UNIVERSITAT DE BARCELONA

Role of Capicua-L in RTK signaling and endocycle regulation

Laura Rodríguez Muñoz

Role of Capicua-L in RTK signaling and endocycle regulation

Laura Rodríguez Muñoz

Doctoral Dissertation
2020

Tesi Doctoral

Universitat de Barcelona

Role of Capicua-L in RTK signaling and endocycle regulation

Memòria presentada per

Laura Rodríguez Muñoz

per optar al grau de

Doctora per la Universitat de Barcelona

Programa de Genètica

Aquesta tesi ha estat realitzada al Departament de Biologia del Desenvolupament de l'Institut de Biologia Molecular de Barcelona (Consejo Superior de Investigaciones Científicas), sota la supervisió del Dr. Gerardo Jiménez.

Director de tesi	Tutor de tesi	Doctoranda
		
Dr. Gerardo Jiménez	Dr. Florenci Serras	Laura Rodríguez

Barcelona, desembre 2020

The HMG-box protein Capicua (Cic) is a conserved transcriptional repressor with key functions downstream of receptor tyrosine kinase (RTK)-Ras-MAPK signaling pathways. In both *Drosophila* and mammals, Cic is expressed as short (Cic-S) and long (Cic-L) isoforms that differ in their N-terminal regions. However, the significance of this difference or whether Cic-S and Cic-L have different functions or regulation is not known. This is because most of the work carried out so far has been done on Cic-S or using approaches that do not discriminate between both isoforms. To address this question, we have compared the expression of both isoforms during *Drosophila* development and obtained CRISPR-induced mutations specifically affecting Cic-L. We find that Cic-L acts redundantly with Cic-S in RTK processes such as wing vein specification and dorsoventral patterning of the embryo. In addition, Cic-L exerts individual functions regulating germline cell growth and development during oogenesis. Specifically, Cic-L accumulates in nurse cell nuclei during mid-oogenesis and is necessary for nurse cell endocycle termination and massive transfer or "dumping" of nurse cells contents into the oocyte in late oogenesis. In addition, we show that endocycle exit precedes the initiation of nurse cell dumping and propose that Cic-L enables nurse cell dumping by triggering endocycle exit. Cic-L exerts this control, at least in part, by promoting stabilization of Cyclin E, a key regulator whose periodic oscillations drive endoreplicative cycles, and downregulation of the Myc cell-growth factor. We also find that these unique functions of Cic-L primarily depend on its specific N-terminal module, which contains three conserved domains –NLS, Tudor-like and N1– that contribute additively to function. In contrast, other domains shared with Cic-S –the HMG-box and C1 DNA binding domains and the C2 motif necessary for MAPK-dependent downregulation– are largely dispensable for Cic-L-specific activity. Finally, we note that basal metazoans including sponges possess truncated "Proto-Cic" variants composed only of Cic-L N-terminal sequences, without the characteristic Cic HMG-box domain. Thus, the Cic-L N-terminal region plays unexpected roles in cell growth and endoreplication that may resemble the ancestral activities of Cic-like proteins in evolution.

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviation	Full name
Ago	Archipelago
aos	*argos*
AP	Anterior-Posterior
ATXN1	ATAXIN-1
ATXN1L	ATAXIN-1-LIKE
Bcd	Bicoid
bp	Base pairs
BSA	Bovine Serum Albumin
Bwk	Bullwinkle
CkII	Casein kinase II
Cas9	CRISPR-Associated Nuclease 9
CBS	CIC Binding Sites
CDK	Cyclin-Dependent Kinase
CIC	Capicua
CIC-L	Capicua-Long
CIC-S	Capicua-Short
CRISPR	Clustered Regulatory Interspaced Short Palindromic Repeats
CycE	Cyclin E
DA	Dorsal-anterior
Dap	Dacapo
DAPI	4′,6-diamidino-2-phenylindole
DFS	Dominant Female Sterile
DNA	Deoxyribonucleic Acid
DUX4	Double Homeobox 4
DV	dorsal-ventral
DYRK1A	Dual-specificity tyrosine-phosphorylation-regulated kinase 1A
EGFR	Epidermal Growth Factor Receptor
EMSA	Electroforetic Mobility Shift Assay
ETS	E-Twenty-Six
ETV	ETS Variant
FBW7	F-box and WD repeat domain-containing 7
Flp	Flipase
FOXO4	Forkhead Box O4
FRT	Flipase Recognition Target
GLCs	Germline Clones
Gro	Groucho
Grk	Gurken
GSK3β	Glycogen Synthase Kinase 3β
ind	*intermediate neuroblasts defective*
HA	Hemagglutinin

HDR	Homology Directed Repair
hkb	*huckebein*
HMG	High Mobility Group
kni	*knirps*
LOH	Loss Of Heterozygosity
M	Molar
MAPK	Mitogen Activated Protein Kinase
MARCM	Mosaic Analysis with a Repressible Cell Marker
mirr	*mirror*
ml	milliliters
MMP	Matrix Metaloproteinase
mRNA	messenger Ribonucleic Acid
µg	micrograms
NF-κB	Nuclear Factor Kappa Beta
NHEJ	Nonhomologous end joining
NSC	Neural Stem Cells
Nos	Nanos
OPC	Oligodendrocyte Progenitor Cell
Osk	Oskar
PBS	Phosphate Buffered Saline
PEA3	Polyoma Enhancer Activator 3
PIP degron	PCNA-interacting-protein degron
p90RSK	p90 Ribosomal S6 Kinase
RNA	Ribonucleic Acid
RTK	Receptor Tyrosine Kinase
SCA1	Spinocerebellar Ataxia type 1
Src	
T-ALL	T-cell Acute Lymphoblastic Lymphoma
tll	*tailles*
tsl	*torso-like*
tub	*tubulin*
twi	*twist*
UAS	Upstream Activator Sequence
UTP	Uridine-5′-Triphosphate
UTR	Unstranslated Region
WT	wildtype
zen	*zerknüllt*

INTRODUCTION

The human body is made up of about 30 trillion cells, and each of them is constantly making decisions. They can decide whether to continue proliferating or stop dividing and initiate differentiation, whether to stay where they are or abandon their position and migrate somewhere else in the body and even whether to remain alive or die. Making the right choices at the right time and place is critical to the survival of the organism, but how does each of these 30 trillion cells know which decision to make at every moment? Working as highly sophisticated information-processing devices, cells have evolved the ability to process and integrate multiple inputs in order to produce a final output. These inputs include a wide range of stimuli such as developmental programs, inner state of the cell, signals from other cells and environmental conditions. For example, nurse cells, which a group of cells that support the maturation of the oocyte in *Drosophila* oogenesis, are developmentally programmed to die at the end of oogenesis. In contrast, under unfavorable environmental conditions, nurse cells initiate cell death in mid-oogenesis. Mechanistically, the information is transmitted by signal transduction proteins, which are organized in signaling pathways. These signaling pathways elicit specific cellular responses by inducing, for example, changes in the genes expressed by the cell or modifications in the activity of proteins. In agreement with this key function in the transmission of information, unregulated activity of signaling pathways is associated with multiple congenital and acquired diseases, as seen in many types of cancer.

In this thesis we have focused on the tumor suppressor Capicua (Cic), which controls gene expression downstream of the Receptor Tyrosine Kinase (RTK)-RAS-MAPK signaling pathway. In particular, we have investigated the functions and mechanism of action of the unexplored long isoform of Cic using *Drosophila melanogaster* as a model.

1. Capicua

CIC is an evolutionary conserved HMG-box transcriptional repressor initially identified in *Drosophila* embryogenesis because of its role downstream of Torso RTK signaling (Jiménez et al., 2000). In this context, loss of Cic function results in embryos that lack most of their trunk and abdominal regions and only maintain the terminal (head and tail) regions; hence the name Capicua, which means "head-and-tail" in Catalan. Subsequent studies in *Drosophila* and mammals have placed CIC as a key effector downstream of the RTK/RAS/MAPK signaling pathway. CIC

functions as a default repressor of genes regulated by RTK/RAS/MAPK signaling. In the absence of signaling, CIC binds to specific DNA sites in the promoters of these genes and keeps them silenced (Fig. 1A). Upon RTK activation, CIC is phosphorylated and downregulated by MAPK. As a result, CIC-mediated repression is relieved, allowing the expression of the target genes in response to tissue-specific or ubiquitous activators (Fig. 1B). Consequently, mutations that eliminate CIC function derepress RTK-responsive genes and simulate a constitutive activation of the pathway (Fig. 1C).

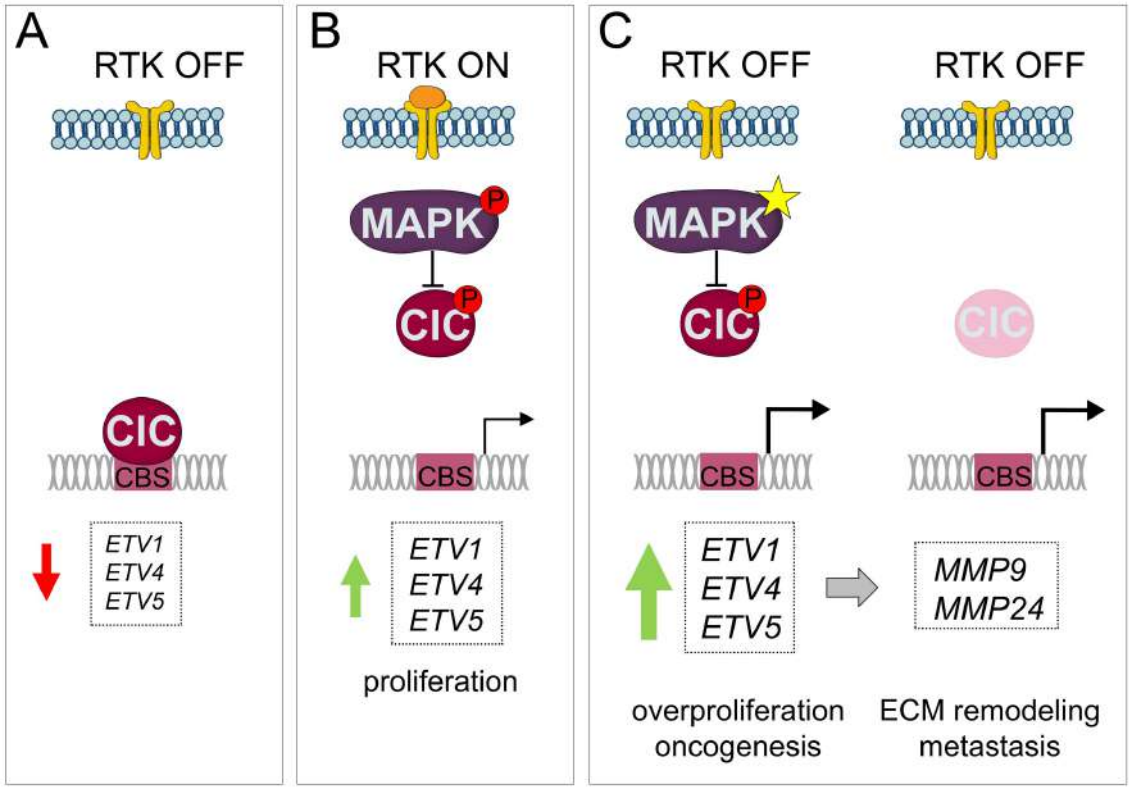

Figure 1. Model of RTK-dependent control of gene repression by CIC. (A) In the absence of RTK signaling, CIC represses RTK-induced genes. **(B) RTK** activation results in CIC downregulation and derepression of its targets, including the *ETV1/4/5* family of proto-oncogenes. **(C)** Both mutations that cause hyperactivation of the RTK signaling pathway and mutations that inactivate CIC lead to upregulation of *ETV1/4/5* genes, which in turn activate transcription of pro-metastatic matrix metalloproteinases (*MMPs*).

The most studied targets of mammalian CIC are the genes that encode the members of the PEA3 subfamily of ETS transcriptional factors: ETV1, ETV4 and ETV5 (Dissanayake et al., 2011; Kawamura-Saito et al., 2006; Lee et al., 2011; Weissmann et al., 2018). The PEA3 transcription factors activate the expression of genes involved in cell proliferation, motility and invasion and they are overexpressed in many different cancer types (de Launoit et al.,

2006; Jané-Valbuena et al., 2010; Monge et al., 2007; Shepherd et al., 2001; Tomlins et al., 2007). Accordingly, mutations that inactivate CIC have been found to promote tumorigenesis and metastasis through overexpression of the *ETV* genes (Choi et al., 2015; Kim et al., 2018; Okimoto and Bivona, 2017; Simón-Carrasco et al., 2017). In addition, hyperactivation of the RTK/RAS/MAPK pathway has also been shown to cause tumorigenic derepression of *ETV* genes due to downregulation of CIC (Bunda et al., 2019).

1.1 CIC functions in mammalian development and homeostasis

CIC is highly conserved from cnidarians to vertebrates (Jiménez et al., 2012a) and shortly after the discovery of *cic* in *Drosophila* in 2000, human and murine orthologues were also described (Lee et al., 2002). They were initially found to be expressed in the cerebellum, hippocampus and olfactory bulb during neural development, but additional studies have revealed that CIC is actually expressed in various embryonic and postnatal tissues in mice (Kim et al., 2015; Lee et al., 2011). Moreover, in both *Drosophila* and mammals, *Cic* encodes two main isoforms of different size, Capicua-Short (CIC-S) and Capicua-Long (CIC-L), which are generated via use of alternative promoters and splicing sites (Forés et al., 2015; Jiménez et al., 2012a; Lam et al., 2006).

The first functional studies on mammalian CIC by the group of Dr. Zoghbi investigated the role of CIC in the pathogenesis of spinocerebellar ataxia type 1. As it is explained below, these studies showed that CIC forms transcriptional repressor complexes with Ataxin-1 (ATXN1) and its paralog Atxn1-like (ATXN1L, also known as Boat) (Bowman et al., 2007; Crespo-Barreto et al., 2010; Lam et al., 2006; Lee et al., 2011). Subsequently, the physiological roles of CIC in mammals have been studied analyzing the phenotype of several CIC loss of function mouse models. First, the group of Dr. Zoghbi generated a *Cic-L*$^{-}$ mouse model based on the insertion of a β-geo genetrap cassette immediately downstream of *Cic-L* exon 1(Fig. 2A). This resulted in complete abolishment of Cic-L expression, but it also led to a 90% reduction of Cic-S expression in the cerebellum. In fact, they described it as a sever hypomorph for the entire *Cic* gene (Fryer et al., 2011). Instead, the other models are conditional *Cic* knockouts that harbor *loxP* sites flanking exons shared by both isoforms ((Lu et al., 2017; Simón-Carrasco et al., 2017; Yang et al., 2017)). Below I explain how the analysis of these mouse models has demonstrated that *Cic* is an essential gene with key roles in development, tissue homeostasis and tumor suppression.

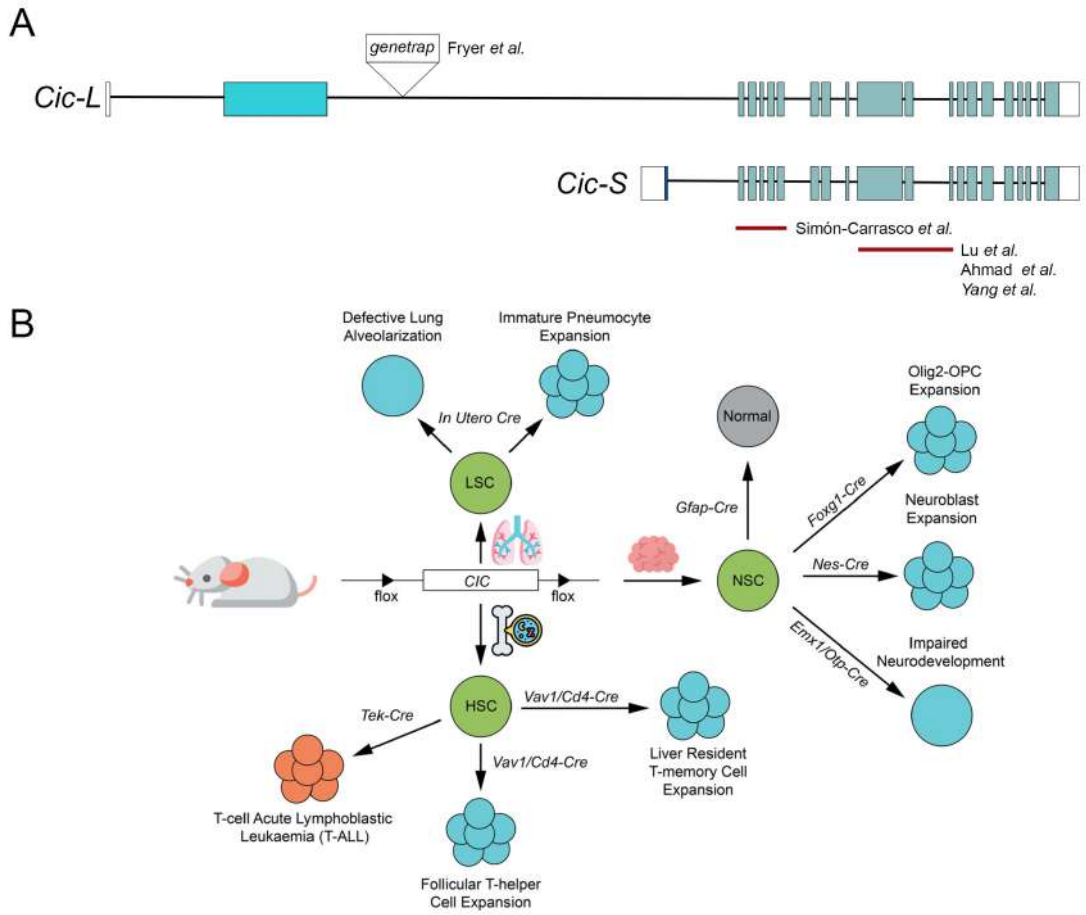

Figure 2. Mouse models of CIC loss of function. (A) Schematic representation of the *Cic-S* and *Cic-L* transcripts in mice. Colored and white boxes indicate coding and untranslated exons, respectively. Note that the two isoforms differ in their first exon. Red lines indicate the sequence deleted in the different conditional knockout models. Simón *et al.* generated an in-frame deletion that eliminates the HMG-box of CIC, resulting in a protein that cannot bind target genes. In contrast, the rest are frameshift deletions that lead to reduced protein levels. **(B)** Developmental consequences of *Cic* loss in different murine tissues. LSC, lung stem cell; NSC, neural stem cell; HSC, hematopoietic stem cell. Adapted from *Wong et al.* 2020.

Several independent studies have reported that homozygous *Cic* knockout mice are present at Mendelian ratios at E18.5, but they die before weaning age (P21) (Lee et al., 2011; Simón-Carrasco et al., 2017; Yang et al., 2017). At E18.5, homozygous *Cic* null embryos are already significantly smaller and most of them present an omphalocele, which is an abdominal wall closure defect (Simón-Carrasco et al., 2017). Omphaloceles occur normally during embryonic development but they are resolved by E16 (Doyonnas et al., 2001). When omphaloceles are not resolved, it usually results in perinatal lethality because the exposed organs are eaten by the mother when removing the placenta after birth (Thumkeo et al., 2005). Interestingly, a similar phenotype has also observed in about 45%

of *Atxn1*$^{-/-}$*; Atxn1L*$^{-/-}$ double mutant mice, suggesting that CIC-ATXN1/ATXN1L complexes play a role in abdominal wall closure regulation (Lee et al., 2011).

CIC-ATXN1/ATXN1L complexes are also necessary for correct late-stage lung development. *Cic* homozygous mutants have been found to present alveolarization defects, which become apparent during the last stage of lung development (P17-P23) (Lee et al., 2011). This phenotype has not been analyzed in *Atxn1*$^{-/-}$*; Atxn1L*$^{-/-}$ double mutants because of perinatal lethality but *Atxn1L* null mice also show similar defects. These defects in lung alveolarization have been attributed to derepression of *Pea3* genes, which leads to overexpression of matrix metalloproteinases (*Mmp*) genes and defects in extracellular matrix (ECM) remodeling. Specifically, Lee *et al*. found that derepression of *Etv4* due to *Cic* knockdown mediates overexpression of MMP9 in a mouse alveolar macrophage cell line(Lee et al., 2011). Subsequently, another study using a stronger allele of *Cic* detected defects in lung maturation at earlier stages of development (Simón-Carrasco et al., 2017). At E18.5 they observed a dramatic increase in cell proliferation and persistent TTF-1 expression, which indicates a delay or alteration in terminal differentiation of the respiratory epithelium. Moreover, these embryos have reduced numbers of type II alveolar cells and do not produce enough surfactant for postnatal survival.

CIC is also required at multiple stages of immune system development and function. First, Zoghbi and colleagues found that deletion of *Cic* in hematopoietic progenitors mediated by the *Tek-Cre* recombinase results in expansion of early T cell precursors in the thymus, which suggests that CIC plays a role in the early steps of T cell development (Tan et al., 2018). Second, CIC has also been implicated in the regulation of hepatic inflammatory responses and bile acid homeostasis. *Cic* knockout mice exhibit increased levels of bile acids in the liver and develop hepatic injury when fed a 1% cholic acid diet (Kim et al., 2015). This phenotype has been attributed to enhanced proinflammatory responses in the liver, which lead to downregulation of several hepatic transcriptional regulators. Interestingly, the enhanced proinflammatory response has been shown to be the result of excessive formation of liver-resident memory T cells (liver T_{RM}) (Park et al., 2019). Specifically, T-cell-specific loss of *Cic* promotes liver T_{RM} formation via derepression of *Etv5*, which results in upregulation of its target gene *Hobit*, a transcription factor that controls T_{RM} cell development. Finally, CIC has also been found to restrict follicular helper T (T_{FH}) cell differentiation through repression of *Etv5*, which in turn, activates the expression of a positive regulator of T_{FH} cell differentiation, *Maf* (Park et al., 2017). Consequently, *Cic*

ablation in the hematopoietic lineage or in T lymphocytes leads to an increase the population of T_{FH} cells, which results in autoimmunity-like phenotypes.

Consistent with the original observation that *Cic* is expressed during neural development, recent studies have revealed that Cic is an important regulator of neurodevelopment (Ahmad et al., 2019; Lu et al., 2017; Yang et al., 2017). However, CIC seems to function during a very specific time window because an initial study reported no major brain alterations after ablation of *Cic* in the entire brain driven by *hGFAP-Cre*, which is expressed in the central nervous system from E13.5 onward (Simón-Carrasco et al., 2017). In contrast, a second study reported that loss of *Cic* in the forebrain was lethal (Ahmad et al., 2019). In this case, they used a different source of Cre recombinase, *Foxg1-Cre*, which is expressed in the forebrain starting at E10.5. They found that at birth, mice were grossly normal but became runt by P7 and died by P22. In addition, *Cic*-null cerebra were smaller, and they suggested that the cause of death could be poor feeding as a consequence of impaired neurologic function.

In the developing brain, CIC expression levels vary according to the differentiation stage: nuclear CIC expression is low in neural stem cells (NSC) and it increases as the cells differentiate (Ahmad et al., 2019). Moreover, *Cic* deletion in NSCs results in increased symmetric divisions and a higher frequency of self-renewing cells. In addition, nuclear CIC levels in the central nervous system are also cell-type specific: they are high in neurons and astrocytes but low in oligodendrocytes (Ahmad et al., 2019). In this sense, both germline and forebrain-specific *Cic* knockout mice exhibit an aberrant expansion of oligodendrocyte progenitor cells (OPC) (Ahmad et al., 2019). The increase in the population of OPCs is caused by enhanced oligodendroglial differentiation at the expense of neural differentiation. Furthermore, *Cic* loss also compromises the differentiation of OPCs into mature oligodendrocytes, which remain arrested in an OPC-like state (Yang et al., 2017). As explained below, *CIC* is frequently mutated in oligodendroglioma and the presence of OPC-like cells is considered to be an early change in gliomagenesis. Additionally, both increased NSC self-renewal and lineage bias caused by loss of CIC function have been shown to be mediated by ETV5 (Ahmad et al., 2019).

On the other hand, a recent study has analyzed the behavioral consequences of eliminating *Cic* in different regions of the developing brain (Lu et al., 2017). Forebrain-specific deletion of *Cic* driven by the *Emx1-Cre* strain caused hyperactivity as well as learning and memory deficits. Interestingly, similar phenotypes were also observed upon deletion of *Atxn1* and

Atxn1l in the forebrain. Instead, *Cic* ablation in the hypothalamus and medial amygdala driven by *Opt-Cre* led to defects in social interaction. Consistent with these phenotypes, patients with heterozygous truncating mutations in *CIC* present several neurodevelopmental disorders, including intellectual disability, attention deficit/hyperactivity disorders and autism spectrum disorders (Lu et al., 2017; Tan and Zoghbi, 2019). Conversely, deletion of *Cic* in the cerebellum, where CIC complexes with polyglutamine-expanded ATXN1 are involved in the pathogenesis of Spinocerebellar Ataxia type 1, does not disrupt motor behavior nor cerebellar morphology (Lu et al., 2017).

Finally, whether mammalian CIC, like *Drosophila* Cic, has a role in restricting proliferation downstream of RTK/RAS/MAPK signaling is still controversial. On the one hand, *Cic* inactivation in mouse embryonic fibroblast leads to derepression of *Pea3* family genes but does not result in increased proliferation (Simón-Carrasco et al., 2017). It does not either restore proliferation in mouse embryonic fibroblasts devoid of Ras proteins. On the other hand, two different studies have reported that loss of *Cic* increases NSC proliferation (Ahmad et al., 2019; Yang et al., 2017). First, Yang and colleagues observed active cell proliferation in the subventricular zone of *Cic* deficient mice at P28, when postnatal neurogenesis has already finished in wildtype mice (Yang et al., 2017). Moreover, Edu labeling incorporation experiments have demonstrated that *Cic* ablation during embryogenesis increases proliferation cell-autonomously. *In vitro*, *Cic* inactivation in NSCs leads to increased proliferation and EGF-independent proliferation under hypoxia conditions. Therefore, it is possible that cell proliferation regulation by CIC in mammals is tissue-specific, but more studies are needed in order to address this question properly.

In summary, mammalian CIC exerts multiple regulatory functions in mammalian development and homeostasis that only recently have begun to be explored. Consistently, CIC alterations have been shown to be implicated in human diseases such as spinocerebellar ataxia 1 and cancer.

1.2 Roles of CIC in disease

1.2.1 Role of CIC in spinocerebellar ataxia 1

As mentioned above, mammalian CIC proteins form nuclear protein complexes with ATXN1, which is a glutamine-rich protein implicated in spinocerebellar ataxia 1 (SCA1) (Lam et al., 2006). SCA1 is a dominantly inherited neurodegenerative disease characterized

by progressive loss of motor coordination and balance. This disease is caused by an abnormal expansion of the polyglutamine tract in ATXN1, which leads to an early atrophy of the cerebellum followed by degeneration of the brainstem and spinal cord (Orr et al., 1993).

Subsequent studies have demonstrated that CIC proteins also interact with ATXN1L, which shares high homology with ATXN1 but lacks its polyglutamine tract (Bowman et al., 2007). ATXN1 and ATXN1L compete with each other to form large stable complexes with CIC (Bowman et al., 2007; Crespo-Barreto et al., 2010). CIC-ATXN1 and CIC-ATXN1L complexes seem to be redundant but their functional significance is not completely understood (Crespo-Barreto et al., 2010; Lee et al., 2011). On the one hand, ATXN1 and ATXN1L have been proposed to function as CIC corepressors because Co-expression of either ATXN1 or ATXN1L with CIC results in synergistic transcriptional repression of a luciferase reporter bearing CIC binding sites (Crespo-Barreto et al., 2010; Lam et al., 2006). However, ATXN1 and ATXN1L have also been shown to stabilize CIC at the protein level (Crespo-Barreto et al., 2010; Lam et al., 2006; Lee et al., 2011).

Interestingly, CIC interacts with polyglutamine-expanded ATXN1 in human cells in culture and in cerebellar extracts from a mouse model of SCA1, *Atxn1*$^{154Q/+}$; and this interaction has been shown to alter gene repression by CIC (Fryer et al., 2011; Lam et al., 2006). In particular, transcriptional studies comparing gene expression in the cerebellum of *Atxn1*$^{154Q/+}$ and *Cic* deficient mice revealed that polyglutamine expansion of ATXN1 causes hyper-repression of a subset of CIC targets genes but also derepression of another group of CIC targets (Fryer et al., 2011). This finding led to the proposal that both gain, and loss of function mechanisms could contribute to the pathogenesis of SCA1.

However, several phenotypic observations suggest that gain of function of the polyglutamine-expanded ATXN1-CIC complexes is the main driver of the pathogenesis. First, polyglutamine-expanded forms of ATXN1 that cannot bind CIC (ATXN1 [82Q] V519A; S602D) or that are not incorporated into CIC complexes *in vivo* (ATXN1 [82Q] S776A) prevent SCA1 pathogenesis (Lam et al., 2006; Rousseaux et al., 2018). Second, genetic reduction of *Cic* levels significantly improves the phenotype of the *Atxn1*$^{154Q/+}$ SCA mouse model. This improvement has also been observed after subjecting *Atxn1*$^{154Q/+}$ mice to an exercise routine, which leads to a reduction of CIC levels through activation of Epidermal Growth Factor Receptor (EGFR) signaling in the brainstem. In addition, polyglutamine-expanded ATXN1 forms toxic oligomers that correlate with disease

progression and the interaction with CIC is necessary for the formation and stabilization of these oligomers (Lasagna-Reeves et al., 2015).

1.2.2 Role of CIC in Cancer

In addition, in recent years CIC has emerged as a validated tumor suppressor (Simón-Carrasco et al., 2018a; Tanaka et al., 2017; Vogelstein et al., 2013; Wong and Yip, 2020). In fact, sequencing studies of tumor samples have revealed that *CIC* is one of the most frequently mutated tumor suppressors across a variety of cancer subtypes [(Campbell et al., 2020), Fig3A]. Next, I describe the role of *CIC* aberrations (point mutations but also chromosomal translocations) in tumorigenesis of three specific tumors: oligodendroglioma, T-cell acute lymphoblastic lymphoma and Ewing-like sarcomas. Moreover, I also explain recent findings showing that *CIC* alterations are implicated in tumor metastasis and therapy resistance to inhibitors of the RAS-MAPK signaling pathway.

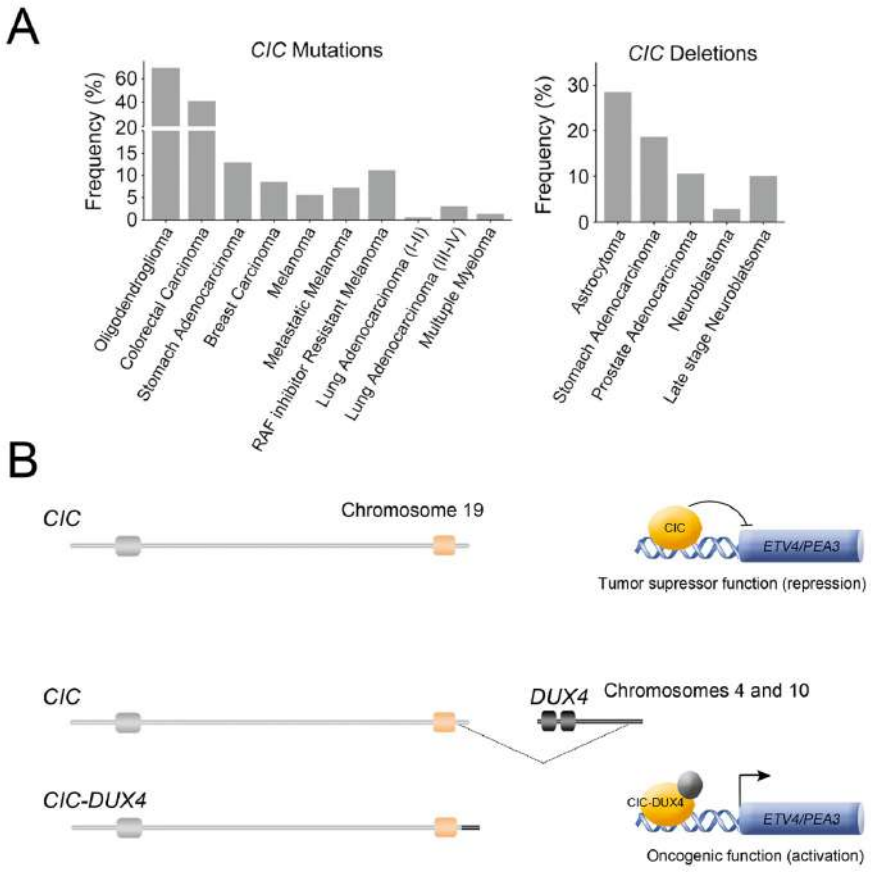

Figure 3. *CIC* aberrations identified in cancer. (A) Charts showing the frequency of *CIC* mutations and deletions (homozygous or heterozygous) in various cancer types. **(B)** Comparison of the structure and function of CIC and the oncogenic CIC-DUX4 fusion. Light grey and orange boxes represent the HMG-box and C1 domains of CIC, whereas the dark grey box represents the double homeodomain region of DUX4.

1.2.2.1 Primary tumors

Oligodendroglioma

As shown in Figure 3A, the highest incidence of *CIC* mutations in cancer is found in oligodendroglioma, a type of brain tumor that has been histologically classified as a diffuse low-grade glioma comprised of neoplastic cells that resemble oligodendroglia. These tumors are characterized by co-deletion of chromosomal arms 1p and 19q, which results in loss of heterozygosity (LOH) of these chromosomal regions (Louis et al., 2016). LOH of a chromosomal region in cancer usually indicates the presence of a tumor suppressor gene and Bettegowda and colleagues set out to identify the putative tumor suppressor gene(s) on chromosomes 1p and 19q (Bettegowda et al., 2011). They analyzed oligodendroglioma samples by exome sequencing and found that 18/26 (69%) oligodendrogliomas with 19q loss also exhibited *CIC* mutations in the remaining allele on chromosome 19q. Later, another study analyzed 29 additional oligodendrogliomas with 1p/19q co-deletion and reported similar results: 20/29 (69%) oligodendrogliomas contained *CIC* mutations (Yip et al., 2012). They showed that *CIC* mutations were highly specific of oligodendroglioma histology and were rarely observed in astrocytic tumors. This finding is consistent with the role of CIC in lineage specification since loss of CIC biases NSCs towards oligodendroglial fate specification (Ahmad et al., 2019). Moreover, oligodendroglioma cells exhibit features of OPCs and *Cic* loss also compromises OPCs differentiation into mature oligodendrocytes (Ahmad et al., 2019; Yang et al., 2017). Indeed, gliomas with *CIC* mutations express more OPC genes than gliomas with wildtype *CIC* (Gleize et al., 2015).

In oligodendroglioma *CIC* mutations are typically associated with *IDH1* mutations, and less frequently with *FUBP1* and *TERT* promoter mutations (Bettegowda et al., 2011; Jiao et al., 2012; Killela et al., 2013; Suzuki et al., 2015; Yip et al., 2012). IDH1 is a cytosolic isocitrate dehydrogenase that is recurrently mutated in gliomas. The most common *IDH1* mutation in gliomas is $IDH1^{R132H}$, which changes the catalytic properties of *IDH1*. Normally, IDH1 converts isocitrate into α-ketoglutarate, but as a result of this mutation IDH1 acquires a new catalytic activity that allows it to further convert α-ketoglutarate into the oncometabolite 2-hydroxiglutarate (2HG) (Dang et al., 2009). Chittaranjan and coworkers have studied the biological consequences of concomitant *CIC* and *IDH* mutations in oligodendroglioma. They found that ectopic expression of CIC^{R1515H} (a recurrent mutation in oligodendroglioma) increases cellular 2HG levels in HEK293 cells expressing $IDH1^{R132H}$ (Chittaranjan et al., 2014). Surprisingly, they also observed that CIC-

S was present in the cytosol in close proximity to the mitochondria in two oligodendroglioma cell lines that carry the 1p/19q co-deletion. In the cytosol, CIC was found to interact with ACLY (ATP-citrate lyase), an enzyme that converts citrate into oxaloacetate and acetyl-CoA. *CIC* mutations were described to affect the stability of ACLY in cell lines and were associated with reduced levels of ACLY in oligodendroglioma patient samples (Chittaranjan et al., 2014). Therefore, although more studies are required, it seems that mutations in *CIC* could contribute to the dysregulation of cellular energetics, a hallmark of cancer cells.

CIC mutations described in oligodendroglioma comprise insertions, deletions, alteration of splice sites, nonsense and missense mutations (Simón-Carrasco et al., 2018b; Yip et al., 2012). Different findings indicate that these mutations confer a selective growth advantage to oligodendroglioma cancer cells. First, different subclones within a single tumor have been found to acquire distinct *CIC* mutations independently, suggesting that there is a selective pressure to inactivate *CIC* (Suzuki et al., 2015). Second, 1p/19q co-deleted gliomas with *CIC* mutations were found to grow faster and had a poorer outcome than those with wildtype *CIC* (Gleize et al., 2015). Hence, CIC seems to function as a tumor suppressor in oligodendroglioma. However, recent studies have suggested that *CIC* inactivation contributes to tumor progression, but it is probably not the initiating event in brain tumorigenesis. Initially, Suzuki et al analyzed the variant allele frequency of different coexisting mutations in gliomas and proposed that while *IDH1* and *TERT* promoter mutations were founder events in gliomagenesis, *CIC* and *FUBP1* mutations were acquired later in tumor development (Suzuki et al., 2015). To directly test if *CIC* inactivation was sufficient to initiate oligodendroglioma development, Simón-Carrasco and coworkers eliminated CIC activity in the entire brain using a *hGFAP-Cre* strain. Animals were followed for 1 year and no tumors or brain alterations were detected (Simón-Carrasco et al., 2017). In contrast, *CIC* inactivation was reported to potentiate tumor development in an orthotopic mouse model of glioma driven by overexpression of PDGFB (Yang et al., 2017), suggesting that *CIC* inactivation can accelerate the growth of brain tumors driven by other cancer drivers, but cannot initiate brain tumorigenesis on its own.

T-cell acute lymphoblastic lymphoma

In contrast to oligodendroglioma, *CIC* inactivation does seem to play a role in the initiation of T-cell acute lymphoblastic lymphoma (T-ALL) (Simón-Carrasco et al., 2017; Tan et al., 2018). Simón-Carrasco and colleagues examined the effects of inactivating *Cic* in adult

mice using a tamoxifen-inducible *Cre* system (*hUBC-CreERT2*). They found that mice exposed to a tamoxifen diet developed T-ALL, which is a neoplasm of immature T-cell precursors (You et al., 2015). A subsequent study using a different allele of *Cic* also corroborated that ubiquitous deletion of *Cic* in adult mice leads to development of T-ALL before one year of age. Regarding the cell population in which Cic activity is required to suppress T-ALL formation, Park *et al*. initially reported that *Cic* deletion in hematopoietic progenitors driven by the *Vav1-Cre* strain did not cause T-ALL (Park et al., 2017). Subsequently though, Tan and coworkers observed that these animals did develop T-ALL, but with a delayed onset and incomplete penetrance (Tan et al., 2018). Moreover, they deleted *Cin* in hematopoietic progenitors using a different *Cre* allele, *Tek-Cre*, and found that mice developed fully penetrant T-ALL, indicating that loss of *Cic* in hematopoietic cells is sufficient to cause T-ALL.

The transcriptional profile of these tumors shows a significant overlap with the transcriptional profile of T-ALL tumors driven by Ras oncogenes and it is characterized by derepression of Cic target genes such as *Etv4* (Simón-Carrasco et al., 2017; Tan et al., 2018). Interestingly, simultaneous loss of *Etv4* dramatically reduces the incidence of T-ALL induced by CIC inactivation, pointing to ETV4 as a key effector in T-ALL transformation (Simón-Carrasco et al., 2017). Furthermore, these tumors display a transcriptional signature indicative of Notch pathway activation, which is known to play a prominent role in T-ALL development and has been shown to regulate *MYC* expression (Belver and Ferrando, 2016). Indeed, Tan *et al*. found that *Cic*-null T-ALL exhibited increased expression of *Myc* and were enriched for MYC targets (Tan et al., 2018).

Regarding the role of CIC in human T-ALL, sequencing analysis of 69 primary T-ALL samples revealed that only 10% of the tumors carried *CIC* mutations (Simón-Carrasco et al., 2017). However, it is also possible that CIC function in T-ALL is compromised through transcriptional or posttranscriptional mechanisms. For instance, hyperactive RAS/MAPK signaling has been associated with reduced CIC nuclear levels in lung adenocarcinoma and glioblastoma (Bunda et al., 2019; Okimoto et al., 2019). In fact, human T-ALLs frequently display aberrant RAS signaling (Von Lintig et al., 2000), and Simón et al. detected that human T-ALL carrying mutations predicted to activate the RAS/MAPK pathway exhibited a gene signature indicative of CIC inactivation (Simón-Carrasco et al., 2017).

Ewing-like sarcomas

Apart from loss of function mutations, *CIC* alterations in cancer also include chromosomal translocations that lead to the formation of chimeric proteins. Recurrent chromosomal translocations involving *CIC* were first identified in undifferentiated small round cell sarcomas where *CIC* was fused to the double homeodomain gene *DUX4* (Kawamura-Saito et al., 2006). Subsequently, it has been shown that *CIC*-rearranged sarcomas define a subset of Ewing-like sarcomas that are biologically and clinically distinct from Ewing sarcomas (Antonescu et al., 2017; Yoshida et al., 2016). *CIC*-rearranged sarcomas are highly aggressive and present an overall survival worse than that of Ewing sarcoma. In fact, Yoshimoto and coworkers recently generated a mouse model of CIC-DUX4 sarcoma and found that expression of a CIC-DUX4 chimera in embryonic mesenchymal cells led to the development of aggressive small round cell sarcomas with shorter latency than that of an Ewing sarcoma mouse model (Yoshimoto et al., 2017).

The CIC-DUX4 fusion protein carries the majority of CIC sequence, including the HMG-box and C1 domains (see Fig. 3B), fused in frame to the C-terminal trans-activating domain of DUX4. The resulting chimera behaves as an oncoprotein that retains the DNA binding specificity of CIC, but instead of repressing, it activates the transcription of CIC target genes, including *ETV1*, *ETV4* and *ETV5* (Forés et al., 2017a; Kawamura-Saito et al., 2006; Specht et al., 2014). Okimoto *et al*. have proposed that the CIC-DUX4 oncoproteins drive tumorigenesis and metastasis through upregulation of different targets of CIC: they found that cyclin E1 (*CCNE1*) upregulation was necessary for tumor growth while *ETV4* upregulation was involved in invasion and metastasis (Okimoto et al., 2019). Additional fusion events involving *CIC* and other partners have been also detected in Ewing-like sarcomas (*CIC-FOXO4*), peripheral neuroectodermal tumors (*CIC-NUTM1*) and angiosarcoma (*CIC-LEUTX*) (Huang et al., 2016; Le Loarer et al., 2019; Sturm et al., 2016; Sugita et al., 2014). However, whether these non-*DUX4* fusions also affect CIC repressor activity is currently unknown.

1.2.2.2 Metastasis

Recent studies have also implicated CIC in invasion and metastasis suppression in different types of cancer. In general, CIC inactivation has been found to drive metastasis through derepression of *PEA3* family genes. PEA3 transcription factors are frequently overexpressed in cancer and have been shown to promote metastasis through transcriptional activation of

various *MMPs*, which are the main enzymes involved in extracellular matrix (ECM) remodeling, a critical step in tumor metastasis (Conlon and Murray, 2019; de Launoit et al., 2006). In particular, CIC was first identified as a metastasis suppressor in a screen for novel mediators of non-small-cell lung cancer metastasis. Using an *in vivo* orthotopic model Okimoto *et al.* demonstrated that CIC inactivation promoted metastasis through upregulation of *ETV4* and subsequent activation of *MMP24* expression (Okimoto et al., 2017). Interestingly, in human tumors they found that reduced expression of CIC in lung adenocarcinoma and *CIC* genetic alterations in gastric adenocarcinoma were associated with high expression of *MMP24*, suggesting that the CIC-ETV4-MMP24 axis is also engaged in these tumors (Okimoto et al., 2017).

Another study investigated the role of CIC inactivation in a mouse model of hepatocellular carcinoma induced by treatment with the chemical carcinogen diethylnitrosamine. They found that *Cic* ablation in the hepatocytes driven by the *Abl-Cre* strain had no effect on tumor formation, but it increased the number of lung metastases (Kim et al., 2018). In addition, experiments in hepatocellular carcinoma cell lines showed that *Cic* deficiency promoted invasion and cell migration through derepression of *ETV4* and subsequent upregulation of *MMP1* (Kim et al., 2018). On the other hand, in prostate cancer, homozygous *CIC* deletions have been reported to be more frequent in metastases than in the primary tumor, suggesting that loss of *CIC* could contribute to metastatic progression of prostate cancer (Seim et al., 2017). In fact, knockdown of *CIC* has been found to increase invasion and cell migration in prostate cancer cell lines (Choi et al., 2015). Finally, as mentioned above, the oncogenic CIC-DUX4 fusion has also been found to drive ETV4-mediated metastasis in a mouse model of Ewing-like sarcoma (Okimoto et al., 2019).

1.2.2.3 Therapy resistance

The RTK/RAS/MAPK pathway is mutated in almost half of human cancers (Sanchez-Vega et al., 2018). Typically, these mutations cause an overactivation of the pathway, which has stimulated the development of several inhibitors that target different components of the pathway (Roberts and Der, 2007). Although some of these inhibitors have already been approved for cancer treatment, their therapeutic efficacy is limited by the development of drug resistance (Sanchez et al., 2018). Consistent with the role of CIC as a downstream effector of MAPK, several studies have shown that CIC inactivation can mediate therapy resistance to RTK/RAS/MAPK pathway inhibition. In relation to EGFR inhibitors, *CIC* knockdown has been found to enhance survival upon treatment with different EGFR

inhibitors in non-small lung cancer cell lines (Liao et al., 2017).. Regarding MEK1/2 inhibitors, Wang and colleagues conducted a genetic screen in *KRAS* mutant pancreatic cell lines and discovered that *CIC* deletion rendered the cells insensitive to MEK inhibition (Wang et al., 2017). They also corroborated this finding in other *RAS* and *BRAF* mutant cancer cell lines from different lineages and using a BRAF inhibitor. In addition, CIC has also been reported to modulate sensitivity to MEK inhibitors in NSCs and T-ALL cells derived from *Cic* knockout mice (Simón-Carrasco et al., 2017; Tan et al., 2018; Yang et al., 2017). Finally, a missense mutation in *CIC* has also been recently identified in a case of acquired resistance to a combined treatment with BRAF and MEK inhibitors, which is a common strategy to try to overcome therapy resistance (Da Vià et al., 2020). This is explained because CIC functions downstream of the signaling cascade.

In summary, during the last few years many studies have begun to investigate the role of *CIC* aberrations in cancer biology. These studies have established CIC as an important tumor and metastasis suppressor in various types of cancer. Still, CIC remains much less studied than other tumor suppressors such as p53 or PTEN.

1.3 Structural and functional conservation of CIC proteins

CIC is highly conserved from cnidarians to vertebrates and, many of the insights into the mechanisms underlying its activity and regulation originate from studies in *Drosophila*. In fact, CIC proteins from *Drosophila* and mammals share numerous functional and structural properties. One common feature whose significance is currently unknown is the existence in both *Drosophila*, and mammals of two main isoforms, CIC-L and CIC-S (Forés et al., 2015; Jiménez et al., 2012a; Lam et al., 2006). The two isoforms are generated through alternative promoter usage, and as shown in Figure 4, share central and C-terminal sequences but differ in their N-terminal regions. The common region includes functionally important domains such as the HMG-box and C1 domains (see below), as well as the MAPK docking site (named C2 in *Drosophila*). In addition, each isoform also has specific domains: the Cic-L-specific Tudor-like and N1 domains, whose function is unknown; and the N2 domain, which is specific of *Drosophila* Cic-S and is necessary for repression in the early embryo (Forés et al., 2015). Although Tudor domains normally bind methylated lysines and arginines, the key methyl-binding residues do not seem to be conserved in the Tudor-like domain present in Cic-L (Faure and Callebaut, 2013). In the next sections, I provide an overview of the molecular mechanisms underlying CIC function and regulation in both *Drosophila* and mammals.

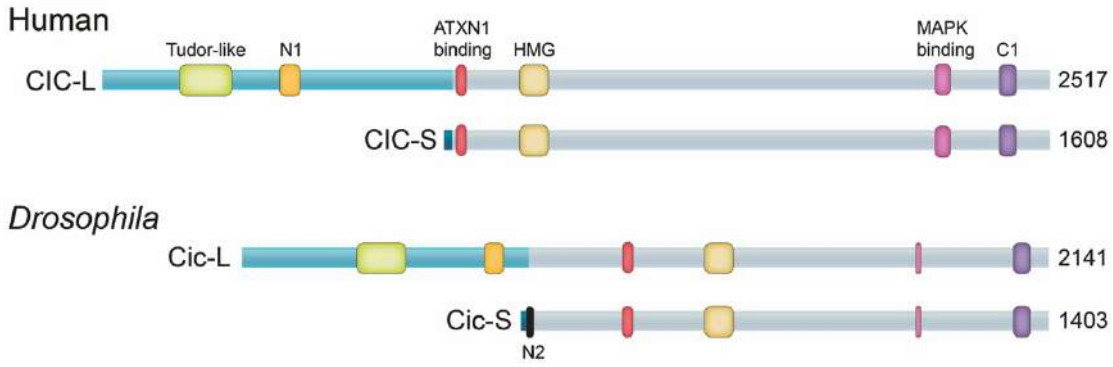

Figure 4. Structure of CIC-L and CIC-S isoforms in *Drosophila* and humans. CIC-L and CIC-S are expressed from alternative promoters and share common C-terminal sequences (in grey) but differ in their N-terminal region (CIC-L in light blue and CIC-S in dark blue). Conserved domains are depicted as colored boxes. Note that the MAPK docking site described for human CIC is different from the domain originally identified in *Drosophila*. Numbers indicate amino acids positions.

1.3.1 Mechanism of CIC DNA binding

Initially, Kawamura-Saito *et al*. reported that human CIC recognized TGAATG(A/G)A octameric sites in the promoters of the *ETV5* and *ETV1* genes (Kawamura-Saito et al., 2006). Subsequently, similar octameric sites were also identified in the promoters and enhancers of several targets of Cic in *Drosophila*: *tailless* (*tll*), *huckebein* (*hkb*), *intermediate neuroblasts defective* (*ind*) and *argos* (*aos*)(Ajuria et al., 2011). Moreover, mutation of these sites resulted in ectopic expression of Cic target genes (see for example *hkb* in Fig. 5B). Therefore, both mammalian and *Drosophila* Cic proteins seem to repress transcription by binding to canonic TGAATGAA-like motifs present in the promoters of target genes. In addition, it has been recently shown that Cic can also bind to non-canonical/suboptimal sites via cooperative binding with Dorsal/NF-κB (Papagianni et al., 2018).

On the other hand, Cic has been found to employ a unique mode of DNA binding that distinguish it from other HMG-box factors (Forés et al., 2017b). Specifically, Cic recognition of octameric sites requires two separate domains of the protein: the HMG-box and the C1 domain. Moreover, in contrast to TCF HMG-box factors, binding of the HMG-box-C1 module to octameric Cic sites does not depend on their flanking sequences. Accordingly, a deletion of 4 amino acids in the C1 domain of the *Drosophila* Cic protein (*cic*^{4} allele) has been reported to cause multiple developmental defects associated with derepression of Cic target genes such as *tll* [(Forés et al., 2017b) and Fig. 5D]. Interestingly, this unique mechanism of DNA binding also explains the pattern of missense mutations found in oligodendroglioma samples, which cluster in both of these domains (Fig. 5F). In

fact, several of these mutations have been shown to impair Cic DNA binding in electrophoretic mobility shift assays (EMSAs) [(Forés et al., 2017b) and Fig. 5E]. Finally, it also explains why the C1 is usually preserved in CIC-DUX4 fusion chimeras (Fig. 5F).

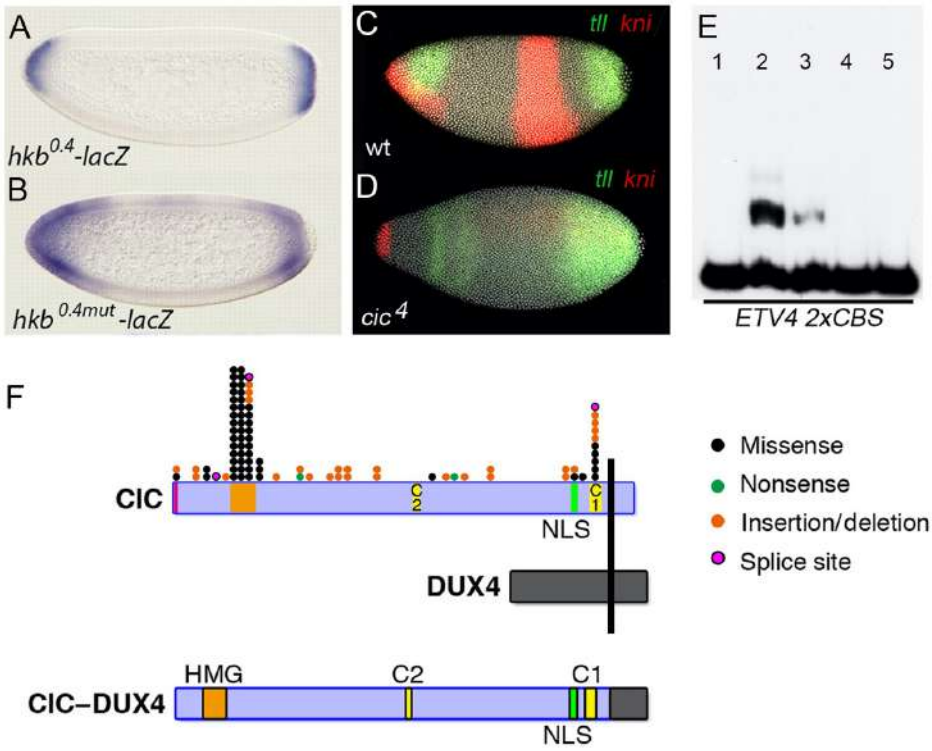

Figure 5. CIC recognizes octameric TGAATGAA-like sites via its HMG-box and C1 domains. (A,B) Embryonic mRNA expression patterns of *hkb-lacZ* reporters containing intact (A) or mutated (B) Cic binding sites. **(C,D)** mRNA transcripts of *tll* (green) and its target *knirps* (red) mRNA in wildtype (C) and maternally mutant *cic*^{4} (D) embryos; nuclei are labeled with DAPI (grey). **(E)** EMSA experiments analyzing how different missense mutations reported in oligodendroglioma affect the binding of the CIC HMG-box-C1 module to an *ETV4* DNA probe containing two CIC Binding Sites (CBSs) Lanes: (1) Negative control without added protein; (2) wildtype construct; (3-5) constructs carrying mutations R201W (3), R215W (4) and R1515L (5). **(F)** Diagrams representing CIC mutations found in oligodendroglioma (note that missense mutations tend to cluster in the HMG-box and C1 domains) and the structure of the CIC-DUX4 fusion protein, which preserves the C1 domain. Adapted from Ajuria *et al.* 2011, Forés *et al.* 2017b and Tanaka *et al.* 2017.

1.3.2 Mechanisms of CIC-mediated repression

Although multiple studies have investigated the mechanism by which DNA-bound CIC represses transcription, a model of CIC-mediated repression has not been clearly established. In fact, a study in *Drosophila* has revealed that Cic represses transcription via distinct mechanisms depending on the context (Forés et al., 2015). Specifically, they showed that Cic repressor activity in the early embryo relies on the Groucho (Gro)

corepressor, which is recruited through the Cic-S-specific motif N2. In contrast, Cic-mediated repression of *aos* in the developing wing and *mirror (mirr)* in the ovary was found to be Gro-independent. Importantly, as the N2 motif is only present in dipteran Cic-S isoforms, probably mammalian CIC proteins also function independently of Gro.

In this sense, several studies have indicated that mammalian CIC proteins form repressor complexes with ATXN1 and ATXN1L (Bowman et al., 2007; Crespo-Barreto et al., 2010; Lam et al., 2006; Lee et al., 2011). However, the precise roles of ATXN1/ATXN1L in CIC-mediated repression have not been defined. On the one hand, ATXN1/ATXN1L could function as CIC co-repressors since they have been shown to associate with nuclear corepressor factors such as nuclear receptor co-repressor 2 (NCOR2, also known as SMRT) and histone deacetylases 3 and 4 (HDAC3/4) [(Tong et al., 2011) and references therein]. Alternatively, ATXN1/ATXN1L might simply contribute to CIC-mediated repression via stabilization of CIC (see CIC regulation). Interestingly, the interaction of CIC with ATXN1/ATXN1L depends on a short motif that is highly conserved in vertebrates and is also recognizable in flies (Kim et al., 2013; Lam et al., 2006). In fact, the *Drosophila* ortholog of mammalian ATXN1 has been found to associate with Cic *in vivo* (Yang et al., 2016), raising the possibility that Cic acts through dAtxn1 in Gro-independent processes.

On the other hand, Weissmann *et al.* have recently showed that CIC interacts with several members of the SIN3 histone deacetylation complex, suggesting that CIC might repress transcription by inducing SIN3-mediated histone deacetylation (Weissmann et al., 2018). In line with this idea, another study has found co-enrichment of CIC, SIN3A and HDAC2 (two components of the SIN3 complex) within the promoters of several targets of CIC. Moreover, loss of CIC resulted in both reduced occupancy of SIN3A and HDAC2 as well as increased H3K9 and H3K27 acetylation (markers of active transcription) at the promoters of CIC target genes (Hwang et al., 2020; Weissmann et al., 2018). Interestingly, they also found that CIC is able to recruit the chromatin remodeling complex mSWI/SNF to the promoters of its target genes(Hwang et al., 2020). However, whether this has an impact on chromatin organization is currently unknown.

1.3.3 CIC regulation

As illustrated above, CIC is a tumor suppressor with key functions in development and homeostasis, and its activities must be under tight control. Below, I explain how CIC is

regulated by RTK-dependent and independent mechanisms in both *Drosophila* and mammals.

Initial studies in *Drosophila* revealed that activation of the Torso-Ras-MAPK pathway induces the degradation of Cic at the poles of the *Drosophila* embryo (see Fig. 6A), which allows restricted expression of its targets *tll* and *hkb* (Astigarraga et al., 2007; Jiménez et al., 2000). In the early embryo Cic degradation occurs primarily in the cytoplasm and it has been proposed that Torso-Ras-MAPK signaling increases the degradation of Cic by triggering its exclusion from the nucleus (Grimm et al., 2012). Subsequent studies have shown that Cic also functions downstream of other RTK pathways at multiple stages of fly development (Ajuria et al., 2011; Andreu et al., 2012b; Astigarraga et al., 2007; Atkey et al., 2006; Goff et al., 2001; Jin et al., 2015; Roch et al., 2002; Tseng et al., 2007). Interestingly, although Cic is also partially relocalized to the cytoplasm in response to EGFR signaling, this is not always accompanied by its degradation [(Astigarraga et al., 2007; Jin et al., 2015) and Fig. 6E]. However, another study found that the Cic target *ind* is induced within minutes of MAPK activation, while Cic is still present in the nucleus (Lim et al., 2013). This observation has been recently explained by Keenan and colleagues through the combination of optogenetic activation of MAPK with time-resolved ChIP-seq (Keenan et al., 2020). Specifically, they show that in the early embryo MAPK signaling inhibits Cic-mediated repression in two-steps: a fast step in which MAPK induces the dissociation of Cic from its target genes; and a slower step in which Cic is exported to the cytoplasm and degraded. It also noteworthy, that competition with other MAPK substrates has been reported to be modulate regulation of Cic by MAPK in *Drosophila* (Kim et al., 2010; Kim et al., 2011).

Several studies have indicated that mammalian CIC also functions in connection with RTK signaling (Bunda et al., 2019; Dissanayake et al., 2011; Fryer et al., 2011; Liao et al., 2017; Okimoto et al., 2017; Simón-Carrasco et al., 2017; Tan et al., 2018; Wang et al., 2017; Yang et al., 2017). Regulation of mammalian CIC downstream of RTK signaling has been found to be mediated by several kinases: MAPK, but also p90RSK and c-SRC (Bunda et al., 2019; Bunda et al., 2020; Dissanayake et al., 2011). First, MAPK-dependent phosphorylation of CIC on serine 1382 and serine 1409 has been reported to prevent binding of CIC to importin-α4 (also known as KPNA3), an adaptor that is required for nuclear import (Dissanayake et al., 2011). However, the biological implication of this mechanism is unclear because the loss of importin-α4 binding does not result in cytoplasmic accumulation of CIC in stimulated cells. Second, both MAPK and p90RSK,

which itself becomes activated downstream of MAPK, have been found to mediate phosphorylation of serine 173 (S173) (Bunda et al., 2019; Dissanayake et al., 2011). Phosphorylation of S173 has been described to serve two main purposes. On the one hand, S173 phosphorylation has been found to promote binding of CIC to 14-3-3 proteins, which, in turn, appears to decrease the interaction of CIC with DNA (Dissanayake et al., 2011). On the other hand, phosphorylation on S173 has been recently found to promote degradation of DNA-bound CIC by the nuclear E3 ligase PJA1(Bunda et al., 2019). In contrast, c-SRC has been proposed to phosphorylate nuclear CIC that is not bound to DNA on tyrosine 1455, which would induce its nuclear export (Bunda et al., 2020).

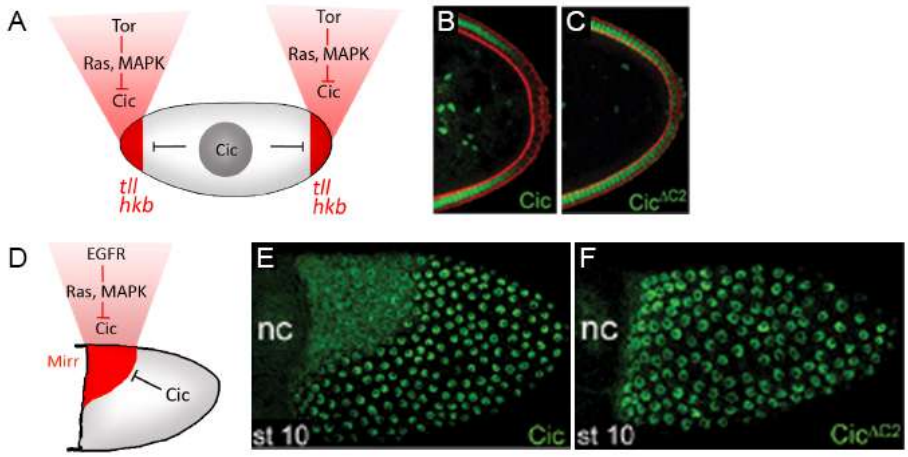

Figure 6. RTK-dependent downregulation of Cic in the early *Drosophila* embryo and ovarian follicle cells. (A) Model of Cic function downstream of the Torso signaling in the early Drosophila embryo. Torso signaling in the embryonic poles induces the degradation of Cic, which allows restricted expression of Cic target genes *tll* and *hkb*. **(B, C)** Protein distribution of Cic (B) and the $Cic^{\Delta C2}$ derivative (C) at the posterior pole of the *Drosophila* embryo. Cic accumulates in central regions of the embryo but not in the poles. In contrast, the MAPK-insensitive $Cic^{\Delta C2}$ is clearly detectable at the poles. **(D)** Model of Cic function downstream of EGFR signaling in ovarian follicle cells. EGFR signaling in dorsal-anterior follicle cells induces partial re-localization of Cic to the cytoplasm, contributing to expression of the Cic target gene *mirror*. **(E, F)** Protein distribution of Cic (E) and the $Cic^{\Delta C2}$ derivative (F) in ovarian follicle cells. Note how, in contrast to Cic, the $Cic^{\Delta C2}$ derivative remains nuclear in dorsal-anterior follicle cells.

Furthermore, MAPK-mediated phosphorylation of *Drosophila* Cic has been found to be direct and depends on a MAPK docking site in Cic named C2 (Astigarraga et al., 2007). Accordingly, mutant forms of Cic that lack the C2 motif are largely insensitive to RTK-mediated inactivation and produce gain of function phenotypes [(Astigarraga et al., 2007)and Fig. 6C, F]. However, the C2 domain is only moderately conserved in vertebrates, and in fact, photocrosslinking studies have identified an alternative MAPK docking site in human CIC (Futran et al., 2015).

On the other hand, Cic is also regulated by different MAPK-independent mechanisms in both *Drosophila* and mammals. For example, several studies have established that ATXN1/ATXN1L proteins control CIC protein stability (Crespo-Barreto et al., 2010; Lam et al., 2006; Lee et al., 2011). The precise mechanism underlying Cic stabilization by ATXN1/ATXN1L has not been defined, but it has been recently described that degradation of CIC upon loss of ATXN1L is mediated by the E3 ubiquitin ligase TRIM25 in an ERK-independent manner (Wong et al., 2020). Moreover, phosphorylation of *Drosophila* Cic by Minibrain, a kinase involved in growth control, has been shown to limit its transcriptional repressor activity (Yang et al., 2016). Interestingly, the ortholog of Minibrain, DYRK1A, has also been found to interact with human CIC in a proteomic screen (Weissmann et al., 2018). In addition, *CIC* has been reported to be downregulated by different microRNAs: bantam in *Drosophila*; miR-106b in renal carcinoma cell lines; miR-106b, miR-93 and miR-375 in prostate cancer cell lines; and miR-1307 in ovarian cancer cells (Choi et al., 2015; Herranz et al., 2012; Miao et al., 2019; Zhou et al., 2019). Finally, it has been suggested that CIC might auto-regulate its transcription because it has been found to bind its own promoter (Weissmann et al., 2018).

Therefore, the combination of *Drosophila* and mammalian studies has facilitated the dissection of many conserved mechanisms underlying CIC function and regulation. However, important questions regarding CIC biology remain still unresolved, most notably, the functional significance of the two CIC isoforms present in both *Drosophila* and mammals. This is because the investigations discussed above focused on CIC-S, on used approaches that do not discriminate between both isoforms (targeting common exons, for example). In fact, although CIC-L is has been shown to represent the ancestral form of CIC in metazoans (Forés et al., 2015), almost nothing is known about its molecular and developmental roles. In this thesis we have addressed this question by exploiting the powerful genetic tools of *Drosophila*. As we have found that Cic-L functions during oogenesis as a regulator of nurse cell endocycle termination, in the next sections I explain the oogenesis process in *Drosophila* and the main mechanisms involved the regulation of endocycles.

2. Overview of *Drosophila* oogenesis

Drosophila ovaries consist of 16-18 parallel ovarioles, which are like assembly lines that contain progressively maturing egg chambers (King, 1970). Egg chambers initiate their development at the anterior tip of the ovariole in a structure called the germarium and progress through the ovariole as they mature, reaching the posterior as mature eggs competent for fertilization. The entire process takes roughly one week and has been arbitrarily divided into 14 stages based on morphological criteria [(reviewed in (Bastock and St Johnston, 2008; Spradling, 1993), Fig. 7].

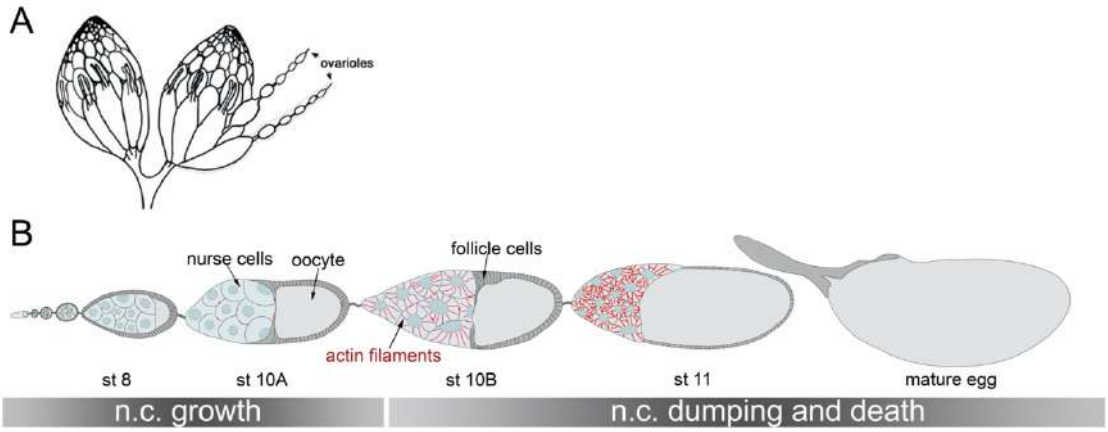

Figure 7. Overview of Drosophila oogenesis. (A) Drawing of the female ovary structure. *Drosophila* females have a pair of ovaries, each of which consists of 16-18 ovarioles. **(B)** Schematic representation showing egg chamber morphology at several stages.

In the germarium, each germline stem cell divides asymmetrically to renew itself, and to produce daughter cell named cystoblast. The cystoblast, in turn, undergoes four rounds of cell division with incomplete cytokinesis. This generates a cyst of 16 germ cells interconnected by cytoplasmic bridges known as ring canals. Among the 16 cells, only one will differentiate as the oocyte and complete meiosis. The other 15 cells, instead, become nurse cells, which adopt a polyploid fate and support the maturation of the developing oocyte. As the cyst progresses into the germarium, the oocyte becomes posteriorly localized and a group of somatic follicle cells encapsulate the oocyte and nurse cells. The cluster of oocyte and nurse cells surrounded by the monolayer of follicle cells constitutes the egg chamber or ovarian follicle.

The oocyte, which is arrested in meiotic prophase I, is basically transcriptionally quiescent and its maturation depends on mRNAs, proteins and organelles provided by the nurse cells and yolk proteins synthesized by the follicle cells and fat bodies. During stages 1-10 (early and mid-oogenesis) the nurse cells grow massively through endoreplication, which is thought to facilitate high levels of biosynthetic activity (discussed below). The products synthesized by the nurse cells during this period are slowly transported into the oocyte through ring canals in a selective manner. Instead, follicle cells in early oogenesis (stages 1-6) undergo mitotic proliferation to accommodate the growing germ cyst. As explained below, during this period the follicle cells present a cuboidal shape and become patterned along the anterior-posterior axis. Follicle cell replication ceases at the end of stage 6, when they also initiate endoreplication. At stage 8, follicle cells begin the synthesis of yolk proteins, which are taken up by the oocyte through endocytosis. As a result of the yolk uptake, the oocyte grows substantially and by stage 10A already occupies half of the egg chamber. Moreover, during this time period, follicle cells undergo extensive cell shape changes and morphogenetic movements [reviewed in (Horne-Badovinac and Bilder, 2005)]. At stage 9, the majority of follicle cells elongate and migrate towards the posterior to form a columnar epithelium covering the oocyte, while the remaining follicle cells stretch flat to cover the nurse cell cluster. At the same time, a group of 6-10 follicle cells, named border cells, delaminate from the anterior of the egg chamber and migrate posteriorly through the nurse cell cluster until they reach the anterior margin of the oocyte.

Late oogenesis is characterized by a dramatic increase in oocyte volume, secretion of the eggshell and finally, degeneration of the nurse and follicle cells. In contrast to early and mid-oogenesis, when mRNAs and proteins are selectively transported from the nurse cells to the oocyte, at the end of stage 10B the nurse cells actively contract and transfer, or dump, their whole cytoplasmic contents into the oocyte (Mahajan-Miklos and Cooley, 1994b). This rapid transport, which is called nurse cell dumping, results in a doubling of the oocyte volume in about 30 minutes and regression of the nurse cell cluster. Nurse cell dumping is accompanied by vigorous cytoplasmic flows in the oocyte (cytoplasmic streaming), which mix the incoming nurse cell cytoplasm with the contents of the oocyte (Quinlan, 2016). In addition, at stage 10B, in preparation for nurse cell dumping, actin filament cables are formed between the plasma membrane and the nucleus of nurse cells. These actin cables anchor the nurse cell nuclei in place during rapid transport to prevent them from blocking the ring canals. At stage 10B, it also begins the non-apoptotic programmed cell death (PCD) of the nurse cells with the permeabilization of the nuclear envelope (Cooley et al., 1992).

Subsequently, once dumping is completed, adjacent stretch cells engulf and eliminate the remnants of nurse nuclei (Mondragon et al., 2019; Timmons et al., 2016).

On the other hand, a group of anterior columnar follicle cells, named centripetal follicle cells, migrate inward during stage 10B to cover the anterior end of the oocyte. This way, the oocyte becomes completely surrounded by follicle cells, which secrete the eggshell between the epithelium and the oocyte membrane. The eggshell is a specialized extracellular matrix that protects the mature egg, and it is constituted by several layers [reviewed in (Waring, 2000)]. The synthesis of the eggshell begins at stage 9 with the secretion of vitelline membrane proteins and finishes at stage 14 when the deposition of the chorion has been completed. In addition, the eggshell presents several specialized structures, the most prominent of which are the dorsal appendages. The dorsal appendages are long tubes of chorion that extend from the anterior of the eggshell and facilitate gas exchange during embryogenesis. They arise from two patches of dorsoanterior columnar follicle cells and their morphogenesis takes places between stages 11 and 14. Finally, once the eggshell has been secreted, follicle cells also degenerate, leaving behind the mature egg.

2.1 Specification of body axes during oogenesis

In *Drosophila*, anterior-posterior (AP) and dorsal-ventral (DV) body axes are initially specified during the differentiation of the oocyte in the ovary and then become fully established in the very early embryo. Polarization of both axes during oogenesis depends on bidirectional signals between the oocyte and the follicle cells. Initially, the oocyte moves to the posterior end of the developing egg chamber, and through action of the EGFR ligand Gurken (Grk), which is associated with the oocyte nucleus, induces adjacent follicle cells to adopt a posterior fate (Cáceres and Nilson, 2005; González-Reyes and St Johnston, 1994; González-Reyes et al., 1995). The follicle cells, in turn, signal back to induce the reorientation of the oocyte's microtubules, which become oriented with their plus-ends at the posterior pole of the oocyte. This reorganization of the microtubule network is critical for the specification of both axes (Becalska and Gavis, 2009; Riechmann and Ephrussi, 2001). On the one hand, it determines the transport of the oocyte nucleus (and *grk* mRNA) in a microtubule-dependent manner to an antero-lateral position, where at stage 9, Grk induces dorsal follicle cell fate. On the other hand, it defines the final localization of maternal mRNAs that control AP polarity. Below, I provide a brief description of the different maternal systems that establish embryonic AP polarity.

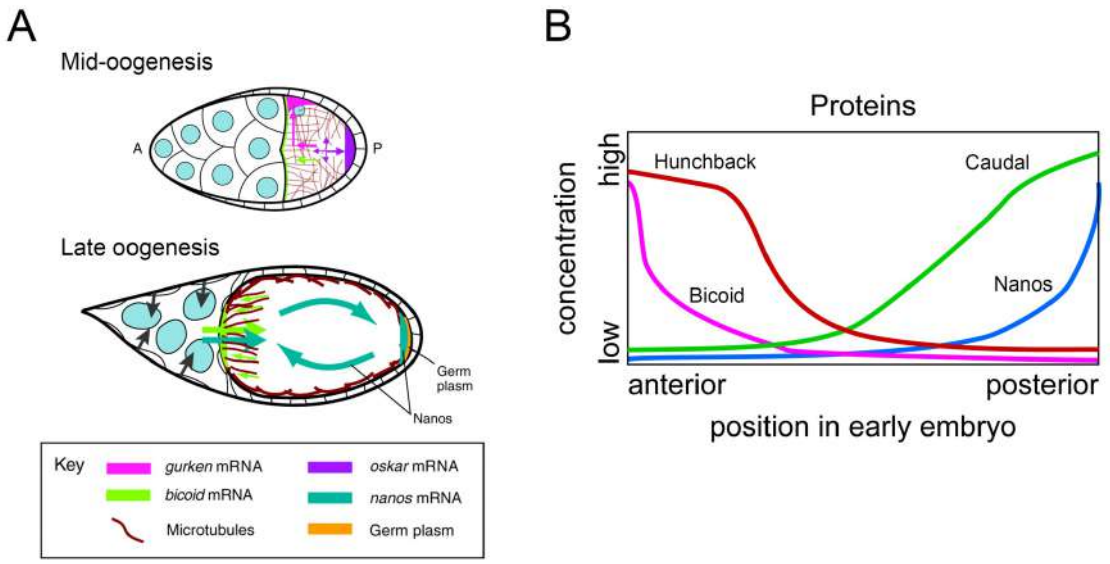

Figure 8. Maternal control of embryonic polarity. (A) Schematic representations showing the localization of patterning mRNAs during mid- and late oogenesis. In mid-oogenesis microtubule-dependent transport results in accumulation of *bcd* and *osk* at the anterior and posterior poles of the oocyte, respectively. *grk* localizes to the anterodorsal corner in close association with the oocyte nucleus, whereas *nos* is not yet localized. In late oogenesis, cytoplasmic streaming coupled with posterior anchoring brings about a further posterior enrichment of *osk* mRNA as well as posterior enrichment of *nos* mRNA. **(B)** Diagram indicating the protein distribution of Bcd, Hunchback, Nanos and Caudal along the AP axis of the early embryo. The Bcd and Hunchback protein form a gradient that extend from anterior to posterior, while the Nanos and Caudal gradients extend from posterior to anterior. Adapted from Becalska *et al*., 2009.

The AP body patterning of the *Drosophila* embryo is initiated through the action of three groups of maternal genes: the anterior group, which is necessary for the specification of the head and thorax; the posterior group, which is required for the formation of abdominal segments; and the terminal group, which is responsible for the specification of the embryonic poles [reviewed in (Manseau and Schüpbach, 1989)]. Bicoid (Bcd) is the main effector among the anterior group. As *bcd* mRNA transport within the oocyte is mediated by the minus-end-directed motor Dynein, the reorientation of the microtubule cytoskeleton results in its accumulation at the anterior pole of the oocyte (Fig. 8A). Then, after fertilization, *bcd* is translated to form an anterior to posterior gradient of Bcd protein (Fig. 8B). Bcd acts as morphogen activating particular zygotic genes at different threshold concentrations. In contrast, the transcript of posterior gene *oskar* (*osk*) is transported via the plus-end-directed motor Kinesin and becomes enriched at the posterior pole upon reorganization of the microtubule network (Fig. 8A). *osk* is translated when it reaches the posterior pole and directs the assembly of the germ plasm, which will give rise to the germline. In addition, the germ plasm anchors the transcript of *nanos* (*nos*) to the posterior

pole during late oogenesis (Fig. 8A). Upon fertilization, *nos* is translated generating a posterior to anterior gradient of Nos protein (Fig. 8B). This results in translation repression of *hunchback* transcripts by Nos in the posterior region of the embryo, allowing the formation of the abdominal segments in this region.

3. Endocycles

Endoreplication is a variation of the cell cycle in which cells replicate their genome, but do not proceed with cell division, resulting in duplication of the cell DNA content. Thus, successive cycles of endoreplication generate polyploid cells that contain multiple copies of the genome [(Edgar and Orr-Weaver, 2001; Fox and Duronio, 2013), Fig. 9]. The degree of polyploidization varies from one cell type to another and it is generally indicated as a chromatin value (C value), which denotes the DNA content as a multiple of the normal haploid genome. For instance, the giant neurons of *Aplysia californica* have been found to reach ploidies higher than 200,000C (Lasek and Dower, 1971). Moreover, in some specialized polyploid cells, the so-called polytene cells, the endoreplicated sister chromatids remain physically attached and produce giant visible chromosomes (Urata et al., 1995).

Endoreplication cycles are very common in nature and have been described in all eukaryotic kingdoms (Edgar et al., 2014). Often, the initiation of endoreplication is associated with the differentiation of mitotic progenitors into more specialized cells. This type of endoreplication is known as developmentally programmed endoreplication because the switch from mitotic to endoreplication cycles is induced by developmental signals (Lee et al., 2009; Orr-Weaver, 2015; Øvrebø and Edgar, 2018; Zielke et al., 2013). Developmentally programmed endoreplication is especially frequent in plants, where, among others, it has been shown to regulate cell fate commitment of *Arabidopsis* trichomes (Bramsiepe et al., 2010; de Veylder et al., 2011). In mammals, polyploidy was initially thought to be rare, but it has been recently shown that polyploidization plays an important role in the development of several mammalian cell types, including trophoblast giant cells, megakaryocytes, endometrial stromal cells, cardiac myocytes, hepatocytes and keratinocytes (Gandarillas et al., 2018). For example, megakaryocytes, the cells responsible for the production of platelets, become polyploid as part of their differentiation program and disruption of polyploidization results in reduced numbers of platelets (Trakala et al., 2015). Endoreplication has also been extensively studied in mammalian trophoblast giant

cells (TGCs), which form the outer layer of the placenta and facilitate the implantation of the embryo. Although TGCs become highly polyploid and reach ploidies of up to 512C, it has not been clearly established that polyploidization of TGCs is required for fetal viability (Chen et al., 2012; Ouseph et al., 2012).

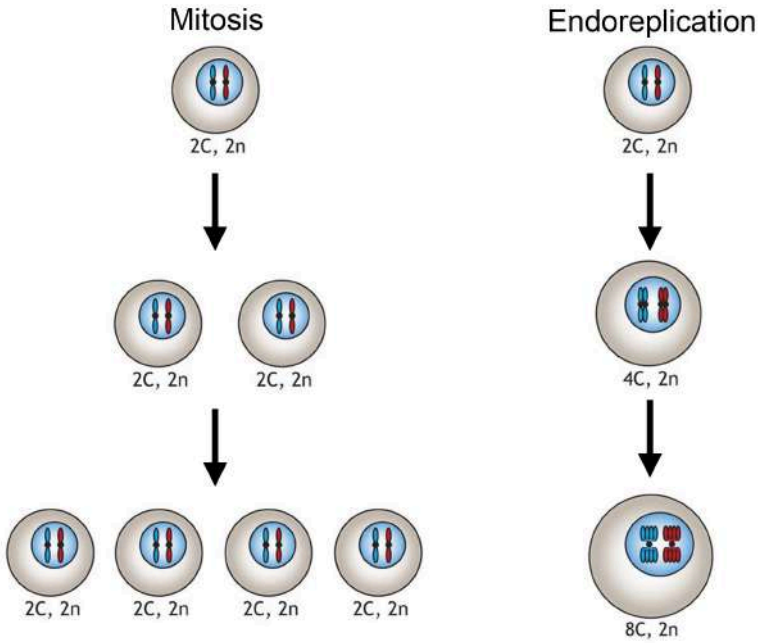

Figure 9. Comparison of mitotic cycles and endocycles. Schematic representation showing how each mitotic cycle results in the duplication of the number of cells whereas endocycles lead to the duplication of the DNA content within one cell.

Regarding the reasons why cells become polyploid, polyploidization has been proposed to serve two main purposes: cell growth and enhancement of cell biosynthetic capacity (Frawley and Orr-Weaver, 2015; Orr-Weaver, 2015). About cell growth, the size of a cell is in general proportional to its amount of nuclear DNA, and thus, increasing the DNA content provides a means to grow. Indeed, polyploidization is a recurrent evolutionary strategy for cell growth in differentiated cells that frequently results in the generation of large cells (Edgar et al., 2014; Orr-Weaver, 2015). With respect to the biosynthetic capacity, polyploidization often occurs in cells with high metabolic activity, like the nurse cells. It has been argued that augmenting the number of gene copies could potentiate gene expression and increase the metabolic output, but this hypothesis has not been formally tested yet (Frawley and Orr-Weaver, 2015).

Endoreplication cycles can be classified in two main groups depending on whether they retain some features of mitosis: endomitotic cycles and endocycles (Fox and Duronio,

2013). In endomitotic cycles, cells enter mitosis after DNA replication, but they fail to complete nuclear division and cytokinesis before re-entering the DNA synthesis (S) phase again. On the other hand, endocycling cells skip mitosis completely and alternate exclusively between gap (G) phases, during which gene expression and growth take place, and S phases. As endocycles are the primary form of endoreplication in arthropods (Smith and Orr-Weaver, 1991), below I focus on what is currently known about the mechanisms that drive this particular endoreplication cycle.

The molecular principles that control progression through the endocycle are essentially the same as the ones that drive the canonical G1-S-G2-M cycle. So, first I will briefly introduce how the canonical cell cycle is regulated to ensure that the DNA is replicated only once per cycle. In mitotic cells, progression through the different phases of the cell cycle is based on periodic activation and inactivation of specific cyclin/cyclin-dependent kinase (CDK) complexes. In animal cells in particular, mitosis is triggered by CDK1 (M-CDK), which is bound and activated by Cyclin B or Cyclin A. Instead, CDK2 (S-CDK) induces DNA synthesis (S phase) and is activated by Cyclin E or Cyclin A. As explained below, DNA replication is actually regulated in two steps, licensing and activation, and CDK2 exerts opposite effects on them.

Chromatin becomes licensed for replication through the assembly of pre-replication complexes (pre-RCs) at replication origins, which takes place before onset of the S phase. Pre-RC assembly involves initial binding of the ORC complex to replication origins followed by recruitment of Cdc6, Cdt1 (known as Double-Parked in *Drosophila*) and the MCM2-7 complex (Arias and Walter, 2007). From yeast to humans, high CDK activity inhibits pre-RC assembly, although different components of the pre-RC are targeted depending on the organism. In metazoans, the primary target for inhibition appears to be Cdt1 (Bell and Kaguni, 2013). Then, once the DNA is licensed, high CDK2 activity induces DNA replication by activating the MCM2-7 complex, which functions as the replicative DNA helicase (Tanaka and Araki, 2013). Finally, MCM2-7 moves away from the replication origins as the DNA is replicated and pre-RCs become dismantled (Arias and Walter, 2007). Therefore, high CDK activity triggers DNA replication from licensed origins but at the same prevents DNA re-replication by blocking the assembly of new pre-RCs. Once mitosis is completed, the levels of CDK activity drop and this allows the assembly of new pre-RCs during G1, licensing the DNA for a new round of replication (Arias and Walter, 2007).

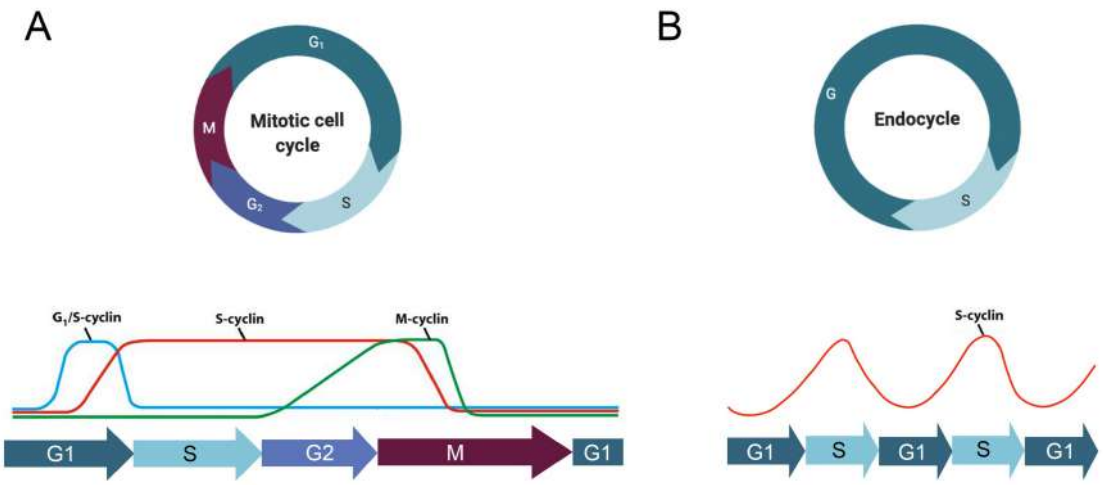

Figure 10. Regulation of the different cell phases in mitotic cycles and endocycles. (A) The mitotic cycle has four distinct phases: the G1 and G2 phases (growth and gene expression), the S phase (DNA replication) and the M phase (mitosis). Progression through the four phases is driven by oscillations of different Cyclin/CDK complexes. **(B)** Endocycling cells alternate between two phases: The G phase (growth and gene expression) and the S phase (DNA replication) but do not undergo cell division. High CYCE/CDK activity induces S phase entry whereas low CYC-CDK activity during the G phase is required for pre-RC assembly.

But how are these mechanisms adapted to convert a canonical mitotic cycle into an endocycle? Since the main difference between mitotic and endocycling cells is the absence of cell division, it is critical that endocycling cells suppress chromosome segregation and cytokinesis. This is generally achieved by selectively downregulating the activity of M-CDK while maintaining the activity of the S-CDK complex. In fact, experimental downregulation of M-CDK activity has been shown to induce endoreplication in *Drosophila* imaginal cells that normally do not endocycle (Hayashi, 1996; Weigmann et al., 1997). On the other hand, although endocycling cells do not divide after DNA replication, they still have to ensure that DNA is replicated only once per cell cycle. How this is accomplished during the endocycle has not been studied in detail, but several findings suggest that the mechanisms that regulate pre-RC assembly during the canonical mitotic cycle are also conserved in endocycling cells. First, all endocycles exhibit G phases, which presumably are necessary for pre-RC reassembly (Edgar and Orr-Weaver, 2001). Besides, the G phase coincides with a period of low Cyclin E (CYCE) protein levels, suggesting that high CYCE/CDK2 activity could inhibit the assembly of pre-RCs (Lilly and Spradling, 1996; Royzman et al., 1997; Weng et al., 2003). Indeed, continuous overexpression of CYCE has been shown to block endocycle progression in *Drosophila* salivary glands (Follette et al., 1998; Weiss et al., 1998). Overall, endocycling cells appear to have simplified the machinery that regulates the canonical G1-S-G2-M cell cycle to drive a G-S cell cycle. So,

progression through the endocycle can be explained by oscillations in CYCE/CDK2 activity: high CYCE/CDK2 activity promotes S phase entry and DNA replication while low CYCE/CDK2 activity during the G phase allows pre-RC assembly to license the next round of DNA replication.

3.1 Endocycle regulation in *Drosophila*

The endocycle is particularly well-characterized in *Drosophila*, where numerous studies have investigated the regulatory mechanisms that initiate and sustain endocycles. After an initial phase of cell proliferation during embryogenesis, most differentiated larval tissues in *Drosophila* enter the endocycle and become polyploid (Smith and Orr-Weaver, 1991). This way, larval growth is primarily achieved via increased cell size rather than increased cell number (Edgar and Orr-Weaver, 2001). Larval tissues that undergo polyploidization include the salivary glands, fat body, epidermis, gut, trachea and renal tubules. Instead, the nervous system and the precursors of adult organs remain diploid as they continue to undergo cell proliferation during larval development (Smith and Orr-Weaver, 1991). In the adult fly, endocycles are also employed for instance by glial cells, sensory bristles, and ovarian nurse and follicle cells (Audibert et al., 2005; Hammond and Laird, 1985a; Unhavaithaya and Orr-Weaver, 2012). Interestingly, each of these cell types is developmentally programmed to achieve a specific final ploidy. For example, while follicle cells undergo 3 endocycles and increase their DNA content to 16C, salivary gland cells and nurse cells reach final ploidies of about 1500C (Hammond and Laird, 1985a; Hammond and Laird, 1985b; Lilly and Spradling, 1996). However, very little is known about how this final ploidy is determined and the mechanisms that regulate endocycle termination.

3.1.1 Switch from mitosis to endoreplication

As discussed above, in order to switch from a mitotic cycle into an endocycle, cells must suppress M-CDK activity while retaining S-CDK oscillations and periodic pre-RC assembly. This can be achieved through multiple strategies that vary widely between organisms and cell types. I will focus on the specific mechanisms that regulate the transition from mitosis to endocycling in the follicle cells because this is the tissue in *Drosophila* where endocycle initiation has been studied the most. During oogenesis follicle cells undergo three different cell cycle variants (Calvi et al., 1998; Deng et al., 2001; Lilly and Spradling, 1996). At the beginning, they divide mitotically, giving rise to approximately 650 follicle cells by stage

6. Then, follicle cells exit the mitotic cycle and execute three endocycles. Finally, at stage 10B, follicle cells stop endoreplicating the whole genome, but continue to replicate specific loci that are necessary for eggshell formation in a process known as gene amplification.

Several studies have revealed that these cell cycle switches are regulated by the Notch pathway (Deng et al., 2001; López-Schier and St. Johnston, 2001; Sun et al., 2008). At stage 6, the oocyte and the nurse cells begin to express the Notch ligand Delta, and this activates Notch in the adjacent follicle cells (Deng et al., 2001; López-Schier and St. Johnston, 2001). Notch signaling, then, initiates the mitosis-to-endocycle transition by inducing the expression of the transcription factor Hindsight (Hnt) (Sun and Deng, 2007). Hnt, on the one hand, represses the expression of String, which is a phosphatase that activates the M-CDK complex (Sun and Deng, 2007). As a consequence, mitosis is blocked and follicle cells arrest in G2 (Deng et al., 2001). On the other hand, Hnt also represses the expression of the transcription factor Cut, which results in the accumulation of Fizzy-related (Fzr; also known as Cdh1) (Sun and Deng, 2005; Sun and Deng, 2007). Fzr is a positive regulatory subunit of the anaphase promoting complex/cyclosome (APC/C), which promotes degradation of mitotic cyclins and the pre-RC assembly inhibitor Geminin (Narbonne-Reveau et al., 2008; Sigrist and Lehner, 1997; Zielke et al., 2008). Therefore, accumulation of Fzr reinforces the mitotic block and also facilitates pre-RC assembly so that follicle cells can re-enter the S phase when CycE/Cdk2 activity reaches the level to fire the pre-formed pre-RCs.

3.1.2 Endocycle progression

Once cells have entered the endocycle, progression through multiple endoreplication cycles depends on oscillations of CycE/Cdk2 activity. The specific mechanisms that generate these oscillations vary between cell types but, in general, CycE/Cdk2 activity can be regulated at three main levels: transcription of *CycE*, proteasomal degradation of CycE and inhibition of Cdk2 activity.

In the salivary glands, where the majority of studies on endocycle progression have been conducted, the principal determinant of CycE/Cdk2 oscillations is periodic transcription of *CycE* [(Zielke et al., 2011) and Fig. 11]. This periodic transcription is achieved through cyclic accumulation and degradation of the transcription factor E2f1, which promotes *CycE* expression (Duronio and O'Farrell, 1995; Duronio et al., 1998; Royzman et al., 1997;

Zielke et al., 2011). During the G phase, E2F1 accumulates and activates the transcription of *CycE*. When CycE levels reach a threshold, the CycE/Cdk2 complex triggers the S-phase, and this, in turn, induces CRL4^{Cdt2}-mediated proteolysis of E2f1. Interestingly activation of the CRL4^{Cdt2} E3 ubiquitin ligase has been shown to require chromatin-bound PCNA, which only exits only at active replication forks (Arias and Walter, 2006; Shibutani et al., 2007; Shibutani et al., 2008). Finally, the destruction of E21 during the S phase leads to a reduction in *CycE* mRNA levels and creates a window of low CycE/Cdk2 activity that allows pre-RC assembly for the next round.

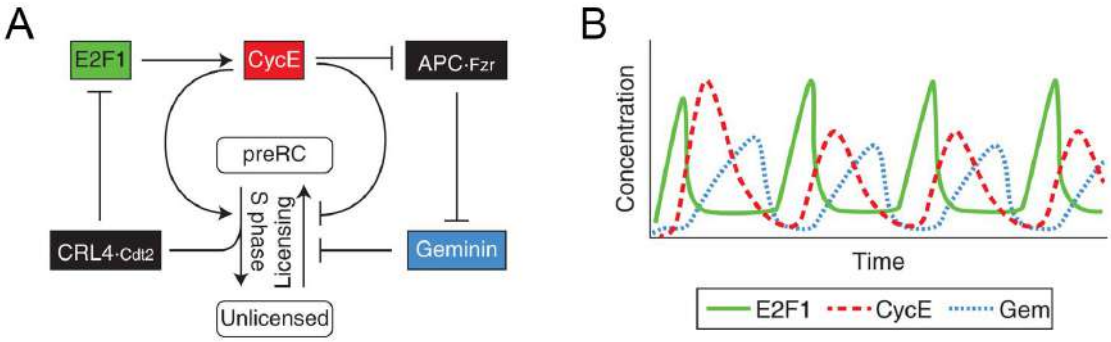

Figure 11. Model of endocycle progression in salivary glands. (A) The E2f1 activator accumulates during G phase and promotes expression of *CycE*. CycE, in turn, binds and activates Cdk2, which triggers DNA replication. Then, the DNA replication fork-dependent E3 ubiquitin ligase CRL4^{Cdt2} is activated and targets E2f1 for degradation. This allows CycE levels to decrease, which is necessary for pre-RC assembly during the G phase. In addition, CycE/Cdk2 suppresses DNA re-replication during the S phase through inhibition of the APC^{Fzr} complex, which allows the accumulation of the pre-RC assembly inhibitor Geminin. **(B)** Oscillations of E2f1, CycE and Geminin in wildtype salivary glands predicted by computational modeling. Adapted from Zielke *et al.* 2011.

Furthermore, studies in yeast, flies and mammals have revealed that CYCE is a conserved target of the SCF^{FBW7} E3 ubiquitin ligase complex (Koepp et al., 2001; Moberg et al., 2001; Strohmaier et al., 2001), which mediates the degradation of many cell cycle regulators. SCF^{FBW7} binds to its substrates through the FBW7 subunit [Archipelago (Ago) in *Drosophila*] and this binding requires previous phosphorylation of the substrates within conserved phospho-degron motifs (Welcker and Clurman, 2008). Loss of *ago* results in CycE accumulation and blocks endocycle progression in salivary glands and follicle cells (Doronkin et al., 2003; Shcherbata et al., 2004; Zielke et al., 2011). Notably, the levels of Ago do not oscillate during the cycle, which suggests that the critical event that regulates CycE degradation is its phosphorylation. Little is known about CycE phosphorylation in *Drosophila*, but studies in human cells have revealed that CYCE is phosphorylated on

threonine 380 (T380) by CDK2 and GSK3 (Koepp et al., 2001; Welcker et al., 2003). Although T380 phosphorylation is sufficient for FBW7 binding, additional phosphorylation of serine 384 (S384) increases the strength of the T380 degron. Interestingly, as S384 is exclusively phosphorylated by CDK2, the efficient degradation of CYCE efficient appears to be dependent on its own activity (Welcker et al., 2003).

Finally, CDK activity is also regulated by CDK inhibitor proteins (CKIs), which bind and inactivate the different Cyclin/CDK complexes. The CYCE/CDK2 complex is specifically inhibited by CKIs of the CIP/KIP family, which contains three members in mammals. In *Drosophila* there is only one Cip/Kip CKI, Dacapo (Dap) (De Nooij et al., 1996; Lane et al., 1996), which has been suggested to promote replication licensing by reinforcing low Cdk activity. Loss of Dap in endocycling cells results in prolonged S phases and reduced pre-RC assembly (Hong et al., 2007), but does not block the endocycle (Hong et al., 2003; Hong et al., 2007; Shcherbata et al., 2004; Zielke et al., 2011). These observations led Swanson and colleagues to propose that rather than being a core component of the machinery that drives endocycle progression, Dap modulates the frequency of endocycling (Swanson et al., 2015). In fact, Dap protein levels have been described to oscillate in several endocycling cell types: low levels in cells undergoing DNA replication and high levels in G phase cells (de Nooij et al., 2000; Hong et al., 2003; Swanson et al., 2015). These oscillations appear to be generated in coordination with endocycle progression by at least two mechanisms. First, CycE has been shown to promote the accumulation of Dap RNA and protein in endocycling nurse cells (de Nooij et al., 2000). Also, similarly to E2F1, Dap is targeted for destruction during the S-phase through a PIP degron that has been previously shown to mediate proteolysis by the $CRL4^{Cdt2}$ ubiquitin ligase (Swanson et al., 2015).

3.1.3 Endocycle termination

In contrast to the mechanisms that control endocycle initiation and progression, the mechanisms that determine the timing of endocycle termination and the final ploidy of a given cell type remain poorly understood. It has been argued, for example, that endocycle exit could be controlled by induction of CKIs or transcriptional downregulation of S-CDK activity, but these hypotheses have not been formally tested (Edgar et al., 2014; Øvrebø and Edgar, 2018). In fact, the regulation of endocycle exit in *Drosophila* has only been studied in detail in the follicle cells, which at stage 10B switch from endoreplication to amplification of specific loci (Calvi et al., 1998). Interestingly, while activation of the Notch

pathway initiates endocycling in the follicle cells, endocycle termination requires Notch downregulation (Sun et al., 2008). Notch downregulation at this stage induces, probably through activation of the Ecdysone Receptor (EcR), the expression of the transcription factor Tramtrack69 (Ttk69). Ttk69 is critical for the endocycle/gene amplification (E/A) switch and its overexpression during mid-oogenesis causes premature exit from the endocycle. Since Ttk has been shown to suppress the expression of CycE in other contexts (Audibert et al., 2005; Badenhorst, 2001), Sun *et al.* proposed that Ttk69 up-regulation during the E/A switch could be important for lowering the levels CycE/Cdk2 activity (Sun et al., 2008). They reasoned that higher levels of CycE/Cdk2 activity are probably required for genomic endoreplication than for gene amplification and Ttk69 would reduce CycE/Cdk2 activity to a level too low to initiate an additional round of endoreplication but sufficient to support gene amplification.

3.1.4 Regulation of final ploidy

In most cases, polyploidy cells appear to be developmentally programmed to achieve a specific ploidy that allows them to perform their function correctly. For example, reduced polyploidization of subperineurial glia (SPG) cells, which surround the neurons in *Drosophila*, compromises the integrity of the blood-brain barrier, while mutants that develop oversized brains show higher levels of SPG ploidy and maintain the functionality of the barrier (Unhavaithaya and Orr-Weaver, 2012). Although the precise molecular mechanisms that determine the final ploidy of each cell type are not well understood, the degree of polyploidization is believed to depend on the period of time during which the cells are endoreplicating and the speed of the endocycle (Edgar et al., 2014; Øvrebø and Edgar, 2018).

Interestingly, the *Drosophila* homolog of Myc has been involved in the regulation of both the time window and the rate of endocycling. For example, the distal cells of the salivary gland have been described to undergo one more endoreplication cycle than the proximal cells and this is associated with longer persistence of Myc in these cells (Pierce et al., 2004). Furthermore, in overexpression experiments Myc has been found to increase the nuclear DNA content of polyploid cells (Demontis and Perrimon, 2009; Pierce et al., 2004; Shcherbata et al., 2004; Unhavaithaya and Orr-Weaver, 2012). In the fat body in particular, where cells usually reach a ploidy of 256C, the overexpression of Myc can drive the cells to a ploidy of about 2048C (3 additional endocycles). Analyzing this effect in more detail by quantifying the DNA content at different time points, Pierce and colleagues showed that

Myc-overexpressing cells contained higher DNA contents at all time points tested (Pierce et al., 2004). Additionally, Myc-overexpressing cells continued endoreplicating after wildtype cells had already stopped. Conversely, loss of Myc function results in reduced endoreplication in several polyploid cell types, including the nurse and follicle cells (Demontis and Perrimon, 2009; Maines et al., 2004; Pierce et al., 2004).

However, very little is known about how Myc promotes endocycling. Considering the well-established role of Myc in promoting cell growth and the strong link between cell growth and endoreplication, one possibility is that Myc stimulates endocycling indirectly through its promotion of growth. In line with this idea, Maines *et al*. reported that *Myc* mutant follicles exhibited growth defects before the onset of endoreplication (Maines et al., 2004). However, in muscle cells Myc seems to be more important for endoreplication than for cell growth. In these cells, the overexpression of Myc results in increased DNA content and nuclear size but it only produces a slight increase in cell size, although it is possible that this small increase is sufficient to drive endoreplication (Demontis and Perrimon, 2009).

Different findings support the idea that the effects of Myc on cell growth and endocycling is mediated by its stimulation of protein synthesis. First, inhibition of protein synthesis with cycloheximide has been found to block endoreplication and larval growth (Britton and Edgar, 1998). Moreover, little growth or endoreplication is induced upon Myc overexpression when larvae are mutant for the translation factors eIf4a or eIf4e (Pierce et al., 2004). Specifically, Myc promotes protein synthesis by increasing ribosome biogenesis and protein translation, although these are energetically expensive processes that need to be tightly controlled (Grewal et al., 2005; Hulf et al., 2005). This control is achieved, in part, through regulation of Myc by the Insulin/Tor signaling pathway according to the nutritional status (Demontis and Perrimon, 2009; Parisi et al., 2011; Teleman et al., 2008). Under optimal conditions, the Insulin/Tor signaling pathway is activated and induces the accumulation of Myc, which leads to increased ribosome biogenesis and growth (Teleman et al., 2008). Conversely, starvation of amino acids in larvae is associated with reduced expression of genes involved in ribosome biogenesis and protein synthesis, and blocks cell growth and endoreplication (Britton and Edgar, 1998; Li et al., 2010; Zielke et al., 2011). Interestingly, starvation also causes reduced accumulation of E2f1 protein, whose overexpression is sufficient to rescue the block in endoreplication induced by starvation, but not the block in cell growth (Zielke et al., 2011). Therefore, Myc might regulate endoreplication (but not cell growth) in response to adequate

In summary, we have a basic understanding of the types of regulators and potential mechanisms that may play a role in signaling endocycle exit once a cell has reached its proper ploidy level. However, exactly how this is achieved and whether it involves a switch in the activities of Myc, the endocycle oscillator or other signaling pathways remains unknown.

OBJECTIVES

CIC has recently emerged as a key RTK signaling effector with important functions in development and human diseases. Studies in *Drosophila* and mammals have defined a conserved mechanism of CIC-mediated transcriptional regulation downstream of RTK signaling. In the absence of signaling, CIC represses its target genes by recognizing specific octameric through the HMG-box and C1 domains. Upon RTK activation, CIC is downregulated in response to phosphorylation by MAPK, and this enables the expression of its target genes. However, the approaches used in these studies did not allow to evaluate potential differences in the function or regulation of CIC-L and CIC-S isoforms. In this thesis we aimed to shed light on these potential differences by using *Drosophila melanogaster* as a model. Specifically, our objectives were the following:

1. Study the functional significance of Cic-S and Cic-L isoforms, comparing their expression and function during *Drosophila* development.

2. Characterize Cic-L-specific functions, focusing on its roles in *Drosophila* oogenesis and their implications for our understanding of this developmental process.

3. Investigate the molecular activity(ies) of Cic-L and its potential interactions with other factors.

RESULTS

1. Functional significance of Cic-S and Cic-L isoforms

1.1 Common functions of Cic isoforms

As mentioned above, in this thesis we have studied the functional significance and molecular activities of the two isoforms of Cic in *Drosophila*. We have focused particularly on the long isoform, Cic-L, because it had not been studied individually in any species before. To begin the characterization of Cic-L, we decided to exploit the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated nuclease 9 (Cas9) technology to tag the endogenous Cic-L isoform. In order to tag specifically this isoform, the tag had to be inserted in its unique N-terminal region. To this end, we identified a poorly conserved sequence between the Tudor-like and the N1 domains where the tag could be inserted without disrupting Cic-L function. Briefly, we designed a guide RNA (gRNA) that targets this sequence and provided a double-stranded DNA donor template that, upon homology-directed repair (HDR), introduced a triple hemagglutinin (HA) epitope at the selected site. (Fig. 12A; for more details see Materials and Methods). Using this approach, we recovered several fly lines that were viable and fertile, suggesting that the tag does not affect Cic-L activity (see the phenotype of *cic-L* mutants below).

Then, we combined this allele with a construct of Cic-S tagged with Venus (Grimm et al., 2012) to compare the expression of the two isoforms in three different tissues where Cic functions downstream of RTK-Ras-MAPK signaling: the early embryo, the wing imaginal disc and the ovarian follicle cells (Ajuria et al., 2011; Andreu et al., 2012b; Astigarraga et al., 2007; Atkey et al., 2006; Jiménez et al., 2000; Roch et al., 2002). In the early embryo, Cic activity is linked to the Torso pathway and we found that only the short isoform is present (Fig. 12B). As previously described, we observed that maternal Cic-S accumulated in the middle regions of the embryo, but it was degraded at the poles in response to Torso activation (Astigarraga et al., 2007; Jiménez et al., 2000). In contrast, in wing imaginal discs and in the follicular epithelium, where Cic functions in connection to EGFR signaling, both isoforms are co-expressed. During wing development, EGFR signaling specifies the formation of veins and induces the downregulation of Cic in presumptive vein cells (Ajuria et al., 2011; Roch et al., 2002). As shown in Figure 12C and F, both isoforms appear to be similarly downregulated by EGFR signaling in presumptive vein cells of wing discs from third instar (LIII) larvae. Instead, although the two isoforms are also co-expressed in the ovarian follicular epithelium, they seem to respond differently to EGFR signaling (Fig. 12D,

G). EGFR activation in dorsal-anterior (DA) follicle cells is known to cause partial relocalization of Cic to the cytoplasm (Astigarraga et al., 2007), but we found that this is only true for Cic-S because Cic-L remains mostly nuclear in these cells (Fig. 12G, G'). Thus, the two isoforms exhibit overlapping but also distinct tissue and subcellular distributions during *Drosophila* development.

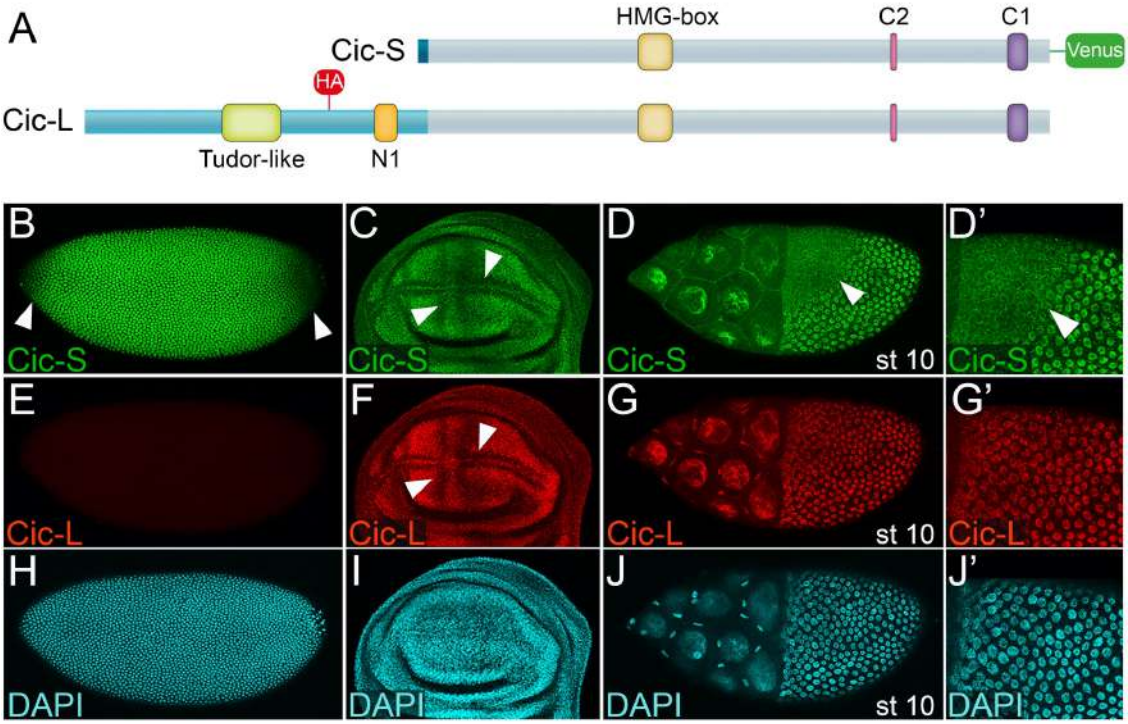

Figure 12. Expression patterns of Cic-S and Cic-L during *Drosophila* development. (A) Diagram of *Drosophila* Cic-S and Cic-L proteins tagged with Venus and HA, respectively. Cic-S-Venus is expressed from a transgene whereas Cic-L-HA has been generated via CRISPR/Cas9. (**B-J)** Immunostainings comparing the expression of Cic-S-Venus (B-D) and Cic-L-HA (E-G) in the early blastoderm embryo (B, E), the wing imaginal disc (C, F) and the ovarian follicular epithelium (D, G). White arrowheads mark regions of RTK-mediated downregulation. DAPI signals are shown in H-J. High magnifications of DA anterior follicle cells are provided in D'-J'. In this and subsequent figures, embryos and egg chambers are oriented with anterior to the left and dorsal up.

Next, we decided to investigate the functional significance of the Cic-L isoform by generating a new allele of *cic* that selectively inactivates this isoform. As illustrated in Figure 13A, the *cic-L* and *cic-S* transcripts originate from alternative promoters and this allowed us to use CRISPR/Cas9-mediated mutagenesis to disrupt Cic-L without affecting Cic-S. In this case, we selected a target sequence just upstream the N1 to generate small mutagenic insertions and deletions (indels) via nonhomologous end joining (NHEJ) DNA repair. One of the isolated mutations, which we refer to as cic^{7}, is a frameshift allele caused by deletion of 5 bp that truncates Cic-L at amino acid 658 of isoform Cic-PD in Flybase, thus removing

the highly conserved N1 domain and all other sequences up to the end of the protein (Fig. 13A). On the other hand, a null allele of *cic-S* was already available, cic^{5}, which is a CRISPR-induced frameshift lesion in the first exon of *cic-S* (Papagianni et al., 2018). Using these isoform-specific alleles, we studied the requirement of each isoform in the specification of wing veins and the establishment of DV polarity in follicle cells. For a list of the different *cic* alleles that we have used in this thesis, see section 3.4 of Materials and Methods.

In the wing primordium, Cic restricts vein formation to appropriate regions by repressing vein-specific genes in intervein cells (Ajuria et al., 2011; Roch et al., 2002). In prospective vein cells, instead, Cic is downregulated in response to EGFR signaling and this allows the expression of target genes such as *aos*, which encodes a feedback inhibitor of EGFR signaling [(Ajuria et al., 2011; Freeman et al., 1992; Golembo et al., 1996; Roch et al., 2002) and Fig. 13B]. Accordingly, reduced Cic function causes ectopic expression of *aos* and other genes such as *ventral veinless* and *decapentaplegic*, and the formation of extra vein tissue [(Ajuria et al., 2011; Goff et al., 2001; Roch et al., 2002) and Fig. 13B]. However, the requirement of each isoform during this process was not known because the initial experiments were performed using alleles that affected both isoforms. So, we analyzed the presence of extra vein tissue in adult wings corresponding to our isoform-specific alleles. Because the cic^{7} mutation is homozygous lethal –which indicates that *cic-L* is essential in *Drosophila*–, we decided to use heteroallelic combinations of each isoform-specific allele with the cic^{4} allele, which consists of a 4 amino acid deletion in the common C1 domain and behaves as a strong hypomorph (Forés et al., 2017b). Interestingly, cic^{7}/cic^{4} flies only exhibited small ectopic veins close to L2 whereas cic^{5}/cic^{4} flies did not show any extra vein tissue (Fig. 13D, E).

One interpretation of these results is that the two isoforms might function redundantly to repress vein-specific genes in intervein cells. Thus, when only one isoform is eliminated, the persisting isoform is sufficient to maintain the correct specification of wing veins. To test this hypothesis, we eliminated both isoforms simultaneously by generating a frameshift mutation similar to cic^{5} in a chromosome carrying cic^{7} (we have named this compound allele cic^{8}). Indeed, inactivation of the two isoforms resulted in abnormal wings with extra vein tissue (Fig. 13F). In addition, we could significantly rescue the wing phenotype of *cic* deficient flies with a Cic-L expression construct (compare Fig. 13G with 13H) (see Fig. 19A for a diagram of the Cic-L expression construct). From these experiments, we conclude that Cic-L and Cic-S act redundantly during the specification of wing veins.

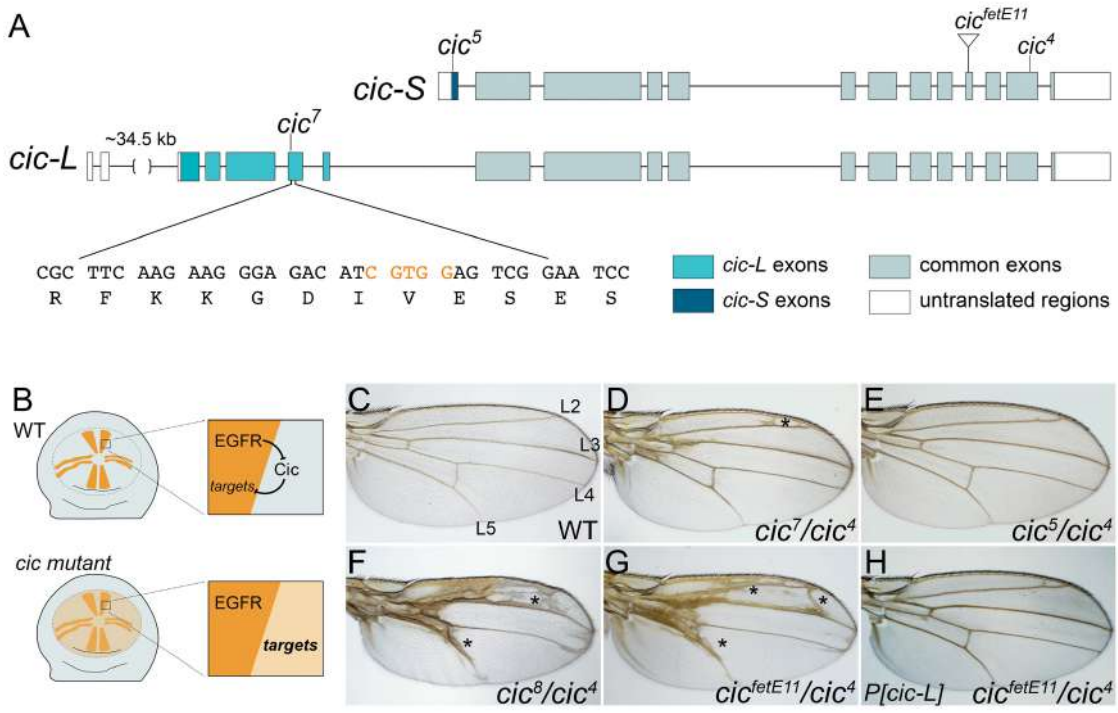

Figure 13. Cic-L and Cic-S function redundantly in the specification of wing veins. (A) Schematic representation of the *cic* locus in *Drosophila* indicating the position of different *cic* alleles: cic^{5} and cic^{7} are CRISPR/Cas9-induced frameshift mutations that selectively disrupt Cic-S and Cic-L, respectively; cic^{fetE11} is a P-element insertion; and cic^{4} is a 4 amino acid deletion in the C1 domain generated by CRISPR/Cas9. **(B)** Model of Cic function in the wing imaginal disc. **(C-H)** Wings from adult flies of the indicated genotypes. Asterisks indicate ectopic vein material.

Next, we tested if Cic-L and Cic-S also function redundantly in the ovarian follicular epithelium. In this context, Cic is required for the establishment of embryonic DV polarity (Andreu et al., 2012b; Astigarraga et al., 2007; Atkey et al., 2006; Goff et al., 2001). In *Drosophila*, embryonic DV patterning relies on positional information generated in the egg chamber that is later transmitted to the fertilized embryo. A key factor in the transmission of this information is the sulfotransferase Pipe, whose expression is restricted to ventral follicle cells and defines the ventral region of the future embryo (Sen et al., 1998). This ventral expression of Pipe requires Cic function. In this context, Cic acts indirectly by repressing *mirr*, which encodes a homeodomain transcription factor that represses *pipe* (Andreu et al., 2012b). *mirr* is normally induced in DA follicle cells by EGFR signaling, and this leads to the repression of *pipe* in dorsal and lateral follicle cells. In addition, downregulation of Cic by EGFR signaling in DA follicle cells has been proposed to reinforce *mirr* induction through derepression (Andreu et al., 2012a).

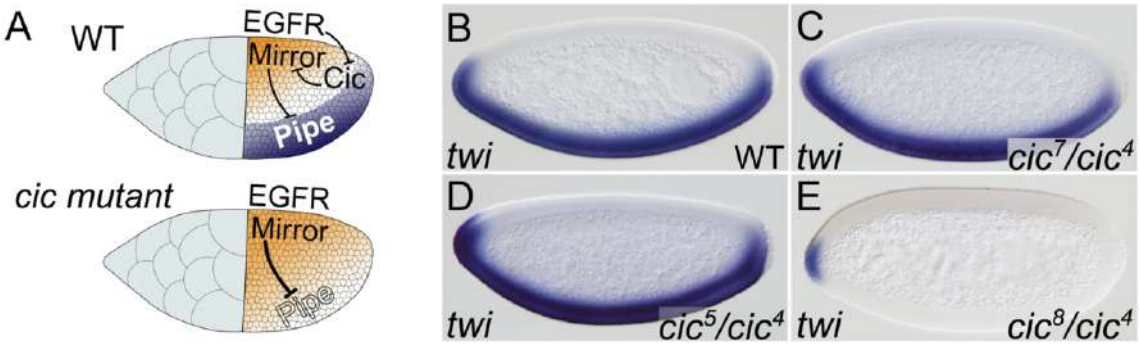

Figure 14. Cic-L and Cic-S function redundantly during the establishment of embryonic DV polarity in follicle cells. (A) Model of Cic function in follicle cells**. (B-D)** mRNA expression pattern of *twi* in blastoderm embryos of the indicated maternal genotype.

In *cic* mutant females, depression of *mirr* causes the loss of *pipe* expression in ventral follicle cells, resulting in complete dorsalization of the embryo (Fig. 14A). So, we asked if this circuit is disrupted in our mutants. We did this by analyzing the expression of the embryonic ventral marker *twist* (*twi*), whose expression is lost in embryos derived from *cic* mutant females (Goff et al., 2001). Similar to what we observed in the wing primordium, the pattern of *twi* expression was not affected upon individual inactivation of Cic-S or Cic-L (Fig. 14C, D). In contrast, the loss of both isoforms led to complete dorsalization of embryos and absence of *twi* expression (Fig. 14E). Thus, the two isoforms seem to be interchangeable also in follicle cells.

The above experiments raised, however, and additional question. If Cic-L contributes to embryonic DV patterning, why does it remain nuclear in DA follicle cells? This observation might be explained by the recent finding that MAPK signaling in the embryo, besides inducing the degradation of Cic-S, also triggers the dissociation of Cic from the regulatory regions of its target genes (Keenan et al., 2020). So, we hypothesized that the nuclear pool of Cic-L in DA follicle cells might be functionally inactivated by EGFR signaling. To test this hypothesis, we compared the repressor activities of Cic-L and Cic-L$^{\Delta C2}$, a Cic-L mutant derivative insensitive to EGFR-mediated downregulation. Specifically, we adopted an assay developed by Andreu *et al.* (Andreu et al., 2012a) where they analyzed the expression of the *mirr^{F7}-lacZ* enhancer trap (McNeill et al., 1997) in follicle cell clones overexpressing Cic-S or Cic-S$^{\Delta C2}$. In our case, we expressed *UAS-Cic-L-HA* and *UAS-Cic-L$^{\Delta C2}$-HA* transgenes in clones using the Mosaic Analysis with a Repressible Cell Marker (MARCM) technique (see Materials and Methods). As shown in Figure 15, although *UAS-Cic-L* and *UAS-Cic-L$^{\Delta C2}$* were both expressed at similar levels, only Cic-L$^{\Delta C2}$ repressed expression of *mirr^{F7}* in DA

follicle cells. This result suggests that despite the nuclear localization of Cic-L, functional inactivation by EGFR signaling prevents it from repressing *mirr* in DA follicle cells.

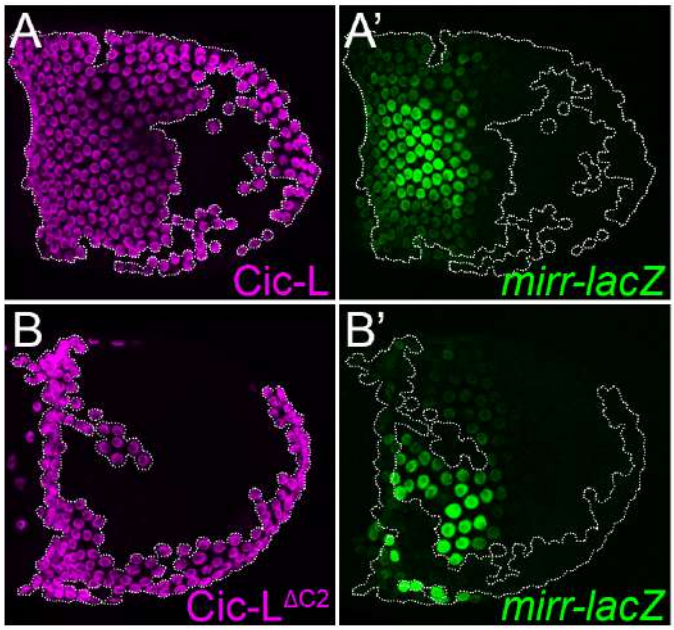

Figure 15. Cic-L is functionally downregulated by EGFR signaling in DA follicle cells. (A-B) Stage 10 mosaic egg chambers carrying the *mirr*^{F7}*-lacZ* enhancer trap and expressing Cic-L-HA (A) or Cic-L$^{\Delta C2}$ -HA (B) proteins in clones. They are stained with anti-HA (A, B) and anti-β-galactosidase (A′, B′).

1.2 Unique functions of Cic isoforms

When we compared the expression of Cic-S and Cic-L in egg chambers, we noticed that they exhibit differential expression patterns in the germline: while Cic-S begins to accumulate in nurse cell nuclei at early stages, Cic-L is not detected in these cells until stage 8 (compare Fig. 16A with 16A′). In addition, we observed Cic-L-specific expression in the germarium, border cells and stretch cells (Fig. 16A′).

With these observations in mind, we set out to investigate potential additional functions of Cic in oogenesis. As a first approach, we analyzed the general morphology of *cic-S and cic*-L mutant egg chambers stained with DAPI and phalloidin, which label DNA and actin, respectively. We dissected ovaries from *cic*^{5} homozygous females, but no major morphological defects were observed (Fig. 17B). Regarding Cic-L, as the *cic*^{7} allele is homozygously lethal, we used the FLP-DFS technique to generate a *cic*^{7} homozygous mutant germline. Notably, females carrying *cic*^{7} germline clones (GLCs) were fully sterile, suggesting a specific requirement of Cic-L in the germline. When we examined the ovaries,

we found that nurse cells persisted attached to the oocyte by the end of oogenesis (Fig. 17C). Moreover, the presence of cytoplasm in these persisting cells suggests that *cic*^{7} nurse cells fail to transfer, or dump, their cytoplasmic contents into the oocyte. This phenotype is generally called *dumpless* and results in the production of small fragile eggs. Also, as the ellipsoid shape of wildtype eggs plays an important role in their transit through the oviduct, alteration of the egg shape by the persisting nurse cells leads to reduced egg laying.

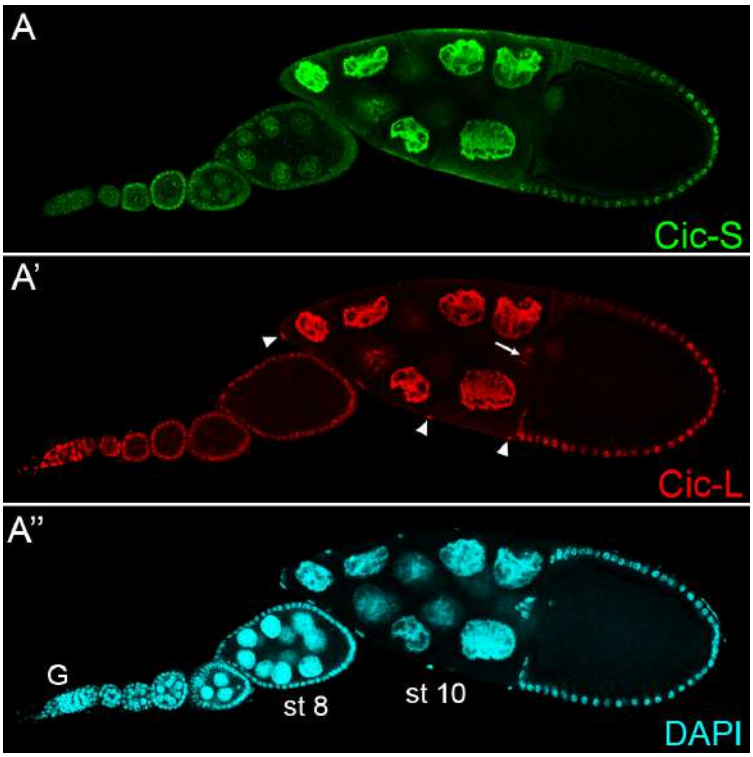

Figure 16. Expression pattern of Cic-S and Cic-L during oogenesis. (A) *cic-S-Venus/cic-L-HA* ovariole stained with anti-GFP (A), anti-HA (A') and DAPI (A''). The arrow indicates Cic-L expression in border cells whereas arrowheads indicate expression in the stretch cells.

In addition, with the idea of uncovering possible maternal effects, we studied the cuticular structures of embryos derived from these females. As expected from the expression pattern in the embryo, only embryos deposited by *cic*^{5} females exhibited the Capicua phenotype (Fig. 17E). On the other hand, the very few eggs laid by females carrying *cic*^{7} GLCs were short, fragile and displayed abnormal dorsal appendage morphology (not shown). Moreover, these embryos do not develop cuticular structures (not shown), which we reasoned could be due the previous dumping defects. So, we decided to use a slightly weaker allele of *cic-L* that we have also generated by CRISPR/Cas9, *cic-L*$^{\Delta T \Delta N}$ (see Results 3.1 for more details about this allele). Embryos laid by *cic-L*$^{\Delta T \Delta N}$ homozygous females are very fragile, and the vitelline membrane is frequently broken at the anterior pole (Fig. 17F).

Most of them do not develop cuticle either, but a small fraction exhibits bicaudal phenotypes (Fig. 17F). The bicaudal phenotype is characterized by the absence of head, thoracic and anterior abdominal segments, which are replaced by a mirror-image duplication of posterior abdominal segments and telson. In the case of the *cic-L*$^{\Delta T \Delta N}$ allele, the bicaudal phenotype is particularly strong because, in general, only one or two segments develop at each end.

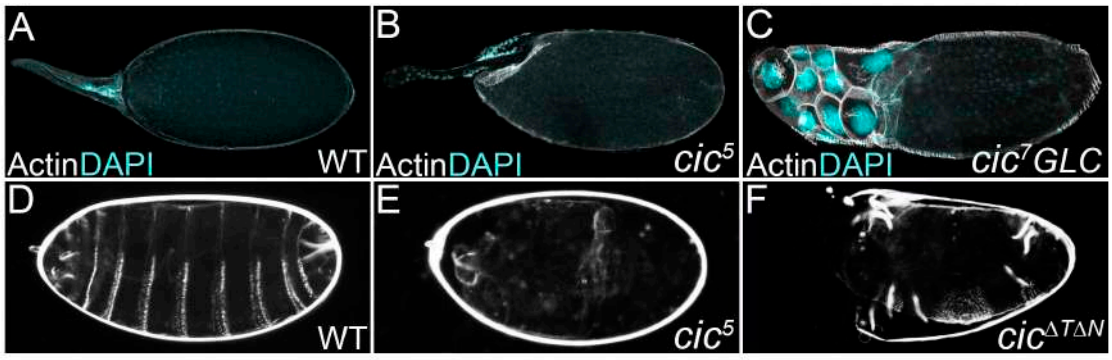

Figure 17. Specific requirement of Cic-L for nurse cell dumping and embryonic AP patterning. (A-C) Staining with DAPI and phalloidin of wildtype (A), *cic*^{5} and *cic*^{7} eggs. (D-F) Cuticle of embryos derived from wildtype (D), *cic*^{5} (E), and *cic-L*$^{\Delta T \Delta N}$ females.

Therefore, these observations indicate that beyond the common activities described above, each isoform also has specific, essential functions. In particular, Cic-S is required during early embryogenesis where, among others, represses zygotic genes of the terminal and DV systems (Ajuria et al., 2011; Jiménez et al., 2000; Papagianni et al., 2018). Instead, Cic-L appears to be necessary for viability, nurse cell dumping, dorsal appendage morphogenesis and establishment of embryonic AP polarity.

Interestingly, the co-called *bullwinkle* (*bwk*) mutations described by the group of Dr. C. Berg 25 years ago is also associated with dumping defects and bicaudal phenotypes (Rittenhouse and Berg, 1995). In fact, one of these mutations is caused by the insertion of a P element, *P*(*PZ*) *bwk*^{8482}, just 300 away from the hobo transposon insertion (*cic*^{1}) that led to the identification of *cic* (Jiménez et al., 2000). However, the initial analysis of the genetic relationship between *bwk*^{8482} and *cic*^{1} revealed that the two mutations complemented each other (Jiménez et al., 2000). This finding suggested that *bwk*^{8482} and *cic*^{1} affected different genetic functions in the same locus, but the molecular relationship between *bwk* and *cic* could not be defined because *bwk* has not been molecularly characterized and thus remains unannotated (see FlyBase). We reasoned that the complementation could be

explained if each mutation disrupts a different isoform of *cic*. Indeed, *cic*^{1} maps to the 5′ untranslated region of Cic-S, whereas *bwk*^{8482} maps to an intron of Cic-L (Flybase, Fig. 18A). Moreover, we tested if *bwk*^{8482} is allelic to *cic*^{7} by dissecting ovaries from *bwk*^{8482}/+, *cic*^{7}/+, and *bwk*^{8482}/*cic*^{7} females and staining them with phalloidin. As expected, only occasional slightly dumpless eggs were observed in ovaries from *bwk*^{8482} or *cic*^{7} heterozygous females (Fig. 18B, C). In contrast, *bwk*^{8482}/*cic*^{7} transheterozygous females were completely sterile and produced a vast majority of dumpless eggs (Fig. 18D). Therefore, these results show that Bwk is, in fact, the long isoform of Cic, clarifying a long-standing question regarding the nature of the *bwk* gene.

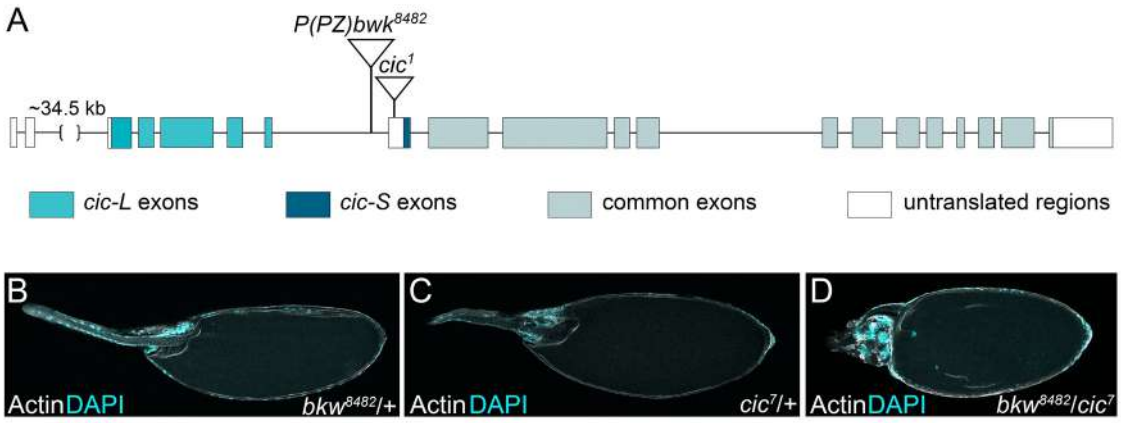

Figure 18. *bwk* mutations target *cic-L*. (A) Schematic representation of the *cic* locus indicating the position of the *P(PZ) bwk*^{8428} (P-element) and *cic*^{1} (hobo) transposon insertions. Note that *bwk*^{8482} targets *cic-L* whereas *cic*^{1} targets *cic-S*. **(B-D)** Eggs of the indicated genotypes stained with DAPI and phalloidin.

In addition, we generated a rescue construct to verify that all the defects associated with Cic-L/Bwk are indeed caused by inactivation of this protein. This construct includes the coding exons, 5′ and 3′ UTR sequences of *cic-L* as well as a 2.4kb upstream fragment as the putative *cic-L* promoter (see Materials and Methods for more details). Although due to technical issues we were not able to include the 34.5kb intron of *cic-L*, the construct was tagged with a triple HA at the C-terminus and we confirmed that it recapitulates the expression in the ovary of the *cic-L* allele tagged via CRISPR/Cas9 (Fig. 19B). We recombined this transgene with the *cic*^{7} mutation and found that it significantly rescues *cic*^{7}-associated lethality as well as the dumpless and bicaudal phenotypes (Fig. 19C, D).

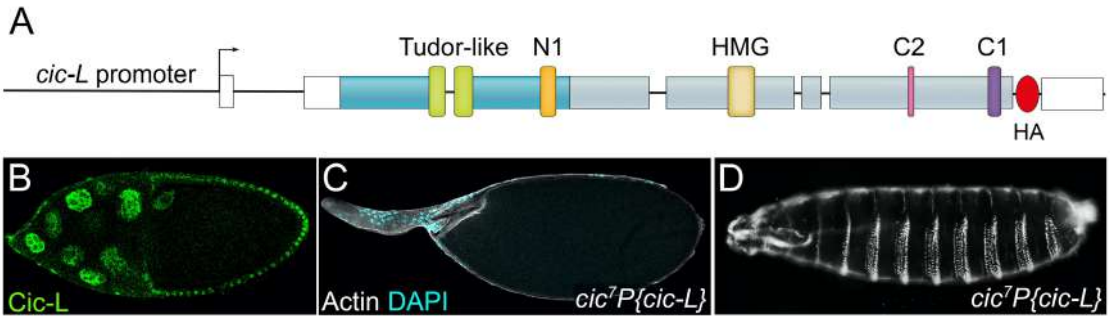

Figure 19. Rescue of *cic-L* phenotypes. (A) Diagram of the *cic-L* rescue construct tagged with a HA**.** The construct includes several introns (black lines), and the 5′ and 3′UTR of *cic-L* (white boxes). **(B)** *cic-L-HA* stage 10 egg chamber stained with anti-HA**. (C)** Double staining with DAPI and phalloidin of a cic^{7} egg rescued by expression of Cic-L. **(D)** Cuticle phenotype of an embryo derived from a cic^{7} female rescued by Cic-L expression.

2. Unique functions of Cic-L in oogenesis

Having established that Cic-L has redundant as well as unique activities, we reasoned that these two types of activities might involve different molecular mechanisms (see Discussion). In redundant functions, Cic-L probably acts as a transcriptional repressor that binds TGAATGAA-like sites in target genes. On the other hand, Cic-L-specific functions may or may not involve similar DNA binding but are likely to depend on other unknown mechanisms (transcriptional or otherwise) mediated by the N-terminal region of the protein. Therefore, we set out to investigate the functions of Cic-L during oogenesis with the idea of identifying specific phenotypes that would provide clues about the underlying molecular mechanisms. As described below, our results show that Cic-L plays a particularly important role in nurse cell dumping, and this phenotype has guided most of our analyses. However, as we have also found that hypomorphic alleles of *cic-L*, which have less severe effects on oogenesis can lead to the development of bicaudal embryos, I present first a set of preliminary analyses to investigate this phenotype.

2.1 Role of Cic-L in the establishment of embryonic AP polarity

In the *Drosophila* embryo, determination of the future abdominal region relies on a posterior-to-anterior gradient of Nos protein, which represses the translation of maternal *hunchback* mRNA (Ephrussi and Lehmann, 1992; Gavis and Lehmann, 1992; Irish et al., 1989; Smith et al., 1992). This gradient of Nos is established through mechanisms that

regulate the localization and translation of *nos* mRNA specifically in the posterior region of the embryo (Kugler and Lasko, 2009). In certain mutants, these processes are affected and cause the accumulation of ectopic Nos activity (either at the anterior pole or throughout the embryo), which suppresses the development of anterior regions and can lead to the production of bicaudal embryos with two abdomens (Ephrussi and Lehmann, 1992; Gavis and Lehmann, 1992; Smith et al., 1992).

nos is synthesized in the nurse cells and is transferred to the oocyte during late oogenesis via nurse cell dumping. Once in the oocyte, *nos* RNA particles diffuse throughout the oocyte with the help of cytoplasmic streaming, and then become trapped at the posterior through association with the germ plasm (Forrest and Gavis, 2003). Importantly, germ plasm assembly, and thus *nos* localization, depends on Osk, which accumulates at the posterior pole (Ephrussi et al., 1991). In fact, mislocalization of *osk* transgenic transcripts to the anterior pole is sufficient to drive anterior accumulation of Nos and development of bicaudal embryos (Ephrussi and Lehmann, 1992). Overexpression of *osk* also causes the development of bicaudal embryos (Ephrussi and Lehmann, 1992; Smith et al., 1992). In this situation, *osk* is dispersed throughout the embryo and this results in ectopic Nos in anterior and central regions of the embryo (Smith et al., 1992). Therefore, we decided to investigate if the *cic-L* bicaudal phenotype is caused by alterations in the regulation of *osk*, which involves three main levels: transport of *osk* mRNA from the nurse cells to the posterior pole of the oocyte, anchoring of the transcript at the posterior pole and translational repression of unlocalized *osk* (Kugler and Lasko, 2009).

To study if Cic-L is involved in the localization of *osk,* we analyzed the distribution of *osk* mRNA in *cic-L*$^{\Delta T \Delta N}$ egg chambers and progeny embryos by in situ hybridization. In wildtype oocytes, *osk* begins to accumulate at the posterior pole during mid-oogenesis via kinesin-mediated transport coupled to a slight bias in the orientation of microtubule plus ends toward posterior [(Ephrussi et al., 1991; Zimyanin et al., 2008), Fig. 20A]. Although in late oogenesis this active transport of *osk* ceases when the microtubules are rearranged, a second mechanism continues to accumulate *osk* at the posterior pole. This second mechanism is less well understood but relies on cytoplasmic streaming and anchorage at the posterior pole (Sinsimer et al., 2011). When we looked at *cic-L* egg-chambers, the initial accumulation of *osk* at the posterior pole appeared to be normal (Fig. 20C). Unfortunately, we could not use in situ hybridizations to analyze the distribution of *osk* in the oocyte during late oogenesis because the presence of the vitelline membrane and the eggshell limits the access of RNA probes. Nevertheless, at those late stages, we did observe

differences in *osk* mRNA accumulation in the nurse cells. Whereas wild-type nurse cells that have completed dumping display very low levels of *osk* mRNA (Fig. 20B), the abnormally persistent nurse cells in *cic-L* egg chambers show significant accumulation of *osk* mRNA (Fig. 20D). Based on this observation, we reasoned that persistent nurse cells might represent an inappropriate source of *osk* mRNA. In this scenario, leakage of *osk* from persisting nurse cells to the oocyte (at a time when the transport of *osk* to the posterior has already ceased), could result in ectopic accumulation of *osk* at the anterior pole of the oocyte.

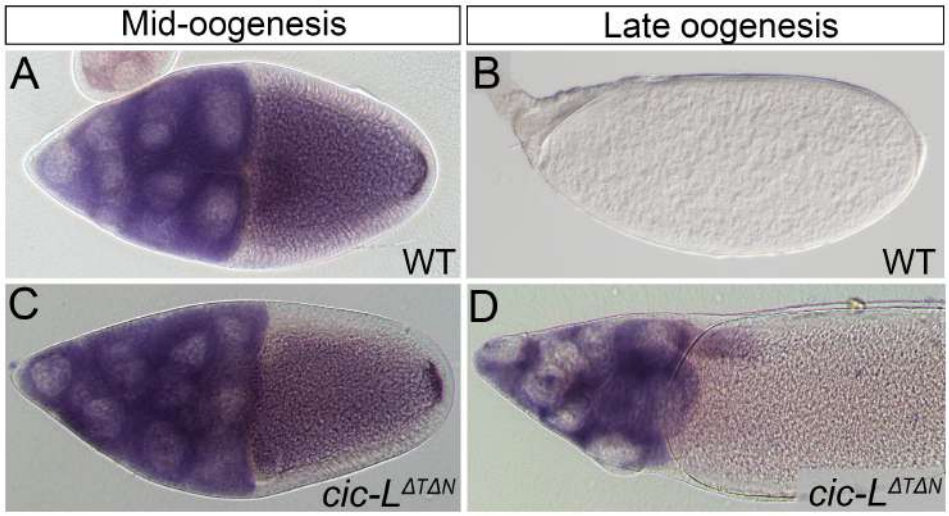

Figure 20. Persisting nurse cells in *cic-L* egg chambers accumulate *osk* transcripts. (A-D) mRNA expression pattern of *osk* in wildtype (A,B) and *cic-L*$^{\Delta T\Delta N}$ (C,D) egg chambers.

To investigate if *osk* accumulates at the anterior pole or other regions of the oocyte during late stages in *cic-L* mutants, we took advantage of an *osk* reporter construct generated by the group of A. Ephrussi [*M1M2-LacZ-osk3'UTR*; (Gunkel et al., 1998)]. This construct contains the *lacZ* reporter gene fused to the 5' and 3'UTR regulatory sequences of *osk* (Fig. 21A) and recapitulates the distribution of endogenous *osk* during mid-oogenesis, both at the RNA and protein levels. First, we used X-Gal staining to examine the pattern of β-galactosidase activity resulting from this reporter during late oogenesis in otherwise wildtype ovaries. β-galactosidase staining was mostly concentrated at the posterior pole of the egg, but weaker staining was also observed emanating from the posterior pole (Fig. 21B). This is probably due to diffusion of the β-galactosidase enzyme due to lack of anchorage. In *cic-L*$^{\Delta T\Delta N}$ mutants, instead, the levels of β-galactosidase staining were high at both poles of the egg and also clearly detectable in the central regions (Fig. 21C).

Interestingly, the distribution of β-galactosidase staining at each pole was different: whereas it appeared tightly concentrated at the posterior, it was more broadly distributed at the anterior and formed stripes that extended toward the center of the egg. We also analyzed the distribution of *osk* mRNA in early embryos and obtained similar results: in wildtype embryos *osk* was only detected at the posterior pole (Fig. 21D), but embryos derived from *cic-L*$^{\Delta T\Delta N}$ females also accumulated *osk* at the anterior pole in a diffuse pattern (Fig. 21E). Therefore, the *cic-L* bicaudal phenotype is most probably caused by mislocalization of *osk* mRNA during late oogenesis. Presumably, the presence of *osk* at the anterior pole results in ectopic accumulation of Nos, which directs the formation of the second abdomen. Finally, the mislocalization of *osk* could represent an indirect consequence of *cic-L* cellular defects during oogenesis, although we cannot rule out a more direct role of Cic-L in *osk* localization during late oogenesis.

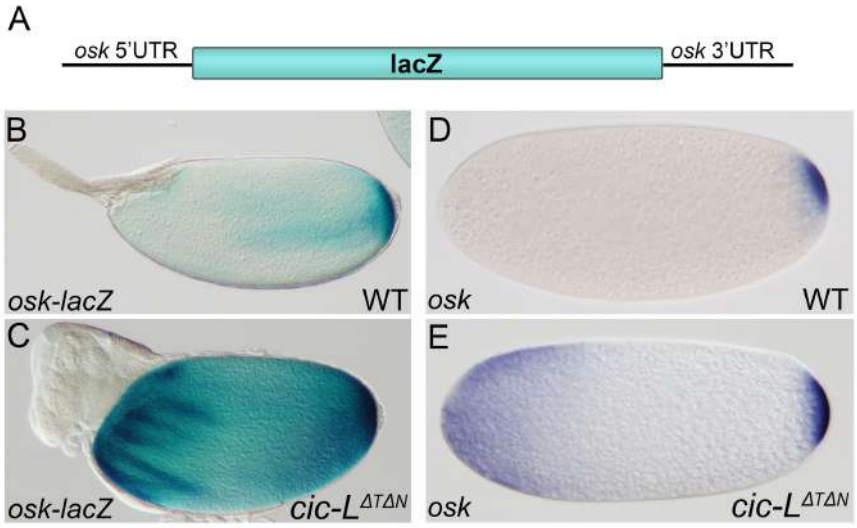

Figure 21. Ectopic accumulation of *osk* during late oogenesis. (A) Diagram of the *M1M2-LacZ-osk'UTR* reporter containing the regulatory regions of *osk*. **(B-C)** β-galactosidase staining of the *M1M2-LacZ-osk'UTR* reporter in late wildtype (B) and *cic-L*$^{\Delta T\Delta N}$ (C) egg chambers. **(D, E)** mRNA expression pattern of *osk* in embryos of the indicated maternal genotype.

2.2 Role of Cic-L in nurse cell dumping

In *cic-L* mutant egg chambers overfull nurse cells remain attached to the oocyte by the end of oogenesis and, since this is the earliest apparent defect in oogenesis, we decided to study it more thoroughly. Normally, at stage 11 nurse cells contract and transfer their cytoplasmic content through the ring canals into to the growing oocyte (Mahajan-Miklos and Cooley,

1994b). In contrast to nurse cell-to-oocyte transport during early phases, dumping is a rapid transport that results in doubling of the oocyte volume (and regression of the nurse cell cluster) in less than 30 minutes (Cooley et al., 1992). Thus, the presence of small eggs and persistent nurse cells suggests that nurse cell dumping is disrupted in *cic-L* mutants. So far, three main origins for the dumpless phenotype have been identified (Hudson and Cooley, 2002). First, incomplete nurse cell dumping can be due to lack of nurse cell contraction, which relies on nonmuscle myosin II. For example, mutations in *spaghetti squash*, which encodes the regulatory light chain of myosin II, have been described to cause dumping defects (Edwards and Kiehart, 1996; Wheatley et al., 1995). Also, the organization and size of the ring canals is key to allow the flow from the nurse cells to the oocyte. As revealed by the loss of function of the F-actin crosslinking proteins Cheerio and Kelch, ring canals with reduced lumens are unable to accommodate normal cytoplasm flow, giving rise to dumpless eggs (Robinson et al., 1994; Robinson et al., 1997). Finally, in other mutants the origin of the dumping phenotype is the obstruction of the ring canals by the large nurse cell nuclei (Bass et al., 2007; Cant et al., 1994; Clough et al., 2014; Cooley et al., 1992; Mahajan-Miklos and Cooley, 1994a; Volpe et al., 2001). In wildtype follicles, this is avoided by actin filament bundles that anchor the nurse cell nuclei and prevent them from being pushed into the ring canals as the nurse cell cytoplasm flows into the oocyte. These actin filament bundles are assembled during stage 10B and create a halo of filamentous-actin (F-actin) around each nurse cell nuclei. Several studies have shown that multiple actin-binding proteins, including Chickadee, Singed and Quail are required to assemble the actin network, but the regulatory mechanism that initiates this major rearrangement of the actin cytoskeleton and nurse cell dumping is still unknown.

In order to investigate if the dumping defects observed in *cic-L* egg chambers were due to any of these three causes, we initially performed a rhodamine-phalloidin staining to visualize F-actin. As previously described, phalloidin labeled the ring canals and the actin filament bundles in wildtype egg chambers. We detected the first actin bundles in stage 10B egg chambers (Fig. 22A'), when the nurse cell cluster and the oocyte have approximately the same volume. However, dumping defects make the morphological identification of stage 10B follicles more complicated because it affects nurse cell regression and oocyte growth. To overcome this challenge and make the staging of follicles more precise, we used a *torso-like-lacZ* (*tsl-lacZ*) reporter that is expressed in border and centripetal follicle cells, which migrate at specific stages (Furriols et al., 2007). We analyzed the timing of centripetal migration and actin cytoskeleton rearrangement in

wildtype egg chambers and observed that actin filament bundles begin to form when the migrating centripetal cells reach the center of the nurse cell-oocyte boundary (Fig. 22A). In contrast, at this same stage, we did not observe actin bundles in *cic*^{7} GLC egg chambers (Fig. 22B'). The network of actin bundles was not assembled later in oogenesis either as only occasional isolated bundles could be detected (see Fig. 17C). The phalloidin staining also revealed that F-actin accumulates in the ring canals of *cic-L* egg chambers but ruling out a morphological defect would require closer inspection. Therefore, these observations indicate that Cic-L is required for the assembly of the actin bundles in nurse cells.

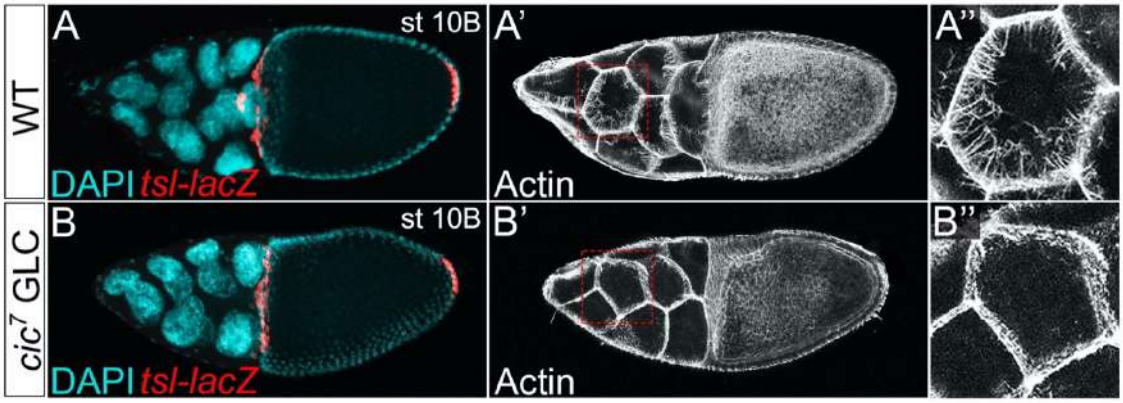

Figure 22. Cic-L is required for nurse cell actin bundles assembly. (A, B) Stage 10B egg chambers of the indicated genotype stained with DAPI and anti-β-galactosidase (A, B) and phalloidin (A', B').

Among the proteins required for the rearrangement of the actin cytoskeleton in the nurse cells, some of them are actin-binding proteins that play a direct role in assembling the actin bundles (Mahajan-Miklos and Cooley, 1994b). However, there are also several nuclear factors required for the assembly of the actin filament bundles whose specific role in the rearrangement of the actin cytoskeleton and nurse cell dumping is unknown (Bass et al., 2007; Clough et al., 2014; Myster et al., 2000; Royzman et al., 2002; Volpe et al., 2001). Since Cic-L is also a nuclear protein, we reasoned that its role in nurse cell dumping is probably indirect. In fact, we hypothesized that Cic-L could function as a signal to initiate nurse cell dumping because it becomes expressed in the nurse cells at stages 8-9. Nevertheless, we also noticed that some nurse cell nuclei in *cic-L* egg chambers appeared to be larger than normal. This observation led us to consider the possibility that Cic-L actually acts at a pre-dumping step to regulate the late stages of nurse cell growth. To explore this idea, we first investigated the relation between nurse cell growth and dumping.

As explained in the introduction, the growth of the nurse cells is based on endoreplication cycles driven by oscillations of CycE/Cdk2 activity. In the nurse cells in particular, the activity of the CycE/Cdk2 complex is principally regulated by oscillations of CycE protein, which accumulates at high levels at the entry of each S phase and then rapidly disappears for the rest of the cycle (Lilly and Spradling, 1996). Thus, we studied the dynamics of the endocycle by analyzing the accumulation of CycE in nurse cell nuclei at different stages. As it had been previously described, we observed that at early stages nurse cells endocycle asynchronously and only a few nurse cells per follicle exhibit high levels of CycE [(Lilly and Spradling, 1996), Fig. 23A']. In contrast, we found that CycE was present uniformly in all nurse cell nuclei at late stages (Fig. 23B',C'). Since high CycE/Cdk2 activity inhibits pre-RC assembly and constitutive CycE accumulation causes endocycle arrest (Follette et al., 1998; Weiss et al., 1998), we inferred that this stabilization of CycE leads to nurse cell endocycle exit. Using the *tsl-lacZ* reporter, we have determined that CycE becomes stabilized at stage 10A, just prior to rearrangement of the actin cytoskeleton at stage 10B (compare Fig. 23B and C). Therefore, these observations suggest that CycE is subjected to a mechanism of stabilization that could lead to nurse cell endoreplication exit at stage 10A before the onset of nurse cell dumping.

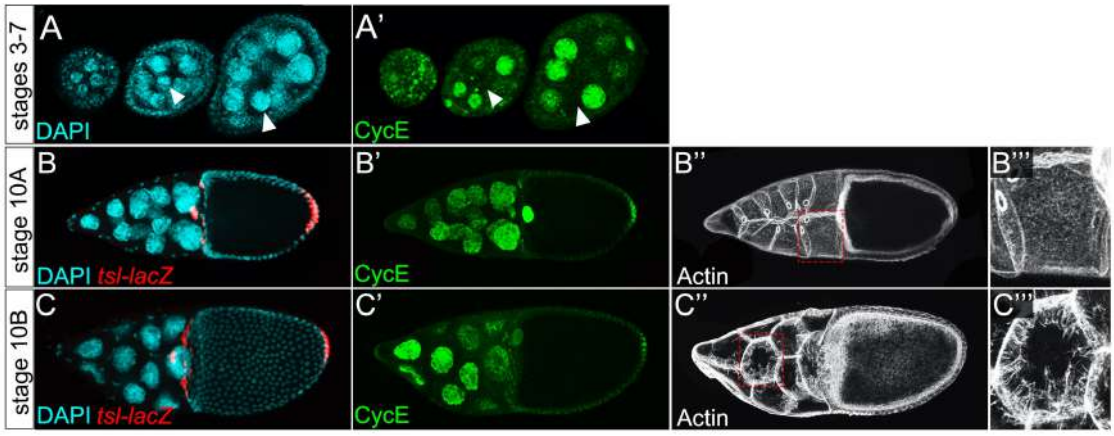

Figure 23. Stabilization of CycE at stage 10A. (A-C) Stage 3-7 (A), 10A (B) and 10B (C) egg chambers carrying the *tsl-lacZ* reporter. Stainings with DAPI (A); DAPI and anti-β-galactosidase (B, C); anti-CycE (A'-C'); and phalloidin (B'', C''). Arrowheads in A and A' indicate nurse cell nuclei that do not exhibit CycE accumulation. (B''', C''') Higher magnification of the indicated areas in B'' and C''. Note that the galactosidase staining (red) is also detected in the green channel in the polar cells due to a very strong signal.

2.3 Role of Cic-L in endocycle exit

The temporal relation between nurse cell growth and dumping led us to hypothesize that nurse cell endocycle exit might be necessary for dumping to occur, and that Cic-L could in fact regulate such exit. Therefore, we asked if CycE stabilization at stage 10A was compromised in *cic-L* ovaries. Indeed, we found that CycE was not synchronously stabilized in *cic*^{7} GLC egg chambers, which by stage 10B still exhibited individual nuclei with either high or low levels of CycE (Fig. 24B'). These results suggest that Cic-L is required for the timely exit of nurse cells from endoreplication and, directly or indirectly, for their subsequent progression into dumping.

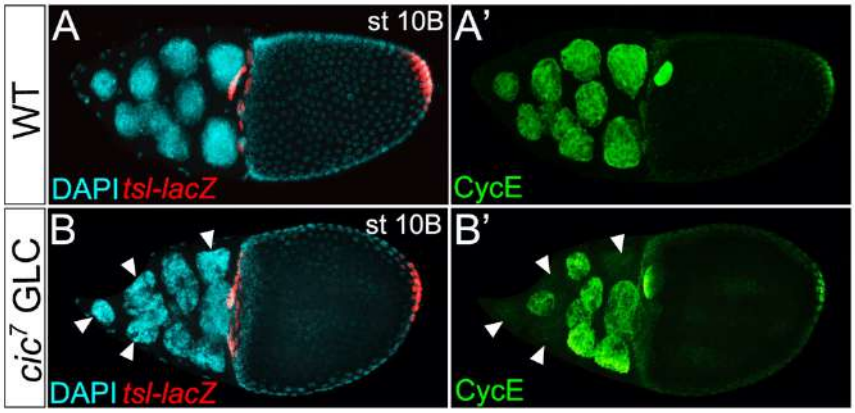

Figure 24. CycE is not stabilized in *cic-L* egg chambers. (A,B) Stage 10B egg chambers of the indicated genotypes stained with DAPI and anti-β-galactosidase (A,B) and anti-CycE (A', B'). Arrowheads in B and B'' indicate nurse cell nuclei that do not present CycE accumulation. Note that the galactosidase staining (red) is also detected in the green channel in the polar cells due to a very strong signal.

In addition, considering that Cic-L is specifically induced during the final stages of nurse cell growth (see Fig. 16A'), we wondered if Cic-L is not only required but sufficient to induce nurse cell growth arrest. To test this idea, we took advantage of the Gal4/UAS system to force the early expression of Cic-L in the nurse cells. Indeed, induction of Cic-L with the *mat-tubulin-Gal4* driver, which directs expression in the germline from stage 3, blocked nurse cell growth at this stage, producing aberrant ovarioles with serially arrested egg chambers (Fig. 25A). Moreover, staining of these egg chambers revealed ubiquitous accumulation of CycE in all nuclei (Fig. 25B'), instead of the characteristic asynchronous fluctuations observed during normal endocycle progression. Thus, premature Cic-L expression is sufficient to induce CycE stabilization and endocycle arrest, suggesting that it

normally functions as a signal that triggers endocycle exit once the nurse cells have completed their growth phase.

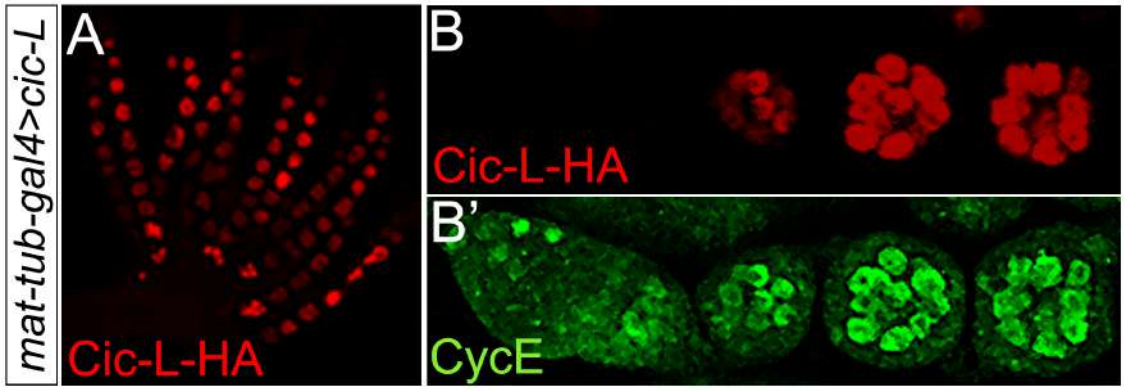

Figure 25. Early expression of Cic-L induces premature endocycle arrest. (A) Ovary prematurely expressing Cic-L-HA under the control of the *maternal-tubulin-Gal4* driver stained with anti-HA. Note that ovarioles consist of serially arrested egg chambers. **(B, B')** Double staining with anti-HA (B) and anti-CycE (B') of the germarium and three consecutive egg chambers.

Next, to begin exploring how Cic-L signals endocycle termination, we turned our attention to the Myc transcription factor, a well-established regulator of post-mitotic growth and endocycling. In particular, we chose Myc as a likely candidate to mediate the effects of Cic-L on endocycle exit based on two principal findings. First, ectopic expression of Myc had been previously found to extend the period of endoreplication in salivary glands and the fat body (Pierce et al., 2004). Second, loss of Myc function in the nurse cells results in reduced endoreplication and impaired cell growth, which resembles the phenotype caused by premature expression of Cic-L (Maines et al., 2004). Initially, we reexamined the distribution of Myc during oogenesis and found that the protein is present in nurse cell nuclei throughout most of the endoreplication period (stages 2-9) but then begins to decline as Cic-L expression is switched on at stages 8-9 (Fig. 26A). To test if this pattern is affected by Cic-L, we then analyzed Myc accumulation in *cic-L* mutants. Notably, cic^{7} mutant egg chambers show persistent Myc accumulation in nurse cell nuclei beyond stage 10A (compare Fig. 26B' and 26C'). In addition, premature Cic-L expression driven by *mat-tubulin-Gal4* leads to complete downregulation of Myc in the resulting growth-arrested follicles (Fig. 26D). Thus, these observations suggest that Cic-L is a negative regulator of Myc.

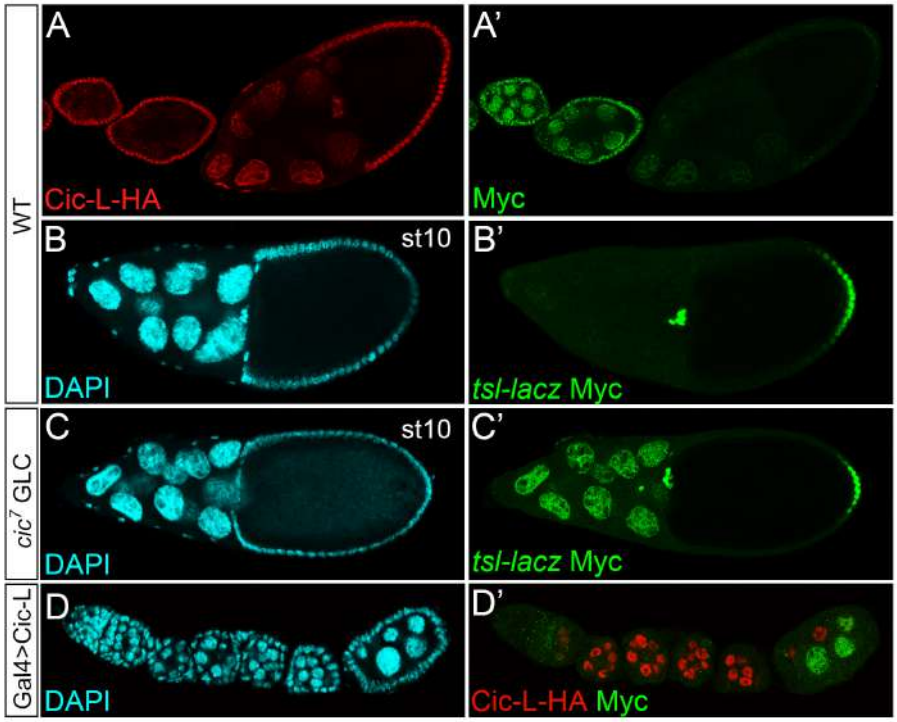

Figure 26. Cic-L promotes downregulation of Myc. (A) Cic-L-HA egg chambers stained with anti-HA (A) and anti-Myc (A'). **(B,C)** Stage 10 egg chambers of the indicated genotype carrying the *tsl-lacZ* reporter triple stained triple stained with DAPI (B,C); and anti-β-galactosidase and anti-Myc. Both the anti-β-galactosidase and the anti-Myc primary antibodies have been generated in mouse and we have been detected them the same anti-mouse secondary antibody (green). However, as it can be observed in Figures 23 and 24, the *tsl-lacZ* reporter is not expressed in the nurse cells. Myc **(D)** Egg chambers expressing Cic-L-HA under the control of the *mat-tubulin-Gal4* driver stained with DAPI (D), anti-HA and anti-Myc (D'). Note that occasional failure to activate Cic-L in some nurse cell nuclei results in larger cells that retain Myc staining.

3. Molecular analysis of Cic-L-specific activities

As explained in the introduction, Cic studies in mammals have used approaches that do not allow to discriminate between the two isoforms and, thus, no isoform-specific functions have been described. However, our findings in *Drosophila* showing that the two isoforms share functions, but they also have their own unique functions, raise the possibility that this is also the case for mammalian CIC proteins. Moreover, as Cic-L is well conserved, the molecular mechanism underlying these potential Cic-L-specific functions could be very similar in *Drosophila* and mammals. Thus, we decided to investigate if Cic-L unique functions in *Drosophila* were mediated by conserved domains. In particular, we studied the role of the Cic-L N-terminal region as well as the contribution of specific conserved domains within this region. In a parallel approach, we also performed a proteomic screen

to identify potential Cic-L-associated factors. Below I summarize the main results that we have obtained with these two approaches.

3.1 Dissection of the Cic-L N-terminal region

In order to determine if the Cic-L-specific N-terminal region is required for Cic-L unique activities, we used CRISPR/Cas9 to generate a new allele of Cic-L that we have named *cic*^{9}. As illustrated in Figure 27A, in this allele we have deleted most of the Cic-L N-terminal sequence and replace it by an in-frame triple HA tag (Fig. 27A). The majority of homozygous *cic*^{9} flies are viable, but the females are fully sterile and lay very few collapsed eggs. To ensure that the sterility was not caused by reduced stability of the Cic^{9} protein, we stained *cic*^{9} ovaries with an anti-HA antibody and confirmed that the protein accumulates normally in the nucleus of the nurse cells (Fig. 27B). Moreover, rhodamine-phalloidin staining revealed that the actin filament bundles were severely reduced and disorganized, resulting in a strong dumpless phenotype (Fig. 27C). Thus, these results suggest that the N-terminal region of Cic-L is necessary for its function in the nurse cells. However, we could not assess if this region is required for the establishment of embryonic AP polarity because embryos deposited by *cic*^{9} females did not develop.

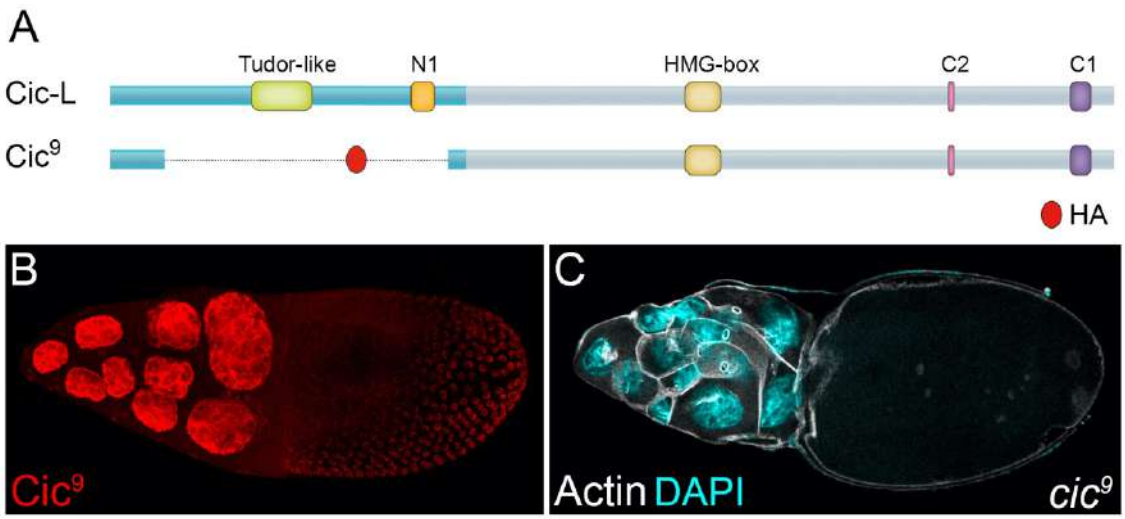

Figure 27. Cic-L-specific function in oogenesis is dependent on its N-terminal region. (A) Diagram indicating the protein structure of Cic-L and the Cic^{9} mutant, which lacks the Cic-L N-terminal sequence. Note that the nurse cell nuclei adjacent to the oocyte appear to be larger than normal. **(B)** Accumulation of the Cic^{9} protein at stage 10B detected by anti-HA staining. (C) Late stage *cic*^{9} egg chamber stained with phalloidin and DAPI.

As previously mentioned, Cic proteins are evolutionarily conserved from cnidarians to mammals. However, we were surprised to find that two more basal organisms, the

sponge *A. queenslandica* and the placozoan *T. adhaerens*, encode in their genomes Cic-L-like proteins containing the highly conserved N1 domain, but lacking the characteristic HMG-box and C1 domains of Cic (Fig. 28A). In support of this observation, we confirmed the structure of the *A. queenslandica* N1-encoding gene (*LOC100637684*) using deep RNA-seq data from this species (Fernandez-Valverde et al., 2015). Accordingly, the predicted stop codon in the genomic sequence lies at the expected position in the *LOC100637684* mRNA, indicating that the protein coding sequence does not extend further downstream. Since these N1-containing proteins appear to represent ancestors of Cic-L proteins before they acquired the HMG-box and C1 domains, we refer to them together as Proto-Cic.

Prompted by these observations, we wondered if a truncated version of Cic-L lacking the HMG-box and other C-terminal domains might still be functional in *Drosophila*, even though these domains are thought to be critical for all known Cic functions in any system. To explore this hypothesis, we generated a new transgenic construct based on the *cic-L* rescue construct but that only expresses the Cic-L N-terminus tagged with HA (Fig. 28A). The tag was particularly important in this case because the Cic-L^{Nter} derivative does not include the HMG-box, which is required for the nuclear localization of Cic-S (Astigarraga et al., 2007). However, the lack of HMG-box did not affect the nuclear localization of the Cic-L^{Nter} derivative, which recapitulated both the expression and subcellular distribution of endogenous Cic-L (Fig. 28B). Next, we combined the Cic-L^{Nter} derivative with the *cic^{9}* allele to analyze its functionality and found that it significantly rescues the dumping phenotype (Fig. 28C). This result suggests that the N-terminal region of Cic-L is not only necessary but sufficient for Cic-L-specific functions in oogenesis.

However, since the *cic^{9}* mutation is an in-frame deletion of the Cic-L N-terminal region that does not affect the common sequence, we could not rule out oligomerization of Cic-L^{Nter} derivative with the endogenous Cic proteins. To overcome this problem, we used CRISPR/Cas9 to generate a new compound allele, *cic^{10}*, that carries the *cic^{9}* deletion as well as an additional mutation in the common region. Specifically, we introduced a frameshift deletion that truncates Cic right before the HMG-box (Fig. 28A). The *cic^{10}* allele is homozygously lethal but generating GLCs we confirmed that it leads to the production of dumpless eggs (Fig. 28D). Interestingly, the Cic-L^{Nter} derivative significantly recues the dumping defects of this mutant background too (Fig. 28E).

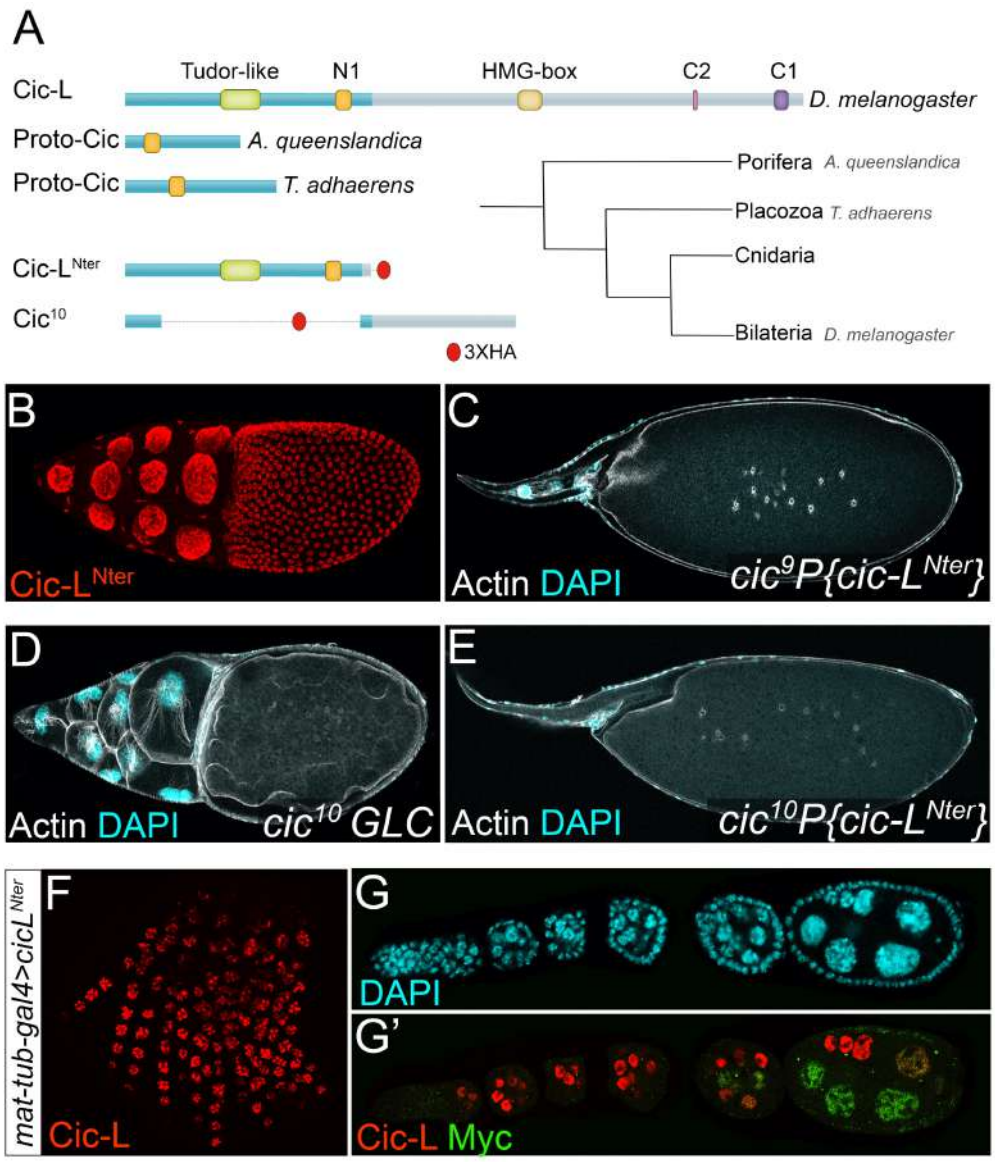

Figure 28. Cic-L controls oogenesis through its ancient, N-terminal module. (A) Schematic diagrams comparing the structure of Cic-L, two Proto-Cic forms from *A. queensladica* and *T. adhaerens*, the cic-L^{Nter} rescue construct, and the resulting protein from the *cic^{10}* mutation. A simplified phylogenetic tree indicating the evolutionary relationship between *A. queensladica*, *T. adhaerens* and *D. melanogaster* is also depicted. **(B)** Expression of the HA-tagged Cic-L^{Nter} protein in a stage 10 egg chamber stained with anti-HA. **(C)** DAPI and phalloidin staining of a *cic^{9}* mutant egg chamber expressing the Cic-L^{Nter} protein. **(D, E)** *cic^{10}* lat-stage egg chambers stained with DAPI and phalloidin. Note that the *cic^{10}* mutation causes strong dumping defects (D) that are rescued by expression of the Cic-L^{Nter} protein (E). **(F, G)** Ovary and ovarioles prematurely expressing HA-tagged Cic-L^{Nter} under the control of the *maternal-tubulin-Gal4* driver. Ovary stained with anti-HA (F) and ovarioles stained with DAPI (G), anti-HA and anti-Myc (G').

Finally, we also studied the effect of expressing the N-terminal region of Cic-L at early stages of nurse cell development. To this end, we assembled a *UAS* construct that only includes the Cic-L N-terminus and induced its expression with the *mat-tubulin-Gal4* driver. In line with the previous results, we found that premature expression of the Cic-L terminal

is sufficient to block nurse cell growth at early stages (Fig. 28F). Moreover, anti-Myc staining of these egg chambers showed complete downregulation of Myc in nurse cell nuclei with high expression of Cic-L$^{N\text{-}ter}$ (Fig. 28G'). Overall, our observations indicate that Cic-L unique activities are mainly dependent on its conserved N-terminal region, whereas the DNA binding HMG-box and C1 domains appear to be largely dispensable in this context. Moreover, the presence in basal metazoans of truncated versions of Cic-L consisting of N-terminal sequences suggests that Cic-L could represent an ancient regulatory module with autonomous activity.

Having established the importance of the N-terminal sequence in Cic-L unique activities, we next focused on the identification of functional domains within this region. To this end, we have generated in-frame deletions by CRISPR/Cas9 of 3 conserved domains. First, we identified a stretch of basic residues well-conserved in insects that we hypothesized that could function as a nuclear localization signal (NLS) (Fig. 29C). So, in order to visualize if deletion of this motif affects the nuclear localization of Cic-L, we conducted the deletion in the context of the *cic-L* allele tagged with the HA epitope. Surprisingly, Cic-L became cytoplasmic in DA follicle cells but remained mostly nuclear in the nurse cells (Fig. 29D). Thus, although this motif appears to function as an NLS, it is probable that additional NLSs (for example in the HMG-box) contribute to the nuclear localization of Cic-L in the nurse cells. Moreover, females that lack this motif are partly sterile and we found that the majority of embryos deposited by these females that do not hatch are bicaudal or exhibit defects in anterior structures. Instead, no apparent dumping defects were observed (data not shown).

We have also mutated the Tudor-like domain, which is located C-terminal to the NLS motif (Fig. 29A). This domain is not particularly well conserved at the sequence level and was initially identified by a new bioinformatic method providing additional information about conserved hydrophobic cores and domain architecture (Faure and Callebaut, 2013). Although it consists of two Tudor domains in tandem, it is not clear if they have methyl-binding activity because the key residues involved in methyl-binding could not be identified. In contrast, we have found that this domain is functional because its deletion causes partial female sterility. As for the deletion of the NLS, most of the embryos deposited by these females exhibit AP patterning defects (Fig. 29F), but the ovaries appear to be morphologically normal.

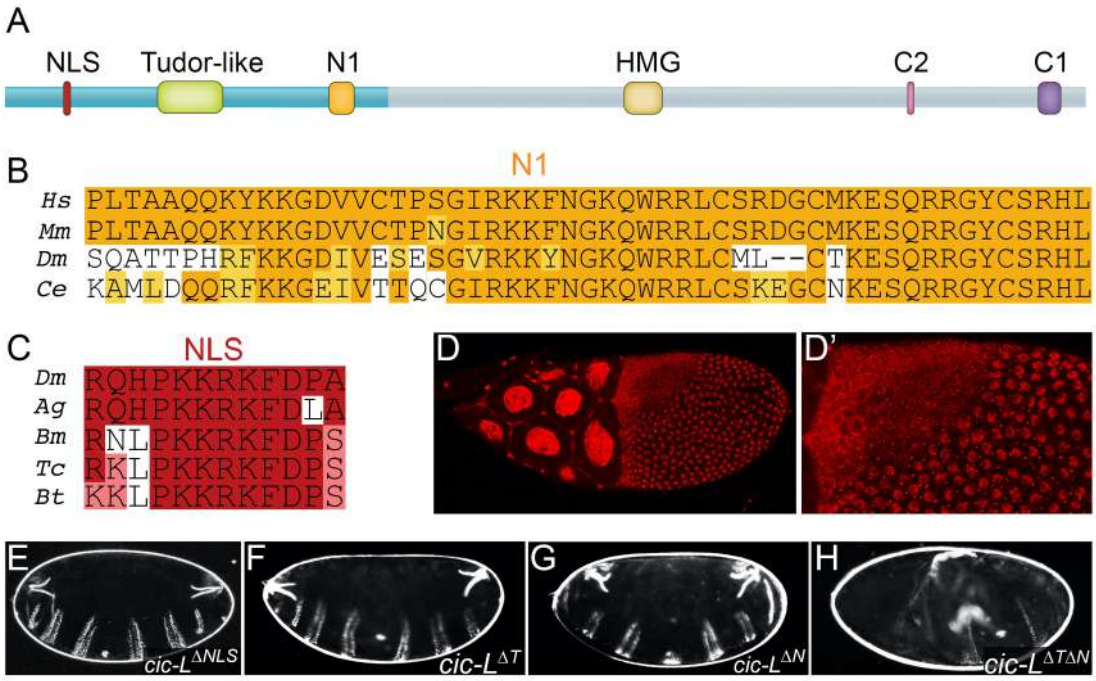

Figure 29. Different conserved motifs in the N-terminal region contribute to Cic-L-specific activity. (A) Diagram of the Cic-L protein indicating the position of different conserved domains. **(B, C)** Alignment of Cic-L N1 (B) and NLS (C) from different species. *Hs*, *Homo sapiens*; *Mm*, *Mus musculus*; *Dm*, *Drosophila melanogaster*; *Ce*, *Caenorhabditis elegans*; *Ag*, *Anopheles gambiae*; *Bm*, *Bombyx mori*; *Tc*, *Tribolium castaneum*; *Bt*, *Bombus terrestris*. **(D)** Expression of the HA-tagged Cic-L$^{\Delta NLS}$ protein in stage 10 egg chambers detected by anti-HA staining. A higher magnification of the DA follicle cells is shown in (D'). **(E-H)** Embryonic cuticle phenotypes of the indicated maternal genotypes.

We identified one final domain of unknown function in the N-terminal region of Cic-L, the N1 domain. This domain is highly conserved (as mentioned above it is detected in Porifera even), but it does not share any obvious similarity with other known domains (Fig. 29B). Due to its high conservation, we expected that loss of the N1 resulted in complete sterility, but it only causes partial sterility and the production of a fraction of bicaudal embryos (Fig. 29G). In fact, the deletions of the NLS, Tudor-like and N1 domains cause similar effects, indicating that the three domains contribute positively to Cic-L activity. In contrast, simultaneous deletion of both the Tudor-like and N1 domains results in complete sterility and much stronger phenotypes. The great majority of embryos deposited by these females do not develop and the occasional bicaudal embryos that can be observed have only one or two duplicated segments (Fig. 29H). Moreover, when we examined the ovaries of *cic-L*$^{\Delta T\Delta N}$ females, we found that dumping was clearly disrupted (data not shown). These observations suggest that none of these domains is essential by itself, but they play additive roles in Cic-L unique activities.

3.2 Identification of putative Cic-L cofactors

On the other hand, as no other study had previously investigated the molecular mechanism of Cic-L function, we performed a proteomic screen to identify Cic-L-associated factors. Specifically, Cic-L-containing complexes were purified from *Drosophila* cultured cells and then analyzed by mass spectrometry (Fig. 30A). For the purification, we took advantage of an improved method of affinity purification developed by the group of Dr. Veraksa based on the use of the streptavidin-binding peptide (SBP)(Yang and Veraksa, 2017). Briefly 1) the protein of interest tagged with SBP is expressed in S2 cells; 2) lysed cells are incubated with streptavidin beads; and 3) the SBP-tagged protein is eluted with biotin. Importantly, this group had already characterized the Cic-S interactome in *Drosophila* S2 cells using this affinity purification protocol combined with liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Yang et al., 2016).

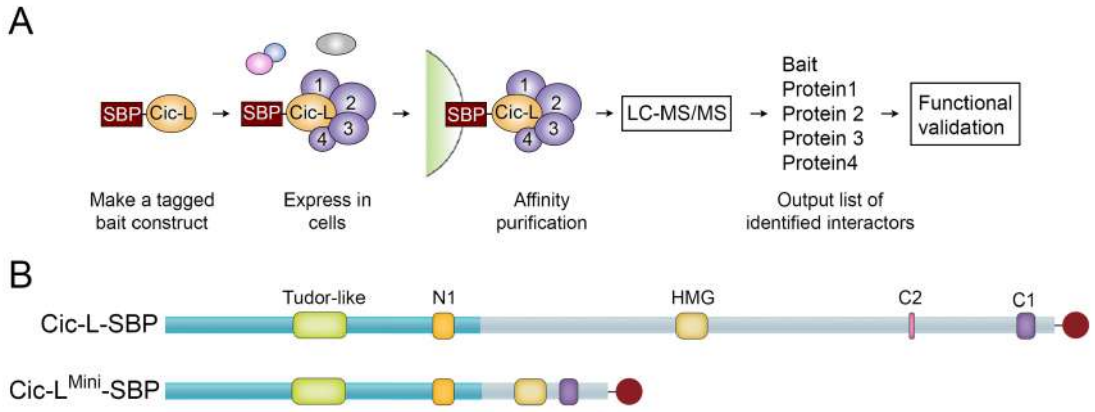

Figure 30. Proteomic approach to identify Cic-L-associated factors. (A) Affinity purification-mass spectrometry (AP-MS) workflow. **(B)** Diagram of the Cic-L-SBP and Cic-L^{Mini}-SBP proteins expressed in *Drosophila* S2 cells for the affinity purification.

We followed the same approach for Cic-L but decided to generate two different constructs in case the high molecular weight of Cic-L had a negative impact on its expression levels. So, we assembled a Cic-L full-length construct and a shorter version (Cic-L^{Mini}) containing the Cic-L-specific region fused to the HMG-box and C1 domains (i.e. lacking other, less conserved regions shared by Cic-L and Cic-S) (Fig. 30B). For each construct, we carried out 3 independent purifications in parallel with untransfected S2 cells as a control and the resulting samples were analyzed by LC-MS/MS. The mass spectrometry results were then processed with the SAINT (significance analysis of interactome) program, which calculates

the probability scores of a *bona fide* protein-protein interaction comparing the number of peptides obtained in the experimental and control samples (Choi et al., 2011). Summaries of Cic-L and Cic-L^{Mini} interacting proteins are provided in Tables 1 and 2, respectively.

Prey	Pep. Num. in exp.	Pep. Num. in ctrl.	Probability
rl	61|299|76	0|1|0	1
mnb	49|63|78	0|0|0	1
cic	651|1200|704	1|7|0	1
s6kII	16|33|56	0|0|0	1
wap	43|39|28	0|0|1	1
Nap1	39|25|16	0|0|9	0.9993
ago	17|5|17	0|0|0	0.9843
CtBP	11|22|30	3|2|4	0.9333
CG7194	2|5|4	0|0|0	0.9043
BEAF-32	2|4|4	0|0|0	0.8920
CkIIalpha	4|7|9	0|0|3	0.8867
Pgam5	3|7|7	0|2|0	0.8670
CG17528	3|7|8	0|0|2	0.8540
CG3714	3|5|10	2|0|0	0.8243
parvin	2|8|4	1|0|0	0.8203
FoxK	2|7|5	0|0|0	0.8200
CG9977	9|8|4	1|1|1	0.8123
CG2982	2|4|9	0|0|0	0.7933
CkIIbeta	2|2|2	0|0|0	0.7633
Dgp-1	2|3|3	0|0|1	0.7490
CG7065	4|6|2	1|0|0	0.7353
CG12301	3|2|3	0|0|0	0.7243
Dhx15	7|12|24	1|5|3	0.6927
Spt6	2|12|19	0|0|0	0.6817
CG3335	3|2|2	0|0|1	0.6703
lid	2|19|31	0|0|0	0.6660
Ltn1	2|3|9	0|0|1	0.6577

Table 1. Selection of interaction partners isolated from S2 cells expressing Cic-L-SBP. The Pep. Num. in exp. column indicates the number of unique peptides identified for each protein in the different experimental samples, whereas the Pep. Num. in ctrl. column indicates the number of peptides for the same protein observed in control samples.

Prey	Pep. Num. in exp.	Pep. Num. in ctrl.	Probability
cic	555|492|985	1|7|0	1
ago	9|7|13	0|0|0	1
CG12433	11|10|10	0|0|0	1
CG9977	9|10|14	1|1|1	0.9993
CG3335	9|13|9	0|0|1	0.9993
Nap1	46|29|74	0|0|9	0.9990
sli	6|10|7	1|0|0	0.9967

FoxK	5|6|9	0|0|0	0.9967
CG12301	6|5|4	0|0|0	0.9947
Xpc	9|4|4	0|0|0	0.9920
CG6701	4|5|4	0|0|0	0.9907
Ku80	4|15|9	0|0|0	0.9817
CG7065	4|4|8	1|0|0	0.9580
Ltn1	4|7|14	0|0|1	0.9570
lid	3|5|10	0|0|0	0.9473
BEAF-32	2|3|5	0|0|0	0.9267
CkIIalpha	4|5|15	0|0|3	0.8763
Spt6	3|4|20	0|0|0	0.8757
CkIIbeta	2|2|7	0|0|0	0.8700
CG17528	3|6|6	0|0|2	0.8650
Dgp-1	2|4|10	0|0|1	0.8397
Dhx15	8|17|22	1|5|3	0.8133
CG2982	2|5|17	0|0|0	0.8090
Pgam5	3|3|8	0|2|0	0.8003
parvin	2|2|3	1|0|0	0.7587
CG3714	3|2|6	2|0|0	0.6857

Table 2. Selection of interaction partners isolated from S2 cells expressing Cic-L^{Mini}-SBP.

With Cic-L full-length we were able to recover with known interactors of Cic-S such as the *Drosophila* MAPK ortholog Rolled (Astigarraga et al., 2007), Minibrain, Wings apart and CtBP (Yang et al., 2016). Importantly, these interactions were not detected with the Cic-L^{Mini} construct, which is consistent with the fact that they map to regions that are not included in this construct (e.g., the C2 MAPK docking site). Thus, these observations highlight the quality of the data that we have obtained with this approach. On the other hand, by comparing our results with the published Cic-S interactome we have also been able to identify putative Cic-L-specific interacting proteins. Moreover, the two independent Cic-L and Cic-L^{Mini} datasets has allowed us to focus only on proteins that are present in both of them. One protein identified with high confidence that caught our attention in particular was Archipelago (Ago). Notably, Ago peptides were not detected in control or Cic-S samples. As mentioned in the introduction, Ago is the substrate recognition component of the SCF^{Ago} ubiquitin ligase complex and has been reported to mediate the degradation of both CycE and Myc (Doronkin et al., 2003; Koepp et al., 2001; Moberg et al., 2001; Moberg et al., 2004; Zielke et al., 2011). Furthermore, Ago is conserved in humans, where its ortholog, FBW7, also participates in the degradation of CYCE and MYC (Koepp et al., 2001; Strohmaier et al., 2001; Welcker et al., 2004; Yada et al., 2004).

So, we set out to investigate the functional relevance of this physical interaction observed in S2 cells. As a first step, we decided to examine the expression pattern of Ago in the ovary by generating an *ago-GFP* expression construct. According to Flybase the *ago* gene is situated between two other genes, *CG1265* and *pav*, so we amplified the whole genomic sequence between these two genes to assemble the Ago construct. A diagram of the construct indicating the position of the GFP and the main conserved domains of Ago (the F-box domain and a stretch of WD40 repeats) can be found in Figure 31A. The F-box domain is involved in the recruitment of the SCF complex, whereas protein substrates are recognized through the WD40 repeats (Moberg et al., 2001; Welcker and Clurman, 2008). As I describe below, we have generated two additional Ago constructs in order to investigate the relation between Ago and Cic-L: a *UAS*-ago construct tagged with V5, and a derivative where the F-box domain has been deleted (Fig. 31A).

Using the *ago-GFP* construct we compared the expression pattern of Ago and Cic-L during oogenesis. Specifically, we had reasoned that the late expression pattern of Cic-L could be due to Ago-mediated degradation of Cic-L during early stages. In fact, Ago has been previously proposed to mediate Cic degradation (Suisse et al., 2017). We found that Ago accumulates in nurse cell nuclei during early oogenesis when Cic-L is absent but is also co-expressed with Cic-L at stage 10 (Fig. 31B). Thus, this observation points that rather than degrading Cic-L, Ago and Cic-L could function together at stage 10. However, to rule out that Cic-L is a degradation target of Ago in early oogenesis, we generated the *UAS-ago*$^{\Delta F\text{-}box}$ construct. As the F-box is required to recruit the SCF complex, we expected this Ago derivative to behave as a dominant negative that recognizes the substrates but cannot degrade them. We reasoned that if Cic-L is a target of Ago in early oogenesis, an Ago dominant negative form should prevent its degradation and lead to earlier accumulation of Cic-L in nurse cell nuclei. To test if this construct behaves as a dominant negative, we first analyzed its capacity to stabilize CycE in early stages of oogenesis. Indeed, expression of the *UAS-ago*$^{\Delta F\text{-}box}$ construct with the *mat-tubulin-Gal4* driver resulted in stabilization of CycE (Fig. 31C). In contrast, Cic-L accumulation in nurse cell nuclei was not affected by expression of Ago$^{\Delta F\text{-box}}$ (Fig. 31D). Therefore, Ago does not seem to mediate degradation of Cic-L during oogenesis.

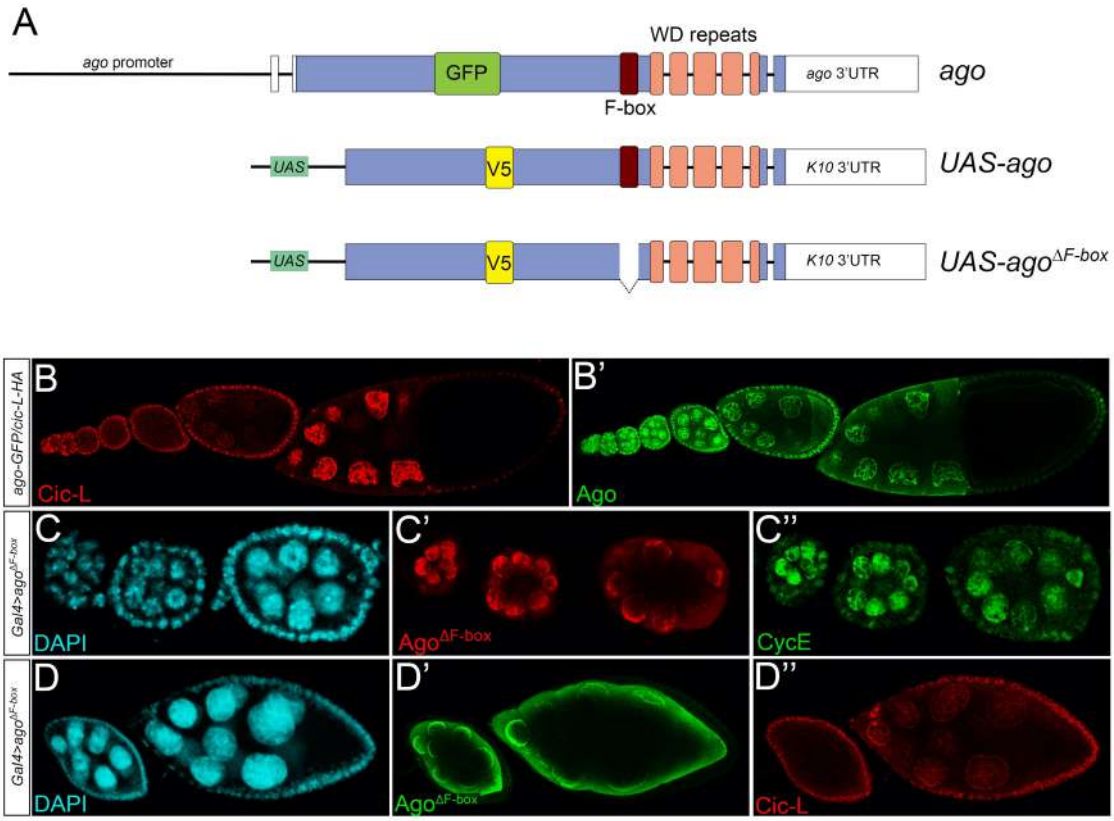

Figure 31. Ago does not mediate Cic-L degradation during oogenesis. (A) Diagrams of the *ago-GFP*, *UAS-ago-V5* and *UAS-ago$^{\Delta F\text{-}box}$-V5* constructs. Note that the construct is expressed under its own genomic regulatory regions. **(B)** *cic-L-HA/ago-GFP* ovariole stained with anti-HA (B) and anti-GFP (B'). **(C, D)** Egg chambers expressing V5-tagged Ago$^{\Delta F\text{-}box}$ under the control of the *mat-tub-Gal4* driver in an otherwise wildtype (C) or *cic-L-HA* background (D). They are stained with DAPI (C, D), anti-V5 (C', D'), anti-CycE (C''), and anti-HA (D''). Note that in early oogenesis Ago and Ago$^{\Delta F\text{-}box}$ accumulate uniformly in nurse cell nuclei but later become concentrated in the periphery of the nucleus and in the cytoplasm of nurse cells.

Next, considering the opposite effects of Cic-L on the accumulation of CycE and Myc, we thought of two alternative working hypotheses. On the one hand, that Cic-L interacts with Ago to inhibit the degradation of CycE, and on the other hand, that Cic-L interacts with Ago to stimulate the degradation of Myc. To address these hypotheses, we designed two different rescue experiments. First, we investigated if overexpression of Ago could rescue the nurse cell growth arrest caused by premature expression Cic-L. The idea behind this experiment was that if Cic-L overexpression leads to the stabilization of CycE via inhibition of Ago, this could be alleviated by providing more Ago. However, overexpression of Ago with the *mat-tubulin-Gal4* driver did not rescue nurse cell growth (Fig. 32A). Initially, we reasoned that the lack of nurse cell growth rescue could be explained if Cic-L induces nurse cell growth arrest through different independent mechanisms. Nevertheless, we found that

overexpression of Ago did not rescue the oscillations of CycE either (Fig. 32B''). On the other hand, to test if Cic-L stimulates the degradation of Myc via Ago, we tried to rescue the downregulation of Myc caused by overexpression of Cic-L with the $Ago^{\Delta\text{-Fbox}}$ dominant negative. However, overexpression of $Ago^{\Delta\text{-Fbox}}$ did not rescue the nurse growth arrest nor the downregulation of Myc caused by overexpression of Cic-L (Fig. 32C, D''). Therefore, from our observations we cannot conclude that the physical interaction observed between Cic-L and Ago in S2 cells is functionally relevant for Cic-L-mediated regulation of nurse cell growth.

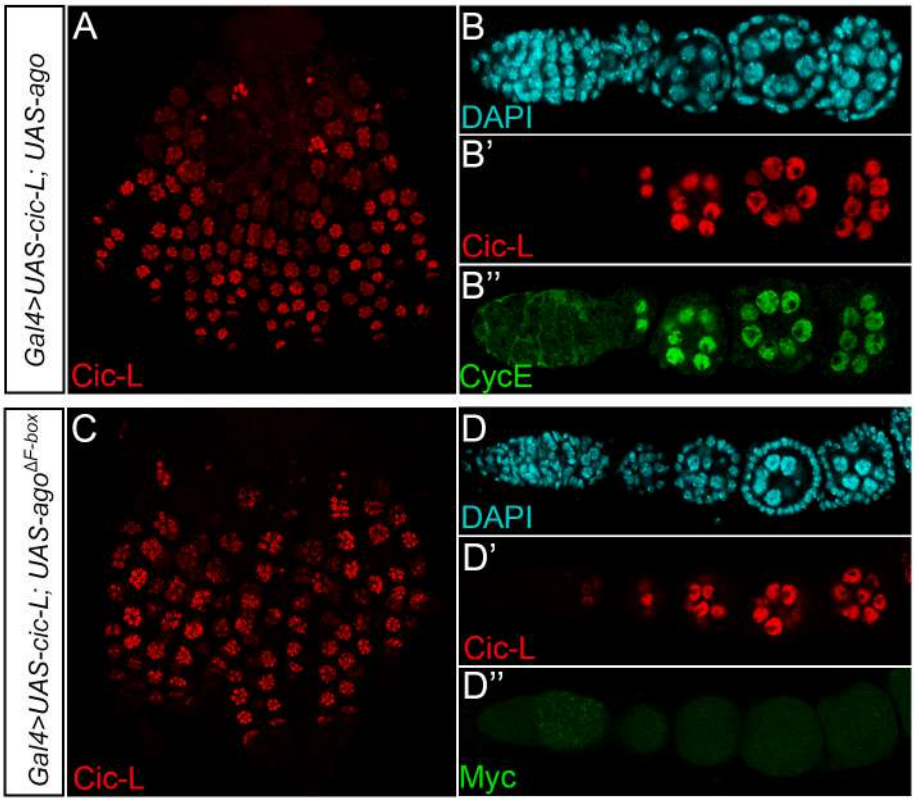

Figure 32. Ago does not mediate Cic-L regulation of CycE and Myc. (A, B) Ovary (A) and egg chambers (B) expressing HA-tagged Cic-L and V5-tagged Ago under the control of the *mat-tubulin-Gal4* driver. They are stained with anti-HA (A, B'), DAPI (B) and anti-CycE (B''). **(C, D)** Ovary (C) and egg chambers (D) expressing HA-tagged Cic-L and V5-tagged $Ago^{\Delta F\text{-box}}$ under the control of the *mat-tubulin-Gal4* driver. They are stained with anti-HA (C, D'), DAPI (D) and anti-Myc (D'').

DISCUSSION

Over the two decades since the discovery of Cic as a regulator of embryonic patterning in *Drosophila*, great progress has been made in understanding this conserved HMG-box protein. For example, we have learnt that Cic is a direct phosphorylation target of MAPK in different developmental contexts, and that this control is critical for the correct interpretation of RTK-Ras-MAPK signaling in vivo (Astigarraga et al., 2007; Futran et al., 2015; Jiménez et al., 2012b). Also, Cic has been shown to regulate RTK signaling responses by binding to similar octameric sites in target enhancers through its conserved HMG-box and C1 domains (Ajuria et al., 2011; Forés et al., 2017b; Kawamura-Saito et al., 2006). Many of these mechanisms have been dissected first in *Drosophila*, where Cic controls well studied regulatory processes that offer a tractable system for genetic and molecular analyses. However, most of the basic principles underlying Cic activity and regulation in *Drosophila* have been found to be conserved in mammals. Furthermore, as the studies on mammalian CIC proteins have been progressing, we are increasingly appreciating their importance in human disease. Indeed, *CIC* is now counted as one of the most frequently mutated tumor suppressors across a variety of cancer subtypes, including oligodendroglioma and lung and gastric adenocarcinomas (Campbell et al., 2020). Moreover, the roles of CIC downstream of mammalian RTK pathways suggest that oncogenic RTK activation is often mechanistically linked to CIC inactivation and derepression of CIC targets such as the as the *ETV1/4/5* family of proto-oncogenes, as seen in the case of glioblastomas (Bunda et al., 2019). Additionally, dysregulation of CIC activity has been also implicated in other clinical disorders, particularly in SCA1 and other neurobehavioral syndromes.

Nevertheless, these advances do not yet deliver a full picture of Cic biology, and several important questions still remain unresolved. In this thesis we have addressed one such question that is frequently cited as one of the most fundamental of all, namely, the functional significance of the two Cic isoforms present in both *Drosophila* and mammals (Jiménez et al., 2012b; Simón-Carrasco et al., 2018a; Tan et al., 2018). By exploiting once again the powerful genetic tools of *Drosophila*, I provide the first functional and molecular characterization of Cic-L in this system. As discussed below, considering the high evolutionary conservation of Cic-L sequence and structure across species, we expect that many of our observations will turn out to be applicable to mammalian CIC-L as well.

Our studies have a number of implications both in terms of Cic-L and Cic-S function and regulation, as well as regarding the mechanisms controlling the growth of nurse cells and their transition to dumping. In the following sections, I discuss each of these topics in detail.

Finally, I consider our results from a broader perspective by discussing their evolutionary significance and their relevance beyond *Drosophila*.

1. Common and unique features of cic isoforms

1.1 Expression pattern and regulation by RTK signaling

Our studies have revealed that the Cic-L and Cic-S exhibit both redundant as well as exclusive functions. Previous studies in mammals had already hinted to such redundant functions of both isoforms, although without offering a firm conclusion in this respect. For example, studies analyzing the expression of the two CIC isoforms in mice revealed that CIC-L and CIC-S are co-expressed in several embryonic and postnatal tissues (Kim et al., 2015; Lee et al., 2011). This observation favored the implicit idea that the two CIC isoforms are functionally equivalent, although this has not been rigorously tested using isoform specific mutations. To address this question more definitively in *Drosophila*, we have first compared the expression of the two isoforms in different tissues and found that they exhibit overlapping but also distinct expression patterns. Among the tissues that we have analyzed, the two isoforms are co-expressed in wing discs from LIII larvae and in the main body follicle cells, where Cic functions in connection with EGFR signaling. In contrast, in two other types of follicle cells, the border and stretch cells, only the long isoform is expressed. Although we have not investigated if Cic-L plays a specific role in these cells, it is worth mentioning that dorsal migration of border cells also depends on EGFR signaling through MAPK (Bianco et al., 2007). Cic-L is also specifically expressed in a group of cells in the germarium, although we have not yet confirmed their germline or somatic origin. Therefore, it is possible that Cic-L functions individually in all these ovarian cell types, and it will be interesting to study if these potential functions involve a canonical, RTK-dependent mechanism or a Cic-L-specific regulatory function as discussed below.

Conversely, in the early embryo we only detect the short isoform. This is surprising because Cic-S is maternally provided as RNA and the transcripts encoding both isoforms are present in nurse cells at stage 10 (data not shown). Furthermore, expression of Cic-L under the regulatory sequences present in the original *cic-S* rescue construct has been shown to drive Cic-L accumulation in the early embryo (Forés et al., 2015), which suggests that this specific accumulation of Cic-S is determined by sequences present in its mRNA rather than resulting from Cic-L protein instability. Because the 3'UTR sequences present in the *cic-S*

and *cic-L* rescue constructs are the same, the 5'UTRs could differentially affect the stability or translation of the two mRNAs. Perhaps more intriguingly, it is unclear if this differential accumulation reflects a biological requirement to prevent Cic-L expression in the blastoderm. If so, we can speculate that the presence of Cic-L in the embryo might disrupt polyploidization of early endoreplicating tissues. For example, large polyploid cells have been detected in the amnioserosa, which is actually an extraembryonic tissue, 220 minutes after egg laying (Foe, 1989).

In addition, in some tissues we have detected differences in the temporal accumulation and subcellular distribution of Cic-L and Cic-S. For example, while Cic-S is expressed in the nurse cells from early stages, Cic-L does not begin to accumulate in these cells until stage 8. One straightforward explanation for the late expression of Cic-L would be that Cic-L is downregulated by RTK signaling during early oogenesis. However, the Cic-L^{Nter} derivative does not contain the C2 MAPK docking site and exhibits the same late accumulation pattern, which suggests that Cic-L is not regulated by RTK signaling in the nurse cells. Still, we cannot rule out the presence of additional MAPK docking sites in the Cic-L-specific N-terminal region. To begin investigating the regulation underlying Cic-L expression, we have generated a *cic-L-GFP* reporter and our preliminary results indicate that Cic-L accumulation in the nurse cells is regulated at the transcriptional level. At the moment we do not know what triggers the expression of *cic-L* but considering the role of Cic-L in nurse cell endocycle termination, one hypothesis could be that the expression of *cic-L* is turned on when the nurse cells reach a DNA content threshold. However, this is only a tentative idea, since any potential mechanisms linking DNA content to gene expression remain poorly studied. Alternatively, *cic-L* expression might be activated by ecdysone signaling, which functions in the germline during mid-oogenesis and allows the progression of egg chambers into the vitellogenic phase of oogenesis (Buszczak et al., 1999).

On the other hand, we have found that Cic-L is mostly nuclear in two groups of cells where Cic-S becomes partially relocalized to the cytoplasm: the DA follicle cells at stage 10A and the nurse cells from stage 10B onwards. In the case of the follicle cells, our results suggest that Cic-L and Cic-S function redundantly in this context by repressing *mirr* (see below). However, their differential subcellular distribution does not seem to reflect different activities of the two isoforms downstream of EGFR signaling. In fact, despite its nuclear accumulation in DA follicle cells, our results indicate that Cic-L-mediated repression of *mirr* in these cells is still inhibited by EGFR signaling. Still, it is possible that the nuclear

localization of Cic-L in follicle cells is related to additional functions not regulated by EGFR signaling. For example, we have noticed that sometimes the chorion of eggs deposited by *cic-L* mutant females appears to be thinner, a phenotype that has been associated with defective gene amplification in follicle cells (Calvi and Spradling, 1999). It is also possible that nuclear Cic-L controls certain aspects of the complex cellular rearrangements required for the formation of the dorsal eggshell appendages, which are also affected in *cic-L* mutants. In addition, we have found that Cic-L-specific nuclear localization in DA follicle cells is related to the presence of a stretch of conserved basic residues (NLS) present in its N-terminal region.

Cic-L also remains nuclear in nurse cells after stage 10B when Cic-S becomes cytoplasmic. In contrast to DA follicle cells, we have found that deletion of the N-terminal NLS does not visibly affect the nuclear localization of Cic-L in the nurse cells, but it leads to partial female sterility and the production of bicaudal embryos. However, at the moment, we do not know if this is because we are not able to detect subtle changes in the subcellular distribution of Cic-L or if this motif has additional functions that contribute to Cic-L-specific functions. In addition, we have investigated if the relocalization of Cic-S to the cytoplasm in nurse cells takes place in response to EGFR signaling, as it happens in DA follicles. Nevertheless, a Cic-S$^{\Delta C2}$ derivative, which is insensitive to MAPK-mediated downregulation, exhibits the same relocalization in late nurse cells (data not shown). Therefore, rather than being specifically relocalized to the cytoplasm, the most probable explanation is that Cic-S becomes cytoplasmic due to the permeabilization of the nurse cell nuclear envelope as they undergo PCD (Buszczak and Cooley, 2000).

Therefore, we have found that both isoforms are co-expressed in several tissues (like in published mammalian studies), but they also exhibit different expression patterns in other tissues. Although these expression differences could already explain isoform-specific functional requirements, below I discuss how the unique functions of each isoform depend on specific domains.

1.2 Functions and mechanism of action

In addition, the functional significance of the two isoforms was not known in any species because all the work on Cic had been done on Cic-S or using approaches that did not allow to discriminate between the two isoforms. Therefore, we have addressed this question by studying the effects of inactivating each isoform specifically,

Our results indicate that Cic-L and Cic-s act redundantly in at least two MAPK-regulated processes: wing vein specification and the establishment of embryonic DV polarity in ovarian follicle cells. Previous studies have shown that during the specification of wing veins Cic represses vein-specific genes such as *aos* (Ajuria et al., 2011; Roch et al., 2002), whereas in follicle cells Cic represses *mirr* (Andreu et al., 2012b). In general, Cic represses its target genes by binding to TGAATGAA-like motifs present in their promoters (Ajuria et al., 2011; Kawamura-Saito et al., 2006). As it has been recently demonstrated, specific binding to the TGAATGAA-like sites requires the HMG-box and the C1 domains (Forés et al., 2017b), which are present in the two isoforms. In contrast, the mechanism of Cic repression is less well understood and has been proposed to be context-dependent (Forés et al., 2015). In the early embryo, Cic-mediated repression relies on its interaction with the Gro corepressor, which requires the Cic-S-specific N2 domain. Instead, loss of Gro function in the developing wing and the follicular epithelium does not recapitulate the effects observed in *cic* mutants such as derepression of *aos* and *mirr* transcription. This led to the proposal that Cic-S might function through other corepressor in those tissues (Forés et al., 2015). However, in light of the redundancy between the two isoforms in these tissues, we can now re-interpret those observations and propose two different scenarios, which are discussed below.

First, it is possible that the two isoforms repress *aos* and *mirr* through a common Gro-independent mechanism. In this regard, studies of mammalian CIC have proposed alternative CIC corepressors such as ATXN1/ATXN1L, the SIN3 deacetylation complex and the SWI/SNF complex (Bowman et al., 2007; Crespo-Barreto et al., 2010; Hwang et al., 2020; Lam et al., 2006; Lee et al., 2011; Weissmann et al., 2018). The interaction with ATXN1 and ATXN1L is particularly interesting because it maps to a common region conserved in *Drosophila* (Lam et al., 2006). Moreover, Atxn1 was recovered as a Cic-S interactor in the early embryo (Yang et al., 2016). However, unpublished results from our lab suggest that Cic functions independently of Atxn1 in *Drosophila*. On the other hand, the SIN3 deacetylation and SWI/SNF complexes were identified through proteomic screens. Nevertheless, we currently do not know if these interactions are common to both isoforms because they have not been mapped to any specific domain of CIC. This also complicates the prediction of whether these interactions could be conserved in *Drosophila*. It is also worth mentioning that we isolated the CtBP corepressor in our screen for Cic-L-associated factors. In addition, the group of Dr. Veraksa also isolated CtBP in their characterization of the Cic-S interactome (Yang et al., 2016), suggesting that the two

isoforms could employ a common mechanism of repression mediated by CtBP. Interestingly, *CtBP* mutant germline clones cause derepression of *tll*; and double *CtBP, gro* mutant germline clones result in uniform expression of *tll* throughout the embryo (Cinnamon et al., 2004). Finally, it is also conceivable that Cic mediates repression by just binding to DNA. During DNA binding, HMG-box proteins are capable of bending DNA through interactions in the minor groove, which could alter chromatin conformation and/or facilitate the binding of other proteins to repress transcription (Giese et al., 1992; Malarkey and Churchill, 2012). However, no conclusive evidence supporting a specific role of HMG-box-mediated DNA bending in transcriptional control has been reported up to this date (Kamachi and Kondoh, 2013).

In the second scenario, each isoform might employ a different mechanism of transcriptional repression. If this correct, then Gro would still mediate Cic-S repressor activity in the wing and the ovary, and it would be the presence of Cic-L that renders both Cic-S and Gro dispensable in these contexts. However, with the evidence available at the moment it is not possible to clearly determine if the mechanism of repression underlying redundant functions is shared between the two isoforms or if each isoform employs a specific mechanism. Thus, to distinguish between these two possibilities I propose an overexpression and a rescue experiment. In the overexpression experiment we could generate a *UAS* construct containing just the sequence common to both isoforms but lacking the C2 domain (to make it insensitive to MAPK-mediated downregulation). By expressing this construct in follicle cell clones we could test if the common domains are sufficient to repress *mirr*. Alternatively, we could generate a construct expressing the common region under the promoter of *cic-L* or *cic-S* to investigate if it rescues the extra vein tissue and DV polarity defects caused by the loss of Cic function.

In addition to these common activities, Cic-L and Cic-S also have unique functions that rely on isoform-specific domains. Previous studies have shown that Cic-S has specific functions in the *Drosophila* embryo, where it represses several genes: the terminal gap genes *tll* and *hkb* (Ajuria et al., 2011; Jiménez et al., 2000); the neuroectodermal gene *ind* (Ajuria et al., 2011; Lim et al., 2013); and the dorsal gene *zerknüllt* (*zen*) through cooperative binding with Dorsal (Jiménez et al., 2000; Papagianni et al., 2018). Importantly, these Cic-S-specific functions do not result from selective accumulation of Cic-S in the embryo because expressing Cic-L maternally in the embryo does not rescue the embryonic *cic* phenotype. In contrast, expression of a Cic-L derivative containing the N2 domain results in significant rescue of the *cic* embryonic phenotype (Forés et al., 2015).

These observations support the notion that the N2 motif endows Cic-S with a repressor activity that is critical for its functions in the early embryo (Forés et al., 2015).

On the other hand, in this thesis we have shown that Cic-L also has unique functions. As it will be discussed in more detail in the next section, Cic-L regulates nurse cell growth and the initiation of the dumping process; morphogenesis of dorsal appendages (Rittenhouse and Berg, 1995; Tran and Berg, 2003); and embryonic AP patterning (directly or indirectly). Our observations indicate that these unique functions of Cic-L primarily depend on its specific N-terminal module. In addition, we have demonstrated the functionality of two highly conserved domains of this region, the Tudor-like and the N1 domains. These two domains appear to play additive roles that contribute to Cic-L-specific activities because individual deletion of each domain only causes partial sterility and no obvious dumping defects. In contrast, simultaneous deletion of both domains leads to complete sterility and persistence of nurse cells. This behavior argues the Tudor-like and N1 domains exert parallel, partly redundant functions (perhaps interacting with different factors) that contribute to full Cic-L activity, rather than acting in a single molecular pathway in which each domain is critical for a given response. If this correct, it might also pose an additional challenge to define the specific mechanisms underlying Cic-L activity, since the analysis of the different contributing mechanisms might reveal relatively subtle effects.

A further question related to the differences between Cic-L and Cic-S is why only Cic-L is essential for viability. We can envision two possibilities: differential expression of the *cic-L* gene or an unknown Cic-L-specific function essential for the viability of larvae and pupae. These two scenarios are not mutually exclusive, and it may well be that *cic-L* lethality arises from cumulative defects in distinct processes where Cic-L is expressed individually or functions via its own molecular mechanism(s).

Strikingly, the common domains of Cic, including the HMG-box and the C1 and C2 domains, are largely dispensable for Cic-L-specific activities. This suggests that Cic-L-specific activities do not involve RTK-regulated repression of target genes through binding to specific TGAATGAA-like sites. Thus, in the context of these functions, Cic-L is best thought of as a novel factor unrelated to Cic-S. In line with these findings, two basal metazoans, the sponge *Amphimedon queenslandica* and the placozoan *Trichoplax adhaerens*, encode truncated versions of Cic-L consisting of N-terminal sequences without an HMG-box or C1 domains. Therefore, the N-terminal region of Cic-L probably represents an ancient regulatory module with autonomous activity. In contrast, the Cic-S-specific N2

domain is not conserved in mammalian CIC proteins and represents a structural innovation that originated in dipteran insects (Forés et al., 2015).

To begin investigating these HMG-box-independent activities of Cic-L we have performed a proteomic screen. In particular, we have isolated Cic-L-associated factors in S2 cells by affinity purification and analyzed them by mass spectrometry. By comparing the factors that we have obtained with the previously published Cic-S interactome we have been able to identify several candidates to mediate Cic-L-specific activities. We have initially focused on the F-box protein Ago, but our current results do not support a critical, Cic-L-associated function of Ago in the nurse cells. However, we cannot rule out that Cic-L and Ago function together in other tissues. Moreover, among the factors that we have isolated, many of them have not been previously studied and their function is unknown. This suggests that Cic-L probably participates in poorly characterized and complex molecular processes, which could be further studied using Cic-L as an entry point. Regarding the other factors with known or predicted functions, we have isolated several chromatin proteins such as Nap1, BEAF-32 and CG2982; and DNA repair proteins such as Xpc and Ku80. So, beyond the canonic activity of Cic-L as transcription factor linked to RTK signaling, we speculate that Cic-L has additional functions in chromatin regulation.

2. Functional analysis of Cic-L in *drosophila* oogenesis

2.1 Cic-L induces nurse cells to exit the endocycle

Oocytes are amazing cells capable, once fertilized, of initiating and in most cases completing by themselves the embryonic development of an animal. This remarkable task requires that the oocyte is loaded with huge supplies of metabolites, macromolecules and organelles which make them one of the largest, if not the largest, cell in the body. This comes at the cost of a very high biosynthetic demand, which in many species, including mammals and *Drosophila*, is solved with dedicated, sister germ cells that "nurse" the oocyte as it differentiates and grows during oogenesis. In *Drosophila*, where the oocyte grows ~5000 times in volume, the nurse cells themselves become giant cells to fuel the demanding oocyte. In particular, nurse cells become polyploid by undergoing 10-12 cycles of DNA endoreplication between stages 2-10 of oogenesis. Although these endocycles have been less studied than those of the salivary gland, it is known that they are similarly driven by oscillations of CycE/Cdk2 activity, which trigger DNA synthesis at high level and

then allow pre-RC assembly during the period of lower activity (G phase) (Edgar et al., 2014).

In the nurse cells, in particular, these oscillations are thought to depend mainly on the periodic degradation of CycE mediated by Ago which, cycle after cycle, would drive the progressive increase in nurse cell volume (Doronkin et al., 2003; Lilly and Spradling, 1996; Moberg et al., 2001). What terminates this cycle once the nurse cells reach their final size, however, has not been investigated and is actually unknown in most other cases of polyploidization. In fact, the only case in which this exit mechanism has been studied in animals is the endocycle-to-amplification transition in follicle cells, which is regulated by Notch and Ecdysone signaling. In particular, these signaling pathways control the up-regulation of Ttk69, a transcription factor whose overexpression stops endoreplication prematurely (Sun et al., 2008). As Ttk69 represses *CycE* transcription in other contexts, the authors proposed that up-regulation of Ttk69 at the endocycle-to-amplification transition could be necessary to lower the levels of CycE/Cdk2 activity. Specifically, they argued that higher levels of CycE/Cdk2 activity are probably necessary for genomic endoreplication than for gene amplification. This model, however, could be specific for the follicle cells, because (i) the hypomorphic $CycE^{01672}$ mutation, which attenuates the oscillations of CycE, causes extended S phases but does not block endocycle progression in the nurse cells (Lilly and Spradling, 1996); and (ii) reduced pulses of *CycE* transcription in late *dDP* and *dE2F* embryos have been found to be sufficient to drive endocycle progression, albeit more slowly (Duronio et al., 1998; Royzman et al., 1997). Therefore, endocycle exit may occur differently in these two contexts.

Indeed, our results indicate that endocycle termination in nurse cells is associated with stabilization of CycE. Although this finding was initially unexpected considering the precedent of the follicle cells just mentioned above, in fact, this result is consistent with previous evidence showing that constitutive CycE expression causes endocycle arrest due to inhibition of pre-RC assembly (Follette et al., 1998; Weiss et al., 1998). In addition, we have demonstrated that this stabilization of CycE depends on Cic-L, which is specifically induced during the final stages of nurse cell growth. Importantly, we have found that forced early expression of Cic-L is sufficient to induce premature nurse cell growth arrest and CycE stabilization, suggesting that Cic-L signals the timely exit of nurse cells from endoreplication. In addition, to further explore how Cic-L controls endocycle termination and CycE stabilization, we have studied the relationship between Cic-L and Myc. In *Drosophila*, Myc promotes endoreplication and growth of multiple cells and tissues,

including the nurse cells, and reduced Myc activity severely compromises their growth rate and size (Gallant, 2013). Besides, overexpression of Myc has been shown to increase the levels of CycE in mitotic cells (Prober and Edgar, 2000). In agreement with its growth-promoting role, we observe that Myc accumulates in nurse cell nuclei during early and mid-oogenesis. Instead, towards the end of the nurse cell growth phase, the levels of Myc begin to decline as the levels of Cic-L increase, suggesting that Cic-L could be a negative regulator of Myc. Indeed, in *cic-L* mutants the nurse cell expression of Myc persists for a longer time; and conversely, premature expression of Cic-L disrupts the early accumulation of Myc. Nevertheless, at the moment we do not know if Cic-L induces the stabilization of CycE via downregulation of Myc. In fact, Maines *et al.* generated mosaic germline cysts of the *Myc* lethal allele, dm^2, and found that CycE oscillations were relatively normal in dm^2/dm^2 nurse cells (Maines et al., 2004). Therefore, it is possible that Cic-L affects the stability of CycE and Myc through independent mechanisms. Moreover, although our observations indicate that Cic-L is a regulator of CycE and Myc, the evidence presented so far does not allow us to discriminate whether this regulation is direct or indirect. Below, I consider different hypothesis about the molecular mechanism underlying Cic-L-mediated regulation of CycE and Myc.

First, the effects that we observe on CycE and dMyc protein accumulation could originate from transcriptional changes. In fact, Cic represses *cycE* transcription in intestinal stem cells (Jin et al., 2015), which is one of the reasons why we first decided to examine CycE expression in the ovary. While the results we obtained (that Cic-L induces stabilization, not downregulation of CycE) do not fit well with that repressor mechanism, Cic-L could cause a positive effect on *CycE* transcription. To begin to address this possibility, we have characterized the expression patterns of *CycE* and *Myc* during oogenesis by in situ hybridization. One initial observation is that whereas the protein levels of CycE differ among the nurse cells in a single follicle during early stages, the mRNA levels are generally uniform (Fig. 33A). This reinforces the idea that, in contrast to the situation in the salivary glands, progression of the endocycle in the nurse cells is not regulated at the level of *CycE* transcription. Moreover, we have noticed that the levels of *CycE* transcript increase as oogenesis proceeds (Fig. 33A). Although we initially thought that this increase could be related to the stabilization of CycE at stage 10, we have observed that other maternal transcripts exhibit a similar pattern, which probably reflects the need to load the embryo with such transcripts. In addition, we have preliminary, non-quantitative results that suggest that the levels of *CycE* mRNA at stage 10 are not affected in *cic-L* mutant egg chambers

(data not shown). We are now trying to confirm this result by analyzing the levels of *CycE* in early egg chambers overexpressing Cic-L. On the other hand, we have found that *Myc* transcripts also accumulate in nurse cells during oogenesis and they are clearly detectable at stage 10 (Fig. 33B). Thus, the downregulation of Myc at this stage does not appear to be caused by transcriptional repression of the gene. Also, we have performed a *Myc* in situ hybridization on *cic*^{7} GLC egg chambers and our preliminary results suggest that the levels of *Myc* RNA are not significantly affected (data not shown). Therefore, it seems unlikely that the primary effect of Cic-L on CycE and Myc occurs at the transcriptional level.

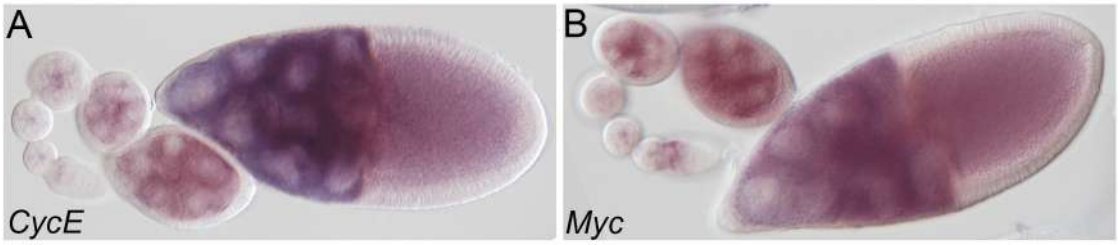

Figure 33. *CycE* and *Myc* transcripts accumulate in nurse cells during oogenesis. (A, B) Wildtype mRNA expression patterns of *CycE* (A) and *Myc* (B) during oogenesis.

Another possibility, although we have not explored it, is that Cic-L directly or indirectly regulates the translation of CycE and/or Myc. Interestingly, the 5'UTR of *CycE* contains several open reading frames (ORFs), which led Prober and Edgar to suggest that CycE translation could be sensitive to the rate of cellular growth (Prober and Edgar, 2000). Based on previous findings about the yeast G1 cyclin Cln3 (Polymenis and Schmidt, 1997), they reasoned that the presence of upstream ORFs could reduce initiation of translation at the downstream CycE translation start site, rendering translation dependent on high abundance of ribosomes in rapidly growing cells. Regarding dMyc translation, a recent study has found that Insulin signaling is required for Myc protein accumulation in the germarium (Wang et al., 2019). Mechanistically, the Insulin pathway has been shown to regulate protein translation by repressing the translation initiation factor 4E binding protein (4E-BP or Thor in *Drosophila*) (Miron et al., 2003), and Wang *et al*. found that knockdown of *thor* partially rescued the reduction in Myc levels observed in Insulin signaling-defective mutants (Wang et al., 2019). Thus, one mechanism controlling Myc activity in the female germline takes place at the level of protein translation, raising the possibility that Cic-L also functions, at least in part, through this or a related pathway.

On the other hand, we have also considered the idea that Cic-L regulates CycE and/or Myc postranslationally. As we identified Ago as a putative Cic-L co-factor, and Ago is known to mediate phosphorylation-dependent degradation of both CycE and dMyc (Doronkin et al., 2003; Moberg et al., 2001; Moberg et al., 2004; Shcherbata et al., 2004; Zielke et al., 2011), we have investigated if Cic-L regulates CycE and/or Myc degradation directly through an interaction with Ago. Such mechanism would be similar to the reported activity of the Minus protein, which interacts with Ago and in this way influences the turnover of CycE, although not of dMyc (Szuplewski et al., 2009). First, we have studied if Cic-L inhibits Ago-mediated degradation of CycE. As early expression of Cic-L causes premature stabilization of CycE, we have attempted to rescue this effect by overexpressing Ago together with Cic-L. However, we did not observe any recovery of CycE oscillations in this experiment. On the other hand, we have also considered the possibility that Cic-L stimulates the activity of Ago in degrading Myc. To test this idea, we have generated a dominant negative form of Ago that lacks the F-box domain, but, again, co-expression of this mutant with Cic-L did not prevent downregulation of Myc. Therefore, our results argue against a direct role of Cic-L in antagonizing or stimulating Ago function in the nurse cells, although Ago could still be part of the network controlling CycE and Myc stability, as discussed further below.

Studies in mammalian cells and *Drosophila* have shown that CycE and Myc are phosphorylated by multiple kinases that control their stability. In the case of CycE, the human protein has been shown to be phosphorylated on threonine 380 (T380) by both CDK2 and GSK3β, which triggers FBW7/Ago-mediated degradation of CYCE. In addition, this control is enhanced by further phosphorylation on serine 384 (S384) by CDK2 (Hao et al., 2007; Welcker et al., 2003). In *Drosophila,* the phosphorylation of CycE has not been studied but we note that the T380-containing phosphodegron is well conserved at the C-terminus of the fly protein (Fig. 34A), suggesting that it plays an important role in Ago-dependent regulation. In the case of Myc, phosphorylation of c-MYC on serine 62 (S62) by MAPK, c-Jun Kinase (JNK) or cyclin-dependent kinase 1 (CDK1) results in its stabilization (Vervoorts et al., 2006). This phosphorylation, however, also primes c-MYC for additional phosphorylation on threonine 58 (T58) by GSK3β (Sears et al., 2000). Then, once both sites are phosphorylated, the phosphatase PP2A can dephosphorylate S62 allowing FBW7 to bind the T58 phosphodegron (Yeh et al., 2004). On the other hand, studies in *Drosophila* have shown that Myc is phosphorylated at least by two kinases, GSK3β and Casein kinase Iα (CkIα), which induces its degradation [(Galletti et al., 2009), Fig. 34B]. Therefore, it

would be interesting to know if Cic-L affects the phosphorylation state of CycE and/or Myc in *Drosophila*. One possibility is that Cic-L regulates the phosphorylation of CycE or Myc through GSK3β. For example, Cic-L could directly interfere with the kinase activity of GSK3β, resulting in stabilization of CycE. If this is correct, then overexpression of GSK3β along with Cic-L should rescue the stabilization of CycE induced by Cic-L alone. Alternatively, Cic-L could also regulate GSK3β indirectly. In this regard, Insulin signaling has been shown to suppress the activity of GSK3β through Akt-mediated phosphorylation on serine 9 (Cross et al., 1995). Interestingly, the pattern of p-Akt accumulation (the active form of Akt) in oogenesis is very similar to the pattern of Myc (Wang et al., 2019). Thus, we could test if premature expression of Cic-L interferes with p-Akt accumulation, as a mechanism that would lead to Myc downregulation via GSK3β.

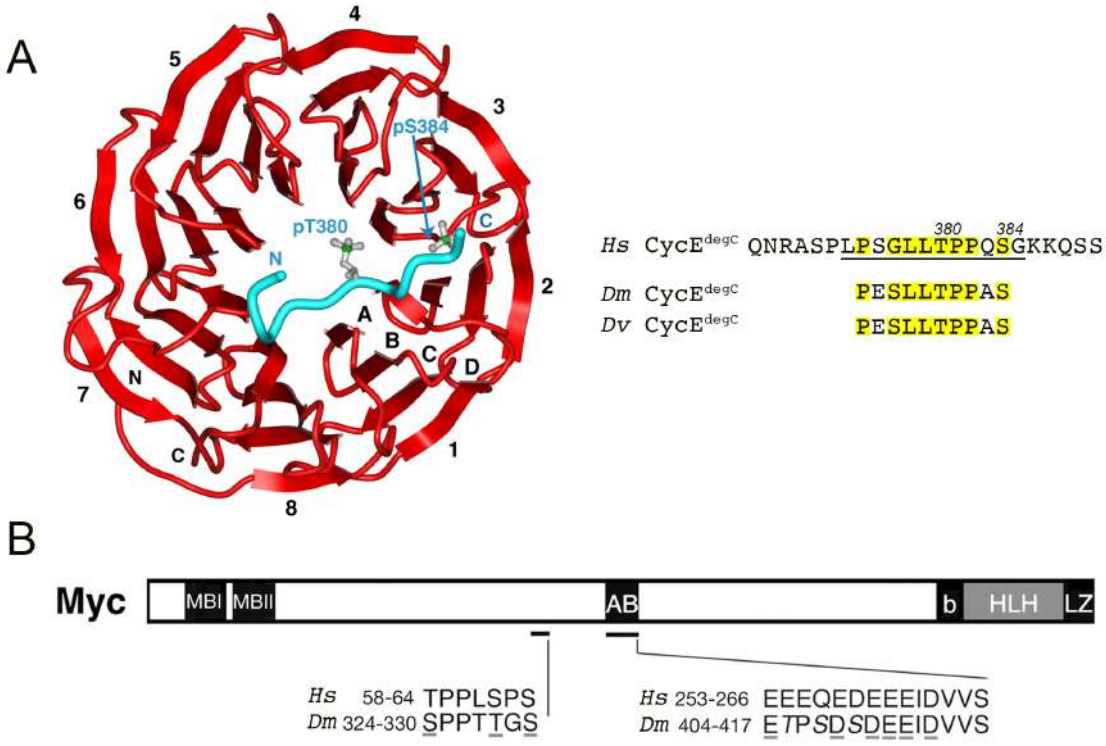

Figure 34. Postranslational regulation of CycE and Myc. (A) Structural model of the FBW7-CYCE complex. The CYCE C-terminal phosphodegron binds across the narrow face of the FBW7 ß-propeller structure. Note that the phosphodegron present at the C-terminal of human CYCE is conserved in *Drosophila*. *Hs*, *Homo sapiens*; *Dm*, *Drosophila melanogaster*; *Dv*, *Drosophila virilis*. **(B)** Structure of the Myc protein in *Drosophila* showing the conservation of phosphorylation sites. AB is an acidic region contained within the PEST domain. Mutation of the underlined residues has been shown to increase the stability of Myc in *Drosophila*. Adapted from Hao *et al*. 2007 (A), and Galletti *et al*. 2009 (B).

On the other hand, in the proteomic screen for Cic-L-associated factors, we identified another kinase that has been linked to MYC phosphorylation, casein kinase II (CKII). Importantly, we isolated the two subunits of CKII with both Cic-L and Cic-L^{Mini} (Tables 1 and 2). Nevertheless, the functional consequences of CKII-mediated phosphorylation of MYC are not well-understood. Initial studies by Lüscher *et al*. showed that CKII phosphorylates vertebrate MYC within the central acidic domain and within a region proximal to the C terminus (Lüscher et al., 1989). Subsequent studies in human cells and *Drosophila* have indicated that phosphorylation of the central acidic domain of Myc induces rapid degradation of the protein (Galletti et al., 2009; Gregory and Hann, 2000). In contrast, another study reported that pharmacologic inhibition of CKII resulted in reduced c-MYC protein levels in a lymphoma cell line (Channavajhala and Seldin, 2002). Thus, it would be interesting to clarify if CkII phosphorylates and regulates the stability of Myc in *Drosophila*. If so, we could then study if Cic-L affects Myc levels via CkII. Interestingly, Weissmann *et al*. also isolated the alpha subunit of CK2 in their proteomic screen using CIC-L as a bait (Weissmann et al., 2018). This raises the possibility that a potential mechanism of Myc regulation via Cic-L and CkII could also be conserved in humans.

Finally, we should also take into consideration that ubiquitinated proteins can be stabilized via removal of ubiquitin polypeptide chains. In fact, the ubiquitin-specific protease Puffyeye has been shown to stabilize CycE and Myc (Li et al., 2013). Perhaps Cic-L acts in a pathway that alters the ubiquitination status of CycE or Myc.

2.2 Transition from nurse cell growth to dumping and programmed cell death

Based on nurse cell function, oogenesis can be divided into two main phases: an early phase of growth (stages 2-10) and a late phase of contraction and elimination (stages10-14). As explained above, during the first phase the nurse cells undergo massive growth through endoreplication. In addition, they also synthesize RNAs and proteins that are selectively (and slowly) transported into the oocyte through the ring canals. Then, at stage 10 the nurse cells stop growing and initiate the rapid dumping of their whole cytoplasm into the oocyte. Moreover, as nurse cells dump their cytoplasmic contents, they also undergo developmental programmed cell death (PCD). During PCD, nurse cells are

engulfed by the overlying stretch cells, which through action of their lysosomal machinery at the plasma membrane promote the extracellular acidification of nurse cells and subsequent DNA fragmentation (Mondragon et al., 2019; Timmons et al., 2016).

In our studies, we have used the migration of border and centripetal cells as a developmental time reference to precisely define the temporal sequence at the transition between the two phases of nurse cell development. Specifically, we have observed that nurse cells exit the endocycle (visualized by stabilization of CycE) at stage 10A before the actin cytoskeleton of nurse cells is rearranged at stage 10B in preparation for nurse cell dumping. Furthermore, we have found that in *cic-L* mutant egg chambers, where nurse cells do not exit the endocycle on schedule, the rearrangement of the actin cytoskeleton does not take place and dumping is disrupted. In addition, our preliminary observations suggest that acidification of nurse cells by the stretch cells is also disrupted in *cic-L* egg chambers. In fact, nurse cell DNA fragmentation has already been found to be delayed or disrupted in other dumpless mutants (Bass et al., 2007; Foley and Cooley, 1998). Thus, the processes of nurse cell growth, dumping and PCD appear to be closely interrelated. Specifically, we propose that nurse cells might need to exit their endoreplicative phase before dumping can be initiated.

Moreover, this relation between nurse cell growth arrest and dumping raises the possibility that other nuclear factors that have been shown to be required for nurse cell dumping, like Rhino, Lola or Eggless (Bass et al., 2007; Clough et al., 2014; Volpe et al., 2001), actually play a role in endocycle dynamics or termination. In this regard, it is interesting to mention that *E2f1* and *dDp* mutant egg chambers exhibit both endocycle progression and dumping defects (Myster et al., 2000; Royzman et al., 2002). In addition, Royzman *et al.* detected persistent endoreplication in the midgut of *E2f1* mutant embryos (Royzman et al., 1997), suggesting the nurse cell endoreplication might also persist beyond stage 10 in *E2f1* egg chambers. Still, considering that Cic-L is induced during the final stages of nurse growth, we think that Cic-L may represent the key factor that triggers the nurse cell growth-to-dumping transition. Alternatively, Cic-L could regulate endocycle termination and nurse dumping independently. To study if persistent endoreplication is the cause of the dumping defects in *cic-L* mutant egg chambers, one could try to rescue the dumpless phenotype by artificially blocking endocycle progression at stage 10 in *cic-L* mutants. To this end, one possibility would be to overexpress CycE at stage 9-10 with the Gal4/UAS system. This is challenging though because no Gal4 lines are currently available that drive expression in the nurse cells specifically at this stage. Thus, we are generating our own Gal4 line using

the promoter of the *lrp2* gene, which is highly expressed in the nurse cells at stage 10 (Parra-Peralbo and Culi, 2011). Finally, we speculate that similar transitions could occur in other polyploid tissues like the salivary glands or the amnioserosa where the actomyosin network plays critical roles (glue secretion in salivary glands and dorsal closure in the amnioserosa) at the end of polyploidization (Hayes and Solon, 2017; Rousso et al., 2016).

2.3 Maternal-effects of Cic-L

As it has been mentioned in the introduction, *Drosophila* embryonic body axes are established during oogenesis through asymmetric localization of four maternal mRNAs: *grk*, *bcd*, *osk* and *nos*. Thus, mutations in factors required for the correct localization of these transcripts can lead to polarity defects later in the embryo. In this context, we have conducted a preliminary investigation on the role of Cic-L in the localization of *osk*, which normally accumulates at the posterior pole of the oocyte. Importantly, Osk controls germ plasm assembly and the posterior localization of *nos*, which upon translation, specifies the abdominal region of the embryo.

On the one hand, we have found that strong alleles of *cic-L* cause dramatic dumping defects that result in the production of very fragile eggs which do not develop further. On the other hand, weak alleles of *cic-L* do not exhibit obvious dumping defects but, in contrast, lead to embryonic AP polarity defects. In particular, these embryos exhibit a mirror-image duplication of posterior abdominal segments and telson that replace the anterior part, a phenotype known as bicaudal. Hence, from these observations, the establishment of embryonic AP polarity seems to be more sensitive to the loss of Cic-L function than nurse cell dumping.

Moreover, as bicaudal phenotypes arise from dysregulation of the posterior system, we have studied the distribution of *osk* transcripts in *cic-L* mutant egg chambers and progeny embryos. Our results suggest that the *cic-L* bicaudal phenotype is caused by ectopic accumulation of *osk* at the anterior pole of the oocyte during late oogenesis. These findings are in agreement with those of Dr. C. Berg and co-workers, who found that *bwk* mutations (allelic to *cic-L* as shown in results section 1.2) also cause ectopic accumulation of *osk* transcripts in anterior regions (Rittenhouse and Berg, 1995). In addition, we have noticed that persistent nurse cells in *cic-L* egg chambers accumulate *osk* transcripts, which has led us to consider the possibility that embryonic AP polarity defects might be, in fact, secondary to primary dumping defects. Specifically, we hypothesize that persistent nurse cells in late

oogenesis could represent an inappropriate source of *osk* that cannot be transported to the posterior pole of the oocyte. To investigate if dumping defects can cause ectopic accumulation of *osk*, we have examined the pattern of β-galactosidase activity of the *M1M2-LacZ-osk3'UTR* reporter in another dumpless mutant defective in *chickadee* (encoding Profilin) function. However, these mutants were not informative because the early posterior localization of *osk* is already disrupted due to premature streaming (Manseau et al., 1996). To explore this question further, we could take advantage of the MS2-MCP system to fluorescently label *osk* mRNA particles. This system is based on the insertion of multiple binding sites for the MS2 coat protein (MCP) in the transcript of interest in parallel with co-expression of the MCP protein fused to a fluorescent protein. Importantly, this system has already been adapted to follow *osk* mRNA particles in living oocytes (Zimyanin et al., 2008), and we could use it to study if *osk* mRNA particles are transported from persisting nurse cells into the oocyte at late stages of oogenesis.

Finally, I would also like to mention that, although ectopic *osk* accumulation has also been associated with the generation of ectopic pole cells (Ephrussi and Lehmann, 1992), embryos deposited by *cic-L* females do not present ectopic pole cells. This observation can be explained by previous *osk* overexpression studies showing that the threshold of *osk* required for abdomen formation is lower that the threshold needed for pole cell formation (Smith et al., 1992). Therefore, the level of *osk* present at the anterior pole of embryos laid by *cic-L* females is probably sufficient to drive the formation of an abdomen, but not pole cells. Alternatively, as most of these embryos cannot be fixed due to their fragility, the lack of ectopic pole cells could be explained by a technical bias towards the embryos with lower expressivity of the phenotype.

3. Evolutionary perspective and open questions

In this last section, I discuss our findings from a broader perspective that considers some of the emerging questions and future challenges regarding Cic-L function, both in *Drosophila* and in other organisms. Because at least some of these ideas have an evolutionary component, I will discuss them in a temporal sequence that spans from the origin of Cic-L-like proteins in early metazoans through their potential functions in mammals and human disease.

We have seen that Cic-L-related proteins are already present in sponges and placozoans (but not in yeast or plants), which are considered two of the earliest extant branches of

multicellular animals. These proteins consist of an N1-like motif without additional recognizable domains, making it likely that they evolved before the standard, HMG-box-containing Cic-L proteins present from cnidarians to vertebrates. Being Cic-L such a complex modular protein, the simple structure of those ancient N1 proteins seems rather surprising. Moreover, it gives us no clues as to their potential molecular functions in those organisms, let alone the processes they could regulate. In this regard, a survey of the literature reveals very few references to the presence of polyploidy cells in either sponges or placozoans (Birstein, 1989), thus preventing speculation on a possible ancient role of ancestral N1 proteins in endocycle control. It is nevertheless intriguing that a *Drosophila* Cic-L N-terminal fragment carrying a mutated Tudor domain has significant growth-arresting activity in the nurse cells (our unpublished results), raising the possibility that the N1 domain alone may be sufficient to regulate some conserved aspects of cell cycle progression in these different species.

Reconstructing how the ancient N1 proteins eventually acquired additional Cic-L domains, and the biological significance of this process, is also beyond our current insight. Protein evolution through fusion, recombination and loss of functional domain is a well-recognized mechanism for the expansion and diversification of protein functions –a mechanism that re-uses existing domains without the need to invent new ones. The emerging features of the resulting proteins may be completely novel and not necessarily reflect the sum of the individual elements [see, for example, (Lees et al., 2016)]. Therefore, we cannot even speculate what new properties emerged after the fusion of the N1 domain to the HMG-box and C1 domains, when a full Cic-L structure was in place. Nor do we know why the resulting proteins became, presumably at approximately the same time, MAPK-binding proteins regulated by RTK signaling. This is particularly puzzling in the case of *Drosophila* Cic-L, which we have shown is clearly sensitive to RTK regulation when acting together with Cic-S but has not revealed a response to such regulation when performing its unique functions. Thus, the N-terminal region of Cic-L alone, separated from the C2 motif and other potential C-terminal docking sites (Astigarraga et al., 2007; Futran et al., 2015), significantly rescues the loss of *cic-L* function in oogenesis, and we also note that the cic^{3} allele, which carries a partially inactivating mutation in C2, does not affect the growth or dumping of the nurse cells (our unpublished results). Related to this question, we have observed significant levels of MAPK activation in nurse cells at stage 10. Perhaps MAPK signals in this or other contexts do have an effect on Cic-L that we have been unable

to see in our experiments. We believe this question deserves further study once the mechanisms of Cic-L activity are better understood.

Concerning CIC-L functions in mammals, the first consideration to be made is the strong structural conservation of this isoform relative to *Drosophila* Cic-L. The N-terminal region contains Tudor-like and N1 domains in the same arrangement as *Drosophila*. This suggests that mammalian and fly CIC-L proteins are likely to function similarly at the molecular level, exerting both redundant and unique functions as described here for the latter. In principle, the potential redundant functions of mammalian CIC-L may be relatively common and linked to the functions that have been characterized in connection to RTK signaling and various disease mechanisms. In support of this idea, different expression studies in mice have revealed similar distributions of CIC-L and CIC-S in multiple tissues and organs, including the brain, the lung, the stomach and the thymus, which suggest a high degree of functional redundancy for both isoforms across those organs and tissues.

As for the unique functions of CIC-L, they are more difficult to assess because no previous studies have investigated the unique requirements of this isoform. Interestingly, this has almost been done by the group of Dr. H. Zoghbi, since they have generated *Cic-L*$^{-/-}$ mutant mice by targeting the *Cic* locus with a β-geo genetrap cassette [(Fryer et al., 2011; Lee et al., 2011), see Fig.2A]. This cassette was inserted immediately downstream of exon 1 of *Cic-L*, which encodes the entire Cic-L-specific N-terminal region, trapping this isoform in homozygous *Cic-L*$^{-/-}$ knockout mice. Two remarks can be made about this design. First, this strategy might have been selected not to study CIC-L in particular but because it would offer the possibility to visualize gene expression from the *Cic* locus via β-galactosidase staining. Second, although their approach should have generated an allele specifically inactivating CIC-L, two subsequent observations dismiss this possibility: (i) these authors found that their mutant also reduced *Cic-S* expression by 90%, and (ii) our finding that the N-terminal region of *Drosophila* Cic-L is largely capable of mediating Cic-L-specific activity in oogenesis raises the possibility that mouse CIC-L trapped by the cassette might still retain similar activities. Therefore, their allele is a strong *Cic* hypomorph that has proved very useful for the study of CIC function in mice (Fryer et al., 2011; Kim et al., 2015; Lee et al., 2011), but cannot inform whether the observed phenotypes reflect a CIC-L specific requirement, a partial loss of CIC-S, or a reduction of total CIC activity (Kim et al., 2015).

Thus, there is no genetic evidence regarding the potential unique functions of CIC-L in mammalian development, and we can only speculate on this point. One idea is that CIC-

L could play a role in regulating endoreplication in mammals. While polyploidization is less common in mammals than in insects or plants, it does normally occur during differentiation of several tissues, including the heart, the placenta, the bone marrow and the liver. These polyploid states arise through different mechanism, with endoreplication being just one of them (Donne et al., 2020; Lee et al., 2009). One of the best-characterized examples of endoreplication takes place in extraembryonic trophoblast giant cells. This polyploidization, whose physiological significance remains somewhat controversial (Chen et al., 2012; Garcí-Higuera et al., 2008; Hu and Cross, 2010), is, as in *Drosophila*, regulated by CYCE and E2F transcription factors (Chen et al., 2012; Geng et al., 2003; Ouseph et al., 2012). Moreover, CIC protein expression is present in trophoblast cells of the placenta (The Human Protein Atlas, www. proteinatlas.org).

One aspect of endoreplication, however, that differs between *Drosophila* and mammals relates to the role of Myc in these processes. In *Drosophila*, Myc plays a major role in growth control by accelerating mass accumulation and is thus critical for endocycle progression (Gallant, 2013; Maines et al., 2004; Pierce et al., 2004). In contrast, the main effect of mammalian c-MYC on body size takes place at the level of cell number, not cell size (Trumpp et al., 2001). This makes it unlikely that CIC-L might affect endocycle function in mammals through a mechanism involving downregulation of MYC, as we see it in the nurse cells. Still, it is conceivable that CIC-L could function in a conserved pathway that affects c-MYC function in other cellular contexts.

Finally, it seems likely that loss or altered function of CIC-L is associated with cancer and other human diseases. First, CIC-L is likely to contribute, along with CIC-S, to the known transcriptional repressor functions of CIC in disease-related processes. This idea is supported by our finding that *Drosophila* Cic-L and Cic-S act redundantly in at least two different developmental contexts, and the similar distribution patterns of both mammalian isoforms across multiple tissues and organs. If this were correct, then most of the cancer driver mutations that map to the common region of CIC-L and CIC-S (for example inactivating the HMG-box in oligodendroglioma tumors) would lead to derepression of CIC targets that are regulated jointly by both isoforms. Similarly, both isoforms share an ATXN1binding motif and are bound to this protein in the cerebellum, implying that polyglutamine-expanded ATXN1 causes SCA1 via aberrant activities in complex with both CIC-L and CIC-S. Nevertheless, as discussed in section 1.2, we have not yet determined if Cic redundant functions in *Drosophila* reflect a single mechanism of repression mediated by the common region of Cic-L and Cic-S or result from two distinct mechanisms requiring

the specific N-terminal portions of each isoform. In this respect, we have observed that vertebrate CIC-S proteins share a highly conserved N-terminal motif that straddles the unique and common sequences of CIC-S. This leaves open the possibility that human CIC-S, despite having only 23 amino acids that are not present in CIC-L, may have unique molecular activities not shared with CIC-L and, therefore, that CIC function connected to known diseases might depend, totally or to a large extent, on just one of the two isoforms.

Less clear is whether CIC-L, by virtue of its potential unique activities related to those described here for *Drosophila* Cic-L, could play any role in disease. To begin with, there are no known mutations in the CIC-L N-terminal region with a proven or likely effect on human pathologies. Intriguingly, Simón-Carrasco *et al*. (Simón-Carrasco et al., 2017) did report two such mutations in T-ALL patient samples, but the two residues affected, R69 and T193, are poorly conserved outside mammals and do not map to any of the conserved functional domains we have characterized in Cic-L. Still, if we assume the CIC-L N-terminal region functions in human endoreplication, then it is worth noting that endoreplication and polyploidization have been implicated in multiple steps and mechanisms of tumorigenesis. Polyploidy is a relatively common feature of cancer cells, which can additionally follow different pathways to cause other oncogenic abnormalities such as genome instability and aneuploidy [reviewed in (Shu et al., 2018; Storchova and Pellman, 2004). Hence, although we are not aware of any study linking a failure in endocycle exit to cancer, a putative role of CIC-L in such a mechanism seems compatible with its possible involvement in a cancer-related process.

In conclusion, Cic-L is a complex molecule with a sort of double life. It acts redundantly with Cic-S in certain contexts and also has unique, more ancient functions that appear to be rooted at the origin of metazoans. The biggest challenge arising from our results is to decipher these functions at the molecular level and to understand their potential impacts on human health, be they related to human endocycle control, MYC regulation or possibly other mechanisms yet to be discovered.

CONCLUSIONS

1. Cic-L and Cic-S display both overlapping and isoform-specific expression patterns through *Drosophila* development.
2. Cic-L and Cic-S act redundantly in different processes regulated by RTK signaling, such as the establishment of embryonic dorsoventral patterning and the specification of wing veins.
3. In addition, just as Cic-S has unique functions in *Drosophila* development, Cic-L also exerts individual functions, in oogenesis and probably in other uncharacterized processes required for viability.
4. *cic-L* corresponds to the *bullwinkle* locus identified by Dr. C. Berg and co-workers in 1995.
5. Cic-L is essential for nurse cell dumping during stage 10B of oogenesis, being specifically required for the assembly of the actin cytoskeleton involved in this process.
6. CycE becomes stabilized in nurse cells at stage 10A, implying that they stop endocycling before they proceed to cytoplasmic dumping.
7. Cic-L signals endocycle exit by the nurse cells, being necessary and sufficient for this developmental switch.
8. Cic-L acts in this process by promoting both stabilization of CycE and downregulation of Myc.
9. Cic-L-specific functions are mainly mediated by its unique N-terminal region and are thus largely independent of the HMG-box and other domains required for Cic-S activity.
10. Correspondingly, truncated polypeptides resembling the Cic-L N-terminal region are found in two species lying at the base of metazoans, the sponge *Amphimedon queeslandica* and the placozoan *Trichoplax adhaerens*, suggesting that this region represents an ancient functional module.
11. The N-terminal region of Cic-L contains three conserved domains, NLS, Tudor-like and N1, that function additively to mediate Cic-L-specific functions in oogenesis.
12. Cic-L potentially interacts in S2 cells with a set of factors specifically associated to this isoform.

MATERIALS AND METHODS

1. Synthetic DNA constructs

1.1 General considerations

Plasmid manipulations were carried out following standard procedures in molecular biology such as amplification of DNA sequences by PCR (Polymerase Chain Reaction), digestion with restriction enzymes, dephosphorylation with Alkaline phosphatase, purification from agarose gel bands (illustra GFX PCR DNA and Gel Band Purification Kit, GE Healthcare) and ligation of DNA fragments. Ligation reactions were transformed into DH5α *Escherichia coli* cells via heat shock. After a recovery period in LB medium without antibiotic, transformed cells were spread on antibiotic selection plates (Ampicillin 100 µg/ml) and allowed to grow overnight. Bacterial colonies were inoculated in LB medium supplemented with 100 µg/ml Ampicillin (3ml for minipreps and 100ml for midipreps) and incubated overnight at 37ºC under strong agitation. Plasmid DNA was purified with commercial kits following the manufacturer's instructions, NucleoSpin Plasmid Mini (MACHERY-NAGEL) for minipreps and Qiafilter Plasmid Midi (Qiagen) for midipreps.

1.2 Transgenic and CRISPR/Cas9 constructs

cic-L: To generate the *cic-L* rescue construct, four different fragments were assembled in the pattB vector. In an upstream to downstream order:

1. A BamHI/BamHI genomic fragment containing nucleotides -2403 to +1084 (counting from the transcription initiation site published for *cic-RD* and *cic-RG* in Flybase).
2. A BamHI/HindIII genomic fragment that includes part of the *cic-L* 5'UTR beginning at position +38,183 and residues 1-380 of Cic-L. We did not include a 34Kb intron in the 5' UTR of *cic-L*.
3. A HindIII/Acc65I fragment from cDNA LD17181 ((GenBank accession number BT100233) encoding amino acids 381-818 of Cic-L.
4. An Acc65I/XbaI fragment from a genomic *cic-S* rescue transgene (Cinnamon et al., 2004) which includes residues 21-1397 of Cic-S (819-2135 of Cic-L) and the natural 3' regulatory sequences of *cic*. It also has three tandem copies of the hemagglutinin (HA) tag (YPYDVPDYA) inserted at position 1398 of Cic-S (amino acid 2136 of Cic-L).

cic-L-^{Nter}: It is a derivative of the previous pattB *cic-L*-HA vector in which the Acc65I/XbaI fragment has been substituted by a fragment that only contains the triple HA tag and the 3'UTR of *cic*.

UAS-cic-L: This construct was assembled in the *pUASp* vector, which contains the regulatory regions necessary for gene expression in the female germline (Rørth, 1998). A BamHI/Acc65I fragment encoding residues 1-818 of Cic-L was isolated from the *cic-L-HA* rescue construct and cloned into a *pUASp* vector in which we had previously modified the multicloning site by cloning an adaptor. We digested the resulting vector Acc65/XbaI and cloned an Acc65I/XbaI fragment from the *pUAStattBcicWT3xHA* construct (Andreu et al., 2012a), which encodes amino acids 819-2135 of Cic-L and a triple HA tag.

***UAS-cic-L$^{\Delta C2}$*:** It is a derivative of the *UAS-cic-L* construct in which residues 1790-1810 of Cic-L (corresponding to the C2 domain) have been deleted.

***UAS-cic-L$^{N\text{-}ter}$*:** This construct was made by replacing residues 819-1236 of the *UAS-cic-L-HA* construct with an Acc65I/XbaI-digested PCR fragment encoding three tandem copies of the HA tag.

***ago*:** We amplified by PCR a genomic fragment of 4.4 Kb that includes *ago* 5' regulatory sequences and the sequence encoding amino acids 1-482. This fragment was digested with Acc65I and NotI, and suncloned in Acc65I/NotI-digested *pCaSpeR4*. A second PCR fragment containing amino acids 483-1326 and the 3' regulatory regions was digested NotI/SalI and cloned into the previous plasmid digested NotI/XhoI. Finally, the resulting plasmid was digested NotI to insert a GFP at position 483.

***UAS-ago*:** The coding region of *ago* was amplified by PCR from the *pCaSpeR4-ago* plasmid and cloned as an Acc65I/XbaI fragment into the *pUASp* vector. Instead of a GFP, this construct has a doble V5 tag (GKPIPNPLLGLDST) inserted at position 483.

***UAS-ago$^{\Delta Fbox}$*:** It is a derivative of the *pUASp-ago* construct but has been assembled in the *pUASz1.0 vector*, which allows efficient expression in both the germline and the soma (Deluca and Spradling, 2018). In this construct we deleted amino acids 878-937 of Ago by inverse PCR (corresponding to the F-box).

CRISPR/Cas9 gRNA plasmids: To generate gRNA expression constructs, we used the *pCDF3* (Port et al., 2014) and the *pBFvU6B* (Kondo and Ueda, 2013) vectors. These

plasmids contain the promoter of the U6:3 spliceosomal snRNA gene that drives expression of the gRNAs and a bacterial attachment sequence (attB) for ΦC31 integrase-mediated recombination.

For single gRNAs, complementary oligonucleotides were annealed overnight and cloned into BbsI-digested *pCFD3*. We used the following complementary oligonucleotides:

gRNA *cic-L*	5'GTCGAAGGGAGACATCGTGGAGT 3'	5' AAACACTCCACGATGTCTCCCTT 3'
gRNA *cic-L-HA*	5' GTCGGATCGGAGCAAGTGGTCC 3'	5' AAACGGACCACTTGCTC CGATC 3'

For the *cic*^{8} allele we used a transgenic line that expressing a gRNA that targets *cic-S* specifically, which had been previously used to isolate the *cic*^{5} allele (Papagianni et al., 2018).

For double gRNAs, two different gRNAs were cloned into the *pBFv-U6.2B* vector, which allows the expression each gRNAs from its own U6 promoter. Initially, the two gRNAs were cloned in separate vectors. For the first gRNA, complementary oligonucleotides were annealed overnight and cloned into BbsI-digested *pBFv-U6.2*. The same procedure was followed for the second gRNA, but it was cloned into BbsI-digested *pBFv-U6.2B*. Next, an EcoRI/NotI fragment containing the U6 promoter and the first gRNA was excised from *the pBFv-U6.2-gRNA#1* plasmid and ligated with the *pBFv-U6.2B-gRNA#2* plasmid to generate the double gRNA vector.

We used the following complementary oligonucleotides:

gRNA#1 NLS	5' CTTCGGGATGCTGCCGCGGCTG 3'	5' AAACCAGCCGCGGCAGCATCCC 3'
gRNA#2 NLS	5' CTTCGGTCTTGTCCAGCTCGGC 3'	5' AAACGCCGAGCTGGACAAGACC 3'
gRNA#1 Tud	5' CTTCGCCCTGCTCGATCTCAGCGAG 3'	5' AAACCTCGCTGAGATCGAGCAGGGC
gRNA#2 Tud	5' CTTCGTCCCACCAAGGCGGCTGC 3'	5' AAACGCAGCCGCCTTGGTGGGAC 3'
gRNA#1 N1	5' CTTCGAAGGGAGACATCGTGGAGT 3'	5' AAACACTCCACGATGTCTCCCTTC 3'
gRNA#2 N1	5' CTTCGGTGTTGCCCTTCTGGTTG 3'	5' AAACCAACCAGAAGGGCAACACC 3'
gRNA#1 SLSA	5' CTTCGCCACTTCGGCTGCCAAG 3'	5' AAACCTTGGCAGCCGAAGTGGC 3'
gRNA#2 SLSA	5' CTTCGCTGCTGCTGTTGCTGCA 3'	5' AAACTGCAGCAACAGCAGCAGC 3'
gRNA#2AAAM	5' CTTCGATCAGGAGGAGACCGATG 3'	5' AAACGATCGGTCTCCTCCTGATC 3'

Moreover, we exploited homology directed repair to tag the endogenous locus of *cic-L* with a triple HA and to generate the *cic*^{9} allele (see section 4.1), and for that we assembled two

donor plasmids: *cic-L-HA* and *cic9*. These plasmids were assembled in a pUC vector and contain a *3xP3-DsRed* marker flanked by loxP recombination sites for its removal, and a triple HA tag. In addition, each plasmid includes 2 specific fragments of about 1Kb homologous to the regions flanking the cleavage site:

***cic-L-HA* donor:** The left homology arm spans nucleotides 36,141-37,122 of the *cic* locus described in Flybase, whereas the right homology arm spans nucleotides 37,123-38,309.

***cic9 donor*:** The left homology arm corresponds to nucleotides 34,685-35,659 of the *cic* locus and the right homology arm to nucleotides 38,300-39,343.

1.3 S2 cell expression constructs

Cic-L-full-length and Cic-L-mini were cloned into the pMK33 vector for the establishment of stable S2 cell lines (Kyriakakis et al., 2008). This vector contains a metallothionein promoter for copper-inducible expression of the tagged protein, and a hygromycin resistance cassette to facilitate the generation of stable cell lines using hygromycin selection.

***cic-L-SBP*:** This construct is based on the *pMK33-Cic-SBP* plasmid generated by Yang *et al.* (Yang et al., 2016). We digested this plasmid XhoI-Acc65I to replace amino acids 1-79 of Cic-S by amino acids 1-817 of Cic-L.

***cic-L^{Mini}-SBP*:** This construct is a derivative of the *cic-L-SBP* constructs in which amino acids 1315-2026 of Cic-L have been deleted (this includes the C2 domain).

2. Synthesis and labeling of antisense RNA probes

In situ hybridizations were carried out using digoxigenin-UTP labeled antisense RNA probes. To generate the antisense RNA probes, a cDNA fragment of the target gene was cloned into a pBluescript vector, which has T7 and T3 promoters for in vitro transcription. The plasmid containing the cDNA fragment was then linearized with a restriction enzyme. Linearized DNA was transcribed *in vitro* with T7 or T3 RNA polymerase by incubating the following mix for 2h at 37ºC:

5µl (1µg) Linearized template

1 µl Transcription buffer 10X

1 µl Digoxigenin UTP mix 10X (Roche)

1µl (50u/µl)RNase inhibitor

1 µl T7 or T3 RNA polymerase (Roche)

1 µl H_2O

After the 2 hours, we added 17µl of H_2O to the mix and 2 µl were run in an agarose to confirm transcription. Next, we added 15µl of carbonate buffer 2X to induce a partial degradation of the RNA. This mix was incubated at 62ºC for a period of time between 15 and 40 minutes depending on the size of the cDNA fragment. To stop this reaction and to precipitate the RNA we added: 50µl of stop solution [0.2M NaAc (pH 6)], 10µl of LiCl 4M, 5µl of tRNA (carrier) 20mg/ml, and 300µl of ethanol. This mix was incubated at -20ºC for at least 15 minutes and centrifuged at 13000 rpm for 20 minutes at 4ºC. We eliminated the supernatant, and the pellet was washed with ethanol 70%. Finally, we resuspended the RNA in 150µl of hybridization solution.

3. *Drosophila* stocks and transgenic lines

3.1 Fly culture and husbandry

Stocks were maintained at 18ºC in vials containing baker's yeast paste. Mating and rescue experiments were done at 25ºC.

3.2 Transgenic fly generation through germline transformation

Transgenic *Drosophila* lines were generated by germline transformation with the P-element and ϕC31 phage integrase systems (Bischof et al., 2007). The *white*$^+$ and *vermillion*$^+$ marker genes were used to screen for positive transgenic lines. Specifically, *UAS-cic-L*, *UAS-cic-L*$^{\Delta C2}$, *UAS-cic-L*^{Nter}, *ago* and *UAS-ago* transgenic lines were obtained by standard P-element transformation. In general, at least two independent insertions were examined for each construct. The *cic-L*, *cic-L*^{Nter} and *UAS-ago*$^{\Delta Fbox}$ transgenes were inserted at position 86F; and gRNAs plasmids were inserted at position 25C.

3.3 Establishment of transgenic lines and mapping of transgenes

Adults from the F0 generation were individually mated with adults of the same mutant injection strain (*white* or *vermillion*) and the offspring were screened for transformants by eye color rescue. The transgenic individuals recovered (F1) were crossed again individually with the mutant strain of injection. In the next generation, transgenic heterozygous females and males were crossed between them. Finally, homozygous progeny from this cross was selected to establish the transgenic line. Moreover, for transgenic lines generated via P-element transformation, transgenes were mapped to determine the chromosome of insertion. Typically, transgenes land on chromosomes X, II or III, since the IV chromosome is rather small and essentially heterochromatic. Transgenes were mapped through segregation analysis using the balancer chromosomes SM6a (for chromosome II) and TM3, Sb (for chromosome III).

3.4 *Cic* alleles and other stocks

Allele	**Description**	**Reference**
cic^{fetE11}	Insertion of a P element in the common region	(Goff et al., 2001)
cic^{4}	Deletion of 4 amino acids in the C1 domain	(Forés et al., 2017b)
cic^{5}	Frameshift deletion in the first exon of *cic-S*	(Papagianni et al., 2018)
cic^{7}	Frameshift deletion in exon 6 of *cic-L-RD*	This thesis
cic^{8}	Compound allele carrying the cic^{7} mutation and a frameshift mutation equivalent to cic^{5}	This thesis
cic^{9}	Replacement of the Cic-L-specific N-terminal region by an in-frame triple HA tag	This thesis
cic^{10}	Compound allele carrying the cic^{9} deletion and a frameshift mutation upstream of the HMG-box	This thesis
$cic^{\Delta NLS}$	In-frame deletion of a potential NLS generated in a *cic-L-HA* chromosome	This thesis
$cic^{\Delta Tud}$	In-frame deletion of the Tudor-like domain	This thesis
$cic^{\Delta N1}$	In-frame deletion of the N1 domain	This thesis
$cic^{\Delta Tud\Delta N1}$	Double deletion of the Tudor-like and N1 domains	This thesis

Other stocks used in this thesis: *cic-S-Venus* (Grimm et al., 2012), *mirr^{F7}-lacZ* (McNeill et al., 1997), *tsl(B)-lacZ* (Furriols et al., 2007), *bwk^{8482}* (Rittenhouse and Berg, 1995), *M1M2-LacZ-osk3'UTR* (Gunkel et al., 1998), *P{αTub84B(FRT.Gal80.y^{+})Gal4.C}* (Zecca and Struhl, 2002), and *mat-tubulin-Gal4* (Bloomington 7063).

4. GENETIC ANALYSES

4.1 CRISPR (Clustered Regular Interspaced Palindromic Repeat)/Cas9 (CRISPR associated) system

In this thesis we have engineered several genome modifications of the *cic* locus using the versatile and cost-effective CRISPR/Cas9 technology. The CRISPR/Cas9 technology facilitates precise genome editing through the generation of double-strand breaks (DSBs) at selected sites in the genome. DSBs are then resolved by two main repair pathways, which can be harnessed for genome editing. The most frequent pathway, nonhomologous end joining (NHEJ), is an error-prone ligation process that often results in the generation of small mutagenic insertions and/or deletions (InDels) at cleavage sites. By targeting open reading frames, this pathway can be used to disrupt genes through frameshift mutations or to produce in-frame deletions of specific amino acids. The second pathway, homology-directed repair (HDR), restores the break by precisely copying a template sequence that bears homology across the DSB site. This pathway can be exploited to precisely edit genomic sequences or insert exogenous DNA (for example a tag) by supplying a donor repair template.

The CRISPR/Cas9 technology is based on the endogenous adaptive immune system of *Streptococcus pyogenes*, which has been simplified to two components for its use in genome engineering: the RNA-guided endonuclease Cas9, and a synthetic guide RNA (gRNA) that determines the target specificity of the Cas9 cleavage. gRNAs can be easily synthesized to recognize a ~20-nucleotide target sequence (protospacer), and the only requirement when selecting a target sequence is the presence of a NGG protospacer adjacent motif (PAM) immediately 3' of the protospacer. Multiple groups have adapted this technology for genome editing in *Drosophila* using slightly different approaches. In general, for highly efficient germline edition we have generated transgenic lines of the different gRNAs and crossed them to a transgenic line that expresses the Cas9 nuclease in the germline under the promoter of *nos* (Kondo and Ueda, 2013). Specifically, to generate

small mutagenic InDels we followed the protocol described by Forés *et al*. (Forés et al., 2017a). For bigger deletions, we proceeded in a similar way, but we expressed two different gRNAs flanking the sequence that we aimed to delete. To introduce the triple HA tag in the *cic-L* locus and to replace the Cic-L-specific N-terminal sequence by a triple HA we have exploited HDR repair. In these cases, we followed the protocol described by Gratz *et al*. (Gratz et al., 2015). Briefly, Cas9 females were crossed to transgenic males expressing the gRNA and progeny embryos were injected with a donor plasmid containing ~1Kb homology arms, the modification of interest and a red fluorescent marker. Recombination events were then identified by red fluorescent eyes and confirmed by PCR and sequencing. Finally, the red fluorescent marker was removed by Cre-mediated recombination.

4.2 Generation of Germline mutant Clones (GLC)

To study if Cic-L is required in the female germline, we have used the Dominant Female Sterile (DFS) technique combined with the Flp/FRT system to generate Germline Clones (GLCs) of the cic^{7} allele (Chou and Perrimon, 1996; Chou et al., 1993). On the one hand, the DFS technique takes advantage of the ovo^{D1} mutation, which is a germline-specific dominant mutation that arrests oogenesis at stage 6. On the other hand, the Flp/FRT system is used to catalyze site-specific mitotic recombination between two homologous chromosomes at specific FRT sites. In brief, hsFLP; FRT[82B] ovo^{D1} males were crossed with FRT[82B] cic^{7} females. To induce the expression of the Flp recombinase, the progeny of this cross was heat shocked for 45 minutes at 37ºC twice when they reached late L2 to L3 larval stages. Then, we selected FRT[82B]ovo^{D1}/ FRT[82B]cic^{7} transheterozygous females in which mitotic recombination catalyzed by Flp might have occurred in the germline. Egg chambers in which recombination has not occurred, remain transheterozygous for the ovo^{D1} and do not progress beyond stage 6. Thus, egg chambers that develop past stage 6 are the ones in which the germ cells have recombined out the copy of *ovoD1*, meaning the germ cells homozygous for the cic^{7} mutation. As a negative control, we used adult females of the same genotype that had not received the heat-shock treatment and they were sterile.

4.3 Ectopic gene expression with the Gal4/UAS system

We have used the Gal4/UAS system to drive ectopic gene expression in tissue-specific patterns (Duffy, 2002). The Gal4/UAS system is based on the ability of the yeast transcriptional activator Gal4 to bind specific upstream activation sequences (UASs) and

activate transcription of a downstream gene. In *Drosophila*, the Gal4/UAS system has been adapted as a bipartite approach consisting of two transgenic lines: a "driver" line, which expresses a Gal4-encoding transgene under the control of a tissue-specific promoter and, a Gal4-responsive "UAS" line containing the gene of intertest. Thus, when the two lines are crossed, the Gal4 protein activates transcription of the *UAS* construct in a transcriptional pattern that reflects the expression pattern of the Gal4.

4.4 Mosaic analysis with a repressible cell marker (MARCM)

The MARCM method combines the FLP/FRT system, the GAL4/UAS system and its repressor GAL80 to generate labelled cell clones (Lee and Luo, 2001). In our case, we have used it to overexpress *UAS-cic-L* and *UAS-cic-L*$^{\Delta C2}$ in follicle cell clones. Before recombination, follicle cells are heterozygous for the *UAS* construct of interest, a *UAS-GFP* construct, and the *Gal4* and *Gal80* transgenes. In this situation the Gal80 inhibits the activity of the Gal4 and the *UAS* constructs are not expressed. Following Flp/FRT-mediated mitotic recombination, the transgene encoding the Gal80 is removed from one of the daughter cells and this allows Gal4-driven expression of GFP as well as Cic-L or Cic-L$^{\Delta C2}$ in this daughter cell and its progeny.

5. Ovary analyses

5.1 Ovary dissection and fixation

To maximize egg chamber production, 1-day old female flies (10-15 individuals) were transferred into fresh fly food vials with several males and dry yeast. Vials were kept at 25ºC for 36-48 prior to ovary dissection. For ovary dissection, a pair of forceps was used to grasp the female between the abdomen and the thorax. Using a second pair of forceps (Dumont #5), the abdominal cuticle was removed, and the ovary pair was transferred to cold PBS. Once all females were dissected, debris were removed and the anterior halves of ovarioles were teased apart using a sharp tungsten needle. Ovaries were fixed in 4% formaldehyde-PBS for 10 minutes. After fixation, ovaries were rinsed 3 times with PBT 0.1% (1X PBS with 0.1% Triton X-100). Egg chambers were disaggregated by pipetting.

5.2 *In situ* mRNA hybridization of ovaries

Fixed ovaries were washed 3 times with PBT 0.1% under rotation for 10 minutes each wash. Next, ovaries were incubated for 2 minutes in PBT: hybridization solution (1:1) (50% deionized formamide, 5X SCC, 50µg/ml heparin, 1% Tween 20 and 1 µg/ml sonicated salmon sperm DNA) and prehybridized in hybridization solution for 2 hours at 55ºC.

The digoxigenin-UTP labelled antisense RNA probe was diluted in hybridization solution (0.5µl of the probe in 60µl of hybridization solution). Before addition, the probe was boiled for 5 minutes to break down secondary structures and rapidly chilled on ice. Ovaries were incubated with the probe overnight at 55ºC.

The next day, the probe was removed, and ovaries were washed with new hybridization solution, PBT: hybridization solution (1:1) and finally 3 times with PBT 0.1% (20 minutes each wash). Ovaries were then incubated with anti-Digoxigenin antibody conjugated with Alkaline Phosphatase (Roche) diluted 1:2000 in PBT at room temperature for 2 hours. To avoid non-specific signal, the α-digoxigenin antibody was previously precleared by incubating with w^{1118} embryos overnight. After the 2 hours, the antibody was removed, and ovaries were washed with PBT 0.1% every 15-20 minutes for 1 hour. Ovaries were transferred to a multiwell plate and incubated in Alkaline Phosphatase buffer (100 mM NaC1,50 mM MgC12, 100 mM Tris pH 9.5, 1 mM, 0.1%Tween 20) containing 4.5 mg/ml NBT BCIP until an intense purple color was developed. Phosphatase alkaline reaction was stopped by washing with PBT several times and ovaries were mounted in 70% glycerol or Fluoromount-G (Southern Biotech). Brightfield images were obtained with a Nikon eclipse 80i microscope.

5.3 Ovary immunostaining

Fixed ovaries were washed once in PBT 0.1% for 15 minutes and blocked with PBT-BSA 0.3% (PBT 0.1% with 0.3% BSA) under rotation at 4ºC for at least 30 minutes. Ovaries were then incubated with primary antibodies diluted in PBT-BSA 0.3% while rotating overnight at 4ºC. The next day, ovaries were washed 4 times for 15 minutes each with PBT-BSA 0.3% and incubated with fluorescently conjugated secondary antibodies in PBT-BSA 0.3% for 2 hours at 4ºC (after the secondary antibody was added, ovaries were protected from light using aluminum foil). If necessary, rhodamine-phalloidin (Sigma-Aldrich) was added during the last hour of secondary antibody incubation and DAPI during the last 10

minutes. After secondary antibody incubation, ovaries were washed with PBT 0.1% every 15-20 minutes for 90 minutes and mounted in Fluoromount-G (southern Biotech). Confocal images were obtained with a Leica TCS SP5 microscope.

For ovary immunostainings we used the following primary antibodies: rat anti-HA (Sigma, dilution 1:250 dilution), rabbit anti-GFP (Thermofisher, 1:500 dilution), mouse anti-V5 (Thermofisher, 1:1000), mouse anti-β-galactosidase (40-1a, Developmental Studies Hybridoma Bank, 1:200), mouse anti-Myc (P4C4-B10, Developmental Studies Hybridoma Bank, 1:500), anti-rabbit anti-CycE [(Richardson et al., 1995), dilution 1:50].

5.4 X-Gal staining of ovaries

Ovaries were dissected in cold PBS and fixed in 4% formaldehyde-phosphate-citrate buffer (Na_2HPO_4 0.2M/citric acid 0.1M/dH_2O (9:1:10)) for 12 minutes. After several rinses with phosphate-citrate buffer, ovaries were transferred to a glass well with incubation buffer (phosphate-citrate buffer containing 5mM each of potassium-ferri and ferro-cyanides and 0.02% Triton X-100) and a saturating amount of X-gal powder was added. The sample was incubated at 30ºC until an intense blue color was developed. Ovaries were washed several times with phosphate-citrate buffer and mounted in 70% glycerol or Fluoromount-G (Southern Biotech). Brightfield imagen were obtained with a Nikon eclipse 80i microscope.

6. Embryo analyses

6.1 Cuticle preparation

Females were allowed to lay eggs in apple juice agar plates at 25ºC. Eggs were collected with deionized water at least 24 hours after being laid. Eggs were dechorionated in 100% bleach for 1 minute, washed with deionized water and mounted in Hoyer's medium: lactic acid (1:1). The sample was incubated at 62ºC overnight. Dark field photographs were obtained with a Nikon eclipse 80i microscope.

6.2 Collection and fixation of embryos

For embryo collection, females were allowed to lay eggs in apple juice agar plates at 25ºC. Embryos were collected with deionized water, dechorionated in bleach 100% for 1 minute and washed with deionized water. Then, embryos were fixed under strong agitation in vials

containing a solution of 4% formaldehyde-PBS: Heptane (1:1). For protein analysis, embryos were fixed for 15 minutes while embryos for mRNA analysis were fixed for 20 minutes. Subsequently, fixed embryos were devitellinized in a heptane: methanol solution (1:1) and rinsed 4 times with methanol 100%. Embryos for immunostaining were stored at -20ºC a maximum of 24 hours while embryos for mRNA analysis can be stored up to several months.

6.3 In situ mRNA hybridization of embryos

In situ hybridizations were performed to visualize mRNA expression in fixed embryos, using digoxigenin-UTP labeled antisense RNA probes. After fixation with formaldehyde, embryos were rinsed 4 times in ethanol 100% and once in ethanol: xylenes (1:1). Embryos were then washed and incubated with xylenes 100% for 90 minutes at room temperature. Embryos were rinsed once with ethanol: xylenes (1:1), 3 times with ethanol 100% and twice with methanol 100%. Embryos were post-fixed in PTW (1X PBS with 0.1% Tween 20) with 4% formaldehyde for 25 minutes at room temperature, washed several times with PTW and incubated for 3 minutes with Proteinase K diluted in PTW at a final concentration of 4 µg/ml. Proteinase K was removed and embryos were washed several times with PTW and post-fixed again. After the second post-fixation, embryos were rinsed 5 times with PTW, incubated with PTW: hybridization solution (1:1) for 2 minutes and prehybridized in hybridization solution for 2 hours at 55ºC. The digoxigenin-UTP labelled antisense RNA probe was diluted in hybridization solution (0.5µl of the probe in 60µl of hybridization solution) and heated at 80ºC for 5 minutes. Next, the probe was cooled down by placing it 1 minute on ice and added to the embryos for an overnight incubation at 55ºC.

The next day, after removal of the probe, embryos were washed 5 times for 20 minutes each wash: once with hybridization solution, once with a PTW: hybridization solution (1:1), and 3 times with PTW. Embryos were then incubated with pre-cleared anti-Digoxigenin antibody conjugated with Alkaline Phosphatase (Roche) diluted 1:2000 in PTW at room temperature for two hours. After the antibody was removed, embryos were washed 3 times with PTW under rotation for 20 minutes each wash. Embryos were then transferred to a multiwell dish and incubated in Alkaline Phosphatase buffer containing 4.5 mg/ml NBT BCIP until an intense purple color was developed (several minutes to hours, depending on the probe and expression of the gene monitored). Phosphatase alkaline reaction was stopped by washing with PTW several times. Embryos were dehydrated with increasing concentrations of ethanol, washed once with ethanol: xylenes (1:1), washed

twice with xylenes 100% and mounted in Permount (Fisher). Brightfield images were obtained with a Nikon eclipse 80i microscope.

6.4 Embryo immunostaining

Fixed embryos were rehydrated with a first wash in PBS: Methanol 1:1 and 3 washes in PBT 0.3% (1X PBS with 0.3% Triton X-100). Then, embryos were blocked with PBT-BSA 2% (PBT 0.3% with 2% BSA) under rotation at 4ºC for at least 45 minutes. Primary antibodies were diluted in PBT-BSA 2% and added to the embryos for an overnight incubation at 4ºC under rotation. After removal of primary antibodies, embryos were washed 4 times for 15 minutes each with PBT-BSA 2% and incubated with fluorescently conjugated secondary antibodies in PBT-BSA 2% for 2 hours at 4ºC (after the secondary antibody was added, samples were protected from light). If nuclei staining was required, DAPI was added during the last 10 minutes of secondary antibody incubation, at a final concentration of 1:10000. Then, embryos were washed with PBT 0.3% every 15-20 minutes for 90 minutes and mounted in Fluoromount-G (Southern Biotech). Confocal images were obtained with a Leica TCS SP5 microscope.

For ovary immunostainings we used the following primary antibodies: rat anti-HA (Sigma, dilution 1:250 dilution), rabbit anti-GFP (Thermofisher, 1:500 dilution).

7. Wing analyses

7.1 Mounting of wings

For analysis of ectopic vein tissue, wings were removed from adult flies and directly transferred to a slide with isopropanol. After evaporation of the isopropanol, we added Euparal mounting medium and the coverslip. Brightfield images were obtained with a E600 Nikon microscope.

7.2 Wing disc immunostaining

Larvae were dissected in cold PBS, fixed for 20 minutes in 4% formaldehyde-PBS and washed 3 times for 15 minutes each wash with PBT 0.1% (1X PBS with 0.1% Triton X-100). Samples were blocked with BBT ((PBT 0.1% with 0.3% BSA and 250 mM NaCl) for at least 30 minutes and incubated overnight at 4ºC with primary antibodies diluted in BBT.

The next day, the samples were washed 4 times for 15 minutes each with BBT and incubated with fluorescently conjugated secondary antibodies diluted in BBT for 2 hours (the samples were protected from light after the secondary antibody was added). DAPI was added during the last 10 minutes of incubation at a final concentration of 1:10000. After secondary antibody incubation, samples were washed with PBT 0.1% every 15-20 minutes for 1 hour and mounted in mounting medium (40ml glycerol + 5ml PBS10X + 400 microlitres N-propyl-gallate 50% diluted in ethanol). Confocal images were obtained with a Leica TCS SP5 microscope.

For wing disc immunostainings we used the following primary antibodies: rat anti-HA (Sigma, dilution 1:250 dilution), rabbit anti-GFP (Thermofisher, 1:500 dilution).

8. Proteomic screen in S2 cells

8.1 Cell culture, transfection and establishment of stable S2 cell lines

Drosophila S2 cells were grown at 25ºC in standard Schneider's S2 medium supplemented with 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin (Life Technologies). S2 cells were transfected using Effectene transfection reagent (Qiagen) according to the manufacturer's instructions. Stable cell lines were selected in the presence of 300 µg/mL hygromycin (Sigma).

8.2 Induction and Cell lysis

For affinity purification, stable cells lines were grown in 75-cm^2 vented flasks, two flasks for each cell line. Untransfected S2 cells were grown in parallel as a negative control for purifications. Expression was induced by treating the cells with 0.07mM $CuSO_4$ overnight (1:1000 dilution from a 0.07M $CuSO_4$ stock solution).

Cells were collected in Falcon tubes, centrifuged at 500g for 3 minutes, washed once with cold PBS and centrifuged a second time. The supernatant was removed, and cells were lysed in 1mL of cold Default Lysis Buffer (DLB) (50 mM Tris pH 7.5, 125 mM NaCl, 5% glycerol, 0.5% IGEPAL, 1.5 mM MgCl2, 1 mM DTT, 25 mM NaF, 1mM Na3VO4, 1mM EDTA and 2x Complete protease inhibitor, Roche).

8.3 Affinity purification

Affinity purification was performed as previously described (Yang and Veraksa, 2017). Briefly, cell lysates were incubated on ice for 15 minutes and centrifuged at maximum speed. Supernatants were filtered and incubated with streptavidin beads for 2 hours at 4ºC under rotation. After incubation, the beads were washed with cold DLB 5 times. Elution was performed by incubating the beads with a 2mM biotin solution. Eluates were divided in two to prepare samples for mass spectrometry and silver-stained gel analysis. Both samples were precipitated with a 100% trichloroacetic acid (TCA) solution and washed 4 times with cold acetone. Pellets were allowed to dry overnight before storing at -20ºC. Three biological replicates were conducted for each construct in parallel with untransfected S2 cells.

8.4 Silver staining and mass spectrometry analysis

Before processing the samples for mass spectrometry analysis, the quality of the samples was assessed by silver staining. The dried pellets were resuspended in 20µl of 2x SDS sample buffer (4% SDS, 80mM Tris pH 6.8, 15% glycerol, 500ng/ml bromophenol blue and 25mM DTT) and heated at 95ºC for 5 minutes. Samples were run in a Novex NuPAGE 4-12% Bis-Tris gradient gel with MOPS running buffer. The gel was allowed to run until the front dye reached the bottom of the gel. The gel was stained with the SilverQuest Staining Kit according to the manufacturer's instructions. We confirmed that the bait protein had been correctly induced and migrated and the expected molecular weight. We also detected additional bands of interacting proteins that were not present in the unstransfected S2 cells.

Protein complexes were analyzed by nanoscale liquid chromatography coupled to tandem mass spectrometry (nanoLC-MS/MS) at the Taplin Mass Spectrometry Facility at Harvard Medical School. Before submission, the samples were run on a short SDS-PAGE gel. The dye front was allowed to migrate in the separating gel up to a distance of 1cm and the gel was stained with Colloidal Coomassie Brilliant Blue. The gel was destained with a solution of 25% methanol and 5% acetic acid, and extensively washed in water. Finally, each lane was cut into two square 5mm x 5mm pieces and submitted for mass spectrometry analysis. The identified Cic-L^{Mini} and Cic-L-interacting proteins by mass spectrometry were then analyzed with the SAINT program, which assigns confidence scores to protein-protein interactions (scores above 0.8 are considered highly significant).

REFERENCES

Ahmad, S. T., Rogers, A. D., Chen, M. J., Dixit, R., Adnani, L., Frankiw, L. S., Lawn, S. O., Blough, M. D., Alshehri, M., Wu, W., et al. (2019). Capicua regulates neural stem cell proliferation and lineage specification through control of Ets factors. *Nat. Commun.* **10**,.

Ajuria, L., Nieva, C., Winkler, C., Kuo, D., Samper, N., José Andreu, M., Helman, A., González-Crespo, S., Paroush, Z., Courey, A. J., et al. (2011). Capicua DNA-binding sites are general response elements for RTK signaling in Drosophila. *Development* **138**, 915–924.

Andreu, M. J., Ajuria, L., Samper, N., González-Pérez, E., Campuzano, S., González-Crespo, S. and Jiménez, G. (2012a). EGFR-Dependent downregulation of capicua and the establishment of drosophila dorsoventral polarity. *Fly (Austin).* **6**,.

Andreu, M. J., González-Pérez, E., Ajuria, L., Samper, N., González-Crespo, S., Campuzano, S. and Jiménez, G. (2012b). Mirror represses pipe expression in follicle cells to initiate dorsoventral axis formation in Drosophila. *Development* **139**, 1110–1114.

Antonescu, C. R., Owosho, A. A., Zhang, L., Chen, S., Deniz, K., Huryn, J. M., Kao, Y. C., Huang, S. C., Singer, S., Tap, W., et al. (2017). Sarcomas with CIC-rearrangements Are a Distinct Pathologic Entity with Aggressive Outcome. *Am. J. Surg. Pathol.* **41**, 941–949.

Arias, E. E. and Walter, J. C. (2006). PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nat. Cell Biol.* **8**, 84–90.

Arias, E. E. and Walter, J. C. (2007). Strength in numbers: Preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev.* **21**, 497–518.

Astigarraga, S., Grossman, R., Díaz-Delfín, J., Caelles, C., Paroush, Z. and Jiménez, G. (2007). A MAPK docking site is critical for downregulation of Capicua by Torso and EGFR RTK signaling. *EMBO J.* **26**, 668–677.

Atkey, M. R., Lachance, J. F. B., Walczak, M., Rebello, T. and Nilson, L. A. (2006). Capicua regulates follicle cell fate in the Drosophila ovary through repression of mirror. *Development* **133**, 2115–2123.

Audibert, A., Simon, F. and Gho, M. (2005). Cell cycle diversity involves differential regulation of cyclin E activity in the Drosophila bristle cell lineage. *Development* **132**, 2287–2297.

Badenhorst, P. (2001). Tramtrack controls glial number and identity in the Drosophila embryonic CNS. *Development* **128**, 4093–4101.

Bass, B. P., Cullen, K. and McCall, K. (2007). The axon guidance gene lola is required for programmed cell death in the Drosophila ovary. *Dev. Biol.* **304**, 771–785.

Bastock, R. and St Johnston, D. (2008). Drosophila oogenesis. *Curr. Biol.* **18**, 1082–1087.

Becalska, A. N. and Gavis, E. R. (2009). Lighting up mRNA localization in Drosophila oogenesis. *Development* **136**, 2493–2503.

Bell, S. P. and Kaguni, J. M. (2013). Helicase loading at chromosomal origins of replication. *Cold Spring Harb. Perspect. Biol.* **5**,.

Belver, L. and Ferrando, A. (2016). The genetics and mechanisms of T cell acute lymphoblastic leukaemia. *Nat. Rev. Cancer* **16**, 494–507.

Bettegowda, C., Agrawal, N., Jiao, Y., Sausen, M., Wood, L. D., Hruban, R. H., Rodriguez, F. J., Cahill, D. P., McLendon, R., Riggins, G., et al. (2011). Mutations in CIC and FUBP1 contribute to human oligodendroglioma. *Science (80-.).* **333**, 1453–1455.

Bianco, A., Poukkula, M., Cliffe, A., Mathieu, J., Luque, C. M., Fulga, T. A. and Rørth, P. (2007). Two distinct modes of guidance signalling during collective migration of border cells. *Nature* **448**, 362–365.

Birstein, V. (1989). On the karyotype of Trichoplax sp. (Placozoa). *Biol. Zent. Bl.* **108**, 63–67.

Bischof, J., Maeda, R. K., Hediger, M., Karch, F. and Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific φC31 integrases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3312–3317.

Bowman, A. B., Lam, Y. C., Jafar-Nejad, P., Chen, H. K., Richman, R., Samaco, R. C., Fryer, J. D., Kahle, J. J., Orr, H. T. and Zoghbi, H. Y. (2007). Duplication of Atxn1l suppresses SCA1 neuropathology by decreasing incorporation of polyglutamine-expanded ataxin-1 into native complexes. *Nat. Genet.* **39**, 373–379.

Bramsiepe, J., Wester, K., Weinl, C., Roodbarkelari, F., Kasili, R., Larkin, J. C., Hülskamp, M. and Schnittger, A. (2010). Endoreplication controls cell fate maintenance. *PLoS Genet.* **6**, 1–14.

Britton, J. S. and Edgar, B. A. (1998). Environmental control of the cell cycle in Drosophila: nutrition activates mitotic and endoreplicative cells by distinct mechanisms. *Development* **125**, 2149–2158.

Bunda, S., Heir, P., Metcalf, J., Li, A. S. C., Agnihotri, S., Pusch, S., Yasin, M., Li, M., Burrell, K., Mansouri, S., et al. (2019). CIC protein instability contributes to tumorigenesis in glioblastoma. *Nat. Commun.* **10**,.

Bunda, S., Heir, P., Li, A. S. C., Mamatjan, Y., Zadeh, G. and Aldape, K. (2020). c-Src Phosphorylates and Inhibits the Function of the CIC Tumor Suppressor Protein. *Mol. Cancer Res.* **18**, 774–786.

Buszczak, M. and Cooley, L. (2000). Eggs to die for: Cell death during Drosophila oogenesis. *Cell Death Differ.* **7**, 1071–1074.

Buszczak, M., Freeman, M. R., Carlson, J. R., Bender, M., Cooley, L. and Segraves, W. A. (1999). Ecdysone response genes govern egg chamber development during mid-oogenesis in Drosophila. *Development* **126**, 4581–4589.

Cáceres, L. and Nilson, L. A. (2005). Production of gurken in the nurse cells is sufficient for axis determination in the Drosophila oocyte. *Development* **132**, 2345–2353.

Calvi, B. R. and Spradling, A. C. (1999). Chorion gene amplification in Drosophila: A model for metazoan origins of DNA replication and S-Phase control. *Methods A Companion to Methods Enzymol.* **18**, 407–417.

Calvi, B. R., Lilly, M. A. and Spradling, A. C. (1998). Cell cycle control of chorion gene amplification. *Genes Dev.* **12**, 734–744.

Campbell, P. J., Getz, G., Korbel, J. O., Stuart, J. M., Jennings, J. L., Stein, L. D., Perry, M. D., Nahal-Bose, H. K., Ouellette, B. F. F., Li, C. H., et al. (2020). Pan-cancer analysis of whole genomes. *Nature* **578**, 82–93.

Cant, K., Knowles, B. A., Mooseker, M. S. and Cooley, L. (1994). Drosophila singed, a fascin homolog, is required for actin bundle formation during oogenesis and bristle extension. *J. Cell Biol.* **125**, 369–380.

Channavajhala, P. and Seldin, D. C. (2002). Functional interaction of protein kinase CK2 and c-Myc in lymphomagenesis. *Oncogene* **21**, 5280–5288.

Chen, H. Z., Ouseph, M. M., Li, J., Pécot, T., Chokshi, V., Kent, L., Bae, S., Byrne, M., Duran, C., Comstock, G., et al. (2012). Canonical and atypical E2Fs regulate the mammalian endocycle. *Nat. Cell Biol.* **14**, 1192–1202.

Chittaranjan, S., Chan, S., Yang, C., Yang, K. C., Chen, V., Moradian, A., Firme, M., Song, J., Go, N. E., Blough, M. D., et al. (2014). Mutations in CIC and IDH1 cooperatively regulate 2-hydroxyglutarate levels and cell clonogenicity. *Oncotarget* **5**, 7960–7979.

Choi, H., Larsen, B., Lin, Z. Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z. S., Tyers, M., Gingras, A. C. and Nesvizhskii, A. I. (2011). SAINT: Probabilistic scoring of affinity purificationg-mass spectrometry data. *Nat. Methods* **8**, 70–73.

Choi, N., Park, J., Lee, J. S., Yoe, J., Park, G. Y., Kim, E., Jeon, H., Cho, Y. M., Roh, T. Y. and Lee, Y. (2015). miR-93/miR-106b/miR-375-CIC-CRABP1: A novel regulatory axis in prostate cancer progression. *Oncotarget* **6**, 23533–23547.

Chou, T. Bin and Perrimon, N. (1996). The autosomal FLP-DFS technique for generating germline mosaics in Drosophila melanogaster. *Genetics* **144**, 1673–1679.

Chou, T. B., Noll, E. and Perrimon, N. (1993). Autosomal P[ovo(D1)]dominant female-sterile insertions in Drosophila and their use in generating germ-line chimeras. *Development* **119**, 1359–1369.

Cinnamon, E., Gur-Wahnon, D., Helman, A., St. Johnston, D., Jiménez, G. and Paroush, Z. (2004). Capicua integrates input from two maternal systems in Drosophila terminal patterning. *EMBO J.* **23**, 4571–4582.

Clough, E., Tedeschi, T. and Hazelrigg, T. (2014). Epigenetic regulation of oogenesis and germ stem cell maintenance by the Drosophila histone methyltransferase Eggless/dSetDB1. *Dev. Biol.* **388**, 181–191.

Conlon, G. A. and Murray, G. I. (2019). Recent advances in understanding the roles of matrix metalloproteinases in tumour invasion and metastasis. *J. Pathol.* **247**, 629–640.

Cooley, L., Verheyen, E. and Ayers, K. (1992). Chickadee Encodes a Profilin Required for Intercellular cytoplasm transport during Drosophila oogenesis. *Dev. Biol.* **69**, 173–184.

Crespo-Barreto, J., Fryer, J. D., Shaw, C. A., Orr, H. T. and Zoghbi, H. Y. (2010). Partial loss of ataxin-1 function contributes to transcriptional dysregulation in spinocerebellar ataxia type 1 pathogenesis. *PLoS Genet.* **6**, 1–17.

Cross, D. A. E., Alessi, D. R., Cohen, P., Andjelkovich, M. and Hemmings, B. A. (1995). Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. *Nature* **378**, 785–789.

Da Vià, M. C., Solimando, A. G., Garitano-Trojaola, A., Barrio, S., Munawar, U., Strifler, S., Haertle, L., Rhodes, N., Teufel, E., Vogt, C., et al. (2020). CIC Mutation as a Molecular Mechanism of Acquired Resistance to Combined BRAF-MEK Inhibition in Extramedullary Multiple Myeloma with Central Nervous System Involvement . *Oncologist* **25**, 112–118.

Dang, L., White, D. W., Gross, S., Bennett, B. D., Bittinger, M. A., Driggers, E. M., Fantin, V. R., Jang, H. G., Jin, S., Keenan, M. C., et al. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744.

de Launoit, Y., Baert, J. L., Chotteau-Lelievre, A., Monte, D., Coutte, L., Mauen, S., Firlej, V., Degerny, C. and Verreman, K. (2006). The Ets transcription factors of the PEA3 group: Transcriptional regulators in metastasis. *Biochim. Biophys. Acta - Rev. Cancer* **1766**, 79–87.

de Nooij, J. C., Graber, K. H. and Hariharan, I. K. (2000). Expression of the cyclin-dependent kinase inhibitor Dacapo is regulated by Cyclin E. *Mech. Dev.* **97**, 73–83.

De Nooij, J. C., Letendre, M. A. and Hariharan, I. K. (1996). A cyclin-dependent kinase inhibitor, dacapo, is necessary for timely exit from the cell cycle during Drosophila embryogenesis. *Cell* **87**, 1237–1247.

de Veylder, L., Larkin, J. C. and Schnittger, A. (2011). Molecular control and function of endoreplication in development and physiology. *Trends Plant Sci.* **16**, 624–634.

Deluca, S. Z. and Spradling, A. C. (2018). Efficient expression of genes in the drosophila germline using a uas promoter free of interference by hsp70 pirnas. *Genetics* **209**, 381–387.

Demontis, F. and Perrimon, N. (2009). Integration of Insulin receptor/Foxo signaling and dMyc activity during muscle growth regulates body size in Drosophila. *Development* **136**, 983–993.

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

Dissanayake, K., Toth, R., Blakey, J., Olsson, O., Campbell, D. G., Prescott, A. R. and Mackintosh, C. (2011). ERK/p90RSK/14-3-3 signalling has an impact on expression of PEA3 Ets transcription factors via the transcriptional repressor capicúa. *Biochem. J.* **433**, 515–525.

Donne, R., Saroul-Aïnama, M., Cordier, P., Celton-Morizur, S. and Desdouets, C. (2020). Polyploidy in liver development, homeostasis and disease. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 391–405.

Doronkin, S., Djagaeva, I. and Beckendorf, S. K. (2003). The COP9 signalosome promotes degradation of Cyclin E during early Drosophila oogenesis. *Dev. Cell* **4**, 699–710.

Doyonnas, R., Kershaw, D. B., Duhme, C., Merkens, H., Chelliah, S., Graf, T. and McNagny, K. M. (2001). Anuria, omphalocele, and perinatal lethality in mice lacking the CD34-related protein podocalyxin. *J. Exp. Med.* **194**, 13–27.

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

Duronio, R. J. and O'Farrell, P. H. (1995). Developmental control of the G1 to S transition in Drosophila: Cyclin E is a limiting downstream target of E2F. *Genes Dev.* **9**, 1456–1468.

Duronio, R. J., Bonnette, P. C. and O'Farrell, P. H. (1998). Mutations of the Drosophila dDP,dE2F, and cyclin E Genes Reveal Distinct Roles for the E2F-DP Transcription Factor and Cyclin E during the G1-S Transition. *Mol. Cell. Biol.* **18**, 141–151.

Edgar, B. A. and Orr-Weaver, T. L. (2001). Endoreplication cell cycles: More for less. *Cell* **105**, 297–306.

Edgar, B. A., Zielke, N. and Gutierrez, C. (2014). Endocycles: A recurrent evolutionary innovation for post-mitotic cell growth. *Nat. Rev. Mol. Cell Biol.* **15**, 197–210.

Edwards, K. A. and Kiehart, D. P. (1996). Drosophila nonmuscle myosin II has multiple essential roles in imaginal disc and egg chamber morphogenesis. *Development* **122**, 1499–1511.

Ephrussi, A. and Lehmann, R. (1992). Induction of germ cell formation by oskar. *Nature* **358**, 387–392.

Ephrussi, A., Dickinson, L. K. and Lehmann, R. (1991). oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* **66**, 37–50.

Faure, G. and Callebaut, I. (2013). Identification of hidden relationships from the coupling of Hydrophobic Cluster Analysis and Domain Architecture information. *Bioinformatics* **29**, 1726–1733.

Fernandez-Valverde, S. L., Calcino, A. D. and Degnan, B. M. (2015). Deep developmental transcriptome sequencing uncovers numerous new genes and enhances gene annotation in the sponge Amphimedon queenslandica. *BMC Genomics* **16**,.

Foe, V. E. (1989). Mitotic domains reveal early commitment of cells in Drosophila embryos. *Trends Genet.* **5**, 322.

Foley, K. and Cooley, L. (1998). Apoptosis in late stage Drosophila nurse cells does not require genes within the H99 deficiency. *Development* **125**, 1075–1082.

Follette, P. J., Duronio, R. J. and O'Farrell, P. H. (1998). Fluctuations in cyclin E levels are required for multiple rounds of endocycle S phase in Drosophila. *Curr. Biol.* **8**, 235–238.

Forés, M., Ajuria, L., Samper, N., Astigarraga, S., Nieva, C., Grossman, R., González-Crespo, S., Paroush, Z. and Jiménez, G. (2015). Origins of Context-Dependent Gene Repression by Capicua. *PLoS Genet.* **11**,.

Forés, M., Papagianni, A., Rodríguez-Muñoz, L. and Jiménez, G. (2017a). Using CRISPR-cas9 to study ERK signaling in Drosophila. In *Methods in Molecular Biology*, pp. 353–365. Humana Press Inc.

Forés, M., Simón-Carrasco, L., Ajuria, L., Samper, N., González-Crespo, S., Drosten, M., Barbacid, M. and Jiménez, G. (2017b). A new mode of DNA binding distinguishes Capicua from other HMG-box factors and explains its mutation patterns in cancer. *PLoS Genet.* **13**,.

Forrest, K. M. and Gavis, E. R. (2003). Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in Drosophila. *Curr. Biol.* **13**, 1159–1168.

Fox, D. T. and Duronio, R. J. (2013). Endoreplication and polyploidy: Insights into development and disease. *Development* **140**, 3–12.

Frawley, L. E. and Orr-Weaver, T. L. (2015). Polyploidy. *Curr. Biol.* **25**, R353–R358.

Freeman, M., Klämbt, C., Goodman, C. S. and Rubin, G. M. (1992). The argos gene encodes a diffusible factor that regulates cell fate decisions in the drosophila eye. *Cell* **69**, 963–975.

Fryer, J. D., Yu, P., Kang, H., Mandel-Brehm, C., Carter, A. N., Crespo-Barreto, J., Gao, Y., Flora, A., Shaw, C., Orr, H. T., et al. (2011). Exercise and genetic rescue of SCA1 via the transcriptional repressor Capicua. *Science (80-.).* **334**, 690–693.

Furriols, M., Ventura, G. and Casanova, J. (2007). Two distinct but convergent groups of cells trigger Torso receptor tyrosine kinase activation by independently expressing torso-like. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11660–11665.

Futran, A. S., Kyin, S., Shvartsman, S. Y. and Link, A. J. (2015). Mapping the binding interface of ERK and transcriptional repressor Capicua using photocrosslinking. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8590–8595.

Gallant, P. (2013). Myc function in drosophila. *Cold Spring Harb. Perspect. Med.* **3**,.

Galletti, M., Riccardo, S., Parisi, F., Lora, C., Saqcena, M. K., Rivas, L., Wong, B., Serra, A., Serras, F., Grifoni, D., et al. (2009). Identification of Domains Responsible for Ubiquitin-Dependent Degradation of dMyc by Glycogen Synthase Kinase 3β and Casein Kinase 1 Kinases. *Mol. Cell. Biol.* **29**, 3424–3434.

Gandarillas, A., Molinuevo, R. and Sanz-Gómez, N. (2018). Mammalian endoreplication emerges to reveal a potential developmental timer. *Cell Death Differ.* **25**, 471–476.

Garcí-Higuera, I., Manchado, E., Dubus, P., Cañamero, M., Méndez, J., Moreno, S. and Malumbres, M. (2008). Genomic stability and tumour suppression by the APC/C cofactor Cdh1. *Nat. Cell Biol.* **10**, 802–811.

Gavis, E. R. and Lehmann, R. (1992). Localization of nanos RNA controls embryonic polarity. *Cell* **71**, 301–313.

Geng, Y., Yu, Q., Sicinska, E., Das, M., Schneider, J. E., Bhattacharya, S., Rideout, W. M., Bronson, R. T., Gardner, H. and Sicinski, P. (2003). Cyclin E ablation in the mouse. *Cell* **114**, 431–443.

Giese, K., Cox, J. and Grosschedl, R. (1992). The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69**, 185–195.

Gleize, V., Alentorn, A., Connen De Kérillis, L., Labussière, M., Nadaradjane, A. A., Mundwiller, E., Ottolenghi, C., Mangesius, S., Rahimian, A., Ducray, F., et al. (2015). CIC inactivating mutations identify aggressive subset of 1p19q codeleted gliomas. *Ann. Neurol.* **78**, 355–374.

Goff, D. J., Nilson, L. A. and Morisato, D. (2001). Establishment of dorsal-ventral polarity of the Drosophila egg requires capicua action in ovarian follicle cells. *Development* **128**, 4553–4562.

Golembo, M., Schweitzer, R., Freeman, M. and Shilo, B. Z. (1996). Argos transcription is induced by the Drosophila EGF receptor pathway to form an inhibitory feedback loop. *Development* **122**, 223–230.

González-Reyes, A. and St Johnston, D. (1994). Role of oocyte position in establishment of anterior-posterior polarity in Drosophila. *Science (80-.).* **266**, 639–642.

González-Reyes, A., Elliott, H. and St Johnston, D. (1995). Polarization of both major body axes in drosophila by gurken-torpedo signalling. *Nature* **375**, 654–658.

Gratz, S. J., Rubinstein, C. D., Harrison, M. M., Wildonger, J. and O'Connor-Giles, K. M. (2015). CRISPR-Cas9 Genome Editing in *Drosophila*. *Curr. Protoc. Mol. Biol.* **111**, 31.2.1-31.2.20.

Gregory, M. A. and Hann, S. R. (2000). c-Myc Proteolysis by the Ubiquitin-Proteasome Pathway: Stabilization of c-Myc in Burkitt's Lymphoma Cells. *Mol. Cell. Biol.* **20**, 2423–2435.

Grewal, S. S., Li, L., Orian, A., Eisenman, R. N. and Edgar, B. A. (2005). Myc-dependent regulation of ribosomal RNA synthesis during Drosophila development. *Nat. Cell Biol.* **7**, 295–302.

Grimm, O., Sanchez Zini, V., Kim, Y., Casanova, J., Shvartsman, S. Y. and Wieschaus, E. (2012). Torso RTK controls Capicua degradation by changing its subcellular localization. *Dev.* **139**, 3962–3968.

Gunkel, N., Yano, T., Markussen, F. H., Olsen, L. C. and Ephrussi, A. (1998). Localization-dependent translation requires a functional interaction between the 5′ and 3′ ends of oskar mRNA. *Genes Dev.* **12**, 1652–1664.

Hammond, M. P. and Laird, C. D. (1985a). Chromosome structure and DNA replication in nurse and follicle cells of Drosophila melanogaster. *Chromosoma* **91**, 267–278.

Hammond, M. P. and Laird, C. D. (1985b). Control of DNA replication and spatial distribution of defined DNA sequences in salivary gland cells of Drosophila melanogaster. *Chromosoma* **91**, 279–86.

Hao, B., Oehlmann, S., Sowa, M. E., Harper, J. W. and Pavletich, N. P. (2007). Structure of a Fbw7-Skp1-Cyclin E Complex: Multisite-Phosphorylated Substrate Recognition by SCF Ubiquitin Ligases. *Mol. Cell* **26**, 131–143.

Hayashi, S. (1996). A Cdc2 dependent checkpoint maintains diploidy in Drosophila. *Development* **122**,.

Hayes, P. and Solon, J. (2017). Drosophila dorsal closure: An orchestra of forces to zip shut the embryo. *Mech. Dev.* **144**, 2–10.

Herranz, H., Hong, X. and Cohen, S. M. (2012). Mutual repression by bantam miRNA and capicua links the EGFR/MAPK and hippo pathways in growth control. *Curr. Biol.* **22**, 651–657.

Hong, A., Lee-Kong, S., Iida, T., Sugimura, I. and Lilly, M. A. (2003). The p27cip/kip ortholog dacapo maintains the Drosophila oocyte in prophase of meiosis I. *Development* **130**, 1235–1242.

Hong, A., Narbonne-Reveau, K., Riesgo-Escovar, J., Fu, H., Aladjem, M. I. and Lilly, M. A. (2007). The cyclin-dependent kinase inhibitor Dacapo promotes replication licensing during Drosophila endocycles. *EMBO J.* **26**, 2071–2082.

Horne-Badovinac, S. and Bilder, D. (2005). Mass transit: Epithelial morphogenesis in the Drosophila egg chamber. *Dev. Dyn.* **232**, 559–574.

Hu, D. and Cross, J. C. (2010). Development and function of trophoblast giant cells in the rodent placenta. *Int. J. Dev. Biol.* **54**, 341–354.

Huang, S. C., Zhang, L., Sung, Y. S., Chen, C. L., Kao, Y. C., Agaram, N. P., Singer, S., Tap, W. D., D'Angelo, S. and Antonescu, C. R. (2016). Recurrent CIC gene abnormalities in angiosarcomas: A molecular study of 120 cases with concurrent investigation of PLCG1, KDR, MYC, and FLT4 gene alterations. *Am. J. Surg. Pathol.* **40**, 645–655.

Hudson, A. M. and Cooley, L. (2002). Understanding the function of actin-binding proteins through genetic analysis of Drosophila oogenesis. *Annu. Rev. Genet.* **36**, 455–488.

Hulf, T., Bellosta, P., Furrer, M., Steiger, D., Svensson, D., Barbour, A. and Gallant, P. (2005). Whole-Genome Analysis Reveals a Strong Positional Bias of Conserved dMyc-Dependent E-Boxes. *Mol. Cell. Biol.* **25**, 3401–3410.

Hwang, I., Pan, H., Yao, J., Elemento, O., Zheng, H. and Paik, J. (2020). CIC is a critical regulator of neuronal differentiation. *JCI Insight* **5**,.

Irish, V., Lehmann, R. and Akam, M. (1989). The Drosophila posterior-group gene nanos functions by repressing hunchback activity. *Nature* **338**, 646–648.

Jané-Valbuena, J., Widlund, H. R., Perner, S., Johnson, L. A., Dibner, A. C., Lin, W. M., Baker, A. C., Nazarian, R. M., Vijayendran, K. G., Sellers, W. R., et al. (2010). An oncogenic role for ETV1 in melanoma. *Cancer Res.* **70**, 2075–2084.

Jiao, Y., Killela, P. J., Reitman, Z. J., Rasheed, B. A., Heaphy, C. M., de Wilde, R. F., Rodriguez, F. J., Rosemberg, S., Oba-Shinjo, S. M., Marie, S. K. N., et al. (2012). Frequent ATRX, CIC, FUBP1 and IDH1 mutations refine the classification of malignant gliomas. *Oncotarget* **3**, 709–722.

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

Jiménez, G., Shvartsman, S. Y. and Paroush, Z. (2012a). The Capicua repressor - A general sensor of RTK signaling in development and disease. *J. Cell Sci.* **125**, 1383–1391.

Jiménez, G., Shvartsman, S. Y. and Paroush, Z. (2012b). The Capicua repressor - A general sensor of RTK signaling in development and disease. *J. Cell Sci.* **125**, 1383–1391.

Jin, Y., Ha, N., Forés, M., Xiang, J., Gläßer, C., Maldera, J., Jiménez, G. and Edgar, B. A. (2015). EGFR/Ras Signaling Controls Drosophila Intestinal Stem Cell Proliferation via Capicua-Regulated Genes. *PLoS Genet.* **11**,.

Kamachi, Y. and Kondoh, H. (2013). Sox proteins: Regulators of cell fate specification and differentiation. *Development* **140**, 4129–4144.

Kawamura-Saito, M., Yamazaki, Y., Kaneko, K., Kawaguchi, N., Kanda, H., Mukai, H., Gotoh, T., Motoi, T., Fukayama, M., Aburatani, H., et al. (2006). Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Hum. Mol. Genet.* **15**, 2125–2137.

Keenan, S. E., Blythe, S. A., Marmion, R. A., Djabrayan, N. J. V., Wieschaus, E. F. and Shvartsman, S. Y. (2020). Rapid Dynamics of Signal-Dependent Transcriptional Repression by Capicua. *Dev. Cell* **52**, 794–801.

Killela, P. J., Reitman, Z. J., Jiao, Y., Bettegowda, C., Agrawal, N., Diaz, L. A., Friedman, A. H., Friedman, H., Gallia, G. L., Giovanella, B. C., et al. (2013). TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6021–6026.

Kim, Y., Coppey, M., Grossman, R., Ajuria, L., Jiménez, G., Paroush, Z. and Shvartsman, S. Y. (2010). MAPK Substrate Competition Integrates Patterning Signals in the Drosophila Embryo. *Curr. Biol.* **20**, 446–451.

Kim, Y., Andreu, M. J., Lim, B., Chung, K., Terayama, M., Jiménez, G., Berg, C. A., Lu, H. and Shvartsman, S. Y. (2011). Gene regulation by MAPK substrate competition. *Dev. Cell* **20**, 880–887.

Kim, E., Lu, H. C., Zoghbi, H. Y. and Song, J. J. (2013). Structural basis of protein complex formation and reconfiguration by polyglutamine disease protein ataxin-1 and Capicua. *Genes Dev.* **27**, 590–595.

Kim, E., Park, S., Choi, N., Lee, J., Yoe, J., Kim, S., Jung, H. Y., Kim, K. T., Kang, H., Fryer, J. D., et al. (2015). Deficiency of Capicua disrupts bile acid homeostasis. *Sci. Rep.* **5**,.

Kim, E., Kim, D., Lee, J. S., Yoe, J., Park, J., Kim, C. J., Jeong, D., Kim, S. and Lee, Y. (2018). Capicua suppresses hepatocellular carcinoma progression by controlling the ETV4–MMP1 axis. *Hepatology* **67**, 2287–2301.

King, R. C. (1970). *Ovarian Development in Drosophila melanogaster*.

Koepp, D. M., Schaefer, L. K., Ye, X., Keyomarsi, K., Chu, C., Harper, J. W. and Elledge, S. J. (2001). Phosphorylation-dependent ubiquitination of cyctin E by the SCFFbw7 ubiquitin ligase. *Science (80-.).* **294**, 173–177.

Kondo, S. and Ueda, R. (2013). Highly Improved gene targeting by germline-specific Cas9 expression in Drosophila. *Genetics* **195**, 715–721.

Kugler, J. M. and Lasko, P. (2009). Localization, anchoring and translational control of oskar, gurken, bicoid and nanos mRNA during drosophila oogenesis. *Fly (Austin).* **3**, 15–28.

Kyriakakis, P., Tipping, M., Abed, L. and Veraksa, A. (2008). Tandem affinity purification in drosophila the advantages of the GS-TAP system. *Fly (Austin).* **2**, 229–235.

Lam, Y. C., Bowman, A. B., Jafar-Nejad, P., Lim, J., Richman, R., Fryer, J. D., Hyun, E. D., Duvick, L. A., Orr, H. T., Botas, J., et al. (2006). ATAXIN-1 Interacts with the Repressor Capicua in Its Native Complex to Cause SCA1 Neuropathology. *Cell* **127**, 1335–1347.

Lane, M. E., Sauer, K., Wallace, K., Jan, Y. N., Lehner, C. F. and Vaessin, H. (1996). Dacapo, a cyclin-dependent kinase inhibitor, stops cell proliferation during Drosophila development. *Cell* **87**, 1225–1235.

Lasagna-Reeves, C. A., Rousseaux, M. W. C., Guerrero-Munoz, M. J., Park, J., Jafar-Nejad, P., Richman, R., Lu, N., Sengupta, U., Litvinchuk, A., Orr, H. T., et al. (2015). A native interactor scaffolds and stabilizes toxic Ataxin-1 oligomers in SCA1. *Elife* **4**, 1–46.

Lasek, R. J. and Dower, W. J. (1971). Aplysia californica: Analysis of nuclear DNA in individual nuclei of giant neurons. *Science (80-.).* **172**, 278–280.

Le Loarer, F., Pissaloux, D., Watson, S., Godfraind, C., Galmiche-Rolland, L., Silva, K., Mayeur, L., Italiano, A., Michot, A., Pierron, G., et al. (2019). Clinicopathologic features of CIC-NUTM1 sarcomas, a new molecular variant of the family of CIC-fused sarcomas. *Am. J. Surg. Pathol.* **43**, 268–276.

Lee, T. and Luo, L. (2001). Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. *Trends Neurosci.* **24**, 251–254.

Lee, C. J., Chan, W. I., Cheung, M., Cheng, Y. C., Appleby, V. J., Orme, A. T. and Scotting, P. J. (2002). CIC, a member of a novel subfamily of the HMG-box superfamily, is transiently expressed in developing granule neurons. *Mol. Brain Res.* **106**, 151–156.

Lee, H. O., Davidson, J. M. and Duronio, R. J. (2009). Endoreplication: Polyploidy with purpose. *Genes Dev.* **23**, 2461–2477.

Lee, Y., Fryer, J. D., Kang, H., Crespo-Barreto, J., Bowman, A. B., Gao, Y., Kahle, J. J., Hong, J. S., Kheradmand, F., Orr, H. T., et al. (2011). ATXN1 Protein Family and CIC Regulate Extracellular Matrix Remodeling and Lung Alveolarization. *Dev. Cell* **21**, 746–757.

Lees, J. G., Dawson, N. L., Sillitoe, I. and Orengo, C. A. (2016). Functional innovation from changes in protein domains and their combinations. *Curr. Opin. Struct. Biol.* **38**, 44–52.

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

Li, L., Anderson, S., Secombe, J. and Eisenman, R. N. (2013). The Drosophila ubiquitin-specific protease Puffyeye regulates dMyc-mediated growth. *Development* **140**, 4776–4787.

Liao, S., Davoli, T., Leng, Y., Li, M. Z., Xu, Q. and Elledge, S. J. (2017). A genetic interaction analysis identifies cancer drivers that modify EGFR dependency. *Genes Dev.* **31**, 184–196.

Lilly, M. A. and Spradling, A. C. (1996). The Drosophila endocycle is controlled by cyclin E and lacks a checkpoint ensuring S-phase completion. *Genes Dev.* **10**, 2514–2526.

Lim, B., Samper, N., Lu, H., Rushlow, C., Jiménez, G. and Shvartsman, S. Y. (2013). Kinetics of gene derepression by ERK signaling. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 10330–10335.

López-Schier, H. and St. Johnston, D. (2001). Delta signaling from the germ line controls the proliferation and differentiation of the somatic follicle cells during Drosophila oogenesis. *Genes Dev.* **15**, 1393–1405.

Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W. K., Ohgaki, H., Wiestler, O. D., Kleihues, P. and Ellison, D. W. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* **131**, 803–820.

Lu, H. C., Tan, Q., Rousseaux, M. W. C., Wang, W., Kim, J. Y., Richman, R., Wan, Y. W., Yeh, S. Y., Patel, J. M., Liu, X., et al. (2017). Disruption of the ATXN1-CIC complex causes a spectrum of neurobehavioral phenotypes in mice and humans. *Nat. Genet.* **49**, 527–536.

Lüscher, B., Kuenzel, E. A., Krebs, E. G. and Eisenman, R. N. (1989). Myc oncoproteins are phosphorylated by casein kinase II. *EMBO J.* **8**, 1111–1119.

Mahajan-Miklos, S. and Cooley, L. (1994a). The villin-like protein encoded by the Drosophila quail gene is required for actin bundle assembly during oogenesis. *Cell* **78**, 291–301.

Mahajan-Miklos, S. and Cooley, L. (1994b). Intercellular Cytoplasm Transport during Drosophila Oogenesis. *Dev. Biol.* **165**, 336–351.

Maines, J. Z., Stevens, L. M., Tong, X. and Stein, D. (2004). Drosophila dMyc is required for ovary cell growth and endoreplication. *Development* **131**, 775–786.

Malarkey, C. S. and Churchill, M. E. A. (2012). The high mobility group box: The ultimate utility player of a cell. *Trends Biochem. Sci.* **37**, 553–562.

Manseau, L. J. and Schüpbach, T. (1989). The egg came first, of course!. Anterior-posterior pattern formation in Drosophila embryogenesis and oogenesis. *Trends Genet.* **5**, 400–405.

Manseau, L., Galley, J. and Phan, H. (1996). Profilin is required for posterior patterning of the Drosophila oocyte. *Development* **122**, 2109–2116.

McNeill, H., Yang, C. H., Brodsky, M., Ungos, J. and Simon, M. A. (1997). Mirror encodes a novel PBX-class homeoprotein that functions in the definition of the dorsal-ventral border in the Drosophila eye. *Genes Dev.* **11**, 1073–1082.

Miao, L. J., Yan, S., Zhuang, Q. F., Mao, Q. Y., Xue, D., He, X. Z. and Chen, J. P. (2019). miR-106b promotes proliferation and invasion by targeting capicua through MAPK signaling in renal carcinoma cancer. *Onco. Targets. Ther.* **12**, 3595–3607.

Miron, M., Lasko, P. and Sonenberg, N. (2003). Signaling from Akt to FRAP/TOR Targets both 4E-BP andS6K in Drosophilamelanogaster. *Mol. Cell. Biol.* **23**, 9117–9126.

Moberg, K. H., Bell, D. W., Wahrer, D. C., Haber, D. A. and Hariharan, I. K. (2001). Archipelago regulates Cyclin E levels in Drosophila and is mutated in human cancer cell lines. *Nature* **413**, 311–316.

Moberg, K. H., Mukherjee, A., Veraksa, A., Artavanis-Tsakonas, S. and Hariharan, I. K. (2004). The Drosophila F box protein archipelago regulates dMyc protein levels in vivo. *Curr. Biol.* **14**, 965–974.

Mondragon, A. A., Yalonetskaya, A., Ortega, A. J., Zhang, Y., Naranjo, O., Elguero, J., Chung, W. S. and McCall, K. (2019). Lysosomal Machinery Drives Extracellular Acidification to Direct Non-apoptotic Cell Death. *Cell Rep.* **27**, 11–19.

Monge, M., Colas, E., Doll, A., Gonzalez, M., Gil-Moreno, A., Planaguma, J., Quiles, M., Arbos, M. A., Garcia, A., Castellvi, J., et al. (2007). ERM/ETV5 up-regulation plays a role during myometrial infiltration through matrix metalloproteinase-2 activation in endometrial cancer. *Cancer Res.* **67**, 6753–6759.

Myster, D. L., Bonnette, P. C. and Duronio, R. J. (2000). A role for the DP subunit of the E2F transcription factor in axis determination during Drosophila oogenesis. *Development* **127**, 3249–3261.

Narbonne-Reveau, K., Senger, S., Pal, M., Herr, A., Richardson, H. E., Asano, M., Deak, P. and Lilly, M. A. (2008). APC/CFzr/Cdh1 promotes cell cycle progression during the Drosophila endocycle. *Development* **135**, 1451–1461.

Okimoto, R. A. and Bivona, T. G. (2017). Metastasis: From head to tail. *Cell Cycle* **16**, 487–488.

Okimoto, R. A., Breitenbuecher, F., Olivas, V. R., Wu, W., Gini, B., Hofree, M., Asthana, S., Hrustanovic, G., Flanagan, J., Tulpule, A., et al. (2017). Inactivation of Capicua drives cancer metastasis. *Nat. Genet.* **49**, 87–96.

Okimoto, R. A., Wu, W., Nanjo, S., Olivas, V., Lin, Y. K., Ponce, R. K., Oyama, R., Kondo, T. and Bivona, T. G. (2019). CIC-DUX4 oncoprotein drives sarcoma metastasis and tumorigenesis via distinct regulatory programs. *J. Clin. Invest.* **129**, 3401–3406.

Orr-Weaver, T. L. (2015). When bigger is better: The role of polyploidy in organogenesis. *Trends Genet.* **31**, 307–315.

Orr, H. T., Chung, M. yi, Banfi, S., Kwiatkowski, T. J., Servadio, A., Beaudet, A. L., McCall, A. E., Duvick, L. A., Ranum, L. P. W. and Zoghbi, H. Y. (1993). Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet.* **4**, 221–226.

Ouseph, M. M., Li, J., Chen, H. Z., Pécot, T., Wenzel, P., Thompson, J. C., Comstock, G., Chokshi, V., Byrne, M., Forde, B., et al. (2012). Atypical E2F Repressors and Activators Coordinate Placental Development. *Dev. Cell* **22**, 849–862.

Øvrebø, J. I. and Edgar, B. A. (2018). Polyploidy in tissue homeostasis and regeneration. *Development* **145**,.

Papagianni, A., Forés, M., Shao, W., He, S., Koenecke, N., Andreu, M. J., Samper, N., Paroush, Z., González-Crespo, S., Zeitlinger, J., et al. (2018). Capicua controls Toll/IL-1 signaling targets independently of RTK regulation. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 1807–1812.

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

Park, S., Lee, S., Lee, C. G., Park, G. Y., Hong, H., Lee, J. S., Kim, Y. M., Lee, S. B., Hwang, D., Choi, Y. S., et al. (2017). Capicua deficiency induces autoimmunity and promotes follicular helper T cell differentiation via derepression of ETV5. *Nat. Commun.* **8**,.

Park, S., Park, J., Kim, E. and Lee, Y. (2019). The Capicua/ETS Translocation Variant 5 Axis Regulates Liver-Resident Memory CD8+ T-Cell Development and the Pathogenesis of Liver Injury. *Hepatology* **70**, 358–371.

Parra-Peralbo, E. and Culi, J. (2011). Drosophila lipophorin receptors mediate the uptake of neutral lipids in oocytes and imaginal disc cells by an endocytosis-independent mechanism. *PLoS Genet.* **7**,.

Pierce, S. B., Yost, C., Britton, J. S., Loo, L. W. M., Flynn, E. M., Edgar, B. A. and Eisenman, R. N. (2004). dMyc is required for larval growth and endoreplication in Drosophila. *Development* **131**, 2317–2327.

Polymenis, M. and Schmidt, E. V. (1997). Coupling of cell division to cell growth by translational control of the G1 cyclin CLN3 in yeast. *Genes Dev.* **11**, 2522–2531.

Port, F., Chen, H. M., Lee, T. and Bullock, S. L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* **111**,.

Prober, D. A. and Edgar, B. A. (2000). Ras1 promotes cellular growth in the Drosophila wing. *Cell* **100**, 435–446.

Quinlan, M. E. (2016). Cytoplasmic Streaming in the Drosophila Oocyte. *Annu. Rev. Cell Dev. Biol.* **32**, 173–195.

Richardson, H., O'Keefe, L. V., Marty, T. and Saint, R. (1995). Ectopic cyclin E expression induces premature entry into S phase and disrupts pattern formation in the Drosophila eye imaginal disc. *Development* **121**, 3371–3379.

Riechmann, V. and Ephrussi, A. (2001). Axis formation during Drosophila oogenesis. *Curr. Opin. Genet. Dev.* **11**, 374–383.

Rittenhouse, K. R. and Berg, C. A. (1995). Mutations in the Drosophila gene bullwinkle cause the formation of abnormal eggshell structures and bicaudal embryos. *Development* **121**, 3023–3033.

Roberts, P. J. and Der, C. J. (2007). Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene* **26**, 3291–3310.

Robinson, D. N., Cant, K. and Cooley, L. (1994). Morphogenesis of Drosophila ovarian ring canals. *Development* **120**, 2015–2025.

Robinson, D. N., Smith-Leiker, T. A., Sokol, N. S., Hudson, A. M. and Cooley, L. (1997). Formation of the Drosophila ovarian ring canal inner rim depends on cheerio. *Genetics* **145**, 1063–1072.

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

Rørth, P. (1998). Gal4 in the Drosophila female germline. *Mech. Dev.* **78**, 113–118.

Rousseaux, M. W. C., Tschumperlin, T., Lu, H. C., Lackey, E. P., Bondar, V. V., Wan, Y. W., Tan, Q., Adamski, C. J., Friedrich, J., Twaroski, K., et al. (2018). ATXN1-CIC Complex Is the Primary Driver of Cerebellar Pathology in Spinocerebellar Ataxia Type 1 through a Gain-of-Function Mechanism. *Neuron* **97**, 1235–1243.

Rousso, T., Schejter, E. D. and Shilo, B. Z. (2016). Orchestrated content release from Drosophila glue-protein vesicles by a contractile actomyosin network. *Nat. Cell Biol.* **18**, 181–190.

Royzman, I., Whittaker, A. J. and Orr-Weaver, T. L. (1997). Mutations in Drosophila DP and E2F distinguish G1-S progression from an associated transcriptional program. *Genes Dev.* **11**, 1999–2011.

Royzman, I., Hayashi-Hagihara, A., Dej, K. J., Bosco, G., Lee, J. Y. and Orr-Weaver, T. L. (2002). The E2F cell cycle regulator is required for Drosophila nurse cell DNA replication and apoptosis. *Mech. Dev.* **119**, 225–237.

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., Luna, A., La, K. C., Dimitriadoy, S., Liu, D. L., Kantheti, H. S., Saghafinia, S., et al. (2018). Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321–337.

Sanchez, J. N., Wang, T. and Cohen, M. S. (2018). BRAF and MEK Inhibitors: Use and Resistance in BRAF-Mutated Cancers. *Drugs* **78**, 549–566.

Sears, R., Nuckolls, F., Haura, E., Taya, Y., Tamai, K. and Nevins, J. R. (2000). Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. *Genes Dev.* **14**, 2501–2514.

Seim, I., Jeffery, P. L., Thomas, P. B., Nelson, C. C. and Chopin, L. K. (2017). Whole-genome sequence of the metastatic PC3 and LNCaP human prostate cancer cell lines. *G3 Genes, Genomes, Genet.* **7**, 1731–1741.

Sen, J., Goltz, J. S., Stevens, L. and Stein, D. (1998). Spatially restricted expression of pipe in the Drosophila egg chamber defines embryonic dorsal-ventral polarity. *Cell* **95**, 471–481.

Shcherbata, H. R., Althauser, C., Findley, S. D. and Ruohola-Baker, H. (2004). The mitotic-to-endrocycle switch in Drosophila follicle cells is executed by Notch-dependent regulation of G1/S, G2/M and M/G1 cell-cycle transitions. *Development* **131**, 3169–3181.

Shepherd, T. G., Kockeritz, L., Szrajber, M. R., Muller, W. J. and Hassell, J. A. (2001). The pea3 subfamily ets genes are required for HER2/Neu-mediated mammary oncogenesis. *Curr. Biol.* **11**, 1739–1748.

Shibutani, S., Swanhart, L. M. and Duronio, R. J. (2007). Rbf1-independent termination of E2f1-target gene expression during early Drosophila embryogenesis. *Development* **134**, 467–478.

Shibutani, S. T., de la Cruz, A. F. A., Tran, V., Turbyfill, W. J., Reis, T., Edgar, B. A. and Duronio, R. J. (2008). Intrinsic Negative Cell Cycle Regulation Provided by PIP Box- and Cul4Cdt2-Mediated Destruction of E2f1 during S Phase. *Dev. Cell* **15**, 890–900.

Shu, Z., Row, S. and Deng, W. M. (2018). Endoreplication: The Good, the Bad, and the Ugly. *Trends Cell Biol.* **28**, 465–474.

Sigrist, S. J. and Lehner, C. F. (1997). Drosophila fizzy-related down-regulates mitotic cyclins and is required for cell proliferation arrest and entry into endocycles. *Cell* **90**, 671–681.

Simón-Carrasco, L., Graña, O., Salmón, M., Jacob, H. K. C., Gutierrez, A., Jiménez, G., Drosten, M. and Barbacid, M. (2017). Inactivation of Capicua in adult mice causes T-cell lymphoblastic lymphoma. *Genes Dev.* **31**, 1456–1468.

Simón-Carrasco, L., Jiménez, G., Barbacid, M. and Drosten, M. (2018a). The Capicua tumor suppressor: a gatekeeper of Ras signaling in development and cancer. *Cell Cycle* **17**, 702–711.

Simón-Carrasco, L., Jiménez, G., Barbacid, M. and Drosten, M. (2018b). The Capicua tumor suppressor: a gatekeeper of Ras signaling in development and cancer. *Cell Cycle* **17**, 702–711.

Sinsimer, K. S., Jain, R. A., Chatterjee, S. and Gavis, E. R. (2011). A late phase of germ plasm accumulation during Drosophila oogenesis requires Lost and Rumpelstiltskin. *Development* **138**, 3431–3440.

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

Smith, J. L., Wilson, J. E. and Macdonald, P. M. (1992). Overexpression of oskar directs ectopic activation of nanos and presumptive pole cell formation in Drosophila embryos. *Cell* **70**, 849–859.

Specht, K., Sung, Y. S., Zhang, L., Richter, G. H. S., Fletcher, C. D. and Antonescu, C. R. (2014). Distinct transcriptional signature and immunoprofile of CIC-DUX4 fusion-positive round cell tumors compared to EWSR1-rearranged ewing sarcomas: Further evidence toward distinct pathologic entities. *Genes Chromosom. Cancer* **53**, 622–633.

Spradling, A. C. (1993). Developmental genetics of oogenesis. In *The development of Drosophila melanogaster.*, pp. 1–70.

Storchova, Z. and Pellman, D. (2004). From polyploidy to aneuploidy, genome instability and cancer. *Nat. Rev. Mol. Cell Biol.* **5**, 45–54.

Strohmaier, H., Spruck, C. H., Kaiser, P., Won, K. A., Sangfelt, O. and Reed, S. I. (2001). Human F-box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line. *Nature* **413**, 316–322.

Sturm, D., Orr, B. A., Toprak, U. H., Hovestadt, V., Jones, D. T. W., Capper, D., Sill, M., Buchhalter, I., Northcott, P. A., Leis, I., et al. (2016). New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. *Cell* **164**, 1060–1072.

Sugita, S., Arai, Y., Tonooka, A., Hama, N., Totoki, Y., Fujii, T., Aoyama, T., Asanuma, H., Tsukahara, T., Kaya, M., et al. (2014). A novel CIC-FOXO4 gene fusion in undifferentiated small round cell sarcoma: A genetically distinct variant of Ewing-like sarcoma. *Am. J. Surg. Pathol.* **38**, 1571–1576.

Suisse, A., He, D. Q., Legent, K. and Treisman, J. E. (2017). COP9 signalosome subunits protect capicua from MAPK-dependent and -independent mechanisms of degradation. *Development* **144**, 2673–2682.

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.

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.

Sun, J., Smith, L., Armento, A. and Deng, W. M. (2008). Regulation of the endocycle/gene amplification switch by Notch and ecdysone signaling. *J. Cell Biol.* **182**, 885–896.

Suzuki, H., Aoki, K., Chiba, K., Sato, Y., Shiozawa, Y., Shiraishi, Y., Shimamura, T., Niida, A., Motomura, K., Ohka, F., et al. (2015). Mutational landscape and clonal architecture in grade II and III gliomas. *Nat. Genet.* **47**, 458–468.

Swanson, C. I., Meserve, J. H., McCarter, P. C., Thieme, A., Mathew, T., Elston, T. C. and Duronio, R. J. (2015). Expression of an S phase-stabilized version of the CDK inhibitor Dacapo can alter endoreplication. *Development* **142**, 4288–4298.

Szuplewski, S., Sandmann, T., Hietakangas, V. and Cohen, S. M. (2009). Drosophila minus is required for cell proliferation and influences Cyclin E turnover. *Genes Dev.* **23**, 1998–2003.

Tan, Q. and Zoghbi, H. Y. (2019). Mouse models as a tool for discovering new neurological diseases. *Neurobiol. Learn. Mem.* **165**,.

Tan, Q., Brunetti, L., Rousseaux, M. W. C., Lu, H. C., Wan, Y. W., Revelli, J. P., Liu, Z., Goodell, M. A. and Zoghbi, H. Y. (2018). Loss of Capicua alters early T cell development and predisposes mice to T cell lymphoblastic leukemia/lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E1511–E1519.

Tanaka, S. and Araki, H. (2013). Helicase activation and establishment of replication forks at chromosomal origins of replication. *Cold Spring Harb. Perspect. Biol.* **5**,.

Tanaka, M., Yoshimoto, T. and Nakamura, T. (2017). A double-edged sword: The world according to Capicua in cancer. *Cancer Sci.* **108**, 2319–2325.

Teleman, A. A., Hietakangas, V., Sayadian, A. C. and Cohen, S. M. (2008). Nutritional Control of Protein Biosynthetic Capacity by Insulin via Myc in Drosophila. *Cell Metab.* **7**, 21–32.

Thumkeo, D., Shimizu, Y., Sakamoto, S., Yamada, S. and Narumiya, S. (2005). ROCK-I and ROCK-II cooperatively regulate closure of eyelid and ventral body wall in mouse embryo. *Genes to Cells* **10**, 825–834.

Timmons, A. K., Mondragon, A. A., Schenkel, C. E., Yalonetskaya, A., Taylor, J. D., Moynihan, K. E., Etchegaray, J. I., Meehan, T. L. and McCall, K. (2016). Phagocytosis genes nonautonomously promote developmental cell death in the Drosophila ovary. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1246–E1255.

Tomlins, S. A., Laxman, B., Dhanasekaran, S. M., Helgeson, B. E., Cao, X., Morris, D. S., Menon, A., Jing, X., Cao, Q., Han, B., et al. (2007). Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **448**, 595–599.

Tong, X., Gui, H., Jin, F., Heck, B. W., Lin, P., Ma, J., Fondell, J. D. and Tsai, C. C. (2011). Ataxin-1 and Brother of ataxin-1 are components of the Notch signalling pathway. *EMBO Rep.* **12**, 428–435.

Trakala, M., Rodríguez-Acebes, S., Maroto, M., Symonds, C. E., Santamaría, D., Ortega, S., Barbacid, M., Méndez, J. and Malumbres, M. (2015). Functional Reprogramming of Polyploidization in Megakaryocytes. *Dev. Cell* **32**, 155–167.

Tran, D. H. and Berg, C. A. (2003). Bullwinkle and shark regulate dorsal-appendage morphogenesis in Drosophila oogenesis. *Development* **130**, 6273–6282.

Trumpp, A., Refaell, Y., Oskarsson, T., Gasser, S., Murphy, M., Martin, G. R. and Bishop, J. M. (2001). C-Myc regulates mammalian body size by controlling cell number but not cell size. *Nature* **414**, 768–773.

Tseng, A. S. K., Tapon, N., Kanda, H., Cigizoglu, S., Edelmann, L., Pellock, B., White, K. and Hariharan, I. K. (2007). Capicua Regulates Cell Proliferation Downstream of the Receptor Tyrosine Kinase/Ras Signaling Pathway. *Curr. Biol.* **17**, 728–733.

Unhavaithaya, Y. and Orr-Weaver, T. L. (2012). Polyploidization of glia in neural development links tissue growth to blood-brain barrier integrity. *Genes Dev.* **26**, 31–36.

Urata, Y., Parmelee, S. J., Agard, D. A. and Sedat, J. W. (1995). A three-dimensional structural dissection of Drosophila polytene chromosomes. *J. Cell Biol.* **131**, 279–295.

Vervoorts, J., Lüscher-Firzlaff, J. and Lüscher, B. (2006). The ins and outs of MYC regulation by posttranslational mechanisms. *J. Biol. Chem.* **281**, 34725–34729.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. and Kinzler, K. W. (2013). Cancer genome landscapes. *Science (80-.).* **340**, 1546–1558.

Volpe, A. M., Horowitz, H., Grafer, C. M., Jackson, S. M. and Berg, C. A. (2001). Drosophila rhino encodes a female-specific chromo-domain protein that affects chromosome structure and egg polarity. *Genetics* **159**, 1117–1134.

Von Lintig, F. C., Huvar, I., Law, P., Diccianni, M. B., Yu, A. L. and Boss, G. R. (2000). Ras activation in normal white blood cells and childhood acute lymphoblastic leukemia. *Clin. Cancer Res.* **6**, 1804–1810.

Wang, B., Krall, E. B., Aguirre, A. J., Kim, M., Widlund, H. R., Doshi, M. B., Sicinska, E., Sulahian, R., Goodale, A., Cowley, G. S., et al. (2017). ATXN1L, CIC, and ETS Transcription Factors Modulate Sensitivity to MAPK Pathway Inhibition. *Cell Rep.* **18**, 1543–1557.

Wang, Z. H., Liu, Y., Chaitankar, V., Pirooznia, M. and Xu, H. (2019). Electron transport chain biogenesis activated by a JNK-insulin-MYC relay primes mitochondrial inheritance in Drosophila. *Elife* **8**,.

Waring, G. L. (2000). Morphogenesis of the eggshell in Drosophila. *Int. Rev. Cytol.* **198**, 67–108.

Weigmann, K., Cohen, S. M. and Lehner, C. F. (1997). Cell cycle progression, growth and patterning in imaginal discs despite inhibition of cell division after inactivation of Drosophila Cdc2 kinase. *Development* **124**, 3555–3563.

Weiss, A., Herzig, A., Jacobs, H. and Lehner, C. F. (1998). Continuous Cyclin E expression inhibits progression through endoreduplication cycles in Drosophila. *Curr. Biol.* **8**, 239–242.

Weissmann, S., Cloos, P. A., Sidoli, S., Jensen, O. N., Pollard, S. and Helin, K. (2018). The tumor suppressor CIC directly regulates MAPK pathway genes via histone deacetylation. *Cancer Res.* **78**, 4114–4125.

Welcker, M. and Clurman, B. E. (2008). FBW7 ubiquitin ligase: A tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat. Rev. Cancer* **8**, 83–93.

Welcker, M., Singer, J., Loeb, K. R., Grim, J., Bloecher, A., Gurien-West, M., Clurman, B. E. and Roberts, J. M. (2003). Multisite phosphorylation by Cdk2 and GSK3 controls cyclin E degradation. *Mol. Cell* **12**, 381–392.

Welcker, M., Orian, A., Jin, J., Grim, J. A., Harper, J. W., Eisenman, R. N. and Clurman, B. E. (2004). The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9085–9090.

Weng, L., Zhu, C., Xu, J. and Du, W. (2003). Critical role of active repression by E2F and Rb proteins in endoreplication during Drosophila development. *EMBO J.* **22**, 3865–3875.

Wheatley, S., Kulkarni, S. and Karess, R. (1995). Drosophila nonmuscle myosin II is required for rapid cytoplasmic transport during oogenesis and for axial nuclear migration in early embryos. *Development* **121**, 1937–1946.

Wong, D. and Yip, S. (2020). Making heads or tails – the emergence of capicua (CIC) as an important multifunctional tumour suppressor. *J. Pathol.* **250**, 532–540.

Wong, D., Sogerer, L., Lee, S. S., Wong, V., Lum, A., Levine, A. B., Marra, M. A. and Yip, S. (2020). TRIM25 promotes Capicua degradation independently of ERK in the absence of ATXN1L. *BMC Biol.* **18**,.

Yada, M., Hatakeyama, S., Kamura, T., Nishiyama, M., Tsunematsu, R., Imaki, H., Ishida, N., Okumura, F., Nakayama, K. and Nakayama, K. I. (2004). Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7. *EMBO J.* **23**, 2116–2125.

Yang, L. and Veraksa, A. (2017). Single-step affinity purification of ERK signaling complexes using the streptavidin-binding peptide (SBP) tag. In *Methods in Molecular Biology*, pp. 113–126. Humana Press Inc.

Yang, L., Paul, S., Trieu, K. G., Dent, L. G., Froldi, F., Forés, M., Webster, K., Siegfried, K. R., Kondo, S., Harvey, K., et al. (2016). Minibrain and Wings apart control organ growth and tissue patterning through down-regulation of Capicua. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 10583–10588.

Yang, R., Chen, L. H., Hansen, L. J., Carpenter, A. B., Moure, C. J., Liu, H., Pirozzi, C. J., Diplas, B. H., Waitkus, M. S., Greer, P. K., et al. (2017). Cic loss promotes gliomagenesis via aberrant neural stem cell proliferation and differentiation. *Cancer Res.* **77**, 6097–6108.

Yeh, E., Cunningham, M., Arnold, H., Chasse, D., Monteith, T., Ivaldi, G., Hahn, W. C., Stukenberg, P. T., Shenolikar, S., Uchida, T., et al. (2004). A signalling pathway controlling c-Myc degradation that impacts oncogenic transformation of human cells. *Nat. Cell Biol.* **6**, 308–318.

Yip, S., Butterfield, Y. S., Morozova, O., Chittaranjan, S., Blough, M. D., An, J., Birol, I., Chesnelong, C., Chiu, R., Chuah, E., et al. (2012). Concurrent CIC mutations, IDH mutations, and 1p/19q loss distinguish oligodendrogliomas from other cancers. *J. Pathol.* **226**, 7–16.

Yoshida, A., Goto, K., Kodaira, M., Kobayashi, E., Kawamoto, H., Mori, T., Yoshimoto, S., Endo, O., Kodama, N., Kushima, R., et al. (2016). CIC-rearranged sarcomas: A study of 20 cases and comparisons with Ewing sarcomas. *Am. J. Surg. Pathol.* **40**, 313–323.

Yoshimoto, T., Tanaka, M., Homme, M., Yamazaki, Y., Takazawa, Y., Antonescu, C. R. and Nakamura, T. (2017). CIC-DUX4 induces small round cell sarcomas distinct from ewing sarcoma. *Cancer Res.* **77**, 2927–2937.

You, M. J., Medeiros, L. J. and Hsi, E. D. (2015). T-lymphoblastic leukemia/lymphoma. *Am. J. Clin. Pathol.* **144**, 411–422.

Zecca, M. and Struhl, G. (2002). Subdivision of the Drosophila wing imaginal disc by EGFR-mediated signaling. *Development* **129**, 1357–1368.

Zhou, Y., Wang, M., Shuang, T., Liu, Y., Zhang, Y. and Shi, C. (2019). MiR-1307 influences the chemotherapeutic sensitivity in ovarian cancer cells through the regulation of the CIC transcriptional repressor. *Pathol. Res. Pract.* **215**,.

Zielke, N., Querings, S., Rottig, C., Lehner, C. and Sprenger, F. (2008). The anaphase-promoting complex/cyclosome (APC/C) is required for rereplication control in endoreplication cycles. *Genes Dev.* **22**, 1690–1703.

Zielke, N., Kim, K. J., Tran, V., Shibutani, S. T., Bravo, M. J., Nagarajan, S., Van Straaten, M., Woods, B., Von Dassow, G., Rottig, C., et al. (2011). Control of Drosophila endocycles by E2F and CRL4 CDT2. *Nature* **480**, 123–127.

Zielke, N., Edgar, B. A. and DePamphilis, M. L. (2013). Endoreplication. *Cold Spring Harb. Perspect. Biol.* **5**,.

Zimyanin, V. L., Belaya, K., Pecreaux, J., Gilchrist, M. J., Clark, A., Davis, I. and St Johnston, D. (2008). In Vivo Imaging of oskar mRNA Transport Reveals the Mechanism of Posterior Localization. *Cell* **134**, 843–853.