Identification of *Phosphoribosyl Pyrophosphate Synthetase* (PRPS) as a direct target of the transcriptional repressor Capicua in *Drosophila melanogaster*

Conor Dellar
Biology Department
McGill University, Montreal
March 2016

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of MSc, Biology

ABSTRACT

Loss of retinoblastoma protein (pRB) function is one of the most common steps in cancer pathogenesis in humans. pRB regulates E2F-family transcription factors required for expression of genes that drive cell cycle progression. Consequently, understanding and developing treatments to reduce unregulated cellular proliferation in pRB-related cancers relies largely on identifying and exploiting factors that cooperate with *RB* mutations. The EGFR-regulated transcriptional repressor Capicua is one such factor which cooperates with loss of the pRB homologue Retinoblastoma-Related Factor-1 (RBF1) in *Drosophila melanogaster.* Loss-of-function *Capicua* mutations in the eye imaginal disc promote ectopic proliferation and survival of *Rbf1* mutant cells specifically. For this reason identifying genes directly targeted by Capicua de-repressed in the context of an *Rbf1* mutant background may identify novel factors responsible for survival and proliferation of cancer cells.

Despite well characterized transcriptional repression during embryogenesis and imaginal disc development, there is little evidence of directly bound Capicua target genes in *Drosophila*. RNA-sequencing data obtained from eye imaginal discs was used to identify possible direct targets of Capicua. These candidate genes were identified by analysis of mRNAs commonly upregulated in *Capicua* mutant discs, in both a wild-type and *Rbf1* mutant background. Investigation of Capicua's DNA-binding activity using Chromatin Immunoprecipitation (ChIP) assays, the gene *Phosphoribosyl Pyrophosphate Synthetase* (PRPS) was identified as a direct target. Direct binding evidence by ChIP was achieved by the generation and Gal4 driven expression of a Capicua cDNA construct. Overexpression of Capicua also provided additional insights into the general effects this transcription factor has on cell cycle progression.

Investigation of Capicua's recruitment to polytene chromosomes concluded that constitutive expression throughout development is necessary to exhibit stable DNA-binding activity. When this requirement is met salivary gland and polytene chromosome size is reduced, consistent with interruption of endocycle progression. This supports prior studies citing the role of *Capicua* in cellular proliferation and is consistent with direct binding and likely repression of the nucleotide biosynthetic gene *PRPS*. Furthermore, identification of a direct interaction with *PRPS* suggests this factor is implicated in the survival and proliferation of RBF1 deficient cells.

ABSTRAIT FRANÇAIS

La perte de la fonction de la protéine du rétinoblastome (pRB) est l'une des étapes les plus courantes dans la pathogenèse du cancer chez les humains. Le pRB régule les facteurs de transcription du E2F-famille requis pour l'expression de gènes qui stimulent la progression du cycle cellulaire. Par conséquence, la compréhension et le développement de traitements pour réduire la prolifération cellulaire non réglementée dans les cancers pRB liés dépend en grande partie sur l'identification et les facteurs qui coopèrent avec les mutations exploitation *RB*.
Le répresseur transcriptionnel EGFR régulé Capicua est un tel facteur qui coopère avec la perte de l'homologue rétinoblastome pRB liés Factor-1 (RBF1) dans le *Drosophila melanogaster*.
La perte de fonction mutation *Capicua* dans le disque imaginal des yeux favorisent la prolifération ectopique et la survie des cellules mutantes *Rbf1* spécifiquement. Pour cette raison, l'identification des gènes directement ciblés par le Capicua de réprimés dans le cadre d'un fond mutant *Rbf1* peut identifier de nouveaux facteurs responsables pour la survie et la prolifération des cellules cancéreuses.

Malgré une excellente répression transcriptionnelle caractérisée pendant l'embryogenèse et le développement du disque imaginal, il y a peu de preuves de gènes cibles Capicua directement liés chez la *Drosophile*. Les données d'ARN-séquençage obtenus à partir de disques imaginaux de l'oeil ont été utilisé pour identifier les cibles directes possibles du Capicua. Ces gènes candidats ont été identifiés par l'analyse des ARNm couramment surexprimés dans les disques mutants Capicua, à la fois de type sauvage et le fond mutant *Rbf1*. Une enquête sur l'activité de liaison à l'ADN du Capicua utilisant l'Immunoprécipitation de la chromatine (ChIP),

le gène phosphoribosylpyrophosphate synthétase (*PRPS*) a été identifié comme une cible directe. La liaison directe par ChIP a été obtenue par la production et l'expression Gal4 d'une construction Capicua ADNc entraînée. La surexpression du Capicua a également fourni des indications supplémentaires sur les effets généraux que ce facteur de transcription a sur la progression du cycle cellulaire.L'enquête sur le recrutement de Capicua aux chromosomes polytènes a conclu que l'expression constitutive au cours du développement est nécessaire pour présenter une activité de liaison à l'ADN stable. Lorsque cette condition est remplie la glande salivaire et la taille des chromosomes polytenes est réduite, conformément à l'interruption de la progression de endocycle. Cela confirme les études antérieures citant le rôle du Capicua dans la prolifération cellulaire et est compatible avec la répression contraignante et susceptible directe des *PRPS*. En outre, l'identification d'une interaction directe avec *PRPS* suggère que ce facteur est impliqué dans la survie et la prolifération des cellules déficientes RBF1.

AUTHOR CONTRIBUTIONS

Dr. Nam-Sung Moon designed and supervised this study, and first investigated the cooperation between *Rbf1* and *Capicua* mutations, promoting survival and proliferation of RBF1 deficient cells in *Drosophila* eye imaginal discs. Gene candidates investigated during this project were identified using RNA sequencing work performed by previous lab member Kate Krivy. I contributed all figures and data presented in this thesis. French translation of my abstract was written by Adrienne Wolfe

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor Dr. Nam-Sung Moon for welcoming me to the lab as his graduate student, pushing me to improve my skills as a researcher, and challenging me with this project. My experience in his lab has instilled confidence and excitement to continue my career in science. I would also like to thank all the members of the Moon lab, Heather Collins, Mary-Rose Bradley-Gill, Josee Houde, Christine Yergeau, Minhee Kim and Dr. Danny Feingold for their assistance and advice on experimental procedures and individual contributions to a exceptionally positive research environment. I also want to extend my thanks to everyone in the Lasko lab for their advisement throughout the course of my thesis work, for the occasional fly stock, and for sharing reagents with myself and the entire Moon lab. Thank you to Mr. Lee for preparing our fly food, to Beili Hu for injection of transgenic constructs, and to the CIAN staff Dr. Elke Küster-Schöck and Dr. Guillaume Lesage for confocal microscopy training. Special thanks to Dr. Frieder Schöck and Dr. Francois Fagotto, members of my graduate committee that provided invaluable advice concerning the course of my research project.

TABLE OF CONTENTS

INTRODUCTION

One of the most well researched cell cycle regulators is pRb, which functions to inhibit E2F-dependant transcription (reviewed in [1,2]). Once bound to E2F, pRb inhibits transcription of targets required for a cell's timely transition from G0/G1 to S phase (reviewed in [3]). Mutations causing inactivation of pRB are central to cancer development, as they lead to a loss of cell cycle regulation through unregulated E2F activity [4,5]. Our lab works to identify factors which genetically interact with *Rbf1* mutations using *Drosophila* development, in particular the eye imaginal disc development, as a tool. Using this model the transcriptional repressor Capicua (*CIC*), a downstream effector of RTK signalling, was identified. Like *RB*, *CIC* has been identified as a possible tumour suppressor gene and is found commonly mutated in oligodendrogliomas [6] and Ewing-like sarcomas [7].

Previous work determined that *Rbf1* hypomorphic mutations coupled with *Capicua* loss-of-function mutations promoted not only survival but also increased ectopic proliferation of cells within the eye disc morphogenetic furrow (MF), which are normally arrested in G0/G1 [8]. This novel genetic interaction situated *Capicua* as a strong candidate for further study to identify what direct targets lead to survival and proliferation of *Rbf1* deficient cells. To identify potential direct targets of Capicua in this genetic context, RNA-sequencing data obtained from eye imaginal discs provided a list of genes regulated by *Capicua* in an *Rbf1 mutant* background. By investigating these candidate genes using the direct DNA-binding assay chromatin immunoprecipitation (ChIP), I sought to determine which genes or gene classes are directly bound by Capicua in order to better understand factors that contribute to the survival and proliferation of *Rbf1* deficient cells.

Background

Retinoblastoma protein regulates cell cycle by inhibiting E2F-dependant transcription

pRB, p130, and p107 are members of the pocket protein family critical to cell proliferation and survival (reviewed in [9]). Pocket family proteins (also known as Rb-family proteins) are conserved thought a wide range of species and are central cell cycle regulators in both mammals and metazoans (reviewed in [3]). These three proteins all contain a "pocket" domain required for interaction and inhibition of E2F family transcription factors, best known to target genes required for G1-S phase transition (reviewed in [3,10]). Rb-family proteins inhibit E2F-dependant transcription by either direct inactivation [1], or by recruitment of chromatin remodelling complexes (reviewed in [11,12]). There are 8 distinct E2F proteins in mammals which can be roughly classified into transcriptional activators (E2F1-3a) and repressors (E2F3b-8), though there is functional redundancy and variation in specific pocket protein affinity within these groups (reviewed in [3,13]). E2Fs bind DNA in a heterodimer with DP, except in the case of atypical E2F7-8, which have been identified in regulating DNA damage response genes, and bind target genes independently or as heterodimers [14,15]. Additionally there is a unique interaction between E2F1 and pRB that is functionally and molecularly different than other pocket protein and E2F complexes [16,17]. This complex is resistant to pRB inactivation by CDK phosphorylation, and changes E2F1's consensus sequence specificity, enabling regulation of genes required for apoptosis separate from cell cycle dependant control [17]. Together these findings outline the complex interactions between pocket proteins and E2F family proteins.

Loss of Retinoblastoma function is central to cancer development

Mutations in *RB* affecting pRB's regulatory control of E2F family proteins is a common genetic abnormality associated with cancer pathogenesis in humans [18]. pRB inactivation either by direct mutation of functional domains, or changes in activity of upstream regulators such as CDKs or CDK inhibitors p16/INK4A results in unregulated E2F-dependant transcription (reviewed in [5]). Because E2Fs regulate genes critical to cellular proliferation, runaway E2F activity commonly leads to tumorigenesis [4,5]. While heterozygous germline mutations in *RB* are not sufficient to induce retinoblastoma in humans, heterozygosity does increase overall cancer susceptibility, as somatic mutation of the remaining wild-type allele can trigger cancer development (reviewed in [19]). This is supported by mouse models where heterozygosity is sufficient to induce both thyroid and pituitary tumours, [4,19] while homozygous *RB* null mutants exhibit widespread cellular proliferation and apoptosis resulting in lethally early in development. Functional inactivation of E2F1 in a *RB* null background reduces this severity, supporting the notion that pRB induced tumorigenesis is a result of upregulated E2F transcriptional activation [19].

Evidence brought forth within recent years questions the simplicity of this model due to the fact that *RB* mutant mice unable to bind E2Fs through the canonical "general" interaction (distinct from specific E2F1 interaction) do not develop cancer [16,20]. This can be explained in part by the function of p107 and p130 which exhibit tumour suppressor properties or regulatory function of pRB outside of E2F transcriptional control [20]. This however has not been fully investigated, and pRB is still described as the strongest tumour suppressor in the pocket protein family. Discrete roles for each E2F activator and repressor must also be taken into account.

These discrete roles are best characterized by E2F1 and E2F3. While E2F1 mutants develop lymphoma independent of pRB, E2F3 inactivation in *RB* heterozygous mutant mice suppresses pituitary tumours while simultaneously inducing pathogenesis of thyroid tumours, suggesting tissue-specific roles (reviewed in [19]).

The RB-E2F regulatory network in *Drosophila*

Clearly the interactions between pocket family proteins and E2Fs are complex, and research into their specific function is hindered by functional redundancy of E2Fs as well as their auto-regulatory effects as is the case with atypical E2F7-8 which act on E2F1 expression (reviewed in [16]). The interaction network between these proteins is simplified in *Drosophila melanogaster*, which contain only two E2F homologs, activator dE2F1 (bound by Retinoblastoma-related factor 1, RBF1) and repressor dE2F2 (bound by RBF2). RBF1 functions to suppress over-proliferation and apoptosis when dE2F1/dDP is ectopically expressed in eye imaginal discs, and retains negative regulation by CycE/CDK kinase activity (reviewed in [13]). *Drosophila* also contains a CDK regulator Dacapo, homologous to p21, suggesting the regulatory network pRB is central to is conserved in flies [13].

Like their mammalians homologues, dE2F1-2, dDP, and RBF1-2 are central to cell cycle control and required for organ development in *Drosophila melanogaster*. dE2F1 functions to promote expression of genes required for S-phase entry including those involved in apoptosis, while these targets are similarly repressed in the presence of dE2F2 (reviewed in [13]). Likewise, early larval lethality via a lack of DNA-synthesis can be prevented to a degree by simultaneously introducing dE2F2 loss of function mutations in dE2F1 mutants, suggesting the function of E2F-transcriptional activators and repressors is conserved in *Drosophila* [21]. Furthermore, like *RB*

homozygous mutant mice, *Drosophila* embryos deficient for RBF1 experience ectopic S-phase entry and increased incidence of apoptosis characteristic of dE2F1 hyper activation, leading to lethality early in larval development (reviewed in [13]).

As stated earlier, the simplified *RB* regulatory network in *Drosophila melanogaster* conserves mammalian function in regulating cell cycle progression and apoptosis. This system is useful in reducing confounding functional redundancy of E2F transcription factors when studying *RB* function, and represents an invaluable system to study relationships between this regulatory network and factors that impact survival of pRB deficient cells.

The Capicua tumour suppressor

Capicua is a transcriptional repressor first identified in a screen for mutations affecting anteroposterior patterning in *Drosophila* embryos [22]. It has since been best characterized in regulating cellular proliferation in both *Drosophila* and humans, and is highly conserved throughout both metazoans and mammals [23-25]. Its mammalian ortholog CIC has been identified as a possible tumour suppressor, as it has been implicated in the pathogenesis of oligodendrogliomas (ODs) [6], the neurodegenerative disease spinocerebellar ataxia type 1 (SCA1) [26], and Ewing-like sarcomas [7]. The most common mutations found in ODs affect CIC's high mobility group box (HMG-Box) domain required for DNA-binding activity, suggesting it's function as a transcriptional repressor is compromised [6]. Furthermore, CIC's implications in Ewing-like sarcomas and several other cancers results in part from a chromosomal translocation resulting in a chimeric CIC-DUX4 (Double homeobox protein) protein, changing it's function to that of a transcriptional activator [7] upregulating the PEA3 subfamily of ETS transcription factors found in several types of cancers. Altered CIC

transcriptional repression is also plays a role in pathogenesis of SCA1 (reviewed in [24]). CIC has been found to form complexes with ataxin-1, (AXTN1) a polyglutamine-repeat protein thought to function as a co-repressor due to it's lack of sequence-specific action [26] and it's association with histone deacetylases (HDAC3-4) [27]. Mutations in mice causing an expanded polyglutamine tract in AXTN1 reduced ATXN1-CIC complex formation, causing derepression of some CIC target genes, while conversely increasing repression of others [27].

Consistent with it's role in tumorigenesis, studies in *Drosophila* have solidified *Capicua* as a regulator of cellular proliferation in addition to it's well described role in embryonic patterning. In the eye imaginal disc, mosaic clones expressing a truncated Capicua protein lacking it's HMG-box domain exhibit a proliferative advantage over wild-type tissue once development is complete [28]. Additionally, loss of Capicua's DNA binding domain, the HMG-Box, in this context does not affect cell fate determination as evidenced from sound ommatidial patterning. This suggests that it acts to repress classes of genes specific to cell cycle progression [28], although there is evidence *Capicua* has a fate determination role in wing discs.

Capicua Transcriptional repression is dependant on RTK signalling

In addition to the HMG-Box DNA binding domain Capicua contains several other highly conserved regulatory motifs that dictate it's activity, including C-terminal motifs C1 and C2. The C1 motif is unique to Capicua and is responsible in part for its function as repressor, however insights into its molecular mechanism of action remain unclear [23,24]. The C2 domain, also unique, is required for regulation by Receptor Tyrosine Kinase (RTK) signalling pathways. When bound by appropriate ligands RTKs Torso and epidermal growth factor receptors (EGFR) transduce their signal through Ras/Raf/MEK and finally MAP kinase (MAPK), which directly

interacts with the C2 motif [23,29]. This interaction is necessary for phosphorylation of Capicua, marking Capicua for degradation and cytoplasmic localization [23,30] and is conserved between both flies and humans [6,31]. Capicua gain of function mutations lacking this domain remain constitutively active and mostly unresponsive to RTK signalling, leading to a variety of patterning defects during embryogenesis [23]. Degradation of Capicua thus leads to de-repression of terminal gap genes important in embryonic patterning such a *tailless* (tll) and *huckebein* (hkb) in the case of RTK (Torso) signalling at the embryonic termini, and *argos* in the case of dorsal EGFR signalling during oogenesis and wing development [22,32,33]. Degradation of Capicua induced by phosphorylation of its C2 domain has also been described in response to activated EGFR and Ras in the imaginal eye disc [28], in intestinal stem cells (ISCs) [25,34] and wing discs [33].

The HMG-Box also plays a role in RTK regulation, however it's exact function has not been completely explored. HMG-Box domains are known to exhibit DNA-binding activity, [35] however Capicua binds altered sequences (T(G/C)AATG(A/G)A) compared to (CC(T/A)TTG(T/A)) bound by other HMG-box proteins (reviewed in [7,24]), and shares low homology [24,35,36]. Moreover, Capicua does not contain a recognizable nuclear localization signal as it is incorporated into the HMG-box [23]. Consistent with this finding, Capicua lacking it's HMG-box domain remains cytoplasmic, non-functional, and resists regulation via torso signalling suggesting it plays a role in RTK regulation in addition to the C2 domain [23] even when an artificial nuclear localization signal is added. These findings situate the HMG-Box as Capicua's most important functional domain.

While not directly related to RTK regulation it is worth noting the function of *Capciua*'s two main isoforms and recently identified N2 motif. The "long" isoform (Capicua-L) contains an extended N-terminal domain and an N1 motif of unidentified function (reviewed in [24]). The "short" isoform (Capicua-S) contains the N2 motif unique to metazoans and is required for repression of targets such as *tll* and *hkb* during embryogenesis [37]. The co-repressor groucho (Gro) is required for Capicua's function in regulating terminal gap genes (*tll* and *hkb*) during the embryonic stage and it has been suggested the N2 domain facilitates this interaction, however the molecular mechanism is not clear, and physical interaction has yet to be demonstrated [38]. Moreover, both the N2 domain and Gro are dispensable for repression of Capicua targets post-embryogenesis, supporting a specific role for Gro association [37]. Despite both isoforms containing an nuclear localization signal in mammals, CIC-S is found proximal to mitochondria in mammalian oligodendrogliomas, while CIC-L is nuclear [39]. In flies both Capicua-L and Capicua-S display nuclear localization (reviewed in [24])

Structure and development of the *Drosophila* eye

The compound adult eye is made up of several hundred repeating units known as ommatidia [40]. Each ommatidia develops to contain four cone cells which secrete the lens, two primary pigment cells [41 and eight photoreceptor neurons (R1-R8) two of which are centrally placed (R7-R8) to create an asymmetrical but precisely patterned array of light-sensitive rhadobmnere [42]. Each ommatidia then shares three mechano-sensory bristles, six secondary pigment cells, and three tertiary pigment cells with its direct neighbour to create a hexagonal lattice (reviewed in [43]). All of these distinct cell types arise from epithelial cells which undergo

a set of highly invariant patterning events during larval development in the eye-antennal disc, also known as the eye imaginal disc (reviewed in [43]).

The eye imaginal disc forms from an invagination of the ectoderm during embryogenesis [44]. This epithelial monolayer consists of cells that proliferate asynchronously throughout larval development until the mid-third instar stage in a process known as the first mitotic wave [43,44]. Cells in the anterior region continue to asynchronously divide while differentiation of cells into ommatidial precursors begin at the posterior-most end of the eye disc [45]. This neural differentiation is marked by a dorsoventral constriction of the retinal epithelium at it's apical surface known as the morphogenetic furrow (MF) [40,43,46]. The MF moves anteriorly across the disc, initiating precursor differentiation beginning during the third instar stage and ending approximately ten hours after puparium formation (reviewed in [43]). As the furrow passes, cells previously undergoing asynchronous division begin to arrest in G1 in a dorsoventral band anterior to the MF referred to as the non-proliferative region (NPR) [40]. Meanwhile, cells entering the furrow itself begin to differentiate into R8 photoreceptor neurons [42,47]. R8 photoreceptors subsequently organize into a five cell pre-cluster by recruiting R2 and R5 followed by R3 and R4, in a pairwise fashion by inducing uncommitted cells as the furrow passes [42,45]. Cells not included in this five cell ommatidial pre-cluster undergo an additional round of mitosis directly posterior to the passing MF known as the second mitotic wave [42,43]. The second mitotic wave is necessary to create additional cells that will become accessory pigment cells, cone cells, the four-cell bristle complex, and the remaining photoreceptor neurons R1/R6 and R7 (reviewed in [48]).

The transition from proliferation to cell cycle arrest and neuronal differentiation throughout the NPR and MF makes the eye imaginal disc an ideal system to study factors required for cell cycle control and transition to arrest in G1. This is particularly true when investigating factors that co-operate with RBF1. Cell cycle arrest always occurs before differentiation in the eye imaginal disc [49]. Preceding terminal differentiation is the MF, where presumptive R8 photoreceptor cell fate is specified [49]. Differentiating R8 cells within the MF secrete several factors including Hedgehog (Hh), and Decapentaplegic (Dpp) [50,51] that induces cell cycle arrest in the NPR [45,52]. Cell cycle arrest is then largely maintained by RBF1, which functions to inhibit dE2F1-dependant transcription of the pro-apoptoic gene *hid* [53] as well as genes required for the G1/S-phase transition. Characteristic of aberrant dE2F1-dependant transcription, loss of RBF1 within the MF leads to cell death and ectopic S-phase cells, similar to phenotypes observed when pRB function is lost in mammalian systems [3,8,54]. While this evidence underscores the importance of RBF1 in maintaining G1 arrest within the furrow, reports by [54] have shown that the EGFR signalling gradient present in the MF is critical in determining whether RBF1 deficient cells undergo cell death or enter S-phase. Active EGFR signalling promotes survival of RBF1 deficient cells in the anterior half of the MF where cells first begin to arrest in G1, in part due to post transcriptional regulation of Hid [3,54,55]. This finding supports a role for EGFR signalling in cooperating with *Rbf1* hypomorphic mutations to predispose cells to either dE2F1-dependant apoptosis [54,56] or exit of cell cycle arrest.

Genetic interaction between *Rbf1* and *Capicua*

Our lab's initial investigation of Capicua stems from the findings published by Moon et al. [54] which investigated the role EGFR/Ras signalling plays in protecting *Rbf1* hypomorphic cells from dE2F1-dependent apoptosis [54]. This led to investigation of factors downstream of EGFR signalling that might affect gene expression responsible for the survival of *Rbf1* mutant cells. Consistent with this finding, mutations in the EGFR-regulated transcriptional repressor Capicua were then identified by our lab to exhibit a novel genetic interaction with *Rbf1* mutations [8]. As stated above, cells undergo developmentally programmed cell cycle arrest in G1 within the MF prior to neuronal differentiation, [49] and de-regulated dE2F1 transcriptional activation resulting from loss of RBF1 function leads to both apoptosis and ectopic S-phase entry within the MF [8,54]. Functional inactivation of Capicua in RBF1 hypomorphic eye discs protect against cell death normally seen in RBF1 hypomorphic tissue, specifically within the anterior half of the MF where cells begin to arrest [8]. Additionally, the incidence of ectopic proliferation within the normally arrested cells in the posterior half of MF increased, representing a previously uninvestigated role for *Capicua* in regulating survival and proliferation of *Rbf1* mutant cells [8]. Furthermore, the effects of unbridled cellular proliferation in these mutants are carried through development, leading to excess inter-ommatial cells observed in pupal eye discs, and lead to ommatidial patterning defects in the mature adult eye [8] more severe than those observed in *Rbf1* single mutant adult eyes.

Function and development of salivary glands

Salivary glands are large tubular organs that actively secrete enzymes required for digestion throughout larval life, as well as glue-like proteins [57] required for attachment during puparium formation. They consist of two major cell types, secretory cells and the duct cells that

connect them to the larval mouth [57]. Arising from two ectodermal plates of about one hundred cells each, salivary glands develop without further cell division, increasing in size coincident with several rounds of endoreduplication [58]. Secretory cells invaginate first and begin to form a tube which migrates posterior of the presumptive mouth and ventral of the nerve cord [57]. Duct cells invaginate last, while secretory cells begin to alternate between G and S-phases without undergoing cytokinesis, creating giant polytene chromosomes necessary to maintain high metabolic activity [57,59]. Secretion ultimately begins in late embryogenesis as the salivary glands reach their most posterior position and continues until pupa formation [57].

Regulation of endoreduplication

In *Drosophila* larvae the majority of tissues grow coincident with successive endocycles including the digestive system, fat bodies, salivary glands, and many cells essential to oogenesis (reviewed in [60]). These cells likely take on a endocyclic program to regulate organ size and organization. Additionally, endocycling cells increase their metabolism through successive rounds of DNA-synthesis, as increases in gene copy number correlate with higher transcriptional activity [58].

Central to endocycle control is regulation of DNA-synthesis. Formation of pre-replication complexes (pre-RCs) in G phase is a carefully regulated process, as reloading of pre-RC proteins onto chromatin can result in what are known as re-replication events [61] which can potentially lead to cancer development. Like mitotic cells, Cyclin E (CycE) and its kinase partner Cyclin-Dependant Kinase 2 (Cdk2) are required for S-phase entry in endocycling cells. Unlike mitotic cells, oscillation of CycE/Cdk2 levels are required for alternating S and G phases (reviewed in [62]). Pre-RCs begin to form in G phase when CycE/Cdk2 levels are low [61]. As CycE/Cdk2 is

transcribed by the dE2F1 activator complex, positive feedback occurs as Cdk2 phosphorylates the dE2F1 inhibitor RBF1 [63]. This creates a peak in CycE/Cdk2 required for the transition to S-phase [61]. High CycE/Cdk2 levels indirectly finalize pre-RCs by regulating factors that recruit the loading of Mcm2-7 DNA helicase onto chromatin (reviewed in [60]), while simultaneously inhibiting pre-RC formation directly (reviewed in [62]). As DNA replication begins, chromatin-bound PCNA activates the E3 ubiquitin ligase CRL4-CDT2 that degrades dE2F1. Subsequently, inherent instability and degradation causes CycE levels to drop in the absence of dE2F1-dependant transcription, setting up conditions necessary for pre-RC formation in the next gap phase [64,65].

Studies of the developmentally programmed switch from mitotic growth to endocyclic growth in follicle cells of the *Drosophila* ovary have determined that Notch signalling is crucial in allowing cells to bypass mitosis altogether [66]. While lack of Notch signalling promotes extra rounds of mitosis [67], expression of Notch ligand delta in germline cells promotes early endocycle entry in follicle cells [67,68]. Activation of Notch signalling promotes activity of the transcription factor Hindsight (Hnt) which in turn represses transcription of pro-mitotic phosphatase String (*Stg*, homologous to mammalian Cdc25) and the transcription factor Cut [69,70]. While *Stg* promotes activity of Cyclin A (CycA) and B (CycB) required for the G2-M transition, Cut inactivation allows for expression of fizzy-related (Fzr), an activator of anaphase promoting complex/cyclosome(APC/C) enzymatic activity, required for phosphorylation and degradation of pro-mitotic CycA and CycB [71-73]. It can be argued that the importance of Notch signalling in endocycling cells is largely due to regulation of APC/C, as loss of APC/C activity via Fzr inhibition leads to a loss of DNA synthesis in secretory cells and results in

severely reduced salivary gland size [74]. The APC/C's role in endocycle regulation is further supported by reports that mutation of subunits APC2 (encoded by *morula*) or Cdc16 subunits result in reduced salivary gland and polytene chromosome size [61,73].

CycE is also tightly linked to APC/C function. High CycE/cdk2 levels also inhibit Fzr activation of APC/C, relieving inhibition of the pre-RC inhibitor Geminin, aiding in stopping re-replication events from occurring as DNA synthesis proceeds (reviewed in [60]). As CycE/cdk2 levels drop off during DNA replication APC/C inhibits Geminin, enabling formation of pre-RCs in the following gap phase [61].

Polytene chromosome structure

Giant polytene chromosomes arise from successive rounds of DNA replication without cell division. In the case of salivary gland cells this process results from approximately ten endocycles (reviewed in [60]). The copy number of each chromosomes can be modelled exponentially, while diploid cells are referred to as 2C (two copies of each chromosome) cells of the salivary glands would be 1024C after ten round of replication or 2^{10} [75], although studies have shown that actual copy number is closer to 1350C [60]. Due to the high copy number these chromosomes are large enough to be visualized by light microscopy and are useful for examining relations between transcriptionally active regions and chromatin-bound proteins. The most commonly used visualization method is a squash method, where nuclei of the salivary gland (or other tissue such as follicle cells) are pressed between a coverslip and glass to spread out the chromosomes prior to staining [76]. This results in an experimentally useful two-dimensional structure, though the complex three-dimensional structure is not easily apparent. Elucidated in [75] these structures have a highly ordered looping structure creating a cylindrical DNA-free axis

at its centre analogous to a hollow tube. Discrete genetic regions are made up of successive loops creating bright bands when visualized with Hoechst or DAPI stains, while DNA linking these loops account for less visible inter-band regions [75]. Furthermore, the tightly packed loops that stain brightly are less transcriptionally active compared to less densely packed inter-band regions, which facilitate transcription [76].

History and use of Chromatin Immunoprecipitation

Determining precise DNA-interactions and genomic localization of transcription factors is central to understanding their effect on gene expression. As is the case with Capicua, several transcription factors have historically been characterized in regulating gene expression without strong biochemical evidence of DNA interaction or data concerning precise DNA binding patterns. This paradigm has changed in recent years with the advent of increasingly precise methods of quantitive DNA-protein interaction assays such as chromatin immunoprecipitation (ChIP), DNA adenine methyltransferase identification (DAMID), and a variety of computational methods used to predict DNA-protein interactions[77,78]. Techniques developed in the 1970s such as DNA-footprinting and electrophoretic mobility-shift assays (EMSA) have been modernized to facilitate quantification of DNA-bound protein, particularly in the case of modified EMSA techniques, although these assays are restricted to providing in vitro evidence which may not completely replicate protein function in vivo [78].

ChIP was originally developed to investigate the relationship between transcriptionally active gene regions and their associated histone modifications (reviewed in [77]). Wide use of ChIP then began in the 1990s to study DNA-binding patterns of transcription factors in virtually all model organisms. While ChIP has been adapted for use with a variety of cell types and

sequencing methods, it follows a common method. The proposed DNA-interacting protein of interest is targeted by a protein or epitope-specific antibody and immunoprecipitated along with its associated DNA fragments from sonicated chromatin extract. The chromatin extract prepared is generally cross-linked using a reversible agent such as formaldehyde, however ChIP can be modified using crosslinkers that allows for immunoprecipitation of more loosely associated proteins, or conversely performed in an absence of crosslinking agents to preferentially immunoprecipitate tight chromatin bound proteins such as histones [79,78]. Once purified, DNA fragments bound by the immunoprecipitated target can be analyzed by a variety of methods. These include genome-wide analysis by microarray hybridization, construction of a ChIP-sequencing library, or by Quantitive-Real-time PCR (qPCR) if a relatively narrow number of proposed target loci are known, as is the case when investigating Capicua regulated genes in an *Rbf1* mutant background.

Direct target genes bound by Capicua have not be determined by ChIP in *Drosophila*. A recent study however has used the DAMID technique to identified several Capicua bound targets in *Drosophila* midgut cells [25]. This technique requires expression of a protein of interest fused to DNA adenine methylase (dam), an enzyme specific to prokaryotic cells [78,81]. Genes bound by the fusion protein are them marked by methylation via dam activity and can be preferentially cleaved, purified and analyzed by either qPCR or genome-wide sequencing techniques as fragments immunoprecipitated in ChIP assays can be. Both ChIP and DAMID have unique advantages. ChIP assays are reliant on use of antibodies specific to a target protein or epitope, but have an advantage over DAMID as native protein function can be conserved *in vivo*. While generation of the required dam fusion protein may affect in vivo function, DAMID is not reliant

on use of antibodies or any crosslinking agents as ChIP assays are. For the purposes of this study, ChIP was determined to be most effective due to it's flexibility in investigating a wider variety of protein and cell types than DAMID [78] and conservation of native protein function.

Project Rationale

The aim of my thesis work was to determine genes directly bound by the transcriptional repressor Capicua in order to better understand what genes or groups of functionally similar genes might be responsible for it's novel genetic interaction with *Rbf1* identified previously by our lab [8]. While mutations compromising activity of Capicua itself are implicated in several disease pathways, investigations into Capicua's function are of particular interest due to it's cooperation with *Rbf1* mutations in promoting survival and proliferation of RBF1 deficient cells when inactivated. Therefore, genes bound by Capicua then de-repressed in the context of an *Rbf1* mutant background may also turn out to be critical therapeutic targets for halting tumour formation and progression. To this end I investigated recruitment of endogenous and epitope tagged Capicua to polytene chromosomes and assayed its direct interaction with candidate genes obtained from RNA-sequencing data, identifying loci bound by Capicua within the candidate target gene PRPS by ChIP-qPCR. I also characterized the effects of Capicua overexpression required for DNA-binding activity on development of endoreduplicating cells. I explain further the identification of candidate target genes, suitability of constructs investigated for use in ChIP, and use of polytene chromosomes and salivary gland cells to investigate requirements for stable DNA-binding activity of Capicua.

Identification of candidate Capicua target genes

In order to identify classes of genes potentially subject to direct transcriptional repression by Capicua we first investigated genes differentially regulated when Capicua is mutated. RNA-sequencing data determined 309 upregulated genes in Capicua mutant eye discs compared to wild-type eye discs (data not shown). Approximately 10% of these genes were determined by gene ontology (GO) term search to be involved in oxidation reduction processes, followed by genes involved in nucleotide biosynthetic processes accounting for 4.5% and production of precursor metabolites for 4.2%. Further analysis of RNA sequencing data determined GO enrichment of genes involved in redox processes and production of precursor metabolites is true for genes commonly regulated by Capicua in both a wild-type and RBF1 hypomorphic background. Many of these genes also contained Torso-Response Element (Tor-RE) like sequences targeted by the HMG-Box domain proximal to transcription start sites, suggesting these may be direct targets of Capicua. Upon inspection of these candidate genes using constructs described below and in Figure 1, the gene *Phosphoribosyl Pyrophosphate Synthetase* (*PRPS*) stood out as a promising Capicua target due to an HMG-Box consensus site within its primary transcription start site, in addition to pseudo-HMG-Box consensus sites within its promoter region (see Appendix A for sequence information). *PRPS* has surfaced in recent investigations of imaginal disc development screens [80], though has not been directly investigated in *Drosophila melanogaster*. It shares sequence and functional similarly with mammalian *PRPS2*, closely related to *PRPS1* responsible for production of Phosphoribosyl-1-pyrophosphate (PRPP) [82] and implicated in a number of human disorders [89]. The production of PRPP represents a rate-limiting step in purine synthesis [83,84] but is also central to pyrimidine biosynthesis [82] and therefore is central to nucleoside biosynthesis. *PRPS* is

unsurprisingly very highly conserved and ubiquitously expressed across a wide range of eukaryotes [85]. For these reasons we hypothesized that derepression of this gene may play a role in promoting ectopic S-phase entry in an RBF1 hypomorphic background by regulating de-novo nucleotide biosynthesis.

Determining construct suitability for use in chromatin immunoprecipitation

Direct Capicua target genes have not been identified in *Drosophila* until very recently [25] through use of DAMID assays. Consistent with a lack of literature reporting direct Capicua bound target genes, studies have cited low DNA-binding affinity of of HMG-Box proteins [19, 22,86]. This reveals a limitation of ChIP in requiring a particular threshold of actively chromatin bound protein in order for immunoprecipitated target genes to be detected by qPCR analysis, despite being one of the most widely used assays for this purpose [78]. Additionally, widespread RTK activity in the developing tissues of the *Drosophila* larvae [87] likely hinder robust DNA-binding of Capicua through MAPK-mediated phosphorylation and degradation [23,29]. Supporting this, EGFR/Ras signalling has been shown to directly reduce Capicua protein levels in several tissue types and developmental stages [28,34]. To address this issue I investigated effects of reducing EGFR function though Gal4 driven expression of dominant negative EGFR and RNAi constructs. Despite these efforts, reducing EGFR signalling activity through both constitutive and heatshock Gal4 drivers did not increase recruitment of endogenous Capicua to polytene chromosomes, or improve Cic-HA target loci enrichment determined through ChIP experiments (data not shown).

It is curious that despite nuclear localization, well characterized response to EGFR signalling, and widespread in vivo evidence of transcriptional repression, Capicua eludes

enrichment for previously reported and candidate target loci containing a high incidence of consensus binding sites during preliminary ChIP assays. In response to these issues I generated an N-terminally HA and Flag (HF) tagged CicRA construct (analogous to Capicua-S) lacking the recently described N2 domain not required for post-embryonic function (HF-CICRA$^{\Delta N}$, Fig.1). This construct contains an upstream activation sequence (UAS) for use with Gal4 drivers in an effort to increase protein expression level, overcome possible dosage-dependant MAPK-mediated degradation, and increase DNA-bound protein available for immunoprecipitation.

Investigating the relationship between Capicua and RBF1 by examining polytene chromosome binding patterns

Polytene chromosomes found in larval salivary glands provide a unique context to validate association of transcription factors with chromatin as they are easily visualized by light microscopy and adaptable to immunofluorescence studies [76]. Furthermore, these endoreduplicating cells develop a large nucleus due to high DNA content, providing easy confirmation of nuclear localization not immediately visible in other cell types. These cells provide an excellent system to not only initially investigate DNA-binding activity of both endogenous and overexpressed Capicua prior to use in ChIP assays, but also to examine the interaction between RBF1 and Capicua. Using a polytene squash method, the banding patterns of both Capicua and RBF1 can be compared to draw conclusions as to whether these transcription factors occupy similar genomic regions. This data would directly support common gene regulation as a mechanism by which *Capicua* mutations cooperate with *Rbf1* mutations to promote survival and cellular proliferation. We hypothesize that there would be at least some overlap in banding patterns between RBF1 and Capicua due to several co-regulated factors

discovered through RNA-sequencing data, however it would be interesting to investigate how common this occurrence actually is.

RESULTS

Cic-HA cannot be detected at target loci of previously reported Cic regulated genes

Mutations causing functional inactivity or loss of RBF1 exhibit well defined patterns of cell death in addition to ectopic S-phase entry within the MF [8]. When *Capicua* is mutated within an *Rbf1* mutant background, cells are protected against dE2F1-dependant apoptosis, while increasing ectopic S-phase entry. To investigate genes directly bound by Capicua, whose de-repression might contribute to survival and proliferation specific to RBF1 deficient cells, I first investigated the expression patterns of two constructs for use in ChIP-qPCR assays (Figs. 1-2). Cic-GFP is expressed in addition to endogenous Capicua and contains a C-terminal GFP tag present in six of the seven discrete isoforms (Fig. 2A). Cic-HA is similarly constructed but is composed of a Capicua-S cDNA clone fused to a genomic promoter. While Cic-GFP is expressed uniformly across the eye imaginal disc (Fig. 2B), it does not recapitulate endogenous Capicua expression patterning anterior to the MF as is seen in larvae expressing Cic-HA (Fig. 2C) [8,28]. As this construct displays the correct expression pattern within the eye disc and was confirmed to rescue lethality of *Capicua* homozygous null mutants used in [28], Chip-qPCR was performed using chromatin samples extracted from *Drosophila* third instar larvae (Fig. 3).We compared target loci enrichment of Cic-HA samples immunoprecipitated with either Anti-HA (Cic-HA-HA) or IgG (Cic-HA-IgG) in addition to control (YW) samples (YW-HA,YW-IgG). Cic-HA-HA samples (orange bars) failed to enrich for promoter loci within *tailless* and previously reported Capicua binding sites within the *argos* 1.0 enhancer element [38] (Fig. 3A). The enrichment

observed at the *argos* 1.0 enhancer locus in YW-HA and IgG samples is non-specific and is in fact non-significant compared to Cic-HA samples when tested by ANOVA ($p > 0.05$). Similarly, Cic-HA-HA samples failed to enrich for transcription start site (TSS) loci within the candidate *PRPS* when compared to mock (IgG), or YW control samples (Fig. 3B). We therefore conclude that Cic-HA does not stably bind consensus sites within the candidate gene *PRPS* or previously reported targets of Capicua repression.

Gal4 driven expression of HF-CICRA$^{\Delta N}$ reduces larval body size primarily in endoreduplicating cells

Cic-HA binding activity was undetectable through ChIP-qPCR, and efforts to reduce possible effects of degradation via EGFR signalling were ineffective (data not shown). This is likely due to the fact that Capicua is tightly regulated by a wide range of RTK signalling pathways inhibiting high protein level and stable DNA-binding activity, general requirements for successful ChIP-qPCR assays using a crosslinking agent such as formaldehyde to preferentially immunoprecipitate direct chromatin bound proteins [77,78,88]. A CicRA construct (analogous to Cic-S) containing an upstream activation sequence was generated to first investigate effects of Gal4 driven overexpression on nuclear localization and larval development (HF-CICRA$^{\Delta N}$, Fig. 4) prior to use in ChIP assays. This construct lacks the N2 domain not required for post-embryonic function, however contains all other functional domains, in addition to N-terminal flag and HA tags (Fig. 4A).

Interestingly, ubiquitin Gal4-driven expression of HF-CICRA$^{\Delta N}$ was found to reduce overall body size compared to control larvae (YW, Fig. 4B), however did not impact development time to third instar stage, or cause pupal lethality. Furthermore, as most larval

tissues develop through endoreduplication (reviewed in [60]), and no difference imaginal eye disc size was observed (data not shown), this observation suggested the effects of HF-CICRA$^{\Delta N}$ overexpression on larval size is largely due to endocycle inhibition. This is further suggested by the size reduction observed in the cells of the salivary glands (Fig. 4C) where HF-CICRA$^{\Delta N}$ displays strong nuclear localization. While duct cells look to develop normally, there is a severe size reduction in secretory cell size (Fig. 4D) consistent with interruption of endoreduplication [61,74]. This suggests ubiquitous HF-CICRA$^{\Delta N}$ expression throughout development impacts S-phase entry or completion, and supports Capicua's role in regulating cell cycle progression. Given this finding we sought to further characterize the effects of HF-CICRA$^{\Delta N}$ overexpression by investigating polytene chromosome recruitment prior to use in ChIP assays.

DNA binding of HF-CICRA$^{\Delta N}$ can be observed when it is expressed by a salivary gland specific Gal4 driver

While investigating recruitment of HF-CICRA$^{\Delta N}$ to polytene chromosomes, we took the opportunity to also investigate the DNA-binding activity of endogenous Capicua (Fig. 5). To facilitate this an affinity purified peptide antibody was acquired, targeting a C-terminal peptide motif common to all isoforms of endogenous Capicua. While displaying strong nuclear localization in wild-type (YW) salivary glands (Fig. 5A) endogenous Capicua staining did not coincide with either chromosomal DNA or RBF1, which display strong banding patterns and recruitment to chromosomes (Fig. 5B). This led to the conclusion that despite strong nuclear localization (Fig. 4C) HF-CICRA$^{\Delta N}$ may also not stably bind DNA. To investigate this HF-CICRA$^{\Delta N}$ was expressed using the salivary gland specific driver VT045325 Gal4 (Fig. 5C-F). Consistent with initial ubiquitin Gal4 driven expression, HF-CICRA$^{\Delta N}$ displayed strong

localization to the nucleus (Figure 5E) and reduced overall size of salivary gland cells compared to the control samples (Figure 5C). Furthermore, VT045325 Gal4 driven expression of HF-CICRA$^{\Delta N}$ co-stains with both DAPI and RBF1 on polytene chromosomes, suggesting this construct stably binds DNA (Figure 5F), however the expression level required for this results in much thinner chromosomes compared to controls (Fig. 5D). For this reason any overlap in the banding patterns of RBF1 and HF-CICRA$^{\Delta N}$ cannot be resolved, and whether these transcription factors globally occupy similar genomic regions cannot be determined by this method. Together this data suggests that Gal4 driven HF-CICRA$^{\Delta N}$ expression is appropriate for further investigation by ChIP-qPCR. Furthermore, a reduction in chromosome size likely due to reduced copy number is consistent thus far with the hypothesis that Capicua represses *PRPS* as identified by multiple HMG-Box consensus sites within its promoter region and is consistent with phenotypes described when endoreduplication is interrupted [61,74].

Gal4 driven HF-CICRA$^{\Delta N}$ expression binds transcription start loci within the candidate gene PRPS

The candidate gene *PRPS* was identified through RNA-sequencing data to determine genes commonly upregulated by Capicua in wild-type and *Rbf1* mutant eye discs. *PRPS* encodes an enzyme required to produce precursor metabolites for de-novo DNA-synthesis. Given the incidence of Capicua HMG-Box consensus sites within its promoter and first transcription start region, this gene stood out as a likely direct target of Capicua. Given prior investigation into DNA-binding activity of HF-CICRA$^{\Delta N}$ identified by its recruitment to polytene chromosomes, ChIP-qPCR was performed on whole third instar larvae expressing this transgene (Fig. 6). Chromatin extracts were prepared from both larvae expressing HF-CICRA$^{\Delta N}$ under the control of

a ubiquitin Gal4 driver, and larvae expressing only the Gal4 driver, identical to those investigated in Figure 4. Primer pairs were designed to assess enrichment of three transcription start sites (amplicons A, C and D, analogous to TSS1,2 and 3 in Fig. 3), two regions unlikely to be bound by Capicua (B and E), and the standard negative ChIP-qPCR control actin (Fig. 6A). TSS-1 contains a pseudo-Tor-RE previously reported as a Capicua HMG-box consensus sequence [24,38] in addition to two additional consensus sequences just upstream of this site, while TSS2 and TSS3 do not contain this consensus sequence. ChIP-qPCR results revealed significant enrichment of TSS-1 and TSS-2 sites when HF-CICRA$^{\Delta N}$ samples were immunoprecipitated with anti-HA (HF-CIC-HA, orange bars) compared to mock (HF-CIC-IgG, yellow bars) and YW controls (YW-IgG, YW-HA blue and green bars). This data, pooled from three biological replicates reveals that HF-CICRA$^{\Delta N}$ directly binds the nucleotide biosynthetic gene *PRPS* at two of three transcription start sites. Non-significant enrichment of additional candidates identified from previously described RNA-sequencing data was detected by qPCR (data not shown), suggesting that binding of Capicua is specific to *PRPS* in an *Rbf1* mutant and may represent preferential binding to nucleotide biosynthetic genes in general, however ChIP-sequencing data would be required to confirm this notion.

Constitutive and widespread expression of HF-CICRA$^{\Delta N}$ is required for recruitment to polytene chromosomes and reduction of salivary gland cell size

When HF-CICRA$^{\Delta N}$ is constitutively overexpressed polytene chromosome size is reduced, making discrete banding patterns of RBF1 and Capicua difficult to resolve as was observed in figure 5. In an effort to conserve polytene chromosome size in order to better resolve the banding pattern and investigate incidence of commonly bound genomic regions by RBF1 and

Capicua, GFP^{+} marked clones over expressing HF-$CICRA^{\Delta N}$ were generated (Fig. 7). These clones were generated by recombining out a CD2 stop cassette via heat-shock driven flippase early during larval development. This allowed for actin-Gal4 driven expression of HF-$CICRA^{\Delta N}$ in a subset of salivary gland cells. Salivary glands and polytene chromosomes retained wild-type size and while there is a lower incidence of GFP^{+} marked clones compared to control cells, there is no observable difference in total cell volume or nucleus size (Fig. 7A). Furthermore, examination of polytene chromosomes within these cells concluded that despite high HF-$CICRA^{\Delta N}$ expression level and nuclear localization, no recruitment to chromatin or robust banding pattern is observed (Fig. 7B). This suggests stable DNA-binding activity of Capicua is dependant on more than simply high expression level, and is subject to constitutive expression throughout larval development.

Construct	Promoter	Tag	Type
Cic-GFP	Genomic	C-terminal GFP	Genomic Fragment (Cic-L/Cic-S)
Cic-HA	Genomic	C-terminal HA	cDNA (Cic-S)
HF-CICRA$^{\Delta N}$	Upstream Activation Sequence	N-terminal Flag/HA	cDNA (Cic-S)

Fig. 1 Summary of Capicua constructs investigated

Cic-GFP and Cic-HA are expressed under the control of a genomic promoter. Cic-GFP is a genomic fragment containing a GFP tag at it's C-terminus, accounting for 6/7 discrete *Capicua* transcripts which include both Capicua-L and Capicua-S isoforms. Cic-HA is composed of a cDNA clone of Capicua-S and contains an HA tag at its C-terminal end. HF-CICRA$^{\Delta N}$ was generated from the LD17181 cDNA clone, analogous to Capicua-S, however lacks the recently characterized N2 domain required for embryonic function. It contains a UAS promoter to facilitate Gal4-driven overexpression.

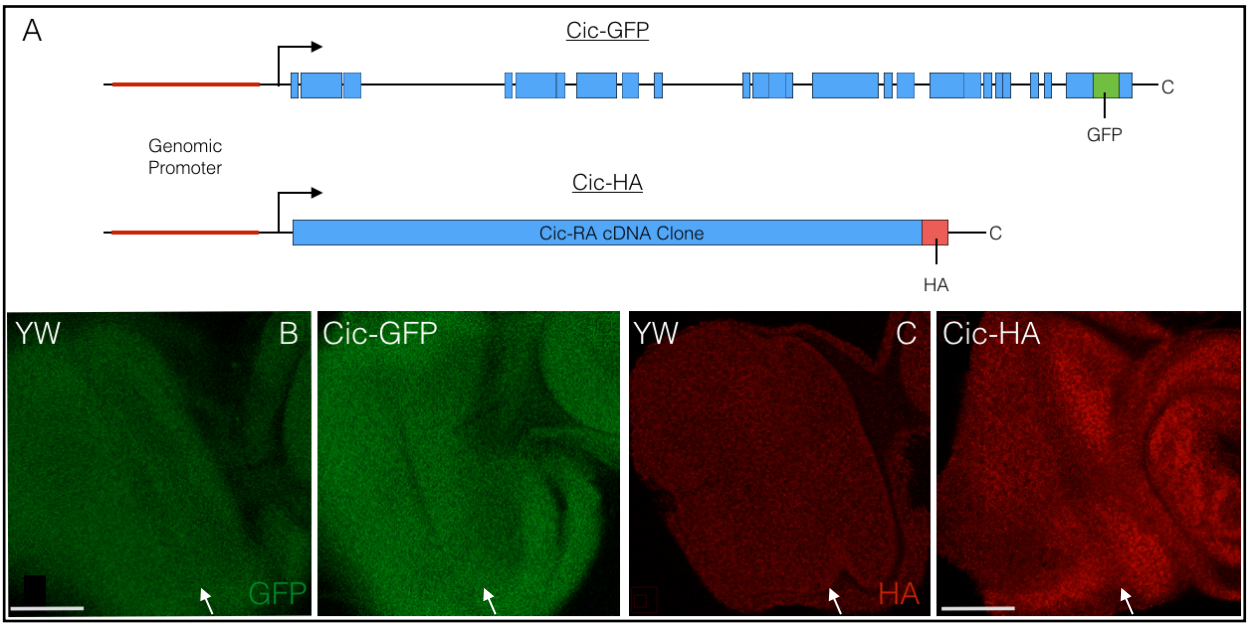

Fig. 2 Cic-HA recapitulates endogenous Cic expression pattern in eye imaginal discs

(**A**) Cic-GFP is a genomic construct whereas Cic-HA is a cDNA construct fused to a Cic genomic promoter (not to scale). Third instar eye imaginal discs expressing tagged Cic were immunostained with either (**B**) anti-GFP or (**C**) anti-HA to examine expression pattern. Cic-HA recapitulates expression pattern of endogenous Cic found anterior to the MF (marked in white). (All scale bars are 40μm).

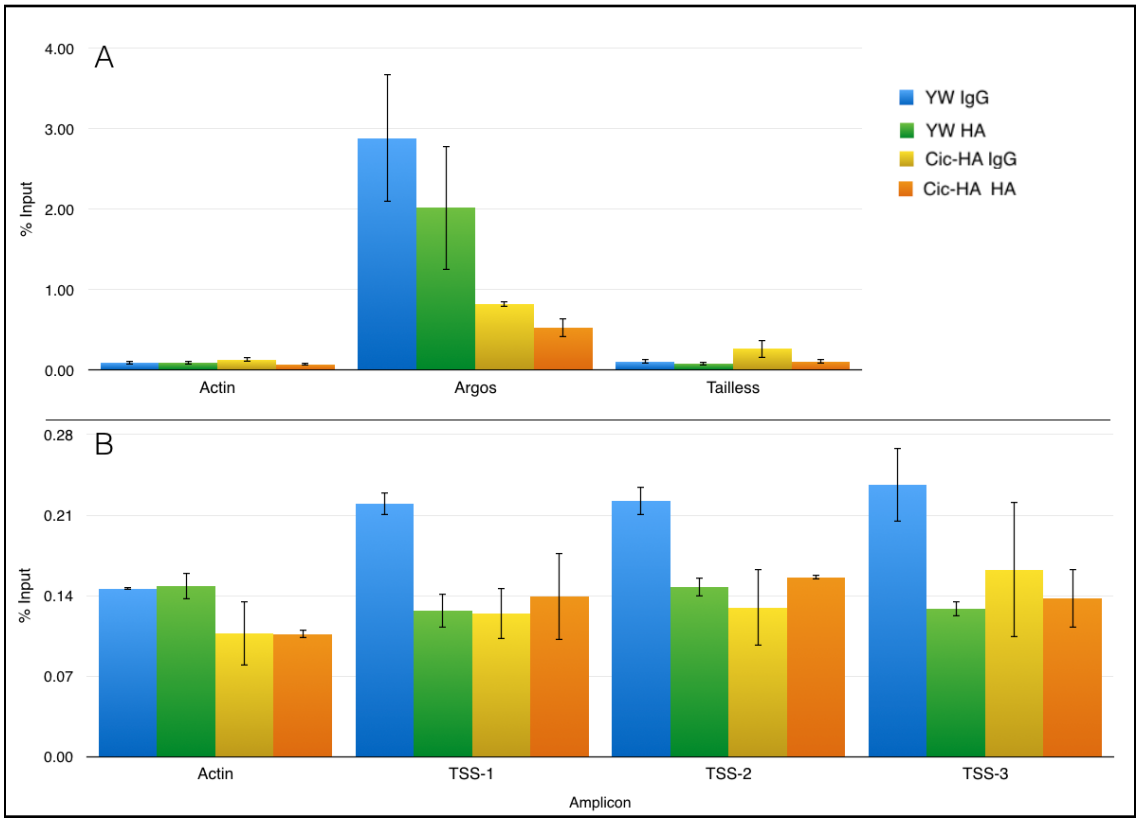

Fig. 3 Cic-HA cannot be detected on the promoters of previously reported Cic regulated genes.

Chromatin Immunoprecipitation (ChIP) using Cic-HA and YW third instar larvae was performed and subsequently analyzed by qPCR(mean ± SEM n=3 biological replicates). (**A**) Non-significant enrichment of Cic target loci was detected when Cic-HA was immunoprecipitated with anti-HA and compared to control samples. (**B**) Similarly, occupancy of Cic-HA at transcription start sites (TSS) of the candidate gene PRPS was not detected. All data analyzed by ANOVA.

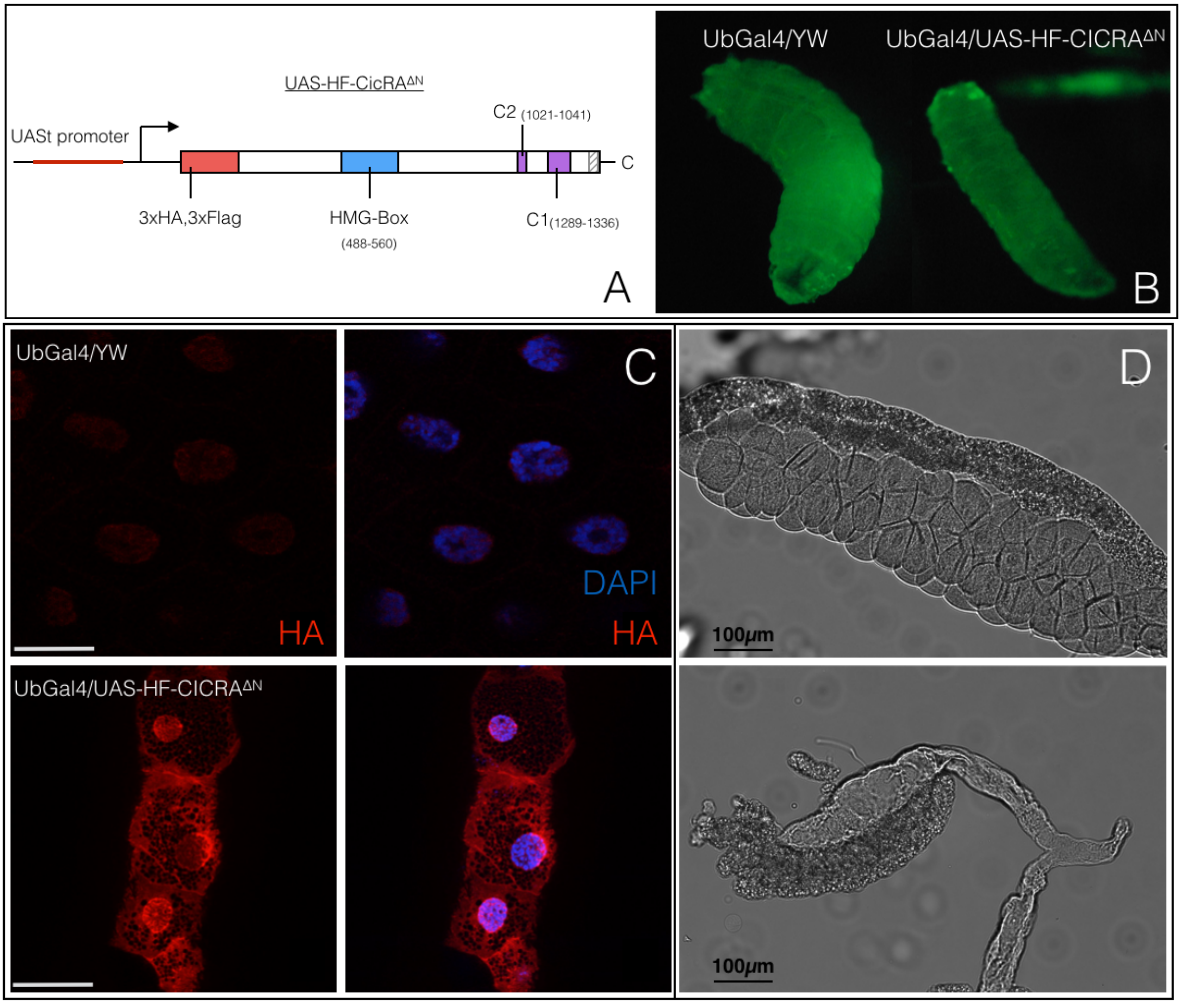

Fig. 4 Gal4 driven expression of HF-CICRA$^{\Delta N}$ reduces total third instar larval body size primarily in endoreduplicating tissue

(**A**) Construct with upstream activation sequences (UAS) for overexpression of HF-CICRA$^{\Delta N}$ containing both HA and Flag epitopes for the purposes of subsequent ChIP analyses is shown. All relevant domains including the HMG-box DNA-binding domain, C2 MAPK phosphorylation site and C1 domain responsible for repressor activity are intact. (**B**) Gal4 driven overexpression of HF-CICRA$^{\Delta N}$ resulted in reduced body size compared to the driver alone. (**C**)Wild-type and transgenic salivary glands were immunostained with anti-HA and DAPI. Note the nuclear localization of HF-CICRA$^{\Delta N}$ (Scale bars 40μm). (**D**) Whole mount salivary glands shown for size comparison (no size difference observed in imaginal eye discs).

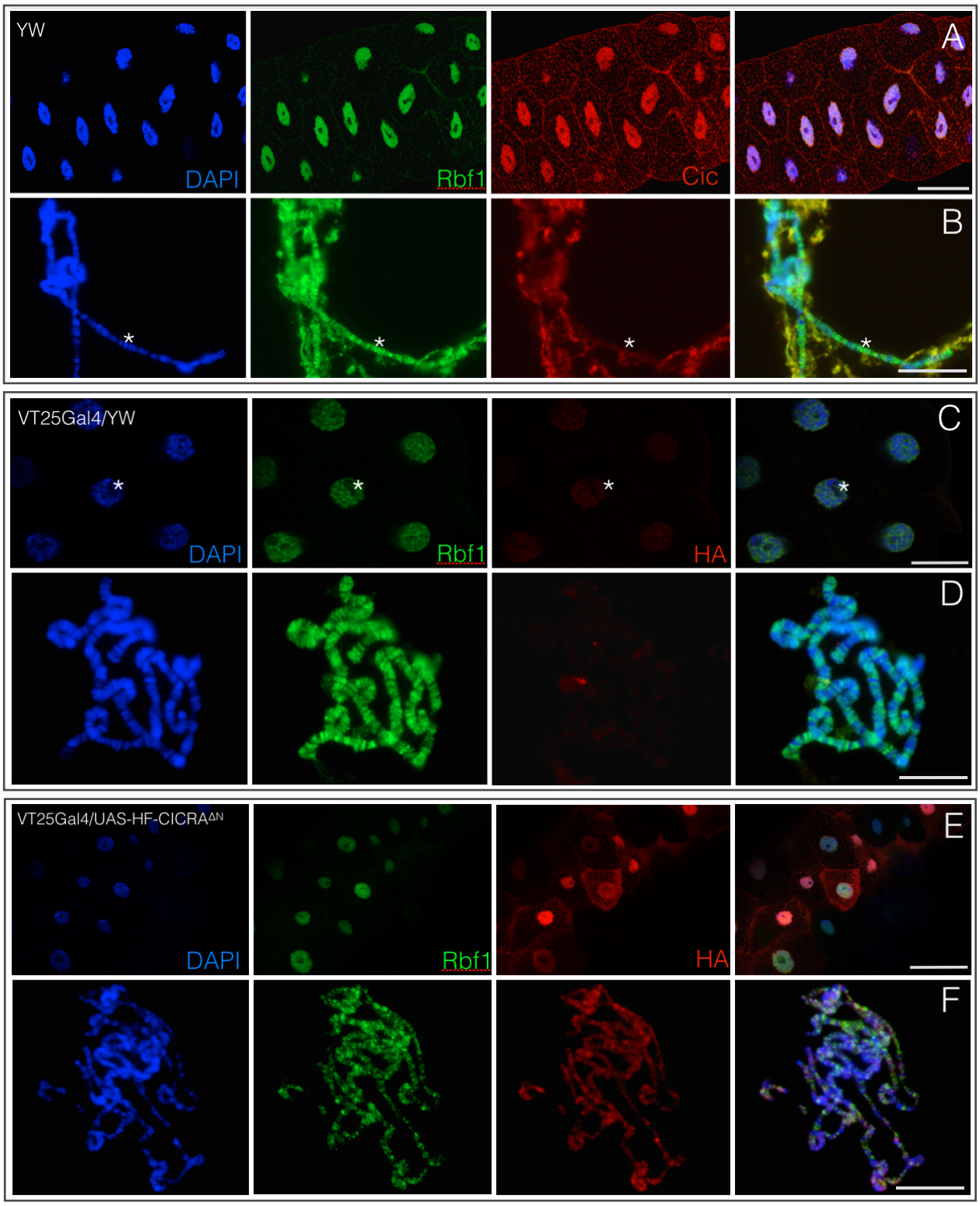

Fig. 5 DNA binding of HF-CICRA$^{\Delta N}$ can be observed when it is expressed by a salivary gland specific Gal4 driver

Wild-type third instar salivary glands (SG) were immunostained with anti-Rbf1(green), and anti-Cic (red). Note that Cic can be detected in the nucleus (**A**) but fails to be detected on polytene chromosomes (**B**). HF-CICRA$^{\Delta N}$ is expressed with a SG specific Gal4 driver VT25. Control (**C**) and HF-CICRA$^{\Delta N}$ (**E**) expressing SGs were immunostained with anti-Rbf1(green) and anti-HA(red). Note that HF-CICRA$^{\Delta N}$ binds polytene chromosomes (**F**) (all scale bars 40μm).

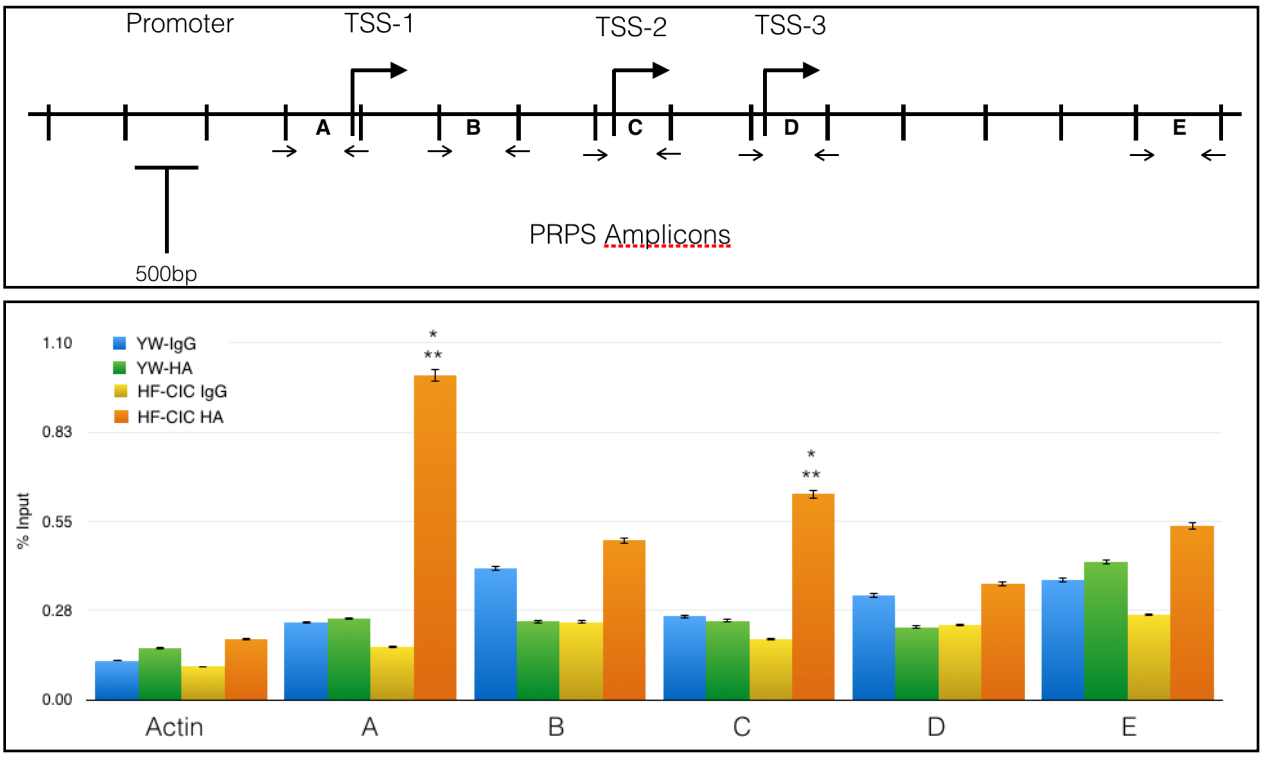

Fig. 6 HF-CICRA$^{\Delta N}$ binds transcription start loci in candidate gene PRPS

(**A**)PRPS contains 3 transcription start sites (TSS). TSS-1 contains an HMG-Box consensus sequence previously shown to be necessary for DNA-binding activity and transcriptional repression by Cic. qPCR primer pairs designed to assess enrichment of Cic using ChIP are shown as arrows. (**B**) ChIP assay against HA and mock (IgG) followed by qPCR are shown. A significant enrichment of amplicons A and C (mean ± SEM n=3 biological replicates) are shown. Analyzed by ANOVA and post-hoc Tukey HSD tests. (**P<.01 between HFCIC-HA and HFCIC mock(IgG). *P<.05 between HFCIC-HA and both YW-HA/YW-IgG.)

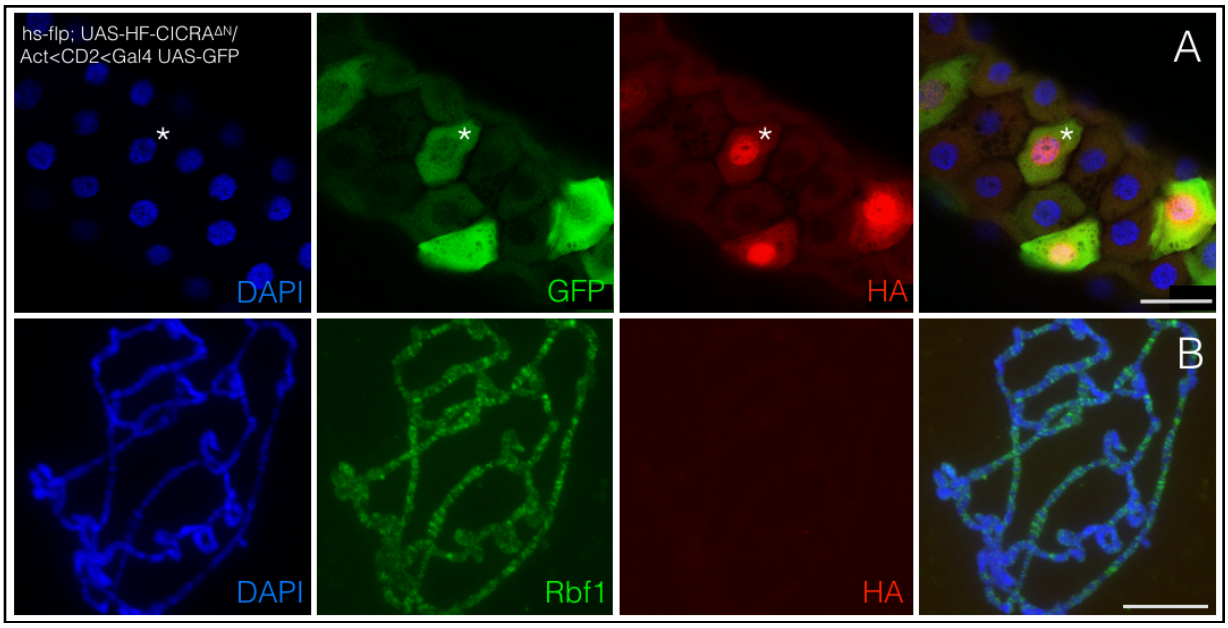

Fig. 7 DNA-binding activity of HF-CICRA$^{\Delta N}$ is likely dependant on temporal expression in addition to expression level

Flippase expression was induced by heat-shock 24hrs post hatching to create GFP^{+} marked clones expressing HF-CICRA$^{\Delta N}$ via actin Gal4. (**A**) Marked by (*) HF-CICRA$^{\Delta N}$ expression does not induce an observable reduction in nuclear or overall cell size and (**B**) staining does not colocalize with polytene chromosomes (All scale bars 40 µm).

DISCUSSION

Loss of Capicua function in an *Rbf1* mutant background has been previously reported to exacerbate the incidence of ectopic S-phase cells within the MF and promote survival of RBF1 deficient cells which normally undergo apoptosis as they begin to arrest and differentiate into photoreceptors [8]. The data presented here support Capicua's role in regulating S-phase entry by demonstrating direct binding to *PRPS*, a gene encoding a key enzyme in the de-novo nucleotide biosynthetic and salvage pathways [82,84,89]. Furthermore, a size reduction of polytene chromosomes and secretory cells observed when Capicua is overexpressed supports a role in cell cycle regulation, though likely does so by interrupting DNA-synthesis via *PRPS* repression rather than regulating Pre-RC formation or endocycle regulators such as APC/C. While the evidence in this thesis supports this notion, the degree of the proposed regulatory effects of Capicua must be viewed in the context of endogenous function; DNA binding activity is undetectable when Capicua is expressed at an endogenous level, consistent with a lack of literature reporting direct interaction with HMG-box consensus sequences in *Drosophila* until very recently [25].

Initial ChIP-qPCR investigations determined Cic-HA did not enrich for previously reported targets or HMG-Box consensus sites within *PRPS* despite recapitulating endogenous expression pattern (Figs. 2 and 3). Upon further investigation in salivary gland cells, endogenous Capicua was not recruited to polytene chromosomes (Fig. 5), supporting the lack of target gene enrichment in initial ChIP-qPCR assays. The strong presence of Capicua within the nucleus of secretory cells was unexpected however, as Capicua has been previously reported to exhibit cytoplasmic localization following MAPK-mediated phosphorylation [6,23,30,31], given the

presence of multiple RTK signalling pathways required for salivary gland development [57,87]. RTK signalling may also explain the lack of polytene chromosome recruitment by Capicua, though again some cytoplasmic localization should be observed if this is the case. It should be noted that nuclear export following MAPK-mediated phosphorylation at Capicua's C2 domain seems to be a conserved mechanism between both metazoans and mammals, though it has not been directly reported in this cell type. Alternately, the lack of observable cytoplasmic Capicua may be due to changes in primary antibody affinity. Phosphorylation of Capicua's C2 domain by MAPK may interrupt the C-terminal epitope targeted by the antibody used for this experiment, causing nuclear Capicua to be preferentially bound. There is some support of this notion in figure 7, where clones expressing N-terminally tagged HF-CICRA$^{\Delta N}$ exhibit some cytoplasmic localization in addition to predominant nuclear staining. Together this data suggests that there is a substantial concentration of Capicua not bound to chromatin within the nucleus of salivary gland cells.

Regulation of Capicua has been best characterized in response to EGFR signalling (reviewed in [24]). While it is advantageous to perform ChIP assays with endogenous protein to preserve expression pattern and activity particular to a developmental stage, preliminary efforts to increase Capicua's DNA-binding activity by interrupting EGFR activity via RNAi and expression of dominant negative EGFR protein had no detectable effect, possibly due to the wide range of other active RTKs during larval development [57,87]. Due to undetectable DNA-binding activity of endogenous Capicua as determined through ChIP-qPCR, overexpression was determined to be necessary for detection of target gene enrichment (Fig. 6) and recruitment to polytene chromosomes (Fig 5).

Consistent with direct binding and likely hyper-repression of *PRPS*, HF-CICRA$^{\Delta N}$ overexpression results in reduced polytene and salivary gland size, as the main determinant of cell volume in endocycling cells is total chromosomes copy number [62] mentioned earlier. The thin appearance of polytene chromosomes is indicative of reduced rounds of endoreduplication, as this phenotype has also been reported in larvae deficient for APC2 or Fzr required for degradation of mitotic cyclins by APC/C [61,73]. While the evidence shown here support a role for Capicua in regulating progression of S-phase by reducing expression of enzymes involved in nucleotide biosynthesis, recent direct Capicua binding evidence put forth by Jin et al., 2015 [25], reveal other factors that may contribute to the effect Capicua overexpression has on cell cycle progression. In addition to *PRPS* this study also reports that *String* as well as *CycE* are directly bound by Capicua. Hyper-repression of either of these targets may also contribute to the size reduction of endorepulicating cells and therefore total larval body size, but also confirms previous reports by our lab that loss of Capicua function may act to increase CycE protein level to overcome cell cycle arrest in RBF1 hypomorphic cells within the MF [8]. While repression of CycE, *String*, or *PRPS* alone would likely contribute to the observed effect on salivary gland growth when Capicua is ubiquitously expressed, *PRPS* expression is most likely to promote ectopic S-phase in RBF1 hypomorphic cells based on it's identification in RNA-sequencing data.

It is clear through investigation of clones overexpressing Capicua (Fig. 7) that the effects observed on salivary gland cell growth require constitutive expression throughout a population of cells. Furthermore, it is possible that factors required for cell growth are subject to down regulation, either directly or indirectly as a result of Capicua overexpression, and these changes can be compensated for by adjacent cells. Data in figure 5 also revealed that strong nuclear

localization of endogenous Capicua is not necessarily coupled to DNA-binding or chromosome colocalization, however there are no other known functions or domains that would facilitate any known protein-protein interaction at this developmental stage. Relief of Capicua repression by widespread RTKs, such as EGFR, has been shown to induce expression of *bantam* microRNA, which itself acts to inhibit Capicua expression directly [90]. It is possible that this mechanism of Capicua inhibition acts to keep protein levels below a certain threshold required for stable DNA-binding activity. Furthermore, the widespread and constitutive expression of Capicua by ubiquitin Gal4 may be required to overcome negative feedback by *bantam* microRNA or dosage dependant effects of RTK signalling during larval development. Altogether these data suggest stable DNA-binding activity of Capicua is relatively low within endocycling cells, which make up the majority of larval tissues (reviewed in [60,61]).

Together the data presented in this thesis support a model where loss of Capicua results in promotion of ectopic S-phase entry, and provides strong ChIP-qPCR evidence consistent with the hypothesis that Capicua directly represses the critical nucleotide biosynthetic gene *PRPS* through HMG-Box consensus sites. Based on this evidence, *PRPS* expression likely contributes to the proliferation of RBF1 deficient cells when Capicua function is compromised. Furthermore, I present evidence supporting Capicua's function as a regulator of cell cycle progression within endoreduplicating tissue, and consistent with the reported functions of the identified target *PRPS*. Finally, I describe a low incidence of chromosome colocalization with both endogenous and overexpressed Capicua, unless constitutively expressed in endoreduplicating cells.

Through the course of my investigation of Capicua regulated genes obtained from RNA sequencing, several highly upregulated genes displayed non-significant enrichment at promoters

and pseudo-Tor-RE sites. While this suggests specific targeting of *PRPS* by Capicua is, it is unlikely that *PRPS* is the sole derepressed factor that contributes to ectopic proliferation in *Rbf1* mutant cells. Furthermore, it is difficult to find evidence of a link between proposed *PRPS* activity and inhibition of cell death in *Rbf1* mutant cells. This effect would be more appropriately attributed to an increase in *bantam* microRNA activity when Capicua is lost, which has been reported to repress the pro-apoptotic gene *hid* [91]. Nevertheless, groundwork laid here detailing Capicua's specific enrichment of transcription start loci supporting transcriptional repression of *PRPS* warrants further investigation into its role in proliferation of *Rbf1* mutant cells. A logical approach to this end would be mutating *PRPS* in such a way to disrupt its functional domains or reduce its expression level via RNAi within *Rbf1* and *Capicua* double mutant eye discs. By quantifying or observing the incidence (or lack there of) of ectopic S-phase cells by BrdU incorporation within the posterior region of the MF, the effect size of *PRPS* in promoting proliferation of *Rbf1* mutant cells can be determined. This effect would be best observed using *PRPS* negative and positive clones, which our lab is currently working to investigate using the MARCM system.

During the course of my investigation a study was published characterizing a previously unidentified motif within Capicua-S dubbed the N2 domain specific to metazoans [37]. While the construct used throughout the course of my investigation did not contain this motif, it is dispensable for transcriptional repression during third instar larval development where the initial candidate gene identification took place. This suggests the lack of the N2 motif would not have impacted direct target gene identification through ChIP-qPCR, however I cannot rule out the possibility that the effects of HF-CICRA$^{\Delta N}$ overexpression affected embryonic development

leading to reduced salivary gland size. This is unlikely due to reports that decreased secretory cell and polytene chromosome size stems primarily from endocycle deregulation such as that associated with CycE and *Stg* fluctuations which have now been confirmed to be direct targets of Capicua [25]. Nevertheless, I prepared a full length CicRA overexpression construct which includes the N2 motif should direct targets or effects of Capicua expression on development be revisited in future studies by our lab.

Lastly, although the focus of our lab's research primarily hinges on better understanding factors that co-operate with *Rbf1* mutations and cell cycle control, the incidence of strong Capicua nuclear localization without DNA binding displayed in figures 5 and 7 suggests there may be previously uninvestigated functions of Capicua. Despite no known protein-protein interactions aside from AXTN1 and Gro during embryogenesis, data reported in Krivy et al. [8] suggests Capicua may regulate CycE post-transcriptionally. This finding however may represent an indirect consequence of Capicua gene regulation, as transcriptional regulation of CycE reported in Jin et al, 2015 was identified in ISCs which are distinct in both tissue type and developmental stage and may alter Capicua's target specificity. Regardless, it is possible that some effects of the loss of Capicua in an *Rbf1* mutant background may be attributable to functions other than transcriptional repression. Investigation to this end however would have to begin at a very basic point, perhaps through co-immunoprecipitation followed by tandem mass spectrometry and may present an avenue of research not directly relevant to better understanding characteristics of *Rbf1* mutant cells. Despite this notion, focus should be first placed on the role of *PRPS* in facilitating S-phase progression and possible effects on survival of *Rbf1* mutant cells.

MATERIALS AND METHODS

Fly Stocks

All fly stocks and crosses were raised at 25°C with the exception of those used to generate HF-CIC overexpressing clones described below (See Cloning of UAS-HF-CICRA$^{\Delta N}$, Generation of HF-CICRA$^{\Delta N}$ overexpression clones) which were raised at 18°C. The alleles Ubiquitin-Gal4 were obtained from Bloomington Drosophila Stock Centre while VT045325-Gal4 used for salivary gland-specific expression was obtained from the Vienna Tile Gal4 Library of the Vienna Drosophila Resource Centre. Capicua transgenes used were HF-CICRA cloned from LD17181 analogous to CIC-RA transcript was generated by TOPO® cloning methods (described below), Cic-HA from Prof. Laura Nilson (McGill University) and Cic-GFP obtained from the Bloomington Stock Centre. Stocks used for generation of HF-CICRA overexpression clones (described below) were hs (heatshock)-flipase;Sp/Sm6, a gift from Prof. Laura Nilson (McGill University) and Act<CD2<Gal4, UAS-GFP.

Generation of HF-CICRA$^{\Delta N}$ overexpression clones

To generate GFP^{+} marked HF-CICRA$^{\Delta N}$ overexpression clones hs-flipase; UAS-HF-CICRA$^{\Delta N}$ males were crossed to virgins homozygous for Act<CD2<Gal4, UAS-GFP at 18°C in bottles. Approximately 24 hours post hatching (at first instar stage) bottles were heat-shocked in a 37°C water bath for 10 minutes to induce flipase expression and removal of the CD2 cassette to enable *Actin*-driven Gal4 expression of GFP and HF-CICRA$^{\Delta N}$ and returned to 18°C. Bottles were subsequently heat shocked approximately 48 hours post hatching (second instar stage) in order to increase flipase expression and produce more HF-CICRA$^{\Delta N}$ overexpression clones.

Eye Disc and Salivary Gland Immunostaining and Microscopy

Antibodies used were Anti-HA (1/500, Roche), Anti-GFP (FITC) (1/200, Abcam), Anti-Capicua (1/500 Custom from Medimaps), Anti-RBF1(1/100 [63]). Third instar larval eye discs or salivary glands were dissected in 1xPBS and fixed with 4% paraformaldehyde for 20 minutes at room temperature (25°C). Following fixation samples were washed twice in 0.3% PBST (0.3% Triton X-100 in PBS) for 10 min. Samples were then incubated with primary antibody in 0.1% PBST (0.1% Triton X-100 in PBS) and 5% normal goat serum (NGS) overnight at 4°C. The following day discs or salivary glands were washed 5 times in 0.1% PBST for 10 minutes. Secondary antibody was then incubated in 0.1% PBST and 5% NGS for 2 hours at room temperature except in the case of Cic-GFP discs (Anti-GFP is conjugated to FITC). Samples were then washed washed 5 times in 0.1% PBST for 10 minutes and mounted for imaging by confocal microscopy (Leica SP8).

Polytene Chromosome Squash and Immunostaining

Antibodies used were Anti-HA (1/500, Roche), Anti-Capicua (1/500 Custom from Medimaps), Anti-RBF1(1/100 [63]). Salivary glands were dissected from third instar larvae in 0.1% PBST. Fat bodies were removed and undamaged glands were transferred to fixation buffer (1.5% paraformaldehyde in 45% acetic acid solution) for 10 minutes. Fixed salivary glands were then transferred and squashed between a lysine coated slide and coverslip. At this point the squash preparation was tapped in a spiral motion to break up cells and spread out polytene chromosomes. The squash preparation was then frozen in liquid nitrogen for 30 seconds at which point the coverslip was removed and the lysine coated slide was transferred to 1xPBS. Slides were washed once in 1xPBS for 5 minutes, once in 0.1%PBST for 10 minutes, and blocked for 30 minutes in PBS-TB (0.1% Triton X-100 and 1% bovine serum albumin (BSA) in PBS). After

removing excess PBS-TB primary antibodies in PBS-TB was added to the squash area and covered with a glass coverslip. Incubation took place overnight at 4°C in a humid chamber. The coverslip was then removed in 1xPBS and slides were then washed 3 times in 1xPBS for 5 minutes then blocked in PBS-TB 2 times for 15 minutes. As before excess PBS-TB was removed and secondary antibody in PBS-TB was added to the squash area and incubated under a coverslip for 1 hour at room temperature in a humid chamber. Coverslips were then removed in 1xPBS after incubation. Slides were washed 3 times in 1xPBS for 5 minutes. Excess PBS was removed and DAPI (1µg/ml) in PBS was added to the squash area and incubated for 5 minutes at room temperature. Slides were then washed 2 times in PBS for 5 minutes and mounted for immunofluorescence imaging (Zeiss Axio Imager.Z2.)

Chromatin Immunoprecipitation

Chromatin samples used in ChIP assays were extracted from staged third instar larvae from homozygous Cic-HA transgenic lines and yw or yw; Ubiquitin-Gal4/+;UAS-HF-CICRA$^{\Delta N}$/+, and yw; Ubiquitin-Gal4/+. 80-100 larvae were homogenized and fixed in 5ml buffer A1 (50mM KCl, 15mM NaCl, 4mM MgCl2, 15mM HEPES pH 7.6, 0.5% Triton X-100, 0.5mM DTT) and 1.8% formaldehyde using an electric homogenizer and glass Dounce and incubated for a total of 15 minutes at room temperature. Formaldehyde cross linking was quenched with 225mM glycine for 5 minutes on ice. Samples were then centrifuged for 5 minutes at 4000g at 4°C and the supernatant was discarded at which point pellets were then washed and resuspended with 3ml A1 buffer. This step was repeated 3 times. Pellets were then resuspended in 0.5ml modified lysis buffer (140mM NaCl, 15mM HEPES pH 7.6, 1mM EDTA, .5mM EGTA, 1% Triton X-100, .5mM DTT, 0.1% DOC, 0.1%SDS, 0.5% N-lauroylsarcosine) and incubated for 10 minutes at

4°C. Chromatin was then solubilized and sheared by sonication on ice to achieve and average size of 500bp and incubated at 4°C for an additional 10 minutes. Insoluble material was removed by centrifugation and the supernatant containing the sheared chromatin was eluted 3 times from the resulting pellet with modified lysis buffer. Protein A and protein G sepharose beads were mixed (50/50) and equilibrated with lysis buffer (140mM NaCl, 15mM HEPES pH 7.6, 1mM EDTA, .5mM EGTA, 1% Triton X-100, .5mM DTT, 0.1% DOC, 0.05%SDS) and used to preclear chromatin samples overnight at 4°C. Once beads were removed samples were then immunoprecipitated using 2.5 µg of anti-HA (monoclonal from mouse, Santa Cruz Biotechnology) or normal mouse IgG (Santa Cruz Biotechnology) per 250µl of sheared chromatin extract. Antibodies were incubated with chromatin samples overnight rocking at 4°C while the protein A and G sepharose bead mixture was blocked in lysis buffer with 1mg/ml sheared salmon sperm DNA and 1mg/ml BSA. Beads were then washed with lysis buffer to create a 50% slurry and added to chromatin samples (50µl slurry per 250µl chromatin extract) and incubated for 6 hours at 4°C. Supernatant was separated and discarded by centrifugation and sepharose beads were washed 4 times in 1ml lysis buffer for 10 minutes, once in LiCl wash buffer (10mM Tris, 250mM LiCl, 1mM EDTA, 0.5% DOC, 0.5% NP-40), and once in standard Tris-EDTA (TE) buffer. At this point the sepharose beads were incubated 2 times in 250µl of a 1%SDS and 0.1M NaHCO3 solution at 37°C for 30 minutes to elute protein-DNA complexes. Beads were again separated by centrifugation after each incubation and supernatant containing eluted protein-DNA complexes (eluates) were combined. Crosslinks were reversed overnight at 65°C and both eluates and pre cleared input samples set aside prior to immunoprecipitation were subject to RNAse A digestion for 30minutes at 37°C and proteinase K treatment for 2 hours at

42°C. DNA was then purified by chloroform:isoamyl alcohol/chloroform extraction and precipitated with 100% ethanol in the presence of 1µl glycogen (20 mg/ml) at -80°C overnight. DNA was subsequently resuspended in H_2O and prepped for enrichment of target loci by quantitive real-time PCR.

Quantitative Real-Time PCR

Results presented are averages of 3 independent chromatin immunoprecipitation experiments analyzed by triplicate qPCR reactions performed using the DyNAmo Flash SYBR Green qPCR Kit (Finnzymes) with a Bio-Rad CFX 96 Real-Time System and C1000 Thermal Cycler. Sample target loci enrichment is determined by percentage of input chromatin not subject to immunoprecipitation and is determined by cycle threshold (Ct) value determined by Bio-Rad CFX Manager software. Percentage input is determined according to the equation in appendix. All data was analyzed by ANOVA and post-hoc Tukey-HSD tests where applicable. Actin is used as a negative control in all cases, and all primers were designed using Primer3[92,93]. Primer pairs used for ChIP-qPCR analysis:

Argos 1.0F: GAATGAGCAGCCGGAAAGAG

Argos 1.0R: CGCACAATTGGCCACTTTTG

TaillessF: ATGGCAGCGCCATAAAATAC

TaillessR: TTTCCGCAGGATTCACTACC

PRPS-A(TSS-1)F: CCCATAAAATAACCAGATAACCAAA

PRPS-A(TSS-1)R: GAAACTCAACTTCCCTCGCTT

PRPS-BF: GACAGTCTTATGGGGCCTTG

PRPS-BR: TTTTCGCTCTGTGTGACG

PRPS-C(TSS-2)F: GAACAGCTGAGTGCGTCTTG

PRPS-C(TSS-2)R: TTCCTTTCATTTCTGTCCGCG

PRPS-D(TSS-3)F: CGCGCTAGAAGTTTGATCCG

PRPS-D(TSS-3)R: GATAGTCGATCGCCTGGGAA

PRPS-EF: GTGCTTGTCTTGTGAGGTCG

PRPS-ER: GCGATTTAGATTCCGCCCA

Actin-F: CCAACTCAAATCGCTTCGAG

Actin-R: CGCACTCACACACCTTTTAG

Cloning and Generation of UAS-HF-CICRA$^{\Delta N}$

The CIC-RA isoform was amplified from the LD17181 clone amplified using the Phusion® Hi-Fidelity PCR kit (Finnzymes) using primer pair (Fwd:CACCATGTCGCTGAGCAGCTCCC) (Rev: GTAATATTGAAAAACATCTGCCGC). As per protocol supplied by the manufacturer 10ng of template in a 6µl TOPO® Cloning (Invitrogen) reaction was incubated for 10 minutes at room temperature and then set on ice prior to bacterial transformation. 2µl of TOPO® Cloning reaction mixture was then transformed with 5-alpha Competent E. coli (NEB) cells. Once verified by restriction digest and Sanger sequencing performed by Génome Québec the LD17181 transcript was transferred from the pENTR-dTOPO® entry vector to the pTFHW gateway vector (Designed by T. Murphy, The Carnegie Institution of Washington, Baltimore, MD) by an LR reaction using LR Clonase^{TM} II enzyme mix kit (Invitrogen) as per manufacture's supplied protocol. The resulting destination vector containing UAS-HF-CICRA$^{\Delta N}$ was transformed in 5-alpha Competent E. coli (NEB) cells and subsequently extracted using the QIAGEN Plasmid

Maxi Kit (Qiagen) following supplied manufacture's protocols and injected into *yw* embryos. Upon re-sequencing of the construct after injection and selection for transgenic adult flies it was found that the C-terminus of UAS-HF-CICRA$^{\Delta N}$ contained an additional 3 amino acids (RDI) likely introduced from an imprecise LR reaction.

REFERENCES

1. Ross, J. F., Näär, A., Cam, H., Gregory, R., & Dynlacht, B. D. (2001). Active repression and E2F inhibition by pRB are biochemically distinguishable. *Genes and Development*, 15(4), 392–397. doi:10.1101/gad.858501

2. Stevaux, O., & Dyson, N. J. (2002). A revised picture of the E2F transcriptional network and RB function. *Current Opinion in Cell Biology*, 14(6), 684–691. doi:10.1016/S0955-0674(02)00388-5

3. van den Heuvel, S., & Dyson, N. J. (2008). Conserved functions of the pRB and E2F families. Nature Reviews. *Molecular Cell Biology*, 9(9), 713–24. doi:10.1038/nrm2469

4. Giacinti, C., & Giordano, A. (2006). RB and cell cycle progression. Oncogene, 25(38), 5220–5227. doi:10.1038/sj.onc.1209615

5. DeGregori, J., & Johnson, D. (2006). Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. *Current Molecular Medicin*e, 6(7), 739–748. doi:10.2174/1566524010606070739

6. Bettegowda, C., Agrawal, N., Jiao, Y., Sausen, M., Wood, L. D., Hruban, R. H., … Kinzler, K. W. (2011). Mutations in *CIC* and *FUBP1* Contribute to Human Oligodendroglioma. *Science (New York, N.Y.)*, *333*(6048), 1453–1455. http://doi.org/10.1126/science.1210557

7. Kawamura-Saito, M., Yamazaki, Y., Kaneko, K., Kawaguchi, N., Kanda, H., Mukai, H., … Nakamura, T. (2006). Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Human Molecular Genetics*, 15(13), 2125–2137. doi:10.1093/hmg/ddl136

8. Krivy, K., Bradley-Gill, M.-R., & Moon, N.-S. (2013). Capicua regulates proliferation and survival of RB-deficient cells in Drosophila. *Biology Open*, 2(2), 183–90. doi:10.1242/bio.20123277

9. Sherr, C. J., & McCormick, F. (2002). The RB and p53 pathways in cancer. *Cancer Cell*, 2(2), 103–112. doi:10.1016/S1535-6108(02)00102-2

10. Classon, M., and Harlow, E. (2002). The retinoblastoma tumour suppressor in development and cancer. *Nat. Rev. Cancer*, 2, 910–917.

11. Harbour, J.W., and Dean, D.C. (2000). Rb function in cell-cycle regula- tion and apoptosis. *Nat. Cell Biol*., 2, E65–E67.

12. Singh, S., Johnson, J., & Chellappan, S. (2010). Small molecule regulators of Rb-E2F pathway as modulators of transcription. *Biochimica et Biophysica Acta*, 1799(10-12), 788–94. doi:10.1016/j.bbagrm.2010.07.004

13. Sutcliffe, J. E., Korenjak, M., & Brehm, A. (2003). Tumour suppressors—a fly's perspective. *European Journal of Cancer*, 39(10), 1355–1362. doi:10.1016/S0959-8049(03)00263-6

14. Zalmas, L. P., Zhao, X., Graham, A. L., Fisher, R., Reilly, C., Coutts, A. S., & La Thangue, N. B. (2008). DNA-damage response control of E2F7 and E2F8. *EMBO Reports*, 9(3), 252–259. doi:10.1038/sj.embor.7401158

15. Zalmas, L. P., Coutts, A. S., Helleday, T., & Thangue, N. B. La. (2013). E2F-7 couples DNA damage-dependent transcription with the DNA repair process. *Cell Cycle*, 12(18), 3037–3051. doi:10.4161/cc.26078

16. Dick, F. a, & Rubin, S. M. (2013). Molecular mechanisms underlying RB protein function. Nature Reviews. *Molecular Cell Biology*, 14(5), 297–306. doi:10.1038/nrm3567

17. 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, 639–649

18. Nevins, J. R. (2001). The Rb/E2F pathway and cancer. *Human Molecular Genetics*, 10(7), 699–703. doi:10.1093/hmg/10.7.699

19. Chen, H.-Z., Tsai, S.-Y., & Leone, G. (2009). Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nature Reviews. Cancer*, *9*(11), 785–797. http://doi.org/10.1038/nrc2696

20. Cecchini, M. J., Thwaites, M. J., Talluri, S., MacDonald, J. I., Passos, D. T., Chong, J.-L., … Dick, F. a. (2014). A retinoblastoma allele that is mutated at its common E2F interaction site inhibits cell proliferation in gene-targeted mice. *Molecular and Cellular Biology*, 34(11), 2029–45. doi:10.1128/MCB.01589-13

21. Frolov, M. V., Huen, D. S., Stevaux, O., Dimova, D., Balczarek-Strang, K., Elsdon, M., & Dyson, N. J. (2001). Functional antagonism between E2F family members. *Genes & development*, *15*(16), 2146-2160.

22. Jiménez, G., Guichet, A., Ephrussi, A., & Casanova, J. (2000). Relief of gene repression by Torso RTK signaling: role of *capicua* in *Drosophila* terminal and dorsoventral patterning. *Genes & Development*, *14*(2), 224–231.

23. Astigarraga, S., Grossman, R., Díaz-Delfín, J., Caelles, C., Paroush, Z., & Jiménez, G. (2007). A MAPK docking site is critical for downregulation of Capicua by Torso and EGFR RTK signaling. *The EMBO Journa*l, 26(3), 668–77. doi:10.1038/sj.emboj.7601532

24. Jiménez, G., Shvartsman, S. Y., & Paroush, Z. (2012). The Capicua repressor--a general sensor of RTK signaling in development and disease. *Journal of Cell Science*, 125(Pt 6), 1383–91. doi:10.1242/jcs.092965

25. Jin, Y., Ha, N., Forés, M., Xiang, J., Gläßer, C., Maldera, J., … Edgar, B. A. (2015). EGFR/Ras Signaling Controls Drosophila Intestinal Stem Cell Proliferation via Capicua-Regulated Genes. *PLoS Genetics*, 11(12), 1–27. doi:10.1371/journal.pgen.1005634

26. Lam, Y. C., Bowman, A. B., Jafar-Nejad, P., Lim, J., Richman, R., Fryer, J. D., … Zoghbi, H. Y. (2006). ATAXIN-1 interacts with the repressor Capicua in its native complex to cause SCA1 neuropathology. *Cell*, 127(7), 1335–47. doi:10.1016/j.cell.2006.11.038

27. Fryer, J. D., Yu, P., Kang, H., Mandel-Brehm, C., Carter, a. N., Crespo-Barreto, J., … Zoghbi, H. Y. (2011). Exercise and Genetic Rescue of SCA1 via the Transcriptional Repressor Capicua. *Science*, 334(6056), 690–693. doi:10.1126/science.1212673

28. Tseng, A.-S. K., Tapon, N., Kanda, H., Cigizoglu, S., Edelmann, L., Pellock, B., … Hariharan, I. K. (2007). Capicua regulates cell proliferation downstream of the receptor tyrosine kinase/ras signaling pathway. Current Biology, 17(8), 728–33. doi:10.1016/j.cub.2007.03.023

29. Lemmon, M. A., & Schlessinger, J. (2010). Cell signaling by receptor-tyrosine kinases. *Cell*, *141*(7), 1117–1134. http://doi.org/10.1016/j.cell.2010.06.011

30. Grimm, O., Zini, V. S., Kim, Y., Casanova, J., Shvartsman, S. Y., & Wieschaus, E. (2012). Torso RTK controls Capicua degradation by changing its subcellular localization. *Development*, 139(21), 3962–3968. doi:10.1242/dev.084327

31. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., & Mann, M. (2006). Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell*, 127(3), 635–648. doi:10.1016/j.cell.2006.09.026

32. Goff, D. J., Nilson, L. a, & Morisato, D. (2001). Establishment of dorsal-ventral polarity of the Drosophila egg requires capicua action in ovarian follicle cells. *Development (Cambridge, England)*, 128(22), 4553–62.

33. Roch, F., Jiménez, G., & Casanova, J. (2002). EGFR signalling inhibits Capicua-dependent repression during specification of Drosophila wing veins. *Development, 129(4): 993–1002.*

34. Jiang, H., Grenley, M. O., Bravo, M. J., Blumhagen, R. Z., & Edgar, B. A. (2011). EGFR/Ras/MAPK signaling mediates adult midgut epithelial homeostasis and regeneration in drosophila. *Cell Stem Cell*, 8(1), 84–95. doi:10.1016/j.stem.2010.11.026

35. Malki, S., Boizet-Bonhoure, B., & Poulat, F. (2010). Shuttling of SOX proteins. International *Journal of Biochemistry and Cell Biology*, 42(3), 411–416. doi:10.1016/j.biocel.2009.09.020

36. Grosschedl, R., K. Giese, and J. Pagel. (1994). HMG domain proteins: Architectural elements in the assembly of nucleoprotein structures. *Trends Genet*. 10: 94–100.

37. Forés, M., Ajuria, L., Samper, N., Astigarraga, S., Nieva, C., Grossman, R., … Jiménez, G. (2015). Origins of Context-Dependent Gene Repression by Capicua. *PLoS Genetics*, 11(1), e1004902. doi:10.1371/journal.pgen.1004902

38. Ajuria, L., Nieva, C., Winkler, C., Kuo, D., Samper, N., Andreu, M. J., … Jiménez, G. (2011). Capicua DNA-binding sites are general response elements for RTK signaling in Drosophila. *Development (Cambridge, England)*, 138(5), 915–24. doi:10.1242/dev.057729

39. Chittaranjan, S., Chan, S., Yang, C., Yang, K. C., Chen, V., Moradian, A., … Marra, M. A. (2014). Mutations in CIC and IDH1 cooperatively regulate 2-hydroxyglutarate levels and cell clonogenicity. *Oncotarget*, *5*(17), 7960–7979.

40. Ready, D. F., Hanson, T. E., & Benzer, S. (1976). Development of the Drosophila retina, a neurocrystalline lattice. *Developmental Biology*, 53(2), 217–240. doi: 10.1016/0012-1606(76)90225-6

41. Cagan, R. L., & Ready, D. F. (1989). The emergence of order in the Drosophila pupal retina. *Developmental Biology*, 136(2), 346–362. doi:10.1016/0012-1606(89)90261-3

42. Baker, N. E., & Yu, S. Y. (2001). The EGF receptor defines domains of cell cycle progression and survival to regulate cell number in the developing Drosophila eye. *Cell*, 104(5), 699–708. doi:10.1016/S0092-8674(02)06076-2

43. Kumar, J. P. (2012). Building an Ommatidium One Cell at a Time. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, 241(1), 136–149. http://doi.org/10.1002/dvdy.23707

44. Baker, N. E. (2001). Cell proliferation, survival, and death in the Drosophila eye. Seminars in *Cell & Developmental Biology*, 12(6), 499–507. doi:10.1006/scdb.2001.0274

45. Greenwood, S., & Struhl, G. (1999). Progression of the morphogenetic furrow in the Drosophila eye: the roles of Hedgehog, Decapentaplegic and the Raf pathway. *Development (Cambridge, England)*, 126, 5795–5808.

46. Tomlinson, A. (1985). The cellular dynamics of pattern formation in the eye of Drosophila, J *Embryol Exp Morphol*., 331, 313–331.

47. Escudero, L. M., & Freeman, M. (2007). Mechanism of G1 arrest in the Drosophila eye imaginal disc. *BMC Developmental Biology*, 7, 13. doi:10.1186/1471-213X-7-13

48. Kumar, J. P. (2011). My What Big Eyes You Have: How the *Drosophila* Retina Grows. *Developmental Neurobiology*, *71*(12), 1133–1152. http://doi.org/10.1002/dneu.20921

49. Baker, N. E. (2007). Patterning signals and proliferation in Drosophila imaginal discs. *Current Opinion in Genetics and Development*, 17(4), 287–293. doi:10.1016/j.gde.2007.05.005

50. Heberlein, U., Wolff, T., & Rubin, G. M. (1993). The TGF homolog dpp and the segment polarity gene hedgehog are required for propagation of a morphogenetic wave in the Drosophila retina. *Cell*, 75(5), 913–926. doi:10.1016/0092-8674(93)90535-X

51. Ma, C., Zhou, Y., Beachy, P. A., & Moses, K. (1993). The segment polarity gene hedgehog is required for progression of the morphogenetic furrow in the developing Drosophila eye. *Cell*, 75(5), 927–938. doi:10.1016/0092-8674(93)90536-Y

52. Baker, N. E., Yu, S., & Han, D. (1996). Evolution of proneural atonal expression during distinct regulatory phases in the developing Drosophila eye. *Current Biology*, 6(10), 1290–1301. doi:10.1016/S0960-9822(02)70715-X

53. Bilak, A., & Su, T. T. (2009). Regulation of *Drosophila melanogaster* pro-apoptotic gene *hid*. *Apoptosis*, *14*(8), 943–949. http://doi.org/10.1007/s10495-009-0374-2

54. Moon, N.-S., Di Stefano, L., & Dyson, N. (2006). A gradient of epidermal growth factor receptor signaling determines the sensitivity of rbf1 mutant cells to E2F-dependent apoptosis. *Molecular and Cellular Biology*, 26(20), 7601–15. doi:10.1128/MCB.00836-06

55. Bergmann, A., Agapite, J., McCall, K., & Steller, H. (1998). The Drosophila gene hid is a direct molecular target of Ras-dependent survival signaling. *Cell*, *95*(3), 331-341.

56. Firth, L. C., & Baker, N. E. (2005). Extracellular Signals Responsible for Spatially Regulated Proliferation in the Differentiating Drosophila Eye. *Developmental Cell*, 8(4), 541–551. doi:10.1016/j.devcel.2005.01.017

57. Andrew, D. J., Henderson, K. D., & Seshaiah, P. (2000). Salivary gland development in Drosophila melanogaster. *Mechanisms of Development*, 92(1), 5–17. doi:10.1016/S0925-4773(99)00321-4

58. Smith, A. V, & Orr-Weaver, T. L. (1991). The regulation of the cell cycle during Drosophila embryogenesis: the transition to polyteny. *Development (Cambridge, England)*, 112(4), 997–1008.

59. Calvi, B. R. (2013). Making big cells: One size does not fit all. *Proceedings of the National Academy of Sciences* , 110 (24), 9621–9622. doi:10.1073/pnas.1306908110

60. Zielke, N., Edgar, B. A., & DePamphilis, M. L. (2013). Endoreplication. *Cold Spring Harbor Perspectives in Biology*, *5*(1), a012948. http://doi.org/10.1101/cshperspect.a012948

61. Zielke, N., Querings, S., Rottig, C., Lehner, C., & Sprenger, F. (2008). The anaphase-promoting complex/cyclosome (APC/C) is required for rereplication control in endoreplication cycles. *Genes & Development*, *22*(12), 1690–1703. http://doi.org/10.1101/gad.469108

62. Lee, H. O., Davidson, J. M., & Duronio, R. J. (2009). Endoreplication: polyploidy with purpose. *Genes & Development*, *23*(21), 2461–2477. http://doi.org/10.1101/gad.1829209

63. Du, W. E. I., Vidal, M., Xie, J.-E., & Dyson, N. (1996). RBF, a novel RB-related gene that regulates E2F activvity and interacts with cyclin E in Drosophila. *Genes and Development*, 10, 1206–1218. doi:10.1101/gad.10.10.1206

64. Havens CG, Walter JC. 2009. Docking of a specialized PIP Box onto chromatin-bound PCNA creates a degron for the ubiquitin ligase CRL4Cdt2. *Mol Cell*., 35: 93–104.

65. Zielke, N., Kim, K. J., Tran, V., Shibutani, S. T., Bravo, M.-J., Nagarajan, S., … Edgar, B. a. (2011). Control of Drosophila endocycles by E2F and CRL4CDT2. *Nature*, 480(7375), 123–127. doi:10.1038/nature10579

66. Shcherbata, H. R., Althauser, C., Findley, S. D., & Ruohola-Baker, H. (2004). The mitotic-to-endocycle switch in Drosophila follicle cells is executed by Notch-dependent regulation of G1/S, G2/M and M/G1 cell-cycle transitions. *Development (Cambridge, England)*, 131(13), 3169–3181. doi:10.1242/dev.01172

67. Deng WM, Althauser C, Ruohola-Baker H. 2001. Notch– Delta signaling induces a transition from mitotic cell cycle to endocycle in Drosophila follicle cells. *Development*, 128: 4737–4746

68. Jordan, K. C., Schaeffer, V., Fischer, K. a, Gray, E. E., & Ruohola-Baker, H. (2006). Notch signaling through tramtrack bypasses the mitosis promoting activity of the JNK pathway in the mitotic-to-endocycle transition of Drosophila follicle cells. *BMC Developmental Biology*, 6, 16. doi:10.1186/1471-213X-6-16

69. Sun, J. and Deng, W. M. (2005). Notch-dependent downregulation of the homeodomain gene cut is required for the mitotic cycle/endocycle switch and cell differentiation in Drosophila follicle cells. *Development*, 132, 4299-4308

70. Sun, J. and Deng, W. M. (2007). Hindsight mediates the role of notch in suppressing hedgehog signaling and cell proliferation. Dev. Cell., 12, 431-442.

71. Sigrist SJ, Lehner CF. 1997. Drosophila fizzy-related down-regulates mitotic cyclins and is required for cell proliferation arrest and entry into endocycles. *Cell*, 90: 671–681.

72. Sørensen, C. S., Lukas, C., Kramer, E. R., Peters, J. M., Bartek, J., & Lukas, J. (2001). A conserved cyclin-binding domain determines functional interplay between anaphase-promoting complex-Cdh1 and cyclin A-Cdk2 during cell cycle progression. *Molecular and Cellular Biology*, 21(11), 3692–703. doi:10.1128/MCB.21.11.3692-3703.2001

73. Kashevsky, H., Wallace, J. a, Reed, B. H., Lai, C., Hayashi-Hagihara, A., & Orr-Weaver, T. L. (2002). The anaphase promoting complex/cyclosome is required during development for modified cell cycles. *Proceedings of the National Academy of Sciences of the United States of America*, 99(17), 11217–11222. doi:10.1073/pnas.172391099

74. Narbonne-Reveau, K., Senger, S., Pal, M., Herr, A., Richardson, H. E., Asano, M., … Lilly, M. a. (2008). APC/CFzr/Cdh1 promotes cell cycle progression during the Drosophila endocycle. *Development (Cambridge, England)*, 135(8), 1451–1461. doi:10.1242/dev.016295

75. Mortin, L. I., & Sedatf, J. W. (1982). Structure of Drosophila Polytene Chromosomes, Evidence for a toroidal organization of the bands. *J Cell Sci*., 113, 73–113.

76. Johansen, K. M., Cai, W., Deng, H., Bao, X., Zhang, W., Girton, J., & Johansen, J. (2009). Polytene chromosome squash methods for studying transcription and epigenetic chromatin modification in Drosophila using antibodies. *Methods*, 48(4), 387–397. doi:10.1016/j.ymeth.2009.02.019

77. Collas, P. (2010). The current state of chromatin immunoprecipitation. *Molecular biotechnology*, *45*(1), 87-100.

78. Dey, B., Thukral, S., Krishnan, S., Chakrobarty, M., Gupta, S., Manghani, C., & Rani, V. (2012). DNA–protein interactions: methods for detection and analysis. *Molecular and cellular biochemistry*, *365*(1-2), 279-299.

79. Porro, A., & Perini, G. (2007). In vivo dual cross-linking chromatin immunoprecipitation: detecting chromatin proteins not directly bound to DNA (PROT29). *The Epigenome-Network of Excellence: Epigenetics protocols database*.

80. Jusiak B, Karandikar UC, Kwak SJ, Wang F, Wang H, et al. (2014) Regulation of Drosophila Eye Development by the Transcription Factor Sine oculis. PLoS ONE 9(2): e89695. doi: 10.1371/journal.pone.0089695

81. Greil F, Moorman C, Steensel BV (2006) DamID: mapping of in vivo protein–genome interactions using tethered DNA adenine methyltransferase. Method Enzymol 410:342–359

82. Kornberg, A., Lieberman, I., and Simms, E.S. (1955). Enzymatic synthesis and properties of 5-phosphoribosylpyro-phosphate. *J. Biol. Chem.* 215, 389–402.

83. Yen, R.C., Raivio, K.O., and Becker, M.A. (1981). Inhibition of phosphoribosylpyrophosphate synthesis in human fibroblasts by 6-methylthioinosinate. *J. Biol. Chem.* 256, 1839–1845

84. Hartman, S.C., and Buchanan, J.M. (1958). Biosynthesis of the purines. XXI. 5-Phosphoribosylpyrophosphate amido- transferase. *J. Biol. Chem.* 233, 451–455

85. Kim, H.-J., Sohn, K.-M., Shy, M. E., Krajewski, K. M., Hwang, M., Park, J.-H., … Kim, J.-W. (2007). Mutations in PRPS1, which encodes the phosphoribosyl pyrophosphate synthetase enzyme critical for nucleotide biosynthesis, cause hereditary peripheral neuropathy with hearing loss and optic neuropathy (cmtx5). *American Journal of Human Genetics*, 81(3), 552–558. doi:10.1086/519529

86. Lee KB, Thomas JO. The effect of the acidic tail on the DNA-binding properties of the HMG1,2 class of proteins: insights from tail switching and tail removal. *J Mol Biol.*, 304:135–149. doi: 10.1006/jmbi.2000.4206

87. Sopko, R., & Perrimon, N. (2013). Receptor Tyrosine Kinases in *Drosophila* Development. *Cold Spring Harbor Perspectives in Biology*, *5*(6), 10.1101/cshperspect.a009050 a009050. http://doi.org/10.1101/cshperspect.a009050

88. Carey, M. F., Peterson, C. L., & Smale, S. T. (2009). Chromatin immunoprecipitation (ChIP). *Cold Spring Harbor Protocols*, *2009*(9), pdb-prot5279.

89. de Brouwer, A. P. M., van Bokhoven, H., Nabuurs, S. B., Arts, W. F., Christodoulou, J., & Duley, J. (2010). PRPS1 Mutations: Four Distinct Syndromes and Potential Treatment. *American Journal of Human Genetics*, 86(4), 506–518. doi:10.1016/j.ajhg.2010.02.024

90. Herranz, H., Hong, X., & Cohen, S. M. (2012). Mutual repression by bantam miRNA and capicua links the EGFR/MAPK and hippo pathways in growth control. Current Biology, 22, 651–657. doi:10.1016/j.cub.2012.02.050

91. Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., & Cohen, S. M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell*, *113*(1), 25-36.

92. Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3 *Bioinformatics* 23(10):1289-91

93. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3 - new capabilities and interfaces. *Nucleic Acids Research* 40(15):e115

94. Attrill H, Falls K, Goodman JL, Millburn GH, Antonazzo G, Rey AJ, Marygold SJ; the FlyBase Consortium. (2016) FlyBase: establishing a Gene Group resource for Drosophila melanogaster. *Nucleic Acids Res*. 44(D1):D786-D792

APPENDIX A

PRPS Gene Region Investigated

FBgn0036030 (3L:9681579) Release: FB2016_04 [94]

HMG-Box Consensus Sequence **Pseudo-HMG-Box Consensus Sequence** **Transcription Start Site**

```
0bp - TTCGCCATGATGTTTCCTCCGTTTCTCGGACTCTCCTTTCTGCTCTCTCCCTGTTCCCTGTTC
      ACCGTTCTCCTTTCTCCGCCGCTCCTCTTCACCACGCTCTTCTTGCTGTGTTCTGTTCTGCC
      TCTGTTTTTATTGTTGCCCCCCCGTGCAATTTTGGATTTTGTGAGTGTTGTGCGTTGCGTAAA
      TTACGCCTGGCTGTTGGCCTACAAGGTTGCTCCTTTCGATCGCTTTTCGAATTGTCGAAGCA
251 - ACGAGCTCGACCAGCGCACGGAAAATAAAAAACTTATTATTCCAATTTCAACACGAATCTG
      ACTTTTTACAGCGCTCAGCACGTCTATTTTGCCAGCGCTATTTCTTTTTTGTTTTTCCGTAGT
      GTGACCGCCGCGGTGCATTCGGATAATTCTGAAAGCTCGCCTATCGATTGACGTATCGATAG
      GGCTTCGGGGAAAGAGCGCGAAAATTGCTAACGAAGTTCGAATGTTCAAATGTTTAGCATA
      TGAATAAATAACTATTACATAGGGACGTAATGAATTGGTGCTACTTGCTGCGAGCCATATATT
561 - GCCTAGATACGCCCTTAATACTACCCAGGTACACCCACTTCACTCAACTCATAACTATTTAAT
      CATATTAAAACCTATTAAACATTTTAAGGAGGAGTCTTCTTCTTAACATTGGTCCTCCAATTA
      TGTCCACTGGATGACATGATGACTTTGTTTCCCTGAAGGCGACATATCAGAATCTGGGAATC
      TTAACCGACACATTTTCCTAACTCGTTTCGATAATACAAATGAAAAGAAACAAAAGCAAAA
      TCTACATAAAAATAACCATATTTACAAACTACTTTCCTAATAAATTCAGCTAAGCCCCATACA
873 - CGAAGTACTCATTCTTACCTCTTTGTTGATAAAATCTATTTACGTTACAATGATTGGCTGAAT
      CAATTCCAAAGTTGGTTTTTGTACCATTGGATAAAAAAGCCGCAGCTTTTCCTTGGCACAC
      GTCAAAGGTTATTTTCCATTGCGTCTCTGATCTTGAATACTGAATACATGGCCTTATTATATAA
      TCGAAGTTCTTTTGACCTCAGTAATATGGTTCCATTACGCTCGCCCGCCTTGTCTGCGGTCG
      GAAATTTGAAACTTGTTTTGCGCCCCATAAACAAAAACAAGAAGTTTAATCTTCTAAGGCG
1184 - TAGGTCTCTCTTGGCAACAGTCCAAAAATAATAGACAGCCATTAATTATGCGCCATACGATC
      CTGTGTTTATAATGGTTATACTTGGAAAAGGGGTACACGAAGTGTTACTTTATTCACATTGC
      AATTTTAAATCAATATGAGAGAACTTTACCTAAAAAATTTTTAAGTGATAAACAATTAAGTG
      GATTACATTTTCAGCAAACTCTTATTAATAAGTTGAATGTTAACAAAATAAAAATATGATCAA
      AAAAATTAGAGGGCATATTTCATACCTGAAAAATCATTAAAGTCGACATTCTCTTTTACTGC
1495 - TCAAATTCACGTTTCGCACAGCTCAATGGCGATCGAGAAAAAAACGTGTCAGCAAAAGCG
      TGTGGGGAAGTTTTGAGAGCCAACGCAACAACAGTTTTCGGATAAGAAATATACCCAATAG
      ACATTCGGAGACTAATTAAAACGGAAGAGCTTCTTGCTTCTCCAACTCTTTGGACACACGT
      GTTTGTGCTCCCCTTTTGTGGCGCTCTTTCGCCACAAAACAAAACGAGTGACCGACACAA
      ACAAATATGCCGGTTCTCTTTTACTCTCCCCGATATTCCCTTTCGAAGCATATGTATTATGACT
1801 - GTGTGAAGAGGGATGCGTTTCCTACCCCATAAAATAACCAGATAACCAAAAGATACATTTA
      TTGTCTTGTAACGTAAGGGTATCCATGCTTCAGCTTGATTAAATCAACTCTCCTCTGCGTTAT
      CTCTTGCTCTCTCGCACCAAACAACTCCATAAACAGAAACTCAACTTCCCTCGCTTAAATT
      CAATTCAAAAGCGTTCTGATTCATTCAGCTTTTGCATTTTCTGAGCAGCGGCTGCATCCAA
      GTGAAATCACAAAGTACAAGTTTTTTTCGACCAACTTGTACAAACCGCATCGTTCAGTGAT
2108 - TAATAGTTTGTAGTTTAAAAGTAATTGCGACACTTTGCTACCTCACATGTTTCAAGTAAGTC
      CAGTATGGTGTACTACTCGATTTTGAAGTGTGCGTGCCTTTGTTTTCGCCTCGACTCAAAAC
      ATGCGTCGCATTCGGCACGTGGTTTTTGCCCCACCAACCCGCTATCCACGGCTCCATTATTC
      CTGTACTTGCAATGCCATTTGCGTCAGTTTTTAGGGATAAAACACGGAGAAGAGTACGTCA
      GCATTTGGCATTAGTTCATTTATATTTAGCCATGGGCACGAAAATAACAATTTAACAAATCG
2417 - CAGTGTTTTTCTGCCGCGTTCGTTTGTTCAGCATCCCAGGTACAGTTGCACAAACAACAAG
      AAAACATCTCGGGGGCAACAAAAGCAAAGACGAGCCCCCGTGAAATGAAAAAGCAAGAA
      CTGTACCTCGGATTTGAGATCTAGAAAACAAAACATTTCTGCTGGTATGTACAAAGTACATA
      GCAGAGATATTTGCAGAGCAACAATTATCCAGAACCAATTGAAAGCGGCATACAGCACAGA
      TACACAGACAAAAAATAAATTTCTAAGTATGTGCACAAAAAGAGCCGCATGCATATAGGTT
2721 - TTAAAAGTGGGAAACGCAGCATTAAATGTTTATTATTGTCTCAGATATTTAGGGTTTTTTTAT
      GTTATTATGAATAGGAAAATATATTGACAGTCTTATGGGGCCTTGGAAATTTATTACCCATTG
      AAATATATATGAAACAATTTTATTGCATATCTGTGTGTTATTAAGAAATTACTGAGTTGTGCA
      GATGTAATGAAACTAAGCCGCAATTTTGTCTTACTTTAGAAATGAGGTCTGTGATTATTTTAT
      AAAAAAAATTAATACTTATTATTGAAATTTTCGCTCTGTGTGACGTTTTACATATCTAAAATT
3036 - GCAACATGACGTATTAAAAATTCTATGGAAATTACCTATCGAAAACAGCTGATTCCGAAAA
      AGAGAGGAAAAGTGCGTCCCGAAAGCTTTTCCTTGTGGCATCAACTTAACTGCGCATCATC
      CCATCCTTGGCGCTCTCTTGCGCTCTCTCGTAAAATATGGGTGTGTGTGTGTGTGGGCAGCC
```

```
      GGTGCGTATGTGTAATAGAAGCATACGCCATGTTGTACGGCGCTCAGTTCCATTCGCTCGCT
      CAATTTCGCCAGATGCGGCGCTCTGCGGCTGCTAATCCCAATCTCCCATTGTCGCTGAAAC
3343 - AGTGCACTTTTCACCTTACCTGGCGCAGAACAGCTGAGTGCGTCTTGTTATTAAACGGGTT
      GTTCAATCAAATAGTCGGCGCTTACCTGTCGCAGAACCTTCATCTCGGGATTTGTTTTGTGA
      ATAATGCCCGATAAGAAATGCACGGAGAACGAGCAGAATGGTGCGTAGATTATATCCTAGA
      ATTCACATTCCTTTCATTTCTGTCCGCGTGTGTGTGTGTGGTGCTAGGTTTTCATGGCCAAC
      TGAGGGACATCCGTGGAACAGCTGTTTACCTCAGCCTTAAAAGCTTTTTTTTTTGGCGCGCG
3650 - CCTGTTGGCATGCCAGCATATTTTATATGTTAAAAGAAAACTCATAGATTTTACGGGTGTGA
      CAACAAGCTGTTTATCTATCGAATATAATCAGTGTCAATTTGCAATTTATATTTATTCTTTTTC
      ATGTTCATATGTAAATTTAATAGCAAACCGAGCCAAAATATTATAAATATAAAATTATTTGCA
      ATGAAATAGGTAAATCCTTGCTTTTTGAAAGGCTAATATTGTAATAGTGCAATCCCATTAAC
      TTGTTTAGATTTATACTGAGAAACACACTAGGAAATTCGAATTGATTAATGAAAATATTTAA
3963 - AATTATAAAAACCTATCGCACCGCTTTAGTAGTCCGGAATCGAAATCGATAGCTTTCCAGTT
      CCCTGGAGAACGCGTAGAAAAATCACATGGTCTGGATGGGAATGGGGGATAATTGGGGTG
      GCTTTACAGAAAGAGCGAGCGAAAGAGAGCGCAGTGCGCAGCAAAAAGCTTCGTTGCTC
      TCTTTTACGCTCTTGCTTTCTCACCGATAAGGCGACGACGTCGACGTCGCTGCCCGCGTTTT
      GGCGAAAAAAATTTATAATAAAAACTTCTACTCATAGATTGTGTGATGCCAGTTCGCTCTG
4267 - CTAATCCGATACGCGCTAGAAGTTTGATCCGAGACAATTTGGAGAAACAGGCTGGTTGTCT
      AAATTTAATCCACTCCAGAATGCCGAACATCAAAGTGTTTTCGGGCACCTCGCATCCGGAT
      TTGGCCCAGCGGATAGTCGATCGCCTGGGAATCGATCTCGGCAAGGTGGTTACGAAGAAG
      TTCAGCAACTTGGAGACCTGGTGAGTAGTCCTGCAGGAGAATTTCCTTCGCAAATTGCGTG
      TGTGACCACCAACAAGTGTTTAACGACAACGGACAAATGTGTTTTTAAAGTGCGCTAATTG
4571 - AGTTTGACGAAAAAGTGTTTGGTCTGCAAACACATGTGCTCAAAAGGTCCTTGACATTCG
      CTGAAGGTTGCTGGCGTTTCGCTGCCACAATCAACAATGGGAATCTATGGGAATTTGTGATT
      ACCTAAAACTGAATGATTTTTCCTGAATTTTTGTGGCCTGCCTTTGGCTTGATTTCTTTTTAT
      TTTTATTGATTTAACAAGTGGACTGTACACCGTCTTTGTCGCTATAGCCGTGCGTCTGTGTG
      TTTGTGTGTGTTTAGTGGCGTTGTGCTGGTATGAGTGAGAGTAACACACACACCAACATAC
4879 - ACGTGAAGCTTGGCCAATGTCGCGTGCAATTTTCACAAAATTTAAATTAAATTTTTTTAAAT
      CTTCCTAGACAAGAAATGCAAAGAAACCGAAAATGATTTTTATTATATATTTTTAATTAAATT
      AGCAGCTGTGTGTGCGTGTGTGTGTATGTCTAAACTTGCGGCTTGCACGCTTTTTGCTTCCC
      TCTCTCTCGCACGCACTCACACAGGCGCAAAACCAGCCAGTCAGCTGCTGATGCAACAAC
      AACAATAAATAACAAATAAAGGCGCCCATAGTTTTTACCCCACCCACTAGCAAGCACCCTT
5187 - TTCAGTTTATTTGGTGTCAACTTTCCATCGGTTTTTTCCAATTTTAATAAAATTAAACAAAAG
      CCCCATAAGCAAACAATTAGAGGAGGTGAAAAGGTGAAAGAAACAAACGAGGCAAACGG
      AAGAAAACCGAAACGAAAATAGAGGTTATAATACGCGCTAAGCTTTCTGACATGATACGCG
      CTCCGTCGGTGTGCGTGTGTTTTTGTTATGCTCGCACTCACTCGCACGCACACCAAGGTAC
      GCGGCAGCGGAAAGAACTAAAGGCGGAGCAGAGTCGAGGCTAGGATGCCCTGTTCCAGG
5490 - TGTTATGCTGTACAAGGTAATAGGAGTCGTGTCAACTTACAATTGTTTACAGTCTGATTTCTT
      ATAGCTTGATATTTTATGATCCTAAGAACCAATTGAAGAAAGTACGACGTTCGACGAATAAA
      TCGTAATAAATATTTTGTGAAAAAAGTTTCTTGGTTATAGAGTTAGACATAGTCTTATCTCTA
      AAAATGCATTATTTTCCAGGTTCCAGTTTTTAACTTTCTAATAATTCTTTACCATTACCGAAA
      GTTGAGACCCAATTTGGCTTACTCGCTTTTATAGTCGACATACCCGACTAAAGGAGGGTACC
5803 - AAATGCCTGCGAGTAGAGAACCACAACATGACTCTGTCACGTTTTTCATTTGCTGACTGAA
      CGGAGAGTAAGAGAACGCTCTCTTAGGGTGAGAGCGAAAGGGGAGCTTGAGCGCCACAC
      AAGAGTCGGCATACGTTGCAACGTCGACTGCGGCAGCGACGCCCCCGCATGGTGGGCATT
      TGGACTCTCTTTGGAGCCCATTCTCCCGCTCTCTCTCTCTTGCTCTCAGACTAAGTGTCAGA
      CCCCCACGTGTGCATAGTGGTTGTTGCTCCACTGTTGTTGTTGCTACAATATTTCTGTTTGC
6107 - ACTTGTCGGCTTTGTTTTTGTTGTTTATCCCATTCTATCACTCTCTGGCATTCTCTAGAAGCA
      GCGGTCAGAAAATCATGTTGACCGTTAAATGCGCTCTTTTTGCTCCAATTTACACTAGGTGT
      TGTTGTTATTGTTGCGGCTGGTGTAACGGTTGAACTTTTGCTTTTATATTTTTCATACTTTGT
      CTTGATTACTTTTGTTGATAATAGCGAACGCACGGCGTTGATGATAATGATTCTGTTGTTGTA
      TTTCTACTCGCTATTGTTGTTGTGTTTGGTGGGGCGACAGGGCTAGAGATCAAGTTTTTAGT
6420 - CTAGATTATGCTCAATTGAAGGGCGCCAAGATGACATCGCGAAATTTTAAACGAATGAGAG
      GTAATATTAAGACTGACACCAATTTCTTTTTGCCTAATTCGCTGATAATGGAAGATATGAGCA
      TAACAAGAGAAGAGAGAAATAGAGAAAAATACCATTTAAGAACTATAAAAACTGGCGATA
      CCCCTTTGTTTGGGTCACACGTGATTGTTCTTTATTTTTAACTTATCACTTCGCAAGCTTAAT
      AAGAGTTTTTTTATTTATTAGTAATGCTGTATAATAATTTTATTGCAACCCTCAAACAATTAAG
6731 - TAATCAATACAGAGGTCATAAAATGTACAAGTATCTAATGTTTTTTATGGCCGCCAACTAAG
```

```
        AAAACCCCCAAACGACCTTCGCCACAGGCCAAGAACAATTTCAGTGGCAGTTCATAATAAT
        TTTTCAATAACCCCAAAGTGAGGCTCTAATTTCCAATTTCGAAGGTTTTGCCACACGCAATT
        CTAAACTGTTCTATACACCACCCCCTCCCTTGAACCCAACCAACCACCCACTGAAAATGTC
        AATTACTAAGCGTTTTTTTTTTGTTTTCCATTTTCGACTCTAAAGTTTGATCGATAAGGAATT
7040 -  TGCGTGCTTGTCTTGTGAGGTCGTTGTCGTATCAGCGAGCTTTTGAAAGGAAGGGGCCCG
        AGTTAATAGTGATAAGGAGCGTATAAATCAAGTGGAACCAACAGAAACTAAGCGAAAGCG
        ATTTAGATTCCGCCCAGACTGAGATAAGCTTAGTGGGAGACAATTGGCAACAGTTTTTTGT
        ACCTACAGTACGGCTTCTATATATAGCACGATATATTTCCTTAATAGTAAACTAAACTACCTTT
        TTAGAATACATGATCTTTGAAACAAAGTAATTAATCTAGATAGGTCCAGGTTTTCAATTTTAT
7348 -  AACATGGCTTTAAATTTGATTATGTTTAATCTACGAAATCCGTACGATAAGCGAATAATAAAA
        GCGAAAAAGAATGTTCTAATCAAACATTTAGGAAAATAAACAAAATCCAAAAAAGTGTGA
        AACTGGTGATTTCAATTAGAGAAGTACGAC
```

APPENDIX B

Chromatin immunoprecipitation % input calculation equations:

$$Sample\,\%\,input = \frac{100}{2^{\Delta Ct[NormalizedChIP]}}$$

$$\triangle Ct[NormalizedChIP] = (Ct[ChIP] - (Ct[Input] - \log(DilutionFactor)))$$

$$DilutionFactor = \frac{VolumeIP}{VolumeInput} \times InputDilutionFactor$$