nature* **412**, 561–5.

Parisi, T., Balsamo, M., Gertler, F. and Lees, J. A. (2018). The Rb tumor suppressor regulates epithelial cell migration and polarity. *Mol. Carcinog.* **57**, 1640–1650.

Payankaulam, S., Yeung, K., McNeill, H., Henry, R. W. and Arnosti, D. N. (2016). Regulation of cell polarity determinants by the Retinoblastoma tumor suppressor protein. *Sci. Rep.* **6**, 22879.

Raj, N., Zhang, L., Wei, Y., Arnosti, D. N. and Henry, R. W. (2012). Ubiquitination of retinoblastoma family protein 1 potentiates gene-specific repression function. *J. Biol. Chem.* **287**, 41835–43.

Robanus-Maandag, E., Dekker, M., van der Valk, M., Carrozza, M. L., Jeanny, J. C., Dannenberg, J. H., Berns, A. and te Riele, H. (1998). p107 is a suppressor of retinoblastoma development in pRb-deficient mice. *Genes Dev.* **12**, 1599–609.

Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L. and Wolffe, A. P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat. Genet.* **25**, 338–42.

Rubin, S. M., Gall, A.-L., Zheng, N. and Pavletich, N. P. (2005). Structure of the Rb C-terminal domain bound to E2F1-DP1: a mechanism for phosphorylation-induced E2F release. *Cell* **123**, 1093–106.

Sadasivam, S. and DeCaprio, J. A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. *Nat. Rev. Cancer* **13**, 585–95.

Saeed, M., Schwarze, F., Loidl, A., Meraner, J., Lechner, M. and Loidl, P. (2012). In vitro phosphorylation and acetylation of the murine pocket protein Rb2/p130. *PLoS One* **7**, e46174.

Seifried, L. A., Talluri, S., Cecchini, M., Julian, L. M., Mymryk, J. S. and Dick, F. A. (2008). pRB-E2F1 complexes are resistant to adenovirus E1A-mediated disruption. *J. Virol.* **82**, 4511–20.

Sengupta, S., Lingnurkar, R., Carey, T. S., Pomaville, M., Kar, P., Feig, M., Wilson, C. A., Knott, J. G., Arnosti, D. N. and Henry, R. W. (2015). The Evolutionarily Conserved C-terminal Domains in the Mammalian Retinoblastoma Tumor Suppressor Family Serve as Dual Regulators of Protein Stability and Transcriptional Potency. *J. Biol. Chem.* **290**, 14462–75.

Sengupta, S. and Henry, R. W. (2015). Regulation of the retinoblastoma–E2F pathway by the ubiquitin–proteasome system. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1849**, 1289–1297.

Stengel, K. R., Thangavel, C., Solomon, D. A., Angus, S. P., Zheng, Y. and Knudsen, E. S. (2009). Retinoblastoma/p107/p130 pocket proteins: protein dynamics and interactions with target gene promoters. *J. Biol. Chem.* **284**, 19265–71.

Stevaux, O., Dimova, D. K., Ji, J.-Y., Moon, N. S., Frolov, M. V and Dyson, N. J. (2005). Retinoblastoma family 2 is required *in vivo* for the tissue-specific repression of dE2F2 target genes. *Cell Cycle* **4**, 1272–80.

Stevaux, O., Dimova, D., Frolov, M. V, Taylor-Harding, B., Morris, E. and Dyson, N. (2002). Distinct mechanisms of E2F regulation by Drosophila RBF1 and RBF2. *EMBO J.* **21**, 4927–37.

Strober, B. E., Dunaief, J. L., Guha and Goff, S. P. (1996). Functional interactions between the hBRM/hBRG1 transcriptional activators and the pRB family of proteins. *Mol. Cell. Biol.* **16**, 1576–83.

Takahashi, Y., Rayman, J. B. and Dynlacht, B. D. (2000). Analysis of promoter binding by the E2F and pRB families *in vivo*: distinct E2F proteins mediate activation and repression. *Genes Dev.* **14**, 804–16.

Taylor-Harding, B., Binné, U. K., Korenjak, M., Brehm, A. and Dyson, N. J. (2004). p55, the Drosophila ortholog of RbAp46/RbAp48, is required for the repression of dE2F2/RBF-regulated genes. *Mol. Cell. Biol.* **24**, 9124–36.

Tedesco, D., Lukas, J. and Reed, S. I. (2002). The pRb-related protein p130 is regulated by phosphorylation-dependent proteolysis via the protein-ubiquitin ligase SCF(Skp2). *Genes Dev.* **16**, 2946–57.

Trimarchi, J. M. and Lees, J. A. (2002). Sibling rivalry in the E2F family. *Nat. Rev. Mol. Cell Biol.* **3**, 11–20.

Trouche, D., Le Chalony, C., Muchardt, C., Yaniv, M. and Kouzarides, T. (1997). RB and hbrm cooperate to repress the activation functions of E2F1. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11268–73.

Uchida, C., Miwa, S., Kitagawa, K., Hattori, T., Isobe, T., Otani, S., Oda, T., Sugimura, H., Kamijo, T., Ookawa, K., et al. (2005). Enhanced Mdm2 activity inhibits pRB function via ubiquitin-dependent degradation. *EMBO J.* **24**, 160–9.

Ullah, Z., Buckley, M. S., Arnosti, D. N. and Henry, R. W. (2007). Retinoblastoma Protein Regulation by the COP9 Signalosome. *Mol. Biol. Cell* **18**, 1179–1186.

Vandel, L., Nicolas, E., Vaute, O., Ferreira, R., Ait-Si-Ali, S. and Trouche, D. (2001). Transcriptional repression by the retinoblastoma protein through the recruitment of a histone methyltransferase. *Mol. Cell. Biol.* **21**, 6484–94.

Viatour, P. and Sage, J. (2011). Newly identified aspects of tumor suppression by RB. *Dis. Model. Mech.* **4**, 581–585.

Wei, Y., Mondal, S. S., Mouawad, R., Wilczyński, B., Henry, R. W. and Arnosti, D. N. (2015). Genome-Wide Analysis of *Drosophila* RBf2 Protein Highlights the Diversity of RB Family Targets and Possible Role in Regulation of Ribosome Biosynthesis. *G3,* **5**, 1503–1515.

Wirt, S. E. and Sage, J. (2010). p107 in the public eye: an Rb understudy and more. *Cell Div.* **5**, 9.

Wu, L., de Bruin, A., Saavedra, H. I., Starovic, M., Trimboli, A., Yang, Y., Opavska, J., Wilson, P., Thompson, J. C., Ostrowski, M. C., et al. (2003). Extra-embryonic function of Rb is essential for embryonic development and viability. *Nature* **421**, 942–7.

Xiao, B., Spencer, J., Clements, A., Ali-Khan, N., Mittnacht, S., Broceño, C., Burghammer, M., Perrakis, A., Marmorstein, R. and Gamblin, S. J. (2003). Crystal structure of the retinoblastoma tumor suppressor protein bound to E2F and the molecular basis of its regulation. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2363–8.

Xin, S., Weng, L., Xu, J. and Du, W. (2002). The role of RBF in developmentally regulated cell proliferation in the eye disc and in Cyclin D/Cdk4 induced cellular growth. *Development* **129**, 1345–56.

Zhang, L., Wei, Y., Pushel, I., Heinze, K., Elenbaas, J., Henry, R. W. and Arnosti, D. N. (2014). Integrated stability and activity control of the Drosophila Rbf1 retinoblastoma protein. *J. Biol. Chem.* **289**, 24863–73.

CHAPTER 2

Role of *Drosophila* retinoblastoma protein instability element in cell growth and proliferation

Abstract

The RB tumor suppressor, a regulator of the cell cycle, apoptosis, senescence, and differentiation, is frequently mutated in human cancers. We recently described an evolutionarily conserved C-terminal "instability element" (IE) of the *Drosophila* Rbf1 retinoblastoma protein that regulates its turnover. Misexpression of wild-type or non-phosphorylatable forms of the Rbf1 protein leads to repression of cell cycle genes. In contrast, overexpression of a defective form of Rbf1 lacking the IE (ΔIE), a stabilized but transcriptionally less active form of the protein, induced ectopic S phase in cell culture. To determine how mutations in the Rbf1 IE may induce dominant effects in a developmental context, we assessed the impact of in vivo expression of mutant Rbf1 proteins on wing development. ΔIE expression resulted in overgrowth of larval wing imaginal discs and larger adult wings containing larger cells. In contrast, a point mutation in a conserved lysine of the IE (K774A) generated severely disrupted, reduced wings. These contrasting effects appear to correlate with control of apoptosis; expression of the pro-apoptotic *reaper* gene and DNA fragmentation measured by acridine orange stain increased in flies expressing the K774A isoform and was suppressed by expression of Rbf1ΔIE. Intriguingly, cancer associated mutations affecting RB homologs p130 and p107 may similarly induce dominant phenotypes.

This work was published as the following manuscript:

Elenbaas, J.S., Mouawad, R., Henry, R.W., Arnosti, D.N. and Payankaulam, S. Role of Drosophila retinoblastoma protein instability element in cell growth and proliferation. *Cell Cycle*. **14**(4):589–597.

My contribution to this study was showing that mutant forms of Rbf1 have impacts on apoptosis and show differential gene regulation when expressed in wing imaginal discs. Wing imaginal discs expressing a mutant form of Rbf1 show distorted tissue architecture.

Introduction

The retinoblastoma (Rb) protein functions as a regulator of cell cycle in multicellular eukaryotes, enabling progression of mitosis in a seamless manner. Rb is also key to the unfolding of developmental programs through its effects on differentiation and apoptosis. In light of its role in these central cellular processes, it is not surprising that the Rb gene or its regulatory pathway is disrupted in most human cancers (Nevins, 2001). The activity of the Rb protein is tightly regulated during the cell cycle. Hyper-phosphorylation of Rb during the late G1 phase by the activity of CDK/cyclin enzymes results in its inactivation throughout S, G2, and M phases (Rubin, 2013; Narasimha et al., 2014). Mammalian Rb and the homologous family members p130 and p107 are also subject to regulated protein turnover by proteasome dependent and independent pathways, a property shared by the *Drosophila* Rbf1 homolog (Ying and Xiao, 2006; Acharya et al., 2010).

We previously showed that the *Drosophila* Rbf1 protein is protected from turnover by the COP9 regulatory complex, and that a C-terminal instability element (IE) of the protein mediates turnover of the protein (Ullah et al., 2007). Deletion of or point mutations in the IE stabilize Rbf1, and recent studies indicate that the IE is a conserved feature in mammalian Rb family proteins (Sengupta et al., 2015). At the same time, the IE appears to be critical for the transcriptional activity of Rbf1; removal of the entire IE inhibits Rbf1 activity on some but not all target genes in cell culture, while mutations that eliminate phosphorylation targets, or a conserved lysine 774, can exhibit marked hypermorphic effects (Acharya et al., 2010; Raj et al., 2012; Zhang et al., 2014). We were particularly interested in 2 classes of mutation; that which eliminated the IE entirely, and mutations affecting K774. The ΔIE mutant protein induces ectopic cell cycles when expressed in cultured cells, and similar forms of proteins may be produced in cancer cells with nonsense mutations that eliminate the C-termini of Rb family proteins. Mutations affecting K774

did not significantly impact transcriptional activity in cell culture, but the mutant protein has dramatically disruptive effects on eye development in the fly (Acharya et al., 2010). Interestingly, mutations in human p130 residue K1083 (homologous to K774 in Rbf1) have been reported in human lung cancer (Claudio et al., 2000), although the frequency of occurrence of this lesion is not known. Because of the potential relevance of IE mutations to cancer, we assessed the developmental importance of both of these classes of mutation to Rbf1 in the wing, a highly sensitive system for quantitative assessment of morphological impacts and molecular effects on gene expression.

Results

Phenotypes induced by expression of mutant RBF1 proteins.

To understand the functional consequence of mutations affecting the Rbf1 IE in a physiological setting, we overexpressed Rbf1, Rbf1ΔIE and K774R/A in larval wing imaginal discs using a *pendulin GAL4* driver (Figure 2-1A). Flies expressing Rbf1 appeared to have slightly smaller wings and had notches along the wing margins as previously noted (Milet et al., 2010). Expression of K774A and K774R had a much more severe effect, inducing significant size reduction and disruption of wing morphology, similar to its dramatic effect on eye development (Zhang et al., 2014) (Figure 2-1B). Expression of Rbf1ΔIE did not induce gross disruption of wing development, but adult wings (Figure 2-1B) and wing imaginal discs (Figure 2-1C) dissected from third instar larvae expressing Rbf1ΔIE appeared to be slightly larger than those expressing wild-type Rbf1 or a control GFP protein. Discs from crosses expressing K774A were significantly smaller with perturbed tissue architecture (Figure 2-1C).

Previous studies with the Rbf1ΔIE mutant had not identified a biological activity of this protein when expressed in developing eyes, but our recent observations that the protein induces S phase entry in cultured cells, together with the transcriptional repression activity on certain promoters led us to quantitatively examine the effect on wing development. We used the WINGMACHINE tool (Rohlf et al., 2003) to measure controls and wings in which Rbf1ΔIE had been overexpressed. We observed a statistically significant ~4% increase in the wing size of both males and females with expression of Rbf1ΔIE (Figure 2-2A, B and Table 2-1). Patterning of the wings was unaffected. Similar increases in wing size were noted when Rbf1ΔIE was expressed with a wing-specific *beadex* driver (not shown).

Expression of Rbf1ΔIE increases cell size.

To determine whether the increase in wing size was due to increase in cell number, size or both, we measured the numbers of trichomes in a defined area of the wing. Single trichomes are produced by individual cells in adult wings. The number of trichomes and area calculations provides a basis to determine cell size and density. Measured cell size was significantly larger in wings of both males and females expressing Rbf1ΔIE (Figure 2-2C and Table 2-1) and Rbf1 (data not shown). The stronger effect in the male may reflect the X-chromosomal location of the endogenous *rbf1* gene; the hemizygosity of the males may lead to stronger perturbations of the Rbf1 regulon upon misexpression of the transgene.

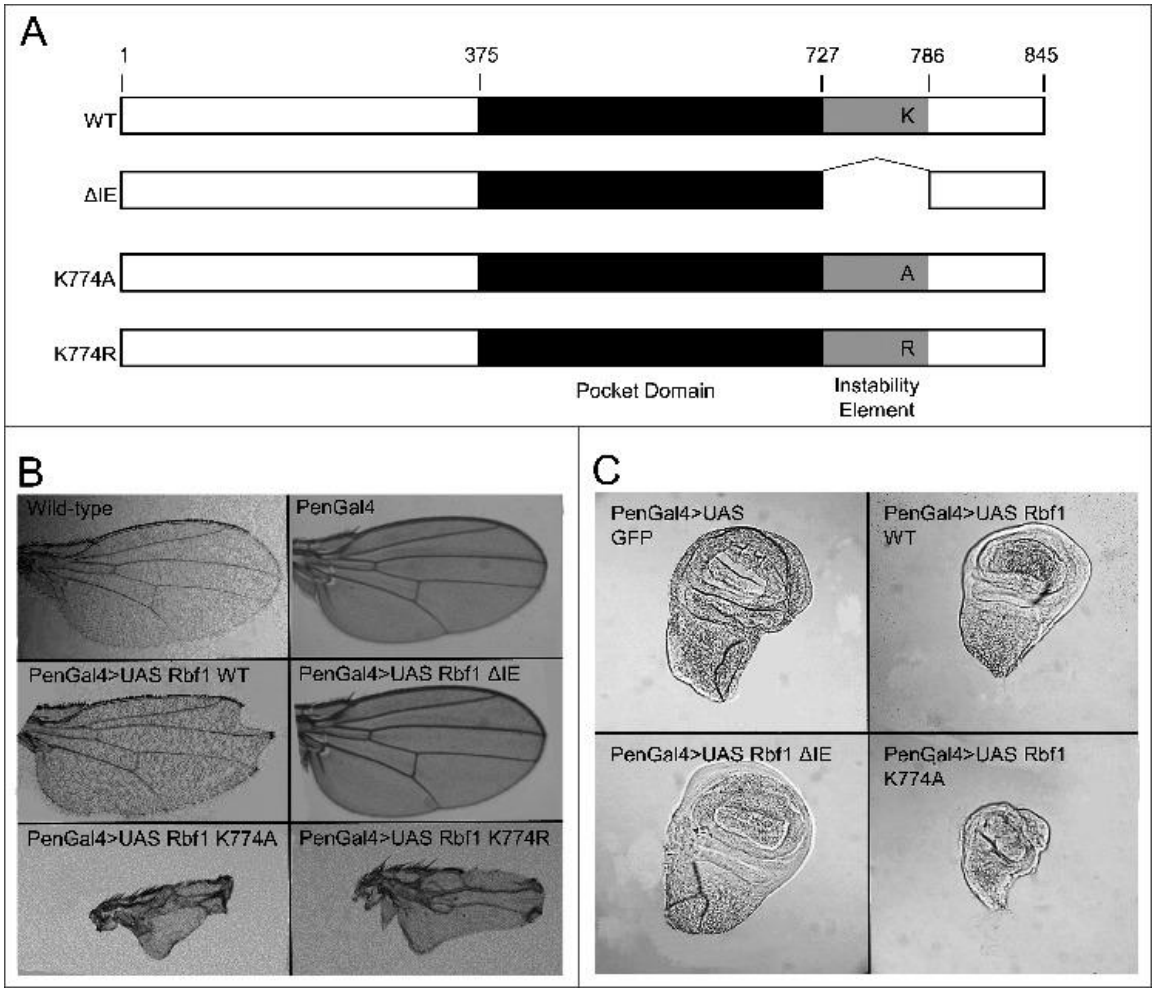


Figure 2-1: Mutant Rbf1 IE isoforms induce dominant and contrasting phenotypes. (A) Schematic diagram of wild type and mutant Rbf1 proteins. The E2F binding domain is shown in black and the instability element in gray. The instability element was excised in the mutant labeled ΔIE. Residue 774 was mutated to either a non-conservative alanine or a conservative arginine in 2 additional mutant proteins. (B) Wing phenotypes of adult flies expressing mutant isoforms. Representative images show the observed phenotype for each of the overexpressed proteins. The line bearing the *PenGal4* alone showed no observable phenotype. The *PenGal4* > UAS *Rbf1* WT flies exhibited a notched phenotype, while the *PenGal4* > UAS *Rbf1ΔIE* exhibited a slight increase in wing size. Wings from crosses expressing *Rbf1K774A* or *K774R* exhibited dramatic

Figure 2-1 (cont'd)

decreases in size among other defects. All images were taken at 4× magnification and in each case more than 30 wings were examined. (C) Third instar larval wing imaginal discs of Rbf1 mutants showing distinct growth response. The wing discs of the mutant flies were dissected from third instar larvae and photographed. The control *PenGal4 >* UAS *GFP* flies had discs that were indistinguishable from wild-type wing discs. Discs expressing wild-type Rbf1 appeared to be slightly reduced in size but showed no obvious defects in gross morphology. The *PenGal4>* UAS *Rbf1ΔIE* discs were noticeably larger in size compared to wild-type discs, while discs expressing Rbf1K774A were much smaller than wild type and showed dramatic morphological defects. Shown are the most commonly observed phenotypes for each transgenic line, representing approximately 75% of at least 100 discs observed for each genotype.

Table 2-1: Tabulated results for wing and cell size measurements from crosses expressing Rbf1ΔIE.

	♀			♂		
Genotype	UAS Rbf1ΔIE	PenGal4	PenGal4 UAS Rbf1ΔIE	UAS Rbf1ΔIE	PenGal4	PenGal4 UAS Rbf1ΔIE
N	25	24	22	24	23	22
Mean surface area (mM^2)	4.74±0.14	4.69±0.08	4.90±0.05	4.00±0.09	4.01±0.08	4.18±0.07
Percent increase in surface area			4.0%			4.4%
N	42	52	50	26	23	21
Mean cell size (pixels/trichome)	53.9±2.6	54.7±2.8	61.1±2.3	46.2±2.3	49.4±2.7	57.3±2.3
Percent increase in cell size			12%			16%

Mean surface area and cell size were tabulated with standard deviations. The increase in surface area was approximately 4% in both females and males, compared to those of the parental lines.

Total wing area was deduced from 22-25 wings and 21-52 for cell size measurements. There was a 16% increase in cell size for males and 12% in females within the area measured.

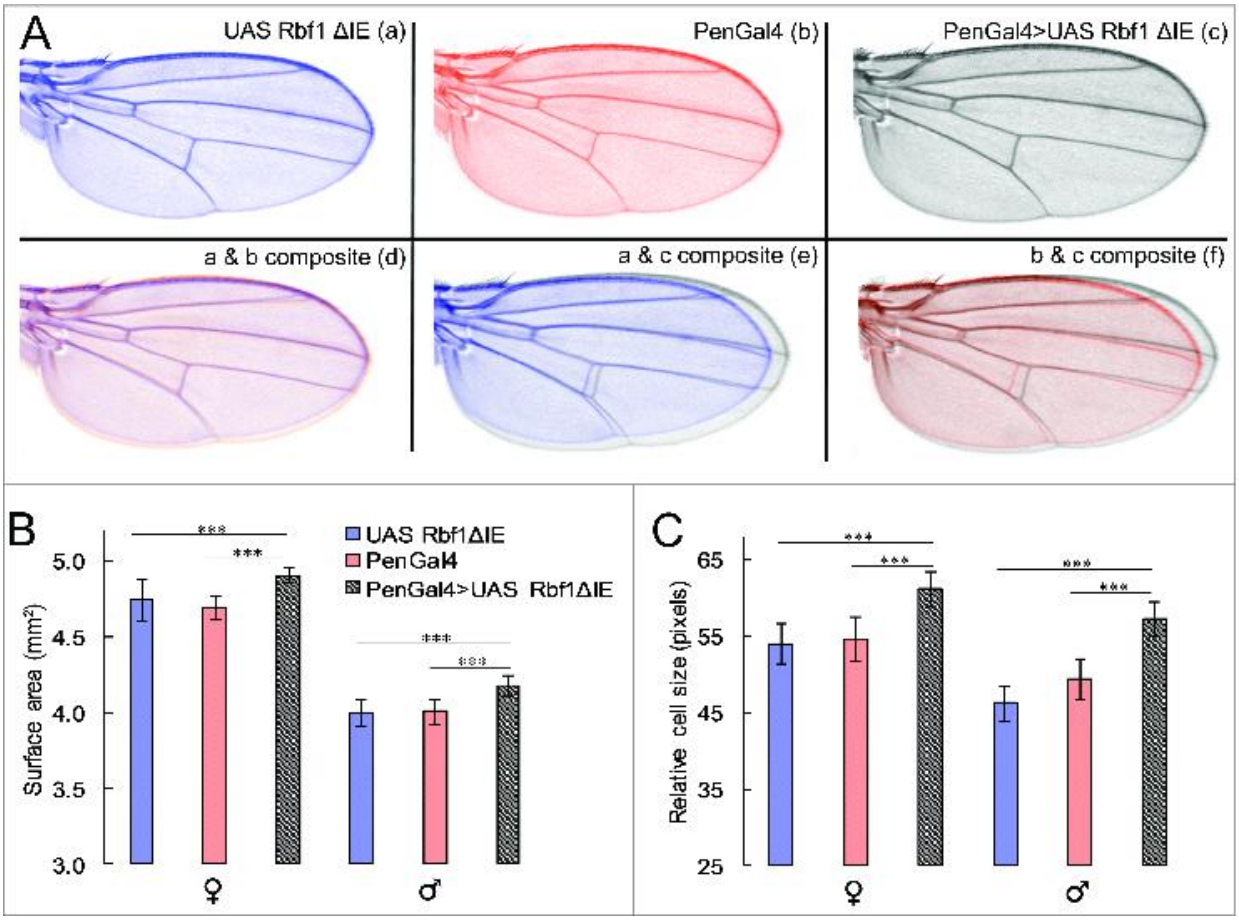


Figure 2-2: Rbf1ΔIE expression causes an increase in wing and cell size (A) Representative wing images were chosen from the test and parental lines. (a,b,c) Representative images for the 2 parental lines and the test line. (d) Composite images of the parental lines. No significant size difference was noticed between the 2 parental lines. (e,f) The composite image of the test line and the parental lines revealed a significant increase in both length and width of the test line compared to the parental lines. (B) Surface area was measured using WINGMACHINE software. Males and

Figure 2-2 (cont'd)

females were evaluated separately due to sex-specific differences in wing size. Only right wings were measured. Both females and males show very significant (***$P < 0.001$, n = 22 to 25 wings) increases in surface area compared to both parental strains. (C) Cell size was measured using Fijiwings. Cell size was calculated using reported values for area measured and number of trichomes counted (n = 21 to 52 wings were used). The males showed greater increase in cell size compared to the females. Error bars represent standard deviation.

Rbf1 isoforms induce contrasting apoptotic responses.

Proliferation of imaginal disc tissue reflects a delicate balance of signaling processes that involve developmentally-regulated cell division and apoptosis. The Rbf1 protein and the mutant forms used in this study have been tested previously for protein expression and stability. Our studies show that transfected S2 cell cultures express Rbf1 and K774A isoforms at comparable levels, and exhibit similar protein stability. In contrast the Rbf1ΔIE protein is expressed at higher levels and has a longer half-life (Acharya et al., 2010; Zhang et al., 2014). The opposing effects on proliferation noted for the Rbf1ΔIE and K774A/R alleles of *Rbf1* therefore may reflect different impacts on apoptosis. We stained wing discs with the vital dye acridine orange, which is a particularly useful tool in identifying apoptotic bodies in the live tissue (Abrams et al., 1993; Arama et al., 2006). Wild-type wing discs in late third instar larvae show low levels of apoptosis, usually restricted to the notum-wing boundary area. In discs with expression of extra wild-type Rbf1, we observed increased acridine orange staining in the wing pouch, consistent with an earlier report (Milet et al., 2010) (Figure 2-3A). These levels were dramatically higher in discs where the K774A isoform was expressed, while discs with Rbf1ΔIE overexpression showed little apoptosis, similar to wild-type wing discs. Consistent with these observations, the pro-apoptotic

gene *reaper* was found to be strongly induced in discs expressing the K774A mutant protein, and suppressed in discs expressing the Rbf1ΔIE protein (Figure 2-3B). *p53,* a likely regulator of this gene, was similarly expressed at lowest levels in the Rbf1ΔIE background, and possibly modestly up regulated in the K774A discs. The expression of other pro-apoptotic genes, including *hid, grim* and *sickle,* did not show any significant changes (data not shown for *grim* and *sickle,* their levels were found to be extremely low in all samples; >14 fold lower than PCNA). *RpL37a* control ribosomal protein gene, showed no significant changes in expression, while 2 canonical Rbf1 cell-cycle related genes, *PCNA* and *polα*, were down regulated in response to all forms of the protein. Thus, the specific effects of overexpression of each form of Rbf1 appear to be associated with differential apoptotic responses.

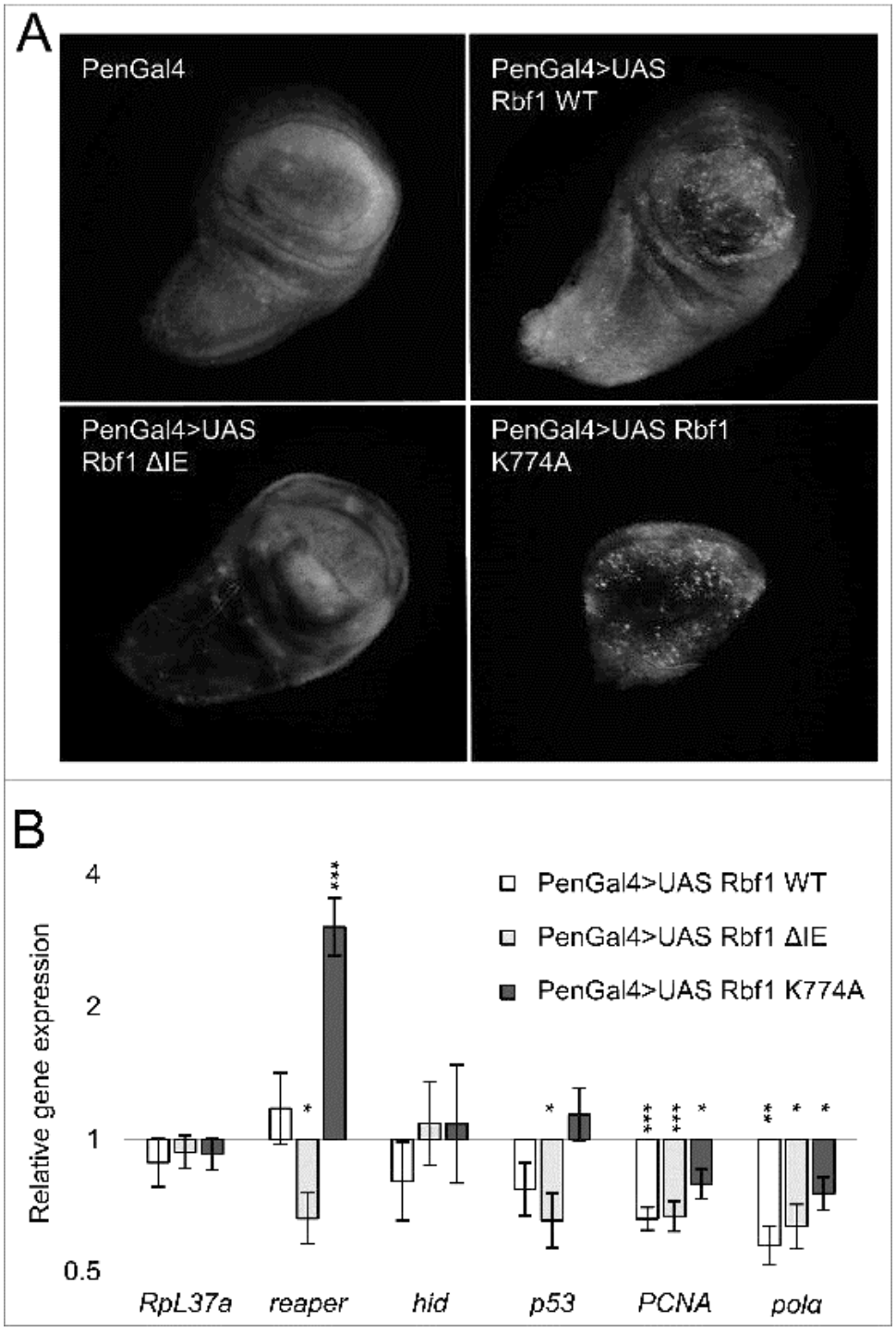


Figure 2-3: Differing apoptotic response to overexpression of Rbf1 isoforms. (A) Visualization of apoptosis in third instar larval wing imaginal discs. Wing discs were stained with acridine orange to examine apoptotic activity. No acridine-positive cells were observed in wild-type discs

Figure 2-3 (cont'd)

expressing only Gal4 or Rbf1ΔIE. Discs from flies expressing Rbf1 and Rbf1K774A showed numerous brightly stained spots, indicating increased apoptosis. Apoptosis was centralized in the wing pouch of flies expressing Rbf1 while the flies expressing Rbf1 K774A showed significant apoptosis throughout the wing disc. In each case 10 wings were stained and analyzed. (B) Distinct changes in the transcript levels of reaper in larval wing imaginal discs. reaper transcripts were reduced in discs from Rbf1ΔIE expressing flies, and strongly elevated in discs from Rbf1K774A expressing flies. hid transcript levels were not significantly different in any of the tested backgrounds; the higher variability reflects the very low expression level of this gene. Rbf1ΔIE expressing discs had significantly lower levels of p53 transcript levels, while changes in levels were not significantly different in other backgrounds. Levels of PCNA and Polα were reduced with expression of all isoforms. RpL37a showed no significant change in expression among the different lines. Transcript levels were normalized to those measured in discs with PenGal4 > GFP. Values represent averages of 5 biological replicates (*P < 0.05; **P < 0.01; ***P < 0.001) and error bars represent standard deviation.

Expression of Rbf1 or Rbf1 ΔIE increases disc cell size with no effect on cell cycle phasing.

To examine the effect of Rbf1 or Rbf1ΔIE overexpression on cell cycle phasing, we dissociated wing imaginal discs from late third instar larvae overexpressing Rbf1 or Rbf1 ΔIE and measured DNA content and cell size by fluorescence activated cell sorting (FACS) (Figure 2-4). Misexpression of Rbf1 or Rbf1ΔIE increased cell size as seen by the rightward shift in the mean of the histogram, indicating larger cell size, while the percentages of cells in each phase of the cell cycle were unaffected in both when compared to control discs. These effects on cell cycle and cell

size are similar to previous observations for Rbf1 overexpression (Neufeld et al., 1998; Prober et al., 2000).

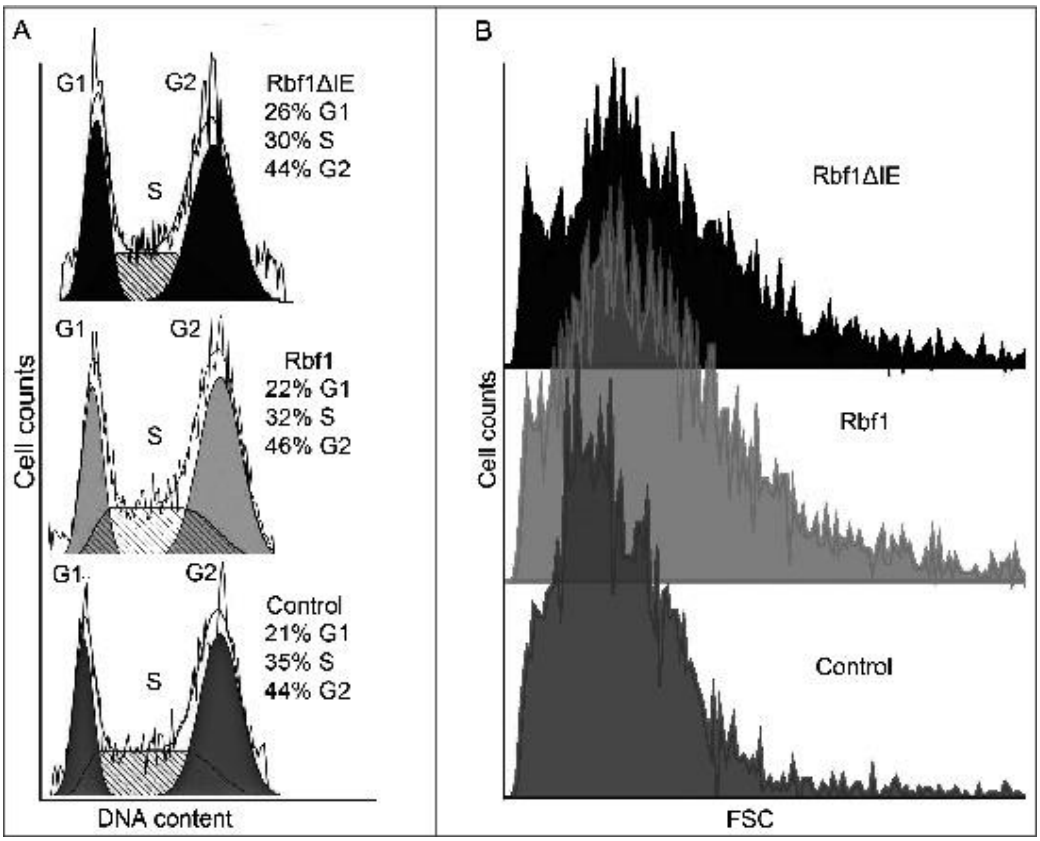


Figure 2-4: Overexpression of Rbf1 and Rbf1ΔIE increases wing imaginal disc cell size with no effect on cell cycle phasing. (A) FACS analysis of Rbf1 and Rbf1ΔIE wing disc cells shows no significant change in cell cycle phasing compared to cells from control discs. Numbers represent the percentage of cells in each phase. A typical cell cycle profile is represented here (n = 3). (B) Cell size as measured by forward scatter (FSC). FSC analyses indicate that misexpression of Rbf1 (Mean = 1.24) and Rbf1ΔIE (Mean = 1.21) increases cell size as observed by the rightward shift in the mean of scatter intensity when compared to control (n = 2). Mean value is obtained by taking the ratio of forward scatter intensity value of Rbf or RbfΔIE to that of the control.

Discussion

Role of Rb/E2F in apoptosis

Although many mutations affecting RB in human cancer are thought to constitute a loss of function, there are specific cases in which elevated RB protein levels positively correlate with disease severity (Yamamoto et al, 1999). Additionally, certain types of mutations in RB, p130, and p107 may not be inactivating, but rather generate hypo- or neomorphic forms of the proteins whose activities may contribute to cellular transformation. Indeed, mutations in *Drosophila* Rbf1 can generate a proliferative phenotype in cell culture (Raj et al., 2012). Our studies show that the *Drosophila* Rbf1 protein C-terminal IE domain affects protein stability and activity, generating gene-specific regulatory effects. This regulation through C-terminal IE-like domains is highly conserved in vertebrate RB family proteins (Sengupta et al., 2015), thus we determined here how changes to IE function impact Rbf-mediated developmental processes.

One of the most striking findings was the opposing effects on apoptosis produced by different lesions in Rbf1; the removal of the IE in its entirety suppressed apoptosis, while the point mutation K774A dramatically enhanced levels of this response. These functional differences are unlikely to be solely due to an overexpression artifact because gene expression analysis shows that all 3 isoforms repress *pcna*, and other canonical target genes in a similar manner. Rbf1 and its binding partner E2F1 have been previously linked to induction of apoptosis in *Drosophila*. Elevated levels of E2F1 induces pro-apoptotic genes such as *p53, Ark/Apaf1, hid*, and *reaper* (Asano et al., 1996; Du et al., 1996; Nicholson et al., 2009). While *hid* appears to be responsible for apoptosis in the eye discs, loss of Rbf1 also causes apoptosis in wing imaginal discs (Duman-Scheel et al., 2004; Moon et al., 2006; Steele et al., 2009). At the same time, expression of Rbf1 can also induce apoptosis in proliferating cells, an effect that is suppressed by ectopic E2F (Asano et al., 1996).

The “threshold” model of E2F activation poses that a precise balance of Rbf1 to E2F1 may be essential to avoid induction of this response (Ziebold et al., 2001). In our case, the most striking effects on apoptosis were those produced by mutant forms of Rbf1, both of which are competent for transcriptional regulation. Rbf1ΔIE, a protein entirely lacking the IE regulatory domain, may suppress apoptosis because this stabilized protein may continue to repress pro-apoptotic genes under circumstances where the wild-type protein is destroyed. Note that the differential impact on *reaper*, a pro-apoptotic gene, is likely to be indirect, as this gene is not found to be bound by Rbf1 in ChIP-Seq experiments (Acharya et al., 2012). The suppression of *p53* expression may be important in this context, as p53 is an activator of *reaper* (Brodsky et al., 2000).

The strong pro-apoptotic effect of Rbf1K774A requires a different explanation; this protein was somewhat less effective in repression of *PCNA, polα*, and *p53*, thus it is possible that weaker effects on a broad range of target genes may induce an apoptotic response. One likely candidate would again be the *p53* gene, which appears to be differentially regulated by Rbf1K774A compared to the effects of wild-type Rbf1 overexpression. Alternatively, or in addition, Rbf1K774A may displace endogenous Rbf1 but fail to effectively repress specific pro-apoptotic genes. Another possibility is that this protein may activate pro-apoptotic genes, similar to the activating role that Rb has on pro-apoptotic genes under conditions of DNA damage (Ianari et al., 2009).

The notion that gene-specific readouts may reflect contributions of different portions of Rbf1, in this case the regulatory IE domain, is supported by structural analysis of the human Rb protein. The mammalian Rb protein makes different types of contacts with members of the E2F family; certain interactions mediated by the Rb pocket domain appear to involve all E2F family members, while other interactions provide discrimination between E2F family members (Dick and Dyson,

2003; Julian et al., 2008; Dick and Rubin, 2013). An illustration of how specific interactions can control apoptotic responses stems from analysis of mutant Rb proteins in cell culture, where disruption of C-terminal interactions does not eliminate repression of cell cycle promoters such as p107 and cyclin E1, but it does abrogate apoptosis. Other mutations in the pocket and the C-terminal domains have an enhanced ability to repress apoptosis (Cecchini and Dick, 2011). Specific interactions between Rb and E2F proteins appear to be conserved. Residues in the C terminus of mammalian Rb make specific contacts with E2F; alanine substitutions in M851A, and V852A, which is conserved in the Drosophila IE ((Sengupta et al., 2015), abolish interaction with E2f1 (Dick and Dyson, 2003; Julian et al., 2008). Thus mutations in the IE are likely to alter E2F interactions, preferentially affecting a subset of Rbf1 targets, with consequent effect on apoptosis. Genome-wide approaches will be helpful to identify such targets.

Consistent with our observations about the effects of Rbf1 overexpression, previous studies have shown that such perturbations increase cell size and cell doubling time (Neufeld et al., 1998). The slow progression through the cell cycle may permit increased accumulation of cell mass, leading to larger cells. Strikingly, only in the case of Rbf1ΔIE overexpression does this result in larger wings presumably because of this protein's anti-apoptotic activity. Increased apoptosis in the case of Rbf overexpression leads to smaller, notched adult wings, despite the larger cell size.

During development, activities of complex signaling pathways normally render imaginal disc growth resistant to perturbation. For instance, experimentally induced apoptosis during early stages of wing development is compensated by increased proliferation, resulting in discs of normal size (Weinkove and Leevers, 2000). Besides cell cycle control programs, a variety of signaling mechanisms including the ecdysone, insulin, Wnt, Dpp, Notch, and Hippo pathways are responsible for coordinated growth rates between and within imaginal discs (Justice et al., 1995;

Burke and Basler, 1996; Serrano and O'Farrell, 1997; Weinkove et al., 1999; Dominguez and de Celis, 1998; Montagne et al., 1999; Scanga et al., 2000; Shingleton, 2005). Expression of Rbf1ΔIE appears to override such controls, resulting in significantly larger cells and measurably larger wings. One possible molecular mechanism may involve changes to Hippo signaling, as the Yorkie effector of this pathway has been demonstrated to co-regulate a number of promoters with E2F1 in *Drosophila* (Nicolay et al., 2011). We previously reported that removal of the IE domain has promoter-specific effects, which may differentially impact signaling pathway genes (Raj et al., 2012). Interestingly, the connection between Rb family proteins and Hippo signaling is evolutionarily conserved, as proliferative controls in mouse hepatocytes are dependent on both E2F/Rb family and Hippo signaling, and may be both affected in hepatocarcinomas (Ehmer et al., 2014).

Despite the tremendous progress in our understanding of Rb's role beyond cell cycle (Nicolay and Dyson, 2013), a major question regarding the function of specific Rb mutations particularly in disease such as cancer remains obscure. Our study argues that lesions affecting Rb family proteins may contribute to cancer in ways beyond simple loss of function. Cancer sequencing projects (Forbes et al., 2011) have identified a number of mutations in p107 and p130 that result in truncation of the C-terminal regions of these proteins and subsequent loss of the conserved IE domains. The molecular activities of Rbf1ΔIE are consistent with this hypothesis, as a dominant, proliferative, anti-apoptotic activity would presumably be selected for in the development of tumors. Additional studies are required to determine whether such mutations can induce similar phenotypes in vertebrates, and whether such activities may present interesting new targets in cancer therapy or diagnosis. Less obvious is the significance of the Rbf1K774A mutation to development of cancer; if this protein were to be expressed in a cell, it appears that its pro-apoptotic

activity would be selected against, unless the perturbed signaling were different in the context of additional mutations accompanying cellular transformation. Determination of the mechanism by which this protein shifts cells to an apoptotic state may be useful for treatment tumors that express Rb family proteins, assuming that Rb, p107, or p130 are similarly affected by such mutations.

Materials and Methods

Fly genetics

The Rbf1 expression lines were constructed as described previously (Acharya et al., 2010; Zhang et al., 2014). Driver lines w[1118] P{*w*[+mW.hs] = GawB}*Bx*[MS1096] (referred to as *Bx*) and *y*[*] *w*[*]; P{*GawB*}NP6333 / CyO,{UAS-*lacZ.UW14*}UW14 (referred to as *Pen*) were obtained from the Bloomington Stock Center. For each experiment, 2 independent crosses were made with 3 virgin females of each genotype with males of the same genotype or differing genotype. The UAS attB *Rbf1δIE* line was balanced over Sm2 marked with *CyO*. Homozygosity was confirmed in previous experiments for females of both driver lines (data not shown). These crosses resulted in offspring with one of the following genotypes: *PenGal4*, *BxGal4*, UAS *Rbf1δIE, PenGal4* > UAS *Rbf1δIE*, and *BxGal4* > UAS *Rbf1δIE*. All crosses were made in parallel and stored at 26°C and 33% humidity. Parent flies were discarded at 9 days after original cross. Adult flies were collected on days 10-19 daily or on alternate days to control population sizes, flies exhibiting the CyO phenotype were discarded. Flies were stored in 80% ethanol in separate vials based on sex and genotype.

Wing photography

Right wings were identified, removed, and washed in 1× PBS. They were then mounted onto slides with mounting solution (70% glycerol, 30% PBS), photographed with an Olympus DP30BW camera mounted on an Olympus BW51 microscope at the same magnification and software settings. Approximately 15 landmarks were identified using tpsDIG software (Rohlf et al., 2003) which was then used to measure the surface area of the wing. To count cells and calculate cell size, wings were photographed at higher resolution with trichomes in focus using an Olympus

BX51 microscope with an Olympus DP30BW camera under the same magnification and settings for all genotypes. Cell density and total trichome number were calculated using the Fijiwings 150px density tool (Dobens et al., 2013). The total number of trichomes were measured in a 150 × 150 pixel area located between L4 and L5 veins and immediately distal to the cross vein. Total numbers of trichomes counted ranged from roughly 300-500 within the area measured. Additional measurements were also made in a 75 × 75 pixel area between L3 and L4 at an equal distance between the intersection of the cross veins, with identical results. ANOVA tests were performed followed by post-hoc T-tests with a Bonferroni correction in Microsoft Excel 2013 to determine statistical significance.

Acridine orange staining

Third instar larvae of similar age were dissected in PBS 1×. Wing imaginal discs were collected and incubated 3 min in 0.6 mg/ml acridine orange/PBS 1× solution. Wing discs (approximately 8 from each line) were then rinsed in PBS 1× and rapidly photographed for fluorescence.

qRT-PCR

Wing imaginal discs were dissected from third-instar larvae, and total RNA was isolated according to the Kreitman (Kreitman, 2012) protocol using TRIzol (Invitrogen) followed by Rneasy Mini kit (Qiagen) for cleanup. 300 ng of total RNA was converted to cDNA using High Capacity cDNA Reverse Transcription kit (Applied Biosystems). The resulting cDNA was diluted 1:10 and 3 µl was used for PCR in a 20µl reaction mixture using SYBR green PCR Master Mix (Applied Biosystems). qPCR was performed on five biological replicates for each Rbf1 isoform. The fold change in gene expression was calculated based on ΔCt analysis method and normalized to *rp49* gene expression. Figure 2-3B represent data from four biological

replicates. Primers used for gene analysis were as follows: *RpL37a* forward CCTTCACGGACCAGTTGTAG, *RpL37a* reverse ACAATAAGACGCACACCCTG, *reaper* forward CCACCGTCGTCCTGGAAAC, *reaper* reverse CCGGTCTTCGGATGACATG, *p53* forward CCGTGGTCCGCTGTCAA, *p53* reverse TGCGTTATTGGCCGTCAAA, *PCNA* forward TGCAGCGACTCCGGCATTCA, *PCNA* reverse CGGAACGCAGGGTCAGCGAG, *polα* forward TGCTCTCAGATGAATGGAAGG, *polα* reverse TGAAGTGCGAAAGATAGTCCC, *RBF1* forward AAGCAGCTGAGCGCCTTCGG, *RBF1* reverse GCAGCTTGGCTATTACCTCTTCGCC, *hid* forward CGAGGATGAGCGCGAGTAC, *hid* reverse CGCCAAACTCGTCCCAAGT, *rp49* forward ATCGGTTACGGATCGAACAAGC, and *rp49* reverse GTAAA CGCGGTTCTGCATGAGC.

Flow cytometry

40–50 Wing imaginal discs from 3rd instar larvae were dissected in PBS and were incubated for 15 min at room temperature in 200 µl of trypsin solution (trypsin-EDTA, Sigma T4299) containing 3 µg/ml of Hoechst (Hoechst 33342, trihydrochloride trihydrate H3570, Molecular Probes) with gentle agitation. Trypsin digestion was stopped by addition of 300 µl of 1% fetal bovine serum (HI FBS, Gibco) in PBS, and after centrifugation at 3500 rpm for 5 min at 4°C the cells were resuspended in 350 µl of 1% FBS (Andrade-Zapata and Baonza, 2014). The cell cycle profile was analyzed on a BD Influx Sorter. Three independent experiments for cell cycle profiles were analyzed using Winlist version8 software. Two independent experiments were used for the analysis of cell size.

Acknowledgments

We thank Will Pitchers and Anne Sonnenschein for help with wing measurements and analysis, Louis King for help with flow cytometry, Satyaki Sengupta for valuable suggestions, and the Bloomington Stock Center for the fly lines. JSE performed the wing, cell size analysis and prepared the figures for this manuscript; RM measured the transcript levels and performed the staining of wing imaginal discs. Project conception was by SP, and the manuscript was written by SP, RWH, and DNA.

REFERENCES

REFERENCES

Abrams, J. M., White, K., Fessler, L. I. and Steller, H. (1993). Programmed cell death during Drosophila embryogenesis. *Development* **117**, 29–43.

Acharya, P., Raj, N., Buckley, M. S., Zhang, L., Duperon, S., Williams, G., Henry, R. W. and Arnosti, D. N. (2010). Paradoxical Instability–Activity Relationship Defines a Novel Regulatory Pathway for Retinoblastoma Proteins. *Mol. Biol. Cell* **21**, 3890–3901.

Acharya, P., Negre, N., Johnston, J., Wei, Y., White, K. P., Henry, R. W. and Arnosti, D. N. (2012). Evidence for Autoregulation and Cell Signaling Pathway Regulation From Genome-Wide Binding of the *Drosophila* Retinoblastoma Protein. *G3: Genes|Genomes|Genetics* **2**, 1459–1472.

Andrade-Zapata, I. and Baonza, A. (2014). The bHLH Factors Extramacrochaetae and Daughterless Control Cell Cycle in Drosophila Imaginal Discs through the Transcriptional Regulation of the cdc25 Phosphatase string. *PLoS Genet.* **10**, e1004233.

Arama, E. and Steller, H. (2006). Detection of apoptosis by terminal deoxynucleotidyl transferase-mediated dUTP nick-end labeling and acridine orange in Drosophila embryos and adult male gonads. *Nat. Protoc.* **1**, 1725–31.

Asano, M., Nevins, J. R. and Wharton, R. P. (1996). Ectopic E2F expression induces S phase and apoptosis in Drosophila imaginal discs. *Genes Dev.* **10**, 1422–32.

Brodsky, M. H., Nordstrom, W., Tsang, G., Kwan, E., Rubin, G. M. and Abrams, J. M. (2000). Drosophila p53 binds a damage response element at the reaper locus. *Cell* **101**, 103–13.

Burke, R. and Basler, K. (1996). Dpp receptors are autonomously required for cell proliferation in the entire developing Drosophila wing. *Development* **122**, 2261–9.

Cecchini, M. J. and Dick, F. A. (2011). The biochemical basis of CDK phosphorylation-independent regulation of E2F1 by the retinoblastoma protein. *Biochem. J.* **434**, 297–308.

Claudio, P. P., Howard, C. M., Pacilio, C., Cinti, C., Romano, G., Minimo, C., Maraldi, N. M., Minna, J. D., Gelbert, L., Leoncini, L., et al. (2000). Mutations in the retinoblastoma-related gene RB2/p130 in lung tumors and suppression of tumor growth in vivo by retrovirus-mediated gene transfer. *Cancer Res.* **60**, 372–82.

Dick, F. A. and Dyson, N. (2003). pRB contains an E2F1-specific binding domain that allows E2F1-induced apoptosis to be regulated separately from other E2F activities. *Mol. Cell* **12**, 639–49.

Dick, F. A. and Rubin, S. M. (2013). Molecular mechanisms underlying RB protein function. *Nat. Rev. Mol. Cell Biol.* **14**, 297–306.

Dobens, A. C. and Dobens, L. L. (2013). FijiWings: an open source toolkit for semiautomated morphometric analysis of insect wings. *G3 (Bethesda).* **3**, 1443–9.

Domínguez, M. and de Celis, J. F. (1998). A dorsal/ventral boundary established by Notch controls growth and polarity in the Drosophila eye. *Nature* **396**, 276–8.

Du, W., Xie, J. E. and Dyson, N. (1996). Ectopic expression of dE2F and dDP induces cell proliferation and death in the Drosophila eye. *EMBO J.* **15**, 3684–92.

Duman-Scheel, M., Johnston, L. A. and Du, W. (2004). Repression of dMyc expression by Wingless promotes Rbf-induced G1 arrest in the presumptive Drosophila wing margin. *Proc. Natl. Acad. Sci.* **101**, 3857–3862.

Ehmer, U., Zmoos, A.-F., Auerbach, R. K., Vaka, D., Butte, A. J., Kay, M. A. and Sage, J. (2014). Organ size control is dominant over Rb family inactivation to restrict proliferation in vivo. *Cell Rep.* **8**, 371–81.

Forbes, S. A., Bindal, N., Bamford, S., Cole, C., Kok, C. Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945-50.

Ianari, A., Natale, T., Calo, E., Ferretti, E., Alesse, E., Screpanti, I., Haigis, K., Gulino, A. and Lees, J. A. (2009). Proapoptotic function of the retinoblastoma tumor suppressor protein. *Cancer Cell* **15**, 184–94.

Julian, L. M., Palander, O., Seifried, L. A., Foster, J. E. G. and Dick, F. A. (2008). Characterization of an E2F1-specific binding domain in pRB and its implications for apoptotic regulation. *Oncogene* **27**, 1572–1579.

Justice, R. W., Zilian, O., Woods, D. F., Noll, M. and Bryant, P. J. (1995). The Drosophila tumor suppressor gene warts encodes a homolog of human myotonic dystrophy kinase and is required for the control of cell shape and proliferation. *Genes Dev.* **9**, 534–46.

Kreitman:RNA extraction from small amount of samples (imaginal discs). *OpenWetWare*, 2012,April9.http://openwetware.org/index.php?title=Kreitman:RNA_extraction_from_small amount_of_samples_%28imaginal_discs%29&oldid=596945

Milet, C., Rincheval-Arnold, A., Mignotte, B. and Guénal, I. (2010). The Drosophila retinoblastoma protein induces apoptosis in proliferating but not in post-mitotic cells. *Cell Cycle* **9**, 97–103.

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

Moon, N.-S., Di Stefano, L. and Dyson, N. (2006). A Gradient of Epidermal Growth Factor Receptor Signaling Determines the Sensitivity of rbf1 Mutant Cells to E2F-Dependent Apoptosis. *Mol. Cell. Biol.* **26**, 7601–7615.

Narasimha, A. M., Kaulich, M., Shapiro, G. S., Choi, Y. J., Sicinski, P. and Dowdy, S. F. (2014). Cyclin D activates the Rb tumor suppressor by mono-phosphorylation. *Elife* **3**.

Neufeld, T. P., de la Cruz, A. F., Johnston, L. A. and Edgar, B. A. (1998). Coordination of growth and cell division in the Drosophila wing. *Cell* **93**, 1183–93.

Nevins, J. R. (2001). The Rb/E2F pathway and cancer. *Hum Mol Genet.* **10**, 699–703.

Nicholson, S. C., Gilbert, M. M., Nicolay, B. N., Frolov, M. V and Moberg, K. H. (2009). The archipelago tumor suppressor gene limits rb/e2f-regulated apoptosis in developing Drosophila tissues. *Curr. Biol.* **19**, 1503–10.

Nicolay, B. N., Bayarmagnai, B., Islam, A. B. M. M. K., Lopez-Bigas, N. and Frolov, M. V (2011). Cooperation between dE2F1 and Yki/Sd defines a distinct transcriptional program necessary to bypass cell cycle exit. *Genes Dev.* **25**, 323–35.

Nicolay, B. N. and Dyson, N. J. (2013). The multiple connections between pRB and cell metabolism. *Curr. Opin. Cell Biol.* **25**, 735–40.

Prober, D. A. and Edgar, B. A. (2000). Ras1 Promotes Cellular Growth in the Drosophila Wing. *Cell* **100**, 435–446.

Raj, N., Zhang, L., Wei, Y., Arnosti, D. N. and Henry, R. W. (2012). Ubiquitination of retinoblastoma family protein 1 potentiates gene-specific repression function. *J. Biol. Chem.* **287**, 41835–43.

Raj, N., Zhang, L., Wei, Y., Arnosti, D. N. and Henry, R. W. (2012). Rbf1 degron dysfunction enhances cellular DNA replication. *Cell Cycle* **11**, 3731–8.

Rohlf, F.J. (2004) TpsDig Version 1.4. Department of Ecology and Evolution. State University of New York at Stony Brook, New York. http://life.bio.sunysb.edu/morph/index.html

Rubin, S. M. (2013). Deciphering the retinoblastoma protein phosphorylation code. *Trends Biochem. Sci.* **38**, 12–19.

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

Sengupta, S., Lingnurkar, R., Carey, T. S., Pomaville, M., Kar, P., Feig, M., Wilson, C. A., Knott, J. G., Arnosti, D. N. and Henry, R. W. (2015). The Evolutionarily Conserved C-terminal Domains in the Mammalian Retinoblastoma Tumor Suppressor Family Serve as Dual Regulators of Protein Stability and Transcriptional Potency. *J. Biol. Chem.* **290**, 14462–75.

Serrano, N. and O'Farrell, P. H. (1997). Limb morphogenesis: connections between patterning and growth. *Curr. Biol.* **7**, R186-95.

Shingleton, A. W. (2005). Body-Size Regulation: Combining Genetics and Physiology. *Curr. Biol.* **15**, R825–R827.

Steele, L., Sukhanova, M. J., Xu, J., Gordon, G. M., Huang, Y., Yu, L. and Du, W. (2009). Retinoblastoma family protein promotes normal R8-photoreceptor differentiation in the absence of rhinoceros by inhibiting dE2F1 activity. *Dev. Biol.* **335**, 228–36.

Ullah, Z., Buckley, M. S., Arnosti, D. N. and Henry, R. W. (2007). Retinoblastoma Protein Regulation by the COP9 Signalosome. *Mol. Biol. Cell* **18**, 1179–1186.

Weinkove, D. and Leevers, S. J. (2000). The genetic control of organ growth: insights from Drosophila. *Curr. Opin. Genet. Dev.* **10**, 75–80.

Weinkove, D., Neufeld, T. P., Twardzik, T., Waterfield, M. D. and Leevers, S. J. (1999). Regulation of imaginal disc cell size, cell number and organ size by Drosophila class I(A) phosphoinositide 3-kinase and its adaptor. *Curr. Biol.* **9**, 1019–29.

Yamamoto, H., Soh, J. W., Monden, T., Klein, M. G., Zhang, L. M., Shirin, H., Arber, N., Tomita, N., Schieren, I., Stein, C. A., et al. (1999). Paradoxical increase in retinoblastoma protein in colorectal carcinomas may protect cells from apoptosis. *Clin. Cancer Res.* **5**, 1805–15.

Ying, H. and Xiao, Z.-X. J. (2006). Targeting Retinoblastoma Protein for Degradation by Proteasomes. *Cell Cycle* **5**, 506–508.

Zhang, L., Wei, Y., Pushel, I., Heinze, K., Elenbaas, J., Henry, R. W. and Arnosti, D. N. (2014). Integrated stability and activity control of the Drosophila Rbf1 retinoblastoma protein. *J. Biol. Chem.* **289**, 24863–73.

Ziebold, U., Reza, T., Caron, A. and Lees, J. A. (2001). E2F3 contributes both to the inappropriate proliferation and to the apoptosis arising in Rb mutant embryos. *Genes Dev.* **15**, 386–91.

CHAPTER 3

Diversification of retinoblastoma protein function associated with cis and trans adaptations

Abstract

Retinoblastoma proteins are eukaryotic transcriptional co-repressors that play central roles in cell cycle control, among other functions. Although most metazoan genomes encode a single retinoblastoma protein, gene duplications have occurred at least twice: in the vertebrate lineage, leading to *Rb*, *p107*, and *p130*, and in Drosophila, an ancestral *Rbf1* gene and a derived *Rbf2* gene. Structurally, Rbf1 resembles p107 and p130, and mutation of the gene is lethal. Rbf2 is more divergent and mutation does not lead to lethality. However, the retention of *Rbf2* over 60 million years in Drosophila points to essential functions, which prior cell-based assays have been unable to elucidate. Here, using genomic approaches, we provide new insights on the function of Rbf2. Strikingly, we show that Rbf2 regulates a set of cell growth related genes and can antagonize Rbf1 on specific genes. These unique properties have important implications for the fly; *Rbf2* mutants show reduced egg laying, and lifespan is reduced in females and males. Structural alterations in conserved regions of *Rbf2* gene suggest that it was sub- or neofunctionalized to develop specific regulatory specificity and activity. We define cis regulatory features of Rbf2 target genes that allow preferential repression by this protein, indicating that it is not a weaker version of *Rbf1* as previously thought. The specialization of retinoblastoma function in Drosophila may reflect a parallel evolution found in vertebrates, and raises the possibility that cell growth control is equally important to cell cycle function for this conserved family of transcriptional corepressors.

This work was submitted as the following manuscript:

Mouawad, R., Prasad, J., Thorley, D., Himadewi, P., Kadiyala, D., Wilson, N., Kapranov, P., Arnosti, DN. (2019). Diversification of retinoblastoma protein function associated with cis and trans adaptations. (*MBE*, in press).

My contribution to this study was showing differential gene regulation by Rbf1 and Rbf2 using RNA-seq analysis and bioinformatics analysis. I also performed reporter assays to show impact of promoter elements on regulation by Rbf1 and Rbf2. I also showed that Rbf2 null ovaries show upregulation of an important signaling gene that may be responsible for the phenotypes noted in the paper.

Introduction

Retinoblastoma proteins are highly conserved transcriptional co-repressors known to be major regulators of cell cycle, differentiation and apoptosis (Burkhart and Sage, 2008). These proteins do not have DNA binding domains but instead have "pocket" domains with which they bind to transcription factors. The well-characterized regulation of cell cycle genes involves the binding and inhibition of E2f/DP1 family transcription factors, and subsequent downregulation of their target genes, a pivotal role conserved in virtually all multicellular organisms.

The mammalian retinoblastoma family includes three paralogs: Rb, p107 and p130 have overlapping and distinct functions in gene regulation. In humans, germline mutations in *RB1*, the gene for Rb, cause retinoblastomas, and numerous cancers involve somatic mutations in *RB1* or associated pathway genes. Mutations in genes encoding p130 and p107 are less common in tumors, but in an *RB1* mutant background, they modify disease outcomes (Wirt and Sage, 2010; Henley and Dick, 2012). At least eight E2f transcription factors are found in humans and classified as activators (E2f1-3) and repressors (E2f4-8). Rb interacts with E2f1-5, p107 preferentially interacts with E2f4, and p130 with E2f4 and E2f5. The specific interactions of Rb with the activator E2fs may contribute to its distinct cellular functions. Genetic and molecular studies have uncovered specific activities of Rb family proteins in different tissues and cell types, including a role for Rb in senescence (Chicas et al., 2010), and p130 in quiescence (Henley and Dick, 2012), but is it is not fully understood how cellular functions are distributed among the Rb members. Furthermore, the cis regulatory information that leads to preferential association of specific E2f factors and Rb family members is poorly understood.

The presence of three retinoblastoma paralogs in vertebrates is a derived feature, since most metazoans rely on a single retinoblastoma protein to perform cellular functions. The expansion of the retinoblastoma family in vertebrates suggests that the genes may have undergone subfunctionalization and/or neofunctionalization. From a structural point of view, Rb itself is the most derived paralog, as it possesses structural aspects that differ from p107 and p130, which are more similar to an inferred ancestral gene (Wirt and Sage, 2010). The distinct functions acquired by Rb may involve gaining new gene targets related to new functional roles in regulation of apoptosis and differentiation. Interestingly, unlike the gene duplications that impact many other families of transcription factors, retinoblastoma genes tend not to be duplicated in metazoan lineages, with the exception of Drosophila, where a gene duplication ca. 60 million years ago resulted in the expression of two retinoblastoma proteins, Rbf1 and Rbf2, which are found in all characterized genomes of this genus. Thus, Drosophila provides a natural system in which to consider the impact of gene duplication in this important family.

Rbf1 and Rbf2 proteins have similar but not identical expression patterns in early embryogenesis, but in adults, Rbf1 is widely expressed, whereas Rbf2 is expressed mainly in the ovaries (Stevaux et al., 2002; Keller et al., 2005). Simpler than the vertebrate system, there are two E2f factors in Drosophila, E2f1 which is an activator and E2f2 which is classified as a repressor. Previous work by Dyson and colleagues suggested that Rbf1 interacts with both E2f factors, whereas Rbf2 interacts mainly with E2f2 (Stevaux et al., 2002). These studies showed that when assayed on cell cycle promoters, Rbf2 is a weaker repressor than Rbf1, and few genes are derepressed upon depletion of Rbf2 in cultured S2 cells (Stevaux et al., 2002; Dimova et al., 2003). *Rbf2* null flies do not have a lethal phenotype, unlike the case for *Rbf1* null alleles (Stevaux et al., 2005). The conservation of the *Rbf2* gene thus poses a conundrum. Here we explore the activities of Rbf2 and

Rbf1 in the context of the intact animal, and show that Rbf2 appears to regulate a large set of genes related to growth control, using unique cis regulatory signals important for specificity. New null alleles of *Rbf2* reveal an important role for Rbf2 in the development and physiological regulation of the ovary. These functions of derived retinoblastoma family members may reflect similar molecular processes that apply to vertebrate paralogs, with application in development and disease.

Results

Rbf2 shows higher divergence than *Rbf1* from ancestral lineage, impacting important functional portions of protein sequence.

The *Rbf1* and *Rbf2* genes were originally identified by their similarities to mammalian retinoblastoma family genes, including a segment encoding the "pocket" domain critical for interactions with E2f/DP1 (Du et al., 1996; Stevaux et al., 2002). We used multiple sequence alignments to understand conservation of specific segments of these genes within the Drosophila lineage, as well as their relative conservation with other metazoan retinoblastoma genes.

To facilitate our analysis, we divided the protein-coding sequences into three segments: the E2f-binding “pocket” domain (including A and B subdomains), all sequences N-terminal to the pocket, and all sequences C-terminal to the pocket, which include the so-called Instability Element (IE) that is important for stability and activity of pocket proteins (Acharya et al., 2010; Raj et al., 2012; Zhang et al., 2014; Sengupta et al., 2015). Considering diverse Drosophila Rbf1 protein sequences, central cyclin fold and pocket domains are more conserved than extreme N and C terminus regions (Figure 3-1A). Furthermore, the level of conservation for all three domains closely mirrors the overall phylogenetic distances, suggesting that gradual changes in *Rbf1* genes may represent neutral or compensated alterations in the protein (Figure 3-2, 3-3, 3-4, 3-5). Rbf2 protein sequences are more divergent overall than Rbf1, especially in sequences of the C-terminus and in the spacer region between the A and B pockets, as well in as the N-terminus (Figure 3-1B, 3-6, 3-7, 3-8, 3-9). Unlike the sequence alignments for Rbf1, Rbf2 sequences can be separated into two clusters; protein sequences from the *melanogaster* subgroup are overall much more similar to each other than those from more distantly related species (*D. ananassae* and others) (Figure 3-6). The Rbf2

sequences from these more divergent lineages exhibit lower conservation than that observed for Rbf1, meaning that Rbf2 sequences are quite malleable in all Drosophila lineages. Considering the *D. melanogaster* sequences, the regions most conserved in Rbf1 and Rbf2 are the E2f/DP1 binding pocket domains (Figure 3-1C).

The pocket regions of retinoblastoma proteins are in general the most conserved; within Drosophila, Rbf1 regions A and B show higher conservation among themselves than do the comparable regions of Rbf2 (Figure 3-2, 3-6). Also impacted is the spacer region between the A and B domains: in mammalian p107 and p130 proteins, the spacer regions have unique cyclin/cdk binding and inhibition activity that is absent from Rb, suggesting that changes in this region have functional consequences (Wirt and Sage, 2010). In the Drosophila counterparts, the spacer between the A and B pocket domains is well conserved among Drosophila Rbf1 homologs, with a constant length of 19 amino acids (Figure 3-1A, 3-4). Rbf2 proteins, in contrast, feature spacer sequences of different lengths and more sequence diversity (Figure 3-1B, 3-8).

In the Rbf1 C-terminus, the IE is the most conserved region, which is consistent with our previous studies that this degron is critical for turnover and function (and is also conserved in p107 and p130). Serines 728, 760, and 771, which represent serine-proline phosphorylation sites (SP) have been shown to mediate regulation by phosphorylation in *D. melanogaster* Rbf1 (Zhang et al., 2014). Lysine 774, which is conserved in p107 and p130, plays an important regulatory role, and is known to be a target of acetylation in the mammalian system (Saeed et al., 2012). These residues are highly conserved in all Rbf1 sequences (Figure 3-5). In addition, conserved blocks of hydrophobic residues within this IE element found in p107, p130 and Rbf1 are absent in Rbf2. Although there are residues of the Rbf2 C-terminus that align with Rbf1 sequences, Rbf2 proteins

appear to lack a canonical IE, and only one of the three conserved SP phosphorylation sites found in Rbf1 can be identified (Figure 3-1D).

Retinoblastoma family proteins contain a cyclin-fold homology domain within the N-terminus. This region is conserved to a similar level in Rbf1 and Rbf2, with more divergence in Rbf2 sequences in the region between the cyclin fold and the pocket (Figure 3-1A, B). Threonine residue 356 of *D. melanogaster* Rbf1 has been shown to be important for regulation of the protein by phosphorylation, similar to Rb (Burke et al., 2012); this residue is absolutely conserved in Rbf1 proteins; this residue is not conserved in Rbf2 (Figure 3-3, 3-7).

The sequence variations for the two gene Drosophila Rbf family may represent a functional interplay between these genes, reflecting neofunctionalization or subfunctionalization. To understand how divergence of Rbf protein sequences in the Drosophila lineage compares with that observed in related arthropod lineages possessing a single *Rbf* gene, we aligned sequences of diverse insect orders, as well as more distantly related chelicerate and crustacean proteins (Figure 3-10, 3-11, 3-12). Conserved features noted in Rbf1 are a general feature of homologous proteins; in these genomes, we see a conservation of the N-terminal cyclin fold, the C-terminal IE element and the phosphorylation sites discussed above. Interestingly, the pocket A-B spacer sequence and length is not highly conserved. The diversification in sequence found in Rbf2 proteins in these generally conserved domains points to relaxed constraints on protein structure, perhaps underlying new cellular roles for this protein. We hypothesize that Rbf2 may have diverged faster than Rbf1 because of specialized roles assumed in the physiology and reproduction of different Drosophila species, as indicated by our genetic studies. In this view, Rbf1 may be responsible for regulation of conserved, general functions that are not subject to marked variation across Drosophila species.

In order to understand how conservation patterns observed for the Drosophila Rbf1 and Rbf2 proteins compare to the other lineage in which this gene family shows duplications, we aligned *D. melanogaster* Rbf1 and Rbf2 with the human Rb, p107, and p130 protein sequences. As previously observed, Rbf1 sequences are more similar to those of p107 and p130 than to those of Rb (Figure 3-1D). In the C-terminal IE region, specific residues shown to be critical for the selectivity of p107 for E2f4 (Liban et al., 2017) are conserved in Rbf1, but not Rbf2. Interestingly, the human Rb protein is more divergent in this area, suggesting that changes in the IE may be a common mechanism for divergent function of duplicated retinoblastoma protein family members (Rubin et al. 2005).

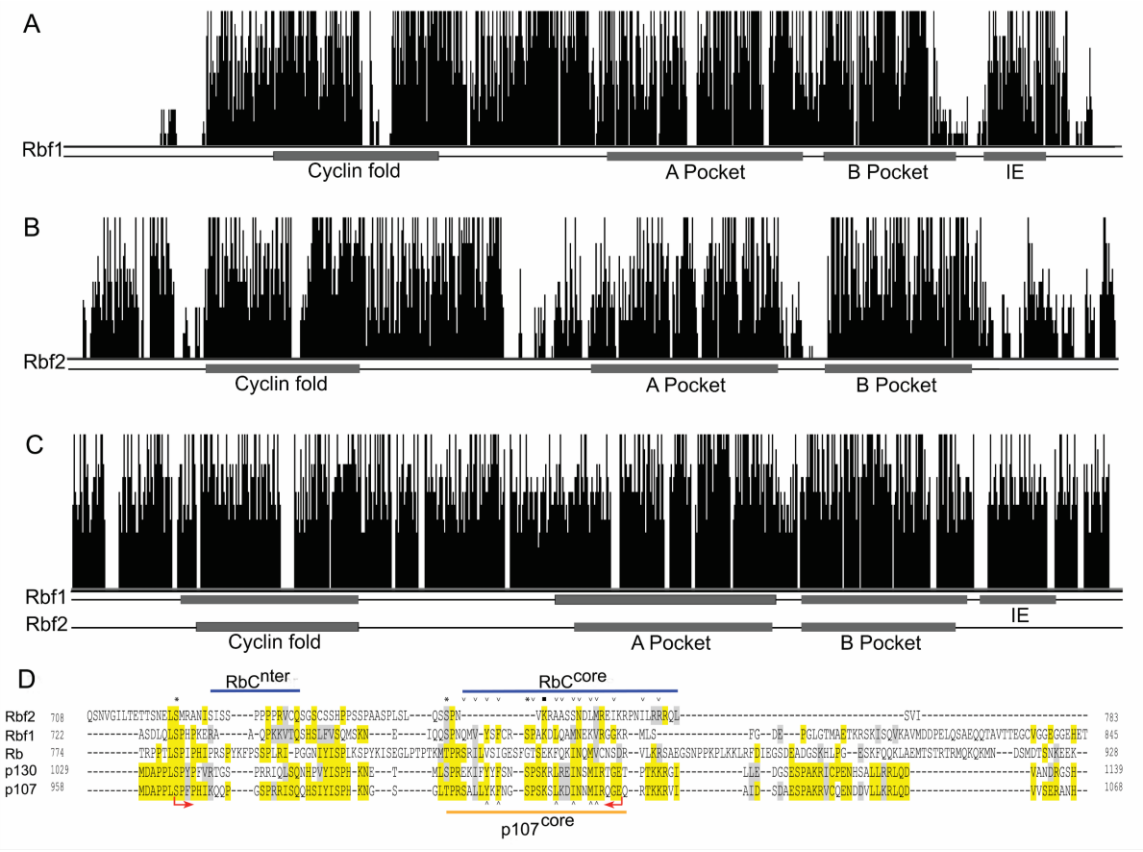


Figure 3-1: Sequence conservation of retinoblastoma proteins in Drosophila and humans. (A) Multiple sequence alignment of Rbf1 in 12 Drosophila species; conservation is observed in C-

Figure 3-1 (cont'd)

terminal IE region and A-B pocket and spacer, as well as N terminal regions. (B) Multiple sequence alignment of Rbf2 in 12 Drosophila species, showing lower conservation in C-terminus and in A-B pocket and spacer, as well as N terminal regions. (C) Pairwise alignment of Drosophila Rbf1 and Rbf2, showing higher conservation in central pocket domains, and lower in C-terminus. For A, B and C, the y-axis represents alignment scores generated by ClustalW, which takes into consideration both identity and chemical similarity (see Methods). Higher bars indicate more conservation. The functional domains are denoted including the cyclin fold domain, A pocket, B pocket and the instability element (IE) in the C-terminus. (D) Multiple sequence alignment of C-termini of Drosophila and mammalian retinoblastoma proteins. The yellow color represents conserved residues and grey represents similar residues with respect to p107. Specific portions of the C-terminus involved in direct contacts with E2F/DP1 proteins are highlighted; the RbC^{nter} and RbC^{core} are shown on top of the figure, and the $p107^{core}$ is shown at the bottom (Rubin et al., 2005; Liban et al., 2017). The IE, from residue 728 to 786 of Rbf1, is denoted by the red arrows. Triangles represent residues that make contacts with E2F/DP marked box domains for both Rb and p107. The asterisk denotes conserved serine residues within SP motifs that are targeted for phosphorylation. The K774 residue within the SPAK motif is denoted by a square.

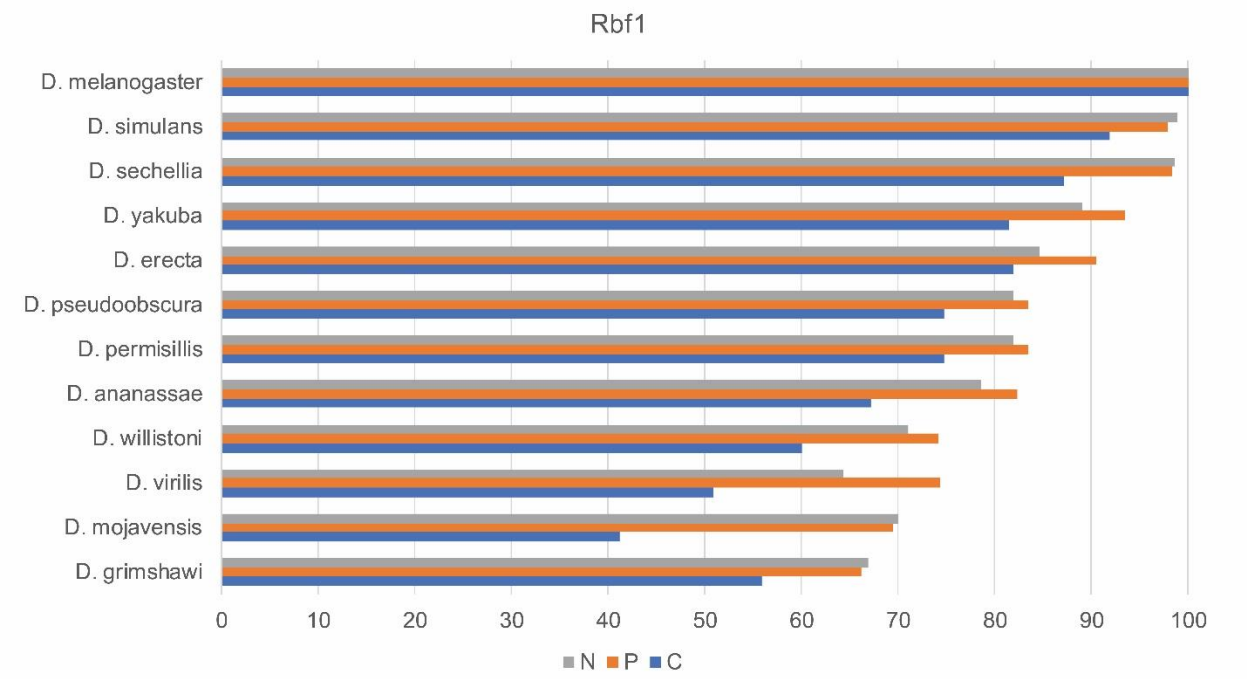


Figure 3-2: Bar graph representing percent identity values from multiple sequence alignments of Rbf1 from Drosophila species against *D. melanogaster* for the N-terminus, pocket and C-terminus. Percent identities calculated from multiple sequence alignments performed using Clustal Omega.

Figure 3-3: Multiple sequence alignments for Rbf1 N-terminus within Drosophila species. Yellow shade represents conserved residues and grey represents similar residues to *D. melanogaster*. The cyclin fold domain is demarcated by a black line. The (*) denotes conserved Thr356 residue.

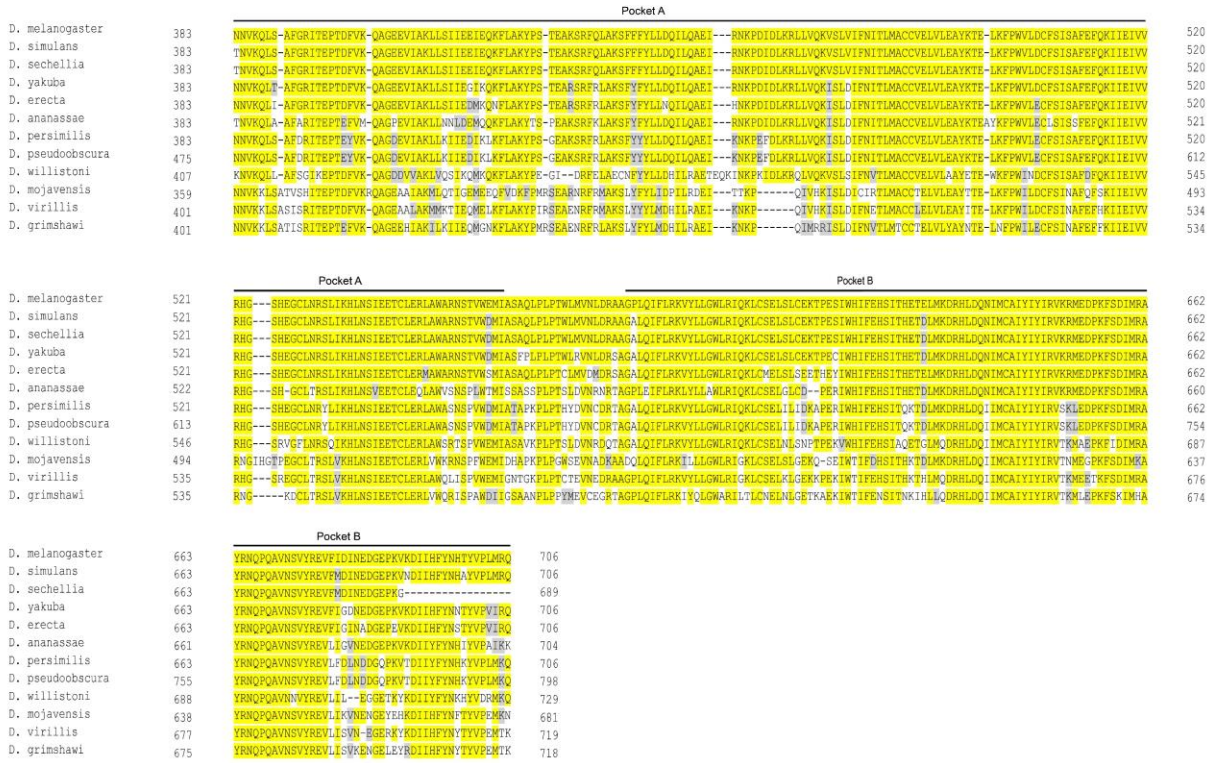


Figure 3-4: Multiple sequence alignment for Rbf1 pocket domain within Drosophila species. Yellow shade represents conserved residues and grey represents similar residues to *D. melanogaster*. A and B pocket subdomains are shown flanking the 19-residue spacer region.

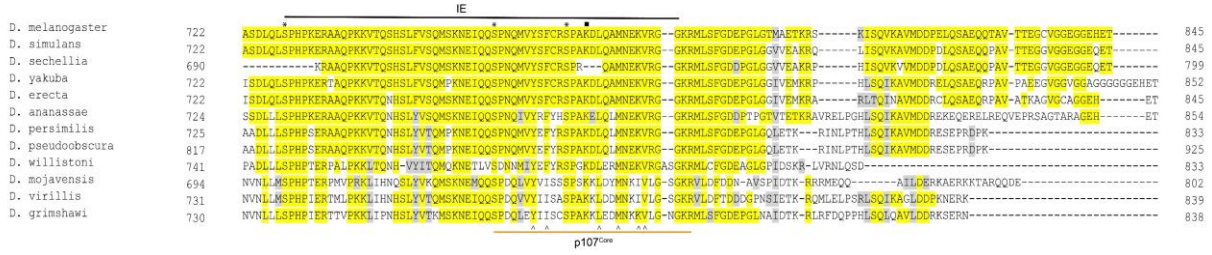


Figure 3-5: Multiple sequence alignments for Rbf1 C-terminus within Drosophila species. Yellow shade represents conserved residues and grey represents similar residues to *D. melanogaster*. The instability element (IE) region is demarcated with a black line. The p107^{core}

region is demarcated with a red line. Triangles indicate residues that interact with E2F/DP1 marked box domain. The asterix denotes conserved serine residues that are targeted for phosphorylation. The K774 residue within the SPAK motif is denoted by a square.

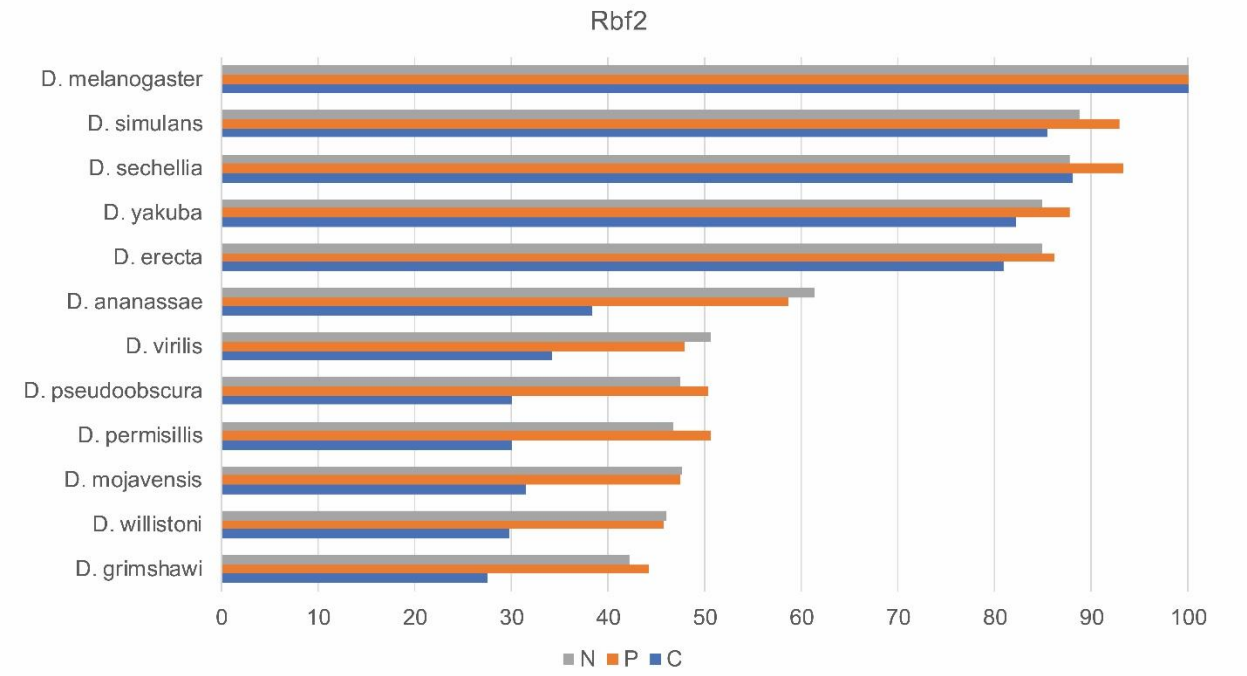


Figure 3-6: Bar graph representing percent identity values from multiple sequence alignment of Rbf2 from Drosophila species against *D. melanogaster* for the N-terminus, pocket and C-terminus. Percent identities calculated from multiple sequence alignments performed using Clustal Omega.

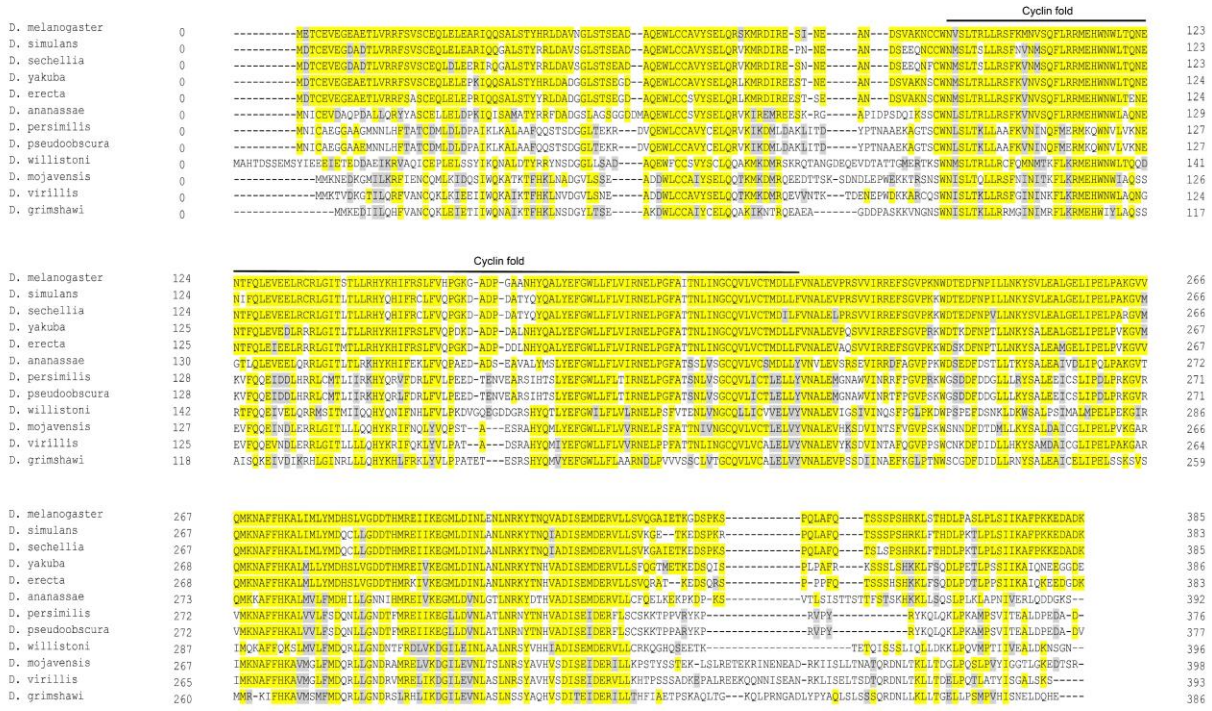


Figure 3-7: Multiple sequence alignment for Rbf2 N-terminus within Drosophila species. Yellow shade represents conserved residues and grey represents similar residues to *D. melanogaster*. The cyclin fold region is demarcated with a black line.

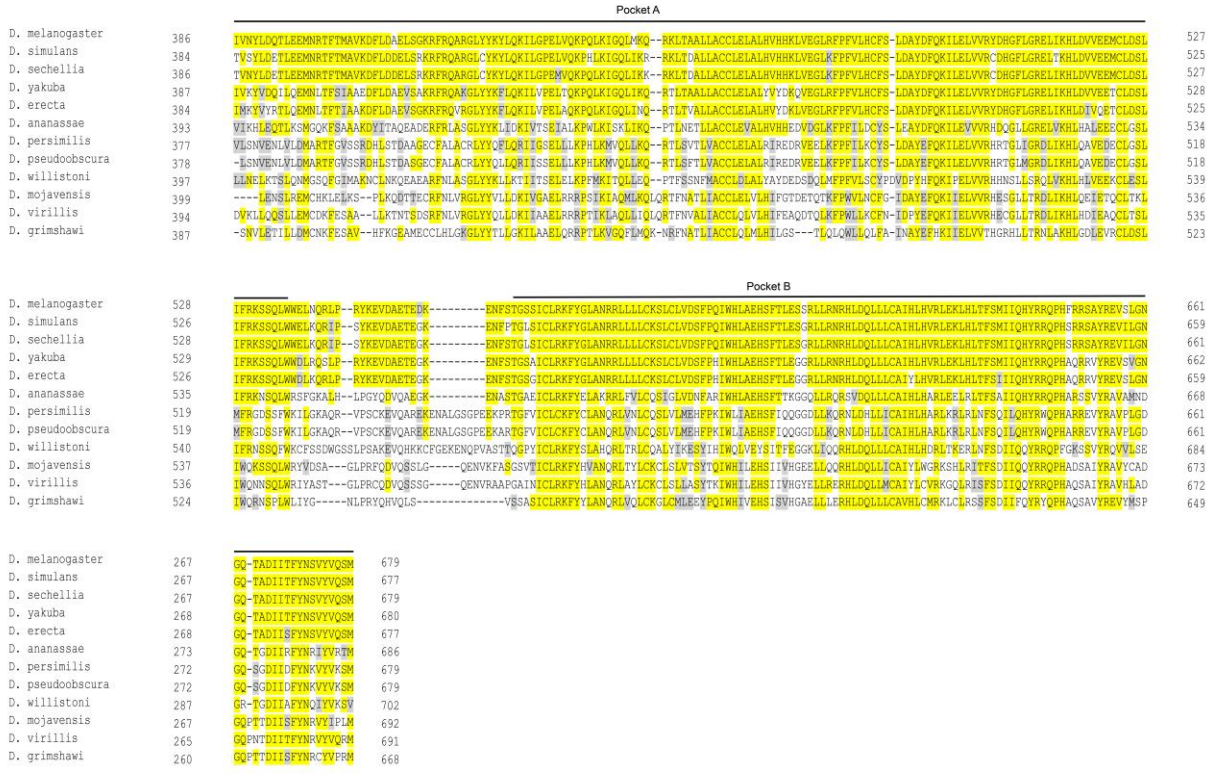


Figure 3-8: Multiple sequence alignment for Rbf2 pocket domain within Drosophila species. Yellow shade represents conserved residues and grey represents similar residues to *D. melanogaster*. A and B pocket subdomains are shown.

Figure 3-9: Multiple sequence alignment for Rbf2 C-terminus within Drosophila species. Yellow shade represents conserved residues and grey represents similar residues to *D. melanogaster*.

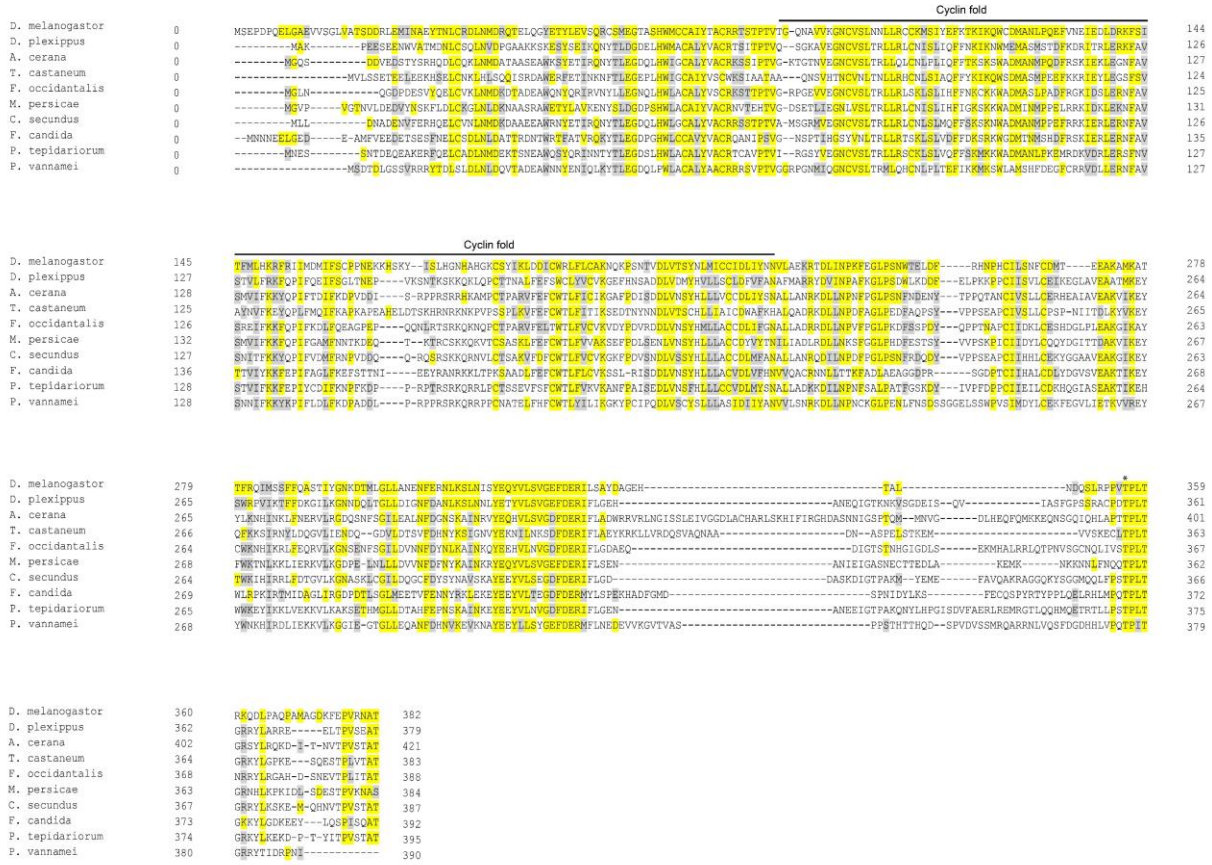


Figure 3-10: Multiple sequence alignment of N-terminus of Rbf1 from *D. melanogaster* and other arthropods. The following arthropod species are: *Danaus plexippus* (Lepidoptera), *Apis cerena* (Lepidoptera), *Tribolium castaneum* (Coleoptera), *Frankliniella occidentalis* (Thysanpotera), *Myzus persicae* (Hemiptera), *Cryptotermes secundus* (Isoptera), *Folsomia candida* (Hymenoptera), *Parasteatoda tepidariorum* (Spider), and *Penaeus vannamei* (White-legged Shrimp).The yellow shade represents conserved residues and grey represents similar residues. The cyclin fold of *D. melanogaster* is demarcated with a black line. The (*) denotes conserved Thr356 residue.

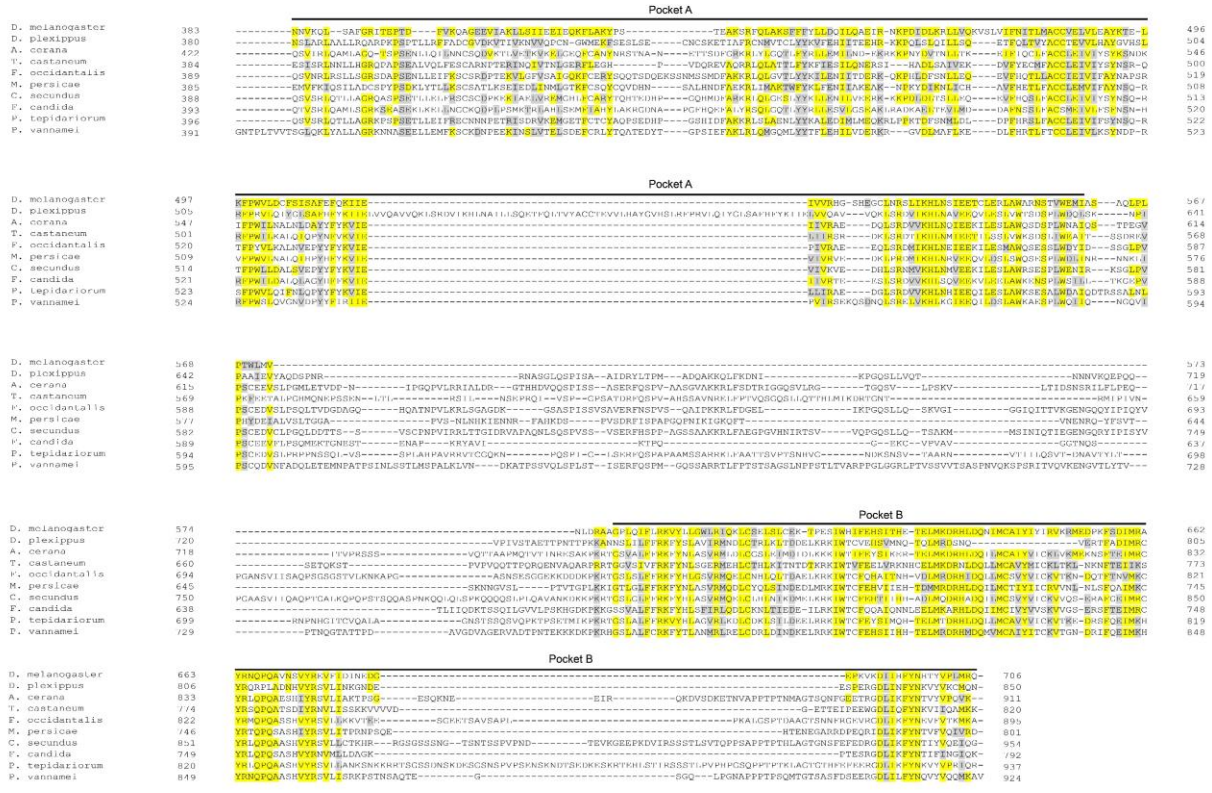


Figure 3-11: Multiple sequence alignment of Rbf1 pocket from *D. melanogaster* and other arthropods. The following arthropod species are: *Danaus plexippus* (Lepidoptera), *Apis cerena* (Lepidoptera), *Tribolium castaneum* (Coleoptera), *Frankliniella occidentalis* (Thysanpotera), *Myzus persicae* (Hemiptera), *Cryptotermes secundus* (Isoptera), *Folsomia candida* (Hymenoptera), *Parasteatoda tepidariorum* (Spider), and *Penaeus vannamei* (White-legged Shrimp). The yellow shade represents conserved residues and grey represents similar residues. The A and B pocket subdomains are indicated.

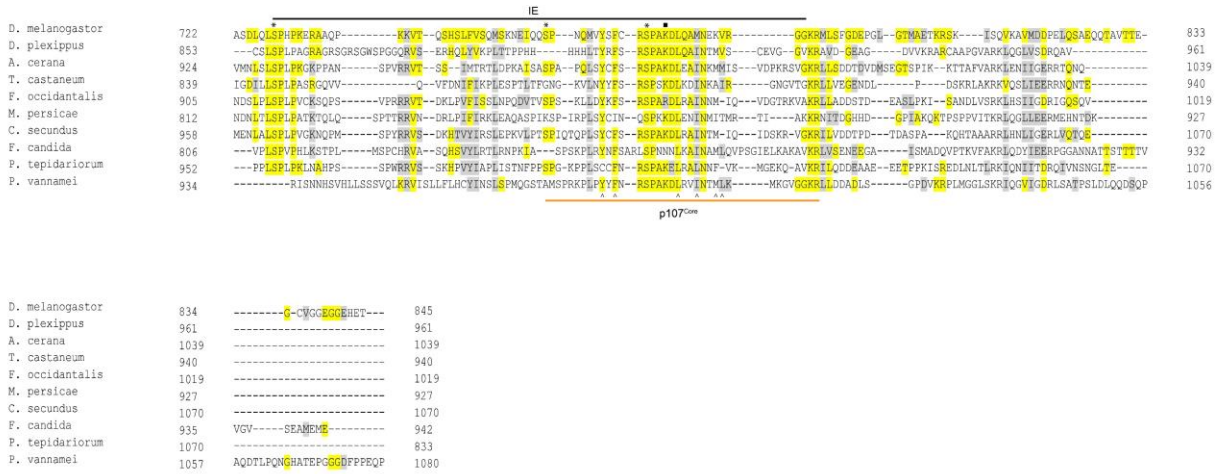


Figure 3-12: Multiple sequence alignment of C-terminus of Rbf1 from *D. melanogaster* and other arthropods. The following arthropod species are: *Danaus plexippus* (Lepidoptera), *Apis cerena* (Lepidoptera), *Tribolium castaneum* (Coleoptera), *Frankliniella occidentalis* (Thysanpotera), *Myzus persicae* (Hemiptera), *Cryptotermes secundus* (Isoptera), *Folsomia candida* (Hymenoptera), *Parasteatoda tepidariorum* (Spider), and *Penaeus vannamei* (White-legged Shrimp). The instability element (IE) region is demarcated with a black line. The p107^{core} region is demarcated with a red line. Triangles indicate residues that interact with E2F/DP1 marked box domain. (*) denotes conserved serine residues that are targeted for phosphorylation. The K774 residue within the SPAK motif is denoted by a square.

In vivo regulation of embryonic genes by Rbf1 and Rbf2.

In previous studies, we mapped in vivo binding profiles of Rbf1 and Rbf2 in the embryo. Rbf2 is found at the promoters of approximately 4,000 genes, while Rbf1 is found at about half that number, in a largely overlapping pattern. The targets of Rbf1 and Rbf2 include ribosomal, cell cycle and signaling genes, however it is not known whether these binding events represent direct regulation in most cases (Acharya et al., 2012, Wei et al., 2015). To determine the effects of Rbf1

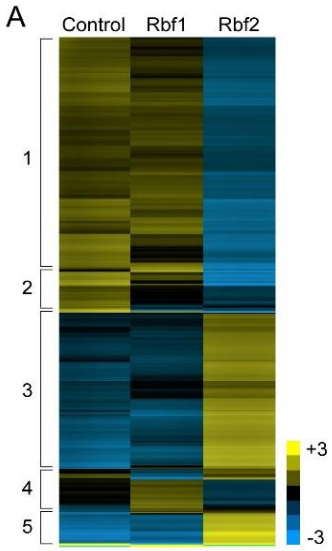
and Rbf2 on gene regulation, we induced the expression of each protein using transgenes under the control of a heat shock promoter and performed RNA-seq analysis on 12-18 hour embryos.

After a brief induction of either the Rbf1 or Rbf2 protein by heatshock, RNA was isolated from embryos after 60 minutes, and RNA-seq libraries were prepared for treated or control (heatshock induction with no Rbf transgene) embryos. We filtered the RNA-seq data to focus on genes directly bound by Rbf1 or Rbf2 based on our previous Chip-seq analysis, and removed genes that had low expression levels in all of the samples. We performed unsupervised clustering on the remaining 3937 genes and analyzed five major clusters, with distinct patterns of gene expression across the samples (Figure 3-13A). Strikingly, Rbf2, which had been characterized as a weak repressor on certain promoters, showed a robust effect on gene expression. All genes in cluster 1 are repressed by Rbf2, with some also exhibiting a weaker repression by Rbf1. On the other hand, Clusters 3 and 5 show a significant upregulation of genes by Rbf2, whereas Rbf1-mediated changes are fewer in number, and weaker in these clusters (Figure 3- 13A, Table 3-1). A number of cell-cycle genes that are repressed by Rbf1 expression in our dataset were found to upregulated in *Rbf1* knock-down cells and *Rbf1* mutant flies (Dimova et al., 2003; Longworth et al., 2012), confirming the physiological relevance of the system used here. The dramatically different effects of Rbf1 and Rbf2 expression point to different functions in gene regulation and cellular processes.

To determine the nature of the genes within each cluster, we performed gene ontology analysis using the DAVID annotation tool. Strikingly, among the most enriched categories of Cluster 1 are ribosomal protein and mitochondrial genes, suggesting that Rbf2 may have an important role in control of genes closely linked to cellular growth control (Figure 3-13B). For specific functional classes of genes, a significant fraction was regulated by Rbf2. For instance, of 93 ribosomal protein genes that are direct targets of Rbf1 or Rbf2, 52 genes show at least 10 % repression by Rbf2, and

15 genes by Rbf1. Out of 80 mitochondrial genes that are direct targets, 70 are repressed at least by 10 % by Rbf2 and 23 genes by Rbf1 (Figure 3-13C). Thus, Rbf2 appears to play a dominant role in regulation of these cell growth-related genes, with Rbf1 playing a secondary role. In the cases where we observe activation by Rbf2, the most enriched categories in Cluster 3 include splicing and transcription regulation, while the Cluster 5 top enriched category is cell cycle including *CycB*, *MCM7* and others. In agreement with previous data, *PCNA* is repressed by Rbf1 but not impacted by Rbf2, and it is present in cluster 3. The positive action of Rbf2 overexpression may represent antagonistic action against Rbf1; notably, Cluster 3 and 5 genes have a somewhat higher fraction of promoters co-bound by both Rbf1 and Rbf2 (Table 3-2).

We considered whether the activation or repression by Rbf2 may relate to the inherent expression levels of targeted genes. Indeed, the majority of genes in Cluster 1 (repressed by Rbf2) were in the top 50% of expression, whereas half of Cluster 5 genes (strongly activated by Rbf2) were in the lowest quartile of expression (Figure 3-14). Under normal circumstances, the targets in Cluster 5 may be kept inactive by endogenous Rbf1, and competition by Rbf2 upon overexpression may cause them to be derepressed, if Rbf2 is less effective as a repressor. Overall, the functional comparison of Rbf1 and Rbf2 activity in the embryo points to a previously unappreciated role for Rbf2 to regulate a pervasive and functionally distinct set of genes linked to growth regulation.



B

Cluster	GO Term	P-value	Count	Total
1	Mitochondrion	7.73E-08	52	263
	Membrane	6.05E-06	356	3266
	Ribosomal protein	2.90E-05	33	168
	Potassium ion transport	4.75E-04	10	27
	Nucleotide-binding	8.67E-04	88	701
2	Oxidative Phosphorylation	0.014	7	139
3	Zinc finger, Transcription regulation	1.02E-13	67	384
	mRNA splicing	3.30E-11	46	226
	ATP binding	3.73E-08	71	558
	Histone acetylation	7.81E-06	13	55
	Protein Phosphorylation	1.08E-04	36	266
4	Zinc finger	0.003	10	225
	Transcription regulation	0.008709	10	275
	Tyrosine Kinase	0.010833	3	16
5	Cell cycle	2.40E-21	24	129
	Kinetochore	1.24E-09	10	40
	P-granule	7.21E-05	6	35
	Oogenesis	1.04E-04	7	66
	ATP-binding	1.68E-04	23	751

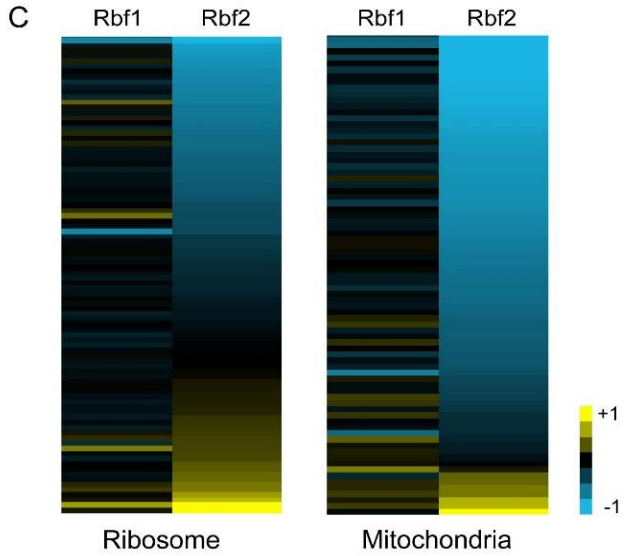


Figure 3-13: Overexpression of Rbf2 results in profound effects on gene expression in embryos. (A) A heatmap generated by unsupervised clustering of RNA-seq data from Rbf1 and Rbf2 overexpressing embryos, and control embryos. Values represent log transformed RPKM reads that are mean centered for each gene. Blue indicates reads below the mean, black equal to the mean, and yellow, above the mean. Values represent average of three biologic replicates. RPKM < 1 were excluded from the analysis. Only genes bound by Rbf1 or Rbf2 in vivo are

Figure 3-13 (cont'd)

included. The heatmap is divided into 5 major clusters based on Euclidean distance. (B) Gene ontology analysis of the five clusters based on the DAVID annotation tool. P-values represent significance of enrichment for each category, the count represents the number of genes in the cluster belonging to each category, and total shows number of genes in the GO category. (C) Relative gene expression of ribosomal and mitochondrial related genes in Rbf1 or Rbf2 overexpressing embryos, relative to control embryos. Values represent average of three biologic replicates.

Table 3-1: RNA-seq analysis results showing gene expression changes after induction of Rbf1 or Rbf2 in embryo.

Rbf1	Rbf2	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
down	down	63	131	0	0	0
down	no	0	5	14	0	0
down	up	0	0	38	0	2
up	down	104	0	0	18	0
up	no	3	0	0	122	0
up	up	0	0	300	48	94
no	down	1058	87	0	0	0
no	no	64	0	0	0	0
no	up	0	0	560	0	69
Total number of genes		1292	223	912	188	165

The changes represent relative expression in comparison to control embryos. Changes in gene expression of more than 20% are counted as up or down, otherwise no change.

Table 3-2: Ratio of genes that are shown to be occupied by Rbf1 or Rbf2 or both in a previous Chip-seq dataset (Wei et al., 2015).

Cluster	Rbf2	Rbf1	Rbf2 and Rbf1	Only Rbf2	Only Rbf1
1	0.87	0.44	0.31	0.55	0.12
2	0.78	0.53	0.31	0.46	0.20
3	0.93	0.44	0.37	0.55	0.06
4	0.85	0.46	0.31	0.53	0.14
5	0.92	0.57	0.49	0.42	0.07

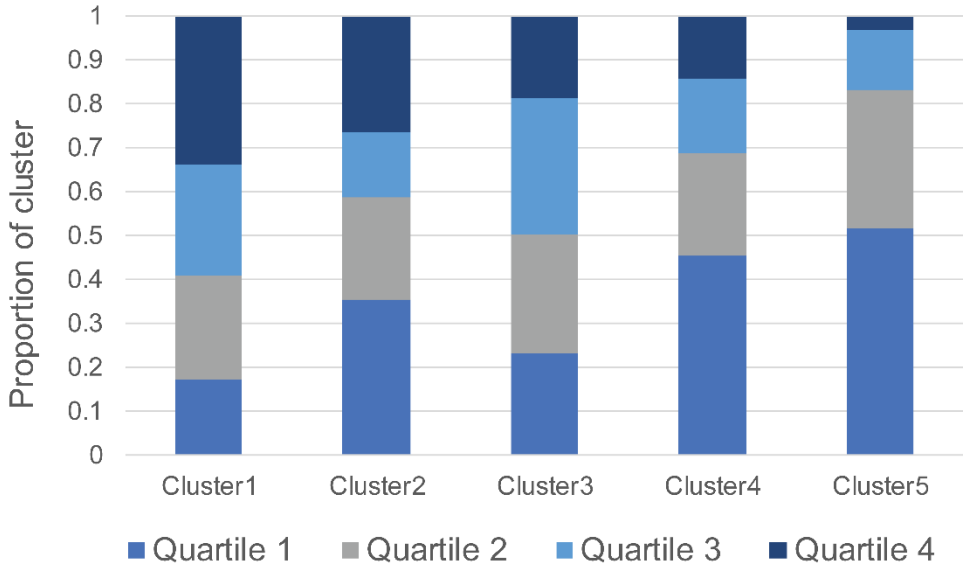


Figure 3-14. Bar graph showing expression levels of genes within each cluster of the heatmap. The expression levels were determined from the RPKM values of the genes in the control samples. The RPKM values were ranked from high to low and divided into four quartiles (Q1 lowest, Q4, highest expression).

Roles for Rbf2 in development and function of the ovary

A previous study generated an *Rbf2* null using deletion of a genomic fragment including the gene, however, this mutation also impacted the neighboring gene, *moira*, and the final rescue construct resulted in expression of a fragment of Rbf2 protein (Stevaux et al., 2005). We generated additional *Rbf2* alleles using CRISPR/Cas9, producing four frameshift alleles that truncate the protein N-terminal to the pocket domain, and two in-frame alleles removing five amino acids in two portions of the N-terminus (Figure 3-15A). Transheterozygous combinations of the presumptive null alleles yielded viable flies, consistent with previous reports for viability of the null mutant. Western blot analysis from ovaries of mutant flies verify the loss of the full length Rbf2 protein (Figure 3-15B). As shown in figure 3-15A, it is possible that small pieces of the N-terminus of Rbf2 are still expressed in mutants carrying the presumptive null alleles. $Rbf2^{\Delta 1}$ would produce a protein with a portion of the N-terminal cyclin fold domain. Levels of *Rbf2* transcripts are reduced in this background, presumably due to destabilization of the mRNA by translational defects (Figure 3-16D).

Although *Rbf2* presumptive null mutants are viable, they exhibited effects on ovarian development and function, as well as survival. Interestingly, the lifespan of both homozygous mutant males and females was significantly shorter than control *yw* flies (Figure 3-15C, D). Rbf2 expression is very low in the adult male, suggesting an earlier developmentally important role for the protein. Homozygous mutants ($Rbf2^{\Delta 1}/Rbf2^{\Delta 38}$, $Rbf2^{\Delta 1}/Rbf2^{\Delta 41}$, $Rbf2^{\Delta 1}/Rbf2^{\Delta 46}$) laid significantly fewer eggs than control *yw* flies (Table 3-3). A large fraction of the ovaries in these homozygous mutant females had an increased number of mature looking oocytes (data not shown), although the number of ovarioles per ovary is not different from control flies (Figure 3-16B, C). Our results for *Rbf2* mutant flies contrast with previous data from Dyson and colleagues, who reported an increased

egg laying for *Rbf2* null females (Stevaux et al. 2005). We considered whether this phenotype resulted from only a partial loss of Rbf2 activity, therefore we tested the egg laying rates from crosses using either male or female *Rbf2* heterozygous flies. We took care to introgress the $Rbf2^{\Delta 1}$ allele into a wild-type *yw* background to control for genetic background effects. These flies did not show a significant difference in egg laying rates, thus this gene appears to be recessive with respect to this trait (Figure 3-16A). However, the average number of ovarioles per ovary was ~30% higher in heterozygous $Rbf2^{\Delta 1}$ females when compared to control *yw* flies, and the number of ovarioles was found to be different between ovaries of the same fly (Figure 3-16B, C). However, the number of ovarioles is similar to control *yw* flies for two other heterozygous mutant alleles that are predicted to produce shorter protein products, indicating that $Rbf2^{\Delta 1}$ may have residual activity.

The $Rbf2^{\Delta 15C}/Rbf2^{\Delta 15C}$ homozygotes bearing an in-frame deletion of five amino acids in the cyclin fold motif within the N-terminus of the protein were female sterile (Table 3-3). The ovaries were very small, with distorted morphology, no discernable germarium or ovariole structures (Figure 3-16E); although the appearance of ovaries from one-day old virgins was similar to wild-type flies (data not shown). Male fertility, on the other hand, was unaffected. $Rbf2^{\Delta 1}/Rbf2^{\Delta 15C}$ transheterozygote females had no obvious defects, suggesting that if the in-frame deletion creates a neomorphic protein, there must be a dosage threshold for this phenotype to be displayed. Another possibility is that the $Rbf2^{\Delta 15C}/Rbf2^{\Delta 15C}$ homozygotes may have a different recessive mutation unlinked to *Rbf2*.

To understand the mutant phenotypes at a molecular level, we assessed expression from select target genes in ovaries of $Rbf2^{\Delta 1}/Rbf2^{\Delta 46}$ transheterozygous mutant flies. As expected, *Rp49*, *Rbf1* and *PCNA* are not affected in the *Rbf2* mutant flies. Interestingly, another Rbf2 direct target gene, *Pi3K92E* is significantly increased in the mutant flies in comparison to controls; the same

induction of *Pi3K92E* was found in the other *Rbf2* transheterozygous backgrounds (data not shown) (Figure 3-16D). Interestingly, the expression of this gene is decreased upon overexpression of Rbf2, but not Rbf1, in embryos, suggesting that this mode of regulation is the same in the two different developmental contexts. This gene encodes the catalytic subunit of class I phosphoinositol-3-kinase that is a component of the insulin signaling pathway and is directly linked to organ growth. Regulation of this gene may contribute to the phenotypes observed.

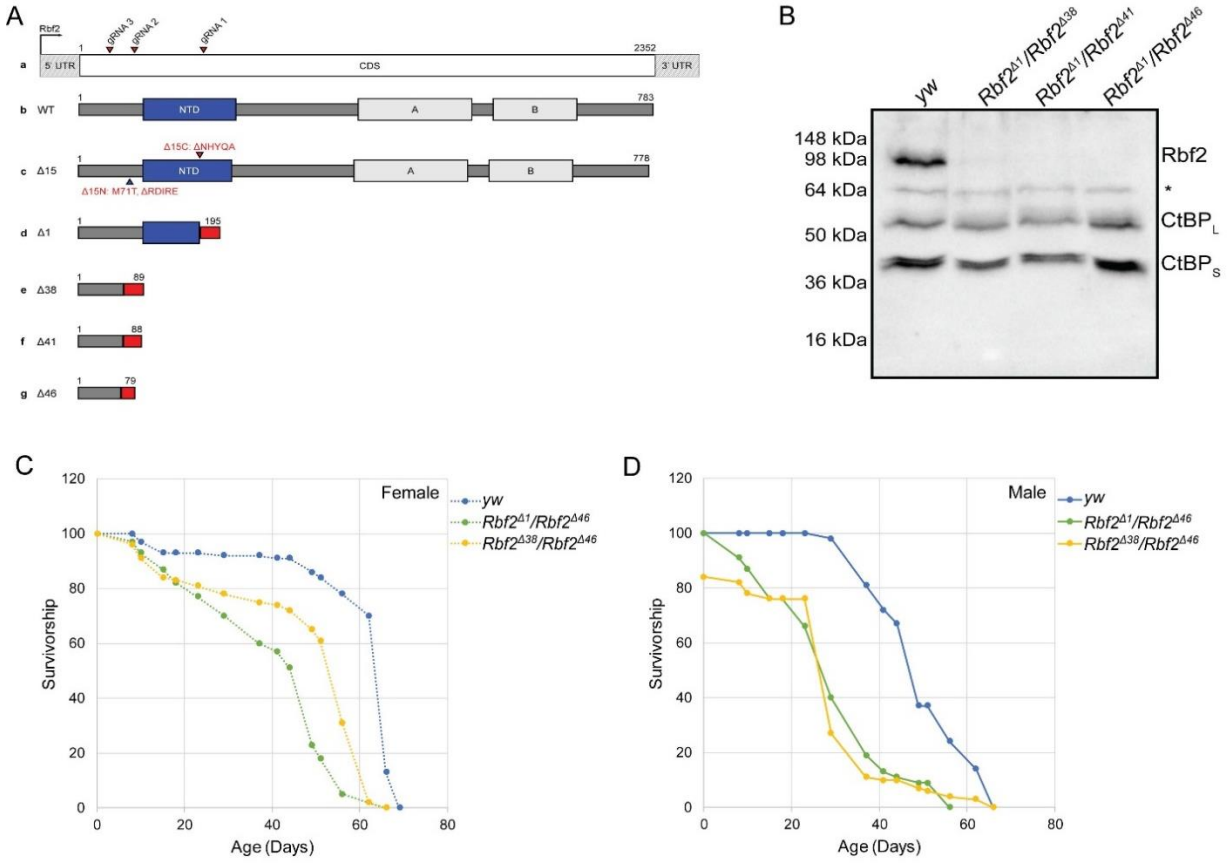


Figure 3-15: *Rbf2* mutant alleles and longevity phenotype. (A) Schematic representation of the CRISPR targeting of *Rbf2* and the alleles generated. gRNA1 produced Δ1 and Δ15C alleles, and gRNA2 produced Δ38, Δ41, Δ46 and Δ15N alleles. (B) Western blot indicating loss of Rbf2

protein from *Rbf2* mutant ovaries from $Rbf2^{\Delta 1}/Rbf2^{\Delta 38}$, $Rbf2^{\Delta 1}/Rbf2^{\Delta 41}$ and $Rbf2^{\Delta 1}/Rbf2^{\Delta 46}$ flies. Anti-CtBP is used as a control. Asterisk indicates nonspecific band. (C) Survivorship curve for $Rbf2^{\Delta 1}/Rbf2^{\Delta 46}$ and $Rbf2^{\Delta 38}/Rbf2^{\Delta 46}$ females and (D) males in comparison to *yw* flies. The curves from the mutants were significantly different from the *yw* flies curves for both females and males according to log-rank test with a p-value <0.0001.

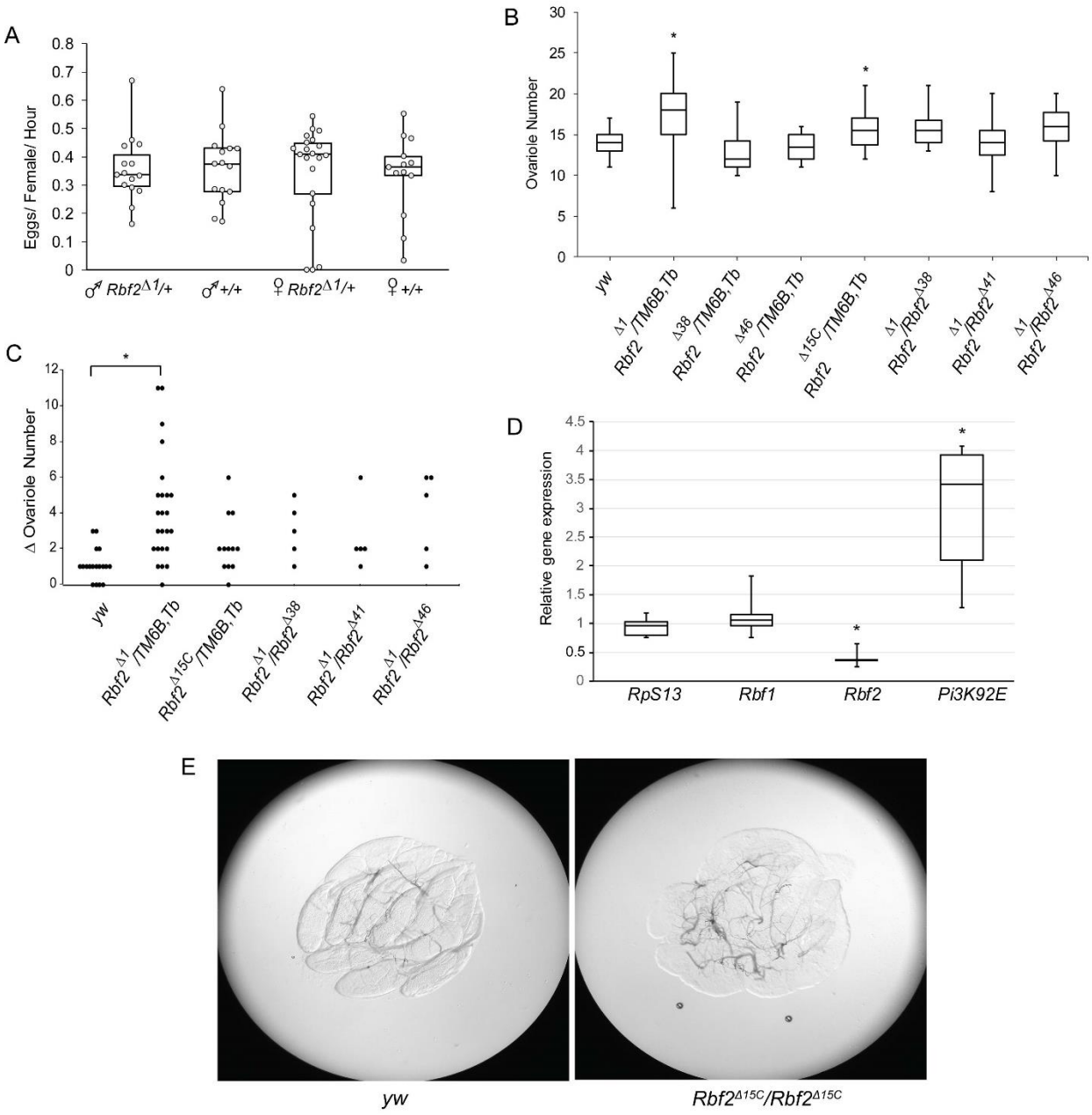


Figure 3-16: Rbf2 effects on egg laying, ovariole numbers, and *Pi3K92E* expression. (A) Egg laying for introgressed $Rbf2^{\Delta 1}$ allele for the mutant males or females. The measurements represent average 24-hour egg count for four days for each single female fly. Mutant or wild-type males were crossed to a single *yw* female, and mutant or wild-type females were crossed to three *yw* males. (B) Ovariole counts of individual adult ovaries for the following genotypes: *yw* (n = 56),

Figure 3-16 (cont'd)

Rbf2$^{\Delta 1}$*/TM6B,Tb* (n = 68), *Rbf2*$^{\Delta 38}$*/TM6B,Tb* (n = 20), *Rbf2*$^{\Delta 46}$*/TM6B,Tb* (n = 20), *Rbf2*$^{\Delta 15C}$*/TM6B,Tb* (n = 24), *Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 38}$ (n = 10), *Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 41}$ (n = 10), and *Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 46}$ (n = 10). (C) Difference in ovariole number between ovaries of each female for the following genotypes: *yw* (n = 36), *Rbf2*$^{\Delta 1}$*/TM6B,Tb* (n = 48), *Rbf2*$^{\Delta 15C}$*/TM6B,Tb* (n = 24), *Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 38}$ (n = 10), *Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 41}$ (n = 10), and *Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 46}$ (n = 10). (D) Box plot representing relative gene expression from ovaries of *Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 46}$ flies in comparison to control *yw* flies. Data represents six biologic replicates. For B, C, D (*) indicates p-value < 0.05. (E) Images of ovary from *yw* and *Rbf2*$^{\Delta 15C}$*/Rbf2*$^{\Delta 15C}$. Images were taken at 10X magnification.

Table 3-3: Rbf2 loss leads to decreased fecundity.

Female	**Male**	**Eggs/Fly/ Hour**	**SD**	**P-value**
yw	*yw*	3.43	0.59	
Rbf2$^{\Delta 15C}$*/Rbf2*$^{\Delta 15C}$	*Rbf2*$^{\Delta 15C}$*/Rbf2*$^{\Delta 15C}$	0.00	0.00	2.09E-04
Rbf2$^{\Delta 1}$*/Rbf2*$^{\Delta 38}$	*Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 38}$	1.36	0.64	7.23E-04
Rbf2$^{\Delta 1}$*/Rbf2*$^{\Delta 41}$	*Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 41}$	1.94	0.56	3.65E-03
Rbf2$^{\Delta 1}$*/Rbf2*$^{\Delta 46}$	*Rbf2*$^{\Delta 1}$*/Rbf2*$^{\Delta 46}$	1.60	0.47	6.36E-04

The table represents the average number of eggs laid by females crossed to males for the shown genotypes. The numbers represent averages of three to five biologic replicates and the corresponding standard deviations (SD). P-values are calculated for each of the crosses in comparison to control *yw* flies.

Specific motifs are associated with different classes of genes.

To identify cis-regulatory elements that may drive the differential gene regulation by Rbf1 and Rbf2, we performed motif analysis on gene promoters in each cluster of the heat map, focusing on regions under Rbf1 or Rbf2 peaks which were described previously (Wei et al. 2015). Certain motifs were enriched only in specific clusters, suggesting that they may represent binding sites for specificity factors that influence the activity of Rbf proteins (Figure 3-17). Cluster 1 possessed a motif with similarities to a cell cycle homology region (CHR) motif, and Cluster 3 was specifically enriched in four motifs for known transcription factors, including the Aef1 repressor protein. Cluster-specific motifs were also noted for Clusters 4 and 5. The disco motif, which also resembles Motif 1 (Ohler et al., 2002) is uniformly distributed across all clusters. Motif 1 is bound by M1BP, and is known to be enriched in growth related genes (Li et al., 2010; Li and Gilmour, 2013). The E2f motif is also uniformly distributed across all clusters, consistent with its important role in mediating E2f/DP1 binding, critical for Rbf recruitment. E2f binding may therefore not be a discriminant for differential Rbf1 and Rbf2. However, the E2f motif is bound by both E2f1 and E2f2; differential binding of these factors may affect regulation by Rbf1 and Rbf2. We referred to E2f1 and E2f2 ChIP datasets (Korenjak et al., 2012), and found that percentage of genes bound by E2f2 is somewhat higher in clusters 3 and 5, while E2f1 bound promoters comprise only a small fraction of each cluster (Table 3-4). Proteins of the Muv/Myb-dREAM complex are known to co-bind promoters with Rb proteins; we note that ChIP data for these proteins (Georlette et al., 2007) identifies a higher fraction of genes in Clusters 3 and 5 (Table 3-4). These results indicate a potential role for E2f2 and the dREAM complexes, but there does not appear to be a simple "code" for differential regulation by Rbf1 and Rbf2.

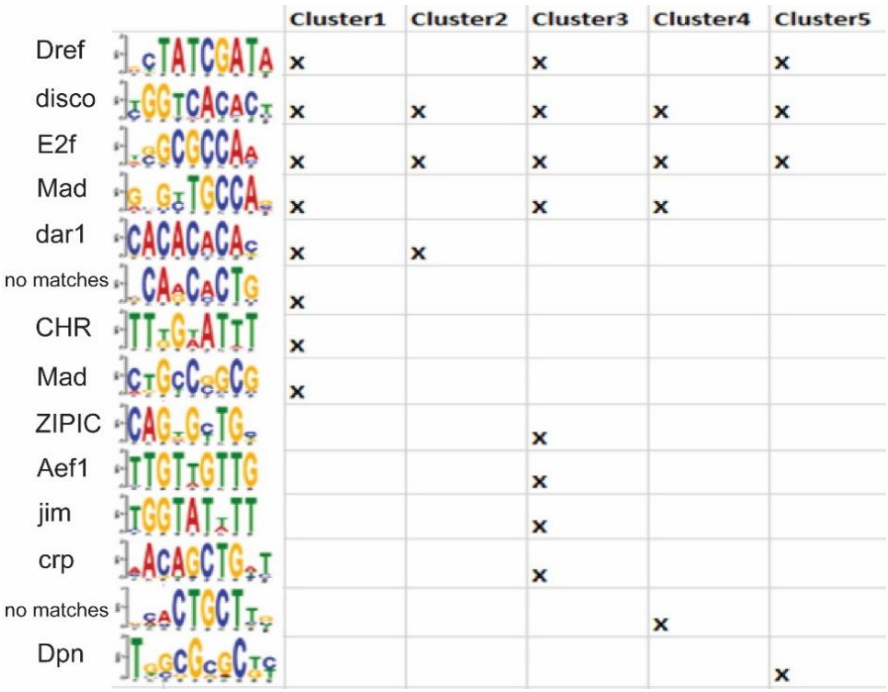


Figure 3-17: Motif analysis of the Rbf1/Rbf2 bound promoter regions of genes within each cluster of the heatmap. The name of the cognate transcription factor to which the motif corresponds to is shown on left of the motif logo.

Table 3-4: Ratio of genes that are shown to be occupied by E2f1, E2f2 or DREAM complex as shown in previous Chip-seq datasets (Korenjak et al., 2012; Georlette et al., 2007).

Cluster	E2f1	E2f2	DREAM
1	0.03	0.37	0.53
2	0.05	0.31	0.42
3	0.05	0.46	0.72
4	0.02	0.41	0.48
5	0.06	0.58	0.81

Cis-regulatory requirements for Rbf2 function.

To understand the effect of promoter structures on Rbf1 and Rbf2 repression, we assayed the activities of promoters from different classes of genes regulated by the corepressors, using luciferase reporters (Figure 3-18A). We tested the effects of expression of Rbf1 or Rbf2 on these promoters in S2 cells. Notably, E2f (as well as Rbf) proteins are endogenously expressed in S2 cells permitting the overexpressed Rbf1 and Rbf2 proteins to interact with promoters (Dimova et al. 2003). The *PCNA* luciferase reporter is strongly repressed by Rbf1, while Rbf2 has no effect on this gene. In contrast, the *CycB* promoter is preferentially repressed by Rbf2, but also shows a strong response to Rbf1 overexpression (Figure 3-18B). These promoter-specific effects indicate that the differential activities of Rbf1 and Rbf2 are not simply a reflection of different expression levels of these corepressors. We hypothesized that differences between the *PCNA* and *CycB* promoters may involve different interactions by E2f1 and E2f2, therefore, we overexpressed E2f1 or E2f2 along with these reporters. *PCNA* is robustly induced by E2f1, but there is little or no effect with E2f2 expression. Strikingly, *CycB* is significantly repressed by E2f1, but induced by E2f2 (Figure 3-18C).

E2f1 induction of *PCNA* (Figure 3-18C) was reversed by co-expression of Rbf1 but not by Rbf2 (Figure 3-18D). In contrast, the weak or nonexistent E2f2 repression was substantially enhanced by either Rbf1 or Rbf2 co-expression (Figure 3-18D). On *CycB*, the promoter that was more sensitive to Rbf2, E2f2 induction (Figure 3-18C) was reversed by co-expression of Rbf1 or Rbf2 (Figure 3-18E). E2f1 repression was little altered by additional Rbf1 or Rbf2 expression (Figure 3-18E). This result suggests that on the *CycB* promoter, E2f2 alone does not act as a repressor unless it is bound by Rbf1 or Rbf2. We propose that on this promoter, E2f2 may compete with endogenous E2f1/Rbf complexes, leading to upregulation. Only when Rbf1 or Rbf2 are expressed

at higher levels does the E2f2 protein become complexed with a corepressor, and form a repressor complex on the *CycB* promoter. In contrast to E2f2, E2f1 would always be recruited to the promoter complexed with Rbf proteins. These data indicate that E2f1 and E2f2 have different impacts on expression of the *PCNA* and *CycB* promoters, and that the E2f activities are differentially regulated by Rbf1 and Rbf2.

The differential responses of these genes to E2f and Rbf proteins is undoubtedly mediated by the distinct sequences of these compact promoters. In order to understand the role of the core promoter region of *PCNA* and *CycB*, we created two chimeric reporters (Figure 3-18A). The first reporter (*CycB-PCNA*) includes *CycB* 5' sequences (-464 to -53) fused to *PCNA* core promoter region (-38 to +23). A complementary reporter, *PCNA-CycB*, includes *PCNA* 5' promoter region (-168 to -38) fused to the *CycB* core promoter region (-53 to +100). Introducing the *CycB* core promoter to the *PCNA* reporter (*PCNA-CycB*) permitted repression by Rbf2, although not as strong as for the wild-type *CycB* construct (Figure 3-19A). Rbf1 repression of this fusion gene was less effective than for the wild-type *PCNA* reporter. Introduction of the *PCNA* core promoter into the *CycB* gene (*CycB-PCNA*) virtually eliminated the strong Rbf2 response; this gene also had weak response to Rbf1 expression (Figure 3-19A). The *CycB* core promoter appears to play a dominant role in sensitivity to E2f1 and E2f2 expression as well; insertion into the *PCNA* gene turns an E2f1-activated gene into an E2f1 repressed gene, while replacement of this core promoter in *CycB* with the corresponding *PCNA* sequences leads to loss of E2f1 repression, and loss of E2f2 activation (Figure 3-19B). On *CycB-PCNA*, co-expression of Rbf1 or Rbf2 along with E2f1 or E2f2 produced a response similar to the expression of the E2f proteins alone (Figure 3-19C). The repression of E2f1 on *PCNA-CycB* is weakened when Rbf1 is coexpressed, while Rbf2 had no impact on the

repression exerted by ectopic E2f1 (Figure 3-19D). The E2f2 induction of *PCNA-CycB* is reversed after coexpression of either Rbf1 or Rbf2 (Figure 3-19D).

Together, these results indicate that both the core promoter region of *CycB* and the more 5' regions are important for optimal repression by Rbf2. The presence of E2f motifs in the *CycB* core promoter suggests that the position or specific sequence of these elements may play a role in differential regulation by the Rbf proteins.

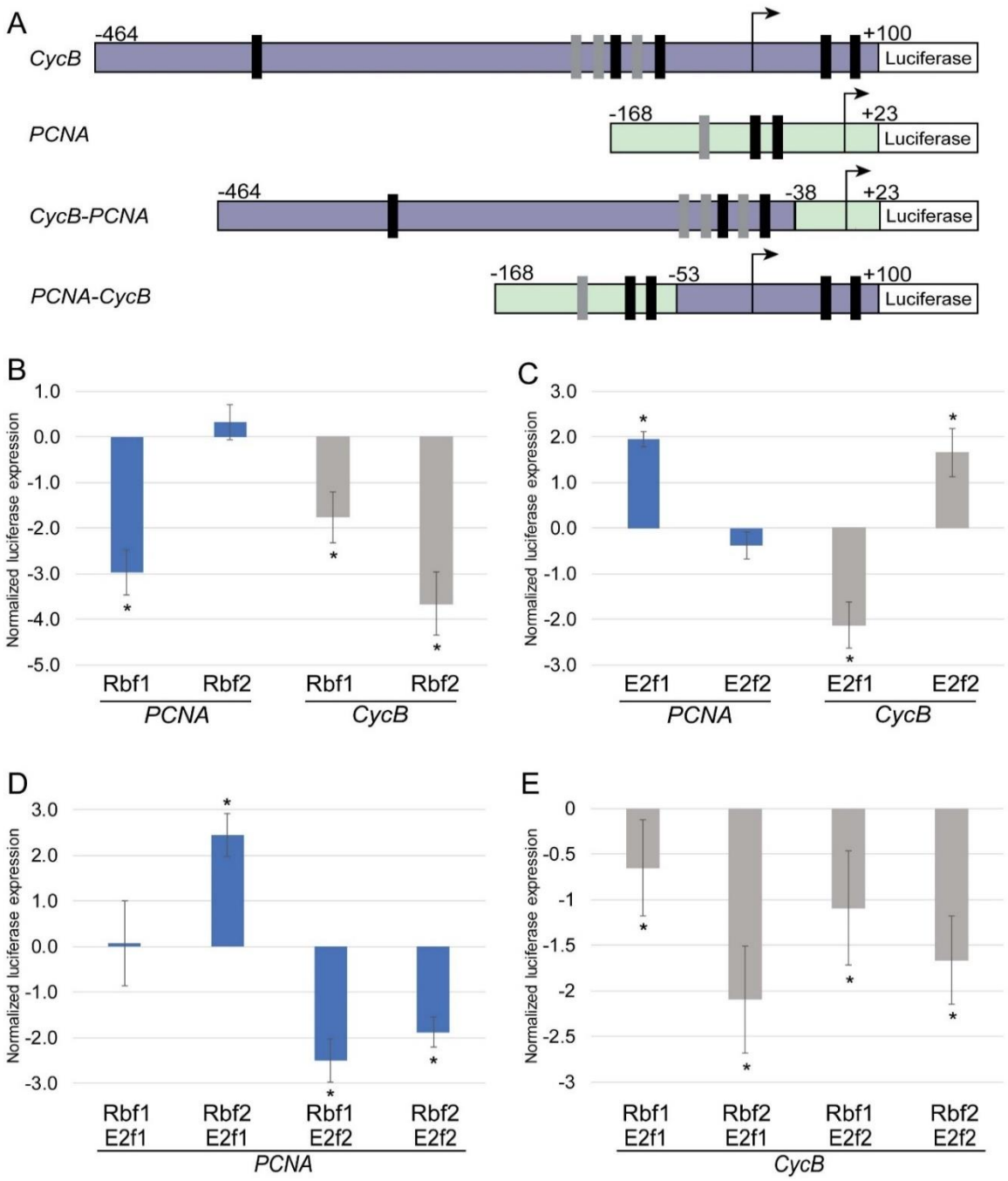


Figure 3-18: Specific regulation of *PCNA* and *CycB* by Rbf1 and Rbf2. (A) Schematic representation of *CycB*, *PCNA*, chimeric *CycB-PCNA* and *PCNA-CycB* luciferase reporter genes. Black bars indicate E2f motifs, and gray bars DREF motifs, often located in cell-cycle related genes. (B) Regulation of *PCNA* and *CycB* by Rbf1 and Rbf2. (C) Regulation of *PCNA* and *CycB*

Figure 3-18 (cont'd)

by E2F1 and E2F2. (D, E) Combined action of Rbf and E2F proteins on *PCNA* and *CycB* promoters. Luciferase measurements were normalized to expression of the reporters in cells cotransfected with the empty expression vectors (no Rbf or E2F genes). Fold changes represented on a $\log_2$ scale plot. Values represent at least three biologic replicates and error bars represent standard deviations. (*) indicates p-value < 0.05.

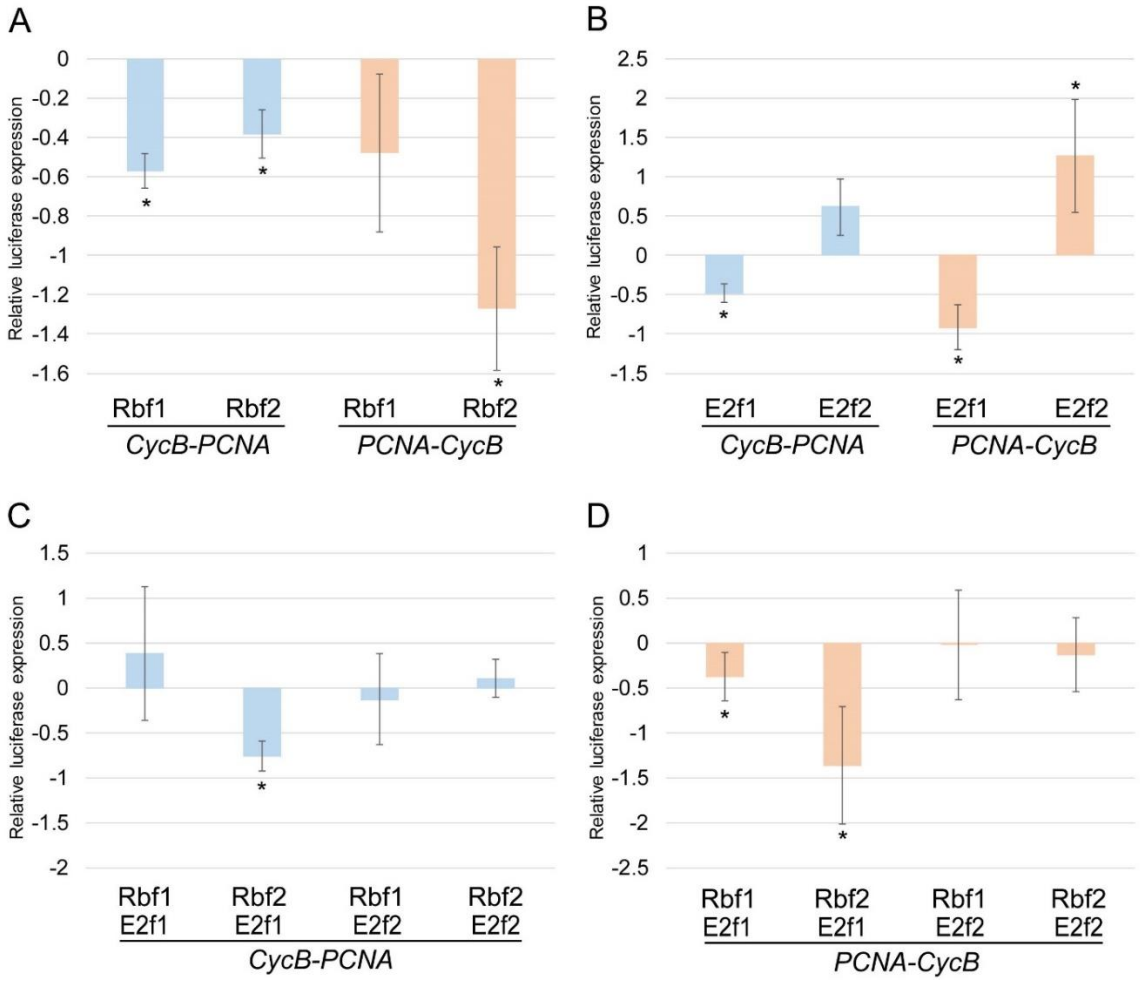


Figure 3-19: The *CycB* core promoter drives responsiveness to Rbf2. (A) Luciferase reporter assays of chimeric reporters *CycB-PCNA* and *PCNA-CycB* in response to expression of Rbf1 or Rbf2 (B) Effect of expression of E2F1 or E2F2 on chimeric reporters. (C, D) Expression of *CycB-*

Figure 3-19 (cont'd)

PCNA or *PCNA-CycB* in response to co-expression of Rbf and E2F proteins. Luciferase measurements are normalized to expression of the reporters in cells cotransfected with the empty expression vector (no Rbf or E2F genes; red horizontal line). Values represent at least three biologic replicates and error bars represent standard deviations. S2(*) indicates p-value < 0.05.

Discussion

Retinoblastoma protein function appears to be indispensable in almost all eukaryotes, however duplication of retinoblastoma genes has only occurred in selected lineages, including in vertebrates and separately in Drosophila. Whether this duplication involves subfunctionalization, neofunctionalization, or both is not currently understood, but our studies of the derived Rbf2 retinoblastoma protein in Drosophila has uncovered features of unique gene targeting, likely linked to rapid evolutionary changes in several conserved parts of the ancestral protein, as well as connection with fertility that may explain why this gene duplication became locked into Drosophila genomes of diverse species. Although the mutant alleles we generated confirm the earlier finding by Dyson that *Rbf2* is not strictly required for viability, the impacts on lifespan and fertility indicate that in fact on an evolutionary scale, the gene is indispensable.

We speculate that *Rbf2* genes have evolved more rapidly than *Rbf1* genes within Drosophila due to specialized functions that are specific to each species. For example, the exact fashion in which transcriptional control is exerted over cell growth-related genes (ribosomal, mitochondrial functions) in response to nutritional signaling may impact the degree to which reproductive strategies are tied to immediate nutritional signals (Terashima and Bownes, 2004). On the other hand, Rbf1, the major regulator of cell cycle genes, may be more conserved within Drosophila, and more widely in metazoa, because of its role in maintaining core cell cycle functions.

Considering functional domains of retinoblastoma proteins, we find parallel changes in mammals and Drosophila. The C-terminus of retinoblastoma proteins is critical for specific binding to E2f transcription factors. Residues in this domain in the mammalian Rb protein permit specific interaction with E2f1, while limiting p107 and p130 to interactions with E2f4-5 (Rubin et al., 2005;

Liban et al., 2017). The C-terminal instability element (IE) region is conserved in the fly Rbf1 as well as the mammalian p107 and p130 proteins; conserved residues in p107 permit specific interaction with E2f4. Strikingly, these residues are conserved in all Drosophila Rbf1 proteins and most arthropods that have a single retinoblastoma protein. The mammalian Rb is divergent in this region; changes in some of the residues allow it to uniquely interact with E2f1 and thus perform Rb-specific functions. Interestingly, Rbf2 is also divergent in this region, perhaps allowing Rbf2 to similarly develop distinct promoter targeting. Indeed, Rbf2 is found at twice the number of promoters as Rbf1, indicating that the binding functions of Rbf1 and Rbf2 are non-identical. Another functional region in retinoblastoma proteins is the spacer region located between the A and B subdomains of the pocket. In mammalian p107 and p130 proteins, the spacer possesses a unique cyclin/cdk binding and inhibition activity that is absent from Rb (Wirt and Sage, 2010). Interestingly, the spacer between the A and B pocket domains is well conserved in Rbf1 among Drosophila, whereas in Rbf2 it is not, possibly affecting the specialized functions of Rbf2 in Drosophila species.

Previous studies of Rbf1 and Rbf2 function focused on these proteins' activities on reporter genes assayed in cultured cells. On specific cell cycle promoters, Rbf2 has only weak effects compared to Rbf1. Using the embryo as a setting for functional tests of Rbf1 and Rbf2, we found that rather than just being a redundant, and less potent version of Rbf1, Rbf2 has unique effects on distinct classes of genes, such as ribosomal and mitochondrial genes in Cluster I, most of which are directly bound by Rbf2. Interestingly, these are genes that are widely expressed and are viewed as "housekeeping" in nature, however, this designation can obscure the dynamic transcriptional regulation that these genes also undergo. It appears that Rbf2 interactions with these genes are geared to effects that are moderate in nature, changing overall output less than twofold in many

cases, a regulation that we deem “soft” repression. Unlike cell cycle target genes that may exhibit complete on/off cycles, these cellular growth-related targets are continuously up and down regulated within specific parameters. Such cybernetic regulation is likely the explanation for the complex transcriptional circuitry found on some Rbf targets, such as the insulin receptor gene, a widely expressed, critical signaling node that includes transcriptional input from Rbf proteins, in addition to a dozen additional genetic elements (Wei et al., 2016). Interestingly, the deployment of specific retinoblastoma proteins to cell growth related genes may be a feature that relates to subfunctionalization of these genes (Figure 3-20A); in human cells, the p130 protein is targeted to many ribosomal protein genes, although the functional relevance remains to be tested.

Regarding the biochemistry of transcriptional regulation, numerous studies have pointed to engagement of mammalian retinoblastoma proteins with a wide spectrum of effectors and targets, including E2f proteins, the basal transcriptional machinery and chromatin regulators (Ross et al., 1999; Dick, 2007; Fiorentino et al., 2013). Similar pathways are likely to be invoked in Drosophila, although this area remains to be explored. It is possible that with divergence of Rbf1 and Rbf2, regulatory mechanisms may also differ, with intrinsic differences in the ability to target basal machinery and recruit histone modifying activities. Alternatively, the finding that Rbf2 regulatory effects appear to be less dramatic than that of Rbf1 may be a function of Rbf2 binding to highly active promoters that are not prone to complete silencing. We explored in depth one instance where Rbf2 exhibits potent repression activity, similar to that found for Rbf1 on its target genes. The ability of Rbf2 to potently inhibit *CycB* reporter appears to be linked to the unique core promoter sequences, which include putative E2f binding sites. The preferential inhibition by Rbf2 is conferred to chimeric reporters containing this unique basal promoter, allowing a switch of preference from Rbf1 to Rbf2. Such activity may point to a preferential interaction with

components of the basal transcription machinery (Figure 3-20B); indeed, the mammalian Rb protein has been shown to interfere with the basal transcription machinery to regulate E2f target genes in vitro (Ross et al., 1999). Interestingly, the core promoter of *CycB* includes a TATA box that is not present in *PCNA* core promoter, while the *PCNA* core promoter includes Initiator and Motif 1; both are absent in *CycB*. Therefore, we think that core promoter elements impact the repression potency of the Rbf proteins. In the embryo, *CycB* is one of the class 5 genes that highly upregulated by Rbf2, in contrast to what we find in our reporter assays in S2 cells. The stage at which we assayed Rbf2 function in the embryo contains many differentiated cells that have exited the cell cycle. In contrast, S2 cells are continuously proliferating, and would activate *CycB* with each cell cycle. The induction in cluster 5 genes observed in the embryo may represent Rbf2 competition with Rbf1 on otherwise silent promoters, whereas repression in S2 cells involves highly active genes.

Preferential action by Rbf1 or Rbf2 may also relate to the type of E2f protein binding to the promoter; certain E2f sites can be bound by either activator or repressor E2fs (Araki et al., 2003). *PCNA* (responsive to Rbf1 and not Rbf2) may be predominantly regulated by E2f1, while *CycB* may be predominantly regulated by E2f2, as supported by ChIP-seq studies (Korenjak et al., 2012). Our coexpression experiments (Figure 3-18) indicate that there may be preferential association with these promoters by unbound E2f proteins, or by E2f associated with Rbf factors. the biochemical basis for such preferential occupancy remains to be elucidated.

A different aspect of Rbf2 function comes from consideration of gene clusters 3 and 5, in which Rbf2 overexpression actually activates genes. Here, antagonism between different retinoblastoma proteins may provide optimal Rbf regulation on certain classes of genes. We found that the most potently induced genes were normally expressed at low levels, possibly silenced by Rbf1. We

hypothesize that Rbf2 overexpression allows the protein to compete with endogenous Rbf1, allowing the genes to be upregulated due to Rbf2's inherently weaker inhibition (Figure 3-20C). This mode of regulation is conceivable, given the normal pattern of expression of Rbf1 and Rbf2. The proteins are co-expressed in many different developmental settings, and it is possible that in addition to its unique roles on certain genes, Rbf2 serves as a moderator of Rbf1 activity through such a competitive mechanism. The alternative occupancy of retinoblastoma target gene promoters by different isoforms is well documented in human cells. For instance, p130 replaces Rb on many promoters in quiescent cells; whether the impact on transcription is equivalent, or whether this poises the genes for alternative regulation is unknown (Chicas et al., 2010).

The developmental role of Rbf2 appears to be tightly, but not uniquely, linked to reproduction. Loss of Rbf2 leads to decreased fecundity. Interestingly, Dyson and colleagues had previously reported the opposite phenotype for a disruption in the *Rbf2* locus that they had engineered (Stevaux et al., 2005). However, in their study, they compared egg laying to control flies possibly without taking into consideration the background effects. In addition, what they assumed was a functional null may in fact have residual activity, as an N-terminal portion of the protein was still expressed.

Gene expression changes in the ovary point to a role for Rbf2 in signaling pathways connected to regulation of oogenesis; we find that *Pi3K92E* is significantly upregulated in *Rbf2* mutant female ovaries. Female fertility is influenced by the InR/Pi3K signaling pathway (Orme et al., 2006; Pritchett and McCall, 2012). We propose that Rbf2 may be optimizing oogenesis by regulating the InR signaling pathway through Pi3K. In addition to this role in regulation of signaling molecules in the adult, a role for Rbf2 in female reproduction would also involve the development of the ovary, as *Rbf2* heterozygotes often possess ovaries with an increased number of ovarioles, which

are specified in larval development. The roles for retinoblastoma proteins in development of the reproductive system appear to be conserved; the conditional knockout of *Rb* in female mice leads to progressive infertility (Andreu-Vieyra et al., 2008), suggesting that it would be important to examine the role of Rb in regulating fertility in humans as well. In addition to a clear connection to female reproduction, it appears that Rbf2 expression has additional roles in fly physiology. We observed that both male and female *Rbf2* mutants had shorter lifespans, suggesting that the expression of Rbf2 in embryo and larva outside of reproductive tissues is likely to have significant consequences in both sexes.

Our study has provided important insights on parallel evolution of retinoblastoma paralogs. Mammalian Rb and Drosophila Rbf2 are the most derived proteins in the retinoblastoma family in their respective lineages, and appear to have acquired indispensable new functions that may represent a process of sub-functionalization or neofunctionalization. The findings that Rbf2 is more subject to evolutionary modifications, has acquired unique gene targeting activities, and may play a role in functional antagonism of the ancestral protein appear to mirror similar processes in mammalian systems. Our study demonstrates how biochemical and physiological activities of these conserved transcriptional corepressors are subject to evolutionary modification, and how diverse retinoblastoma protein functions in humans may be better understood in model systems such as Drosophila.

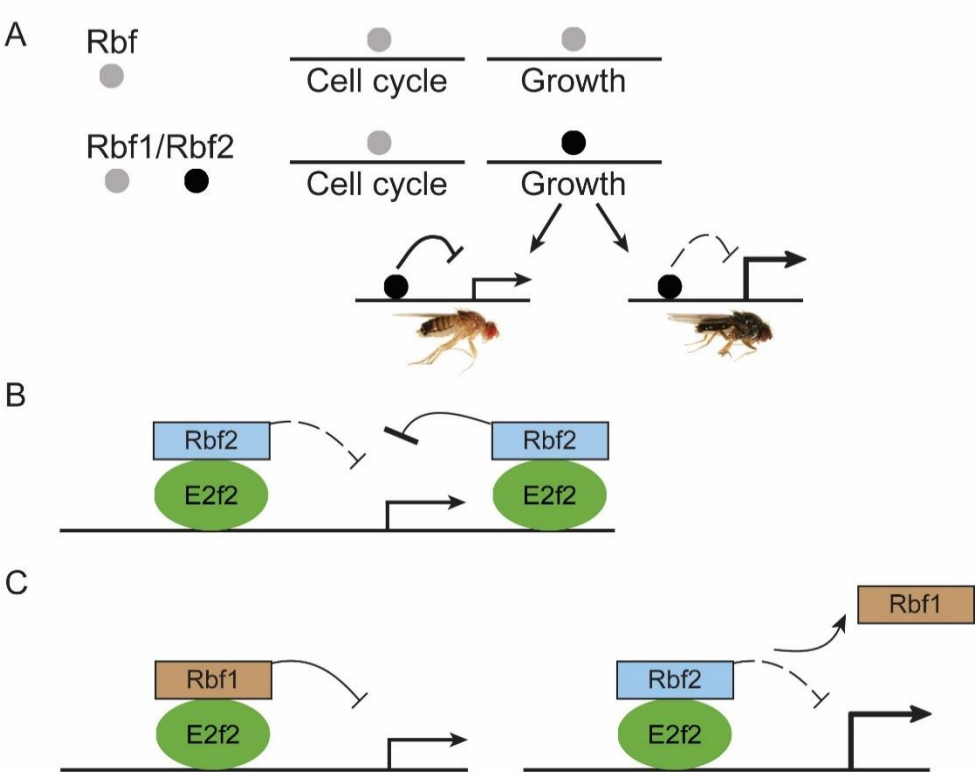


Figure 3-20: Model for evolved functions of Rbf proteins. (A) Subfunctionalization of Rbf proteins into cell cycle and cell growth control. Ancestral Rbf proteins (gray circle) regulate cell cycle and cell growth related genes. Duplication of Rbf proteins resulted in subfunctionalization where the more derived Rbf2 protein (black circle) assumes regulation of cell growth related genes. Rbf2 protein is subject to more rapid evolution within different Drosophila species to provide optimal growth control and fecundity. (B) Model for specific action of Rbf2 from the core promoter position, as in *CycB*; repression is weaker from 5' positions. (C) Rbf2 competes with Rbf1 binding on E2F2 regulated genes, leading to partial derepression and optimal gene regulation on certain classes of genes.

Materials and Methods

Protein sequence alignments

Protein sequences of retinoblastoma genes were obtained from FlyBase and NCBI databases. Multiple sequence alignments were generated with Clustal Omega v1.2.4 and ClustalW v2.1 (European Bioinformatics Institute, EBI) using default settings and manual adjustments. Output files from ClustalW were visualized with Jalview v2.10.5 (EBI).

Creation of transgenic lines to express Rbf1 and Rbf2 proteins, and generation of novel *Rbf2* alleles with CRISPR

For expression of proteins in the embryo, FLAG-tagged *Rbf1* and *Rbf2* cDNAs were subcloned from the *pAX* vector (Acharya et al., 2010; Wei et al., 2015) into the *pattB* heatshock vector (Kok et al., 2015) using *Hind*III and *Xba*I restriction sites for *Rbf1*. For *Rbf2*, a bridge oligonucleotide containing *Bgl*I and *Not*I sites was cloned between *Hind*III and *Xba*I sites, and *Rbf2* cDNA was inserted into *pattB* using these new restriction sites. The plasmids were injected by Rainbow Transgenics into the 51D site on the second chromosome of *yw* flies to generate the homozygous transgenic lines.

Genomic *Rbf2* target sites were identified at http://tools.flycrispr.molbio.wisc.edu/targetFinder/ (Gratz et al. 2014). Three sites near the 5' end of the *Rbf2* coding region were selected. Guide-RNAs targeting *ebony* (gRNA-e) and *Rbf2* (gRNA-1, gRNA-2, and gRNA-3) were inserted in vector *pU6-BbsI-chiRNA* (Addgene plasmid #45946) as described (Gratz et al., 2014). The sequences for *ebony* gRNA is: 5'-CTTCGCCACAATTGTCGATCGTCA-3' and 5'-AAACTGACGATCGACAATTGTGGC. The sequences for each *Rbf2* gRNA are as follows: 5'-

CTTCGCAGCGCTTGGTAGTGATTCG-3’ and 5’-AAACCGAATCACTACCAAGCGCTGC-3’ for gRNA-1; 5’-CTTCGCTCGAAGATGCGCGATATTA-3’ and 5’-AAACTAATATCGCGCATCTTCGAGC-3’ for gRNA-2; and 5’-CTTCGTCTGTCCACCTACCATCGCT-3’ and 5’-AAACAGCGATGGTAGGTGGACAGAC-3’ for gRNA-3. *y[1] M{w[+mC]=nos-Cas9.P}ZH-2A w[*]* embryos were injected with each gRNA along with gRNA-e by BestGene Inc. Fly crosses and screening were accomplished using a co-CRISPR strategy adapted from Kane and colleagues (Kane et al., 2017). Injected adult flies, whose germlines potentially contained mutated *Rbf2* alleles, were crossed to the double balancer stock *w[1118]/Dp(1;Y)y[+]; CyO/Bl[1]; TM2, e/TM6B, e, Tb[1]* (Bloomington Drosophila Stock Center #3704) in the parental F(0) generation. Flies in the F (1) generation were scored for ebony body color and tubby pupal shape. These progeny were then crossed to the third chromosome balancer stock *w[1118]; InR[GC25]/TM6B, e, Tb[1]*. Ebony and tubby phenotypes were again scored in the F(2) generation and the flies were crossed inter se to produce homozygous (non-tubby) and balanced (TM6B,Tb) fly lines. Genomic DNA extraction and PCR-amplification of the target-regions was performed on all homozygote and heterozygote lines. Mutations were confirmed by Sanger sequencing. $Rbf2^{\Delta 1}$ represents a one base pair deletion that disrupts the codon for N168 of Rbf2 and creates an out of frame, truncated protein with an additional 28 amino acids. $Rbf2^{\Delta 15C}$ represents a 15 base pair deletion that eliminates codons 168-172 (NHYQA). $Rbf2^{\Delta 15N}$ represents a 15 base pair deletion that disrupts codon 71 (M71T) and eliminates codons 72-76 (RDIRE). $Rbf2^{\Delta 38}$ represents a 38 base pair deletion that disrupts the codon for S64 and creates an out of frame, truncated protein with an additional 26 amino acids. $Rbf2^{\Delta 41}$ represents a 41 base pair deletion that disrupts the codon for Y63 and creates an out of frame, truncated protein with an additional 26 amino acids. $Rbf2^{\Delta 46}$ represents a

46 base pair deletion that disrupts the codon for C60 and creates an out of frame, truncated protein with an additional 20 amino acids. Introgression of the *Rbf2*$^{\Delta 1}$ allele into a lab stock of *yw* flies was performed over 5 generations.

Measurements of gene expression

RT-qPCR

Rbf1 or *Rbf2* transgenes were induced in 12-18 hour embryos by means of a 20-minute heat shock, floating 35 mm apple juice plates with freshly laid embryos on a covered water bath. After 20, 40 or 60 minutes recovery time, RNA was extracted using Total RNA Kit (OMEGA). Control embryos lacking the heat shock transgene were treated similarly to control for nonspecific heat shock effects. cDNA synthesis was performed on total RNA using high capacity cDNA reverse transcription kit (Applied Biosystems). qPCR analysis was performed using Perfecta SYBR Green Fastmix (Quanta Bio). Three biologic replicates were done for both control and transgenic flies (Figure 3-21). To analyze gene expression in *Rbf2* mutants, RNA was extracted from ovaries using Trizol followed by cDNA synthesis and qPCR analysis, with six biologic replicates from control flies and the mutants. Sequence of primers used are available upon request.

RNA-seq analysis

Rbf1 or *Rbf2* transgenes were induced with a 20-minute heat shock in 12-18 hour embryos. After 60 minute-recovery, RNA was extracted using Total RNA Kit (OMEGA). Control flies were treated similarly to control for the effect of heat shock on gene expression. Poly-A+ RNAs were purified from the total RNA using Oligotex mRNA Mini kit (Qiagen) and were prepared for the SMS essentially as described previously (Kapranov et al. 2010). Sequencing was performed at the

SeqLL, LLC facility (Woburn, MA). The SMS reads were processed basically as described before (Kapranov et al., 2010) and aligned to the DM6 version of the *Drosophila melonagaster* genome using indexDPgenomic aligner (Giladi et al., 2010). Uniquely aligned reads were used to generate RPKM values for each transcript annotated in the RefSeq Genes database of the UCSC Genome browser (http://hgdownload.soe.ucsc.edu/goldenPath/dm6/database/refGene.txt.gz) (Kent et al., 2002). Three biologic replicates were done for each sample. Reads lower than RPKM of 1 were removed, which reduced the number of genes to 12,060. Only genes bound by either Rbf1 or Rbf2 based on ChIP-seq dataset (Wei el al., 2015) were further analyzed, for a total of 3937 genes. Unsupervised clustering was performed using Cluster3.0 software, and the heatmap was visualized using JAVA TreeView v1.1.6r4. The counts were log transformed and mean centered, and filtered at 0.3 SD to remove genes with little variation across samples. The heatmap that includes 2795 genes and we decided to analyze the data at the level of five major clusters. Gene ontology analysis was performed for each cluster using DAVID v6.8. Motif analysis was performed on Rbf2 or Rbf1 peak regions of genes in each cluster using MEME-ChIP v4.12.0. Rbf1 and Rbf2 ChIP peak regions are described previously; Rbf1 peak region is 98 bp long, and Rbf2 is 160 bp (Wei et al., 2015). Matched motifs were obtained using Tomtom (MEME Suite). Core promoter elements described in Ohler et al. (2002) were identified in *PCNA* and *CycB* core promoters using MAST (MEME Suite). The list of Drosophila ribosomal genes was obtained as previously described (Wei el al., 2015), and the list of mitochondrial genes was obtained from MitoDrome database (Sardiello et al., 2003).

Lifespan and fertility assays

Measurement of the *Rbf2* mutant flies' lifespans was done as previously described (Linford et al., 2013). 100 females and 100 males from each genotype were separated, and 10 flies were placed

in separate vials and maintained at 25°C. After transferring flies into new vials, dead flies were counted and recorded. Log-rank test was used to assess significance of differences in lifespans (Yang et al., 2011). Assays for fertility were adapted from Stevaux et al. 2005. For transheterozygous *Rbf2* mutants, eggs were counted after crossing about 15 staged virgin females with 15 males in laying bottles. After two 30-minute preclearing steps, 3-hour collections were made to determine the egg-laying rate. For the infertile homozygous $Rbf2^{\Delta 15C}$ lines, individual virgin females and virgin males were crossed with *yw* flies to assess female or male sterility. For the introgressed $Rbf2^{\Delta 1}$ allele, fertility was measured using single female $rbf2^{\Delta 1}$ flies crossed to *yw* males, and single male $Rbf2^{\Delta 1}$ flies crossed to single *yw* females. Eggs were counted after 24 hours of laying. For all the crosses, the egg count was conducted 6 to 9 days after eclosion, and egg laying rates averaged for four measurements. Egg laying experiments were done in parallel with control flies under same conditions.

Analysis of mutant ovaries

Ovaries were dissected from staged females on ice-cold PBS with 0.1% Triton X-100. Whole ovaries were directly mounted in 75% glycerol and imaged on a Leica compound microscope under 10X magnification. Ovarioles were split apart and isolated with microsurgical forceps and a fine needle and their number was recorded.

Western blot analysis

Ovaries were dissected from staged females and homogenized with a polypropylene pestle in lysis buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 1% Triton X-100). The concentration of the extracts was determined via Bradford protein assay, and 50 μg of protein was run per lane in 10% SDS-PAGE gels. Gels were analyzed by Western blot on PVDF membrane using anti-Rbf2 rabbit

antibodies (Keller et al., 2005) that bind to the C-terminal end of the protein. Primary antibodies were diluted 1:5000 in TBST (20 mM Tris-Cl, pH 7.5, 120 mM NaCl, 0.1% Tween-20) with 5% nonfat dry milk, and incubation was done overnight at 4°C. Blots were developed using HRP-conjugated goat anti-rabbit secondary antibodies (1:10,000) (30-minute incubation at room temperature) and SuperSignal West Pico chemiluminescent substrate (Pierce).

Reporter constructs

The *CycB* promoter region (-464 to +100) was cloned into *Asc*I and *Sal*I sites in the pAC2T-luciferase vector (Acharya et al., 2010). The *PCNA*-luciferase reporter (a gift from the Nick Dyson laboratory) was previously described (Yamaguchi et al., 1995; Acharya et al., 2010). The *pIE-E2F1* and *pIE-E2F2* vectors were a gift of the Maxim Frolov laboratory (Frolov et al., 2001). *PCNA-CycB* and *CycB-PCNA* hybrid constructs were synthesized as Gblock gene fragments by IDT (IDTDNA.com) and cloned into the *pAC2T-luciferase* vector (Wei et al., 2015) using *Asc*I and *Sal*I sites.

Luciferase reporter assays

Drosophila SL2 cells were cultured in Schneider's medium (Gibco) supplied with 10% HI-FBS and penicillin-streptomycin (100 units/mL penicillin and 100 µg/mL streptomycin, Gibco). 1.5 million cells were transfected using Effectene transfection reagent (Qiagen) with 250 ng each of reporter vector, pAX-*Rbf1* or pAX-*Rbf2*, pAX vector as control, and pRL-CMV *Renilla* luciferase reporter. Co-transfection with 250ng of *pIE-E2F1* or *pIE-E2F2* along with *pAX-Rbf1* or *pAX-Rbf2* was also performed, compared to equal amount (500ng) of *pAX* vector as control. Cells were harvested 72 hours post-transfection, and luciferase assays were conducted as described previously (Acharya et al., 2010*;* Wei et al., 2015*).*

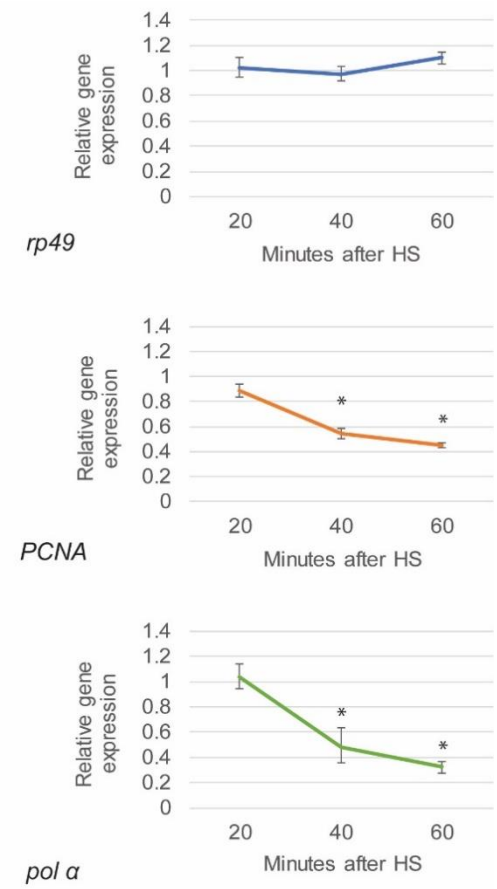


Figure 3-21: Kinetics of gene expression after induction of Rbf1 protein in 12-18hr embryos. *Rp49* is used as control. Data represents average of three biologic replicates. Error bars indicate standard deviation. (*) indicates p-value <0.05.

Acknowledgements

We thank J. Ravi, J. Rennhack, R.W. Henry, S. Payankaulam, Y. Wei and members of the Arnosti lab for technical assistance and advice. We thank Nick Dyson for sharing the *PCNA* and *Polα* luciferase reporter genes and Maxim Frolov for providing the pIE-Myc-*E2f1* and pIE-Myc-*E2f2* constructs. We thank the Bloomington Stock Center for Drosophila lines. This research was supported by NIH grant GM124137 to DNA. and the Michigan State University BEACON consortium to RM.

REFERENCES

REFERENCES

Acharya, P., Negre, N., Johnston, J., Wei, Y., White, K. P., Henry, R. W. and Arnosti, D. N. (2012). Evidence for autoregulation and cell signaling pathway regulation from genome wide binding of the Drosophila retinoblastoma protein. *G3 (Bethesda).* **2**, 1459–72.

Acharya, P., Raj, N., Buckley, M. S., Zhang, L., Duperon, S., Williams, G., Henry, R. W. and Arnosti, D. N. (2010). Paradoxical Instability–Activity Relationship Defines a Novel Regulatory Pathway for Retinoblastoma Proteins. *Mol. Biol. Cell* **21**, 3890–3901.

Andreu-Vieyra, C., Chen, R. and Matzuk, M. M. (2008). Conditional Deletion of the Retinoblastoma (*Rb*) Gene in Ovarian Granulosa Cells Leads to Premature Ovarian Failure. *Mol. Endocrinol.* **22**, 2141–2161.

Araki, K., Nakajima, Y., Eto, K. and Ikeda, M.-A. (2003). Distinct recruitment of E2F family members to specific E2F-binding sites mediates activation and repression of the E2F1 promoter. *Oncogene* **22**, 7632–41.

Burke, J. R., Hura, G. L. and Rubin, S. M. (2012). Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control. *Genes Dev.* **26**, 1156–66.

Burkhart, D. L. and Sage, J. (2008). Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat. Rev. Cancer* **8**, 671–82.

Chicas, A., Wang, X., Zhang, C., McCurrach, M., Zhao, Z., Mert, O., Dickins, R. A., Narita, M., Zhang, M. and Lowe, S. W. (2010). Dissecting the Unique Role of the Retinoblastoma Tumor Suppressor during Cellular Senescence. *Cancer Cell* **17**, 376–387.

Dick, F. A. (2007). Structure-function analysis of the retinoblastoma tumor suppressor protein – is the whole a sum of its parts? *Cell Div.* **13**, 2–26.

Dimova, D. K., Stevaux, O., Frolov, M. V and Dyson, N. J. (2003). Cell cycle-dependent and cell cycle-independent control of transcription by the Drosophila E2F/RB pathway. *Genes Dev.* **17**, 2308–20.

Du, W., Vidal, M., Xie, J. E., Dyson, N. (1996). RBF, a novel RB-related gene that regulates E2F activity and interacts with cyclin E in Drosophila. *Genes Dev.* **10**, 1206–18.

Fiorentino, F. P., Marchesi, I. and Giordano, A. (2013). On the role of retinoblastoma family proteins in the establishment and maintenance of the epigenetic landscape. *J. Cell. Physiol.* **228**, 276–84.

Frolov, M. V, Huen, D. S., Stevaux, O., Dimova, D., Balczarek-Strang, K., Elsdon, M. and Dyson, N. J. (2001). Functional antagonism between E2F family members. *Genes Dev.* **15**, 2146–60.

Georlette, D., Ahn, S., MacAlpine, D. M., Cheung, E., Lewis, P. W., Beall, E. L., Bell, S. P., Speed, T., Manak, J. R. and Botchan, M. R. (2007). Genomic profiling and expression studies reveal both positive and negative activities for the Drosophila Myb MuvB/dREAM complex in proliferating cells. *Genes Dev.* **21**, 2880–96.

Giladi, E., Healy, J., Myers, G., Hart, C., Kapranov, P., Lipson, D., Roels, S., Thayer, E. and Letovsky, S. (2010). Error Tolerant Indexing and Alignment of Short Reads with Covering Template Families. *J. Comput. Biol.* **17**, 1397–1411.

Gratz, S. J., Ukken, F. P., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. M. and O'Connor-Giles, K. M. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics* **196**, 961–71.

Henley, S. A. and Dick, F. A. (2012). The retinoblastoma family of proteins and their regulatory functions in the mammalian cell division cycle. *Cell Div.* **7**, 10.

Kane, N. S., Vora, M., Varre, K. J. and Padgett, R. W. (2017). Efficient Screening of CRISPR/Cas9-Induced Events in Drosophila Using a Co-CRISPR Strategy. *G3 (Bethesda).* **7**, 87–93.

Kapranov, P., St Laurent, G., Raz, T., Ozsolak, F., Reynolds, C. P., Sorensen, P. H., Reaman, G., Milos, P., Arceci, R. J., Thompson, J. F., et al. (2010). The majority of total nuclear-encoded non-ribosomal RNA in a human cell is "dark matter" un-annotated RNA. *BMC Biol.* **8**, 149.

Keller, S. A., Ullah, Z., Buckley, M. S., Henry, R. W. and Arnosti, D. N. (2005). Distinct developmental expression of Drosophila retinoblastoma factors. *Gene Expr. Patterns* **5**, 411–21.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. and Haussler, D. (2002). The human genome browser at UCSC. *Genome Res.* **12**, 996–1006.

Kok, K., Ay, A., Li, L. M. and Arnosti, D. N. (2015). Genome-wide errant targeting by Hairy. *Elife* **4**.

Korenjak, M., Anderssen, E., Ramaswamy, S., Whetstine, J. R. and Dyson, N. J. (2012). RBF binding to both canonical E2F targets and noncanonical targets depends on functional dE2F/dDP complexes. *Mol. Cell. Biol.* **32**, 4375–87.

Li, L., Edgar, B. A. and Grewal, S.S. (2010). Nutritional control of gene expression in Drosophila larvae via TOR, Myc and a novel cis-regulatory element. *BMC Cell Biol.* **11**, 7.

Li, J. and Gilmour, D.S. (2013). Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *EMBO J.* **32**, 1829–1841.

Liban, T. J., Medina, E. M., Tripathi, S., Sengupta, S., Henry, R. W., Buchler, N. E. and Rubin, S. M. (2017). Conservation and divergence of C-terminal domain structure in the retinoblastoma protein family. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4942–4947.

Linford, N. J., Bilgir, C., Ro, J. and Pletcher, S. D. (2013). Measurement of Lifespan in Drosophila melanogaster. *J. Vis. Exp.* **7**, 71.

Longworth, M. S., Walker, J. A., Anderssen, E., Moon, N.-S., Gladden, A., Heck, M. M. S., Ramaswamy, S. and Dyson, N. J. (2012). A shared role for RBF1 and dCAP-D3 in the regulation of transcription with consequences for innate immunity. *PLoS Genet.* **8**, e1002618.

Ohler, U., Liao, G., Niemann, H. and Rubin G. M. (2002). Computational analysis of core promoters in the Drosophila genome. *Genome Biol.* **3**, 12.

Orme, M. H., Alrubaie, S., Bradley, G. L., Walker, C. D. and Leevers, S. J. (2006). Input from Ras is required for maximal PI(3)K signalling in Drosophila. *Nat. Cell Biol.* **8**, 1298–302.

Pritchett, T. L. and McCall, K. (2012). Role of the insulin/Tor signaling network in starvation-induced programmed cell death in Drosophila oogenesis. *Cell Death Differ.* **19**, 1069–79.

Raj, N., Zhang, L., Wei, Y., Arnosti, D. N. and Henry, R. W. (2012). Rbf1 degron dysfunction enhances cellular DNA replication. *Cell Cycle* **11**, 3731–3738.

Ross, J. F., Liu, X. and Dynlacht, B. D. (1999). Mechanism of transcriptional repression of E2F by the retinoblastoma tumor suppressor protein. *Mol. Cell* **3**, 195–205.

Rubin, S. M., Gall, A.-L., Zheng, N. and Pavletich, N. P. (2005). Structure of the Rb C-terminal domain bound to E2F1-DP1: a mechanism for phosphorylation-induced E2F release. *Cell* **123**, 1093–106.

Saeed, M., Schwarze, F., Loidl, A., Meraner, J., Lechner, M. and Loidl, P. (2012). In vitro phosphorylation and acetylation of the murine pocket protein Rb2/p130. *PLoS One* **7**, e46174.

Sardiello, M., Licciulli, F., Catalano, D., Attimonelli, M. and Caggese, C. (2003). MitoDrome: a database of Drosophila melanogaster nuclear genes encoding proteins targeted to the mitochondrion. *Nucleic Acids Res.* **31**, 322–4.

Sengupta, S., Lingnurkar, R., Carey, T. S., Pomaville, M., Kar, P., Feig, M., Wilson, C. A., Knott, J. G., Arnosti, D. N. and Henry, R. W. (2015). The Evolutionarily Conserved C-terminal Domains in the Mammalian Retinoblastoma Tumor Suppressor Family Serve as Dual Regulators of Protein Stability and Transcriptional Potency. *J. Biol. Chem.* **290**, 14462–75.

Stevaux, O., Dimova, D. K., Ji, J.-Y., Moon, N. S., Frolov, M. V and Dyson, N. J. (2005). Retinoblastoma family 2 is required in vivo for the tissue-specific repression of dE2F2 target genes. *Cell Cycle* **4**, 1272–80.

Stevaux, O., Dimova, D., Frolov, M. V, Taylor-Harding, B., Morris, E. and Dyson, N. (2002). Distinct mechanisms of E2F regulation by Drosophila RBF1 and RBF2. *EMBO J.* **21**, 4927–37.

Terashima, J. and Bownes, M. (2004). Translating available food into the number of eggs laid by Drosophila melanogaster. *Genetics* **167**, 1711–9.

Wei, Y., Gokhale, R. H., Sonnenschein, A., Montgomery, K. M., Ingersoll, A. and Arnosti, D. N. (2016). Complex cis-regulatory landscape of the insulin receptor gene underlies the broad expression of a central signaling regulator. *Development* **143**, 3591–3603.

Wei, Y., Mondal, S. S., Mouawad, R., Wilczyński, B., Henry, R. W. and Arnosti, D. N. (2015). Genome-Wide Analysis of Drosophila RBf2 Protein Highlights the Diversity of RB Family Targets and Possible Role in Regulation of Ribosome Biosynthesis. *G3 (Bethesda).* **5**, 1503–15.

Wirt, S. E. and Sage, J. (2010). p107 in the public eye: an Rb understudy and more. *Cell Div.* **5**, 9.

Yamaguchi, M., Hayashi, Y. and Matsukage, A. (1995). Essential role of E2F recognition sites in regulation of the proliferating cell nuclear antigen gene promoter during Drosophila development. *J. Biol. Chem.* **270**, 25159–65.

Yang, J.-S., Nam, H.-J., Seo, M., Han, S. K., Choi, Y., Nam, H. G., Lee, S.-J. and Kim, S. (2011). OASIS: online application for the survival analysis of lifespan assays performed in aging research. *PLoS One* **6**, e23525.

Zhang, L., Wei, Y., Pushel, I., Heinze, K., Elenbaas, J., Henry, R. W. and Arnosti, D. N. (2014). Integrated stability and activity control of the Drosophila Rbf1 retinoblastoma protein. *J. Biol. Chem.* **289**, 24863–73.

CHAPTER 4

Selective repression of the Drosophila *Cyclin B* promoter by retinoblastoma and E2f proteins

Abstract

The Cyclin B1 gene encodes a G2/M cyclin that is deregulated in various human cancers, however, the transcriptional regulation of this gene is incompletely understood. The E2F and retinoblastoma family of proteins are clearly involved in this gene's regulation, but there is disagreement on which of the different E2F and retinoblastoma proteins interact with the promoter to regulate this gene. Here, we dissect the promoter region of the Drosophila *CycB* gene, and study the role of Rbf and E2F factors in its regulation. This gene exhibits several remarkable features that distinguish it from G1/S regulated promoters, such as *PCNA*. The promoter is comprised of modular elements with dedicated repressor and activator functions, including a segment spanning the first intron that interferes with a 5' activator element. A highly active minimal promoter (-464, +100) is repressed by the Rbf1 retinoblastoma protein, but much more potently repressed by the Rbf2 retinoblastoma protein, which has been linked in other studies to control of cell growth genes. Unlike many other cell-cycle related genes, which are activated by E2F1 and repressed by E2F2, *CycB* is potently activated by E2F2, and repressed by E2F1. Although the bulk of Rbf binding is associated with a region 5' of the core promoter, E2F and retinoblastoma proteins functionally interact with the basal promoter region, in part through a conserved E2F site at -80 bp. The specific regulatory requirements of this late cell cycle promoter appear to be linked to the unique activities of E2F and retinoblastoma family members acting on a complex cis-regulatory circuit.

This work was submitted as the following manuscript:

Mouawad, R., Himadewi, P., Kadiyala, D., Arnosti, DN. (2019). Selective repression of the Drosophila Cyclin B promoter by retinoblastoma and E2F proteins. (submitted).

My contribution to this study was execution of all luciferase reporter assays, and motif search on the *CycB* promoters.

Introduction

Cyclin-CDK complexes are core regulators of cell cycle progression from the quiescent (G_0) phase to the mitosis (M) phase. At least five major types of cyclins are present in mammals based on structural similarities and expression during cell cycle, and each cyclin is further subdivided into multiple subtypes (Ito, 2000; Satyanarayana and Kaldis, 2009). Cyclins are differentially expressed during each phase of the cell cycle, in a manner that is conserved from lower eukaryotes to humans (Ito, 2000). Cyclins are in some cases apparently functionally redundant, but deletions in cyclin A2 and cyclin B1 are embryonic lethal in the mouse, indicating unique roles in cell cycle control and development (Satyanarayana and Kaldis, 2009). In addition to transcriptional regulation, cyclins are regulated by degradation through the ubiquitin-proteosome pathway (Bloom and Cross, 2007).

The first human cyclin to be identified was cyclin B1 (Pines and Hunter, 1989), a G2/M cyclin; its mRNA level peaks at the G2/M phase of the cell cycle and is necessary for entry into mitosis (Ito 2000). In association with Cdk1, cyclin B1 promotes nuclear envelope breakdown, chromosome condensation, and mitotic spindle assembly (Satyanarayana and Kaldis, 2009). Several studies have provided insights into the transcriptional regulation of the human cyclin B1 gene *CCNB1* through the use of reporter genes (Hwang et al., 1995; Cogswell et al., 1995; Piaggio et al., 1995). However, despite the identification of multiple cis elements in the *CCNB1* promoter, none of these were necessary for the G2/M-specific activation (Ito, 2000). Therefore, many questions remain regarding the core promoter cis elements involved in regulating this gene, a topic of considerable interest in light of the elevated expression observed in human tumors. For instance, upregulation of *CCNB1* is associated with poor prognosis in breast cancer (Casimiro et al., 2012).

Regarding the trans-acting factors linked to *CCNB1*, various studies have indicated that p53, FOXM1 and others control transcription of this gene, and that the mRNA is posttranscriptionally regulated by miRNAs (Innocente et al., 1999; Laoukili et al., 2005; Huang et al., 2012; Khan et al., 2013; Shi et al., 2016). At a posttranslational level, the Rb/E2F pathway has been shown to regulate the levels of cyclin B protein indirectly by transcriptional inhibition of the gene for the cyclin A protein which, when complexed with cdk2, phosphorylates and inactivates the anaphase promoting complex (APC), the ubiquitin ligase that triggers degradation of cyclin B (Lukas et al., 1999).

The Rb/E2F pathway has been shown to be directly involved in regulating cyclin B1 expression, but a clear molecular mechanism is still lacking (Markey et al., 2002; Jackson et al., 2005; Jackson and Pereira-Smith 2006; Li et al., 2012). Three retinoblastoma family proteins are found in mammalian genomes, Rb, p107 and p130, which possess overlapping yet distinct roles in gene regulation (Dyson, 1998; Henley and Dick, 2012; Wirt and Sage, 2010). Rb proteins interact with five E2F transcription factors that are canonically divided into two groups: E2F1-3, mainly involved in transcriptional activation, and E2F4-5, mainly involved in repression (Du and Pogoriler, 2006). In doxorubicin treated MCF-7 cells, p130 binds to the Cyclin B gene but is not required for repression; only when all three Rb proteins are absent is Cyclin B deregulated (Jackson and Pereira-Smith, 2006). Consistent with this result, overexpression of Rb in A2-4 and A5-1 rat fibroblasts leads to downregulation of the Cyclin B1 gene (Markey et al., 2002). However, another study showed that the Cyclin B gene in mouse embryonic fibroblasts is upregulated in p107 and p130 double knockout cells after DNA damage, despite the presence of Rb (Jackson et al., 2005). In addition to the different role of pocket proteins, the impact of different E2F transcription factors on the regulation of the mammalian cyclin B1 promoter is controversial. In one study mammalian

E2F2 was shown to be a repressor of the Cyclin B1 gene in quiescent mouse T cells (Infante et al., 2008), while another study showed that overexpression of E2F2 (by intrathoracic injection of E2F2-adenovirus) induces the expression of Cyclin B1 gene in differentiated mouse cardiomyocytes (Ebelt et al., 2008). Analysis of a region of human *CCNB1* promoter revealed that E2F sites within the promoter can have a positive or negative effect on expression (Zhu et al., 2004). Therefore, more studies are called for to address the molecular mechanisms and the promoter elements by which E2F factors regulate cyclin B1 expression, especially considering that the Rb/E2F pathway is deregulated in many cancers, which may impact cyclin B1 expression and tumor progression.

Drosophila provides a powerful system for Rb/E2F studies. Drosophila contains two pocket proteins, Rbf1 and Rbf2, and two E2F proteins, E2f1, which plays an activator role and E2f2, which is associated with repressor functions. Our Chip-seq from Drosophila embryos showed that the cyclin B *(CycB)* gene is bound by both Rbf1 and Rbf2, implicating these proteins in a direct regulation of the promoter (Wei et al., 2015). Recently, we showed that the Drosophila *CycB* promoter is potently and preferentially repressed by the Rbf2 protein, in contrast to the lack of effect on the cell-cycle controlled *PCNA* promoter (Mouawad et al., in press). Therefore, to understand the impact of *CycB* promoter elements on regulation by the Rbf/E2f pathway, we characterized the regulatory elements of a minimal *CycB* promoter and tested the impact of Rbf and E2f proteins in Drosophila S2 cells.

Results

Interacting functional modules of the *CycB* promoter

Based on published ChIP-seq data (Wei et al., 2015), we found that Rbf1 and Rbf2 are bound to the promoter proximal region of the *CycB* gene (Figure 4-1A). To identify cis regulatory regions important for expression of this gene, we created six luciferase reporter constructs extending ~1kbp 5' and 3' of the transcriptional start site (TSS). We included a portion of the gene containing the first large intron, because transcription factor binding sites are often located 3' of the TSS (Figure 4-1A, B). Using a motif identification algorithm, we identified putative E2F sites within this region; sites of higher and lower affinity are present in 5' and 3' regions of the TSS (Figure 4-1B). All constructs were assayed in transfected Drosophila S2 cells, and luciferase activity was assessed in parallel with a control *PCNA* reporter gene.

Successive deletions from the 5' end of these promoter constructs uncovered repressor activity in the most distal portion of the promoter, with expression increasing approximately five-fold as regions from -794 to -118 were removed (Figure 4-1C). Removal of the promoter sequences from -118 to -53 led to a significant loss of activity, indicating that activators interact with this portion of the gene. To determine the impact of 3' sequences, we modified the gene containing the activation region between -118 and -53, and removed all of the 3' sequences outside of the core promoter region, that is +100 to +965. This reporter, -118 to +100, exhibited activity very similar to -118 to +965 (Figure 4-1C). To assess the entire region that includes binding by Rbf1 and Rbf2, we also tested -464 to +100. Strikingly, this construct exhibited activity more than twenty-fold higher than the other constructs. The activity of this reporter construct was very similar to that of the previously studied *PCNA* promoter (Mouawad et al., in press). This result indicates that there

are additional activators present in the region -118 to -464 that are functional in a context-dependent manner, when +100 to +965 is absent. Significantly, the similar activities of -118 to +965 and -118 to +100 indicate that the negative effects of the 3' region is not a simple post-transcriptional one, such as translational inhibition of the luciferase gene. These results show that the *CycB* promoter region contains both activation and repression functions 5' and 3' of the TSS, and that these elements do not function in a strictly additive manner but may interact in a manner similar to regulatory modules found on developmental genes (Yuh et al., 2001; Wei et al., 2016).

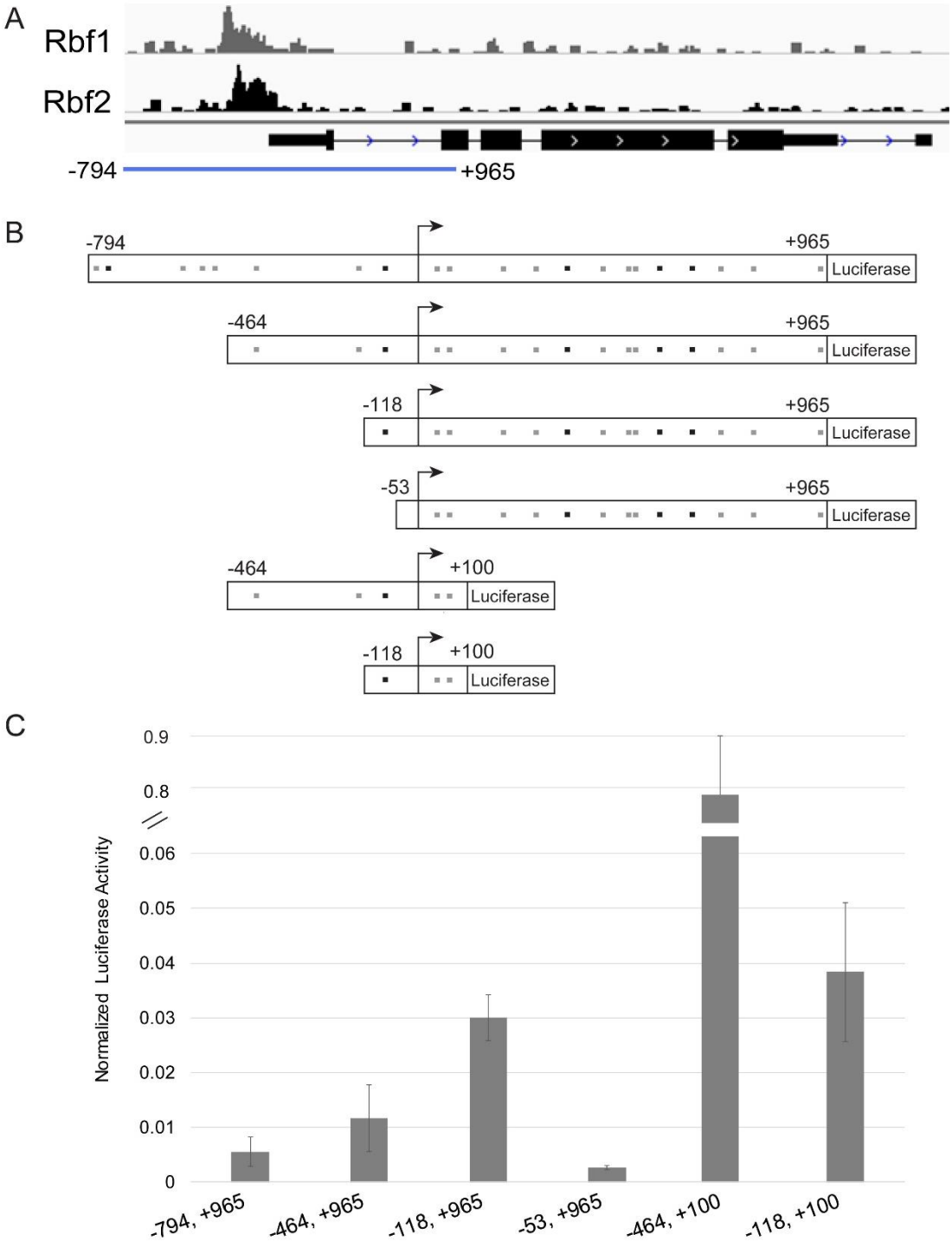


Figure 4-1: Functional modules of the *CycB* promoter. (A) Rbf1 and Rbf2 binding regions on *CycB* promoter based on published ChIP-seq data. The peaks are right upstream of the transcription

Figure 4-1 (cont'd)

start site. The blue line represents the largest piece of the *CycB* promoter cloned and tested. (B) A schematic representation of the six *CycB* luciferase promoter constructs. Black squares represent high affinity E2F sites (P<0.001), and grey squares represent low affinity E2F sites (P<0.005). (C) Normalized luciferase activity of the six *CycB* promoter constructs in S2 cells. Luciferase readings are normalized to *PCNA* reporter done in parallel on the same day. Values represent averages of at least three biologic replicates done on different days, and error bars represent standard deviations. The expression levels of the *CycB* reporters were significantly different from each other (P<0.05) except for -794, + 965 in comparison to -53, +965, and -118, +965 in comparison to -118, +100.

Impact of Rbf and E2f proteins on *CycB* promoter activity

Previous studies indicated that Rbf2 is a weak repressor on the *PCNA* promoter, in contrast to Rbf1 (Stevaux et al., 2002). However, we recently showed that Rbf2 is more potent repressor than Rbf1 on a *CycB* (-464, +100) reporter (Mouawad et al., in press). In order to understand the impact of the promoter regions defined in Figure 4-1 on this differential response to Rbf1 and Rbf2 proteins, we tested the impact of these corepressors on the variant *Cycb* promoters (Figure 4-2A). Rbf2 mediates a robust ten-fold reduction in expression of the most active promoter (-464, +100), while Rbf1 is able to mediate a weaker but still substantial three-fold effect, as previously noted (Mouawad et al., in press). Constructs lacking the activation region between -464 to -118, or constructs in which this activation region is neutralized by the presence of downstream interfering segments (+100 to +965) exhibited a weaker response to Rbf2 expression (Figure 4-2A). In these cases, the effects of Rbf1 or Rbf2 expression were similar; approximately 2-3 fold, with marginally

greater effects of Rbf2 on several constructs. Significantly, the only reporter that had no response to Rbf2 expression was -53, +965, a gene that had less than 1% of the activity of the -464, +100 reporter.

The selective action of retinoblastoma proteins on the *CycB* promoter is likely associated with the interactions with E2f proteins, therefore we assessed the effect of E2f1 and E2f2 expression. We tested two promoters that include the peak of Rbf1 and Rbf2 binding (-464 to -118), as well as two counterparts in which this segment was removed. On all the constructs, E2f1 expression resulted in decrease of activity by 50-70%, regardless of whether the activator-containing -464 to -118 region was present. The repression fold change was similar for highly active and less active constructs. Thus, although E2f1 is traditionally thought of as an activating transcription factor, and indeed stimulates the activity of *PCNA* (Mouawad et al., in press) the effect in this context is uniformly repressive (Figure 4-2B). Interestingly, the activity of E2f2 on the reporter constructs is more variable than E2f1 and is influenced by the 3' region (+100 to +965). In the presence of this element, E2f2 has no significant effect on *CycB* expression. In contrast, E2f2 activated expression from constructs that lacked this 3' inhibitory element. It appears that the +100 to +965 region is occupied with repressors, and removing these repressors allows E2f2 to act as an activator on this reporter, in contrast to its generally described role in repression.

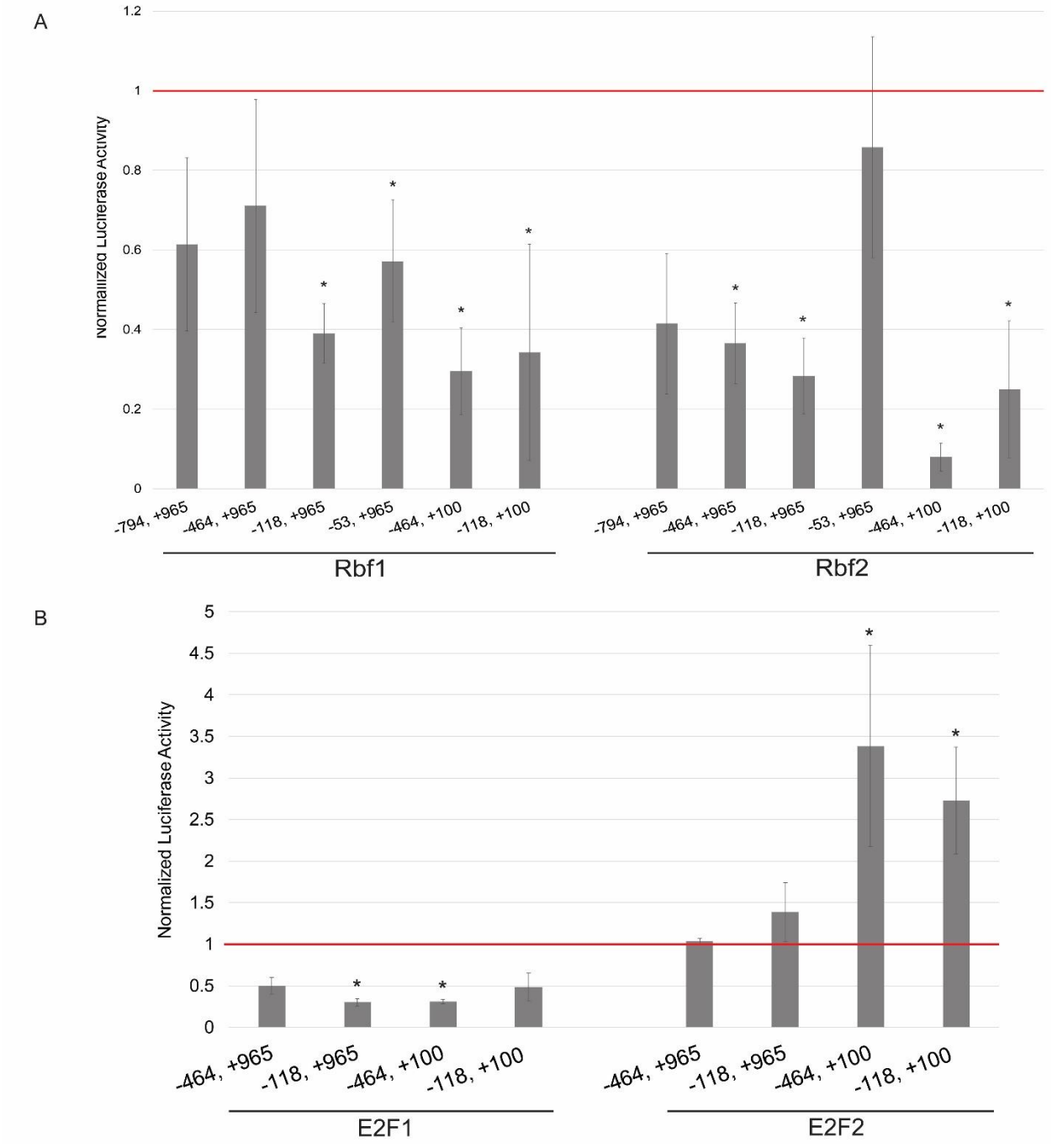


Figure 4-2: Impact of Rbf and E2F proteins on *CycB* promoter activity. (A) Normalized luciferase activity of six *CycB* constructs in response to expression of Rbf1 or Rbf2 in S2 cells. Rbf2 repression activity is most evident on the -464, +100 reporter. Rbf2 repression on the least active reporter, -53, +100 is abolished unlike significant repression by Rbf1. (B) Normalized

Figure 4-2 (cont'd)

luciferase activity of four *CycB* constructs in response to expression of E2F1 or E2F2 in S2 cells. E2F1 results in repression, whereas E2F2 activates two of the *CycB* reporters. For A and B, values represent average of at least three biologic replicates done on different days and normalized to *PCNA* luciferase vector then re-normalized to pAX empty vector. Error bars represent standard deviation. (*) represents a $P<0.05$.

Role of E2F and DREF sites

To test the importance of specific regulatory motifs within the promoter, we mutated the highest predicted affinity E2F site 5' of the TSS, as well as an associated DREF site in the context of *CycB* (-118, +965) (Figure 4-3A). We found that the activity of the reporter decreases significantly when the E2F site is mutated and appears to decrease further when both E2F and DREF sites are mutated, reaching levels similar to the -53, +965 construct (Figure 4-3B). These results indicate that the activity of this reporter is dependent on the high affinity E2F site at -80 bp, and that DREF may also contribute to its activity.

We tested the importance of the E2F and DREF sites on the response of this reporter to Rbf protein expression. Both the wild type and the mutant reporters have similar levels of repression by Rbf1, however, Rbf2 repression is reduced in the absence of the E2F and DREF sites (Figure 4-3C). These results indicate that the E2F and DREF sites are required for full repression of *CycB* by Rbf2 but are not necessary for repression by Rbf1.

The repression of these mutant forms of the -118, +965 promoter by expression of E2f1 was somewhat attenuated, but not abolished, indicating that E2f1 can work through other elements of

this promoter as well (Figure 4-3D). Interestingly, expression of E2f2 resulted in a higher level of induction for the mutant promoters, which themselves have significantly weaker activity than the wild-type promoter. This weak activation may involve interaction between E2f2 and the basal promoter regions, as discussed below.

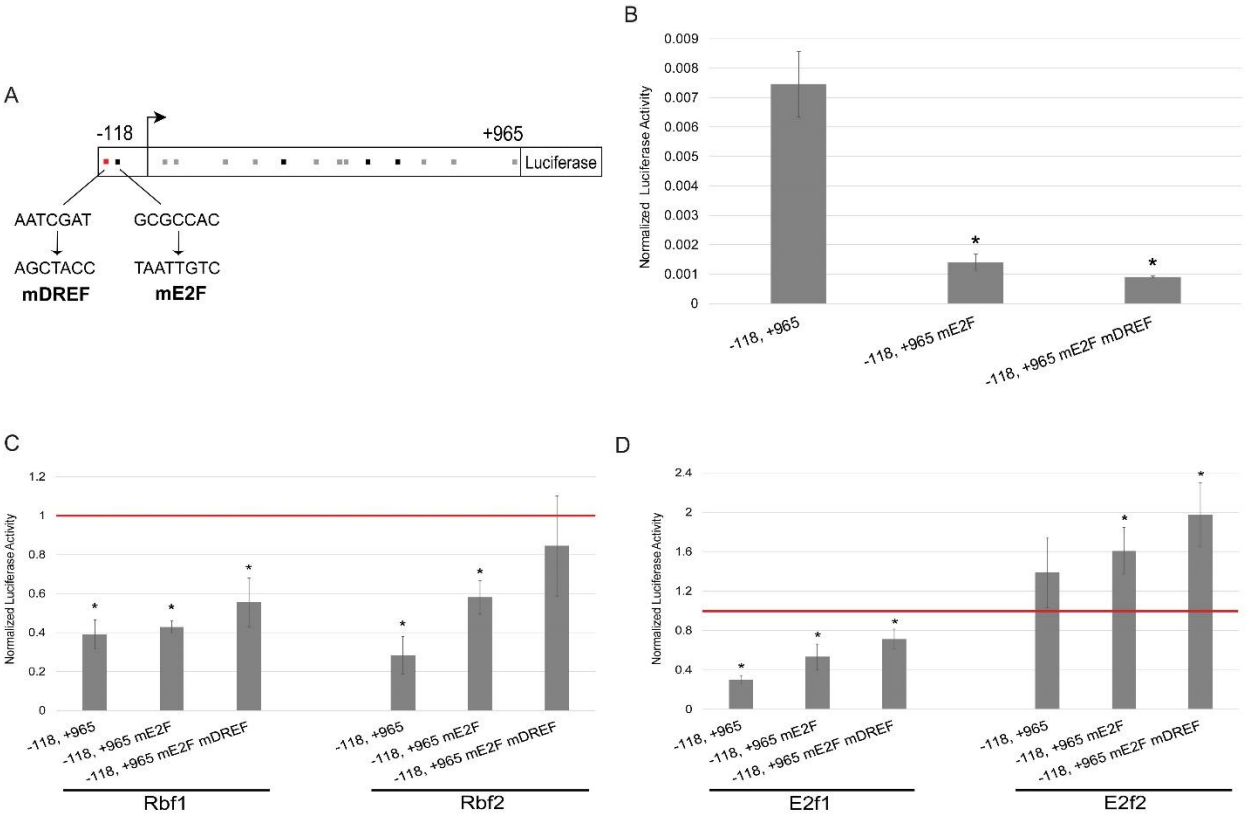


Figure 4-3: Impact of E2F and DREF sites on *CycB* promoter activity and response to Rbf and E2F proteins. (A) Schematic of *CycB* (-118, +965) and the DREF and E2F sites that are mutated. (B) Normalized luciferase activity of the wild type construct, the E2F mutant construct and the E2F and DREF mutant construct. Values are normalized to *PCNA* reporter vector which was transfected in cells in parallel. (C) Normalized luciferase activity of the wild type and mutant constructs in response to Rbf1 and Rbf2. Mutation of E2F and DREF completely abolishes Rbf2

Figure 4-3 (cont'd)

repression, with no or slight effect on Rbf1 repression. (D) Normalized luciferase activity of the wild type and mutant constructs in response to E2F1 and E2F2. E2F1 still represses the mutated reporters, and E2F2 activates them. For C and D values are normalized to *PCNA* reporter and re-normalized to pAX vectors. For B, C and D, values represent average of at least three biologic replicates done on different days, and error bars represent standard deviation. (*) represents a $P<0.05$

Functional redundancy of E2F sites in the *CycB* promoter

The E2F and DREF sites located in the promoter proximal region clearly have roles in repression and activation of the reporter, but it was not clear if these activities are relevant to the high level of expression possible when more 5' activation sequences are present, in the absence of the inhibitory 3' element. Earlier we showed that the core promoter of *CycB* (-53, +100) is essential for Rbf2 to efficiently repress the *CycB* reporter (-464, +100) (Mouawad et al., in press). Therefore, we tested the impact of mutating E2F sites present within this core promoter region on this highly active form of the promoter.

We created two additional constructs: one having a mutation in the -80 bp 5' E2F site (m1E2F), and the other having in addition mutations in the 3' E2F sites (m2E2F and m3E2F) (Figure 4-4A). The activity of the m1E2F construct was very similar to wild type, indicating that the E2F site at -80 bp is not necessary for the activity of this *CycB* reporter. We noted a modest increase in expression of the reporter in which all three E2Fs sites were mutated, which may indicate that these E2F sites can contribute to repression (Figure 4-4B).

We expressed both Rbf and E2f proteins to test the importance of these basal promoter proximal E2F sites in regulation of the gene. We found that Rbf1 and Rbf2 repress all three constructs in a similar fashion, with Rbf2 demonstrating a greater effect, as observed before. These results indicate that Rbf repression on these reporters may be largely mediated by more 5' E2F sites, consistent with measured ChIP occupancy over this region, or that there is some level of redundancy among the E2F sites for repression (Figure 4-4C). E2f1 represses all the constructs, with an increased effect on the triple mutant. E2f2 induction is attenuated in the single E2F mutant, and eliminated in the triple E2F knockout reporter, suggesting that this effect is directly mediated through basal promoter E2F sites.

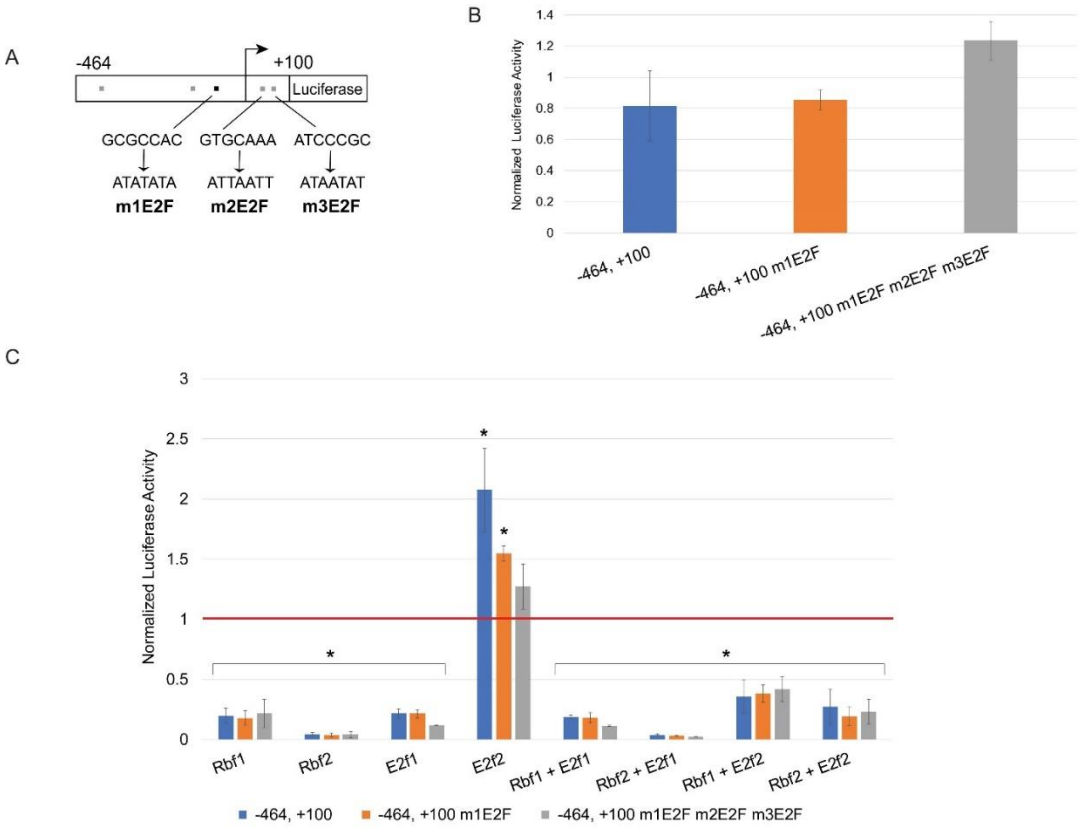


Figure 4-4: Function of core promoter E2F sites on *CycB* promoter activity and response to Rbf and E2F proteins. (A) Schematic of *CycB* (-464, +100) and the E2F sites that are mutated.

Figure 4-4 (cont'd)

(B) Normalized luciferase activity of the wild type CycB construct, the one E2F mutant (m1E2F) construct and three E2Fs mutant (m1E2F, m2E2F and m3E2F) construct. Values are normalized to *PCNA* reporter vector done in parallel. (C) Normalized luciferase activity of the wild type and mutant constructs in response to transfection with Rbf1, Rbf2, E2F1, E2F2, Rbf1 + E2F1, Rbf2 + E2F1, Rbf1 + E2F2 and Rbf2 + E2F2. Values are normalized to *PCNA* reporter and re-normalized to pAX vectors. For B and C, values represent average of at least three biologic replicates done on different days, and error bars represent standard deviation. (*) represents a $P<0.05$.

Conservation of *CycB* promoter and first intronic region in Drosophila

The distinct functional properties of the promoter sequences suggest that these may be have been selected to endow proper dynamic activity on *CycB.* To ascertain whether these sequences show evidence of conservation, we aligned *CycB* promoter segments from Drosophila species of different phylogenetic distances (Figure 4-5). We observe variable levels of conservation in the promoter regions that we analyzed including blocks of sequences 5' of the TSS as well as the coding region, as expected. Significantly, the first intron shows blocks of highly conserved sequences which include several of the predicted E2F sites.

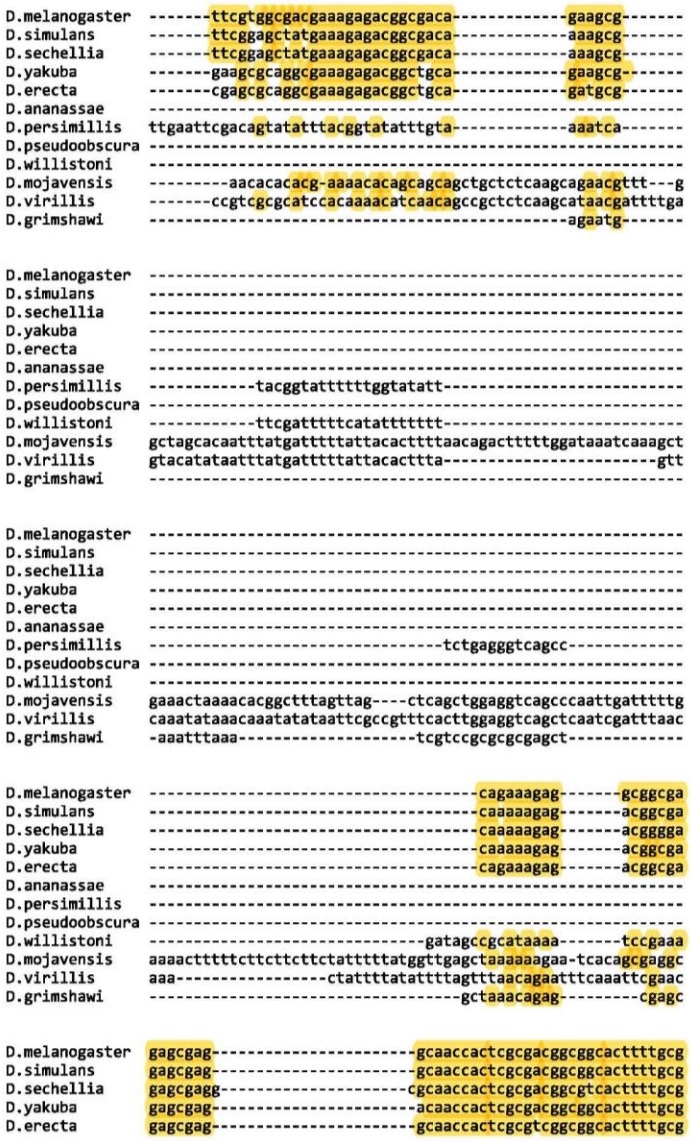


Figure 4-5: Conservation of *CycB* promoter regions and first intron within Drosophila species. Multiple sequence alignment for *CycB* promoter and 3' regions in Drosophila species (-1000 to +1000 bp). Yellow highlighted bases represent conservation with respect to *D. melanogaster*. Orange lines represent putative high affinity E2F motifs (P<0.001), and green lines represent lower affinity E2F motifs (P<0.005). The TSS is denoted by +1. The regions of the promoter segments that were analyzed are denoted as -794, -464, -118, -53, +100 and +965. Uppercase letters indicate the open reading frame.

Figure 4-5 (cont'd)

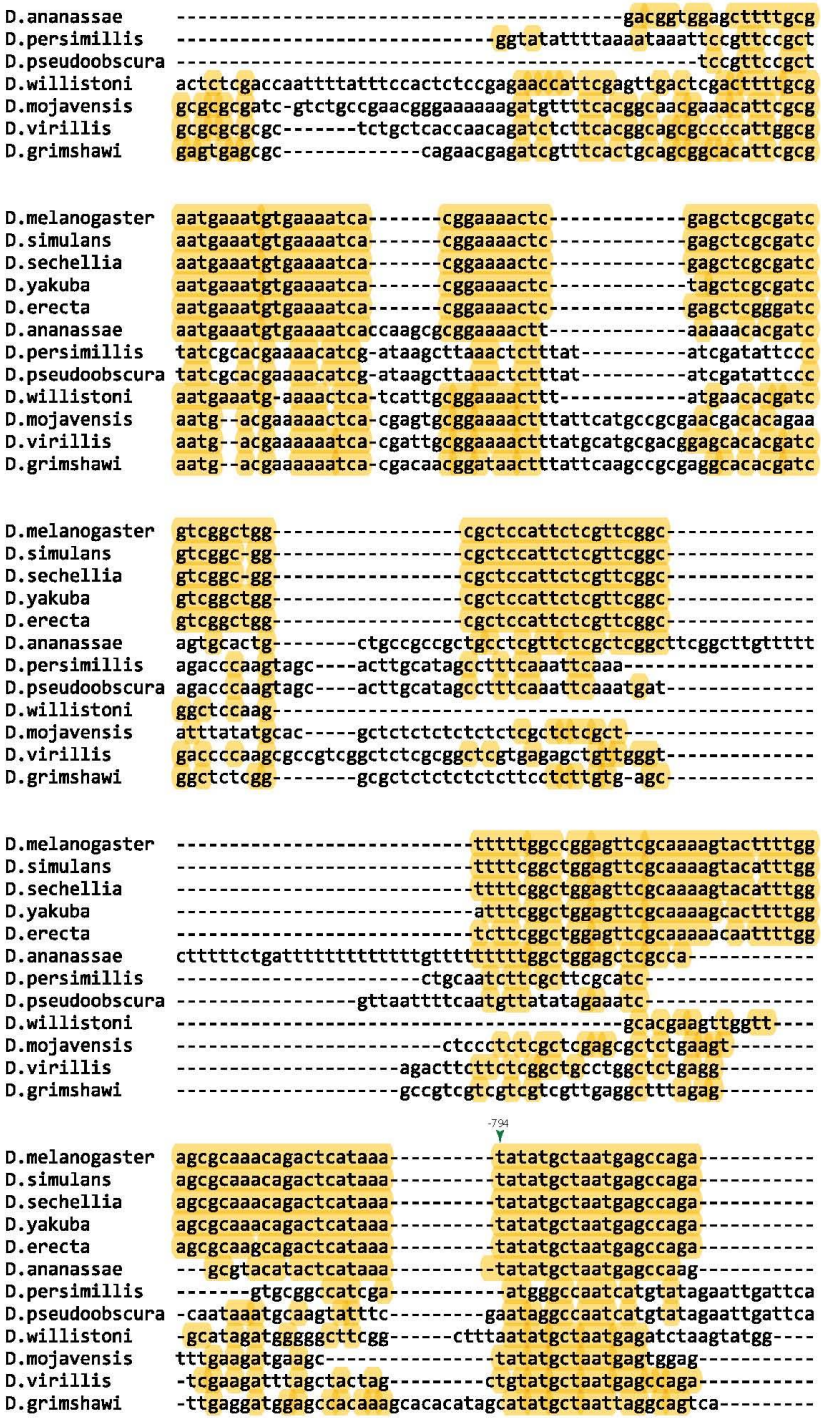


Figure 4-5 (cont'd)

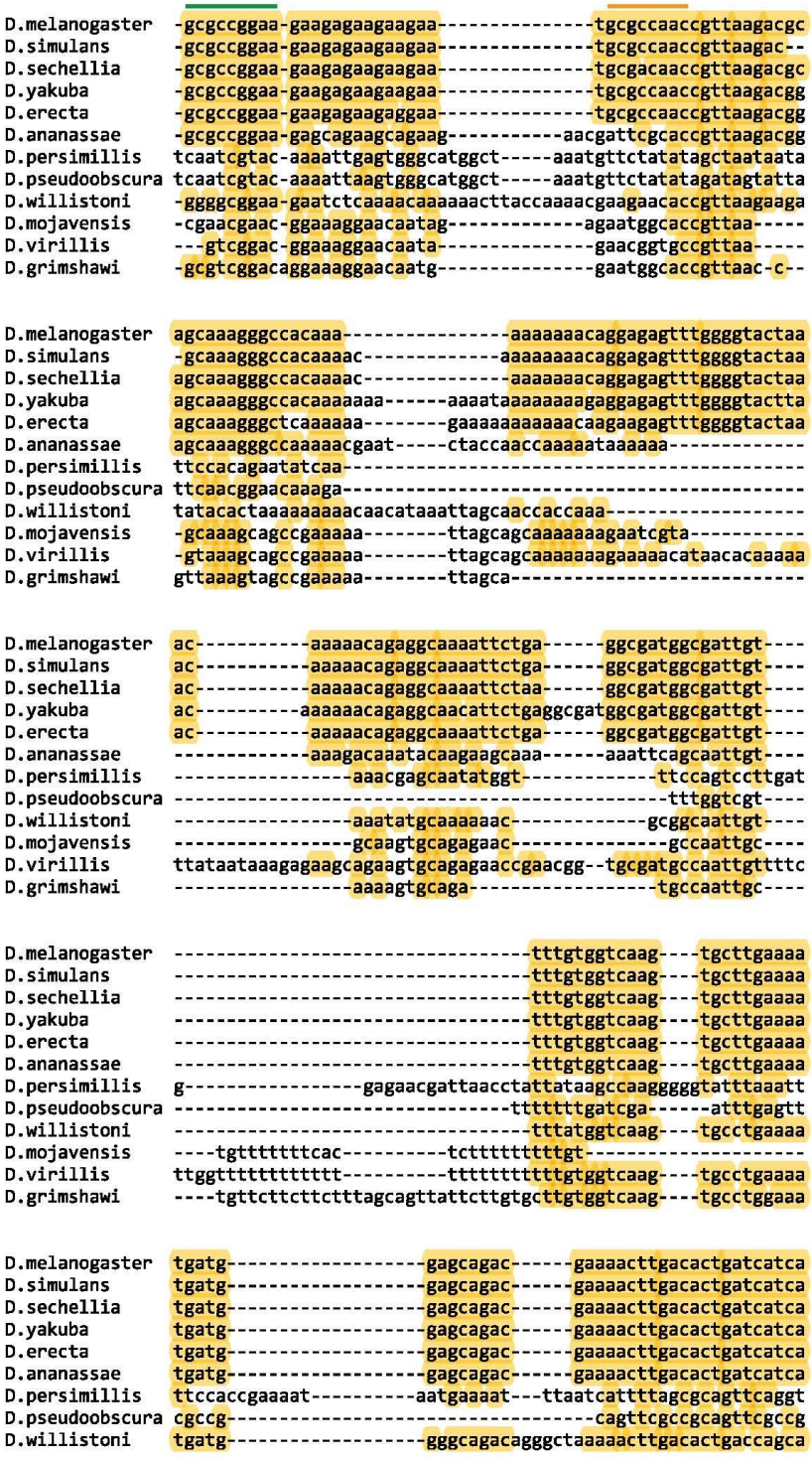


Figure 4-5 (cont'd)

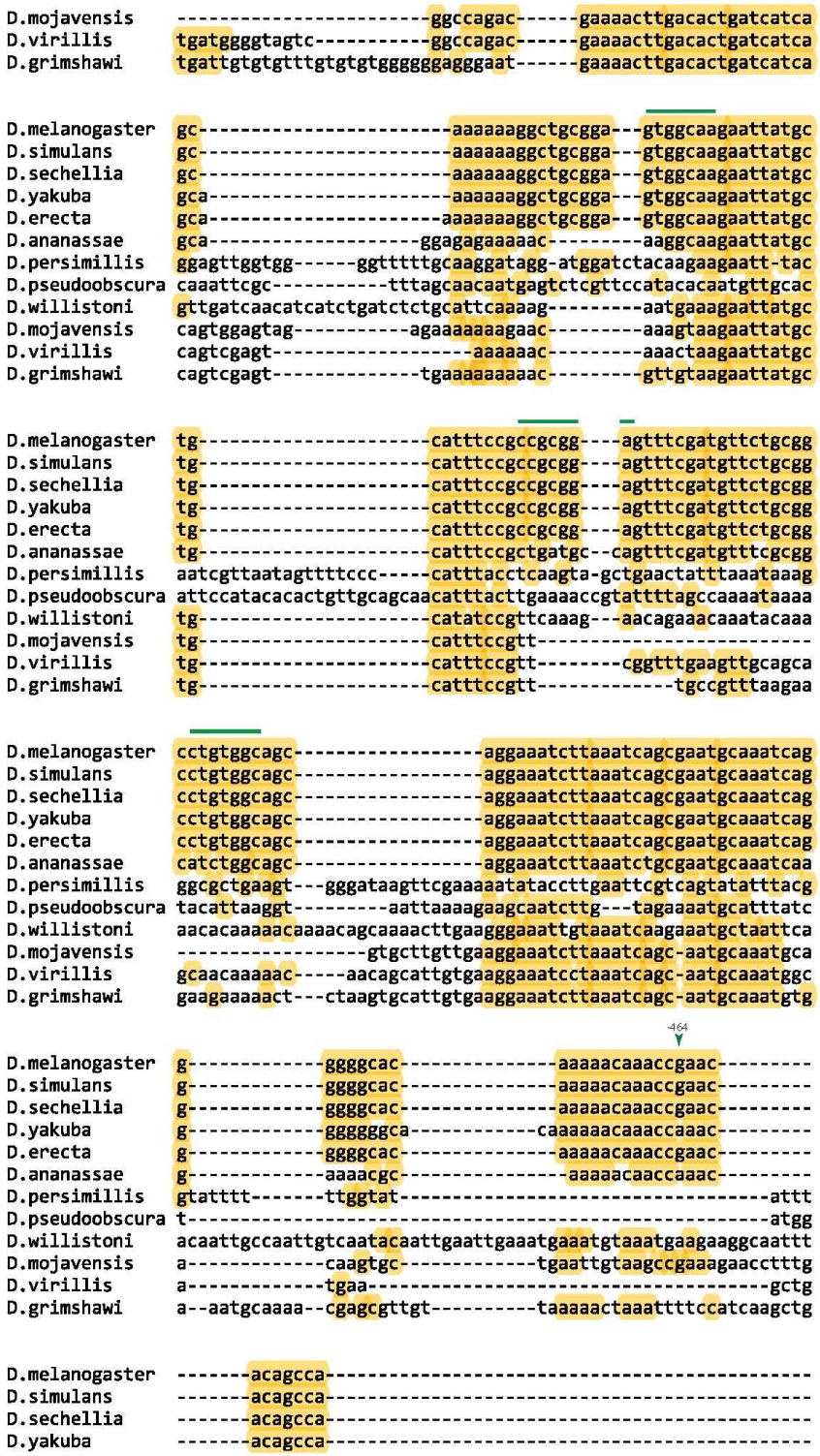


Figure 4-5 (cont'd)

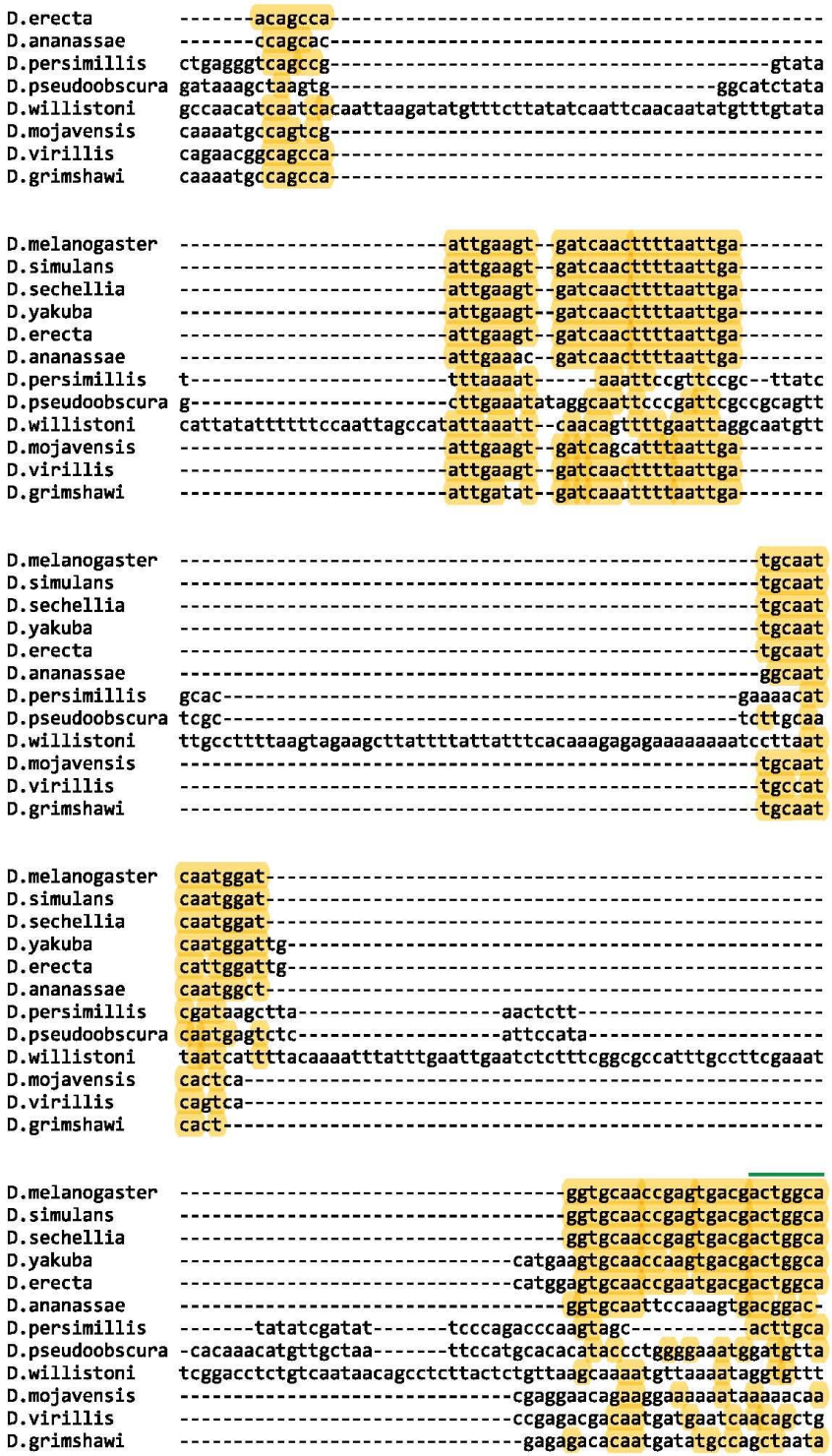


Figure 4-5 (cont'd)

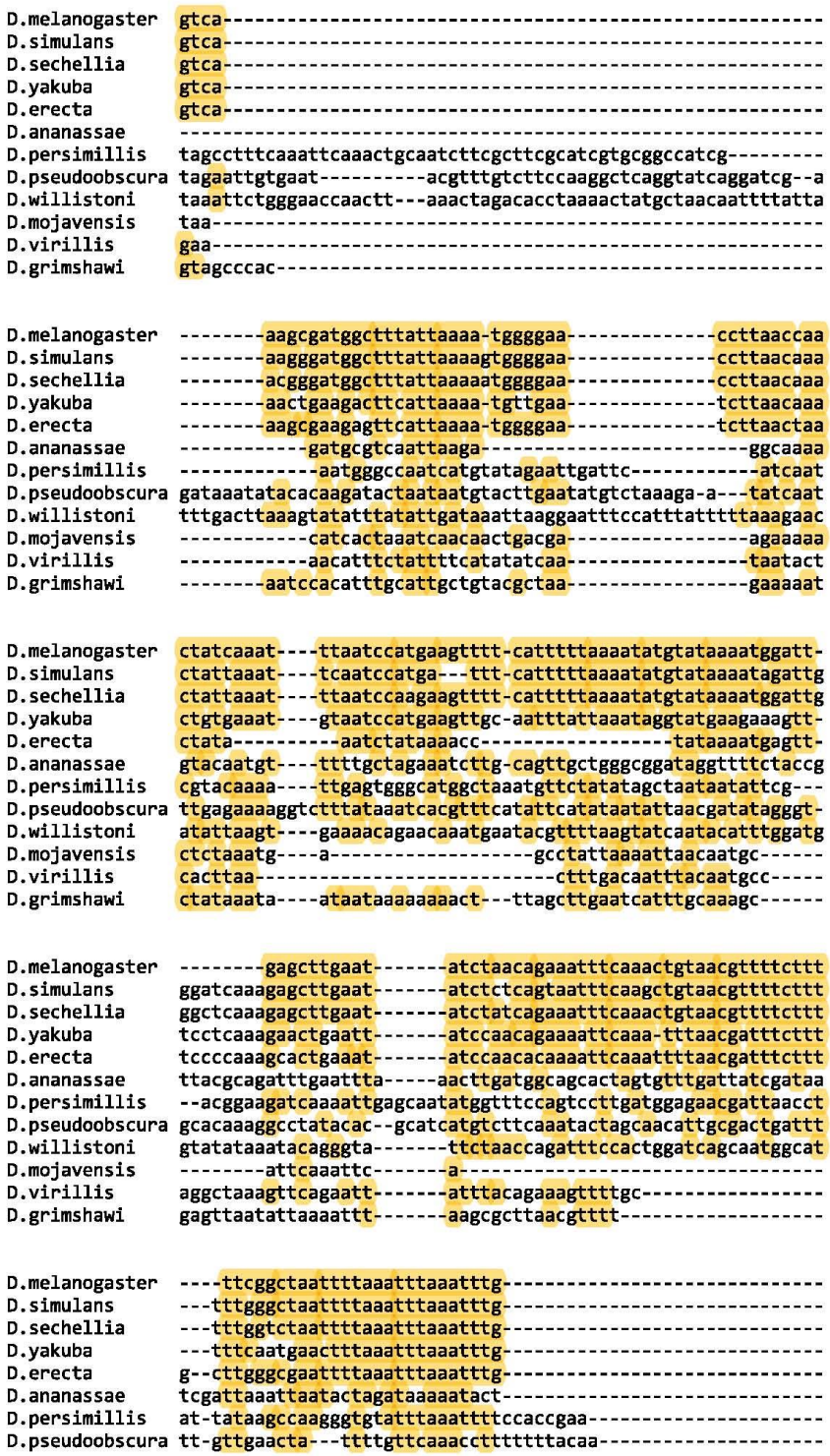


Figure 4-5 (cont'd)

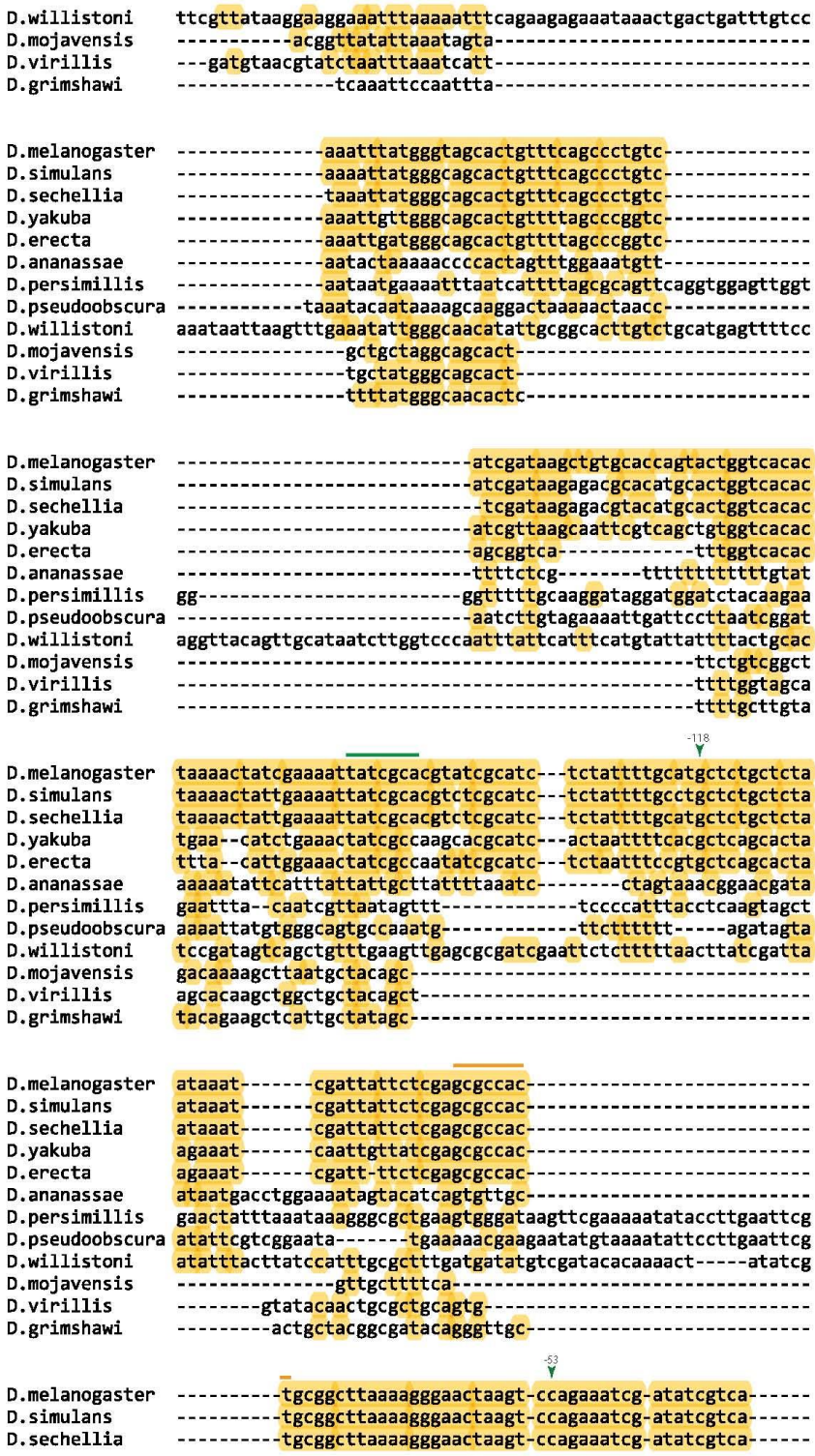


Figure 4-5 (cont'd)

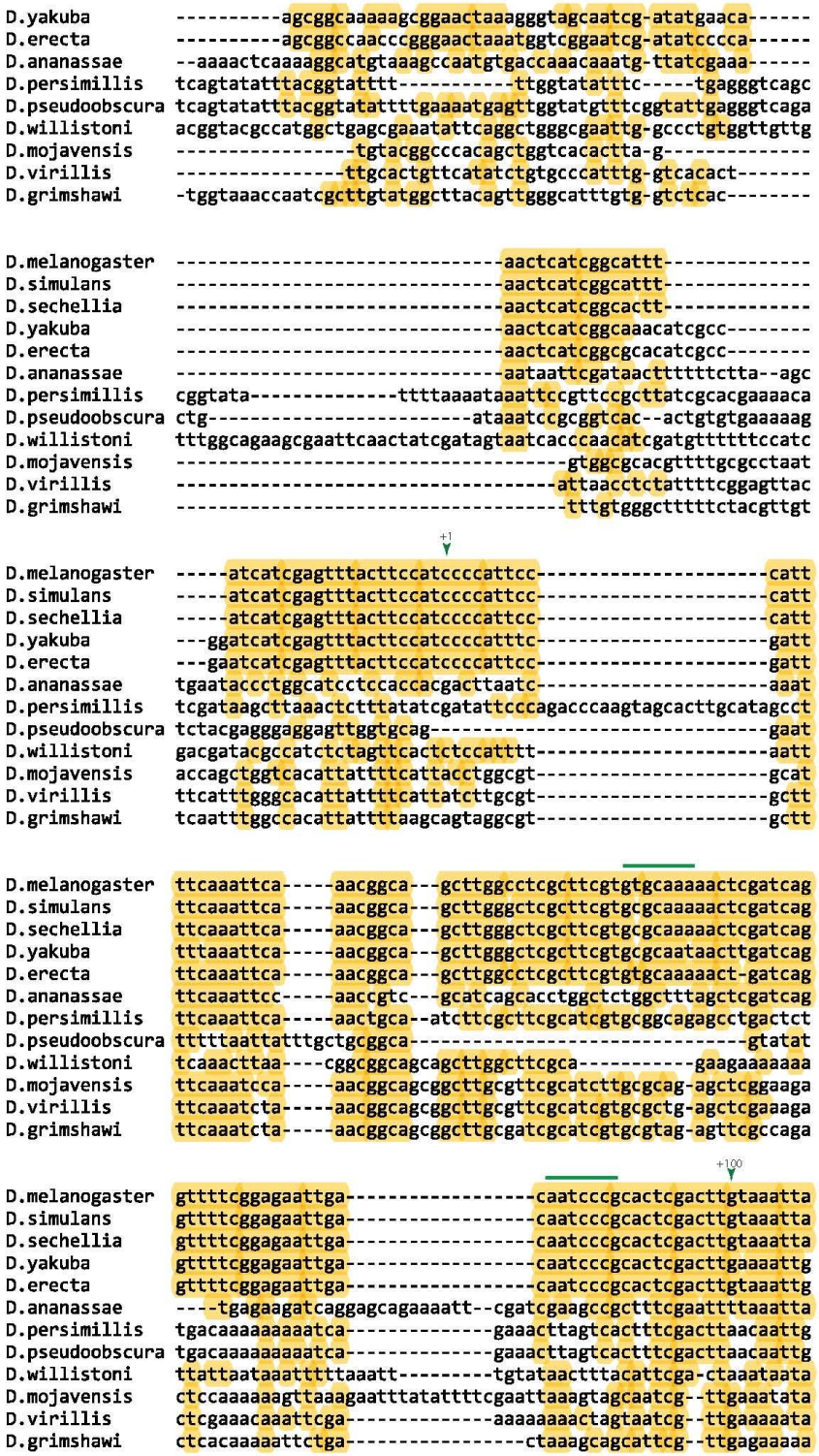


Figure 4-5 (cont'd)

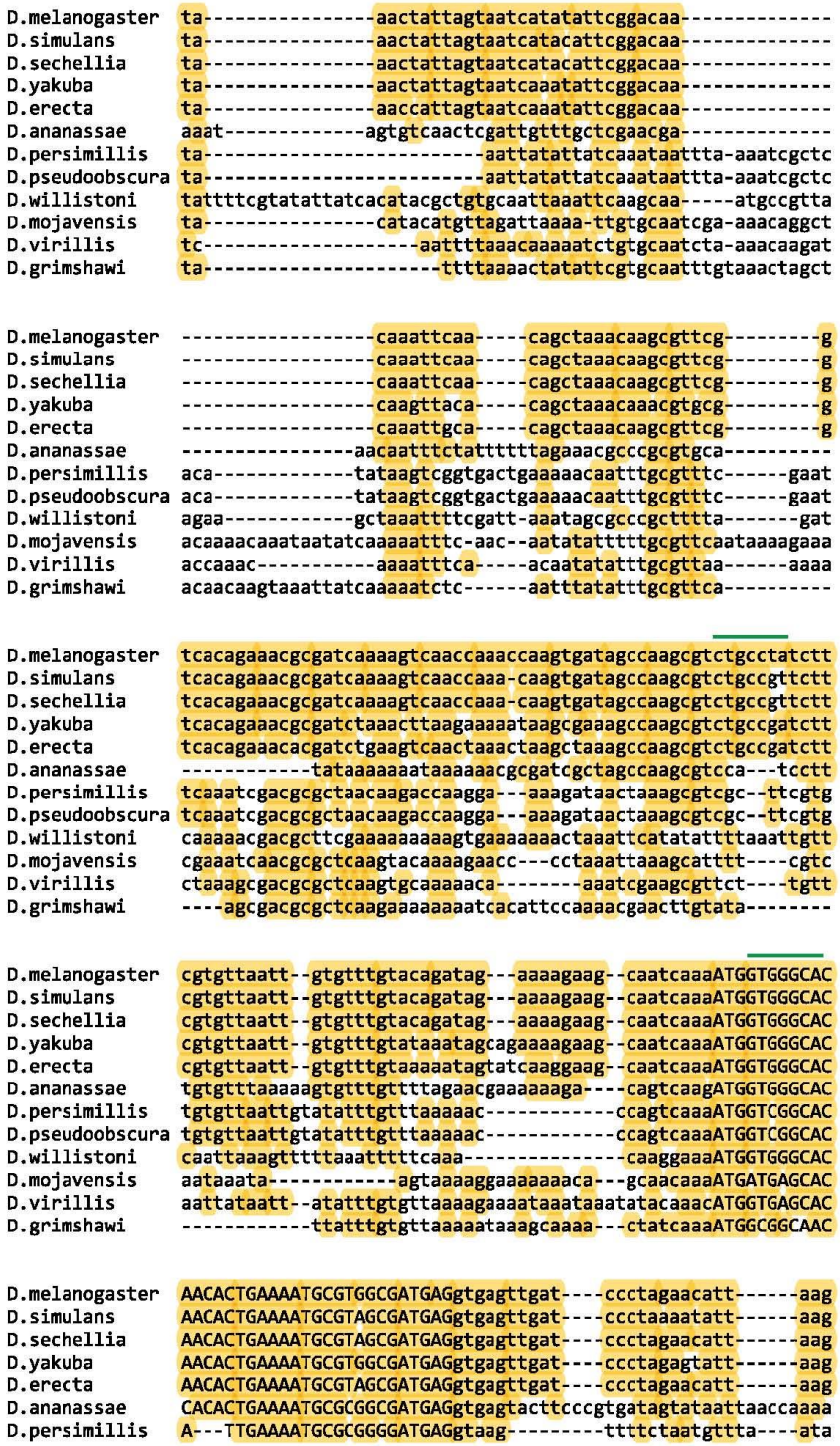


Figure 4-5 (cont'd)

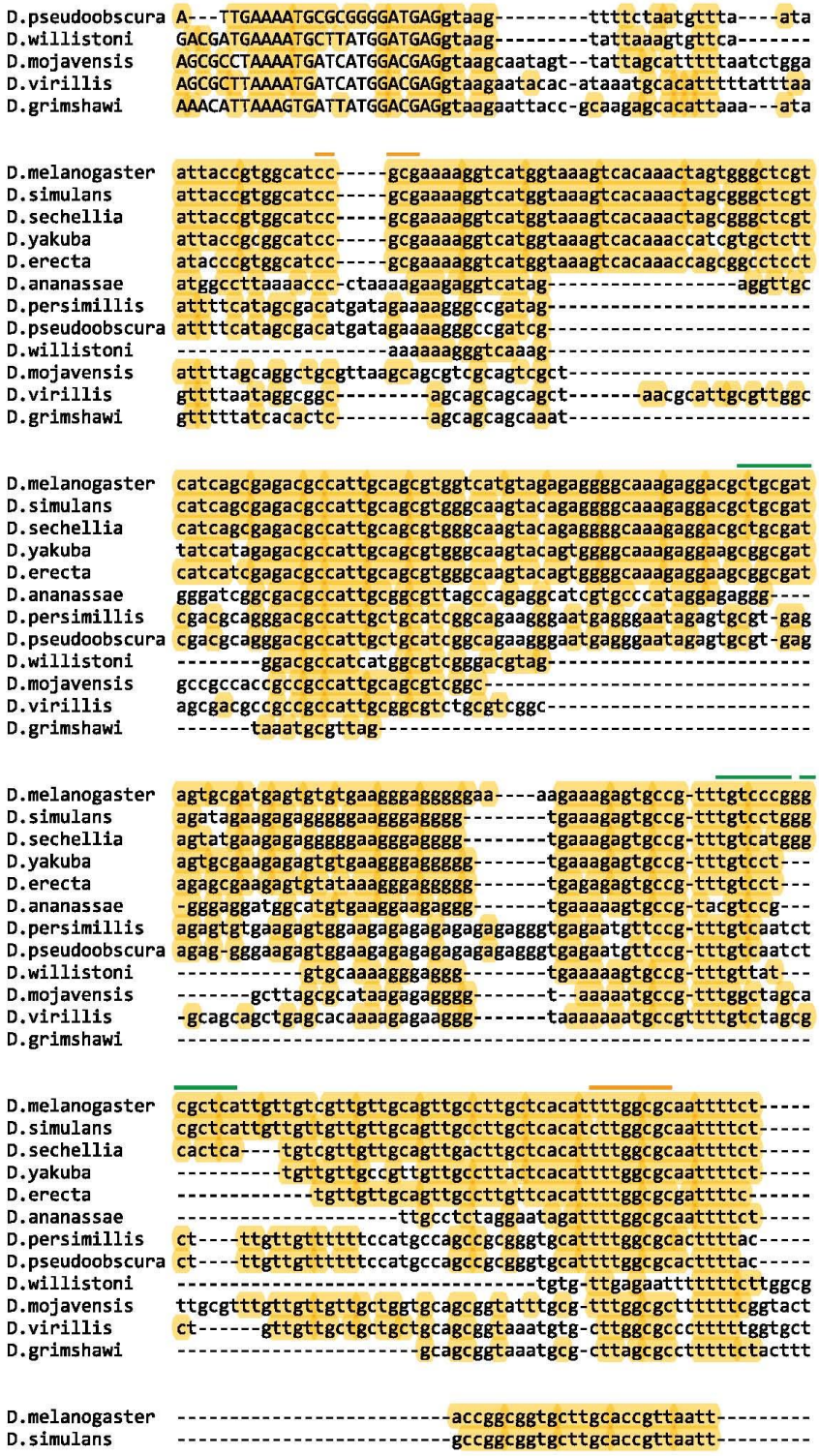


Figure 4-5 (cont'd)

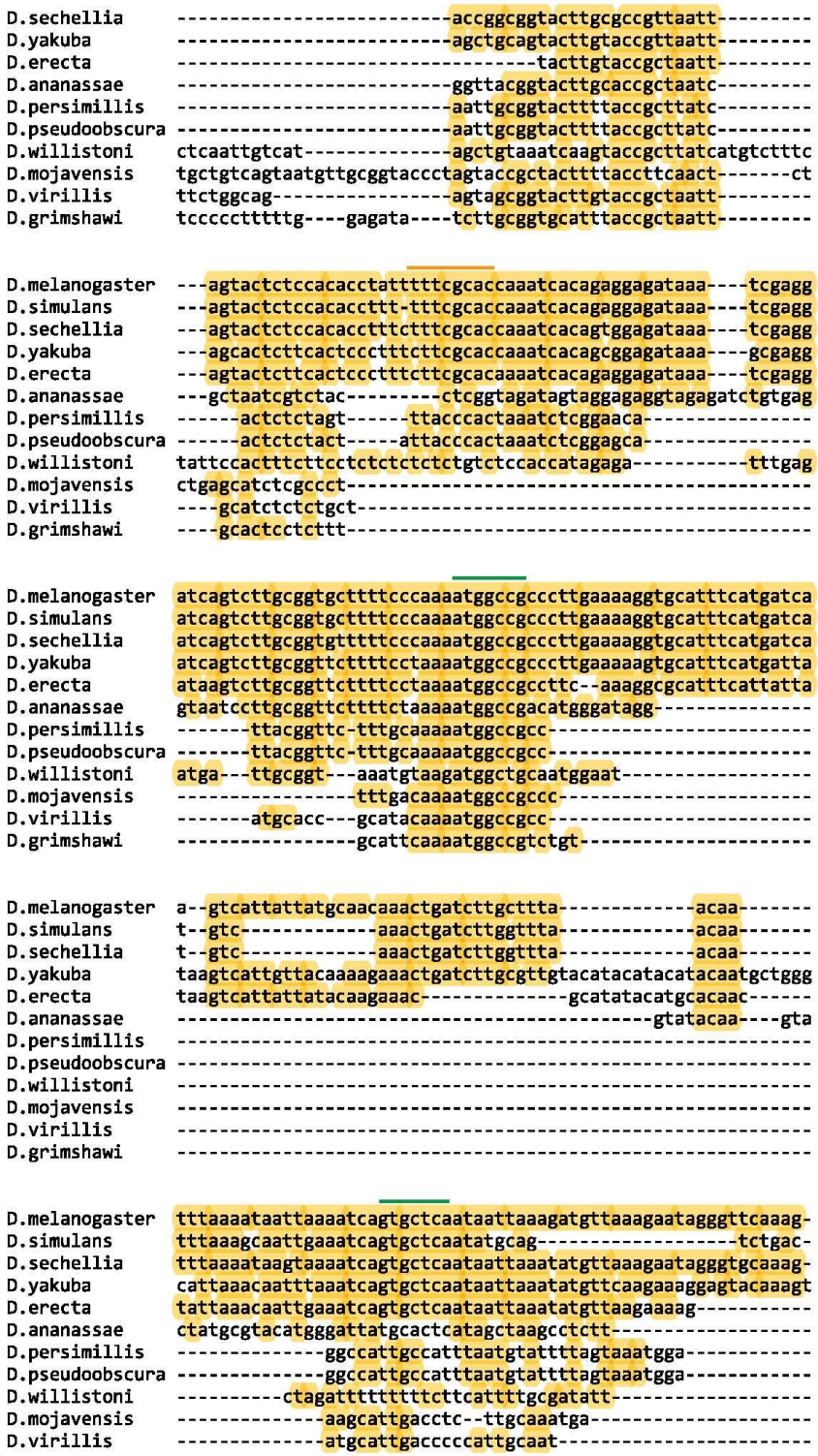


Figure 4-5 (cont'd)

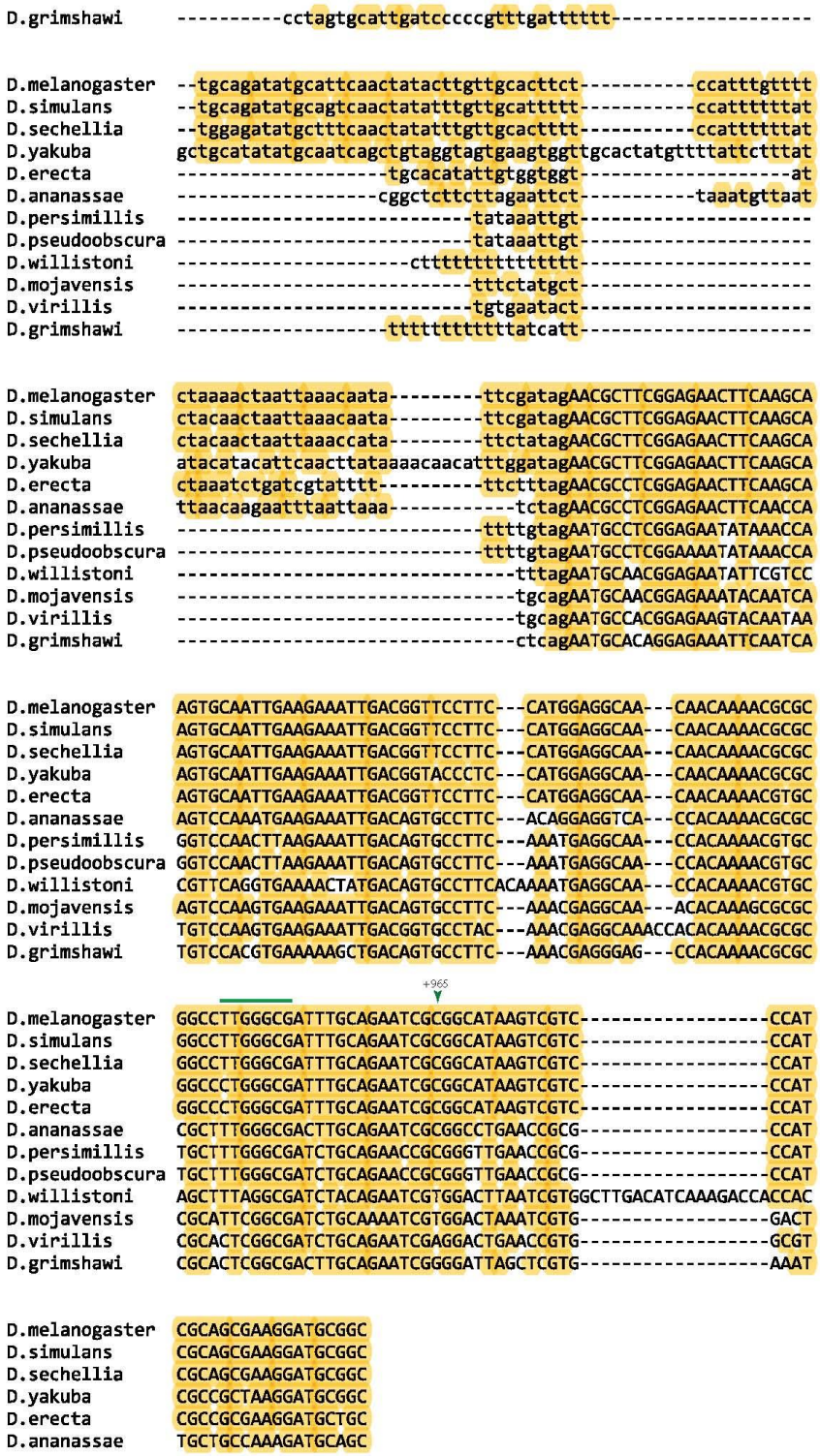
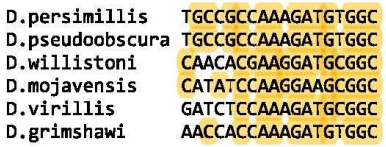


Figure 4-5 (cont'd)



Discussion

Our analysis of the *CycB* promoter reveals several unexpected properties of this cis regulatory region. Unlike the more compact design of the G1/S specific *PCNA* gene, which appears to be largely dependent on two E2f sites found 5' of the TSS (Thacker et al., 2003), sequences found both 5' and 3' of the promoter appear to interact to influence *CycB* activity (Figure 4-6A). The dependence on 5' activating sequences for 3' repression to be manifested is reminiscent of other well-studied developmental promoters such as the *endo16* promoter from *S. pupuratus,* in which specific modules combine their outputs to generate non-linear responses in developmental time (Yuh et al., 2001). *CycB* is likely not the only cell cycle gene that uses larger promoter regions to generate proper regulatory patterns; genes such as *cyclin E* were found to be associated with distal open chromatin regions that are closed as cells of the wing disc enter terminal differentiation. This promoter can be stimulated to enter one or two ectopic cell cycles by misexpression of E2F1, but then are no more responsive, presumably due to loss of input from these distal regions (Ma et al., in press).

The consistent repressive effect of E2f1 expression on different versions of this promoter, and activation by E2f2, is especially intriguing. A previous survey of endogenous genes responsive to depletion of E2f1 or E2f2 showed that a number of genes silent in cycling S2 cells are upregulated upon loss of E2f2, consistent with its previously assigned role in repression (Dimova et al., 2003). However, E2f2 proteins contain a domain that is similar to the previously characterized transcriptional activation domain of E2f1 proteins (Sawado et al., 1998), thus if there are specific genes on which E2f2 can bind in the absence of associated retinoblastoma proteins, this latent activity may be revealed. Similarly, depletion of E2f1 protein from S2 cells generally is associated with the reduction of activity of cell-cycle active genes, indicating a role for activation (Dimova

et al., 2003). However, binding of retinoblastoma proteins to E2f1 inhibits their function, thus if *CycB* presents a specific cis-regulatory context in which E2f proteins preferentially associate without (E2f2) or with (E2f1) complexed Rbf corepressor proteins, novel regulatory outputs may be the result (Figure 4-6B). The specific nature of the cis regulatory context that would specify such "role reversals" for E2f1 and E2f2 remains to be identified. The lack of E2f2 activation for constructs in which the 3' repressive region is present suggests that in the context of the endogenous gene, where these sequences are naturally present, E2f2 may not exhibit this activity, although it is not known if regulation would be different in quiescent cells.

The effects of Rbf1 and Rbf2 overexpression are also revealing. By assessing different forms of this promoter, containing different regulatory regions, we were able to identify contexts in which the effect of Rbf2 are significant, and sometimes greater than that of Rbf1. Notably, the region important for high level activity of this gene (-464 to -118) was not essential for Rbf2 regulation, although this portion coincides with the center of the ChIP peak. We found that the Rbf2 is able to mediate effective repression of constructs containing the promoter-proximal area, which is consistent with our previous finding that the core promoter region of *CycB* appears to be important for mediating robust repression by Rbf2. Distinct basal promoter architecture has been shown to affect responsiveness to enhancer activation, possibly because the core transcriptional machinery experiences different rate-limiting allosteric changes in formation and release of the pre-initiation complex (Arnold et al., 2017). We have previously suggested that Rbf2 may have evolved to generate a different biochemical response compared to that of Rbf1 on genes that are common targets (Mouawad et al., in press). This Rbf1/2 distinction may include relative stability of complex formation or actual transcriptional regulatory effect. These results indicate that properties of the basal promoter may play a key role in enabling this differential function (Figure 4-6C). We did not

comprehensively explore the roles of all E2f-like sequences in promoter proximal and distal areas, thus it is unclear whether Rbf2 functions are entirely mediated through E2f2 binding. It is notable that human Rb protein has been described to directly target basal machinery of RNA polymerase III genes (Gjidoda and Henry, 2013).

An additional aspect of Rbf2 regulation is the observation that when this corepressor is overexpressed in 12-18 hour embryos, a point in development where many cell types have entered into terminal differentiation, a number of lowly expressed genes, including *CycB*, are actually activated (Mouawad et al., in press). We propose that stable long-term repression may involve formation of Rbf1 complexes on these promoters, where transient replacement with Rbf2 would disrupt existing complexes and allow a burst of expression. The differences in protein complex stability and actual Rbf corepressor repression mechanisms in continuously cycling cells vs. differentiated cells remains to be fully explored.

Our study of *CycB* regulation includes 3' intronic sequences. Considering the phylogenetic conservation of the gene, it is apparent that certain promoter-proximal as well as intronic elements are highly conserved, suggesting that these regions may be important for proper expression of the gene. It is significant that the study of Rb family proteins on the human *CCNB1* promoter has been focused exclusively on basal promoter and 5' regions (Zhu et al., 2004), although there is a high level of sequence conservation within the first intron in mammalian *CCNB1* genes (Figure 4-7). As foreshadowed by the presence of developmentally active distal enhancer-like sequences for certain Drosophila cell cycle genes (Ma et al., in press), the complete understanding of these genes will require a more comprehensive consideration of relevant cis elements, which may reveal novel properties for regulation by E2f and Rb family members.

The complex and specific interactions between elements of the *CycB* promoter and Drosophila retinoblastoma proteins doubtlessly reflect a strong selection to yield proper cell cycle regulation. It is interesting that the *Rbf2* gene has evolved significantly from the ancestral retinoblastoma gene, with loss of several conserved features, including the instability element in the C-terminal region of the protein (Mouawad et al., in press). The unique functional responses of *CycB* take advantage of the distinct regulatory potentials of the Drosophila retinoblastoma proteins; similar functional differentiation is likely the case in vertebrates, where a parallel evolutionary divergence of retinoblastoma proteins has occurred.

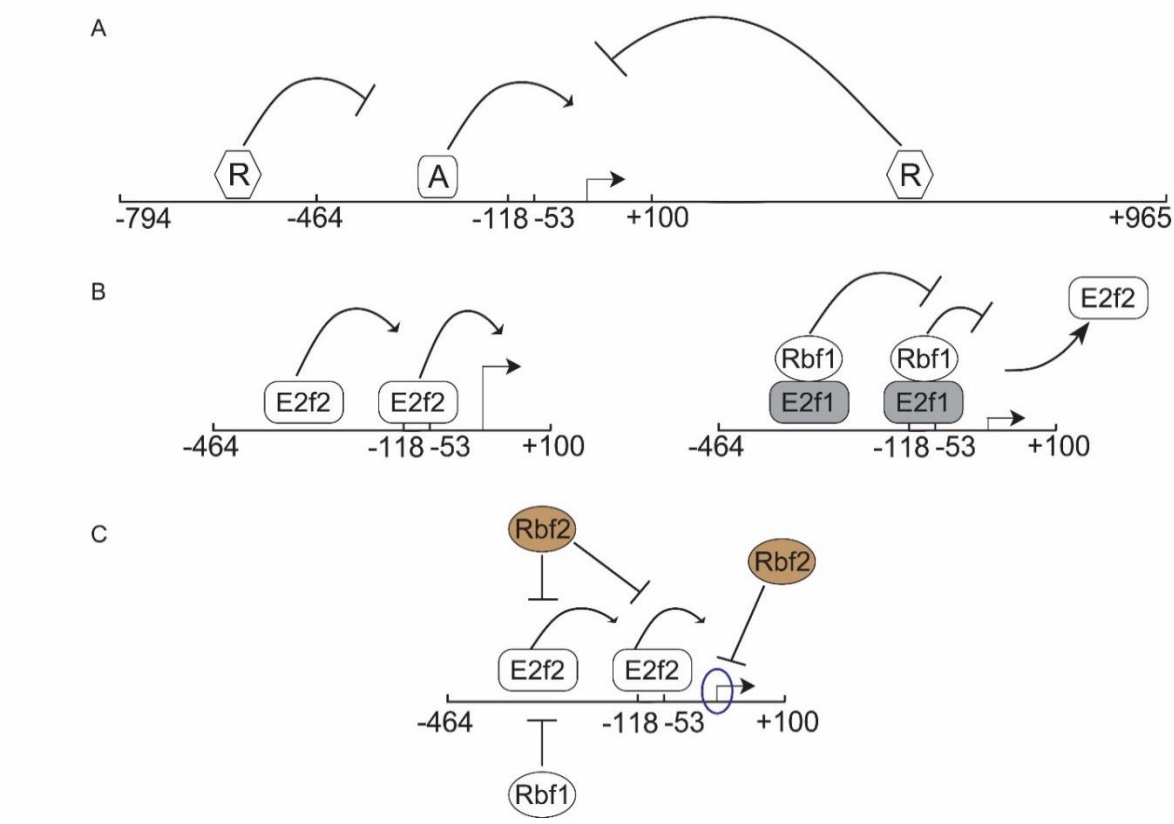


Figure 4-6: Specific *CycB* promoter elements impact gene activity and transcriptional response to Rbf and E2f factors. (A) Potential repressors in the 3' region of the promoter interfere with activators on the 5' -118 to-464 region. Additional repressors present in the 5' end region also

Figure 4-6 (cont'd)

inhibit the activation of the gene in a different mechanism. (B) E2f2 is an activator of *CycB*, and overexpression of E2f1 antagonizes E2f2 and binds to the promoter as a complex with Rbf1 leading to repression. (C) Rbf2 potently represses *CycB* through interacting with E2f2 at the -53 to -118, and -118 to -464 regions, and independently inhibiting factors at the basal promoter. Rbf1 primarily inhibits the gene through interacting with the -118 to -464 region only.

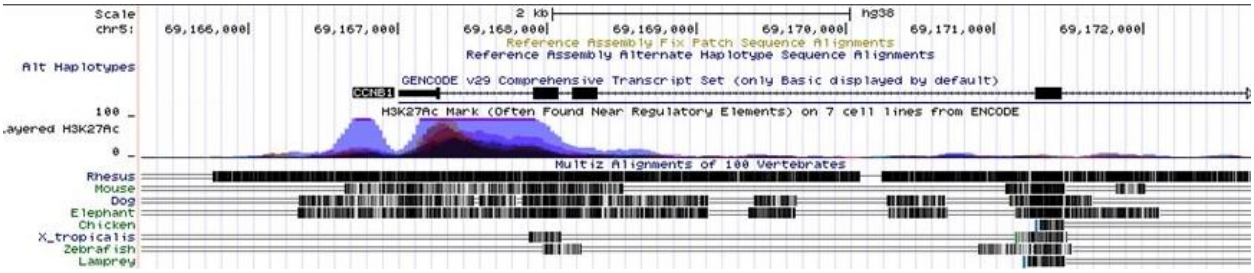


Figure 4-7: Conservation of promoter region and first intron of Cyclin B1 gene among mammals. Schematic representation of the *CCNB1* gene taken from the UCSC genome browser. Blue peaks indicate the H3K27Ac chromatin mark. Cyclin B1 genes from the following organisms are shown: Rhesus, mouse, dog, elephant, chicken, frog, zebrafish and lamprey. The conservation is indicated by black bars. Notably, the first intron is highly conserved in mammals including rhesus, mouse, dog and elephant.

Materials and methods

Reporter constructs

The C*ycB* promoter regions were cloned into *Asc*I and *Sal*I sites in the pAC2T-luciferase vector (Acharya et al., 2010). The The *CycB* (-464, +100) m1E2F mutant construct, and the CycB (-118, +965) E2F and DREF mutant constructs were created using site-directed mutagenesis. *CycB* (-464, +100) triple m1E2F, m2E2F and m3E2F mutant construct was synthesized as Gblock gene fragments by IDT (IDTDNA.com) and cloned into the *pAC2T-luciferase* vector using *Asc*I and *Sal*I sites. The *PCNA*-luciferase reporter (a gift from the Nick Dyson laboratory) was previously described (Acharya et al., 2010; Yamaguchi et al., 1995). The *pIE-E2f1* and *pIE-E2f2* vectors were a gift of the Maxim Frolov laboratory (Frolov et al., 2001). The pAX-*Rbf1* and pAX-*Rbf2* were previously described (Acharya et al., 2010).

Luciferase reporter assays

Drosophila SL2 cells were cultured in Schneider's medium (Gibco) supplied with 10% HI-FBS and penicillin-streptomycin (100 units/mL penicillin and 100 µg/mL streptomycin, Gibco). 1.5 million cells were transfected using Effectene transfection reagent (Qiagen) with 250 ng of each *CycB* reporter vector. An equal amount of *PCNA*-luciferase reporter vector was transfected in separate wells on the same day. Expression of *CycB* reporters were normalized to *PCNA* reporter. Transfection of *CycB* reporters along with pAX-*Rbf1*, pAX-*Rbf2*, *pIE-E2f1*, *pIE-E2f2* or pAX control vector was done similarly. Co-transfection with 250ng of *pIE-E2f1* or *pIE-E2f2* along with *pAX-Rbf1* or *pAX-Rbf2* was also performed, compared to equal amount (500ng) of *pAX* vector as control. Luciferase values were normalized to *PCNA* reporter and then re-normalized to pAX

control vector. Cells were harvested 72 hours post-transfection, and luciferase assays were conducted as described previously (Wei et al., 2015; Acharya et al., 2010).

Motif search

The E2F and DREF motifs, described previously (Acharya et al., 2012), were identified using MAST (MEME-suite v.5.0.5) using P<0.001 and P<0.005 cutoffs.

Multiple sequence alignments

D. melanogaster CycB (-1000 to +1000) sequence was acquired from FlyBase. Sequences from other 11 Drosophila species (*D. simulans, D. sechellia, D. yakuba, D.erecta, D. anannassae, D. persimillis, D. pseudoobscura, D. mojavensis, D. virillis,* and *D. grimshawi*) were retrieved from BLAST tool from NCBI. Multiple sequence alignment was performed using Multiple Alignment using Fast Fourier Transform (MAFFT v.7.397) software by European Bioinformatics Institute (EBI). Conserved residues with respect to *D. melanogaster* were highlighted in yellow shade.

Acknowledgments

We thank R.W. Henry, S. Payankaulam and members of the Arnosti lab for technical assistance and advice. We thank Nick Dyson for sharing the *PCNA* luciferase reporter gene and Maxim Frolov for providing the pIE-Myc-*E2F1* and pIE-Myc-*E2F2* constructs. This research was supported by NIH grant GM124137 to DNA.

REFERENCES

REFERENCES

Acharya, P., Negre, N., Johnston, J., Wei, Y., White, K. P., Henry, R. W. and Arnosti, D. N. (2012). Evidence for Autoregulation and Cell Signaling Pathway Regulation From Genome-Wide Binding of the *Drosophila* Retinoblastoma Protein. *G3* **2**, 1459–1472.

Acharya, P., Raj, N., Buckley, M. S., Zhang, L., Duperon, S., Williams, G., Henry, R. W. and Arnosti, D. N. (2010). Paradoxical Instability–Activity Relationship Defines a Novel Regulatory Pathway for Retinoblastoma Proteins. *Mol. Biol. Cell* **21**, 3890–3901.

Arnold, C. D., Zabidi, M. A., Pagani, M., Rath, M., Schernhuber, K., Kazmar, T. and Stark, A. (2017). Genome-wide assessment of sequence-intrinsic enhancer responsiveness at single-base-pair resolution. *Nat. Biotechnol.* **35**, 136–144.

Bloom, J. and Cross, F. R. (2007). Multiple levels of cyclin specificity in cell-cycle control. *Nat. Rev. Mol. Cell Biol.* **8**, 149–160.

Casimiro, M. C., Crosariol, M., Loro, E., Li, Z. and Pestell, R. G. (2012). Cyclins and Cell Cycle Control in Cancer and Disease. *Genes Cancer* **3**, 649–657.

Cogswell, J. P., Godlevski, M. M., Bonham, M., Bisi, J. and Babiss, L. (1995). Upstream stimulatory factor regulates expression of the cell cycle-dependent cyclin B1 gene promoter. *Mol. Cell. Biol.* **15**, 2782–2790.

Dimova, D. K., Stevaux, O., Frolov, M. V and Dyson, N. J. (2003). Cell cycle-dependent and cell cycle-independent control of transcription by the Drosophila E2F/RB pathway. *Genes Dev.* **17**, 2308–20.

Du, W. and Pogoriler, J. (2006). Retinoblastoma family genes. *Oncogene* **25**, 5190–5200.

Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes Dev.* **12**, 2245–2262.

Ebelt, H., Zhang, Y., Kampke, A., Xu, J., Schlitt, A., Buerke, M., Muller-Werdan, U., Werdan, K. and Braun, T. (2008). E2F2 expression induces proliferation of terminally differentiated cardiomyocytes in vivo. *Cardiovasc. Res.* **80**, 219–226.

Frolov, M. V. (2001). Functional antagonism between E2F family members. *Genes Dev.* **15**, 2146–2160.

Gjidoda, A. and Henry, R. W. (2013). RNA polymerase III repression by the retinoblastoma tumor suppressor protein. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1829**, 385–392.

Henley, S. A. and Dick, F. A. (2012). The retinoblastoma family of proteins and their regulatory functions in the mammalian cell division cycle. *Cell Div.* **7**, 10.

Huang, V., Place, R. F., Portnoy, V., Wang, J., Qi, Z., Jia, Z., Yu, A., Shuman, M., Yu, J. and Li, L.-C. (2012). Upregulation of Cyclin B1 by miRNA and its implications in cancer. *Nucleic Acids Res.* **40**, 1695–1707.

Hwang, A., Maity, A., McKenna, W. G. and Muschel, R. J. (1995). Cell cycle-dependent regulation of the cyclin B1 promoter. *J. Biol. Chem.* **270**, 28419–24.

Infante, A., Laresgoiti, U., Fernández-Rueda, J., Fullaondo, A., Galán, J., Díaz-Uriarte, R., Malumbres, M., Field, S. J. and M. Zubiaga, A. (2008). E2F2 represses cell cycle regulators to maintain quiescence. *Cell Cycle* **7**, 3915–3927.

Innocente, S. A., Abrahamson, J. L. A., Cogswell, J. P. and Lee, J. M. (1999). p53 regulates a G2 checkpoint through cyclin B1. *Proc. Natl. Acad. Sci.* **96**, 2147–2152.

Ito, M. (2000). Factors controlling cyclin B expression. *Plant Mol. Biol.* **43**, 677–90.

Jackson, J. G. and Pereira-Smith, O. M. (2006). Primary and Compensatory Roles for RB Family Members at Cell Cycle Gene Promoters That Are Deacetylated and Downregulated in Doxorubicin-Induced Senescence of Breast Cancer Cells. *Mol. Cell. Biol.* **26**, 2501–2510.

Jackson, M. W. (2005). p130/p107/p105Rb-dependent transcriptional repression during DNA-damage-induced cell-cycle exit at G2. *J. Cell Sci.* **118**, 1821–1832.

Khan, S., Brougham, C. L., Ryan, J., Sahrudin, A., O'Neill, G., Wall, D., Curran, C., Newell, J., Kerin, M. J. and Dwyer, R. M. (2013). miR-379 Regulates Cyclin B1 Expression and Is Decreased in Breast Cancer. *PLoS One* **8**, e68753.

Laoukili, J., Kooistra, M. R. H., Brás, A., Kauw, J., Kerkhoven, R. M., Morrison, A., Clevers, H. and Medema, R. H. (2005). FoxM1 is required for execution of the mitotic programme and chromosome stability. *Nat. Cell Biol.* **7**, 126–136.

Li, Y., Zhang, D. Y., Ren, Q., Ye, F., Zhao, X., Daniels, G., Wu, X., Dynlacht, B. and Lee, P. (2012). Regulation of a Novel Androgen Receptor Target Gene, the Cyclin B1 Gene, through Androgen-Dependent E2F Family Member Switching. *Mol. Cell. Biol.* **32**, 2454–2466.

Lukas, C., Sørensen, C. S., Kramer, E., Santoni-Rugiu, E., Lindeneg, C., Peters, J.-M., Bartek, J. and Lukas, J. (1999). Accumulation of cyclin B1 requires E2F and cyclin-A-dependent rearrangement of the anaphase-promoting complex. *Nature* **401**, 815–818.

Ma, Y., McKay, D. J. and Buttitta, L. (2019). Changes in chromatin accessibility ensure robust cell cycle exit in terminally differentiated cells. (in press)

Markey, M. P., Angus, S. P., Strobeck, M. W., Williams, S. L., Gunawardena, R. W., Aronow, B. J. and Knudsen, E. S. (2002). Unbiased analysis of RB-mediated transcriptional repression identifies novel targets and distinctions from E2F action. *Cancer Res.* **62**, 6587–97.

Mouawad, R., Himadewi, P., Kadiyala, D. and Arnosti, D. N. (2019). Selective repression of the Drosophila Cyclin B promoter by retinoblastoma and E2F proteins. (in press)

Piaggio, G., Farina, A., Perrotti, D., Manni, I., Fuschi, P., Sacchi, A. and Gaetano, C. (1995). Structure and Growth-Dependent Regulation of the Human Cyclin B1 Promoter. *Exp. Cell Res.* **216**, 396–402.

Pines, J. and Hunter, T. (1989). Isolation of a human cyclin cDNA: Evidence for cyclin mRNA and protein regulation in the cell cycle and for interaction with p34cdc2. *Cell* **58**, 833–846.

Satyanarayana, A. and Kaldis, P. (2009). Mammalian cell-cycle regulation: several Cdks, numerous cyclins and diverse compensatory mechanisms. *Oncogene* **28**, 2925–2939.

Sawado, T., Yamaguchi, M., Nishimoto, Y., Ohno, K., Sakaguchi, K. and Matsukage, A. (1998). dE2F2, a Novel E2F-Family Transcription Factor inDrosophila melanogaster. *Biochem. Biophys. Res. Commun.* **251**, 409–415.

Shi, Q., Wang, W., Jia, Z., Chen, P., Ma, K. and Zhou, C. (2016). ISL1, a novel regulator of CCNB1, CCNB2 and c-MYC genes, promotes gastric cancer cell proliferation and tumor growth. *Oncotarget* **7**, 36489–36500.

Stevaux, O. (2002). Distinct mechanisms of E2F regulation by Drosophila RBF1 and RBF2. *EMBO J.* **21**, 4927–4937.

Wei, Y., Gokhale, R. H., Sonnenschein, A., Montgomery, K. M., Ingersoll, A. and Arnosti, D. N. (2016). Complex *cis* -regulatory landscape of the insulin receptor gene underlies the broad expression of a central signaling regulator. *Development* **143**, 3591–3603.

Wei, Y., Mondal, S. S., Mouawad, R., Wilczyński, B., Henry, R. W. and Arnosti, D. N. (2015). Genome-Wide Analysis of *Drosophila* RBf2 Protein Highlights the Diversity of RB Family Targets and Possible Role in Regulation of Ribosome Biosynthesis. *G3* **5**, 1503–1515.

Wirt, S. E. and Sage, J. (2010). p107 in the public eye: an Rb understudy and more. *Cell Div.* **5**, 9.

Yamaguchi, M., Hayashi, Y. and Matsukage, A. (1995). Essential Role of E2F Recognition Sites in Regulation of the Proliferating Cell Nuclear Antigen Gene Promoter during *Drosophila* Development. *J. Biol. Chem.* **270**, 25159–25165.

Yuh, C. H., Bolouri, H. and Davidson, E. H. (2001). Cis-regulatory logic in the endo16 gene: switching from a specification to a differentiation mode of control. *Development* **128**, 617–29.

Zhu, W., Giangrande, P. H. and Nevins, J. R. (2004). E2Fs link the control of G1/S and G2/M transcription. *EMBO J.* **23**, 4615–4626.

CHAPTER 5

The function of the Rbf1 instability element in gene regulation

Abstract

Turnover and activity of retinoblastoma proteins are impacted by a conserved C-terminal region termed the instability element (IE). The IE is conserved in Rbf1, p107 and p130, and conserved residues of this region are changed or absent from the more derived retinoblastoma proteins: Rb in humans and Rbf2 in Drosophila. Therefore, understanding the function of the IE may provide insights on the functional diversification of the retinoblastoma proteins. Here, we studied the impact of Rbf1 and IE-mutant Rbf1 isoforms on gene regulation using RNA-seq analysis in wing imaginal discs and embryos. We showed that in both systems, IE mutations affect the function of Rbf1 in a different manner. Removing the IE region (Rbf1-ΔIE) is usually associated with weaker activity of the protein, indicating that this region is either crucial for gene targeting or associating with cofactors for a potent repression activity. Interestingly, point mutations of specific residues within the IE impact gene regulation in different manners, indicating gene specific functions.

This manuscript is in preparation.

Mouawad, R. and Arnosti, N. D. The function of the Rbf1 instability element in gene regulation. (In prep)

My contribution to this study was the execution of independent RNA-seq experiments from Drosophila wing imaginal discs and embryos, and the data analysis.

Introduction

In addition to regulation by phosphorylation, retinoblastoma protein levels are regulated by the turnover through the ubiquitin proteasome pathway. A conserved C-terminal region (58 residues), termed the instability element (IE) was identified as a degron responsible for degradation of the ancestral retinoblastoma proteins: Rbf1, p107 and p130 (Acharya et al., 2010; Sengupta et al., 2015). Interestingly, the IE also affects repression function as shown in reporter assays for E2F target genes (Acharya et al., 2010; Sengupta et al., 2015). Rbf1-ΔIE (deletion of the IE) mutant is more stable than full length WT-Rbf1 protein but a weaker repressor on canonical E2F target genes such as *PCNA* (Acharya et al., 2010; Raj et al., 2012).

Conserved lysine and phosphorylation serine residues within the IE affect Rbf1 stability and activity in distinct manners (Acharya et al., 2010; Zhang et al., 2014). A mutant form of Rbf1 with a mutation in a conserved lysine residue (K774) was a stronger repressor of *PCNA* reporter and induced severe phenotypes when expressed in fly eyes and wings (Acharya et al., 2010; Zhang et al., 2014; Elenbaas et al., 2015). Within the IE, there are three conserved serine phosphorylation residues (S728, S760 and S771) that represent canonical serine-proline phosphorylation motifs. Overexpression of a mutant form of Rbf1 in which these serines are mutated to alanines (3SA), results in phenotypes as severe as those produced by the K774A mutant (Zhang et al., 2014). Interestingly, both Rbf1-3SA and Rbf1-K774A mutant isoforms were not inactivated by Cyclin-Cdk overexpression, as shown by their ability to repress *PCNA* reporter (Zhang et al., 2014).

In vitro studies showed that acetylation of the homologous lysine to K774 residue is important for in vitro Cdk4-mediated phosphorylation of the human Rb protein (Saeed et al., 2012), pointing to important roles in regulation of retinoblastoma proteins by phosphorylation. It is not fully

understood whether mutation of this residue has the same consequences for Rbf1 function as the conversion of three serines to alanine in the IE, which are phosphorylation target sites. Interestingly, conserved residues of the IE are modified or lost in the more derived retinoblastoma proteins, Rb in mammals and Rbf2 in Drosophila, pointing to important functions that are associated with functional diversification of retinoblastoma proteins. Considering the functional role of the IE, and the evident evolutionary changes that impact its structure, more studies have to address the role of conserved functional residues within the IE in a physiological context. In addition to evolutionary perspectives, it is interesting that specific point mutations and deletions within p107 and p130 IE region have been reported in cancer patients, suggesting that altered function of this segment of Rb family proteins may be relevant to understanding disease (Forbes et al., 2011; Gao et al., 2013).

Results and Discussion

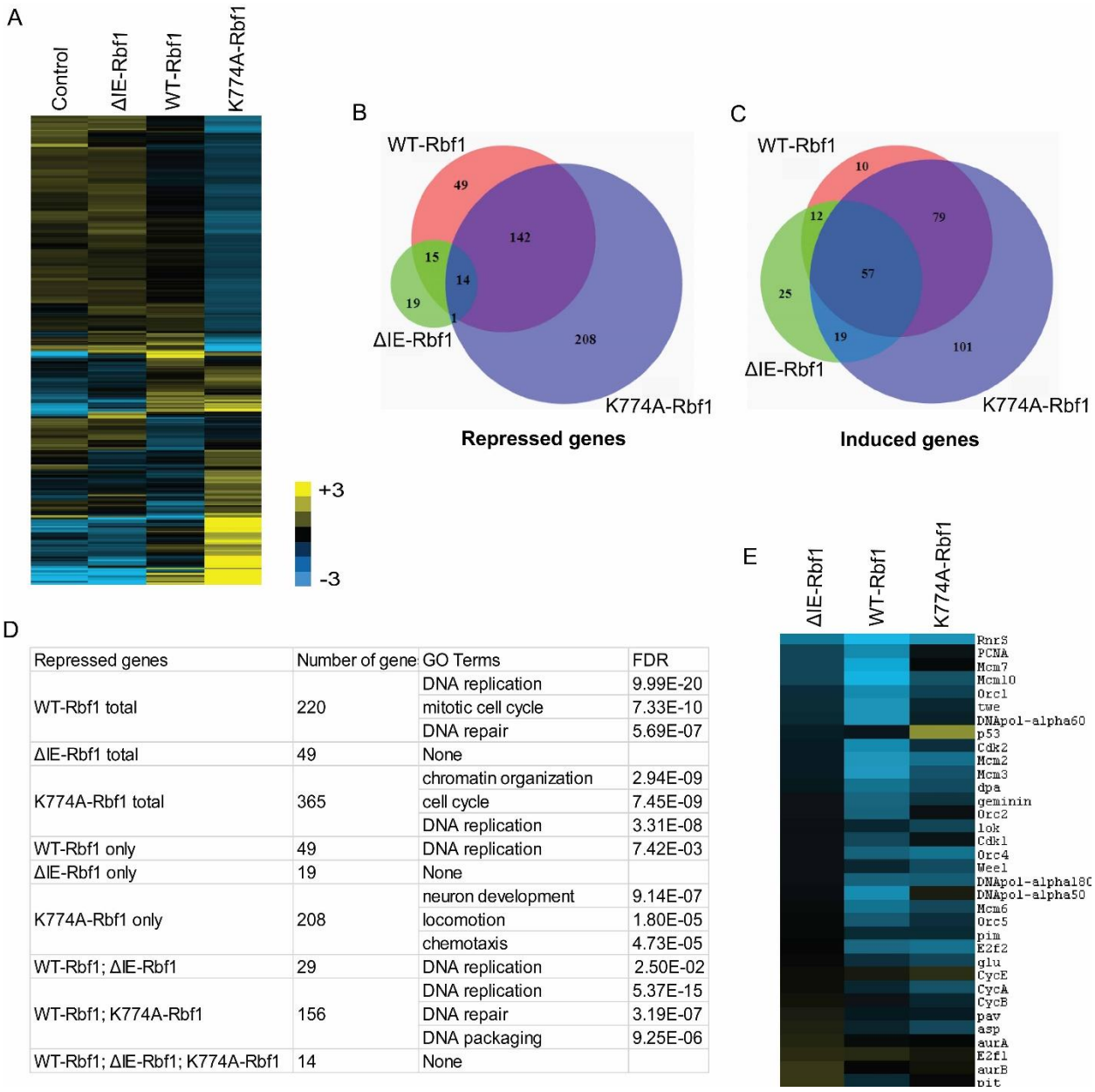
Regulation of gene expression by Rbf1 and Rbf1 mutant proteins in wing imaginal discs.

We showed in earlier studies that mutations in the Rbf1 IE region resulted in distinct wing phenotypes when these mutants were expressed in fly wing imaginal discs; wings with overexpressed Rbf1-ΔIE were only subtly affected, while overexpression of the wild-type Rbf1 protein induced apoptosis and notching of the wing. Expression of the Rbf1-K774A mutant induced high levels of apoptosis and severe wing developmental defects (Elenbaas et al., 2015). To understand the impact of mutations of the Rbf1 IE region on gene regulation in this system, we performed an RNA-seq analysis of wing imaginal discs overexpressing either WT-Rbf1, Rbf1-ΔIE or Rbf1-K774A proteins. The relevant cDNAs were expressed by a GAL4 activator protein expressed under control of the *Pen* promoter, which is active in the wing imaginal discs of the third instar larvae, at a developmental stage when the wing is patterned from a sheet of epithelial cells to become the differentiated structure. We focused on genes that are directly occupied by Rbf1, based on previous ChIP-seq data for the endogenous protein in embryos (Wei et al., 2015). Overall, expression of the Rbf1-ΔIE resulted in a minimal impact on gene expression, with only 162 genes' expression changed more than 20% from the control levels. In contrast, overexpression the Rbf1-K774A mutant resulted in a more pervasive effect on gene expression, with 621 genes showing up or down regulation. The overexpression of the WT Rbf1 protein resulted in moderate changes to gene expression, with 378 genes' expression affected (Figure 5-1A, B, C). These results are consistent with the different phenotypes that resulted from overexpression of these two different mutants in wing imaginal discs (Elenbaas et al., 2015).

Although we focused on potential Rbf1 direct targets, the significant number of genes induced after the expression of the Rbf1 isoforms contrasts with the expected suppression of transcription

by this corepressor. We hypothesize that these inductions may reflect indirect effects on gene regulation, as in this experimental setup, the expression of these proteins continues for approximately 72 hours, long enough for secondary and tertiary responses to occur. Gene ontology analysis for genes induced by all the Rbf1 isoforms did not reveal significant enrichment for any biological process (data not shown). In contrast, gene ontology for genes repressed by the WT-Rbf1 and Rbf1-K774A overexpression showed enrichment for cell cycle related genes. Interestingly, genes that are uniquely repressed by the Rbf1-K774A protein are enriched for neuron development and nervous system processes (Figure 5-1D).

We focused on cell cycle genes that are known to be regulated by Rbf1 protein to determine how they are affected by the IE mutations (Figure 5-1E). Expression of the Rbf1-ΔIE protein had little effect on this group of genes, in contrast to the effects of expression of the WT protein. Interestingly, unlike the overall pattern of stronger repression activity on many genes, for cell cycle-related genes, the Rbf1-K774A mutant had a much weaker effect on cell cycle genes, compared to that of the WT protein. Therefore, in this setting, the Rbf1-K774A mutant does not seem to be a uniformly more potent repressor than the WT-Rbf1 protein, although the possibility of secondary and tertiary effects makes a direct comparison difficult. In any event, it is clear that mutations affecting this lysine residue have a gene specific function. Consistent with our previous findings that this mutation induces apoptosis in wing imaginal discs, we find that *p53* is induced (Figure 5-1E).



Repressed genes	Number of gene	GO Terms	FDR
WT-Rbf1 total	220	DNA replication	9.99E-20
		mitotic cell cycle	7.33E-10
		DNA repair	5.69E-07
ΔIE-Rbf1 total	49	None	
K774A-Rbf1 total	365	chromatin organization	2.94E-09
		cell cycle	7.45E-09
		DNA replication	3.31E-08
WT-Rbf1 only	49	DNA replication	7.42E-03
ΔIE-Rbf1 only	19	None	
K774A-Rbf1 only	208	neuron development	9.14E-07
		locomotion	1.80E-05
		chemotaxis	4.73E-05
WT-Rbf1; ΔIE-Rbf1	29	DNA replication	2.50E-02
WT-Rbf1; K774A-Rbf1	156	DNA replication	5.37E-15
		DNA repair	3.19E-07
		DNA packaging	9.25E-06
WT-Rbf1; ΔIE-Rbf1; K774A-Rbf1	14	None	

Figure 5-1: RNA-seq analysis from wing imaginal discs expressing WT and mutant Rbf1 proteins. (A) A heatmap generated by unsupervised clustering of RNA-seq data from wing imaginal discs overexpressing WT-Rbf1, ΔIE-Rbf1 and K774A-Rbf1. Control samples represent wing imaginal discs overexpressing GFP. Values represent average of three biologic replicates, and counts lower than 10 are filtered out, and only genes bound by Rbf1 based on ChIP-seq data are included. Values represent log transformed RNA-seq counts that are mean centered for each

Figure 5-1 (cont'd)

gene. Blue indicates reads below the mean, black equal to the mean, and yellow, above the mean. (B) Venn diagram showing the overlap between genes repressed after overexpression of each Rbf1 isoform. Those genes were extracted from the heatmap, and repression indicates at least 20% downregulation of the gene with respect to control. (C) Venn diagram showing the overlap between genes induced after overexpression of each Rbf1 isoform. Those genes were extracted from the heatmap, and induction indicates at least 20% upregulation of the gene with respect to control. (D) Gene ontology analysis on unique and overlapping repressed genes shown in the Venn diagram in B. The top three categories based on FDR were selected and displayed in the table. "None" indicated no enrichment of any biologic process. (E) Relative gene expression of cell cycle genes that have counts > 10 and are bound by Rbf1 in vivo. Values represent average of three biologic replicates.

Regulation of gene expression by Rbf1 and Rbf1 mutant proteins in embryos.

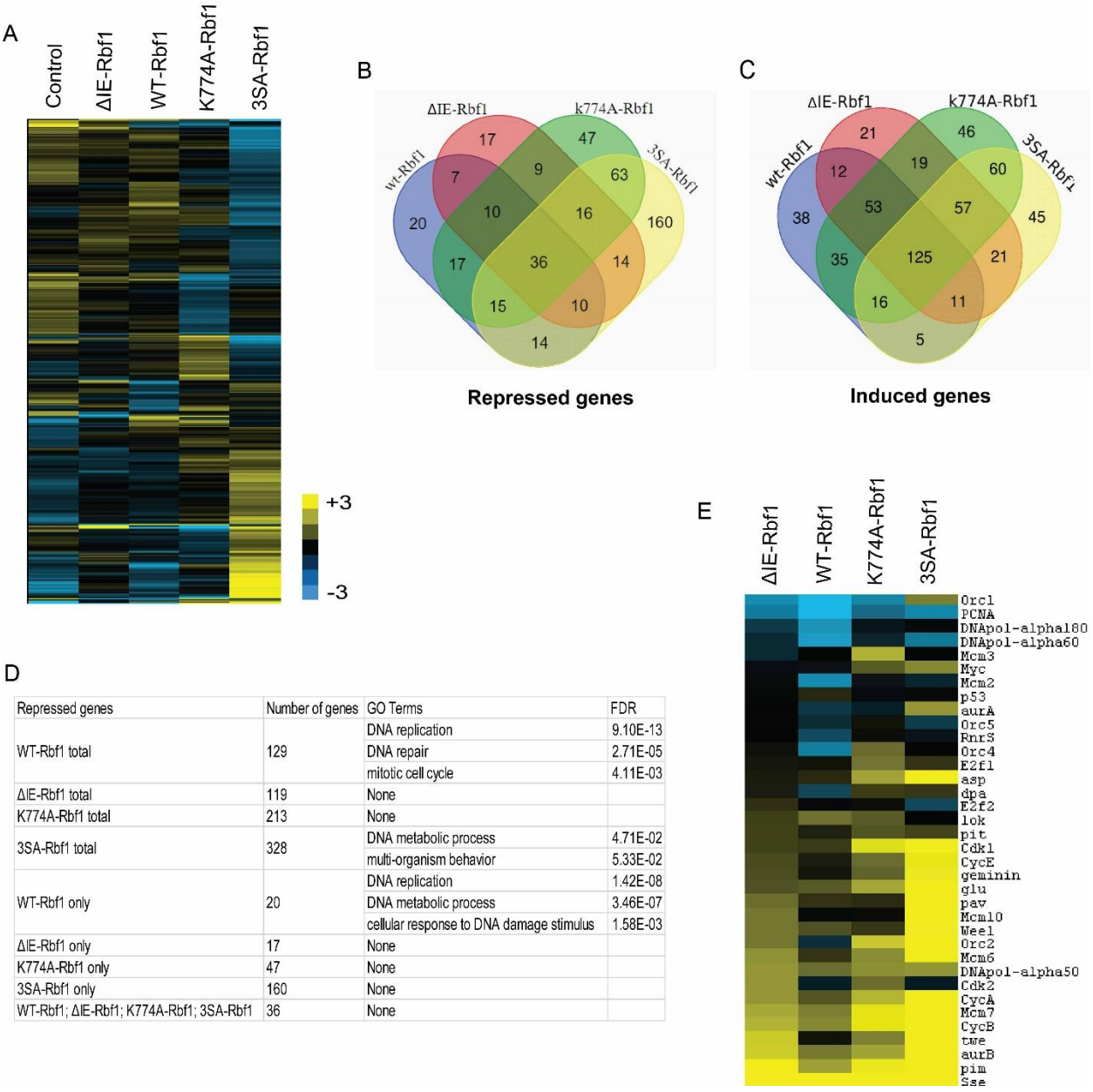
As noted above, the long-term expression of Rbf1 proteins in the wing disc allows for secondary and tertiary effects to be manifested, which can complicate interpretation of possible direct effects on promoter activity. To obtain a more direct insight into transcriptional regulation, as well as assessing the activities of these proteins in an earlier developmental period when Rbf1 is also active, we expressed the WT-Rbf1 protein and different IE mutants in 12-18 hr. embryos using a heat shock inducible system. After a brief 20-minute heat induction, embryos were aged 60 minutes before RNA collection and RNA-seq analysis. Here again, we focused on genes directly bound by the endogenous Rbf1 protein (Wei et al., 2015). In these experiments we included an additional IE mutant, 3SA-Rbf1, in which three conserved serine residues in the IE

were changed to alanine. Similar to Rbf1-K774A, this mutant protein was also shown to induce severe phenotypes when expressed in Drosophila wings and eyes (Zhang et al., 2014).

In this system, the ΔIE-Rbf1 resulted in more changes to gene expression, affecting the expression of 438 genes, which was similar to the level of perturbation resulting from the WT-Rbf1 protein (Figure 5-2A, B, C). As expected, the Rbf1-K774A protein resulted in a greater impact on gene expression. Overexpression of the Rbf1-3SA resulted in the most pervasive effect on gene expression, and showed a pattern that was distinct from that of Rbf1-K774A. In our earlier studies, we found that Rbf1-3SA and Rbf1-K774A mutant isoforms had similar properties; they were not inactivated by Cyclin-Cdk overexpression for repression of a *PCNA* reporter, and overexpression of these mutants resulted in similar developmental phenotypes (Zhang et al., 2014). One hypothesis explaining these similarities is that modification of K774 regulates phosphorylation of neighboring serine residues. Therefore, it was unexpected that overexpression of Rbf1-3SA and Rbf1-K774A in embryos resulted in different effects on gene expression ((Figure 5-2A, B, C), indicating that the function of the lysine K774 residue may not simply control modification of the adjacent serine residues present in the IE. We hypothesize that the IE impacts Rbf1 repression potency and gene targeting in a context-specific manner.

As expected, gene ontology analysis showed enrichment for cell cycle related genes among genes repressed after overexpression of the WT-Rbf1 protein (Figure 5-2D). There was no significant enrichment for any biological process in the genes that were repressed by the other Rbf1 isoforms. Surprisingly, for the induced genes, GO analysis on the commonly induced genes (n=125) showed enrichment for cell cycle and mitosis genes (data not shown). Indeed, when we looked closely at cell cycle genes, we found a number of them, including *CycB*, being induced by all the Rbf1

isoforms (Figure 5-2E). Interestingly, *PCNA* is repressed by all the Rbf1 isoforms, indicating that even if IE mutations are present, *PCNA* is still sensitive to Rbf1 overexpression.



D

Repressed genes	Number of genes	GO Terms	FDR
WT-Rbf1 total	129	DNA replication	9.10E-13
		DNA repair	2.71E-05
		mitotic cell cycle	4.11E-03
ΔIE-Rbf1 total	119	None	
K774A-Rbf1 total	213	None	
3SA-Rbf1 total	328	DNA metabolic process	4.71E-02
		multi-organism behavior	5.33E-02
WT-Rbf1 only	20	DNA replication	1.42E-08
		DNA metabolic process	3.46E-07
		cellular response to DNA damage stimulus	1.58E-03
ΔIE-Rbf1 only	17	None	
K774A-Rbf1 only	47	None	
3SA-Rbf1 only	160	None	
WT-Rbf1; ΔIE-Rbf1; K774A-Rbf1; 3SA-Rbf1	36	None	

Figure 5-2: RNA-seq analysis from embryos overexpressing WT and mutant Rbf1 proteins. (A) A heatmap generated by unsupervised clustering of RNA-seq data from embryos overexpressing WT-Rbf1, ΔIE-Rbf1, K774A-Rbf1 and 3SA-Rbf1 proteins, or control embryos.

Figure 5-2 (cont'd)

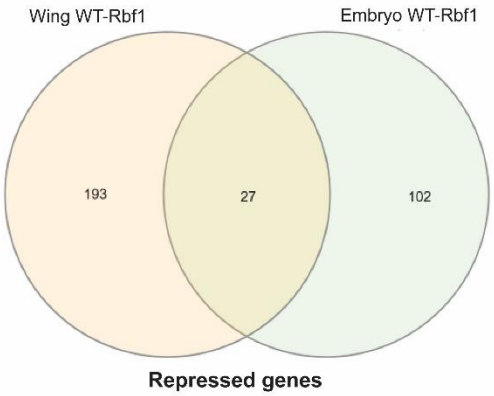
Values represent average of three biologic replicates, and RPKMs <1 are filtered out, and only genes bound by Rbf1 based on ChIP-seq data are included. Values represent log transformed RNA-seq counts that are mean centered for each gene. Blue indicates reads below the mean, black equal to the mean, and yellow, above the mean. (B) Venn diagram showing the overlap between genes repressed after overexpression of each Rbf1 isoform. Those genes were extracted from the heatmap, and repression indicates at least 20% downregulation of the gene with respect to control. (C) Venn diagram showing the overlap between genes induced after overexpression of each Rbf1 isoform. Those genes were extracted from the heatmap, and induction indicates at least 20% upregulation of the gene with respect to control. (D) Gene ontology analysis on unique and overlapping repressed genes shown in the Venn diagram in B. The top three categories based on FDR were selected and displayed in the table. "None" indicated no enrichment of any biologic process. (E) Relative gene expression of cell cycle genes that have RPKMs> 1 and are bound by Rbf1 in vivo. Values are log-transformed and represent average of three biologic replicates.

The impact on gene expression by Rbf1 is context-dependent.

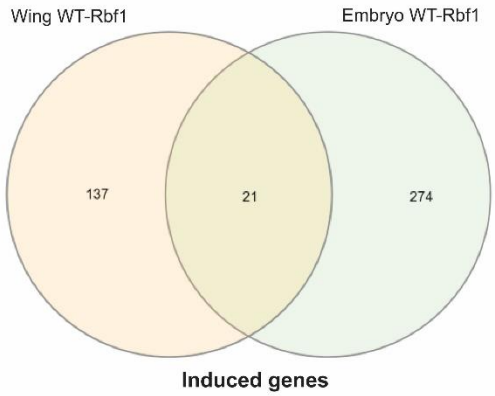
To determine whether similar changes are observed in both wing imaginal discs and embryos, we compared overlap of the two RNA-seq data sets. There was small overlap between genes repressed and induced by the overexpression of the WT-Rbf1 protein (Figure 5-3A, B). Only 27 genes were similarly repressed in the two systems, and these genes were enriched for cell cycle related genes (Figure 5-3C). Similarly for the induced genes, the overlap was only 21 genes and no enrichment for any biological process (Figure 5-3D). Similarly, overexpression of ΔIE-Rbf1 and K774A-Rbf1 resulted in small overlap between wing and embryo datasets (Figure 5-4, 5-5).

Interestingly, in both datasets, *PCNA* was repressed by overexpression of ΔIE-Rbf1 (Figure 5-4A, B) indicating the IE is not necessary for repression of this gene. K774A-Rbf1 overexpression in the wing imaginal discs resulted in repression of genes involved in cell cycle and neurogenesis. These biological processes were absent from the embryo data.

A



B



C

Repressed genes	Number of genes	GO Terms	FDR
WT-Rbf1 wing only	193	cell cycle	1.80E-07
		gamete generation	3.05E-06
		regulation of chromatin organization	8.29E-05
WT-Rbf1 embryo only	102	None	
WT-Rbf1 wing and embryo	27	DNA replication	6.58E-19
		mitotic cell cycle	2.09E-08
		DNA repair	1.08E-07

D

Induced genes	Number of gene	GO Terms	FDR
WT-Rbf1 wing only	137	neuron development	4.59E-02
WT-Rbf1 embryo only	274	mitotic cell cycle	6.32E-09
		organelle fission	3.99E-07
		gamete generation	9.15E-05
WT-Rbf1 wing and embryo	21	None	

Figure 5-3: Comparison of RNA-seq data from wing imaginal discs and embryos expressing WT-Rbf1 protein. (A) Venn diagram showing intersection of repressed genes in wing discs and embryo. Repression indicates at least 20% downregulation of the gene with respect to control in each dataset. RPKM<1 and counts<10 are removed from the wing and embryo datasets

Figure 5-3 (cont'd)

respectively. Only genes bound by Rbf1 based on previous ChIP-seq data are included. (B) Venn diagram showing intersection of induced genes in wing discs and embryo. Induction indicates at least 20% upregulation of the gene with respect to control in each dataset. RPKM<1 and counts<10 are removed from the wing and embryo datasets respectively. Only genes bound by Rbf1 based on previous ChIP-seq data are included. (C) Gene ontology analysis on unique and overlapping repressed and (D) induced genes shown in the Venn diagram in B. The top three categories based on FDR were selected and displayed in the table. "None" indicated no enrichment of any biologic process.

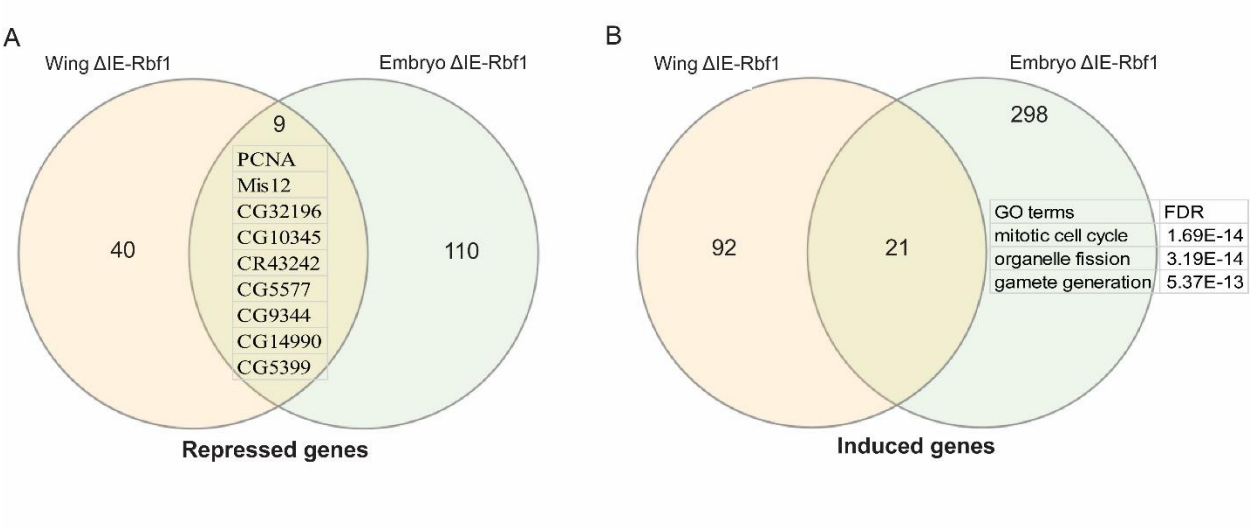
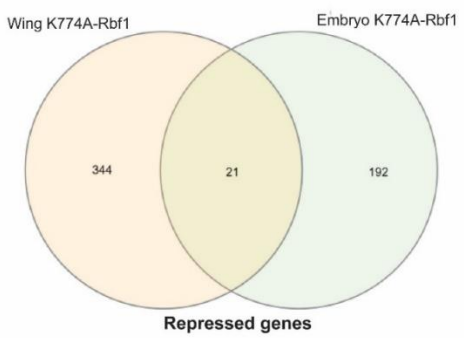


Figure 5-4: Comparison of RNA-seq data from wing imaginal discs and embryos expressing ΔIE-Rbf1 protein. (A) Venn diagram showing intersection of repressed genes in wing discs and embryo. Repression indicates at least 20% downregulation of the gene with respect to control in each dataset. RPKM<1 and counts<10 are removed from the wing and embryo datasets respectively. Only genes bound by Rbf1 based on previous ChIP-seq data are included. The nine

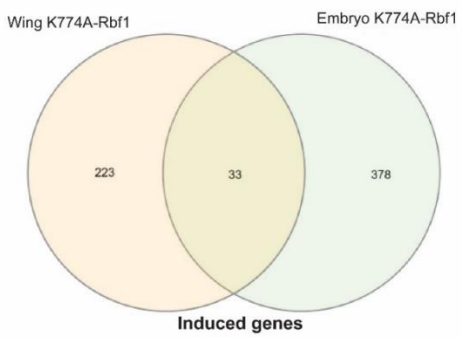
Figure 5-4 (cont'd)

commonly repressed genes are listed. (B) Venn diagram showing intersection of induced genes in wing discs and embryo. Induction indicates at least 20% upregulation of the gene with respect to control in each dataset. RPKM<1 and counts<10 are removed from the wing and embryo datasets respectively. Only genes bound by Rbf1 based on previous ChIP-seq data are included. The GO term for uniquely induced genes (298) in embryo are shown in the table.

A



B



C

Repressed genes	Number of genes	GO Terms	FDR
K774A-Rbf1 wing only	344	chromosome organization	2.17E-15
		cell cycle	1.45E-11
		neurogenesis	7.64E-10
K774A-Rbf1 embryo only	192	None	
K774A-Rbf1 wing and embryo	21	None	

D

Induced genes	Number of genes	GO Terms	FDR
K774A-Rbf1 wing only	223	imaginal disc development	3.65E-03
K774A-Rbf1 embryo only	378	cell cycle	2.08E-26
		chromosome organization	3.18E-17
		nuclear division	5.60E-16
K774A-Rbf1 wing and embryo	33	female gamete generation	4.09E-02
		regulation of cell size	4.35E-02
		oogenesis	5.41E-02

Figure 5-5: Comparison of RNA-seq data from wing imaginal discs and embryos expressing K774A-Rbf1 protein. (A) Venn diagram showing intersection of repressed genes in wing discs and embryo. Repression indicates at least 20% downregulation of the gene with respect to control

Figure 5-5 (cont'd)

in each dataset. RPKM<1 and counts<10 are removed from the wing and embryo datasets respectively. Only genes bound by Rbf1 based on previous ChIP-seq data are included. (B) Venn diagram showing intersection of induced genes in wing discs and embryo. Induction indicates at least 20% upregulation of the gene with respect to control in each dataset. RPKM<1 and counts<10 are removed from the wing and embryo datasets respectively. Only genes bound by Rbf1 based on previous ChIP-seq data are included. (C) Gene ontology analysis on unique and overlapping repressed and (D) induced genes shown in the Venn diagram in B. The top three categories based on FDR were selected and displayed in the table. "None" indicated no enrichment of any biologic process.

Collectively, these data indicate that in each tissue, Rbf1 and IE isoforms have different impacts on gene regulation. However, as noted, the differences in how these experiments are conducted is likely to impact the outcome of RNA-seq analysis; transient embryo expression allows for better assessment of direct transcriptional impacts, while the sustained expression in the wing disc can reveal gene pathways that include direct and indirect impacts.

In conclusion, we were able to provide insights on the function of the IE *in vivo* in a whole animal and in a specific tissue. We found that depending on the tissue or cell type, Rbf1 may have a different effect on gene regulation, and the IE affects this regulation. More studies are needed to address the molecular mechanisms by which the IE region influence gene targeting and repression activity of Rbf1.

Material and methods

RNA-seq analysis in wing imaginal discs

The UAS-*WT-RBf1*, UAS-*ΔIE-RBf1* and UAS-*K774A-RBf1* fly lines were constructed as previously described (Zhang et al. 2014, Elenbaas et al., 2015). *Pendulin*-Gal4 driver line (Stock Number: 113920) and UAS-*GFP* line which is used as control (Stock Numbers: 35786) were obtained from Bloomington Stock Center. Wing imaginal discs were dissected from third-instar larvae generated from crossing *Pen*Gal4 flies to UAS-*WT-RBf1*, UAS-*ΔIE-RBf1*, UAS-*K774A RBf1* or UAS-*GFP* flies. Total RNA was isolated using TRIzol (Invitrogen) and RNeasy Mini kit (QIAGEN). Three biologic replicates from each cross were submitted for sequencing using Illumina HiSeq2500 platform as previously described (Wei et al., 2015). RNA-seq reads were mapped using TopHat v2.1.0 and counted using HTSeq v0.6.0. Only genes that have counts >10 and are occupied by Rbf1 based on previous ChIP-seq data (Wei et al., 2015) were included in the following analysis. Unsupervised clustering was performed on 1615 genes using Cluster3.0 software based on Euclidean distance, and the heatmap was visualized using JAVA TreeView v1.1.6r4. The average counts from three biologic replicates were log transformed, mean centered and filtered at 0.2 SD to remove genes with little variation across samples (703 genes remained in the heatmap). Genes repressed or induced (at least 20% change) by each Rbf isoform were picked from the heatmap to generate Venn diagrams.

RNA-seq analysis in embryos

Embryos (12-18 hr age) carrying heat shock inducible *WT-Rbf1*, *ΔIE-RBf1*, *K774A-Rbf1* or *3SA-Rbf1* transgenes, as well as control flies that don't carry any *Rbf* transgene, were briefly heat shocked for 20-minutes. After 60 minute-recovery time, total RNA was extracted using Total RNA

Kit (OMEGA). Poly-A+ RNAs were purified from the total RNA using Oligotex mRNA Mini kit (Qiagen) and were prepared for the single molecule sequencing (SMS) as described previously (Kapranov et al. 2010). Sequencing was performed at the SeqLL, LLC facility (Woburn, MA). Reads were processed and aligned to the Dm6 version of the *Drosophila melonagaster* genome using indexDP genomic aligner as described previously (Giladi et al., 2010; Kapranov et al., 2010). RPKM values were generated from uniquely aligned reads for each transcript annotated in the RefSeq Genes database of the UCSC Genome browser (Kent et al. 2002). Three biologic replicates were done for each sample. Only genes with RPKM>1 and bound by either Rbf1 based on previous ChIP-seq dataset (Wei el al., 2015) were further analyzed. Unsupervised clustering was performed on 1756 genes using Cluster3.0 software, and the heatmap was visualized using JAVA TreeView v1.1.6r4. The counts were log transformed, mean centered and filtered at 0.2 SD to generate a heatmap that includes a total of 893 genes. Genes repressed or induced (at least 20% change) by each Rbf isoform were picked from the heatmap to generate Venn diagrams.

Gene ontology

Gene ontology analysis was performed for repressed and induced genes in both wing discs and embyos using PANTHER14.1. Overrepresentation test based on GO biological process was calculated using Fisher's exact test with FDR correction.

REFERENCES

REFERENCES

Acharya, P., Raj, N., Buckley, M. S., Zhang, L., Duperon, S., Williams, G., Henry, R. W. and Arnosti, D. N. (2010). Paradoxical Instability–Activity Relationship Defines a Novel Regulatory Pathway for Retinoblastoma Proteins. *Mol. Biol. Cell* **21**, 3890–3901.

Elenbaas, J. S., Mouawad, R., Henry, R. W., Arnosti, D. N. and Payankaulam, S. (2015). Role of Drosophila retinoblastoma protein instability element in cell growth and proliferation. *Cell Cycle* **14**, 589–97.

Forbes, S. A., Bindal, N., Bamford, S., Cole, C., Kok, C. Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945-50.

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* **6**, pl1–pl1.

Giladi, E., Healy, J., Myers, G., Hart, C., Kapranov, P., Lipson, D., Roels, S., Thayer, E. and Letovsky, S. (2010). Error Tolerant Indexing and Alignment of Short Reads with Covering Template Families. *J. Comput. Biol.* **17**, 1397–1411.

Kapranov, P., St Laurent, G., Raz, T., Ozsolak, F., Reynolds, C. P., Sorensen, P. H., Reaman, G., Milos, P., Arceci, R. J., Thompson, J. F., et al. (2010). The majority of total nuclear-encoded non-ribosomal RNA in a human cell is "dark matter" un-annotated RNA. *BMC Biol.* **8**, 149.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. and Haussler, D. (2002). The human genome browser at UCSC. *Genome Res.* **12**, 996–1006.

Raj, N., Zhang, L., Wei, Y., Arnosti, D. N. and Henry, R. W. (2012). Ubiquitination of retinoblastoma family protein 1 potentiates gene-specific repression function. *J. Biol. Chem.* **287**, 41835–43.

Saeed, M., Schwarze, F., Loidl, A., Meraner, J., Lechner, M. and Loidl, P. (2012). In vitro phosphorylation and acetylation of the murine pocket protein Rb2/p130. *PLoS One* **7**, e46174.

Sengupta, S., Lingnurkar, R., Carey, T. S., Pomaville, M., Kar, P., Feig, M., Wilson, C. A., Knott, J. G., Arnosti, D. N. and Henry, R. W. (2015). The Evolutionarily Conserved C-terminal Domains in the Mammalian Retinoblastoma Tumor Suppressor Family Serve as Dual Regulators of Protein Stability and Transcriptional Potency. *J. Biol. Chem.* **290**, 14462–75.

Wei, Y., Mondal, S. S., Mouawad, R., Wilczyński, B., Henry, R. W. and Arnosti, D. N. (2015). Genome-Wide Analysis of *Drosophila* RBf2 Protein Highlights the Diversity of RB Family Targets and Possible Role in Regulation of Ribosome Biosynthesis. *G3* **5**, 1503–1515.

Zhang, L., Wei, Y., Pushel, I., Heinze, K., Elenbaas, J., Henry, R. W. and Arnosti, D. N. (2014). Integrated stability and activity control of the Drosophila Rbf1 retinoblastoma protein. *J. Biol. Chem.* **289**, 24863–73.

CHAPTER 6

Future Directions

Here, I discuss future projects designed to learn more about the function of the retinoblastoma family of proteins. My research has transformed our understanding of the genomic-level role of the Drosophila retinoblastoma protein Rbf2, and the insights from this work point to the utility of using the fly Rbf1/Rbf2 system to help us understand the roles of the human retinoblastoma proteins.

In Chapter 3, we show that Rbf2 plays a role in regulating female fertility (Mouawad et al., in press). To understand the molecular mechanisms by which Rbf2 regulates ovary function, I propose to perform an RNA-seq experiment on *Rbf2* mutant ovaries. This assay will provide a genome-wide perspective on the expression of genes impacted by the loss of Rbf2 in the ovaries, and would potentially indicate a molecular mechanism for the egg laying phenotypes that we observe. I expect to find genes related to the Pi3K pathway, since *Pi3K92E* was one of the genes significantly induced in *Rbf2* mutant ovaries. In addition, since *Rbf2* mutant males and females show a decrease in lifespan, functions that impact both sexes also remain to be elucidated. Therefore, it will be interesting to analyze the transcriptome in embryos, where Rbf2 is expressed in both sexes. I hypothesize that Rbf2 may be impacting processes in embryonic development that are leading to a decrease in lifespan when Rbf2 is lost. The widespread expression of *Rbf2* leaves open the possibility that the observed lifespan phenotype may be related to function of this gene product in one or more tissues, including the central nervous system. The critical genes that are impacted may include those that we identified as “Rbf2-exclusive” targets, or they may include

genes in which a proposed interplay between Rbf1 and Rbf2 is critical. The advantages of a general picture from RNA-seq assays is that one can assess the impacts of Rbf2 across many genes and systems. Of course, these experiments will not be able to directly differentiate direct from indirect effects, although our knowledge about the physical targeting of genes in the embryo from ChIP-seq experiments provides a first-level estimation of likely direct transcriptional effects. Another experiment that would be revealing is studying the expression of Rbf2 within the ovary to determine in which cells it is predominantly present, which will give us hints about its function. In our first assessment of Rbf2 mRNA and protein expression in development, we did not examine the expression in adult tissues in detail. A first-level assessment would involve carrying out immunohistochemical studies of ovaries using antibodies to Rbf2. A more fine-grained analysis would involve examining the impact of Rbf2 loss on individual cell types using single-cell RNA-seq approaches. Recent studies from the Frolov laboratory have described how such sc-RNA-seq analysis can provide a detailed picture of Rbf1 function in the eye imaginal disc (Ariss et al., 2018). It is useful that the *Rbf2* mutant phenotype is not lethal, allowing us to test the activity of the gene in different developmental stages. For instance, we can study the impact of presumably complete *Rbf2* loss on ovary function; it would be interesting to study the impact of *Rbf2* heterozygosity as well. I expect that the phenotypes may not be the same as the *Rbf2* null phenotype, as the gross phenotypes of the heterozygotes are less severe as the homozygous mutant.

My studies describe a parallel evolution of Drosophila and human retinoblastoma proteins emphasizing the divergence of the C-terminal domain in human Rb and Drosophila Rbf2. Interestingly, we showed that the C-terminal instability element, which is lost in the derived Rb proteins, has important impacts on cell growth and apoptosis (Elenbaas et al., 2015). Conserved serine and lysine residues within the IE have significant impact on the normal function of the

protein (Zhang et al., 2015). In order to study the function of the IE and how it contributed to the functional diversification of the retinoblastoma proteins, it would be very interesting to understand molecular mechanisms in which the IE impacts gene targeting and repression potency of Rbf1. As shown in Chapter 5, when overexpressed in embryo and wing imaginal discs, Rbf1 IE-mutants impact gene regulation in distinct manners. Currently, it is not known whether the differential effects of the different isoforms is due to changes in the binding of the Rbf1 proteins to target genes, or whether their repression activities are changed due to differential association with chromatin modifying/remodeling complexes, for example.

Therefore, using these datasets, it will be interesting to pick groups of genes and study the impact of IE mutations on binding to target genes and on changes to chromatin structure. We are currently using a modified CRISPR/Cas-9 system to drive IE-mutant and WT Rbf1 proteins to specific genes and study chromatin changes. This analysis will provide insights on the molecular mechanisms by which the IE impacts Rbf1 function and will help us understand why the more derived Rb proteins lack the IE element, and how this alters their function. I have also used my RNA-seq datasets to select genes that were differentially regulated by the overexpression of different IE-mutant Rbf1 proteins, and we have created transgenic flies containing reporters generated from those genes. We are currently using these transgenic flies and inducing the expression of the WT and the IE-mutant Rbf1 proteins to measure changes on the reporter gene in vivo. The goal of these experiments is to test whether the defined promoter-proximal areas are sufficient to mediate the differential effects of Rbf1 isoforms (wild-type protein, and mutants with changes to the C-terminal domain). If regulation of these transgenes is the same as for their endogenous counterparts, it will be interesting to mutate cis elements on promoters of these reporter genes to understand the impact of promoter structure on repression by Rbf1 and the IE-mutant proteins. This analysis will provide insights on

the role of the IE in gene specific repression and targeting and how promoter contexts influence this process.

In Chapter 4, I change my focus from genome-wide perspectives to a detailed promoter-specific analysis of Rbf1 and Rbf2 function. This work has provided insights on the role of promoter contexts on gene regulation by Rbf1 and Rbf2. I show that Rbf2 is a potent repressor of *CycB* in contrast to the weak activity on *PCNA*. I mutated putative E2F sites within the core promoter region, and found that other E2F sites are playing redundant roles in regulating the *CycB* gene. Therefore, it would be interesting to extend the analysis of cis regulatory sequences within this promoter, including the mutation of additional E2F sites, including ones that lie in the conserved 3' intronic region that I show to have a repressive role on the expression of the *CycB* reporter. Some of these E2F sites are highly conserved in all Drosophila lineages and are therefore of special interest for further exploration. An interesting point from these studies is that E2f1 and E2f2 functions on *CycB* reporter are opposite to the canonical picture which holds that E2f1 is an activator and E2f2 is associated with repression function. Additional E2F mutations on the *CycB* reporter will provide insights on mechanisms by which E2f factors change from activators to repressors or vice versa on specific genes. An unsolved question from these studies of the *CycB* promoter is that in contrast to the repression by Rbf2 seen in actively cycling cultured S2 cells, the gene is activated by overexpression of Rbf2 in the context of 12-18 hour old embryos, possibly because of differences between the cell cycle status and occupancy of Rbf1 between these two conditions.

In summary, my research has provided experimental evidence for a different type of function from the simple model describing Rb proteins as a mechanism for on/off switching of cell cycle genes. Previous work by Dyson and colleagues has hinted at the role for Rb proteins on developmentally

regulated genes, where cell-cycle phosphorylation and inactivation is not part of the regulation (Dimova et al., 2003). With my studies, we now have a richer picture that indicates pocket proteins can play more nuanced roles, modulating the expression of specific sets of cell growth control genes, either counterbalancing Rbf1 controls, or working independently. This differentiated function appears to be a field of gene network regulation in which differential evolutionary scenarios are explored, impacting the structure and function of retinoblastoma proteins. Future molecular and evolutionary studies will provide us a richer understanding of the complexity of retinoblastoma protein function in metazoa.

REFERENCES

REFERENCES

Ariss, M. M., Islam, A. B. M. M. K., Critcher, M., Zappia, M. P. and Frolov, M. V. (2018). Single cell RNA-sequencing identifies a metabolic aspect of apoptosis in Rbf mutant. *Nat. Commun.* **9**, 5024.

Dimova, D. K., Stevaux, O., Frolov, M. V and Dyson, N. J. (2003). Cell cycle-dependent and cell cycle-independent control of transcription by the Drosophila E2F/RB pathway. *Genes Dev.* **17**, 2308–20.

Elenbaas, J. S., Mouawad, R., Henry, R. W., Arnosti, D. N. and Payankaulam, S. (2015). Role of Drosophila retinoblastoma protein instability element in cell growth and proliferation. *Cell Cycle* **14**, 589–97.

Mouawad, R., Himadewi, P., Kadiyala, D. and Arnosti, D. N. (2019). Selective repression of the Drosophila Cyclin B promoter by retinoblastoma and E2F proteins. (in press)

Zhang, L., Wei, Y., Pushel, I., Heinze, K., Elenbaas, J., Henry, R. W. and Arnosti, D. N. (2014). Integrated stability and activity control of the Drosophila Rbf1 retinoblastoma protein. *J. Biol. Chem.* **289**, 24863–73.